Kalman Filtering with Intermittent Observations: Weak Convergence to a Stationary Distribution

Soummya Kar, Bruno Sinopoli, and José M. F. Moura*

Abstract

The paper studies the asymptotic behavior of discrete time Random Riccati Equations (RRE) arising in Kalman filtering when the arrival of the observations is described by a Bernoulli i.i.d. process. We model the RRE as an order-preserving, strongly sublinear random dynamical system (RDS). Under a sufficient condition, stochastic boundedness, and using a limit-set dichotomy result for order-preserving, strongly sublinear RDS, we establish the asymptotic properties of the RRE: the sequence of random prediction error covariance matrices converges weakly to a unique invariant distribution, whose support exhibits fractal behavior. For stabilizable and detectable systems, stochastic boundedness (and hence weak convergence) holds for any non-zero observation packet arrival probability and, in particular, we can establish weak convergence at operating arrival rates well below the critical probability for mean stability (the resulting invariant measure in that situation does not possess a first moment.) We apply the weak-Feller property of the Markov process governing the RRE to characterize the support of the limiting invariant distribution as the topological closure of a countable set of points, which, in general, is not dense in the set of positive semi-definite matrices. We use the explicit characterization of the support of the invariant distribution and the almost sure ergodicity of the sample paths to easily compute statistics of the invariant distribution. A one dimensional example illustrates that the support is a fractured subset of the non-negative reals with self-similarity properties.

I. Introduction

A. Background and Motivation

Named after Count Jacopo Francesco Riccati, the man who conceived and studied it first, the Riccati equation has received great interest in science and engineering. In particular, its applications to control theory are widespread, ranging from optimal to robust and stochastic control. In Kalman filtering, [1], the Riccati equation describes the evolution of the state error covariance for linear Gaussian systems. We focus in the paper on the discrete version of the Riccati equation. Kalman showed that, for a linear time-invariant system, under detectability conditions, the Riccati equation converges to a fixed point,

The authors are with the Dep. Electrical and Computer Engineering, Carnegie Mellon University, Pittsburgh, PA 15213, USA (e-mail: soummyak@andrew.cmu.edu, brunos@ece.cmu.edu, moura@ece.cmu.edu, ph: (412)268-6341, fax: (412)268-3890.)

Work supported by NSF under grants # ECS-0225449 and # CNS-0428404.

which is also unique if certain stabilizability conditions are satisfied. The result is very powerful as it asserts that the estimation steady state error is constant. As a consequence, the steady state estimator gain is also constant, providing a very practical result for implementation. The problem is more involved when the system matrices are time-varying, and it is further complicated if, in addition, they are random.

Initial study of Random Riccati Equations (RRE)[1] was motivated by the Linear Quadratic Gaussian (LQG) in optimal control when the system parameters are random. This leads naturally to a RRE. In adaptive control, where the parameters of the system are unknown and need to be identified, RRE also arises. Initial studies of RRE are in [2], where the authors consider linear stochastic systems with additive white Gaussian noise, with the added generality that the system matrices are random and adapted to the observation process. The paper shows that the sufficient conditions for the Kalman Filter to provide mean and covariance of the conditionally Gaussian state estimate are that the random matrices are finite with probability one at each time instant. This result applies to control problems of a linear stochastic system in the case its parameters need to be identified recursively. More recently, Wang and Guo [3] provide sufficient conditions on the stochastic Grammian to guarantee stability of RREs.

In the past few years, RRE has received renewed interest in the area of networked control systems. This is concerned with estimation and control where components, namely, sensors, controllers, and actuators are connected via general purpose communication channels, such as ethernet, W-LANs, or Personal area networks (PANs), e.g., IEEE 802.15.4-based networks. In this realm, the stochastic characteristics of the channels introduce additional sources of randomness, non Gaussian, in the control problem. Special interest has been given to analog erasure channels. Under this model, the observation packet is either dropped with probability $\bar{\gamma}$, or reaches the receiver with probability $1-\bar{\gamma}$. One limitation of this model is that it does not take into account quantization. Limits of control in the presence of quantization have been addressed in [4], [5], [6], [7], [8]. Fundamental results show that systems can be stabilized with quantization level easily achievable by common off-the-shelf A/D converters. This makes the infinite precision assumption realistic. The results provided in this paper therefore neglect quantization effects.

In Matveev and Savkin [9], the authors consider Kalman filtering where observations can suffer bounded delay in communication between the sensors and the estimator. Sinopoli et al. [10] consider a discrete-time system in which the arrival of an observation at the estimator is modeled as a Bernoulli i.i.d. random process γ_t. The observation is received by the estimator with probability $\bar{\gamma}$. They show that under this model the Kalman Filter is still the optimal estimator and study the time evolution of the error covariance. Differently from the standard Kalman Filter, the error covariance is now a random matrix, depending on the realization of the process $\{\gamma_t\}$. This is described by a RRE. They study the asymptotic behavior (in time) of its mean to determine stability of the filter and show that, depending on the eigenvalues of the matrix and on the structure of the matrix, there exists a critical value $\overline{\gamma}^{\mathrm{bim}}$, such that, if the probability of arrival of an observation at time t is $\bar{\gamma} > \overline{\gamma}^{\mathrm{bim}}$, then the expectation of the estimation error covariance is always finite (under stabilizability and detectability hypotheses.) The authors provide upper and lower

[1]In the sequel the term RRE refers to the discrete time Random Riccati Equations considered in this paper.

bounds for this critical probability and compute it in closed form for a few special cases. Subsequent work [11] characterizes the critical value for a large class of linear systems, showing a direct relationship with the spectral radius of the dynamic matrix A.

The model proposed in [10] has been widely adopted and extended by several authors [12], [13], [14], [15], [16], [17], [18]. Although many present extensions to general Markov chains and account for smart sensors sending local estimates instead of observations, all the results are established with respect to mean stability, i.e., boundedness of the mean covariance. This metric is unsatisfactory in many applications, as it does not provide information about the fluctuations of the error covariance that could grow and be unusable for long time intervals. We would like to characterize the asymptotic behavior of its distribution–the goal of this paper.

In this work, we characterize the asymptotics of the state error covariance for a linear Gaussian system where observations are lost according to a Bernoulli process, as in [10]. Based on stochastic boundedness (see Subsection II-B) of the sequence of random prediction error covariance matrices, we provide a sufficient condition (which is also necessary under broad assumptions, including stabilizability and detectability of the system in question) for the existence and uniqueness of an attracting invariant (stationary) distribution for the RRE. We show that stochastic boundedness implies weak convergence of the sequence of random prediction error covariance matrices to a unique invariant distribution, irrespective of the initial condition. We show that the mean stability considered in [10] implies stochastic boundedness and hence it is possible to operate at packet arrival probabilities below the threshold for mean stability and converge to an invariant distribution. In particular, for stabilizable and detectable systems, stochastic boundedness is ensured by operating at any non-zero packet arrival probability leading to weak convergence, whereas, the critical probability for boundedness in mean can be very high, depending on the instability of the system. However, operating above the critical probability for mean stability ensures that the invariant distribution has a finite mean, which may not hold if operated below. Our approach is based on modelling the RRE as an order-preserving random dynamical system (RDS) (see [19]), possessing the property of strong sublinearity (to be explained later.) We use a limit-set dichotomy result for such order-preserving, strongly sublinear RDS to establish asymptotic properties of the RRE concerning existence and uniqueness of invariant distributions. We contrast our work with Vakili and Hassibi [20] and Censi [21]. In [20], the authors take a completely different and very interesting approach. They use the Stieltjes transform to compute a fixed point for the RRE associated with intermittent loss of observation due to a Bernoulli process. Although this is numerically sound, it assumes the existence of a stationary distribution for the error covariance, and it is applicable only to large matrices, i.e., as N tends to infinity, which are also asymptotically free [22]. When the first draft of our paper was complete, we came across [21], which studies weak convergence of the RRE using the theory of Iterated Function Systems (IFS) (e.g., [23].) When the system matrix A is invertible and a non-overlapping condition is satisfied, the RRE satisfies a mean contraction property, leading to existence and uniqueness of an attracting invariant distribution (see [23]). Reference [21] uses these results to show weak convergence of the RRE to a unique invariant distribution if the system is operated above the critical probability for mean stability and the resulting invariant distribution has fractal support. By contrast, our paper shows weak convergence to an attracting invariant distribution for the general case and at operating points below the critical probability for mean stability.

The weak-Feller property of the Markov process governing the RRE enables us to explicitly characterize its support of the resulting invariant distribution. We show that its support is the topological closure (in the metric space of positive semidefinite matrices) of a countable set of points (given explicitly as functionals of the deterministic fixed point of the corresponding algebraic Riccati equation.) The above set of points is not, in general, a dense subset of the set of positive semidefinite matrices. A detailed study of a scalar example shows that the support is a highly fractured subset of the non-negative reals with self-similarity properties, thus exhibiting the characteristics of a fractal set. Finally, the explicit identification of the support of the invariant distribution in the general case and almost sure (a.s.) ergodicity of the sample paths enable us to easily compute numerically the moments (and probabilities) of the invariant distribution. In this context, we note that a complete analytic characterization of the resulting invariant measures (for example, probabilities of large excursions under the invariant measures) has been addressed more recently in the follow-up paper ([24]), which characterizes moderate deviations properties of the invariant measures as the packet arrival probability $\overline{\gamma}$ approaches 1.

The paper is organized as follows. Subsection I-B sets notation and summarizes preliminary results. Section II presents a rigorous formulation of the weak convergence problem and the main results of the paper are stated in Section III. The RDS formulation of the RRE is carried out in Section IV, while Section V establishes various properties of the RRE in the context of RDS theory. The proofs of the main results are presented in Section VI. Subsection VII-A analyzes a scalar example in detail, while numerical studies on the invariant distribution for the general case are presented in Subsection VII-B. Finally Section VIII concludes the paper.

B. *Notation and Preliminaries*

Denote by: $\mathbb{R}$, the reals; $\mathbb{R}^M$, the M-dimensional Euclidean space; $\mathbb{T}$, the integers; $\mathbb{T}_+$, the non-negative integers; $\mathbb{N}$, the natural numbers; and $\mathcal{X}$, a generic space. For a subset $B \subset \mathcal{X}$, $\mathbb{I}_B : \mathcal{X} \longmapsto \{0,1\}$ is the indicator function, which is 1 when the argument is in B and zero otherwise; and $id_{\mathcal{X}}$ is the identity function on $\mathcal{X}$.

Cones in partially ordered Banach spaces. We summarize facts and definitions on the structure of cones in partially ordered Banach spaces. Let V be a Banach space (over the field of the reals) with a closed (w.r.t. the Banach space norm) convex cone V_+ and assume $V_+ \cap (-V_+) = \{0\}$. The cone V_+ induces a partial order in V, namely, for $X, Y \in V$, we write $X \preceq Y$, if $Y - X \in V_+$. In case $X \preceq Y$ and $X \neq Y$, we write $X \prec Y$. The cone V_+ is called solid, if it has a non-empty interior $\operatorname{int} V_+$; in that case, V_+ defines a strong ordering in V, and we write $X \ll Y$, if $Y - X \in \operatorname{int} V_+$. The cone V_+ is normal if the norm $\|\cdot\|$ of V is semi-monotone, i.e., $\exists\, c > 0$, s.t. $0 \preceq X \preceq Y \Rightarrow \|X\| \leq c\|Y\|$. There are various equivalent characterizations of normality, of which we note that the normality of V_+ ensures that the topology in V induced by the Banach space norm is compatible with the ordering induced by

V_+, in the sense that any norm-bounded set $B \subset V$ is contained in a conic interval of the form $[X, Y]$, where $X, Y \in V$. Finally, a cone is said to be minihedral, if every order-bounded (both upper and lower bounded) finite set $B \subset V$ has a supremum (here bounds are w.r.t. the partial order.)

We focus on the separable Banach space of symmetric $N \times N$ matrices, $\mathbb{S}^N$, equipped with the induced 2-norm. The subset $\mathbb{S}_+^N$ of positive semidefinite matrices is a closed, convex, solid, normal, minihedral cone in $\mathbb{S}^N$, with non-empty interior $\mathbb{S}_{++}^N$, the set of positive definite matrices. The conventions above denote the partial and strong ordering in $\mathbb{S}^N$ induced by $\mathbb{S}_+^N$. For example, we use the notation $X \gg 0$ to denote that the matrix $X \in \mathbb{S}^N$ is positive definite, whereas $X \succeq 0$ denotes positive semidefiniteness and $X \succ 0$ indicates that X is positive semidefinite and different from the zero matrix.

Operator theoretic preliminaries. We review operator-theoretic concepts needed to analyze the Markov process generated by the random covariance equations, details in, for example, [25]. Let: $(\mathcal{X}, d)$ a locally compact separable metric space $\mathcal{X}$ with metric d; $\mathbb{B}(\mathcal{X})$ its Borel algebra; $B(\mathcal{X})$ the Banach space of real-valued bounded functions on $\mathcal{X}$, equipped with the sup-norm, i.e., $f \in B(\mathcal{X})$, $\|f\| = \sup_{x \in \mathcal{X}} |f(x)|$; and $C_b(\mathcal{X})$ the subspace of $B(\mathcal{X})$ of continuous functions.

Let $\mathcal{M}(\mathcal{X})$ be the space of finite Borel measures on $\mathcal{X}$. It is a Banach space under the total variation norm (see [25] for details.) For $\mu \in \mathcal{M}(\mathcal{X})$, we define the support of μ, supp(μ), by

$$\text{supp}(\mu) = \{x \in \mathcal{X} \mid \mu(B_\varepsilon(x)) > 0, \ \forall \varepsilon > 0\} \tag{1}$$

where $B_\varepsilon(x)$ is the open ball of radius ε centered at x. It follows that supp(μ) is a closed set. An element $\mu \in \mathcal{M}(\mathcal{X})$ is positive, i.e., $\mu \geq 0$, if

$$\mu(A) \geq 0, \forall\ A \in \mathbb{B}(\mathcal{X}) \tag{2}$$

For $f \in B(\mathcal{X})$, $\mu \in \mathcal{M}(\mathcal{X})$, we define the bilinear form $< \cdot, \cdot >: B(\mathcal{X}) \times \mathcal{M}(\mathcal{X}) \longmapsto \mathbb{R}$

$$< f, \mu > = \int_{\mathcal{X}} f(x) \mu(dx) \tag{3}$$

A linear operator $T : \mathcal{M}(\mathcal{X}) \longmapsto \mathcal{M}(\mathcal{X})$ is positive if $T\mu \geq 0$ for $\mu \geq 0$. It can be shown that such positive operators are necessarily bounded.

A linear operator $T : \mathcal{M}(\mathcal{X}) \longmapsto \mathcal{M}(\mathcal{X})$ is a contraction if $\|T\| \leq 1$.

A positive contraction $T : \mathcal{M}(\mathcal{X}) \longmapsto \mathcal{M}(\mathcal{X})$ is a Markov operator if $\|T\mu\| = \|\mu\|, \forall \mu \geq 0$.

Definition 1 (Markov-Feller Operator) Consider the linear operator $L : C_b(\mathcal{X}) \longmapsto C_b(\mathcal{X})$ and the Markov operator $T : \mathcal{M}(\mathcal{X}) \longmapsto \mathcal{M}(\mathcal{X})$. The pair (L, T) is a Markov-Feller pair if

$$< Lf, \mu > = < f, T\mu >, \forall f \in C_b(\mathcal{X}), \mu \in \mathcal{M}(\mathcal{X}) \tag{4}$$

A Markov operator $T : \mathcal{M}(\mathcal{X}) \longmapsto \mathcal{M}(\mathcal{X})$ is a Markov-Feller operator if there exists a linear operator $L : C_b(\mathcal{X}) \longmapsto C_b(\mathcal{X})$ such that (L, T) is a Markov-Feller pair.

Weak Convergence and Invariant probabilities. Assume $(\mathcal{X}, d)$ is a locally compact separable metric space. Let $\mathcal{P}(\mathcal{X})$ be the subset of probability measures in $\mathcal{M}(\mathcal{X})$. The sequence $\{\mu_t\}_{t \in \mathbb{T}_+}$ in $\mathcal{P}(\mathcal{X})$

converges weakly to $\mu \in \mathcal{P}(\mathcal{X})$ if

$$\lim_{t\to\infty} < f, \mu_t > = < f, \mu >, \quad \forall \; f \in C_b(\mathcal{X}) \tag{5}$$

Weak convergence is denoted by $\mu_t \Longrightarrow \mu$ and is also referred to as convergence in distribution. The weak topology on $\mathcal{P}(\mathcal{X})$ generated by weak convergence can be metrized. In particular, e.g., [26], one has the Prohorov metric d_p on $\mathcal{P}(\mathcal{X})$, such that the metric space $(\mathcal{P}(\mathcal{X}), d_p)$ is complete, separable, and a sequence $\{\mu_t\}_{t\in\mathbb{T}_+}$ in $\mathcal{P}(\mathcal{X})$ converges weakly to μ in $\mathcal{P}(\mathcal{X})$ *iff*

$$\lim_{t\to\infty} d_p(\mu_t, \mu) = 0 \tag{6}$$

Let (L, T) be a Markov-Feller pair on $(\mathcal{X}, d)$. A probability measure $\mu \in \mathcal{P}(\mathcal{X})$ is an invariant probability for T if $T\mu = \mu$. The operator T is uniquely ergodic if T has exactly one invariant probability. A probability measure μ^* is an attracting probability for T if, for any $\mu \in \mathcal{P}(\mathcal{X})$, the sequence $\{T^t\mu\}_{t\in\mathbb{T}_+}$ converges weakly to μ^*. In other words,

$$\lim_{t\to\infty} < f, T^t\mu > \;\; = \;\; < f, \mu^* > \;\; \forall \; f \in C_b(\mathcal{X}), \; \mu \in \mathcal{P}(\mathcal{X}) \tag{7}$$

It follows that, if T has an attracting probability μ^*, then T is uniquely ergodic ([25].)

II. Problem Formulation

A. System Model

We review the model of Kalman filtering with intermittent observations in [10]. Let

$$\mathbf{x}_{t+1} = A\mathbf{x}_t + \mathbf{w}_t \tag{8}$$

$$\mathbf{y}_t = C\mathbf{x}_t + \mathbf{v}_t \tag{9}$$

Here $\mathbf{x}_t \in \mathbb{R}^N$ is the signal (state) vector, $\mathbf{y}_t \in \mathbb{R}^M$ is the observation vector, $\mathbf{w}_t \in \mathbb{R}^N$ and $\mathbf{v}_t \in \mathbb{R}^M$ are Gaussian random vectors with zero mean and covariance matrices $Q \succeq 0$ and $R \gg 0$, respectively. The sequences $\{\mathbf{w}_t\}_{t\in\mathbb{T}_+}$ and $\{\mathbf{v}_t\}_{t\in\mathbb{T}_+}$ are uncorrelated and mutually independent. Also, assume that the initial state $\mathbf{x}_0$ is a zero-mean Gaussian vector with covariance P_0. The m.m.s.e. predictor $\widehat{\mathbf{x}}_{t|t-1}$ of the signal vector $\mathbf{x}_t$ given the observations $\{\mathbf{y}_s\}_{0\leq s<t}$ is the conditional mean. It is recursively implemented by the Kalman filter. The sequence of conditional prediction error covariances, $\{P_t\}_{t\in\mathbb{T}_+}$, is then given by

$$P_t \;\; = \;\; \mathbb{E}\left[\left(\mathbf{x}_t - \widehat{\mathbf{x}}_{t|t-1}\right)\left(\mathbf{x}_t - \widehat{\mathbf{x}}_{t|t-1}\right)^T \mid \{\mathbf{y}(s)\}_{0\leq s<t}\right] \tag{10}$$

$$P_{t+1} \;\; = \;\; AP_tA^T + Q - AP_tC^T\left(CP_tC^T + R\right)^{-1}CP_tA^T \tag{11}$$

Under the hypothesis of stabilizability of the pair (A, Q) and detectability of the pair (A, C), the deterministic sequence $\{P_t\}_{t\in\mathbb{T}_+}$ converges to a unique value P^* (which is a fixed point of the algebraic Riccati equation (11)) from any initial condition P_0.

This corresponds to the classical perfect observation scenario, where the estimator has complete knowledge of the observation packet $\mathbf{y}_t$ at every time t. With intermittent observations, the observation packets are dropped randomly (across the communication channel to the estimator), and the estimator receives observations at random times. We study the intermittent observation model considered in [10], where the channel randomness is modeled by a sequence $\{\gamma_t\}_{t\in\mathbb{T}_+}$ of i.i.d. Bernoulli random variables with mean $\overline{\gamma}$ (note, $\overline{\gamma}$ then denotes the arrival probability.) Here, $\gamma_t = 1$ corresponds to the arrival of the observation packet $\mathbf{y}_t$ at time t to the estimator, whereas a packet dropout corresponds to $\gamma_t = 0$. Denote by $\widetilde{\mathbf{y}}_t$ the pair

$$\widetilde{\mathbf{y}}_t = \left(\mathbf{y}_t \mathbb{I}_{(\gamma_t=1)}, \gamma_t\right) \tag{12}$$

Under the TCP packet acknowledgement protocol in [10] (the estimator knows at each time whether the observation packet arrived or not), the m.m.s.e. predictor of the signal is given by:

$$\widehat{\mathbf{x}}_{t|t-1} = \mathbb{E}\left[\mathbf{x}_t \mid \{\widetilde{\mathbf{y}}_s\}_{0\leq s<t}\right] \tag{13}$$

A modified form of the Kalman filter giving a recursive implementation of the estimator in eqn. (13) is in [10]. The sequence of conditional prediction error covariance matrices, $\{P_t\}_{t\in\mathbb{T}_+}$, is updated according to the following random algebraic Riccati equation (RRE):

$$P_t = \mathbb{E}\left[\left(\mathbf{x}_t - \widehat{\mathbf{x}}_{t|t-1}\right)\left(\mathbf{x}_t - \widehat{\mathbf{x}}_{t|t-1}\right)^T \mid \{\widetilde{\mathbf{y}}(s)\}_{0\leq s<t}\right] \tag{14}$$

$$P_{t+1} = AP_tA^T + Q - \gamma_t AP_tC^T\left(CP_tC^T + R\right)^{-1}CP_tA^T \tag{15}$$

Unlike the classical case, the sequence $\{P_t\}_{t\in\mathbb{T}_+}$ is now random (because of its dependence on the random sequence $\{\gamma_t\}_{t\in\mathbb{T}_+}$.) Thus, for each t, P_t is a random element of S_+^N, and we denote by $\mu_t^{\overline{\gamma},P_0}$ its distribution (the measure it induces on S_+^N.) The superscripts $\overline{\gamma}$, P_0 emphasize the dependence of $\mu_t^{\overline{\gamma},P_0}$ on the packet arrival probability and the initial condition.

In the subsequent sections, we analyze the random sequence $\{P_t\}_{t\in\mathbb{T}_+}$ governed by the RRE, eqn. (15), and establish its asymptotic properties including the weak convergence of the corresponding sequence $\left\{\mu_t^{\overline{\gamma},P_0}\right\}_{t\in\mathbb{T}_+}$ to a unique invariant distribution $\mu^{\overline{\gamma}}$ on S_+^N.

Before that, we set notation. Define the functions, $f_i : \mathbb{S}_+^N \longmapsto \mathbb{S}_+^N$, $i = 0, 1$, by

$$f_0(X) = AXA^T + Q \tag{16}$$

$$f_1(X) = AXA^T + Q - AXC^T\left(CXC^T + R\right)^{-1}CXA^T \tag{17}$$

Also, define $f : \{0,1\} \times \mathbb{S}_+^N \longmapsto \mathbb{S}_+^N$ by

$$\begin{aligned} f(\gamma, X) &= \mathbb{I}_0(\gamma)f_0(X) + \mathbb{I}_1(\gamma)f_1(X) \\ &= AXA^T + Q - \gamma AXC^T\left(CXC^T + R\right)^{-1}CXA^T \end{aligned} \tag{18}$$

Proposition 2 For a fixed $\gamma \in \{0,1\}$, if $R \gg 0$, the function $f(\gamma, X) : \mathbb{S}_+^N \longmapsto \mathbb{S}_+^N$ is continuous[2] in X. Also, $f(\cdot)$ is jointly measurable in γ, X.

For a fixed $\overline{\gamma}$, define the transition probability operator $\mathbb{Q}^{\overline{\gamma}} : \mathbb{S}_+^N \times \mathcal{B}(\mathbb{S}_+^N) \longmapsto [0,1]$ on the locally compact separable metric space $\mathbb{S}_+^N$ by

$$\mathbb{Q}^{\overline{\gamma}}(X, B) = (1-\overline{\gamma})\mathbb{I}_B\left(f_0(X)\right) + \overline{\gamma}\mathbb{I}_B\left(f_1(X)\right), \ \forall X \in \mathbb{S}_+^N, B \in \mathcal{B}\left(\mathbb{S}_+^N\right) \tag{19}$$

Now, consider the canonical path space of the sequence $\{P_t\}_{t \in \mathbb{T}_+}$,

$$\Omega^c = \times_{t=1}^{\infty} \mathbb{S}_+^N \tag{20}$$

and $\mathcal{F}^c$ be the corresponding product σ-algebra on Ω^c. For fixed $\overline{\gamma}$ and P_0, denote $\mathbb{P}^{\overline{\gamma}, P_0}$ to be the probability measure induced on $(\Omega^c, \mathcal{F}^c)$ by $\{P_t\}_{t \in \mathbb{T}_+}$. Then, in the sense of distribution induced on path space, the RRE generates a Markov process $\{P_t\}_{t \in \mathbb{T}_+}$ on $\left(\Omega^c, \mathcal{F}^c, \mathbb{P}^{\overline{\gamma}, P_0}\right)$, such that

$$\mathbb{P}^{\overline{\gamma}, P_0}\left(P_{t+1} \in B \mid P_t = X\right) = Q^{\overline{\gamma}}(X, B) \tag{21}$$

We denote the expectation operator associated with $\mathbb{P}^{\overline{\gamma}, P_0}$ by $\mathbb{E}^{\overline{\gamma}, P_0}$. For a fixed $\overline{\gamma}$, the family of measures $\left\{\mathbb{P}^{\overline{\gamma}, P_0}\right\}_{P_0 \in \mathbb{S}_+^N}$ on $(\Omega^c, \mathcal{F}^c)$ is called a Markov family.

Let $B\left(\mathbb{S}_+^N\right)$ be the Banach space of real-valued bounded functions on $\mathbb{S}_+^N$. For fixed $\overline{\gamma}$, define:

$$L^{\overline{\gamma}} : B\left(\mathbb{S}_+^N\right) \longmapsto B\left(\mathbb{S}_+^N\right) : \ \left(L^{\overline{\gamma}} g\right)(X) = \int_{\mathbb{S}_+^N} g(Y)\mathbb{Q}^{\overline{\gamma}}(X, dY), \ \forall g \in B\left(\mathbb{S}_+^N\right), X \in \mathbb{S}_+^N \tag{22}$$

$$T^{\overline{\gamma}} : \mathcal{M}\left(\mathbb{S}_+^N\right) \longmapsto \mathcal{M}\left(\mathbb{S}_+^N\right) : \ \left(T^{\overline{\gamma}} \mu\right)(B) = \int_{\mathbb{S}_+^N} \mathbb{Q}^{\overline{\gamma}}(Y, B)\mu(dY), \ \forall \mu \in \mathcal{M}\left(\mathbb{S}_+^N\right), B \in \mathcal{B}\left(\mathbb{S}_+^N\right) \tag{23}$$

We then have the following proposition (for a proof see the Appendix):

Proposition 3 For every $\overline{\gamma}$, $\left(L^{\overline{\gamma}}, T^{\overline{\gamma}}\right)$ is a Markov-Feller pair on $\mathbb{S}_+^N$.

Finally, we note, that

$$\mu_t^{\overline{\gamma}, P_0} = \left(T^{\overline{\gamma}}\right)^t \delta_{P_0}, \ \forall t \tag{24}$$

where δ_{P_0} denotes the Dirac probability measure concentrated at P_0.

B. Stability notions and critical probabilities

There are various stability notions for the random sequence $\{P_t\}_{t \in \mathbb{T}_+}$. In this work, we consider two: the first is stochastic boundedness (uniform boundedness in probability); and the second is the more stronger bounded in mean stability.

[2]As stated in Subsection IB, we may assume throughout that $\mathbb{S}^N$ is equipped with the induced 2-norm. However, as far as topological properties like continuity etc. are considered, the exact norm is not important as long as it makes $\mathbb{S}^N$ complete, because all norms on a finite dimensional linear space are equivalent, i.e., generate the same topology.

Definition 4 (Stochastic boundedness) Consider fixed $\overline{\gamma}$ and $P_0 \in \mathbb{S}_+^N$. The sequence $\{P_t\}_{t\in\mathbb{T}_+}$ is stochastically bounded (s.b.) if

$$\lim_{N\to\infty} \sup_{t\in\mathbb{T}_+} \mathbb{P}^{\overline{\gamma},P_0}\left(\|P_t\| > N\right) = 0 \tag{25}$$

The corresponding sequence of measures $\left\{\mu_t^{\overline{\gamma},P_0}\right\}_{t\in\mathbb{T}_+}$ is said to be tight (see [26].)

Definition 5 (Boundedness in mean) Consider fixed $\overline{\gamma}$ and $P_0 \in \mathbb{S}_+^N$. The sequence $\{P_t\}_{t\in\mathbb{T}_+}$ is bounded in mean (b.i.m.) if there exists $M^{\overline{\gamma},P_0}$, such that,

$$\sup_{t\in\mathbb{T}_+} \mathbb{E}^{\overline{\gamma},P_0}\left[P_t\right] \preceq M^{\overline{\gamma},P_0} \tag{26}$$

(Note, that the supremum above is taken w.r.t. the partial order in $\mathbb{S}^n$.)

We note here, that the above stability notions, applies to all systems irrespective of properties like stabilizability, detectability. Stochastic boundedness provides a trade-off between the permissible estimation error margin and performance guarantee uniformly over all time t. Stochastic boundedness is weaker than bounded in mean stability, as indicated by the following proposition (proof in the Appendix.)

Proposition 6 Consider fixed $\overline{\gamma}$ and $P_0 \in \mathbb{S}_+^N$. If the sequence $\{P_t\}_{t\in\mathbb{T}_+}$ is b.i.m., then it is s.b.

For most interesting cases, the stability of the sequence $\{P_t\}_{t\in\mathbb{T}_+}$ depends on $\overline{\gamma}$. Indeed, as exhibited in [10], there exist critical probabilities marking sharp transitions in the stability behavior. For the above stability notions, consider the following critical probabilities:

$$\overline{\gamma}^{\text{sb}} = \inf\left\{\overline{\gamma} \in [0,1] : \{P_t\}_{t\in\mathbb{T}_+} \text{ is s.b., } \forall P_0 \in \mathbb{S}_+^N\right\} \tag{27}$$

$$\overline{\gamma}^{\text{bim}} = \inf\left\{\overline{\gamma} \in [0,1] : \{P_t\}_{t\in\mathbb{T}_+} \text{ is b.i.m., } \forall P_0 \in \mathbb{S}_+^N\right\} \tag{28}$$

Thus, $\overline{\gamma}^{\text{sb}}$ marks a transition in stochastic boundedness of the sequence $\{P_t\}_{t\in\mathbb{T}_+}$, in the sense that, if the operating[3] $\overline{\gamma} > \overline{\gamma}^{\text{sb}}$, the sequence $\{P_t\}_{t\in\mathbb{T}_+}$ is s.b. for all initial conditions P_0, whereas it explodes if operated below $\overline{\gamma}^{\text{sb}}$. A similar interpretation holds for $\overline{\gamma}^{\text{bim}}$.

In [10], upper and lower bounds for $\overline{\gamma}^{\text{bim}}$ were obtained. Precisely, the following was shown:

Result 7 ([10]) For $\left(A, Q^{1/2}\right)$ stabilizable, (A, C) detectable, and A unstable, then $\exists\, \overline{\gamma}^{\text{bim}} \in [0,1)$, s.t.

$$\lim_{t\to\infty} \mathbb{E}^{\overline{\gamma},P_0}\left[P_t\right] = \infty, \text{ for } 0 \le \overline{\gamma} \le \overline{\gamma}^{\text{bim}} \text{ and } \exists\ P_0 \succeq 0 \tag{29}$$

$$\{P_t\}_{t\in\mathbb{T}_+} \text{ is b.i.m., for } \overline{\gamma}^{\text{bim}} < \overline{\gamma} \le 1 \text{ and } \forall\ P_0 \succeq 0 \tag{30}$$

[3]Note for stochastic boundedness, one has to operate strictly above $\overline{\gamma}^{\text{sb}}$, because the infimum in eqn. (27) may not be attainable.

Also, $\overline{\gamma}_l^{\text{bim}} \leq \overline{\gamma}^{\text{bim}} \leq \overline{\gamma}_u^{\text{bim}}$ where

$$\overline{\gamma}_l^{\text{bim}} = 1 - \frac{1}{\alpha^2} \tag{31}$$

$$\overline{\gamma}_u^{\text{bim}} = \inf\left\{\overline{\gamma} \in [0,1] : \exists(\widehat{K}, \widehat{X}), \text{ s.t. } \widehat{X} \gg \phi^{\overline{\gamma}} \left(\widehat{K}, \widehat{X}\right)\right\} \tag{32}$$

where α is the absolute value of the largest eigenvalue of A, and the operator $\phi^{\overline{\gamma}}$ is

$$\phi^{\overline{\gamma}}(K, X) = (1 - \overline{\gamma})\left(AXA^T + Q\right) + \overline{\gamma}\left(FXF^T + V\right) \tag{33}$$

and $F = A + KC$, $V = Q + KRK^T$.

The above result provides computable upper and lower bounds on the critical probability $\overline{\gamma}^{\text{bim}}$.

The following proposition relates the two critical probabilities:

Proposition 8 For $\left(A, Q^{1/2}\right)$ stabilizable, (A, C) detectable, and A unstable, then

i) For a general system $\overline{\gamma}^{\text{sb}} \leq \overline{\gamma}^{\text{bim}}$.
ii) If, in addition, $\left(A, Q^{1/2}\right)$ is stabilizable and (A, C) is detectable, then $\overline{\gamma}^{\text{sb}} = 0$.

Proof: The proofs of part [i] and part [ii] under the additional assumption of invertible C are provided in the Appendix. The proof of part [ii] for the general case of stabilizable and detectable systems can be found in the follow-up paper [24]. ■

From the above it is clear that, in general, any upper bound on $\overline{\gamma}^{\text{bim}}$ is also an upper bound on $\overline{\gamma}^{\text{sb}}$. Part ii) of the above proposition shows that under most reasonable assumptions $\overline{\gamma}^{\text{sb}} = 0$. For such systems, operating at any $\overline{\gamma} > 0$ guarantees stochastic boundedness[4], whereas if one needs to remain bounded in mean, one has to operate at $\overline{\gamma} > 1 - \frac{1}{\alpha^2}$, which can be large if A is highly unstable.This is important to the system designer, because, if the design criterion is stochastic boundedness (i.e., boundedness in probability) rather than boundedness in mean, one may operate at a value of $\overline{\gamma}$ strictly lower than $\overline{\gamma}^{\text{bim}}$.

In the next section, we state and discuss the main results of this paper. Among others, we show that operating above $\overline{\gamma}^{\text{sb}}$ guarantees the existence of a unique invariant distribution (independent of the initial condition P_0) to which the random sequence $\{P_t\}_{t \in \mathbb{T}_+}$ converges weakly.

III. MAIN RESULTS: INVARIANT DISTRIBUTION

The first result concerns the weak convergence properties of $\{P_t\}_{t \in \mathbb{T}_+}$ generated by the RRE.

Theorem 9 Assume: $\left(A, Q^{1/2}\right)$ stabilizable; (A, C) detectable; Q positive definite; fixed $\overline{\gamma}$ and $P_0 \in \mathbb{S}_+^N$. Then:

[4]We note here that for stable systems, the infimum in the definition of $\overline{\gamma}^{\text{sb}}$ is attained and the sequence $\{P_t\}$ is stochastically bounded for $\overline{\gamma} = 0$. However, for unstable systems to achieve stochastic boundedness, it is necessary to operate at $\overline{\gamma} > 0$, as $\overline{\gamma} = 0$ would imply that no observations arrive in the infinite horizon leading to a.s. unboundedness of the sequence $\{P_t\}$.

i) If $\overline{\gamma} > 0$, the sequence $\{P_t\}_{t\in\mathbb{T}_+}$ is stochastically bounded and there exists a unique invariant distribution $\mu^{\overline{\gamma}}$ s.t. the sequence $\{P_t\}_{t\in\mathbb{T}_+}$ (or sequence $\left\{\mu_t^{\overline{\gamma},P_0}\right\}_{t\in\mathbb{T}_+}$ of measures) converges weakly to $\mu^{\overline{\gamma}}$ from any initial condition P_0.

In other words, the operator $T^{\overline{\gamma}}$ is uniquely ergodic with attracting probability $\mu^{\overline{\gamma}}$.

ii) If, in addition, $\overline{\gamma} > \overline{\gamma}^{\text{bim}}$, the corresponding unique invariant measure $\mu^{\overline{\gamma}}$ has finite mean

$$\int_{\mathbb{S}_+^N} Y\mu^{\overline{\gamma}}(dY) < \infty \tag{34}$$

Theorem 9 states that, for stabilizable and detectable systems, if $\overline{\gamma} > 0$, the sequence $\{P_t\}_{t\in\mathbb{T}_+}$ converges in distribution to a unique invariant distribution, irrespective of the initial condition. In particular, one may operate below $\overline{\gamma}^{\text{bim}}$ and still converge to a unique invariant distribution. However, operating at $0 < \overline{\gamma} < \overline{\gamma}^{\text{bim}}$ may not guarantee that the corresponding invariant distribution $\mu^{\overline{\gamma}}$ has finite mean.

We discuss several implications of Theorem III. First, we note from the discussion following Proposition 8 (especially the footnote) that for stable systems, $\overline{\gamma} = 0$ also leads to stochastic boundedness of the sequence $\{P_t\}$, which does not hold for unstable systems. Thus for stable systems, the conclusions of Theorem III hold not only for $\overline{\gamma} > 0$, but alos for $\overline{\gamma} = 0$.

Finally, we note that, Theorem III as stated above in the context of stabilizable and detectable systems is, in fact, more general. As can be noted from the proof of Theorem III (Subsection VI-A), in general, a sufficient condition for the existence and uniqueness of an attracting invariant measure, is stochastic boundedness. Thus, for a general system (for which stabilizability, detectability may not be verified), operating above the critical probability $\overline{\gamma}^{\text{sb}}$ of stochastic boundedness is sufficient to guarantee weak convergence to a unique invariant distribution. In the case of stabilizable and detectable systems, by Proposition 8, $\overline{\gamma}^{\text{sb}} = 0$ and hence stochastic boundedness is ensured by operating under any $\overline{\gamma} > 0$.

The second result explicitly determines the support of the invariant measure $\mu^{\overline{\gamma}}$.

Theorem 10 Assume: $\left(A, Q^{1/2}\right)$ stabilizable; (A, C) detectable; Q positive definite. Define the set $\mathcal{S} \subset \mathbb{S}_+^N$ by

$$\mathcal{S} = \{f_{i_1} \circ f_{i_2} \circ \cdots \circ f_{i_s}(P^*) \mid i_r \in \{0,1\},\ 1 \le r \le s,\ s \in \mathbb{T}_+\} \tag{35}$$

where P^* is the fixed point of the (deterministic) Riccati equation Then[5], if $0 < \overline{\gamma} < 1$,

$$\text{supp}\left(\mu^{\overline{\gamma}}\right) = \text{cl}(\mathcal{S}) \tag{36}$$

where cl$(\mathcal{S})$ denotes the topological closure of $\mathcal{S}$ in $\mathbb{S}_+^N$. In particular, we have

$$\mu^{\overline{\gamma}}\left(\left\{Y \in \mathbb{S}_+^N \mid Y \succeq P^*\right\}\right) = 1 \tag{37}$$

Theorem 10 states that, for $0 < \overline{\gamma} < 1$, supp$\left(\mu^{\overline{\gamma}}\right)$ is independent of $\overline{\gamma}$ and is given by the closure of the countable set $\mathcal{S}$ (but the distribution is dependent on the value of $\overline{\gamma}$.) If $\overline{\gamma} = 1$, it reduces to the deterministic Kalman filtering, and the invariant measure is a Dirac mass at P^*.

[5] In eqn. 35 s can take the value 0, implying $P^* \in \mathcal{S}$.

The fact that the invariant measure is concentrated on the conic interval $[P^*, \infty)$, where P^* is the fixed point of the algebraic Riccati equation, eqn. (11), is quite natural (but not obvious), as one cannot expect to obtain better performance with intermittent observations.

The set $\mathcal{S}$ is not generally dense in $[P^*, \infty)$ and the support is an unbounded fractured (many holes) subset of $\mathbb{S}_+^N$. We study a scalar example to show, both analytically and numerically, that the invariant measure exhibits fractal properties, i.e., the support of the measure is a highly fractured subset of the positive reals and exhibits self-similarity.

The next three sections are devoted to the proofs of Theorems 9,10. The proof of Theorem 9 relies on the theory of random dynamical systems (RDS), and Theorem 10 uses the Markov-Feller property of the transition operator. Section IV summarizes results on RDS and models the RRE as an RDS. Section V establishes properties of the RRE as an RDS. We complete the proof of Theorem 9 in Subsection VI-A, whereas Theorem VI-B is proved in Subsection VI-B.

IV. Random Dynamical System Formulation

We start by defining a random dynamical system (RDS). We follow the notation in [19], [27].

Definition 11 (RDS) A RDS with (one-sided) time $\mathbb{T}_+$ and state space $\mathcal{X}$ is the pair (θ, φ):

A) A metric dynamical system $\theta = (\Omega, \mathcal{F}, \mathbb{P}, \{\theta_t, t \in \mathbb{T}\})$ with two-sided time $\mathbb{T}$, i.e., a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with a family of transformations $\{\theta_t : \Omega \longmapsto \Omega\}_{t \in \mathbb{T}}$ such that[6]

- **A.1)** $\theta_0 = id_\Omega, \ \theta_t \circ \theta_s = \theta_{t+s}, \ \forall t, s \in \mathbb{T}$
- **A.2)** $(t, \omega) \longmapsto \theta_t \omega$ is measurable.
- **A.3)** $\theta_t \mathbb{P} = \mathbb{P} \, \forall t \in \mathbb{T}$, i.e., $\mathbb{P}(\theta_t B) = \mathbb{P}(B)$ for all $B \in \mathcal{F}$ and all $t \in \mathbb{T}$.

B) A cocycle φ over θ of continuous mappings of $\mathcal{X}$ with time $\mathbb{T}_+$, i.e., a measurable mapping

$$\varphi : \mathbb{T}_+ \times \Omega \times \mathcal{X} \to \mathcal{X}, \ (t, \omega, X) \longmapsto \varphi(t, \omega, X) \tag{38}$$

- **B.1)** The mapping $X \longmapsto \varphi(t, \omega, X) \equiv \varphi(t, \omega) X$ is continuous in $X \ \forall \, t \in \mathbb{T}_+, \ \omega \in \Omega$.
- **B.2)** The mappings $\varphi(t, \omega) \doteq \varphi(t, \omega, \cdot)$ satisfy the cocycle property:

$$\varphi(0, \omega) = id_{\mathcal{X}}, \ \ \varphi(t+s, \omega) = \varphi(t, \theta_s \omega) \circ \varphi(s, \omega), \ \ \forall \, t, s \in \mathbb{T}_+, \ \omega \in \Omega \tag{39}$$

In a RDS, randomness is captured by the space $(\Omega, \mathcal{F}, \mathbb{P})$. Iterates indexed by ω indicate pathwise construction. For example, if X_0 is the deterministic state at $t = 0$, the state at $t \in \mathbb{T}_+$ is

$$X_t(\omega) = \varphi(t, \omega, X_0) \tag{40}$$

The measurability assumptions guarantee that the state X_t is a well-defined random variable. Also, the iterates are defined for non-negative (one-sided) time; however, the family of transformations $\{\theta_t\}$ is two-sided, which is purely for technical convenience, as will be seen later.

[6]The function id_Ω denotes the identity map on Ω, i.e., for all $\omega \in \Omega$, $id_\Omega(\omega) = \omega$.

We now show that the sequence $\{P_t\}$ generated by the RRE can be modeled as the sequence of iterates (in the sense of distributional equivalence) of a suitably defined RDS.

Fix $\overline{\gamma}$ and define: $\left(\widetilde{\Omega}, \widetilde{\mathcal{F}}, \widetilde{\mathbb{P}}^{\overline{\gamma}}\right)$, where $\widetilde{\Omega} = \{0,1\}$, $\widetilde{\mathcal{F}} = 2^{\{0,1\}}$ and $\widetilde{\mathbb{P}}^{\overline{\gamma}}(\{1\}) = \overline{\gamma}$; and the product space, $\left(\Omega, \mathcal{F}, \mathbb{P}^{\overline{\gamma}}\right)$, where $\Omega = \times_{t\in\mathbb{T}}\widetilde{\Omega}$ and $\mathcal{F}$ and $\mathbb{P}^{\overline{\gamma}}$ are the product σ-algebra and the product measure[7]. From the construction, a sample point $\omega \in \Omega$ is a two-sided binary sequence and, since $\mathbb{P}^{\overline{\gamma}}$ is the product of $\widetilde{\mathbb{P}}^{\overline{\gamma}}$, the projections are i.i.d. binary random variables with probability of one being $\overline{\gamma}$. Define the family of transformations $\left\{\theta_t^R\right\}_{t\in\mathbb{T}}$ on Ω as the family of left-shifts

$$\theta_t^R\omega(\cdot) = \omega(t+\cdot),\ \forall t \in \mathbb{T} \tag{41}$$

With this, the space $\left(\Omega, \mathcal{F}, \mathbb{P}^{\overline{\gamma}}, \left\{\theta_t^R, t\in\mathbb{T}\right\}\right)$ is the canonical path space of a two-sided stationary (in fact, i.i.d.) sequence equipped with the left-shift operator; hence, (e.g., [28]) it satisfies the Assumptions **A.1)-A.3)** to be a metric dynamical system; in fact, it is ergodic.

Recall the Riccati iterates $f(\gamma, X)$ in eqn. (18). Define the function $\widetilde{f} : \Omega \times \mathbb{S}_+^N \longmapsto \mathbb{S}_+^N$ by

$$\widetilde{f}(\omega, X) = f(\omega(0), X) \tag{42}$$

Since the projection map from ω to $\omega(0)$ is measurable (continuous) and $f(\cdot)$ is jointly measurable in γ, X (Proposition 2), it follows that $\widetilde{f}(\cdot)$ is jointly measurable in ω, X. Define the function φ^R : $\mathbb{T}_+ \times \Omega \times \mathbb{S}_+ \longmapsto \mathbb{S}_+$ by

$$\varphi^R(0,\omega,X) = X, \qquad \forall\ \omega, X \tag{43}$$

$$\varphi^R(1,\omega,X) = \widetilde{f}(\omega, X), \qquad \forall\ \omega, X \tag{44}$$

$$\varphi^R(t,\omega,X) = \widetilde{f}\left(\theta_{t-1}^R\omega, \varphi^R(t-1,\omega,X)\right),\ \forall\ t>1, \omega, X \tag{45}$$

It follows from the measurability of the transformations $\left\{\theta_t^R\right\}$, the measurability of $\widetilde{f}(\cdot)$, and the fact that $\mathbb{T}_+$ is countable that the function $\varphi^R(t,\omega,X)$ is jointly measurable in t, ω, X. Finally, $\varphi^R(\cdot)$ defined above satisfies Assumption **B.2)** by virtue of Proposition 2, and Assumption **B.3)** follows by the construction given by eqns. (43-45). Thus, the pair $\left(\theta^R, \varphi^R\right)$ is an RDS over $\mathbb{S}_+^N$. Given a deterministic initial condition $P_0 \in \mathbb{S}_+^N$, it follows that the sequence $\{P_t\}_{t\in\mathbb{T}_+}$ generated by the RRE eqn. (15) is equivalent in the sense of distribution to the sequence $\left\{\varphi^R(t,\omega,P_0)\right\}_{t\in\mathbb{T}_+}$ generated by the iterates of the above constructed RDS, i.e.,

$$P_t \stackrel{d}{=} \varphi^R(t,\omega,P_0),\ \forall t \in \mathbb{T}_+ \tag{46}$$

Indeed, by studying eqn. (45), we note that the iterate $\varphi^R(t,\omega,P_0)$ at time t is obtained by applying the map $f_{\omega_{t-1}}$ to $\varphi^R(t-1,\omega,P_0)$ and by construction, the random variable ω_{t-1} is 1 with probability $\overline{\gamma}$ and 0 with probability $1-\overline{\gamma}$. Thus, investigating the distributional properties of $\{P_t\}_{t\in\mathbb{T}_+}$ is equivalent to analyzing the distributional properties of $\left\{\varphi^R(t,\omega,P_0)\right\}_{t\in\mathbb{T}_+}$, which we carry out in the rest of the paper.

[7]Note the difference between the measure $\mathbb{P}^{\overline{\gamma}}$ and the measures $\mathbb{P}^{\overline{\gamma},P_0}$ defined in Subsection II-A.

In the sequel, we use the pair (θ, φ) to denote a generic RDS and $\left(\theta^R, \varphi^R\right)$ for the one constructed above for the RRE.

V. Properties of the RDS (θ, φ)

A. Facts about generic RDS

We review concepts on RDS (see [19], [27] for details.) Consider a generic RDS (θ, φ) with state space $\mathcal{X}$ as in Definition 11. Assume that $\mathcal{X}$ is a non-empty subset of a real Banach space V with a closed, convex, solid, normal (w.r.t. the Banach space norm), minihedral cone V_+. Denote by $\preceq$ the partial order induced by V_+ in $\mathcal{X}$ and $<<$ denotes the corresponding strong order. Although the development that follows may hold for arbitrary $\mathcal{X} \subset V$, in the sequel, we assume $\mathcal{X} = V_+$ (which is true for the RDS $\left(\theta^R, \varphi^R\right)$ modeling the RRE.)

Definition 12 (Order-Preserving RDS) A RDS (θ, φ) with state space V_+ is order-preserving if

$$X \preceq Y \implies \varphi(t, \omega, X) \preceq \varphi(t, \omega, Y),\ \forall t \in \mathbb{T}_+,\ \omega \in \Omega,\ X, Y \in V_+ \tag{47}$$

Definition 13 (Sublinearity) An order-preserving RDS (θ, φ) with state space V_+ is sublinear if for every $X \in V_+$ and $\lambda \in (0,1)$ we have

$$\lambda\varphi(t, \omega, X) \preceq \varphi(t, \omega, \lambda X),\ \forall t > 0,\ \omega \in \Omega \tag{48}$$

The RDS is strongly sublinear if in addition to eqn. (48), we have

$$\lambda\varphi(t, \omega, X) \ll \varphi(t, \omega, \lambda X),\ \forall t > 0,\ \omega \in \Omega,\ X \in \operatorname{int} V_+ \tag{49}$$

Definition 14 (Equilibrium) A random variable $u : \Omega \longmapsto V_+$ is called an equilibrium (fixed point, stationary solution) of the RDS (θ, φ) if it is invariant under φ, i.e.,

$$\varphi\left(t, \omega, u(\omega)\right) = u\left(\theta_t \omega\right),\ \forall t \in \mathbb{T}_+,\ \omega \in \Omega \tag{50}$$

If eqn. (50) holds $\forall\, \omega \in \Omega$, except on set of $\mathbb{P}$ measure zero, u is an almost equilibrium.

Since the transformations $\{\theta_t\}$ are measure-preserving, i.e., $\theta_t \mathbb{P} = \mathbb{P}$, $\forall t$, we have

$$u\left(\theta_t \omega\right) \stackrel{d}{=} u(\omega),\ \forall t \tag{51}$$

By eqn. (50), for an almost equilibrium u, the iterates in the sequence $\{\varphi\left(t, \omega, u(\omega)\right)\}_{t \in \mathbb{T}_+}$ have the same distribution, which is the distribution of u.

Definition 15 (Orbit) For a random variable $u : \Omega \longmapsto V_+$, we define the *forward* orbit $\eta_u^f(\omega)$ emanating from $u(\omega)$ as the random set $\{\varphi\left(t, \omega, u(\omega)\right)\}_{t \in \mathbb{T}_+}$. The forward orbit gives the sequence of iterates of the RDS starting at u.

Although η_u^f is the object of interest, for technical convenience (as will be seen later), we also define the *pull-back* orbit $\eta_u^b(\omega)$ emanating from u as the random set $\{\varphi\left(t, \theta_{-t}\omega, u\left(\theta_{-t}\omega\right)\right)\}_{t \in \mathbb{T}_+}$.

We establish asymptotic properties for the pull-back orbit η_u^b. This is because it is more convenient and because analyzing η_u^b leads to understanding the asymptotic distributional properties for η_u^f. In fact, the random sequences $\{\varphi(t,\omega,u(\omega))\}_{t\in\mathbb{T}_+}$ and $\{\varphi(t,\theta_{-t}\omega,u(\theta_{-t}\omega))\}_{t\in\mathbb{T}_+}$ are equivalent in distribution. In other words,

$$\varphi(t,\omega,u(\omega)) \stackrel{d}{=} \varphi(t,\theta_{-t}\omega,u(\theta_{-t}\omega)), \ \forall t\in\mathbb{T}_+ \tag{52}$$

This follows from $\theta_t\mathbb{P}=\mathbb{P}$, $\forall t\in\mathbb{T}$ (hence the random objects ω and $\theta_t\omega$ possess the same distribution.) Thus, in particular, we have the following assertion.

Lemma 16 Let the sequence $\{\varphi(t,\theta_{-t}\omega,u(\theta_{-t}\omega))\}_{t\in\mathbb{T}_+}$ converge in distribution to a measure μ on V_+, where $u:\Omega\longmapsto V_+$ is a random variable. Then the sequence $\{\varphi(t,\omega,u(\omega))\}_{t\in\mathbb{T}_+}$ also converges in distribution to the measure μ.

We now introduce some notions of boundedness of RDS, which will be used in the sequel.

Definition 17 (Boundedness) Let $a:\Omega\longmapsto V_+$ be a random variable. The pull-back orbit $\eta_a^b(\omega)$ emanating from a is bounded on $U\in\mathcal{F}$ if there exists a random variable C on U s.t.

$$\|\varphi(t,\theta_{-t}\omega,a(\theta_{-t}\omega))\|\leq C(\omega), \ \forall t\in\mathbb{T}_+, \ \omega\in U \tag{53}$$

Definition 18 (Conditionally Compact RDS) An RDS (θ,φ) in V_+ is conditionally compact if for any $U\in\mathcal{F}$ and pull-back orbit $\eta_a^b(\omega)$ that is bounded on U there exists a family of compact sets $\{K(\omega)\}_{\omega\in U}$ s.t.

$$\lim_{t\to\infty}\operatorname{dist}\left(\varphi(t,\theta_{-t}\omega,a(\theta_{-t}\omega)),K(\omega)\right)=0, \ \omega\in U \tag{54}$$

It is to be noted that conditionally compact is a topological property of the space V_+. In particular, an RDS in a finite dimensional space V_+ is conditionally compact.

We now state a limit set dichotomy result for a class of sublinear, order-preserving RDS.

Theorem 19 (Corollary 4.3.1. in [27]) Let V be a separable Banach space with a normal solid cone V_+. Assume that (θ,φ) is a strongly sublinear conditionally compact order-preserving RDS over an ergodic metric dynamical system θ. Suppose that $\varphi(t,\omega,0)\gg 0$ for all $t>0$ and $\omega\in\Omega$. Then precisely one of the following applies:

(a) For any $X\in V_+$ we have

$$\mathbb{P}\left(\lim_{t\to\infty}\|\varphi(t,\theta_{-t}\omega,X)\|=\infty\right)=1 \tag{55}$$

(b) There exists a unique almost equilibrium $u(\omega)\gg 0$ defined on a θ-invariant set[8] $\Omega^*\in\mathcal{F}$ with $\mathbb{P}(\Omega^*)=1$ such that, for any random variable $v(\omega)$ possessing the property $0\preceq v(\omega)\preceq \alpha u(\omega)$ for all $\omega\in\Omega^*$ and deterministic $\alpha>0$, the following holds:

$$\lim_{t\to\infty}\varphi(t,\theta_{-t}\omega,v(\theta_{-t}\omega))=u(\omega), \ \omega\in\Omega^* \tag{56}$$

[8]A set $A\in\mathcal{F}$ is called θ-invariant if $\theta_tA=A$ for all $t\in\mathbb{T}$.

B. Properties of the Riccati RDS

In this subsection we establish some properties of the RDS $\left(\theta^{R}, \varphi^{R}\right)$ modeling the RRE.

Lemma 20 The RDS $\left(\theta^{R}, \varphi^{R}\right)$ with state space $\mathbb{S}_{+}^{N}$ is order-preserving. In other words,

$$X \preceq Y \implies \varphi^{R}(t, \omega, X) \preceq \varphi^{R}(t, \omega, Y),\ \forall t \in \mathbb{T}_{+},\ \omega \in \Omega,\ X, Y \in \mathbb{S}_{+}^{N} \tag{57}$$

Also, if Q is positive definite, i.e., $Q \gg 0$, it is strongly sublinear.

Proof: We establish order-preserving. Eqn. (57) holds for $t=0$, because by definition

$$\varphi^{R}(0, \omega, \cdot) = id_{S_{+}^{N}},\ \forall \omega \in \Omega \tag{58}$$

Consider $t=1$. From eqn. (44) we have

$$\varphi^{R}(1, \omega, X) = f(\omega(0), X),\ \forall \omega \in \Omega,\ X \in \mathbb{S}_{+}^{N} \tag{59}$$

where $f(\cdot)$ is defined in eqn. (18). From [10] (Lemma 1, part (c)), we note that, for fixed $\gamma \in \{0,1\}$, the function $f_{\gamma}(\cdot) = f(\gamma, \cdot)$ is order-preserving in X, i.e.,

$$X \preceq Y \implies f_{\gamma}(X) \preceq f_{\gamma}(Y),\ \forall X, Y \in \mathbb{S}_{+}^{N} \tag{60}$$

Hence, for a given $\omega \in \Omega$, we have from eqns. (59,60), if $X \preceq Y$,

$$\varphi^{R}(1, \omega, X) = f_{\omega(0)}(X) \preceq f_{\omega(0)}(Y) = \varphi^{R}(1, \omega, Y) \tag{61}$$

Thus, the order-preserving property is established for $t=1$. For $t>1$, we have from eqn. (45)

$$\varphi^{R}(t, \omega, X) = f_{\omega(t-1)} \circ f_{\omega(t-2)} \circ \cdots \circ f_{\omega(0)}(X) \tag{62}$$

For $\omega \in \Omega$, the functions $\{f_{\omega(i)}(\cdot)\}_{0 \leq t-1}$ are order-preserving by eqn. (60). Since the composition of order-preserving functions remains order-preserving, from eqn. (62) the function $\varphi^{R}(t, \omega, \cdot)$ is order-preserving in X. This establishes the order-preserving of the RDS $\left(\theta^{R}, \varphi^{R}\right)$.

We now establish strong sublinearity when $Q \gg 0$. Fix $\gamma \in \{0,1\}$, $\lambda \in (0,1)$. Then, from the concavity of $f_{\gamma}(\cdot)$ (Lemma 1, part (e) in [10]), we have

$$\lambda f_{\gamma}(X) + (1-\lambda) f_{\gamma}(0) \preceq f_{\gamma}(\lambda X),\ \ \forall\ X \in \mathbb{S}_{+}^{N} \tag{63}$$

Again, from [10] (Lemma 1 part (f)), we have

$$Q \preceq f_{\gamma}(0) \tag{64}$$

Under the assumption $Q \gg 0$ and $\lambda \in (0,1)$, we have from eqn. (64)

$$(1-\lambda) f_{\gamma}(0) \succeq (1-\lambda) Q \gg 0$$

From eqns. (63,65), we then have for every $\lambda \in (0,1)$ and $\gamma \in \{0,1\}$

$$\lambda f_{\gamma}(X) \ll f_{\gamma}(\lambda X),\ X \in \mathbb{S}_{+}^{N} \tag{65}$$

We then have from eqn. (65) for all $X \in \mathbb{S}_{+}^{N}$, $\lambda \in (0,1)$, $\omega \in \Omega$, and $t = 1$

$$\lambda \varphi^{R}(1, \omega, X) = \lambda f_{\omega(0)}(X) \ll f_{\omega(0)}(\lambda X) = \varphi^{R}(1, \omega, \lambda X) \tag{66}$$

which establishes strong sublinearity for $t = 1$ (the above is *stronger* than strong sublinearity, as given by Definition 13, since the latter requires $\ll$ to hold only for $X \in \mathbb{S}_{++}^{N}$.) We extend it to $t > 1$ by induction. Assume that the property in eqn. (66) (which implies strong sublinearity) holds for $t = s > 0$. We now show that it holds for $t = s+1$. Indeed, for $X \in \mathbb{S}_{+}^{N}$

$$\begin{aligned} \lambda \varphi^{R}(s+1, \omega, X) &= \lambda f_{\omega(s)}\left(\varphi^{R}(s, \omega, X)\right) \ll f_{\omega(s)}\left(\lambda \varphi^{R}(s, \omega, X)\right) \\ &\preceq f_{\omega(s)}\left(\varphi^{R}(s, \omega, \lambda X)\right) = \varphi^{R}(s+1, \omega, \lambda X) \end{aligned} \tag{67}$$

where the second step follows from eqn. (65) and the third from the induction step

$$\lambda \varphi^{R}(s, \omega, X) \ll \varphi^{R}(s, \omega, \lambda X) \tag{68}$$

and the fact that $f_{\omega(s)}(\cdot)$ is order-preserving. Thus we have strong sublinearity. ■

VI. Proofs of Theorems 9,10

A. Proof of Theorem 9

Lemma 21 Consider the RDS $\left(\theta^{R}, \varphi^{R}\right)$. Assume: $\overline{\gamma} \in (\overline{\gamma}^{\mathrm{sb}}, 1]$; Q positive definite. Then there exists unique almost equilibrium $u^{\overline{\gamma}}(\omega) \gg 0$ defined on a θ^{R}-invariant set $\Omega^{*} \in \mathcal{F}$ with $\mathbb{P}^{\overline{\gamma}}\left(\Omega^{*}\right) = 1$ s.t. for any random variable $v(\omega)$ possessing the property $0 \preceq v(\omega) \preceq \alpha u^{\overline{\gamma}}(\omega)\, \forall\, \omega \in \Omega^{*}$ and deterministic $\alpha > 0$, the following holds:

$$\lim_{t \to \infty} \varphi^{R}\left(t, \theta_{-t}^{R}\omega, v\left(\theta_{-t}^{R}\omega\right)\right) = u^{\overline{\gamma}}(\omega),\ \omega \in \Omega^{*} \tag{69}$$

Proof: From Lemma 20, $\left(\theta^{R}, \varphi^{R}\right)$ is strongly sublinear and order-preserving. It is conditionally compact because the space $\mathbb{S}_{+}^{N}$ is finite dimensional. Also, the cone $\mathbb{S}_{+}^{N}$ satisfies the conditions required in the hypothesis of Theorem 19. From Lemma 1f) in [10], we note for $t > 0$

$$\varphi^{R}\left(t, \omega, 0\right) = f_{\omega(t-1)}\left(\varphi^{R}(t-1, \omega, 0)\right) \succeq Q \gg 0 \tag{70}$$

Thus the hypotheses of Theorem 19 are satisfied and precisely one of the assertions **a)** or **b)** holds. We show assertion **a)** does not hold. Assume that **a)** holds on the contrary. Then, there exists $P_0 \in \mathbb{S}_{+}^{N}$ such that

$$\mathbb{P}^{\overline{\gamma}}\left(\lim_{t \to \infty} \left\|\varphi^{R}\left(t, \theta_{-t}^{R}\omega, P_0\right)\right\| = \infty\right) = 1 \tag{71}$$

Then, for every $N \in \mathbb{T}_{+}$, we have

$$\lim_{t \to \infty} \left\|\varphi^{R}\left(t, \theta_{-t}^{R}\omega, P_0\right)\right\| > N,\ \ \mathbb{P}^{\overline{\gamma}}\ \text{a.s.} \tag{72}$$

In other words, the sequence $\left\{\mathbb{I}_{(N,\infty)}\left(\left\|\varphi^{R}\left(t,\theta^{R}_{-t}\omega,P_{0}\right)\right\|\right)\right\}_{t\in\mathbb{T}_{+}}$ satisfies

$$\mathbb{P}^{\overline{\gamma}}\left(\lim_{t\to\infty}\mathbb{I}_{(N,\infty)}\left(\left\|\varphi^{R}\left(t,\theta^{R}_{-t}\omega,P_{0}\right)\right\|\right)=1\right)=1 \tag{73}$$

Since convergence a.s. implies convergence in probability, we have for every $\varepsilon>0$,

$$\lim_{t\to\infty}\mathbb{P}^{\overline{\gamma}}\left(\left|\mathbb{I}_{(N,\infty)}\left(\left\|\varphi^{R}\left(t,\theta^{R}_{-t}\omega,P_{0}\right)\right\|\right)-1\right|\leq\varepsilon\right)=1 \tag{74}$$

Since $\mathbb{I}_{(N,\infty)}(\cdot)$ takes the values $\{0,1\}$, eqn. (74) implies

$$\lim_{t\to\infty}\mathbb{P}^{\overline{\gamma}}\left(\mathbb{I}_{(N,\infty)}\left(\left\|\varphi^{R}\left(t,\theta^{R}_{-t}\omega,P_{0}\right)\right\|\right)=1\right)=1 \tag{75}$$

We thus have

$$\lim_{t\to\infty}\mathbb{P}^{\overline{\gamma}}\left(\left\|\varphi^{R}\left(t,\theta^{R}_{-t}\omega,P_{0}\right)\right\|>N\right)=1 \tag{76}$$

Since the above holds for every $N\in\mathbb{T}_{+}$, we have

$$\lim_{N\to\infty}\sup_{t\in\mathbb{T}_{+}}\mathbb{P}^{\overline{\gamma}}\left(\left\|\varphi^{R}\left(t,\theta^{R}_{-t}\omega,P_{0}\right)\right\|>N\right)=1 \tag{77}$$

On the other hand, $\overline{\gamma}>\overline{\gamma}^{\text{sb}}$ and Lemma 16 both imply

$$\begin{aligned}\lim_{N\to\infty}\sup_{t\in\mathbb{T}_{+}}\mathbb{P}^{\overline{\gamma}}\left(\left\|\varphi^{R}\left(t,\theta^{R}_{-t}\omega,P_{0}\right)\right\|>N\right) &= \lim_{N\to\infty}\sup_{t\in\mathbb{T}_{+}}\mathbb{P}^{\overline{\gamma}}\left(\left\|\varphi^{R}\left(t,\omega,P_{0}\right)\right\|>N\right)\\ &= \lim_{N\to\infty}\sup_{t\in\mathbb{T}_{+}}\mathbb{P}^{\overline{\gamma},P_{0}}\left(\left\|P_{t}\right\|>N\right)\\ &= 0\end{aligned} \tag{78}$$

This contradicts eqn. (72) and **a)** does not hold. Thus **b)** holds, and we have the result. ■

Lemma 21 establishes the existence of a unique almost equilibrium $u^{\overline{\gamma}}$ if $\overline{\gamma}^{\text{sb}}<\overline{\gamma}\leq 1$. From the distributional equivalence of pull-back and forward orbits, it follows that, for the Markov-Feller pair $(L^{\overline{\gamma}},T^{\overline{\gamma}})$, $T^{\overline{\gamma}}$ is uniquely ergodic. However, to show that the measure induced by $u^{\overline{\gamma}}$ on $\mathbb{S}^{N}_{+}$ is attracting for $T^{\overline{\gamma}}$, eqn. (69) must hold for all initial v. Lemma 21 establishes convergence for a restricted class of initial conditions v. We need to extend it to general initial conditions.

Lemma 22 For $\overline{\gamma}^{\text{sb}}<\overline{\gamma}\leq 1$ let $u^{\overline{\gamma}}$ be an almost equilibrium of the RDS $\left(\theta^{R},\varphi^{R}\right)$. Then

$$\mathbb{P}^{\overline{\gamma}}\left(\omega:u^{\overline{\gamma}}(\omega)\succeq Q\right)=1 \tag{79}$$

Proof: By the definition of an almost equilibrium (see Definition 14) we have

$$\mathbb{P}^{\overline{\gamma}}\left(\omega:\varphi^{R}\left(1,\omega,u^{\overline{\gamma}}(\omega)\right)\succeq Q\right)=\mathbb{P}^{\overline{\gamma}}\left(\omega:u^{\overline{\gamma}}(\theta^{R\overline{\gamma}}_{1}\omega)\succeq Q\right)=\mathbb{P}^{\overline{\gamma}}\left(\omega:u^{\overline{\gamma}}(\omega)\succeq Q\right) \tag{80}$$

Again, by Lemma 1f) in [10], we have $\mathbb{P}^{\overline{\gamma}}$ a.s.

$$\varphi^{R}\left(1,\omega,u^{\overline{\gamma}}(\omega)\right)=f_{\omega(0)}(u^{\overline{\gamma}}(\omega))\succeq Q \tag{81}$$

Since eqn. (81) holds $\mathbb{P}^{\overline{\gamma}}$ a.s., we have

$$\mathbb{P}^{\overline{\gamma}}\left(\omega : \varphi^{R}\left(1,\omega,u^{\overline{\gamma}}(\omega)\right) \succeq Q\right) = 1 \tag{82}$$

The Lemma then follows from eqns. (80,82). ■

Proof of Theorem 9: We now complete the proof of Theorem 9. The key step consists of finding a suitable modification $\widetilde{X}(\omega)$ of the initial condition P_0, such that $\widetilde{X}(\omega) = P_0$ a.s. and there exists a deterministic $\alpha > 0$ satisfying $0 \preceq \widetilde{X}(\omega) \preceq \alpha u^{\overline{\gamma}}(\omega)$. In that case, we can invoke Lemma 21 to establish weak convergence of the sequence $\{\varphi^{R}\left(t,\omega,\widetilde{X}(\omega)\right)\}_{t\in\mathbb{T}_+}$ with initial condition $\widetilde{X}(\omega)$ to $\mu^{\overline{\gamma}}$. Since $\widetilde{X}(\omega)$ is a.s. equal to P_0, this would allow us to deduce the weak convergence of the desired sequence $\{\varphi^{R}\left(t,\omega,P_0\right)\}_{t\in\mathbb{T}_+}$. We detail such a construction in the following.

For $\overline{\gamma} > \overline{\gamma}^{\mathrm{sb}}$ (note that $\overline{\gamma}^{\mathrm{sb}} = 0$ under the assumptions of Theorem III) let $\mu^{\overline{\gamma}}$ be the distribution of the unique almost equilibrium in Lemma 21. By Lemma 22 we have

$$\mu^{\overline{\gamma}}\left(\mathbb{S}^{N}_{++}\right) = 1 \tag{83}$$

since $u^{\overline{\gamma}}(\omega) \preceq Q \gg 0$ a.s. Let $P_0 \in \mathbb{S}^{N}_{+}$ be an arbitrary initial state. By construction of the RDS $\left(\theta^{R},\varphi^{R}\right)$, the sequences $\{P_t\}_{t\in\mathbb{T}_+}$ and $\{\varphi^{R}\left(t,\omega,P_0\right)\}_{t\in\mathbb{T}_+}$ are distributionally equivalent, i.e.,

$$P_t \stackrel{d}{=} \varphi^{R}\left(t,\omega,P_0\right) \tag{84}$$

Recall Ω^{*} as the θ^{R}-invariant set with $\mathbb{P}^{\overline{\gamma}}(\Omega^{*}) = 1$ in Lemma 21 on which the almost equilibrium $u^{\overline{\gamma}}$ is defined. By Lemma 22, there exists $\Omega_1 \subset \Omega^{*}$ with $\mathbb{P}^{\overline{\gamma}}(\Omega_1) = 1$, such that

$$u^{\overline{\gamma}}(\omega) \succeq Q,\ \omega \in \Omega_1 \tag{85}$$

Define the random variable $\widetilde{X} : \Omega \longmapsto \mathbb{S}^{N}_{+}$ by

$$\begin{cases} P_0 & \text{if } \omega \in \Omega_1 \\ 0 & \text{if } \omega \in \Omega_1^{c} \end{cases} \tag{86}$$

Now choose $\alpha > 0$ sufficiently large, such that,

$$P_0 \preceq \alpha Q \tag{87}$$

This is possible because $Q \gg 0$. Then

$$0 \preceq \widetilde{X}(\omega) \preceq \alpha u^{\overline{\gamma}}(\omega),\ \omega \in \Omega^{*} \tag{88}$$

Indeed, we have

$$0 \preceq P_0 \;=\; \widetilde{X}(\omega) \preceq \alpha Q \preceq \alpha u^{\overline{\gamma}}(\omega),\ \omega \in \Omega_1 \tag{89}$$

$$0 = \widetilde{X}(\omega) \;\preceq\; \alpha u^{\overline{\gamma}}(\omega),\ \omega \in \Omega\backslash\Omega_1 \tag{90}$$

Then, by Lemma 21

$$\lim_{t\to\infty} \varphi^R\left(t, \theta^R_{-t}\omega, \widetilde{X}\left(\theta^R_{-t}\omega\right)\right) = u^{\overline{\gamma}}(\omega),\ \omega \in \Omega^* \tag{91}$$

Since convergence $\mathbb{P}^{\overline{\gamma}}$ a.s. implies convergence in distribution, we have

$$\varphi^R\left(t, \theta^R_{-t}\omega, \widetilde{X}\left(\theta^R_{-t}\omega\right)\right) \Longrightarrow \mu^{\overline{\gamma}} \tag{92}$$

as $t \to \infty$, where $\Longrightarrow$ denotes weak convergence or convergence in distribution. Then by Lemma 16, the sequence $\left\{\varphi^R\left(t, \omega, \widetilde{X}(\omega)\right)\right\}_{t\in\mathbb{T}_+}$ also converges in distribution to the unique stationary distribution $\mu^{\overline{\gamma}}$, i.e., as $t \to \infty$

$$\varphi^R\left(t, \omega, \widetilde{X}(\omega)\right) \Longrightarrow \mu^{\overline{\gamma}} \tag{93}$$

Now, since $\mathbb{P}^{\overline{\gamma}}(\Omega_1) = 1$, by eqn. (86)

$$\varphi^R\left(t, \omega, P_0\right) = \varphi^R\left(t, \omega, \widetilde{X}(\omega)\right),\ \mathbb{P}^{\overline{\gamma}}\ a.s.,\ t \in \mathbb{T}_+ \tag{94}$$

which implies

$$\varphi^R\left(t, \omega, P_0\right) \stackrel{d}{=} \varphi^R\left(t, \omega, \widetilde{X}(\omega)\right),\ t \in \mathbb{T}_+ \tag{95}$$

From eqns. (93,95), we then have as $t \to \infty$

$$\varphi^R\left(t, \omega, P_0\right) \Longrightarrow \mu^{\overline{\gamma}} \tag{96}$$

which together with eqn. (84) implies

$$P_t \Longrightarrow \mu^{\overline{\gamma}} \tag{97}$$

as $t \to \infty$. This completes part i) of Theorem 9.

For part ii), we note that, if $\overline{\gamma} > \overline{\gamma}^{\text{bim}}$, by [10] (see Result 7), there exists $M^{\overline{\gamma}, P_0}$, such that

$$\sup_{t\in\mathbb{T}_+} \mathbb{E}^{\overline{\gamma}, P_0}[P_t] \preceq M^{\overline{\gamma}, P_0} \tag{98}$$

From eqn. (91), we have by Fatou's lemma

$$\liminf_{t\to\infty} \mathbb{E}^{\overline{\gamma}}\left[\left\|\varphi^R\left(t, \theta^R_{-t}\omega, \widetilde{X}\left(\theta^R_{-t}\omega\right)\right)\right\|\right] \geq \mathbb{E}^{\overline{\gamma}}\left[\|u(\omega)\|\right] \tag{99}$$

Using Lemma 16 and standard manipulations

$$\begin{aligned}
\int_{S^n_+} \|Y\|\, \mu^{\overline{\gamma}}(dY) &= \mathbb{E}^{\overline{\gamma}}\left[\|u(\omega)\|\right] \leq \liminf_{t\to\infty} \mathbb{E}^{\overline{\gamma}}\left[\left\|\varphi^R\left(t, \theta^R_{-t}\omega, \widetilde{X}\left(\theta^R_{-t}\omega\right)\right)\right\|\right] \\
&= \liminf_{t\to\infty} \mathbb{E}^{\overline{\gamma}}\left[\left\|\varphi^R\left(t, \omega, \widetilde{X}(\omega)\right)\right\|\right] = \liminf_{t\to\infty} \mathbb{E}^{\overline{\gamma}}\left[\left\|\varphi^R\left(t, \omega, P_0\right)\right\|\right] \\
&= \liminf_{t\to\infty} \mathbb{E}^{\overline{\gamma}}\left[\|P_t\|\right] \leq \liminf_{t\to\infty} \mathbb{E}^{\overline{\gamma}}\left[\text{Tr}\ (P_t)\right] \leq \text{Tr}\, M^{\overline{\gamma}, P_0} < \infty
\end{aligned}$$

which establishes part ii). ∎

B. Proof of Theorem 10

We state a result on the properties of invariant probabilities of Markov-Feller operators needed for the proof. Consider a locally compact separable metric space $(\mathcal{X}, d)$ and define the topological lower limit of a sequence $\{B_n\}_{n\in\mathbb{T}_+}$ of subsets of $\mathcal{X}$ by

$$\mathrm{Li}_{n\to\infty}B_n = \{x \in \mathcal{X} \mid \exists \text{ a sequence } \{x_n\}\text{, s.t. } x_n \in B_n,\ \forall n, \text{ and } \{x_n\} \text{ converges to } x\} \tag{100}$$

which, by definition, is closed. We then have the following result from [25].

Theorem 23 (Theorem 1.3.1 [25]) Let (L, T) be a Markov-Feller pair defined on $(\mathcal{X}, d)$. For all $x \in \mathcal{X}$, consider the sequence of measures $\left\{T^t\delta_x\right\}_{t\in\mathbb{T}_+}$ and define

$$\sigma(x) = \mathrm{Li}_{t\to\infty}\mathrm{supp}\left(T^t\delta_x\right) \tag{101}$$

$$\sigma = \cap_{x\in\mathcal{X}}\sigma(x) \tag{102}$$

Then, if (L, T) has an attractive probability μ, we have

$$\mathrm{supp}(\mu) = \sigma \tag{103}$$

We now complete the proof of Theorem 10.

Proof of Theorem 10: Fix $\overline{\gamma}^{\mathrm{sb}} < \overline{\gamma} < 1$ and recall the Markov-Feller pair $\left(L^{\overline{\gamma}}, T^{\overline{\gamma}}\right)$ (eqns. (22,23) and Proposition 3.) By Theorem 9, $\mu^{\overline{\gamma}}$ is an attractive probability for the pair $\left(L^{\overline{\gamma}}, T^{\overline{\gamma}}\right)$. We now use Theorem 23 to obtain the support of $\mu^{\overline{\gamma}}$.

For $X \in \mathbb{S}_+^N$, let

$$\sigma(X) = \mathrm{Li}_{t\to\infty}\mathrm{supp}\left(\left(T^{\overline{\gamma}}\right)^t\delta_X\right) \tag{104}$$

Then

$$\mathrm{supp}\left(\mu^{\overline{\gamma}}\right) = \cap_{X\in\mathbb{S}_+^N}\sigma(X) \tag{105}$$

We first show that

$$\mathrm{cl}(\mathcal{S}) \subset \mathrm{supp}\left(\mu^{\overline{\gamma}}\right) \tag{106}$$

where

$$\mathcal{S} = \{f_{i_1}\circ f_{i_2}\circ\cdots\circ f_{i_s}\left(P^*\right) \mid i_r \in \{0,1\},\ 1 \le r \le s,\ s \in \mathbb{T}_+\} \tag{107}$$

To this end, consider $X \in \mathbb{S}_+^N$. It follows from the properties of $T^{\overline{\gamma}}$ that

$$\mathrm{supp}\left(\left(T^{\overline{\gamma}}\right)^t\delta_X\right) = \{f_{i_1}\circ f_{i_2}\circ\cdots\circ f_{i_t}\left(X\right) \mid i_r \in \{0,1\},\ 1 \le r \le t\} \tag{108}$$

Indeed, starting from X, the only set of points reached with non-zero probability (and exhaustively) are the ones obtained by applying t arbitrary compositions of f_0 and f_1 on X. Since a set with finite cardinality is closed, we have the R.H.S. of eqn. (108) as the support of $\left(T^{\overline{\gamma}}\right)^t$.

Recall $P^* \in \mathbb{S}_{++}^N$ to be the deterministic fixed point of the algebraic Riccati equation. Consider the sequence $\left\{f_1^t(X)\right\}_{t\in\mathbb{T}_+}$ in $\mathbb{S}_+^N$. It follows from eqn. (108) that

$$f_1^t(X) \in \text{supp}\left(\left(T^{\overline{\gamma}}\right)^t \delta_X\right),\ \forall t \tag{109}$$

Also, from the properties of the algebraic Riccati equation, we have

$$\lim_{t\to\infty} f_1^t(X) = P^* \tag{110}$$

Hence, by the definition of topological lower limit,

$$P^* \in \sigma(X) \tag{111}$$

We now show that, for every $s \in \mathbb{T}_+$ and $i_r \in \{0,1\},\ 1 \leq r \leq s$,

$$f_{i_1} \circ f_{i_2} \circ \cdots \circ f_{i_s}\left(P^*\right) \in \sigma(X) \tag{112}$$

Indeed, define the sequence, $\{X_t\}_{t\in\mathbb{T}_+}$ as

$$X_t = \begin{cases} f_{i_1} \circ f_{i_2} \circ \cdots \circ f_{i_t}\left(X\right) & \text{if } t \leq s \\ f_{i_1} \circ f_{i_2} \circ \cdots \circ f_{i_s} \circ f_1^{t-s}(X) & \text{if } t > s \end{cases} \tag{113}$$

Clearly, $X_t \in \text{supp}\left(\left(T^{\overline{\gamma}}\right)^t \delta_X\right)$, $\forall t$. Since the sequence $\left\{f_1^{t-s}(X)\right\}_{t>s}$ converges to P^*, we have

$$\begin{aligned} \lim_{t\to\infty} X_t &= \lim_{t>s,t\to\infty} X_t = \lim_{t>s,t\to\infty} f_{i_1} \circ f_{i_2} \circ \cdots \circ f_{i_s}\left(f_1^{t-s}(X)\right) \\ &= f_{i_1} \circ f_{i_2} \circ \cdots \circ f_{i_s}\left(\lim_{t>s,t\to\infty} f_1^{t-s}(X)\right) = f_{i_1} \circ f_{i_2} \circ \cdots \circ f_{i_s}\left(P^*\right) \end{aligned}$$

where the continuity of $f_{i_1} \circ f_{i_2} \circ \cdots \circ f_{i_s}$ (being the composition of continuous functions, Proposition 2) permits bringing the limit inside.

Thus, the sequence $\{X_t\}_{t\in\mathbb{T}_+}$ converges to $f_{i_1} \circ f_{i_2} \circ \cdots \circ f_{i_s}\left(P^*\right)$ and $X_t \in \text{supp}\left(\left(T^{\overline{\gamma}}\right)^t \delta_X\right)$, $\forall t$. Hence, $f_{i_1} \circ f_{i_2} \circ \cdots \circ f_{i_s}\left(P^*\right) \in \sigma(X)$. It then follows

$$\mathcal{S} \subset \sigma(X),\ \forall X \in \mathbb{S}_+^N \tag{114}$$

Since the set $\sigma(X)$ is closed, we have

$$\text{cl}\left(\mathcal{S}\right) \subset \sigma(X),\ \forall X \in \mathbb{S}_+^N \tag{115}$$

which implies by eqn. (105)

$$\text{cl}(\mathcal{S}) \subset \text{supp}\left(\mu^{\overline{\gamma}}\right) \tag{116}$$

To obtain the reverse inclusion, we note that

$$\sigma\left(P^*\right) = \text{Li}_{t\to\infty}\text{supp}\left(\left(T^{\overline{\gamma}}\right)^t \delta_{P^*}\right) \subset \text{cl}\left(\cup_{t\in\mathbb{T}_+}\text{supp}\left(\left(T^{\overline{\gamma}}\right)^t \delta_{P^*}\right)\right) = \text{cl}(\mathcal{S})$$

Here the first step follows from the fact that, if $Y \in \mathrm{Li}_{t\to\infty}\mathrm{supp}\left(\left(T^{\overline{\gamma}}\right)^t \delta_{P^*}\right)$, then Y is a limit point of $\cup_{t\in\mathbb{T}_+}\mathrm{supp}\left(\left(T^{\overline{\gamma}}\right)^t \delta_{P^*}\right)$ and hence belongs to its closure. The last step is obvious from eqn. (108). We thus have

$$\mathrm{supp}\left(\mu^{\overline{\gamma}}\right) \subset \sigma\left(P^*\right) \subset \mathrm{cl}(\mathcal{S}) \tag{117}$$

which establishes the other inclusion and we have

$$\mathrm{supp}\left(\mu^{\overline{\gamma}}\right) = \mathrm{cl}(\mathcal{S}) \tag{118}$$

It remains to establish eqn. (37). To this end, we first show that

$$f_{i_1} \circ f_{i_2} \circ \cdots \circ f_{i_s}\left(P^*\right) \succeq P^* \;\; \forall\; s \in \mathbb{T}_+,\; i_r \in \{0,1\},\; 1 \leq r \leq s \tag{119}$$

We prove this by an inductive argument on s. Clearly, the above holds for $s = 1$, as

$$f_0\left(P^*\right) = AP^*A^T + Q \succeq P^* \tag{120}$$

(A is unstable and $Q >> 0$) and

$$f_1\left(P^*\right) = P^* \tag{121}$$

Assume the claim holds for $s \leq t$. We now show it holds for $s = t+1$. By the induction step,

$$f_{i_2} \circ \cdots \circ f_{i_{t+1}}\left(P^*\right) \succeq P^* \tag{122}$$

If $i_1 = 0$

$$f_{i_1} \circ f_{i_2} \circ \cdots \circ f_{i_{t+1}}\left(P^*\right) = f_0\left(f_{i_2} \circ \cdots \circ f_{i_{t+1}}\left(P^*\right)\right) \succeq f_0\left(P^*\right) \succeq P^* \tag{123}$$

which follows from the order preserving property of f_1 and eqn. (120). If $i_1 = 1$

$$f_{i_1} \circ f_{i_2} \circ \cdots \circ f_{i_{t+1}}\left(P^*\right) = f_1\left(f_{i_2} \circ \cdots \circ f_{i_{t+1}}\left(P^*\right)\right) \succeq f_1\left(P^*\right) = P^*$$

which follows from the order preserving property of f_0 and eqn. (121).

We thus have for $i_r \in \{0,1\},\; 1 \leq r \leq t+1$

$$f_{i_1} \circ f_{i_2} \circ \cdots \circ f_{i_{t+1}}\left(P^*\right) \succeq P^* \tag{124}$$

and the claim in eqn. (119) follows.

To complete the proof, we note from the above,

$$\mathcal{S} \subset [P^*, \infty) \tag{125}$$

Since the conic interval $[P^*, \infty)$ is closed, we have

$$\mathrm{cl}(\mathcal{S}) \subset [P^*, \infty) \tag{126}$$

and eqn. (37) follows as

$$\mu^{\overline{\gamma}}\left(\mathrm{cl}(\mathcal{S})\right)=1 \tag{127}$$

$\mathrm{cl}(\mathcal{S})$ being the support of $\mu^{\overline{\gamma}}$.

■

VII. A SCALAR EXAMPLE AND NUMERICAL STUDIES

A. Scalar Example

We investigate in detail a scalar system, for which we qualitatively characterize the structure of the support of the invariant distributions. In the general (non-scalar) case, Theorem 10 explicitly characterizes the support set of the invariant distributions and shows, in particular, that $\mathrm{supp}\left(\mu^{\overline{\gamma}}\right)$ is independent of $\overline{\gamma}$ as long as $\overline{\gamma}^{\mathrm{sb}}<\overline{\gamma}<1$. In this section, by studying a scalar example, we show that the set $\mathcal{S}$ is, in general, not a dense subset of the conic interval $[P^{*},\infty)$. In fact, for the example we consider, the support set $\mathcal{S}$ is a highly fractured subset of the interval $[P^{*},\infty)$ with a self-similar structure (to be explained below), thus exhibiting fractal-like properties.

Consider the scalar system, $A=\sqrt{2}$, $C=Q=R=1$. The functions f_0,f_1 then reduce to

$$f_0(X)=2X+1 \tag{128}$$

$$f_1(X)=3-\frac{2}{X+1} \tag{129}$$

and the fixed point of the algebraic Riccati equation is given by

$$P^{*}=1+\sqrt{2} \tag{130}$$

The next proposition shows that $\mathrm{supp}\left(\mu^{\overline{\gamma}}\right)$ is not dense in $[1+\sqrt{2},\infty)$ and exhibits self-similarity.

Proposition 24 Define $\mathcal{S}_0=\mathrm{cl}\left(\mathcal{S}\right)\cap[1+\sqrt{2},3]$ and, recursively, $\mathcal{S}_n=\{2Y+1,\ Y\in\mathcal{S}_{n-1}\}\,,\ n\geq 1$. We then have $\mathrm{cl}\,\mathcal{S}=\cup_{n\geq 0}\mathcal{S}_n$

Before proving Proposition 24, we interpret it. First, it reflects the self-similarity of $\mathrm{supp}\left(\mu^{\overline{\gamma}}\right)$, i.e., it suffices to know the structure of $\mathrm{supp}\left(\mu^{\overline{\gamma}}\right)$ in the interval $[1+\sqrt{2},3]$; this structure (with proper scaling) is repeated over space. In particular, if $\mathcal{S}_0$ is the restriction of $\mathrm{supp}\left(\mu^{\overline{\gamma}}\right)$ to $[1+\sqrt{2},3]$, the restriction to $[3+2\sqrt{2},7]$ is given by $\mathcal{S}_1$ (which can be written alternatively as $f_0\left(\mathcal{S}_0\right)$) and is a stretched version of $\mathcal{S}_0$, the stretching factor being 2. More generally, for $n\geq 1$, the restriction of $\mathrm{supp}\left(\mu^{\overline{\gamma}}\right)$ to $\left[2^{n}(1+\sqrt{2})+2^{n}-1,2^{n}.3+2^{n}-1\right]$ is given by $\mathcal{S}^{n}$, which is a stretched version of $\mathcal{S}_0$, the stretching factor being 2^{n}. Thus, $\mathrm{supp}\left(\mu^{\overline{\gamma}}\right)$ consists of stretching the set $\mathcal{S}_0$ by factors of 2^{n} and placing them over the real line.

Proposition 24 also shows that $\mathrm{supp}\left(\mu^{\overline{\gamma}}\right)$ is not dense in $\left[1+\sqrt{2},\infty\right)$ and contains holes. In fact, for every $n\geq 0$, there is a hole of length $2^{n+1}\sqrt{2}$ between the sets $\mathcal{S}_n$ and $\mathcal{S}_{n+1}$, corresponding to the interval $\left(2^{n}.3+2^{n}-1,2^{n+1}(1+\sqrt{2})+2^{n+1}-1\right)$. In other words, for every $n\geq 0$, the open interval $\left(2^{n}.3+2^{n}-1,2^{n+1}(1+\sqrt{2})+2^{n+1}-1\right)$ does not belong to $\mathrm{supp}\left(\mu^{\overline{\gamma}}\right)$.

Proof of Proposition 24.: Let $Y \in \text{supp}\left(\mu^{\overline{\gamma}}\right)$ and $Y > 3$.[9] Then, there exists a sequence $\{Y_n\}_{n\in\mathbb{T}_+}$ converging to Y, s.t. $Y_n \in \mathcal{S}$ for all n. For every $n \in \mathbb{T}_+$, Y_n can be represented as

$$Y_n = f_{i_1} \circ \cdots \circ f_{i_{s_n}}\left(P^*\right) \tag{131}$$

for some $s_n \in \mathbb{T}_+$ (depending on n) and $i_r \in \{0,1\}$, $1 \le r \le s_n$.

Since the sequence $\{Y_n\}_{n\in\mathbb{T}_+}$ is convergent, it is Cauchy, and there exists $n_0 \in \mathbb{T}_+$ such that

$$|\, Y_n - Y_{n_0} \,| < 2\sqrt{2},\ \forall\ \ n \ge n_0 \tag{132}$$

Without loss of generality, assume $n_0 = 0$ (otherwise, work with the sequence starting at n_0.) For every n, define

$$\widetilde{s}_n = \min\{r \mid i_r = 1,\ 1 \le r \le s_n\} \tag{133}$$

$$\widetilde{Y}_n = f_{i_{\widetilde{s}_n}} \circ \cdots \circ f_{i_{s_n}}\left(P^*\right) \tag{134}$$

Clearly, $1+\sqrt{2} \le \widetilde{Y}_n \le 3$ and, by definition, $\widetilde{Y}_n \in \mathcal{S}_0$, for all n. By basic manipulations and using eqn. (132), it can be shown

$$\widetilde{s}_n = \widetilde{s}_0,\ \forall n \tag{135}$$

We can then represent the sequence $\{Y_n\}_{n\in\mathbb{T}_+}$ as

$$Y_n = f_0^{\widetilde{s}_0 - 1}\left(\widetilde{Y}_n\right) \tag{136}$$

Since $Y_n \to Y$, and the function $f_0^{\widetilde{s}_0-1}(\cdot)$ is one-to-one and continuous, the sequence $\left\{\widetilde{Y}_n\right\}_{n\in\mathbb{T}_+}$ must converge to some $\widetilde{Y}$, i.e.,

$$\lim_{n\to\infty} \widetilde{Y}_n = \widetilde{Y} \tag{137}$$

It also follows that $1+\sqrt{2} \le \widetilde{Y} \le 3$ and $\widetilde{Y} \in \mathcal{S}_0$, the set $\mathcal{S}_0$ being closed. We then have

$$Y = f_0^{\widetilde{s}_0 - 1}\left(\widetilde{Y}\right) \tag{138}$$

which implies $Y \in \mathcal{S}_{\widetilde{s}_0 - 1}$. We thus showed the inclusion

$$\text{supp}\left(\mu^{\overline{\gamma}}\right) \subset \cup_{n\ge 0}\mathcal{S}_n \tag{139}$$

The reverse inclusion is obvious, and we have the claim. ■

Proposition 24 shows the self-similarity of supp $\left(\mu^{\overline{\gamma}}\right)$ at scales of 2^n, where $n \in \mathbb{N}$. A rigorous definition of fractal (see, for example, [29]) requires self-similarity at every scale, which we do not pursue here. This explains why we use 'fractal like' to describe the structure of supp $\left(\mu^{\overline{\gamma}}\right)$. The fractal nature of supp $\left(\mu^{\overline{\gamma}}\right)$, though not obvious, is not very surprising. In fact, it is known (see, for example, [30], [23]) that a large class of iterated function systems (systems, which generate a Markov process by random

[9] If $Y \le 3$, then $Y \in \mathcal{S}_0$ trivially by construction and hence $Y \in \cup_{n\ge 0}\mathcal{S}_n$.

switching between a set of at most countable functions)[10] leads to fractal invariant distributions.

Apart from the holes (fractures) in $\operatorname{supp}\left(\mu^{\overline{\gamma}}\right)$ explained by Proposition 24, the set $\operatorname{supp}\left(\mu^{\overline{\gamma}}\right)$ contains much more fractures, as observed in the numerical plots of $\operatorname{supp}\left(\mu^{\overline{\gamma}}\right)$ (see Fig. 1.) It follows from Proposition 24 that a thorough study of $\operatorname{supp}\left(\mu^{\overline{\gamma}}\right)$ requires studying only one of the sets, $\{\mathcal{S}_n\}_{n\in\mathbb{T}_+}$, as the pattern is repeated over the real line.

In Fig. 1 on the top left, the blue region corresponds to the set $\mathcal{S}_0$. The figure shows that the set contains many fractures (these fractures are internal to $\mathcal{S}_0$ and different from the holes between consecutive elements of $\{\mathcal{S}_n\}_{n\in\mathbb{T}_+}$ as explained by Proposition 24.) The blue blobs appearing in the figure are fractured more finely, but the visualization software limits the resolution by coalescing disconnected components separated by small distances into one large blob. A better visualization is obtained by looking at the set $\mathcal{S}_1$, see Fig. 1 on the top right, which is a stretched version (by a factor of 2) of $\mathcal{S}_0$, and more fractures are resolved. Finally, in Fig. 1 bottom, we plot the set $\mu^{\overline{\gamma}}$ restricted to $[1+\sqrt{2}, 15]$, i.e., the set $\mathcal{S}_0\cup\mathcal{S}_1\cup\mathcal{S}_2$. The figure demonstrates the self-similarity and the inter-set holes (holes between consecutive elements of $\{\mathcal{S}_n\}_{n\in\mathbb{T}_+}$) as explained by Proposition 24.

B. *Numerical Studies*

We study numerically the eigenvalue distribution from $\mu^{\overline{\gamma}}$ ($\overline{\gamma}^{\text{sb}} < \overline{\gamma} < 1$) for a 10-dimensional system. The matrices A and C (of dimensions 10×10 and 5×10, respectively) are generated randomly, so that the assumptions of Theorem 9 are satisfied.

In Fig. 2 on the left, we plot the cumulative distribution function (c.d.f.) of the largest eigenvalue $\lambda_{10}(\cdot)$ from $\mu^{\overline{\gamma}}$ for different values of $\overline{\gamma}^{\text{sb}} < \overline{\gamma} < 1$. From the figure, we see (as expected) that, as $\overline{\gamma}$ increases to 1, the distributions approach the Dirac distribution $\delta_{\lambda_{10}(P^*)}$ (the distribution with entire mass concentrated at $\lambda_{10}\left(P^*\right)$, the largest eigenvalue of the deterministic fixed point P^* of the algebraic Riccati equation.) The fact that, for $\overline{\gamma}^{\text{sb}} < \overline{\gamma} < 1$, the support of the eigenvalue distribution is a subset of $\left[\lambda_{10}\left(P^*\right),\infty\right)$ follows from Theorem 10 (eqn. (37).

Similarly, in Fig. (2) on the right, we plot the c.d.f. of the trace (which is the conditional mean-squared error) from $\mu^{\overline{\gamma}}$ for different values of $\overline{\gamma}^{\text{sb}} < \overline{\gamma} < 1$. From the figure, we see (as expected) that, as $\overline{\gamma}$ increases to 1, the distributions approach the Dirac distribution $\delta_{\text{Tr}(P^*)}$.

The eigenvalue distributions can be used in system design for control and estimation problems. Since the RRE converges in distribution to $\mu^{\overline{\gamma}}$, the system designer can tune the operating $\overline{\gamma}$ to ensure satisfactory system performance.

We end this subsection by remarking on the numerical computation of moments from the invariant distribution $\mu^{\overline{\gamma}}$ $\left(\overline{\gamma}^{\text{sb}} < \overline{\gamma} < 1.\right)$ We set notation first.

Define $L^1\left(\mu^{\overline{\gamma}}\right)$ to be the set of integrable functions on $\mathbb{S}_+^N$ w.r.t. the measure $\mu^{\overline{\gamma}}$. We then have the following assertion:

[10]The RRE can be viewed as an iterated function system, where the iterations are randomly switched between the Lyapunov (f_0) and Riccati (f_1) functions.

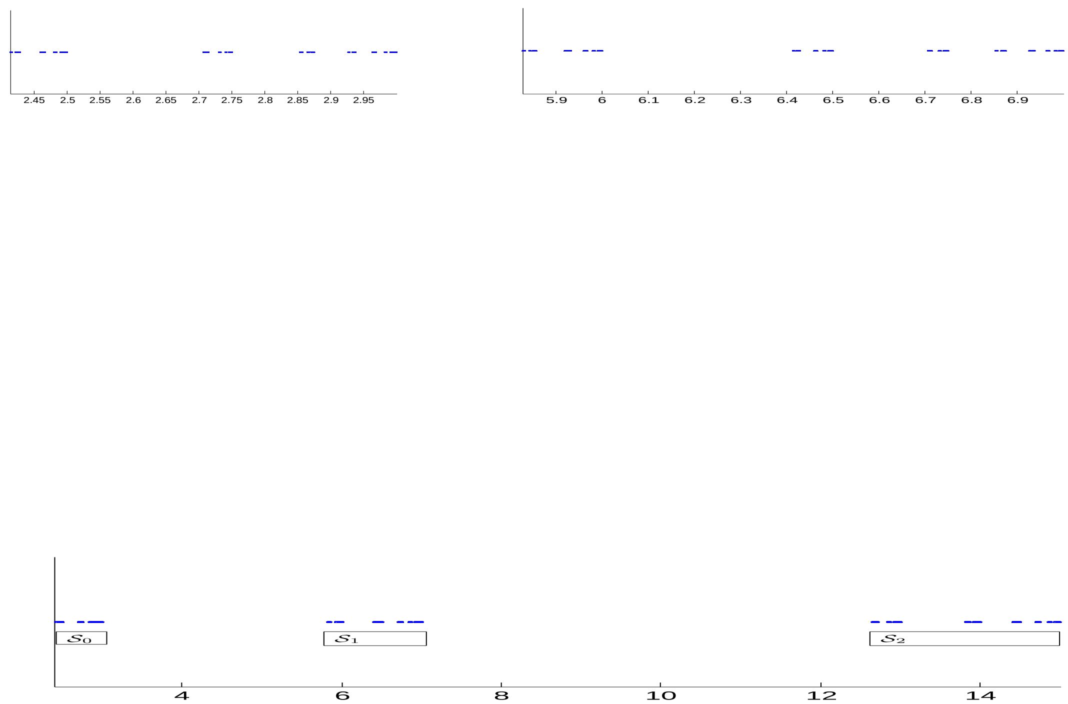

Fig. 1. Top left: Restriction of $\mu^{\overline{\gamma}}$ to $[1+\sqrt{2}, 3]$, i.e., the set $\mathcal{S}_0$. Top right: Restriction of $\mu^{\overline{\gamma}}$ to $[3+2\sqrt{2}, 7]$, i.e., the set $\mathcal{S}_1$. Bottom: Restriction of $\mu^{\overline{\gamma}}$ to $[1+\sqrt{2}, 15]$, i.e., the set $\mathcal{S}_0 \cup \mathcal{S}_1 \cup \mathcal{S}_2$.

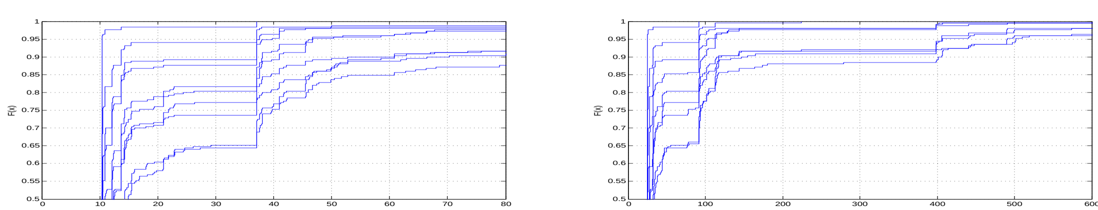

Fig. 2. Left: CDF of largest eigenvalue from $\mu^{\overline{\gamma}}$ for varying $\overline{\gamma}$, as $\overline{\gamma}$ approaches 1 ($\overline{\gamma}$ increases from right to left.) Right: CDF of trace from $\mu^{\overline{\gamma}}$ for varying $\overline{\gamma}$, as $\overline{\gamma}$ approaches 1 ($\overline{\gamma}$ increases from right to left.)

Proposition 25 Fix $\overline{\gamma}^{\text{sb}} < \overline{\gamma} < 1$ and let $h \in L^1\left(\mu^{\overline{\gamma}}\right)$. Then there exists a set $\mathcal{S}_h^{\overline{\gamma}} \subset \mathbb{S}_+^N$ with $\mu^{\overline{\gamma}}\left(\mathcal{S}_h^{\overline{\gamma}}\right) = 1$, such that, for every $P_0 \in \mathcal{S}_h^{\overline{\gamma}}$, the following holds:

$$\lim_{t\to\infty} \frac{1}{t} \sum_{s=0}^{t-1} h(P_t) = \int_{\mathbb{S}_+^N} h(Y)\mu^{\overline{\gamma}} d(Y),\ \mathbb{P}^{\overline{\gamma},P_0} \text{ a.s.} \tag{140}$$

where $\{P_t\}_{t\in\mathbb{T}_+}$ is the sequence generated by the RRE with initial condition P_0.

Proof: It follows from the fact that $\mu^{\overline{\gamma}}$ is an ergodic measure (being attractive) for the transition probability operator $\mathbb{Q}^{\overline{\gamma}}$ of the Markov process $\{P_t\}_{t\in\mathbb{T}_+}$ (see, for example, [31].) ■

Proposition 25 has important consequences in computing moments and probabilities (for example, the probability of escape from a set can be obtained by using h to be the complementary indicator function) from the invariant distribution $\mu^{\overline{\gamma}}$. It says that, for a function $h : \mathbb{S}_+^N \longmapsto \mathbb{R}$ with finite $\mu^{\overline{\gamma}}$-moment, there exists a set $\mathcal{S}_h^{\overline{\gamma}}$ with $\mu^{\overline{\gamma}}$-probability one, such that, if the initial condition belongs to $\mathcal{S}_h^{\overline{\gamma}}$, the empirical moment converges to the $\mu^{\overline{\gamma}}$-moment for every sample path a.s. Thus, in order to compute a $\mu^{\overline{\gamma}}$-moment, generating a single instance of the Markov process suffices, as long as the initial condition belongs to the set $\mathcal{S}_h^{\overline{\gamma}}$. This has important consequences in moment computation from the invariant distribution as, under the assumptions of Proposition 25, one does not need to run costly simulations to generate the distribution $\mu^{\overline{\gamma}}$ empirically; rather, the generation of a single path would suffice.

The difficulty in using Proposition 25 is that the set $\mathcal{S}_h^{\overline{\gamma}}$ is known only up to a.s. equivalence and further depends on h. In general, $\mathcal{S}_h^{\overline{\gamma}}$ is not the entire $\mathbb{S}_+^N$.[11] However, Theorem 10 provides us with the support of $\mu^{\overline{\gamma}}$ and implies, in particular, that for every $\varepsilon > 0$ and $Y \in \mathcal{S}$ the open ball $B_\varepsilon(Y)$ has positive measure. Thus, by choosing initial conditions P_0 arbitrarily close to (and including) a $Y \in \mathcal{S}$, one is likely to get the convergence in Proposition 25.

Also, we note that the set of functions $h \in L^1(\mu^{\overline{\gamma}})$ is not empty. As a matter of fact, all bounded measurable functions $h : \mathbb{S}_+^N \longmapsto \mathbb{R}$ are contained in $L^1(\mu^{\overline{\gamma}})$, for every $\overline{\gamma}$ which guarantees the existence and uniqueness of $\mu^{\overline{\gamma}}$ (for example, $\overline{\gamma} > 0$ for stabilizable and detectable systems.) In some situations, it may be possible to determine the moments of unbounded functionals by approximating them by a sequence of suitable truncations and then invoking some form of dominated convergence. An important case is the mean evaluation corresponding to $h(Y) = Y$. In that case, by Theorem III we note that if we operate at $\overline{\gamma} > \overline{\gamma}^{\text{bim}}$, the integral w.r.t. the corresponding $\mu^{\overline{\gamma}}$ exists and hence, one may invoke Proposition 25 to compute the mean under the invariant measure.

VIII. Conclusions and Future Work

The paper presents a new analysis of the Random Riccati Equation. It studies the evolution of the state error covariance arising from a Kalman filter where observations can be lost according to an i.i.d. Bernoulli process. This process can model an analog erasure channel between the sensor and the

[11] In fact, $\mathcal{S}_h^{\overline{\gamma}} = \mathbb{S}_+^N$ *iff* the Markov process is positive Harris recurrent, a property that iterated function systems do not possess generally (see [31].)

estimator. Following a novel approach based on random dynamical systems, we provide an exhaustive analysis of the steady state behavior of the filter.

We showed the existence of critical arrival probabilities $\overline{\gamma}^{\mathrm{sb}}, \overline{\gamma}^{\mathrm{bim}}$, such that the error covariance converges in distribution to a unique steady state distribution if the arrival probability $\overline{\gamma} > \overline{\gamma}^{\mathrm{sb}}$; this distribution has finite mean for $\overline{\gamma} > \overline{\gamma}^{\mathrm{bim}}$. Additionally, we provided a characterization of the support of the steady state distribution, showing its fractal characteristics. The latter result, combined with ergodicity arguments, provides a method to numerically evaluate the error covariance steady state distribution. We feel that our approach is particularly amenable to addressing general problems of networked control problems, as they provide a theoretical framework to combine stochastic processes used in the modeling of communication networks with differential equations describing the evolution of dynamical systems.

The approach and results in the paper will easily extend to problems of control over erasure channels. Further, we plan to describe more complex interactions and tradeoffs between communication and control via random dynamical systems.

Appendix A
Proofs of Propositions 3,6,8

Proposition 3: The fact that $T^{\overline{\gamma}}$ is a Markov operator is a standard consequence of $\mathcal{Q}^{\overline{\gamma}}$ being a transition probability (see, for example, [25].) Also, $L^{\overline{\gamma}}$ is linear. Thus, we only need to verify that $L^{\overline{\gamma}}$ maps $C_b\left(\mathbb{S}_+^N\right)$ to $C_b\left(S_n^+\right)$. For linear operators generated by eqn. (22), such a property is called the weak-Feller property of the transition probability $\mathbb{Q}^{\overline{\gamma}}$ (see [31].) It can be shown (see Proposition 7.2.1. in [31]) that $\mathbb{Q}^{\overline{\gamma}}$ is weak-Feller *iff* for every sequence $\{Y_n\}_{n\in\mathbb{T}_+}$ in $\mathbb{S}_+^N$ that converges to some $Y\in\mathbb{S}_+^N$, and every open set $\mathcal{O}\in\mathcal{B}\left(\mathbb{S}_+^N\right)$

$$\liminf_{n\to\infty}\mathbb{Q}^{\overline{\gamma}}\left(Y_n,\mathcal{O}\right)\geq\mathbb{Q}^{\overline{\gamma}}\left(Y,\mathcal{O}\right) \tag{141}$$

To verify eqn. (141), we note that by eqn. (19)

$$\mathbb{Q}^{\overline{\gamma}}\left(Y_n,\mathcal{O}\right)=\left(1-\overline{\gamma}\right)\mathbb{I}_{\mathcal{O}}\left(f_0\left(Y_n\right)\right)+\overline{\gamma}\mathbb{I}_{\mathcal{O}}\left(f_1\left(Y_n\right)\right),\ \forall n\in\mathbb{T}_+ \tag{142}$$

If $f_0(Y)\in\mathcal{O}$, from the continuity of f_0 (Lemma 2) and $\mathcal{O}$ open, $\exists\, n_0\in\mathbb{T}_+$, s.t.

$$f_0\left(Y_n\right)\in\mathcal{O},\ n\geq n_0 \tag{143}$$

which implies

$$\lim_{n\to\infty}\mathbb{I}_{\mathcal{O}}\left(f_0\left(Y_n\right)\right)=1=\mathbb{I}_{\mathcal{O}}\left(f_0(Y)\right) \tag{144}$$

On the other hand, if $f_0(Y)\notin\mathcal{O}$, we have

$$\liminf_{n\to\infty}\mathbb{I}_{\mathcal{O}}\left(f_0\left(Y_n\right)\right)\geq 0=\mathbb{I}_{\mathcal{O}}\left(f_0(Y)\right) \tag{145}$$

Similarly, we have

$$\liminf_{n\to\infty}\mathbb{I}_{\mathcal{O}}\left(f_1\left(Y_n\right)\right)\geq\mathbb{I}_{\mathcal{O}}\left(f_1(Y)\right) \tag{146}$$

and eqn. (141) follows. ∎

Proposition 6: Assume $\{P_t\}_{t\in\mathbb{T}_+}$ is b.i.m. for some $\overline{\gamma}$ and P_0, i.e., $\exists\, M^{\overline{\gamma},P_0}$ such that

$$\sup_{t\in\mathbb{T}_+} \mathbb{E}^{\overline{\gamma},P_0}\left[P_t\right] \preceq M^{\overline{\gamma},P_0} \tag{147}$$

By the positive semidefinitess of the matrices and linearity of the trace,

$$\forall t\in\mathbb{T}_+ :\ \mathbb{E}^{\overline{\gamma},P_0}\left[\|P_t\|\right] \leq \mathbb{E}^{\overline{\gamma},P_0}\left[\operatorname{Tr} P_t\right] = \operatorname{Tr}\left(\mathbb{E}^{\overline{\gamma},P_0}\left[P_t\right]\right) \leq \operatorname{Tr} M^{\overline{\gamma},P_0} \tag{148}$$

Chebyshev's inequality then implies

$$\begin{aligned} \mathbb{P}^{\overline{\gamma},P_0}\left(\|P_t\| > N\right) &\leq \frac{\operatorname{Tr} M^{\overline{\gamma},P_0}}{N} \\ \limsup_{N\to\infty}\ \sup_{t\in\mathbb{T}_+} \mathbb{P}^{\overline{\gamma},P_0}\left(\|P_t\| > N\right) &\leq \lim_{N\to\infty} \frac{\operatorname{Tr} M^{\overline{\gamma},P_0}}{N} = 0 \end{aligned} \tag{149}$$

■

Proposition 8: Part i) is obvious from Proposition 6.

For part ii) we consider the case of unstable A. For stable A, the proposition is trivial and follows from the fact, that, the unconditional variance of the state sequence reaches a steady state (hence bounded), and a suboptimal estimate $\widehat{x}_t \equiv 0$ for all t leads to pathwise boundedness of the corresponding error covariance. In fact, as pointed earlier, in this case, even $\overline{\gamma} = 0$ leads to stochastic boundedness of the sequence $\{P_t\}$ from every initial condition.

We now prove part [ii] for unstable A, under the additional assumption of invertibility of C (the general case being considered in Lemma 13 of [24].) Fix $\overline{\gamma} > 0$, $P_0 \in \mathbb{S}_+^N$ and recall the function $f_1(\cdot)$ (eqn. (17))

$$f_1(X) = AXA^T + Q - AXC^T\left(CXC^T + R\right)^{-1} CXA^T \tag{150}$$

Since C is invertible, we use the matrix inversion lemma to obtain

$$\left(CXC^T + R\right)^{-1} = C^{-T}X^{-1}C^{-1} - C^{-T}X^{-1}C^{-1}\left(R^{-1} + C^{-T}X^{-1}C^{-1}\right)^{-1} C^{-T}X^{-1}C^{-1} \tag{151}$$

By substituting into eqn. (150), we have

$$f_1(X) = AC^{-1}\left(R^{-1} + C^{-T}X^{-1}C^{-1}\right)^{-1} C^{-T}A^T + Q \tag{152}$$

Using $\left\|\left(R^{-1} + C^{-T}X^{-1}C^{-1}\right)^{-1}\right\| \leq \|R\|$, we have

$$\|f_1(X)\| \leq \|A\|\|C\|\|C^{-T}\|\|A^T\|\|R\| + \|Q\| \tag{153}$$

For $N \in \mathbb{T}_+$ and sufficiently large, define

$$k(N) = \min\left\{k\in\mathbb{T}_+ \mid M\left(\alpha^{2(k-1)} - 1\right) + \frac{\|Q\|\,\alpha^{2(k-1)}}{\alpha^2 - 1} \geq N\right\} \tag{154}$$

where α is the absolute value of the largest eigenvalue of A and M is

$$M = \max\left\{\|A\|\|C\|\|C^{-T}\|\|A^T\|\|R\| + \|Q\|\,,\|P_0\|\right\} \tag{155}$$

As $N\uparrow\infty$, $k(N)\uparrow\infty$. To estimate $\mathbb{P}^{\overline{\gamma},P_0}\left(\|P_t\| > N\right)$, define the random time $\widetilde{t}$ by

$$\widetilde{t} = \max\left\{0 < s \leq t \mid \gamma_{s-1} = 1\right\} \tag{156}$$

where the maximum of an empty set is taken to be zero. Using the above arguments, we have

$$\|P_{\widetilde{t}}\| \leq M \tag{157}$$

Indeed, if $\widetilde{t}=0$, eqn. (157) clearly holds by definition of M. On the contrary, if $\widetilde{t}>0$, we have

$$\|P_{\widetilde{t}}\| = \left\|f_1\left(P_{\widetilde{t}-1}\right)\right\| \leq \|A\|\|C\|\|C^{-T}\|\|A^T\|\|R\| + \|Q\| \leq M \tag{158}$$

We also have $P_t = f_0^{t-\widetilde{t}-1}\left(P_{\widetilde{t}}\right)$, which implies

$$\|P_t\| = \left\|f_0^{t-\widetilde{t}-1}\left(P_{\widetilde{t}}\right)\right\| \leq M\alpha^{2(t-\widetilde{t}-1)} + \|Q\| \sum_{k=1}^{t-\widetilde{t}-1} \alpha^{2(k-1)} = M\alpha^{2(t-\widetilde{t}-1)} + \|Q\| \frac{\alpha^{2(t-\widetilde{t}-1)}-1}{\alpha^2-1}$$

From this inequality and eqn. (154), it follows

$$\mathbb{P}^{\overline{\gamma},P_0}\left(\|P_t\|>N\right) \leq \mathbb{P}^{\overline{\gamma},P_0}\left(t-\widetilde{t} \geq k(N)\right) \tag{159}$$

We observe that the random time $0 \leq \widetilde{t} \leq t$. In case, $0 \leq k \leq t-1$,

$$\mathbb{P}^{\overline{\gamma},P_0}\left(t-\widetilde{t}=k\right) = \mathbb{P}^{\overline{\gamma},P_0}\left(\gamma_{t-k-1}=1, \{\gamma_s=0\}_{t-k\leq s<t}\right) = \overline{\gamma}\left(1-\overline{\gamma}\right)^k \leq \left(1-\overline{\gamma}\right)^k \tag{160}$$

If $k=t$, we have

$$\mathbb{P}^{\overline{\gamma},P_0}\left(t-\widetilde{t}=k\right) = \mathbb{P}^{\overline{\gamma},P_0}\left(\{\gamma_s=0\}_{0\leq s<t}\right) = \left(1-\overline{\gamma}\right)^k \tag{161}$$

We thus have the upper bound (possibly loose, but sufficient for our purposes)

$$\mathbb{P}^{\overline{\gamma},P_0}\left(t-\widetilde{t} \geq k(N)\right) \leq \sum_{k=k(N)}^{\infty} \left(1-\overline{\gamma}\right)^k = \frac{\left(1-\overline{\gamma}\right)^{k(N)}}{\overline{\gamma}} \tag{162}$$

From eqns. (159,162), we have for all t and sufficiently large N

$$\mathbb{P}^{\overline{\gamma},P_0}\left(\|P_t\|>N\right) \leq \frac{\left(1-\overline{\gamma}\right)^{k(N)}}{\overline{\gamma}} \tag{163}$$

Since $\overline{\gamma}>0$ and $k(N)\uparrow\infty$ as $N\uparrow\infty$, it then follows

$$\lim_{N\to\infty}\sup_{t\in\mathbb{T}_+}\mathbb{P}^{\overline{\gamma},P_0}\left(\|P_t\|>N\right) = 0 \tag{164}$$

Thus, $\{P_t\}_{t\in\mathbb{T}_+}$ is s.b. (for all initial conditions P_0) for every $\overline{\gamma}>0$. The Proposition follows. ■

References

[1] R. E. Kalman, "A new approach to linear filtering and prediction problems," *Transactions of the ASME - Journal of Basic Engineering on Automatic Control*, vol. 82(D), pp. 35–45, March 1960.

[2] H. Chen, P. Kumar, and J. van Schuppen, "On Kalman filtering for conditionally Gaussian systems with random matrices," *Systems & Control Letters*, vol. 13, pp. 397–404, 1989.

[3] Y. Wang and L. Guo, "On stability of random Riccati equations," *Science in China Series E: Technological Sciences*, vol. 42, no. 2, pp. 136–148, April 1999.

[4] S. Tatikonda and S. Mitter, "Control under communication constraints," *IEEE Transactions on Automatic Control*, vol. 49, no. 7, pp. 1056 – 1068, July 2004.

[5] ——, "Control over noisy channels," *IEEE Transactions on Automatic Control*, vol. 49, no. 7, pp. 1196 – 1201, July 2004.

[6] P. Minero, M. Franceschetti, S. Dey, and G. Nair, "Data rate theorem for stabilization over time-varying feedback channels," *IEEE Transactions on Automatic Control*, vol. 54, no. 2, pp. 243–255, February 2009.

[7] A. S. Matveev and A. V. Savkin, "An analogue of Shannon information theory for networked control systems: State estimation via a noisy discrete channel," in *IEEE Conference on Decision and Control*, Dec 2004, pp. 4485–4490.

[8] N. Martins, M. Daleh, and N. Elia, "Feedback stabilization of uncertain systems in the presence of a direct link," *IEEE Transactions on Automatic Control*, vol. 51, no. 3, pp. 438–447, March 2006.

[9] A. Matveev and A. Savkin, "The problem of state estimation via asynchronous communication channels with irregular transmission times," *IEEE Transactions on Automatic Control*, vol. 48, no. 4, pp. 670–676, April 2006.

[10] B. Sinopoli, L. Schenato, M. Franceschetti, K. Poolla, M. Jordan, and S. Sastry, "Kalman filtering with intermittent observations," *IEEE Transactions of Automatic Control*, vol. 49, no. 9, pp. 1453–1464, 2004.

[11] Y. Mo and B. Sinopoli, "A characterization of the critical value for Kalman filtering with intermittent observations," in *IEEE Conference on Decision and Control*, Cancun, Mexico, Dec 2008.

[12] X. Liu and A. Goldsmith, "Kalman filtering with partial observation losses," in *Proceedings of IEEE Conference on Decision and Control*, vol. 4, Bahamas, December 2004, pp. 4180–4186.

[13] V. Gupta, T. Chung, B. Hassibi, and R. M. Murray, "On a stochastic sensor selection algorithm with applications in sensor scheduling and sensor coverage," *Submitted to Automatica*, 2005.

[14] Y. Xu and J. Hespanha, "Estimation under controlled and uncontrolled communications in networked control systems," in *IEEE Conference on Decision and Control*, Sevilla, Spain, December 2005, pp. 842–847.

[15] M. Huang and S. Dey, "Stability of Kalman filtering with Markovian packet losses," *Automatica*, vol. 43, no. 4, pp. 598–607, 2007.

[16] K. Plarre and F. Bullo, "On Kalman filtering for detectable systems with intermittent observations," *IEEE Transactions on Automatic Control*, May 2007, submitted for publication.

[17] C. Robinson and P. R. Kumar, "Sending the most recent observation is not optimal in networked control: Linear temporal coding and towards the design of a control specific transport protocol," in *IEEE Conference on Decision and Control*, Dec 2007, pp. 334–339.

[18] L. Xie and L. Xie, "Stability of a random Riccati equation with Markovian binary switching," in *IEEE Conference on Decision and Control*, Dec 2007.

[19] L. Arnold and I. Chueshov, "Order-preserving random dynamical systems: Equilibria, attractors, applications," *Dynamics and Stability of Systems*, vol. 13, pp. 265–280, 1998.

[20] A. Vakili and B. Hassibi, "A Stieltjes transform approach for studying the steady-state behavior of random Lyapunov and Riccati recursions," in *IEEE Conference on Decision and Control*, Dec 2008.

[21] A. Censi, "On the performance of Kalman filtering with intermittent observations: a geometric approach with fractals," in *American Control Conference*, 2009, submitted.

[22] D.-V. Voiculescu, *Free Probability Theory*. American Mathematical Society, 1997.

[23] P. Diaconis and D. Freedman, "Iterated random functions," *SIAM Review*, vol. 41, no. 1, pp. 45–76, March 1999.

[24] S. Kar and J. M. F. Moura, "Kalman filtering with intermittent observations: Weak convergence and moderate deviations," October 2009, submitted for journal publication. [Online]. Available: http://arxiv.org/abs/0910.4686

[25] R. Zaharopol, *Invariant Probabilities of Markov-Feller Operators and their Supports*. Birkhauser, 2005.

[26] J. Jacod and A. Shiryaev, *Limit Theorems for Stochastic Processes*. Berlin Heidelberg: Springer-Verlag, 1987.

[27] I. Chueshov, *Monotone Random Systems - Theory and Applications*. Springer, 2002.

[28] O. Kallenberg, *Foundations of Modern Probability*, 2nd ed. Springer Series in Statistics., 2002.

[29] B. Mandelbrot, *Fractals: Form, Chance, and Dimension*. San Francisco, CA: Freeman, 1977.

[30] A. Lsota and M. Mackey, *Chaos, Fractals, and Noise*. New York: Springer Verlag, 1994.

[31] O. Hernadez-Lerma and J. Lasserre, *Markov Chains and Invariant Probabilities*. Birkhauser, 2003.