SIAM J. CONTROL OPTIM.
Vol. 55, No. 6, pp. 4015–4047
© 2017 Society for Industrial and Applied Mathematics

Downloaded 08/19/18 to 138.25.168.33. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

ON THE STABILITY OF KALMAN–BUCY DIFFUSION PROCESSES*

ADRIAN N. BISHOP[†] AND PIERRE DEL MORAL[‡]

Abstract. The Kalman–Bucy filter is the optimal state estimator for an Ornstein–Uhlenbeck diffusion given that the system is partially observed via a linear diffusion-type (noisy) sensor. Under Gaussian assumptions, it provides a finite-dimensional exact implementation of the optimal Bayes filter. It is generally the only such finite-dimensional exact instance of the Bayes filter for continuous state-space models. Consequently, this filter has been studied extensively in the literature since the seminal 1961 paper of Kalman and Bucy. The purpose of this work is to review, re-prove and refine existing results concerning the dynamical properties of the Kalman–Bucy filter so far as they pertain to filter stability and convergence. The associated differential matrix Riccati equation is a focal point of this study with a number of bounds, convergence, and eigenvalue inequalities rigorously proven. New results are also given in the form of exponential and comparison inequalities for both the filter and the Riccati flow.

Key words. differential Riccati equations, diffusion flows, Kalman–Bucy diffusion, Kalman–Bucy filter, transition semigroups

AMS subject classifications. 62M20, 93E11, 60G99

DOI. 10.1137/16M1102707

1. Introduction.

The aim of this study is to review, re-prove, and also refine a number of existing stability results on the Kalman–Bucy filter and the associated Riccati equation. We correct prior work where necessary. New results are also given in the form of exponential and comparison inequalities for both the stochastic flow of the filter and the Riccati flow. This work is intended to be a complete and self-contained analysis on the stability and convergence of Kalman–Bucy filtering, with detailed proofs of each necessary result.

Consider a linear-Gaussian filtering model of the following form:

$$(1)\qquad \begin{cases} dX_t = A_t X_t dt + R_1^{1/2} dW_t, \\ dY_t = C_t X_t dt + R_2^{1/2} dV_t. \end{cases}$$

Here, (W_t, V_t) is an $(r_1 + r_2)$-dimensional standard Brownian motion. We let $\mathcal{F}_t = \sigma(Y_s, s \leqslant t)$ be the filtration generated by the observation and $Y_0 = 0$. We assume that X_0 is a r_1-valued independent random vector with mean $\mathbb{E}(X_0)$ and finite covariance P_0. Note that X_0 is not necessarily Gaussian.

Further, A_t is a square $(r_1 \times r_1)$-matrix, C_t is an $(r_2 \times r_1)$-matrix, and $R_1^{1/2}$ and $R_2^{1/2}$ are symmetric $(r_1 \times r_1)$-matrices. The eigenvalues of A, C, R_1, R_2 are bounded above and below (uniformly in time) and those of R_1, R_2 are uniformly bounded positive.

We consider both time-varying (e.g. A_t) and time-invariant signal models (e.g., $A_t = A$) and the convergence properties of the respective filters and associated Riccati

*Received by the editors November 8, 2016; accepted for publication (in revised from) July 6, 2017; published electronically December 19, 2017.

http://www.siam.org/journals/sicon/55-6/M110270.html

Funding: The first author is supported by the Australian Research Council (ARC) via a Discovery Early Career Researcher Award (DE-120102873).

[†]University of Technology Sydney, Ultimo NSW 2007, Australia (adrian.bishop@anu.edu.au).

[‡]INRIA, 200 Avenue de la Vieille Tour, 33405 Talence, France (pierre.del_moral@inria.fr).

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 08/19/18 to 138.25.168.33. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

equations. Typically, we state general results for the time-varying signal first, and then follow this with more quantitative results in the time-invariant case.

When X_0 is Gaussian, it is well known that the conditional distribution of the signal state X_t given $\mathcal{F}_t$ is an r_1-dimensional Gaussian distribution with a mean and covariance matrix

$$\widehat{X}_t := \mathbb{E}(X_t|\mathcal{F}_t) \quad \text{and} \quad P_t := \mathbb{E}\left((X_t - \mathbb{E}(X_t|\mathcal{F}_t))\,(X_t - \mathbb{E}(X_t|\mathcal{F}_t))'\right)$$

given by the Kalman–Bucy and the Riccati equations

$$d\widehat{X}_t = A_t\widehat{X}_t dt + P_t C_t' R_2^{-1}\left(dY_t - C_t\widehat{X}_t dt\right), \tag{2}$$

$$\partial_t P_t = \mathrm{Ricc}(P_t) \tag{3}$$

with the Riccati drift function from $\mathbb{S}_{r_1}^+$ into $\mathbb{S}_{r_1}$ defined for any $Q \in \mathbb{S}_{r_1}^+$ by

$$\mathrm{Ricc}(Q) = A_t Q + Q A_t' - Q S_t Q + R_1 \quad \text{with} \quad S_t := C_t' R_2^{-1} C_t.$$

Note that S_t is time-varying whenever, e.g., C_t is time-varying.

The Kalman–Bucy filter is the $\mathbb{L}_2$-optimal state estimator for an Ornstein–Uhlenbeck diffusion given that the system is partially observed via a linear diffusion-type (noisy) sensor; see [27, 8].

The stability of this filter was initially studied by Kalman and Bucy in their seminal paper [27], with related prior work by Kalman [23, 24, 26] and later work by Bucy [4]. In [2] the stability of the filter was analyzed under a relaxed controllability condition. It was analyzed again in [35] for systems with non-Gaussian initial state via a Kallianpur–Striebel-type change of probability measures. An alternative approach is to consider the following conditional nonlinear McKean–Vlasov-type diffusion process:

$$d\overline{X}_t = A_t\overline{X}_t dt + R_1^{1/2} d\overline{W}_t + \mathcal{P}_{\eta_t} C_t' R_2^{-1}\left[dY_t - \left(C_t\overline{X}_t dt + R_2^{1/2} d\overline{V}_t\right)\right], \tag{4}$$

where $(\overline{W}_t, \overline{V}_t, \overline{X}_0)$ are independent copies of (W_t, V_t, X_0) (thus independent of the signal and the observation path). In the above displayed formula, $\mathcal{P}_{\eta_t}$ stands for the covariance matrix

$$\mathcal{P}_{\eta_t} = \eta_t\left[(e - \eta_t(e))(e - \eta_t(e))'\right] \quad \text{with} \quad \eta_t := \mathrm{Law}(\overline{X}_t|\mathcal{F}_t) \quad \text{and} \quad e(x) := x.$$

We shall call this probabilistic model the Kalman–Bucy (nonlinear) diffusion process.

In contrast to conventional nonlinear diffusions, the interaction does not take place only on the drift part but also on the diffusion matrix functional. In addition, the nonlinearity does not depend on the distribution of the random states $\pi_t = \mathrm{Law}(\overline{X}_t)$ but on their conditional distributions $\eta_t := \mathrm{Law}(\overline{X}_t|\mathcal{F}_t)$. The well-posedness of this nonlinear diffusion is discussed in [17].

The nonlinear Kalman–Bucy diffusion is, in some sense, a generalized description of the Kalman–Bucy filter. The Riccati equation (3) is encapsulated in the nonlinear term of the diffusion. More precisely, the conditional expectations of the random states $\overline{X}_t$ and their conditional covariance matrices $\mathcal{P}_{\eta_t}$ w.r.t. $\mathcal{F}_t$ satisfy the Kalman–Bucy and the Riccati equations (2) and (3), *even when the initial state is not Gaussian.* That is, if we *redefine*

$$\widehat{X}_t := \mathbb{E}\left(\overline{X}_t|\mathcal{F}_t\right) \qquad \text{and} \qquad P_t := \mathcal{P}_{\eta_t}, \tag{5}$$

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 08/19/18 to 138.25.168.33. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

then the flow of this (conditional) mean and covariance satisfy (2) and (3) regardless of the distribution of X_0. We assume this more general definition of $\widehat{X}_t$ and P_t when referring to (2) and (3) going forward.

Note that the flow of matrices $\mathcal{P}_{\eta_t}$ depends only on the covariance matrix of the initial state $\overline{X}_0$. This property follows from the specially designed structure of the nonlinear diffusion, which ensures that the mean and covariance matrices satisfy the Kalman–Bucy filter and Riccati equations. This structure simplifies the stability analysis of this diffusion. Given $\mathcal{P}_{\eta_0}$, the Kalman–Bucy diffusion (4) can be interpreted as a nonhomogeneous Ornstein–Uhlenbeck-type diffusion with a conditional covariance matrix $P_t = \mathcal{P}_{\eta_t}$ that satisfies the Riccati equation (3) starting from $P_0 = \mathcal{P}_{\eta_0}$. In this sense, the nonlinearity of the process is encapsulated by the Riccati equation.

Analysis of this diffusion allows one to capture non-Gaussian initial states even for time-varying signal models. This class of nonlinear diffusion also arises in the mathematical and numerical foundations of ensemble Kalman–Bucy filters and data assimilation [19]. In this context, the stability properties of the Kalman–Bucy diffusion are essential for analyzing the long-time behavior of this class of algorithm.

Reiterating, in this work we revisit the stability of the Kalman–Bucy filter and we study for the first time the stability properties of the Kalman–Bucy diffusion. We derive new exponential inequalities detailing the convergence of the filter and the diffusion with arbitrary initial conditions, as well as the convergence properties of the associated differential Riccati equation. The classical study of Riccati equations in control and estimation theory is motivated by their relationship with Kalman–Bucy filtering and linear-quadratic optimal control theory [27, 23, 24]. Indeed, the two topics are dual, and the two relevant differential Riccati equations are (mostly) equivalent up to a time reversal. We deal here primarily with the forward-type equation associated with the evolution of the Kalman–Bucy filtering error.

We review now some of the key literature on the (deterministic) matrix Riccati-type differential equation, i.e., quadratic matrix differential equations [27]. Our interest in this equation follows because it describes the covariance flow of the Kalman–Bucy state estimation error. However, the properties and behavior of this equation are of interest in their own right. Bucy [4] originally studied a number of global properties of the differential matrix Riccati equation. In particular, he proved that solutions exist for all time when the initial condition is positive semidefinite, and he proved a number of important monotonicity properties, along with bounds on the solution stated in terms of the controllability and observability Gramians. Bucy [4] also studied the case when the solution of the autonomous Riccati equation converges to a solution of an associated (fixed-point) algebraic Riccati equation, and finally he proved exponential stability of the time-varying Kalman–Bucy filter along with an exponential forgetting property of the associated Riccati equation. We review and re-prove these results here via novel methods. We also refine quantitative estimates.

It is worth noting some history concerning Bucy's uniform bounds. The original upper and lower bounds on the (time-varying) Riccati equation given in [4] were particularly elegant in appearance, being given in terms of the relevant observability and controllability Gramians. However, as noted in [22], there was a crucial (yet commonly made) error in the proof which invalidated the result as given. This error was repeated (and/or overlooked) in numerous subsequent works, including by the current authors in the first writing of this work. A correction [5] was noted in a reply to [22]; see Bucy's reply [6] and a separate reply by Kalman [25]. However, a complete reworking of the result did not appear in entirety, it seems, until much later, in [18]. We remark that, in some sense, the qualitative nature of the Kalman–Bucy filter's

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 08/19/18 to 138.25.168.33. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

stability was not jeopardized, as noted by Bucy [6, 5] and Kalman [25]. However, given time-varying signal models, the lack of a complete proof on the uniform boundedness of the Riccati equation in quantitative terms was somewhat unsatisfactory.

Associated with the differential Riccati equation is a (fixed-point) algebraic equation whose solution(s) correspond to the equilibrium point(s) of the corresponding differential equation. This (fixed-point) algebraic equation was studied by Bucy in [9] and it was shown that there exists an unstable negative definite solution (in addition to the desired positive-definite equilibrium). A more detailed study of the algebraic Riccati equation was given by Willems [44], who considered characterizing every possible solution. Bucy [7] later considered the so-called structural stability of these solutions. See also [10, 45] and the early review paper [30] for related literature. We do not delve deeply into the fixed-point solutions here, as we are concerned mostly with the dynamical properties of the filter and the associated Riccati equation. When considering convergence to a fixed point of the Riccati flow, we are content to work under those assumptions that ensure an appropriate positive-definite equilibrium. Indeed, a stabilizing solution of the algebraic Riccati equation is shown to exist in [36] under mild detectability conditions.

Returning to the differential Riccati equation [4], convergence was studied extensively in [10] in the autonomous case, where a number of generalized conditions for convergence were given. We also note the early paper [46] that studied convergence and dealt further with a generalized version of the Riccati equation with a linear perturbation term. A geometric analysis of the differential Riccati equation and its solution(s) is given in [41] under time homogeneity. In [21] sufficient conditions are given such that the solution of the differential Riccati equation at any time is stabilizing; see also [34]. In later work [12, 11] convergence to a stabilizing positive-definite solution was studied again with further relaxations and where necessary conditions were addressed.

Finally, we point to the texts [38, 3, 1], dedicated to the Riccati equation, for further background and results (many of which are tangent to the discussion relevant here).

Given convergence of the differential Riccati flow and some associated semigroups, one typically concludes, in a straightforward way, the corresponding stability of the Kalman–Bucy filter; see the work of Bucy [27, 4] and the studies [2, 35]. However, we refine this conclusion in this work with exponential inequalities and some related results.

1.1. Statement of the main results and paper organization. Let $\|.\|_2$ be the Euclidean norm on $\mathbb{R}^r$, or the spectral norm on $\mathbb{R}^{r\times r}$, for some $r \geqslant 1$. We denote by $\mathbb{S}_r$ the set of $(r \times r)$ real symmetric matrices, and by $\mathbb{S}_r^+$ the subset of positive-definite matrices. To describe our main results with some precision we need to introduce some notation. For any $0 \leqslant s \leqslant t$ and $(x, Q) \in (\mathbb{R}^{r_1} \times \mathbb{S}_{r_1}^+)$ we let

$$\left(\varphi_{s,t}(x), \psi_{s,t}(x,Q), \overline{\psi}_{s,t}(x,Q), \phi_{s,t}(Q)\right) \in \left(\mathbb{R}^{r_1} \times \mathbb{R}^{r_1} \times \mathbb{R}^{r_1} \times \mathbb{S}_{r_1}^+\right)$$

be, respectively, the flow of the signal (1), the Kalman–Bucy filter (2), the Kalman–Bucy diffusion (4), and the Riccati equation (3). We take here the conventional observability/controllability conditions as holding; see the standing assumption (9) in section 2.

In section 2 we introduce the relevant signal model, the Kalman–Bucy filter, and an associated nonlinear diffusion process. This diffusion offers a novel interpretation of the Kalman–Bucy filter, i.e., as the conditional mean and covariance of an associated nonlinear McKean–Vlasov-type diffusion. This interpretation is interesting in its own

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

right, and allows one to *avoid Gaussian assumptions* on the initial state in a systematic way. We also introduce the concepts of observability and controllability, which are signal-related properties but which are relevant to the coming stability analysis. We also introduce the notion of a steady-state limit of the Riccati flow.

In section 3 we outline the relevant exponential and so-called Kalman–Bucy semigroups, associated largely with the Riccati flow and the stochastic flow of the Kalman–Bucy diffusion. We show how trajectories/solutions of the Riccati flow and Kalman–Bucy diffusion are defined in terms of these semigroups. Some preliminary technical lemmas are given concerning semigroup estimates and a number of invariance relationships are introduced. Both time-varying and homogeneous signal models are considered.

Section 4 is dedicated to the deterministic Riccati flow. Our first main result concerns the boundedness of the solution to the Riccati equation.

THEOREM 1.1. *There exist some $\upsilon > 0$, assuming standard uniform observability and controllability conditions, and some $\Lambda_{min}, \Lambda_{max} \in \mathbb{S}^+_{r_1}$ such that for any $t \geqslant \upsilon$ and any $Q \in \mathbb{S}^+_{r_1}$ we have*

$$\Lambda_{min} \leqslant \phi_t(Q) \leqslant \Lambda_{max}.$$

This result is stated precisely in section 4 as Theorem 4.4 where the upper- and lower-bounds are given in terms of the observability and controllability Gramians. Indeed, this theorem correctly upper- and lower-bounds the Riccati flow in terms of these Gramians and it corrects Bucy's erroneous bounds given in [4].

Section 4 is largely inspired by the seminal paper of Bucy [4] and we review, re-prove, correct (where necessary), and refine those major results here. For example, under basic conditions, we consider the Lipschitz continuity and existence of solutions to the Riccati matrix differential equation. Following Bucy's original work [4], a detailed proof of uniform convergence for the associated Kalman–Bucy semigroup is derived based on the corrected uniform bounds on the Riccati flow (and its inverse). This leads to a number of qualitative and *quantitative* contraction estimates for the semigroup and the Riccati flow, both with time-varying and time-invariant models (where convergence to the fixed point of the Riccati operator is then considered).

The first main result of this type is of the following form and is stated precisely in section 4 as Theorem 4.8.

THEOREM 1.2. *There exists some $\upsilon > 0$ such that for any $t \geqslant \upsilon$ and $Q_1, Q_2 \in \mathbb{S}^+_{r_1}$ we have*

$$\|\phi_t(Q_1) - \phi_t(Q_2)\|_2 \leqslant \alpha \exp\{-\beta t\} \|Q_1 - Q_2\|_2$$

for some parameters (α, β) whose values only depend on $(\Lambda_{min}, \Lambda_{max})$. The same inequality holds for any time $t \geqslant 0$ for some $\alpha = \alpha(Q_1, Q_2)$ that also depends on (Q_1, Q_2).

In section 5 we initiate a novel analysis on the convergence of Kalman–Bucy stochastic flows, both in the classical filtering form and the novel nonlinear diffusion form. The first main result of this type is the classical filtering stability result.

THEOREM 1.3. *There exists some $\upsilon > 0$ such that for any $t \geqslant s \geqslant \upsilon$ it follows that*

$$\sup_{Q \in \mathbb{S}^+_{r_1}} \left\| \mathbb{E}\left(\psi_{s,t}(x, Q) - \varphi_{s,t}(X_s) | X_s\right) \right\|_2 \leqslant \alpha \exp\{-\beta(t-s)\} \|x - X_s\|_2$$

with the parameters $\alpha, \beta > 0$ as given in Theorem 1.2.

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 08/19/18 to 138.25.168.33. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

This result is stated precisely in section 5 as Theorem 5.1 and it is interesting because it shows that the bias between the filter and the signal is exponentially stable regardless of the stability properties of the (time-varying) true signal. Much more is true, and we study exponential and comparison inequalities that bound with dedicated probability (at any time), the stochastic flow of the filter *sample paths* with respect to the underlying signal. That is, the next theorem shows that all the sample paths of the Kalman filter remain bounded close to the true signal with a large exponential probability. This result is stated precisely in section 5 as Theorem 5.2.

THEOREM 1.4. *The conditional probability of the events*

$$\|\psi_{s,t}(x,Q) - \varphi_{s,t}(X_s)\|_2 \leqslant \alpha_1(Q)e^{-\beta(t-s)}\|x - X_s\|_2 + \alpha_2(Q)\left[1 + \delta + \sqrt{\delta}\right],$$

given the state variable X_s, is greater than $1 - e^{-\delta}$ for any $\delta \geqslant 0$ and any $t \in [s, \infty[$, and some parameter $\beta > 0$ and some parameters $\alpha_i(Q)$ whose values only depend on Q with $i = 1, 2$.

In addition to this probabilistic convergence result, we give almost sure contraction-type estimates on the mean squared stochastic flow of the filter, conditioned on the underlying signal of interest. The next result is stated precisely in section 5 as Theorem 5.4.

THEOREM 1.5. *For any $t \geqslant s \geqslant 0$, $x_1, x_2 \in \mathbb{R}^{r_1}$, $Q_1, Q_2 \in \mathbb{S}^+_{r_1}$, and $n \geqslant 1$ we have the almost sure local contraction estimate*

$$\mathbb{E}\left(\|\psi_{s,t}(x_1,Q_1) - \psi_{s,t}(x_2,Q_2)\|_2^{2n}|X_s\right)^{\frac{1}{2n}} \leqslant \alpha_1(Q_1,Q_2)e^{-\beta(t-s)}\|x_1 - x_2\|_2$$
$$+ e^{-\beta(t-s)}\alpha_2(Q_1,Q_2)\left\{\|x_2 - X_s\|_2 + \sqrt{n}\right\}\|Q_1 - Q_2\|_2$$

for some $\beta > 0$ and some parameters $\alpha_i(Q_1, Q_2)$ whose values only depend on (Q_1, Q_2) with $i = 1, 2$.

The preceding two results concern the Kalman–Bucy filter. We also have analogous results for the nonlinear Kalman–Bucy diffusion, i.e., we show that all sample paths of the Kalman–Bucy diffusion follow the true signal with a large exponential probability and we provide an almost sure contraction-type estimate on the mean squared stochastic flow of both diffusions.

THEOREM 1.6. *The conditional probability of the events*

$$\|\overline{\psi}_{s,t}(x,Q) - \varphi_{s,t}(X_s)\|_2 \leqslant \alpha_1(Q)e^{-\beta(t-s)}\|x - X_s\|_2 + \alpha_2(Q)\left[1 + \delta + \sqrt{\delta}\right],$$

given the state variable X_s, is greater than $1 - e^{-\delta}$ for any $\delta \geqslant 0$ and any $t \in [s, \infty[$, and some parameter $\beta > 0$ and some parameters $\alpha_i(Q)$ whose values only depend on Q with $i = 1, 2$.

THEOREM 1.7. *For any $t \geqslant s \geqslant 0$, $x_1, x_2 \in \mathbb{R}^{r_1}$, $Q_1, Q_2 \in \mathbb{S}^+_{r_1}$, and $n \geqslant 1$ we have the almost sure local contraction estimate*

$$\mathbb{E}\left(\|\overline{\psi}_{s,t}(x_1,Q_1) - \overline{\psi}_{s,t}(x_2,Q_2)\|_2^{2n}|X_s\right)^{\frac{1}{2n}} \leqslant \alpha_1(Q_1,Q_2)e^{-\beta(t-s)}\|x_1 - x_2\|_2$$
$$+ e^{-\beta(t-s)}\alpha_2(Q_1,Q_2)\left\{\|x_2 - X_s\|_2 + \sqrt{n}\right\}\|Q_1 - Q_2\|_2$$

for some $\beta > 0$ and some parameters $\alpha_i(Q_1, Q_2)$ whose values only depend on (Q_1, Q_2) with $i = 1, 2$.

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 08/19/18 to 138.25.168.33. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

Downloaded 08/19/18 to 138.25.168.33. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

The preceding two results are stated precisely in section 5 as Theorems 5.5 and 5.6, respectively. Both offer a general notion of filter stability. To the best of our knowledge, this approach to studying the stability of Kalman–Bucy stochastic flows is novel, and indeed this is certainly true so far as the nonlinear Kalman–Bucy diffusion is concerned.

Throughout this article, attention is paid to the quantitative nature of the convergence and stability results. For example, we study exponential rates in the autonomous case in terms of different estimates on the relevant semigroups, and we track closely the related constants in front of these exponential terms. Our estimates are explicitly expressed with local Lipschitz contraction inequalities, dependent on the relevant signal matrix norms, etc. This contrasts with the classical analysis in [27, 8, 2, 35], which is purely qualitative in nature.

Carefully tracking constants is important for many applications, e.g., when studying the stability of ensemble Kalman filters [31, 42, 17] or extended Kalman filters [39, 16, 15], or when it comes to understanding general approximations of the Kalman filter and its error dependence on the state-space dimension [13, 37, 32].

1.2. Some basic notation. This section details some basic notation and terms used throughout the article.

With a slight abuse of notation, we denote by Id the identity matrix (with the size obvious from the context). The matrix transpose is denoted by $'$.

Denote by $\lambda_i(A)$, with $1 \leqslant i \leqslant r$, the nonincreasing sequence of eigenvalues of an $(r \times r)$-matrix A and let $\mathrm{Spec}(A)$ be the set of all eigenvalues. We often denote by $\lambda_{min}(A) = \lambda_r(A)$ and $\lambda_{max}(A) = \lambda_1(A)$ the minimal and the maximal eigenvalue, respectively. We set $A_{sym} := (A + A')/2$ for any $(r \times r)$-square matrix A. We define the logarithmic norm $\mu(A)$ of an $(r_1 \times r_1)$-square matrix A by

$$\begin{aligned}\mu(A) &:= \inf\left\{\alpha : \forall x, \langle x, Ax\rangle \leqslant \alpha \|x\|_2^2\right\} \\ &= \lambda_{max}(A_{sym}) \\ &= \inf\{\alpha : \forall t \geqslant 0, \|\exp(At)\|_2 \leqslant \exp(\alpha t)\}.\end{aligned} \tag{6}$$

The above equivalent formulations show that

$$\mu(A) \geqslant \varsigma(A) := \max\{\mathrm{Re}(\lambda) : \lambda \in \mathrm{Spec}(A)\},$$

where $\mathrm{Re}(\lambda)$ stands for the real part of the eigenvalues λ. The parameter $\varsigma(A)$ is often called the spectral abscissa of A. Also notice that A_{sym} is negative semidefinite as soon as $\mu(A) < 0$. The Frobenius matrix norm of a given $(r_1 \times r_2)$-matrix A is defined by

$$\|A\|_F^2 = \mathrm{tr}(A'A) \qquad \text{with the trace operator } \mathrm{tr}(\,.\,).$$

If A is a matrix $(r \times r)$, then we have $\|A\|_F^2 = \sum_{1\leqslant i,j\leqslant r} A(i,j)^2 \geqslant \|A\|^2$. For any $(r \times r)$-matrix A, we recall the norm equivalence formulae

$$\|A\|_2^2 = \lambda_{max}(A'A) \leqslant \mathrm{tr}(A'A) = \|A\|_F^2 \leqslant r\|A\|_2^2.$$

The Hoffmann–Wielandt theorem (Theorem 9.21 in [20]) also tells us that for any symmetric matrices A, B we have

$$\sum_{1\leqslant i\leqslant r_1} (\lambda_i(A) - \lambda_i(B))^2 \leqslant \|A - B\|_F^2 = \sum_{1\leqslant i\leqslant r_1} (\lambda_i(A - B))^2\,.$$

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 08/19/18 to 138.25.168.33. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

Now, given some random variable Z with some probability measure or distribution η and some measurable function f on some product space $\mathbb{R}^r$, we let

$$\eta(f) = \mathbb{E}(f(Z)) = \int f(x)\ \eta(dx)$$

be the integral of f w.r.t. η or the expectation of $f(Z)$. As a rule, any multivariate variable, say Z, is represented by a column vector and we use the transposition operator Z' to denote the row vector (similarly for matrices).

We also recall as background information that for any nonnegative random variable Z such that

$$\mathbb{E}\left(Z^{2n}\right)^{1/n} \leqslant z^2 n \quad \text{for some parameter } z \neq 0$$

and for any $n \geqslant 1$ we have

$$\mathbb{E}\left(Z^{2n}\right) \leqslant (z^2 n)^n \leqslant \frac{e}{\sqrt{2}} \left(\frac{e}{2} z^2\right)^n \mathbb{E}(V^{2n})$$

for some Gaussian and centered random variable V with unit variance. We check this claim using the Stirling approximation

$$\begin{aligned}\mathbb{E}(V^{2n}) &= 2^{-n}\frac{(2n)!}{n!}\\ &\geqslant e^{-1}2^{-n}\frac{\sqrt{4\pi n}(2n)^{2n}e^{-2n}}{\sqrt{2\pi n}n^n e^{-n}} = \sqrt{2}e^{-1}\left(\frac{2}{e}\right)^n n^n.\end{aligned}$$

By [14, Proposition 11.6.6], the probability of the event

$$(Z/z)^2 \leqslant \frac{e^2}{\sqrt{2}}\left[\frac{1}{2} + \left(\delta + \sqrt{\delta}\right)\right] \tag{7}$$

is greater than $1 - e^{-\delta}$ for any $\delta \geqslant 0$.

Given a real-valued continuous martingale M_t starting at the origin $M_0 = 0$, for any $n \geqslant 1$ and any time horizon $t \geqslant 0$ we have

$$\mathbb{E}\left(\sup_{0\leqslant s\leqslant t} |M_s|^n\right)^{1/n} \leqslant 2\sqrt{2}\sqrt{n}\mathbb{E}\left(\langle M\rangle_t^{n/2}\right)^{1/n}. \tag{8}$$

Proofs of these Burkholder–Davis–Gundy inequalities can be found in [40]; see also [28, Theorem B.1, p. 97].

2. Description of the models.

2.1. The Kalman–Bucy filter.

In general (i.e., not assuming that X_0 is Gaussian), for any $0 \leqslant s \leqslant t$ we define the stochastic flow

$$\Phi_{s,t} : (x, Q) \in (\mathbb{R}^{r_1} \times \mathbb{S}^+_{r_1}) \mapsto \Phi_{s,t}(x, Q) = (\psi_{s,t}(x, Q), \phi_{s,t}(Q)) \in (\mathbb{R}^{r_1} \times \mathbb{S}^+_{r_1})$$

as describing the Kalman–Bucy filter, where for any horizon s and any time $t \in [s, \infty[$ we have

$$\begin{cases} d\psi_{s,t}(x, Q) = \left[A_t - \phi_{s,t}(Q)S_t\right]\psi_{s,t}(x, Q)dt + \phi_{s,t}(Q)C_t'R_2^{-1}dY_t,\\ \partial_t\phi_{s,t}(Q) = \operatorname{Ricc}\left(\phi_{s,t}(Q)\right) \quad \text{with} \quad \Phi_{s,s}(x, Q) = (x, Q).\end{cases}$$

With similar notation, we also denote by $\varphi_{s,t}(x)$ the stochastic flow of the signal process

$$d\varphi_{s,t}(x) = A_t\varphi_{s,t}(x)dt + R_1^{1/2}dW_t \quad \text{with} \quad \varphi_{s,s}(x) = x$$

for any $t \in [s, \infty[$ and any $x \in \mathbb{R}^{r_1}$.

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 08/19/18 to 138.25.168.33. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

Note that, in general, $\phi_{s,t}(\phi_s(Q)) = \phi_t(Q)$,

$$\phi_{s+t}(Q) = \phi_{s,s+t}(\phi_s(Q)), \quad \text{and} \quad \phi_{s,t}(Q) = \phi_{s+u,t}(\phi_{s,s+u}(Q)), 0 \leqslant u \leqslant t-s.$$

Observe that when the signal is time invariant, then so is the Riccati equation and thus

$$\phi_{s,s+t}(Q) = \phi_t(Q) =: \phi_{0,t}(Q) \text{ or } \phi_{s,t}(Q) = \phi_{t-s}(Q) = \phi_{u,t-s}(\phi_u(Q)), \quad 0 \leqslant u \leqslant t-s,$$

along with numerous other (equivalent) combinations of stationary shifts.

2.2. Nonlinear Kalman–Bucy diffusions. For any $0 \leqslant s \leqslant t$ we let

$$\overline{\Phi}_{s,t} : (x,Q) \in (\mathbb{R}^{r_1} \times \mathbb{S}^+_{r_1}) \mapsto \overline{\Phi}_{s,t}(x,Q) = \left(\overline{\psi}_{s,t}(x,Q), \phi_{s,t}(Q)\right) \in (\mathbb{R}^{r_1} \times \mathbb{S}^+_{r_1})$$

be the stochastic flow of the Kalman–Bucy diffusion; that is, for any time horizon s and any time $t \in [s,\infty[$ we have

$$\begin{aligned} d\overline{\psi}_{s,t}(x,Q) &= [A_t - \phi_{s,t}(Q)S_t]\,\overline{\psi}_{s,t}(x,Q)dt + \phi_{s,t}(Q)C_t'R_2^{-1}dY_t \\ &\quad + R_1^{1/2}d\overline{W}_t - \phi_{s,t}(Q)C_t'R_2^{-1/2}d\overline{V}_t \end{aligned}$$

with $\overline{\psi}_{s,s}(x,Q) = x$ for $t = s$.

2.3. Observability and controllability conditions. We consider the observability and controllability Gramians $\mathcal{O}_{s,t}$ and $\mathcal{C}_{s,t}$ defined by

$$\mathcal{C}_{s,t} := \int_s^t \exp\left[\oint_r^t A_u\ du\right] R_1 \exp\left[\oint_r^t A_u\ du\right]' dr$$

and

$$\mathcal{O}_{s,t} := \int_s^t \exp\left[\oint_t^r A_u\ du\right]' S_r \exp\left[\oint_t^r A_u\ du\right] dr$$

for all $t \geqslant s \geqslant 0$. We let $\mathcal{C}_t := \mathcal{C}_{0,t}$ and $\mathcal{O}_t := \mathcal{O}_{0,t}$. Here $\exp\left[\oint_s^t A_u du\right]$ defines a semigroup associated with the matrix flow. We return to this semigroup and its properties later; see (12) for a more precise definition.

We take the following assumption as holding in the statement of all results.

Standing assumption. *There exist parameters* $\upsilon, \varpi_\pm^{o,c} > 0$ *such that*

$$\varpi_-^c Id \leqslant \mathcal{C}_{t,t+\upsilon} \leqslant \varpi_+^c Id \quad \text{and} \quad \varpi_-^o Id \leqslant \mathcal{O}_{t,t+\upsilon} \leqslant \varpi_+^o Id \tag{9}$$

uniformly for all $t \geqslant 0$. *The parameter* υ *is called the interval of observability/controllability.*

Note that if the signal matrices are time invariant, then the pair $(A, R_1^{1/2})$ is a controllable and (A,C) is observable if

$$\left[R_1^{1/2}, A\left(R_1^{1/2}\right), \dots, A^{r_1-1}R_1^{1/2}\right] \quad \text{and} \quad \begin{bmatrix} C \\ CA \\ \vdots \\ CA^{r_1-1} \end{bmatrix} \tag{10}$$

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

both have rank r_1. Under these conditions, there always exist parameters $\upsilon, \varpi_{\pm}^{o,c} > 0$ ensuring that (9) holds. For example, whenever the signal drift matrix A is diagonalizable we can choose

$$\varpi_{-}^{c} = \lambda_{min}(R) \min_{\lambda \in \mathrm{Spec}(A)} \frac{e^{2\lambda\upsilon} - 1}{2\lambda} \leqslant \varpi_{+}^{c} = \lambda_{max}(R_1) \max_{\lambda \in \mathrm{Spec}(A)} \frac{e^{2\lambda\upsilon} - 1}{2\lambda}$$

as well as

$$\varpi_{-}^{o} = \lambda_{min}(S) \min_{\lambda \in \mathrm{Spec}(A)} \frac{1 - e^{-2\lambda\upsilon}}{2\lambda} \leqslant \varpi_{+}^{o} = \lambda_{max}(S) \max_{\lambda \in \mathrm{Spec}(A)} \frac{1 - e^{-2\lambda\upsilon}}{2\lambda}$$

for any $\upsilon > 0$.

In the time-invariant case, these conditions (10) are sufficient (but not necessary [36, 11]) to ensure that there exists a (unique) positive-definite fixed-point matrix $P = \phi_t(P)$ solving the so-called algebraic Riccati equation

$$\mathrm{Ricc}(P) := AP + PA' - PSP + R_1 = 0. \tag{11}$$

In this case, the matrix difference $A - PS$ is asymptotically stable even when the signal matrix A is unstable. Relaxed conditions for this solution to exist are discussed widely in the literature and we highlight the important works [44, 36, 11]. In this setting, our results are concerned with convergence to the fixed point P. We note that the stability of $A - PS$ follows from Theorems 4.8 and 5.1 and it follows that $\phi_t(Q)$ for $Q \in \mathbb{S}_{r_1}^{+}$ converges to the fixed point P due to Theorem 4.8 and the Banach fixed-point theorem.

In the time-varying case, we are interested in asymptotic stability results and results that tend to bound the Riccati flow $\phi_t(Q)$ uniformly on both sides by the controllability and observability Gramians. In this setting, there is typically no fixed point for the flow $\phi_t(Q)$ and the difference $A - \phi_t(Q)S$ need not be a stable matrix at any instant in general; see also [21]. Actually, $A - \phi_t(Q)S$ need not be stable in both the time-invariant and time-varying settings.

Since we switch between time-invariant signal models and time-varying models (in which no fixed point of (3) generally exists), we choose not to relax our observability/controllability assumptions, e.g., [44, 36, 11]. In fact, as discussed in [11], there is really no generality lost by assuming observability over the (arguably) weaker detectability condition [36], even in the time-invariant case. One may substitute a generalized controllability condition such as that studied in [2]; however, only weaker results are achievable (as shown by example in [2]).

3. Semigroups of the Riccati flow and the Kalman–Bucy filter.

3.1. Exponential semigroups. The transition matrix associated with a smooth flow of $(r \times r)$-matrices $A : u \mapsto A_u$ is denoted by

$$\mathcal{E}_{s,t}(A) = \exp\left[\oint_s^t A_u du\right] \Longleftrightarrow \partial_t \mathcal{E}_{s,t}(A) = A_t \mathcal{E}_{s,t}(A) \quad \text{and} \quad \partial_s \mathcal{E}_{s,t}(A) = -\,\mathcal{E}_{s,t}(A) A_s \tag{12}$$

for any $s \leqslant t$ with $\mathcal{E}_{s,s} = Id$, the identity matrix.

The following technical lemma gives a pair of semigroup estimates for the state transition matrices associated with a sum of drift-type matrices.

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 08/19/18 to 138.25.168.33. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

Downloaded 08/19/18 to 138.25.168.33. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

LEMMA 3.1. *Let $A : u \mapsto A_u$ and $B : u \mapsto B_u$ be the smooth flows of $(r \times r)$-matrices. For any $s \leqslant t$ we have*

$$\|\mathcal{E}_{s,t}(A+B)\|_2 \leqslant \exp\left(\int_s^t \mu(A_u)du + \int_s^t \|B_u\|_2 du\right).$$

In addition, for the matrix spectral, or Frobenius, norm $\|\cdot\|$ we have

$$\|\mathcal{E}_{s,t}(A+B)\| \leqslant \alpha_A \exp\left[-\beta_A(t-s) + \alpha_A \int_s^t \|B_u\|du\right]$$

as soon as

$$\forall 0 \leqslant s \leqslant t, \qquad \|\mathcal{E}_{s,t}(A)\| \leqslant \alpha_A \exp\left(-\beta_A(t-s)\right).$$

Proof. The above estimate is a direct consequence of the matrix log-norm inequality

$$\mu(A_t + B_t) \leqslant \mu(A_t) + \mu(B_t) \quad \text{and the fact that} \quad \mu(B_t) \leqslant \|B_t\|_2.$$

This ends the proof of the first assertion. To check the second assertion we observe that

$$\partial_t \mathcal{E}_{s,t}(A+B) = (\partial_t \mathcal{E}_t(A+B))\, \mathcal{E}_s(A+B)^{-1} = A_t \mathcal{E}_{s,t}(A+B) + B_t \mathcal{E}_{s,t}(A+B).$$

This implies that

$$\mathcal{E}_{s,t}(A+B) = \mathcal{E}_{s,t}(A) + \int_s^t \mathcal{E}_{u,t}(A) B_u \mathcal{E}_{s,u}(A+B) du$$

for any $s \leqslant t$, from which we prove that

$$\begin{aligned} e^{\beta_A(t-s)}\|\mathcal{E}_{s,t}(A+B)\| &\leqslant \alpha_A + \alpha_A \int_s^t e^{\beta_A(t-s)} e^{-\beta_A(t-u)} \|B_u\| \|\mathcal{E}_{s,u}(A+B)\| du \\ &= \alpha_A + \alpha_A \int_s^t \|B_u\| e^{\beta_A(u-s)} \|\mathcal{E}_{s,u}(A+B)\| du. \end{aligned}$$

By Grönwall's lemma this implies that

$$e^{\beta_A(t-s)}\|\mathcal{E}_{s,t}(A+B)\| \leqslant \alpha_A \exp\left[\int_s^t \alpha_A \|B_u\| du\right].$$

This ends the proof of the lemma. ∎

3.1.1. Time-invariant exponential semigroups. For time-invariant matrices $A_t = A$, the state transition matrix reduces to the conventional matrix exponential

$$\mathcal{E}_{s,t}(A) = e^{(t-s)A} = \mathcal{E}_{t-s}(A).$$

In this subsection we are interested in estimating the norm of $\mathcal{E}_t(A)$. We state the following general convergence result on the time-invariant semigroup generated by the matrix difference $A - PS$, where P is defined by (11).

LEMMA 3.2. *Under the time-invariant observability/controllability rank conditions* (10)*, it follows that*

$$\exists \nu > 0, \exists \kappa < \infty : \forall t \geqslant 0, \quad \|e^{t(A-PS)}\|_2 \leqslant \kappa\, e^{-\nu t}. \tag{13}$$

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Proof. The observability/controllability rank conditions (10) are sufficient to ensure the existence of a (unique) positive-definite solution P of (11) and that $\varsigma(A-PS) < 0$; see [36]. We know that

$$\|e^{t(A-PS)}\|_2 \leqslant e^{\mu(A-PS)t},$$

which applies here whenever $\mu(A-PS) < 0$. Otherwise, we can use any of the estimates presented below in (14), (15), (16), and (17). ∎

The norm of $\mathcal{E}_t(A)$ can be estimated in a number of ways. The first is based on the Jordan decomposition $T^{-1}AT = J$ of the matrix A in terms of k Jordan blocks associated with the eigenvalues with multiplicities m_i, with $1 \leqslant i \leqslant k$. In this situation, we have the Jordan type estimate

$$e^{\varsigma(A)t} \leqslant \|\mathcal{E}_t(A)\|_2 \leqslant \kappa_{Jor,t}(T)\, e^{\varsigma(A)t} \tag{14}$$

with

$$\kappa_{Jor,t}(T) := \left(\max_{0\leqslant j<n} \frac{t^j}{j!}\right) \|T\|_2 \|T^{-1}\|_2 \quad \text{and} \quad n := \max_{1\leqslant i\leqslant k} m_i.$$

Note that $\kappa_{Jor,t}(T)$ depends on the time horizon t as soon as A is not of full rank. In addition, whenever A is close to singular, the condition number $\text{cond}(T) := \|T\|_2\|T^{-1}\|_2$ tends to be very large. When A is diagonalizable, the above estimate becomes

$$e^{\varsigma(A)t} \leqslant \|\mathcal{E}_t(A)\|_2 \leqslant \text{cond}(T)\, e^{\varsigma(A)t}. \tag{15}$$

Another method is based on the Schur decomposition $U'AU = D + T$ in terms of a unitary matrix U with $D = \text{diag}(\lambda_1(A), \ldots, \lambda_r(A))$ and a strictly triangular matrix T such that $T_{i,j} = 0$ for any $i \geqslant j$. In this case we have the Schur-type estimate

$$\|\mathcal{E}_t(A)\|_2 \leqslant \kappa_{Sch,t}(T)\, e^{\varsigma(A)t} \quad \text{with} \quad \kappa_{Sch,t}(T) := \sum_{0\leqslant i\leqslant r} \frac{(\|T\|_2 t)^i}{i!}. \tag{16}$$

The proof of these estimates can be found in [43, 33]. In both cases, for any $\epsilon \in]0,1]$ and any $t \geqslant 0$ we have

$$e^{\varsigma(A)t} \leqslant \|\mathcal{E}_t(A)\|_2 \leqslant \kappa_A(\epsilon)\, e^{(1-\epsilon)\varsigma(A)t} \tag{17}$$

for some constants $\kappa_A(\epsilon)$ whose values only depend on the parameters ϵ. When A is asymptotically stable, that is, all its eigenvalues have negative real parts, for any positive definite matrix B we have

$$e^{\varsigma(A)t} \leqslant \|\mathcal{E}_t(A)\|_2 \leqslant \text{cond}(T)\, \exp\left[-t/\|B^{-1/2}TB^{-1/2}\|_2\right]$$

with the positive-definite matrix

$$T = \int_0^\infty e^{A't}Be^{At}dt \Longleftrightarrow A'T + TA = -B.$$

The proof of these estimates can be found in [29, see, e.g., Theorem 13.6].

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 08/19/18 to 138.25.168.33. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

Downloaded 08/19/18 to 138.25.168.33. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

3.2. Kalman–Bucy semigroups. For any $s \leqslant t$ and $Q_1, Q_2 \in \mathbb{S}^+_{r_1}$ we set

$$E_{s,t}(Q_1,Q_2) := \exp\left[\oint_s^t \left(A_u - \frac{\phi_u(Q_1)+\phi_u(Q_2)}{2} S_u\right) du\right] \quad \text{and}$$
$$E_{s,t}(Q_1) := E_{s,t}(Q_1,Q_1).$$

When $s = 0$ we write $E_t(Q_1)$ and $E_t(Q_1,Q_2)$ in place of $E_{0,t}(Q_1)$ and $E_{0,t}(Q_1,Q_2)$. In this notation we have

$$E_{s,t}(Q_1,Q_2) = E_t(Q_1,Q_2)E_s(Q_1,Q_2)^{-1} \quad \text{and} \quad E_{s,t}(Q_1) = E_t(Q_1)E_s(Q_1)^{-1}.$$

We have the following important result.

PROPOSITION 3.3. *For any $s \leqslant t$ and $Q_1, Q_2 \in \mathbb{S}^+_{r_1}$ we have*

$$\phi_t(Q_1) - \phi_t(Q_2) = E_{s,t}(Q_1)\left[\phi_s(Q_1) - \phi_s(Q_2)\right] E_{s,t}(Q_2)' \tag{18}$$
$$= E_{s,t}(Q_1,Q_2)\left[\phi_s(Q_1) - \phi_s(Q_2)\right] E_{s,t}(Q_1,Q_2)' \tag{19}$$

as well as

$$\begin{aligned} \phi_t(Q_1) - \phi_t(Q_2) = {} & E_{s,t}(Q_2)\left[\phi_s(Q_1) - \phi_s(Q_2)\right] E_{s,t}(Q_2)' \\ & - \int_s^t E_{u,t}(Q_2)\left[\phi_u(Q_1) - \phi_u(Q_2)\right] S_u \\ & \times \left[\phi_u(Q_1) - \phi_u(Q_2)\right] E_{u,t}(Q_2)' du. \end{aligned} \tag{20}$$

Proof. These semigroup formulae are direct consequences of the following three polarization-type formulae:

$$\begin{aligned} \mathrm{Ricc}(Q_1) - \mathrm{Ricc}(Q_2) &= (A_t - Q_1S_t)(Q_1 - Q_2) + (Q_1 - Q_2)(A_t - Q_2S_t)' \\ &= \left[A_t - \frac{1}{2}(Q_1+Q_2)S_t\right](Q_1 - Q_2) + (Q_1 - Q_2)\left[A_t - \frac{1}{2}(Q_1+Q_2)S_t\right]' \\ &= (A_t - Q_2S_t)(Q_1 - Q_2) + (Q_1 - Q_2)(A_t - Q_2S_t)' - (Q_1 - Q_2)S_t(Q_1 - Q_2), \end{aligned} \tag{21}$$

where the first line implies (18), the second line implies (19), and the third line implies (20). We check these polarization-type formulae using the decompositions

$$\begin{aligned} Q_1S_tQ_1 - Q_2S_tQ_2 &= Q_1S_t(Q_1 - Q_2) + (Q_1 - Q_2)S_tQ_2 \\ &= \frac{1}{2}(Q_1+Q_2)S_t(Q_1 - Q_2) + \frac{1}{2}(Q_1 - Q_2)S_t(Q_1+Q_2) \\ &= (Q_1 - Q_2)S_t(Q_1 - Q_2) + Q_2S(Q_1 - Q_2) + (Q_1 - Q_2)S_tQ_2. \end{aligned}$$

The proofs of (18), (19), and (20) follow from basic calculations; e.g., as for (19) we have,

$$\begin{aligned} \partial_t(\phi_t(Q_1) - \phi_t(Q_2)) &= \mathrm{Ricc}(\phi_t(Q_1)) - \mathrm{Ricc}(\phi_t(Q_2)) \\ &= \left(A_t - \frac{\phi_t(Q_1)+\phi_t(Q_2)}{2}S_t\right)(\phi_t(Q_1) - \phi_t(Q_2)) \\ &\quad + (\phi_t(Q_1) - \phi_t(Q_2))\left(A_t - \frac{\phi_t(Q_1)+\phi_t(Q_2)}{2}S_t\right)'. \end{aligned}$$

Now the solution of this linear equation is given by $E_{s,t}(Q_1,Q_2)\left[\phi_s(Q_1) - \phi_s(Q_2)\right]$ $E_{s,t}(Q_1,Q_2)'$, which is (19). This ends the proof of the proposition. ∎

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 08/19/18 to 138.25.168.33. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

Given a time-varying signal model, it is useful to define some additional notation. For any $s \leqslant u \leqslant t$ and $Q \in \mathbb{S}_{r_1}^+$ we set

$$E_{t|s}(Q) := \exp\left[\oint_s^t (A_u - \phi_{s,u}(Q)S_u)\,du\right] \quad \text{and}$$
$$E_{u,t|s}(Q) := \exp\left[\oint_u^t (A_r - \phi_{s,r}(Q)S_r)\,dr\right]$$

with $E_{u,t|s}(Q) = E_{t|s}(Q)E_{u|s}(Q)^{-1}$. Note that there is a relationship between $E_{t|s}$ and $E_{s,t}$ in the following sense:

$$\begin{aligned} E_{t|s}(\phi_s(Q)) &= \exp\left[\oint_s^t (A_u - \phi_{s,u}(\phi_s(Q))S_u)\,du\right] \\ &= \exp\left[\oint_s^t (A_u - \phi_u(Q)S_u)\,du\right] = E_{s,t}(Q), \end{aligned}$$

where we simply used $\phi_{s,t}(\phi_s(Q)) = \phi_t(Q)$.

4. Riccati flows. We start this section with a preliminary result concerning the monotonicity of the Riccati operator, some basic boundedness results, and a Lipschitz estimate.

PROPOSITION 4.1. *The Riccati flow $Q \mapsto \phi_t(Q)$ is a nondecreasing function w.r.t. the Loewner partial order; that is, we have*

$$Q_1 \leqslant Q_2 \Longleftrightarrow \phi_t(Q_1) \leqslant \phi_t(Q_2).$$

For any $Q_1, Q_2 \in \mathbb{S}_{r_1}^+$ we have the local Lipschitz inequality

$$\|\phi_t(Q_1) - \phi_t(Q_2)\|_F \leqslant l_{Q_1,Q_2}(\phi_t)\|Q_1 - Q_2\|_F \tag{22}$$

for some Lipschitz constant

$$l_{Q_1,Q_2}(\phi_t) \leqslant [\|E_t(Q_1)\|_2 \, \|E_t(Q_2)\|_2] \wedge \|E_t(Q_1,Q_2)\|_2 < \infty.$$

Proof. Using Proposition 3.3 we prove that $Q \mapsto \phi_t(Q)$ is an nondecreasing function w.r.t. the Loewner partial order. The Lipschitz estimate (22) is a direct consequence of the implicit semigroup formulae (18) and (19). ∎

It follows that for any $Q \in \mathbb{S}_{r_1}^+$ the time-varying Riccati flow $\phi_t(Q)$ is well defined and a unique solution exists for all $t \geqslant 0$, since the local Lipschitz estimate is "global" on any finite interval.

The Riccati flow $Q \mapsto \phi_t(Q)$ also depends monotonically on the parameters S and R_1.

COROLLARY 4.2. *Let $R_1^2 \geqslant R_1^1 \in \mathbb{S}_{r_1}^+$ and $S_t^1 \geqslant S_t^2 \in \mathbb{S}_{r_1}^+$ for all $t \geqslant 0$. Then $\phi_{s,t}(Q,2) \geqslant \phi_{s,t}(Q,1)$, where*

$$\partial_t \phi_{s,t}(Q,i) = A_t Q + Q A_t' - Q S_t^i Q + R_1^i, \; i \in \{1,2\}.$$

Now define

$$\|\phi(Q)\|_2 := \sup_{t \geqslant 0} \|\phi_t(Q)\|_2 < \infty, \tag{23}$$

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 08/19/18 to 138.25.168.33. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

which is always uniformly bounded for small enough t as a result of the Lipschitz estimate on $[0,t]$.

In the time-invariant setting, when the desired solution P of (11) exists, the following result characterizes a uniform upper bound on the Riccati flow and a bound on its growth.

Proposition 4.3. *The Riccati flow obeys*

$$Q \mapsto \phi_t(Q) \leqslant P + E_t(P)(Q-P)E_t(P)'.$$

In addition, for any $Q \in \mathbb{S}^+_{r_1}$ *we have the uniform estimates*

$$\|\phi(Q)\|_2 \leqslant \|P\|_2 + \kappa^2\|Q-P\|_2 \quad \text{and} \quad \sup_{t>0} t^{-1}\|\phi_t(Q)-Q\|_2 < \infty, \tag{24}$$

where κ *is defined in Lemma* 3.2.

Proof. Choosing $Q_2 = P$ and $s = 0$ in (20), we find that

$$\begin{aligned}\phi_t(Q) - P = &\ E_t(P)\,[Q-P]\,E_t(P)' \\ &- \int_0^t E_{u,t}(P)\ [\phi_u(Q)-P]\ S\ [\phi_u(Q)-P]\ E_{u,t}(P)'du.\end{aligned}$$

This implies that

$$0 \leqslant \phi_t(Q) \leqslant P + E_t(P)(Q-P)E_t(P)' \Rightarrow \|\phi_t(Q)\|_2 \leqslant \|P\|_2 + \|E_t(P)\|_2^2\|Q-P\|_2.$$

It then follows that $\|\phi(Q)\|_2 \leqslant \|P\|_2 + \kappa^2\|Q-P\|_2$, from which we conclude that

$$\phi_t(Q) = Q + \int_0^t \mathrm{Ricc}(\phi_s(Q))ds \Rightarrow \|\phi_t(Q)-Q\|_F \leqslant c_Q t$$

for some finite constant c_Q whose values only depend on Q. This completes the proof. □

4.1. Uniform bounds on the Riccati flow. We let $\mathcal{C}_t(\mathcal{O})$ and $\mathcal{O}_t(\mathcal{C})$ be the Gramian matrices defined by

$$\begin{aligned}\mathcal{O}_t(\mathcal{C}) &:= \mathcal{C}_t^{-1}\left[\int_0^t \mathcal{E}_{s,t}(A)\mathcal{C}_s S_s \mathcal{C}_s \mathcal{E}_{s,t}(A)'ds\right]\mathcal{C}_t^{-1}, \\ \mathcal{C}_t(\mathcal{O}) &:= \mathcal{O}_t^{-1}\left[\int_0^t \mathcal{E}'_{s,t}(A)^{-1}\mathcal{O}_s R_1 \mathcal{O}_s \mathcal{E}_{s,t}(A)^{-1}ds\right]\mathcal{O}_t^{-1}.\end{aligned}$$

Under our standard observability and controllability assumptions (9), there exist some parameters $\varpi^c_\pm(\mathcal{O}), \varpi^o_\pm(\mathcal{C}) > 0$ such that

$$\varpi^c_-(\mathcal{O})Id \leqslant \mathcal{C}_\upsilon(\mathcal{O}) \leqslant \varpi^c_+(\mathcal{O})Id \qquad \text{and} \qquad \varpi^o_-(\mathcal{C})Id \leqslant \mathcal{O}_\upsilon(\mathcal{C}) \leqslant \varpi^o_+(\mathcal{C})Id$$

hold uniformly on the interval $\upsilon > 0$ of observability/controllability.

The main objective of this section is to prove the following theorem.

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 08/19/18 to 138.25.168.33. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

THEOREM 4.4. *For any $t \geqslant \upsilon$ and any $Q \in \mathbb{S}^+_{r_1}$ we have*

$$\left(\mathcal{O}_\upsilon(\mathcal{C}) + \mathcal{C}_\upsilon^{-1}\right)^{-1} \leqslant \phi_t(Q) \leqslant \mathcal{O}_\upsilon^{-1} + \mathcal{C}_\upsilon(\mathcal{O}).$$

In addition, this implies

$$\left(\mathcal{O}_\upsilon(\mathcal{C}) + \mathcal{C}_\upsilon^{-1}\right)^{-1} \leqslant P \leqslant \mathcal{O}_\upsilon^{-1} + \mathcal{C}_\upsilon(\mathcal{O}) \quad \textit{and}$$
$$\left(\mathcal{O}_\upsilon^{-1} + \mathcal{C}_\upsilon(\mathcal{O})\right)^{-1} \leqslant P^{-1} \leqslant \mathcal{O}_\upsilon(\mathcal{C}) + \mathcal{C}_\upsilon^{-1}.$$

The following corollary is immediate.

COROLLARY 4.5. *For any $Q \in \mathbb{S}^+_{r_1}$ and any $t \geqslant \upsilon$ we have*

$$\mathrm{Spec}(\phi_t(Q)) \quad \textit{and} \quad \mathrm{Spec}(P) \quad \subset \quad \left[\left(\varpi_+^o(\mathcal{C}) + 1/\varpi_-^c\right)^{-1}, \varpi_+^c(\mathcal{O}) + 1/\varpi_-^o\right]$$

Note that this corollary, together with the definition (23), following the proof of Proposition 4.3, yields the growth estimate

$$\sup_{t \geqslant \upsilon} t^{-1} \|\phi_t(Q) - Q\|_2 < \infty$$

in the general setting with time-varying signal models.

The proof of the theorem is based on comparison inequalities between the Riccati flow and the flow of matrices defined below.

We let

$$Q \mapsto \phi_t^o(Q) \quad \text{and} \quad Q \mapsto \phi_t^c(Q),$$

with the flows associated with the Riccati equation with drift functions Ricc^o and Ricc^c defined by

$$\mathrm{Ricc}^c(Q) := A_t Q + Q A_t' + R_1,$$
$$\mathrm{Ricc}^o(Q) := A_t Q + Q A_t' - Q S_t Q = \mathrm{Ricc}(Q) - R_1.$$

LEMMA 4.6. *For any $t \geqslant \upsilon$ we have*

$$\mathcal{C}_t \leqslant \phi_t^c(Q) = \mathcal{E}_t(A) Q \mathcal{E}_t(A)' + \mathcal{C}_t \quad \textit{and} \tag{25}$$
$$\phi_t^o(Q) = \mathcal{E}_t(A)\left(Q^{-1} + \overline{\mathcal{O}}_t\right)^{-1} \mathcal{E}_t(A)' \leqslant \mathcal{O}_t^{-1}$$

with

$$\overline{\mathcal{O}}_t := \mathcal{E}_t(A)' \mathcal{O}_t \mathcal{E}_t(A) = \int_0^t \mathcal{E}_s(A)' S_s \mathcal{E}_s(A) ds.$$

In addition, for any $t \geqslant \upsilon$ we have the estimates

$$\phi_t^o(Q) \leqslant \phi_t(Q) \leqslant \mathcal{O}_t^{-1} + \mathcal{C}_t(\mathcal{O})$$

as well as

$$\sup_{t \geqslant \upsilon} \phi_t(Q) \leqslant \mathcal{O}_\upsilon^{-1} + \mathcal{C}_\upsilon(\mathcal{O}) \quad \textit{and} \quad 0 < \left(\mathcal{O}_\upsilon^{-1} + \mathcal{C}_\upsilon(\mathcal{O})\right)^{-1} \leqslant \inf_{t \geqslant \upsilon} \phi_t(Q)^{-1}.$$

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 08/19/18 to 138.25.168.33. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

Proof. The left-hand side inequality of (25) is immediate. We check the right-hand side inequality of (25) using the fact that

$$\begin{aligned}\partial_t\phi^o_t(Q) &= (\partial_t\mathcal{E}_t(A))\left(Q^{-1}+\overline{\mathcal{O}}_t\right)^{-1}\mathcal{E}_t(A)' \\ &\quad + \mathcal{E}_t(A)\left(Q^{-1}+\overline{\mathcal{O}}_t\right)^{-1}\left(\partial_t\mathcal{E}_t(A)'\right) + \mathcal{E}_t(A)\left[\partial_t\left(Q^{-1}+\overline{\mathcal{O}}_t\right)^{-1}\right]\mathcal{E}_t(A)' \\ &= A_t\phi^o_t(Q)+\phi^o_t(Q)A_t' + \mathcal{E}_t(A)\left[\partial_t\left(Q^{-1}+\overline{\mathcal{O}}_t\right)^{-1}\right]\mathcal{E}_t(A)'.\end{aligned}$$

On the other hand, recalling the inverse derivation formula

$$\partial_t M_t^{-1} = -M_t^{-1}\left(\partial_t M_t\right)M_t^{-1},$$

we find via Leibniz's rule that

$$\begin{aligned}\partial_t\left(Q^{-1}+\overline{\mathcal{O}}_t\right)^{-1} &= -\left(Q^{-1}+\overline{\mathcal{O}}_t\right)^{-1}\underbrace{\left[\partial_t\left(Q^{-1}+\overline{\mathcal{O}}_t\right)\right]}_{=\mathcal{E}_t(A)'S_t\mathcal{E}_t(A)}\left(Q^{-1}+\overline{\mathcal{O}}_t\right)^{-1} \\ &= -\left\{\left(Q^{-1}+\overline{\mathcal{O}}_t\right)^{-1}\mathcal{E}_t(A)'\right\}S_t\left\{\mathcal{E}_t(A)\left(Q^{-1}+\overline{\mathcal{O}}_t\right)^{-1}\right\}.\end{aligned}$$

This implies that

$$\begin{aligned}&\mathcal{E}_t(A)\left[\partial_t\left(Q^{-1}+\overline{\mathcal{O}}_t\right)^{-1}\right]\mathcal{E}_t(A)' \\ &\quad = -\left\{\mathcal{E}_t(A)\left(Q^{-1}+\overline{\mathcal{O}}_t\right)^{-1}\mathcal{E}_t(A)'\right\}S_t\left\{\mathcal{E}_t(A)\left(Q^{-1}+\overline{\mathcal{O}}_t\right)^{-1}\mathcal{E}_t(A)'\right\} \\ &\quad = -\phi^o_t(Q)S_t\phi^o_t(Q).\end{aligned}$$

We also have

$$\begin{aligned}\mathcal{E}_t(A)\left(Q^{-1}+\overline{\mathcal{O}}_t\right)^{-1}\mathcal{E}_t(A)' &\leqslant \mathcal{E}_t(A)\,\overline{\mathcal{O}}_t^{-1}\mathcal{E}_t(A)' \\ &= \mathcal{E}_t(A)\,\mathcal{E}_t(A)^{-1}\mathcal{O}_t^{-1}(\mathcal{E}_t(A)^{-1})'\,\mathcal{E}_t(A)' = \mathcal{O}_t^{-1}.\end{aligned}$$

This ends the proof of (25). Also observe that

$$\mathrm{Ricc}(Q_1)-\mathrm{Ricc}^o(Q_2) = \mathrm{Ricc}(Q_1)-\mathrm{Ricc}(Q_2)+R_1.$$

Using the polarization formulae (21), we conclude that

$$\phi_t(Q)-\phi^o_t(Q) = \int_0^t \mathcal{E}_{s,t}(M(Q))R_1\mathcal{E}_{s,t}(M(Q))'ds \geqslant 0$$

with the flow of matrices

$$u\mapsto M_u(Q) := A_u - \frac{\phi^o_u(Q)+\phi_u(Q)}{2}S_u.$$

We have the decomposition

$$\begin{aligned}\overline{\phi}_t(Q) &:= \mathcal{E}_t(A)^{-1}\phi_t(Q)\mathcal{E}_t'(A)^{-1} \\ &= Q + \int_0^t \mathcal{E}_s(A)^{-1}R_1\mathcal{E}_s'(A)^{-1}ds \\ &\quad - \int_0^t \mathcal{E}_s(A)^{-1}\phi_s(Q)\mathcal{E}_s'(A)^{-1}\partial_s\overline{\mathcal{O}}_s\mathcal{E}_s(A)^{-1}\phi_s(Q)\mathcal{E}_s'(A)^{-1}ds.\end{aligned}$$

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 08/19/18 to 138.25.168.33. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

In differential form this equation resumes to

$$\partial_t\overline{\phi}_t(Q) = \overline{R}_t - \overline{\phi}_t(Q)\left[\partial_t\overline{\mathcal{O}}_t\right]\overline{\phi}_t(Q) \quad \text{with} \quad \overline{R}_t = \mathcal{E}_t(A)^{-1}R_1\mathcal{E}'_t(A)^{-1}.$$

On the other hand, we have

$$\begin{aligned}\partial_t\left\{\overline{\mathcal{O}}_t\overline{\phi}_t(Q)\overline{\mathcal{O}}_t\right\} &= \left[\partial_t\overline{\mathcal{O}}_t\right]\overline{\phi}_t(Q)\overline{\mathcal{O}}_t + \overline{\mathcal{O}}_t\overline{\phi}_t(Q)\left[\partial_t\overline{\mathcal{O}}_t\right] \\ &\quad + \overline{\mathcal{O}}_t\left\{\overline{R}_t - \overline{\phi}_t(Q)\left[\partial_t\overline{\mathcal{O}}_t\right]\overline{\phi}_t(Q)\right\}\overline{\mathcal{O}}_t \\ &= \overline{\mathcal{O}}_t\overline{R}_t\overline{\mathcal{O}}_t + \partial_t\overline{\mathcal{O}}_t - \left[Id - \overline{\mathcal{O}}_t\overline{\phi}_t(Q)\right]\partial_t\overline{\mathcal{O}}_t\left[Id - \overline{\mathcal{O}}_t\overline{\phi}_t(Q)\right]' \\ &\leqslant \partial_t\overline{\mathcal{O}}_t + \overline{\mathcal{O}}_t\overline{R}_t\overline{\mathcal{O}}_t := \partial_t\left[\overline{\mathcal{O}}_t + \overline{\mathcal{R}}_t\right]\end{aligned}$$

with

$$\overline{\mathcal{R}}_t = \int_0^t \overline{\mathcal{O}}_s\overline{R}_s\overline{\mathcal{O}}_s ds = \int_0^t \mathcal{E}_s(A)'\mathcal{O}_sR_1\mathcal{O}_s\mathcal{E}_s(A)ds.$$

This implies that

$$\begin{aligned}\mathcal{E}_t(A)'\mathcal{O}_t\phi_t(Q)\mathcal{O}_t\mathcal{E}_t(A) &= \overline{\mathcal{O}}_t\overline{\phi}_t(Q)\overline{\mathcal{O}}_t \leqslant \overline{\mathcal{O}}_t + \overline{\mathcal{R}}_t \\ &= \overline{\mathcal{O}}_t + \int_0^t \overline{\mathcal{O}}_s\overline{R}_s\overline{\mathcal{O}}_s ds \\ &= \mathcal{E}_t(A)'\mathcal{O}_t\mathcal{E}_t(A) + \int_0^t \mathcal{E}_s(A)'\mathcal{O}_sR_1\mathcal{O}_s\mathcal{E}_s(A)ds,\end{aligned}$$

from which we conclude that

$$\phi_t(Q) \leqslant \mathcal{O}_t^{-1} + \mathcal{C}_t(\mathcal{O}).$$

Since $\mathcal{C}_t(\mathcal{O})$ and $\mathcal{O}_t$ do not depend on the initial state Q, for any $t \geqslant \upsilon$ we have

$$\phi_t(Q) = \phi_\upsilon(\phi_{t-\upsilon}(Q)) \leqslant \mathcal{O}_\upsilon^{-1} + \mathcal{C}_\upsilon(\mathcal{O}).$$

The inverse of both sides exists due to our observability/controllability assumption. This ends the proof of the lemma. ∎

Whenever it exists, the inverse $\phi_t(Q)^{-1}$ of the positive-definite symmetric matrices $\phi_t(Q) > 0$ satisfies the eigenvalue relationships

$$\phi_t(Q)^{-1} \geqslant \left(\mathcal{O}_\upsilon^{-1} + \mathcal{C}_\upsilon(\mathcal{O})\right)^{-1} \quad \Longrightarrow \inf_{Q\in\mathbb{S}^+_{r_1}} \lambda_{min}\left(\phi_t(Q)^{-1}\right) \geqslant \left(\varpi^c_+(\mathcal{O}) + 1/\varpi^o_-\right)^{-1} > 0$$

and

$$\sup_{Q\in\mathbb{S}^+_{r_1}} \varsigma(\phi_t(Q)) \leqslant \varsigma\left(\mathcal{O}_\upsilon^{-1} + \mathcal{C}_\upsilon(\mathcal{O})\right) \leqslant \varsigma\left(\mathcal{O}_\upsilon^{-1}\right) + \varsigma(\mathcal{C}_\upsilon(\mathcal{O})) \leqslant \varpi^c_+(\mathcal{O}) + (1/\varpi^o_-).$$

We also have

$$\begin{aligned}\inf_{Q\in\mathbb{S}^+_{r_1}} \lambda_{min}\left(\phi_t(Q)^{-1}\right) = \frac{1}{\sup_{Q\in\mathbb{S}^+_{r_1}} \lambda_{max}\left(\phi_t(Q)\right)} &\Rightarrow \sup_{Q\in\mathbb{S}^+_{r_1}} \lambda_{max}\left(\phi_t(Q)\right) \\ &\leqslant \varpi^c_+(\mathcal{O}) + 1/\varpi^o_- < \infty.\end{aligned}$$

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 08/19/18 to 138.25.168.33. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

The inverse matrices $\phi_t(Q)^{-1}$ satisfy the equation

$$\partial_t \phi_t(Q)^{-1} = \mathrm{Ricc}_-\big(\phi_t(Q)^{-1}\big)$$

for any $t \geqslant \upsilon$, with the drift function

$$\mathrm{Ricc}_-(Q) = -A_t' Q - QA_t - QR_1Q + S_t.$$

We denote by $\phi_t^{-o}(Q^{-1})$ the flow starting at Q^{-1} associated with the drift function

$$\mathrm{Ricc}_{-o}(Q) := -QA_t - A_t'Q - QR_1Q = \mathrm{Ricc}_-(Q) - S_t.$$

The next lemma concerns the uniform boundedness of the inverse $\phi_t(Q)^{-1}$ of the positive-definite symmetric matrices $\phi_t(Q) > 0$ w.r.t. the time horizon $t \geqslant \upsilon$.

LEMMA 4.7. *For any $t \geqslant \upsilon$ we have*

$$\mathcal{C}_t \leqslant \big[\phi_t^{-o}\big(Q^{-1}\big)\big]^{-1} = \mathcal{E}_t(A)Q^{-1}\mathcal{E}_t(A)' + \mathcal{C}_t = \phi_t^c\big(Q^{-1}\big).$$

In addition, we have

$$\phi_t^{-o}\big(Q^{-1}\big) \leqslant \phi_t(Q)^{-1} \leqslant \mathcal{O}_t(\mathcal{C}) + \mathcal{C}_t^{-1}$$

as well as

$$\sup_{t\geqslant\upsilon} \phi_t(Q)^{-1} \leqslant \mathcal{O}_\upsilon(\mathcal{C}) + \mathcal{C}_\upsilon^{-1} \quad \textit{and} \quad 0 < \big(\mathcal{O}_\upsilon(\mathcal{C}) + \mathcal{C}_\upsilon^{-1}\big)^{-1} \leqslant \inf_{t\geqslant\upsilon} \phi_t(Q).$$

Proof. For brevity we write $\mathcal{E}_{s,t} := \mathcal{E}_{s,t}(A)$. Arguing as in the proof of the preceding lemma, the flow $\phi_t^{-o}(Q^{-1})$ associated with Ricc_{-o} is given by

$$\begin{aligned}
\phi_t^{-o}\big(Q^{-1}\big) &= (\mathcal{E}_t')^{-1}\left(Q^{-1} + \int_0^t \mathcal{E}_s^{-1} R_1 (\mathcal{E}_s')^{-1} ds\right)^{-1} \mathcal{E}_t^{-1} \\
&= \left(\mathcal{E}_t Q^{-1}\mathcal{E}_t' + \int_0^t \mathcal{E}_{s,t} R_1 \mathcal{E}_{s,t}' ds\right)^{-1} \\
&= \big(\mathcal{E}_t Q^{-1}\mathcal{E}_t' + \mathcal{C}_t\big)^{-1} = \phi_t^c\big(Q^{-1}\big)^{-1} \leqslant \mathcal{C}_t^{-1}.
\end{aligned}$$

This can be checked via differentiation as in the preceding proof. In addition, arguing as before,

$$\phi_t(Q)^{-1} \geqslant \phi_t^{-o}\big(Q^{-1}\big).$$

Also observe that the Riccati flow $\widehat{\phi}_t(Q) := \phi_t(Q)^{-1}$ satisfies a Riccati equation defined similarly to that of $\phi_t(Q)$ but with a replacement on the matrices (A_t, R_1, S_t) given by $(\widehat{A}_t, \widehat{R}_t, \widehat{S})$ with

$$\widehat{A}_t := -A_t', \qquad \widehat{R}_t = S_t, \quad \text{and} \quad \widehat{S} = R_1.$$

That is, we have that

$$\widehat{\phi}_t(Q) = \widehat{A}_t\widehat{\phi}_t(Q) + \widehat{\phi}_t(Q)\widehat{A}_t' + \widehat{R}_t - \widehat{\phi}_t(Q)\widehat{S}\widehat{\phi}_t(Q)$$

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

with the initial condition $\widehat{\phi}_0(Q) = Q^{-1}$. This follows the inverse derivation formula also used in the preceding proof. Now it follows from the preceding lemma that

$$\widehat{\phi}_t(Q) \leqslant \widehat{\mathcal{O}}_t^{-1} + \widehat{\mathcal{C}}_t(\widehat{\mathcal{O}})$$

with

$$\widehat{\mathcal{O}}_t = \int_0^t \mathcal{E}'_s\big(\widehat{A}\big)^{-1} R_1 \mathcal{E}_s\big(\widehat{A}\big)^{-1} ds = \int_0^t \mathcal{E}_s R_1 \mathcal{E}'_s ds = \mathcal{C}_t$$

and

$$\begin{aligned}\widehat{\mathcal{C}}_t(\widehat{\mathcal{O}}) &= \widehat{\mathcal{O}}_t^{-1}\left[\int_0^t \mathcal{E}'_{s,t}\big(\widehat{A}\big)^{-1}\widehat{\mathcal{O}}_s \widehat{R}_t \widehat{\mathcal{O}}_s \mathcal{E}_{s,t}\big(\widehat{A}\big)^{-1} ds\right]\widehat{\mathcal{O}}_t^{-1} \\ &= \mathcal{C}_t^{-1}\left[\int_0^t \mathcal{E}_{s,t}\mathcal{C}_s S_t \mathcal{C}_s \mathcal{E}'_{s,t} ds\right]\mathcal{C}_t^{-1} = \mathcal{O}_t(\mathcal{C}).\end{aligned}$$

We conclude that

$$\phi_t^{-o}\big(Q^{-1}\big) \leqslant \phi_t(Q)^{-1} \leqslant \mathcal{C}_t^{-1} + \mathcal{O}_t(\mathcal{C}).$$

We also have

$$\sup_{t\geqslant 0} \phi_t(Q)^{-1} = \sup_{t\geqslant v} \phi_v\left(\phi_{t-v}(Q)\right)^{-1} \leqslant \mathcal{O}_v(\mathcal{C}) + \mathcal{C}_v^{-1},$$

and therefore

$$\inf_{t\geqslant v} \phi_t(Q) \geqslant \big(\mathcal{O}_v(\mathcal{C}) + \mathcal{C}_v^{-1}\big)^{-1}.$$

This ends the proof of the lemma. ∎

Combining this pair of lemmas, we readily prove Theorem 4.4.

4.2. Bucy's convergence theorem for Kalman–Bucy semigroups. We now prove the exponential convergence of a time-varying semigroup generated by a time-varying matrix difference of the form $A_t - \phi_t(Q)S_t$. This significantly generalizes (13) of Lemma 3.2.

THEOREM 4.8 (Bucy [4]). *For any $t \geqslant s \geqslant v$ we have*

$$\sup_{Q\in\mathbb{S}^+_{r_1}} \|E_{s,t}(Q)\|_2 \leqslant \alpha \exp\{-\beta(t-s)\}$$

with the parameters

$$\begin{aligned}\alpha^2 &:= \frac{\varpi_+^o(\mathcal{C}) + 1/\varpi_-^c}{\varpi_+^c(\mathcal{O}) + 1/\varpi_-^o} \quad \textit{and} \\ 2\beta &:= \frac{1}{(\varpi_+^o(\mathcal{C}) + 1/\varpi_-^c)}\left[\inf_{t\geqslant 0}\lambda_{min}(S_t) + \frac{\lambda_{min}(R_1)}{(\varpi_+^c(\mathcal{O}) + 1/\varpi_-^o)^2}\right].\end{aligned}$$

Proof. Observe that

$$\begin{aligned}&\phi_t(Q)^{-1}(A_t - \phi_t(Q)S_t) + (A_t - \phi_t(Q)S_t)'\phi_t(Q)^{-1} + \mathrm{Ricc}_-\left(\phi_t(Q)^{-1}\right) \\ &\quad = -\left[S_t + \phi_t(Q)^{-1}R_1\phi_t(Q)^{-1}\right].\end{aligned}$$

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 08/19/18 to 138.25.168.33. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

Downloaded 08/19/18 to 138.25.168.33. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

This implies that

$$\begin{aligned}
&\partial_t\left(E_{s,t}(Q)'\phi_t(Q)^{-1}E_{s,t}(Q)\right)\\
&\quad= E_{s,t}(Q)'\left\{(A_t-\phi_t(Q)S_t)'\phi_t(Q)^{-1}+\phi_t(Q)^{-1}(A_t-\phi_{t-s}(Q)S_t)\right.\\
&\qquad\left.+\ \mathrm{Ricc}_-\left(\phi_t(Q)^{-1}\right)\right\}E_{s,t}(Q),
\end{aligned}$$

from which we conclude that

$$\partial_t\left(E_{s,t}(Q)'\phi_t(Q)^{-1}E_{s,t}(Q)\right)=-E_{s,t}(Q)'\left[S_t+\phi_t(Q)^{-1}R_1\phi_t(Q)^{-1}\right]E_{s,t}(Q).$$

By Theorem 4.4, we also have

$$\begin{aligned}
S_t+\phi_t(Q)^{-1}R_1\phi_t(Q)^{-1}&\geqslant\left[\inf_{t\geqslant0}\lambda_{min}(S_t)+\frac{\lambda_{min}(R_1)}{\lambda^2_{max}(\phi_t(Q))}\right]Id\\
&\geqslant\left[\inf_{t\geqslant0}\lambda_{min}(S_t)+\frac{\lambda_{min}(R_1)}{(\varpi^c_+(\mathcal{O})+1/\varpi^o_-)^2}\right]Id
\end{aligned}$$

and

$$\begin{aligned}
&(\varpi^o_+(\mathcal{C})+1/\varpi^c_-)^{-1}Id\leqslant\phi_t(Q)\leqslant(\varpi^c_+(\mathcal{O})+1/\varpi^o_-)Id\\
&\quad\Longleftrightarrow(\varpi^o_+(\mathcal{C})+1/\varpi^c_-)Id\geqslant\phi_t(Q)^{-1}\geqslant(\varpi^c_+(\mathcal{O})+1/\varpi^o_-)^{-1}Id.
\end{aligned}$$

This implies that

$$S_t+\phi_t(Q)^{-1}R_1\phi_t(Q)^{-1}\geqslant\beta\phi_t(Q)^{-1}$$

with

$$\beta:=\frac{1}{(\varpi^o_+(\mathcal{C})+1/\varpi^c_-)}\left[\inf_{t\geqslant0}\lambda_{min}(S_t)+\frac{\lambda_{min}(R_1)}{(\varpi^c_+(\mathcal{O})+1/\varpi^o_-)^2}\right],$$

from which we conclude that

$$E_{s,t}(Q)'\phi_t(Q)^{-1}E_{s,t}(Q)\geqslant(\varpi^c_+(\mathcal{O})+1/\varpi^o_-)E_{s,t}(Q)'E_{s,t}(Q).$$

This implies that

$$\partial_t\langle E_{s,t}(Q)x,\phi_t(Q)^{-1}E_{s,t}(Q)x\rangle\leqslant-\beta\langle E_{s,t}(Q)x,\phi_t(Q)^{-1}E_{s,t}(Q)x\rangle.$$

By Grönwall's inequality we prove that

$$\begin{aligned}
(\varpi^c_+(\mathcal{O})+1/\varpi^o_-)\langle E_{s,t}(Q)x,E_{s,t}(Q)x\rangle&\leqslant\left\langle E_{s,t}(Q)x,\phi_t(Q)^{-1}E_{s,t}(Q)x\right\rangle\\
&\leqslant e^{-\beta(t-s)}\left\langle x,\phi_s(Q)^{-1}x\right\rangle\\
&\leqslant(\varpi^o_+(\mathcal{C})+1/\varpi^c_-)e^{-\beta(t-s)}\langle x,x\rangle,
\end{aligned}$$

from which we conclude that

$$\|E_{s,t}(Q)\|_2^2\leqslant\frac{\varpi^o_+(\mathcal{C})+1/\varpi^c_-}{\varpi^c_+(\mathcal{O})+1/\varpi^o_-}e^{-\beta(t-s)}.$$

This ends the proof of the theorem. ∎

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

We also have the following corollary.

Corollary 4.9. *For any* $0 \leqslant s \leqslant t$ *and any* $Q \in \mathbb{S}^+_{r_1}$ *we have*

$$\|E_{s,t}(Q)\|_2 \leqslant \rho(Q)\, \exp\{-\beta(t-s)\} \tag{26}$$

with the function

$$Q \mapsto \rho(Q) := (\alpha \vee 1) \exp\left[\left(\beta + \sup_{t\geqslant 0}\|A_t\|_2 + \|\phi(Q)\|_2 \sup_{t\geqslant 0}\|S_t\|_2\right)\upsilon\right]$$

and the uniform norm $\|\phi(Q)\|_2$ *introduced in* (23).

Proof. The estimate is immediate when $\upsilon \leqslant s \leqslant t$. By Lemma 3.1, for any $0 \leqslant s \leqslant t \leqslant \upsilon$ we have

$$\begin{aligned}\|E_{s,t}(Q)\|_2 &\leqslant \exp\{-\beta(t-s)\}\\ &\quad\times \exp\left[\beta(t-s) + \left(\sup_{t\geqslant 0}\|A_t\|_2 + \|\phi(Q)\|_2 \sup_{t\geqslant 0}\|S_t\|_2\right)\upsilon\right] \Longrightarrow (26).\end{aligned}$$

In the same vein, when $0 \leqslant s \leqslant \upsilon \leqslant t$, we use Theorem 4.8 to check that

$$\begin{aligned}E_{s,t}(Q) = E_{\upsilon,t}(Q)E_{s,\upsilon}(Q) \Longrightarrow \|E_{s,t}(Q)\|_2 &\leqslant \alpha \exp\{-\beta(t-\upsilon)\}\\ &\quad\times \exp\left[\left(\sup_{t\geqslant 0}\|A_t\|_2 + \|\phi(Q)\|_2 \sup_{t\geqslant 0}\|S_t\|_2\right)(\upsilon - s)\right].\end{aligned}$$

This implies that

$$\|E_{s,t}(Q)\|_2 \leqslant \alpha \exp\{-\beta(t-s)\} \exp\left[\left(\beta + \sup_{t\geqslant 0}\|A_t\|_2 + \|\phi(Q)\|_2 \sup_{t\geqslant 0}\|S_t\|_2\right)(\upsilon - s)\right].$$

This ends the proof of the corollary. ∎

Using (18), we readily check the following contraction estimate.

Corollary 4.10. *For any* $t \geqslant 0$ *and any* $Q_1, Q_2 \in \mathbb{S}^+_{r_1}$ *we have*

$$\|\phi_t(Q_1) - \phi_t(Q_2)\|_2 \leqslant \rho(Q_1, Q_2)\, \exp\{-2\beta t\} \|Q_1 - Q_2\|_2$$

with

$$\rho(Q_1, Q_2) := \rho(Q_1)\rho(Q_2),$$

with $Q \mapsto \rho(Q)$ *defined in Corollary* 4.9.

4.3. Quantitative contraction estimates for time-invariant signal models. We now consider time-invariant signal models and note that satisfaction of the observability and controllability rank conditions (10) is sufficient to ensure the existence of a (unique) positive-definite solution P of (11). We thus assume that the time-invariant matrix $A - PS$ satisfies (13) of Lemma 3.2 for some $\nu > 0$ and some $\kappa < \infty$.

Bucy's theorem, discussed in section 4.2, yields more or less directly several contraction inequalities. Note that because of (24) it follows that

$$\begin{aligned}Q \mapsto \rho(Q) &:= (\alpha \vee 1)\, \exp\left[(\beta + \|A\|_2 + \|\phi(Q)\|_2\|S\|_2)\,\upsilon\right]\\ &\leqslant (\alpha \vee 1)\, \exp\left[\left(\beta + \|A\|_2 + \left(\|P\|_2 + \kappa^2\|Q - P\|_2\right)\|S\|_2\right)\upsilon\right]\end{aligned} \tag{27}$$

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 08/19/18 to 138.25.168.33. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

Downloaded 08/19/18 to 138.25.168.33. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

and thus Corollaries 4.9 and 4.10 immediately deliver crude (in terms of the rate) quantitative contraction results in the time-invariant setting.

The first result of this section concerns the exponential rate of the convergence of the Riccati flow towards the steady state.

COROLLARY 4.11. *For any* $t \geqslant 0$ *and any* $Q \in \mathbb{S}_{r_1}^+$ *we have*

$$\|\phi_t(Q) - P\|_2 \leqslant \kappa_\phi(Q) e^{-2\nu t} \|Q - P\|_2$$

with the parameters

$$\kappa_\phi(Q) := \kappa^2 \exp\left\{(2\beta)^{-1} \|S\|_2 \kappa^2 \rho(P,Q) \|Q - P\|_2\right\}.$$

In the above, $\rho(P,Q)$ *is defined in Corollary* 4.9, *noting the upper bound* (27).

Proof. Combining (20) and (13) with Corollary 4.9, we prove that

$$\begin{aligned} Z_t &:= e^{-2\nu t} \|\phi_t(Q) - P\|_2 \\ &\leqslant \kappa^2 Z_0 + \kappa^2 \|S\|_2 \rho(P,Q) \|Q - P\|_2 \int_0^t \exp\{-2\beta u\} Z_u du \end{aligned}$$

for any $\epsilon \in]0,1]$ and any $t \geqslant 0$. A direct application of Grönwall's inequality yields

$$Z_t \leqslant \kappa^2 Z_0 \exp\left\{(2\beta)^{-1} \kappa^2 \|S\|_2 \rho(P,Q) \|Q - P\|_2 \int_0^t 2\beta \exp\{-2\beta u\} du\right\}.$$

This ends the proof of the corollary. ∎

We also have the following quantitative exponential convergence result on the Kalman–Bucy semigroup and a contraction-type inequality on the Riccati flow.

COROLLARY 4.12. *For any* $0 \leqslant s \leqslant t$, $\epsilon \in]0,1]$, *and* $Q \in \mathbb{S}_{r_1}^+$ *we have*

$$\|E_{s,t}(Q)\|_2 \leqslant \kappa_E(Q) \exp[-\nu(t-s)] \Longrightarrow \|E(Q)\|_2 := \sup_{0 \leqslant s \leqslant t} \|E_{s,t}(Q)\|_2 \leqslant \kappa_E(Q)$$

with the parameters

$$\kappa_E(Q) := \kappa \exp\left(\frac{\kappa}{2\nu} \kappa_\phi(Q) \|S\|_2 \|Q - P\|_2\right),$$

where $\kappa_\phi(Q)$ *is the parameter defined in Corollary* 4.11.

In addition, for any $Q_1, Q_2 \in \mathbb{S}_{r_1}^+$ *we have the local Lipschitz property*

$$\|\phi_t(Q_1) - \phi_t(Q_2)\|_2 \leqslant \kappa_\phi(Q_1, Q_2) \exp[-2\nu t] \|Q_1 - Q_2\|_2 \tag{28}$$

with

$$\kappa_\phi(Q_1, Q_2) := \kappa_E(Q_1) \kappa_E(Q_2).$$

Proof. For any $0 \leqslant s \leqslant t$ we have

$$E_{s,t}(Q) = \exp\left[\oint_s^t [(A - PS) + (P - \phi_u(Q))S]\, du\right].$$

By Corollary 4.11 we have

$$\int_s^t \|\phi_u(Q) - P)S\|_2 du \leqslant \kappa_\phi(Q) \|S\|_2 \|Q - P\|_2 \int_s^t e^{-2\nu u} du \leqslant \frac{\kappa_\phi(Q) \|S\|_2}{2\nu} \|Q - P\|_2.$$

The first estimate is a direct consequence of Lemma 3.1. Estimate (28) is a direct consequence of (18). The end of the proof is now a direct consequence of Lemma 3.1. ∎

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 08/19/18 to 138.25.168.33. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

The preceding result can be contrasted with Corollaries 4.9 and 4.10 that concern the more general time-varying Kalman–Bucy semigroup and Riccati flow (noting the upper bound given by (27)). In the time-invariant domain one may be able to improve the exponential rate significantly, e.g., via $\nu > 0$ in (13) of Lemma 3.2, as compared to Bucy's general theorem (Theorem 4.8). For example, ν may (optimally) be given by the negative log-norm or spectral abscissa of $A - PS$. However, this improvement in the rate may come at the expense of a (much) larger constant.

Tracking constants carefully is important when applying these results in practice, e.g., when analyzing the stability of ensemble Kalman filters [42, 17] or extended Kalman filters [39, 16, 15]. These constants are also typically related to the underlying state-space dimension. In this sense, one should carefully follow the form and the source of these terms, in order to understand accurately the dimensional error dependence of any approximation (in, e.g., high dimensions) [13, 37, 32].

Finally, we have a quantitative contraction inequality on the Kalman–Bucy semigroup with time-invariant signal models.

Corollary 4.13. *For any $0 \leqslant s \leqslant t$, $\epsilon \in]0,1]$, and $Q_1, Q_2 \in \mathbb{S}_{r_1}^+$ we have*

$$\|E_{s,t}(Q_2) - E_{s,t}(Q_1)\|_2 \leqslant \kappa_E(Q_1,Q_2)\exp\left[-\nu(t-s)\right]\|Q_2 - Q_1\|_2$$

with

$$\kappa_E(Q_1,Q_2) := \kappa_E(Q_2) + \kappa_E^2(Q_2)\kappa_\phi(Q_1,Q_2)\|S\|_2(2\nu)^{-1}.$$

Proof. We have

$$\begin{aligned}\partial_t\left(E_{s,t}(Q_2) - E_{s,t}(Q_1)\right) &= (A - \phi_t(Q_2)S)E_{s,t}(Q_2) - (A - \phi_t(Q_1)S)E_{s,t}(Q_1)\\ &= (A - \phi_t(Q_2)S)(E_{s,t}(Q_2) - E_{s,t}(Q_1))\\ &\quad + (\phi_t(Q_1) - \phi_t(Q_2))SE_{s,t}(Q_1).\end{aligned}$$

This implies that

$$\begin{aligned}E_{s,t}(Q_2) - E_{s,t}(Q_1) &= E_{s,t}(Q_2)(Q_2 - Q_1)\\ &\quad + \int_s^t E_{u,t}(Q_2)(\phi_u(Q_1) - \phi_u(Q_2))SE_{s,u}(Q_1)du.\end{aligned}$$

By Corollary 4.12, this yields the estimate

$$\begin{aligned}&\exp\left[\nu(t-s)\right]\|E_t(Q_2) - E_t(Q_1)\|_2\\ &\quad\leqslant \kappa_E(Q_2)\|Q_2 - Q_1\|_2 + \kappa_E^2(Q_2)\|S\|_2\int_s^t \|\phi_u(Q_1) - \phi_u(Q_2)\|_2 du.\end{aligned}$$

Using (28), we check that

$$\begin{aligned}&\exp\left[\nu(t-s)\right]\|E_{s,t}(Q_2) - E_{s,t}(Q_1)\|_2\\ &\quad\leqslant \|Q_2 - Q_1\|_2\left[\kappa_E(Q_2) + \kappa_E^2(Q_2)\kappa_\phi(Q_1,Q_2)\|S\|_2(2\nu)^{-1}\right].\end{aligned}$$

This ends the proof of the corollary. ∎

5. Contraction of Kalman–Bucy-type stochastic flows. We first review a straightforward qualitative stability result for the time-varying Kalman–Bucy filter that follows from the uniform boundedness of the Riccati flow.

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 08/19/18 to 138.25.168.33. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

THEOREM 5.1. *For any $t \geqslant s \geqslant \upsilon$ we have the uniform estimate*

$$\sup_{Q\in\mathbb{S}^+_{r_1}} \| \mathbb{E}\left(\psi_{s,t}(x,Q) - \varphi_{s,t}(X_s) \,|\, X_s\right) \|_2 \leqslant \alpha \exp\{-\beta(t-s)\} \| x - X_s\|_2$$

with the parameters $\alpha, \beta > 0$ defined as in Theorem 4.8.

Proof. The proof (and result) follow those of Theorem 4.8. By uniform observability/controllability we suppose that $Q = \phi_{s-\upsilon,s}(Q_0)$ without loss of generality. Let $\xi_{s,t} := \mathbb{E}\left(\psi_{s,t}(x,Q) - \varphi_{s,t}(X_s) \,|\, X_s\right)$. Consider the functional

$$\frac{\|\xi_{s,t}\|_2^2}{\left(\varpi^c_+(\mathcal{O}) + 1/\varpi^o_-\right)} \leqslant \xi'_{s,t}\, \phi_{s,t}(Q)^{-1}\xi_{s,t} \leqslant \left(\varpi^o_+(\mathcal{C}) + 1/\varpi^c_-\right) \|\xi_{s,t}\|_2^2 .$$

Then

$$\begin{aligned} \partial_t \xi_{s,t}{}'\phi_{s,t}(Q)^{-1}\xi_{s,t} &= -\, \xi'_{s,t}\, \left(\phi_{s,t}(Q)^{-1}R_1\phi_{s,t}(Q)^{-1} + S_t\right) \xi_{s,t} \\ &\leqslant -\,\beta\, \xi_{s,t}{}'\phi_{s,t}(Q)^{-1}\xi_{s,t}. \end{aligned}$$

By Grönwall's inequality we find that

$$\begin{aligned} (\varpi^c_+(\mathcal{O}) + 1/\varpi^o_-)\xi_{s,t}{}'\xi_{s,t} &\leqslant \xi_{s,t}{}'\phi_{s,t}(Q)^{-1}\xi_{s,t} \\ &\leqslant e^{-\beta(t-s)}\xi_{s,s}{}'\phi_{s,s}(Q)^{-1}\xi_{s,s} \\ &\leqslant (\varpi^o_+(\mathcal{C}) + 1/\varpi^c_-)e^{-\beta(t-s)}\xi_{s,s}{}'\xi_{s,s} \end{aligned}$$

and the result follows with $\alpha, \beta > 0$ defined as in Theorem 4.8. ∎

Given this classical, qualitative stability result, we now study in more precise terms the convergence of Kalman–Bucy stochastic flows, both in the classical filtering form and in the novel nonlinear diffusion form. We study exponential inequalities that bound, with dedicated probability, the stochastic flow of the sample paths at any time, with respect to the underlying signal. We also provide almost sure contraction-type estimates. Both types of result offer a notion of filter stability and the analysis in this section is novel.

We assume that the signal models are time invariant throughout the remainder of this section, and we build on the quantitative estimates of the previous section.

In further development of this section, going forward we consider the function

$$Q \mapsto \sigma(Q) := 2\sqrt{2}\kappa_E(Q)\left[\left(\|\phi(Q)\|_2^2\|S\|_2 + \|R\|_2\right) r_1/\nu\right]^{1/2}$$

as well as

$$\begin{aligned} \chi_0(Q_1,Q_2) &= \nu^{-1}\kappa_E(Q_1)\kappa_\phi(Q_1,Q_2), \\ \chi_1(Q) &= \|S\|_2\kappa_E(Q)/2, \quad \text{and} \\ \chi_2(Q) &= \|S\|_2\sigma(Q) + 2\sqrt{2r_1\|S\|_2\nu} \end{aligned}$$

with the parameters ν, $\|\phi(Q)\|_2$, $\kappa_E(Q)$, and $\kappa_\phi(Q_1,Q_2)$ defined, respectively, in the exponential rate of (13), Proposition 4.3, Corollary 4.12, and Corollary 4.12 again.

5.1. Time-invariant Kalman–Bucy filter.

THEOREM 5.2. *The conditional probability of the events*

$$\|\psi_{s,t}(x,Q) - \varphi_{s,t}(X_s) - E_{t-s}(Q)\left[x - X_s\right]\|_2 \leqslant \frac{e^2}{\sqrt{2}}\left[\frac{1}{2} + \left(\delta + \sqrt{\delta}\right)\right]\sigma^2(Q), \tag{29}$$

given the state variable X_s, is greater than $1 - e^{-\delta}$ for any $\delta \geqslant 0$ and any $t \in [s,\infty[$.

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

By (29), the conditional probability of the event

$$\|\psi_{s,t}(x,Q)-\varphi_{s,t}(X_s)\|_2 \leqslant \kappa_E(Q)e^{-\nu(t-s)}\|x-X_s\|_2 + \frac{e^2}{\sqrt{2}}\left[\frac{1}{2}+\left(\delta+\sqrt{\delta}\right)\right]\sigma^2(Q), \tag{30}$$

given the state variable X_s, is greater than $1-e^{-\delta}$ for any $\delta \geqslant 0$ and any $t \in [s,\infty[$.

The above theorem is a direct consequence of (7) and the following technical lemma.

LEMMA 5.3. *For any $x \in \mathbb{R}^{r_1}$, $Q \in \mathbb{S}^+_{r_1}$, and $n \geqslant 1$ we have the uniform estimate*

$$\sup_{t\geqslant s}\mathbb{E}\left(\left\|\psi_{s,t}(x,Q)-\varphi_{s,t}(X_s)-E_{t|s}(Q)\left[x-X_s\right]\right\|_2^{2n}|X_s\right)^{1/n} \leqslant n\sigma^2(Q).$$

In particular, for any $t \geqslant s \geqslant 0$ we have

$$\mathbb{E}\left(\|\psi_{s,t}(x,Q)-\varphi_{s,t}(X_s)\|_2^{2n}|X_s\right)^{\frac{1}{2n}} \leqslant \sqrt{n}\sigma(Q)+\kappa_E(Q)e^{-\nu(t-s)}\left\|x-X_s\right\|_2.$$

Proof. For any given $s \geqslant 0$ and for any $t \in [s,\infty[$ and any $x \in \mathbb{R}^{r_1}$ we have

$$d\left[\psi_{s,t}(x,Q)-\varphi_{s,t}(X_s)\right] = \left[A-\phi_{s,t}(Q)S\right]\left[\psi_{s,t}(x,Q)-\varphi_{s,t}(X_s)\right]dt + dM_{s,t}$$

with the r_1-multivariate martingale $(M_{s,t})_{t\in[s,\infty[}$ given by

$$t \in [s,\infty[\mapsto M_{s,t} = \int_s^t \phi_{s,u}(Q)C'R_2^{-1}dV_u - R_1^{1/2}(W_t-W_s) \tag{31}$$
$$\Longrightarrow \partial_t\langle M_{s,\cdot}(k), M_{s,\cdot}(l)\rangle_t = \phi_{s,t}(Q)S\phi_{s,t}(Q)+R_1.$$

This yields the formula

$$N_{s,t} := \left[\psi_{s,t}(x,Q)-\varphi_{s,t}(X_s)\right]-E_{t|s}(Q)\left[x-X_s\right] = \int_s^t E_{u-s,t-s}(Q)dM_{s,u}.$$

On the other hand, we have

$$\mathbb{E}\left(\|N_{s,t}\|_2^{2n}\right)^{1/n} = \left[\mathbb{E}\left(\left[\sum_{1\leqslant k\leqslant r_1} N_{s,t}(k)^2\right]^n\right)\right]^{1/n} \leqslant \sum_{1\leqslant k\leqslant r_1}\mathbb{E}\left(N_{s,t}(k)^{2n}\right)^{1/n}$$
$$= \sum_{1\leqslant k\leqslant r_1}\mathbb{E}\left(\left[\sum_{1\leqslant l\leqslant r_1}\int_s^t E_{u-s,t-s}(Q)(k,l)dM_{s,u}(l)\right]^{2n}\right)^{1/n}.$$

By the Burkholder–Davis–Gundy inequality (8), we have

$$\mathbb{E}\left(\left[\sum_{1\leqslant l\leqslant r_1}\int_s^t E_{u,t|s}(Q)(k,l)dM_{s,u}(l)\right]^{2n}\right)^{1/n}$$
$$\leqslant 4^2 n \sum_{1\leqslant l,l'\leqslant r_1}\int_s^t E_{u,t|s}(Q)(k,l)E_{u,t|s}(Q)(k,l')(\phi_{s,u}(Q)S\phi_{s,u}(Q)+R_1)(l,l')du$$
$$\leqslant 4^2 n\int_s^t\left[E_{u,t|s}(Q)(\phi_{s,u}(Q)S\phi_{s,u}(Q)+R_1)E_{u,t|s}(Q)'\right](k,k)du.$$

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 08/19/18 to 138.25.168.33. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

This implies that

$$\mathbb{E}\left(\|N_{s,t}\|_2^{2n}\right)^{1/n} \leqslant 4^2 n \int_s^t \operatorname{tr}\left[(\phi_{s,u}(Q)S\phi_{s,u}(Q) + R_1)E_{u,t|s}(Q)'E_{u,t|s}(Q)\right] du$$
$$\leqslant 4^2 n \int_s^t \|\phi_{s,u}(Q)S\phi_{s,u}(Q) + R_1\|_2 \|E_{u,t|s}(Q)\|_F^2 du.$$

This yields

$$\mathbb{E}\left(\left\|\int_s^t E_{u,t|s}(Q) dM_{s,u}\right\|_2^{2n}\right)^{1/n} \leqslant 4^2 n \left(\|\phi(Q)\|_2^2\|S\|_2 + \|R_1\|_2\right) r_1 \int_s^t \|E_{u-s,t-s}(Q)\|_2^2 du.$$

This implies that

$$\mathbb{E}\left(\left\|\psi_{s,t}(x,Q) - \varphi_{s,t}(X_s) - E_{t|s}(Q)\left[x - X_s\right]\right\|_2^{2n}\right)^{1/n}$$
$$\leqslant 4^2 n \left(\|\phi(Q)\|_2^2\|S\|_2 + \|R_1\|_2\right) r_1 \int_s^t \|E_{u-s,t-s}(Q)\|_2^2 du.$$

Using Proposition 4.3 and Corollary 4.12, we have

$$\mathbb{E}\left(\left\|\psi_{s,t}(x,Q) - \varphi_{s,t}(X_s) - E_{t|s}(Q)\left[x - X_s\right]\right\|_2^{2n}\right)^{1/n}$$
$$\leqslant 8n\nu^{-1}\kappa_E^2(Q)\left(\|\phi(Q)\|_2^2\|S\|_2 + \|R_1\|_2\right) r_1 \int_s^t (2\nu)\exp\left[-2\nu(t-u)\right] du.$$

This ends the proof of the lemma. ◻

THEOREM 5.4. *For any $t \geqslant s \geqslant 0$, $x_1, x_2 \in \mathbb{R}^{r_1}$, $Q_1, Q_2 \in \mathbb{S}_{r_1}^+$, and $n \geqslant 1$ we have the almost sure local contraction estimate*

$$\mathbb{E}\left(\|\psi_{s,t}(x_1,Q_1) - \psi_{s,t}(x_2,Q_2)\|_2^{2n} | X_s\right)^{\frac{1}{2n}} \leqslant \kappa_E(Q_1)e^{-\nu(t-s)}\|x_1 - x_2\|_2$$
$$+ e^{-\nu(t-s)}\chi_0(Q_1,Q_2)\left\{\chi_1(Q_2)\left\|x_2 - X_s\right\|_2 + \sqrt{n}\chi_2(Q_1)\right\}\|Q_1 - Q_2\|_2.$$

Proof. We have

$$\begin{aligned}
& d\left(\psi_{s,t}(x_1,Q_1) - \psi_{s,t}(x_2,Q_2)\right) \\
&\quad = \left\{\left[A - \phi_{s,t}(Q_1)S\right]\psi_{s,t}(x_1,Q_1) - \left[A - \phi_{s,t}(Q_2)S\right]\psi_{s,t}(x_2,Q_2)\right\} dt \\
&\qquad + \left[\phi_{s,t}(Q_1) - \phi_{s,t}(Q_2)\right] C' R_2^{-1} dY_t \\
&\quad = \left[A - \phi_{s,t}(Q_1)S\right]\left(\psi_{s,t}(x_1,Q_1) - \psi_{s,t}(x_2,Q_2)\right) dt \\
&\qquad - \left[\phi_{s,t}(Q_1) - \phi_{s,t}(Q_2)\right] S\psi_{s,t}(x_2,Q_2) dt \\
&\qquad + \left[\phi_{s,t}(Q_1) - \phi_{s,t}(Q_2)\right]\left(S\varphi_{s,t}(X_s) ds + C' R_2^{-1/2} dV_t\right).
\end{aligned}$$

This yields the decomposition

$$\begin{aligned}
d\left[\psi_{s,t}(x_1,Q_1) - \psi_{s,t}(x_2,Q_2)\right] &= \left[A - \phi_{s,t}(Q_1)S\right]\left[\psi_{s,t}(x_1,Q_1) - \psi_{s,t}(x_2,Q_2)\right] dt \\
&\quad + \left[\phi_{s,t}(Q_1) - \phi_{s,t}(Q_2)\right] S\left[\varphi_{s,t}(X_s) - \psi_{s,t}(x_2,Q_2)\right] dt \\
&\quad + \left[\phi_{s,t}(Q_1) - \phi_{s,t}(Q_2)\right] dM_t
\end{aligned}$$

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 08/19/18 to 138.25.168.33. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

Downloaded 08/19/18 to 138.25.168.33. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

with $M_t = C' R_2^{-1/2} V_t$. This then yields the decomposition

$$\begin{aligned}\psi_{s,t}(x_1, Q_1) - \psi_{s,t}(x_2, Q_2) &= E_{t|s}(Q_1)(x_1 - x_2) \\ &+ \int_s^t E_{u,t|s}(Q_1)\left[\phi_{s,u}(Q_1) - \phi_{s,u}(Q_2)\right] S\left[\varphi_{s,u}(X_s) - \psi_{s,u}(x_2, Q_2)\right] du \\ &+ \int_s^t E_{u,t|s}(Q_1)\left[\phi_{s,u}(Q_1) - \phi_{s,u}(Q_2)\right] dM_u.\end{aligned}$$

Arguing as in the proof of Lemma 5.3, we have

$$\begin{aligned}&\mathbb{E}\left(\left\|\int_s^t E_{u,t|s}(Q_1)\left[\phi_{s,u}(Q_1) - \phi_{s,u}(Q_2)\right] dM_u\right\|_2^{2n}\right)^{1/n} \\ &\quad\leqslant 4^2 n r_1 \|S\|_2 \int_s^t \|\phi_{s,u}(Q_1) - \phi_{s,u}(Q_2)\|_2^2 \|E_{u,t|s}(Q)\|_2^2 du.\end{aligned}$$

By Corollary 4.12, we have the estimate

$$\begin{aligned}&\mathbb{E}\left(\left\|\int_s^t E_{u,t|s}(Q_1)\left[\phi_{s,u}(Q_1) - \phi_{s,u}(Q_2)\right] dM_u\right\|_2^{2n}\right)^{1/n} \\ &\quad\leqslant 4^2 n r_1 \|S\|_2 \kappa_E^2(Q_1)\kappa_\phi^2(Q_1, Q_2)\|Q_1 - Q_2\|_2^2 \exp\left[-2\nu(t-s)\right] \int_s^t \exp\left[-2\nu(u-s)\right] du \\ &\quad\leqslant 8 n r_1 (\|S\|_2/\nu)(\kappa_E(Q_1)\kappa_\phi(Q_1, Q_2)\|Q_1 - Q_2\|_2)^2 \exp\left[-2\nu(t-s)\right],\end{aligned}$$

from which we find that

$$\begin{aligned}&\mathbb{E}\left(\left\|\int_s^t E_{u,t|s}(Q_1)\left[\phi_{s,u}(Q_1) - \phi_{s,u}(Q_2)\right] dM_u\right\|_2^{2n}\right)^{\frac{1}{2n}} \\ &\quad\leqslant \sqrt{n} 2 \sqrt{2 r_1 (\|S\|_2/\nu)} \kappa_E(Q_1)\kappa_\phi(Q_1, Q_2)\|Q_1 - Q_2\|_2 \exp\left[-\nu(t-s)\right].\end{aligned}$$

On the other hand, we have the almost sure estimate

$$\begin{aligned}\|\psi_{s,t}(x_1, Q_1) - \psi_{s,t}(x_2, Q_2)\|_2 &\leqslant \|E_{t|s}(Q_1)\|_2 \|x_1 - x_2\|_2 \\ &+ \|S\|_2 \int_s^t \|E_{u,t|s}(Q_1)\|_2 \|\phi_{s,u}(Q_1) - \phi_{s,u}(Q_2)\|_2 \|\varphi_{s,u}(X_s) - \psi_{s,u}(x_2, Q_2)\|_2 du \\ &+ \left\|\int_s^t E_{u,t|s}(Q_1)\left[\phi_{s,u}(Q_1) - \phi_{s,u}(Q_2)\right] dM_u\right\|_2.\end{aligned}$$

Combining Corollary 4.12 and the estimate (28), we prove that

$$\begin{aligned}&\|\psi_{s,t}(x_1, Q_1) - \psi_{s,t}(x_2, Q_2)\|_2 \\ &\quad\leqslant \kappa_E(Q_1) \exp\left[-\nu(t-s)\right]\|x_1 - x_2\|_2 + \|S\|_2 \kappa_E(Q_1)\kappa_\phi(Q_1, Q_2)\|Q_1 - Q_2\|_2 \\ &\qquad\times \int_s^t \exp\left[-\nu(t-u)\right]\exp\left[-2\nu(u-s)\right]\|\varphi_{s,u}(X_s) - \psi_{s,u}(x_2, Q_2)\|_2 du \\ &\qquad+ \left\|\int_s^t E_{u-s,t-s}(Q_1)\left[\phi_{s,u}(Q_1) - \phi_{s,u}(Q_2)\right] dM_u\right\|_2.\end{aligned}$$

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 08/19/18 to 138.25.168.33. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

This implies that

$$\begin{aligned}
&\|\psi_{s,t}(x_1,Q_1)-\psi_{s,t}(x_2,Q_2)\|_2\\
&\quad\leqslant \kappa_E(Q_1)\exp\left[-\nu(t-s)\right]\|x_1-x_2\|_2+\|S\|_2\kappa_E(Q_1)\kappa_\phi(Q_1,Q_2)\|Q_1-Q_2\|_2\\
&\qquad\times\exp\left[-\nu(t-s)\right]\int_s^t\exp\left[-\nu(u-s)\right]\|\varphi_{s,u}(X_s)-\psi_{s,u}(x_2,Q_2)\|_2du\\
&\qquad+\left\|\int_s^t E_{u-s,t-s}(Q_1)\left[\phi_{s,u}(Q_1)-\phi_{s,u}(Q_2)\right]dM_u\right\|_2.
\end{aligned}$$

Using the generalized Minkowski inequality, we check that

$$\begin{aligned}
&\mathbb{E}\left(\|\psi_{s,t}(x_1,Q_1)-\psi_{s,t}(x_2,Q_2)\|_2^{2n}|X_s\right)^{\frac{1}{2n}}\\
&\quad\leqslant \kappa_E(Q_1)\exp\left[-\nu(t-s)\right]\|x_1-x_2\|_2+\|S\|_2\kappa_E(Q_1)\kappa_\phi(Q_1,Q_2)\|Q_1-Q_2\|_2\\
&\quad\times\exp\left[-\nu(t-s)\right]\int_s^t\exp\left[-\nu(u-s)\right]\mathbb{E}\left(\|\varphi_{s,u}(X_s)-\psi_{s,u}(x_2,Q_2)\|_2^{2n}|X_s\right)^{\frac{1}{2n}}du\\
&\quad+\mathbb{E}\left(\left\|\int_s^t E_{u-s,t-s}(Q_1)\left[\phi_{s,u}(Q_1)-\phi_{s,u}(Q_2)\right]dM_u\right\|_2^{2n}\right)^{\frac{1}{2n}}.
\end{aligned}$$

Lemma 5.3, combined with the Burkholder–Davis–Gundy estimates stated above, implies that

$$\begin{aligned}
&\exp\left[\nu(t-s)\right]\mathbb{E}\left(\|\psi_{s,t}(x_1,Q_1)-\psi_{s,t}(x_2,Q_2)\|_2^{2n}|X_s\right)^{\frac{1}{2n}}\\
&\quad\leqslant \kappa_E(Q_1)\|x_1-x_2\|_2+\|S\|_2\kappa_E(Q_1)\kappa_\phi(Q_1,Q_2)\|Q_1-Q_2\|_2\\
&\qquad\times\int_s^t\exp\left[-\nu(u-s)\right]\left\{\sqrt{n}\sigma(Q_2)+\kappa_E(Q_2)e^{-\nu(u-s)}\left\|x_2-X_s\right\|_2\right\}du\\
&\qquad+\sqrt{n}2\sqrt{2r_1(\|S\|_2/\nu)}\kappa_E(Q_1)\kappa_\phi(Q_1,Q_2)\|Q_1-Q_2\|_2.
\end{aligned}$$

Observe that

$$\begin{aligned}
&\int_s^t\exp\left[-\nu(u-s)\right]\left\{\sqrt{n}\sigma(Q_2)+\kappa_E(Q_2)e^{-\nu(u-s)}\left\|x_2-X_s\right\|_2\right\}du\\
&\quad=\sqrt{n}(\sigma(Q_2)/\nu)+(\kappa_E(Q_2)/(2\nu))\left\|x_2-X_s\right\|_2.
\end{aligned}$$

This yields

$$\begin{aligned}
&\exp\left[\nu(t-s)\right]\mathbb{E}\left(\|\psi_{s,t}(x_1,Q_1)-\psi_{s,t}(x_2,Q_2)\|_2^{2n}|X_s\right)^{\frac{1}{2n}}\\
&\quad\leqslant \kappa_E(Q_1)\|x_1-x_2\|_2+\|S\|_2\kappa_E(Q_1)\kappa_\phi(Q_1,Q_2)\|Q_1-Q_2\|_2\\
&\qquad\times\left[\sqrt{n}(\sigma(Q_2)/\nu)+(\kappa_E(Q_2)/(2\nu))\left\|x_2-X_s\right\|_2\right]\\
&\qquad+\sqrt{n}2\sqrt{2r_1(\|S\|_2/\nu)}\kappa_E(Q_1)\kappa_\phi(Q_1,Q_2)\|Q_1-Q_2\|_2,
\end{aligned}$$

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 08/19/18 to 138.25.168.33. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

from which we conclude that

$$\begin{aligned}&\exp\left[\nu(t-s)\right]\mathbb{E}\left(\|\psi_{s,t}(x_1,Q_1)-\psi_{s,t}(x_2,Q_2)\|_2^{2n}|X_s\right)^{\frac{1}{2n}}\leqslant\kappa_E(Q_1)\|x_1-x_2\|_2\\&\quad+\left\{\left[\|S\|_2\kappa_E(Q_2)/2\right]\|x_2-X_s\|_2+\sqrt{n}\left[\|S\|_2\sigma(Q_2)+2\sqrt{2r_1\|S\|_2\nu}\right]\right\}\\&\quad\times\nu^{-1}\kappa_E(Q_1)\kappa_\phi(Q_1,Q_2)\|Q_1-Q_2\|_2.\end{aligned}$$

This ends the proof of the theorem. ∎

5.2. Nonlinear time-invariant Kalman–Bucy diffusions.

THEOREM 5.5. *For any $t\geqslant s\geqslant 0$, $x\in\mathbb{R}^{r_1}$, $Q\in\mathbb{S}^+_{r_1}$, and $n\geqslant 1$ we have*

$$\mathbb{E}\left(\left\|\overline{\psi}_{s,t}(x,Q)-\varphi_{s,t}(X_s)\right\|_2^{2n}|X_s\right)^{\frac{1}{2n}}\leqslant\sqrt{2n}\sigma(Q)+\kappa_E(Q)e^{-\nu(t-s)}\left\|x-X_s\right\|_2.$$

The conditional probability of the events

$$\left\|\overline{\psi}_{s,t}(x,Q)-\varphi_{s,t}(X_s)-E_{t-s}(Q)\left[x-X_s\right]\right\|_2\leqslant\sqrt{2}e^2\left[\frac{1}{2}+\left(\delta+\sqrt{\delta}\right)\right]\sigma^2(Q),$$

given the state variable X_s, is greater than $1-e^{-\delta}$ for any $\delta\geqslant 0$ and any $t\in[s,\infty[$.

Proof. Observe that

$$d\left[\overline{\psi}_{s,t}(x,Q)-\varphi_{s,t}(X_s)\right]=\left[A-\phi_{s,t}(Q)S\right]\left[\overline{\psi}_{s,t}(x,Q)-\varphi_{s,t}(X_s)\right]dt+d\overline{M}_{s,t}$$

with an r_1-valued martingale $(\overline{M}_{s,t})_{t\geqslant s}$ defined by

$$\begin{aligned}t\in[s,\infty[\mapsto\overline{M}_{s,t}&=\int_s^t\phi_{s,u}(Q)C'R_2^{-1/2}d(V_u-\overline{V}_u)\\&\quad+R_1^{1/2}\left[(\overline{W}_t-W_t)-(\overline{W}_s-W_s)\right]\\&\overset{law}{=}\sqrt{2}M_{s,t}\end{aligned}$$

with the martingale $(M_{s,t})_{t\in[s,\infty[}$ discussed in (31). The proof now follows the same arguments as the proofs of Theorem 5.2 and Lemma 5.3, thus it is skipped. ∎

In the same vein, recalling that

$$\begin{aligned}d\overline{\psi}_{s,t}(x,Q)&=\left[\left[A-\phi_{s,t}(Q)S\right]\overline{\psi}_{s,t}(x,Q)+\phi_{s,t}(Q)S\varphi_{s,t}(X_s)\right]dt\\&\quad+R_1^{1/2}d\overline{W}_t+\phi_{s,t}(Q)C'R_2^{-1/2}d(V_t-\overline{V}_t),\end{aligned}$$

we find the decomposition

$$\begin{aligned}&\overline{\psi}_{s,t}(x_1,Q_1)-\overline{\psi}_{s,t}(x_2,Q_2)=E_{t|s}(Q_1)(x_1-x_2)\\&\quad+\int_s^t E_{u,t|s}(Q_1)\left[\phi_{s,u}(Q_1)-\phi_{s,u}(Q_2)\right]S\left[\varphi_{s,u}(X_s)-\overline{\psi}_{s,u}(x_2,Q_2)\right]du\\&\quad+\int_s^t E_{u,t|s}(Q_1)\left[\phi_{s,u}(Q_1)-\overline{\psi}_{s,u}(Q_2)\right]d\overline{M}_u\end{aligned}$$

with $\overline{M}_t=\sqrt{2}C'R_2^{-1/2}(V_t-\overline{V}_t)/\sqrt{2}$.

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 08/19/18 to 138.25.168.33. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

THEOREM 5.6. *For any $t \geqslant s \geqslant 0$, $x_1, x_2 \in \mathbb{R}^{r_1}$, $Q_1, Q_2 \in \mathbb{S}^+_{r_1}$, and $n \geqslant 1$ we have the almost sure local contraction estimate*

$$\mathbb{E}\left(\|\overline{\psi}_{s,t}(x_1,Q_1) - \overline{\psi}_{s,t}(x_2,Q_2)\|_2^{2n} | X_s\right)^{\frac{1}{2n}} \leqslant \kappa_E(Q_1) e^{-\nu(t-s)} \|x_1 - x_2\|_2$$
$$+ \sqrt{2} e^{-\nu(t-s)} \chi_0(Q_1,Q_2) \left\{\chi_1(Q_2)\left\|x_2 - X_s\right\|_2 + \sqrt{n}\chi_2(Q_1)\right\} \|Q_1 - Q_2\|_2.$$

The proof of this theorem readily follows that of Theorem 5.4.

The analysis in this section encapsulates and extends the existing convergence and stability results for the Kalman–Bucy filter, as studied, e.g., in [27, 8, 2, 35]. We capture the properties of this convergence in a more quantitative manner than was previously considered. The use of our nonlinear Kalman–Bucy diffusion provides a novel interpretation of the Kalman–Bucy filter that allows one to consider a more general class of signal models in a natural manner.

In particular, stability of the nonlinear Kalman–Bucy diffusion implies convergence of the filter, given arbitrary initial conditions, to the conditional mean of the signal given the observation filtration. Moreover, it implies convergence of the conditional distribution to a Gaussian defined by the conditional mean of the Kalman–Bucy diffusion and its covariance. Similar results were considered by Ocone and Pardoux in [35] but with no quantitative analysis.

Note that our analysis further provides exponential relationships between the actual sample paths of the filter and the signal (with dedicated probability).

This analysis completes our review of the Kalman–Bucy filter and its stability properties.

Acknowledgments. Adrian N. Bishop is with the University of Technology Sydney (UTS) and Data61 (CSIRO), Analytics Research Group. He is also an adjunct Fellow at the Australian National University (ANU).

Pierre Del Moral is with INRIA, Bordeaux Research Center (France) and the University of New South Wales (UNSW), School of Mathematics and Statistics (Australia).

REFERENCES

[1] H. Abou-Kandil, G. Freiling, V. Ionescu, and G. Jank, *Matrix Riccati Equations in Control and Systems Theory*, Birkhäuser, Basel, 2003.
[2] B. D. O. Anderson, *Stability properties of Kalman–Bucy filters*, J. Franklin Inst., 291 (1971), pp. 137–144.
[3] S. Bittanti, A. J. Laub, and J. C. Willems, *The Riccati Equation*, Springer-Verlag, Berlin, 1991.
[4] R. S. Bucy, *Global theory of the Riccati equation*, J. Comput. System Sci., 1 (1967), pp. 349–361.
[5] R. S. Bucy, *The Riccati equation and its bounds*, J. Comput. System Sci., 6 (1972), pp. 342–353.
[6] R. S. Bucy, *Remarks on "A note on bounds on solutions of the Riccati equation"*, IEEE Trans. Automat. Control, 17 (1972), p. 179.
[7] R. S. Bucy, *Structural stability for the Riccati equation*, SIAM J. Control, 13 (1975), pp. 749–753.
[8] R. S. Bucy and P. D. Joseph, *Filtering for Stochastic Processes with Applications to Guidance*, Interscience, New York, 1968.
[9] R. S. Bucy and J. Rodriguez-Canabal, *A negative definite equilibrium and its induced cone of global existence for the Riccati equation*, SIAM J. Math. Anal., 3 (1972), pp. 644–646.
[10] F. M. Callier and J. L. Willems, *Criterion for the convergence of the solution of the Riccati differential equation*, IEEE Trans. Automat. Control, 26 (1981), pp. 1232–1242.

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 08/19/18 to 138.25.168.33. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

[11] F. M. Callier and J. Winkin, *Convergence of the time-invariant Riccati differential equation towards its strong solution for stabilizable systems*, J. Math. Anal. Appl., 192 (1995), pp. 230–257.

[12] G. de Nicolao and M. Gevers, *Difference and differential Riccati equations: A note on the convergence to the strong solution*, IEEE Trans. Automat. Control, 37 (1992), pp. 1055–1057.

[13] P. Del Moral, *Feynman-Kac Formulae*, Springer-Verlag, New York, 2004.

[14] P. Del Moral, *Mean Field Simulation for Monte Carlo Integration*, Chapman & Hall/CRC Monographs on Statistics & Applied Probability 126, CRC Press, Boca Raton, FL, 2013.

[15] P. Del Moral, A. Kurtzmann, and J. Tugaut, *On the Stability and the Uniform Propagation of Chaos of Extended Ensemble Kalman–Bucy Filters*, preprint, arXiv:1606.08256, 2016.

[16] P. Del Moral, A. Kurtzmann, and J. Tugaut, *On the Stability and the Exponential Concentration of Extended Kalman–Bucy Filters*, preprint, arXiv:1606.08251, 2016.

[17] P. Del Moral and J. Tugaut, *On the Stability and the Uniform Propagation of Chaos Properties of Ensemble Kalman–Bucy Filters*, preprint, arXiv:1605.09329, 2016.

[18] B. Delyon, *A note on uniform observability*, IEEE Trans. Automat. Control, 46 (2001), pp. 1326–1327.

[19] G. Evensen, *The Ensemble Kalman Filter: Theoretical formulation and practical implementation*, Ocean Dyn., 53 (2003), pp. 343–367.

[20] M. Fiedler, *Special Matrices and Their Applications in Numerical Mathematics*, 2nd ed., Dover, Mineola, NY, 2008.

[21] M. Gevers, R. R. Bitmead, I. R. Petersen, and R. J. Kaye, *When is the solution of the Riccati equation stabilizing at every iteration?*, in Frequency Domain and State Space Methods for Linear Systems, C. I. Byrnes and A. Lindquist, eds., North-Holland, Amsterdam, 1986, pp. 531–540.

[22] K. L. Hitz, T. E. Fortmann, and B. D. O. Anderson, *A note on bounds on solutions of the Riccati equation*, IEEE Trans. Automat. Control, 17 (1972), p. 178.

[23] R. E. Kalman, *Contributions to the theory of optimal control*, Bol. Soc. Mat. Mex., 5 (1960), pp. 102–119.

[24] R. E. Kalman, *A new approach to linear filtering and prediction problems*, J. Basic Eng., 82 (1960), pp. 35–45.

[25] R. E. Kalman, *Further remarks on "A note on bounds on solutions of the Riccati equation"*, IEEE Trans. Automat. Control, 17 (1972), pp. 179–180.

[26] R. E. Kalman and J. E. Bertram, *Control systems analysis and design via the "second method" of Lyapunov*, J. Basic Eng., 82 (1960), pp. 371–393.

[27] R. E. Kalman and R. S. Bucy, *New results in linear filtering and prediction theory*, J. Basic Eng., 83 (1961), pp. 95–108.

[28] D. Khoshnevisan, *Analysis of Stochastic Partial Differential Equations*, American Mathematical Society, Providence, RI, 2014.

[29] V. Krešimir, *Damped Oscillations of Linear Systems: A Mathematical Introduction*, Lecture Notes in Mathematics 2023, Springer-Verlag, Berlin, 2011.

[30] V. Kucera, *A review of the matrix Riccati equation*, Kybernetika, 9 (1973), pp. 42–61.

[31] F. le Gland, V. Monbet, and V.-D. Tran, *Large Sample Asymptotics for the Ensemble Kalman Filter*, Research report RR-7014, INRIA, 2009, inria-00409060.

[32] A. J. Majda and X. T. Tong, *Robustness and Accuracy of Finite Ensemble Kalman Filters in Large Dimensions*, preprint, arXiv:1606.09321, 2016.

[33] C. Moler and C. Van Loan, *Nineteen dubious ways to compute the exponential of a matrix, twenty-five years later*, SIAM Rev., 45 (2003), pp. 3–49.

[34] R. T. O'Brien, *On stabilizing solutions to the Riccati differential equation*, in Proceedings of the 30th Southeastern Symposium on System Theory, 1998, pp. 80–84.

[35] D. Ocone and E. Pardoux, *Asymptotic stability of the optimal filter with respect to its initial condition*, SIAM J. Control Optim., 34 (1996), pp. 226–243.

[36] M-A. Poubelle, I. R. Petersen, M. R. Gevers, and R. R. Bitmead, *A miscellany of results of an equation of Count J. F. Riccati*, IEEE Trans. Automat. Control, 31 (1986), pp. 651–654.

[37] P. Rebeschini and R. van Handel, *Can local particle filters beat the curse of dimensionality?*, Ann. Appl. Probab., 25 (2015), pp. 2809–2866.

[38] W. T. Reid, *Riccati Differential Equations*, Academic Press, New York, 1972.

[39] K. Reif, S. Gunther, E. Yaz, and R. Unbehauen, *Stochastic stability of the continuous-time extended Kalman filter*, IEE Proc. Control Theory Appl., 147 (2000), pp. 45–52.

[40] Y. F. Ren, *On the Burkholder–Davis–Gundy inequalities for continuous martingales*, Statist. Probab. Lett., 78 (2008), pp. 3034–3039.

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

[41] M. A. Shayman, *Phase portrait of the matrix Riccati equation*, SIAM J. Control Optim., 24 (1986), pp. 1–65.

[42] X. T. Tong, A. J. Majda, and D. Kelly, *Nonlinear stability and ergodicity of ensemble based Kalman filters*, Nonlinearity, 29 (2016), pp. 657–691.

[43] C. Van Loan, *The sensitivity of the matrix exponential*, SIAM J. Numer. Anal., 14 (1977), pp. 971–981.

[44] J. C. Willems, *Least squares stationary optimal control and the algebraic Riccati equation*, IEEE Trans. Automat. Control, 16 (1971), pp. 621–634.

[45] H. K Wimmer, *Monotonicity of maximal solutions of algebraic Riccati equations*, Systems Control Lett., 5 (1985), pp. 317–319.

[46] W. M. Wonham, *On a matrix Riccati equation of stochastic control*, SIAM J. Control, 6 (1968), pp 681–697.

Downloaded 08/19/18 to 138.25.168.33. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.