Karvonen, Toni; Bonnabel, Silvere; Särkkä, Simo; Moulines, Eric

# Bounds on the Covariance Matrix of a Class of Kalman-Bucy Filters for Systems with Non-Linear Dynamics

*Published in:*
Proceedings of 57th IEEE Conference on Decision and Control, CDC 2018

*DOI:*
10.1109/CDC.2018.8619726

Published: 18/01/2019

*Document Version*
Peer reviewed version

*Please cite the original version:*
Karvonen, T., Bonnabel, S., Särkkä, S., & Moulines, E. (2019). Bounds on the Covariance Matrix of a Class of Kalman-Bucy Filters for Systems with Non-Linear Dynamics. In Proceedings of 57th IEEE Conference on Decision and Control, CDC 2018 (Vol. 2018-December, pp. 7176-7181). [8619726] (Proceedings of the IEEE Conference on Decision & Control). Institute of Electrical and Electronics Engineers. https://doi.org/10.1109/CDC.2018.8619726

# Bounds on the covariance matrix of a class of Kalman–Bucy filters for systems with non-linear dynamics

Toni Karvonen, Silvère Bonnabel, Eric Moulines, and Simo Särkkä

This article has been accepted for publication in *Proceedings of 57th IEEE Conference on Decision and Control*, to be held in Miami Beach, Florida, USA, on December 17–19, 2018. When citing this article, use the following reference:

> T. Karvonen, S. Bonnabel, E. Moulines, and S. Särkkä, "Bounds on the covariance matrix of a class of Kalman–Bucy filters for systems with non-linear dynamics," in *57th IEEE Conference on Decision and Control*, 2018.

# Bounds on the covariance matrix of a class of Kalman–Bucy filters for systems with non-linear dynamics

Toni Karvonen, Silvère Bonnabel, Eric Moulines, and Simo Särkkä

***Abstract*— We consider a broad class of Kalman–Bucy filter extensions for continuous-time systems with non-linear dynamics and linear measurements. This class contains, for example, the extended Kalman–Bucy filter, the unscented Kalman–Bucy filter, and most other numerical integration filters. We provide simple upper and lower bounds for the trace of the error covariance, as solved from a matrix Riccati equation, for this class of filters. The upper bounds require assuming that the state is fully observed. The bounds are applied to a simple simultaneous localisation and mapping problem and numerically demonstrated on a two-dimensional trigonometric toy model.**

## I. INTRODUCTION

In classical Kalman filters for continuous and discrete-time linear systems, evolution of the error covariance matrix is controlled by either a matrix Riccati differential or difference equation, respectively. In the literature, numerous results based on various controllability and observability conditions provide upper and lower bounds on solutions of the differential [1]–[5] and difference [6], [7] equations and, when the system is time-invariant, guarantee convergence to the solution of the corresponding algebraic Riccati equation [8]–[10].

Unfortunately, none of these results generalise for the various extensions of the Kalman filter for non-linear systems [11], such as the extended Kalman filter (EKF) or the more recent unscented Kalman filter (UKF) [12] and other similar filters based on numerical integration of Gaussian expectations [13]. This is partially due to the enormous difficulties in easily generalising the notions of observability and controllability to non-linear systems. The only bounds we are aware of that are applicable to, for example, the EKF, appear in [14] (see also [15]). However, although they have seen some applications [16], [17], these bounds, being based on existence of a bounded stabilising gain matrix, are difficult to verify and work with.

We derive upper and lower bounds for the trace of the error covariance of a large class (including, e.g., the EKF and the UKF) of commonly used extensions of the Kalman–Bucy filter to continuous-time systems with non-linear dynamics and linear measurements. It appears that these bounds, simple as they are, have not appeared explicitly in the literature before. In particular, for the UKF and other similar filters we are not aware of any prior work on the topic. Lower bounds require few assumptions besides uniform boundedness of the Jacobian matrix of the dynamics, but to obtain upper bounds we need to assume that the system is fully observed (see Definition 1).

Our main motivation comes from stability and convergence analysis of Kalman-like filters and observers for non-linear systems, where one of the key requirements is that the error covariance remains uniformly bounded [16], [18]–[21]. Although the assumption that the system is fully observed is extremely stringent, we nevertheless believe the bounds we derive are of interest and useful as existing results mainly concern the EKF only and, furthermore, at the moment stability analysis of Kalman filters for non-linear systems using assumptions that are a priori verifiable is actually restricted to fully observed models [22].

The main results of this article are the upper and lower bounds of Propositions 4 and 5 in Section III. We also (qualitatively) compare these bounds to those appearing in [14]. In Section V, we numerically validate the bounds on a two-dimensional toy model and apply them to a simple simultaneous localisation and mapping model from [17].

### A. Some preliminaries

The spectral norm of a matrix $A$ is $\|A\|^2 = \lambda_{\max}(A^\mathsf{T} A)$. In addition to this norm, we make frequent use of the *logarithmic norms* (that, being possibly negative, are not really norms)

$$\mu(A) = \frac{1}{2}\lambda_{\max}(A + A^\mathsf{T})$$

and

$$\nu(A) = \frac{1}{2}\lambda_{\min}(A + A^\mathsf{T}) = -\mu(-A).$$

The logarithmic norms satisfy the triangle inequalities $\mu(A+B) \leq \mu(A) + \mu(B)$ and $\nu(A+B) \geq \nu(A) + \nu(B)$. Recall also the trace and eigenvalue inequalities [23, Section 8.4]

$$\lambda_{\min}(A)\operatorname{tr}(B) \leq \operatorname{tr}(AB) \leq \lambda_{\max}(A)\operatorname{tr}(B)$$

for a symmetric $A$ and a positive-semidefinite $B$ and

$$\nu(A)\operatorname{tr}(B) \leq \operatorname{tr}(AB) \leq \mu(A)\operatorname{tr}(B) \tag{1}$$

for a positive-semidefinite $B$.

For a differentiable function $g \colon \mathbb{R}^n \to \mathbb{R}^n$, $[J_g(x)]_{ij} = (\partial g_i / \partial x_j)(x)$ stands for the Jacobian matrix of $g$ evaluated at $x \in \mathbb{R}^n$. We define the *logarithmic Lipschitz constants*

$$M(g) = \frac{1}{2}\sup_{x\in\mathbb{R}^n} \lambda_{\max}\big[J_g(x) + J_g(x)^\mathsf{T}\big] = \sup_{x\in\mathbb{R}^n} \mu[J_g(x)]$$

T. Karvonen was supported by Aalto ELEC Doctoral School and the Foundation for Aalto University Science and Technology.

T. Karvonen and S. Särkkä are with the Department of Electrical Engineering and Automation, Aalto University, Espoo, Finland. toni.karvonen@aalto.fi & simo.sarkka@aalto.fi

S. Bonnabel is with the Centre for Robotics, MINES ParisTech, Paris, France. silvere.bonnabel@mines-paristech.fr

E. Moulines is with Centre de Mathématiques Appliquées, École Polytechnique, Palaiseau, France. eric.moulines@polytechnique.edu

and

$$N(g) = \frac{1}{2} \inf_{x\in\mathbb{R}^n} \lambda_{\min}\big[J_g(x) + J_g(x)^\mathsf{T}\big] = \inf_{x\in\mathbb{R}^n} \nu[J_g(x)].$$

The following inequality for the Euclidean inner product and norm will be useful:

$$N(g)\,\|x-y\|^2 \leq \langle x-y, g(x)-g(y)\rangle \leq M(g)\,\|x-y\|^2 \tag{2}$$

for any $x, y \in \mathbb{R}^n$.

## II. KALMAN–BUCY FILTER EXTENSIONS FOR NON-LINEAR SYSTEMS

This section briefly introduces stochastic dynamic models with non-linear dynamics and extensions of the classical Kalman–Bucy filter [24] for estimation of the latent state of such systems.

### A. Models

We consider continuous-time models formulated as stochastic differential equations of the (Itô) form

$$\begin{aligned} \mathrm{d}X_t &= f(X_t)\,\mathrm{d}t + Q^{1/2}\,\mathrm{d}W_t, \\ \mathrm{d}Y_t &= HX_t\,\mathrm{d}t + R^{1/2}\,\mathrm{d}V_t, \end{aligned} \tag{3}$$

where $X_t \in \mathbb{R}^n$ stands for the latent *state*, initialised from $X_0 \sim \mathcal{N}(\widehat{X}_0, P_0)$, that is to be estimated from the incomplete and noisy *measurements* $Y_t \in \mathbb{R}^m$. The state evolution is dictated by the continuously differentiable *dynamics* $f\colon \mathbb{R}^n \to \mathbb{R}^n$ and the measurements are obtained through the linear *measurement model* $H \in \mathbb{R}^{m\times n}$. The noise terms $W_t$ and $V_t$ are independent standard Brownian motions with positive-definite covariance matrices $Q$ and $R$.

In Section III we derive upper bounds on the error covariance matrix under the assumption that the model is fully observed.

*Definition 1:* The model (3) is *fully observed* if $m \geq n$ and the matrix $H$ is of full rank.

A model is typically fully observed when $m = n$ and $H$ is the identity (or invertible). The case $m > n$ may arise in the presence of redundant sensor information. In any case, a model being fully observed is a fairly restrictive condition. Indeed, knowledge of the dynamics $f$, while often beneficial, is then not strictly necessary to estimate the state (but it helps reducing the variance of the estimate). Nevertheless, fully observed models sometimes appear in applications [17], [25].

### B. Kalman–Bucy filtering

Let $\mathcal{F}_t = \sigma(Y_s \mid 0 \leq s \leq t)$ stand for the sigma-algebra generated by the continuous-time measurements. In filtering, the aim is to compute or approximate the filtering distributions $X_t \mid \mathcal{F}_t$. When the function $f$ is not affine, these distributions are in general intractable, necessitating use of approximations.

For each $t \geq 0$, let $\ell_t\colon C^1(\mathbb{R}^n) \to \mathbb{R}^n$ be a vector-valued (random) functional acting on continuously differentiable functions. A *Kalman–Bucy filter* for the model (3) produces the Gaussian approximation

$$\mathcal{N}(\widehat{X}_t, P_t) \approx X_t \mid \mathcal{F}_t$$

whose mean, used as an estimate to the latent state $X_t$, is computed from

$$\mathrm{d}\widehat{X}_t = \ell_t(f)\,\mathrm{d}t + P_t H^\mathsf{T} R^{-1}(\mathrm{d}Y_t - H\widehat{X}_t\,\mathrm{d}t), \tag{4}$$

where $P_t \in \mathbb{R}^{n\times n}$ is the positive-semidefinite *error covariance matrix*. This matrix is solved from the *Riccati differential equation*

$$\partial_t P_t = L_{t,P_t}(f) + L_{t,P_t}(f)^\mathsf{T} + Q - P_t S P_t, \tag{5}$$

where $S = H^\mathsf{T} R^{-1} H$ (note that this matrix is invertible when the model is fully observed) and $L_{t,P_t}\colon C^1(\mathbb{R}^n) \to \mathbb{R}^{n\times n}$ are matrix-valued functionals (dependency on $P_t$ is purposely made explicit) that depend on $t$ and $P_t$. Moreover, we will always assume that the functionals $L_{t,P_t}$ are sufficiently regular to ensure that (5) is well-defined and that a positive-definite solution exists (for our purposes, the functionals will anyway be chosen such that (6) below holds). The equation is initialised from a positive-semidefinite $P_0 = \mathrm{Cov}(X_0)$.

*Assumption 2:* Assume that

- The Lipschitz constants $M(f)$ and $N(f)$ are bounded;
- The functionals $L_{t,P_t}$ safisfy
$$N(f)\operatorname{tr}(P_t) \leq \operatorname{tr}[L_{t,P_t}(f)] \leq M(f)\operatorname{tr}(P_t) \tag{6}$$
for every $t \geq 0$.

The first assumption is classical, and its analogue for linear systems often appears in stability analysis of linear Kalman and Kalman–Bucy filters [5], [7]. Indeed, considering the EKF, for instance, in the light of the calculations to follow, we see that it is hopeless to try to derive bounds on the solution if we allow the estimate $\widehat{X}_t$ to drive the Jacobian $J_f(\widehat{X}_t)$ into regions where it can grow unboundedly[1]. The second assumption is painless, as it holds for all sensible extensions of the Kalman–Bucy filter. We next provide two examples of such extensions.

### C. The extended Kalman–Bucy filter

With the choice $\ell_t(f) = f(\widehat{X}_t)$ and $L_{t,P_t} = J_f(\widehat{X}_t)P_t$, we obtain the classical extended Kalman–Bucy filter. In this case, Equation (1) implies that

$$N(f)\operatorname{tr}(P_t) \leq \operatorname{tr}[J_f(\widehat{X}_t)P_t] \leq M(f)\operatorname{tr}(P_t).$$

That is, Assumption 2 holds.

### D. Numerical integration filters (e.g., the UKF)

Let $\xi_1, \ldots, \xi_N \in \mathbb{R}^n$ and $w_1, \ldots, w_N \geq 0$ be *unit sigma-points* and *weights* and $\lambda > 0$ a constant such that

$$\sum_{i=1}^N w_i p(\mu + \lambda\sqrt{P}\xi_i) = \int_{\mathbb{R}^n} p(x)\mathcal{N}(x \mid \mu, P)\,\mathrm{d}x \tag{7}$$

for any multivariate polynomial $p$ of degree at most two and any $\mu \in \mathbb{R}^n$ and $P \in \mathbb{R}^{n\times n}$. In particular, as (7) holds for

[1] This does not mean that the condition is necessary, but it means that it cannot be avoided if we want to consider the Riccati equation on its own, without specifying the behavior of the estimate $\widehat{X}_t$. Of course the assumption could be relaxed by, for instance, assuming the state space to be bounded.

the polynomials $p(x)=1$ and $p(x)=x$, this implies that $\sum_{i=1}^N w_i = 1$ and $\sum_{i=1}^N w_i \lambda \sqrt{P} \xi_i = 0$ for any positive-definite matrix $P$. Denote $\chi_{t,i} = \lambda \sqrt{P_t} \xi_i$ and set

$$\begin{aligned}\ell_t(f) &= \sum_{i=1}^N w_i f(\chi_{t,i}), \\ L_{t,P_t}(f) &= \sum_{i=1}^N w_i \chi_{t,i} f(\widehat{X}_t + \chi_{t,i})^\mathsf{T}.\end{aligned}$$

For instance, the unscented transform, the tensor product Gauss–Hermite rule, and other popular numerical integration rules used in Kalman filtering fit this framework. A concrete example is (a particular version of) the unscented Kalman–Bucy filter that uses the unit sigma-points

$$\xi_1 = 0, \quad \xi_{i+1} = e_i, \quad \xi_{n+i+1} = -e_i$$

for $i=1,\ldots,n$, where $e_i$ are the unit coordinate vectors, and the weights

$$w_0 = \frac{1}{n+1}, \qquad w_{i+1} = w_{n+i+1} = \frac{1}{2(n+1)}$$

for $i=1,\ldots,n$. Of course, different unit sigma-points and weights could—and often are—be used in $\ell_t$ and $L_{t,P_t}$; this has no effect on our analysis.

To verify that (6) holds for this class of filters, observe that

$$\begin{aligned}&\operatorname{tr}[L_{t,P_t}(f)] \\ &\quad= \sum_{i=1}^N w_i \big\langle \chi_{t,i}, f(\widehat{X}_t + \chi_{t,i}) \big\rangle \\ &\quad= \sum_{i=1}^N w_i \big\langle (\widehat{X}_t + \chi_{t,i}) - \widehat{X}_t, f(\widehat{X}_t + \chi_{t,i}) - f(\widehat{X}_t) \big\rangle,\end{aligned}$$

where we have exploited the facts that the weights sum to one and $\sum_{i=1}^N w_i \chi_{t,i} = 0$. An application of (2) and the fact that (7) holds for second-degree polynomials yield

$$\begin{aligned}\operatorname{tr}\left[L_{t,P_t}(f)\right] &\leq M(f) \sum_{i=1}^N w_i \left\|\chi_{t,i}\right\|^2 \\ &= \int_{\mathbb{R}^n} \|x\|^2 \mathcal{N}(x \mid 0, P_t) \,\mathrm{d}x \\ &= M(f) \operatorname{tr}(P_t).\end{aligned}$$

That $\operatorname{tr}[L_{t,P_t}(f)] \geq N(f)\operatorname{tr}(P_t)$ follows along similar lines.

## III. Bounds on the Error Covariance

The main results of this section are Propositions 4 and 5 that contain upper (under the full observation assumption) and lower bounds on the trace of $P_t$ as solved from the Riccati differential equation (5). Bounds somewhat similar to ours, albeit without a proof and less general, have appeared in [26] in the context of stability analysis of the ensemble Kalman–Bucy filter.

### A. An auxiliary lemma

There is an extensive theory on scalar Riccati differential equations [27, Chapter 3]. The following basic result on exponential convergence to the equilibrium of solutions of time-invariant scalar Riccati differential equations will be useful to us. For completeness, we also present its proof. See for example [9], [10], [28] for similar results regarding matrix Riccati differential equations.

*Lemma 3:* Let $a, c > 0$, $b$, and $x_0 \geq 0$ be constants and define

$$\begin{aligned}\alpha(a,b,c) = \alpha &= \sqrt{ac+b^2}, \\ \beta(a,b,c,x_0) = \beta &= \frac{ax_0 - \alpha - b}{ax_0 + \alpha - b}.\end{aligned} \tag{8}$$

Consider the scalar Riccati differential equation

$$\partial_t x_t = -a x_t^2 + 2bx_t + c \tag{9}$$

with the initial condition $x_0$. The solution $x_t$ satisfies

$$x^+ - \frac{2\alpha}{a} |\beta| \,\mathrm{e}^{-2\alpha t} \leq x_t \leq x^+ + \frac{2\alpha}{a} \frac{1}{1-\beta} \,\mathrm{e}^{-2\alpha t}$$

where

$$x^+ := \frac{b+\alpha}{a} > 0$$

is the equilibrium point and $\beta \leq 0$ when $x_0 \leq x^+$ and $0 \leq \beta < 1$ when $x_0 \geq x^+$.

*Proof:* There exist two solutions to the quadratic equation (or the *algebraic Riccati equation*) $-ax^2 + 2bx + c = 0$:

$$x^+ = \frac{b+\alpha}{a} > 0 \quad \text{ and } \quad x^- = \frac{b-\alpha}{a} < 0.$$

Since $x_0 \geq 0$ and $\partial_t x_t = c$ if $x_t = 0$, any solution to (9) must be non-negative. It follows that there is an equilibrium at $x^+$ and that $x_t < x^+$ if $x_0 < x^+$ and $x_t > x^+$ if $x_0 > x^+$.

Suppose that $x_t \neq x^+$ and denote $\varepsilon_t = x_t - x^+ \neq 0$. The error $\varepsilon_t$ satisfies the differential equation

$$\partial_t \varepsilon_t = \varepsilon_t(-a\varepsilon_t - 2\alpha).$$

The change of variables $z_t = 1/\varepsilon_t$, valid because $\varepsilon_t \neq 0$, leads to the affine differential equation

$$\partial_t z_t = 2\alpha z_t + a,$$

the solution to which is

$$z_t = -\frac{a}{2\alpha} + \frac{a}{2\alpha\beta} \,\mathrm{e}^{2\alpha t}.$$

We thus obtain the solution

$$x_t = \frac{1}{z_t} + x^+ = x^+ + \frac{2\alpha}{a} \frac{\beta \,\mathrm{e}^{-2\alpha t}}{1 - \beta \,\mathrm{e}^{-2\alpha t}}$$

to the scalar Riccati equation (9). When $x_0 \leq x^+$, $\beta \leq 0$ and

$$x_t \geq x^+ - \frac{2\alpha}{a} |\beta| \,\mathrm{e}^{-2\alpha t},$$

whereas $1 > \beta \geq 0$ and $1 - \beta \,\mathrm{e}^{-2\alpha t} \geq 1 - \beta > 0$ when $x_0 \geq x^+$, yielding

$$x_t \leq x^+ + \frac{2\alpha}{a} \frac{1}{1-\beta} \,\mathrm{e}^{-2\alpha t}.$$

∎

Note that the simpler bounds

$$\min(x_0, x^+) \le x_t \le \max(x_0, x^+)$$

are established at the beginning of the proof and do not required explicitly solving the differential equation (9).

### B. Upper and lower bounds on $P_t$

Lemma 3 can be now used to derive lower and upper bounds for the solution $P_t$ of the continuous-time matrix Riccati differential equation (5). The proofs make use of the standard comparison theorem (see e.g. [29, Appendix E]) that allows for deducing inequalities of solutions to differential equations from differential inequalities (e.g., $\partial_t x_t \le \partial_t y_t$ implies $x_t \le y_t$).

*Proposition 4 (Upper bound):* Suppose that the model is fully observed (see Definition 1) and denote

$$\begin{aligned}\alpha &= \alpha\big(\lambda_{\min}(S)/n, M(f), \operatorname{tr}(Q)\big),\\ \beta &= \beta\big(\lambda_{\min}(S)/n, M(f), \operatorname{tr}(Q), \operatorname{tr}(P_0)\big),\end{aligned}$$

where $\alpha(a,b,c)$ and $\beta(a,b,c,x_0)$ are defined in (8). If Assumption 2 holds and $P_t$ is the solution to the Riccati differential equation (5), then

$$\begin{aligned}\operatorname{tr}(P_t) &\le \frac{M(f)+\alpha}{\lambda_{\min}(S)/n} + \frac{2\alpha}{(1-\beta)\lambda_{\min}(S)/n}\,\mathrm{e}^{-2\alpha t}\\ &\xrightarrow[t\to\infty]{} \frac{M(f)+\alpha}{\lambda_{\min}(S)/n}\end{aligned} \tag{10}$$

for every $t \ge 0$.

*Proof:* Since $H \in \mathbb{R}^{m\times n}$ is of full rank and $m \ge n$, the matrix $S = H^\mathsf{T} R^{-1} H$ is positive-definite. This implies that $\lambda_{\min}(S) > 0$. Evolution of the trace of $P_t$ is

$$\begin{aligned}\partial_t \operatorname{tr}(P_t) &= \operatorname{tr}\left[L_{t,P_t}(f) + L_{t,P_t}(f)^\mathsf{T}\right] + \operatorname{tr}(Q) - \operatorname{tr}(SP_t^2)\\ &\le 2M(f)\operatorname{tr}(P_t) + \operatorname{tr}(Q) - \lambda_{\min}(S)\operatorname{tr}(P_t^2)\\ &\le 2M(f)\operatorname{tr}(P_t) + \operatorname{tr}(Q) - [\lambda_{\min}(S)/n]\operatorname{tr}(P_t)^2,\end{aligned}$$

where we have used Jensen's inequality in the form $(\sum_{i=1}^n a_i)^2/n \le \sum_{i=1}^n a_i^2$. The upper bound of Lemma 3 establishes the claim. ∎

*Proposition 5 (Lower bound):* Denote

$$\begin{aligned}\alpha &= \alpha\big(\lambda_{\max}(S), N(f), \operatorname{tr}(Q)\big),\\ \beta &= \beta\big(\lambda_{\max}(S), N(f), \operatorname{tr}(Q), \operatorname{tr}(P_0)\big),\end{aligned}$$

where $\alpha(a,b,c)$ and $\beta(a,b,c,x_0)$ are defined in (8). If Assumption 2 holds and $P_t$ is the solution to the Riccati differential equation (5), then

$$\begin{aligned}\operatorname{tr}(P_t) &\ge \frac{N(f)+\alpha}{\lambda_{\max}(S)} - \frac{2\alpha\,|\beta|}{\lambda_{\max}(S)}\,\mathrm{e}^{-2\alpha t}\\ &\xrightarrow[t\to\infty]{} \frac{N(f)+\alpha}{\lambda_{\max}(S)}\end{aligned} \tag{11}$$

for every $t \ge 0$.

*Proof:* This time, the trace satisfies

$$\begin{aligned}\partial_t \operatorname{tr}(P_t) &= \operatorname{tr}\left[L_{t,P_t}(f) + L_{t,P_t}(f)^\mathsf{T}\right] + \operatorname{tr}(Q) - \operatorname{tr}(SP_t^2)\\ &\ge 2N(f)\operatorname{tr}(P_t) + \operatorname{tr}(Q) - \lambda_{\max}(S)\operatorname{tr}(P_t^2)\\ &\ge 2N(f)\operatorname{tr}(P_t) + \operatorname{tr}(Q) - \lambda_{\max}(S)\operatorname{tr}(P_t)^2,\end{aligned}$$

where we have used the inequality $\sum_{i=1}^n a_i^2 \le (\sum_{i=1}^n a_i)^2$. The claim follows from the lower bound of Lemma 3. ∎

## IV. DISCUSSION

We now discuss some properties of the bounds proved in the previous section.

### A. Qualitative properties

Suppose that $H = I$. Consequently, $S = R^{-1}$. When written out, the limiting bounds in (10) and (11) are

$$\begin{aligned}&\operatorname{tr}(P_t)\\ &\quad\le \frac{nM(f) + \sqrt{n\lambda_{\min}(S)\operatorname{tr}(Q) + n^2M(f)^2}}{\lambda_{\min}(S)}\\ &\quad= \lambda_{\max}(R)\Big(nM(f) + \sqrt{n\operatorname{tr}(Q)/\lambda_{\max}(R) + n^2M(f)^2}\Big)\end{aligned}$$

and

$$\begin{aligned}\operatorname{tr}(P_t) &\ge \frac{N(f) + \sqrt{\lambda_{\max}(S)\operatorname{tr}(Q) + N(f)^2}}{\lambda_{\max}(S)}\\ &\quad= \lambda_{\min}(R)\Big(N(f) + \sqrt{\operatorname{tr}(Q)/\lambda_{\min}(R) + N(f)^2}\Big).\end{aligned}$$

Some intuitive behaviour can be observed:

- Larger model and measurement noises lead to larger bounds since both bounds grow linearly with $R$ and as a square of root of $Q$;
- For $M(f) < 0$ (i.e., the homogeneous system $\partial_t x_t = f(x_t)$ is exponentially stable), we obtain from $\sqrt{a+b} \le \sqrt{a} + \sqrt{b}$ the simpler bound

  $$\operatorname{tr}(P_t) \le \sqrt{n\operatorname{tr}(Q)\lambda_{\max}(R)}.$$

  That is, if there is little model noise, the filter is accurate because it correctly expects that $X_t$ remains close to the origin. In fact, the model does not need to be assumed fully observed in this case since, in the proof of Proposition 4, the term $-\operatorname{tr}(SP_t^2)$ can be removed and boundedness concluded from Grönwall's inequality.

### B. Comparison to previous bounds

In this section we compare the bounds of Propositions 4 and 5 to those derived in [14] for the EKF by a control-theoretic argument. Recall the EKF Riccati equation

$$\partial_t P_t = J_f(\widehat{X}_t)P_t + P_tJ_f(\widehat{X}_t)^\mathsf{T} + Q - P_tSP_t.$$

The model (3) is *uniformly detectable* if there is a matrix-valued function $\Lambda_d(x)$ such that

$$\sup_{x\in\mathbb{R}^n} \|\Lambda_d(x)\| = \|\Lambda_d\| < \infty$$

and

$$\sup_{x\in\mathbb{R}^n} \mu\big[J_f(x) - \Lambda_d(x)H\big] \le -\lambda_d$$

for $\lambda_d > 0$. It is *uniformly controllable* if there exists a matrix-valued function $\Lambda_c(x)$ such that

$$\sup_{x\in\mathbb{R}^n} \|\Lambda_c(x)\| = \|\Lambda_c\| < \infty$$

and

$$\inf_{x\in\mathbb{R}^n} \mu\big[J_f(x) + Q^{1/2}\Lambda_c(x)\big] \ge \lambda_c > 0$$

for $\lambda_c > 0$.

*Proposition 6 (Theorem 7 in [14]):* If the model (3) is uniformly detectable and controllable, then

$$\|P_t\| \leq \|P_0\| + \big(\,\|Q\| + \|R\|\,\|\Lambda_d\|^2\,\big)/(2\lambda_d),$$
$$\|P_t^{-1}\| \leq \|P_0^{-1}\| + \big(\,\|S\| + \|\Lambda_c\|^2\,\big)/(2\lambda_c).$$

For an easy comparison with our bounds, consider the case $H = I$ and $Q = qI$ for some $q > 0$. Then

$$\sup_{x \in \mathbb{R}^n} \mu\big[J_f(x) - \Lambda_d(x)\big] \leq M(f) + \sup_{x \in \mathbb{R}} \mu[-\Lambda(x)],$$

which means that $\Lambda_d(x) = [M(f) + \alpha_d]I$ for some $\alpha_d > 0$ yields $\lambda_d = \alpha_d$. Similarly, $\Lambda_c(x) = q^{-1/2}(|N(f)| + \alpha_c)$ yields $\lambda_c = \alpha_c$. Then the bounds of Proposition 6 become

$$\|P_t\| \leq \|P_0\| + \big(q + \|R\|\,[M(f) + \alpha_d]^2\big)/(2\alpha_d),$$
$$\|P_t^{-1}\| \leq \|P_0^{-1}\| + \big(\,\|R^{-1}\| + q^{-1}(|N(f)| + \alpha_c)^2\big)/(2\alpha_c).$$

The upper bound is linear in $q$ and $R$ and quadratic in $M(f)$ while, as discussed in Section IV-A, our upper bound only grows linearly in $M(f)$ and as a square root of $Q$. Furthermore, unlike our bounds where dependency on $P_0$ decays exponentially, these bounds retain $P_0$.

## V. EXAMPLES

In this section we apply the error covariance bounds of Section III to a fully observed simultaneous localisation and mapping (SLAM) model from [17] and conduct numerical simulations on a simple two-dimensional toy model.

### A. A SLAM application example

As a simple example application of our bounds we use a continuous-time fully observed SLAM model from [17]. In this article, the authors consider a SLAM problem, where a wheeled robot observes fixed landmarks of the environment in its frame and tries to estimate their positions as well as its position and orientation. In robocentric mapping, all equations are written in the robot's frame. This yields at all times a (closed-loop) estimate of the landmarks' positions in the robot frame (the map), and the trajectory of the robot and the landmarks in the global (fixed) frame can be recovered by open loop integration. We assume that the wheeled robot's dynamics are described by the unicycle equations (also known as the nonholonomic car equations). If there are $p$ landmarks, the state $X_t \in \mathbb{R}^{2p}$ at time $t$ consists of ranges $d_{t,i} \geq 0$ (distance between the landmark $i$ and the robot) and bearings $\vartheta_{t,i} \in [0, 2\pi[$ (direction of landmark $i$ viewed from the robot), as follows:

$$X_t = \begin{bmatrix} d_{t,1} & \vartheta_{t,1} & \cdots & d_{t,p} & \vartheta_{t,p} \end{bmatrix}^\mathsf{T}.$$

The dynamics then write [17]

$$f_t(X_t) = \begin{bmatrix} f_{t,1}(X_t) & \cdots & f_{t,p}(X_t) \end{bmatrix}^\mathsf{T} \tag{12}$$

with

$$f_{t,i}(X_t) = \begin{bmatrix} -v_{t,r} \cos \vartheta_{t,i} \\ \frac{v_{t,r}}{d_{t,i}} \sin \vartheta_{t,i} - w_{t,r} \end{bmatrix},$$

where $v_{t,r} \geq 0$ and $w_{t,r} \in \mathbb{R}$ are the robot translational and angular velocities, respectively. The measurement model is $H = I$, which corresponds to ranges and bearings to all of the $p$ landmarks being constantly measured (in other terms, all landmarks remain visible during the experiments, and if this is not the case one must work with subsets of landmarks). The full model with noise-free dynamics is therefore

$$\mathrm{d}X_t = f_t(X_t)\,\mathrm{d}t,$$
$$\mathrm{d}Y_t = X_t\,\mathrm{d}t + R^{1/2}\,\mathrm{d}V_t$$

for some measurement noise covariance $R$. As in [17], we make the additional (reasonable) assumptions that $v_{t,r}$ and $w_{t,r}$ remain uniformly bounded in time and $d_{t,i} \geq 1$ for each $i = 1, \ldots, p$ and $t \geq 0$. From this it follows that

$$-\infty < N \leq \inf_{t \geq 0} N(f_t) \leq \sup_{t \geq 0} M(f_t) \leq M < \infty. \tag{13}$$

for some constants $N$ and $M$.

For estimating the latent state $X_t$, we consider a modified EKF whose Riccati equation is

$$\partial_t P_t = J_{f_t}(\widehat{X}_t) P_t + P_t J_{f_t}(\widehat{X}_t)^\mathsf{T} + Q_{\text{tu}} - P_t S P_t,$$

where $Q_{\text{tu}}$ is a positive-definite tuning parameter. Despite time-varying dynamics, the error covariance bounds derived in Section III are valid if $M(f)$ and $N(f)$ are replaced with $M$ and $N$ from (13). At $t \to \infty$, we obtain the bounds

$$\operatorname{tr}(P_t) \leq r_2\Big(2pM + \sqrt{2p\operatorname{tr}(Q_{\text{tu}})/r_2 + 4p^2M^2}\Big),$$
$$\operatorname{tr}(P_t) \geq r_1\Big(N + \sqrt{\operatorname{tr}(Q_{\text{tu}})/r_1 + N^2}\Big),$$

where $r_1 = \lambda_{\min}(R)$ and $r_2 = \lambda_{\max}(R)$.

### B. A two-dimensional toy model

In this section we consider the two-dimensional fully observed trigonometric model

$$\mathrm{d}\begin{pmatrix} X_{t,1} \\ X_{t,2} \end{pmatrix} = \begin{pmatrix} -\sin(X_{t,1}) + \cos(X_{t,2}) \\ \cos(X_{t,2}) - \sin(X_{t,1}) \end{pmatrix} \mathrm{d}t + Q^{1/2}\,\mathrm{d}W_t,$$
$$\mathrm{d}Y_t = X_t\,\mathrm{d}t + R^{1/2}\,\mathrm{d}V_t$$

with diagonal noise covariances $Q = 0.2I$ and $R = 0.2I$. The state is initialised from $X_0 = (1, 1)$ and the Riccati equation from $P_0 = 0.3I$. The logarithmic Lipschitz constants are $M(f) = \sqrt{2}$ and $N(f) = -\sqrt{2}$. We experiment with (i) the extended Kalman–Bucy filter and (ii) the unscented Kalman–Bucy filter (as formulated in Section II-D). The stochastic differential equations are solved up to time $t = 100$ using the Euler–Maruyama method with step-size 0.01. Figures 1 and 2 display a particular realisation of the state and the estimates by the EKF and the UKF. Figure 3 depicts the upper and lower bounds for $\operatorname{tr}(P_t)$, as computed from Propositions 4 and 5 and five actual trace trajectories for both the EKF and the UKF. We observe that the theoretical bounds are valid.

## VI. CONCLUSIONS

In Propositions 4 and 5 we have proved bounds on trace of the error covariance of Kalman–Bucy filters for systems with non-linear dynamics and linear (or fully observed) measurements. As proved in Section II, the bounds are valid for a large class of prevalent filters. Possibly the most straightforward generalisation would be to systems whose partially observed state components are asymptotically stable.

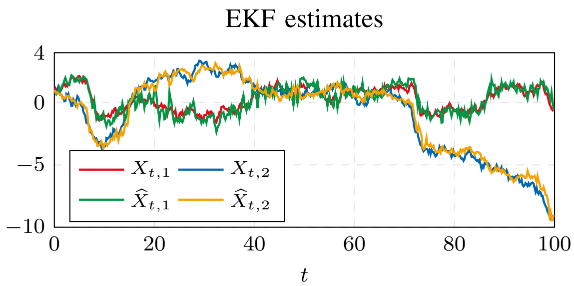


**Fig. 1:** State components and their EKF estimates for a particular realisation of the system of Section V-B.

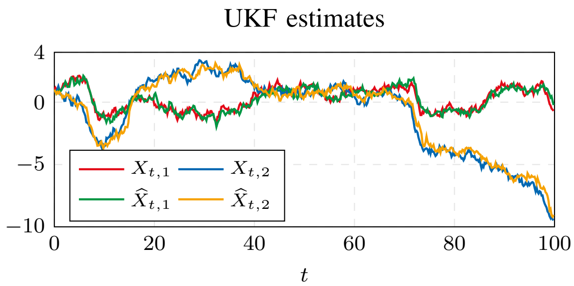


**Fig. 2:** State components and their UKF estimates for a particular realisation of the system of Section V-B.

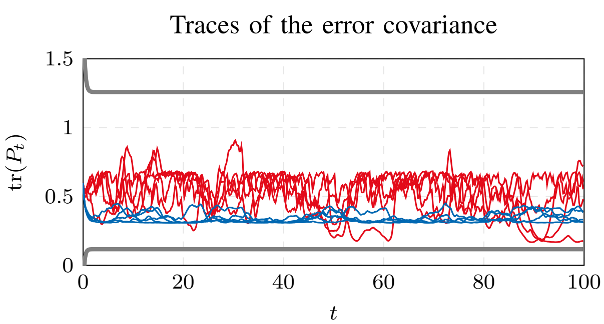


**Fig. 3:** The theoretical upper and lower bounds of Propositions 4 and 5 (gray) and five actual trace trajectories for the EKF (red) and the UKF (blue). The norm bounds of Proposition 6, applicable only to the EKF, are (using optimal values of $\alpha_d$ and $\alpha_c$) $\|P_t\| \leq 0.93$ and $\|P_t^{-1}\| \leq 19.1$ while the empirically observed bounds are $\|P_t\| \leq 0.63$ and $\|P_t^{-1}\| \leq 15.7$.

## VII. Acknowledgements

We thank Marc Jungers for commentary and the present proof of Lemma 3.

## References

[1] R. S. Bucy, "Global theory of the Riccati equation," *Journal of Computer and System Sciences*, vol. 1, no. 4, pp. 349–361, 1967.

[2] K. N. Murty, K. R. Prasad, and M. A. S. Srinivas, "Upper and lower bounds for the solution of the general matrix Riccati differential equation," *Journal of Mathematical Analysis and Applications*, vol. 147, no. 1, pp. 12–21, 1990.

[3] B. Delyon, "A note on uniform observability," *IEEE Transactions on Automatic Control*, vol. 46, no. 8, pp. 1326–1327, 2001.

[4] M. Pengov, E. Richard, and J.-C. Vivalda, "On the boundedness of the solutions of the continuous Riccati equation," *Journal of Inequalities and Applications*, vol. 6, pp. 641–649, 2001.

[5] A. N. Bishop and P. Del Moral, "On the stability of Kalman–Bucy diffusion processes," *SIAM Journal on Control and Optimization*, vol. 55, no. 6, pp. 4015–4047, 2017.

[6] J. J. Deyst, Jr. and C. F. Price, "Conditions for asymptotic stability of the discrete minimum-variance linear estimator," *IEEE Transactions on Automatic Control*, vol. 13, no. 6, pp. 702–705, 1968.

[7] B. D. O. Anderson and J. B. Moore, "Detectability and stabilizability of time-varying discrete-time linear systems," *SIAM Journal on Control and Optimization*, vol. 19, no. 1, pp. 20–32, 1981.

[8] S. Chan, G. Goodwin, and K. Sin, "Convergence properties of the Riccati difference equation in optimal filtering of nonstabilizable systems," *IEEE Transactions on Automatic Control*, vol. 29, no. 2, pp. 110–118, 1984.

[9] F. M. Callier, J. Winkin, and J. L. Willems, "On the exponential convergence of the time-invariant matrix Riccati differential equation," in *31st IEEE Conference on Decision and Control*, 1992, pp. 1536–1537.

[10] F. M. Callier and J. Winkin, "Convergence of the time-invariant Riccati differential equation towards its strong solution for stabilizable systems," *Journal of Mathematical Analysis and Applications*, vol. 192, no. 1, pp. 230–257, 1995.

[11] S. Särkkä, *Bayesian Filtering and Smoothing*, ser. Institute of Mathematical Statistics Textbooks. Cambridge University Press, 2013.

[12] S. J. Julier, J. K. Uhlmann, and H. F. Durrant-Whyte, "A new method for the nonlinear transformation of means and covariances in filters and estimators," *IEEE Transactions on Automatic Control*, vol. 45, no. 3, pp. 477–482, 2000.

[13] Y. Wu, D. Hu, M. Wu, and X. Hu, "A numerical-integration perspective on Gaussian filters," *IEEE Transactions on Automatic Control*, vol. 54, no. 8, pp. 2910–2921, 2006.

[14] J. S. Baras, A. Bensoussan, and M. R. James, "Dynamic observers as asymptotic limits of recursive filters: Special cases," *SIAM Journal on Applied Mathematics*, vol. 48, no. 5, pp. 1147–1158, 1988.

[15] A. J. Krener, *The Convergence of the Extended Kalman Filter*, ser. Lecture Notes in Control and Information Sciences, 2003, vol. 286, pp. 173–182.

[16] K. Reif, S. Günther, E. Yaz, and R. Unbehauen, "Stochastic stability of the continuous-time extended Kalman filter," *IEE Proceedings - Control Theory and Applications*, vol. 147, no. 1, pp. 45–52, 2000.

[17] A. N. Bishop and P. Jensfelt, "A stochastically stable solution to the problem of robocentric mapping," in *IEEE International Conference on Robotics and Automation*, 2009, pp. 1615–1622.

[18] K. Reif, S. Günther, E. Yaz, and R. Unbehauen, "Stochastic stability of the discrete-time extended Kalman filter," *IEEE Transactions on Automatic Control*, vol. 44, no. 4, pp. 714–728, Apr 1999.

[19] K. Xiong, H. Zhang, and C. W. Chan, "Performance evaluation of UKF-based nonlinear filtering," *Automatica*, vol. 42, no. 2, pp. 261–270, 2006.

[20] S. Bonnabel and J.-J. Slotine, "A contraction theory-based analysis of the stability of the deterministic extended Kalman filter," *IEEE Transactions on Automatic Control*, vol. 60, no. 2, pp. 565–569, 2015.

[21] T. Karvonen and S. Särkkä, "Fourier–Hermite series for stochastic stability analysis of non-linear Kalman filters," in *19th International Conference on Information Fusion*, 2016, pp. 1829–1836.

[22] P. Del Moral, A. Kurtzmann, and J. Tugaut, "On the stability and the exponential concentration of extended Kalman-Bucy filters," 2016, preprint, arXiv:1606.08251v2.

[23] D. S. Bernstein, *Matrix Mathematics: Theory, Facts, and Formulas*, 2nd ed. Princeton University Press, 2009.

[24] R. E. Kalman and R. S. Bucy, "New results in linear filtering and prediction theory," *Journal of Basic Engineering*, vol. 83, no. 1, pp. 95–108, 1961.

[25] S. Särkkä, V. Tolvanen, J. Kannala, and E. Rahtu, "Adaptive Kalman filtering and smoothing for gravitation tracking in mobile systems," in *International Conference on Indoor Positioning and Navigation*, 2015.

[26] J. de Wiljes, S. Reich, and W. Stannat, "Long-time stability and accuracy of the ensemble Kalman-Bucy filter for fully observed processes and small measurement noise," 2017, preprint, arXiv:1612.06065v3.

[27] H. T. Davis, *Introduction to nonlinear differential and integral equations*. Courier Corporation, 1962.

[28] R. E. Kalman, "Contributions to the theory of optimal control," *Boletín de la Sociedad Matemática Mexicana*, vol. 5, no. 2, pp. 102–119, 1960.

[29] W. J. Terrell, *Stability and Stabilization: An Introduction*. Princeton University Press, 2009.