Fourier–Hermite series for stochastic stability analysis of non-linear Kalman filters

Toni Karvonen and Simo Särkkä

This article has been published in *Proceedings of 19th International Conference on Information Fusion*, held in Heidelberg, Germany, on July 5–8, 2016. When citing this article, use the following reference:

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

Fourier–Hermite Series for Stochastic Stability Analysis of Non-Linear Kalman Filters

Toni Karvonen
Dept. of Electrical Engineering and Automation
Aalto University, Espoo, Finland
Email: toni.karvonen@aalto.fi

Simo Särkkä
Dept. of Electrical Engineering and Automation
Aalto University, Espoo, Finland
Email: simo.sarkka@aalto.fi

***Abstract*—Stochastic stability results for the extended Kalman filter and some other non-linear filters have been available for some time now. In this context stochastic stability refers to mean square boundedness of the estimation error. In this article we use Fourier–Hermite series expansion to derive novel stability results for general discrete-time non-linear Kalman filters that can be interpreted as numerical integration rules of Gaussian integrals arising from moment-matching. We also provide an upper bound for the Kalman gain matrix that is not explicitly dependent on the measurement model Jacobian, eliminating thus the need to assume boundedness of this Jacobian. Furthermore, we formulate the system non-linearity assumptions so that it is possible to verify them when the model functions are Lipschitz continuous. We use these results for *a priori* assessment of the stability of a univariate non-linear filter and verify the results numerically.**

I. Introduction

It has been long known that the celebrated Kalman filter possesses strong exponential stability properties when the underlying dynamic system is linear (see [1] and [2]). However, stability properties of different non-linear extensions of the Kalman filter have begun to receive some attention only during the past 15 years. The first results of some generality were derived by Reif *et al.* [3] for the discrete-time extended Kalman filter (EKF) and later somewhat generalised by Kluge *et al.* [4]. These results provide mean square estimation error bounds with regrettably restrictive assumptions that are rarely verifiable. A major drawback is also that in practice stability can be assessed only after the filter has been run, not beforehand as would be desirable. The proofs are based on the use of a simple stochastic Lyapunov type lemma, frequently termed *stochastic stability lemma* (see Lemma 3), the very nature of which seems to be somewhat unsuitable for obtaining strong results [5].

With the help of residual-correcting random diagonal matrices [6], the same approach was shown to be applicable to the unscented Kalman filter (UKF) with linear measurement model by Xiong *et al.* [7] and later extended to systems with non-linear measurement model as well as a wider class of certain non-linear Kalman filters [8]–[11]. One problem with this approach is that instead of quantitative estimation error bounds such as those obtained for the EKF, only qualitative results about the effect of noise covariance matrix tuning [12] can be obtained. UKF stability has also been studied with interesting contraction theoretic methods by Maree *et al.* [13].

With minor modifications, the stability results have been extended for a number of different filters similar to the EKF or the UKF, see for example [14] and [15]. The continuous-time case has also been investigated (see [16] and [17]) with methods analogous to those used in the discrete-time case. Results have been applied to a few problems [18]–[20] but such applications are challenging presently.

In this article we use Fourier–Hermite series expansion to analyse non-linear Kalman filter stability. We use this expansion, coupled with the interpretation of non-linear Kalman filters as numerical integration rules for Gaussian integrals [21], to derive stability results for a wide class of non-linear filters. Our other contributions include elimination of the unintuitive assumption of boundedness of the measurement model Jacobian. We also give a practical example of rigorous *a priori* stability verification of a non-linear Kalman filter. We control the non-linearities of the system slightly differently from how they have been controlled before. This imposes some new conditions (while removing others) on the non-linearities but makes it easier to verify that the systems considered satisfy these conditions. We find the Fourier–Hermite expansion based results of this article more powerful, applicable, quantititative and more intuitive in derivation than the Taylor series and random matrix approach in [8] and [9].

Throughout this article $\|\cdot\|$ denotes the usual Euclidean norm of vectors or the spectral norm of matrices and $\|\cdot\|_2$ the L^2 norm of random variables. Expectation is denoted by $\mathbb{E}$ and covariance matrix by Cov. For symmetric square matrices $\mathbf{A} > \mathbf{B}$ ($\mathbf{A} \geq \mathbf{B}$) means that $\mathbf{A} - \mathbf{B}$ is positive-definite (positive-semidefinite). This partial ordering of symmetric matrices is known as *Löwner ordering* and it exhibits some, but not all, of the familiar properties of the ordering of real numbers (see [22, Chapter 8]). The identity matrix is denoted by $\mathbf{I}$. The element of a matrix $\mathbf{A}$ on ith row and jth column is denoted by $A_{i,j}$ or $[\mathbf{A}]_{i,j}$. Similarly, the ith component of a vector $\mathbf{x}$ is x_i and that of a vector-valued function $\mathbf{f}$ is f_i.

The article is structured as follows. In Sections II, III and IV we provide the necessary preliminaries for stability analysis and introduce our notation and terminology of non-linear Kalman filters. Sections V and VI include proofs of stability. A rigorous univariate example is given in Section VII. Finally, conclusions, with some discussion, are drawn in Section VIII.

II. Fourier–Hermite Series Expansion

This is a very brief summary of properties of multidimensional Hermite polynomials and Fourier–Hermite series. For a fuller and more detailed treatment, see for example [23]–[25].

Multidimensional Hermite polynomials, orthogonal with respect to $\mathcal{N}(\mathbf{0}, \mathbf{I})$, are defined as

$$\mathrm{H}_{[i_1,\ldots,i_k]}(\mathbf{x}) = (-1)^k \, \mathrm{e}^{\|\mathbf{x}\|^2/2} \, \frac{\partial^k}{\partial x_{i_1} \cdots \partial x_{i_k}} \, \mathrm{e}^{-\|\mathbf{x}\|^2/2}$$

for $\mathbf{x} \in \mathbb{R}^p$. For notational simplicity in forming a Fourier–Hermite series with respect to any Gaussian distribution $\mathcal{N}(\mathbf{m}, \mathbf{P})$, we denote scaled versions of Hermite polynomials by $\mathrm{H}_{[i_1,\ldots,i_k]}(\mathbf{x}; \mathbf{m}, \mathbf{P}) \coloneqq \mathrm{H}_{[i_1,\ldots,i_k]}[\mathbf{L}^{-1}(\mathbf{x} - \mathbf{m})]$, with $\mathbf{L} = \sqrt{\mathbf{P}}$. Then, any function $\mathbf{g} \colon \mathbb{R}^p \to \mathbb{R}^q$, square-integrable with respect to $\mathcal{N}(\mathbf{m}, \mathbf{P})$, can be expressed as a *Fourier–Hermite series*

$$\begin{aligned} \mathbf{g}(\mathbf{x}) &= \sum_{k=0}^{\infty} \frac{1}{k!} \sum_{\substack{i_1,\ldots,i_k=1 \\ j_1,\ldots,j_k=1}}^{p} \mathbb{E}\left(\frac{\partial^k \mathbf{g}(\mathbf{x})}{\partial x_{j_1} \cdots \partial x_{j_k}} \right) \\ &\qquad \times \prod_{m=1}^{k} L_{j_m, i_m} \, \mathrm{H}_{[i_1,\ldots,i_k]}(\mathbf{x}; \mathbf{m}, \mathbf{P}) \\ &= \mathbb{E}[\mathbf{g}(\mathbf{x})] + \mathbb{E}[\mathbf{J}_{\mathbf{g}}(\mathbf{x})](\mathbf{x} - \mathbf{m}) + \boldsymbol{\omega}_2\big[\mathbf{g}(\mathbf{x}), \mathbf{m}, \mathbf{P}\big], \end{aligned} \tag{1}$$

where $\mathbf{x} \sim \mathcal{N}(\mathbf{m}, \mathbf{P})$, $\mathbf{J}_{\mathbf{g}}(\mathbf{x})$ is the Jacobian matrix of $\mathbf{g}$ at $\mathbf{x}$ and $\boldsymbol{\omega}_2\big[\mathbf{g}(\mathbf{x}), \mathbf{m}, \mathbf{P}\big]$ is the second order remainder term of this Fourier–Hermite series expansion. Furthermore, by orthogonality and Parseval's identity we have the simple covariance equation

$$\begin{aligned} \mathrm{Cov}[\mathbf{g}(\mathbf{x})] &= \sum_{k=1}^{\infty} \frac{1}{k!} \sum_{\substack{i_1,\ldots,i_k=1 \\ j_1,\ldots,j_k=1}}^{p} \mathbb{E}\left(\frac{\partial^k \mathbf{g}(\mathbf{x})}{\partial x_{j_1} \cdots \partial x_{j_k}} \right) \\ &\qquad \times \prod_{m=1}^{k} P_{j_m, i_m} \mathbb{E}\left(\frac{\partial^k \mathbf{g}(\mathbf{x})}{\partial x_{i_1} \cdots \partial x_{i_k}} \right)^{\mathsf{T}}. \end{aligned}$$

III. Non-linear Kalman Filters

In this article we consider discrete-time stochastic dynamic systems of the form

$$\begin{aligned} \mathbf{x}_k &= \mathbf{f}(\mathbf{x}_{k-1}) + \mathbf{q}_{k-1}, \\ \mathbf{y}_k &= \mathbf{h}(\mathbf{x}_k) + \mathbf{r}_k, \end{aligned} \tag{2}$$

where $\mathbf{x}_k \in \mathbb{R}^n$ is the *state* of the system and $\mathbf{y}_k \in \mathbb{R}^m$ the *measurement*. The function $\mathbf{f} \colon \mathbb{R}^n \to \mathbb{R}^n$ is the *dynamic model function* and $\mathbf{h} \colon \mathbb{R}^n \to \mathbb{R}^m$ the *measurement model function*, both assumed differentiable. The noise processes are distributed as $\mathbf{q}_{k-1} \sim \mathcal{N}(\mathbf{0}, \mathbf{Q}_{k-1})$ and $\mathbf{r}_k \sim \mathcal{N}(\mathbf{0}, \mathbf{R}_k)$ with the noise covariance matrices $\mathbf{Q}_{k-1}$ and $\mathbf{R}_k$ positive-definite. It is assumed that the noise processes are uncorrelated and independent of the initial state $\mathbf{x}_0 \sim p(\mathbf{x}_0)$.

The system admits the optimal Bayesian filter, the *predictive distribution* and *filtering distribution* given by the *Bayesian filtering equations* [26, Theorem 4.1]. In general those equations are intractable and approximative schemes must be used. A natural way to obtain such an approximation is to assume that the predictive and filtering distribution are Gaussian and employ moment-matching. This approach yields the *exact non-linear Kalman filter* [26, Chapter 6] of Algorithm 1.

Henceforth expectations with respect to $\mathcal{N}(\mathbf{m}_k, \mathbf{P}_k)$ and $\mathcal{N}(\mathbf{m}_k^-, \mathbf{P}_k^-)$ are denoted by $\mathbb{E}_k$ and $\mathbb{E}_k^-$, respectively. Analogous notation is adopted for covariance matrices. The following notation is employed throughout this article:

$$\begin{aligned} \mathbf{F}_k &\coloneqq \mathbb{E}_k[\mathbf{J}_{\mathbf{f}}(\mathbf{x})], \\ \mathbf{H}_k &\coloneqq \mathbb{E}_k^-[\mathbf{J}_{\mathbf{h}}(\mathbf{x})], \\ \boldsymbol{\Omega}_k^{\mathbf{f}} &\coloneqq \mathbb{E}_k \left(\boldsymbol{\omega}_2\big[\mathbf{f}(\mathbf{x}), \mathbf{m}_k, \mathbf{P}_k\big] \boldsymbol{\omega}_2\big[\mathbf{f}(\mathbf{x}), \mathbf{m}_k, \mathbf{P}_k\big]^{\mathsf{T}} \right), \\ \boldsymbol{\Omega}_k^{\mathbf{h}} &\coloneqq \mathbb{E}_k^- \left(\boldsymbol{\omega}_2\big[\mathbf{h}(\mathbf{x}), \mathbf{m}_k^-, \mathbf{P}_k^-\big] \boldsymbol{\omega}_2\big[\mathbf{h}(\mathbf{x}), \mathbf{m}_k^-, \mathbf{P}_k^-\big]^{\mathsf{T}} \right). \end{aligned}$$

With this notation

$$\begin{aligned} \mathrm{Cov}_k[\mathbf{f}(\mathbf{x})] &= \mathbf{F}_k \mathbf{P}_k \mathbf{F}_k^{\mathsf{T}} + \boldsymbol{\Omega}_k^{\mathbf{f}}, \\ \mathrm{Cov}_k^-[\mathbf{h}(\mathbf{x})] &= \mathbf{H}_k \mathbf{P}_k^- \mathbf{H}_k^{\mathsf{T}} + \boldsymbol{\Omega}_k^{\mathbf{h}}, \end{aligned} \tag{3}$$

where it is to be noted that $\boldsymbol{\Omega}_k^{\mathbf{f}}$ and $\boldsymbol{\Omega}_k^{\mathbf{h}}$ are positive-semidefinite.

In the following filtering algorithm $\widehat{\mathbf{Q}}_{k-1}$ and $\widehat{\mathbf{R}}_k$ are some positive-definite matrices that need not equal the noise covariance matrices $\mathbf{Q}_{k-1}$ and $\mathbf{R}_k$. Appropriately selecting $\widehat{\mathbf{Q}}_{k-1}$ and $\widehat{\mathbf{R}}_k$ is referred to as *tuning* [12]. Magnifying these matrices often leads to improved stability properties with the drawback of degraded estimation accuracy [7]. Unfortunately, a thorough discussion of this tuning is out of the scope of this article. With the help the Fourier–Hermite series the exact non-linear Kalman filter algorithm can be written in the following form.

Algorithm 1 (Exact non-linear Kalman filter)**.** *The exact non-linear Kalman filter for the non-linear system* (2) *approximates predictive distributions and filtering distributions with Gaussian distributions* $\mathcal{N}(\mathbf{m}_k^-, \mathbf{P}_k^-)$ *and* $\mathcal{N}(\mathbf{m}_k, \mathbf{P}_k)$*, respectively. The parameters of these distributions are computed recursively by the* prediction step

$$\begin{aligned} \mathbf{m}_k^- &= \mathbb{E}_{k-1}[\mathbf{f}(\mathbf{x})], \\ \mathbf{P}_k^- &= \mathrm{Cov}_{k-1}[\mathbf{f}(\mathbf{x})] + \widehat{\mathbf{Q}}_{k-1} \\ &= \mathbf{F}_{k-1} \mathbf{P}_{k-1} \mathbf{F}_{k-1}^{\mathsf{T}} + \boldsymbol{\Omega}_{k-1}^{\mathbf{f}} + \widehat{\mathbf{Q}}_{k-1} \\ &\eqqcolon \mathbf{F}_{k-1} \mathbf{P}_{k-1} \mathbf{F}_{k-1}^{\mathsf{T}} + \widehat{\mathbf{Q}}'_{k-1} \end{aligned}$$

and the update step

$$\begin{aligned} \mathbf{S}_k &= \mathrm{Cov}_k^-[\mathbf{h}(\mathbf{x})] + \widehat{\mathbf{R}}_k \\ &= \mathbf{H}_k \mathbf{P}_k^- \mathbf{H}_k^{\mathsf{T}} + \boldsymbol{\Omega}_k^{\mathbf{h}} + \widehat{\mathbf{R}}_k \\ &\eqqcolon \mathbf{H}_k \mathbf{P}_k^- \mathbf{H}_k^{\mathsf{T}} + \widehat{\mathbf{R}}'_k, \\ \mathbf{K}_k &= \mathrm{Cov}_k^-[\mathbf{x}, \mathbf{h}(\mathbf{x})] \mathbf{S}_k^{-1} \\ &= \mathbf{P}_k^- \mathbf{H}_k^{\mathsf{T}} \mathbf{S}_k^{-1}, \\ \mathbf{m}_k &= \mathbf{m}_k^- + \mathbf{K}_k \big(\mathbf{y}_k - \mathbb{E}_k^-[\mathbf{h}(\mathbf{x})] \big), \\ \mathbf{P}_k &= \mathbf{P}_k^- - \mathbf{K}_k \mathbf{S}_k \mathbf{K}_k^{\mathsf{T}}. \end{aligned}$$

The recursion is started from the initial distribution $\mathcal{N}(\mathbf{m}_0, \mathbf{P}_0)$.

The equality $\mathbf{K}_k = \mathbf{P}_k^- \mathbf{H}_k^{\mathsf{T}} \mathbf{S}_k^{-1}$ for the Kalman gain matrix follows from the well-known Stein's identity, see for example [27]. Although $\mathbf{P}_k^-$ and $\mathbf{P}_k$ are not the real predicted error covariance and error covariance, respectively, they will be called such in an analogy to the linear filtering problem.

Because $\mathbb{E}_k[\mathbf{f}(\mathbf{x})]$, $\mathbb{E}_k[\mathbf{h}(\mathbf{x})]$ and (3) cannot be in most cases computed analytically, further approximations are necessary. The class of filters employing different approximations to these Gaussian integrals is variedly known as the class of non-linear Kalman filters, *Gaussian filters* [28] or *local filters* [29]. This class includes, for example, the sigma-point based UKF, Gauss–Hermite Kalman filter based on classical Gauss–Hermite quadrature and the Fourier–Hermite Kalman filter [25] that merely truncates $\mathbf{\Omega}_k^{\mathbf{f}}$ and $\mathbf{\Omega}_k^{\mathbf{h}}$.[1] In these algorithms the expectations and covariances are replaced by their numerical approximations, indicated in this article by a tilde (e.g. $\widetilde{\mathbb{E}}_k^-[\mathbf{h}(\mathbf{x})]$ and $\widetilde{\mathrm{Cov}}_{k-1}[\mathbf{f}(\mathbf{x})]$). Of course, for these approximate filters the Fourier–Hermite series or Stein's identity cannot be used to obtain expression for covariance matrices and Kalman gain reminiscent of those of the linear Kalman filter as in Algorithm 1.

IV. Stochastic Stability

Stability of a non-linear filter in this article is taken to mean boundedness of mean square estimation error.

Definition 2. A stochastic process $\boldsymbol{\xi}_k \in \mathbb{R}^n$ for $k \geq 0$ is said to be *bounded in mean square* if there is a non-negative scalar M such that $\|\boldsymbol{\xi}_k\|_2 \leq M$ for all k.

The following lemma is usually called *stochastic stability lemma*. Our version is an amalgamation of the versions in [31, Theorem 2] and [32, Satz IX.9]. In fact, in our formulation the lemma has nothing to do with stochastics and the stochastic processes involved could be, with no further modifications, replaced with sequences of real numbers.

Lemma 3 (Stochastic stability lemma)**.** *Let $\boldsymbol{\xi}_k \in \mathbb{R}^n$, for $k \geq 0$, be a discrete-time stochastic process. Suppose there is a scalar-valued stochastic process V_k, positive scalars v_1, v_2, μ and ε, and $0 < \alpha \leq 1$ such that*

(A1) The inequalities

$$v_1 \|\boldsymbol{\xi}_k\|_2^2 \leq \mathbb{E}(V_k) \leq v_2 \|\boldsymbol{\xi}_k\|_2^2$$

hold for all $k \geq 0$.

(A2) The bounds

$$\|\boldsymbol{\xi}_0\|_2^2 \leq \frac{v_1\varepsilon^2}{2v_2} \quad \textit{and} \quad \mu \leq \frac{\alpha v_1 \varepsilon^2}{2} \quad \textit{hold.}$$

(A3) The inequality

$$\mathbb{E}(V_{k+1}) \leq (1-\alpha)\mathbb{E}(V_k) + \mu \tag{4}$$

holds if $\|\boldsymbol{\xi}_k\|_2 \leq \varepsilon$.[2]

Then $\boldsymbol{\xi}_k$ is bounded in mean square by ε.

Proof: The proof is by induction. Since $\|\boldsymbol{\xi}_0\|_2 \leq \varepsilon$ by Assumptions (A1) and (A2), we have

$$\begin{aligned} \|\boldsymbol{\xi}_1\|_2^2 &\leq \frac{\mathbb{E}(V_1)}{v_1} \leq (1-\alpha)\frac{\mathbb{E}(V_0)}{v_1} + \frac{\mu}{v_1} \\ &\leq (1-\alpha)\frac{\varepsilon^2}{2} + \frac{\alpha\varepsilon^2}{2} \leq \varepsilon^2. \end{aligned}$$

Suppose then that $\|\boldsymbol{\xi}_i\|_2 \leq \varepsilon$ for $0 \leq i \leq k-1$. Then, by (4) and properties of geometric sum,

$$\begin{aligned} \|\boldsymbol{\xi}_k\|_2^2 &\leq (1-\alpha)^k \frac{\mathbb{E}(V_0)}{v_1} + \frac{\mu}{v_1}\sum_{i=0}^{k-1}(1-\alpha)^i \\ &\leq (1-\alpha)\frac{\varepsilon^2}{2} + \frac{\mu}{v_1\alpha} \leq \varepsilon^2, \end{aligned}$$

and therefore the claim holds. ■

It is often claimed that also almost sure boundedness (i.e. $\sup_{k\geq 0}\|\boldsymbol{\xi}_k\| < \infty$ almost surely) follows from the assumptions of this lemma (see e.g. [3] and [4]) on the basis of [33, Section 4.1, Theorem 1] when (4) is replaced with an analogous inequality involving conditional expectations. However, as has already been noted [34, Lemma 8], this claim is not strictly true.

V. Stability of the Exact Non-Linear Kalman Filter

In this article the stability of non-linear Kalman filters is studied using the Fourier–Hermite series expansion of $\mathbf{f}$ and $\mathbf{h}$. Based on the expansion (1),

$$\begin{aligned} \mathbf{f}(\mathbf{x}_k) - \mathbb{E}_k[\mathbf{f}(\mathbf{x})] &= \mathbf{F}_k(\mathbf{x}_k - \mathbf{m}_k) + \boldsymbol{\omega}_2\big[\mathbf{f}(\mathbf{x}_k), \mathbf{m}_k, \mathbf{P}_k\big], \\ \mathbf{h}(\mathbf{x}_k) - \mathbb{E}_k^-[\mathbf{h}(\mathbf{x})] &= \mathbf{H}_k(\mathbf{x}_k - \mathbf{m}_k^-) + \boldsymbol{\omega}_2\big[\mathbf{h}(\mathbf{x}_k), \mathbf{m}_k^-, \mathbf{P}_k^-\big]. \end{aligned}$$

Using the non-linear Kalman filter equations of Algorithm 1, the *predicted estimation error* $\boldsymbol{\xi}_k^- := \mathbf{x}_k - \mathbf{m}_k^-$ can be then written recursively as

$$\boldsymbol{\xi}_{k+1}^- = \mathbf{F}_k\mathbf{A}_k\boldsymbol{\xi}_k^- + \boldsymbol{\rho}_k + \boldsymbol{\sigma}_k, \tag{5}$$

and the relation between this and $\boldsymbol{\xi}_k := \mathbf{x}_k - \mathbf{m}_k$, the *estimation error*, is

$$\boldsymbol{\xi}_k = \mathbf{A}_k\boldsymbol{\xi}_k^- - \mathbf{K}_k\mathbf{r}_k - \mathbf{K}_k\boldsymbol{\omega}_2\big[\mathbf{h}(\mathbf{x}_k), \mathbf{m}_k^-, \mathbf{P}_k^-\big], \tag{6}$$

where

$$\begin{aligned} \mathbf{A}_k &= \mathbf{I} - \mathbf{K}_k\mathbf{H}_k, \\ \boldsymbol{\rho}_k &= \boldsymbol{\omega}_2[\mathbf{f}(\mathbf{x}_k), \mathbf{m}_k, \mathbf{P}_k] - \mathbf{F}_k\mathbf{K}_k\boldsymbol{\omega}_2\big[\mathbf{h}(\mathbf{x}_k), \mathbf{m}_k^-, \mathbf{P}_k^-\big], \\ \boldsymbol{\sigma}_k &= \mathbf{q}_k - \mathbf{F}_k\mathbf{K}_k\mathbf{r}_k. \end{aligned}$$

For convenience, the notation $\boldsymbol{\varphi}_k := \boldsymbol{\omega}_2\big[\mathbf{f}(\mathbf{x}_k), \mathbf{m}_k, \mathbf{P}_k\big]$ and $\boldsymbol{\chi}_k := \boldsymbol{\omega}_2\big[\mathbf{h}(\mathbf{x}_k), \mathbf{m}_k^-, \mathbf{P}_k^-\big]$ is used for the remainder terms of

[1] Note that a strict interpretation of our definition does not include, for example, the EKF although it is of the exactly same form because it is based on approximating the non-linear functions $\mathbf{f}$ and $\mathbf{h}$ themselves, not the means and covariances. However, the proof of Theorem 4, our main result, works also for the EKF. In fact, the EKF can be interpreted as a sigma-point filter that uses only one point [30].

[2] This may be a delicate point. We are not making a circular argument of proving $\|\boldsymbol{\xi}_k\|_2 \leq \varepsilon$ by assuming the same thing. What we assume here is that we know *a priori* that the inequality (4) holds only for those k for which $\|\boldsymbol{\xi}_k\|_2 \leq \varepsilon$. Combining this with the other assumptions then leads to the conclusion that $\|\boldsymbol{\xi}_k\|_2 \leq \varepsilon$ for all $k \geq 0$.

Fourier–Hermite series and $\mathbf{\Pi}_k := (\mathbf{P}_k^-)^{-1}$ for the inverse of the predicted error covariance matrix.

Theorem 4. *Consider the non-linear dynamic system* (2) *and the exact non-linear Kalman filter of Algorithm 1. Suppose that the following conditions hold:*

(A1) There exist positive scalars f, $\hat{q}'$, $\hat{r}'$, p_1^-, p_2^- and p_2 such that

$$\begin{aligned} \|\mathbf{F}_{k-1}\| &\le f, \\ \hat{q}'\mathbf{I} \le \widehat{\mathbf{Q}}'_{k-1}, &\quad \hat{r}'\mathbf{I} \le \widehat{\mathbf{R}}'_k, \\ p_1^-\mathbf{I} \le \mathbf{P}_k^- \le p_2^-\mathbf{I}, &\quad \mathbf{P}_{k-1} \le p_2\mathbf{I} \end{aligned}$$

for all $k \ge 1$.

(A2) There exist non-negative scalars $\kappa_{\boldsymbol{\varphi}}$, $\kappa_{\boldsymbol{\varphi}}^+$, $\kappa_{\boldsymbol{\chi}}$ and $\kappa_{\boldsymbol{\chi}}^+$ such that

$$\begin{aligned} \|\boldsymbol{\varphi}_{k-1}\|_2 &\le \kappa_{\boldsymbol{\varphi}} \|\mathbf{x}_{k-1} - \mathbf{m}_{k-1}\|_2 + \kappa_{\boldsymbol{\varphi}}^+, \\ \|\boldsymbol{\chi}_k\|_2 &\le \kappa_{\boldsymbol{\chi}} \|\mathbf{x}_k - \mathbf{m}_k^-\|_2 + \kappa_{\boldsymbol{\chi}}^+ \end{aligned}$$

for all $k \ge 1$.

Then, given any $\varepsilon \ge 0$ for which $\|\boldsymbol{\xi}_1^-\|_2 \le \varepsilon$, there exists $\delta \ge 0$ such that the conditions $\mathbf{Q}_k, \mathbf{R}_k \le \delta^2\mathbf{I}$ guarantee that the predicted estimation error $\boldsymbol{\xi}_k^-$ is bounded in mean square if $\kappa_{\boldsymbol{\varphi}}$, $\kappa_{\boldsymbol{\varphi}}^+$, $\kappa_{\boldsymbol{\chi}}$ and $\kappa_{\boldsymbol{\chi}}^+$ are sufficiently small.

Remarks:

(1) The assumption $\mathbf{P}_k \le p_2\mathbf{I}$ is only for obtaining less conservative bounds since $\mathbf{P}_k \le \mathbf{P}_k^-$ anyway. Similarly, $\mathbf{P}_k^-$ has the trivial lower bound $\mathbf{P}_k^- \ge \hat{q}'\mathbf{I}$ which is non-optimal unless f is small compared to $\hat{q}'$.

(2) In practice, the lower bounds for $\widehat{\mathbf{Q}}'_{k-1}$ and $\widehat{\mathbf{R}}'_k$ are those for the tuning matrices $\widehat{\mathbf{Q}}_{k-1}$ and $\widehat{\mathbf{R}}_k$.

(3) This theorem could be easily extended for systems with non-additive noise and intermittent observations as has been done for the EKF [4].

(4) With minor changes the proof is applicable to the EKF as well.

(5) Assumption (A2) is satisfied by Lipschitz functions, for if $\mathbf{g}$ is such a function and $L = \mathrm{Lip}(\mathbf{g})$ its Lipschitz constant, then it can be easily shown that

$$\left\|\boldsymbol{\omega}_2\big[\mathbf{g}(\mathbf{x}), \mathbf{m}, \mathbf{P}\big]\right\|_2 \le 2L\,\|\mathbf{x} - \mathbf{m}\|_2 + L\,\|\mathbf{P}\|.$$

(6) Assumption (A2) is usually given in a form that contains an exponent $1 < \gamma \le 2$ on the right-hand side (as well as lacking the constant term) instead of $\gamma = 1$ here. Such an assumption provides stronger stability results but we have not found a way to actually prove that there exist non-linear functions for which this assumption holds.

(7) Because the matrices in Assumption (A1) are random, some L^p boundedness condition combined with Hölder's inequality could be used in the proofs to follow.

(8) An upper bound for $\mathbf{A}_k$ is needed. A bound independent of $\mathbf{H}_k$ can be obtained by noting that $\mathbf{P}_k = \mathbf{A}_k\mathbf{P}_k^-$ and so $\|\mathbf{A}_k\| \le p_2/p_1^- := A$.

The proof utilises a few lemmas. Lemma 5 provides an upper bound for $\mathbf{K}_k$ independent of $\mathbf{H}_k$, a bound we have been unable to find in the literature (see [35] for a bound involving the ratio of singular values of $\mathbf{H}_k$). Lemma 6 follows an analogous lemma in [4] and Lemmas 7 and 8 are based on the ones in [3].

Lemma 5. *Under the assumptions of Theorem 4, for all $k \ge 1$*

$$\|\mathbf{K}_k\| \le \frac{p_2^-}{\sqrt{p_1^-\hat{r}'}} := K. \tag{7}$$

Proof: It suffices to show that all eigenvalues of $\mathbf{K}_k^\mathsf{T}\mathbf{K}_k$ remain uniformly bounded. It can be seen that

$$\begin{aligned} \mathbf{K}_k^\mathsf{T}\mathbf{K}_k &= \left(\mathbf{H}_k\mathbf{P}_k^-\mathbf{H}_k^\mathsf{T} + \widehat{\mathbf{R}}'_k\right)^{-1}\mathbf{H}_k\mathbf{P}_k^- \\ &\quad \times \mathbf{P}_k^-\mathbf{H}_k^\mathsf{T}\left(\mathbf{H}_k\mathbf{P}_k^-\mathbf{H}_k^\mathsf{T} + \widehat{\mathbf{R}}'_k\right)^{-1} \\ &\le (p_2^-)^2\left(\mathbf{H}_k\mathbf{P}_k^-\mathbf{H}_k^\mathsf{T} + \widehat{\mathbf{R}}'_k\right)^{-1}\mathbf{H}_k \\ &\quad \times \mathbf{H}_k^\mathsf{T}\left(\mathbf{H}_k\mathbf{P}_k^-\mathbf{H}_k^\mathsf{T} + \widehat{\mathbf{R}}'_k\right)^{-1} \\ &\le \frac{(p_2^-)^2}{p_1^-}\left(\mathbf{H}_k\mathbf{P}_k^-\mathbf{H}_k^\mathsf{T} + \widehat{\mathbf{R}}'_k\right)^{-1}\left(\mathbf{H}_k\mathbf{P}_k^-\mathbf{H}_k^\mathsf{T} + \widehat{\mathbf{R}}'_k\right) \\ &\quad \times \left(\mathbf{H}_k\mathbf{P}_k^-\mathbf{H}_k^\mathsf{T} + \widehat{\mathbf{R}}'_k\right)^{-1} \\ &\le \frac{(p_2^-)^2}{p_1^-\hat{r}'}\mathbf{I}, \end{aligned}$$

which implies the claim. The second inequality follows because $p_1\mathbf{I} \le \mathbf{P}_k^-$ implies that $\mathbf{H}_k\mathbf{H}_k^\mathsf{T} \le \mathbf{H}_k\mathbf{P}_k^-\mathbf{H}_k^\mathsf{T}/p_1^-$. ∎

The bound is not probably the strictest possible as for scalars one has

$$\sup_{H_k \in \mathbb{R}}\left|\frac{P_k^- H_k}{H_k^2 P_k^- + \widehat{R}'_k}\right| = \frac{1}{2}\sqrt{\frac{P_k^-}{\widehat{R}'_k}}, \tag{8}$$

a bound we have found to be supported by numerical evidence also in higher dimensions.

Lemma 6. *Under the assumptions of Theorem 4 there exists $0 < \alpha' < 1$ such that*

$$\mathbf{A}_k^\mathsf{T}\mathbf{F}_k^\mathsf{T}\mathbf{\Pi}_{k+1}\mathbf{F}_k\mathbf{A}_k \le (1 - \alpha')\mathbf{\Pi}_k$$

for all $k \ge 1$. The constant α' is given by $\alpha' = \hat{q}'/(f^2p_2 + \hat{q}')$.

Proof: Let $a > 1$. By the definition of $\mathbf{P}_k^-$ and the assumptions of Theorem 4,

$$\mathbf{P}_{k+1}^- = \mathbf{F}_k\mathbf{P}_k\mathbf{F}_k^\mathsf{T} + \widehat{\mathbf{Q}}'_k > \left(1 + \frac{\hat{q}'}{af^2p_2}\right)\mathbf{F}_k\mathbf{P}_k\mathbf{F}_k^\mathsf{T}. \tag{9}$$

The Joseph form [36, p. 108]

$$\mathbf{P}_k = \mathbf{A}_k\mathbf{P}_k^-\mathbf{A}_k^\mathsf{T} + \mathbf{K}_k\widehat{\mathbf{R}}'_k\mathbf{K}_k^\mathsf{T}$$

and (9) yield

$$\mathbf{P}_{k+1}^- > \left(1 + \frac{\hat{q}'}{af^2p_2}\right)\mathbf{F}_k\mathbf{A}_k\mathbf{P}_k^-\mathbf{A}_k^\mathsf{T}\mathbf{F}_k^\mathsf{T}.$$

Then, [4, Lemma 6.1] implies that

$$\mathbf{A}_k^\mathsf{T}\mathbf{F}_k^\mathsf{T}\mathbf{\Pi}_{k+1}\mathbf{F}_k\mathbf{A}_k \le \frac{af^2p_2}{af^2p_2 + \hat{q}'}\mathbf{\Pi}_k.$$

Since this holds for every $a > 1$, the claim follows. ∎

Lemma 7. *Under the assumptions of Theorem 4 there exist non-negative scalars $\kappa_i^{\boldsymbol{\rho}}$ for $1 \le i \le 6$ such that*

$$\mathbb{E}\Big(\boldsymbol{\rho}_k^\mathsf{T}\boldsymbol{\Pi}_{k+1}\big[2\mathbf{F}_k\mathbf{A}_k\boldsymbol{\xi}_k^- + \boldsymbol{\rho}_k\big]\Big) \le \kappa_1^{\boldsymbol{\rho}}\|\boldsymbol{\xi}_k^-\|_2^2 + \kappa_2^{\boldsymbol{\rho}}\delta\|\boldsymbol{\xi}_k^-\|_2 + \kappa_3^{\boldsymbol{\rho}}\delta^2 + \kappa_4^{\boldsymbol{\rho}}\|\boldsymbol{\xi}_k^-\|_2 + \kappa_5^{\boldsymbol{\rho}}\delta + \kappa_6^{\boldsymbol{\rho}}$$

for all $k \ge 1$.

Proof: By (6) and the assumptions of Theorem 4,

$$\|\boldsymbol{\xi}_k\|_2 \le \big(A + K\kappa_{\boldsymbol{\chi}}\big)\|\boldsymbol{\xi}_k^-\|_2 + K\sqrt{m}\delta + K\kappa_{\boldsymbol{\chi}}^+, \tag{10}$$

and hence

$$\begin{aligned}\|\boldsymbol{\rho}_k\|_2 &\le \|\boldsymbol{\varphi}_k\|_2 + \|\mathbf{F}_k\mathbf{K}_k\boldsymbol{\chi}_k\|_2 \\ &\le \kappa_{\boldsymbol{\varphi}}\|\boldsymbol{\xi}_k\|_2 + Kf\kappa_{\boldsymbol{\chi}}\|\boldsymbol{\xi}_k^-\|_2 + \kappa_{\boldsymbol{\varphi}}^+ + Kf\kappa_{\boldsymbol{\chi}}^+ \\ &\le \big(A\kappa_{\boldsymbol{\varphi}} + K\kappa_{\boldsymbol{\chi}}\kappa_{\boldsymbol{\varphi}} + Kf\kappa_{\boldsymbol{\chi}}\big)\|\boldsymbol{\xi}_k^-\|_2 \\ &\quad + K\kappa_{\boldsymbol{\varphi}}\sqrt{m}\delta + \kappa_{\boldsymbol{\varphi}}^+ + (\kappa_{\boldsymbol{\varphi}} + f)K\kappa_{\boldsymbol{\chi}}^+ \\ &:= \kappa_1\|\boldsymbol{\xi}_k^-\|_2 + \kappa_2\delta + \kappa_+. \end{aligned}\tag{11}$$

Now,

$$\begin{aligned}&\mathbb{E}\Big(\boldsymbol{\rho}_k^\mathsf{T}\boldsymbol{\Pi}_{k+1}\boldsymbol{\rho}_k\Big) \\ &\le \frac{1}{p_1^-}\Big(\kappa_1^2\|\boldsymbol{\xi}_k^-\|_2^2 + \kappa_2^2\delta^2 + \kappa_+^2 + 2\kappa_1\kappa_2\delta\|\boldsymbol{\xi}_k^-\|_2 \\ &\quad + 2\kappa_1\kappa_+\|\boldsymbol{\xi}_k^-\|_2 + 2\kappa_2\kappa_+\delta\Big),\end{aligned}\tag{12}$$

and utilisation of the Cauchy–Schwarz inequality yields

$$\begin{aligned}&2\mathbb{E}\left(\boldsymbol{\rho}_k^\mathsf{T}\boldsymbol{\Pi}_{k+1}\mathbf{F}_k\mathbf{A}_k\boldsymbol{\xi}_k^-\right) \\ &\le \frac{2Af}{p_1^-}\|\boldsymbol{\rho}_k\|_2\|\boldsymbol{\xi}_k^-\|_2 \\ &\le \frac{2Af}{p_1^-}\left(\kappa_1\|\boldsymbol{\xi}_k^-\|_2^2 + \kappa_2\delta\|\boldsymbol{\xi}_k^-\|_2 + \kappa_+\|\boldsymbol{\xi}_k^-\|_2\right).\end{aligned}\tag{13}$$

By combining inequalities (12) and (13) we get the claim. ■

Lemma 8. *Under the assumptions of Theorem 4 there exists $\kappa^{\boldsymbol{\sigma}} \ge 0$ such that for all $k \ge 1$*

$$\mathbb{E}\left(\boldsymbol{\sigma}_k^\mathsf{T}\boldsymbol{\Pi}_{k+1}\boldsymbol{\sigma}_k\right) \le \kappa^{\boldsymbol{\sigma}}\delta^2.$$

Proof: Since the noise terms are uncorrelated, all cross-terms vanish in

$$\mathbb{E}\left(\boldsymbol{\sigma}_k^\mathsf{T}\boldsymbol{\Pi}_{k+1}\boldsymbol{\sigma}_k\right) = \mathbb{E}\left(\mathbf{q}_k^\mathsf{T}\boldsymbol{\Pi}_{k+1}\mathbf{q}_k\right) + \mathbb{E}\left(\mathbf{r}_k^\mathsf{T}\mathbf{K}_k^\mathsf{T}\mathbf{F}_k^\mathsf{T}\boldsymbol{\Pi}_{k+1}\mathbf{F}_k\mathbf{K}_k\mathbf{r}_k\right).$$

Therefore, because both sides are just scalars, trace operations produce the bound

$$\begin{aligned}\mathbb{E}\left(\boldsymbol{\sigma}_k^\mathsf{T}\boldsymbol{\Pi}_k\boldsymbol{\sigma}_k\right) &\le \frac{1}{p_1^-}\mathbb{E}\left(\operatorname{tr}\mathbf{q}_k^\mathsf{T}\mathbf{q}_k\right) + \frac{1}{\hat{r}'}\left(\frac{fp_2^-}{p_1^-}\right)^2\mathbb{E}\left(\operatorname{tr}\mathbf{r}_k^\mathsf{T}\mathbf{r}_k\right) \\ &= \frac{1}{p_1^-}\operatorname{tr}\mathbf{Q}_k + \frac{1}{\hat{r}'}\left(\frac{fp_2^-}{p_1^-}\right)^2\operatorname{tr}\mathbf{R}_k \\ &\le \frac{1}{p_1^-}\left(n + \frac{(fp_2^-)^2m}{\hat{r}'}\right)\delta^2.\end{aligned}$$

■

With these lemmas we are in a position to provide a proof of Theorem 4.

Proof of Theorem 4: The idea of the proof is to apply Lemma 3 to the process $\boldsymbol{\xi}_k^-$. Choose a stochastic Lyapunov function $V_k = (\boldsymbol{\xi}_k^-)^\mathsf{T}\boldsymbol{\Pi}_k\boldsymbol{\xi}_k^-$. Assumption (A1) of Lemma 3 is satisfied with $v_1 = 1/p_2^-$ and $v_2 = 1/p_1^-$.

Now, using the predicted estimation error recursion (5) and grouping the terms appropriately, one obtains

$$\begin{aligned}V_{k+1} = &(\boldsymbol{\xi}_k^-)^\mathsf{T}\mathbf{A}_k^\mathsf{T}\mathbf{F}_k^\mathsf{T}\boldsymbol{\Pi}_{k+1}\mathbf{F}_k\mathbf{A}_k\boldsymbol{\xi}_k^- \\ &+ \boldsymbol{\rho}_k^\mathsf{T}\boldsymbol{\Pi}_{k+1}\big(2\mathbf{F}_k\mathbf{A}_k\boldsymbol{\xi}_k^- + \boldsymbol{\rho}_k\big) + \boldsymbol{\sigma}_k^\mathsf{T}\boldsymbol{\Pi}_{k+1}\boldsymbol{\sigma}_k \\ &+ 2\boldsymbol{\sigma}_k^\mathsf{T}\boldsymbol{\Pi}_{k+1}\big(\mathbf{F}_k\mathbf{A}_k\boldsymbol{\xi}_k^- + \boldsymbol{\rho}_k\big).\end{aligned}\tag{14}$$

By independence, the last term on the right-hand side vanishes when expectations are taken. The three remaining terms are evaluated by Lemmas 6–8. For $k \ge 1$, this yields the inequality

$$\begin{aligned}\mathbb{E}(V_{k+1}) \le (1-\alpha')\mathbb{E}(V_k) &+ \kappa_1^{\boldsymbol{\rho}}\|\boldsymbol{\xi}_k^-\|_2^2 + \kappa_2^{\boldsymbol{\rho}}\delta\|\boldsymbol{\xi}_k^-\|_2 \\ &+ (\kappa_3^{\boldsymbol{\rho}} + \kappa^{\boldsymbol{\sigma}})\delta^2 + \kappa_4^{\boldsymbol{\rho}}\|\boldsymbol{\xi}_k^-\|_2 + \kappa_5^{\boldsymbol{\rho}}\delta + \kappa_6^{\boldsymbol{\rho}}.\end{aligned}\tag{15}$$

If $\kappa_1^{\boldsymbol{\rho}} \le \alpha'/p_2^-\beta$ with $\beta > 1$, then

$$\begin{aligned}\mathbb{E}(V_{k+1}) \le &\left(1 - \frac{\beta-1}{\beta}\alpha'\right)\mathbb{E}(V_k) + (\kappa_2^{\boldsymbol{\rho}}\tilde{\varepsilon} + \kappa_5^{\boldsymbol{\rho}})\delta \\ &+ (\kappa_3^{\boldsymbol{\rho}} + \kappa^{\boldsymbol{\sigma}})\delta^2 + \kappa_4^{\boldsymbol{\rho}}\tilde{\varepsilon} + \kappa_6^{\boldsymbol{\rho}}\end{aligned}$$

whenever $\|\boldsymbol{\xi}_k^-\|_2 \le \tilde{\varepsilon}$, with $\tilde{\varepsilon}$ any positive scalar. Thus Lemma 3 implies that the predicted estimation error is bounded in mean square by $\tilde{\varepsilon}$ if

$$\|\boldsymbol{\xi}_1^-\|_2^2 \le \frac{p_1^-\tilde{\varepsilon}^2}{2p_2^-}, \qquad \kappa_1^{\boldsymbol{\rho}} \le \frac{\alpha'}{p_2^-\beta}$$

and

$$(\kappa_2^{\boldsymbol{\rho}}\tilde{\varepsilon} + \kappa_5^{\boldsymbol{\rho}})\delta + (\kappa_3^{\boldsymbol{\rho}} + \kappa^{\boldsymbol{\sigma}})\delta^2 + \kappa_4^{\boldsymbol{\rho}}\tilde{\varepsilon} + \kappa_6^{\boldsymbol{\rho}} \le \frac{(\beta-1)\alpha'\tilde{\varepsilon}^2}{2p_2^-\beta}. \tag{16}$$

Because $\kappa_4^{\boldsymbol{\rho}}$ and $\kappa_6^{\boldsymbol{\rho}}$ do not depend on δ and $\kappa_4^{\boldsymbol{\rho}}, \kappa_6^{\boldsymbol{\rho}} \to 0$ as $\kappa_{\boldsymbol{\varphi}}^+, \kappa_{\boldsymbol{\chi}}^+ \to 0$, (16) may hold only if $\kappa_{\boldsymbol{\varphi}}^+$, $\kappa_{\boldsymbol{\chi}}^+$ and δ are sufficiently small.

Inspection of inequalities (12) and (13) and substitution of the values of A and K shows that the upper bound $\kappa_1^{\boldsymbol{\rho}} \le \alpha'/p_2^-\beta$ implies the inequality

$$\begin{aligned}&\frac{p_2^-}{p_1^-}\left(\frac{p_2\kappa_{\boldsymbol{\varphi}}}{p_1^-} + \frac{p_2^-\kappa_{\boldsymbol{\chi}}\kappa_{\boldsymbol{\varphi}}}{\sqrt{p_1^-\hat{r}'}} + \frac{p_2^-f\kappa_{\boldsymbol{\chi}}}{\sqrt{p_1^-\hat{r}'}}\right) \\ &\quad\times\left(\frac{p_2\kappa_{\boldsymbol{\varphi}}}{p_1^-} + \frac{p_2^-\kappa_{\boldsymbol{\chi}}\kappa_{\boldsymbol{\varphi}}}{\sqrt{p_1^-\hat{r}'}} + \frac{p_2^-f\kappa_{\boldsymbol{\chi}}}{\sqrt{p_1^-\hat{r}'}} + \frac{2p_2f}{p_1^-}\right) \\ &\le \frac{\alpha'}{\beta} < 1,\end{aligned}\tag{17}$$

which can only hold if $\kappa_{\boldsymbol{\varphi}}$ and $\kappa_{\boldsymbol{\chi}}$ are significantly smaller than one. ■

VI. Stability of Approximative Non-Linear Kalman Filters

As explained in Section III, an approximative non-linear Kalman filter uses numerical integration to compute the means and covariances if they cannot be evaluated analytically. This section sketches a stability proof for any such filter based on the one in the preceding section.

Consider any approximative non-linear Kalman filter. The predicted estimation error recursion (5) for such a filter can be written as

$$\boldsymbol{\xi}_{k+1}^{-} = \mathbf{F}_k \mathbf{A}_k \boldsymbol{\xi}_k^{-} + \boldsymbol{\rho}_k + \boldsymbol{\sigma}_k + \boldsymbol{\tau}_k\,, \tag{18}$$

where the additional term $\boldsymbol{\tau}_k$ incorporates the errors due to numerical integration:

$$\boldsymbol{\tau}_k = \mathbf{e}_k^{\mathbf{f}} - \mathbf{F}_k \mathbf{K}_k \mathbf{e}_k^{\mathbf{h}},$$

with $\mathbf{e}_k^{\mathbf{f}}$ and $\mathbf{e}_k^{\mathbf{h}}$ the integration errors

$$\begin{aligned}
\mathbf{e}_k^{\mathbf{f}} &= \mathbb{E}_k[\mathbf{f}(\mathbf{x})] - \widetilde{\mathbb{E}}_k[\mathbf{f}(\mathbf{x})], \\
\mathbf{e}_k^{\mathbf{h}} &= \mathbb{E}_k^{-}[\mathbf{h}(\mathbf{x})] - \widetilde{\mathbb{E}}_k^{-}[\mathbf{h}(\mathbf{x})].
\end{aligned}$$

Advantageous forms for covariance matrices can be obtained in a similar manner, namely

$$\begin{aligned}
\mathbf{P}_k^{-} &= \mathrm{Cov}_{k-1}[\mathbf{f}(\mathbf{x})] + \mathbf{C}_{k-1}^{\mathbf{f}} + \widehat{\mathbf{Q}}_{k-1} \\
&= \mathbf{F}_{k-1}\mathbf{P}_{k-1}\mathbf{F}_{k-1}^{\mathsf{T}} + \boldsymbol{\Omega}_{k-1}^{\mathbf{f}} + \widehat{\mathbf{Q}}_{k-1} + \mathbf{C}_{k-1}^{\mathbf{f}}, \\
\mathbf{S}_k &= \mathrm{Cov}_k^{-}[\mathbf{h}(\mathbf{x})] + \mathbf{C}_k^{\mathbf{h}} + \widehat{\mathbf{R}}_k \\
&= \mathbf{H}_k \mathbf{P}_k^{-} \mathbf{H}_k^{\mathsf{T}} + \boldsymbol{\Omega}_k^{\mathbf{h}} + \widehat{\mathbf{R}}_k + \mathbf{C}_k^{\mathbf{h}},
\end{aligned}$$

where

$$\begin{aligned}
\mathbf{C}_{k-1}^{\mathbf{f}} &= \widetilde{\mathrm{Cov}}_{k-1}[\mathbf{f}(\mathbf{x})] - \mathrm{Cov}_{k-1}[\mathbf{f}(\mathbf{x})], \\
\mathbf{C}_k^{\mathbf{h}} &= \widetilde{\mathrm{Cov}}_k^{-}[\mathbf{h}(\mathbf{x})] - \mathrm{Cov}_k^{-}[\mathbf{h}(\mathbf{x})].
\end{aligned}$$

It cannot be guaranteed that $\mathbf{C}_{k-1}^{\mathbf{f}}$ and $\mathbf{C}_k^{\mathbf{h}}$ are positive-semidefinite. However, positive-definiteness of $\boldsymbol{\Omega}_{k-1}^{\mathbf{f}} + \widehat{\mathbf{Q}}_{k-1} + \mathbf{C}_{k-1}^{\mathbf{f}}$ and $\boldsymbol{\Omega}_k^{\mathbf{h}} + \widehat{\mathbf{R}}_k + \mathbf{C}_k^{\mathbf{h}}$, necessary for carrying out the stability proof, can be guaranteed if $\widehat{\mathbf{Q}}_{k-1}$ and $\widehat{\mathbf{R}}_k$ are chosen large enough. In practice this means that $\mathbf{C}_{k-1}^{\mathbf{f}}$ and $\mathbf{C}_k^{\mathbf{h}}$ need to remain bounded.

Similar reasoning cannot be applied to the approximated cross-covariance $\widetilde{\mathrm{Cov}}_k^{-}[\mathbf{x}, \mathbf{h}(\mathbf{x})]$ but this would be in fact unnecessary and it suffices to assume this matrix uniformly bounded. However, since $\widetilde{\mathrm{Cov}}_k^{-}[\mathbf{x}, \mathbf{h}(\mathbf{x})] = \mathbf{P}_k^{-}\mathbf{H}_k^{\mathsf{T}}$ does not hold generally for approximative non-linear Kalman filters, it cannot be concluded that $\mathbf{P}_k = \mathbf{A}_k \mathbf{P}_k^{-}$ in this case. Consequently, the bound $\|\mathbf{A}_k\| \leq 1 + \|\mathbf{K}_k \mathbf{H}_k\|$ has to be used. This requires the unfortunate assumption of bounded $\mathbf{H}_k$.

With this notation, we have the following more general theorem on stability of non-linear Kalman filters.

Theorem 9. *Consider the non-linear dynamic system* (2) *and any approximative non-linear Kalman filter as formulated in Section III. Suppose that the following conditions hold:*

(A1) There exist positive scalars f, h, p_1^{-}, p_2^{-}, p_2, g, $\hat{q}_{\mathbf{C}}$ and $\hat{r}_{\mathbf{C}}$ such that

$$\begin{aligned}
&\|\mathbf{F}_{k-1}\| \leq f, && \|\mathbf{H}_k\| \leq h, \\
&p_1^{-}\mathbf{I} \leq \mathbf{P}_k^{-} \leq p_2^{-}\mathbf{I}\,, && \mathbf{P}_{k-1} \leq p_2 \mathbf{I}, \\
&\big\|\widetilde{\mathrm{Cov}}_k^{-}[\mathbf{x}, \mathbf{h}(\mathbf{x})]\big\| \leq g, \\
&\hat{q}_{\mathbf{C}}\mathbf{I} \leq \boldsymbol{\Omega}_{k-1}^{\mathbf{f}} + \widehat{\mathbf{Q}}_{k-1} + \mathbf{C}_{k-1}^{\mathbf{f}}, \\
&\hat{r}_{\mathbf{C}}\mathbf{I} \leq \boldsymbol{\Omega}_k^{\mathbf{h}} + \widehat{\mathbf{R}}_k + \mathbf{C}_k^{\mathbf{h}},
\end{aligned} \tag{19}$$

for all $k \geq 1$.

(A2) There exist non-negative scalars $\varepsilon^{\mathbf{f}}$ and $\varepsilon^{\mathbf{h}}$ such that $\|\mathbf{e}_{k-1}^{\mathbf{f}}\|_2 \leq \varepsilon^{\mathbf{f}}$ and $\|\mathbf{e}_k^{\mathbf{h}}\|_2 \leq \varepsilon^{\mathbf{h}}$ for all $k \geq 1$.

(A3) There exist non-negative scalars $\kappa_{\boldsymbol{\varphi}}$, $\kappa_{\boldsymbol{\varphi}}^{+}$, $\kappa_{\boldsymbol{\chi}}$ and $\kappa_{\boldsymbol{\chi}}^{+}$ such that

$$\begin{aligned}
\|\boldsymbol{\varphi}_{k-1}\|_2 &\leq \kappa_{\boldsymbol{\varphi}}\, \|\mathbf{x}_{k-1} - \mathbf{m}_{k-1}\|_2 + \kappa_{\boldsymbol{\varphi}}^{+}, \\
\|\boldsymbol{\chi}_k\|_2 &\leq \kappa_{\boldsymbol{\chi}}\, \|\mathbf{x}_k - \mathbf{m}_k^{-}\|_2 + \kappa_{\boldsymbol{\chi}}^{+}
\end{aligned}$$

for all $k \geq 1$.

Then, given any $\varepsilon \geq 0$ for which $\|\boldsymbol{\xi}_1^{-}\|_2 \leq \varepsilon$, there exists $\delta \geq 0$ such that the conditions $\mathbf{Q}_k, \mathbf{R}_k \leq \delta^2 \mathbf{I}$ guarantee that the predicted estimation error $\boldsymbol{\xi}_k^{-}$ is bounded in mean square if $\kappa_{\boldsymbol{\varphi}}$, $\kappa_{\boldsymbol{\varphi}}^{+}$, $\kappa_{\boldsymbol{\chi}}$, $\kappa_{\boldsymbol{\chi}}^{+}$, $\varepsilon^{\mathbf{f}}$ and $\varepsilon^{\mathbf{h}}$ are sufficiently small.

Proof: The proof goes as that of Theorem 4 and accompanying lemmas. That the proof of Lemma 6 is identical follows from (19) and the proofs of Lemmas 7 and 8 require only minor tweaking of the upper bounds.

The proof of Lemma 8 includes a few modifications. First of all, the constant A is this time $1 + gh/\hat{r}_{\mathbf{C}}$. Secondly, because

$$\boldsymbol{\xi}_k = \mathbf{A}_k \boldsymbol{\xi}_k^{-} - \mathbf{K}_k \mathbf{r}_k - \mathbf{K}_k \boldsymbol{\omega}_2\big[\mathbf{h}(\mathbf{x}_k), \mathbf{m}_k^{-}, \mathbf{P}_k^{-}\big] + \mathbf{K}_k \mathbf{e}_k^{\mathbf{h}},$$

additional terms that include $\varepsilon^{\mathbf{h}}$ will appear in the upper bound of the lemma.

By (18), Equation (14) for the stochastic Lyapunov function now includes some additional terms containing $\boldsymbol{\tau}_k$ and not vanishing when the expectation is taken, namely

$$2\boldsymbol{\tau}_k^{\mathsf{T}}\boldsymbol{\Pi}_{k+1}\big(\mathbf{F}_k \mathbf{A}_k \boldsymbol{\xi}_k^{-} + \boldsymbol{\rho}_k\big) + \boldsymbol{\tau}_k^{\mathsf{T}}\boldsymbol{\Pi}_{k+1}\boldsymbol{\tau}_k.$$

As in Lemma 7, these can be bounded from above by

$$\begin{aligned}
&2\mathbb{E}\left(\boldsymbol{\tau}_k^{\mathsf{T}}\boldsymbol{\Pi}_{k+1}\big[\mathbf{F}_k \mathbf{A}_k \boldsymbol{\xi}_k^{-} + \boldsymbol{\rho}_k\big]\right) \\
&\quad \leq \frac{2}{p_1^{-}}\left(\left(\varepsilon^{\mathbf{f}} + \frac{fg\varepsilon^{\mathbf{h}}}{\hat{r}_{\mathbf{C}}}\right)\left(\kappa_1 + f + \frac{fgh}{\hat{r}_{\mathbf{C}}}\right)\right)\|\boldsymbol{\xi}_k^{-}\|_2 \\
&\qquad + \frac{2}{p_1^{-}}\left(\varepsilon^{\mathbf{f}} + \frac{fg\varepsilon^{\mathbf{h}}}{\hat{r}_{\mathbf{C}}}\right)\kappa_2 \delta
\end{aligned}$$

and

$$\mathbb{E}\left(\boldsymbol{\tau}_k^{\mathsf{T}}\boldsymbol{\Pi}_{k+1}\boldsymbol{\tau}_k\right) \leq \frac{1}{p_1^{-}}\left(\left(\varepsilon^{\mathbf{f}}\right)^2 + \frac{2fg\varepsilon^{\mathbf{f}}\varepsilon^{\mathbf{h}}}{\hat{r}_{\mathbf{C}}} + \left(\frac{fg\varepsilon^{\mathbf{h}}}{\hat{r}_{\mathbf{C}}}\right)^2\right)$$

with κ_1 and κ_2 analogous to those in (11). These terms then appear on the right-hand side of (15) and, with $\varepsilon^{\mathbf{f}}$, $\varepsilon^{\mathbf{h}}$, $\kappa_{\boldsymbol{\varphi}}$, $\kappa_{\boldsymbol{\varphi}}^{+}$, $\kappa_{\boldsymbol{\chi}}$ and $\kappa_{\boldsymbol{\chi}}^{+}$ sufficiently small, Lemma 3 can be applied. ∎

VII. A NUMERICAL SIMULATION

This section provides an example of a stable non-linear Kalman filter for a one-dimensional system. Stability of the filter is verified before any measurements have arrived. Application of Theorems 4 and 9 is difficult in higher dimensions because of the difficulties in obtaining upper bounds for $\mathbf{P}_k^-$ and $\mathbf{P}_k$. As such, the numerical simulation here illustrates conservativeness of our theoretical results.

Consider the system (2) in univariate setting with functions f and h of the form

$$\begin{aligned} f(x) &= a_f x + g_f(x), \\ h(x) &= a_h x + g_h(x), \end{aligned}$$

where a_f and a_h are some constants and g_f and g_h some Lipschitz functions such that $h_1 := \inf_{x\in\mathbb{R}} h'(x) > 0$. We use the exact non-linear Kalman filter presented in Algorithm 1. Now, by Chernoff's inequality

$$\mathbb{E}\big[f'(x)\big]^2 P \leq \mathrm{Cov}[f(x)] \leq \mathbb{E}\big[f'(x)^2\big] P,$$

where $x \sim \mathcal{N}(m, P)$ (see e.g. [37]). We therefore get

$$\begin{aligned} P_k^- &= \mathrm{Cov}_{k-1}[f(x)] + \widehat{Q}_{k-1} \\ &\leq \mathbb{E}_{k-1}\big[f'(x)^2\big] P_{k-1} + \widehat{Q}_{k-1} \\ &\leq \mathrm{Lip}(f)^2 P_{k-1} + \widehat{Q}_{k-1}. \end{aligned}$$

Note that a multi-dimensional form of the latter inequality does not regrettably hold (see [38] for similar matrix inequalities). Chernoff's inequality also yields the error covariance bound

$$\begin{aligned} P_k &= \left(1 - \frac{\mathbb{E}_k^-\big[h'(x)\big]^2 P_k^-}{\mathrm{Cov}_k^-[h(x)] + \widehat{R}_k}\right) P_k^- \\ &\leq \left(1 - \frac{h_1^2 P_k^-}{h_1^2 P_k^- + \widehat{R}_k}\right) P_k^-. \end{aligned}$$

Hence upper bounds for P_k^- and P_k are given by the upper bounds $\sup_{k\geq 0} \mathrm{Var}(x_k \mid y_{1:k})$ and $\sup_{k\geq 0} \mathrm{Var}(x_k \mid y_{1:k-1})$ of the linear Kalman filter variances for the system

$$\begin{aligned} x_k &= \mathrm{Lip}(f) x_{k-1} + q_{k-1}, \\ y_k &= h_1 x_k + r_k. \end{aligned}$$

The steady-state Kalman filter prediction variance P^- of this system can be easily calculated. When the non-linear filter is initialised such that $P_1^- \leq P^-$, we obtain the upper bounds $p_2^- = p_2 = P^-$.

For a numerical simulation we use a modified univariate non-stationary growth model (see e.g. [39]) with linear measurement model

$$\begin{aligned} x_{k+1} &= x_k + C\frac{x_k}{1+x_k^2} + q_k, \\ y_{k+1} &= H x_{k+1} + r_{k+1}, \end{aligned} \tag{20}$$

where C is a positive constant, $H = 2$ and $q_k, r_{k+1} \sim \mathcal{N}(0, \delta^2)$ with δ to be determined. For this model $\mathrm{Lip}(g_f) = C$ so κ_φ and κ_φ^+ are obtained by Remark (5) after Theorem 4 (by linearity $\kappa_\chi = \kappa_\chi^+ = 0$). We set $p_1^- = \widehat{Q}_k = \widehat{R}_{k+1} = 1$. The parameters β and $\tilde{\varepsilon}$ of the proof of Theorem 4 are set as $\beta = \tilde{\varepsilon} = 2$. By testing different values of C we find that inequalities (16) and (17) hold for $C \leq 0.028$. This value yields $P^- \approx 1.22$, $\delta \leq 0.2004$ and $\|\boldsymbol{\xi}_1^-\|_2 \leq 2p_1^-/p_2^- \approx 1.64$. We can conclude that the predicted estimation error obeys the bound $\|\boldsymbol{\xi}_k^-\|_2 \leq 2$ if $C = 0.028$, $\|\boldsymbol{\xi}_1^-\|_2 \leq 1.64$ and $q_k, r_{k+1} \sim \mathcal{N}(0, 0.2004^2)$. Note that we have used the general upper bounds derived in the course of the proof of Theorem 4. Some of these bounds can be replaced with ones optimised for the univariate case. Using (8) instead of (7) and $A = 1$ yields $C \leq 0.040$ and $\delta \leq 0.186$. Figure 1 illustrates the effect of increasing H.

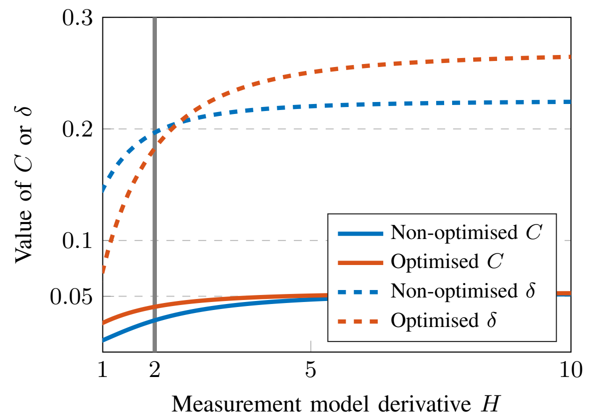

Fig. 1. The effect of increasing the measurement model derivative H on maximal C and δ for the system (20) with and without univariate optimisation. The value of H used in the example simulation is marked with grey. For most values of $H < 1$ stability of the filter cannot be guaranteed.

However, filtering a large number of simulated realisations over 200 time-steps with $x_0 = m_0 = 0$ and minimal initial uncertainty results to $\max_{1\leq k\leq 200} \|\boldsymbol{\xi}_k^-\|_2 \approx 0.22$. Theoretical results derived in this article are thus somewhat conservative.

VIII. CONCLUSIONS AND DISCUSSION

We studied mean square boundedness of estimation error of non-linear Kalman filters that use Gaussian approximations to the true filtering distributions and match the first two moments. The analysis was done with the stochastic stability lemma, the staple of non-linear Kalman filter stability analysis, and Fourier–Hermite series expansion, better suited for the Gaussian integration approach than the customarily employed Taylor series expansion.

We were able to discard the non-intuitive assumption of bounded measurement model Jacobian for the exact non-linear Kalman filter, this result applying also to the EKF. In order to perform rigorous stability assessment beforehand our emphasis was on models with dynamic and measurement model functions Lipschitz continuous. Our results do not require small initial estimation error; this is instead a tunable parameter. However, as seen in our univariate numerical example, constraints imposed on non-linearity of the system are severe and the resulting error bound rather conservative. Application of the results remains challenging, this being mostly attributable to the difficulty of proving boundedness of the filter error covariance matrices.

We find two points in the present methodology that particularly call for improvement. The inequalities used in the proof of the main results, specifically those in Lemma 7, are very crude. Especially unintuitive is the bound (10) as it is usually expected that the error diminishes over the filter update step. Secondly, in bounding the filter error covariance matrix some new innovations are required.

Stability of non-linear Kalman filters is intimately dependent on stability properties of the corresponding optimal Bayesian filter [26, Theorem 4.1] because mean square boundedness of $\boldsymbol{\xi}_k^-$ implies that of the optimal filter covariance. In future research application of ideas and concepts of optimal filtering stability theory (see e.g. [40]) should be attempted in the context of non-linear Kalman filters.

Acknowledgment

This work was supported by Aalto ELEC Doctoral School and the Academy of Finland projects 266940 and 273475. We thank Jakub Prüher for proofreading and comments and the anonymous reviewers for suggestions that helped in improving many points in this article.

References

[1] 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, Jan 1981.

[2] D. Ocone and E. Pardoux, "Asymptotic stability of the optimal filter with respect to its initial condition," *SIAM Journal on Control and Optimization*, vol. 34, no. 1, pp. 226–243, Jan 1996.

[3] 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.

[4] S. Kluge, K. Reif, and M. Brokate, "Stochastic stability of the extended Kalman filter with intermittent observations," *IEEE Transactions on Automatic Control*, vol. 55, no. 2, pp. 514–518, Feb 2010.

[5] K. Rapp and P.-O. Nyman, "On stability of the extended Kalman filter," in *Proceedings of the 10th IEEE International Conference on Methods and Models in Automation and Robotics*, R. Kaszyński and S. Domek, Eds. Wydawnictwo Uczelniane Politechniki Szczecińskiej, 2004.

[6] M. Boutayeb, H. Rafaralahy, and M. Darouach, "Convergence analysis of the extended Kalman filter used as an observer for nonlinear deterministic discrete-time systems," *IEEE Transactions on Automatic Control*, vol. 42, no. 4, pp. 581–586, Apr 1997.

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

[8] Y. Wu, D. Hu, and X. Hu, "Comments on "Performance evaluation of UKF-based nonlinear filtering"," *Automatica*, vol. 43, no. 3, pp. 567–568, Mar 2007.

[9] K. Xiong, H. Zhang, and C. W. Chan, "Author's reply to "Comments on 'Performance evaluation of UKF-based nonlinear filtering'"," *Automatica*, vol. 43, no. 3, pp. 569–570, Mar 2007.

[10] K. Xiong, L. D. Liu, and H. Zhang, "Modified unscented Kalman filtering and its application in autonomous satellite navigation," *Aerospace Science and Technology*, vol. 13, no. 4–5, pp. 238–246, 2009.

[11] L. Li and Y. Xia, "Stochastic stability of the unscented Kalman filter with intermittent observations," *Automatica*, vol. 48, no. 5, pp. 978–981, May 2012.

[12] S. Bolognani, L. Tubiana, and M. Zigliotto, "Extended Kalman filter tuning in sensorless PMSM drives," *IEEE Transactions on Industry Applications*, vol. 39, no. 6, pp. 1741–1747, Nov–Dec 2003.

[13] J. P. Maree, L. Imsland, and J. Jouffroy, "On convergence of the unscented Kalman–Bucy filter using contraction theory," *International Journal of Systems Science*, Aug 2014.

[14] L. Cao, Q. Zhang, X. Chen, and Y. Zhao, "Stochastic stability of center difference predictive filter," *Journal of the Franklin Institute*, vol. 351, no. 11, pp. 5204–5233, Nov 2014.

[15] L. Cao, Y. Tang, X. Chen, and Y. Zhao, "Stochastic stability of sigma-point unscented predictive filter," *ISA Transactions*, vol. 57, pp. 189–204, Jul 2015.

[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, Jan 2000.

[17] J. Xu, S. Wang, G. M. Dimirovski, and Y. Jing, "Stochastic stability of the continuous-time unscented Kalman filter," in *Proceedings of the 47th IEEE Conference on Decision and Control*, 2008, pp. 5110–5115.

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

[19] L. Shang and G. Liu, "Sensor and actuator fault detection and isolation for a high performance aircraft engine bleed air temperature control system," *IEEE Transactions on Control Systems Technology*, vol. 19, no. 5, pp. 1260–1268, Sep 2011.

[20] M. Yazdanian, A. Mehrizi-Sani, and M. Mojiri, "Estimation of electromechanical oscillation parameters using an extended Kalman filter," *IEEE Transactions on Power Systems*, vol. 30, no. 6, pp. 2994–3002, 2015.

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

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

[23] E. Kreyszig, *Introductory Functional Analysis with Applications*, ser. Wiley Classics Library. John Wiley & Sons, 1989.

[24] P. Malliavin, *Stochastic Analysis*, ser. Grundlehren der mathematischen Wissenschaften. Springer, 1997, no. 313.

[25] J. Sarmavuori and S. Särkkä, "Fourier–Hermite Kalman filter," *IEEE Transactions on Automatic Control*, vol. 57, no. 6, pp. 1511–1515, 2012.

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

[27] Z. Wei, X. Zhang, and T. Li, "On Stein identity, Chernoff inequality, and orthogonal polynomials," *Communications in Statistics – Theory and Methods*, vol. 39, no. 14, pp. 2573–2593, 2010.

[28] K. Ito and K. Xiong, "Gaussian filters for nonlinear filtering problems," *IEEE Transactions on Automatic Control*, vol. 45, no. 5, pp. 910–927, May 2000.

[29] J. Duník, O. Straka, and M. Šimandl, "Nonlinearity and non-Gaussianity measures for stochastic dynamic systems," in *16th International Conference on Information Fusion (FUSION)*, 2013, pp. 204–211.

[30] J. Prüher and S. Särkkä, "On the use of gradient information in Gaussian process quadratures," 2016. Submitted.

[31] T.-J. Tarn and Y. Rasis, "Observers for nonlinear stochastic systems," *IEEE Transactions on Automatic Control*, vol. 21, no. 34, pp. 441–448, Aug 1976.

[32] R. Unbehauen, *Systemtheorie 2: Mehrdimensionale, adaptive und nichtlineare Systeme*. Oldenbourg Wissenschaftsverlag, Feb 1998.

[33] R. G. Agniel and E. I. Jury, "Almost sure boundedness of randomly sampled systems," *SIAM Journal on Control*, vol. 9, no. 3, pp. 372–384, Aug 1971.

[34] L. Xie and P. P. Khargonekar, "Lyapunov-based adaptive state estimation for a class of nonlinear stochastic systems," *Automatica*, vol. 48, no. 7, pp. 1423–1431, Jul 2012.

[35] K. Rapp and P.-O. Nyman, "Stability properties of the discrete-time extended Kalman filter," in *Nonlinear Control Systems: A proceedings volume from 6th IFAC Symposium*, F. Allgöwer and M. Zeitz, Eds., vol. 3. Elsevier, 2005, pp. 1377–1382.

[36] A. Gelb, *Applied Optimal Estimation*. MIT Press, 1974.

[37] T. Cacoullos, "On upper and lower bounds for the variance of a function of a random variable," *The Annals of Probability*, vol. 10, no. 3, pp. 799–809, 1982.

[38] J. K. Baksalary, B. Schipp, and G. Trenkler, "Some further results on Hermitian-matrix inequalities," *Linear Algebra and its Applications*, vol. 160, pp. 119–129, Jan 1992.

[39] G. Kitagawa, "Monte Carlo filter and smoother for non-Gaussian nonlinear state space models," *Journal of Computational and Graphical Statistics*, vol. 5, no. 1, pp. 1–25, 1996.

[40] R. Atar and O. Zeitouni, "Exponential stability for nonlinear filtering," *Annales de l'institut Henri Poincaré (B) Probabilités et Statistiques*, vol. 33, no. 6, pp. 697–725, 1997.