University of Helsinki

Master's thesis

Stability of linear and non-linear Kalman filters

Author:
Toni Karvonen

Advisor:
Simo Särkkä

December 4, 2014

HELSINGIN YLIOPISTO — HELSINGFORS UNIVERSITET — UNIVERSITY OF HELSINKI

Tiedekunta/Osasto — Fakultet/Sektion — Faculty	Laitos — Institution — Department
Matemaattis-luonnontieteellinen	Matematiikan ja tilastotieteen laitos

Tekijä — Författare — Author

Toni Karvonen

Työn nimi — Arbetets titel — Title

Stability of linear and non-linear Kalman filters

Oppiaine — Läroämne — Subject

Matematiikka

Työn laji — Arbetets art — Level	Aika — Datum — Month and year	Sivumäärä — Sidoantal — Number of pages
Pro gradu -tutkielma	Joulukuu 2014	93 s.

Tiivistelmä — Referat — Abstract

Kalmanin suodin ja sen approksimatiiviset yleistykset epälineaarisille systeemeille ovat stokastisten dynaamisten systeemien tilaestimoinnin perustyökaluja. Oleellinen kysymys näitä suotimia käytettäessä on, ovatko ne jossakin mielessä stabiileja. Lineaarisella Kalmanin suotimella on vahvoja eksponentiaalisia stabiilisuusominaisuuksia, mutta epälineaarisille Kalmanin suotimille osoitetut stabiilisuustulokset ovat hyvin heikkoja. Näiden tulosten tarkentaminen ja vahvistaminen on avoin tutkimuskohde.

Tässä tutkielmassa esitellään lineaarisen Kalmanin suotimen merkittävimmät stabiilisuustulokset ja olemassa olevat tulokset epälineaarisille Kalmanin suotimille. Lineaarisen Kalmanin suotimen stabiilisuus vaatii säätötekniset oletukset systeemin havaittavuudesta (englanniksi detectability) ja stabiloituvuudesta. Nämä oletukset mahdollistavat epäoptimaalisen lineaarisen tilaestimaattorin, johon liittyvä virhekovarianssimatriisi on rajoitettu, konstruoimisen, jolloin Kalmanin suotimen lineaarisen minimivarianssin ominaisuutta voidaan hyödyntää. Varsinaisessa stabiliisuustodistuksessa käytetään erästä Ljapunovin stabiilisuusteorian yleistystä.

Epälineaarisista Kalmanin suotimista tutkielmassa käsitellään pääasiassa laajennettua Kalmanin suodinta ja hajustamatonta Kalmanin suodinta. Molemmille suotimille todistetaan stokastisia stabiilisuustuloksia erittäin tiukin oletuksin, jotka eivät mahdollista stabiilisuuden toteamista etukäteen. Tulokset saadaan erään stokastista stabiilisuutta koskevan lemman melko suoraviivaisella soveltamisella, vaikkakin hajustamattomalle Kalmanin suotimille todistettaviin tuloksiin vaaditaan myös eräiden approksimaatioiden muuntamista yhtälöiksi diagonaalisten satunnaismatriisien avulla. Hajustamatonta Kalmanin suodinta koskevat tulokset voidaan yleistää kaikilla epälineaarisille Kalmanin suotimille (tai gaussisille suotimille). Nämä tulokset ovat kuitenkin hyvin kvalitatiivisia ja niiden ainoa konkreettinen anti on kohinakovarianssimatriisien virittämisen vaikutuksen selventäminen, mistä tutkielmassa esitetään muutamia yksinkertaisia numeerisia esimerkkejä.

Tutkielman lopussa kartoitetaan eräitä mahdollisesti lupaavia menetelmiä, joita ei ole tähän mennessä käytetty epälineaaristen Kalmanin suotimien stabiilisuuden tutkimiseen. Näitä menetelmiä ovat Fourier'n–Hermiten sarjakehitelmä ja teleskooppisummamenetelmä, jonka avulla on aikaisemmin tutkittu partikkelisuotimien tasaista konvergenssia.

Avainsanat — Nyckelord — Keywords

Kalmanin suodin, epälineaarinen Kalmanin suodin, stabiilisuus

Säilytyspaikka — Förvaringsställe — Where deposited

Kumpulan kampuskirjasto

Muita tietoja — Övriga uppgifter — Additional information

Contents

Notation Conventions

The notation follows for the most part that of Särkkä (2013).

Vectors and Matrices

Throughout this thesis matrices are denoted by boldface upper-case letters and (column) vectors by boldface lower-case letters. The same convention is taken for matrix or vector valued functions. Naturally, scalars and scalar valued functions are then denoted by non-bolded letters. However, if the situation so dictates, matrices can be vectors and vectors scalars. Identity matrix is $\mathbf{I}$ and null matrix $\mathbf{0}$, dimensions of which are not stated unless not self-evident. In this case a subscript revealing the dimensionality is added: $\mathbf{I}_n, \mathbf{0}_n$. Vectors are sometimes denoted by n-tuples.

The notions of positive-definite and positive-semidefinite (or non-negative-definite) matrices are frequently employed. For such matrices $\mathbf{A} > \mathbf{B}$ ($\mathbf{A} \geq \mathbf{B}$) means that $\mathbf{A} - \mathbf{B}$ is positive-definite (positive-semidefinite). Every positive-definite matrix $\mathbf{A}$ admits a unique square root $\sqrt{\mathbf{A}}$ which is a symmetric positive-definite matrix such that $\sqrt{\mathbf{A}}\sqrt{\mathbf{A}} = \mathbf{A}$.

The transpose of a matrix $\mathbf{A}$ is $\mathbf{A}^\mathsf{T}$ and the inverse of the transpose (or the transpose of the inverse) is $\mathbf{A}^{-\mathsf{T}}$. The usual trace operation for a p dimensional square matrix $\mathbf{A}$ is $\operatorname{tr}(A) = \sum_{i=1}^{p} [\mathbf{A}]_{i,i}$, where $[\mathbf{A}]_{i,j}$ are the elements of $\mathbf{A}$. The smallest and largest eigenvalues of a matrix $\mathbf{A}$ are denoted by $\lambda_{\min}(\mathbf{A})$ and $\lambda_{\max}(\mathbf{A})$, respectively.

Norms

Unless otherwise specified, the norm $\|\mathbf{x}\|$ of a vector $\mathbf{x} \in \mathbb{R}^p$ with elements $x_1, x_2, \ldots, x_p$ is the usual Euclidean norm

$$\|\mathbf{x}\| = \sqrt{\sum_{i=1}^{p} x_i^2} = \sqrt{\mathbf{x}^\mathsf{T}\mathbf{x}}.$$

Similarly, the norm $\|\mathbf{A}\|$ of a matrix $\mathbf{A} \in \mathbb{R}^{r \times p}$ is

$$\|\mathbf{A}\| = \sup\left\{ \frac{\|\mathbf{A}\mathbf{x}\|}{\|\mathbf{x}\|} \, : \, \mathbf{x} \in \mathbb{R}^p, \|\mathbf{x}\| \neq 0 \right\}.$$

For square matrices this norm is the same as the spectral norm given by $\sqrt{\lambda_{\max}(\mathbf{A}^\mathsf{T}\mathbf{A})}$. In fact, in the end it is of little importance which matrix norm is chosen.

The L^2-norm of a function $\mathbf{f}\colon \Omega \to \mathbb{R}^n$ is defined as usual (with μ a measure in a σ-algebra of Ω) by

$$\|\mathbf{f}\|_2 = \left(\int_\Omega \|\mathbf{f}\|^2 \, \mathrm{d}\mu \right)^{1/2} .$$

When $\sup_{i \in \mathcal{I}} \|\mathbf{A}_i\| < \infty$, where $\mathcal{I}$ is any index set, $\mathbf{A}_i$ is called uniformly bounded. The definition applies to scalars also.

Probability Theory

The terms probability distribution and probability density are frequently used interchangeable and the notation for probability density functions abused. The context is such that there should be no risk of serious ambiguity. The underlying probability measure is $\mathbb{P}$.

The expectation of a random vector $\boldsymbol{\xi}$ is $\mathbb{E}(\boldsymbol{\xi})$. The conditional expectation of a random vector $\boldsymbol{\xi}$ with respect to another random vector $\boldsymbol{\eta}$ (or the σ-algebra generated by it) is $\mathbb{E}(\boldsymbol{\xi} \mid \boldsymbol{\eta})$. The covariance matrix of a random vector $\boldsymbol{\xi}$ with mean $\mathbf{m}$ is $\mathrm{Cov}(\boldsymbol{\xi}) = \mathbb{E}((\boldsymbol{\xi} - \mathbf{m})(\boldsymbol{\xi} - \mathbf{m})^\mathsf{T})$.

Gaussian (normal) distribution with mean $\mathbf{m}$ and covariance matrix $\mathbf{P}$ is denoted by $\mathcal{N}(\mathbf{m}, \mathbf{P})$. This is the only special distribution explicitly used.

Miscellaneous Matters

The Hessian matrix of second partial derivatives of a smooth enough function $\mathbf{f}$ is $\mathrm{Hess}(f)$.

The thesis is mostly concerned with dynamic systems that have some system state $\mathbf{x}_k$ of which measurements $\mathbf{y}_k$ are obtained. It will be throughout assumed that $\mathbf{x}_k \in \mathbb{R}^n$, $\mathbf{y}_k \in \mathbb{R}^m$ and $1 \le m \le n$, which is no restriction.

1 Introduction

The optimal linear Kalman filters (Kalman, 1960b) and their approximative non-linear extensions (non-linear Kalman filters) for state estimation of stochastic dynamic systems have been around since the 1960s. During the last 20 years a considerable number of new approximative non-linear filters have emerged (Särkkä, 2013). Through experience it has been seen that most of these filters seem to be stable in some sense under moderate conditions.

Fortunately, one is not left to trust mere experience in all cases. For linear Kalman filters, very strong stability results, obtained by Lyapunov stability techniques, have existed almost since their inception (Jazwinski, 1970) requiring only weak assumptions that have been further relaxed (Anderson and Moore, 1981). However, most practical models are not linear. The stability of optimal non-linear filters has received considerable attention (van Handel, 2007), but as the optimal non-linear filter is rarely implementable, these results are of little use in practice.

The stability of most approximative non-linear filters was not backed by any rigorous analysis until late 1990s when it was proven (Reif et al., 1999) by the use of a certain supermartingale boundedness lemma that under very strict conditions, rarely satisfied in practice and impossible to verify beforehand, boundedness of the estimation error can be concluded for the extended Kalman filter that is one of the most used of non-linear filters. Similar analysis has been since then extended to a large number of different filters, without essential improvements in the results however. Most important of these extensions are those for the unscented Kalman filter and the wide class of Gaussian filters (Xiong et al., 2006, 2009a; Li and Xia, 2012) even though they provide only qualitative insight. As such, the stability of approximative non-linear filters remains an open research problem that would significantly benefit from new methods.

This thesis reviews the classical stability results of the linear Kalman filter and the existing stability results for some non-linear filters. The non-linear case is exhaustively referenced. The results are critically discussed and some of the proofs are given minor improvements. Also, some new tools that might in the future contribute to better stability results for Gaussian filters are introduced. These are partly based on the ideas of Simo Särkkä.

The treatment is limited to the discrete-time case, which is to curb the length of the thesis as well as to keep the content somewhat consistent as the stability of continuous-time non-linear Kalman filters has generated very little research. Detailed proofs are presented selectively and moderate amount of historical commentary is included. Sections 2 and 3 briefly cover the basics of Bayesian filtering theory and general stability notions and results that are used in the sequel. Section 4 is about the stability of the time-varying linear Kalman filter and Section 5 contains the most important stability results for the non-linear Kalman filters. Some simple numerical examples, that mostly investigate the effects of the practice known as covariance matrix tuning, are included. Section 6 contains some suggestions for new methods to approach the stability of Gaussian filters. Also attached is a short appendix that contains some elementary matrix results and a proof of the optimality of the linear Kalman filter.

1.1 Literature Review

The general stability theory for deterministic dynamic systems is covered in Willems (1970) and for difference equations in Lakshmikantham and Trigiante (2002). For stocastic stability there seems to be not a single great source, but some material can be found in Gard (1988) which serves also as an introduction to stochastic differential equations.

The classical source on Kalman filtering containing the early stability results is Jazwinski (1970) and another one, though lacking detailed discussion on stability, is Maybeck (1979). An excellent reference, where stability is treated in the time-invariant case, is Anderson and Moore (1979). A more recent and corrected work is Kamen and Su (1999). Generalizations of the early linear Kalman filter stability can be found in Moore and Anderson (1980) and Anderson and Moore (1981).

Indispensable sources on different non-linear filtering algorithms are Särkkä (2013), which also serves as a Bayesian introduction to filtering and smoothing, and Simon (2006) where treatment is less extensive but more detailed. Stability results for these filters in English exist currently only in research papers, most important of which are Reif et al. (1999) and Kluge et al. (2010) for the extended Kalman filter and Xiong et al. (2006, 2009a) Li and Xia (2012) for the unscented Kalman filter and related Gaussian filters.

1.2 Acknowledgements

This thesis began as a report on the stability of non-linear Kalman filters serving as my summer project in the Bayesian Statistical Methods group in the Department of Biomedical Engineering and Computational Science (BECS) at Aalto University.

I would like to thank my advisor Simo Särkkä from BECS for his guidance and providing me this opportunity to explore for me hitherto unknown world of filtering. I would also like to thank Arno Solin from BECS for his support and advice during this project.

2 Bayesian Filtering

In every measurement of any conceivable quantity there is some uncertainty introduced by measuring devices, always at least to some degree inherently inaccurate, and the lack of thorough knowledge of measuring situation and environment. Also, if the measured quantity is modeled as some kind of time-varying system, there may lurk inaccuracies due to incomplete understanding of the underlying model as well as the sheer complexity of taking into account every tiny variable capable of causing disturbance in the system. Consequently, one can never know exactly what the quantity measured in reality is. In such a regrettable situation one must then settle for the best one can hope for, that is, the best possible estimate — in some sense — of the quantity given all available knowledge of the model and the measurements obtained. To tackle this problem the *filtering theory* was developed. The precursor of the filtering theory is least-squares method of Gauss in the 19th century but its modern origins lie in the work of Norbert Wiener and Andrey Kolmogorov in the 1940s (Sorenson, 1970) that led to the discovery of the recursive linear optimal mean square estimator by Rudolf Kalman in 1960 (Kalman, 1960b), still one of the most celebrated achievement of the field.

This section gives a brief introduction to filtering theory in Bayesian framework. The presentation is rather light in mathematical terms as the emphasis of this thesis is in investigating discrete-time linear filtering and approximative non-linear filtering, neither of which requires advanced stochastic machinery or stochastic calculus. The section is primarily based on Särkkä (2013). A more general account can be found for example in Bain and Crisan (2009).

2.1 Overview of Bayesian Filtering

Suppose the unknown quantity one is interested in determining with as much accuracy as possible is a vector-valued (of vectors of dimension n) time-series $\{\mathbf{x}_0, \mathbf{x}_1, \mathbf{x}_2, \ldots\}$ of *states* of which one is only able to acquire noisy *measurements* $\{\mathbf{y}_1, \mathbf{y}_2, \ldots\}$. Figure 2.1 contains an illustration of this situation in a simple setting. The problem of determining the states $\mathbf{x}_{0:T} = \{\mathbf{x}_0, \ldots, \mathbf{x}_T\}$ up to time-step T from the noisy measurements $\mathbf{y}_{1:T} = \{\mathbf{y}_1, \ldots, \mathbf{y}_T\}$ is then that of *statistical inversion* which in the Bayesian framework amounts to computing the *joint posterior distribution* of all the states given the measurements. The joint posterior distribution can be, in principle, computed

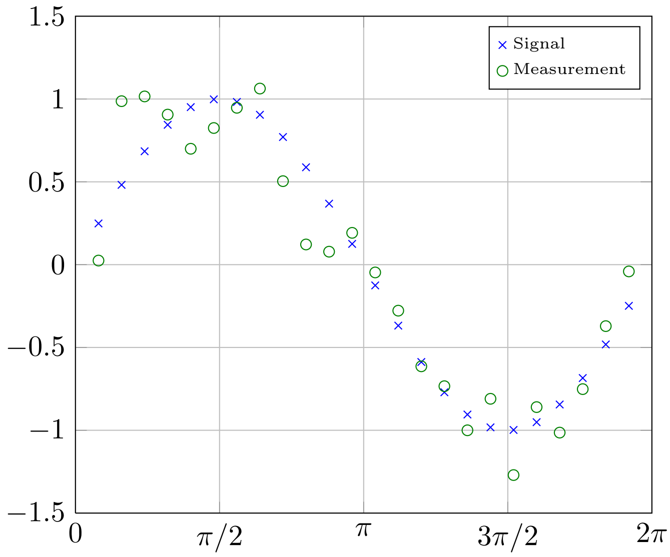

Figure 2.1: An example of a discrete-time sine signal. The true signal state remains unknown but noisy measurements are obtained of it. The objective of filtering is to infer the unknown state when only the measurements are known.

by an application of Bayes' rule, result being the posterior distribution

$$p(\mathbf{x}_{0:T} \mid \mathbf{y}_{1:T}) = \frac{p(\mathbf{y}_{1:T} \mid \mathbf{x}_{0:T})p(\mathbf{x}_{0:T})}{p(\mathbf{y}_{1:T})},$$

where $p(\mathbf{x}_{0:T})$ is the *prior distribution* determined by the underlying model and the normalization constant in the denominator is

$$p(\mathbf{y}_{1:T}) = \int_{\mathbb{R}^{n\times T}} p(\mathbf{y}_{1:T} \mid \mathbf{x}_{0:T})p(\mathbf{x}_{0:T})\,\mathrm{d}\mathbf{x}_{0:T}.$$

However, the burden of computing this full posterior distribution quickly becomes unbearable, particularly so if the measurements are obtained one at a time and an estimate for the state corresponding to the measurement just obtained is to be computed immediately. This amounts to computing the whole joint posterior distribution of ever increasing dimensionality every time a new measurement is obtained. This issue of computational complexity is circumvented if one consents to be satisfied only with certain marginal distributions of the states and supposes that the states form a *Markov sequence* and that each measurement depends only on the corresponding state. The Markov assumption $p(\mathbf{x}_k \mid \mathbf{x}_{0:k-1}, \mathbf{y}_{1:k-1}) = p(\mathbf{x}_k \mid \mathbf{x}_{k-1})$, which means that the states are dependent only on the very previous state instead of possibly all the preceding ones, provides in the end the luxury of recursive computations. In this way the states and the measurements can be presented by the probabilistic model

$$\begin{aligned}\mathbf{x}_0 &\sim p(\mathbf{x}_0),\\ \mathbf{x}_k &\sim p(\mathbf{x}_k \mid \mathbf{x}_{k-1}),\\ \mathbf{y}_k &\sim p(\mathbf{y}_k \mid \mathbf{x}_k).\end{aligned}$$

What are then the marginal posterior distributions to settle for? The three such distributions of interest are the following:

- *Filtering distributions* are the marginal distributions of the current state $\mathbf{x}_k$ given the measurements up to and including the one corresponding to the current state,

$$p(\mathbf{x}_k \mid \mathbf{y}_{1:k}). \tag{2.1}$$

- *Prediction distributions* are the marginal distributions of the future state $\mathbf{x}_{k+m}$ given only the measurements up to some preceding time-step,

$$p(\mathbf{x}_{k+m} \mid \mathbf{y}_{1:k}). \tag{2.2}$$

- *Smoothing distributions* are the marginal distributions of the state $\mathbf{x}_k$ given also measurements after the time-step k up to Tth time-step,

$$p(\mathbf{x}_k \mid \mathbf{y}_{1:T}), \qquad T > k. \tag{2.3}$$

The computation of filtering distributions is called *filtering* and that of smoothing distributions *smoothing*. The corresponding terms for equations to achieve these tasks are *filter* and *smoother*. One-step prediction distributions arise naturally in the Bayesian filtering equations of Theorem 2.1 and as such are intrisically connected to the filtering distributions. In this thesis smoothing is not addressed in its own right, although a certain smoothed estimate will be constructed in a technical proof of Theorem 4.14.

Having in some way calculated the filtering distribution, one is faced with the task of choosing a single suitable state estimator. The conventional choice is the conditional expectation $\mathbb{E}(\mathbf{x}_k \mid \mathbf{y}_{1:k})$ which is the best state estimator in the sense of minimizing the mean square estimation error. This follows from the orthogonality property of the conditional expectation in the space of square integrable random variables (Williams, 1991, Chapter 9).

2.2 Bayesian Filtering Equations

Of course, none of these definitions is worthwhile if there is no way to actually compute the distributions (2.1)–(2.3) in a meaningful recursive way. That the filtering distributions can be computed recursively arises from Bayes' rule

$$p(\mathbf{x}_k \mid \mathbf{y}_{1:k}) = p(\mathbf{x}_k \mid \mathbf{y}_k, \mathbf{y}_{1:k-1}) \propto p(\mathbf{y}_k \mid \mathbf{x}_k)p(\mathbf{x}_k \mid \mathbf{y}_{1:k-1}).$$

Therefore, an expression for the predictive distribution $p(\mathbf{x}_k \mid \mathbf{y}_{1:k-1})$ in terms of the preceding filtering distribution $p(\mathbf{x}_{k-1} \mid \mathbf{y}_{1:k-1})$ would provide one with a recursive formula for the filtering distributions. This is easily achieved, the following theorem yielding the general recursive equations for doing this. Similar equations are easily obtained for the smoothing problem also (Särkkä, 2013, Theorem 8.1).

Theorem 2.1 (Bayesian filtering equations). *The Bayesian filtering equations for computing the filtering distribution $p(\mathbf{x}_k \mid \mathbf{y}_{1:k})$ at the time-step k consist of the* prediction step *for computing the prediction distribution $p(\mathbf{x}_k \mid \mathbf{y}_{1:k-1})$ by the* Chapman–Kolmogorov equation

$$p(\mathbf{x}_k \mid \mathbf{y}_{1:k-1}) = \int_{\mathbb{R}^n} p(\mathbf{x}_k \mid \mathbf{x}_{k-1})p(\mathbf{x}_{k-1} \mid \mathbf{y}_{1:k-1}) \,\mathrm{d}\mathbf{x}_{k-1} \tag{2.4}$$

and the update step *that utilizes the prediction distribution for computing the filtering distribution by Bayes' rule*

$$p(\mathbf{x}_k \mid \mathbf{y}_{1:k}) = \frac{1}{Z_k} p(\mathbf{y}_k \mid \mathbf{x}_k)p(\mathbf{x}_k \mid \mathbf{y}_{1:k-1}), \tag{2.5}$$

where the normalization constant Z_k is

$$Z_k = \int_{\mathbb{R}^n} p(\mathbf{y}_k \mid \mathbf{x}_k)p(\mathbf{x}_k \mid \mathbf{y}_{1:k-1}) \,\mathrm{d}\mathbf{x}_k. \tag{2.6}$$

In case of discrete state components the corresponding integrals are replaced by sums.

Proof. (Särkkä, 2013, Theorem 4.1) The joint distribution of $\mathbf{x}_k$ and $\mathbf{x}_{k-1}$ given $\mathbf{y}_{1:k-1}$ is

$$\begin{aligned} p(\mathbf{x}_k, \mathbf{x}_{k-1} \mid \mathbf{y}_{1:k-1}) &= p(\mathbf{x}_k \mid \mathbf{x}_{k-1}, \mathbf{y}_{1:k-1})p(\mathbf{x}_{k-1} \mid \mathbf{y}_{1:k-1}) \\ &= p(\mathbf{x}_k \mid \mathbf{x}_{k-1})p(\mathbf{x}_{k-1} \mid \mathbf{y}_{1:k-1}), \end{aligned}$$

since the measurement history does not affect the state by the Markov property of the model. Then the prediction distribution, that is the marginal distribution of $\mathbf{x}_k$ given $\mathbf{y}_{1:k-1}$, can be computed by integrating out $\mathbf{x}_{k-1}$. This way one obtains the Chapman–Kolmogorov equation

$$p(\mathbf{x}_k \mid \mathbf{y}_{1:k-1}) = \int_{\mathbb{R}^n} p(\mathbf{x}_k \mid \mathbf{x}_{k-1}) p(\mathbf{x}_{k-1} \mid \mathbf{y}_{1:k-1}) \, \mathrm{d}\mathbf{x}_{k-1}.$$

Then the distribution of $\mathbf{x}_k$ given $\mathbf{y}_{1:k}$, or in other words, given $\mathbf{y}_k$ and $\mathbf{y}_{1:k-1}$ is by Bayes' rule

$$\begin{aligned} p(\mathbf{x}_k \mid \mathbf{y}_{1:k}) &= p(\mathbf{x}_k \mid \mathbf{y}_k, \mathbf{y}_{1:k-1}) \\ &= \frac{1}{Z_k} p(\mathbf{y}_k \mid \mathbf{x}_k, \mathbf{y}_{k-1}) p(\mathbf{x}_k \mid \mathbf{y}_{k-1}) \\ &= \frac{1}{Z_k} p(\mathbf{y}_k \mid \mathbf{x}_k) p(\mathbf{x}_k \mid \mathbf{y}_{k-1}), \end{aligned}$$

where the normalization constant Z_k is given by (2.6) and the past measurement history $\mathbf{y}_{1:k-1}$ can be dropped from the last equality as the current measurement $\mathbf{y}_k$ is conditionally independent, given $\mathbf{x}_k$, of the past measurements. □

In the prediction step the filtering distribution is propagated through the system dynamics and in the update step this propagated distribution is conditioned on the new measurement to obtain a new filtering distribution.

Unfortunately, Equations (2.4) and (2.5) can be solved in a closed form only if the dependencies between the current and the preceding state as well as that of the current state and measurement of it are linear and noise is additive and Gaussian. The resulting recursion is known as the *Kalman filter* and is treated in Section 4.

If the dependencies are non-linear, no simple solution in closed form exists in general. For an efficient approximative solution of this non-linear filtering problem a great multitude of different *approximative non-linear filters* have been proposed. Some of the most widely applied of these are the topic of Section 5. Of course, the Bayesian filtering equations can be studied as such without employing any approximation. Such filters are called *optimal filters* and naturally include the Kalman filter. In this thesis optimal non-linear filters are rarely considered even though there is extensive theory concerning them.

3 Concepts of Stability

This section discusses several different concepts of stability that will be of use later. For linear dynamic systems a few strong deterministic notions of Lyapunov stability are introduced. Since the non-linear Kalman filters whose stability is investigated are approximative and sub-optimal, the corresponding stability notions are far weaker. The much more complex nature of non-linear systems also forces one to have the stability notions stochastic. As a small sidenote, one notion of optimal filter stability is also discussed.

3.1 Lyapunov Stability of Linear Systems

First, the stability of linear dynamic systems is considered. A *discrete-time linear dynamic system* is of the form

$$\mathbf{x}_k = \mathbf{A}_{k-1}\mathbf{x}_{k-1} + \mathbf{B}_{k-1}\mathbf{u}_{k-1} \tag{3.1}$$

for $k \geq 1$. The vector $\mathbf{x}_k \in \mathbb{R}^n$ is the *system state*, $\mathbf{u}_{k-1} \in \mathbb{R}^p$ the *input*, *control* or the *forcing function* and $\mathbf{A}_{k-1}$ and $\mathbf{B}_{k-1}$ possibly time-varying coefficient matrices of appropriate dimensions. This system is commonly called *forced*, referring to the dynamics being disturbed be $\mathbf{u}_{k-1}$. The *homogeneous part* of (3.1) is the same system but with the forcing term disregarded. The homogeneous part of the system has the *equilibrium state* at the origin as having reached this state it is impossible to escape. The starting time of the system is taken to be $k = 0$ and the *initial state* is thus $\mathbf{x}_0$. It is very much possible for distinct initial states to induce totally different behavior in the system; the interest lies in pinning down the conditions that ensure that the initial state is essentially forgotten in the long run.

The *state transition matrix* of the system (3.1) from time-step j to $i \geq j$ is the matrix

$$\mathbf{\Phi}_{i,j} \coloneqq \prod_{m=j}^{i-1} \mathbf{A}_m$$

with properties $\mathbf{\Phi}_{j,j} = \mathbf{I}$ and $\mathbf{\Phi}_{i,j} = \mathbf{\Phi}_{j,i}^{-1}$ if the matrices $\mathbf{A}_{k-1}$ are non-singular.

Definition 3.1 (Linear system stability)**.** The homogeneous part of the linear dynamic system (3.1) is called *uniformly asymptotically stable* if $\|\mathbf{x}_k\| \to 0$ as $k \to \infty$ regardless of the initial state $\mathbf{x}_0$. An equivalent definition (Bucy and Joseph, 1968, Theorem 1.4) is that there are $c_1, c_2 > 0$, independent of k, such that

$$\|\mathbf{\Phi}_{k,0}\| \leq c_1 \exp(-c_2 k)$$

for all $k \geq 0$. This inspires the terms *exponentially stable*[1] and *uniformly exponentially stable*, former of which shall be used in this thesis. The forced system (3.1) is *bounded-input-bounded-output* (BIBO) stable if bounded inputs produce bounded outputs, that is, $\sup_{k\geq 1} \|\mathbf{u}_{k-1}\| < \infty$ implies $\sup_{k\geq 0} \|\mathbf{x}_k\| < \infty$.

Definitions of stability can be frequently given in terms of either the system state or the state transition matrix as can be seen in the equivalence of uniform asymptotic stability and exponential stability (Willems, 1970, Chapter 4). An important property of linear systems is that exponential stability implies BIBO stability if the matrix $\mathbf{B}_{k-1}$ is bounded.

Theorem 3.2. *An exponentially stable linear dynamic system* (3.1) *is BIBO stable if* $\mathbf{B}_{k-1}$ *is uniformly bounded.*

Proof. (Willems, 1970, p. 12 and Theorem 3.1) □

These two notions of linear system stability, though only a fraction of all existing notions, are sufficient for this thesis. Large number of other definitions in different forms are usually given in the literature covering the subject. Especially for exponential stability there are numerous distinct definitions and characterizations arising from the fact that the concept is not so straightforward as it seems by this exposition if the dynamic system is allowed to be non-linear.

Then, under what conditions is a system stable? The simplest situation is that when $\mathbf{A}_{k-1}$ is contractive, that is, $\|\mathbf{A}_{k-1}\| \leq c < 1$ for all $k \geq 1$ and some positive c and $\mathbf{B}_{k-1}\mathbf{u}_{k-1}$ remains uniformly bounded from above. In this case the convergence of the geometric series implies the boundedness of $\mathbf{x}_k$ and the homogeneous part is clearly exponentially stable. However, such contractive systems are rarities and one thus needs some more general criteria.

[1]Exponential stability can of course be extended to any sequence of scalars with an analogous condition. Furthermore, a condition equivalent to the exponential stability of a sequence a_k of positive scalars is that

$$\lim_{n\to\infty} \frac{1}{k} \log a_k < 0.$$

This form (or limes superior version of it) is often used in the stability analysis of optimal filters that is briefly discussed in Section 3.3

In giving sufficient more general conditions for the stability of linear systems the work of Lyapunov (1907), originally published in Russian in 1892 and translated into French in 1907, is utilized. Specifically, the so-called second method of Lyapunov will be used. The point of this method is to infer about stability of a system without explicit knowledge of the solutions. The idea of the second method is in the analogy of energy: if the energy of a physical system is decreasing for every possible state except for a single equilibrium state, then the the energy continues to decrease until it assumes its minimum value at that equilibrium state. Translation of this physical analogy to precise mathematical formulation means roughly that one needs to find a scalar function V of the system state, termed Lyapunov function, that is positive and strictly decreasing outside the equilibrium state such that $V(\mathbf{x}) = V'(\mathbf{x}) = 0$ if and only if $\mathbf{x}$ is the equilibrium state. Such a function is called *positive-definite*.

Books devoted to the stability theory of dynamic systems are the classical, though somewhat outdated, one by Hahn (1963), containing a short treatment of the discrete-time case, Willems (1970), an excellent source on the relations of different notions of stability, and Michel et al. (2008) with a good exposition of Lyapunov functions. Lakshmikantham and Trigiante (2002) covers the discrete-time case. The articles of Kalman and Bertram (1960a,b) also cover most of the material needed.

The following theorem will be of use in proving the exponential stability of the linear Kalman filter. As the discrete-time versions of stability theorems can usually be obtained as easy consequences of continuous-time counterparts, there is not much literature dealing explicitly with discrete-time systems.

Theorem 3.3 (Lyapunov stability theorem)**.** *The homogeneous part of the discrete-time linear dynamic system* (3.1) *is exponentially stable if there exists a scalar function* $V(\mathbf{x}_k, k)$ *with* $V(\mathbf{0}, k) \equiv 0$*, continuous non-decreasing scalar functions* $\gamma_1(\|\mathbf{x}_k\|)$ *and* $\gamma_2(\|\mathbf{x}_k\|)$ *and a continuous scalar function* $\gamma_3(\|\mathbf{x}_k\|)$ *with* $\gamma_1(0) = \gamma_2(0) = \gamma_3(0) = 0$ *and* $\gamma_1(\|\mathbf{x}_k\|) \to \infty$ *when* $\|\mathbf{x}_k\| \to \infty$*, such that, for some integers* $N, M > 0$*,*

$$0 < \gamma_1(\|\mathbf{x}_k\|) \leq V(\mathbf{x}_k, k) \leq \gamma_2(\|\mathbf{x}_k\|) \tag{3.2}$$

for all $\mathbf{x}_k \neq \mathbf{0}$*,* $k \geq M$ *and*

$$V(\mathbf{x}_k, k) - V(\mathbf{x}_{k-N}, k-N) \leq \gamma_3(\|\mathbf{x}_k\|) < 0 \tag{3.3}$$

for all $\mathbf{x}_k \neq \mathbf{0}$ *and* $k \geq M$*. The function* V *is called a* Lyapunov function *for the system* (3.1).

Proof. Let $\varepsilon > 0$ be arbitrary and choose $\delta > 0$ such that $\gamma_2(\delta) < \gamma_1(\varepsilon)$. This is possible since $\gamma_1(0) = \gamma_2(0) = 0$, γ_1 and γ_2 are continuous and γ_1 acquires every non-negative real value. Let $k \geq M$ and suppose that $\|\mathbf{x}_k\| \leq \delta$. Then the assumptions (3.2) and (3.3) imply that

$$\gamma_1(\varepsilon) > \gamma_2(\delta) \geq V(\mathbf{x}_k, k) > V(\mathbf{x}_{k+N}, k+N) \geq \gamma_1(\|\mathbf{x}_{k+N}\|),$$

and γ_1 being non-decreasing, it follows, with a shift of k, that $\|\mathbf{x}_k\| < \varepsilon$ when $k \geq M + N$ and $\|\mathbf{x}_{k-N}\| \leq \delta$.

Let then $c_1 > 0$ be arbitrary and choose $r > 0$ such that $\gamma_2(r) < \gamma_1(c_1)$, which is again possible. Fix $l \geq M+N$ and suppose that $\|\mathbf{x}_{l-N}\| \leq r$. Then, by the preceding part of the proof and Equation (3.3), $\|\mathbf{x}_{l+kN}\| < c_1$ for all $k \geq 0$. Take any $0 < \mu \leq \|\mathbf{x}_{l-N}\|$ and $\nu > 0$ such that $\gamma_2(\nu) < \gamma_1(\mu)$. Denote $c_2 := \max_{\nu \leq \|\mathbf{x}\| \leq c_1} \gamma_3(\|\mathbf{x}\|) < 0$ and define $m = -\lfloor \gamma_2(r)/c_2 \rfloor > 0$, where $\lfloor a \rfloor$ is the biggest integer smaller than or equal to $a \in \mathbb{R}$. Now, suppose that $\|\mathbf{x}_{l+kN}\| > \nu$ for all $-1 \leq k \leq m$. By the definition of c_2 and an iterated use of (3.3), it follows that

$$\begin{aligned} 0 < \gamma_1(\nu) \leq V(\mathbf{x}_{l+mN}, l+mN) &\leq \sum_{k=0}^{m} \gamma_3(\|\mathbf{x}_{l+kN}\|) + V(\mathbf{x}_{l-N}, l-N) \\ &\leq (m+1)c_2 + \gamma_2(r) \leq \gamma_2(r) - c_2 \left\lfloor \frac{\gamma_2(r)}{c_2} \right\rfloor + c_2 \leq c_2 < 0, \end{aligned}$$

which is a contradiction. Therefore $\|\mathbf{x}_{l+k_1 N}\| \leq \nu$ for some $-1 \leq k_1 \leq m$. Thus, for all $k \geq k_1$, one has

$$0 < \gamma_1(\nu) \leq V(\mathbf{x}_{l+kN}, l+kN) \leq V(\mathbf{x}_{l+k_1 N}, l+k_1 N) \leq \gamma_2(\nu) < \gamma_1(\mu)$$

and hence $\|\mathbf{x}_{l+kN}\| < \mu$ for all $k \geq m \geq k_1$. Since $\mu > 0$ can be chosen arbitrarily small, $\|\mathbf{x}_{l+kN}\| \to 0$ as $k \to \infty$. Having assumed that $\gamma_1(\|\mathbf{x}_k\|) \to \infty$ when $\|\mathbf{x}_k\| \to \infty$, one can choose r to be arbitrarily large, provided that c_1 is first chosen to be large enough. Therefore there is no need to impose an upper bound on $\|\mathbf{x}_{l-N}\|$. Furthermore, since no other restriction than that of being bounded from below by $N + M$ was placed on l, the convergence of $\|\mathbf{x}_{l+kN}\|$ to zero holds for any such l. Thus $\|\mathbf{x}_k\| \to 0$ as $k \to \infty$ which proves the exponential stability. □

The above theorem and its proof are slight modifications of Kalman and Bertram (1960a, Theorem 1) as a consequence of which it the theorem is stated in Jazwinski (1970, p. 240); Deyst and Price (1968) and

Kamen and Su (1999, Appendix C). In this version the time domain is discrete and the Lyapunov function is required to decrease only over N steps instead of being decreasing at every step. In this form the theorem is particularly suited for the assumptions that will be made when the stability of the Kalman filter is being proven in Section 4.4. An extended version of this theorem is needed in Section 4.5 but as it is intrinsically connected with certain definitions from control theory and somewhat non-classical, it is not presented yet.

3.2 Stochastic Stability

In this section the boundedness of stochastic processes is investigated. For the purposes of this thesis this amounts to finding conditions that guarantee the mean square and stochastic boundedness of the process. So, under consideration is a discrete-time stochastic process $\boldsymbol{\xi}_k, k \geq 0$ with state space $\mathbb{R}^n$. No dynamic structure as in the preceeding section is assumed for the process.

An intuitive extension of the Lyapunov stability ideas to the stochastic framework would be to consider the conditional expectations of some Lyapunov function (which in this case is too a stochastic process) and infer something about stochastic stability from their behavior. An important observation that positive supermartingales correspond to Lyapunov functions was made by Bucy (1965). This is well supported by intuition since a supermartingale tends to decrease in time and is in analogy with Lyapunov functions that are decreasing. By the aid of this observation a stochastic stability theory similar to that of the ordinary Lyapunov stability theory was developed (Bucy and Joseph, 1968, Chapter VI). However, the most elegant part of this theory will not be needed, for the stability analysis of the linear Kalman filter can be done with classical Lyapunov stability theory and with non-linear Kalman filters the assumptions of the stochastic stability theory are rarely satisfied.

Further treatment of the stochastic stability can be found in Kushner (1967); Aoki (1967) and Gard (1988), where the treatment is geared towards the stability questions of stochastic differential equations. This subject is taken up also by Khasminskii (2011). Dragan et al. (2006, Section 2.5) discuss exponential mean square convergence to zero for linear systems.

The following four modes of stochastic stability are relevant in the setting of non-linear Kalman filters, although it appears that it can only be proven that two of them hold.

Definition 3.4. A stochastic process $\boldsymbol{\xi}_k$ is *exponentially bounded in mean square* if there are constants $\eta, \nu > 0$ and $0 < \vartheta < 1$ such that

$$\mathbb{E}(\|\boldsymbol{\xi}_k\|^2) \leq \eta \mathbb{E}(\|\boldsymbol{\xi}_0\|^2)\vartheta^k + \nu$$

for all $k \geq 0$.

Definition 3.5. A stochastic process $\boldsymbol{\xi}_k$ is *stochastically bounded* if for any $\varepsilon > 0$ there is $M > 0$ such that $\sup_{k\geq 0} \mathbb{P}(\|\boldsymbol{\xi}_k\| \geq M) \leq \varepsilon$.

Definition 3.6. A stochastic process $\boldsymbol{\xi}_k$ is *stochastically sample path bounded* if for any $\varepsilon > 0$ there is $M > 0$ such that $\mathbb{P}(\sup_{k\geq 0} \|\boldsymbol{\xi}_k\| \geq M) \leq \varepsilon$.

Definition 3.7. A stochastic process $\boldsymbol{\xi}_k$ is *almost surely bounded* if $\sup_{k\geq 0} \|\boldsymbol{\xi}_k\| < \infty$ almost surely.

Clearly, a stochastic process is exponentially bounded in mean square if and only if it is uniformly bounded in mean square, which means that there exists a bound $M > 0$ such that $\mathbb{E}(\|\boldsymbol{\xi}_k\|^2) \leq M$ for all $k \geq 0$. As will be seen, mean square boundedness implies stochastic boundedness.

The following lemma and similar forms of it have been essential in almost every stability proof for non-linear Kalman filters. The version here is only concerned with exponential mean square boundedness and stochastic boundedness, rest of the notions are for a brief discussion.

Lemma 3.8 (Stochastic stability lemma). *Let $\boldsymbol{\xi}_k$ be a Markov process. Suppose there is a scalar-valued stochastic process $V_k(\boldsymbol{\xi}_k)$, adapted to the same σ-algebra, and positive real numbers v_1, v_2, μ and $0 < \alpha \leq 1$ such that*

$$v_1 \|\boldsymbol{\xi}_k\|^2 \leq V_k(\boldsymbol{\xi}_k) \leq v_2 \|\boldsymbol{\xi}_k\|^2 \tag{3.4}$$

and

$$\mathbb{E}[V_{k+1}(\boldsymbol{\xi}_{k+1}) \mid \boldsymbol{\xi}_k] - V_k(\boldsymbol{\xi}_k) \leq \mu - \alpha V_k(\boldsymbol{\xi}_k) \tag{3.5}$$

hold for every $k \geq 0$. Then the stochastic process $\boldsymbol{\xi}_k$ is exponentially bounded in mean square. Specifically,

$$\begin{aligned}\mathbb{E}(\|\boldsymbol{\xi}_k\|^2) &\leq \frac{v_2}{v_1}\mathbb{E}(\|\boldsymbol{\xi}_0\|^2)(1-\alpha)^k + \frac{\mu}{v_1}\sum_{i=0}^{k-1}(1-\alpha)^i \\ &\leq \frac{v_2}{v_1}\mathbb{E}(\|\boldsymbol{\xi}_0\|^2)(1-\alpha)^k + \frac{\mu}{v_1\alpha}\end{aligned}$$

for every $k \geq 0$. Furthermore, $\boldsymbol{\xi}_k$ is stochastically bounded.

Clearly $V_k(\boldsymbol{\xi}_k)$ is not generally a supermartingale. The meaning of (3.5) is then that if $V_k(\boldsymbol{\xi}_k)$ increases enough ($V_k(\boldsymbol{\xi}_k) \geq \mu/\alpha$), so that the left hand side of the inequality becomes non-positive, the supermartingale inequality is satisfied.

Proof of Lemma 3.8. (Tarn and Rasis, 1976, Theorem 2) The assumption (3.5) implies that

$$\mathbb{E}[V_{k+1}(\boldsymbol{\xi}_{k+1}) \mid \boldsymbol{\xi}_k] \leq \mu + (1-\alpha)V_k(\boldsymbol{\xi}_k)$$

and noting that

$$\mathbb{E}[V_k(\boldsymbol{\xi}_k) \mid \boldsymbol{\xi}_{k-2}] = \mathbb{E}[\mathbb{E}(V_k(\boldsymbol{\xi}_k) \mid \boldsymbol{\xi}_{k-1}) \mid \boldsymbol{\xi}_{k-2}],$$

it follows that

$$\begin{aligned}\mathbb{E}[V_k(\boldsymbol{\xi}_k) \mid \boldsymbol{\xi}_{k-2}] &\leq \mathbb{E}[\mu + (1-\alpha)V_{k-1}(\boldsymbol{\xi}_{k-1}) \mid \boldsymbol{\xi}_{k-2}] \\ &= \mu + (1-\alpha)\mathbb{E}[V_{k-1}(\boldsymbol{\xi}_{k-1}) \mid \boldsymbol{\xi}_{k-2}] \\ &\leq \mu + \mu(1-\alpha) + (1-\alpha)^2 V_{k-2}(\boldsymbol{\xi}_{k-2}).\end{aligned}$$

Continuing this finally results to

$$\mathbb{E}[V_k(\boldsymbol{\xi}_k) \mid \boldsymbol{\xi}_0] \leq (1-\alpha)^k V_0(\boldsymbol{\xi}_0) + \mu \sum_{i=0}^{k-1} (1-\alpha)^i$$

and taking of expectations yields

$$\mathbb{E}[V_k(\boldsymbol{\xi}_k)] \leq (1-\alpha)^k \mathbb{E}[V_0(\boldsymbol{\xi}_0)] + \mu \sum_{i=0}^{k-1} (1-\alpha)^i. \tag{3.6}$$

Then, using (3.4), one obtains

$$\mathbb{E}(\|\boldsymbol{\xi}_k\|^2) \leq \frac{v_2}{v_1} \mathbb{E}(\|\boldsymbol{\xi}_0\|^2)(1-\alpha)^k + \frac{\mu}{v_1} \sum_{i=0}^{k-1} (1-\alpha)^i.$$

To see that mean square boundedness implies stochastic boundedness, Chebyshev's inequality is used. According to it, for any $a > 0$,

$$\mathbb{P}\Big(\sqrt{V_k(\boldsymbol{\xi}_k)} \geq \sqrt{a}\Big) \leq \frac{\mathbb{E}[V_k(\boldsymbol{\xi}_k)]}{a}.$$

The inequalities (3.4) and (3.6) then yield

$$\mathbb{P}\Big(\|\boldsymbol{\xi}_k\| \geq \sqrt{a/v_2}\Big) \leq \frac{\mathbb{E}[V_0(\boldsymbol{\xi}_0)] + \mu/\alpha}{a}.$$

Therefore $\boldsymbol{\xi}_k$ is stochastically bounded. □

Usually, the much more powerful almost sure boundedness of the process is stated instead of stochastic boundedness. However, the proof cited for this fact (Agniel and Jury, 1971, Section 4.1, Theorem 1) may contain an error in asserting that the process $W_k(\boldsymbol{\xi}_k) := V_k(\boldsymbol{\xi}_k) - \mu/\alpha$ is supermartingale. Martingale convergence theorems (Loève, 1978, pp. 59–60) would then imply the almost sure boundedness of this process and hence that of $\boldsymbol{\xi}_k$. Stochastic sample path boundedness is claimed to hold in the analogous version for continuous-time processes (Reif et al., 2000).

In this lemma the stochastic process $V_k(\boldsymbol{\xi}_k)$ is a stochastic analogy, in some weak sense, of the deterministic Lyapunov function of the preceding section. Consequently, it will be called a *stochastic Lyapunov function.* As seen, the mean square of the process is completely determined by the mean square at the initial time-step and the constants v_1, v_2, μ and α. When the lemma is used in Section 5, some conditions, such as that $\|\boldsymbol{\xi}_0\| \leq \varepsilon$ for some positive ε, are imposed. Then the stability statements are of course to be understood in the sense of being restricted to the set of outcomes for which the conditions given are satisfied.

3.3 Stability of Filters with Respect to Initial Conditions

This section summarizes some stability results and methods used in the general optimal filtering theory of Markov processes. Nothing in this section will be used in the sequel; the sole purpose is to provide some context, especially for the treatment of the approximative non-linear filters. Good general surveys of these results can be found in Chigansky (2006); van Handel (2007) and Heine (2007).

In optimal filtering theory one is concerned with a discrete-time Markov process $(X_k)_{k\geq 0}$ and its observation process $(Y_k)_{k\geq 0}$ (such that the joint process is Markov) whose state spaces can be any suitable abstract spaces. The filtering densities are of course obtained from the Bayesian filtering equations of Theorem 2.1. As one does not always know the initial distribution of X_k, a central question becomes whether filters with distinct initial distributions converge towards each other with respect to some suitable metric. This is referred to as the *stability of the optimal filter with respect to its initial conditions.* The classical choice is the metric induced by the *total variation norm*

$\|\cdot\|_{TV}$ that is defined in the space of probability distributions (measures) $\mathcal{P}(E)$, where $(E, \mathcal{F})$ is some probability space, by setting

$$\|\mu - \nu\|_{TV} \coloneqq \sup_{A \in \mathcal{F}} |\mu(A) - \nu(A)| = \frac{1}{2} \int_E |p_\mu - p_\nu| \, \mathrm{d}\pi, \tag{3.7}$$

for $\mu, \nu \in \mathcal{P}(E)$. The functions p_μ, p_ν are the Radon–Nikodym derivatives of μ and ν with respect to some reference measure $\pi \in \mathcal{P}(E)$ that must of course be chosen such that μ and ν are absolutely continuous with respect to it. In the case $E = \mathbb{R}^n$ and probability distributions absolutely continuous with respect to the Lebesque measure this definition is spelled out as

$$\|\mu - \nu\|_{TV} = \frac{1}{2} \int_{\mathbb{R}^n} |p_\mu(\mathbf{x}) - p_\nu(\mathbf{x})| \, \mathrm{d}\mathbf{x}.$$

The filtering distribution with initial distribution μ corresponding to (2.1) at time-step k is customarily denoted by π_k^μ. So the problem is to find sufficient conditions for

$$\lim_{k \to \infty} \mathbb{E} \, \|\pi_k^\mu - \pi_k^\nu\|_{TV} = 0.$$

Under certain assumptions that include the ergodicity of X_k it can be shown that the above convergence holds and is exponential (Afar and Zeitouni, 1997). This is achieved by utilizing the *Hilbert projective metric* on the set $\mathcal{P}(E)$. The advantage of this metric, defined for $\mu, \nu \in \mathcal{P}(E)$ that satisfy a certain comparability condition, by

$$h(\mu, \nu) = \log \frac{\sup_{A \in \mathcal{F}, \nu(A) > 0} \mu(A)/\nu(A)}{\inf_{A \in \mathcal{F}, \nu(A) > 0} \mu(A)/\nu(A)},$$

is that it is invariant under scaling. This property simplifies the treatment of the update equation (2.5) as the normalization constant can be disregarded. Finally, the Hilbert projective metric is linked to the total variation norm by the inequality

$$\|\mu - \nu\|_{TV} \leq \frac{2}{\log 3} h(\mu, \nu),$$

that is, convergence with respect to the Hilbert projective metric implies convergence with respect to the total variation norm. However, the use of the Hilbert projective metric usually requires a compact state space or that the Markov transition kernel of the filtering process satisfy certain restrictive mixing condition.

Another approach is to use the *Dobrushin ergodic coefficient* (Del Moral and Guionnet, 2001). For Markov transition kernels K on E the Dobrushin ergodic coefficient $\alpha(K)$ is defined as

$$\alpha(K) \coloneqq 1 - \sup |K(x, A) - K(y, A)| \, ,$$

where the supremum is taken over all $x, y \in E$ and $A \in \mathcal{F}$. Unfortunately, also this technique is dependent on mixing conditions of some form.

4 Stability of the Linear Kalman Filter

In this section the stability of Kalman filters for linear dynamic systems is treated. A *discrete-time linear dynamic system* disturbed by zero-mean Gaussian noise processes can be written as

$$\begin{aligned} \mathbf{x}_k &= \mathbf{A}_{k-1}\mathbf{x}_{k-1} + \mathbf{q}_{k-1}, \\ \mathbf{y}_k &= \mathbf{H}_k\mathbf{x}_k + \mathbf{r}_k, \end{aligned} \tag{4.1}$$

where $\mathbf{x}_k \in \mathbb{R}^n$ is the system state, $\mathbf{y}_k \in \mathbb{R}^m$ is the *measurement*, $\mathbf{q}_{k-1} \sim \mathcal{N}(\mathbf{0}, \mathbf{Q}_{k-1})$ is the *process noise* and $\mathbf{r}_k \sim \mathcal{N}(\mathbf{0}, \mathbf{R}_k)$ is the *measurement noise*. The system is initialized from $k = 0$ with *initial system state distribution* $\mathbf{x}_0 \sim \mathcal{N}(\mathbf{m}_0, \mathbf{P}_0)$. The noise processes are assumed uncorrelated and independent of the initial state. The matrices $\mathbf{A}_{k-1}$ are *dynamic model matrices* and $\mathbf{H}_k$ *measurement model matrices.* The state transition matrix of $\mathbf{x}_k$ is denoted as in the preceding section. Figure 4.1 illustrates the workings of this system. In probabilistic terms the model can be written as

$$\begin{aligned} p(\mathbf{x}_k \mid \mathbf{x}_{k-1}) &= \mathcal{N}(\mathbf{x}_k \mid \mathbf{A}_{k-1}\mathbf{x}_{k-1}, \mathbf{Q}_{k-1}), \\ p(\mathbf{y}_k \mid \mathbf{x}_k) &= \mathcal{N}(\mathbf{y}_k \mid \mathbf{H}_k\mathbf{x}_k, \mathbf{R}_k). \end{aligned} \tag{4.2}$$

Many linear dynamic systems of interest have constant dynamic and measurement model matrices as well as constant noise covariance matrices. Such systems are called *time-invariant* and for them strong stability and convergence results are obtainable with significantly reduced complexity (Anderson and Moore, 1979, Chapter 4).

Often a deterministic *control signal* $\mathbf{u}_{k-1}$ is added to the state part of (4.1). This term is omitted here as it does not affect the stability analysis and would simply lengthen the already lengthy formulas. Also, it is not imperative to have the noise processes uncorrelated with each other, but this too would make the notation more cumbersome.

4.1 The Kalman Filter Equations

For linear dynamic systems one has the celebrated Kalman filter equations that provide closed form solutions to the Bayesian filtering equations (2.4) and (2.5).

Theorem 4.1 (Kalman filter)**.** *The Bayesian filtering equations of Theorem 2.1 for linear dynamic system* (4.1) *can be evaluated in closed form.*

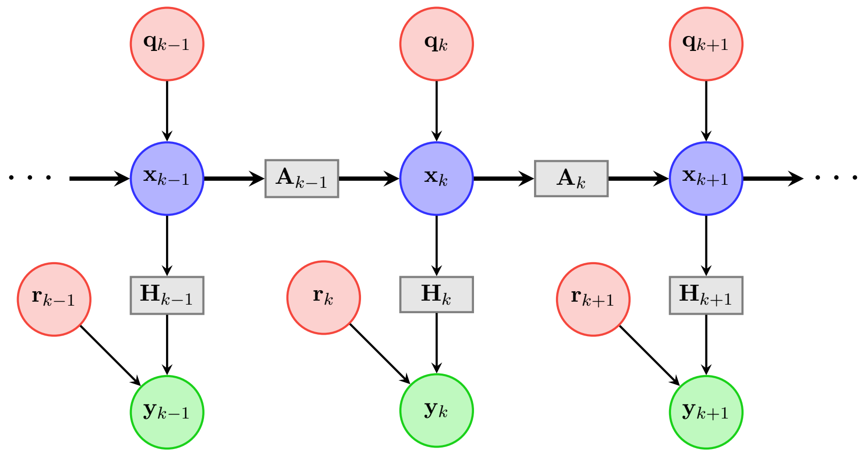

Figure 4.1: Diagram of the discrete-time dynamic system (4.1). At each time-step k the previous system state $\mathbf{x}_{k-1}$ is propagated through matrix $\mathbf{A}_{k-1}$ and augmented with a noise term $\mathbf{q}_{k-1}$ to obtain the next state $\mathbf{x}_k$. However, the state is not directly observed but measurements are instead received of $\mathbf{x}_k$ by a linear transformation by $\mathbf{H}_k$ with an added unknown noise term $\mathbf{r}_k$.

The resulting distributions are

$$\begin{aligned} p(\mathbf{x}_k \mid \mathbf{y}_{1:k-1}) &= \mathcal{N}(\mathbf{x}_k \mid \mathbf{m}_k^-, \mathbf{P}_k^-), \\ p(\mathbf{x}_k \mid \mathbf{y}_{1:k}) &= \mathcal{N}(\mathbf{x}_k \mid \mathbf{m}_k, \mathbf{P}_k), \end{aligned}$$

where the parameters are calculated with the following prediction step

$$\begin{aligned} \mathbf{m}_k^- &= \mathbf{A}_{k-1}\mathbf{m}_{k-1}, \\ \mathbf{P}_k^- &= \mathbf{A}_{k-1}\mathbf{P}_{k-1}\mathbf{A}_{k-1}^\mathsf{T} + \mathbf{Q}_{k-1} \end{aligned} \tag{4.3}$$

and the update step

$$\mathbf{v}_k = \mathbf{y}_k - \mathbf{H}_k\mathbf{m}_k^-,$$

$$\mathbf{S}_k = \mathbf{H}_k\mathbf{P}_k^-\mathbf{H}_k^\mathsf{T} + \mathbf{R}_k, \tag{4.4}$$

$$\mathbf{K}_k = \mathbf{P}_k^-\mathbf{H}_k^\mathsf{T}\mathbf{S}_k^{-1}, \tag{4.5}$$

$$\mathbf{m}_k = \mathbf{m}_k^- + \mathbf{K}_k\mathbf{v}_k, \tag{4.6}$$

$$\mathbf{P}_k = \mathbf{P}_k^- - \mathbf{K}_k\mathbf{S}_k\mathbf{K}_k^\mathsf{T}. \tag{4.7}$$

Proof. (See Appendix A.2) □

In the above algorithm the matrix $\mathbf{K}_k$ is called *Kalman gain matrix*. The vector $\mathbf{m}_k$ is the optimal mean-square estimate of the state $\mathbf{x}_k$ and the matrix $\mathbf{P}_k$ is the covariance matrix of *estimation error* $\boldsymbol{\xi}_k := \mathbf{x}_k - \mathbf{m}_k$. The vector $\mathbf{m}_k^-$ is the *predicted state estimate* and $\mathbf{P}_k^-$ the *predicted error covariance*, which are obtained by propagating the state through the system dynamics but not yet conditioning it on the latest measurement. The vector $\mathbf{v}_k$ is sometimes called the *innovation* and $\mathbf{S}_k$ is the covariance matrix of the innovation. The recursion is started from the *initial estimate* $\mathbf{m}_0$ and *initial error covariance* $\mathbf{P}_0$.

The discrete-time Kalman filter equations were first derived by Kalman (1960b). In his paper, Kalman used the concept of orthogonal projections by considering the space spanned by measurements and observing that the optimal estimate (in the sense of minimizing the quadratic loss function) of the current state is the orthogonal projection of the state into this space. Unlike the derivation presented in the appendices, that of Kalman does not require Gaussian noise processes — them having mean zero is enough. However, such a relaxation of assumptions provides one only with the optimal estimate and error covariance, not closed form solutions to the Bayesian filtering equations as with the Gaussian assumption. Jazwinski (1970, Chapter 7) gives a wealth of other possible derivations. Yet, it is important to have a Gaussian initial distribution as with a non-Gaussian initial distribution the Kalman filter is not the optimal filter. For the much more complicated form of the filtering distribution in this case, see Beneš and Karatzas (1983, Theorems 4.1 and 5.1). Some stability results concerning such a situation are referenced at the end of Section 4.5.

The state estimates $\mathbf{m}_k$ of the Kalman filter constitute a linear dynamic system of their own, since Equation (4.6) can easily be modified into a recursive form

$$\mathbf{m}_k = (\mathbf{I} - \mathbf{K}_k\mathbf{H}_k)\mathbf{A}_{k-1}\mathbf{m}_{k-1} + \mathbf{K}_k\mathbf{y}_k. \tag{4.8}$$

For the predicted state estimates the analogous recursive form is

$$\mathbf{m}_{k+1}^- = \mathbf{A}_k(\mathbf{I} - \mathbf{K}_k\mathbf{H}_k)\mathbf{m}_k^- + \mathbf{A}_k\mathbf{K}_k\mathbf{y}_k, \tag{4.9}$$

and the relationship between these estimates can be written as

$$\mathbf{m}_k = (\mathbf{I} - \mathbf{K}_k\mathbf{H}_k)\mathbf{m}_k^- + \mathbf{K}_k\mathbf{y}_k. \tag{4.10}$$

An important property of the Kalman filter is that it is the *linear minimum variance state estimator* (Anderson and Moore, 1979, Chapter 5). This means that among the state estimators given by a linear system as that above the Kalman filter has the smallest error covariance, that is, any $\overline{\mathbf{m}}_k$ used to estimate $\mathbf{x}_k$ and obtained recursively by $\overline{\mathbf{m}}_k = \mathbf{F}_{k-1}\overline{\mathbf{m}}_{k-1} + \mathbf{z}_{k-1}$ for some $\mathbf{F}_{k-1}$ and $\mathbf{z}_{k-1}$ satisfies $\mathbb{E}[(\mathbf{x}_k - \overline{\mathbf{m}}_k)(\mathbf{x}_k - \overline{\mathbf{m}}_k)^\mathsf{T}] \geq \mathbf{P}_k$. This property is an essential one in proving the stability of the Kalman filter.

The stability of the Kalman filter is to be understood as the stability of the homogeneous part of the dynamic system (4.8). Clearly one cannot obtain a result to the end that $\|\mathbf{x}_k - \mathbf{m}_k\| \to 0$ unless the measurement noise vanishes. The best one could hope for then is that this difference is dominated by the measurements in the long run, which means that the homogeneous part must decay. Rigorously,

$$\begin{aligned} \boldsymbol{\xi}_k = \mathbf{x}_k - \mathbf{m}_k &= \mathbf{A}_{k-1}(\mathbf{x}_{k-1} - \mathbf{m}_{k-1}) - \mathbf{K}_k \mathbf{v}_k + \mathbf{q}_{k-1} \\ &= (\mathbf{I} - \mathbf{K}_k \mathbf{H}_k)\mathbf{A}_{k-1}\boldsymbol{\xi}_{k-1} + (\mathbf{I} - \mathbf{K}_k \mathbf{H}_k)\mathbf{q}_{k-1} - \mathbf{K}_k \mathbf{r}_k, \end{aligned} \tag{4.11}$$

and so it is seen that the expected estimation error is

$$\mathbb{E}(\boldsymbol{\xi}_k) = (\mathbf{I} - \mathbf{K}_k \mathbf{H}_k)\mathbf{A}_{k-1}\mathbb{E}(\boldsymbol{\xi}_{k-1}),$$

and hence governed by exactly the same dynamics as the homogeneous part of (4.8).

Reformulation of Equation (4.3) for $\mathbf{P}_k^-$ into a recursive form gives rise to a *discrete-time Riccati equation*

$$\mathbf{P}_{k+1}^- = \mathbf{A}_k \mathbf{P}_k^- \mathbf{A}_k^\mathsf{T} - (\mathbf{A}_k \mathbf{P}_k^- \mathbf{H}_k^\mathsf{T})(\mathbf{H}_k \mathbf{P}_k^- \mathbf{H}_k^\mathsf{T} + \mathbf{R}_k)^{-1}(\mathbf{H}_k \mathbf{P}_k^- \mathbf{A}_k^\mathsf{T}) + \mathbf{Q}_k. \tag{4.12}$$

There exists an extensive theory covering the behavior of equations of this type partly motivated by the prevalence of the equation in filtering theory. The time-invariant Riccati equation is comprehensively covered in Lancaster and Rodman (1995). More commentary and results, with the focus on the context of filtering, can be found in De Nicolao (1992); Ahlbrandt and Heifetz (1995) and Costa and Astolfi (2008). For the purposes of a stability analysis a more useful recursive form of the predicted error covariance matrix is

$$\mathbf{P}_{k+1}^- = \mathbf{A}_k(\mathbf{I} - \mathbf{K}_k \mathbf{H}_k)\mathbf{P}_k^-(\mathbf{I} - \mathbf{K}_k \mathbf{H}_k)^\mathsf{T}\mathbf{A}_k^\mathsf{T} + \mathbf{Q}_k + \mathbf{A}_k \mathbf{K}_k \mathbf{R}_k \mathbf{K}_k^\mathsf{T} \mathbf{A}_k^\mathsf{T}, \tag{4.13}$$

readily derived from the definition of this matrix and the Kalman filter equations. As for the updated error covariance matrix, the matrix inversion lemma (found as Lemma A.1 in the appendices) yields a useful form

$$\mathbf{P}_k = \left[(\mathbf{P}_k^-)^{-1} + \mathbf{H}_k^\mathsf{T} \mathbf{R}_k^{-1} \mathbf{H}_k\right]^{-1}, \tag{4.14}$$

inverse of which is known as the *information matrix* (Anderson and Moore, 1979, Section 6.3). As can be seen, the error covariance matrices do not depend on the measurements and therefore, being completely deterministic, can be calculated offline before the filter is being run. Another relation for the error covariance matrices is the *Joseph form*

$$\mathbf{P}_k = (\mathbf{I} - \mathbf{K}_k\mathbf{H}_k)\mathbf{P}_k^-(\mathbf{I} - \mathbf{K}_k\mathbf{H}_k)^\top + \mathbf{K}_k\mathbf{R}_k\mathbf{K}_k^\top. \tag{4.15}$$

Related to Riccati equation is the slightly simpler *Lyapunov equation* (Gajić and Qureshi, 2008) that is encountered for example when the measurement noise is singular (Halevi, 1989). An example of this equation is (4.23) and its properties will be exploited in Theorem 4.7 that is of central importance in proving the most general results on the stability of the Kalman filter.

4.2 Some Concepts from Control Theory

In control theory one frequently employs the concepts of *observability* and *controllability* and their generalizations *detectability* and *stabilizability*. Assuming either pair of these two suffices for the exponential stability of the Kalman filter. In the following all the matrices are assumed to be of appropriate dimensions.

Definition 4.2 (Observability)**.** The pair $[\mathbf{F}_k, \mathbf{G}_k]$ of matrices for $k > 0$ is said to be *completely observable* if the *observability Gramian*

$$\boldsymbol{\mathcal{O}}_{k,l} := \sum_{i=l}^{k} \left(\prod_{j=l}^{i-1} \mathbf{F}_j \right)^\top \mathbf{G}_i^\top \mathbf{G}_i \left(\prod_{j=l}^{i-1} \mathbf{F}_j \right), \tag{4.16}$$

defined for $l < k$, is positive-definite for some such k and l. Furthermore, the pair is said to be *uniformly observable* (or *uniformly completely observable*) if there exists a positive integer N and positive constants α_1 and α_2 such that

$$\mathbf{0} \leq \alpha_1 \mathbf{I} \leq \boldsymbol{\mathcal{O}}_{k,k-N} \leq \alpha_2 \mathbf{I}$$

for all $k \geq N$. Complete or uniform observability of the system (4.1) means that of the pair $[\mathbf{A}_k, \mathbf{R}_k^{-1/2}\mathbf{H}_k]$. In this case, the observability Gramian is simply

$$\boldsymbol{\mathcal{O}}_{k,l} = \sum_{i=l}^{k} \mathbf{\Phi}_{i,l}^\top \mathbf{H}_i^\top \mathbf{R}^{-1} \mathbf{H}_i \mathbf{\Phi}_{i,l}.$$

Definition 4.3 (Controllability)**.** The pair $[\mathbf{F}_k, \mathbf{G}_k]$ of matrices for $k > 0$ is said to be *completely controllable* if the *controllability Gramian* (or *control-*

lability matrix)

$$\mathcal{C}_{k,l} \coloneqq \sum_{i=l}^{k-1} \Big(\prod_{j=i+1}^{k-1} \mathbf{F}_j \Big) \mathbf{G}_i \mathbf{G}_i^{\mathsf{T}} \Big(\prod_{j=i+1}^{k-1} \mathbf{F}_j \Big)^{\mathsf{T}}, \tag{4.17}$$

defined for $l < k$, is positive-definite for some such k and l. Furthermore, the pair is said to be *uniformly controllable* (or *uniformly completely controllable*) if there exists a positive integer N and positive constants β_1 and β_2 such that

$$\mathbf{0} \leq \beta_1 \mathbf{I} \leq \mathcal{C}_{k,k-N} \leq \beta_2 \mathbf{I}$$

for all $k \geq N$. Complete or uniform controllability of the system (4.1) means that of the pair $[\mathbf{A}_k, \mathbf{Q}_k^{1/2}]$. In this case, the controllability Gramian is simply

$$\mathcal{C}_{k,l} \coloneqq \sum_{i=l}^{k-1} \mathbf{\Phi}_{k,i+1} \mathbf{Q}_i \mathbf{\Phi}_{k,i+1}^{\mathsf{T}}.$$

A related concept to the observability Gramian is the *observability matrix* that is defined when $\mathbf{F}_k$ are non-singular by

$$\widetilde{\mathcal{O}}_{k,l} \coloneqq \sum_{i=l}^{k} \Big(\prod_{j=i}^{k-1} \mathbf{F}_j \Big)^{-\mathsf{T}} \mathbf{G}_i^{\mathsf{T}} \mathbf{G}_i \Big(\prod_{j=i}^{k-1} \mathbf{F}_j \Big)^{-1}. \tag{4.18}$$

For the system (4.1) this is

$$\widetilde{\mathcal{O}}_{k,l} = \sum_{i=l}^{k} \mathbf{\Phi}_{l,i}^{\mathsf{T}} \mathbf{H}_i^{\mathsf{T}} \mathbf{R}^{-1} \mathbf{H}_i \mathbf{\Phi}_{l,i}.$$

This matrix is closely related to the observability Gramian as one is able to observe that (Aoki, 1967, p. 215)

$$\mathcal{O}_{k,l} = \Big(\prod_{j=l}^{k-1} \mathbf{F}_j \Big)^{\mathsf{T}} \widetilde{\mathcal{O}}_{k,l} \Big(\prod_{j=l}^{k-1} \mathbf{F}_j \Big),$$

which implies that if $\mathbf{F}_k$ is uniformly bounded from above, one can equivalently use the observability matrix in the definitions of complete observability and uniform observability. In the literature this connection is rarely made explicit and depending on the authors the results are formulated for either the observability Gramian or the observability matrix.

The discussion on the meaning of these notions is postponed until their generalizations detectability and stabilizability have been presented.

Definition 4.4 (Detectability)**.** The pair $[\mathbf{F}_k, \mathbf{G}_k]$ of matrices for $k > 0$ is said to be *uniformly detectable* if there are non-negative integers s and t with $s \geq t$ and constants $0 \leq d < 1$ and $b > 0$ such that when

$$\left\| \prod_{i=k}^{k+t-1} \mathbf{F}_i \mathbf{z} \right\| \geq d \|\mathbf{z}\| \tag{4.19}$$

for some vector $\mathbf{z}$ and k, then

$$\mathbf{z}^\mathsf{T} \boldsymbol{\mathcal{O}}_{k+s,k} \mathbf{z} \geq b \mathbf{z}^\mathsf{T} \mathbf{z}.$$

Similarly to observability, uniform detectability of the system (4.1) means that of the pair $[\mathbf{A}_k, \mathbf{R}_k^{-1/2} \mathbf{H}_k]$ and in this case (4.19) is

$$\|\mathbf{\Phi}_{k+t-1,k} \mathbf{z}\| \geq d \|\mathbf{z}\|. \tag{4.20}$$

Definition 4.5 (Stabilizability)**.** The pair $[\mathbf{F}_k, \mathbf{G}_k]$ of matrices for $k > 0$ is said to be *uniformly stabilizable* if there are non-negative integers s and t with $s \geq t$ and constants $0 \leq d < 1$ and $b > 0$ such that when

$$\left\| \prod_{i=k+1-t}^{k} \mathbf{F}_i \mathbf{z} \right\| \geq d \|\mathbf{z}\| \tag{4.21}$$

for some vector $\mathbf{z}$ and k, then

$$\mathbf{z}^\mathsf{T} \boldsymbol{\mathcal{C}}_{k,k-s} \mathbf{z} \geq b \mathbf{z}^\mathsf{T} \mathbf{z}.$$

Similarly to controllability, uniform stabilizability of the system (4.1) means that of the pair $[\mathbf{A}_k, \mathbf{Q}_k^{1/2}]$, (4.21) having a form analogous to (4.20).

Uniform observability clearly implies the uniform detectability, for then by the definition there are positive integer N and positive scalar α_1 such that

$$\boldsymbol{\mathcal{O}}_{k,k-N} - \alpha_1 \mathbf{I} \geq \mathbf{0}$$

for all $k \geq N$. For any vector $\mathbf{z}$ multiplication from left by $\mathbf{z}^\mathsf{T}$ and right by $\mathbf{z}$ produces, by the definition of positive-semidefiniteness,

$$\mathbf{z}^\mathsf{T} \boldsymbol{\mathcal{O}}_{k,k-N} \mathbf{z} \geq \alpha_1 \mathbf{z}^\mathsf{T} \mathbf{z}.$$

An identical argument establishes that uniform controllability implies uniform stabilizability.

The notions of controllability and observability were introduced by

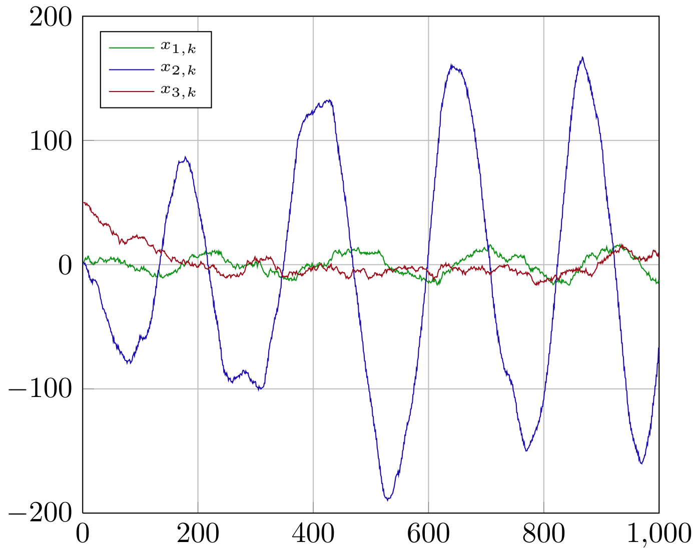

Figure 4.2: An example a linear system uniformly detectable but not uniformly observable with dynamics (4.22). Depicted is a sample path with $\mathbf{x}_0 = (0, 0, 50)$. It is seen that this system is not uniformly observable because the observability Gramian is never positive-definite. This is clear from the choice of measurement model matrix as no information is carried in the measurements about the third component $\mathbf{x}_{k,3}$ of the system state. However, this component decays exponentially to zero, as can be easily seen, and so it is intuitive that the system should be uniformly detectable. Through some rather straightforward calculations it can be verified that this indeed is the case.

Kalman (1960a) and further elaborated in Kalman (1961). Detectability and stabilizability were first defined for time-invariant systems only (Wonham, 1967, 1968) but the definitions were soon extended to cover also time-varying systems (Hager and Horowitz, 1976; Anderson and Moore, 1981). In addition to these articles, much discussion and interpretation on these four notions can be found in Kwakernaak and Sivan (1972).

Complete controllability of the system (3.1) means that the system can be transferred from the initial state $\mathbf{x}_0$ to any possible state in a finite number of time-steps with a suitable construction of the input function $\mathbf{u}_{k-1}$. A system with measurement part is completely observable if its state can be inferred from the measurements. These physical aspects are more discussed in, for example, Kalman et al. (1963).

Detectability and stabilizability generalize observability and controllability in that they do not require all the modes of the system be observable or controllable, but all such modes must be stable so that insufficient knowledge or inability to control them causes no complications. In Figure 4.2 this distinction is demonstrated for observability and detectability. The system is (4.1) with dynamic model matrices

$$\mathbf{A}_{2k} = \begin{pmatrix} 1.001 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0.99 \end{pmatrix}, \quad \mathbf{A}_{2k+1} = \begin{pmatrix} 1.001 & 0.005 & 0 \\ -0.7 & 1 & 0 \\ 0 & 0 & 0.99 \end{pmatrix} \tag{4.22}$$

and measurement model matrix

$$\mathbf{H}_k = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}.$$

The noise terms are set to be distributed according to $\mathcal{N}(\mathbf{0}, \mathbf{I})$ and the system is initialized from $\mathbf{x}_0 = (0, 0, 50)$.

The following property of uniform stabilizability, called *invariance under feedback*, plays an important part in the proof Kalman filter stability.

Theorem 4.6 (Invariance under feedback)**.** *Suppose that the matrices* $\mathbf{F}_k, \mathbf{G}_k$ *and* $\mathbf{T}_k$ *are uniformly bounded. Then the pair* $[\mathbf{F}_k, \mathbf{G}_k]$ *is uniformly stabilizable if and only if the pair* $[\mathbf{F}_k + \mathbf{G}_k\mathbf{T}_k, \mathbf{G}_k]$ *is.*

Proof. (Anderson and Moore, 1981, Section 3) □

The next theorem is a specific kind of generalization of Lyapunov stability theorem. It gives a strong connection between the boundedness of the solutions of a Lyapunov equation and the exponential stability of certain system.

Theorem 4.7. *Suppose that the pair* $[\mathbf{F}_k, \mathbf{G}_k]$ *is uniformly stabilizable, that* $\mathbf{F}_k$ *and* $\mathbf{G}_k$ *are uniformly bounded and that there exists a bounded sequence* $\mathbf{\Pi}_k$ *of positive-semidefinite matrices satisfying the Lyapunov equation*

$$\mathbf{\Pi}_{k+1} = \mathbf{F}_k\mathbf{\Pi}_k\mathbf{F}_k^\top + \mathbf{G}_k\mathbf{G}_k^\top \tag{4.23}$$

for $k \geq 0$*. Then the system* $\mathbf{z}_{k+1} = \mathbf{F}_k\mathbf{z}_k$ *is exponentially stable.*

Proof. (Anderson and Moore, 1981, Theorems 4.2 and 4.3) □

There also exist versions of the two preceding theorems with uniform stabilizability replaced by uniform detectability.

For time-invariant systems there exist simple matrix rank conditions for uniform observability and controllability. The equivalence of these rank conditions to the ones previously given seems not to have been clearly stated in the literature but follows rather uncomplicatedly. Some versions can be found in (Lancaster and Rodman, 1995, Corollary 4.3.2; Bucy and Joseph, 1968, Corollaries 3.1 and 3.2).

Theorem 4.8. *Let* $\mathbf{F} \in \mathbb{R}^{p\times p}$*. The pair* $[\mathbf{F}, \mathbf{G}]$ *is uniformly observable if and only if it satisfies the* observability rank condition *according to which the matrix*

$$\boldsymbol{\mathcal{O}} := \begin{pmatrix} \mathbf{G} \\ \mathbf{GF} \\ \mathbf{GF}^2 \\ \vdots \\ \mathbf{GF}^{n-1} \end{pmatrix}$$

has rank p*. The pair* $[\mathbf{F}, \mathbf{G}]$ *is uniformly controllable if and only if it satisfies the* controllability rank condition *according to which the matrix*

$$\boldsymbol{\mathcal{C}} := \begin{pmatrix} \mathbf{G} & \mathbf{GF} & \mathbf{G}^2\mathbf{F} & \cdots & \mathbf{F}^{n-1}\mathbf{G} \end{pmatrix}$$

has rank p*.*

Proof. The equivalences follow from the remarks that $\boldsymbol{\mathcal{O}}_{k,k-p} = \boldsymbol{\mathcal{O}}^\mathsf{T}\boldsymbol{\mathcal{O}}$ and $\boldsymbol{\mathcal{C}}_{k,k-p} = \boldsymbol{\mathcal{C}}\boldsymbol{\mathcal{C}}^\mathsf{T}$ and the use of the properties of matrix rank. □

A similar condition will be encountered in the study of non-linear Kalman filters.

4.3 Overview of Linear Stability Results

The main result of this section is Theorem 4.14 that asserts that the exponential stability of the Kalman filter is guaranteed if the dynamic system is uniformly detectable and stabilizable. This result was established by Anderson and Moore (1981) but was preceded by one requiring the somewhat more restrictive assumptions of uniform observability and controllability (Deyst and Price, 1968). Despite this earlier result being less general its proof is also presented. This is mainly for two reasons. The first is that it is interesting to see the development of the reasoning used as tedious matrix manipulations filling multiple pages are replaced by arguments surprisingly straightforward and more general (albeit less trivial in many aspects). The second reason, equally interesting, is that the original proof of Deyst and Price (1968) contained a small error, only affecting certain bounds for the error covariance matrices, but still sometimes overlooked despite having been pointed out by Hitz (1970).

Now a brief sketch of the stability proofs is presented. The proofs are intimately connected with the time-varying Riccati equation (4.12) for the error covariance matrix. It is possible to prove that the solution to time-invariant Riccati equation converges to a certain matrix. For the time-varying version the situation is not that simple, but nevertheless under uniform detectability and stabilizability it can be shown that the solution is uniformly bounded. The proof of this relies on the linear minimum variance property of the Kalman filter. One can construct such a sub-optimal state estimator that the associated error covariance matrix is seen to admit an upper bound which then is also the upper bound for the error covariance of the Kalman filter. The construction of such an estimator by Anderson and Moore (1981) is the most complicated part of the stability proof and a task far from trivial as even the authors themselves note. The original sub-optimal estimator of Deyst and Price (1968) is much simpler and even provides one with an explicit upper bound. This was the part in the original proof that contained an error. The boundedness of the error covariance matrix is then used to show that a certain function constructed out of it is a Lyapunov function from which the exponential stability of the homogeneous part of the filter equation (4.8) can be concluded.

Next follows a historical account on the development of these stability results, beginning from the conception of the Kalman filter by Kalman (1960b) and culminating in the generalized results of Anderson and Moore (1981).

Unlike the continuous-time Kalman–Bucy filter for which stability results existed from the beginning (Kalman and Bucy, 1961), it took some some time for the analogous results to develop in the discrete-time case for systems uniformly observable and controllable. The first attempt to establish bounds for the error covariance matrix was by Sorenson (1967) who used a decomposition property of linear systems, but his bounds were not of much use in proving the stability of the filter. The first proof was finally given by Deyst and Price (1968) and incorporated into the influential book by Jazwinski (1970). However, Deyst and Price (1968) committed a small error, duplicated by Jazwinski (1970), in the proof of an upper bound for the error covariance matrix, pointed out in Hitz (1970, Section 2.2) and Hitz et al. (1972). Following the exposure of this error, Tse (1973) was quick to pronounce the stability an open question. The error does not invalidate the stability results; Deyst (1973) keenly provided a correction that supplied a different upper bound. Some, but not all, authors seem to be aware of the error and Jazwinski (1970) is frequently cited as the only source on the subject. Discussion or references about the error are included, as a part of presentation of the stability of the Kalman filter, in, for example, Maybeck (1979, Section 5.8) and Rhudy and Gu (2013). In Kim et al. (2006, Appendix A) a very terse version of the correct proof is given. Kerr (1985, Section V) warns about this error among other problems and flaws in the field one should be cautious about. The early proofs (Kalman, 1963; Bucy and Joseph, 1968, Theorem 5.4) of the stability of the continuous-time linear optimal filter, the *Kalman-Bucy filter* developed by Kalman and Bucy (1961), contained an analogous error in the bounds for the error covariance matrix and it is notable that a corrected version was not provided until by Delyon (2001).

During the 1970s there emerged a desire to generalize the results for uniformly detectable and stabilizable systems as it is quite intuitive that unobservable or uncontrollable states that decay fast should have no effect on the stability of the filter. Although not directly to this end, Anderson (1971) was able to pinpoint conditions for stability weaker than exponential. In the beginning of the 1980s the assumption of the non-singularity of the state transition matrices of the dynamic system (having been assumed in all the preceeding proofs) was first discarded (Moore and Anderson, 1980) and then uniform observability and controllability of the system were weakened to uniform detectability and stabilizability (Anderson and Moore, 1981), proofs being more direct than previously. Curiously, Engwerda (1990) claimed that this proof too was erroneous by constructing an alleged counterexample that turned out not to be uniformly detectable (Anderson and Moore, 1992; Bittanti et al., 1992).

In the expository presentation of the earlier results, Kamen and Su (1999, Appendix C) is followed. Kamen and Su (1999) in turn follow Jazwinski (1970, Chapter 7); Deyst and Price (1968) and Deyst (1973). The more general proof is solely taken from Anderson and Moore (1981).

The Kalman filter being the optimal filter for linear systems, it is no wonder that is possesses these strong stability properties. Nevertheless, it is of course necessary to actually prove that these properties hold instead of merely assuming stability would follow trivially from optimality as has sometimes been done in a similar setting (Kalman, 1960a, p. 160). Although the Kalman filter should usually be stable, modeling and numerical errors can easily cause filters to diverge (Schlee et al., 1967; Fitzgerald, 1971). This is a matter outside the scope of this thesis.

4.4 Original Stability Proof

This section presents the original exponential stability results of the Kalman filter in the corrected form. In this section it is assumed that $\mathbf{A}_{k-1}$ is non-singular and uniformly bounded. The non-singularity of the dynamic model matrices is required in several matrix manipulations of this original proof. The proof utilizes the observability matrix instead of observability Gramian and the boundedness assumption guarantees the equivalence of using either of these matrices. Originally this boundedness assumption was missing and the proof done using a definition of uniform observability that had the observability Gramian replaced with the observability matrix (4.18).

One begins the proof by obtaining bounds for the error covariance matrix via construction of a sub-optimal linear estimate of the system state. The proofs are given here rather sketchily so as to avoid lengthy chains of matrix manipulations and present only the core idea of the proofs.

Lemma 4.9. *If the dynamic system* (4.1) *is uniformly observable and uniformly controllable, then*

$$\mathbf{P}_k \leq \left(\frac{\alpha_1 + N\alpha_2^2\beta_2}{\alpha_1^2} \right) \mathbf{I}$$

for all $k \geq N$.

Proof. (Deyst, 1973; Kamen and Su, 1999, Lemma C.2) The proof relies on the linear minimum variance property of the Kalman filter. For $k \geq N$, a state estimate

$$\widetilde{\mathbf{m}}_k = \widetilde{\boldsymbol{\mathcal{O}}}_{k,k-N}^{-1} \sum_{i=k-N}^{k} \boldsymbol{\Phi}_{i,k}^\mathsf{T} \mathbf{H}_i^\mathsf{T} \mathbf{R}_i^{-1} \mathbf{y}_i \tag{4.24}$$

is defined and since this estimate is not the optimal one and thus not of minimum variance, one has the inequality

$$\mathbf{P}_k \leq \mathrm{Cov}(\mathbf{m}_k - \widetilde{\mathbf{m}}_k).$$

Through somewhat tedious evaluations, the error covariance matrix for this sub-optimal estimate is reduced to

$$\begin{aligned} \mathrm{Cov}(\mathbf{m}_k - \widetilde{\mathbf{m}}_k) = {}& \widetilde{\boldsymbol{\mathcal{O}}}_{k,k-N}^{-1} + \widetilde{\boldsymbol{\mathcal{O}}}_{k,k-N}^{-1} \sum_{j=k-N}^{k-1} \left(\sum_{i=k-N}^{j} \boldsymbol{\Phi}_{i,k}^\mathsf{T} \mathbf{H}_i^\mathsf{T} \mathbf{R}_i^{-1} \mathbf{H}_i \boldsymbol{\Phi}_{i,k} \right) \\ & \times \boldsymbol{\Phi}_{k,j+1} \mathbf{Q}_j \boldsymbol{\Phi}_{k,j+1}^\mathsf{T} \left(\sum_{i=k-N}^{j} \boldsymbol{\Phi}_{i,k}^\mathsf{T} \mathbf{H}_i^\mathsf{T} \mathbf{R}_i^{-1} \mathbf{H}_i \boldsymbol{\Phi}_{i,k} \right) \widetilde{\boldsymbol{\mathcal{O}}}_{k,k-N}^{-1}. \end{aligned}$$

For this, and thus for $\mathbf{P}_k$, an upper bound is obtained owing to the assumptions of uniform complete observability and uniform complete controllability. □

Originally, an error was committed by Deyst and Price (1968) in the purported proof of the above lemma. They claimed a particularly simple upper bound, namely

$$\mathbf{P}_k \leq \widetilde{\boldsymbol{\mathcal{O}}}_{k,k-N}^{-1} + \boldsymbol{\mathcal{C}}_{k,k-N}, \tag{4.25}$$

but in course of the proof committed an error, pointed out in Hitz (1970, Section 2.2) and Hitz et al. (1972), in asserting that

$$\begin{aligned}
&\operatorname{Cov}\left(\widetilde{\boldsymbol{\mathcal{O}}}_{k,k-N}^{-1} \sum_{i=k-N}^{k} \boldsymbol{\Phi}_{i,k}^{\mathsf{T}} \mathbf{H}_i^{\mathsf{T}} \mathbf{R}_i^{-1} \mathbf{H}_i \boldsymbol{\Phi}_{i,k} \sum_{j=i}^{k-1} \boldsymbol{\Phi}_{k,j+1} \mathbf{Q}_j\right) \\
&\qquad \leq \operatorname{Cov}\left(\widetilde{\boldsymbol{\mathcal{O}}}_{k,k-N}^{-1} \sum_{i=k-N}^{k} \boldsymbol{\Phi}_{i,k}^{\mathsf{T}} \mathbf{H}_i^{\mathsf{T}} \mathbf{R}_i^{-1} \mathbf{H}_i \boldsymbol{\Phi}_{i,k} \sum_{j=k-N}^{k-1} \boldsymbol{\Phi}_{k,j+1} \mathbf{Q}_j\right).
\end{aligned}$$

The incorrect bound (4.25) still appears sometimes in the literature. A counterexample to it seems to have not been given when $\mathbf{A}_k$ is non-singular. Nevertheless, only the existence of the bound is important, not its exact form.

Lemma 4.10. *If the dynamic system* (4.1) *is uniformly observable and uniformly controllable, and if* $\mathbf{P}_0$ *is positive-definite, then*

$$\mathbf{P}_k \geq \left(\frac{\beta_1^2}{\beta_1 + N\alpha_2\beta_2^2}\right)\mathbf{I},$$

for all $k \geq N$.

Proof. (Kamen and Su, 1999, Lemma C.3) Matrices

$$\begin{aligned}
\bar{\mathbf{S}}_k &\coloneqq \mathbf{P}_k^{-1} - \mathbf{H}_k^{\mathsf{T}} \mathbf{R}_k^{-1} \mathbf{H}_k, \\
\bar{\mathbf{S}}_k^- &\coloneqq \mathbf{A}_{k-1}^{-\mathsf{T}} \mathbf{P}_{k-1}^{-1} \mathbf{A}_{k-1}^{-1}
\end{aligned}$$

are defined. Using (4.3) and (4.14) these can be manipulated into

$$\begin{aligned}
\bar{\mathbf{S}}_k &= \left((\bar{\mathbf{S}}_k^-)^{-1} + \mathbf{Q}_{k-1}\right)^{-1}, \\
\bar{\mathbf{S}}_k^- &= \mathbf{A}_{k-1}^{-\mathsf{T}} \bar{\mathbf{S}}_{k-1} \mathbf{A}_{k-1}^{-1} + \mathbf{A}_{k-1}^{-\mathsf{T}} \mathbf{H}_{k-1}^{\mathsf{T}} \mathbf{R}_{k-1}^{-1} \mathbf{H}_{k-1} \mathbf{A}_{k-1}^{-1}.
\end{aligned}$$

Comparison with (4.14) and (4.3) shows that $\bar{\mathbf{S}}_k$ can be thought as the error covariance matrix for the system

$$\begin{aligned}
\bar{\mathbf{x}}_k &= \mathbf{A}_{k-1}^{-\mathsf{T}} \bar{\mathbf{x}}_{k-1} + \bar{\mathbf{q}}_{k-1}, \\
\bar{\mathbf{y}}_k &= \bar{\mathbf{x}}_k + \bar{\mathbf{r}}_k,
\end{aligned}$$

where the system noise is $\bar{\mathbf{q}}_{k-1} \sim \mathcal{N}(\mathbf{0}, \mathbf{A}_{k-1}^{-\mathsf{T}} \mathbf{H}_{k-1}^{\mathsf{T}} \mathbf{R}_{k-1}^{-1} \mathbf{H}_{k-1} \mathbf{A}_{k-1}^{-1})$ and the measurement noise $\bar{\mathbf{r}}_k \sim \mathcal{N}(\mathbf{0}, \mathbf{H}_k \mathbf{Q}_{k-1} \mathbf{H}_k^{\mathsf{T}})$. It is easily seen that this system is uniformly observable and uniformly controllable whenever the system (4.1) is. Thus the proof of Lemma 4.9 can be applied to obtain the claimed lower bound for $\mathbf{P}_k$. □

Lemma 4.11. *If the dynamic system* (4.1) *is uniformly observable and* $\mathbf{P}_0$ *is positive-semidefinite, then* $\mathbf{P}_k$ *is positive-definite for all* $k \geq N$.

Proof. (Jazwinski, 1970, Theorem 7.3; Kamen and Su, 1999, Lemma C.1) From (4.3) and (4.14) it is seen that if $\mathbf{P}_N$ is positive-definite, then so are $\mathbf{P}_k$ for $k \geq N$. Since from the same equations $\mathbf{P}_N$ is seen to be positive-semidefinite, it suffices to prove that $\mathbf{P}_N$ is non-singular. Assuming to the contrary, an equation contradicting the assumption of uniform controllability is derived by tedious matrix manipulations. □

Theorem 4.12. *If the dynamic system* (4.1) *is uniformly observable and uniformly controllable, and if* $\mathbf{P}_0$ *is positive-semidefinite, then the Kalman filter of Theorem 4.1 is exponentially stable.*

Proof. (Deyst and Price, 1968; Jazwinski, 1970, Theorem 7.4; Kamen and Su, 1999, Lemma C.4 and Theorem C.2) The final proof utilizes Theorem 3.3. A Lyapunov function

$$V(\mathbf{m}_k, k) = \mathbf{m}_k^{\mathsf{T}} \mathbf{P}_k^{-1} \mathbf{m}_k$$

is defined. Lemmas 4.9 and 4.10 enable the construction of the bounding functions γ_1 and γ_2 of V. The construction of γ_3 is not so easy. Again, through tedious manipulations, one obtains the inequality

$$\begin{aligned} &V(\mathbf{m}_k, k) - V(\mathbf{m}_{k-N-1}, k-N-1) \\ &\qquad \leq - \sum_{i=k-N}^{k} \left(\mathbf{m}_i^{\mathsf{T}} \mathbf{H}_i^{\mathsf{T}} \mathbf{R}_i^{-1} \mathbf{H}_i \mathbf{m}_i + \mathbf{r}_i^{\mathsf{T}} (\mathbf{P}_i^{-})^{-1} \mathbf{r}_i \right). \end{aligned}$$

By minimizing the right-hand side of this inequality with respect to sequence $\{\mathbf{r}_{n-N}, \dots, \mathbf{r}_n\}$ an upper bound

$$V(\mathbf{m}_k, k) - V(\mathbf{m}_{k-N-1}, k-N-1) \leq -\beta \left\| \mathbf{m}_{k-N-1} \right\|^2$$

is obtained for some $\beta > 0$. Furthermore, $\|\mathbf{m}_{k-N-1}\|$ can be proven to be bounded below by $\delta \|\mathbf{m}_k\|$ for some $\delta > 0$, and so

$$\gamma_3(\|\mathbf{m}_k\|) \coloneqq -\beta\delta \left\| \mathbf{m}_k \right\| < 0$$

is a suitable choice. □

4.5 Generalization to Detectable and Stabilizable Systems

This sections contains a generalization, due to Anderson and Moore (1981), of the results of the preceding section to systems that are only uniformly detectable and stabilizable. Useful references on the matter, in addition to the one already cited, are Moore and Anderson (1980) and Anderson (1982). The latter one is particularly helpful in providing some additional clarity to the construction of a sub-optimal estimator needed to bound the error covariance matrices is Anderson and Moore (1981) is occasionally rather terse.

In this section the matrices $\mathbf{A}_{k-1}, \mathbf{H}_k, \mathbf{Q}_{k-1}, \mathbf{R}_k$ and $\mathbf{R}_k^{-1}$ are assumed uniformly bounded. Non-singularity assumptions as in the preceding section are not needed. The proof is based on elegant use of general Theorems 4.6 and 4.7.

First, a counterpart to Lemmas 4.9 and 4.10 is proven. The same fact of the Kalman filter being the linear minimum variance estimator is used but the construction of a suitable sub-optimal estimate is considerably more convoluted.

Lemma 4.13. *If the dynamic system* (4.1) *is uniformly detectable, then the matrices* $\mathbf{P}_k^-$ *and* $\mathbf{P}_k$ *remain uniformly bounded.*

Proof. (Anderson and Moore, 1981, Lemma 5.1; Anderson, 1982, Lemma 3.1 and Proposition 3.1) Let s, t, d and b have the same meaning as in the definition of uniform detectability. Defining a new system $\bar{\mathbf{x}}_k = \mathbf{T}_k \mathbf{x}_k$ for orthogonal $\mathbf{T}_k$, it is seen that the corresponding observability Gramian is $\bar{\boldsymbol{\mathcal{O}}}_{k+s,k} = \mathbf{T}_k^\mathsf{T} \boldsymbol{\mathcal{O}}_{k+s,k} \mathbf{T}_k$. As the observability Gramian is symmetric, it is possible to choose $\mathbf{T}_k$ such that it orthogonally diagonalizes $\boldsymbol{\mathcal{O}}_{k+s,k}$, which means that $\bar{\boldsymbol{\mathcal{O}}}_{k+s,k}$ is diagonal (Murty, 2014, Chapter 6). Therefore, as orthogonal transformations preserve the norm, one may assume that $\boldsymbol{\mathcal{O}}_{k+s,k}$ is diagonal for each k. Further, as seen in the orthogonal diagonalization algorithm, it can be assumed that the diagonal elements of $\boldsymbol{\mathcal{O}}_{k+s,k}$ are ordered in decreasing magnitude. Thus this matrix can be written as

$$\boldsymbol{\mathcal{O}}_{k+s,k} = \boldsymbol{\mathcal{O}}_{k+s,k}^1 \oplus \boldsymbol{\mathcal{O}}_{k+s,k}^2,$$

where $\oplus$ is the direct sum. The matrices $\boldsymbol{\mathcal{O}}_{k+s,k}^1$ and $\boldsymbol{\mathcal{O}}_{k+s,k}^2$ are diagonal matrices of dimensions n_1 and n_2, respectively, such that $\boldsymbol{\mathcal{O}}_{k+s,k}^1 \geq b\mathbf{I}$ and $\boldsymbol{\mathcal{O}}_{k+s,k}^2 < b\mathbf{I}$.

Define then $\overline{\mathbf{m}}_{k+t|k+s}$, a smoothed estimate of $\mathbf{x}_{k+t}$ given the measurements up to time-step $k+s$ by

$$\begin{aligned}\overline{\mathbf{m}}_{k+t|k+s} = \mathbf{\Phi}_{k+t,k}\Big[&\big([\mathcal{O}^1_{k+s,k}]^{-1} \oplus \mathbf{0}_{n_2}\big)\sum_{i=k}^{k+s}\mathbf{\Phi}^\mathsf{T}_{i,k}\mathbf{H}^\mathsf{T}_i\mathbf{R}^{-1}_i\mathbf{y}_i \\ &+ \big(\mathbf{0}_{n_1} \oplus \mathbf{I}_{n_2}\big)\overline{\mathbf{m}}_{k|k+s-t}\Big],\end{aligned} \tag{4.26}$$

initialization being $\overline{\mathbf{m}}_{k|k+s-t} = \mathbf{0}$ for $k < t$. For $i \geq k$ the measurement $\mathbf{y}_i$ can be written as

$$\mathbf{y}_i = \mathbf{H}_i\mathbf{\Phi}_{i,k}\mathbf{x}_k + \mathbf{r}_i + \mathbf{H}_i\sum_{j=1}^{i-k}\mathbf{\Phi}_{i,i-j}\mathbf{q}_{i-j}.$$

Denoting by $\mathbf{B}_i$ the last two terms in the above expression and inserting this into (4.26) one obtains

$$\begin{aligned}&\overline{\mathbf{m}}_{k+t|k+s} \\ &\quad= \mathbf{\Phi}_{k+t,k}\Big[\big([\mathcal{O}^1_{k+s,k}]^{-1} \oplus \mathbf{0}_{n_2}\big)\big(\mathcal{O}^1_{k+s,k} \oplus \mathcal{O}^2_{k+s,k}\big)\mathbf{x}_k + \big(\mathbf{0}_{n_1} \oplus \mathbf{I}_{n_2}\big)\overline{\mathbf{m}}_{k|k+s-t}\Big] \\ &\qquad+ \mathbf{\Phi}_{k+t,k}\big[(\mathcal{O}^1_{k+s,k})^{-1} \oplus \mathbf{0}_{n_2}\big]\sum_{i=k}^{k+s}\mathbf{\Phi}^\mathsf{T}_{i,k}\mathbf{H}^\mathsf{T}_i\mathbf{R}^{-1}_i\mathbf{B}_i.\end{aligned}$$

Further, by using $\mathbf{C}_k$ for the second term in this equality and noting that

$$\mathbf{x}_{k+t} = \mathbf{\Phi}_{k+t,k}\mathbf{x}_k + \mathbf{D}_k,$$

where $\mathbf{D}_k = \sum_{i=1}^t \mathbf{\Phi}_{k+t,k+i}\mathbf{q}_{k+i-1}$, a slight rearrangement yields the form

$$\overline{\mathbf{m}}_{k+t|k+s} - \mathbf{x}_{k+t} = \mathbf{\Phi}_{k+t,k}\big(\mathbf{0}_{n_1} \oplus \mathbf{I}_{n_2}\big)(\overline{\mathbf{m}}_{k|k+s-t} - \mathbf{x}_k) + \mathbf{C}_k - \mathbf{D}_k. \tag{4.27}$$

Choose $\mathbf{z} = (0, 0, \ldots, 0, 1)$. Then $\mathbf{z}^\mathsf{T}\mathcal{O}_{k+s,k}\mathbf{z} = \mathbf{z}^\mathsf{T}\mathcal{O}^2_{k+s,k}\mathbf{z} < b\mathbf{z}^\mathsf{T}\mathbf{z} = b$. Therefore by the definition of uniform detectability,

$$\|\mathbf{\Phi}_{k+t,k}\big(\mathbf{0}_{n_1} \oplus \mathbf{I}_{n_2}\big)\| = \|\mathbf{\Phi}_{k+t,k}\big(\mathbf{0}_{n_1} \oplus \mathbf{I}_{n_2}\big)\mathbf{z}\| = \|\mathbf{\Phi}_{k+t,k}\mathbf{z}\| < d\,\|\mathbf{z}\| = d < 1.$$

Then, since by the uniform boundedness assumption of system matrices, $\mathbf{C}_k - \mathbf{D}_k$ is merely a linear combination of the noise terms with uniformly bounded coefficients and by the above inequality the system (4.27) is contractive, it follows that $\mathbb{E}(\|\overline{\mathbf{m}}_{k+t|k+s} - \mathbf{x}_{k+t}\|^2)$ remains uniformly bounded.

Error dynamics of a non-optimal filter estimate defined by

$\overline{\mathbf{m}}_{k+s} \coloneqq \mathbf{\Phi}_{k+s,k+t}\overline{\mathbf{m}}_{k+t|k+s}$ are given as

$$\overline{\mathbf{m}}_{k+s} - \mathbf{x}_{k+s} = \mathbf{\Phi}_{k+s,k+t}(\overline{\mathbf{m}}_{k+t|k+s} - \mathbf{x}_{k+t}) - \sum_{i=1}^{s-t} \mathbf{\Phi}_{k+s,k+i}\mathbf{q}_{k+i-1},$$

the covariance of which is uniformly bounded as a consequence of the preceeding remarks. By the optimality of the Kalman filter it now follows that $\mathbf{P}_k$ is uniformly bounded and by the uniform boundedness of $\mathbf{A}_{k-1}$ and $\mathbf{Q}_{k-1}$ it is seen from (4.3) that so is $\mathbf{P}_k^-$. □

Adding the assumption of uniform stabilizability then leads to the main result on the exponential stability of the Kalman filter.

Theorem 4.14. *If the dynamic system* (4.1) *is uniformly detectable and uniformly stabilizable and* $\mathbf{P}_0$ *is positive-semidefinite, then the Kalman filter of Theorem 4.1 is exponentially stable.*

Proof. (Anderson and Moore, 1981, Theorem 5.3) The form (4.13) can be equivalently written as

$$\mathbf{P}_{k+1}^- = \mathbf{A}_k(\mathbf{I} - \mathbf{K}_k\mathbf{H}_k)\mathbf{P}_k^-(\mathbf{I} - \mathbf{K}_k\mathbf{H}_k)^\mathsf{T}\mathbf{A}_k^\mathsf{T} + \mathbf{G}_k\mathbf{G}_k^\mathsf{T},$$

where

$$\mathbf{G}_k = \begin{pmatrix} \mathbf{Q}_k^{1/2} & \mathbf{A}_k\mathbf{K}_k\mathbf{R}_k^{1/2} \end{pmatrix}.$$

As the pair $[\mathbf{A}_k, \mathbf{Q}_k^{1/2}]$ has been assumed uniformly stabilizable, it readily follows that so is the pair $[\mathbf{A}_k, \mathbf{G}_k]$ since (4.21) is not changed and the controllability matrix of the latter pair grows, compared to that of the former, in terms of positive-definitess. Now, one can write

$$\mathbf{A}_k(\mathbf{I} - \mathbf{K}_k\mathbf{H}_k) = \mathbf{A}_k + \mathbf{G}_k\mathbf{T}_k,$$

where

$$\mathbf{T}_k = \begin{pmatrix} \mathbf{0}_n \\ -\mathbf{R}_k^{-1/2}\mathbf{H}_k \end{pmatrix}.$$

From Lemma 4.13 and the boundedness assumptions of the system matrices it follows that the matrices $\mathbf{A}_k, \mathbf{P}_k^-, \mathbf{B}_k$ and $\mathbf{T}_k$ are uniformly bounded from above. Therefore, applying invariance under feedback of Theorem 4.6 the pair $[\mathbf{A} + \mathbf{G}\mathbf{T}_k, \mathbf{G}_k]$ is found uniformly stabilizable. Thus, by Theorem 4.7 the system $\mathbf{z}_{k+1} = \mathbf{A}_k(\mathbf{I} - \mathbf{K}_k\mathbf{H}_k)\mathbf{z}_k$ is exponentially stable. This is exactly the homogeneous part of the predicted state estimate equation (4.9). Finally, since the matrix $\mathbf{I}-\mathbf{K}_k\mathbf{H}_k$ is uniformly bounded from above, the relationship (4.10) shows that the Kalman filter is exponentially stable. □

By interpreting the measurements in (4.8) as input it is seen by Theorem 3.2 that the filter is also BIBO stable. In fact, there is a strong connection between exponential stability and BIBO stability (Anderson, 1982).

An important and immediate consequence of the above theorem is that the Kalman filter forgets its initial state: the sequences of error covariances initialized with different $\mathbf{P}_0$ converge to each other and the same is true of the state estimates $\mathbf{m}_k$.

Theorem 4.15. *Suppose that the dynamic system* (4.1) *is uniformly detectable and uniformly stabilizable. If* $\mathbf{m}_k^1$ *and* $\mathbf{P}_k^1$ *are the state estimate and error covariance of the Kalman filter with initial distribution* $\mathcal{N}(\mathbf{m}_0^1, \mathbf{P}_0^1)$ *and* $\mathbf{m}_k^2$ *and* $\mathbf{P}_k^2$ *those of the filter with initial distribution* $\mathcal{N}(\mathbf{m}_0^2, \mathbf{P}_0^2)$*, then*

$$\lim_{k\to\infty} \|\mathbf{P}_k^1 - \mathbf{P}_k^2\| = 0,$$
$$\lim_{k\to\infty} \mathbb{E}(\|\mathbf{m}_k^1 - \mathbf{m}_k^2\|) = 0.$$

Proof. The first part of the proof on the convergence of the error covariance matrices can be found in Jazwinski (1970, Theorem 7.5) and Kamen and Su (1999, Theorem C.4).

The difference $\|\mathbf{P}_k^1 - \mathbf{P}_k^2\|$ can be written as

$$\|\mathbf{P}_k^1 - \mathbf{P}_k^2\| = \|\mathbf{\Psi}_{k,0}^1(\mathbf{P}_0^1 - \mathbf{P}_0^2)(\mathbf{\Psi}_{k,0}^2)^{\mathsf{T}}\| \leq \|\mathbf{\Psi}_{k,0}^1\| \, \|(\mathbf{P}_0^1 - \mathbf{P}_0^2)\| \, \|(\mathbf{\Psi}_{k,0}^2)^{\mathsf{T}}\| \,,$$

where $\mathbf{\Psi}_{k,0}^1$ and $\mathbf{\Psi}_{k,0}^2$ are the state transition matrices of the filter (4.8), corresponding to $\mathbf{P}_k^1$ and $\mathbf{P}_k^2$, respectively. Since the Kalman filter is exponentially stable with these assumptions, both these state transition matrices converge to zero exponentially and the first claim of the theorem follows. This was particularly easy as the error covariance matrix does not depend on the mean of the initial distribution. The state estimates however depend on the measurements and the covariance matrices.

Let $\mathbf{m}_k^{2,1}$ be a state estimate at time-step k of a Kalman filter with initial distribution $\mathcal{N}(\mathbf{m}_0^2, \mathbf{P}_0^1)$. Then

$$\mathbf{m}_k^1 - \mathbf{m}_k^2 = (\mathbf{m}_k^1 - \mathbf{m}_k^{2,1}) + (\mathbf{m}_k^{2,1} - \mathbf{m}_k^2).$$

The first term on the right-hand side is merely

$$\mathbf{m}_k^1 - \mathbf{m}_k^{2,1} = \mathbf{\Psi}_{k,0}^1(\mathbf{m}_0^1 - \mathbf{m}_0^2)$$

which converges to zero. The second term, a somewhat trickier one, is

$$\begin{aligned}
\mathbf{m}_k^{2,1} - \mathbf{m}_k^2 &= \mathbf{\Psi}_{k,0}^1 \mathbf{m}_0^2 - \mathbf{\Psi}_{k,0}^2 \mathbf{m}_0^2 + \sum_{i=1}^{k} (\mathbf{\Psi}_{k,i}^1 \mathbf{K}_i^1 - \mathbf{\Psi}_{k,i}^2 \mathbf{K}_i^2) \mathbf{y}_i \\
&= (\mathbf{\Psi}_{k,0}^1 - \mathbf{\Psi}_{k,0}^2) \mathbf{m}_0^2 \\
&\quad + \sum_{i=1}^{k} (\mathbf{\Psi}_{k,i}^1 \mathbf{K}_i^1 - \mathbf{\Psi}_{k,i}^2 \mathbf{K}_i^2) \mathbf{H}_i \left(\mathbf{\Phi}_{i,0} \mathbf{x}_0 + \sum_{j=0}^{i-1} \mathbf{\Phi}_{i,j+1} \mathbf{q}_j \right) \\
&\quad + \sum_{i=1}^{k} (\mathbf{\Psi}_{k,i}^1 \mathbf{K}_i^1 - \mathbf{\Psi}_{k,i}^2 \mathbf{K}_i^2) \mathbf{r}_i,
\end{aligned}$$

where $\mathbf{K}_i^1$ and $\mathbf{K}_i^2$ are the Kalman gains of the filters initialized with $\mathbf{P}_0^1$ and $\mathbf{P}_0^2$, respectively. Because $\mathbf{\Psi}_{k,i}^1$ and $\mathbf{\Psi}_{k,i}^2$ converge exponentially to each other, it follows that $\mathbb{E}(\|\mathbf{m}_k^{2,1} - \mathbf{m}_k^2\|) \to 0$ as $k \to \infty$. Hence

$$\lim_{k \to \infty} \mathbb{E}(\|\mathbf{m}_k^1 - \mathbf{m}_k^2\|) = 0.$$

□

Note that only the case with a Gaussian initial distribution has been treated. Results allowing for more general initial distributions are few. Sowers and Makowski (1992) show that the estimates produced by the linear Kalman filter and the optimal filter for a linear time-invariant system with non-Gaussian initial distribution converge to each other exponentially under certain stabilizability conditions. They limit the complete analysis to scalar systems, however. A more general approach can be found in Ocone and Pardoux (1996) where almost sure asymptotic stability results are proved for the time-invariant Kalman filter with possibly non-Gaussian initial distribution. It seems there are no analogous results for systems not time-invariant.

5 Stability of Non-Linear Kalman Filters

In this section *non-linear dynamic systems* are treated. A *general discrete-time non-linear dynamic system* can be written as

$$\begin{aligned} \mathbf{x}_k &= \mathbf{f}(\mathbf{x}_{k-1}, \mathbf{q}_{k-1}), \\ \mathbf{y}_k &= \mathbf{h}(\mathbf{x}_k, \mathbf{r}_k). \end{aligned} \tag{5.1}$$

In this general system the manner by which the process and measurement noises enter the system may be *non-additive* as opposed to additive linear systems considered in the previous section. As such, a special case of this system is a system in which the process and measurement noises are assumed additive. The focus will be on this kind of additive systems as the stability results for them were proven first and the proofs of analogous results for general systems are not essentially different. So, an *additive discrete-time non-linear dynamic system* can be written as

$$\begin{aligned} \mathbf{x}_k &= \mathbf{f}(\mathbf{x}_{k-1}) + \mathbf{q}_{k-1}, \\ \mathbf{y}_k &= \mathbf{h}(\mathbf{x}_k) + \mathbf{r}_k. \end{aligned} \tag{5.2}$$

In both of these systems (5.1) and (5.2) $\mathbf{x}_k \in \mathbb{R}^n$ is the system state, $\mathbf{y}_k \in \mathbb{R}^m$ is the measurement, $\mathbf{q}_{k-1} \sim \mathcal{N}(\mathbf{0}, \mathbf{Q}_{k-1})$ is the process noise and $\mathbf{r}_k \sim \mathcal{N}(\mathbf{0}, \mathbf{R}_k)$ is the measurement noise. The noise processes are assumed uncorrelated and independent of the initial state. Function $\mathbf{f}$ is the *dynamic model function* and $\mathbf{h}$ the *measurement model function*, both assumed continuously differentiable. In addition to the system state, the the functions $\mathbf{f}$ and $\mathbf{h}$ can very well depend on time also. As this causes no changes in the proofs such notation is suppressed.

Similarly to the linear case, inclusion of a control signal $\mathbf{u}_{k-1}$ to the model is possible, causing no effect to the proofs. In this case, the state part of (5.1) would take the form

$$\mathbf{x}_k = \mathbf{f}(\mathbf{x}_{k-1}, \mathbf{u}_{k-1}, \mathbf{q}_{k-1})$$

or an analogous additive term would appear in (5.2).

5.1 Approximative Non-Linear Filters

Section 3.3 briefly discussed the available stability results for general optimal filters. In practice these results cannot be applied for two reasons. Firstly, the assumptions made are far too restrictive to be satisfied by any practical model. The second reason, and the more important one, is that, as previously remarked, it is not possible to efficiently implement an optimal filter when the model is non-linear. Some minor exceptions do exist, though (Chen, 2003). Therefore one is forced to resort to approximative sub-optimal filters. There exists a great number of such filters, more being constantly developed. Perhaps the most widely used are *the extended Kalman filter* (EKF), *the unscented Kalman filter* (UKF) and *particle filters*. This section is concerned with the stability properties, theory of which is still very limited, of the EKF and the UKF as well as a more general class of *Gaussian filters* or *non-linear Kalman filters*.

Consider Gaussian approximations to the filtering distributions. To consider just a single step of filtering, one has a random variable $\mathbf{x} \sim \mathcal{N}(\mathbf{m}, \mathbf{P})$ and its non-linear transformation $\mathbf{y} = \mathbf{g}(\mathbf{x}) + \mathbf{q}$, where $\mathbf{g}$ is some non-linear function and $\mathbf{q} \sim \mathcal{N}(\mathbf{0}, \mathbf{Q})$. To obtain a Gaussian approximation to the distribution of $\mathbf{x} \mid \mathbf{y}$, the joint distribution of $\mathbf{x}$ and $\mathbf{y}$ is given a Gaussian approximation,

$$\begin{pmatrix} \mathbf{x} \\ \mathbf{y} \end{pmatrix} \sim \mathcal{N}\left(\begin{pmatrix} \mathbf{m} \\ \boldsymbol{\mu} \end{pmatrix}, \begin{pmatrix} \mathbf{P} & \mathbf{C} \\ \mathbf{C}^\mathsf{T} & \mathbf{S} \end{pmatrix} \right). \tag{5.3}$$

Then by Lemma A.5 $\mathbf{x} \mid \mathbf{y}$ has the approximative Gaussian distribution $\mathcal{N}(\mathbf{m} + \mathbf{C}\mathbf{S}^{-1}(\mathbf{y} - \boldsymbol{\mu}), \mathbf{P} - \mathbf{C}\mathbf{S}^{-1}\mathbf{C}^\mathsf{T})$. It is up to the particular approximative filtering algorithm to obtain approximations to the joint mean and covariance of $\mathbf{x}$ and $\mathbf{y}$. This procedure can then be iterated to obtain the particular approximative non-linear filter. See Särkkä (2013) for a more thorough treatment.

Naturally, there are multiple ways to determine the parameters of (5.3). The ones considered in this thesis use the following three methods:

- The extended Kalman filter that uses a local first-order linearization at the latest estimate is somewhat of a standard in non-linear estimation but can easily diverge if the model is too non-linear or the initial estimate is poor.

- The unscented Kalman filter utilizes a sigma-point approximation in which a set of deterministically chosen points is used to approximate the mean and covariance. The UKF is a more recent algorithm and works usually better than the EKF.

- Gaussian filters are based on matching the first and second moments of the distribution. This provides the linear minimum mean square estimator (García-Fernández et al., 2014). However, the algorithm necessitates the computation of integrals that may not have a solution in closed form so that approximations to these integrals (of which the UKF is in fact one) must be used.

Approximating the filtering distribution with a single Gaussian distribution usually works well enough if non-linearity of the model is sufficiently restricted, but multi-modality of the filtering distribution may cause serious complications for these unimodal approximations. In such a case, multiple weighted Gaussian distributions could naturally be used instead of a single one as any density function can be arbitrarily accurately approximated by such a sum of weighted Gaussian densities. Variants of this idea are known as *Gaussian sum filters* (Alspach and Sorenson, 1972). A more popular choice in this situation is to use Monte Carlo simulation based particle filters (Särkkä, 2013, Chapter 7).

The non-linear Kalman filters of this section, in contrast to the Kalman filter of the preceding section, are in general sub-optimal. As such, the state estimate $\mathbf{m}_k$ produced by them loses its probabilistic interpretation as the conditional mean $\mathbb{E}(\mathbf{x}_k \mid \mathbf{y}_{1:k})$ and $\mathbf{P}_k$ can no longer be regarded as the error covariance matrix. However, for convenience and convention $\mathbf{P}_k$ will also in sequel called the error covariance matrix.

5.2 Overview of the Non-Linear Stability Results

None of the stability results for the non-linear filters considered are comparable in the simplicity of assumptions or in strength of implications to those obtained in the linear case. In every stability theorem one has to assume numerous matrix inequalities, some of which cannot be verified beforehand. This is coupled with the requirement that the initial estimation error be sufficiently small. In fact, the theorems for the EKF produce a quantitative bound for the initial estimation error — a bound so small and conservative that the results are of little practical use. In the linear case the initial estimate is of no importance as by Theorem 4.15 the Kalman filters initialized with different initial distributions convergence towards each other.

The theorems for the UKF and Gaussian filters in general are based on the use of random matrices that are introduced to transform certain approximations to equalities. The use of this method does not provide any quantitative bounds and the assumptions become very difficult to verify. The results are very qualitative in nature and mainly tell that noise covariance matrix tuning may induce filter stability.

These constraints stem from the fact that all these non-linear filters are based on some kind of heuristic approximation, losing the optimality of the linear Kalman filter, and hence no minimum-variance based arguments can be employed as in the linear case. To form some kind of convenient recursive form for the filter estimates as (4.8) of the linear case one has to perform a linearization of some type. But linearization (be it Taylor series based or statistical) is dependent on the filter estimates to be of any use and so there are not many quantities that do not depend on the observations. This is a serious problem, particularly for the error covariance matrix that just has to be assumed uniformly bounded from above and below. To this day, no one has been able to perform analysis of its behavior.

Since the filter performance is not anyway optimal, there is no reason not to discard the use of the actual noise covariance matrices $\mathbf{Q}_k$ and $\mathbf{R}_k$ in the filtering algorithms. In fact, by tuning these matrices one can often achieve better filter performance (Bolognani et al., 2003; Xiong et al., 2006), which is not by any means a recent discovery (Anderson, 1973; Maybeck, 1979, Section 6.8). Tuning (usually enlargement) of these matrices in order to achieve stability, with the possible drawback of deteriorating accuracy, is discussed in several occasions in this section. By this procedure surprisingly useful, although still difficult to verify, stability results are possible. In practice, it is often impossible to determine the real covariances of the noise terms and hence $\mathbf{Q}_k$ and $\mathbf{R}_k$ often remain out of necessity mere guesses even in the linear case.

Regrettably, not all stability results for non-linear Kalman filters can be discussed here, for there exist a body of theory known as *contraction theory* (Jouffroy and Fossen, 2010) that utilizes differential geometric arguments to study stability. Using contraction theory, Maree et al. (2014) have recently been able to prove some stability results for the UKF.

As the general stability results for approximative non-linear filters began to appear only in the 1990s, there is not much general literature available commenting them. This is amplified by the uselessness of the current results in applications. Some commentary is given by Simon (2006, Section 13.5).

5.3 The Extended Kalman Filter

The extended Kalman filter, obtained by local first-order linearization, is the standard tool for practical non-linear filtering, having been around since the 1960s when it was proposed by Stanley Schmidt (Simon, 2006, p. 400). However, the results on the stability of the EKF have been very lacking, owing to its highly heuristic nature.

Until recently there have been no results for the EKF in the general setting of stochastic non-linear systems. Special cases, such as the EKF used as a parameter estimator for linear systems have been treated (Ljung, 1979;

Ursin, 1980). These seem to be one of the first stability results concerning the EKF, given remarks on the absence of such results by Ljung (1979) and Anderson and Moore (1979, p. 206). The stability of the constant gain EKF and the modified gain EKF were considered by Safonov and Athans (1978) and Song and Speyer (1985), respectively. Using the total stability theorem (Anderson et al., 1986, Section 1.3.2), stability results were obtained by La Scala et al. (1995) for a system with linear measurement model.

The stability of the EKF as an observer for non-linear deterministic systems, that is, for systems without any noise, has received relatively much attention. Baras et al. (1988) constructed observers for continuous-time non-linear systems with a linear measurement model and provided some stability results. Song and Grizzle (1992, 1995) were able to prove that the EKF is a local asymptotic observer for discrete-time non-linear deterministic systems if the initial estimation error or the non-linearities of the system are sufficiently small and the system satisfies a non-linear observability rank condition related to that of Theorem 4.8. The most advanced work to this end is that of Reif and Unbehauen (1999) who proved that the EKF is an exponential observer, that is, the difference equation for the EKF estimation error admits an exponentially stable equilibrium at the origin. They do not need sufficiently small initial estimation error to establish their result which they achieve by standard results on Lyapunov functions. Krener (2003) shows by an extension of the methods of Baras et al. (1988) the exponential convergence to zero of the estimation error of the EKF observer for continuous-time systems with non-linear measurement model.

Reif et al. (1999) were first to prove general results about the stability of the EKF in the general stochastic case. They proved that the estimation error is exponentially bounded in mean square and almost surely bounded (though latter of the claims may not hold given the discussion in Section 3.2). This they accomplished by using stochastic stability lemma as a substitute of some variant of deterministic Lyapunov stability theorem, and the one-step formulation of the EKF equations. However, the practicality of these results is hindered by the fact that the noise covariance matrices and initial errors have very conservative bounds and the non-linearities of the dynamic model and measurement model functions permitted by the result are severely bounded. Kluge et al. (2010) were able to drop the assumption of additive noises and consider the general non-linear dynamic model (5.1) using a proof very similar to that of Reif et al. (1999), but the problem of very conservative bounds remains even though Rhudy et al. (2012) have been able to relax them slightly.

It would be interesting to know more closely the exact history behind these results, for Reif et al. (1999) was not the first general EKF stability results to appear in print. Before it similar were published in Unbehauen (1998, Section 6.2) and Reif et al. (1998), both in German.

After a brief introduction of the EKF equations the proof of Reif et al. (1999) for the boundedness of the EKF estimation error is presented. Then the same result is presented in terms of a certain non-linear observability condition, applying the preceding theorem. Finally the generalization of Kluge et al. (2010) to non-additive system is given.

5.3.1 The Extended Kalman Filter Equations

Consider the additive non-linear dynamic system (5.2). As the dynamic model function $\mathbf{f}$ and the measurement model function $\mathbf{h}$ are assumed to be at least once continuously differentiable, a Taylor series based linearization can be used. Let $\mathbf{F_x}$ and $\mathbf{H_x}$ denote the Jacobians $\partial \mathbf{f}/\partial \mathbf{x}$ and $\partial \mathbf{h}/\partial \mathbf{x}$ of the functions $\mathbf{f}$ and $\mathbf{h}$ with respect to $\mathbf{x}$, respectively, and let $\widehat{\mathbf{Q}}_{k-1}$ and $\widehat{\mathbf{R}}_k$ for $k \geq 1$ be some possibly time-varying symmetric positive-definite matrices. Then the extended Kalman filter equations are given by the following algorithms.

Algorithm 5.1 (Two-step extended Kalman filter)**.** *The two-step extended Kalman filter equations for the non-linear dynamic system* (5.2) *consists of the prediction step*

$$\begin{aligned} \mathbf{m}_k^- &= \mathbf{f}(\mathbf{m}_{k-1}), \\ \mathbf{P}_k^- &= \mathbf{F_x}(\mathbf{m}_{k-1})\mathbf{P}_{k-1}\mathbf{F}_\mathbf{x}^\mathsf{T}(\mathbf{m}_{k-1}) + \widehat{\mathbf{Q}}_{k-1} \end{aligned}$$

and the update step

$$\begin{aligned} \mathbf{v}_k &= \mathbf{y}_k - \mathbf{h}(\mathbf{m}_k^-), \\ \mathbf{S}_k &= \mathbf{H_x}(\mathbf{m}_k^-)\mathbf{P}_k^-\mathbf{H}_\mathbf{x}^\mathsf{T}(\mathbf{m}_k^-) + \widehat{\mathbf{R}}_k, \\ \mathbf{K}_k &= \mathbf{P}_k^-\mathbf{H}_\mathbf{x}^\mathsf{T}(\mathbf{m}_k^-)\mathbf{S}_k^{-1}, \\ \mathbf{m}_k &= \mathbf{m}_k^- + \mathbf{K}_k\mathbf{v}_k, \\ \mathbf{P}_k &= \mathbf{P}_k^- - \mathbf{K}_k\mathbf{S}_k\mathbf{K}_k^\mathsf{T}. \end{aligned}$$

These equations are obtained by local first-order linearization of $\mathbf{f}$ and $\mathbf{h}$ at the latest estimate, which is to be understood as calculating their Jacobians at $\mathbf{m}_{k-1}$ and $\mathbf{m}_k^-$ and using them to approximate these non-linear functions. Then one just applies the usual linear Kalman filter of Theorem 4.1 on these linear estimates on each step. Therefore, as can readily be seen, the EKF algorithm is just the usual linear Kalman filter with dynamic and measurement model matrices replaced by the Jacobians of the dynamic model and measurement functions, respectively. A statistical linearization based derivation can be found in Särkkä (2013, Chapter 5). Other sources, from the vast number of possibilities, are Jazwinski (1970, Chapter 9) and Anderson and Moore (1979, Chapter 8).

Of course, one could use a linearization of higher order (assuming of course the existence of such derivatives of higher order), including more than the first term of the Taylor series, though this is rarely done. The second-order EKF (Särkkä, 2013, Algorithm 5.6) is sometimes use, but the more involved nature of Taylor series in higher dimensions complicates the use of such enchantments.

The usual choice is to take $\widehat{\mathbf{Q}}_{k-1}$ and $\widehat{\mathbf{R}}_k$ to be the covariance matrices of the noise terms, $\widehat{\mathbf{Q}}_{k-1} = \mathbf{Q}_{k-1}$ and $\widehat{\mathbf{R}}_k = \mathbf{R}_k$, but this is not by any means necessary since the EKF is not an optimal filtering algorithm even to start with. By choosing these matrices appropriately one can hope to tune the filter stability and performance.

In dealing with the stability of the EKF the following one-step formulation of the EKF, found for example in Goodwin and Sin (1984, Section 7.7), will be of use. The practical effectivity of the two-step and one-step algorithms may be different (Ljung, 1979) but the convergence properties are the same. The use of this one-step formulation somewhat simplifies the inequalities involved in analysis of a Lyapunov function in the proof of Theorem 5.4 (Rapp, 2004, Section 3.2.2), though it is still possible to achieve similar results with the two-step formulation.

Note that in this algorithm the state and covariance estimates are obtained as estimates to the predicted mean and predicted covariance, but this poses no problem as the EKF estimates are approximative even to begin with.

Algorithm 5.2 (One-step extended Kalman filter)**.** *The one-step extended Kalman filter equations for the non-linear dynamic system* (5.2) *are*

$$\begin{aligned}\mathbf{m}_{k+1} &= \mathbf{f}(\mathbf{m}_k) + \mathbf{K}_k[\mathbf{y}_k - \mathbf{h}(\mathbf{m}_k)],\\ \mathbf{P}_{k+1} &= \mathbf{F}_{\mathbf{x}}(\mathbf{m}_k)\mathbf{P}_k\mathbf{F}_{\mathbf{x}}^\mathsf{T}(\mathbf{m}_k) + \widehat{\mathbf{Q}}_k\\ &\quad - \mathbf{F}_{\mathbf{x}}(\mathbf{m}_k)\mathbf{K}_k[\mathbf{H}_{\mathbf{x}}(\mathbf{m}_k)\mathbf{P}_k\mathbf{H}_{\mathbf{x}}^\mathsf{T}(\mathbf{m}_k) + \widehat{\mathbf{R}}_k]\mathbf{K}_k^\mathsf{T}\mathbf{F}_{\mathbf{x}}^\mathsf{T}(\mathbf{m}_k),\end{aligned} \tag{5.4}$$

where the Kalman gain is

$$\mathbf{K}_k = \mathbf{P}_k\mathbf{H}_{\mathbf{x}}^\mathsf{T}(\mathbf{m}_k)[\mathbf{H}_{\mathbf{x}}(\mathbf{m}_k)\mathbf{P}_k\mathbf{H}_{\mathbf{x}}^\mathsf{T}(\mathbf{m}_k) + \widehat{\mathbf{R}}_k]^{-1}. \tag{5.5}$$

Finally, there is an analogous two-step version for the non-additive system (5.3), found in Särkkä (2013, Algorithm 5.5).

Algorithm 5.3 (Non-additive extended Kalman filter)**.** *The extended Kalman filter equations for the non-linear dynamic system* (5.1) *consist of the prediction step*

$$\begin{aligned}
\mathbf{m}_k^- &= \mathbf{f}(\mathbf{m}_{k-1}, \mathbf{0}),\\
\mathbf{P}_k^- &= \mathbf{F}_\mathbf{x}(\mathbf{m}_{k-1})\mathbf{P}_{k-1}\mathbf{F}_\mathbf{x}^\mathsf{T}(\mathbf{m}_{k-1}) + \mathbf{F}_\mathbf{q}(\mathbf{m}_{k-1})\widehat{\mathbf{Q}}_{k-1}\mathbf{F}_\mathbf{q}^\mathsf{T}(\mathbf{m}_{k-1})
\end{aligned}$$

and the update step

$$\begin{aligned}
\mathbf{v}_k &= \mathbf{y}_k - \mathbf{h}(\mathbf{m}_k^-, \mathbf{0}),\\
\mathbf{S}_k &= \mathbf{H}_\mathbf{x}(\mathbf{m}_k^-)\mathbf{P}_k^-\mathbf{H}_\mathbf{x}^\mathsf{T}(\mathbf{m}_k^-) + \mathbf{H}_\mathbf{r}(\mathbf{m}_k^-)\widehat{\mathbf{R}}_k\mathbf{H}_\mathbf{r}^\mathsf{T}(\mathbf{m}_k^-),\\
\mathbf{K}_k &= \mathbf{P}_k^-\mathbf{H}_\mathbf{x}^\mathsf{T}(\mathbf{m}_k^-)\mathbf{S}_k^{-1},\\
\mathbf{m}_k &= \mathbf{m}_k^- + \mathbf{K}_k\mathbf{v}_k,\\
\mathbf{P}_k &= \mathbf{P}_k^- - \mathbf{K}_k\mathbf{S}_k\mathbf{K}_k^\mathsf{T}.
\end{aligned}$$

5.3.2 Stability of the Additive EKF

First, the standard result on the EKF stability of Reif et al. (1999) is presented. Since $\mathbf{f}$ and $\mathbf{h}$ are continuously differentiable, for any points $\mathbf{y}$ and $\mathbf{z}$ in their domains it is possible to write

$$\begin{aligned}
\mathbf{f}(\mathbf{y}) - \mathbf{f}(\mathbf{z}) &= \mathbf{F}_\mathbf{x}(\mathbf{z})(\mathbf{y} - \mathbf{z}) + \boldsymbol{\varphi}(\mathbf{y}, \mathbf{z}),\\
\mathbf{h}(\mathbf{y}) - \mathbf{h}(\mathbf{z}) &= \mathbf{H}_\mathbf{x}(\mathbf{z})(\mathbf{y} - \mathbf{z}) + \boldsymbol{\chi}(\mathbf{y}, \mathbf{z}),
\end{aligned}$$

where the remainder terms of Taylor series $\boldsymbol{\varphi}$ and $\boldsymbol{\chi}$ are some, generally non-linear, functions. For notational convenience, denote $\mathbf{F}_k = \mathbf{F}_\mathbf{x}(\mathbf{m}_k)$ and $\mathbf{H}_k = \mathbf{H}_\mathbf{x}(\mathbf{m}_k)$.

The estimation error $\boldsymbol{\xi}_k = \mathbf{x}_k - \mathbf{m}_k$ can be, in the light of Algorithm 5.2, written recursively as

$$\boldsymbol{\xi}_{k+1} = \mathbf{F}_k(\mathbf{I} - \mathbf{K}_k\mathbf{H}_k)\boldsymbol{\xi}_k + \boldsymbol{\rho}_k + \boldsymbol{\sigma}_k, \tag{5.6}$$

where

$$\begin{aligned}
\boldsymbol{\rho}_k &:= \boldsymbol{\varphi}(\mathbf{x}_k, \mathbf{m}_k) - \mathbf{K}_k\boldsymbol{\chi}(\mathbf{x}_k, \mathbf{m}_k), \qquad (5.7)\\
\boldsymbol{\sigma}_k &:= \mathbf{q}_k - \mathbf{K}_k\mathbf{r}_k.
\end{aligned}$$

Theorem 5.4. *Consider the discrete-time non-linear dynamic system* (5.2) *and the extended Kalman filter of Algorithms 5.1 and 5.2. Suppose that the following conditions hold:*

(1) There exist $f, h, p_1, p_2, q, r > 0$ *such that the following bounds hold for every* $k \geq 0$*:*

$$\|\mathbf{F}_k\| \leq f, \quad \|\mathbf{H}_k\| \leq h, \tag{5.8}$$

$$p_1 \mathbf{I} \leq \mathbf{P}_k \leq p_2 \mathbf{I}, \tag{5.9}$$

$$\hat{q}\mathbf{I} \leq \widehat{\mathbf{Q}}_k, \qquad \hat{r}\mathbf{I} \leq \widehat{\mathbf{R}}_k.$$

(2) $\mathbf{F}_k$ *is non-singular for every* $k \geq 0$.

(3) There exist $\varepsilon_{\boldsymbol{\varphi}}, \varepsilon_{\boldsymbol{\chi}}, \kappa_{\boldsymbol{\varphi}}, \kappa_{\boldsymbol{\chi}} > 0$ *such that the functions* $\boldsymbol{\varphi}$ *and* $\boldsymbol{\chi}$ *are bounded from above by*

$$\begin{aligned} \|\boldsymbol{\varphi}(\mathbf{y}, \mathbf{z})\| &\leq \kappa_{\boldsymbol{\varphi}} \|\mathbf{y} - \mathbf{z}\|^2, \\ \|\boldsymbol{\chi}(\mathbf{y}, \mathbf{z})\| &\leq \kappa_{\boldsymbol{\chi}} \|\mathbf{y} - \mathbf{z}\|^2 \end{aligned} \tag{5.10}$$

for all points $\mathbf{y}$ *and* $\mathbf{z}$ *in their domain with* $\|\mathbf{y} - \mathbf{z}\| \leq \varepsilon_{\varphi}$ *and* $\|\mathbf{y} - \mathbf{z}\| \leq \varepsilon_{\chi}$*, respectively.*

Then there exist $\varepsilon, \delta > 0$ *such that the conditions* $\|\boldsymbol{\xi}_0\| \leq \varepsilon$ *and* $\mathbf{Q}_k, \mathbf{R}_k \leq \delta \mathbf{I}$ *guarantee that the estimation error* $\boldsymbol{\xi}_k$ *is exponentially bounded in mean square and stochastically bounded.*

The proof of this theorem requires three technical lemmas that provide necessary upper bounds for terms that will in appear the stochastic Lyapunov function to be constructed. For notational convenience, denote in the following $\mathbf{\Pi}_k := \mathbf{P}_k^{-1}$.

Note the resemblance of (5.10) to local Hölder continuity and that the proof to follow can be carried out with any exponent on the interval $(1, 2]$ on the right-hand side of these equations.

Lemma 5.5. *Under the assumptions of Theorem 5.4 there exists* $0 < \alpha < 1$ *such that*

$$(\mathbf{I} - \mathbf{K}_k \mathbf{H}_k)^\mathsf{T} \mathbf{F}_k^\mathsf{T} \mathbf{\Pi}_{k+1} \mathbf{F}_k (\mathbf{I} - \mathbf{K}_k \mathbf{H}_k) \leq (1 - \alpha) \mathbf{\Pi}_k$$

for all $k \geq 0$.

Proof. (Reif et al., 1999, Lemma 3.1) From the one-step EKF equations of

Algorithm 5.2 it follows that

$$\mathbf{P}_{k+1} = \mathbf{F}_k \mathbf{P}_k \mathbf{F}_k^\mathsf{T} + \widehat{\mathbf{Q}}_k - \mathbf{F}_k \mathbf{P}_k \mathbf{H}_k^\mathsf{T} \mathbf{K}_k^\mathsf{T} \mathbf{F}_k^\mathsf{T},$$

which can be rearranged to

$$\begin{aligned} \mathbf{P}_{k+1} = \mathbf{F}_k(\mathbf{I} - \mathbf{K}_k\mathbf{H}_k)\mathbf{P}_k(\mathbf{I} - \mathbf{K}_k\mathbf{H}_k)^\mathsf{T}\mathbf{F}_k^\mathsf{T} + \widehat{\mathbf{Q}}_k \\ + \mathbf{F}_k\mathbf{K}_k\mathbf{H}_k\mathbf{P}_k(\mathbf{I} - \mathbf{K}_k\mathbf{H}_k)^\mathsf{T}\mathbf{F}_k^\mathsf{T}. \end{aligned} \tag{5.11}$$

First, it will be shown that the last term in this equation is positive-semidefinite. This amounts to showing that $\mathbf{K}_k\mathbf{H}_k\mathbf{P}_k(\mathbf{I}-\mathbf{K}_k\mathbf{H}_k)^\mathsf{T}$ is positive-semidefinite. By the definition of the gain matrix and the matrix inversion lemma

$$\begin{aligned} &\mathbf{K}_k\mathbf{H}_k\mathbf{P}_k(\mathbf{I} - \mathbf{K}_k\mathbf{H}_k)^\mathsf{T} \\ &\quad = \mathbf{P}_k\mathbf{H}_k^\mathsf{T}(\mathbf{H}_k\mathbf{P}_k\mathbf{H}_k^\mathsf{T} + \widehat{\mathbf{R}}_k)^{-1}\mathbf{H}_k[\mathbf{P}_k - \mathbf{P}_k\mathbf{H}_k^\mathsf{T}(\mathbf{H}_k\mathbf{P}_k\mathbf{H}_k^\mathsf{T} + \widehat{\mathbf{R}}_k)^{-1}\mathbf{H}_k\mathbf{P}_k]^\mathsf{T} \\ &\quad = \mathbf{P}_k\mathbf{H}_k^\mathsf{T}(\mathbf{H}_k\mathbf{P}_k\mathbf{H}_k^\mathsf{T} + \widehat{\mathbf{R}}_k)^{-1}\mathbf{H}_k(\mathbf{P}_k^{-1} + \mathbf{H}_k^\mathsf{T}\widehat{\mathbf{R}}_k^{-1}\mathbf{H}_k)^{-\mathsf{T}} \\ &\quad = (\mathbf{P}_k^{-1} + \mathbf{H}_k^\mathsf{T}\widehat{\mathbf{R}}_k^{-1}\mathbf{H}_k)^{-1}\mathbf{H}_k^\mathsf{T}\widehat{\mathbf{R}}_k^{-1}\mathbf{H}_k(\mathbf{P}_k^{-1} + \mathbf{H}_k^\mathsf{T}\widehat{\mathbf{R}}_k^{-1}\mathbf{H}_k)^{-\mathsf{T}}, \end{aligned}$$

which is positive-semidefinite since $\widehat{\mathbf{R}}_k$ has been assumed positive-definite. Therefore

$$\mathbf{P}_{k+1} \geq \mathbf{F}_k(\mathbf{I} - \mathbf{K}_k\mathbf{H}_k)\mathbf{P}_k(\mathbf{I} - \mathbf{K}_k\mathbf{H}_k)^\mathsf{T}\mathbf{F}_k^\mathsf{T} + \widehat{\mathbf{Q}}_k. \tag{5.12}$$

Now, from the non-singularity of $\mathbf{F}_k$ and the matrix inversion lemma it can be seen that $\mathbf{F}_k(\mathbf{I} - \mathbf{K}_k\mathbf{H}_k)$ is non-singular. Hence (5.12) can be written as

$$\begin{aligned} \mathbf{P}_{k+1} \geq (\mathbf{F}_k - \mathbf{F}_k\mathbf{K}_k\mathbf{H}_k)[\mathbf{P}_k + (\mathbf{F}_k - \mathbf{F}_k\mathbf{K}_k\mathbf{H}_k)^{-1}\widehat{\mathbf{Q}}_k(\mathbf{F}_k - \mathbf{F}_k\mathbf{K}_k\mathbf{H}_k)^{-\mathsf{T}}] \\ \times (\mathbf{F}_k - \mathbf{F}_k\mathbf{K}_k\mathbf{H}_k)^\mathsf{T}. \end{aligned}$$

From the various boundedness assumptions it then follows that

$$\mathbf{P}_{k+1} \geq \mathbf{F}_k(\mathbf{I} - \mathbf{K}_k\mathbf{H}_k)\left(\mathbf{P}_k + \frac{\hat{q}}{(f + fp_2h^2/\hat{r})^2}\mathbf{I}\right)(\mathbf{I} - \mathbf{K}_k\mathbf{H}_k)^\mathsf{T}\mathbf{F}_k^\mathsf{T}.$$

Taking inverses of the both sides and multiplying from left by $\mathbf{F}_k(\mathbf{I}-\mathbf{K}_k\mathbf{H}_k)$ and from right by $(\mathbf{I} - \mathbf{K}_k\mathbf{H}_k)^\mathsf{T}\mathbf{F}_k^\mathsf{T}$ yields

$$(\mathbf{I} - \mathbf{K}_k\mathbf{H}_k)^\mathsf{T}\mathbf{F}_k^\mathsf{T}\mathbf{\Pi}_{k+1}\mathbf{F}_k(\mathbf{I} - \mathbf{K}_k\mathbf{H}_k) \leq \left(1 + \frac{\hat{q}}{p_2(f + fp_2h^2/\hat{r})^2}\right)^{-1}$$

and so the claim is obtained by setting

$$1 - \alpha = \left(1 + \frac{\hat{q}}{p_2(f + fp_2h^2/\hat{r})^2}\right)^{-1}. \tag{5.13}$$

□

Only in the proof of the preceding lemma is the non-singularity of the Jacobians $\mathbf{F}_k$ of the dynamic model function $\mathbf{f}$ exploited. The part of the proof where it is shown that the last term of (5.11) is positive-semidefinite has been streamlined a bit. Also, the use of Reif et al. (1999) of the dubious fact that positive-semidefiniteness is preserved under matrix multiplication has been gotten rid of.

Lemma 5.6. *Under the assumptions of Theorem 5.4 there exist* $\varepsilon', \kappa_{nonl} > 0$ *such that*

$$\boldsymbol{\rho}_k^\top \boldsymbol{\Pi}_k [2\mathbf{F}_k(\mathbf{I} - \mathbf{K}_k\mathbf{H}_k)\boldsymbol{\xi}_k + \boldsymbol{\rho}_k] \leq \kappa_{nonl} \|\mathbf{x}_k - \mathbf{m}_k\|^3 \tag{5.14}$$

when $\|\boldsymbol{\xi}_k\| \leq \varepsilon'$.

Proof. (Reif et al., 1999, Lemma 3.2) From the assumptions and the definition of $\mathbf{K}_k$ it follows that $\|\mathbf{K}_k\| \leq p_2h/\hat{r}$. So, (5.7) yields the bound

$$\|\boldsymbol{\rho}_k\| \leq \|\boldsymbol{\varphi}(\mathbf{x}_k, \mathbf{m}_k)\| + \frac{fp_2h}{\hat{r}} \|\boldsymbol{\chi}(\mathbf{x}_k, \mathbf{m}_k)\|$$

and with the choice $\varepsilon' = \min(\varepsilon_{\boldsymbol{\varphi}}, \varepsilon_{\boldsymbol{\chi}})$ it follows that

$$\|\boldsymbol{\rho}_k\| \leq \kappa_{\boldsymbol{\varphi}} \|\boldsymbol{\xi}_k\|^2 + \frac{fp_2h}{\hat{r}} \kappa_{\boldsymbol{\chi}} \|\boldsymbol{\xi}_k\|^2 := \kappa' \|\boldsymbol{\xi}_k\|^2$$

for $\|\boldsymbol{\xi}_k\| \leq \varepsilon'$. Taking the norm of the left-hand side of (5.14) it can be concluded that for $\|\boldsymbol{\xi}_k\| \leq \varepsilon'$

$$\begin{aligned}
&\boldsymbol{\rho}_k^\top \boldsymbol{\Pi}_k [2\mathbf{F}_k(\mathbf{I} - \mathbf{K}_k\mathbf{H}_k)\boldsymbol{\xi}_k + \boldsymbol{\rho}_k] \\
&\qquad \leq 2 \|\boldsymbol{\rho}_k^\top \boldsymbol{\Pi}_k \mathbf{F}_k(\mathbf{I} - \mathbf{K}_k\mathbf{H}_k)\boldsymbol{\xi}_k\| + \|\boldsymbol{\rho}_k^\top \boldsymbol{\Pi}_k \boldsymbol{\rho}_k\| \\
&\qquad \leq 2 \|\boldsymbol{\rho}_k\| \|\boldsymbol{\Pi}_k \mathbf{F}_k(\mathbf{I} - \mathbf{K}_k\mathbf{H}_k)\| \|\boldsymbol{\xi}_k\| + \frac{1}{p_1} \|\boldsymbol{\rho}_k\|^2 \\
&\qquad \leq 2\kappa' \|\boldsymbol{\xi}_k\|^3 \left(f + \frac{fp_2h^2}{p_1\hat{r}}\right) + \frac{1}{p_1}(\kappa')^2 \varepsilon' \|\boldsymbol{\xi}_k\|^3 \\
&\qquad = \kappa_{\text{nonl}} \|\boldsymbol{\xi}_k\|^3,
\end{aligned}$$

where

$$\kappa_{\text{nonl}} = \kappa' \frac{1}{p_1} \left(2 \left(f + \frac{f p_2 h^2}{p_1 \hat{r}} \right) + \kappa' \varepsilon' \right). \tag{5.15}$$

□

Lemma 5.7. *Under the assumptions of Theorem 5.4 there exists $\kappa_{noise} > 0$ independent of δ such that*

$$\mathbb{E}(\boldsymbol{\sigma}_k^\mathsf{T} \mathbf{\Pi}_{k+1} \boldsymbol{\sigma}_k) \leq \kappa_{noise} \delta$$

for all $k \geq 0$.

Proof. (Reif et al., 1999, Lemma 3.3) Since the noise terms are uncorrelated all the cross-terms vanish in

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

As in the proof of the preceeding lemma, $\|\mathbf{K}_k\| \leq p_2 h / \hat{r}$ so that

$$\mathbb{E}(\boldsymbol{\sigma}_k^\mathsf{T} \mathbf{\Pi}_{k+1} \boldsymbol{\sigma}_k) \leq \frac{1}{p_1} \mathbb{E}(\mathbf{q}_k^\mathsf{T} \mathbf{q}_k) + \frac{f^2 h^2 p_2^2}{p_1 \hat{r}^2} \mathbb{E}(\mathbf{r}_k^\mathsf{T} \mathbf{r}_k).$$

Because both sides of this inequality are scalars, taking trace changes nothing and using the properties of trace it follows that

$$\begin{aligned}
\mathbb{E}(\boldsymbol{\sigma}_k^\mathsf{T} \mathbf{\Pi}_{k+1} \boldsymbol{\sigma}_k) &\leq \frac{1}{p_1} \mathbb{E}[\operatorname{tr}(\mathbf{q}_k^\mathsf{T} \mathbf{q}_k)] + \frac{f^2 h^2 p_2^2}{p_1 \hat{r}^2} \mathbb{E}[\operatorname{tr}(\mathbf{r}_k^\mathsf{T} \mathbf{r}_k)] \\
&= \frac{1}{p_1} \mathbb{E}[\operatorname{tr}(\mathbf{q}_k \mathbf{q}_k^\mathsf{T})] + \frac{f^2 h^2 p_2^2}{p_1 \hat{r}^2} \mathbb{E}[\operatorname{tr}(\mathbf{r}_k \mathbf{r}_k^\mathsf{T})] \\
&= \frac{1}{p_1} \mathbb{E}(\operatorname{tr} \mathbf{Q}_k) + \frac{f^2 h^2 p_2^2}{p_1 \hat{r}^2} \mathbb{E}(\operatorname{tr} \mathbf{R}_k) \\
&\leq \frac{1}{p_1} \delta \mathbb{E}(\operatorname{tr} \mathbf{I}_n) + \frac{f^2 h^2 p_2^2}{p_1 \hat{r}^2} \delta \mathbb{E}(\operatorname{tr} \mathbf{I}_m) \\
&= \left(\frac{n}{p_1} + \frac{f^2 h^2 p_2^2 m}{p_1 \hat{r}^2} \right) \delta,
\end{aligned}$$

which is the claim with

$$\kappa_{\text{noise}} := \left(\frac{n}{p_1} + \frac{f^2 h^2 p_2^2 m}{p_1 \hat{r}^2} \right). \tag{5.16}$$

□

Proof of Theorem 5.4. (Reif et al., 1999, Theorem 3.1) The idea of the proof is to construct a suitable stochastic Lyapunov function to meet the conditions of Lemma 3.8. Choose a stochastic Lyapunov function

$$V_k(\boldsymbol{\xi}_k) = \boldsymbol{\xi}_k^\mathsf{T} \boldsymbol{\Pi}_k \boldsymbol{\xi}_k$$

and note that $\boldsymbol{\xi}_k$ satisfies the Markov property. From (5.9) follow the bounds

$$\frac{1}{p_2} \|\boldsymbol{\xi}_k\|^2 \leq V_k(\boldsymbol{\xi}_k) \leq \frac{1}{p_1} \|\boldsymbol{\xi}_k\|^2 \tag{5.17}$$

so the condition (3.5) of Lemma 3.8 is met. Condition (3.4) is a bit trickier to meet, requiring as it does the bounds of Lemmas 5.5 to 5.7. Equation (5.6) yields

$$V_{k+1}(\boldsymbol{\xi}_{k+1}) = [\boldsymbol{\xi}_k^\mathsf{T}(\mathbf{I}-\mathbf{K}_k\mathbf{H}_k)^\mathsf{T}\mathbf{F}_k^\mathsf{T} + \boldsymbol{\rho}_k^\mathsf{T} + \boldsymbol{\sigma}_k^\mathsf{T}]\boldsymbol{\Pi}_{k+1}[\mathbf{F}_k(\mathbf{I}-\mathbf{K}_k\mathbf{H}_k)\boldsymbol{\xi}_k + \boldsymbol{\rho}_k + \boldsymbol{\sigma}_k]$$

and by appying Lemma 5.5 one obtains

$$\begin{aligned} V_{k+1}(\boldsymbol{\xi}_{k+1}) \leq & (1-\alpha)V_k(\boldsymbol{\xi}_k) + \boldsymbol{\rho}_k^\mathsf{T}\boldsymbol{\Pi}_{k+1}[\mathbf{F}_k(\mathbf{I}-\mathbf{K}_k\mathbf{H}_k)\boldsymbol{\xi}_k + \boldsymbol{\rho}_k] \\ & + 2\boldsymbol{\sigma}_k^\mathsf{T}\boldsymbol{\Pi}_{k+1}[\mathbf{F}_k(\mathbf{I}-\mathbf{K}_k\mathbf{H}_k)\boldsymbol{\xi}_k + \boldsymbol{\rho}_k] + \boldsymbol{\sigma}_k^\mathsf{T}\boldsymbol{\Pi}_{k+1}\boldsymbol{\sigma}_k. \end{aligned} \tag{5.18}$$

Taking conditional expectation with respect to $\boldsymbol{\xi}_k$ and using Lemmas 5.6 and 5.7 yields

$$\mathbb{E}[V_{k+1}(\boldsymbol{\xi}_{k+1}) \mid \boldsymbol{\xi}_k] - V_k(\boldsymbol{\xi}_k) \leq -\alpha V_k(\boldsymbol{\xi}_k) + \kappa_{\text{nonl}} \|\boldsymbol{\xi}_k\|^3 + \kappa_{\text{noise}}\delta \tag{5.19}$$

for $\|\boldsymbol{\xi}_k\| \leq \varepsilon' = \min(\varepsilon_{\boldsymbol{\varphi}}, \varepsilon_{\boldsymbol{\chi}})$. The term $2\boldsymbol{\sigma}_k^\mathsf{T}\boldsymbol{\Pi}_{k+1}[\mathbf{F}_k(\mathbf{I}-\mathbf{K}_k\mathbf{H}_k)\boldsymbol{\xi}_k + \boldsymbol{\rho}_k]$ in (5.18) vanishes under conditional expectation since the noise terms are independent of $\boldsymbol{\xi}_k, \boldsymbol{\rho}_k$ as well as the matrices $\mathbf{F}_k, \mathbf{K}_k$ and $\mathbf{H}_k$. Defining

$$\varepsilon = \min\left(\varepsilon', \frac{\alpha}{2p_2\kappa_{\text{nonl}}}\right), \tag{5.20}$$

one has for $\|\boldsymbol{\xi}_k\| \leq \varepsilon$ by (5.17) that

$$\kappa_{\text{nonl}} \|\boldsymbol{\xi}_k\| \|\boldsymbol{\xi}_k\|^2 \leq \frac{\alpha}{2p_2} \|\boldsymbol{\xi}_k\|^2 \leq \frac{\alpha}{2} V_k(\boldsymbol{\xi}_k).$$

Using this in (5.19) one obtains the inequality

$$\mathbb{E}[V_{k+1}(\boldsymbol{\xi}_{k+1}) \mid \boldsymbol{\xi}_k] - V_k(\boldsymbol{\xi}_k) \leq \kappa_{\text{noise}}\delta - \frac{\alpha}{2} V_k(\boldsymbol{\xi}_k) \tag{5.21}$$

for $\|\boldsymbol{\xi}_k\| \leq \varepsilon$. Now, choosing

$$\delta = \frac{\alpha \tilde{\varepsilon}^2}{2p_2 \kappa_{\text{noise}}} \tag{5.22}$$

for some $\tilde{\varepsilon} < \varepsilon$ one has for $\tilde{\varepsilon} \leq \|\boldsymbol{\xi}_k\| \leq \varepsilon$ the supermartingale inequality

$$\begin{aligned}\mathbb{E}[V_{k+1}(\boldsymbol{\xi}_{k+1}) \mid \boldsymbol{\xi}_k] - V_k(\boldsymbol{\xi}_k) &\leq \kappa_{\text{noise}}\delta - \frac{\alpha}{2} V_k(\boldsymbol{\xi}_k) \\ &\leq \frac{\alpha}{2p_2} \|\boldsymbol{\xi}_k\|^2 - \frac{\alpha}{2} V_k(\boldsymbol{\xi}_k) \leq 0\end{aligned}$$

so as to keep the estimation error bounded for all k (Gard, 1988, Theorem 5.2). Thus Lemma 3.8 can be applied when $\|\boldsymbol{\xi}_0\| \leq \varepsilon$. □

Of course, nothing can be said about the bounds (5.8) and (5.9) before the filter is being run. Therefore, it should be understood that the state estimates produced by the filter can be regarded as reliable ones in the sense of the this theorem if the numerically calculated values for $\mathbf{F}_k, \mathbf{H}_k$ and $\mathbf{P}_k$ during the estimation process remain bounded. Of course, one can never be certain that they will always remain uniformly bounded and thus whether or not the filter will begin to diverge at some point in the future. So, applicability of this theorem is very limited as one possesses no mean of ensuring the uniform boundedness of these matrices beforehand. Particularly problematic is the assumption (5.9) on the bounds for $\mathbf{P}_k$ since the boundedness of this matrix is intrisically connected with stability properties as was seen in the linear case where the least trivial part was the verification of the existence of such bounds under uniform detectability assumption in Lemma 4.13.

It is of importance for applying and testing the theorem numerically and for analyzing the relationships of values of different parameters and bounds that Equations (5.20) and (5.22) contain explicit formulas for estimating ε and δ. As will be seen in Section 5.3.5, these estimates tend to be very conservative, meaning that the estimation error does not seem to diverge even with considerably larger initial estimation error and noise terms than permitted by the theorem. This is yet another reason for the small practical value of the theorem.

Even though the bounds are very conservative, the practical utility of enlarging $\widehat{\mathbf{Q}}_k$ and $\widehat{\mathbf{R}}_k$ is foreshadowed in that with these matrices enlarged $\hat{q}$ and $\hat{r}$ can be chosen larger, which through (5.13), (5.15) and (5.16) implies larger — though still extremely conservative — ε and δ if $\varepsilon_{\boldsymbol{\varphi}}$ and $\varepsilon_{\boldsymbol{\chi}}$, influencing ε and δ through (5.20) and (5.22) and dependent on the degree of non-linearity of $\mathbf{f}$ and $\mathbf{h}$, can be chose arbitrarily large. However, as noted in Ford and Coulter (2001) and Ford (2002), where the EKF stability is discussed in the context of a military precision guidance problem, the enlargement of the noise covariance matrices comes with the cost of degraded filter accuracy. The tuning of the covariance matrices will be discussed with more length in Section 5.4.5, general theme of which also applies to this theorem.

Some improvement has been achieved by Rhudy et al. (2012) who consider the EKF in an aircraft attitude estimation problem and are able to relax the bounds for ε and δ, albeit the bounds still remain too strict for this to be a noteworthy improvement. This relaxation of bounds is not limited to the specific application of theirs.

5.3.3 Role of Non-Linear Stability Rank Condition

Reif et al. (1999) also present a theorem linking the non-linear observability rank condition (Nijmeijer, 1982), a certain block matrix rank condition analogous to that of Theorem 4.8, to the stability of the EKF. They consider only an autonomous version of (5.2) without process noise, namely

$$\begin{aligned} \mathbf{x}_k &= \mathbf{f}(\mathbf{x}_{k-1}), \\ \mathbf{y}_k &= \mathbf{h}(\mathbf{x}_k) + \mathbf{r}_k. \end{aligned} \tag{5.23}$$

However, a generalization of the following theorem to additive system with process noise is easily made, but with a very restrictive additional assumption.

Definition 5.8. The discrete-time non-linear dynamic system (5.2) (or the pair $[\mathbf{F}_k, \mathbf{H}_k]$) is said to satisfy the *non-linear observability rank condition* at $\mathbf{x}_k \in \mathbb{R}^n$ if the *non-linear observability matrix*

$$\mathbf{U}(\mathbf{x}_k) := \begin{pmatrix} \frac{\partial \mathbf{h}}{\partial \mathbf{x}}(\mathbf{x}_k) \\ \frac{\partial \mathbf{h}}{\partial \mathbf{x}}(\mathbf{x}_{k+1}) \frac{\partial \mathbf{f}}{\partial \mathbf{x}}(\mathbf{x}_k) \\ \vdots \\ \frac{\partial \mathbf{h}}{\partial \mathbf{x}}(\mathbf{x}_{k+n-1}) \frac{\partial \mathbf{f}}{\partial \mathbf{x}}(\mathbf{x}_{k+n-2}) \cdots \frac{\partial \mathbf{f}}{\partial \mathbf{x}}(\mathbf{x}_k) \end{pmatrix}$$

has rank n.

Theorem 5.9. *Consider the non-linear autonomous dynamic system* (5.23) *and the extended Kalman filter of Algorithm 5.1 and Algorithm 5.2. Suppose* *that there are* $\hat{r}, \hat{q} > 0$ *such that*

$$\begin{aligned} \hat{q}\mathbf{I} &\leq \widehat{\mathbf{Q}}_k, \\ \hat{r}\mathbf{I} &\leq \widehat{\mathbf{R}}_k \end{aligned} \tag{5.24}$$

for all $k \geq 0$ *and a compact set* $K \subset \mathbb{R}^n$ *such that the following conditions hold:*

(1) The system (5.23) *satisfies the non-linear observability rank condition for every* $\mathbf{x}_k \in K$*.*

(2) The functions $\mathbf{f}$ *and* $\mathbf{h}$ *are twice continuously differentiable and the Jacobian of* $\mathbf{f}$ *is non-zero in every point of* K*.*

(3) The sample paths of $\mathbf{x}_k$ *are almost surely bounded and* K *contains the* ε_K*-neighbourhoods of such sample paths, where* ε_K *is some positive real number independent of* k*.*

Then there exist $\varepsilon, \delta > 0$ *such that the conditions* $\|\boldsymbol{\xi}_0\| \leq \varepsilon$ *and* $\mathbf{R}_k \leq \delta\mathbf{I}$ *guarantee that the estimation error* $\boldsymbol{\xi}_k$ *is exponentially bounded in mean square and stochastically bounded.*

Observe that (5.24) implies that $\widehat{\mathbf{Q}}_k$ must be positive-definite, although under consideration is an autonomous system without process noise, meaning that $\mathbf{Q}_k = \mathbf{0}$. Therefore in applying this theorem it is not even possible to choose $\widehat{\mathbf{Q}}_k = \mathbf{Q}_k$.

Only a brief outline of the proof is presented here. For this, the following lemma of Song and Grizzle (1995) is needed.

Lemma 5.10. *Suppose that the non-linear autonomous system* (5.23) *satisfies the non-linear observability rank condition for every* $\mathbf{x}_k$ *in a compact subset* K *of* $\mathbb{R}^n$ *satisfying the assumption (3) of Theorem 5.9. Then there exists* $\varepsilon_{obs} > 0$ *such that the pair* $[\mathbf{F}_k, \mathbf{H}_k]$ *is uniformly observable, given that* $\|\mathbf{x}_k - \mathbf{m}_k\| \leq \varepsilon_{obs}$*.*

Proof. (Song and Grizzle, 1995, Proposition 4.1) Consider a continuous function $\mathbf{J}\colon K \to \mathbb{R}^{n\times n}$ defined by setting $\mathbf{J}(\mathbf{x}) = \mathbf{U}(\mathbf{x})^\mathsf{T}\mathbf{U}(\mathbf{x})$, that is, $\mathbf{J}(\mathbf{x}_k)$ is the observability Gramian associated with the pair $[\mathbf{F}_k, \mathbf{H}_k]$. Since K is compact, $\mathbf{J}$ is uniformly continuous, and thus for every $\delta > 0$ there exists $0 < \varepsilon \leq \varepsilon_K$ such that $\|\mathbf{J}(\mathbf{x}_k) - \mathbf{J}(\mathbf{m}_k)\| \leq \delta$ when $\|\mathbf{x}_k - \mathbf{m}_k\| \leq \varepsilon$. Defining $\alpha_1 := \inf_{\mathbf{x}_0\in K} \lambda_{\min}[\mathbf{J}(\mathbf{x}_k)]$ and $\alpha_2 := \inf_{\mathbf{x}_0\in K} \lambda_{\max}[\mathbf{J}(\mathbf{x}_k)]$ and choosing $\delta = \alpha_1/2$, it follows that

$$\frac{\alpha_1}{2}\mathbf{I} \leq \mathbf{J}(\mathbf{m}_k) \leq \left(\alpha_2 + \frac{\alpha_1}{2}\right)\mathbf{I}$$

whenever $\|\mathbf{x}_k - \mathbf{m}_k\| \leq \varepsilon_{\text{obs}}$ for some sufficiently small $0 < \varepsilon_{\text{obs}} \leq \varepsilon_K$. □

With this lemma, an outline of the proof is plausible.

Proof of Theorem 5.9. (Reif et al., 1999, Theorem 4.1) The assumption (3) implies that $[\mathbf{F}_k, \mathbf{H}_k]$ satisfies the non-linear observability rank condition by Lemma 5.10. Therefore it follows from results in Sections 4.4 and 4.5 that (5.9) holds.

Functions $\mathbf{f}$ and $\mathbf{h}$ being twice differentiable in a compact K, their Taylor series can be used to estimate upper bounds for $\boldsymbol{\varphi}$ and $\boldsymbol{\chi}$ (Nikolsky, 1977, Section 7.13; Hauser, 2004, Section 168). This yields values

$$\begin{aligned} \kappa_{\boldsymbol{\varphi}} &= \max_{1 \leq i \leq n} \sup_{\mathbf{x} \in K} \|\text{Hess}\, \mathbf{f}_i(\mathbf{x})\|, \\ \kappa_{\boldsymbol{\chi}} &= \max_{1 \leq i \leq m} \sup_{\mathbf{x} \in K} \|\text{Hess}\, \mathbf{h}_i(\mathbf{x})\| \end{aligned} \tag{5.25}$$

for (5.10), where $\mathbf{f}_i$ and $\mathbf{h}_i$ are the ith components of $\mathbf{f}$ and $\mathbf{h}$. These formulas can always be used to obtain values for $\kappa_{\boldsymbol{\varphi}}$ and $\kappa_{\boldsymbol{\chi}}$ in this kind of a setting.

The remaining bounds follow by a certain inductive argument similar to that in Song and Grizzle (1995) that links together the boundedness of $\mathbf{F}_k, \mathbf{H}_k$ and $\mathbf{P}_k$ to that of $\boldsymbol{\xi}_k$, allowing one to apply Theorem 5.4 with

$$\varepsilon = \min\left(\varepsilon_{\text{obs}}, \varepsilon', \frac{\alpha}{2p_2\kappa_{\text{nonl}}}\right).$$

□

The proof of Lemma 5.10 is the only part in the proof of Theorem 5.9 where the assumption of autonomous system state equation is needed. The introduction of noise could easily cause disturbances big enough to cast $\mathbf{x}_k$ out of K. The situation can be fixed by requiring the solution $\tilde{\mathbf{x}}_k$ of the non-autonomous system (5.2) to remain close enough to the solution $\mathbf{x}_k$ of the autonomous system. That is, $\|\mathbf{x}_k - \tilde{\mathbf{x}}_k\| \leq \varepsilon_q$ for some sufficiently small $\varepsilon_q > 0$. However, this is a very restrictive assumption. The proof of Lemma 5.10 can be then modified by using the triangle inequality

$$\|\mathbf{x}_k - \mathbf{m}_k\| \leq \|\mathbf{x}_k - \tilde{\mathbf{x}}_k\| + \|\tilde{\mathbf{x}}_k - \mathbf{m}_k\|.$$

5.3.4 Stability of the Non-Additive EKF

A slight generalization, both in terms of applying to a larger class of dynamic systems and in relaxing the assumptions, of Theorem 5.4 is achieved by Kluge et al. (2010) who consider a general non-linear dynamic system (5.1).

They are able to drop the assumption (2) of Theorem 5.4, that is, the non-singularity of the Jacobians of the system model function. They also allow observations to be intermittent, following Sinopoli et al. (2004) who have considered the stability of the time-invariant Kalman filter with intermittent observations. Here this property is ignored.

Therefore, ignoring the possibility of intermittent observations, one has the following, currently the most general, theorem on the stability of the EKF. In proving this theorem, the two-step formulation of Algorithm 5.3 for the non-additive EKF is used. However, this introduces some complications as it is not anymore possible to use conditional expecations due to dependencies inherent to the two-step EKF. Also, instead of the usual Euclidean norm of, the L^2 norm (denoted by $\|\cdot\|_2$) is used in some places.

The functions $\mathbf{f}$ and $\mathbf{h}$ in (5.1) are assumed to be continuously differentiable with respect to their first argument in $\mathbb{R}^n \times \{\mathbf{0}\}$. For their various derivatives the following notation is used:

$$\mathbf{F}_k = \frac{\partial \mathbf{f}}{\partial \mathbf{x}}(\mathbf{m}_k, \mathbf{0}) \qquad \mathbf{H}_k = \frac{\partial \mathbf{h}}{\partial \mathbf{x}}(\mathbf{m}_k^-, \mathbf{0})$$
$$\mathbf{G}_k = \frac{\partial \mathbf{f}}{\partial \mathbf{q}}(\mathbf{m}_k, \mathbf{0}) \qquad \mathbf{D}_k = \frac{\partial \mathbf{h}}{\partial \mathbf{r}}(\mathbf{m}_k^-, \mathbf{0}).$$

Furthermore, in addition to the estimation error $\boldsymbol{\xi}_k$, also the predicted estimation error $\boldsymbol{\xi}_k^- = \mathbf{x}_k - \mathbf{m}_k^-$ is made use of. Using Taylor expansions of $\mathbf{f}$ and $\mathbf{h}$ at $(\mathbf{m}_k, \mathbf{0})$ it is possible to write

$$\begin{aligned} \mathbf{f}(\mathbf{x}_k, \mathbf{q}_k) &= \mathbf{f}(\mathbf{m}_k, \mathbf{0}) + \mathbf{F}_k(\mathbf{x}_k - \mathbf{m}_k) + \mathbf{G}_k \mathbf{q}_k + \boldsymbol{\varphi}(\mathbf{x}_k, \mathbf{m}_k, \mathbf{q}_k), \\ \mathbf{h}(\mathbf{x}_k, \mathbf{r}_k) &= \mathbf{h}(\mathbf{m}_k^-, \mathbf{0}) + \mathbf{H}_k(\mathbf{x}_k - \mathbf{m}_k^-) + \mathbf{D}_k \mathbf{r}_k + \boldsymbol{\chi}(\mathbf{x}_k, \mathbf{m}_k^-, \mathbf{r}_k), \end{aligned}$$

where $\boldsymbol{\varphi}$ and $\boldsymbol{\chi}$ are the remainder functions. Using these expansions and the non-additive extended Kalman filter of Algorithm 5.3 the estimation errors can be expressed recursively as

$$\begin{aligned} \boldsymbol{\xi}_k &= (\mathbf{I} - \mathbf{K}_k \mathbf{H}_k)\boldsymbol{\xi}_k^- - \mathbf{K}_k \mathbf{D}_k \mathbf{r}_k - \mathbf{K}_k \boldsymbol{\chi}(\mathbf{x}_k, \mathbf{m}_k^-, \mathbf{r}_k), \\ \boldsymbol{\xi}_{k+1}^- &= \mathbf{F}_k(\mathbf{I} - \mathbf{K}_k \mathbf{H}_k)\boldsymbol{\xi}_k^- + \boldsymbol{\rho}_k + \boldsymbol{\sigma}_k, \end{aligned} \tag{5.26}$$

where

$$\begin{aligned} \boldsymbol{\rho}_k &= \boldsymbol{\varphi}(\mathbf{x}_k, \mathbf{m}_k, \mathbf{q}_k) - \mathbf{F}_k \mathbf{K}_k \boldsymbol{\chi}(\mathbf{x}_k, \mathbf{m}_k^-, \mathbf{r}_k), \\ \boldsymbol{\sigma}_k &= \mathbf{G}_k \mathbf{q}_k - \mathbf{F}_k \mathbf{K}_k \mathbf{D}_k \mathbf{r}_k. \end{aligned}$$

Theorem 5.11. *Consider the non-linear dynamic system* (5.1) *and the ex-*

tended Kalman filter of Algorithm 5.3. Suppose that the following conditions hold:

(1) There exist $f, g, h, d, \hat{q}_1, \hat{q}_2, \hat{r}_1, \hat{r}_2, p_1, p_2 > 0$ *such that*

$$\begin{aligned} &\|\mathbf{F}_k\| \le f, \quad \|\mathbf{G}_k\| \le g, \quad \|\mathbf{H}_k\| \le h, \quad \|\mathbf{D}_k\| \le d, \\ &\hat{q}_1\mathbf{I} \le \widehat{\mathbf{Q}}_k \le \hat{q}_2\mathbf{I}, \qquad \hat{r}_1\mathbf{I} \le \widehat{\mathbf{R}}_k \le \hat{r}_2\mathbf{I}, \\ &p_1\mathbf{I} \le \mathbf{P}_k \le \mathbf{P}_k^- \le p_2\mathbf{I} \end{aligned}$$

for all $k \ge 0$*.*

(2) For all $\varepsilon_{\boldsymbol{\varphi}}, \varepsilon_{\boldsymbol{\chi}} > 0$ *there are* $\kappa_{\boldsymbol{\varphi}}, \kappa_{\boldsymbol{\chi}} > 0$ *such that if* $\|\mathbf{x}_k - \mathbf{m}_k\|_2 < \varepsilon_{\boldsymbol{\varphi}}$ *and* $\|\mathbf{x}_k - \mathbf{m}_k^-\|_2 < \varepsilon_{\boldsymbol{\chi}}$*, respectively, then*

$$\begin{aligned} \|\boldsymbol{\varphi}(\mathbf{x}_k, \mathbf{m}_k, \mathbf{q}_k)\|_2 &\le \kappa_{\boldsymbol{\varphi}} \|\mathbf{x}_k - \mathbf{m}_k\|_2^2, \\ \|\boldsymbol{\chi}(\mathbf{x}_k, \mathbf{m}_k^-, \mathbf{r}_k)\|_2 &\le \kappa_{\boldsymbol{\chi}} \|\mathbf{x}_k - \mathbf{m}_k^-\|_2^2. \end{aligned}$$

Then there exist $\varepsilon, \delta > 0$ *such that the conditions* $\|\boldsymbol{\xi}_1^-\| \le \varepsilon$ *and* $\mathbf{Q}_k, \mathbf{R}_k \le \delta\mathbf{I}$ *guarantee that the predicted estimation error* $\boldsymbol{\xi}_k^-$ *is exponentially bounded in mean square and stochastically bounded.*

Proof. (Unbehauen, 1998, Satz IX.9; Kluge et al., 2010, Theorem 3.2) Denote $\boldsymbol{\Pi}_k = (\mathbf{P}_k^-)^{-1}$ and define a Lyapunov function by

$$V_k(\boldsymbol{\xi}_k^-) = (\boldsymbol{\xi}_k^-)^\mathsf{T}\boldsymbol{\Pi}_k\boldsymbol{\xi}_k^-.$$

Then if follows from (5.26) that

$$\begin{aligned} V_{k+1}(\boldsymbol{\xi}_{k+1}^-) = (\boldsymbol{\xi}_k^-)^\mathsf{T}&(\mathbf{I} - \mathbf{K}_k\mathbf{H}_k)^\mathsf{T}\mathbf{F}_k^\mathsf{T}\boldsymbol{\Pi}_{k+1}\mathbf{F}_k(\mathbf{I} - \mathbf{K}_k\mathbf{H}_k)\boldsymbol{\xi}_k^- \\ &+ \boldsymbol{\rho}_k^\mathsf{T}\boldsymbol{\Pi}_{k+1}[2\mathbf{F}_k(\mathbf{I} - \mathbf{K}_k\mathbf{H}_k)\boldsymbol{\xi}_k^- + \boldsymbol{\rho}_k] \\ &+ 2\boldsymbol{\sigma}_k^\mathsf{T}\boldsymbol{\Pi}_{k+1}[\mathbf{F}_k(\mathbf{I} - \mathbf{K}_k\mathbf{H}_k)\boldsymbol{\xi}_k^- + \boldsymbol{\rho}_k] + \boldsymbol{\sigma}_k^\mathsf{T}\boldsymbol{\Pi}_{k+1}\boldsymbol{\sigma}_k. \end{aligned}$$

The error covariance matrix $\mathbf{P}_k$ can be manipulated into

$$\begin{aligned} \mathbf{P}_k &= (\mathbf{I} - \mathbf{K}_k\mathbf{H}_k)\mathbf{P}_k^-(\mathbf{I} - \mathbf{K}_k\mathbf{H}_k)^\mathsf{T} + \mathbf{K}_k\mathbf{D}_k\widehat{\mathbf{R}}_k\mathbf{D}_k^\mathsf{T}\mathbf{K}_k^\mathsf{T} \\ &\ge \left(1 + \frac{q_1}{2f^2p_2}\right)\mathbf{F}_k(\mathbf{I} - \mathbf{K}_k\mathbf{H}_k)\mathbf{P}_k^-(\mathbf{I} - \mathbf{K}_k\mathbf{H}_k)^\mathsf{T}\mathbf{F}_k^\mathsf{T} \end{aligned}$$

on which Lemma A.2 can be utilized to yield

$$(\mathbf{I} - \mathbf{K}_k\mathbf{H}_k)\mathbf{F}_k^\mathsf{T}\boldsymbol{\Pi}_{k+1}\mathbf{F}_k(\mathbf{I} - \mathbf{K}_k\mathbf{H}_k) \le (1 - \alpha)\boldsymbol{\Pi}_k,$$

where

$$0 < \alpha = \frac{q_2}{2f^2 p_2 + q_2} < 1.$$

As opposed to the proof of Theorem 5.4, the non-singularity of $\mathbf{F}_k$ is not needed this time.

Thus one has

$$\begin{aligned} V_{k+1}(\boldsymbol{\xi}_{k+1}^-) \leq (1-\alpha)V_k(\boldsymbol{\xi}_k^-) &+ \boldsymbol{\rho}_k^\mathsf{T}\boldsymbol{\Pi}_{k+1}[2\mathbf{F}_k(\mathbf{I}-\mathbf{K}_k\mathbf{H}_k)\boldsymbol{\xi}_k^- + \boldsymbol{\rho}_k] \\ &+ 2\boldsymbol{\sigma}_k^\mathsf{T}\boldsymbol{\Pi}_{k+1}[\mathbf{F}_k(\mathbf{I}-\mathbf{K}_k\mathbf{H}_k)\boldsymbol{\xi}_k^- + \boldsymbol{\rho}_k] + \boldsymbol{\sigma}_k^\mathsf{T}\boldsymbol{\Pi}_{k+1}\boldsymbol{\sigma}_k, \end{aligned}$$

which is analogous to (5.18). From this point one can proceed by taking the expectation and use results similar to Lemmas 5.6 and 5.7 to obtain an inequality similar to (5.21). Then the exponential boundedness in mean square and stochastic boundedness can be concluded as in the proof of Lemma 3.8. This has been carried out detailedly in Unbehauen (1998, Satz IX.9), regrettably in German, though. □

5.3.5 Numerical Simulations

This section briefly examines the effect of covariance matrix tuning on the maximum initial estimation error of the EKF numerically. The system for which simulations are carried out is the non-linear system apparently first investigated by Reif et al. (1999). As such, this system has become somewhat a standard system for demonstrating the stability of the EKF and the UKF. Such numerical simulations have already been carried out in Rapp and Nyman (2005); Xiong et al. (2006); Dymirkovsky (2012) and Li and Xia (2012). Since basic stability simulations have been carried out for this system, this section concentrates on more explicit analysis of the effects of tuning the covariance matrices.

Consider the system (5.2) with the system and measurement model functions given by

$$\begin{aligned} \mathbf{f}(\mathbf{x}_k) &= \begin{pmatrix} x_{1,k} + \tau x_{2,k} \\ x_{2,k} + \tau[-x_{1,k} + (x_{1,k}^2 + x_{2,k}^2 - 1)x_{2,k}] \end{pmatrix}, \\ \mathbf{h}(\mathbf{x}_k) &= \mathbf{x}_k, \end{aligned} \tag{5.27}$$

where $x_{1,k}$ and $x_{2,k}$ are the components of $\mathbf{x}_k$ and $\tau = 10^{-3}$. The components of this system are plotted with 10,000 time-steps in Figure 5.1. The noise enters to the system additively with $\mathbf{Q}_{k-1} = 0.01\mathbf{I}$ and $\mathbf{R}_k = \mathbf{I}$. The initial system state is taken to be $\mathbf{x}_0 = (0.8, 0.2)$ and initial error covariance $\mathbf{P}_0 = \mathbf{I}$.

Reif et al. (1999) were able to calculate bounds for the initial estima-

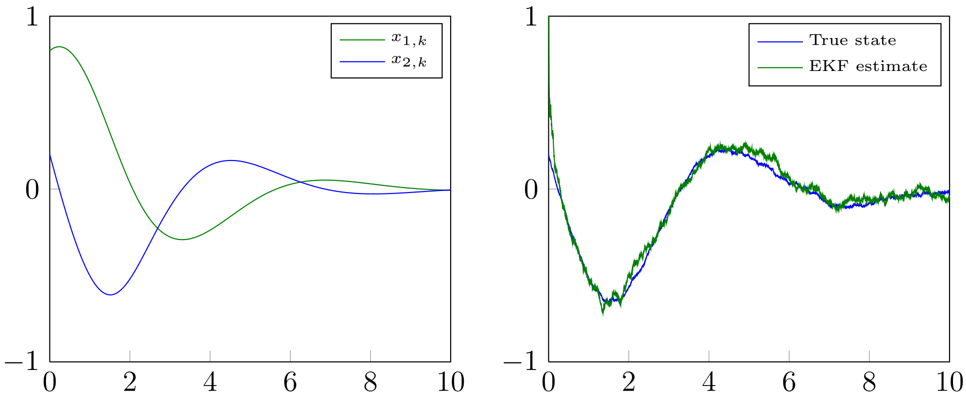

Figure 5.1: On the left are plotted the trajectories of the components of the system with (5.27) without process noise. On the right, with process and measurement noise, are $x_{2,k}$ and its EKF estimate with $\widehat{\mathbf{Q}}_{k-1} = \tau\mathbf{I}$ and $\widehat{\mathbf{R}}_k = \tau^{-1}\mathbf{I}$. The initial estimate is $\mathbf{m}_0 = (2, 2)$, giving $\|\boldsymbol{\xi}_0\| = 2.16$. MSE of the estimation error in $x_{2,k}$ is 0.0041 and after 500th time-step 0.0011. Time-steps on the x-axis are in thousands.

tion error and the noise covariance matrices as permitted by Theorem 5.4. For $\widehat{\mathbf{Q}}_{k-1} = \tau\mathbf{I}$ and $\widehat{\mathbf{R}}_k = \tau^{-1}\mathbf{I}$ the stability is guaranteed only when $\|\boldsymbol{\xi}_0\| \leq 5 \cdot 10^{-3}$ and $\mathbf{Q}_{k-1}, \mathbf{R}_k \leq 10^{-10}\mathbf{I}$. These are very strict bounds, given that even with the noise covariances of the preceding paragraph and an initial estimation error of significantly greater magnitude there is no problem of filter divergence as evidenced by Figure 5.1.

In particular, the interest is here in examining how tuning the noise covariance matrices $\widehat{\mathbf{Q}}_{k-1}$ and $\widehat{\mathbf{R}}_k$ affects the magnitude of possible initial estimation error and the filter accuracy. First, the effect of enlarging the matrix $\widehat{\mathbf{Q}}_k$ in the EKF algorithm on the upper bound for the initial estimation error $\boldsymbol{\xi}_0$ is tested, taking $\widehat{\mathbf{R}}_k = \tau^{-1}\mathbf{I}$. The result for a particular noise realization is given in Figure 5.2 with $\widehat{\mathbf{Q}}_{k-1}$ ranging from $10^{-3}\mathbf{I}$ to $5\mathbf{I}$. The increasing MSE renders the filter rather useless for large initial estimation error, strongly suggesting decreasing the magnitude of $\widehat{\mathbf{Q}}_{k-1}$ after the large initial estimation error has been eliminated. However, the matrix $\widehat{\mathbf{R}}_k$ has an opposite effect for this model, contradicting the theoretical results of the preceding section. Fixing $\widehat{\mathbf{Q}}_{k-1} = 0.01\mathbf{I}$, Figure 5.3 shows that reducing the size of $\widehat{\mathbf{R}}_k$ leads to larger maximum initial estimation error. It is not clear what causes this phenonmenon. Overally then, the effect on the MSE is a bit mixed.

The unsatisfactory behavior of the filter estimate is illustrated in Figure 5.4. Also illustrated is an improvement obtained when letting $\widehat{\mathbf{Q}}_{k-1} \to \mathbf{Q}_{k-1}$ as k increases. This approach is more detailedly described in Section 5.4.5.

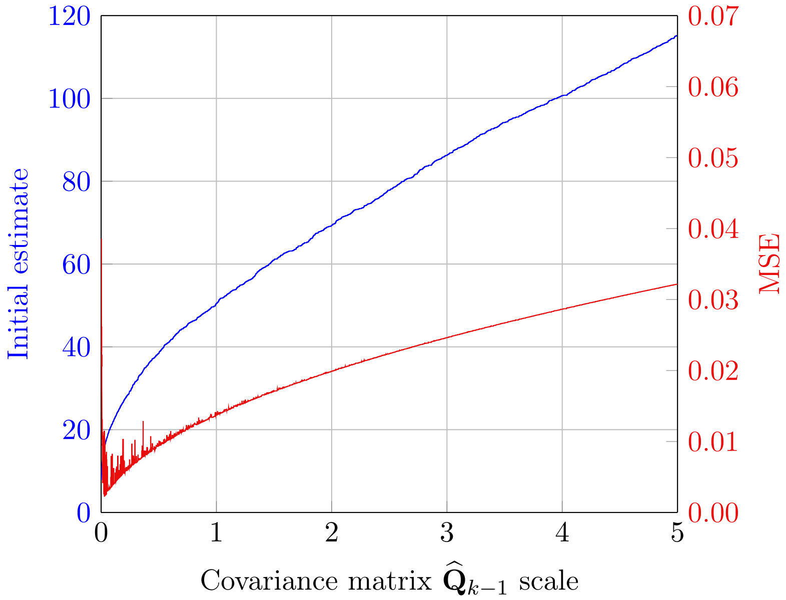

Figure 5.2: Plot of the effect of enlarging $\widehat{\mathbf{Q}}_{k-1} = a\mathbf{I}$ from $10^{-3}\mathbf{I}$ to $5\mathbf{I}$ to the maximum of the initial estimate $\mathbf{m}_0 = (b, b)$ such that the estimation error remains bounded for the system with (5.27). On the x-axis is a. On the y-axis the blue plot is b and the red is the MSE of the estimates with $\mathbf{m}_0 = (b, b)$ and $\widehat{\mathbf{Q}}_{k-1} = b\mathbf{I}$. The MSE is calculated from 500th time-step so as to not let large initial estimation error skew it too much.

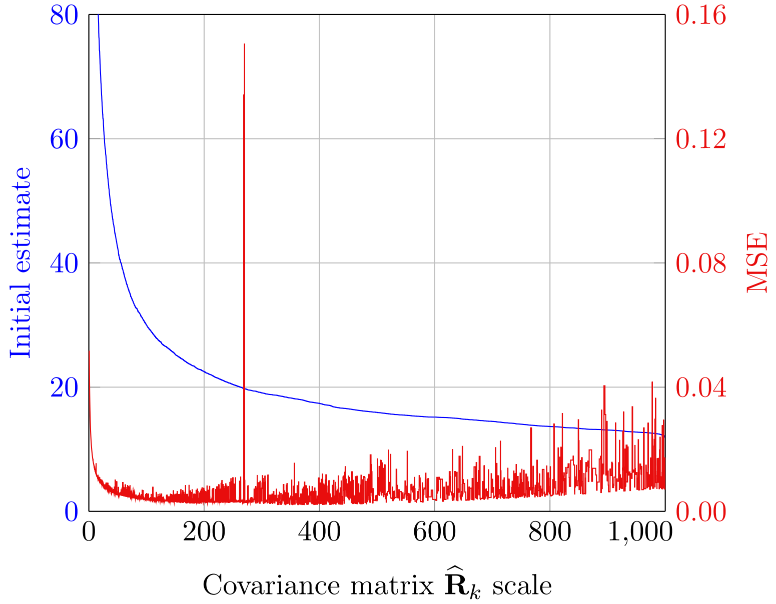

Figure 5.3: Plot of the effect of enlarging $\widehat{\mathbf{R}}_k = a\mathbf{I}$ from $\mathbf{I}$ to $1000\mathbf{I}$ to the maximum of the initial estimate $\mathbf{m}_0 = (b, b)$ such that the estimation error remains bounded for the system with (5.27). On the x-axis is a. On the y-axis the blue plot is b and the red is the MSE of the estimates with $\mathbf{m}_0 = (b, b)$ and $\widehat{\mathbf{R}}_k = b\mathbf{I}$, calculated again from 500th time-step. The rapid growth of b continues when nearing zero covariance matrix, for $\widehat{\mathbf{R}}_k = \mathbf{I}$ one has $b = 623$. The remarks for Figure 5.2 apply to these plots too.

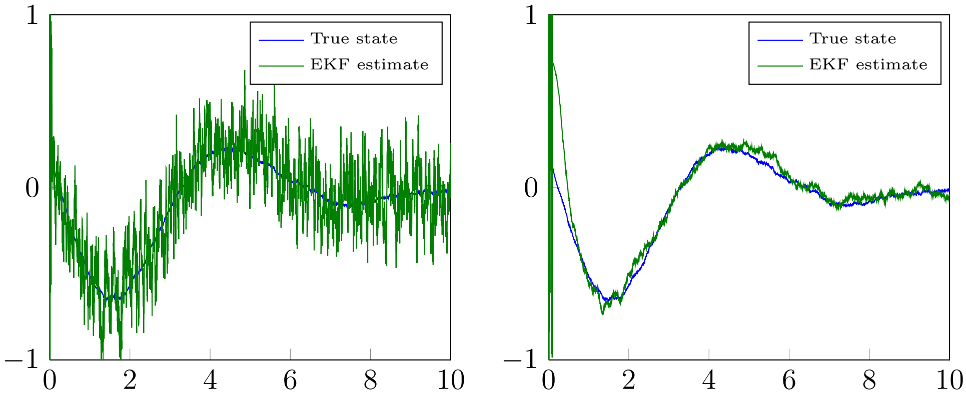

Figure 5.4: On the left is the EKF estimate of $x_{2,k}$ of the system with (5.27) when $\widehat{\mathbf{Q}}_{k-1} = 3\mathbf{I}$, $\widehat{\mathbf{R}}_k = \tau^{-1}\mathbf{I}$ and $\mathbf{m}_0 = (82, 82)$. On the right is an estimate in the same situation but with $\widehat{\mathbf{Q}}_{k-1} = \mathbf{Q}_{k-1} + 10\mathbf{P}_{k-1}e^{-(k-1)}$. As can be seen, the first estimate fluctuates rapidly and its MSE after the 500th time-step is 0.0246, whereas in the second case this MSE is 0.0013. Time-steps on the x-axis are in thousands.

5.4 Gaussian Filters and the Unscented Kalman Filter

The unscented Kalman filter was introduced by Julier et al. (1995) as an alternative to the EKF for non-linear filtering problems. The UKF employs the unscented transform (UT) which is a method for calculating the mean and covariance of a non-linearly transformed Gaussian distribution. The UT is based on propagating through the non-linearity a finite number of deterministically chosen sigma points that capture the original mean and covariance, resulting thus to an approximation of the mean and covariance of the transformed distribution. As the UKF is not based on a linearization by Taylor series, the evaluation of Jacobian matrices, prone to error, need not be done. The approximated mean generated by the UT agrees with the true mean up to the third order, whereas the linearization used in the EKF is accurate only up to the first order (Simon, 2006, Section 14.2.1). This usually results into improved accuracy (Julier et al., 2000; Ristic et al., 2003) of the UKF over the EKF. The UT and linearization both yield approximated covariances correct up to the third order but the UT can be expected to introduce less error in terms of higher order (Simon, 2006, Section 14.2.2).

It did not take long for Ito and Xiong (2000) to remark that the UKF belongs to a wider class of Gaussian filters. These filters are based on Gaussian density approximations to the prediction and filtering densities and encompass a large number of different filters (Nørgaard et al., 2000; Wu et al., 2006).

The first stability results for the UKF were derived by Xiong et al. (2006, 2007a). The more useful results of Xiong et al. (2006) concerned a non-linear system with linear measurement model and were soon observed to extend with virtually no modifications for Gaussian filters also (Wu et al., 2007; Xiong et al., 2007b). In the latter paper a theorem from Xiong (2006) was presented, giving conditions for the stability of a limited class of Gaussian filters for non-linear systems with measurement model also non-linear. An accessible proof of this theorem is was finally given in Xiong et al. (2009a). Li and Xia (2012) extended this in the case of the UKF to the scenario of intermittent observations and provide a different proof. The assumptions of these theorems are mainly upper and lower bounds for certain the residual-correcting random matrices introduced and implicitly contain the requirement of sufficiently small initial estimation error. However, as opposed to the results for the EKF, the proofs indicate no quantitative bounds for the initial estimation error or the mean square estimation error.

All results on the stability of the UKF and Gaussian filters are based on the similar use of Lemma 3.8 as those involving the EKF. However, since it is not generally possible to obtain a simple recursive expression similar to (5.6), one is forced to introduce random diagonal matrices to transform a first-order Taylor series based recursive approximation for the estimation error into an equation, that is, correcting the residuals in this first-order approximation. This approach originated in Boutayeb et al. (1997) and Boutayeb and Aubry (1999) where it was used to prove certain stability results for the EKF as an observer for deterministic non-linear systems.

The main contribution of this stability analysis has not been an introduction of easily verifiable conditions for stability or the strength of the stability results so obtained. Rather, the most important achievement seems to have been showing that by tuning the noise covariance matrices stability may be induced even with considerable initial estimation error. This aspect is also present in the stability results for the EKF, as remarked, although this is not explicitly made clear by Reif et al. (1999) and Kluge et al. (2010). As no quantitative bounds are provided by the proofs and the verification of assumptions is exceptionally hard, the stability analysis for the UKF and Gaussian filters remains very qualitative in nature.

Because the most general proofs of Xiong et al. (2009a) and Li and Xia (2012) contain some errors, the main stability theorem as well as its proof presented here are somewhat modified from what has been published previously. In addition to this theorem some discussion is given to the effect of covariance matrix tuning.

5.4.1 Gaussian Filter Equations

The idea of Gaussian filters is to assume that the filtering and prediction distributions $p(\mathbf{x}_k \mid \mathbf{y}_{1:k})$ and $p(\mathbf{x}_k \mid \mathbf{y}_{1:k-1})$ are Gaussian and use moment matching to approximate the non-linearly transformed densities (Särkkä, 2013, Chapter 6). The matrices are $\widehat{\mathbf{Q}}_{k-1}$ and $\widehat{\mathbf{R}}_k$ are again some time-varying symmetric positive-definite matrices as with the EKF.

Algorithm 5.12 (Gaussian filter)**.** *A Gaussian filter computes the Gaussian approximations to the prediction and filtering distributions of a non-linear dynamic system* (5.2)*. The prediction distribution* (2.2) *is approximated by* $\mathcal{N}(\mathbf{m}_k^-, \mathbf{P}_k^-)$ *and the filtering distribution* (2.1) *by* $\mathcal{N}(\mathbf{m}_k, \mathbf{P}_k)$*. The parameters of these Gaussian distributions are computed by the prediction step*

$$\mathbf{m}_k^- = \int_{\mathbb{R}^n} \mathbf{f}(\mathbf{x}_{k-1}) \mathcal{N}(\mathbf{x}_{k-1} \mid \mathbf{m}_{k-1}, \mathbf{P}_{k-1}) \, \mathrm{d}\mathbf{x}_{k-1}, \tag{5.28}$$

$$\begin{aligned} \mathbf{P}_k^- = \int_{\mathbb{R}^n} & [\mathbf{f}(\mathbf{x}_{k-1}) - \mathbf{m}_k^-][\mathbf{f}(\mathbf{x}_{k-1}) - \mathbf{m}_k^-]^\mathsf{T} \\ & \times \mathcal{N}(\mathbf{x}_{k-1} \mid \mathbf{m}_{k-1}, \mathbf{P}_{k-1}) \, \mathrm{d}\mathbf{x}_{k-1} + \widehat{\mathbf{Q}}_{k-1} \end{aligned} \tag{5.29}$$

and the update step

$$\begin{aligned} \boldsymbol{\mu}_k &= \int_{\mathbb{R}^n} \mathbf{h}(\mathbf{x}_k) \mathcal{N}(\mathbf{x}_k \mid \mathbf{m}_k^-, \mathbf{P}_k^-) \, \mathrm{d}\mathbf{x}_k, \\ \mathbf{S}_k &= \int_{\mathbb{R}^n} [\mathbf{h}(\mathbf{x}_k) - \boldsymbol{\mu}_k][\mathbf{h}(\mathbf{x}_k) - \boldsymbol{\mu}_k]^\mathsf{T} \mathcal{N}(\mathbf{x}_k \mid \mathbf{m}_k^-, \mathbf{P}_k^-) \, \mathrm{d}\mathbf{x}_k + \widehat{\mathbf{R}}_k, \\ \mathbf{C}_k &= \int_{\mathbb{R}^n} [\mathbf{x}_k - \mathbf{m}_k^-][\mathbf{h}(\mathbf{x}_k) - \boldsymbol{\mu}_k]^\mathsf{T} \mathcal{N}(\mathbf{x}_k \mid \mathbf{m}_k^-, \mathbf{P}_k^-) \, \mathrm{d}\mathbf{x}_k, \\ \mathbf{K}_k &= \mathbf{C}_k \mathbf{S}_k^{-1}, \\ \mathbf{m}_k &= \mathbf{m}_k^- + \mathbf{K}_k(\mathbf{y}_k - \boldsymbol{\mu}_k), \\ \mathbf{P}_k &= \mathbf{P}_k^- - \mathbf{K}_k \mathbf{S}_k \mathbf{K}_k^\mathsf{T}. \end{aligned} \tag{5.30}$$

The above general Gaussian filter algorithm is not a practical solution to the non-linear filtering problem as it is in many cases impossible to evaluate the integrals in closed form. Luckily, there are several efficient numerical methods for computing the integrals that include powerful methods of Gauss–Hermite integration and spherical cubature integration (Särkkä, 2013, Sections 6.4 and 6.5).

5.4.2 The Unscented Kalman Filter Equations

Next, the equations constituting the unscented Kalman filter are presented although their exact form is of little importance to the stability analysis to be carried out. In the UKF the unscented transform is used to approximate the five integrals whose computation is a prerequisite in using the Gaussian filter. In the unscented transform this approximation is done by selecting a set of deterministically chosen sigma-points that are then propagated through the non-linear transformation. After this, the required approximate means and covariances are calculated as a weighted sums of the transformed sigma-points.

Algorithm 5.13 (Unscented Kalman filter)**.** *The unscented Kalman filter equations for the non-linear dynamic system* (5.2) *start with the prediction step:*

(1) Form the sigma points $\boldsymbol{\mathcal{X}}_{k-1}^{(j)}$ *for* $j = 0, 1, \ldots, 2n$*:*

$$\begin{aligned} \boldsymbol{\mathcal{X}}_{k-1}^{(0)} &= \mathbf{m}_{k-1}, \\ \boldsymbol{\mathcal{X}}_{k-1}^{(i)} &= \mathbf{m}_{k-1} + \sqrt{n+\lambda}\left[\sqrt{\mathbf{P}_{k-1}}\right]_i, \\ \boldsymbol{\mathcal{X}}_{k-1}^{(i+n)} &= \mathbf{m}_{k-1} - \sqrt{n+\lambda}\left[\sqrt{\mathbf{P}_{k-1}}\right]_i, \end{aligned}$$

where $i = 1, 2, \ldots, n$ *and* $\left[\sqrt{\mathbf{P}_{k-1}}\right]_i$ *is the ith column of the square root of the matrix* $\mathbf{P}_{k-1}$*. The scalar* λ *is a scaling parameter determining the spread of sigma points around the mean.*

(2) Propagate the sigma points through the dynamic model:

$$\widehat{\boldsymbol{\mathcal{X}}}_k^{(i)} = \mathbf{f}\big(\mathcal{X}_{k-1}^{(i)}\big)$$

for $i = 0, 1, \ldots, 2n$*.*

(3) Compute the predicted mean $\mathbf{m}_k^-$ *and the predicted covariance* $\mathbf{P}_k^-$*:*

$$\begin{aligned} \mathbf{m}_k^- &= \sum_{i=0}^{2n} W_i^{(m)} \widehat{\boldsymbol{\mathcal{X}}}_k^{(i)}, \\ \mathbf{P}_k^- &= \sum_{i=0}^{2n} W_i^{(c)} \big(\widehat{\boldsymbol{\mathcal{X}}}_k^{(i)} - \mathbf{m}_k^-\big)\big(\widehat{\boldsymbol{\mathcal{X}}}_k^{(i)} - \mathbf{m}_k^-\big)^\mathsf{T} + \widehat{\mathbf{Q}}_{k-1}, \end{aligned} \tag{5.31}$$

where $W_i^{(m)}$ *and* $W_i^{(c)}$ *are suitably selected scalar weights.*

After prediction, the update step is performed:

(1) *Form the sigma points:*

$$\begin{aligned}
\boldsymbol{\mathcal{X}}_k^{-(0)} &= \mathbf{m}_k^-, \\
\boldsymbol{\mathcal{X}}_k^{-(i)} &= \mathbf{m}_k^- + \sqrt{n+\lambda}\left[\sqrt{\mathbf{P}_{k-1}}\right]_i, \\
\boldsymbol{\mathcal{X}}_k^{-(i+n)} &= \mathbf{m}_k^- - \sqrt{n+\lambda}\left[\sqrt{\mathbf{P}_{k-1}}\right]_i,
\end{aligned}$$

where $i = 1, 2, \ldots, n$.

(2) *Propagate the sigma points through the measurement model:*

$$\widehat{\boldsymbol{\mathcal{Y}}}_k^{(i)} = \mathbf{h}(\boldsymbol{\mathcal{X}}_k^{-(i)})$$

for $i = 0, 1, \ldots, 2n$.

(3) *Compute the predicted mean* $\boldsymbol{\mu}_k$*, the predicted covariance of the measurement* $\mathbf{S}_k$ *and the cross-covariance of the state and the measurement* $\mathbf{C}_k$*:*

$$\begin{aligned}
\boldsymbol{\mu}_k &= \sum_{i=0}^{2n} W_i^{(m)} \widehat{\boldsymbol{\mathcal{Y}}}_k^{(i)}, \\
\mathbf{S}_k &= \sum_{i=0}^{2n} W_i^{(c)} \big(\widehat{\boldsymbol{\mathcal{Y}}}_k^{(i)} - \boldsymbol{\mu}_k\big)\big(\widehat{\boldsymbol{\mathcal{Y}}}_k^{(i)} - \boldsymbol{\mu}_k\big)^\mathsf{T} + \widehat{\mathbf{R}}_k, && (5.32) \\
\mathbf{C}_k &= \sum_{i=0}^{2n} W_i^{(c)} \big(\boldsymbol{\mathcal{X}}_k^{-(i)} - \mathbf{m}_k^-\big)\big(\widehat{\boldsymbol{\mathcal{Y}}}_k^{(i)} - \boldsymbol{\mu}_k\big)^\mathsf{T}. && (5.33)
\end{aligned}$$

(4) *Compute the filter gain* $\mathbf{K}_k$*, the filtered state mean estimate* $\mathbf{m}_k$ *and the error covariance* $\mathbf{P}_k$*:*

$$\begin{aligned}
\mathbf{K}_k &= \mathbf{C}_k \mathbf{S}_k^{-1}, && (5.34) \\
\mathbf{m}_k &= \mathbf{m}_k^- + \mathbf{K}_k(\mathbf{y}_k - \boldsymbol{\mu}_k), \\
\mathbf{P}_k &= \mathbf{P}_k^- - \mathbf{K}_k \mathbf{S}_k \mathbf{K}_k^\mathsf{T}. && (5.35)
\end{aligned}$$

The algorithm, including the exact form of the parameter λ and weights $W_i^{(m)}$ and $W_i^{(c)}$, can be found in Wan and van der Merwe (2001); Julier and Uhlmann (2004, Chapter 5) and Särkkä (2013). A detailed and instructive treatment is in Simon (2006, Chapter 14). Computational needs can be lessened with sacrificed accuracy by reusing the sigma points generated in the prediction step instead of generating new in the beginning of the update step.

5.4.3 Stability of the UKF

Consider the additive non-linear discrete-time system (5.2). As assumed, $\mathbf{f}$ and $\mathbf{h}$ are continuously differentiable. The classical approach is to use Taylor series to form a first-order approximation of innovation $\mathbf{v}_k := \mathbf{y}_k - \boldsymbol{\mu}_k$ and the predicted estimation error $\boldsymbol{\xi}_{k+1}^-$ as

$$\begin{aligned}\boldsymbol{\xi}_{k+1}^- &\approx \mathbf{F}_k\boldsymbol{\xi}_k + \mathbf{q}_k,\\ \mathbf{v}_k &\approx \mathbf{H}_k\boldsymbol{\xi}_k^- + \mathbf{r}_k,\end{aligned}$$

where $\mathbf{F}_k$ is the Jacobian of $\mathbf{f}$ evaluated at $\mathbf{m}_k$ and $\mathbf{H}_k$ is the Jacobian of $\mathbf{h}$ evaluated at $\mathbf{m}_k^-$. There exist residuals of $\boldsymbol{\xi}_k^-$ and $\mathbf{v}_k$, the magnitude of which depends on the validity of the local first-order linearization about the given point and the non-linearities of the functions. As such, these approximations cannot be used in any rigorous treatment of the boundedness of the estimation error. However, following the approach demonstrated in Boutayeb et al. (1997) and Boutayeb and Aubry (1999), this problem can be circumvented and a qualitative analysis performed. The residuals can be taken into account and an equality obtained by introducing unknown random diagonal matrices $\boldsymbol{\Gamma}_k$ and $\mathbf{Z}_k$ such that

$$\boldsymbol{\xi}_{k+1}^- = \boldsymbol{\Gamma}_k\mathbf{F}_k\boldsymbol{\xi}_k + \mathbf{q}_k, \tag{5.36}$$

$$\mathbf{v}_k = \mathbf{Z}_k\mathbf{H}_k\boldsymbol{\xi}_k^- + \mathbf{r}_k. \tag{5.37}$$

The equations can be modified into a recursive form for the estimation error by recognizing that

$$\boldsymbol{\xi}_k = \mathbf{x}_k - \mathbf{m}_k = \mathbf{x}_k - \mathbf{m}_k^- - \mathbf{K}_k\mathbf{v}_k = \boldsymbol{\xi}_k^- - \mathbf{K}_k\mathbf{v}_k.$$

Inserting (5.36) into this and using (5.37) then yields a recursive equation

$$\begin{aligned}\boldsymbol{\xi}_k &= \boldsymbol{\Gamma}_{k-1}\mathbf{F}_{k-1}\boldsymbol{\xi}_{k-1} + \mathbf{q}_{k-1} - \mathbf{K}_k\mathbf{v}_k\\ &= (\mathbf{I} - \mathbf{K}_k\mathbf{Z}_k\mathbf{H}_k)\boldsymbol{\Gamma}_{k-1}\mathbf{F}_{k-1}\boldsymbol{\xi}_{k-1} + (\mathbf{I} - \mathbf{K}_k\mathbf{Z}_k\mathbf{H}_k)\mathbf{q}_{k-1} - \mathbf{K}_k\mathbf{r}_k,\end{aligned} \tag{5.38}$$

pivotal to the stability analysis to come. With this in mind, the error covariance matrix can be written as

$$\mathbf{P}_{k+1} = (\mathbf{I} - \mathbf{K}_k\mathbf{Z}_k\mathbf{H}_k)\boldsymbol{\Gamma}_{k-1}\mathbf{F}_{k-1}\mathbf{P}_k([\mathbf{I} - \mathbf{K}_k\mathbf{Z}_k\mathbf{H}_k]\boldsymbol{\Gamma}_{k-1}\mathbf{F}_{k-1})^\mathsf{T} + \boldsymbol{\Xi}_k, \tag{5.39}$$

where

$$\mathbf{\Xi}_k = \mathbf{P}_{k+1} - (\mathbf{I} - \mathbf{K}_k\mathbf{Z}_k\mathbf{H}_k)\mathbf{\Gamma}_{k-1}\mathbf{F}_{k-1}\mathbf{P}_k([\mathbf{I} - \mathbf{K}_k\mathbf{Z}_k\mathbf{H}_k]\mathbf{\Gamma}_{k-1}\mathbf{F}_{k-1})^\mathsf{T}.$$

With an introduction of a new stochastic matrix $\mathbf{\Lambda}_k$ to absorb the remainder terms the cross-covariance $\mathbf{C}_k$ in (5.33) can be expressed as

$$\mathbf{C}_k = \mathbf{P}_k^- \mathbf{\Lambda}_k(\mathbf{Z}_k\mathbf{H}_k)^\mathsf{T}.$$

and the predicted measurement covariance $\mathbf{S}_k$ in (5.32) as

$$\mathbf{S}_k = \mathbf{Z}_k\mathbf{H}_k\mathbf{\Lambda}_k\mathbf{P}_k^-\mathbf{\Lambda}_k\mathbf{H}_k^\mathsf{T}\mathbf{Z}_k + \mathbf{\Sigma}_k, \tag{5.40}$$

where $\mathbf{\Sigma}_k$ is a matrix analogous to $\mathbf{\Xi}_k$. Using these, the filter gain $\mathbf{K}_k$ from (5.34) has the form

$$\mathbf{K}_k = \mathbf{P}_k^-\mathbf{\Lambda}_k\mathbf{H}_k^\mathsf{T}\mathbf{Z}_k(\mathbf{Z}_k\mathbf{H}_k\mathbf{\Lambda}_k\mathbf{P}_k^-\mathbf{\Lambda}_k\mathbf{H}_k^\mathsf{T}\mathbf{Z}_k + \mathbf{\Sigma}_k)^{-1}. \tag{5.41}$$

To simplify the notation a little, denote $\mathbf{A}_k \coloneqq \mathbf{\Gamma}_k\mathbf{F}_k, \mathbf{B}_k \coloneqq \mathbf{Z}_k\mathbf{H}_k$ and $\mathbf{G}_k \coloneqq \mathbf{Z}_k\mathbf{H}_k\mathbf{\Lambda}_k$. Then some of the equations are transformed into the following less cumbersome form:

$$\begin{aligned}
\boldsymbol{\xi}_k &= (\mathbf{I} - \mathbf{K}_k\mathbf{B}_k)\mathbf{A}_{k-1}\boldsymbol{\xi}_{k-1} + (\mathbf{I} - \mathbf{K}_k\mathbf{B}_k)\mathbf{q}_{k-1} - \mathbf{K}_k\mathbf{r}_k, \\
\mathbf{P}_{k+1} &= (\mathbf{I} - \mathbf{K}_k\mathbf{B}_k)\mathbf{A}_{k-1}\mathbf{P}_k\mathbf{A}_{k-1}^\mathsf{T}(\mathbf{I} - \mathbf{K}_k\mathbf{B}_k)^\mathsf{T} + \mathbf{\Xi}_k, \\
\mathbf{K}_k &= \mathbf{P}_k^-\mathbf{G}_k^\mathsf{T}(\mathbf{G}_k\mathbf{P}_k^-\mathbf{G}_k^\mathsf{T} + \mathbf{\Sigma}_k)^{-1}.
\end{aligned} \tag{5.42}$$

With this machinery introduced, it is possible to prove a result of very qualitative nature. The following theorem is based on the one provided by Li and Xia (2012) but with a clarified and partially corrected proof. The formulation here is partially inspired by Xiong et al. (2009a). Li and Xia (2012) employ recursions for the predicted state estimate and covariance matrix but as the proof does not seem to completely follow by their formulation, the updated estimate and covariance are used here. Their paper is in fact concerned with the version of the UKF allowing for intermittent observations. This aspect is omitted here as it does not really affect the proof.

In fact, the theorem applies to Gaussian filters also as in the above formulation the UKF equations have not been really used. See the remarks after Theorem 5.15.

Theorem 5.14. *Consider the non-linear dynamic system* (5.2) *and the unscented Kalman filter of Algorithm 5.13. Suppose that there are positive* *scalars* $a, b, g_1, g_2, p_1, p_2, q, r, \Sigma_1, \Sigma_2$ *and* Ξ *such that*

$$\begin{aligned} &a^2\mathbf{I} \le \mathbf{A}_k\mathbf{A}_k^\mathsf{T}, && \mathbf{B}_k\mathbf{B}_k^\mathsf{T} \le b\mathbf{I}, && g_1 \le \|\mathbf{G}_k\| \le g_2, && \mathbf{R}_k \le r\mathbf{I}, \\ &p_1\mathbf{I} \le \mathbf{P}_k \le p_2\mathbf{I}, && \Xi\mathbf{I} \le \Xi_k, && \Sigma_1\mathbf{I} \le \mathbf{\Sigma}_k \le \Sigma_2\mathbf{I}, && \mathbf{Q}_k \le q\mathbf{I} \end{aligned}$$

for every $k \ge 0$*. Then the estimation error* $\boldsymbol{\xi}_k$ *is bounded in mean square and stochastically bounded.*

Proof. Lemma 3.8 will be used to prove the claims of boundedness of $\boldsymbol{\xi}_k$. Choose a stochastic Lyapunov function

$$V_k(\boldsymbol{\xi}_k) = \boldsymbol{\xi}_k^\mathsf{T}\mathbf{P}_k^{-1}\boldsymbol{\xi}_k.$$

From the assumption of the boundedness of the error covariance matrix it follows that

$$\frac{1}{p_2}\|\boldsymbol{\xi}_k\|^2 \le V_k(\boldsymbol{\xi}_k) \le \frac{1}{p_1}\|\boldsymbol{\xi}_k\|^2,$$

fulfilling the condition (3.4).

Equation (5.38) and the observation that the cross-terms vanish imply that

$$\begin{aligned} \mathbb{E}[V_k(\boldsymbol{\xi}_k) \mid \boldsymbol{\xi}_{k-1}] = \boldsymbol{\xi}_{k-1}^\mathsf{T}&\mathbf{A}_{k-1}^\mathsf{T}(\mathbf{I}-\mathbf{K}_k\mathbf{B}_k)^\mathsf{T}\mathbf{P}_k^{-1}(\mathbf{I}-\mathbf{K}_k\mathbf{B}_k)\mathbf{A}_{k-1}\boldsymbol{\xi}_{k-1} \\ &+ \mathbb{E}[\mathbf{q}_{k-1}^\mathsf{T}(\mathbf{I}-\mathbf{K}_k\mathbf{B}_k)^\mathsf{T}\mathbf{P}_k^{-1}(\mathbf{I}-\mathbf{K}_k\mathbf{B}_k)\mathbf{q}_{k-1} \\ &\quad + \mathbf{r}_k^\mathsf{T}\mathbf{K}_k^\mathsf{T}\mathbf{P}_k^{-1}\mathbf{K}_k\mathbf{r}_k \mid \boldsymbol{\xi}_{k-1}]. \end{aligned} \tag{5.43}$$

First, evaluate the first term on the right-hand side of this expression. Note that the assumptions imply that $\mathbf{A}_k$ is non-singular. Then, by the same matrix inversion lemma argument as in Jazwinski (1970, p. 234) the matrix $(\mathbf{I}-\mathbf{K}_k\mathbf{B}_k)\mathbf{A}_{k-1}$ is non-singular, validating the form

$$\begin{aligned} \mathbf{P}_k = (\mathbf{I}-\mathbf{K}_k\mathbf{B}_k)\mathbf{A}_{k-1}[\mathbf{P}_{k-1} + ((\mathbf{I}-\mathbf{K}_k\mathbf{B}_k)\mathbf{A}_{k-1})^{-1}\Xi_{k-1} \\ \times ((\mathbf{I}-\mathbf{K}_k\mathbf{B}_k)\mathbf{A}_{k-1})^{-\mathsf{T}}]\mathbf{A}_{k-1}^\mathsf{T}(\mathbf{I}-\mathbf{K}_k\mathbf{B}_k)^\mathsf{T}. \end{aligned} \tag{5.44}$$

The assumptions and the matrix inversion lemma imply an upper bound

$$\begin{aligned} \|\mathbf{K}_k\| &= \|\mathbf{P}_k^-\mathbf{G}_k^\mathsf{T}(\mathbf{G}_k\mathbf{P}_k^-\mathbf{G}_k^\mathsf{T} + \mathbf{\Sigma}_k)^{-1}\| \\ &= \|[(\mathbf{P}_k^-)^{-1} + \mathbf{G}_k^\mathsf{T}\mathbf{\Sigma}_k^{-1}\mathbf{G}_k]^{-1}\mathbf{G}_k^\mathsf{T}\mathbf{\Sigma}_k^{-1}\| \\ &\le \frac{g_2}{\Sigma_1\Sigma_2 g_1^2} := K \end{aligned}$$

for the gain matrix, whereupon

$$[(\mathbf{I} - \mathbf{K}_k\mathbf{B}_k)\mathbf{A}_{k-1}]^{-1}\boldsymbol{\Xi}_{k-1}[(\mathbf{I} - \mathbf{K}_k\mathbf{B}_k)\mathbf{A}_{k-1}]^{-\mathsf{T}} \geq \frac{\Xi}{[a(1+Kb)]^2}\mathbf{I} := \kappa\mathbf{I}.$$

Taking the inverse of (5.44), the first term on the right-hand side of (5.43) becomes

$$\begin{aligned}&\boldsymbol{\xi}_{k-1}^{\mathsf{T}}\mathbf{A}_{k-1}^{\mathsf{T}}(\mathbf{I} - \mathbf{K}_k\mathbf{B}_k)^{\mathsf{T}}\mathbf{P}_k^{-1}(\mathbf{I} - \mathbf{K}_k\mathbf{B}_k)\mathbf{A}_{k-1}\boldsymbol{\xi}_{k-1}\\ &\quad = \boldsymbol{\xi}_{k-1}^{\mathsf{T}}[\mathbf{P}_{k-1} + ((\mathbf{I} - \mathbf{K}_k\mathbf{B}_k)\mathbf{A}_{k-1})^{-1}\boldsymbol{\Xi}_{k-1}((\mathbf{I} - \mathbf{K}_k\mathbf{B}_k)\mathbf{A}_{k-1})^{-\mathsf{T}}]^{-1}\boldsymbol{\xi}_{k-1}.\end{aligned}$$

Hence, by Lemma A.3, the first term has the upper bound

$$\begin{aligned}&\boldsymbol{\xi}_{k-1}^{\mathsf{T}}\mathbf{A}_{k-1}^{\mathsf{T}}(\mathbf{I} - \mathbf{K}_k\mathbf{B}_k)^{\mathsf{T}}\mathbf{P}_k^{-1}(\mathbf{I} - \mathbf{K}_k\mathbf{B}_k)\mathbf{A}_{k-1}\boldsymbol{\xi}_{k-1}\\ &\quad \leq \left(1 + \frac{p_2}{\kappa}\right)^{-1}\boldsymbol{\xi}_{k-1}^{\mathsf{T}}\boldsymbol{\Pi}_{k-1}\boldsymbol{\xi}_{k-1} = (1-\lambda)V_{k-1}(\boldsymbol{\xi}_{k-1}),\end{aligned} \tag{5.45}$$

where $0 < (1 + p_2/\kappa)^{-1} := 1 - \lambda < 1$. Note the similarity of this argument to that in Lemma 5.5.

As the matrices inside the conditional expectation in the second term of (5.43) are just scalars and $\boldsymbol{\xi}_{k-1}$ is independent of $\mathbf{q}_{k-1}$ and $\mathbf{r}_k$, matrix traces can be used to yield the upper bound

$$\begin{aligned}&\mathbb{E}[\mathbf{q}_{k-1}^{\mathsf{T}}(\mathbf{I} - \mathbf{K}_k\mathbf{B}_k)^{\mathsf{T}}\mathbf{P}_k^{-1}(\mathbf{I} - \mathbf{K}_k\mathbf{B}_k)\mathbf{q}_{k-1} + \mathbf{r}_k^{\mathsf{T}}\mathbf{K}_k^{\mathsf{T}}\mathbf{P}_k^{-1}\mathbf{K}_k\mathbf{r}_k \mid \boldsymbol{\xi}_{k-1}]\\ &\quad \leq \mathbb{E}\left(\frac{(1+Kb)^2}{p_1}\operatorname{tr}(\mathbf{q}_{k-1}^{\mathsf{T}}\mathbf{q}_{k-1}) + \frac{K^2}{p_1}\operatorname{tr}(\mathbf{r}_k^{\mathsf{T}}\mathbf{r}_k)\right)\\ &\quad \leq \frac{nq(1+Kb)^2}{p_1} + \frac{mrK^2}{p_1} := \mu.\end{aligned} \tag{5.46}$$

From (5.45) and (5.46) it thus follows that

$$\mathbb{E}[V_{k+1}(\boldsymbol{\xi}_{k+1}) \mid \boldsymbol{\xi}_k] - V_k(\boldsymbol{\xi}_k) \leq \mu - \lambda V_k(\boldsymbol{\xi}_k),$$

meaning that Lemma 3.8 can by applied in order to conclude that $\boldsymbol{\xi}_k$ is bounded in mean square as well as stochastically bounded. □

5.4.4 Stability of Gaussian Filters

In Xiong et al. (2007b) the following general result from Xiong (2006) about the stability of the EKF, the UKF and particle filters for non-linear dynamic systems (5.2) is presented without a proof. A proof is later given in Xiong et al. (2009a), but seems to contain an error. Nonetheless, it is presented here for a comparison.

The estimation error and the corresponding covariance matrices for the aforementioned filters can be written uniformly as

$$\begin{aligned} \boldsymbol{\xi}_{k+1} &= (\mathbf{I} - \mathbf{K}_k \mathbf{B}_k)\mathbf{A}_{k-1}\boldsymbol{\xi}_k + \mathbf{C}_k \mathbf{q}_k - \mathbf{K}_k \mathbf{r}_k, \\ \mathbf{P}_k^- &= \mathbf{A}_{k-1}\mathbf{P}_{k-1}\mathbf{A}_{k-1}^\mathsf{T} + \Xi_{k-1}, \\ \mathbf{P}_k &= (\mathbf{I} - \mathbf{K}_k \mathbf{G}_k)\mathbf{P}_k^-, \end{aligned} \tag{5.47}$$

where $\mathbf{K}_k = \mathbf{P}_k^- \mathbf{G}_k^\mathsf{T} (\mathbf{G}_k \mathbf{P}_k^- \mathbf{G}_k^\mathsf{T} + \boldsymbol{\Sigma}_k)^{-1}$ and $\boldsymbol{\Xi}_k$ and $\boldsymbol{\Sigma}_k$ are random matrices similar to the ones in (5.39) and (5.40), respectively. The matrices $\mathbf{A}_{k-1}, \mathbf{B}_k, \mathbf{C}_k$ and $\mathbf{G}_k$ are random and determined by the system and the particular filtering algorithm under consideration. Comparison with the analogous expressions for in Section 5.4.3 is useful. With this formulation the following theorem has been claimed to hold by Xiong et al. (2009a).

Theorem 5.15. *Consider the non-linear dynamic system* (5.2) *and the EKF, the UKF and particle filter for which the estimation error and error covariance matrices can be formulated as* (5.47)*. Suppose that the following conditions hold:*

(1) There are non-zero scalars a_1, a_2, b, c, g_1 *and* g_2 *such that*

$$\begin{aligned} &a_1^2 \mathbf{I} \le \mathbf{A}_k \mathbf{A}_k^\mathsf{T} \le a_2^2 \mathbf{I}, && \mathbf{B}_k \mathbf{B}_k^\mathsf{T} \le b^2 \mathbf{I}, \\ &g_1^2 \mathbf{I} \le \mathbf{G}_k \mathbf{G}_k^\mathsf{T} \le g_2^2 \mathbf{I}, && \mathbf{C}_k \mathbf{C}_k^\mathsf{T} \le c^2 \mathbf{I}, \\ &(\mathbf{G}_k - \mathbf{B}_k)(\mathbf{G}_k - \mathbf{B}_k)^\mathsf{T} \le (g_2 - b)^2 \mathbf{I} \end{aligned}$$

for every $k \ge 0$*;*

(2) There are positive scalars $p_1, p_2, q, r, \Xi_1, \Xi_2$ *and* Σ *such that*

$$\begin{aligned} &p_1 \mathbf{I} \le \mathbf{P}_k \le p_2 \mathbf{I}, && \mathbf{Q}_k \le q\mathbf{I}, && \mathbf{R}_k \le r\mathbf{I}, \\ &\Xi_1 \mathbf{I} < \boldsymbol{\Xi}_k \le \Xi_2 \mathbf{I}, && \Sigma \mathbf{I} < \boldsymbol{\Sigma}_k \end{aligned}$$

for every $k \ge 0$*, where* Σ *is the maximum of* $a_2^2 (g_2 - b)^2 (p_2 + p_2^2 a_2^2 \Xi_1^{-1})$ *and* $b^2(a_2^2 p_2 + \Xi_2) - g_2^2(a_1^2 p_1 + \Xi_1)$.

Then the estimation error $\boldsymbol{\xi}_k$ *is bounded in mean square and stochastically bounded.*

Proof. (Xiong, 2006; Xiong et al., 2009a, Appendix A) The proof is an application of Lemma 3.8 with a stochastic Lyapunov function $V_k(\boldsymbol{\xi}_k) = \boldsymbol{\xi}_k^\mathsf{T} \mathbf{P}_k^{-1} \boldsymbol{\xi}_k$. □

The proof of above theorem seems to contain a small error as Xiong et al. (2009a) assert that from the positivity of λ_k for $k \geq 0$, it follows that $\inf_{k\geq 0} \lambda_k > 0$ (in fact, they speak incorrectly of minimum). This number, $\inf_{k\geq 0} \lambda_k$, they use as λ in Lemma 3.8 but as its positivity cannot be guaranteed their proof is not valid, at least not in uniform sense.

However, what has been done in Section 5.4.3 does not really depend on the UKF as the Gaussian filtering algorithm used. Consequently, Theorem 5.14 applies to any Gaussian filter whose estimation error and error covariance matrix can be presented in the form (5.42).

5.4.5 Diagonal Matrices and Covariance Matrix Tuning

The discussion here is concentrated on matrices $\mathbf{\Gamma}_k$ and Ξ_k. Similar reasoning can be applied to the other unknown matrices introduced. The magnitudes of the diagonal matrices $\mathbf{\Gamma}_k, \mathbf{Z}_k$ and $\mathbf{\Lambda}_k$ are influenced by the method used to handle the non-linear transformations (for example, in the UKF this method is the unscented transformation) and the nature of the non-linearities. Residuals remain small if these matrices stay near identity matrices and vice versa. Even though it might seem that in Theorem 5.14 it is not explicitly required that the initial estimation error be sufficiently small, the requirement is still there, for with a large initial estimation error the diagonal matrix $\mathbf{\Gamma}_k$ will also be large. Recalling that the matrix Ξ_k was defined as

$$\Xi_k = \mathbf{P}_{k+1} - (\mathbf{I} - \mathbf{K}_k\mathbf{Z}_k\mathbf{H}_k)\mathbf{\Gamma}_{k-1}\mathbf{F}_{k-1}\mathbf{P}_k([\mathbf{I} - \mathbf{K}_k\mathbf{Z}_k\mathbf{H}_k]\mathbf{\Gamma}_{k-1}\mathbf{F}_{k-1})^{\mathsf{T}}.$$

and a lower bound $\mathbf{0} < \Xi\mathbf{I} \leq \Xi_k$ for it assumed, it is seen that with $\mathbf{\Gamma}_k$ sufficiently large there is no guarantee that Ξ_k is always positive-definite, violating the lower bound. But, if $\mathbf{P}_{k+1}$ is sufficiently enlarged, it can be ensured that Ξ_k remains positive-definite even with considerable initial estimation error. To enlarge $\mathbf{P}_{k+1}$, one has to enlarge $\widehat{\mathbf{Q}}_k$ as seen from (5.31) and (5.35). Thus, here the focus is on the effects and design of enlarging $\widehat{\mathbf{Q}}_k = \mathbf{Q}_k + \Delta\mathbf{Q}_k$, where $\Delta\mathbf{Q}_k$ is an additional positive-semidefinite matrix.

The addition of $\Delta\mathbf{Q}_k$ sacrifices some of the accuracy of the filter for its stability. Xiong et al. (2006) evaluate this sacrifice in the case of a non-linear system with linear measurement model and the UKF by using Cramér-Rao lower bound (Tichavský et al., 1998) which in this case is given by the following theorem. The matrix $\widehat{\mathbf{R}}_k$ is taken to be $\mathbf{R}_k$. Note that, as the measurement model is linear, in the two following theorems $\mathbf{H}_k$ are deterministic measurement model matrices, not the Jacobian of the non-linear function $\mathbf{h}$.

Theorem 5.16. *The Cramér–Rao lower bound in the case of the UKF ap-*

plied to the non-linear dynamic system (5.2) *with linear measurement model is*

$$\mathrm{Cov}(\boldsymbol{\xi}_k) \geq \boldsymbol{\mathcal{I}}_k^{-1},$$

where the Fisher information matrix $\boldsymbol{\mathcal{I}}_k$ *can in this particular case expressed recursively as*

$$\boldsymbol{\mathcal{I}}_k = \mathbf{Q}_k^{-1} + \mathbf{H}_k^{\mathsf{T}}\mathbf{R}_k^{-1}\mathbf{H}_k - \mathbf{Q}_k^{-1}\mathbf{F}_k(\boldsymbol{\mathcal{I}}_{k-1} + \mathbf{F}_k^{\mathsf{T}}\mathbf{Q}_k^{-1}\mathbf{F}_k)^{-1}\mathbf{F}_k^{\mathsf{T}}\mathbf{Q}_k^{-1}.$$

The matrix $\mathrm{Cov}(\boldsymbol{\xi}_k)$ *is also the mean squared error (MSE) of the filter.*

Proof. (Xiong et al., 2006, Appendix B) □

With this, some evaluation of lost accuracy is possible, amounting mostly to seeing that MSE may deviate from the Cramér–Rao lower bound.

Theorem 5.17. *Consider the non-linear dynamic system* (5.2) *with linear measurement model and the UKF of Algorithm 5.13. Then the mean squared error of the filter is*

$$\begin{aligned}\mathrm{Cov}(\boldsymbol{\xi}_{k+1}) =& (\widehat{\mathbf{Q}}_k^{-1} + \mathbf{H}_k^{\mathsf{T}}\mathbf{R}_k^{-1}\mathbf{H}_k - \widehat{\mathbf{Q}}_k^{-1}\boldsymbol{\Gamma}_k\mathbf{F}_k \\ & \times (\mathbf{P}_k^{-1} + \mathbf{F}_k^{\mathsf{T}}\boldsymbol{\Gamma}_k\widehat{\mathbf{Q}}_k^{-1}\boldsymbol{\Gamma}_k\mathbf{F}_k)^{-1}\mathbf{F}_k^{\mathsf{T}}\boldsymbol{\Gamma}_k\widehat{\mathbf{Q}}_k^{-1})^{-1} + \Delta\mathbf{P}_k,\end{aligned}$$

where

$$\Delta\mathbf{P}_k = (\mathbf{I} - \mathbf{K}_k\mathbf{H}_k)(\mathrm{Cov}(\boldsymbol{\xi}_k^-) - \mathbf{P}_k^-)(\mathbf{I} - \mathbf{K}_k\mathbf{H}_k)^{\mathsf{T}}.$$

Proof. (Xiong et al., 2006, Appendix C) □

From these theorems it can be seen that the MSE of UKF is near to the Cramér–Rao lower bound if high order error is negligible and the noise covariance matrix is not enlarged, that is $\boldsymbol{\Gamma}_k \approx \mathbf{I}$ and $\widehat{\mathbf{Q}}_k = \mathbf{Q}_k$. If this is not the case, that is to say that the non-linearities are more than minuscule or the process noise covariance matrix is tuned, the difference between Cramér–Rao lower bound and the MSE may be enlarged.

Naturally, under the possibility of filter divergence with the covariance matrix too small and lost accuracy with the matrix overly enlarged, there arises the question of good design. A brief investigation of this matter is carried out by Xiong et al. (2009a) who study the design of $\Delta\mathbf{Q}_k$, letting $\Delta\mathbf{R}_k = 0$. They propose a heuristic method, namely

$$\Delta\mathbf{Q}_k = \gamma\mathbf{P}_k e^{-k} \tag{5.48}$$

for $k \geq 0$, where γ is a positive real number determined case by case (this method has already been demonstrated in Figure 5.4). The idea behind

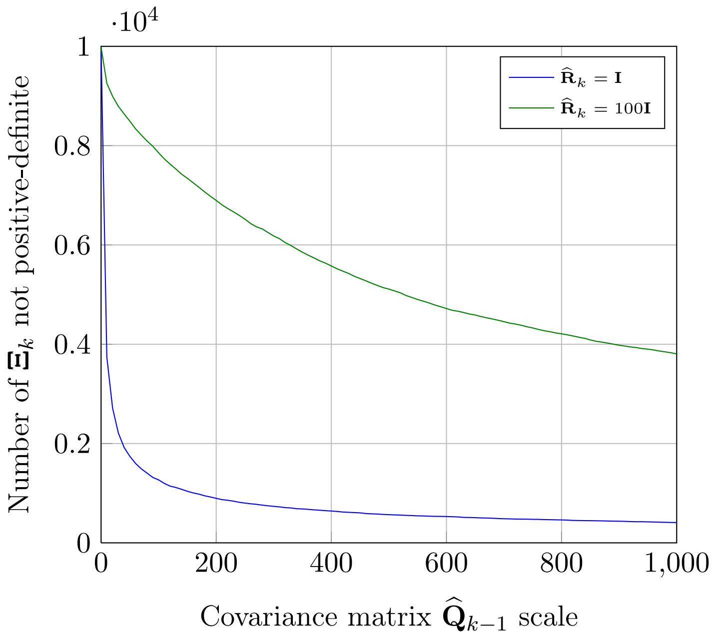

Figure 5.5: Plot of the effect of enlarging $\widehat{\mathbf{Q}}_{k-1} = b\mathbf{I}$ from $\mathbf{0}$ to $1000\mathbf{I}$ to the number of $\boldsymbol{\Xi}_k$ that are not positive-definite. The initial state and estimate are $\mathbf{m}_0 = \mathbf{x}_0 = (0.8, 0.2)$ and $\widehat{\mathbf{R}}_k = \mathbf{R} = \mathbf{I}$ in the first plot and $\widehat{\mathbf{R}}_k = 100\mathbf{I}$ in the second plot.

this method is to use larger additional matrices in the beginning, when the filter estimate may not be very close to the true state due to large initial estimation error, in order guarantee the positive-definiteness of $\boldsymbol{\Xi}_k$ but to get rid of $\Delta\mathbf{Q}_k$ later when the estimates can be trusted to be reasonably close to the true state so as not to lose accuracy unnecessarily. To help in determining γ, Xiong et al. (2009a) further give a certain approximate inequality that is necessary, but not sufficient, for $\boldsymbol{\Xi}_k > 0$. This inequality leads to selecting γ by

$$\gamma = \lambda_{\max}((\mathbf{H}_1\mathbf{P}_0\mathbf{H}_1^\mathsf{T})^{-1}\operatorname{Cov}(\mathbf{v}_1)).$$

The real innovation covariance matrix $\operatorname{Cov}(\mathbf{v}_1)$ can be approximated by averaging over n independent trials,

$$\operatorname{Cov}(\mathbf{v}_1) = \frac{1}{n}\sum_{i=1}^{n}(\mathbf{y}_1^{(i)} - \boldsymbol{\mu}_1^{(i)})(\mathbf{y}_1^{(i)} - \boldsymbol{\mu}_1^{(i)})^\mathsf{T},$$

where the superscript denotes the value obtained in the ith trial.

To test these ideas numerically, consider the system (5.27) for which some EKF simulations have been performed previously. From a number of test, it seems difficult to induce the UKF to diverge for this system. However, this does not mean that the assumptions of Theorem 5.14 are fulfilled. Enlarging $\widehat{\mathbf{Q}}_{k-1}$ decreases the number of Ξ_k that are not positive-definite but such Ξ_k still exist even for $\widehat{\mathbf{Q}}_{k-1} = 1000\mathbf{I}$. This is presented in Figure 5.5, from where it is readily seen that enlarging $\widehat{\mathbf{R}}_k$ hinders the efforts to have all Ξ_k positive-definite.

5.5 Similar Results for Special Cases and Other Filters

There are many special cases of the non-linear system (5.2) and different filters similar to the EKF and or the UKF whose stability has been addressed with stochastic stability lemma or the method of residual-correcting random matrices since the publication of Reif et al. (1999) and Xiong et al. (2006). Most of these results are very similar to the ones presented previously, although there are some exceptions of using somewhat different methods.

Rapp and Nyman (2004) prove stability results for a system with linear time-varying state model and non-linear measurement model. They also consider tuning of the covariance matrix $\widehat{\mathbf{Q}}_k$. In Rapp and Nyman (2005) they are able to relax the assumption of boundedness from above of the norm of the measurement model function Jacobian to only require a finite ratio between its largest and smallest values. However, in doing this they have to assume the noise processes bounded. See also Rapp (2004, Chapter 3). Naturally, as their object of study is the EKF, they introduce some new methods of minor importance.

Özbek et al. (2010) prove the stability result of Reif et al. (1999) assuming that the system is constrained. In such a system on the system state a constraint

$$\mathbf{D}\mathbf{x}_k = \mathbf{d}_k,$$

where $\mathbf{D}$ is some known matrix $\mathbf{d}_k$ and some known vectors, is imposed. Their proof reflects that of Reif et al. (1999) very closely. They only use a slightly different Lyapunov function, this modification resulting from the constrained nature of the system, namely

$$V_k(\boldsymbol{\vartheta}_k) = \boldsymbol{\vartheta}_k^{\mathsf{T}}(\mathrm{Cov}(\boldsymbol{\vartheta}_k))^{-1}\boldsymbol{\vartheta}_k,$$

where $\boldsymbol{\vartheta}_k = (\mathbf{I} - \mathbf{W}^{-1}\mathbf{D}^{\mathsf{T}}(\mathbf{D}\mathbf{W}^{-1}\mathbf{D}^{\mathsf{T}})^{-1}\mathbf{D})\boldsymbol{\xi}_k$. Here, $\mathbf{W}$ is some positive-definite matrix. This work is an extension of their earlier in Köksal Babacan et al. (2008) where analogous result for a deterministic system without any noise is proven. The constrained Kalman filter is treated, for example, in Simon and Chia (2002) and Simon (2010).

Xiong et al. (2008) analyze the stability of the robust extended Kalman filter (Einicke and White, 1999) for a non-linear system with a linear measurement model using arguments and obtaining results similar to those of Reif et al. (1999). In Xiong et al. (2009b) analysis is carried out by the diagonal random matrix method. Analogous results for the adaptive two-stage extended Kalman filter (Kim et al., 2009) are obtained by Kim et al. (2008). In Wang et al. (2013) stability of the EKF for systems with measurement packet losses is studied and results similar to those of Kluge et al. (2010) are obtained. Cao et al. (2014) use the same methods to prove results for the center difference predictive filter.

Xu et al. (2007) prove results analogous to Theorem 5.14 for systems with correlated noises when the measurement model is linear. The result cannot be directly extended for systems with both system and measurement model non-linear (Wang et al., 2012). Dymirkovsky (2012) gives a theorem for UKF stability, similar to Theorem 5.9 for the EKF. Beikzadeh and Taghirab (2009) prove the stability of the state-dependent Riccati equation filter (Mracek et al., 1996).

From a practical point of view the stability is considered in Shang and Liu (2011) where the convergence of the UKF for a certain bleed air system is demonstrated. In Tønne (2007) stability of the EKF utilized to satellite attitude determination is analyzed.

Analogous results have also been obtained in the continuous-time setting of stochastic differential equations, see Reif et al. (2000) for the EKF and Xu et al. (2008) for the UKF, for which the continuous-time version was introduced by Särkkä (2007).

6 Discussion and Future Research

As has been seen in the preceding section, the current stability results for non-linear Kalman filters are very poor. The biggest problem with them is that they can only be used to assess the past filter stability, nothing in them provides ways to determine if the filter will also be stable in the future. Therefore, it would be very beneficial to find conditions for stability that only depend on the dynamic and measurement model functions and the magnitude of noise terms. The stochastic Lyapunov technique seems to be inadequate to provide strong results (Rapp and Nyman, 2004). Consequently, alternative approaches should be investigated. This section records some possible avenues to prove the stability of the Gaussian filter of Algorithm 5.12.

Up to now, only the estimation error $\mathbf{x}_k - \mathbf{m}_k$ has been considered. However, it might be easier to study the difference in the optimal mean square estimate $\mathbb{E}(\mathbf{x}_k \mid \mathbf{y}_{1:k})$ and the Gaussian filter estimate $\mathbf{m}_k$, disregarding the use of stochastic stability lemma altogether. Most of the difficulties still lie in the fact that the error covariance matrix $\mathbf{P}_k$ is a random matrix and depends on the measurements.

6.1 Fourier–Hermite Series Expansion

For non-negative integers k, define the *Hermite polynomials* as

$$H_k(x) = (-1)^k e^{x^2/2} \frac{\mathrm{d}^k}{\mathrm{d}x^k} e^{-x^2/2}.$$

Then, consider the Hilbert space of functions $f, g \colon \mathbb{R} \to \mathbb{R}$ integrable with respect to the standard normal distribution with an inner product

$$\langle f, g \rangle = \int_{\mathbb{R}} f(x) g(x) \mathcal{N}(x \mid 0, 1)\, \mathrm{d}x = \mathbb{E}[f(x) g(x)],$$

where $x \sim \mathcal{N}(0, 1)$. Now, the Hermite polynomials are orthogonal with respect to each other (Kreyszig, 1989, pp. 182-183), that is $\langle H_k, H_l \rangle = 0$ if $n \neq m$. For $k = l$ one has $\langle H_k, H_k \rangle = k!$, giving that $(k!)^{-1/2} H_k$ form an orthonormal basis (Malliavin, 1997, Theorem 2.5). This means that it is possible to express any appropriate f as a *Fourier–Hermite series*

$$\begin{aligned} f(x) &= \sum_{k=0}^{\infty} \langle f, (k!)^{-1/2} H_k \rangle (k!)^{-1/2} H_k = \sum_{k=0}^{\infty} \frac{1}{k!} \mathbb{E}[f(x) H_k(x)] H_k(x) \\ &= \sum_{k=0}^{\infty} \frac{1}{k!} \mathbb{E}[f^{(k)}(x)] H_k(x), \end{aligned} \tag{6.1}$$

where the last equality follows from repeated integration by parts of

$$\mathbb{E}[f(x) H_k(x)] = \frac{(-1)^k}{\sqrt{2\pi}} \int_{\mathbb{R}} f(x) \frac{\mathrm{d}^k}{\mathrm{d}x^k} e^{-x^2/2} \, \mathrm{d}x.$$

The standard normal distribution $\mathcal{N}(0,1)$ of the above formulas can be replaced with an arbitrary normal distribution $\mathcal{N}(\mu, \sigma^2)$, in which case (6.1) becomes

$$\begin{aligned} f(x) &= \sum_{k=0}^{\infty} \frac{1}{k!} \mathbb{E}\left[f(x) H_k\left(\frac{x-\mu}{\sigma}\right)\right] H_k\left(\frac{x-\mu}{\sigma}\right) \\ &= \sum_{k=0}^{\infty} \frac{1}{k!} \mathbb{E}[f^{(k)}(x)] H_k\left(\frac{x-\mu}{\sigma}\right) \sigma^k. \end{aligned} \tag{6.2}$$

Multivariate extensions arise naturally (Sarmavuori and Särkkä, 2012a,b).

This Fourier–Hermite series expansion of functions may provide a necessary and more tractable tool to bound the integrals that need to be computed in the Gaussian filter or to obtain a convenient recursive form than the Taylor series expansion. It is easier to take expectation of this expansion than of Taylor series. Tractability of the Fourier–Hermite series comes partially from Parseval's identity

$$\mathbb{E}(f(x)^2) = \sum_{k=0}^{\infty} \frac{1}{k!} \big(\mathbb{E}[f^{(k)}(x)] \sigma^k\big)^2$$

that follows from the orthogonality of Fourier–Hermite polynomials.

6.2 Telescoping Sum Approach

The following telescoping sum approach has been used in convergence analysis of particle filters (Del Moral and Guionnet, 2001; Rebeschini and van Handel, 2013) and may be of use in proving the stability of Gaussian filters if needed modifications can be made.

Denote by π_k the optimal filter and by $\widehat{\pi}_k$ the Gaussian filter at time-step k. In other words, π_k is the distribution with density $p(\mathbf{x}_k \mid \mathbf{y}_{1:k})$ and $\widehat{\pi}_k$ the distribution with density $\mathcal{N}(\mathbf{m}_k, \mathbf{P}_k)$ that has been obtained by Algorithm 5.12. The corresponding prediction distributions are π_k^- and $\widehat{\pi}_k^-$. The filtering recursion is started from π_0, assumed known here (although in the final reasoning this is of no importance). If necessary, the initial distribution is indicated by a superscript, $\pi_k^{\pi_0}$. Then the objective is to obtain a bound for $d(\pi_k, \widehat{\pi}_k)$ where d is some suitable metric in the space $\mathcal{P}(\mathbb{R}^n)$ of probability measures of $\mathbb{R}^n$ to be determined. Analyzing this difference directly is rather difficult as one would probably need a recursive relationship of some sort. Fortunately, under certain conditions the problem can be reduced to the analysis of difference produced in a single time-step.

Use the following notation for different operators acting on the space $\mathcal{P}(\mathbb{R}^n)$. The *filtering operator* F_k takes π_{k-1} to π_k, that is, $\mathrm{F}_k\,\pi_{k-1} = \pi_k$. This operator is composed of two parts, the *prediction operator* P and the *update operator* U_k that perform the prediction and update steps. The prediction operator maps π_{k-1} to π_k^- and the update operator π_k^- to π_k: $\mathrm{P}\,\pi_{k-1} = \pi_k^-$, $\mathrm{U}_k\,\pi_k^- = \pi_k$. Thus $\mathrm{F}_k\,\pi_{k-1} = \mathrm{U}_k\,\mathrm{P}\,\pi_{k-1}$. The corresponding operators for the Gaussian filter are denoted by a hat. In addition to these, denote by Pr_G the Gaussian projection of probability measures. This projection maps every probability measure to a Gaussian measure with the same mean and covariance. An important remark to be made is that if π_{k-1} is Gaussian, then $\widehat{\mathrm{P}}_k\pi_{k-1} = \mathrm{Pr}_G(\mathrm{P}\,\pi_{k-1})$.

$$\pi_{k-1} \xrightarrow{\text{prediction}} \pi_k^- = \mathrm{P}\,\pi_{k-1} \xrightarrow{\text{update}} \pi_k = \mathrm{U}\,\pi_k^-$$

By employing the above operator notation and using the triangle inequality one can write

$$\begin{aligned} d(\pi_k, \widehat{\pi}_k) &\leq \sum_{s=1}^{k} d\big(\mathrm{F}_k \cdots \mathrm{F}_{s+1}\,\mathrm{F}_s\,\widehat{\mathrm{F}}_{s-1} \cdots \widehat{\mathrm{F}}_1\pi_0, \mathrm{F}_k \cdots \mathrm{F}_{s+1}\,\widehat{\mathrm{F}}_s\widehat{\mathrm{F}}_{s-1} \cdots \widehat{\mathrm{F}}_1\pi_0\big) \\ &= \sum_{s=1}^{k} d\big(\mathrm{F}_k \cdots \mathrm{F}_{s+1}\,\mathrm{F}_s\,\widehat{\pi}_{s-1}, \mathrm{F}_k \cdots \mathrm{F}_{s+1}\,\widehat{\mathrm{F}}_s\widehat{\pi}_{s-1}\big), \end{aligned}$$

Now, any single term on the right-hand side of this inequality can be interpreted as the distance of two optimal filters after $k - s$ steps, one initialized from the distribution $\mathrm{F}_s\,\widehat{\pi}_{s-1}$ and the other from $\widehat{\mathrm{F}}_s\widehat{\pi}_{s-1}$. With this interpretation the inequality can be written as

$$d(\pi_k, \widehat{\pi}_k) \leq \sum_{s=1}^{k} d\left(\pi_{k-s}^{\mathrm{F}_s \widehat{\pi}_{s-1}}, \pi_{k-s}^{\widehat{\mathrm{F}}_s \widehat{\pi}_{s-1}}\right). \tag{6.3}$$

Therefore, if one was able to show that terms on the right-hand side converge to zero uniformly and exponentially as k increases then one would have that $d(\pi_k, \widehat{\pi}_k)$ remains bounded. In precise terms, the condition is that there exist a positive a and $0 < b < 1$ such that for any probability measures μ and ν one has

$$d(\pi_{k-s}^{\mu}, \pi_{k-s}^{\nu}) \leq a b^{k-s} d(\mu - \nu). \tag{6.4}$$

In addition to this, a uniform upper bound on $d(\mathrm{F}_s \widehat{\pi}_{s-1}, \widehat{\mathrm{F}}_s \widehat{\pi}_{s-1})$ would be needed, that is, a positive c independent of s such that $d(\mathrm{F}_s \widehat{\pi}_{s-1}, \widehat{\mathrm{F}}_s \widehat{\pi}_{s-1}) < c$. These conditions would then imply that

$$d(\pi_k, \widehat{\pi}_k) \leq \sum_{s=1}^{k} d\left(\pi_{k-s}^{\mathrm{F}_s \widehat{\pi}_{s-1}}, \pi_{k-s}^{\widehat{\mathrm{F}}_s \widehat{\pi}_{s-1}}\right) \leq ac \sum_{s=0}^{k-1} b^s \leq \frac{ac}{1-b}.$$

The distance $d(\mathrm{F}_s \widehat{\pi}_{s-1}, \widehat{\mathrm{F}}_s \widehat{\pi}_{s-1})$ is the approximation error introduced by the Gaussian filter in one time-step when the preceding distribution is Gaussian.

The metric d should of course chosen to be such that its boundedness implied that of the difference of means. Probably the most well-known such metric is the *Wasserstein metric*, defined for $\mu, \nu \in \mathcal{P}(\mathbb{R}^n)$ by

$$d_W(\mu, \nu)^2 := \inf_{X \sim \mu, Y \sim \nu} \mathbb{E}\left(\|X - Y\|^2\right) \in [0, \infty). \tag{6.5}$$

Another natural possibility would be to use the pseudometric defined as the difference of means. For an introduction and comparison of a number of different probability metrics, see Gibbs and Su (2002).

Showing that $d(\mathrm{F}_s \widehat{\pi}_{s-1}, \widehat{\mathrm{F}}_s \widehat{\pi}_{s-1})$ is uniformly bounded is probably not too difficult under appropriate assumptions. However, the uniform exponential convergence of the terms on the right-hand side of (6.3) is very likely difficult to prove given that the only such existing uniformity results that have been obtained (having total variation norm as the metric) require essentially that the state space be compact (van Handel, 2009). Assuming the state space compact makes the problem of bounding $\mathbb{E}(\mathbf{x}_k \mid \mathbf{y}_{1:k}) - \mathbf{m}_k$ quite trivial.

6.3 Statistical Linearization Framework

The Gaussian filter can be interpreted as the statistically linearized filter (Morelande and García-Fernández, 2013; García-Fernández et al., 2014) in which a non-linear transformation $\mathbf{g}(\mathbf{x})$ of a n-dimensional Gaussian random variable $\mathbf{x} \sim \mathcal{N}(\mathbf{m}, \mathbf{P})$ is approximated by the *enabling approximation*

$$\mathbf{g}(\mathbf{x}) \approx \widetilde{\mathbf{g}}(\mathbf{x}) = \mathbf{A}\mathbf{x} + \mathbf{b} + \mathbf{e},$$

where $\mathbf{e} \sim \mathcal{N}(0, \boldsymbol{\Omega})$ with $\mathbf{A}, \boldsymbol{\Omega}$ matrices and $\mathbf{b}$ a vector given by (compare with the Gaussian filter equations of Algorithm 5.12)

$$\begin{aligned}
\boldsymbol{\mu} &= \int_{\mathbb{R}^n} \mathbf{g}(\mathbf{x}) \mathcal{N}(\mathbf{x} \mid \mathbf{m}, \mathbf{P}) \, \mathrm{d}\mathbf{x}, \\
\mathbf{A} &= \left(\int_{\mathbb{R}^n} [\mathbf{x} - \mathbf{m}][\mathbf{g}(\mathbf{x}) - \boldsymbol{\mu}] \mathcal{N}(\mathbf{x} \mid \mathbf{m}, \mathbf{P}) \, \mathrm{d}\mathbf{x} \right)^{\mathsf{T}} \mathbf{P}^{-1}, \\
\mathbf{b} &= \boldsymbol{\mu} - \mathbf{A}\mathbf{m}, \\
\boldsymbol{\Omega} &= \int_{\mathbb{R}^n} [\mathbf{g}(\mathbf{x}) - \boldsymbol{\mu}][\mathbf{g}(\mathbf{x}) - \boldsymbol{\mu}]^{\mathsf{T}} \mathcal{N}(\mathbf{x} \mid \mathbf{m}, \mathbf{P}) \, \mathrm{d}\mathbf{x} - \mathbf{A}\mathbf{P}\mathbf{A}^{\mathsf{T}}.
\end{aligned}$$

This enables a straightforward presentation of the Gaussian filter as a linear filter and may open new possibilities for analysis. Moreover, the linear minimum mean square estimator property of the Gaussian filter could perhaps be used in a similar linear minimum variance based reasoning that was employed in Lemmas 4.9 and 4.13 where the error covariance matrix of the linear Kalman filter was shown uniformly bounded.

A Appendices

A.1 Useful Matrix Relations

This appendix collects some useful non-trivial matrix identities.

Lemma A.1 (Matrix inversion lemma)**.** *Suppose that* $\mathbf{A}, \mathbf{B}$ *and* $\mathbf{C}$ *are matrices of appropriate dimensions such that inverses in the following exist. Then*

$$(\mathbf{A} + \mathbf{B}\mathbf{C})^{-1} = \mathbf{A}^{-1} - \mathbf{A}^{-1}\mathbf{B}(\mathbf{I} + \mathbf{C}\mathbf{A}^{-1}\mathbf{B})^{-1}\mathbf{C}\mathbf{A}^{-1}.$$

Proof. (Boyd and Vandenberghe, 2009, Appendix C.4.3) □

For the purposes of this thesis the most useful formulations of this identity are (see Anderson and Moore, 1979, pp. 138–139)

$$(\mathbf{I}+\mathbf{A}\mathbf{B}^\mathsf{T}\mathbf{C}^{-1}\mathbf{B})^{-1}\mathbf{A} = (\mathbf{A}^{-1}+\mathbf{B}^\mathsf{T}\mathbf{C}^{-1}\mathbf{B})^{-1} = \mathbf{A}-\mathbf{A}\mathbf{B}^\mathsf{T}(\mathbf{B}\mathbf{A}\mathbf{B}^\mathsf{T}+\mathbf{C})^{-1}\mathbf{B}\mathbf{A}$$

and

$$\begin{aligned}(\mathbf{I} + \mathbf{A}\mathbf{B}^\mathsf{T}\mathbf{C}^{-1}\mathbf{B})^{-1}\mathbf{A}\mathbf{B}^\mathsf{T}\mathbf{C}^{-1} &= (\mathbf{A}^{-1} + \mathbf{B}^\mathsf{T}\mathbf{C}^{-1}\mathbf{B})^{-1}\mathbf{B}^\mathsf{T}\mathbf{C}^{-1} \\ &= \mathbf{A}\mathbf{B}^\mathsf{T}(\mathbf{B}\mathbf{A}\mathbf{B}^\mathsf{T} + \mathbf{C})^{-1}.\end{aligned}$$

Lemma A.2. *Suppose that the matrices* $\mathbf{A}, \mathbf{B}$ *and* $\mathbf{C}$ *are of appropriate dimensions, that* $\mathbf{B}$ *and* $\mathbf{C}$ *are positive-definite and that* $\mathbf{C}-\mathbf{A}\mathbf{B}\mathbf{A}^\mathsf{T}$ *is positive-definite. Then* $\mathbf{B}^{-1} - \mathbf{A}^\mathsf{T}\mathbf{C}^{-1}\mathbf{A}$ *is positive-definite.*

Proof. (Kluge et al., 2010) Since $\mathbf{C} - \mathbf{A}\mathbf{B}\mathbf{A}^\mathsf{T}$ is positive-definite, it is non-singular. Then, by a form of matrix inversion lemma,

$$0 < \mathbf{B} + \mathbf{B}\mathbf{A}^\mathsf{T}(\mathbf{C} - \mathbf{A}\mathbf{B}\mathbf{A}^\mathsf{T})^{-1}\mathbf{A}\mathbf{B} = (\mathbf{B}^{-1} - \mathbf{A}^\mathsf{T}\mathbf{C}^{-1}\mathbf{A})^{-1}.$$

As $(\mathbf{B}^{-1}-\mathbf{A}^\mathsf{T}\mathbf{C}^{-1}\mathbf{A})^{-1}$ is positive-definite, so is its inverse $\mathbf{B}^{-1}-\mathbf{A}^\mathsf{T}\mathbf{C}^{-1}\mathbf{A}$. □

Lemma A.3. *Suppose that the matrices* $\mathbf{A}$ *and* $\mathbf{B}$ *are of appropriate dimensions, positive-definite and satisfy* $a\mathbf{I} \leq \mathbf{A}$ *and* $\mathbf{B} \leq b\mathbf{I}$ *for some positive* a *and* b*. Then* $(\mathbf{A} + \mathbf{B})^{-1} \leq (1 + b/a)^{-1}\mathbf{A}^{-1}$.

Proof. As the matrices are positive-definite, their inverses exist and it is therefore possible to write $(\mathbf{A} + \mathbf{B})^{-1} = (\mathbf{I} + \mathbf{A}^{-1}\mathbf{B})^{-1}\mathbf{A}^{-1}$ so that $\mathbf{I} + \mathbf{A}^{-1}\mathbf{B} \leq (1 + b/a)^{-1}\mathbf{I}$. From this the claim follows. □

A.2 Derivation of the Kalman Filter Equations

In this section a simple derivation of the Kalman filter equations of Theorem 4.1 is presented. The derivation, that follows the one in Särkkä (2013), merely utilizes some basic properties of joint distribution of Gaussian random variables.

Lemma A.4. *The joint distribution of a Gaussian random variable* $\mathbf{x} \sim \mathcal{N}(\mathbf{m}, \mathbf{P})$ *and its linear transformation* $\mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{r}$*, where* $\mathbf{A}$ *is of appropriate dimensions and* $\mathbf{r} \sim \mathcal{N}(\mathbf{0}, \mathbf{R})$*, is*

$$\begin{pmatrix} \mathbf{x} \\ \mathbf{y} \end{pmatrix} \sim \mathcal{N}\left(\begin{pmatrix} \mathbf{m} \\ \mathbf{A}\mathbf{m} \end{pmatrix}, \begin{pmatrix} \mathbf{P} & \mathbf{P}\mathbf{A}^\mathsf{T} \\ \mathbf{A}\mathbf{P} & \mathbf{A}\mathbf{P}\mathbf{A}^\mathsf{T} + \mathbf{R} \end{pmatrix} \right).$$

Proof. The random variable $\mathbf{y}$, being a linear transformation of a Gaussian random variable, is Gaussian (Helstrom, 1984, pp. 211–212). Then the joint distribution of $\mathbf{x}$ and $\mathbf{y}$ is Gaussian, its mean and covariance easily calculatable. □

Lemma A.5. *Let random variables* $\mathbf{x}$ *and* $\mathbf{y}$ *have the joint Gaussian probability distribution*

$$\begin{pmatrix} \mathbf{x} \\ \mathbf{y} \end{pmatrix} \sim \mathcal{N}\left(\begin{pmatrix} \mathbf{a} \\ \mathbf{b} \end{pmatrix}, \begin{pmatrix} \mathbf{A} & \mathbf{C} \\ \mathbf{C}^\mathsf{T} & \mathbf{B} \end{pmatrix} \right).$$

Then the marginal and conditional distributions of $\mathbf{x}$ *and* $\mathbf{y}$ *are given by*

$$\begin{aligned} \mathbf{x} &\sim \mathcal{N}(\mathbf{a}, \mathbf{A}), \\ \mathbf{y} &\sim \mathcal{N}(\mathbf{b}, \mathbf{B}), \\ \mathbf{x} \mid \mathbf{y} &\sim \mathcal{N}\Big(\mathbf{a} + \mathbf{C}\mathbf{B}^{-1}(\mathbf{y} - \mathbf{b}), \mathbf{A} - \mathbf{C}\mathbf{B}^{-1}, \mathbf{C}^\mathsf{T}\Big), \\ \mathbf{y} \mid \mathbf{x} &\sim \mathcal{N}\Big(\mathbf{b} + \mathbf{C}^\mathsf{T}\mathbf{A}^{-1}(\mathbf{x} - \mathbf{a}), \mathbf{B} - \mathbf{C}^\mathsf{T}\mathbf{A}^{-1}\mathbf{C}\Big). \end{aligned}$$

Proof. (Helstrom, 1984, pp. 213–216) □

These two simple lemmas enable one to see that prediction and filtering distributions remain Gaussian and derive the filter equations.

Proof of Theorem 4.1. (Särkkä, 2013, Theorem 4.2) According to the previous lemmas, the joint distribution of $\mathbf{x}_{k-1}$ and $\mathbf{x}_k = \mathbf{A}_{k-1}\mathbf{x}_{k-1} + \mathbf{q}_{k-1}$ given

$\mathbf{y}_{1:k-1}$ is

$$\mathbf{x}_{k-1}, \mathbf{x}_k \mid \mathbf{y}_{1:k-1} \sim \mathcal{N}\left(\begin{pmatrix}\mathbf{m}_{k-1} \\ \mathbf{m}_{k-1}^-\end{pmatrix}, \begin{pmatrix}\mathbf{P}_{k-1} & \mathbf{P}_{k-1}\mathbf{A}_{k-1}^\mathsf{T} \\ \mathbf{A}_{k-1}\mathbf{P}_{k-1} & \mathbf{P}_{k-1}^-\end{pmatrix}\right),$$

where $\mathbf{m}_{k-1}^- = \mathbf{A}_{k-1}\mathbf{m}_{k-1}$ and $\mathbf{P}_{k-1}^- = \mathbf{A}_{k-1}\mathbf{P}_{k-1}\mathbf{A}_{k-1}^\mathsf{T} + \mathbf{Q}_{k-1}$. The marginal distribution of $\mathbf{x}_k$ is then

$$\mathbf{x}_k \mid \mathbf{y}_{1:k-1} \sim \mathcal{N}(\mathbf{A}_{k-1}\mathbf{m}_{k-1}, \mathbf{A}_{k-1}\mathbf{P}_{k-1}\mathbf{A}_{k-1}^\mathsf{T} + \mathbf{Q}_{k-1}).$$

This is the predictive distribution (2.4). Similarly, the joint distribution of $\mathbf{x}_k$ and $\mathbf{y}_k = \mathbf{H}_k\mathbf{x}_k + \mathbf{r}_k$ given $\mathbf{y}_{1:k-1}$ is

$$\mathbf{x}_k, \mathbf{y}_k \mid \mathbf{y}_{1:k-1} \sim \mathcal{N}\left(\begin{pmatrix}\mathbf{m}_k^- \\ \mathbf{H}_k\mathbf{m}_k^-\end{pmatrix}, \begin{pmatrix}\mathbf{P}_k^- & \mathbf{P}_k^-\mathbf{H}_k \\ \mathbf{H}_k\mathbf{P}_k^- & \mathbf{H}_k\mathbf{P}_k^-\mathbf{H}_k^\mathsf{T} + \mathbf{R}_k\end{pmatrix}\right).$$

Therefore, by Lemma A.5, the conditional distribution of $\mathbf{x}_k$ given the observations is

$$\mathbf{x}_k \mid \mathbf{y}_{1:k} = \mathbf{x}_k \mid \mathbf{y}_k, \mathbf{y}_{1:k-1} \sim \mathcal{N}(\mathbf{m}_k, \mathbf{P}_k),$$

where

$$\begin{aligned}\mathbf{m}_k &= \mathbf{m}_k^- + \mathbf{P}_k^-\mathbf{H}_k^\mathsf{T}(\mathbf{H}_k\mathbf{P}_k^-\mathbf{H}_k^\mathsf{T} + \mathbf{R}_k)^{-1}(\mathbf{y}_k - \mathbf{H}_k\mathbf{m}_k^-), \\ \mathbf{P}_k &= \mathbf{P}_k^- - \mathbf{P}_k^-\mathbf{H}_k^\mathsf{T}(\mathbf{H}_k\mathbf{P}_k^-\mathbf{H}_k^\mathsf{T} + \mathbf{R}_k)^{-1}\mathbf{H}_k\mathbf{P}_k^-,\end{aligned}$$

and the update part of the filter equations is obtained by denoting

$$\begin{aligned}\mathbf{v}_k &\coloneqq \mathbf{y}_k - \mathbf{H}_k\mathbf{m}_k^-, \\ \mathbf{S}_k &\coloneqq \mathbf{H}_k\mathbf{P}_k^-\mathbf{H}_k^\mathsf{T} + \mathbf{R}_k, \\ \mathbf{K}_k &\coloneqq \mathbf{P}_k^-\mathbf{H}_k^\mathsf{T}\mathbf{S}_k^{-1}.\end{aligned}$$

□

References

Afar, R. and Zeitouni, O. (1997). Exponential stability for nonlinear filtering. *Annales de l'institut Henri Poincaré (B) Probabilités et Statistiques*, 33(6):697–725.

Agniel, R. G. and Jury, E. I. (1971). Almost sure boundedness of randomly sampled systems. *SIAM Journal on Control*, 9(3):372–384.

Ahlbrandt, C. D. and Heifetz, M. (1995). Discrete Riccati equations of filtering and control. In Elaydi, S. N., Graef, J. R., Ladas, G., and Peterson, A. C., editors, *Proceedings of the First International Conference on Difference Equations, Trinity University, San Antonio, Texas, 25–29 May 1994*, pages 1–16.

Alspach, D. S. and Sorenson, H. W. (1972). Nonlinear Bayesian estimation using Gaussian sum approximations. *IEEE Transactions on Automatic Control*, 17(4):439–448.

Anderson, B. D. O. (1971). Stability properties of Kalman–Bucy filters. *Journal of the Franklin Institute*, 291(2):137–144.

Anderson, B. D. O. (1973). Exponential data weighting in the Kalman–Bucy filter. *Information Sciences*, 5:217–230.

Anderson, B. D. O. (1982). Internal and external stability of linear time-varying systems. *SIAM Journal on Control and Optimization*, 20(3):408–413.

Anderson, B. D. O., Bitmead, R. F., Johnson, C. R., Kokotovic, P. V., Kosut, R. L., Mareels, M. Y., Praly, L., and Riedle, B. D. (1986). *Stability of Adaptive Systems: Passivity and Averaging Analysis.* Number 8 in The MIT Press Series in Signal Processing, Optimization, and Control. MIT Press.

Anderson, B. D. O. and Moore, J. B. (1979). *Optimal Filtering.* Information and System Sciences Series. Prentice-Hall.

Anderson, B. D. O. and Moore, J. B. (1981). Detectability and stabilizability of time-varying discrete-time linear systems. *SIAM Journal on Control and Optimization*, 19(1):20–32.

Anderson, B. D. O. and Moore, J. B. (1992). Comments on "stabilizability and detectability of discrete-time time-varying systems" by J.C. Engwerder. *IEEE Transactions on Automatic Control*, 37(3):409–410.

Aoki, M. (1967). *Optimization of Stochastic Systems.* Mathematics in Science and Engineering. Academic Press.

Bain, A. and Crisan, D. (2009). *Fundamentals of Stochastic Filtering.* Number 60 in Stochastic Modelling and Applied Probability. Springer.

Baras, J. S., Bensoussan, A., and James, M. R. (1988). Dynamic observers as asymptotic limits of recursive filters: Special cases. *SIAM Journal on Applied Mathematics*, 48(5):1147–1158.

Beikzadeh, H. and Taghirab, H. D. (2009). Stability analysis of the discrete-time difference SDRE state estimator in a noisy environment. In *Proceedings of 2009 IEEE International Conference on Control and Automation Christchurch, New Zealand, December 9–11, 2009*, pages 1751–1756.

Beneš, V. E. and Karatzas, I. (1983). Estimation and control for linear, partially observable systems with non-gaussian initial distribution. *Stochastic Processes and their Applications*, 14(3):233–248.

Bittanti, S., Bolzern, P., and De Nicolao, G. (1992). Comments on "stabilizability and detectability of discrete-time time-varying systems" by J.C. Engwerder. *IEEE Transactions on Automatic Control*, 37(8):1274–1275.

Bolognani, S., Tubiana, L., and Zigliotto, M. (2003). Extended Kalman filter tuning in sensorless PMSM drives. *IEEE Transactions on Industry Applications*, 39(6):1741–1747.

Boutayeb, M. and Aubry, D. (1999). A strong tracking extended Kalman observer for nonlinear discrete-time systems. *IEEE Transactions on Automatic Control*, 44(8):1550–1556.

Boutayeb, M., Rafaralahy, H., and Darouach, M. (1997). Convergence analysis of the extended Kalman filter used as an observer for nonlinear deterministic discrete-time systems. *IEEE Transactions on Automatic Control*, 42(4):581–586.

Boyd, S. P. and Vandenberghe, L. (2009). *Convex Optimization*. Cambridge University Press, 7th edition.

Bucy, R. S. (1965). Stability and positive supermartingales. *Journal of Differential Equations*, 1(2):151–155.

Bucy, R. S. and Joseph, P. D. (1968). *Filtering for Stochastic Processes with Applications to Guidance*. Number 23 in Interscience Tracts in Pure and Applied Mathematics. Interscience Publishers.

Cao, L., Zhang, Q., Chen, X., and Zhao, Y. (2014). Stochastic stability of center difference predictive filter. *Journal of the Franklin Institute*, 351(11):5204–5233.

Chen, Z. (2003). Bayesian filtering: From Kalman filters to particle filters, and beyond.

Chigansky, P. (2006). Stability of nonlinear filter: A survey. Lecture notes, Petropolis, Brazil.

Costa, E. F. and Astolfi, A. (2008). On the stability of the recursive Kalman filter for linear time-invariant systems. In *Proceedings of the 2008 American Control Conference, Seattle, Washington, USA, 11–13 June 2008*, pages 1286–1291.

De Nicolao, G. (1992). On the time-varying Riccati difference equation of optimal filtering. *SIAM Journal on Control and Optimization*, 30(6):1251–1269.

Del Moral, P. and Guionnet, A. (2001). On the stability of interacting processes with applications to filtering and genetic algorithms. *Annales de l'institut Henri Poincaré (B) Probabilités et Statistiques*, 37(2):155–194.

Delyon, B. (2001). A note on uniform observability. *IEEE Transactions on Automatic Control*, 46(8):1326–1327.

Deyst, Jr., J. J. (1973). Correction to "conditions for asymptotic stability of the discrete minimum-variance linear estimator". *IEEE Transactions on Automatic Control*, 18(5):562–563.

Deyst, Jr., J. J. and Price, C. F. (1968). Conditions for asymptotic stability of the discrete minimum-variance linear estimator. *IEEE Transactions on Automatic Control*, 13(6):702–705.

Dragan, V., Morozan, T., and Stoica, A.-M. (2006). *Mathematical Methods in Robust Control of Linear Stochastic Systems*. Number 50 in Mathematical Concepts and Methods in Science and Engineering. Springer.

Dymirkovsky, G. (2012). New results on stochastic stability of discrete-time unscented Kalman filter. In *7th IEEE Conference on Industrial Electronics and Applications, Singapore, 18–20 July 2012*, pages 1543–1548.

Einicke, G. A. and White, L. B. (1999). Robust extended Kalman filtering. *IEEE Transactions on Signal Processing*, 47(9):2596–2599.

Engwerda, J. (1990). Stabilizability and detectability of discrete-time time-varying systems. *IEEE Transactions on Automatic Control*, 35(4):425–429.

Fitzgerald, R. J. (1971). Divergence of the Kalman filter. *IEEE Transactions on Automatic Control*, 16(6):738–747.

Ford, J. J. (2002). Non-linear and robust filtering: From the Kalman filter to the particle filter. Technical report, DSTO Aeronautical and Maritime Research Laboratory, 506 Lorimer St, Fishermans Bend, Victoria, Australia 3207.

Ford, J. J. and Coulter, A. S. (2001). Filtering for precision guidance: The extended Kalman filter. Technical report, DSTO Aeronautical and Maritime Research Laboratory, 506 Lorimer St, Fishermans Bend, Victoria, Australia 3207.

Gajić, Z. and Qureshi, M. T. J. (2008). *Lyapunov Matrix Equation in System Stability and Control*. Dover Civil and Mechanical Engineering. Dover.

García-Fernández, A. F., Svensson, L., and Morelande, M. R. (2014). Iterated statistical linear regression for bayesian updates. In *17th International Conference on Information Fusion (FUSION)*.

Gard, T. C. (1988). *Introduction to Stochastic Differential Equations*. Pure and Applied Mathematics. Marcel Dekker.

Gibbs, A. L. and Su, F. E. (2002). On choosing and bounding probability metrics. *International Statistical Review*, 70(3):419–435.

Goodwin, G. C. and Sin, K. S. (1984). *Adaptive Filtering, Prediction and Control*. Information and System Sciences Series. Prentice-Hall.

Hager, W. W. and Horowitz, L. L. (1976). Convergence and stability properties of the discrete Riccati operator equation and the associated optimal control and filtering problems. *SIAM Journal on Control and Optimization*, 14(2):295–312.

Hahn, W. (1963). *Theory and Application of Liapunov's Direct Method*. Prentice-Hall.

Halevi, Y. (1989). The optimal reduced-order estimator for systems with singular measurement noise. *IEEE Transactions on Automatic Control*, 34(7):777–781.

Hauser, H. (2004). *Lehrbuch der Analysis*, volume 2 of *Mathematische Leitfäden*. B.G. Teubner, 13th edition.

Heine, K. (2007). *On the Stability of the Discrete Time Filter and the Uniform Convergence of Its Approximations*. PhD thesis, Tampere University of Technology.

Helstrom, C. W. (1984). *Probability and Stochastic Processes for Engineers*. Macmillan.

Hitz, K. L. (1970). *Relations Between Continuous-Time and Discrete-Time Quadratic Minimization*. PhD thesis, The University of Newcastle.

Hitz, K. L., Fortmann, T., and Anderson, B. D. O. (1972). A note on bounds on solutions of the Riccati equation. *IEEE Transactions on Automatic Control*, 17(1):178.

Ito, K. and Xiong, K. (2000). Gaussian filters for nonlinear filtering problems. *IEEE Transactions on Automatic Control*, 45(5):910–927.

Jazwinski, A. H. (1970). *Stochastic Processes and Filtering Theory*. Academic Press.

Jouffroy, J. and Fossen, T. I. (2010). A tutorial on incremental stability analysis using contraction theory. *Modeling, Identification and Control*, 31(3):93–106.

Julier, S. J. and Uhlmann, J. K. (2004). Unscented filtering and nonlinear estimation. In *Proceedings of the IEEE*, volume 92, pages 401–422.

Julier, S. J., Uhlmann, J. K., and Durrant-Whyte, H. F. (1995). A new approach for filtering nonlinear systems. In *Proceedings of the 1995 American Control Conference, Seattle, Washington, 21–23 June 1995*, pages 1628–1632.

Julier, S. J., Uhlmann, J. K., and Durrant-Whyte, H. F. (2000). A new method for the nonlinear transformation of means and covariances in filters and estimators. *IEEE Transactions on Automatic Control*, 45(3):477–482.

Kalman, R. E. (1960a). Contributions to the theory of optimal control. *Boletín de la Sociedad Matemática Mexicana*, 5(1):102–119.

Kalman, R. E. (1960b). A new approach to linear filtering and prediction problems. *Transactions of the ASME, Journal of Basic Engineering*, 82(1):35–45.

Kalman, R. E. (1961). On the general theory of control systems. In Coales, J. F., Ragazzini, J. R., and Thomas, F. A., editors, *Automatic and Remote Control: Proceedings of the First International Congress of the International Federation of Automatic Control (I.F.A.C.): Moscow, 1960*, volume 1, pages 481–491. Butterworth Scientific Publications.

Kalman, R. E. (1963). New methods in Wiener filtering theory. In Bodganoff, J. L. and Kozin, F., editors, *Proceedings of the First Symposium of Engineering Applications of Random Function Theory and Probability*, pages 270–388. John Wiley and Sons.

Kalman, R. E. and Bertram, J. E. (1960a). Control system analysis and design via the "second method" of Lyapunov, I continuous-time systems. *Transactions of the ASME, Journal of Basic Engineering*, 82(2):371–393.

Kalman, R. E. and Bertram, J. E. (1960b). Control system analysis and design via the "second method" of Lyapunov, II discrete-time systems. *Transactions of the ASME, Journal of Basic Engineering*, 82(2):394–400.

Kalman, R. E. and Bucy, R. S. (1961). New results in linear filtering and prediction theory. *Transactions of the ASME, Journal of Basic Engineering*, 83(1):95–108.

Kalman, R. E., Ho, Y.-C., and Narendra, K. S. (1963). Controllability of linear dynamical systems. In Lasalle, J. P. and Diaz, J. B., editors, *Contributions to Differential Equations*, volume 1, pages 189–213. Interscience Publishers.

Kamen, E. W. and Su, J. K. (1999). *Introduction to Optimal Estimation*. Advanced Textbooks in Control and Signal Processing. Springer.

Kerr, T. H. (1985). The proper computation of the matrix pseudoinverse and its impact in MVRO filtering. *IEEE Transactions on Aerospace and Electronic Systems*, 21(5):711–724.

Khasminskii, R. (2011). *Stochastic Stability of Differential Equations*. Number 66 in Stochastic Modelling and Applied Probability. Springer, 2nd edition.

Kim, K.-H., Jee, G.-I., and Song, J.-H. (2008). The stability of the adaptive two-stage extended Kalman filter. In *International Conference on Control, Automation and Systems, Seoul, Korea, 14–17 October 2008*, pages 1378–1383.

Kim, K. H., Lee, J. G., and Park, C. G. (2006). Adaptive two-stage Kalman filter in the presence of unknown random bias. *International Journal of Adaptive Control and Signal Processing*, 20(7):305–319.

Kim, K. H., Lee, J. G., and Park, C. G. (2009). Adaptive two-stage extended Kalman filter for a fault-tolerant INS-GPS loosely coupled system. *IEEE Transactions on Aerospace and Electronic Systems*, 45(1):125–137.

Kluge, S., Reif, K., and Brokate, M. (2010). Stochastic stability of the extended Kalman filter with intermittent observations. *IEEE Transactions on Automatic Control*, 55(2):514–518.

Krener, A. J. (2003). The convergence of the extended Kalman filter. In Rantzer, A. and Byrnes, C. I., editors, *Directions in Mathematical Systems Theory and Optimization*, volume 286 of *Lecture Notes in Control and Information Sciences*, chapter 12, pages 173–182. Springer-Verlag Berlin Heidelberg.

Kreyszig, E. (1989). *Introductory Functional Analysis with Applications*. Wiley Classics Library. John Wiley & Sons.

Kushner, H. J. (1967). *Stochastic Stability and Control*. Mathematics in Science and Engineering. Academic Press.

Kwakernaak, H. and Sivan, R. (1972). *Linear Optimal Control Systems*. Wiley-Interscience.

Köksal Babacan, E., Özbek, L., and Efe, M. (2008). Stability of the extended Kalman filter when the states are constrained. *IEEE Transactions on Automatic Control*, 53(11):2707–2711.

La Scala, B. F., Bitmead, R. F., and James, M. R. (1995). Conditions for stability of the extended Kalman filter and their application to the frequency tracking problem. *Mathematics of Control, Signals and Systems*, 8(1):1–26.

Lakshmikantham, V. and Trigiante, D. (2002). *Theory Of Difference Equations: Numerical Methods And Applications*. Marcel Dekker, second edition.

Lancaster, P. and Rodman, L. (1995). *Algebraic Riccati Equations*. Oxford Science Publications. Oxford University Press.

Li, L. and Xia, Y. (2012). Stochastic stability of the unscented Kalman filter with intermittent observations. *Automatica*, 48(5):978–981.

Ljung, L. (1979). Asymptotic behavior of the extended Kalman filter as a parameter estimator for linear systems. *IEEE Transactions on Automatic Control*, 24(1):36–50.

Loève, M. (1978). *Probability Theory*, volume 2. Springer.

Lyapunov, A. M. (1907). Problème général de la stabilité du mouvement. *Annales de la faculté des sciences de Toulouse*, 9:203–474.

Malliavin, P. (1997). *Stochastic Analysis*. Number 313 in Grundlehren der mathematischen Wissenschaften. Springer.

Maree, J. P., Imsland, L., and Jouffroy, J. (2014). On convergence of the unscented Kalman–Bucy filter using contraction theory. *International Journal of Systems Science*.

Maybeck, P. S. (1979). *Stochastic Models, Estimation and Control*, volume 1. Academic Press.

Michel, A. N., Hou, L., and Liu, D. (2008). *Stability of Dynamical Systems: Continuous, Discontinuous, and Discrete Systems*. Systems & Control: Foundations & Applications. Birkhäuser.

Moore, J. B. and Anderson, B. D. O. (1980). Coping with singular transition matrices in estimation and control stability theory. *International Journal of Control*, 31(3):571–586.

Morelande, M. R. and García-Fernández, A. F. (2013). Analysis of Kalman filter approximations for nonlinear measurements. *IEEE Transactions on Signal Processing*, 61(22):5477–5484.

Mracek, C. P., Cloutier, J. R., and D'Souza, C. A. (1996). A new technique for nonlinear estimation. In *Proceedings of the 1996 IEEE International Conference on Control Applications, Dearborn, MI, USA, 15 – 18 September 1996*, pages 338–343.

Murty, K. G. (2014). *Computational and Algorithmic Linear Algebra and n-Dimensional Geometry*. Institute of Mathematical Statistics Textbooks. World Scientific Publishing Company.

Nijmeijer, H. (1982). Observability of autonomous discrete time non-linear systems: a geometric approach. *International Journal of Control*, 36(5):867–874.

Nikolsky, S. M. (1977). *A Course of Mathematical Analysis*, volume 1. Mir Publishers.

Nørgaard, M., Poulsen, N. K., and Ravn, O. (2000). New developments in state estimation for nonlinear systems. *Automatica*, 36(11):1627–1638.

Ocone, D. and Pardoux, E. (1996). Asymptotic stability of the optimal filter with respect to its initial condition. *SIAM Journal on Control and Optimization*, 34(1):226–243.

Rapp, K. (2004). *Nonlinear Estimation and Control in the Iron Ore Pelletizing Process: An Application and Analysis of the Extended Kalman Filter*. PhD thesis, Norwegian University of Science and Technology.

Rapp, K. and Nyman, P.-O. (2004). On stability of the extended Kalman filter. In Kaszyński, R. and Domek, S., editors, *Proceedings of the 10th IEEE International Conference on Methods and Models in Automation and Robotics, Międzyzdroje, Poland, 30 August – 2 September 2004*. Wydawnictwo Uczelniane Politechniki Szczecińskiej. Extended version included in Rapp's thesis (2004).

Rapp, K. and Nyman, P.-O. (2005). Stability properties of the discrete-time extended kalman filter. In Allgöwer, F. and Zeitz, M., editors, *Nonlinear Control Systems: A proceedings volume from 6th IFAC Symposium, Stuttgart, Germany, 1–3 September 2004*, volume 3, pages 1377–1382. Elsevier. Extended version included in Rapp's thesis (2004).

Rebeschini, P. and van Handel, R. (2013). Can local particle filters beat the curse of dimensionality? *Annals of Applied Probability*. To appear.

Reif, K., Günther, S., Yaz, E., and Unbehauen, R. (1998). Stabilität des zeitkontinuierlichen erweiterten Kaiman–Filters. *at - Automatisierungstechnik*, 46(12):592–601.

Reif, K., Günther, S., Yaz, E., and Unbehauen, R. (1999). Stochastic stability of the discrete-time extended Kalman filter. *IEEE Transactions on Automatic Control*, 44(4):714–728.

Reif, K., Günther, S., Yaz, E., and Unbehauen, R. (2000). Stochastic stability of the continuous-time extended Kalman filter. *IEE Proceedings - Control Theory and Applications*, 147(1):45–52.

Reif, K. and Unbehauen, R. (1999). The extended Kalman filter as an exponential observer for nonlinear systems. *IEEE Transactions on Signal Processing*, 47(8):2324–2328.

Rhudy, M., Gu, Y., and Napolitano, M. M. (2012). Relaxation of initial error and noise bounds for stability of GPS/INS attitude estimation. In *AIAA Guidance, Navigation, and Control Conference, Minneapolis, Minnesota, USA, 13–16 August 2012*.

Rhudy, M. B. and Gu, Y. (2013). Online stochastic convergence analysis of the Kalman filter. *International Journal of Stochastic Analysis*, 2013.

Ristic, B., Farina, A., Benvenuti, D., and Arulampalam, S. (2003). Performance bounds and comparison of nonlinear filters for tracking a ballistic object on re-entry. *IEE Proceedings - Radar, Sonar and Navigation*, 150(2):65–70.

Safonov, M. G. and Athans, M. (1978). Robustness and computational aspects of nonlinear stochastic estimators and regulators. *IEEE Transactions on Automatic Control*, 23(4):717–725.

Sarmavuori, J. and Särkkä, S. (2012a). Fourier–Hermite Kalman filter. *IEEE Transactions on Automatic Control*, 57(6):1511–1515.

Sarmavuori, J. and Särkkä, S. (2012b). Fourier–Hermite Rauch–Tung–Striebel smoother. In *Proceedings of EUSIPCO*, pages 2109–2113.

Schlee, F. H., Standish, C. J., and Toda, N. F. (1967). Divergence in the Kalman filter. *AIAA Journal*, 5(6):1114–1120.

Shang, L. and Liu, G. (2011). Sensor and actuator fault detection and isolation for a high performance aircraft engine bleed air temperature control system. *IEEE Transactions on Control Systems Technology*, 19(5):1260–1268.

Simon, D. (2006). *Optimal State Estimation: Kalman, H_∞, and Nonlinear Approaches*. Wiley-Interscience.

Simon, D. (2010). Kalman filtering with state constraints: a survey of linear and nonlinear algorithms. *IET Control Theory & Applications*, 4(8):1303–1318.

Simon, D. and Chia, T. L. (2002). Kalman filtering with state equality constraints. *IEEE Transactions on Aerospace and Electronic Systems*, 38(1):128–136.

Sinopoli, B., Schenato, L., Franceschetti, M., Poolla, K., Jordan, M. I., and Sastry, S. S. (2004). Kalman filtering with intermittent observations. *IEEE Transactions on Automatic Control*, 49(9):1453–1464.

Song, T. L. and Speyer, J. L. (1985). A stochastic analysis of a modified gain extended Kalman filter with applications to estimation with bearings only measurements. *IEEE Transactions on Automatic Control*, 30(10):940–949.

Song, Y. and Grizzle, J. W. (1992). The extended Kalman filter as a local asymptotic observer for nonlinear discrete-time systems. In *Proceedings of the 1992 American Control Conference: The Westin Hotel, Chicago, Illinois, 24–26 June 1992*, pages 3365–3369.

Song, Y. and Grizzle, J. W. (1995). The extended Kalman filter as a local asymptotic observer for discrete-time nonlinear systems. *Journal of Mathematical Systems, Estimation, and Control*, 5(1):59–78.

Sorenson, H. W. (1967). On the error behavior in linear minimum variance estimation problems. *IEEE Transactions on Automatic Control*, 12(5):557–562.

Sorenson, H. W. (1970). Least-squares estimation: from Gauss to Kalman. *IEEE Spectrum*, 7(7):63–68.

Sowers, R. B. and Makowski, A. M. (1992). Discrete-time filtering for linear systems with non-Gaussian initial conditions: asymptotic behavior of the difference between the MMSE and LMSE estimates. *IEEE Transactions on Automatic Control*, 37(1):114–120.

Särkkä, S. (2007). On unscented Kalman filtering for state estimation of continuous-time nonlinear systems. *IEEE Transactions on Automatic Control*, 52(9):1631–1641.

Särkkä, S. (2013). *Bayesian Filtering and Smoothing*. Institute of Mathematical Statistics Textbooks. Cambridge University Press.

Tarn, T.-J. and Rasis, Y. (1976). Observers for nonlinear stochastic systems. *IEEE Transactions on Automatic Control*, 21(34):441–448.

Tichavský, P., Muravchik, C. H., and Nehorai, A. (1998). Posterior Cramér–Rao bounds for discrete-time nonlinear filtering. *IEEE Transactions on Signal Processing*, 46(5):1386–1396.

Tse, E. (1973). Observer-estimators for discrete-time systems. *IEEE Transactions on Automatic Control*, 18(1):10–16.

Tønne, K. K. (2007). Stability analysis of EKF-based attitude determination and control. Master's thesis, Norwegian University of Science and Technology.

Unbehauen, R. (1998). *Systemtheorie 2: Mehrdimensionale, adaptive und nichtlineare Systeme*. Oldenbourg Wissenschaftsverlag.

Ursin, B. (1980). Asymptotic convergence properties of the extended Kalman filter using filtered state estimates. *IEEE Transactions on Automatic Control*, 25(6):1207–1211.

van Handel, R. (2007). *Filtering, Stability and Robustness*. PhD thesis, California Institute of Technology.

van Handel, R. (2009). Uniform time average consistency of Monte Carlo particle filters. *Stochastic Processes and their Applications*, 119:3835–3861.

Wan, E. A. and van der Merwe, R. (2001). The unscented Kalman filtering. In Haykin, S., editor, *Kalman Filterin and Neural Networks*, chapter 7, pages 221–280. John Wiley & Sons.

Wang, G., Chen, J., and Sun, J. (2013). Stochastic stability of extended filtering for nonlinear systems with measurement packet losses. *IET Control Theory & Applications*, 8(17):2048–2055.

Wang, X., Liang, Y., Pan, Q., and Yang, F. (2012). A Gaussian approximation recursive filter for nonlinear systems with correlated noises. *Automatica*, 48(9):2290–2297.

Willems, J. L. (1970). *Stability Theory of Dynamical Systems*. Studies in Dynamical Systems. Thomas Nelson and Sons Ltd.

Williams, D. (1991). *Probability with Martingales*. Cambridge Mathematical Textbooks. Cambridge University Press.

Wonham, W. M. (1967). On pole assignment in multi-input controllable linear systems. *IEEE Transactions on Automatic Control*, 12(6):660–665.

Wonham, W. M. (1968). On a matrix Riccati equation of stochastic control. *SIAM Journal on Control*, 6(4):681–697.

Wu, Y., Hu, D., and Hu, X. (2007). Comments on "Performance evaluation of UKF-based nonlinear filtering". *Automatica*, 43(3):567–568.

Wu, Y., Hu, D., Wu, M., and Hu, X. (2006). A numerical-integration perspective on Gaussian filters. *IEEE Transactions on Signal Processing*, 54(8):2910–2921.

Xiong, K. (2006). *Nonlinear Filter and Its Application in Satellite Attitude*. PhD thesis, Beihang University.

Xiong, K., Chan, C. W., and Zhang, H. (2007a). Detection of satellite attitude sensor faults using the UKF. *IEEE Transactions on Aerospace and Electronic Systems*, 43(2):480–491.

Xiong, K., Liu, L. D., and Zhang, H. (2009a). Modified unscented Kalman filtering and its application in autonomous satellite navigation. *Aerospace Science and Technology*, 13(4–5):238–246.

Xiong, K., Zhang, H., and Chan, C. W. (2006). Performance evaluation of UKF-based nonlinear filtering. *Automatica*, 42(2):261–270.

Xiong, K., Zhang, H., and Chan, C. W. (2007b). Author's reply to "Comments on 'Performance evaluation of UKF-based nonlinear filtering"'. *Automatica*, 43(3):569–570.

Xiong, K., Zhang, H., and Liu, L. (2008). Adaptive robust extended Kalman filter for nonlinear stochastic systems. *IET Control Theory & Applications*, 2(3):239–250.

Xiong, K., Zhang, H., and Liu, L. (2009b). Adaptive robust extended Kalman filter. In Moreno, V. M. and Pigazo, A., editors, *Kalman Filter: Recent Advances and Applications*, pages 121–142. InTech.

Xu, J., Dimirovski, G. M., Jing, Y., and Shen, C. (2007). UKF design and stability for nonlinear stochastic systems with correlated noises. In *Proceedings of the 46th IEEE Conference on Decision and Control, New Orleans, Louisiana, USA, 12–14 December 2007*, pages 6226–6231.

Xu, J., Wang, S., Dimirovski, G. M., and Jing, Y. (2008). Stochastic stability of the continuous-time unscented Kalman filter. In *Proceedings of the 47th IEEE Conference on Decision and Control, Cancún, Mexico, 9–11 December 2008*, pages 5110–5115.

Özbek, L., Köksal Babacan, E., and Efe, M. (2010). Stochastic stability of the discrete-time constrained extended Kalman filter. *Turkish Journal of Electrical Engineering & Computer Sciences*, 18(2):211–224.