universität wien

MASTERARBEIT / MASTER'S THESIS

Titel der Masterarbeit / Title of the Master's Thesis

„Stochastic Filtering and Parameter Estimation in Infinite Dimensions“

verfasst von / submitted by

Fabian Kai Parzer, B.Sc.

angestrebter akademischer Grad / in partial fulfilment of the requirements for the degree of

Master of Science (MSc)

Wien, 2020 / Vienna 2020

Studienkennzahl lt. Studienblatt / degree programme code as it appears on the student record sheet:	UA 066 821
Studienrichtung lt. Studienblatt / degree programme as it appears on the student record sheet:	Masterstudium Mathematik
Betreut von / Supervisor:	Univ.-Prof. Dipl.-Ing. Dr. Otmar Scherzer

Abstract

In this thesis, we present aspects of the theory of infinite-dimensional stochastic filtering, with an additional focus on parameter estimation. After giving an overview of the necessary background on probability distributions and stochastic calculus in Banach and Hilbert spaces, we derive the infinite-dimensional versions of the Kalman filter, the Kalman-Bucy filter, and the Kalman smoother.

Furthermore, we present approximate filters for the nonlinear filtering problem, in particular the (iterated) extended Kalman filter and the ensemble Kalman filter/smoother. We also discuss their applications to real-time parameter estimation in dynamical systems.

Moreover, we discuss recent applications of filtering methods to the solution of inverse problems, in form of the ensemble Kalman inversion method. We also show that the reduced approach for parameter estimation in partial differential equations using a regularized Gauss-Newton method can be interpreted as a dual iterated extended Kalman filter.

Finally, we present two possible applications of the methods presented in this thesis. First, we show how a dual ensemble Kalman filter can be applied to a stochastic potential identification problem. Second, we estimate the initial conditon in a stochastic wave equation using the ensemble Kalman smoother.

Zusammenfassung

Diese Masterarbeit stellt Auszüge aus der Theorie der stochastischen Filter in unendlichdimensionalen Räumen dar, mit einem zusätzlichen Fokus auf Parameterschätzung. Nach einer Zusammenfassung der notwendigen Vorkenntnisse zu Wahrscheinlichkeitsverteilungen und stochastischer Analysis in Banach- und Hilberträumen leiten wir die unendlichdimensionalen Versionen des Kalman-Filters, des Kalman-Bucy Filters und des Kalman-Glättungsfilters her.

Außerdem präsentieren wir näherungsweise Filter, insbesondere das (iterierte) Extended Kalman Filter sowie das Enbsemble Kalman Filter/Glättungsfilter. Weiters erörtern wir die Anwendung dieser Methoden für die Echtzeit-Parameterschätzung in dynamischen Systemen.

Zusätzlich erläutern wir neuere Anwendungen von Filtermethoden zur Lösung von Inversen Problemen, die Ensemble Kalman Inversion. Darüber hinaus zeigen wir, dass man die reduzierte Methode zur Parameterschätzung in Partiellen Differentialgleichungen mittels Gauß-Newton-Verfahren als duales iteriertes Extended Kalman Filter auffassen kann.

Schlussendlich stellen wir zwei mögliche Anwendungen der vorgestellten Verfahren dar. Als Erstes zeigen wir, wie das duale Ensemble Kalman Filter auf ein stochastisches Potentialidentifikationsproblem angewendet werden kann. Zweitens schätzen wir den Anfangswert in einer stochastischen Wellengleichung mittels Ensemble Kalman Glättungsfilter.

Acknowledgements

I would like to express my deepest gratitute to my supervisor Otmar Scherzer for his guidance, support, patience, and availability through the time of this thesis. Furthermore, I would like to thank my parents for their unfailing support, and my dear friends Chris Kilian, Mark Strempel, and Johanna Marbach.

Contents

1 Introduction and basic probability theory **1**
- 1.1 Introduction . . . 1
- 1.2 Notation and basic terminology . . . 2
- 1.3 Abstract random elements . . . 4
- 1.4 Random elements of separable Banach spaces . . . 10
- 1.5 Fractional powers of operators and pseudoinverses . . . 13
- 1.6 Gaussian measures on separable Banach spaces . . . 14
- 1.7 Stochastic calculus in Hilbert spaces . . . 19
- 1.8 Comments and further references . . . 23

2 Stochastic filtering **25**
- 2.1 The general filtering problem . . . 25
- 2.2 The Kalman filter . . . 26
- 2.3 The optimal linear filter for non-Gaussian noise . . . 28
- 2.4 The Kalman-Bucy filter . . . 32
- 2.5 Linear smoothers . . . 34
- 2.6 Comments and further references . . . 37

3 Approximate filters **41**
- 3.1 The extended Kalman filter . . . 41
- 3.2 The iterated extended Kalman filter . . . 43
- 3.3 The ensemble Kalman filter . . . 45
- 3.4 Ensemble smoothers . . . 49
- 3.5 Ensemble Kalman inversion . . . 50
- 3.6 Comments and further references . . . 54

4 Real-time parameter estimation **56**
- 4.1 Filtering of uncertain systems . . . 56
- 4.2 The joint extended Kalman filter for parameter estimation . . . 57
- 4.3 A dual iterated extended Kalman filter . . . 57
- 4.4 A dual ensemble Kalman filter . . . 59
- 4.5 Comments and further references . . . 62

5 Comparison with deterministic methods **64**
- 5.1 Regularized Gauss-Newton method and the extended Kalman filter . . . 64
- 5.2 Reduced approach and the dual iterated extended Kalman filter . . . 67
- 5.3 Comments and further references . . . 68

6 Applications **70**
- 6.1 Potential identification . . . 70
- 6.2 Stochastic wave equation . . . 72

7 Conclusion **77**

Bibliography **79**

Chapter 1

Introduction and basic probability theory

1.1 Introduction

Most real-world problems can be formulated as inference or control of dynamical systems in the presence of uncertainty. Filtering is concerned with the former problem. In all areas of science and technology, it is important to estimate the state of a system from observations. These observations are usually incomplete and noisy. Furthermore, there usually is uncertainty about the system itself. This uncertainty might either be classified as *parametric uncertainty*, meaning that the exact behavior of the system is unknown and any model is only a likely candidate among many others, or *intrinsic uncertainty*, which means that the system shows behavior which cannot be consistently predicted.

Filtering methods try to estimate the state of a system in the presence of uncertainty, given a set of noisy and incomplete observations. Since this problem is of an inherently stochastic nature, it is approached using probability theory.

Classically, filtering theory was developed in finite dimensions where both the state X_k of the system and its observation Y_k at a discrete time k can be represented by vectors

$$X_k = \begin{bmatrix} X_k^1 \\ \vdots \\ X_k^d \end{bmatrix}, \qquad Y_k = \begin{bmatrix} Y_k^1 \\ \vdots \\ Y_k^m \end{bmatrix}.$$

First mathematical treatments of the filtering problem considered the linear case, where the future state X_k of the system is a linear function of the past X_{k-1}, and each observation Y_k is a linear function of the current state X_k:

$$\begin{aligned} X_k &= F_k X_{k-1}, \\ Y_k &= M_k X_k, \qquad k \in \mathbb{N}. \end{aligned}$$

Intrinsic noise in the system is modelled by adding a random vector W_k, such that the future state is now a *random* function of the past. Noise in the observation (for example resulting from an inexact measurement procedure) can be modeled similarly by an additive random noise vector V_k, such that the observed dynamical system is modeled by

$$\begin{aligned} X_k &= F_k X_{k-1} + W_k, \\ Y_k &= M_k X_k + V_k. \end{aligned}$$

The seminal article by Kalman [Kal60] then showed that the optimal linear estimator of the state X_k given the available measurements $Y_1, \ldots, Y_k$ of this system can be computed in a recursive way:

The optimal estimator at time k is a linear function of the optimal estimator at time $k-1$ and the new measurement Y_k. This allows for a very efficient real-time computation of optimal estimates. The simplicity of this method, which came to be known as the Kalman filter, revolutionized the field of control theory and engineering. The Kalman filter played in particular a vital role in the success of the Apollo Moon project (see [GA10]).

Moreover, extensions of the Kalman filter to alternative filtering problems have been studied extensively. On a practical side, the extended Kalman filter and the more recent ensemble Kalman filter are two classes of methods that are based on the Kalman filter and yield approximate solutions to the nonlinear filtering problem. On the theoretical side, consideration of more and more general problems led to the development of the theory of stochastic filtering, which culminated in the work of Stratonovich [Str60], Kushner [Kus67] and Zakai [Zak69].

In this thesis, we will discuss how the Kalman filter and its continuous version, the Kalman-Bucy filter, can be formulated for filtering problems in infinite-dimensional spaces. This possibility was suggested quite soon after the conception of the Kalman filter [Fal67].

Infinite-dimensional filtering problems are relevant when the dynamical system of interest is modelled by a partial differential equation. If we also want to model intrinsic noise in the system, this leads to stochastic partial differential equations. In the functional-analytic approach, partial differential equations are treated as evolution equations where the state takes values in an infinite-dimensional space of functions. Similarly, stochastic partial differential equations can be understood as stochastic differential equations in Hilbert spaces.

In the rest of this chapter, we will present the necessary prerequisites from probability theory to allow the abstract formulation of stochastic filtering with partial differential equations that we will present in chapter 2. In particular, we will then see that the infinite-dimensional Kalman filter can be derived as solution to the linear discrete-time filtering problem in separable Banach spaces. From then on, we will concentrate on the continuous-discrete filtering problem, where the state evolves continuously but is observed at discrete timepoints. We will derive a variant of the Kalman-Bucy filter that solves the linear continuous-discrete filtering problem in Hilbert spaces. Furthermore, we will show that this filter also leads to a solution of the smoothing problem, i.e. the problem of sequentially updating the estimate of a past state given new measurements of the future trajectory.

In chapter 3, we will then further discuss the extended Kalman filter and the ensemble Kalman filter, show how they can be formulated in Hilbert spaces, and present further extensions of these basic approaches. In particular, we will see how the ensemble Kalman filter can be applied even to inverse problems that are not of a sequential nature.

The problem of parametric uncertainty in our state model is adressed in chapter 4, where we will see how the methods of the previous chapter lead to general algorithms for real-time simultaneous state and parameter estimation in dynamical systems. We will compare these approaches with deterministic, optimization-based methods in chapter 5, and in fact demonstrate that both the extended Kalman filter and its ensemble version can be interpreted as regularized optimization methods.

Finally, we will demonstrate in chapter 6 how the application of ensemble-based filtering methods to stochastic partial differential equations could look like in principle. In particular, we observe that these methods lead to derivative-free algorithms that can cope with high-dimensional problems.

1.2 Notation and basic terminology

Usually, we will consider separable Banach or Hilbert spaces. Given a Banach space E, we will use the symbol E^* to denote its topological dual, i.e. the space of all continuous linear functionals on E. Elements of the dual space will usually be denoted by $u^*, v^* \in E^*$. We will use the duality brackets

$$\langle u, u^* \rangle := u^*(u),$$

to denote the action of a linear functional $u^* \in E^*$ on an element $u \in E$. For the inner product in a Hilbert space H, we will use round brackets $(\cdot, \cdot) : H \times H \to \mathbb{R}$.

We will denote with $\mathrm{Lin}(E_1; E_2)$ the set of linear maps from a Banach space E_1 to a Banach space E_2, and denote with $\mathcal{L}(E_1; E_2) \subset \mathrm{Lin}(E_1; E_2)$ the space of all bounded linear operators, equipped with

the operator norm. We will furthermore use the shorthand $\mathcal{L}(E) := \mathcal{L}(E; E)$. For an unbounded linear operator A from E_1 to E_2, we will denote its domain by $\operatorname{dom} A$ and its range by $\operatorname{ran} A$.

We will call an operator $R : E^* \to E$ **symmetric** if $\langle Ru^*, v^* \rangle = \langle Rv^*, u^* \rangle$, for all $u^*, v^* \in E^*$, and **positive** if $\langle Ru^*, u^* \rangle \geq 0$, for all $u^* \in E^*$. For an operator $R : H \to H$ in a Hilbert space, this definition reduces to the conditions $(Rh, g) = (Rg, h)$ and $(Rh, h) \geq 0$ for $h, g \in H$.

Recall the definition of the **Fréchet derivative** (or total derivative) of a function $F : E_1 \to E_2$ between Banach spaces [Lan93]: If existent, $DF(x)$ is the unique linear operator such that

$$\lim_{\|h\| \to 0} \frac{\|F(x+h) - F(x) - DF(x)h\|}{\|h\|} = 0.$$

It can be computed throught the identity

$$DF(x)h = \frac{d}{d\epsilon} F(x + \epsilon h)\Big|_{\epsilon=0}.$$

We call F a C^1 function if $DF : E_1 \to \mathcal{L}(E_1; E_2)$ is continuous. The term $DF(x)h$ can also be interpreted as the directional (or Gâteaux) derivative of f along h at the point x. Recall also that classical results from multivariate calculus (like Taylor approximation or the mean value theorem) also hold for general Fréchet derivatives (see e.g. [Lan93]).

In the following, we will use the notation $D_x[f(x, y)](x_0)$ to denote the total derivative of the function $F_y(x) := f(x, y)$ at the point x_0, meaning

$$D_x[f(x, y)](x_0) := DF_y(x_0).$$

Given a measure space $(\Omega, \mathcal{A}, \mu)$, $L^p(\mu) := L^p(\Omega, \mathcal{A}, \mu)$ will denote the space of p-integrable functions, modulo equality μ-almost everywhere. For a given Banach space E, $L^p(\mu; E) := L^p(\Omega, \mathcal{A}, \mu; E)$ will denote the Bochner space of all Bochner measurable functions $f : \Omega \to E$ with

$$\|f\|_{L^p(\mu;E)} := \left(\int_\Omega \|f(x)\|^p \mu(dx) \right)^{1/p} < \infty,$$

again modulo equality μ-almost everywhere (cf. [DS88] or [Kre15]). We will also use the notation $M(\mu; E)$ for the space of all $\mathcal{A}$-measurable functions from Ω to E, modulo equality μ-almost everywhere, and $M(E_1; E_2)$ to denote the space of all Borel measurable functions between the spaces E_1 and E_2. Recall also that on separable Banach spaces, the notions of Bochner measurability and Borel measurability are equivalent.

Furthermore, we will use the tensor product notation. Let E_1 and E_2 be Banach spaces. For $u \in E_1, v \in E_2$, we define a linear operator $u \otimes v \in \mathcal{L}(E_2^*; E_1)$ by

$$\langle (u \otimes v)v^*, u^* \rangle := \langle u, u^* \rangle \langle v, v^* \rangle.$$

Clearly, $\|u \otimes v\| \leq \|u\| \|v\|$, so that the tensor product always yields a bounded operator. In the Hilbert space case, $h \otimes g$ yields for all $h \in H_1$ and $g \in H_2$ an operator in $\mathcal{L}(H_2; H_1)$, given by

$$h \otimes g(f) := h \cdot (g, f), \qquad f \in H_2.$$

Finally, for finite sequences of natural numbers we will sometimes use the notation

$$[N] := \{1, 2, 3, ..., N\} \qquad N \in \mathbb{N}.$$

And we will use the symbol

$$\mathbb{R}^+ := [0, \infty)$$

to denote the set of nonnegative real numbers.

1.3 Abstract random elements

Let $\mathcal{P}(\Omega, \mathcal{A})$ denote the space of probability measures on a measurable space $(\Omega, \mathcal{A})$. If Ω is a topological space and $\mathcal{A} = \mathcal{B}(\Omega)$ is its Borel σ-algebra, we will use the brief notation $\mathcal{P}(\Omega) := \mathcal{P}(\Omega, \mathcal{B}(\Omega))$ to denote the same space. Next, we define what we mean by calling an element of a measurable space *random*.

Definition 1.3.1 (Random element)**.** *Let $(S, \mathcal{S})$ be a measurable space. A **random element** X of S consists of a probability space $(\Omega, \mathcal{A}, \mathbb{P})$ and a measurable map $X : \Omega \to S$. We define its **distribution** $\mathbb{P}^X \in \mathcal{P}(S, \mathcal{S})$ as the pushforward of $\mathbb{P}$ by X,*

$$\mathbb{P}^X(B) := \mathbb{P}(X^{-1}(B)), \qquad B \in \mathcal{S}.$$

*For a Bochner measurable function $f : S \to E$ into a Banach space E, we define the **expectation** of $f(X)$ as the Bochner integral*

$$\mathbb{E}f(X) := \int_\Omega f(X(\omega))\mathbb{P}(d\omega) = \int_S f(x)\mathbb{P}^X(dx).$$

By Bochner's integrability theorem [Kre15, theorem 2.5], the expectation exists if and only if $\mathbb{E}\|f(X)\| < \infty$.

It is common to identify a random element with its defining map, and hence we use the notation

$$X : (\Omega, \mathcal{A}, \mathbb{P}) \longrightarrow (S, \mathcal{S})$$

to mean that "X is a random element of S defined on the probability space $(\Omega, \mathcal{A}, \mathbb{P})$". Usually, however, we will simply write "Let X be a random element of S..." without explicit reference to the associated probability space, and where we will assume that S is a topological space equipped with its Borel σ-algebra, $\mathcal{S} = \mathcal{B}(S)$, if not stated otherwise. Nevertheless, *we will always assume that all random elements are defined on a common probability space.*

Since all relevant properties of X are uniquely determined by its law, we will also use the notation

$$X \sim \mu$$

with a probability measure $\mu \in \mathcal{P}(S)$ to mean that X is a random element of S with distribution $\mathbb{P}^X = \mu$. It is a basic fact of probability theory that given any probability measure, such an element can always be constructed. We will also denote equality in distribution of two random elements X and Y with

$$X \overset{d}{=} Y,$$

meaning that $\mathbb{P}^X = \mathbb{P}^Y$.

Often, we will work with $S = \mathbb{R}$ or $S = \mathbb{R}^d$ equipped with the respective Borel σ-algebras. A random element of $\mathbb{R}$ is usually called **random variable**, and we will reserve the term **random vector** for random elements of $\mathbb{R}^d$. Recall that we say a measure μ on $(\Omega, \mathcal{A})$ is **absolutely continuous** with respect to another measure ν on (Ω, A), and write

$$\mu \ll \nu,$$

if

$$\nu(N) = 0 \Longrightarrow \mu(N) = 0,$$

for all $N \in \mathcal{A}$. If μ and ν are σ-finite, the Radon-Nikodym theorem (see e.g. [Kal02, theorem 2.10]) states that there exists a measurable function $\frac{d\mu}{d\nu} \in M(\nu; \mathbb{R}^+)$ such that

$$\int f(x)\mu(dx) = \int f(x)\frac{d\mu}{d\nu}\nu(dx),$$

for all functions $f \in L^1(\mu)$. If μ is a probability measure, $\frac{d\mu}{d\nu}$ is called the **probability density function** of μ with respect to ν. The most relevant application is of course the case where $S = \mathbb{R}^d$ and ν is taken to be the d-dimensional Lebesgue measure λ^d. In that case, we will also use differential notation, where

$$\mu(dx) = p(x)dx$$

means that $p \in L^1(\lambda^d)$ is the probability density of μ with respect to the Lebesgue measure. If X is a random vector and $\mathbb{P}^X << \lambda^d(x)$, we define the **probability density function of** X by

$$p_X(x) := \frac{\mathbb{P}^X(dx)}{dx}.$$

Moreover, given a probability space $(\Omega, \mathcal{A}, \mathbb{P})$, we use the word **almost surely** if an event holds with probability 1 or, equivalently, $\mathbb{P}$-almost everywhere. For example, for a random element $X : \Omega \to \mathbb{R}$, the formula

$$X > 0,$$

has to be read as

$$\mathbb{P}(\{\, \omega \in \Omega :\ X(\omega) > 0 \,\}) = 1.$$

Note that random processes also fall under definition 1.3.1: A random process is a random element of a function space. To be more precise:

Definition 1.3.2 (Random process)**.** *Let $(S, \mathcal{S})$ be a measurable space and T a totally ordered set. Then, an S-**valued random process** on T is any random element X of $(S^T, \mathcal{S}^T)$, where $\mathcal{S}^T$ is the σ-algebra on S^T generated by the point evaluations $\pi_t : S^T \to S, \quad \omega \mapsto \omega(t)$. We will denote the point evaluations of the random process X with $X(t) := \pi_t X$.*

*If X is almost surely in some subset $F \subset S^T$, we also say that X **has paths in** F.*

Random processes (more frequently called "stochastic processes") are useful models in a multitude of problems in science and engineering. They also provide a theoretical framework for the probabilistic approach to parameter identification in dynamical systems.

Example 1.3.3 (Wave equation with uncertain parameters). Consider the wave equation on a domain $D \subset \mathbb{R}^3$.

$$\begin{aligned}
&\partial_t^2 u(x,t) = \Delta u(x,t) && x \in D, t \in [0,T],\\
&u(x,t) = 0, && x \in \partial D, t \in [0,T],\\
&u(x,0) = f(x), && x \in D,\\
&\partial_t u(x,0) = 0, && x \in D.
\end{aligned}$$

Suppose the initial pressure f is unknown. We can model it as a random variable $F : \Omega \to H_0^1(D)$ on some probability space $(\Omega, \mathcal{A}, \mathbb{P})$. In this model, the solution of the wave equation is a $H_0^1(D)$-valued stochastic process $U = U(t, \omega)$ on $[0, T]$, which satisfies for $\mathbb{P}$-almost every $\omega \in \Omega$

$$\begin{aligned}
&\partial_t U(t,\omega) = \Delta U(t,\omega), && t \in [0,T],\\
&U(t,\omega)\Big|_{\partial D} = 0, && t \in [0,T],\\
&U(0,\omega) = F(\omega),\\
&\partial_t U(0,\omega) = 0.
\end{aligned}$$

Usually, we suppress the dependence of U on ω and simply write

$$\begin{aligned}
&\partial_t U(t,\omega) = \Delta U(t,\omega), && t \in [0,T],\\
&U(t,\omega)\Big|_{\partial D} = 0, && t \in [0,T],\\
&U(0,\omega) = F(\omega),\\
&\partial_t U(0,\omega) = 0,
\end{aligned}$$

keeping in mind that $U(t)$ is still a random element.

In example 1.3.3, one could then try to infer the unknown parameter F from measurements of the solution U. In the probabilistic approach, this question would be formulated as the problem of determining the conditional distribution of F given U. In order to deal with this rigorously, we need to formalize the notion of conditional probabilites and conditonal expectation for abstract random elements that may take values in infinite-dimensional spaces. The traditional approach constructs the conditional expectations first, and then defines the conditional probability through the basic identity

$$\mathbb{E}\mathbb{1}_A = \mathbb{P}(A),$$

where $\mathbb{1}_A$ is interpreted as a random variable on the probability space $(\Omega, \mathcal{A}, \mathbb{P})$.

Theorem 1.3.4 (Conditional expectation)**.** *Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space, E a Banach space, and $\mathcal{B} \subset \mathcal{A}$ a sub-σ-algebra. Then there is a unique linear operator $\mathbb{E}^{\mathcal{B}} : L^1(\Omega, \mathcal{A}, \mathbb{P}; E) \to L^1(\Omega, \mathcal{B}, \mathbb{P}; E)$, called the **conditional expectation given** $\mathcal{B}$, such that*

$$\int_B X(\omega)\mathbb{P}(d\omega) = \int_B (\mathbb{E}^{\mathcal{B}}X)(\omega)\mathbb{P}(d\omega), \qquad \forall B \in \mathcal{B},$$

for all random elements $X \in L^1(\Omega, \mathcal{A}, \mathbb{P}; E)$.

Proof. [VTC87, proposition 4.1]. □

It is important to keep in mind that $\mathbb{E}^{\mathcal{B}}X$ is again a random element, but now on the smaller probability space $(\Omega, \mathcal{B}, \mathbb{P})$.

Suppose we are given a random element Y of some measurable space $(T, \mathcal{T})$.

Then, we define the conditional expectation of X given Y as

$$\mathbb{E}[X|Y] := \mathbb{E}^{\sigma(Y)}X,$$

where $\sigma(Y)$ denotes the σ-algebra generated by the sets $Y^{-1}(B)$, with $B \in \mathcal{T}$. Furthermore, we can give formal meaning to "the expectation of X given $Y = y$". This is achieved through Doob's functional representation lemma, which states that any $\sigma(Y)$-measurable function can be written as a deterministic function of Y (cf. [Kle13, chapter 8]).

Lemma 1.3.5. *Let $(\Omega, \mathcal{A})$ and $(T, \mathcal{T})$ be a measurable spaces. and let $Y : \Omega \to T$ be a measurable function. Suppose $X : \Omega \to S$ is a function into a measurable space $(S, \mathcal{S})$. Then X is $\sigma(Y) - \mathcal{S}$-measurable if and only if there exists a measurable function $F : T \to S$ with*

$$X(\omega) = F(Y(\omega)),$$

for all $\omega \in \Omega$.

Proof. See [Kal02, theorem 1.13] □

Applying this result to the function $\mathbb{E}[X|Y]$ leads to a functional representation.

Corollary 1.3.6. *There exists a measurable function $F : \Omega \to E$ such that*

$$\mathbb{E}[X|Y](\omega) = F(Y(\omega)),$$

*for all $\omega \in \Omega$. We then define for $y \in T$ the **conditonal expectation of** X**, given** $Y = y$, by*

$$\mathbb{E}[X|Y = y] := F(y).$$

Proof. By theorem 1.3.4, $\mathbb{E}[X|Y]$ is $\sigma(Y)$-measurable. Hence, by lemma 1.3.5, there exists a $\mathcal{T}$-measurable function $F : T \to E$ such that $\mathbb{E}[X|Y] = F(Y)$. □

This definition corresponds to intuition, since we have indeed for any measurable function $f : E \times E \to \mathbb{R}$

$$\mathbb{E}[f(X,Y)|Y=y] = \mathbb{E}[f(X,y)], \qquad \forall y \in E,$$

if the integrals exist (see [PZ14, chapter 1]).

Now, we can list some basic properties that are useful for computing conditional expectations in practice.

Proposition 1.3.7 (Properties of the conditional expectation)**.** *Let X and Y be random elements of a Banach space E, with $\mathbb{E}\|X\| < \infty$. Then:*

i) $\|\mathbb{E}[X|Y]\| \leq \mathbb{E}\|X\|$.

ii) If X is independent of Y, then $\mathbb{E}[X|Y=y] = \mathbb{E}X$, for all $y \in E$.

iii) If X is $\sigma(Y)$-measurable, then $\mathbb{E}[X|Y] = X$.

iv) Tower property: $\mathbb{E}[\mathbb{E}[X \mid Y]] = \mathbb{E}X$.

v) $X - \mathbb{E}[X|Y]$ is independent of Y.

Proof. See [VTC87, p. II.4.1]. □

The next example shall help to demonstrate the distinction between expectation and conditional expectation.

Example 1.3.8. Let X be a random element of a Banach space E with expectation $\mathbb{E}X = 0$. Consider the random vector Y defined by

$$Y = MX + b,$$

where $M \in \mathcal{L}(E)$ and $b \in E$. Then, the expectation of Y is

$$\mathbb{E}Y = M\mathbb{E}[X] + b = b.$$

But the conditional expectation of Y, given X, is

$$\mathbb{E}[Y|X] = M\mathbb{E}[X|X] + b = MX + b,$$

and we furthermore have for a given $x \in E$

$$\mathbb{E}[Y|X=x] = Mx + b.$$

In particular, $\mathbb{E}[Y|X]$ is again a random vector, while $\mathbb{E}Y$ and $\mathbb{E}[Y|X=y]$ are determistic.

Next, we define the **conditional probability of an event** $\{X \in A\} \in \mathcal{A}$ for a random element X of a measurable space S, given $Y = y$.

$$\mathbb{P}(X \in A \mid Y) := \mathbb{E}[\mathbb{1}_A(X)|Y].$$

However, in general this does not yield a probability measure (cf. [Kal02, chapter 6]). But if we assume a little additional topological structure on the space S, the following theorem states that we can always find a version of $\mathbb{P}(X \in \cdot \mid Y)$ that yields a probability measure when evaluated at a point $\omega \in \Omega$. Such a measure (or rather, the associated probability kernel) is called a regular conditional distribution. A sufficient condition on Ω is that it is a **Polish space**. Recall that we call a topological space Polish if it is separable and completely metrizable (see [Bog07, chapter 6]).

Theorem 1.3.9 (Regular conditional distribution)**.** *Let $X : (\Omega, \mathcal{A}, \mathbb{P}) \to (S, \mathcal{B}(S))$ be a random element of a Polish space S and let $\mathcal{B}$ be a sub-σ-algebra of $\mathcal{A}$. Then, there exists a unique function*

$$\begin{aligned}\mathbb{P}^{X|\mathcal{B}} :& \mathcal{B}(S) \times \Omega \to [0,1],\\ &(A,\omega) \longmapsto \mathbb{P}^{X|\mathcal{B}}(A)(\omega),\end{aligned}$$

such that

i) for any $\omega \in \Omega$, $\mathbb{P}^{X|\mathcal{B}}(dx)(\omega)$ is a probability measure on $(S, \mathcal{B}(S))$,

ii) for any $A \in \mathcal{B}S$, $\mathbb{P}^{X|\mathcal{B}}(A)$ is a $\mathcal{B}$-measurable function, and

iii) $\mathbb{P}^{X|\mathcal{B}}(A) = \mathbb{E}^{\mathcal{B}} \mathbb{1}_A(X)$ almost surely, for all $A \in \mathcal{B}(S)$.

*We call $\mathbb{P}^{X|\mathcal{B}}$ the **regular conditional probability distribution** of X given $\mathcal{B}$.*

Proof. Follows from [Kal02, theorem 6.3]. □

Here, the notation $\mathbb{P}^{X|\mathcal{B}}(dx)(\omega)$ denotes the measure $A \mapsto \mathbb{P}^{X|\mathcal{B}}(A)(\omega)$. Objects of this form (meaning bivariate functions that are measurable in one argument, and measures in the other) are called **probability kernels** (or also "Markov kernels" or "transition kernels").

Similar to the conditional expectation, we define for a random element Y of an arbitrary measurable space $(T, \mathcal{T})$ the **conditional distribution of X, given Y**, as

$$\mathbb{P}^{X|Y} := \mathbb{P}^{X|\sigma(Y)}.$$

Example 1.3.10. In the situation of example 1.3.3, the distribution of the solution $U(t)$ is a probability measure on $H_0^1(D)$, and the distribution of the whole process U is a probability measure on the space $L^1([0,T]; H_0^1(D))$. The unknown initial pressure F is a random element of $H_0^1(D)$, which is certainly Polish. Thus, a regular conditional distribution $\mathbb{P}^{F|U}$ of F given U exists.

Then, as before, it follows from theorem 1.3.9 above and lemma 1.3.5 (Doob) that $\mathbb{P}^{X|Y}$ can be written as a deterministic function of Y. This means that we can always find a map $\mathfrak{M} : S \to \mathcal{P}(S)$ into the space of probability measures, such that

$$\mathbb{P}^{X|Y} = \mathfrak{M}(Y).$$

It might be helpful to keep in mind that both sides of the identity denote regular conditional probabilites (i.e probability kernels). For a given $y \in T$, we can then define the **conditional distribution of X, given $Y = y$** by

$$\mathbb{P}^{X|Y}(\cdot|y) := \mathfrak{M}(Y).$$

This functional representation motivates the following shorthand:

$$X \mid Y \sim \mathfrak{M}(Y)$$

means that for all $y \in S$, $\mathbb{P}^{X|Y}(\cdot|y)$ is identical to the probability measure $\mathfrak{M}(y)$.

Example 1.3.11. For the one-dimensional normal distribution $\mathcal{N} : \mathbb{R} \times (0, \infty) \to \mathcal{P}(\mathbb{R})$,

$$\mathcal{N}(m, \sigma^2) := \frac{1}{\sqrt{2\pi\sigma^2}} e^{(x-m)^2/2\sigma^2} dx,$$

the expression $X|Y \sim \mathcal{N}(Y, \sigma^2)$ means that for any $y \in \mathbb{R}$, the conditional distribution of X, given $Y = y$, is normal with mean y and variance σ^2.

Bayes theorem

The most important formula for conditional probabilities is Bayes' theorem, which relates a conditional probability $\mathbb{P}(A|B)$ with its likelihood $\mathbb{P}(B|A)$, for any two events A, B with $\mathbb{P}(A), \mathbb{P}(B) > 0$:

$$\mathbb{P}(A|B) = \mathbb{P}(A) \frac{\mathbb{P}(B|A)}{\mathbb{P}(B)}.$$

For random vectors, there is a well-known density version.

Proposition 1.3.12 (Bayes' theorem for probability densities)**.** *Let X and Y be random vectors with joint probability density function $p_{X,Y}$. Then the conditional probability density functions $p_{X \mid Y}$ and $p_{Y \mid X}$ exist and are related by*

$$p_{X \mid Y}(x|y) = p_X(x) \frac{p_{Y \mid X}(x|y)}{p_Y(y)}. \tag{1.3.1}$$

Proof. As in [Kle13, example 8.31], we have

$$p_{Y|X}(y|x) = \frac{p_{X,Y}(x,y)}{p_X(x)},$$
$$p_{X|Y}(x|y) = \frac{p_{X,Y}(x,y)}{p_Y(y)}.$$

Combining the two identities yields

$$p_{X|Y}(x,y) = \frac{p_X(x)p_{X,Y}(x,y)}{p_Y(y)},$$

for $\mathbb{P}^X$-almost every x and $\mathbb{P}^Y$-almost every y. □

Now, we are ready to give an extension of Bayes' rule that is suitable to our general setting (see [Hel18]).

Theorem 1.3.13 (Bayes)**.** *Let X and Y be random elements of Polish spaces S and T respectively. Let, moreover, μ be a σ-finite measure on T and assume that*

$$\mathbb{P}^{Y|X}(dy|x) \ll \nu(dy), \qquad \text{for } \mathbb{P}^X\text{-a.e. } x \in S,$$

and denote the corresponding density with

$$p_{Y|X}(y|x) := \frac{\mathbb{P}^{Y|X}(dy|x)}{\nu(dy)}.$$

Then, $\mathbb{P}^{X|Y}(dx|y) \ll \mathbb{P}^X(dx)$ for $\mathbb{P}^Y$-a.e. $y \in T$, and

$$\frac{\mathbb{P}^{X|Y}(dx|y)}{\mathbb{P}^X(dx)} = \frac{p_{Y|X}(y|x)}{Z(y)}, \tag{1.3.2}$$

where $Z(y) := \int p_{Y|X}(y|x)\mathbb{P}^X(dx) \in (0,\infty)$ $\mathbb{P}^Y$-a.e.

The proof of this result is essentially analogous to the proof of the most basic form of Bayes rule, where one simply notes that

$$\mathbb{P}(B)\mathbb{P}(A|B) = \mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B|A), \tag{1.3.3}$$

and then divides by $\mathbb{P}(B)$ if it is nonzero. The measure-theoretic analogon of this identity is provided by the following lemma:

Lemma 1.3.14 (Product rule for measures)**.** *Let X and Y be random elements of Polish spaces S and T respectively. Then,*

$$\begin{aligned}\mathbb{P}^{X,Y}(A \times B) &= \int_A \mathbb{P}^{Y|X}(B|x)\mathbb{P}^X(dx) \\ &= \int_B \mathbb{P}^{X|Y}(A|y)\mathbb{P}^Y(dy),\end{aligned}$$

for all $A \in \mathcal{B}(S), B \in \mathcal{B}(T)$.

Proof. See [Bog07, chapter 10]. □

In differential notation, the similarity to (1.3.3) becomes even more obvious:

$$\mathbb{P}^Y(dy)\mathbb{P}^{X|Y}(dx|y) = \mathbb{P}^{X,Y}(dx,dy) = \mathbb{P}^X(dx)\mathbb{P}^{Y|X}(dy|x).$$

Then, similarly, one wants to "divide by $\mathbb{P}^Y(dy)$", which means formally that we desire to take the Radon-Nikodym derivative. The proof then executes this informal argument:

Proof of theorem 1.3.13. We follow [Hel18]. First, we note using Fubini's theorem that we have for any $B \in \mathcal{B}(T)$:

$$\begin{aligned}\mathbb{P}^Y(B) &= \mathbb{P}^{X,Y}(S \times B) \\ &= \int \mathbb{P}^{Y|X}(B|x)\mathbb{P}^X(dx) \\ &= \int \int_B p(y|x)\nu(dy)\mathbb{P}^X(dx) \\ &= \int_B Z(y)\nu(dy). \end{aligned} \tag{1.3.4}$$

and so in particular $\mathbb{P}^Y \ll \nu$. We first need to show that $Z(y) \in (0, \infty)$ for P^Y-almost every $y \in T$. This follows, since

$$\mathbb{P}^Y(Z = 0) = \int_{Z=0} Z(y)\nu(dy) = 0,$$

which means $Z > 0$ ν-almost everywhere and thus $\mathbb{P}^Y$-almost everywhere. Also

$$1 \geq \mathbb{P}^Y(Z = \infty) = \int_{Z=\infty} Z(y)\nu(dy),$$

which implies first $\nu(Z = \infty) = 0$ and then $\mathbb{P}^Y(Z = \infty) = 0$. Having established that, it remains to prove (1.3.2). We have on one hand

$$\begin{aligned}\mathbb{P}^{X,Y}(A \times B) &= \int_A \mathbb{P}^{Y|X}(B|x)\mathbb{P}^X(dx) \\ &= \int_A \int_B p_{Y|X}(y|x)\nu(dy)\mathbb{P}^X(dx) \\ &= \int_B \int_A p_{Y|X}(y|x)\mathbb{P}^X(dx)\nu(dy), \end{aligned} \tag{1.3.5}$$

and on the other hand, using equation 1.3.4,

$$\begin{aligned}\mathbb{P}^{X,Y}(A \times B) &= \int_B \mathbb{P}^{X|Y}(A|y)\mathbb{P}^Y(dy) \\ &= \int_B Z(y)\mathbb{P}^{X|Y}(A|y)\nu(dy). \end{aligned} \tag{1.3.6}$$

Comparing (1.3.5) and (1.3.6) which hold for arbitrary sets $B \in \mathcal{B}(T)$, we necessarily have

$$\int_A p_{Y|X}(y|x)\mathbb{P}^X(dx) = Z(y)\mathbb{P}^{X|Y}(A|y)$$

for ν-almost every $y \in T$, and thus $\mathbb{P}^Y$-almost every $y \in T$. Dividing by $Z(y)$ and taking the Radon-Nikodym derivative then yields equation (1.3.2). □

Note that in the setting of Bayesian inference, $\mathbb{P}^X$ would be called the prior, $\mathbb{P}^{Y|X}$ the likelihood, and $\mathbb{P}^{X|Y}$ the posterior. Note also that the σ-finite measure ν in the proof plays the role of a base measure with respect to which we derive densities for the other measures, similar to the Lebesgue measure in $\mathbb{R}^d$. And indeed, if one applies this result with $\nu = \lambda^d$, one can quickly recover proposition 1.3.12.

1.4 Random elements of separable Banach spaces

The characterization of probability measures is a central question for both pure and applied probability theory. A central concept related to this problem is the characteristic functional of a measure [VTC87, chapter 4]:

Definition 1.4.1 (Characteristic functional)**.** *Let E be a Banach space and $\mu \in \mathcal{P}(E)$. Then, we define* ***the characteristic functional of*** *μ as*

$$\hat{\mu}(u^*) := \int e^{i\langle x, u^*\rangle} \mu(dx), \qquad u^* \in E^*.$$

Let X be a random element of E. Then, we define the ***characteristic functional of*** *X as:*

$$\phi_X(u^*) := \mathbb{E}[e^{i\langle X, u^*\rangle}], \qquad u^* \in E^*.$$

Obviously, for any random element X, we have $\phi_X = \widehat{\mathbb{P}^X}$.

A large class of measures can be uniquely characterized by their characteristic functionals. In particular, this holds for the special case that will be relevant to us, namely Borel measures on separable Banach spaces.

Lemma 1.4.2. *Let E be a separable Banach space and $\mu_1, \mu_2 \in \mathcal{P}(E)$. If $\hat{\mu}_1 = \hat{\mu}_2$, then $\mu_1 = \mu_2$.*

Proof. [VTC87, IV.2.1, corollary 1] states that all Radon measures are uniquely determined by their characteristic functionals. But Borel measures on separable Banach spaces are always Radon (see [Sch74]). □

The definition of the covariance of a random element is again an extension of the finite-dimensional case, but less straightforward than the definition of the expectation. It is based on the fact that given a probability measure μ on a Banach space E we can regard elements of E^* as random variables on $(E, \mathcal{B}(E), \mu)$. Then, the covariance is defined as the bilinear form (which can then be identified with a linear operator) that maps any two dual elements $u^*, v^* \in E^*$ to their (one-dimensional) cross-covariance:

Proposition 1.4.3 (Covariance and cross-covariance)**.** *Let $X : (\Omega, \mathcal{A}, \mathbb{P}) \to (E_1, \mathcal{B}(H_1)$ and $Y : (\Omega, \mathcal{A}, \mathbb{P}) \to (E_2, \mathcal{B}(E_2)$ be random elements of separable Banach spaces E_1 and E_2, with $\mathbb{E}\|X\|^2 < \infty$ and $\mathbb{E}\|Y\|^2 < \infty$. Let $\mathcal{B}$ be a sub-σ-algebra of $\mathcal{A}$. Then, there exists an essentially unique random linear operator $\Sigma_{XY} : (\Omega, \mathcal{B}, \mathbb{P}) \to \mathcal{L}(E_2; E_1)$ such that*

$$\langle \Sigma_{XY|\mathcal{B}} v^*, u^*\rangle = \int \langle x, u^*\rangle\langle y, v^*\rangle \mathbb{P}^{X,Y|\mathcal{B}}(dx, dy) - \langle \mathbb{E}^{\mathcal{B}} X, u^*\rangle\langle \mathbb{E}^{\mathcal{B}} Y, v^*\rangle, \qquad u^* \in E_1^*, v^* \in E_2.$$

We call $\Sigma_{XY|\mathcal{B}}$ the ***conditional cross-covariance of*** *X* ***and*** *Y****, given*** *$\mathcal{B}$ and also write*

$$\mathrm{Cov}[X, Y|\mathcal{B}] := \Sigma_{XY|\mathcal{B}}.$$

In the case $E_1 = E_2$ and $X = Y$, we call $\mathrm{Cov}(X|\mathcal{B}) := \Sigma_{XX|\mathcal{B}}$ the ***conditional covariance of*** *X****, given*** *$\mathcal{B}$.*

Proof. Follows from [VTC87, III.2.1, proposition 2.1 and corollary 2]. □

In the Hilbert space case, we can also express the cross-covariance operator in terms of the tensor product. This leads to the more elegant formula

$$\Sigma_{XY|\mathcal{B}} = \mathbb{E}^{\mathcal{B}}[X \otimes Y] - \mathbb{E}^{\mathcal{B}}[X] \otimes \mathbb{E}^{\mathcal{B}}[Y].$$

Given a further random element Z, we can repeat the procedure that we are already familiar with from the conditonal expectation and define

$$\mathrm{Cov}[X, Y|Z] := \mathrm{Cov}[X, Y|\sigma(Z)].$$

And by setting $\mathcal{B} = \mathcal{A}$, we recover the (unconditional) cross-covariance

$$\mathrm{Cov}[X, Y] := \mathrm{Cov}[X, Y|\mathcal{A}].$$

Covariance and cross-covariance operators tend to be very regular. For example, they are always nuclear.

Definition 1.4.4 (Nuclear)**.** *Let E_1 and E_2 be separable Banach spaces. We call an operator $A \in \mathcal{L}(E_2^*; E_1)$* ***nuclear*** *if there are sequences (x_n) in E_1 and (y_n) in E_2^{**} with $\sum_n \|x_n\|\|y_n^{**}\| < \infty$ such that A has the representation*

$$Ay^* = \sum_{n=1}^{\infty} \langle y^*, y_n^{**} \rangle x_n, \qquad y^* \in E_2^*.$$

We then define the ***nuclear norm*** *$\|A\|_{\mathrm{nuc}}$ as the infimum of $\sum_n \|y_n^{**}\|\|x_n\|$ over all such representations.*

Proposition 1.4.5 (Properties of the cross-covariance)**.** *Let X, X_1, X_2 be random elements of a separable Banach space E_1, and Y, Y_1, Y_2 be random elements of a separable Banach space E_2, and assume that they all have finite second moments. Then, for any random element Z:*

i) $\mathrm{Cov}[X, Y|Z]$ *is a nuclear operator with* $\|\mathrm{Cov}[X, Y|Z]\|_{\mathrm{nuc}} \leq \mathbb{E}[\|X\|\|Y\||Z]$.

ii) $\mathrm{Cov}(X|Z)$ *is positive and symmetric.*

iii) $\mathrm{Cov}[AX_1 + X_2, BY_1 + Y_2|Z] = A\mathrm{Cov}[X_1, Y_1|Z]B^* + A\mathrm{Cov}[X_1, Y_2|Z] + \mathrm{Cov}[X_2, Y_2|Z]B^* + \mathrm{Cov}[X_2, Y_2|Z]$, *for any $A \in \mathcal{L}(H_1), B \in \mathcal{L}(H_2)$.*

iv) Let $U := \begin{bmatrix} X \\ Y \end{bmatrix}$. *Then:*

$$\mathrm{Cov}(U|Z) = \begin{bmatrix} \mathrm{Cov}(X|Z) & \mathrm{Cov}[X, Y|Z] \\ \mathrm{Cov}[Y, X|Z] & \mathrm{Cov}(Y|Z) \end{bmatrix}.$$

Proof. The first statement is [VTC87, III.2.4, theorem 2.5]. The second statement is [VTC87, III.2.2, proposition 2.2].

The third statement follows directly from proposition 1.4.3. Assuming without loss of generality that $\mathbb{E}[X|Z] = 0 = \mathbb{E}[Y|Z]$, we have

$$\begin{aligned} \langle \mathrm{Cov}[AX, BY|Z]v^*, u^* \rangle &= \int \langle Ax, u^* \rangle \langle By, v^* \rangle \mathbb{P}^{X,Y|Z}(dx, dy) \\ &= \int \langle x, A^*u^* \rangle \langle y, B^*v^* \rangle \mathbb{P}^{X,Y|Z}(dx, dy) \\ &= \langle \mathrm{Cov}[X, Y|Z]B^*v^*, A^*u^* \rangle = \langle (A\mathrm{Cov}[X, Y|Z]B^*)v^*, u^* \rangle, \end{aligned}$$

and thus, $\mathrm{Cov}[AX, BY|Z] = A\mathrm{Cov}[X, Y|Z]B^*$ by uniqueness. The more general case follows analogously.

The fourth statement can be found in [VTC87, section III.2.4]. □

Before we go on, we look at the Hilbert space case. Here, the concept of a nuclear operator reduces to what is more frequently called a trace-class operator:

Definition 1.4.6 (Trace class)**.** *Let H and U be separable Hilbert spaces with orthonormal bases (h_n) and (u_n). Then, for an operator $A \in \mathcal{L}(H; U)$, we define its* ***trace***

$$\mathrm{tr}\, A := \sum_n \langle Ah_n, u_n \rangle.$$

We say that A is in the ***trace class*** *if $\|A\|_1 := \mathrm{tr}((A^*A)^{1/2}) < \infty$.*

And indeed, the two concepts nuclearity and trace-class are equivalent.

Lemma 1.4.7. *Let H be a separable Hilbert space and $\Sigma \in \mathcal{L}(H)$. Then Σ is nuclear if and only if it is in the trace class, and we have*

$$\|\Sigma\|_{\mathrm{nuc}} = \|\Sigma\|_1.$$

Proof. See [VTC87, section III.1., proposition 1.9] □

A related concept that will also be extremely useful in later chapters is that of a Hilbert-Schmidt operator:

Definition 1.4.8 (Hilbert-Schmidt operator)**.** *We call an operator $A \in \mathcal{L}(H;U)$ **Hilbert-Schmidt** if*

$$\|A\|_2 := \sum_n \|Ah_n\|^2 < \infty.$$

We denote the space of Hilbert-Schmidt operators by $\mathcal{L}_2(H;U)$.

Clearly, definition 1.4.6 reduces in finite dimensions to the familiar concept of the trace of a matrix. And indeed, the trace as defined above shares a lot of properties with its finite-dimensional analogon:

Proposition 1.4.9 (Properties of the trace)**.** *Let H and U be separable Hilbert spaces.*

i) The trace is linear: $\operatorname{tr}(\alpha A + \beta B) = \alpha \operatorname{tr} A + \beta \operatorname{tr} B$, for all $\alpha, \beta \in \mathbb{R}$, $A, B \in \mathcal{L}_1(H;U)$.

ii) For $h, g \in H$, $\langle h, g\rangle = \operatorname{tr}(h \otimes g)$.

iii) Let $A \in \mathcal{L}_1(H)$ and $T \in \mathcal{L}(H)$. Then $\operatorname{tr}(TB) = \operatorname{tr}(BA)$.

iv) $|\operatorname{tr}(TA)| \leq \|T\|\|A\|_1$, for all $A \in BL_1(H)$ and $T \in BL(H)$.

v) $\|\cdot\|_1$ defines a norm on $BL_1(H;U)$.

vi) $\|\cdot\|_2$ defines a norm on $\mathcal{L}_2(H;U)$.

vii) Every trace-class operator is compact.

Proof. [Con94, chapter 9.2]. □

Finally, in the Hilbert space case we can strenghten proposition 1.4.5 a bit.

Proposition 1.4.10 (Properties of the cross-covariance)**.** *Let X, X_1, X_2 be random elements of a separable Hilbert space H_1, and Y, Y_1, Y_2 be random elements of a separable Hilbert space H_2, and assume that they all have finite second moments. Then, for any random element Z:*

i) $\operatorname{Cov}[X, Y|Z]$ is a trace-class operator with $\|\operatorname{Cov}[X, Y|Z]\|_1 \leq \mathbb{E}[\|X\|\|Y\||Z]$.

ii) $\operatorname{tr}(\operatorname{Cov}(X|Z)) = \mathbb{E}[\|X\|^2|Z]$.

Proof. The first statement is a follows from proposition 1.4.5 and lemma 1.4.7. The second statement is given in [VTC87, section III.2.4]. □

1.5 Fractional powers of operators and pseudoinverses

Let $A : \operatorname{dom} A \subset H \to H$ be an unbounded self-adjoint linear operator. For the subsequent analysis, it will be necessary to give a rigorous meaning to the objects $A^{1/2}$, $A^{-1/2}$, and more generally, to A^s for any fractional power $s \in \mathbb{R}$. This will be useful in the case when A is the covariance operator of a Gaussian measure on a Hilbert space, and in chapter 6 when A is a partial differential operator. We give a construction that is based on the following consequence of the spectral theorem for self-adjoint unbounded operators.

Theorem 1.5.1 (Spectral representation)**.** *Let H be a separable Hilbert space and $A : \operatorname{dom} H \subset H \to H$ a self-adjoint linear operator. Then there is a σ-finite measure space $(\Omega, \mathcal{A}, \mu)$, a unitary map operator $U : H \to L^2(\mu)$, and a measurable function $\phi \in M(\mu)$ such that*

$$A = U^{-1} M_\phi U,$$

where $M_\phi : \operatorname{dom} M_\phi \subset M(\mu) \to L^2(\mu)$, $M_\phi f := \phi f$, is the multiplication operator with domain

$$\operatorname{dom} M_\phi = \{\, f \in M(\mu) :\ \phi f \in L^2(\mu) \,\}.$$

Proof. By [Con94, theorem X.4.19], the statement holds for normal operators with a possibly complex-valued function ϕ. However, if A is self adjoint, M_ϕ must be self-adjoint, which is the case if and only if $\phi = \bar{\phi}$. Hence, ϕ must necessarily be real-valued. □

The theorem says that every linear self-adjoint operator is unitarily equivalent to multiplication with a corresponding measurable function ϕ. We can then define positive powers of A as the operators corresponding to positive powers of ϕ. If in addition A is positive, we necessarily have $\phi \geq 0$, so that this construction can be extended to negative powers. We summarize in the following definition (cf. [Tay10, chapter 8]).

Definition 1.5.2. *Let $A : \operatorname{dom} A \subset H \to H$ be a self-adjoint unbounded linear operator.*

i) For all $s \geq 0$, we define

$$A^s := U^{-1} M_{\phi^s} U,$$
$$\operatorname{dom} A^s = U^{-1} \operatorname{dom} M_{\phi^s},$$

where $M_{\phi^s} f = \phi^s f$.

ii) If A is positive, we define in addition

$$A^{-s} := U^{-1} M_{\phi^{-s}} U,$$
$$\operatorname{dom} A^{-s} = U^{-1}(\operatorname{dom} M_{\phi^{-s}}),$$

for all $s > 0$.

In particular, note that if $A \in \mathcal{L}(H)$ is a bounded self-adjoint positive operator, then $A^{-s} : H \to H$ is the Moore-Penrose pseudoinverse of A^s, for all $s \in \mathbb{R}$, because it is self-adjoint and we have

$$A^s A^{-s} A^s = U^{-1} M_{\phi^s} M_{\phi^{-s}} M_{\phi^s} U = A^s,$$
$$A^{-s} A^s A^{-s} = U^{-1} M_{\phi^{-s}} M_{\phi^s} M_{\phi^{-s}} U = A^{-s}.$$

Moreover, the construction of $A^{-1/2}$ for self-adjoint positive operators allows for the definition of operator-weighted norms on Hilbert spaces.

Corollary 1.5.3 (Covariance-weighted norm)**.** *Let $A : H \to H$ be a self-adjoint positive operator. Then, $\operatorname{dom} A^{-s} = \operatorname{ran} A^{-s}$ is a Hilbert space under the inner product*

$$(u, v)_s := (A^s u, A^s v), \qquad u, v \in \operatorname{dom} A^s.$$

In particular, if Γ is a covariance operator, we define

$$\|u\|_\Gamma := \begin{cases} \|\Gamma^{-1/2} u\|, & \text{if } u \in \operatorname{dom} \Gamma^{-1/2}, \\ \infty, & \text{else.} \end{cases} \tag{1.5.1}$$

1.6 Gaussian measures on separable Banach spaces

Gaussian measures play also in infinite-dimensional probability an as central role as they do in finite dimensions. Maybe even more so because one cannot define a Lebesgue measure on an infinite-dimensional space[1], and so the Gaussian measures takes its place as basic reference measure. Recall the definition of Gaussian measures on the real line.

Definition 1.6.1. *We call a probability measure $\gamma \in \mathcal{P}(\mathbb{R})$ a **Gaussian measure on** $\mathbb{R}$ if it is Dirac, i.e. $\gamma = \delta_x$ for some $x \in \mathbb{R}$, or if it is absolutely continuous with respect to the Lebesgue measure λ and there exist $m \in \mathbb{R}$ and $\sigma^2 > 0$ such that its density can be written as*

$$\frac{\gamma(dx)}{\lambda(dx)} = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-m)^2}{2\sigma^2}}.$$

*We call a random variable a **Gaussian random variable** (or normally distributed) if its law is a Gaussian probability measure.*

[1] The only translation-invariant and locally finite measure on an infinite-dimensional Banach space is the trivial measure $\mu = 0$.

In infinite dimensions, this definition through the probability density function is substituted with a functional definition. For a probability measure μ on a topological vector space E, we can consider any continuous linear functional u^* on E as a random variable $u^* : (E, \mathcal{B}(E), \mu) \to (\mathbb{R}, \mathcal{B}(\mathbb{R}))$. In the finite-dimensional case where $E = \mathbb{R}^d$, it can be shown that μ is a multivariate Gaussian distribution if and only if all $u^* : (\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d), \mu) \to (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ are Gaussian random variables (see [Gal16, chapter 1]). In infinite dimensions, this property is used as a definition.

Definition 1.6.2. *i) We call a measure $\gamma \in \mathcal{P}(E)$ on a topological vector space E a **Gaussian measure** if, for all $u^* \in E^*$, the random variable $X(x) := u^*(x)$ on $(E, \mathcal{B}(E), \gamma)$ is Gaussian.*

*ii) We call a random element X of a topological vector space E a **Gaussian random element** if for all $u^* \in E$, $u^*(X)$ is a Gaussian random variable.*

On a Hilbert space, this definition reduces to the following.

Proposition 1.6.3. *Let X be a random element of a Hilbert space H. Then X is Gaussian if and only if for all $h \in H$, (X, h) is a Gaussian random variable.*

Proof. Follows from definition 1.6.2 and the Riesz representation theorem. □

It is immediate that a random element is Gaussian if and only if its law is a Gaussian probability measure, and vice versa. One of the most important regularity theorems for abstract Gaussian measures is *Fernique's theorem*.

Theorem 1.6.4 (Fernique)**.** *Let X be a Gaussian random element of a separable Banach space E. Then, there exists $\lambda > 0$ such that*

$$\mathbb{E}e^{\lambda \|X\|^2} < \infty.$$

In particular, X has finite moments: $\mathbb{E}\|X\|^p < \infty$, for all $p \geq 0$.

Proof. See [PZ14, theorem 2.7]. □

In particular, this result guarantees that every Gaussian measure has a mean and a covariance, and that the characteristic functional is well-defined.

Proposition 1.6.5 (Characteristic functional)**.** *Let X be a Gaussian random element of a separable Banach space E, with mean $\bar{x}$ and covariance Σ_{XX}. Then, its characteristic functional is*

$$\phi_X(u^*) = e^{i\langle \bar{x}, u^* \rangle - \frac{1}{2}\langle \Sigma_{XX} u^*, v^* \rangle}.$$

Proof. [VTC87, IV.2.4, proposition 2.8]. □

Hence, by lemma 1.4.2, any Gaussian measure is uniquely determined by its characteristic functional. This motivates the notation

$$X \sim \mathcal{N}(\bar{x}, \Sigma_{XX}), \tag{1.6.1}$$

if X is a Gaussian random element of a separable Banach space with mean $\bar{x} \in E$ and covariance $\Sigma_{XX} \in \mathcal{L}(E^*; E)$. Similarly, we will use the notation m_γ and Σ_γ to denote the mean and covariance of a Gaussian measure γ. Furthermore, we recover the multivariate Gaussian distribution.

Corollary 1.6.6. *Let $\gamma \in \mathcal{P}(\mathbb{R}^d)$ be a Gaussian measure with mean $m \in \mathbb{R}^d$ and positive definite covariance Σ. Then, γ is absolutely continuous with respect to the d-dimensional Lebesgue measure λ^d, and its density is given by:*

$$\frac{\gamma(dx)}{\lambda^d(dx)} = (2\pi \det(\Sigma))^{-d/2} e^{-\frac{1}{2}(x-m)^\top \Sigma (x-m)}.$$

Proof. [Bog15, proposition 1.2.2] □

Moreover, there is a complete characterization of the class of covariance operators of Gaussian measures on a separable Hilbert space.

Theorem 1.6.7 (Mourier)**.** *Suppose H is a separable Hilbert space. Let $m \in H$, and let $\Sigma \in \mathcal{L}(H)$ be a symmetric positive trace-class operator. Then there exists a Gaussian element X of H with $\mathbb{E}X = m$ and $\operatorname{Cov}(X) = \Sigma$.*

Proof. See [PZ14] or [VTC87, theorem IV.2.4]. □

Motivated by this result, we say that a linear operator $C \in \mathcal{L}(H)$ is a **covariance operator** if it is symmetric, positive, and in the trace-class. There is, however, no corresponding generalization of Mourier's theorem to Banach spaces. It is not clear whether every positive symmetric nuclear operator on a Banach space is the covariance operator of a Gaussian random element. However, in our case this problem will not be relevant since we will always know beforehand that the involved covariance operators correspond to Gaussian distributions.

We now present a list of useful properties of Gaussian measures that will be vital for the filtering theory we develop in chapter 2. At first, we note that as in the finite-dimensional case, independence of Gaussian random elements is equivalent to uncorrelatedness.

Lemma 1.6.8 (Independence)**.** *Let X and Y be Gaussian random elements of separable Banach spaces H_1 and H_2. Then, X and Y are independent if and only if $\operatorname{Cov}[X, Y] = 0$.*

Proof. Follows from proposition 1.6.5 and the fact that two random elements X and Y are independent if and only if

$$\phi_{X,Y} = \phi_X \otimes \phi_Y,$$

a statement that holds in great generality (e.g. for Radon measures on normed spaces [VTC87, IV.2.2, proposition 2.2]). □

One of the most useful properties of Gaussian distributions is their compatibility with linear transformation, a property that generalizes to arbitrary Hilbert spaces.

Lemma 1.6.9 (Linear combinations of Gaussian random elements)**.** *Suppose $X \sim \mathcal{N}(\bar{x}, \Sigma_{XX})$ and $Y \sim \mathcal{N}(\bar{y}, \Sigma_{YY})$ are independent Gaussian random elements of separable Banach spaces E_1 and E_2. Let $A \in \mathcal{L}(E_1; E_3), B \in \mathcal{L}(E_2; E_3)$ and $c \in E_3$, where E_3 is a separable Banach space. Then,*

$$AX + BY + c \sim \mathcal{N}(A\bar{x} + B\bar{y} + c, A\Sigma_{XX}A^* + B\Sigma_{YY}B^*).$$

Proof. This follows again from the characteristic functional (proposition 1.6.5). Letting without loss of generality $c = 0$, it follows by independence:

$$\begin{aligned}\phi_{AX+BY}(x) &= \phi_X(A^*x)\phi_Y(B^*x) \\ &= e^{i\langle \bar{x}, A^*x\rangle - \frac{1}{2}\langle \Sigma_{XX}A^*x, A^*x\rangle} e^{i\langle \bar{y}, B^*x\rangle - \frac{1}{2}\langle \Sigma_{YY}B^*x, B^*x\rangle} \\ &= e^{i\langle A\bar{x} + B\bar{y}, x\rangle - \frac{1}{2}\langle A\Sigma_{XX}A^* + B\Sigma_{YY}B^*x, x\rangle},\end{aligned}$$

and thus the statement follows by lemma 1.4.2. □

Lemma 1.6.9 leads to the useful property that many probabilistic operations on Gaussian measures, like computing the joint distribution, marginalization over a nuisance variable, and computing the conditional distribution with respect to another Gaussian random element, reduce to simple operator formulas.

Lemma 1.6.10 (Joint normal distribution)**.** *Suppose $X \sim \mathcal{N}(\bar{x}, \Sigma_{XX})$ and $Y \mid X \sim \mathcal{N}(MX + b, \Gamma)$ are Gaussian random elements of separable Banach spaces E_1 and E_2, respectively, where $M \in \mathcal{L}(E_1; E_2)$, and $b \in E_2$. Then*

$$\begin{bmatrix} X \\ Y \end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix} \bar{x} \\ M\bar{x} + b \end{bmatrix}, \begin{bmatrix} \Sigma_{XX} & \Sigma_{XX}M^* \\ M\Sigma_{XX} & M\Sigma_X M^* + \Gamma \end{bmatrix}\right).$$

Proof. First, note that Y can be written as $Y = MX + \eta$, with a random element $\eta \sim \mathcal{N}(b, \Gamma)$ of E_2 that is independent of X. $[X, \eta]$ is then a Gaussian random element of $E_1 \times E_2$. By independence, we have

$$\begin{bmatrix} X \\ \eta \end{bmatrix} \sim \mathcal{N}(\begin{bmatrix} \bar{x} \\ b \end{bmatrix}, \begin{bmatrix} \Sigma_{XX} & 0 \\ 0 & \Gamma \end{bmatrix}).$$

Thus, using lemma 1.6.9, we arrive at

$$\begin{bmatrix} X \\ Y \end{bmatrix} = \begin{bmatrix} \mathrm{Id} & 0 \\ M & \mathrm{Id} \end{bmatrix} \begin{bmatrix} X \\ \eta \end{bmatrix} \sim \mathcal{N}(\begin{bmatrix} \bar{x} \\ M\bar{x} + b \end{bmatrix}, \begin{bmatrix} \Sigma_{XX} & \Sigma_{XX}M^* \\ M\Sigma_{XX} & M\Sigma_{XX}M^* + \Gamma \end{bmatrix}).$$

□

Lemma 1.6.11 (Marginalization over nuisance variable)**.** *Suppose $X \sim \mathcal{N}(\bar{x}, \Sigma_{XX})$ and $Y \mid X \sim \mathcal{N}(MX + b, \Gamma)$ are Gaussian random elements of separable Banach spaces E_1 and E_2, respectively, where $M \in \mathcal{L}(E_1; E_2)$, and $b \in E_2$. Then*

$$Y \sim \mathcal{N}(M\bar{x} + b, M\Sigma_{XX}M^* + \Gamma).$$

Proof. Applying lemma 1.6.10, we have

$$\begin{bmatrix} X \\ Y \end{bmatrix} \sim \mathcal{N}(\begin{bmatrix} \bar{x} \\ M\bar{x} + b \end{bmatrix}, \begin{bmatrix} \Sigma_{XX} & \Sigma_{XX}M^* \\ M\Sigma_{XX} & M\Sigma_X M^* + \Gamma \end{bmatrix}).$$

Taking the projection $\Pi_2 : E_1 \times E_2 \to E_2$, it follows from lemma 1.6.9 that

$$Y = \Pi_2(X, Y) \sim \mathcal{N}(M\bar{x} + b, M\Sigma_{XX}M^* + \Gamma).$$

□

Before we can describe the conditional distribution of a Gaussian random elements with respect to another Gaussian random element, we need the following elementary lemma.

Lemma 1.6.12. *Let $X, Y_1, Y_2, \ldots, Y_m$ be random variables with finite second moments, where $Y_1, \ldots, Y_n$ are independent and $\mathrm{Cov}(Y_i) > 0$ for all $i \in [m]$. Then,*

$$\mathbb{E}[X|Y_1, \ldots, Y_n] = \sum_{=1}^{m} \frac{\mathbb{E}[XY_j]}{\mathrm{Cov}[Y_j^2]} Y_i.$$

Proof. Note that the conditional expectation $\mathbb{E}[\cdot|Y_1, \ldots, Y_n]$ is the orthogonal projection in $L^2(\mathbb{P})$ on the space spanned by $\{Y_1, \ldots, Y_n\} \subset L^2(\mathbb{P})$. Hence, there must exist $x_1, \ldots, x_n \in \mathbb{R}$ such that

$$\mathbb{E}[X|Y_1, \ldots, Y_n] = \sum_{i=1}^{n} x^i Y_i.$$

Now, independence of Y_n implies L^2-orthogonality. Hence, we have

$$x^i = \frac{(X, Y_i)_{L^2(\mathbb{P})}}{\|Y\|^2} = \frac{\mathbb{E}[XY_i]}{\mathbb{E}[Y_i^2]}.$$

□

This lemma is directly applied in the proof of the next assertion, which is fundamental for the rest of this thesis.

Lemma 1.6.13 (Conditioning of Gaussian measures)**.** *Suppose X is a random element of a separable Banach space E and Y is an m-dimensional random vector, and suppose that X and Y are jointly Gaussian distributed with*

$$\begin{bmatrix} X \\ Y \end{bmatrix} \sim \mathcal{N}(\begin{bmatrix} \bar{x} \\ \bar{y} \end{bmatrix}, \begin{bmatrix} \Sigma_{XX} & \Sigma_{XY} \\ \Sigma_{YX} & \Sigma_{YY} \end{bmatrix}).$$

Then, the regular conditional distribution of X, given Y, is

$$X \mid Y \sim \mathcal{N}(\bar{x} + \Sigma_{XY}\Sigma_{YY}^{-1}(Y - \bar{y}), \Sigma_{XX} - \Sigma_{XY}\Sigma_{YY}^{-1}\Sigma_{YX}). \tag{1.6.2}$$

Proof. The fact that $(X|Y)$ is Gaussian distributed is easy to see. Let $x^* \in E^*$. Then, $\langle X, x^* \rangle$ and Y are finite-dimensional random vectors with joint Gaussian distribution,

$$\begin{bmatrix} \langle X, x^* \rangle \\ Y \end{bmatrix} \sim \mathcal{N}(\begin{bmatrix} \langle \bar{x}, x^* \rangle \\ \bar{y} \end{bmatrix}, \begin{bmatrix} \langle \Sigma_{XX} x^*, x^* \rangle & x^* \Sigma_{XY} \\ \Sigma_{YX} x^* & \Sigma_{YY} \end{bmatrix}).$$

Hence, $(\langle X, x^* \rangle | Y)$ has a conditional Gaussian distribution, and since this holds true for all $x^* \in E^*$, we have that $(X|Y)$ is conditional Gaussian distributed. Thus, it remains to show that the conditional expectation and covariance are as in (1.6.2).

We assume without loss of generality $\mathbb{E}X = 0$ and $\mathbb{E}Y = 0$. Excluding the degenerate case $Y = 0$, let $\lambda_1 \geq \lambda_l > \lambda_{l+1} = \cdots = \lambda_m = 0$ be the eigenvalues of $\Sigma_Y Y$ with corresponding eigenvectors $e_1, \ldots, e_m$. Let

$$Y^i := (Y, e_i),$$

so that $\mathrm{Cov}(Y^i) = \mathbb{E}(Y^i)^2 = \lambda_i$, and $Y_1, ..., Y_l$ are uncorrelated, and thus by lemma 1.6.8 independent. Then, for any $x^* \in E^*$, we have

$$\begin{aligned} \mathbb{E}[X|Y] &= \mathbb{E}[X|Y^1, \ldots Y^m] \\ &= \mathbb{E}[X|Y^1, \ldots Y^l] \\ &= \sum_{i=1}^{l} \frac{\mathbb{E}[\langle X, x^* \rangle Y^i]}{\lambda^i} Y^i, \end{aligned}$$

where the last equality is due to lemma 1.6.12. From the definition of the conditional covariance, we then obtain

$$\begin{aligned} \mathbb{E}[\langle X, x^* \rangle | Y^1, \ldots, Y^m] &= \sum_{i=1}^{l} \frac{\langle \Sigma_{XY} e_i, x^* \rangle}{\lambda^i} Y^i \\ &= \sum_{i=1}^{l} \langle \Sigma_{XY} \Sigma_{YY}^{-1} e_i Y^i, x^* \rangle \\ &= \langle \sum_{i=1}^{l} \Sigma_{XY} \Sigma_{YY}^{-1} e_i Y^i, x^* \rangle. \end{aligned} \tag{1.6.3}$$

Now, since the conditional expectation is linear, we have

$$\mathbb{E}[\langle X, x^* \rangle | Y^1, \ldots, Y^m] = \langle \mathbb{E}[X|Y^1, \ldots, Y^m], x^* \rangle.$$

Therefore, since (1.6.3) holds for all $x^* \in E^*$, we obtain

$$\begin{aligned} \mathbb{E}[X|Y^1, \ldots, Y^m] &= \sum_{i=1}^{l} \Sigma_{XY} \Sigma_{YY}^{-1} e_i Y^i \\ &= \Sigma_{XY} \Sigma_{YY}^{-1} \sum_{i=1}^{l} Y^i e_i \\ &= \Sigma_{XY} \Sigma_{YY}^{-1} Y, \end{aligned} \tag{1.6.4}$$

where the reader should keep in mind that Σ_{YY}^{-1} denotes the Moore-Penrose pseudoinverse if $l < m$. For the covariance, we define

$$\eta := X - \mathbb{E}[X|Y].$$

Then, by proposition 1.3.7, η is independent of Y, and so by proposition 1.4.5, we have

$$\Sigma_{XX} = \mathrm{Cov}(\mathbb{E}[X|Y]) + \mathrm{Cov}(\eta).$$

Using (1.6.4) and proposition 1.4.5, we obtain

$$\begin{aligned}\mathrm{Cov}(\eta) &= \Sigma_{XX} - \Sigma_{XY}\Sigma_{YY}^{-1}\Sigma_{YY}\Sigma_{YY}^{-1}\Sigma_{XY}^{*} \\ &= \Sigma_{XX} - \Sigma_{XY}\Sigma_{YY}^{-1}\Sigma_{YX}.\end{aligned}$$

This finishes the proof. □

Lemma 1.6.13 is important because it yields the following version of Bayes' theorem that provides the fundamental relation on which our later derivations of the Kalman filter will be based on.

Corollary 1.6.14 (Bayes' theorem for Gaussian random elements)**.** *Suppose X is a Gaussian random element of a separable Banach space E, with $X \sim \mathcal{N}(\bar{x}, \Sigma_{XX})$, and Y is an m-dimensional Gaussian random vector with $Y \mid X \sim \mathcal{N}(MX + b, \Gamma)$, where $M \in \mathcal{L}(E; \mathbb{R}^m)$ and $b \in \mathbb{R}^m$). Then:*

$$X \mid Y \sim \mathcal{N}(\bar{x} + K(Y - M\bar{x} - b), \Sigma_{XX} - KM\Sigma_{XX}),$$

where $K := \Sigma_{XX}M^*(M\Sigma_{XX}M^* + \Gamma)^{-1}$.

Proof. First, apply lemma 1.6.10 to obtain the joint distribution of $\begin{bmatrix} X \\ Y \end{bmatrix}$. Then, we obtain the conditional distribution of X, given Y, through use of lemma 1.6.13. □

1.7 Stochastic calculus in Hilbert spaces

Continuous-time filtering problems are defined through stochastic differential equations. Recall that a stochastic differential equation (SDE) of the form

$$dX(t) = f(X(t), t)dt + G(X(t), t)dW(t), \tag{1.7.1a}$$
$$X(0) = X_0 \tag{1.7.1b}$$

is just a formal notation for a corresponding stochastic integral equation, in this case (see [Kal02, chapter 21] or [Jaz70])

$$X(t) = X_0 + \int_0^t f(X(s), s)ds + \int_0^t G(X(s), s)dW(s). \tag{1.7.2}$$

Equation (1.7.1) is called an equation with **additive noise** if G is independent of X, and else an equation with **multiplicative noise**. In this thesis, we will restrict ourselves on the former case.

Stochastic partial differential equations (SPDEs) can then be formulated as stochastic differential equations in infinite-dimensional spaces, e.g. where $X(t)$ is a random element of a function space, and where f is a partial differential operator with respect to space. We will restrict ourselves to the case where X takes values in a separable Hilbert space, and not deal in this thesis with stochastic differential equations in Banach spaces. The reason for this choice is that, while there exist constructions of stochastic differential equations on Banach spaces (see section 1.8, the Hilbert space theory is at the moment much more developed, and suffices already for a lot of applications.

First of all, we need to clarify how to construct the Hilbert-space-valued noise process W in (1.7.1). In this thesis, W will always denote Hilbert-space-valued Wiener process defined as follows:

Definition 1.7.1 (Hilbert-space-valued Wiener process)**.** *Let H be a Hilbert space and Q a symmetric positive trace-class (see definition 1.4.6) operator. We call an H-valued random process W on $\mathbb{R}^+$ a* ***Q-Wiener process*** *if*

a) W has paths in $C^0(\mathbb{R}^+; H)$ (cf. definition 1.3.2),

b) $W(0) = 0$ almost surely,

c) W has independent increments, i.e. for all $t \geq 0$ and $\epsilon > 0$, $W(t+\epsilon) - W(t)$ is independent of $(W(s))_{s \leq t}$, and

d) $W(t) - W(s) \sim \mathcal{N}(0, |t-s|Q)$, for all $t, s \geq 0$.

On separable Hilbert spaces we can guarantee existence of Q-Wiener processes for arbitrary covariance operators.

Proposition 1.7.2 (Existence of Q-Wiener processes)**.** *Let H be a separable Hilbert space and Q a symmetric positive trace-class operator. Then, there exists an H-valued Q-Wiener process.*

Proof. [PZ14, proposition 4.4]. □

Note that in the case $H = \mathbb{R}$ and $Q = 1$, the process W coincides with the standard Brownian motion (cf. [Kal02, chapter 13] or [Gal16, chapte 2]). However, there strictly does not exist a "standard" Wiener process on infinite-dimensional Hilbert spaces, since in that case Id is not in the trace-class, because it is not compact (see proposition 1.4.9). So, there cannot exist a Wiener process with covariance Id. A remedy for this is the concept of a generalized Wiener process.

Definition 1.7.3 (Generalized Wiener process)**.** *Let H be a Hilbert space. We call a random map $W : H \to C^0(\mathbb{R}^+), h \mapsto W_h$ a* ***generalized Wiener process*** *if for all $h \in H$, W_h is a real-valued Wiener process with*

$$\lim_{h_n \to h} \mathbb{E}|W_h(t) - W_{h_n}(t)|^2 = 0, \qquad \forall t \geq 0.$$

We call W a cylindrical standard Wiener process if

$$\mathbb{E}[W_h(t)W_g(s)] = \min(s,t)(h,g), \qquad \forall t, s \geq 0, h, g \in H.$$

We also have a corresponding existence theorem.

Proposition 1.7.4 (Existence of generalized Wiener processes)**.** *Let H be a separable Hilbert space and $Q \in \mathcal{L}(H)$ a self-adjoint positive operator. Then, there exists a generalized H-valued Wiener process W with*

$$\mathbb{E}[W_h(t)W_g(s)] = \min(s,t)(Qh,g), \qquad \forall t, s \geq 0, h, g \in H.$$

Proof. See [PZ14, section 4.1.2]. □

In this thesis, however, we will solely deal with Wiener processes in the strict sense. However, extensions of our results to generalized and even spatially homogeneous Wiener processes (see [PZ14, section 4.1] should be possible mutas mutandis.

Having defined the noice process in (1.7.1), it is necessary to extend the stochastic integral in (1.7.2) to the Hilbert space setting. In the following, we will only list some useful results from the theory of Hilbert space stochastic calculus and stochastic partial differential equations that will be used in the formulation of the infinite-dimensional Kalman-Bucy filter (section 2.4). For more details, see [PZ14], [Cho07] or [Hai09].

Lemma 1.7.5 (Properties of the stochastic integral)**.** *Let U and H be separable Hilbert spaces and let W be a U-valued Q-Wiener process on $[0,T]$. We define the inner product*

$$(u,v)_0 := (Q^{-1/2}u, Q^{-1/2}v), \qquad u, v \in Q^{1/2}U,$$

under which $Q^{1/2}U$ is a Hilbert space. Let $G \in M([0,T]; \mathcal{L}_2(Q^{1/2}U; H))$ with

$$\int_0^T \|G(t)\|_2^2 < \infty$$

almost surely. Then:

i) $\mathbb{E}\int_0^t G(s)dW(s) = 0$*, for all* $t \in [0,T]$*.*

ii) $\mathbb{E}\|\int_0^t G(s)dW(s)\|^2 = \int_0^t \|G(s)\|_2^2 ds$*, for all* $t \in [0,T]$*.*

iii) The stochastic integral is linear: Let $A \in \mathcal{L}(H)$*. Then*

$$A\int_0^T G(s)dW(s) = \int_0^T AG(s)dW(s),$$

almost surely.

iv) If X *and* W *are independent, so are* X *and* $\int_0^t G(s)dW(s)$*, for all* $t \in [0,T]$*.*

Proof. See [PZ14, chapter 4.3] □

Definition 1.7.6. *Let* $(\Omega, \mathcal{A}, \mathbb{P})$ *be a probability space.*

i) We call a collection $\mathcal{F} = (\mathcal{F})_{t\in[0,T]}$ *of* σ*-algebras a* ***filtration*** *if*

$$\mathcal{F}_s \subset \mathcal{F}_t \subset \mathcal{A}$$

for all $0 \le s \le t \le T$*. A random process* X *on* $[0,T]$ *is called* ***adapted*** *if* $X(t)$ *is* $\mathcal{F}_t$*-measurable for all* $t \in [0,T]$*.*

ii) We define the ***predictable*** σ***-algebra*** $\mathcal{P}_T$ *to be the* σ*-algebra generated by sets of the form* $(s,t) \times F$*, where* $F \in \mathcal{F}_s$*, i.e.*

$$\mathcal{P}_T := \sigma\left(\{(s,t] \times F : F \in \mathcal{F}_s\} \cup \{\{0\} \times F : F \in \mathcal{F}_0\}\right).$$

A random process X *on* $[0,T]$ *is called* ***predictable*** *if the map* $(t,\omega) \mapsto X(t)(\omega)$ *is* $\mathcal{P}_T$*-measurable.*

iii) We say a Wiener process W *is a* ***Wiener process with respect to the filtration*** $\mathcal{F}$ *if it is adapted and if* $W(t+h) - W(t)$ *is independent of* $\mathcal{F}_t$*, for all* $h \ge 0$ *and* $t \in [0,T]$*.*

iv) We call a filtration a ***normal filtration*** *or say that it satisfies the* usual conditions *if it is* ***complete****, i.e. have* $A \in \mathcal{F}_0$ *for all* $A \in \mathcal{A}$ *with* $\mathbb{P}(A) = 0$*, and it is* ***right-continuous****, i.e.*

$$F_t = \bigcap_{h>0} \mathcal{F}_{t+h}.$$

The notion of a filtration is important for modeling how much information we have availabe at a certain point in time. For example, it would not be desirable that the solution of the SPDE (1.7.1) at time t depends on the noise at time $t+\epsilon$ for some $\epsilon > 0$. This motivates the notion of predictability as defined above. It is more intuitive to think of the σ-algebra $\mathcal{F}_t$ as the knowledge of a set of random elements. These are precisely those random elements that are $\mathcal{F}_t$-measurable. Similarly, for a process to be adapted to a filtration means that at time t, we have knowledge past trajectory of the process and its present value, simply all the information that would be available to use from observing the process in the time interval $[0,t]$.

We can also define a filtration in terms of a given random process. For example, a Wiener process W generates a filtration by letting $\mathcal{F}_t^W$ represent knowledge of the trajectory of W up to and including time t. Formally, this means

$$\mathcal{F}_t^W := \sigma(W(s) : s \le t), \qquad t \in [0,T].$$

For this thesis, it will suffice to always take the filtration

$$\mathcal{G}_t := \sigma(X_0, W(s) : s \le t),$$

so that X_0 is always included.

It is obvious from definitions 1.7.1 and 1.7.6 that W is a Wiener process with respect to $\mathcal{G}$. In order to obtain a normal filtration, we have to make the *usual augmentations*.

Lemma 1.7.7. *For every filtration $\mathcal{F}$ there exists a smallest complete and right-continuous extension, called the* ***augmented filtration****.*

Proof. See [Kal02, lemma 7.8]. □

Hence, we can assume without loss of generality that $\mathcal{G}$ is normal.

Next, we introduce the classical concept of a strong solution. It will usually be necessary for the operator G in to be a Hilbert-Schmidt operator. However

Definition 1.7.8 (Strong solution)**.** *Let U and H be separable Hilbert spaces and $T > 0$. Suppose we are given an U-valued Q-Wiener process W, a map $F \in M(H \times [0,T]; H)$ and an operator $G \in L^2([0,T]; \mathcal{L}_2(U;H))$. Suppose $A : \operatorname{dom} A \subset H \to H$ is an unbounded linear operator. Finally, let X_0 be a random element of H. Then, we call an H-valued predictable (with respect to $\mathcal{G}$) process X on $[0,T]$ a* ***strong solution*** *to*

$$\begin{aligned} &X(t) = (AX(t) + F(X(t),t))\,dt + G(t)dW(t), \qquad t \in [0,T],\\ &X(0) = X_0, \end{aligned}$$

if

$$\begin{aligned} &X(t) \in \operatorname{dom} A,\\ &\operatorname{tr} G(X(t),t) < \infty, \qquad\qquad \int_0^T \|AX(s) + F(X(s),s)ds\|ds < \infty, \end{aligned}$$

and

$$X(t) = X_0 + \int_0^t (AX(s) + F(X(s),s))\,ds + \int_0^t G(X(s),s)ds$$

almost surely, for all $t \in [0,T]$.

Similarly to the theory of partial differential equations, the concept of a strong solution is usually too restrictive for a lot of examples. Hence, in analogy to the deterministic case, one introduces the concept of a *weak solution*. As a preparation, recall the definition of a strongly continuous semigroup.

Definition 1.7.9. *A family $(S(t))_{t\geq 0}$ of bounded linear operators on a Banach space E is called a* ***semigroup*** *if*

$$\begin{aligned} &S(t+s) = S(t)S(s), \qquad \forall t,s \geq 0,\\ &S(0) = \mathrm{Id}. \end{aligned}$$

It is called ***strongly continuous****, if for every $x \in E$ the map $t \mapsto S(t)x$ is continuous.*

Definition 1.7.10 (Weak solution)**.** *Let U and H be separable Hilbert spaces and $T > 0$. Suppose we are given an U-valued Q-Wiener process W, a map $F \in M(H \times [0,T]; H)$ and an operator $G \in L^2([0,T]; \mathcal{L}_2(U;H))$. Suppose $A : \operatorname{dom} A \subset H \to H$ is an unbounded linear operator that generates a strongly continuous semigroup. Finally, let X_0 be a random element of H. Then, we call an H-valued predictable random process X with paths in $L^1([0,T]; H)$ a* ***weak solution*** *to*

$$\begin{aligned} &dX(t) = (AX(t) + F(X(t),t))\,dt + G(t)dW(t), \qquad t \in [0,T],\\ &X(0) = X_0, \end{aligned}$$

if

$$\begin{aligned} \int_0^T \|F(X(t),t)\|dt < \infty \quad &\text{almost surely,}\\ \int_0^T \|G(t)\|_{\mathcal{L}_2(Q^{1/2}H;H)}dt < \infty \quad &\text{almost surely,} \end{aligned}$$

and

$$(X(t), w) = (X_0, w) + \int_0^t (X(s), A^* w) + (F(X(s), s), w)ds + (\int_0^t G(s)dW(s), w),$$

almost surely, for all $t \in [0, T]$ *and all* $w \in \operatorname{dom} A^*$.

Note that there is also the slightly different concept of a mild solution (see [PZ14]) of a SPDE. However, in the linear case we do not have to worry about the distinction (see [Hai09]).

Similarly to the semigroup theory of parabolic PDEs [EN06], we have the following variation of constants formula for linear SPDEs.

Theorem 1.7.11 (Weak existence and uniqueness for linear problems)**.** *Let* U *and* H *be separable Hilbert spaces and* $T > 0$*. Consider the SPDE*

$$dX(t) = AX(t)dt + f(t)dt + G(t)dW(t), \qquad t \in [0, T], \tag{1.7.3a}$$

$$X(0) = X_0 \tag{1.7.3b}$$

Suppose that

- W *is an* U*-valued* Q*-Wiener process and* $G \in L^\infty([0, T]; \mathcal{L}_2(Q^{1/2}U; H))$.
- X_0 *is a random element of* H.
- $A : \operatorname{dom} A \subset H \to H$ *is linear and generates a strongly continuous semigroup* S.
- $f \in L^1([0, T]; H)$.

Then, there exists exactly one weak solution to equation (1.7.3)*, given by*

$$X(t) = S(t)X_0 + \int_0^t S(t-s)f(s)ds + \int_0^t S(t-s)G(s)dW(s), \qquad t \in [0, T].$$

Proof. See [PZ14, theorem 7.1]. □

1.8 Comments and further references

As we will see in the next chapter (see theorem 2.2.2), the key ingredient for deriving the Kalman filter is corollary 1.6.14, which is in turn an immediate consequence of lemma 1.6.13 (conditioning). For both results, we assumed the measurement Y to be finite-dimensional. However, both results can be generalized to the case where the measurement variable Y is also a Gaussian random element of a separable Banach space. This was shown for Hilbert spaces in [Man84], and then for Banach spaces in [Lus96]. The key ingredient is the following factorization lemma (see [Vak94] and [VTC87, section III.1.2]), which allows to apply Hilbert-space techniques to Gaussian measures on more general spaces.

Lemma 1.8.1 (Canonical factorization)**.** *Let* $X : (\Omega, \mathcal{A}, \mathbb{P}) \to (E, \mathcal{B}(E))$ *be a Gaussian random element of a separable Banach space* E*. Then there exists a separable Hilbert space* H_X *and an operator* $A_X \in \mathcal{L}(E; H_X)$ *with dense image, such that*

$$\Sigma_{XX} = A_X A_X^*. \tag{1.8.1}$$

In that case, $\operatorname{ran} A_X^* \subset E$*, and* A_X *is unique up to unitary equivalence. The space* H_X *is called the* ***Cameron-Martin space*** *of* X.

Proof. Sketch. One defines the map $A_X : E \to L^2(\mathbb{P})$,

$$A_X u^*(\omega) := \langle X(\omega) - \mathbb{E}X, u^* \rangle, \qquad \omega \in \Omega,$$

and then defines H_X to be the closure of $\operatorname{ran} A_X$ under the $L^2(\mathbb{P})$-norm. □

The operator A_X is a so-called γ-radonifying operator (see [Lus96]). An operator $A : E \to H$ into a separable Hilbert space is called γ-radonifying if AA^* is the covariance of a Gaussian random element. If E is Hilbert, it follows from theorem 1.6.7 (Mourier) that A is γ-radonifying if and only if it is Hilbert-Schmidt.

A major complication in the case of infinite-dimensional measurements is that since covariance matrices are compact, Σ_{YY}^{-1} will always be unbounded in the infinite-dimensional case. The operator K in corollary 1.6.14 will then only be bounded on a set of probability zero (see [LPS89]). It will not even be linear anymore, but only linear on a set of probability 1. In this thesis, we circumvented these problems by assuming finite-dimensional measurements. Note that we did not have to rely on regularity properties of the distribution of the state X in the proof of lemma 1.6.13. This suggests that the result might hold in even greater generality, for example if X is a random element of an arbitrary locally convex space. The finite dimensionality of the measurements has a naturally regularizing effect for the inference problem.

Regarding section 1.7, there exists a theory of stochastic integration on Banach spaces extending some of the results presented there. The objects of integration are in that case opertor-valued measurable functions $\Phi : [0, T] \to \mathcal{L}(H; E)$ that satisfy

$$\int_0^T \|\Phi(t)^* u^*\|_H^2 dt < \infty,$$

for all $u^* \in E^*$, and for which the map $I_\phi \in \mathcal{L}(L^2([0, T]; H); E)$,

$$\langle I_\Phi f, u^* \rangle := \int_0^T (\Phi(t)^* u^*, f(t))_H dt, \qquad f \in L^2([0, T]; H),$$

is a well-defined, γ-radonifying operator. A corresponding formulation of optimal linear filters in Banach spaces seems to be missing from the literature.

Chapter 2

Stochastic filtering

In this chapter, we present some parts of the theory of stochastic filtering on infinite-dimensional spaces. We present some general aspects in section 2.1. Then, we derive the optimal filter for the linear Gaussian filtering problem in discrete time — the Kalman filter — in section 2.2. In section 2.3 we then rederive the Kalman filter as the solution to the linear filtering problem with non-Gaussian noise. Then, in section 2.4 introduce the continuous-time version of the Kalman filter, the Kalman-Bucy filter. At the end of this chapter we turn in section 2.5 to the smoothing problem and derive the linear fixed-point smoother and the Rauch-Tung-Striebel smoother.

2.1 The general filtering problem

In general, the filtering problem can be subsumed as the problem of inferring the state of a dynamical system from noisy measurements. Such a system can be deterministic or stochastic. Since the latter approach is more general, we will assume it in the following. In any case, the problem is governed by two models. First, the **state model** which describes the time-evolution of the state. In discrete-time, we will study the Markov state space model with additive noise

$$X_k = f_k(X_{k-1}) + W_k, \qquad k \in \mathbb{N},$$

where X_k is a random element of a Banach space E, $f_1, f_2, ... : E \to E$ each are measurable functions of the state, and $W_1, W_2, ...$ are random elements of E, representing the state noise, or intrinsic uncertainty, of the system. In the following, we will always assume that $W_1, W_2, ...$ are mutually independent.

Second, it is often not possible to measure the whole state of a system. Instead, one measures an observable Y_k which is a (possibly random) function of the state X_k. We will assume that Y_k only depends on the current state X_k, not on the whole trajectory, and that the randomness can be represented by additive, independent noise vectors $V_1, V_2, ...$ (the *measurement noise*). Furthermore, we will assume finite-dimensional measurements. Thus, we arrive at the discrete **measurement model**

$$Y_k = m_k(X_k) + V_k, \qquad k \in \mathbb{N},$$

where $m_k \in M(E; \mathbb{R}^m)$, for $k \in \mathbb{N}$. In continuous time and given a separable Hilbert space H, the corresponding state model can be written as a stochastic differential equation

$$dX(t) = f(X(t), t) + G(X(t), t)dW(t), \quad t \geq 0, \tag{2.1.1}$$
$$X(0) = X_0, \tag{2.1.2}$$

where X is an H-valued stochastic process on $\mathbb{R}^+$,$f \in M(H \times \mathbb{R}^+; H)$ and $G \in M(H \times \mathbb{R}^+; \mathrm{Lin}(U; H))$ denote deterministic functions, while W is an U-valued stochastic process on $\mathbb{R}^+$, where U is some separable Hilbert space. Classically, this situation is only treated in the finite-dimensional special case

where X is assumed to be a strong solution of equation (2.1.1) (see for example [Jaz70]). Since we will be interested in applications to (stochastic) partial differential equations, we will aim to derive results for the case where X is a weak solution (see definition 1.7.10) that takes values in a Hilbert space.

2.2 The Kalman filter

We will start with a fairly elementary situation, the linear discrete-time filtering problem with Gaussian noice. Since we are in discrete time, we can still stay in the more general Banach space setting. The fundamental algorithm is the famous **Kalman filter**, an on-line method that yields a complete solution of the Gaussian linear discrete filtering problem that is defined as follows.

Problem 2.2.1 (Gaussian linear discrete filtering)**.** *Let E be a separable Banach space. Consider the linear state-space and measurement models*

$$X_k = F_k X_{k-1} + W_k, \tag{2.2.1a}$$
$$Y_k = M_k X_k + V_k, \tag{2.2.1b}$$

where for all $k \in \mathbb{N}$ we assume $F_k \in \mathcal{L}(E)$, $M_k \in \mathrm{Lin}(E; \mathbb{R}^m)$ and Gaussian noise $W_k \sim \mathcal{N}(0, Q_k)$ in E and $V_k \sim \mathcal{N}(0, R_k)$ in $\mathbb{R}^m$, and an initial distribution $X_0 \sim \mathcal{N}(x_0, P_0)$. We also assume that $X_0, W_1, V_1, W_2, V_2, \dots$ are all independent.

Given a sequence of measurements $Y_1 = y_1, Y_2 = y_2, ...$, the filtering problem is to determine the conditional probability distribution

$$\mathbb{P}^{X_k | \boldsymbol{Y}_k = \boldsymbol{y}_k}$$

of X_k, given $\boldsymbol{Y}_k = \boldsymbol{y}_k$, for all $k \in \mathbb{N}$. Here

$$\boldsymbol{Y}_k := \begin{bmatrix} Y_1 & \dots & Y_k \end{bmatrix},$$
$$\boldsymbol{y}_k := \begin{bmatrix} y_1 & \dots & y_k \end{bmatrix} \in \mathbb{R}^{m \times k},$$

denote the collections of all measurements and their realizations, respectively, up to and including time k.

Given the setup (2.2.1), we can formulate all inference on the state X as the problem of determinining the conditional distributions

$$\mathbb{P}^{X_n | \boldsymbol{Y}_k = \boldsymbol{y}_k}, \qquad n \in \mathbb{N}_0, k \in \mathbb{N}.$$

Usually, one distinguishes between the case $n > k$ (called **prediction**), the case $n = k$ (called **filtering**) and the case $n < k$ (called **smoothing**).

For problem 2.2.1, the Kalman filter gives a recursive algorithm to calculate the conditional expectations $\mathbb{E}[X_k | \boldsymbol{Y}_k]$ and the conditional covariances $\mathrm{Cov}(X_k | \boldsymbol{Y}_k)$. Since $X_1, X_2, ...$ are Gaussian by lemma 1.6.9 and a Gaussian measure is uniquely determined by its mean and covariance (cf. section 1.6), this is all the information one needs to construct the probability distributions $\mathbb{P}^{X_n | \boldsymbol{Y}_k = \boldsymbol{y}_k}$. This is what we mean when we say that *the Kalman-filter provides a complete solution to the linear discrete filtering problem 2.2.1.*

The derivation of the Kalman filter equations in the setting of problem 2.2.1 is based on the fundamental facts about Gaussian measures in Banach spaces, presented in section 1.4.

We define

$$\boxed{x_{n|k} := \mathbb{E}[X_n | \boldsymbol{Y}_k = \boldsymbol{y}_k],} \qquad n \in \mathbb{N}_0, k \in \mathbb{N}. \tag{2.2.2}$$

We call the special case $x_{k|k}$ the **analysis estimate** and $x_{k+1|k}$ the **forecast estimate** [1] to denote a

[1] We will call a random element $\hat{X} = f(Y)$ that estimates X, given a random measurement Y, an **estimator** at time k, while we call the value $\hat{x} = f(y)$ of the estimator given a particular realization of the measurement, $Y = y$, an **estimate**.

random element. We denote the corresponding covariance matrix by

$$\boxed{P_{n|k} := \mathrm{Cov}(X_n|\boldsymbol{Y}_k = \boldsymbol{y}_k)}, \qquad n \in \mathbb{N}_0, k \in naturals. \tag{2.2.3}$$

(2.2.4)

Sometimes, it will be useful for the theoretical analysis to consider the measurements random and correspondingly treat the analysis and forecast estimates as random elements. We define

$$\boxed{X_{n|k} := \mathbb{E}[X_n|\boldsymbol{Y}_k],} \qquad n \in \mathbb{N}_0, k \in \mathbb{N}. \tag{2.2.5}$$

It will be important to keep in mind that $x_{n|k}$ and $P_{n|k}$ are deterministic, while $X_{n|k}$ is random, because it depends on the random measurements $\boldsymbol{Y}_k$.

With all these preparations, the proof of the following theorem becomes rather simple.

Theorem 2.2.2. *(Kalman filtering) Given the linear discrete filtering problem 2.2.1, we have for $k \in \mathbb{N}$ the following recursive relations (with $x_{0|0} = x_0$, $P_{0|0} = P_0$):*

$$\boxed{\begin{aligned} x_{k|k-1} &= F_k x_{k-1|k-1}, \\ P_{k|k-1} &= F_k P_{k-1|k-1} F_k^* + Q_k, \end{aligned}} \tag{2.2.6}$$

(2.2.7)

*called the **time update**, and*

$$\boxed{\begin{aligned} x_{k|k} &= x_{k|k-1} + K_k(y_k - M_k x_{k|k-1}), \\ P_{k|k} &= (\mathrm{Id} - K_k M_k) P_{k|k-1}, \end{aligned}} \tag{2.2.8}$$

(2.2.9)

*called the **measurement update**,*

$$\boxed{K_k := P_{k|k-1} M_k^* (M_k P_{k|k-1} M_k^* + R_k)^{-1}.} \tag{2.2.10}$$

*is called the **Kalman gain**.*

Proof. By lemma 1.6.9, all X_k and Y_k are normally distributed, and by lemma 1.6.13, the same holds for the distribution of $(Y_k \mid \boldsymbol{Y}_{k-1}, X_k)$ and $(X_k \mid \boldsymbol{Y}_k)$. Taking into account the definitions of the forecast and analysis estimate ((2.2.2) and (2.2.3)), we then have by construction

$$(X_k \mid \boldsymbol{Y}_{k-1} = \boldsymbol{y}_{k-1}) \sim \mathcal{N}(x_{k|k-1}, P_{k|k-1}), \tag{2.2.11}$$

$$(X_k \mid \boldsymbol{Y}_k = \boldsymbol{y}_k) \sim \mathcal{N}(x_{k|k}, P_{k|k}). \tag{2.2.12}$$

To derive the time update, recall that $X_k = F_k X_{k-1} + W_k$, with $W_k \sim \mathcal{N}(0, Q_k)$. From lemma 1.6.9 applied with $X \overset{d}{=} (X_{k-1} \mid \boldsymbol{Y}_{k-1})$ and $Y = W_{k-1}$, we obtain the conditional distribution

$$(X_k \mid \boldsymbol{Y}_{k-1} = \boldsymbol{y}_{k-1}) \sim \mathcal{N}(F_k x_{k-1|k-1}, F_k P_{k-1|k-1} F_k^* + Q_k).$$

Comparing with (2.2.11), this yields the recursive identities

$$\begin{aligned} x_{k|k-1} &= F_k x_{k-1|k-1}, \\ P_{k|k-1} &= F_k P_{k-1|k-1} F_k^* + Q_k. \end{aligned}$$

For the measurement update, we have since $Y_k = M_k X_k + V_k$ and $V_k \sim \mathcal{N}(0, R_k)$ that

$$(Y_k \mid \boldsymbol{Y}_{k-1}, X_k) \overset{d}{=} (Y_k \mid X_k) \sim \mathcal{N}(M_k X_k, R_k). \tag{2.2.13}$$

Using Bayes' theorem for Gaussian distributions (corollary 1.6.14) with $X = X_{k|k-1}$ (and thus $\bar{x} = x_{k|k-1}$ and $\Sigma_X = P_{k|k-1}$), $Y = Y_k, M = M_k, b = 0$ and $\Gamma = R_k$, we obtain

$$(X_k \mid \boldsymbol{Y}_{k-1} = \boldsymbol{y}_{k-1}, Y_k) \sim \mathcal{N}(x_{k|k-1} + K_k(Y_k - M_k x_{k|k-1} k) + x_{k|k-1}, P_{k|k-1} - K_k M_k P_{k|k-1}),$$

with $K_k = P_{k|k-1}M_k^*(M_kP_{k|k-1}M_k^* + R_k)^{-1}$. Hence, inserting the realization $Y_k = y_k$, and comparing again with the definitions of $x_{k|k}$ and $P_{k|k}$ ((2.2.2) and (2.2.3)), we have

$$\begin{aligned} x_{k|k} &= x_{k|k-1} + K_k(y_k - M_k x_{k|k-1}), \\ P_{k|k} &= P_{k|k-1} - K_k M_k P_{k|k-1}. \end{aligned}$$

The statement then follows by induction over $k \in \mathbb{N}$. □

The time update is also known under the terms "forecast update" or "state propagation". Similarly, the measurement update is sometimes called "analysis update" or "state update" (see [Eve03]).

It is also remarkable that at time k, the Kalman filter not only provides us with an optimal estimate $x_{k|k}$ of the current state x_k given all available measurements at this time, but through the time-update equation (2.2.6) also gives optimal predictions $x_{k+1|k}, x_{k+2|k}, \dots$ of the future evolution of the system.

2.3 The optimal linear filter for non-Gaussian noise

Originally (see [Kal60]), the Kalman filter was not derived in the simple setting of problem 2.2.1, but for the more challenging problem of filtering with not necessarily Gaussian noise. In that case, there is in general no analytic formula for the optimal filter. However, it can be shown that the Kalman filter still provides the optimal *linear* filter. For the rest of this section, we will assume for simplicity that the state space is a separable Hilbert space. We then formulate the problem of finding the optimal linear filter in the discrete-time case:

Problem 2.3.1 (Minimum mean-square linear filtering)**.** *Let H be a separable Hilbert space. Consider the linear discrete filtering problem*

$$\begin{aligned} X_k &= F_k X_{k-1} + W_k, \\ Y_k &= M_k X_k + V_k, \end{aligned}$$

where $F_k \in \mathcal{L}(H)$, $M_k \in \mathcal{L}(H; \mathbb{R}^m)$, W_k is a random element of H, and V_k is an m-dimensional random vector. Suppose that X_0, W_k and V_k all have finite second moments, with

$$\begin{aligned} &\mathbb{E}X_0 = \bar{x}_0, \\ &\mathbb{E}V_k = 0, \\ &\mathbb{E}W_k = 0, \\ &\mathrm{Cov}(X_0) = P_0, \\ &\mathrm{Cov}[W_k, X_0] = 0, \\ &\mathrm{Cov}[V_k, X_0] = 0, \\ &\mathrm{Cov}[W_k, W_n] = Q_k\delta_{kn}, \\ &\mathrm{Cov}[V_k, V_n] = R_k\delta_{kn}, \\ &\mathrm{Cov}[W_k, V_n] = 0, \end{aligned}$$

for all $k, n \in \mathbb{N}$. Let

$$\mathfrak{L}(X_k|\boldsymbol{Y}_k) := \{\, A\boldsymbol{Y}_k + b : \, A \in \mathcal{L}(\mathbb{R}^{m\times k}; H), b \in H \,\},$$

denote the set of linear estimators for X_k given measurements $\boldsymbol{Y}_k = \begin{bmatrix} Y_1 & \dots & Y_k \end{bmatrix}$. Then, the minimum mean-square linear filtering problem is to find the ***optimal linear filter*** *$\hat{X}_k \in \mathfrak{L}(X_k|\boldsymbol{Y}_k)$ in the sense that*

$$\boxed{\hat{X}_k = \operatorname*{argmin}_{X \in \mathfrak{L}(X_k|\boldsymbol{Y}_k)} \mathbb{E}\|X_k - X\|^2.}$$

In order to solve this problem, we need to build on the results from the elementary theory of linear estimation. On Hilbert spaces, we have the following central result.

Lemma 2.3.2 (Optimal linear estimator)**.** *Let X be a random element of a separable Hilbert space H, and Y an m-dimensional random vector, with*

$$\begin{aligned}&\mathbb{E}X = \bar{x}, \quad \mathbb{E}Y = \bar{y},\\ &\mathrm{Cov}(X) = \Sigma_{XX}, \quad \mathrm{Cov}(Y) = \Sigma_{YY}, \quad \mathrm{Cov}(X,Y) = \Sigma_{XY},\end{aligned}$$

(cf. proposition 1.4.3) and let

$$\mathfrak{L}(X|Y) := \{\, AY + b :\ A \in \mathcal{L}(\mathbb{R}^m; H), b \in H \,\} \tag{2.3.1}$$

denote the set of linear estimators of X, given Y. Then, the optimal linear estimator

$$E_L(X|Y) := \operatorname*{argmin}_{\hat{X}\in\mathcal{L}(X|Y)} \mathbb{E}\|X - \hat{X}\|^2,$$

is given by

$$\boxed{E_L(X|Y) = \bar{x} + \Sigma_{XY}\Sigma_{YY}^{-1}(Y - \bar{y}).} \tag{2.3.2}$$

Proof. The proof is straightforward. We simply take derivatives with respect to b and A and set them equal to zero, in order to find the optimum. For b, this gives

$$D_b\left[\mathbb{E}\|X - AY - b\|^2\right](b) = 2\mathbb{E}(X - AY - b),$$

and thus equating with zero yields $b = \mathbb{E}X - A\mathbb{E}Y = \bar{x} - A\bar{y}$.

For determining A, recall that

$$\begin{aligned}D_A[(h, Ag)](A)B &= (h, Bg),\\ D_A[(Ah, Ah)](A)B &= 2(Ah, Bh).\end{aligned}$$

With this, we have

$$\begin{aligned}D_A[\mathbb{E}|X - AY - b|^2](A)B &= D_A\left[\mathbb{E}[(X,X) - 2(X,AY) - 2(b,X) + 2(AY,AY) + 2(b,AY) + (b,b)]\right]B\\ &= -2\mathbb{E}(X,BY) + \mathbb{E}(AY,BY) + 2\mathbb{E}(b,BY).\end{aligned}$$

Inserting $b = \bar{x} - A\bar{y}$ then yields

$$\begin{aligned}D_A[\mathbb{E}\|X - AY - b\|^2](A)B &= -2\mathbb{E}(X,BY) + \mathbb{E}(AY,BY) + 2\mathbb{E}(\bar{x} - A\bar{y}, BY)\\ &= -\mathrm{Cov}[X,BY] + \mathrm{Cov}[AY,BY]\\ &= -2\Sigma_{XY}B^* + A\Sigma_{YY}B^*\end{aligned}$$

where the last equality is due to lemma 1.6.9. Equating this expression with zero for all $B \in \mathcal{L}(H_2; H_1)$ implies the necessary condition

$$A\Sigma_{YY} - \Sigma_{XY} = 0,$$

and hence $A = \Sigma_{XY}\Sigma_{YY}^{-1}$. □

The estimator from lemma 2.3.2 is also called the **minimum mean-square error (MMSE) estimator**. Its covariance is equal to the minimum of the mean-square error it attains, and is given in the following corollary.

Corollary 2.3.3. $\mathrm{Cov}(X - E_L(X|Y)) = \Sigma_{XX} - \Sigma_{XY}\Sigma_{YY}^{-1}\Sigma_{XY}^*$.

Proof. Follows by inserting the identity (2.3.2) in the left-hand side. □

Note also the obvious but important property that E_L is linear,

$$E_L(X_1 + X_2|Y) = E_L(X_1|Y) + E_L(X_2|Y).$$

Lemma 2.3.4 (Orthogonality principle)**.** $\mathbb{E}[(X - E_L(X|Y), \hat{X})] = 0$ *for all* $\hat{X} \in \mathfrak{L}(X|Y)$ *(as defined in* (2.3.1)*).*

Proof. Consider $\mathbb{E}\|X - \hat{X}\|^2$ for $\hat{X} \in \mathcal{L}(X|Y)$. Since $E_L(X|Y)$ is by definition the minimum of this expression, its directional derivative in any direction $\hat{X}$ must necessarily equal zero. This means, that for all $\hat{X}_0 \in \mathcal{L}(X|Y)$ we have

$$\begin{aligned} 0 &= D_{\hat{X}} \left[\mathbb{E}\|X - \hat{X}\|^2\right](E_L(X|Y))\hat{X}_0 \\ &= \mathbb{E}\left[(X - E_L(X|Y), \hat{X}_0)\right]. \end{aligned}$$

□

Lemma 2.3.5 (Properties of E_L)**.** *Let* $Y = [Y_1, Y_2]^\top$ *with*

$$\begin{aligned} &\mathrm{Cov}[Y_i, Y_j] = Q_i \delta_{ij}, \qquad i, j \in [2], \\ &\mathrm{Cov}[X, Y] = \Sigma_{XY} := \begin{bmatrix}\Sigma_{XY_1} & \Sigma_{XY_2}\end{bmatrix}. \end{aligned}$$

Then,

i) $E_L(X|Y_1, Y_2) = E_L(X|Y_1) + E_L(X|Y_2) - \bar{x}$.

ii) $\mathrm{Cov}(X - E_L(X|Y_1, Y_2)) = \mathrm{Cov}(X - E_L(X|Y_1)) - \Sigma_{XY_2}\Sigma^{-1}_{Y_2Y_2}\Sigma_{Y_2X}$.

Proof. (i): Note that for $\Sigma_{YY} := \mathrm{Cov}(Y)$, we have by the above assumption

$$\Sigma^{-1}_{YY} = \begin{bmatrix}\Sigma^{-1}_{Y_1Y_1} & 0 \\ 0 & \Sigma^{-1}_{Y_2Y_2}\end{bmatrix}.$$

Then, applying lemma 2.3.2, we obtain

$$\begin{aligned} E_L(X|Y_1, Y_2) &= E_L(X|Y) = \Sigma_{XX} - \Sigma_{XY}\Sigma^{-1}_Y(Y - \bar{y}) \\ &= \bar{x} - \Sigma_{XY_1}\Sigma^{-1}_{Y_1Y_1}(Y_1 - \bar{y}_1) - \Sigma_{XY_2}\Sigma^{-1}_{Y_2Y_2}(Y_2 - \bar{y}_2) \\ &= E_L(X|Y_1) + E_L(X|Y_2) - \bar{x}, \end{aligned}$$

again by lemma 2.3.2.

(ii) Follows similarly from corollary 2.3.3. □

Recall our specific dynamical setting:

$$\begin{aligned} X_k &= F_k X_{k-1} + W_k, \\ Y_k &= H_k X_k + V_k. \end{aligned}$$

In order to use lemma 2.3.5, we would like to have independent (or at least uncorrelated) measurements. This can be achieved by defining the so-called **innovation vectors**

$$\boxed{\tilde{Y}_k := Y_k - E_L(Y_k|\boldsymbol{Y}_{k-1}).}$$

With this definition, note that $\mathbb{E}\tilde{Y}_k = 0$ and $E_L(X_k|\boldsymbol{Y}_n) = E_L(X_k|\tilde{\boldsymbol{Y}}_n)$.

In analogy to theorem 2.2.2, we define

$$\begin{aligned} \hat{X}_{n|k} &:= E_L(X_n|\boldsymbol{Y}_k), \\ \tilde{X}_{n|k} &:= X_n - \hat{X}_{n|k}, \\ \hat{P}_{k|n} &:= \mathrm{Cov}(\tilde{X}_{n|k}). \end{aligned}$$

Then, we can formulate the Hilbert-space version of Kalman's original solution to problem 2.3.1.

Theorem 2.3.6. *The solution to problem 2.3.1 is given by the Kalman filter equations. For $k \in \mathbb{N}$, with initialization $\hat{X}_{0|0} = \mathbb{E}X_0$ and $\hat{P}_{0|0} = \mathrm{Cov}(X_0)$,*

$$\begin{aligned}
\hat{X}_{k|k-1} &= F_k X_{k-1|k-1}, \\
\hat{P}_{k|k-1} &= F_k \hat{P}_{k-1|k-1} F_k^* + Q_k, \\
K_k &= \hat{P}_{k|k-1} M_k^* (M_k \hat{P}_{k|k-1} M_k^* + R_k)^{-1}, \\
\hat{X}_{k|k} &= \hat{X}_{k|k-1} - K_k (Y_k - M_k \hat{X}_{k|k-1}), \\
\hat{P}_{k|k} &= (\mathbb{I} - K_k H_k) \hat{P}_{k|k-1}.
\end{aligned}$$

The optimal linear filter is then given by $\hat{X}_k = \hat{X}_{k|k}$.

Proof. Similar to theorem 2.2.2, we determine $\hat{X}_{k|k}$ in a recursive way: Given $\hat{X}_{k-1|k-1}$, note that

$$\begin{aligned}
\hat{X}_{k|k-1} &= E_L(X_k|\boldsymbol{Y}_{k-1}) = E_L(X_k|\tilde{\boldsymbol{Y}}_{k-1}) \\
&= E_L(F_k X_{k-1} + W_k|\tilde{\boldsymbol{Y}}_{k-1}) = F_k \hat{X}_{k|k-1} + E_L(W_k|\tilde{\boldsymbol{Y}}_{k-1}) \\
&= F_k \hat{X}_{k-1|k-1},
\end{aligned}$$

since W_k is independent of $\tilde{\boldsymbol{Y}}_{k-1}$. For the covariance matrix, we have

$$\begin{aligned}
\hat{P}_{k|k-1} &= \mathrm{Cov}(X_k - E_L(X_k|\boldsymbol{Y}_{k-1})) \\
&= \mathrm{Cov}(F_k X_k + W_k - E_L(F_k X_{k-1} + W_k|\boldsymbol{Y}_{k-1})) \\
&= \mathrm{Cov}(F_k(X_k - \hat{X}_{k-1|k-1}) + W_k) \\
&= F_k \mathrm{Cov}(X_k - \hat{X}_{k-1|k-1}) F_k^* + \mathrm{Cov}(W_k) \\
&= F_k \hat{P}_{k-1|k-1} F_k^* + Q_k.
\end{aligned}$$

Then, for the measurement update, we proceed in the same way

$$\begin{aligned}
\hat{X}_{k|k-1} &= E_L(X_k|\boldsymbol{Y}_{k-1}) = E_L(X_k|\tilde{\boldsymbol{Y}}_{k-1}) \\
&= E_L(X_k|\tilde{\boldsymbol{Y}}_k, \tilde{Y}_k) = E_L(X_k|\tilde{\boldsymbol{Y}}_k) + E_L(X_k|\tilde{Y}_k) - \mathbb{E}X_k,
\end{aligned}$$

by lemma 2.3.5. Inserting the identity (2.3.2), we have

$$\hat{X}_{k|k} = E_L(X_k|\tilde{\boldsymbol{Y}}_k) + \mathrm{Cov}(X_k, \tilde{Y}_k)\mathrm{Cov}(\tilde{Y}_k)^{-1}\tilde{Y}_k.$$

Straightforward computations and an application of the orthogonality principle (lemma 2.3.4) yield

$$\mathrm{Cov}[X_k, \tilde{Y}_k] = \hat{P}_{k|k-1} M_k^*, \tag{2.3.3}$$

and

$$\mathrm{Cov}(\tilde{Y}_k) = M_k \hat{P}_{k|k-1} M_k^* + R_k, \tag{2.3.4}$$

which gives

$$\hat{X}_{k|k} = \hat{X}_{k|k-1} + \hat{P}_{k|k-1} M_k^* (M_k \hat{P}_{k|k-1} M_k^* + R_k)^{-1} (Y_k - M_k \hat{X}_{k|k-1}),$$

since $E_L(Y_k|\tilde{\boldsymbol{Y}}_{k-1}) = M_k \hat{X}_{k|k-1}$. Finally, the recursion formula for $\hat{P}_{k|k}$ follows through an application of lemma 2.3.5, ii.

$$\begin{aligned}
\hat{P}_{k|k} &= \mathrm{Cov}(X_k - E_L(X_k|\tilde{\boldsymbol{Y}}_k)) \\
&= \mathrm{Cov}(X_k - E_L(X_k|\tilde{\boldsymbol{Y}}_{k-1}, \tilde{Y}_k)) \\
&= \mathrm{Cov}(X_k - E_L(X_k|\tilde{\boldsymbol{Y}}_{k-1})) - \mathrm{Cov}[X_k, \tilde{Y}_k]\mathrm{Cov}(Y_k)^{-1}\mathrm{Cov}[\tilde{Y}_k, X_k] \\
&= \hat{P}_{k|k-1} - \hat{P}_{k|k-1} M_k^* (M_k \hat{P}_{k|k-1} M_k^* + R_k)^{-1} M_k \hat{P}_{k|k-1},
\end{aligned}$$

where the last line is due to equations (2.3.3) and (2.3.4). □

2.4 The Kalman-Bucy filter

We consider now the problem of filtering a continuous-time dynamical system given discrete measurements. As discussed at the beginning of section 1.7, we will only consider the case where the state takes values in a separable Hilbert space. The corresponding *continuous-discrete filtering* problem is formulated below.

Problem 2.4.1 (Nonlinear continuous-discrete filtering problem)**.** *Let H and U be separable Hilbert spaces. Consider the state model*

$$\begin{aligned} X(t) &= f(X(t),t)dt + G(t)dW(t), \qquad t \in [0,T], \\ X(0) &= X_0, \end{aligned}$$

where the initial value $X_0 \sim \mathcal{N}(x_0, P_0)$ is a Gaussian random element of H, $f \in M(H \times \mathbb{R}^+; H)$, $G \in M(\mathbb{R}^+ \to \mathcal{L}(Q^{1/2}U; H))$ and W is an U-valued Q-Wiener process. We consider discrete finite-dimensional measurements

$$Y_k = m_k(X(t_k)) + V_k, \qquad k \in \mathbb{N},$$

with measurement function $m_k : H \to \mathbb{R}^m$ and Gaussian measurement noise $V_k \sim \mathcal{N}(0, R_k)$.

Due to the generality of this problem, we can not expect to find a simple solution formula for the nonlinear continuous-discrete filtering problem. Therefore, as in section 2.2, we concentrate on the linear case.

In analogy to the discrete-time case, we define

$$X(t|\tau) := \mathbb{E}[X(t) \mid \boldsymbol{Y}_\tau], \tag{2.4.1}$$
$$x(t|\tau) := \mathbb{E}[X(t) \mid \boldsymbol{Y}_\tau = \boldsymbol{y}_\tau], \tag{2.4.2}$$
$$P(t|\tau) := \mathrm{Cov}(X(t) \mid \boldsymbol{Y}_\tau = \boldsymbol{y}_\tau), \tag{2.4.3}$$

for $t \geq 0$ and $\tau \in \{t_1, t_2, \ldots\}$. We will also denote the state estimates at the times of measurement $t_1, t_2, ...$ by

$$X_{n|k} := X(t_n|t_k), \tag{2.4.4}$$
$$x_{n|k} := x(t_n|t_k), \tag{2.4.5}$$
$$P_{n|k} := P(t_n|t_k),, \qquad n \in \mathbb{N}_0, k \in \mathbb{N}. \tag{2.4.6}$$

We can then derive equations for the evolution of $x(\cdot|\cdot)$ and $P(\cdot|\cdot)$ for the linear Gaussian special case of problem 2.4.1, defined as problem 2.4.2 below.

Problem 2.4.2 (Linear continuous-discrete filtering)**.** *Consider the linear SPDE*

$$dX(t) = AX(t)dt + G(t)dW(t), \qquad t \in [0,T] \tag{2.4.7a}$$
$$X(0) = X_0, \tag{2.4.7b}$$

with measurements

$$Y_k = M_k X(t_k) + V_k, \qquad k \in \mathbb{N}.$$

We assume the following:

- *A is a linear unbounded operator that generates a stronlgy continuous semigroup.*
- *$G \in L^\infty([0,T]; \mathcal{L}_2(Q^{1/2}U; H))$..*
- *W is a U-valued Q-Wiener process.*
- *$M_k \in \mathrm{Lin}(H; \mathbb{R}^m)$.*

- *X_0, W and $V_1, ..., V_K$ are independent.*

It is clear from the definition of problem 2.4.2 that the process X is Gaussian, and so the linear continuous-discrete filtering problem is solved by determining $X(t|\tau)$ and $P(t|\tau)$ for all $t, \tau \geq 0$, as in the linear discrete-time case.

Since the system is now assumed to be linear, $X(t_k)$ is a linear function of $X(t_{k-1})$. Hence, the problem could be reformulated as a linear discrete filtering problem of the form of problem 2.2.1 by taking $F_k = \Phi(t_k, t_{k-1})$, where Φ is the fundamental matrix generated by F. Theorem 2.2.2 then implies that the measurement update will again be given by equation 2.2.8. For the time update, however, instead of using equation 2.2.6 with the evolution semigroup, we can simply apply Itô' s lemma (section 1.7) and arrive at a more elegant continuous-time description of the evolution of the conditional expectation $x(\cdot|\tau)$ between measurement times.

This leads to the *Kalman-Bucy filter* for linear continuous-discrete filtering problems:

Theorem 2.4.3 (Kalman-Bucy)**.** *Suppose we are given problem 2.4.2.Then, between observations, the conditional expectation $x(t|t_{k-1})$ and covariance $P(\cdot|t_{k-1})$ (cf.* (2.4.1)-(2.4.6)*) satisfy the initial value problems*

$$\boxed{\begin{aligned} &D_t x(t|t_{k-1}) = Ax(t|t_{k-1}), \qquad t \in (t_{k-1}, t_k), \\ &x(t_{k-1}|t_{k-1}) = x_{k-1|k-1}, \end{aligned}} \tag{2.4.8}$$

and

$$\boxed{\begin{aligned} &D_t P(t|t_{k-1}) = AP(t|t_{k-1}) + P(t|t_{k-1})A^* + G(t)QG(t)^*, \qquad t \in (t_{k-1}, t_k), \quad (2.4.9a) \\ &P(t_{k-1}|t_{k-1}) = P_{k-1|k-1}, \quad (2.4.9b) \end{aligned}}$$

where $x_{0|0} = x_0$ and $P_{0|0} = P_0$. The measurement update is

$$x_{k|k} = x_{k|k-1} + K_k(y_k - M_k X_{k|k-1}), \tag{2.4.10a}$$
$$P_{k|k} = (\mathbb{I} - K_k M_k)P_{k|k-1}, \tag{2.4.10b}$$

where

$$K_k := P_{k|k-1}M_k^*(M_k P_{k|k-1} M_k^* + R_k)^{-1}.$$

Proof. Recall that by theorem 1.7.11, the unique weak solution to (2.4.7) satisfies

$$X(t) = S(t - t_{k-1})X(t_{k-1}) + \int_{t_{k-1}}^{t} S(t-s)G(s)dW(s), \tag{2.4.11}$$

where S is the strongly continuous semigroup generated by A. Taking the conditional expectation $\mathbb{E}[\cdot|\boldsymbol{Y}_{k-1} = \boldsymbol{y}_{k-1}]$ of both sides, this yields:

$$x(t|t_{k-1}) = S(t - t_{k-1})x_{k-1|k-1}, \tag{2.4.12}$$

where the second term on the right-hand side vanished by lemma 1.7.5, because W is assumed to be independent. Equation (2.4.12) can then again be brought into differential form, leading to the ODE (2.4.8).

Similarly, recalling that the conditional covariance is a bilinear form (see proposition 1.4.3), we have by (2.4.11):

$$\begin{aligned} P(t|t_{k-1}) &= \mathrm{Cov}[X(t), X(t)|\boldsymbol{Y}_{k-1} = \boldsymbol{y}_{k-1}] \\ &= \mathrm{Cov}[S(t - t_{k-1})X(t_{k-1}) + \int_{t_{k-1}}^{t} S(t-s)G(s)dW(s), S(t - t_{k-1})X(t_{k-1})] \quad (2.4.13) \\ &+ \int_{t_{k-1}}^{t} S(t-s)G(s)dW(s) \mid \boldsymbol{Y}_{k-1} = \boldsymbol{y}_{k-1}]. \quad (2.4.14) \end{aligned}$$

Using again lemma 1.7.5 and the independence of W, (2.4.14) reduces to

$$P(t|t_{k-1}) = S(t-t_{k-1})\mathrm{Cov}[X(t_{k-1}, X(t_{k-1})|\boldsymbol{Y}_{k-1} = \boldsymbol{y}_{k-1}]S(t-t_{k-1})^* \tag{2.4.15}$$
$$+ \int_0^t S(t-s)G(s)QG(s)^*S(t-s)^*ds$$
$$= S(t-t_{k-1})P_{k-1|k-1}S(t-t_{k-1})^* + \int_0^t S(t-s)G(s)QG(s)^*S(t-s)^*ds. \tag{2.4.16}$$

Taking the derivative with respect to t of (2.4.16) and using Leibniz' integral rule, we obtain a differential equation for $P(\cdot|t_{k-1})$

$$D_tP(t|t_{k-1}) = AP(t|t_{k-1}) + P(t|t_{k-1})A^* + G(t)QG(t)^*,$$

with the initial condition $P(t_{k-1}|t_{k-1}) = P_{k-1|k-1}$.

After this, the measurement equations follow in the same way as in theorem 2.2.2, namely as a simple application of Bayes' theorem. Given measurement $Y_k = y_k$, we apply corollary 1.6.14 with $X = X_{k|k-1} \overset{\text{def}}{=} X(t_k|t_{k-1})$ (such that $\bar{x} = x_{k|k-1} = x(t_k|t_{k-1})$ and $\Sigma_{XX} = P_{k|k-1} \overset{\text{def}}{=} P(t_k|t_{k-1})$) and $Y = Y_k$ (such that $M = M_k$ and $b = 0$). Then, we have exactly as in theorem 2.2.2

$$\big(X_{k|k-1} \mid Y_k = y_k\big) \sim \mathcal{N}(x_{k|k-1} + K_k(Y_k - M_kX_{k|k-1}), P_{k|k-1} - K_kM_kP_{k|k-1}),$$

with $K_k = P_{k|k-1}M_k^*\left(M_kP_{k|k-1}M_k^* + R_k\right)^{-1}$. The update (2.4.10) then follows from the definitions of $x_{k|k}$ and $P_{k|k}$.

The assertion then follows by induction over k. □

2.5 Linear smoothers

While the Kalman-Bucy filter solves the filtering and the prediction problem in the continuous-discrete setting, there are situations where we also want to infer past states of the system. For example, one might be interested in estimating only the initial state $X(0)$ in real-time based on incoming measurements $Y_1, Y_2, \ldots$. Solutions of this smoothing problem are called **smoothers**, and we will present two versions in the following. The first variant is the **fixed-point smoothing problem**, where one wants to estimate a state $X(\tau)$ at a fixed time $\tau \geq 0$ given measurements $\boldsymbol{Y}_t$ with $t \geq \tau$. In fact, a simple modification of the Kalman-Bucy filter suffices such that it solves the linear smoothing problem.

Proposition 2.5.1 (Fixed-point smoother)**.** *Consider the setup of problem 2.2.1, and let $\tau \in [0,T]$. Then, for all $t > \tau$, the smoothed estimates $x(\tau|t)$ and $P(\tau|t)$ can be determined through the recursion*

$$x(\tau|t_k) = x(\tau|t_{k-1}) + \tilde{K}_k(y_k - M_kx_{k|k-1}),$$
$$P(\tau|t_k) = P(\tau|t_{k-1}) - \tilde{K}_kM_k\Gamma(\tau, t_k|t_{k-1})^*,$$

with

$$\tilde{K}_k = \Gamma(\tau, t_k|t_{k-1})^*M_k^*(M_kP_{k|k-1}M_k^* + R_k)^{-1}.$$

$\Gamma(\tau, \cdot|t_{k-1})$ satisfies for $t_{k-1} \leq \tau < t_k$

$$D_t\Gamma(\tau, t|t_{k-1}) = \Gamma(\tau, t|t_{k-1})A^*, \qquad t \in (\tau, t_k),$$
$$\Gamma(\tau, \tau|t_{k-1}) = P(\tau|t_{k-1}),$$

where $P(\tau|t_{k-1})$ is as in (2.4.9)*. For $\tau < t_{k-1}$, $\Gamma(\tau, \cdot|t_{k-1})$ satisfies*

$$D_t\Gamma(\tau, t|t_{k-1}) = \Gamma(\tau, t|t_{k-1})A^* \qquad t \in (t_{k-1}, t_k),$$

and

$$\Gamma(\tau, t|t_k) = \Gamma(\tau, t_k|t_{k-1}) - \tilde{K}_k M_k P_{k|k-1}.$$

The values of $x_{k|k-1}$ and $P_{k|k-1}$ are given by the Kalman-Bucy recursion (theorem 2.4.3).

Proof. We derive the fixed-point smoother through an augmented-state approach, as in [Sim06, chapter 9]. We keep track of our estimates for the state $X(\tau)$ by adding an additional state variable $\tilde{X}(t)$ with the trivial dynamics

$$\begin{aligned} D_t\tilde{X}(t) &= 0, \qquad t \in (\tau, T), \\ \tilde{X}(\tau) &= X(\tau), \end{aligned}$$

such that $\tilde{X}(t) = X(\tau)$ for all $t > \tau$. Then, we define the augmented state

$$Z(t) = \begin{bmatrix} X(t) \\ \tilde{X}(t) \end{bmatrix}.$$

It has the dynamics

$$\begin{aligned} dZ(t) &= A^z Z(t)dt + G^z(t)dW(t) \\ &= \begin{bmatrix} A & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} X(t) \\ \tilde{X}(t) \end{bmatrix} + \begin{bmatrix} G(t) & 0 \\ 0 & 0 \end{bmatrix} dW(t). \end{aligned}$$

Correspondingly, we can represent the measurements by

$$\begin{aligned} Y_k &= M_k^z Z(t_k) + V_k \\ &= \begin{bmatrix} M_k & 0 \end{bmatrix} \begin{bmatrix} X(t_k) \\ \tilde{X}(t_k) \end{bmatrix} + V_k. \end{aligned}$$

Applying then the time-update of the Kalman filter, we obtain the forecast estimates

$$\begin{bmatrix} x_{k|k-1} \\ \tilde{x}_{k|k-1} \end{bmatrix} = \begin{bmatrix} x(t_k|t_{k-1}) \\ \tilde{x}_{k-1|k-1} \end{bmatrix}, \tag{2.5.1}$$

where $x(t_k|t_{k-1})$ is the solution to

$$\begin{aligned} D_t x(t|t_{k-1}) &= Ax(t|t_{k-1}), \qquad t \in (t_{k-1}, t_k), \\ x(t_{k-1}|t_{k-1}) &= x_{k-1|k-1}. \end{aligned}$$

Writing for the covariance estimates

$$P^z(t|t_k) = \begin{bmatrix} P(t|t_k) & \Gamma(\tau, t|t_k)^* \\ \Gamma(\tau, t|t_k) & \tilde{P}(t|t_k) \end{bmatrix},$$

we can write the time update of the covariance operator as

$$\begin{aligned} D_t P^z(t|t_{k-1} &= A^z P^z(t|t_{k-1}) + P^z(t|t_{k-1})A^{z*} + G^z(t)QG^z(t)^* \\ &= \begin{bmatrix} AP(t|t_{k-1}) + P(t|t_{k-1})A^* + G(\tau, t|t_{k-1})QG(t)^* & \Gamma(\tau, t|t_{k-1})A^* \\ A\Gamma(\tau, t|t_{k-1}^* & 0 \end{bmatrix}. \end{aligned} \tag{2.5.2}$$

Note that $\Gamma(\tau, t|t_{k-1})$ is therefore (proposition 1.4.5) precisely the conditional cross-covariance of $\tilde{X}(t)$ with $X(t)$, given $\boldsymbol{Y}_{k-1} = \boldsymbol{y}_{k-1}$. Recalling that by construction $\tilde{X}(t) = X(\tau)$, we have in particular

$$\Gamma(\tau, t|t_{k-1}) = \mathrm{Cov}[X(\tau), X(t)|\boldsymbol{Y}_{k-1} = \boldsymbol{y}_{k-1}], \tag{2.5.3}$$

(2.5.2) reduces to the component operator updates

$$
\begin{aligned}
&D_t P(t|t_{k-1}) = AP(t|t_{k-1}) + P(t|t_{k-1})A^* + G(t)QG(t)^*, t \in (t_{k-1}, t_k)\\
&P(t_{k-1}|t_{k-1}) = P_{k-1|k-1},\\
&\ddot{\Gamma}(t|t_{k-1}) = \Gamma(t|t_{k-1})A^*,\\
&\Gamma(t|t_{k-1}) = \Gamma_{k-1},\\
&\tilde{P}_{k|k-1} = \tilde{P}_{k-1|k-1}.
\end{aligned}
$$

The Kalman measurement update is

$$
\begin{bmatrix} x_{k|k} \\ \tilde{x}_{k|k} \end{bmatrix} = \begin{bmatrix} x_{k|k-1} \\ \tilde{x}_{k|k-1} \end{bmatrix} + K_k^z (y_k - M_k x_{k|k-1}), \tag{2.5.4}
$$

where the Kalman gain K_k^z for the augmented state is given by

$$
\begin{aligned}
K_k^z &= P_{k|k-1}^z M_k^{z*} (M_k^z P_{k|k-1}^z M_k^{z*} + R_k) - 1\\
&= \begin{bmatrix} P_{k|k-1} M_k^* \\ \Gamma(\tau, t_k|t_{k-1})^* M_k^* \end{bmatrix} (M_k P_{k|k-1} M_k^* + R_k)^{-1}\\
&=: \begin{bmatrix} K_k \\ \tilde{K}_k \end{bmatrix}.
\end{aligned}
$$

Thus, we obtain from (2.5.4):

$$
\tilde{x}_{k|k-1} = \tilde{x}_{k|k-1} + \Gamma(\tau, t_k|t_{k-1})(M_k P_{k|k-1} M_k^* + R_k)^{-1}(y_k - M_k x_{k|k-1}).
$$

The measurement update for the covariance is

$$
\begin{aligned}
P_{k|k}^z &= (\mathrm{Id} - K_k^z M_k^z) P_{k|k-1}^z\\
&= \begin{bmatrix} Id - K_k M_k & 0 \\ -\tilde{K}_k M_k & \mathrm{Id} \end{bmatrix} \begin{bmatrix} P_{k|k-1} & \Gamma(\tau, t_k|t_{k-1})^* \\ \Gamma(\tau, t_k|t_{k-1}) & \tilde{P}_{k|k-1} \end{bmatrix}\\
&= \begin{bmatrix} (\mathrm{Id} - K_k M_k) P_{k|k-1} & (\mathrm{Id} - K_k M_k)\Gamma(\tau, t_k|t_{k-1})^* \\ \Gamma(\tau, t_k|t_{k-1}) - \tilde{K}_k M_k P_{k|k-1} & \tilde{P}_{k|k-1} - \tilde{K}_k M_k \Gamma(\tau, t_k|t_{k-1})^* \end{bmatrix}.
\end{aligned}
$$

Hence, we have the component updates

$$
\begin{aligned}
&P_{k|k} = (\mathrm{Id} - K_k M_k) P_{k|k-1},\\
&\tilde{P}_{k|k} = \tilde{P}_{k|k-1} - \tilde{K}_k M_k \Gamma(\tau, t_k|t_{k-1})^*,\\
&\Gamma(\tau, t_k|t_k) = \Gamma(\tau, t_k|t_{k-1}) - \tilde{K}_k M_K P_{k|k-1}.
\end{aligned}
$$

Note that the recursion for the estimates $x_{k|k}$ and $P_{k|k}$ coincides, as expected, with the Kalman updates from theorem 2.2.2. Now, we recall that, by construction, $\tilde{X}(t) = X(\tau)$ for all $t \geq \tau$. Hence, we have in fact

$$
\begin{aligned}
\tilde{x}(t_k|t_k) &= x(\tau|t_k),\\
\tilde{P}(t_k|t_k) &= P(\tau|t_k).
\end{aligned}
$$

And so we arrive at the desired smoothed estimates for $X(\tau)$. □

We will call the corresponding algorithm the **linear fixed-point smoother**.

As a second variant of the smoothing problem, we consider the case where we are interested in a complete smoothed trajectory, which means that given measurements $Y_1, \ldots, Y_K$, we want to estimate the whole evolution $(X(t))_{t \in [0, t_K]}$ of the system up to the time t_K of the latest measurement. One possible solution of this problem in the linear case is provided by the **Rauch-Tung-Striebel smoother**. It solves this problem in two passes: In the forward pass, it applies the Kalman-Bucy filter up to time t_K to obtain an estimate. Then, in a backward pass, it computes recursively the smoothed estimate $x_{k|K}$ for $k = K-1, K-2, \ldots, 0$, using the following backward recursion (cf. [Sar13]):

Theorem 2.5.2 (Rauch-Tung-Striebel)**.** *Given problem 2.2.1, we have the backward recursion*

$$x_{k-1|K} = x_{k-1|k-1} + L_k(x_{k|K} - x_{k|k-1}),$$
$$P_{k-1|K} = P_{k-1|k-1} + L_k(P_{k|K} - P_{k|k-1})L_k^*,$$

for all $k \in \mathbb{N}$*, where*

$$L_k := \Gamma_k P_{k|k-1}^{-1},$$

and $\Gamma_k := \Gamma(t_{k-1}, t_k|t_{k-1})$ *is as in proposition 2.5.1.*

Proof. Firstly, observe from (2.5.3) in the proof of proposition 2.5.1 that the conditional cross-covariance of $X(t_{k-1})$ with $X(t_k)$, given $\boldsymbol{Y}_{k-1} = \boldsymbol{y}_{k-1}$, is simply

$$\mathrm{Cov}[X_{k-1}, X_k|\boldsymbol{Y}_{k-1} = \boldsymbol{y}_{k-1}] = \Gamma(t_{k-1}, t_k|t_{k-1}).$$

For brevity, let us set $\Gamma_k := \Gamma(t_{k-1}, t_k|t_{k-1})$. We then have

$$(\begin{bmatrix} X_{k-1} \\ X_k \end{bmatrix} |\boldsymbol{Y}_{k-1} = \boldsymbol{y}_{k-1}) \sim \mathcal{N}\left(\begin{bmatrix} x_{k-1|k-1} \\ x_{k|k-1} \end{bmatrix}, \begin{bmatrix} P_{k-1|k-1} & \Gamma_k^* \\ \Gamma_k & P_{k|k-1} \end{bmatrix}\right). \tag{2.5.5}$$

Now, we note that because of the assumed independence of the noise, the state X_{k-1} is conditionally independent of $Y_k, Y_{k+1}, \ldots, Y_K$, given X_k. Thus

$$(X_{k-1}|X_k, \boldsymbol{Y}_K = \boldsymbol{y}_K) \overset{d}{=} (X_{k-1}|X_k, \boldsymbol{Y}_{k-1} = \boldsymbol{y}_{k-1})$$

Applying lemma 1.6.13 (conditioning) to 2.5.5, we obtain, with the definition $L_k = \Gamma_k P_{k|k-1}^{-1}$,

$$\begin{aligned}(X_{k-1}|X_k, \boldsymbol{Y}_{k-1} = \boldsymbol{y}_{k-1}) &\sim \mathcal{N}(x_{k-1|k-1} + L_k(X_k - x_{k|k-1}), P_{k-1|k-1} - \Gamma_k P_{k|k-1}^{-1}\Gamma_k^*) \\ &= \mathcal{N}(x_{k-1|k-1} + L_k(X_k - x_{k|k-1}), P_{k-1|k-1} - L_k P_{k|k-1} L_k^*). \end{aligned} \tag{2.5.6}$$

We also have by (2.4.5) and (2.4.6):

$$(X_k|\boldsymbol{Y}_K = \boldsymbol{y}_K) \sim \mathcal{N}(x_{k|K}, P_{k|K}). \tag{2.5.7}$$

Applying lemma 1.6.11 (marginalization) to the distributions in (2.5.6) and (2.5.7) we obtain

$$(X_{k-1}|\boldsymbol{Y}_K = \boldsymbol{y}_K) \sim \mathcal{N}(x_{k-1|k-1} + L_k(x_{k|K} - x_{k|k-1}), L_k P_{k|K} L_k^* + P_{k-1|k-1} - L_k P_{k|k-1} L_k^*).$$

Thus, by definition of $x_{k-1|K}$ and $P_{k-1|K}$, we obtain the backward recursion

$$x_{k-1|K} = x_{k-1|k-1} + L_k(x_{k|K} - x_{k|k-1}),$$
$$P_{k-1|K} = P_{k-1|k-1} + L_k(P_{k|K} - P_{k|k-1})L_k^*.$$

Lemma 1.6.13 (conditioning) guarantees that all resulting operators are operations are well-defined. □

2.6 Comments and further references

Most of our presentation, in particular the proof of theorem 2.4.3, follows the classic treatment by Jazwinski [Jaz70], which gives a thorough description of discrete- and continuous-time stochastic filtering in finite dimensions. In general, it is hard to find modern treatments of stochastic filtering in infinite dimensions, although considerable work in this direction has been done (see [Cur75a], [Cur75b] and [OS89]).

However, the extension in sections 2.2 and 2.4 of the finite-dimensional results from [Jaz70] to separable Banach or Hilbert spaces were rather straightforward and follow the spirit of [Cur75b] and [Fal67]. Section 2.3 is based on Kalman's original paper [Kal60], where the results were proved in finite dimensions.

The derivations of the linear fixed-point smoother and the Rauch-Tung-Striebel smoother are simple modifications of the proofs in [Sim06, chapter 9] and [Sar13], where the finite-dimensional discrete-time case was treated.

Note that we did not formulate an optimal filter for the nonlinear filtering problem 2.4.1. In theory, however, it is possible to describe such a filter. Consider the finite-dimensional case first. Then, under sufficient regularity assumptions, the information on the state can be represented by a conditional probability density function. A description of the evolution of this conditional density can be obtained from the Fokker-Planck equation (see [Jaz70]), as the following result shows.

Theorem 2.6.1 (Evolution of the densities)**.** *Consider problem 2.4.1 with $U = H = \mathbb{R}^d$. Assume that the state equation*

$$X(t) = (AX(t) + f(X(t)))\,dt + G(t)dW(t), \tag{2.6.1}$$

admits a unique strong solution X, and suppose that there exist densities

$$\begin{aligned}
&\mathbb{P}^{X(t)|\boldsymbol{Y}_k}(dx|\boldsymbol{y}_k) = p(x,t|t_k)dx,\\
&\mathbb{P}^{X(0)}(dx) = p_0(x)dx,\\
&\mathbb{P}^{Y_k|X(t_k)}(dy|x) = q(y|x)dy.
\end{aligned}$$

with $p_0 \in C^1$ and p twice continuously differentiable in space and once continuously differentiable in time. Then, p satisfies

$$D_t p(x,t|t_{k-1}) = \frac{1}{2}\sum_{i,j}\partial_i\partial_j\left[p(x,t|t_{k-1})\left(G(t)QG(t)^\top\right)^i_j\right] - \sum_i \partial_i\left[\left(f(x,t)p(x,t|t_{k-1})\right)^i\right], \qquad t \in (t_{k-1},t_k), \tag{2.6.2}$$

and

$$p(x,0|0) = p_0(x), \tag{2.6.3}$$

$$p(x,t_k|t_k) = p(x,t_{k-1}|t_{k-1})\frac{q(y_k|x)}{\int_{\mathbb{R}^d} p(\tilde{x},t_{k-1}|t_{k-1})q(y_k|\tilde{x})d\tilde{x}}, \tag{2.6.4}$$

for $k \in \mathbb{N}$.

Proof. The proof follows simply from the Fokker-Planck equation for the SDE (2.6.1) combined with proposition 1.3.12. See [Jaz70, theorem 6.1]. □

This result provides the optimal nonlinear filter for the continuous-discrete filtering problem in finite-dimensions, since all the necessary information is contained in the conditional probability densities p. For example, if we want to estimate the state of $X(t)$, at time t_k, we simply have to compute the integral

$$x(t|t_k) = \int_{\mathbb{R}^d} xp(x,t|t_k)dx. \tag{2.6.5}$$

However, the actual computation of this filter might be very difficult in practice, since it involves, even in this finite-dimensional case, the sequential solution of the partial differential equation (2.6.2). Furthermore, the actual computation of estimators for the state, either for its mean or its covariance, involve the numerical computation of d-dimensional integrals. This can be very computationally expensive in practice, especially if the state dimension d is large.

In infinite dimensions, the situation is similar. It is possible to show that the distribution $\mu_t := \mathbb{P}^{X(t)}$ of a weak solution X of a stochastic partial differential equation

$$\begin{aligned} dX(t) &= (AX(t) + F(X(t), t))dt + G(t)dW(t), \\ X(0) &= X_0, \end{aligned} \tag{2.6.6}$$

satisfies a corresponding Fokker-Planck equation (see [BPR11]),

$$D_t\mu_t = L_0^*\mu_t, \qquad t \in (0, T), \tag{2.6.7}$$

$$\mu_0 = \mathbb{P}^{X_0}, \tag{2.6.8}$$

where L_0 is the Kolmogorov operator associated to the SPDE (2.6.6), given by

$$L_0\phi(x, t) = D_t\phi(x, t) + A^*D_x\phi(x, t)x + D_x\phi(x, t)f(x, t), \qquad x \in H.$$

Equation (2.6.7) has to be intepreted in a weak sense, meaning that we have for all functions ϕ in an appropriate test space (see [BDR09]),

$$\int_H \phi(x, t)\mu_t(dx) = \int_H \phi(x, 0)\mathbb{P}^{X_0}(dx) + \int_0^t \int_H L_0\phi(x, s)\mu_s(dx)ds, \qquad t \in [0, T].$$

Based on this equation, one could then derive a description of the evolution of the conditional distribution $\mathbb{P}^{X(t)|\boldsymbol{Y}_k}$, analogous to theorem 2.6.1. However, this again leads to untractable evolution equations. Therefore, practical approaches do not try to calculate the optimal nonlinear filter, but rather find good approximations for which computation is feasible. One of the most successful of these methods is the extended Kalman filter, presented in chapter 3.

Of course, there are still a lot of possible formulations of stochastic filtering problems we did not address. Similar to the discrete-time case (section 2.3), there exist results on the linear optimality of the Kalman-Bucy filter (see [KB61] or [AT67]). Furthermore, we did not address the continuous filtering problem, viz. where the state and the measurement are both continous stochastic processes (see [Cur75b], [OS89] and [BC09]). In that case, there is for example a Hilbert space result by R. Curtain [Cur75b], which considers the following filtering problem.

Problem 2.6.2 (Optimal linear continuous filtering)**.** *Let H be a Hilbert space. Consider the state model*

$$\begin{aligned} dX(t) &= A(t)X(t)dt + G(t)dW(t), \qquad t \in [0, T], \\ X(0) &= X_0, \end{aligned}$$

where W is an U-valued Q-Wiener process, and X_0 is a random element of H with

$$\begin{aligned} \mathbb{E}X_0 &= x_0, \\ \mathrm{Cov}(X_0) &= P_0, \end{aligned}$$

where $x_0 \in H$, and $P_0 \in \mathcal{L}(H; H)$ is a symmetric positive trace-class operator. We are given finite-dimensional continuous-time measurements

$$\begin{aligned} dY(t) &= M(t)X(t)dt + C(t)dV(t), \quad t \in [0, T], \\ Y(0) &= 0, \end{aligned}$$

where $Y(t)$ is m-dimensional, and V is an m-dimensional R-Wiener process. Assume furthermore:

- *For all $t \in [0, T]$, $A(t)$ is a linear, closed, unbounded operator.*
- *A generates a semigroup Φ s.t.*
  - *$\Phi : [0, T] \times [0, T] \to \mathcal{L}(H)$ is strongly continuous.*

– $\Phi(\cdot, s)$ *is strongly differentiable, with*

$$\frac{d}{dt}\Phi(t,s) = A(t)\Phi(t,s).$$

– $\Phi(t,s) = e^{-(t-s)A(t)} + N(t,s)$*, and there exist* $c > 0$ *and* $\theta \in (0,1)$ *such that*

$$\|A(t)N(t,s)\| \leq c|t-s|^{-\theta} \qquad \forall t, s \in [0,T].$$

- $G \in L^\infty([0,T]; \mathcal{L}_2(Q^{1/2}U; H))$.
- $M \in L^\infty([0,T]; \mathcal{L}(H; \mathbb{R}^m))$*, and* $C, C^{-1} \in L^\infty([0,T]; \mathbb{R}^{m\times m})$.
- X_0*,* W *and* V *are independent.*

Then, given a realization $Y = y \in C^0([0,T]; \mathbb{R}^m)$*, let*

$$\mathfrak{L}(X|y) := \{\, \hat{x} : [0,T] \to H : \hat{x}(t) = \int_0^t K(t,s)y(s)ds, \quad K(t,\cdot) \in L^2(I; \mathcal{L}(\mathbb{R}^m; H)), \forall t \in I \,\}$$

denote the class of all linear estimators for X.

Then, the problem is to find a linear estimator that is pointwise optimal in the sense that

$$\hat{x}(t) = \underset{\hat{x}\in\mathfrak{L}(X|y)}{\operatorname{argmin}} \mathbb{E}\|X(t) - \hat{x}(t)\|^2, \qquad t \in [0,T]. \tag{2.6.9}$$

Curtain showed that under all these assumptions the above problem is well-posed, meaning that there always exists a linear estimator $\hat{x}$ that is optimal in the sense of equation 2.6.9.

Theorem 2.6.3 (Curtain, 1975)**.** *Problem 2.6.2 admits a unique solution given by*

$$\hat{x}(t) := \int_0^t U(t,s)K(s)\dot{y}(s)ds,$$

where K *is given by*

$$K(t) := P(t)M(t)^*(C(t)RC(t))^{-1},$$

with P *being the solution of the Riccati equation*

$$\begin{aligned} &D_tP(t) = A(t)P(t) + P(t)A(t)^* + G(t)QG(t)^* - P(t)M(t)^*(C(t)RC(t)^*)^{-1}C(t)P(t), \\ &P(0) = P_0, \end{aligned} \tag{2.6.10}$$

and U *being the evolution semigroup generated by* $A(t) - K(t)M(t)$.

Equivalently, $\hat{x}$ *satisfies the initial value problem*

$$\begin{aligned} &D_t\hat{x}(t) = A(t)\hat{x}(t) + K(t)(D_ty(t) - M(t)\hat{x}(t)), \\ &\hat{x}(0) = x_0 \end{aligned}$$

Chapter 3

Approximate filters

In this chapter, we return to the nonlinear Hilbert space continuous-discrete filtering problem (problem 2.4.1) and discuss practical algorithms. First, we derive an approximate nonlinear filter to adress problem 2.4.1. Then, we discuss the practical application of the Kalman filter to high-dimensional problems. This will lead us to ensemble-based versions of the Kalman filter and the linear fixed-point smoother.

3.1 The extended Kalman filter

Recall the governing equations of the general continuous-discrete filtering problem:

$$dX(t) = f(X(t), t)dt + G(t)dW(t), \tag{3.1.1a}$$
$$X(0) = X_0, \tag{3.1.1b}$$
$$Y_k = m_k(X(t_k)) + V_k, \tag{3.1.1c}$$

where $X_0 \sim \mathcal{N}(x_0, P_0)$. A direct approach to this problem would be to use the general theory of nonlinear filtering (see [Jaz70] and [BC09]) to derive optimal nonlinear filters. Such filters are usually analytically intractable, so one has to work with numerical approximations, often of Monte Carlo type [Kus11]. A different approach is to approximate not the resulting filter but the nonlinear system (3.1.1), and then to apply an exact linear filter. The most popular example of this type of methods is the **extended Kalman filter** (EKF). This filter is derived by linearizing the state and measurement functions f and m around a deterministic trajectory $\bar{x} = \bar{x}(t)$ for $t \in (0, t_1)$. Then, the Kalman-Bucy filter from theorem 2.4.3 is applied to the linearized system to give an estimate $x_{1|1}$ with covariance $P_{1|1}$. A new reference solution is computed on (t_1, t_2) based on this estimate, and the procedure is repeated to produce a sequence of estimates $(x_{k|k}, P_{k|k})_{k=1}^{\infty}$.

In more detail, we start with a **reference trajectory** $\bar{x}$ around which we will develop a Taylor approximation of the process X. The idea is to start with the initial guess x_0, and then to simulate the deterministic dynamics up to the first measurement time t_1. Using the first measurement $Y_1 = y_1$, we then compute an approximate estimate for the state, given y_1, denoted by $\tilde{x}_{1|1} \approx \mathbb{E}[X(t_1)|Y_1 = y_1]$, and compute the reference trajectory up to time t_2 using that estimator as an initial guess, and so on.

This means we want $\bar{x}$ on $[0, t_1]$ to be the solution of the deterministic initial value problem

$$D_t\bar{x}(t) = f(\bar{x}(t), t), \qquad t \in (0, t_1), \tag{3.1.2a}$$
$$\bar{x}(0) = x_0. \tag{3.1.2b}$$

Then, we define the random **state deviation**

$$\delta X(t) := X(t) - \bar{x}(t), \qquad t \in [0, T]. \tag{3.1.3}$$

It then follows from the definitions in problem 2.4.1 and equation 3.1.2 that δX satisfies the SDE

$$\begin{aligned} d(\delta X)(t) &= (f(X(t),t) - f(\bar{x}(t),t))dt + G(t)dW(t) \\ &\approx \partial_1 f(\bar{x}(t),t)\delta X(t)dt + G(t)dW(t), \qquad t \in (0,t_1), \\ \delta X(0) &= X(0) - x_0 \sim \mathcal{N}(0,P_0), \end{aligned}$$

where the second line is a Taylor approximation around $\bar{x}$. This motivates us to define the approximate state deviation $\delta\tilde{X}$ as the solution of the linearized system

$$\begin{aligned} d(\delta\tilde{X})(t) &= F(t;0)\delta\tilde{X}(t)dt + G(t)dW(t), \qquad t \in (0,t_1), \\ \delta\tilde{X}(t_0) &\sim \mathcal{N}(0,P_0), \end{aligned}$$

where $F(t;t_0) := \partial_1 f(\bar{x}(t),t)$. Since we have $X = \bar{x} + \delta X$ by definition, this in turn leads to the state approximation

$$\tilde{X} := \bar{x} + \delta\tilde{X} \tag{3.1.4}$$

We then define

$$\begin{aligned} \tilde{P}(t|t_k) &:= \mathrm{Cov}(\tilde{X}(t)|\boldsymbol{Y}_k = \boldsymbol{y}_k) \\ &= \mathrm{Cov}(\delta\tilde{X}(t)|\boldsymbol{Y}_k = \boldsymbol{y}_k). \end{aligned} \tag{3.1.5}$$

Defining the **nominal measurement**

$$\bar{y}_1 := m(\bar{x}(t_1),t_1),$$

and the **measurement deviation**

$$\begin{aligned} \delta Y_1 := Y_1 - \bar{y}_1 &= m_1(X(t_1)) - m_1(\bar{x}(t_1)) + V_1 \\ &\approx \underbrace{Dm_1(\bar{x}(t_1))}_{=:M_1}\delta X(t_1) + V_1, \\ &=: \delta\tilde{Y}_1, \end{aligned} \tag{3.1.6}$$

we can approximate the measurement function at t_1 by its linearization M_1. To summarize, the linearized system is then

$$d(\delta\tilde{X})(t) = F(t;0)\delta\tilde{X}(t)dt + G(t)dW(t), \tag{3.1.7a}$$

$$\delta\tilde{X}(0) \sim \mathcal{N}(0,P_0), \tag{3.1.7b}$$

$$\delta\tilde{Y}_1 = M_1\delta\tilde{X}_1(t_1) + V_1. \tag{3.1.7c}$$

In order to compute the estimates $\tilde{x}_{1|1}$ and $P_{1|1}$, we apply one time update and one measurement update of the Kalman-Bucy filter (theorem 2.4.3) to the linearized filtering problem (3.1.7) and obtain the measurement update equations

$$\begin{aligned} \delta\tilde{x}_{1|1} &= \delta\tilde{x}_{1|0} + K_1(\delta\tilde{y}_1 - M_1\delta x_{1|0}), \\ \tilde{P}_{1|1} &= (\mathbb{I} - K_1M_1)\tilde{P}_{1|0}, \end{aligned} \tag{3.1.8}$$

where $\tilde{P}_{1|0} := \tilde{P}(t_1|t_0)$ is determined by the initial value problem

$$\begin{aligned} D_t\tilde{P}(t|0) &= F(t;t_0)\tilde{P}(t|t_0) + \tilde{P}(t|t_0)F(t;t_0)^* + G(t)QG(t)^*, t \in (0,t_1), \\ \tilde{P}(0|0) &= P_0, \end{aligned}$$

and $K_1 := P_{1|0}M_1^*(M_1P_{1|0}M_1^* + R_1)^{-1}$. Using the definitions of $\tilde{X}$ (see (3.1.4)) and $\delta\tilde{Y}$ (see (3.1.6)), and using the approximation $\delta\tilde{y}_1 \approx \delta y_1 = y_1 - \bar{y}_1$ we can rewrite equation 3.1.8 as

$$\tilde{x}_{1|1} = \tilde{x}_{1|0} + K_1(y_1 - m_1(\bar{x}(t_1)),$$

and since $\mathrm{Cov}(\tilde{X}) = \mathrm{Cov}(\delta\tilde{X})$, we have for the covariance estimate

$$\mathrm{Cov}(\tilde{X}(t_1)|Y_1 = y_1) = \tilde{P}_{1|1}.$$

For the time update of the approximate state estimator $\tilde{X}$ we simply have

$$\mathbb{E}\tilde{X}(t|0) = \bar{x}(t), \qquad t \in [0, t_1], \tag{3.1.9}$$

so that in particular $\tilde{x}_{1|0} \overset{\text{def}}{=} \mathbb{E}\tilde{X}(t_1|0) = \bar{x}(t_1)$. For the next interval $[t_1, t_2]$, one repeats this procedure using the current estimate $\tilde{x}_{1|1}$ as initial value for the reference trajectory, so that $\bar{x}$ is given on $[t_1, t_2]$ as the solution of the initial value problem

$$\begin{aligned} &D_t\bar{x}(t) = f(\bar{x}(t), t), \qquad t \in (t_1, t_2), \\ &\bar{x}(t_1) = \tilde{x}_{1|1}, \end{aligned}$$

and then proceeds as above. Repeating for $k = 2, 3, ...$ leads to algorithm 1.

Algorithm 1 `Extended Kalman filter`

Given problem 2.4.1, with measurements $Y_1 = y_1, Y_2 = y_2, ...$

1: Initialize $x_{0|0} = x_0$, $P_{0|0} = P_0$;
2: **for** k = 1,2,... **do**
`Time-update`
3: Solve

$$\begin{aligned} &D_t x(t|t_{k-1}) = f(x(t|t_{k-1}), \qquad t \in (t_{k-1}, t_k), \\ &x(t_{k-1}|t_{k-1}) = x_{k-1|k-1}; \end{aligned}$$

4: Set $x_{k|k-1} = x(t_k|t_{k-1})$;
5: Solve

$$\begin{aligned} &D_t P(t|t_{k-1}) = F(t|t_{k-1})P(t|t_{k-1}) + P(t|t_{k-1})F(t|t_{k-1})^* + G(t)QG(t)^*, \qquad t \in (t_{k-1}, t_k), \\ &P(t_{k-1}|t_{k-1}) = P_{k-1|k-1}, \end{aligned}$$

with $F(t|t_{k-1}) := \partial_1 f(x(t|t_k), t)$;
6: Set $P_{k|k-1} = P(t_k|t_{k-1})$;
`Measurement-update`
7: Set $M_k = Dm_k(X_{k|k-1})$;
8: Set $K_k = P_{k|k-1}M_k^*(M_k P_{k|k-1} M_k^* + R_k)^{-1}$;
9: Set $x_{k|k} = x_{k|k-1} + K_k(y_k - m_k(x_{k|k-1}))$;
10: Set $P_{k|k} = (\mathrm{Id} - K_k M_k)P_{k|k-1}$;

In practice, the solution of the evolution equations (lines 3 and 5) in Hilbert space would be achieved through a numerical approximation. In principle, any numerical method for partial differential equations (see e.g. [Gro07]) could be applied here, depending on the particular problem. For more discussion of the EKF and its variants introduced below, see section 3.6.

3.2 The iterated extended Kalman filter

Since the extended Kalman filter is only a first-order approximation, the theoretical optimality results of the Kalman-Bucy filter (theorem 2.4.3) do not apply. Hence, the estimate $x_{k|k}$ from the extended Kalman filter can be improved upon. For example, we can relinearize the measurement equation around $x_{k|k}$ and then reapply the measurement update in order to get a new estimate $x^1_{k|k}$. Iterating this procedure leads to the **iterated extended Kalman filter** (IEKF). In the following, let $x_{k|l}$ and

$P_{k|l}$ denote the mean and covariance estimates generated by the EKF applied to the filtering problem 2.4.1. Suppose we are given a deterministic guess $x^n_{k|k} \in H$ of the state $X(t_k)$. As in the derivation of the EKF, we can linearize the measurement equation around this estimate: Defining

$$\delta Y_k := Y_k - m_k(x^n_{k|k}),$$

we have

$$\begin{aligned} \delta Y^n_k &= m(X(t_k)) - m_k(x^n_{k|k}) + V_k && (3.2.1)\\ &\approx M^n_k(\tilde{X}(t_k) - x^n_{k|k}) + V_k =: \delta\tilde{Y}^n_k, && (3.2.2)\end{aligned}$$

where

$$\boxed{M^n_k := Dm_k(x^n_{k|k})}$$

and

$$\tilde{X} := \bar{x} + \delta\tilde{X} \tag{3.2.3}$$

is the linearization of the state around the reference trajectory $\bar{x}$, where $\delta\bar{X}$ is the solution to

$$\begin{aligned} d(\delta\tilde{X}(t)) &= F(t;t_{k-1})\delta\tilde{X}(t)dt + G(t)dW(t), \qquad t \in (t_{k-1}, t_k), && (3.2.4)\\ \delta\tilde{X}(t_{k-1}) &\sim \mathcal{N}(0, P_{k|k-1}), && (3.2.5)\end{aligned}$$

with $F(t;t_{k-1}) := \partial_1 f(\bar{x}(t), t)$ (see section 3.1). Let us set

$$\begin{aligned} \Delta X^n_k &:= \tilde{X}(t_k) - x^n_{k|k}, && (3.2.6)\\ \Delta x^n_{k|l} &:= \mathbb{E}[\Delta X^n_k | \boldsymbol{Y}_l = \boldsymbol{y}_l]. \end{aligned}$$

Note that ΔX^n_k is by definition Gaussian (since $\tilde{X}(t_k)$ is Gaussian), and for all $n \in \mathbb{N}_0$:

$$\begin{aligned} \Delta x^{n+1}_{k|k-1} &= \mathbb{E}[\tilde{X}(t_k) - x^n_{k|k} | \boldsymbol{Y}_{k-1} = \boldsymbol{y}_{k-1}]\\ &= \bar{x}(t_k) - \mathbb{E}[\delta\tilde{X}(t_k) | \boldsymbol{Y}_{k-1} = \boldsymbol{y}_{k-1}] - x^n_{k|k}\\ &= \bar{x}(t_k) - x^n_{k|k}\\ &= x_{k|k-1} - x^n_{k|k}. && (3.2.7)\end{aligned}$$

where the second line follows from the definition of $\tilde{X}$ in (3.2.3), the third line follows from (3.2.5) and the fourth line follows from the definition $x_{k|k-1} = \bar{x}(t_k)$ in the extended Kalman filter (see (3.1.9)). Let furthermore

$$P^{n+1}_{k|l} := \mathrm{Cov}(\Delta X^{n+1}_k | \boldsymbol{Y}_l = \boldsymbol{y}_l). \tag{3.2.8}$$

Note similarly to (3.2.7) that

$$\begin{aligned} P^{n+1}_{k|k-1} &= \mathrm{Cov}(\tilde{X}(t_k) | \boldsymbol{Y}_{k-1} = \boldsymbol{y}_{k-1})\\ &= \mathrm{Cov}(\delta\tilde{X}(t_k) | \boldsymbol{Y}_{k-1} = \boldsymbol{y}_{k-1})\\ &= P_{k|k-1}, && (3.2.9)\end{aligned}$$

where the second line follows again from (3.2.3) and the third line follows from the definition $P_{k|k-1} = \mathrm{Cov}(\delta\tilde{X}_k | \boldsymbol{Y}_{k-1} = \boldsymbol{y}_{k-1})$ in the extended Kalman filter (see (3.1.5)). Treating the measurement $Y_k = y_k$ as new, we can do a Bayesian update for ΔX^n_k. This means we apply corollary 1.6.14 with $X \overset{d}{=} (\Delta X^{n+1}_k | \boldsymbol{Y}_{k-1} = \boldsymbol{y}_{k-1}), Y = \delta\tilde{Y}^n_k$ (and thus $M = M^n_k$) and obtain

$$\begin{aligned} \Delta x^{n+1}_{k|k} &= \Delta x^n_{k|k-1} + K^n_k(\delta Y^n_k - M^n_k \Delta x^n_{k|k-1}),\\ P^{n+1}_{k|k} &= (\mathrm{Id} - K^n_k M^n_k) P_{k|k-1}, \end{aligned}$$

where

$$K_k^n = P_{k|k-1}(M_k^n)^*(M_k^n P_{k|k-1}(M_k^n)^* + R_k)^{-1} \tag{3.2.10}$$

(we used that $P_{k|k-1}^{n+1} = P_{k|k-1}$, see (3.2.9)). Using $y_k - m_k(x_{k|k}^n)$ as realization of δY_k^n (see (3.2.1) and (3.2.2)), we obtain for the estimate $\Delta x_{k|k}^{n+1}$:

$$\begin{aligned}\Delta x_{k|k}^{n+1} &:= \Delta x_{k|k-1}^{n+1} + K_k^n(y_k - m_k(x_{k|k}^n) - M_k^n \Delta x_{k|k-1}^n) \\ &= x_{k|k-1} - x_{k|k}^n + K_k^n(y_k - m_k(x_{k|k}^n) - M_k^n(x_{k|k-1} - x_{k|k}^n),\end{aligned} \tag{3.2.11}$$

where the second line is due to equaton (3.2.7).

Since $\Delta X_k^{n+1} \overset{\text{def}}{=} \tilde{X}(t_k) - x_{k|k}^n$, this suggests the updated state estimate

$$x_{k|k}^{n+1} := \mathbb{E}[\tilde{X}(t_k)|\boldsymbol{Y}_k = \boldsymbol{y}_k] = x_{k|k}^n + \Delta x_{k|k}^{n+1}, \tag{3.2.12}$$

for $\mathbb{E}[X(t_k)|\boldsymbol{Y}_k = \boldsymbol{y}_k]$. If we insert (3.2.11) in (3.2.12), we obtain

$$\boxed{x_{k|k}^{n+1} = x_{k|k-1} + K_k^n(y_k - m_k(x_{k|k}^n) - M_k^n(x_{k|k-1} - x_{k|k}^n))} \tag{3.2.13}$$

Correspondingly, by (3.2.6) there holds $\mathrm{Cov}(\Delta X_k^{n+1}) = \mathrm{Cov}(\tilde{X}(t_k))$. This suggests through (3.2.8) to estimate the covariance of $X(t_k)$, given $\boldsymbol{Y}_k = \boldsymbol{y}_k$, by $P_{k|k}^{n+1}$.

The iteration (3.2.13) can then be repeated until convergence (see algorithm 2). Furthermore, it can be checked that if we only compute one iteration of (3.2.13), the estimates $x_{k|k}^1$ and $P_{k|k}^1$ coincide with the EKF estimates $x_{k|k}$ and $P_{k|k}$ as computed through algorithm 1.

3.3 The ensemble Kalman filter

While the EKF and its variants have experienced widespread use in practical applications (see for example [Aug+13]) there are still a lot of problems where they are computationally infeasible. In the case where the state space H is very high-dimensional (or even infinite-dimensional), the need to store and compute the covariance estimates makes the EKF as described above impractical. In the following section, we discuss a Monte Carlo approximation that circumvents this problem, called the ensemble Kalman filter. We will introduce it in the simple situation of linear discrete filtering, and then present a few extensions and modifications to more challenging problems.

The **ensemble Kalman filter** (EnKF) is a Monte Carlo approximation of the classical Kalman filter. We will focus on the continuous-discrete case. However, the purely discrete and purely continuous cases are quite similar (see [Eve03] and [Tag+18], respectively). Recall the linear continuous-discrete filtering problem (see problem 2.4.2)

$$X(t) = AX(t) + G(t)dW(t), \tag{3.3.1}$$
$$X(0) = X_0,$$
$$Y_k = M_k X(t_k) + V_k, \tag{3.3.2}$$

where X_0 is a Gaussian random element of a separable Hilbert space H, W is an independent (generalized) Q-Wiener process, and $V_1, V2, \ldots$ are independent Gaussian random vectors. Suppose the Kalman-Bucy filter (theorem 2.4.3) is applied in a practical, finite-dimensional setting. Explicitly calculating the covariance $P_{k|k}$ with formula (2.4.9) can become very costly if the problem is high-dimensional. For state dimension d (i.e. $H = \mathbb{R}^d$ in problem 2.4.2) and measurement dimension m, the computational complexity will be of order $O(md^2)$ [Man06]. The ensemble Kalman filter circumvents this problem by generating an ensemble $(X^j)_{j=1}^N$ of random elements such that its mean empirical distribution converges towards the distribution of $X(t_k)$, given $\boldsymbol{Y}_k = \boldsymbol{y}_k$, in the limit. Even though the ensemble Kalman filter was originally [Eve94] developed in the setting of finite-dimensional filtering, it

Algorithm 2 `Iterated extended Kalman filter`

Given problem 2.4.1, with measurements $Y_1 = y_1, Y_2 = y_2, ...$

Initialize $x_{0|0} = \bar{x}_0$, $P_{0|0} = P_0$;
for k = 1,2,... **do**
    `Time-update:`
    Solve

$$D_t x(t|t_{k-1}) = f(x(t|t_{k-1}), t), \qquad t \in (t_{k-1}, t_k),$$
$$x(t_{k-1}|t_{k-1}) = x_{k-1|k-1};$$

    Set $x_{k|k-1} = x(t_k|t_{k-1})$;
    Solve

$$D_t P(t|t_{k-1}) = F(t|t_{k-1})P(t|t_{k-1}) + P(t|t_{k-1})F(t|t_{k-1})^* + G(t)QG(t)^*, \qquad t \in (t_{k-1}, t_k),$$
$$P(t_{k-1}|t_{k-1}) = P_{k-1|k-1},$$

    where $F(t|t_{k-1}) = \partial_1 f(x(t|t_{k-1}), t)$;
    Set $P_{k|k-1} = P(t_k|t_{k-1})$;
    `Measurement-update`
    Set $x^0_{k|k} = x_{k|k-1}$;
    Set $n = 0$;
    repeat
        Set $M^n_k = Dm_k(x^n_{k|k})$;
        Set $K^n_k = P_{k|k-1}(M^n_k)^*(M^n_k P_{k|k-1}(M^n_k)^* + R_k)^{-1}$;
        Set $x^{n+1}_{k|k} = x_{k|k-1} + K^n_k(y_k - m_k(x^n_{k|k}) - M^n_k(x_{k|k-1} - x^n_{k|k}))$;
        $n \leftarrow n + 1$;
    until convergence
    Set $x_{k|k} = x^n_{k|k}$;
    Set $P_{k|k} = (\mathrm{Id} - K^n_k M^n_k)P_{k|k-1}$;

can be formulated in our more abstract, infinite dimensional Hilbert space setting (see also [KM17]), as we will show now.

The core idea of the EnKF is to store the information on the conditional distribution of the state X, given measurements $\boldsymbol{Y}_k$, in form of an evolving ensemble $(X^j(t))$. This means that we aim to define an *interacting stochastic dynamical system of particles* in such a way that we have (at least approximately) the equality in distribution

$$X^j(t|t_k) \overset{d}{=} (X(t) \mid \boldsymbol{Y}_k = \boldsymbol{y}_k).$$

Let us assume inductively that $X^j_{k-1|k-1} \overset{d}{=} (X(t_k)|\boldsymbol{Y}_{k-1} = \boldsymbol{y}_{k-1})$. Then, let $X^j_{k|k} = X^j(t|t_k)$ be the solution to the SPDE

$$\begin{aligned} dX^j(t|t_{k-1}) &= AX^j(t|t_{k-1})dt + G(t)dW^j(t), \\ X^j(t|t_{k-1}) &= X^j, \end{aligned}$$

at time t_k, where for each $j \in [N]$, W^j is an independent copy of the Wiener process W. We have by (3.3.1) that

$$X^j_{k|k-1} \overset{d}{=} (X(t_k)|\boldsymbol{Y}_{k-1} = \boldsymbol{y}_{k-1}),$$

as desired. If we set

$$Y^j_k := M_k X^j_{k|k-1} + V^j_k, \tag{3.3.3}$$

with $V^1_k \overset{d}{=} \cdots \overset{d}{=} V^N_k \overset{d}{=} V_k$ mutually independent, we have by equation (3.3.2):

$$Y^j_k \overset{d}{=} (Y_k|\boldsymbol{Y}_{k-1} = \boldsymbol{y}_{k-1}).$$

By linearity, Y_k and $X(t_k)$ are jointly Gaussian distributed. Then, recall from lemma 1.6.13 (conditioning) and definition 2.4.5 of $x_{k|k-1}$ that we have

$$\begin{aligned} (X(t_k)|\boldsymbol{Y}_{k-1} = \boldsymbol{y}_{k-1}, Y_k = y_k) \sim \mathcal{N}(&x_{k|k-1} - \Sigma_{X(t_k),Y_k|k-1}(\Sigma_{Y_k,Y_k|k-1})^{-1}(y_k - \bar{y}_k), \\ &\Sigma_{X(t_k),X(t_k)|k-1} - \Sigma_{X(t_k),Y_k|k-1}(\Sigma_{Y_k,Y_k|k-1})^{-1}\Sigma_{Y_k,X(t_k)|k-1}), \end{aligned} \tag{3.3.4}$$

where we defined

$$\Sigma_{X,Y|k-1} := \mathrm{Cov}[X, Y|\boldsymbol{Y}_{k-1} = \boldsymbol{y}_{k-1}]. \tag{3.3.5}$$

Motivated by this, we would like to update the ensemble by

$$X^j_{k|k} = X^j_{k|k-1} + \Sigma_{X(t_k),Y_k|k-1}\Sigma^{-1}_{Y_k,Y_k|k-1}(y_k - Y^j_k), \tag{3.3.6}$$

because then we would have from (3.3.4):

$$\begin{aligned} \mathbb{E}X^j_{k|k} &= x_{k|k-1} + \Sigma_{X(t_k),Y_k|k-1}\Sigma^{-1}_{Y_k,Y_k|k-1}(y_k - \bar{y}_k) \\ &= \mathbb{E}[X(t_k)|\boldsymbol{Y}_k = y_k]. \end{aligned}$$

However, in practice we do not have access to the covariances, so we need to estimate them. We use the standard estimator

$$\Sigma_{XY} \approx \frac{1}{N-1}\sum_{j=1}^{N}(X^j - \bar{X}) \otimes (Y^j - \bar{Y}).$$

To this end, let

$$\bar{X}_{k|k-1} := \frac{1}{N}\sum_{j=1}^{N} X^j_{k|k-1}$$

be the (random) ensemble mean, and let

$$D_{k|k-1} := \left[X^1_{k|k-1} - \bar{X}_{k|k-1} \quad \dots \quad X^N_{k|k-1} - \bar{X}_{k|k-1} \right],$$

denote the collection of ensemble deviations[1] and

$$B_k := \left[Y^1_k - \bar{Y}_k \quad \dots \quad Y^N_k - \bar{Y}_k \right] \tag{3.3.7}$$

the corresponding measurement deviations, with

$$\bar{Y}_k := \frac{1}{N} \sum_{j=1}^{N} Y^j_k.$$

Then, we estimate

$$\Sigma_{X(t_k),Y_k|k-1} \approx \frac{1}{N-1} D_{k|k-1} \otimes B_k,$$

and

$$\Sigma_{Y_k,Y_k|k-1} \approx \frac{1}{N-1} B_k \otimes B_k.$$

Thus, the ideal update (3.3.6) can be approximated by

$$X^j_{k|k} = X^j_{k|k-1} + \frac{1}{N-1} D_{k|k-1} \otimes B_k (\frac{1}{N-1} B_k \otimes B_k + R_k)^{-1} (y_k - Y^j_k). \tag{3.3.8}$$

Note, however, that this introduces an interaction in the ensemble: The updated ensemble member $X^j_{k|k}$ depends not only on $X^j_{k|k-1}$ and Y^j_k, but through $D_{k|k-1}$ also on all other ensemble members.

A very convenient property of the EnKF is that, in contrast to the Kalman-Bucy filter, it does not involve direct computations with the evolution operator A and the measurement operator M_k. It only requires direct evaluations of these operators on concrete ensemble members. This has two advantages: On one hand, it makes it suited for high-dimensional problems in which direct manipulation of large operators might not be practical (see [HM01]). On the other hand, it means that the algorithm can be applied to nonlinear problems without relying on any linearizations. In that case, the time update is simply computed by solving the nonlinear SPDE

$$\begin{aligned} dX^j(t|t_{k-1}) &= f(X^j(t|t_{k-1}), t)dt + G(t)dW^j(t), \qquad t \in (t_{k-1}, t_k), \\ X^j(t_{k-1}|t_{k-1}) &= X^j_{k-1|k-1}, \end{aligned}$$

for each ensemble members. Correspondingly, the simulated measurements are now simply generated using now the nonlinear measurement function. This means that (3.3.3) is substituted with

$$Y^j_k := m_k(X^j_{k|k-1}) + V^j_k.$$

Apart from that, the same algorithm applies without change (see 3).

In order to guarantee that the expression

$$\frac{1}{N-1} D_{k|k-1} \otimes B_k (\frac{1}{N-1} B_k \otimes B_k + R_k)^{-1} (y_k - Y^j_k)$$

in (3.3.8) is always well-defined, it is recommended to choose R_k to be a positive definite matrix. In the case of Hilbert-space-valued measurements (which we do not treat in this thesis), R_k would have to be lower-bounded in order to guarantee well-definedness of this term (see [KM17]). This requirement leads to conceptual difficulties, since covariance operators are trace-class and thereby not lower-bounded.

[1] These deviations are usually called "anomalies" in the EnKF literature and denoted by A. However, since A already denotes the linear operator in (3.3.1), we use an alternative notation.

Algorithm 3 `Ensemble Kalman filter, EnKF`

Given problem 2.4.1, measurements $Y_1 = y_1, Y_2 = y_2, ...$, and ensemble size $N \in \mathbb{N}$.

1: Generate independent samples $x^1_{0|0}, ..., x^J_{0|0} \sim \mathcal{N}(x_0, P_0)$;
2: **for** $k = 1, 2, ...$ **do**
3: For all $j \in [N]$, solve

$$dx^j(t) = f(x^j(t), t)dt + dw^j(t), \qquad t \in (t_{k-1}, t_k),$$
$$x^j(t_{k-1}) = x^j_{k-1|k-1},$$

and set $x^j_{k|k-1} = x^j(t_k)$;
4: Compute $\bar{x}_{k|k-1} = \frac{1}{N}\sum_j x^j_{k|k-1}$;
5: Assemble $D_{k|k-1} = [x^1_{k|k-1} - \bar{x}_{k|k-1}, \dots, x^N_{k|k-1} - \bar{x}_{k|k-1}]$;
6: Generate independent samples $v^1_k, \dots, v^N_k \sim \mathcal{N}(0, R_k)$;
7: Set $y^j_k = m_k(x^j_{k|k-1})$ for all $j \in [N]$;
8: Compute $\bar{y}_k = \frac{1}{N}\sum_j y^j_k$;
9: Assemble $B_k = [y^1_k - \bar{y}_k, \dots, y^N_k - \bar{y}_k]$;
10: Set $x^j_{k|k} = x^j_{k|k-1} + D_{k|k-1} \otimes B_k \left(B_k \otimes B_k + (N-1)R_k\right)^{-1} (y_k - y^j_k)$ for all $j \in [N]$;

One can implement the finite-dimensional EnKF with a computational complexity of $O(m^3 + m^2N + mN^2 + dN^2)$. If the measurement dimension m is large, an alternative implementation that has complexity $O(N^3 + mN^2 + dN^2)$ might be better suited [Man06].

Since the whole point of the ensemble Kalman filter is to provide a Monte Carlo approximation of its exact counterpart, one is automatically confronted with the question of convergence. Theoretical analysis is then complicated by the fact that, in general, the random vectors $X^1_{k|k}, ..., X^N_{k|k}$ are neither independent nor normally distributed. This can be seen by comparing equation (3.3.8) to equation (2.2.8). In the former, $D_{k|k-1}$ and B_k are now also random elements with nontrivial distributions, such that the linear combination lemma 1.6.9 does not apply anymore. Nevertheless, convergence can be proven with rather elementary methods, for example in the following result:

Proposition 3.3.1 (Convergence of the EnKF in the large-ensemble limit)**.** *Consider the linear Gaussian discrete filtering problem 2.2.1 with $H = \mathbb{R}^d$. Let $k \in \mathbb{N}$ and $(X^j_{k|k})_{j=1}^N$ constructed by the ensemble Kalman filter method above. Then, we have:*

$$\bar{X}_{k|k} \xrightarrow{L^p} X_{k|k},$$
$$\bar{P}_{k|k} \xrightarrow{L^p} P_{k|k},$$

as $N \to \infty$, where $X_{k|k}$ and $P_{k|k}$ are as in theorem 2.2.2.

Proof. See [MCB11] and [KM17]. □

Of course, one can not expect that a similar result holds in the general nonlinear case. For once, the fact that the EnKF is derived in the linear Gaussian setting means that applying it without modification to nonlinear problems correspond to consecutive Gaussian approximations of the probability distribution of the state. In complex scenarios, such an approximation may of course be very poor.

3.4 Ensemble smoothers

The ensemble approach can also be applied to the smoothing problem from section 2.5. For example, we can revisit the linear fixed-point smoother (proposition 2.5.1) with the recursion

$$x(\tau|t_k) = x(\tau|t_{k-1}) + \tilde{K}_k(y_k - M_k x_{k|k-1}). \tag{3.4.1}$$

Recall furthermore that we had

$$\tilde{K}_k = \Gamma(\tau, t_k|t_{k-1})^* M_k \left(M_k P_{k|k-1} M_k^* + R_k\right)^{-1}. \tag{3.4.2}$$

Since we had (see the proof of propositon 2.5.1)

$$\Gamma(\tau, t_k|t_{k-1}) = \mathrm{Cov}[X(\tau), X(t_k)|\boldsymbol{Y}_{k-1} = \boldsymbol{y}_{k-1}],$$

we can rewrite (3.4.2) as

$$\tilde{K}_k = \mathrm{Cov}[X(\tau), Y_k|\boldsymbol{Y}_{k-1} = \boldsymbol{y}_{k-1}] \left(\mathrm{Cov}(Y_k|\boldsymbol{Y}_{k-1} = \boldsymbol{y}_{k-1}) + R_k\right)^{-1}.$$

Applying the same reasoning as in the previous section, we can make an ensemble approximation of (3.4.1). We define an ensemble $(X^j(\tau|t_{k-1}))_{j=1}^N$ to represent our knowledge on $X(\tau)$, given $\boldsymbol{Y}_{k-1} = \boldsymbol{y}_{k-1}$, and set

$$X^j(\tau|t_k) = X^j(\tau|t_{k-1}) + D_{\tau|k-1} \otimes B_k \left(B_k \otimes B_k + (N-1)R_k\right)^{-1} (y_k - Y_k^j), \tag{3.4.3}$$

where Y_k^j and B_k are as in (3.3.3) and (3.3.7), respectively, and $D_{\tau|k-1}$ is defined as

$$D_{\tau|k-1} := \left[X^1(\tau|t_{k-1}) - \bar{X}(\tau|t_{k-1}) \quad \cdots \quad X^N(\tau|t_{k-1}) - \bar{X}(\tau|t_{k-1})\right].$$

In order to have access to all the quantities that are needed in the computation of (3.4.3), we have to keep track of two ensembles: On one hand, we have to evolve a fixed-time, smoothed ensemble $((X^j(\tau|t_k))$ with respect to new measurements. On the other hand, we have to evolve a filtering ensemble $(X_{k|k}^j)$ in order to generated new innovations $(y_k - y_k^j)$. We call the resulting algorithm the **ensemble Kalman smoother** (cf. [Raa16]). Similar to the EnKF, it can also be applied as an approximate algorithm for nonlinear filtering problems (see section 3.1). The details of the method can be taken from algorithm 4. Similarly to proposition 3.3.1, it can be shown (see [BGM19]) that in the linear case, the ensemble Kalman smoother converges to the linear fixed-point smoother (from proposition 2.5.1) in L^p , for all $p \in [1, \infty)$, as $N \to \infty$.

3.5 Ensemble Kalman inversion

The ensemble Kalman filter can also be applied to inverse problems that do not directly fit into the dynamic systems framework from chapter 2. Note that it was observed quite early (e.g. in [Fra70] and [Lar72]) that ill-posed inverse problems can be reformulated as parameter estimation problems in infinite-dimensional spaces, which gave rise to the probabilistic approach to inversion (see e.g. [Stu10]). For example, consider the following linear inverse problem in Hilbert spaces:

Problem 3.5.1. *Let H_1 and H_2 be separable real Hilbert spaces, $A : H_1 \to H_2$ a linear operator, $\mu_0 \in \mathcal{P}(H_1)$, and let $\eta \sim \mathcal{N}(0, \Gamma)$ be a Gaussian random element of H_2. Given the forward problem,*

$$Y = AU + \eta, \tag{3.5.1}$$

with $U \sim \mu_0$, determine $\mathbb{P}^{U|Y=y}$.

The basic strategy for applying the ensemble Kalman filter (or any other filtering technique) to this problem is to define an artificial dynamical system, for example like the following [ILS13]:

Definition 3.5.2 (Artificial dynamics for problem 3.5.1)**.** *Given problem 3.5.1, we define $H = H_1 \times H_2$, and denote elements $x \in H$ with*

$$x = \begin{bmatrix} u \\ p \end{bmatrix}, \qquad u \in H_1, p \in H_2.$$

Algorithm 4 `Ensemble Kalman smoother, EnKS`

Given problem 2.4.1, measurements $Y_1 = y_1, Y_2 = y_2, ...$, and ensemble size $N \in \mathbb{N}$.

1: For all $k \in \mathbb{N}_0$ with $t_k < \tau$, evolve the ensemble $(x^j_{k|k})$ using the EnKF (algorithm 3);

2: for all $j \in [N]$, solve

$$\begin{aligned} dx^j(t) &= f(x(t), t)dt + G(t)dw^j(t), \qquad t \in (t_{k-1}, \tau), \\ x^j(t_{k-1}) &= x^j_{k-1|k-1}, \end{aligned}$$

and set $x^j(\tau|t_{k-1}) = x^j(\tau)$;

3: **for** $t_{k-1} > \tau$ **do**

`EnKF step`

4: For all $j \in [N]$, solve

$$\begin{aligned} dx^j(t) &= f(x(t), t)dt + G(t)dw^j(t), \qquad t \in (t_{k-1}, t_k), \\ x^j(t_{k-1}) &= x^j_{k-1|k-1}. \end{aligned}$$

and set $x^j_{k|k-1} = x^j(t_k)$;

5: Compute $\bar{x}_{k|k-1} = \frac{1}{N}\sum_j x^j_{k|k-1}$;

6: Assemble $D_{k|k-1} = [x^1_{k|k-1} - \bar{x}_{k|k-1}, \ldots, x^N_{k|k-1} - \bar{x}_{k|k-1}]$;

7: Generate independent samples $v^1_k, \ldots, v^N_k \sim \mathcal{N}(0, R_k)$;

8: Set $y^j_k = m_k(x^j_{k|k-1})$ for all $j \in [N]$;

9: Compute $\bar{y}_k = \frac{1}{N}\sum_j y^j_k$;

10: Assemble $B_k = [y^1_k - \bar{y}_k, \ldots, y^N_k - \bar{y}_k]$;

11: Set $x^j_{k|k} = x^j_{k|k-1} + D_{k|k-1} \otimes B_k \left(B_k \otimes B_k + (N-1)R_k\right)^{-1} (y_k - y^j_k)$ for all $j \in [N]$;

`Smoothing step`

12: Compute $\bar{x}(\tau|t_{k-1}) = \frac{1}{N}\sum_j x^j(\tau|t_{k-1})$;

13: Set $D_{\tau|k-1} = [x^1(\tau|t_{k-1}) - \bar{x}(\tau|t_{k-1}), \ldots, x^N(\tau|t_{k-1}) - \bar{x}(\tau|t_{k-1})]$;

14: Set $x^j(\tau|t_k) = x^j(\tau|t_{k-1}) + D_{\tau|k-1} \otimes B_k(B_k \otimes B_k + (N-1)R_k)^{-1}(y_k - y^j_k)$, for all $j \in [N]$;

Let

$$\Xi : H \to H, \quad \Xi(x) := \Xi(u,p) := \begin{bmatrix} u \\ Au \end{bmatrix},$$

and, given the initial point $x_0 = \begin{bmatrix} u_0 \\ Au_0 \end{bmatrix} \in H$*, define the corresponding linear dynamical system*

$$\begin{aligned} X_k &= \Xi X_{k-1} \\ Y_k &= MX_k + V_k, \end{aligned}$$

where $M := \begin{bmatrix} 0 & \mathrm{Id} \end{bmatrix}$*, such that* $M \begin{bmatrix} u \\ p \end{bmatrix} = p$*, and with* $V_0, V_1, ... \sim \mathcal{N}(0,R)$ *independent.*

Note that we have $U_k = U$ for all $k \in \mathbb{N}$, so that the dynamical system is mainly a tool to provide a sequence of randomized measurements $(Y)_{k=1}^{\infty}$. A tuning parameter is the choice of the covariance R for the perturbation of the artificial measurements. Usually, one sets $R = \Gamma$. However different choices are possible and lead to algorithms with different properties (see [SS17] [SS18]). Applying the discrete version of the EnKF algorithm from section 3.3 to this dynamical system, using however the prior μ_0 on u for the generation of the initial ensemble, leads to the following algorithm (cf. [ILS13]):

Algorithm 5 `Ensemble Kalman inversion`

Given ensemble size $N \in \mathbb{N}$, a prior probability measure μ_0 on H_1, and a measurement $y \in H_2$.

1: sample $u_{0|0}^1, ..., u_{0|0}^N \sim \mu_0$ independently;
2: $x_{0|0}^j := \begin{bmatrix} u_{0|0}^j \\ Au_{0|0}^j \end{bmatrix}$, for all $j \in [N]$;
3: Set $P_0 = \frac{1}{N-1} \sum_{j=1}^N (x_{0|0}^j - \bar{x}_{0|0}) \otimes (x_{0|0}^j - \bar{x}_{0|0})$;
4: **for** $k = 1, 2, ...$ **do**
5: Set $x_{k|k-1}^j = \Xi(x_{k|k-1}^j)$, for all $j \in [N]$;
6:
7: set $P_{k|k-1} := \frac{1}{N-1} \sum_{j=1}^N (x_{k|k-1}^j - \bar{x}_{k|k-1}) \otimes (x_{k|k-1}^j - \bar{x}_{k|k-1})$;
8: Generate independent samples $y_k^1, ..., y_k^N \sim \mathcal{N}(y,R)$;
9: $K_n = P_n M^* (M P_n M^* + R)^{-1}$;
10: $x_{n|n}^j = x_{n|n-1} + K_n(y_n^j - M x_{n|n}^j)$, for $j = 1, ..., N$;
11: Assemble $D_{k|k-1} = [x_{k|k-1}^1 - \bar{x}_{k|k-1}, \ldots, x_{k|k-1}^N - \bar{x}_{k|k-1}]$;
12: Generate independent samples $v_k^1, ..., v_k^N \sim \mathcal{N}(y,R)$;
13: Set $y_k^j = M x_{k|k-1}^j + v_k^j$ for all $j \in [N]$;
14: Compute $\bar{y}_k = \frac{1}{N} \sum_j y_k^j$;
15: Assemble $B_k = [y_k^1 - \bar{y}_k, \ldots, y_k^N - \bar{y}_k]$;
16: Set $x_{k|k}^j = x_{k|k-1}^j + D_{k|k-1} \otimes B_k \left(B_k \otimes B_k + (N-1)R\right)^{-1} (y_k - y_k^j)$ for all $j \in [N]$;

An important property for the theoretical analyis of algorithm 5 is that all generated estimates remain in the linear subspace spanned by the initial ensemble $(u_{0|0}^j)$. Write

$$\begin{aligned} x_{n|n-1}^j &= \begin{bmatrix} u_{n|n-1}^j \\ p_{n|n-1}^j \end{bmatrix}, \\ x_{n|n}^j &= \begin{bmatrix} u_{n|n}^j \\ p_{n|n}^j \end{bmatrix}, \end{aligned}$$

Then, we have:

Theorem 3.5.3 (Invariant subspace property)**.** *Let* $G \subset H_1$ *be a linear subspace. Suppose that in algorithm 5, we have* $u_{0|0}^1, ..., u_{0|0}^N \in G$*. Then, we have* $u_{k|k-1}^j, u_{k|k}^j \in G$*, for all* $j \in [N]$ *and all* $k \in \mathbb{N}$*.*

Proof. Induction, see [ILS13]. □

As a consequence, if G is a finite-dimensional subspace, this theorem guarantees (by linearity of G), that all ensemble members $x^j_{k|k}$ stay in a finite-dimensional subspace of H. For the present algorithm, we can trivially guarantee that the initial ensemble lies in a convenient finite-dimensional subspace G, simply by setting

$$G = \operatorname{span}\{\, u^j_{0|0} : j \in [N] \,\}. \tag{3.5.2}$$

After iteration n, one can estimate $U_k = U$ by the current ensemble mean, and hence we obtain a sequence of estimators

$$\bar{u}_k := \frac{1}{N} \sum_{j=1}^{N} u^j_{k|k}. \tag{3.5.3}$$

The invariant subspace property then trivially yields a first lower bound.

Corollary 3.5.4. *Assume that in algorithm 5 the initial ensemble satisfies $u^1_{0|0}, ..., u^N_{0|0} \in G$. Then:*

$$\|u - \bar{u}_n\| \geq \inf_{v \in G} \|u - v\|,$$

for all $n \in \mathbb{N}$.

Proof. Obvious from theorem 3.5.3. □

Moreover, ensemble Kalman inversion can be shown to converge in the sense of classical regularization (for example as in [EHN00]). Consider the Tikhonov-Phillips regularized error functional (cf. [ILS13]):

$$I(u) := \|y - Au\|^2_\Gamma + \|u - \bar{u}\|^2_\Sigma, \tag{3.5.4}$$

where Γ is as in problem 3.5.1, and where $\bar{u}$ and Σ denote the mean and covariance of μ_0 (which we assume to exist in the following). We denote the corresponding minimizer with

$$u_{TP} := \operatorname*{argmin}_{u \in H_1} I(u).$$

Then, there exists a relation between u_{TP} and the ensemble Kalman estimators $\bar{u}_k$. One has the following result:

Theorem 3.5.5 (Large ensemble limit)**.** *Given problem 3.5.1, let $\mu_0 = \mathcal{N}(\bar{u}, \Sigma)$, and let $u^j_{k|k}$ denote the ensemble members generated by a realization of algorithm 5 using $R = \Gamma$, for $j \in [N]$ and $k \in \mathbb{N}_0$. Let $\bar{u}_k := N^{-1} \sum_{j=1}^N u^j_{k|k}$. Then*

$$\bar{u}_1 \longrightarrow u_{TP}, \text{ as } N \to \infty,$$

in L^p, for all $p \in [1, \infty)$.

Proof. One can show that in the linear Gaussian case, the Kalman filter estimate $u_{1|1}$ minimizes the regularized cost functional I (see [ILS13], and section 5.1), meaning that

$$u_{1|1} = u_{TP}.$$

In our case, the Kalman filter estimate is given by (cf. (2.2.8)

$$u_{1|1} = \bar{u} + \Sigma A^*(A\Sigma A^* + \Gamma)^{-1}(y - A\bar{u}).$$

On the other hand, we have by proposition 3.3.1

$$\bar{u}_1 \underset{N \to \infty}{\longrightarrow} u_{1|1} = u_{TP},$$

in L^p, for all $p \in [1, \infty)$. □

3.6 Comments and further references

Our presentation of the EKF and the IEKF is largely based on [Jaz70] (where the iterated EKF is also called "local iteration"). There exists a whole number of modifications and further extensions of the Kalman filter that we did not address: The *iterated linear filter-smoother* [Jaz70] further builds on the idea of the IEKF by improving also the time-update $x_{k|k-1}$ in every iteration through a Rauch-Tung-Striebel-smoothing step (cf. theorem 2.5.2), which yields better performance in the presence of strong state nonlinearities.

Theoretical analysis of the extended Kalman filter and its relatives is still an active area of research. It is not clear how to give a rigorous interpretation to the Taylor-like approximations made in the derivation of the EKF (section 3.1) and the IEKF (section 3.2). This prohibits a clear interpretation of the EKF beyond the fact that it coincides with the Kalman-Bucy filter if both the state and the measurement model are linear. General convergence results for the iterated extended Kalman filter are still missing from the literature. Most articles on the convergence analysis of the extended Kalman filter do this only restricted to special cases (see e.g. [Kre03], [Gou+18], [SG07], [AF05], or [Alo+15]). On the theoretical side, it has been shown in [Pic91] that the EKF is a good approximation of the optimal nonlinear filter if knowledge on the initial state is precise, and if the system is "almost linear" in a specific sense. Similarly, it was shown in [Rei+99] and [Rei+00] that the estimation error of the EKF remains bounded if the initial guess is good and the disturbing noise is small. Recently, the extended Kalman filter has also been investigated from an information-geometric point of view, where it has been shown to be equivalent to online natural gradient descent [Oll18].

However, if one considers the iterated EKF purely as an optimization method, one can show that it is equivalent to the regularized Gauss-Newton algorithm (see section 5.1). Thus, if one is only interested in a purely deterministic analysis, the convergence theory of the Gauss-Newton method applies.

The lack of rigorous theoretical results notwithstanding, the EKF is an extremely successful numerical technique. Its use is widespread in electrical engineering and especially in control and automatization. For example, the EKF has been applied to satellite control [CTL00], autonomous driving [], and sensorless control ([Jan06] and [Alo+15]), to name a few. Further applications exist to economics (see e.g. [Wal02]), finance [Nir], computational biology [CP16], chemical engineering [KK15], and even to tomography [Pin+18] and other branches of medicine [MA08].

Finally, an alternative to the EKF is the **unscented Kalman filter** [JU97]. In practice, it has been observed to perform better for nonlinear systems while having the same computational complexity [WV00]. It is based on the idea of "sigma points". A finite-dimensional Gaussian probability density can be fully represented by a finite number of well-chosen vectors. Instead of evolving expectation and covariance estimates using a linear approximation of the system, the unscented Kalman filter applies the nonlinear state and measurement equations to the sigma-points and then estimates the expectation and covariance by sample averages. In particular, differentiation of the state and measurement functions is not needed.

The ensemble Kalman filter was developed for oceanographic problems (see [Eve94]), where the underlying state system is notoriously high-dimensional. Since its conception, the EnKF has been very successful in the numerical weather prediction literature, where filtering methods can be found under the name "data assimilation" (see e.g. [HZ16] and [BEW03]). However, due to its generality and practicality, the EnKF had soon left its original field of application. Nowadays, there exist applications in the oil industry (see [AS06] and [CTS10]), signal processing [Rot+17], economics [KS12] and tomography [But+09].

For highly nonlinear systems, combinations of the ensemble Kalman filter with variational methods have been studied (see [SOB12], [SHB18], [Oli17], [Eve+19], [ER13], [CO12], and [RSE19]). Usually, the basic ensemble approach is combined with an iteration such that state and measurement nonlinearites can better be accounted for. Recall the measurement iteration for the IEKF,

$$x_{k|k}^{n+1} = x_{k|k-1} + K_k^n \left(y_k - m_k(x_{k|k}^n) - M_k^n (x_{k|k-1} - x_{k|k}^n) \right),$$

with

$$K_k^n = P_{k|k-1}(M_k^n)^*(M_k^n P_{k|k-1}(M_k^n)^* + R_k)^{-1}.$$

A first idea would be to use the IEKF iteration (see section 3.2) on each ensemble member separately, meaning that we introduce the iteration

$$x_{k|k}^{n+1,j} = x_{k|k-1} + K_k^{n,j}(y_k^j - m_k(x_{k|k}^{n,j}) - M_k^{n,j}(x_{k|k-1} - x_{k|k}^{n,j})), \tag{3.6.1}$$

for each $j \in [N]$ separately, where

$$K_k^{n,j} := \bar{P}_{k|k-1}(M_k^{n,j})^*(M_k^{n,j} P_{k|k-1} M_k^{n,j} + R_k)^{-1},$$
$$\bar{P}_{k|k-1} := \frac{1}{N-1} D_{k|k-1} \otimes D_{k|k-1},$$

(see section 3.3) and

$$M_k^{n,j} := Dm_k(x_{k|k-1}^{n,j}).$$

This method is originally known as "randomized maximum likelihood". This name stems from the interpretation of the IEKF as Gauss-Newton method as described in section 5.1. For the finite-dimensional Gaussian filtering problem (problem 2.4.1 with $H = \mathbb{R}^d$), sequential minimization of the regularized cost functional

$$\Phi_k(x) = \|y_k - m_k(x)\|_R^2 + \|x - x_{k|k-1}\|_{P_{k|k-1}}^2$$

can be interpreted as finding the maximum a-posteriori (MAP) estimator $\hat{x}_k := \operatorname{argmin} \Phi_k(x)$ of the state $X(t_k)$ given measurements $\boldsymbol{Y}_k = \boldsymbol{y}_k$. The cost functional Φ_k itself corresponds to the logarithmic probability density of X_k given $\boldsymbol{Y}_k = \boldsymbol{y}_k$.

The iteration (3.6.1) can then be interpreted as Gauss-Newton iteration for the "randomized log posterior"

$$\tilde{\Phi}_k(x) = \|y_k^j - m_k(x)\|_R^2 + \|x - x_{k|k-1}^j\|_{\bar{P}_{k|k-1}},$$

where we use the perturbed measurement $y_k^j \sim (y_k, R_k)$ instead of the actual measurement y and the ensemble forecast $x_{k|k-1}^j$ instead of the conditional mean $x_{k|k-1} = \mathbb{E}[X(t_k)|\boldsymbol{Y}_k = \boldsymbol{y}_k]$.

In an application, the iteration (3.6.1) would then be repeated until convergence for each ensemble member separately. Note that this includes having to evaluate the gradient $M_k^{n,j} = Dm_k(x_{k|k-1}^{n,j})$ in each iteration for each ensemble member separately. If the computational cost of evaluating Dm_k is an issue, this might be impractical. Due to this, Gu and Reynolds [GO07] proposed to use in each update (3.6.1) the same ensemble-based approximation $\bar{M}_k^n$ (to be defined below) instead of $M_k^{n,j}$. They called the resulting iterative ensemble Kalman filter the *ensemble randomized maximum likelihood* (EnRML) method.

Suppose after the $(n-1)$st iteration we have the ensemble $(x_{k|k-1}^{n,j})$. Define

$$\boldsymbol{m}_k^n := \begin{bmatrix} m_k(x_{k|k}^{n,1}) & \dots & m_k(x_{k|k}^{n,N}) \end{bmatrix},$$

and

$$\boldsymbol{X}_{k|k}^n := \begin{bmatrix} x_{k|k}^{n-1,1} - \bar{x}_{k|k}^n & \dots & x_{k|k}^{n,N} - \bar{x}_{k|k}^n \end{bmatrix},$$

where $\bar{x}_{k|k}^n := N^{-1} \sum_j x_{k|k}^{n,j}$ is the corresponding ensemble average. We then define the estimator

$$\bar{M}_k^n := \boldsymbol{M}_k(\boldsymbol{X}_{k|k}^n)^+, \tag{3.6.2}$$

where for a matrix A, A^+ denotes the Moore-Penrose pseudoinverse (see e.g. [TI97]).

It can be shown that this estimator is asymptotically consistent [RSE19, theorem 1].

Chapter 4

Real-time parameter estimation

4.1 Filtering of uncertain systems

In the preceding chapters, we always assumed that our uncertainty about the system

$$\begin{aligned} dX(t) &= f(X(t), t)dt + G(t)dW(t), \\ X(0) &= X_0, \end{aligned}$$

only comes from the uncertainty about the initial state X_0 and the intrinsic noise $G(t)dW(t)$. However, we are often faced with a situation where the we have an additional model uncertainty, represented by an unknown parameter θ. That is, we assume that our state model is

$$\begin{aligned} dX(t) &= f(X(t), \theta, t)dt + G(t)dW(t), \\ X(0) &= X_0. \end{aligned}$$

Similarly, we can also assume uncertainty in our measurement model by letting the measurement functions m_k also depend on θ, such that the measurement equation becomes

$$Y_k = m_k(X(t_k), \theta) + V_k.$$

Hence, in this new system all inference on the state X is influenced by our uncertainty on the parameter θ. And all inference on the parameter θ is influenced by the uncertainty in the state, that comes for example from the intrinsic noise $G(t)dW(t)$. Thus, we are faced with the problem of simultaneous estimation of the state and parameter in dynamical systems.

Problem 4.1.1 (Filtering in an uncertain system)**.** *Let H, U be separable Hilbert spaces, and Θ be a separable Banach space. Consider the observed system*

$$U(t) = f(U(t), \theta, t)dt + G(t)dW(t), \tag{4.1.1a}$$

$$U(0) = U_0, \tag{4.1.1b}$$

$$Y_k = m_k(U(t_k), \theta) + V_k, \qquad k \in \mathbb{N} \tag{4.1.1c}$$

where the initial value $U(0) \sim \mathcal{N}(x_0, P_0)$ is a Gaussian random element of H, $\theta \sim \mathcal{N}(\bar{\vartheta}, \Gamma)$ is a Gaussian random element of Θ, $f \in M(H \times \Theta \times [0,T]; H)$ is a measurable map, $G \in M([0,T]; \mathcal{L}_2(Q^{1/2}U; H))$ is a function of operators, W is an independent U-valued Q-Wiener process on $[0,T]$, and for each $k \in \mathbb{N}$, $m_k : H \times \Theta \to \mathbb{R}^m$ denotes the measurement function, with independent Gaussian noise $V_k \sim \mathcal{N}(0, R_k)$. The filtering problem is then to determine the conditional distributions of $U(t_k)$ and θ, given measurements $\boldsymbol{Y}_k = \boldsymbol{y}_k$,

$$\mathbb{P}^{U(t_k), \theta | \boldsymbol{Y}_k = \boldsymbol{y}_k},$$

for $k \in \mathbb{N}$.

As in the standard filtering problem, exact solution of the dual filtering problem is usually infeasible, except for a few special cases. Therefore, in general one settles for finding suboptimal filters, for example by computing approximations for the mean and covariance. In the following sections, we will present how the methods from chapter 3 can be modified in order to yield approximate solutions to the dual filtering problem 4.1.1.

4.2 The joint extended Kalman filter for parameter estimation

With the extended filtering techniques like the EKF (section 3.1) and the IEKF (section 3.2) in mind, an obvious approach is to consider the parameter simply as an additional, constant state, with the trivial dynamics $D_t\theta(t) = 0$ (cf. [Jaz70, section 8.4]). This means we define a new state variable

$$X(t) = \begin{bmatrix} U(t) \\ \theta(t) \end{bmatrix}$$

and then consider the extended system:

$$dX(t) = \begin{bmatrix} f(X^1(t), X^2(t), t) & 0 \\ 0 & 0 \end{bmatrix} dt + \begin{bmatrix} G(t) \\ 0 \end{bmatrix} dW(t), \tag{4.2.1a}$$

$$X(t_0) = \begin{bmatrix} U_0 \\ \theta_0 \end{bmatrix}, \tag{4.2.1b}$$

$$Y_k = m_k(X^1(t_k), X^2(t_k)) + V_k. \tag{4.2.1c}$$

Note that while we assumed in the derivation of the EKF and IEKF that the state space is Hilbert. However, this was only in order to deal with stochastic evolution of the state in continuous time. The discrete-time Kalman update holds for separable Banach spaces (see section 2.2. Since the parameter θ has the deterministic dynamics $D_t\theta(t) = 0$, the algorithms from chapter 3 still apply, even though θ takes values in a Banach space.

Thus, we can simply apply the EKF or IEKF to the system (4.2.1). This is also known as the *augmented state approach* [NS76]. After every step, it provides approximations of the mean and the covariance for the state and parameter

$$x_{k|k} = \begin{bmatrix} u_{k|k} \\ \theta_{k|k} \end{bmatrix},$$
$$P_{k|k} = \begin{bmatrix} \Sigma_{k|k} & C_{k|k} \\ C^*_{k|k} & \Gamma_{k|k} \end{bmatrix},$$

where

$$\begin{aligned} u_{k|k} &\approx \mathbb{E}[U(t_k)|\boldsymbol{Y}_k = \boldsymbol{y}_k], \\ \theta_{k|k} &\approx \mathbb{E}[\theta|\boldsymbol{Y}_k = \boldsymbol{y}_k], \\ \Sigma_{k|k} &\approx \mathrm{Cov}(U(t_k)|\boldsymbol{Y}_k = \boldsymbol{y}_k), \\ C_{k|k} &\approx \mathrm{Cov}[U(t_k), \theta|\boldsymbol{Y}_k = \boldsymbol{y}_k], \\ \Gamma_{k|k} &\approx \mathrm{Cov}(\theta|\boldsymbol{Y}_k = \boldsymbol{y}_k). \end{aligned}$$

4.3 A dual iterated extended Kalman filter

A different approach is provided by so-called dual methods, which solve the dual filtering problem by separating it into two coupled filtering problems (see [NS76], [Nel00], and [WN01]).

In the following, we present a dual iterated extended Kalman filter. In this method, we consider the system (4.1.1) as an inference problem purely on the unknown parameter θ. We choose a strictly sequential approach (in contrast to the dual EKF in [Nel00] that leads to a global recursion). Suppose after the first $k-1$ measurements, we have a deterministic estimate ϑ_{k-1} for the parameter (in the

following, we will use θ to denote the parameter as a random element of Θ, and $\vartheta \in \Theta$ for its possible realizations), and $u_{k-1|k-1}$ for the state U at time t_{k-1}, together with respective covariances T_{k-1} and $\Sigma_{k-1|k-1}$. The state $U(t_k)$ at the next timepoint can be considered simply as a random function of θ:

$$U(t_k|\theta) = U(t_{k-1}) + \int_{t_{k-1}}^{t_k} f(U(t|\theta), \theta, t)dt + \int_{t_{k-1}}^{t_k} G(t)dW(t). \tag{4.3.1}$$

For a practical algorithm, we need a deterministic estimate (an estimate that is a deterministic function of θ) of $U(t_k|\theta)$. This can be achieved by applying the time update of the extended Kalman filter to the filtering problem

$$\begin{aligned} dU(t) &= f(U(t), \vartheta, t)dt + G(t)dW(t), \\ U(t_{k-1}) &= U_{k-1}, \\ Y_k &= m_k(U(t_k), \vartheta) + V_k, \end{aligned}$$

where $\vartheta \in \Theta$ is considered as a given, fixed parameter, and $U_{k-1} \sim \mathcal{N}(u_{k-1|k-1}, \Sigma_{k-1|k-1})$ represents our information on $U(t_{k-1})$ based on the first $k-1$ measurements. The EKF time update then yields an estimate $u_{k|k-1}(\vartheta)$ of $U(t_k|\vartheta)$, depending on ϑ, with corresponding covariance $\Sigma_{k|k-1}(\vartheta) \approx \mathrm{Cov}(U(t_k)|\boldsymbol{Y}_{k-1} = \boldsymbol{y}_{k-1}, \theta = \vartheta)$ (see section 3.1). Substituting $U(t_k|\theta)$ with $u_{k|k-1}(\theta)$, we obtain a deterministic (given θ) approximation of (4.1.1c):

$$Y_k \approx m_k(u_{k|k-1}(\theta), \theta) + V_k.$$

Thus, we can formulate a continuous-discrete filtering problem in θ:

$$\begin{aligned} &\frac{d}{dt}\theta(t) = 0, \\ &\theta(0) = \theta_0, \\ &Y_k = m_k(u_{k|k-1}(\theta(t)), \theta(t)) + V_k, \end{aligned} \tag{4.3.2}$$

with $\theta_0 \sim \mathcal{N}(\bar{\vartheta}, \Gamma)$. Similarly to section 2.2, we set

$$\begin{aligned} \vartheta_{n|k} &:= \mathbb{E}[\theta(t_n)|\boldsymbol{Y}_k = \boldsymbol{y}_k] = \mathbb{E}[\theta|\boldsymbol{Y}_k = \boldsymbol{y}_k], \\ \Gamma_{n|k} &:= \mathrm{Cov}(\theta(t_n)|\boldsymbol{Y}_k = \boldsymbol{y}_k) = \mathrm{Cov}(\theta|\boldsymbol{Y}_k = \boldsymbol{y}_k), \end{aligned}$$

where in both lines the second equalities are due to (4.3.2).

Due to (4.3.2), the time update for the parameter is trivial:

$$\begin{aligned} \vartheta_{k|k-1} &= \vartheta_{k-1|k-1}, \\ \Gamma_{k|k-1} &= \Gamma_{k-1|k-1}. \end{aligned} \tag{4.3.3}$$

Applying the IEKF (see 3.2.13) to the filtering problem 4.1.1, we obtain the iteration

$$\vartheta_{k|k}^{n+1} = \vartheta_{k|k-1} + G_k^n \left(y_k - m_k(u_{k|k-1}(\vartheta_{k|k}^n), \vartheta_{k|k}^n) - N_k^n(\vartheta_{k-1} - \vartheta_{k|k}^n) \right), \tag{4.3.4}$$

where

$$G_k^n := \Gamma_{k-1|k-1}(N_k^n)^*(N_k^n\Gamma_{k-1|k-1}(N_k^n)^* + R_k)^{-1} \tag{4.3.5}$$

denotes the Kalman gain for the parameter θ, and

$$N_k^n := Dm_k(u_{k|k-1}(\vartheta), \vartheta)\Big|_{\vartheta=\vartheta_{k|k}^n} \tag{4.3.6}$$

denotes the linearization of the measurement function at $\theta = \vartheta_{k|k}^n$ (see below). After convergence, the covariance estimate can be updated through the formula

$$\Gamma_{k|k} = (\mathrm{Id} - G_k^n N_k^n)\Gamma_{k|k-1},$$

using the highest available n.

Since it is also necessary to calculate the EKF estimates $u_{k|k}$, this leads to a two-stage method, where one alternates between estimating the state U and the parameter θ. Of course, the dual filter can be combined with an iterated EKF for the state. The resulting **dual iterated extended Kalman filter** (dual IEKF) is described in detail in algorithm 6. The algorithm looks complicated but is able to adress nonlinear dependence of the measurements both on the state and parameter. For a new measurement $Y_k = y_k$, the parameter is updated using an IEKF. Since the k-th measurement depends implicitly on the parameter through the state $U(t_k)$, it is necessary to recompute the state prediction $u_{k|k-1}(\vartheta^n_{k|k})$ for each parameter update $\vartheta^n_{k|k}$. Even if the measurement function m_k is linear, the state may depend on the parameter in a nonlinear way, which is a reason for why it might be a good idea to do multiple EKF-updates for the parameter. Then, if the parameter estimate has converged, the state is updated through a second IEKF, where the latest parameter estimate is fixed. If the measurement function is linear in the state, only one iteration is necessary here.

Let us take a closer look at how to compute the matrices N^n_k needed in the parameter update (4.3.4). We have by the chain rule

$$D_\vartheta m_k(u_{k|k-1}(\vartheta), \vartheta)\Big|_{\vartheta=\vartheta^n_{k|k}} = \partial_1 m_k(u_{k|k-1}(\vartheta^n_{k|k}), \vartheta^n_{k|k}) D_\theta u_{k|k-1}(\vartheta^n_{k|k}) + \partial_2 m_k(u_{k|k-1}(\vartheta^n_k), \vartheta^n_k).$$

In order to compute the derivative of the state forecast $u_{k|k-1}$ with respect to the parameter, note that we have from the EKF time update (3.1.9) that $u_{k|k-1}(\vartheta) = u(t_k|t_{k-1}, \vartheta)$ is given by

$$u(t_k|t_{k-1}, \vartheta) = u_{k-1|k-1} + \int_{t_{k-1}}^{t_k} f(u(t|t_{k-1}, \vartheta), \vartheta, t)dt.$$

Thus, we have

$$D_\vartheta u(t_k|\vartheta, t_{k-1})\Big|_{\vartheta=\vartheta^n_{k|k}} = \int_{t_{k-1}}^{t_k} \left[\partial_1 f(u(t|\vartheta^n_{k|k}, t_{k-1}), \vartheta^n_{k|k}, t) D_\vartheta u(t_k|\vartheta^n_{k|k}, t_{k-1}) + \partial_2 f(u(t|\vartheta^n_{k|k}, t_{k-1}), \vartheta^n_{k|k}, t)\right] dt,$$

Equivalently, $\Lambda(t_k|\vartheta^n_{k|k}, t_{k-1}) := D_\vartheta u(t_k|\vartheta, t_{k-1})\Big|_{\vartheta=\vartheta^n_{k|k}}$ can be obtained as solution of the initial value problem

$$D_t\Lambda(t|\vartheta^n_{k|k}, t_{k-1}) = \partial_1 f(u(t|\vartheta^n_{k|k}, t_{k-1}), \vartheta^n_{k|k}, t)\Lambda(t|\vartheta^n_{k|k}, t_{k-1}) + \partial_2 f(u(t|\vartheta^n_{k|k}, t_{k-1}), \vartheta^n_{k|k}, t), \qquad t \in (t_{k-1}, t_k), \tag{4.3.7a}$$

$$\Lambda(t_{k-1}|\vartheta^n_{k|k}, t_{k-1}) = 0. \tag{4.3.7b}$$

Note that the simplifications in this approach lead to the behavior that the update $\vartheta_{k-1|k-1} \to \vartheta_{k|k}$ of the parameter only depends on the change of the system in the interval $[t_{k-1}, t_k]$. Thus, this approach only works if the parameter θ affects the behavior of the system in this time interval. Otherwise, we have $\vartheta_{k|k} = \vartheta_{k|k-1} = \vartheta_{k-1|k-1}$. *This makes the approach for example ill-suited for the estimation of initial conditions.* However, the dual IEKF (or its single-update version, the dual EKF) has the advantage that by modifying the parameter dynamics (4.3.2) it is even applicable to time-varying parameters (see also section 4.5).

4.4 A dual ensemble Kalman filter

The **dual ensemble Kalman filter** (dual EnKF) tries to apply the ideas from section 3.3 to the dual filtering problem. The difference to the standard ensemble Kalman filter as described in algorithm 3 is that we now have to evolve two ensembles $(U^j_{k|k})_{j=1}^N$ and $(\theta^j_{k|k})_{j=1}^N$ to represent the conditional distributions of the state and the parameter, respectively. Apart from that, the idea is mostly the same.

Algorithm 6 `Dual iterated extended Kalman filter`

Given problem 2.4.1 with measurements $Y_1 = y_1, Y_2 = y_2, ...$

Initialize $u_{0|0} = u_0$, $\Sigma_{0|0} = \Sigma_0$, $\vartheta_0 = \bar{\theta}$, $\Gamma_{0|0} = \Gamma$;
for k = 1,2,... **do**
  Set $\vartheta^0_{|k} = \vartheta_{k-1|k-1}$;
  Set $n = 0$;
  repeat
  `EKF time update for the state:`
    Solve

$$D_t u(t|t_{k-1}) = f(u(t|t_{k-1}), \vartheta^n_k, t), \qquad t \in (t_{k-1}, t_k),$$
$$u(t_{k-1}|t_{k-1}) = u_{k-1|k-1}.$$

  and

$$D_t \Lambda(t|\vartheta^n_{k|k}, t_{k-1}) = \partial_1 f(u(t|\vartheta^n_{k|k}, t_{k-1}), \vartheta^n_{k|k}, t)\Lambda(t|\vartheta^n_{k|k}, t_{k-1}) + \partial_2 f(u(t|\vartheta^n_{k|k}, t_{k-1}), \vartheta^n_{k|k}, t), \qquad t \in [t_{k-1}, t_k],$$
$$\Lambda(t_{k-1}|\vartheta^n_{k|k}, t_{k-1}) = 0,$$

  simultaneously, and set $u^n_{k|k-1} = u(t_k|t_{k-1})$ and $\Lambda^n_{k|k-1} = \Lambda(t_k|\vartheta^n_{k|k}, t_{k-1})$;
  `IEKF measurement update for the parameter:`
    Compute $N^n_k = \partial_1 m_k(u^n_{k|k-1}, \vartheta^n_{k|k})\Lambda^n_{k|k} + \partial_2 m_k(u^n_{k|k-1}, \vartheta^n_{k|k})$;
    Set $G^n_k = \Gamma_{k|k-1}(N^n_k)^*(N^n_k\Gamma_{k|k-1}(N^n_k)^* + R_k)^{-1}$;
    Set $\vartheta^{n+1}_{k|k} = \vartheta^n_{k|k} + G^n_k\left(y_k - m_k(u^n_{k|k-1}, \vartheta^n_{k|k}) - N^n_k(\vartheta_{k|k-1} - \vartheta^n_{k|k})\right)$;
    $n \leftarrow n + 1$;
  until convergence
  Set $\vartheta_{k|k} = \vartheta^n_k$;
  Set $\Gamma_{k|k} = (\mathrm{Id} - G^n_k N^n_k)\Gamma_{k|k-1}$;
  `Recompute the time-update for the state using the final parameter estimate:`
  Solve

$$D_t u(t|t_{k-1}) = f(u(t|t_{k-1}), \vartheta_{k|k}, t), \qquad t \in (t_{k-1}, t_k),$$
$$u(t_{k-1}|t_{k-1}) = u_{k-1|k-1},$$

  and set $u_{k|k-1} = u(t_k|t_{k-1})$;
  Solve

$$D_t \Sigma(t|t_{k-1}) = F_1(t|t_{k-1})\Sigma(t|t_{k-1}) + \Sigma(t|t_{k-1})F_1(t|t_{k-1})^* + G(t)QG(t)^*, \qquad t \in (t_{k-1}, t_k),$$
$$\Sigma(t_{k-1}|t_{k-1}) = \Sigma_{k-1|k-1},$$

  with $F_1(t|t_{k-1}) = \partial_1 f(u(t|t_k), \theta_k, t)$, and set $\Sigma_{k|k-1} = \Sigma(t_k|t_{k-1})$;
  `IEKF measurement update for the state:`
  Set $u^0_{k|k} = u_{k|k-1}$;
  Set $n = 0$;
  repeat
    Set $M^n_k = \partial_1 m_k(u^n_{k|k}, \vartheta_{k|k})$;
    Set $K^n_k = \Sigma_{k|k-1}(M^n_k)^*(M^n_k\Sigma_{k|k-1}(M^n_k)^* + R_k)^{-1}$;
    Set $u^{n+1}_{k|k} = u_{k|k-1} + K^n_k\left(y_k - m_k(u^n_{k|k}, \vartheta_{k|k}) - M^n_k(u_{k|k-1} - u^n_{k|k})\right)$;
    $n \leftarrow n + 1$;
  until convergence
  Set $u_{k|k} = u^n_{k|k}$;
  Set $\Sigma_{k|k} = (\mathrm{Id} - K^n_k M^n_k)\Sigma_{k|k-1}$;

As in the dual IEKF, we assume the model

$$dU(t) = f(U(t), \theta(t), t)dt + G(t)dW(t), \tag{4.4.1a}$$
$$U(0) = U_0, \tag{4.4.1b}$$
$$\frac{d}{dt}\theta(t) = 0, \tag{4.4.1c}$$
$$\theta(0) = \theta_0, \tag{4.4.1d}$$
$$Y_k = m_k(U(t), \theta(t)) + V_k. \tag{4.4.1e}$$

Suppose we have

$$U^j_{k-1|k-1} \overset{d}{=} (U(t_{k-1})|\boldsymbol{Y}_{k-1} = \boldsymbol{y}_{k-1}),$$
$$\theta^j_{k-1|k-1} \overset{d}{=} (\theta|\boldsymbol{Y}_{k-1} = \boldsymbol{y}_{k-1}).$$

Then, if we evolve U^j according to (4.1.1a) by solving

$$dU^j(t) = f(U^j(t), \theta^j, t)dt + G(t)dW^j(t),$$
$$U^j(t_{k-1}) = U^j_{k-1|k-1},$$

and set $U^j_{k|k-1} := U^j(t_k)$, we have

$$U^j_{k|k-1} \overset{d}{=} (U(t_k)|\boldsymbol{Y}_{k-1} = \boldsymbol{y}_{k-1}).$$

Similarly, we set

$$\theta^j_{k|k-1} = \theta^j_{k-1|k-1}$$

as in (4.3.3).

In correspondence to section 3.3, we generate perturbed measurements

$$Y^j_k = m_k(U^j_{k|k-1}, \theta^j_{k|k-1}) + V^j_k,$$

where $V^1_k \overset{d}{=} \cdots \overset{d}{=} V^N_k \overset{d}{=} V_k$ are mutually independent. Now, we make the strong assumption that θ, Y_k and $U(t_k)$ are jointly Gaussian distributed.

Then, we have in analogy to (3.3.4):

$$(\theta|\boldsymbol{Y}_{k-1} = \boldsymbol{y}_{k-1}, Y_k = y_k) \sim \mathcal{N}(\vartheta_{k|k-1} - \Sigma_{\theta,Y_k|k-1}(\Sigma_{Y_k,Y_k|k-1})^{-1}(y_k - \bar{y}_k),$$
$$\Sigma_{\theta,\theta|k-1} - \Sigma_{\theta,Y_k|k-1}(\Sigma_{Y_k,Y_k|k-1})^{-1}\Sigma_{Y_k,\theta|k-1}),$$

where we used definition (3.3.5) and recalled $\vartheta_{k|k-1} = \mathbb{E}[\theta|\boldsymbol{Y}_{k-1} = \boldsymbol{y}_{k-1}]$. This motivates us to define a Monte Carlo approximated ensemble update (cf. (3.3.8))

$$\theta^j_{k|k} = \theta^j_{k|k-1} + D^\theta_{k|k-1} \otimes B_k \left(B_k \otimes B_k + (N-1)R_k\right)^{-1},$$

where the collection of measurement deviations B_k is defined as in (3.3.7) and

$$D^\theta_{k|k-1} := \begin{bmatrix} \theta^1_{k|k-1} - \bar{\theta}_{k|k-1} & \cdots & \theta^N_{k|k-1} - \bar{\theta}_{k|k-1} \end{bmatrix}$$

are the deviations of the parameter ensemble, with the ensemble mean

$$\bar{\theta}_{k|k-1} := \frac{1}{N-1}\sum_{j=1}^{N} \theta^j_{k|k-1}.$$

Combining this with an EnKF update for the state ensemble $(U^j_{k|k-1})$ leads to a dual version of the ensemble Kalman filter, described in algorithm 7.

Algorithm 7 `Dual Ensemble Kalman filter, dual EnKF`

Given problem 2.4.1, measurements $\boldsymbol{Y}_k = \boldsymbol{y}_k$, and ensemble size $N \in \mathbb{N}$.

Generate independen samples $u^1_{0|0}, ..., u^N_{0|0} \sim \mathcal{N}(u_0, \Sigma_0)$;
Generate independent samples $\vartheta^1_{0|0}, \ldots, \vartheta^N_{0|0} \sim \mathcal{N}(\bar{\vartheta}, \Gamma)$;
for $k = 1, 2, ...$ **do**
    For all $j \in [N]$, solve

$$\begin{aligned} du^j(t) &= f(u^j(t), \theta^j, t)dt + G(t)dw^j(t), \qquad t \in (t_{k-1}, t_k), \\ u^j(t_{k-1}) &= u^j_{k-1|k-1}, \end{aligned}$$

    and set $u^j_{k|k-1} = u^j(t_k)$;
    Set $\vartheta^j_{k|k-1} = \vartheta^j_{k-1|k-1}$ for all $j \in [N]$;
    Compute $\bar{u}_{k|k-1} = \frac{1}{N}\sum_j u^j_{k|k-1}$;
    Compute $\bar{\vartheta}_{k|k-1} = \frac{1}{N}\sum_j \vartheta^j_{k|k-1}$;
    Set $D^u_{k|k-1} = [u^1_{k|k-1} - \bar{u}_{k|k-1}, \ldots, u^N_{k|k-1} - \bar{u}_{k|k-1}]$;
    Set $D^\vartheta_{k|k-1} = [\vartheta^1_{k|k-1} - \bar{\vartheta}_{k|k-1}, \ldots, \vartheta^N_{k|k-1} - \bar{\vartheta}_{k|k-1}]$;
    Generate independent samples $v^1_k, \ldots, v^j_k \sim \mathcal{N}(0, R_k)$;
    Set $y^j_k = m_k(u^j_{k|k-1}, \vartheta^j_{k|k-1}) + v^j_k$ for $j \in [N]$;
    Compute $\bar{y}_k = \frac{1}{N}\sum_j y^j_k$;
    Set $B_k = [y^1_k - \bar{y}_k, \ldots, y^N_k - \bar{y}_k]$;
    Update $u^j_{k|k} = u^j_{k|k-1} + D^u_{k|k-1} \otimes B_k (B_k \otimes B_k + (N-1)R_k)^{-1}(y_k - y^j_k)$;
    Update $\vartheta^j_{k|k} = \vartheta^j_{k|k-1} + D^\vartheta_{k|k-1} \otimes B_k (B_k \otimes B_k + (N-1)R_k)^{-1}(y_k - y^j_k)$;

4.5 Comments and further references

We have presented only a few possible methods to approach the dual filtering problem. Of course, one could design a wide range of methods for dual filtering by combining any of the ideas from section 3. For example, one could obtain a simple joint ensemble Kalman filter by applying the EnKF to the augmented state (4.2.1). For high-dimensional and nonlinear systems, it might be necessary to combine the dual EnKF with an iterative procedure (cf. section 3.6).

The idea of solving the dual filtering problem using two separate filters seems to have originated in adaptive control [NS76]. The dual IEKF as given in algorithm 6 is a modification of the dual EKF as described in [Nel00] and [WN01], where it is discussed in the context of nonlinear time series analysis and neural networks.

As noted above, the filtering approach to parameter estimation also allows to treat parameters that might be time-varying in a random way, i.e. by letting the parameter evolve in time according to a stochastic equation,

$$d\theta(t) = G_2(t)dW_2(t).$$

Some authors (e.g. [Gel74]) recommend to use this technique in non-Gaussian scenarios even if the parameter is constant, in order to avoid underestimating parameter uncertainty.

Note, however, that the dual filters in this chapter have been designed with applications to on-line filtering in mind. If one is only interested in post-hoc parameter estimation, there are approaches which are better suited. For example, if we have a deterministic system with parametric uncertainty,

$$\begin{aligned} &\frac{d}{dt}u(t) = f(u(t), \theta), \qquad t \in [0, T], \\ &u(0) = u_0, \\ &Y_k = m_k(u(t_k)) + V_k, \qquad 0 \le t_1 < \ldots < t_K, \end{aligned}$$

one can use for example a Markov chain Monte Carlo method (see e.g. [BGJ11], [LLC10], [RR04], [Mar+12]), to approximate the moments of the posterior distribution $\mathbb{P}^{\theta|\boldsymbol{Y}_K}$ by Birkhoff averages (see e.g. [Dou+19] or [Mey09]) over a Markov chain $(\theta_n)_{n=1}^{\infty}$, i.e.

$$\mathbb{E}[f(\theta)|\boldsymbol{Y}_K = \boldsymbol{y}_K] \approx \frac{1}{n}\sum_{i=1}^{n} f(\theta_i),$$

for some $f \in L^1(\mathbb{P}^{\theta|\boldsymbol{Y}_K=\boldsymbol{y}_K})$. This type of methods can be more flexible for uncertainty quantification (since it allows to estimate properties of the posterior distribution beyond the mean and covariance) and nonlinear problems, where the involved probability distributions can be multimodal or have an otherwise complex geometrical structure.

Chapter 5

Comparison with deterministic methods

We illustrate that the iterated extended Kalman filter is equivalent to a regularized Gauss-Newton method. First, we look at a simplified setting, that was originally treated in [BC93]. Then, we extend the result to the inverse problem of parameter identification for PDEs, by showing that the reduced approach using a regularized Gauss-Newton method is equivalent to the dual iterated extended Kalman filter. Note that all results in this chapter hold for separable Hilbert spaces.

5.1 Regularized Gauss-Newton method and the extended Kalman filter

Consider the following Gaussian discrete-time filtering problem:

$$\begin{aligned} &X_k = f(X_{k-1}), \\ &X_0 \sim \mathcal{N}(x_0, P_0), \\ &Y_k = m(X_k) + V_k, \qquad V_k \sim \mathcal{N}(0, R_k), \end{aligned}$$

where the state space is a separable Hilbert space H and we have a differentiable measurement function $m \in C^1(H; \mathbb{R}^m)$.

Let us look at the first IEKF update for this problem. The time update is simply

$$x_{1|0} = f(x_0), \tag{5.1.1}$$
$$P_{1|0} = F_1 P_0 F_1^*, \tag{5.1.2}$$

where $F_1 := Df(x_0)$. The iteration for the measurement update is then

$$\begin{aligned} x_{1|1}^n &= x_{1|0} + K_1^n(y_1 - m(x_{1|1}^{n-1}) - M_1^n(x_{1|0} - x_{1|1}^{n-1})) \\ &= f(x_0) + K_1^n(y_1 - m(x_{1|1}^{n-1}) - M_1^n(f(x_0) - x_{1|1}^{n-1})) \\ &= f(x_0) + P_{1|0}(M_1^n)^*(M_1^n P_{1|0}(M_1^n)^* + R_1)^{-1}(y_1 - m(x_{1|1}^{n-1}) - M_1^n(f(x_0) - x_{1|1}^{n-1})) \\ &= f(x_0) + F_1 P_0 F_1^*(M_1^n)^*(M_1^n F_1 P_0 F_1^*(M_1^n)^* + R)^{-1}(y_1 - m(x_{1|1}^{n-1}) - M_1^n(f(x_0) - x_{1|1}^{n-1})), \end{aligned} \tag{5.1.3}$$

with $M_1^n := Dm(x_{1|1}^{n-1})$, where we used (5.1.1) in the second line, (3.2.10) in the third line, and (5.1.2) in the fourth line.

On the other hand, consider the inverse problem of estimating X_1 given $Y_1 = y_1$, where

$$y_1 = m(x_1) + V_1.$$

Our initial guess x_0 induces the initial guess $x_0^1 := f(x_0)$ for x_1. We introduce the cost functional

$$\Phi(x) = \|y_1 - m(x)\|_R^2 + \|x - f(x_0)\|_\Sigma^2. \tag{5.1.4}$$

where $R \in \mathbb{R}^{m\times m}$ is a symmetric positive definite matrix and $\Sigma \in \mathcal{L}(H)$ is a symmetric positive trace-class operator. Recall the Gauss-Newton update.

Lemma 5.1.1. *Let $F : H \to \mathbb{R}^m$ be a C^1 map between separable Hilbert spaces, let $R \in \mathbb{R}^{m\times m}$ be a symmetric positive definite matrix, and $\Sigma \in \mathcal{L}(H)$ be symmetric positive trace-class operators. Then*

$$x^{n+1} := x^n + \Sigma^{1/2}(\Sigma^{1/2}DF(x)^*R^{-1}DF(x)\Sigma^{1/2} + \mathrm{Id})^{-1}\left[\Sigma^{1/2}DF(x)^*\Gamma^{-1}(y - F(x)) + \Sigma^{-1/2}(x_0 - x^n)\right],$$

minimizes the functional

$$\Phi(x) = \|y - F(x^n) - DF(x^n)(x - x^n)\|_R^2 + \|x - x_0\|_\Sigma^2,$$

where we assume $x^n - x_0 \in \operatorname{dom}\Sigma^{-1/2}$.

Proof. First of all, we recall from the definition of the covariance-weighted norm in (1.5.1) that we have

$$\|x - x_0\|_\Sigma < \infty \iff x - x_0 \in \operatorname{dom}\Sigma^{-1/2} = \operatorname{ran}\Sigma^{1/2}.$$

Hence, it suffices to minimize over all $x = \Sigma^{1/2}z + x_0$, with $z \in H$. We insert this in the cost functional and obtain

$$\tilde{\Phi}(z) = \Phi(\Sigma^{1/2}z + x_0) = \|R^{-1/2}(y - F(x^n) - DF(x^n)(\Sigma^{1/2}z + x_0 - x^n))\|^2 + \|z\|^2.$$

Taking the Fréchet derivative, we have for $h \in H$:

$$\begin{aligned} D\tilde{\Phi}(z)h &= -(R^{-1/2}(y - F(x^n) - DF(x^n)(\Sigma^{1/2}z + x_0 - x^n)), R^{-1/2}DF(x^n)\Sigma^{1/2}h) + (z, h) \\ &= (z - \Sigma^{1/2}DF(x^n)^*R^{-1}(y - F(x^n) - DF(x^n)(\Sigma^{1/2}z + x_0 - x^n)), h). \end{aligned} \tag{5.1.5}$$

Since (5.1.5) holds for all $h \in H$, we have necessarily

$$z = \Sigma^{1/2}DF(x^n)^*R^{-1}(y - F(x^n) - DF(x^n)(\Sigma^{1/2}z + x_0 - x^n)).$$

Reordering yields

$$\begin{aligned} (\Sigma^{1/2}DF(x^n)^*R^{-1}DF(x^n)\Sigma^{1/2} + \mathrm{Id})z &= \Sigma^{1/2}DF(x^n)^*R^{-1}(y - F(x^n)) \\ &\quad - \Sigma^{1/2}DF(x^n)^*R^{-1}DF(x^n)(x_0 - x^n) \\ &= \Sigma^{1/2}DF(x^n)^*R^{-1}(y - F(x^n)) \\ &\quad - (\Sigma^{1/2}DF(x^n)^*R^{-1}DF(x^n)\Sigma^{1/2} + \mathrm{Id})\Sigma^{-1/2}(x_0 - x^n) \\ &\quad + \Sigma^{-1/2}(x_0 - x^n). \end{aligned}$$

Since the operator on the left-hand side is invertible, we obtain

$$\begin{aligned} z = &- \Sigma^{-1/2}(x_0 - x^n) + (\Sigma^{1/2}DF(x^n)^*R^{-1}DF(x^n)\Sigma^{1/2} + \mathrm{Id})^{-1} \\ &\cdot \left[\Sigma^{1/2}DF(x^n)^*R^{-1}(y - F(x^n)) + \Sigma^{-1/2}(x_0 - x^n)\right]. \end{aligned}$$

Finally, inserting back into $x = \Sigma^{1/2}z + x_0$ gives

$$x = x^n + \Sigma^{1/2}(\Sigma^{1/2}DF(x^n)^*R^{-1}DF(x^n)\Sigma^{1/2} + \mathrm{Id})^{-1}\left[\Sigma^{1/2}DF(x^n)^*R^{-1}(y - F(x^n)) + \Sigma^{-1/2}(x_0 - x^n)\right].$$

□

Hence, in our case, this means that the Gauss-Newton iteration corresponding to the cost functional (5.1.4) is given by

$$x_1^{n+1} = x_1^n + \Sigma^{1/2}(\Sigma^{1/2}(M_1^n)^* R^{-1} M_1^n \Sigma^{1/2} + \mathrm{Id})^{-1} \left[\Sigma^{1/2}(M_1^n)^* R^{-1}(y - m(x_1^n)) + \Sigma^{-1/2}(f(x_0) - x_1^n)\right], \tag{5.1.6}$$

where $M_1^n := Dm(x_1^n)$. If we choose $x^0 = f(x_0)$, we have $x_1^n - f(x_0) \in \operatorname{dom} \Sigma^{-1/2}$ by induction, so that the assumptions of lemma 5.1.1 are always satisfied.

We now use the following elementary operator identity which provides us with an alternative way to write the Kalman gain.

Lemma 5.1.2. *Let Σ be a symmetric positive operator and $R \in \mathbb{R}^{m\times m}$ a symmetric positive definite matrix. Then*

$$\Sigma M^*(M\Sigma M^* + R)^{-1} = \Sigma^{1/2}(\Sigma^{1/2} M^* R^{-1} M \Sigma^{1/2} + \mathrm{Id})^{-1} \Sigma^{1/2} M^* R^{-1} \tag{5.1.7}$$

and both expressions are well-defined.

Proof. Since R is positive definite and Σ is positive, all the involved inverses exist. Then, we simply note that

$$\begin{aligned}
&\Sigma^{1/2}(\Sigma^{1/2} M^* R^{-1} M \Sigma^{1/2} + \mathrm{Id})^{-1} \Sigma^{1/2} M^* R^{-1} (M\Sigma M^* + R)^{-1} \\
&= \Sigma^{1/2}(\Sigma^{1/2} M^* R^{-1} M \Sigma^{1/2} + \mathrm{Id})^{-1} (\Sigma^{1/2} M^* R^{-1} M \Sigma M^* + \Sigma^{1/2} M^*) \\
&= \Sigma^{1/2}(\Sigma^{1/2} M^* R^{-1} M \Sigma^{1/2} + \mathrm{Id})^{-1} (\Sigma^{1/2} M^* R^{-1} M \Sigma^{1/2} + \mathrm{Id}) \Sigma^{1/2} M^* \\
&= \Sigma M^*.
\end{aligned}$$

□

Corollary 5.1.3. *Let Σ be a symmetric positive operator and $R \in \mathbb{R}^{m\times m}$ a symmetric positive definite matrix. Then*

$$x^n + \Sigma^{1/2}(\Sigma^{1/2} M R^{-1} M \Sigma^{1/2} + \mathrm{Id})^{-1} \Sigma^{-1/2}(x_0 - x^n) = x_0 - \Sigma M^*(M\Sigma M^* + R)^{-1} M(x_0 - x^n).$$

Proof. We rewrite equation (5.1.8):

$$x_0 - x^n = \Sigma^{1/2}(\Sigma^{1/2} M R^{-1} M \Sigma^{1/2} + \mathrm{Id})^{-1} \Sigma^{-1/2}(x_0 - x^n) + \Sigma M^*(M\Sigma M^* + R)^{-1} M(x_0 - x^n).$$

This holds true if we have

$$\mathrm{Id} = \Sigma^{1/2}(\Sigma^{1/2} M R^{-1} M \Sigma^{1/2} + \mathrm{Id})^{-1} \Sigma^{-1/2} + \Sigma M^*(M\Sigma M^* + R)^{-1} M.$$

Applying lemma 5.1.2, the right-hand side becomes

$$\begin{aligned}
&\Sigma^{1/2}(\Sigma^{1/2} M R^{-1} M \Sigma^{1/2} + \mathrm{Id})^{-1} \Sigma^{-1/2} + \Sigma^{1/2}(\Sigma^{1/2} M R^{-1} M \Sigma^{1/2} + \mathrm{Id})^{-1} \Sigma^{1/2} M^* R^{-1} M \\
&= \Sigma^{1/2}(\Sigma^{1/2} M R^{-1} M \Sigma^{1/2} + \mathrm{Id})^{-1} (\Sigma^{-1/2} + \Sigma^{1/2} M^* R^{-1} M) \\
&= \Sigma^{1/2}(\Sigma^{1/2} M R^{-1} M \Sigma^{1/2} + \mathrm{Id})^{-1} (\mathrm{Id} + \Sigma^{1/2} M^* R^{-1} M \Sigma^{1/2}) \Sigma^{-1/2} \\
&= \mathrm{Id},
\end{aligned}$$

as desired. □

First using lemma 5.1.2, the Gauss-Newton update (5.1.6) can be rewritten, as

$$x_1^{n+1} = x_1^n + \Sigma(M_1^n)^*(M_1^n \Sigma (M_1^n)^* + R)^{-1} (y - m(x_1^n)) \tag{5.1.8}$$

$$+ \Sigma(M_1^n)^*(M_1^n \Sigma (M_1^n)^* + R)^{-1} \Sigma^{-1/2} (f(x_0) - x_1^n). \tag{5.1.9}$$

Then, using corollary 5.1.3, this becomes

$$x_1^{n+1} = x_0 + \Sigma(M_1^n)^*(M_1^n \Sigma (M_1^n)^* + R)^{-1} [y - m(x_1^n) - M_1^n(x_0 - x_1^n)].$$

Comparing the expressions (**??**) and (5.1.3), we see that they coincide if we put $\Sigma = F_1 P_0 F_1^*$ and $R = R_1$. Hence, by induction over the measurement number k, the IEKF is equivalent to the regularized Gauss-Newton method.

5.2 Reduced approach and the dual iterated extended Kalman filter

Now, we consider the more advanced problem of parameter identification in PDEs. We are given

$$D_t u(t) = f(u(t), \theta), \tag{5.2.1a}$$
$$u(0) = u_0, \tag{5.2.1b}$$
$$Y_k = m_k(u(t_k)) + V_k, \qquad k \in \mathbb{N}, \tag{5.2.1c}$$

where $u \in \mathcal{U}$ is an element of some space of functions that take values in a separable Hilbert space U, and has derivatives in a space $\mathcal{W}$ of functions that take values in another separable Hilbert space W. Furthermore, assume $u_0 \in U$, that $\theta \in \Theta$ is an unknown parameter that takes values in a separable Hilbert space Θ, that $f \in M(\mathcal{U} \times \Theta; W)$ is a deterministic function, and $m_k \in C^1(U; \mathbb{R}^m)$ is a Fréchet differentiable measurement function. We assume measurement noise given by independent Gaussian vectors $V_k \sim \mathcal{N}(0, R_k)$. Furthermore, suppose we have the initial guess $\theta \sim \mathcal{N}(\theta, \Gamma)$.

In the following we consider incremental approaches. Instead of solving the parameter identification problem for the whole sequence of measurements $(y_1, y_2, ...)$ at once, we consider sequentially solving multiple inverse problems, one for each new measurement. The estimate of θ from the previous measurements is then used as the initial guess for the estimator of θ based on the new measurement. The advantage of such an approach is that it lends itself to on-line estimation and to problems where the number of measurements is so large that it might be impractical to incorporate all datapoints at once in the parameter update. The reduced approach [KNS19] formulates then an associated inverse problem by considering the solution map

$$S : \Theta \to \mathcal{U},$$
$$S(t; \vartheta) = u(t) :\Longleftrightarrow \begin{cases} D_t u = f(u, \vartheta), \\ u(0) = u_0. \end{cases}$$

which maps a solution $u \in \mathcal{U}$ to each parameter value ϑ. Then, the forward problem (for the k-th measurement) is given by

$$Y_k = F_k(\theta) + V_k,$$

where $F_k(\theta) := m_k(S(t_k; \theta))$. For each measurement y_k, we define a regularized cost functional

$$\Phi(\vartheta) := \|y_k - m_k(S(t_k; \vartheta))\|^2_{R_k} + \|\vartheta - \vartheta_{k-1}\|^2_{\Gamma_k},$$

with $\Gamma_k \in \mathcal{L}_1(\Theta)$ being symmetric positive definite trace-class operators, and $\vartheta_{k-1} \in \Theta$ the best guess for θ after the first $k-1$ measurements. The regularized Gauss-Newton iteration (see lemma 5.1.1) for this problem is then

$$\vartheta_k^{n+1} = \vartheta_k^n + \Gamma_k^{1/2}(DF(\vartheta_k^n)^* R_k^{-1} DF(\vartheta_k^n)\Gamma_k^{1/2} + \mathrm{Id})^{-1}(\Gamma_k^{1/2} DF(\vartheta_k^n)^* R_k^{-1}(y_k - m_k(S(t_k; \vartheta_k^n))) \tag{5.2.2}$$
$$+ \Gamma_k^{-1/2}(\vartheta_{k-1} - \vartheta_k^n)), \tag{5.2.3}$$
$$\vartheta_k^0 = \vartheta_{k-1}.$$

For calculating the required Jacobians $DF(\vartheta) = Dm_k(S(t_k; \vartheta) D_\vartheta S(t_k; \vartheta)$, we need to calculate the parameter derivative of the solution operator. We have

$$S(t_k; \vartheta) = u_0 + \int_0^{t_k} f(S(t; \vartheta), \vartheta) dt.$$

If we do not want to solve the forward problem for the whole time horizon $[0, t_k]$ in every step, we can make the approximation

$$S(t_k; \vartheta) \approx u_{k-1|k-1} + \int_{t_{k-1}}^{t_k} \left[\partial_1 f(S(t; \vartheta, \vartheta) D_\vartheta S(t; \vartheta) + \partial_2 f(S(t; \vartheta), \vartheta)\right] dt,$$

where $u_{k-1|k-1}$ is our estimate of U given $\boldsymbol{Y}_{k-1} = \boldsymbol{y}_{k-1}$. We obtain after differentiating

$$D_\theta S(t_k;\theta) = \int_0^{t_k} \left[\partial_1 f(S(t;\theta),\theta) D_\theta S(t;\theta) + \partial_2 f(S(t;\theta),\theta)\right] dt.$$

Comparing this with (4.3.7) from section 4.3, we note that $DF(\vartheta_k^n)$ corresponds to N_k^n. Hence we can rewrite the update (5.2.3) as

$$\vartheta_k^{n+1} = \vartheta_k^n + \Gamma_k^{1/2}(\Gamma_k^{1/2}(N_k^n)^*\Gamma_k^{-1}N_k^n\Gamma_k^{1/2} + \mathrm{Id})^{-1}\left(\Gamma_k^{1/2}(N_k^n)^*\Gamma_k^{-1}(y_k - m_k(S(t_k;\vartheta_k^n)) + \Gamma_k^{-1/2}(\vartheta_{k-1} - \vartheta_k^n)\right). \tag{5.2.4}$$

Using lemma 5.1.2 and corollary 5.1.3, this is equivalent to

$$\vartheta_k^{+1} = \vartheta_{k-1} + \Gamma_k(N_k^n)^*((N_k^n)^*\Sigma_k N_k^n + \Gamma_k)^{-1}\left(y_k - m_k(S(t_k;\theta_k^{n-1})) - N_k^n(\vartheta_{k-1} - \vartheta_k^n)\right). \tag{5.2.5}$$

Recall now the measurement update (4.3.4) for the parameter from the dual IEKF:.

$$\vartheta_{k|k}^{n+1} = \vartheta_{k|k-1} + G_k^n\left(y_k - m_k(u_{k|k}^n, \vartheta_{k|k}^n) - N_k^n(\vartheta_{k|k-1} - \vartheta_{k|k}^n)\right), \tag{5.2.6}$$

where in our case

$$\begin{aligned} G_k^n &= \Gamma_{k|k-1}(N_k^n)^*(N_k^n\Gamma_{k|k-1}(N_k^n)^* + R)^{-1}, \\ m(u_{k|k}^n, \vartheta_{k|k}^n, t_k) &= m_k(S(t_k;\vartheta_{k|k}^n)) \\ u_{k|k}^n &= S(t_k;\vartheta_{k|k}^n), \end{aligned}$$

so that the update (5.2.6) becomes

$$\vartheta_{k|k}^{n+1} = \vartheta_{k|k-1} + \Gamma_{k|k-1}(N_k^n)^*(N_k^n\Gamma_{k|k-1}(N_k^n)^* + R_k)^{-1}\left(y_k - m_k(S(t_k;\theta_k^n)) - N_k^n(\vartheta_{k|k-1} - \vartheta_k^n)\right), \tag{5.2.7}$$

which coincides with (5.2.5) if we identify $\Gamma_k = \Gamma_{k|k-1}$. *This shows that the reduced approach to parameter identification using a Gauss-Newton method is equivalent to the dual iterated extended Kalman filter.*

5.3 Comments and further references

The equivalence of the IEKF and the Gauss-Newton method was originally shown for the finite-dimensional discrete-time setting in [BC93], but already seems to have been noticed in [Jaz70, chapter 9.7]. The equivalence of the reduced approach and the dual IEKF is a straightforward extension of this argument.

In general, variational regularization techniques for filtering or inverse problems can often be given a probabilistic interpretation. Any cost functional can be exponentiated and interpreted as an unnormalized probability density. Consider for example a finite-dimensional, nonlinear inverse problem

$$y = F(u) + v,$$

with $F : \mathbb{R}^d \to \mathbb{R}^m$. We can give this problem a probabilistic treatment by modelling the noise as Gaussian random vector $V \sim \mathcal{N}(0, R)$ and assigning a prior probability $U \sim \mathcal{N}(\bar{u}, \Sigma)$. Assuming that R and Σ are positive definite, all densities exist and we have by proposition 1.3.12 (Bayes)

$$\begin{aligned} p^{U|Y}(u|y) &= p^U(u)\frac{p^{Y|U}(y|u)}{p^Y(y)} \\ &\propto e^{-\frac{1}{2}(\|y-F(u)\|_R^2 + \|u-\bar{u}\|_\Sigma^2)}. \end{aligned}$$

Since maximization of a function is equivalent to minimizing its negative logarithm, we have that

$$\begin{aligned}\hat{u}(y) &:= \operatorname*{argmax}_u p^{U|Y}(u|y) \\ &= \operatorname*{argmin}_u \|y - F(u)\|_R^2 + \|u - \bar{u}\|_\Sigma^2.\end{aligned}$$

$\hat{u}(y)$ is called the **maximum a posteriori** (MAP) estimate of U given $Y = y$ (see [Gel+13, chapter 13]). Thus, any instance of Tikhonov regularization can be given a probabilistic interpretation as finding the maximum of the conditional probability density (called the posterior density), given $Y = y$. In linear Gaussian problems, the MAP estimator coincides with the posterior mean $\mathbb{E}[U|Y = y]$. Since the EKF and the IEKF approximate the posterior mean of the state, this provides an intuitive explanation for the equivalency of Kalman-type methods and variational regularization techniques.

However, in nonlinear or non-Gaussian problems, the MAP estimator and the posterior mean do not coincide, because in that case, the posterior might be multimodal or not symmetric at all. Therefore, the MAP estimator has *no probabilistic meaning* in general [Gel+13, chapter 13]. Nevertheless, its use — whether explicitly or implicitly — is widespread and has seen a lot of success even for mildly nonlinear problems (see e.g. [Fat+12]).

Chapter 6

Applications

In the following, we discuss two applications. First, we demonstrate how the dual EKF and the joint EKF look on a concrete potential identification problem. After that, we will consider the inverse problem of estimating the initial condition of a wave equation from consecutive measurements. For this problem, we will first look at a direct method based on a forward-backward approach. Then we contrast this to simply applying the dual EnKF from section 4.4.

6.1 Potential identification

As a first application, we look at a potential identification problem. Classically, one is posed with determining an unknown potential $\theta : D \to \mathbb{R}$ from measuring the solution of the diffusive Malthus equation on a domain $D \subset \mathbb{R}^n$:

$$\begin{aligned} &\partial_t u(t,x) - \Delta u(t,x) + \theta(x)u(t,x) = \varphi(t,x), && t \in (0,T), x \in D, \\ &u(x,t) = 0, && t \in (0,T), x \in \partial D, \\ &u(x,0) = u_0(x), && x \in D, \end{aligned} \tag{6.1.1}$$

We will consider the noisy version, i.e. where we assume that there is also intrinsic noise in the evolution equation (6.1.1). Formally, we consider the stochastic partial differential equation:

$$\partial_t u = \Delta u - \theta u + \varphi + \partial_t W,$$

where $W = W(x,t)$ is some random noise process. We will assume in the following that W is a Q-Wiener process with values in a Hilbert space H that we will specify later.

Let us formulate the problem as a linear SPDE

$$\begin{aligned} &dU(t) = (\Delta + \theta)U(t)dt + dW(t), \\ &U_0 = 0, \end{aligned}$$

where $U_0 \in H$ and we identify θ with the multiplication operator $u \mapsto \theta u$. Before considering potential identification, we have to adress well-posedness of the forward problem. This follows from the basic existence theorem 1.7.11. The conditions of that theorem can easily be verified in our case, such that existence and uniqueness follow.

Proposition 6.1.1. *For all $\theta \in C_b(D)$, equation (6.1.2) admits a unique weak solution in $H_0^1(D)$.*

Proof. We apply theorem 1.7.11 with $H = L^2(D)$. The linear operator $Au := \Delta u - \theta u$ defines for all $\theta \in C_b(D)$ an analytic (and in particular strongly continuous) semigroup S (see [Lun95, chapter 3.1]). Moreover, we have by proposition 1.4.9:

$$|\int_0^T \operatorname{tr}[S(t)GQG^*S(t)^*]dt| \leq \|G\|_2^2 \|Q\|_1 \int_0^T \|S(t)\|^2 dt.$$

By assumption and because $S(t)$ is contractive for all $t \in [0,T]$ [Eva10, section 7.4], the right-hand side is finite, so that all conditions for theorem 1.7.11 are satisfied. □

Since $C_b(D)$ is a separable Banach space, the methods from chapter 4 can be applied. Finally, we can formulate the potential identification problem.

Problem 6.1.2 (Potential identification)**.** *Given a bounded domain $D \subset \mathbb{R}^n$ with C^2 boundaries, let $\varphi \in L^1([0,T]; L^2(D))$ and $u_0 \in H_0^1(D)$. Given the noisy Malthus equation*

$$dU(t) = (\Delta U(t) - \theta U(t) + \varphi(t))dt + dW(t), \qquad t \in (0,T), \tag{6.1.2a}$$

$$U\Big|_{\partial D} = 0, \tag{6.1.2b}$$

$$U(0) = u_0, \tag{6.1.2c}$$

where W is a $L^2(D)$-valued Q-Wiener process (whose existence is guaranteed by proposition 1.7.2). We assume discrete measurements at times $0 < t_1 < ... < t_K \leq T$ given by

$$Y_k = m(U(t_k)) + V_k, \quad k \in [K],$$

where $m \in M(L^2(D); \mathbb{R}^m)$ and $V_k \sim \mathcal{N}(0, R_k)$. The problem is then to estimate the potential θ in $C_b(D)$, $s > n/2$, with prior mean $\mathbb{E}\theta = \bar{\vartheta}$ and covariance $\mathrm{Cov}(\theta) = \Gamma$, given measurements $Y_1 = y_1, ..., Y_K = y_K$.

Now, we can apply any of the methods from chapter 4 to this problem. In order to get a practical algorithm, we have two choices:

1. We can apply the Hilbert space version of the joint or dual EKF to the infinite-dimensional filtering problem 6.1.2, and then discretize by projection to finite-dimensional subspaces in order to obtain a practical algorithm.

2. We can discretize first, and then apply the finite-dimensional versions of the joint or dual EKF to the discretized problem.

We will follow the latter approach and discretize first. We choose a simple Galerkin approximation (see [DL00] or also [Tho06] for the deterministic, and [Yan04] for the stochastic case):

Consider the weak formulation of equation (6.1.2).

$$\begin{aligned}&(dU(t), \phi) = (-\nabla U(t), \nabla\phi) - (\theta U(t), \phi) + (\varphi, \phi))\, dt + dW(t), \qquad \phi \in H_0^1(D),\\ &U_0 = u_0.\end{aligned}$$

Let (S_h) be a sequence of finite-dimensional subspaces $S_h \subset H_0^1(D)$. Suppose S_h has a basis $(\phi_i)_{i=1}^{d_h}$ of piecewise polynomials of degree 1 or less, and let $\Pi_h : L^2(D) \to S_h$ be the L^2-projection on S_h, with basis expansion

$$\Pi_h u = \sum_{i=1}^{d_h} u_h^i \phi^i, \qquad u \in L^2(D).$$

Furthermore, let (V_h) be a sequence of finite-dimensional subspaces $P_h \subset C_b(D)$, with corresponding basis $(p_i)_{i=1}^{r_h}$. We then make the finite-dimensional approximations

$$U(t) \approx U_h(t) := \Pi_h U(t) = \sum_{i=1}^{d_h} U_h^i(t)\phi_i,$$

$$\theta \approx \theta_h := \Pi_h \theta = \sum_{i=1}^{r_h} \theta_h^i p_i.$$

Inserting the expressions for U_h and θ_h in the weak formulation and putting $\phi = \phi_k$ leads to the discretized problem

$$\sum_i dU_h^i(t)(\phi_i, \phi_k) = \left(-\sum_i U_h^i(t)(\nabla\phi_i, \nabla\phi_k) - \sum_i \sum_j \theta_h^j U_h^i(t)(p_i\phi_i, \phi_k) + (\varphi, \phi_k) \right) dt + (dW^i(t), \phi_k),$$

$\Pi_h U(0) = \Pi_h u_0$.

for $k = 1, \ldots, d_h$. Together with the initial condition, this can be rewritten in matrix form:

$$\begin{aligned} B_h dU_h(t) &= A_h(\theta_h)U_h(t)dt + \varphi_h(t)dt + dW_h(t), \\ U_h(0) &= u_{0,h}, \end{aligned} \tag{6.1.3}$$

where

$$\begin{aligned} A_h(\theta_h)_j^i &:= -(\phi_i, \phi_j - \sum_k \theta_h^k (p_k\phi_i, \phi_j), \\ (B_h)_j^i &= (\phi_i, \phi_j,), \\ \varphi_h^i(t) &:= (\varphi(t), \phi_i), \\ W_h(t)^i &= (W(t), \phi_i), \\ u_0 \approx u_{0,h}(t) &:= \Pi_h u_0 = \sum_{i=1}^{d_h} u_{0,h}^i \phi_i. \end{aligned}$$

Note that the terms $(p_k\phi_i, \phi_j)$ are finite because we assumed $V_h \subset C_b(D)$. It also follows from the definition of Gaussian random elements (definition 1.6.2) that W_h is a $\mathbb{R}^{d_h}$-valued Wiener process. As a last step, we note that B is a Gram matrix and therefore invertible. Hence, we can rewrite (6.1.3) as

$$dU_h(t) = \tilde{A}_h(\theta_h)U_h(t) + \tilde{\varphi}_h(t)dt + d\tilde{W}_h(t),$$

where $\tilde{A} := B^{-1}\tilde{A}$, $\tilde{\varphi} := B^{-1}\varphi_h$ and $\tilde{W}_h(t) := B^{-1}W_h(t)$. Furthermore, let $\tilde{Q}_h$ be the covariance of $\tilde{W}_h$, and let $\bar{\vartheta}_h = \Pi_h\bar{\vartheta}$ and $T_h = \Pi_h T\Pi_h^*$.

Finally, we have brought the problem into the form of a finite-dimensional filtering problem. Furthermore, it has been shown in [Yan04] that the Galerkin approximation converges. For the homogeneous case $\varphi = 0$, the following estimate is available:

Theorem 6.1.3. *Let $A := \Delta - \theta\mathrm{Id}$ and $\dot{H}^1 := \operatorname{dom} A^{1/2}$, with inner product*

$$\|v\|_{\dot{H}^1} := \|A^{1/2}v\|_2.$$

Then, we have

$$\|U_h(t) - U(t)\|_{L^2(\mathbb{P};L^2(D))} \leq Ch\left(\|u_0\|_{\dot{H}(D)} + \operatorname{tr}(Q)^{1/2}\right).$$

Proof. This is a consequence of [Yan04, theorem 1.1] (c.f. [Bar10]). □

We can then simply apply the dual EnKF (algorithm 7) from above. This leads to algorithm 8. It produces with each incoming measurement $Y_k = y_k$ an updated parameter estimate $\bar{\vartheta}_{k|k}$ with a corresponding covariance estimate given by $\frac{1}{N-1}D_k^\vartheta D_k^{\vartheta *}$.

6.2 Stochastic wave equation

Consider the wave equation on a domain $D \subset \mathbb{R}^3$.

$$\begin{aligned} \partial_t u(x,t) &= \Delta u(x,t), && x \in D, t \geq 0, \\ u(x,t) &= 0, && x \in \partial D, t \geq 0, \\ u(0,x) &= \theta(x), \quad \partial_t u(x,0) = 0, && x \in D, \end{aligned}$$

Algorithm 8 `Dual EnKF for real-time potential identification`

Given u_0, φ, measurements $Y_1 = y_1, Y_2 = y_2, ..., K$, and inital guess $\theta \sim \mathcal{N}(\hat{\vartheta}, \Gamma)$

1: Compute discretizations $\tilde{\varphi}_h$, $u_{0,h}, \bar{\vartheta}_h, \Gamma_h$ and $\tilde{Q}_h$;
2: Initialize $\vartheta_0^1, \dots, \vartheta_0^N \sim \mathcal{N}(\bar{\vartheta}_h, T_h)$;
3: **for** $k = 1, 2, \dots$ **do**

`Time update`

4: For all $j \in [N]$, solve

$$du^j(t) = (\tilde{A}_h(\vartheta_{k-1}^j)u^j(t) + \tilde{\varphi}_h(t))dt + dw^j(t), t \in (t_{k-1}, t_k),$$
$$u^j(t_{k-1}) = u_{k-1|k-1}^j,$$

with w^j being independent $\tilde{Q}_h$-Wiener processes, and set $u_{k|k-1}^j = u(t_k|t_{k-1})$;

`Measurement update`

5: Sample $v_k^1, \dots, v_k^N \sim \mathcal{N}(0, R_k)$;
6: Set $y_k^j = m(u_{k|k-1}^j) + v_k^j$ for $j \in [N]$;
7: Compute $\bar{\vartheta}_{k-1} = \frac{1}{N}\sum_j \vartheta_{k-1}^j$;
8: Set $D_{k-1}^\theta = [\vartheta_{k-1}^1 - \bar{\vartheta}_{k-1}, \dots, \vartheta_{k-1}^N - \bar{\vartheta}_{k-1}]$;
9: Compute $\bar{y}_k = \frac{1}{N}\sum_j y_k^j$;
10: Compute $\bar{u}_{k|k-1} = \frac{1}{N}\sum_j u_{k|k-1}^j$;
11: Assemble $D_{k|k-1}^u = [u_{k|k-1}^1 - \bar{u}_{k|k-1}, \dots, u_{k|k-1}^N - \bar{u}_{k|k-1}]$;
12: Set $B_k = [y_k^1 - \bar{y}_k, \dots, y_k^N - \bar{y}_k]$;
13: Set $\vartheta_k^j = \vartheta_{k-1}^j + D_{k-1}^\theta \otimes B_k(B_k \otimes B_k + (N-1)R_k)^{-1}(y_k - y_k^j)$ for all $j \in [N]$;
14: Set $u_{k|k}^j = u_{k|k-1}^j + D_{k|k-1}^u \otimes B_k(B_k \otimes B_k + (N-1)R_k)^{-1}(y_k - y_k^j)$ for all $j \in [N]$;

with the initial pressure $\theta \in H_0^1(D)$ being the unknown quantity to be estimated (cf. example 1.3.3). We will assume that our initial information about θ is well represented by a Gaussian distribution on $H_0^1(D)$, $\theta \sim \mathcal{N}(\bar{\theta}, T_0)$ (see section 1.6). For brevity, we will write $u(t) := u(\cdot, t)$, and correspondingly $\dot{u}(t) := \partial_t u(\cdot, t)$, $\Delta u(t) := \Delta u(\cdot, t)$.

In the following, we want to investigate the noisy counterpart of the wave equation, which might be written informally as

$$\partial_t u(x,t) - \Delta u(x,t) = \partial_t N(x,t),$$

for some appropriate function-valued noise process N. In order to write this as a proper stochastic equation, we transform this in a first-order system by defining:

$$z(t) := \begin{bmatrix} u(t) \\ D_t u(t) \end{bmatrix}.$$

Furthermore, we will assume that the noise process is of the form $N(t) = \int_0^t G(s)dW(s)$, for an appropriate Wiener process W.

For the observation process, we assume boundary measurements at points $x_1, ..., x_m \in \partial D$ and times $t_1, t_2, ...$ of the form

$$Y_k^i = \int_{\mathbb{R}^3} U(x, t_k)h(x_i - x)dx + V_k^i,$$

where $V_k \sim \mathcal{N}(0, R_k)$ are independent m-dimensional Gaussian random vectors, and the measurement kernel h is a Gaussian function with variance $\sigma^2 > 0$:

$$h(x) = (2\pi\sigma^2)^{-3/2}e^{|x|^2/2\sigma^2}.$$

We will write this in vector form:

$$Y_k = MU(t_k) + V_k, \tag{6.2.1}$$

where Y_k is now a m-dimensional random vector, and $M : H_0^1(D) \times L^2(D) \to \mathbb{R}^m$, $M(u,v)^i := u * h(x_i)$, for $i = 1, ..., m$.

The rigorous formulation of this problem is then the following.

Problem 6.2.1 (Estimation of the initial condition)**.** *Let $D \in \mathbb{R}^3$ be a domain with smooth boundary, W be a $L^2(D)$-valued Q-Wiener process and let $G : [0,T] \to \mathcal{L}_2(L^2(D))$. Let θ be a random element of $H_0^1(D)$ with Gaussian distribution $\theta \sim \mathcal{N}(\bar{\vartheta}, T)$. Given the stochastic wave equation,*

$$dZ(t) = A(t)Z(t)dt + d\tilde{W}(t), \tag{6.2.2a}$$
$$Z|_{\partial\Omega} = 0, \tag{6.2.2b}$$
$$Z(0) = f, \tag{6.2.2c}$$

where

$$A(t) := \begin{bmatrix} 0 & \mathrm{Id} \\ \Delta & 0 \end{bmatrix},$$
$$\tilde{W}(t) := \begin{bmatrix} 0 \\ \int_0^t G(s)dW(s) \end{bmatrix},$$
$$f := \begin{bmatrix} \theta \\ 0 \end{bmatrix} \in H_0^1(D) \times L^2(D),$$

and measurements

$$Y_k = MZ(t_k) + V_k, \qquad k \in [K],$$

with $M : H_0^1(D) \times L^2(D) \to \mathbb{R}^m$ defined under equation (6.2.1) above and $V_k \sim \mathcal{N}(0, R_k)$, estimate the conditional expectation and covariance of θ,

$$\theta_k := \mathbb{E}[\theta | \boldsymbol{Y}_k = \boldsymbol{y}_k],$$
$$T_k := \mathrm{Cov}(\theta | \boldsymbol{Y}_k = \boldsymbol{y}_k),$$

for $k \in [K]$.

First, we observe well-posedness of the forward problem.

Proposition 6.2.2 (Well-posedness of the stochastic wave equation)**.** *Assume $G \in L^2([0,T]; \mathcal{L}_2(L^2(D)))$ and W is a $L^2(D)$-valued Q-Wiener process. Let $\theta = \vartheta \in H_0^1(D)$. Then, there exists a unique weak solution $Z = \begin{bmatrix} U & D_t U \end{bmatrix}^\top$ of the stochastic wave equation (6.2.2) with $U \in L^2(\mathbb{P}; C([0,T]; H^1(D)))$ and $D_t U \in L^2(\mathbb{P}; C([0,T]; L^2(D)))$ which satisfies:*

$$\mathbb{E} \sup_{0 \le t \le T} \|U(t)\|_{H_0^1}^2 \le C(T) \left(\|\vartheta\|_{H_0^1}^2 + 2 \int_0^T \|G(t)Q\|_2^2 dt \right).$$

Proof. Follows from [Cho07, chapter 5, theorem 3.5]. □

For the further analysis, we bring the system in weak form.

$$d(Z(t), w) = a(Z(t), w)dt + (G(t)dW(t), w), \qquad \forall w \in H_0^1(D) \times L^2(D),$$

where

$$a(v, w) = \begin{bmatrix} (v^2, w^2) \\ -(\nabla v^1 \nabla w^1) \end{bmatrix}, \qquad v, w \in H_0^1(D) \times L^2(D).$$

From here, we proceed analogously to section 6.1 by choosing a finite element approximation (S_h) of suitable (e.g. piecewise constant) polynomials, with a basis $(\phi_i)_{i=1}^{d_h}$ and such that $S_h \overset{h\to 0}{\longrightarrow} H_0^1(D)$. Eventually, we obtain the space-discretized problem

$$\begin{aligned} dZ_h(t) &= \tilde{A}_h Z_h(t)dt + d\tilde{W}_h(t), \\ Z_h(0) &= f_h, \end{aligned}$$

where Z_h is a $2d_h$-dimensional random process and

$$\begin{aligned} f_h &:= \begin{bmatrix} \theta_h \\ 0 \end{bmatrix}, \\ \theta_h &:= \Pi_h \theta, \\ \tilde{A}_h &:= B^{-1} A_h, \\ B_j^i &:= (\phi_i, \phi_j), \qquad i,j \in [d_h], \\ A_h &:= \begin{bmatrix} 0 & \mathbb{I} \\ A_2 & 0 \end{bmatrix} \in \mathbb{R}^{2d_h \times 2d_h}, \\ (A_2)_j^i &:= -(\nabla\phi_i, \nabla\phi_j), \qquad i,j \in [d_h], \\ \tilde{W}_h &:= B^{-1} W_h, \\ W_h(t)^i &:= (\int_0^t G(s)dW(s), \phi_i), \qquad i \in [d_h]. \end{aligned}$$

Similarly, for the measurement equation we have

$$Y_k^i = \sum_{j=1}^{d_h} Z^j(t_k)\phi_j * h(x_i) + V_l^i, \qquad j \in [d_h], i \in [m],$$

which can again be written in vector form,

$$Y_k = M_h Z_h(t_k) + V_k,$$

with $(M_h)_j^i = \phi_j * h(x_i)$ for $j \in [d_h], i \in [m]$, and $(M_h)_j^{d_h+k} = 0$ for $k \in [d_h], i \in [m]$. Finally, we discretize the initial guess

$$\bar{\vartheta}_h := \Pi_h \bar{\vartheta},$$

and the covariance operator

$$T_h := \Pi_h T \Pi_h^*,$$

so that $\Pi_h \in \mathbb{R}^{d_h \times d_h}$.

Eventually, we can formulate the corresponding finite-dimensional stochastic system

$$\begin{aligned} dZ_h(t) &= \tilde{A}_h Z_h(t)dt + d\tilde{W}_h(t), \qquad tin(0,T), \\ Z_h(0) &= f_h, \\ Y_k &= M_h Z_h(t_k) + V_k, \end{aligned}$$

with $f_h \sim \mathcal{N}(\bar{f}_h, \tilde{T}_h)$, where $\bar{f} = [\bar{\vartheta}_h, 0]^\top$ and

$$\tilde{T}_h = \begin{bmatrix} T_h & 0 \\ 0 & 0 \end{bmatrix}.$$

Similar to section 6.1, it has been shown in [KLS10] that the Galerkin approximation converges.

Moreover, note that the relation of f_h and $Z_h(t)$ is linear. In fact, it is given by the formula (see [PZ14]):

$$Z_h(t) = S(t)f + \int S(t)d\tilde{W}_h(t),$$

where S is the semigroup generated by A. Taking expectations, we have

$$z_{k|k} := \mathbb{E}[Z(t_k)|\boldsymbol{Y}_k = \boldsymbol{y}_k] = S(t)\theta_k \tag{6.2.3}$$

For a post-hoc estimate of θ we can simply compute $z_{K,K}$ using the Kalman-Bucy method, and then solve the deterministic system

$$D_t z(t) = A_h(t)z(t), \tag{6.2.4a}$$
$$z(t_K) = z_{K|K}, \tag{6.2.4b}$$

backwards. Then, $\theta_K = z(0)$. Similarly, the posterior covariance T_K can be calculated from $\Sigma_{K|K} = \mathrm{Cov}(Z(t_K)|\boldsymbol{Y}_k = \boldsymbol{y}_k)$ by solving backwards the system

$$D_t\Sigma(t) = A_h(t)\Sigma(t) + \Sigma(t)A_h(t)^\top + Q_l(t), \tag{6.2.5a}$$
$$\Sigma(t_K) = \Sigma_{K|K}, \tag{6.2.5b}$$

and then setting $T_K = \Sigma(0)$.

For a practical approach that can cope with large d_h, we use the ensemble Kalman smoother. The resulting algorithm is described in algorithm 9. It provides an on-line algorithm for updating the mean and covariance estimates for the initial value $z(0)$ given consecutive, real-time measurements $Y_1 = y_1, \ldots, Y_K = y_k$. Note that θ_h can then be estimated by projecting on the first d_h components.

Algorithm 9 `Ensemble Kalman smoother for the wave equation`

Given measurements $Y_1 = y_1, Y_2 = y_2, \ldots$ and ensemble size N.

Compute discretizations $\bar{\vartheta}_h$, T_h;
Sample $z^1_{0|0}, \ldots, z^N_{0|0} \sim \mathcal{N}(\bar{f}_h, \tilde{f}_h)$;
for $k = 1, 2, \ldots$ **do**
`EnKF step`
For all $j \in [N]$, solve

$$dz^j(t) = A_h z^j(t)dt + dw^j(t), \qquad t \in (t_{k-1}, t_k),$$
$$z^j(t_{k-1}) = z^j_{k-1|k-1},$$

where w^j are independent samples of $\tilde{W}_h$, and set $z^j_{k|k-1} = z^j(t_k)$.
Compute $\bar{z}_{k|k-1} = \frac{1}{N}\sum_j z^j_{k|k-1}$;
Set $D^z_{k|k-1} = [z^1_{k|k-1} - \bar{z}_{k|k-1}, \ldots, z^N_{k|k-1} - \bar{z}_{k|k-1}]$;
Sample $v^1_k, \ldots, v^j_k \sim \mathcal{N}(0, R_k)$;
Set $y^j_k = M_h z_{k|k-1} + v^j_k$ for $j \in [N]$;
Compute $\bar{y}_k = \frac{1}{N}\sum_j y^j_k$;
Set $B_k = [y^1_k - \bar{y}_k, \ldots, y^N_k - \bar{y}_k]$;
Update $z^j_{k|k} = z^j_{k|k-1} + D^z_{k|k-1} \otimes B_k(B_k \otimes B_k + (N-1)R_k)^{-1}(y_k - y^j_k)$;
`Smoothing step`
Compute $\bar{z}_{0|k-1} = \frac{1}{N-1}\sum_j z^j_{0|k-1}$;
Set $D^z_{0|k-1} = [z^1_{0|k-1} - \bar{z}_{0|k-1}, \ldots, z^N_{0|k-1} - \bar{z}_{0|k-1}]$;
Update $z^j_{0|k} = z^j_{0|k-1} + D^z_{0|k-1} \otimes B_k(B_k \otimes B_k + (N-1)R_k)^{-1}(y_k - y^j_k)$;

Chapter 7

Conclusion

In this thesis, we have demonstrated how a selection of methods from stochastic filtering can be generalized in Hilbert spaces. This allows them to be applied to very general models in a wide range of applications.

The theoretical foundation for our infinite-dimensional filtering results in chapter 2 were given by the observation that all the important properties of Gaussian distributions on which classical, finite-dimensional linear Gaussian filtering is based, generalize to infinite-dimensional Hilbert spaces. This gave rise to a simple derivation of the infinite-dimensional Kalman filter as the solution to the linear discrete filtering problem.

Similarly, using the definition of the stochastic integral in Hilbert spaces, a simple derivation of the Kalman-Bucy filter for continuous-discrete filtering was presented. The Kalman-Bucy filter was then used as a basic inspiration for all the other filtering methods we presented in this thesis.

In chapter 3, we presented a few approximate filters for nonlinear filtering problems. They can be organized into two main ideas: linearization and ensemble approximation. The extended Kalman filter and its iterated version fall in the former category. They are based on sequential linearization of the state and measurement dynamics. The resulting methods have the advantage that they are conceptually simple, derive their intuition from the Kalman filter, and are in practice compatible with any deterministic solver of the state equation. Next, we presented the ensemble approach to filtering, which represents the posterior distribution of the state using a finite ensemble. Historically, the main motivation for this approach was that the Kalman filter is impractical for high-dimensional systems. But as we showed in sections 3.3 to 3.5, ensemble-based methods also have the advantage that they lead to derivative-free algorithms, and that they are very flexible and can be adapted to a wide range of problems. In particular, we discussed the method of ensemble Kalman inversion, which applies the ensemble Kalman filter to a very general class of estimation problems. This turns the ensemble Kalman filter basically into a black-box Monte Carlo method. This type of application is still an area of active research [Gar+20].

The fact that filtering can also be applied to dynamical parameter estimation was demonstrated in chapter 4. We presented a modification of the dual extended Kalman filter to continuous-discrete filtering problems in Hilbert spaces, and also outlined an ensemble-based version that solves the dual problem in a very economical way, without requiring derivatives with respect to the parameter and thereby circumventing the need to solve an adjoint equation.

Furthermore, in chapter 5, we noted that if we extend the classical equivalence of the iterated extended Kalman filter with the regularized Gauss-Newton method to dual filtering in Hilbert spaces, we can recover already existing approaches for parameter estimation in partial differential equations. In particular, we noted that the dual iterated extended Kalman filter from section 4.3 is actually equivalent to the sequential version of the reduced approach to parameter identification, as described

for example in [KNS19]. This correspondence can take fruits in two directions: On one hand, by interpreting iterative filtering methods as minimization of a suitable cost functional, we can apply other iterative optimization methods, like the Levenberg-Marquardt method or the Landweber iteration, to solve filtering problems. On the other hand, the probabilistic interpretation of regularization methods yields natural suggestions for the choice of regularization parameters. Furthermore, it allows to combine deterministic regulariziation methods with stochastic filtering techniques, for example with ensemble-based methods. This can, and does (see e.g. [MBG16]) lead to effective methods that combine the "best of two worlds".

Bibliography

[AF05] Azad Shademan and F. Janabi-Sharifi. "Sensitivity analysis of EKF and iterated EKF pose estimation for position-based visual servoing". In: *Proceedings of 2005 IEEE Conference on Control Applications, 2005. CCA 2005.* 2005 IEEE Conference on Control Applications, 2005. CCA 2005. Toronto, Canada: IEEE, 2005, pp. 755–760.

[Alo+15] Francesco Alonge, Tommaso Cangemi, Filippo D'Ippolito, Adriano Fagiolini, and Antonino Sferlazza. "Convergence Analysis of Extended Kalman Filter for Sensorless Control of Induction Motor". In: *IEEE Transactions on Industrial Electronics* 62.4 (Apr. 2015), pp. 2341–2352.

[AS06] Ariel Almendral-Vazquez and Anne Randi Syversveen. *The Ensemble Kalman Filter - theory and applications in oil industry.* SAND/05/06. Norwegian Computing Center, Sept. 12, 2006, p. 25.

[AT67] M. Athans and E. Tse. "A direct derivation of the optimal linear filter using the maximum principle". In: *IEEE Transactions on Automatic Control* 12.6 (Dec. 1967), pp. 690–698.

[Aug+13] Francois Auger, Mickael Hilairet, Josep M. Guerrero, Eric Monmasson, Teresa Orlowska-Kowalska, and Seiichiro Katsura. "Industrial Applications of the Kalman Filter: A Review". In: *IEEE Transactions on Industrial Electronics* 60.12 (Dec. 2013), pp. 5458–5471.

[Bar10] Andrea Barth. "A finite element method for martingale-driven stochastic partial differential equations". In: *Communications on Stochastic Analysis* 4.3 (Sept. 1, 2010).

[BC09] Alan Bain and Dan Crisan. *Fundamentals of Stochastic Filtering.* Stochastic Modelling and Applied Probability. New York: Springer-Verlag, 2009.

[BC93] B.M. Bell and F.W. Cathey. "The iterated Kalman filter update as a Gauss-Newton method". In: *IEEE Transactions on Automatic Control* 38.2 (Feb. 1993), pp. 294–297.

[BDR09] V.I. Bogachev, G. Da Prato, and M. Röckner. "Fokker–Planck equations and maximal dissipativity for Kolmogorov operators with time dependent singular drifts in Hilbert spaces". In: *Journal of Functional Analysis* 256.4 (Feb. 2009), pp. 1269–1298.

[BEW03] Laurent Bertino, Geir Evensen, and Hans Wackernagel. "Sequential Data Assimilation Techniques in Oceanography". In: *International Statistical Review* 71.2 (2003), pp. 223–241.

[BGJ11] Steve Brooks, Andrew Gelman, and Galin Jones. *Handbook of Markov Chain Monte Carlo.* Boca Raton: CRC Press, May 13, 2011. 619 pp.

[BGM19] El Houcine Bergou, Serge Gratton, and Jan Mandel. "On the convergence of a non-linear ensemble Kalman smoother". In: *Applied Numerical Mathematics* 137 (Mar. 2019), pp. 151–168.

[Bog07] Vladimir I. Bogachev. *Measure Theory.* Berlin Heidelberg: Springer-Verlag, 2007.

[Bog15] Vladimir I. Bogachev. *Gaussian Measures.* Providence, RI: American Mathematical Society, Mar. 13, 2015. 433 pp.

[BPR11] Vladimir Bogachev, Giuseppe Da Prato, and Michael Röckner. "Uniqueness for Solutions of Fokker–Planck Equations on Infinite Dimensional Spaces". In: *Communications in Partial Differential Equations* 36.6 (Jan. 17, 2011), pp. 925–939.

[But+09] Mark D. Butala, Richard A. Frazin, Yuguo Chen, and Farzad Kamalabadi. "Tomographic Imaging of Dynamic Objects With the Ensemble Kalman Filter". In: *IEEE Transactions on Image Processing* 18.7 (July 2009), pp. 1573–1587.

[Cho07] Pao-Liu Chow. *Stochastic Partial Differential Equations*. 1st ed. Boca Raton, FL: Chapman and Hall/CRC, Mar. 19, 2007.

[CO12] Yan Chen and Dean S. Oliver. "Ensemble Randomized Maximum Likelihood Method as an Iterative Ensemble Smoother". In: *Mathematical Geosciences* 44.1 (Jan. 2012), pp. 1–26.

[Con94] John B. Conway. *A Course in Functional Analysis*. 2nd. New York: Springer, Jan. 25, 1994.

[CP16] Michal Capinski and Andrzej Polanski. "Parameter Estimation in Systems Biology Models by Using Extended Kalman Filter". In: *Man–Machine Interactions 4*. Ed. by Aleksandra Gruca, Agnieszka Brachman, Stanisław Kozielski, and Tadeusz Czachórski. Advances in Intelligent Systems and Computing. Cham: Springer International Publishing, 2016, pp. 195–204.

[CTL00] R. Clements, P. Tavares, and P. Lima. "Small satellite attitude control based on a Kalman filter". In: *Proceedings of the 2000 IEEE International Symposium on Intelligent Control. Held jointly with the 8th IEEE Mediterranean Conference on Control and Automation (Cat. No.00CH37147)*. Proceedings of the 2000 IEEE International Symposium on Intelligent Control. Held jointly with the 8th IEEE Mediterranean Conference on Control and Automation (Cat. No.00CH37147). ISSN: 2158-9860. July 2000, pp. 79–84.

[CTS10] Saneej B. Chitralekha, Japan J. Trivedi, and Sirish Shah. "Application of the Ensemble Kalman Filter for Characterization and History Matching of Unconventional Oil Reservoirs". In: Canadian Unconventional Resources and International Petroleum Conference. Society of Petroleum Engineers, Jan. 1, 2010.

[Cur75a] Ruth Curtain. "A Survey of Infinite-Dimensional Filtering". In: *SIAM Review* 17.3 (July 1975), pp. 395–411.

[Cur75b] Ruth F. Curtain. "Infinite-Dimensional Filtering". In: *SIAM Journal on Control* 13.1 (Jan. 1, 1975), pp. 89–104.

[DL00] Robert Dautray and Jacques-Louis Lions. *Mathematical Analysis and Numerical Methods for Science and Technology: Volume 5 Evolution Problems I*. Vol. 5. Dautray,R.;Lions,J-L.:Mathematical Analysis,Numerical Methods. Berlin Heidelberg: Springer-Verlag, 2000.

[Dou+19] Randal Douc, Eric Moulines, Pierre Priouret, and Philippe Soulier. *Markov Chains*. 1st ed. 2018. New York, NY: Springer, Jan. 3, 2019. 775 pp.

[DS88] Nelson Dunford and Jacob T. Schwartz. *Linear Operators, Part 1: General Theory*. New York: Wiley-Interscience, Feb. 1988. 872 pp.

[EHN00] Heinz Werner Engl, Martin Hanke, and Andreas Neubauer. *Regularization of Inverse Problems*. Mathematics and Its Applications. Springer Netherlands, 2000.

[EN06] Klaus-Jochen Engel and Rainer Nagel. *A Short Course on Operator Semigroups*. Universitext. New York: Springer-Verlag, 2006.

[ER13] Alexandre A. Emerick and Albert C. Reynolds. "Investigation of the sampling performance of ensemble-based methods with a simple reservoir model". In: *Computational Geosciences* 17.2 (Apr. 1, 2013), pp. 325–350.

[Eva10] Lawrence C Evans. *Partial Differential Equations: Second Edition*. 2nd ed. Graduate Studies in Mathematics 19. American Mathematical Society, 2010.

[Eve+19] Geir Evensen, Patrick N. Raanes, Andreas S. Stordal, and Joakim Hove. "Efficient Implementation of an Iterative Ensemble Smoother for Data Assimilation and Reservoir History Matching". In: *Frontiers in Applied Mathematics and Statistics* 5 (2019).

[Eve03] Geir Evensen. "The Ensemble Kalman Filter: theoretical formulation and practical implementation". In: *Ocean Dynamics* 53.4 (Nov. 1, 2003), pp. 343–367.

[Eve94] Geir Evensen. “Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics”. In: *Journal of Geophysical Research* 99 (C5 1994), p. 10143.

[Fal67] Peter L. Falb. “Infinite-dimensional filtering: The Kalman-Bucy filter in Hilbert space”. In: *Information and Control* 11.1 (July 1967), pp. 102–137.

[Fat+12] Maryam Fatemi, Lennart Svensson, Lars Hammarstrand, and Mark Morelande. “A Study of MAP Estimation Techniques for Nonlinear Filtering”. In: 15th interantional conference on information fusion. Singapore, July 9, 2012, p. 9.

[Fra70] Joel N Franklin. “Well-posed stochastic extensions of ill-posed linear problems”. In: *Journal of Mathematical Analysis and Applications* 31.3 (Sept. 1970), pp. 682–716.

[GA10] Mohinder S. Grewal and Angus P. Andrews. “Applications of Kalman Filtering in Aerospace 1960 to the Present [Historical Perspectives]”. In: *IEEE Control Systems Magazine* 30.3 (June 2010), pp. 69–78.

[Gal16] Jean-François Le Gall. *Brownian Motion, Martingales, and Stochastic Calculus*. Graduate Texts in Mathematics. Springer International Publishing, 2016.

[Gar+20] Alfredo Garbuno-Inigo, Franca Hoffmann, Wuchen Li, and Andrew M. Stuart. “Interacting Langevin Diffusions: Gradient Structure and Ensemble Kalman Sampler”. In: *SIAM Journal on Applied Dynamical Systems* (Jan. 1, 2020), pp. 412–441.

[Gel+13] Andrew Gelman, John B. Carlin, Hal S. Stern, David B. Dunson, Vehtari Aki, and Donald B. Rubin. *Bayesian Data Analysis*. 3 New edition. Boca Raton: Chapman & Hall/CRC, Nov. 7, 2013. 675 pp.

[Gel74] Arthur Gelb. *Applied Optimal Estimation*. Cambridge, Mass: MIT Press, July 15, 1974. 374 pp.

[GO07] Yaqing Gu and Dean S. Oliver. “An Iterative Ensemble Kalman Filter for Multiphase Fluid Flow Data Assimilation”. In: *SPE Journal* 12.4 (Nov. 1, 2007), pp. 438–446.

[Gou+18] R. Gouveia, J. Villanueva, F. Santos, and J. Silva. “Convergence analysis of the extended Kalman filter used in the ultrasonic time-of-flight estimation”. In: *Journal of Physics: Conference Series* 1044 (June 2018), p. 012018.

[Gro07] Christian Grossmann. *Numerical Treatment of Partial Differential Equations*. 2007th ed. Berlin ; New York: Springer, Oct. 4, 2007. 612 pp.

[Hai09] Martin Hairer. “An Introduction to Stochastic PDEs”. Available for download from www.hairer.org. July 24, 2009.

[Hel18] Tapio Helin. “Bayesian inversion: theoretical perspective”. 2018.

[HM01] P L Houtekamer and Herschel L Mitchell. “A Sequential Ensemble Kalman Filter for Atmospheric Data Assimilation”. In: *Monthly Weather Review* 129 (Jan. 2001).

[HZ16] P. L. Houtekamer and Fuqing Zhang. “Review of the Ensemble Kalman Filter for Atmospheric Data Assimilation”. In: *Monthly Weather Review* 144.12 (Dec. 2016), pp. 4489–4532.

[ILS13] Marco A. Iglesias, Kody J. H. Law, and Andrew M. Stuart. “Ensemble Kalman methods for inverse problems”. In: *Inverse Problems* 29.4 (Mar. 2013).

[Jan06] Dariusz Janiszewski. “Extended Kalman Filter Based Speed Sensorless PMSM Control with Load Reconstruction”. In: *IECON 2006 - 32nd Annual Conference on IEEE Industrial Electronics*. IECON 2006 - 32nd Annual Conference on IEEE Industrial Electronics. Nov. 2006, pp. 1465–1468.

[Jaz70] Andrew H. Jazwinski. *Stochastic Processes and Filtering Theory*. New York: Academic Press Inc, 1970. 376 pp.

[JU97] Simon J. Julier and Jeffrey K. Uhlmann. “A New Extension of the Kalman Filter to Nonlinear Systems”. In: *Proceedings, Signal Processing, Sensor Fucion, And Target Recognition VI* 3068 (1997), p. 12.

[Kal02] Olav Kallenberg. *Foundations of Modern Probability*. 2nd ed. Probability and Its Applications. New York: Springer-Verlag, 2002.

[Kal60] Rudolf Emil Kalman. "A New Approach to Linear Filtering and Prediction Problems". In: *Journal of Basic Engineering*. D 82 (1960), pp. 35–45.

[KB61] R. E. Kalman and R. S. Bucy. "New Results in Linear Filtering and Prediction Theory". In: *Journal of Basic Engineering* 83.1 (Mar. 1, 1961), pp. 95–108.

[KK15] Gennady Yu. Kulikov and Maria V. Kulikova. "High-order accurate continuous-discrete extended Kalman filter for chemical engineering". In: *European Journal of Control* 21 (Jan. 1, 2015), pp. 14–26.

[Kle13] Achim Klenke. *Wahrscheinlichkeitstheorie*. 3., überarb. u. erg. Aufl. 2013. Berlin Heidelberg: Springer Spektrum, Mar. 16, 2013. 664 pp.

[KLS10] Mihály Kovács, Stig Larsson, and Fardin Saedpanah. "Finite Element Approximation of the Linear Stochastic Wave Equation with Additive Noise". In: *SIAM Journal on Numerical Analysis* 48.2 (Jan. 2010), pp. 408–427.

[KM17] Ivan Kasanický and Jan Mandel. "On well-posedness of Bayesian data assimilation and inverse problems in Hilbert space". In: (Jan. 28, 2017). arXiv: `1701.08298`.

[KNS19] Barbara Kaltenbacher, Tram Thi Ngoc Nguyen, and Otmar Scherzer. "The tangential cone condition for some coefficient identification model problems in parabolic PDEs". In: (Aug. 3, 2019). arXiv: `1908.01239`.

[Kre03] Arthur J. Krener. "The Convergence of the Extended Kalman Filter". In: *Directions in Mathematical Systems Theory and Optimization*. Ed. by Anders Rantzer and Christopher I. Byrnes. Lecture Notes in Control and Information Sciences. Berlin, Heidelberg: Springer, 2003, pp. 173–182.

[Kre15] Marcel Kreuter. "Sobolev Spaces of Vector-Valued Functions". Master Thesis. Ulm: Ulm University, Apr. 15, 2015.

[KS12] Sturla F. Kvamsdal and Leif K. Sandal. *The Ensemble Kalman Filter in Bioeconomics*. 2012/5. Norwegian School of Economics, Department of Business and Management Science, May 14, 2012.

[Kus11] Harold J Kushner. "Numerical Approximations to Optimal Nonlinear Filters". In: *The Oxford Handbook on Nonlinear Filtering*. Ed. by Dan Crisan and Boris L. Rozovskii. Oxford Handbooks. Oxford University Press, 2011.

[Kus67] H. Kushner. "Nonlinear filtering: The exact dynamical equations satisfied by the conditional mode". In: *IEEE Transactions on Automatic Control* 12.3 (June 1967), pp. 262–267.

[Lan93] Serge Lang. *Real and Functional Analysis*. 3rd ed. Graduate Texts in Mathematics. New York: Springer-Verlag, 1993.

[Lar72] F.M. Larkin. "Gaussian measure in Hilbert space and applications in numerical analysis". In: *Rocky Mountain Journal of Mathematics* 2.3 (Sept. 1972), pp. 379–422.

[LLC10] Faming Liang, Chuanhai Liu, and Raymond Carroll. *Advanced Markov Chain Monte Carlo Methods: Learning from Past Samples*. 1. Chichester, West Sussex, U.K: Wiley, July 16, 2010. 378 pp.

[LPS89] M. S. Lehtinen, L. Paivarinta, and E. Somersalo. "Linear inverse problems for generalised random variables". In: *Inverse Problems* 5.4 (Aug. 1, 1989), pp. 599–612.

[Lun95] Alessandra Lunardi. *Analytic Semigroups and Optimal Regularity in Parabolic Problems*. Modern Birkhäuser Classics. Birkhäuser Basel, 1995.

[Lus96] H. Luschgy. "Linear Estimators and Radonifying Operators". In: *Theory of Probability & Its Applications* 40.1 (Jan. 1996), pp. 167–175.

[MA08] James McNames and Mateo Aboy. "Statistical Modeling of Cardiovascular Signals and Parameter Estimation Based on the Extended Kalman Filter". In: *IEEE Transactions on Biomedical Engineering* 55.1 (Jan. 2008), pp. 119–129.

[Man06] Jan Mandel. *Efficient Implementation of the Ensemble Kalman Filter*. CCM Report No 231. University of Colorado at Denver and Health Sciences Center, May 2006, p. 9.

[Man84] Avi Mandelbaum. "Linear estimators and measurable linear transformations on a Hilbert space". In: *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete* 65.3 (1984), pp. 385–397.

[Mar+12] James Martin, Lucas C. Wilcox, Carsten Burstedde, and Omar Ghattas. "A Stochastic Newton MCMC Method for Large-Scale Statistical Inverse Problems with Application to Seismic Inversion". In: *SIAM Journal on Scientific Computing* 34.3 (Jan. 2012), A1460–A1487.

[MBG16] Jan Mandel, Elhoucine Bergou, and Serge Gratton. "4DVAR by ensemble Kalman smoother". In: *Nonlinear Processes in Geophysics* 23.2 (Mar. 11, 2016), pp. 59–73. arXiv: 1304.5271.

[MCB11] Jan Mandel, Loren Cobb, and Jonathan D. Beezley. "On the convergence of the ensemble Kalman filter". In: *Applications of Mathematics* 56.6 (Dec. 2011), pp. 533–541.

[Mey09] Sean Meyn. *Markov Chains and Stochastic Stability*. 2nd ed. Cambridge ; New York: Cambridge University Press, Apr. 2, 2009. 624 pp.

[Nel00] Alex Tremain Nelson. "Nonlinear estimation and modeling of noisy time-series by dual Kalman filtering methods". PhD thesis. Oregon Graduate Institue of Science and Technology, Sept. 2000.

[Nir] Mahesan Niranjan. "Sequential Tracking in Pricing Financial Options using Model Based and Neural Network Approaches". In: *Advances in Neural Information Processing Systems 9*. NIPS 1996, p. 7.

[NS76] L. Nelson and E. Stear. "The simultaneous on-line estimation of parameters and states in linear systems". In: *IEEE Transactions on Automatic Control* 21.1 (Feb. 1976), pp. 94–98.

[Oli17] Dean S. Oliver. "Metropolized Randomized Maximum Likelihood for Improved Sampling from Multimodal Distributions". In: *SIAM/ASA Journal on Uncertainty Quantification* 5.1 (Jan. 1, 2017), pp. 259–277.

[Oll18] Yann Ollivier. "Online natural gradient as a Kalman filter". In: *Electronic Journal of Statistics* 12.2 (2018), pp. 2930–2961.

[OS89] Sigeru Omatu and John H. Seinfeld. *Distributed parameter systems, theory and applications*. Vol. 4. Oxford Mathematical Monographs, Clarendon Press, 1989. 430 pp.

[Pic91] Jean Picard. "Efficiency of the Extended Kalman Filter for Nonlinear Systems with Small Noise". In: *SIAM Journal on Applied Mathematics* 51.3 (1991), pp. 843–885.

[Pin+18] Basile Pinsard, Arnaud Boutin, Julien Doyon, and Habib Benali. "Integrated fMRI Preprocessing Framework Using Extended Kalman Filter for Estimation of Slice-Wise Motion". In: *Frontiers in Neuroscience* 12 (2018).

[PZ14] Giuseppe Da Prato and Jerzy Zabczyk. *Stochastic Equations in Infinite Dimensions*. 2nd ed. Cambridge, United Kingdom: Cambridge University Press, Apr. 17, 2014. 512 pp.

[Raa16] Patrick Nima Raanes. "On the ensemble Rauch-Tung-Striebel smoother and its equivalence to the ensemble Kalman smoother: On the EnRTS and Its Equivalence to the EnKS". In: *Quarterly Journal of the Royal Meteorological Society* 142.696 (Apr. 2016), pp. 1259–1264.

[Rei+00] K. Reif, S. Gunther, E. Yaz, and R. Unbehauen. "Stochastic stability of the continuous-time extended Kalman filter". In: *IEEE Proceedings - Control Theory and Applications* 147.1 (Jan. 2000), pp. 45–52.

[Rei+99] K. Reif, S. Gunther, E. Yaz, and R. Unbehauen. "Stochastic stability of the discrete-time extended Kalman filter". In: *IEEE Transactions on Automatic Control* 44.4 (Apr. 1999), pp. 714–728.

[Rot+17] Michael Roth, Gustaf Hendeby, Carsten Fritsche, and Fredrik Gustafsson. "The Ensemble Kalman filter: a signal processing perspective". In: *EURASIP Journal on Advances in Signal Processing* 2017.1 (Dec. 2017), p. 56.

[RR04] Gareth O. Roberts and Jeffrey S. Rosenthal. "General state space Markov chains and MCMC algorithms". In: *Probability Surveys* 1 (2004), pp. 20–71.

[RSE19] Patrick Nima Raanes, Andreas Størksen Stordal, and Geir Evensen. "Revising the stochastic iterative ensemble smoother". In: *Nonlinear Processes in Geophysics* 26.3 (Sept. 17, 2019), pp. 325–338.

[Sar13] Simo Sarkka. *Bayesian Filtering and Smoothing*. Cambridge: Cambridge University Press, 2013.

[Sch74] Laurent Schwartz. *Radon Measures on Arbitrary Topological Spaces and Cylindrical Measures*. London: Oxford University Press, Nov. 21, 1974. 408 pp.

[SG07] Shoudong Huang and Gamini Dissanayake. "Convergence and Consistency Analysis for Extended Kalman Filter Based SLAM". In: *IEEE Transactions on Robotics* 23.5 (Oct. 2007), pp. 1036–1049.

[SHB18] Pavel Sakov, Jean-Matthieu Haussaire, and Marc Bocquet. "An iterative ensemble Kalman filter in the presence of additive model error". In: *Quarterly Journal of the Royal Meteorological Society* 144.713 (2018), pp. 1297–1309.

[Sim06] Dan Simon. *Optimal State Estimation: Kalman, H Infinity, and Nonlinear Approaches*. 1. Hoboken, N.J: Wiley-Interscience, July 21, 2006. 552 pp.

[SOB12] Pavel Sakov, Dean S. Oliver, and Laurent Bertino. "An Iterative EnKF for Strongly Nonlinear Systems". In: *Monthly Weather Review* 140.6 (Feb. 9, 2012), pp. 1988–2004.

[SS17] Claudia Schillings and Andrew M. Stuart. "Analysis of the Ensemble Kalman Filter for Inverse Problems". In: *SIAM Journal on Numerical Analysis* 55.3 (Jan. 2017), pp. 1264–1290.

[SS18] C. Schillings and A. M. Stuart. "Convergence analysis of ensemble Kalman inversion: the linear, noisy case". In: *Applicable Analysis* 97.1 (Jan. 2, 2018), pp. 107–123.

[Str60] R. L. Stratonovich. "Conditional Markov Processes". In: *Theory of Probability & Its Applications* 5.2 (Jan. 1, 1960), pp. 156–178.

[Stu10] A. M. Stuart. "Inverse problems: A Bayesian perspective". In: *Acta Numerica* 19 (May 2010), pp. 451–559.

[Tag+18] Amirhossein Taghvaei, Jana de Wiljes, Prashant G. Mehta, and Sebastian Reich. "Kalman Filter and Its Modern Extensions for the Continuous-Time Nonlinear Filtering Problem". In: *Journal of Dynamic Systems, Measurement, and Control* 140.3 (Mar. 1, 2018).

[Tay10] Michael E. Taylor. *Partial Differential Equations II: Qualitative Studies of Linear Equations*. 2nd ed. 2011. New York: Springer, Nov. 10, 2010. 614 pp.

[Tho06] Vidar Thomee. *Galerkin Finite Element Methods for Parabolic Problems*. 2nd ed. 2006. Berlin ; New York: Springer, July 7, 2006. 376 pp.

[TI97] Lloyd N. Trefethen and David Bau III. *Numerical Linear Algebra*. New. Philadelphia: Society for Industrial and Applied Mathematics, June 1, 1997. 184 pp.

[Vak94] N. N. Vakhania. "Canonical Factorization of Gaussian Covariance Operators and Some of Its Applications". In: *Theory of Probability & Its Applications* 38.3 (Sept. 1994), pp. 498–505.

[VTC87] N. Vakhania, Vazha Tarieladze, and S. Chobanyan. *Probability Distributions on Banach Spaces*. Mathematics and its Applications. Springer Netherlands, 1987.

[Wal02] David M. Walker. "Kalman Filtering of Time Series Data". In: *Modelling and Forecasting Financial Data: Techniques of Nonlinear Dynamics*. Ed. by Abdol S. Soofi and Liangyue Cao. Studies in Computational Finance. Boston, MA: Springer US, 2002, pp. 137–157.

[WN01] Eric A. Wan and Alex T. Nelson. “Dual Extended Kalman Filter Methods”. In: Simon Haykin. *Kalman filtering and neural networks*. OCLC: 849927375. New York: Wiley, 2001.

[WV00] E.A. Wan and R. Van Der Merwe. “The unscented Kalman filter for nonlinear estimation”. In: *Proceedings of the IEEE 2000 Adaptive Systems for Signal Processing, Communications, and Control Symposium (Cat. No.00EX373)*. Symposium on Adaptive Systems for Signal Processing Communications and Control. Lake Louise, Alta., Canada: IEEE, 2000, pp. 153–158.

[Yan04] Yubin Yan. “Semidiscrete Galerkin Approximation for a Linear Stochastic Parabolic Partial Differential Equation Driven by an Additive Noise”. In: *BIT Numerical Mathematics* 44.4 (Dec. 2004), pp. 829–847.

[Zak69] Moshe Zakai. “On the optimal filtering of diffusion processes”. In: *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete* 11.3 (Sept. 1, 1969), pp. 230–243.