FACULTY
OF MATHEMATICS
AND PHYSICS
Charles University

MASTER THESIS

Ondřej Vostal

Continuous Time Linear Quadratic Optimal Control

Department of Probability and Mathematical Statistics

Supervisor of the master thesis:	prof. RNDr. Bohdan Maslowski, DrSc.
Study programme:	Mathematics
Study branch:	Probability, Mathematical Statistics, and Econometrics

Prague 2017

I declare that I carried out this master thesis independently, and only with the cited sources, literature and other professional sources.

I understand that my work relates to the rights and obligations under the Act No. 121/2000 Sb., the Copyright Act, as amended, in particular the fact that the Charles University has the right to conclude a license agreement on the use of this work as a school work pursuant to Section 60 subsection 1 of the Copyright Act.

In .... date .... signature of the author

Název práce: Lineárně kvadratické optimální řízení ve spojitém čase

Autor: Bc. Ondřej Vostal

Katedra: Katedra pravděpodobnosti a matematické statistiky

Vedoucí práce: prof. RNDr. Bohdan Maslowski, DrSc.,
Katedra pravděpodobnosti a matematické statistiky

Abstrakt: Podáváme řešení problému adaptivního ergodického stochastického optimálního řízení kdy je budícím procesem frakcionální Brownův pohyb s Hurstovým parametrem $H > 1/2$. Předkládáme vzorec pro výpočet optimálního zpětnovazebného řízení v případě, že je k dispozici silně konzistentní odhad parametrů řízeného systému. Od odhadu rovněž vyžadujeme splnění speciálních podmínek, kvůli čemuž není náš výsledek zcela obecný. Platí např. v případě odhadu nezávislém na budícím frakcionálním Brownově pohybu. V práci rovněž konstruujeme stochastický integrál vhodných deterministických funkcí vzhledem k frakcionálnímu Brownovu pohybu s Hurstovým parametrem $H > 1/2$ přes neomezenou kladnou část reálné osy.

Klíčová slova: Adaptivní ergodické stochastické optimální řízení, frakcionální Brownův pohyb, lineární stochastické diferenciální rovnice

Title: Continuous Time Linear Quadratic Optimal Control

Author: Bc. Ondřej Vostal

Department: Department of Probability and Mathematical Statistics

Supervisor: prof. RNDr. Bohdan Maslowski, DrSc.,
Department of Probability and Mathematical Statistics

Abstract: We partially solve the adaptive ergodic stochastic optimal control problem where the driving process is a fractional Brownian motion with Hurst parameter $H > 1/2$. A formula is provided for an optimal feedback control given a strongly consistent estimator of the parameters of the controlled system is available. There are some special conditions imposed on the estimator which means the results are not completely general. They apply, for example, in the case where the estimator is independent of the driving fractional Brownian motion. In the course of the thesis, construction of stochastic integrals of suitable deterministic functions with respect to fractional Brownian motion with Hurst parameter $H > 1/2$ over the unbounded positive real half-line is presented as well.

Keywords: Adaptive Ergodic Stochastic Optimal Control, Fractional Brownian Motion, Linear Stochastic Differential Equations

My sincere gratitude goes to my supervisor, prof. Bohdan Maslowski. Despite his busy schedule, he always had time for a meeting to discuss any problems I might have had with the topic. His remarks were essential for me to be able to grasp mathematical theory of stochastic optimal control. There is undisputed timeless usefulness in what he helped me learn for which I will be forever grateful.

A great deal of my thanks goes to my wife Martina, who provided me with love essential in overcoming the despair one is struck down with when his work does not seem to bear fruit.

Contents

Introduction

There is no doubt that a vast amount of real-world problems can be formulated in terms of an optimization problem. If one wishes to track the optimal behaviour not only at a singular moment, but through time, the formulation usually changes to optimal control.

Classical control theory is to date well established and developed. A gentle introduction can be found in Zabczyk (2009). In real-world applications, usually, nothing is certain and some applications benefit from accounting for noise which may be present in the controlled system or in the control.

For a long time, the semimartingale processes were the best models of noise available. The theory of semimartingales and classical Itô integration is well developed which makes it approachable in applications. Control theory for stochastic systems with Brownian perturbations is now mature and many of the results may be found in Yong & Zhou (1999).

In recent years, semimartingales turned out to be inadequate for description of natural phenomena that have "long memory". Long memory has been observed to occur in telecommunication connections and asset prices, to name a few. Long memory phenomena cannot be described by the Wiener process, which has independent increments which makes it memoryless. On the other hand, the concept of turbulence in hydrodynamics can be described by self-similar fields with stationary and dependent increments, e.g. Yaglom (2012). This is where the ideas of Kolmogorov (1940) and Mandelbrot & Van Ness (1968) come into play. The fractional Brownian motion they described as an example of long-memory processes can be used to model long memory in many cases. This means that stochastic control of systems with fractional Brownian perturbations becomes of interest.

For the study of optimal control problems, stochastic calculus and the theory of stochastic differential equations are instrumental. For the case of fractional Brownian motion, stochastic integration is developed in Pipiras & Taqqu (2000), Alòs & Nualart (2003), stochastic differential equations are treated in Guerra & Nualart (2008) and the comprehensive monograph of Mishura (2008) containing many topics concerning fractional Brownian motions.

To date, many works in the field of stochastic optimal control were published where the ordinary Brownian motion was replaced by some kind of fractional Brownian motion. Solution of a linear quadratic stochastic optimal control problem with fractional noise with finite horizon in a Hilbert space was treated in Duncan et al. (2012). Solution of linear quadratic stochastic ergodic optimal control problem with fractional noise was given in Kleptsyna et al. (2005) for a single dimension and in Duncan et al. (2015) for problems in Hilbert spaces.

In this thesis, an adaptive ergodic control problem is formulated and solved that is described by a system of linear stochastic differential equations driven by multidimensional fractional Brownian motion and an ergodic quadratic cost functional. For the solution to exist it is required that a strongly consistent parameter estimator be given that is asymptotically independent of the system evolution. In full generality, this can be only assured if the estimator is independent of the Brownian motion driving the system.

The problem consists of finding optimal feed-back control, described by a random process $u = (u(t), t > 0)$. The state of the system is described by a random process $X = (X(t), t > 0)$. The dynamics of the system is modeled by a linear stochastic differential equation driven by a multidimensional fractional Brownian motion. For illustration, we can formally write

$$\begin{aligned} dX(t) &= [A(\theta_0)X(t) + G(t)u(t)]\,dt + \sigma\,dB(t), \quad t > 0 \\ X(0) &= x_0 \end{aligned} \tag{1}$$

for a continuous matrix-valued map $A(\theta)$, matrices G, σ of correct dimensions, a fractional Brownian motion B, a true value of the parameter θ_0 and an initial condition x_0. The optimality of behavior is measured by a functional $J(x_0, u)$ which depends on the initial state of the system x_0 and the exercised control u. The values $J(x_0, u)$ are real and deterministic and hence provide a kind of summarization of the effect of u on the system in all possible situation. To solve the linear quadratic problem, we have to describe how to calculate value of the control $u(t)$ at every time $t > 0$ in order to minimize $J(y, u)$. The control may depend on the random state of the system X, but at every time $t > 0$, we can only use values of $X(s)$ for $s < t$.

In practice, the precise evolution equations (1) are rarely known exactly. This is why it is most practical to have a result which describes how the optimal control looks like if we have to estimate the evolution equations form the behavior of the controlled system, and at the same time we want to control it optimally. In other words, the control cannot depend on the value of θ_0. This is what we strive to deliver in this thesis.

In a single dimension a solution to the adaptive ergodic control problem with fractional Brownian noise was worked out in Duncan et al. (2002). In the narrowest sense, this thesis is an attempt to generalize the results of Duncan et al. (2002) to a multidimensional case. In contrast to Duncan et al. (2002) we did not, however, specify the parameter estimator precisely which in turn cost us the additional requirement of its asymptotic negative quadrant dependence.

The thesis is organized as follows. In Section 1 we formulate classical results needed further in the thesis. Notably, in Section 1.8 construction of stochastic calculus for the fractional Brownian motion with Hurst parameter $H > 1/2$ on the real half line $[0, \infty)$ is presented. In contrast to Mishura (2008), we use the representation of Alòs & Nualart (2003) to establish the representation of the fractional Brownian motion as a specific integral of an ordinary Brownian motion on the real positive half-line and not on the whole real line. The Section 1.9 is rather original as we did not find treatment of stochastic differential equations driven by fractional Brownian motion which would allow random coefficients. The only non-trivial proof, however, is that of Theorem 33 which is a small variation of the proof found in Mishura (2008). In Section 2 the linear quadratic ergodic control problem with fractional noise is formulated and known results are recalled. Most notably, in Section 2.4 results concerning solvability and continuity of the algebraic Riccati equation are formulated and adapted to the main problem of the thesis. In Section 2.5, solution of the ergodic linear quadratic optimal control problem with fractional noise as proved by Duncan et al. (2015) is given and discussed. The main section of the thesis is Section 3. The whole section consists of original results. Important is the main Theorem 56 along with all the lemmas of the section. The original propositions in Section 2.6 also proved to be crucial for our approach to work successfully. Propositions and proofs in Section 2.8 are original as well.

1 Preliminaries

This section is supposed to serve as a collecting place for results established elsewhere which we will use in the course of the thesis. Care has been taken so that all citations are accurate. Proofs are mostly not included. Exceptions to this rule are definition of stochastic integration with respect to fractional Brownian motion given in Section 1.8 and the existence theorem for solutions of stochastic differential equations driven by fractional Brownian motion with random coefficients given in Theorem 33 in Section 1.9.

1.1 Notation

We work with various normed spaces. To simplify notation we overload the absolute value symbol and write $|x|$ for the norm of an object x in the norm of the space into which x belongs. If it is not clear into which normed space x belongs, we clarify the used norm by subscripting the symbol representing the normed space as in $|x|_{\mathbb{R}}$. We use an analogous notational convention for scalar products, i.e. $\langle x, y\rangle$ denotes the scalar product of objects x and y.

We use $\|M\|$ to denote the operator norm of an operator M. It is defined as

$$\|M\| := \sup\{|Mx|;\ |x| \le 1,\, x \in \operatorname{Dom} M\}.$$

We use x^* to denote the adjoint if x is an operator and a transpose if x is a matrix or vector.

For a space with measure $(X, \mathcal{X}, \mu)$, a natural number m and a number $q \ge 1$ we let $L^q(X, \mathbb{R}^m)$ denote the standard Lebesgue space of measurable mappings $f\colon X \to Y$ satisfying $|f|_q < \infty$ where $|f|_q := \left(\int_X |f(x)|^q_{\mathbb{R}^m} \mu(dx)\right)^{\frac{1}{q}}$. If $m = 1$ we write more tersely $L^q(X)$.

The symbol $\mathbb{1}_A$ is used to denote the indicator function of a set A. The symbol $\operatorname{rank} M$ is used to denote rank of a matrix M. The symbol $\mathbb{B}(X)$ is used to denote a Borel σ-algebra generated by a metric space X. The symbols $\sigma(\mathcal{A})$ and $\sigma(Z)$ denote the σ-algebra generated by the set system $\mathcal{A}$ and the random variable Z respectively. The symbol $\mathbb{R}_+$ denotes the nonnegative real axis, i.e. the interval $[0, \infty)$.

1.2 Convergence of functions

This section includes miscellaneous results about convergence of integrals of functions.

Lemma 1. *Fix $a \in \mathbb{R} \cup \{\pm\infty\}$, $K \ge 0$ and $f \in L^1_{loc}(\mathbb{R}_+)$. If $\lim_{t\to\infty} f(t) = a$ then*

$$\lim_{t\to\infty} \frac{1}{t}\int_0^t f(s)\, ds = a.$$

and

$$\lim_{t\to\infty} \int_0^t e^{-K(t-s)} f(s)\, ds = \frac{a}{K}.$$

Proof. We may use the L'Hôpital rule as stated in Theorem 5.13 on p. 109 in Rudin et al. (1964) for the case "c/∞" and obtain the results. □

1.3 Random variables, processes and measurability

Let Ω be an abstract set of elementary events. Let $\mathcal{F}$ be a sigma-algebra on Ω and let $\mathcal{F}(t),\ t \geq 0$ be a filtration of $\mathcal{F}$, meaning that $\mathcal{F}(t)$ is σ-algebra and $\mathcal{F}(t) \subseteq \mathcal{F}$ for all $t \geq 0$ and $\mathcal{F}(s) \subseteq \mathcal{F}(t)$ for all $s < t$.

The quadruplet $(\Omega, \mathcal{F}, \mathcal{F}(t)_{t\geq 0}, \mathbb{P})$ is called a stochastic basis.

Definition 1. We say, that the stochastic basis $(\Omega, \mathcal{F}, \mathcal{F}(t)_{t\geq 0}, \mathbb{P})$ satisfies the usual conditions if

$$\bigcap_{t>s} \mathcal{F}(t) = \mathcal{F}(s), \text{ for all } s \geq 0.$$

and the sigma-algebra of null sets

$$\mathcal{N} := \{N \subseteq \Omega; \exists M \in \mathcal{F}, N \subseteq M, \mathbb{P}M = 0\}$$

is contained in $\mathcal{F}(t)$ for all $t \geq 0$.

The filtered probability space satisfying the usual conditions $(\Omega, \mathcal{F}, \mathcal{F}(t)_{t\geq 0}, \mathbb{P})$ will underlie all our considerations in the main part of the thesis.

1.3.1 Measurability

But first, let us begin with some basic definitions. In line with Kallenberg (2001) p. 4 we define measurable mappings in the following definition.

Definition 2 (Measurable mappings)**.** Let $(X, \mathcal{X})$ and $(Y, \mathcal{Y})$ be a spaces with measure. A mapping $f\colon X \to Y$ is said to be measurable if $f^{-1}(A) \in \mathcal{X}$ holds for all $A \in \mathcal{Y}$. In cases where the spaces are not clear from context, we say that f is measurable as a map from $(X, \mathcal{X})$ to $(Y, \mathcal{Y})$.

Let $\mathcal{A}$ be a sub σ-algebra of $\mathcal{X}$. The mapping f is called $\mathcal{A}$-measurable if it is measurable as a map from $(X, \mathcal{A})$ to $(Y, \mathcal{Y})$.

Let $(Z, \mathcal{Z})$ be a measurable space and $g\colon X \to Z$ be measurable. The mapping f is said to be g-measurable if it is $\sigma(g)$-measurable, where $\sigma(g)$ is the sigma-field generated by the mapping g defined as $\sigma(g) := \{g^{-1}(A); A \in \mathcal{Z}\}$.

Definition 3 (Random variable)**.** Let m be a natural number and let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. A measurable mapping $X\colon (\Omega, \mathcal{F}, \mathbb{P}) \to (\mathbb{R}^m, \mathbb{B}(\mathbb{R}^m))$ is called a random variable on $(\Omega, \mathcal{F}, \mathbb{P})$ with values in $\mathbb{R}^m$, or simply a random variable if the underlying probability space and the dimension of the space of values can be easily inferred from context.

In line with Yong & Zhou (1999), Definition 2.1 on p. 15 the notion of a random process is given by the following definition.

Definition 4 (Random process)**.** Fix a natural number m and let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. A mapping $X\colon \mathbb{R}_+ \times \Omega \to \mathbb{R}^m$ such that $X(t) := X(t, \cdot)$ is a random variable with values in $\mathbb{R}^m$ for all $t > 0$ is called a random process on $(\Omega, \mathcal{F}, \mathbb{P})$ with values in $\mathbb{R}^m$, or simply a random process if the underlying probability space and the dimension of the space of values can be easily inferred from context.

We give a notion of progressive measurability in the following definition in line with Kallenberg (2001), p. 122.

Definition 5 (Progressive measurability)**.** Let $(\Omega, \mathcal{F}, (\mathcal{F}(t))_{t>0}, \mathbb{P})$ be a stochastic basis. A random process $(X(t),\ t \geq 0)$ on $(\Omega, \mathcal{F}, \mathbb{P})$ is called $[\mathcal{F}(t)_{t\geq 0}]$-progressive if the restriction of the process X to $[0,t]\times\Omega$, i.e. the process $(X(s),\ 0 \leq s \leq t)$, is $\mathbb{B}([0,t])\otimes\mathcal{F}(t)$-measurable for every $t \geq 0$. We omit the filtration and talk simply about progressive processes if the filtration can be easily inferred from context. If it is necessary to specify the filtration, we will also write just "$\mathcal{F}(t)$-progressive" instead of quite complex $[\mathcal{F}(t)_{t\geq 0}]$-progressive if no confusion can arise.

1.3.2 Inequalities

The following Theorem contains the statement of the classical Hölder inequality which we use heavily. A proof can be found for example in Kallenberg (2001) as Lemma 1.29 on p. 15.

Theorem 2 (Hölder)**.** *Fix a space with measure $(\Omega, \mathcal{F}, \mu)$. For any measurable functions f and g and numbers $p, q \geq 1$ satisfying $pq = p + q$ we have*

$$\int fg = \left(\int f^p\right)^{\frac{1}{p}} \left(\int g^q\right)^{\frac{1}{q}}.$$

Remark 3. Let $n \in \mathbb{N}$, $p \geq 1$ and $a_1, \ldots, a_n \in \mathbb{R}$. In the Hölder inequality let $\Omega := \{1, \ldots, n\}$, $\mathcal{F}$ be the power set of M and μ be the discrete measure concentrated on the points of Ω satisfying $\mu(\{i\}) = 1$ for all $i \in M$. Then Theorem 2 asserts that we have

$$(a_1 + \ldots + a_n)^p \leq n^{p-1}(a_1^p + \ldots + a_n^p).$$

Clearly, as $n \geq 1$ we have that $n^p \geq n^{p-1}$.

1.3.3 Convergence of random variables

The propositions in this section show how a.s. convergence translates into L^p convergence.

The following definition can be found in Billingsley (2008) on p. 230.

Definition 6. Let $(M, \mathcal{M}, \mu)$ be a space with measure and let $f_n\colon (M, \mathcal{M}, \mu) \to \mathbb{R}$ be measurable functions for all $n \in \mathbb{N}$. The sequence of functions f_n is said to be uniformly integrable if

$$\lim_{c\to\infty} \sup_{n\in\mathbb{N}} \int_{[|f_n|\geq c]} |f_n| d\mu = 0.$$

The following Theorem can be found in Billingsley (2008) as Theorem 16.14 on p. 230.

Theorem 4. *Let $(M, \mathcal{M}, \mu)$ be a space with measure and let $f_n, f\colon (M, \mathcal{M}, \mu) \to \mathbb{R}$ be measurable functions for all $n \in \mathbb{N}$. Let the underlying space with measure be finite, i.e. satisfying $\mu M < \infty$. Let the sequence f_n be uniformly integrable and such that $f_n \to f$ as $n \to \infty$ μ-a.e. in M. Then f is integrable and*

$$\int_M f_n d\mu \to \int_M f d\mu.$$

Note that the dominated convergence theorem is a particular case of Theorem 4 for if there exists g integrable so that $|f_n| < g$ then f_n is uniformly integrable and Theorem 4 applies.

1.4 Linear ordinary differential equations

In this section we present an exposition of basic definitions and theorems in the domain of linear ordinary differential equations. The preliminaries are mainly taken from Sideris (2013). We constrain ourselves to equations defined for non-negative times.

Fix a natural number m. Let $A\colon \mathbb{R}_+ \to \mathbb{R}^{m\times m}$ and $b\colon \mathbb{R}_+ \to \mathbb{R}^m$ be continuous. Let $x_0 \in \mathbb{R}^m$ and $s \geq 0$. Consider a system of linear differential equations written in matrix form

$$\dot{x}(t) = A(t)x(t) + b(t), \qquad \text{for } t > s, \tag{2}$$

$$x(s) = x_0. \tag{3}$$

The condition (3) is called the initial condition. The system (2) is called homogeneous if $b(t) = 0$ for all $t \geq 0$, it is called autonomous if $b(t) = b$ and $A(t) = A$ are constant with respect to the time variable t.

Next, we define the concept of solution to the ordinary differential equation. The definition is a standard one, cf. Section 2.1 in Sideris (2013) on p. 5 or Section 1.1 in Braun & Golubitsky (1983) on p. 1.

Definition 7. Let $I \subseteq \mathbb{R}_+$ be an open interval such that $s \in I$. An absolutely continuous function $x\colon I \to \mathbb{R}^m$ is said to be a solution of the ordinary differential equation (2) with initial condition (3) on the interval I if it is differentiable on I a.e., satisfies $x(s) = x_0$ and

$$\dot{x}(t) = A(t)x(t) + b(t), \qquad \text{for a.a. } t \in I,$$

where by $\dot{x}(t)$ we mean the time derivative of the function x at the point t; it can also be written as $\frac{d}{dt}x(t)$.

The following theorem can be found as Lemma 3.3 in Sideris (2013) on p. 26. It is the classical Grönwall lemma.

Theorem 5 (Grönwall)**.** *Let ξ, φ be nonnegative functions which are continuous on a open interval $(a, b) \subseteq \mathbb{R}$ containing the point $s \in \mathbb{R}$ and let $c \geq 0$ be a real constant. Let*

$$\varphi(t) \leq c + \left| \int_s^t \xi(r)\varphi(r)\, dr \right| \qquad \textit{for all } a < t < b.$$

Then

$$\varphi(t) \leq c \exp \left| \int_s^t \xi(r)\, dr \right| \qquad \textit{for all } a < t < b.$$

We will need a generalization of the Grönwall lemma for dealing with convolutional integrals. We present one such generalization next. The proof can be found in Pachpatte (1998) as Theorem 1.3.2 on p. 13.

Theorem 6. *Let ξ, φ, c and k be nonnegative functions continuous on a closed interval $[a, b] \subseteq \mathbb{R}$, and*

$$\varphi(t) \leq c(t) + k(t) \int_a^t \xi(r)\varphi(r)\, dr \qquad \textit{for all } a \leq t \leq b.$$

Then

$$\varphi(t) \leq c(t) + k(t) \int_a^t c(s)\xi(s) \exp\left(\int_s^t k(r)\xi(r)\, dr \right) ds \qquad \textit{for all } a \leq t \leq b.$$

Remark 7. It is well known that every solution extends to the maximal interval, cf. Theorem 3.4 in Sideris (2013) on p. 27. It is trivial to check that a solution to the equation (2) with initial condition (3) also solves the integral equation

$$x(t) = x_0 + \int_s^t \big(A(r)x(r) + b(r)\big)dr \quad \text{for all } t \geq s.$$

Thus we may write

$$|x(t)| \leq |x_0| + \int_s^t \big(\|A(r)\||x(r)| + |b(r)|\big)dr \quad \text{for all } t \geq s.$$

Notice that $\int_s^t |b(r)|dr$ is nondecreasing in t. Fix arbitrary $T > s$. We can estimate

$$|x(t)| \leq c_T + \int_s^t \|A(r)\||x(r)|\, dr \quad \text{for all } t \geq s.$$

where

$$c_T = |x_0| + \int_s^T |b(r)|\, dr.$$

Using Theorem 5, we obtain

$$|x(t)| \leq c_T \exp \int_s^t \|A(r)\|\, dr \quad \text{for all } T \geq t \geq s.$$

Thanks to the bound we obtained, the maximal interval I satisfies $[s, T] \subseteq I$ for all $T \geq 0$ and hence $[s, \infty) \subseteq I$. This implies that the equation has a global solution for every initial condition x_0 and $s \geq 0$.

The result sketched in Remark 7 together with the local existence of solution proved in Sideris (2013) is formulated as Corollary 3.3. in Sideris (2013) on p. 40. We formulate it in the next corollary.

Corollary 8. *Fix a natural number m. Let $A\colon \mathbb{R}_+ \to \mathbb{R}^{m\times m}$ and $b\colon \mathbb{R}_+ \to \mathbb{R}^m$ be continuous. Then for every $s \in \mathbb{R}_+$ and $x_0 \in \mathbb{R}^m$, the linear ordinary differential equation* (2) *with the initial condition* (3) *has a unique solution x on $\mathbb{R}_+$, i.e. the solution x is unique and global.*

In what follows, we omit specification of the domain of definition of the solutions to a linear ordinary differential equation. Thanks to Corollary 8, we can do this without any loss of generality, since every solution can be extended to $\mathbb{R}_+$ and thus we may implicitly assume that the domain of every solution is $\mathbb{R}_+$. It alleviates us from complexity we would not benefit from.

Definition 8. Fix a natural number m. A two-parameter matrix-valued map $S\colon \mathbb{R}_+^2 \to \mathbb{R}^{m\times m}$ is called the flow map of the system (2), or simply the flow map of A, if it satisfies

$$x^{r,x_0}(t) = S(t, r)x_0, \text{ for all } t \geq r \geq 0 \text{ and all } x_0 \in \mathbb{R}^m,$$

where x^{r,x_0} is a solution to the ordinary differential equation

$$\dot{x}(t) = A(t)x(t), \quad t \geq 0$$

with the initial condition $x^{r,x_0}(r) = x_0$. We write $S(t) := S(t,0)$ and call $S(t)$ the fundamental matrix solution.

The matrix $S(t,s)$ is called the state transition matrix. The special case $S(t,0)$ is called the fundamental matrix solution. In the sequel we will omit the second parameter and write the fundamental matrix solution as a single parameter map $S(t) := S(t,0)$.

The following lemma presents basic properties of the flow map. It is based on Lemma 3.6. in Sideris (2013) on p. 34 which is formulated for the case of a general ordinary differential equation which means that everything it asserts holds also for the flow map of a linear ordinary differential equation with the additional simplification, that we can ignore the domain of the flow map as it can be always assumed to equal $\mathbb{R}_+$. We present the simplified version of the lemma.

Lemma 9. *Let S be a flow map of A. Then*

(i) $S(t,t_0) = S(t,r) \circ S(r,t_0)$ *for all* $t, r, t_0 \geq 0$.

(ii) $S(t,t) = \mathrm{Id}_{m\times m}$ *for all* $t \geq 0$.

(iii) $S(r,t) \circ S(t,r) = \mathrm{Id}_{m\times m}$ *for all* $t, r \geq 0$.

(iv) $S(t,t_0)$ *is a homeomorphism from* $\mathbb{R}^m$ *to* $\mathbb{R}^m$.

Note that property (iii) says that $S(t,t_0)$ is invertible and that the inverse can be calculated as $S(t,r)^{-1} = S(r,t)$ for all $t, r \geq 0$.

It is easy to verify, that

$$\frac{d}{dt} S(t,s) = A(t)S(t), \quad t > s. \tag{4}$$

The following theorem, found in Sideris (2013) as Theorem 4.1 on p 54, is the most fundamental reason why studying linear ordinary differential equation is very fruitful. It gives us means of expressing solutions of the inhomogeneous ordinary differential equation (2) in term of the flow map of A.

Theorem 10 (Variation of constants)**.** *Fix $m \in \mathbb{N}$. Let $A\colon \mathbb{R}_+ \to \mathbb{R}^{m\times m}$ and $b\colon \mathbb{R}_+ \to \mathbb{R}^m$ be continuous. Let S be the flow map of A. Fix $x_0 \in \mathbb{R}^m$ and $s \geq 0$. The function*

$$x(t) = S(t,s)x_0 + \int_s^t S(t,r)b(r)\,dr \qquad \text{for } t \geq s$$

is the unique global solution of (2) *with initial condition* (3)*.*

If $m = 1$, we can readily calculate the flow map of A using a rather simple formula. This is the subject of the next Proposition 11.

Proposition 11. *Let $a\colon \mathbb{R}_+ \to \mathbb{R}$ be continuous. The function*

$$S(t,s) = \exp\left\{\int_s^t a(r)\,dr\right\} \qquad \text{for } t \geq s \geq 0$$

is the flow map of a in the sense of Definition 8.

1.5 Linear ordinary differential equations with constant coefficients

If the system (2) is autonomous, i.e. the coefficients A and b are constant with respect to the time variable t, the unique global solution of the system can be expressed explicitly in terms of the matrix exponential. Formulating this result precisely is the topic of this section.

Based on Definition 2.2 in Sideris (2013) on p. 6 we define the matrix exponential. It plays a fundamental role in the explicit formula for the fundamental matrix solution of a system of ordinary differential equations with constant coefficients.

Definition 9. Fix a natural number m. Given a square matrix $A \in \mathbb{R}^{m\times m}$ we define the matrix exponential of A to be the square-matrix-valued map

$$\exp\{At\} := \sum_{k=0}^{\infty} \frac{1}{k!} A^k t^k.$$

The statement and proof of the following proposition can be found in Sideris (2013) as Lemma 2.1 on p. 6.

Proposition 12. *Fix a natural number m. Given square matrices $A, B \in \mathbb{R}^{m\times m}$ we have*

(i) $\exp At$ *is defined for all* $t \in \mathbb{R}$, *moreover* $\|\exp At\| \leq \exp \|A\||t|$.

(ii) $\exp\{A+B\} = \exp A \exp B = \exp B \exp A$ *provided* A *and* B *commute, i.e.* $AB = BA$.

(iii) $\exp A(t+s) = \exp At \exp As = \exp As \exp At$ *for all* $t, s \in \mathbb{R}$.

(iv) $\exp At$ *is invertible for all* $t \in \mathbb{R}$, *and* $(\exp At)^{-1} = \exp\{-At\}$.

(v) $\frac{d}{dt} \exp At = A \exp At = (\exp At)A$.

We assess the measurability properties of the matrix exponential in the following lemma.

Lemma 13. *Fix a natural number m. The matrix exponential* $\exp\{At\}$ *for* $A \in \mathbb{R}^{m\times m}$ *and* $t \in \mathbb{R}$ *is a* $[\mathbb{B}\,\mathbb{R}^{m\times m} \otimes \mathbb{B}\mathbb{R}]$*-measurable function*

Proof. The matrix exponential is defined as a series, cf. Definition 9

$$\exp A := \sum_{k=0}^{\infty} \frac{1}{k!} A^k t^k.$$

The partial sums

$$s_N = \sum_{k=0}^{N} \frac{1}{k!} A^k t^k.$$

are measurable since the coordinate functions s_N^{ij} are polynomials in $a_{ij} \in \mathbb{R}$ and $t \in \mathbb{R}$ for $i, j = 1, \ldots, m$ of finite order and hence measurable functions. By Lemma 1.10 in Kallenberg (2001) on p. 6 the matrix exponential is measurable since it is a limit of measurable functions s_N as $N \to \infty$. □

For ordinary differential equations with constant coefficients, the fundamental matrix solution can be obtained explicitly via the matrix exponential as the following proposition describes. It follows immediately from Theorem 2.1 on p. 8 in Sideris (2013).

Proposition 14. *Fix a natural number m and a matrix $A \in \mathbb{R}^{m\times m}$. The fundamental solution matrix of A is $\exp\{At\}$ for $t \geq 0$. The flow map of A is thus given by $S(t,s) := \exp\{A(t-s)\}$ for all $t \geq s \geq 0$.*

Corollary 15 (Variation of constants for systems with constant coefficients)**.** *Fix a natural number m. Let $A \in \mathbb{R}^{m\times m}$ and $b \in \mathbb{R}^m$ be constant. Fix $x_0 \in \mathbb{R}^m$ and $s \geq 0$. The function*

$$x(t) = \exp\{A(t-s)\}x_0 + \int_s^t \exp\{A(t-r)\}\, b\, dr \qquad \text{for } t \geq s \tag{5}$$

is the unique global solution of (2) *with initial condition* (3)*.*

Proof. By Proposition 14 we know that $S(t,r) = \exp\{A(t-r)\}$ is the flow map of A. Using Theorem 10 we obtain (5). □

Thanks to the variation of constants formula for systems with constant coefficients and continuity of solutions with respect to initial conditions, the following important proposition holds which states that the matrix exponential is continuous.

Proposition 16. *Fix a natural number m. The matrix exponential*

$$\mathbb{R}^{m\times m} \times \mathbb{R} \to \mathbb{R}^{m\times m} \colon (A,t) \mapsto \exp\{At\}$$

is continuous.

Proof. The proof simply follows from the continuity of solutions of ordinary differential equations on initial conditions which is described in Theorem 3.5 on p. 29 in Sideris (2013). □

The following proposition estimates distance between solutions of stable equations using the distance of their system matrices and hence establishes a continuity property of uniformly stable matrices.

Proposition 17. *Fix a natural number m. Let two stable matrices $A, B \in \mathbb{R}^{m\times m}$ be given, i.e. there exist positive constants M_1, K_1 and M_2, K_2 so that*

$$\|\exp At\| \leq M_1 e^{-K_1 t} \quad \text{and} \quad \|\exp Bt\| \leq M_2 e^{-K_2 t} \quad \text{for all } t \geq 0.$$

If $K_2 > K_1$, then

$$\|\exp\{At\} - \exp\{Bt\}\| \leq \frac{M_1 M_2}{K_2 - K_1}\|A - B\| \left(e^{-K_1 t} - e^{-K_2 t}\right) \quad \text{for all } t \geq 0.$$

If $K := K_1 = K_2$, then

$$\|\exp\{At\} - \exp\{Bt\}\| \leq M_1 M_2 \|A - B\| t e^{-Kt} \quad \text{for all } t \geq 0.$$

Proof. The proof follows from continuous dependence of the solution on parameter can be found in Sideris (2013) as Theorem 6.2 on p. 92. It can be, however, also easily directly proved by using variation of constants on the difference of equations having as system matrices A and B. We give such proof now.

Fix arbitrary $x_0 \in \mathbb{R}^m$. Let two differential equations be given

$$\begin{aligned} \dot{x}(t) &= Ax(t), \qquad \text{for all } t > 0, \quad x(0) = x_0, \\ \dot{y}(t) &= By(t), \qquad \text{for all } t > 0, \quad y(0) = x_0. \end{aligned} \tag{6}$$

The equations have unique global solution by Corollary 8. They are given by Corollary 15 as

$$x(t) = \exp\{At\}x_0, \quad y(t) = \exp\{Bt\}x_0 \quad \text{for } t \geq 0. \tag{7}$$

We can denote $z(t) = x(t) - y(t)$ for $t \geq 0$ and write

$$\dot{z}(t) = (A - B)x(t) + Bz(t) \quad \text{for all } t > 0, \quad z(0) = 0.$$

By Corollary 15 the unique solution z can also be expressed as

$$z(t) = \int_0^t \exp\{B(t-s)\}(A-B)x(s)\,ds, \quad t \geq 0.$$

By assumed stability, we may estimate

$$\begin{aligned} |z(t)| \leq \int_0^t M_2 e^{-K_2(t-s)} \|A - B\| M_1 e^{-K_1 s} |x_0|\,ds \\ = |x_0| e^{-K_1 t} \int_0^t M_2 e^{-(K_2 - K_1)(t-s)} \|A - B\| M_1\,ds. \end{aligned}$$

Let now $K_2 > K_1$. Then obviously

$$|z(t)| \leq \frac{M_1 M_2}{K_2 - K_1}(e^{-K_1 t} - e^{-K_2 t})\|A - B\|\,|x_0|, \quad t \geq 0. \tag{8}$$

If, on the other hand, $K := K_1 = K_2$, we obtain

$$|z(t)| \leq M_1 M_2 e^{-K_1 t} t\,\|A - B\|\,|x_0|, \quad t \geq 0. \tag{9}$$

We now only have to realize that $z(t) = (\exp\{At\} - \exp\{Bt\})x_0$ since the estimate (8) does not depend on x_0 which was arbitrary. The estimates have the form $|z(t)| = |(\exp\{At\} - \exp\{Bt\})x_0| \leq f(t)|x_0|$ for some function f and by definition of the operator norm, this is enough for the norm to satisfy

$$\|\exp\{At\} - \exp\{Bt\}\| = f(t) \quad \text{for } t \geq 0.$$

$\square$

1.6 Classical control theory

In this section, we introduce some concepts of classical control theory. They are important for us since utilizing these concepts, conditions on the parameters of the system (43) can be imposed so that the related optimal control problem has a solution.

Our main source for this chapter is Zabczyk (2009). Equivalent definitions of the concepts can, however, be found in many control theory related monographs, as is for example Lancaster & Rodman (1995).

Fix natural numbers m and k. For matrices $a \in \mathbb{R}^{m\times m}$, $b \in \mathbb{R}^{m\times k}$, a control $u \in L^1_{loc}(\mathbb{R}_+;\mathbb{R}^k)$ and an initial condition $x_0 \in \mathbb{R}^m$ let the linear ordinary differential equation

$$\begin{aligned} \dot{x}(t) &= ax(t) + bu(t), \quad t > 0 \\ x(0) &= x_0 \end{aligned} \tag{10}$$

be given. By setting the values of the function u one can excercise control over the trajectory of the system x. This is why u is called the control.

By Corollary 15 we know that the fundamental matrix solution of a is $\exp at$ for $t \geq 0$ and that

$$x(t) := \exp\{at\}x_0 + \int_0^t \exp\{a(t-r)\}bu(r)\,dr \quad \text{for } t \geq 0 \tag{11}$$

is the unique global solution of the equation (10). For a control $u \in L^1_{loc}(\mathbb{R}_+;\mathbb{R}^k)$ and an initial condition $x_0 \in \mathbb{R}^m$ we denote $x^{x_0,u}(t)$ the solution of (10) given by (11).

1.6.1 Controllability

In line with Zabczyk (2009), section 1.2 on p. 14 ff. we state the following definitions.

Definition 10. We say that a control $u \in L^1_{loc}(\mathbb{R}_+;\mathbb{R}^k)$ transfers a state $x_0 \in \mathbb{R}^m$ to state $x_1 \in \mathbb{R}^m$ in time $T > 0$ if

$$x^{x_0,u}(T) = x_1.$$

Definition 11. We say that the pair (a, b) is controllable if for each initial condition $x_0 \in \mathbb{R}^m$ and each terminal condition $x_1 \in \mathbb{R}^m$ there exists a control $u \in L^1_{loc}(\mathbb{R}_+;\mathbb{R}^k)$ and time $T > 0$ so that u transfers x_0 to x_1 in time T.

The following theorem along with proof can be found as Theorem 1.2 in Zabczyk (2009) on p. 17.

We will denote as $[a|b]$ the $m\times mk$ matrix $(b\ ab\ a^2b\ \dots\ a^{m-1}b)$, i.e. the matrix composed by placing matrices b, ab, ..., $a^{m-1}b$ next to each other.

Theorem 18. *The following conditions are equivalent.*

(i) The pair (a, b) is controllable.

(ii) $\operatorname{rank}[a|b] = m$.

The condition (ii) is called the Kalman rank condition.

1.6.2 Observability

Set $b = 0$ in (10) and augment it with an observation equation. This way, we obtain a system

$$\begin{aligned} \dot{z}(t) &= az(t), \quad t > 0 \\ z(0) &= x_0, \quad t \geq 0 \\ w(t) &= cz(t) \end{aligned} \tag{12}$$

where $c \in \mathbb{R}^{m \times k}$ is the observation matrix. Let $z^{x_0}(t)$ denote its solution for $t \geq 0$.

Definition 12. The pair (c, a) is said to be observable if for all nonzero initial conditions $x_0 \in \mathbb{R}^m$, $x_0 \neq 0$ there exists a time $t > 0$ such that

$$w(t) = cz^{x_0}(t) \neq 0.$$

The notion of observability we present here is different from Zabczyk (2009). We say that a pair (a, c) is observable if and only if the pair (c, a) is observable by the definition in Zabczyk (2009). Our notion complies with the notion Lancaster & Rodman (1995) use.

For observability a theorem similar to Theorem 18 holds. It can be found in Zabczyk (2009) as Theorem 1.6 on p. 25.

Theorem 19. *The following conditions are equivalent.*

(i) The pair (c, a) is observable.

(ii) $\operatorname{rank}[a^*|c^*] = m$.

Remark 20. It follows from Theorems 18 and 19 that the pair (c, a) is observable if and only if the pair (a^*, c^*) is controllable.

1.6.3 Stability

Let us study the linear system

$$\begin{aligned} \dot{z}(t) &= az(t), \quad t > 0 \\ z(0) &= x_0 \end{aligned} \tag{13}$$

for $t \geq 0$. Let $z^{x_0}(t)$ denote its solution for $t \geq 0$.

Definition 13. We say that the matrix a is stable if for arbitrary $x_0 \in \mathbb{R}^m$ we have that $z^{x_0}(t) \to 0$ as $t \to \infty$.

The following theorem is of paramount importance. It states that the stability of a linear system is equivalent with exponential stability and the speed of convergence depends only on the real parts of eigenvalues of the matrix a. It is a combination of Theorem 2.3 and Lemma 2.1 from Zabczyk (2009) on p. 30.

Theorem 21. *Let $a \in \mathbb{R}^{m \times m}$. Denote $\omega^* := \sup\{\Re\lambda; \lambda \text{ is an eigenvalue of } a\}$. The following conditions are equivalent.*

(i) We have $z^{x_0}(t) \to 0$ as $t \to \infty$.

(ii) *There exist positive constants M and K, depending only on ω^*, so that $\|\exp at\| \le Me^{-Kt}$ for all $t \ge 0$.*

(iii) *We have $\omega^* < 0$.*

We will study systems (13) that depend on a parameter. The interest will thus lie in the stability properties of system of the form (13) for $a := a(\theta)$ with $\theta \in \Theta$ for some parameter set Θ. The concept of stability can be strengthened by requiring the stability to be uniform with respect to the parameter as in the following definition.

Definition 14. Let Θ be an abstract set and $a\colon \Theta \to \mathbb{R}^{m\times m}$. We say, that the matrix $a(\theta)$ is stable uniformly with respect to $\theta \in \Theta$ if there exists a negative constant ω^* so that $\Re\lambda \le \omega^*$ for all the eigenvalues λ of all the matrices $a(\theta)$, $\theta \in \Theta$.

We will abuse notation and say that $\Lambda(t;\theta)$ is uniformly stable with respect to $\theta \in \Theta$ if $a(\theta)$ is uniformly stable with respect to $\theta \in \Theta$ and $\Lambda(t;\theta) = \exp\{a(\theta)t\}$, $t \ge 0$, $\theta \in \Theta$. Similarly for stable matrices and stable matrix exponentials.

Corollary 22. *Let Θ be an abstract set and $a\colon \Theta \to \mathbb{R}^{m\times m}$ and let $a(\theta)$ be stable uniformly with respect to θ. Then there exist positive constants M and K so that*

$$\|\exp\{a(\theta)t\}\| \le Me^{-Kt} \quad \text{for all } t \ge 0 \text{ and } \theta \in \Theta.$$

Proof. The statement simply follows from Theorem 21. □

1.6.4 Stabilizability

Consider again the controlled system (10) which reads

$$\begin{aligned} \dot{x}(t) &= ax(t) + bu(t), \quad t > 0 \\ x(0) &= x_0 \end{aligned} \tag{14}$$

for $a \in \mathbb{R}^{m\times m}$ and $b \in \mathbb{R}^{m\times k}$.

Definition 15. We say that the pair (a, b) is stabilizable if there exists a matrix $\kappa \in \mathbb{R}^{k\times m}$ so that $a + b\kappa$ is a stable matrix.

If (a, b) is stabilizable, κ is the stabilizing matrix from Definition 15 and the control is given in the so called feedback form

$$u(t) := \kappa x(t), \quad t \ge 0,$$

then for all $x_0 \in \mathbb{R}^m$ there exist positive constants M and K so that the solution of (14) satisfies

$$|x^{x_0,u}(t)| = |x^{x_0,\kappa x}(t)| \le Me^{-Kt} \quad \text{for all } t \ge 0.$$

Definition 16. The pair (a, b) is called completely stabilizable if for all positive K there exists a matrix $\kappa \in \mathbb{R}^{k\times m}$ and a positive constant M so that the solution of (14) for arbitrary initial condition $x_0 \in \mathbb{R}^m$ satisfies

$$|x^{x_0,\kappa x}(t)| \le Me^{-Kt}|x|_0 \quad \text{for all } t \ge 0.$$

It is immediately clear that if the pair (a, b) is completely stabilizable, it is also stabilizable.

The following theorem, can be found in Zabczyk (2009) as Theorem 2.9, connects controllability and stabilizability.

Theorem 23. *The following conditions are equivalent.*

(i) The pair (a, b) is completely stabilizable.

(ii) The pair (a, b) is controllable.

For stabilizability we would like to have a notion which connects stabilizability properties with a set of parametrized matrices. This is what we introduce in the following definition.

Definition 17. Let Θ be an abstract set, $a\colon \Theta \to \mathbb{R}^{m\times m}$ and $b \in \mathbb{R}^{m\times k}$. We say that the pairs $(a(\theta), b)$ are stabilizable uniformly with respect to $\theta \in \Theta$ if there exists a matrix $\kappa \in \mathbb{R}^{k\times m}$ so that the matrices $a(\theta) + b\kappa$ are stable uniformly with respect to $\theta \in \Theta$.

Remark 24. It follows easily from definition that (a, b) is stabilizable if and only if $(a, -b)$ is.

The following useful proposition can be found as Lemma 4.5.3. in Lancaster & Rodman (1995) on p. 93. It shows that the stability of a pair is invariant to certain transformations of the matrices a and b. Remark 24 is a special case of this fact.

Proposition 25. *Let $a \in \mathbb{R}^{m\times m}$ and $b \in \mathbb{R}^{m\times k}$. Then (a, b) is stabilizable if and only if $(a + b\kappa, bl)$ is stabilizable for any $(k \times m)$-matrix κ and any $(m \times p)$-matrix l, for arbitrary $p \in \mathbb{N}$, for which the linear subspace generated by bl is equal to the linear subspace generated by b.*

1.6.5 Detectability

Detectability is a dual counterpart of stabilizability. Here, as for observability, we use the notions from Lancaster & Rodman (1995) rather than the definitions from Zabczyk (2009). The meaning is analogous, but notation may differ.

Definition 18. We say that the pair (c, a) is detectable if there exists a matrix $l \in \mathbb{R}^{m\times k}$ so that the matrix $a + lc$ is stable.

By Theorem 23 it follows that controllability implies stabilizability. In a similar way, observability implies detectability: Let the pair (c, a) be observable. Then (a^*, c^*) is controllable by definition. Hence by Theorem 23 there exists a matrix κ such that $a^* + c^*\kappa$ is stable. Therefore the matrix $a + \kappa^* c$ is stable. We see that it is enough to set $l = k^*$ and we obtain detectability of the pair (c, a).

We will need notion of uniform detectability, analogous to uniform stabilizability. We define this notion next.

Definition 19. Let Θ be an abstract set, $a: \Theta \to \mathbb{R}^{m\times m}$ and $c \in \mathbb{R}^{k\times m}$. We say that the pairs $(c, a(\theta))$ are detectable uniformly with respect to $\theta \in \Theta$ if there exists a matrix $l \in \mathbb{R}^{m\times k}$ so that the matrices $a(\theta) + lc$ are stable uniformly with respect to $\theta \in \Theta$.

Remark 26. Note that it follows trivially from definition that the pair (a, c) is detectable if and only if the pair (c^*, a^*) is stabilizable.

1.7 Fractional Brownian motion

Fix $H \in (0, 1)$ and a stochastic basis $(\Omega, \mathcal{F}, \mathcal{F}(t)_{t\geq 0}, \mathbb{P})$. A Gaussian stochastic process $B = (B(t), t \geq 0)$ on $(\Omega, \mathcal{F}, \mathcal{F}(t)_{t\geq 0}, \mathbb{P})$ with $B(0) = 0$, zero mean and covariance function given by

$$R(s,t) := \mathbb{E}[B(s)B(t)] = \frac{1}{2}\left[s^{2H} + t^{2H} - |s-t|^{2H}\right] \qquad \text{for } s,t > 0$$

is called a fractional Brownian motion with Hurst parameter H. It was introduced by Kolmogorov (1940) and later studied by Mandelbrot & Van Ness (1968) as Guerra & Nualart (2008) note. For a review of some applications the reader may be interested in peeking into Decreusefond & Üstünel (1998).

Calculate $\mathbb{E}|B(t) - B(s)|^2 = |t-s|^{2H}$ for all $t, s \geq 0$. By the equivalency of the moments of Gaussian distribution, cf. Kallenberg (2001), and the Kolmogorov continuity criterion, to be found e.g. as Theorem 3.23 in Kallenberg (2001) on p. 57, we may deduce that B has a version with all trajectories continuous. We will always assume, without loss of generality, that we work with the version of B which has all trajectories continuous.

Definition 20. Fix a natural number n and a Hurst parameter $0 \leq H \leq 1$. Let $B_1, \ldots, B_n$ be independent fractional Brownian motions with a common Hurst parameter H. The random process $B = (B_1, \ldots, B_n)^*$ is called an n-dimensional fractional Brownian motion with Hurst parameter H.

1.8 Stochastic integrals of deterministic functions

The integration with respect to fractional Brownian motion poses a major challenge. It is mainly because of the peculiar nature of fractional Brownian motion which has even more unpleasant properties than the ordinary Brownian motion. On one hand, much like for an ordinary Brownian motion, the paths of fractional Brownian motion have infinite variation a.s., hence standard methods of Lebesgue-Stieltjes integration cannot be used. What is more, fractional Brownian motion is not a semi-martingale either, hence one cannot use the classical Itô's stochastic calculus. Numerous attempts to develop some kind of stochastic integration for the case of fractional Brownian motion have been investigated and many of those are summarized in the work of Pipiras & Taqqu (2000).

In this section, we present a construction that enables us to integrate with respect to fractional Brownian motion with Hurst parameter $H > 1/2$ certain class of deterministic functions, i.e. we will define the meaning of the symbol $\int f dB$ for functions f from this class. We proceed in accord with Pipiras & Taqqu (2000). The resulting construction agrees with the extension for integration of stochastic processes which can be found in Alòs & Nualart (2003). An agreeing construction can also be found in Mémin et al. (2001) or Mishura (2008).

In what follows, we fix $H > 1/2$. Let $B := (B(t),\ t \geq 0)$ be a fractional Brownian motion with Hurst parameter H. Require also that B be $\mathcal{F}(t)$-progressive.

As Pipiras & Taqqu (2000) note on p. 2: In applications, one likes to view the integral $\int f dB$ to be approximated by

$$\sum_{i=0}^{n-1} f(t_i)[B(t_{i+1}) - B(t_i)], \tag{15}$$

where $0 = t_0 < \ldots < t_n < \infty$.

To construct a stochastic integral of a deterministic integrand, we may proceed along the requirements put forward by (15). For simple functions, we can define the integral so as to agree with (15) and then extend it to a broad class of functions which include $L^2(\mathbb{R}_+)$.

A simple function is a function of the form

$$f(t) = \sum_{i=0}^{n-1} a_i \mathbb{1}_{[t_i,t_{i+1})}(t), \tag{16}$$

where $n \in \mathbb{N}$, $0 = t_0 < t_1 < \ldots < t_n < \infty$, $a_i \in \mathbb{R}$. Let $\mathcal{E}$ denote the linear space of simple functions. For a simple function $f \in \mathcal{E}$ of the form (16), we define the stochastic integral with respect to B to be the random variable $I(f)$ given by

$$I(f) := \sum_{i=0}^{n-1} a_i [B(t_{i+1}) - B(t_i)].$$

Let $\mathcal{I} := \{I(f); f \in \mathcal{E}\}$, i.e. $\mathcal{I}$ denotes the linear space of Gaussian variables of the special form $I(f)$ for some simple function f. Let $\mathcal{B}$ denote the closure of $\mathcal{I}$ in $L^2(\Omega)$, i.e.

$$\mathcal{B} := \{X : (\Omega, \mathcal{F}, \mathbb{P}) \to \mathbb{R};\ I(f_n) \to X \text{ in } L^2(\Omega) \text{ for some sequence } f_n \text{ in } \mathcal{E}\}$$

The map I maps the space of simple functions $\mathcal{E}$ into the linear space of Gaussian variables $\mathcal{I}$ which is a subspace of $\mathcal{B}$ which in turn is a subspace of a complete space $L^2(\Omega)$.

The construction of the integral of deterministic functions we employ is based on the following proposition of Pipiras & Taqqu (2000, Proposition 2.1, p. 5). The original formulation is made for functions on the real line. Modifying the proof to functions on the positive real half-line is trivial which is why we omit it.

Proposition 27. *Let $\mathcal{C}$ be a set of deterministic functions on the $\mathbb{R}_+$ such that*

(i) $\mathcal{C}$ is an inner product space with an inner product $\langle f, g\rangle_{\mathcal{C}}$, for $f, g \in \mathcal{C}$;

(ii) $\mathcal{E} \subseteq \mathcal{C}$ and $\langle f, g\rangle_{\mathcal{C}} = \mathbb{E}\, I(f)I(g)$ for $f, g \in \mathcal{E}$;

(iii) $\mathcal{E}$ is dense in $\mathcal{C}$.

Then

(a) there is an isometry between the spaces $\mathcal{C}$ and a linear subspace of $\mathcal{B}$;

(b) the spaces $\mathcal{C}$ and $\mathcal{B}$ are isometric if and only if $\mathcal{C}$ is a complete space.

Let us present a classical example of the use of the proposition. In the standard Brownian motion case, we can take for $\mathcal{C}$ the space $L^2(\mathbb{R}_+)$. The space $L^2(\mathbb{R}_+)$ is known to be complete. Hence Proposition 27 implies that in the case $H = 1/2$, the space $\mathcal{B}$ is isometric to $L^2(\mathbb{R}_+)$.

The case $H > 1/2$ is by far not that simple. As Pipiras & Taqqu (2000) note on p. 6

> Observe that the isometry map I might depend on the inner product space $\mathcal{C}$. In other words, if $\mathcal{C}_1$ and $\mathcal{C}_2$ are two different classes of functions that satisfy Proposition 27, then, a priori, it is not clear, whether the corresponding isometry maps, say I_1 and I_2 are equal on $\mathcal{C}_1 \cap \mathcal{C}_2$ (A necessary and sufficient condition for this fact is $\langle f, g\rangle_{\mathcal{C}_1} = \langle f, g\rangle_{\mathcal{C}_2}$ for all f, g in $\mathcal{C}_1 \cap \mathcal{C}_2$.)

We choose the space $\mathcal{C}$ and the scalar product so as to agree with standard literature. As we already noted, although the construction is not identical in all the steps, the construction presented here agrees with Alòs & Nualart (2003), Mémin et al. (2001) or Mishura (2008). This is because all the constructions use identical scalar product on step functions given by (17) and the space $\mathcal{C}$ is the largest space of deterministic real functions which fits into the closure of $\mathcal{E}$ in this scalar product. Hence in both constructions both sets $\mathcal{C}$ and the scalar products agree which is bound to lead to identical results eventually.

Let $\alpha_H = H(2H-1)$. We have that

$$R(s,t) = \alpha_H \int_0^t \int_0^s |u-v|^{2H-2}\, du\, dv \qquad \text{for all } s,t > 0$$

On the space of simple functions we define a scalar product

$$\langle f, g\rangle_{\mathcal{H}} = \mathbb{E}\, I(f) I(g) \qquad \text{for } f, g \in \mathcal{E}. \tag{17}$$

We can require, without the loss of generality, by adding additional cutpoints t_i if necessary, that

$$f(t) = \sum_{i=0}^{n-1} a_i \mathbb{1}_{[t_i, t_{i+1})}(t), \quad g(t) = \sum_{i=0}^{n-1} b_i \mathbb{1}_{[t_i, t_{i+1})}(t) \qquad \text{for all } t \geq 0,$$

where $n \in \mathbb{N}$, $0 = t_0 < t_1 < \ldots < t_n < \infty$, $a_i, b_i \in \mathbb{R}$, $i = 0, \ldots, n-1$. We can write

$$\begin{aligned}\mathbb{E}\, I(f) I(g) &= \sum_{i=0}^{n-1} \sum_{j=0}^{n-1} a_i b_j [R(t_{i+1}, t_{j+1}) - R(t_i, t_{j+1}) - R(t_{i+1}, t_j) + R(t_i, t_j)] \\ &= \int_0^\infty \int_0^\infty f(u) g(v) R(du, dv) = \alpha_H \int_0^\infty \int_0^\infty f(u) g(v) |u-v|^{2H-2}\, du\, dv. \end{aligned} \tag{18}$$

The reader may want to note that in (18) the integrals are just another way of expressing a linear combination of the terms $R(u,v)$ for various u and v.

For a function $\varphi : \mathbb{R}_+ \to \mathbb{R}$ define

$$|\varphi|^2_{|\mathcal{H}|} := \alpha_H \int_0^\infty \int_0^\infty |\varphi(r)||\varphi(s)||r-s|^{2H-2}\, dr\, ds.$$

The last term of (18) suggests that we may be able to integrate with respect to B the functions from the space $|\mathcal{H}| := \{\varphi : \mathbb{R}_+ \to \mathbb{R};\ |\varphi|_{|\mathcal{H}|} < \infty\}$.

It is an immediate consequence of p. 17 in Pipiras & Taqqu (2000), that the space $|\mathcal{H}|$ is a Banach space with the norm $|\cdot|_{|\mathcal{H}|}$, in which the set of simple functions $\mathcal{E}$ is dense, by Theorem 3.2 ibid. on p. 9.

The space $|\mathcal{H}|$ is, however, not complete with respect to the norm $|\cdot|_{\mathcal{H}} := \sqrt{\langle \cdot, \cdot \rangle_{\mathcal{H}}}$, ibid. From Proposition 27 it follows that there exists an isometry from the space $|\mathcal{H}|$ into a proper subspace of $\mathcal{B}$.

Let $\mathcal{H}$ denote the completion of the space of simple functions $\mathcal{E}$ in the scalar product $\langle \cdot, \cdot \rangle_{\mathcal{H}}$. From Proposition 27 it follows that $\mathcal{H}$ is isometric to $\mathcal{B}$ itself.

Definition 21 (Stochastic integral)**.** Let $\varphi \in \mathcal{H}$. Let $I \colon \mathcal{H} \to \mathcal{B}$ be the isometry warranted by Proposition 27 using the scalar product (17). A stochastic integral of φ with respect to the fractional Brownian motion is defined to be the random variable $I(\varphi)$ and is denoted as $\int \varphi dB$. We introduce the standard notation $\int_s^t \varphi(s) dB(s) = \int \mathbb{1}_{[s,t]} \varphi dB$ for $0 \le s \le t \le \infty$.

It is trivial to see that $|\mathcal{H}| \subseteq \mathcal{H}$. Hence the isometry I also acts on the space of functions $|\mathcal{H}|$.

Proposition 28. *Let $\varphi, \psi \in \mathcal{H}$. Then*

$$\mathbb{E}\left(\int \varphi\, dB\right)\left(\int \psi\, dB\right) = \langle \varphi, \psi \rangle_{\mathcal{H}}.$$

Proof. The proof follows from the fact that the stochastic integral is induced by the isometry of $\mathcal{H}$ and $\mathcal{B}$. □

Remark 29. From the proof of Theorem 1.2 in Mémin et al. (2001) on p. 201, or also Mishura (2008), Lemma 1.6.6, p. 20, we have that for every function $\varphi \in |\mathcal{H}|$ there exists a constant b_H such that the estimate

$$|\varphi|_{|\mathcal{H}|} \le b_H |\varphi|_{\frac{1}{H}}$$

holds. This estimate implies the inclusion of $L^{1/H} \subseteq |\mathcal{H}|$. By trivial properties of the Lebesgue integral, it also holds $|\varphi|_{\mathcal{H}} \le |\varphi|_{|\mathcal{H}|}$ Hence we can write by Proposition 28

$$\mathbb{E}\left(\int_0^\infty \varphi\, dB\right)^2 = |\varphi|^2_{\mathcal{H}} \le b_H^2 |\varphi|^2_{\frac{1}{H}}$$

for some positive constant b_H independent of φ for any function $\varphi \in |\mathcal{H}|$.

Let us now analyze the inner product $\langle \cdot, \cdot \rangle_{\mathcal{H}}$ in greater detail. Once done, we will be able to calculate $\langle \cdot, \cdot \rangle_{\mathcal{H}}$ for all functions in $|\mathcal{H}|$ which will prove useful. We proceed in accord with p. 3 ff. of Alòs & Nualart (2003), we only have $T = \infty$

and replace $\mathcal{E}$ with $|\mathcal{H}|$. This makes no harm since all the integrals still converge. First, we have

$$|r-u|^{2H-2} = \frac{(ru)^{H-\frac{1}{2}}}{\beta\left(2-2H, H-\frac{1}{2}\right)} \int_0^u v^{1-2H}(r-v)^{H-\frac{3}{2}}(u-v)^{H-\frac{3}{2}}\,dv, \quad \text{for all } r,u \geq 0. \quad (19)$$

where β denotes the beta function. Suppose $r > u > 0$. The formula (19) can be obtained by the change of variables $z = (r-v)/(u-v)$ and $x = r/uz$ we obtain

$$\begin{aligned}\int_0^u v^{1-2H}(r-v)^{H-\frac{3}{2}}(u-v)^{H-\frac{3}{2}}dv &= (r-v)^{2H-2}\int_{\frac{r}{u}}^{\infty}(zu-r)^{1-2H}z^{H-\frac{3}{2}}dz \\ &= (ru)^{\frac{1}{2}-H}(r-u)^{2H-2}\int_0^1(1-x)^{1-2H}x^{H-\frac{3}{2}}dx \\ &= \beta\left(2-2H, H-\frac{1}{2}\right)(ru)^{\frac{1}{2}-H}(r-u)^{2H-2}.\end{aligned}$$

Consider the square integrable kernel

$$\begin{aligned}K_H(t,s) &= c_H s^{\frac{1}{2}-H}\int_s^t (u-s)^{H-\frac{3}{2}}u^{H-\frac{1}{2}}\,du, \quad \text{where } t > s \geq 0, \text{ and} \\ c_H &= \left[\frac{H(2H-1)}{\beta\left(2-2H, H-\frac{1}{2}\right)}\right]^{\frac{1}{2}}.\end{aligned}$$

Notice that

$$\dot{K}_H(t,s) = c_H\left(\frac{t}{s}\right)^{H-\frac{1}{2}}(t-s)^{H-\frac{3}{2}} \geq 0 \quad \text{for all } t > s \geq 0. \quad (20)$$

where $\dot{K}_H(r,s)$ is the partial derivative of $K_H(r,s)$ in the first variable r, or put differently $\dot{K}_H(r,s) := \frac{\partial}{\partial r}K_H(r,s)$. We may verify that for all $t,s \geq 0$

$$\begin{aligned}&\int_0^{t\wedge s} K_H(t,u)K_H(s,u)du \\ &= c_H^2\int_0^{t\wedge s}\left(\int_u^t (y-u)^{H-\frac{3}{2}}y^{H-\frac{1}{2}}dy\right)\left(\int_u^s (z-u)^{H-\frac{3}{2}}z^{H-\frac{1}{2}}dz\right)u^{1-2H}du \\ &= c_H^2\beta\left(2-2H, H-\frac{1}{2}\right)\int_0^t\int_0^s |y-z|^{2H-2}dzdy = R(t,s). \quad (21)\end{aligned}$$

From (21) we see that $R(t,s)$ is nonnegative definite. Consider now, the linear operator K_H^* mapping $\mathcal{E}$ into $L^2(\mathbb{R}_+)$ defined as

$$K_H^* f(s) := \int_s^{\infty} f(r)\dot{K}_H(r,s)\,dr, \quad s \geq 0, \quad (22)$$

for a simple function $f \in \mathcal{E}$. Using (21) and (18) we may write for simple functions f and g from $\mathcal{E}$

$$
\begin{aligned}
&\langle K_H^* f, K_H^* g\rangle_{L^2(\mathbb{R}_+)} \\
&\qquad = \int_0^\infty \left(\int_s^\infty f(u)\dot{K}_H(u,s)\,du\right)\left(\int_s^\infty g(v)\dot{K}_H(v,s)\,dv\right) ds \\
&\qquad = \int_0^\infty \int_0^\infty \left(\int_0^{u\wedge v} \dot{K}_H(u,s)\dot{K}_H(v,s)\,ds\right) f(u)g(v)\,du\,dv \\
&= \int_0^\infty \int_0^\infty \frac{\partial^2}{\partial u \partial v} R(u,v) f(u)g(v)\,du\,dv = \alpha_H \int_0^\infty \int_0^\infty |u-v|^{2H-2} f(u)g(v)\,du\,dv \\
&\qquad = \langle f, g\rangle_{\mathcal{H}}. \quad (23)
\end{aligned}
$$

From the isometry relation (23) and the fact that $\mathcal{E}$ is dense in $\mathcal{H}$ it is apparent that the operator K_H^* can be extended from $\mathcal{E}$ to $\mathcal{H}$, cf. Alòs & Nualart (2003); in the rest of the paragraph we complete their terse note on the bottom of p. 8. Let $f_n \in \mathcal{E}$ be simple so that they converge to $\varphi \in \mathcal{H}$ in $|\cdot|_{\mathcal{H}}$. By the isometry relation and completeness of $L^2(\mathbb{R}_+)$ there is an element $g \in L^2(\mathbb{R}_+)$ such that $K_H^* f_n \to g$ in $L^2(\mathbb{R}_+)$. We define $K_H^*\varphi := g$. Note that this way, $K_H^*\varphi$ is defined only a.e. on $\mathbb{R}_+$ with respect to the Lebesgue measure. It is also clear that $\{f\colon \mathbb{R}_+ \to \mathbb{R},\ \text{measurable};\ K_H^* f \in L^2(\mathbb{R}_+)\} \subseteq \mathcal{H}$ and that the set $\{K_H^* f;\ f \in \mathcal{H}\}$ is dense in $L^2(\mathbb{R}_+)$.

Let $\varphi \in |\mathcal{H}|$ and define

$$\tilde{K}_H^*\varphi(s) := \int_s^\infty \varphi(r)\dot{K}_H(r,s)\,dr, \quad s \geq 0.$$

Then the exchange of the order of integration in (23) is possible by the Fubini theorem and the fact that $\varphi \in |\mathcal{H}|$, and we obtain by the exactly same calculations as in (23) that the representation

$$\langle \tilde{K}_H^*\varphi, \tilde{K}_H^*\xi\rangle_{L^2(\mathbb{R}_+)} = \langle \varphi, \xi\rangle_{\mathcal{H}} \tag{24}$$

holds for all $\varphi, \xi \in |\mathcal{H}|$. Since $|\tilde{K}_H^*\varphi|^2_{L^2(\mathbb{R}_+)} = |\varphi|^2_{\mathcal{H}} \leq |\varphi|^2_{|\mathcal{H}|} < \infty$ we also have that $\tilde{K}_H^*\varphi(s) < \infty$ for a.a. $s \in \mathbb{R}_+$. Let now $f_n \in \mathcal{E}$ be simple functions converging to $\varphi \in |\mathcal{H}|$ in $|\cdot|_{\mathcal{H}}$. Then $K_H^* f_n = \tilde{K}_H^* f_n$ for all $n \in \mathbb{N}$ and hence by (24) we can write

$$
\begin{aligned}
&\left|K_H^*\varphi - \tilde{K}_H^*\varphi\right|_{L^2(\mathbb{R}_+)} \\
&\quad \leq |K_H^*\varphi - K_H^* f_n|_{L^2(\mathbb{R}_+)} + \left|K_H^* f_n - \tilde{K}_H^* f_n\right|_{L^2(\mathbb{R}_+)} + \left|\tilde{K}_H^* f_n - \tilde{K}_H^*\varphi\right|_{L^2(\mathbb{R}_+)} \\
&\quad \leq 2|f_n - \varphi|_{\mathcal{H}} \to 0 \quad \text{as } n \to \infty \quad (25)
\end{aligned}
$$

and hence

$$K_H^*\varphi(t) = \tilde{K}_H^*\varphi(t) \quad \text{for a.a. } t \geq 0 \text{ in the Lebesgue sense.} \tag{26}$$

The operator K_H^* also gives an important representation of B in terms of a particular Wiener process. The process

$$N(t) := \int_0^\infty (K_H^*)^{-1}(\mathbb{1}_{[0,t]})\,dB \tag{27}$$

is a Wiener process, cf. Alòs & Nualart (2003), (8) on p. 5. Trivially $N(0) = B(0) = 0$. The covariance function can be obtained by direct calculation

$$
\begin{aligned}
\mathbb{E}N(t)N(s) &= \mathbb{E}\int_0^\infty (K_H^*)^{-1}(\mathbb{1}_{[0,t]})\,dB \int_0^\infty (K_H^*)^{-1}(\mathbb{1}_{[0,s]})\,dB \\
&= \big\langle (K_H^*)^{-1}(\mathbb{1}_{[0,t]}), (K_H^*)^{-1}(\mathbb{1}_{[0,s]})\big\rangle_{\mathcal{H}} = \big\langle \mathbb{1}_{[0,t]}, \mathbb{1}_{[0,s]}\big\rangle_{L^2(\mathbb{R}_+)} \\
&= t \wedge s \qquad \text{for all } t, s \geq 0. \qquad (28)
\end{aligned}
$$

We will call this Wiener process the Wiener process associated with the fractional Brownian motion B. What is more, B has an integral representation in terms of N

$$
B(t) = \int_0^t K_H(t,s)\,dN(s), \quad t \geq 0 \text{ a.s.} \tag{29}
$$

This is easily obtained for we have $K_H^*\mathbb{1}_{[0,t]}(s) = K_H(t,s)$ for all $t \geq s \geq 0$. From (23) and comments following we obtain that for every function $f \in \mathcal{H}$ we have

$$
\int_0^\infty f(t)\,dB(t) = \int_0^\infty K_H^* f(t)\,dN(t) \quad \text{a.s.} \tag{30}
$$

From (26) for every $f \in |\mathcal{H}|$ we have

$$
\int_0^\infty f(t)\,dB(t) = \int_0^\infty \int_t^\infty f(r)\dot{K}_H(r,t)\,dr\,dN(t) \quad \text{a.s.} \tag{31}
$$

Working in multiple dimensions, we will use the following definition.

Definition 22. Let B be an n-dimensional fractional Brownian motion with Hurst parameter $H > 1/2$ in the sense of Definition 20 and $M\colon \mathbb{R}_+ \to \mathbb{R}^{m\times n}$ be a deterministic matrix-valued map so that all its components M_{ij} are in $\mathcal{H}$. We define

$$
\int_0^\infty M(t)\,dB(t) := \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \text{ where } x_i := \sum_{k=1}^n \int_0^\infty M_{ik}(t)\,dB_k(t), \; i = 1, \dots, m.
$$

The following theorem is used in the proof of the variations of constants formula for linear stochastic differential equations. It can be found in Mishura (2008) as Theorem 1.13.1 on p. 57. We present a slight generalization for the case of a matrix valued map and an n-dimensional fractional Brownian motion. It can be readily proved using the one-dimensional theorem by rewriting the matrix products as sums and realizing that $a_{ij} = |e_i^* a e_j| \leq \|a\|$ for all $i, j = 1, \dots, m$ if $a \in \mathbb{R}^{m\times m}$, for a fixed natural number m.

Theorem 30 (Stochastic Fubini's theorem)**.** *Fix $T > 0$, natural numbers n and m, and an n-dimensional fractional Brownian motion B with Hurst parameter $H > 1/2$. Let $f\colon [0,T]^2 \to \mathbb{R}^{m\times n}$ be a measurable function such that*

$$
\int_0^T \int_0^T \int_0^T \|f(t,u)\| \|f(t,s)\| |s-u|^{2H-1}\,ds\,du\,dt < \infty
$$

and

$$\int_0^T \int_0^T \int_0^T \int_0^T \|f(t_1,u)\|\|f(t_2,s)\||s-u|^{2H-1}\,ds\,du\,dt_1dt_2 < \infty.$$

Then the following integrals exist and we have

$$\int_0^T \int_0^T f(t,s)\,dB(s)\,dt = \int_0^T \int_0^T f(t,s)\,dt\,dB(s) \qquad a.s.$$

Remark 31. At the end of the chapter, let us note that from the representation (30) we obtain that $\mathbb{E}\int \varphi dB = 0$ for all $\varphi \in \mathcal{H}$ and that the stochastic integral $\int \cdot\, dB$ is linear by the standard properties of the Wiener integral.

1.9 Stochastic differential equations with fractional noise

Fix natural numbers m and n and a stochastic basis $(\Omega, \mathcal{F}, \mathcal{F}(t)_{t\geq 0}, \mathbb{P})$ satisfying the usual conditions. Let A be an $\mathcal{F}(t)$-progressive continuous process with values in $\mathbb{R}^{m\times m}$, b be an $\mathcal{F}(t)$-progressive process with trajectories in $L^1_{loc}(\mathbb{R}_+;\mathbb{R}^m)$ a.s. and $\sigma \in \mathbb{R}^{m\times n}$ be a constant deterministic matrix. Let B be an n-dimensional fractional Brownian motion with Hurst parameter $H > 1/2$ adapted to the filtration $\mathcal{F}(t)_{t\geq 0}$ and X_0 an $\mathcal{F}(0)$-measurable random variable with values $\mathbb{R}^m$ representing the initial condition. Consider the equation

$$dX(t) = [A(t)X(t) + b(t)]\,dt + \sigma\,dB(t), \quad t > 0, \tag{32}$$

$$X(0) = X_0. \tag{33}$$

Definition 23. Let $(X(t),\ t \geq 0)$ be an $\mathcal{F}(t)$-progressive stochastic process defined on the stochastic basis $(\Omega, \mathcal{F}, \mathcal{F}(t)_{t>0}, \mathbb{P})$ and let $I \subseteq \mathbb{R}_+$ be an open interval. The process $(X(t),\ t \in I)$ is said to be a strong solution of the linear stochastic differential equation driven by fraction Brownian motion (32) with initial condition (33) on interval I if

$$\int_0^t |A(s)X(s)|\,ds < \infty \qquad \text{for all } t \in I \ \text{ a.s.} \tag{34}$$

and it verifies

$$X(t) = X_0 + \int_0^t [A(s)X(s) + b(s)]\,ds + \int_0^t \sigma\,dB(s) \qquad \text{for all } t \in I \ \text{ a.s.} \tag{35}$$

We will omit the word strong in the sequel and only call X the solution of (32), (33) as we will not deal with other notions of solutions in this thesis. We will omit the interval I if talking about solutions when the interval is clear from context or is the whole $\mathbb{R}_+$.

If the solution to the linear equation exists then it is continuous and unique. This is what we prove next.

Theorem 32. *Let a solution of the equation* (32) *with the initial condition* (33) *exist. Then it is continuous and unique in the following sense. Let X and Y be solutions on some intervals I_1 and I_2 of* (32) *with the initial condition* (33). *Then they are continuous and* $\mathbb{P}\{X(t) = Y(t),\ \text{for all } t \in I_1 \cap I_2\} = 1$.

Proof. Since X is a solution by Definition 23 it satisfies

$$X(t) = X_0 + \int_0^t [A(s)X(s) + b(s)]\, ds + \sigma B(t) \quad \text{for all } t \in I \;\; \mathbb{P}\text{-a.s.}$$

Recall that all trajectories of $B(t)$ are continuous by definition. The integral $\int_0^t A(s)X(s) + b(s)\, ds$ is defined pathwise in the Lebesgue sense and its value is finite on compact intervals by definition of the solution and assumptions on b and A. Hence it is continuous in t a.s. This makes all trajectories of the solution X continuous.

For all $\omega \in \Omega'$ we have

$$|X(t,\omega) - Y(t,\omega)| \leq \int_0^t \|A(s,\omega)\| |X(s,\omega) - Y(s,\omega)|\, ds$$

using Grönwalls lemma in Theorem 5, we obtain $X(t,\omega) = Y(t,\omega)$ for all $t \in I'$ for all $\omega \in \Omega'$ for all $I' \subseteq I_1 \cap I_2$ compact intervals containing 0. $\square$

Our equation (32) is rather simple. The existence and uniqueness of solutions would be warranted by Theorem 1.12.1 in Mishura (2008) on p. 55 had we not have a random coefficient matrix A. In the next theorem, we prove the existence of solutions of stochastic differential equation (32) for the case of random coefficients if the random matrix $A(t)$ is bounded uniformly a.s. for all $t \geq 0$. The proof is basically the one in Mishura (2008) referenced above.

Theorem 33. *Fix $p \geq 1$. Let $C_p := \operatorname{esssup}_{\omega\in\Omega} \sup_{s\geq 0} \|A(s,\omega)\|^p$. Require $C_p < \infty$, $\mathbb{E}|X_0|^p < \infty$ and $\mathbb{E}\int_0^T |b(t)|^p dt < \infty$ for all $T \geq 0$. Then the equation* (32) *with the initial condition* (33) *has a unique global solution.*

Moreover, for every $T > 0$ we have for all $0 \leq t \leq T$.

$$\mathbb{E} \sup_{0\leq s\leq t} |X(s)|^p \leq 4^{p-1} e^{4^{p-1} C_p T^{p-1}}\, \mathbb{E}\left(|X_0|^p + \left| \int_0^T b(s)\, ds \right|^p + \sigma \sup_{0\leq s\leq T} |B(s)|^p \right).$$

Proof. Choose $T < C_p^{-1}$. Define the successive approximation operator

$$\mathcal{L}X(t) := X_0 + \int_0^t A(s)X(s) + b(s)\, ds + \sigma B(t) \quad \text{for all } 0 \leq t \leq T.$$

We will show that the operator is a contraction in the space of random processes

$$S_p := \left\{ (\xi(t),\ T \geq t \geq 0);\ \sup_{0\leq t\leq 1} \mathbb{E}|\xi(t)|^p < \infty \right\}.$$

with the norm

$$|\xi|_{S_p} := \sup_{0\leq t\leq 1} \left(\mathbb{E}|\xi(t)|^p\right)^{\frac{1}{p}}.$$

Calculate using Remark 3 and the Hölder inequality

$$\mathbb{E}|\mathcal{L}X(t)|^p \leq 3^p \Bigg[\mathbb{E}|X_0|^p + \mathbb{E}|\sigma B(t)|^p + 2^p t^{p-1} C_p \left(\int_0^t \mathbb{E}|X(s)|^p\, ds + \int_0^t \mathbb{E}|b(s)|^p\, ds \right) \Bigg] \quad \text{for all } 0 \leq t \leq T.$$

This means that $\mathcal{L}X \in S_p$ if $X \in S_p$. Further, for $0 \le t \le T$ we have

$$\begin{aligned}\mathbb{E}|\mathcal{L}X(t) - \mathcal{L}Y(t)|^p &\le \mathbb{E}\left(\int_0^t \|A(s)\||X(s) - Y(s)|\,ds\right)^p \\ &\le C_p\,\mathbb{E}\left(\int_0^t |X(s) - Y(s)|\,ds\right)^p \le t^{p-1}C_p\int_0^t \mathbb{E}|X(s) - Y(s)|^p\,ds.\end{aligned}$$

Hence

$$|\mathcal{L}X - \mathcal{L}Y|_{S_p} \le T^{p-1}C_p\,|X - Y|_{S_p}.$$

Let us now prove that the space S_p is complete. Let $X_n \in S_p$, $n \in \mathbb{N}$ be a Cauchy sequence in S_p. Then $X_n(t)$ is Cauchy in $L^p(\Omega)$ for every $0 \le t \le T$. Since it is well known that the space $L^p(\Omega)$ is complete, for the case of $p = 2$ cf. Theorem 11.42 in Rudin et al. (1964), p. 329, we can choose $X(t)$ for every $0 \le t \le T$ to be the limit of $X_n(t)$ in $L^p(\Omega)$. Since the convergence was uniform in t we have that $X_n \to X$ in S_p and we obtained that the space S_p is complete.

Observe that $T^{p-1}C_p(T) < 1$ by construction. By the fixed-point theorem can be found for example in Rudin et al. (1964) as Theorem 9.23 on p. 220, there exists a unique process $X \in S_p$ so that $\mathcal{L}X(t) = X(t)$ for all $0 \le t \le 1$ which means it solves the equation (32) on $[0,1]$. By construction it also holds $X(0) = x$ and hence the initial condition (33) is also met by X.

The solution can be extended to the whole real line by repeating the construction for every $k \in \mathbb{N}$ using the extending equation

$$X(t) = X(kT) + \int_{kT}^t A(s)X(s) + b(s)\,ds + \sigma[B(t) - B(kT)]$$

for $kT \le t \le (k+1)T$.

We proved that there is a global solution X of the equation (32). By Theorem 32, the solution X is continuous. Denote $X^\star(t) := \sup_{0\le s\le t} X(s)$ for all $t \ge 0$. Fix $T \ge 0$. Almost every trajectory of the solution is continuous, hence $X^\star(t)$ is continuous a.s. and satisfies for all $0 \le t \le T$

$$\begin{aligned}|X^\star(t)|^p \le 4^{p-1}\Bigg(|X_0|^p &+ \left|\int_0^T b(s)\,ds\right|^p \\ &+ \sigma \sup_{0\le s\le T}|B(s)|^p + C_p\,T^{p-1}\int_0^t |X^\star(s)|^p\,ds\Bigg).\end{aligned}$$

Using the Grönwall lemma in Theorem 5 we obtain for all $0 \le t \le T$

$$|X^\star(t)|^p \le 4^{p-1}e^{4^{p-1}C_p\,T^{p-1}}\left(|X_0|^p + \left|\int_0^T b(s)\,ds\right|^p + \sigma\sup_{0\le s\le T}|B(s)|^p\right).$$

Uniqueness follows from Theorem 32. This concludes the proof. □

Proposition 34. *Let a solution* $(X(t),\ t \ge 0)$ *of the equation* (32) *with the initial condition* (33) *exist. Then the function* $\mu(t) := \mathbb{E}X(t)$ *for* $t \ge 0$ *is finite for all* $t \ge 0$*. Moreover, if* A *is deterministic,* μ *satisfies an ordinary differential equation*

$$\dot\mu(t) = A(t)\mu(t) + \mathbb{E}b(t) \quad \text{for } t \ge 0$$

with the initial condition

$$\mu(0) = \mathbb{E}X_0.$$

and μ is thus continuous.

Proof. The finiteness follows from Theorem 33.

Let A be deterministic. Taking expectation of both sides of the equation (35), realizing that $\mathbb{E}B(t) = 0$ for all $t \geq 0$ and that thanks to the assumptions put on the solution in Definition 23 in (34) we can use the Fubini theorem and exchange integral and expectation we obtain

$$\mu(t) = \mathbb{E}X_0 + \int_0^t A(s)\mu(s) + \mathbb{E}b(s)\, ds \quad t \geq 0$$

which once differentiated concludes the proof. □

The following proposition provides means for writing the solution of the equation (32) with initial condition (33) in a rather explicit form. If the fundamental matrix solution is known, the form becomes an explicit stochastic integral with respect to the fractional Brownian motion B. The formula has the form of a "variation of constants" formula which holds for linear ordinary differential equations. We constrain ourselves to deterministic matrices A. This is mainly because of the fact that we did not develop theory of integration of stochastic processes with respect to fractional Brownian motion.

Proposition 35. *Let A be deterministic. Let S be the flow map of A in the spirit of Definition 8. Then the process $(X(t),\ t \geq 0)$ defined as*

$$X(t) = S(t,0)x + \int_0^t S(t,r)b(r)\, dr + \int_0^t S(t,r)\sigma\, dB(r), \quad t \geq 0 \tag{36}$$

is the solution of the linear stochastic differential equation (32) *with the initial condition* (33) *on $\mathbb{R}_+$.*

Proof. We take the solution defined in (36) and substitute in into the right hand side of equation (32). If we succeed to show that the right hand side equals $X(t)$ a.s. we will have proved the theorem. After the substitution, we obtain

$$\begin{aligned}
x + \int_0^t A(s) \left(S(s,0)x + \int_0^s S(s,r)b(r)\, dr + \int_0^s S(s,r)\sigma\, dB(r) \right) ds& \\
+ \int_0^t b(s)\, ds + \int_0^t \sigma\, dB(s)& \\
= x + \int_0^t A(s)S(s,0)x\, ds + \int_0^t \int_0^s A(s)S(s,r)b(r)\, dr\, ds& \\
+ \int_0^t \int_0^s A(s)S(s,r)\sigma\, dB(r)\, ds + \int_0^t b(s)\, ds + \int_0^t \sigma\, dB(s)&
\end{aligned}$$

for all $t \geq 0$. To use the stochastic the Fubini Theorem 30 we have to check the two conditions. Clearly for all $t \geq 0$

$$\int_0^t \int_0^t \int_0^t \|A(t)\|^2 \|\sigma\|^2 \|S(t,s)\| \|S(t,r)\| |s-r|^{2H-1}\, ds\, dr\, dt < \infty$$

and

$$\int_0^t \int_0^t \int_0^t \int_0^t \|A(t_1)\| \|A(t_2)\| \|\sigma\|^2 \|S(t_1, s)\| \|S(t_2, r)\| |s - r|^{2H-1} \, ds \, dr \, dt_1 dt_2 < \infty$$

since the integrands are bounded, because they are continuous and the integral is over a finite interval. Note that the flow map S is continuous in both parameters by Lemma 9 (iv).

Using the stochastic the Fubini Theorem 30, the ordinary Fubini theorem and the fundamental property of the semigroup from p. 6, namely that $S(\cdot, s)$ solves the ordinary differential equation $\dot{x}(t) = A(t)x(t),\ x(s) = \mathrm{Id}_{m\times m}$ for all $t > s \geq 0$, we may continue

$$\begin{aligned}
&= x + \int_0^t \dot{S}(s,0)x \, ds + \int_0^t \int_r^t \dot{S}(s,r)b(r) \, ds \, dr \\
&\qquad + \int_0^t \int_r^t \dot{S}(s,r)\sigma \, ds \, dB(r) + \int_0^t b(s) \, ds + \int_0^t \sigma \, dB(s) \\
&= x + (S(t,0) - S(0,0))x \, ds + \int_0^t (S(t,r) - S(r,r))\sigma \, dB(r) \\
&\qquad + \int_0^t (S(t,r) - S(r,r))b(r) \, dr + \int_0^t b(s) \, ds + \int_0^t \sigma \, dB(s) \\
&= S(t,0)x + \int_0^t S(t,r)b(r) \, dr + \int_0^t S(t,r)\sigma \, dB(r)
\end{aligned}$$

for all $t \geq 0$ a.s. where the formula after the last equality sign matches X as defined in (36). We proved that X solves the equation (32), (33) which concludes the proof. □

Proposition 36. *If A is a constant $(m \times m)$-matrix, the flow map of A is known to be expressible in terms of the matrix exponential as $S(t,s) = \exp\{A(t-s)\}$ for $t \geq s \geq 0$. The formula (36) simplifies to*

$$X(t) = \exp\{At\}x + \int_0^t \exp\{A(t-r)\}\sigma \, dB(r), \quad t \geq 0. \tag{37}$$

Proof. The proof follows directly from Propositions 14 and 35. □

2 Problem formulation

In this section, we first define an ergodic control problem with fractional noise, then extend it to a parameter-dependent system. Finally, we formulate the main problem of the thesis.

We will first fix the general settings for what follows.

Fix a stochastic basis $(\Omega, \mathcal{F}, \mathcal{K}(t)_{t\geq 0}, \mathbb{P})$ satisfying the usual conditions and such that there are two filtrations of $\mathcal{F}$, $\mathcal{F}(t)$ and $\mathcal{G}(t)_{t\geq 0}$ so that $\mathcal{K}(t) := \mathcal{F}(t) \vee \mathcal{G}(t)$ for all $t \geq 0$. Require that $\mathcal{F}(t)$ and $\mathcal{G}(t)$ are independent for all $t \geq 0$.

Fix an $\mathcal{F}(t)$-progressive fractional Brownian motion B with Hurst parameter $H > 1/2$. We will consider the continuous version of B which exists by Section 1.7.

It is well known that an $\mathcal{F}(t)$-progressive process is also $[\mathcal{F}(t) \vee \mathcal{J}(t)]$-progressive for any filtration $\mathcal{J}(t)_{t\geq 0}$ of $\mathcal{F}$ such that $\mathcal{J}(t)$ and $\mathcal{F}(t)$ are independent for all $t \geq 0$. This is exactly our situation since $\mathcal{K}(t) = \mathcal{F}(t) \vee \mathcal{G}(t)$ and $\mathcal{F}(t)$ is independent of $\mathcal{G}(t)$ for all $t \geq 0$. Noting this, it is clear that B is also $\mathcal{K}(t)$-progressive.

2.1 Ergodic linear quadratic problem with fractional noise

We begin by giving a precise mathematical formulation of a simpler, non-adaptive, problem and comment on the meaning of the symbols that appear in the formulation. The formulation of the problem is analogous to a standard one for the case with ordinary Brownian motion as it is to be found e.g. in Chapter 6 in Yong & Zhou (1999) on p. 300.

Fix natural numbers k, l, m and n. Let $A \in \mathbb{R}^{m\times m}, G \in \mathbb{R}^{m\times k}$, $\sigma \in \mathbb{R}^{m\times n}$ be deterministic matrices and $x \in \mathbb{R}^m$ a deterministic initial condition. We denote $\mathcal{U}$ the set of all $\mathcal{K}(t)-$progressive processes $u : \mathbb{R}_+ \times \Omega \to \mathbb{R}^k$ which satisfy $\mathbb{E}\int_0^T |u(t)|^2 dt < \infty$ for all $T > 0$. The set $\mathcal{U}$ represents the set of admissible controls for the problem and a $u \in \mathcal{U}$ represents a control. The state of the system of interest is described by an m-dimensional random process $(X(t),\ t > 0)$. The system evolves according to the following equation

$$\begin{aligned} dX(t) &= [AX(t) + Gu(t)]\, dt + \sigma\, dB(t) \qquad \text{for } t > 0, \\ X(0) &= x \end{aligned} \tag{38}$$

for a fixed initial condition $x \in \mathbb{R}^m$.

It is not important that x is deterministic. It only frees us from some tedious considerations. If the cost functional below is updated accordingly, x can be random, but has to be $\mathcal{K}_0$-measurable so that the equation (38) makes sense and with bounded moments so that the following manipulations do not break down.

In order to be able to speak about optimality, we need some means of measuring it. This is why we define the cost functional $J(x, u)$. Let $Q \in \mathbb{R}^{m\times m}$ and $R \in \mathbb{R}^{k\times k}$ be symmetric positive semidefinite matrices. The matrices Q and R embody our preference for how the system should be controlled. For a process $u : \mathbb{R}_+ \times \Omega \to \mathbb{R}^k$ and a vector $x \in \mathbb{R}^m$ we define the cost functional to be

$$J(x, u) \quad = \quad \limsup_{T\to\infty} \frac{1}{T} J_T(y, u), \tag{39}$$

where the finite time cost is given by

$$J_T(x, u) \quad = \quad \frac{1}{2}\mathbb{E}\left[\int_0^T \langle QX(t), X(t)\rangle + \langle Ru(t), u(t)\rangle\, dt\right] \qquad \text{for } T \geq 0. \tag{40}$$

Definition 24. We say, that an admissible control $\bar{u} \in \mathcal{U}$ is an optimal control for the ergodic control problem (38) with cost (39) if

$$J(x, \bar{u}) = \inf_{u \in \mathcal{U}} J(x, u). \tag{41}$$

Definition 25. A control is said to be in feedback form, or to be a feedback control, if it can be written as

$$u(t) = \gamma(t) X(t) + V(t) \quad t \geq 0 \tag{42}$$

for some random processes γ with values in $\mathbb{R}^{k \times m}$ and V with values in $\mathbb{R}^k$, where X is a solution of the system equations (38) with u substituted from (42), i.e. a solution of

$$\begin{aligned} dX(t) &= \{[A + G\gamma(t)]X(t) + GV(t)\}\, dt + \sigma\, dB(t) \quad t > 0 \\ X(0) &= x \end{aligned}$$

Feedback controls incorporate true trajectory of the system and hence have a greater chance at being more robust to fluctuations in the system dynamics should the model not match reality completely. That is why we concentrate on finding optimal controls in feedback form.

To reiterate the most important aspects of the formulation we have given: At each time $t \geq 0$, the system has a state described by a random process $X = (X(t), t > 0)$ with dynamics described by a system of linear stochastic differential equations (38) driven by fractional Brownian motion. By setting the values of $u(t)$ one is provided by means of control he can exercise to make the system behave as he would like it to.

By defining the functional $J(x, u)$ one can set what is meant by "optimal". It is worth noting that the functional takes on real deterministic values and hence provides a kind of summarization of the effect of u on the system in all possible situations.

To solve the linear quadratic problem, we have to describe how to calculate the value of the control $u(t)$ at every time $t > 0$ in order to minimize $J(x, u)$. The control uses the state of the system X, but at every time $t > 0$, we can only use values of $X(s)$ for $s < t$, since we required adaptiveness of the control u.

2.2 Adaptive ergodic linear quadratic problem with fractional noise

In practice, the coefficients in equations are rarely known exactly. This is why it is most practical to have a result which describes how the optimal control looks like if we have to estimate the system equations form the behavior of the controlled system, and at the same time we want it controlled optimally.

In the setting of the ergodic linear quadratic problem, this means that there is some uncertainty in the parameters of the system equations (38). This is why in this section, we reformulate the problem to allow for variations in the equations describing its dynamics.

First, as in Section 2.1, fix natural numbers k, l, m, n and let $G \in \mathbb{R}^{m \times k}$, $\sigma \in \mathbb{R}^{m \times n}$. In addition fix a compact set $\Theta \subseteq \mathbb{R}^l$ to represent the possible values of an unknown parameter. The variations in the dynamics of the system will be allowed by replacing the matrix A by a continuous matrix-valued map defined on a compact set Θ, i.e. by a continuous map $A\colon \Theta \to \mathbb{R}^{m\times m}$. We let $\mathcal{U}$, as before, denote the set of all $\mathcal{K}(t)-$progressive processes $u : \mathbb{R}_+ \times \Omega \to \mathbb{R}^k$ which satisfy $\mathbb{E}\int_0^T |u(t)|^2 dt < \infty$ for all $T > 0$. The state of the parametrized system X evolves according to the following true equation

$$\begin{aligned} dX(t) &= [A(\theta_0)X(t) + Gu(t)]\,dt + \sigma\,dB(t) \quad \text{for } t > 0, \\ X(0) &= x, \end{aligned} \tag{43}$$

where $\theta_0 \in \Theta$ is an unknown parameter to be estimated jointly with control and $x \in \mathbb{R}^m$ is the fixed initial condition.

Let as before $Q \in \mathbb{R}^{m\times m}$ be a symmetric positive semidefinite matrix and $R \in \mathbb{R}^{n\times n}$ be a symmetric positive definite and hence regular matrix. In the case of a parametrized system, the cost functional remains the same, namely for the matrices Q and R defining our preference on how the system should be controlled, for a process $u : \mathbb{R}_+ \times \Omega \to \mathbb{R}^k$ and a vector $x \in \mathbb{R}^m$ we define the cost functional to be

$$J(x,u) \;=\; \limsup_{T\to\infty} \frac{1}{T} J_T(x,u), \tag{44}$$

where the finite-time cost is given by

$$J_T(x,u) \;=\; \frac{1}{2} E\left\{ \int_0^T \langle QX(t), X(t)\rangle + \langle Ru(t), u(t)\rangle\, dt \right\}. \tag{45}$$

Definition 26. Fix $\mathcal{V} \subseteq \mathcal{U}$. We say, that an admissible control $\bar{u} \in \mathcal{V}$ is an optimal adaptive control for the parametrized ergodic control problem (43) with cost (44) in class $\mathcal{V}$ if

$$J(x,\bar{u}) = \inf_{u\in\mathcal{V}} J(x,u)$$

and $\bar{u}(t)$ can be calculated without the knowledge of the true parameter θ_0 for all $t \geq 0$.

The adaptive control problem consists of finding an adaptive optimal feedback control for the controlled system (38) in the sense of Definition 26 and 25.

2.3 The parameter estimator

The evolution of system (43) depends on an unknown parameter $\theta_0 \in \Theta$. In order to be able to derive an optimal control for the system, we require some estimator for the parameter to be available.

Assume we have access to an estimator $\hat{\theta}$ that is $\mathcal{K}(t)$-progressive, continuous, takes on values only from the set of feasible parameters Θ and is a strongly consistent estimator of the true parameter, i.e. a.s. $\hat{\theta}(t) \to \theta_0$ as $t \to \infty$.

We can now explain why we chose $\mathcal{K}(t) := \mathcal{F}(t) \vee \mathcal{G}(t)$ such that $\mathcal{F}(t)$ and $\mathcal{G}(t)$ independent for all $t \geq 0$. This way, we will be able to prove some theorems in full generality, while in others we may require that the estimator is independent of the driving fractional Brownian motion.

An independent parameter estimator would be realizable, if we had access to two identical but independent systems with equal unknown parameter θ_0. We would estimate θ_0 in one and use the same control to control both systems. Our main theorem then states, that we are able to optimally control one of the systems while using the other one to estimate the parameter.

Since the parameter space Θ is compact, by Theorem 4.15 on p. 89 in Rudin et al. (1964) the estimator process $(\hat{\theta}(t)^p,\ t > 0)$ is bounded for every $p \geq 1$ and hence by the dominated convergence theorem we have $\hat{\theta}(t) \to \theta_0$ in L_p as $t \to \infty$ for every $p \geq 1$.

2.4 The parametrized algebraic Riccati equation

The solution of the optimal control problem is formulated in a feedback form in terms of the instantaneous system state $X(t)$ and the optimal gain $\gamma(t)$. To recover the optimal gain for an ergodic control problem one has to know the solution of a special equation, the so called algebraic Riccati equation. The availability of the solution of the control then inherits properties of the solution to the algebraic Riccati equation.

The algebraic Riccati equation is defined in this section and the properties of solutions we will need are stated. The most notable result is continuous dependence of the solution of algebraic Riccati equation on parameters given suitable conditions are satsfied.

Let the settings of Section 2.2 be in force and let us have access to a strongly consistent estimator $\hat{\theta}(t)$ of the true parameter θ_0 with properties described in Section 2.3. We specify the following assumptions to ease further formulations, cf. Chojnowska-Michalik et al. (1992).

(A1) The pairs $(A(\theta), G)$ are stabilizable uniformly with respect to $\theta \in \Theta$.

(A2) The pairs $(Q, A(\theta))$ are detectable uniformly with respect to $\theta \in \Theta$.

Since the matrix-valued map A is assumed to be continuous, we have by Proposition 16

(A3) the function $\theta \mapsto \exp\{A(\theta)t\}$ is continuous for all $t \geq 0$.

Note that in Chojnowska-Michalik et al. (1992), (A3) is assumed, whereas in our finite-dimensional case we obtained it as a consequence of the continuity of the matrix-valued map A.

For a parameter $\theta \in \Theta$ consider the parametrized algebraic Riccati equation

$$0 = A(\theta)^* p + pA(\theta) + Q - pGR^{-1}G^* p \tag{46}$$

for an unknown symmetric matrix $p \in \mathbb{R}^{m \times m}$.

The next theorem states that there is a unique solution of (46) given suitable conditions are satisfied. It may be found Zabczyk (1975) as Theorem 1 on p. 252. We reformulate it for the finite-dimensional case.

Theorem 37. *Fix $\theta \in \Theta$ and let the pair $(A(\theta), G)$ be stabilizable and $(Q, A(\theta)$ be detectable. Then there exists a unique symmetric nonnegative solution $P(\theta)$ of (46). Moreover, the matrix $A(\theta) - GR^{-1}G^* P(\theta)$ is stable.*

In the sequel, let $P(\theta)$ denote the unique symmetric nonnegative solution of (46) and

$$A_P(\theta) := A(\theta) - GR^{-1}G^*P(\theta), \quad \theta \in \Theta \tag{47}$$

if the Riccati equation (46) is uniquely solvable for the respective $\theta \in \Theta$.

The next theorem states that the unique solution of (46) is continuous, cf. Theorem 1 in Chojnowska-Michalik et al. (1992) on p. 177. We reformulate it for the finite-dimensional case.

Theorem 38. *Let (A1) and (A2) be satisfied. Let $P(\theta)$, $\theta \in \Theta$ be the unique nonnegative solution of the Riccati equation* (46) *warranted by Theorem 37. Then P is continuous in the operator norm $\|\cdot\|$.*

The following proposition is a restatement of Lemma 2 in Chojnowska-Michalik et al. (1992) on p. 177.

Proposition 39. *Let (A1) and (A2) be satisfied. Then the matrix $A_P(\theta)$ is stable uniformly with respect to $\theta \in \Theta$.*

Recall from Section 2.3 that $\hat{\theta}$ is a strongly consistent continuous parameter estimator of $\theta_0 \in \Theta$ which among other things satisfies $\hat{\theta}(t) \in \Theta$ for all $t \geq 0$ and $\hat{\theta}(t) \to \theta_0$ a.s. We have the following corollary.

Corollary 40. *Let $f\colon \mathbb{R}^l \to \mathbb{R}$ be a continuous function.*

Then $\operatorname{esssup}_{\omega\in\Omega} \sup_{t\geq 0} f(\hat{\theta}(t,\omega)) < \infty$ *and* $\mathbb{E}|f(\hat{\theta}(t)) - f(\theta_0)| \to 0$ *as* $t \to \infty$.

Especially, if $P(\theta)$ is the solution of (46) *that is continuous in θ, we have $P(\hat{\theta}(t)) \to P(\theta_0)$ a.s. as $t \to \infty$. Moreover, for $p \geq 1$, we have that*

$$\operatorname*{esssup}_{\omega\in\Omega} \sup_{t\geq 0} \|P(\hat{\theta}(t,\omega))\|^p < \infty$$

and $\mathbb{E}\|P(\hat{\theta}(t)) - P(\theta_0)\|^p \to 0$ *as* $t \to \infty$.

Proof. Since f is continuous and $\hat{\theta}$ is strongly consistent, $f(\hat{\theta}(t)) \to f(\theta_0)$ a.s. as $t \to \infty$.

Recall now, that $\hat{\theta}(t) \in \Theta$ for all $t \geq 0$ and that Θ was assumed to be a compact subset of $\mathbb{R}^l$. By Theorem 4.15 on p. 89 in Rudin et al. (1964) we have that the function $\theta \mapsto f(\theta)$ is bounded. We can thus write

$$\sup_{t\geq 0} \left| f(\hat{\theta}(t)) - f(\theta_0) \right| \leq \sup_{\theta\in\Theta} |f(\theta) - f(\theta_0)| < \infty$$

hence the first statement of the corollary follows now from the dominated convergence theorem.

Since P is continuous, and from the just proved part, we obtain that $P(\hat{\theta}(t)) \to P(\theta_0)$ a.s. as $t \to \infty$.

In the same spirit, since the operator norm $M \mapsto \|M\|$ is continuous as well as the power function, we obtain the rest of the statement. $\square$

2.5 Solution of the ergodic linear quadratic optimal control problem

The form of optimal control of the system (43) with cost (44) for the case when the parameter $\theta_0 \in \Theta$ is known and fixed is well known even for the infinite-dimensional case. The solution was presented in a paper by Duncan et al. (2015). In this section, we will present their main theorem which is an indispensable building block of our main result.

We present the result of Duncan et al. (2015) in a slightly modified manner and consider a solution for a parametrized system. The only difference is, however, in notation, where we instead of A write $A(\theta_0)$. We will also skip the presentation of the original result for the infinite-dimensional case and present directly a simplified version which holds in the finite dimension. We can do this by invoking comments in Example 4.1 on p. 101 of Duncan et al. (2015) which justifies that the result of the main theorem hold also for the finite-dimensional case.

There are some discrepancies between the notation of Duncan et al. (2015) and ours which is the result of the fact, that in the infinite-dimensional case many considerations are more complicated and usually stronger assumptions are imposed. In order to transfer the results of Duncan et al. (2015) into our work we please the reader to acknowledge that a finite-dimensional linear operator is always trace class. This means, that in (3) in Duncan et al. (2015) we can safely set the covariance of the cylindrical fractional Brownian motion $\tilde{Q} := \mathrm{Id}_n$ which is why the theorem is applicable to system (43).

Let the settings of Section 2.2 be in force.

Duncan et al. (2015) treat a control problem (38) and define the cost functional to be $J(x,u) = \limsup_{T\to\infty} J_T(x,u)/T$ where

$$J_T(x,u) = \frac{1}{2}\mathbb{E}\left\{ \int_0^T |LX(t)|^2 + \langle Ru(t), u(t)\rangle \, dt \right\}. \tag{48}$$

For a positive semidefinite and symmetric matrix $Q \in \mathbb{R}^{m\times m}$ we may always write $Q = LL$ for some symmetric positive-definite matrix $L \in \mathbb{R}^{m\times m}$, cf. Olive (2017) the comment above Definition 11.11 on p. 317. The matrix L is then called the square root of Q. Using this fact, we see that (48) is equivalent to (45) when we put $Q = L^*L$. It is also trivial to see that the results which holds for (38) also hold for (43) when θ_0 is assumed to be known.

Impose the following conditions, cf. Duncan et al. (2015), conditions (A1) and (A2) on p. 94-95.

(D1) The pair $(Q, A(\theta_0))$ is detectable.

(D2) The pair $(A(\theta_0), G)$ is stabilizable.

The condition (D1) is equivalent to requiring that $(L, A(\theta_0))$ is detectable by Proposition 25 and Remark 26 since L and Q span the same linear subspace of $\mathbb{R}^{m\times m}$.

If the conditions (D1) and (D2) are satisfied, by Theorem 37 there exists a unique symmetric solution $P(\theta_0)$ of the Riccati equation (46) with $\theta := \theta_0$ and for the matrix $A_P(\theta_0)$ defined in (47) to be $A(\theta_0) - GR^{-1}G^*P(\theta_0)$ there exist positive constants $K_P(\theta_0)$ and $M_P(\theta_0)$ so that

$$\|\Phi(t;\theta_0)\| \leq M_P(\theta_0)e^{-K_P(\theta_0)t} \quad \text{for all } t \geq 0. \tag{49}$$

where $\Phi(t;\theta_0) := \exp\{A_P(\theta_0)t\}$ for $t \geq 0$.

Let

$$\Psi(t;\theta) := \exp\{A_P^*(\theta)t\} \quad \text{for } t \geq 0 \text{ and } \theta \in \Theta. \tag{50}$$

where $A_P(\theta)$ is defined in (47). Observe that $\Psi(t;\theta_0)$ is the adjoint of $\Phi(t;\theta_0)$ for all $t \geq 0$. This follows from the Definition 9 of the matrix exponential.

For $\theta \in \Theta$ we define the random process

$$W(t;\theta) := \mathbb{E}\left[\int_t^\infty \Psi(s-t;\theta)P(\theta)\sigma\, dB(s) \,\Big|\, \mathcal{F}(t)\right], \quad t \geq 0. \tag{51}$$

Note that the process represents an estimation of future evolution of the noise component in the optimally controlled system based on the observations until time t. Note also, that θ is constant in the variable we integrate over and that the integrand is deterministic since θ is.

As noted, the integrand in (51) is deterministic. On top of this, from (49) we can estimate for $i = 1, \ldots, m,\ j = 1, \ldots, n$

$$\begin{aligned}
&\int_t^\infty |(\Psi(s-t;\theta_0)P(\theta_0)\sigma)_{ij}|^{\frac{1}{H}}\, ds \leq \int_t^\infty \|\Psi(s-t;\theta_0)P(\theta_0)\sigma\|^{\frac{1}{H}}\, ds \\
&\leq \|P(\theta_0)\sigma\|^{\frac{1}{H}} \int_t^\infty \|\Psi(s-t;\theta_0)\|^{\frac{1}{H}}\, ds \leq \|P(\theta_0)\sigma\|^{\frac{1}{H}} M(\theta_0)^{\frac{1}{H}} \int_t^\infty e^{-\frac{K(\theta_0)}{H}(s-t)}\, ds \\
&\qquad \leq \|P(\theta_0)\sigma\|^{\frac{1}{H}} \frac{HM(\theta_0)^{\frac{1}{H}}}{K(\theta_0)} < \infty
\end{aligned}$$

and hence the integrand has all coordinates in $L^{\frac{1}{H}}(\mathbb{R}_+)$ which means that it is B-integrable by Remark 29 and the process $W(\cdot;\theta_0)$ is well defined.

Let $X^{x,u}$ denote the trajectory of the system with initial condition $x \in \mathbb{R}^m$ and under control $u \in \mathcal{U}$.

We restrict the class of controls in which we search for the optimal control by the following condition, cf. Duncan et al. (2015), condition (15) on p. 95.

$$\lim_{T\to\infty} \frac{1}{T}\mathbb{E}\langle P(\theta_0)X^{x,u}(T), X^{x,u}(T)\rangle = 0. \tag{52}$$

An explicit solution to the control problem (43), (44) in terms of the symmetric solution $P(\theta_0)$ of the Riccati equation (46) is given in the following theorem which can be found as Theorem 3.2 of Duncan et al. (2015) on p. 95 and reads:

Theorem 41. *Let (D1) and (D2) be satisfied, and let $u \in \mathcal{U}$ be a control satisfying (52). Then*

$$J_\infty \leq J(x,u). \tag{53}$$

where

$$\begin{aligned}
J_\infty := &- \left[\limsup_{T\to\infty} \frac{1}{2T}\mathbb{E}\int_0^T |R^{-1/2}G^*W(s;\theta_0)|^2\, ds\right] \\
&+ \alpha_H \int_0^\infty \mathrm{Tr}[P(\theta_0)\Phi(s;\theta_0)]\, |s|^{2H-2}\, ds.
\end{aligned}$$

Moreover, the feedback control

$$\bar{u}(t) = -R^{-1}G^*(P(\theta_0)X^{x,\bar{u}}(t) + W(t;\theta_0)) \tag{54}$$

is admissible, satisfies (52) and

$$J(x,\bar{u}) = J_\infty \tag{55}$$

hence it is a solution of, and J_∞ is the optimal cost for, the ergodic linear quadratic control problem (43) *with initial condition x in the class of admissible controls from $\mathcal{U}$ satisfying* (52).

Moreover, according to the proof of the main theorem in Duncan et al. (2015) on p. 9, the following propositions hold.

Proposition 42. *Fix $q > 0$ and assume (D1) and (D2). Then there exists a positive constant C_q so that*

$$\sup_{t\geq 0} \mathbb{E}|\bar{X}(t)|^{2q} \leq C_q. \tag{56}$$

Proposition 43. *Fix $q > 0$ and assume (D1) and (D2). Then there exists a positive constant D_q so that*

$$\sup_{t\geq 0} \mathbb{E}|W(t;\theta_0)|^{2q} \leq D_q.$$

2.6 Representation of W

In this section, we establish elementary representations of conditional expectations of stochastic integrals with respect to fractional Brownian motion. This will be one of the main building blocks in the construction of the adaptive ergodic control.

The following propositions are based on the work of Kleptsyna et al. (2005). Let $B(s\,|\,t) := \mathbb{E}[B(s)\,|\,\mathcal{F}(t)]$ for $s \geq t \geq 0$ which could be called the conditional fractional Brownian motion and let $N(t) = \int_0^\infty (K_H^*)^{-1}\mathbb{1}_{[0,t]}dB$, $t \geq 0$ be the Wiener process associated with B as in (27), this means it is an $\mathcal{F}(t)$-Wiener process.

Proposition 44. *We have*

$$B(s\,|\,t) = \int_0^t K_H(s,r)\,dN(r) \qquad \text{for all } s \geq t \geq 0. \tag{57}$$

Proof. Let us observe that by (29) we can write

$$\mathbb{E}\left[B(s)\,|\,\mathcal{F}(t)\right] = \mathbb{E}\left[\int_0^s K_H(s,r)\,dN(r)\,\Big|\,\mathcal{F}(t)\right] \quad \text{a.s.}$$

Realizing that N is a Wiener process and hence $(N(r) - N(t),\ r > t)$ and $\mathcal{F}(t)$ are independent for all $t \geq 0$, using the fact that the Wiener integral has zero expectation and the assumption $0 \leq t \leq s$ we may write

$$\mathbb{E}\left[\int_0^s K_H(s,r)\,dN(r)\,\Big|\,\mathcal{F}(t)\right] = \mathbb{E}\left[\int_0^t K_H(s,r)\,dN(r)\,\Big|\,\mathcal{F}(t)\right].$$

The integral on the right hand side is $\mathcal{F}(t)$-measurable and hence the conditional expectation can be removed and we obtain the statement of the lemma. □

Proposition 45. *Let $f \in |\mathcal{H}|$. Then for all $t \geq 0$*

$$\mathbb{E}\left[\int_t^\infty f(s)\,dB(s)\,\Big|\,\mathcal{F}(t)\right] = \int_0^t \int_t^\infty f(r)\dot{K}_H(r,s)\,dr\,dN(s) \quad a.s. \tag{58}$$

Proof. Let us observe that by (29) we can write

$$\mathbb{E}\left[\int_t^\infty f(s)\,dB(s)\,\Big|\,\mathcal{F}(t)\right] = \mathbb{E}\left[\int_0^\infty K_H^*(\mathbb{1}_{[t,\infty)}f)\,dN\,\Big|\,\mathcal{F}(t)\right] \quad \text{a.s.}$$

Realizing that N is a Wiener process and hence $(N(r) - N(t),\ r > t)$ and $\mathcal{F}(t)$ are independent for all $t \geq 0$ and using the fact that the Wiener integral has zero expectation we may write

$$\mathbb{E}\left[\int_0^\infty K_H^*(\mathbb{1}_{[t,\infty)}f)\,dN\,\Big|\,\mathcal{F}(t)\right] = \mathbb{E}\left[\int_0^t K_H^*(\mathbb{1}_{[t,\infty)}f)\,dN\,\Big|\,\mathcal{F}(t)\right] \quad \text{a.s.}$$

The integral on the right hand side is $\mathcal{F}(t)$-measurable and hence the conditional expectation can be removed. If we, in addition, use (22) we obtain

$$\int_0^t \int_s^\infty \mathbb{1}_{[t,\infty)}(r)f(r)\dot{K}_H(r,s)\,dr\,dN(s) = \int_0^t \int_t^\infty f(r)\dot{K}_H(r,s)\,dr\,dN(s) \quad \text{a.s.}$$

□

Remark 46 (Very informal)**.** Notice that since by Proposition 44

$$dB(s\,|\,t) = \int_0^t \dot{K}_H(s,r)\,dN(r)\,ds \quad \text{a.s.}$$

Proposition 45 can be symbolically written after exchanging the order of integration as

$$\mathbb{E}\left[\int_t^\infty f(s)\,dB(s)\,\Big|\,\mathcal{F}(t)\right] = \int_t^\infty f(s)\,dB(s\,|\,t), \text{ for all } t \geq 0 \quad \text{a.s.}$$

and hence provides an extension of Proposition 44.

It will prove useful to obtain the following estimate.

Proposition 47. *Fix a function $f \in |\mathcal{H}|$. Then there exists a constant b_H so that*

$$\int_0^t \left(\int_t^\infty f(s)\dot{K}_H(s,r)\,ds\right)^2 dr \leq \left|\mathbb{1}_{[t,\infty]}f\right|^2_{|\mathcal{H}|} \leq b_H^2\left|\mathbb{1}_{[t,\infty]}f\right|^2_{L^{\frac{1}{H}}} \quad \textit{for all } t \geq 0.$$

Proof. By Proposition 44, the Itô isometry and the Jensen inequality, we may write

$$\begin{aligned}\int_0^t \left(\int_t^\infty f(s)\dot{K}_H(s,r)\,ds\right)^2 dr &= \mathbb{E}\left(\int_0^t \int_t^\infty f(s)\dot{K}_H(s,r)\,ds\,dN(r)\right)^2 \\ = \mathbb{E}\left\{\mathbb{E}\left[\int_t^\infty f(s)\,dB(s)\,\Big|\,\mathcal{F}(t)\right]\right\}^2 &\leq \mathbb{E}\left(\int_t^\infty f(s)\,dB(s)\right)^2 \quad \text{for all } t \geq 0.\end{aligned}$$

From Remark 29 we know that there exists a positive constant b_H so that

$$\mathbb{E}\left(\int_t^\infty f(s)\,dB(s)\right)^2 \leq \left|\mathbb{1}_{[t,\infty]}f\right|^2_{|\mathcal{H}|} \leq b_H^2\left|\mathbb{1}_{[t,\infty]}f\right|^2_{L^{\frac{1}{H}}} \quad \text{for all } t \geq 0.$$

This concludes the proof. □

2.7 Oracle's control

Let the settings of Section 2.2 be in force and let, similarily as in Section 2.5, $(\bar{X}(t),\ t \geq 0) = (X^{\bar{u},x}(t),\ t \geq 0)$ denote the solution of the system equations (43) with control $u := \bar{u}$, initial condition $X^{\bar{u},x}(0) = x$ and the true parameter $\theta_0 := \theta_0 \in \Theta$. This means that $\bar{X}$ satisfies the stochastic differential equation driven by fractional Brownian motion

$$\begin{aligned} d\bar{X}(t) &= [A(\theta_0)\bar{X}(t) + G\bar{u}(t)]\,dt + \sigma\,dB(t) \quad \text{for } t > 0, \\ \bar{X}(0) &= x. \end{aligned} \tag{59}$$

The feedback control $\bar{u}$ is defined as in formula (54) only with θ replaced by the true parameter θ_0 resulting in

$$\bar{u}(t) := -R^{-1}G^*[P(\theta_0)\bar{X}(t) + W(t;\theta_0))] \quad \text{for all } t \geq 0. \tag{60}$$

If we knew the value of the true parameter θ_0, and (A1) and (A2) were satisfied, we could actually use the control (60) to control the process. Based on Theorem 41 we know that the feedback control (60) is the solution of the ergodic optimal control problem (43) with the cost functional (44) and hence would provide an optimal control for the system.

We may derive an explicit solution for the trajectory $\bar{X}(t)$ of the optimally controlled system. We can expand $\bar{u}$ to obtain a stochatic differential equation driven by fractional Brownian motion for the trajectory of the system that reads

$$\begin{aligned} d\bar{X}(t) &= \left\{[A(\theta_0) - GR^{-1}G^*P(\theta_0)]\bar{X}(t) - GR^{-1}G^*W(t;\theta_0)\right\}\,dt + \sigma\,dB(t) \\ &\qquad \text{for } t \geq 0, \\ \bar{X}(0) &= x. \end{aligned} \tag{61}$$

From Theorem 41 it also follows that the optimally controlled process trajectory $\bar{X}$ satisfies (52), i.e.

$$\lim_{t\to\infty} \frac{1}{t} E\left\langle P(\theta_0)\bar{X}(t), \bar{X}(t)\right\rangle = 0.$$

The problem with the control $\bar{u}$ is, however, that we do not know the value of the true parameter and hence such a control is infeasible in reality. We also excuded such controls which use the value of θ_0 in Definition 26. In the next section, we define a control, that we can calculate with no knowledge of θ_0, except for Θ, while still being optimal.

2.8 An adaptive control

Let the settings of Section 2.2 be in force. On top of this require (A1) and (A2) from Section 2.4 to hold.

In order to obtain a useful feedback control, we have to be able to calculate the control even if we remain ignorant of the value of the true parameter θ_0. A natural candidate for the adaptive ergodic control (43) would be the feedback control obtained by replacing θ by the on-line parameter estimate $\hat{\theta}(t)$, described in Section 2.3, in the defining formula for control (54). This leads to the feedback control

$$\tilde{u}(t) := -R^{-1}B^*[P(\hat{\theta}(t))\tilde{X}(t) + W(t;\hat{\theta}(t))] \quad \text{for } t \geq 0, \tag{62}$$

where $\tilde{X}$ is the solution of (43) with the control $u := \tilde{u}$. The hope with this approach is that we will be able to prove that such a control is a solution of the optimal ergodic control problem (43) with the cost functional (39) by showing that

$$J(x, \bar{u}) = J(x, \tilde{u}).$$

Duncan et al. (2002), on p. 4096, use a similar idea. It is called the "certainty equivalence principle". They are, however, in a different position for they use the certainty equivalence principle to construct an implementable on-line control based on a stochastic differential equation the process W has to satisfy. Then they replace the optimally controlled trajectory with the adaptively controlled one and derive an estimate of the driving Brownian motion from the adaptively controlled trajectory. We choose a different approach here. We aim at obtaining an optimal adaptive control candidate from the formula for W that is available in (51).

The expression $W(t; \hat{\theta}(t))$ present in control (62), however, demands more careful analysis before it can be used. At this point, due to the presence of a stochastic integral in the definition of W, is it not even clear whether $W(t; \hat{\theta}(t))$ makes any sense at all. A well defined process inspired by the form of $W(t; \hat{\theta}(t))$ is defined in Section 2.8.1 which can be used further.

2.8.1 The predictable noise effect on the trajectory

Recall from Section 2.3 that among other properties the estimator $\hat{\theta}$ is $\mathcal{K}(t)$-progressive, where $\mathcal{K}(t) = \mathcal{F}(t) \vee \mathcal{G}(t)$ and $\mathcal{F}(t)$ is independent of $\mathcal{G}(t)$ for every $t \geq 0$ and $\mathcal{F}(t)_{t \geq 0}$ and $\mathcal{G}(t)_{t \geq 0}$ are filtrations of $\mathcal{F}$. The estimator also satisfies $\hat{\theta}(t) \in \Theta$ for every $t \geq 0$.

Recall that W was defined in (51) as

$$W(t; \theta) := \mathbb{E}\left[\int_t^\infty \Psi(s - t; \theta) P(\theta) \sigma \, dB(s) \,\Big|\, \mathcal{F}(t) \right] \qquad \text{for } t \geq 0 \text{ and } \theta \in \Theta. \quad (63)$$

In the following proposition we establish a useful representation of W as an Itô integral.

Proposition 48. *Fix $\theta \in \Theta$. For the coordinates of the process $(W(t; \theta),\ t \geq 0)$ we have a representation*

$$W_i(t; \theta) = \sum_{j=1}^m \sum_{k=1}^m \sum_{l=1}^n \sigma_{kl} \int_0^t w_{ijk}(s, t) \, dN_l(s) \qquad \textit{for } t \geq 0 \ \textit{a.s.} \quad (64)$$

for every $i = 1, \ldots, m$ where

$$w_{ijk}(s, t) := \int_t^\infty \Psi_{ij}(r - t; \theta) P_{jk}(\theta) \dot{K}_H(r, s) \, dr.$$

Proof. We may rewrite $W(t, \theta)$ by coordinates, using linearity of stochastic integral and conditional expectation to obtain

$$W_i(t; \theta) = \sum_{j=1}^m \sum_{k=1}^m \sum_{l=1}^n P_{jk}(\theta) \sigma_{kl} \, \mathbb{E}\left[\int_t^\infty \Psi(s - t; \theta)_{ij} \, dB_l(s) \,\Big|\, \mathcal{F}(t) \right]$$

for all $t \geq 0$, $\theta \in \Theta$ and $i = 1, \dots, m$. By arguments similar to those of (49) we have

$$\|\Psi(t;\theta)\| \leq M_P(\theta) e^{-K_P(\theta)t} \quad \text{for all } t \geq 0 \tag{65}$$

for some positive constants $M_P(\theta)$ and $K_P(\theta)$ for arbitrary $\theta \in \Theta$. Observe that

$$\begin{aligned}\int_t^\infty \Psi(s-t;\theta)_{ij}^{\frac{1}{H}}\, ds \leq \int_t^\infty \|\Psi(s-t;\theta)\|^{\frac{1}{H}}\, ds &\leq \int_t^\infty M_P^{\frac{1}{H}}(\theta) e^{-\frac{K_P(\theta)}{H}(s-t)}\, ds \\ &\leq \frac{H M_P^{\frac{1}{H}}(\theta)}{K_P(\theta)} \quad \text{for all } t \geq 0\end{aligned}$$

and hence $\Psi(\cdot - t;\theta) \in L^{\frac{1}{H}}(t,\infty)$ for all $t \geq 0$. This means we may use Proposition 45 and rewrite

$$\mathbb{E}\left[\int_t^\infty \Psi(s-t;\theta)_{ij}\, dB_l(s) \,\Big|\, \mathcal{F}(t)\right] = \int_0^t \int_t^\infty \Psi(r-t;\theta)_{ij} \dot{K}_H(r,s)\, dr\, dN_l(s) \quad \text{a.s.}$$

where $N(t) = \int_0^\infty (K_H^*)^{-1}(\mathbb{1}_{[0,t]}) dB$, $t \geq 0$ is the Wiener process associated with B as in (27). □

We use the representation as an inspiration and define a random process $(V(t),\ t \geq 0)$ by coordinates for $t \geq 0$ and $i = 1, \dots, m$ using the equality

$$V_i(t) := \sum_{j=1}^m \sum_{k=1}^m \sum_{l=1}^n \sigma_{kl} \int_0^t v_{ijk}(s,t)\, dN_l(s) \tag{66}$$

where

$$v_{ijk}(s,t) := \int_t^\infty \Psi(r-t;\hat{\theta}(s))_{ij} P_{jk}(\hat{\theta}(s)) \dot{K}_H(r,s)\, dr, \quad r \geq t \geq 0 \text{ and } 0 \leq s \leq t$$

and the integral in (66) is understood in the Itô sense.

Proposition 49. *The process V is well defined, $\mathcal{K}(t)$-progressive and continuous.*

Proof. It is well known that an $\mathcal{F}(t)$-Wiener process is also an $[\mathcal{F}(t)\vee\mathcal{J}(t)]$-Wiener process for a filtration $\mathcal{J}(t)_{t\geq 0}$ of $\mathcal{F}$ such that $\mathcal{F}(t)$ and $\mathcal{J}(t)$ are independent for all $t \geq 0$. This is exactly our situation since $\mathcal{K}(t) = \mathcal{F}(t) \vee \mathcal{G}(t)$, and $\mathcal{F}(t)$ is independent of $\mathcal{G}(t)$ for all $t \geq 0$. Noting this, it is clear that in order to show that V is well defined, we have to show that for every $t \geq 0$ the processes $v_{ijk}(\cdot,t)$ are $[\mathcal{K}(s)_{s\geq 0}]$-progressive, or have a progressive modification, and that $v_{ijk}(\cdot,t) \in L^2(0,t)$ for all $t \geq 0$ a.s. for $i,j,k = 1,\dots,m$.

First, we will show progressiveness. Recall (50), defining $\Psi(t;\theta) = \exp\{A_P(\theta)t\}$ for $t \geq 0$ and $\theta \in \Theta$. Observe that $\Psi(t;\theta)$ is measurable by Lemma 13. On top of this, $P(\theta)$ is continuous by the assumptions and Theorem 38 and hence measurable. Fix $t \geq 0$ and $i,j,k = 1,\dots,m$.

Put

$$\rho(s,\theta) := \int_t^\infty \Psi(r-t;\theta)_{ij} P_{jk}(\theta) \dot{K}_H(r,s)\, dr.$$

By Proposition 39, $\|\Psi(t;\theta)\| \le Me^{-Kt}$ for all $t \ge 0$ and $\theta \in \Theta$. By (20) we may estimate

$$\left|\int_t^\infty \Psi(r-t;\theta)_{ij}P_{jk}(\theta)\dot K_H(r,s)\,dr\right|$$
$$\le c_H s^{\frac12 - H}\|P(\theta)\| \int_t^\infty \|\Psi(r-t;\theta)\| r^{H-\frac12}(r-s)^{H-\frac32}\,dr$$
$$\le Mc_H s^{\frac12 - H}\|P(\theta)\| \int_t^\infty e^{-K(r-t)} r^{H-\frac12}(r-s)^{H-\frac32}\,dr. \quad (67)$$

The integral on the right converges for all $t > s > 0$ without any problems. Realize that by the assumption $H > 1/2$ we have $H - 1/2 > 0$ and $-1 < H - 3/2 < 0$. Hence, for $t = s > 0$ we may compare the integral $\int_t^{t+1} e^{-K(r-t)} r^{H-\frac12}(r-s)^{H-\frac32}\,dr$ with $(2t)^{H-\frac12}\int_0^1 r^{H-\frac32}\,dr$ which converges. Hence the integral (67) converges and may be approximated by sums of approximating simple functions. The integral thus preserves measurability and the function ρ is measurable. We have $v_{ijk}(s,t) = \rho(s,\hat\theta(s))$ is $[\mathcal{K}(\tau)_{\tau\ge 0}]$-adapted and measurable for any fixed $t > 0$. This means $(v_{ijk}(s,t),\ s > 0)$ has a $[\mathcal{K}(\tau)_{\tau\ge0}]$-progressive modification $(v_{ijk}^{\mathrm{mod}}(s,t),\ s > 0)$. We extend $v_{ijk}(s,t)$ by $v_{ijk}^{\mathrm{mod}}(0,t) = 0$ and obtain that that $(v_{ijk}(s,t),\ s \ge 0)$ has a $[\mathcal{K}(\tau)_{\tau\ge0}]$-progressive modification, the extended v_{ijk}^{mod}.

By Proposition 39, Ψ is stable uniformly with respect to θ, and realizing that $\dot K(r,s) \ge 0$ for all $r > s \ge 0$, cf. (20), we may write for $i,j,k = 1,\dots,m$

$$\int_0^t v_{ijk}^2(s,t)\,ds = \int_0^t \left(\int_t^\infty \Psi(r-t;\hat\theta(s))_{ij}P_{jk}(\hat\theta(s))\dot K_H(r,s)\,dr\right)^2 ds$$
$$\le \int_0^t \left(\int_t^\infty \left\|\Psi(r-t;\hat\theta(s))\right\|\left\|P(\hat\theta(s))\right\|\left|\dot K_H(r,s)\right|dr\right)^2 ds$$
$$\le C_P\int_0^t \left(\int_t^\infty Me^{-K(r-t)}\dot K_H(r,s)\,dr\right)^2 ds \quad \text{for all } t\ge 0 \text{ a.s.} \quad (68)$$

where we used the fact that by Corollary 40, $C_P := \operatorname{esssup}_{\omega\in\Omega}\sup_{t\ge0}\|P(\hat\theta(t,\omega))\|$ is finite. Using Proposition 47, there is a positive constant b_H such that

$$\int_0^t \left(\int_t^\infty Me^{-K(r-t)}\dot K_H(r,s)\,dr\right)^2 ds \le b_H^2\left|\mathbb{1}_{[t,\infty]}Me^{-K(r-t)}\right|^2_{L^{\frac1H}}$$
$$= b_H^2\left(\int_t^\infty M^{\frac1H}e^{-\frac{K}{H}(r-t)}\,dr\right)^{2H} \le b_H^2\frac{M^2H^{2H}}{K^{2H}} < \infty \quad \text{for all } t\ge 0. \quad (69)$$

It is now clear that V is well defined, $\mathcal{K}(t)$-progressive and continuous. □

Let us turn to the estimation of moments of V.

Proposition 50. *There exists a positive constant D_1 such that we can estimate*

$$\mathbb{E}|V(t)|^2 \le D_1 \quad \textit{for all } t \ge 0.$$

Proof. Using Remark 3 and the definition (66), we may write

$$\mathbb{E}|V_i(t)|^2 \le m^4n^2\sum_{j=1}^m\sum_{k=1}^m\sum_{l=1}^n \sigma_{kl}^2\,\mathbb{E}\left(\int_0^t v_{ijk}(s,t)\,dN_l(s)\right)^2 \quad \text{for } t\ge 0.$$

Using the Itô isometry and the exact same arguments as in (68) we obtain for $i, j, k = 1, \ldots, m$

$$\begin{aligned}\mathbb{E}\left(\int_0^t v_{ijk}(s,t)\,dN(s)\right)^2 &= \mathbb{E}\int_0^t v_{ijk}^2(s,t)\,ds \\ &\leq C_P \int_0^t \left[\int_t^\infty Me^{-K(r-t)}\dot{K}_H(r,s)\,dr\right]^2 ds \quad \text{for all } t \geq 0.\end{aligned}$$

Exactly as in (69) we obtain existence of a positive constant b_H so that

$$\begin{aligned}C_P \int_0^t \left[\int_t^\infty Me^{-K(r-t)}\dot{K}_H(r,s)\,dr\right]^2 ds &\leq b_H^2 \left|\mathbb{1}_{[t,\infty]} v_{ijk}\right|^2_{L^{\frac{1}{H}}} \\ &\leq b_H^2 \frac{M^2 H^{2H}}{K^{2H}} < \infty \quad (70)\end{aligned}$$

for all $t \geq 0$. Setting $D_1 := b_H^2 M^2 H^{2H}/K^{2H}$ concludes the proof. □

Proposition 51. *For every $q \geq 1$, there exists a positive constant D_q such that we can estimate*

$$\mathbb{E}|V(t)|^{2q} \leq D_q \quad \text{for all } t \geq 0.$$

Proof. Observe that being Itô stochastic integrals, the processes

$$w_{ijk}(t,T) := \int_0^t v_{ijk}(s,T)\,dN(s), \quad 0 \leq t \leq T$$

are continuous local L^2-martingales by definition for all $T \geq 0$ and $i, j, k = 1, \ldots, m$. We can use the Burkholder-Davis-Gundy inequality, cf. Theorem 17.7 in Kallenberg (2001) on p. 333 to obtain for $i, j, k = 1, \ldots, m$ the estimate

$$\mathbb{E}|w_{ijk}(T,T)|^{2q} \leq \mathbb{E}\sup_{0\leq t\leq T}|w_{ijk}(t,T)|^{2q} \leq c_q\,\mathbb{E}\left[\int_0^T v_{ijk}^2(s,T)\,ds\right]^q$$

for some positive constant c_q and all $T \geq 0$. By arguments similar to (68) and (70) we obtain

$$c_q\,\mathbb{E}\left[\int_0^T v_{ijk}^2(s,T)\,ds\right]^q \leq c_q\left(b_H^2\frac{M^2H^{2H}}{K^{2H}}\right)^q < \infty \quad \text{for } i,j,k = 1,\ldots,m$$

where the bound does not depend on $T \geq 0$. Using Remark 3 the proof is easily concluded. □

2.8.2 The on-line control trajectory

Let $(\hat{X}(t),\ t \geq 0) = (X^{\hat{u},x}(t),\ t \geq 0)$, similarily as in Section 2.7, denote the solution of the system equations (43) with control $u := \hat{u}$, initial condition $X^{\hat{u},x}(0) = x$ and the true parameter $\theta_0 := \theta_0 \in \Theta$. This means that $\hat{X}$ satisfies the stochastic differential equation driven by a fractional Brownian motion

$$\begin{aligned}d\hat{X}(t) &= [A(\theta_0)\hat{X}(t) + G\hat{u}(t)]\,dt + \sigma\,dB(t) \quad \text{for } t > 0, \\ \hat{X}(0) &= x.\end{aligned} \quad (71)$$

The feedback control $\hat{u}$ is defined similarily as in the formula (54) only with θ replaced by the true parameter θ_0 and by replacing W with V resulting in

$$\hat{u}(t) := -R^{-1}B^*[P(\hat{\theta}(t))\hat{X}(t) + V(t)] \quad \text{for all } t \geq 0. \tag{72}$$

The trajectory $\hat{X}(t)$ of the adaptively controlled system evolves according to (43) in which we substitute $\hat{u}$ for u and obtain

$$\begin{aligned} d\hat{X}(t) &= \left\{ \left[A(\theta_0) - GR^{-1}G^*P(\hat{\theta}(t)) \right] \hat{X}(t) - GR^{-1}G^*V(t) \right\} dt + \sigma \, dB(t) \\ &\qquad \text{for all } t > 0, \\ \hat{X}(0) &= x. \end{aligned} \tag{73}$$

The following proposition is essential for the proof of the main theorem of the thesis. It solves a difficulty that is at the heart of the problem of adaptive control, namely, the question of the stability of the adaptively controlled system. The main task is to cope with the fact that the matrix

$$A(\theta_0) - GR^{-1}G^*P(\hat{\theta}(t)), \quad \text{for } t \geq 0$$

governing the evolution of the adaptively controlled system, cf. (73), is not necessarily exponentially stable. We obtain an estimate of the same quality as for the optimally controlled system which we formulated in Proposition 42.

We first prove that the adaptively controlled trajectory is well defined and has continuous moments.

Proposition 52. *The process $\hat{X}$ defined as a solution to the equation* (73) *is a well defined continuous and $\mathcal{K}(t)$-progressive process. The process is unique in the sense that if there is another process $\tilde{X}$ satisfying* (73)*, then $\mathbb{P}\{\hat{X}(t) = \tilde{X}(t), \text{ for all } t \geq 0\} = 1$.*

Moreover, $\mathbb{E}|\hat{X}(t)|^{2q}$ is continuous in t on $\mathbb{R}_+$ for all $q \geq 1$.

Proof. Recall, that we observed at the start of Section 2 that the fractional Brownian motion B is $\mathcal{K}(t)$-progressive.

Fix $q > 0$. By assumptions, $\hat{\theta}$ is continuous. By Theorem 38, $P(\theta)$ is continuous in $\theta \in \Theta$. Since A was assumed to be continuous, we obtain that $A_P(\theta) = A(\theta) - GR^{-1}G^*P(\theta)$ is continuous in $\theta \in \Theta$. Moreover, since Θ is compact, we obtain that

$$\operatorname*{esssup}_{\omega \in \Omega} \sup_{t \geq 0} \|A_P(\hat{\theta}(t, \omega))\|^{2q} < \infty.$$

By Proposition 51 using the Fubini theorem we have $\mathbb{E} \int_0^t |V(r)|^{2q} dr \leq tD_q < \infty$ for all $t \geq 0$. This implies that $V \in L^1_{loc}(\mathbb{R}_+; \mathbb{R}^{m \times m})$ a.s. We can thus use Theorem 33 to conclude that $\hat{X}$ is continuous global and unique.

Fix $T > 0$ and $0 \leq s \leq T$. The random variable $2|\sup_{0 \leq r \leq T} \hat{X}(s)|^{2q}$ is an integrable majorant by previous considerations and the estimate in Theorem 33. The process $\hat{X}$, and $\hat{X}^{2q}$ in turn, was already proved to be continuous. We may use the dominated convergence theorem and write

$$\mathbb{E}\left|\hat{X}(t)\right|^{2q} - \mathbb{E}\left|\hat{X}(s)\right|^{2q} \leq \mathbb{E}\left[\left|\hat{X}(t)\right|^{2q} - \left|\hat{X}(s)\right|^{2q}\right] \to 0 \quad \text{as } t \to s \text{ so that } t < T.$$

Hence $\mathbb{E}|\hat{X}|^{2q}$ is continuous. This concludes the proof. $\square$

We will need one rather technical lemma.

Lemma 53. *Fix $q \geq 1$, natural numbers m and n, a stable matrix $A \in \mathbb{R}^{m\times m}$, a matrix $c \in \mathbb{R}^{m\times n}$ and an n-dimensional fractional Brownian motion B with Hurst parameter $H > 1/2$. Let S be the flow map of A. Then $\int_0^t S(t,r)cdB(r)$ is well defined for all $t \geq 0$ and there exists a positive constant F_q so that*

$$\mathbb{E}\left|\int_0^t S(t,r)c\,dB(r)\right|^{2q} < F_q \quad \textit{for all } t \geq 0.$$

Proof. It is well known, cf. e.g. Section 12.14 in Lin & Bai (2010) on p. 176, that all moments of a Gaussian variable are equivalent. It is hence enough to consider $q = 1$.

Let us first calculate

$$\int_0^t ae^{-b(t-r)}\,dr = \frac{a}{b}\left(1 - e^{-bt}\right) \leq \frac{a}{b} \quad \text{for all } t \geq 0 \tag{74}$$

where a and b are arbitrary constants.

By linearity of the stochastic integral and triangle inequality, we first rewrite for every coordinate $i = 1, \ldots, m$

$$\left(\mathbb{E}\int_0^t S(t,r)c\,dB(r)\right)_i = \sum_{j=1}^m \sum_{k=1}^m c_{jk}\,\mathbb{E}\int_0^t S_{ij}(t,r)\,dB_k(r) \quad \text{for all } t \geq 0. \tag{75}$$

Since A is stable, there exist positive constants M and K so that $|S_{ij}(t,r)| \leq \|S(t,s)\| \leq Me^{-K(t-s)}$ for all $t \geq s \geq 0$, $i, j = 1, \ldots, m$. This means that

$$\int_0^t S_{ij}^{\frac{1}{H}}(t,r)\,dr \leq M^{\frac{1}{H}} \int_0^t e^{-\frac{K}{H}(t-r)}\,dr \leq \frac{HM^{\frac{1}{H}}}{2K}.$$

We obtained that $S_{ij}(t,r) \in L^{\frac{1}{H}}(\mathbb{R}_+)$. By Remark 29 it means that $S_{ij}(t,r) \in |\mathcal{H}|$ and hence all the stochastic integrals in (75) are well defined for all $t \geq 0$ and hence also $\int_0^t S(t,r)cdB(r)$ is.

By Remark 29 we have that there exists a positive constant b_H so that

$$\mathbb{E}\left(\int_0^t S_{ij}(t,r)\,dB(r)\right)^2 \leq b_H^2\left(\int_0^t S_{ij}^{\frac{1}{H}}(t,r)\,dr\right)^2 \leq \left(\frac{HM^{\frac{1}{H}}}{2K}\right)^2$$

for all $t \geq 0$. Using (75) and Remark 3 we obtain for all $t \geq 0$

$$\mathbb{E}\left|\int_0^t S(t,r)c\,dB(r)\right|^2 \leq m^3 \sum_{j=1}^m \sum_{k=1}^m c_{ik}\,\mathbb{E}\left|\int_0^t S_{ij}(t,r)\,dB_k(r)\right|^2 \leq m^5\|c\|\frac{M^2}{2K}.$$

The proof is thus concluded. $\square$

Proposition 54. *Let $\hat{X}$ be the process defined in (73) and let $q \geq 1$. Assume there exist nonnegative continuous functions $g, h\colon \mathbb{R}_+ \to \mathbb{R}_+$ such that $g(t) \to 0$ as $t \to \infty$ and h is bounded satisfying*

$$\mathbb{E}\left(\left\|P(\theta_0) - P(\hat{\theta}(t))\right\|\left|\hat{X}(t)\right|\right)^{2q} \leq g(t)\,\mathbb{E}\left|\hat{X}(t)\right|^{2q} + h(t) \quad \textit{for all } t \geq 0. \tag{76}$$

Then there exists a positive constant E_q so that for each $t > 0$ the estimate

$$\mathbb{E}\left|\hat{X}(t)\right|^{2q} \leq E_q. \tag{77}$$

is satisfied.

Remark 55. Denote

$$g(t) := \mathbb{E}\left\|P(\theta_0) - P(\hat{\theta}(t))\right\| \quad \text{for } t \geq 0.$$

By Corollary 40 we have that $g(t)^{2q} \to 0$ as $t \to \infty$.

The condition (76) is satisfied in two basic cases. If there exists $t_0 \geq 0$ so that the estimate $\hat{\theta}(t)$ is independent of $\mathcal{F}(t)$ for all $t \geq t_0$, i.e. if $\mathcal{F}(t)$ is independent of $\mathcal{G}(t)$ for all $t \geq 0$.

Or, in the case when $\|P(\theta_0) - P(\hat{\theta}(t))\|$ and $|\hat{X}(t)|$ are negatively quadrant dependent for all $t \geq t_0$ for some $t_0 \geq 0$ i.e.

$$\begin{aligned}\mathbb{P}\left(\left\|P(\theta_0) - P(\hat{\theta}(t))\right\| \geq x, \left|\hat{X}(t)\right| \geq y\right)&\\ \leq \mathbb{P}\left(\left\|P(\theta_0) - P(\hat{\theta}(t))\right\| \geq x\right) \mathbb{P}\left(\left|\hat{X}(t)\right| \geq y\right)&\end{aligned} \tag{78}$$

for all real x, y and all $t \geq t_0$. Since both $\|P(\theta_0) - P(\hat{\theta}(t))\|$ and $|\hat{X}(t)|$ are nonnegative, it is enough to check condition (78) for nonnegative x and y only.

Let $\|P(\theta_0) - P(\hat{\theta}(t))\|$ and $|\hat{X}(t)|$ are negatively quadrant dependent for all $t \geq t_0$. Choose $x', y' \in \mathbb{R}_+$ and set $x = x'^{\frac{1}{2q}}$ and $y' = y'^{\frac{1}{2q}}$. Then by (78) we have

$$\begin{aligned}\mathbb{P}\left(\left\|P(\theta_0) - P(\hat{\theta}(t))\right\|^{2q} \geq x', \left|\hat{X}(t)\right|^{2q} \geq y'\right)&\\ \leq \mathbb{P}\left(\left\|P(\theta_0) - P(\hat{\theta}(t))\right\|^{2q} \geq x'\right) \mathbb{P}\left(\left|\hat{X}(t)\right|^{2q} \geq y'\right) \quad &\text{for all } t \geq t_0.\end{aligned}$$

We obtained that $\|P(\theta_0) - P(\hat{\theta}(t))\|^{2q}$ and $|\hat{X}(t)|^{2q}$ are negatively quadrant dependent for all $t \geq t_0$ as well.

If $\|P(\theta_0) - P(\hat{\theta}(t))\|$ and $|\hat{X}(t)|$ are negatively quadrant dependent for all $t \geq t_0$ then by Chapter 11.1 in Lin & Bai (2010) on p. 149 we have

$$\mathbb{E}\left(\left\|P(\theta_0) - P(\hat{\theta}(t))\right\|^{2q}\left|\hat{X}(t)\right|^{2q}\right) \leq \mathbb{E}\left\|P(\theta_0) - P(\hat{\theta}(t))\right\|^{2q}\mathbb{E}\left|\hat{X}(t)\right|^{2q}$$

for all $t \geq t_0$ and the condition (76) is satisfied.

In this view, the condition (76) can be understood as asymptotic negative quadrant dependence of $\|P(\theta_0) - P(\hat{\theta}(t))\|$ and $|\hat{X}(t)|$ as $t \to \infty$.

Proof of Proposition 54. Let us rewrite the defining equation of $\hat{X}$, the equation (73), by adding and subtracting $GR^{-1}G^*P(\theta_0)$ to obtain

$$\begin{aligned}d\hat{X}(t) &= \Big\{ \left[A(\theta_0) - GR^{-1}G^*P(\theta_0)\right]\hat{X}(t)\\ &\qquad + GR^{-1}G^*\left[P(\theta_0) - P(\hat{\theta}(t))\right]\hat{X}(t) - GR^{-1}G^*V(t)\Big\}dt + \sigma dB(t)\\ &\qquad \text{for all } t > 0,\\ \bar{X}(0) &= x.\end{aligned} \tag{79}$$

By Proposition 14 the map defined by $S(t,s) := \exp\{A_P(\theta_0)(t-s)\}$ for $t \geq s \geq 0$ is the flow map of the deterministic matrix $A_P(\theta_0)$. By the stochastic variation of constants formula in Proposition 35 and the uniqueness of $\hat{X}$ warranted by Proposition 52 we have

$$\begin{aligned}\hat{X}(t) = S(t,0)x &+ \int_0^t S(t,r)GR^{-1}G^* \left[P(\theta_0) - P(\hat{\theta}(r))\right] \hat{X}(r)\,dr \\ &- \int_0^t S(t,r)GR^{-1}G^* V(r)\,dr \\ &+ \int_0^t S(t,r)\sigma\,dB(r) \quad \text{for all } t \geq 0 \text{ a.s.}\end{aligned} \tag{80}$$

Using Remark 3, and properties of the integral we may calculate

$$\begin{aligned}\mathbb{E}\left|\hat{X}(t)\right|^{2q} \leq 4^{2q}\Bigg\{ &|x|^{2q}\|S(t,0)\|^{2q} \\ &+ \mathbb{E}\left(\int_0^t \|S(t,r)\|^{\frac{1}{2}}\|S(t,r)\|^{\frac{1}{2}}\left\|GR^{-1}G^*\right\|\left\|P(\theta_0) - P(\hat{\theta}(r))\right\|\left|\hat{X}(r)\right|dr\right)^{2q} \\ &+ \mathbb{E}\left(\int_0^t \|S(t,r)\|^{\frac{1}{2}}\|S(t,r)\|^{\frac{1}{2}}\left\|GR^{-1}G^*\right\||V(r)|\,dr\right)^{2q} \\ &+ \mathbb{E}\left|\int_0^t S(t,r)\sigma dB(r)\right|^{2q}\Bigg\} \quad \text{for all } t \geq 0.\end{aligned}$$

Denote

$$\mu(q) := \frac{2q}{2q-1} \quad \text{for } q \geq 1.$$

The matrix $GR^{-1}G^*$ is deterministic and thus $\|GR^{-1}G^*\|$ can be taken out of the expectations. Denote

$$C := \left\|GR^{-1}G^*\right\|.$$

We use the Hölder inequality in Theorem 2 with $p := 2q$ and $q := \mu(q)$ to obtain

$$\begin{aligned}\mathbb{E}\left|\hat{X}(t)\right|^{2q} \leq 4^{2q}\Bigg\{ &|x|^{2q}\|S(t,0)\|^{2q} \\ &+ \left[\left(\int_0^t \|S(t,r)\|^{\frac{\mu(q)}{2}}\,dr\right)^{2q-1} \mathbb{E}\int_0^t \|S(t,r)\|^q C^{2q}\left\|P(\theta_0) - P(\hat{\theta}(r))\right\|^{2q}\left|\hat{X}(r)\right|^{2q} dr\right] \\ &+ \left[\left(\int_0^t \|S(t,r)\|^{\frac{\mu(q)}{2}}\,dr\right)^{2q-1} \mathbb{E}\int_0^t \|S(t,r)\|^q C^{2q}|V(r)|^{2q}\,dr\right] \\ &+ \mathbb{E}\left(\int_0^t S(t,r)\sigma dB(r)\right)^{2q}\Bigg\} \quad \text{for all } t \geq 0.\end{aligned} \tag{81}$$

By Theorem 37, $A_P(\theta_0)$ is a stable matrix and there exist positive constants M and K so that $\|S(t,s)\| = \exp\{A_P(\theta_0)(t-s)\} \leq Me^{-K(t-s)}$ for all $t \geq s \geq 0$. Put

$$M_q := \frac{M^q}{qK} \quad \text{for } q \in \mathbb{N}.$$

From (74) and the stability of $A_P(\theta_0)$, we obtain for all $t \geq 0$

$$\int_0^t \|S(t,r)\|^q \, dr \leq \int_0^t M^q e^{-qK(t-r)} \, dr \leq M_q \tag{82}$$

and

$$\left(\int_0^t \|S(t,r)\|^{\frac{\mu(q)}{2}} \, dr \right)^{2q-1} \leq \left[\frac{M^{\frac{\mu(q)}{2}}}{\frac{\mu(q)}{2} K} \left(1 - e^{-\frac{\mu(q)}{2} K t} \right) \right]^{2q-1} \leq M_{\mu(q)/2}^{2q-1}. \tag{83}$$

To simplify notation further, we denote for $q \geq 1$

$$\tilde{M}_q := M_{\mu(q)/2}^{2q-1}.$$

By the estimate of moments of V, established in Proposition 51, there exists a positive constant D_q such that $\mathbb{E}|V(t)|^{2q} \leq D_q$ for all $t \geq 0$. Using the stability of $A_P(\theta_0)$ again and calculation (82) we may estimate for all $t \geq 0$

$$\int_0^t \|S(t,r)\|^q \, \mathbb{E}|V(r)|^{2q} \, dr \leq D_q \int_0^t M^q e^{-qK(t-r)} \, dr \leq M_q D_q. \tag{84}$$

By Proposition 53 we have that there exists a positive constant F_q so that

$$\mathbb{E}\left| \int_0^t S(t,r) \sigma dB(r) \right|^{2q} \leq F_q \quad \text{ for all } t \geq 0. \tag{85}$$

Realize that $\|S(t,r)\|$ is deterministic and calculate using (84), (85), the assumption (76) and the Fubini theorem

$$\begin{aligned}
\mathbb{E}\left| \hat{X}(t) \right|^{2q} \leq 4^{2q} \Bigg\{ & M^{2q} |x|^{2q} \\
& + \tilde{M}_q C^{2q} \, \mathbb{E} \int_0^t \|S(t,r)\|^q \left\| P(\theta_0) - P(\hat{\theta}(r)) \right\|^{2q} \left| \hat{X}(r) \right|^{2q} dr \\
& + \tilde{M}_q C^{2q} \, \mathbb{E} \int_0^t \|S(t,r)\|^q |V(r)|^{2q} \, dr + F_q \Bigg\} \\
\leq 4^{2q} \Bigg\{ & M^{2q} |x|^{2q} + \tilde{M}_q C^{2q} \int_0^t \|S(t,r)\|^q \left[g(r) \, \mathbb{E}\left| \hat{X}(r) \right|^{2q} + h(r) \right] dr \\
& + \tilde{M}_q C^{2q} M_q D_q + F_q \Bigg\} \\
& \leq \tilde{c}(t) + \tilde{d} \int_0^t e^{-qK(t-r)} g(r) \, \mathbb{E}\left| \hat{X}(r) \right|^{2q} dr
\end{aligned} \tag{86}$$

for all $t \geq 0$ where

$$\tilde{c}(t) := 4^{2q} (M^{2q} |x|^{2q} + \tilde{M}_q C^{2q} M_q D_q + F_q) + \tilde{d} \int_0^t e^{-qK(t-r)} h(r) \, dr$$

and $\tilde{d} := 4^{2q} \tilde{M}_q C^{2q} M^q$ is a positive constant.

By assumptions, h is bounded, hence there exists a positive constant $\tilde{h}$ so that $h(t) \leq \tilde{h}$ for all $t \geq 0$. By the estimate (74) we may write $\tilde{c}(t) \leq \tilde{c}$ for all $t \geq 0$ where

$$\tilde{c} := 4^{2q}(M^{2q}|x|^{2q} + \tilde{M}_q C^{2q} M_q D_q + F_q) + \frac{\tilde{d}\tilde{h}}{qK}.$$

Observe that (86) is in the form suitable for the usage of the generalized Grönwall lemma. Fix $T \geq 0$ and write

$$\varphi(t) \leq \tilde{c} + \tilde{d}e^{-qKt} \int_0^t e^{qKr} g(r)\varphi(r)\, dr \quad \text{for all } 0 < t < T. \tag{87}$$

Put $\varphi(t) := \mathbb{E}|\hat{X}(t)|^{2q}$, $t \geq 0$. Then the inequality (87) is equivalent with (86). We know that $\mathbb{E}|\hat{X}(t)|^{2q}$ is continuous in t by Proposition 52. Hence the functions $\varphi, g, \tilde{c}$ and the exponential are all continuous. We may use Theorem 6 and obtain for all $0 < t < T$

$$\varphi(t) \leq \tilde{c} + \tilde{d}e^{-qKt} \int_0^t \tilde{c}e^{qKs} g(s) \exp\left(\int_s^t e^{-qKr} e^{qKr} g(r)\, dr\right) ds,$$

which simplifies to

$$\varphi(t) \leq \tilde{c}\left[1 + \tilde{d}\int_0^t e^{-qK(t-s)} g(s) \exp\left(\int_s^t g(r)\, dr\right) ds\right] \quad \text{for all } 0 < t < T. \tag{88}$$

Fix $\varepsilon < qK$. Since $g(s) \to 0$ as $t \to \infty$, we can find such $t_0 \geq 0$ that satisfies $g(t) < \varepsilon$ for all $t \geq t_0$. Moreover, there exists a positive constant $\tilde{C}$ so that $g(t) \leq \tilde{C}$ for all $t \geq 0$. For all $t \geq t_0$

$$\begin{aligned}
&\int_0^t e^{-qK(t-s)} g(s) \exp\left(\int_s^t g(r)\, dr\right) ds \\
&= \int_0^{t_0} e^{-qK(t-s)} g(s) \exp\left(\int_s^{t_0} g(r)\, dr\right) e^{\varepsilon(t-t_0)}\, ds + \int_{t_0}^t e^{-qK(t-s)} g(s) e^{\varepsilon(t-s)}\, ds \\
&\qquad = \tilde{C}\tilde{K}e^{-(qK-\varepsilon)t-\varepsilon t_0} \int_0^{t_0} e^{qKs}\, ds + \int_{t_0}^t e^{-(qK-\varepsilon)(t-s)} g(s)\, ds \\
&\qquad = \frac{e^{(qK-\varepsilon)t_0} - e^{-\varepsilon t_0}}{qK} \tilde{C}\tilde{K}e^{-(qK-\varepsilon)t} + \int_{t_0}^t e^{-(qK-\varepsilon)(t-s)} g(s)\, ds \\
&\qquad \leq e^{-(qK-\varepsilon)t}\left[\frac{e^{(qK-\varepsilon)t_0} - e^{-\varepsilon t_0}}{qK}\tilde{C}\tilde{K} + \varepsilon\frac{e^{(qK-\varepsilon)t} - e^{(qK-\varepsilon)t_0}}{qK-\varepsilon}\right] \\
&\qquad\qquad \leq \frac{1}{qK}\tilde{C}\tilde{K} + \frac{\varepsilon}{qK-\varepsilon} < \infty
\end{aligned} \tag{89}$$

where $\tilde{K} := \exp\left(\int_0^{t_0} g(r)\, dr\right)$. Notice that the estimate does not depend on T. Since obviously

$$\tilde{S} := \sup_{0 \leq t \leq t_0} \tilde{c}\left[1 + \tilde{d}\int_0^t e^{-qK(t-s)} g(s) \exp\left(\int_s^t g(r)\, dr\right) ds\right] < \infty$$

we may set $E_q := \max(\tilde{S}, \tilde{C}\tilde{K}/qK + \varepsilon/(qK - \varepsilon))$ and by (88) conclude that $\sup_{0 \leq t} \varphi(t) = \sup_{0 \leq t} \mathbb{E}|\hat{X}(t)|^{2q} \leq E_q$ for some positive constant E_q. □

3 Adaptive ergodic linear quadratic control problem solution

In this section, we turn to the formulation of the main result of the thesis. The settings of Section 2.2, Section 2.4 and Sections 2.7 and 2.8 are in force.

Theorem 56. *Let $\hat{\theta}$ be the strongly consistent estimator described in Section 2.3 and let (A1) and (A2) from Section 2.4 be satisfied.*

Let $\hat{u}$ be the feedback control defined in formula (72)*, i.e. defined as*

$$\hat{u}(t) := -R^{-1}G^*[P(\hat{\theta}(t))\hat{X}(t) + V(t)] \quad \text{for all } t \geq 0 \tag{90}$$

where V is the process defined in (66) *and $\hat{X}$ is the trajectory of the system satisfying the equation* (43) *controlled with the control $u := \hat{u}$. Let $\bar{u}$ be Oracle's feedback control defined in formula* (60)*, i.e. defined as*

$$\bar{u}(t) := -R^{-1}G^*[P(\theta_0)\bar{X}(t) + W(t;\theta_0))] \quad \text{for all } t \geq 0 \tag{91}$$

where $(W(t;\theta_0),\ t \geq 0)$ is the process defined in (51) *and $\bar{X}$ is the trajectory of the system satisfying the equation* (43) *controlled with the control $u := \bar{u}$. Recall also that in both formulas $P(\theta)$, $\theta \in \Theta$, is the unique nonnegative symmetric solution of* (46) *which exists thanks to (A1) and (A2) by Theorem 37.*

Let there exist nonnegative continuous functions $g_q, h_q \colon \mathbb{R}_+ \to \mathbb{R}_+$ such that $g_q(t) \to 0$ as $t \to \infty$ and h_q is bounded satisfying for $q = 1, 2$

$$\mathbb{E}\left(\left\|P(\theta_0) - P(\hat{\theta}(t))\right\|\left|\hat{X}(t)\right|\right)^{2q} \leq g_q(t)\,\mathbb{E}\left|\hat{X}(t)\right|^{2q} + h_q(t) \quad \text{for all } t \geq 0. \tag{92}$$

Then

$$J(y,\hat{u}) = J(y,\bar{u}).$$

and $\hat{u}$ is the optimal control, in the sense of Definition 24, for the ergodic control problem (43) *with the cost functional* (39) *in the class of controls $u \in \mathcal{U}$ satisfying*

$$\lim_{t\to\infty} \frac{1}{t}\mathbb{E}\langle P(\theta_0)X^{u,x}(t), X^{u,x}(t)\rangle = 0. \tag{93}$$

To prove the theorem it is sufficient to show

$$J(x,\bar{u}) - J(x,\hat{u}) = 0 \tag{94}$$

and

$$\lim_{t\to\infty} \frac{1}{t}\mathbb{E}\langle P(\theta_0)\hat{X}(t), \hat{X}(t)\rangle = 0. \tag{95}$$

Let us first calculate

$$
\begin{aligned}
J(x,\bar{u}) - J(x,\hat{u}) &= \limsup_{T\to\infty} \frac{1}{T}\mathbb{E}\int_0^T \left(\langle Q\bar{X}(t),\bar{X}(t)\rangle + \langle R\bar{u}(t),\bar{u}(t)\rangle\right) \\
&\qquad - \Big(\langle Q\hat{X}(t),\hat{X}(t)\rangle + \langle R\hat{u}(t),\hat{u}(t)\rangle\Big)\,dt \\
&= \limsup_{T\to\infty} \frac{1}{T}\mathbb{E}\int_0^T \Big(\Big\langle Q(\bar{X}(t)-\hat{X}(t)),\bar{X}(t)\Big\rangle + \Big\langle Q\hat{X}(t),\bar{X}(t)-\hat{X}(t)\Big\rangle\Big) \\
&\qquad + \Big(\Big\langle R(\bar{u}(t)-\hat{u}(t)),\bar{u}(t)\Big\rangle + \Big\langle R\hat{u}(t),\bar{u}(t)-\hat{u}(t)\Big\rangle\Big)\,dt \\
&= \limsup_{T\to\infty} \frac{1}{T}\mathbb{E}\int_0^T \Big\langle Q(\bar{X}(t)-\hat{X}(t)),\bar{X}(t)+\hat{X}(t)\Big\rangle \\
&\qquad + \Big\langle R(\bar{u}(t)-\hat{u}(t)),\bar{u}(t)+\hat{u}(t)\Big\rangle\,dt. \quad (96)
\end{aligned}
$$

where in the last equality, we used the symmetry of the matrices Q and R and the fact that the scalar products are in $\mathbb{R}^m$ and hence are symmetric. We denote for $t \geq 0$

$$
\begin{aligned}
A_1(t) &:= \Big\langle Q(\bar{X}(t)-\hat{X}(t)),\bar{X}(t)+\hat{X}(t)\Big\rangle, \\
A_2(t) &:= \Big\langle R(\bar{u}(t)-\hat{u}(t)),\bar{u}(t)+\hat{u}(t)\Big\rangle.
\end{aligned} \quad (97)
$$

Realize first that the integral in (96) can be calculated path-wise as a Lebesgue integral, an ordinary Fubini theorem is in force. Using the Jensen inequality and the ordinary Fubini theorem, we can write further

$$
\begin{aligned}
|J(y,\bar{u}) - J(y,\hat{u})| &\leq \limsup_{T\to\infty} \frac{1}{T}\mathbb{E}\int_0^T |A_1(t)+A_2(t)|\,dt \\
&\leq \limsup_{T\to\infty} \frac{1}{T}\int_0^T \mathbb{E}|A_1(t)| + \mathbb{E}|A_2(t)|\,dt. \quad (98)
\end{aligned}
$$

The proof of the main theorem follows. It rests on auxiliary results proved further in the thesis.

Proof of Theorem 56. We prove (94). Lemma 66 shows that $\mathbb{E}|A_1(t)| \to 0$ as $t \to \infty$. Lemma 68 shows that $\mathbb{E}|A_2(t)| \to 0$ as $t \to \infty$. Using Lemma 1, we conclude by (98) that (94) is satisfied.

Only (95) is left to prove now. Since by Proposition 54 there is a positive constant E_1 such that $\sup_{t\geq 0}\mathbb{E}\big|\hat{X}(t)\big|^2 \leq E_1$ we may write

$$
\frac{1}{t}\mathbb{E}\Big\langle P(\theta_0)\hat{X}(t),\hat{X}(t)\Big\rangle \leq \frac{1}{t}\|P(\theta_0)\|\,\mathbb{E}\Big|\hat{X}(t)\Big|^2 \to 0 \quad \text{as } t \to \infty.
$$

The main theorem is thus proved. □

In what follows, we strive to prove that $|A_1(t)|$ as well as $|A_2(t)|$ vanish in the mean as $t \to \infty$. We ascertain these essential preconditions for the proof of Theorem 56 to work in lemmas that follow. First, we will work towards breaking down what is to be proved into small pieces.

For $t \geq 0$ denote

$$\begin{aligned} b_1(t) &:= P(\hat{\theta}(t))\hat{X}(t) - P(\theta_0)\bar{X}(t). \\ b_2(t) &:= V(t) - W(t;\theta_0). \end{aligned} \tag{99}$$

Observe that then

$$\bar{u}(t) - \hat{u}(t) = R^{-1}G^*(b_1(t) + b_2(t)) \quad \text{for all } t \geq 0. \tag{100}$$

Next, let us denote for $t \geq 0$

$$\begin{aligned} b_{11}(t) &:= P(\hat{\theta}(t))(\hat{X}(t) - \bar{X}(t)), \\ b_{12}(t) &:= (P(\hat{\theta}(t)) - P(\theta_0))\bar{X}(t) \end{aligned} \tag{101}$$

and calculate for $t \geq 0$

$$\begin{aligned} b_1(t) &= P(\hat{\theta}(t))\hat{X}(t) - P(\theta_0)\bar{X}(t) \\ &\qquad = P(\hat{\theta}(t))(\hat{X}(t) - \bar{X}(t)) + (P(\hat{\theta}(t)) - P(\theta_0))\bar{X}(t) = b_{11}(t) + b_{12}(t). \end{aligned}$$

The following Lemma provides an estimate of the limit behavior of $b_{12}(t)$.

Lemma 57. *We have that* $\mathbb{E}|b_{12}(t)|^2 \to 0$ *as* $t \to \infty$.

Proof. Using the Hölder inequality stated in Theorem 2. By Corollary 40, we have $\mathbb{E}\|P(\hat{\theta}(t)) - P(\theta_0)\|^4 \to 0$ as $t \to \infty$. We also have $\mathbb{E}\left|\bar{X}(t)\right|^4 \leq C_2$ for all $t \geq 0$ for some positive constant C_2 by Proposition 42. We may thus write

$$\begin{aligned} \mathbb{E}|b_{12}(t)|^2 = \mathbb{E}\left|(P(\hat{\theta}(t)) - P(\theta_0))\bar{X}(t)\right|^2 &\leq \left(\mathbb{E}\left\|P(\hat{\theta}(t)) - P(\theta_0)\right\|^4 \mathbb{E}\left|\bar{X}(t)\right|^4\right)^{1/2} \\ &\leq C_2\, \mathbb{E}\left\|P(\hat{\theta}(t)) - P(\theta_0)\right\|^4 \to 0 \text{ as } t \to \infty. \end{aligned}$$

□

We are now ready to investigate how the difference $\bar{X}(t) - \hat{X}(t)$ which arises in $A_1(t)$ as well as, through $\bar{u}(t) - \hat{u}(t)$, in $A_2(t)$ behaves for $t \to \infty$. Once done, we will be able to establish rather easily the proofs of Lemmas 66 and 68.

3.1 Controlled system trajectory convergence

The aim of this section is to prove Lemma 58. It estimates the mean-square difference of the trajectory under Oracle's optimal control and the trajectory under the feasible on-line control at infinity.

Lemma 58. *Recall that* $\bar{X}$ *is the process satisfying equation* (59) *and* $\hat{X}$ *is the process satisfying equation* (71). *We have that*

$$\lim_{t \to \infty} \mathbb{E}\left|\bar{X}(t) - \hat{X}(t)\right|^2 = 0.$$

Denote

$$Y(t) := \hat{X}(t) - \bar{X}(t) \qquad \text{for } t \geq 0. \tag{102}$$

Our goal is to show that

$$\lim_{t\to\infty} \mathbb{E}|Y(t)|^2 = 0.$$

We outline the structure of the proof that follows. We proceed by breaking Y into parts for which we prove the convergence separately. We obtain that

$$dY(t) = [A_P(\theta_0)Y(t) - c_1(t) - c_2(t)]dt \quad t > 0, \quad Y(0) = 0$$

for

$$\begin{aligned} c_1(t) &:= GR^{-1}G^*[P(\hat{\theta}(t)) - P(\theta_0)]\hat{X}(t), \\ c_2(t) &:= GR^{-1}G^*[V(t) - W(t;\theta_0)], \end{aligned} \tag{103}$$

and $A_P = A_P(\theta)$ which was defined in (47) as $A_P(\theta) := A(\theta_0) - GR^{-1}G^*P(\theta_0)$. Solving the equation pathwise and estimating the absolute value of the solution we further obtain in Lemma 59 that

$$\mathbb{E}|Y(t)|^2 \leq 2\frac{M}{K}\int_0^t Me^{-K(t-r)}(\mathbb{E}|c_1(r)|^2 + \mathbb{E}|c_2(r)|^2)dr.$$

We then turn to proving that c_1 and c_2 vanish in mean-square as $t \to \infty$. To prove this for c_2 turns out to be rather technical and requires multiple technical lemmas. The convergence of c_2 in mean-square is finally proved in Lemma 65. Having proved that, it is sufficient to use L'Hôpital rule and obtain that $\mathbb{E}|Y(t)| \to 0$ as $t \to \infty$.

We start by noting that $\bar{X}$ and $\hat{X}$ are well defined since they are solutions of stochastic differential equations with coefficients satisfying proper requirements of Theorem 33. For the process $\bar{X}$ this follows from Duncan et al. (2015) and for $\hat{X}$ this follows from Proposition 52

Calculate symbolically, using the equations which $\bar{X}$ and $\hat{X}$ satisfy, namely (61) and (73). We write differentials only, omitting the integral signs to enhance readability. For every $t \geq 0$ it holds a.s.

$$\begin{aligned} dY(t) &= d\hat{X}(t) - d\bar{X}(t) \\ &= A(\theta_0)(\hat{X}(t) - \bar{X}(t))\,dt - GR^{-1}G^*[P(\hat{\theta}(t))\,\hat{X}(t) - P(\theta_0)\,\bar{X}(t)]\,dt \\ &\qquad - GR^{-1}G^*(V(t) - W(t;\theta_0))\,dt \\ &= A(\theta_0)[\hat{X}(t) - \bar{X}(t)]\,dt \\ &\qquad - GR^{-1}G^*\left\{P(\theta_0)[\hat{X}(t) - \bar{X}(t)] + [P(\hat{\theta}(t)) - P(\theta_0)]\hat{X}(t)\right\}dt \\ &\qquad - GR^{-1}G^*[V(t) - W(t;\theta_0)]\,dt \\ &= [A(\theta_0) - GR^{-1}G^*P(\theta_0)][\hat{X}(t) - \bar{X}(t)]\,dt \\ &\quad - GR^{-1}G^*\left\{[P(\hat{\theta}(t)) - P(\theta_0)]\hat{X}(t) + [V(t) - W(t;\theta_0)]\right\}dt. \end{aligned} \tag{104}$$

Note that the diffusion term $\sigma\,dB(t)$ is present and equal in both differentials, in $d\hat{X}(t)$ and in $d\bar{X}(t)$, and hence cancels.

Note that for every $t \geq 0$, A_P is a deterministic stable matrix in $\mathbb{R}^{m\times m}$ whereas $c_1(t)$ and $c_2(t)$ are random processes with values in $\mathbb{R}^m$ that are continuous by

Proposition 52 and the fact that $P(\hat{\theta}(t))$ and $W(t;\theta_0)$ are continuous in t. In the notation introduced in (102) and (103), the formula obtained in (104) can be expressed as

$$dY(t) = [A_P Y(t) - c_1(t) - c_2(t)]dt \quad t > 0, \quad Y(0) = 0, \tag{105}$$

It will be convenient to reformulate the fact in the language of ordinary differential equations. We will use this approach to prove the following lemma.

Lemma 59. *There exist positive constant M and K so that for every $t \geq 0$ we have*

$$\mathbb{E}|Y(t)|^2 \leq 2\frac{M}{K}\int_0^t Me^{-K(t-r)}(\mathbb{E}|c_1(r)|^2 + \mathbb{E}|c_2(r)|^2)dr.$$

Proof. As the first step, we will study a particular linear ordinary differential equation with random coefficients. Observe, that A_P is a deterministic matrix which simplifies theoretical considerations considerably. Study the equation

$$\dot{Z}(t) = [A_P Z(t) - c_1(t) - c_2(t)] \quad t > 0, \quad Z(0) = 0, \tag{106}$$

where the unknown is a random process $Z\colon \mathbb{R}_+ \times \Omega \to \mathbb{R}^m$ and the derivative is understood pathwise. A process $Z\colon \mathbb{R}_+ \times \Omega \to \mathbb{R}^m$ is a solution of (106) if $Z(\cdot,\omega)$ is a solution of the ordinary differential equation

$$\dot{z}(t) = [A_P z(t) - c_1(t,\omega) - c_2(t,\omega)] \quad t > 0, \quad z(0) = 0,$$

for every $\omega \in \Omega$ in the sense of Definition 7. According to (105) the process Y is a solution to the differential equation with random coefficients (106).

Let $L(t,s)$ be the flow map of A_P and Ω' be the set of full probability where both processes c_1 and c_2 are continuous. By Theorem 10 applied for every $\omega \in \Omega'$ we obtain that

$$\tilde{Y}(t,\omega) = -\int_0^t L(t,r)(c_1(r,\omega) + c_2(r,\omega))\,dr \quad \text{for all } t \geq 0 \text{ and all } \omega \in \Omega'.$$

Since the equation (106) is linear, by Corollary 8 its solutions are global and unique. Since the process Y solves the equation as well as $\tilde{Y}$, we obtain that there exists a version of Y for which $Y = \tilde{Y}$, or put differently there exists Ω'' of full probability so that

$$Y(t,\omega) = -\int_0^t L(t,r)(c_1(r,\omega) + c_2(r,\omega))\,dr \quad \text{for all } t \geq 0 \text{ and } \omega \in \Omega'' \cap \Omega'. \tag{107}$$

The rest of the proof is done pathwise, we omit the randomness parameter ω for clarity. All formulas are understood to hold for all $\omega \in \Omega' \cap \Omega''$.

Using triangle inequality and properties of the Lebesgue integral on (107) we obtain

$$|Y(t)| \leq \int_0^t \|L(t,r)\|(|c_1(r)| + |c_2(r)|)\,dr \quad \text{for all } t \geq 0.$$

By the detectability and stabilizability assumptions and Theorem 37 the flow map L is stable, i.e. there exist positive constants M and K such that they satisfy

$$\|L(t,s)\| \leq Me^{-K(t-s)} \quad \text{for all } t \geq s \geq 0.$$

Hence using the Hölder inequality we can write

$$\begin{aligned}\mathbb{E}|Y(t)|^2 &\leq \int_0^t Me^{-K(t-r)}dr \int_0^t Me^{-K(t-r)}\mathbb{E}(|c_1(r)|+|c_2(r)|)^2\,dr\\ &\leq 2\frac{M}{K}\int_0^t Me^{-K(t-r)}(\mathbb{E}|c_1(r)|^2+\mathbb{E}|c_2(r)|^2)\,dr\end{aligned}$$

for all $t \geq 0$. We have that $\Omega' \cap \Omega''$ is of full probability since $\mathbb{P}(\Omega \setminus (\Omega' \cap \Omega'')) \leq \mathbb{P}(\Omega \setminus \Omega') + \mathbb{P}(\Omega \setminus \Omega'') = 0$, by the fact that Ω' and Ω'' are of full probability which concludes the proof. □

It is now easy to see that the proof of Lemma 58 will follow if we prove that c_1 and c_2 vanish at infinity in mean square. This is what we do next.

Lemma 60. *We have that* $\mathbb{E}|c_1(t)|^2 \to 0$ *as* $t \to \infty$.

Proof. By Proposition 54 we have $\sup_{t\geq 0}\mathbb{E}|\hat{X}(t)|^4 \leq E_2$ for some positive constant E_2. By Corollary 40, $\mathbb{E}\|P(\hat{\theta}(t)) - P(\theta_0)\|^4 \to 0$ as $t \to \infty$. Using the Hölder inequality stated in Theorem 2 we may write

$$\begin{aligned}\mathbb{E}\Big|(P(\hat{\theta}(t)) - P(\theta_0))\hat{X}(t)\Big|^2 &\leq \Big(\mathbb{E}\Big\|P(\hat{\theta}(t)) - P(\theta_0)\Big\|^4\mathbb{E}\Big|\hat{X}(t)\Big|^4\Big)^{1/2}\\ &\leq D_2\,\mathbb{E}\Big\|P(\hat{\theta}(t)) - P(\theta_0)\Big\|^4 \to 0 \quad \text{as } t \to \infty.\end{aligned}$$

Combining these facts and realizing that $GR^{-1}G^*$ is a deterministic matrix, we obtain

$$\mathbb{E}|c_1(t)|^2 \leq D_2\big\|GR^{-1}G^*\big\|\mathbb{E}\Big\|P(\hat{\theta}(t)) - P(\theta_0)\Big\|^4 \to 0 \quad \text{as } t \to \infty,$$

which concludes the proof. □

Proving that c_2 vanishes in mean square at infinity requires some fine estimates. It will be convenient to study fist the properties of the difference of the processes $V(t)$ and $W(t;\theta_0)$ as $t \to \infty$. Lemma 65 provides an essential result concerning a mean square behavior of the difference at infinity.

To prove it, we will need a set of rather technical lemmas. Their meaning will become clear first in the proof of Lemma 65. It may be beneficial for the reader to first look into the proof of Lemma 65 before studying the technical lemmas which follow now.

Lemma 61. *There is a positive constant C such that*

$$\int_t^\infty e^{-K(u-t)}(u-t)u^{H-\frac{1}{2}}du \leq Ct^{H-\frac{1}{2}} \quad \textit{for all } t \geq 0.$$

Proof. Realize that $t^{H-\frac{1}{2}}e^{-Kt} \to 0$ as $t \to \infty$ and calculate

$$\int_t^\infty e^{-Ku}(u-t)u^{H-\frac{1}{2}}du = \int_t^\infty e^{-Ku}u^{H+\frac{1}{2}}du - t\int_t^\infty e^{-Ku}u^{H-\frac{1}{2}}du \quad (108)$$

The first integral is a partial gamma function. The gamma function is hence the integrable majorant of the integrand and by the continuity of the integral $\int_t^\infty e^{-Ku}u^{H+\frac{1}{2}}du \to 0$ as $t \to \infty$. By the same arguments we have that $\int_t^\infty e^{-Ku}u^{H-\frac{1}{2}}du \to 0$ as $t \to \infty$. Using the L'Hôpital rule

$$\lim_{t\to\infty} \frac{\int_t^\infty e^{-Ku}u^{H-\frac{1}{2}}du}{\frac{1}{t}} = \lim_{t\to\infty} \frac{e^{-Kt}t^{H-\frac{1}{2}}}{\frac{1}{t^2}} = 0.$$

Using the L'Hôpital rule once more calculate

$$\begin{aligned}
\lim_{t\to\infty} \frac{1}{t^{H-\frac{1}{2}}} \int_t^\infty e^{-K(u-t)}(u-t)u^{H-\frac{1}{2}}du &= \lim_{t\to\infty} \frac{\int_t^\infty e^{-Ku}(u-t)u^{H-\frac{1}{2}}du}{t^{H-\frac{1}{2}}e^{-Kt}} \\
&= \lim_{t\to\infty} \frac{\int_t^\infty ue^{-Ku}u^{H-\frac{1}{2}}du - \int_t^\infty te^{-Ku}u^{H-\frac{1}{2}}du}{Ke^{-Kt}\,t^{H-\frac{1}{2}} - (H-\frac{1}{2})e^{-Kt}\,t^{H-\frac{3}{2}}} \\
&= \lim_{t\to\infty} \frac{e^{-Kt}t^{H+\frac{1}{2}} - e^{-Kt}t^{H+\frac{1}{2}} - \int_t^\infty e^{-Ku}u^{H-\frac{1}{2}}}{Ke^{-Kt}\,t^{H-\frac{1}{2}} - (H-\frac{1}{2})e^{-Kt}\,t^{H-\frac{3}{2}}} \\
&= \lim_{t\to\infty} \frac{\int_t^\infty e^{-Ku}u^{H-\frac{1}{2}}}{Ke^{-Kt}\,t^{H-\frac{1}{2}} - (H-\frac{1}{2})e^{-Kt}\,t^{H-\frac{3}{2}}} \\
&= \lim_{t\to\infty} \frac{e^{-Kt}t^{H-\frac{1}{2}}}{K^2e^{-Kt}\,t^{H-\frac{1}{2}} - 2(H-\frac{1}{2})Ke^{-Kt}\,t^{H-\frac{3}{2}} + (H-\frac{1}{2})(H-\frac{3}{2})e^{-Kt}\,t^{H-\frac{5}{2}}} \\
&= \lim_{t\to\infty} \frac{1}{K^2 - 2(H-\frac{1}{2})Kt^{-1} + (H-\frac{1}{2})(H-\frac{3}{2})t^{-2}} = \frac{1}{K^2}.
\end{aligned}$$

This means that $t^{\frac{1}{2}-H}\int_t^\infty e^{-K(u-t)}(u-t)u^{H-\frac{1}{2}}du$ is bounded in t by some positive constant C because it converges and the proof is thus concluded. □

Lemma 62. *There exists a positive constant D such that for all $t \geq s \geq 0$, $t \geq 1$ we have*

$$t^{\frac{1}{2}-H}\int_t^\infty (r-t)e^{-K(r-t)}\dot{K}_H(r,s)\,dr \leq Ds^{\frac{1}{2}-H}. \tag{109}$$

Proof. To simplify notation, all formulas in the proof hold for all $t \geq s \geq 0$, $t \geq 1$ if nothing else is specified.

Recall that by definition in (20) we have

$$\dot{K}_H(r,s) = c_H\left(\frac{r}{s}\right)^{H-\frac{1}{2}}(r-s)^{H-\frac{3}{2}} \geq 0 \quad \text{for all } r > s \geq 0.$$

Hence we can write the left hand side of (109) as

$$I(t) := \int_t^\infty t^{\frac{1}{2}-H}(r-t)e^{-K(r-t)}c_H\left(\frac{r}{s}\right)^{H-\frac{1}{2}}(r-s)^{H-\frac{3}{2}}\,dr.$$

Realizing that $t \geq s$ by assumptions and that $H - 3/2 < 0$ since $H < 1$, and hence $(r-s)^{H-3/2}$ is increasing in s we can estimate

$$I(t) \leq c_H t^{\frac{1}{2}-H}s^{\frac{1}{2}-H}\int_t^\infty e^{-K(r-t)}r^{H-\frac{1}{2}}(r-t)^{H-\frac{1}{2}}\,dr. \tag{110}$$

We split the integral on the right hand side of (110) so that we are able to estimate it from above and start estimating one of the parts: We use substitution $x := r - t$ on the integral

$$\int_t^{t+1} e^{-K(r-t)} r^{H-\frac{1}{2}} (r-t)^{H-\frac{1}{2}}\, dr$$

and realize that $x \leq 1$ and $H - 1/2 > 0$ to obtain

$$\int_t^{t+1} e^{-K(r-t)} r^{H-\frac{1}{2}} (r-t)^{H-\frac{1}{2}}\, dr \leq \int_0^1 e^{-Kx} t^{H-\frac{1}{2}} \left(\frac{x}{t}+1\right)^{H-\frac{1}{2}} dx.$$

Now since $0 < H - 1/2 < 1$ and $x/t + 1 \geq 1$ we have that

$$t^{H-\frac{1}{2}} \int_0^1 e^{-Kx} \left(\frac{x}{t}+1\right)^{H-\frac{1}{2}} dx \leq t^{H-\frac{1}{2}} \int_0^1 e^{-Kx} \left(\frac{x}{t}+1\right) dx$$

The integral converges and is bounded for all $t \geq 1$. Hence we obtained a positive constant D' so that

$$t^{H-\frac{1}{2}} \int_0^1 e^{-Kx} \left(\frac{x}{t}+1\right) dx = D'\, t^{H-\frac{1}{2}} \qquad \text{for all } t \geq 1. \tag{111}$$

We turn to estimate the second one of the two parts, i.e.

$$\int_{t+1}^{\infty} e^{-K(r-t)} r^{H-\frac{1}{2}} (r-t)^{H-\frac{1}{2}}\, dr$$

Realizing that $r - t \geq 1$ and $0 < H - \frac{1}{2} < 1$ in the integral

$$\begin{aligned}\int_{t+1}^{\infty} e^{-K(r-t)} r^{H-\frac{1}{2}} (r-t)^{H-\frac{1}{2}}\, dr &\leq \int_{t+1}^{\infty} e^{-K(r-t)} r^{H-\frac{1}{2}} (r-t)\, dr \\ &\leq \int_t^{\infty} e^{-K(r-t)} r^{H-\frac{1}{2}} (r-t)\, dr\end{aligned}$$

we can use Lemma 61 to obtain the existence of a positive constant C' so that

$$\int_{t+1}^{\infty} e^{-K(r-t)} r^{H-\frac{1}{2}} (r-t)^{H-\frac{1}{2}}\, dr \leq C' t^{H-\frac{1}{2}}. \tag{112}$$

Combining (111) and (112) and looking back at (110) we obtain that for some positive constant D we have

$$I(t) \leq c_H t^{\frac{1}{2}-H} s^{\frac{1}{2}-H} (D' + C') t^{H-\frac{1}{2}} = D s^{\frac{1}{2}-H} \qquad \text{for all } t \geq s > 0 \text{ and } t \geq 1.$$

$\square$

Lemma 63. *Fix $\delta < 3/2 - H$. Let f be a nonnegative function such that $t^\delta f(t)$ vanishes at infinity. Then for every $\varepsilon > 0$ there exists $t_0 > 0$ so that for all $r \geq t \geq t_0$ we have*

$$\int_0^t s^{\frac{1}{2}-H} f(s) \dot{K}_H(r,s)\, ds \leq c_H r^{H-\frac{1}{2}} t^{\frac{1}{2}-H-\delta} \varepsilon. \tag{113}$$

Proof. All equalities hold for $r \geq t \geq 0$. By Definition 20 we first write

$$\int_0^t s^{\frac{1}{2}-H} f(s) \dot{K}_H(r,s)\, ds = c_H \int_0^t f(s) \left(\frac{r}{s^2}\right)^{H-\frac{1}{2}} (r-s)^{H-\frac{3}{2}}\, ds.$$

Realizing that $r \geq t$ in the integral and that $(r-s)^{H-3/2}$ is decreasing in r, we obtain

$$c_H\, r^{H-\frac{1}{2}} \int_0^t f(s) s^{1-2H} (r-s)^{H-\frac{3}{2}}\, ds \leq c_H\, r^{H-\frac{1}{2}} \int_0^t f(s) s^{1-2H} (t-s)^{H-\frac{3}{2}}\, ds.$$

Substituting $x := s/t$ we obtain

$$\begin{aligned} c_H\, r^{H-\frac{1}{2}} t^{\frac{1}{2}-H} &\int_0^1 f(tx)\, x^{1-2H} (1-x)^{H-\frac{3}{2}}\, dx \\ &= c_H\, r^{H-\frac{1}{2}} t^{\frac{1}{2}-H-\delta} \int_0^1 (tx)^\delta f(tx)\, x^{1-2H-\delta} (1-x)^{H-\frac{3}{2}}\, dx \end{aligned}$$

Since the integral $\int_0^1 x^{1-2H-\delta}(1-x)^{H-\frac{3}{2}}$ has a value of a beta function, which is assured by the assumption $H > 1/2$ and $\delta < 3/2 - H$ and $(tx)^\delta f(tx)$ converges by assumptions, and hence is bounded, we can use the dominated convergence theorem to obtain that for every $\varepsilon > 0$ there exists $t_0 > 0$ so that for all $t \geq t_0$ we have

$$\int_0^1 (tx)^\delta f(tx)\, x^{1-2H-\delta} (1-x)^{H-\frac{3}{2}}\, dx < \varepsilon.$$

Hence we have

$$c_H\, r^{H-\frac{1}{2}} t^{\frac{1}{2}-H} \int_0^1 f(tx)\, x^{1-2H} (1-x)^{H-\frac{3}{2}}\, dx \leq c_H r^{H-\frac{1}{2}} t^{\frac{1}{2}-H-\delta} \varepsilon \quad \text{for all } r \geq t \geq t_0.$$

□

Lemma 64. *Let f be a nonnegative function vanishing at infinity. Then*

$$\int_0^t f(s) \left\{ \int_t^\infty e^{-K(r-t)} (r-t) \dot{K}_H(r,s)\, dr \right\}^2 ds \to 0 \quad \textit{as } t \to \infty. \tag{114}$$

Proof. Let $I(t)$ for $t \geq 0$ denote the integral in (114). Rewrite the integral using the Fubini theorem as

$$\int_t^\infty e^{-K(v-t)} (v-t) \int_0^t f(s) \dot{K}_H(v,s) \left(\int_t^\infty e^{-K(u-t)} (u-t) \dot{K}_H(u,s)\, du \right) ds\, dv$$

for all $t \geq 0$.

By Lemma 62 we obtain

$$I(t) \leq t^{H-\frac{1}{2}} \int_t^\infty e^{-K(v-t)} (v-t) \int_0^t f(s) \dot{K}_H(u,s) C s^{\frac{1}{2}-H}\, ds\, dv$$

for all $t \geq 1$.

By Lemma 63 we obtain that for every $\varepsilon > 0$ there is a $t_0' > 0$ so that for all $t \geq t_0 := \max(t_0', 1)$ we may write

$$I(t) \leq t^{\frac{1}{2}-H} \int_t^\infty e^{-K(v-t)}(v-t)FCu^{H-\frac{1}{2}}\varepsilon\, dv$$

for all $t \geq 1$.

By Lemma 61 there exists a constant C' so that $I(t) \leq C'\varepsilon$ for all $t \geq t_0$. Hence $I(t) \to 0$ as $t \to \infty$ and the proof is concluded. □

Lemma 65. *Let $(W(t;\theta),\ t \geq 0)$ be the process defined in (63), i.e.*

$$W(t;\theta) := \mathbb{E}\left[\int_t^\infty \Psi(s-t;\theta)P(\theta)\sigma\, dB(s) \,\middle|\, \mathcal{F}_t\right] \qquad \text{for } t \geq 0 \text{ and } \theta \in \Theta \quad (115)$$

and $(V(t),\ t \geq 0)$ be the process defined in (66) by coordinates as

$$V_i(t) := \sum_{j=1}^m \sum_{k=1}^m \sum_{l=1}^n \sigma_{kl} \int_0^t v_{ijk}(s,t)\, dN_l(s) \qquad \text{for } t \geq 0 \quad (116)$$

where

$$v_{ijk}(s,t) := \int_t^\infty \Psi_{ij}(r-t;\hat{\theta}(s))P_{jk}(\hat{\theta}(s))\dot{K}_H(r,s)\, dr.$$

Recall that $b_2(t) := V(t) - W(t;\theta_0)$ for $t \geq 0$. We have that $\mathbb{E}|b_2(t)|^2 \to 0$ as $t \to \infty$.

Proof. By Proposition 48 we have the representation

$$W_i(t;\theta_0) = \sum_{j=1}^m \sum_{k=1}^m \sum_{l=1}^n \sigma_{kl} \int_0^t w_{ijk}(s,t)\, dN_l(s) \qquad \text{for } t \geq 0 \text{ a.s.} \quad (117)$$

where

$$w_{ijk}(s,t) := \int_t^\infty \Psi_{ij}(r-t;\theta_0)P_{jk}(\theta_0)\dot{K}_H(r,s)\, dr.$$

In the whole proof the numbers i, j, k and l are reserved for coordinate indexes and are quantified in all formulas, if they are left free, to take on arbitrary values $i, j, k = 1, \ldots, m$ and $l = 1, \ldots, n$.

Denote

$$f_{ijk}(r;\theta) := \Psi_{ij}(r;\theta)P_{jk}(\theta) - \Psi_{ij}(r;\theta_0)P_{jk}(\theta_0) \quad r \geq 0 \text{ and } \theta \in \Theta.$$

Using linearity of the stochastic integral, expectation and matrix multiplication we obtain for all $t \geq 0$

$$V_i(t) - W_i(t;\theta_0) = \sum_{j=1}^m \sum_{k=1}^m \sum_{l=1}^n \sigma_{kl} \int_0^t \int_t^\infty f_{ijk}(r-t;\hat{\theta}(s))\dot{K}_H(r,s)\, dr\, dN_l(s)$$

Using Remark 3 we may estimate for all $t \geq 0$

$$\mathbb{E}|V(t) - W(t;\theta_0)|^2 \leq m^2 \sum_{i=1}^m \mathbb{E}|V_i(t) - W_i(t;\theta_0)|^2 \leq m^6 n^2 \sum_{i=1}^m \sum_{j=1}^m \sum_{k=1}^m \sum_{l=1}^n \sigma_{kl}$$
$$\mathbb{E}\left|\int_0^t \int_t^\infty f_{ijk}(r-t;\hat{\theta}(s))\dot{K}_H(r,s)\, dr\, dN_l(s)\right|^2. \quad (118)$$

Let us now decompose f_{ijk} so that for all $r \geq 0$ and $\theta \in \Theta$

$$f_{ijk}(r;\theta) = [\Psi_{ij}(r;\theta) - \Psi_{ij}(r;\theta_0)]\, P_{jk}(\theta_0) \;-\Psi_{ij}(r;\theta)\,[P_{jk}(\theta_0) - P_{jk}(\theta)]\,.$$

This way, using linearity again we arrive at

$$\int_0^t \int_t^\infty f_{ijk}(r-t;\hat{\theta}(s))\dot{K}_H(r,s)\,dr\,dN_l(s) = I_{ijkl}(t) - J_{ijkl}(t) \tag{119}$$

for all $t \geq 0$ where

$$I_{ijkl}(t) := \int_0^t \int_t^\infty \Big(\Psi_{ij}(r-t;\hat{\theta}(s)) - \Psi_{ij}(r-t;\theta_0)\Big)\, P_{jk}(\theta_0)\dot{K}(r,s)\,dr\,dN_l(s)$$

$$J_{ijkl}(t) := \int_0^t \int_t^\infty \Psi_{ij}(r-t;\hat{\theta}(s))\,\Big(P_{jk}(\theta_0) - P_{jk}(\hat{\theta}(s))\Big)\,\dot{K}(r,s)\,dr\,dN_l(s)$$

Using Remark 3 we may write

$$\mathbb{E}|I_{ijkl}(t) - J_{ijkl}(t)|^2 \leq 2\mathbb{E}I_{ijkl}^2(t) + 2\mathbb{E}J_{ijkl}^2(t) \quad \text{for all } t \geq 0. \tag{120}$$

First concentrate on I_{ijkl}. By the Itô isometry and realizing that $\dot{K}_H(r,s)$ is nonnegative for $r > s \geq 0$, cf. (20), we obtain for all $t \geq 0$

$$\begin{aligned}\mathbb{E}\,I_{ijkl}^2(t) &= \mathbb{E}\int_0^t \left\{\int_t^\infty \Big(\Psi_{ij}(r-t;\hat{\theta}(s)) - \Psi_{ij}(r-t;\theta_0)\Big)\, P_{jk}(\theta_0)\dot{K}(r,s)\,dr\right\}^2 ds \\ &\leq \|P(\theta_0)\|\;\mathbb{E}\int_0^t \left\{\int_t^\infty \Big\|\Psi(r-t;\hat{\theta}(s)) - \Psi(r-t;\theta_0)\Big\|\dot{K}(r,s)\,dr\right\}^2 ds.\end{aligned}$$

Denote $f(t) := \mathbb{E}\Big\|A_P^*(\hat{\theta}(t)) - A_P^*(\theta_0)\Big\|^2$, $t \geq 0$. By Proposition 39, $A_P^*(\theta)$ is stable uniformly in $\theta \in \Theta$ so that there exist positive constants K and M so that $\|\exp\{A_P^*(\theta)t\}\| \leq Me^{-Kt}$ for all $t \geq 0$ and $\theta \in \Theta$. We can use the Fubini theorem and apply Proposition 17 to continue as

$$\mathbb{E}\,I_{ijkl}^2(t) \leq M^2\|P(\theta_0)\| \int_0^t f(s)\left\{\int_t^\infty (r-t)e^{-K(r-t)}\dot{K}(r,s)\,dr\right\}^2 ds$$

for all $t \geq 0$. Thanks to stabilizability and detectability assumptions we may use Corollary 40 to obtain $f(t) \to 0$ as $t \to \infty$. The assumption of Lemma 64 is thus satisfied and we may conclude that

$$\mathbb{E}\,I_{ijkl}^2(t) \to 0 \text{ as } t \to \infty.$$

We proceed by establishing an estimate for $\mathbb{E}\,J_{ijkl}^2(t)$. Similarly as for I, using the Itô isometry and realizing that $\dot{K}_H(r,s)$ is nonnegative for $r > s \geq 0$, cf. (20), we obtain for all $t \geq 0$

$$\mathbb{E}\,J_{ijkl}^2(t) \;\;\leq\;\; \mathbb{E}\int_0^t \left\{\int_t^\infty \Big\|\Psi(r-t;\hat{\theta}(s))\,\Big(P(\theta_0) - P(\hat{\theta}(s))\Big)\Big\|\dot{K}(r,s)\,dr\right\}^2 ds$$

By Proposition 39, $A_P(\theta)$ is stable uniformly with respect to $\theta \in \Theta$, that is there exist positive constants M and K so that $\|\Psi(r;\theta)\| = \exp\{A_P(\theta)r\} \leq Me^{-Kr}$ for all $r \geq 0$ and every $\theta \in \Theta$. By assumptions imposed on the estimator, specifically that $\hat{\theta}(t) \in \Theta$ for all $t \geq 0$, we have that $\Psi(r; \hat{\theta}(t)) = \exp\{A_P(\hat{\theta}(t))r\} \leq Me^{-Kr}$ for all $r \geq t \geq 0$. Letting $f(t) = \mathbb{E}\|P(\theta_0) - P(\hat{\theta}(s))\|^2$, $t \geq 0$, we can use the Fubini theorem and write

$$\mathbb{E}\, J_{ijkl}^2(t) \leq \int_0^t f(s) \left\{ \int_t^\infty Me^{-K(r-t)} \dot{K}(r,s)\, dr \right\}^2 ds \text{ for all } t \geq 0.$$

Realizing that by Corollary 40 again $f(t) \to 0$ as $t \to \infty$, we may use Lemma 64 once more to obtain $\mathbb{E}\, J_{ijkl}^2(t) \to 0$ as $t \to \infty$.

It is only necessary now, to combine the estimates back together. By (120) we as well obtained $\mathbb{E}|I_{ijkl}(t) - J_{ijkl}(t)|^2 \to 0$ as $t \to \infty$. This means, looking at (118) and (119), that $\mathbb{E}|V(t) - W(t;\theta_0)|^2 \to 0$ as $t \to \infty$. The proof of the lemma is thus concluded. □

Proof of Lemma 58. Recall that by (103),

$$c_2(t) := GR^{-1}G^*[V(t) - W(t;\theta_0)] \quad \text{ for } t \geq 0.$$

Realizing that $\mathbb{E}|c_2(t)|^2 \leq \|GR^{-1}G^*\|^2 \mathbb{E}|V(t) - W(t;\theta_0)|^2$ since $GR^{-1}G^*$ is a deterministic matrix, we may use the just proved Lemma 65 to obtain

$$\mathbb{E}|c_2(t)|^2 \to 0 \quad \text{ as } t \to \infty.$$

Recall that we defined $c_1(t) := GR^{-1}G^*[P(\hat{\theta}(t)) - P(\theta_0)]\hat{X}(t)$ for $t \geq 0$ in (103). Using the fact that $GR^{-1}G$ is a deterministic matrix and the Hölder inequality we may write

$$\mathbb{E}|c_1(t)|^2 \leq \left\|GR^{-1}G^*\right\|^2 \mathbb{E}\left|P(\hat{\theta}(t)) - P(\theta_0)\right|^4 \mathbb{E}\left|\hat{X}(t)\right|^4 \quad \text{ for all } t \geq 0.$$

In Proposition 54 we established that there exists a positive constant E_2 so that $\mathbb{E}|\hat{X}(t)|^4 \leq E_2$ for all $t \geq 0$. Hence using Corollary 40, we obtain $\mathbb{E}|c_1(t)|^2 \to 0$ as $t \to \infty$.

Recall that in Lemma 59, we established that

$$\mathbb{E}|Y(t)|^2 \leq 2\frac{M}{K} \int_0^t Me^{-K(t-r)} (\mathbb{E}|c_1(r)|^2 + \mathbb{E}|c_2(r)|^2) dr.$$

Putting the partial results together using the L'Hôpital rule on the convolutional integral, we obtain

$$\mathbb{E}\left|\bar{X}(t) - \hat{X}(t)\right|^2 \to 0 \text{ as } t \to \infty.$$

□

Lemma 66. *We have that* $\mathbb{E}|A_1(t)| \to 0$ *as* $t \to \infty$*, where* A_1 *is defined in* (97).

Proof. We use the Schwarz inequality and the fact that $\|Q\| < \infty$ to obtain

$$\begin{aligned}\mathbb{E}|A_1(t)| &= \mathbb{E}\left|\left\langle Q(\bar{X}(t) - \hat{X}(t)), \bar{X}(t) + \hat{X}(t) \right\rangle\right| \\ &\leq \|Q\| \left(\mathbb{E}\left|\bar{X}(t) - \hat{X}(t)\right|^2 \mathbb{E}\left|\bar{X}(t) + \hat{X}(t)\right|^2 \right)^{1/2}.\end{aligned}$$

We can now use Remark 3 to get

$$\mathbb{E}|A_1(t)| \leq 2\|Q\| \left[\mathbb{E}\left|\bar{X}(t) - \hat{X}(t)\right|^2 \left(\mathbb{E}\left|\bar{X}(t)\right|^2 + \mathbb{E}\left|\hat{X}(t)\right|^2\right)\right]^{1/2}.$$

Using Proposition 42, Proposition 54 and Lemma 58 provides estimates of bounds on all terms. This concludes the proof of the lemma. □

Lemma 67. *Recall that for $t \geq 0$ we defined in* (99) *and* (101)

$$b_1(t) = P(\hat{\theta}(t))\hat{X}(t) - P(\theta_0)\bar{X}(t) = b_{11}(t) + b_{12}(t)$$

and

$$\begin{aligned} b_{11}(t) &:= P(\hat{\theta}(t))(\hat{X}(t) - \bar{X}(t)), \\ b_{12}(t) &:= (P(\hat{\theta}(t)) - P(\theta_0))\bar{X}(t). \end{aligned}$$

We have that $\mathbb{E}|b_1(t)|^2 \to 0$ as $t \to \infty$.

Proof. By Corollary 40, $\|P(\hat{\theta}(t))\|^2$ is bounded for $t \in \mathbb{R}_+$. We can thus write

$$\begin{aligned} \mathbb{E}|b_{11}(t)|^2 = \mathbb{E}\left|P(\hat{\theta}(t))(\hat{X}(t) - \bar{X}(t))\right|^2 &\leq \mathbb{E}\left\|P(\hat{\theta}(t))\right\|^2 \left|\hat{X}(t) - \bar{X}(t)\right|^2 \\ &\leq \tilde{K}\,\mathbb{E}\left|\hat{X}(t) - \bar{X}(t)\right|^2 \to 0 \text{ as } t \to \infty \quad (121) \end{aligned}$$

by Lemma 58 for the positive constant $\tilde{K} := \sup_{\theta\in\Theta} \|P(\theta)\|^2$.

Using Remark 3, we may estimate

$$\mathbb{E}|b_1(t)|^2 \leq 2\mathbb{E}|b_{11}(t)|^2 + 2\mathbb{E}|b_{12}(t)|^2$$

By the result established in Lemma 57, $\mathbb{E}|b_{12}(t)|^2 \to 0$ as $t \to \infty$. Finally, based on (121) we may conclude that $\mathbb{E}|b_1(t)|^2 \to 0$ as $t \to \infty$. □

Lemma 68. *We have that $\mathbb{E}|A_2(t)| \to 0$ as $t \to \infty$, where A_2 is defined in* (97).

Proof. We use, similarily as in the estimation of $A_1(t)$, the Schwarz inequality, the Hölder inequaly as in Remark 3 and the fact that $\|R\| < \infty$, to obtain

$$\begin{aligned} \mathbb{E}|A_2(t)| = \mathbb{E}\Big\langle R(\bar{u}(t) - \hat{u}(t)), \bar{u}(t) + \hat{u}(t)\Big\rangle& \\ \leq 2\|R\| \left[\mathbb{E}|\bar{u}(t) - \hat{u}(t)|^2 \left(\mathbb{E}|\bar{u}(t)|^2 + \mathbb{E}|\hat{u}(t)|^2\right)\right]^{1/2}&. \end{aligned}$$

We first show the boundedness of second moments of $\hat{u}(t)$ and $\bar{u}(t)$ for all $t \geq 0$. Since by definition in (60), $\bar{u}(t) = -R^{-1}G^*(P(\theta_0)\bar{X}(t) + W(t;\theta_0))$ for $t \geq 0$, we may write for all $t \geq 0$ using Remark 3

$$\mathbb{E}|\bar{u}(t)|^2 \leq 2\left\|R^{-1}G^*\right\|^2 \left(\|P(\theta_0)\|^2\mathbb{E}\left|\bar{X}(t)\right|^2 + \mathbb{E}|W(t;\theta_0)|^2\right). \quad (122)$$

The second moment of $\bar{X}(t)$ is bounded uniformly for all $t \geq 0$ based on Proposition 42. The second moment of the process $W(\cdot;\theta_0)$ is bounded uniformly for all $t \geq 0$ as well based on Proposition 43.

Similarily, since $\hat{u}(t) = -R^{-1}G^*(P(\hat{\theta}(t))\hat{X}(t) + V(t))$ we may write

$$\mathbb{E}|\hat{u}(t)|^2 \leq 2\|R^{-1}G^*\|^2 \left(\left\|P(\hat{\theta}(t))\right\|^2 \mathbb{E}\left|\hat{X}(t)\right|^2 + \mathbb{E}|V(t)|^2 \right) \qquad \text{for all } t \geq 0. \quad (123)$$

The process $\hat{X}$ is bounded uniformly in mean square based on Proposition 54. The process V is bounded uniformly in mean square based on Proposition 51. Hence the process $\hat{u}$ is uniformly bounded in mean square as well. The process $\|P(\hat{\theta}(t))\|^2$ is bounded for $t \in \mathbb{R}_+$ by Corollary 40.

Based on (122) and (123) we obtain that there exists a positive constant K such that

$$\left(\mathbb{E}|\bar{u}(t)|^2 + \mathbb{E}|\hat{u}(t)|^2\right)^{1/2} < K \qquad \text{for all } t \geq 0.$$

To prove the statement, we recall that according to (100) we can write $\bar{u}(t) - \hat{u}(t) = R^{-1}G^*(b_1(t) + b_2(t))$. Since $\|R^{-1}G^*\| < \infty$, and from Lemmas 67 and 65 it follows that $R^{-1}G^*(b_1(t) + b_2(t))$ converges to zero in mean square, we have that $\mathbb{E}|\bar{u}(t) - \hat{u}(t)|^2 \to 0$ as $t \to \infty$ and the lemma is proved. $\square$

Conclusion

We studied adaptive linear quadratic stochastic ergodic control problem with regular fractional Brownian noise, i.e. with Hurst parameter $H > 1/2$. The aim was to optimally control the system

$$
\begin{aligned}
dX(t) &= [A(\theta_0)X(t) + G(t)u(t)]\, dt + \sigma\, dB(t), \quad t > 0 \\
X(0) &= x_0
\end{aligned}
$$

with the cost functional

$$
J(x,u) \;=\; \limsup_{T\to\infty} \frac{1}{T}\mathbb{E}\left[\int_0^T \langle QX(t), X(t)\rangle + \langle Ru(t), u(t)\rangle\, dt\right].
$$

Provided a strongly consistent parameter estimator of θ_0 that is asymptotically negatively quadrant dependent of the controlled systems trajectory exists, we have proved that there is an optimal feedback control which is realizable with no knowledge of the value of the true parameter. Our results generalize those of Duncan et al. (2002) into multiple dimensions. Notable is that in contrast to Duncan et al. (2002), we did not require the estimator to have a concrete form.

To suggest further direction of research, estimators which satisfy the required criteria would have to be found. The only case we know of is the trivial case where a strongly consistent estimator is independent of the trajectory of the system. This is only realizable if there are two identical systems with independent disturbances.

It would be interesting to see whether results similar to ours also hold for the infinite-dimensional case. Since many of the results we used were actually directed to infinite dimensions it seems, that there is enough theory which could enable one to proof what we did here for the infinite-dimensional case.

References

Alòs, Elisa & David Nualart (2003). Stochastic Integration with Respect to the Fractional Brownian Motion, *Stochastics and Stochastic Reports*, Vol. 75, pp. 129–152.

Billingsley, Patrick (2008). *Probability and Measure*: John Wiley & Sons.

Braun, Martin & Martin Golubitsky (1983). *Differential Equations and Their Applications*: Springer, 4th edition.

Chojnowska-Michalik, Anna, Tyrone E Duncan, and Bozenna Pasik-Duncan (1992). Uniform operator continuity of the stationary Riccati equation in Hilbert space, *Applied Mathematics & Optimization*, Vol. 25, pp. 171–187.

Decreusefond, Laurent & Ali Üstünel (1998). Fractional Brownian Motion: Theory and Applications, in *ESAIM: proceedings*, Vol. 5, pp. 75–86.

Duncan, Tyrone E, Bohdan Maslowski, and Bozenna Pasik-Duncan (2012). Linear-Quadratic Control for Stochastic Equations in a Hilbert Space with Fractional Brownian Motions, *SIAM Journal of Control and Optimization*, Vol. 50, pp. 507–531.

——— (2015). Ergodic Control of Linear Stochastic Equations in a Hilbert Space with Fractional Brownian Motion, *Banach Center Publications*: Instytut Matematyczny Polskiej Akademii Nauk, Vol. 105, pp. 91–102.

Duncan, Tyrone E, Bozenna Pasik-Duncan, and Bohdan Maslowski (2002). Fractional Brownian Motion and Stochastic Equations in Hilbert Spaces, *Stochastics and Dynamics*, Vol. 2, pp. 225–250.

Guerra, João & David Nualart (2008). Stochastic Differential Equations Driven by Fractional Brownian Motion and Standard Brownian Motion, *Stochastic Analysis and Applications*, Vol. 26, pp. 1053–1075.

Kallenberg, Olav (2001). *Foundations of Modern Probability*: Springer Science & Business Media, 2nd edition.

Kleptsyna, Marina L, Alain Le Breton, and Michel Viot (2005). On the Infinite Time Horizon Linear-Quadratic Regulator Problem Under a Fractional Brownian Perturbation, *ESAIM: Probability and Statistics*, Vol. 9, pp. 185–205.

Kolmogorov, Andrei N (1940). Wienersche Spiralen und einige andere interessante Kurven im Hilbertschen Raum, in *CR (Dokl.) Acad. Sci. URSS*, Vol. 26, pp. 115–118.

Lancaster, Peter & Leiba Rodman (1995). *Algebraic Riccati Equations*, New York: Oxford University Press, Inc. 1st edition.

Lin, Zhengyan & Zhidong Bai (2010). *Probability Inequalities*: Springer.

Mandelbrot, Benoit B & John W Van Ness (1968). Fractional Brownian Motions, Fractional Noises and Applications, *SIAM Review*, Vol. 10, pp. 422–437.

Mishura, Yuliya S (2008). *Stochastic Calculus for Fractional Brownian Motion and Related Processes*: Springer Science & Business Media.

Mémin, Jean, Yulia Mishura, and Esko Valkeila (2001). Inequalities for the Moments of Wiener Integrals with Respect to a Fractional Brownian Motion, *Statistics & Probability Letters*, Vol. 51, No. 2, pp. 197 – 206.

Olive, Davig J (2017). *Linear Regression*: Springer International Publishing.

Pachpatte, Baburao G (1998). *Inequalities for Differential and Integral Equations*: Academic press.

Pipiras, Vladas & Murad S. Taqqu (2000). Integration Questions Related to Fractional Brownian Motion, *Probability Theory and their Applications*, Vol. 11, No. 3.

Rudin, Walter et al. (1964). *Principles of Mathematical Analysis*, Vol. 3: McGraw-Hill New York.

Sideris, Thomas C (2013). *Ordinary Differential Equations and Dynamical Systems*: Springer.

Yaglom, Akira M (2012). *Correlation Theory of Stationary and Related Random Functions: Supplementary Notes and References*: Springer Science & Business Media.

Yong, Jiongmin & Xun Yu Zhou (1999). *Stochastic controls: Hamiltonian systems and HJB equations*, Vol. 43: Springer Science & Business Media.

Zabczyk, Jerzy (1975). Remarks on the algebraic Riccati equation in Hilbert space, *Applied Mathematics & Optimization*, Vol. 2, pp. 251–258.

——— (2009). *Mathematical Control Theory: an Introduction*: Springer Science & Business Media.