# Solutions to the SEP and Position Control Problems using FBSDEs and Simulation of super-linear MV-SDEs

## Dissertation

zur Erlangung des akademischen Grades
doctor rerum naturalium (Dr. rer. nat.)

vorgelegt dem Rat der
Fakultät für Mathematik und Informatik
der Friedrich-Schiller-Universität Jena

von M.Sc. Stefan Engelhardt
geboren am 10.08.1990 in Schleiz

**Gutachter**

1. Prof. Dr. Stefan Ankirchner, Friedrich-Schiller-Universität Jena, Germany
2. Prof. Dr. Anthony Réveillac, INSA de Toulouse, France
3. Dr. Gonçalo dos Reis, University of Edinburgh, United Kingdom

**Tag der öffentlichen Verteidigung:** 24.09.2020

# Zusammenfassung


Wir lösen das Skorokhod-Einbettungsproblem für eine Klasse von stochastischen Prozessen, die eine inhomogene stochastische Differentialgleichung (SDE) der Form $\mathrm{d}A_t = \mu(t, A_t)\,\mathrm{d}t + \sigma(t, A_t)\,\mathrm{d}W_t$ erfüllen. Wir leiten hinreichende Bedingungen her, die garantieren, dass für ein gegebenes Wahrscheinlichkeitsmaß $\nu$ auf $\mathbb{R}$ eine Stoppzeit $\tau$ und eine reelle Zahl $a$ existieren, sodass die Lösung $(A_t)$ der SDE mit Startwert $A_0 = a$ die Bedingung $A_\tau \sim \nu$ erfüllt. Dabei unterscheiden wir die Fälle, in denen $(A_t)$ die Lösung der SDE im schwachen oder im starken Sinn ist. Unsere Konstruktion der einbettenden Stoppzeit basiert auf der Lösung einer voll gekoppelten Vorwärts-Rückwärts-Differentialgleichung (FBSDE). Wir benutzen die sogenannte „Method of Decoupling Fields", um zu verifizieren, dass die FBSDE eine eindeutige Lösung hat. Schließlich stellen wir einen Algorithmus vor, der unsere theoretischen Ergebnisse in die Praxis umsetzt und illustrieren ihn mit einem numerischen Experiment.

Außerdem untersuchen wir eindimensionale, zeitlich inhomogene Positionskontrollprobleme, deren Drift-Term kontrolliert wird. Wir stellen zwei hinreichende Mengen von Bedingungen bereit, sodass hier Lösungen existieren und geben jeweils eine optimale Kontrolle an. Im Spezialfall der linear-quadratischen Kontrollprobleme leiten wir die optimale Feedback-Kontrolle und die Wertfunktion für sowohl den endlichen Zeithorizont als auch für den ergodischen Fall her. Unsere Methode basiert auf Pontryagins Maximumsprinzip, das das Kontrollproblem in eine voll gekoppelte FBSDE überführt, deren Existenz und Eindeutigkeit wir mit Hilfe der „Method of Decoupling Fields" verifizieren.

Des Weiteren präsentieren wir zwei stochastische Euler-Schemata, ein explizites und ein implizites, für die Simulation von stochastischen McKean-Vlasov Differentialgleichungen (MV-SDEs) mit einer zufälligen Startbedingung und einem Drift, der stärker als linear wachsen kann. Wir zeigen ein pfadweises Resultat für das sogenannte „Propagation of Chaos" und zeigen die starke Konvergenz beider Schemata für die resultierenden Partikelsysteme. Das explizite Schema konvergiert mit der Standardrate von $1/2$ in der Schrittlänge. Für das implizite Schema verwenden wir erfolgreich Stoppzeitargumente zusammen mit einem Partikelsystem. In numerischen Tests weisen wir die theoretischen Konvergenzraten nach und illustrieren den Rechenzeitvorteil des expliziten Schemas gegenüber dem impliziten. Wir wenden unseren Algorithmus auf eine nicht Lipschitz MV-SDE aus [GPV19] und auf das Modell eines neuronalen Netzes aus [BFFT12] an und vergleichen unsere Resultate mit den dortigen. Wir weisen numerisch den Effekt der „Particle Corruption" nach, bei dem ein einziger Partikel divergiert und so das gesamte System korrumpiert.

## Abstract

We solve the Skorokhod embedding problem for a class of stochastic processes satisfying an inhomogeneous stochastic differential equation (SDE) of the form $\mathrm{d}A_t = \mu(t, A_t)\,\mathrm{d}t + \sigma(t, A_t)\,\mathrm{d}W_t$. We provide sufficient conditions guaranteeing that for a given probability measure $\nu$ on $\mathbb{R}$ there exists a bounded stopping time $\tau$ and a real $a$ such that the solution $(A_t)$ of the SDE with initial value $a$ satisfies $A_\tau \sim \nu$. We hereby distinguish the cases where $(A_t)$ is a solution of the SDE in a weak or strong sense. Our construction of embedding stopping times is based on the solution of a fully coupled forward-backward stochastic differential equation (FBSDE). We use the so-called method of decoupling fields to verify that the FBSDE has a unique solution. Finally, we sketch an algorithm for putting our theoretical construction into practice and illustrate it with a numerical experiment.

We also provide two sets of sufficient conditions for the existence of a solution to one-dimensional, time inhomogeneous position targeting problems, where the drift of the state process can be controlled and derive optimal controls. For the special case of linear-quadratic control problems we derive the optimal linear feedback control and value function, for the finite time horizon and in the ergodic version. Our method is based on Pontryagin's maximum principle transforming the control problem into a fully coupled FBSDE, whose existence and uniqueness we verify with the method of decoupling fields.

Furthermore, we present two fully probabilistic Euler schemes, one explicit and one implicit, for the simulation of McKean-Vlasov Stochastic Differential Equations (MV-SDEs) with drifts of super-linear growth and random initial condition. We provide a pathwise propagation of chaos result and show strong convergence for both schemes on the consequent particle system. The explicit scheme attains the standard $1/2$ rate in stepsize. For the implicit scheme we successfully use stopping times in combination to the particle system. Numerical tests recover the theoretical convergence rates and illustrate a computational complexity advantage of the explicit over the implicit scheme. Comparative analysis is carried out on a stylized non Lipschitz MV-SDE from [GPV19] and the neuron network model proposed in [BFFT12]. We provide numerical tests illustrating a *particle corruption* effect where one single diverging particle can "corrupt" the whole particle system. Moreover, the more particles in the system the more likely this divergence is to occur.

# Contents

# 1 Introduction

The aim of the Skorokhod Embedding Problem (SEP) is to find, for a given random process and a given probability distribution, a stopping time with nice properties such that the stopped process is distributed in the desired way. A simple example is stopping a standard Browian motion such that we obtain the standard normal distribution, for which the deterministic stopping time $\tau = 1$ is a solution. Although the first formulation of this problem by Skorokhod [Sko65] was already provided in the 1960's, most applications were only found quite recently. Besides the usage in some theoretical numerical results, there are also applications in finance and control theory. For example Hobson and Klimmek [HK12] derived model independent bounds on options and in [AKTKK19] the constrained problem gets transformed into another unconstrained one, which is easier to solve.

After the formulation by Skorokhod, there have been countless approaches to the SEP. A good survey was done by Obłój [Obł04] in 2004. Since then there have been importent advances like [AHS15], where the authors give a characterisation of all distributions which can be embedded into processes solving time homogeneous stochastic differential equations (SDEs) in integrable and in bounded time. This characterisation, for example, enabled the mentioned result in [AKTKK19].

One approach used in many recent works goes back to Bass [Bas83] and is based on a time change argument. We too use this approach. In Chapter 3 we consider a forward-backward stochastic differential equation (FBSDE), show that it has a unique solution with the method of decoupling fields from Chapter 2, and then prove that this solution can be transformed into a solution for the SEP. Finally we propose a numerical scheme for the simulation of our solution.

To the best of our knowlege, the paper [AEFR18], on which Chapter 3 is based, is the first one presenting a solution to the SEP for processes with inhomogeneous and non-deterministic coefficents. For a more detailed comparison of our results to other works, see Section 3.6.

In optimal position control the aim is to steer a process which fulfills an SDE, such that the generated costs of the process and steering are minimized. When solving such position targeting problems, the commonly used approaches trace back to Bellman and Pontryagin. The Dynamic Programming Principle, developed by Bellman, together with the Hamiltion-Jacobi-Bellman equation make use of PDE theory. On the other hand, Pontryagin's Maximum Principle states the equivalence of the control problem to the solution of a backward stochastic differential equation (BSDE) or respectively an FBSDE. Most works making use of Pontryagin's Maximum Principle either make assumptions that decouple the FBSDE making it a BSDE, or, in the linear-quadratic case, exploite some dualities, which then allow to solve the control problem. In the latter case the solution of the FBSDE is only a byproduct.

Our approach, presented in Chapter 4, is different in that we directly solve the coupled FBSDE and thereby obtain a solution to the control problem. For this we apply the relatively new method of decoupling fields (see Chapter 2). After deriving two sets of sufficient conditions for an optimal control, we turn to the special case of linear-quadratic control problems. Here we can derive some explicit formulas for the optimal control and the value function, which we show to fulfill the Hamilton-Jacobi-Bellman equation. In a final section we then consider the ergodic case of an infinite horizon.

In the last part of this thesis, which is based on [dRES18], we are concerned with the simulation of McKean Vlasov Stochastic Differential Equations (MV-SDEs) with super linear growth. MV-SDEs are SDEs in which the drift and diffusion coefficients are allowed to depend on the distribution of the process. In our case super linear growth means that the drift coefficent is not globally Lipschitz continuous. Such MV-SDEs with super linear growth appear for example in the simulation of neuronal activity (see e.g. [BFFT12], [BCC11], [BFT15]) or in biology and physics (see e.g. [DGG$^+$11], [GGM$^+$18]). For a more detailed motivation we refer to [BFFT12].

The problem for simulating MV-SDEs with super linear growth is threefold. Firstly, the super linear growth has to be dealt with. Secondly, the distribution of the process has to be approximated. And thirdly, the combination has to converge.

The super linearity poses a problem because for standard SDEs it is known that the explicit Euler scheme runs into difficulties, see [HJK11]. We confirm this for MV-SDEs with a numerical experiment. The usage of an implicit scheme as in [HMS02] is impractical, since this would require to solve a fixed point equation at every time-step, which is computationally expensive. To circumvent this problem we apply a so-called *Tamed Euler* scheme, which was developed in [HJK12] and has already been successfully used by several authors (see e.g. [CJM16], [Sab13], [FG16]) to deal with coefficients that grow super-linearly.

Although there are other techniques (see [GP18]) to approximate the distribution of the process, we use the most common one. A so-called interacting particle system consists of simulating many paths simultaneously and using the averaged sum of Dirac measures at the points of the paths as distribution. In the Lipschitz setting this system is known to converge pathwise to the true solution of the MV-SDE (see [Szn91], [Mél96]).

We refer to the Chapters 3, 4 and 5 for more detailed introductions.

# 2 The Method of Decoupling Fields

In this chapter we briefly summarize the key results of the abstract theory of decoupling fields, we rely on later. We present the SLC theory (standing for Standard Lipschitz Conditions) of Chapter 3 of [Fro15] and and the MLLC theory (standing for Markov Local Lipschitz Conditions), which is derived from SLC (also see [Fro15]).

We consider families of measurable functions $M, \Sigma, F, \xi$, more precisely,

$$\begin{aligned} M &: \Omega \times [0,T] \times \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^{m\times d} \longrightarrow \mathbb{R}^n, \\ \Sigma &: \Omega \times [0,T] \times \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^{m\times d} \longrightarrow \mathbb{R}^{n\times d}, \\ F &: \Omega \times [0,T] \times \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^{m\times d} \longrightarrow \mathbb{R}^m, \\ \xi &: \Omega \times \mathbb{R}^n \to \mathbb{R}^m, \end{aligned}$$

where $n, m, d \in \mathbb{N}$ and $T > 0$. Let further $(\Omega, \mathcal{F}, \mathbf{P})$ be a probability space with a $d$-dimensional Brownian motion $(W_t)_{t\in[0,T]}$ and denote by $(\mathcal{F}_t)_{t\in[0,T]}$ the augmented Brownian filtration.

For $x \in \mathbb{R}^n$ we consider the FBSDE

$$\begin{aligned} X_t &= x + \int_0^t M(s, X_s, Y_s, Z_s)\,\mathrm{d}s + \int_0^t \Sigma(s, X_s, Y_s, Z_s)\,\mathrm{d}W_s \\ Y_t &= \xi(X_T) + \int_t^T F(s, X_s, Y_s, Z_s)\,\mathrm{d}s - \int_t^T Z_s\,\mathrm{d}W_s. \end{aligned} \tag{2.1}$$

The aim is to study existence and uniqueness of solutions of the above FBSDE. It is a longstanding challenge to find conditions guaranteeing that the fully coupled FBSDE (2.1) possesses a solution. Sufficient conditions are provided e.g. in [MPY94], [PT99], [MY99], [PW99], [Del02], [MWZZ15] (see also the references therein). The method of decoupling fields, developed in [Fro15] (see also the precursor articles [MYZ12], [FI13] and [MWZZ15]), is convenient for determining whether a solution exists. A decoupling field describes the functional dependence of the backward part $Y$ on the forward component $X$. Roughly speaking, a decoupling field is a function $u$ such that for every time $s \in [0,T]$

$$u(s, X_s) = Y_s.$$

Under some nice conditions on the parameters of the FBSDE, there exists a maximal non-empty interval possessing a solution triple $(X, Y, Z)$ and a decoupling field with nice regularity properties. The method of decoupling fields consists in analyzing the dynamics of the decoupling field's gradient in order to determine whether the FBSDE has a solution on the whole considered time interval.

At first, we have to fix some notation, which we also use in the subsequent chapters. For a stochastic process $A : \Omega \times I \to \mathbb{R}^N$, where $I$ is an interval in $[0,\infty)$ and $N \in \mathbb{N}$, we introduce for $J \subset I$ the norm

$$\|A\|_{\infty,J} := \operatorname*{ess\,sup}_{(s,\omega)\in J\times\Omega} |A_s(\omega)|$$

with regard to the product measure $\lambda \times \mathbf{P}$. For a function $f : \Omega \times I \times \mathbb{R}^n \to \mathbb{R}^m$ with $n, m \in \mathbb{N}$ we likewise define for a non-empty subinterval $J \subset I$

$$\|f\|_{\infty,J} := \operatorname*{ess\,sup}_{(\omega,s,x)\in\Omega\times J\times\mathbb{R}^n} |f(\omega,s,x)|$$

with regard to the product of $\mathbf{P}$ and the Lebesgue measure. We simply write $\|A\|_\infty$ and $\|f\|_\infty$ if $J = I$.

For a measurable map $g : \Omega \times \mathbb{R}^n \to \mathbb{R}^m$, with $n, m \in \mathbb{N}$ we define

$$L_g := \inf\left\{L \geqslant 0 \,|\, |g(x) - g(x')| \leqslant L|x - x'| \text{ for all } x, x' \in \mathbb{R}^n\right\},$$

where $\inf \varnothing := \infty$. We also set $L_g := \infty$ if $g$ is not measurable. $L_g < \infty$ implies that $g$ is Lipschitz continuous. For a map $u : \Omega \times [t,T] \times \mathbb{R}^n \to \mathbb{R}^m$, with $[t,T] \subset [0,\infty)$, we define $L_{u,x} := \sup_{s\in[t,T]} L_{u(s,\cdot)}$.

We denote by $L_{\Sigma,z}$ the Lipschitz constant of $\Sigma$ with respect to the dependence on the last component $z$ (and with respect to the Frobenius norms on $\mathbb{R}^{m\times d}$ and $\mathbb{R}^{n\times d}$), by which we mean the minimum of all Lipschitz constants or $\infty$ in case that $\Sigma$ is not Lipschitz continuous in $z$. If $L_{\Sigma,z} < \infty$, we denote by $L_{\Sigma,z}^{-1} = \frac{1}{L_{\Sigma,z}}$ the value $\frac{1}{L_{\Sigma,z}}$ if $L_{\Sigma,z} > 0$ and $\infty$ otherwise.

The following two assumptions form the basis for Chapter 3 and 4. In the assumptions in those chapters we suppose that at least one of the two is fulfilled in order to apply the theory of this chapter.

**Assumption 2.1** (SLC)
*The functions $M, \Sigma, F, \xi$ satisfy Standard Lipschitz Conditions (SLC) if*

1. *$M, \Sigma, F$ are Lipschitz continuous in $(x, y, z)$ with some Lipschitz constant $L$,*
2. *$\|(|M| + |\sigma| + |F|)(\cdot,\cdot,0,0,0)\|_\infty < \infty$,*
3. *$\xi$ is measurable such that $\|\xi(\cdot,0)\|_\infty < \infty$ and $L_{\xi,x} < L_{\Sigma,z}^{-1}$.*

**Assumption 2.2** (MLLC)
*The functions $M, \Sigma, F, \xi$ fulfill Modified Local Lipschitz Conditions (MLLC) if*

1. *the functions $M, \Sigma, F$ are*
   (a) *deterministic,*
   (b) *Lipschitz continuous in $x, y, z$ on sets of the form $[0,T] \times \mathbb{R}^n \times \mathbb{R}^m \times B$, where $B \subset \mathbb{R}^{m\times d}$ is an arbitrary bounded set,*
   (c) *and fulfill $\|M(\cdot,0,0,0)\|_\infty$, $\|F(\cdot,0,0,0)\|_\infty$, $\|\Sigma(\cdot,\cdot,\cdot,0)\|_\infty$, $L_{\Sigma,z} < \infty$,*
2. *$\xi : \mathbb{R}^n \to \mathbb{R}^m$ is also deterministic and satisfies $L_{\xi,x} < L_{\Sigma,z}^{-1}$.*

In contrast to SLC, there are only deterministic mappings $M, \Sigma, F, \xi$ allowed in the MLLC theory. In this so-called Markovian case the Lipschitz continuity assumptions of Chapter 3 of [Fro15] get relaxed a bit and we still obtain local existence together with uniqueness. In the Markovian case the property

$$\text{“}Z_s = \partial_x u(s, X_s) \cdot \Sigma(s, X_s, Y_s, Z_s)\text{”},$$

which comes from the fact that $u$ will also be deterministic, gets exploited. This allows to bound $Z$ by a constant if $\Sigma$ and $\partial_x u$ are assumed to be bounded.

**Definition 2.3**
Let $M, \Sigma, F, \xi$ fulfill SLC and $t \in [0,T]$. We call a function $u : \Omega \times [t,T] \times \mathbb{R}^n \to \mathbb{R}^m$ with $u(\omega, T, \cdot) = \xi(\omega, \cdot)$ for a.a. $\omega \in \Omega$ a *decoupling field* for $M, \Sigma, F, \xi$ on $[t,T]$ if for all $t_1, t_2 \in [t,T]$ with $t_1 \leqslant t_2$ and any $\mathcal{F}_{t_1}$–measurable $X_{t_1} : \Omega \to \mathbb{R}^n$ there exist progressively measurable processes $X, Y, Z$ on $[t_1, t_2]$ such that

- $X_s = X_{t_1} + \int_{t_1}^s M(r, X_r, Y_r, Z_r)\,\mathrm{d}r + \int_{t_1}^s \Sigma(r, X_r, Y_r, Z_r)\,\mathrm{d}W_r$,
- $Y_s = Y_{t_2} + \int_s^{t_2} F(r, X_r, Y_r, Z_r)\,\mathrm{d}r - \int_s^{t_2} Z_r\,\mathrm{d}W_r$,
- $Y_s = u(s, X_s)$,

a.s. for all $s \in [t_1, t_2]$. In particular, we want all integrals to be well-defined and $X, Y, Z$ to have values in $\mathbb{R}^n$, $\mathbb{R}^m$ and $\mathbb{R}^{m \times d}$ respectively.
Furthermore, we call a function $u : \Omega \times (t,T] \times \mathbb{R}^n \to \mathbb{R}^m$ a decoupling field for $M, \Sigma, F, \xi$ on $(t,T]$ if $u$ restricted to $[t', T]$ is a decoupling field for all $t' \in (t,T]$.

**Definition 2.4**
Let $M, \Sigma, F, \xi$ fulfill MLLC and let $t \in [0,T]$. We call a deterministic function $u : [t,T] \times \mathbb{R}^n \to \mathbb{R}^m$ with $u(T, \cdot) = \xi$ a *Markovian decoupling field* for $M, \Sigma, F, \xi$ on $[t,T]$ if for all $t_1, t_2 \in [t,T]$ with $t_1 \leqslant t_2$ and any $\mathcal{F}_{t_1}$–measurable $X_{t_1} : \Omega \to \mathbb{R}^n$ there exist progressively measurable processes $X, Y, Z$ on $[t_1, t_2]$ such that

- $X_s = X_{t_1} + \int_{t_1}^s M(r, X_r, Y_r, Z_r)\,\mathrm{d}r + \int_{t_1}^s \Sigma(r, X_r, Y_r, Z_r)\,\mathrm{d}W_r$,
- $Y_s = Y_{t_2} - \int_s^{t_2} f(r, X_r, Y_r, Z_r)\,\mathrm{d}r - \int_s^{t_2} Z_r\,\mathrm{d}W_r$,
- $Y_s = u(s, X_s)$,

a.s. for all $s \in [t_1, t_2]$ and such that $\|Z\|_{\infty, [t_1, t_2]} < \infty$ holds. In particular, we want all integrals to be well-defined and $X, Y, Z$ to have values in $\mathbb{R}^n$, $\mathbb{R}^m$ and $\mathbb{R}^{m \times d}$ respectively.
In addition, we call a function $u \colon (t,T] \times \mathbb{R}^n \to \mathbb{R}^m$ a Markovian decoupling field for $M, \Sigma, F, \xi$ on $(t,T]$ if $u$ restricted to $[t', T]$ is a Markovian decoupling field for all $t' \in (t,T]$.

We refer in both cases to the stated property, that $Y_s = u(s, X_s)$ a.s., as the *decoupling condition*.

In the following we work with weak derivatives. This allows us to obtain variational differentiability (i.e. w.r.t. the initial value $x \in \mathbb{R}^n$) of the processes $X, Y, Z$ for Lipschitz (or locally Lipschitz) continuous $M, \Sigma, F, \xi$. We start by fixing notation and giving some definitions:

For $x \in \mathbb{R}^{m \times d}$ or $x \in \mathbb{R}^{n \times d}$ the expression $|x|$ denotes the Frobenius norm of the linear operator $x$, i.e. the square root of the sum of the squares of its matrix coefficients.
We denote by $S^{n-1} := \{x \in \mathbb{R}^n \,|\, |x| = 1\}$ the $(n-1)$–dimensional sphere. If $x \in \mathbb{R}^{n \times n}$, $x \in \mathbb{R}^{m \times n}$, $x \in \mathbb{R}^{m \times d \times n}$ or $x \in \mathbb{R}^{n \times d \times n}$, we define $|x|_v := |x \cdot v|$ for all $v \in S^{n-1}$, where $\cdot$ is the application of the linear operator $x$ to the vector $v$ such that $x \cdot v$ is in $\mathbb{R}^n$, $\mathbb{R}^m$, $\mathbb{R}^{m \times d}$ or $\mathbb{R}^{n \times d}$ respectively. We refer to $\sup_{v \in S^{n-1}} |x|_v$ as the operator norm of $x$.

Now, consider a mapping $X : \mathcal{M} \times \Lambda \to \mathbb{R}$, where $(\mathcal{M}, \mathcal{A}, \rho)$ is some measure space with finite measure $\rho$ and $\Lambda \subseteq \mathbb{R}^N$ is open, $N \in \mathbb{N}$. We say that $X$ is *weakly differentiable* w.r.t. the parameter $\lambda \in \Lambda$, if for almost all $\omega \in \mathcal{M}$ the mapping $X(\omega, \cdot) : \Lambda \to \mathbb{R}$ is weakly differentiable. This means that there exists a mapping $\partial_\lambda X : \mathcal{M} \times \Lambda \to \mathbb{R}^{1 \times N}$ such that

$$\int_\Lambda \varphi(\lambda) \partial_\lambda X(\omega, \lambda)\,\mathrm{d}\lambda = -\int_\Lambda X(\omega, \lambda) \partial_\lambda \varphi(\lambda)\,\mathrm{d}\lambda$$

for any real valued test function $\varphi \in C_c^\infty(\Lambda)$, for almost all $\omega \in \mathcal{M}$. In particular, $X(\omega,\cdot)$ and the *weak derivative* $\partial_\lambda X(\omega,\cdot)$ have to be locally integrable for a.a. $\omega$. This of course includes measurability w.r.t. $\lambda$ for almost every *fixed* $\omega$.

We remark that weak differentiability for vector valued mappings is defined component-wise. We refer to Section 2.1.2 of [Fro15] for more details on weak derivatives.

Note that if $L_{u,x} < \infty$ and, therefore, $u$ is Lipschitz continuous in $x$, then $u$ is weakly differentiable in $x$ (see e.g. Lemma A.3.1. of [Fro15]) and even classically differentiable almost everywhere. If not otherwise specified we refer to $\partial_x u : [t,T] \times \mathbb{R}^n \to \mathbb{R}^{m\times n}$ as the particular version of the weak derivative which coincides with the classical derivative in all points for which a classical derivative exists and is zero in all other points. See for instance the statement and proof of Lemma A.3.1. in [Fro15] for details.

We write $\mathbb{E}_{t,\infty}[X]$ for $\operatorname{ess\,sup} \mathbb{E}[X|\mathcal{F}_t]$ in the following definition:

**Definition 2.5**
Let $u$ be a decoupling field or Markovian decoupling field for $M,\Sigma,F,\xi$. We call $u$ *weakly regular* if $L_{u,x} < L_{\Sigma,z}^{-1}$ and $\sup_{s\in[t,T]} \|u(s,0)\|_\infty < \infty$.

Furthermore, we call a weakly regular $u$ *strongly regular* if for all fixed $t_1,t_2 \in [t,T]$, $t_1 \leqslant t_2$, the processes $X,Y,Z$ arising in the defining property of a decoupling field or a Markovian decoupling field, respectively, are a.e. unique for each *constant* initial value $X_{t_1} = x \in \mathbb{R}^n$ and satisfy

$$\sup_{s\in[t_1,t_2]} \mathbb{E}_{t_1,\infty}[|X_s|^2] + \sup_{s\in[t_1,t_2]} \mathbb{E}_{t_1,\infty}[|Y_s|^2] + \mathbb{E}_{t_1,\infty}\left[\int_{t_1}^{t_2} |Z_s|^2\,\mathrm{d}s\right] < \infty \quad \forall x \in \mathbb{R}^n.$$

In addition $X,Y,Z$ must be measurable as functions of $(x,s,\omega)$ such that for every $s\in[t_1,t_2]$ the mappings $X_s$ and $Y_s$ are measurable functions of $(x,\omega)$. Moreover, $X,Y,Z$ have to be weakly differentiable w.r.t. $x \in \mathbb{R}$ such that

$$\operatorname*{ess\,sup}_{x\in\mathbb{R}^n} \sup_{v\in S^{n-1}} \sup_{s\in[t_1,t_2]} \mathbb{E}_{t_1,\infty}\left[\left|\frac{\partial}{\partial x}X_s\right|_v^2\right] < \infty,$$

$$\operatorname*{ess\,sup}_{x\in\mathbb{R}^n} \sup_{v\in S^{n-1}} \sup_{s\in[t_1,t_2]} \mathbb{E}_{t_1,\infty}\left[\left|\frac{\partial}{\partial x}Y_s\right|_v^2\right] < \infty,$$

$$\operatorname*{ess\,sup}_{x\in\mathbb{R}^n} \sup_{v\in S^{n-1}} \mathbb{E}_{t_1,\infty}\left[\int_{t_1}^{t_2}\left|\frac{\partial}{\partial x}Z_s\right|_v^2 \mathrm{d}s\right] < \infty.$$

We say that a decoupling field or Markovian decoupling field $u$ on $[t,T]$ is *strongly regular* on a subinterval $[t_1,t_2] \subseteq [t,T]$ if $u$ restricted to $[t_1,t_2]$ is a strongly regular (Markovian) decoupling field for $M,\Sigma,F,u(t_2,\cdot)$.
Furthermore, we say that a decoupling field or Markovian decoupling field $u$ is

- weakly regular if $u$ restricted to $[t',T]$ is weakly regular for all $t' \in (t,T]$,
- strongly regular if $u$ restricted to $[t',T]$ is strongly regular for all $t' \in (t,T]$.

The following natural concept introduces a type of Markovian decoupling field for non-Lipschitz problems (non-Lipschitz in $z$), to which nevertheless standard Lipschitz results can be applied.

**Definition 2.6**
Let $u$ be a Markovian decoupling field for $M, \Sigma, F, \xi$. We call $u$ *controlled in* $z$ if there exists a constant $C > 0$ such that for all $t_1, t_2 \in [t, T]$, $t_1 \leqslant t_2$, and all initial values $X_{t_1}$, the corresponding processes $X, Y, Z$ from the definition of a Markovian decoupling field satisfy $|Z_s(\omega)| \leqslant C$, for almost all $(s, \omega) \in [t, T] \times \Omega$. If for a fixed triple $(t_1, t_2, X_{t_1})$ there are different choices for $X, Y, Z$, then all of them are supposed to satisfy the above control.

We say that a Markovian decoupling field $u$ on $[t, T]$ is *controlled in* $z$ on a subinterval $[t_1, t_2] \subseteq [t, T]$ if $u$ restricted to $[t_1, t_2]$ is a Markovian decoupling field for $M, \Sigma, F, u(t_2, \cdot)$ that is controlled in $z$.

Furthermore, we call a Markovian decoupling field on an interval $(s, T]$ *controlled in* $z$ if it is controlled in $z$ on every compact subinterval $[t, T] \subseteq (s, T]$ (with $C$ possibly depending on $t$).

**Definition 2.7**
Let $I_{\max} \subseteq [0, T]$ be the union of all intervals $[t, T] \subseteq [0, T]$ such that there exists a weakly regular decoupling field or a Markovian decoupling field $u$ on $[t, T]$ for $M, \Sigma, F, \xi$.

**Theorem 2.8** (Theorem 3.1.12 and Theorem 4.2.28 in [Fro15].)
*Let $M, \Sigma, F, \xi$ satisfy SLC or MLLC. Then there exists a unique weakly regular decoupling field resp. a weakly regular Markovian decoupling field $u$ on $I_{\max}$. This $u$ is also strongly regular, continuous and, if MLLC is fulfilled, controlled in $z$.*
*Furthermore, either $I_{\max} = [0, T]$ or $I_{\max} = (t_{\min}, T]$, where $0 \leqslant t_{\min} < T$.*

Theorem 2.8 is fundamental for the theory of decoupling fields. First of all, it gives the existence of a decoupling field on a non-empty interval. And secondly, it narrows the possibilities down to two cases. Either we have existence on the whole interval $I_{\max} = [0, T]$, meaning that the FBSDE has a solution, or there is some $t_{\min}$, where the Lipschitz constant of the decoupling field "explodes" (for a precise statement see Lemma 2.10 below).

The next lemma states that existence of weakly regular decoupling fields implies existence and uniqueness of classical solutions:

**Lemma 2.9** (Corollary 3.1.9 and Theorem 4.2.25 in [Fro15].)
*Let $M, \Sigma, F, \xi$ satisfy SLC or MLLC and assume that there exists a weakly regular decoupling field or resp. a weakly regular Markovian decoupling field $u$ on some interval $[t, T]$.*
*Then for any initial condition $X_t = x \in \mathbb{R}^n$ there is a unique solution $(X, Y, Z)$ of the FBSDE* (2.1) *on $[t, T]$ such that*

$$\sup_{s \in [t,T]} \mathbb{E}[|X_s|^2] + \sup_{s \in [t,T]} \mathbb{E}[|Y_s|^2] + \mathbb{E}\left[\int_t^T |Z_s|^2 \, \mathrm{d}s\right] < \infty.$$

The following result basically states that for a singularity in $t_{\min}$ to occur $\partial_x u$ has to "explode" at $t_{\min}$, as mentioned above. It is the key for showing well-posedness for particular problems via contradiction.

**Lemma 2.10** (Lemma 3.1.15 and Lemma 4.2.29 in [Fro15].)
*Let $M, \Sigma, F, \xi$ satisfy SLC or MLLC. If $I_{\max} = (t_{\min}, T]$, then*

$$\lim_{t \searrow t_{\min}} L_{u(t,\cdot)} = L_{\Sigma,z}^{-1},$$

*where $u$ is the unique weakly regular decoupling field or resp. weakly regular Markovian decoupling field from Theorem 2.8.*

For all $s \in I_{\max}$ we call

$$U_s := \partial_x u(s, X_s)$$

the *gradient process* corresponding to FBSDE (2.1), where $X$ is part of the solution $(X, Y, Z)$ of FBSDE (2.1) and $u$ is the corresponding (Markovian) decoupling field. As the following theorem states, we can use the gradient process $U$ to show that FBSDE (2.1) has a solution.

**Theorem 2.11**
*Let $M, \Sigma, F, \xi$ fulfill SLC or MLLC. If for every initial value $X_t = x \in \mathbb{R}^n$ with $t \in I_{\max}$ the gradient process $U$ fulfills $|U_t| \leqslant C < L_{\Sigma,z}^{-1}$ then $I_{\max} = [0, T]$; and for every initial value $X_0 = x \in \mathbb{R}^n$ there exists a unique solution $(X, Y, Z)$ of FBSDE* (2.1) *such that*

$$\sup_{s \in [0,T]} \mathbb{E}[|X_s|^2] + \sup_{s \in [0,T]} \mathbb{E}[|Y_s|^2] + \mathbb{E}\left[\int_0^T |Z_s|^2 \,\mathrm{d}s\right] < \infty.$$

*Proof.* By Theorem 2.8 we have that there is a non-empty interval $I_{\max} \subset [0, T]$ on which a unique and strongly regular (Markovian) decoupling field exists and either $I_{\max} = [0, T]$ or $I_{\max} = (t_{\min}, T]$ for some $t_{\min} \in [0, T)$. Lemma 2.10 states that the case $I_{\max} = (t_{\min}, T]$ can only happen if $L_{u(t,\cdot)}$ converges towards $L_{\Sigma,z}^{-1}$ for $t \searrow t_{\min}$, which we loosely call an "explosion". Thus, we want to bound $\partial_x u$ away from $L_{\Sigma,z}^{-1}$. If this is the case the "explosion" can not happen and the only remaining case, $I_{\max} = [0, T]$, has to be true.

Since for every pair $(t, x) \in I_{\max} \times \mathbb{R}^n$ of initial values we have, by the assumptions made, that $\partial_x u(t, x) = \partial_x u(t, X_t) = U_t$ is bounded by $C < L_{\Sigma,z}^{-1}$, we get that the whole function $\partial_x u$ is bounded by $\|\partial_x u\|_\infty \leqslant C < L_{\Sigma,z}^{-1}$. Hence, we obtain with Lemma 2.10 that $I_{\max} = (t_{\min}, T]$ can not hold true. Thus, by Theorem 2.8, the only other possible case is $I_{\max} = [0, T]$. Finally, with Lemma 2.9 we obtain existence and uniqueness of the solution $(X, Y, Z)$ of FBSDE (2.1). ■

By applying Theorem 2.8 and Lemma 2.10, Theorem 2.11 finally gives us sufficient conditions, which are relatively easy to verify, for a solution of FBSDE (2.1) to exist.

In Chapter 3 and Chapter 4 we make use of Theorem 2.11 in the following way: First we derive the dynamics of the gradient process $U$ by differentiating $Y_s$ and $u(s, X_s)$ with respect to the initial value $x \in \mathbb{R}^n$ of the forward component $X$. This results in $U$ being the solution of a BSDE which is quadratic in $U$ itself. Then we apply the standard BSDE theory to conclude that the gradient process is bounded. Thus, we can apply Theorem 2.11 to obtain that the considered FBSDE has a unique solution.

# 3 The Skorokhod Embedding Problem for general diffusions

Let $\nu$ be a probability measure on $\mathbb{R}$, let $\mu, \sigma : [0,\infty) \times \mathbb{R} \to \mathbb{R}$ be continuous in both arguments and let $(A_t)_{t \geqslant 0}$ be a stochastic process satisfying the inhomogeneous stochastic differential equation (SDE)

$$\mathrm{d}A_t = \mu(t, A_t)\,\mathrm{d}t + \sigma(t, A_t)\,\mathrm{d}W_t, \tag{3.1}$$

where $W$ is a Brownian motion. In this chapter we consider the Skorokhod embedding problem (SEP) for $\nu$ in $(A_t)$. More precisely, we provide sufficient conditions on $\mu$, $\sigma$ and $\nu$ guaranteeing the existence of a stopping time $\tau$ and a real number $a$ such that the solution of the SDE (3.1), in a weak or strong sense, with initial condition $A_0 = a$, satisfies $A_\tau \sim \nu$.

We solve the embedding problem by reducing it to the forward-backward stochastic differential equation (FBSDE)

$$\begin{aligned} X_s^{(1)} &= x^{(1)} + W_s \\ X_s^{(2)} &= x^{(2)} + \int_0^s \frac{Z_r^2}{\sigma^2(X_r^{(2)}, Y_r + X_r^{(3)})} dr \\ X_s^{(3)} &= x^{(3)} + \int_0^s \mu(X_r^{(2)}, Y_r + X_r^{(3)}) \frac{Z_r^2}{\sigma^2(X_r^{(2)}, Y_r + X_r^{(3)})} dr \\ Y_s &= g(X_1^{(1)}) - X_1^{(3)} - \int_s^1 Z_r\,\mathrm{d}W_r \end{aligned} \tag{3.2}$$

for $s \in [0,1]$ and $(x^{(1)}, x^{(2)}, x^{(3)}) \in \mathbb{R}^3$, where $g$ is a real function chosen such that $g(W_1) \sim \nu$. Notice that the FBSDE (3.2) is fully coupled, i.e. the second and third forward equation depend on the solution components $Y$ and $Z$ of the backward equation; and, vice versa, the backward equation depends on the forward components $X^{(1)}$ and $X^{(3)}$.

We use the method of decoupling fields to prove that, under some suitable conditions on $\mu$, $\sigma$ and $g$, the FBSDE (3.2) has a unique solution on $[0,1]$ for every initial value. By using the particular solution with initial value $(x^{(1)}, x^{(2)}, x^{(3)}) = 0$, we then construct a weak solution of the SDE (3.1) and a stopping time $\tau$ embedding $\nu$. Indeed, the second component $X^{(2)}$ of the forward part in (3.2) can be interpreted as a random time change. One can show that the time change is invertible, say with inverse clock $\gamma(t)$. Moreover, there exists a filtration $(\mathcal{G}_t)$ and a $(\mathcal{G}_t)$-Brownian motion $B$ such that, first, $X_1^{(2)}$ is a $(\mathcal{G}_t)$-stopping time and, second, under the inverse clock the solution component $Y$ together with $B$ solve the SDE (3.1) in a weak sense. In the following we refer to the tuple $\big((\mathcal{G}_t), (B_t), \tau, a\big)$ as a weak solution of the SEP. By the very construction the time changed process $Y_{\gamma(\cdot)}$ at $X_1^{(2)}$ is equal to $g(W_1)$, and hence $X_1^{(2)}$ is a stopping time embedding $\nu$ into a weak solution of (3.1).

In a further step we characterize the embedding stopping time $X_1^{(2)}$ in terms of a four dimensional Lipschitz SDE driven by the constructed Brownian motion $B$. The SDE establishes a mapping from the paths of $B$ to $X_1^{(2)}$, and hence allows to find stopping times embedding $\nu$ into *strong* solutions of the SDE (3.1), where we refer to the pair $(\tau, a)$ as a strong solution of the SEP.

In Section 3.5 we show that solving the system

$$\begin{aligned} W_s =& \int_0^s \frac{\sigma(X_r^{(2)}, Y_r + X_r^{(3)})}{Z_r} \, \mathrm{d}B_{X_r^{(2)}} \\ X_s^{(2)} =& \int_0^s \frac{Z_r^2}{\sigma^2(X_r^{(2)}, Y_r + X_r^{(3)})} \, \mathrm{d}r \\ X_s^{(3)} =& \int_0^s \mu(X_r^{(2)}, Y_r + X_r^{(3)}) \frac{Z_r^2}{\sigma^2(X_r^{(2)}, Y_r + X_r^{(3)})} \, \mathrm{d}r \\ Y_s =& g(W_1) - X_1^{(3)} - \int_s^1 Z_r \, \mathrm{d}W_r \end{aligned} \tag{3.3}$$

for all $s \in [0, 1]$ and setting $\tau := X_1^{(2)}$ also yields a strong solution. Furthermore, we propose a scheme, based on the system (3.3), to numerically simulate a solution of the SEP.

A major idea of our approach for solving the SEP is to change the time of a stochastic process that has the wanted distribution at the deterministic time 1. This idea goes back to Bass [Bas83] who solves the SEP for Brownian motion. Indeed, our approach generalizes Bass' solution method. If $\mu$ is zero and $\sigma$ constant equal to one, then the component $X^{(3)}$ of (3.2) vanishes and the solution part $Y$ of the backward equation coincides with the martingale of conditional expectations of $g(W_1)$, which is the process used by Bass. Moreover, the time change $X^{(2)}$ coincides with the quadratic variation of $Y$, the time change used in [Bas83].

The time change idea has been employed in several further articles. In [AHI08] the solution of a quadratic BSDE is time changed in order to solve the SEP for the Brownian motion with drift. The FBSDE (3.2) simplifies to the BSDE of [AHI08] if $A$ is a Brownian motion with drift. [AHS15] uses a time change argument to construct stopping times embedding a given distribution into a stochastic process solving a homogeneous SDE. In [FIP15] a fully coupled FBSDE is solved and then time changed to obtain a stopping time embedding a distribution into a Gaussian process satisfying an SDE with deterministic coefficients. [FIP15] also relies on the method for decoupling fields for proving existence of a solution of the FBSDE.

There are more recent articles that are inspired by or related to Bass' time-change approach for solving the SEP for the Brownian motion. E.g. the article [BCHK17] proves optimality of the Bass solution, among all solutions of the SEP for Brownian motion, for some minimization problems formulated in terms of associated measure-valued martingales. [DGPR17] solve the SEP for a class of Levy processes via an analytic approach and by extending Bass' time-change arguments. The process of conditional expectations of $g(X_1^{(1)})$, used by Bass, is shown in [VBHK19] to minimize a martingale transport problem.

To the best of our knowledge there do not exist any other articles than [AEFR18], on which this whole chapter is based, that consider the SEP for general inhomogeneous diffusions of the type (3.1). There are various contributions to the SEP for homogeneous diffusions. The article [PP01] classifies the distributions that can be embedded into homogeneous diffusions. The survey [Obł04] collects results on the SEP, including results for homogeneous diffusions. We remark that in the homogeneous case in which the coefficients of the SDE (3.1) do not depend on time, the FBSDE (3.2) can be decoupled. We explain this in Section 3.5 below.

This chapter is organized as follows: In Sections 3.1 and 3.2 we compute the dynamics of the decoupling field gradient process and derive some estimates allowing to conclude with via the method of decoupling fields on the existence and uniqueness of a solution to FBSDE (3.2) on the whole interval. In Section 3.3 we present the weak solution, meaning that there exists a

Brownian motion such that we can give a solution to the SEP, and in Section 3.4 we present the strong solution, where for every Brownian motion we give a solution for the SEP. Illustrative numerical results can be found in Section 3.5. Finally in Section 3.6 we revisite our main results and compare our assumptions to other existing works.

## 3.1 Gradient dynamics of the decoupling field

In this section we investigate the dynamics of the spatial gradient of the decoupling field for FBSDE (3.2), which we call its grandient process. Based on the findings of this section we derive, in the subsequent section, a uniform bound for this gradient prosess, allowing us to apply Theorem 2.11.

Let $W$ be a Brownian motion on a probability space $(\Omega, \mathcal{F}, \mathbf{P})$ and denote by $(\mathcal{F}_t)_{t \geqslant 0}$ the associated augmented Brownian filtration. Also, denote by $\nu$ the probability measure on $\mathbb{R}$, which is to be embedded, and let $F_\nu$ be the cumulative distribution function of $\nu$. We set

$$g := g_\nu := F_\nu^{-1} \circ \Phi,$$

where $\Phi$ is the cumulative distribution function of the standard normal distribution and $F_\nu^{-1}$ the right-continuous generalized inverse of $F_\nu$. In the following, for a differentiable function $f : \mathbb{R}^n \to \mathbb{R}$ we denote by $\partial_{x_i} f$ its partial derivative with respect to the $i$th coordinate.

Furthermore, let $g$, $\mu$ and $\sigma$ be differentiable, $\sigma \geqslant \varepsilon > 0$ and $g'$, $\frac{\mu}{\sigma^2}$, $\frac{\partial_t \mu}{\sigma^2}$, $\frac{\partial_a \mu}{\sigma^2}$, $\frac{\partial_t \sigma}{\sigma}$ as well as $\frac{\partial_a \sigma}{\sigma}$ be bounded. Under these conditions, which are assumed to hold true everywhere in this chapter, it is straightforward to verify that the FBSDE (3.2), which is

$$\begin{aligned} X_s^{(1)} &= x^{(1)} + W_s \\ X_s^{(2)} &= x^{(2)} + \int_0^s \frac{Z_r^2}{\sigma^2(X_r^{(2)}, Y_r + X_r^{(3)})} dr \\ X_s^{(3)} &= x^{(3)} + \int_0^s \mu(X_r^{(2)}, Y_r + X_r^{(3)}) \frac{Z_r^2}{\sigma^2(X_r^{(2)}, Y_r + X_r^{(3)})} dr \\ Y_s &= g(X_1^{(1)}) - X_1^{(3)} - \int_s^1 Z_r \, \mathrm{d}W_r \end{aligned}$$

for $s \in [0,1]$ and $(x^{(1)}, x^{(2)}, x^{(3)}) \in \mathbb{R}^3$, satisfies MLLC. Hence the theory of Chapter 2 is applicable. By Theorem 2.8 the maximal interval $I_{\max}$ contains an interval $[t, 1]$ with $t < 1$. Let $x \in \mathbb{R}^3$ and denote by $X = (X^{(1)}, X^{(2)}, X^{(3)})^\top, Z, Y$ the solution of the FBSDE (3.2) on $[t, 1]$ with initial condition $(X_t^{(1)}, X_t^{(2)}, X_t^{(3)}) = x$. Moreover, denote by $u$ the decoupling field associated to the FBSDE (3.2). From Theorem 2.8 we also know that the partial derivatives $\partial_{x_1} u$, $\partial_{x_2} u$, $\partial_{x_3} u$ and the process $Z$ are bounded on $[t, 1]$.

For shorter notation we define for all $s \in [t, 1]$

$$\begin{aligned} \sigma_s &:= \sigma(X_s^{(2)}, Y_s + X_s^{(3)}), & \mu_s &:= \mu(X_s^{(2)}, Y_s + X_s^{(3)}), \\ \sigma_{t,s} &:= \partial_t \sigma(X_s^{(2)}, Y_s + X_s^{(3)}), & \sigma_{a,s} &:= \partial_a \sigma(X_s^{(2)}, Y_s + X_s^{(3)}), \\ \mu_{t,s} &:= \partial_t \mu(X_s^{(2)}, Y_s + X_s^{(3)}), & \mu_{a,s} &:= \partial_a \mu(X_s^{(2)}, Y_s + X_s^{(3)}) \end{aligned}$$

and

$$\begin{aligned} u_s^{(1)} &:= \partial_{x_1} u(s, X_s^{(1)}, X_s^{(2)}, X_s^{(3)}), \\ u_s^{(2)} &:= \partial_{x_2} u(s, X_s^{(1)}, X_s^{(2)}, X_s^{(3)}), \\ u_s^{(3)} &:= \partial_{x_3} u(s, X_s^{(1)}, X_s^{(2)}, X_s^{(3)}). \end{aligned}$$

In the following we refer to $u^{(1)}, u^{(2)}, u^{(3)}$ as the gradient processes associated to the initial value $x$ at time $t$. The next result describes the dynamics of the gradient processes. For its derivation we first argue that the processes are Itô processes and then match the coefficients appropriately. In contrast to the approach of [FIP15], we do not explicitly compute the dynamics of the inverse of the Jacobi matrix of $X$.

**Lemma 3.1**
*Let $g$, $\mu$ and $\sigma$ be differentiable, $\sigma \geqslant \varepsilon > 0$ and $g'$, $\frac{\mu}{\sigma^2}$, $\frac{\partial_t \mu}{\sigma^2}$, $\frac{\partial_a \mu}{\sigma^2}$, $\frac{\partial_t \sigma}{\sigma}$ as well as $\frac{\partial_a \sigma}{\sigma}$ be bounded. Then the gradient processes $u^{(1)}$, $u^{(2)}$ and $u^{(3)}$ have the dynamics*

$$
\begin{aligned}
u_s^{(1)} &= g'\left(X_1^{(1)}\right) + \int_s^1 u_r^{(1)} \frac{Z_r^2}{\sigma_r^2}\left(u_r^{(3)}\left(\mu_{a,r} - 2\mu_r \frac{\sigma_{a,r}}{\sigma_r}\right) - 2u_r^{(2)}\frac{\sigma_{a,r}}{\sigma_r}\right)\mathrm{d}r - \int_s^1 \tilde{Z}_r^{(1)}\,\mathrm{d}\widetilde{W}_r \\
u_s^{(2)} &= \int_s^1 u_r^{(3)}\frac{Z_r^2}{\sigma_r^2}\left(u_r^{(2)}\mu_{a,r} + \mu_{t,r}\right) - 2\frac{Z_r^2}{\sigma_r^2}\left(\frac{\sigma_{t,r}}{\sigma_r} + u_r^{(2)}\frac{\sigma_{a,r}}{\sigma_r}\right)\left(u_r^{(2)} + u_r^{(3)}\mu_r\right)\mathrm{d}r - \int_s^1 \tilde{Z}_r^{(2)}\,\mathrm{d}\widetilde{W}_r \\
u_s^{(3)} &= -1 + \int_s^1 \left(u_r^{(3)} + 1\right)\frac{Z_r^2}{\sigma_r^2}\left(u_r^{(3)}\mu_{a,s} - 2\frac{\sigma_{a,r}}{\sigma_r}\left(u_r^{(2)} + u_r^{(3)}\mu_r\right)\right)\mathrm{d}r - \int_s^1 \tilde{Z}_r^{(3)}\,\mathrm{d}\widetilde{W}_r, \qquad (3.4)
\end{aligned}
$$

*for all $s \in [t,1]$, where $\tilde{Z}^{(1)}$, $\tilde{Z}^{(2)}$, $\tilde{Z}^{(3)}$ are locally square integrable processes. Moreover, the process*

$$\widetilde{W}_s := W_s - \int_t^s 2\frac{Z_r}{\sigma_r^2}\left(u_r^{(2)} + u_r^{(3)}\mu_r\right)\mathrm{d}r$$

*is a Brownian motion under an equivalent probability measure, and the Jacobi matrix*

$$\partial_x X_s := \begin{pmatrix} \partial_{x_1} X_s^{(1)} & \partial_{x_2} X_s^{(1)} & \partial_{x_3} X_s^{(1)} \\ \partial_{x_1} X_s^{(2)} & \partial_{x_2} X_s^{(2)} & \partial_{x_3} X_s^{(2)} \\ \partial_{x_1} X_s^{(3)} & \partial_{x_2} X_s^{(3)} & \partial_{x_3} X_s^{(3)} \end{pmatrix}$$

*is invertible for every $s \in [t,1]$ almost surely.*

*Proof.* For $x' = (x_1', x_2', x_3')^\top \in \mathbb{R}^3$, $y, z \in \mathbb{R}$ we define

$$M\left(x', y, z\right) := \begin{pmatrix} 0 \\ \frac{z^2}{\sigma^2(x_2', y + x_3')} \\ \mu\left(x_2', y + x_3'\right)\frac{z^2}{\sigma^2(x_2', y + x_3')} \end{pmatrix}, \qquad \Sigma := \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$$

and

$$\xi\left(x'\right) := g(x_1') - x_3'.$$

Then, for a starting value $x_0 \in \mathbb{R}$ at time $t$, i.e. $X_t = x_0$, we can write FBSDE (3.2) as

$$
\begin{aligned}
X_s &= x_0 + \int_t^s M\left(X_r, Y_r, Z_r\right)\mathrm{d}r + \int_t^s \Sigma\,\mathrm{d}W_r \\
Y_s &= \xi\left(X_1\right) - \int_s^1 Z_r\,\mathrm{d}W_r.
\end{aligned}
$$

Now, define a stopping time $\tau$ via

$$\tau := \inf\{s \in [t,1] \mid \det\left(\partial_{x_0} X_s\right) \leqslant 0\} \wedge 1.$$

Notice that $\tau > t$ since $\det(\partial_{x_0} X_t) = 1$ and $\partial_{x_0} X_t$ is an Itô process and in particular continuous in time (see Lemma A.2.5 and Lemma A.2.6 in [Fro15]). For all $s \in [t, \tau)$ we have that $\partial_{x_0} X_s$ is invertible with $(\partial_{x_0} X_s)^{-1}$ being an Itô process , too. By setting

$$U_s := \partial_x u\,(s, X_s) = (\partial_{x_1} u, \partial_{x_2} u, \partial_{x_3} u)\,(s, X_s)$$

which is the gradient process we get

$$\partial_{x_0} Y_s = U_s \cdot \partial_{x_0} X_s$$

for all $s \in [t, \tau)$ by applying the chain rule in Lemma A.3.1 in [Fro15] to the decoupling condition. Hence, $U_s = \partial_{x_0} Y_s \cdot (\partial_{x_0} X_s)^{-1}$ is an Itô process and thus there exist $(b_s)$ and $(\tilde{Z}_s)$ such that

$$U_s = U_1 + \int_s^\tau b_r \,\mathrm{d}r - \int_s^\tau \tilde{Z}_r \,\mathrm{d}W_r$$

for all $s \in [t, \tau)$.

For the following we also introduce for an Itô process $I_s = I_0 + \int_0^s i_r \,\mathrm{d}r + \int_0^s j_r \,\mathrm{d}W_r$ the two operators $\mathrm{D}^t$ and $\mathrm{D}^w$ defined via $(\mathrm{D}^t\, I)_s := i_s$ and $(\mathrm{D}^w\, I)_s := j_s$. Note that due to Lemma A.2.5 and Lemma A.2.6 in [Fro15] the operators $\mathrm{D}^w$ and $\mathrm{D}^t$ can be interchanged with the weak differentiation $\partial_{x_0}$. Using this notation we have

$$\begin{aligned} \partial_{x_0} Z_s &= \mathrm{D}^w\, \partial_{x_0} Y_s \\ &= \mathrm{D}^w\, (U_s \cdot \partial_{x_0} X_s) \\ &= U_s \cdot \mathrm{D}^w\, \partial_{x_0} X_s + \mathrm{D}^w\, U_s \cdot \partial_{x_0} X_s. \end{aligned}$$

Since $\mathrm{D}^w\, \partial_{x_0} X_s = 0$, we further obtain $\partial_{x_0} Z_s = \tilde{Z}_s \cdot \partial_{x_0} X_s$ and thus we get

$$\tilde{Z}_s = \partial_{x_0} Z_s \cdot (\partial_{x_0} X_s)^{-1}$$

for all $s \in [t, \tau)$. Also,

$$\begin{aligned} &\partial_{x_0} \left[ M\,(X_s, Y_s, Z_s) \right] \\ &\quad = \partial_x M\,(X_s, Y_s, Z_s)\, \partial_{x_0} X_s + \partial_y M\,(X_s, Y_s, Z_s)\, \partial_{x_0} Y_s + \partial_z M\,(X_s, Y_s, Z_s)\, \partial_{x_0} Z_s \\ &\quad = \partial_x M\,(X_s, Y_s, Z_s)\, \partial_{x_0} X_s + \partial_y M\,(X_s, Y_s, Z_s)\, U_s \partial_{x_0} X_s + \partial_z M\,(X_s, Y_s, Z_s)\, \tilde{Z}_s \partial_{x_0} X_s \end{aligned}$$

and

$$0 = \mathrm{D}^t\, \partial_{x_0} Y_s = \mathrm{D}^t (U_s \partial_{x_0} X_s) = -b_s \cdot \partial_{x_0} X_s + U_s \cdot \partial_{x_0} \left[ M\,(X_s, Y_s, Z_s) \right]$$

yielding

$$b_s = U_s \left[ \partial_x M\,(X_s, Y_s, Z_s) + \partial_y M\,(X_s, Y_s, Z_s)\, U_s + \partial_z M\,(X_s, Y_s, Z_s)\, \tilde{Z}_s \right]$$

for all $s \in [t, \tau)$ with

$$\partial_x M\,(x, y, z) = \begin{pmatrix} 0 & 0 & 0 \\ 0 & -2z^2 \frac{\partial_t \sigma(x_2, y+x_3)}{\sigma^3(x_2, y+x_3)} & \frac{\partial_t \mu(x_2, y+x_3) \cdot z^2}{\sigma^2(x_2, y+x_3)} - 2z^2 \frac{\partial_t \sigma(x_2, y+x_3)}{\sigma(x_2, y+x_3)} \frac{\mu(x_2, y+x_3)}{\sigma^2(x_2, y+x_3)} \\ 0 & -2z^2 \frac{\partial_a \sigma(x_2, y+x_3)}{\sigma^3(x_2, y+x_3)} & \frac{\partial_a \mu(x_2, y+x_3) \cdot z^2}{\sigma^2(x_2, y+x_3)} - 2z^2 \frac{\partial_a \sigma(x_2, y+x_3)}{\sigma(x_2, y+x_3)} \frac{\mu(x_2, y+x_3)}{\sigma^2(x_2, y+x_3)} \end{pmatrix}^T,$$

$$\partial_y M\,(x, y, z) = \begin{pmatrix} 0 \\ -2z^2 \frac{\partial_a \sigma(x_2, y+x_3)}{\sigma^3(x_2, y+x_3)} \\ \frac{\partial_a \mu(x_2, y+x_3) \cdot z^2}{\sigma^2(x_2, y+x_3)} - 2z^2 \frac{\partial_a \sigma(x_2, y+x_3)}{\sigma(x_2, y+x_3)} \frac{\mu(x_2, y+x_3)}{\sigma^2(x_2, y+x_3)} \end{pmatrix},$$

$$\partial_z M(x,y,z) = \begin{pmatrix} 0 \\ \frac{2z}{\sigma^2(x_2,y+x_3)} \\ 2z\frac{\mu(x_2,y+x_3)}{\sigma^2(x_2,y+x_3)} \end{pmatrix}$$

being the derivatives of $M$.

Next we turn our attention to the question whether $\partial_{x_0} X$ is invertible. We use that on the interval $[t,1]$ the processes $U$ and $Z$ as well as the functions $\frac{1}{\sigma}$, $\frac{\mu}{\sigma^2}$, $\frac{\partial_t \mu}{\sigma^2}$, $\frac{\partial_a \mu}{\sigma^2}$, $\frac{\partial_t \sigma}{\sigma}$ and $\frac{\partial_a \sigma}{\sigma}$ are bounded, giving that $\partial_x M(X_r,Y_r,Z_r)$, $\partial_y M(X_r,Y_r,Z_r)\,U_r$ and $\partial_z M(X_r,Y_r,Z_r)$ are bounded, too. Thus, there exist some bounded processes $\alpha$ and $\beta$ depending on $U$, $X$, $Y$ and $Z$, such that for every stopping time $\tilde{\tau} < \tau$, $i = 1,2,3$ and $s \in [t,1]$ the process $u^{(i)}_{\cdot\wedge\tilde{\tau}}$ has dynamics

$$u^{(i)}_{s\wedge\tilde{\tau}} = u^{(i)}_t + \int_t^s \left(\alpha^{(i)}_r + \beta^{(i)}_r \cdot \tilde{Z}^{(i)}_r\right) \mathbb{1}_{\{r<\tilde{\tau}\}}\,\mathrm{d}r + \int_t^s \tilde{Z}^{(i)}_r \mathbb{1}_{\{r<\tilde{\tau}\}}\,\mathrm{d}W_r.$$

Standard results on linear BSDEs (see e.g. Theorem A.1.11 in [Fro15]) yield, for every stopping time $\tilde{\tau} < \tau$ and $i = 1,2,3$, that $\tilde{Z}^{(i)}$ has a bounded BMO($\mathbf{P}$)-norm which is independent of $\tilde{\tau}$. Hence,

$$\mathbb{E}\left[\int_t^\tau |\tilde{Z}_r|^2\,\mathrm{d}r\right] < \infty. \tag{3.5}$$

Now observe that

$$\begin{aligned}\partial_{x_0} X_s &= \mathrm{Id} + \int_t^s \partial_{x_0}\left[M(X_r,Y_r,Z_r)\right]\mathrm{d}r \\ &= \mathrm{Id} + \int_t^s \left[\partial_x M(X_r,Y_r,Z_r) + \partial_y M(X_r,Y_r,Z_r)\,U_r + \partial_z M(X_r,Y_r,Z_r)\,\tilde{Z}_r\right]\partial_{x_0} X_r\,\mathrm{d}r,\end{aligned}$$

which yields with Theorem 1 and Conclusion 1 of [Vrk78] that

$$\det\left(\partial_{x_0} X_s\right) = \exp\left(\int_t^s \mathrm{tr}\left[\partial_x M(X_r,Y_r,Z_r) + \partial_y M(X_r,Y_r,Z_r)\,U_r + \partial_z M(X_r,Y_r,Z_r)\,\tilde{Z}_r\right]\mathrm{d}r\right).$$

Together with Inequality (3.5) this implies that $\tau = 1$ and $\partial_{x_0} X$ is invertible on the whole interval $[t,1]$.

What remains to do is to calculate the explicit dynamics of $U$. Observe that

$$
\begin{aligned}
&b_s\\
&= U_s\left[\partial_x M\left(X_s,Y_s,Z_s\right)+\partial_y M\left(X_s,Y_s,Z_s\right)U_s+\partial_z M\left(X_s,Y_s,Z_s\right)\tilde{Z}_s\right]\\
&= \left(u_s^{(1)},u_s^{(2)},u_s^{(3)}\right)\left[\begin{pmatrix} 0 & 0 & 0\\ 0 & -2\frac{Z_s^2}{\sigma_s^2}\frac{\sigma_{t,s}}{\sigma_s} & -2\frac{Z_s^2}{\sigma_s^2}\frac{\sigma_{a,s}}{\sigma_s}\\ 0 & \mu_{t,s}\frac{Z_s^2}{\sigma_s^2}-2\mu_s\frac{Z_s^2}{\sigma_s^2}\frac{\sigma_{t,s}}{\sigma_s} & \mu_{a,s}\frac{Z_s^2}{\sigma_s^2}-2\mu_s\frac{Z_s^2}{\sigma_s^2}\frac{\sigma_{a,s}}{\sigma_s}\end{pmatrix}\right.\\
&\qquad\left.+\begin{pmatrix} 0\\ -2\frac{Z_s^2}{\sigma_s^2}\frac{\sigma_{a,s}}{\sigma_s}\\ \mu_{a,s}\frac{Z_s^2}{\sigma_s^2}-2\mu_s\frac{Z_s^2}{\sigma_s^2}\frac{\sigma_{a,s}}{\sigma_s}\end{pmatrix}\left(u_s^{(1)},u_s^{(2)},u_s^{(3)}\right)+\begin{pmatrix}0\\ \frac{2Z_s}{\sigma_s^2}\\ \frac{2Z_s}{\sigma_s^2}\mu_s\end{pmatrix}\left(\tilde{Z}_s^{(1)},\tilde{Z}_s^{(2)},\tilde{Z}_s^{(3)}\right)\right]\\
&= \begin{pmatrix} -2u_s^{(1)}u_s^{(2)}\frac{Z_s^2}{\sigma_s^2}\frac{\sigma_{a,s}}{\sigma_s}+u_s^{(1)}u_s^{(3)}\frac{Z_s^2}{\sigma_s^2}\left(\mu_{a,s}-2\mu_s\frac{\sigma_{a,s}}{\sigma_s}\right)\\ -2u_s^{(2)}\frac{Z_s^2}{\sigma_s^2}\frac{\sigma_{t,s}}{\sigma_s}+u_s^{(3)}\frac{Z_s^2}{\sigma_s^2}\left(\mu_{t,s}-2\mu_s\frac{\sigma_{t,s}}{\sigma_s}\right)-2\left(u_s^{(2)}\right)^2\frac{Z_s^2}{\sigma_s^2}\frac{\sigma_{a,s}}{\sigma_s}+u_s^{(2)}u_s^{(3)}\frac{Z_s^2}{\sigma_s^2}\left(\mu_{a,s}-2\mu_s\frac{\sigma_{a,s}}{\sigma_s}\right)\\ -2u_s^{(2)}\frac{Z_s^2}{\sigma_s^2}\frac{\sigma_{a,s}}{\sigma_s}+u_s^{(3)}\frac{Z_s^2}{\sigma_s^2}\left(\mu_{a,s}-2\mu_s\frac{\sigma_{a,s}}{\sigma_s}\right)-2u_s^{(2)}u_s^{(3)}\frac{Z_s^2}{\sigma_s^2}\frac{\sigma_{a,s}}{\sigma_s}+\left(u_s^{(3)}\right)^2\frac{Z_s^2}{\sigma_s^2}\left(\mu_{a,s}-2\mu_s\frac{\sigma_{a,s}}{\sigma_s}\right)\end{pmatrix}^T\\
&\qquad+\begin{pmatrix}\frac{2Z_s}{\sigma_s^2}\left(u_s^{(2)}+u_s^{(3)}\mu_s\right)\tilde{Z}_s^{(1)}\\ \frac{2Z_s}{\sigma_s^2}\left(u_s^{(2)}+u_s^{(3)}\mu_s\right)\tilde{Z}_s^{(2)}\\ \frac{2Z_s}{\sigma_s^2}\left(u_s^{(2)}+u_s^{(3)}\mu_s\right)\tilde{Z}_s^{(3)}\end{pmatrix}^T
\end{aligned}
$$

Using that $Y_1=\xi(X_1)$ and hence $U_1=\nabla\xi(X_1)$ we obtain for the gradient processes the dynamics

$$
\begin{aligned}
u_s^{(1)} &= g'\left(X_1^{(1)}\right)+\int_s^1 u_r^{(1)}\frac{Z_r^2}{\sigma_r^2}\left(u_r^{(3)}\left(\mu_{a,r}-2\mu_r\frac{\sigma_{a,r}}{\sigma_r}\right)-2u_r^{(2)}\frac{\sigma_{a,r}}{\sigma_r}\right)\mathrm{d}r-\int_s^1\tilde{Z}_r^{(1)}\,\mathrm{d}\widetilde{W}_r\\
u_s^{(2)} &= \int_s^1 u_r^{(3)}\frac{Z_r^2}{\sigma_r^2}\left(u_r^{(2)}\mu_{a,r}+\mu_{t,r}\right)-2\frac{Z_r^2}{\sigma_r^2}\left(\frac{\sigma_{t,r}}{\sigma_r}+u_r^{(2)}\frac{\sigma_{a,r}}{\sigma_r}\right)\left(u_r^{(2)}+u_r^{(3)}\mu_r\right)\mathrm{d}r-\int_s^1\tilde{Z}_r^{(2)}\,\mathrm{d}\widetilde{W}_r\\
u_s^{(3)} &= -1+\int_s^1\left(u_r^{(3)}+1\right)\frac{Z_r^2}{\sigma_r^2}\left(u_r^{(3)}\mu_{a,s}-2\frac{\sigma_{a,r}}{\sigma_r}\left(u_r^{(2)}+u_r^{(3)}\mu_r\right)\right)\mathrm{d}r-\int_s^1\tilde{Z}_r^{(3)}\,\mathrm{d}\widetilde{W}_r,
\end{aligned}
$$

where $\widetilde{W}_s:=W_s-\int_t^s\frac{2Z_r}{\sigma_r^2}\left(u_r^{(2)}+u_r^{(3)}\mu_r\right)\mathrm{d}r$ for all $s\in[t,1]$. Since $\frac{2Z_s}{\sigma_r^2}\left(u_s^{(2)}+u_s^{(3)}\mu_s\right)$ is bounded for all $s\in[t,1]$, where $t\in I_{\max}^M$, we get by Girsanov's theorem that $\widetilde{W}$ is a Brownian motion for an equivalent probability measure. ■

## 3.2 Bounding the gradient of the decoupling field

In this section we use the notations and definitions of Section 3.1.

In the following we derive bounds for the gradient processes that do not depend on the starting time $t\in I_{\max}$ and initial value $x\in\mathbb{R}^3$. In particular, we obtain global estimates for the space derivatives $\partial_{x_i}u$, $i\in\{1,2,3\}$, of the decoupling field $u$. By applying Theorem 2.11 we can conclude that FBSDE (3.2) has a solution on the whole interval $[0,1]$.

**Lemma 3.2**
*Assume that $g$, $\mu$ and $\sigma$ are differentiable, $\sigma\geqslant\varepsilon>0$ and $g'$, $\frac{\mu}{\sigma^2}$, $\frac{\partial_t\mu}{\sigma^2}$, $\frac{\partial_a\mu}{\sigma^2}$, $\frac{\partial_t\sigma}{\sigma}$, $\frac{\partial_a\sigma}{\sigma}$ are bounded. Let $u$ be the unique decoupling field to FBSDE* (3.2) *on* $I_{\max}$.

*Furthermore, let $t \in I_{\max}$, $x \in \mathbb{R}^3$ and $(X^{(1)}, X^{(2)}, X^{(3)}, Y, Z)$ be the solution of FBSDE* (3.2) *with initial condition $x$ at time $t$, and let $u^{(1)}, u^{(2)}, u^{(3)}$ be the associated gradient processes. Then for almost all $(\omega, s) \in \Omega \times [t, 1]$*

$$|Z_s| \leqslant \sup_{r \in (s,1]} \sup_{x \in \mathbb{R}^3} |\partial_{x_1} u(r, x)|$$

*and in particular $\|Z\|_{\infty,t} \leqslant \|\partial_{x_1} u\|_{\infty,t}$.*

*Furthermore, if the weak derivative $\partial_{x_1} u$ has a version whose restriction to the set $[t, 1) \times \mathbb{R}^3$ is continuous in the first two components $t$ and $x_1$, and $\partial_{x_1} u$ is bounded, then*

$$Z_s(\omega) = \partial_{x_1} u\left(s, X_s^{(1)}(\omega), X_s^{(2)}(\omega), X_s^{(3)}(\omega)\right) = u_s^{(1)}(\omega)$$

*for almost all $(\omega, s) \in \Omega \times [t, 1]$.*

*Proof.* Observe that with Itô's formula we get for a.a. $(\omega, s) \in [t, 1)$ and $h > 0$ such that $s + h \in [t, 1]$

$$\begin{aligned} \frac{1}{h}\mathbb{E}\left[Y_{s+h}(W_{s+h} - W_s) \middle| \mathcal{F}_s\right] &= \frac{1}{h}\mathbb{E}\left[\int_s^{s+h} Y_r \, \mathrm{d}W_r + \int_s^{s+h} (W_r - W_s) Z_r \, \mathrm{d}W_r + \int_s^{s+h} Z_r \, \mathrm{d}r \middle| \mathcal{F}_s\right] \\ &= \frac{1}{h}\mathbb{E}\left[\int_s^{s+h} Z_r \, \mathrm{d}r \middle| \mathcal{F}_s\right] \\ &\to Z_s \quad \text{for} \quad h \to 0. \end{aligned}$$

On the other hand we get, using the decoupling condition $Y_r = u\left(r, X_r^{(1)}, X_r^{(2)}, X_r^{(3)}\right)$, that

$$\begin{aligned} &Y_{s+h}(W_{s+h} - W_s) \\ &\quad = u\left(s+h, X_{s+h}^{(1)}, X_{s+h}^{(2)}, X_{s+h}^{(3)}\right)(W_{s+h} - W_s) \\ &\quad = u\left(s+h, X_{s+h}^{(1)}, X_s^{(2)}, X_s^{(3)}\right)(W_{s+h} - W_s) \\ &\qquad + \left(u\left(s+h, X_{s+h}^{(1)}, X_{s+h}^{(2)}, X_s^{(3)}\right) - u\left(s+h, X_{s+h}^{(1)}, X_s^{(2)}, X_s^{(3)}\right)\right)(W_{s+h} - W_s) \\ &\qquad + \left(u\left(s+h, X_{s+h}^{(1)}, X_{s+h}^{(2)}, X_{s+h}^{(3)}\right) - u\left(s+h, X_{s+h}^{(1)}, X_{s+h}^{(2)}, X_s^{(3)}\right)\right)(W_{s+h} - W_s). \end{aligned} \tag{3.6}$$

At first let us take a look at the third summand on the right hand side of (3.6). Since $u$ is Lipschitz continuous in its fourth argument on $[t, 1]$ with some constant $L_{u,x_3}^t$ that might depend on $t$ and since furthermore $X_{s+h}^{(3)} = X_s^{(3)} + \int_s^{s+h} \mu_r \frac{Z_r^2}{\sigma^2(X_r^{(2)}, Y_r + X_r^{(3)})} \, \mathrm{d}r$ we can estimate the absolute value of the third summand against

$$\begin{aligned} &\frac{1}{h}\left|\mathbb{E}\left[\left(u\left(s+h, X_{s+h}^{(1)}, X_{s+h}^{(2)}, X_{s+h}^{(3)}\right) - u\left(s+h, X_{s+h}^{(1)}, X_{s+h}^{(2)}, X_s^{(3)}\right)\right)(W_{s+h} - W_s) \middle| \mathcal{F}_s\right]\right| \\ &\quad \leqslant \frac{1}{h}\mathbb{E}\left[\left|u\left(s+h, X_{s+h}^{(1)}, X_{s+h}^{(2)}, X_{s+h}^{(3)}\right) - u\left(s+h, X_{s+h}^{(1)}, X_{s+h}^{(2)}, X_s^{(3)}\right)\right| |W_{s+h} - W_s| \middle| \mathcal{F}_s\right] \\ &\quad \leqslant \frac{1}{h}\mathbb{E}\left[L_{u,x_3}^t \left|\int_s^{s+h} \mu_r \frac{Z_r^2}{\sigma^2(X_r^{(2)}, Y_r + X_r^{(3)})} \, \mathrm{d}r\right| |W_{s+h} - W_s| \middle| \mathcal{F}_s\right] \\ &\quad \leqslant \frac{1}{h} L_{u,x_3}^t h \left\|\frac{\mu}{\sigma^2}\right\|_\infty \|Z\|_{\infty,t}^2 \mathbb{E}\left[|W_{s+h} - W_s| \middle| \mathcal{F}_s\right], \end{aligned}$$

which clearly goes to $0$ as $h \to 0$ because $\|\frac{\mu}{\sigma^2}\|_\infty$ and $\|Z\|_{\infty,t}$ are finite on $[t,1]$.

With analogous arguments we also get that

$$\begin{aligned}
&\frac{1}{h}\left|\mathbb{E}\left[\left(u\left(s+h, X^{(1)}_{s+h}, X^{(2)}_{s+h}, X^{(3)}_s\right) - u\left(s+h, X^{(1)}_{s+h}, X^{(2)}_s, X^{(3)}_s\right)\right)(W_{s+h}-W_s)\middle|\mathcal{F}_s\right]\right| \\
&\leqslant \frac{1}{h}\mathbb{E}\left[\left|u\left(s+h, X^{(1)}_{s+h}, X^{(2)}_{s+h}, X^{(3)}_s\right) - u\left(s+h, X^{(1)}_{s+h}, X^{(2)}_s, X^{(3)}_s\right)\right||W_{s+h}-W_s|\middle|\mathcal{F}_s\right] \\
&\leqslant \frac{1}{h}\mathbb{E}\left[L^t_{u,x_2}\left|\int_s^{s+h}\frac{Z_r^2}{\sigma^2(X^{(2)}_r, Y_r + X^{(3)}_r)}\,\mathrm{d}r\right||W_{s+h}-W_s|\middle|\mathcal{F}_s\right] \\
&\leqslant \frac{1}{h}L^t_{u,x_2}h\|Z\|^2_{\infty,t}\varepsilon^{-2}\mathbb{E}\left[|W_{s+h}-W_s|\right|\mathcal{F}_s] \\
&\to 0 \;\; a.s. \quad \text{for} \quad h \to 0,
\end{aligned}$$

where $L^t_{u,x_2}$ is the Lipschitz constant of $u$ in the third argument on the time interval $[t,1]$.

Now consider the remaining first term on the right hand side of Equation (3.6). For this remember

- $X^{(1)}_s$, $X^{(2)}_s$, $X^{(3)}_s$ are $\mathcal{F}_s$ measurable,
- $X^{(1)}_{s+h} = X^{(1)}_s + (W_{s+h} - W_s)$,
- $W_{s+h} - W_s$ is independent of $\mathcal{F}_s$,
- $u$ is deterministic, i.e. is a function of $(s, x^{(1)}, x^{(2)}, x^{(3)}) \in [t,1] \times \mathbb{R} \times \mathbb{R} \times \mathbb{R}$ only.

Using integration by parts these properties imply

$$\begin{aligned}
&\mathbb{E}\left[u\left(s+h, X^{(1)}_{s+h}, X^{(2)}_s, X^{(3)}_s\right)(W_{s+h}-W_s)\middle|\mathcal{F}_s\right] \\
&\qquad = \int_{\mathbb{R}} u\left(s+h, X^{(1)}_s + z\sqrt{h}, X^{(2)}_s, X^{(3)}_s\right) z\sqrt{h}\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}z^2}\,\mathrm{d}z \\
&\qquad = \int_{\mathbb{R}} \partial_{x_1}u\left(s+h, X^{(1)}_s + z\sqrt{h}, X^{(2)}_s, X^{(3)}_s\right) h\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}z^2}\,\mathrm{d}z.
\end{aligned}$$

Hence

$$\begin{aligned}
&\left|\frac{1}{h}\mathbb{E}\left[u\left(s+h, X^{(1)}_{s+h}, X^{(2)}_s, X^{(3)}_s\right)(W_{s+h}-W_s)\middle|\mathcal{F}_s\right]\right| \\
&\qquad = \left|\int_{\mathbb{R}} \partial_{x_1}u\left(s+h, X^{(1)}_s + z\sqrt{h}, X^{(2)}_s, X^{(3)}_s\right)\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}z^2}\,\mathrm{d}z\right| \\
&\qquad \leqslant \int_{\mathbb{R}} \sup_{x\in\mathbb{R}^3}|\partial_{x_1}u(s+h,x)|\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}z^2}\,\mathrm{d}z \\
&\qquad = \sup_{x\in\mathbb{R}^3}|\partial_{x_1}u(s+h,x)|.
\end{aligned}$$

Putting everything together we get

$$
\begin{aligned}
|Z_s| &= \lim_{h\searrow 0}\left|\frac{1}{h}\mathbb{E}\left[\int_s^{s+h} Z_r\,\mathrm{d}r\middle|\,\mathcal{F}_s\right]\right| \\
&= \lim_{h\searrow 0}\left|\frac{1}{h}\mathbb{E}\left[Y_{s+h}(W_{s+h}-W_s)|\,\mathcal{F}_s\right]\right| \\
&= \lim_{h\searrow 0}\left|\frac{1}{h}\mathbb{E}\left[u\left(s+h,X^{(1)}_{s+h},X^{(2)}_s,X^{(3)}_s\right)(W_{s+h}-W_s)\middle|\,\mathcal{F}_s\right]\right. \\
&\quad+\frac{1}{h}\mathbb{E}\left[\left(u\left(s+h,X^{(1)}_{s+h},X^{(2)}_{s+h},X^{(3)}_s\right)-u\left(s+h,X^{(1)}_{s+h},X^{(2)}_s,X^{(3)}_s\right)\right)(W_{s+h}-W_s)\middle|\,\mathcal{F}_s\right] \\
&\quad\left.+\frac{1}{h}\mathbb{E}\left[\left(u\left(s+h,X^{(1)}_{s+h},X^{(2)}_{s+h},X^{(3)}_{s+h}\right)-u\left(s+h,X^{(1)}_{s+h},X^{(2)}_{s+h},X^{(3)}_s\right)\right)(W_{s+h}-W_s)\middle|\,\mathcal{F}_s\right]\right| \\
&\leqslant \limsup_{h\searrow 0}\sup_{x\in\mathbb{R}^3}|\partial_{x_1}u(s+h,x)|+|0|+|0| \\
&\leqslant \sup_{r\in(s,1]}\sup_{x\in\mathbb{R}^3}|\partial_{x_1}u(r,x)|.
\end{aligned}
$$

If we have that $\partial_{x_1}u$ is continuous in the first two arguments, we can derive, by using dominated convergence since $u^{(1)}$ is bounded on $[t,1]$, the more precise result

$$
\begin{aligned}
Z_s &= \lim_{h\searrow 0}\frac{1}{h}\mathbb{E}\left[u\left(s+h,X^{(1)}_{s+h},X^{(2)}_s,X^{(3)}_s\right)(W_{s+h}-W_s)\middle|\,\mathcal{F}_s\right] \\
&= \int_{\mathbb{R}}\lim_{h\searrow 0}\partial_{x_1}u\left(s+h,X^{(1)}_s+z\sqrt{h},X^{(2)}_s,X^{(3)}_s\right)\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}z^2}\,\mathrm{d}z \\
&= \partial_{x_1}u\left(s,X^{(1)}_s,X^{(2)}_s,X^{(3)}_s\right)
\end{aligned}
$$

almost surely. ■

To obtain estimates for the gradient processes we use the following result.

**Lemma 3.3** (See [MPF12], pg. 362)
*Let the function $f$ be continuous and non-negative on $J=[\alpha,\beta]$, $a,b\geqslant 0$, and $n$ be a positive integer ($n\geqslant 2$). If*

$$f(t)\leqslant a+b\int_\alpha^t f^n(s)\,\mathrm{d}s,\quad t\in J,$$

*then*

$$f(t)\leqslant a\left[1-(n-1)\int_\alpha^t a^{n-1}b\,\mathrm{d}s\right]^{\frac{1}{1-n}},\quad \alpha\leqslant t\leqslant\beta_n,$$

*where $\beta_n=\sup\left\{t\in J:(n-1)\int_\alpha^t a^{n-1}b\,\mathrm{d}s<1\right\}$.*

**Lemma 3.4**
*Assume that $g$, $\mu$ and $\sigma$ are differentiable, $\sigma\geqslant\varepsilon>0$ and $g'$, $\frac{\mu}{\sigma^2}$, $\frac{\partial_t\mu}{\sigma^2}$, $\frac{\partial_a\mu}{\sigma^2}$, $\frac{\partial_t\sigma}{\sigma}$, $\frac{\partial_a\sigma}{\sigma}$ are bounded. Let $u$ be the unique decoupling field of the FBSDE (3.2). Then for any $t\in I_{\max}$ and initial condition $(X^{(1)}_t,X^{(2)}_t,X^{(3)}_t)=x\in\mathbb{R}^3$ the associated gradient process $u^{(3)}$ satisfies for all $s\in[t,1]$*

$$u^{(3)}_s=-1.$$

*If we additionally assume that* $\sigma_{a,s} \cdot u_s^{(2)} \geqslant 0$ *a.s. for all* $s \in [t,1]$ *and*

$$\inf_{(\theta,x)\in\mathbb{R}_+\times\mathbb{R}} \frac{\sigma \cdot \partial_a \mu - 2\partial_a \sigma \cdot \mu}{\sigma^3}(\theta,x) > -\frac{1}{2\|g'\|_\infty^2},$$

*then it also holds that*

$$0 \leqslant u_s^{(1)} \leqslant \left( \frac{1}{\|g'\|_\infty^2} + 2 \min\left\{0, \inf_{(\theta,x)\in\mathbb{R}_+\times\mathbb{R}} \left( \frac{\sigma \cdot \partial_a \mu - 2\partial_a \sigma \cdot \mu}{\sigma^3} \right)(\theta,x) \right\} \right)^{-\frac{1}{2}} < \infty$$

*for all* $s \in [t,1]$.

*Proof.* By interpreting (3.4) as a system of BSDEs we get for $u^{(3)}$ the trivial solution $u_s^{(3)} = -1$ for all $s \in [t,1]$ as the unique bounded solution of this BSDE.

Also note that $g' \geqslant 0$ since $g = F_\nu^{-1} \circ \Phi$ and $F_\nu$ as well as $\Phi$ are non-decreasing. Thus $\check{u}_s = 0$ is the trivial and unique solution to

$$\check{u}_s = 0 + \int_s^1 -\check{u}_r \frac{Z_r^2}{\sigma_r^2} \left( -\mu_{a,r} + 2\mu_r \frac{\sigma_{a,r}}{\sigma_r} - 2u_r^{(2)} \frac{\sigma_{a,r}}{\sigma_r} \right) \mathrm{d}r - \int_s^1 \tilde{Z}_r^{(1)} \,\mathrm{d}\widetilde{W}_r,$$

which implies by comparison (see e.g. Theorem 6.2.2 in [Pha09]) that $0 = \check{u}_s \leqslant u_s^{(1)}$ for all $s \in [t,1]$.

For the upper bound of $u^{(1)}$ remember that $u_s^{(1)} = \partial_{x_1} u(s, X_s^{(1)}, X_s^{(2)}, X_s^{(3)})$ for all $s \in [t,1]$ and in particular for any fixed $t \in I_{\max}$ and all starting conditions $x = (x^{(1)}, x^{(2)}, x^{(3)}) \in \mathbb{R}^3$ we have

$$\partial_{x_1} u(t,x) = u_t^{(1)} = g'\left(X_1^{(1)}\right) - \int_t^1 u_r^{(1)} \frac{Z_r^2}{\sigma_r^2} \left( \mu_{a,r} - 2\mu_r \frac{\sigma_{a,r}}{\sigma_r} + 2u_r^{(2)} \frac{\sigma_{a,r}}{\sigma_r} \right) \mathrm{d}r - \int_t^1 \tilde{Z}_r^{(1)} \,\mathrm{d}\widetilde{W}_r.$$

Using this and that $Z$ is bounded on every interval $[t,1] \subset I_{\max}$, we get

$$\begin{aligned}
u_t^{(1)} &= \mathbb{E}\left[ u_t^{(1)} \middle| \mathcal{F}_t \right] \\
&= \mathbb{E}\left[ g'\left(X_1^{(1)}\right) - \int_t^1 u_r^{(1)} \frac{Z_r^2}{\sigma_r^2} \left( \mu_{a,r} - 2\mu_r \frac{\sigma_{a,r}}{\sigma_r} + 2u_r^{(2)} \frac{\sigma_{a,r}}{\sigma_r} \right) \mathrm{d}r \middle| \mathcal{F}_t \right] \\
&\leqslant \mathbb{E}\left[ g'\left(X_1^{(1)}\right) - \int_t^1 u_r^{(1)} \frac{Z_r^2}{\sigma_r^2} \left( \mu_{a,r} - 2\frac{\sigma_{a,r}}{\sigma_r}\mu_r \right) \mathrm{d}r \middle| \mathcal{F}_t \right]
\end{aligned}$$

for all $t \in I_{\max}$ and $(x^{(1)}, x^{(2)}, x^{(3)}) \in \mathbb{R}^3$, where we use that $\sigma_{a,r} \cdot u_r^{(2)} \geqslant 0$. Next we use the inequality

$$-\frac{\sigma_s \mu_{a,s} - 2\sigma_{a,s}\mu_r}{\sigma_s^3} \leqslant \max\left\{0, -\inf_{(\theta,x)\in\mathbb{R}_+\times\mathbb{R}} \left( \frac{\sigma \cdot \partial_a \mu - 2\partial_a \sigma \cdot \mu}{\sigma^3} \right)(\theta,x) \right\} =: \beta$$

and the estimate from Lemma 3.2 for $Z$ to obtain

$$u_t^{(1)} \leqslant \|g'\|_\infty + \beta \int_t^1 \sup_{x\in\mathbb{R}^3} \partial_{x_1} u(r,x) \sup_{\theta\in[r,1]} \sup_{x\in\mathbb{R}^3} \left(\partial_{x_1} u\right)^2(\theta,x) \,\mathrm{d}r.$$

Thus we can derive the inequality

$$\begin{aligned}
\sup_{\rho\in[t,1]} \sup_{x\in\mathbb{R}^3} \partial_{x_1} u(\rho,x) &\leqslant \|g'\|_\infty + \beta \sup_{\rho\in[t,1]} \left\{ \int_\rho^1 \sup_{x\in\mathbb{R}^3} \partial_{x_1} u(r,x) \sup_{\theta\in[r,1]} \sup_{x\in\mathbb{R}^3} \left(\partial_{x_1} u\right)^2(\theta,x) \,\mathrm{d}r \right\} \\
&\leqslant \|g'\|_\infty + \beta \int_t^1 \sup_{\theta\in[r,1]} \sup_{x\in\mathbb{R}^3} \left(\partial_{x_1} u\right)^3(\theta,x) \,\mathrm{d}r.
\end{aligned}$$

Note that $\inf_{(\theta,x)\in\mathbb{R}_+\times\mathbb{R}} \frac{\sigma\cdot\partial_a\mu - 2\partial_a\sigma\cdot\mu}{\sigma^3}(\theta,x) > -\frac{1}{2\|g'\|_\infty^2}$ implies $\beta < \frac{1}{2\|g'\|_\infty^2}$. Hence, we obtain by setting $f(t) = \sup_{\rho\in[t,1]}\sup_{x\in\mathbb{R}^3}\partial_{x_1}u(\rho,x)$ and applying Lemma 3.3 that

$$\sup_{\rho\in[t,1]}\sup_{x\in\mathbb{R}^3}\partial_{x_1}u(\rho,x) \leqslant \left(\frac{1}{\|g'\|_\infty^2} - 2\beta(1-t)\right)^{-\frac{1}{2}}$$

and thus,

$$\|u^{(1)}\|_{\infty,t} \leqslant \|\partial_{x_1}u\|_{\infty,t} \leqslant \left(\frac{1}{\|g'\|_\infty^2} - 2\beta\right)^{-\frac{1}{2}} < \infty.$$

■

**Assumption 3.5**
*Let $g$, $\mu$ and $\sigma$ be differentiable, $\sigma \geqslant \varepsilon > 0$ and $g'$, $\frac{\mu}{\sigma^2}$, $\frac{\partial_t\mu}{\sigma^2}$, $\frac{\partial_a\mu}{\sigma^2}$, $\frac{\partial_t\sigma}{\sigma}$ as well as $\frac{\partial_a\sigma}{\sigma}$ be bounded. Furthermore, let*

$$\inf_{(\theta,x)\in\mathbb{R}_+\times\mathbb{R}} \frac{\sigma\cdot\partial_a\mu - 2\partial_a\sigma\cdot\mu}{\sigma^3}(\theta,x) > -\frac{1}{2\|g'\|_\infty^2} \tag{3.7}$$

*and one of the following conditions be satisfied:*

*i)* $\partial_a\sigma \equiv 0$

*ii)* $\partial_a\sigma \geqslant 0$, $2\partial_t\sigma\cdot\mu - \sigma\cdot\partial_t\mu \geqslant 0$ *or*

*iii)* $\partial_a\sigma \leqslant 0$, $2\partial_t\sigma\cdot\mu - \sigma\cdot\partial_t\mu \leqslant 0$.

**Theorem 3.6**
*Let $g$, $\mu$ and $\sigma$ fulfill Assumption 3.5. Then, for FBSDE* (3.2)*, we have $I_{\max} = [0,1]$ and there exists a unique, strongly regular Markovian decoupling field $u$ on the whole interval $[0,1]$. This $u$ is a continuous function on $[0,1]\times\mathbb{R}^3$.*

*Furthermore let $(X^{(1)}, X^{(2)}, X^{(3)}, Y, Z)$ be the solution of FBSDE* (3.2) *with an arbitrary initial condition $x \in \mathbb{R}^3$ and $u^{(1)}, u^{(2)}, u^{(3)}$ be the associated gradient processes on $[0,1]$. Then we have $u^{(3)} \equiv -1$ and the finite estimates*

$$0 \leqslant u^{(1)} \leqslant \left(\frac{1}{\|g'\|_\infty^2} + 2\min\left\{0, \inf_{(\theta,x)\in\mathbb{R}_+\times\mathbb{R}}\left(\frac{\sigma\cdot\partial_a\mu - 2\partial_a\sigma\cdot\mu}{\sigma^3}\right)(\theta,x)\right\}\right)^{-\frac{1}{2}}, \tag{3.8}$$

$$\begin{aligned}\left\|u^{(2)}\right\|_\infty \leqslant \exp\left[\|Z\|_\infty^2\left(\left\|\frac{\partial_a\mu}{\sigma^2}\right\|_\infty + 2\left(\left\|\frac{\partial_a\sigma}{\sigma}\right\|_\infty\left\|\frac{\mu}{\sigma^2}\right\|_\infty + \frac{1}{\varepsilon^2}\left\|\frac{\partial_t\sigma}{\sigma}\right\|_\infty\right)\right)\right] \\ \cdot\|Z\|_\infty^2\left(2\left\|\frac{\partial_t\sigma}{\sigma}\right\|_\infty\left\|\frac{\mu}{\sigma^2}\right\|_\infty + \left\|\frac{\partial_t\mu}{\sigma^2}\right\|_\infty\right)\end{aligned} \tag{3.9}$$

*and*

$$\|Z\|_\infty \leqslant \left\|u^{(1)}\right\|_\infty \leqslant \left(\frac{1}{\|g'\|_\infty^2} + 2\min\left\{0, \inf_{(\theta,x)\in\mathbb{R}_+\times\mathbb{R}}\left(\frac{\sigma\cdot\partial_a\mu - 2\partial_a\sigma\cdot\mu}{\sigma^3}\right)(\theta,x)\right\}\right)^{-\frac{1}{2}}. \tag{3.10}$$

*Proof.* Using Lemma 2.10 we only need to show that the weak derivative of $u$ with regard to the initial value $x \in \mathbb{R}^3$ is bounded by some constant which is independent of the time interval $[t,1] \subset I_{\max}$ on which it is defined. Then it follows that $I_{\max} = [0,1]$ and hence $t$ can be chosen to equal $0$ and the estimates (3.8), (3.9) and (3.10) hold true for corresponding processes on the whole interval $[0,1]$.

For now fix $t \in I_{\max}$ and $x \in \mathbb{R}^3$ and let $u^{(1)}, u^{(2)}, u^{(3)}$ be the associated gradient processes. Lemma 3.4 yields $u^{(3)} \equiv -1$. In order to derive Estimate (3.8) we show that $\sigma_{a,s} \cdot u_s^{(2)} \geqslant 0$ a.s. for all $s \in [t,1]$ which then allows us to apply Lemma 3.4 yielding the estimate. Consider the three cases $i)$, $ii)$ and $iii)$ of Assumption 3.5: With $\partial_a \sigma \equiv 0$ of case $i)$ this is obviously true. For the remaining two cases observe that

$$u_s^{(2)} = \int_s^1 \frac{Z_r^2}{\sigma_r^2} \left[ \left(u_r^{(2)}\right)^2 \left(-2\frac{\sigma_{a,r}}{\sigma_r}\right) + u_r^{(2)} \left(-\mu_{a,r} + 2\frac{\sigma_{a,r}}{\sigma_r}\mu_r - 2\frac{\sigma_{t,r}}{\sigma_r}\right) + \left(2\frac{\sigma_{t,r}}{\sigma_r}\mu_r - \mu_{t,r}\right)\right] \mathrm{d}r - \int_s^1 \tilde{Z}_r^{(2)} \,\mathrm{d}\widetilde{W}_r.$$

Because $u_r^{(2)}$ is bounded on every interval $[t,1] \subset I_{\max}$, we can view $u^{(2)}$ as fulfilling a Lipschitz BSDE. This allows us to use the comparison theorem by changing $2\frac{\sigma_{t,r}}{\sigma_r}\mu_r - \mu_{t,r}$ to zero and hence compare with the trivial solution which is constantly $0$. Thus in the case $ii)$ we have $u^{(2)} \geqslant 0$ and in case $iii)$ $u^{(2)} \leqslant 0$. Therefore, we have $\partial_a \sigma \cdot u^{(2)} \geqslant 0$ for the cases $ii)$ and $iii)$ as well. Hence we can apply Lemma 3.4 to obtain, for $s \in [t,1]$,

$$0 \leqslant u_s^{(1)} \leqslant \left( \frac{1}{\|g'\|_\infty^2} + 2\min\left\{0, \inf_{(\theta,x)\in\mathbb{R}_+\times\mathbb{R}} \left(\frac{\sigma\cdot\partial_a\mu - 2\partial_a\sigma\cdot\mu}{\sigma^3}\right)(\theta,x)\right\}\right)^{-\frac{1}{2}}.$$

In addition with Lemma 3.2 this yields

$$\|Z\|_{\infty,t} \leqslant \|u^{(1)}\|_{\infty,t} \leqslant \left( \frac{1}{\|g'\|_\infty^2} + 2\min\left\{0, \inf_{(\theta,x)\in\mathbb{R}_+\times\mathbb{R}} \left(\frac{\sigma\cdot\partial_a\mu - 2\partial_a\sigma\cdot\mu}{\sigma^3}\right)(\theta,x)\right\}\right)^{-\frac{1}{2}} < \infty.$$

Since, as stated before, in case $ii)$ we have $u^{(2)} \geqslant 0$ and $\partial_a\sigma \geqslant 0$ and in case $iii)$ $u^{(2)} \leqslant 0$ and $\partial_a\sigma \leqslant 0$, we again can apply the comparison theorem to see that in case $ii)$ we have $0 \leqslant u^{(2)} \leqslant \bar{u}$ and in case $iii)$ $\bar{u} \leqslant u^{(2)} \leqslant 0$, where $\bar{u}$ is the solution of the linear BSDE

$$\bar{u}_s = \int_s^1 \bar{u}_r \frac{Z_r^2}{\sigma_r^2}\left(-\mu_{a,r} + 2\frac{\sigma_{a,r}}{\sigma_r}\mu_r - 2\frac{\sigma_{t,r}}{\sigma_r}\right) + \frac{Z_r^2}{\sigma_r^2}\left(2\frac{\sigma_{t,r}}{\sigma_r}\mu_r - \mu_{t,r}\right)\mathrm{d}r - \int_s^1 \bar{Z}_r \,\mathrm{d}\widetilde{W}_r.$$

In case $i)$ we have that $u^{(2)} = \bar{u}$ giving that $u^{(2)}$ is bounded by $\bar{u}$ as well.

By estimating

$$\begin{aligned}
|\bar{u}_s| &= \left|\mathbb{E}\left[\int_s^1 \exp\left(\int_s^r \frac{Z_r^2}{\sigma_r^2}\left(-\mu_{a,r} + 2\frac{\sigma_{a,r}}{\sigma_r}\mu_r - 2\frac{\sigma_{t,r}}{\sigma_r}\right)\mathrm{d}\rho\right)\frac{Z_r^2}{\sigma_r^2}\left(2\frac{\sigma_{t,r}}{\sigma_r}\mu_r - \mu_{t,r}\right)\mathrm{d}r \,\middle|\, \mathcal{F}_s\right]\right| \\
&\leqslant \exp\left[\|Z\|_\infty^2\left(\left\|\frac{\partial_a\mu}{\sigma^2}\right\|_\infty + 2\left(\left\|\frac{\partial_a\sigma}{\sigma}\right\|_\infty \left\|\frac{\mu}{\sigma^2}\right\|_\infty + \frac{1}{\varepsilon^2}\left\|\frac{\partial_t\sigma}{\sigma}\right\|_\infty\right)\right)\right] \\
&\qquad \cdot \|Z\|_\infty^2\left(2\left\|\frac{\partial_t\sigma}{\sigma}\right\|_\infty\left\|\frac{\mu}{\sigma^2}\right\|_\infty + \left\|\frac{\partial_t\mu}{\sigma^2}\right\|_\infty\right)
\end{aligned}$$

we have found a finite bound for $u^{(2)}$ that is independent of $t$.

Thus $u^{(1)}$, $u^{(2)}$ and $u^{(3)}$ are bounded independently of $t$. Thus, Theorem 2.11 gives that $I_{\max} = [0,1]$, FBSDE (3.2) has a unique solution and we also have that all bounds are valid on this interval. ■

## 3.3 Weak solution

In this section we show that a weak solution of the SEP can be obtained from the solution of the FBSDE (3.2). This means that we construct the Brwonian motion driving the stopped process. Recall that if Assumption 3.5 is fulfilled, then by Theorem 3.6 FBSDE (3.2) has a solution on the whole interval $[0, 1]$ and the gradient processes are bounded.

In the following we sometimes use the fact that for two Itô processes $A$ and $B$ and a time change $\gamma$, in the sense of Definition 1.2 in Chapter V, [RY13], it holds that

$$\int_0^{\gamma(t)} A_r \,\mathrm{d}B_r = \int_0^t A_{\gamma(r)} \,\mathrm{d}B_{\gamma(r)}$$

(see e.g. Proposition 1.4, Chapter V, [RY13]).

The next theorem is a version of Theorem 3.21 with an explicit weak solution of the SEP.

**Theorem 3.7**
*Let $g$, $\mu$ and $\sigma$ fulfill Assumption 3.5. Furthermore let $(X^{(1)}, X^{(2)}, X^{(3)}, Y, Z)$ be the solution of the FBSDE* (3.2) *with initial value $(X_0^{(1)}, X_0^{(2)}, X_0^{(3)}) = (0, 0, 0)$. Define the random time*

$$\tilde{\tau} := X_1^{(2)},$$

*the time change*

$$\gamma(t) := \begin{cases} \inf\left\{ s \geqslant 0 | X_s^{(2)} > t \right\} & \text{if } 0 \leqslant t < \tilde{\tau}, \\ 1 & \text{if } t \geqslant \tilde{\tau}, \end{cases}$$

*the filtration $\mathcal{G}_t := \mathcal{F}_{\gamma(t)}$ and the process $A_t := Y_{\gamma(t)} + X^{(3)}_{\gamma(t)}$ on $[0, \tilde{\tau}]$.*

*Then $\tilde{\tau}$ is a $(\mathcal{G}_t)$-stopping time satisfying*

$$\tilde{\tau} \leqslant \varepsilon^{-2} \left( \frac{1}{\|g'\|_\infty^2} + 2 \min\left\{ 0, \inf_{(\theta,x) \in \mathbb{R}_+ \times \mathbb{R}} \left( \frac{\sigma \cdot \partial_a \mu - 2\partial_a \sigma \cdot \mu}{\sigma^3} \right)(\theta, x) \right\} \right)^{-1} \text{ a.s.}$$

*Furthermore, on $[0, \tilde{\tau}]$, the process $B_t := \int_0^t \frac{1}{\sigma(X^{(2)}_{\gamma(r)}, Y_{\gamma(r)} + X^{(3)}_{\gamma(r)})} \,\mathrm{d}Y_{\gamma(r)}$ is a $(\mathcal{G}_t)$-Brownian motion, $A$ fulfills the SDE*

$$A_t = Y_0 + \int_0^t \mu(r, A_r) \,\mathrm{d}r + \int_0^t \sigma(r, A_r) \,\mathrm{d}B_r$$

*and we have*

$$A_{\tilde{\tau}} \sim \nu.$$

*Proof.* By standard results it follows that $\tilde{\tau}$ is a $(\mathcal{G}_t)$-stopping time (see e.g. Proposition 1.1, Chapter V, [RY13]). With

$$\gamma^{-1}(s) := X_s^{(2)} \tag{3.11}$$

for all $s \in [0, 1]$ we have for all $t \in [0, \tilde{\tau}]$ that $X^{(2)}_{\gamma(t)} = \gamma^{-1}(\gamma(t)) = t$. Therefore, and because $\mathrm{d}Y_r = Z_r \,\mathrm{d}W_r$, we obtain

$$\langle B, B \rangle_t = \int_0^{\gamma(t)} \frac{Z_r^2}{\sigma^2(X_r^{(2)}, Y_r + X_r^{(3)})} \,\mathrm{d}r = \gamma^{-1}(\gamma(t)) = t.$$

By Levy's characterisation of Brownian motion we get that $(B_t)$ is a $(\mathcal{G}_t)$-Brownian motion on $[0, \tilde{\tau}]$.

Note that for all $\omega \in \Omega$ the function $\gamma$ is $\lambda$-a.e. differentiable on $[0, \tilde{\tau}]$ with

$$\gamma'(t) = ((\gamma^{-1})^{-1})'(t) = \frac{1}{(\gamma^{-1})'(\gamma(t))} = \frac{\sigma^2(X^{(2)}_{\gamma(t)}, Y_{\gamma(t)} + X^{(3)}_{\gamma(t)})}{Z^2_{\gamma(t)}} \tag{3.12}$$

and hence

$$\begin{aligned} A_t &= X^{(3)}_{\gamma(t)} + Y_{\gamma(t)} - Y_0 + Y_0 \\ &= Y_0 + \int_0^{\gamma(t)} \mu\left(X^{(2)}_r, Y_r + X^{(3)}_r\right) \frac{Z_r^2}{\sigma^2(X^{(2)}_r, Y_r + X^{(3)}_r)}\,\mathrm{d}r + \int_0^{\gamma(t)} \frac{\sigma(X^{(2)}_r, Y_r + X^{(3)}_r)}{\sigma(X^{(2)}_r, Y_r + X^{(3)}_r)}\,\mathrm{d}Y_r \\ &= Y_0 + \int_0^t \mu\left(X^{(2)}_{\gamma(r)}, Y_{\gamma(r)} + X^{(3)}_{\gamma(r)}\right)\mathrm{d}r + \int_0^t \sigma\left(X^{(2)}_{\gamma(r)}, Y_{\gamma(r)} + X^{(3)}_{\gamma(r)}\right)\mathrm{d}B_r \\ &= Y_0 + \int_0^t \mu\left(r, A_r\right)\mathrm{d}r + \int_0^t \sigma(r, A_r)\,\mathrm{d}B_r \end{aligned}$$

for all $t \in [0, \tilde{\tau}]$. Also

$$A_{\tilde{\tau}} = Y_{\gamma(\tilde{\tau})} + X^{(3)}_{\gamma(\tilde{\tau})} = Y_1 + X^{(3)}_1 = g(W_1) \sim \nu.$$

The bound for $\tilde{\tau}$ follows with the bound for $\|Z\|_\infty$ stated in Theorem 3.6 and by $\sigma \geqslant \varepsilon$. ■

The next lemma characterizes the stopping time $\tilde{\tau} = \gamma^{-1}(1)$ of Theorem 3.7 in terms of the solution of an FBSDE driven by the Brownian motion $B$. We use the lemma later to show existence of strong solutions of the SEP.

**Lemma 3.8**
*Assume $g$, $\mu$ and $\sigma$ to fulfill Assumption 3.5. Let the decoupling field $u$ of the FBSDE* (3.2) *have a continuous weak derivative $\partial_{x_1} u > 0$. Also let $(X^{(1)}, X^{(2)}, X^{(3)}, Y, Z)$, $\gamma$ and $B$ be defined as in Theorem 3.7. Moreover, let $\hat{B}$ be any Brownian motion coinciding with $B$ on $[0, X^{(2)}_1]$. Then $\gamma$, $W$, $X^{(3)}$ and $Y$ solve the system*

$$\begin{aligned} \gamma(t) &= \int_0^t \frac{\sigma^2\left(r, Y_{\gamma(r)} + X^{(3)}_{\gamma(r)}\right)}{(\partial_{x_1} u)^2\,(\gamma(r), W_{\gamma(r)}, r, X^{(3)}_{\gamma(r)})}\,\mathrm{d}r \\ W_{\gamma(t)} &= \int_0^t \frac{\sigma\left(r, Y_{\gamma(r)} + X^{(3)}_{\gamma(r)}\right)}{\partial_{x_1} u(\gamma(r), W_{\gamma(r)}, r, X^{(3)}_{\gamma(r)})}\,\mathrm{d}\hat{B}_r \\ X^{(3)}_{\gamma(t)} &= \int_0^t \mu\left(r, Y_{\gamma(r)} + X^{(3)}_{\gamma(r)}\right)\mathrm{d}r \\ Y_{\gamma(t)} &= Y_0 + \int_0^t \sigma\left(r, Y_{\gamma(r)} + X^{(3)}_{\gamma(r)}\right)\mathrm{d}\hat{B}_r \end{aligned} \tag{3.13}$$

*for all $t \geqslant 0$ such that $\gamma(t) \leqslant 1$. Additionally, for $\gamma^{-1}$ defined as in* (3.11) *we have*

$$\gamma^{-1}(1) \leqslant \frac{\|\partial_{x_1} u\|^2_\infty}{\varepsilon^2} \leqslant \varepsilon^{-2}\left(\frac{1}{\|g'\|^2_\infty} + 2\min\left\{0, \inf_{(\theta,x)\in\mathbb{R}_+\times\mathbb{R}} \left(\frac{\sigma\cdot\partial_a\mu - 2\partial_a\sigma\cdot\mu}{\sigma^3}\right)(\theta, x)\right\}\right)^{-1}. \tag{3.14}$$

*Proof.* Note that Theorem 3.6 implies the bound (3.14). Since $\partial_{x_1} u$ is continuous we get with Lemma 3.2 that $Z_s = \partial_{x_1} u(s, X_s^{(1)}, X_s^{(2)}, X_s^{(3)}) > 0$ for all $s \in [0,1]$ and hence both $\gamma$ and $\gamma^{-1}$ are strict monotone increasing and continuous. Moreover, Lemma 3.2, Equation (3.12) and the fact that $X^{(2)}_{\gamma(t)} = t$ yield

$$\gamma'(t) = \frac{\sigma^2\left(X^{(2)}_{\gamma(t)}, Y_{\gamma(t)} + X^{(3)}_{\gamma(t)}\right)}{Z^2_{\gamma(t)}} = \frac{\sigma^2\left(t, Y_{\gamma(t)} + X^{(3)}_{\gamma(t)}\right)}{(\partial_{x_1} u)^2\left(\gamma(t), X^{(1)}_{\gamma(t)}, X^{(2)}_{\gamma(t)}, X^{(3)}_{\gamma(t)}\right)}$$

for all $0 \leqslant t \leqslant \gamma^{-1}(1)$.

Furthermore, $X_s^{(1)} = W_s$ yields that

$$\begin{aligned} W_{\gamma(t)} = \int_0^t 1 \,\mathrm{d}W_{\gamma(r)} &= \int_0^t \frac{\sigma\left(X^{(2)}_{\gamma(r)}, Y_{\gamma(r)} + X^{(3)}_{\gamma(r)}\right)}{Z_{\gamma(r)}} \,\mathrm{d}\hat{B}_r \\ &= \int_0^t \frac{\sigma\left(r, Y_{\gamma(r)} + X^{(3)}_{\gamma(r)}\right)}{\partial_{x_1} u\left(\gamma(r), X^{(1)}_{\gamma(r)}, X^{(2)}_{\gamma(r)}, X^{(3)}_{\gamma(r)}\right)} \,\mathrm{d}\hat{B}_r \\ &= \int_0^t \frac{\sigma\left(r, Y_{\gamma(r)} + X^{(3)}_{\gamma(r)}\right)}{\partial_{x_1} u\left(\gamma(r), W_{\gamma(r)}, r, X^{(3)}_{\gamma(r)}\right)} \,\mathrm{d}\hat{B}_r. \end{aligned}$$

Also

$$\begin{aligned} Y_{\gamma(t)} = Y_0 + \int_0^{\gamma(t)} Z_r \,\mathrm{d}W_r &= Y_0 + \int_0^t Z_{\gamma(r)} \,\mathrm{d}W_{\gamma(r)} \\ &= Y_0 + \int_0^t \sigma\left(r, Y_{\gamma(r)} + X^{(3)}_{\gamma(r)}\right) \mathrm{d}\hat{B}_r, \end{aligned}$$

$$\begin{aligned} X^{(3)}_{\gamma(t)} &= \int_0^{\gamma(t)} \mu\left(X_r^{(2)}, Y_r + X_r^{(3)}\right) \frac{Z_r^2}{\sigma^2(X_r^{(2)}, Y_r + X_r^{(3)})} \,\mathrm{d}r \\ &= \int_0^t \mu\left(X^{(2)}_{\gamma(r)}, Y_{\gamma(r)} + X^{(3)}_{\gamma(r)}\right) \frac{Z^2_{\gamma(r)}}{\sigma^2\left(X^{(2)}_{\gamma(r)}, Y_{\gamma(r)} + X^{(3)}_{\gamma(r)}\right)} \frac{\sigma^2\left(X^{(2)}_{\gamma(r)}, Y_{\gamma(r)} + X^{(3)}_{\gamma(r)}\right)}{Z^2_{\gamma(r)}} \,\mathrm{d}r \\ &= \int_0^t \mu\left(r, Y_{\gamma(r)} + X^{(3)}_{\gamma(r)}\right) \mathrm{d}r \end{aligned}$$

and

$$\begin{aligned} \gamma(t) = \int_0^t \gamma'(r) \,\mathrm{d}r &= \int_0^t \frac{\sigma^2\left(r, Y_{\gamma(r)} + X^{(3)}_{\gamma(r)}\right)}{(\partial_{x_1} u)^2\left(\gamma(r), X^{(1)}_{\gamma(r)}, X^{(2)}_{\gamma(r)}, X^{(3)}_{\gamma(r)}\right)} \,\mathrm{d}r \\ &= \int_0^t \frac{\sigma^2\left(r, Y_{\gamma(r)} + X^{(3)}_{\gamma(r)}\right)}{(\partial_{x_1} u)^2\left(\gamma(r), W_{\gamma(r)}, r, X^{(3)}_{\gamma(r)}\right)} \,\mathrm{d}r \end{aligned}$$

for all $t \in [0, \gamma^{-1}(1)]$. ■

## 3.4 Strong solution

We use the definitions and constructions of the former sections. In particular let $u$ be the unique strongly regular decoupling field of the FBSDE (3.2) which exists on the whole interval $[0,1]$ if Assumption 3.5 is fulfilled.

**Theorem 3.9**
*Let $g$, $\mu$ and $\sigma$ fulfill Assumption 3.5 and $\mu$, $\sigma$ and their derivatives be bounded. Denote by $u$ the decoupling field of FBSDE* (3.2) *and assume the partial derivative $\partial_{x_1} u$ with respect to the first space variable to be Lipschitz continuous in every argument and $\partial_{x_1} u \geqslant \delta > 0$. Let $B$ be an arbitrary Brownian motion and denote by $(\mathcal{F}^B) = (\mathcal{F}^B_s)_{s\in[0,\infty)}$ the augmented filtration generated by $B$. Then there exists a bounded stopping time $\tau$ with respect to the filtration $\mathcal{F}^B$ such that for the process $A$ given by*

$$A_t = Y_0 + \int_0^t \mu\left(r, A_r\right) \mathrm{d}r + \int_0^t \sigma(r, A_r) \,\mathrm{d}B_r,$$

*for all $t \in [0,\tau]$, we have that $A_\tau \sim \nu$ and the stopping time $\tau$ satisfies*

$$\tau \leqslant \varepsilon^{-2}\left(\frac{1}{\|g'\|_\infty^2} + 2\min\left\{0, \inf_{(\theta,x)\in\mathbb{R}_+\times\mathbb{R}}\left(\frac{\sigma\cdot\partial_a\mu - 2\partial_a\sigma\cdot\mu}{\sigma^3}\right)(\theta,x)\right\}\right)^{-1} \quad a.s.$$

*By solving the Lipschitz SDE*

$$\begin{aligned} \gamma(r) &= \int_0^r \frac{\sigma^2(s,\Theta_s+\Delta_s)}{\left(\partial_{x_1}u(\gamma(s),\Gamma_s,s,\Delta_s)\right)^2}\,\mathrm{d}s \\ \Gamma_r &= \int_0^r \frac{\sigma(s,\Theta_s+\Delta_s)}{\partial_{x_1}u(\gamma(s),\Gamma_s,s,\Delta_s)}\,\mathrm{d}B_s \\ \Delta_r &= \int_0^r \mu(s,\Theta_s+\Delta_s)\,\mathrm{d}s \\ \Theta_r &= Y_0 + \int_0^r \sigma(s,\Theta_s+\Delta_s)\,\mathrm{d}B_s \end{aligned} \tag{3.15}$$

*for all $r \geqslant 0$ such that $\gamma(r) \leqslant 1$ and where $Y_0$ is the starting value of the process $Y$ in the FBSDE* (3.2) *and setting $\tau := \inf\{r \geqslant 0 | \gamma(r) = 1\}$ we can obtain such a stopping time.*

*Proof.* Since any solution of FBSDE (3.2) has a unique distribution independent of the driving Brownian motion, we know that the constant $Y_0$ is always the same and does not depend on the driving Brownian motion.

Let us take a look at the system (3.15). Note that for all $a, b \in [0,1]\times\mathbb{R}^3$

$$\left|\frac{1}{\partial_{x_1}u(a)} - \frac{1}{\partial_{x_1}u(b)}\right| = \left|\frac{\partial_{x_1}u(b) - \partial_{x_1}u(a)}{\partial_{x_1}u(a)\cdot\partial_{x_1}u(b)}\right| \leqslant \frac{L_{u,x_1}}{\delta^2}|b-a|,$$

yielding that $(\partial_{x_1}u)^{-1}$ is Lipschitz continuous. Since hence both $(\partial_{x_1}u)^{-1}$ and $\sigma$ are Lipschitz continuous and bounded we get that $\sigma\cdot(\partial_{x_1}u)^{-1}$ and $\sigma^2\cdot(\partial_{x_1}u)^{-2}$ are Lipschitz and bounded as well. Thus, we have that all coefficients of the system (3.15) are Lipschitz continuous. Therefore there exists a unique solution $(\gamma,\Gamma,\Delta,\Theta)$ of (3.15) which is progressively measurable w.r.t. $(\mathcal{F}^B_t)$. Hence $\tau := \inf\{r \geqslant 0 | \gamma(r) = 1\}$ is a stopping time w.r.t. $(\mathcal{F}^B_t)$ because $\gamma$ is continuous.

Furthermore, the systems (3.13) and (3.15) just differ by notation and the driving Brownian motion. By the principle of causality (see [KS91]) the distributions of $(\gamma, W_\gamma, X^{(3)}_\gamma, Y_\gamma)$ from

Lemma 3.8 and $(\gamma, \Gamma, \Delta, \Theta)$ are the same. Hence, we immediately have the bound for $\tau$ as stated in Lemma 3.8 and also for $A_t := \Delta_t + \Theta_t$ that

$$A_\tau = \Delta_\tau + \Theta_\tau = \Delta_{\gamma^{-1}(1)} + \Theta_{\gamma^{-1}(1)} \sim X^{(3)}_{\gamma(\gamma^{-1}(1))} + Y_{\gamma(\gamma^{-1}(1))} = X^{(3)}_1 + Y_1 = g(W_1) \sim \nu$$

and

$$\begin{aligned} A_t = \Delta_t + \Theta_t &= Y_0 + \int_0^t \mu(s, \Delta_s + \Theta_s)\,\mathrm{d}s + \int_0^t \sigma(s, \Delta_s + \Theta_s)\,\mathrm{d}B_s \\ &= Y_0 + \int_0^t \mu(s, A_s)\,\mathrm{d}s + \int_0^t \sigma(s, A_s)\,\mathrm{d}B_s. \end{aligned}$$

■

What remains to do is to find sufficient conditions for the assumptions of Theorem 3.9 to hold true. For this we use that the decoupling field $u$ of FBSDE (3.2) is three times weakly differentiable. To show this we extend FBSDE (3.2) by the dynamics of the gradient processes and view this system as a extended FBSDE, for which we can show the weak differentiability of its decoupling field.

Let $a := \max\left(\|\partial_{x_1} u\|_\infty, \|\partial_{x_2} u\|_\infty, \|\partial_{x_3} u\|_\infty\right)$ and define the truncation operator $T : \mathbb{R} \to \mathbb{R}$ by $T(z) := \min(\max(z, -a), a)$. Note that the map $T$ is uniformly Lipschitz. Assume that $g$, $\mu$, $\sigma$ and their first derivatives are Lipschitz continuous and consider the FBSDE

$$\begin{aligned}
X^{(1)}_s &= x^{(1)} + \int_t^s 1\,\mathrm{d}W_r, \\
X^{(2)}_s &= x^{(2)} + \int_t^s \frac{(Z^{(0)}_r)^2}{\sigma_r^2}\,\mathrm{d}r, \\
X^{(3)}_s &= x^{(3)} + \int_t^s \mu_r \frac{(Z^{(0)}_r)^2}{\sigma_r^2}\,\mathrm{d}r, \\
Y^{(0)}_s &= g(X^{(1)}_1) - X^{(3)}_1 - \int_s^1 Z^{(0)}_r\,\mathrm{d}W_r, \\
Y^{(1)}_s &= g'(X^{(1)}_1) + \int_s^1 T\left(Y^{(1)}_r\right) \frac{(Z^{(0)}_r)^2}{\sigma_r^2} \left( T\left(Y^{(3)}_r\right)\left(\mu_{a,r} - 2\mu_r \frac{\sigma_{a,r}}{\sigma_r}\right) - 2T\left(Y^{(2)}_r\right)\frac{\sigma_{a,r}}{\sigma_r}\right)\mathrm{d}r \\
&\qquad + \int_s^1 2\frac{Z^{(0)}_r}{\sigma_r^2} T\left(\left(Y^{(2)}_r\right) + T\left(Y^{(3)}_r\right)\mu_r\right) Z^{(1)}_r\,\mathrm{d}r - \int_s^1 Z^{(1)}_r\,\mathrm{d}W_r \\
Y^{(2)}_s &= 0 + \int_s^1 -2\frac{(Z^{(0)}_r)^2}{\sigma_r^2}\left(\frac{\sigma_{t,r}}{\sigma_r} + T\left(Y^{(2)}_r\right)\frac{\sigma_{a,r}}{\sigma_r}\right)\left(T\left(Y^{(2)}_r\right) + T\left(Y^{(3)}_r\right)\mu_r\right)\mathrm{d}r \\
&\qquad + \int_s^1 T\left(Y^{(3)}_r\right)\frac{(Z^{(0)}_r)^2}{\sigma_r^2}\left(T\left(Y^{(2)}_r\right)\mu_{a,r} + \mu_{t,r}\right)\mathrm{d}r \\
&\qquad + \int_s^1 2\frac{Z^{(0)}_r}{\sigma_r^2}\left(T\left(Y^{(2)}_r\right) + T\left(Y^{(3)}_r\right)\mu_r\right) Z^{(2)}_r\,\mathrm{d}r - \int_s^1 Z^{(2)}_r\,\mathrm{d}W_r \\
Y^{(3)}_r &= -1 + \int_s^1 \left(T\left(Y^{(3)}_r\right) + 1\right)\frac{(Z^{(0)}_r)^2}{\sigma_r^2}\left(T\left(Y^{(3)}_r\right)\mu_{a,s} - 2\frac{\sigma_{a,r}}{\sigma_r}\left(T\left(Y^{(2)}_r\right) + T\left(Y^{(3)}_r\right)\mu_r\right)\right)\mathrm{d}r \\
&\qquad + \int_s^1 2\frac{Z^{(0)}_r}{\sigma_r^2}\left(T\left(Y^{(2)}_r\right) + T\left(Y^{(3)}_r\right)\mu_r\right) Z^{(3)}_r\,\mathrm{d}r - \int_s^1 Z^{(3)}_r\,\mathrm{d}W_r
\end{aligned} \tag{3.16}$$

with the decoupling condition

$$\begin{aligned}
Y_s^{(0)} &= u^{(0)}(s, X_s^{(1)}, X_s^{(2)}, X_s^{(3)}),\\
Y_s^{(1)} &= u^{(1)}(s, X_s^{(1)}, X_s^{(2)}, X_s^{(3)}),\\
Y_s^{(2)} &= u^{(2)}(s, X_s^{(1)}, X_s^{(2)}, X_s^{(3)}),\\
Y_s^{(3)} &= u^{(3)}(s, X_s^{(1)}, X_s^{(2)}, X_s^{(3)}),
\end{aligned}$$

where

$$\mu_r := \mu\left(X_r^{(2)}, Y_r^{(0)} + X_r^{(3)}\right), \qquad \sigma_r := \sigma\left(X_r^{(2)}, Y_r^{(0)} + X_r^{(3)}\right)$$

and

$$\begin{aligned}
\mu_{t,r} &:= \partial_t \mu\left(X_r^{(2)}, Y_r^{(0)} + X_r^{(3)}\right), & \mu_{a,r} &:= \partial_a \mu\left(X_r^{(2)}, Y_r^{(0)} + X_r^{(3)}\right),\\
\sigma_{t,r} &:= \partial_t \sigma\left(X_r^{(2)}, Y_r^{(0)} + X_r^{(3)}\right), & \sigma_{a,r} &:= \partial_a \sigma\left(X_r^{(2)}, Y_r^{(0)} + X_r^{(3)}\right).
\end{aligned}$$

**Lemma 3.10**
*Let $g$, $\mu$ and $\sigma$ fulfill Assumption 3.5. In addition, suppose that $g$, $\mu$ and $\sigma$ are twice differentiable and that the second derivatives are bounded. Then, for the FBSDE* (3.16)*, we have $I_{\max}^M = [0,1]$ and there exists a unique, strongly regular Markovian decoupling field $(u^{(0)}, u^{(1)}, u^{(2)}, u^{(3)})$ on the whole interval $[0,1]$. Furthermore,*

$$u^{(0)} = u, \quad u^{(1)} = \partial_{x_1} u, \quad u^{(2)} = \partial_{x_2} u \quad \textit{and} \quad u^{(3)} = \partial_{x_3} u,$$

*a.e., where $u$ is the unique decoupling field to FBSDE* (3.2)*. In particular, $u$ is twice weakly differentiable w.r.t. the initial value $x$ with uniformly bounded derivatives.*

*Proof.* It is straightforward to verify that FBSDE (3.16) satisfies (MLLC), and hence Theorem 2.8 is applicable. Let $u^{(i)}$, $i = 0,1,2,3$ be the corresponding unique weakly regular Markovian decoupling field on $I_{\max}$. $u^{(i)}$, $i = 0,1,2,3$, are continuous functions on $I_{\max} \times \mathbb{R}^3$. In order to show that $I_{\max} = [0,1]$ we again need to prove that every partial derivative of $u^{(i)}$ for $i = 0,1,2,3$ is bounded independently with regard to the interval $[t,1] \subset I_{\max}$ where we consider it.

Let $t \in I_{\max}$. For an arbitrary initial condition $\bar{x} \in \mathbb{R}^3$ consider the corresponding processes

$$X^{(1)}, X^{(2)}, X^{(3)}, Y^{(0)}, Y^{(1)}, Y^{(2)}, Y^{(3)}, Z^{(0)}, Z^{(1)}, Z^{(2)}, Z^{(3)}$$

on $[t,1]$. Note that $X^{(1)}, X^{(2)}, X^{(3)}, Y^{(0)}, Z^{(0)}$ solve FBSDE (3.2), which implies that they coincide with the processes $X^{(1)}, X^{(2)}, X^{(3)}, Y, Z$ from (3.2) since strong regularity of Markovian decoupling fields guarantees uniqueness. Now $Y^{(0)} = Y$ implies $u(t', x') = u^{(0)}(t', x')$ for all $t' \in [t,1]$, $x' \in \mathbb{R}^3$.

Note that a truncation with $T$ does not effect any gradient process of FBSDE (3.2). Thus, $(Y_s^{(1)})$, $(Y_s^{(2)})$, $(Y_s^{(3)})$ fulfill the same dynamics resp. BSDEs as the gradient processes $(u_s^{(1)})$, $(u_s^{(2)})$, $(u_s^{(3)})$ in (3.4). Therefore, we can apply the same arguments and conclude that they also satisfy the estimates (3.8), (3.9) and (3.10) (see Theorem 3.6). In particular $Y_s^{(3)} = -1 = u_s^{(3)}$

for all $s \in [t,1]$ and therefore also $Z_s^{(3)} = 0 = \tilde{Z}_s^{(3)}$. Hence,

$$\begin{aligned}
Y_s^{(2)} - u_s^{(2)} &= \int_s^1 \left(\left(Y_r^{(2)}\right)^2 - \left(u_r^{(2)}\right)^2\right) \frac{(Z_r^{(0)})^2}{\sigma_r^2} \left(-2\frac{\sigma_{a,r}}{\sigma_r}\right) \mathrm{d}r \\
&\quad + \int_s^1 \left(Y_r^{(2)} - u_r^{(2)}\right) \frac{(Z_r^{(0)})^2}{\sigma_r^2} \left(-\mu_{a,r} + 2\frac{\sigma_{a,r}}{\sigma_r}\mu_r - 2\frac{\sigma_{t,r}}{\sigma_r}\right) \mathrm{d}r \\
&\quad + \int_s^1 \frac{2Z_r^{(0)}}{\sigma_r^2} \left(\left(Y_r^{(2)} - \mu_r\right) Z_r^{(2)} - \left(u_r^{(2)} - \mu_r\right) \tilde{Z}_r^{(2)}\right) \mathrm{d}r - \int_s^1 \left(Z_r^{(2)} - \tilde{Z}_r^{(2)}\right) \mathrm{d}W_r \\
&= \int_s^1 \left(\left(Y_r^{(2)}\right)^2 - \left(u_r^{(2)}\right)^2\right) \frac{(Z_r^{(0)})^2}{\sigma_r^2} \left(-2\frac{\sigma_{a,r}}{\sigma_r}\right) \mathrm{d}r - \int_s^1 \left(Z_r^{(2)} - \tilde{Z}_r^{(2)}\right) \mathrm{d}W_r \\
&\quad + \int_s^1 \left(Y_r^{(2)} - u_r^{(2)}\right) \frac{(Z_r^{(0)})^2}{\sigma_r^2} \left(-\mu_{a,r} + 2\frac{\sigma_{a,r}}{\sigma_r}\mu_r - 2\frac{\sigma_{t,r}}{\sigma_r}\right) \mathrm{d}r \\
&\quad + \int_s^1 \frac{2Z_r^{(0)}}{\sigma_r^2} \left(\left(Y_r^{(2)} - u_r^{(2)}\right) Z_r^{(2)} + \left(u_r^{(2)} - \mu_r\right)\left(Z_r^{(2)} - \tilde{Z}_r^{(2)}\right)\right) \mathrm{d}r
\end{aligned}$$

Since $\frac{2Z_r^{(0)}}{\sigma_r^2}\left(u_r^{(2)} - \mu_r\right)$ is bounded we have that $\widetilde{W}_s := W_s - W_t - \int_t^s \frac{2Z_r^{(0)}}{\sigma_r^2}\left(u_r^{(2)} - \mu_r\right) \mathrm{d}r$, $s \in [t,1]$, is a Brownian motion under some probability measure equivalent to $\mathbf{P}$. Under the new measure the process pair $(Y_s^{(2)} - u_s^{(2)}, Z_s^{(2)} - \tilde{Z}_s^{(2)})$ is a solution of the following linear BSDE with bounded coefficients

$$\begin{aligned}
\hat{Y}_s &= \int_s^1 \hat{Y}_r \left(Y_r^{(2)} + u_r^{(2)}\right) \frac{(Z_r^{(0)})^2}{\sigma_r^2} \left(-2\frac{\sigma_{a,r}}{\sigma_r}\right) \mathrm{d}r \\
&\quad + \int_s^1 \hat{Y}_r \frac{(Z_r^{(0)})^2}{\sigma_r^2} \left(-\mu_{a,r} + 2\frac{\sigma_{a,r}}{\sigma_r}\mu_r - 2\frac{\sigma_{t,r}}{\sigma_r}\right) \mathrm{d}r \\
&\quad + \int_s^1 \hat{Y}_r \frac{2Z_r^{(0)}}{\sigma_r^2} Z_r^{(2)} \, \mathrm{d}r - \int_s^1 \hat{Z}_r \, \mathrm{d}\widetilde{W}_r.
\end{aligned}$$

Note that $(0,0)$ is the unique solution of the previous BSDE. Consequently, $Y^{(2)}$ and $u^{(2)}$ are indistinguishable and $Z^{(2)} = \tilde{Z}^{(2)}$, $\lambda \otimes P$-almost everywhere on $[t,1] \times \Omega$.

Similarly we can show that $Y^{(1)}$ and $u^{(1)}$ as well as $Z^{(1)}$ and $\tilde{Z}^{(1)}$ coincide. Thus we have

$$\begin{aligned}
\partial_{x_1} u^{(0)}(s, X_s^{(1)}, X_s^{(2)}, X_s^{(3)}) &= \partial_{x_1} u(s, X_s^{(1)}, X_s^{(2)}, X_s^{(3)}) = u_s^{(1)} = Y_s^{(1)}, \\
\partial_{x_2} u^{(0)}(s, X_s^{(1)}, X_s^{(2)}, X_s^{(3)}) &= \partial_{x_2} u(s, X_s^{(1)}, X_s^{(2)}, X_s^{(3)}) = u_s^{(2)} = Y_s^{(2)}, \\
\partial_{x_3} u^{(0)}(s, X_s^{(1)}, X_s^{(2)}, X_s^{(3)}) &= \partial_{x_3} u(s, X_s^{(1)}, X_s^{(2)}, X_s^{(3)}) = u_s^{(3)} = Y_s^{(3)}
\end{aligned}$$

a.e. on $[t,1]$.

It remains to show that $I_{\max}^M = [0,1]$. Define for $x = (x_1, x_2, x_3)^T \in \mathbb{R}^3$, $y = (y_0, y_1, y_2, y_3)^T \in \mathbb{R}^4$, $z = (z_0, z_1, z_2, z_3)^T \in \mathbb{R}^4$

$$\bar{X}_s := \begin{pmatrix} X_s^{(1)} \\ X_s^{(2)} \\ X_s^{(3)} \end{pmatrix}, \qquad \bar{Y}_s := \begin{pmatrix} Y_s^{(0)} \\ Y_s^{(1)} \\ Y_s^{(2)} \\ Y_s^{(3)} \end{pmatrix}, \qquad \bar{Z}_s := \begin{pmatrix} Z_s^{(0)} \\ Z_s^{(1)} \\ Z_s^{(2)} \\ Z_s^{(3)} \end{pmatrix}$$

$$\bar{M}(x,y,z) := \begin{pmatrix} 0 \\ \frac{z_0^2}{\sigma^2(x_2, y_0 + x_3)} \\ \mu(x_2, y_0 + x_3) \frac{z_0^2}{\sigma^2(x_2, y_0 + x_3)} \end{pmatrix}, \qquad \bar{\Sigma} := \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \qquad \bar{\xi}(x) := \begin{pmatrix} g(x_1) - x_3 \\ g'(x_1) \\ 0 \\ -1 \end{pmatrix}$$

and

$$\bar{F}(x,y,z)$$

$$:= \begin{pmatrix} 0 \\ y_1 \frac{(z_0)^2}{\sigma^2(x_2,y_0+x_3)} \left( \partial_a \mu(x_2, y_0+x_3) - 2\mu(x_2,y_0+x_3) \frac{\partial_a \sigma(x_2,y_0+x_3)}{\sigma(x_2,y_0+x_3)} + 2y_2 \frac{\partial_a \sigma(x_2,y_0+x_3)}{\sigma(x_2,y_0+x_3)} \right) \\ 2\frac{(z_0)^2}{\sigma^2(x_2,y_0+x_3)} \left( \frac{\partial_t \sigma(x_2,y_0+x_3)}{\sigma(x_2,y_0+x_3)} + y_2 \frac{\partial_a \sigma(x_2,y_0+x_3)}{\sigma(x_2,y_0+x_3)} \right) (y_2 - \mu(x_2,y_0+x_3)) \\ 0 \end{pmatrix}$$

$$+ \begin{pmatrix} 0 \\ 0 \\ \frac{(z_0)^2}{\sigma^2(x_2,y_0+x_3)} \left( y_2 \partial_a \mu(x_2,y_0+x_3) + \partial_t \mu(x_2,y_0+x_3) \right) \\ 0 \end{pmatrix}$$

$$+ \begin{pmatrix} 0 \\ -\frac{2z_0}{\sigma^2(x_2,y_0+x_3)} (y_2 - \mu(x_2,y_0+x_3)) z_1 \\ -\frac{2z_0}{\sigma^2(x_2,y_0+x_3)} (y_2 - \mu(x_2,y_0+x_3)) z_2 \\ 0 \end{pmatrix}.$$

Then

$$\bar{X}_s = \bar{x} + \int_t^s \bar{M}\left(\bar{X}_r, \bar{Y}_r, \bar{Z}_r\right) \mathrm{d}r + \int_t^s \bar{\Sigma} \, \mathrm{d}W_r$$

and

$$\bar{Y}_s = \bar{\xi}\left(\bar{X}_1\right) - \int_s^1 \bar{F}\left(\bar{X}_r, \bar{Y}_r, \bar{Z}_r\right) \mathrm{d}r - \int_s^1 \bar{Z}_r \, \mathrm{d}W_r.$$

By setting

$$\bar{U}_s := \partial_x \begin{pmatrix} u^{(0)} \\ u^{(1)} \\ u^{(2)} \\ u^{(3)} \end{pmatrix} \left(s, \bar{X}_s\right) = \begin{pmatrix} u^{(1)} & u^{(2)} & u^{(3)} \\ \partial_{x_1} u^{(1)} & \partial_{x_2} u^{(1)} & \partial_{x_3} u^{(1)} \\ \partial_{x_1} u^{(2)} & \partial_{x_2} u^{(2)} & \partial_{x_3} u^{(2)} \\ \partial_{x_1} u^{(3)} & \partial_{x_2} u^{(3)} & \partial_{x_3} u^{(3)} \end{pmatrix} \left(s, \bar{X}_s\right)$$

we get

$$\partial_x \bar{Y}_s = \bar{U}_s \cdot \partial_x \bar{X}_s.$$

Since $(\partial_x \bar{X}_s)^{-1}$ is a multidimensional Itô process on $[t,1]$ (see Lemma 3.1 and its proof) we get that $\bar{U}_s = \partial_x \bar{Y}_s \cdot (\partial_x \bar{X}_s)^{-1}$ is also an Itô process and hence there exist $(b_s)$ and $(\hat{Z}_s)$ such that

$$\bar{U}_s = \bar{U}_1 - \int_s^1 b_r \, \mathrm{d}r - \int_s^1 \hat{Z}_r \, \mathrm{d}W_r.$$

For the following we also introduce for an Itô process $I_s = I_0 - \int_0^s i_r \, \mathrm{d}r - \int_0^s j_r \, \mathrm{d}W_r$ the two operators $\mathrm{D}^t$ and $\mathrm{D}^w$ via $(\mathrm{D}^t I)_s := i_s$ and $(\mathrm{D}^w I)_s := j_s$. Using this notation we have

$$\begin{aligned} \partial_x \bar{Z}_s &= \mathrm{D}^w \, \partial_x \bar{Y}_s \\ &= \mathrm{D}^w \left(\bar{U}_s \cdot \partial_x \bar{X}_s\right) \\ &= \bar{U}_s \cdot \mathrm{D}^w \, \partial_x \bar{X}_s + \mathrm{D}^w \, \bar{U}_s \cdot \partial_x \bar{X}_s \\ &= \hat{Z}_s \cdot \partial_x \bar{X}_s, \end{aligned}$$

$$\begin{aligned} \partial_x \left[\bar{F}\left(\bar{X}_s, \bar{Y}_s, \bar{Z}_s\right)\right] &= \mathrm{D}^t \, \partial_x \bar{Y}_s \\ &= \mathrm{D}^t \left(\bar{U}_s \cdot \partial_x \bar{X}_s\right) \\ &= \bar{U}_s \cdot \mathrm{D}^t \, \partial_x \bar{X}_s + \mathrm{D}^t \, \bar{U}_s \cdot \partial_x \bar{X}_s + \mathrm{D}^w \, \bar{U}_s \cdot \mathrm{D}^w \, \partial_x \bar{X}_s \\ &= \bar{U}_s \cdot \partial_x \left[\bar{M}\left(\bar{X}_s, \bar{Y}_s, \bar{Z}_s\right)\right] + b_s \cdot \partial_x \bar{X}_s, \end{aligned}$$

where we can further specify

$$\begin{aligned}
&\partial_x \left[\bar{M}\left(\bar{X}_s, \bar{Y}_s, \bar{Z}_s\right)\right] \\
&\quad = \partial_x \bar{M}\left(\bar{X}_s, \bar{Y}_s, \bar{Z}_s\right) \partial_x \bar{X}_s + \partial_y \bar{M}\left(\bar{X}_s, \bar{Y}_s, \bar{Z}_s\right) \partial_x \bar{Y}_s + \partial_z \bar{M}\left(\bar{X}_s, \bar{Y}_s, \bar{Z}_s\right) \partial_x \bar{Z}_s \\
&\quad = \partial_x \bar{M}\left(\bar{X}_s, \bar{Y}_s, \bar{Z}_s\right) \partial_x \bar{X}_s + \partial_y \bar{M}\left(\bar{X}_s, \bar{Y}_s, \bar{Z}_s\right) \bar{U}_s \partial_x \bar{X}_s + \partial_z \bar{M}\left(\bar{X}_s, \bar{Y}_s, \bar{Z}_s\right) \hat{Z}_s \partial_x \bar{X}_s
\end{aligned}$$

and likewise

$$\begin{aligned}
&\partial_x \left[\bar{F}\left(\bar{X}_s, \bar{Y}_s, \bar{Z}_s\right)\right] \\
&\quad = \partial_x \bar{F}\left(\bar{X}_s, \bar{Y}_s, \bar{Z}_s\right) \partial_x \bar{X}_s + \partial_y \bar{F}\left(\bar{X}_s, \bar{Y}_s, \bar{Z}_s\right) \bar{U}_s \partial_x \bar{X}_s + \partial_z \bar{F}\left(\bar{X}_s, \bar{Y}_s, \bar{Z}_s\right) \hat{Z}_s \partial_x \bar{X}_s .
\end{aligned}$$

Thus we get

$$\hat{Z}_s = \partial_x \bar{Z}_s \cdot \left(\partial_x \bar{X}_s\right)^{-1}$$

and

$$\begin{aligned}
b_s = {} & \partial_x \bar{F}\left(\bar{X}_s, \bar{Y}_s, \bar{Z}_s\right) + \partial_y \bar{F}\left(\bar{X}_s, \bar{Y}_s, \bar{Z}_s\right) \bar{U}_s + \partial_z \bar{F}\left(\bar{X}_s, \bar{Y}_s, \bar{Z}_s\right) \hat{Z}_s \\
& + \bar{U}_s \left[\partial_x \bar{M}\left(\bar{X}_s, \bar{Y}_s, \bar{Z}_s\right) + \partial_y \bar{M}\left(\bar{X}_s, \bar{Y}_s, \bar{Z}_s\right) \bar{U}_s + \partial_z \bar{M}\left(\bar{X}_s, \bar{Y}_s, \bar{Z}_s\right) \hat{Z}_s\right],
\end{aligned} \tag{3.17}$$

where the derivatives of $\bar{M}$ and $\bar{F}$ are bounded due to the assumptions made. Therefore, we see that the dynamics of $\bar{U}$ are linear with exception to the quadratic terms $\bar{U}_s \partial_y \bar{M}(\bar{X}_s, \bar{Y}_s, \bar{Z}_s) \bar{U}_s$ and $\partial_z \bar{M}\left(\bar{X}_s, \bar{Y}_s, \bar{Z}_s\right) \hat{Z}_s$. However, we claim that we can reduce the dynamics of $\bar{U}$ to a linear BSDE.

It is straightforward to see that

$$\partial_y \bar{M}(\bar{X}_s, \bar{Y}_s, \bar{Z}_s) = \begin{pmatrix} 0 & 0 & 0 & 0 \\ -2\frac{(Z_s^{(0)})^2}{\sigma_s^2}\frac{\sigma_{a,s}}{\sigma_s} & 0 & 0 & 0 \\ \frac{(Z_s^{(0)})^2}{\sigma_s^2}\mu_{a,s} - 2\frac{(Z_s^{(0)})^2}{\sigma_s^2}\frac{\sigma_{a,s}}{\sigma_s}\mu_s & 0 & 0 & 0 \end{pmatrix}.$$

Note that $\alpha := -2\frac{(Z_s^{(0)})^2}{\sigma_s^2}\frac{\sigma_{a,s}}{\sigma_s}$ and $\beta := \frac{(Z_s^{(0)})^2}{\sigma_s^2}\mu_{a,s} - 2\frac{(Z_s^{(0)})^2}{\sigma_s^2}\frac{\sigma_{a,s}}{\sigma_s}\mu_s$ are both uniformly bounded, and we have

$$\partial_y \bar{M}\left(\bar{X}_s, \bar{Y}_s, \bar{Z}_s\right) \bar{U}_s = \begin{pmatrix} 0 & 0 & 0 \\ \alpha \cdot u_s^{(1)} & \alpha \cdot u_s^{(2)} & \alpha \cdot u_s^{(3)} \\ \beta \cdot u_s^{(1)} & \beta \cdot u_s^{(2)} & \beta \cdot u_s^{(3)} \end{pmatrix},$$

which is bounded independently of $[t, 1]$ (cf. in Theorem 3.6).

Moreover, note that

$$\partial_z \bar{M}\left(\bar{X}_s, \bar{Y}_s, \bar{Z}_s\right) = \begin{pmatrix} 0 & 0 & 0 & 0 \\ \frac{2Z_s^{(0)}}{\sigma_s^2} & 0 & 0 & 0 \\ \mu_s \frac{2Z_s^{(0)}}{\sigma_s^2} & 0 & 0 & 0 \end{pmatrix}$$

only depends on the solution components $(X^{(2)}, X^{(3)}, Y^{(0)}, Z^{(0)})$. Hence, together with the estimates of Theorem 3.6, we conclude that $\partial_x \bar{M}(\bar{X}_s, \bar{Y}_s, \bar{Z}_s)$ is bounded. Since $\bar{U}$ is bounded on $[t, 1]$, the term $\bar{U}_s \partial_z \bar{M}\left(\bar{X}_s, \bar{Y}_s, \bar{Z}_s\right) \hat{Z}_s$ in Equation (3.17) can be shifted, via a Girsanov measure change, into the Brownian motion $W$. Similary, the term $\partial_z \bar{F}\left(\bar{X}_s, \bar{Y}_s, \bar{Z}_s\right) \hat{Z}_s$ in Equation (3.17) can be shifted into $W$. To sum up, there exists a Brownian motion $\hat{W}$ under an equivalent probability measure such that $(\bar{U}, \hat{Z})$ solves the BSDE on $[t, 1]$ driven by $\hat{W}$ with linear driver

$$f(s, y, z) = \partial_x \bar{F}\left(\bar{X}_s, \bar{Y}_s, \bar{Z}_s\right) + \partial_y \bar{F}\left(\bar{X}_s, \bar{Y}_s, \bar{Z}_s\right) y + y \left[\partial_x \bar{M}\left(\bar{X}_s, \bar{Y}_s, \bar{Z}_s\right) + \partial_y \bar{M}\left(\bar{X}_s, \bar{Y}_s, \bar{Z}_s\right) \bar{U}_s\right]$$

and terminal condition $\nabla\bar{\xi}(\bar{X}_1)$. Observe that the terminal condition and all coefficients are bounded by some constant independent of $t$ and $x$. Therefore, also $\bar{U}$ is bounded independently of $t$ and $x$. By Lemma 2.9 this yields that $I_{\max} = [0, 1]$. ■

**Remark 3.11**
The second and third derivatives do not have to be bounded. It would suffice if the second and third derivatives of $\mu$ divided by $\sigma^2$ and the second and third derivatives of $\sigma$ divided by $\sigma$ are bounded.

**Lemma 3.12**
*Let $g$, $\mu$ and $\sigma$ fulfill Assumption 3.5 and their second and third derivatives be bounded. Then the decoupling field $u$ of FBSDE* (3.2) *is three times weakly differentiable w.r.t. to the initial condition $x \in \mathbb{R}^3$ with uniformly bounded derivatives.*

*Proof.* This proof is completely analogous the proof of Lemma 3.10. Therefore, we only give a sketch.

Extend the system (3.16) by the dynamics of $\bar{Y}^{(ij)} := u^{(ij)} := \partial_{x_j} u^{(i)}$ for all $i, j \in \{1, 2, 3\}$ as obtained in the proof of Lemma 3.10 and by the corresponding entries in the decoupling field. Then argue analogously to the proof of Lemma 3.10 that for every $i \in \{0, 1, 2, 3\}$ the $u^{(i)}$ of FBSDE (3.16) coincides with the $u^{(i)}$ of the extended system. Redefine, if necessary, the vectors $\bar{X}$, $\bar{Y}$,$\bar{Z}$ and the functions $\bar{M}$, $\bar{\Sigma}$, $\bar{\xi}$, $\bar{F}$ such that for the extended system we have

$$\bar{X}_s = x + \int_t^s \bar{M}\left(\bar{X}_r, \bar{Y}_r, \bar{Z}_r\right) \mathrm{d}r + \int_t^s \bar{\Sigma}\, \mathrm{d}W_r$$

and

$$\bar{Y}_s = \bar{\xi}\left(\bar{X}_1\right) - \int_s^1 \bar{F}\left(\bar{X}_r, \bar{Y}_r, \bar{Z}_r\right) \mathrm{d}r - \int_s^1 \bar{Z}_r\, \mathrm{d}W_r.$$

Also define $\bar{U}_s$ as the partial derivatives of the decoupling field $u(s, \bar{X}_s)$ of the extended system for all $s \in [t, 1]$. Again there exist $(b_s)$ and $(\hat{Z}_s)$ such that

$$\bar{U}_s = \bar{U}_1 - \int_s^1 b_r\, \mathrm{d}r - \int_s^1 \hat{Z}_r\, \mathrm{d}W_r.$$

By the same calculation as in the proof of Lemma 3.10 we obtain that

$$\hat{Z}_s = \partial_x \bar{Z}_s \cdot \left(\partial_x \bar{X}_s\right)^{-1}$$

and

$$\begin{aligned} b_s = {} & \partial_x \bar{F}\left(\bar{X}_s, \bar{Y}_s, \bar{Z}_s\right) + \partial_y \bar{F}\left(\bar{X}_s, \bar{Y}_s, \bar{Z}_s\right) \bar{U}_s + \partial_z \bar{F}\left(\bar{X}_s, \bar{Y}_s, \bar{Z}_s\right) \hat{Z}_s \\ & + \bar{U}_s \left[\partial_x \bar{M}\left(\bar{X}_s, \bar{Y}_s, \bar{Z}_s\right) + \partial_y \bar{M}\left(\bar{X}_s, \bar{Y}_s, \bar{Z}_s\right) \bar{U}_s + \partial_z \bar{M}\left(\bar{X}_s, \bar{Y}_s, \bar{Z}_s\right) \hat{Z}_s\right]. \end{aligned}$$

Analogous to the proof above, $\partial_x \bar{F}$, $\partial_y \bar{F}$, $\partial_z \bar{F}$, $\partial_x \bar{M}$, $\partial_y \bar{M}$ and $\partial_z \bar{M}$ are bounded while additionally $\partial_y \bar{M}$ only has entries in the first column which allows us to conclude that $\partial_y \bar{M}(\bar{X}_s, \bar{Y}_s, \bar{Z}_s)\bar{U}_s$ is bounded. Furthermore every coefficient in front of $\hat{Z}$ is bounded on every Interval $[t, 1] \subset I^M_{\max}$ and can therefore be transformed away with Girsanov's Theorem. Hence we have linear dynamics for $\bar{U}$ with bounded coefficients which yields that it is bounded independently of the interval $[t, 1]$, giving $I^M_{\max} = [0, 1]$. ■

**Lemma 3.13**
*Let $g$, $\mu$ and $\sigma$ fulfill Assumption 3.5, their first and second derivatives be bounded and $g' \geqslant \delta > 0$. Then the weak derivative $\partial_{x_1} u$ of the decoupling field $u$ from the FBSDE* (3.2) *fulfills*

$$\left\| \frac{1}{\partial_{x_1} u} \right\|_\infty \leqslant \left\| \frac{1}{g'} \right\|_\infty + \|\partial_{x_1} u\|_\infty \left( \left\| \frac{\partial_a \mu}{\sigma^2} \right\|_\infty + 2 \left\| \frac{\mu}{\sigma^2} \right\|_\infty \left\| \frac{\partial_a \sigma}{\sigma} \right\|_\infty + \frac{2}{\varepsilon^2} \|\partial_{x_2} u\|_\infty \left\| \frac{\partial_a \sigma}{\sigma} \right\|_\infty \right) \tag{3.18}$$

*and in particular $\partial_{x_1} u$ is bounded away from* $0$.

*Proof.* By Lemma 3.10 the decoupling field of the FBSDE (3.2) exists on the whole interval $[0,1]$ and is twice weakly differentiable. In particular $\partial_{x_1} u$ is continuous (see e.g. Theorem 4.2.17 in [Fro15]), and hence we can apply Lemma 3.2 yielding $Z_r^{(0)} = \partial_{x_1} u \left( r, X_r^{(1)}, X_r^{(2)}, X_r^{(3)} \right)$ for all $r \in [0,1]$. Also using Lemma 3.4 we know that $u^{(1)}$ is bounded by some constant for every starting time $t \in I_{\max}^M = [0,1]$ and every initial value $x \in \mathbb{R}^3$.

Now we set $V_r := \frac{1}{\partial_{x_1} u \left( r, X_r^{(1)}, X_r^{(2)}, X_r^{(3)} \right)}$ for all $r \in (t_0, 1]$ where

$$t_0 := \inf\{ t \geqslant 0 | \partial_{x_1} u(t,x) = 0 \text{ for at least one } x \in \mathbb{R}^3 \}$$

with the convention that $\inf \varnothing = 0$. We immediately get that $\frac{1}{V_r} \leqslant \|\partial_{x_1} u\|_\infty < \infty$ and the dynamics

$$\begin{aligned} V_s &= \frac{1}{g'(X_1^{(1)})} - \int_s^1 \left( V_r^3 \left( Z_r^{(1)} \right)^2 - \frac{1}{V_r} \frac{\mu_{a,r} - 2\mu_r \frac{\sigma_{a,r}}{\sigma_r} + 2u_r^{(2)} \frac{\sigma_{a,r}}{\sigma_r}}{\sigma_r^2} \right) \mathrm{d}r - \int_s^1 -Z_r^{(1)} V_r^2 \, \mathrm{d}\widetilde{W}_r \\ &= \frac{1}{g'(X_1^{(1)})} - \int_s^1 \frac{1}{V_r} \left( \left( \hat{Z}_r \right)^2 - \frac{\mu_{a,r} - 2\mu_r \frac{\sigma_{a,r}}{\sigma_r} + 2u_r^{(2)} \frac{\sigma_{a,r}}{\sigma_r}}{\sigma_r^2} \right) \mathrm{d}r - \int_s^1 \hat{Z}_r \, \mathrm{d}\widetilde{W}_r, \end{aligned}$$

where $u_r^{(2)} := \partial_{x_2} u \left( r, X_r^{(1)}, X_r^{(2)}, X_r^{(3)} \right)$, $\hat{Z}_r := -\frac{Z_r^{(1)}}{V_r^2}$ and $\tilde{W}$ is defined as in the proof of Lemma 3.10.

Using that $\frac{1}{V_s} \leqslant \|\partial_{x_1} u\|_\infty$ we can apply Corollary 2.2 of [Kob00] to obtain

$$\|V\|_\infty \leqslant \left\| \frac{1}{g'} \right\|_\infty + \|\partial_{x_1} u\|_\infty \left( \left\| \frac{\partial_a \mu}{\sigma^2} \right\|_\infty + 2 \left\| \frac{\mu}{\sigma^2} \right\|_\infty \left\| \frac{\partial_a \sigma}{\sigma} \right\|_\infty + \frac{2}{\varepsilon^2} \|\partial_{x_2} u\|_\infty \left\| \frac{\partial_a \sigma}{\sigma} \right\|_\infty \right) < \infty$$

because $\partial_{x_1} u$ and $\partial_{x_2} u$ are bounded by Theorem 3.6. Since this bound is independent of $s$ we also get that

$$\partial_{x_1} u \left( s, X_s^{(1)}, X_s^{(2)}, X_s^{(3)} \right) = \frac{1}{V_s} \geqslant \frac{1}{\|V\|_\infty} > 0$$

for all $s$ where $V$ is defined. Because, as stated above, $\partial_{x_1} u$ is continuous, we get that $t_0 = 0$ and that hence Equation (3.18) holds true. ∎

**Lemma 3.14**
*Let $g$, $\mu$ and $\sigma$ fulfill Assumption 3.5 and their second derivatives be bounded. Then for the problem* (3.16) *it holds for all $s \in [0,1]$ almost surely that*

$$|Z_s^{(1)}| \leqslant \|\partial_{x_1} u^{(1)}\|_\infty < \infty.$$

*Proof.* Note that this proof runs on similar lines as the proof of Lemma 3.2.

Remember that Lemma 3.10 yields that for problem (3.16) there exists a unique solution on the whole interval $[0,1]$ for every initial condition in $\mathbb{R}^3$. Observe that with Itô's formula we get for $h>0$ and $s, s+h \in [0,1]$

$$\begin{aligned}
&\frac{1}{h}\mathbb{E}\left[ Y^{(1)}_{s+h}(W_{s+h}-W_s)\middle| \mathcal{F}_s\right] \\
=&\frac{1}{h}\mathbb{E}\left[\int_s^{s+h} Y^{(1)}_r \,\mathrm{d}W_r + \int_s^{s+h}(W_r-W_s)\,\mathrm{d}Y^{(1)}_r + \int_s^{s+h} 1\cdot Z^{(1)}_r\,\mathrm{d}r\middle| \mathcal{F}_s\right] \\
=&\frac{1}{h}\mathbb{E}\left[\int_s^{s+h} Z^{(1)}_r\,\mathrm{d}r + \int_s^{s+h} Y^{(1)}_r\,\mathrm{d}W_r + \int_s^{s+h}(W_r-W_s)Z^{(1)}_r\,\mathrm{d}W_r\right.\\
&\quad + \int_s^{s+h}(W_r-W_s)\left(-\frac{(Z^{(0)}_r)^2}{\sigma_r^2}Y^{(1)}_r\left(Y^{(3)}_r\left(\mu_{a,r}-2\mu_r\frac{\sigma_{a,r}}{\sigma_r}\right)-2Y^{(2)}_r\frac{\sigma_{a,r}}{\sigma_r}\right)\right)\mathrm{d}r \\
&\quad \left.+ \int_s^{s+h}(W_r-W_s)\left(-\frac{2Z^{(0)}_r}{\sigma_r^2}\left(Y^{(2)}_r+Y^{(3)}_r\mu_r\right)Z^{(1)}_r\right)\mathrm{d}r\middle| \mathcal{F}_s\right] \\
\to& Z^{(1)}_s \quad a.s. \quad \text{as} \quad h\to 0.
\end{aligned}$$

On the other hand we get by using the decoupling condition that

$$\begin{aligned}
&Y^{(1)}_{s+h}(W_{s+h}-W_s) \\
=&u^{(1)}\left(s+h, X^{(1)}_{s+h}, X^{(2)}_{s+h}, X^{(3)}_{s+h}\right)(W_{s+h}-W_s) \\
=&u^{(1)}\left(s+h, X^{(1)}_{s+h}, X^{(2)}_{s}, X^{(3)}_{s}\right)(W_{s+h}-W_s) \\
&+\left(u^{(1)}\left(s+h, X^{(1)}_{s+h}, X^{(2)}_{s+h}, X^{(3)}_{s}\right)-u^{(1)}\left(s+h, X^{(1)}_{s+h}, X^{(2)}_{s}, X^{(3)}_{s}\right)\right)(W_{s+h}-W_s) \\
&+\left(u^{(1)}\left(s+h, X^{(1)}_{s+h}, X^{(2)}_{s+h}, X^{(3)}_{s+h}\right)-u^{(1)}\left(s+h, X^{(1)}_{s+h}, X^{(2)}_{s+h}, X^{(3)}_{s}\right)\right)(W_{s+h}-W_s).
\end{aligned} \tag{3.19}$$

At first let us take a look at the third summand at the right hand side of (3.19). Since $u^{(1)}$ is Lipschitz continuous in its fourth argument with some constant $L^t_{u^{(1)},x_3}$ and since furthermore

$$X^{(3)}_{s+h} = X^{(3)}_s + \int_s^{s+h}\mu_r\frac{(Z^{(0)}_r)^2}{\sigma_r^2}\,\mathrm{d}r$$

we can estimate

$$\begin{aligned}
&\frac{1}{h}\left|\mathbb{E}\left[\left(u^{(1)}\left(s+h, X^{(1)}_{s+h}, X^{(2)}_{s+h}, X^{(3)}_{s+h}\right)-u^{(1)}\left(s+h, X^{(1)}_{s+h}, X^{(2)}_{s+h}, X^{(3)}_{s}\right)\right)(W_{s+h}-W_s)\middle| \mathcal{F}_s\right]\right| \\
&\leqslant \frac{1}{h}\mathbb{E}\left[L^t_{u^{(1)},x_3}\left|\int_s^{s+h}\mu_r\frac{(Z^{(0)}_r)^2}{\sigma_r^2}\,\mathrm{d}r\right||W_{s+h}-W_s|\middle| \mathcal{F}_s\right] \\
&\leqslant \frac{1}{h}L^t_{u^{(1)},x_3}h\left\|\frac{\mu}{\sigma^2}\right\|_\infty\|Z^{(0)}\|^2_\infty\mathbb{E}\left[|W_{s+h}-W_s|\,|\,\mathcal{F}_s\right],
\end{aligned}$$

which clearly goes to 0 as $h\to 0$ because $\frac{\mu}{\sigma^2}$ and $Z^{(0)}$ are bounded by Theorem 3.6. Analogously we get, with $L^t_{u^{(1)},x_2}$ being the Lipschitz constant of $u^{(1)}$ in the third argument, that

$$\begin{aligned}
&\frac{1}{h}\left|\mathbb{E}\left[\left(u^{(1)}\left(s+h, X^{(1)}_{s+h}, X^{(2)}_{s+h}, X^{(3)}_{s}\right)-u^{(1)}\left(s+h, X^{(1)}_{s+h}, X^{(2)}_{s}, X^{(3)}_{s}\right)\right)(W_{s+h}-W_s)\middle| \mathcal{F}_s\right]\right| \\
&\leqslant \frac{1}{h}L^t_{u^{(1)},x_2}h\|Z^{(0)}\|^2_\infty\varepsilon^{-2}\mathbb{E}\left[|W_{s+h}-W_s|\,|\,\mathcal{F}_s\right] \\
&\to 0 \quad a.s. \quad \text{for} \quad h\to 0.
\end{aligned}$$

Now consider the remaining first term on the right hand side of Equation (3.19). Using integration by parts we obtain

$$\begin{aligned}\mathbb{E}&\left[u^{(1)}\left(s+h, X^{(1)}_{s+h}, X^{(2)}_s, X^{(3)}_s\right)(W_{s+h}-W_s)\Big|\mathcal{F}_s\right]\\ &= \int_{\mathbb{R}} u^{(1)}\left(s+h, X^{(1)}_s+z\sqrt{h}, X^{(2)}_s, X^{(3)}_s\right) z\sqrt{h}\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}z^2}\,\mathrm{d}z\\ &= \int_{\mathbb{R}} \partial_{x_1}u^{(1)}\left(s+h, X^{(1)}_s+z\sqrt{h}, X^{(2)}_s, X^{(3)}_s\right) h\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}z^2}\,\mathrm{d}z.\end{aligned}$$

Since $\partial_{x_1}u^{(1)}$ is bounded as proved in Lemma 3.10 we have

$$\begin{aligned}&\left|\frac{1}{h}\mathbb{E}\left[u^{(1)}\left(s+h, X^{(1)}_{s+h}, X^{(2)}_s, X^{(3)}_s\right)(W_{s+h}-W_s)\Big|\mathcal{F}_s\right]\right|\\ &\qquad= \left|\int_{\mathbb{R}} \partial_{x_1}u^{(1)}\left(s+h, X^{(1)}_s+z\sqrt{h}, X^{(2)}_s, X^{(3)}_s\right)\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}z^2}\,\mathrm{d}z\right|\\ &\qquad\leqslant \|\partial_{x_1}u^{(1)}\|_\infty.\end{aligned}$$

Putting the derived estimates together we get

$$\left|Z^{(1)}_s\right| = \left|\lim_{h\searrow 0}\frac{1}{h}\mathbb{E}\left[Y^{(1)}_{s+h}(W_{s+h}-W_s)\Big|\mathcal{F}_s\right]\right| \leqslant \left\|\partial_{x_1}u^{(1)}\right\|_\infty.$$

By Lemma 3.10, $\|\partial_{x_1}u^{(1)}\|_\infty < \infty$, which further implies the result. ■

**Proposition 3.15**
*Let $g$, $\mu$ and $\sigma$ fulfill Assumption 3.5, let their first, second and third derivatives as well as $\sigma$ and $\frac{1}{g'}$ be bounded. Then the requirements of Theorem 3.9 are fulfilled.*

*Proof.* Remember that the derivative $\partial_{x_1}u$ of the decoupling field of FBSDE (3.2) equals $u^{(1)}$ of the decoupling field of FBSDE (3.16) by Lemma 3.10 and which, by Lemma 3.13, is bounded from below by a $\delta > 0$. Hence, it only remains to show that $\partial_{x_1}u$ which equals $u^{(1)}$ is Lipschitz continuous. Since we already know that the derivatives w.r.t. the space variables are bounded (by Lemma 3.10) we only need to prove that $u^{(1)}$ is Lipschitz continuous in the time variable.

Consider FBSDE (3.16) for a starting time $t \in [0,1)$ on the interval $[t,1]$ with initial condition $\left(X^{(1)}_t, X^{(2)}_t, X^{(3)}_t\right) = (x^{(1)}, x^{(2)}, x^{(3)}) = x \in \mathbb{R}^3$. Let $s \in (t,1]$. Using the triangle inequality several times gives

$$\begin{aligned}\left|u^{(1)}(s,x)-u^{(1)}(t,x)\right| \leqslant& \left|u^{(1)}(s,x)-\mathbb{E}\left[u^{(1)}\left(s, X^{(1)}_s, x^{(2)}, x^{(3)}\right)\right]\right|\\ &+\left|\mathbb{E}\left[u^{(1)}\left(s, X^{(1)}_s, x^{(2)}, x^{(3)}\right)\right]-\mathbb{E}\left[u^{(1)}\left(s, X^{(1)}_s, X^{(2)}_s, x^{(3)}\right)\right]\right|\\ &+\left|\mathbb{E}\left[u^{(1)}\left(s, X^{(1)}_s, X^{(2)}_s, x^{(3)}\right)\right]-\mathbb{E}\left[u^{(1)}\left(s, X^{(1)}_s, X^{(2)}_s, X^{(3)}_s\right)\right]\right|\\ &+\left|\mathbb{E}\left[u^{(1)}\left(s, X^{(1)}_s, X^{(2)}_s, X^{(3)}_s\right)-u^{(1)}\left(t, X^{(1)}_t, X^{(2)}_t, X^{(3)}_t\right)\right]\right|.\end{aligned}$$

We take a closer look at every summand on the right hand side starting with the first one. By defining

$$\varphi(z) := u^{(1)}(s, x^{(1)}, x^{(2)}, x^{(3)}) - u^{(1)}(s, x^{(1)}+z, x^{(2)}, x^{(3)})$$

we see that the first summand equals $|\mathbb{E}[\varphi(W_s - W_t)]|$. Furthermore, $\varphi(0) = 0$ and by Lemma 3.12, $\varphi$ is two times weakly differentiable with derivatives bounded by some constant

$L_{\partial_x u^{(1)}} < \infty$. Hence, the inequality $\left|\int_{\mathbb{R}} \varphi(a \cdot z)\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}z^2}\,\mathrm{d}z\right| \leqslant \frac{1}{2}a^2\|\varphi''\|_\infty$ holds true (see e.g. Lemma 4.3.11 in [Fro15]). Therefore,

$$\left|u^{(1)}(s,x) - \mathbb{E}\left[u^{(1)}\left(s, X_s^{(1)}, x^{(2)}, x^{(3)}\right)\right]\right| = \left|\mathbb{E}\left[\varphi(W_s - W_t)\right]\right| \leqslant \frac{(s-t)}{2} \cdot L_{\partial_x u^{(1)}}.$$

For the second summand we use the Lipschitz constant of $u^{(1)}$ denoted by $L_{u^{(1)}}$ to get

$$\begin{aligned}\left|\mathbb{E}\left[u^{(1)}\left(s, X_s^{(1)}, x^{(2)}, x^{(3)}\right) - u^{(1)}\left(s, X_s^{(1)}, X_s^{(2)}, x^{(3)}\right)\right]\right| \leqslant & L_{u^{(1)}}\mathbb{E}\left|X_s^{(2)} - x^{(2)}\right| \\ = & L_{u^{(1)}}\mathbb{E}\left|\int_t^s \frac{(Z_r^{(0)})^2}{\sigma_r^2}\,\mathrm{d}r\right| \\ \leqslant & L_{u^{(1)}}\|u^{(1)}\|_\infty^2 \varepsilon^{-2}(s-t)\end{aligned}$$

since $|Z^{(0)}| \leqslant \|u^{(1)}\|_\infty < \infty$ by Theorem 3.6.

The third summand can be estimated similarly by

$$\begin{aligned}&\left|\mathbb{E}\left[u^{(1)}\left(s, X_s^{(1)}, X_s^{(2)}, x^{(3)}\right) - u^{(1)}\left(s, X_s^{(1)}, X_s^{(2)}, X_s^{(3)}\right)\right]\right| \\ &\qquad \leqslant L_{u^{(1)}}\mathbb{E}\left|\int_t^s \mu_r \frac{(Z_r^{(0)})^2}{\sigma_r^2}\,\mathrm{d}r\right| \\ &\qquad \leqslant L_{u^{(1)}}\left\|\frac{\mu}{\sigma^2}\right\|_\infty \|u^{(1)}\|_\infty^2 (s-t).\end{aligned}$$

For the last summand we use the decoupling condition and $Y_\cdot^{(3)} = -1$ to obtain

$$\begin{aligned}&\left|\mathbb{E}\left[u^{(1)}\left(s, X_s^{(1)}, X_s^{(2)}, X_s^{(3)}\right) - u^{(1)}\left(t, X_t^{(1)}, X_t^{(2)}, X_t^{(3)}\right)\right]\right| \\ &\quad \leqslant \left|\mathbb{E}\left[Y_s^{(1)} - Y_t^{(1)}\right]\right| \\ &\quad = \left|\mathbb{E}\left[\int_t^s Y_r^{(1)}\frac{(Z_r^{(0)})^2}{\sigma_r^2}\left(\mu_{a,r} - 2\mu_r\frac{\sigma_{a,r}}{\sigma_r} + 2Y_r^{(2)}\frac{\sigma_{a,r}}{\sigma_r}\right) - \frac{2Z_r^{(0)}}{\sigma_r^2}\left(Y_r^{(2)} - \mu_r\right)Z_r^{(1)}\,\mathrm{d}r\right]\right| \\ &\quad \leqslant \left[\|u^{(1)}\|_\infty^3\left(\left\|\frac{\partial_a\mu}{\sigma^2}\right\|_\infty + 2\left\|\frac{\mu}{\sigma^2}\right\|_\infty\left\|\frac{\partial_a\sigma}{\sigma}\right\|_\infty + \frac{2}{\varepsilon^2}\|u^{(2)}\|_\infty\left\|\frac{\partial_a\sigma}{\sigma}\right\|_\infty\right)\right. \\ &\qquad\qquad \left. +2\|u^{(1)}\|_\infty\left(\varepsilon^{-2}\|u^{(2)}\|_\infty + \left\|\frac{\mu}{\sigma^2}\|_\infty\right)\|\partial_{x_1}u^{(1)}\|_\infty\right](s-t)\end{aligned}$$

where we applied Theorem 3.6 and Lemma 3.14. Thus, the last summand is Lipschitz continuous by Theorem 3.6 and Lemma 3.10, too.

Putting all estimates together we arrive at $|u^{(1)}(s,x) - u^{(1)}(r,x)| \leqslant L(s-t)$ for some finite constant $L$ which is independent of $s$ and $t$. Hence $u^{(1)}$ is Lipschitz continuous in the time variable.

■

Observe that Proposition 3.15 and Theorem 3.9 imply Theorem 3.22.

## 3.5 Numerics

We now illustrate numerically an example of an embedding using the methodology developed. This is done by numerically approximating the solution of the FBSDE

$$\begin{aligned}
W_s &= \int_0^s \frac{\sigma(X_r^{(2)}, Y_r + X_r^{(3)})}{Z_r} \, \mathrm{d}B_{X_r^{(2)}} \\
X_s^{(2)} &= \int_0^s \frac{Z_r^2}{\sigma^2(X_r^{(2)}, Y_r + X_r^{(3)})} \, \mathrm{d}r \\
X_s^{(3)} &= \int_0^s \mu(X_r^{(2)}, Y_r + X_r^{(3)}) \frac{Z_r^2}{\sigma^2(X_r^{(2)}, Y_r + X_r^{(3)})} \, \mathrm{d}r \\
Y_s &= g(W_1) - X_1^{(3)} - \int_s^1 Z_r \, \mathrm{d}W_r.
\end{aligned} \tag{3.20}$$

To the best of our knowledge no literature exists able to deal directly with approximations of (3.20) and hence, inspired by known literature, we propose a numerical scheme whose rigorous study is left for future research. FBSDE (3.20) is a fully coupled quadratic growth FBSDE which we deal with as follows: from [IDRZ10] we inject the theoretical a priori hard bounds in the coefficients, reducing FBSDE (3.20) to a uniformly Lipschitz fully-coupled one, then apply a decoupling technique based on Picard iterations [BZ08] to reduce the problem to the iterative simulation of uniformly Lipschitz fully-decoupled FBSDE. The final approximation step is carried out using a classic explicit Euler scheme discretization [BZ08] while the approximation of the conditional expectations is done via projection over basis functions [GLW05]. The final outcome is the approximation of the embedding stopping time and the verification that the stopped process does embed the target distribution.

From a mathematical point of view, the only step of the described numerical approximation that cannot be fully justified is the convergence of the Picard iteration step. The results of [BZ08] do not apply if the diffusion coefficient $\sigma$ depends on $Z$. We stress, however, that for some special cases the algorithm outlined below can be shown to converge, e.g. in the homogeneous case (see Remark 3.20 below).

### The problem, its conditions and the hard bounds

At first we show that FBSDE (3.20) has a unique solution from which we can construct a strong solution of the SEP.

**Proposition 3.16**
*Let the assumptions of Theorem 3.9 or Proposition 3.15 be satisfied. Denote by $u$ the decoupling field of FBSDE* (3.2). *Let $B$ be an arbitrary Brownian motion and denote by $(\mathcal{F}^B) = (\mathcal{F}_s^B)_{s\in[0,\infty)}$ the augmented filtration generated by $B$. Then there exist unique square-integrable processes $(W, X^{(2)}, X^{(3)}, Y)$ solving the FBSDE* (3.20). *Moreover, $\tau := X_1^{(2)}$ is an $(\mathcal{F}_t^B)$-stopping time bounded as in* (3.22)*, $W$ is a Brownian motion on $[0,1]$ and the pair $(\tau, Y_0)$ is a strong solution of the SEP.*

*Proof.* Remember that by Theorem 3.9 the SDE (3.15) has a unique solution $(\gamma, \Gamma, \Delta, \Theta)$. We introduce the time change $\gamma^{-1}(t) = \inf\{r \geqslant 0 : \gamma(r) \geqslant t\}$ for $t \in [0,1]$. Observe that $\gamma^{-1}$ has the dynamics

$$\gamma^{-1}(t) = \int_0^t \frac{\left(\partial_{x_1} u(s, \Gamma_{\gamma^{-1}(s)}, \gamma^{-1}(s), \Delta_{\gamma^{-1}(s)})\right)^2}{\sigma^2(\gamma^{-1}(s), \Theta_{\gamma^{-1}(s)} + \Delta_{\gamma^{-1}(s)})} \, \mathrm{d}s.$$

By setting $Z_s := \partial_{x_1} u(s, \Gamma_{\gamma^{-1}(s)}, \gamma^{-1}(s), \Delta_{\gamma^{-1}(s)})$ for $s \in [0,1]$, replacing the dynamics of $\gamma$ by the dynamics of $\gamma^{-1}$ and applying the time change $\gamma^{-1}$ to all other processes, we can rewrite the system (3.15) as

$$\begin{aligned}
\gamma^{-1}(t) &= \int_0^t \frac{(Z_s)^2}{\sigma^2(\gamma^{-1}(s), \Theta_{\gamma^{-1}(s)} + \Delta_{\gamma^{-1}(s)})} \, \mathrm{d}s \\
\Gamma_{\gamma^{-1}(t)} &= \int_0^t \frac{\sigma(\gamma^{-1}(s), \Theta_{\gamma^{-1}(s)} + \Delta_{\gamma^{-1}(s)})}{Z_s} \, \mathrm{d}B_{\gamma^{-1}(s)} \\
\Delta_{\gamma^{-1}(t)} &= \int_0^t \mu(\gamma^{-1}(s), \Theta_{\gamma^{-1}(s)} + \Delta_{\gamma^{-1}(s)}) \frac{(Z_s)^2}{\sigma^2(\gamma^{-1}(s), \Theta_{\gamma^{-1}(s)} + \Delta_{\gamma^{-1}(s)})} \, \mathrm{d}s \\
\Theta_{\gamma^{-1}(t)} &= Y_0 + \int_0^t Z_s \, \mathrm{d}\Gamma_{\gamma^{-1}(s)}
\end{aligned}$$

for all $t \in [0,1]$. Here it is straightforward to see that with $\gamma^{-1}(t) = X_t^{(2)}$, $\Gamma_{\gamma^{-1}(t)} = W_t$, $\Delta_{\gamma^{-1}(t)} = X_t^{(3)}$ and $\Theta_{\gamma^{-1}(t)} = Y_t$ we exactly have the system (3.20). Thus the system (3.20) has a solution $(W, X^{(2)}, X^{(3)}, Y, Z)$ which fulfills that $\tau := X_1^{(2)} = \gamma^{-1}(1) = \inf\{r \geqslant 0 | \gamma(r) = 1\}$ is a stopping time with regard to $(\mathcal{F}_t^B)$ bounded as in (3.22) and that $A_\tau \sim \nu$.

It remains to show the uniqueness of this solution. Now take an arbitrary square integrable solution $(W, X^{(2)}, X^{(3)}, Y, Z)$ of (3.20). Define the time change

$$\bar{\gamma}(t) := \begin{cases} \inf\{s \geqslant 0 : X_s^{(2)} \geqslant t\}, & t \leqslant X_1^{(2)} \\ 1, & t > X_1^{(2)} \end{cases}$$

and observe that by

$$\langle W, W \rangle_t = \int_0^{X_t^{(2)}} \frac{\sigma^2\left(r, Y_{\bar{\gamma}(r)} + X^{(3)}_{\bar{\gamma}(r)}\right)}{Z^2_{\bar{\gamma}(r)}} \, \mathrm{d}r = \int_0^t \frac{\sigma^2\left(X_r^{(2)}, Y_r + X_r^{(3)}\right)}{Z_r^2} \, \mathrm{d}X_r^{(2)} = \int_0^t 1 \, \mathrm{d}r = t$$

$W$ is a Brownian motion on $[0,1]$. Thus the processes $(W, X^{(2)}, X^{(3)}, Y, Z)$ solve FBSDE (3.2) for the initial value $0$. Due to Theorem 2.8 and Lemma 2.9 this solution of FBSDE (3.2) is unique. ■

**Remark 3.17**

If one is only interested in a weak solution, then only FBSDE (3.2) needs to be solved, where $W$ is given, and the Brownian motion $B$ can be calculated afterwards, as described in Theorem 3.7. Aside from simplifying the system that needs to be simulated, this also has the advantage of being valid for more general coefficients $\mu$ and $\sigma$ (compare the assumptions of Theorem 3.21 and Theorem 3.22).

By the combination of Lemma 3.13, Lemma 3.2 and Theorem 3.6 we have for $Z$ the $\lambda \times \mathbf{P}$ a.e. bounds $0 < \check{Z} \leqslant Z \leqslant \hat{Z} < \infty$, which are

$$\hat{Z} = \left( \frac{1}{\|g'\|_\infty^2} + 2 \min\left\{0, \inf_{(\theta,x) \in \mathbb{R}_+ \times \mathbb{R}} \left( \frac{\sigma \cdot \partial_a \mu - 2 \partial_a \sigma \cdot \mu}{\sigma^3} \right)(\theta, x) \right\} \right)^{-\frac{1}{2}} \quad \text{and}$$

$$\check{Z} = \left( \left\| \frac{1}{g'} \right\|_\infty + \check{Z} \left( \left\| \frac{\partial_a \mu}{\sigma^2} \right\|_\infty + 2 \left\| \frac{\mu}{\sigma^2} \right\|_\infty \left\| \frac{\partial_a \sigma}{\sigma} \right\|_\infty + \frac{2}{\varepsilon^2} \|\partial_{x_2} u\|_\infty \left\| \frac{\partial_a \sigma}{\sigma} \right\|_\infty \right) \right)^{-1}$$

with

$$\|\partial_{x_2} u\|_\infty \leqslant \exp\left[\widehat{Z}^2\left(\left\|\frac{\partial_a\mu}{\sigma^2}\right\|_\infty + 2\left(\left\|\frac{\partial_a\sigma}{\sigma}\right\|_\infty \left\|\frac{\mu}{\sigma^2}\right\|_\infty + \frac{1}{\varepsilon^2}\left\|\frac{\partial_t\sigma}{\sigma}\right\|_\infty\right)\right)\right] \cdot \widehat{Z}^2\left(2\left\|\frac{\partial_t\sigma}{\sigma}\right\|_\infty \left\|\frac{\mu}{\sigma^2}\right\|_\infty + \left\|\frac{\partial_t\mu}{\sigma^2}\right\|_\infty\right).$$

Therefore, we have that

$$\frac{\check{Z}^2}{\|\sigma\|_\infty^2} \leqslant \frac{Z_s^2}{\sigma^2(X_s^{(2)}, Y_s + X_s^{(3)})} \leqslant \frac{\widehat{Z}^2}{\varepsilon^2}, \qquad \lambda \times \mathbf{P} \text{ a.e.}$$

and in particular

$$\frac{\check{Z}^2}{\|\sigma\|_\infty^2} \leqslant \tau = X_1^{(2)} \leqslant \frac{\widehat{Z}^2}{\varepsilon^2} \qquad \text{a.s.} \tag{3.21}$$

**Example 3.18** (Embedding a Normal distribution into a Brownian motion with drift)
For $\mu \equiv m \in \mathbb{R}$, $\sigma \equiv 1$ and $\nu = \mathcal{N}(0, \alpha^2)$ for $\alpha > 0$ we know that $\tau = \alpha^2$ and $A_0 = -m \cdot \alpha^2$ solves the SEP. In this case we have that $g(x) = \alpha x$ and the above bounds for $Z$ become the explicit values $\alpha \leqslant Z \leqslant \alpha$ and the system (3.20) simplifies to

$$W_s = \int_0^s \frac{1}{\alpha}\,\mathrm{d}B_{X_r^{(2)}}, \quad X_s^{(2)} = \int_0^s \alpha^2 dr, \quad X_s^{(3)} = \int_0^s m \cdot \alpha^2 dr,$$
$$Y_s = \alpha W_1 - X_1^{(3)} - \left(B_{X_1^{(2)}} - B_{X_s^{(2)}}\right)$$

giving that $\tau = X_1^{(2)} = \alpha^2$ a.s. which equals the above mentioned stopping time. We immediately find the correct value for $A_0$ since

$$A_0 = Y_0 = \mathbb{E}\left[Y_1 \middle| \mathcal{F}_0\right] = \mathbb{E}\left[\alpha W_1 - \int_0^1 m\alpha^2\,\mathrm{d}r - B_{X_1^{(2)}} + B_{X_0^{(2)}} \middle| \mathcal{F}_0\right] = -m\alpha^2.$$

**Example 3.19**
Again let $\nu = \mathcal{N}(0, \alpha^2)$ for $\alpha > 0$. Furthermore, set

$$\sigma(t,a) = p_1^\sigma + \frac{p_2^\sigma}{1+e^{-t}} + \frac{p_3^\sigma}{1+e^{-a}} \qquad \text{and} \qquad \mu(t,a) = p_1^\mu + \frac{p_2^\mu}{1+e^{-t}} + \frac{p_3^\mu}{1+e^{-a}}$$

for the vectors $p^\sigma, p^\mu \in \mathbb{R}^3$ containing parameters such that

$$\varepsilon := p_1^\sigma + \min(0, p_2^\sigma) + \min(0, p_3^\sigma) > 0,$$
$$2p_2^\sigma p_3^\sigma p_1^\mu - p_1^\sigma p_2^\mu + \min(0, p_2^\sigma p_3^\sigma p_2^\mu) + \min(0, 2p_2^\sigma p_3^\sigma p_3^\mu - (p_3^\sigma)^2 p_2^\mu) > 0$$

and

$$\frac{1}{\alpha^2} + \frac{p_1^\sigma p_3^\mu - 2p_3^\sigma p_1^\mu + \min(0, p_2^\sigma p_3^\mu - 2p_3^\sigma p_2^\mu) - \max(0, p_3^\sigma p_3^\mu)}{2\varepsilon^3} > 0.$$

Then observe that all conditions of Proposition 3.15 and therefore also of Proposition 3.16 are fulfilled,

$$\widehat{Z} \leqslant \left(\frac{1}{\alpha^2} + \frac{p_1^\sigma p_3^\mu - 2p_3^\sigma p_1^\mu + \min(0, p_2^\sigma p_3^\mu - 2p_3^\sigma p_2^\mu) - \max(0, p_3^\sigma p_3^\mu)}{2\varepsilon^3}\right)^{-\frac{1}{2}} < \infty$$

and also $\check{Z}$ can be directly obtained since

$$\|\sigma\|_\infty = p_1^\sigma + \max(0, p_2^\sigma) + \max(0, p_3^\sigma),$$
$$\|\mu\|_\infty = \max\left(p_1^\mu + \max(0, p_2^\mu) + \max(0, p_3^\mu), -p_1^\mu - \min(0, p_2^\mu) - \min(0, p_3^\mu)\right),$$
$$\|\partial_a\sigma\|_\infty = |p_3^\sigma|, \quad \|\partial_t\sigma\|_\infty = |p_2^\sigma|, \quad \|\partial_a\mu\|_\infty = |p_3^\mu|, \quad \|\partial_t\mu\|_\infty = |p_2^\mu|.$$

## Iterative procedure

To numerically approximate (3.20) we first embed the hard bounds for $Z$, as found above, in the system, then create a Picard-type approximative sequence converging to (3.20) and numerically approximate the terms of said sequence. Since we have a coupled system of FBSDEs with a truncated quadratic growth component, we combine [IDRZ10] and [BZ08].

Since $X^{(2)}$ is increasing and

$$X_1^{(2)} \leqslant \varepsilon^{-2}\left(\frac{1}{\|g'\|_\infty^2} + 2\min\left\{0, \inf_{(\theta,x)\in\mathbb{R}_+\times\mathbb{R}}\left(\frac{\sigma\cdot\partial_a\mu - 2\partial_a\sigma\cdot\mu}{\sigma^3}\right)(\theta,x)\right\}\right)^{-1}$$

a.s. as stated in Equation (3.21), we only need a trajectory of $B$ untill this point.

Furthermore, choose any starting value for $Z$ between the lower and upper bounds $\check{Z}, \widehat{Z}$ respectively. Here we set the starting value $Z^{(0)} = \|g'\|_\infty$ since $\check{Z} \leqslant \|\frac{1}{g'}\|_\infty^{-1} \leqslant \|g'\|_\infty \leqslant \widehat{Z}$. Moreover, we define a truncation operator to incorporate the hard bounds for $Z$, namely, let $T:\mathbb{R}\to\mathbb{R}$ such that given $\check{Z}, \widehat{Z}$, we define $T(z) := \min(\max(z,\check{Z}),\widehat{Z})$. The map $T$ is uniformly Lipschitz.

For the other starting conditions we choose $Y^{(0)} = X^{(2),(0)} = X^{(3),(0)} = 0$. Then we do the following iterations for $k \in \mathbb{N}_0$:

$$\begin{aligned}
X_s^{(2),(k+1)} &= \int_0^s \frac{\left(T\big(Z_r^{(k)}\big)\right)^2}{\sigma^2\left(X_r^{(2),(k+1)}, Y_r^{(k)} + X_r^{(3),(k+1)}\right)}\,\mathrm{d}r\\
X_s^{(3),(k+1)} &= \int_0^s \mu\left(X_r^{(2),(k+1)}, Y_r^{(k)} + X_r^{(3),(k+1)}\right)\frac{\left(T\big(Z_r^{(k)}\big)\right)^2}{\sigma^2\left(X_r^{(2),(k+1)}, Y_r^{(k)} + X_r^{(3),(k+1)}\right)}\,\mathrm{d}r\\
W_s^{(k+1)} &= \int_0^s \frac{\sigma\left(X_r^{(2),(k+1)}, Y_r^{(k)} + X_r^{(3),(k+1)}\right)}{T\big(Z_r^{(k)}\big)}\,\mathrm{d}B_{X_r^{(2),(k+1)}}\\
Y_s^{(k+1)} &= g(W_1^{(k+1)}) - X_1^{(3),(k+1)} - \int_s^1 \sigma\left(X_r^{(2),(k+1)}, Y_r^{(k)} + X_r^{(3),(k+1)}\right)\mathrm{d}B_{X_r^{(2),(k+1)}}.
\end{aligned}$$

Under the conditions imposed on $\mu,\sigma$ (Lipschitz and bounded) and $T$, all the coefficient maps of the truncated FBSDE system are Lipschitz continuous. It is currently not clear how to show that the iterative system converges to the solution of (3.20) where one could possibly use a result similar to [BZ08, Theorem 2.1]; this difficulty stems from the fact that the [BZ08] methodology does not allow for either random drift or diffusion coefficients or $\sigma$ depending on $Z$. Note that in the limit ($k\to\infty$) the truncation does not affect the system as $\check{Z} \leqslant Z \leqslant \widehat{Z}$ .

### Numerical procedure (time discretization)

We introduce the time discretization $\pi = \{0 = t_0, \dots, t_n = 1\}$ for $n \in \mathbb{N}$ and define $|\pi| := \max_{i=0,\cdots,n} |t_{i+1} - t_i|$ as the mesh's modulus. The numerical approximation of the iterative system, for each $k \in \mathbb{N}$ follows [BT04] (or [BZ08]). We apply an explicit Euler type approximation to the integrals and let throughout $t_i \in \pi \backslash \{t_0\}$. At first

$$\begin{aligned}
X^{(2),(k+1)}_{t_0} =&0, \qquad X^{(3),(k+1)}_{t_0} = 0\\
X^{(2),(k+1)}_{t_{i+1}} =&X^{(2),(k+1)}_{t_i} + (t_{i+1} - t_i)\left(\frac{T\big(Z^{(k)}_{t_i}\big)}{\sigma(X^{(2),(k+1)}_{t_i}, Y^{(k)}_{t_i} + X^{(3),(k+1)}_{t_i})}\right)^2\\
X^{(3),(k+1)}_{t_{i+1}} =&X^{(3),(k+1)}_{t_i} + (t_{i+1} - t_i)\frac{\mu\left(X^{(2),(k+1)}_{t_i}, Y^{(k)}_{t_i} + X^{(3),(k+1)}_{t_i}\right)\left(T\big(Z^{(k)}_{t_i}\big)\right)^2}{\sigma^2\left(X^{(2),(k+1)}_{t_i}, Y^{(k)}_{t_i} + X^{(3),(k+1)}_{t_i}\right)},
\end{aligned}$$

then

$$W^{(k+1)}_{t_0} = 0, \quad W^{(k+1)}_{t_{i+1}} =W^{(k+1)}_{t_i} + \frac{\sigma(X^{(2),(k+1)}_{t_i}, Y^{(k)}_{t_i} + X^{(3),(k+1)}_{t_i})}{T\big(Z^{(k)}_{t_i}\big)}\left(B_{X^{(2),(k+1)}_{t_{i+1}}} - B_{X^{(2),(k+1)}_{t_i}}\right)$$

and

$$\begin{aligned}
Y^{(k+1)}_{t_n} =&g\left(W^{(k+1)}_1\right) - X^{(3),(k+1)}_1\\
Y^{(k+1)}_{t_{i-1}} =&\mathbb{E}\left[\, Y^{(k+1)}_{t_i}\Big|\, \mathcal{F}_{t_{i-1}}\right]\\
Z^{(k+1)}_{t_{i-1}} =&\frac{1}{t_i - t_{i-1}}\mathbb{E}\left[\left(Y^{(k+1)}_{t_i} - \mathbb{E}\left[Y^{(k+1)}_{t_i}\Big|\, \mathcal{F}_{t_{i-1}}\right]\right)\left(W_{t_i} - W_{t_{i-1}}\right)\Big|\, \mathcal{F}_{t_{i-1}}\right].
\end{aligned}$$

The time discretization expression for $Z^{(k+1)}_{t_{i-1}}$ is somewhat non-standard when compared with the [BT04] scheme. The inner term with the conditional expectation of $Y^{(k+1)}_{t_i}$ is a variance reduction trick which has been discussed in several places, e.g. [LdRS15, Section 5.4.2]; independently, the scheme's convergence (for fixed $k$ as $h \searrow 0$) follows via [BT04, Theorem 3.1] yielding a convergence rate of order $h^{1/2}$ (the formulation associated to [BZ08, Theorem 2.2] would deliver the same convergence). In the calculation of $Z$ we use that

$$\int_s^1 \sigma(X^{(2)}_r, Y_r + X^{(3)}_r)\,\mathrm{d}B_{X^{(2)}_r} = \int_s^1 Z_r\,\mathrm{d}W_r$$

for all $s \in [0,1]$ and hence for small $h > 0$

$$\begin{aligned}
Z_t \approx&\frac{1}{h}\mathbb{E}\left[\int_t^{t+h} Z_r\,\mathrm{d}r\Big|\, \mathcal{F}_t\right]\\
=&\frac{1}{h}\mathbb{E}\left[\left(Y_{t+h} - Y_t\right)\left(W_{t+h} - W_t\right) - \int_t^{t+h}\left(Y_r - Y_t + \left(W_r - W_t\right) Z_r\right)\mathrm{d}W_r\Big|\, \mathcal{F}_t\right]\\
=&\frac{1}{h}\mathbb{E}\left[Y_{t+h}\left(W_{t+h} - W_t\right)\big|\, \mathcal{F}_t\right]\\
=&\frac{1}{h}\mathbb{E}\left[\left(Y_{t+h} - \mathbb{E}\left[Y_{t+h}\big|\, \mathcal{F}_t\right]\right)\left(W_{t+h} - W_t\right)\big|\, \mathcal{F}_t\right].
\end{aligned}$$

For the calculation of $W$ we implicitly assume that the value of $B$ is known for every $X^{(2),(k)}_{t_i}$ for all $k \geqslant 0$ and $t_i \in \pi$. This problem is more involved if the trajectory of $B$ is to be calculated at the beginning of the simulation. However, it can be eliminated by calculating the trajectory of $B$ just in time for the points needed by the method of Brownian bridge and storing all thereby obtained points. It is well known that the distribution of a Brownian bridge $B$ at time $t_1$ under the condition of the values of $B$ at the times $t_0 < t_1$ and $t_2 > t_1$ is

$$B_{t_1}|B_{t_0}, B_{t_2} \sim \mathcal{N}\left(B_{t_0} \cdot \frac{t_2 - t_1}{t_2 - t_0} + B_{t_2} \cdot \frac{t_1 - t_0}{t_2 - t_0}\ ,\ \ \frac{(t_2 - t_1)(t_1 - t_0)}{t_2 - t_0}\right),$$

see e.g. [KS91]. Thus the simulation of $B$ at the exact points of time is straightforward as well. Lastly, the conditional expectations are computed via Least-Squares regression functions as shown in [GLW05]; we project over 3-dimensional polynomials up to degree 2.

After finishing the simulation of the FBSDE we can use the simulated trajectory of $B$ to simulate our process $A$ and apply the stopping time $\tau$ to see if $A_\tau$ has the desired distribution.

**Remark 3.20**
For time homogeneous coefficients $\mu$ and $\sigma$ the FBSDE (3.2) simplifies to the decoupled FBSDE

$$X_s^{(2)} = \int_0^s \frac{Z_r^2}{\sigma^2(\bar{Y}_r)}\,\mathrm{d}r, \qquad \bar{Y}_s = g(W_1) - \int_s^1 \mu(\bar{Y}_r)\frac{Z_r^2}{\sigma^2(\bar{Y}_r)}\,\mathrm{d}r - \int_s^1 Z_r\,\mathrm{d}W_r.$$

For this decoupled system one can use the same trick as above and inject in the BSDE the hard bounds on $Z$. Once truncated and using the condition on $\mu, \sigma$, the driver of the BSDE, say $f_R(y,z) = T^2(z)\mu(y)/\sigma^2(y)$ using the notation from before, is a standard uniformly Lipschitz driver in $y, z$ for which it is known ([BT04], [BZ08], [GLW05]) that the Euler explicit scheme converges to the true solution. For weak solutions (see Remark 3.17) of the SEP this explicit scheme is equivalent to the scheme we propose here. Hence, we have a special case where the convergence of our scheme is known.

## Numerical testing for Example 3.19

For the parameters $\alpha = 1$, $p^\sigma = (2, 0.5, 2)$ and $p^\mu = (1.5, -2.5, 0.5)$ such that $\nu = \mathcal{N}(0,1)$,

$$\sigma(t,a) = 2 + \frac{0.5}{1+e^{-t}} + \frac{2}{1+e^{-a}} \qquad \text{and} \qquad \mu(t,a) = 1.5 + \frac{-2.5}{1+e^{-t}} + \frac{0.5}{1+e^{-a}}$$

we get $\varepsilon = 2$, $\|\sigma\|_\infty = 4.5$, $\widehat{Z} \leqslant \sqrt{\frac{8}{5}}$ and $\check{Z} \geqslant 0.111$ giving $6 \times 10^{-4} \leqslant \tau \leqslant 0.4$. A simulation with $10^5$ paths, 20 time steps and 50 iterations yielded values for $\tau$ in the interval $[0.061; 0.161]$ and the starting value $Y_0 = -0.042$.

We simulated $A_\tau$ with initial condition $A_0 = Y_0 = -0.042$. In Figure 3.1 one finds the histogram of the simulated values of the $A_\tau$ (left) and the stopping time $\tau$ (right). The histogram of $A_\tau$ indicates that our algorithm generates the sought normal distribution (with the appropriate characteristics). Also, D'Agostino and Pearson's [D'A71, DP73] test for normality, applied to the simulated data $A_\tau$, yielded a $p$-value of 0.37. Given such a high $p$-value we do not reject the hypothesis of normality at any reasonable significance level.

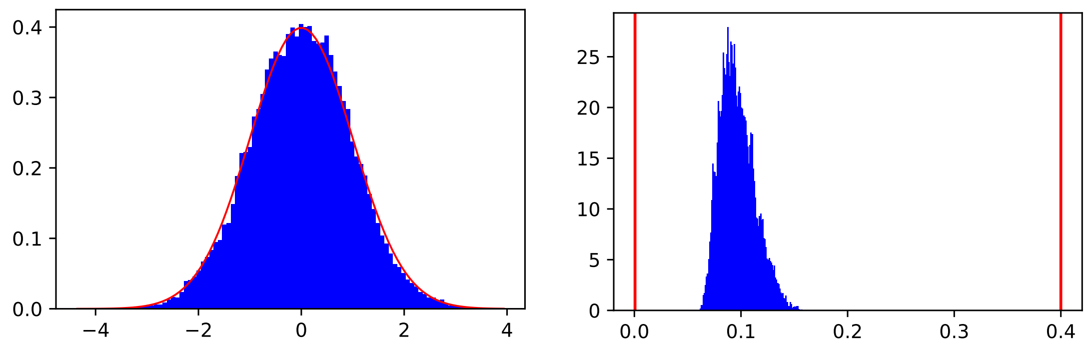



Figure 3.1: On the left, Histogram of $10^5$ samples of $A_\tau$ against the density of $\mathcal{N}(0,\alpha)$; on the right, the Histogram of the corresponding samples of $\tau$ and at $x = 0.0055$ and $x = 0.4$ the a priori hard bounds for the stopping time.

## 3.6 Discussion of the results

Here, we shortly restate our main results and compare our assumptions to the assumptions in other works. We start with our weak solution to the SEP.

**Theorem 3.21** (Weak Solution (see Theorem 3.7))
*Let Assumption 3.5 be satisfied. Then there exists a complete filtered probability space $(\Omega, \mathcal{F}, (\mathcal{G}_t)_{t\geqslant 0}, \mathbf{P})$, a $(\mathcal{G}_t)$-Brownian motion $(B_t)$, a bounded $(\mathcal{G}_t)$-stopping time $\tau$ and a real number $a$ such that for the strong solution $A$ of the SDE* (3.1) *with driving Brownian motion $B$ and initial condition $A_0 = a$ we have $A_\tau \sim \nu$. Furthermore, $\tau$ can be chosen such that*

$$\tau \leqslant \varepsilon^{-2}\left(\frac{1}{\|g'\|_\infty^2} + 2\min\left\{0, \inf_{(\theta,x)\in\mathbb{R}_+\times\mathbb{R}}\left(\frac{\sigma\cdot\partial_a\mu - 2\partial_a\sigma\cdot\mu}{\sigma^3}\right)(\theta,x)\right\}\right)^{-1} \quad a.s. \tag{3.22}$$

Theorem 3.21 basically states that if $\mu$, $\sigma$, $g$ fulfill Assumption 3.5, then we can give a filtration together with a Brownian motion and a stopping time which solve the SEP. We call it a weak solution since the filtration and Brownian motion are part of the solution instead of being given up front. In contrast, Theorem 3.22 states existance of a strong solution. This means that for given $\mu$, $\sigma$, $g$ and a Brownian motion we can construct a stopping time solving the SEP.

**Theorem 3.22** (Strong Solution (see Theorem 3.9 and Proposition 3.15))
*Let Assumption 3.5 be satisfied and assume furthermore that $\sigma$, $\frac{1}{g'}$ the first, second and third derivatives of $g$, $\mu$ and $\sigma$ are bounded. Let $B$ be a Brownian motion on a probability space $(\Omega, \mathcal{F}, P)$ and denote by $(\mathcal{F}_t)$ the augmented Brownian filtration. Then there exists $a \in \mathbb{R}$ and a bounded $(\mathcal{F}_t)$-stopping time $\tau$ satisfying* (3.22) *such that for the strong solution $A$ of the SDE* (3.1) *with driving Brownian motion $B$ and initial condition $A_0 = a$ we have $A_\tau \sim \nu$.*

In the following we shed some light on what some of the assumptions mean and how they compare to already existing works. For this, in the next lemma we give some necessary conditions for $g'$ to be bounded and bounded away from zero.

**Lemma 3.23**
*For $x \in \mathbb{R}$ define $g(x) := F_\nu^{-1}(\Phi(x))$ for $F_\nu$ and $\Phi$ being the cumulative distribution functions of $\nu$ and the standard normal distribution, and additionally define $\Phi_{0,\sigma}(x) = \Phi(\frac{x}{\sigma})$ for any $\sigma > 0$. If $\|g'\|_\infty < \infty$, then there exist $K > 0$ and $\sigma > 0$ such that*

- *for all $x < -K$ we have $F_\nu(x) \leqslant \Phi_{0,\sigma}(x) = \Phi(\frac{x}{\sigma})$ and*
- *for all $x > K$ we have $F_\nu(x) \geqslant \Phi_{0,\sigma}(x) = \Phi(\frac{x}{\sigma})$.*

*If additionally there exists a constant $c > 0$ such that $0 < c \leqslant g'$ then there exist $K > 0$ and $\sigma_1, \sigma_2 > 0$ such that*

- *for all $x > K$ we have $\quad \Phi_{0,\sigma_1}(x) = \Phi(\frac{x}{\sigma_1}) \leqslant F_\nu(x) \leqslant \Phi_{0,\sigma_2}(x) = \Phi(\frac{x}{\sigma_2}) \quad$ and*
- *for all $x < -K$ we have $\quad \Phi_{0,\sigma_2}(x) = \Phi(\frac{x}{\sigma_2}) \leqslant F_\nu(x) \leqslant \Phi_{0,\sigma_1}(x) = \Phi(\frac{x}{\sigma_1})$.*

*Proof.* Select $K, \sigma, \varepsilon > 0$ such that for all $x > K$ we have $g(\frac{x}{\sigma}) \leqslant x$ and for all $x < -K$ we have $g(\frac{x}{\sigma}) - \varepsilon \geqslant x$, which is possible since $0 \leqslant g' \leqslant C < \infty$. Then

$$\begin{aligned}
&\text{for } x > K: && F_\nu(x) = F_\nu(\tfrac{\sigma x}{\sigma}) \geqslant F_\nu(g(\tfrac{x}{\sigma})) = F_\nu(F_\nu^{-1}(\Phi(\tfrac{x}{\sigma}))) \geqslant \Phi(\tfrac{x}{\sigma}) = \Phi_{0,\sigma}(x),\\
&\text{for } x < -K: && F_\nu(x) = F_\nu(\tfrac{\sigma x}{\sigma}) \leqslant F_\nu(g(\tfrac{x}{\sigma}) - \varepsilon) = F_\nu(F_\nu^{-1}(\Phi(\tfrac{x}{\sigma})) - \varepsilon) \leqslant \Phi(\tfrac{x}{\sigma}) = \Phi_{0,\sigma}(x).
\end{aligned}$$

If additionally $0 < c \leqslant g'$ then we can choose $K_2 > 0$ and some $\sigma_2 > 0$ such that for all $x > K_2$ we have $g(\frac{x}{\sigma_2}) - \varepsilon \geqslant x$ and for all $x < -K_2$ we have $g(\frac{x}{\sigma_2}) \leqslant x$. By an analogous argumentation as above we then obtain the remaining estimates. Setting $K$ as the maximum of $K$ from above and $K_2$ and furthermore $\sigma_1 := \sigma$ we have proved the statement. ∎

**Remark 3.24**
We now comment on Assumption 3.5. In particular, we relate the assumption to some conditions appearing in the literature that have been shown to be sufficient for a bounded solution of the SEP to exist.

a) By Lemma 3.23 we get that the assumption of $g'$ being bounded entails that there exists a compact set outside of which the tails of $\nu$ are dominated by the tails of a normal distribution. If, as in Theorem 3.22, we additionally have that $g'$ is bounded from below by a positive constant, then the tails of $\nu$ also dominates the tails of a normal distribution.

   Furthermore, observe that the left hand side of Condition (3.7) is equal to $\partial_a \left(\frac{\mu}{\sigma^2}\right)$ and in the cases $ii)$ and $iii)$ the term $2\partial_t\sigma \cdot \mu - \sigma \cdot \partial_t\mu$ equals $-\sigma^3\partial_t \left(\frac{\mu}{\sigma^2}\right)$; hence Assumption 3.5 imposes conditions on the growth of $\frac{\mu}{\sigma^2}$.

b) Theorem 3.1 in [AS11] states that the boundedness of $g'$ is sufficient for the SEP for the BM, possibly with a constant drift, to possess a bounded solution. Notice that for $\sigma \equiv 1$ and constant $\mu$ Inequality (3.22) simplifies to

$$\tau \leqslant \|g'\|_\infty^2,$$

   and hence coincides with the estimate on the embedding stopping time provided in Theorem 3.1 in [AS11]. Moreover, observe that if $\sigma$ and $\mu$ are constant, then all the other properties of Assumption 3.5 are satisfied trivially.

c) The ratio on the left-hand side of (3.7) is equal to $\partial_a\left(\frac{\mu}{\sigma^2}\right)$. Thus, (3.7) is somewhat weaker than requiring that $\frac{\mu}{\sigma^2}$ is non-decreasing in $x$. For some mean-reversion processes, e.g. the Ornstein-Uhlenbeck process, $\partial_a\left(\frac{\mu}{\sigma^2}\right)$ is unbounded from below. A mean reversion effect can imply that at any time the tails of the diffusion $A$ are lighter than the tails of $\nu$; in this case $\nu$ can not be embedded into $A$ in bounded time.

   A condition related to (3.7) appears in Theorem 6 of the article [AHS15] studying the SEP in the special case where $\mu$ and $\sigma$ do only depend on $x$. The theorem states that if $-\frac{2\mu}{\sigma}+\sigma'$ is non-increasing and $\frac{g'}{\sigma(g)}$ is bounded, then there exists a bounded solution of the SEP. Note that if, in addition, $\sigma$ is constant, the assumption of Theorem 6, [AHS15], coincides with our Assumption 3.5.

d) In [FIP15] the authors consider the special case when $\mu, \sigma$ do not depend on $a$, but on time only. To obtain weak solutions for the SEP using the FBSDE approach the authors of that work assume that $\sigma$ is bounded away from zero as well as that $g'$ and $\delta'$ are bounded, where $\delta'(r) = \frac{\mu(H^{-1}(r))}{\sigma^2(H^{-1}(r))}$ and where $H^{-1}$ is the inverse of the mapping $t \mapsto \int_0^t \sigma^2(s)ds$. This boundedness of $\delta'(r)$ is equivalent to our assumption that $\frac{\mu}{\sigma^2}$ is bounded.

# 4 A decoupling field approach to position control problems

The aim in position control problems is to steer a process such that the generated costs of the process and steering are minimized. There are many applications in economics, engineering and management. However, we start with space flight as a graphic illustration of the aim of such an optimal control.

Our first example is the steering of a rocket that is supposed to land on the moon. A deviation from the desired path might cause unforeseeable problems. Hence, fuel has to be expended to make corrections to the course while flying. In the end the rocket should land on a specified place and the ground gets rougher proportionally with the distance to the center, increasing the risk of a crash while landing. And everywhere on the way there are unknown influences randomly diverting from the flight path. While designing the rocket, the engineers need to have a good estimate of how much fuel it will need. And while in the air, it has to be steered with a minimal expenditure of fuel. After the landing no more fuel is needed, which represents a time horizon for the problem.

Another example is keeping a satellite on its orbit and knowing how long it can stay there. The steering stays the same as with the rocket. However, using a finite time horizon is not well suited for this problem. It is much more convenient to assume an infinite time horizon and an unlimited amount of fuel. After obtaining the minimal amount of fuel to be used and averaging over time, one gets an estimate of how long the satellite can be kept on its orbit. This infinite horizon problem motivates the study of so called *ergodic* control problems.

As a current example for a control problem we want to highlight the outbreak of a contagious disease. The number of newly infected people is stochastic and depends on time and the number of already infected people. The governmental measures against spreading can be viewed as the control and generate costs of a social and economic nature, while the costs of infected people arise e.g. in health care. Beyond the point where all hospitals are at their limits, either the spending has to be increased radically or more people will die. Thus, it is apparent purely quadratic cost functions are sometimes insufficient for modeling.

The common approaches for solving position control problems trace back to Bellman and Pontryagin. Bellman developed the Dynamic Programming Principle, which together with the Hamilton-Jacobi-Bellman (HJB) equation makes use of PDE theory. On the other hand, Pontryagin's Maximum Principle derives the equivalence of the control problem to an FBSDE. When using Pontryagin's Maximum Principle, most works either make assumptions such that the resulting FBSDE decouples to a BSDE or, in the case of linear-quadratic control problems, exploit further dualities, which then allow to solve the control problem (see e.g. [YZ99], [SXY18]). In this latter case the solution to the FBSDE is obtained only as a byproduct. Our approach is different in that we directly solve the coupled FBSDE by using the method of decoupling fields (see Chapter 2). In essence, this means that we derive an adjoint BSDE with quadratic dynamics and show that it has a bounded solution. From there we then derive the optimal control and value function. This approach has already been applied in [AFKP18] to which we draw a short comparison in Remark 4.12.

We present two sets of sufficient conditions for the existence of an optimal control of the drift coefficient. The first set (see Assumption 4.8) contains processes with stochastic linear dynamics and convex cost functions that grow at most quadratically in the position of the process. This first set contains the linear-quadratic setting. The second set (see Assumption 4.13) consists of a Lipschitz drift and linear diffusion coefficient but has stronger restrictions on the growth of the cost functions.

In Section 4.2 we concentrate on the linear-quadratic case, which is a special case of Assumtion 4.8. Our findings conform with the known results from [YZ99] and [SXY18] (see Remark 4.23 for details).

Then, in Section 4.3, we derive some convergence results for deterministic functions with quadratic dynamics, which are used in Section 4.4 about ergodic linear-quadratic control problems. To the best of our knowlege other works about ergodic control deal either only with time homogeneous dynamics (see e.g. [BF92] and [Bor06]) or assume dissipativity of the process to be controlled (see e.g. [CFP16] and [OTV19]). We, on the other hand, restrict to a 1-dimensional linear-quadratic setting with bounded, deterministic, time inhomogeneous factors, without any other restrictions to the controlled process.

## 4.1 The problem and general solutions

Let $W$ be a standard 1-dimensional Brownian Motion on a probability space $(\Omega, \mathcal{F}, \mathbf{P})$ and let $\mathbb{F} = (\mathcal{F}_t)_{0 \leqslant t < \infty}$ be the augmented natural filtration of $W$.

**Assumption 4.1**
*Let $T > 0$ be a time horizon and the functions $\mu, \sigma : \Omega \times [0,T] \times \mathbb{R} \to \mathbb{R}$ be progressively measurable and in the last argument two times differentiable with bounded derivatives. Also, let $g : \Omega \times \mathbb{R} \to \mathbb{R}$ and $f : \Omega \times [0,T] \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ be two times differentiable and convex in all space components, $\lambda \times \mathbf{P}$ a.e. and progressively measurable. Furthermore, assume $f$ to be strictly convex in its last argument $\lambda \times \mathbf{P}$ a.e. Finally, let*

$$\| |\mu(\cdot,\cdot,0)| + |\sigma(\cdot,\cdot,0)| + |f(\cdot,\cdot,0,0)| + |g(\cdot,0)| \|_\infty < \infty.$$

Let Assumption 4.1 be fulfilled and denote by $\mathcal{A}$ the set of progressively measurable controls $\alpha : \Omega \times [0,T] \to \mathbb{R}$ with $\mathbb{E} \int_0^T \alpha_s^2 \, \mathrm{d}s < \infty$. For $x_0 \in \mathbb{R}$ and $\alpha \in \mathcal{A}$ we define $X^\alpha$ as the solution of the integral equation

$$X_t^\alpha = x_0 + \int_0^t (\mu(s, X_s^\alpha) - \alpha_s) \, \mathrm{d}s + \int_0^t \sigma(s, X_s^\alpha) \, \mathrm{d}W_s.$$

**Remark 4.2**
If Assumption 4.1 is fulfilled, $\mu$ and $\sigma$ are Lipschitz-continuous in the space argument uniformly in time. Since each admissible control is square integrable we obtain by the standard theory (see e.g. Theorem 6.3 in [YZ99]) that $X$ has a unique strong solution, which satisfies

$$\mathbb{E}\left[\sup_{t\in[0,T]} |X_t|^2\right] < \infty.$$

This moreover implies that for any term, which grows at most linear in $X$, we can interchange the order of expectation and integration by Fubini's Theorem. Since furthermore the $Z$-component of a solution of a BSDE fulfills $\mathbb{E}\left[\int_0^T Z_s^2 \,\mathrm{d}s\right] < \infty$ by definition, we likewise obtain that every stochastic integral with respect to a Brownian motion of any Lipschitz continuous function of $X$ and $Z$ is a true martingale and has thereby an expected value of $0$. In the following we will use those facts without mentioning them.

Our aim is to solve the control problem that consists in minimizing the cost functional

$$J(T, x_0, \alpha) := \mathbb{E}\left[\int_0^T f(s, X_s^\alpha, \alpha_s)\,\mathrm{d}s + g(X_T^\alpha)\right] \tag{4.1}$$

over all controls $\alpha \in \mathcal{A}$. For now we treat $T$ as a fixed constant, but in Section 4.4 we let $T$ go to infinity. The so-called Hamiltonian of the above control problem is, for $t \in [0, T]$ and $x, y, \alpha, z \in \mathbb{R}$, defined by

$$H(t, x, \alpha, y, z) := (\mu(t, x) - \alpha)y + \sigma(t, x)z + f(t, x, \alpha). \tag{4.2}$$

Pontryagin's Maximum Principle is one of the standard methods used to solve control problems. Basically it states that the control problem is equivalent to an FBSDE.

**Theorem 4.3** (see e.g. Theorem 6.4.6 in [Pha09])
*Suppose that Assumption 4.1 is fulfilled. Let $x_0 \in \mathbb{R}$, $\hat\alpha \in \mathcal{A}$ and $\hat X = X^{\hat\alpha}$ the associated controlled diffusion. Suppose that there exists a solution $(\hat Y, \hat Z)$ to the associated BSDE*

$$-\,\mathrm{d}Y_t = \partial_x H\big(t, \hat X_t, \hat\alpha_t, \hat Y_t, \hat Z_t\big)\,\mathrm{d}t - \hat Z_t\,\mathrm{d}W_t, \qquad \hat Y_T = \partial_x g\big(\hat X_T\big)$$

*such that*

$$H\big(t, \hat X_t, \hat\alpha_t, \hat Y_t, \hat Z_t\big) = \min_{\alpha\in\mathbb{R}} H\big(t, \hat X_t, \alpha, \hat Y_t, \hat Z_t\big), \qquad 0 \leqslant t \leqslant T, \quad a.s.$$

*and*

$$(x, \alpha) \mapsto H\big(t, x, \alpha, \hat Y_t, \hat Z_t\big)$$

*is a $\lambda \times \mathbf{P}$ a.e. convex function for every $t \in [0, T]$. Then $\hat\alpha$ is an optimal control, which means that*

$$J(T, x_0, \hat\alpha) = \inf_{\alpha\in\mathcal{A}} J(T, x_0, \alpha).$$

Note that, since $f$ is strictly convex in $\alpha$ whenever Assumption 4.1 is fulfilled, we get that $\partial_\alpha f$ is strictly monotone increasing. Hence, there exists a (random) inverse function of $\partial_\alpha f$ with respect to $\alpha$, which we denote with $f_\alpha^{-1}$. Furthermore, we denote by $\mathcal{D}(f_\alpha^{-1})$ the domain of $f_\alpha^{-1}$ and for constants $a, b \in [-\infty, \infty]$ with $a < b$ we define the truncation operator $\mathcal{T}_a^b$ as $\mathcal{T}_a^b(x) := \max(a, \min(x, b))$ for all $x \in \mathbb{R}$.

**Proposition 4.4**
*Let Assumption 4.1 be fulfilled and $x_0 \in \mathbb{R}$. Moreover, let $a, b, c, d \in [-\infty, \infty]$ with $a < b$, $c < d$ and $[0,T] \times \mathbb{R} \times ([a,b] \cap \mathbb{R}) \subset \mathcal{D}(f_\alpha^{-1})$ such that the FBSDE*

$$\begin{aligned} X_t &= x_0 + \int_0^t \Big[\mu(s, X_s) - f_\alpha^{-1}\big(s, X_s, \mathcal{T}_a^b(Y_s)\big)\Big]\,\mathrm{d}s + \int_0^t \sigma(s, X_s)\,\mathrm{d}W_s \\ Y_t &= \partial_x g(X_T) - \int_t^T Z_s\,\mathrm{d}W_s \\ &\quad + \int_t^T \Big[\partial_x \mu(s, X_s)\mathcal{T}_a^b(Y_s) + \partial_x \sigma(s, X_s)\mathcal{T}_c^d(Z_s) + \partial_x f\left(s, X_s, f_\alpha^{-1}\left(s, X_s, \mathcal{T}_a^b(Y_s)\right)\right)\Big]\,\mathrm{d}s \end{aligned} \tag{4.3}$$

*fulfills SLC (see Assuption 2.1) or MLLC (see Assumption 2.2). If the gradient process $U$ is bounded independently of the interval $[t,T] \subset I_{\max}$, $\lambda \times \mathbf{P}$ a.e., then it has an a.e. unique solution $(X, Y, Z)$ on the whole interval $[0,T]$. If furthermore $a \leqslant Y \leqslant b$, $c \leqslant Z \leqslant d$, $\lambda \times \mathbf{P}$ a.e.,*

$$(x, \alpha) \mapsto H(t, x, \alpha, Y_t, Z_t) \tag{4.4}$$

*is a $\lambda \times \mathbf{P}$ a.e. convex function for $t \in [0,T]$, then*

$$\big(\hat{\alpha}_t := f_\alpha^{-1}(t, X_t, Y_t)\big)_{t\in[0,T]}$$

*is an admissible, optimal control.*

*Proof.* Since FBSDE (4.3) fulfills SLC or MLLC and the gradient process is bounded, we obtain by Theorem 2.11 that FBSDE (4.3) has a unique solution on the whole interval $[0,T]$. Next, remember that $f_\alpha^{-1}$, which is the inverse of $\partial_\alpha f$ with respect to $\alpha$, is well defined because $f$ is strictly convex in $\alpha$. Since furthermore $H$ is also strictly convex in $\alpha$ and $\partial_\alpha H(t,x,\alpha,y,z) = -y + \partial_\alpha f(t,x,\alpha)$ we get for all $y$ for which $f_\alpha^{-1}$ is defined and all $(t,x,z) \in [0,T] \times \mathbb{R} \times \mathbb{R}$ that

$$\min_{\alpha\in\mathbb{R}} H(t,x,\alpha,y,z) = \big(\mu(t,x) - f_\alpha^{-1}(t,x,y)\big)\,y + \sigma(t,x)z + f\left(t, x, f_\alpha^{-1}(t,x,y)\right). \tag{4.5}$$

Hence, for specific processes $X$, $Y$ such that $(t, X_t, Y_t) \in [0,T] \times \mathbb{R} \times ([a,b] \cap \mathbb{R}) \subset \mathcal{D}(f_\alpha^{-1})$, we have with $\hat{\alpha}_t = f_\alpha^{-1}(t, X_t, Y_t)$ a candidate for the optimal control. Since FBSDE (4.3) fulfills SLC or MLLC, which implies that $f_\alpha^{-1}$ is Lipschitz-continuous in $x$ and $y$, we also get that $\hat{\alpha}$ is admissible.

Now, note that if $a \leqslant Y_t \leqslant b$ for all $t \in [0,T]$ then $\mathcal{T}_a^b(Y_t) = Y_t$ and likewise for $Z$. Thus, we get that FBSDE (4.3) is equivalent to the FBSDE

$$\begin{aligned} X_t &= x_0 + \int_0^t \Big[\mu(s, X_s) - f_\alpha^{-1}(s, X_s, Y_s)\Big]\,\mathrm{d}s + \int_0^t \sigma(s, X_s)\,\mathrm{d}W_s \\ Y_t &= \partial_x g(X_T) + \int_t^T \Big[\partial_x \mu(s, X_s)Y_s + \partial_x \sigma(s, X_s)Z_s + \partial_x f\left(s, X_s, f_\alpha^{-1}(s, X_s, Y_s)\right)\Big]\,\mathrm{d}s. \\ &\quad - \int_t^T Z_s\,\mathrm{d}W_s \end{aligned}$$

This in turn yields, together with Equation (4.5), that $\hat{\alpha} := f_\alpha^{-1}(\cdot, X, Y)$ and $(Y, Z)$ fulfill the requirements of Theorem 4.3. Hence $\hat{\alpha}$, as defined above, is an admissible, optimal control. ∎

In the remainder of this section we derive conditions such that the assumptions in Proposition 4.4 are fulfilled, which allows us to obtain a solution to the control problem.

For this section let $(X, Y, Z)$ always be the solution of FBSDE (4.3) on $[t,T] \subset I_{\max}$ with initial value $x_0 \in \mathbb{R}$.

**Lemma 4.5**
*Assume that $\mu, \sigma, f, g$ fulfill Assumption 4.1 and that $\mu, \sigma, f, g, a, b, c, d$ are such that FBSDE* (4.3) *fulfills SLC or MLLC. Then the gradient process $U$ of FBSDE* (4.3) *solves for all $s \in [t,T] \subset I_{\max}$ the BSDE*

$$U_s = g''(X_T) + \int_s^T h\big(r, U_r, \tilde{Z}_r\big)\,\mathrm{d}r - \int_s^T \tilde{Z}_r\,\mathrm{d}W_r, \tag{4.6}$$

*where*

$$\begin{aligned} h(r,v,z) := \Big[ &- v^2 \partial_y (f_\alpha^{-1})(r, X_r, \mathcal{T}_a^b(Y_r)) \mathbb{1}_{[a,b]}(Y_r) \\ &+ v \left( \partial_x \mu(r, X_r)(1 + \mathbb{1}_{[a,b]}(Y_r)) - \partial_x (f_\alpha^{-1})(r, X_r, \mathcal{T}_a^b(Y_r)) + (\partial_x \sigma)^2 (r, X_r) \mathbb{1}_{[c,d]}(Z_r) \right) \\ &+ v \left( \partial_{x\alpha} f(r, X_r, f_\alpha^{-1}(r, X_r, \mathcal{T}_a^b(Y_r))) \partial_y (f_\alpha^{-1})(r, X_r, \mathcal{T}_a^b(Y_r)) \mathbb{1}_{[a,b]}(Y_r) \right) \\ &+ \partial_{xx} \mu(r, X_r) \mathcal{T}_a^b(Y_r) + \partial_{xx} \sigma(r, X_r) \mathcal{T}_c^d(Z_r) + \partial_{xx} f(r, X_r, f_\alpha^{-1}(r, X_r, \mathcal{T}_a^b(Y_r))) \\ &+ \partial_{x\alpha} f(r, X_r, f_\alpha^{-1}(r, X_r, \mathcal{T}_a^b(Y_r))) \partial_x (f_\alpha^{-1})(r, X_r, \mathcal{T}_a^b(Y_r)) \\ &+ z \partial_x \sigma(r, X_r)(1 + \mathbb{1}_{[c,d]}(Z_r)) \Big] \end{aligned}$$

*for $(r, v, z) \in [t,T] \times \mathbb{R} \times \mathbb{R}$ and with $(X, Y, Z)$ from FBSDE* (4.3).

*Proof.* For $x, y, z \in \mathbb{R}$, $s \in [t,T] \subset I_{\max}$ define

$$M(s,x,y) := \mu(s,x) - f_\alpha^{-1}\left(s, x, \mathcal{T}_a^b(y)\right),$$

and

$$F(s,x,y,z) := \partial_x \mu(s,x) \mathcal{T}_a^b(y) + \partial_x \sigma(s,x) \mathcal{T}_c^d(z) + \partial_x f\left(s, x, f_\alpha^{-1}\left(s, x, \mathcal{T}_a^b(y)\right)\right).$$

Then, for an initial value $x_0 \in \mathbb{R}$ at time $t$, i.e. $X_t = x_0$, FBSDE (4.3) can be written as

$$\begin{aligned} X_s &= x_0 + \int_t^s M(r, X_r, Y_r)\,\mathrm{d}r + \int_t^s \sigma(r, X_r)\,\mathrm{d}W_r, \\ Y_s &= \partial_x g(X_T) + \int_s^T F(r, X_r, Y_r, Z_r)\,\mathrm{d}r - \int_s^T Z_r\,\mathrm{d}W_r. \end{aligned}$$

Let $u$ be the decoupling field of FBSDE (4.3). With the decoupling condition and the chain rule in Lemma A.3.1 of [Fro15] we get

$$\partial_{x_0} Y_s = \partial_{x_0}\big[u(s, X_s)\big] = \partial_x u(s, X_s) \cdot \partial_{x_0} X_s = U_s \cdot \partial_{x_0} X_s.$$

Now, define a stopping time $\tau$ via

$$\tau := \inf\{s \in [t,T] | \partial_{x_0} X_s \leqslant 0\} \wedge T.$$

Notice that $\tau > t$ since $\partial_{x_0} X_t = 1$ and $\partial_{x_0} X_t$ is an Itô process and in particular continuous in time (see Lemma A.2.5 and Lemma A.2.6 in [Fro15]). For all $s \in [t, \tau)$ we have that $\partial_{x_0} X_s$ is invertible with $(\partial_{x_0} X_s)^{-1}$ being an Itô process, too. Hence, $U_s = \partial_{x_0} Y_s \cdot (\partial_{x_0} X_s)^{-1}$ is an Itô process and thus there exist $(b_s)$ and $(\tilde{Z}_s)$ such that

$$U_s = U_T + \int_s^T b_r\,\mathrm{d}r - \int_s^T \tilde{Z}_r\,\mathrm{d}W_r$$

for all $s \in [t, \tau)$.

We also introduce for an Itô process $I_s = I_0 - \int_0^s i_r \, \mathrm{d}r - \int_0^s j_r \, \mathrm{d}W_r$ the two operators $\mathrm{D}^t$ and $\mathrm{D}^w$ defined via $(\mathrm{D}^t \, I)_s := i_s$ and $(\mathrm{D}^w \, I)_s := j_s$. Note that due to Lemma A.2.5 and Lemma A.2.6 in [Fro15] the integrals and hence also the operators $\mathrm{D}^w$ and $\mathrm{D}^t$ can be interchanged with the weak differentiation $\partial_{x_0}$. Using this notation we have

$$\begin{aligned} \partial_{x_0} Z_s &= \mathrm{D}^w \, \partial_{x_0} Y_s \\ &= \mathrm{D}^w \left( U_s \cdot \partial_{x_0} X_s \right) \\ &= U_s \cdot \mathrm{D}^w \, \partial_{x_0} X_s + \mathrm{D}^w \, U_s \cdot \partial_{x_0} X_s \\ &= U_s \cdot \partial_x \sigma \left( s, X_s \right) \partial_{x_0} X_s + \tilde{Z}_s \partial_{x_0} X_s . \end{aligned}$$

Thus we get

$$\tilde{Z}_s = \partial_{x_0} Z_s \left( \partial_{x_0} X_s \right)^{-1} - U_s \, \partial_x \sigma \left( s, X_s \right)$$

for all $s \in [t, \tau)$. Also,

$$\begin{aligned} \partial_{x_0} \left[ M \left( s, X_s, Y_s \right) \right] &= \partial_x M \left( s, X_s, Y_s \right) \partial_{x_0} X_s + \partial_y M \left( s, X_s, Y_s \right) \partial_{x_0} Y_s \\ &= \partial_x M \left( s, X_s, Y_s \right) \partial_{x_0} X_s + \partial_y M \left( s, X_s, Y_s \right) U_s \partial_{x_0} X_s \end{aligned}$$

and using the dynamics of $Y$ yields

$$\begin{aligned} \mathrm{D}^t \, \partial_{x_0} Y_s &= -\partial_x F \left( s, X_s, Y_s, Z_s \right) \partial_{x_0} X_s - \partial_y F \left( s, X_s, Y_s, Z_s \right) \partial_{x_0} Y_s - \partial_z F \left( s, X_s, Y_s, Z_s \right) \partial_{x_0} Z_s \\ &= -\partial_x F \left( s, X_s, Y_s, Z_s \right) \partial_{x_0} X_s - \partial_y F \left( s, X_s, Y_s, Z_s \right) U_s \partial_{x_0} X_s \\ &\qquad - \partial_z F \left( s, X_s, Y_s, Z_s \right) \left( U_s \cdot \partial_x \sigma \left( s, X_s \right) + \tilde{Z}_s \right) \partial_{x_0} X_s , \end{aligned}$$

while we obtain with the decoupling condition that

$$\begin{aligned} \mathrm{D}^t \, \partial_{x_0} Y_s &= \mathrm{D}^t (U_s \partial_{x_0} X_s) \\ &= -b_s \cdot \partial_{x_0} X_s + U_s \cdot \left( \partial_x M \left( s, X_s, Y_s \right) \partial_{x_0} X_s + \partial_y M \left( s, X_s, Y_s \right) U_s \partial_{x_0} X_s \right) \\ &\qquad + \tilde{Z}_s \partial_x \sigma (s, X_s) \partial_{x_0} X_s . \end{aligned}$$

Equating the two representations of $\mathrm{D}^t \, \partial_{x_0} Y$ and rearranging yields

$$\begin{aligned} b_s = {} & U_s \left[ \partial_x M \left( s, X_s, Y_s, \right) + \partial_y M \left( s, X_s, Y_s, \right) U_s \right] + \tilde{Z}_s \partial_x \sigma (s, X_s) \\ & + \partial_x F \left( s, X_s, Y_s, Z_s \right) + \partial_y F \left( s, X_s, Y_s, Z_s \right) U_s + \partial_z F \left( s, X_s, Y_s, Z_s \right) \left( U_s \cdot \partial_x \sigma \Big( s, X_s \Big) + \tilde{Z}_s \right) \end{aligned}$$

for all $s \in [t, \tau)$ with

$$\begin{aligned} \partial_x M(s,x,y) &= \partial_x \mu(s,x) - \partial_x (f_\alpha^{-1}) \left( s, x, \mathcal{T}_a^b(y) \right) \\ \partial_y M(s,x,y) &= -\partial_y (f_\alpha^{-1}) \left( s, x, \mathcal{T}_a^b(y) \right) \mathbb{1}_{[a,b]}(y) \\ \partial_x F(s,x,y,z) &= \partial_{xx} \mu(s,x) \mathcal{T}_a^b(y) + \partial_{xx} \sigma(s,x) \mathcal{T}_c^d(z) + \partial_{xx} f \left( s, x, f_\alpha^{-1} \left( s, x, \mathcal{T}_a^b(y) \right) \right) \\ &\qquad + \partial_{x\alpha} f \left( s, x, f_\alpha^{-1} \left( s, x, \mathcal{T}_a^b(y) \right) \right) \partial_x (f_\alpha^{-1}) \left( s, x, \mathcal{T}_a^b(y) \right) \\ \partial_y F(s,x,y,z) &= \partial_x \mu(s,x) \mathbb{1}_{[a,b]}(y) + \partial_{x\alpha} f \left( s, x, f_\alpha^{-1} \left( s, x, \mathcal{T}_a^b(y) \right) \right) \partial_y (f_\alpha^{-1}) \left( s, x, \mathcal{T}_a^b(y) \right) \mathbb{1}_{[a,b]}(y) \\ \partial_z F(s,x,y,z) &= \partial_x \sigma(s,x) \mathbb{1}_{[c,d]}(z) . \end{aligned}$$

Next we turn our attention to the question whether $\partial_{x_0} X$ is invertible on the whole interval $[t, T]$. Observe that

$$\begin{aligned}\partial_{x_0} X_s &= \mathrm{Id} + \int_t^s \partial_{x_0} \left[M\left(r, X_r, Y_r\right)\right] \mathrm{d}r + \int_t^s \partial_{x_0} \left[\sigma(r, X_r)\right] \mathrm{d}W_r \\ &= \mathrm{Id} + \int_t^s \left[\partial_x M\left(r, X_r, Y_r\right) + \partial_y M\left(r, X_r, Y_r\right) U_r\right] \partial_{x_0} X_r \,\mathrm{d}r + \int_t^s \partial_x \sigma\left(r, X_r\right) \partial_{x_0} X_r \,\mathrm{d}W_r\end{aligned}$$

implying that

$$\begin{aligned}\partial_{x_0} X_s = \exp\Bigg( \int_t^s &\left[\partial_x M\left(r, X_r, Y_r\right) + \partial_y M\left(r, X_r, Y_r\right) U_r - \frac{1}{2}\left(\partial_x \sigma\right)^2 \left(r, X_r\right)\right] \mathrm{d}r \\ &+ \int_t^s \partial_x \sigma(r, X_r) \,\mathrm{d}W_r \Bigg). \qquad (4.7)\end{aligned}$$

Note that all coefficients in (4.7) are bounded on $[t, T]$ giving that $\partial_{x_0} X_s > 0$ for all $s \in [t, T]$. Therefore, $\partial_{x_0} X_s$ is invertible on the whole interval $[t, T]$ and $\tau = T$.

What remains to do is to calculate the explicit dynamics of $U$. Observe that

$$\begin{aligned}b_s &= U_s \left[\partial_x M\left(s, X_s, Y_s\right) + \partial_y M\left(s, X_s, Y_s\right) U_s\right] + \tilde{Z}_s \partial_x \sigma(s, X_s) \\ &\quad + \partial_x F\left(s, X_s, Y_s, Z_s\right) + \partial_y F\left(s, X_s, Y_s, Z_s\right) U_s + \partial_z F\left(s, X_s, Y_s, Z_s\right) \left(U_s \cdot \partial_x \sigma\left(s, X_s\right) + \tilde{Z}_s\right) \\ &= U_s \left[\partial_x \mu(s, X_s) - \partial_x (f_\alpha^{-1}) \left(s, X_s, \mathcal{T}_a^b(Y_s)\right)\right] - U_s \partial_y (f_\alpha^{-1}) \left(s, X_s, \mathcal{T}_a^b(Y_s)\right) \mathbb{1}_{[a,b]}(Y_s) U_s \\ &\quad + \tilde{Z}_s \partial_x \sigma(s, X_s) + \partial_{xx} \mu(s, X_s) \mathcal{T}_a^b(Y_s) + \partial_{xx} \sigma(s, X_s) \mathcal{T}_c^d(Z_s) + \partial_{xx} f\big(s, X_s, f_\alpha^{-1}\big(s, X_s, \mathcal{T}_a^b(Y_s)\big)\big) \\ &\quad + \partial_{x\alpha} f\left(s, X_s, f_\alpha^{-1}\left(s, X_s, \mathcal{T}_a^b(Y_s)\right)\right) \partial_x (f_\alpha^{-1}) \left(s, X_s, \mathcal{T}_a^b(Y_s)\right) \\ &\quad + \left[\partial_x \mu(s, X_s) \mathbb{1}_{[a,b]}(Y_s) + \partial_{x\alpha} f\big(s, X_s, f_\alpha^{-1}\big(s, X_s, \mathcal{T}_a^b(Y_s)\big)\big) \partial_y (f_\alpha^{-1})\big(s, X_s, \mathcal{T}_a^b(Y_s)\big) \mathbb{1}_{[a,b]}(Y_s)\right] U_s \\ &\quad + \partial_x \sigma(s, X_s) U_s \partial_x \sigma(s, X_s) \mathbb{1}_{[c,d]}(Z_s) + \partial_x \sigma(s, X_s) \mathbb{1}_{[c,d]}(Z_s) \tilde{Z}_s \\ &= -U_s^2 \partial_y (f_\alpha^{-1})(s, X_s, \mathcal{T}_a^b(Y_s)) \mathbb{1}_{[a,b]}(Y_s) \\ &\quad + U_s \left[\partial_x \mu(s, X_s)(1 + \mathbb{1}_{[a,b]}(Y_s)) - \partial_x (f_\alpha^{-1})\big(s, X_s, \mathcal{T}_a^b(Y_s)\big) + \left(\partial_x \sigma\right)^2 \left(s, X_s\right) \mathbb{1}_{[c,d]}(Z_s)\right] \\ &\quad + U_s \left[\partial_{x\alpha} f\big(s, X_s, f_\alpha^{-1}\big(s, X_s, \mathcal{T}_a^b(Y_s)\big)\big) \partial_y (f_\alpha^{-1})(s, X_s, \mathcal{T}_a^b(Y_s)) \mathbb{1}_{[a,b]}(Y_s)\right] \\ &\quad + \partial_{xx} \mu(s, X_s) \mathcal{T}_a^b(Y_s) + \partial_{xx} \sigma(s, X_s) \mathcal{T}_c^d(Z_s) + \partial_{xx} f(s, X_s, f_\alpha^{-1}(s, X_s, \mathcal{T}_a^b(Y_s))) \\ &\quad + \partial_{x\alpha} f(s, X_s, f_\alpha^{-1}(s, X_s, \mathcal{T}_a^b(Y_s))) \partial_x (f_\alpha^{-1})(s, X_s, \mathcal{T}_a^b(Y_s)) + (1 + \mathbb{1}_{[c,d]}(Z_s)) \partial_x \sigma(s, X_s) \tilde{Z}_s.\end{aligned}$$

Finally note that

$$U_T = \partial_{x_0} Y_T (\partial_{x_0} X_T)^{-1} = \left[\partial_{x_0} g'(X_T)\right] (\partial_{x_0} X_T)^{-1} = g''(X_T).$$

■

**Remark 4.6**
Note that by Lemma 4.5 $U$ solves some kind of stochastic Riccati equation. Under sufficent conditions (see Remark 4.20) this Riccati equation coincides with the Equation (6.1) considered in [SXY18].

For shorter notation we define the function $\varphi : \mathcal{D}(f_\alpha^{-1}) \to [0, T] \times \mathbb{R} \times \mathbb{R}$ as

$$\varphi(s, x, y) := (s, x, f_\alpha^{-1}(s, x, y)).$$

**Lemma 4.7**
*Assume that $\mu, \sigma, f, g$ fulfill Assumption 4.1, $\mu, \sigma, f, g, a, b, c, d$ are such that FBSDE* (4.3) *fulfills SLC or MLLC and $[t, T] \subset I_{\max}$. If additionally $\partial_{\alpha\alpha} f\left(\varphi\left(s, X_s, \mathcal{T}_a^b(Y_s)\right)\right) > 0$ for all $s \in [t, T]$, then for all $s \in [t, T]$ the gradient process $U_s$ of FBSDE* (4.3) *solves the BSDE*

$$
\begin{aligned}
U_s = g''(X_T) + \int_s^T \Bigg[ & - U_r^2 \frac{\mathbb{1}_{[a,b]}(Y_r)}{\partial_{\alpha\alpha} f\left(\varphi\left(r, X_r, \mathcal{T}_a^b(Y_r)\right)\right)} + U_r \left( \left(\partial_x \sigma\right)^2 (r, X_r) \mathbb{1}_{[c,d]}(Z_s) \right) \\
& + U_r \left( (1 + \mathbb{1}_{[a,b]}(Y_r)) \left( \frac{\partial_{x\alpha} f}{\partial_{\alpha\alpha} f} \left( \varphi\left(r, X_r, \mathcal{T}_a^b(Y_r)\right)\right) + \partial_x \mu(r, X_r) \right) \right) \\
& + \partial_{xx} \mu(r, X_r) \mathcal{T}_a^b(Y_r) + \partial_{xx} \sigma(r, X_r) \mathcal{T}_c^d(Z_r) + \partial_{xx} f \left( \varphi\left(r, X_r, \mathcal{T}_a^b(Y_r)\right)\right) \\
& - \frac{\left(\partial_{x\alpha} f\right)^2}{\partial_{\alpha\alpha} f} \left( \varphi\left(r, X_r, \mathcal{T}_a^b(Y_r)\right)\right) + (1 + \mathbb{1}_{[c,d]}(Z_r)) \partial_x \sigma(r, X_r) \tilde{Z}_r \Bigg] \mathrm{d}r \\
& - \int_s^T \tilde{Z}_r \, \mathrm{d}W_r
\end{aligned}
$$

*and*

$$
\begin{aligned}
\partial_x (f_\alpha^{-1})(s, X_s, \mathcal{T}_a^b(Y_s)) &= - \frac{\partial_{x\alpha} f}{\partial_{\alpha\alpha} f} \left( \varphi\left(r, X_r, \mathcal{T}_a^b(Y_r)\right)\right), \\
\partial_y (f_\alpha^{-1})(s, X_s, \mathcal{T}_a^b(Y_s)) &= \frac{1}{\partial_{\alpha\alpha} f} \left( \varphi\left(r, X_r, \mathcal{T}_a^b(Y_r)\right)\right).
\end{aligned}
$$

*Proof.* Take a look at the derivatives of $f_\alpha^{-1}$. Observe that by definition $y = \partial_\alpha f(s, x, f_\alpha^{-1}(s, x, y))$ and hence

$$
0 = \partial_x \left[ \partial_\alpha f(s, x, f_\alpha^{-1}(s, x, y)) \right] = \partial_{x\alpha} f(s, x, f_\alpha^{-1}(s, x, y)) + \partial_{\alpha\alpha} f(s, x, f_\alpha^{-1}(s, x, y)) \partial_x (f_\alpha^{-1})(s, x, y)
$$

and

$$
1 = \partial_y \left[ \partial_\alpha f(s, x, f_\alpha^{-1}(s, x, y)) \right] = \partial_{\alpha\alpha} f(s, x, f_\alpha^{-1}(s, x, y)) \partial_y (f_\alpha^{-1})(s, x, y)
$$

yielding with $\partial_{\alpha\alpha} f\left(\varphi\left(s, X_s, \mathcal{T}_a^b(Y_s)\right)\right) > 0$ that

$$
\begin{aligned}
\partial_x (f_\alpha^{-1})(s, X_s, \mathcal{T}_a^b(Y_s)) &= - \frac{\partial_{x\alpha} f}{\partial_{\alpha\alpha} f} \left( \varphi\left(s, X_s, \mathcal{T}_a^b(Y_s)\right)\right), \\
\partial_y (f_\alpha^{-1}) \left(s, X_s, \mathcal{T}_a^b(Y_s)\right) &= \frac{1}{\partial_{\alpha\alpha} f} \left( \varphi\left(s, X_s, \mathcal{T}_a^b(Y_s)\right)\right).
\end{aligned}
$$

Plugging these two identities into BSDE (4.6) (given by Lemma 4.5), yields the desired result. ■

By $\mathcal{H}(f)$ we denote the Hessian of $f$ but only with respect to the space arguments. I.e.

$$
\mathcal{H}(f)(s, x, a) = \begin{pmatrix} \partial_{xx} f & \partial_{x\alpha} f \\ \partial_{x\alpha} f & \partial_{\alpha\alpha} f \end{pmatrix} (s, x, a)
$$

and

$$
\det(\mathcal{H}(f))(s, x, a) = \partial_{xx} f(s, x, a) \partial_{\alpha\alpha} f(s, x, a) - \left(\partial_{x\alpha} f\right)^2 (s, x, a)
$$

for $(s, x, a) \in [0, T] \times \mathbb{R} \times \mathbb{R}$.

**Assumption 4.8**
*Let $T > 0$, $\mu$ and $\sigma$ be of the form*

$$\mu(t,x) = b_t + B_t x, \qquad \sigma(t,x) = c_t + C_t x$$

*for $b, B, c, C : \Omega \times [0,T] \to \mathbb{R}$ being progressively measurable and bounded processes. Also assume that*

1. *$g : \Omega \times \mathbb{R} \to \mathbb{R}$ and $f : \Omega \times [0,T] \times \mathbb{R} \times \mathbb{R}$ are twice differentiable and convex in the space arguments, $\lambda \times \mathbf{P}$ a.e., and progressively measurable,*
2. *$\left\|f_\alpha^{-1}(\cdot,\cdot,0,0)\right\|_\infty, \|\partial_x f(\cdot,\cdot,0,f_\alpha^{-1}(\cdot,\cdot,0,0))\|_\infty < \infty$,*
3. *$\partial_{\alpha\alpha} f \geqslant \varepsilon > 0$, $\mathbf{P}$ a.s.,*
4. *$\frac{\det(\mathcal{H}(f))}{\partial_{\alpha\alpha} f}$ and $\frac{\partial_{x\alpha} f}{\partial_{\alpha\alpha} f}$ are bounded, $\mathbf{P}$ a.s.,*
5. *$\|g(\cdot,0)\|_\infty, \|\partial_x g(\cdot,0)\|_\infty, \|\partial_{xx} g\|_\infty < \infty$.*

Note that Assumption 4.8 is a more specific case of Assumption 4.1, allowing us to apply the results from e.g. Proposition 4.4 and Lemma 4.7.

**Lemma 4.9**
*Let $\sigma$, $\mu$, $f$ and $g$ fulfill Assumption 4.8. Then FBSDE* (4.3) *fulfills SLC for $a = c = -\infty$, $b = d = \infty$ and the gradient process $U$ is bounded by*

$$0 \leqslant U \leqslant \left(\|g''\|_\infty + \left\|\partial_{xx} f - \frac{(\partial_{x\alpha} f)^2}{\partial_{\alpha\alpha} f}\right\|_\infty\right) \exp\left(T\left(\|C^2\|_\infty + 2\|B\|_\infty + 2\left\|\frac{\partial_{x\alpha} f}{\partial_{\alpha\alpha} f}\right\|_\infty\right)\right) < \infty,$$

*$\lambda \times \mathbf{P}$ a.e.*

*Proof.* Observe that with Lemma 4.7, for all $s \in [0,T]$, $x', y \in \mathbb{R}$,

$$\partial_x(f_\alpha^{-1})(s,x',y) = -\frac{\partial_{x\alpha} f}{\partial_{\alpha\alpha} f}\left(s,x',f_\alpha^{-1}(s,x',y)\right) \quad \text{and} \quad \partial_y(f_\alpha^{-1})(s,x',y) = \frac{1}{\partial_{\alpha\alpha} f}\left(s,x',f_\alpha^{-1}(s,x',y)\right)$$

giving

$$\partial_{x'}\left[\partial_x f(s,x',f_\alpha^{-1}(s,x',y))\right] = \frac{\det(\mathcal{H}(f))}{\partial_{\alpha\alpha} f}(s,x',f_\alpha^{-1}(s,x',y))$$

and

$$\partial_y\left[\partial_x f(s,x',f_\alpha^{-1}(s,x',y))\right] = \frac{\partial_{x\alpha} f}{\partial_{\alpha\alpha} f}(s,x',f_\alpha^{-1}(s,x',y)),$$

which are both bounded, by Assumption 4.8. Because for all $x' \in \mathbb{R}$ we have $\partial_x \mu(t,x') = B_t$ and $\partial_x \sigma(t,x') = C_t$, which are also bounded, FBSDE (4.3) reduces for $a = c = -\infty$, $b = d = \infty$ and $x_0 \in \mathbb{R}$ to

$$\begin{aligned}
X_t &= x_0 + \int_0^t \left[b_s + B_s X_s - f_\alpha^{-1}(s,X_s,Y_s)\right] \mathrm{d}s + \int_0^t [c_s + C_s X_s]\, \mathrm{d}W_s \\
Y_t &= \partial_x g(X_T) + \int_t^T \left[B_s Y_s + C_s Z_s + \partial_x f(s,X_s,f_\alpha^{-1}(s,X_s,Y_s))\right] \mathrm{d}s - \int_t^T Z_s\, \mathrm{d}W_s,
\end{aligned}$$

which is Lipschitz in $X$, $Y$, and $Z$. Since furthermore $\|b\|_\infty$, $\|c\|_\infty$, $\|\partial_x g(\cdot,0)\|_\infty$, $\|\partial_{xx} g\|_\infty$, $\left\|f_\alpha^{-1}(\cdot,\cdot,0,0)\right\|_\infty$ and $\|\partial_x f(\cdot,\cdot,0,f_\alpha^{-1}(\cdot,\cdot,0,0))\|_\infty$ are all finite, we obtain that SLC is fulfilled.

Also the dynamics of the gradient process $U$ given by Lemma 4.7 simplify to

$$\begin{aligned}U_s = g''(X_T) + \int_s^T \Bigg[ &- U_r^2 \frac{1}{\partial_{\alpha\alpha} f(r,X_r,f_\alpha^{-1}(r,X_r,Y_r))} + U_r\left(C_r^2 + 2B_r\right) \\ &+ U_r\left(2\frac{\partial_{x\alpha} f(r,X_r,f_\alpha^{-1}(r,X_r,Y_r))}{\partial_{\alpha\alpha} f(r,X_r,f_\alpha^{-1}(r,X_r,Y_r))}\right) \\ &+ \left(\frac{1}{\partial_{\alpha\alpha} f}\left(\partial_{xx} f \partial_{\alpha\alpha} f - (\partial_{x\alpha} f)^2\right)\right)(r,X_r,f_\alpha^{-1}(r,X_r,Y_r))\Bigg] \mathrm{d}r - \int_s^T \tilde{Z}_r \, \mathrm{d}\widetilde{W}_r,\end{aligned}$$

where $\widetilde{W}_s = W_s - \int_s^T 2\partial_x \sigma(r,X_r)\,\mathrm{d}r$ for all $s \in [t,T]$. Since $\partial_x \sigma$ is bounded we get by Girsanov's theorem that there exists an equivalent probability measure $\widetilde{\mathbf{P}}$ under which $\widetilde{W}$ is a Brownian motion. By $\widetilde{\mathbb{E}}$ we denote the corresponding expectation operator.

Observe that, since $U$ is bounded on every closed subinterval of $I_{\max}$ because the corresponding decoupling field $u$ is weakly regular by Theorem 2.8, we can interpret the dynamics of $U$ as a Lipschitz BSDE allowing us to apply the Comparison Theorem (see e.g. Theorem 6.2.2 in [Pha09]). Note that the Hessian of a convex function is positive-semidefinite and hence its determinant is greater than or equal to zero. Thus and since in addition $\partial_{\alpha\alpha} f \geqslant \varepsilon > 0$, we have $\left(\frac{1}{\partial_{\alpha\alpha} f}\left(\partial_{xx} f \partial_{\alpha\alpha} f - (\partial_{x\alpha} f)^2\right)\right) \geqslant 0$. Because $g'' \geqslant 0$, due to convexity, and

$$\begin{aligned}\check{U}_s = \int_s^T &\left[-U_r^2 \frac{1}{\partial_{\alpha\alpha} f(r,X_r,f_\alpha^{-1}(r,X_r,Y_r))} + U_r\left(C_r^2 + 2B_r + 2\frac{\partial_{x\alpha} f(r,X_r,f_\alpha^{-1}(r,X_r,Y_r))}{\partial_{\alpha\alpha} f(r,X_r,f_\alpha^{-1}(r,X_r,Y_r))}\right)\right] \mathrm{d}r \\ &- \int_s^T \check{Z}_r \, \mathrm{d}\widetilde{W}_r\end{aligned}$$

has the trivial solution $(\check{U},\check{Z}) = (0,0)$, we get via the Comparison Theorem that $U$ is bounded from below by $0$. Using the Comparison Theorem again on $U$ and

$$\begin{aligned}\hat{U}_s = g''(X_T) + \int_s^T \Bigg[ &U_r\left(C_r^2 + 2B_r + 2\frac{\partial_{x\alpha} f(r,X_r,f_\alpha^{-1}(r,X_r,Y_r))}{\partial_{\alpha\alpha} f(r,X_r,f_\alpha^{-1}(r,X_r,Y_r))}\right) \\ &+ \left(\frac{1}{\partial_{\alpha\alpha} f}\left(\partial_{xx} f \partial_{\alpha\alpha} f - (\partial_{x\alpha} f)^2\right)\right)(r,X_r,f_\alpha^{-1}(r,X_r,Y_r))\Bigg] \mathrm{d}r - \int_s^T \hat{Z}_r \, \mathrm{d}\widetilde{W}_r\end{aligned}$$

yields that $U$ is also bounded from above since

$$\begin{aligned}\hat{U}_s &= \widetilde{\mathbb{E}}\left[ g''(X_T)\Gamma(s,T) + \int_s^T \left(\frac{1}{\partial_{\alpha\alpha} f}\left(\partial_{xx} f \partial_{\alpha\alpha} f - \partial_{x\alpha} f^2\right)\right)(r,X_r,f_\alpha^{-1}(r,X_r,Y_r))\Gamma(s,r)\,\mathrm{d}r \,\middle|\, \mathcal{F}_s\right] \\ &\leqslant \left(\|g''\|_\infty + T\left\|\frac{\det(\mathcal{H}(f))}{\partial_{\alpha\alpha} f}\right\|_\infty\right)\exp\left(T\left(\|C^2\|_\infty + 2\|B\|_\infty + 2\left\|\frac{\partial_{x\alpha} f}{\partial_{\alpha\alpha} f}\right\|_\infty\right)\right),\end{aligned}$$

where

$$\Gamma(t,s) := \exp\left(\int_t^s C_r^2 + 2B_r + 2\frac{\partial_{x\alpha} f(r,X_r,f_\alpha^{-1}(r,X_r,Y_r))}{\partial_{\alpha\alpha} f(r,X_r,f_\alpha^{-1}(r,X_r,Y_r))}\,\mathrm{d}r - \frac{1}{2}\int_t^s \tilde{Z}_r^2\,\mathrm{d}r\right).$$

To sum up, we have that

$$0 \leqslant U \leqslant \left(\|g''\|_\infty + \left\|\frac{\det(\mathcal{H}(f))}{\partial_{\alpha\alpha} f}\right\|_\infty\right)\exp\left(T\left(\|C^2\|_\infty + 2\|B\|_\infty + 2\left\|\frac{\partial_{x\alpha} f}{\partial_{\alpha\alpha} f}\right\|_\infty\right)\right) < \infty,$$

$\lambda \times \mathbf{P}$ a.e. ■

Now we come to our first main result of this section.

**Theorem 4.10**
*Let $\sigma$, $\mu$, $f$ and $g$ fulfill Assumption 4.8. Then the control problem of minimizing* (4.1) *has the optimal control*

$$\hat{\alpha}_t := f_\alpha^{-1}(t, X_t, Y_t),$$

*where $X$, $Y$ solve FBSDE* (4.3)*, which has a unique solution on the whole interval* $[0,T]$.

*Proof.* Since Assumption 4.8 is a more specific case of Assumption 4.1 and Lemma 4.9 shows that for $a = c = -\infty$ and $b = d = \infty$ FBSDE (4.3) fulfills SLC and the gradient process $U$ is $\lambda \times \mathbf{P}$ a.e. bounded, we only need to show the following two points in order to obtain the claimed result by Proposition 4.4. Firstly that $H$ is convex in $(x, a)$, which is straightforward since its only nonlinear part is the convex function $f$. And secondly that $[0,T] \times \mathbb{R} \times ([a,b] \cap \mathbb{R}) \subset \mathcal{D}(f_\alpha^{-1})$. This follows since $\partial_{\alpha\alpha} f \geqslant \varepsilon > 0$ and hence $\partial_\alpha f$ has range $\mathbb{R}$. Thus, the domain of $f_\alpha^{-1}$, which is the inverse of $\partial_\alpha f$ in $a$, has domain $\mathcal{D}(f_\alpha^{-1}) = [0,T] \times \mathbb{R} \times \mathbb{R} = [0,T] \times \mathbb{R} \times ([a,b] \cap \mathbb{R})$. ∎

**Example 4.11**
Let $b, B, c, C, r, F^a, F^x$ be bounded processes and $G_0, G_x$ bounded random variables, such that $F^a \geqslant \varepsilon > 0$, $F^x \in [0.5, 1] \cup \{0\}$ and $G_0, G_x > 0$. Define

$$\mu(s,x) := b_s + B_s \cdot x, \qquad \sigma(s,x) := c_s + C_s \cdot x,$$
$$f(s,x,a) := e^{-r_s \cdot s}\left(\left(1+x^2\right)^{F_s^x} + \cosh\left(F_s^a \cdot a\right)\right), \qquad g(x) := \left(G_0 + G_x \cdot x^2\right)^{2/3}$$

for $(s,x) \in [0,T] \times \mathbb{R}$. Then Theorem 4.10 states that minimizing the cost fuctional $J$ from Equation (4.1) over all admissible controls has the optimal solution

$$\hat{\alpha}_s = \frac{\sinh^{-1}\left(\frac{Y_s \exp(r_s \cdot s)}{F_s^a}\right)}{F_s^a},$$

where $Y$ is part of the unique solution of FBSDE (4.3).

**Remark 4.12**
Theorem 4.10 gives the same representation of the optimal control as found in [AFKP18], where the autors also use the method of decoupling fields. However, they do not allow for a drift and diffusion term ($b = B = c = C = 0$), set $g(x) = Lx^2$ and have further smaller differences in the assumptions, which are sometimes more general on their side and sometimes on ours.

The argruments applied in Lemma 4.9 heavily rely on the linearity of $\mu$ and $\sigma$ in Assumption 4.8. This property however restricts the dynamics of the controlled processes. In the following we introduce another set of assumptions which does not need the linearity of $\mu$ but relies on other properties.

**Assumption 4.13**
*Let $\mu, \sigma : \Omega \times [0,T] \times \mathbb{R} \to \mathbb{R}$, $f : \Omega \times [0,T] \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ and $g : \Omega \times \mathbb{R} \to \mathbb{R}$ be twice differentiable in the space arguments and progressively measurable. Furthermore, assume*

1. *$\partial_x \mu$, $\partial_{xx}\mu$ are bounded,*
2. *there are bounded, progressively measurable processes $c, C$ such that $\sigma(s,x) = c + Cx$,*
3. *$\partial_x f$, $\frac{\det(\mathcal{H}(f))}{\partial_{\alpha\alpha} f}$ and $\frac{\partial_{x\alpha} f}{\partial_{\alpha\alpha} f}$ are bounded,*
4. *$\partial_{\alpha\alpha} f(s,x,f_\alpha^{-1}(s,x,y)) \geqslant \varepsilon > 0$ a.s., for all $(s,x,y) \in [0,T] \times \mathbb{R} \times [-\hat{y}, \hat{y}]$, where $\hat{y} := (\|g'\|_\infty + T\|\partial_x f\|_\infty) \exp(T\|\partial_x \mu\|_\infty)$,*
5. *$\|\mu(\cdot,\cdot,0)\|_\infty$, $\|(\partial_x f \circ \varphi)(\cdot,\cdot,0,0)\|_\infty$, $\|f_\alpha^{-1}(\cdot,\cdot,0,0)\|_\infty$ and $\|g(\cdot,0)\|_\infty$ are finite,*
6. *$g$ and $f$ are convex in the space arguments, $\lambda \times \mathbf{P}$ a.e.,*
7. *$g$ is monotone in the space argument with a bounded first and second derivative,*
8. *at least one of the following two cases is fulfilled*
   i) *$g' \geqslant 0$, $\mu$ is convex in $x$ and $\partial_x f \geqslant 0$,*
   ii) *$g' \leqslant 0$, $\mu$ is concave in $x$ and $\partial_x f \leqslant 0$.*

**Lemma 4.14**
*Let $\sigma$, $\mu$, $f$ and $g$ fulfill Assumption 4.13 and set $a = -\hat{y}$, $b = \hat{y}$, $c = -\infty$, $d = \infty$. Then FBSDE* (4.3) *fulfills SLC and the gradient process $U$ is bounded by*

$$0 \leqslant U \leqslant \left[\|g''\|_\infty + T\left(\|\partial_{xx}\mu\|_\infty \hat{y} + \left\|\frac{\det(H(f))}{\partial_{\alpha\alpha} f}\right\|_\infty\right)\right] \cdot \exp\left(T\left(\|\partial_x \sigma\|_\infty^2 + 2\|\partial_x \mu\|_\infty + 2\left\|\frac{\partial_{x\alpha} f}{\partial_{\alpha\alpha} f}\right\|_\infty\right)\right),$$

*$\lambda \times \mathbf{P}$ a.e. Moreover, $-\hat{y} \leqslant Y \leqslant \hat{y}$ and $\partial_{xx}\mu \cdot Y \geqslant 0$, $\lambda \times \mathbf{P}$ a.e., and $0 \leqslant Y \leqslant \hat{y}$, $\lambda \times \mathbf{P}$ a.e., if $g' \geqslant 0$ and $-\hat{y} \leqslant Y \leqslant 0$, $\lambda \times \mathbf{P}$ a.e., if $g' \leqslant 0$.*

*Proof.* By the assumptions made and the identities in Lemma 4.7, FBSDE (4.3) fulfills SLC. The first thing we have to prove is that $Y$ does not exceed the truncation bounds. To this end observe that we can rewrite the dynamics of $Y$ as

$$Y_s = g'(X_T) + \int_s^T \left[\partial_x \mu(r, X_r)\mathcal{T}_a^b(Y_r) + \partial_x f\left(r, X_r, f_\alpha^{-1}\left(r, X_r, \mathcal{T}_a^b(Y_r)\right)\right)\right] \mathrm{d}r - \int_s^T Z_r \,\mathrm{d}\widetilde{W}_r,$$

where $\widetilde{W}_s := W_s - \int_s^T \partial_x \sigma(r, X_r)\,\mathrm{d}r$ is a Brownian motion with respect to some measure $\widetilde{\mathbf{P}}$, due to Girsanov's theorem since $\partial_x \sigma(s, X_s) = C_s$ is bounded. Consider, for now, the process $\bar{Y}$ given by the BSDE

$$\hat{Y}_s = g'(X_T) + \int_s^T \left[\partial_x \mu(r, X_r)\hat{Y}_r + \partial_x f\left(r, X_r, f_\alpha^{-1}\left(r, X_r, \mathcal{T}_a^b(Y_r)\right)\right)\right] \mathrm{d}r - \int_s^T Z_r \,\mathrm{d}\widetilde{W}_r.$$

For any given process $X$ and all $s \in [0, T]$ for which $Y$ exists on $[s, T]$, the solution formula for linear BSDEs (see e.g. Proposition 6.2.1 in [Pha09]) states that

$$\hat{Y}_s = \tilde{\mathbb{E}}\left[ g'(X_T)\Gamma(s,T) + \int_s^T \partial_x f\left(r, X_r, f_\alpha^{-1}\left(r, X_r, \mathcal{T}_a^b(Y_r)\right)\right)\Gamma(s,r)\,\mathrm{d}r \middle| \widetilde{\mathcal{F}}_s \right],$$

where

$$\Gamma(t,r) := \exp\left( \int_t^r \partial_x \mu(u, X_u)\,\mathrm{d}u - \frac{1}{2}\int_t^r Z_u^2\,\mathrm{d}u \right)$$

for all $s \leqslant t \leqslant r \leqslant T$. Therefore, and since $\partial_x f$ is bounded, we conclude that $\hat{Y}$ is bounded by $\|\hat{Y}\|_\infty \leqslant (\|g'\|_\infty + T\,\|\partial_x f\|_\infty)\exp(T\|\partial_x\mu\|_\infty) = \hat{y}$. Thus, $\hat{Y}$ does not exceed bounds $a = -\hat{y}$, $b = \hat{y}$ of the truncation and hence coincides with $Y$. This means that the truncation of $Y$ can be omitted.

Next, observe that the BSDE

$$\bar{Y}_t = 0 + \int_t^T \partial_x\mu(s, X_s)\bar{Y}_s\,\mathrm{d}s - \int_t^T \bar{Z}_s\,\mathrm{d}\widetilde{W}_s$$

has the trivial solution $\bar{Y} \equiv 0$, $\bar{Z} \equiv 0$. Thus, we can use the comparison theorem and obtain in the case where $g' \geqslant 0$ and $\partial_x f \geqslant 0$ that $0 \leqslant Y \leqslant \hat{y}$ and in the case where $g' \leqslant 0$ and $\partial_x f \leqslant 0$ that $-\hat{y} \leqslant Y \leqslant 0$. Hence, by the assumptions made, either $Y \geqslant 0$ and $\partial_{xx}\mu \geqslant 0$ or $Y \leqslant 0$ and $\partial_{xx}\mu \leqslant 0$, yielding in any case that $\partial_{xx}\mu(s, X_s)Y_s \geqslant 0$.

It remains to show that the bounds for $U$ hold true. Have a look at the dynamics of $U$, they are

$$\begin{aligned} U_s = g''(X_T) + \int_s^T \Bigg[ &- U_r^2 \frac{1}{\partial_{\alpha\alpha} f(r, X_r, f_\alpha^{-1}(r, X_r, Y_r))} + U_r\left(\partial_x\sigma^2(r, X_r) + 2\partial_x\mu(r, X_r)\right) \\ &+ U_r\left( 2\frac{\partial_{x\alpha} f(r, X_r, f_\alpha^{-1}(r, X_r, Y_r))}{\partial_{\alpha\alpha} f(r, X_r, f_\alpha^{-1}(r, X_r, Y_r))} \right) \\ &+ \partial_{xx}\mu(r, X_r)Y_r + \partial_{xx}\sigma(r, X_r)Z_r + \partial_{xx} f(r, X_r, f_\alpha^{-1}(r, X_r, Y_r)) \\ &- \frac{\left(\partial_{x\alpha} f(r, X_r, f_\alpha^{-1}(r, X_r, Y_r))\right)^2}{\partial_{\alpha\alpha} f(r, X_r, (f_\alpha^{-1})(r, X_r, Y_r))} \Bigg]\mathrm{d}r \\ - \int_s^T \tilde{Z}_r\,\mathrm{d}\widetilde{W}_r, & \end{aligned}$$

where $\widetilde{W}_s = W_s - \int_s^T 2\partial_x\sigma(r, X_r)\,\mathrm{d}r$ is a Brownian motion with respect to $\widetilde{\mathbf{P}}$ by Girsanov's theorem. Because $U$ is bounded on every interval $[t, T] \subset I_{\max}$, we can interpret its dynamics as being Lipschitz allowing us to apply the Comparison Theorem. Note that $g'' \geqslant 0$, $\partial_{xx}\mu(s, X_s)Y_s \geqslant 0$, $\partial_{xx}\sigma = 0$ and $\partial_{xx} f - \frac{(\partial_{x\alpha} f)^2}{\partial_{\alpha\alpha} f} = \frac{\det(\mathcal{H}(f))}{\partial_{\alpha\alpha} f} \geqslant 0$. Hence, we get for all $s \in [t, T]$ by the Comparison Theorem that $U \geqslant \check{U}$, where

$$\begin{aligned} \check{U}_s = 0 + \int_s^T \Bigg[ &\check{U}_r\left( (\partial_x\sigma)^2(r, X_r) + 2\partial_x\mu(r, X_r) + 2\frac{\partial_{x\alpha} f(r, X_r, f_\alpha^{-1}(r, X_r, Y_r))}{\partial_{\alpha\alpha} f(r, X_r, f_\alpha^{-1}(r, X_r, Y_r))} \right) \\ &- \check{U}_r^2 \frac{1}{\partial_{\alpha\alpha} f(r, X_r, f_\alpha^{-1}(r, X_r, Y_r))} \Bigg]\mathrm{d}r - \int_s^T \tilde{Z}_r\,\mathrm{d}\widetilde{W}_r, \end{aligned}$$

which has the trivial solution $\check{U} \equiv 0$, $\tilde{Z} \equiv 0$. Thus $U \geqslant 0$.

For the upper bound we apply the Comparison Theorem again. By dropping the quadratic term, obtaining

$$\begin{aligned}\hat{U}_s = g''(X_T) + \int_s^T \Bigg[ &\hat{U}_r \left( \partial_x \sigma^2(r, X_r) + 2\partial_x \mu(r, X_r) + 2\frac{\partial_{x\alpha} f(r, X_r, f_\alpha^{-1}(r, X_r, Y_r))}{\partial_{\alpha\alpha} f(r, X_r, f_\alpha^{-1}(r, X_r, Y_r))} \right) \\ &+ \partial_{xx}\mu(r, X_r) Y_r + \partial_{xx} f(r, X_r, f_\alpha^{-1}(r, X_r, Y_r)) \\ &- \frac{\left(\partial_{x\alpha} f(r, X_r, f_\alpha^{-1}(r, X_r, Y_r))\right)^2}{\partial_{\alpha\alpha} f(r, X_r, (f_\alpha^{-1})(r, X_r, Y_r))} \Bigg] \mathrm{d}r - \int_s^T \tilde{Z}_r \, \mathrm{d}\widetilde{W}_r,\end{aligned}$$

we get that $U$ is bounded from above by the solution $\hat{U}$ of this linear BSDE, which again is bounded by

$$\begin{aligned}\hat{U}_s \leqslant \Bigg[ \|g''\|_\infty + T \left( \|\partial_{xx}\mu\|_\infty \hat{y} + \left\| \frac{\det(H(f))}{\partial_{\alpha\alpha} f} \right\|_\infty \right) \Bigg] \\ \cdot \exp\left( T \left( \|\partial_x \sigma\|_\infty^2 + 2\|\partial_x \mu\|_\infty + 2\left\| \frac{\partial_{x\alpha} f}{\partial_{\alpha\alpha} f} \right\|_\infty \right) \right).\end{aligned}$$

■

The following theorem is the second main result of this section, stating a solution to a control problem with non-linear dynamics.

**Theorem 4.15**
*Let $\sigma$, $\mu$, $f$ and $g$ fulfill Assumption 4.13. Then the control problem of minimizing* (4.1) *has the optimal control*

$$\hat{\alpha}_t := f_\alpha^{-1}(t, X_t, Y_t),$$

*where $X$, $Y$ solve FBSDE* (4.3)*, which has an a.e. unique solution on the whole interval $[0, T]$. Furthermore, $0 \leqslant Y \leqslant \hat{y}$, $\lambda \times \mathbf{P}$ a.e., if $g' \geqslant 0$ and $-\hat{y} \leqslant Y \leqslant 0$, $\lambda \times \mathbf{P}$ a.e., if $g' \leqslant 0$.*

*Proof.* With the statement of Lemma 4.14 and since $[0, T] \times \mathbb{R} \times [-\hat{y}, \hat{y}] \subset \mathcal{D}(f_\alpha^{-1})$ is implicitly given by Assumption 4.13, for the assumptions of Proposition 4.4 it only remains to show that $(x, a) \mapsto H(t, x, a, Y_t, Z_t)$ is convex for all $t \in [0, T]$. To this end we define for $t \in [0, T]$ the functions

$$H^\mu(t, x, a) := \mu(t, x) \cdot Y_t, \quad H^\sigma(t, x, a) := \sigma(t, x) \cdot Z_t \quad \text{and} \quad H^f(t, x, a) := f(t, x, a)$$

such that $H(t, x, a, Y_t, Z_t) = \left(H^\mu + H^\sigma + H^f\right)(t, x, a)$. Observe that by Lemma 4.14 we get that

$$\partial_{xx} H^\mu(t, x, a) = \partial_{xx}\mu(t, x) \cdot Y_t \geqslant 0$$

and furthermore, that

$$\partial_{xx} H^\sigma(t, x, a) = \partial_{xx}\sigma(t, x) \cdot Z_t = 0.$$

Since furthermore $H^\mu$ and $H^\sigma$ are independent of the argument $a$, they are convex in $(x, a)$. Therefore and because $H^f$ is convex by assumption, we obtain that $H$ is the sum of convex functions and hence convex in $(x, a)$ itself. Thus, Proposition 4.4 can be applied and we obtain that $\hat{\alpha}$ is an admissible optimal control. The remaining statement about $Y$ is given by Lemma 4.14. ■

Unlike Theorem 4.10 the method for Theorem 4.15 relies on the boundedness of $Y$. This however comes with the price of other restrictions like $g$ and $f$ having a bounded derivative with respect to $x$.

**Example 4.16**
Let $b, c, C, r$ be bounded progressively measurable processes and

$$\mu(s,x) := \sqrt{1 + b_s + x^2}, \qquad \sigma(s,x) := c_s + C_s \cdot x \qquad f(s,x,a) := \exp(-r_s \cdot s) \cdot \cosh(a),$$
$$g(x) := \begin{cases} 0, & x < 0, \\ \frac{x^3}{4}\left(1 - \frac{x}{4}\right), & 0 \leqslant x < 2, \\ x - 1, & x \geqslant 2, \end{cases}$$

for $(s,x) \in [0,T] \times \mathbb{R}$. Then Theorem 4.15 states that the control problem consisting of minimizing the cost fuctional $J$ from Equation (4.1) over all admissible controls has the optimal solution

$$\hat{\alpha}_s = \sinh^{-1}\left(Y_s \cdot \exp\left(r_s \cdot s\right)\right),$$

where $Y$ is part of the unique solution of FBSDE (4.3).

## 4.2 Linear-quadratic control problems

In this section we turn our attention to a special case of Assumption 4.8, the linear-quadratic case. We make the following assumption.

**Assumption 4.17**
*Let $T > 0$ and $\mu, \sigma : \Omega \times [0,T] \times \mathbb{R} \to \mathbb{R}$, $f : \Omega \times [0,T] \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}$, $g : \mathbb{R} \to \mathbb{R}$ be of the form*

$$\mu(t,x) = b_t + B_t x, \qquad \sigma(t,x) = c_t + C_t x, \qquad g(x) = G_0 + G_1 x + G_2 x^2,$$
$$f(t,x,a) = \beta_{xx}(t)x^2 + \beta_x(t)x + \beta_{xa}(t)ax + \beta_{aa}(t)a^2 + \beta_a(t)a + \beta_0(t)$$

*for $b, B, c, C, \beta_{xx}, \beta_x, \beta_{xa}, \beta_{aa}, \beta_a, \beta_0$ being essentially bounded processes on $[0,T]$, such that*

- $\det(\mathcal{H}(f))(t,\cdot,\cdot) = \beta_{aa}(t)\beta_{xx}(t) - \beta_{ax}^2(t) \geqslant \varepsilon_1 > 0$ *for $t \in [0,T]$ and some constant $\varepsilon_1 > 0$*

- $\beta_{aa}(t) \geqslant \varepsilon_2 > 0$ *for $t \in [0,T]$ and some constant $\varepsilon_2 > 0$*

*and $G_0, G_1 \in \mathbb{R}$, $G_2 \geqslant 0$.*

Note that hence $f_\alpha^{-1}(t,x,y) = \frac{y-\beta_a(t)-\beta_{xa}(t)x}{2\beta_{aa}(t)}$. Furthermore, for the processes in Assumption 4.17 we denote by $\check{\cdot}$ the essential infimum over time and by $\hat{\cdot}$ the essential supremum over time, e.g. $\check{\beta}_{aa} := \operatorname{ess\,inf}_{s\in[0,T]} \beta_{aa}(s)$ and $\hat{\beta}_{aa} := \operatorname{ess\,sup}_{s\in[0,T]} \beta_{aa}(s)$.

**Corollary 4.18**
*Let $\sigma$, $\mu$, $f$ and $g$ fulfill Assumption 4.17. Then, for the truncation parameters $a, c = -\infty$, $b, d = \infty$, FBSDE* (4.3) *has a solution on the whole time interval $[0,T]$ and is equivalent to*

$$\begin{aligned} X_t &= x_0 + \int_0^t \left[\mu(s,X_s) - f_\alpha^{-1}(s,X_s,Y_s)\right] \mathrm{d}s + \int_0^t \sigma(s,X_s)\,\mathrm{d}W_s \\ Y_t &= \partial_x g(X_T) - \int_t^T Z_s\,\mathrm{d}W_s \\ &\quad + \int_t^T \left[\partial_x\mu(s,X_s)Y_s + \partial_x\sigma(s,X_s)Z_s + \partial_x f\left(s,X_s,f_\alpha^{-1}(s,X_s,Y_s)\right)\right] \mathrm{d}s. \end{aligned} \tag{4.8}$$

*Proof.* Observe that Assumption 4.17 is a special case of Assumption 4.8. Hence Lemma 4.9 is applicable and yields that FBSDE (4.3) fulfills SLC for $a, c = -\infty$, $b, d = \infty$ and has a solution on the whole time interval $[0,T]$. ∎

**Corollary 4.19**
*Let $\sigma$, $\mu$, $f$ and $g$ fulfill Assumption 4.17. Then the gradient process $U^T$ of FBSDE* (4.8) *is indistinguishable from the process that solves the quadratic BSDE*

$$\begin{aligned} U_t^T = 2G_2 + \int_t^T \left[-\frac{\left(U_r^T\right)^2}{2\beta_{aa}(r)} + U_r^T\left(C_r^2 + 2B_r + \frac{\beta_{xa}(r)}{\beta_{aa}(r)}\right) + 2\beta_{xx}(r) - \frac{\beta_{xa}^2(r)}{2\beta_{aa}(r)} + 2C_r Z_r^U\right] \mathrm{d}r \\ - \int_t^T Z_r^U\,\mathrm{d}W_r \end{aligned}$$

*for $t \in [0,T]$.*

*Proof.* Remember that Corollary 4.18 states that there is a solution to (4.8) on the whole time interval. Furthermore, by Lemma 4.7 we get that $U^T$ has the dynamics

$$\begin{aligned} U_t^T &= g''(X_T) + \int_s^T \Bigg[ -\left(U_r^T\right)^2 \frac{1}{\partial_{\alpha\alpha} f\left(\varphi\left(r,X_r,Y_r\right)\right)} + U_r^T\left(\partial_x\sigma^2(r,X_r)\right) \\ &\qquad + U_r^T\left(2\left(\frac{\partial_{x\alpha} f}{\partial_{\alpha\alpha} f}\left(\varphi\left(r,X_r,Y_r\right)\right) + \partial_x\mu(r,X_r)\right)\right) \\ &\qquad + \partial_{xx}\mu(r,X_r)Y_r + \partial_{xx}\sigma(r,X_r)Z_r + \partial_{xx} f\left(\varphi\left(r,X_r,Y_r\right)\right) \\ &\qquad - \frac{\left(\partial_{x\alpha} f\right)^2}{\partial_{\alpha\alpha} f}\left(\varphi\left(r,X_r,Y_r\right)\right) + 2\partial_x\sigma(r,X_r)Z_r^U \Bigg] \mathrm{d}r - \int_s^T Z_r^U\,\mathrm{d}W_r \\ &= 2G_2 + \int_t^T \left[-\frac{\left(U_r^T\right)^2}{2\beta_{aa}(r)} + U_r^T\left(C_r^2 + 2B_r + \frac{\beta_{xa}(r)}{\beta_{aa}(r)}\right) + 2\beta_{xx}(r) - \frac{\beta_{xa}^2(r)}{2\beta_{aa}(r)} + 2C_r Z_r^U\right] \mathrm{d}r \\ &\quad - \int_t^T Z_r^U\,\mathrm{d}W_r. \end{aligned}$$

∎

**Remark 4.20**
Note that the BSDE for $U^T$ in Corollary 4.19 is for $\beta_x = \beta_a = b = c = 0$ a special case of Equation (6.1) in [SXY18]. Thus, in some sense, we have an alternative proof for the existence of a solution. We obtain this as a special case of the BSDE in Lemma 4.7, for which we proved the existence under the more general Assumption 4.8.

Using this representation of $U^T$ and in particular that it does not depend on $X$, we can derive a representation of the decoupling field. This then allows us to give a formula for the optimal control.

**Proposition 4.21**
*Let $\sigma$, $\mu$, $f$ and $g$ fulfill Assumption 4.17. Then the decoupling field $u$ of FBSDE* (4.8) *is equal to*

$$u(t,x) = \varphi_t^T + U_t^T \cdot x,$$

*where $U^T$ is the gradient process given in Corollary 4.19 and $\varphi^T$ is indistinguishable from the solution of the BSDE*

$$\begin{aligned}\varphi_t^T = G_1 - \int_t^T Z_s^\varphi \,\mathrm{d}W_s + \int_t^T \left(C_s Z_s^\varphi + c_s Z_s^U\right) \mathrm{d}s \\ + \int_t^T \left[\varphi_s^T \left(B_s + \frac{\beta_{xa}(s) - U_s^T}{2\beta_{aa}(s)}\right) + C_s c_s U_s^T + b_s U_s^T + \beta_x(s) + \beta_a(s) \frac{U_s^T - \beta_{xa}(s)}{2\beta_{aa}(s)}\right] \mathrm{d}s,\end{aligned}$$

*where $Z^U$ is the diffusion part of $U^T$ (see Corollary 4.19). In particular, both $U^T$ and $\varphi^T$, and hence also the decoupling field $u$, do not depend on the process $X$ from FBSDE* (4.8)*. The component $Z^\varphi$ is given by $Z_s^\varphi = Z_s - X_s Z_s^U - (c_s + C_s X_s) U_s^T$ for all $s \in [0,T]$, where $X$ and $Z$ are from FBSDE* (4.8).

*Proof.* Note that Corollary 4.19 already states that for every $t \in [0,T]$ we have $\partial_x u(t, X_t) = U_t^T$, which is independent of $X_t$, since $U^T$ does not depend on the starting value of $X$. Hence, $u$ has to be of the form $u(t,x) = \varphi_t^T + U_t^T \cdot x$ for some process $\varphi^T$, which does not depend on $X$ neither. Using Itô's formula, the dynamics of $Y$, $X$, $U$ and the decoupling condition, a straightforward calculation yields

$$\begin{aligned}
\varphi_t^T &= Y_t - U_t^T X_t \\
&= G_1 + 2G_2 X_T - 2G_2 X_T \\
&\quad + \int_t^T \left[ B_s Y_s + C_s Z_s + 2\beta_{xx}(s) X_s + \beta_x(s) + \beta_{xa}(s) \frac{Y_s - \beta_a(s) - \beta_{xa}(s) X_s}{2\beta_{aa}(s)} \right] \mathrm{d}s - \int_t^T Z_s \,\mathrm{d}W_s \\
&\quad - \int_t^T \left[ -\frac{\left(U_s^T\right)^2}{2\beta_{aa}(s)} + U_s^T \left( C_s^2 + 2B_s + \frac{\beta_{xa}(s)}{\beta_{aa}(s)} \right) + 2\beta_{xx}(s) - \frac{\beta_{xa}^2(s)}{2\beta_{aa}(s)} + 2C_s Z_s^U \right] X_s \,\mathrm{d}s \\
&\quad - \int_t^T -Z_s^U X_s \,\mathrm{d}W_s - \int_t^T -Z_s^U \left( c_s + C_s X_s \right) \mathrm{d}s \\
&\quad - \int_t^T -U_s^T \left[ b_s + B_s X_s - \frac{Y_s - \beta_a(s) - \beta_{xa}(s) X_s}{2\beta_{aa}(s)} \right] \mathrm{d}s - \int_t^T -U_s^T (c_s + C_s X_s) \,\mathrm{d}W_s \\
&= G_1 - \int_t^T \left[ Z_s - X_s Z_s^U - \left( c_s + C_s X_s \right) U_s^T \right] \mathrm{d}W_s \\
&\quad + \int_t^T \left[ Z_s - X_s Z_s^U - \left( c_s + C_s X_s \right) U_s^T \right] C_s \,\mathrm{d}s + \int_t^T C_s U_t^T \left( c_s + C_s X_s \right) \mathrm{d}s \\
&\quad + \int_t^T \left[ B_s Y_s + 2\beta_{xx}(s) X_s + \beta_x(s) + \beta_{xa}(s) \frac{\varphi_s^T + U_s^T X_s - \beta_a(s) - \beta_{xa}(s) X_s}{2\beta_{aa}(s)} \right] \mathrm{d}s \\
&\quad + \int_t^T \left[ \frac{\left(U_s^T\right)^2}{2\beta_{aa}(s)} - U_s^T \left( C_s^2 + 2B_s + \frac{\beta_{xa}(s)}{\beta_{aa}(s)} \right) - 2\beta_{xx}(s) + \frac{\beta_{xa}^2(s)}{2\beta_{aa}(s)} \right] X_s \,\mathrm{d}s \\
&\quad + \int_t^T Z_s^U c_s \,\mathrm{d}s + \int_t^T U_s^T \left[ b_s + B_s X_s - \frac{\varphi_s^T + U_s^T X_s - \beta_a(s) - \beta_{xa}(s) X_s}{2\beta_{aa}(s)} \right] \mathrm{d}s
\end{aligned}$$

and further that

$$\begin{aligned}
\varphi_t^T &= G_1 - \int_t^T \left[ Z_s - X_s Z_s^U - \left( c_s + C_s X_s \right) U_s^T \right] \mathrm{d}W_s \\
&\quad + \int_t^T \left[ Z_s - X_s Z_s^U - \left( c_s + C_s X_s \right) U_s^T \right] C_s + c_s Z_s^U \,\mathrm{d}s + \int_t^T X_s \cdot 0 \,\mathrm{d}s \\
&\quad + \int_t^T \left[ \varphi_s^T \left( B_s + \frac{\beta_{xa}(s) - U_s^T}{2\beta_{aa}(s)} \right) + C_s c_s U_s^T + b_s U_s^T + \beta_x(s) + \beta_a(s) \frac{U_s^T - \beta_{xa}(s)}{2\beta_{aa}(s)} \right] \mathrm{d}s \\
&= G_1 - \int_t^T Z_s^\varphi \,\mathrm{d}W_s + \int_t^T C_s Z_s^\varphi + c_s Z_s^U \,\mathrm{d}s \\
&\quad + \int_t^T \left[ \varphi_s^T \left( B_s + \frac{\beta_{xa}(s) - U_s^T}{2\beta_{aa}(s)} \right) + C_s c_s U_s^T + b_s U_s^T + \beta_x(s) + \beta_a(s) \frac{U_s^T - \beta_{xa}(s)}{2\beta_{aa}(s)} \right] \mathrm{d}s,
\end{aligned}$$

where $Z_s^\varphi := Z_s - X_s Z_s^U - \left( c_s + C_s X_s \right) U_s^T$ for all $s \in [0, T]$. ■

**Corollary 4.22**
*Let $\sigma$, $\mu$, $f$ and $g$ fulfill Assumption 4.17. Then the optimal control is*

$$\alpha_t^T = \frac{\varphi_t^T - \beta_a(t) + (U_t^T - \beta_{xa}(t)) X_t}{2\beta_{aa}(t)} = \alpha_t^T(X_t)$$

*for $t \in [0, T]$, which is a linear feedback control with the random function $\alpha^T : \Omega \times [0, T] \times \mathbb{R} \to \mathbb{R}$*

$$\alpha_t^T(x) := \frac{\varphi_t^T - \beta_a(t) + (U_t^T - \beta_{xa}(t)) x}{2\beta_{aa}(t)}. \tag{4.9}$$

*Proof.* Remember that FBSDE (4.8) has a solution on the whole interval $[0,T]$ and its decoupling field is explicitly given in Proposition 4.21. Thus, as a particular case of Theorem 4.10, the optimal control fulfills

$$\alpha_t^T = f_\alpha^{-1}(t, X_t, Y_t) = \frac{Y_t - \beta_a(t) - \beta_{xa}(t)X_t}{2\beta_{aa}(t)} = \frac{\varphi_t^T - \beta_a(t) + (U_t^T - \beta_{xa}(t))X_t}{2\beta_{aa}(t)}.$$

Since everything on the right hand side, aside from $X$ itself, does not depend on $X$, as stated in Proposition 4.21, we can view this as a linear random function of $X_t$. Hence, we can view the optimal control as a linear feedback control. ■

**Remark 4.23**
In the case of $\beta_x(s) = \beta_a(s) = b_s = c_s = 0$ for all $s \in [0,T]$ the result in Corollary 4.22 follows by Theorem 6.7 in [SXY18]. Likewise, for all coefficents being deterministic and $\beta_x(s) = \beta_a(s) = 0$ for all $s \in [0,T]$ this result follows by Theorem 6.1 in [YZ99]. Furthermore, note that in this latter case the process $P$ in [YZ99] equals $\frac{U}{2}$ from our work, while their $\varphi$ equals $\frac{\varphi}{2}$ here.

With those representations of the components of the decoupling field and the optimal control, we derive a formula for the value function of the linear-quadratic control problem, which does not depend on the solution of FBSDE (4.8). This is our main result of this section.

**Theorem 4.24**
*Let Assumption 4.17 be fulfilled. Then for all $t \in [0,T]$ and $x \in \mathbb{R}$*

$$\begin{aligned} V(t,T,x) :=& \mathbb{E}\left[\frac{1}{2}U_t^T x^2 + \varphi_t^T x + G_0 + \int_t^T \left(\varphi_s^T b_s + U_s^T \frac{c_s^2}{2} + \beta_0(s) - \frac{\left(\varphi_s^T - \beta_a(s)\right)^2}{4\beta_{aa}(s)} + c_s Z_s^\varphi\right) \mathrm{d}s\right] \\ =& \mathbb{E}\left[\int_t^T f\left(s, X_s^{t,x}, \alpha_s^T(X_s^{t,x})\right) \mathrm{d}s + g(X_T^{t,x})\right], \end{aligned}$$

*where $X^{t,x}$ is the process $X$ conditioned on $X_t = x$ and $(\alpha_s^T)_{s\in[0,T]}$ is the optimal feedback control from Equation* (4.9). *In particular,*

$$V(0,T,x) = \inf_{\alpha\in\mathcal{A}} J(T,x,\alpha).$$

*Proof.* Recall that the dynamics of $U^T$ and $\varphi^T$ are given in Corollary 4.19 and Proposition 4.21. Hence, by Itô's formula,

$$\begin{aligned} &\frac{1}{2}U_t^T X_t^2 \\ &= \frac{1}{2}U_T^T X_T^2 \\ &\quad - \int_t^T U_s^T X_s \left(b_s + B_s X_s - \frac{\varphi_s^T - \beta_a(s)}{2\beta_{aa}(s)} - \frac{U_s^T - \beta_{xa}(s)}{2\beta_{aa}(s)} X_s\right) \mathrm{d}s - \int_t^T U_s^T X_s \left(c_s + C_s X_s\right) \mathrm{d}W_s \\ &\quad + \int_t^T \frac{X_s^2}{2}\left[-\frac{\left(U_s^T\right)^2}{2\beta_{aa}(s)} + U_s^T\left(C_s^2 + 2B_s + \frac{\beta_{xa}(s)}{\beta_{aa}(s)}\right) + 2\beta_{xx}(s) - \frac{\beta_{xa}^2(s)}{2\beta_{aa}(s)} + 2C_s Z_s^U\right] \mathrm{d}s \\ &\quad - \int_t^T \frac{X_s^2}{2} Z_s^U \,\mathrm{d}W_s - \int_t^T \frac{1}{2}U_s^T \left(c_s + C_s X_s\right)^2 \mathrm{d}s - \int_t^T X_s Z_s^U \left(c_s + C_s X_s\right) \mathrm{d}s \end{aligned}$$

and

$$
\begin{aligned}
&\varphi_t^T X_t \\
&= \varphi_T^T X_T \\
&\quad - \int_t^T \varphi_s^T \left( b_s + B_s X_s - \frac{\varphi_s^T - \beta_a(s)}{2\beta_{aa}(s)} - \frac{U_s^T - \beta_{xa}(s)}{2\beta_{aa}(s)} X_s \right) \mathrm{d}s - \int_t^T \varphi_s^T \left( c_s + C_s X_s \right) \mathrm{d}W_s \\
&\quad + \int_t^T X_s \left[ \varphi_s^T \left( B_s + \frac{\beta_{xa}(s) - U_s^T}{2\beta_{aa}(s)} \right) + C_s c_s U_s^T + b_s U_s^T + \beta_x(s) + \beta_a(s) \frac{U_s^T - \beta_{xa}(s)}{2\beta_{aa}(s)} \right] \mathrm{d}s \\
&\quad + \int_t^T X_s \left[ c_s Z_s^U + C_s Z_s^\varphi \right] \mathrm{d}s - \int_t^T X_s Z_s^\varphi \, \mathrm{d}W_s - \int_t^T Z_s^\varphi \left( c_s + C_s X_s \right) \mathrm{d}s.
\end{aligned}
$$

Therefore, it is straightforward to verify that

$$
\begin{aligned}
&\frac{1}{2} U_t^T X_t^2 + \varphi_t^T X_t + G_0 + \int_t^T \left[ \varphi_s^T b_s + U_s^T \frac{c_s^2}{2} + \beta_0(s) - \frac{\left(\varphi_s^T - \beta_a(s)\right)^2}{4\beta_{aa}(s)} + c_s Z_s^\varphi \right] \mathrm{d}s \\
&= G_2 X_T^2 + G_1 X_T + G_0 \\
&\quad + \int_t^T \Bigg\{ X_s^2 \Bigg[ \frac{1}{2} \left( -\frac{\left(U_s^T\right)^2}{2\beta_{aa}(s)} + U_s^T \left( 2B_s + \frac{\beta_{xa}(s)}{\beta_{aa}(s)} + C_s^2 \right) + 2\beta_{xx}(s) - \frac{\beta_{xa}^2(s)}{2\beta_{aa}(s)} \right) \\
&\qquad\qquad - U_s^T \left( B_s - \frac{U_s^T - \beta_{xa}(s)}{2\beta_{aa}(s)} + \frac{C_s^2}{2} \right) \Bigg] \\
&\qquad + X_s \Bigg[ \left( \left( U_s^T \left( b_s + c_s C_s + \frac{\beta_a(s)}{2\beta_{aa}(s)} \right) - \frac{\beta_a(s)\beta_{xa}(s)}{2\beta_{aa}(s)} + \beta_x(s) \right) + \left( B_s + \frac{\beta_{xa}(s) - U_s^T}{2\beta_{aa}(s)} \right) \varphi_s^T \right. \\
&\qquad\qquad \left. - \varphi_s^T \left( B_s - \frac{U_s^T - \beta_{xa}(s)}{2\beta_{aa}(s)} \right) - U_s^T \left( b_s - \frac{\varphi_s^T - \beta_a(s)}{2\beta_{aa}(s)} + c_s C_s \right) \right] \\
&\qquad + \left[ \beta_0(s) + \beta_a(s) \left( \frac{\varphi_s^T - \beta_a(s)}{2\beta_{aa}(s)} \right) + \beta_{aa}(s) \left( \frac{\varphi_s^T - \beta_a(s)}{2\beta_{aa}(s)} \right)^2 \right] \Bigg\} \mathrm{d}s \\
&\quad - \int_t^T \left[ \left( \varphi_s^T + U_s^T X_s \right) \left( c_s + C_s X_s \right) + \frac{1}{2} X_s^2 Z_s^U + X_s Z_s^\varphi \right] \mathrm{d}W_s \\
&= G_2 X_T^2 + G_1 X_T + G_0 \\
&\quad + \int_t^T \Bigg[ \beta_0(s) + \beta_{xx}(s) X_s^2 + \beta_x(s) X_s + \beta_{xa}(s) X_s \left( \frac{\varphi_s^T - \beta_a(s) + \left(U_s^T - \beta_{xa}(s)\right) X_s}{2\beta_{aa}(s)} \right) \\
&\qquad + \beta_{aa}(s) \left( \left( \frac{\varphi_s^T - \beta_a(s)}{2\beta_{aa}(s)} \right)^2 + 2 \frac{\left(\varphi_s^T - \beta_a(s)\right)\left(U_s^T - \beta_{xa}(s)\right) X_s}{\left(2\beta_{aa}(s)\right)^2} + \left( \frac{\left(U_s^T - \beta_{xa}(s)\right) X_s}{2\beta_{aa}(s)} \right)^2 \right) \\
&\qquad + \beta_a(s) \frac{\varphi_s^T - \beta_a(s) + \left(U_s^T - \beta_{xa}(s)\right) X_s}{2\beta_{aa}(s)} \Bigg] \mathrm{d}s \\
&\quad - \int_t^T \left[ \left( \varphi_s^T + U_s^T X_s \right) \left( c_s + C_s X_s \right) + \frac{1}{2} X_s^2 Z_s^U + X_s Z_s^\varphi \right] \mathrm{d}W_s \\
&= g(X_T) + \int_t^T f\left(s, X_s, \alpha_s^T\left(X_s\right)\right) \mathrm{d}s - \int_t^T \left[ \left( \varphi_s^T + U_s^T X_s \right) \left( c_s + C_s X_s \right) + \frac{1}{2} X_s^2 Z_s^U + X_s Z_s^\varphi \right] \mathrm{d}W_s.
\end{aligned}
$$

This yields that

$$V(t,T,x) = \mathbb{E}\left[\int_t^T f\left(s, X_s^{t,x}, \alpha_s^T\left(X_s^{t,x}\right)\right) \mathrm{d}s + g(X_T^{t,x})\right].$$

Since $\alpha^T$ is the optimal control we furthermore obtain

$$V(0,T,x) = \mathbb{E}\left[\int_0^T f\left(s, X_s^{0,x}, \alpha_s^T\left(X_s^{0,x}\right)\right) \mathrm{d}s + g(X_T^{t,0})\right] = \inf_{\alpha \in \mathcal{A}} J(T,x,\alpha).$$

■

**Remark 4.25**
Using the essential bounds for the parameter processes $b, B, c, C, \beta_{xx}, \beta_x, \beta_{xa}, \beta_{aa}, \beta_a, \beta_0$ and the results from Section 4.3, we are able to derive bounds for $U$ which do not depend on the time horizon $T$, similar to the estimate in Proposition 4.42 (i).

However, corresponding bounds for $\varphi^T$ would require estimates similar to the ones in Section 4.3 for quadratic BSDEs of the type $U^T$ solves in Section 4.2. Obtaining such estimates and thereby generalizing Section 4.4 to random parameter functions is left for future research.

## 4.3 Some results on functions with quadratic dynamics

In this section we investigate the properties of the solution of a deterministic quadratic ODE. First, we derive them for piecewise constant parameter functions and then generalize to right-continuous functions. The resulting estimates are used in Section 4.4 to show the convergence of the decoupling field in an ergodic control problem.

**Assumption 4.26**
*Let $p, q, a : [0,\infty) \to \mathbb{R}$ be deterministic right-continuous functions such that for all $s \in [0,\infty)$*

$$-\infty < \check{p} \leqslant p_s \leqslant \hat{p} < \infty, \qquad 0 < \check{q} \leqslant q_s \leqslant \hat{q} < \infty, \qquad 0 < \check{a} \leqslant a_s \leqslant \hat{a} < \infty$$

*for constants $\check{p}, \hat{p}, \check{q}, \hat{q}, \check{a}, \hat{a} \in \mathbb{R}$.*

We define the constants $\check{Y} := \check{p} + \sqrt{\check{p}^2 + \check{q}}$ and $\hat{Y} := \hat{p} + \sqrt{\hat{p}^2 + \hat{q}}$.

**Lemma 4.27**
*Let Assumption 4.26 be fulfilled and $t \geqslant 0$. Then the integral equation*

$$Y_s = Y_t + \int_t^s -a_r\left(Y_r^2 - 2p_r Y_r - q_r\right) \mathrm{d}r \tag{4.10}$$

*for $s \in [t,\infty)$ with starting value $Y_t \in [0,\hat{Y}]$ has a unique solution. Also, the solution $(Y_s)_{s \geqslant t}$ is bounded by*

$$\min\left\{Y_t, \check{Y}\right\} \leqslant Y_s \leqslant \hat{Y}$$

*for all $s \in [t,\infty)$.*

*Proof.* We define the auxiliary process $\tilde{Y}$ as the solution of the Lipschitz ODE

$$\partial_t \tilde{Y}_s = -a_s\left(\left(\mathcal{T}_0^{\hat{Y}}(\tilde{Y}_s)\right)^2 - 2p_s\tilde{Y}_s - q_s\right), \qquad \tilde{Y}_t = Y_t,$$

where $\mathcal{T}$ is the truncation operator defined in Section 4.1. Observe that for $\tilde{Y}_s \in [0, \check{Y})$ we have $-a_s\big(\big(\mathcal{T}_0^{\hat{Y}}(\tilde{Y}_s)\big)^2 - 2p_s\tilde{Y}_s - q_s\big) > 0$ and for $\tilde{Y}_s \in [\check{Y}, \infty)$ that $-a_s\big(\big(\mathcal{T}_0^{\hat{Y}}(\tilde{Y}_s)\big)^2 - 2p_s\tilde{Y}_s - q_s\big) \leqslant 0$. Hence, for $\tilde{Y}_t < \check{Y}$ we have that $\tilde{Y}_s \geqslant \tilde{Y}_t$ for all $s \in [t, \infty)$, since $Y$ is continuous. By the same argument we also obtain for $\tilde{Y}_t \geqslant \check{Y}$ that $\tilde{Y}_s$ cannot reach any value below $\check{Y}$ and likewise because $\tilde{Y}_t \leqslant \hat{Y}$ that $\tilde{Y}_s \leqslant \hat{Y}$. Thus, the truncation of the quadratic term has no consequence and can be omitted without changing the solution. Hence, the bounds are also valid for $Y$ and the solution of (4.10) is also unique. ■

In the following we denote by $Y$ the solution of Equation (4.10).

**Remark 4.28**
In the proofs of this section we make use of the following hyperbolic identities without explicitely mentioning it:

- $\tanh^{-1}(x) = \frac{1}{2}\ln\left(\frac{1+x}{1-x}\right)$ for $x \in (-1, 1)$,
- $\coth^{-1}(x) = \frac{1}{2}\ln\left(\frac{x+1}{x-1}\right)$ for $|x| > 1$,
- $\cosh(\tanh^{-1}(x)) = (1 - x^2)^{-1/2}$ for $x \in (-1, 1)$,
- $\sinh(\coth^{-1}(x)) = (1 - x^2)^{-1/2}$ for $x > 1$.

**Lemma 4.29**
*Let Assumption 4.26 be fulfilled and $t \geqslant 0$. Furthermore, assume that $Y_t \in [0, \hat{Y}]$ and for some $s \in (t, \infty)$ that the functions $p, q, a$ are constant on the interval $[t, s)$, i.e. there are $\bar{p}, \bar{q}, \bar{a} \in \mathbb{R}$ such that $p_r = \bar{p}$, $q_r = \bar{q}$ and $a_r = \bar{a}$ for all $r \in [t, s)$. Then*

$$Y_r = \begin{cases} \bar{p} + \sqrt{\bar{p}^2 + \bar{q}}\tanh\left(\bar{a}\sqrt{\bar{p}^2 + \bar{q}}(r - t) + \tanh^{-1}\left(\frac{Y_t - \bar{p}}{\sqrt{\bar{p}^2 + \bar{q}}}\right)\right), & Y_t \in \big[0, \bar{p} + \sqrt{\bar{p}^2 + \bar{q}}\big) \\ \bar{p} + \sqrt{\bar{p}^2 + \bar{q}}, & Y_t = \bar{p} + \sqrt{\bar{p}^2 + \bar{q}} \\ \bar{p} + \sqrt{\bar{p}^2 + \bar{q}}\coth\left(\bar{a}\sqrt{\bar{p}^2 + \bar{q}}(r - t) + \coth^{-1}\left(\frac{Y_t - \bar{p}}{\sqrt{\bar{p}^2 + \bar{q}}}\right)\right), & Y_t \in \big(\bar{p} + \sqrt{\bar{p}^2 + \bar{q}}, \infty\big) \end{cases} \tag{4.11}$$

*for all $r \in [t, s]$. In particular, $Y$ is monotone on the interval $[t, s]$.*

*Proof.* Observe that the dynamics of $Y$ state that it solves for $r \in [t, s)$ the separable ODE

$$Y_r' = -a_t\left((Y_r - p_t)^2 - p_t^2 - q_t\right).$$

The three cases follow by straightforward calculations. Also, Lemma 4.27 provides uniqueness. The remaining monotonicity follows from the monotonicity of $\tanh$ and $\coth$. ■

**Lemma 4.30**
*Let Assumption 4.26 be fulfilled and $[t_1, t_2] \subset [0, \infty)$ with $t_1 < t_2$. Furthermore, assume that $Y_{t_1} \in [0, \hat{Y}]$ and that the functions $p, q, a$ are constant on the interval $[t_1, t_2)$, i.e. there are $\bar{p}, \bar{q}, \bar{a} \in \mathbb{R}$ such that $p_r = \bar{p}$, $q_r = \bar{q}$ and $a_r = \bar{a}$ for all $r \in [t_1, t_2)$. Then, for $t_1 \leqslant t \leqslant s \leqslant t_2$,*

$$\int_t^s -a_r\left(Y_r - p_r\right)\mathrm{d}r = \begin{cases} -\bar{a}\sqrt{\bar{p}^2 + \bar{q}}(s - t), & Y_t = \bar{p} + \sqrt{\bar{p}^2 + \bar{q}} \\ \frac{1}{2}\ln\left(\frac{Y_t^2 - 2\bar{p}Y_t - \bar{q}}{Y_s^2 - 2\bar{p}Y_s - \bar{q}}\right), & Y_t \neq \bar{p} + \sqrt{\bar{p}^2 + \bar{q}} \end{cases} \tag{4.12}$$

*and for* $Y_t \neq \bar{p} + \sqrt{\bar{p}^2 + \bar{q}}$ *we moreover have*

$$s - t = \frac{1}{2\bar{a}\sqrt{\bar{p}^2+\bar{q}}} \ln\left(\frac{\bar{p}^2+\bar{q}-(Y_t-\bar{p})^2}{\bar{p}^2+\bar{q}-(Y_s-\bar{p})^2}\right) + \frac{1}{\bar{a}\sqrt{\bar{p}^2+\bar{q}}} \ln\left(\frac{\sqrt{\bar{p}^2+\bar{q}}+(Y_s-\bar{p})}{\sqrt{\bar{p}^2+\bar{q}}+(Y_t-\bar{p})}\right).$$

*Proof.* Rearranging the formula in (4.11) we obtain for $Y_t < \bar{p} + \sqrt{\bar{p}^2 + \bar{q}}$

$$\begin{aligned} s - t &= \frac{1}{\bar{a}\sqrt{\bar{p}^2+\bar{q}}}\left(\tanh^{-1}\left(\frac{Y_s-\bar{p}}{\sqrt{\bar{p}^2+\bar{q}}}\right) - \tanh^{-1}\left(\frac{Y_t-\bar{p}}{\sqrt{\bar{p}^2+\bar{q}}}\right)\right) \\ &= \frac{1}{2\bar{a}\sqrt{\bar{p}^2+\bar{q}}} \ln\left(\frac{\sqrt{\bar{p}^2+\bar{q}}+(Y_s-\bar{p})}{\sqrt{\bar{p}^2+\bar{q}}-(Y_s-\bar{p})}\frac{\sqrt{\bar{p}^2+\bar{q}}-(Y_t-\bar{p})}{\sqrt{\bar{p}^2+\bar{q}}+(Y_t-\bar{p})}\right) \\ &= \frac{1}{2\bar{a}\sqrt{\bar{p}^2+\bar{q}}} \ln\left(\frac{\bar{p}^2+\bar{q}-(Y_t-\bar{p})^2}{\bar{p}^2+\bar{q}-(Y_s-\bar{p})^2}\frac{\left(\sqrt{\bar{p}^2+\bar{q}}+(Y_s-\bar{p})\right)^2}{\left(\sqrt{\bar{p}^2+\bar{q}}+(Y_t-\bar{p})\right)^2}\right) \\ &= \frac{1}{2\bar{a}\sqrt{\bar{p}^2+\bar{q}}} \ln\left(\frac{\bar{p}^2+\bar{q}-(Y_t-\bar{p})^2}{\bar{p}^2+\bar{q}-(Y_s-\bar{p})^2}\right) + \frac{1}{\bar{a}\sqrt{\bar{p}^2+\bar{q}}} \ln\left(\frac{\sqrt{\bar{p}^2+\bar{q}}+(Y_s-\bar{p})}{\sqrt{\bar{p}^2+\bar{q}}+(Y_t-\bar{p})}\right), \end{aligned} \tag{4.13}$$

and for $Y_t > \bar{p} + \sqrt{\bar{p}^2 + \bar{q}}$

$$\begin{aligned} s - t &= \frac{1}{\bar{a}\sqrt{\bar{p}^2+\bar{q}}}\left(\coth^{-1}\left(\frac{Y_s-\bar{p}}{\sqrt{\bar{p}^2+\bar{q}}}\right) - \coth^{-1}\left(\frac{Y_t-\bar{p}}{\sqrt{\bar{p}^2+\bar{q}}}\right)\right) \\ &= \frac{1}{2\bar{a}\sqrt{\bar{p}^2+\bar{q}}} \ln\left(\left(-\frac{\sqrt{\bar{p}^2+\bar{q}}+(Y_s-\bar{p})}{\sqrt{\bar{p}^2+\bar{q}}-(Y_s-\bar{p})}\right)\left(-\frac{\sqrt{\bar{p}^2+\bar{q}}-(Y_t-\bar{p})}{\sqrt{\bar{p}^2+\bar{q}}+(Y_t-\bar{p})}\right)\right) \\ &= \frac{1}{2\bar{a}\sqrt{\bar{p}^2+\bar{q}}} \ln\left(\frac{\bar{p}^2+\bar{q}-(Y_t-\bar{p})^2}{\bar{p}^2+\bar{q}-(Y_s-\bar{p})^2}\right) + \frac{1}{\bar{a}\sqrt{\bar{p}^2+\bar{q}}} \ln\left(\frac{\sqrt{\bar{p}^2+\bar{q}}+(Y_s-\bar{p})}{\sqrt{\bar{p}^2+\bar{q}}+(Y_t-\bar{p})}\right). \end{aligned}$$

Now we have a look at the integral in (4.12). For $Y_t < \bar{p} + \sqrt{\bar{p}^2 + \bar{q}}$ we get

$$\begin{aligned} \int_t^s -\bar{a}\left(Y_r - \bar{p}\right) \mathrm{d}r &= \int_t^s -\bar{a}\sqrt{\bar{p}^2+\bar{q}} \tanh\left(\bar{a}\sqrt{\bar{p}^2+\bar{q}}\,(r-t) + \tanh^{-1}\left(\frac{Y_t-\bar{p}}{\sqrt{\bar{p}^2+\bar{q}}}\right)\right) \mathrm{d}r \\ &= -\ln\left(\frac{\cosh\left(\tanh^{-1}\left(\frac{Y_t-\bar{p}}{\sqrt{\bar{p}^2+\bar{q}}}\right)\right)}{\cosh\left(\bar{a}\sqrt{\bar{p}^2+\bar{q}}\,(s-t) + \tanh^{-1}\left(\frac{Y_t-\bar{p}}{\sqrt{\bar{p}^2+\bar{q}}}\right)\right)}\right) \\ &= -\ln\left(\frac{\cosh\left(\tanh^{-1}\left(\frac{Y_t-\bar{p}}{\sqrt{\bar{p}^2+\bar{q}}}\right)\right)}{\cosh\left(\tanh^{-1}\left(\frac{Y_s-\bar{p}}{\sqrt{\bar{p}^2+\bar{q}}}\right)\right)}\right) \\ &= \frac{1}{2}\ln\left(\frac{\bar{p}^2+\bar{q}-\left(Y_t-\bar{p}\right)^2}{\bar{p}^2+\bar{q}-\left(Y_s-\bar{p}\right)^2}\right), \end{aligned}$$

where we use Equation (4.13) in the second to last step. In the case of $Y_t > \bar{p} + \sqrt{\bar{p}^2 + \bar{q}}$ we obtain similarly

$$\begin{aligned}\int_t^s -\bar{a}\left(Y_r-\bar{p}\right)\mathrm{d}r &= \int_t^s -\bar{a}\sqrt{\bar{p}^2+\bar{q}}\coth\left(\bar{a}\sqrt{\bar{p}^2+\bar{q}}\,(r-t)+\coth^{-1}\left(\frac{Y_t-\bar{p}}{\sqrt{\bar{p}^2+\bar{q}}}\right)\right)\mathrm{d}r\\ &= -\ln\left(\frac{\sinh\left(\coth^{-1}\left(\frac{Y_t-\bar{p}}{\sqrt{\bar{p}^2+\bar{q}}}\right)\right)}{\sinh\left(\bar{a}\sqrt{\bar{p}^2+\bar{q}}\,(s-t)+\coth^{-1}\left(\frac{Y_t-\bar{p}}{\sqrt{\bar{p}^2+\bar{q}}}\right)\right)}\right)\\ &= \frac{1}{2}\ln\left(\frac{\bar{p}^2+\bar{q}-(Y_t-\bar{p})^2}{\bar{p}^2+\bar{q}-(Y_s-\bar{p})^2}\right).\end{aligned}$$

■

Lemma 4.30 gives us the value of the integral in (4.12) when the parameters are constant all the way. Next, we want to find the value of that integral when the process $Y$ goes up and down ending at the value where it started, which we later call an excursion.

**Lemma 4.31**
*Let Assumption 4.26 be fulfilled. Furthermore, let $[t_1,t_2],[t_3,t_4]\subset[0,\infty)$, $Y_{t_1}=Y_{t_4}\in[0,\hat{Y}]$, $Y_{t_2}=Y_{t_3}\in[0,\hat{Y}]$ and $p,q,a$ be constant on $[t_1,t_2)$ and also on $[t_3,t_4)$. Then*

$$\int_{t_1}^{t_2} -a_s(Y_s-p_s)\,\mathrm{d}s+\int_{t_3}^{t_4} -a_s(Y_s-p_s)\,\mathrm{d}s \leqslant -\min\left\{\check{a}\sqrt{\check{q}},\frac{\hat{a}(\check{Y}^2+\check{q})}{2\hat{Y}}\right\}((t_2-t_1)+(t_4-t_3)).$$

*Proof.* First we define $p_1:=p_{t_1}$, $q_1:=q_{t_1}$, $a_1:=a_{t_1}$ and $p_2:=p_{t_3}$, $q_2:=q_{t_3}$, $a_2:=a_{t_3}$ since $p$, $q$ and $a$ are constant on the intervals $[t_1,t_2)$, $[t_3,t_4)$. Now note that, due to the monotonicity of $Y$ stated in Lemma 4.29, we have one of the three cases

(i) $Y_{t_1}=Y_{t_2}=Y_{t_3}=Y_{t_4}=p_1+\sqrt{p_1^2+q_1}=p_2+\sqrt{p_2^2+q_2}$,

(ii) $p_1+\sqrt{p_1^2+q_1}<Y_{t_1}=Y_{t_4}<Y_{t_2}=Y_{t_3}<p_2+\sqrt{p_2^2+q_2}$ or

(iii) $p_2+\sqrt{p_2^2+q_2}<Y_{t_2}=Y_{t_3}<Y_{t_1}=Y_{t_4}<p_1+\sqrt{p_1^2+q_1}$.

In Case (i) it is straighforward that

$$\begin{aligned}\int_{t_1}^{t_2} -a_s(Y_s-p_s)\,\mathrm{d}s+\int_{t_3}^{t_4} -a_s(Y_s-p_s)\,\mathrm{d}s &= -a_1\sqrt{p_1^2+q_1}(t_2-t_1)-a_2\sqrt{p_2^2+q_2}(t_4-t_3)\\ &\leqslant -\check{a}\sqrt{\check{q}}(t_2-t_1+t_4-t_3).\end{aligned}$$

Now observe for Cases (ii) and (iii) that by Lemma 4.30 and since $Y_{t_1} = Y_{t_4}$, $Y_{t_2} = Y_{t_3}$ we get

$$\begin{aligned}
&\int_{t_1}^{t_2} -a_s(Y_s - p_s)\,\mathrm{d}s + \int_{t_3}^{t_4} -a_s(Y_s - p_s)\,\mathrm{d}s \\
&\quad= \frac{1}{2}\ln\left(\frac{Y_{t_1}^2 - 2p_1Y_{t_1} - q_1}{Y_{t_2}^2 - 2p_1Y_{t_2} - q_1}\right) + \frac{1}{2}\ln\left(\frac{Y_{t_3}^2 - 2p_2Y_{t_3} - q_2}{Y_{t_4}^2 - 2p_2Y_{t_4} - q_2}\right) \\
&\quad= \int_{Y_{t_1}}^{Y_{t_2}} -\frac{x - p_1}{x^2 - 2p_1x - q_1} + \frac{x - p_2}{x^2 - 2p_2x - q_2}\,\mathrm{d}x \\
&\quad= \int_{Y_{t_1}}^{Y_{t_2}} \frac{1}{2x}\left(-\frac{x^2 - 2p_1x - q_1}{x^2 - 2p_1x - q_1} + \frac{x^2 - 2p_2x - q_2}{x^2 - 2p_2x - q_2} - \frac{x^2 + q_1}{x^2 - 2p_1x - q_1} + \frac{x^2 + q_2}{x^2 - 2p_2x - q_2}\right)\mathrm{d}x \\
&\quad= \int_{Y_{t_1}}^{Y_{t_2}} \frac{1}{2x}\left(-\frac{x^2 + q_1}{x^2 - 2p_1x - q_1} + \frac{x^2 + q_2}{x^2 - 2p_2x - q_2}\right)\mathrm{d}x.
\end{aligned}$$

Furthermore, note that Case (ii) implies that $0 < x^2 - 2p_1x - q_1$ and $x^2 - 2p_2x - q_2 < 0$, while Case (iii) implies $0 > x^2 - 2p_1x - q_1$ and $x^2 - 2p_2x - q_2 > 0$ for $x$ between $Y_{t_1}$ and $Y_{t_2}$. Hence we obtain

$$\begin{aligned}
\int_{t_1}^{t_2} -a_s(Y_s - p_s)\,\mathrm{d}s + \int_{t_3}^{t_4} -a_s(Y_s - p_s)\,\mathrm{d}s &= \int_{Y_{t_1}}^{Y_{t_2}} \frac{1}{2x}\left(-\frac{x^2 + q_1}{x^2 - 2p_1x - q_1} + \frac{x^2 + q_2}{x^2 - 2p_2x - q_2}\right)\mathrm{d}x \\
&\leqslant -\left|Y_{t_2} - Y_{t_1}\right| \frac{1}{2\hat{Y}}\left(\frac{\check{Y}^2 + \check{q}}{\hat{Y}^2 - 2\check{p}\hat{Y} - \check{q}} + \frac{\check{Y}^2 + \check{q}}{\hat{q} + 2\hat{p}\check{Y} - \check{Y}^2}\right) \\
&\leqslant -\left|Y_{t_2} - Y_{t_1}\right| \frac{1}{\hat{Y}}\frac{\check{Y}^2 + \check{q}}{\hat{Y}^2 - 2\check{p}\hat{Y} - \check{q}} \qquad (4.14)
\end{aligned}$$

in Case (ii) and (iii).

It remains to estimate the term $|Y_{t_2} - Y_{t_1}|$ with an expression of time difference. To this end, remember the second result from Lemma 4.30 which gives

$$\begin{aligned}
t_2 - t_1 &= \frac{1}{2a_1\sqrt{p_1^2 + q_1}}\left(\ln\left(\frac{p_1^2 + q_1 - (Y_{t_1} - p_1)^2}{p_1^2 + q_1 - (Y_{t_2} - p_1)^2}\right) + 2\ln\left(\frac{\sqrt{p_1^2 + q_1} + (Y_{t_2} - p_1)}{\sqrt{p_1^2 + q_1} + (Y_{t_1} - p_1)}\right)\right) \\
&= \frac{1}{a_1\sqrt{p_1^2 + q_1}}\int_{Y_{t_1}}^{Y_{t_2}} -\frac{x - p_1}{p_1^2 + q_1 - (x - p_1)^2} - \frac{1}{\sqrt{p_1^2 + q_1} - p_1 + x}\,\mathrm{d}x \\
&= \frac{1}{a_1\sqrt{p_1^2 + q_1}}\int_{Y_{t_1}}^{Y_{t_2}} -\frac{x - p_1}{p_1^2 + q_1 - (x - p_1)^2} - \frac{\sqrt{p_1^2 + q_1} - (x - p_1)}{p_1^2 + q_1 - (x - p_1)^2}\,\mathrm{d}x \\
&= \frac{1}{a_1}\int_{Y_{t_1}}^{Y_{t_2}} \frac{1}{(x - p_1)^2 - p_1^2 - q_1}\,\mathrm{d}x
\end{aligned}$$

and analogously

$$t_4 - t_3 = \frac{1}{a_2}\int_{Y_{t_3}}^{Y_{t_4}} \frac{1}{(x - p_2)^2 - p_2^2 - q_2}\,\mathrm{d}x = -\frac{1}{a_2}\int_{Y_{t_1}}^{Y_{t_2}} \frac{1}{(x - p_2)^2 - p_2^2 - q_2}\,\mathrm{d}x.$$

Hence, by simular arguments as above, we get

$$\begin{aligned}
t_2 - t_1 + t_4 - t_3 &= \int_{Y_{t_1}}^{Y_{t_2}} \frac{1}{a_1((x - p_1)^2 - p_1^2 - q_1)} - \frac{1}{a_2((x - p_2)^2 - p_2^2 - q_2)}\,\mathrm{d}x \\
&\geqslant \left|Y_{t_2} - Y_{t_1}\right| \frac{2}{\hat{a}(\hat{Y}^2 - 2\check{p}\hat{Y} - \check{q})}
\end{aligned}$$

and therefore

$$|Y_{t_2} - Y_{t_1}| \leqslant (t_4 - t_3 + t_2 - t_1)\frac{\hat{a}(\hat{Y}^2 - 2\check{p}\hat{Y} - \check{q})}{2}.$$

Pluging this into Estimate (4.14) we finally obtain in Case $(ii)$ and $(iii)$

$$\begin{aligned}\int_{t_1}^{t_2} -a_s(Y_s - p_s)\,\mathrm{d}s + \int_{t_3}^{t_4} -a_s(Y_s - p_s)\,\mathrm{d}s \\ \leqslant -\frac{1}{\hat{Y}}\frac{\check{Y}^2 + \check{q}}{\hat{Y}^2 - 2\check{p}\hat{Y} - \check{q}}(t_4 - t_3 + t_2 - t_1)\frac{\hat{a}(\hat{Y}^2 - 2\check{p}\hat{Y} - \check{q})}{2} \\ = -\hat{a}\frac{\check{Y}^2 + \check{q}}{2\hat{Y}}\big(t_4 - t_3 + t_2 - t_1\big).\end{aligned}$$

■

**Lemma 4.32**
*Let Assumption 4.26 be fulfilled and assume that on the interval* $[t_0, t_1)$ *with* $0 \leqslant t_0 < t_1 < \infty$ *the functions* $p, q, a$ *are constant and* $Y_{t_0} \in [0, \hat{Y}]$. *Then*

$$\int_{t_0}^{t_1} -a_s(Y_s - p_s)\,\mathrm{d}s \leqslant -\check{a}\frac{\sqrt{\check{q}}}{\sqrt{2}}(t_1 - t_0) + \frac{2\hat{Y}}{\check{q}}\,(Y_{t_1} - Y_{t_0})\,\mathbb{1}_{\{Y_{t_1} - Y_{t_0} > 0\}}. \tag{4.15}$$

*Proof.* To shorten notation we write $\bar{p}, \bar{q}, \bar{a}$ for the constants $p_s, q_s, a_s$ with $s \in [t_0, t_1)$. Also, we set $\delta := \sqrt{\frac{\bar{q}}{2}}$. We derive estimates for the integrand of the integral in (4.15) and for the duration of the "bad" time, where those estimates do not hold true.

First, note that, since $Y$ is monotone and getting nearer to $\bar{p} + \sqrt{\bar{p}^2 + \bar{q}}$ (see Lemma 4.29), we get for any $s \in [t_0, t_1]$ with $-(Y_s - \bar{p}) \leqslant -\delta$ that for all $r \in [s, t_1]$ we have

i) for $Y_s - \left(\bar{p} + \sqrt{\bar{p}^2 + \bar{q}}\right) < 0$

$$-\left(Y_r - \bar{p}\right) = -\left(Y_r - \bar{p} - \sqrt{\bar{p}^2 + \bar{q}}\right) - \sqrt{\bar{p}^2 + \bar{q}} \leqslant 0 - \sqrt{\bar{p}^2 + \bar{q}} \leqslant -\sqrt{\frac{\bar{q}}{2}},$$

ii) for $Y_s - \left(\bar{p} + \sqrt{\bar{p}^2 + \bar{q}}\right) = 0$

$$-\left(Y_r - \bar{p}\right) = -\left(Y_r - \bar{p} - \sqrt{\bar{p}^2 + \bar{q}}\right) - \sqrt{\bar{p}^2 + \bar{q}} = 0 - \sqrt{\bar{p}^2 + \bar{q}} \leqslant -\sqrt{\frac{\bar{q}}{2}},$$

iii) for $Y_s - \left(\bar{p} + \sqrt{\bar{p}^2 + \bar{q}}\right) > 0$

$$-\left(Y_r - \bar{p}\right) \leqslant -\left(Y_s - \bar{p}\right) \leqslant -\sqrt{\frac{\bar{q}}{2}}$$

and hence in every case $-(Y_r - \bar{p}) \leqslant -\delta$. Thus, we then obtain

$$\int_s^{t_1} -\bar{a}(Y_r - \bar{p})\,\mathrm{d}r \leqslant \int_s^{t_1} -\bar{a}\delta\,\mathrm{d}r \leqslant -\bar{a}\delta(t_1 - s).$$

Now we have a closer look at the case where $-(Y_{t_0} - \bar{p}) > -\delta$. For this, remember the dynamics of $Y$ which are

$$Y_s = Y_t + \int_t^s -\bar{a}\left((Y_r - \bar{p})^2 - \bar{p}^2 - \bar{q}\right)\mathrm{d}r$$

for $s, t \in [t_0, t_1]$. There are two cases we have to consider. Firstly, $-\delta < -(Y_s - \bar{p}) < \delta$, which implies

$$-\bar{a}\left((Y_s - \bar{p})^2 - \bar{p}^2 - \bar{q}\right) > -\bar{a}\left(\left(\sqrt{\frac{\bar{q}}{2}}\right)^2 - \bar{p}^2 - \bar{q}\right) = \bar{a}\left(\bar{p}^2 + \frac{1}{2}\bar{q}\right) \geqslant \frac{\bar{a}\bar{q}}{2}.$$

And secondly the case of $-(Y_s - \bar{p}) \geqslant \delta$. Note that then $|Y_s - \bar{p}| = -Y_s + \bar{p} \leqslant \bar{p}$ since $Y \geqslant 0$ by Lemma 4.27. This gives us

$$-\bar{a}\left((Y_s - \bar{p})^2 - \bar{p}^2 - \bar{q}\right) \geqslant -\bar{a}\left((-\bar{p})^2 - \bar{p}^2 - \bar{q}\right) = \bar{a}\bar{q}.$$

Hence, for all $s \in [t_0, t_1]$ with $-(Y_s - p) \geqslant \delta$ we have $Y_s' \geqslant \frac{\bar{a}\bar{q}}{2} > 0$. Let

$$\tau := \inf\{t \in [t_0, t_1] | -(Y_t - \bar{p}) \leqslant -\delta\} \wedge t_1$$

be the first time in $[t_0, t_1]$, where $-(Y_\cdot - p) \leqslant -\delta$ or $t_1$ if there is no such time. Then we obtain

$$Y_\tau - Y_{t_0} = \int_{t_0}^{\tau} Y_t' \,\mathrm{d}t \geqslant \int_{t_0}^{\tau} \frac{\bar{a}\bar{q}}{2} \,\mathrm{d}t = \frac{\bar{a}\bar{q}}{2}(\tau - t_0)$$

and thus

$$\tau - t_0 \leqslant \frac{2}{\bar{a}\bar{q}}(Y_\tau - Y_{t_0}) = \frac{2}{\bar{a}\bar{q}}(Y_\tau - Y_{t_0}) \mathbb{1}_{\{Y_{t_1} - Y_{t_0} > 0\}} \leqslant \frac{2}{\bar{a}\bar{q}}(Y_{t_1} - Y_{t_0}) \mathbb{1}_{\{Y_{t_1} - Y_{t_0} > 0\}}, \tag{4.16}$$

where we use that if $Y_{t_1} - Y_{t_0} \leqslant 0$ we know that $Y_{t_0} \geqslant \bar{p} + \sqrt{\bar{p}^2 + \bar{q}} > \bar{p} + \delta$ and therefore $\tau = t_0$.

Hence, we have the following estimates.

- For the times where $-\bar{a}(Y - \bar{p}) \leqslant -\bar{a}\delta$ we directly estimate the integrand of the left hand side of (4.15) by $-\bar{a}\delta$.
- For the times where $-\bar{a}(Y - \bar{p}) > -\bar{a}\delta$ we can estimate the integrand of the left hand side of (4.15) by $-\bar{a}(Y_{t_0} - \bar{p})$ and the length of this time interval by Estimate (4.16).

To sum up, using that $0 \leqslant \hat{p} + \sqrt{\hat{p}^2 + \hat{q}} \geqslant \bar{p} + \sqrt{\frac{\bar{q}}{2}} - Y_{t_0}$, we derive

$$\begin{aligned}\int_{t_0}^{t_1} -\bar{a}(Y_r - \bar{p}) \,\mathrm{d}r &\leqslant -\bar{a}\delta(t_1 - \tau) - \bar{a}(Y_{t_0} - \bar{p})(\tau - t_0)\\ &= -\bar{a}\delta(t_1 - t_0) + (\bar{a}\delta - \bar{a}(Y_{t_0} - \bar{p}))(\tau - t_0)\\ &\leqslant -\check{a}\frac{\sqrt{\check{q}}}{\sqrt{2}}(t_1 - t_0) + \frac{2}{\check{q}}\left(\hat{p} + \sqrt{\hat{p}^2 + \hat{q}}\right)(Y_{t_1} - Y_{t_0}) \mathbb{1}_{\{Y_{t_1} - Y_{t_0} > 0\}}.\end{aligned}$$

■

**Proposition 4.33**
*Let Assumption 4.26 be fulfilled,* $0 \leqslant t_0 \leqslant t_1 < \infty$ *and* $Y_{t_0} \in [0, \hat{Y}]$. *Then there exist constants* $\delta_1, \delta_2 > 0$ *independent of* $t_0$ *and* $t_1$ *such that*

$$\int_{t_0}^{t_1} -a_s(Y_s - p_s) \,\mathrm{d}s \leqslant -\delta_1(t_1 - t_0) + \delta_2.$$

*Proof.* First, we have a look at functions that are piecewise constant to $p, q, a$. We split the path of $Y$ into many excursions (as described in Lemma 4.31) and left over time intervals which can not be put together to excursions. Those left over time intervals have to be such that either $Y$ is monotone decreasing or $Y$ is monotone increasing on all of them. Since $0 \leqslant Y \leqslant \hat{Y}$ (see Lemma 4.27) we get from Lemma 4.32 that the contributions of the left over monotone intervals in the estimate are bounded by $2\frac{\hat{Y}}{\check{q}}\hat{Y} =: \delta_2$. Now we set

$$\delta_1 := \min\left(\hat{a}\frac{\check{Y}^2 + \check{q}}{2\hat{Y}}, \frac{\check{a}\sqrt{\check{q}}}{\sqrt{2}}\right),$$

which is the minimum of the factors that get multiplied with the time increments, given in Lemma 4.31 and Lemma 4.32. Hence, the result holds for all piecewise constant functions $p, q, a$ uniformly.

Since $Y$ depends continuously on $a$, $p$ and $q$, for every $\varepsilon_1 > 0$ we can choose piecewise constant approximations $\tilde{a}$, $\tilde{p}$, $\tilde{q}$ fulfilling Assumption 4.26 for the same bounds as $a, p, q$ and generating a $\tilde{Y}$ such that $\max\left(\|a-\tilde{a}\|_{\infty,[t_0,t_1]}, \|p-\tilde{p}\|_{\infty,[t_0,t_1]}, \|q-\tilde{q}\|_{\infty,[t_0,t_1]}\|Y-\tilde{Y}\|_{\infty,[t_0,t_1]}\right) < \varepsilon_1$. Now observe that

$$\begin{aligned}
&\left|\int_{t_0}^{t_1} -a_s(Y_s - p_s)\,\mathrm{d}s - \int_{t_0}^{t_1} -\tilde{a}_s(\tilde{Y}_s - \tilde{p}_s)\,\mathrm{d}s\right| \\
&= \left|\int_{t_0}^{t_1} -(a_s - \tilde{a}_s)(\tilde{Y}_s - \tilde{p}_s) - a_s(Y_s - \tilde{Y}_s - (p_s - \tilde{p}_s))\,\mathrm{d}s\right| \\
&\leqslant \|a-\tilde{a}\|_\infty \left|\int_{t_0}^{t_1} \tilde{Y}_s - \tilde{p}_s\,\mathrm{d}s\right| + (\|Y-\tilde{Y}\|_\infty + \|p-\tilde{p}\|_\infty)\int_{t_0}^{t_1} a_s\,\mathrm{d}s \\
&\leqslant \|a-\tilde{a}\|_\infty T(\hat{Y} + \max\{|\hat{p}|, |\check{p}|\}) + (\|Y-\tilde{Y}\|_\infty + \|p-\tilde{p}\|_\infty)T\hat{a}.
\end{aligned}$$

Hence, we can choose for every $\varepsilon_2 > 0$ our $\varepsilon_1$ as $\varepsilon_1 = \frac{\varepsilon_2}{3T}\frac{1}{\hat{Y}+\max\{|\hat{p}|,|\check{p}|\}+\hat{a}}$ and obtain

$$\left|\int_{t_0}^{t_1} -a_s(Y_s - p_s)\,\mathrm{d}s - \int_{t_0}^{t_1} -\tilde{a}_s(\tilde{Y}_s - \tilde{p}_s)\,\mathrm{d}s\right| \leqslant \varepsilon_1 T(\hat{Y} + \max\{|\hat{p}|, |\check{p}|\}) + 2\varepsilon_1 T\hat{a} < \varepsilon_2.$$

Thus, the result for piecewise constant functions holds also true for all allowed functions $a$, $p$ and $q$. ∎

**Theorem 4.34**
*Let Assumption 4.26 be fulfilled. Denote with $Y^{t,x}$ the solution of the ODE*

$$Y_s^{t,x} = x + \int_t^s -a_r\left(\left(Y_r^{t,x}\right)^2 - 2p_r Y_r^{t,x} - q_r\right)\mathrm{d}r$$

*for $0 \leqslant t \leqslant s < \infty$. Then there are constants $K_1, K_2 > 0$ such that for all $x_1, x_2 \in [0, \hat{Y}]$ we have that*

$$\left|Y_s^{t,x_1} - Y_s^{t,x_2}\right| \leqslant |x_1 - x_2|\, K_1 e^{-K_2(s-t)}$$

*for all $0 \leqslant t \leqslant s < \infty$.*

*Proof.* First, note that for $x_0 \in \mathbb{R}$ the dynamics of $Y^{t,x_0}$ are the same as of $Y$ above. Furthermore, by introducing the function $h(r,x) := -a_r\left(x^2 - 2p_r x - q_r\right)$ for $(r,x) \in [0,\infty) \times \mathbb{R}$ and using differentiation in its weak sense, we can write the dynamics as

$$\partial_s Y_s^{t,x_0} = h(s, Y_s^{t,x_0}), \quad Y_t^{t,x_0} = x_0.$$

By standard theory (see e.g. Theorem 1 in Chapter 2.5 of [Per91]) it is known that $Y^{t,x_0}$ is also differentiable with respect to its initial value $x_0$ and that $\partial_{x_0} Y_s^{t,x_0}$ solves the differential equation

$$y'(s) = \partial_x h(s, Y_s^{t,x_0}) y(s), \qquad y_t = 1,$$

which has the solution

$$\partial_{x_0} Y_s^{t,x_0} = y(s) = \exp\left(\int_t^s \partial_x h(r, Y_s^{t,x_0})\,\mathrm{d}r\right) = \exp\left(\int_t^s -2a_r\left(Y_r^{t,x_0} - p_r\right)\mathrm{d}r\right).$$

Therefore,

$$\partial_{x_0} Y_s^{t,x_0} \leqslant \exp\left(-2\delta_1(s-t) + 2\delta_2\right)$$

for some constants $\delta_1, \delta_2 > 0$ by Proposition 4.33. Hence,

$$\begin{aligned}\left|Y_s^{t,x_1} - Y_s^{t,x_2}\right| &= \left|\int_{x_2}^{x_1} \partial_x Y_s^{t,x}\,\mathrm{d}x\right| \\ &\leqslant \left|\int_{x_2}^{x_1} \exp\left(-2\delta_1(s-t) + 2\delta_2\right)\mathrm{d}x\right| \\ &= |x_1 - x_2| \exp\left(-2\delta_1(s-t) + 2\delta_2\right).\end{aligned}$$

Thus, defining $K_1 := e^{2\delta_2}$ and $K_2 := 2\delta_1$ we obtain the claimed result. ■

**Remark 4.35**
It should be possible to generalize the results of Proposition 4.33 and Theorem 4.34 to the much more general setting, where the derivative of $Y$ is a strictly concave function having a strictly negative and a strictly positive zero. Also the starting value of $Y$ can be generalized to be greater than any negative zero. However, a proof for this claim, using abstract arguments instead of the tedious calculations as presented here, is left for future reasearch.

## 4.4 Ergodic linear-quadratic control problems

In this section we consider the ergodic linear-quadratic case. This means that we are interested in

$$\inf_{\alpha \in \mathcal{A}} \limsup_{T \to \infty} \frac{1}{T} J(T, x, \alpha), \tag{4.17}$$

which we call the optimal ergodic cost, where $x \in \mathbb{R}$ is the starting value of the controlled process. It is appearent that for this aim we need to have $\mu$, $\sigma$, $f$ and the space of admissable controls to be defined on the whole positive timeline $[0, \infty)$ and not just on $[0, T]$ for some finite $T > 0$ as before. More specifically we define

$$\mathcal{A} := \left\{\alpha : \Omega \times [0, \infty) \to \mathbb{R} \;\middle|\; \mathbb{E}\int_0^T \alpha_s^2\,\mathrm{d}s < \infty \text{ for all } T > 0\right\}$$

and make the following assumption.

**Assumption 4.36**
*Let $\mu, \sigma : [0,\infty) \times \mathbb{R} \to \mathbb{R}$ and $f : [0,\infty) \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ be of the form*

$$\mu(t,x) = b_t + B_t x, \qquad \sigma(t,x) = c_t + C_t x,$$
$$f(t,x,a) = \beta_{xx}(t)x^2 + \beta_x(t)x + \beta_{xa}(t)ax + \beta_{aa}(t)a^2 + \beta_a(t)a + \beta_0(t)$$

*for $b, B, c, C, \beta_{xx}, \beta_x, \beta_{xa}, \beta_{aa}, \beta_a, \beta_0 : [0,\infty) \to \mathbb{R}$ being deterministic, right-continuous, bounded processes, such that*

- $\det(\mathcal{H}(f))(t,\cdot,\cdot) = 4\beta_{aa}(t)\beta_{xx}(t) - \beta_{ax}^2(t) \geqslant \varepsilon_1 > 0$ *for $t \in [0,\infty)$ and some constant $\varepsilon_1 > 0$,*
- $\beta_{aa}(t) \geqslant \varepsilon_2 > 0$ *for $t \in [0,\infty)$ and some constant $\varepsilon_2 > 0$,*
- $\limsup_{t\to\infty} \frac{1}{t}\int_0^t |\beta_0(s)| \,\mathrm{d}s < \infty$.

*Also let $g \equiv 0$.*

Note that on a finite horizon $[0,T]$, $T > 0$, Assumption 4.36 is a special case of Assumption 4.17. Hence, we can use the results from Section 4.2. Moreover, we restrict ourselves to deterministic parameter functions in order to be able to apply the results from Section 4.3. Choosing $g \equiv 0$ is just a matter of convenience, which does not effect the results of the ergodic case.

**Lemma 4.37**
*Let $\sigma$, $\mu$, $f$ and $g$ fulfill Assumption 4.36. Then, for the time horizon $T > 0$, the gradient process $U^T$ of FBSDE* (4.8) *is indistinguishable from the deterministic function that fulfills the Riccati-type integral equation*

$$U_t^T = \int_t^T \left[ -\frac{\left(U_r^T\right)^2}{2\beta_{aa}(r)} + U_r^T \left( C_r^2 + 2B_r + \frac{\beta_{xa}(r)}{\beta_{aa}(r)} \right) + 2\beta_{xx}(r) - \frac{\beta_{xa}^2(r)}{2\beta_{aa}(r)} \right] \mathrm{d}r,$$

*for $t \in [0,T]$.*

*Proof.* Note that $g \equiv 0$. Corollary 4.19 gives that FBSDE (4.8) has a solution on the whole time interval $[0,T]$ and that the gradient process $U^T$ solves the BSDE

$$U_t^T = \int_t^T \left[ -\frac{\left(U_r^T\right)^2}{2\beta_{aa}(r)} + U_r^T \left( C_r^2 + 2B_r + \frac{\beta_{xa}(r)}{\beta_{aa}(r)} \right) + 2\beta_{xx}(r) - \frac{\beta_{xa}^2(r)}{2\beta_{aa}(r)} + 2C_r Z_r^U \right] \mathrm{d}r$$
$$- \int_s^T Z_r^U \,\mathrm{d}W_r.$$

Since the drift and final condition are completely deterministic except for the $Z^U$, we get by standard BSDE theory that $Z^U \equiv 0$ and hence, $U^T$ is deterministic as well. Thus, we obtain that $U^T$ solves the Riccati-type integral equation

$$U_t^T = \int_t^T \left[ -\frac{\left(U_r^T\right)^2}{2\beta_{aa}(r)} + U_r^T \left( C_r^2 + 2B_r + \frac{\beta_{xa}(r)}{\beta_{aa}(r)} \right) + 2\beta_{xx}(r) - \frac{\beta_{xa}^2(r)}{2\beta_{aa}(r)} \right] \mathrm{d}r.$$

■

**Lemma 4.38**
*Let $p, q : [0,\infty) \to \mathbb{R}$ be measurable and bounded. The integral equation*

$$h(t) = h(0) + \int_0^t [p(s) \cdot h(s) + q(s)] \, \mathrm{d}s,$$

*for $h(0) \in \mathbb{R}$ and $t \geqslant 0$, has the unique, explicit solution*

$$h(t) = e^{\int_0^t p(s)\,\mathrm{d}s} \left( h(0) + \int_0^t q(s) e^{-\int_0^s p(r)\,\mathrm{d}r} \, \mathrm{d}s \right) = h(0) e^{\int_0^t p(s)\,\mathrm{d}s} + \int_0^t q(s) e^{\int_s^t p(r)\,\mathrm{d}r} \, \mathrm{d}s.$$

*Proof.* That $h$ solves the integral equation is straightforward by weak differentiation. The uniqueness follows since the integral equation is linear in $h$ with bounded coefficents, which makes it a Lipschitz ODE. ■

**Proposition 4.39**
*Let $\sigma$, $\mu$, $f$ and $g$ fulfill Assumption 4.36 and $T > 0$. Then, the decoupling field $u$ of FBSDE* (4.8) *is equal to*

$$u(t,x) = \varphi_t^T + U_t^T \cdot x,$$

*where $U^T$ is the deterministic gradient process given in Lemma 4.37 and $\varphi^T$ is indistinguishable from the deterministic process*

$$\varphi_t^T := \int_t^T \left[ U_s^T (b_s + c_s C_s) + \beta_a(s) \frac{U_s^T - \beta_{xa}(s)}{2\beta_{aa}(s)} + \beta_x(s) \right] \exp\left( \int_t^s B_r + \frac{\beta_{xa}(r) - U_r^T}{2\beta_{aa}(r)} \, \mathrm{d}r \right) \mathrm{d}s,$$

*which solves the integral equation*

$$\varphi_t^T = \int_t^T \left[ \varphi_s^T \left( B_s + \frac{\beta_{xa}(s) - U_s^T}{2\beta_{aa}(s)} \right) + U_s^T (b_s + c_s C_s) + \beta_a(s) \frac{U_s^T - \beta_{xa}(s)}{2\beta_{aa}(s)} + \beta_x(s) \right] \mathrm{d}s.$$

*Proof.* Proposition 4.21 gives that $\varphi^T$ solves the BSDE

$$\begin{aligned} \varphi_t^T = &\int_t^T \left[ \varphi_s^T \left( B_s + \frac{\beta_{xa}(s) - U_s^T}{2\beta_{aa}(s)} \right) + C_s c_s U_s^T + b_s U_s^T + \beta_x(s) + \beta_a(s) \frac{U_s^T - \beta_{xa}(s)}{2\beta_{aa}(s)} \right] \mathrm{d}s \\ &+ \int_t^T C_s Z_s^\varphi + c_s Z_s^U \, \mathrm{d}s - \int_t^T Z_s^\varphi \, \mathrm{d}W_s. \end{aligned}$$

Lemma 4.37 gives that $Z^U \equiv 0$. Since therefore $Z^\varphi$ is the only stochastic component, we get by standard BSDE theory that $\varphi^T$ is the unique solution of the deterministic linear integral equation

$$\varphi_t^T = \int_t^T \left[ \varphi_s^T \left( B_s + \frac{\beta_{xa}(s) - U_s^T}{2\beta_{aa}(s)} \right) + U_s^T (b_s + c_s C_s) + \beta_a(s) \frac{U_s^T - \beta_{xa}(s)}{2\beta_{aa}(s)} + \beta_x(s) \right] \mathrm{d}s.$$

Solving this linear integral equation (see Lemma 4.38) yields

$$\varphi_t^T = \int_t^T \left[ U_s^T (b_s + c_s C_s) + \beta_a(s) \frac{U_s^T - \beta_{xa}(s)}{2\beta_{aa}(s)} + \beta_x(s) \right] \exp\left( \int_t^s B_r + \frac{\beta_{xa}(r) - U_r^T}{2\beta_{aa}(r)} \, \mathrm{d}r \right) \mathrm{d}s.$$

■

**Corollary 4.40**
*Let Assumption 4.36 be fulfilled and $T > 0$. Then, for the gradient process $U^T$ and $Z$ of FBSDE* (4.8) *we have for every $t \in [0,T]$*

$$Z_t = U_t^T \sigma(t, X_t) \ \text{ a.s.}$$

*Proof.* By Proposition 4.39 and Proposition 4.21 we know that

$$0 = Z_t^{\varphi} = Z_t - X_t^{\alpha} Z_t^U - \sigma\left(t, X_t^{\alpha}\right) U_t^T$$

for all $t \in [0,T]$. On the other hand, Lemma 4.37 gives $Z^U \equiv 0$. Hence, $Z_t = U_t^T \sigma(t, X_t)$ a.s. for all $t \in [0,T]$. ■

**Corollary 4.41**
*Let $\sigma$, $\mu$, $f$ and $g$ fulfill Assumption 4.36 and $T > 0$. Then, for $t \in [0,T]$ the optimal control is*

$$\alpha_t^T = \frac{\varphi_t^T - \beta_a(t) + (U_t^T - \beta_{xa}(t))X_t}{2\beta_{aa}(t)} = \alpha_t^T(X_t),$$

*which is a feedback control with the function $\alpha^T : [0,T] \times \mathbb{R} \to \mathbb{R}$*

$$\alpha_t^T(x) := \frac{\varphi_t^T - \beta_a(t) + (U_t^T - \beta_{xa}(t))x}{2\beta_{aa}(t)}.$$

*Proof.* Since Assumption 4.36 is a special case of Assumption 4.17, this is just a special case of Corollary 4.22. ■

Since it is more convenient, we make use of the feedback notation in the following.

**Proposition 4.42**
*Let Assumption 4.36 be fulfilled and $U^T$ be the gradient process with time horizon $T > 0$. Then the following holds true.*

*(i) $U^T$ is bounded independently of $T$ by $0 \leqslant U^T \leqslant \hat{U} := P + \sqrt{P^2 + Q}$, where*

$$P := \sup_{s \in [0,\infty)} \left(2B_s\beta_{aa}(s) + \beta_{xa}(s) + C_s^2\beta_{aa}(s)\right) \ \textit{and} \ Q := \sup_{s \in [0,\infty)} \left(4\beta_{xx}(s)\beta_{aa}(s) - \beta_{xa}^2(s)\right).$$

*(ii) $U^T$ converges pointwise for $T \to \infty$ to a process $U^\infty$, with the same bounds, which fulfills*

$$U_t^\infty = U_T^\infty + \int_t^T -\left(\frac{(U_r^\infty)^2}{2\beta_{aa}(r)} - U_r^\infty\left(2B_r + \frac{\beta_{xa}(r)}{\beta_{aa}(r)} + C_r^2\right) + 2\beta_{xx}(r) - \frac{\beta_{xa}^2(r)}{2\beta_{aa}(r)}\right) \mathrm{d}r$$

*for all $0 \leqslant t \leqslant T < \infty$. Furthermore, there are constants $K_1, K_2 > 0$ such that*

$$\left|U_t^\infty - U_t^T\right| \leqslant K_1 e^{-K_2(T-t)}$$

*for all $0 \leqslant t \leqslant T < \infty$.*

*(iii) There are constants* $\delta_1, \delta_2 > 0$ *such that*

$$\int_s^t \left( B_r + \frac{\beta_{xa}(r) - U_r^T}{2\beta_{aa}(r)} \right) \mathrm{d}r \leqslant \int_s^t \left( B_r + \frac{\beta_{xa}(r) - U_r^T}{2\beta_{aa}(r)} + \frac{C_r^2}{2} \right) \mathrm{d}r \leqslant -\delta_1(t-s) + \delta_2$$

*and*

$$\int_s^t \left( B_r + \frac{\beta_{xa}(r) - U_r^\infty}{2\beta_{aa}(r)} \right) \mathrm{d}r \leqslant \int_s^t \left( B_r + \frac{\beta_{xa}(r) - U_r^\infty}{2\beta_{aa}(r)} + \frac{C_r^2}{2} \right) \mathrm{d}r \leqslant -\delta_1(t-s) + \delta_2$$

*for all* $0 \leqslant s \leqslant t \leqslant T < \infty$,

*Proof.* First, we define for $s \in [0, T]$

$$\tilde{p}_s := 2B_s\beta_{aa}(s) + \beta_{xa}(s) + C_s^2\beta_{aa}(s), \qquad \tilde{q}_s := 4\beta_{xx}(s)\beta_{aa}(s) - \beta_{xa}^2(s), \qquad \tilde{a}_s := \frac{1}{2\beta_{aa}(s)}.$$

Furthermore, $p_s^T := \tilde{p}_{T-s}$, $q_s^T := \tilde{q}_{T-s}$ and $a_s^T := \tilde{a}_{T-s}$. Observe that the dynamics of $U^T$, as given in Lemma 4.37, can be written as

$$U_t^T = \int_t^T -\tilde{a}_r \left( \left(U_r^T\right)^2 - 2\tilde{p}_r U_r^T - \tilde{q}_r \right) \mathrm{d}r.$$

Now we define for $0 \leqslant s \leqslant t \leqslant T$ and $x \in [0, \hat{U}]$ the process $Y^{T,s,x}$ as the solution of the ODE

$$y'(t) = -a_t^T \left( (y(t))^2 - 2p_t^T y(t) - q_t^T \right), \qquad y(s) = x,$$

which exists and is unique by Lemma 4.27 and Assumption 4.36. Note that due to the construction of $Y$ we have for $0 \leqslant r \leqslant t \leqslant T$ that

$$Y_t^{T,0,0} = Y_t^{T,r,Y_r^{T,0,0}} = Y_{t-r}^{T-r,0,Y_r^{T,0,0}}$$

and hence for $0 \leqslant t \leqslant T \leqslant \tau$ also

$$Y_{\tau-t}^{\tau,0,0} = Y_{\tau-t}^{\tau,\tau-T,Y_{\tau-T}^{\tau,0,0}} = Y_{T-t}^{T,0,Y_{\tau-T}^{\tau,0,0}}. \tag{4.18}$$

Furthermore, a straightforward calculation yields that $(Y_t^{T,0,0})_{t\in[0,T]}$ has exactly the same dynamics as $(U_{T-t}^T)_{t\in[0,T]}$ and hence, by uniqueness, they are equal. Thus, we can apply all results from Section 4.3. Using this, Lemma 4.27 yields that $U^T$ is bounded independently of $T$ by

$$0 \leqslant U^T \leqslant \hat{U} = \sup_{s\in[0,T]} \tilde{p}_s + \sqrt{\left(\sup_{s\in[0,T]} \tilde{p}_s\right)^2 + \sup_{s\in[0,T]} \tilde{q}_s}.$$

Also, for any $0 \leqslant t_0 \leqslant t_1 \leqslant T$ we obtain

$$\int_{t_0}^{t_1} -\tilde{a}_r \left(U_r^T - \tilde{p}_r\right) \mathrm{d}r = \int_{t_0}^{t_1} -a_{T-r}^T \left(Y_{T-r}^{T,0,0} - p_{T-r}^T\right) \mathrm{d}r = \int_{T-t_1}^{T-t_0} -a_r^T \left(Y_r^{T,0,0} - p_r^T\right) \mathrm{d}r$$
$$\leqslant -\delta_1(t_1 - t_0) + \delta_2$$

by applying Proposition 4.33. Replacing $\tilde{a}$ and $\tilde{p}$ by their long forms gives the right hand estimate in $(iii)$ and noting that $\frac{C^2}{2} \geqslant 0$ the left hand one.

Next, observe that due to Equation (4.18) $U_t^\tau = Y_{\tau-t}^{\tau,0,0} = Y_{T-t}^{T,0,U_T^\tau}$ for all $0 \leqslant t \leqslant T \leqslant \tau$ and hence, by Theorem 4.34,there are constants $K_1, K_2 > 0$ such that

$$\left|U_t^\tau - U_t^T\right| = \left|Y_{T-t}^{T,0,U_T^\tau} - Y_{T-t}^{T,0,0}\right| \leqslant |U_T^\tau - 0|\, K_1 e^{-K_2(T-t)} \leqslant \hat{U} K_1 e^{-K_2(T-t)}.$$

Thus, for all $s \geqslant 0$ the sequence $\left(U_s^T\right)_{T\geqslant s}$ is a Cauchy sequence, which implies that $U^T$ converges pointwise exponentially fast to a function that we call $U^\infty$. Because $U$ is bounded independently of the time horizon we get that $U^\infty$ is bounded by the same constants.

Next, we want to show that $U^\infty$ fulfills the claimed dynamics, which can be rewritten as

$$U_t^\infty = U_T^\infty + \int_t^T -\tilde{a}_r \left( (U_r^\infty)^2 - 2\tilde{p}_r U_r^\infty - \tilde{q}_r \right) \mathrm{d}r$$

for all $0 \leqslant t \leqslant T < \infty$. To this end, define $\widetilde{U}^{T,\infty}$ as the solution of

$$y(t) = U_T^\infty + \int_t^T -\tilde{a}_r \left( (y(t))^2 - 2\tilde{p}_r y(t) - \tilde{q}_r \right) \mathrm{d}r.$$

Since the solutions of ODEs depend continuously on their starting value (see e.g. Theorem 1 in Chapter 2.5 of [Per91]) we can, using Equation (4.18), for all $0 \leqslant t \leqslant T < \infty$ derive

$$U_t^\infty = \lim_{\tau \to \infty} U_t^\tau = \lim_{\tau \to \infty} Y_{\tau - t}^{\tau,0,0} = \lim_{\tau > T, \tau \to \infty} Y_{T-t}^{T,0,U_T^\tau} = Y_{T-t}^{T,0,U_T^\infty} = \widetilde{U}_t^{T,\infty},$$

which proves this result. Since therefore $U_t^\infty = Y_{T-t}^{T,0,U_T^\infty}$ for all $0 \leqslant t \leqslant T$, we can furthermore estimate with Proposition 4.33

$$\int_{t_0}^{t_1} -\tilde{a}_r \left( U_r^\infty - \tilde{p}_r \right) \mathrm{d}r = \int_{T-t_1}^{T-t_0} -a_r^T \left( Y_r^{T,0,U_T^\infty} - p_r^T \right) \mathrm{d}r \leqslant -\delta_1 (t_1 - t_0) + \delta_2.$$

■

**Remark 4.43**

(i) It might seem a little strange that we describe the dynamics of $U^\infty$ via a backward equation although there is no real final value. On the other hand, the value of $U_0^\infty$ is unknown and a small error would increase over time, due to the dynamics. Furthermore, the backward notation is similar to the notation of $U^T$ and in a computation the influence of an error in the final value would decrease over time (see Theorem 4.34).

(ii) In the proof of Proposition 4.42 it becomes clear why we choose the parameter functions to be deterministic in contrast to Section 4.2, where they are stochastic. In order to obtain the convergence of $U^T$ to $U^\infty$ we need to apply the results from Section 4.3, which are only available for deterministic parameter functions.

(iii) After extending our parameter functions to negative time e.g. by mirroring them, we can view $\left( Y_t^{T,0,U_T^\infty} \right)$ from the proof of Proposition 4.42 as a single point pullback attractor as in Definition 2.3 from [Sch00]. Since the theory for pullback attractors also extends to random processes (see e.g. [Sch99]), we hope that this theory will enable future research to overcome the restrictions described in Point (ii).

Although Proposition 4.42 states that the gradient process $U^T$ converges to a process $U^\infty$ and gives estimates we use later on, we are not able to give explicit solutions due to the time inhomogenous parameters. In the following corollary we have a look at the time homogenous case, where the formulas become explicit.

**Corollary 4.44**
*Let Assumption 4.36 be fulfilled and additionally all parameter functions be constant. Then, for this time-homogeneous problem, the gradient process $U^T$ is explicitly given for all every $T > 0$ and $t \in [0, T]$ by*

$$U_t^T = p + \sqrt{p^2+q}\tanh\left(a\sqrt{p^2+q}(T-t) + \tanh^{-1}\left(-\frac{p}{\sqrt{p^2+q}}\right)\right),$$

*where $p := 2B_0\beta_{aa}(0)+\beta_{xa}(0)+C_0^2\beta_{aa}(0)$, $q := 4\beta_{xx}(0)\beta_{aa}(0)-\beta_{xa}^2(0)$, $a := \frac{1}{2\beta_{aa}(0)}$. Furthermore,*

$$U_t^\infty = p + \sqrt{p^2+q}$$

*for all $t \in [0, \infty)$, which means that it is constant.*

*Proof.* The first statement can be verified by using the construction of $Y$ in the proof of Proposition 4.42 together with Lemma 4.29. The second statement follows directly by considering the limit $\lim_{T\to\infty} U_t^T$. ■

The next lemma shows that not just $U^T$ but also $\varphi^T$ converges, which means that the decoupling field and the optimal feedback control converge.

**Lemma 4.45**
*Let Assumption 4.36 be fulfilled. Then $\varphi^T$ is bounded by some constant $\hat{\varphi} > 0$, independently of $T > 0$, and converges for $T \to \infty$ pointwise to the bounded function $\varphi^\infty$ which is defined by*

$$\varphi_t^\infty := \int_t^\infty \left[U_s^\infty\left(b_s + c_sC_s + \frac{\beta_a(s)}{2\beta_{aa}(s)}\right) - \frac{\beta_a(s)\beta_{xa}(s)}{2\beta_{aa}(s)} + \beta_x(s)\right] \cdot \exp\left(\int_t^s B_r + \frac{\beta_{xa}(r)-U_r^\infty}{2\beta_{aa}(r)}\,\mathrm{d}r\right)\mathrm{d}s.$$

*Moreover, $\varphi^\infty$ solves for all $0 \leqslant t \leqslant T < \infty$ the integral equation*

$$\varphi_t^\infty = \varphi_T^\infty - \int_t^T \left[\varphi_s^\infty\left(B_s + \frac{\beta_{xa}(s)-U_s^\infty}{2\beta_{aa}(s)}\right) + U_s^\infty\left(b_s + c_sC_s + \frac{\beta_a(s)}{2\beta_{aa}(s)}\right) - \frac{\beta_a(s)\beta_{xa}(s)}{2\beta_{aa}(s)} + \beta_x(s)\right]\mathrm{d}s.$$

*Also, there are constants $k_1, k_2 > 0$ such that*

$$\left|\varphi_t^\infty - \varphi_t^T\right| \leqslant k_1 e^{-k_2(T-t)},$$

*which further yields*

$$\int_0^T \left|\varphi_t^\infty - \varphi_t^T\right|\mathrm{d}t \leqslant \frac{k_1}{k_2}$$

*for all $0 \leqslant t \leqslant T < \infty$.*

*Proof.* Proposition 4.39 gives us $\varphi^T$ as

$$\varphi_t^T = \int_t^T \left[U_s^T\left(b_s + c_sC_s + \frac{\beta_a(s)}{2\beta_{aa}(s)}\right) - \frac{\beta_a(s)\beta_{xa}(s)}{2\beta_{aa}(s)} + \beta_x(s)\right]\exp\left(\int_t^s B_r + \frac{\beta_{xa}(r)-U_r^T}{2\beta_{aa}(r)}\,\mathrm{d}r\right)\mathrm{d}s.$$

By Proposition 4.42 there are $\delta_1, \delta_2 > 0$ such that for all $s \geqslant t \geqslant 0$

$$\exp\left(\int_t^s B_r + \frac{\beta_{xa}(r) - U_r^T}{2\beta_{aa}(r)}\,\mathrm{d}r\right) \leqslant \exp\big(-\delta_1(s-t) + \delta_2\big).$$

Since furthermore $U^T$ is bounded by Proposition 4.42 and the processes

$$\left(b_s + c_s C_s + \frac{\beta_a(s)}{2\beta_{aa}(s)}\right)_{s\in[0,\infty)}, \qquad \left(-\frac{\beta_a(s)\beta_{xa}(s)}{2\beta_{aa}(s)} + \beta_x(s)\right)_{s\in[0,\infty)}$$

are bounded by Assumption 4.36, we get that $\varphi^T$ is bounded independently of $T$.

Also by Proposition 4.42 we obtain

$$\begin{aligned}
&\int_t^\infty \left|\left[U_s^\infty\left(b_s + c_s C_s + \frac{\beta_a(s)}{2\beta_{aa}(s)}\right) - \frac{\beta_a(s)\beta_{xa}(s)}{2\beta_{aa}(s)} + \beta_x(s)\right]\exp\left(\int_t^s B_r + \frac{\beta_{xa}(r) - U_r^\infty}{2\beta_{aa}(r)}\,\mathrm{d}r\right)\right|\mathrm{d}s\\
&\quad\leqslant \sup_{s\in[0,\infty)}\left[\hat{U}\left|b_s + c_s C_s + \frac{\beta_a(s)}{2\beta_{aa}(s)}\right| + \left|-\frac{\beta_a(s)\beta_{xa}(s)}{2\beta_{aa}(s)} + \beta_x(s)\right|\right]\int_t^\infty e^{-\delta_1(s-t)+\delta_2}\,\mathrm{d}s\\
&\quad= \sup_{s\in[0,\infty)}\left[\hat{U}\left|b_s + c_s C_s + \frac{\beta_a(s)}{2\beta_{aa}(s)}\right| + \left|-\frac{\beta_a(s)\beta_{xa}(s)}{2\beta_{aa}(s)} + \beta_x(s)\right|\right]\frac{e^{\delta_2}}{\delta_1}\\
&\quad< \infty,
\end{aligned}$$

which means that $\varphi^\infty$ is well defined and bounded. Furthermore, we get for $T \geqslant t \geqslant 0$

$$\begin{aligned}
&|\varphi_t^\infty - \varphi_t^T|\\
&\quad= \Bigg|\int_t^T \left[U_s^\infty\left(b_s + c_s C_s + \frac{\beta_a(s)}{2\beta_{aa}(s)}\right) - \frac{\beta_a(s)\beta_{xa}(s)}{2\beta_{aa}(s)} + \beta_x(s)\right]\\
&\qquad\qquad \left[\exp\left(\int_t^s B_r + \frac{\beta_{xa}(r) - U_r^\infty}{2\beta_{aa}(r)}\,\mathrm{d}r\right) - \exp\left(\int_t^s B_r + \frac{\beta_{xa}(r) - U_r^T}{2\beta_{aa}(r)}\,\mathrm{d}r\right)\right]\mathrm{d}s\\
&\qquad + \int_t^T (U_s^\infty - U_s^T)\left(b_s + c_s C_s + \frac{\beta_a(s)}{2\beta_{aa}(s)}\right) e^{\int_t^s B_r + \frac{\beta_{xa}(r) - U_r^T}{2\beta_{aa}(r)}\,\mathrm{d}r}\,\mathrm{d}s\\
&\qquad + \int_T^\infty \left[U_s^\infty\left(b_s + c_s C_s + \frac{\beta_a(s)}{2\beta_{aa}(s)}\right) - \frac{\beta_a(s)\beta_{xa}(s)}{2\beta_{aa}(s)} + \beta_x(s)\right] e^{\int_t^s B_r + \frac{\beta_{xa}(r) - U_r^\infty}{2\beta_{aa}(r)}\,\mathrm{d}r}\,\mathrm{d}s\Bigg|\\
&\quad\leqslant \left[\hat{U}\sup_{s\in[0,\infty)}\left|b_s + c_s C_s + \frac{\beta_a(s)}{2\beta_{aa}(s)}\right| + \sup_{s\in[0,\infty)}\left|-\frac{\beta_a(s)\beta_{xa}(s)}{2\beta_{aa}(s)} + \beta_x(s)\right|\right]\\
&\qquad\qquad \int_t^T\left|\exp\left(\int_t^s B_r + \frac{\beta_{xa}(r) - U_r^\infty}{2\beta_{aa}(r)}\,\mathrm{d}r\right) - \exp\left(\int_t^s B_r + \frac{\beta_{xa}(r) - U_r^T}{2\beta_{aa}(r)}\,\mathrm{d}r\right)\right|\mathrm{d}s\\
&\qquad + \sup_{s\in[0,\infty)}\left|b_s + c_s C_s + \frac{\beta_a(s)}{2\beta_{aa}(s)}\right|\left|\int_t^T K_1 e^{-K_2(T-s)} e^{-\delta_1(s-t)+\delta_2}\,\mathrm{d}s\right.\\
&\qquad + \left[\hat{U}\sup_{s\in[0,\infty)}\left|b_s + c_s C_s + \frac{\beta_a(s)}{2\beta_{aa}(s)}\right| + \sup_{s\in[0,\infty)}\left|\frac{\beta_a(s)\beta_{xa}(s)}{2\beta_{aa}(s)} - \beta_x(s)\right|\right]\int_T^\infty e^{-\delta_1(s-t)+\delta_2}\,\mathrm{d}s
\end{aligned}$$

by Proposition 4.42 with some constants $\delta_1, \delta_2, K_1, K_2 > 0$. Next we estimate the three integrals on the right hand side separately. In order to estimate the first summand, we obtain by again using Proposition 4.42

$$
\begin{aligned}
&\int_t^T \left|\exp\left(\int_t^s B_r + \frac{\beta_{xa}(r) - U_r^\infty}{2\beta_{aa}(r)}\,\mathrm{d}r\right) - \exp\left(\int_t^s B_r + \frac{\beta_{xa}(r) - U_r^T}{2\beta_{aa}(r)}\,\mathrm{d}r\right)\right| \mathrm{d}s \\
&\quad\leqslant \int_t^{\frac{T+t}{2}} \exp\left(\int_t^s B_r + \frac{\beta_{xa}(r) - U_r^\infty}{2\beta_{aa}(r)}\,\mathrm{d}r\right) \left|1 - \exp\left(\int_t^s \frac{|U_r^T - U_r^\infty|}{2\beta_{aa}(r)}\,\mathrm{d}r\right)\right| \mathrm{d}s \\
&\qquad\qquad + \int_{\frac{T+t}{2}}^T \exp\left(\int_t^s B_r + \frac{\beta_{xa}(r) - U_r^\infty}{2\beta_{aa}(r)}\,\mathrm{d}r\right) + \exp\left(\int_t^s B_r + \frac{\beta_{xa}(r) - U_r^T}{2\beta_{aa}(r)}\,\mathrm{d}r\right) \mathrm{d}s \\
&\quad\leqslant \int_t^{\frac{T+t}{2}} e^{-\delta_1(s-t)+\delta_2} \left(\exp\left(\frac{1}{2\check{\beta}_{aa}} \int_t^s K_1 e^{-K_2(T-r)}\,\mathrm{d}r\right) - 1\right) \mathrm{d}s + \int_{\frac{T+t}{2}}^T 2e^{-\delta_1(s-t)+\delta_2}\,\mathrm{d}s \\
&\quad\leqslant \int_t^{\frac{T+t}{2}} e^{-\delta_1(s-t)+\delta_2} \left(\exp\left(\frac{K_1}{2\check{\beta}_{aa}} e^{-K_2\left(T-\frac{T+t}{2}\right)}(s-t)\right) - 1\right) \mathrm{d}s + \int_{\frac{T+t}{2}}^T 2e^{-\delta_1(s-t)+\delta_2}\,\mathrm{d}s \\
&\quad= e^{\delta_2} \frac{\frac{K_1}{2\check{\beta}_{aa}} e^{-K_2\frac{T-t}{2}} \left[1 - e^{-\delta_1\frac{T-t}{2}}\right] + \delta_1 e^{-\delta_1\frac{T-t}{2}} \left[1 - e^{\frac{K_1}{2\check{\beta}_{aa}} e^{-K_2\frac{T-t}{2}}\frac{T-t}{2}}\right]}{\delta_1(\delta_1 - \frac{K_1}{2\check{\beta}_{aa}} e^{-K_2\frac{T-t}{2}})} \\
&\qquad\qquad + \frac{2e^{\delta_2}}{\delta_1}\left[e^{-\delta_1\frac{T-t}{2}} - e^{-\delta_1(T-t)}\right] \\
&\quad\leqslant \frac{e^{\delta_2} e^{-\frac{\min\{\delta_1,K_2\}}{2}(T-t)}}{\delta_1(\delta_1 - \frac{K_1}{2\check{\beta}_{aa}} e^{-\frac{K_2}{2}(T-t)})} \left[\frac{K_1}{2\check{\beta}_{aa}}\left(1 - e^{-\frac{\delta_1}{2}(T-t)}\right) + \delta_1\left(1 - \exp\left(\frac{K_1}{4\check{\beta}_{aa}}(T-t)e^{-\frac{K_2}{2}(T-t)}\right)\right)\right].
\end{aligned}
$$

For the second summand we get

$$
\begin{aligned}
\int_t^T K_1 e^{-K_2(T-s)} e^{-\delta_1(s-t)+\delta_2}\,\mathrm{d}s &\leqslant \int_t^T K_1 e^{\delta_2} e^{-\min\{K_2,\delta_1\}(T-s+s-t)}\,\mathrm{d}s \\
&= K_1 e^{\delta_2}(T-t)e^{-\min\{K_2,\delta_1\}(T-t)}
\end{aligned}
$$

and for the third one

$$
\int_T^\infty e^{-\delta_1(s-t)+\delta_2}\,\mathrm{d}s = \frac{e^{\delta_2}}{\delta_1} e^{-\delta_1(T-t)}.
$$

Remember that $(T-t)e^{-k(T-t)}$ goes to $0$ for $T \to \infty$ and is furthermore bounded for all $k > 0$. Thus, by putting the three integrals back together, we can find constants $k_1, k_2 > 0$ depending only on the bounds in Assumption 4.36 such that

$$
\left|\varphi_t^\infty - \varphi_t^T\right| \leqslant k_1 e^{-k_2(T-t)},
$$

which means that $\varphi^T$ converges to $\varphi^\infty$ exponentially fast.

Now we take a look at $\int_0^T \left|\varphi_t^\infty - \varphi_t^T\right| \mathrm{d}t$. Integrating the above estimate yields

$$
\int_0^T \left|\varphi_t^\infty - \varphi_t^T\right| \mathrm{d}t \leqslant \frac{k_1}{k_2}\left(1 - e^{-k_2 T}\right) \leqslant \frac{k_1}{k_2}
$$

independently of $T$.

Finally, we turn to proving the integral equation. Note that Proposition 4.39 states the $\varphi^T$ solves a similar integral equation. Using this and dominated convergence, we obtain for

$0 \leqslant t \leqslant T < \infty$

$$\begin{aligned}\varphi_t^\infty &= \lim_{\substack{\hat{T}\to\infty\\ T\leqslant \hat{T}}} \left[\varphi_T^{\hat{T}} + \int_t^T \left(\varphi_s^{\hat{T}}\left(B_s + \frac{\beta_{xa}(s) - U_s^{\hat{T}}}{2\beta_{aa}(s)}\right)\right.\right.\\ &\qquad \left.\left. + U_s^{\hat{T}}\left(b_s + c_s C_s + \frac{\beta_a(s)}{2\beta_{aa}(s)}\right) - \frac{\beta_a(s)\beta_{xa}(s)}{2\beta_{aa}(s)} + \beta_x(s)\right) \mathrm{d}s\right]\\ &= \varphi_T^\infty + \int_t^T \left[\lim_{\substack{\hat{T}\to\infty\\ T\leqslant \hat{T}}} \varphi_s^{\hat{T}}\left(B_s + \lim_{\substack{\hat{T}\to\infty\\ T\leqslant \hat{T}}} \frac{\beta_{xa}(s) - U_s^{\hat{T}}}{2\beta_{aa}(s)}\right)\right.\\ &\qquad \left. + \lim_{\substack{\hat{T}\to\infty\\ T\leqslant \hat{T}}} U_s^{\hat{T}}\left(b_s + c_s C_s + \frac{\beta_a(s)}{2\beta_{aa}(s)}\right) - \frac{\beta_a(s)\beta_{xa}(s)}{2\beta_{aa}(s)} + \beta_x(s)\right] \mathrm{d}s\\ &= \varphi_T^\infty + \int_t^T \left[\varphi_s^\infty\left(B_s + \frac{\beta_{xa}(s) - U_s^\infty}{2\beta_{aa}(s)}\right)\right.\\ &\qquad \left. + U_s^\infty\left(b_s + c_s C_s + \frac{\beta_a(s)}{2\beta_{aa}(s)}\right) - \frac{\beta_a(s)\beta_{xa}(s)}{2\beta_{aa}(s)} + \beta_x(s)\right] \mathrm{d}s.\end{aligned}$$

■

**Corollary 4.46**
*Let Assumption 4.36 be fulfilled and additionally all parameter functions be constant. Then, using the definition of $p, q, a$ from Corollary 4.44, for all $t \geqslant 0$*

$$\varphi_t^\infty = \frac{\left(p + \sqrt{p^2+q}\right)(b_0 + c_0 C_0 + a\beta_a(0)) - a\beta_a(0)\beta_{xa}(0) + \beta_x(0)}{a\sqrt{p^2+q} + \frac{C_0^2}{2}},$$

*which in particular means that $\varphi^\infty$ is constant.*

*Proof.* Using the definition of $\varphi^\infty$ in Lemma 4.45 and the result from Corollary 4.44 we obtain

$$\begin{aligned}\varphi_t^\infty &= \int_t^\infty \left[\left(p + \sqrt{p^2+q}\right)\left(b_0 + c_0 C_0 + a\beta_a(0)\right) - a\beta_a(0)\beta_{xa}(0) + \beta_x(0)\right] e^{\int_t^s \left(-a\sqrt{p^2+q} - \frac{C_0^2}{2}\right) \mathrm{d}r} \mathrm{d}s\\ &= \left[\frac{\left(p + \sqrt{p^2+q}\right)\left(b_0 + c_0 C_0 + a\beta_a(0)\right) - a\beta_a(0)\beta_{xa}(0) + \beta_x(0)}{-a\sqrt{p^2+q} - \frac{C_0^2}{2}}\right.\\ &\qquad \left. \exp\left((s-t)\left(-a\sqrt{p^2+q} - \frac{C_0^2}{2}\right)\right)\right]_{s=t}^\infty\\ &= \frac{\left(p + \sqrt{p^2+q}\right)\left(b_0 + c_0 C_0 + a\beta_a(0)\right) - a\beta_a(0)\beta_{xa}(0) + \beta_x(0)}{a\sqrt{p^2+q} + \frac{C_0^2}{2}}.\end{aligned}$$

■

**Corollary 4.47**
*Let Assumption 4.36 be fulfilled. Then the optimal feedback controls $\alpha^T$ converge for $T \to \infty$ pointwise to the feedback control*

$$\alpha_t^\infty(x) = \frac{\varphi_t^\infty - \beta_a(t) + (U_t^\infty - \beta_{xa}(t))x}{2\beta_{aa}(t)}$$

*Proof.* Since $\varphi^T$ and $U^T$ converge for $T \to \infty$ and all other components do not depend on $T$, this is straightforward. ■

We have to distinguish between the process controlled by the control $\alpha^T$, which is optimal for horizon $T$, and the ones controlled by the ergodic control $\alpha^\infty$. Therefore, recall the notation from the beginning of Section 4.1: By $X^\alpha$ we denote the state process controlled by the feedback control $\alpha$. This means that $X^\alpha$ solves the SDE

$$X_t^\alpha = x_0 + \int_0^t (\mu(s, X_s^\alpha) - \alpha_s(X_s^\alpha))\,\mathrm{d}s + \int_0^t \sigma(s, X_s^\alpha)\,\mathrm{d}W_s.$$

In order to show that $\alpha^\infty$ solves the ergodic control problem (4.17), we have do derive some estimates for the first moments of the controlled process.

**Remark 4.48**
Since $X^{\alpha^T}$ and $X^{\alpha^\infty}$ both have Lipschitz continuous dynamics, we know by standard theory (see e.g. [Kun97]) that $\mathbb{E}\big[|X_\cdot^{\alpha^T}|^{2m}\big]$ and $\mathbb{E}\big[|X_\cdot^{\alpha^\infty}|^{2m}\big]$ for all $m \in \mathbb{N}$ are finite and integrable. This in particular implies that terms like $\int_0^\cdot \big(X_s^{\alpha^T}\big)^2\,\mathrm{d}W_s$ or $\int_0^\cdot \big(X_s^{\alpha^\infty}\big)^2\,\mathrm{d}W_s$ are true martingales and hence have an expectation of 0. We use this fact in the following lemmas without mentoning it.

**Lemma 4.49**
*Let Assumption 4.36 be fulfilled. Then*

$$\mathbb{E}\left[X_t^{\alpha^T}\right] = x_0 e^{\int_0^t B_s + \frac{\beta_{xa}(s) - U_s^T}{2\beta_{aa}(s)}\,\mathrm{d}s} + \int_0^t \left(b_s + \frac{-\varphi_s^T + \beta_a(s)}{2\beta_{aa}(s)}\right) e^{\int_s^t B_r + \frac{\beta_{xa}(r) - U_r^T}{2\beta_{aa}(r)}\,\mathrm{d}r}\,\mathrm{d}s,$$

$$\begin{aligned}
\mathbb{E}\left[\left(X_t^{\alpha^T}\right)^2\right] &= x_0^2 e^{\int_0^t B_s + \frac{\beta_{xa}(s) - U_s^T}{2\beta_{aa}(s)} + \frac{C_s^2}{2}\,\mathrm{d}s} \\
&\quad + \int_0^t \left[c_s^2 + 2\mathbb{E}\left[X_s^{\alpha^T}\right]\left(b_s + \frac{-\varphi_s^T + \beta_a(s)}{2\beta_{aa}(s)} + 2c_sC_s\right)\right] e^{\int_s^t B_r + \frac{\beta_{xa}(r) - U_r^T}{2\beta_{aa}(r)} + \frac{C_s^2}{2}\,\mathrm{d}r}\,\mathrm{d}s \\
&= x_0^2 e^{\int_0^t B_s + \frac{\beta_{xa}(s) - U_s^T}{2\beta_{aa}(s)} + \frac{C_s^2}{2}\,\mathrm{d}s} + \int_0^t c_s^2 e^{\int_s^t B_r + \frac{\beta_{xa}(r) - U_r^T}{2\beta_{aa}(r)} + \frac{C_s^2}{2}\,\mathrm{d}r}\,\mathrm{d}s \\
&\quad + \int_0^t 2\left[x_0 e^{\int_0^s B_s + \frac{\beta_{xa}(r) - U_r^T}{2\beta_{aa}(r)}\,\mathrm{d}r} + \int_0^s \left(b_r + \frac{-\varphi_r^T + \beta_a(r)}{2\beta_{aa}(r)}\right) e^{\int_r^s B_v + \frac{\beta_{xa}(v) - U_v^T}{2\beta_{aa}(v)}\,\mathrm{d}v}\,\mathrm{d}r\right] \\
&\qquad \cdot \left[b_s + \frac{-\varphi_s^T + \beta_a(s)}{2\beta_{aa}(s)} + 2c_sC_s\right] e^{\int_s^t B_r + \frac{\beta_{xa}(r) - U_r^T}{2\beta_{aa}(r)} + \frac{C_s^2}{2}\,\mathrm{d}r}\,\mathrm{d}s
\end{aligned}$$

*and the moments* $\mathbb{E}\Big[X_t^{\alpha^T}\Big]$, $\mathbb{E}\Big[X_t^{\alpha^\infty}\Big]$, $\mathbb{E}\Big[\big(X_t^{\alpha^T}\big)^2\Big]$ *and* $\mathbb{E}\Big[\big(X_t^{\alpha^\infty}\big)^2\Big]$ *are bounded independently of* $0 \leqslant t \leqslant T < \infty$ *for every initial value* $x_0 \in \mathbb{R}$.

*Proof.* Observe that by Corollary 4.22

$$\begin{aligned}
\mathbb{E}\left[X_t^{\alpha^T}\right] &= x_0 + \mathbb{E}\left[\int_0^t \left(b_s + B_s X_s^{\alpha^T} - \frac{\varphi_s^T - \beta_a(s) + (U_s^T - \beta_{xa}(s))X_s^{\alpha^T}}{2\beta_{aa}(s)}\right)\mathrm{d}s\right] \\
&= x_0 + \int_0^t \left(b_s + \frac{-\varphi_s^T + \beta_a(s)}{2\beta_{aa}(s)}\right)\mathrm{d}s + \int_0^t \left(B_s + \frac{\beta_{xa}(s) - U_s^T}{2\beta_{aa}(s)}\right)\mathbb{E}\left[X_s^{\alpha^T}\right]\mathrm{d}s.
\end{aligned}$$

By Lemma 4.38 we get

$$\mathbb{E}\left[X_t^{\alpha^T}\right] = x_0 e^{\int_0^t B_s + \frac{\beta_{xa}(s) - U_s^T}{2\beta_{aa}(s)}\,\mathrm{d}s} + \int_0^t \left(b_s + \frac{-\varphi_s^T + \beta_a(s)}{2\beta_{aa}(s)}\right) e^{\int_s^t B_r + \frac{\beta_{xa}(r) - U_r^T}{2\beta_{aa}(r)}\,\mathrm{d}r}\,\mathrm{d}s$$

and hence, using Proposition 4.42 and that

$$\left|b_s + \frac{-\varphi_s^T + \beta_a(s)}{2\beta_{aa}(s)}\right| \leqslant \sup_{r\in[0,\infty)} |b_r| + \frac{\hat{\varphi} + \sup_{r\in[0,\infty)} |\beta_a(r)|}{2\check{\beta}_{aa}} < \infty,$$

we obtain

$$\begin{aligned}
\left|\mathbb{E}\left[X_t^{\alpha^T}\right]\right| &\leqslant |x_0| e^{-\delta_1(t-0)+\delta_2} + \sup_{r\in[0,\infty)} \left|b_r + \frac{-\varphi_r^T + \beta_a(r)}{2\beta_{aa}(r)}\right| \int_0^t e^{-\delta_1(t-r)+\delta_2}\,\mathrm{d}s \\
&= |x_0| e^{\delta_2} e^{-\delta_1 t} + \left(1 - e^{-\delta_1 t}\right) \frac{e^{\delta_2}}{\delta_1} \sup_{r\in[0,\infty)} \left|b_r + \frac{-\varphi_r^T + \beta_a(r)}{2\beta_{aa}(r)}\right| \\
&\leqslant \max\left(|x_0| e^{\delta_2}, \frac{e^{\delta_2}}{\delta_1}\left(\sup_{r\in[0,\infty)} |b_r| + \frac{\hat{\varphi} + \sup_{r\in[0,\infty)} |\beta_a(r)|}{2\check{\beta}_{aa}}\right)\right).
\end{aligned}$$

Completely analogous we also have that

$$\left|\mathbb{E}\left[X_t^{\alpha^\infty}\right]\right| \leqslant \max\left(|x_0| e^{\delta_2}, \frac{e^{\delta_2}}{\delta_1}\left(\sup_{r\in[0,\infty)} |b_r| + \frac{\hat{\varphi} + \sup_{r\in[0,\infty)} |\beta_a(r)|}{2\check{\beta}_{aa}}\right)\right).$$

Furthermore, using Itô's Formula

$$\begin{aligned}
&\mathbb{E}\left[\left(X_t^{\alpha^T}\right)^2\right] \\
&= x_0^2 + \mathbb{E}\left[\int_0^t 2X_s^{\alpha^T}\left(b_s + B_s X_s^{\alpha^T} - \frac{\varphi_s^T - \beta_a(s) + (U_s^T - \beta_{xa}(s)) X_s^{\alpha^T}}{2\beta_{aa}(s)}\right) + \left(c_s + C_s X_s^{\alpha^T}\right)^2 \mathrm{d}s\right] \\
&= x_0^2 + \int_0^t \left(c_s^2 + 2\mathbb{E}\left[X_s^{\alpha^T}\right]\left(b_s + \frac{-\varphi_s^T + \beta_a(s)}{2\beta_{aa}(s)} + 2c_s C_s\right)\right) \mathrm{d}s \\
&\quad + \int_0^t 2\left(B_s + \frac{\beta_{xa}(s) - U_s^T}{2\beta_{aa}(s)} + \frac{C_s^2}{2}\right) \mathbb{E}\left[\left(X_s^{\alpha^T}\right)^2\right] \mathrm{d}s \\
&= x_0^2 \exp\left(\int_0^t B_s + \frac{\beta_{xa}(s) - U_s^T}{2\beta_{aa}(s)} + \frac{C_s^2}{2}\,\mathrm{d}s\right) \\
&\quad + \int_0^t \left[c_s^2 + 2\mathbb{E}\left[X_s^{\alpha^T}\right]\left(b_s + \frac{-\varphi_s^T + \beta_a(s)}{2\beta_{aa}(s)} + 2c_s C_s\right)\right] \exp\left(\int_s^t B_r + \frac{\beta_{xa}(r) - U_r^T}{2\beta_{aa}(r)} + \frac{C_s^2}{2}\,\mathrm{d}r\right) \mathrm{d}s
\end{aligned}$$

due to Lemma 4.38. Plugging in the formula for $\mathbb{E}\left[X_s^{\alpha^T}\right]$ we get the second claimed equation for $\mathbb{E}\left[\left(X_t^{\alpha^T}\right)^2\right]$. By Proposition 4.42 and the above result we can moreover estimate

$$\begin{aligned}
&\left|\mathbb{E}\left[\left(X_t^{\alpha^T}\right)^2\right]\right| \\
&\quad \leqslant x_0^2 e^{-\delta_1(t-0)+\delta_2} \\
&\qquad + \left(\sup_{s\in[0,\infty)}\left|c_s^2\right| + 2\sup_{s\in[0,\infty)}\left|\mathbb{E}\left[X_s^{\alpha^T}\right]\right|\sup_{s\in[0,\infty)}\left|b_s + \frac{-\varphi_s^T+\beta_a(s)}{2\beta_{aa}(s)} + 2c_sC_s\right|\right)\int_0^t e^{-\delta_1(t-r)+\delta_2}\,\mathrm{d}s \\
&\quad = x_0^2 e^{\delta_2}e^{-\delta_1 t} \\
&\qquad + \left(1-e^{-\delta_1 t}\right)\frac{e^{\delta_2}}{\delta_1}\left(\sup_{r\in[0,\infty)}\left|c_s^2\right| + 2\sup_{r\in[0,\infty)}\left|\mathbb{E}\left[X_s^{\alpha^T}\right]\right|\sup_{r\in[0,\infty)}\left|b_s + \frac{-\varphi_s^T+\beta_a(s)}{2\beta_{aa}(s)} + 2c_sC_s\right|\right) \\
&\quad < \infty.
\end{aligned}$$

Again, we completely analogously obtain $\left|\mathbb{E}\left[\left(X_t^{\alpha^\infty}\right)^2\right]\right| < \infty$. ■

We have shown that the first moments of $X$ are bounded. The next lemma states that the first and second moment of $X^{\alpha^T}$ converge with exponential speed towards the ones of $X^{\alpha^\infty}$.

**Lemma 4.50**
*Let Assumption 4.36 be fulfilled. Then there are constants* $K_1,\ldots,K_6 > 0$ *independent of* $0 \leqslant t \leqslant T < \infty$ *and* $x_0 \in \mathbb{R}$ *such that*

$$\begin{aligned}
\left|\mathbb{E}\left[X_t^{\alpha^\infty} - X_t^{\alpha^T}\right]\right| &\leqslant K_1 e^{-K_2(T-t)} \\
\mathbb{E}\left[\left(X_t^{\alpha^\infty} - X_t^{\alpha^T}\right)^2\right] &\leqslant K_3 e^{-K_4(T-t)} \qquad \textit{and} \\
\left|\mathbb{E}\left[\left(X_t^{\alpha^\infty}\right)^2 - \left(X_t^{\alpha^T}\right)^2\right]\right| &\leqslant K_5 e^{-K_6(T-t)}.
\end{aligned}$$

*Proof.* Using Corollary 4.22 and Corollary 4.47 we can calculate

$$\begin{aligned}
&\mathbb{E}\left[X_t^{\alpha^\infty} - X_t^{\alpha^T}\right] \\
&\quad = \mathbb{E}\left[\int_0^t B_s\left(X_s^{\alpha^\infty} - X_s^{\alpha^T}\right) - \frac{\varphi_s^\infty - \varphi_s^T + U_s^\infty X_s^{\alpha^\infty} - U_s^T X_s^{\alpha^T} - \beta_{xa}(s)\left(X_s^{\alpha^\infty} - X_s^{\alpha^T}\right)}{2\beta_{aa}(s)}\,\mathrm{d}s\right] \\
&\quad = \int_0^t -\frac{\varphi_s^\infty - \varphi_s^T + \left(U_s^\infty - U_s^T\right)\mathbb{E}\left[X_s^{\alpha^\infty}\right]}{2\beta_{aa}(s)} + \left(B_s + \frac{\beta_{xa}(s) - U_s^T}{2\beta_{aa}(s)}\right)\mathbb{E}\left[X_s^{\alpha^\infty} - X_s^{\alpha^T}\right]\mathrm{d}s.
\end{aligned}$$

Lemma 4.38 furthermore yields

$$\mathbb{E}\left[X_t^{\alpha^\infty} - X_t^{\alpha^T}\right] = \int_0^t \left(-\frac{\varphi_s^\infty - \varphi_s^T + \left(U_s^\infty - U_s^T\right)\mathbb{E}\left[X_s^{\alpha^\infty}\right]}{2\beta_{aa}(s)}\right) e^{\int_s^t\left(B_r + \frac{\beta_{xa}(r) - U_r^T}{2\beta_{aa}(r)}\right)\mathrm{d}r}\,\mathrm{d}s.$$

Using Proposition 4.42, Lemma 4.45, Lemma 4.49 and Assumption 4.36 we get that there are constants $K_1, K_2 > 0$ such that

$$\begin{aligned}\left|\mathbb{E}\left[X_t^{\alpha^\infty} - X_t^{\alpha^T}\right]\right| &\leqslant \int_0^t K_1 K_2 e^{-K_2(T-s)} e^{-K_2(t-s)}\,\mathrm{d}s \\ &= K_1\left(e^{-K_2(T-t)} - e^{-K_2(T+t)}\right) \\ &\leqslant K_1 e^{-K_2(T-t)}. \end{aligned} \tag{4.19}$$

Again using the representation of $\alpha^T$ and $\alpha^\infty$, we obtain

$$\begin{aligned}&\mathbb{E}\left[\left(X_t^{\alpha^\infty} - X_t^{\alpha^T}\right)^2\right] \\ &\quad= \mathbb{E}\left[\int_0^t \left\{2\left(X_s^{\alpha^\infty} - X_s^{\alpha^T}\right)\left(\mu\left(s, X_s^{\alpha^\infty}\right) - \alpha_s^\infty\left(X_s^{\alpha^\infty}\right) - \mu\left(s, X_s^{\alpha^T}\right) + \alpha_s^T\left(X_s^{\alpha^T}\right)\right)\right.\right. \\ &\qquad\qquad \left.\left. + \left(\sigma\left(s, X_s^{\alpha^\infty}\right) - \sigma\left(s, X_s^{\alpha^T}\right)\right)^2\right\}\mathrm{d}s\right] \\ &\quad= \mathbb{E}\Bigg[\int_0^t \Bigg\{\left(X_s^{\alpha^\infty} - X_s^{\alpha^T}\right)\left(\frac{\varphi_s^T - \varphi_s^\infty}{\beta_{aa}(s)} + 2X_s^{\alpha^T}\left(U_s^T - U_s^\infty\right)\right) \\ &\qquad\qquad + \left(X_s^{\alpha^\infty} - X_s^{\alpha^T}\right)^2\left(2B_s + \frac{\beta_{xa}(s)}{\beta_{aa}(s)} - U_s^\infty + C_s^2\right)\Bigg\}\mathrm{d}s\Bigg] \\ &\quad\leqslant \int_0^t \Bigg\{\left|\mathbb{E}\left[X_s^{\alpha^\infty} - X_s^{\alpha^T}\right]\right|\left|\frac{\varphi^T - \varphi_s^\infty}{\beta_{aa}(s)}\right| + \left(\mathbb{E}\left[\left(X_s^{\alpha^\infty} - X_s^{\alpha^T}\right)^2\right] + \mathbb{E}\left[\left(X_s^{\alpha^T}\right)^2\right]\right)\left(U_s^T - U_s^\infty\right) \\ &\qquad\qquad + \mathbb{E}\left[\left(X_s^{\alpha^\infty} - X_s^{\alpha^T}\right)^2\right]\left(2B_s + \frac{\beta_{xa}(s)}{\beta_{aa}(s)} - U_s^\infty + C_s^2\right)\Bigg\}\mathrm{d}s.\end{aligned}$$

With Gronwall's inequality and the estimates in Inequality (4.19), Proposition 4.42, Lemma 4.45 and Lemma 4.49 we derive that there are constants $k_1, \ldots, k_8 > 0$ such that

$$\begin{aligned}\mathbb{E}\left[\left(X_t^{\alpha^\infty} - X_t^{\alpha^T}\right)^2\right] &\leqslant \int_0^t \left[\left|\mathbb{E}\left[X_s^{\alpha^\infty} - X_s^{\alpha^T}\right]\right|\left|\frac{\varphi_s^\infty - \varphi^T}{\beta_{aa}(s)}\right| + \mathbb{E}\left[\left(X_s^{\alpha^T}\right)^2\right]\left(U_s^\infty - U_s^T\right)\right]\mathrm{d}s \\ &\quad\cdot\exp\left(\int_0^t \left\{\left(U_s^\infty - U_s^T\right) + 2B_s + \frac{\beta_{xa}(s)}{\beta_{aa}(s)} - U_s^\infty + C_s^2\right\}\mathrm{d}s\right) \\ &\leqslant \int_0^t \left[k_1 e^{-k_2(T-s)} + k_3 e^{-k_4(T-s)}\right]\mathrm{d}s\,\exp\left(\int_0^t k_5 e^{-k_6(T-s)}\,\mathrm{d}s - k_7 t + k_8\right) \\ &\leqslant \frac{\max\{k_1, k_3\}}{\min\{k_2, k_4\}}\exp\Big(-\min\{k_2, k_4\}(T-t)\Big)\exp\left(-k_7 t + \frac{k_5}{k_6} + k_8\right).\end{aligned}$$

Hence, there are constants $K_3, K_4 > 0$ such that

$$\mathbb{E}\left[\left(X_t^{\alpha^\infty} - X_t^{\alpha^T}\right)^2\right] \leqslant K_3 e^{-K_4(T-t)}.$$

Once more using Corollary 4.41, Corollary 4.47 and Lemma 4.38

$$\begin{aligned}\mathbb{E}\left[\left(X_t^{\alpha^\infty}\right)^2-\left(X_t^{\alpha^T}\right)^2\right] &= \int_0^t \Bigg(-\frac{\left(\varphi_s^\infty-\varphi_s^T\right)\mathbb{E}\left[X_s^{\alpha^\infty}\right]+\left(U_s^\infty-U_s^T\right)\mathbb{E}\left[\left(X_s^{\alpha^\infty}\right)^2\right]}{2\beta_{aa}(s)}\\
&\qquad +2\left(b_s+\frac{\beta_a(s)-\varphi_s^T}{2\beta_{aa}(s)}+c_sC_s\right)\mathbb{E}\left[X_s^{\alpha^\infty}-X_s^{\alpha^T}\right]\\
&\qquad +2\left(B_s+\frac{\beta_{xa}(s)-U_s^T}{2\beta_{aa}(s)}+\frac{C_s^2}{2}\right)\mathbb{E}\left[\left(X_s^{\alpha^\infty}\right)^2-\left(X_s^{\alpha^T}\right)^2\right]\Bigg)\,\mathrm{d}s\\
&= \int_0^t \Bigg(-\frac{\left(\varphi_s^\infty-\varphi_s^T\right)\mathbb{E}\left[X_s^{\alpha^\infty}\right]+\left(U_s^\infty-U_s^T\right)\mathbb{E}\left[\left(X_s^{\alpha^\infty}\right)^2\right]}{2\beta_{aa}(s)}\\
&\qquad +2\left(b_s+\frac{\beta_a(s)-\varphi_s^T}{2\beta_{aa}(s)}+c_sC_s\right)\mathbb{E}\left[X_s^{\alpha^\infty}-X_s^{\alpha^T}\right]\Bigg)\\
&\qquad \cdot e^{\int_s^t 2\left(B_r+\frac{\beta_{xa}(r)-U_r^T}{2\beta_{aa}(r)}+\frac{C_r^2}{2}\right)\mathrm{d}r}\,\mathrm{d}s.\end{aligned}$$

Thus, by Lemma 4.49, the bound on the parameter functions and Proposition 4.42, there are constants $K_5, K_6 > 0$ such that

$$\left|\mathbb{E}\left[\left(X_t^{\alpha^\infty}\right)^2-\left(X_t^{\alpha^T}\right)^2\right]\right| \leqslant \int_0^t K_5K_6e^{-K_6(T-s)}e^{-K_6(t-s)}\,\mathrm{d}s \leqslant K_5e^{-K_6(T-t)}$$

analogously to Inequality (4.19). ■

**Remark 4.51**
Lemma 4.50 implies that $X^{\alpha^T}$ converges for $T\to\infty$ pointwise in probability to $X^{\alpha^\infty}$. Since moreover the decoupling field $u$ of FBSDE (4.8) converges pointwise (see Proposition 4.39, Proposition 4.42 point $(ii)$ and Lemma 4.45), we get that the backward process $Y_t = u(t, X_t^{\alpha^T})$ converges pointwise with exponential speed, too. Furthermore, Corollary 4.40 states that $Z_t = U_t^T\sigma(t, X_t^{\alpha^T})$. Since again all components converge, we get that $Z$ converges pointwise as well.

Summing up, we obtain that the solution $(X^{\alpha^T}, Y, Z)$ of FBSDE (4.8) converges for $T\to\infty$ pointwise to some tuple $(X^{\alpha^\infty}, Y^\infty, Z^\infty)$ which solves for every $\tau \geqslant 0$ and all $t\in[0,\tau]$ the FBSDE

$$\begin{aligned}X_t^{\alpha^\infty} &= x_0+\int_0^t\left[\mu(s,X_s^{\alpha^\infty})-f_\alpha^{-1}\left(s,X_s^{\alpha^\infty},Y_s^\infty\right)\right]\mathrm{d}s+\int_0^t\sigma(s,X_s^{\alpha^\infty})\,\mathrm{d}W_s\\
Y_t^\infty &= Y_\tau^\infty-\int_t^\tau Z_s^\infty\,\mathrm{d}W_s\\
&\quad+\int_t^T\left[\partial_x\mu(s,X_s^{\alpha^\infty})Y_s^\infty+\partial_x\sigma(s,X_s^{\alpha^\infty})Z_s^\infty+\partial_xf\left(s,X_s^{\alpha^\infty},f_\alpha^{-1}\left(s,X_s^{\alpha^\infty},Y_s^\infty\right)\right)\right]\mathrm{d}s.\end{aligned}$$

After knowing that the first and second moments are bounded and converge for a fixed starting value, we now turn to an estimate of the difference of the first and second moment for optimally controlled processes with different starting values.

**Lemma 4.52**
*Let Assumption 4.36 be fulfilled and $x_1, x_2 \in \mathbb{R}$. Then, for the processes $X^{\alpha^\infty,x_1}$, $X^{\alpha^\infty,x_2}$ fulfilling*

$$X_t^{\alpha^\infty,x_1} = x_1 + \int_0^t \left[\mu\left(s, X_s^{\alpha^\infty,x_1}\right) - \alpha_s^\infty\left(X_s^{\alpha^\infty,x_1}\right)\right] \mathrm{d}s + \int_0^t \sigma\left(s, X_s^{\alpha^\infty,x_1}\right) \mathrm{d}W_s$$

*resp.*

$$X_t^{\alpha^\infty,x_2} = x_2 + \int_0^t \left[\mu\left(s, X_s^{\alpha^\infty,x_2}\right) - \alpha_s^\infty\left(X_s^{\alpha^\infty,x_2}\right)\right] \mathrm{d}s + \int_0^t \sigma\left(s, X_s^{\alpha^\infty,x_2}\right) \mathrm{d}W_s,$$

*there exist constants $\delta_1, \delta_2, \delta_3, \delta_4 > 0$ such that*

$$\left|\mathbb{E}\left[X_t^{\alpha^\infty,x_1} - X_t^{\alpha^\infty,x_2}\right]\right| \leqslant |x_1 - x_2|\, \delta_1 e^{-\delta_2 t}$$

*and*

$$\left|\mathbb{E}\left[\left(X_t^{\alpha^\infty,x_1}\right)^2 - \left(X_t^{\alpha^\infty,x_2}\right)^2\right]\right| \leqslant \left(\left|x_1^2 - x_2^2\right| + |x_1 - x_2|\right) \delta_3 e^{-\delta_4 t}.$$

*Proof.* We define $\Xi_t^{x_1,x_2} := X_t^{\alpha^\infty,x_1} - X_t^{\alpha^\infty,x_2}$ and $P_t^{x_1,x_2} := \left(X_t^{\alpha^\infty,x_1}\right)^2 - \left(X_t^{\alpha^\infty,x_2}\right)^2$ for $t \geqslant 0$. By Corollary 4.47 we obtain

$$\Xi_t^{x_1,x_2} = x_1 - x_2 + \int_0^t \left(B_s - \frac{U_s^\infty - \beta_{xa}(s)}{2\beta_{aa}(s)}\right) \Xi_s^{x_1,x_2} \,\mathrm{d}s + \int_0^t C_s \Xi_s^{x_1,x_2} \,\mathrm{d}W_s,$$

$$\begin{aligned} P_t^{x_1,x_2} = x_1^2 - x_2^2 &+ \int_0^t 2\left(b_s + \frac{\varphi_s^\infty - \beta_a(s)}{2\beta_{aa}(s)} + c_s C_s\right) \Xi_s^{x_1,x_2} \,\mathrm{d}s \\ &+ \int_0^t 2\left(B_s - \frac{U_s^\infty - \beta_{xa}(s)}{2\beta_{aa}(s)} + \frac{C_s^2}{2}\right) P_s^{x_1,x_2} \,\mathrm{d}s + \int_0^t \left[c_s \Xi_s^{x_1,x_2} + C_s P_s^{x_1,x_2}\right] \mathrm{d}W_s, \end{aligned}$$

and hence, by Lemma 4.38,

$$\mathbb{E}\left[\Xi_t^{x_1,x_2}\right] = x_1 - x_2 + \int_0^t \left(B_s - \frac{U_s^\infty - \beta_{xa}(s)}{2\beta_{aa}(s)}\right) \mathbb{E}\left[\Xi_s^{x_1,x_2}\right] \mathrm{d}s = (x_1 - x_2)\, e^{\int_0^t B_s - \frac{U_s^\infty - \beta_{xa}(s)}{2\beta_{aa}(s)} \,\mathrm{d}s},$$

$$\begin{aligned}\mathbb{E}\left[P_t^{x_1,x_2}\right] &= x_1^2 - x_2^2 + \int_0^t 2\left(b_s + \frac{\varphi_s^\infty - \beta_a(s)}{2\beta_{aa}(s)} + c_sC_s\right)\mathbb{E}\left[\Xi_s^{x_1,x_2}\right]\mathrm{d}s \\ &\quad + \int_0^t 2\left(B_s - \frac{U_s^\infty - \beta_{xa}(s)}{2\beta_{aa}(s)} + \frac{C_s^2}{2}\right)\mathbb{E}\left[P_s^{x_1,x_2}\right]\mathrm{d}s \\ &= \left(x_1^2 - x_2^2\right)e^{\int_0^t 2\left(B_s - \frac{U_s^\infty - \beta_{xa}(s)}{2\beta_{aa}(s)} + \frac{C_s^2}{2}\right)\mathrm{d}s} \\ &\quad + \int_0^t 2\left(b_s + \frac{\varphi_s^\infty - \beta_a(s)}{2\beta_{aa}(s)} + c_sC_s\right)\mathbb{E}\left[\Xi_s^{x_1,x_2}\right]e^{\int_s^t 2\left(B_r - \frac{U_r^\infty - \beta_{xa}(r)}{2\beta_{aa}(r)} + \frac{C_r^2}{2}\right)\mathrm{d}r}\mathrm{d}s.\end{aligned}$$

Using Proposition 4.42 we obtain that there are constants $\delta_1, \delta_2 > 0$ such that

$$\left|\mathbb{E}\left[\Xi_t^{x_1,x_2}\right]\right| \leqslant |x_1 - x_2|\,\delta_1 e^{-\delta_2 t}.$$

Therefore and by Lemma 4.45 and Assumption 4.36 we also get that there exist other constants $\delta_3, \delta_4, \delta_5 > 0$ such that

$$\begin{aligned}\left|\mathbb{E}\left[P_t^{x_1,x_2}\right]\right| &\leqslant \left|x_1^2 - x_2^2\right|\delta_1^2 e^{-2\delta_2 t} + \int_0^t 2\delta_5\,|x_1 - x_2|\,\delta_1 e^{-\delta_2 s}\delta_1^2 e^{-2\delta_2(t-s)}\,\mathrm{d}s \\ &\leqslant \left|x_1^2 - x_2^2\right|\delta_1^2 e^{-2\delta_2 t} + \int_0^t 2\delta_5\,|x_1 - x_2|\,\delta_1^3 e^{-\delta_2 t}\,\mathrm{d}s \\ &= \delta_1^2 e^{-\delta_2 t}\left(\left|x_1^2 - x_2^2\right|e^{-\delta_2 t} + |x_1 - x_2|\,2\delta_5\delta_1^2 t\right) \\ &\leqslant \left(\left|x_1^2 - x_2^2\right| + |x_1 - x_2|\right)\delta_3 e^{-\delta_4 t}.\end{aligned}$$

■

Now we have all necessary tools to show that $\alpha^\infty$ is indeed an optimal ergodic control.

**Theorem 4.53**
*Let Assumption 4.36 be fulfilled. Then the optimal ergodic costs are*

$$\inf_{\alpha\in\mathcal{A}}\limsup_{T\to\infty}\frac{1}{T}J(T,x_0,\alpha) = \limsup_{T\to\infty}\frac{1}{T}J(T,x_0,\alpha^\infty) = \limsup_{T\to\infty}\frac{1}{T}J(T,x_0,\alpha^T) =: \eta \in \mathbb{R}$$

*for all $x_0 \in \mathbb{R}$. In particular $\eta$ is a constant and does not depend on the starting value $x_0 \in \mathbb{R}$ of the controlled process. Likewise, the minimal ergodic costs are*

$$\inf_{\alpha\in\mathcal{A}}\liminf_{T\to\infty}\frac{1}{T}J(T,x_0,\alpha) = \liminf_{T\to\infty}\frac{1}{T}J(T,x_0,\alpha^\infty) = \liminf_{T\to\infty}\frac{1}{T}J(T,x_0,\alpha^T) =: \check{\eta} \in \mathbb{R}$$

*for all $x_0 \in \mathbb{R}$.*

*Furthermore, there is a constant $K > 0$ such that for all $x_1, x_2 \in \mathbb{R}$ and $T > 0$*

$$\begin{aligned}&\left|\frac{1}{T}J(T,x_1,\alpha^\infty) - \frac{1}{T}J(T,x_1,\alpha^T)\right| \leqslant \frac{K}{T} \qquad \textit{and} \\ &\left|\frac{1}{T}J(T,x_1,\alpha^\infty) - \frac{1}{T}J(T,x_2,\alpha^\infty)\right| \leqslant \left(\left|x_1^2 - x_2^2\right| + |x_1 - x_2|\right)\frac{K}{T}.\end{aligned}$$

*Proof.* Since for each time horizon $T > 0$ and $x_0 \in \mathbb{R}$ the control $\alpha^T$ is optimal, we know that for the optimal ergodic cost holds

$$\inf_{\alpha\in\mathcal{A}} \limsup_{T\to\infty} \frac{1}{T} J(T,x,\alpha) \geqslant \limsup_{T\to\infty} \inf_{\alpha\in\mathcal{A}} \frac{1}{T} J(T,x,\alpha) = \limsup_{T\to\infty} \frac{1}{T} J(T,x,\alpha^T).$$

Thus, $\limsup_{T\to\infty} \frac{1}{T} J(T,x_0,\alpha^T)$ is a lower bound for the optimal ergodic cost. Our aim is to show that this lower bound is reached by the ergodic costs of the control $\alpha^\infty$. To this end, remember Corollary 4.22 and Corollary 4.47 and observe that

$$\begin{aligned}
&\limsup_{T\to\infty} \left| \frac{1}{T} J(T,x_0,\alpha^\infty) - \frac{1}{T} J(T,x_0,\alpha^T) \right| \\
&= \limsup_{T\to\infty} \left| \frac{1}{T} \mathbb{E} \int_0^T \left[ f\left(s, X_s^{\alpha^\infty}, \alpha_s^\infty\left(X_s^{\alpha^\infty}\right)\right) - f\left(s, X_s^{\alpha^T}, \alpha_s^T\left(X_s^{\alpha^T}\right)\right)\right] \mathrm{d}s \right| \\
&= \limsup_{T\to\infty} \left| \frac{1}{T} \mathbb{E} \int_0^T \left[ \beta_{xx}(s) \left( \left(X_s^{\alpha^\infty}\right)^2 - \left(X_s^{\alpha^T}\right)^2 \right) + \beta_x(s) \left( X_s^{\alpha^\infty} - X_s^{\alpha^T} \right) \right.\right. \\
&\qquad + \beta_{xa}(s) \left( X_s^{\alpha^\infty} \left( \frac{\varphi_s^\infty - \beta_a(s) + (U_s^\infty - \beta_{xa}(s)) X_s^{\alpha^\infty}}{2\beta_{aa}(s)} \right) \right. \\
&\qquad\qquad \left. - \left( X_s^{\alpha^T} \frac{\varphi_s^T - \beta_a(s) + \left(U_s^T - \beta_{xa}(s)\right) X_s^{\alpha^T}}{2\beta_{aa}(s)} \right) \right) \\
&\qquad + \beta_a(s) \left( \frac{\varphi_s^\infty - \beta_a(s) + (U_s^\infty - \beta_{xa}(s)) X_s^{\alpha^\infty}}{2\beta_{aa}(s)} \right. \\
&\qquad\qquad \left. - \frac{\varphi_s^T - \beta_a(s) + \left(U_s^T - \beta_{xa}(s)\right) X_s^{\alpha^T}}{2\beta_{aa}(s)} \right) \\
&\qquad + \beta_{aa}(s) \left( \left( \frac{\varphi_s^\infty - \beta_a(s) + (U_s^\infty - \beta_{xa}(s)) X_s^{\alpha^\infty}}{2\beta_{aa}(s)} \right)^2 \right. \\
&\qquad\qquad \left.\left.\left. - \left( \frac{\varphi_s^T - \beta_a(s) + \left(U_s^T - \beta_{xa}(s)\right) X_s^{\alpha^T}}{2\beta_{aa}(s)} \right)^2 \right) \right] \mathrm{d}s \right| \\
&= \limsup_{T\to\infty} \left| \frac{1}{T} \int_0^T \left( \mathbb{E}\left[ X_s^{\alpha^\infty} - X_s^{\alpha^T} \right] D_s^1 + \mathbb{E}\left[ \left(X_s^{\alpha^\infty}\right)^2 - \left(X_s^{\alpha^T}\right)^2 \right] D_s^2 \right.\right. \\
&\qquad + \mathbb{E}\left[ \varphi_s^\infty X_s^{\alpha^\infty} - \varphi_s^T X_s^{\alpha^T} \right] D_s^3 + \mathbb{E}\left[ U_s^\infty X_s^{\alpha^\infty} - U_s^T X_s^{\alpha^T} \right] D_s^4 \\
&\qquad + \mathbb{E}\left[ \varphi_s^\infty U_s^\infty X_s^{\alpha^\infty} - \varphi_s^T U_s^T X_s^{\alpha^T} \right] D_s^5 + \mathbb{E}\left[ U_s^\infty \left(X_s^{\alpha^\infty}\right)^2 - U_s^T \left(X_s^{\alpha^T}\right)^2 \right] D_s^6 \\
&\qquad + \mathbb{E}\left[ \left(U_s^\infty X_s^{\alpha^\infty}\right)^2 - \left(U_s^T X_s^{\alpha^T}\right)^2 \right] D_s^7 + \mathbb{E}\left[ \varphi_s^\infty - \varphi_s^T \right] D_s^8 \\
&\qquad \left.\left. + \mathbb{E}\left[ (\varphi_s^\infty)^2 - \left(\varphi_s^T\right)^2 \right] D_s^9 \right) \mathrm{d}s \right|
\end{aligned}$$

for some bounded and deterministic processes $D^i$ for $i = 1, \ldots, 9$ which are independent of $T$. Using that $ab - cd = (a-c)d + a(b-d)$ and likewise $abc - def = (a-d)ef + a(b-e)f + ab(c-f)$

we can rewrite this as

$$
\begin{aligned}
&\limsup_{T\to\infty}\left|\frac{1}{T}J(T,x_0,\alpha^\infty)-\frac{1}{T}J(T,x_0,\alpha^T)\right|\\
&=\limsup_{T\to\infty}\left|\frac{1}{T}\int_0^T\Bigg(\mathbb{E}\left[X_s^{\alpha^\infty}-X_s^{\alpha^T}\right]\left(D_s^1+\varphi_s^\infty D_s^3+U_s^\infty D_s^4+\varphi_s^\infty U_s^\infty D_s^5\right)\right.\\
&\qquad\qquad+\mathbb{E}\left[\left(X_s^{\alpha^\infty}\right)^2-\left(X_s^{\alpha^T}\right)^2\right]\left(D_s^2+U_s^\infty D_s^6+\left(U_s^\infty\right)^2D_s^7\right)\\
&\qquad\qquad+\left(\varphi_s^\infty-\varphi_s^T\right)\left(\mathbb{E}\left[X_s^{\alpha^T}\right]D_s^3+U_s^T\mathbb{E}\left[X_s^{\alpha^T}\right]D_s^5+D_s^8+\left(\varphi_s^\infty+\varphi_s^T\right)D_s^9\right)\\
&\qquad\qquad+\left(U_s^\infty-U_s^T\right)\Bigg(\mathbb{E}\left[X_s^{\alpha^T}\right]D_s^4+\varphi_s^\infty\mathbb{E}\left[X_s^{\alpha^T}\right]D_s^5\\
&\qquad\qquad\qquad\left.+\mathbb{E}\left[\left(X_s^{\alpha^T}\right)^2\right]D_s^6+\left(U_s^T+U_s^\infty\right)\mathbb{E}\left[\left(X_s^{\alpha^T}\right)^2\right]D_s^7\Bigg)\Bigg)\,\mathrm{d}s\right|.
\end{aligned}
\tag{4.20}
$$

By Lemma 4.50 we have for some constants $K_1,K_2,K_3,K_4>0$ and every $T>0$

$$\frac{1}{T}\int_0^T\left|\mathbb{E}\left[X_t^{\alpha^\infty}-X_t^{\alpha^T}\right]\right|\mathrm{d}t\leqslant\frac{1}{T}\int_0^T K_1e^{-K_2(T-t)}\,\mathrm{d}t=\frac{1}{T}\frac{K_1}{K_2}\left(1-e^{-K_2T}\right)\leqslant\frac{1}{T}\frac{K_1}{K_2}\tag{4.21}$$

and

$$\frac{1}{T}\int_0^T\left|\mathbb{E}\left[\left(X_t^{\alpha^\infty}\right)^2-\left(X_t^{\alpha^T}\right)^2\right]\right|\mathrm{d}t\leqslant\frac{1}{T}\int_0^T K_3e^{-K_4(T-t)}\,\mathrm{d}t\leqslant\frac{1}{T}\frac{K_3}{K_4}.\tag{4.22}$$

Likewise, Proposition 4.42 and Lemma 4.45 imply that there are constants $K_5,K_6,K_7,K_8>0$ such that for all $T>0$

$$\frac{1}{T}\int_0^T\left|U_t^\infty-U_t^T\right|\mathrm{d}t\leqslant\frac{1}{T}\int_0^T K_5e^{-K_6(T-t)}\,\mathrm{d}t\leqslant\frac{1}{T}\frac{K_5}{K_6}\tag{4.23}$$

and

$$\frac{1}{T}\int_0^T\left|\varphi_t^\infty-\varphi_t^T\right|\mathrm{d}t\leqslant\frac{1}{T}\int_0^T K_7e^{-K_8(T-t)}\,\mathrm{d}t\leqslant\frac{1}{T}\frac{K_7}{K_8}.\tag{4.24}$$

Note that $D^i$ for $i=1,\dots,9$ are bounded and by Proposition 4.42, Lemma 4.45 and Lemma 4.49 the processes $U^T$, $U^\infty$, $\varphi^T$, $\varphi^\infty$, $\mathbb{E}\left[X_\cdot^{\alpha^T}\right]$ and $\mathbb{E}\left[\left(X_\cdot^{\alpha^T}\right)^2\right]$ are bounded, too. Hence, applying the estimates of the integrals in the equations (4.21), (4.22), (4.23), (4.24) to Equation (4.20) we obtain

$$\left|\frac{1}{T}J(T,x_0,\alpha^\infty)-\frac{1}{T}J(T,x_0,\alpha^T)\right|\leqslant\frac{K}{T}$$

for some constant $K>0$. Thus,

$$\limsup_{T\to\infty}\left|\frac{1}{T}J(T,x_0,\alpha^\infty)-\frac{1}{T}J(T,x_0,\alpha^T)\right|=0$$

and the ergodic control $\alpha^\infty$ yields ergodic costs equal to the lower bound

$$\limsup_{T\to\infty}\inf_{\alpha\in\mathcal{A}}\frac{1}{T}J(T,x_0,\alpha)=\limsup_{T\to\infty}\frac{1}{T}J(T,x_0,\alpha^\infty)$$

implying that the costs of $\alpha^\infty$ are optimal. Using the same arguments we also obtain the minimal ergodic costs.

Now let $x_1, x_2 \in \mathbb{R}$ be two starting values. For the corresponding processes $X^{\alpha^\infty,x_1}$ and $X^{\alpha^\infty,x_2}$ we obtain by Lemma 4.52 that there exist constants $\delta_1, \delta_2, \delta_3, \delta_4 > 0$ such that

$$\left|\mathbb{E}\left[X_t^{\alpha^\infty,x_1} - X_t^{\alpha^\infty,x_2}\right]\right| \leqslant |x_1 - x_2|\, e^{-\delta_1 t + \delta_2}$$

and

$$\left|\mathbb{E}\left[\left(X_t^{\alpha^\infty,x_1}\right)^2 - \left(X_t^{\alpha^\infty,x_2}\right)^2\right]\right| \leqslant \left(\left|x_1^2 - x_2^2\right| + |x_1 - x_2|\right)\delta_3 e^{-\delta_4 t}.$$

Due to the structure of $f$ and $\alpha^\infty$ (see Assumption 4.17 and Corollary 4.47), we know that there are constants $K_1, K_2 > 0$ such that we obtain for any $T \geqslant 0$

$$\begin{aligned}
&\left|J\left(T, x_1, \alpha^\infty\right) - J\left(T, x_2, \alpha^\infty\right)\right| \\
&\quad \leqslant \int_0^T K_1 \left|\mathbb{E}\left[\left(X_s^{\alpha^\infty,x_1}\right)^2 - \left(X_s^{\alpha^\infty,x_2}\right)^2\right]\right| + K_2 \left|\mathbb{E}\left[X_s^{\alpha^\infty,x_1} - X_s^{\alpha^\infty,x_2}\right]\right| \mathrm{d}s \\
&\quad \leqslant K_1 \frac{\delta_3}{\delta_4}\left(\left|x_1^2 - x_2^2\right| + |x_1 - x_2|\right)\left(1 - e^{-\delta_4 T}\right) + K_2 \frac{e^{\delta_2}}{\delta_1}|x_1 - x_2|\left(1 - e^{-\delta_2 T}\right) \\
&\quad \leqslant K_1 \frac{\delta_3}{\delta_4}\left|x_1^2 - x_2^2\right| + \left(K_1 \frac{\delta_3}{\delta_4} + K_2 \frac{e^{\delta_2}}{\delta_1}\right)|x_1 - x_2|.
\end{aligned}$$

Hence, the difference of ergodic costs for those two starting values can be estimated by

$$\begin{aligned}
&\limsup_{T\to\infty}\left|\frac{1}{T}J\left(T, x_1, \alpha^\infty\right) - \frac{1}{T}J\left(T, x_2, \alpha^\infty\right)\right| \\
&\qquad \leqslant \limsup_{T\to\infty}\frac{1}{T}\left(K_1\frac{\delta_3}{\delta_4}\left|x_1^2 - x_2^2\right| + \left(K_1\frac{\delta_3}{\delta_4} + K_2\frac{e^{\delta_2}}{\delta_1}\right)|x_1 - x_2|\right) \\
&\qquad = 0,
\end{aligned}$$

giving us that for all starting values the optimal ergodic costs are equal, making them a constant. ■

**Remark 4.54**
Note that in general the optimal ergodic costs $\eta$ and the minimal ergodic costs $\check{\eta}$ are not necessarily equal. It only holds true that

$$\check{\eta} = \inf_{\alpha\in\mathcal{A}}\liminf_{T\to\infty}\frac{1}{T}J(T, x_0, \alpha) \leqslant \inf_{\alpha\in\mathcal{A}}\limsup_{T\to\infty}\frac{1}{T}J(T, x_0, \alpha) = \eta.$$

In the following we examine the properties of the optimal control more closely and draw some connections to the Hamilton-Jacobi-Bellman approach.

**Lemma 4.55**
*Let Assumption 4.36 be fulfilled. Define for $T \in (0,\infty) \cup \{\infty\}$ the function*

$$\Theta^T(t,x) := \frac{1}{2}U_t^T \cdot x^2 + \varphi_t^T \cdot x - \int_0^t \varphi_s^T b_s + U_s^T \frac{c_s^2}{2} + \beta_0(s) - \frac{\left(\varphi_s^T - \beta_a(s)\right)^2}{4\beta_{aa}(s)}\,\mathrm{d}s$$

*for all $t \in [0,T] \cap [0,\infty)$ and $x \in \mathbb{R}$. Then for $T \in (0,\infty]$ and $x \in \mathbb{R}$*

$$-f\left(t, x, \alpha_t^T(x)\right) = \partial_t \Theta^T(t,x) + \left(\mu(t,x) - \alpha_t^T(x)\right)\partial_x \Theta^T(t,x) + \frac{1}{2}\sigma^2(t,x)\partial_{xx}\Theta^T(t,x).$$

*Proof.* Remember that by Proposition 4.39 and Lemma 4.45

$$\partial_t\varphi_t^T = -\left[U_t^T\left(b_t + c_tC_t + \frac{\beta_a(t)}{2\beta_{aa}(t)}\right) - \frac{\beta_a(t)\beta_{xa}(t)}{2\beta_{aa}(t)} + \beta_x(t)\right] - \left[B_t + \frac{\beta_{xa}(t) - U_t^T}{2\beta_{aa}(t)}\right]\varphi_t^T$$

and by Lemma 4.37 and Proposition 4.42

$$\partial_t U_t^T = \frac{\left(U_t^T\right)^2}{2\beta_{aa}(t)} - U_t^T\left(2B_t + \frac{\beta_{xa}(t)}{\beta_{aa}(t)} + C_t^2\right) - 2\beta_{xx}(t) + \frac{\beta_{xa}^2(t)}{2\beta_{aa}(t)}.$$

Hence, with Corollary 4.22 and Corollary 4.47

$$\begin{aligned}
&\partial_t\Theta^T(t,x) + \left(\mu(t,x) - \alpha_t^T(x)\right)\partial_x\Theta^T(t,x) + \frac{1}{2}\sigma^2(t,x)\partial_{xx}\Theta^T(t,x)\\
&= \frac{1}{2}x^2\left(\frac{\left(U_t^T\right)^2}{2\beta_{aa}(t)} - U_t^T\left(2B_t + \frac{\beta_{xa}(t)}{\beta_{aa}(t)} + C_t^2\right) - 2\beta_{xx}(t) + \frac{\beta_{xa}^2(t)}{2\beta_{aa}(t)}\right)\\
&\quad - x\left(\left[U_t^T\left(b_t + c_tC_t + \frac{\beta_a(t)}{2\beta_{aa}(t)}\right) - \frac{\beta_a(t)\beta_{xa}(t)}{2\beta_{aa}(t)} + \beta_x(t)\right] + \left[B_t + \frac{\beta_{xa}(t) - U_t^T}{2\beta_{aa}(t)}\right]\varphi_t^T\right)\\
&\quad - \varphi_t^T b_t - U_t^T\frac{c_t^2}{2} - \beta_0(t) + \left(\varphi_t^T - \beta_a(t)\right)\left(\frac{\varphi_t^T - \beta_a(t)}{2\beta_{aa}(t)}\right) - \beta_{aa}(t)\left(\frac{\varphi_t^T - \beta_a(t)}{2\beta_{aa}(t)}\right)^2\\
&\quad + \left(b_t + B_tx - \frac{\varphi_t^T - \beta_a(t)}{2\beta_{aa}(t)} - \frac{U_t^T - \beta_{xa}(t)}{2\beta_{aa}(t)}x\right)\left(\varphi_t^T + U_t^Tx\right) + \frac{1}{2}\left(c_t + C_tx\right)^2U_t^T\\
&= x^2\left[-\frac{\left(U_t^T\right)^2}{4\beta_{aa}(t)} - \beta_{xx}(t) + \frac{\beta_{xa}^2(t)}{4\beta_{aa}(t)}\right]\\
&\quad + x\left[-\left(U_t^T\frac{\beta_a(t)}{2\beta_{aa}(t)} - \frac{\beta_a(t)\beta_{xa}(t)}{2\beta_{aa}(t)} + \beta_x(t)\right) + U_t^T\left(-\frac{\varphi_t^T - \beta_a(t)}{2\beta_{aa}(t)}\right)\right]\\
&\quad + \left[-\beta_0(t) - \beta_a(t)\left(\frac{\varphi_t^T - \beta_a(t)}{2\beta_{aa}(t)}\right) - \beta_{aa}(t)\left(\frac{\varphi_t^T - \beta_a(t)}{2\beta_{aa}(t)}\right)^2\right]
\end{aligned}$$

On the other hand

$$\begin{aligned}
&- f\left(t, x, \alpha_t^T(x)\right)\\
&= -\Bigg\{\beta_0(t) + \beta_{xx}(t)x^2 + \beta_x(t)x + \beta_{xa}(t)x\left(\frac{\varphi_t^T - \beta_a(t) + \left(U_t^T - \beta_{xa}(t)\right)x}{2\beta_{aa}(t)}\right)\\
&\quad + \beta_{aa}(t)\left(\left(\frac{\varphi_t^T - \beta_a(t)}{2\beta_{aa}(t)}\right)^2 + 2\frac{\left(\varphi_t^T - \beta_a(t)\right)\left(U_t^T - \beta_{xa}(t)\right)x}{\left(2\beta_{aa}(t)\right)^2} + \left(\frac{\left(U_t^T - \beta_{xa}(t)\right)x}{2\beta_{aa}(t)}\right)^2\right)\\
&\quad + \beta_a(t)\frac{\varphi_t^T - \beta_a(t) + \left(U_t^T - \beta_{xa}(t)\right)x}{2\beta_{aa}(t)}\Bigg\}\\
&= x^2\left[-\frac{\left(U_t^T\right)^2}{4\beta_{aa}(t)} - \beta_{xx}(t) + \frac{\beta_{xa}^2(t)}{4\beta_{aa}(t)}\right]\\
&\quad + x\left[-\left(U_t^T\frac{\beta_a(t)}{2\beta_{aa}(t)} - \frac{\beta_a(t)\beta_{xa}(t)}{2\beta_{aa}(t)} + \beta_x(t)\right) + U_t^T\left(-\frac{\varphi_t^T - \beta_a(t)}{2\beta_{aa}(t)}\right)\right]\\
&\quad + \left[-\beta_0(t) - \beta_a(t)\left(\frac{\varphi_t^T - \beta_a(t)}{2\beta_{aa}(t)}\right) - \beta_{aa}(t)\left(\frac{\varphi_t^T - \beta_a(t)}{2\beta_{aa}(t)}\right)^2\right].
\end{aligned}$$

Thus, we obtain that

$$\partial_t \Theta^T(t,x) + \left(\mu(t,x) - \alpha_t^T(x)\right)\partial_x \Theta^T(t,x) + \frac{1}{2}\sigma^2(t,x)\partial_{xx}\Theta^T(t,x) = -f\left(t,x,\alpha_t^T(x)\right).$$

■

**Proposition 4.56**
*Let Assumption 4.36 be fulfilled and $T \in (0,\infty) \cup \{\infty\}$. Then $\Theta^T$, as defined in Lemma 4.55, solves the HJB-equation*

$$0 = \inf_{a \in \mathbb{R}} \left\{ \partial_t \Theta^T(t,x) + (\mu(t,x) - a)\partial_x \Theta^T(t,x) + \frac{1}{2}\sigma^2(t,x)\partial_{xx}\Theta^T(t,x) + f(t,x,a) \right\}$$

*for all $t \in [0,T] \cap [0,\infty)$ and $x \in \mathbb{R}$.*

*Proof.* Note that $f$ is strictly convex in $a$ and hence the whole function inside the infimum is strictly convex in $a$. To find the global minimum it suffices to determine for which $a$ the derivative equals $0$. To this end observe that

$$\begin{aligned}
&\partial_a \left[ \partial_t \Theta^T(t,x) + (\mu(t,x) - a)\partial_x \Theta^T(t,x) + \frac{1}{2}\sigma^2(t,x)\partial_{xx}\Theta^T(t,x) + f(t,x,a) \right]\Bigg|_{a=\alpha_t^T(x)} \\
&\quad = \left[-\partial_x \Theta^T(t,x) + \beta_a(t) + \beta_{xa}(t)x + 2\beta_{aa}(t)a\right]\Big|_{a=\alpha_t^T(x)} \\
&\quad = -U_t^T x - \varphi_t^T + \beta_a(t) + \beta_{xa}(t)x + 2\beta_{aa}(t)\alpha_t^T(x) \\
&\quad = 0
\end{aligned}$$

by Corollary 4.41 if $T \in (0,T)$ or by Corollary 4.47 if $T = \infty$. Since furthermore Lemma 4.55 yields that

$$0 = \left[ \partial_t \Theta^T(t,x) + (\mu(t,x) - a)\partial_x \Theta^T(t,x) + \frac{1}{2}\sigma^2(t,x)\partial_{xx}\Theta^T(t,x) + f(t,x,a) \right]\Bigg|_{a=\alpha_t^T(x)},$$

we know that $\alpha^T$ is a minimizer and hence $\Theta^T$ solves the HJB-equation. ■

Now we have a function that solves the HJB-equation. The standard HJB-theory suggests that the value function also solves HJB-equation. Hence, we already have a guess for the value function. In the next theorem we give for a finite time horizon an explicit representation of the value function and show that the value function differs from $\Theta^T$ only by a constant.

**Theorem 4.57**
*Let Assumption 4.36 be fulfilled and $T \in (0,\infty)$. Then, for $0 \leqslant t \leqslant T$ and $x \in \mathbb{R}$, the value function is given by*

$$\begin{aligned} V(t,T,x) &:= \inf_{\alpha\in\mathcal{A}^T} \mathbb{E}\int_t^T f\left(s, X_s^{\alpha,t,x}, \alpha_s\right) \mathrm{d}s \\ &= \frac{1}{2}U_t^T \cdot x^2 + \varphi_t^T \cdot x + \int_t^T \left[\varphi_s^T b_s + U_s^T \frac{c_s^2}{2} + \beta_0(s) - \frac{\left(\varphi_s^T - \beta_a(s)\right)^2}{4\beta_{aa}(s)}\right] \mathrm{d}s, \end{aligned}$$

*where $X^{\alpha,t,x}$ is the process controlled by $\alpha$ and $X_t^{\alpha,t,x} = x$. In particular*

$$V(0,T,x) = \inf_{\alpha\in\mathcal{A}^T} J(T,x,\alpha).$$

*Furthermore, for all $(t,x) \in [0,T]\times\mathbb{R}$ the value function $V$ solves the HJB-equation*

$$0 = \inf_{a\in\mathbb{R}} \left\{ \partial_t V(t,T,x) + (\mu(t,x) - a)\partial_x V(t,T,x) + \frac{1}{2}\sigma^2(t,x)\partial_{xx}V(t,T,x) + f(t,x,a) \right\}.$$

*Proof.* First, note that $\varphi_t^T$ and $U_t^T$ are deterministic and depend only on the parameter functions from time $t$ until $T$. Hence, it is straightforward to see that shifting time together with the parameter functions does not change $\varphi^T$ or $U^T$. Thus, we can also conclude that the optimal control $\alpha^T$ does not change by a timeshift of the problem neither. Therefore, Corollary 4.22 still yields the optimal control regardless whether we start at time 0 or at time $t$. Using this, we have by Corollary 4.41, Lemma 4.55, Itô's formula and the fact $U_T^T = \varphi_T^T = 0$ that

$$\begin{aligned} V(t,T,x) &= \mathbb{E}\left[\int_t^T f\left(s, X_s^{\alpha^T,t,x}, \alpha_s^T\left(X_s^{\alpha^T,t,x}\right)\right) \mathrm{d}s\right] \\ &= \mathbb{E}\Bigg[-\int_t^T \Bigg(\partial_t\Theta^T\left(s, X_s^{\alpha^T,t,x}\right) + \left(\mu\left(s, X_s^{\alpha^T,t,x}\right) - \alpha_s^T\left(X_s^{\alpha^T,t,x}\right)\right)\partial_x\Theta\left(s, X_s^{\alpha^T,t,x}\right) \\ &\qquad\qquad + \frac{1}{2}\sigma^2\left(s, X_s^{\alpha^T,t,x}\right)\partial_{xx}\Theta\left(s, X_s^{\alpha^T,t,x}\right)\Bigg) \mathrm{d}s\Bigg] \\ &= \mathbb{E}\left[\Theta^T\left(t, X_t^{\alpha^T,t,x}\right) - \Theta^T\left(T, X_T^{\alpha^T,t,x}\right)\right] \\ &= \frac{1}{2}U_t^T \cdot x^2 + \varphi_t^T \cdot x + \int_t^T \left(\varphi_s^T b_s + U_s^T \frac{c_s^2}{2} + \beta_0(s) - \frac{\left(\varphi_s^T - \beta_a(s)\right)^2}{4\beta_{aa}(s)}\right) \mathrm{d}s, \end{aligned}$$

which proves the first result. Now, we turn to the HJB-equation. Note that

$$V(t,T,x) - \Theta^T(t,x) = \int_0^T \left[\varphi_s^T b_s + U_s^T \frac{c_s^2}{2} + \beta_0(s) - \frac{\left(\varphi_s^T - \beta_a(s)\right)^2}{4\beta_{aa}(s)}\right] \mathrm{d}s,$$

which does not depend on $t$ or $x$. Therefore, the derivatives with respect to $t$ and $x$ of the functions $V$ and $\Theta^T$ are the same. Since only those derivatives appear in the HJB-equation we get that $V$ solves it, exactly as $\Theta^T$ does by Propositon 4.56. ■

Theorem 4.57 allows us to give a more explicit formula for the optimal ergodic cost.

**Corollary 4.58**
*Let Assumption 4.36 be fulfilled. Then, for any $x_0 \in \mathbb{R}$, the optimal ergodic cost $\eta$ equals*

$$\begin{aligned}\eta &= \limsup_{T\to\infty} \frac{1}{T} V(0,T,0) \\ &= \limsup_{T\to\infty} \frac{1}{T}\int_0^T \left[ \varphi_s^T b_s + U_s^T \frac{c_s^2}{2} + \beta_0(s) - \frac{\left(\varphi_s^T - \beta_a(s)\right)^2}{4\beta_{aa}(s)} \right] \mathrm{d}s \\ &= \limsup_{T\to\infty} \frac{1}{T}\int_0^T \left[ \varphi_s^\infty b_s + U_s^\infty \frac{c_s^2}{2} + \beta_0(s) - \frac{(\varphi_s^\infty - \beta_a(s))^2}{4\beta_{aa}(s)} \right] \mathrm{d}s\end{aligned}$$

*and the minimal ergodic cost $\check{\eta}$ likewise equals*

$$\begin{aligned}\check{\eta} &= \liminf_{T\to\infty} \frac{1}{T} V(0,T,0) \\ &= \liminf_{T\to\infty} \frac{1}{T}\int_0^T \left[ \varphi_s^T b_s + U_s^T \frac{c_s^2}{2} + \beta_0(s) - \frac{\left(\varphi_s^T - \beta_a(s)\right)^2}{4\beta_{aa}(s)} \right] \mathrm{d}s \\ &= \liminf_{T\to\infty} \frac{1}{T}\int_0^T \left[ \varphi_s^\infty b_s + U_s^\infty \frac{c_s^2}{2} + \beta_0(s) - \frac{(\varphi_s^\infty - \beta_a(s))^2}{4\beta_{aa}(s)} \right] \mathrm{d}s.\end{aligned}$$

*Proof.* By Theorem 4.53 we know that each starting value gives the same ergodic cost. Hence, we can choose without loss of generality $0$ as starting value. Again, Theorem 4.53 yields that

$$\eta = \limsup_{T\to\infty} \frac{1}{T} J(T,0,\alpha^T)$$

and therefore, by Theorem 4.57 we obtain

$$\begin{aligned}\eta &= \limsup_{T\to\infty} \frac{1}{T} V(0,T,0) \\ &= \limsup_{T\to\infty} \frac{1}{T}\int_0^T \left[ \varphi_s^T b_s + U_s^T \frac{c_s^2}{2} + \beta_0(s) - \frac{\left(\varphi_s^T - \beta_a(s)\right)^2}{4\beta_{aa}(s)} \right] \mathrm{d}s.\end{aligned}$$

Also, $U^T, U^\infty, \varphi^T, \varphi^\infty$ and all parameter functions are bounded and by Proposition 4.42 and Lemma 4.45 there are positive constants $K_1, K_2$ such that $|U_t^\infty - U_t^T| \leqslant K_1 \exp(-K_2(T-t))$ and $|\varphi_t^\infty - \varphi_t^T| \leqslant K_1 \exp(-K_2(T-t))$ for all $0 \leqslant t \leqslant T < \infty$. Hence, we obtain that

$$\begin{aligned}\lim_{T\to\infty} \Bigg| \frac{1}{T}\int_0^T &\left[ \varphi_s^T b_s + U_s^T \frac{c_s^2}{2} + \beta_0(s) - \frac{\left(\varphi_s^T - \beta_a(s)\right)^2}{4\beta_{aa}(s)} \right] \mathrm{d}s \\ &- \frac{1}{T}\int_0^T \left[ \varphi_s^\infty b_s + U_s^\infty \frac{c_s^2}{2} + \beta_0(s) - \frac{(\varphi_s^\infty - \beta_a(s))^2}{4\beta_{aa}(s)} \right] \mathrm{d}s \Bigg| = 0\end{aligned}$$

and therefore,

$$\begin{aligned}\limsup_{T\to\infty} \frac{1}{T}\int_0^T &\left[ \varphi_s^T b_s + U_s^T \frac{c_s^2}{2} + \beta_0(s) - \frac{\left(\varphi_s^T - \beta_a(s)\right)^2}{4\beta_{aa}(s)} \right] \mathrm{d}s \\ &= \limsup_{T\to\infty} \frac{1}{T}\int_0^T \left[ \varphi_s^\infty b_s + U_s^\infty \frac{c_s^2}{2} + \beta_0(s) - \frac{(\varphi_s^\infty - \beta_a(s))^2}{4\beta_{aa}(s)} \right] \mathrm{d}s.\end{aligned}$$

The result for the minimal ergodic cost $\check{\eta}$ follows completely analogously. ■

In the next proposition we state a result for the ergodic version of the value function, which actually does not represent the value (or in our case rather the costs) but can be used to derive the optimal control, as we show in the following corollary.

**Proposition 4.59**
*Let Assumption 4.36 be fulfilled. Define for $\tilde{\eta} \in \mathbb{R}$ and $(t,x) \in [0,\infty) \times \mathbb{R}$*

$$V_{\tilde{\eta}}^{\infty}(t,x) := \frac{1}{2}U_t^{\infty} \cdot x^2 + \varphi_t^{\infty} \cdot x + \int_0^t \tilde{\eta} - \left(\varphi_s^{\infty} b_s + U_s^{\infty}\frac{c_s^2}{2} + \beta_0(s) - \frac{(\varphi_s^{\infty} - \beta_a(s))^2}{4\beta_{aa}(s)}\right) \mathrm{d}s.$$

*Then, for all $\tilde{\eta} \in \mathbb{R}$ and $(t,x) \in [0,\infty) \times \mathbb{R}$, the function $V_{\tilde{\eta}}^{\infty}$ fulfills*

$$0 = \inf_{a \in \mathbb{R}} \left\{ \partial_t V_{\tilde{\eta}}^{\infty}(t,x) + (\mu(t,x) - a)\partial_x V_{\tilde{\eta}}^{\infty}(t,x) + \frac{1}{2}\sigma^2(t,x)\partial_{xx} V_{\tilde{\eta}}^{\infty}(t,x) + f(t,x,a) - \tilde{\eta} \right\},$$

*which we call the ergodic Hamilton-Jacobi-Bellmann equation (eHJB). Moreover, $\eta$ is the only real number with*

$$\begin{aligned}
&\liminf_{t\to\infty} \frac{1}{t} V_{\eta}^{\infty}(t,x) = 0 \qquad \textit{and} \\
&V_{\eta}^{\infty}(t,x) = \lim_{T\to\infty}\Big(V(t,T,x) - V(0,T,0)\Big) + \limsup_{T\to\infty} \frac{t}{T} V(0,T,0)
\end{aligned}$$

*and $\breve{\eta}$ is the only real number with*

$$\begin{aligned}
&\limsup_{t\to\infty} \frac{1}{t} V_{\breve{\eta}}^{\infty}(t,x) = 0 \qquad \textit{and} \\
&V_{\breve{\eta}}^{\infty}(t,x) = \lim_{T\to\infty}\Big(V(t,T,x) - V(0,T,0)\Big) + \liminf_{T\to\infty} \frac{t}{T} V(0,T,0).
\end{aligned}$$

*Proof.* Observe that $V_{\tilde{\eta}}^{\infty}(t,x) = \Theta^{\infty}(t,x) + t \cdot \tilde{\eta}$. Since $\Theta^{\infty}$ solves the HJB and the eHJB and HJB differ only by the term $-\tilde{\eta}$, we obtain

$$\begin{aligned}
&\inf_{a \in \mathbb{R}} \left\{ \partial_t V_{\tilde{\eta}}^{\infty}(t,x) + (\mu(t,x) - a)\partial_x V_{\tilde{\eta}}^{\infty}(t,x) + \frac{1}{2}\sigma^2(t,x)\partial_{xx} V_{\tilde{\eta}}^{\infty}(t,x) + f(t,x,a) - \tilde{\eta} \right\} \\
&\quad = \inf_{a \in \mathbb{R}} \Big\{ \partial_t \Theta^{\infty}(t,x) + (\mu(t,x) - a)\partial_x \Theta^{\infty}(t,x) + \frac{1}{2}\sigma^2(t,x)\partial_{xx}\Theta^{\infty}(t,x) + f(t,x,a) \\
&\qquad\qquad + \partial_t(t\cdot\tilde{\eta}) + (\mu(t,x) - a)\partial_x(t\cdot\tilde{\eta}) + \frac{1}{2}\sigma^2(t,x)\partial_{xx}(t\cdot\tilde{\eta}) - \tilde{\eta} \Big\} \\
&\quad = \inf_{a \in \mathbb{R}} \left\{ \partial_t \Theta^{\infty}(t,x) + (\mu(t,x) - a)\partial_x \Theta^{\infty}(t,x) + \frac{1}{2}\sigma^2(t,x)\partial_{xx}\Theta^{\infty}(t,x) + f(t,x,a) \right\} \\
&\quad = 0.
\end{aligned}$$

Next,

$$
\begin{aligned}
&\lim_{T\to\infty}\Big(V(t,T,x)-V(0,T,0)\Big)+\limsup_{T\to\infty}\frac{t}{T}V(0,T,0)\\
&=\lim_{T\to\infty}\Bigg(\frac{1}{2}U_t^T\cdot x^2+\varphi_t^T\cdot x+\int_t^T\left[\varphi_s^T b_s+U_s^T\frac{c_s^2}{2}+\beta_0(s)-\frac{\big(\varphi_s^T-\beta_a(s)\big)^2}{4\beta_{aa}(s)}\right]\mathrm{d}s\\
&\qquad\qquad -0-0\quad-\int_0^T\left[\varphi_s^T b_s+U_s^T\frac{c_s^2}{2}+\beta_0(s)-\frac{\big(\varphi_s^T-\beta_a(s)\big)^2}{4\beta_{aa}(s)}\right]\mathrm{d}s\\
&\qquad\qquad +t\cdot\limsup_{T\to\infty}\frac{1}{T}V(0,T,0)\Bigg)\\
&=\lim_{T\to\infty}\left(\frac{1}{2}U_t^T\cdot x^2+\varphi_t^T\cdot x\right)+t\cdot\eta-\lim_{T\to\infty}\int_0^t\left[\varphi_s^T b_s+U_s^T\frac{c_s^2}{2}+\beta_0(s)-\frac{\big(\varphi_s^T-\beta_a(s)\big)^2}{4\beta_{aa}(s)}\right]\mathrm{d}s\\
&=\frac{1}{2}U_t^\infty\cdot x+\varphi_t^\infty\cdot x+\int_0^t\eta\,\mathrm{d}s-\int_0^t\left[\varphi_s^\infty b_s+U_s^\infty\frac{c_s^2}{2}+\beta_0(s)-\frac{(\varphi_s^\infty-\beta_a(s))^2}{4\beta_{aa}(s)}\right]\mathrm{d}s\\
&=V_\eta^\infty(t,x)
\end{aligned}
$$

since everything inside the integral is bounded, in particular with respect to $T$, and hence dominated convergence can be applied together with Proposition 4.42 and Lemma 4.45. Analogously we obtain

$$\lim_{T\to\infty}\Big(V(t,T,x)-V(0,T,0)\Big)+\liminf_{T\to\infty}\frac{t}{T}V(0,T,0)=V_{\check{\eta}}^\infty(t,x).$$

Finally, observe that

$$
\begin{aligned}
&\liminf_{t\to\infty}\frac{1}{t}V_{\tilde{\eta}}^\infty(t,x)\\
&\quad=\liminf_{t\to\infty}\frac{1}{t}\left[\frac{1}{2}U_t^\infty\cdot x+\varphi_t^\infty\cdot x+\int_0^t\tilde{\eta}-\left(\varphi_s^\infty b_s+U_s^\infty\frac{c_s^2}{2}+\beta_0(s)-\frac{(\varphi_s^\infty-\beta_a(s))^2}{4\beta_{aa}(s)}\right)\mathrm{d}s\right]\\
&\quad=0+\tilde{\eta}-\limsup_{t\to\infty}\frac{1}{t}\int_0^t\left[\varphi_s^\infty b_s+U_s^\infty\frac{c_s^2}{2}+\beta_0(s)-\frac{(\varphi_s^\infty-\beta_a(s))^2}{4\beta_{aa}(s)}\right]\mathrm{d}s\\
&\quad=\tilde{\eta}-\eta
\end{aligned}
$$

and likewise

$$\limsup_{t\to\infty}\frac{1}{t}V_{\tilde{\eta}}^\infty(t,x)=\tilde{\eta}-\check{\eta}.$$

■

**Corollary 4.60**
*Let Assumption 4.36 be fulfilled. Then, for all $T>0$ and $(t,x)\in[0,T]\times\mathbb{R}$, we have*

$$\alpha_t^T(x)=\frac{\partial_x V(t,T,x)-\beta_a(t)-\beta_{xa}(t)x}{2\beta_{aa}(t)}$$

*and for all $(t,x)\in[0,\infty)\times\mathbb{R}$, $\tilde{\eta}\in\mathbb{R}$*

$$\alpha_t^\infty(x)=\frac{\partial_x V_{\tilde{\eta}}^\infty(t,x)-\beta_a(t)-\beta_{xa}(t)x}{2\beta_{aa}(t)}.$$

*Proof.* Using the explicit formulas of $V$ and $V^\infty$ as given in Theorem 4.57 and Proposition 4.59 and the explicit formulas for $\alpha^T$ and $\alpha^\infty$ given in Corollary 4.41 and Corollary 4.47 this is straightforward. ∎

In the next corollary we take a look at time-homogeneous parameter functions as a special case.

**Corollary 4.61**
*Let Assumption 4.36 be fulfilled and additionally the parameter functions be constant. Then*

$$V_{\tilde\eta}^\infty(t,x) = \frac{1}{2}U_0^\infty \cdot x^2 + \varphi_0^\infty \cdot x + t \cdot (\tilde\eta - \eta)$$

*for every $t \geqslant 0$, $x, \tilde\eta \in \mathbb{R}$ and the optimal ergodic cost is given by*

$$\eta = \check\eta = \varphi_0^\infty b_0 + U_0^\infty \frac{c_0^2}{2} + \beta_0(0) - \frac{(\varphi_0^\infty - \beta_a(0))^2}{4\beta_{aa}(0)},$$

*where for $p := 2B_0\beta_{aa}(0) + \beta_{xa}(0) + C_0^2\beta_{aa}(0)$ and $q := 4\beta_{xx}(0)\beta_{aa}(0) - \beta_{xa}^2(0)$*

$$U_0^\infty = p + \sqrt{p^2+q} \quad \text{and} \quad \varphi_0^\infty = \frac{2\beta_{aa}(0)\left[U_0^\infty\left[b_0 + c_0C_0 + \frac{\beta_a(0)}{2\beta_{aa}(0)}\right] - \frac{\beta_a(0)\beta_{xa}(0)}{2\beta_{aa}(0)} + \beta_x(0)\right]}{\sqrt{p^2+q} + C_0^2\beta_{aa}(0)}.$$

*In particular, only for the parameter $\eta$ does $V_\eta^\infty$ not depend on time and is equal to*

$$V_\eta^\infty(t,x) = \frac{1}{2}U_0^\infty \cdot x^2 + \varphi_0^\infty \cdot x = \lim_{T\to\infty} V(0,T,x) - V(0,T,0).$$

*Proof.* First, observe that $\varphi_\cdot^\infty b + U_\cdot^\infty \frac{c^2}{2} - \beta_0 - \frac{(\varphi_\cdot^\infty)^2 - \beta_a^2}{2\beta_{aa}} - \frac{(\varphi_\cdot^\infty - \beta_a)^2}{4\beta_{aa}}$ is constant by Corollary 4.44 and Corollary 4.46. Hence, for every $T > 0$

$$\frac{1}{T}\int_0^T \left[\varphi_s^\infty b_s + U_s^\infty \frac{c_s^2}{2} + \beta_0(s) - \frac{(\varphi_s^\infty - \beta_a(s))^2}{4\beta_{aa}(s)}\right] \mathrm{d}s = \varphi_0^\infty b_0 + U_0^\infty \frac{c_0^2}{2} + \beta_0(0) - \frac{(\varphi_0^\infty - \beta_a(0))^2}{4\beta_{aa}(0)}$$

and thus, by Corollary 4.58,

$$\begin{aligned}\varphi_0^\infty b_0 + U_0^\infty \frac{c_0^2}{2} + \beta_0(0) - \frac{(\varphi_0^\infty - \beta_a(0))^2}{4\beta_{aa}(0)} \\ = \lim_{T\to\infty}\frac{1}{T}\int_0^T \left[\varphi_s^\infty b_s + U_s^\infty \frac{c_s^2}{2} + \beta_0(s) - \frac{(\varphi_s^\infty - \beta_a(s))^2}{4\beta_{aa}(s)}\right] \mathrm{d}s = \eta = \check\eta.\end{aligned}$$

Using this together with the definition of $V_{\tilde\eta}^\infty$, we already have

$$V_{\tilde\eta}^\infty(t,x) = \frac{1}{2}U_0^\infty \cdot x^2 + \varphi_0^\infty \cdot x + t \cdot (\tilde\eta - \eta).$$

Finally, observe that

$$\begin{aligned}
\lim_{T\to\infty} V(0,T,x) - V(0,T,0) \\
&= \lim_{T\to\infty} \frac{1}{2}U_t^T \cdot x^2 + \varphi_t^T \cdot x + \int_0^T \left[ \varphi_s^T b_s + U_s^T \frac{c_s^2}{2} + \beta_0(s) - \frac{\left(\varphi_s^T - \beta_a(s)\right)^2}{4\beta_{aa}(s)} \right] \mathrm{d}s \\
&\qquad - \int_0^T \left[ \varphi_s^T b_s + U_s^T \frac{c_s^2}{2} + \beta_0(s) - \frac{\left(\varphi_s^T - \beta_a(s)\right)^2}{4\beta_{aa}(s)} \right] \mathrm{d}s \\
&= \lim_{T\to\infty} \frac{1}{2}U_t^T \cdot x^2 + \varphi_t^T \cdot x \\
&= V_\eta^\infty(t,x).
\end{aligned}$$

■

**Remark 4.62**
Restricting to time-homogeneous parameter functions and $\beta_{xa} = \beta_0 = 0$ and furthermore plugging the optimal control $\alpha_t^\infty(x) = \frac{\partial_x V_\eta^\infty(t,x) - \beta_a(t) - \beta_{xa}(t)x}{2\beta_{aa}(t)}$ into the eHJB we obtain that the pair $(V_\eta^\infty(0,\cdot),\eta)$ is the unique solution of

$$\begin{aligned}
0 &= \left( b_0 + B_0 x - \frac{\partial_x V_\eta^\infty(0,x) - \beta_a(0)}{2\beta_{aa}(0)} \right) \partial_x V_\eta^\infty(0,x) + \frac{1}{2}\left(c_0 + C_0 x\right)^2 \partial_{xx} V_\eta^\infty(0,x) \\
&\quad + \beta_{aa}(0)x^2 + \beta_x(0)x + \beta_{aa}(0) \left( \frac{\partial_x V_\eta^\infty(0,x) - \beta_a(0)}{2\beta_{aa}(0)} \right)^2 + \beta_a(0)\frac{\partial_x V_\eta^\infty(0,x) - \beta_a(0)}{2\beta_{aa}(0)} - \eta \\
&= \left(b_0 + B_0 x\right) \partial_x V_\eta^\infty(0,x) + \frac{1}{2}\left(c_0 + C_0 x\right)^2 \partial_{xx} V_\eta^\infty(0,x) \\
&\quad + \beta_{aa}(0)x^2 + \beta_x(0)x - \frac{\left(\partial_x V_\eta^\infty(0,x) - \beta_a(0)\right)^2}{4\beta_{aa}(0)} - \eta. \qquad (4.25)
\end{aligned}$$

Note that Equation (4.25) replicates the 1-dimensional version of the result in [BF92].

**Remark 4.63**
It is possible to weaken Assumption 4.36 a little bit. If the parameter functions are bounded by a slowly growing function, then they would still be bounded on every interval $[0,T] \subset [0,\infty)$, enabling us to still apply the results from Theorem 4.10 and Section 4.3. However, then the constants given in Proposition 4.33 and Theorem 4.34 depend on time since the growth of the parameter functions is involved. Choosing the growth small enough still yields some exponential decay allowing a polynomial growth for all parameters not yet restricted.

# 5 Simulation of McKean Vlasov SDEs with super linear growth

The aim of this chapter, which is based on [dRES18], is to develop a numerical scheme for simulating McKean-Vlasov Stochastic Differential Equations (MV-SDEs) with drifts of super-linear growth and Lipschitz diffusion coefficients (with linear growth). MV-SDEs differ from standard SDEs by means of the presence of the law of the solution process in the coefficients:

$$\mathrm{d}X_t = b(t, X_t, \mu_t^X)\,\mathrm{d}t + \sigma(t, X_t, \mu_t^X)\,\mathrm{d}W_t, \quad X_0 \in L_0^m(\mathbb{R}^d),$$

where $\mu_t^X$ denotes the law of the process $X$ at time $t$. Similar to standard SDEs, MV-SDEs have been shown to have a unique strong solution in the super-linear growth in spatial parameter setting, see [dRST19]. Of course, many mean-field models exhibit non-global Lipschitz growth, for example mean-field models for neuronal activity (e.g. stochastic mean-field FitzHugh-Nagumo models or the network of Hodgkin-Huxley neurons) [BFFT12], [BCC11], [BFT15] appearing in biology or physics [DGG$^+$11], [GGM$^+$18]. We refer to the review in [BFFT12] for further motivation of the problem.

In general, closed form solutions for such equations are rare. Hence, to fully utilize MV-SDEs as a modelling tool, one needs a reliable way in which to simulate them. It is well known that for SDEs the explicit Euler scheme runs into difficulties in the super-linear growth setting, see [HJK11], even though the SDE is known to have a unique strong solution. The original solution to this problem was to consider an implicit (or backward) Euler scheme developed in [HMS02]. Although implicit schemes allowed one to tackle more general SDEs they are slower especially in higher dimensions. The reason for this boils down to the fact that one is required to solve a fixed point equation at every time-step, which can be computationally expensive. To solve this problem an explicit scheme was then developed in [HJK12], a so-called *Tamed Euler* scheme. Since then several authors have built on this result and developed algorithms to deal with coefficients that grow super-linearly, see [CJM16], [Sab13], [FG16] for example. There has been some work on improved Monte Carlo methods for MV-SDEs with super-linear drift, see e.g. [dRST18].

An extra complication MV-SDEs offer over standard SDEs is the requirement to approximate the law $\mu$ at each time step. Although there are other techniques (see [GP18]), the most common is the so-called interacting particle system

$$\mathrm{d}X_t^{i,N} = b\big(t, X_t^{i,N}, \mu_t^{X,N}\big)\,\mathrm{d}t + \sigma\big(t, X_t^{i,N}, \mu_t^{X,N}\big)\,\mathrm{d}W_t^i,$$

where $\mu_t^{X,N}(\mathrm{d}x) := \frac{1}{N}\sum_{j=1}^N \delta_{X_t^{j,N}}(\mathrm{d}x)$ and $\delta_{X_t^{j,N}}$ is the Dirac measure at point $X_t^{j,N}$, and $W^i, i = 1, \ldots, N$ are independent Brownian motions. Under Lipschitz type conditions this particle system is known to converge pathwise to the true solution of the MV-SDE (see [Szn91], [Mél96]). However, this convergence (with corresponding rate) in a super-linear growth setting has thus far not been considered in full generality.

In this chapter we show that the above particle scheme converges (propagation of chaos) in the super-linear growth case without coercivity/dissipativity. This result is crucial in showing convergence of the numerical scheme to the particle system rather than to the original MV-SDE, with corresponding rate.

Furthermore, we develop an explicit scheme and prove strong convergence to the MV-SDE, inspired by the explicit scheme originally developed in [HJK12], [Sab13]. We also obtain the classical $1/2$ rate of convergence in the stepsize. Combining this with the propagation of chaos result gives an overall convergence rate for the explicit scheme.

The final contribution is to show strong convergence of an implicit scheme. This turns out to be a challenging problem since results involving implicit schemes rely on stopping time arguments. This causes several issues when generalizing results to the MV-SDE setting and we have to make stronger assumptions on the coefficients in this setting in order for the arguments to continue to hold. On the other hand, we allow for random initial conditions and time dependent coefficients that, to the best of our knowledge, have not been fully treated in the standard SDE setting. We discuss these issues in Remark 5.7. We only focus on strong convergence of this scheme and not the rate, mainly because the explicit scheme is shown to work under more general assumptions, scales better (as our numerical testing shows) and such proof would lead to lengthy statements without substantially enhancing the scope of applications. The question is left for future research with a tentative methodology discussed in Remark 5.11 below.

Other works, which are close to ours, are the following: [BF17] develop an explicit Euler scheme to deal with a specific MV-SDE type equation from a chemotaxis model; convergence is given but under Lipschitz conditions and constant diffusion coefficient. [Mal03] studies an implicit Euler scheme in order to approximate a specific equation and requires a constant diffusion coefficient, symmetry and uniform convexity of the interaction potential. Lastly, in [GPV19] Section 3.5 the authors are only able to justify their simulation for the Lipschitz case and the results we propose would allow for more general potentials.

This chapter is structured in the following way: In Section 5.1 we introduce the notation and our tamed particle scheme. In Section 5.2, we state our main result, namely, propagation of chaos and convergence results for the two schemes. Following that, in Section 5.3 we provide several numerical examples and highlight the *particle corruption* phenomena. This analysis implies one cannot hope to build a reliable scheme based on a standard Euler scheme. We further show the increased computational complexity associated with a MV-SDE makes the implicit scheme a less viable option than the explicit (tamed) scheme. Finally, the proofs are given in Section 5.4.

This chapter is based on [dRES18] in which most of my contribution is to the proof of the convergence of explicit Euler scheme and the implementation.

## 5.1 Preliminaries

Throughout the chapter we work on a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geqslant 0}, \mathbf{P})$ satisfying the usual conditions, where $\mathcal{F}_t$ is the augmented filtration of a standard multidimensional Brownian motion $W$. We work with $\mathbb{R}^d$, the $d$-dimensional Euclidean space of real numbers, and for $a = (a_1, \cdots, a_d) \in \mathbb{R}^d$ and $b = (b_1, \cdots, b_d) \in \mathbb{R}^d$ we denote by $|a|^2 = \sum_{i=1}^d a_i^2$ the usual Euclidean distance on $\mathbb{R}^d$ and by $\langle a, b\rangle = \sum_{i=1}^d a_i b_i$ the usual scalar product. For matrices $V \in \mathbb{R}^{k\times\ell}$ we define $|V| = \sup_{u\in\mathbb{R}^\ell,\ |u|\leqslant 1} |Vu|$.

We consider some finite terminal time $T < \infty$ and use the following notation for spaces, which are standard in the McKean-Vlasov literature (see [Car16]): We define $\mathbb{S}^p$ for $p \geqslant 1$, as the space of $\mathbb{R}^d$-valued, $\mathcal{F}_\cdot$-adapted processes $Z$, that satisfy $\mathbb{E}[\sup_{0\leqslant t\leqslant T} |Z(t)|^p]^{1/p} < \infty$.

Similarly, $L^p_t(\mathbb{R}^d)$, defines the space of $\mathbb{R}^d$-valued, $\mathcal{F}_t$-measurable random variables $X$, that satisfy $\mathbb{E}[|X|^p]^{1/p} < \infty$.

Given the measurable space $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$, we denote by $\mathcal{P}(\mathbb{R}^d)$ the set of probability measures on this space, and for $p \geqslant 1$ write $\mu \in \mathcal{P}_p(\mathbb{R}^d)$ if $\mu \in \mathcal{P}(\mathbb{R}^d)$ and for some $x \in \mathbb{R}^d$, $\int_{\mathbb{R}^d} |x - y|^p \mu(\mathrm{d}y) < \infty$. We then have the following metric on the space $\mathcal{P}_p(\mathbb{R}^d)$ (Wasserstein metric) for $\mu,\ \nu \in \mathcal{P}_p(\mathbb{R}^d)$ (see [Vil08], [dRST19] among others),

$$W^{(p)}(\mu,\nu) = \inf_{\pi} \Big\{ \Big( \int_{\mathbb{R}^d \times \mathbb{R}^d} |x-y|^p \pi(\mathrm{d}x, \mathrm{d}y) \Big)^{\frac{1}{p}} \ : \ \pi \in \mathcal{P}(\mathbb{R}^d \times \mathbb{R}^d) \text{ with marginals } \mu \text{ and } \nu \Big\}.$$

The most common choice in the McKean-Vlasov setting is, $p = 2$, and is what we shall use throughout most of this paper. As $W^{(2)}$ is a metric (see [Vil08] Chapter 6), we have for $\mu_1, \mu_2, \mu_3 \in \mathcal{P}_2(\mathbb{R}^d)$

$$W^{(2)}(\mu_1, \mu_3) \leqslant W^{(2)}(\mu_1, \mu_2) + W^{(2)}(\mu_2, \mu_3).$$

As in [Car16], we introduce the empirical measure constructed from i.i.d. samples of some process $X$ by $\mu^{X,N}_s := \frac{1}{N}\sum_{j=1}^N \delta_{X^j_s}$. Another standard result for the Wasserstein metric for two such empirical measures $\mu^{X,N}_s$, $\mu^{Y,N}_s$ is that

$$W^{(2)}(\mu^{X,N}_s, \mu^{Y,N}_s) \leqslant \Big( \frac{1}{N} \sum_{j=1}^N |X^j_s - Y^j_s|^2 \Big)^{1/2}.$$

### 5.1.1 McKean-Vlasov stochastic differential equations

Let $W$ be an $l$-dimensional Brownian motion and take the progressively measurable maps $b : [0,T] \times \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d) \to \mathbb{R}^d$ and $\sigma : [0,T] \times \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d) \to \mathbb{R}^{d \times l}$. MV-SDEs are typically written in the form,

$$\mathrm{d}X_t = b(t, X_t, \mu^X_t)\, \mathrm{d}t + \sigma(t, X_t, \mu^X_t)\, \mathrm{d}W_t, \quad X_0 \in L^p_0(\mathbb{R}^d), \tag{5.1}$$

where $\mu^X_t$ denotes the law of the process $X$ at time $t$, i.e. $\mu^X_t = \mathbf{P} \circ X_t^{-1}$. We make the following assumption on the coefficients throughout.

**Assumption 5.1**
*Assume that $\sigma$ is Lipschitz in the sense that there exists $L_\sigma > 0$ such that for all $t \in [0,T]$ and all $x, x' \in \mathbb{R}^d$ and $\forall \mu, \mu' \in \mathcal{P}_2(\mathbb{R}^d)$ we have that*

$$|\sigma(t,x,\mu) - \sigma(t,x',\mu')| \leqslant L_\sigma(|x-x'| + W^{(2)}(\mu,\mu')),$$

*and let $b$ satisfy*

1. *One-sided Lipschitz in $x$ and Lipschitz in law: there exist $L_b,\ L > 0$ such that for all $t \in [0,T]$, all $x, x' \in \mathbb{R}^d$ and all $\mu, \mu' \in \mathcal{P}_2(\mathbb{R}^d)$ we have that*

$$\langle x - x', b(t,x,\mu) - b(t,x',\mu) \rangle \leqslant L_b |x-x'|^2$$
$$\text{and} \quad |b(t,x,\mu) - b(t,x,\mu')| \leqslant L W^{(2)}(\mu,\mu').$$

2. *Locally Lipschitz with polynomial growth in $x$: there exist $L > 0$ and $q \in \mathbb{N}$ with $q > 1$ such that for all $t \in [0,T]$, $\forall \mu \in \mathcal{P}_2(\mathbb{R}^d)$ and all $x,\ x' \in \mathbb{R}^d$*

$$|b(t,x,\mu) - b(t,x',\mu)| \leqslant L(1 + |x|^q + |x'|^q)|x-x'|.$$

**Assumption 5.2**
*Assume that $b$ and $\sigma$ are $1/2$-Hölder continuous in time, uniformly in $x$ and $\mu$.*

Using the one-sided Lipschitz drift, a particularized version of Theorem 3.3 in [dRST19] provides a result for existence and uniqueness. Assumption 5.2 is not needed here.

**Theorem 5.3** ([dRST19])
*Suppose that $b$ and $\sigma$ satisfy Assumption 5.1 and 5.2. Further, assume for some $m \geqslant 2$, $X_0 \in L_0^m(\mathbb{R}^d)$. Then there exists a unique solution $X \in \mathbb{S}^m([0,T])$ to the MV-SDE (5.1). For some positive constant $C$ we have*

$$\mathbb{E}\big[\sup_{t\in[0,T]} |X_t|^m\big] \leqslant C\left(\mathbb{E}[|X_0|^m]+1\right)e^{CT}.$$

If the law $\mu^X$ is known beforehand, then the MV-SDE reduces to a "standard" SDE with added time-dependency. Typically this is not the case and usually the MV-SDE is approximated by a particle system.

**The interacting particle system approximation.** We approximate (5.1) (driven by the Brownian motion $W$), using an $N$-dimensional system of interacting particles. Let $i = 1, \dots, N$ and consider $N$ particles $X^{i,N}$ satisfying the SDE with i.i.d. $X_0^{i,N} = X_0^i$ (the initial condition is random, but independent of other particles)

$$\mathrm{d}X_t^{i,N} = b\Big(t, X_t^{i,N}, \mu_t^{X,N}\Big)\,\mathrm{d}t + \sigma\Big(t, X_t^{i,N}, \mu_t^{X,N}\Big)\,\mathrm{d}W_t^i, \tag{5.2}$$

where $\mu_t^{X,N}(\mathrm{d}x) := \frac{1}{N}\sum_{j=1}^N \delta_{X_t^{j,N}}(\mathrm{d}x)$ and $\delta_{X_t^{j,N}}$ is the Dirac measure at point $X_t^{j,N}$, and the independent Brownian motions $W^i, i = 1, \dots, N$ (also independent of the BM $W$ appearing in (5.1); with a slight abuse of notation to avoid re-defining the probability space's Filtration).

**Propagation of chaos.** In order to show that the particle approximation is of use, one shows a pathwise propagation of chaos result. Although different types exist we are interested in the strong error. Hence a pathwise convergence result is needed and we consider the system of non interacting particles

$$\mathrm{d}X_t^i = b(t, X_t^i, \mu_t^{X^i})\,\mathrm{d}t + \sigma(t, X_t^i, \mu_t^{X^i})\,\mathrm{d}W_t^i, \quad X_0^i = X_0^i, \quad t \in [0,T], \tag{5.3}$$

which are of course just MV-SDEs and since the $X^i$s are independent, $\mu_t^{X^i} = \mu_t^X$ for all $i$. Under global Lipschitz conditions, one can then prove the following convergence result (see Theorem 1.10 in [Car16] for example)

$$\lim_{N\to\infty} \sup_{1\leqslant i\leqslant N} \mathbb{E}\big[\sup_{0\leqslant t\leqslant T} |X_t^{i,N} - X_t^i|^2\big] = 0.$$

Several propagation of chaos results have been shown over the years under varying conditions, see [Szn91], [Mél96] and [Lac18] among others. All SDEs appearing below have initial condition $X_0^i$ and we work on the interval $[0,T]$.

**Standard Euler scheme particle system.** In general one cannot simulate (5.2) directly and therefore turns to a numerical scheme such as Euler. We partition the time interval $[0,T]$ into $M$ steps of size $h := T/M$, we then define $t_k := kh$ and recursively define the particle system for $k \in \{0, \dots, M-1\}$ as,

$$\bar{X}_{t_{k+1}}^{i,N,M} = \bar{X}_{t_k}^{i,N,M} + b\Big(t_k, \bar{X}_{t_k}^{i,N,M}, \bar{\mu}_{t_k}^{X,N}\Big)h + \sigma\Big(t_k, \bar{X}_{t_k}^{i,N,M}, \bar{\mu}_{t_k}^{X,N}\Big)\Delta W_{t_k}^i,$$

where $\bar{\mu}^{X,N}_{t_k}(\mathrm{d}x) := \frac{1}{N}\sum_{j=1}^{N}\delta_{\bar{X}^{j,N,M}_{t_k}}(\mathrm{d}x)$, $\Delta W^i_{t_k} := W^i_{t_{k+1}} - W^i_{t_k}$ and $\bar{X}^{i,N,M}_0 := X^i_0$. Under Lipschitz regularity it is well known that this scheme converges, see [BT97] or [KHO97] (here a weak rate of convergence is shown under an additional regularity assumption).

**Euler particle system for the super-linear case: Explicit and Implicit.** However, as discussed in works such as [HJK11], [HJK12], [Sab13] one does not have convergence of the Euler scheme when we move away from the global Lipschitz setting. The goal of this chapter is to therefore construct a suitable numerical scheme which converges. Inspired by the above works we consider a so-called *tamed* Euler scheme. With the notation above we consider the following scheme

$$\bar{X}^{i,N,M}_{t_{k+1}} = \bar{X}^{i,N,M}_{t_k} + \frac{b\Big(t_k, \bar{X}^{i,N,M}_{t_k}, \bar{\mu}^{X,N}_{t_k}\Big)}{1 + M^{-\alpha}\Big|b\Big(t_k, \bar{X}^{i,N,M}_{t_k}, \bar{\mu}^{X,N}_{t_k}\Big)\Big|} h + \sigma\Big(t_k, \bar{X}^{i,N,M}_{t_k}, \bar{\mu}^{X,N}_{t_k}\Big)\Delta W^i_{t_k}, \tag{5.4}$$

where $\bar{\mu}^{X,N}_{t_k}(\mathrm{d}x) = \frac{1}{N}\sum_{j=1}^{N}\delta_{\bar{X}^{j,N,M}_{t_k}}(\mathrm{d}x)$ and $\alpha \in (0, 1/2]$ with $\bar{X}^{i,N,M}_0 = X^i_0$.

Of course, explicit schemes are not the only method one can deploy to solve this problem, we also consider the following implicit scheme

$$\tilde{X}^{i,N,M}_{t_{k+1}} = \tilde{X}^{i,N,M}_{t_k} + b\Big(t_k, \tilde{X}^{i,N,M}_{t_{k+1}}, \tilde{\mu}^{X,N,M}_{t_k}\Big)h + \sigma\Big(t_k, \tilde{X}^{i,N,M}_{t_k}, \tilde{\mu}^{X,N,M}_{t_k}\Big)\Delta W^i_{t_k}, \tag{5.5}$$

where $\tilde{\mu}^{X,N,M}_{t_k}(\mathrm{d}x) := \frac{1}{N}\sum_{j=1}^{N}\delta_{\tilde{X}^{j,N,M}_{t_k}}(\mathrm{d}x)$ and $\tilde{X}^{i,N,M}_0 = X^i_0$.

## 5.2 Main Results

We state our main results and assumption here, the proofs are postponed to Section 5.4. Recall that we want to associate a particle system to the MV-SDE and show its convergence, so-called *propagation of chaos*. We have the following result that holds under weaker assumptions than those in Theorem 5.6.

**Proposition 5.4** (Propagation of chaos)
*Let the assumption in Theorem 5.3 hold for $m > 4$. Let $X^i$ be the solution to* (5.3)*, and $X^{i,N}$ be the solution to* (5.2).

*Then we have the following convergence result.*

$$\sup_{1\leqslant i\leqslant N}\mathbb{E}\Big[\sup_{0\leqslant t\leqslant T}|X^i_t - X^{i,N}_t|^2\Big] \leqslant C\begin{cases} N^{-1/2} & \textit{if } d < 4, \\ N^{-1/2}\log(N) & \textit{if } d = 4, \\ N^{-2/d} & \textit{if } d > 4. \end{cases}$$

This result shows the particle scheme will converge to the MV-SDE with a given rate. Therefore, to show convergence between our numerical scheme and the MV-SDE, we only need to show that the “true” particle scheme and numerical version of the particle scheme converge.

## Explicit scheme

We first introduce the continuous time version of the explicit scheme (5.4). Denote by $\kappa(t) := \sup\{s \in \{0, h, 2h, \ldots, Mh\} : s \leqslant t\}$ for all $t \in [0,T]$, $b_M(t,x,\nu) := \frac{b(t,x,\nu)}{1+M^{-\alpha}|b(t,x,\nu)|}$ with $\alpha \in (0, 1/2]$ for all $t \in [0,T]$, $x \in \mathbb{R}^d$, $\nu \in \mathcal{P}_2(\mathbb{R}^d)$

$$
\begin{aligned}
X_t^{i,N,M} &= X_0^i + \int_0^t b_M\left(\kappa(s), X_{\kappa(s)}^{i,N,M}, \mu_{\kappa(s)}^{X,N,M}\right) \mathrm{d}s \\
&\quad + \int_0^t \sigma\left(\kappa(s), X_{\kappa(s)}^{i,N,M}, \mu_{\kappa(s)}^{X,N,M}\right) \mathrm{d}W_s^i, \quad \mu_t^{X,N,M}(\mathrm{d}x) = \frac{1}{N}\sum_{j=1}^N \delta_{X_t^{j,N,M}}(\mathrm{d}x). \qquad (5.6)
\end{aligned}
$$

Note that $|b_M(t,x,\nu)| \leqslant \min\left(M^\alpha, |b(t,x,\nu)|\right)$ and that $\bar{X}_{t_k}^{i,N,M} = X_{t_k}^{i,N,M}$ for all $k \in \{0,1,\ldots,M\}$ and hence $X^{i,N,M}$ is a continuous version of $\bar{X}^{i,N,M}$ from (5.4). We then obtain the following convergence result.

**Proposition 5.5**
*Let the assumptions in Theorem 5.6 (see below) hold. Then it holds that*

$$\sup_{1\leqslant i\leqslant N} \mathbb{E}[\sup_{0\leqslant t\leqslant T} |X_t^{i,N} - X_t^{i,N,M}|^2] \leqslant Ch.$$

This then leads to our main explicit scheme convergence result.

**Theorem 5.6** (Strong Convergence of Explicit)
*Let Assumption 5.1 and 5.2 hold, further let $X_0 \in L^m(\mathbb{R}^d)$ for $m \geqslant 4(1+q)$ (note $q > 1$) and set $\alpha = 1/2$. Let $X^i$ be the solution to* (5.3)*, and $X^{i,N,M}$ be that for* (5.6).
*Then we obtain the following convergence result*

$$\sup_{1\leqslant i\leqslant N} \mathbb{E}\big[\sup_{0\leqslant t\leqslant T} |X_t^i - X_t^{i,N,M}|^2\big] \leqslant C \begin{cases} N^{-1/2} + h & \text{if } d < 4, \\ N^{-1/2}\log(N) + h & \text{if } d = 4, \\ N^{-2/d} + h & \text{if } d > 4. \end{cases}$$

*Proof of Theorem 5.6.* Theorem 5.6 is a consequence of Propositions 5.4 and 5.5. ∎

**Remark 5.7** (Issues using stopping times)
The technique of using the stopping time $\tau_R^i := \inf\{t \geqslant 0 : |X_t^{i,N,M}| \geqslant R\}$ to control the particles is suboptimal here and several problems appear by introducing them. Namely, one can only consider stopping times that stop one particle since otherwise the convergence speed would decrease with a higher number of particles. However, applying a stopping time to a single particle does not allow us to fully bound the coefficients and moreover destroys the result of all particles being identically distributed.

The stopping time arguments used for the implicit scheme below require stronger assumptions in order to make the theory hold.

## Implicit scheme

As alternative to the explicit scheme we now discuss the implicit or backward Euler scheme. That being said, the implicit scheme has some well documented disadvantages, namely it is expensive compared to its explicit counterpart, we discuss this issue further in Section 5.3. One can consult, [MS13] for example on the implicit scheme (and extensions) for standard SDEs.

Standard implicit scheme convergence results rely on the so called monotone growth condition, we therefore proceed with the following assumption.

**Assumption 5.8**

*(H1). There exists a constant $C$ such that, for all $\mu \in \mathcal{P}_2(\mathbb{R}^d)$,*

$$|b(0,0,\mu)| + |\sigma(0,0,\mu)| \leqslant C.$$

*(H2). The drift and diffusion coefficient satisfy the stronger bound in measure condition, for all $t \in [0,T]$, all $x \in \mathbb{R}^d$ and all $\mu, \mu' \in \mathcal{P}_2(\mathbb{R}^d)$*

$$|b(t,x,\mu) - b(t,x,\mu')| + |\sigma(t,x,\mu) - \sigma(t,x,\mu')| \leqslant LW^{(1)}(\mu,\mu'),$$

*where $W^{(1)}(\cdot,\cdot)$ is the Wasserstein-1 distance.*

Although the main convergence theorem requires both H1 and H2, we only use H2 at the end of the proof of convergence. We present our auxiliary results requiring only H1 as we believe them to be of general independent interest.

We now state the strong convergence of the implicit scheme (5.5) to (5.2).

**Proposition 5.9**

*Let Assumption 5.1, 5.2 and 5.8 hold. Fix a timestep $h^* < 1/\max(L_b, 2\beta)$ and assume that $X_0 \in L^{4(q+1)}(\mathbb{R}^d)$. Let $X^{i,N}$ be the solution to* (5.2)*, and $\tilde{X}^{i,N,M}$ be that for* (5.5)*. Then, for any $h$ and $M$ with $T = hM$ and $s \in [1,2)$*

$$\sup_{1 \leqslant i \leqslant N} \lim_{h \to 0} \mathbb{E}[|X_T^{i,N} - \tilde{X}_T^{i,N,M}|^s] = 0.$$

**Theorem 5.10** (Strong Convergence of Implicit Scheme)

*Let the assumption in Proposition 5.9 hold and let $X^i$ be the solution to* (5.3)*, and $\tilde{X}^{i,N,M}$ be that for* (5.5)*. Then, for any $h$ and $M$ with $T = hM$ and $s \in [1,2)$ one has*

$$\lim_{N \to \infty} \sup_{1 \leqslant i \leqslant N} \lim_{h \to 0} \mathbb{E}[|X_T^{i} - \tilde{X}_T^{i,N,M}|^s] = 0.$$

*Proof.* The proof of this result follows by combing Proposition 5.4 and 5.9 and noting that the assertion in Proposition 5.9 is independent of $N$. ■

**Remark 5.11** (On the convergence rate of the implicit scheme)

Theorem 5.10 shows the convergence of the implicit scheme but without establishing a rate. Methodologically speaking, the approach proposed in [HMS02] seems applicable here, where the convergence rate of the implicit scheme would be shown by defining an intermediate process and considering the convergence of the implicit scheme to the intermediate process and then that of the intermediate process to the original equation, see [HMS02]. We suspect that such proof is not straightforward with several extra constraints appearing due to the presence of the law. As it stands, the convergence of our implicit scheme requires stronger assumptions (see Assumption 5.8) than the explicit one so we leave establishing the rate for future. Our numerical experiments hint that the convergence rate should be the same as the explicit, which is consistent with the case of standard SDEs.

## 5.3 Numerical testing and Examples

We illustrate our results with numerical examples. We highlight the issues of using the standard Euler scheme in this setting and also compare the computational time and complexity of the explicit and implicit scheme. We juxtapose our findings to those in [BFFT12].

### 5.3.1 Particle Corruption

It is well known that the Euler scheme fails (diverges) when one moves outside the realm of linear growing coefficients, see [HJK11]. We claim that this divergence is worse in the setting of MV-SDEs and associated particle system due to an effect we refer to as *particle corruption*.

The basic idea is that one particle becomes influential on all other particles, thus we are no longer in the setting of "weakly interacting". This is of course not a problem for standard SDE simulation. We show two aspects of particle corruption in a simple example. Firstly, one particle can cause the whole system to crash. Secondly and perhaps more profoundly, the more particles one has the more likely this occurs. This is of course a devastating issue when simulating a MV-SDE since accurately approximating the measure depends on having a large number of interacting particles.

To show this example we take a classical non-globally Lipschitz SDE, the stochastic Ginzburg Landau equation (see [Tie13]) and add a simple mean field term to it,

$$\mathrm{d}X_t = \Big(\frac{\sigma^2}{2}X_t - X_t^3 + c\mathbb{E}[X_t]\Big)\,\mathrm{d}t + \sigma X_t\,\mathrm{d}W_t, \quad X_0 = x.$$

This MV-SDE clearly satisfies the assumption to have a unique strong solution in $\mathbb{S}^p$ for all $p > 1$, hence in theory one could calculate $\varphi(t) := \mathbb{E}[X_t]$ and have a standard SDE with one-sided Lipschitz drift. The analysis carried out in [HJK11] then implies that the Euler scheme diverges here.

**Showing particle corruption exists.** For our example we simulate $N = 5000$ particles with a time step $h = 0.05$, $T = 2$ and $X_0 = 1$, we also take $\sigma = 3/2$ and $c = 1/2$. We rerun this example until we observed a blow up and plotted the particle paths in Figure 5.1.

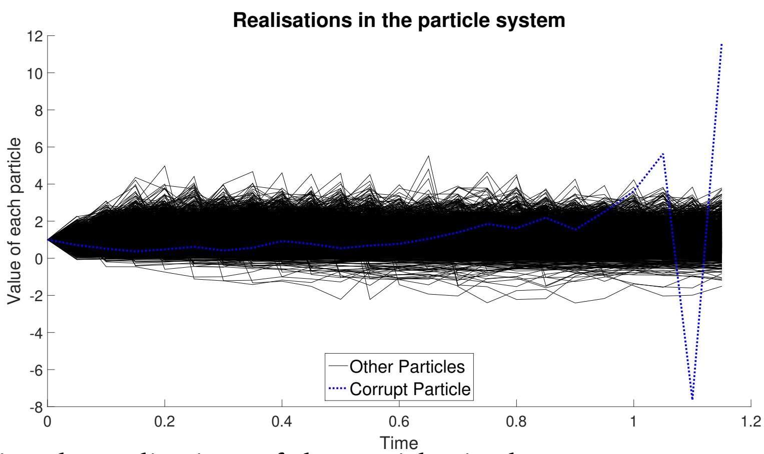



Figure 5.1: Showing the realizations of the particles in the system. We note that the particle given by the dashed line is starting to oscillate and is taking larger values than its surrounding particles.

Figure 5.1 shows the first part of the divergence, namely all particles are reasonably well behaved until one starts to oscillate rapidly. We have stopped plotting before the time boundary since this particle diverges shortly after this. We refer to this particle as the *corrupt particle* and it is fairly straightforward to see it will diverge. However, due to the interaction this single particle influences all the remaining particles and the whole system diverges shortly after.

**Remark 5.12** (Why is particle corruption so pronounced?)
The reason this effect is so dramatic is a simple consequence of the mean-field interaction. Typically, one observes divergence of the Euler scheme via a handful of Monte Carlo simulations that return extremely large (or infinite) values. When one then looks to calculate the expected value of the SDEs at the terminal time for example, these few events completely dominate the other results. This is summed up in a statement of [HJK11], where an exponentially small probability event has a double exponential impact.

The difference in the MV-SDE (weakly interacting particle) case is that the expectation appears inside the simulation, hence a divergence of a single particle influences multiple particles simultaneously during the simulation and not just at the final time.

**Convergence of Euler and propagation of chaos is impossible.** The above shows that one particle diverging can cause the whole system to diverge. One may argue that using more particles would reduce the dependency between them and hence influence the system less. In fact as we shall see the opposite is true, the more particles the more likely a divergence is. To test this we use the same example as above but use $N = [1000,\ 5000,\ 10000,\ 20000]$ particles and rerun each case $1000$ times and record the total number of times we observe a divergence over the ensemble.

| Number of particles | 1000 | 5000 | 10000 | 20000 |
|---|---|---|---|---|
| Number of blow ups | 3 | 32 | 43 | 108 |

Table 5.1: Number of divergences recorded at each particle level out of $1000$ simulations.

The results in Table 5.1 show conclusively that the more particles the more likely a divergence is to occur. This is a real problem in this setting since in order to minimize the propagation of chaos error one should take $N$ as large as possible, but doing so makes the Euler scheme approximation (likelier to) diverge.

**Remark 5.13** (Euler cannot work)
We have shown that naively applying the standard Euler scheme in the MV-SDE setting with non globally Lipschitz coefficient has issues. However, for standard SDEs there are some simple fixes one can apply and still obtain convergence e.g. removing paths that leave some ball as considered in [MT05]. Methods like this cannot work here since, we either take the ball "small" and therefore our approximation to the law is poor. Or we take a large ball, but then as the particles head towards the boundary they can "drag" other particles with them which again makes the system unstable.

The dependence on the measure (other particles) implies that the cruder approximation techniques cannot yield the strong convergence results we obtain with the more sophisticated techniques presented in this chapter. In [BFFT12] the authors have a non-globally Lipschitz MV-SDE and simulate using standard Euler scheme. Since no divergence was observed in their simulations they conjectured that the Euler scheme works in their setting. However, they used a "small" diffusion coefficient ($\sigma \in [0, 0.5]$) and small particle number (in the order of hundreds), which makes divergence unlikely to be observed (but not impossible) and yields poorer approximation results. Again, our methods provide certainty in terms of convergence (and convergence rate).

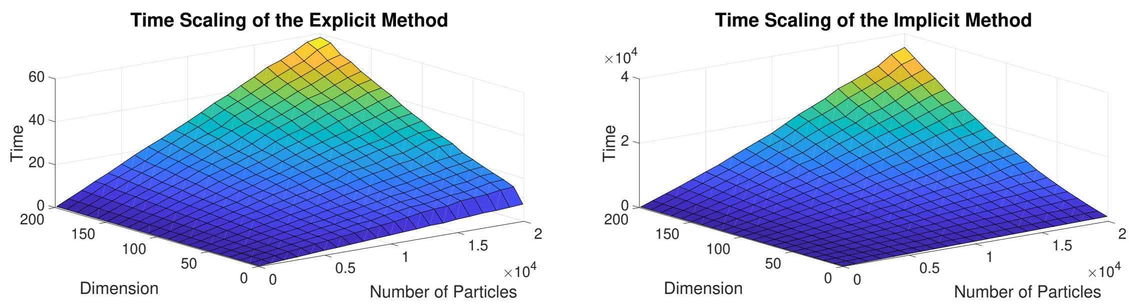
**Phase transition and particle systems within a bistable potential.** We have applied our algorithms to the problem highlighted in [GPV19] (see their equation (2.1) and the setup of their Section 3.5) and shortly report that we recover the same findings as above to their problem when dealing with the bistable potential $V_\eta(\eta) = \eta^4/4 - \eta^2/2$. Divergence of the explicit Euler scheme in [GPV19, Section 3.5] when using $V_\eta'(\eta)$ while both schemes we propose behave as we have described. We do not provide the numerical experiments as it would be a repetition of the results above.

### 5.3.2 Timing of Implicit vs Explicit: Size of cloud and spatial dimension

It is well documented that implicit schemes are slower than explicit ones, mainly because one must solve a fixed point equation at each step. This operation is not "cheap" and moreover scales $d^2$ in dimension, see [HJK12]. Of course this analysis is carried out for standard SDEs. What we wish to consider is how the particle system affects the timing of both methods.

We consider the same example as previous (but take $T = 1$), we then consider a set of dimensions from $1$ to $200$ and number of particles from $100$ to $20000$. Plotting the time taken for both methods is given in Figure 5.2.

Firstly, we observe that the explicit scheme is two to three orders of magnitude faster than the implicit scheme. At the highest dimensional and particle number this difference is very apparent with the tamed scheme taking approximately 1 minute and the implicit 10 hours. Another note to make is the scaling of each method: both methods scale similarly with particle number, but the tamed scheme scales linearly with dimension; this is superior to the $d^2$ scaling of the implicit scheme.




(a) Explicit Scheme (b) Implicit Scheme

Figure 5.2: Showing how the time (in seconds) of the explicit scheme (left; timescale $\approx 60$ seconds) and implicit scheme (right; timescale $\approx 10^4$ seconds) changes with particles and dimension.

Even for the case $d = 1$, $N = 20000$ the tamed scheme takes approximately 7 seconds while the implicit scheme takes approximately 23 minutes. For many practical applications $N = 20000$ is not enough for an acceptable level of accuracy, with this in mind and the dimension scaling, this makes the implicit scheme a very expensive method in this setting.

### 5.3.3 Explicit Vs Implicit Convergence: the Neuron Network Model

We compare the convergence of the explicit and the implicit scheme. To this end we use the system in [BFFT12] where the authors develop a non globally Lipschitz MV-SDE to model neuron activity. In our notation their system with $b : [0,T] \times \mathbb{R}^3 \times \mathcal{P}_2(\mathbb{R}^3) \to \mathbb{R}^3$, $\sigma : [0,T] \times \mathbb{R}^3 \times \mathcal{P}_2(\mathbb{R}^3) \to \mathbb{R}^{3\times 3}$ reads for $x = (x_1, x_2, x_3), z = (z_1, z_2, z_3) \in \mathbb{R}^3$ as

$$b\,(t,x,\mu) := \begin{pmatrix} x_1 - (x_1)^3/3 - x_2 + I - \int_{\mathbb{R}^3} J\,(x_1 - V_{rev})\, z_3 \,\mathrm{d}\mu(z) \\ c\,(x_1 + a - bx_2) \\ a_r \frac{T_{max}(1-x_3)}{1+\exp(-\lambda(x_1 - V_T))} - a_d x_3 \end{pmatrix}$$

$$\sigma\,(t,x,\mu) := \begin{pmatrix} \sigma_{ext} & 0 & -\int_{\mathbb{R}^3} \sigma_J\,(x_1 - V_{rev})\, z_3 \,\mathrm{d}\mu(z) \\ 0 & 0 & 0 \\ 0 & \sigma_{32}(x) & 0 \end{pmatrix}$$

with

$$\sigma_{32}(x) := \mathbb{1}_{\{x_3 \in (0,1)\}} \sqrt{a_r \frac{T_{max}(1-x_3)}{1+\exp(-\lambda(x_1 - V_T))} + a_d x_3}\; \Gamma \exp(-\Lambda/(1-(2x_3-1)^2)),$$

$T = 2$ is chosen as the final time and

$$X_0 \sim \mathcal{N}\left( \begin{pmatrix} V_0 \\ w_0 \\ y_0 \end{pmatrix}, \begin{pmatrix} \sigma_{V_0} & 0 & 0 \\ 0 & \sigma_{w_0} & 0 \\ 0 & 0 & \sigma_{y_0} \end{pmatrix} \right),$$

where the parameters have the values

$$\begin{array}{ccccccc} V_0 = 0 & \sigma_{V_0} = 0.4 & a = 0.7 & b = 0.8 & c = 0.08 & I = 0.5 & \sigma_{ext} = 0.5 \\ w_0 = 0.5 & \sigma_{w_0} = 0.4 & V_{rev} = 1 & a_r = 1 & a_d = 1 & T_{max} = 1 & \lambda = 0.2 \\ y_0 = 0.3 & \sigma_{y_0} = 0.05 & J = 1 & \sigma_J = 0.2 & V_T = 2 & \Gamma = 0.1 & \Lambda = 0.5. \end{array}$$

As the true solution is unknown to compare the convergence rates, we use as proxy the output of the explicit scheme with $2^{23}$ steps. Since the explicit scheme has convergence rate $\sqrt{h}$ we know that $2^{16}$ steps and below yields one order of magnitude larger errors. The simulation for 1000 particles and average root mean square error of each particle is given in Figure 5.3.

One can observe that although initially the implicit scheme has a better rate of convergence, it levels off to yield the expected $1/2$ rate[1]. Making the explicit scheme the more computationally efficient. Of course our "true" was calculated from the explicit scheme, hence we additionally carried out a similar test with a "true" from the implicit, and the results were almost identical.

**Remark 5.14** (Small Diffusion Setting)
Above, we have taken $\sigma_{ext} = 0.5$, this goes against the example in [BFFT12] where $\sigma_{ext} = 0$. As it turns out, in the case $\sigma_{ext} = 0$, the implicit scheme has a convergence rate close to $1$ (up to an error of around $10^{-4}$), while the explicit scheme maintains the standard $1/2$ rate. It is our belief that this is due to the fact that when $\sigma_{ext} = 0$ the diffusion coefficient makes little difference, hence both scheme revert close to their deterministic convergence rate. The explicit scheme of course still rate of order $1/2$, while the implicit is order $1$. It may therefore be that in the setting of small diffusion terms the implicit can yield superior results, of course though this is a special case and is not true in general.

[1] One can note that the x-axis is written in terms of runtime rather than number of time-steps. As there is a one to one correspondence between the time-steps and the time taken we can still determine the rate. However, this scale allows one to compare both the rate and the time-taken to achieve a given error.

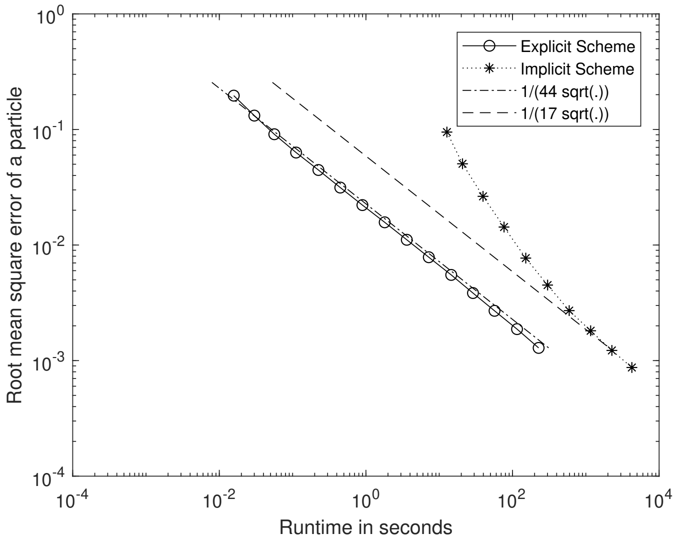



Figure 5.3: Root mean square error of the explicit and implicit (see Footnote 1). The number of steps of the explicit scheme are $M \in \{2^2, 2^3, \ldots, 2^{16}\}$ and of the implicit scheme are $M \in \{2^2, 2^3, \ldots, 2^{11}\}$. We used $1000$ particles and the true is calculated from the explicit with $2^{23}$ steps. Both schemes converge with rate $1/2$.

## Obtaining the Density

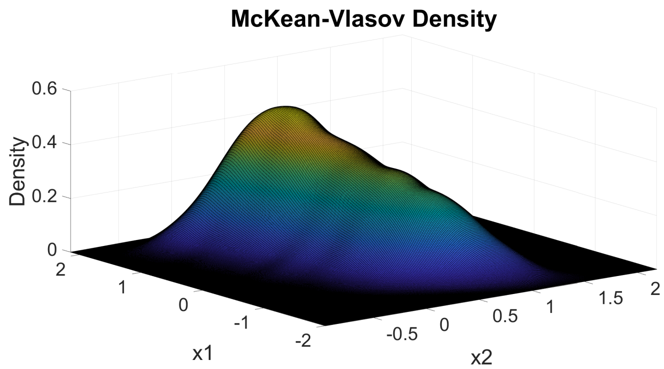



Figure 5.4: Approximate density of the first and second component of the MV-SDE at time $T = 1.2$. We used $10000$ particles, $2^{20}$ steps and a bandwidth of $0.15$ in the kernel smoothing.

In some applications as well as the value of the MV-SDE at the terminal time, one may also be interested in the density (law). In [BFFT12, Section 4] the authors compare density estimation using both the Fokker-Plank equation and the histogram from the particle system. The approach using PDEs becomes computationally expensive here if one considers multiple populations of MV-SDE and hence the authors take a simple case (see [BFFT12, Section 4.3]). There are of course other drawbacks such as dimension scaling which often make stochastic techniques more favorable in this setting. Moreover, using the PDE one will only obtain the density. If one is further interested in calculating a “payoff” i.e. $\mathbb{E}[G(X_T)]$ for some function $G$, then we would require an additional integral approximation or Metropolis Hastings style sampling scheme to calculate this expectation. While [BFFT12] apply a basic histogram approach when using MV-SDEs, this does not yield particularly nice results, namely, the resultant density is not a smooth surface. There are however, many statistical techniques one can use to improve this, see [Kee11, Chapter 18.4] for further results and discussion. Taking the example in [BFFT12] (with $\sigma_{ext} = 0$) and applying MATLAB's `ksdensity` function we obtain Figure 5.4.

One can observe the similarity between our result using SDEs and the one obtained in [BFFT12, pg 31] using the (expensive) PDE approach.

### Conclusions and future work

We have shown how one can apply the techniques from SDEs to the MV-SDE setting and some of its pitfalls and challenges that arise. The numerical testing carried out shows that the explicit scheme yields superior results (over the implicit scheme) in general.

Although we have been able to obtain convergence for the implicit scheme it is under stronger assumptions than the explicit scheme (the implicit scheme works very well in Section 5.3.3). The reason for these assumptions is that the implicit scheme is more challenging to bound than the explicit. The standard approach around this problem is to use stopping time arguments. However, as described in Remark 5.7, stopping times are harder to handle in the MV-SDE framework. Caution is needed to account for the extra technicalities that arise.

It is our belief that Assumption 5.8, although sufficient, is not necessary to guarantee the implicit scheme converges. As research is carried out into stopping times and MV-SDEs, future theoretical developments in this direction may allow this assumption to be weakened. We also leave open a proof for the convergence rate of the implicit scheme. Showing such a convergence rate in our framework is clearly possible but adds little in scope given the gains of the explicit over the implicit scheme. We leave the question open until a time a more resourceful implicit scheme can be designed.

Another interesting area which we have not discussed is sign preservation and the impact it has on the law. For example a MV-SDE may be known to be positive. However, if the numerical scheme takes the solution into the negative region how does the law dependence influence the remaining particles? One can consider the special case of $L_b < 0$ in Assumption 5.1, even though the MV-SDE could have a nonnegative solution, the numerical scheme may not preserve this feature.

## 5.4 Proof of Main Results

We shall use $C$ to denote a constant that can changes from line to line, but only depends on known quantities, $T$, $d$, the one-sided Lipschitz coefficients etc.

### 5.4.1 Propagation of Chaos

Let us show the propagation of chaos result.

*Proposition 5.4.* Let us fix $1 \leqslant i \leqslant N$, we then approach the proof in the usual way for dealing with one-sided Lipschitz coefficients, namely we apply Itô's formula to the difference (note the

$X_0^i$ cancel out),

$$|X_t^i - X_t^{i,N}|^2 = \int_0^t 2\langle X_s^i - X_s^{i,N}, b(s, X_s^i, \mu_s) - b(s, X_s^{i,N}, \mu_s^{X,N})\rangle \,\mathrm{d}s \\ + \int_0^t 2\langle X_s^i - X_s^{i,N}, (\sigma(s, X_s^i, \mu_s) - \sigma(s, X_s^{i,N}, \mu_s^{X,N}))\,\mathrm{d}W_s^i\rangle \\ + \sum_{a=1}^{l} \int_0^t |\sigma_a(s, X_s^i, \mu_s) - \sigma_a(s, X_s^{i,N}, \mu_s^{X,N})|^2 \,\mathrm{d}s, \tag{5.7}$$

where $\sigma_a$ is the $a$th column of matrix $\sigma$, hence $\sigma_a$ is a $d$-dimensional vector. Considering the first integral in (5.7),

$$\langle X_s^i - X_s^{i,N}, b(s, X_s^i, \mu_s) - b(s, X_s^{i,N}, \mu_s^{X,N})\rangle \\ = \langle X_s^i - X_s^{i,N}, b(s, X_s^i, \mu_s) - b(s, X_s^{i,N}, \mu_s)\rangle + \langle X_s^i - X_s^{i,N}, b(s, X_s^{i,N}, \mu_s) - b(s, X_s^{i,N}, \mu_s^{X,N})\rangle.$$

Applying the one-sided Lipschitz property in space and $W^{(2)}$ in measure along with Cauchy-Schwarz we obtain,

$$\langle X_s^i - X_s^{i,N}, b(s, X_s^i, \mu_s) - b(s, X_s^{i,N}, \mu_s^{X,N})\rangle \leqslant C|X_s^i - X_s^{i,N}|^2 + C|X_s^i - X_s^{i,N}|W^{(2)}(\mu_s, \mu_s^{X,N}).$$

As in [Car16], we introduce the empirical measure constructed from i.i.d. samples of the true solution $\mu_s^N := \frac{1}{N}\sum_{j=1}^N \delta_{X_s^j}$. As $W^{(2)}$ is a metric (see [Vil08, Chapter 6]), we have

$$W^{(2)}(\mu_s, \mu_s^{X,N}) \leqslant W^{(2)}(\mu_s, \mu_s^N) + W^{(2)}(\mu_s^N, \mu_s^{X,N}).$$

Since $\mu_s^N$, $\mu_s^{X,N}$ are empirical measures a standard result for the Wasserstein metric is

$$W^{(2)}(\mu_s^N, \mu_s^{X,N}) \leqslant \Big(\frac{1}{N}\sum_{j=1}^N |X_s^j - X_s^{j,N}|^2\Big)^{1/2}.$$

We leave the other $W^{(2)}$ term for the moment and consider the diffusion coefficient in the time integral. Since $\sigma$ is globally Lipschitz and $W^{(2)}$ for each $a$ (by definition $\sigma_a = \sigma e_a$, with $e_a$ the basis vector, global Lipschitz follows from our norm), we get

$$|\sigma_a(s, X_s^i, \mu_s) - \sigma_a(s, X_s^{i,N}, \mu_s^{X,N})|^2 \\ \leqslant C\big(|\sigma_a(s, X_s^i, \mu_s) - \sigma_a(s, X_s^{i,N}, \mu_s)|^2 + |\sigma_a(s, X_s^{i,N}, \mu_s) - \sigma_a(s, X_s^{i,N}, \mu_s^{X,N})|^2\big) \\ \leqslant C\big(|X_s^i - X_s^{i,N}|^2 + W^{(2)}(\mu_s, \mu_s^{X,N})^2\big) \\ \leqslant C\big(|X_s^i - X_s^{i,N}|^2 + \frac{1}{N}\sum_{j=1}^N |X_s^j - X_s^{j,N}|^2 + W^{(2)}(\mu_s, \mu_s^N)^2\big).$$

One can note this is independent of $a$. The final term to bound is the stochastic integral term. To do this we apply the supremum and expectation opterator to (5.7)

$$\mathbb{E}\Big[\sup_{t\in[0,T]} |X_t^i - X_t^{i,N}|^2\Big] \\ \leqslant C\mathbb{E}\Big[\sup_{t\in[0,T]} \int_0^t |X_s^i - X_s^{i,N}|^2 + |X_s^i - X_s^{i,N}|W^{(2)}(\mu_s, \mu_s^{X,N})\,\mathrm{d}s\Big] \\ + \mathbb{E}\Big[\sup_{t\in[0,T]} \int_0^t 2\langle X_s^i - X_s^{i,N}, (\sigma(s, X_s^i, \mu_s) - \sigma(s, X_s^{i,N}, \mu_s^{X,N}))\,\mathrm{d}W_s^i\rangle\Big] \\ + Cl\mathbb{E}\Big[\sup_{t\in[0,T]} \int_0^t |X_s^i - X_s^{i,N}|^2 + \frac{1}{N}\sum_{j=1}^N |X_s^j - X_s^{j,N}|^2 + W^{(2)}(\mu_s, \mu_s^N)^2\,\mathrm{d}s\Big]. \tag{5.8}$$

For the stochastic integral,

$$\begin{aligned}
&\mathbb{E}\Big[\sup_{t\in[0,T]}\int_0^t 2\langle X_s^i - X_s^{i,N}, (\sigma(s,X_s^i,\mu_s) - \sigma(s,X_s^{i,N},\mu_s^{X,N}))\,\mathrm{d}W_s^i\rangle\Big] \\
&\qquad\leqslant \mathbb{E}\Big[\sup_{t\in[0,T]}\Big|\int_0^t 2\langle X_s^i - X_s^{i,N}, (\sigma(s,X_s^i,\mu_s) - \sigma(s,X_s^{i,N},\mu_s^{X,N}))\,\mathrm{d}W_s^i\rangle\Big|\Big] \\
&\qquad\leqslant C\mathbb{E}\Big[\Big(\int_0^T \Big(\sum_{a=1}^{l}|\sigma_a(s,X_s^i,\mu_s) - \sigma_a(s,X_s^{i,N},\mu_s^{X,N})|^2\Big)|X_s^i - X_s^{i,N}|^2\,\mathrm{d}s\Big)^{1/2}\Big] \\
&\qquad\leqslant \mathbb{E}\Big[\Big(\sup_{t\in[0,T]}|X_t^i - X_t^{i,N}|^2 C\int_0^T\sum_{a=1}^{l}|\sigma_a(s,X_s^i,\mu_s) - \sigma_a(s,X_s^{i,N},\mu_s^{X,N})|^2\,\mathrm{d}s\Big)^{1/2}\Big],
\end{aligned}$$

where we have applied Burkholder-Davis-Gundy to remove the stochastic integral. Using Young's inequality $ab \leqslant a^2/2 + b^2/2$ we can bound this term by

$$\mathbb{E}\Big[\frac{1}{2}\sup_{t\in[0,T]}|X_t^i - X_t^{i,N}|^2 + \frac{C}{2}\int_0^T\sum_{a=1}^{l}|\sigma_a(s,X_s^i,\mu_s) - \sigma_a(s,X_s^{i,N},\mu_s^{X,N})|^2\,\mathrm{d}s\Big].$$

Substituting into (5.8) yields,

$$\begin{aligned}
&\mathbb{E}\Big[\sup_{t\in[0,T]}|X_t^i - X_t^{i,N}|^2\Big] \\
&\quad\leqslant C\mathbb{E}\Big[\sup_{t\in[0,T]}\int_0^t |X_s^i - X_s^{i,N}|^2 + |X_s^i - X_s^{i,N}|W^{(2)}(\mu_s,\mu_s^{X,N})\,\mathrm{d}s\Big] \\
&\qquad + \mathbb{E}\Big[\frac{1}{2}\sup_{t\in[0,T]}|X_t^i - X_t^{i,N}|^2 + \frac{C}{2}\int_0^T\sum_{a=1}^{l}|\sigma_a(s,X_s^i,\mu_s) - \sigma_a(s,X_s^{i,N},\mu_s^{X,N})|^2\,\mathrm{d}s\Big] \\
&\qquad + C\mathbb{E}\Big[\sup_{t\in[0,T]}\int_0^t |X_s^i - X_s^{i,N}|^2 + \frac{1}{N}\sum_{j=1}^{N}|X_s^j - X_s^{j,N}|^2 + W^{(2)}(\mu_s,\mu_s^N)^2\,\mathrm{d}s\Big].
\end{aligned}$$

Taking the $\frac{1}{2}\sup_{t\in[0,T]}|X_t^i - X_t^{i,N}|^2$ to the other side, noting that the supremum value over the integrals is $t = T$ and using the bound for the difference in $\sigma$ we obtain,

$$\begin{aligned}
\mathbb{E}\left[\sup_{t\in[0,T]}|X_t^i - X_t^{i,N}|^2\right] &\leqslant C\mathbb{E}\left[\int_0^T |X_s^i - X_s^{i,N}|^2 + |X_s^i - X_s^{i,N}|W^{(2)}(\mu_s,\mu_s^{X,N})\,\mathrm{d}s\right] \\
&\quad + C\mathbb{E}\left[\int_0^T |X_s^i - X_s^{i,N}|^2 + \frac{1}{N}\sum_{j=1}^{N}|X_s^j - X_s^{j,N}|^2 + W^{(2)}(\mu_s,\mu_s^N)^2\,\mathrm{d}s\right].
\end{aligned}$$

To deal with the summation term, observe that since all $j$ are identically distributed,

$$\mathbb{E}\Big[\frac{1}{N}\sum_{j=1}^{N}|X_s^j - X_s^{j,N}|^2\Big] = \mathbb{E}\big[|X_s^i - X_s^{i,N}|^2\big].$$

Therefore, applying Young's inequality to $|X_s^i - X_s^{i,N}|W^{(2)}(\mu_s,\mu_s^N)$ and taking the supremum over $i$,

$$\begin{aligned}
\sup_{1\leqslant i\leqslant N}\mathbb{E}\left[\sup_{t\in[0,T]}|X_t^i - X_t^{i,N}|^2\right] &\leqslant C\int_0^T \sup_{1\leqslant i\leqslant N}\mathbb{E}\left[|X_s^i - X_s^{i,N}|^2\right] + \mathbb{E}[W^{(2)}(\mu_s,\mu_s^N)^2]\,\mathrm{d}s \\
&\leqslant C\int_0^T \mathbb{E}\left[W^{(2)}(\mu_s,\mu_s^N)^2\right]\mathrm{d}s,
\end{aligned}$$

where the final step follows from Gronwall's inequality. At this point, one could conclude a pathwise propagation of chaos result, see [Car16, Lemma 1.9], however, here we are interested in the rate of convergence. We use the improved version [CD17, Theorem 5.8] of the classical convergence result [RR98, Chapter 10.2]. Provided $X^i_\cdot \in L^p_\cdot(\mathbb{R}^d)$ for any $p > 4$, which follows from [dRST19, Theorem 3.3] then for any $s$,

$$\mathbb{E}\left[W^{(2)}(\mu_s, \mu_s^N)^2\right] \leqslant C \begin{cases} N^{-1/2} & \text{if } d < 4, \\ N^{-1/2}\log(N) & \text{if } d = 4, \\ N^{-2/d} & \text{if } d > 4. \end{cases}$$

Using the result in Theorem 5.3 with our assumption then completes the proof. ■

### 5.4.2 Proof of Explicit Convergence

We prove Proposition 5.5 by establishing first a few auxiliary results. To keep expressions compact we introduce

$$\Delta X_s^{i,N,M} := X_s^{i,N} - X_s^{i,N,M} \quad \text{for } s \in [0,T].$$

Further, we will use throughout and without mentioning the following result

$$\mathbb{E}\Big[\frac{1}{N}\sum_{j=1}^{N}|\Delta X_s^{j,N,M}|^2\Big] = \mathbb{E}\left[|\Delta X_s^{i,N,M}|^2\right] = \sup_{1\leqslant j\leqslant N}\mathbb{E}\left[|\Delta X_s^{j,N,M}|^2\right],$$

which holds because for every $i$ the RVs are identically distributed.

**Lemma 5.15**
*Suppose Assumption 5.1 and 5.2 are fulfilled, then there exists a constant $C$ which is independent of $N$ and $M$ such that*

$$\langle X_t^{i,N,M}, b_M\left(t, X_t^{i,N,M}, \mu_t^{X,N,M}\right)\rangle \leqslant C\left(1 + |X_t^{i,N,M}|^2 + \frac{1}{N}\sum_{j=1}^{N}|X_t^{j,N,M}|^2\right)$$

*and*

$$\left|\sigma\left(t, X_t^{i,N,M}, \mu_t^{X,N,M}\right)\right|^2 \leqslant C\left(1 + |X_t^{i,N,M}|^2 + \frac{1}{N}\sum_{j=1}^{N}|X_t^{j,N,M}|^2\right).$$

*Proof.* First, observe for $x, x' \in \mathbb{R}^d$ and $\mu \in \mathcal{P}(\mathbb{R}^d)$ that

$$\begin{aligned}
&\langle x - x', b_M(t,x,\mu) - b_M(t,x',\mu)\rangle \\
&\quad = \frac{\langle x - x', b(t,x,\mu) - b(t,x',\mu)\rangle}{1 + M^{-\alpha}|b(t,x,\mu)|} + \langle x - x', \frac{b(t,x',\mu)(|b(t,x,\mu)| - |b(t,x',\mu))|}{(M^\alpha + |b(t,x,\mu)|)(M^\alpha + |b(t,x',\mu)|)}\rangle \\
&\quad \leqslant \frac{\langle x - x', b(t,x,\mu) - b(t,x',\mu)\rangle}{1 + M^{-\alpha}|b(t,x,\mu)|} + |x - x'|^2 + \left|\frac{|b(t,x',\mu)|^2 - |b(t,x,\mu)||b(t,x',\mu)|}{|b(t,x,\mu)||b(t,x',\mu)|}\right|^2.
\end{aligned}$$

Assuming without loss of generality (otherwise just switch $x$ and $x'$) that $|b(t,x,\mu)| \geqslant |b(t,x',\mu)|$ we get by Assumption 5.1

$$\langle x - x', b_M(t,x,\mu) - b_M(t,x',\mu)\rangle \leqslant (L_b + 1)|x - x'|^2 + 1.$$

Similarily we obtain for all $x \in \mathbb{R}^d$ and $\mu, \mu' \in \mathcal{P}(\mathbb{R}^d)$

$$|b_M(t,x,\mu) - b_M(t,x,\mu')| \leqslant |b(t,x,\mu) - b(t,x,\mu')| + 1 \leqslant LW^{(2)}(\mu,\mu') + 1.$$

Using this, we have

$$\begin{aligned}
&\langle X_t^{i,N,M}, b_M\left(t, X_t^{i,N,M}, \mu_t^{X,N,M}\right)\rangle \\
&\quad = \langle X_t^{i,N,M} - 0, b_M\left(t, X_t^{i,N,M}, \mu_t^{X,N,M}\right) - b_M\left(t,0,\delta_0\right)\rangle + \langle X_t^{i,N,M}, b_M\left(t,0,\delta_0\right)\rangle \\
&\quad \leqslant L_b |X_t^{i,N,M}|^2 + 2|X_t^{i,N,M}|^2 + LW^{(2)}\left(\mu_t^{X,N,M}, \delta_0\right)^2 + 1 + |X_t^{i,N,M}|^2 + |b_M\left(t,0,\delta_0\right)|^2 \\
&\quad \leqslant C\Big(1 + |X_t^{i,N,M}|^2 + \frac{1}{N}\sum_{j=1}^N |X_t^{j,N,M}|^2\Big)
\end{aligned}$$

by the 1/2-Hölder-continuity in Assumption 5.2. Again with Assumption 5.1 and 5.2 we get

$$\begin{aligned}
\left|\sigma\left(t, X_t^{i,N,M}, \mu_t^{X,N,M}\right)\right|^2 &\leqslant \left|\sigma\left(t, X_t^{i,N,M}, \mu_t^{X,N,M}\right) - \sigma\left(t,0,\delta_0\right)\right|^2 + \left|\sigma\left(t,0,\delta_0\right)\right|^2 \\
&\leqslant L_\sigma\left(|X_t^{i,N,M}|^2 + W^{(2)}\left(\mu_t^{X,N,M}, \delta_0\right)^2\right) + \left|\sigma\left(t,0,\delta_0\right)\right|^2 \\
&\leqslant C\left(1 + |X_t^{i,N,M}|^2 + \frac{1}{N}\sum_{j=1}^N |X_t^{j,N,M}|^2\right).
\end{aligned}$$

■

**Lemma 5.16**
*Suppose Assumption 5.1 and 5.2 are fulfilled and $X_0 \in L^2(\mathbb{R}^d)$, then there exists a constant $C$ which is independent of $N$ and $M$ such that*

$$\sup_{M \geqslant 1} \sup_{1 \leqslant i \leqslant N} \sup_{0 \leqslant t \leqslant T} \mathbb{E}\left[|X_t^{i,N,M}|^2\right] < C.$$

*Proof.* Applying Itô's formula and restructuring the terms gives

$$\begin{aligned}
&\left|X_t^{i,N,M}\right|^2 \\
&\quad = \left|X_0^i\right|^2 + \int_0^t 2\langle X_{\kappa(s)}^{i,N,M}, b_M\left(\kappa(s), X_{\kappa(s)}^{i,N,M}, \mu_{\kappa(s)}^{X,N,M}\right)\rangle + \sum_{a=1}^l \left|\sigma_a\left(\kappa(s), X_{\kappa(s)}^{i,N,M}, \mu_{\kappa(s)}^{X,N,M}\right)\right|^2 \mathrm{d}s \\
&\qquad + \int_0^t 2\langle X_s^{i,N,M}, \sigma\left(\kappa(s), X_{\kappa(s)}^{i,N,M}, \mu_{\kappa(s)}^{X,N,M}\right) \mathrm{d}W_s^i\rangle \\
&\qquad + \int_0^t 2\langle X_s^{i,N,M} - X_{\kappa(s)}^{i,N,M}, b_M\left(\kappa(s), X_{\kappa(s)}^{i,N,M}, \mu_{\kappa(s)}^{X,N,M}\right)\rangle \mathrm{d}s.
\end{aligned}$$

We start with the expectations of the last term (using $|s-\kappa(s)| \leqslant T/M$ and $\alpha \in (0,1/2]$)

$$\begin{aligned}
&\left|\mathbb{E}\left[\int_0^t \langle X_s^{i,N,M} - X_{\kappa(s)}^{i,N,M}, b_M\left(\kappa(s), X_{\kappa(s)}^{i,N,M}, \mu_{\kappa(s)}^{X,N,M}\right)\rangle \mathrm{d}s\right]\right| \\
&= \Big|\mathbb{E}\Big[\int_0^t \langle \int_{\kappa(s)}^s b_M\left(\kappa(r), X_{\kappa(r)}^{i,N,M}, \mu_{\kappa(r)}^{X,N,M}\right)\mathrm{d}r \\
&\quad + \int_{\kappa(s)}^s \sigma\left(\kappa(r), X_{\kappa(r)}^{i,N,M}, \mu_{\kappa(r)}^{X,N,M}\right)\mathrm{d}W_r^i, b_M\left(\kappa(s), X_{\kappa(s)}^{i,N,M}, \mu_{\kappa(s)}^{X,N,M}\right)\rangle \mathrm{d}s\Big]\Big| \\
&= \Bigg|\sum_{k=0}^{M-1}\int_{t_k}^{t_{k+1}} \mathbb{1}_{\{s\leqslant t\}}\mathbb{E}\Bigg[\Big\langle b_M\Big(\kappa(s), X_{\kappa(s)}^{i,N,M}, \mu_{\kappa(s)}^{X,N,M}\Big), \\
&\quad \mathbb{E}\Big[\int_{t_k}^s b_M\Big(\kappa(r), X_{\kappa(r)}^{i,N,M}, \mu_{\kappa(r)}^{X,N,M}\Big)\mathrm{d}r + \int_{t_k}^s \sigma\Big(\kappa(r), X_{\kappa(r)}^{i,N,M}, \mu_{\kappa(r)}^{X,N,M}\Big)\mathrm{d}W_r^i\Big|\mathcal{F}_{t_k}\Big]\Big\rangle\Bigg]\mathrm{d}s\Bigg| \\
&\leqslant \mathbb{E}\left[\int_0^t \left|b_M\left(\kappa(s), X_{\kappa(s)}^{i,N,M}, \mu_{\kappa(s)}^{X,N,M}\right)\right| \int_{\kappa(s)}^s \left|b_M\left(\kappa(r), X_{\kappa(r)}^{i,N,M}, \mu_{\kappa(r)}^{X,N,M}\right)\right|\mathrm{d}r\ \mathrm{d}s\right] \\
&\leqslant tTM^{2\alpha-1} \\
&\leqslant tT.
\end{aligned}$$

Putting this together and using Lemma 5.15 we obtain

$$\begin{aligned}
\mathbb{E}\big[|X_t^{i,N,M}|^2\big] &\leqslant \mathbb{E}\big[|X_0^i|^2\big] + C\Big(1 + \mathbb{E}\Big[\int_0^t \left|X_{\kappa(s)}^{i,N,M}\right|^2 + \frac{1}{N}\sum_{j=1}^N \left|X_{\kappa(s)}^{j,N,M}\right|^2 \mathrm{d}s\Big]\Big) \\
&\leqslant \mathbb{E}\big[|X_0^i|^2\big] + C\Big(1 + \int_0^t \sup_{0\leqslant u\leqslant s}\mathbb{E}\Big[\left|X_u^{i,N,M}\right|^2 + \frac{1}{N}\sum_{j=1}^N \left|X_u^{j,N,M}\right|^2\Big]\mathrm{d}s\Big),
\end{aligned}$$

which furthermore yields

$$\sup_{1\leqslant i\leqslant N}\sup_{0\leqslant u\leqslant t}\mathbb{E}\left[\left|X_u^{i,N,M}\right|^2\right] \leqslant C\left(1 + \mathbb{E}\left[|X_0|^2\right] + \int_0^t \sup_{1\leqslant i\leqslant N}\sup_{0\leqslant u\leqslant s}\mathbb{E}\left[\left|X_u^{i,N,M}\right|^2\right]\mathrm{d}s\right) < \infty,$$

and hence by Gronwall's lemma

$$\sup_{1\leqslant i\leqslant N}\sup_{0\leqslant u\leqslant t}\mathbb{E}\left[\left|X_u^{i,N,M}\right|^2\right] < C,$$

where $C$ is a constant which is independent of $N$ and $M$. ■

**Lemma 5.17**
*If Assumption 5.1 and 5.2 are fulfilled and $X_0 \in L^2(\mathbb{R}^d)$, then for all $p \in (0,2]$ we have*

$$\sup_{1\leqslant i\leqslant N}\sup_{0\leqslant t\leqslant T}\mathbb{E}\left[\left|X_t^{i,N,M} - X_{\kappa(t)}^{i,N,M}\right|^p\right] \leqslant CM^{-p/2}, \tag{5.9}$$

*and*

$$\sup_{1\leqslant i\leqslant N}\sup_{0\leqslant t\leqslant T}\mathbb{E}\left[\left|X_t^{i,N,M} - X_{\kappa(t)}^{i,N,M}\right|^p \left|b_M\left(\kappa(t), X_{\kappa(t)}^{i,N,M}, \mu_{\kappa(t)}^{X,N,M}\right)\right|^p\right] \leqslant C, \tag{5.10}$$

*where $C$ is a positive constant independent of $N$ and $M$. Furthermore, if for $p > 2$*

$$\sup_{M\geqslant 1}\sup_{1\leqslant i\leqslant N}\mathbb{E}\left[\sup_{0\leqslant t\leqslant T}\left|X_t^{i,N,M}\right|^p\right] < \infty,$$

*then the estimates (5.9) and (5.10) hold for those $p$ as well.*

*Proof.* We obtain for any $p \geqslant 2$

$$\left|\int_{\kappa(t)}^{t} b_M\left(\kappa(s), X_{\kappa(s)}^{i,N,M}, \mu_{\kappa(s)}^{X,N,M}\right) \mathrm{d}s\right|^p \leqslant T^p M^{-p/2}, \tag{5.11}$$

since $|b_M| \leqslant M^\alpha$ and $\alpha \leqslant 1/2$. It is easy to see that in the case of $p \in (0, 2]$

$$\begin{aligned}
&\mathbb{E}\left[\left|X_t^{i,N,M} - X_{\kappa(t)}^{i,N,M}\right|^p\right] \\
&\quad \leqslant \mathbb{E}\Big[\Big|\int_{\kappa(t)}^{t} b_M\big(\kappa(s), X_{\kappa(s)}^{i,N,M}, \mu_{\kappa(s)}^{X,N,M}\big)\,\mathrm{d}s + \int_{\kappa(t)}^{t} \sigma\big(\kappa(s), X_{\kappa(s)}^{i,N,M}, \mu_{\kappa(s)}^{X,N,M}\big)\,\mathrm{d}W_s^i\Big|^2\Big]^{\frac{p}{2}} \\
&\quad \leqslant 2^{p/2}\mathbb{E}\Big[\Big|\int_{\kappa(t)}^{t} b_M\big(\kappa(s), X_{\kappa(s)}^{i,N,M}, \mu_{\kappa(s)}^{X,N,M}\big)\,\mathrm{d}s\Big|^2 + \Big|\int_{\kappa(t)}^{t} \sigma\big(\kappa(s), X_{\kappa(s)}^{i,N,M}, \mu_{\kappa(s)}^{X,N,M}\big)\,\mathrm{d}W_s^i\Big|^2\Big]^{\frac{p}{2}},
\end{aligned}$$

and due to Itô's isometry , Lemma 5.15 and Lemma 5.16 for $C$ independent of $M$ and $i$

$$\begin{aligned}
&\mathbb{E}\left[\left|\int_{\kappa(t)}^{t} \sigma\left(\kappa(s), X_{\kappa(s)}^{i,N,M}, \mu_{\kappa(s)}^{X,N,M}\right) \mathrm{d}W_s^i\right|^2\right] \\
&\qquad \leqslant \mathbb{E}\left[\int_{\kappa(t)}^{t} K\left(1 + \left|X_{\kappa(s)}^{i,N,M}\right|^2 + \frac{1}{N}\sum_{j=1}^{N}\left|X_{\kappa(s)}^{j,N,M}\right|^2\right) \mathrm{d}s\right] \\
&\qquad \leqslant \sup_{1\leqslant i\leqslant N}\sup_{s\in[\kappa(t),t]} \mathbb{E}\left[\frac{T}{M}K\left(1 + \left|X_s^{i,N,M}\right|^2 + \left|X_s^{i,N,M}\right|^2\right)\right] \leqslant CM^{-1},
\end{aligned}$$

which gives, combined with (5.11), that

$$\sup_{0\leqslant t\leqslant T} \mathbb{E}\left[\left|X_t^{i,N,M} - X_{\kappa(t)}^{i,N,M}\right|^p\right] \leqslant CM^{-p/2}, \quad \text{for all } p \in (0, 2].$$

If additionally $\sup_{M\geqslant 1}\sup_{1\leqslant i\leqslant N} \mathbb{E}\big[\sup_{0\leqslant t\leqslant T}|X_t^{i,N,M}|^p\big] < \infty$ for some $p > 2$, then

$$\begin{aligned}
&\mathbb{E}\left[\left|X_t^{i,N,M} - X_{\kappa(t)}^{i,N,M}\right|^p\right] \\
&\quad \leqslant C\mathbb{E}\left[\Big|\int_{\kappa(t)}^{t} b_M\big(\kappa(s), X_{\kappa(s)}^{i,N,M}, \mu_{\kappa(s)}^{X,N,M}\big)\,\mathrm{d}s\Big|^p + \Big|\int_{\kappa(t)}^{t} \sigma\big(\kappa(s), X_{\kappa(s)}^{i,N,M}, \mu_{\kappa(s)}^{X,N,M}\big)\,\mathrm{d}W_s^i\Big|^p\right] \\
&\quad \leqslant C\mathbb{E}\left[T^p M^{p/2} + \Big|\int_{\kappa(t)}^{t} \sigma\big(\kappa(s), X_{\kappa(s)}^{i,N,M}, \mu_{\kappa(s)}^{X,N,M}\big)^2\,\mathrm{d}s\Big|^{p/2}\right],
\end{aligned}$$

by the estimate (5.11) and the Burkholder-Davis-Gundy inequality. Since furthermore,

$$\begin{aligned}
&\mathbb{E}\left[\left|\int_{\kappa(t)}^{t} \sigma\left(\kappa(s), X_{\kappa(s)}^{i,N,M}, \mu_{\kappa(s)}^{X,N,M}\right)^2 \mathrm{d}s\right|^{p/2}\right] \\
&\qquad \leqslant \mathbb{E}\left[\left(\frac{T}{M}\right)^{p/2} \sup_{s\in[\kappa(t),t]} K\left(1 + |X_s^{i,N,M}|^p + \left(\frac{1}{N}\sum_{j=1}^{N}|X_s^{j,N,M}|^2\right)^{p/2}\right)\right] \\
&\qquad \leqslant \left(\frac{T}{M}\right)^{p/2} K\left(1 + \mathbb{E}\left[\sup_{0\leqslant t\leqslant T}\left|X_t^{i,N,M}\right|^p\right] + \sup_{1\leqslant j\leqslant N}\mathbb{E}\left[\sup_{0\leqslant t\leqslant T}\left|X_t^{j,N,M}\right|^p\right]\right) \\
&\qquad \leqslant CM^{-p/2},
\end{aligned}$$

we get the desired result here as well.

Finally, using the above results and that $\alpha \leqslant 1/2$, we obtain for any $p \geqslant 0$ for which $\mathbb{E}[|X_t^{i,N,M} - X_{\kappa(t)}^{i,N,M}|^p] \leqslant CM^{-p/2}$, that

$$\mathbb{E}\left[\left|X_t^{i,N,M} - X_{\kappa(t)}^{i,N,M}\right|^p \left|b_M\big(\kappa(t), X_{\kappa(t)}^{i,N,M}, \mu_{\kappa(t)}^{X,N,M}\big)\right|^p\right] \leqslant \mathbb{E}\left[\left|X_t^{i,N,M} - X_{\kappa(t)}^{i,N,M}\right|^p\right] M^{p\alpha} \leqslant C,$$

holds for any $t \in [0,T]$ and $1 \leqslant i \leqslant N$, which completes the proof. ∎

**Lemma 5.18**
*Suppose that Assumption 5.1 and 5.2 are fulfilled, then for every $p \geqslant 2$ with $X_0 \in L^p(\mathbb{R}^d)$ there exists a constant $C$ such that*

$$\sup_{M\geqslant 1} \sup_{1\leqslant i\leqslant N} \mathbb{E}\left[\sup_{0\leqslant t\leqslant T} \left|X_t^{i,N,M}\right|^p\right] < C.$$

*Proof.* Define $\hat{p} \geqslant 2$ such that $\mathbb{E}[|X_0|^{\hat{p}}] < \infty$ and note that for $p < 2$ Lemma 5.17 yields immediately the result.

We use an inductive argument and start with $p = 2$. In every step we set $q = 2p \wedge \hat{p}$. By Itô's formula and Lemma 5.15 we have

$$\begin{aligned}\left|X_t^{i,N,M}\right|^2 \leqslant \left|X_0^{i,N,M}\right|^2 &+ \int_0^t \left|X_s^{i,N,M} - X_{\kappa(s)}^{i,N,M}\right| \left|b_M\big(\kappa(s), X_{\kappa(s)}^{i,N,M}, \mu_{\kappa(s)}^{X,N,M}\big)\right| \mathrm{d}s \\ &+ \int_0^t C\Big(1 + \left|X_{\kappa(s)}^{i,N,M}\right|^2 + \frac{1}{N}\sum_{j=1}^N \left|X_{\kappa(s)}^{j,N,M}\right|^2\Big) \mathrm{d}s \\ &+ \left|\int_0^t X_u^{i,N,M} \sigma\big(\kappa(u), X_{\kappa(u)}^{i,N,M}, \mu_{\kappa(u)}^{X,N,M}\big) \mathrm{d}W_u^i\right|.\end{aligned}$$

With the inequality $|a+b|^{q/2} \leqslant C(|a|^{q/2} + |b|^{q/2})$ and Jensen's inequality we therefore obtain

$$\begin{aligned}\mathbb{E}\big[\sup_{0\leqslant s\leqslant t} |X_s^{i,N,M}|^q\big] \leqslant C\Big(1 &+ \mathbb{E}\Big[|X_0^{i,N,M}|^q\Big] + \int_0^t \mathbb{E}\Big[|X_{\kappa(s)}^{i,N,M}|^q\Big] \mathrm{d}s \\ &+ \int_0^t \mathbb{E}\Big[|X_s^{i,N,M} - X_{\kappa(s)}^{i,N,M}|^{q/2} |b_M\big(\kappa(s), X_{\kappa(s)}^{i,N,M}, \mu_{\kappa(s)}^{X,N,M}\big)|^{q/2}\Big] \mathrm{d}s \\ &+ \mathbb{E}\Big[\sup_{0\leqslant s\leqslant t} |\int_0^s X_u^{i,N,M} \sigma\big(\kappa(u), X_{\kappa(u)}^{i,N,M}, \mu_{\kappa(u)}^{X,N,M}\big) \mathrm{d}W_u^i|^{q/2}\Big]\Big).\end{aligned}$$

The application of the Burkholder-Davis-Gundy inequality and Lemma 5.17 with[2] $q/2$ yields

$$\begin{aligned}\mathbb{E}\left[\sup_{0\leqslant s\leqslant t} |X_s^{i,N,M}|^q\right] \leqslant C\Bigg(1 &+ \mathbb{E}\left[\left|X_0^{i,N,M}\right|^q\right] + \int_0^t \mathbb{E}\left[\sup_{0\leqslant u\leqslant s} |X_u^{i,N,M}|^q\right] \mathrm{d}s \\ &+ \mathbb{E}\left[\left(\int_0^t |X_s^{i,N,M}|^2 \left|\sigma\left(\kappa(s), X_{\kappa(s)}^{i,N,M}, \mu_{\kappa(s)}^{X,N,M}\right)\right|^2 \mathrm{d}s\right)^{q/4}\right]\Bigg),\end{aligned}$$

where $C$ denotes in each case a constant that is independent of $M$. With Young's inequality in

[2] Observe that Lemma 5.17 holds for the current value of $p$ and since $q = 2p \wedge \hat{p}$ it implies that it holds for $q/2$.

the form $ab \leqslant \frac{1}{2C}a^2 + \frac{C}{2}b^2$, Hölder's inequality and the estimate for $\sigma$ we have

$$\begin{aligned}
&\mathbb{E}\left[\sup_{0\leqslant s\leqslant t}|X_s^{i,N,M}|^q\right]\\
&\leqslant C\left(1+\mathbb{E}\left[\left|X_0^{i,N,M}\right|^q\right]+\int_0^t\mathbb{E}\left[\sup_{0\leqslant u\leqslant s}|X_u^{i,N,M}|^q\right]\mathrm{d}s+\frac{1}{2C}\mathbb{E}\left[\sup_{0\leqslant s\leqslant t}|X_s^{i,N,M}|^q\right]\right.\\
&\qquad\left.+\frac{C}{2}\mathbb{E}\left[\int_0^t\left|\sigma\left(\kappa(s),X_{\kappa(s)}^{i,N,M},\mu_{\kappa(s)}^{X,N,M}\right)\right|^q\mathrm{d}s\right]\right)\\
&\leqslant C\Bigg(1+\mathbb{E}\left[\left|X_0^{i,N,M}\right|^q\right]+\int_0^t\mathbb{E}\left[\sup_{0\leqslant u\leqslant s}|X_u^{i,N,M}|^q\right]\mathrm{d}s+\frac{1}{2C}\mathbb{E}\left[\sup_{0\leqslant s\leqslant t}|X_s^{i,N,M}|^q\right]\\
&\qquad+\frac{C}{2}\int_0^t\mathbb{E}\Bigg[\sup_{0\leqslant u\leqslant s}K\Big(1+|X_u^{i,N,M}|^q+\Big(\frac{1}{N}\sum_{j=1}^N|X_u^{j,N,M}|^2\Big)^{q/2}\Big)\Bigg]\mathrm{d}s\Bigg).
\end{aligned}$$

Taking the $\frac{1}{2}\mathbb{E}[\sup_{0\leqslant s\leqslant t}|X_s^{i,N,M}|^q]$ term to the LHS taking the sup over $i$ on both sides we obtain

$$\sup_{1\leqslant i\leqslant N}\mathbb{E}\left[\sup_{0\leqslant s\leqslant t}|X_s^{i,N,M}|^q\right]\leqslant C\left(1+\mathbb{E}\left[|X_0^{i,N,M}|^q\right]+\int_0^t\sup_{1\leqslant i\leqslant N}\mathbb{E}\left[\sup_{0\leqslant u\leqslant s}|X_u^{i,N,M}|^q\right]\mathrm{d}s\right)<\infty,$$

and thus the application of Gronwall's lemma yields that

$$\sup_{1\leqslant i\leqslant N}\mathbb{E}\Big[\sup_{0\leqslant t\leqslant T}|X_t^{i,N,M}|^q\Big]<C, \tag{5.12}$$

for some positive constant $C$ which depends on $\mathbb{E}[|X_0^i|^q]$ but is independent of $N$ and $M$.

Since (5.12) is proven for $q$ we can set $p=q$ and use this result in the next step of the iteration. Since the new $q$ is at most twice as much as $p$, Lemma 5.17 can again be applied for $q/2$. This iteration gets repeated until $q=\hat{p}$. ∎

Now we can complete the proof of Proposition 5.5.

*Proof of Proposition 5.5.* Using Itô's formula we observe,

$$\begin{aligned}
\left|\Delta X_t^{i,N,M}\right|^2=&\int_0^t 2\langle\Delta X_s^{i,N,M},\left(b\left(s,X_s^{i,N},\mu_s^{X,N}\right)-b_M\left(\kappa(s),X_{\kappa(s)}^{i,N,M},\mu_{\kappa(s)}^{X,N,M}\right)\right)\rangle\,\mathrm{d}s\\
&+\sum_{a=1}^l\int_0^t|\sigma_a\left(s,X_s^{i,N},\mu_s^{X,N}\right)-\sigma_a\left(\kappa(s),X_{\kappa(s)}^{i,N,M},\mu_{\kappa(s)}^{X,N,M}\right)|^2\,\mathrm{d}s\\
&+\int_0^t 2\langle\Delta X_s^{i,N,M},\left(\sigma\left(s,X_s^{i,N},\mu_s^{X,N}\right)-\sigma\left(\kappa(s),X_{\kappa(s)}^{i,N,M},\mu_{\kappa(s)}^{X,N,M}\right)\right)\mathrm{d}W_s^i\rangle.
\end{aligned}$$

Furthermore observe that

$$\begin{aligned}
&\langle X_s^{i,N}-X_s^{i,N,M},b\big(s,X_s^{i,N},\mu_s^{X,N}\big)-b_M\big(\kappa(s),X_{\kappa(s)}^{i,N,M},\mu_{\kappa(s)}^{X,N,M}\big)\rangle\\
&=\langle\Delta X_s^{i,N,M},b\big(s,X_s^{i,N},\mu_s^{X,N}\big)-b\big(s,X_s^{i,N,M},\mu_s^{X,N}\big)\rangle\\
&\quad+\langle\Delta X_s^{i,N,M},b\big(s,X_s^{i,N,M},\mu_s^{X,N}\big)-b\big(s,X_s^{i,N,M},\mu_s^{X,N,M}\big)\rangle\\
&\quad+\langle\Delta X_s^{i,N,M},b\big(s,X_s^{i,N,M},\mu_s^{X,N,M}\big)-b\big(\kappa(s),X_s^{i,N,M},\mu_s^{X,N,M}\big)\rangle\\
&\quad+\langle\Delta X_s^{i,N,M},b\big(\kappa(s),X_s^{i,N,M},\mu_s^{X,N,M}\big)-b\big(\kappa(s),X_{\kappa(s)}^{i,N,M},\mu_s^{X,N,M}\big)\rangle\\
&\quad+\langle\Delta X_s^{i,N,M},b\big(\kappa(s),X_{\kappa(s)}^{i,N,M},\mu_s^{X,N,M}\big)-b\big(\kappa(s),X_{\kappa(s)}^{i,N,M},\mu_{\kappa(s)}^{X,N,M}\big)\rangle\\
&\quad+\langle\Delta X_s^{i,N,M},b\big(\kappa(s),X_{\kappa(s)}^{i,N,M},\mu_{\kappa(s)}^{X,N,M}\big)-b_M\big(\kappa(s),X_{\kappa(s)}^{i,N,M},\mu_{\kappa(s)}^{X,N,M}\big)\rangle,
\end{aligned}$$

where we estimate every term on the right hand side as follows. Due to Assumption 5.1 we have

$$\langle \Delta X_s^{i,N,M}, b\left(s, X_s^{i,N}, \mu_s^{X,N}\right) - b\left(s, X_s^{i,N,M}, \mu_s^{X,N}\right)\rangle \leqslant L_b \left|\Delta X_s^{i,N,M}\right|^2,$$

and

$$\begin{aligned}\langle \Delta X_s^{i,N,M}, b\left(s, X_s^{i,N,M}, \mu_s^{X,N}\right) - b\left(s, X_s^{i,N,M}, \mu_s^{X,N,M}\right)\rangle &\leqslant \left|\Delta X_s^{i,N,M}\right| \left|W^{(2)}\left(\mu_s^{X,N}, \mu_s^{X,N,M}\right)\right| \\ &\leqslant \left|\Delta X_s^{i,N,M}\right| \frac{1}{\sqrt{N}}\Big(\sum_{j=1}^N |\Delta X_s^{j,N,M}|^2\Big)^{1/2} \\ &\leqslant \frac{1}{2}\left|\Delta X_s^{i,N,M}\right|^2 + \frac{1}{2}\frac{1}{N}\sum_{j=1}^N |\Delta X_s^{j,N,M}|^2,\end{aligned}$$

and with Assumption 5.2

$$\begin{aligned}&\langle \Delta X_s^{i,N,M}, b\left(s, X_s^{i,N,M}, \mu_s^{X,N,M}\right) - b\left(\kappa(s), X_s^{i,N,M}, \mu_s^{X,N,M}\right)\rangle \\ &\qquad \leqslant C\left|\Delta X_s^{i,N,M}\right| |s-\kappa(s)|^{1/2} \leqslant \frac{1}{2}\left|\Delta X_s^{i,N,M}\right|^2 + CM^{-1}.\end{aligned}$$

Further,

$$\begin{aligned}&\Big\langle \Delta X_s^{i,N,M}, b\big(\kappa(s), X_s^{i,N,M}, \mu_s^{X,N,M}\big) - b\big(\kappa(s), X_{\kappa(s)}^{i,N,M}, \mu_s^{X,N,M}\big)\Big\rangle \\ &\qquad \leqslant \frac{1}{2}\left|\Delta X_s^{i,N,M}\right|^2 + \frac{1}{2}\left|b\left(\kappa(s), X_s^{i,N,M}, \mu_s^{X,N,M}\right) - b\left(\kappa(s), X_{\kappa(s)}^{i,N,M}, \mu_s^{X,N,M}\right)\right|^2,\end{aligned}$$

which we can furthermore dominate by using the polynomial growth of $b$ with rate $q$, Cauchy-Schwarz, Lemma 5.18 and Lemma 5.17, to have

$$\begin{aligned}&\mathbb{E}\left[\sup_{u\in[0,t]}\int_0^u \left|b\left(\kappa(s), X_s^{i,N,M}, \mu_s^{X,N,M}\right) - b\left(\kappa(s), X_{\kappa(s)}^{i,N,M}, \mu_s^{X,N,M}\right)\right|^2 \mathrm{d}s\right] \\ &\qquad \leqslant \int_0^t \mathbb{E}\left[L\left(1+\left|X_s^{i,N,M}\right|^q + \left|X_{\kappa(s)}^{i,N,M}\right|^q\right)^2 \left|X_s^{i,N,M} - X_{\kappa(s)}^{i,N,M}\right|^2\right]\mathrm{d}s \\ &\qquad \leqslant \int_0^t \sqrt{\mathbb{E}\left[L\left(1+\left|X_s^{i,N,M}\right|^q + \left|X_{\kappa(s)}^{i,N,M}\right|^q\right)^4\right]\mathbb{E}\left[\left|X_s^{i,N,M} - X_{\kappa(s)}^{i,N,M}\right|^4\right]}\,\mathrm{d}s \\ &\qquad \leqslant \int_0^t \sqrt{CM^{-2}}\,\mathrm{d}s \leqslant CM^{-1}\end{aligned}$$

since

$$\sup_{M\geqslant 1}\sup_{1\leqslant i\leqslant N}\mathbb{E}\left[\sup_{0\leqslant t\leqslant T}\left|X_t^{i,N,M}\right|^{4q}\right] \leqslant 1 + \sup_{M\geqslant 1}\sup_{1\leqslant i\leqslant N}\mathbb{E}\left[\sup_{0\leqslant t\leqslant T}\left|X_t^{i,N,M}\right|^{4(1+q)}\right] < \infty.$$

Again Assumption 5.1 yields

$$\begin{aligned}&\langle \Delta X_s^{i,N,M}, b\left(\kappa(s), X_{\kappa(s)}^{i,N,M}, \mu_s^{X,N,M}\right) - b\left(\kappa(s), X_{\kappa(s)}^{i,N,M}, \mu_{\kappa(s)}^{X,N,M}\right)\rangle \\ &\qquad \leqslant \left|\Delta X_s^{i,N,M}\right| \frac{L}{\sqrt{N}}\Big(\sum_{j=1}^N |X_s^{j,N,M} - X_{\kappa(s)}^{j,N,M}|^2\Big)^{1/2} \\ &\qquad \leqslant \frac{1}{2}\left|\Delta X_s^{i,N,M}\right|^2 + \frac{1}{2}\frac{L^2}{N}\sum_{j=1}^N \left|X_s^{j,N,M} - X_{\kappa(s)}^{j,N,M}\right|^2,\end{aligned}$$

and the definition of $b_M$ together with $|a - \frac{a}{1+M^{-\alpha}|a|}| = |a\frac{M^{-\alpha}|a|}{1+M^{-\alpha}|a|}| \leqslant |a|^2 M^{-\alpha}$ that

$$\begin{aligned}
&\langle \Delta X_s^{i,N,M}, b\big(\kappa(s), X_{\kappa(s)}^{i,N,M}, \mu_{\kappa(s)}^{X,N,M}\big) - b_M\big(\kappa(s), X_{\kappa(s)}^{i,N,M}, \mu_{\kappa(s)}^{X,N,M}\big)\rangle \\
&\qquad \leqslant \frac{1}{2}\left|\Delta X_s^{i,N,M}\right|^2 + \frac{1}{2}\left|b\big(\kappa(s), X_{\kappa(s)}^{i,N,M}, \mu_{\kappa(s)}^{X,N,M}\big) - b_M\big(\kappa(s), X_{\kappa(s)}^{i,N,M}, \mu_{\kappa(s)}^{X,N,M}\big)\right|^2 \\
&\qquad \leqslant \frac{1}{2}\left|\Delta X_s^{i,N,M}\right|^2 + \frac{1}{2}M^{-2\alpha}|b\left(\kappa(s), X_{\kappa(s)}^{i,N,M}, \mu_{\kappa(s)}^{X,N,M}\right)|^4 \\
&\qquad \leqslant \frac{1}{2}\left|\Delta X_s^{i,N,M}\right|^2 + CM^{-2\alpha}\Big(1 + \left|X_{\kappa(s)}^{i,N,M}\right|^{4(1+q)} + \Big(\frac{1}{N}\sum_{j=1}^{N}\left|X_{\kappa(s)}^{j,N,M}\right|^2\Big)^2\Big),
\end{aligned}$$

where $q$ is again the polynomial growth rate of $b$. Also the Burkholder-Davis-Gundy inequality yields

$$\begin{aligned}
&\mathbb{E}\left[\sup_{u\in[0,t]}\int_0^u 2\langle \Delta X_s^{i,N,M}, \Big(\sigma\big(s, X_s^{i,N}, \mu_s^{X,N}\big) - \sigma\big(\kappa(s), X_{\kappa(s)}^{i,N,M}, \mu_{\kappa(s)}^{X,N,M}\big)\Big)\,\mathrm{d}W_s^i\rangle\right] \\
&\qquad \leqslant \mathbb{E}\left[\Big(C\int_0^t \Big(\sum_{a=1}^{l}|\sigma_a(s, X_s^{i,N}, \mu_s^{X,N}) - \sigma_a(s, X_{\kappa(s)}^{i,N,M}, \mu_{\kappa(s)}^{X,N,M})|^2\Big)|\Delta X_s^{i,N,M}|^2\,\mathrm{d}s\Big)^{\frac{1}{2}}\right] \\
&\qquad \leqslant \mathbb{E}\Big[\frac{1}{2}\sup_{u\in[0,t]}|\Delta X_u^{i,N,M}|^2 + C\int_0^t\sum_{a=1}^{l}\left|\sigma_a(s, X_s^{i,N}, \mu_s^{X,N}) - \sigma_a(\kappa(s), X_{\kappa(s)}^{i,N,M}, \mu_{\kappa(s)}^{X,N,M})\right|^2\mathrm{d}s\Big].
\end{aligned}$$

and

$$\begin{aligned}
&\left|\sigma_a\left(s, X_s^{i,N}, \mu_s^{X,N}\right) - \sigma_a\left(\kappa(s), X_{\kappa(s)}^{i,N,M}, \mu_{\kappa(s)}^{X,N,M}\right)\right|^2 \\
&\qquad \leqslant C\,|s-\kappa(s)| + C\left|X_s^{i,N} - X_{\kappa(s)}^{i,N,M}\right|^2 + CW^{(2)}\left(\mu_s^{X,N}, \mu_{\kappa(s)}^{X,N,M}\right)^2 \\
&\qquad \leqslant CM^{-1} + C\left|X_s^{i,N} - X_{\kappa(s)}^{i,N,M}\right|^2 + \frac{C}{N}\sum_{j=1}^{N}\left|X_s^{j,N} - X_{\kappa(s)}^{j,N,M}\right|^2 \\
&\qquad \leqslant CM^{-1} + C\left|X_s^{i,N} - X_{\kappa(s)}^{i,N,M}\right|^2 + \frac{C}{N}\sum_{j=1}^{N}\left(\left|\Delta X_s^{j,N,M}\right|^2 + \left|X_s^{j,N,M} - X_{\kappa(s)}^{j,N,M}\right|^2\right).
\end{aligned}$$

By putting those estimates together we obtain

$$\begin{aligned}
&\mathbb{E}\left[\sup_{0\leqslant u\leqslant t}\left|\Delta X_u^{i,N,M}\right|^2\right] \\
&\quad \leqslant C\mathbb{E}\Big[M^{-1} + \int_0^t\left|\Delta X_s^{i,N,M}\right|^2 + \frac{1}{N}\sum_{j=1}^{N}\left|X_s^{j,N,M} - X_{\kappa(s)}^{j,N,M}\right|^2 + M^{-1} + \frac{1}{N}\sum_{j=1}^{N}\left|\Delta X_s^{j,N,M}\right|^2 \\
&\qquad + \left|X_s^{i,N,M} - X_{\kappa(s)}^{i,N,M}\right|^2 + M^{-2\alpha}\Big(1 + \left|X_{\kappa(s)}^{i,N,M}\right|^{4(1+q)}\Big) + M^{-2\alpha}\Big(\frac{1}{N}\sum_{j=1}^{N}\left|X_{\kappa(s)}^{j,N,M}\right|^2\Big)^2\mathrm{d}s\Big] \\
&\qquad + \mathbb{E}\Big[\frac{1}{2}\sup_{u\in[0,t]}|\Delta X_u^{i,N,M}|^2\Big]
\end{aligned}$$

and therefore

$$\mathbb{E}\left[\sup_{0\leqslant u\leqslant t}\left|\Delta X_u^{i,N,M}\right|^2\right] \leqslant C\Big(\int_0^t\mathbb{E}\Big[\sup_{0\leqslant u\leqslant s}\left|\Delta X_u^{i,N,M}\right|^2\Big]\mathrm{d}s + M^{-2\alpha} + M^{-1}\Big),$$

by Lemma 5.18 and since $X^{i,N}$ are identically distributed and $X^{i,N,M}$ are identically distributed for all $i \in \{1, \ldots, N\}$. This estimate holds for every $i$ hence we can insert $\sup_{1 \leqslant i \leqslant N}$ on both sides giving

$$\sup_{1\leqslant i\leqslant N} \mathbb{E}\left[\sup_{0\leqslant u\leqslant t} \left|\Delta X_u^{i,N,M}\right|^2\right] \leqslant C\Big(\int_0^t \sup_{1\leqslant i\leqslant N} \mathbb{E}\Big[\sup_{0\leqslant u\leqslant s} \left|\Delta X_u^{i,N,M}\right|^2\Big] \mathrm{d}s + M^{-2\alpha} + M^{-1}\Big) < \infty,$$

and finally by Gronwall's lemma (using that $\alpha = 1/2$),

$$\sup_{1\leqslant i\leqslant N} \mathbb{E}\left[\sup_{0\leqslant u\leqslant t} \left|X_u^{i,N} - X_u^{i,N,M}\right|^2\right] \leqslant CM^{-1}.$$

■

### 5.4.3 Proof of Implicit Convergence

The main goal here is to prove Proposition 5.9. We loosely follow [MS13], however, due to the extra dependencies on time and measure and further allowing for random initial conditions we require more refined arguments. We take $N$ as some fixed positive integer. Before considering the implicit scheme, let us make a remark and show a result on the particle system (5.2).

**Remark 5.19** (Monotone Growth)
The combination of Assumption 5.1, 5.2 and H1, imply the monotone growth condition. Namely, there exist constants $\alpha, \beta \in \mathbb{R}$ such that $\forall t \in [0,T], \mu \in \mathcal{P}_2(\mathbb{R}^d)$ with $l$ being the dimension of the Brownian motion,

$$\langle x, b(t,x,\mu)\rangle + \frac{1}{2}\sum_{a=1}^{l} |\sigma_a(t,x,\mu)|^2 \leqslant \alpha + \beta|x|^2 \quad \forall x \in \mathbb{R}^d.$$

**Proposition 5.20**
*Let Assumption 5.1, 5.2 and H1 (in Assumption 5.8) hold, further, let $X_0 \in L^2(\mathbb{R}^d)$. Then the following bounds hold,*

$$\sup_{1\leqslant i\leqslant N} \mathbb{E}[|X_T^{i,N}|^2] \leqslant \big(\mathbb{E}[|X_0|^2] + 2\alpha T\big)\exp(2\beta T),$$

*and for $\tau_m^i = \inf\{t \geqslant 0 : |X_t^{i,N}| > m\}$ we have*

$$\sup_{1\leqslant i\leqslant N} \mathbf{P}(\tau_m^i \leqslant T) \leqslant \frac{1}{m^2}\big(\mathbb{E}[|X_0|^2] + 2\alpha T\big)\exp(2\beta T).$$

*Proof.* Firstly, let us consider the stopped process $X^{i,N}_{T\wedge\tau_m^i}$. Applying Itô to the square of this process and taking expectations yields

$$\begin{aligned}\mathbb{E}[|X^{i,N}_{T\wedge\tau_m^i}|^2] &= \mathbb{E}[|X_0^i|^2] + \mathbb{E}\Big[\int_0^{T\wedge\tau_m^i} 2\langle X_s^{i,N}, b(s, X_s^{i,N}, \mu_s^{X,N})\rangle + \sum_{a=1}^{l} |\sigma_a(s, X_s^{i,N}, \mu_s^{X,N})|^2 \,\mathrm{d}s\Big] \\ &\leqslant \mathbb{E}[|X_0^i|^2] + 2\alpha T + \int_0^T 2\beta\mathbb{E}[|X^{i,N}_{s\wedge\tau_m^i}|^2]\,\mathrm{d}s \leqslant \big(\mathbb{E}[|X_0^i|^2] + 2\alpha T\big)e^{2\beta T},\end{aligned}$$

where we have used the growth and stopping condition to remove the martingale term, then Remark 5.19, uniform boundedness of $b$ in the measure component and Gronwall's inequality to obtain the result.

Noting that the following lower bound also holds,

$$\mathbb{E}[|X^{i,N}_{T\wedge\tau^i_m}|^2] \geqslant m^2\mathbf{P}(\tau^i_m \leqslant T), \quad \text{we obtain} \quad \mathbf{P}(\tau^i_m \leqslant T) \leqslant \frac{1}{m^2}\big(\mathbb{E}[|X^i_0|^2] + 2\alpha T\big)\exp(2\beta T).$$

Further, since $\lim_{m\to\infty}|X^{i,N}_{T\wedge\tau^i_m}| = |X^{i,N}_T|$, we obtain by Fatou's lemma,

$$\mathbb{E}[|X^{i,N}_T|^2] \leqslant \liminf_{m\to\infty}\mathbb{E}[|X^{i,N}_{T\wedge\tau^i_m}|^2] \leqslant \big(\mathbb{E}[|X^i_0|^2] + 2\alpha T\big)\exp(2\beta T).$$

The result then follows by noting that $\mathbb{E}[|X^i_0|^2] = \mathbb{E}[|X_0|^2]$ and hence the bounds are independent of $i$, so we obtain the result for the supremum over $i$. ■

Let us now return to the implicit scheme. At each time step $t_i$ and for each particle $i$ one needs to solve the fixed point equation

$$\tilde{X}^{i,N,M}_{t_{k+1}} - b\Big(t_k, \tilde{X}^{i,N,M}_{t_{k+1}}, \tilde{\mu}^{X,N,M}_{t_k}\Big)h = \tilde{X}^{i,N,M}_{t_k} + \sigma\Big(t_k, \tilde{X}^{i,N,M}_{t_k}, \tilde{\mu}^{X,N,M}_{t_k}\Big)\Delta W^i_{t_k}.$$

This leads us to consider a function $F$

$$F(t,x,\mu) := x - b(t,x,\mu)h. \tag{5.13}$$

For the implicit scheme to have a solution the function $F$ must have a unique inverse. The following lemma is crucial in proving convergence of the implicit scheme.

**Lemma 5.21**
*Let Assumption 5.1, 5.2 and H1 (in Assumption 5.8) hold and fix $h^* < 1/\max(L_b, 2\beta)$. Further, let $0 < h \leqslant h^*$ and take any $t \in [0,T]$ and $\mu \in \mathcal{P}_2(\mathbb{R}^d)$ fixed. Then for all $y \in \mathbb{R}^d$, there exists a unique $x$ such that $F(t,x,\mu) = y$. Hence the fixed point problem in* (5.5) *is well defined.*
*Moreover, for all $t \in [0,T]$ and $\mu \in \mathcal{P}_2(\mathbb{R}^d)$ the following bound holds,*

$$|x|^2 \leqslant (1-2h\beta)^{-1}(|F(t,x,\mu)|^2 + 2h\alpha), \tag{5.14}$$

*and for any $k \geqslant 1$ the following recursive bound holds,*

$$\begin{aligned} &|F(t_k, \tilde{X}^{i,N,M}_{t_{k+1}}, \tilde{\mu}^{X,N,M}_{t_k})|^2 \\ &\qquad \leqslant |F(t_{k-1}, \tilde{X}^{i,N,M}_{t_k}, \tilde{\mu}^{X,N,M}_{t_{k-1}})|^2 + \Big(\sum_{a=1}^{l}|\sigma_a(t_k, \tilde{X}^{i,N,M}_{t_k}, \tilde{\mu}^{X,N,M}_{t_k})||\big(\Delta W^i_{t_k}\big)_a|\Big)^2 \\ &\qquad\quad + 2h\alpha + 2h\beta|\tilde{X}^{i,N,M}_{t_k}|^2 + 2\langle \tilde{X}^{i,N,M}_{t_k}, \sigma(t_k, \tilde{X}^{i,N,M}_{t_k}, \tilde{\mu}^{X,N,M}_{t_k})\Delta W^i_{t_k}\rangle, \end{aligned} \tag{5.15}$$

*where $\big(\Delta W^i_{t_k}\big)_a$ is the $a$th entry of the vector.*

*Proof.* Let us first prove there exists a unique solution to (5.13), in the sense that for all $t \in [0,T]$ and $\mu \in \mathcal{P}_2(\mathbb{R}^d)$ fixed, then there exists a unique $x \in \mathbb{R}^d$ such that $F(t,x,\mu) = y$ for a given $y \in \mathbb{R}^d$, provided $0 < h < h^*$. This is a classical problem considered in [Zei90, p.557] or see [LdRS15, p.2596], which requires $F$ to be continuous, monotone and coercive (in $x$). The continuity of $b$ yields that of $F$. For the monotonicity of $F$, we have

$$\begin{aligned} \langle x - x', F(t,x,\mu) - F(t,x',\mu)\rangle &= |x-x'|^2 - \langle x - x', b(t,x,\mu)h - b(t,x',\mu)h\rangle \\ &\geqslant |x-x'|^2(1 - L_b h), \end{aligned}$$

and provided $h < 1/L_b$, the final constant is strictly positive. Coercivity follows similarly by the monotone growth condition in $b$,

$$\langle x, F(t,x,\mu)\rangle \geqslant |x|^2 - h(\alpha + \beta|x|^2),$$

therefore,

$$\lim_{|x|\to\infty} \frac{\langle x, F(t,x,\mu)\rangle}{|x|} = \infty, \quad \text{for } h < 1/\beta.$$

Hence $F(t,x,\mu) = y$ has a unique solution for $F$ defined in (5.13) and therefore the numerical scheme (5.5) is well defined.

To show $x$ is bounded by $F(\cdot,x,\cdot)$, again fix some $t \in [0,T]$ and $\mu \in \mathcal{P}_2(\mathbb{R}^d)$. Then,

$$\begin{aligned}|F(t,x,\mu)|^2 &= |x|^2 - 2\langle x, b(t,x,\mu)\rangle h + |b(t,x,\mu)|^2 h^2 \\ &\geqslant |x|^2 - 2\langle x, b(t,x,\mu)\rangle h \geqslant (1-2h\beta)|x|^2 - 2h\alpha,\end{aligned}$$

by Remark 5.19. Since $h < 1/(2\beta)$, we obtain

$$|x|^2 \leqslant (1-2h\beta)^{-1}(|F(t,x,\mu)|^2 + 2h\alpha).$$

This result is useful since it holds for all $t \in [0,T]$ and $\mu \in \mathcal{P}_2(\mathbb{R}^d)$. For the recursive bound it is useful to note

$$\begin{aligned}F(t_k, \tilde{X}^{i,N,M}_{t_{k+1}}, \tilde{\mu}^{X,N,M}_{t_k}) &= \tilde{X}^{i,N,M}_{t_{k+1}} - b(t_k, \tilde{X}^{i,N,M}_{t_{k+1}}, \tilde{\mu}^{X,N,M}_{t_k})h \\ &= \tilde{X}^{i,N,M}_{t_k} + \sigma(t_k, \tilde{X}^{i,N,M}_{t_k}, \tilde{\mu}^{X,N,M}_{t_k})\Delta W^i_{t_k} \\ &= F(t_{k-1}, \tilde{X}^{i,N,M}_{t_k}, \tilde{\mu}^{X,N,M}_{t_{k-1}}) + b(t_{k-1}, \tilde{X}^{i,N,M}_{t_k}, \tilde{\mu}^{X,N,M}_{t_{k-1}})h \\ &\quad + \sigma(t_k, \tilde{X}^{i,N,M}_{t_k}, \tilde{\mu}^{X,N,M}_{t_k})\Delta W^i_{t_k}.\end{aligned} \tag{5.16}$$

This recursion is only valid for $k \geqslant 1$ due to the appearance of $t_{k-1}$. Using this relation observe the following,

$$\begin{aligned}|F(t_k, \tilde{X}^{i,N,M}_{t_{k+1}}, \tilde{\mu}^{X,N,M}_{t_k})|^2 &= |F(t_{k-1}, \tilde{X}^{i,N,M}_{t_k}, \tilde{\mu}^{X,N,M}_{t_{k-1}})|^2 + |b(t_{k-1}, \tilde{X}^{i,N,M}_{t_k}, \tilde{\mu}^{X,N,M}_{t_{k-1}})|^2 h^2 \\ &\quad + |\sigma(t_k, \tilde{X}^{i,N,M}_{t_k}, \tilde{\mu}^{X,N,M}_{t_k})\Delta W^i_{t_k}|^2 \\ &\quad + 2\langle F(t_{k-1}, \tilde{X}^{i,N,M}_{t_k}, \tilde{\mu}^{X,N,M}_{t_{k-1}}), b(t_{k-1}, \tilde{X}^{i,N,M}_{t_k}, \tilde{\mu}^{X,N,M}_{t_{k-1}})\rangle h \\ &\quad + 2\langle F(t_{k-1}, \tilde{X}^{i,N,M}_{t_k}, \tilde{\mu}^{X,N,M}_{t_{k-1}}) \\ &\qquad + b(t_{k-1}, \tilde{X}^{i,N,M}_{t_k}, \tilde{\mu}^{X,N,M}_{t_{k-1}})h, \sigma(t_k, \tilde{X}^{i,N,M}_{t_k}, \tilde{\mu}^{X,N,M}_{t_k})\Delta W^i_{t_k}\rangle.\end{aligned}$$

We now look to bound these various terms. By definition of $F$,

$$\begin{aligned}&2\langle F(t_{k-1}, \tilde{X}^{i,N,M}_{t_k}, \tilde{\mu}^{X,N,M}_{t_{k-1}}), b(t_{k-1}, \tilde{X}^{i,N,M}_{t_k}, \tilde{\mu}^{X,N,M}_{t_{k-1}})\rangle\, h + |b(t_{k-1}, \tilde{X}^{i,N,M}_{t_k}, \tilde{\mu}^{X,N,M}_{t_{k-1}})|^2 h^2 \\ &\qquad \leqslant 2\langle \tilde{X}^{i,N,M}_{t_k}, b(t_{k-1}, \tilde{X}^{i,N,M}_{t_k}, \tilde{\mu}^{X,N,M}_{t_{k-1}})\rangle\, h \leqslant 2h\alpha + 2h\beta|\tilde{X}^{i,N,M}_{t_k}|^2.\end{aligned}$$

Similarly,

$$\begin{aligned}&2\langle F(t_{k-1}, \tilde{X}^{i,N,M}_{t_k}, \tilde{\mu}^{X,N,M}_{t_{k-1}}) + b(t_{k-1}, \tilde{X}^{i,N,M}_{t_k}, \tilde{\mu}^{X,N,M}_{t_{k-1}})h, \sigma(t_k, \tilde{X}^{i,N,M}_{t_k}, \tilde{\mu}^{X,N,M}_{t_k})\Delta W^i_{t_k}\rangle \\ &\qquad = 2\langle \tilde{X}^{i,N,M}_{t_k}, \sigma(t_k, \tilde{X}^{i,N,M}_{t_k}, \tilde{\mu}^{X,N,M}_{t_k})\Delta W^i_{t_k}\rangle.\end{aligned}$$

In order to obtain the desired form we note

$$\sigma(t,x,\mu)\Delta W_t = \sum_{a=1}^{l} \sigma_a(t,x,\mu)(\Delta W_t)_a.$$

Crucially $(\Delta W_t)_a$ is a scalar and standard properties of norms yield,

$$|\sigma(t_k, \tilde{X}_{t_k}^{i,N,M}, \tilde{\mu}_{t_k}^{X,N,M})\Delta W_{t_k}^i| \leqslant \sum_{a=1}^{l} |\sigma_a(t_k, \tilde{X}_{t_k}^{i,N,M}, \tilde{\mu}_{t_k}^{X,N,M})||\big(\Delta W_{t_k}^i\big)_a|.$$

The bound on $F$ then follows immediately from these results. ■

Let us now show the first moment bound result. As is standard with implicit schemes we firstly do this up to a stopping time, hence we define

$$\lambda_m^i = \inf\{k : \ |\tilde{X}_{t_k}^{i,N,M}| > m\}. \tag{5.17}$$

One should note that this stopping time does not actually bound $\tilde{X}$ at that point $i$, the best one can do is bound the previous point i.e. for $\lambda_m^i > 0$, we have $|\tilde{X}_{\lambda_m^i - 1}^{i,N,M}| \leqslant m$.

**Lemma 5.22**
*Let Assumption 5.1, 5.2 and H1 (in Assumption 5.8) hold and fix $h^* < 1/\max(L_b, 2\beta)$. Then for any $p \geqslant 2$ such that $\mathbb{E}[|X_0|^p] = C(p) < \infty$, we also have,*

$$\sup_{1\leqslant i\leqslant N} \mathbb{E}\big[|\tilde{X}_{t_k}^{i,N,M}|^p \mathbb{1}_{\{k\leqslant \lambda_m^i\}}\big] \leqslant C(p,m) \quad \forall k \leqslant M \ \ \textit{and}\ 0 < h \leqslant h^*.$$

Using standard notation, $C(a)$ denotes a constant that can depend on variable $a$.

*Proof.* As it turns out the function $F$ in (5.13) gives us a useful bound. From (5.16) we obtain,

$$|F(t_k, \tilde{X}_{t_{k+1}}^{i,N,M}, \tilde{\mu}_{t_k}^{X,N,M})|^p \leqslant 2^{p-1}\big(|\tilde{X}_{t_k}^{i,N,M}|^p + |\sigma(t_k, \tilde{X}_{t_k}^{i,N,M}, \tilde{\mu}_{t_k}^{X,N,M})\Delta W_{t_k}^i|^p\big).$$

Hence, multiplying with the indicator and taking expected values yields,

$$\begin{aligned}&\mathbb{E}[|F(t_k, \tilde{X}_{t_{k+1}}^{i,N,M}, \tilde{\mu}_{t_k}^{X,N,M})|^p \mathbb{1}_{\{k+1\leqslant \lambda_m^i\}}] \\ &\qquad \leqslant C(p)\Big(m^p + \mathbb{E}\big[|\sigma(t_k, \tilde{X}_{t_k}^{i,N,M}, \tilde{\mu}_{t_k}^{X,N,M})\Delta W_{t_k}^i|^p \mathbb{1}_{\{k+1\leqslant \lambda_m^i\}}\big]\Big).\end{aligned}$$

Then we estimate

$$\begin{aligned}&\mathbb{E}\big[|\sigma(t_k, \tilde{X}_{t_k}^{i,N,M}, \tilde{\mu}_{t_k}^{X,N,M})\Delta W_{t_k}^i|^p \mathbb{1}_{\{k+1\leqslant \lambda_m^i\}}\big] \\ &\qquad \leqslant \sum_{a=1}^{l} \mathbb{E}[|\sigma_a(t_k, \tilde{X}_{t_k}^{i,N,M}, \tilde{\mu}_{t_k}^{X,N,M})|^{2p} \mathbb{1}_{\{k+1\leqslant \lambda_m^i\}}] + \mathbb{E}[|(\Delta W_{t_k}^i)_a|^{2p}].\end{aligned}$$

Using the bounds on each coefficient of $\sigma$, it is straightforward to observe,

$$|\sigma_a(t_k, \tilde{X}_{t_k}^{i,N,M}, \tilde{\mu}_{t_k}^{X,N,M})|^{2p} \leqslant C(p)\big(1 + |\tilde{X}_{t_k}^{i,N,M}|^{2p}\big).$$

Using this bound we obtain,

$$\mathbb{E}[|F(t_k, \tilde{X}_{t_{k+1}}^{i,N,M}, \tilde{\mu}_{t_k}^{X,N,M})|^p \mathbb{1}_{\{k+1\leqslant \lambda_m^i\}}] \leqslant C(p,m).$$

Rewriting the quantity we wish to bound as

$$\mathbb{E}\big[|\tilde{X}_{t_k}^{i,N,M}|^p \mathbb{1}_{\{k\leqslant \lambda_m^i\}}\big] = \mathbb{E}\big[|\tilde{X}_{t_k}^{i,N,M}|^p \mathbb{1}_{\{k\leqslant \lambda_m^i, k>0\}}\big] + \mathbb{E}\big[|\tilde{X}_{t_0}^{i,N,M}|^p \mathbb{1}_{\{k=0,\ \lambda_m^i=0\}}\big] \leqslant C(p,m),$$

where the inequality follows from Estimate (5.14), our bound on $F$, and the assumption that $X_0 \in L^p(\mathbb{R}^d)$. Again, the corresponding bound is independent of the choice of $i$ and hence the result holds for the supremum over $i$. ■

Although the previous bound is useful, the presence of the stopping time is inconvenient. We therefore remove it and show the second moment is bounded.

**Proposition 5.23**
*Let Assumption 5.1, 5.2 and H1 (in Assumption 5.8) hold and fix $h^* < 1/\max(L_b, 2\beta)$. Further assume that $X_0 \in L^4(\mathbb{R}^d)$. Then,*

$$\sup_{1\leqslant i\leqslant N}\ \sup_{h\leqslant h^*}\ \sup_{0\leqslant k\leqslant M} \mathbb{E}[|\tilde{X}_{t_k}^{i,N,M}|^2] \leqslant C.$$

*Proof.* Firstly let us take a nonnegative integer $K$, such that $Kh \leqslant T$. Now let us consider (5.15). One can note that this bound still holds where the $F$ terms are multiplied by $\mathbb{1}_{\{\lambda_m^i>0\}}$ (since both sides are nonnegative and the indicator is bounded above by one). Summing both sides from $k=1$ to $K\wedge\lambda_m^i$, noting that the $F$ terms cancel, we obtain,

$$\begin{aligned}
&|F(t_{K\wedge\lambda_m^i}, \tilde{X}_{t_{(K\wedge\lambda_m^i)+1}}^{i,N,M}, \tilde{\mu}_{t_{K\wedge\lambda_m^i}}^{X,N,M})|^2 \mathbb{1}_{\{\lambda_m^i>0\}} \\
&\quad \leqslant |F(t_0, \tilde{X}_{t_1}^{i,N,M}, \tilde{\mu}_{t_0}^{X,N,M})|^2 \mathbb{1}_{\{\lambda_m^i>0\}} + \sum_{k=1}^{K\wedge\lambda_m^i} \big(2h\alpha + 2h\beta|\tilde{X}_{t_k}^{i,N,M}|^2 \mathbb{1}_{\{\lambda_m^i>0\}}\big) \\
&\qquad + \sum_{k=1}^{K\wedge\lambda_m^i} \Big(\sum_{a=1}^{l} |\sigma_a(t_k, \tilde{X}_{t_k}^{i,N,M}, \tilde{\mu}_{t_k}^{X,N,M})||(\Delta W_{t_k}^i)_a|\Big)^2 \mathbb{1}_{\{\lambda_m^i>0\}} \\
&\qquad + \sum_{k=1}^{K\wedge\lambda_m^i} 2\langle \tilde{X}_{t_k}^{i,N,M}, \sigma(t_k, \tilde{X}_{t_k}^{i,N,M}, \tilde{\mu}_{t_k}^{X,N,M})\Delta W_{t_k}^i\rangle \mathbb{1}_{\{\lambda_m^i>0\}},
\end{aligned}$$

where we use the convention $\sum_{k=1}^{0} \cdot = 0$. Although the stopping time is useful it is not ideal that it appears on the sum. However, for nonnegative terms it is straightforward to take the stopping time into the coefficients and the stochastic term can be rewritten as

$$\begin{aligned}
&\sum_{k=1}^{K\wedge\lambda_m^i} 2\langle \tilde{X}_{t_k}^{i,N,M}, \sigma(t_k, \tilde{X}_{t_k}^{i,N,M}, \tilde{\mu}_{t_k}^{X,N,M})\Delta W_{t_k}^i\rangle \mathbb{1}_{\{\lambda_m^i>0\}} \\
&\qquad = \sum_{k=1}^{K} 2\langle \tilde{X}_{t_k}^{i,N,M}, \sigma(t_k, \tilde{X}_{t_k}^{i,N,M}, \tilde{\mu}_{t_k}^{X,N,M})\Delta W_{t_k}^i\rangle \mathbb{1}_{\{k\leqslant\lambda_m^i\}}.
\end{aligned}$$

Taking expectations and noting, by Lemma 5.22, that $\tilde{X}_{t_k}^{i,N,M}\mathbb{1}_{\{k\leqslant\lambda_m^i\}} \in L_{t_k}^4(\mathbb{R}^d)$ we conclude this term to be a martingale. We therefore obtain the following bound,

$$\begin{aligned}
&\mathbb{E}[|F(t_{K\wedge\lambda_m^i}, \tilde{X}_{t_{(K\wedge\lambda_m^i)+1}}^{i,N,M}, \tilde{\mu}_{t_{K\wedge\lambda_m^i}}^{X,N,M})|^2 \mathbb{1}_{\{\lambda_m^i>0\}}] \\
&\quad \leqslant \mathbb{E}\big[|F(t_0, \tilde{X}_{t_1}^{i,N,M}, \tilde{\mu}_{t_0}^{X,N,M})|^2\big] + 2\alpha T + \sum_{k=1}^{K} 2h\beta \mathbb{E}\big[|\tilde{X}_{t_{k\wedge\lambda_m^i}}^{i,N,M}|^2 \mathbb{1}_{\{\lambda_m^i>0\}}\big] \\
&\qquad + \sum_{k=1}^{K} \mathbb{E}\Big[\Big(\sum_{a=1}^{l} |\sigma_a(t_{k\wedge\lambda_m^i}, \tilde{X}_{t_{k\wedge\lambda_m^i}}^{i,N,M}, \tilde{\mu}_{t_{k\wedge\lambda_m^i}}^{X,N,M})||(\Delta W_{t_{k\wedge\lambda_m^i}}^i)_a|\Big)^2 \mathbb{1}_{\{\lambda_m^i>0\}}\Big].
\end{aligned}$$

The idea is to apply the discrete version of Gronwall's inequality to this (see for example [MPF12, pg. 436] or [MS13, Lemma 3.4]), which requires our bound to be in terms of $F$. Using arguments similar to previous ones

$$\begin{aligned}\mathbb{E}\Big[\Big(\sum_{a=1}^{l}|\sigma_a(t_{k\wedge\lambda_m^i},\tilde{X}^{i,N,M}_{t_{k\wedge\lambda_m^i}},\tilde{\mu}^{X,N,M}_{t_{k\wedge\lambda_m^i}})||(\Delta W^i_{t_{k\wedge\lambda_m^i}})_a|\Big)^2\mathbb{1}_{\{\lambda_m^i>0\}}\Big]&\\ \leqslant C\sum_{a=1}^{l}\mathbb{E}\Big[|\sigma_a(t_{k\wedge\lambda_m^i},\tilde{X}^{i,N,M}_{t_{k\wedge\lambda_m^i}},\tilde{\mu}^{X,N,M}_{t_{k\wedge\lambda_m^i}})|^2|(\Delta W^i_{t_{k\wedge\lambda_m^i}})_a|^2\mathbb{1}_{\{\lambda_m^i>0\}}\Big]&\\ \leqslant C\sum_{a=1}^{l}h\big(1+\mathbb{E}\big[|\tilde{X}^{i,N,M}_{t_{k\wedge\lambda_m^i}}|^2\mathbb{1}_{\{\lambda_m^i>0\}}\big]\big),&\end{aligned}$$

where we have used independence of $\sigma(\cdot)\mathbb{1}_{\{\lambda_m^i>0\}}$ and $\Delta W$ along with the growth bounds on $\sigma$ to obtain the final inequality. Combining this with our previous bounds and appealing again to Lemma 5.21 (to bound $\tilde{X}$ by $F$) we obtain,

$$\begin{aligned}\mathbb{E}[|F(t_{K\wedge\lambda_m^i},\tilde{X}^{i,N,M}_{t_{(K\wedge\lambda_m^i)+1}},\tilde{\mu}^{X,N,M}_{t_{K\wedge\lambda_m^i}})|^2\mathbb{1}_{\{\lambda_m^i>0\}}]&\\ \leqslant \mathbb{E}\big[|F(t_0,\tilde{X}^{i,N,M}_{t_1},\tilde{\mu}^{X,N,M}_{t_0})|^2\big]+C+\sum_{k=1}^{K}Ch\mathbb{E}\big[|\tilde{X}^{i,N,M}_{t_{k\wedge\lambda_m^i}}|^2\mathbb{1}_{\{\lambda_m^i>0\}}\big]&\\ \leqslant \mathbb{E}\big[|F(t_0,\tilde{X}^{i,N,M}_{t_1},\tilde{\mu}^{X,N,M}_{t_0})|^2\big]+C(1+\frac{h}{1-2h\beta})&\\ +\sum_{k=1}^{K}C\frac{h}{1-2h\beta}\mathbb{E}\big[|F(t_{(k\wedge\lambda_m^i)-1},\tilde{X}^{i,N,M}_{t_{k\wedge\lambda_m^i}},\tilde{\mu}^{X,N,M}_{t_{(k\wedge\lambda_m^i)-1}})|^2\mathbb{1}_{\{\lambda_m^i>0\}}\big].&\end{aligned}$$

Applying a discrete version of the Gronwall inequality and noting $\sum_{k=1}^{K}1\leqslant T/h$ yields

$$\begin{aligned}\mathbb{E}[|F(t_{K\wedge\lambda_m^i},\tilde{X}^{i,N,M}_{t_{(K\wedge\lambda_m^i)+1}},\tilde{\mu}^{X,N,M}_{t_{K\wedge\lambda_m^i}})|^2\mathbb{1}_{\{\lambda_m^i>0\}}]&\\ \leqslant \Big(\mathbb{E}\big[|F(t_0,\tilde{X}^{i,N,M}_{t_1},\tilde{\mu}^{X,N,M}_{t_0})|^2\big]+C\big(1+\frac{h}{1-2h\beta}\big)\Big)\exp\Big(\frac{C}{1-2h\beta}\Big).&\end{aligned}$$

Recalling (5.16), we can apply the same arguments as before to obtain the bound

$$\mathbb{E}\big[|F(t_0,\tilde{X}^{i,N,M}_{t_1},\tilde{\mu}^{X,N,M}_{t_0})|^2\big]\leqslant C(1+(1+h)\mathbb{E}[|\tilde{X}^{i,N,M}_{t_0}|^2]).$$

Noting that our bound for $F$ is now independent of $m$, we can use Fatou's lemma to take the limit and obtain (for $K\geqslant 1$),

$$\mathbb{E}[|F(t_K,\tilde{X}^{i,N,M}_{t_{K+1}},\tilde{\mu}^{X,N,M}_{t_K})|^2]\leqslant C\Big(1+(1+h)\mathbb{E}[|\tilde{X}^{i,N,M}_{t_0}|^2]+\frac{h}{1-2h\beta}\Big)\exp\Big(\frac{C}{1-2h\beta}\Big).$$

Again by Lemma 5.21, the LHS of the latter inequality bounds $\tilde{X}^{i,N,M}_{t_{K+1}}$ (with some constant), hence we obtain a bound for $\tilde{X}^{i,N,M}_{t_k}$ for $k\geqslant 2$. By assumption $\tilde{X}^{i,N,M}_{t_0}$ has second moment therefore we need to obtain a bound for $\tilde{X}^{i,N,M}_{t_1}$. This is not difficult to obtain using again that we can bound $\tilde{X}$ as follows,

$$\mathbb{E}\big[|\tilde{X}^{i,N,M}_{t_1}|^2\big]\leqslant\big(1-2h\beta\big)^{-1}\Big(2h\alpha+\mathbb{E}\big[|F(t_0,\tilde{X}^{i,N,M}_{t_1},\tilde{\mu}^{X,N,M}_{t_0})|^2\big]\Big),$$

then we can apply the same bound on $F$ as above.

In order to complete the proof, we also need to show that this bound exists for all $i$ and $0 < h \leqslant h^*$. One can see immediately that all bounds decrease as $h$ decreases, hence the supremum value is to set $h = h^*$, which is also finite since $h^* < 1/(2\beta)$. The supremum over $i$ follows from the fact that all bounds are independent of $i$. ∎

Now that we have established a bound on the second moment, we look to show convergence of this scheme to the true particle system solution. As always with discrete schemes it is beneficial to introduce their continuous counterpart. As it turns out doing it naively for implicit schemes leads to measurability problems, hence one introduces the so-called forward backward scheme

$$\hat{X}^{i,N,M}_{t_{k+1}} = \hat{X}^{i,N,M}_{t_k} + b\left(t_{k-1\vee 0}, \tilde{X}^{i,N,M}_{t_k}, \tilde{\mu}^{X,N,M}_{t_{k-1\vee 0}}\right) h + \sigma\left(t_k, \tilde{X}^{i,N,M}_{t_k}, \tilde{\mu}^{i,N,M}_{t_k}\right) \Delta W^i_{t_k},$$

where $\hat{X}^{i,N,M}_0 = X^i_0$ and $\vee$ denotes the maximum. The scheme's continuous time version is

$$\hat{X}^{i,N,M}_t = X^i_0 + \int_0^t b\left((\kappa(s)-h)\vee 0, \tilde{X}^{i,N,M}_{\kappa(s)}, \tilde{\mu}^{X,N,M}_{(\kappa(s)-h)\vee 0}\right) \mathrm{d}s + \int_0^t \sigma\left(\kappa(s), \tilde{X}^{i,N,M}_{\kappa(s)}, \tilde{\mu}^{i,N,M}_{\kappa(s)}\right) \mathrm{d}W^i_s. \tag{5.18}$$

The first result we present is that the discrete and continuous versions stay close to one another, up to the stopping time (5.17).

**Lemma 5.24**
*Let Assumption 5.1, 5.2 and H1 (in Assumption 5.8) hold and fix $h^* < 1/\max(L_b, 2\beta)$. Further assume $X_0 \in L^{4(q+1)}(\mathbb{R}^d)$. Then for $1 \leqslant p \leqslant 4$ the following holds for $0 < h \leqslant h^*$,*

$$\sup_{1\leqslant i\leqslant N}\sup_{0\leqslant k\leqslant M} \mathbb{E}\big[|\hat{X}^{i,N,M}_{t_k} - \tilde{X}^{i,N,M}_{t_k}|^p \mathbb{1}_{\{k\leqslant \lambda^i_m\}}\big] \leqslant C(m,p)h^p.$$

*Moreover, we also have the following relation between $\hat{X}$ and $F$ for all $1 \leqslant k \leqslant M$,*

$$|\hat{X}^{i,N,M}_{t_k}|^2 \geqslant \frac{1}{2}|F(t_{k-1}, \tilde{X}^{i,N,M}_{t_k}, \tilde{\mu}^{X,N,M}_{t_{k-1}})|^2 - |b(t_0, \tilde{X}^{i,N,M}_{t_0}, \tilde{\mu}^{X,N,M}_{t_0})h|^2. \tag{5.19}$$

*Proof.* To show the first part we start by noting the following useful relation between (5.5) and (5.18), namely for $1 \leqslant k \leqslant M$,

$$\hat{X}^{i,N,M}_{t_k} - \tilde{X}^{i,N,M}_{t_k} = \big(b(t_0, \tilde{X}^{i,N,M}_{t_0}, \tilde{\mu}^{X,N,M}_{t_0}) - b(t_{k-1}, \tilde{X}^{i,N,M}_{t_k}, \tilde{\mu}^{X,N,M}_{t_{k-1}})\big)h.$$

Noting that one can bound

$$|b(t_0, \tilde{X}^{i,N,M}_{t_0}, \tilde{\mu}^{X,N,M}_{t_0}) - b(t_{k-1}, \tilde{X}^{i,N,M}_{t_k}, \tilde{\mu}^{X,N,M}_{t_{k-1}})| \leqslant C\big(1 + |t_k|^{1/2} + |\tilde{X}^{i,N,M}_{t_0}|^{q+1} + |\tilde{X}^{i,N,M}_{t_k}|^{q+1}\big),$$

where we have used the polynomial growth, Hölder-continuity on the coefficient $b$ and in particular Assumption H1. Hence,

$$\begin{aligned}&\mathbb{E}\big[|\hat{X}^{i,N,M}_{t_k} - \tilde{X}^{i,N,M}_{t_k}|^p \mathbb{1}_{\{k\leqslant\lambda^i_m\}}\big] \\ &\qquad\leqslant C(p)h^p\big(1 + |t_k|^{p/2} + \mathbb{E}\big[|\tilde{X}^{i,N,M}_{t_0}|^{p(q+1)}\mathbb{1}_{\{k\leqslant\lambda^i_m\}}\big] + \mathbb{E}\big[|\tilde{X}^{i,N,M}_{t_k}|^{p(q+1)}\mathbb{1}_{\{k\leqslant\lambda^i_m\}}\big]\big).\end{aligned}$$

One observes that the terms on the RHS are bounded by $C(p,m)$ for $p \leqslant 4$ since $X_0 \in L^{4(q+1)}(\mathbb{R}^d)$ and Lemma 5.22. This completes the first part of the proof.

For the second part, recall from the relation between (5.5) and (5.18), one has,

$$\begin{aligned}\hat{X}_{t_k}^{i,N,M} &= b(t_0, \tilde{X}_{t_0}^{i,N,M}, \tilde{\mu}_{t_0}^{X,N,M})h + \tilde{X}_{t_k}^{i,N,M} - b(t_{k-1}, \tilde{X}_{t_k}^{i,N,M}, \tilde{\mu}_{t_{k-1}}^{X,N,M})h \\ &= b(t_0, \tilde{X}_{t_0}^{i,N,M}, \tilde{\mu}_{t_0}^{X,N,M})h + F(t_{k-1}, \tilde{X}_{t_k}^{i,N,M}, \tilde{\mu}_{t_{k-1}}^{X,N,M}).\end{aligned}$$

Using the reverse triangle inequality we obtain,

$$|\hat{X}_{t_k}^{i,N,M}|^2 \geqslant -|b(t_0, \tilde{X}_{t_0}^{i,N,M}, \tilde{\mu}_{t_0}^{X,N,M})h| + |F(t_{k-1}, \tilde{X}_{t_k}^{i,N,M}, \tilde{\mu}_{t_{k-1}}^{X,N,M})|.$$

The result follows from squaring both sides and applying the generalisation of Young's inequality, namely,

$$\begin{aligned}&|b(t_0, \tilde{X}_{t_0}^{i,N,M}, \tilde{\mu}_{t_0}^{X,N,M})h||F(t_{k-1}, \tilde{X}_{t_k}^{i,N,M}, \tilde{\mu}_{t_{k-1}}^{X,N,M})| \\ &\qquad \leqslant |b(t_0, \tilde{X}_{t_0}^{i,N,M}, \tilde{\mu}_{t_0}^{X,N,M})h|^2 + \frac{1}{4}|F(t_{k-1}, \tilde{X}_{t_k}^{i,N,M}, \tilde{\mu}_{t_{k-1}}^{X,N,M})|^2.\end{aligned}$$

■

The next result we wish to present is that both schemes do not blow up in finite time, for this we define a new stopping time,

$$\eta_m^i := \inf\big\{t \geqslant 0 : |\hat{X}_t^{i,N,M}| \geqslant m, \text{ or } |\tilde{X}_{\kappa(t)}^{i,N,M}| > m\big\}.$$

Note in particular that $\eta_m^i$ is smaller than or equal to $\lambda_m^i$ in (5.17).

**Lemma 5.25**
*Let Assumption 5.1, 5.2 and H1 (in Assumption 5.8) hold, fix $h^* < 1/\max(L_b, 2\beta)$ and assume $X_0 \in L^{4(q+1)}(\mathbb{R}^d)$. Then, for any $\varepsilon > 0$, there exists a $m^*$ such that, for any $m \geqslant m^*$ we can find a $h_0^*(m)$ (note the dependence on $m$) so that,*

$$\sup_{1\leqslant i\leqslant N} \mathbf{P}(\eta_m^i < T) \leqslant \varepsilon, \ \text{for any } 0 < h \leqslant h_0^*(m).$$

*Proof.* Note due to the initial condition being random we must be careful with how we set $m$, we shall come back to this later. Let us start by applying Itô to the stopped version of (5.18),

$$\begin{aligned}&|\hat{X}_{T\wedge\eta_m^i}^{i,N,M}|^2 \\ &\quad =|X_0^i|^2 + \int_0^{T\wedge\eta_m^i} 2\langle \hat{X}_s^{i,N,M}, b\left((\kappa(s)-h)\vee 0, \tilde{X}_{\kappa(s)}^{i,N,M}, \tilde{\mu}_{(\kappa(s)-h)\vee 0}^{X,N,M}\right)\rangle \\ &\qquad + \sum_{a=1}^{l} |\sigma_a\left(\kappa(s), \tilde{X}_{\kappa(s)}^{i,N,M}, \tilde{\mu}_{\kappa(s)}^{i,N,M}\right)|^2 \,\mathrm{d}s + \int_0^{T\wedge\eta_m^i} 2\langle \hat{X}_s^{i,N,M}, \sigma\left(\kappa(s), \tilde{X}_{\kappa(s)}^{i,N,M}, \tilde{\mu}_{\kappa(s)}^{i,N,M}\right) \mathrm{d}W_s^i\rangle.\end{aligned}$$

We now look to bound the various integrands. Firstly one can observe

$$\begin{aligned}&\langle \hat{X}_t^{i,N,M}, b\big((\kappa(s)-h)\vee 0, \tilde{X}_{\kappa(s)}^{i,N,M}, \tilde{\mu}_{(\kappa(s)-h)\vee 0}^{X,N,M}\big)\rangle + \sum_{a=1}^{l} |\sigma_a\big(\kappa(s), \tilde{X}_{\kappa(s)}^{i,N,M}, \tilde{\mu}_{\kappa(s)}^{i,N,M}\big)|^2 \\ &\quad = \langle \hat{X}_t^{i,N,M} - \tilde{X}_{\kappa(s)}^{i,N,M}, b\big((\kappa(s)-h)\vee 0, \tilde{X}_{\kappa(s)}^{i,N,M}, \tilde{\mu}_{(\kappa(s)-h)\vee 0}^{X,N,M}\big)\rangle \\ &\qquad + \langle \tilde{X}_{\kappa(s)}^{i,N,M}, b\big((\kappa(s)-h)\vee 0, \tilde{X}_{\kappa(s)}^{i,N,M}, \tilde{\mu}_{(\kappa(s)-h)\vee 0}^{X,N,M}\big)\rangle + \sum_{a=1}^{l} |\sigma_a\big(\kappa(s), \tilde{X}_{\kappa(s)}^{i,N,M}, \tilde{\mu}_{\kappa(s)}^{i,N,M}\big)|^2 \\ &\quad \leqslant C|\hat{X}_t^{i,N,M} - \tilde{X}_{\kappa(s)}^{i,N,M}|(1 + |\tilde{X}_{\kappa(s)}^{i,N,M}|^{q+1}) + 2\alpha + \beta|\tilde{X}_{\kappa(s)}^{i,N,M}|^2,\end{aligned}$$

where we used Cauchy-Schwarz, polynomial growth bound, Hölder-continuity and monotone growth to obtain the final inequality.

Taking expectations and noting that due to the stopping time the stochastic integral is square integrable and hence a martingale, we obtain,

$$\begin{aligned}&\mathbb{E}[|\hat{X}^{i,N,M}_{T\wedge\eta^i_m}|^2]\\&\quad\leqslant \mathbb{E}[|X^i_0|^2]+\mathbb{E}\Big[\int_0^{T\wedge\eta^i_m} C|\hat{X}^{i,N,M}_s-\tilde{X}^{i,N,M}_{\kappa(s)}|(1+|\tilde{X}^{i,N,M}_{\kappa(s)}|^{q+1})+2\alpha+\beta|\tilde{X}^{i,N,M}_{\kappa(s)}|^2\,\mathrm{d}s\Big].\end{aligned}$$

To proceed we note the following, $|\tilde{X}^{i,N,M}_{\kappa(s)}|^2\leqslant 2(|\tilde{X}^{i,N,M}_{\kappa(s)}-\hat{X}^{i,N,M}_s|^2+|\hat{X}^{i,N,M}_s|^2)$ and also that

$$\int_0^{T\wedge\eta^i_m}|\hat{X}^{i,N,M}_s-\tilde{X}^{i,N,M}_{\kappa(s)}|^2\,\mathrm{d}s\leqslant C(m)\int_0^{T\wedge\eta^i_m}|\hat{X}^{i,N,M}_s-\tilde{X}^{i,N,M}_{\kappa(s)}|\,\mathrm{d}s,$$

where we used the fact that the stopping time ensures $\tilde{X}$ and $\hat{X}$ are $\leqslant m$ for $s<\eta^i_m$ and $s=\eta^i_m$ has measure zero. The same reasoning also implies,

$$\int_0^{T\wedge\eta^i_m}C|\hat{X}^{i,N,M}_s-\tilde{X}^{i,N,M}_{\kappa(s)}|(1+|\tilde{X}^{i,N,M}_{\kappa(s)}|^{q+1})\,\mathrm{d}s\leqslant C(m)\int_0^{T\wedge\eta^i_m}|\hat{X}^{i,N,M}_s-\tilde{X}^{i,N,M}_{\kappa(s)}|\,\mathrm{d}s.$$

Hence the following result holds,

$$\mathbb{E}[|\hat{X}^{i,N,M}_{T\wedge\eta^i_m}|^2]\leqslant \mathbb{E}[|X^i_0|^2]+C\mathbb{E}\Big[\int_0^{T\wedge\eta^i_m}C(m)|\hat{X}^{i,N,M}_s-\tilde{X}^{i,N,M}_{\kappa(s)}|+1+\beta|\hat{X}^{i,N,M}_s|^2\,\mathrm{d}s\Big].$$

The next step is of course to take the expectation inside the integral. Let us start by noting the difference term can be bounded as

$$\begin{aligned}&\mathbb{E}\Big[\int_0^{T\wedge\eta^i_m}|\hat{X}^{i,N,M}_s-\tilde{X}^{i,N,M}_{\kappa(s)}|\,\mathrm{d}s\Big]\\&\qquad\leqslant \mathbb{E}\Big[\int_0^{T\wedge\eta^i_m}|\hat{X}^{i,N,M}_s-\hat{X}^{i,N,M}_{\kappa(s)}|\,\mathrm{d}s+\int_0^{T\wedge\eta^i_m}|\hat{X}^{i,N,M}_{\kappa(s)}-\tilde{X}^{i,N,M}_{\kappa(s)}|\,\mathrm{d}s\Big]\\&\qquad\leqslant \mathbb{E}\Big[h\int_0^{T\wedge\eta^i_m}|b\left((\kappa(s)-h)\vee 0,\tilde{X}^{i,N,M}_{\kappa(s)},\tilde{\mu}^{X,N,M}_{(\kappa(s)-h)\vee 0}\right)|\,\mathrm{d}s\Big]\\&\qquad\quad+\mathbb{E}\Big[\int_0^{T\wedge\eta^i_m}|\sigma\left(\kappa(s),\tilde{X}^{i,N,M}_{\kappa(s)},\tilde{\mu}^{i,N,M}_{\kappa(s)}\right)(W^i_s-W^i_{\kappa(s)})|\,\mathrm{d}s\Big]+C(m)h,\end{aligned}$$

where we have used Lemma 5.24 for the final inequality. For the other terms, one can note due to the growth assumptions on $b$ and Lemma 5.22, that

$$\mathbb{E}\Big[h\int_0^{T\wedge\eta^i_m}|b\left((\kappa(s)-h)\vee 0,\tilde{X}^{i,N,M}_{\kappa(s)},\tilde{\mu}^{X,N,M}_{(\kappa(s)-h)\vee 0}\right)|\,\mathrm{d}s\Big]\leqslant C(m)h.$$

The term involving $\sigma$ is more complex. However, we can bound it as follows:

$$\begin{aligned}&\mathbb{E}\Big[\int_0^{T\wedge\eta^i_m}|\sigma\left(\kappa(s),\tilde{X}^{i,N,M}_{\kappa(s)},\tilde{\mu}^{i,N,M}_{\kappa(s)}\right)(W^i_s-W^i_{\kappa(s)})|\,\mathrm{d}s\Big]\\&\qquad\leqslant C\int_0^T\sum_{a=1}^{l}\mathbb{E}\Big[|\sigma_a\left(\kappa(s),\tilde{X}^{i,N,M}_{\kappa(s)},\tilde{\mu}^{i,N,M}_{\kappa(s)}\right)||(W^i_s-W^i_{\kappa(s)})_a|\mathbb{1}_{\{\kappa(s)\leqslant t_{\lambda^i_m}\}}\Big]\,\mathrm{d}s\\&\qquad\leqslant C\int_0^T\sum_{a=1}^{l}h^{1/2}(1+\mathbb{E}[|\tilde{X}_{\kappa(s)\wedge t_{\lambda^i_m}}|^2])\,\mathrm{d}s\leqslant C(m)h^{1/2}.\end{aligned}$$

Further, since $|\hat{X}_s^{i,N,M}| \geqslant 0$, we obtain,

$$\mathbb{E}\Big[\int_0^{T\wedge\eta_m^i} |\hat{X}_s^{i,N,M}|^2 \,\mathrm{d}s\Big] \leqslant \int_0^T \mathbb{E}\big[|\hat{X}_{s\wedge\eta_m^i}^{i,N,M}|^2\big]\,\mathrm{d}s.$$

Hence,

$$\begin{aligned}\mathbb{E}[|\hat{X}_{T\wedge\eta_m^i}^{i,N,M}|^2] &\leqslant \mathbb{E}[|X_0^i|^2] + C(m)h^{1/2} + C\int_0^T 1 + \beta\mathbb{E}\big[|\hat{X}_{s\wedge\eta_m^i}^{i,N,M}|^2\big]\,\mathrm{d}s \\ &\leqslant \big(\mathbb{E}[|X_0^i|^2] + C + C(m)h^{1/2}\big)\exp(C\beta T), \qquad (5.20)\end{aligned}$$

where the final inequality follows from Gronwall.

In order to obtain an upper bound on the probability of the stopping time occurring we look to obtain a lower bound for (5.18) at the stopping time. For the moment let us take $|X_0^i| < m$, hence $\eta_m^i > 0$. There are now two possible ways the stopping time can be reached: if $\hat{X}$ hits the boundary first, then we have $|\hat{X}_{\eta_m^i}^{i,N,M}| = m$ and if $\tilde{X}$ hits the boundary first we have $|\tilde{X}_{\eta_m^i}^{i,N,M}| > m$.

In the case that $\hat{X}$ hits the boundary first, the lower bound is obvious, namely $|\hat{X}_{\eta_m^i}^{i,N,M}| = m$. For the second case it is less obvious. Recalling (5.19) and (5.14) we obtain lower bound

$$|\hat{X}_{t_k}^{i,N,M}|^2 \geqslant \frac{1}{2}\big((1-2h\beta)|\tilde{X}_{t_k}^{i,N,M}|^2 - 2h\alpha\big) - |b(t_0, \tilde{X}_{t_0}^{i,N,M}, \tilde{\mu}_{t_0}^{X,N,M})h|^2,$$

where again we are taking $k \geqslant 1$ here, but this is not a problem since we are assuming for the moment $|X_0^i| < m$. Observing that this lower bound holds independently of which process triggers the stopping condition we have on $\{|\tilde{X}_{\eta_m^i}^{i,N,M}| > m\}$ that

$$\begin{aligned} m^2 &\geqslant |\hat{X}_{\eta_m^i}^{i,N,M}|^2 \mathbb{1}_{\{|X_0^i|<m\}} \\ &\geqslant \frac{1}{2}\big((1-2h\beta)m^2 - 2h\alpha\big)\mathbb{1}_{\{|X_0^i|<m\}} - |b(t_0, \tilde{X}_{t_0}^{i,N,M}, \tilde{\mu}_{t_0}^{X,N,M})h|^2\mathbb{1}_{\{|X_0^i|<m\}}.\end{aligned}$$

Thus, for constants $C_1, C_2 > 0$,

$$|\hat{X}_{\eta_m^i}^{i,N,M}|^2\mathbb{1}_{\{|X_0^i|<m\}} \geqslant (C_1m^2 - C_2h)\mathbb{1}_{\{|X_0^i|<m\}} - C(m)h^2\mathbb{1}_{\{|X_0^i|<m\}},$$

where $|b(t_0, \tilde{X}_{t_0}^{i,N,M}, \tilde{\mu}_{t_0}^{X,N,M})|\mathbb{1}_{\{|X_0^i|<m\}} \leqslant C(m)\mathbb{1}_{\{|X_0^i|<m\}}$ via the growth condition on $b$. Let us now combine these results to obtain an upper bound for the probability of the stopping time. Notice that

$$\begin{aligned}\mathbb{E}[|\hat{X}_{T\wedge\eta_m^i}^{i,N,M}|^2] &\geqslant \mathbb{E}[|X_0^i|^2\mathbb{1}_{\{|X_0^i|\geqslant m\}}] + \mathbb{E}[|\hat{X}_{\eta_m^i}^{i,N,M}|^2\mathbb{1}_{\{|X_0^i|<m\}}\mathbb{1}_{\{0<\eta_m^i<T\}}] \\ &\geqslant \mathbf{P}(\eta_m^i = 0) + \big((C_1m^2 - C_2h) - C(m)h^2\big)\,\mathbf{P}(\{|X_0^i| < m\} \cap \{0 < \eta_m^i < T\}).\end{aligned}$$

Leaving the second term for the moment, and noting that $X_0^i$ is uniformly integrable, then for any $\varepsilon > 0$ there exists an $m^* > 0$ such that for all $m \geqslant m^*$

$$\mathbf{P}(\eta_m^i = 0) \leqslant m\mathbf{P}(|X_0^i| \geqslant m) \leqslant \mathbb{E}[|X_0^i|\mathbb{1}_{\{|X_0^i|\geqslant m\}}] \leqslant \frac{\varepsilon}{3}.$$

It is also useful to note that

$$\mathbf{P}(\{|X_0^i| < m\} \cap \{0 < \eta_m^i < T\}) = \mathbf{P}(\{0 < \eta_m^i < T\}).$$

From our previous analysis it is clear that for $m$ large enough and some contant $C(m)$, by using (5.20), the probability can be bounded by

$$\mathbf{P}(0<\eta_m^i<T) \leqslant \frac{\mathbb{E}[|\hat{X}^{i,N,M}_{T\wedge\eta_m^i}|^2]}{(C_1m^2-C_2h-C(m)h^2)} \leqslant \frac{\big(\mathbb{E}[|X_0^i|^2]+C+C(m)h^{1/2}\big)\exp(C\beta T)}{C_1m^2-C_2h-C(m)h^2}.$$

Now the goal is to bound this by $2\varepsilon/3$. We already have taken $m$ sufficiently large to obtain the last inequality. Now consider for any given $m$ a factor $h^*_{01}(m)$ such that $C_2h^*_{01}(m)+C(m)h^*_{01}(m)^2 \leqslant 1$. It is clear for $0<h<h^*_{01}(m)$ the same bound holds. Then for the same $\varepsilon$ as before choose $m$ large enough such that,

$$\frac{\big(\mathbb{E}[|X_0^i|^2]+C\big)\exp(C\beta T)}{C_1m^2-1} \leqslant \frac{\varepsilon}{3}.$$

Redefine $m^*$ as the corresponding maximum of this $m$ and $m^*$. Now for any $m\geqslant m^*$, define $h^*_{02}(m)$ such that,

$$\frac{C(m)(h^*_{02})^{1/2}\exp(C\beta T)}{C_1m^2-1} \leqslant \frac{\varepsilon}{3}.$$

Again for $0<h<h^*_{02}(m)$ the above inequality holds. Hence for any $m\geqslant m^*$ and any $0<h<\min(h^*_{01}(m),h^*_{02}(m))$, we have, $\mathbf{P}(\eta_m^i<T)\leqslant \mathbf{P}(\eta_m^i=0)+\mathbf{P}(0<\eta_m^i<T)\leqslant\varepsilon$. ■

We now look towards showing our strong convergence result, firstly by showing convergence between (5.18) and (5.2) and then (5.5) and (5.2). From this point onwards we require H2 (in Assumption 5.8).

Recalling the stopping time in Proposition 5.20, we now define $\theta_m^i := \tau_m^i \wedge \eta_m^i$ and have the following convergence result.

**Lemma 5.26**
*Let Assumption 5.1, 5.2, the full Assumption 5.8 hold, fix $h^*<1/\max(L_b,2\beta)$ and assume $X_0\in L^{4(q+1)}(\mathbb{R}^d)$. Then, for all $h\in(0,h^*)$,*

$$\sup_{1\leqslant i\leqslant N}\mathbb{E}\big[\sup_{0\leqslant t\leqslant T}|\hat{X}^{i,N,M}_{t\wedge\theta_m^i}-X^{i,N}_{t\wedge\theta_m^i}|^2\big]\leqslant C(m)h+C\mathbb{E}\big[\mathbb{1}_{\{T>\theta_m^i\}}\big]^{1/2}.$$

*Proof.* For ease of presentation we denote by $\overline{\kappa}(s):=(\kappa(s)-h)\vee 0$. As is standard we start by applying Itô to the difference to obtain

$$\begin{aligned}
&|X^{i,N}_{t\wedge\theta_m^i}-\hat{X}^{i,N,M}_{t\wedge\theta_m^i}|^2\\
&\quad=\int_0^{t\wedge\theta_m^i} 2\langle X_s^{i,N}-\hat{X}_s^{i,N,M}, b(s,X_s^{i,N},\mu_s^{X,N})-b(\overline{\kappa}(s),\tilde{X}^{i,N,M}_{\kappa(s)},\tilde{\mu}^{X,N,M}_{\overline{\kappa}(s)})\rangle\\
&\qquad\qquad+\sum_{a=1}^{l}|\sigma_a(s,X_s^{i,N},\mu_s^{X,N})-\sigma_a(\kappa(s),\tilde{X}^{i,N,M}_{\kappa(s)},\tilde{\mu}^{X,N,M}_{\kappa(s)})|^2\,\mathrm{d}s\\
&\qquad+\int_0^{t\wedge\theta_m^i}2\langle X_s^{i,N}-\hat{X}_s^{i,N,M},\big(\sigma(s,X_s^{i,N},\mu_s^{X,N})-\sigma(\kappa(s),\tilde{X}^{i,N,M}_{\kappa(s)},\tilde{\mu}^{X,N,M}_{\kappa(s)})\big)\,\mathrm{d}W_s^i\rangle.
\end{aligned}$$

By writing out the drift term we have that

$$
\begin{aligned}
&\langle X_s^{i,N} - \hat{X}_s^{i,N,M}, b(s, X_s^{i,N}, \mu_s^{X,N}) - b(\overline{\kappa}(s), \tilde{X}_{\kappa(s)}^{i,N,M}, \tilde{\mu}_{\overline{\kappa}(s)}^{X,N,M})\rangle \\
&\quad = \langle X_s^{i,N} - \hat{X}_s^{i,N,M}, b(s, X_s^{i,N}, \mu_s^{X,N}) - b(s, \hat{X}_s^{i,N,M}, \mu_s^{X,N})\rangle \\
&\qquad + \langle X_s^{i,N} - \hat{X}_s^{i,N,M}, b(s, \hat{X}_s^{i,N,M}, \mu_s^{X,N}) - b(\overline{\kappa}(s), \hat{X}_s^{i,N,M}, \mu_s^{X,N})\rangle \\
&\qquad + \langle X_s^{i,N} - \hat{X}_s^{i,N,M}, b(\overline{\kappa}(s), \hat{X}_s^{i,N,M}, \mu_s^{X,N}) - b(\overline{\kappa}(s), \hat{X}_{\kappa(s)}^{i,N,M}, \mu_s^{X,N})\rangle \\
&\qquad + \langle X_s^{i,N} - \hat{X}_s^{i,N,M}, b(\overline{\kappa}(s), \hat{X}_{\kappa(s)}^{i,N,M}, \mu_s^{X,N}) - b(\overline{\kappa}(s), \tilde{X}_{\kappa(s)}^{i,N,M}, \mu_s^{X,N})\rangle \\
&\qquad + \langle X_s^{i,N} - \hat{X}_s^{i,N,M}, b(\overline{\kappa}(s), \tilde{X}_{\kappa(s)}^{i,N,M}, \mu_s^{X,N}) - b(\overline{\kappa}(s), \tilde{X}_{\kappa(s)}^{i,N,M}, \tilde{\mu}_{\overline{\kappa}(s)}^{X,N,M})\rangle \\
&\quad \leqslant C\Big(|X_s^{i,N} - \hat{X}_s^{i,N,M}|^2 + h + \big(C \wedge CW^{(1)}(\mu_s^{X,N}, \tilde{\mu}_{\overline{\kappa}(s)}^{X,N,M})\big)^2 \\
&\qquad\quad + (1 + |\hat{X}_s^{i,N,M}|^{2q} + |\hat{X}_{\kappa(s)}^{i,N,M}|^{2q})|\hat{X}_s^{i,N,M} - \hat{X}_{\kappa(s)}^{i,N,M}|^2 \\
&\qquad\qquad + (1 + |\hat{X}_{\kappa(s)}^{i,N,M}|^{2q} + |\tilde{X}_{\kappa(s)}^{i,N,M}|^{2q})|\hat{X}_{\kappa(s)}^{i,N,M} - \tilde{X}_{\kappa(s)}^{i,N,M}|^2\Big),
\end{aligned}
$$

where we have used the growth bounds on $b$ along with several applications of Cauchy-Schwarz and Young's inequality. In particular we have used the fact that $b$ is both globally and $W^{(1)}$ bounded in measure to obtain the $C \wedge CW^{(1)}(\mu_s^{X,N}, \tilde{\mu}_{\overline{\kappa}(s)}^{X,N,M})$ bound. Using similar arguments to earlier proofs and to the drift term above, we get the following bound for the diffusion

$$
\begin{aligned}
&|\sigma_a(s, X_s^{i,N}, \mu_s^{X,N}) - \sigma_a(\kappa(s), \tilde{X}_{\kappa(s)}^{i,N,M}, \tilde{\mu}_{\kappa(s)}^{X,N,M})| \\
&\quad \leqslant C\big(h^{1/2} + |X_s^{i,N} - \hat{X}_s^{i,N,M}| + |\hat{X}_s^{i,N,M} - \hat{X}_{\kappa(s)}^{i,N,M}| \\
&\qquad + |\hat{X}_{\kappa(s)}^{i,N,M} - \tilde{X}_{\kappa(s)}^{i,N,M}| + 1 \wedge W^{(1)}(\mu_s^{X,N}, \tilde{\mu}_{\kappa(s)}^{X,N,M})\big).
\end{aligned}
$$

Ultimately we need to take supremum and expected values. Hence, we wish to bound

$$
\mathbb{E}\Big[\sup_{0\leqslant r\leqslant t\wedge\theta_m^i} \int_0^r 2\langle X_s^{i,N} - \hat{X}_s^{i,N,M}, \big(\sigma(s, X_s^{i,N}, \mu_s^{X,N}) - \sigma(\kappa(s), \tilde{X}_{\kappa(s)}^{i,N,M}, \tilde{\mu}_{\kappa(s)}^{X,N,M})\big)\, \mathrm{d}W_s^i\rangle\Big].
$$

We use the Burkholder Davis Gundy inequality, however care is needed since the terminal time is a stopping time. It turns out the usual upper bound still holds (see for example [Pro05, pg. 226]), hence we obtain, by using Young's inequality,

$$
\begin{aligned}
&\mathbb{E}\Big[\sup_{0\leqslant r\leqslant t\wedge\theta_m^i} \int_0^r 2\langle X_s^{i,N} - \hat{X}_s^{i,N,M}, \big(\sigma(s, X_s^{i,N}, \mu_s^{X,N}) - \sigma(\kappa(s), \tilde{X}_{\kappa(s)}^{i,N,M}, \tilde{\mu}_{\kappa(s)}^{X,N,M})\big)\, \mathrm{d}W_s^i\rangle\Big] \\
&\quad \leqslant C\mathbb{E}\Big[\Big(\int_0^{t\wedge\theta_m^i} |X_s^{i,N} - \hat{X}_s^{i,N,M}|^2 \sum_{a=1}^{l} |\sigma_a(s, X_s^{i,N}, \mu_s^{X,N}) - \sigma_a(\kappa(s), \tilde{X}_{\kappa(s)}^{i,N,M}, \tilde{\mu}_{\kappa(s)}^{X,N,M})|^2\, \mathrm{d}s\Big)^{1/2}\Big] \\
&\quad \leqslant \frac{1}{2}\mathbb{E}\Big[\sup_{0\leqslant s\leqslant t\wedge\theta_m^i} |X_s^{i,N} - \hat{X}_s^{i,N,M}|^2\Big] \\
&\qquad\qquad + C\mathbb{E}\Big[\int_0^{t\wedge\theta_m^i} \sum_{a=1}^{l} |\sigma_a(s, X_s^{i,N}, \mu_s^{X,N}) - \sigma_a(\kappa(s), \tilde{X}_{\kappa(s)}^{i,N,M}, \tilde{\mu}_{\kappa(s)}^{X,N,M})|^2\, \mathrm{d}s\Big].
\end{aligned}
$$

Taking the supremum over time and expectations of our original difference and using these bounds we obtain the inequality

$$
\begin{aligned}
&\frac{1}{2}\mathbb{E}\left[\sup_{0\leqslant t\leqslant T}|X^{i,N}_{t\wedge\theta^i_m}-\hat{X}^{i,N,M}_{t\wedge\theta^i_m}|^2\right]\\
&\quad\leqslant \mathbb{E}\bigg[\int_0^{T\wedge\theta^i_m} C\Big(|X^{i,N}_s-\hat{X}^{i,N,M}_s|^2+\big(1\wedge W^{(1)}(\mu^{X,N}_s,\tilde{\mu}^{X,N,M}_{\overline{\kappa}(s)})\big)^2\\
&\qquad\quad+(1+|\hat{X}^{i,N,M}_s|^{2q}+|\hat{X}^{i,N,M}_{\kappa(s)}|^{2q})|\hat{X}^{i,N,M}_s-\hat{X}^{i,N,M}_{\kappa(s)}|^2\\
&\qquad\quad+h+(1+|\hat{X}^{i,N,M}_{\kappa(s)}|^{2q}+|\tilde{X}^{i,N,M}_{\kappa(s)}|^{2q})|\hat{X}^{i,N,M}_{\kappa(s)}-\tilde{X}^{i,N,M}_{\kappa(s)}|^2\Big)\\
&\qquad\quad+C\sum_{a=1}^{l}\Big(h+|X^{i,N}_s-\hat{X}^{i,N,M}_s|^2+|\hat{X}^{i,N,M}_s-\hat{X}^{i,N,M}_{\kappa(s)}|^2\\
&\qquad\qquad\qquad+|\hat{X}^{i,N,M}_{\kappa(s)}-\tilde{X}^{i,N,M}_{\kappa(s)}|^2+\big(1\wedge W^{(1)}(\mu^{X,N}_s,\tilde{\mu}^{X,N,M}_{\kappa(s)})\big)^2\Big)\,\mathrm{d}s\bigg].
\end{aligned}
$$

Let us now concentrate on the measure terms $1\wedge W^{(1)}(\mu^{X,N}_s,\tilde{\mu}^{X,N,M}_{\overline{\kappa}(s)})$ and $1\wedge W^{(1)}(\mu^{X,N}_s,\tilde{\mu}^{X,N,M}_{\kappa(s)})$. The goal in the end is to use a Grönwall type inequality. Hence, we want to obtain terms of a similar form. The standard argument in this case is to remove the average sum of other particles using the fact that they are identically distributed, unfortunately the presence of the stopping time breaks this argument and forces us to argue a different way. We start by noting the following bound

$$
W^{(1)}(\mu^{X,N}_s,\tilde{\mu}^{X,N,M}_{\overline{\kappa}(s)})\leqslant\frac{1}{N}\sum_{j=1}^{N}|X^{j,N}_s-\tilde{X}^{j,N,M}_{\overline{\kappa}(s)}|\mathbb{1}_{\{s\leqslant\theta^j_m\}}+\frac{1}{N}\sum_{j=1}^{N}|X^{j,N}_s-\tilde{X}^{j,N,M}_{\overline{\kappa}(s)}|\mathbb{1}_{\{s>\theta^j_m\}}.
$$

By using the fact that for $a$, $b$, $c>0$, $\min(a,b+c)\leqslant\min(a,b)+\min(b,c)$ and $\min(a,b)\leqslant\sqrt{a}\sqrt{b}$ alongside Hölder inequality for sums, we obtain

$$
1\wedge W^{(1)}(\mu^{X,N}_s,\tilde{\mu}^{X,N,M}_{\overline{\kappa}(s)})\leqslant\sqrt{\frac{1}{N}\sum_{j=1}^{N}|X^{j,N}_s-\tilde{X}^{j,N,M}_{\overline{\kappa}(s)}|^2\mathbb{1}_{\{s\leqslant\theta^j_m\}}}+\sqrt{\frac{1}{N}\sum_{j=1}^{N}|X^{j,N}_s-\tilde{X}^{j,N,M}_{\overline{\kappa}(s)}|\mathbb{1}_{\{s>\theta^j_m\}}}\,.
$$

Let us further define $\hat{\mu}^{X,N,M}_s:=\frac{1}{N}\sum_{j=1}^{N}\delta_{\hat{X}^{j,N,M}_s}$. Then using the triangle inequality we get

$$
\begin{aligned}
&\frac{1}{N}\sum_{j=1}^{N}|X^{j,N}_s-\tilde{X}^{j,N,M}_{\overline{\kappa}(s)}|^2\mathbb{1}_{\{s\leqslant\theta^j_m\}}\\
&\quad\leqslant\frac{C}{N}\sum_{j=1}^{N}|X^{j,N}_s-\hat{X}^{j,N,M}_s|^2\mathbb{1}_{\{s\leqslant\theta^j_m\}}+\frac{C}{N}\sum_{j=1}^{N}|\hat{X}^{j,N,M}_s-\hat{X}^{j,N,M}_{\kappa(s)}|^2\mathbb{1}_{\{s\leqslant\theta^j_m\}}\\
&\qquad+\frac{C}{N}\sum_{j=1}^{N}|\hat{X}^{j,N,M}_{\kappa(s)}-\hat{X}^{j,N,M}_{\overline{\kappa}(s)}|^2\mathbb{1}_{\{s\leqslant\theta^j_m\}}+\frac{C}{N}\sum_{j=1}^{N}|\hat{X}^{j,N,M}_{\overline{\kappa}(s)}-\tilde{X}^{j,N,M}_{\overline{\kappa}(s)}|^2\mathbb{1}_{\{s\leqslant\theta^j_m\}}\,.
\end{aligned}
$$

Hence, we can bound the measure terms by

$$\begin{aligned}
&\mathbb{E}\Big[\int_0^{T\wedge\theta_m^i}\big(1\wedge W^{(1)}(\mu_s^{X,N},\tilde{\mu}_{\overline{\kappa}(s)}^{X,N,M})\big)^2\,\mathrm{d}s\Big]\\
&\leqslant \mathbb{E}\Big[\int_0^{T}\big(1\wedge W^{(1)}(\mu_s^{X,N},\tilde{\mu}_{\overline{\kappa}(s)}^{X,N,M})\big)^2\,\mathrm{d}s\Big]\\
&\leqslant \mathbb{E}\Big[\int_0^{T}\frac{C}{N}\Big(\sum_{j=1}^N|X_s^{j,N}-\hat{X}_s^{j,N,M}|^2\mathbb{1}_{\{s\leqslant\theta_m^j\}}+\sum_{j=1}^N|\hat{X}_s^{j,N,M}-\hat{X}_{\kappa(s)}^{j,N,M}|^2\mathbb{1}_{\{s\leqslant\theta_m^j\}}\\
&\qquad+\sum_{j=1}^N|\hat{X}_{\kappa(s)}^{j,N,M}-\hat{X}_{\overline{\kappa}(s)}^{j,N,M}|^2\mathbb{1}_{\{s\leqslant\theta_m^j\}}+\sum_{j=1}^N|\hat{X}_{\overline{\kappa}(s)}^{j,N,M}-\tilde{X}_{\overline{\kappa}(s)}^{j,N,M}|^2\mathbb{1}_{\{s\leqslant\theta_m^j\}}\\
&\qquad+\sum_{j=1}^N|X_s^{j,N}-\tilde{X}_{\overline{\kappa}(s)}^{j,N,M}|\mathbb{1}_{\{s>\theta_m^j\}}\,\mathrm{d}s\Big)\Big]
\end{aligned}$$

and likewise also

$$\begin{aligned}
&\mathbb{E}\Big[\int_0^{T\wedge\theta_m^i}\big(1\wedge W^{(1)}(\mu_s^{X,N},\tilde{\mu}_{\kappa(s)}^{X,N,M})\big)^2\,\mathrm{d}s\Big]\\
&\quad\leqslant \mathbb{E}\Big[\int_0^{T}\frac{C}{N}\Big(\sum_{j=1}^N|X_s^{j,N}-\hat{X}_s^{j,N,M}|^2\mathbb{1}_{\{s\leqslant\theta_m^j\}}+\sum_{j=1}^N|\hat{X}_s^{j,N,M}-\hat{X}_{\kappa(s)}^{j,N,M}|^2\mathbb{1}_{\{s\leqslant\theta_m^j\}}\\
&\qquad+\sum_{j=1}^N|\hat{X}_{\kappa(s)}^{j,N,M}-\tilde{X}_{\kappa(s)}^{j,N,M}|^2\mathbb{1}_{\{s\leqslant\theta_m^j\}}+\sum_{j=1}^N|X_s^{j,N}-\tilde{X}_{\kappa(s)}^{j,N,M}|\mathbb{1}_{\{s>\theta_m^j\}}\,\mathrm{d}s\Big)\Big].
\end{aligned}$$

Therefore, taking the expectation inside the integral and supremum over the particle index; noting particles are identically distributed, we obtain

$$\begin{aligned}
&\sup_{1\leqslant i\leqslant N}\mathbb{E}\left[\sup_{0\leqslant t\leqslant T\wedge\theta_m^i}|X_{t\wedge\theta_m^i}^{i,N}-\hat{X}_{t\wedge\theta_m^i}^{i,N,M}|^2\right]\\
&\leqslant C\Big(hT+\int_0^T\sup_{1\leqslant i\leqslant N}\mathbb{E}\Big[\sup_{0\leqslant r\leqslant s}|X_{r\wedge\theta_m^i}^{i,N}-\hat{X}_{r\wedge\theta_m^i}^{i,N,M}|^2\Big]+\sup_{1\leqslant i\leqslant N}\mathbb{E}\Big[|\hat{X}_{\kappa(s)}^{i,N,M}-\hat{X}_{\overline{\kappa}(s)}^{i,N,M}|^2\mathbb{1}_{\{s\leqslant\theta_m^i\}}\Big]\\
&\quad+\sup_{1\leqslant i\leqslant N}\mathbb{E}\Big[|\hat{X}_{\overline{\kappa}(s)}^{i,N,M}-\tilde{X}_{\overline{\kappa}(s)}^{i,N,M}|^2\mathbb{1}_{\{s\leqslant\theta_m^i\}}\Big]\\
&\qquad\qquad+\sup_{1\leqslant i\leqslant N}\mathbb{E}\Big[|X_s^{i,N}-\tilde{X}_{\overline{\kappa}(s)}^{i,N,M}|\mathbb{1}_{\{s>\theta_m^i\}}\Big]+\sup_{1\leqslant i\leqslant N}\mathbb{E}\Big[|X_s^{i,N}-\tilde{X}_{\kappa(s)}^{i,N,M}|\mathbb{1}_{\{s>\theta_m^i\}}\Big]\\
&\quad+\sup_{1\leqslant i\leqslant N}\mathbb{E}\Big[(1+|\hat{X}_s^{i,N,M}|^{2q}+|\hat{X}_{\kappa(s)}^{i,N,M}|^{2q})|\hat{X}_s^{i,N,M}-\hat{X}_{\kappa(s)}^{i,N,M}|^2\mathbb{1}_{\{s\leqslant\theta_m^i\}}\Big]\\
&\qquad\qquad+\sup_{1\leqslant i\leqslant N}\mathbb{E}\Big[(1+|\hat{X}_{\kappa(s)}^{i,N,M}|^{2q}+|\tilde{X}_{\kappa(s)}^{i,N,M}|^{2q})|\hat{X}_{\kappa(s)}^{i,N,M}-\tilde{X}_{\kappa(s)}^{i,N,M}|^2\mathbb{1}_{\{s\leqslant\theta_m^i\}}\Big]\,\mathrm{d}s\Big).
\end{aligned}$$

where we have further used that if $Y_\cdot\geqslant 0$, then $Y_\cdot\mathbb{1}_{\{\cdot\leqslant t\}}\leqslant Y_{\{\cdot\wedge t\}}$. Noting $\mathbb{1}_{\{\cdot\}}=\mathbb{1}_{\{\cdot\}}^2$, we obtain via Cauchy-Schwarz inequality

$$\begin{aligned}
&\mathbb{E}\Big[(1+|\hat{X}_s^{i,N,M}|^{2q}+|\hat{X}_{\kappa(s)}^{i,N,M}|^{2q})|\hat{X}_s^{i,N,M}-\hat{X}_{\kappa(s)}^{i,N,M}|^2\mathbb{1}_{\{s\leqslant\theta_m^i\}}\Big]\\
&\qquad\qquad\leqslant C(m)\mathbb{E}\Big[|\hat{X}_s^{i,N,M}-\hat{X}_{\kappa(s)}^{i,N,M}|^4\mathbb{1}_{\{s\leqslant\theta_m^i\}}\Big]^{1/2}.
\end{aligned}$$

Noting that

$$|\hat{X}_s^{i,N,M} - \hat{X}_{\kappa(s)}^{i,N,M}| \leqslant |b\left(\overline{\kappa}(s), \tilde{X}_{\kappa(s)}^{i,N,M}, \tilde{\mu}_{(\kappa(s)-h)\vee 0}^{X,N,M}\right)|h + |\sigma\left(\kappa(s), \tilde{X}_{\kappa(s)}^{i,N,M}, \tilde{\mu}_{\kappa(s)}^{X,N,M})\right)(W_s^i - W_{\kappa(s)}^i)|,$$

which implies

$$\begin{aligned}
&\mathbb{E}\Big[|\hat{X}_s^{i,N,M} - \hat{X}_{\kappa(s)}^{i,N,M}|^4 \mathbb{1}_{\{s \leqslant \theta_m^i\}}\Big] \\
&\quad \leqslant Ch^4 \mathbb{E}\Big[(1 + |\tilde{X}_{\kappa(s)}^{i,N,M}|^{4(q+1)})\mathbb{1}_{\{s \leqslant \theta_m^i\}}\Big] + C\mathbb{E}\Big[(1 + |\tilde{X}_{\kappa(s)}^{i,N,M}|^4)\mathbb{1}_{\{s \leqslant \theta_m^i\}}\Big]\mathbb{E}\Big[(W_s^i - W_{\kappa(s)}^i)^4\Big] \\
&\quad \leqslant C(m)h^2,
\end{aligned}$$

where we used Lemma 5.22 to obtain the final inequality (note that by assumption $X_0 \in L^{4(q+1)}(\mathbb{R}^d)$). Arguing in the exact same fashion along with Lemma 5.24 also yields

$$\mathbb{E}\Big[(1 + |\hat{X}_{\kappa(s)}^{i,N,M}|^{2q} + |\tilde{X}_{\kappa(s)}^{i,N,M}|^{2q})|\hat{X}_{\kappa(s)}^{i,N,M} - \tilde{X}_{\kappa(s)}^{i,N,M}|^2 \mathbb{1}_{\{s \leqslant \theta_m^i\}}\Big] \leqslant C(m)h.$$

The remaining terms can be bounded using the same arguments as above. Substituting these bounds then implies

$$\begin{aligned}
&\sup_{1 \leqslant i \leqslant N} \mathbb{E}\Big[\sup_{0 \leqslant t \leqslant T} |X_{t\wedge\theta_m^i}^{i,N} - \hat{X}_{t\wedge\theta_m^i}^{i,N,M}|^2\Big] \\
&\leqslant C(m)h + C\int_0^T \sup_{1 \leqslant i \leqslant N} \mathbb{E}\Big[|X_s^{i,N} - \tilde{X}_{\overline{\kappa}(s)}^{i,N,M}|\mathbb{1}_{\{s > \theta_m^i\}}\Big]\mathrm{d}s \\
&\quad + C\int_0^T \sup_{1 \leqslant i \leqslant N} \mathbb{E}\Big[|X_s^{i,N} - \tilde{X}_{\kappa(s)}^{i,N,M}|\mathbb{1}_{\{s > \theta_m^i\}}\Big]\mathrm{d}s + C\int_0^T \sup_{1 \leqslant i \leqslant N} \mathbb{E}\Big[\sup_{0 \leqslant r \leqslant s} |X_{r\wedge\theta_m^i}^{i,N} - \hat{X}_{r\wedge\theta_m^i}^{i,N,M}|^2\Big]\mathrm{d}s.
\end{aligned}$$

Hence, by Gronwall's inequality we obtain,

$$\begin{aligned}
&\sup_{1 \leqslant i \leqslant N} \mathbb{E}\Big[\sup_{0 \leqslant t \leqslant T} |X_{t\wedge\theta_m^i}^{i,N} - \hat{X}_{t\wedge\theta_m^i}^{i,N,M}|^2\Big] \\
&\leqslant C(m)h + C\int_0^T \sup_{1 \leqslant i \leqslant N} \mathbb{E}\Big[|X_s^{i,N} - \tilde{X}_{\overline{\kappa}(s)}^{i,N,M}|\mathbb{1}_{\{s > \theta_m^i\}}\Big] + \sup_{1 \leqslant i \leqslant N} \mathbb{E}\Big[|X_s^{i,N} - \tilde{X}_{\kappa(s)}^{i,N,M}|\mathbb{1}_{\{s > \theta_m^i\}}\Big]\mathrm{d}s.
\end{aligned}$$

We then complete the proof by applying Cauchy-Schwarz to the expectations in the integrand along with Propositions 5.20 and 5.23. ■

We now can prove our main implicit scheme result.

*Proof of Proposition 5.9.* Recall that $s \in [1, 2)$. Define the error term as $E_r(T)^i = X_T^{i,N} - \tilde{X}_T^{i,N,M}$ and also let us note a more general version of Young's inequality

$$x^s y \leqslant \frac{\delta s}{2}x^2 + \frac{2-s}{2\delta^{s/(2-s)}}y^{2/(2-s)}, \quad \forall\, x,\ y,\ \delta > 0.$$

Hence,

$$\begin{aligned}
\mathbb{E}[|X_T^{i,N} - \tilde{X}_T^{i,N,M}|^s] \leqslant 2^{s-1}\big(&\mathbb{E}[|X_T^{i,N} - \hat{X}_T^{i,N,M}|^s \mathbb{1}_{\{\tau_m^i > T,\ \eta_m^i > T\}}] \\
&+ \mathbb{E}[|\hat{X}_T^{i,N,M} - \tilde{X}_T^{i,N,M}|^s \mathbb{1}_{\{\tau_m^i > T,\ \eta_m^i > T\}}]\big) \\
&+ \frac{\delta s}{2}\mathbb{E}[|E_r(T)^i|^2] + \frac{2-s}{2\delta^{s/(2-s)}}\mathbb{E}[\mathbb{1}_{\{\tau_m^i \leqslant T \text{ or } \eta_m^i \leqslant T\}}].
\end{aligned}$$

From Lemma 5.24 we obtain,

$$\mathbb{E}[|\hat{X}_T^{i,N,M} - \tilde{X}_T^{i,N,M}|^s \mathbb{1}_{\{\tau_m^i > T,\ \eta_m^i > T\}}] \leqslant C(m,s)h^s.$$

Also let us note,

$$\mathbb{E}[|E_r(T)^i|^2] \leqslant 2\mathbb{E}[|X_T^{i,N}|^2 + |\tilde{X}_T^{i,N,M}|^2] \leqslant 2C,$$

where we have used Propositions 5.20 and 5.23. Hence for any $\varepsilon > 0$, we can choose $\delta$ such that,

$$\frac{\delta s}{2}\mathbb{E}[|E_r(T)^i|^2] \leqslant \frac{\varepsilon}{3}.$$

By subadditivity of measures, $\mathbb{E}[\mathbb{1}_{\{\tau_m^i \leqslant T \text{ or } \eta_m^i \leqslant T\}}] \leqslant \mathbf{P}(\tau_m^i \leqslant T) + \mathbf{P}(\eta_m^i \leqslant T)$. By Proposition 5.20, there exists $m^*$ (dependent on $\delta$), such that for $m \geqslant m^*$,

$$\frac{2-s}{2\delta^{s/(2-s)}}\mathbf{P}(\tau_m^i \leqslant T) \leqslant \frac{\varepsilon}{3}.$$

Then, noting by Lemma 5.26,

$$\begin{aligned}\mathbb{E}[|\hat{X}_T^{i,N,M} - X_T^{i,N}|^s \mathbb{1}_{\{\tau_m^i > T,\ \eta_m^i > T\}}] &\leqslant \mathbb{E}\Big[\sup_{0 \leqslant t \leqslant T}|\hat{X}_{t\wedge\theta_m^i}^{i,N,M} - X_{t\wedge\theta_m^i}^{i,N}|^2\Big]^{s/2} \\ &\leqslant C(m)h^{s/2} + C\mathbb{E}\big[\mathbb{1}_{\{T > \theta_m^i\}}\big]^{s/4}.\end{aligned}$$

Again by subadditivity of measures we can bound

$$\mathbb{E}\big[\mathbb{1}_{\{T > \theta_m^i\}}\big]^{s/4} \leqslant \mathbf{P}(\tau_m^i \leqslant T)^{s/4} + \mathbf{P}(\eta_m^i \leqslant T)^{s/4}.$$

By the same argument as before we can define a new $m^*$, greater than or equal to the previous such that $C\mathbf{P}(\tau_m^i \leqslant T)^{s/4}$ is sufficiently small. By Lemma 5.25, by taking $h$ small enough for any $\tilde{\varepsilon} > 0$, $\mathbf{P}(\eta_m^i < T) \leqslant \tilde{\varepsilon}$, and by extension, there exists an $h$ small enough such that $\mathbf{P}(\eta_m^i < T)^{s/4} \leqslant \tilde{\varepsilon}$. Hence, for any $m$, we can take $h$ small enough such that

$$\begin{aligned}&2^{s-1}\big(\mathbb{E}[|X_T^{i,N} - \hat{X}_T^{i,N,M}|^s \mathbb{1}_{\{\tau_m^i > T,\ \eta_m^i > T\}}] \\ &\quad + \mathbb{E}[|\hat{X}_T^{i,N,M} - \tilde{X}_T^{i,N,M}|^s \mathbb{1}_{\{\tau_m^i > T,\ \eta_m^i > T\}}]\big) + \frac{2-s}{2\delta^{s/(2-s)}}\mathbf{P}(\eta_m^i \leqslant T) \leqslant \frac{\varepsilon}{3}\end{aligned}$$

and hence $\mathbb{E}[|X_T^{i,N} - \tilde{X}_T^{i,N,M}|^s] \leqslant \varepsilon$. Since $\varepsilon > 0$ was arbitrary, we have the result. ■

# Bibliography


[AEFR18] S. Ankirchner, S. Engelhardt, A. Fromm, and G. d. Reis, *The Skorokhod embedding problem for inhomogeneous diffusions*, arXiv preprint arXiv:1810.05098 (2018).

[AFKP18] S. Ankirchner, A. Fromm, T. Kruse, and A. Popier, *Optimal position targeting via decoupling fields*, 2018. working paper or preprint.

[AHI08] S. Ankirchner, G. Heyne, and P. Imkeller, *A BSDE approach to the Skorokhod embedding problem for the Brownian motion with drift*, Stochastics and Dynamics **8** (2008), no. 01, 35–46.

[AHS15] S. Ankirchner, D. Hobson, and P. Strack, *Finite, integrable and bounded time embeddings for diffusions*, Bernoulli **21** (2015), no. 2, 1067–1088.

[AKTKK19] S. Ankirchner, N. Kazi-Tani, M. Klein, and T. Kruse, *Stopping with expectation constraints: 3 points suffice*, Electronic Journal of Probability **24** (2019).

[AS11] S. Ankirchner and P. Strack, *Skorokhod embeddings in bounded time*, Stoch. Dyn. **11** (2011), no. 2-3, 215–226. MR2836522

[Bas83] R. F. Bass, *Skorokhod imbedding via stochastic integrals*, Séminaire de probabilités de Strasbourg **17** (1983), 221–224 (eng).

[BCC11] F. Bolley, J. A. Cañizo, and J. A. Carrillo, *Stochastic mean-field limit: non-Lipschitz forces and swarming*, Math Models Methods Appl Sci **21** (2011).

[BCHK17] M. Beiglböck, A. M. Cox, M. Huesmann, and S. Källblad, *Measure-valued martingales and optimality of Bass-type solutions to the Skorokhod embedding problem*, arXiv preprint arXiv:1708.07071 (2017).

[BF17] A. Budhiraja and W.-T. Fan, *Uniform in time interacting particle approximations for nonlinear equations of Patlak-Keller-Segel type*, Electron. J. Probab. **22** (2017), Paper No. 8, 37. MR3613701

[BF92] A. Bensoussan and J. Frehse, *On Bellman equations of ergodic control in R n*, Applied stochastic analysis, 1992, pp. 21–29.

[BFFT12] J. Baladron, D. Fasoli, O. Faugeras, and J. Touboul, *Mean-field description and propagation of chaos in networks of Hodgkin-Huxley and FitzHugh-Nagumo neurons*, The Journal of Mathematical Neuroscience (2012).

[BFT15] M. Bossy, O. Faugeras, and D. Talay, *Clarification and complement to "mean-field description and propagation of chaos in networks of Hodgkin–Huxley and FitzHugh–Nagumo neurons"*, The Journal of Mathematical Neuroscience (JMN) **5** (2015), no. 1, 19.

[Bor06] V. S Borkar, *Ergodic control of diffusion processes*, Proceedings icm, 2006.

[BT04] B. Bouchard and N. Touzi, *Discrete-time approximation and Monte-Carlo simulation of backward stochastic differential equations*, Stochastic Processes and their applications **111** (2004), no. 2, 175–206.

[BT97] M. Bossy and D. Talay, *A stochastic particle method for the McKean-Vlasov and the Burgers equation*, Mathematics of Computation of the American Mathematical Society **66** (1997), no. 217, 157–192.

[BZ08] C. Bender and J. Zhang, *Time discretization and Markovian iteration for coupled FBSDEs*, The Annals of Applied Probability **18** (2008), no. 1, 143–177.

[Car16] R. Carmona, *Lectures on BSDEs, stochastic control, and stochastic differential games with financial applications*, Financial Mathematics, vol. 1, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2016. MR3629171

[CD17] R. Carmona and F. Delarue, *Probabilistic theory of mean field games with applications I*, 1st ed., Probability Theory and Stochastic Modelling, vol. 84, Springer International Publishing, 2017.

[CFP16] A. Cosso, M. Fuhrman, and H. Pham, *Long time asymptotics for fully nonlinear Bellman equations: a Backward SDE approach*, Stochastic Processes and their Applications **126** (2016), no. 7, 1932–1973.

[CJM16] J.-F. Chassagneux, A. Jacquier, and I. Mihaylov, *An explicit Euler scheme with strong rate of convergence for financial SDEs with non-Lipschitz coefficients*, SIAM Journal on Financial Mathematics **7** (2016), no. 1, 993–1021.

[D'A71] R. B. D'Agostino, *An omnibus test of normality for moderate and large size samples*, Biometrika **58** (1971), no. 2, 341–348.

[Del02] F. Delarue, *On the existence and uniqueness of solutions to FBSDEs in a non-degenerate case*, Stochastic Process. Appl. **99** (2002), no. 2, 209–286.

[DGG$^+$11] W. Dreyer, M. Gaberšček, C. Guhlke, R. Huth, and J. Jamnik, *Phase transition in a rechargeable lithium battery*, European Journal of Applied Mathematics **22** (2011), no. 3, 267–290.

[DGPR17] L. Doering, L. Gonon, D. J Prömel, and O. Reichmann, *On Skorokhod embeddings and Poisson equations*, arXiv preprint arXiv:1703.05673 (2017).

[DP73] R. D'Agostino and E. S. Pearson, *Tests for departure from normality*, Biometrika **60** (1973), no. 3, 613–622.

[dRES18] G. dos Reis, S. Engelhardt, and G. Smith, *Simulation of McKean Vlasov SDEs with super linear growth*, arXiv preprint arXiv:1808.05530 (2018).

[dRST18] G. dos Reis, G. Smith, and P. Tankov, *Importance sampling for McKean-Vlasov SDEs*, arXiv:1803.09320 (2018).

[dRST19] G. dos Reis, W. Salkeld, and J. Tugaut, *Freidlin-Wentzell LDP in path space for McKean-Vlasov equations and the functional iterated logarithm law*, Ann. Appl. Probab. **29** (2019), no. 3, 1487–1540. MR3914550

[FG16] W. Fang and M. B. Giles, *Adaptive Euler-Maruyama method for SDEs with non-globally Lipschitz drift: Part I, finite time interval*, arXiv preprint arXiv:1609.08101 (2016).

[FI13] A. Fromm and P. Imkeller, *Existence, Uniqueness and Regularity of Decoupling Fields to Multidimensional Fully Coupled FBSDEs*, Preprint arXiv:1310.0499 (2013).

[FIP15] A. Fromm, P. Imkeller, and D. Prömel, *An FBSDE approach to the Skorokhod embedding problem for Gaussian processes with non-linear drift*, Electron. J. Probab. **20** (2015), 38 pp.

[Fro15] A. Fromm, *Theory and applications of decoupling fields for forward-backward stochastic differential equations*, Ph.D. Thesis, 2015.

[GGM$^+$18] C. Guhlke, P. Gajewski, M. Maurelli, P. K. Friz, and W. Dreyer, *Stochastic many-particle model for LFP electrodes*, Contin. Mech. Thermodyn. **30** (2018), no. 3, 593–628. MR3783112

[GLW05] E. Gobet, J.-P. Lemor, and X. Warin, *A regression-based Monte Carlo method to solve backward stochastic differential equations*, Ann. Appl. Probab. **15** (2005), no. 3, 2172–2202. MR2152657 (2006c:60078)

[GP18] E. Gobet and S. Pagliarani, *Analytical approximations of non-linear SDEs of McKean-Vlasov type*, Journal of Mathematical Analysis and Applications (2018).

[GPV19] S. Gomes, G. Pavliotis, and U Vaes, *Mean-field limits for interacting diffusions with colored noise: phase transitions and spectral numerical methods*, arXiv preprint arXiv:1904.05973 (2019).

[HJK11] M. Hutzenthaler, A. Jentzen, and P. E. Kloeden, *Strong and weak divergence in finite time of Euler's method for stochastic differential equations with non-globally Lipschitz continuous coefficients*, Proc. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci. **467** (2011), no. 2130, 1563–1576. MR2795791

[HJK12] M. Hutzenthaler, A. Jentzen, and P. E Kloeden, *Strong convergence of an explicit numerical method for SDEs with nonglobally Lipschitz continuous coefficients*, The Annals of Applied Probability **22** (2012), no. 4, 1611–1641.

[HK12] D. Hobson and M. Klimmek, *Model-independent hedging strategies for variance swaps*, Finance and Stochastics **16** (2012), no. 4, 611–649.

[HMS02] D. J Higham, X. Mao, and A. M Stuart, *Strong convergence of Euler-type methods for nonlinear stochastic differential equations*, SIAM Journal on Numerical Analysis **40** (2002), no. 3, 1041–1063.

[IDRZ10] P. Imkeller, G. Dos Reis, and J. Zhang, *Results on numerics for FBSDE with drivers of quadratic growth*, Contemporary quantitative finance, 2010, pp. 159–182.

[Kee11] R. W Keener, *Theoretical statistics: Topics for a core course*, Springer, 2011.

[KHO97] A. Kohatsu-Higa and S. Ogawa, *Weak rate of convergence for an Euler scheme of nonlinear SDE's*, Monte Carlo Methods and Applications **3** (1997), 327–345.

[Kob00] M. Kobylanski, *Backward stochastic differential equations and partial differential equations with quadratic growth*, Annals of Probability (2000), 558–602.

[KS91] I. Karatzas and S. E. Shreve, *Brownian Motion and Stochastic Calculus*, Graduate Texts in Mathematics, Springer New York, 1991.

[Kun97] H. Kunita, *Stochastic flows and stochastic differential equations*, Vol. 24, Cambridge university press, 1997.

[Lac18] D. Lacker, *On a strong form of propagation of chaos for McKean-Vlasov equations*, Electronic Communications in Probability **23** (2018).

[LdRS15] A. Lionnet, G. dos Reis, and L. Szpruch, *Time discretization of FBSDE with polynomial growth drivers and reaction–diffusion PDEs*, Ann. Appl. Probab. **25** (2015), no. 5, 2563–2625. MR3375884

[Mal03] F. Malrieu, *Convergence to equilibrium for granular media equations and their Euler schemes*, Ann. Appl. Probab. **13** (2003), no. 2, 540–560. MR1970276

[Mél96] S. Méléard, *Asymptotic behaviour of some interacting particle systems; McKean-Vlasov and boltzmann models*, Probabilistic models for nonlinear partial differential equations, 1996, pp. 42–95.

[MPF12] D. S. Mitrinovic, J. Pecaric, and A. M. Fink, *Inequalities involving functions and their integrals and derivatives*, Vol. 53, Springer Science & Business Media, 2012.

[MPY94] J. Ma, P. Protter, and J. Yong, *Solving forward-backward stochastic differential equations explicitly – a four step scheme*, Probab. Theory Relat. Fields **98** (1994), no. 3, 339–359.

[MS13] X. Mao and L. Szpruch, *Strong convergence and stability of implicit numerical methods for stochastic differential equations with non-globally Lipschitz continuous coefficients*, Journal of Computational and Applied Mathematics **238** (2013), 14–28.

[MT05] G. Milstein and M. V Tretyakov, *Numerical integration of stochastic differential equations with nonglobally Lipschitz coefficients*, SIAM journal on numerical analysis **43** (2005), no. 3, 1139–1154.

[MWZZ15] J. Ma, Z. Wu, D. Zhang, and J. Zhang, *On well-posedness of forward-backward SDEs—a unified approach*, Ann. Appl. Probab. **25** (2015), no. 4, 2168–2214. MR3349005

[MY99] J. Ma and J. Yong, *Forward-backward stochastic differential equations and their applications*, Lecture Notes in Mathematics, vol. 1702, Springer-Verlag, Berlin, 1999. MR1704232

[MYZ12] J. Ma, H. Yin, and J. Zhang, *On non-Markovian forward-backward SDEs and backward stochastic PDEs*, Stochastic Process. Appl. **122** (2012), no. 12, 3980–4004. MR2971722

[Obł04] J. Obłój, *The Skorokhod embedding problem and its offspring*, Probab. Surv. **1** (2004), 321–390. MR2068476

[OTV19] C. Orrieri, G. Tessitore, and P. Veverka, *Ergodic maximum principle for stochastic systems*, Applied Mathematics & Optimization **79** (2019), no. 3, 567–591.

[Per91] L. Perko, *Differential equations and dynamical systems*, Springer-Verlag, 1991.

[Pha09] H. Pham, *Continuous-time stochastic control and optimization with financial applications*, Vol. 61, Springer Science & Business Media, 2009.

[PP01] J. L. Pedersen and G. Peskir, *The Azéma-Yor embedding in non-singular diffusions*, Stochastic Process. Appl. **96** (2001), no. 2, 305–312. MR1865760

[Pro05] P. E Protter, *Stochastic differential equations*, Springer, 2005.

[PT99] E. Pardoux and S. Tang, *Forward-backward stochastic differential equations and quasilinear parabolic PDEs*, Probab. Theory Relat. Fields **114** (1999), no. 2, 123–150.

[PW99] S. Peng and Z. Wu, *Fully coupled forward-backward stochastic differential equations and applications to optimal control*, SIAM J. Control Optim. **37** (1999), no. 3, 825–843. MR1675098

[RR98] S. T. Rachev and L. Rüschendorf, *Mass transportation problems. Vol. II*, Probability and its Applications (New York), Springer-Verlag, New York, 1998. Applications. MR1619171

[RY13] D. Revuz and M. Yor, *Continuous martingales and Brownian motion*, Vol. 293, Springer Science & Business Media, 2013.

[Sab13] S. Sabanis, *A note on tamed Euler approximations*, Electron. Commun. Probab. **18** (2013), no. 47, 10. MR3070913

[Sch00] B. Schmalfuss, *Attractors for the non–autonomous dynamical systems*, Equadiff 99: (in 2 volumes), 2000, pp. 684–689.

[Sch99] B. Schmalfuss, *Measure attractors and random attractors for stochastic partial differential equations*, Stochastic analysis and applications **17** (1999), no. 6, 1075–1101.

[Sko65] A. V. Skorokhod, *Studies in the theory of random processes*, Translated from the Russian by Scripta Technica, Inc, Addison-Wesley Publishing Co., Inc., Reading, Mass., 1965. MR0185620 (32 #3082b)

[SXY18] J. Sun, J. Xiong, and J. Yong, *Stochastic linear-quadratic optimal control problems with random coefficients: Closed-loop representation of open-loop optimal controls*, arXiv preprint arXiv:1809.00261 (2018).

[Szn91] A.-S. Sznitman, *Topics in propagation of chaos*, Ecole d'Eté de Probabilités de Saint-Flour XIX — 1989 (1991), 165–251.

[Tie13] D. N. Tien, *A stochastic Ginzburg-Landau equation with impulsive effects*, Physica A: Statistical Mechanics and its Applications **392** (2013), no. 9, 1962–1971.

[VBHK19] J. B. Veraguas, M. Beiglböck, M. Huesmann, and S. Källblad, *Martingale Benamou–Brenier: a probabilistic perspective*, arXiv preprint arXiv:1708.04869 (2019).

[Vil08] C. Villani, *Optimal transport: old and new*, Vol. 338, Springer Science & Business Media, 2008.

[Vrk78] I. Vrkoč, *Liouville formula for systems of linear homogeneous Itô stochastic differential equations*, Commentationes Mathematicae Universitatis Carolinae **19** (1978), no. 1, 141–146.

[YZ99] J. Yong and X. Y. Zhou, *Stochastic controls: Hamiltonian systems and HJB equations*, Vol. 43, Springer Science & Business Media, 1999.

[Zei90] E. Zeidler, *Nonlinear functional analysis and its applications. II/B*, Springer-Verlag, New York, 1990. Nonlinear monotone operators, Translated from the German by the author and Leo F. Boron. MR1033498

# Ehrenwörtliche Erklärung

Hiermit erkläre ich,

- dass mir die Promotionsordnung der Fakultät bekannt ist,
- dass ich die Dissertation selbst angefertigt habe, keine Textabschnitte oder Ergebnisse eines Dritten oder eigener Prüfungsarbeiten ohne Kennzeichnung übernommen und alle von mir benutzten Hilfsmittel, persönliche Mitteilungen und Quellen in meiner Arbeit angegeben habe,
- dass ich die Hilfe eines Promotionsberaters nicht in Anspruch genommen habe und dass Dritte weder unmittelbar noch mittelbar geldwerte Leistungen von mir für Arbeiten erhalten haben, die in Zusammenhang mit dem Inhalt der vorgelegten Dissertation stehen,
- dass ich die Dissertation noch nicht als Prüfungsarbeit für eine staatliche oder andere wissenschaftliche Prüfung eingereicht habe,
- dass ich bei der Auswahl und Auswertung des Materials sowie bei der Herstellung des Manuskripts nicht von Dritten unterstützt wurde,
- dass ich weder die gleiche noch eine in wesentlichen Teilen ähnliche bzw. eine andere Abhandlung bereits bei einer anderen Hochschule als Dissertation eingereicht habe.

Jena, den 21.04.2020 Stefan Engelhardt