Dissertation

PATHWISE UNIQUENESS FOR STOCHASTIC DIFFERENTIAL EQUATIONS WITH SINGULAR DRIFT AND NONCONSTANT DIFFUSION

Katharina von der Lühe

Fakultät für Mathematik
Universität Bielefeld

August 2016

Contents

1. Introduction

1.1. Brief survey of the problem in finite dimensions

We consider the following ordinary stochastic differential equation (SDE):

$$X_t = x + \int_0^t b(s, X_s)\, ds + \int_0^t \sigma(s, X_s)\, dW_s, \qquad t \in [0, T],$$

where $x \in \mathbb{R}^d$, b, σ are measurable functions from $[0, T] \times \mathbb{R}^d$ to $\mathbb{R}^d$, respectively $\mathbb{R}^{d\times m}$, and W_t is an m-dimensional standard Wiener process.

It is well known that we have strong existence and uniqueness for this equation under Lipschitz continuity of the coefficients, which was shown by K. Itô, see [Itô46], who first rigorously developed the theory of stochastic integration. Since Lipschitz continuity is a rather strong assumption and this kind of SDE arises in many settings that do not necessarily provide Lipschitz continuous coefficients, e.g. interacting particles, it is natural to ask if it is also possible to get a unique strong solution under weaker properties. It turns out that this holds under much more general assumptions on the drift term b, neither continuity nor the absence of singularities is necessary.

So two questions have to be answered, namely first whether there is a strong or at least a weak solution and second if there is some solution whether it is unique at least in some sense. A great tool in this theory was found by T. Yamada and S. Watanabe. They proved, that existence of a weak solution and pathwise uniqueness imply the existence of a unique strong solution, see [YW71].

There are many works which investigate the problem of existence or uniqueness under weaker assumptions than Lipschitz continuity. Beside [KR05], [FF11] and [Zha11] on which we will have a closer look later, we want to mention here some of these results. Strong existence and uniqueness could be obtained for example under local weak monotonicity and weak coercivity conditions on the coefficients. A proof can be found in the book *Stochastic Partial Differential Equations: An Introduction* of W. Liu and M. Röckner [LR15], which is based on [Kry99]. Furthermore, in their work *A study of a class of stochastic differential equations with non-Lipschitzian coefficients*, [FZ05], S. Fang and T. Zhang relaxed the Lipschitzian conditions mainly by a logarithmic factor. This means that the Lipschitz constant is multiplied with a function, depending on the distance, with special properties which are typically fulfilled by $\log(1/s)$, $\log(1/s) \cdot \log\log(1/s)$ and so on. Moreover, A. Yu. Veretennikov proved strong existence and uniqueness for bounded measurable coefficients if the diffusion matrix is nondegenerated, continuous and Lipschitz continuous in x, see [Ver78]. In [GM01] I. Gyöngy and T. Martínez relaxed this to locally unbounded drift, namely $b \in L^{2(d+1)}_{loc}(\mathbb{R}_+ \times \mathbb{R}^d)$ and b almost everywhere bounded by a constant plus some nonnegative function in $L^{d+1}(\mathbb{R}_+ \times \mathbb{R}^d)$.

In their work *Strong solutions of stochastic equations with singular time dependent drift* ([KR05]) N. Krylov and M. Röckner proved the existence of a unique strong solution in the white noise case, i.e. the diffusion coefficient σ is the unit matrix. The drift coefficient, defined on an open set $Q \subset \mathbb{R}^{d+1}$, is supposed to fulfill

$$\int_{\mathbb{R}} \left(\int_{\{x\in\mathbb{R}^d:(t,x)\in Q^n\}} |b(t,x)|^{p_n}\,dx \right)^{\frac{q_n}{p_n}} dt < \infty$$

for some $p_n \geq 2$, $q_n > 2$ such that

$$\frac{d}{p_n} + \frac{2}{q_n} < 1$$

and a sequence $(Q^n)_n$ of bounded open subsets of Q with $\overline{Q^n} \subset Q^{n+1}$ and $\bigcup_n Q^n = Q$. In 2011 E. Fedrizzi and F. Flandoli, [FF11], introduced a new method to prove the pathwise uniqueness under such conditions. The aim of this thesis is to extend their result to nonconstant diffusion. Therefore, we will have a look on this method in detail in the next section.

Also if the diffusion is not constant it is possible to get existence and uniqueness results under similar conditions on the drift. The most general result can be found in the work of X. Zhang *Stochastic homeomorphism flows of SDEs with singular drifts and Sobolev diffusion coefficients* [Zha11], respectively [Zha05] for the case $p = q$. There, the drift is in $L^q_{\text{loc}}(\mathbb{R}_+, L^p(\mathbb{R}^d))$ for some $p, q > 1$, fulfilling

$$\frac{d}{p} + \frac{2}{q} < 1. \tag{1}$$

The diffusion coefficient is uniformly continuous in x, locally uniformly with respect to t, nondegenerated, bounded and the gradient is also in $L^q_{loc}(\mathbb{R}_+, L^p(\mathbb{R}^d))$. The idea of the proof is to remove the drift by the so-called Zvonkin transformation, see [Zvo74], and use known results for SDEs with zero drift. This transformation is based on the solution u to the equation

$$\partial_t u + \sum_{i=1}^d b^i \partial_{x_i} u + \frac{1}{2}\sum_{i=1}^d \sum_{j=1}^d (\sigma\sigma^*)_{ij} \partial^2_{x_i x_j} u = 0, \qquad u(T,x) = x.$$

One difficulty is to show that this is a diffeomorphism to get a one-to-one correspondence between the solution X_t for the original SDE and the solution $u(t, X_t)$ for the transformed equation.

The method of E. Fedrizzi and F. Flandoli to prove pathwise uniqueness for constant diffusion is more intuitive. A central point of this work is to extend this proof and some results of [FF11] to nonconstant diffusion coefficients. Therefore, the following section is devoted to present their method in some detail.

1.2. Method of E. Fedrizzi and F. Flandoli

Let $X_t^{(1)}$, $X_t^{(2)}$ be two strong solutions to the equation

$$X_t = x + \int_0^t b(s, X_s)\,ds + W_t, \qquad t \in [0,T].$$

For $b \in L^q((0,T), L^p(\mathbb{R}^d))$ there exists a unique solution, see [Kry01], to the equation

$$\partial_t u + \frac{1}{2}\Delta u = -b \quad \text{on } [0,T], \quad u(T,x) = 0. \tag{2}$$

Denote this solution by U_b and apply Itô's formula to $U_b(t, X_t^{(i)})$. Since U_b is a solution to the above equation we get the following expression for the drift term:

$$\begin{aligned}\int_0^t b(s, X_s^{(i)})\, ds = U_b(0,x) - U_b(t, X_t^{(i)}) + \int_0^t \partial_x U_b(s, X_s^{(i)}) b(s, X_s^{(i)})\, ds \\ + \int_0^t \partial_x U_b(s, X_s^{(i)})\, dW_s.\end{aligned}$$

Now, the SDE may be rewritten by replacing the drift:

$$\begin{aligned}X_t^{(i)} &= x + U_b(0,x) - U_b(t, X_t^{(i)}) \\ &\quad + \int_0^t \partial_x U_b(s, X_s^{(i)}) b(s, X_s^{(i)})\, ds + \int_0^t \partial_x U_b(s, X_s^{(i)}) + I\, dW_s.\end{aligned} \tag{3}$$

The advantage of this reformulation is that the new drift term $\partial_x U_b \cdot b$ is in some way more regular than before. The solution U_b of (2) is an element of the Sobolev space $W^{1,q}((0,T), W^{2,p}(\mathbb{R}^d))$ and therefore has nice properties, e.g. $\partial_x U_b$ is Hölder continuous. If we define

$$\mathcal{T}(b)(t,x) := \partial_x U_b(t,x) b(t,x), \quad (t,x) \in [0,T] \times \mathbb{R}^d,$$

and take a solution $U_{\mathcal{T}(b)}$ of the equation

$$\partial_t u + \frac{1}{2}\Delta u = -\mathcal{T}(b) \quad \text{on } [0,T], \quad u(T,x) = 0,$$

an application of Itô's formula for $U_{\mathcal{T}(b)}$, yields an expression for the transformed drift term:

$$\begin{aligned}\int_0^t \mathcal{T}(b)(s, X_s^{(i)})\, ds &= U_{\mathcal{T}(b)}(0,x) - U_{\mathcal{T}(b)}(t, X_t^{(i)}) \\ &\quad + \int_0^t \partial_x U_{\mathcal{T}(b)}(s, X_s^{(i)}) b(s, X_s^{(i)})\, ds + \int_0^t \partial_x U_{\mathcal{T}(b)}(s, X_s^{(i)})\, dW_s.\end{aligned}$$

By replacing this term in equation (3), we get

$$\begin{aligned} X_t^{(i)} &= x + U_b(0,x) + U_{\mathcal{T}(b)}(0,x) - U_b(t, X_t^{(i)}) - U_{\mathcal{T}(b)}(t, X_t^{(i)}) \\ &\quad + \int_0^t \partial_x U_{\mathcal{T}(b)}(s, X_s^{(i)}) b(s, X_s^{(i)})\, ds \\ &\quad + \int_0^t \partial_x U_{\mathcal{T}(b)}(s, X_s^{(i)}) + \partial_x U_b(s, X_s^{(i)}) + I\, dW_s \end{aligned}$$

and by iteration and the convention $\mathcal{T}^{k+1}(b) = \partial_x U_{\mathcal{T}^k(b)} \cdot b$, $\mathcal{T}^0(b) = b$,

$$\begin{aligned} \underbrace{X_t^{(i)} + \sum_{k=0}^{n} U_{\mathcal{T}^k(b)}(t, X_t^{(i)})}_{=:Y_t^{(i,n)}} &= x + \sum_{k=0}^{n} U_{\mathcal{T}^k(b)}(0,x) + \int_0^t \mathcal{T}^{n+1}(b)(s, X_s^{(i)})\, ds \\ &\quad + \int_0^t \underbrace{\sum_{k=0}^{n} \partial_x U_{\mathcal{T}^k(b)}(s, X_s^{(i)}) + I}_{=:\sigma^{(n)}(s, X_s^{(i)})}\ dW_s. \end{aligned}$$

Then one can prove that

$$\begin{aligned} &\mathbb{E}\left[|X_t^{(1)} - X_t^{(2)}|\right] \\ &\le C\mathbb{E}\left[e^{A_t^{(n)}}\right]^{\frac{1}{2}} \left(\underbrace{\mathbb{E}\left[\int_0^t |X_s^{(1)} - X_s^{(2)}||\mathcal{T}^{n+1}(b)(s, X_s^{(1)}) - \mathcal{T}^{n+1}(b)(s, X_s^{(2)})|\, ds\right]}_{I_1} \right. \\ &\quad \left. + \mathbb{E}\left[\underbrace{\int_0^t e^{-A_s^{(n)}} \left\langle Y_s^{(1,n)} - Y_s^{(2,n)}, \left(\sigma^{(n)}(s, X_s^{(1)}) - \sigma^{(n)}(s, X_s^{(2)})\right)\, dW_s \right\rangle}_{I_2}\right]\right)^{\frac{1}{2}}, \end{aligned} \tag{4}$$

where

$$A_t^{(n)} := \int_0^t \frac{\left|\sigma^{(n)}(s, X_s^{(1)}) - \sigma^{(n)}(s, X_s^{(2)})\right|^2}{\left|Y_s^{(1,n)} - Y_s^{(2,n)}\right|^2} \mathbb{1}_{\{Y_s^{(1,n)} \neq Y_s^{(2,n)}\}}\, ds.$$

By proving that $\mathbb{E}[e^{A_t^{(n)}}]$ is uniformly bounded in n, that I_1 converges to 0 for $n \to \infty$ and that I_2 is a martingale, one gets pathwise uniqueness.

1.3. Aim and progress of the thesis

The aim of this thesis is to generalize the method of E. Fedrizzi and F. Flandoli to time and space dependent diffusion. Instead of

$$\partial_t u + \frac{1}{2}\Delta u = -f$$

one considers equations of the form

$$\partial_t u + \frac{1}{2}\sum_{i=1}^{d}\sum_{j=1}^{d}(\sigma\sigma^*)_{ij}\partial^2_{x_i x_j} u = -f$$

to transform the SDE. Along with some other technical issues, the nonconstant diffusion leads to additional terms in the stochastic integral of the reformulated SDE. Nevertheless the core of the proof remains the same as in [FF11] where we have to handle the fact that a solution to the SDE is in general not a Brownian motion. This is the case if $\sigma = I$ and it was a crucial point in the proof of E. Fedrizzi and F. Flandoli. This property enabled them to use Girsanov's formula and exponential estimates for Brownian motion which are not applicable in our generalization. As a compensation, we successfully use Krylov estimates. We therefore prove a version of Lemma 5.1 from [Kry86] for different integrability in time and space stated as Lemma 3.1 and proved in Section A.3. The price we have to pay is that we have to assume $p, q > 2(d+1)$. Since the estimates are based on solutions to PDEs it should be possible to extend it, maybe up to the case p, q fulfilling condition (1), but in this thesis we restrict to these stonger assumptions on p and q.

Beside the ordinary Krylov-type estimates we also need similar ones on conditional expectations, which we formulate and prove in Section 4.1. We only have to assume that the diffusion coefficient is bounded, nondegenerated, the drift is in $L^q((0,T), L^p(\mathbb{R}^d))$ and

$$\mathbb{P}\left(\int_0^T |b(t, X_t)|dt < \infty\right) = 1.$$

Up to our knowledge this has not been done yet under these general assumptions. For bounded b a version can be found in [Kry09] and for σ uniformly continuous in x, local uniformly continuous with respect to t in [Zha11]. Only for an estimate on the linear combination of two solutions as in Proposition 4.4 continuity of the diffusion term is required.

For simplicity we will state our result under global assumptions, but there are no difficulties to extend it by localization techniques, e.g. in the same way as in [Zha11].

The result of X. Zhang, [Zha11], is close to ours. The assumptions are more general with respect to the integrability of b and $\partial_x\sigma$ since we have to assume $p, q > 2(d+1)$, which comes from our Krylov estimates, but could be possible extended to p, q fulfilling (1), which also X. Zhang requires. Furthermore, the assumptions on drift and diffusion coefficients are the same except the continuity condition on σ. Instead of requiring uniform continuity in x, locally uniformly with respect to t, which gives directly (see Remark 10.4 of [KR05]) the solvability of equations of the form

$$\partial_t u + \frac{1}{2}\sum_{i=1}^{d}\sum_{j=1}^{d}(\sigma\sigma^*)_{ij}\partial^2_{x_i x_j} u + \sum_{i=1}^{d} b^i \partial_{x_i} u = f,$$

we assume σ to be continuous and such that there exists a solution to the equation

$$\partial_t u + \frac{1}{2}\sum_{i=1}^{d}\sum_{j=1}^{d}(\sigma\sigma^*)_{ij}\partial^2_{x_i x_j} u = f,$$

see Assumptions 2.2 and 2.3. Beside these similarities as far as assumptions are concerned, our method of proof is completely different and more probabilistic and also much simpler at least from our point of view. We are able to prove pathwise uniqueness directly by estimating $\mathbb{E}[|X_t^{(1)} - X_t^{(2)}|]$ in a similar way as in (4).

1.4. Future directions

One step in further research could be an optimization of the proof method. Maybe it is possible to avoid the exponential estimates on the transformed diffusion by stopping time arguments. A technique in this direction was recently developed by G. Da Prato, F. Flandoli, M. Röckner and A. Yu. Veretennikov in [DPFRV16].
Another very interesting issue is the generalization to infinite dimensions. This was a strong motivation for this thesis, since it seems achievable and would be a great step forward in the theory of stochastic partial differential equations. In contrast to the finite dimensional case we do not have elliptic regularity for partial differential equations on Hilbert spaces. To avoid difficulties it could be a good approach to start with exponential integrable coefficients.

1.5. Structure

The second chapter is devoted to basic definitions, especially the involved spaces are introduced. Then we state our assumptions on the coefficients of the stochastic differential equation and the result about pathwise uniqueness.
The third chapter deals with the transformation of the SDE which is necessary to prove pathwise uniqueness. Since the transformation is based on Itô's formula, we first show that under our assumptions it is applicable for functions in the mixed norm Sobolev space $W^{1,q}((0,T), W^{2,p}(\mathbb{R}^d))$ before we explain the transformation in detail.
The fourth chapter states all necessary tools to prove pathwise uniqueness on a small interval. First we give some useful facts of the involved functions and the relation between a solution to the original SDE and the transformed equation. Then we prove the Krylov estimates for conditional expectations, which then give us a uniform exponential estimate for the transformed diffusion. After this we show the convergence of the difference between the transformed drift terms of two solutions and in the end, we prove that

under our assumptions solutions of (5) have finite first and second moments.
Chapter 5 is devoted to the proof of pathwise uniqueness. Since we stated some necessary tools only up to some possible small T, we first prove it on $[0,T]$, before we show that it is extendable to arbitrarily large intervals. Therefore, in the end we get pathwise uniqueness on the whole interval of the original SDE.
In the appendix we list some small lemmas which we need in the proofs before. For the sake of completeness they are all given with proofs also if some of them are easy and just little generalizations of well known results. We start with some facts about our mixed norm spaces, especially approximation by smooth functions. It seems that this has not been done yet rigorously and therefore we prove them in detail. Then we state an easy mean-value inequality and prove a Sobolev embedding. In the end we give some Krylov estimates. The proofs for the inequalities of expectations are very similar to the ones for conditional expectations. But since we need some of these results for the proofs in the conditional case, we also give here the proofs in detail.

Acknowledgment

First and foremost, I thank Prof. Dr. Michael Röckner for supervising this work, for his continuous support and advice.
Furthermore, I am grateful to Nora Müller and Daniel Altemeier for giving me some valuable feedback on this thesis as well as to Sarah Gertz for her advice on grammar and phrasing. I would also like to thank my friends and family for their support in the last few years.
Financial support of the DFG through the IRTG 1132 *Stochastic and Real World Models* and the CRC 701 *Spectral Structures and Topological Methods in Mathematics* is gratefully acknowledged, as well as of the Bielefeld Young Researchers' Fund through the doctorate completion scholarship.

2. Preliminaries and main result

In this chapter, basic concepts for this thesis as mixed-norm Sobolev spaces are introduced. In particular, we formulate the required assumptions on the coefficients of the SDE and present our result about pathwise uniqueness. Besides, there will be some short notations introduced which will be used throughout the thesis.

Definition 2.1. *For $p, q \in (1,\infty)$ we define*

$$\|f\|_{L_p^q(T)} := \left(\int_0^T \left(\int_{\mathbb{R}^d} |f(t,x)|^p \, dx \right)^{\frac{q}{p}} dt \right)^{\frac{1}{q}},$$

where $|\cdot|$ denotes the Hilbert-Schmidt norm.
We define $L_p^q(T)$ to be the space of measurable functions $f : [0,T] \times \mathbb{R}^d \to \mathbb{R}^d$ (respectively $\mathbb{R}^{d\times m}$) such that $\|f\|_{L_p^q(T)} < \infty$.
Furthermore

$$W^{1,2}_{q,p}(T) := \left\{ f : [0,T] \times \mathbb{R}^d \to \mathbb{R}^d \;\middle|\; f, \partial_t f, \partial_x f, \partial_x^2 f \in L_p^q(T) \right\},$$

where ∂_t, ∂_x, ∂_x^2 denote weak derivatives with respect to time, respectively space. The associated norm is given by

$$\|f\|_{W^{1,2}_{q,p}(T)} := \|f\|_{L_p^q(T)} + \|\partial_t f\|_{L_p^q(T)} + \|\partial_x f\|_{L_p^q(T)} + \|\partial_x^2 f\|_{L_p^q(T)}.$$

If we omit the T, we mean that we take $\mathbb{R}_+$ instead of $[0,T]$.

We consider the SDE

$$X_t = x + \int_0^t b(s, X_s)\, ds + \int_0^t \sigma(s, X_s)\, dW_s, \quad t \in [0,T], \tag{5}$$

where W is an m-dimensional standard Wiener process on a filtered probability space $(\Omega, (\mathcal{F}_t)_t, \mathbb{P})$, with $(\mathcal{F}_t)_t$ fulfilling the usual conditions, $x \in \mathbb{R}^d$ and $b : [0,T] \times \mathbb{R}^d \to \mathbb{R}^d$, $\sigma : [0,T] \times \mathbb{R}^d \to \mathbb{R}^{d\times m}$ are measurable functions with the following properties:

Assumption 2.2. *For some $p, q > 2(d+1)$ we have*

(c1) $b \in L_p^q(T)$,

(c2) *σ is continuous in (t,x),*

(c3) *σ is nondegenerated, i.e. there exists a constant $c_\sigma > 0$ such that*

$$\langle \sigma\sigma^*(t,x)\xi, \xi \rangle \geq c_\sigma \langle I\xi, \xi \rangle \qquad \forall\ \xi \in \mathbb{R}^d \quad \forall\ (t,x) \in [0,T] \times \mathbb{R}^d,$$

where σ^* denotes the transposed matrix of σ, (c4) *σ is bounded by a constant $\tilde{c}_\sigma$,*

(c5) *$\partial_x \sigma \in L_p^q(T)$.*

Assumption 2.3. *Let σ be such that for all $f \in L_p^q(T)$ there is a solution $u \in W_{q,p}^{1,2}(T)$ to the partial differential equation*

$$\partial_t u + \frac{1}{2} \sum_{i=1}^d \sum_{j=1}^d (\sigma\sigma^*)_{ij} \partial^2_{x_i x_j} u = f \quad \text{on } [0,T], \quad u(T,x) = 0,$$

such that

$$\|u\|_{L_p^q(T)} \leq C \|f\|_{L_p^q(T)},$$

where C is independent of f and increasing in T.

Remark 2.4. *Assumption 2.3 seems to be massive restrictive, but in fact it is proven for a large class of functions. If σ is independent of t the result can be found in [Kry01]. Based on this one can prove that it holds also for σ uniformly continuous in $x \in \mathbb{R}^d$, uniformly continuous with respect to t, see Remark 10.4 in [KR05]. If $q \geq p$ and σ satisfies a vanishing mean oscillation condition the assumption is also fulfilled, see [Kry07].*

Definition 2.5 (weak/strong solution)**.** *A weak solution for equation* (5) *is a pair (X, W) on a filtered probability space $(\Omega, (\mathcal{F}_t)_t, \mathbb{P})$ such that X_t is $\mathcal{F}_t$-adapted, W_t is an $\mathcal{F}_t$-Brownian motion and (X, W) solves equation* (5)*.*
Given a Brownian motion W on a probability space, a strong solution for equation (5) *is a continuous process which is adapted to the filtration generated by W and solves equation* (5)*.*

Definition 2.6 (Pathwise Uniqueness)**.** *We say that pathwise uniqueness holds for equation* (5) *if for two weak solutions (X, W), $(\tilde{X}, \tilde{W})$, defined on the same probability space, we have that $X_0 = \tilde{X}_0$ and $W = \tilde{W}$ imply*

$$\mathbb{P}\left(X_t = \tilde{X}_t \quad \forall\, t \in [0,T]\right) = 1.$$

Theorem 2.7 (Main result)**.** *Under Assumptions 2.2 and 2.3, we have pathwise uniqueness in the set of continuous processes which fulfill*

$$\mathbb{P}\left(\int_0^T |b(s, X_s)|\, ds < \infty\right) = 1. \tag{6}$$

Notation 2.8. *For two solutions $X_t^{(1)}$, $X_t^{(2)}$ to SDE* (5)*, defined on the same probability space, with initial values $x^{(1)}$, $x^{(2)}$ and the same Brownian motion, we define for all*

$\lambda \in [0,1]$, $R > 0$ *and* $t \in [0,T]$

$$\begin{aligned}
X_t^\lambda &:= \lambda X_t^{(1)} + (1-\lambda) X_t^{(2)}, \\
x^\lambda &:= \lambda x^{(1)} + (1-\lambda) x^{(2)}, \\
b^\lambda(t, X_t^{(1)}, X_t^{(2)}) &:= \lambda b(t, X_t^{(1)}) + (1-\lambda) b(t, X_t^{(2)}), \\
\sigma^\lambda(t, X_t^{(1)}, X_t^{(2)}) &:= \lambda \sigma(t, X_t^{(1)}) + (1-\lambda) \sigma(t, X_t^{(2)}), \\
\tau_R^\lambda &:= \inf\left\{ t \geq 0 \ : \ \left| X_t^\lambda \right| > R \right\}, \\
\tau_R &:= \inf\left\{ t \geq 0 \ : \ |X_t^{(1)}| > R \ \mathit{or} \ |X_t^{(2)}| > R \right\},
\end{aligned}$$

and

$$a_t^\lambda := \frac{1}{2} \left(\lambda \sigma(t, X_t^{(1)}) + (1-\lambda) \sigma(t, X_t^{(2)}) \right) \cdot \left(\lambda \sigma(t, X_t^{(1)}) + (1-\lambda) \sigma(t, X_t^{(2)}) \right)^* .$$

In the following, whenever we speak of two solutions, we mean two weak solutions defined on the same probability space with the same Brownian motion.
Furthermore by $C > 0$ we always denote various finite constants, where we often indicate the dependence of parameters by writing them in brackets.

3. Transformation of the SDE

This chapter consists of a detailed study of our before mentioned transformation of the SDE. To this end, an appropriate version of Itô's formula for functions in $W^{1,2}_{q,p}(T)$ will be established in the first section. The second covers the formulation and study of the transformation.

3.1. Itô's formula for mixed-norm Sobolev functions

We formulate a version of Itô's formula for functions in $W^{1,2}_{q,p}(T)$ in Proposition 3.3. The proof relies on two auxiliaries – a Krylov-type estimate and a Sobolev embedding. Those two are formulated as Lemmas 3.1 and 3.2 with both proofs deferred to the appendix.

Lemma 3.1. *Let* (c1), (c3), (c4) *of Assumption 2.2 be fulfilled and X_t be a solution to* (5) *such that condition* (6) *holds. Then we have for every $v, r \geq d+1$ and any nonnegative measurable function $f : [0,T] \times \mathbb{R}^d \to \mathbb{R}$*

$$\mathbb{E}\left[\int_0^T f(t, X_t)\,dt\right] \leq C(T, d, p, q, v, r, c_\sigma, \tilde{c}_\sigma, \|b\|_{L^q_p(T)})\|f\|_{L^r_v(T)}.$$

Lemma 3.2. *For all $u \in W^{1,2}_{q,p}(T)$, there exists a version of u such that*

$$\sup_{(t,x)\in[0,T]\times\mathbb{R}^d} |u(t,x)| \leq C\|u\|_{W^{1,2}_{q,p}(T)},$$

where C is independent of u. In particular this version is continuous.

Proposition 3.3. *(Itô's formula) Let $u \in W^{1,2}_{q,p}(T)$,* (c1), (c3), (c4) *of Assumption 2.2 be fulfilled and X_t a solution to* (5) *such that condition* (6) *holds. Then there exists a version of u such that for $0 \leq s \leq t \leq T$ we have*

$$\begin{aligned} u(t, X_t) = u(s, X_s) &+ \int_s^t \partial_t u(r, X_r)\,dr + \int_s^t \partial_x u(r, X_r) b(r, X_r)\,dr \\ &+ \int_s^t \partial_x u(r, X_r)\sigma(r, X_r)\,dW_r \\ &+ \frac{1}{2}\int_s^t \sum_{i=1}^d \sum_{j=1}^d (\sigma\sigma^*(r, X_r))_{ij}\partial^2_{x_i x_j} u(r, X_r)\,dr \qquad \mathbb{P}\text{-almost surely.} \end{aligned}$$

Proof. By Lemma A.5 there exists a sequence $(u_n)_n$ in $\mathcal{C}^\infty((0,T) \times \mathbb{R}^d)$ which converges to u in $W^{1,2}_{q,p}(T)$. Then Itô's formula, see e.g. [KS91] Chapter 3 Theorem 3.6, yields

$$\begin{aligned} u_n(t,X_t) = u_n(s,X_s) &+ \int_s^t \partial_t u_n(r,X_r)\,dr + \int_s^t \partial_x u_n(r,X_r)b(r,X_r)\,dr \\ &+ \int_s^t \partial_x u_n(r,X_r)\sigma(r,X_r)\,dW_r \\ &+ \frac{1}{2}\int_s^t \sum_{i=1}^d \sum_{j=1}^d (\sigma\sigma^*(r,X_r))_{ij}\partial^2_{x_i x_j} u_n(r,X_r)\,dr \end{aligned}$$

$\mathbb{P}$-almost surely for every $n \in \mathbb{N}$. From Lemma 3.2 we know that there exists a version of u such that

$$|u(t,X_t) - u_n(t,X_t)| \le C\|u - u_n\|_{W^{1,2}_{q,p}(T)}.$$

Therefore, $u_n(t,X_t)$ converges $\mathbb{P}$-almost surely to $u(t,X_t)$ as $n \to \infty$. With Lemma 3.1 we obtain

$$\begin{aligned} &\mathbb{E}\left[\left|\int_s^t \partial_t u(r,X_r)\,dr - \int_s^t \partial_t u_n(r,X_r)\,dr\right|\right] \\ &\le \mathbb{E}\left[\int_s^t |\partial_t u(r,X_r) - \partial_t u_n(r,X_r)|\,dr\right] \\ &\le C\|\partial_t u - \partial_t u_n\|_{L^q_p(T)}. \end{aligned}$$

And using Hölders inequality twice leads to

$$\begin{aligned} &\mathbb{E}\left[\left|\int_s^t \partial_x u(r,X_r)b(r,X_r)\,dr - \int_s^t \partial_x u_n(r,X_r)b(r,X_r)\,dr\right|\right] \\ &\le \mathbb{E}\left[\int_s^t |\partial_x u(r,X_r)b(r,X_r) - \partial_x u_n(r,X_r)b(r,X_r)|\,dr\right] \\ &\le \mathbb{E}\left[\int_s^t |b(r,X_r)|\cdot|\partial_x u(r,X_r) - \partial_x u_n(r,X_r)|\,dr\right] \\ &\le \mathbb{E}\left[\int_s^t |b(r,X_r)|^2\,dr\right]^{\frac{1}{2}} \cdot \mathbb{E}\left[\int_s^t |\partial_x u(r,X_r) - \partial_x u_n(r,X_r)|^2\,dr\right]^{\frac{1}{2}}. \end{aligned}$$

One more application of Lemma 3.1 yields then

$$
\begin{aligned}
&\mathbb{E}\left[\left|\int_s^t \partial_x u(r,X_r)b(r,X_r)\,dr - \int_s^t \partial_x u_n(r,X_r)b(r,X_r)\,dr\right|\right] \\
&\le C\||b|^2\|^{\frac{1}{2}}_{L^{q/2}_{p/2}(T)} \cdot \||\partial_x u - \partial_x u_n|^2\|^{\frac{1}{2}}_{L^{q/2}_{p/2}(T)} \\
&\le C\|\partial_x u - \partial_x u_n\|_{L^q_p(T)}.
\end{aligned}
$$

For the last deterministic integral, we receive a similar estimate:

$$
\begin{aligned}
&\mathbb{E}\left[\left|\frac{1}{2}\int_s^t \sum_{i=1}^d \sum_{j=1}^d (\sigma\sigma^*(r,X_r))_{ij}\left(\partial^2_{x_i x_j} u(r,X_r) - \partial^2_{x_i x_j} u_n(r,X_r)\right)\,dr\right|\right] \\
&\le \mathbb{E}\left[\frac{1}{2}\int_s^t |\sigma(r,X_r)|^2 \cdot |\partial^2_x u(r,X_r) - \partial^2_x u_n(r,X_r)|\,dr\right] \\
&\le \frac{1}{2}\tilde{c}^2_\sigma \mathbb{E}\left[\int_s^t |\partial^2_x u(r,X_r) - \partial^2_x u_n(r,X_r)|\,dr\right] \\
&\le C\|\partial^2_x u - \partial^2_x u_n\|_{L^q_p(T)}.
\end{aligned}
$$

Finally, for the stochastic integral, we have by similar estimates and the Itô-Isometry

$$
\begin{aligned}
&\mathbb{E}\left[\left|\int_s^t (\partial_x u(r,X_r) - \partial_x u_n(r,X_r))\sigma(r,X_r)\,dW_r\right|\right] \\
&\le \mathbb{E}\left[\left|\int_s^t (\partial_x u(r,X_r) - \partial_x u_n(r,X_r))\sigma(r,X_r)\,dW_r\right|^2\right]^{\frac{1}{2}} \\
&= \mathbb{E}\left[\int_s^t |(\partial_x u(r,X_r) - \partial_x u_n(r,X_r))\sigma(r,X_r)|^2\,dr\right]^{\frac{1}{2}} \\
&\le \tilde{c}^2_\sigma \mathbb{E}\left[\int_s^t |\partial_x u(r,X_r) - \partial_x u_n(r,X_r)|^2\,dr\right]^{\frac{1}{2}} \\
&\le C\||\partial_x u - \partial_x u_n|^2\|^{\frac{1}{2}}_{L^{q/2}_{p/2}(T)} \\
&= C\|\partial_x u - \partial_x u_n\|_{L^q_p(T)}.
\end{aligned}
$$

Therefore, there exist a subsequence $(u_{n_k})_k$ such that

$$\begin{aligned}
&\int\limits_s^t \partial_t u_{n_k}(r,X_r)\,dr + \int\limits_s^t \partial_x u_{n_k}(r,X_r)b(r,X_r)\,dr + \int\limits_s^t \partial_x u_{n_k}(r,X_r)\sigma(r,X_r)\,dW_r \\
&\qquad + \frac{1}{2}\int\limits_s^t \sum_{i=1}^d \sum_{j=1}^d (\sigma\sigma^*(r,X_r))_{ij}\partial^2_{x_i x_j} u_{n_k}(r,X_r)\,dr \\
&\xrightarrow{k\to\infty} \int\limits_s^t \partial_t u(r,X_r)\,dr + \int\limits_s^t \partial_x u(r,X_r)b(r,X_r)\,dr + \int\limits_s^t \partial_x u(r,X_r)\sigma(r,X_r)\,dW_r \\
&\qquad + \frac{1}{2}\int\limits_s^t \sum_{i=1}^d \sum_{j=1}^d (\sigma\sigma^*(r,X_r))_{ij}\partial^2_{x_i x_j} u(r,X_r)\,dr \qquad \mathbb{P}\text{-a.s.}
\end{aligned}$$

and therefore, we have Itô's formula for functions in $W^{1,2}_{q,p}(T)$. $\square$

3.2. Transformation of the SDE

We may transform SDE (5) by means of solutions to a particular PDE, which is stated below in (7). Then an application of Itô's formula is used to replace the drift term. By iteration, we are able to reformulate the equation as stated in (14).

Assume that b and σ fulfill Assumptions 2.2 and 2.3, then for every $f \in L^q_p(T)$ there exists a solution $u \in W^{1,2}_{q,p}(T)$ to the equation

$$\partial_t u + \frac{1}{2}\sum_{i=1}^d \sum_{j=1}^d (\sigma\sigma^*)_{ij}\partial^2_{x_i x_j} u = f, \quad \text{on } [0,T], \quad u(T,x) = 0 \tag{7}$$

such that

$$\|u\|_{W^{1,2}_{q,p}(T)} \leq C\|f\|_{L^q_p(T)}, \tag{8}$$

where C does not depend on f and is increasing in T. Then we have

$$\begin{aligned}
\sup_{(t,x)\in[0,T]\times\mathbb{R}^d} |\partial_x u(t,x)| &\leq \sup_{(t,x)\in[0,T]\times\mathbb{R}^d} |\partial_x u(t,x) - \partial_x u(T,x)| + |\partial_x u(T,x)| \\
&= \sup_{(t,x)\in[0,T]\times\mathbb{R}^d} |\partial_x u(t,x) - \partial_x u(T,x)|.
\end{aligned}$$

By the Hölder continuity of $\partial_x u$, see [KR05] Lemma 10.2, this together with (8) leads to

$$\begin{aligned}
\sup_{(t,x)\in[0,T]\times\mathbb{R}^d} |\partial_x u(t,x)| &\leq \sup_{(t,x)\in[0,T]\times\mathbb{R}^d} C(p,q,\varepsilon)|T-t|^{\frac{\varepsilon}{2}}\|u\|_{W^{1,2}_{q,p}(T)} \\
&\leq C(p,q,\varepsilon,T)T^{\frac{\varepsilon}{2}}\|f\|_{L^q_p(T)} \qquad (9)
\end{aligned}$$

for every $\varepsilon \in (0,1)$, which fulfills

$$\varepsilon + \frac{d}{p} + \frac{2}{q} < 1.$$

Since $C(p,q,\varepsilon,T)$ is increasing in T, we can assume the constant in front of $\|f\|_{L_p^q(T)}$ to be as small as we want by choosing T appropriate. This will be of importance in Lemma 4.1. Now, let U_b a solution to the equation

$$\partial_t u + \frac{1}{2}\sum_{i=1}^{d}\sum_{j=1}^{d}(\sigma\sigma^*)_{ij}\partial^2_{x_i x_j} u = -b \quad \text{on } [0,T], \qquad u(T,x) = 0. \tag{10}$$

Using Itô's formula for functions in $W^{1,2}_{q,p}(T)$ (Proposition 3.3), we get

$$\begin{aligned} U_b(t,X_t) = U_b(0,x) &+ \int_0^t \partial_x U_b(s,X_s) b(s,X_s)\,ds + \int_0^t \partial_x U_b(s,X_s)\sigma(s,X_s)\,dW_s \\ &+ \int_0^t \partial_t U_b(s,X_s) + \frac{1}{2}\sum_{i=1}^{d}\sum_{j=1}^{d}(\sigma\sigma^*(s,X_s))_{ij}\partial^2_{x_i x_j} U_b(s,X_s)\,ds. \end{aligned}$$

Here, we use that U_b is a solution to PDE (10), to obtain

$$\begin{aligned} U_b(t,X_t) = U_b(0,x) &+ \int_0^t \partial_x U_b(s,X_s) b(s,X_s)\,ds + \int_0^t \partial_x U_b(s,X_s)\sigma(s,X_s)\,dW_s \\ &- \int_0^t b(s,X_s)\,ds. \end{aligned}$$

That implies

$$\begin{aligned} \int_0^t b(s,X_s)\,ds = U_b(0,x) - U_b(t,X_t) &+ \int_0^t \partial_x U_b(s,X_s) b(s,X_s)\,ds \\ &+ \int_0^t \partial_x U_b(s,X_s)\sigma(s,X_s)\,dW_s. \end{aligned}$$

Now, we define

$$\mathcal{T}(b) := \partial_x U_b \cdot b$$

and transform SDE (5) by replacing the drift term:

$$\begin{aligned} X_t = x + U_b(0,x) - U_b(t,X_t) &+ \int_0^t \mathcal{T}(b)(s,X_s)\,ds \\ &+ \int_0^t \partial_x U_b(s,X_s)\sigma(s,X_s) + \sigma(s,X_s)\,dW_s. \end{aligned} \tag{11}$$

Note, that $\mathcal{T}(b) \in L_p^q(T)$ since $\partial_x U_b$ is bounded and $b \in L_p^q(T)$. Next, let $U_{\mathcal{T}(b)}$ be a solution to the equation

$$\partial_t u + \frac{1}{2}\sum_{i=1}^{d}\sum_{j=1}^{d}(\sigma\sigma^*)_{ij}\partial^2_{x_i x_j} u = -\mathcal{T}(b) \quad \text{on } [0,T], \qquad u(T,x) = 0.$$

Using again Itô's formula (Proposition 3.3), we get

$$\begin{aligned}
U_{\mathcal{T}(b)}(t,X_t) = U_{\mathcal{T}(b)}(0,x) &+ \int_0^t \partial_x U_{\mathcal{T}(b)}(s,X_s) b(s,X_s)\,ds + \int_0^t \partial_x U_{\mathcal{T}(b)}(s,X_s)\sigma(s,X_s)\,dW_s \\
&+ \int_0^t \partial_t U_{\mathcal{T}(b)}(s,X_s) + \frac{1}{2}\sum_{i=1}^{d}\sum_{j=1}^{d}(\sigma\sigma^*(s,X_s))_{ij}\partial^2_{x_i x_j} U_{\mathcal{T}(b)}(s,X_s)\,ds \\
&= U_{\mathcal{T}(b)}(0,x) + \int_0^t \partial_x U_{\mathcal{T}(b)}(s,X_s) b(s,X_s)\,ds + \int_0^t \partial_x U_{\mathcal{T}(b)}(s,X_s)\sigma(s,X_s)\,dW_s \\
&- \int_0^t \mathcal{T}(b)(s,X_s)\,ds,
\end{aligned}$$

and therefore

$$\begin{aligned}
\int_0^t \mathcal{T}(b)(s,X_s)\,ds = U_{\mathcal{T}(b)}(0,x) - U_{\mathcal{T}(b)}(t,X_t) &+ \int_0^t \partial_x U_{\mathcal{T}(b)}(s,X_s) b(s,X_s)\,ds \\
&+ \int_0^t \partial_x U_{\mathcal{T}(b)}(s,X_s)\sigma(s,X_s)\,dW_s.
\end{aligned}$$

Again, we define

$$\mathcal{T}^2(b) := \partial_x U_{\mathcal{T}(b)} \cdot b$$

and replace the drift in the transformed SDE (11):

$$\begin{aligned}
X_t = x + U_b(0,x) + U_{\mathcal{T}(b)}(0,x) - U_b(t,X_t) - U_{\mathcal{T}(b)}(t,X_t) + \int_0^t \mathcal{T}^2(b)(s,X_s)\,ds \\
+ \int_0^t \partial_x U_b(s,X_s)\sigma(s,X_s) + \partial_x U_{\mathcal{T}(b)}(s,X_s)\sigma(s,X_s) + \sigma(s,X_s)\,dW_s.
\end{aligned}$$

Iteration yields after $n+1$ steps

$$X_t = x + \sum_{k=0}^{n} U_{\mathcal{T}^k(b)}(0,x) - \sum_{k=0}^{n} U_{\mathcal{T}^k(b)}(t,X_t) + \int_0^t \mathcal{T}^{n+1}(b)(s,X_s)\,ds$$
$$+ \int_0^t \sum_{k=0}^{n} \partial_x U_{\mathcal{T}^k(b)}(s,X_s)\sigma(s,X_s) + \sigma(s,X_s)\,dW_s \tag{12}$$

with the convention

$$\mathcal{T}^0(b) = b \quad \text{and} \quad \mathcal{T}^{k+1}(b) = \partial_x U_{\mathcal{T}^k(b)} \cdot b.$$

We define

$$U^{(n)}(t,x) := \sum_{k=0}^{n} U_{\mathcal{T}^k(b)}(t,x)$$

and therefore, SDE (12) becomes

$$X_t = x + U^{(n)}(0,x) - U^{(n)}(t,X_t) + \int_0^t \mathcal{T}^{n+1}(b)(s,X_s)\,ds$$
$$+ \int_0^t \left(\partial_x U^{(n)}(s,X_s) + I\right)\sigma(s,X_s)\,dW_s. \tag{13}$$

For two solutions $X_t^{(1)}$, $X_t^{(2)}$ we define

$$\begin{aligned} Y_t^{(i,n)} &:= X_t^{(i)} + U^{(n)}(t,X_t^{(i)}), \\ b^{(n)}(t,X_t^{(i)}) &:= \mathcal{T}^{n+1}(b)(t,X_t^{(i)}), \\ \sigma^{(n)}(t,X_t^{(i)}) &:= \left(\partial_x U^{(n)}(t,X_t^{(i)}) + I\right)\sigma(t,X_t^{(i)}). \end{aligned}$$

Then equation (13) reads

$$Y_t^{(i,n)} = Y_0^{(i,n)} + \int_0^t b^{(n)}(s,X_s^{(i)})\,ds + \int_0^t \sigma^{(n)}(s,X_s^{(i)})\,dW_s. \tag{14}$$

4. Some helpful lemmas

In this chapter we present the necessary tools to prove our main result. First, we give some useful properties of the involved functions and a contraction property between $X_t^{(i)}$ and $Y_t^{(i,n)}$ in Lemma 4.1. Then we prove two Krylov-type estimates for conditional expectations. A version of Lemma 5.1 from [Kry86] is stated in Lemma 4.2 under very general assumptions on the coefficients. It is sufficient to assume $b \in L_p^q(T)$ and σ to be bounded and nondegenerated. The price we have to pay is that it is only applicable to functions in $L_v^r(T)$ with $r, v \geq d+1$. The second Krylov-type estimate, namely Proposition 4.4, requires also the continuity of the diffusion coefficient. Based on this inequality we prove an exponential estimate for the transformed diffusion before we show convergence of the difference between the transformed drift terms of two solutions. In the end, we will prove that under our conditions every solution to (5) has finite first and second moments.
The following Lemma is similar to Lemma 7 in [FF11] and so is the proof. But we give it in detail to make clear that it works in the same way for our extended transformation.

Lemma 4.1. *Let (c1), (c3), (c4) of Assumption 2.2 and Assumption 2.3 be fulfilled and $X_t^{(1)}$, $X_t^{(2)}$ be two solutions to (5) such that (6) holds. Then there exists $T_0 \leq T$ such that for all $T' \in (0, T_0]$ we have*

(i) $\|\mathcal{T}^n(b)\|_{L_p^q(T')} \leq \dfrac{1}{2^n}\|b\|_{L_p^q(T')}$,

(ii) $\displaystyle\sum_{k=0}^{n} \sup_{(t,x)\in[0,T']\times\mathbb{R}^d} |\partial_x U_{\mathcal{T}^k(b)}(t,x)| \leq \frac{1}{2}$,

(iii) $\|\partial_x^2 U^{(n)}\|_{L_p^q(T')} \leq C$ *for some constant $C > 0$, independent of n, and*

(iv) $\left|Y_t^{(1,n)} - Y_t^{(2,n)}\right| \leq \dfrac{3}{2}\left|X_t^{(1)} - X_t^{(2)}\right|$,

$\left|X_t^{(1)} - X_t^{(2)}\right| \leq 2\left|Y_t^{(1,n)} - Y_t^{(2,n)}\right|$ *for all $t \in (0, T']$.*

Proof. (i) Set

$$\varepsilon = \frac{1}{4(\|b\|_{L_p^q(T)} + 1)}$$

and choose T_0 such that

$$\sup_{(t,x)\in[0,T']\times\mathbb{R}^d} |\partial_x U_f(t,x)| \leq \varepsilon \|f\|_{L_p^q(T')} \tag{15}$$

for all $T' \in (0, T_0]$ and $f \in L_p^q(T')$, where U_f denotes a solution to

$$\partial_t u + \frac{1}{2}\sum_{i=1}^{d}\sum_{j=1}^{d}(\sigma\sigma^*)_{ij}\partial_{x_i x_j}^2 u = -f \quad \text{on } [0,T'], \quad u(T',x) = 0.$$

The possibility of choosing such a T_0 is given by (9). The $+1$ in the denominator of ε is just to avoid issues in case $b = 0$. Then we have

$$\sup_{(t,x)\in[0,T']\times\mathbb{R}^d} |\partial_x U_b(t,x)| \leq \varepsilon \|b\|_{L_p^q(T')},$$

$$\|\mathcal{T}^1(b)\|_{L_p^q(T')} \leq \sup_{(t,x)\in[0,T']\times\mathbb{R}^d} |\partial_x U_b(t,x)| \cdot \|b\|_{L_p^q(T')} \leq \varepsilon \|b\|^2_{L_p^q(T')},$$

$$\sup_{(t,x)\in[0,T']\times\mathbb{R}^d} |\partial_x U_{\mathcal{T}^1(b)}(t,x)| \leq \varepsilon \|\mathcal{T}^1(b)\|_{L_p^q(T')} \leq \varepsilon^2 \|b\|^2_{L_p^q(T')},$$

$$\|\mathcal{T}^2(b)\|_{L_p^q(T')} \leq \sup_{(t,x)\in[0,T']\times\mathbb{R}^d} |\partial_x U_{\mathcal{T}^1(b)}(t,x)| \cdot \|b\|_{L_p^q(T')} \leq \varepsilon^2 \|b\|^3_{L_p^q(T')}$$

and by iterating

$$\|\mathcal{T}^k(b)\|_{L_p^q(T')} \leq \varepsilon^k \|b\|^{k+1}_{L_p^q(T')} \leq \frac{1}{4^k(\|b\|_{L_p^q(T')}+1)^k} \|b\|^{k+1}_{L_p^q(T')} \leq \frac{1}{4^k} \|b\|_{L_p^q(T')} \tag{16}$$

which proves (i).

(ii) Applying (15) and (16) yields

$$\sup_{(t,x)\in[0,T']\times\mathbb{R}^d} |\partial_x U_{\mathcal{T}^k(b)}(t,x)| \leq \varepsilon \|\mathcal{T}^k(b)\|_{L_p^q(T')} \leq \frac{\varepsilon}{4^k} \|b\|_{L_p^q(T')} \leq \frac{1}{4^{k+1}}.$$

Therefore, we get

$$\sum_{k=0}^n \sup_{(t,x)\in[0,T']\times\mathbb{R}^d} |\partial_x U_{\mathcal{T}^k(b)}(t,x)| \leq \sum_{k=0}^n \frac{1}{4^{k+1}} = \frac{1}{4} \sum_{k=0}^n \frac{1}{4^k} \leq \frac{1}{4} \sum_{k=0}^\infty \frac{1}{4^k} \leq \frac{1}{2}$$

and so the second inequality is proved.

(iii) We have with (8)

$$\|\partial_x^2 U^{(n)}\|_{L_p^q(T')} \leq \sum_{k=0}^n \|\partial_x^2 U_{\mathcal{T}^k(b)}\|_{L_p^q(T')} \leq \sum_{k=0}^n \|U_{\mathcal{T}^k(b)}\|_{W^{1,2}_{q,p}(T')} \leq \sum_{k=0}^n C \|\mathcal{T}^k(b)\|_{L_p^q(T')}.$$

And using (i) leads to

$$\|\partial_x^2 U^{(n)}\|_{L_p^q(T')} \leq C \sum_{k=0}^n \frac{1}{2^k} \|b\|_{L_p^q(T')} \leq C \|b\|_{L_p^q(T)} \sum_{k=0}^n \frac{1}{2^k} \leq C \sum_{k=0}^\infty \frac{1}{2^k} \leq C$$

for some $C > 0$, independent of n.

(iv) To prove the contraction between X and Y we use the mean-value inequality from Lemma A.7 for $U_{\mathcal{T}^k(b)}$. Thus

$$\begin{aligned}\left|Y_t^{(1,n)} - Y_t^{(2,n)}\right| &= \left|X_t^{(1)} + U^{(n)}(t, X_t^{(1)}) - X_t^{(2)} - U^{(n)}(t, X_t^{(2)})\right| \\ &\leq \left|X_t^{(1)} - X_t^{(2)}\right| + \left|\sum_{k=0}^{n} U_{\mathcal{T}^k(b)}(t, X_t^{(1)}) - U_{\mathcal{T}^k(b)}(t, X_t^{(2)})\right| \\ &\leq \left|X_t^{(1)} - X_t^{(2)}\right| + \sum_{k=0}^{n} \left|U_{\mathcal{T}^k(b)}(t, X_t^{(1)}) - U_{\mathcal{T}^k(b)}(t, X_t^{(2)})\right| \\ &\leq \left|X_t^{(1)} - X_t^{(2)}\right| + \sum_{k=0}^{n} \sup_{(t,x)\in[0,T']\times\mathbb{R}^d} \left|\partial_x U_{\mathcal{T}^k(b)}(t,x)\right| \left|X_t^{(1)} - X_t^{(2)}\right|.\end{aligned}$$

Then (ii) provides

$$\left|Y_t^{(1,n)} - Y_t^{(2,n)}\right| \leq \frac{3}{2}\left|X_t^{(1)} - X_t^{(2)}\right|.$$

On the other hand, we have with the same arguments

$$\begin{aligned}\left|X_t^{(1)} - X_t^{(2)}\right| &= \left|Y_t^{(1,n)} - U^{(n)}(t, X_t^{(1)}) - Y_t^{(2,n)} + U^{(n)}(t, X_t^{(2)})\right| \\ &\leq \left|Y_t^{(1,n)} - Y_t^{(2,n)}\right| + \left|\sum_{k=0}^{n} U_{\mathcal{T}^k(b)}(t, X_t^{(1)}) - U_{\mathcal{T}^k(b)}(t, X_t^{(2)})\right| \\ &\leq \left|Y_t^{(1,n)} - Y_t^{(2,n)}\right| + \frac{1}{2}\left|X_t^{(1)} - X_t^{(2)}\right|,\end{aligned}$$

which is equivalent to

$$\left|X_t^{(1)} - X_t^{(2)}\right| \leq 2\left|Y_t^{(1,n)} - Y_t^{(2,n)}\right|.$$

□

From now on we denote T_0 by T.

4.1. Krylov-type estimates for conditional expectation

To get an exponential estimate on the transformed diffusion, which we are going to state in the next section, we need a Krylov-type estimate on the linear combination of two solutions of SDE (5) as stated in Proposition 4.4. For the proof we have to do some preparation, first Lemma 4.2 which is a version of Lemma 5.1 from [Kry86] for conditional expectation and different integrability in time and space, and second Lemma 4.3, where we prove that the terms on the right-hand side of the inequality are bounded.

Lemma 4.2. *Let the conditions* $(c1)$, $(c3)$, $(c4)$ *of Assumption 2.2 be fulfilled and* $X_t^{(1)}$, $X_t^{(2)}$ *be two solutions of* (5) *such that* (6) *holds. Then, for any nonnegative function*

$f : [0,T] \times \mathbb{R}^d \to \mathbb{R}$ with $\|f\|_{L^r_v(T)} < \infty$, any stopping time γ, $0 \le t_0 \le T$, $R > 0$ and $r, v \ge d+1$ the following holds:

$$\mathbb{1}_{\{t_0 \le \tau_R^\lambda \wedge \gamma\}} \mathbb{E}\left[\int_{t_0}^{T \wedge \tau_R^\lambda \wedge \gamma} \det(a_t^\lambda)^{\frac{1}{d+1}} f(t, X_t^\lambda)\, dt \,\middle|\, \mathcal{F}_{t_0} \right] \le \mathbb{1}_{\{t_0 \le \tau_R^\lambda \wedge \gamma\}} C(T,d,v,r)(\mathbb{B}^2 + \mathbb{A})^{\frac{d}{d+1} - \frac{d}{2v}} \|f\|_{L^r_v(T)} \qquad \mathbb{P}\text{-a. s.}.$$

Here we denote

$$\mathbb{B} := \mathbb{E}\left[\int_{t_0}^{T \wedge \tau_R^\lambda \wedge \gamma} \left| b^\lambda(t, X_t^{(1)}, X_t^{(2)}) \right| dt \,\middle|\, \mathcal{F}_{t_0} \right], \quad \mathbb{A} := \mathbb{E}\left[\int_{t_0}^{T \wedge \tau_R^\lambda \wedge \gamma} \operatorname{tr}(a_t^\lambda)\, dt \,\middle|\, \mathcal{F}_{t_0} \right].$$

Note, that $\mathbb{A}$ and $\mathbb{B}$ depend on $t_0, T, R, \lambda, \gamma$. We refrain from denoting this in indices since it will always be clear what we mean and thus, would be more confusing than helpful.

The proof is structured as follows. First we prove the inequality for nonnegative functions in $\mathcal{C}_0^\infty$ such that $f > 0$ on $[0,T] \times \overline{B_R}$, B_R denotes here the open ball in $\mathbb{R}^d$ around the origin with radius R and $\overline{B_R}$ the closure of it. This will be done by using Lemma A.8, Itô's formula and the martingale property of the stochastic integral. Then, we extend this to nonnegative functions in $\mathcal{C}_0^\infty$. After that we prove that for these functions the inequality holds also for $|f|$. The statement is extended to measurable bounded functions by a monotone class argument and finally also to unbounded measurable functions.

Proof. Note, that all the conditional expectations exist, since we always integrate nonnegative functions. Fix a $\mu > 0$ and take a nonnegative $f \in \mathcal{C}_0^\infty(\mathbb{R}^{d+1})$ with $f > 0$ on $[0,T] \times \overline{B_R}$. Obviously there exist T', R' such that $f = 0$ for $|t| \ge T'$ or $|x| > R'$. Then Lemma A.8 ensures the existence of a nonnegative function φ with bounded weak derivatives $\partial_t \varphi, \partial_x \varphi, \partial_x^2 \varphi$ such that for any symmetric, positive semidefinite $d \times d$ matrix α the following holds:

$$\begin{aligned}
&\partial_t \varphi + \sum_{i=1}^d \sum_{j=1}^d \alpha_{ij} \partial^2_{x_i x_j} \varphi - \mu(1 + \operatorname{tr}(\alpha))\varphi + \det(\alpha)^{\frac{1}{d+1}} f e^{\mu t} \le 0, \\
&|\partial_x \varphi| \le \sqrt{\mu} \varphi, \\
&\varphi(t,x) \le C(d,v) \mu^{\frac{d}{2v} - \frac{d}{d+1}} (T' - t)^{\frac{1}{d+1} - \frac{1}{r}} e^{\mu t} \|f\|_{L^r_v}.
\end{aligned}$$

Define $\psi := e^{-\mu t} \varphi$. Then we have

$$\partial_t \psi + \sum_{i=1}^d \sum_{j=1}^d \alpha_{ij} \partial^2_{x_i x_j} \psi - \mu \operatorname{tr}(\alpha) \psi + \det(\alpha)^{\frac{1}{d+1}} f \le 0, \tag{17}$$

$$|\partial_x \psi| \le \sqrt{\mu} \psi, \tag{18}$$

$$\psi(t,x) \le C(d,v) \mu^{\frac{d}{2v} - \frac{d}{d+1}} (T' - t)^{\frac{1}{d+1} - \frac{1}{r}} \|f\|_{L^r_v}. \tag{19}$$

From [Kry87] Example 6.4.6, we know that $\partial_t\psi, \partial_x\psi, \partial_x^2\psi$ are continuous on $[0,T]\times\overline{B_R}$. Therefore, we may apply Itô's formula and get

$$\begin{aligned}\psi(t,X_t^\lambda)-\psi(0,x^\lambda) &= \int_0^t \partial_t\psi(s,X_s^\lambda)\,ds + \int_0^t \partial_x\psi(s,X_s^\lambda)b^\lambda(s,X_s^{(1)},X_s^{(2)})\,ds \\ &\quad + \int_0^t \partial_x\psi(s,X_s^\lambda)\sigma^\lambda(s,X_s^{(1)},X_s^{(2)})\,dW_s \\ &\quad + \frac{1}{2}\int_0^t \sum_{i=1}^d\sum_{j=1}^d (\sigma^\lambda\sigma^{\lambda^*}(s,X_s^{(1)},X_s^{(2)}))_{ij}\partial^2_{x_ix_j}\psi(s,X_s^\lambda)\,ds\end{aligned}$$

which shows that

$$\begin{aligned}\kappa_t := \psi(t,X_t^\lambda) - \int_0^t &\partial_x\psi(s,X_s^\lambda)b^\lambda(s,X_s^{(1)},X_s^{(2)}) \\ &+\partial_t\psi(s,X_s^\lambda)+\sum_{i=1}^d\sum_{j=1}^d (a_s^\lambda)_{ij}\partial^2_{x_ix_j}\psi(s,X_s^\lambda)\,ds\end{aligned}$$

is a martingale on $[0,T\wedge\tau_R^\lambda\wedge\gamma)$. Then, by applying inequality (17) we have for all $t\in[0,T\wedge\tau_R^\lambda\wedge\gamma)$ on the set $\{t_0\le t\}$

$$\begin{aligned}&\mathbb{E}\left[\int_{t_0}^t \det(a_s^\lambda)^{\frac{1}{d+1}}f(s,X_s^\lambda)\,ds\;\middle|\;\mathcal{F}_{t_0}\right] \\ &\le \mathbb{E}\left[\int_{t_0}^t \mu\,\mathrm{tr}(a_s^\lambda)\psi(s,X_s^\lambda)-\partial_t\psi(s,X_s^\lambda)-\sum_{i=1}^d\sum_{j=1}^d (a_s^\lambda)_{ij}\partial^2_{x_ix_j}\psi(s,X_s^\lambda)\,ds\;\middle|\;\mathcal{F}_{t_0}\right] \\ &= \mathbb{E}\Bigg[\int_{t_0}^t \mu\,\mathrm{tr}(a_s^\lambda)\psi(s,X_s^\lambda)\,ds+\kappa_t-\kappa_{t_0}-\psi(t,X_t^\lambda)+\psi(t_0,X_{t_0}^\lambda) \\ &\qquad + \int_{t_0}^t \partial_x\psi(s,X_s^\lambda)b^\lambda(s,X_s^{(1)},X_s^{(2)})\,ds\;\Bigg|\;\mathcal{F}_{t_0}\Bigg] \qquad \mathbb{P}\text{-a.s.}.\end{aligned}$$

Since ψ is nonnegative and κ_t is a martingale, we obtain

$$\begin{aligned}
&\mathbb{E}\left[\int_{t_0}^{t} \det(a_s^\lambda)^{\frac{1}{d+1}} f(s, X_s^\lambda)\, ds \,\middle|\, \mathcal{F}_{t_0}\right] \\
&\quad \le \mathbb{E}\Bigg[\int_{t_0}^{t} \mu \operatorname{tr}(a_s^\lambda)\psi(s, X_s^\lambda)\, ds + \kappa_t - \kappa_{t_0} + \psi(t_0, X_{t_0}^\lambda) \\
&\qquad\qquad + \int_{t_0}^{t} \partial_x \psi(s, X_s^\lambda) b^\lambda(s, X_s^{(1)}, X_s^{(2)})\, ds \,\Bigg|\, \mathcal{F}_{t_0}\Bigg] \\
&\quad = \mathbb{E}\left[\psi(t_0, X_{t_0}^\lambda) + \int_{t_0}^{t} \mu \operatorname{tr}(a_s^\lambda)\psi(s, X_s^\lambda) + \partial_x \psi(s, X_s^\lambda) b^\lambda(s, X_s^{(1)}, X_s^{(2)})\, ds \,\middle|\, \mathcal{F}_{t_0}\right] \\
&\quad \le \mathbb{E}\left[\psi(t_0, X_{t_0}^\lambda) + \int_{t_0}^{t} \mu \operatorname{tr}(a_s^\lambda)\psi(s, X_s^\lambda) + \left|\partial_x \psi(s, X_s^\lambda)\right| \left|b^\lambda(s, X_s^{(1)}, X_s^{(2)})\right|\, ds \,\middle|\, \mathcal{F}_{t_0}\right].
\end{aligned}$$

Then with (18), we receive that

$$\begin{aligned}
&\mathbb{E}\left[\int_{t_0}^{t} \det(a_s^\lambda)^{\frac{1}{d+1}} f(s, X_s^\lambda)\, ds \,\middle|\, \mathcal{F}_{t_0}\right] \\
&\le \mathbb{E}\left[\psi(t_0, X_{t_0}^\lambda) + \int_{t_0}^{t} \mu \operatorname{tr}(a_s^\lambda)\psi(s, X_s^\lambda) + \sqrt{\mu}\psi(s, X_s^\lambda) \left|b^\lambda(s, X_s^{(1)}, X_s^{(2)})\right|\, ds \,\middle|\, \mathcal{F}_{t_0}\right] \\
&\le \mathbb{E}\left[\psi(t_0, X_{t_0}^\lambda) + \sup_{s\in[t_0,t]} \psi(s, X_s^\lambda) \int_{t_0}^{t} \mu \operatorname{tr}(a_s^\lambda) + \sqrt{\mu} \left|b^\lambda(s, X_s^{(1)}, X_s^{(2)})\right|\, ds \,\middle|\, \mathcal{F}_{t_0}\right] \\
&\le C(d, v, r, T')\|f\|_{L_v^r} \mathbb{E}\left[\mu^{\frac{d}{2v} - \frac{d}{d+1}} \left(1 + \int_{t_0}^{t} \mu \operatorname{tr}(a_s^\lambda) + \sqrt{\mu} \left|b^\lambda(s, X_s^{(1)}, X_s^{(2)})\right|\, ds\right) \,\middle|\, \mathcal{F}_{t_0}\right],
\end{aligned}$$

where the last inequality follows with (19). This inequality is independent of ψ and holds for all $\mu > 0$, therefore it is also true for

$$\mu := \mathbb{1}_{\{0 \le \mathbb{A} < \mathbb{B}^2\}} \mathbb{B}^{-2} + \mathbb{1}_{\{\mathbb{A} > 0, \mathbb{A} \ge \mathbb{B}^2\}} \mathbb{A}^{-1} + \mathbb{1}_{\{\mathbb{A} = \mathbb{B} = 0\}} c, \qquad c > 0.$$

By Lemma A.11 $\mathbb{A}$ and $\mathbb{B}$ are $\mathbb{P}$-almost surely finite, which prevents us from technical issues, e.g. dividing by infinity.
Since all the indicator functions and $\mathbb{A}$, $\mathbb{B}$ are measurable with respect to $\mathcal{F}_{t_0}$, we have

for the conditional expectation

$$\mathbb{E}\left[\mu^{\frac{d}{2v}-\frac{d}{d+1}}\left(1+\sqrt{\mu}\int\limits_{t_0}^{t}\left|b^\lambda(s,X_s^{(1)},X_s^{(2)})\right|\,ds+\mu\int\limits_{t_0}^{t}\operatorname{tr}(a_s^\lambda)\,ds\right)\,\middle|\,\mathcal{F}_{t_0}\right]$$

$$\begin{aligned}
&=\mathbb{E}\Bigg[\mathbb{1}_{\{0\le\mathbb{A}<\mathbb{B}^2\}}\mathbb{B}^{\frac{2d}{d+1}-\frac{d}{v}}\Bigg(1+\mathbb{B}^{-1}\int\limits_{t_0}^{t}\left|b^\lambda(s,X_s^{(1)},X_s^{(2)})\right|\,ds\\
&\qquad\qquad\qquad+\mathbb{B}^{-2}\int\limits_{t_0}^{t}\operatorname{tr}(a_s^\lambda)\,ds\Bigg)\,\Bigg|\,\mathcal{F}_{t_0}\Bigg]\\
&\quad+\mathbb{E}\Bigg[\mathbb{1}_{\{\mathbb{A}>0,\mathbb{A}\ge\mathbb{B}^2\}}\mathbb{A}^{\frac{d}{d+1}-\frac{d}{2v}}\Bigg(1+\mathbb{A}^{-\frac{1}{2}}\int\limits_{t_0}^{t}\left|b^\lambda(s,X_s^{(1)},X_s^{(2)})\right|\,ds\\
&\qquad\qquad\qquad+\mathbb{A}^{-1}\int\limits_{t_0}^{t}\operatorname{tr}(a_s^\lambda)\,ds\Bigg)\,\Bigg|\,\mathcal{F}_{t_0}\Bigg]\\
&\quad+\mathbb{E}\Bigg[\mathbb{1}_{\{\mathbb{A}=\mathbb{B}=0\}}c^{\frac{d}{2v}-\frac{d}{d+1}}\Bigg(1+\sqrt{c}\int\limits_{t_0}^{t}\left|b^\lambda(s,X_s^{(1)},X_s^{(2)})\right|\,ds\\
&\qquad\qquad\qquad+c\int\limits_{t_0}^{t}\operatorname{tr}(a_s^\lambda)\,ds\Bigg)\,\Bigg|\,\mathcal{F}_{t_0}\Bigg]\\
&=\mathbb{1}_{\{0\le\mathbb{A}<\mathbb{B}^2\}}\mathbb{B}^{\frac{2d}{d+1}-\frac{d}{v}}\Bigg(1+\mathbb{B}^{-1}\mathbb{E}\Bigg[\int\limits_{t_0}^{t}\left|b^\lambda(s,X_s^{(1)},X_s^{(2)})\right|\,ds\,\Bigg|\,\mathcal{F}_{t_0}\Bigg]\\
&\qquad\qquad\qquad+\mathbb{B}^{-2}\mathbb{E}\Bigg[\int\limits_{t_0}^{t}\operatorname{tr}(a_s^\lambda)\,ds\,\Bigg|\,\mathcal{F}_{t_0}\Bigg]\Bigg)\\
&\quad+\mathbb{1}_{\{\mathbb{A}>0,\mathbb{A}\ge\mathbb{B}^2\}}\mathbb{A}^{\frac{d}{d+1}-\frac{d}{2v}}\Bigg(1+\mathbb{A}^{-\frac{1}{2}}\mathbb{E}\Bigg[\int\limits_{t_0}^{t}\left|b^\lambda(s,X_s^{(1)},X_s^{(2)})\right|\,ds\,\Bigg|\,\mathcal{F}_{t_0}\Bigg]\\
&\qquad\qquad\qquad+\mathbb{A}^{-1}\mathbb{E}\Bigg[\int\limits_{t_0}^{t}\operatorname{tr}(a_s^\lambda)\,ds\,\Bigg|\,\mathcal{F}_{t_0}\Bigg]\Bigg)\\
&\quad+\mathbb{1}_{\{\mathbb{A}=\mathbb{B}=0\}}c^{\frac{d}{2v}-\frac{d}{d+1}}\Bigg(1+\sqrt{c}\,\mathbb{E}\Bigg[\int\limits_{t_0}^{t}\left|b^\lambda(s,X_s^{(1)},X_s^{(2)})\right|\,ds\,\Bigg|\,\mathcal{F}_{t_0}\Bigg]\\
&\qquad\qquad\qquad+c\,\mathbb{E}\Bigg[\int\limits_{t_0}^{t}\operatorname{tr}(a_s^\lambda)\,ds\,\Bigg|\,\mathcal{F}_{t_0}\Bigg]\Bigg).
\end{aligned}$$

This leads to

$$\begin{aligned}
&\mathbb{E}\left[\mu^{\frac{d}{2v}-\frac{d}{d+1}}\left(1+\sqrt{\mu}\int_{t_0}^{t}\left|b^\lambda(s,X_s^{(1)},X_s^{(2)})\right|\,ds+\mu\int_{t_0}^{t}\operatorname{tr}(a_s^\lambda)\,ds\right)\,\middle|\,\mathcal{F}_{t_0}\right]\\
&\quad\le \mathbb{1}_{\{0\le\mathbb{A}<\mathbb{B}^2\}}\mathbb{B}^{\frac{2d}{d+1}-\frac{d}{v}}\left(2+\mathbb{B}^{-2}\mathbb{A}\right)+\mathbb{1}_{\{\mathbb{A}>0,\mathbb{A}\ge\mathbb{B}^2\}}\mathbb{A}^{\frac{d}{d+1}-\frac{d}{2v}}\left(2+\mathbb{A}^{-\frac{1}{2}}\mathbb{B}\right)\\
&\qquad+\mathbb{1}_{\{\mathbb{A}=\mathbb{B}=0\}}c^{\frac{d}{2v}-\frac{d}{d+1}}\left(1+\sqrt{c}\mathbb{B}+c\mathbb{A}\right)\\
&\quad\le \mathbb{1}_{\{0\le\mathbb{A}<\mathbb{B}^2\}}3\mathbb{B}^{\frac{2d}{d+1}-\frac{d}{v}}+\mathbb{1}_{\{\mathbb{A}>0,\mathbb{A}\ge\mathbb{B}^2\}}3\mathbb{A}^{\frac{d}{d+1}-\frac{d}{2v}}+\mathbb{1}_{\{\mathbb{A}=\mathbb{B}=0\}}c^{\frac{d}{2v}-\frac{d}{d+1}}\\
&\quad\le \mathbb{1}_{\{\mathbb{A}>0\text{ or }\mathbb{B}>0\}}3(\mathbb{B}^2+\mathbb{A})^{\frac{d}{d+1}-\frac{d}{2v}}+\mathbb{1}_{\{\mathbb{A}=\mathbb{B}=0\}}c^{\frac{d}{2v}-\frac{d}{d+1}}\\
&\quad\overset{c\to\infty}{\to} 3(\mathbb{B}^2+\mathbb{A})^{\frac{d}{d+1}-\frac{d}{2v}}.
\end{aligned}$$

So, we proved for $t\in[0,T\wedge\tau_R^\lambda\wedge\gamma)$

$$\begin{aligned}
&\mathbb{1}_{\{t_0\le t\}}\mathbb{E}\left[\int_{t_0}^{t}\det(a_s^\lambda)^{\frac{1}{d+1}}f(s,X_s^\lambda)\,ds\,\middle|\,\mathcal{F}_{t_0}\right]\\
&\qquad\qquad\le \mathbb{1}_{\{t_0\le t\}}C(d,v,r,T')(\mathbb{B}^2+\mathbb{A})^{\frac{d}{d+1}-\frac{d}{2v}}\|f\|_{L_v^r}\qquad\mathbb{P}\text{-a. s.}.
\end{aligned}$$

With Fatou's Lemma for conditional expectation, we get for $0\le t_0\le T$

$$\begin{aligned}
&\mathbb{1}_{\{t_0\le\tau_R^\lambda\wedge\gamma\}}\mathbb{E}\left[\int_{t_0}^{T\wedge\tau_R^\lambda\wedge\gamma}\det(a_s^\lambda)^{\frac{1}{d+1}}f(s,X_s^\lambda)\,ds\,\middle|\,\mathcal{F}_{t_0}\right]\\
&\quad=\mathbb{E}\left[\mathbb{1}_{\{t_0\le T\wedge\tau_R^\lambda\wedge\gamma\}}\int_{t_0}^{T\wedge\tau_R^\lambda\wedge\gamma}\det(a_s^\lambda)^{\frac{1}{d+1}}f(s,X_s^\lambda)\,ds\,\middle|\,\mathcal{F}_{t_0}\right]\\
&\quad=\mathbb{E}\left[\liminf_{n\to\infty}\mathbb{1}_{\{t_0\le T\wedge\tau_R^\lambda\wedge\gamma-\frac{1}{n}\}}\int_{t_0}^{T\wedge\tau_R^\lambda\wedge\gamma-\frac{1}{n}}\det(a_s^\lambda)^{\frac{1}{d+1}}f(s,X_s^\lambda)\,ds\,\middle|\,\mathcal{F}_{t_0}\right]\\
&\quad\le\liminf_{n\to\infty}\mathbb{E}\left[\mathbb{1}_{\{t_0\le T\wedge\tau_R^\lambda\wedge\gamma-\frac{1}{n}\}}\int_{t_0}^{T\wedge\tau_R^\lambda\wedge\gamma-\frac{1}{n}}\det(a_s^\lambda)^{\frac{1}{d+1}}f(s,X_s^\lambda)\,ds\,\middle|\,\mathcal{F}_{t_0}\right]\\
&\quad=\liminf_{n\to\infty}\mathbb{1}_{\{t_0\le T\wedge\tau_R^\lambda\wedge\gamma-\frac{1}{n}\}}\mathbb{E}\left[\int_{t_0}^{T\wedge\tau_R^\lambda\wedge\gamma-\frac{1}{n}}\det(a_s^\lambda)^{\frac{1}{d+1}}f(s,X_s^\lambda)\,ds\,\middle|\,\mathcal{F}_{t_0}\right]\\
&\quad\le\liminf_{n\to\infty}\mathbb{1}_{\{t_0\le T\wedge\tau_R^\lambda\wedge\gamma-\frac{1}{n}\}}C(d,v,r,T')(\mathbb{B}^2+\mathbb{A})^{\frac{d}{d+1}-\frac{d}{2v}}\|f\|_{L_v^r}\\
&\quad=\mathbb{1}_{\{t_0\le\tau_R^\lambda\wedge\gamma\}}C(d,v,r,T')(\mathbb{B}^2+\mathbb{A})^{\frac{d}{d+1}-\frac{d}{2v}}\|f\|_{L_v^r}\qquad\mathbb{P}\text{-a. s.}
\end{aligned}$$

for all nonnegative $f \in \mathcal{C}_0^\infty(\mathbb{R}^{d+1})$ with $f > 0$ on $[0,T] \times \overline{B_R}$. Now, let $f \in \mathcal{C}_0^\infty(\mathbb{R}^{d+1})$ with $f \geq 0$. Take a smooth function $\chi : \mathbb{R}^{d+1} \to [0,1]$ with

$$\chi > 0 \quad \text{on} \quad [0,T] \times \overline{B_R},$$

for example χ from Lemma A.9. Then we have on the set $\{t_0 \leq \tau_R^\lambda \wedge \gamma\}$

$$\begin{aligned}
&\mathbb{E}\left[\int\limits_{t_0}^{T\wedge\tau_R^\lambda\wedge\gamma} \det(a_t^\lambda)^{\frac{1}{d+1}} f(t, X_t^\lambda)\, dt \,\middle|\, \mathcal{F}_{t_0} \right] \\
&= \mathbb{E}\left[\int\limits_{t_0}^{T\wedge\tau_R^\lambda\wedge\gamma} \det(a_t^\lambda)^{\frac{1}{d+1}} \lim_{\varepsilon\searrow 0} (f + \varepsilon\chi)(t, X_t^\lambda)\, dt \,\middle|\, \mathcal{F}_{t_0} \right] \\
&= \lim_{\varepsilon\searrow 0} \mathbb{E}\left[\int\limits_{t_0}^{T\wedge\tau_R^\lambda\wedge\gamma} \det(a_t^\lambda)^{\frac{1}{d+1}} (f + \varepsilon\chi)(t, X_t^\lambda)\, dt \,\middle|\, \mathcal{F}_{t_0} \right]
\end{aligned}$$

by dominated convergence. As $f + \varepsilon\chi$ is strictly positive on $[0,T] \times \overline{B_R}$, we have, for a suitable $T' > 0$

$$\begin{aligned}
\mathbb{E}\left[\int\limits_{t_0}^{T\wedge\tau_R^\lambda\wedge\gamma} \det(a_t^\lambda)^{\frac{1}{d+1}} f(t, X_t^\lambda)\, dt \,\middle|\, \mathcal{F}_{t_0} \right] &\leq \lim_{\varepsilon\searrow 0} C(d, v, r, T')(\mathbb{B}^2 + \mathbb{A})^{\frac{d}{d+1} - \frac{d}{2v}} \|f + \varepsilon\chi\|_{L_v^r} \\
&= C(d, v, r, T')(\mathbb{B}^2 + \mathbb{A})^{\frac{d}{d+1} - \frac{d}{2v}} \|f\|_{L_v^r} \qquad \mathbb{P}\text{-a. s..}
\end{aligned}$$

The next step is to get rid of the dependence on T'. To this end, consider the smooth function

$$g(t) := \begin{cases} c \exp\left(-\frac{1}{1-|2t|^2} \right) & \text{if } |t| < \frac{1}{2}, \\ 0 & \text{else,} \end{cases}$$

where c is chosen such that

$$\int\limits_{\mathbb{R}} g(t) dt = 1.$$

Then we have for the convolution

$$\left(\mathbb{1}_{[-\frac{1}{2}, T+\frac{1}{2}]} * g \right)(t) = \int\limits_{\mathbb{R}} \mathbb{1}_{[-\frac{1}{2}, T+\frac{1}{2}]}(t-s) g(s)\, ds = \int\limits_{-\frac{1}{2}}^{\frac{1}{2}} \mathbb{1}_{[-\frac{1}{2}, T+\frac{1}{2}]}(t-s) g(s)\, ds$$

which is 1 for $t \in [0,T]$ and 0 for $t \notin [-1, T+1]$. Since $(\mathbb{1}_{[-\frac{1}{2}, T+\frac{1}{2}]} * g) \cdot f$ is smooth and equal to f on $[t_0, T]$, we have on the set $\{t_0 \leq \tau_R^\lambda \wedge \gamma\}$

$$\begin{aligned}
&\mathbb{E}\left[\int\limits_{t_0}^{T\wedge\tau_R^\lambda\wedge\gamma} \det(a_t^\lambda)^{\frac{1}{d+1}} f(t, X_t^\lambda)\, dt \,\middle|\, \mathcal{F}_{t_0}\right] \\
&= \mathbb{E}\left[\int\limits_{t_0}^{T\wedge\tau_R^\lambda\wedge\gamma} \det(a_t^\lambda)^{\frac{1}{d+1}} \left(\mathbb{1}_{[-\frac{1}{2}, T+\frac{1}{2}]} * g\right)(t) f(t, X_t^\lambda)\, dt \,\middle|\, \mathcal{F}_{t_0}\right] \\
&\leq C(d, v, r, T+1)(\mathbb{B}^2 + \mathbb{A})^{\frac{d}{d+1} - \frac{d}{2v}} \left(\int\limits_{\mathbb{R}_+} \left(\int\limits_{\mathbb{R}^d} \left|\left(\mathbb{1}_{[-\frac{1}{2}, T+\frac{1}{2}]} * g\right)(t) f(t, x)\right|^v dx\right)^{\frac{r}{v}} dt\right)^{\frac{1}{r}} \\
&\leq C(d, v, r, T)(\mathbb{B}^2 + \mathbb{A})^{\frac{d}{d+1} - \frac{d}{2v}} \|f\|_{L_v^r} \qquad \mathbb{P}\text{-a. s..}
\end{aligned}$$

Now, let $f \in \mathcal{C}_0^\infty(\mathbb{R}^{d+1})$. Since $|f|$ is continuous and compactly supported, there exists a sequence $(f_n)_n$ of nonnegative functions in $\mathcal{C}_0^\infty(\mathbb{R}^{d+1})$ which converges uniformly to $|f|$ (let ψ be a mollifier on $\mathbb{R}^{d+1}$ and take $f_n := \psi_{1/n} * |f|$, see Appendix for the definition of mollifier and $\psi_{1/n}$). Therefore, on the set $\{t_0 \leq \tau_R^\lambda \wedge \gamma\}$ we have

$$\mathbb{E}\left[\int\limits_{t_0}^{T\wedge\tau_R^\lambda\wedge\gamma} \det(a_t^\lambda)^{\frac{1}{d+1}} |f(t, X_t^\lambda)|\, dt \,\middle|\, \mathcal{F}_{t_0}\right] = \lim_{n\to\infty} \mathbb{E}\left[\int\limits_{t_0}^{T\wedge\tau_R^\lambda\wedge\gamma} \det(a_t^\lambda)^{\frac{1}{d+1}} f_n(t, X_t^\lambda)\, dt \,\middle|\, \mathcal{F}_{t_0}\right]$$

and since the inequality is true for nonnegative functions in $\mathcal{C}_0^\infty(\mathbb{R}^{d+1})$

$$\begin{aligned}
\mathbb{E}\left[\int\limits_{t_0}^{T\wedge\tau_R^\lambda\wedge\gamma} \det(a_t^\lambda)^{\frac{1}{d+1}} |f(t, X_t^\lambda)|\, dt \,\middle|\, \mathcal{F}_{t_0}\right] &\leq \lim_{n\to\infty} C(d, v, r, T)(\mathbb{B}^2 + \mathbb{A})^{\frac{d}{d+1} - \frac{d}{2v}} \|f_n\|_{L_v^r} \\
&= C(d, v, r, T)(\mathbb{B}^2 + \mathbb{A})^{\frac{d}{d+1} - \frac{d}{2v}} \|f\|_{L_v^r} \qquad \mathbb{P}\text{-a. s..}
\end{aligned}$$

To prove that the inequality is also valid for bounded measurable functions, define

$$\begin{aligned}
\mathcal{X} := \Bigg\{ f : \mathbb{R}^{d+1} \to \mathbb{R} \,\Bigg|\, & f \text{ is measurable, bounded and fulfills } \mathbb{P}\text{-almost surely} \\
& \mathbb{1}_{\{t_0 \leq \tau_R^\lambda \wedge \gamma\}} \mathbb{E}\left[\int\limits_{t_0}^{T\wedge\tau_R^\lambda\wedge\gamma} \det(a_t^\lambda)^{\frac{1}{d+1}} |f(t, X_t^\lambda)|\, dt \,\middle|\, \mathcal{F}_{t_0}\right] \\
& \qquad \leq \mathbb{1}_{\{t_0 \leq \tau_R^\lambda \wedge \gamma\}} C(d, v, r, T)(\mathbb{B}^2 + \mathbb{A})^{\frac{d}{d+1} - \frac{d}{2v}} \|f\|_{L_v^r} \Bigg\}.
\end{aligned}$$

Note, that the left-hand side exists, since we integrate nonnegative functions. The right-hand side of the inequality maybe be infinite, which is feasible since the inequality is then trivially fulfilled. Let $0 \leq f_1 \leq f_2 \leq ... \leq f_n \leq ...$ in $\mathcal{X}$ with $f_n \to f$ pointwise and f bounded, then the inequality holds for f, because with monotone convergence we obtain

$$\begin{aligned}
&\mathbb{E}\left[\int\limits_{t_0}^{T\wedge\tau_R^\lambda\wedge\gamma} \det(a_t^\lambda)^{\frac{1}{d+1}} |f(t, X_t^\lambda)| \, dt \,\middle|\, \mathcal{F}_{t_0} \right] \\
&= \mathbb{E}\left[\int\limits_{t_0}^{T\wedge\tau_R^\lambda\wedge\gamma} \det(a_t^\lambda)^{\frac{1}{d+1}} \lim_{n\to\infty} |f_n(t, X_t^\lambda)| \, dt \,\middle|\, \mathcal{F}_{t_0} \right] \\
&= \lim_{n\to\infty} \mathbb{E}\left[\int\limits_{t_0}^{T\wedge\tau_R^\lambda\wedge\gamma} \det(a_t^\lambda)^{\frac{1}{d+1}} |f_n(t, X_t^\lambda)| \, dt \,\middle|\, \mathcal{F}_{t_0} \right] \\
&\leq \lim_{n\to\infty} C(d, v, r, T)(\mathbb{B}^2 + \mathbb{A})^{\frac{d}{d+1} - \frac{d}{2v}} \|f_n\|_{L_v^r} \\
&= C(d, v, r, T)(\mathbb{B}^2 + \mathbb{A})^{\frac{d}{d+1} - \frac{d}{2v}} \|\lim_{n\to\infty} f_n\|_{L_v^r} \\
&= C(d, v, r, T)(\mathbb{B}^2 + \mathbb{A})^{\frac{d}{d+1} - \frac{d}{2v}} \|f\|_{L_v^r}
\end{aligned}$$

$\mathbb{P}$-almost surely on the set $\{t_0 \leq \tau_R^\lambda \wedge \gamma\}$. Since f is again measurable, we have $f \in \mathcal{X}$. Therefore $\mathcal{X}$ is closed under bounded monotone convergence. And by similar means it can be also shown that $\mathcal{X}$ is closed under uniform convergence. Since $\mathcal{C}_0^\infty(\mathbb{R}^{d+1})$ is an algebra and there exists a sequence f_n in $\mathcal{C}_0^\infty(\mathbb{R}^{d+1})$ such that $f_n \nearrow 1$, the monotone class theorem is applicable in the version of [Del78] (22.2) and this yields that $\mathcal{X}$ contains all measurable bounded functions. Now, let f be a nonnegative measurable function with $\|f\|_{L_v^r(T)} < \infty$. Since $\mathbb{1}_{[0,T]}(f \wedge n) \in \mathcal{X}$ we obtain on the set $\{t_0 \leq \tau_R^\lambda \wedge \gamma\}$ with monotone convergence

$$\begin{aligned}
&\mathbb{E}\left[\int\limits_{t_0}^{T\wedge\tau_R^\lambda\wedge\gamma} \det(a_t^\lambda)^{\frac{1}{d+1}} f(t, X_t^\lambda) \, dt \,\middle|\, \mathcal{F}_{t_0} \right] \\
&= \lim_{n\to\infty} \mathbb{E}\left[\int\limits_{t_0}^{T\wedge\tau_R^\lambda\wedge\gamma} \det(a_t^\lambda)^{\frac{1}{d+1}} \mathbb{1}_{[0,T]}(t)(f \wedge n)(t, X_t^\lambda) dt \,\middle|\, \mathcal{F}_{t_0} \right] \\
&\leq \lim_{n\to\infty} C(d, v, r, T)(\mathbb{B}^2 + \mathbb{A})^{\frac{d}{d+1} - \frac{d}{2v}} \|f \wedge n\|_{L_v^r(T)} \\
&= C(d, v, r, T)(\mathbb{B}^2 + \mathbb{A})^{\frac{d}{d+1} - \frac{d}{2v}} \|f\|_{L_v^r(T)} \qquad \mathbb{P}\text{-a. s.}.
\end{aligned}$$

□

Lemma 4.3. *Let (c1), (c3), (c4) of Assumption 2.2 be fulfilled and $X_t^{(1)}$, $X_t^{(2)}$ be two solutions to* (5) *such that condition* (6) *holds. Then we have for all* $0 \le t_0 \le T$, $\lambda \in [0,1]$,

$$\mathbb{E}\left[\int_{t_0}^{T} \operatorname{tr}(a_t^\lambda)\, dt \,\middle|\, \mathcal{F}_{t_0} \right] \le C(T, \tilde{c}_\sigma)$$

and

$$\mathbb{E}\left[\int_{t_0}^{T} \left| b^\lambda(t, X_t^{(1)}, X_t^{(2)}) \right| dt \,\middle|\, \mathcal{F}_{t_0} \right] \le C(d, p, q, T, c_\sigma, \tilde{c}_\sigma, \|b\|_{L_p^q(T)})$$

$\mathbb{P}$*-almost surely.*

The idea of proving this, especially the second estimate is taken from [GM01], Proof of Corollary 3.2.

Proof. Using (c4) we may estimate the trace of a_t^λ as in (41):

$$\operatorname{tr}(a_t^\lambda) \le 2\tilde{c}_\sigma^2.$$

Then, monotonicity of the conditional expectation results in

$$\mathbb{E}\left[\int_{t_0}^{T} \operatorname{tr}(a_t^\lambda)\, dt \,\middle|\, \mathcal{F}_{t_0} \right] \le 2\tilde{c}_\sigma^2 T.$$

To prove that the second conditional expectation is $\mathbb{P}$-almost surely finite, we will use Lemma 4.2 for $X_t^{(1)}$ and $X_t^{(2)}$. Note that all the eigenvalues of $\sigma\sigma^*$ are bounded from below by c_σ because of (c3). Since a symmetric matrix has only real eigenvalues and the determinant is the product of them, we have in case $\lambda = 1$

$$\det(a_t^1) = \frac{1}{2^d} \det(\sigma\sigma^*(t, X_t^{(1)})) \ge \frac{1}{2^d} c_\sigma^d.$$

And the same holds for $\det(a_t^0)$. Define

$$\gamma_n := \inf \left\{ t \ge t_0 \;:\; \mathbb{E}\left[\int_{t_0}^{t} \left| b(s, X_s^{(1)}) \right| ds \,\middle|\, \mathcal{F}_{t_0} \right] > \; n \right\}$$

and

$$\mathbb{B}^{(n)} := \mathbb{E}\left[\int_{t_0}^{T\wedge\tau_R^1\wedge\gamma_n} \left| b(t, X_t^{(1)}) \right| dt \,\middle|\, \mathcal{F}_{t_0} \right],$$

$$\mathbb{A}^{(n)} := \mathbb{E}\left[\int_{t_0}^{T\wedge\tau_R^1\wedge\gamma_n} \operatorname{tr}(a_t^1)\, dt \,\middle|\, \mathcal{F}_{t_0} \right].$$

With Jensen's inequality for the conditional expectation and the Lebesgue measure on $[0,T]$, we receive on the set $\{t_0 \le \tau_R^1 \wedge \gamma_n\}$

$$\begin{aligned}
\left(\mathbb{B}^{(n)}\right)^2 &\le \mathbb{E}\left[\left(\int_{t_0}^{T\wedge\tau_R^1\wedge\gamma_n} \left|b(t,X_t^{(1)})\right| dt\right)^2 \,\middle|\, \mathcal{F}_{t_0}\right] \\
&\le T\mathbb{E}\left[\int_{t_0}^{T\wedge\tau_R^1\wedge\gamma_n} \left|b(t,X_t^{(1)})\right|^2 dt \,\middle|\, \mathcal{F}_{t_0}\right] \\
&= T\mathbb{E}\left[\int_{t_0}^{T\wedge\tau_R^1\wedge\gamma_n} \left(\frac{\det(a_t^1)}{\det(a_t^1)}\right)^{\frac{1}{d+1}} \left|b(t,X_t^{(1)})\right|^2 dt \,\middle|\, \mathcal{F}_{t_0}\right] \\
&\le \left(\frac{2}{c_\sigma}\right)^{\frac{d}{d+1}} T\mathbb{E}\left[\int_{t_0}^{T\wedge\tau_R^1\wedge\gamma_n} \det(a_t^1)^{\frac{1}{d+1}} \left|b(t,X_t^{(1)})\right|^2 dt \,\middle|\, \mathcal{F}_{t_0}\right].
\end{aligned}$$

Applying the inequality from Lemma 4.2 with $v = \frac{p}{2}$, $r = \frac{q}{2}$, provides

$$\begin{aligned}
\left(\mathbb{B}^{(n)}\right)^2 &\le \left(\frac{2}{c_\sigma}\right)^{\frac{d}{d+1}} TC(d,p,q,T)\left(\left(\mathbb{B}^{(n)}\right)^2 + \mathbb{A}^{(n)}\right)^{\frac{d}{d+1}-\frac{d}{p}} \|b\|^2_{L_p^q(T)} \\
&\le \left(\frac{2}{c_\sigma}\right)^{\frac{d}{d+1}} TC(d,p,q,T)\left(\left(\mathbb{B}^{(n)}\right)^{\frac{2d}{d+1}-\frac{2d}{p}} + (2\tilde{c}_\sigma^2 T)^{\frac{d}{d+1}-\frac{d}{p}}\right) \|b\|^2_{L_p^q(T)}
\end{aligned}$$

$\mathbb{P}$-almost surely. With Young's inequality we have for $\varepsilon > 0$ and $z := \frac{d}{d+1} - \frac{d}{p} < 1$,

$$\begin{aligned}
\left(\mathbb{B}^{(n)}\right)^{2z} &= \frac{1}{\varepsilon}\cdot\varepsilon\left(\mathbb{B}^{(n)}\right)^{2z} \\
&\le (1-z)\varepsilon^{-\frac{1}{1-z}} + z\varepsilon^{\frac{1}{z}}\left(\mathbb{B}^{(n)}\right)^2 \\
&\le \varepsilon^{\frac{1}{z-1}} + \varepsilon^{\frac{1}{z}}\left(\mathbb{B}^{(n)}\right)^2.
\end{aligned}$$

Let ε be small enough such that

$$\left(\frac{2}{c_\sigma}\right)^{\frac{d}{d+1}} TC(d,p,q,T)\varepsilon^{\frac{1}{z}}\|b\|^2_{L_p^q(T)} < 1.$$

Note, that we may choose ε independent of ω, n and R. Then we get

$$\left(\mathbb{B}^{(n)}\right)^2 \le \left(\frac{2}{c_\sigma}\right)^{\frac{d}{d+1}} TC(d,p,q,T)\left((2\tilde{c}_\sigma^2 T)^{\frac{d}{d+1}-\frac{d}{p}} + \varepsilon^{\frac{1}{z-1}} + \varepsilon^{\frac{1}{z}}\left(\mathbb{B}^{(n)}\right)^2\right)\|b\|^2_{L_p^q(T)} \qquad \mathbb{P}\text{-a.s.}$$

which is equivalent to

$$\left(\mathbb{B}^{(n)}\right)^2 \le \frac{\left(\frac{2}{c_\sigma}\right)^{\frac{d}{d+1}} TC(d,p,q,T)\left((2\tilde{c}_\sigma^2 T)^{\frac{d}{d+1}-\frac{d}{p}} + \varepsilon^{\frac{1}{z-1}}\right)\|b\|^2_{L_p^q(T)}}{1 - \left(\frac{2}{c_\sigma}\right)^{\frac{d}{d+1}} TC(d,p,q,T)\varepsilon^{\frac{1}{z}}\|b\|^2_{L_p^q(T)}} \qquad \mathbb{P}\text{-a.s.}$$

on the set $\{t_0 \leq \tau_R^1 \wedge \gamma_n\}$, which is finite and independent of n and ω. If we take the limit $n \to \infty$ we obtain that

$$\mathbb{E}\left[\int\limits_{t_0}^{T\wedge\tau_R^1} |b(t, X_t^{(1)})|\, dt \,\middle|\, \mathcal{F}_{t_0}\right] \leq C(d, p, q, T, c_\sigma, \tilde{c}_\sigma, \|b\|_{L_p^q(T)}) \qquad \mathbb{P}\text{-a. s.}$$

on the set $\{t_0 \leq \tau_R^1\}$. Analogously, we can prove that the same holds for $X_t^{(2)}$. Furthermore, the bound is also independent of R. If we take the limit $R \to \infty$ we get

$$\mathbb{E}\left[\int\limits_{t_0}^{T} \left|b(t, X_t^{(i)})\right|\, dt \,\middle|\, \mathcal{F}_{t_0}\right] < C(d, p, q, T, c_\sigma, \tilde{c}_\sigma, \|b\|_{L_p^q(T)}) \qquad \mathbb{P}\text{-a. s..}$$

Therefore, we obtain

$$\begin{aligned}
&\mathbb{E}\left[\int\limits_{t_0}^{T} \left|\lambda b(t, X_t^{(1)}) + (1-\lambda) b(t, X_t^{(2)})\right|\, dt \,\middle|\, \mathcal{F}_{t_0}\right] \\
&\leq \lambda \mathbb{E}\left[\int\limits_{t_0}^{T} \left|b(t, X_t^{(1)})\right|\, dt \,\middle|\, \mathcal{F}_{t_0}\right] + (1-\lambda)\mathbb{E}\left[\int\limits_{t_0}^{T} \left|b(t, X_t^{(2)})\right|\, dt \,\middle|\, \mathcal{F}_{t_0}\right] \\
&\leq C(d, p, q, T, c_\sigma, \tilde{c}_\sigma, \|b\|_{L_p^q(T)})
\end{aligned}$$

$\mathbb{P}$-almost surely, for every $\lambda \in [0, 1]$. □

Proposition 4.4. *Let* (c1)-(c4) *of Assumptions 2.2 be fulfilled and* $X_t^{(1)}$, $X_t^{(2)}$ *be two solutions to* (5) *such that* (6) *holds. Then for arbitrary* $R > 0$ *there exists an* $\varepsilon > 0$ *such that for every nonnegative measurable function* $f : [0, T] \times \mathbb{R}^d \to \mathbb{R}$ *with* $\|f\|_{L_v^r(T)} < \infty$, $r, v \geq d + 1$, *and every* $0 \leq t_0 \leq T$, $\lambda \in [0, 1]$ *we have on the set* $\{t_0 \leq \tau_R \wedge \tau_\varepsilon\}$

$$\mathbb{E}\left[\int\limits_{t_0}^{T\wedge\tau_R\wedge\tau_\varepsilon} f(t, X_t^\lambda)\, dt \,\middle|\, \mathcal{F}_{t_0}\right] \leq C(d, p, v, r, T, \|b\|_{L_p^q(T)}, c_\sigma, \tilde{c}_\sigma)\|f\|_{L_v^r(T)} \quad \mathbb{P}\textit{-almost surely,}$$

where

$$\tau_\varepsilon := \inf\left\{t \geq 0 \;:\; |X_t^{(1)} - X_t^{(2)}| > \varepsilon\right\}.$$

Proof. Since σ is uniformly continuous on $[0, T] \times \overline{B_R}$ there exists an $\varepsilon > 0$ such that

$$|\sigma(t, x) - \sigma(s, y)| < \frac{c_\sigma}{4\tilde{c}_\sigma} \quad \forall\ (t, x), (s, y) \in [0, T] \times \overline{B_R} \text{ with } |(t, x) - (s, y)| \leq \varepsilon.$$

That implies for all $\xi \in \mathbb{R}^d$, $0 \le t \le T \wedge \tau_R \wedge \tau_\varepsilon$

$$\begin{aligned}
&\left|\left\langle \left(\sigma(t, X_t^{(1)}) - \sigma(t, X_t^{(2)})\right) \sigma^*(t, X_t^{(2)})\xi, \xi \right\rangle\right| \\
&\quad \le \left|\left(\sigma(t, X_t^{(1)}) - \sigma(t, X_t^{(2)})\right) \sigma^*(t, X_t^{(2)})\xi\right| |\xi| \\
&\quad \le \left|\sigma(t, X_t^{(1)}) - \sigma(t, X_t^{(2)})\right| \left|\sigma^*(t, X_t^{(2)})\right| |\xi|^2 \\
&\quad \le \frac{c_\sigma}{4\tilde{c}_\sigma} \tilde{c}_\sigma |\xi|^2 \\
&\quad = \frac{1}{4} c_\sigma |\xi|^2
\end{aligned} \tag{20}$$

and therefore,

$$\begin{aligned}
\left\langle \sigma^\lambda \sigma^{\lambda^*}(t, X_t^{(1)}, X_t^{(2)})\xi, \xi \right\rangle &= \lambda^2 \left\langle \sigma\sigma^*(t, X_t^{(1)})\xi, \xi \right\rangle + (1-\lambda)^2 \left\langle \sigma\sigma^*(t, X_t^{(2)})\xi, \xi \right\rangle \\
&\quad + 2\lambda(1-\lambda) \left\langle \sigma(t, X_t^{(1)})\sigma^*(t, X_t^{(2)})\xi, \xi \right\rangle \\
&= \lambda^2 \left\langle \sigma\sigma^*(t, X_t^{(1)})\xi, \xi \right\rangle + (1 - 2\lambda + \lambda^2) \left\langle \sigma\sigma^*(t, X_t^{(2)})\xi, \xi \right\rangle \\
&\quad + 2\lambda(1-\lambda) \left\langle \sigma(t, X_t^{(1)})\sigma^*(t, X_t^{(2)})\xi, \xi \right\rangle \\
&= \lambda^2 \left\langle \sigma\sigma^*(t, X_t^{(1)})\xi, \xi \right\rangle + (1 - \lambda^2) \left\langle \sigma\sigma^*(t, X_t^{(2)})\xi, \xi \right\rangle \\
&\quad + (2\lambda^2 - 2\lambda) \left\langle \sigma\sigma^*(t, X_t^{(2)})\xi, \xi \right\rangle \\
&\quad + 2\lambda(1-\lambda) \left\langle \sigma(t, X_t^{(1)})\sigma^*(t, X_t^{(2)})\xi, \xi \right\rangle \\
&= \lambda^2 \left\langle \sigma\sigma^*(t, X_t^{(1)})\xi, \xi \right\rangle + (1 - \lambda^2) \left\langle \sigma\sigma^*(t, X_t^{(2)})\xi, \xi \right\rangle \\
&\quad + 2\lambda(1-\lambda) \left\langle \left(\sigma(t, X_t^{(1)}) - \sigma(t, X_t^{(2)})\right) \sigma^*(t, X_t^{(2)})\xi, \xi \right\rangle .
\end{aligned}$$

Together with estimate (20) and (c3) we obtain that

$$\left\langle \sigma^\lambda \sigma^{\lambda^*}(t, X_t^{(1)}, X_t^{(2)})\xi, \xi \right\rangle \ge \lambda^2 c_\sigma |\xi|^2 + (1-\lambda^2) c_\sigma |\xi|^2 - 2\lambda(1-\lambda)\frac{1}{4} c_\sigma |\xi|^2 \ge \frac{1}{2} c_\sigma |\xi|^2.$$

This shows that for $0 \le t \le T \wedge \tau_R \wedge \tau_\varepsilon$ all the eigenvalues of $\sigma^\lambda \sigma^{\lambda^*}$ are bounded from below by $\frac{1}{2} c_\sigma$ and therefore, we can estimate the determinant:

$$\det(a_t^\lambda) = \frac{1}{2^d} \det(\sigma^\lambda \sigma^{\lambda^*}(t, X_t^{(1)}, X_t^{(2)})) \ge \frac{1}{2^{2d}} c_\sigma^d.$$

Note, that $\tau_R \le \tau_R^\lambda$ since $|X_t^{(1)}| \le R$ and $|X_t^{(2)}| \le R$ imply that

$$|\lambda X_t^{(1)} + (1-\lambda) X_t^{(2)}| \le \lambda |X_t^{(1)}| + (1-\lambda)|X_t^{(2)}| \le R.$$

So, we obtain on the set $\{t_0 \leq \tau_R \wedge \tau_\varepsilon\}$

$$\begin{aligned}
&\mathbb{E}\left[\int_{t_0}^{T\wedge\tau_R\wedge\tau_\varepsilon} f(t, X_t^\lambda)\, dt \,\middle|\, \mathcal{F}_{t_0}\right] \\
&= \mathbb{E}\left[\int_{t_0}^{T\wedge\tau_R\wedge\tau_\varepsilon} \left(\frac{\det(a_t^\lambda)}{\det(a_t^\lambda)}\right)^{\frac{1}{d+1}} f(t, X_t^\lambda)\, dt \,\middle|\, \mathcal{F}_{t_0}\right] \\
&\leq \left(\frac{4}{c_\sigma}\right)^{\frac{d}{d+1}} \mathbb{E}\left[\int_{t_0}^{T\wedge\tau_R\wedge\tau_\varepsilon} \det(a_t^\lambda)^{\frac{1}{d+1}} f(t, X_t^\lambda)\, dt \,\middle|\, \mathcal{F}_{t_0}\right] \\
&\leq \left(\frac{4}{c_\sigma}\right)^{\frac{d}{d+1}} \mathbb{E}\left[\int_{t_0}^{T\wedge\tau_R^\lambda\wedge\tau_\varepsilon} \det(a_t^\lambda)^{\frac{1}{d+1}} f(t, X_t^\lambda)\, dt \,\middle|\, \mathcal{F}_{t_0}\right] \qquad \mathbb{P}\text{-a. s..}
\end{aligned}$$

With Lemma 4.2 and Lemma 4.3, we deduce that

$$\mathbb{E}\left[\int_{t_0}^{T\wedge\tau_R\wedge\tau_\varepsilon} f(t, X_t^\lambda)\, dt \,\middle|\, \mathcal{F}_{t_0}\right] \leq C(d, p, q, v, r, T, \|b\|_{L_p^q(T)}, c_\sigma, \tilde{c}_\sigma)\|f\|_{L_v^r(T)} \qquad \mathbb{P}\text{-a. s.}$$

on the set $\{t_0 \leq \tau_R \wedge \tau_\varepsilon\}$. □

4.2. Uniform exponential estimate for the transformed diffusion

In this section we prove that for

$$A_t^{(n)} := \int_0^t \frac{\left|\sigma^{(n)}(s, X_s^{(1)}) - \sigma^{(n)}(s, X_s^{(2)})\right|^2}{\left|Y_s^{(1,n)} - Y_s^{(2,n)}\right|^2} \mathbb{1}_{\{Y_s^{(1,n)} \neq Y_s^{(2,n)}\}}\, ds$$

we have that $\mathbb{E}[e^{A_T^{(n)}}]$ is uniformly bounded in n. To this end we need a Khasminski-type estimate, as stated in Lemma 4.5, to get the exponential estimate via conditional expectations. This is done in Proposition 4.6.

Lemma 4.5. *Let $f : [0, T] \times \mathbb{R}^d \to \mathbb{R}$ be a nonnegative measurable function and γ an arbitrary stopping time. Assume that X_t is an adapted process such there exists a constant $\alpha < 1$ with*

$$\mathbb{1}_{\{t_0 \leq \gamma\}} \mathbb{E}\left[\int_{t_0}^{T\wedge\gamma} f(t, X_t)\, dt \,\middle|\, \mathcal{F}_{t_0}\right] \leq \alpha \qquad \mathbb{P}\text{-a. s. } \forall\ 0 \leq t_0 \leq T.$$

Then we have

$$\mathbb{E}\left[\exp\left(\int_0^{T\wedge\gamma} f(t, X_t)\, dt\right)\right] \leq \frac{1}{1-\alpha}.$$

Proof. We have

$$\mathbb{E}\left[\exp\left(\int_0^{T\wedge\gamma} f(t,X_t)\,dt\right)\right] = \mathbb{E}\left[\sum_{n=0}^{\infty}\frac{1}{n!}\left(\int_0^{T\wedge\gamma} f(t,X_t)\,dt\right)^n\right].$$

By induction one can prove that

$$\left(\int_0^T f(t,X_t)\,dt\right)^n = n!\int_0^T\int_{s_1}^T \dots \int_{s_{n-1}}^T f(s_1,X_{s_1})f(s_2,X_{s_2})\cdot\ldots\cdot f(s_n,X_{s_n})\,ds_n\ldots ds_2\,ds_1,$$

and thus

$$\begin{aligned}
&\mathbb{E}\left[\exp\left(\int_0^{T\wedge\gamma} f(t,X_t)\,dt\right)\right]\\
&= \sum_{n=0}^{\infty}\mathbb{E}\Bigg[\int_0^{T\wedge\gamma}\int_{s_1}^{T\wedge\gamma}\cdots\int_{s_{n-2}}^{T\wedge\gamma}\int_{s_{n-1}}^{T\wedge\gamma} f(s_1,X_{s_1})f(s_2,X_{s_2})\cdot\ldots\\
&\qquad\qquad\cdot f(s_{n-1},X_{s_{n-1}})f(s_n,X_{s_n})\,ds_n\,ds_{n-1}\ldots ds_2\,ds_1\Bigg]\\
&= \sum_{n=0}^{\infty}\mathbb{E}\Bigg[\int_0^T\int_{s_1}^T\cdots\int_{s_{n-2}}^T\int_{s_{n-1}}^T \mathbb{1}_{\{s_1\le\gamma\}}f(s_1,X_{s_1})\mathbb{1}_{\{s_2\le\gamma\}}f(s_2,X_{s_2})\cdot\ldots\\
&\qquad\qquad\cdot\mathbb{1}_{\{s_{n-1}\le\gamma\}}f(s_{n-1},X_{s_{n-1}})\mathbb{1}_{\{s_n\le\gamma\}}f(s_n,X_{s_n})\,ds_n\,ds_{n-1}\ldots ds_2\,ds_1\Bigg]\\
&= \sum_{n=0}^{\infty}\int_0^T\int_{s_1}^T\cdots\int_{s_{n-2}}^T \mathbb{E}\Bigg[\mathbb{1}_{\{s_1\le\gamma\}}f(s_1,X_{s_1})\mathbb{1}_{\{s_2\le\gamma\}}f(s_2,X_{s_2})\cdot\ldots\\
&\qquad\qquad\cdot\mathbb{1}_{\{s_{n-1}\le\gamma\}}f(s_{n-1},X_{s_{n-1}})\int_{s_{n-1}}^T \mathbb{1}_{\{s_n\le\gamma\}}f(s_n,X_{s_n})\,ds_n\Bigg]\,ds_{n-1}\ldots ds_2\,ds_1\\
&= \sum_{n=0}^{\infty}\int_0^T\int_{s_1}^T\cdots\int_{s_{n-2}}^T \mathbb{E}\Bigg[\mathbb{E}\Bigg[\mathbb{1}_{\{s_1\le\gamma\}}f(s_1,X_{s_1})\mathbb{1}_{\{s_2\le\gamma\}}f(s_2,X_{s_2})\cdot\ldots\\
&\qquad\cdot\mathbb{1}_{\{s_{n-1}\le\gamma\}}f(s_{n-1},X_{s_{n-1}})\int_{s_{n-1}}^T \mathbb{1}_{\{s_n\le\gamma\}}f(s_n,X_{s_n})\,ds_n\;\Bigg|\;\mathcal{F}_{s_{n-1}}\Bigg]\Bigg]\,ds_{n-1}\ldots ds_2\,ds_1.
\end{aligned}$$

Since all the terms except for the last integral are measurable with respect to $\mathcal{F}_{s_{n-1}}$ we have

$$\begin{aligned}
&\mathbb{E}\left[\exp\left(\int_0^{T\wedge\gamma} f(t,X_t)\,dt\right)\right] \\
&= \sum_{n=0}^{\infty}\int_0^T\int_{s_1}^T\cdots\int_{s_{n-2}}^T \mathbb{E}\Bigg[\mathbb{1}_{\{s_1\le\gamma\}}f(s_1,X_{s_1})\mathbb{1}_{\{s_2\le\gamma\}}f(s_2,X_{s_2})\cdot\ldots \\
&\qquad\cdot\mathbb{1}_{\{s_{n-1}\le\gamma\}}f(s_{n-1},X_{s_{n-1}})\mathbb{E}\left[\int_{s_{n-1}}^{T\wedge\gamma} f(s_n,X_{s_n})\,ds_n \,\middle|\, \mathcal{F}_{s_{n-1}}\right]\Bigg]\,ds_{n-1}\ldots ds_2\,ds_1 \\
&\le \sum_{n=0}^{\infty}\alpha\int_0^T\int_{s_1}^T\cdots\int_{s_{n-3}}^T \mathbb{E}\Bigg[\mathbb{1}_{\{s_1\le\gamma\}}f(s_1,X_{s_1})\mathbb{1}_{\{s_2\le\gamma\}}f(s_2,X_{s_2})\cdot\ldots \\
&\qquad\cdot\mathbb{1}_{\{s_{n-2}\le\gamma\}}f(s_{n-2},X_{s_{n-2}})\int_{s_{n-2}}^T \mathbb{1}_{\{s_{n-1}\le\gamma\}}f(s_{n-1},X_{s_{n-1}})\,ds_{n-1}\Bigg]\,ds_{n-2}\ldots ds_2\,ds_1.
\end{aligned}$$

So, by iteration we get

$$\mathbb{E}\left[\exp\left(\int_0^{T\wedge\gamma} f(t,X_t)\,dt\right)\right] \le \sum_{n=0}^{\infty}\alpha^n = \frac{1}{1-\alpha}.$$

$\square$

Proposition 4.6. *Let Assumptions 2.2, 2.3 be fulfilled and $X_t^{(1)}$, $X_t^{(2)}$ be two solutions to* (5) *such that* (6) *holds. For*

$$A_t^{(n)} := \int_0^t \frac{\left|\sigma^{(n)}(s,X_s^{(1)})-\sigma^{(n)}(s,X_s^{(2)})\right|^2}{\left|Y_s^{(1,n)}-Y_s^{(2,n)}\right|^2}\mathbb{1}_{\{Y_s^{(1,n)}\neq Y_s^{(2,n)}\}}$$

and ε from Proposition 4.4, there exists a constant $C>0$ such that

$$\mathbb{E}\left[e^{A^{(n)}_{T\wedge\tau_R\wedge\tau_\varepsilon}}\right] \le C \qquad \textit{uniformly for all } n\in\mathbb{N}.$$

Proof. Considering $\sigma^{(n)}$ we find that:

$$\partial_{x_i}\sigma^{(n)} = \left(\partial_{x_i}\partial_x U^{(n)}\right)\sigma + \partial_x U^{(n)}\partial_{x_i}\sigma + \partial_{x_i}\sigma.$$

We use that σ is bounded and $\partial_x\sigma \in L_p^q(T)$, that $\partial_x U^{(n)}$ is uniformly bounded by $\frac{1}{2}$ and $\partial_x^2 U^{(n)}$ is equibounded in $L_p^q(T)$ (see Lemma 4.1) to deduce that

$$\|\partial_x\sigma^{(n)}\|_{L_p^q(T)} \le C \text{ uniformly in } n.$$

Additionally, $\sigma^{(n)}$ is continuous, since $\partial_x U^{(n)}$ is Hölder continuous. Then by Lemma A.6 there exists a sequence of continuous functions $(u_m)_m$, which are differentiable with respect to x in the ordinary sense, such that

$$u_m \to \sigma^{(n)} \text{ uniformly on } [0,T] \times \overline{B_R}$$

and

$$\|\partial_x u_m\|_{L_p^q(T)} \leq \|\partial_x \sigma^{(n)}\|_{L_p^q(T)} \quad \forall\, m \in \mathbb{N}.$$

Then we have with Lemma 4.1 (iv)

$$\begin{aligned}
&\mathbb{E}\left[\exp\left(\int_0^{T\wedge\tau_R\wedge\tau_\varepsilon} \frac{\left|\sigma^{(n)}(t,X_t^{(1)}) - \sigma^{(n)}(t,X_t^{(2)})\right|^2}{\left|Y_t^{(1,n)} - Y_t^{(2,n)}\right|^2} \mathbb{1}_{\{Y_t^{(1,n)} \neq Y_t^{(2,n)}\}}\, dt\right)\right] \\
&\leq \mathbb{E}\left[\exp\left(4 \int_0^{T\wedge\tau_R\wedge\tau_\varepsilon} \frac{\left|\sigma^{(n)}(t,X_t^{(1)}) - \sigma^{(n)}(t,X_t^{(2)})\right|^2}{\left|X_t^{(1)} - X_t^{(2)}\right|^2} \mathbb{1}_{\{X_t^{(1)} \neq X_t^{(2)}\}}\, dt\right)\right].
\end{aligned}$$

By uniform convergence, we receive that

$$\begin{aligned}
&\mathbb{E}\left[\exp\left(\int_0^{T\wedge\tau_R\wedge\tau_\varepsilon} \frac{\left|\sigma^{(n)}(t,X_t^{(1)}) - \sigma^{(n)}(t,X_t^{(2)})\right|^2}{\left|Y_t^{(1,n)} - Y_t^{(2,n)}\right|^2} \mathbb{1}_{\{Y_t^{(1,n)} \neq Y_t^{(2,n)}\}}\, dt\right)\right] \\
&\leq \lim_{m\to\infty} \mathbb{E}\left[\exp\left(4 \int_0^{T\wedge\tau_R\wedge\tau_\varepsilon} \frac{\left|u_m(t,X_t^{(1)}) - u_m(t,X_t^{(2)})\right|^2}{\left|X_t^{(1)} - X_t^{(2)}\right|^2} \mathbb{1}_{\{X_t^{(1)} \neq X_t^{(2)}\}}\, dt\right)\right] \\
&= \lim_{m\to\infty} \mathbb{E}\left[\exp\left(4 \int_0^{T\wedge\tau_R\wedge\tau_\varepsilon} \frac{\left|\int_0^1 \partial_x u_m(t,X_t^{\lambda})(X_t^{(1)} - X_t^{(2)})\, d\lambda\right|^2}{\left|X_t^{(1)} - X_t^{(2)}\right|^2} \mathbb{1}_{\{X_t^{(1)} \neq X_t^{(2)}\}}\, dt\right)\right] \\
&\leq \lim_{m\to\infty} \mathbb{E}\left[\exp\left(4 \int_0^{T\wedge\tau_R\wedge\tau_\varepsilon} \int_0^1 |\partial_x u_m(t,X_t^{\lambda})|^2\, d\lambda\, dt\right)\right].
\end{aligned}$$

An application of Fubini's Theorem and Jensen's inequality yields

$$\begin{aligned}
&\mathbb{E}\left[\exp\left(\int_0^{T\wedge\tau_R\wedge\tau_\varepsilon} \frac{\left|\sigma^{(n)}(t,X_t^{(1)}) - \sigma^{(n)}(t,X_t^{(2)})\right|^2}{\left|Y_t^{(1,n)} - Y_t^{(2,n)}\right|^2} \mathbb{1}_{\{Y_t^{(1,n)} \neq Y_t^{(2,n)}\}}\, dt\right)\right] \\
&\leq \lim_{m\to\infty} \int_0^1 \mathbb{E}\left[\exp\left(4 \int_0^{T\wedge\tau_R\wedge\tau_\varepsilon} |\partial_x u_m(t,X_t^{\lambda})|^2\, dt\right)\right] d\lambda.
\end{aligned}$$

Now, choose $\mu > 0$ so small that $p, q \geq 2(d+1)(1+\mu)$ which exists since $p, q > 2(d+1)$. Then we have for $\beta > 0$ with Young's inequality

$$
\begin{aligned}
&\mathbb{E}\left[\exp\left(\int_0^{T\wedge\tau_R\wedge\tau_\varepsilon} \frac{\left|\sigma^{(n)}(t, X_t^{(1)}) - \sigma^{(n)}(t, X_t^{(2)})\right|^2}{\left|Y_t^{(1,n)} - Y_t^{(2,n)}\right|^2} \mathbb{1}_{\{Y_t^{(1,n)} \neq Y_t^{(2,n)}\}}\, dt\right)\right] \\
&\leq \lim_{m\to\infty} \int_0^1 \mathbb{E}\left[\exp\left(\frac{4\beta}{\beta} \int_0^{T\wedge\tau_R\wedge\tau_\varepsilon} |\partial_x u_m(t, X_t^\lambda)|^2\, dt\right)\right] d\lambda \\
&\leq \lim_{m\to\infty} \int_0^1 \mathbb{E}\left[\exp\left(\frac{1}{\mu+1}\left(\beta \int_0^{T\wedge\tau_R\wedge\tau_\varepsilon} |\partial_x u_m(t, X_t^\lambda)|^2\, dt\right)^{1+\mu} + \frac{\mu}{1+\mu}\left(\frac{4}{\beta}\right)^{\frac{1+\mu}{\mu}}\right)\right] d\lambda.
\end{aligned}
$$

And with Hölder's inequality

$$
\begin{aligned}
&\mathbb{E}\left[\exp\left(\int_0^{T\wedge\tau_R\wedge\tau_\varepsilon} \frac{\left|\sigma^{(n)}(t, X_t^{(1)}) - \sigma^{(n)}(t, X_t^{(2)})\right|^2}{\left|Y_t^{(1,n)} - Y_t^{(2,n)}\right|^2} \mathbb{1}_{\{Y_t^{(1,n)} \neq Y_t^{(2,n)}\}}\, dt\right)\right] \\
&\leq \exp\left(\frac{\mu}{1+\mu}\left(\frac{4}{\beta}\right)^{\frac{1+\mu}{\mu}}\right) \\
&\quad \cdot \lim_{m\to\infty} \int_0^1 \mathbb{E}\left[\exp\left(\int_0^{T\wedge\tau_R\wedge\tau_\varepsilon} \frac{\beta^{1+\mu}}{1+\mu} T^{\frac{\mu}{1+\mu}} |\partial_x u_m(t, X_t^\lambda)|^{2(1+\mu)}\, dt\right)\right] d\lambda.
\end{aligned}
\tag{21}
$$

Furthermore, we have with Proposition 4.4 for all $0 \leq t_0 \leq T$ on the set $\{t_0 \leq \tau_R \wedge \tau_\varepsilon\}$

$$
\begin{aligned}
&\mathbb{E}\left[\int_{t_0}^{T\wedge\tau_R\wedge\tau_\varepsilon} \frac{\beta^{1+\mu}}{1+\mu} T^{\frac{\mu}{1+\mu}} \left|\partial_x u_m(t, X_t^\lambda)\right|^{2(1+\mu)}\, dt \,\middle|\, \mathcal{F}_{t_0}\right] \\
&\leq C(d, p, q, \mu, T, c_\sigma, \tilde{c}_\sigma, \|b\|_{L_p^q(T)}) \frac{\beta^{1+\mu}}{1+\mu} T^{\frac{\mu}{1+\mu}} \||\partial_x u_m|^{2(1+\mu)}\|_{L^{\frac{q}{2(1+\mu)}}_{\frac{p}{2(1+\mu)}}(T)} \\
&= C(d, p, q, \mu, T, c_\sigma, \tilde{c}_\sigma, \|b\|_{L_p^q(T)}) \beta^{1+\mu} \|\partial_x u_m\|_{L_p^q(T)}^{2(1+\mu)} \\
&\leq C(d, p, q, \mu, T, c_\sigma, \tilde{c}_\sigma, \|b\|_{L_p^q(T)}) \beta^{1+\mu} \|\partial_x \sigma^{(n)}\|_{L_p^q(T)}^{2(1+\mu)} \\
&=: \alpha.
\end{aligned}
$$

Since $\|\partial_x \sigma^{(n)}\|_{L_p^q(T)}$ is equibounded, we can choose β so small that this is less than 1 for all $n \in \mathbb{N}$. Then we have by Lemma 4.5 and inequality (21) that

$$
\begin{aligned}
\mathbb{E}\left[e^{A^{(n)}_{T\wedge\tau_R\wedge\tau_\varepsilon}}\right] &\leq \exp\left(\frac{\mu}{1+\mu}\left(\frac{4}{\beta}\right)^{\frac{1+\mu}{\mu}}\right) \frac{1}{1-\alpha} \\
&\leq C,
\end{aligned}
$$

where C does not depend on n. □

4.3. Convergence of the transformed drift

In the following we prove that

$$\mathbb{E}\left[\int_0^T \left|b^{(n)}(t, X_t^{(1)}) - b^{(n)}(t, X_t^{(2)})\right|^2 dt\right]$$

converges to 0 for $n \to \infty$. The proof is much simpler than in [FF11] since we are able to apply Krylov's estimate. The price to pay is that we have to assume $p, q > 2(d+1)$.

Lemma 4.7. *Let* $(c1)$*,* $(c3)$*,* $(c4)$ *of Assumption 2.2 and Assumption 2.3 be fulfilled and* $X_t^{(1)}$*,* $X_t^{(2)}$ *be two solutions of* (5) *such that condition* (6) *holds. Then we have*

$$\lim_{n\to\infty} \mathbb{E}\left[\int_0^T \left|b^{(n)}(t, X_t^{(1)}) - b^{(n)}(t, X_t^{(2)})\right|^2 dt\right] = 0.$$

Proof. Young's inequality, Lemma 3.1 with $v = \frac{p}{2}$, $r = \frac{q}{2}$ and an application of Lemma 4.1 on the arising $\|\mathcal{T}^{n+1}(b)\|^2$ term yields

$$\begin{aligned}
&\mathbb{E}\left[\int_0^T \left|b^{(n)}(t, X_t^{(1)}) - b^{(n)}(t, X_t^{(2)})\right|^2 dt\right] \\
&\leq 2\mathbb{E}\left[\int_0^T \left|b^{(n)}(t, X_t^{(1)})\right|^2 dt\right] + 2\mathbb{E}\left[\int_0^T \left|b^{(n)}(t, X_t^{(2)})\right|^2 dt\right] \\
&\leq C(d, p, q, T, c_\sigma, \tilde{c}_\sigma, \|b\|_{L_p^q(T)})\|b^{(n)}\|^2_{L_p^q(T)} \\
&= C(d, p, q, T, c_\sigma, \tilde{c}_\sigma, \|b\|_{L_p^q(T)})\|\mathcal{T}^{n+1}(b)\|^2_{L_p^q(T)} \\
&\leq C(d, p, q, T, c_\sigma, \tilde{c}_\sigma, \|b\|_{L_p^q(T)})\frac{1}{2^{2(n+1)}}\|b\|^2_{L_p^q(T)} \\
&\xrightarrow{n\to\infty} 0.
\end{aligned}$$

□

4.4. Bounded first and second moments

In this section we show that $|X_t|$ and $|X_t|^2$ are integrable which is the last necessary tool to prove pathwise uniqueness. Furthermore, we also obtain the finiteness of $\mathbb{E}[\sup_{t\in[0,T]}|X_t|]$ and therefore, there is no explosion for our SDE. All we need is Lemma 3.1 and the inequality of Burkholder, Davis and Gundy.

Lemma 4.8. *Let* $(c1)$, $(c3)$ *and* $(c4)$ *of Assumption 2.2 be fulfilled. If* X_t *is a solution to SDE* (5) *such that condition* (6) *holds, we have*

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

Proof. We have

$$\begin{aligned}
\mathbb{E}\left[\sup_{t\in[0,T]}|X_t|\right] &= \mathbb{E}\left[\sup_{t\in[0,T]}\left| x + \int_0^t b(s,X_s)\,ds + \int_0^t \sigma(s,X_s)\,dW_s \right|\right] \\
&\le |x| + \mathbb{E}\left[\sup_{t\in[0,T]}\int_0^t |b(s,X_s)|\,ds\right] + \mathbb{E}\left[\sup_{t\in[0,T]}\left|\int_0^t \sigma(s,X_s)\,dW_s\right|\right] \\
&\le |x| + \mathbb{E}\left[\int_0^T |b(s,X_s)|\,ds\right] + \mathbb{E}\left[\sup_{t\in[0,T]}\left|\int_0^t \sigma(s,X_s)\,dW_s\right|\right].
\end{aligned}$$

Then applications of Lemma 3.1 to the first expectation term and of the inequality of Burkholder, Davis and Gundy (see e.g. [RY05] Corollary IV.4.2) to the second yield

$$\mathbb{E}\left[\sup_{t\in[0,T]}|X_t|\right] \le |x| + C\|b\|_{L_p^q(T)} + C\mathbb{E}\left[\left(\int_0^T \sigma(s,X_s)^2\,ds\right)^{\frac{1}{2}}\right].$$

Since σ is bounded and $b \in L_p^q(T)$, this is finite. Furthermore,

$$\begin{aligned}
&\sup_{t\in[0,T]}\mathbb{E}\left[|X_t|^2\right] \\
&= \sup_{t\in[0,T]}\mathbb{E}\left[\left| x + \int_0^t b(s,X_s)\,ds + \int_0^t \sigma(s,X_s)\,dW_s \right|^2\right] \\
&\le 2|x|^2 + 2\sup_{t\in[0,T]}\mathbb{E}\left[\left| \int_0^t b(s,X_s)\,ds + \int_0^t \sigma(s,X_s)\,dW_s \right|^2\right] \\
&\le 2|x|^2 + 4\sup_{t\in[0,T]}\mathbb{E}\left[\left| \int_0^t b(s,X_s)\,ds \right|^2\right] + 4\sup_{t\in[0,T]}\mathbb{E}\left[\left| \int_0^t \sigma(s,X_s)\,dW_s \right|^2\right].
\end{aligned}$$

We apply Hölder's inequality to the first expectation and the multidimensional Itô Isometry to the second one to receive

$$\sup_{t\in[0,T]}\mathbb{E}\left[|X_t|^2\right] \le 2|x|^2 + 4T\mathbb{E}\left[\int_0^T |b(s,X_s)|^2\,ds\right] + 4\sup_{t\in[0,T]}\mathbb{E}\left[\int_0^t |\sigma(s,X_s)|^2\,ds\right].$$

Again, we use Lemma 3.1, (c4) and (c1) to obtain

$$\sup_{t\in[0,T]} \mathbb{E}\left[|X_t|^2\right] \leq 2|x|^2 + 4TC\|b\|^2_{L^q_p(T)} + 4T\tilde{c}^2_\sigma < \infty.$$

$\square$

5. Pathwise uniqueness

This section is devoted to the proof of Theorem 2.7. We start on a small interval $[0,T]$ which is given by Lemma 4.1 and show pathwise uniqueness up to this T just by estimating the difference of two solutions with the same initial values. In the second part we then conclude that it is possible to extend this to arbitrarily large T.

5.1. On small intervals

To prove pathwise uniqueness we show that the expectation of the difference of two solutions with the same initial values is zero. After the preparation in the previous sections this can be done easily, when we pass over from X_t to $Y_t^{(n)}$, which is given by our transformation. Using Itô's formula, Lemma 4.1 and Proposition 4.6 we find that the expectation of the difference of two solutions is bounded by a term depending on n. By taking the limit $n \to \infty$ and applying Lemma 4.7 and Lemma 4.8 we then finally conclude that pathwise uniqueness holds.

Proof of Theorem 2.7 for small T. In the following, we denote by x^i the i-th entry of a vector $x \in \mathbb{R}^d$. Let Assumptions 2.2 and 2.3 be fulfilled and $X_t^{(1)}$, $X_t^{(2)}$ be two solutions to (5), such that condition (6) holds. Furthermore, let $T := T_0$ from Lemma 4.1 and $Y_t^{(i,n)}$ given by (14) in Section 3.2. By Itô's formula we then have

$$
\begin{aligned}
d\left|Y_t^{(1,n)} - Y_t^{(2,n)}\right|^2 &= \sum_{i=1}^{d} 2\left(Y_t^{(1,n)} - Y_t^{(2,n)}\right)^i d\left(Y_t^{(1,n)} - Y_t^{(2,n)}\right)^i \\
&\quad + \frac{1}{2}\sum_{i=1}^{d} 2\left(d\left(Y_t^{(1,n)} - Y_t^{(2,n)}\right)^i\right)^2 \\
&= 2\sum_{i=1}^{d}\left(Y_t^{(1,n)} - Y_t^{(2,n)}\right)^i \left(b^{(n)}(t, X_t^{(1)}) - b^{(n)}(t, X_t^{(2)})\right)^i dt \\
&\quad + 2\sum_{i=1}^{d}\left(Y_t^{(1,n)} - Y_t^{(2,n)}\right)^i \left(\left(\sigma^{(n)}(t, X_t^{(1)}) - \sigma^{(n)}(t, X_t^{(2)})\right) dW_t\right)^i \\
&\quad + \sum_{i=1}^{d}\left(\sum_{j=1}^{m}\left(\sigma^{(n)}(t, X_t^{(1)}) - \sigma^{(n)}(t, X_t^{(2)})\right)_{ij} dW_t^j\right)^2 \\
&= 2\sum_{i=1}^{d}\left(Y_t^{(1,n)} - Y_t^{(2,n)}\right)^i \left(b^{(n)}(t, X_t^{(1)}) - b^{(n)}(t, X_t^{(2)})\right)^i dt \\
&\quad + 2\left\langle Y_t^{(1,n)} - Y_t^{(2,n)}, \left(\sigma^{(n)}(t, X_t^{(1)}) - \sigma^{(n)}(t, X_t^{(2)})\right) dW_t\right\rangle \\
&\quad + \sum_{i=1}^{d}\sum_{j=1}^{m}\left(\sigma^{(n)}(t, X_t^{(1)}) - \sigma^{(n)}(t, X_t^{(2)})\right)_{ij}^2 dt,
\end{aligned}
$$

and therefore

$$\begin{aligned} d\left|Y_t^{(1,n)} - Y_t^{(2,n)}\right|^2 = {} & 2\sum_{i=1}^{d}\left(Y_t^{(1,n)} - Y_t^{(2,n)}\right)^i \left(b^{(n)}(t, X_t^{(1)}) - b^{(n)}(t, X_t^{(2)})\right)^i dt \\ & + 2\left\langle Y_t^{(1,n)} - Y_t^{(2,n)}, \left(\sigma^{(n)}(t, X_t^{(1)}) - \sigma^{(n)}(t, X_t^{(2)})\right) dW_t\right\rangle \\ & + \left|\sigma^{(n)}(t, X_t^{(1)}) - \sigma^{(n)}(t, X_t^{(2)})\right|^2 dt. \end{aligned}$$

Applying the inequality of Cauchy and Schwarz yields

$$\begin{aligned} d\left|Y_t^{(1,n)} - Y_t^{(2,n)}\right|^2 \le {} & 2\left|Y_t^{(1,n)} - Y_t^{(2,n)}\right| \left|b^{(n)}(t, X_t^{(1)}) - b^{(n)}(t, X_t^{(2)})\right| dt \\ & + 2\left\langle Y_t^{(1,n)} - Y_t^{(2,n)}, \left(\sigma^{(n)}(t, X_t^{(1)}) - \sigma^{(2)}(t, X_t^{(2)})\right) dW_t\right\rangle \\ & + \left|\sigma^{(n)}(t, X_t^{(1)}) - \sigma^{(n)}(t, X_t^{(2)})\right|^2 dt. \end{aligned} \tag{22}$$

Moreover, with

$$A_t^{(n)} = \int_0^t \frac{\left|\sigma^{(n)}(s, X_s^{(1)}) - \sigma^{(n)}(s, X_s^{(2)})\right|^2}{\left|Y_s^{(1,n)} - Y_s^{(2,n)}\right|^2} \mathbb{1}_{\{Y_s^{(1,n)} \neq Y_s^{(2,n)}\}} ds,$$

from Proposition 4.6, we have

$$\begin{aligned} d\left(e^{-A_t^{(n)}}\left|Y_t^{(1,n)} - Y_t^{(2,n)}\right|^2\right) = {} & e^{-A_t^{(n)}} d\left|Y_t^{(1,n)} - Y_t^{(2,n)}\right|^2 + \left|Y_t^{(1,n)} - Y_t^{(2,n)}\right|^2 de^{-A_t^{(n)}} \\ & + d\left[e^{-A_{(\cdot)}^{(n)}}, \left|Y_{(\cdot)}^{(1,n)} - Y_{(\cdot)}^{(2,n)}\right|^2\right]_t, \end{aligned}$$

where $[\,\cdot\,,\,\cdot\,]$ denotes the quadratic covariation which is zero due to the monotonicity of $e^{-A_t^{(n)}}$. So, we deduce that

$$d\left(e^{-A_t^{(n)}}\left|Y_t^{(1,n)} - Y_t^{(2,n)}\right|^2\right) = e^{-A_t^{(n)}} d\left|Y_t^{(1,n)} - Y_t^{(2,n)}\right|^2 - \left|Y_t^{(1,n)} - Y_t^{(2,n)}\right|^2 e^{-A_t^{(n)}} dA_t^{(n)}.$$

Now, we use inequality (22) to conclude that

$$\begin{aligned} & d\left(e^{-A_t^{(n)}}\left|Y_t^{(1,n)} - Y_t^{(2,n)}\right|^2\right) \\ & \le 2e^{-A_t^{(n)}}\left|Y_t^{(1,n)} - Y_t^{(2,n)}\right| \left|b^{(n)}(t, X_t^{(1)}) - b^{(n)}(t, X_t^{(2)})\right| dt \\ & \quad + 2e^{-A_t^{(n)}}\left\langle Y_t^{(1,n)} - Y_t^{(2,n)}, \left(\sigma^{(n)}(t, X_t^{(1)}) - \sigma^{(n)}(t, X_t^{(2)})\right) dW_t\right\rangle \\ & \quad + e^{-A_t^{(n)}}\left|\sigma^{(n)}(t, X_t^{(1)}) - \sigma^{(n)}(t, X_t^{(2)})\right|^2 dt \\ & \quad - e^{-A_t^{(n)}}\left|Y_t^{(1,n)} - Y_t^{(2,n)}\right|^2 dA_t^{(n)}. \end{aligned}$$

For the last term we find that

$$\begin{aligned}
&e^{-A_t^{(n)}} \left|Y_t^{(1,n)} - Y_t^{(2,n)}\right|^2 dA_t^{(n)} \\
&\quad = e^{-A_t^{(n)}} \left|Y_t^{(1,n)} - Y_t^{(2,n)}\right|^2 \frac{\left|\sigma^{(n)}(t, X_t^{(1)}) - \sigma^{(n)}(t, X_t^{(2)})\right|^2}{\left|Y_t^{(1,n)} - Y_t^{(2,n)}\right|^2} \mathbb{1}_{\left\{Y_t^{(1,n)} \neq Y_t^{(2,n)}\right\}} \, dt \\
&\quad = e^{-A_t^{(n)}} \left|\sigma^{(n)}(t, X_t^{(1)}) - \sigma^{(n)}(t, X_t^{(2)})\right|^2 \, dt.
\end{aligned}$$

Therefore, we get

$$\begin{aligned}
&d\left(e^{-A_t^{(n)}} \left|Y_t^{(1,n)} - Y_t^{(2,n)}\right|^2\right) \\
&\quad \leq 2e^{-A_t^{(n)}} \left|Y_t^{(1,n)} - Y_t^{(2,n)}\right| \left|b^{(n)}(t, X_t^{(1)}) - b^{(n)}(t, X_t^{(2)})\right| \, dt \\
&\qquad + 2e^{-A_t^{(n)}} \left\langle Y_t^{(1,n)} - Y_t^{(2,n)}, \left(\sigma^{(n)}(t, X_t^{(1)}) - \sigma^{(n)}(t, X_t^{(2)})\right) \, dW_t \right\rangle
\end{aligned}$$

and thus,

$$\begin{aligned}
&\mathbb{E}\left[e^{-A_t^{(n)}} \left|Y_t^{(1,n)} - Y_t^{(2,n)}\right|^2\right] \\
&\quad \leq \mathbb{E}\left[\left|Y_0^{(1,n)} - Y_0^{(2,n)}\right|^2\right] \\
&\qquad + 2\mathbb{E}\left[\int_0^t e^{-A_s^{(n)}} \left|Y_s^{(1,n)} - Y_s^{(2,n)}\right| \left|b^{(n)}(s, X_s^{(1)}) - b^{(n)}(s, X_s^{(2)})\right| \, ds\right] \\
&\qquad + 2\mathbb{E}\left[\int_0^t e^{-A_s^{(n)}} \left\langle Y_s^{(1,n)} - Y_s^{(2,n)}, \left(\sigma^{(n)}(s, X_s^{(1)}) - \sigma^{(n)}(s, X_s^{(2)})\right) \, dW_s \right\rangle\right].
\end{aligned}$$

With the help of Lemma 4.1, we get

$$\begin{aligned}
&\mathbb{E}\left[e^{-A_t^{(n)}} \left|Y_t^{(1,n)} - Y_t^{(2,n)}\right|^2\right] \\
&\quad \leq \frac{9}{4} \left|x^{(1)} - x^{(2)}\right|^2 \\
&\qquad + 3\mathbb{E}\left[\int_0^t \left|X_s^{(1)} - X_s^{(2)}\right| \left|b^{(n)}(s, X_s^{(1)}) - b^{(n)}(s, X_s^{(2)})\right| \, ds\right] \qquad (23) \\
&\qquad + 2\mathbb{E}\left[\int_0^t e^{-A_s^{(n)}} \left\langle Y_s^{(1,n)} - Y_s^{(2,n)}, \left(\sigma^{(n)}(s, X_s^{(1)}) - \sigma^{(n)}(s, X_s^{(2)})\right) \, dW_s \right\rangle\right].
\end{aligned}$$

Summarizing, for two solutions with the same initial values, $R > 0$, and ε from Proposition 4.4, we have for all $t \leq T$

$$\begin{aligned}
\mathbb{E}\left[\left|X^{(1)}_{t\wedge\tau_R\wedge\tau_\varepsilon} - X^{(2)}_{t\wedge\tau_R\wedge\tau_\varepsilon}\right|\right] &\leq \mathbb{E}\left[2\left|Y^{(1,n)}_{t\wedge\tau_R\wedge\tau_\varepsilon} - Y^{(2,n)}_{t\wedge\tau_R\wedge\tau_\varepsilon}\right|\right] \\
&= 2\mathbb{E}\left[e^{\frac{1}{2}A^{(n)}_{t\wedge\tau_R\wedge\tau_\varepsilon}} e^{-\frac{1}{2}A^{(n)}_{t\wedge\tau_R\wedge\tau_\varepsilon}}\left|Y^{(1,n)}_{t\wedge\tau_R\wedge\tau_\varepsilon} - Y^{(2,n)}_{t\wedge\tau_R\wedge\tau_\varepsilon}\right|\right] \\
&\leq 2\mathbb{E}\left[e^{A^{(n)}_{t\wedge\tau_R\wedge\tau_\varepsilon}}\right]^{\frac{1}{2}} \mathbb{E}\left[e^{-A^{(n)}_{t\wedge\tau_R\wedge\tau_\varepsilon}}\left|Y^{(1,n)}_{t\wedge\tau_R\wedge\tau_\varepsilon} - Y^{(2,n)}_{t\wedge\tau_R\wedge\tau_\varepsilon}\right|^2\right]^{\frac{1}{2}}.
\end{aligned}$$

With Proposition 4.6 and inequality (23) we obtain

$$\begin{aligned}
&\mathbb{E}\left[\left|X^{(1)}_{t\wedge\tau_R\wedge\tau_\varepsilon} - X^{(2)}_{t\wedge\tau_R\wedge\tau_\varepsilon}\right|\right] \\
&\leq C\left(\mathbb{E}\left[\int_0^T \left|X^{(1)}_s - X^{(2)}_s\right| \left|b^{(n)}(s, X^{(1)}_s) - b^{(n)}(s, X^{(2)}_s)\right| \, ds\right]\right. \\
&\qquad \left. + \mathbb{E}\left[\int_0^{t\wedge\tau_R\wedge\tau_\varepsilon} e^{-A^{(n)}_s} \left\langle Y^{(1,n)}_s - Y^{(2,n)}_s, \left(\sigma^{(n)}(s, X^{(1)}_s) - \sigma^{(n)}(s, X^{(2)}_s)\right) \, dW_s\right\rangle\right]\right)^{\frac{1}{2}}.
\end{aligned}$$

The second expectation term vanishes due to the martingale property of the stochastic integral which is well defined as $\sigma^{(n)}$ is bounded and $|Y^{(1,n)}_t - Y^{(2,n)}_t|^2$ is integrable by Lemma 4.8. That leads to

$$\begin{aligned}
&\mathbb{E}\left[\left|X^{(1)}_{t\wedge\tau_R\wedge\tau_\varepsilon} - X^{(2)}_{t\wedge\tau_R\wedge\tau_\varepsilon}\right|\right] \\
&\qquad \leq C\mathbb{E}\left[\int_0^T \left|X^{(1)}_s - X^{(2)}_s\right| \left|b^{(n)}(s, X^{(1)}_s) - b^{(n)}(s, X^{(2)}_s)\right| \, ds\right]^{\frac{1}{2}} \quad \forall\, n \in \mathbb{N}.
\end{aligned}$$

And therefore,

$$\begin{aligned}
&\mathbb{E}\left[\left|X^{(1)}_{t\wedge\tau_R\wedge\tau_\varepsilon} - X^{(2)}_{t\wedge\tau_R\wedge\tau_\varepsilon}\right|\right] \\
&\leq C \limsup_{n\to\infty} \mathbb{E}\left[\int_0^T \left|X^{(1)}_s - X^{(2)}_s\right| \left|b^{(n)}(s, X^{(1)}_s) - b^{(n)}(X^{(2)}_s)\right| \, ds\right]^{\frac{1}{2}} \\
&\leq C\mathbb{E}\left[\int_0^T \left|X^{(1)}_s - X^{(2)}_s\right|^2 \, ds\right]^{\frac{1}{4}} \\
&\qquad \cdot \limsup_{n\to\infty} \mathbb{E}\left[\int_0^T \left|b^{(n)}(s, X^{(1)}_s) - b^{(n)}(s, X^{(2)}_s)\right|^2 \, ds\right]^{\frac{1}{4}}.
\end{aligned}$$

With Fubini's Theorem, we then obtain

$$\begin{aligned}
&\mathbb{E}\left[\left|X^{(1)}_{t\wedge\tau_R\wedge\tau_\varepsilon} - X^{(2)}_{t\wedge\tau_R\wedge\tau_\varepsilon}\right|\right] \\
&\leq C\left(\int_0^T \mathbb{E}\left[\left|X^{(1)}_s - X^{(2)}_s\right|^2\right]\, ds\right)^{\frac{1}{4}} \\
&\qquad \cdot \limsup_{n\to\infty} \mathbb{E}\left[\int_0^T \left|b^{(n)}(s, X^{(1)}_s) - b^{(n)}(s, X^{(2)}_s)\right|^2\, ds\right]^{\frac{1}{4}} \\
&= 0
\end{aligned}$$

since by Lemma 4.8 the first term is bounded and by Lemma 4.7 the second one is zero. So, we have

$$X^{(1)}_t = X^{(2)}_t \quad \mathbb{P}\text{-a.s.} \quad \forall\, t \leq T \wedge \tau_R \wedge \tau_\varepsilon.$$

By the definition of τ_ε, the equality holds true for all $t \leq T \wedge \tau_R$. Since we can take R arbitrarily large, we have

$$X^{(1)}_t = X^{(2)}_t \quad \mathbb{P}\text{-a.s.} \quad \forall\, t \in [0, T].$$

Thus,

$$\mathbb{P}\left(X^{(1)}_t = X^{(2)}_t \ \forall\, t \in \mathbb{Q} \cap [0, T]\right) = 1$$

and by continuity of the solutions

$$\mathbb{P}\left(X^{(1)}_t = X^{(2)}_t \ \forall\, t \in [0, T]\right) = 1.$$

□

5.2. Extension from small to arbitrarily large intervals

Until now, we only proved pathwise uniqueness up to some possibly small T. Now, let T be arbitrarily large and take T_0 from Lemma 4.1. Let us shortly remind how T_0 was chosen. We assumed that σ is such that for all $f \in L^q_p(T)$ the equation

$$\partial_t u + \frac{1}{2}\sum_{i=1}^d \sum_{j=1}^d (\sigma\sigma^*)_{ij}\partial^2_{x_i x_j} u = -f, \qquad u(T, x) = 0$$

has a solution $u \in W^{1,2}_{q,p}(T)$ such that

$$\|u\|_{W^{1,2}_{q,p}(T)} \leq C\|f\|_{L^q_p(T)}, \tag{24}$$

where C is independent of f and increasing in T. In particular, we can find a uniform upper bound for C for all $T' \leq T$. By Hölder continuity of $\partial_x u$ we get

$$\sup_{(t,x)\in[0,T']\times\mathbb{R}^d} |\partial_x u(t, x)| \leq C(d, p, \varepsilon)(T')^{\frac{\varepsilon}{2}}\|u\|_{W^{1,2}_{q,p}(T')}$$

for every $\varepsilon \in (0,1)$ which fulfills

$$\varepsilon + \frac{d}{p} + \frac{2}{q} < 1.$$

By (24) we have therefore,

$$\sup_{(t,x)\in[0,T']\times\mathbb{R}^d} |\partial_x u(t,x)| \leq C(d,p,\varepsilon,T) T'^{\frac{\varepsilon}{2}} \|f\|_{L^q_p(T')}.$$

Note, that C depends on T but not on T'.
Because of the factor $T'^{\frac{\varepsilon}{2}}$, for fixed ε we may choose T_0 so small that

$$\sup_{(t,x)\in[0,T']\times\mathbb{R}^d} |\partial_x u(t,x)| \leq \frac{1}{4(\|b\|_{L^q_p(T)}+1)} \|f\|_{L^q_p(T')} \quad \text{uniformly } \forall\, T' \leq T_0.$$

For this small T_0 the statements of Lemma 4.1 are fulfilled and we have pathwise uniqueness on the interval $[0,T_0]$. We will show that the interval of pathwise uniqueness can be extended by means of a time-shift argument from $[0,T_0]$ to any $[kT_0,(k+1)T_0]$. For that purpose we define

$$\tilde{\sigma}(t,x) := \sigma(t+T_0,x) \qquad \text{for } t \in [0, T-T_0],$$

then for any $f \in L^q_p(T_0)$ there exists a solution $u \in W^{1,2}_{q,p}(T_0)$ to the equation

$$\partial_t u + \frac{1}{2}\sum_{i=1}^d \sum_{j=1}^d (\tilde{\sigma}\tilde{\sigma}^*)_{ij} \partial^2_{x_i x_j} u = -f \ \text{ on } [0,T_0], \quad u(T_0,x) = 0,$$

such that

$$\|u\|_{W^{1,2}_{q,p}(T_0)} \leq C\|f\|_{L^q_p(T_0)} \quad \text{and} \quad \sup_{(t,x)\in[0,T_0]\times\mathbb{R}^d} |\partial_x u(t,x)| \leq \frac{1}{4(\|b\|_{L^q_p(T)}+1)} \|f\|_{L^q_p(T_0)}.$$

If we take $\tilde{b}(t,x) := b(t+T_0,x)$ for $t \in [0,T_0]$ we have $\tilde{b} \in L^q_p(T_0)$ and a solution $u_{\tilde{b}}$ to the equation with $f = \tilde{b}$. Define $\tilde{u}_{\tilde{b}}(t,x) := u_{\tilde{b}}(t-T_0)$ for $t \in [T_0, 2T_0]$. Then we have for all $t \in [T_0,2T_0]$, $x \in \mathbb{R}^d$

$$\begin{aligned}
&\partial_t \tilde{u}_{\tilde{b}}(t,x) + \frac{1}{2}\sum_{i=1}^d \sum_{j=1}^d (\sigma\sigma^*(t,x))_{ij} \partial^2_{x_i x_j} \tilde{u}_{\tilde{b}}(t,x) \\
&\qquad = \partial_t u_{\tilde{b}}(t-T_0,x) + \frac{1}{2}\sum_{i=1}^d \sum_{j=1}^d (\tilde{\sigma}\tilde{\sigma}^*(t-T_0,x))_{ij} \partial^2_{x_i x_j} u_{\tilde{b}}(t-T_0,x) \\
&\qquad = -\tilde{b}(t-T_0,x) \\
&\qquad = -b(t,x)
\end{aligned}$$

and by Itô's formula55

$$\begin{aligned}
\tilde{u}_{\tilde{b}}(t, X_t) - \tilde{u}_{\tilde{b}}(T_0, X_{T_0}) &= \int_{T_0}^{t} \partial_t \tilde{u}_{\tilde{b}}(s, X_s)\, ds \\
&+ \frac{1}{2} \int_{T_0}^{t} \sum_{i=1}^{d} \sum_{j=1}^{d} (\sigma\sigma^*(s, X_s))_{ij} \partial^2_{x_i x_j} \tilde{u}_{\tilde{b}}(s, X_s)\, ds \\
&+ \int_{T_0}^{t} \partial_x \tilde{u}_{\tilde{b}}(s, X_s) b(s, X_s)\, ds + \int_{T_0}^{t} \partial_x \tilde{u}_{\tilde{b}}(s, X_s) \sigma(s, X_s)\, dW_s \\
&= - \int_{T_0}^{t} b(s, X_s)\, ds \\
&+ \int_{T_0}^{t} \partial_x \tilde{u}_{\tilde{b}}(s, X_s) b(s, X_s)\, ds + \int_{T_0}^{t} \partial_x \tilde{u}_{\tilde{b}}(s, X_s) \sigma(s, X_s)\, dW_s
\end{aligned}$$

which shows that we can transform the SDE for values in $[T_0, 2T_0]$ in a similar way as before. This simple time shift gives us all the necessary properties to prove pathwise uniqueness on the interval $[T_0, 2T_0]$, especially a version of Lemma 4.1. If we iterate this, we get pathwise uniqueness for the whole interval $[0, T]$.

A. Appendix

Throughout the thesis several minor results have been claimed and applied, but left unproved. Now, we catch up this. At first we show that all functions in the mixed-norm space $L^q_p(T)$, respectively $W^{1,2}_{q,p}(T)$, may be approximated by continuous, respectively smooth functions. These are well-known facts in the case of identical time and space integrability, but up to our best knowledge a rigorous proof has not been carried out so far for $p \neq q$. Therefore, we state all these more or less simple results and give the proofs in detail. Moreover, we provide the proofs of a mean-value inequality and a Sobolev embedding. In the end of this section we prove some Krylov-type estimates.

A.1. Approximation by continuous functions

In this section we establish the denseness of $\mathcal{C}_0$ in $L^q_p(T)$ through Lemma A.1, as well as the denseness of $\mathcal{C}^\infty$ in $W^{1,2}_{q,p}(T)$ through Lemma A.5. Lemmas A.3 and A.4 serve as auxiliary results for Lemma A.5 and they concern mollifiers (see Definition A.2 below) in $L^q_p(T)$ and $W^{1,2}_{q,p}(T)$ respectively. Finally Lemma A.6 reveals an approximation result on $\sigma^{(n)}$ which was applied in Proposition 4.6.

Lemma A.1. *$\mathcal{C}_0((0,T)\times\mathbb{R}^d)$ is dense in $L^q_p(T)$.*

Proof. Let $f \in L^q_p(T)$ and $\varepsilon > 0$ be arbitrary. Interpret f as a function from $(0,T)$ to $L^p(\mathbb{R}^d)$. Then $f \in L^q((0,T);L^p(\mathbb{R}^d))$ and therefore, there exist an $m \in \mathbb{N}$, disjoint sets $I_j \in \mathcal{B}((0,T))$, where $\mathcal{B}$ denotes the Borel σ-algebra, $g_j \in L^p(\mathbb{R}^d)$, $j = 1,...,m$ such that

$$\|f - \sum_{j=1}^m \mathbb{1}_{I_j} g_j\|_{L^q_p(T)} = \|f - \sum_{j=1}^m \mathbb{1}_{I_j} g_j\|_{L^q((0,T);L^p(\mathbb{R}^d))} < \frac{\varepsilon}{2}, \tag{25}$$

see e.g. [Alt16] Lemma 3.26. Since $\mathcal{C}_0(\mathbb{R}^d)$ is dense in $L^p(\mathbb{R}^d)$, for every $j \in \{1,...,m\}$ there exists a $\tilde{g}_j \in \mathcal{C}_0(\mathbb{R}^d)$ such that

$$\|\tilde{g}_j - g_j\|_{L^p(\mathbb{R}^d)} < \frac{\varepsilon}{4mT^q}.$$

And by Lusin's theorem, see e. g. Theorem 7.10 in [Fol99], for every $j \in \{1,...,m\}$ there exists $h_j \in \mathcal{C}_0((0,T))$ such that

$$\sup_{t\in(0,T)} |h_j(t)| \leq \sup_{t\in(0,T)} \mathbb{1}_{I_j}(t) \leq 1$$

and

$$dt(\{t \in \mathbb{R} : h_j(t) \neq \mathbb{1}_{I_j}(t)\}) < \left(\frac{\varepsilon}{8m\|\tilde{g}_j\|_{L^p(\mathbb{R}^d)}}\right)^q.$$

Define

$$\varphi(t,x) := \sum_{j=1}^m h_j(t)\tilde{g}_j(x).$$

Then we have $\varphi \in \mathcal{C}_0((0,T)\times\mathbb{R}^d)$ and

$$\begin{aligned}
\|\sum_{j=1}^m \mathbb{1}_{I_j} g_j - \varphi\|_{L_p^q(T)} &= \|\sum_{j=1}^m \mathbb{1}_{I_j} g_j - h_j \tilde{g}_j\|_{L_p^q(T)} \\
&\leq \sum_{j=1}^m \|\mathbb{1}_{I_j} g_j - h_j \tilde{g}_j\|_{L_p^q(T)} \\
&\leq \sum_{j=1}^m \|\mathbb{1}_{I_j} g_j - \mathbb{1}_{I_j}\tilde{g}_j\|_{L_p^q(T)} + \|\mathbb{1}_{I_j}\tilde{g}_j - h_j\tilde{g}_j\|_{L_p^q(T)} \\
&\leq \sum_{j=1}^m \|g_j - \tilde{g}_j\|_{L_p^q(T)} + \|\mathbb{1}_{I_j}\tilde{g}_j - h_j\tilde{g}_j\|_{L_p^q(T)} \\
&= \sum_{j=1}^m T^q \|g_j - \tilde{g}_j\|_{L^p(\mathbb{R}^d)} + \|\mathbb{1}_{I_j}\tilde{g}_j - h_j\tilde{g}_j\|_{L_p^q(T)} \\
&< \frac{\varepsilon}{4} + \sum_{j=1}^m \|\mathbb{1}_{I_j}\tilde{g}_j - h_j\tilde{g}_j\|_{L_p^q(T)} \\
&= \frac{\varepsilon}{4} + \sum_{j=1}^m \left(\int_0^T \left(\int_{\mathbb{R}^d} |\mathbb{1}_{I_j}(t) - h_j(t)|^p |\tilde{g}_j(x)|^p \, dx \right)^{\frac{q}{p}} dt \right)^{\frac{1}{q}} \\
&= \frac{\varepsilon}{4} + \sum_{j=1}^m \left(\int_0^T |\mathbb{1}_{I_j}(t) - h_j(t)|^q \|\tilde{g}_j\|^q_{L^p(\mathbb{R}^d)} \, dt \right)^{\frac{1}{q}} \\
&\leq \frac{\varepsilon}{4} + \sum_{j=1}^m 2\|\tilde{g}_j\|_{L^p(\mathbb{R}^d)} \, dt(\{t\in(0,T) : \mathbb{1}_{I_j}(t) \neq h_j(t)\})^{\frac{1}{q}} \\
&< \frac{\varepsilon}{2}.
\end{aligned}$$

Together with (25) this leads to

$$\|f - \varphi\|_{L_p^q(T)} < \varepsilon$$

and therefore, $\mathcal{C}_0((0,T)\times\mathbb{R}^d)$ is dense in $L_p^q(T)$. □

Definition A.2. *A nonnegative function $\varphi \in \mathcal{C}_0^\infty(\mathbb{R}^n)$ is called mollifier if*

(i) $\varphi(x) = 0$ *if* $|x| \geq 1$, *and*

(ii) $\int_{\mathbb{R}^n} \varphi(x)\, dx = 1$.

Let φ be a mollifier on $\mathbb{R}$, ψ a mollifier on $\mathbb{R}^d$. For $\varepsilon > 0$, $\delta > 0$ and a function f on $\mathbb{R}^{d+1}$, we define

$$\varphi_\varepsilon(t) := \varepsilon^{-1}\varphi(t/\varepsilon), \quad \psi_\delta(x) := \delta^{-d}\psi(x/\delta),$$

and

$$(\varphi_\varepsilon *_t f)(t,x) := \int_{\mathbb{R}} \varphi_\varepsilon(t-s) f(s,x)\, ds, \quad (\psi_\delta *_x f)(t,x) := \int_{\mathbb{R}^d} \psi_\delta(x-y) f(t,y)\, dy.$$

The next Lemma provides approximability of functions in $L_p^q(T)$ through mollification. It is in fact a generalization of Theorem 2.29 (c) from [AF09] for different integrability in time and space.

Lemma A.3. *Let f be a function on $\mathbb{R}^{d+1}$ that vanishes identically outside of $(0,T)\times\mathbb{R}^d$, φ be a mollifier on $\mathbb{R}$ and ψ a mollifier on $\mathbb{R}^d$. If $f \in L_p^q(T)$, then $\varphi_\varepsilon *_t \psi_\delta *_x f \in L_p^q(T)$,*

$$\|\varphi_\varepsilon *_t \psi_\delta *_x f\|_{L_p^q(T)} \leq \|f\|_{L_p^q(T)} \quad \text{and} \quad \lim_{\varepsilon,\delta\searrow 0} \|\varphi_\varepsilon *_t \psi_\delta *_x f - f\|_{L_p^q(T)} = 0.$$

Proof. We have, with Hölder's inequality and the fact that ψ is a mollifier

$$\begin{aligned}
&\|\varphi_\varepsilon *_t \psi_\delta *_x f\|_{L_p^q(T)} \\
&\leq \left(\int_0^T \left(\int_{\mathbb{R}^d} \left(\int_{\mathbb{R}^d} \psi_\delta(x-y) \int_{\mathbb{R}} \varphi_\varepsilon(t-s) |f(s,y)|\, ds\, dy \right)^p dx \right)^{\frac{q}{p}} dt \right)^{\frac{1}{q}} \\
&= \left(\int_0^T \left(\int_{\mathbb{R}^d} \left(\int_{\mathbb{R}^d} \psi_\delta(x-y)^{\frac{p-1}{p}} \psi_\delta(x-y)^{\frac{1}{p}} \int_{\mathbb{R}} \varphi_\varepsilon(t-s) |f(s,y)|\, ds\, dy \right)^p dx \right)^{\frac{q}{p}} dt \right)^{\frac{1}{q}} \\
&\leq \left(\int_0^T \left(\int_{\mathbb{R}^d} \left(\int_{\mathbb{R}^d} \psi_\delta(x-y)\, dy \right)^{p-1} \right.\right. \\
&\qquad \left.\left. \cdot \left(\int_{\mathbb{R}^d} \psi_\delta(x-y) \left(\int_{\mathbb{R}} \varphi_\varepsilon(t-s) |f(s,y)|\, ds \right)^p dy \right) dx \right)^{\frac{q}{p}} dt \right)^{\frac{1}{q}} \\
&= \left(\int_0^T \left(\int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \psi_\delta(x-y) \left(\int_{\mathbb{R}} \varphi_\varepsilon(t-s) |f(s,y)|\, ds \right)^p dy\, dx \right)^{\frac{q}{p}} dt \right)^{\frac{1}{q}} \\
&= \left(\int_0^T \left(\int_{\mathbb{R}^d} \left(\int_{\mathbb{R}} \varphi_\varepsilon(t-s) |f(s,y)|\, ds \right)^p \int_{\mathbb{R}^d} \psi_\delta(x-y)\, dx\, dy \right)^{\frac{q}{p}} dt \right)^{\frac{1}{q}} \\
&= \left(\int_0^T \left(\int_{\mathbb{R}^d} \left(\int_{\mathbb{R}} \varphi_\varepsilon(t-s) |f(s,y)|\, ds \right)^p dy \right)^{\frac{q}{p}} dt \right)^{\frac{1}{q}}.
\end{aligned} \tag{26}$$

Applying again Hölder's inequality and Fubini's theorem leads to

$$
\begin{aligned}
&\int\limits_{\mathbb{R}^d} \left(\int\limits_{\mathbb{R}} \varphi_\varepsilon(t-s)|f(s,y)|\,ds \right)^p dy \\
&\quad = \int\limits_{\mathbb{R}^d} \int\limits_{\mathbb{R}} \varphi_\varepsilon(t-r)|f(r,y)|\,dr \left(\int\limits_{\mathbb{R}} \varphi_\varepsilon(t-s)|f(s,y)|\,ds \right)^{p-1} dy \\
&\quad = \int\limits_{\mathbb{R}^d} \int\limits_{\mathbb{R}} \varphi_\varepsilon(t-r)|f(r,y)| \left(\int\limits_{\mathbb{R}} \varphi_\varepsilon(t-s)|f(s,y)|\,ds \right)^{p-1} dr\,dy \\
&\quad = \int\limits_{\mathbb{R}} \varphi_\varepsilon(t-r) \int\limits_{\mathbb{R}^d} |f(r,y)| \left(\int\limits_{\mathbb{R}} \varphi_\varepsilon(t-s)|f(s,y)|\,ds \right)^{p-1} dy\,dr \\
&\quad \le \int\limits_{\mathbb{R}} \varphi_\varepsilon(t-r) \left(\int\limits_{\mathbb{R}^d} |f(r,y)|^p\,dy \right)^{\frac{1}{p}} \left(\int\limits_{\mathbb{R}^d} \left(\int\limits_{\mathbb{R}} \varphi_\varepsilon(t-s)|f(s,y)|\,ds \right)^p dy \right)^{\frac{p-1}{p}} dr \\
&\quad = \left(\int\limits_{\mathbb{R}^d} \left(\int\limits_{\mathbb{R}} \varphi_\varepsilon(t-s)|f(s,y)|\,ds \right)^p dy \right)^{\frac{p-1}{p}} \int\limits_{\mathbb{R}} \varphi_\varepsilon(t-r) \left(\int\limits_{\mathbb{R}^d} |f(r,y)|^p\,dy \right)^{\frac{1}{p}} dr
\end{aligned}
$$

which is equivalent to

$$
\left(\int\limits_{\mathbb{R}^d} \left(\int\limits_{\mathbb{R}} \varphi_\varepsilon(t-s)|f(s,y)|\,ds \right)^p dy \right)^{\frac{1}{p}} \le \int\limits_{\mathbb{R}} \varphi_\varepsilon(t-s) \left(\int\limits_{\mathbb{R}^d} |f(s,y)|^p\,dy \right)^{\frac{1}{p}} ds.
$$

Using this in (26) leads to

$$
\begin{aligned}
&\|\varphi_\varepsilon *_t \psi_\delta *_x f\|_{L_p^q(T)} \\
&\quad \le \left(\int\limits_0^T \left(\int\limits_{\mathbb{R}} \varphi_\varepsilon(t-s) \left(\int\limits_{\mathbb{R}^d} |f(s,y)|^p\,dy \right)^{\frac{1}{p}} ds \right)^q dt \right)^{\frac{1}{q}} \\
&\quad = \left(\int\limits_0^T \left(\int\limits_{\mathbb{R}} \varphi_\varepsilon(t-s)^{\frac{q-1}{q}} \varphi_\varepsilon(t-s)^{\frac{1}{q}} \left(\int\limits_{\mathbb{R}^d} |f(s,y)|^p\,dy \right)^{\frac{1}{p}} ds \right)^q dt \right)^{\frac{1}{q}}.
\end{aligned}
$$

Using once more Hölder's inequality yields

$$\begin{aligned}
&\|\varphi_\varepsilon *_t \psi_\delta *_x f\|_{L_p^q(T)} \\
&\leq \left(\int_0^T \left(\int_{\mathbb{R}} \varphi_\varepsilon(t-s)\, ds \right)^{q-1} \left(\int_{\mathbb{R}} \varphi_\varepsilon(t-s) \left(\int_{\mathbb{R}^d} |f(s,y)|^p\, dy \right)^{\frac{q}{p}} ds \right) dt \right)^{\frac{1}{q}} \\
&= \left(\int_0^T \int_{\mathbb{R}} \varphi_\varepsilon(t-s) \left(\int_{\mathbb{R}^d} |f(s,y)|^p\, dy \right)^{\frac{q}{p}} ds\, dt \right)^{\frac{1}{q}} \\
&\leq \left(\int_{\mathbb{R}} \left(\int_{\mathbb{R}^d} |f(s,y)|^p\, dy \right)^{\frac{q}{p}} \int_{\mathbb{R}} \varphi_\varepsilon(t-s)\, dt\, ds \right)^{\frac{1}{q}} \\
&= \|f\|_{L_p^q(T)}.
\end{aligned}$$

Now, let $\eta > 0$ be given. By Lemma A.1 there exists $\phi \in \mathcal{C}_0((0,T) \times \mathbb{R}^d)$ such that

$$\|f - \phi\|_{L_p^q(T)} < \frac{\eta}{3}$$

and therefore also

$$\|\varphi_\varepsilon *_t \psi_\delta *_x f - \varphi_\varepsilon *_t \psi_\delta *_x \phi\|_{L_p^q(T)} < \frac{\eta}{3}.$$

Finally,

$$\begin{aligned}
&\|\varphi_\varepsilon *_t \psi_\delta *_x \phi - \phi\|_{L_p^q(T)} \\
&\leq \left(\int_0^T \left(\int_{\mathbb{R}^d} \left(\int_{\{s \in \mathbb{R} : |t-s| < \varepsilon\}} \varphi_\varepsilon(t-s) \right.\right.\right. \\
&\qquad \left.\left.\left. \cdot \int_{\{y \in \mathbb{R}^d : |x-y| < \delta\}} \psi(x-y) |\phi(s,y) - \phi(t,x)|\, dy\, ds \right)^p dx \right)^{\frac{q}{p}} dt \right)^{\frac{1}{q}},
\end{aligned}$$

where we may choose ε, δ so small that this is less than $\eta/3$. This is possible since ϕ is compactly supported and uniformly continuous. Then, by triangle inequality we have

$$\begin{aligned}
&\|\varphi_\varepsilon *_t \psi_\delta *_x f - f\|_{L_p^q(T)} \\
&\leq \|\varphi_\varepsilon *_t \psi_\delta *_x f - \varphi_\varepsilon *_t \psi_\delta *_t \phi\|_{L_p^q(T)} + \|\varphi_\varepsilon *_t \psi_\delta *_x \phi - \phi\|_{L_p^q(T)} + \|\phi - f\|_{L_p^q(T)} \\
&< \eta,
\end{aligned}$$

which proves the last claim. $\square$

Now we are able to receive approximation by mollification of functions in the mixed-norm Sobolev space on subsets of $[0,T]$. Again it is a generalization of a well-known result, see e. g. Lemma 3.16 in [AF09], allowing different integrability in time and space.

Lemma A.4. *Let φ be a mollifier on $\mathbb{R}$, ψ be a mollifier on $\mathbb{R}^d$, $f \in W^{1,2}_{q,p}(T)$. Then we have for all $0 < s < r < T$*

$$\lim_{\varepsilon,\delta \searrow 0} \|\varphi_\varepsilon *_t \psi_\delta *_x f - f\|_{W^{1,2}_{q,p}(s,r)} = 0,$$

where

$$\|f\|_{W^{1,2}_{q,p}(s,r)} = \|f\|_{L^q_p(s,r)} + \|\partial_t f\|_{L^q_p(s,r)} + \|\partial_x f\|_{L^q_p(s,r)} + \|\partial_x^2\|_{L^q_p(s,r)}$$

and

$$\|f\|_{L^q_p(s,r)} = \left(\int_s^r \left(\int_{\mathbb{R}^d} |f(t,x)|^p \, dx \right)^{\frac{q}{p}} dt \right)^{\frac{1}{q}}.$$

Proof. Let $\varepsilon < \min(s, T-r)$ and $\tilde{f} = \mathbb{1}_{[0,T]} f$. Then, we have for all test functions $\phi \in \mathcal{C}_0^\infty((s,r) \times \mathbb{R}^d)$

$$\begin{aligned}
&\int_s^r \int_{\mathbb{R}^d} \varphi_\varepsilon *_t \psi_\delta *_x f(t,x) \partial_t \phi(t,x) \, dx \, dt \\
&\quad = \int_s^r \int_{\mathbb{R}^d} \int_{\mathbb{R}} \varphi_\varepsilon(t') \int_{\mathbb{R}^d} \psi_\delta(y) f(t-t', x-y) \, dy \, dt' \partial_t \phi(t,x) \, dx \, dt \\
&\quad = \int_{\mathbb{R}} \int_{\mathbb{R}^d} \int_{\mathbb{R}} \varphi_\varepsilon(t') \int_{\mathbb{R}^d} \psi_\delta(y) \tilde{f}(t-t', x-y) \, dy \, dt' \partial_t \phi(t,x) \, dx \, dt \\
&\quad = \int_{\mathbb{R}} \varphi_\varepsilon(t') \int_{\mathbb{R}^d} \psi_\delta(y) \int_{\mathbb{R}} \int_{\mathbb{R}^d} \tilde{f}(t-t', x-y) \partial_t \phi(t,x) \, dx \, dt \, dy \, dt' \\
&\quad = -\int_{\mathbb{R}} \varphi_\varepsilon(t') \int_{\mathbb{R}^d} \psi_\delta(y) \int_{\mathbb{R}} \int_{\mathbb{R}^d} \partial_t \tilde{f}(t-t', x-y) \phi(t,x) \, dx \, dt \, dy \, dt' \\
&\quad = -\int_s^r \int_{\mathbb{R}^d} \varphi_\varepsilon *_t \psi_\delta *_x \partial_t f(t,x) \phi(t,x) \, dx \, dt.
\end{aligned}$$

And therefore $\partial_t \varphi_\varepsilon *_t \psi_\delta *_x f = \varphi_\varepsilon *_t \psi_\delta *_x \partial_t f$ in the weak sense on $(s,r) \times \mathbb{R}^d$. Analogous equalities hold true if we replace ∂_t by the derivatives with respect to space. Since $\partial_t f \in L^q_p(T)$ we obtain

$$\lim_{\varepsilon,\delta \searrow 0} \|\partial_t \varphi_\varepsilon *_t \psi_\delta *_x f - \partial_t f\|_{L^q_p(s,r)} = \lim_{\varepsilon,\delta \searrow 0} \|\varphi_\varepsilon *_t \psi_\delta *_x \partial_t f - \partial_t f\|_{L^q_p(s,r)} = 0$$

with Lemma A.3. And analogously for $\partial_x f, \partial_x^2 f$. $\square$

Lemma A.5. *Let $f \in W^{1,2}_{q,p}(T)$ and $\varepsilon > 0$ be arbitrary. Then there exists a function $g \in \mathcal{C}^\infty((0,T) \times \mathbb{R}^d)$ such that*

$$\|f - g\|_{W^{1,2}_{q,p}(T)} < \varepsilon.$$

The proof follows ideas from the proof of [AF09] Theorem 3.17.

Proof. Let $f \in W^{1,2}_{q,p}(T)$. For $k = 1, 2, ...$ define

$$I_k := \left\{ t \in (0,T) : t < k,\ \frac{1}{k} < t < T - \frac{1}{k} \right\}$$

and $I_0 = I_{-1} = \emptyset$. Furthermore, set

$$U_k := I_{k+1} \setminus \overline{I_{k-1}}.$$

Then $(U_k)_k$ is a locally finite open cover of $(0,T)$ and there exists a partition of unity $(\phi_k)_k$ (see e.g. [Alt16] 4.20), i.e.

$$\phi_k \in \mathcal{C}_0^\infty(U_k), \quad \phi_k \geq 0, \quad \text{and} \quad \sum_{k=1}^{\infty} \phi_k(t) = 1 \ \ \forall\, t \in (0,T).$$

Where locally only finitely many terms in the sum are nonzero. If $0 < \varepsilon < \frac{1}{(k+1)(k+2)}$, then $\varphi_\varepsilon *_t \psi_\delta *_x (\phi_k f)$ has support in

$$V_k = (I_{k+2} \setminus \overline{I_{k-2}}) \times \mathbb{R}^d.$$

Since $\phi_k f \in W^{1,2}_{q,p}(V_k)$ we may choose ε_k, satisfying $0 < \varepsilon_k < \frac{1}{(k+1)(k+2)}$ and $\delta_k > 0$ such that by Lemma A.4

$$\|\varphi_{\varepsilon_k} *_t \psi_{\delta_k} *_x (\phi_k f) - \phi_k f\|_{W^{1,2}_{q,p}(T)} = \|\varphi_{\varepsilon_k} *_t \psi_{\delta_k} *_x (\phi_k f) - \phi_k f\|_{W^{1,2}_{q,p}(V_k)} < \frac{\varepsilon}{2^k}.$$

Define

$$g := \sum_{k=1}^{\infty} \varphi_{\varepsilon_k} *_t \psi_{\delta_k} *_x (\phi_k f).$$

Since locally only finitely many terms in the sum can be nonzero, $g \in \mathcal{C}^\infty((0,T) \times \mathbb{R}^d)$. For $t \in I_k$, $x \in \mathbb{R}^d$ we have

$$f(t,x) = \sum_{j=1}^{k+2} \phi_j(t) f(t,x) \quad \text{and} \quad g(t,x) = \sum_{j=1}^{k+2} \varphi_{\varepsilon_j} *_t \psi_{\delta_j} *_x (\phi_j f)(t,x).$$

Thus

$$\|f - g\|_{W^{1,2}_{q,p}(I_k)} \leq \sum_{j=1}^{k+2} \|\varphi_{\varepsilon_j} *_t \psi_{\delta_j} *_x (\phi_j f) - \phi_j f\|_{W^{1,2}_{q,p}(T)} < \varepsilon,$$

which leads to

$$\|f - g\|_{W^{1,2}_{q,p}(T)} = \lim_{k \to \infty} \|f - g\|_{W^{1,2}_{q,p}(I_k)} < \varepsilon.$$

$\square$

Lemma A.6. *Let $f : [0,T] \times \mathbb{R}^d \to \mathbb{R}^d$ (or $\mathbb{R}^{d\times m}$) be continuous and differentiable with respect to x in the weak sense, ψ be a mollifier on $\mathbb{R}^d$. Then we have*

(i) $\psi_\varepsilon *_x f \to f$ *for* $\varepsilon \searrow 0$ *pointwise on* $[0,T] \times \mathbb{R}^d$,

(ii) $\psi_\varepsilon *_x f \to f$ *for* $\varepsilon \searrow 0$ *uniformly on* $[0,T] \times \overline{B_R}$ *for all* $R > 0$,

(iii) $\psi_\varepsilon *_x f$ *is differentiable with respect to x in the ordinary sense, and* $\partial_x(\psi_\varepsilon *_x f) = \partial_x\psi_\varepsilon *_x f = \psi_\varepsilon *_x \partial_x f$,

(iv) $\|\partial_x(\psi_\varepsilon *_x f)\|_{L^r_v(T)} \leq \|\partial_x f\|_{L^r_v(T)}$ *for all* $\varepsilon > 0$, $r, v > 1$.

Proof. (i) Let $(t,x) \in [0,T] \times \mathbb{R}^d$, $\varepsilon > 0$ be arbitrary. By continuity of f there exists a $\delta > 0$ such that

$$|f(t,x) - f(t,y)| < \varepsilon \ \ \forall\, y \in \mathbb{R}^d \text{ with } |x-y| < \delta.$$

Therefore, we have for all $\eta \leq \delta$

$$\begin{aligned}
|(\psi_\eta *_x f)(t,x) - f(t,x)| &= \left| \int\limits_{\mathbb{R}^d} \psi_\eta(x-y) f(t,y)\,dy - f(t,x) \int\limits_{\mathbb{R}^d} \psi_\eta(x-y)\,dy \right| \\
&\leq \int\limits_{\mathbb{R}^d} \psi_\eta(x-y) |f(t,y) - f(t,x)|\,dy \\
&= \int\limits_{\{y\in\mathbb{R}^d : |x-y|<\eta\}} \psi_\eta(x-y) |f(t,y) - f(t,x)|\,dy \\
&< \varepsilon \int\limits_{\{y\in\mathbb{R}^d : |x-y|<\eta\}} \psi_\eta(x-y)\,dy \\
&= \varepsilon,
\end{aligned}$$

which means pointwise convergence on $[0,T] \times \mathbb{R}^d$.

(ii) f is uniformly continuous on $[0,T] \times \overline{B_{R+1}}$ and therefore, for arbitrary $\varepsilon > 0$, there is a $0 < \delta < 1$ such that

$$|f(t,x) - f(s,y)| < \varepsilon \ \ \forall\, (t,x), (s,y) \in [0,T] \times \overline{B_{R+1}} \text{ with } |(t,x) - (s,y)| < \delta.$$

Then we get for all $(t,x) \in [0,T] \times \overline{B_R}$ and all $\eta \leq \delta$

$$|(\psi_\eta *_x f)(t,x) - f(t,x)| \leq \int\limits_{\{y\in\mathbb{R}^d : |x-y|<\eta\}} \psi_\eta(x-y) |f(t,x) - f(t,y)|\,dy.$$

Since $\eta < 1$, the domain of integration is included in $\overline{B_{R+1}}$, so

$$|(\psi_\eta *_x f)(t,x) - f(t,x)| < \varepsilon,$$

which proves uniform convergence on $[0,T] \times \overline{B_R}$.

(iii) Clearly, we have

$$\begin{aligned}\partial_x(\psi_\varepsilon *_x f)(t,x) &= \partial_x \int\limits_{\mathbb{R}^d} \psi_\varepsilon(x-y) f(t,y)\,dy \\ &= \int\limits_{\mathbb{R}^d} \partial_x \psi_\varepsilon(x-y) f(t,y)\,dy \\ &= (\partial_x \psi_\varepsilon) *_x f(t,x)\end{aligned}$$

and for all test functions $\phi \in \mathcal{C}_0^\infty((0,T)\times\mathbb{R}^d)$

$$\begin{aligned}&\int\limits_0^T \int\limits_{\mathbb{R}^d} \partial_x \phi(t,x)(\psi_\varepsilon *_x f)(t,x)\,dx\,dt \\ &= \int\limits_0^T \int\limits_{\mathbb{R}^d} \partial_x \phi(t,x) \int\limits_{\mathbb{R}^d} \psi_\varepsilon(y) f(t,x-y)\,dy\,dx\,dt \\ &= \int\limits_{\mathbb{R}^d} \psi_\varepsilon(y) \int\limits_0^T \int\limits_{\mathbb{R}^d} \partial_x \phi(t,x) f(t,x-y)\,dx\,dt\,dy \\ &= (-1)^d \int\limits_{\mathbb{R}^d} \psi_\varepsilon(y) \int\limits_0^T \int\limits_{\mathbb{R}^d} \phi(t,x)\partial_x f(t,x-y)\,dx\,dt\,dy \\ &= (-1)^d \int\limits_0^T \int\limits_{\mathbb{R}^d} \phi(t,x) \int\limits_{\mathbb{R}^d} \psi_\varepsilon(y)\partial_x f(t,x-y)\,dy\,dx\,dt \\ &= (-1)^d \int\limits_0^T \int\limits_{\mathbb{R}^d} \phi(t,x)(\psi_\varepsilon *_x \partial_x f)(t,x)\,dx\,dt\end{aligned}$$

and therefore, $\partial_x\psi_\varepsilon *_x f = \partial_x(\psi_\varepsilon *_x f) = \psi_\varepsilon *_x \partial_x f$.

(iv) We have

$$\begin{aligned}&\|\partial_x(\psi_\varepsilon *_x f)\|_{L_v^r(T)} \\ &= \|\psi_\varepsilon *_x (\partial_x f)\|_{L_v^r(T)} \\ &= \left(\int\limits_0^T \left(\int\limits_{\mathbb{R}^d} \left| \int\limits_{\mathbb{R}^d} \psi_\varepsilon(x-y)\partial_x f(t,y)\,dy \right|^v dx \right)^{\frac{r}{v}} dt \right)^{\frac{1}{r}} \\ &\le \left(\int\limits_0^T \left(\int\limits_{\mathbb{R}^d} \left(\int\limits_{\mathbb{R}^d} \psi_\varepsilon(x-y)^{\frac{v-1}{v}} \psi_\varepsilon(x-y)^{\frac{1}{v}} \left|\partial_x f(t,y)\right|\,dy \right)^v dx \right)^{\frac{r}{v}} dt \right)^{\frac{1}{r}}.\end{aligned}$$

An application of Hölder's inequality yields

$$\|\partial_x(\psi_\varepsilon *_x f)\|_{L^r_v(T)}$$
$$\leq \left(\int_0^T \left(\int_{\mathbb{R}^d} \left(\int_{\mathbb{R}^d} \psi_\varepsilon(x-y)\,dy \right)^{v-1} \cdot \left(\int_{\mathbb{R}^d} \psi_\varepsilon(x-y)\,|\partial_x f(t,y)|^v \,dy \right) dx \right)^{\frac{r}{v}} dt \right)^{\frac{1}{r}}$$

and since ψ is a mollifier, we obtain

$$\begin{aligned} \|\partial_x(\psi_\varepsilon *_x f)\|_{L^r_v(T)} &\leq \left(\int_0^T \left(\int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \psi_\varepsilon(x-y)\,|\partial_x f(t,y)|^v \,dy\,dx \right)^{\frac{r}{v}} dt \right)^{\frac{1}{r}} \\ &= \left(\int_0^T \left(\int_{\mathbb{R}^d} |\partial_x f(t,y)|^v \int_{\mathbb{R}^d} \psi_\varepsilon(x-y)\,dx\,dy \right)^{\frac{r}{v}} dt \right)^{\frac{1}{r}} \\ &= \left(\int_0^T \left(\int_{\mathbb{R}^d} |\partial_x f(t,y)|^v \,dy \right)^{\frac{r}{v}} dt \right)^{\frac{1}{r}} \\ &= \|\partial_x f\|_{L^r_v(T)}. \end{aligned}$$

□

A.2. Mean-value inequality for weakly differentiable functions and Sobolev embedding

In this section we first prove a simple mean-value inequality for weakly differentiable functions stated as Lemma A.7. The second part consists of a Sobolev embedding for $W^{1,2}_{q,p} \to \mathcal{C}_b$, which is a straight forward generalization of a known result, see e.g. [AF09] Theorem 4.12. Since we consider different integrability and also different order of integration in time and space, we give the proof in detail although a few modifications are sufficient.

Lemma A.7. *For $u \in W^{1,2}_{q,p}(T)$, $t \in [0,T]$, $x, y \in \mathbb{R}^d$, we have*

$$|u(t,x) - u(t,y)| \leq |x-y| \sup_{z \in \mathbb{R}^d} |\partial_x u(t,z)|.$$

Proof. Let ψ be a mollifier on $\mathbb{R}^d$. Since u is continuous, see [KR05] Lemma 10.2, we have for all $t \in [0,T]$, $x, y \in \mathbb{R}^d$ with Lemma A.6

$$\begin{aligned}
&|u(t,x) - u(t,y)| \\
&\quad = \lim_{\varepsilon \searrow 0} |\psi_\varepsilon *_x u(t,x) - \psi_\varepsilon *_x u(t,y)| \\
&\quad = \lim_{\varepsilon \searrow 0} \left| \int_0^1 \partial_x (\psi_\varepsilon *_x u)(t, \lambda x + (1-\lambda) y)(x-y)\, d\lambda \right| \\
&\quad = \lim_{\varepsilon \searrow 0} \left| \int_0^1 \psi_\varepsilon *_x \partial_x u(t, \lambda x + (1-\lambda) y)(x-y)\, d\lambda \right| \\
&\quad = \lim_{\varepsilon \searrow 0} \left| (x-y) \int_0^1 \int_{\mathbb{R}^d} \psi_\varepsilon(\lambda x + (1-\lambda) y - z) \partial_x u(t,z)\, dz\, d\lambda \right| \\
&\quad \le |x-y| \sup_{z \in \mathbb{R}^d} |\partial_x u(t,z)| \lim_{\varepsilon \searrow 0} \int_0^1 \int_{\mathbb{R}^d} \psi_\varepsilon(\lambda x + (1-\lambda) y - z)\, dz\, d\lambda \\
&\quad = |x-y| \sup_{z \in \mathbb{R}^d} |\partial_x u(t,z)|.
\end{aligned}$$

□

Lemma 3.2. *For all $u \in W^{1,2}_{q,p}(T)$, there exists a version of u such that*

$$\sup_{(t,x) \in [0,T] \times \mathbb{R}^d} |u(t,x)| \le C \|u\|_{W^{1,2}_{q,p}(T)},$$

where C is independent of u. In particular this version is continuous.

Since the lemma is a simple generalization of [AF09] Theorem 4.12 Part I Case A (1), we follow the ideas of the proof therein.

Proof. First note, that $[0,T] \times \mathbb{R}^d$ fulfills the so called *cone condition*, i.e. there exists a finite cone

$$K = \left\{ z \in \mathbb{R}^{d+1} : z = 0 \text{ or } 0 < |z| \le \rho,\ \measuredangle(z,w) \le \frac{\kappa}{2} \right\}$$

with height $\rho > 0$, axis direction $w \in \mathbb{R}^{d+1}_+$, $w \neq 0$ and aperture angle $0 < \kappa \le \pi$, $\measuredangle(z,w)$ denotes here the angle between z and w, such that each $(t,x) \in [0,T] \times \mathbb{R}^d$ is the vertex of a finite cone $K_{(t,x)}$ contained in $[0,T] \times \mathbb{R}^d$ and congruent to K. Note that $K_{(t,x)}$ need not be obtained from K by parallel translation, but simply by rigid motion. Let $u \in \mathcal{C}^\infty([0,T] \times \mathbb{R}^d) \cap W^{1,2}_{q,p}(T)$, $(t,x) \in [0,T] \times \mathbb{R}^d$, $0 < r \le \rho$ and $(s,y) \in K_{(t,x),r}$, where

$$K_{(t,x),r} := \left\{ (s,y) \in K_{(t,x)} \;\middle|\; |(t,x) - (s,y)| \le r \right\}.$$

Define

$$f(\lambda) := u(\lambda(t,x) + (1-\lambda)(s,y)), \qquad \lambda \in [0,1],$$

then we have

$$f(1) = f(0) + \int_0^1 f'(\lambda)\, d\lambda.$$

With

$$\begin{aligned} f'(\lambda) &= \frac{\partial}{\partial\lambda} u(\lambda(t,x) + (1-\lambda)(s,y)) \\ &= Du(\lambda(t,x) + (1-\lambda)(s,y)) \cdot ((t,x) - (s,y)), \end{aligned}$$

where Du denotes the $d \times (d+1)$ matrix $(\partial_t u, \partial_x u)$, we have therefore,

$$u(t,x) = u(s,y) + ((t,x) - (s,y)) \int_0^1 Du(\lambda(t,x) + (1-\lambda)(s,y))\, d\lambda.$$

Applying triangle inequality yields then

$$|u(t,x)| \leq |u(s,y)| + |(t,x) - (s,y)| \int_0^1 |Du(\lambda(t,x) + (1-\lambda)(s,y))|\, d\lambda.$$

If K has the volume $c\rho^{d+1}$, $K_{(t,x),r}$ has the volume cr^{d+1} which by integrating reveals

$$\begin{aligned} cr^{d+1}|u(t,x)| \leq &\int_{K_{(t,x),r}} |u(s,y)|\, d(s,y) \\ &+ \int_{K_{(t,x),r}} |(t,x) - (s,y)| \int_0^1 |Du(\lambda(t,x) + (1-\lambda)(s,y))|\, d\lambda\, d(s,y). \end{aligned} \tag{27}$$

By Fubini's theorem the second integral term may be rewritten as

$$\begin{aligned} &\int_{K_{(t,x),r}} |(t,x) - (s,y)| \int_0^1 |Du(\lambda(t,x) + (1-\lambda)(s,y))|\, d\lambda\, d(s,y) \\ &\quad = \int_0^1 \int_{K_{(t,x),r}} |(t,x) - (s,y)|\, |Du(\lambda(t,x) + (1-\lambda)(s,y))|\, d(s,y)\, d\lambda. \end{aligned}$$

Substituting $z = \lambda(t,x) + (1-\lambda)(s,y)$ leads to

$$\begin{aligned} &\int_{K_{(t,x),r}} |(t,x) - (s,y)| \int_0^1 |Du(\lambda(t,x) + (1-\lambda)(s,y))|\, d\lambda\, d(s,y) \\ &\quad = \int_0^1 \int_{K_{(t,x),(1-\lambda)r}} (1-\lambda)^{-(d+2)} |(t,x) - z|\, |Du(z)|\, dz\, d\lambda. \end{aligned}$$

One more application of Fubini's theorem provides

$$\begin{aligned}
&\int\limits_{K_{(t,x),r}} |(t,x)-(s,y)| \int\limits_0^1 |Du(\lambda(t,x)+(1-\lambda)(s,y))| \, d\lambda \, d(s,y) \\
&\quad = \int_{K_{(t,x),r}} \int\limits_0^{1-\frac{|(t,x)-z|}{r}} (1-\lambda)^{-(d+2)} |(t,x)-z| \, |Du(z)| \, d\lambda \, dz \\
&\quad = \int\limits_{K_{(t,x),r}} |(t,x)-z| \, |Du(z)| \int\limits_0^{1-\frac{|(t,x)-z|}{r}} (1-\lambda)^{-(d+2)} \, d\lambda \, dz \\
&\quad = \int\limits_{K_{(t,x),r}} |(t,x)-z| \, |Du(z)| \left[\frac{1}{d+1} (1-\lambda)^{-(d+1)} \right]_0^{1-\frac{|(t,x)-z|}{r}} dz \\
&\quad = \int\limits_{K_{(t,x),r}} |(t,x)-z| \, |Du(z)| \frac{1}{d+1} \left(\frac{r^{d+1}}{|(t,x)-z|^{d+1}} - 1 \right) dz \\
&\quad \le \int\limits_{K_{(t,x),r}} \frac{r^{d+1}}{d+1} |(t,x)-z|^{-d} |Du(z)| \, dz.
\end{aligned}$$

Therefore, with (27) we get

$$|u(t,x)| \le \frac{1}{cr^{d+1}} \int\limits_{K_{(t,x),r}} |u(z)| \, dz + \frac{1}{c(d+1)} \int\limits_{K_{(t,x),r}} |(t,x)-z|^{-d} |Du(z)| \, dz. \tag{28}$$

With Hölder's inequality we deduce for the first integral

$$\begin{aligned}
\int\limits_{K_{(t,x),r}} |u(z)| \, dz &= \int\limits_0^T \int\limits_{\{y \in \mathbb{R}^d : (s,y) \in K_{(t,x),r}\}} |u(s,y)| \, dy \, ds \\
&\le \int\limits_0^T \left(\int\limits_{\{y \in \mathbb{R}^d : (s,y) \in K_{(t,x),r}\}} |u(s,y)|^p \, dy \right)^{\frac{1}{p}} \left(\int\limits_{\{y \in \mathbb{R}^d : (s,y) \in K_{(t,x),r}\}} 1 \, dy \right)^{\frac{p-1}{p}} ds \\
&\le \int\limits_0^T \left(\int\limits_{\mathbb{R}^d} |u(s,y)|^p \, dy \right)^{\frac{1}{p}} \left(\int\limits_{\{y \in \mathbb{R}^d : |x-y| \le r\}} 1 \, dy \right)^{\frac{p-1}{p}} ds \\
&\le (2r)^{\frac{d(p-1)}{p}} \int\limits_0^T \left(\int\limits_{\mathbb{R}^d} |u(s,y)|^p \, dy \right)^{\frac{1}{p}} ds,
\end{aligned}$$

just by estimating the volume of B_r by $(2r)^d$. Using once more Hölder's inequality leads to

$$\int\limits_{K_{(t,x),r}} |u(z)|\,dz \leq (2r)^{\frac{d(p-1)}{p}} T^{\frac{q-1}{q}} \left(\int\limits_0^T \left(\int\limits_{\mathbb{R}^d} |u(s,y)|^p\,dy \right)^{\frac{q}{p}} ds \right)^{\frac{1}{q}}$$
$$= (2r)^{\frac{d(p-1)}{p}} T^{\frac{q-1}{q}} \|u\|_{L_p^q(T)}. \tag{29}$$

For the second integral in (28) we have

$$\begin{aligned}
&\int\limits_{K_{(t,x),r}} |(t,x)-z|^{-d}|Du(z)|\,dz \\
&= \int\limits_0^T \int\limits_{\{y\in\mathbb{R}^d:(s,y)\in K_{(t,x),r}\}} |(t,x)-(s,y)|^{-d}|(\partial_t u(s,y),\partial_x u(s,y))|\,dy\,ds \\
&\leq \int\limits_0^T \left(\int\limits_{\{y\in\mathbb{R}^d:(s,y)\in K_{(t,x),r}\}} |(\partial_t u(s,y),\partial_x u(s,y))|^p\,dy \right)^{\frac{1}{p}} \\
&\qquad \cdot \left(\int\limits_{\{y\in\mathbb{R}^d:(s,y)\in K_{(t,x),r}\}} |(t,x)-(s,y)|^{\frac{-dp}{p-1}}\,dy \right)^{\frac{p-1}{p}} ds \\
&\leq \left(\int\limits_0^T \left(\int\limits_{\{y\in\mathbb{R}^d:(s,y)\in K_{(t,x),r}\}} |(\partial_t u(s,y),\partial_x u(s,y))|^p\,dy \right)^{\frac{q}{p}} ds \right)^{\frac{1}{q}} \\
&\qquad \cdot \left(\int\limits_0^T \left(\int\limits_{\{y\in\mathbb{R}^d:(s,y)\in K_{(t,x),r}\}} |(t,x)-(s,y)|^{\frac{-dp}{p-1}}\,dy \right)^{\frac{(p-1)q}{p(q-1)}} ds \right)^{\frac{q-1}{q}} \\
&\leq \left(\int\limits_0^T \left(\int\limits_{\{y\in\mathbb{R}^d:(s,y)\in K_{(t,x),r}\}} \big||\partial_t u(s,y)|+|\partial_x u(s,y)|\big|^p\,dy \right)^{\frac{q}{p}} ds \right)^{\frac{1}{q}} \\
&\qquad \cdot \left(\int\limits_0^T \left(\int\limits_{\{y\in\mathbb{R}^d:(s,y)\in K_{(t,x),r}\}} |(t,x)-(s,y)|^{\frac{-dp}{p-1}}\,dy \right)^{\frac{(p-1)q}{p(q-1)}} ds \right)^{\frac{q-1}{q}}.
\end{aligned}$$

Using Minkowski's inequality twice for the first integral yields

$$
\begin{aligned}
&\int\limits_{K_{(t,x),r}} |(t,x)-z|^{-d}|Du(z)|\,dz \\
&\qquad \le \left(\|\partial_t u\|_{L_p^q(T)} + \|\partial_x u\|_{L_p^q(T)}\right) \qquad (30) \\
&\qquad\qquad \cdot \left(\int\limits_0^T \left(\int\limits_{\{y\in\mathbb{R}^d:(s,y)\in K_{(t,x),r}\}} |(t,x)-(s,y)|^{\frac{-dp}{p-1}}\,dy\right)^{\frac{(p-1)q}{p(q-1)}} ds\right)^{\frac{q-1}{q}}.
\end{aligned}
$$

The integral does not depend u and on (t,x) since we only integrate distances. Furthermore, it is finite because of

$$
\begin{aligned}
&\left(\int\limits_0^T \left(\int\limits_{\{y\in\mathbb{R}^d:(s,y)\in K_{(t,x),r}\}} |(t,x)-(s,y)|^{\frac{-dp}{p-1}}\,dy\right)^{\frac{(p-1)q}{p(q-1)}} ds\right)^{\frac{q-1}{q}} \\
&= \left(\int\limits_0^T \left(\int\limits_{\{y\in\mathbb{R}^d:(s,y)\in K_{(t,x),r}\}} \left((t-s)^2+\sum_{i=1}^d (x_i-y_i)^2\right)^{\frac{-dp}{2(p-1)}} dy\right)^{\frac{(p-1)q}{p(q-1)}} ds\right)^{\frac{q-1}{q}} \\
&= \left(\int\limits_0^T \left(\int\limits_{\{y\in\mathbb{R}^d:(s,y)\in K_{(t,x),r}\}} \left((t-s)^2+\sum_{i=1}^d (x_i-y_i)^2\right)^{\frac{-dp}{2(d+1)(p-1)}}\right.\right. \\
&\qquad\qquad \left.\left.\cdot \left((t-s)^2+\sum_{i=1}^d (x_i-y_i)^2\right)^{\frac{-d^2p}{2(d+1)(p-1)}} dy\right)^{\frac{(p-1)q}{p(q-1)}} ds\right)^{\frac{q-1}{q}} \\
&\le \left(\int\limits_0^T \left(\int\limits_{\{y\in\mathbb{R}^d:(s,y)\in K_{(t,x),r}\}} |t-s|^{\frac{-dp}{(d+1)(p-1)}} \cdot |x-y|^{\frac{-d^2p}{(d+1)(p-1)}}\,dy\right)^{\frac{(p-1)q}{p(q-1)}} ds\right)^{\frac{q-1}{q}} \\
&= \left(\int\limits_0^T |t-s|^{\frac{-dq}{(d+1)(q-1)}} \left(\int\limits_{\{y\in\mathbb{R}^d:(s,y)\in K_{(t,x),r}\}} |x-y|^{\frac{-d^2p}{(d+1)(p-1)}}\,dy\right)^{\frac{(p-1)q}{p(q-1)}} ds\right)^{\frac{q-1}{q}}
\end{aligned}
$$

and

$$
\frac{dq}{(d+1)(q-1)} < 1 \quad \text{and} \quad \frac{d^2p}{(d+1)(p-1)} < d.
$$

So, with (28), (29) and (30) we have

$$\begin{aligned}|u(t,x)| &\leq C\left(\|u\|_{L^q_p(T)} + \|\partial_t u\|_{L^q_p(T)} + \|\partial_x u\|_{L^q_p(T)}\right)\\ &\leq C\|u\|_{W^{1,2}_{q,p}(T)}\end{aligned}$$

for all $u \in \mathcal{C}^\infty([0,T]\times\mathbb{R}^d)\cap W^{1,2}_{q,p}(T)$ and $(t,x)\in[0,T]\times\mathbb{R}^d$, where C does not depend on u. Now, let $u \in W^{1,2}_{q,p}(T)$ and $(u_n)_n$ be a sequence of smooth functions with

$$\lim_{n\to\infty}\|u-u_n\|_{W^{1,2}_{q,p}(T)} = 0.$$

For all $(t,x)\in[0,T]\times\mathbb{R}^d$ we have that $(u_n(t,x))_n$ is a Cauchy sequence in $\mathbb{R}^d$, because of

$$|u_n(t,x)-u_m(t,x)| \leq C\|u_n-u_m\|_{W^{1,2}_{q,p}(T)}.$$

Therefore $(u_n(t,x))_n$ converges for all $(t,x)\in[0,T]\times\mathbb{R}^d$ and we can define

$$\tilde{u}(t,x) := \lim_{n\to\infty} u_n(t,x).$$

Since $(u_n)_n$ converges uniformly to $\tilde{u}$, $\tilde{u}$ is continuous. Furthermore we have

$$|\tilde{u}(t,x)| = \lim_{n\to\infty}|u_n(t,x)| \leq \lim_{n\to\infty} C\|u_n\|_{W^{1,2}_{q,p}(T)} \leq C\|u\|_{W^{1,2}_{q,p}(T)}$$

and

$$\begin{aligned}\|u-\tilde{u}\|_{L^q_p(T)} &= \left(\int_0^T\left(\int_{\mathbb{R}^d}|u(t,x)-\tilde{u}(t,x)|^p\,dx\right)^{\frac{q}{p}}dt\right)^{\frac{1}{q}}\\ &= \left(\int_0^T\left(\int_{\mathbb{R}^d}\liminf_{n\to\infty}|u(t,x)-u_n(t,x)|^p\,dx\right)^{\frac{q}{p}}dt\right)^{\frac{1}{q}}\\ &\leq \liminf_{n\to\infty}\|u-u_n\|_{L^q_p(T)}\\ &= 0\end{aligned}$$

and therefore $\tilde{u}$ is a version of u with the claimed properties. □

A.3. Krylov-type estimates

In this section we prove Lemma 3.1 which was needed several times, e. g. to show that Itô's formula is applicable to functions in $W^{1,2}_{q,p}(T)$. To this end, we first generalize Theorem 3.2.4 from [Kry87], see Lemma A.8, to different integrability in time and space. In fact, the original proof remains principally intact except small modifications. Then we prove a version of Lemma 5.1 of [Kry86] for functions in $L^q_p(T)$ statet as Lemma A.10. Further, Lemma A.11 provides useful estimates on terms arising through Krylov's estimate. This finally enables us to prove Lemma 3.1.

Lemma A.8. *Let $f : \mathbb{R}^{d+1} \to \mathbb{R}$ be a nonnegative function which is equal to zero if $|t| \geq T$ or $|x| > R$ for some constants T and R. Suppose that f and all its derivatives with respect to t and x up to the second order are bounded and uniformly continuous. Then there exists a nonnegative function $\varphi : \mathbb{R}^{d+1} \to \mathbb{R}$ such that all weak derivatives $\partial_t\varphi, \partial_x\varphi, \partial_x^2\varphi$ exist, are bounded on $\mathbb{R}^{d+1}$ and for any symmetric, positive semidefinite $d \times d$ matrix α and arbitrary $\beta \geq 0$, $\mu > 0$ and $v, r \geq d+1$ we have*

(i) $\beta\partial_t\varphi + \sum_{i=1}^{d}\sum_{j=1}^{d} \alpha_{ij}\partial^2_{x_ix_j}\varphi - \mu(\beta + \operatorname{tr}(\alpha))\varphi + (\beta\det(\alpha))^{\frac{1}{d+1}} f \leq 0,$

(ii) $|\partial_x\varphi| \leq \sqrt{\mu}\varphi,$

(iii) $\varphi(t,x) \leq C(d,v)\mu^{\frac{d}{2v}-\frac{d}{d+1}}(T-t)^{\frac{1}{d+1}-\frac{1}{r}} \left(\int_0^\infty e^{-\mu r s} \left(\int_{\mathbb{R}^d} |f(t+s,x)|^v \, dx \right)^{\frac{r}{v}} ds \right)^{\frac{1}{r}}.$

Proof. [Kry87] Theorem 3.2.4 covers the existence of a function φ fulfilling (i) and (ii). Further, from the proof of this theorem we carry over the estimates

$$\kappa_d d!(v\sqrt{\mu})^{-d}(\varphi(t,x))^v \leq \int_{\mathbb{R}^d} |\varphi(t,x)|^v \, dx \tag{31}$$

and

$$\left(e^{-\mu t} \left(\int_{\mathbb{R}^d} |\varphi(t,x)|^v \, dx \right)^{\frac{1}{v}} \right)^{d+1} \leq \mu^{-d}(d+1)^{-d} \int_t^\infty \left(e^{-\mu s} \left(\int_{\mathbb{R}^d} |f(s,x)|^v \, dx \right)^{\frac{1}{v}} \right)^{d+1} ds \tag{32}$$

for every $v > d+1$ and $t \in (-\infty, \infty)$, where κ_d is the volume of a ball in $\mathbb{R}^d$ with radius 1. (31) implies that

$$\begin{aligned} e^{-\mu v t}\kappa_d d!(v\sqrt{\mu})^{-d}(\varphi(t,x))^v &\leq e^{-\mu v t} \int_{\mathbb{R}^d} |\varphi(t,x)|^v \, dx \\ &= \left(\left(e^{-\mu t} \left(\int_{\mathbb{R}^d} |\varphi(t,x)|^v \, dx \right)^{\frac{1}{v}} \right)^{d+1} \right)^{\frac{v}{d+1}}, \end{aligned}$$

and (32) that

$$\begin{aligned} &e^{-\mu vt}\kappa_d d!(v\sqrt{\mu})^{-d}(\varphi(t,x))^v \\ &\qquad \le (\mu(d+1))^{-\frac{dv}{d+1}} \left(\int\limits_t^\infty \left(e^{-\mu s} \left(\int\limits_{\mathbb{R}^d} |f(s,x)|^v \, dx \right)^{\frac{1}{v}} \right)^{d+1} ds \right)^{\frac{v}{d+1}} . \end{aligned}$$

Since f is zero on $[T,\infty)$, we have

$$\begin{aligned} &e^{-\mu vt}\kappa_d d!(v\sqrt{\mu})^{-d}(\varphi(t,x))^v \\ &\qquad \le (\mu(d+1))^{-\frac{dv}{d+1}} \left(\int\limits_t^T \left(e^{-\mu s} \left(\int\limits_{\mathbb{R}^d} |f(s,x)|^v \, dx \right)^{\frac{1}{v}} \right)^{d+1} ds \right)^{\frac{v}{d+1}} . \end{aligned}$$

In contrast to the former L_p-proof in [Kry87], we apply Hölder's inequality for $\frac{r}{d+1}$ which originally was applied for $\frac{v}{d+1}$. And therefore,

$$\begin{aligned} &e^{-\mu vt}\kappa_d d!(v\sqrt{\mu})^{-d}(\varphi(t,x))^v \\ &\qquad \le (\mu(d+1))^{-\frac{dv}{d+1}} (T-t)^{\frac{v}{d+1}-\frac{v}{r}} \left(\int\limits_t^T e^{-\mu rs} \left(\int\limits_{\mathbb{R}^d} |f(s,x)|^v \, dx \right)^{\frac{r}{v}} ds \right)^{\frac{v}{r}} \\ &\qquad = (\mu(d+1))^{-\frac{dv}{d+1}} (T-t)^{\frac{v}{d+1}-\frac{v}{r}} \left(\int\limits_t^\infty e^{-\mu rs} \left(\int\limits_{\mathbb{R}^d} |f(s,x)|^v \, dx \right)^{\frac{r}{v}} ds \right)^{\frac{v}{r}} \\ &\qquad = (\mu(d+1))^{-\frac{dv}{d+1}} (T-t)^{\frac{v}{d+1}-\frac{v}{r}} \left(\int\limits_0^\infty e^{-\mu r(t+s)} \left(\int\limits_{\mathbb{R}^d} |f(t+s,x)|^v \, dx \right)^{\frac{r}{v}} ds \right)^{\frac{v}{r}} . \end{aligned}$$

Multiplying this inequality with $e^{\mu vt}(\kappa_d d!)^{-1}(v\sqrt{\mu})^d$ we get

$$\begin{aligned} &(\varphi(t,x))^v \\ &\quad \le \frac{v^d}{\kappa_d d!} \mu^{\frac{d}{2}} (\mu(d+1))^{-\frac{dv}{d+1}} (T-t)^{\frac{v}{d+1}-\frac{v}{r}} e^{\mu vt} \left(\int\limits_0^\infty e^{-\mu r(t+s)} \left(\int\limits_{\mathbb{R}^d} |f(t+s,x)|^v \, dx \right)^{\frac{r}{v}} ds \right)^{\frac{v}{r}} \\ &\quad = C(d,v)\mu^{\frac{d}{2}-\frac{dv}{d+1}} (T-t)^{\frac{v}{d+1}-\frac{v}{r}} \left(\int\limits_0^\infty e^{-\mu rs} \left(\int\limits_{\mathbb{R}^d} |f(t+s,x)|^v \, dx \right)^{\frac{r}{v}} ds \right)^{\frac{v}{r}} \end{aligned}$$

and therefore,

$$\varphi(t,x) \le C(d,v)\mu^{\frac{d}{2v}-\frac{d}{d+1}}(T-t)^{\frac{1}{d+1}-\frac{1}{r}} \left(\int_0^\infty e^{-\mu r s} \left(\int_{\mathbb{R}^d} |f(t+s,x)|^v \, dx \right)^{\frac{r}{v}} ds \right)^{\frac{1}{r}}$$

for all $r \ge d+1$, $v > d+1$. By letting $v \searrow d+1$, we obtain the inequality also for $v = d+1$. The critical term in this limit argument is

$$\lim_{v \searrow d+1} \int_0^\infty e^{-\mu r s} \left(\int_{\mathbb{R}^d} |f(t+s,x)|^v \, dx \right)^{\frac{r}{v}} ds.$$

Since f is bounded and zero outside of B_R, we have with dominated convergence

$$\begin{aligned} \lim_{v \searrow d+1} \int_{\mathbb{R}^d} |f(t+s,x)|^v \, dx &= \int_{\mathbb{R}^d} \lim_{v \searrow d+1} |f(t+s,x)|^v \, dx \\ &= \int_{\mathbb{R}^d} |f(t+s,x)|^{d+1} \, dx. \end{aligned} \tag{33}$$

Furthermore, f is also zero outside of $[0,T]$, so we obtain for all $s \ge 0$

$$\begin{aligned} e^{-\mu r s} \left(\int_{\mathbb{R}^d} |f(t+s,x)|^v \, dx \right)^{\frac{r}{v}} &\le \mathbb{1}_{[0,T]}(t+s) \left(\int_{\mathbb{R}^d} |f(t+s,x)|^v \, dx \right)^{\frac{r}{v}} \\ &\le \mathbb{1}_{[0,T]}(t+s) \left(\int_{B_R} C^v \, dx \right)^{\frac{r}{v}} \\ &= \mathbb{1}_{[0,T]}(t+s) |B_R|^{\frac{r}{v}} C^r \\ &\le C \mathbb{1}_{[0,T]}(t+s) \end{aligned}$$

and therefore, again by dominated convergence

$$\begin{aligned} &\lim_{v \searrow d+1} \int_0^\infty e^{-\mu r s} \left(\int_{\mathbb{R}^d} |f(t+s,x)|^v \, dx \right)^{\frac{r}{v}} ds \\ &= \int_0^\infty e^{-\mu r s} \lim_{v \searrow d+1} \left(\int_{\mathbb{R}^d} |f(t+s,x)|^v \, dx \right)^{\frac{r}{v}} ds. \end{aligned}$$

If we use the properties of the exponential and the logarithmic functions, we then obtain

$$\begin{aligned}
&\lim_{v\searrow d+1}\int_0^\infty e^{-\mu rs}\left(\int_{\mathbb{R}^d}|f(t+s,x)|^v\,dx\right)^{\frac{r}{v}}ds\\
&=\int_0^\infty e^{-\mu rs}\lim_{v\searrow d+1}\exp\left(\frac{r}{v}\ln\left(\int_{\mathbb{R}^d}|f(t+s,x)|^v\,dx\right)\right)ds\\
&=\int_0^\infty e^{-\mu rs}\exp\left(\frac{r}{d+1}\ln\left(\lim_{v\searrow d+1}\int_{\mathbb{R}^d}|f(t+s,x)|^v\,dx\right)\right)ds
\end{aligned}$$

and with (33)

$$\begin{aligned}
&\lim_{v\searrow d+1}\int_0^\infty e^{-\mu rs}\left(\int_{\mathbb{R}^d}|f(t+s,x)|^v\,dx\right)^{\frac{r}{v}}ds\\
&=\int_0^\infty e^{-\mu rs}\exp\left(\frac{r}{d+1}\ln\left(\int_{\mathbb{R}^d}|f(t+s,x)|^{d+1}\,dx\right)\right)ds\\
&=\int_0^\infty e^{-\mu rs}\left(\int_{\mathbb{R}^d}|f(t+s,x)|^{d+1}\,dx\right)^{\frac{r}{d+1}}ds.
\end{aligned}$$

□

The following lemma is just to prove the existence of a smooth function from $\mathbb{R}^{d+1}$ to $[0,1]$ which is strictly positive on $[0,T]\times\overline{B_R}$. Such a function is needed in the proofs of Lemma 4.2 and Lemma A.10.

Lemma A.9. *Set*

$$g(t,x) := \begin{cases} c\exp\left(-\frac{1}{1-|(t,x)|^2}\right) & \textit{if } \ |(t,x)|<1,\\ 0 & \textit{else,}\end{cases}$$

where c is chosen such that

$$\int_{\mathbb{R}^{d+1}} g(t,x)\,d(t,x) = 1. \tag{34}$$

Then

$$\chi := g * \left(\mathbb{1}_{[0,T]}\cdot\mathbb{1}_{\overline{B_R}}\right)$$

fulfills

$$0<\chi\le 1 \quad \textit{on} \ \ [0,T]\times\overline{B_R}.$$

Proof. We have

$$\begin{aligned}\chi(t,x) &= \int\limits_{\mathbb{R}^{d+1}} g(t-s,x-y)\mathbb{1}_{[0,T]}(s)\mathbb{1}_{\overline{B_R}}(y)\,d(s,y)\\ &= \int\limits_0^T \int\limits_{\overline{B_R}} g(t-s,x-y)\,dy\,ds\\ &= \int\limits_0^T \int\limits_{\overline{B_{x,R}}} g(t-s,y)\,dy\,ds,\end{aligned}$$

where $B_{x,R}$ denotes the ball with radius R and center x. Obviously, by (34)

$$\chi(t,x) = \int\limits_{t-T}^{t} \int\limits_{\overline{B_{x,R}}} g(s,y)\,dy\,ds \le 1.$$

To prove that χ is strictly positive for all $(t,x) \in [0,T] \times \overline{B_R}$, it is enough to show that

$$dt \times dx\left(\left([t-T,t] \times \overline{B_{x,R}}\right) \cap B_1^{(d+1)}\right) > 0,$$

where the $(d+1)$ indicates that we mean here a ball in $\mathbb{R}^{d+1}$. For $R + |x| < \frac{1}{2}$, we have

$$\left(t - T \vee -\frac{1}{2}, t \wedge \frac{1}{2}\right) \times \overline{B_{x,R}} \ \subset \ B_1^{(d+1)},$$

since for all $(s,y) \in \left(t - T \vee -\frac{1}{2}, t \wedge \frac{1}{2}\right) \times \overline{B_{x,R}}$

$$\begin{aligned}|(s,y)| &\le |s| + |y|\\ &< \frac{1}{2} + |y-x| + |x|\\ &\le \frac{1}{2} + R + |x|\\ &< 1,\end{aligned}$$

and therefore,

$$\begin{aligned}dt \times dx&\left(\left([t-T,t] \times \overline{B_{x,R}}\right) \cap B_1^{(d+1)}\right)\\ &\ge dt \times dx\left(\left(t - T \vee -\frac{1}{2}, t \wedge \frac{1}{2}\right) \times \overline{B_{x,R}}\right)\\ &> 0.\end{aligned}$$

If $R+|x|\geq \frac{1}{2}$ we have

$$\left(t-T\vee -\frac{1}{2}, t\wedge \frac{1}{2}\right)\times B_{\frac{x}{4|x|},\frac{1}{4}} \ \subset \ B_1^{(d+1)},$$

since for all $(s,y)\in \left(t-T\vee -\frac{1}{2}, t\wedge \frac{1}{2}\right)\times B_{\frac{x}{4|x|},\frac{1}{4}}$

$$\begin{aligned}|(s,y)| &\leq |s|+|y|\\ &< \frac{1}{2}+\left|y-\frac{x}{4|x|}\right|+\left|\frac{x}{4|x|}\right|\\ &< 1\end{aligned}$$

and

$$B_{\frac{x}{4|x|},\frac{1}{4}} \ \subset \ \overline{B_{x,R}},$$

since for all $y\in B_{\frac{x}{4|x|},\frac{1}{4}}$ we have

$$\begin{aligned}|y-x| &\leq \left|y-\frac{x}{4|x|}\right|+\left|\frac{x}{4|x|}-x\right|\\ &< \frac{1}{4}+|x|\left|\frac{1}{4|x|}-1\right|\\ &= \frac{1}{4}+|x|\cdot\left\{\begin{array}{ccc}\frac{1}{4|x|}-1 & , & \text{if } \ 4|x|\leq 1\\ 1-\frac{1}{4|x|} & , & \text{if } \ 4|x|>1\end{array}\right\}\\ &= \left\{\begin{array}{ccc}\frac{1}{2}-|x| & , & \text{if } \ 4|x|\leq 1\\ |x| & , & \text{if } \ 4|x|>1\end{array}\right\}\\ &\leq R.\end{aligned}$$

Therefore, we have also in this case

$$\begin{aligned}&dt\times dx\left(\left([t-T,t]\times\overline{B_{x,R}}\right)\cap B_1^{(d+1)}\right)\\ &\qquad\geq dt\times dx\left(\left(t-T\vee -\frac{1}{2}, t\wedge\frac{1}{2}\right)\times B_{\frac{x}{4|x|},\frac{1}{4}}\right)\\ &\qquad> 0.\end{aligned}$$

□

Lemma A.10. *Let* $(c3)$, $(c4)$ *of Assumption 2.2 be fulfilled and* $X_t^{(1)}$, $X_t^{(2)}$ *be two solutions to* (5) *such that condition* (6) *holds. Then we have, for any nonnegative Borel function* $f:[0,T]\times\mathbb{R}^d\to\mathbb{R}$, *any stopping time* γ, $\lambda\in[0,1]$ *and* $r,v\geq d+1$,

$$\mathbb{E}\left[\int_0^{T\wedge\tau_R^\lambda\wedge\gamma}\det(a_t^\lambda)^{\frac{1}{d+1}}f(t,X_t^\lambda)\,dt\right]\leq C(d,v,r,T)(\mathbb{B}^2+\mathbb{A})^{\frac{d}{d+1}-\frac{d}{2v}}\|f\|_{L_v^r(T)}$$

where

$$\mathbb{A} := \mathbb{E}\left[\int_0^{T\wedge\tau_R^\lambda\wedge\gamma} \operatorname{tr}(a_t^\lambda)\,dt\right], \qquad \mathbb{B} := \mathbb{E}\left[\int_0^{T\wedge\tau_R^\lambda\wedge\gamma} \left|b^\lambda(t, X_t^{(1)}, X_t^{(2)})\right|\,dt\right].$$

The proof follows the ideas of the proof of Lemma 5.1 in [Kry86].

Proof. We assume $\mathbb{A} < \infty$ and $\mathbb{B} < \infty$, otherwise the inequality is trivially fulfilled. Fix a number $\mu > 0$ and take a nonnegative $f \in \mathcal{C}_0^\infty(\mathbb{R}^{d+1})$ such that $f > 0$ on $[0,T] \times \overline{B_R}$. Note, that there exists T', R' such that $f(t,x) = 0$ for $|t| \geq T'$ or $|x| > R'$. Then Lemma A.8, applied on $e^{\mu t} f$, ensures the existence of a nonnegative function φ with bounded weak derivatives $\partial_t\varphi, \partial_x\varphi, \partial_x^2\varphi$ such that for any symmetric, positive semidefinite $d \times d$ matrix α

$$\begin{aligned}
&\partial_t\varphi + \sum_{i=1}^d\sum_{j=1}^d \alpha_{ij}\partial^2_{x_ix_j}\varphi - \mu(1+\operatorname{tr}(\alpha))\varphi + \det(\alpha)^{\frac{1}{d+1}} f e^{\mu t} \leq 0,\\
&|\partial_x\varphi| \leq \sqrt{\mu}\varphi,\\
&\varphi(t,x) \leq C(d,v)\mu^{\frac{d}{2v}-\frac{d}{d+1}}(T'-t)^{\frac{1}{d+1}-\frac{1}{r}} e^{\mu t}\|f\|_{L_v^r}.
\end{aligned}$$

Define $\psi := \varphi e^{-\mu t}$. Then we have

$$\partial_t\psi + \sum_{i=1}^d\sum_{j=1}^d \alpha_{ij}\partial^2_{x_ix_j}\psi - \mu\operatorname{tr}(\alpha)\psi + \det(\alpha)^{\frac{1}{d+1}} f \leq 0, \tag{35}$$

$$|\partial_x\psi| \leq \sqrt{\mu}\psi, \tag{36}$$

$$\psi(t,x) \leq C(d,v)\mu^{\frac{d}{2v}-\frac{d}{d+1}}(T'-t)^{\frac{1}{d+1}-\frac{1}{r}}\|f\|_{L_v^r}. \tag{37}$$

From [Kry87], Example 6.4.6 we know, that $\partial_t\psi, \partial_x\psi, \partial_x^2\psi$ are continuous on $[0,T]\times\overline{B_R}$. Therefore, we may apply Itô's formula and get for all $t \in [0, T\wedge\tau_R^\lambda\wedge\gamma)$

$$\begin{aligned}
\psi(t, X_t^\lambda) - \psi(0, x^\lambda) &= \int_0^t \partial_t\psi(s, X_s^\lambda)\,ds + \int_0^t \partial_x\psi(s, X_s^\lambda) b^\lambda(s, X_s^{(1)}, X_s^{(2)})\,ds\\
&\quad + \int_0^t \partial_x\psi(s, X_s^\lambda)\sigma^\lambda(s, X_s^{(1)}, X_s^{(2)})\,dW_s\\
&\quad + \frac{1}{2}\int_0^t \sum_{i=1}^d\sum_{j=1}^d \left(\sigma^\lambda{\sigma^\lambda}^*(s, X_s^{(1)}, X_s^{(2)})\right)_{ij}\partial^2_{x_ix_j}\psi(s, X_s^\lambda)\,ds
\end{aligned}$$

which shows that

$$\kappa_t := \psi(t, X_t^\lambda) - \int_0^t \partial_x \psi(s, X_s^\lambda) b^\lambda(s, X_s^{(1)}, X_s^{(2)}) + \partial_t \psi(s, X_s^\lambda) + \sum_{i=1}^d \sum_{j=1}^d (a_s^\lambda)_{ij} \partial^2_{x_i x_j} \psi(s, X_s^\lambda)\, ds$$

is a martingale on $[0, T \wedge \tau_R^\lambda \wedge \gamma)$. Since for all $A \in \mathbb{R}^{d \times m}$ the matrix AA^* is positive semidefinite, we can use (35) and get

$$\kappa_t \geq \psi(t, X_t^\lambda) - \int_0^t \partial_x \psi(s, X_s^\lambda) b^\lambda(s, X_s^{(1)}, X_s^{(2)}) + \mu \operatorname{tr}(a_s^\lambda) \psi(s, X_s^\lambda) - \det(a_s^\lambda)^{\frac{1}{d+1}} f(s, X_s^\lambda)\, ds.$$

Since ψ is nonnegative we conclude that

$$\begin{aligned} \kappa_t &\geq - \int_0^t \partial_x \psi(s, X_s^\lambda) b^\lambda(s, X_s^{(1)}, X_s^{(2)}) + \mu \operatorname{tr}(a_s^\lambda) \psi(s, X_s^\lambda) - \det(a_s^\lambda)^{\frac{1}{d+1}} f(s, X_s^\lambda)\, ds \qquad (38)\\ &\geq - \int_0^t \partial_x \psi(s, X_s^\lambda) b^\lambda(s, X_s^{(1)}, X_s^{(2)}) + \mu \operatorname{tr}(a_s^\lambda) \psi(s, X_s^\lambda)\, ds \\ &\geq - \int_0^t \left|\partial_x \psi(s, X_s^\lambda)\right| \left|b^\lambda(s, X_s^{(1)}, X_s^{(2)})\right| + \mu \operatorname{tr}(a_s^\lambda) \psi(s, X_s^\lambda)\, ds. \end{aligned}$$

Then we may apply estimate (36) on $|\partial_x \psi|$ to obtain

$$\begin{aligned} \kappa_t &\geq - \int_0^t \sqrt{\mu} \psi(s, X_s^\lambda) \left|b^\lambda(s, X_s^{(1)}, X_s^{(2)})\right| + \mu \operatorname{tr}(a_s^\lambda) \psi(s, X_s^\lambda)\, ds \\ &\geq - \sup_{s \in [0,t]} \psi(s, X_s^\lambda) \int_0^t \sqrt{\mu} \left|b^\lambda(s, X_s^{(1)}, X_s^{(2)})\right| + \mu \operatorname{tr}(a_s^\lambda)\, ds. \end{aligned}$$

And an application of (37) yields

$$\kappa_t \geq -C(d, v) \mu^{\frac{d}{2v} - \frac{d}{d+1}} (T')^{\frac{1}{d+1} - \frac{1}{r}} \|f\|_{L_v^r} \int_0^t \sqrt{\mu} \left|b^\lambda(s, X_s^{(1)}, X_s^{(2)})\right| + \mu \operatorname{tr}(a_s^\lambda)\, ds.$$

Since $\mathbb{A}$ and $\mathbb{B}$ are finite, the right-hand side is bounded from below by an integrable function, justifying an application of Fatou's lemma providing

$$\mathbb{E}\Big[\ \kappa_{T\wedge\tau_R^\lambda\wedge\gamma}\ \Big] = \mathbb{E}\Big[\ \liminf_{n\to\infty}\kappa_{T\wedge\tau_R^\lambda\wedge\gamma-\frac{1}{n}}\ \Big] \le \liminf_{n\to\infty}\mathbb{E}\Big[\ \kappa_{T\wedge\tau_R^\lambda\wedge\gamma-\frac{1}{n}}\ \Big] = \mathbb{E}\Big[\ \kappa_0\ \Big].$$

And therefore, by (38)

$$\begin{aligned}
&\mathbb{E}\left[\int_0^{T\wedge\tau_R^\lambda\wedge\gamma}\det(a_t^\lambda)^{\frac{1}{d+1}}f(t,X_t^\lambda)\,dt\right]\\
&\qquad-\mathbb{E}\left[\int_0^{T\wedge\tau_R^\lambda\wedge\gamma}\partial_x\psi(t,X_t^\lambda)b^\lambda(t,X_t^{(1)},X_t^{(2)})+\mu\,\mathrm{tr}(a_t^\lambda)\psi(t,X_t^\lambda)\,dt\right]\\
&\le\mathbb{E}\Big[\ \kappa_{T\wedge\tau_R^\lambda\wedge\gamma}\ \Big]\\
&\le\psi(0,x^\lambda).
\end{aligned}$$

Further, with the help of the estimates (36) and (37) we deduce that

$$\begin{aligned}
&\mathbb{E}\left[\int_0^{T\wedge\tau_R^\lambda\wedge\gamma}\det(a_t^\lambda)^{\frac{1}{d+1}}f(t,X_t^\lambda)\,dt\right]\\
&\le\mathbb{E}\left[\int_0^{T\wedge\tau_R^\lambda\wedge\gamma}\partial_x\psi(t,X_t^\lambda)b^\lambda(t,X_t^{(1)},X_t^{(2)})+\mu\,\mathrm{tr}(a_t^\lambda)\psi(t,X_t^\lambda)\,dt\right]+\psi(0,x^\lambda)\\
&\le\mathbb{E}\left[\int_0^{T\wedge\tau_R^\lambda\wedge\gamma}\left|\partial_x\psi(t,X_t^\lambda)\right|\left|b^\lambda(t,X_t^{(1)},X_t^{(2)})\right|+\mu\,\mathrm{tr}(a_t^\lambda)\psi(t,X_t^\lambda)\,dt\right]+\psi(0,x^\lambda)\\
&\le\mathbb{E}\left[\int_0^{T\wedge\tau_R^\lambda\wedge\gamma}\sqrt{\mu}\psi(t,X_t^\lambda)\left|b^\lambda(t,X_t^{(1)},X_t^{(2)})\right|+\mu\,\mathrm{tr}(a_t^\lambda)\psi(t,X_t^\lambda)\,dt\right]+\psi(0,x^\lambda)\\
&\le\mathbb{E}\left[\sup_{t\in[0,T\wedge\tau_R^\lambda\wedge\gamma]}\psi(t,X_t^\lambda)\int_0^{T\wedge\tau_R^\lambda\wedge\gamma}\sqrt{\mu}\left|b^\lambda(t,X_t^{(1)},X_t^{(2)})\right|+\mu\,\mathrm{tr}(a_t^\lambda)\,dt\right]+\psi(0,x^\lambda)\\
&\le C(d,v)\mu^{\frac{d}{2v}-\frac{d}{d+1}}(T')^{\frac{1}{d+1}-\frac{1}{r}}\|f\|_{L_v^r}\left(\sqrt{\mu}\mathbb{B}+\mu\mathbb{A}+1\right)
\end{aligned}$$

which leaves us in need of a case distinction to handle the term $\sqrt{\mu}\mathbb{B}+\mu\mathbb{A}+1$. There are three cases to consider, namely $0\le\mathbb{A}<\mathbb{B}^2$, $\mathbb{A}>0\wedge\mathbb{A}\ge\mathbb{B}^2$ and $\mathbb{A}=\mathbb{B}=0$. If $0\le\mathbb{A}<\mathbb{B}^2$ take $\mu^{-1}=\mathbb{B}^2$, then we have

$$\mu^{\frac{d}{2v}-\frac{d}{d+1}}\left(\sqrt{\mu}\mathbb{B}+\mu\mathbb{A}+1\right)\ \le\ 3\mathbb{B}^{\frac{2d}{d+1}-\frac{2d}{2v}}\ \le\ 3(\mathbb{B}^2+\mathbb{A})^{\frac{d}{d+1}-\frac{d}{2v}}.$$

In the case of $\mathbb{A} > 0$ and $\mathbb{A} \geq \mathbb{B}^2$ take $\mu^{-1} = \mathbb{A}$ which leads to the same estimate:

$$\mu^{\frac{d}{2v}-\frac{d}{d+1}}(\sqrt{\mu}\mathbb{B} + \mu\mathbb{A} + 1) \;\leq\; 3\mathbb{A}^{\frac{d}{d+1}-\frac{d}{2v}} \;\leq\; 3(\mathbb{B}^2 + \mathbb{A})^{\frac{d}{d+1}-\frac{d}{2v}}.$$

Finally, for $\mathbb{A} = \mathbb{B}^2 = 0$ we have

$$\lim_{\mu\to\infty} \mu^{\frac{d}{2v}-\frac{d}{d+1}}(\sqrt{\mu}\mathbb{B} + \mu\mathbb{A} + 1) \;=\; 0 \;=\; 3(\mathbb{B}^2 + \mathbb{A})^{\frac{d}{d+1}-\frac{d}{2v}}.$$

So, we proved

$$\mathbb{E}\left[\int_0^{T\wedge\tau_R^\lambda\wedge\gamma} \det(a_t^\lambda)^{\frac{1}{d+1}} f(t, X_t^\lambda)\, dt\right] \leq C(d, v, r, T')(\mathbb{B}^2 + \mathbb{A})^{\frac{d}{d+1}-\frac{d}{2v}} \|f\|_{L_v^r} \tag{39}$$

for all nonnegative $f \in \mathcal{C}_0^\infty(\mathbb{R}^{d+1})$ with $f > 0$ on $[0, T] \times \overline{B_R}$. Now, let $f \in \mathcal{C}_0^\infty(\mathbb{R}^{d+1})$ be nonnegative. Take a smooth function $\chi : \mathbb{R}^{d+1} \to [0, 1]$ with compact support and

$$\chi > 0 \quad \text{on} \quad [0, T] \times \overline{B_R},$$

for example χ from Lemma A.9. Then we have

$$\mathbb{E}\left[\int_0^{T\wedge\tau_R^\lambda\wedge\gamma} \det(a_t)^{\frac{1}{d+1}} f(t, X_t^\lambda)\, dt\right] = \mathbb{E}\left[\int_0^{T\wedge\tau_R^\lambda\wedge\gamma} \det(a_t)^{\frac{1}{d+1}} \lim_{\varepsilon\searrow 0} (f + \varepsilon\chi)(t, X_t^\lambda)\, dt\right].$$

Uniform convergence of $\varepsilon\chi \searrow 0$ implies

$$\mathbb{E}\left[\int_0^{T\wedge\tau_R^\lambda\wedge\gamma} \det(a_t)^{\frac{1}{d+1}} f(t, X_t^\lambda)\, dt\right] = \lim_{\varepsilon\searrow 0} \mathbb{E}\left[\int_0^{T\wedge\tau_R^\lambda\wedge\gamma} \det(a_t)^{\frac{1}{d+1}} (f + \varepsilon\chi)(t, X_t^\lambda)\, dt\right].$$

As $f + \varepsilon\chi$ is strictly positive on $[0, T] \times \overline{B_R}$, we have by (39)

$$\begin{aligned}\mathbb{E}\left[\int_0^{T\wedge\tau_R^\lambda\wedge\gamma} \det(a_t)^{\frac{1}{d+1}} f(t, X_t^\lambda)\, dt\right] &\leq \lim_{\varepsilon\searrow 0} C(d, v, r, T')(\mathbb{B}^2 + \mathbb{A})^{\frac{d}{d+1}-\frac{d}{2v}} \|f + \varepsilon\chi\|_{L_v^r} \\ &= C(d, v, r, T')(\mathbb{B}^2 + \mathbb{A})^{\frac{d}{d+1}-\frac{d}{2v}} \|f\|_{L_v^r}\end{aligned}$$

for a suitable $T' > 0$.

The next step is to prove that C depends only on T not on T'. To this end define the smooth function

$$g(t) := \begin{cases} c\exp\left(-\frac{1}{1-|2t|^2}\right) & \text{for } |t| < \frac{1}{2}, \\ 0 & \text{else,} \end{cases}$$

where c is chosen such that

$$\int_{\mathbb{R}} g(t)dt = 1.$$

Observe that

$$\begin{aligned}\left(\mathbb{1}_{[-\frac{1}{2},T+\frac{1}{2}]} * g\right)(t) &= \int_{\mathbb{R}} \mathbb{1}_{[-\frac{1}{2},T+\frac{1}{2}]}(t-s)g(s)\,ds \\ &= \int_{-\frac{1}{2}}^{\frac{1}{2}} \mathbb{1}_{[-\frac{1}{2},T+\frac{1}{2}]}(t-s)g(s)\,ds \\ &= \begin{cases} 1, & t \in [0,T], \\ 0, & t \in [-1,T+1]^c, \\ \in [0,1], & \text{else}, \end{cases} \end{aligned} \tag{40}$$

because for all $s \in [-\frac{1}{2}, \frac{1}{2}]$ we have

$$\begin{aligned} 0 \le t \le T \quad &\Rightarrow -\frac{1}{2} \le (t-s) \le T + \frac{1}{2}, \\ t < -1 \quad &\Rightarrow (t-s) < -\frac{1}{2} \\ \text{and} \qquad t > T+1 \quad &\Rightarrow (t-s) > T + \frac{1}{2}. \end{aligned}$$

Smoothness of $(\mathbb{1}_{[-\frac{1}{2},T+\frac{1}{2}]} * g) \cdot f$ and (40) imply that

$$\begin{aligned} &\mathbb{E}\left[\int_0^{T\wedge\tau_R^\lambda\wedge\gamma} \det(a_t^\lambda)^{\frac{1}{d+1}} f(t, X_t^\lambda)\,dt\right] \\ &= \mathbb{E}\left[\int_0^{T\wedge\tau_R^\lambda\wedge\gamma} \det(a_t^\lambda)^{\frac{1}{d+1}} \left(\mathbb{1}_{[-\frac{1}{2},T+\frac{1}{2}]} * g\right)(t) f(t, X_t^\lambda)\,dt\right] \\ &\le C(d,v,r,T+1)(\mathbb{B}^2 + \mathbb{A})^{\frac{d}{d+1} - \frac{d}{2v}} \|(\mathbb{1}_{[-\frac{1}{2},T+\frac{1}{2}]} * g) \cdot f\|_{L_v^r} \\ &\le C(d,v,r,T)(\mathbb{B}^2 + \mathbb{A})^{\frac{d}{d+1} - \frac{d}{2v}} \|f\|_{L_v^r}. \end{aligned}$$

Now, let $f \in \mathcal{C}_0^\infty(\mathbb{R}^{d+1})$. Since $|f|$ is continuous and compactly supported, there exists a sequence $(f_n)_n$ of nonnegative functions in $\mathcal{C}_0^\infty(\mathbb{R}^{d+1})$ converging uniformly to $|f|$. (Just take a mollifier ϕ on $\mathbb{R}^{d+1}$ (see Definition A.2), then clearly $\phi_\varepsilon * |f| \in \mathcal{C}_0^\infty$ and $\phi_\varepsilon * |f| \to |f|$ uniformly.) Therefore, we have

$$\mathbb{E}\left[\int_0^{T\wedge\tau_R^\lambda\wedge\gamma} \det(a_t^\lambda)^{\frac{1}{d+1}} |f(t, X_t^\lambda)|\,dt\right] = \lim_{n\to\infty} \mathbb{E}\left[\int_0^{T\wedge\tau_R^\lambda\wedge\gamma} \det(a_t^\lambda)^{\frac{1}{d+1}} f_n(t, X_t^\lambda)\,dt\right].$$

Using inequality (39), we obtain

$$\mathbb{E}\left[\int_0^{T\wedge\tau_R^\lambda\wedge\gamma} \det(a_t^\lambda)^{\frac{1}{d+1}}|f(t,X_t^\lambda)|\,dt\right] \leq \lim_{n\to\infty} C(d,v,r,T)(\mathbb{B}^2+\mathbb{A})^{\frac{d}{d+1}-\frac{d}{2v}}\|f_n\|_{L_v^r}$$
$$= C(d,v,r,T)(\mathbb{B}^2+\mathbb{A})^{\frac{d}{d+1}-\frac{d}{2v}}\|f\|_{L_v^r}.$$

To extend the validity of this estimate to bounded measurable functions we denote

$$\mathcal{X} := \left\{ f:\mathbb{R}^{d+1}\to\mathbb{R} \;\middle|\; f \text{ is measurable, bounded and} \right.$$
$$\left. \mathbb{E}\left[\int_0^{T\wedge\tau_R^\lambda\wedge\gamma} \det(a_t^\lambda)^{\frac{1}{d+1}}|f(t,X_t^\lambda)|\,dt\right] \leq C(d,v,r,T)(\mathbb{B}^2+\mathbb{A})^{\frac{d}{d+1}-\frac{d}{2v}}\|f\|_{L_v^r} \right\}.$$

This set is closed under bounded monotone convergence, because for every sequence $0\leq f_1\leq f_2\leq ...\leq f_n\leq ...$ in $\mathcal{X}$ with $f_n\to f$ pointwise and f bounded, by monotone convergence we achieve

$$\begin{aligned}
&\mathbb{E}\left[\int_0^{T\wedge\tau_R^\lambda\wedge\gamma} \det(a_t^\lambda)^{\frac{1}{d+1}}|f(t,X_t^\lambda)|\,dt\right]\\
&= \mathbb{E}\left[\int_0^{T\wedge\tau_R^\lambda\wedge\gamma} \det(a_t^\lambda)^{\frac{1}{d+1}}\lim_{n\to\infty}|f_n(t,X_t^\lambda)|\,dt\right]\\
&= \lim_{n\to\infty}\mathbb{E}\left[\int_0^{T\wedge\tau_R^\lambda\wedge\gamma} \det(a_t^\lambda)^{\frac{1}{d+1}}|f_n(t,X_t^\lambda)|\,dt\right]\\
&\leq \lim_{n\to\infty} C(d,v,r,T)(\mathbb{B}^2+\mathbb{A})^{\frac{d}{d+1}-\frac{d}{2v}}\|f_n\|_{L_v^r}\\
&= C(d,v,r,T)(\mathbb{B}^2+\mathbb{A})^{\frac{d}{d+1}-\frac{d}{2v}}\|\lim_{n\to\infty} f_n\|_{L_v^r}\\
&= C(d,v,r,T)(\mathbb{B}^2+\mathbb{A})^{\frac{d}{d+1}-\frac{d}{2v}}\|f\|_{L_v^r}.
\end{aligned}$$

And since measurability is preserved by pointwise convergence, we have $f\in\mathcal{X}$. Replacing the monotone convergence by uniform convergence in the above computation serves that $\mathcal{X}$ is also closed under uniform convergence. $\mathcal{C}_0^\infty(\mathbb{R}^{d+1})$ is an algebra and there exists a sequence f_n in $C_0^\infty(\mathbb{R}^{d+1})$ such that $f_n\nearrow 1$. So we may apply a version of the monotone-class theorem, which can be found in [Del78], stated as a variant of Theorem 21, labeled as (22.2). Therefore, $\mathcal{X}$ contains all measurable bounded functions.

For unbounded measurable functions we deduce with an application of the monotone convergence theorem:

$$\begin{aligned}
&\mathbb{E}\left[\int_0^{T\wedge\tau_R^\lambda\wedge\gamma} \det(a_t^\lambda)^{\frac{1}{d+1}} |f(t,X_t^\lambda)|\,dt\right] \\
&= \mathbb{E}\left[\int_0^{T\wedge\tau_R^\lambda\wedge\gamma} \det(a_t^\lambda)^{\frac{1}{d+1}} \lim_{n\to\infty} |f(t,X_t^\lambda)| \wedge n \;\; dt\right] \\
&= \lim_{n\to\infty} \mathbb{E}\left[\int_0^{T\wedge\tau_R^\lambda\wedge\gamma} \det(a_t^\lambda)^{\frac{1}{d+1}} |f(t,X_t^\lambda)| \wedge n \;\; dt\right] \\
&\leq \lim_{n\to\infty} C(d,v,r,T)(\mathbb{B}^2+\mathbb{A})^{\frac{d}{d+1}-\frac{d}{2v}} \|f\wedge n\|_{L_v^r} \\
&= C(d,v,r,T)(\mathbb{B}^2+\mathbb{A})^{\frac{d}{d+1}-\frac{d}{2v}} \|f\|_{L_v^r}.
\end{aligned}$$

Summarizing we have for any nonnegative measurable function f and any $\lambda \in [0,1]$

$$\begin{aligned}
&\mathbb{E}\left[\int_0^{T\wedge\tau_R^\lambda\wedge\gamma} \det(a_t^\lambda)^{\frac{1}{d+1}} f(t,X_t^\lambda)\,dt\right] \\
&= \mathbb{E}\left[\int_0^{T\wedge\tau_R^\lambda\wedge\gamma} \det(a_t^\lambda)^{\frac{1}{d+1}} \mathbb{1}_{[0,T]}(t) f(t,X_t^\lambda)\,dt\right] \\
&\leq C(d,v,r,T)(\mathbb{B}^2+\mathbb{A})^{\frac{d}{d+1}-\frac{d}{2v}} \|f\|_{L_v^r(T)}.
\end{aligned}$$

□

Lemma A.11. *Let* (c1), (c3), (c4) *of Assumption 2.2 be fulfilled and* $X_t^{(1)}$, $X_t^{(2)}$ *be two solutions to* (5) *such that condition* (6) *holds. Then we have for every* $\lambda \in [0,1]$

$$\mathbb{E}\left[\int_0^T \operatorname{tr}(a_t^\lambda)\,dt\right] \leq C(T,\tilde{c}_\sigma)$$

and

$$\mathbb{E}\left[\int_0^T \left|b^\lambda(t,X_t^{(1)},X_t^{(2)})\right|\,dt\right] \leq C(d,p,q,T,c_\sigma,\tilde{c}_\sigma,\|b\|_{L_p^q(T)}).$$

The idea of the proof, in particular for the second estimate is borrowed from [GM01], Proof of Corollary 3.2.

Proof. Rewriting the trace of a_t^λ and applying Young's inequality and the boundedness of σ yields

$$\begin{aligned}
\operatorname{tr}(a_t^\lambda) &= \sum_{i=1}^{d} \frac{1}{2}\left(\sigma^\lambda \sigma^{\lambda^*}(t, X_t^{(1)}, X_t^{(2)})\right)_{ii} \\
&= \frac{1}{2}\sum_{i=1}^{d}\sum_{j=1}^{m}\left(\lambda\sigma(t, X_t^{(1)})_{ij} + (1-\lambda)\sigma(t, X_t^{(2)})_{ij}\right)^2 \\
&= \frac{1}{2}\left|\lambda\sigma(t, X_t^{(1)}) + (1-\lambda)\sigma(t, X_t^{(2)})\right|^2 \\
&\le \lambda^2\left|\sigma(t, X_t^{(1)})\right|^2 + (1-\lambda)^2\left|\sigma(t, X_t^{(2)})\right|^2 \\
&\le 2\tilde{c}_\sigma^2
\end{aligned} \tag{41}$$

and therefore,

$$\mathbb{E}\left[\int_0^T \operatorname{tr}(a_t^\lambda)\,dt\right] \le 2\tilde{c}_\sigma^2 T.$$

To prove that the second expectation is finite, we use Lemma A.10 for $X_t^{(1)}$ and $X_t^{(2)}$. Note, that all the eigenvalues of $\sigma\sigma^*$ are bounded from below by c_σ because of the nondegenerateness of σ (Assumption 2.2 (c3)). Since a symmetric matrix has solely real valued eigenvalues, and the determinant is the product of these, we have in case of $\lambda = 1$

$$\det(a_t^1) = \frac{1}{2^d}\det(\sigma\sigma^*(t, X_t^{(1)})) \ge \frac{1}{2^d}c_\sigma^d. \tag{42}$$

And the same holds true for $\det(a_t^0)$. Define

$$\gamma_n := \inf\left\{\, t \ge 0 \;\middle|\; \int_0^t |b(s, X_s^{(1)})|\,ds \;>\; n \,\right\},$$

$$\mathbb{B}^{(n)} := \mathbb{E}\left[\int_0^{T\wedge\tau_R^1\wedge\gamma_n} \left|b(t, X_t^{(1)})\right| dt\right] \quad \text{and} \quad \mathbb{A}^{(n)} := \mathbb{E}\left[\int_0^{T\wedge\tau_R^1\wedge\gamma_n} \operatorname{tr}(a_t^1)dt\right].$$

Then we have

$$
\begin{aligned}
\left(\mathbb{B}^{(n)}\right)^2 &\leq \mathbb{E}\left[\left(\int\limits_0^{T\wedge\tau_R^1\wedge\gamma_n} \left|b(t,X_t^{(1)})\right| \,dt\right)^2\right] \\
&\leq T\mathbb{E}\left[\int\limits_0^{T\wedge\tau_R^1\wedge\gamma_n} \left|b(t,X_t^{(1)})\right|^2 \,dt\right] \\
&= T\mathbb{E}\left[\int\limits_0^{T\wedge\tau_R^1\wedge\gamma_n} \left(\frac{\det(a_t^1)}{\det(a_t^1)}\right)^{\frac{1}{d+1}} \left|b(t,X_t^{(1)})\right|^2 \,dt\right] \\
&\leq \left(\frac{2}{c_\sigma}\right)^{\frac{d}{d+1}} T\mathbb{E}\left[\int\limits_0^{T\wedge\tau_R^1\wedge\gamma_n} \det(a_t^1)^{\frac{1}{d+1}} \left|b(t,X_t^{(1)})\right|^2 \,dt\right].
\end{aligned}
$$

Now, we may use the inequality from Lemma A.10 with $v = \frac{p}{2}$, $r = \frac{q}{2}$, to receive

$$
\begin{aligned}
\left(\mathbb{B}^{(n)}\right)^2 &\leq \left(\frac{2}{c_\sigma}\right)^{\frac{d}{d+1}} TC(d,p,q,T)\left(\left(\mathbb{B}^{(n)}\right)^2 + \mathbb{A}^{(n)}\right)^{\frac{d}{d+1}-\frac{d}{p}} \|b\|^2_{L_p^q(T)} \\
&\leq \left(\frac{2}{c_\sigma}\right)^{\frac{d}{d+1}} TC(d,p,q,T)\left(\left(\mathbb{B}^{(n)}\right)^{\frac{2d}{d+1}-\frac{2d}{p}} + \left(2\tilde{c}_\sigma^2 T\right)^{\frac{d}{d+1}-\frac{d}{p}}\right) \|b\|^2_{L_p^q(T)}.
\end{aligned} \tag{43}
$$

With Young's inequality we have for $\varepsilon > 0$ and $z := \frac{d}{d+1} - \frac{d}{p} < 1$

$$
\begin{aligned}
\left(\mathbb{B}^{(n)}\right)^{2z} &= \frac{1}{\varepsilon}\cdot\varepsilon\left(\mathbb{B}^{(n)}\right)^{2z} \\
&\leq (1-z)\varepsilon^{-\frac{1}{1-z}} + z\varepsilon^{\frac{1}{z}}\left(\mathbb{B}^{(n)}\right)^2 \\
&< \varepsilon^{\frac{1}{z-1}} + \varepsilon^{\frac{1}{z}}\left(\mathbb{B}^{(n)}\right)^2.
\end{aligned}
$$

Let ε be so small, that

$$
\left(\frac{2}{c_\sigma}\right)^{\frac{d}{d+1}} TC(d,p,q,T)\varepsilon^{\frac{1}{z}}\|b\|^2_{L_p^q(T)} < 1.
$$

Note, that the choice of ε is independent of n and R. Then we get with (43)

$$
\left(\mathbb{B}^{(n)}\right)^2 \leq \left(\frac{2}{c_\sigma}\right)^{\frac{d}{d+1}} TC(d,p,q,T)\left(\left(2\tilde{c}_\sigma^2 T\right)^{\frac{d}{d+1}-\frac{d}{p}} + \varepsilon^{\frac{1}{z-1}} + \varepsilon^{\frac{1}{z}}\left(\mathbb{B}^{(n)}\right)^2\right)\|b\|^2_{L_p^q(T)}
$$

which is equivalent to

$$
\left(\mathbb{B}^{(n)}\right)^2 \leq \frac{\left(\frac{2}{c_\sigma}\right)^{\frac{d}{d+1}} TC(d,p,q,T)\left((2\tilde{c}_\sigma^2 T)^{\frac{d}{d+1}-\frac{d}{p}} + \varepsilon^{\frac{1}{z-1}}\right)\|b\|^2_{L_p^q(T)}}{1-\left(\frac{2}{c_\sigma}\right)^{\frac{d}{d+1}} TC(d,p,q,T)\varepsilon^{\frac{1}{z}}\|b\|^2_{L_p^q(T)}},
$$

where the right-hand side is finite and independent of n. If we take the limit $n \to \infty$ we obtain that

$$\mathbb{E}\left[\int_0^{T\wedge\tau_R^1} \left|b(t, X_t^{(1)})\right| dt\right] \le C(d, p, q, T, c_\sigma, \tilde{c}_\sigma, \|b\|_{L_p^q(T)}).$$

Furthermore, the right-hand side is also independent of R. If we take the limit $R \to \infty$, we get

$$\mathbb{E}\left[\int_0^{T} \left|b(t, X_t^{(1)})\right| dt\right] \le C(d, p, q, T, c_\sigma, \tilde{c}_\sigma, \|b\|_{L_p^q(T)})$$

and analogously the same estimate for $X_t^{(2)}$. Therefore we obtain

$$\begin{aligned}
&\mathbb{E}\left[\int_0^{T} \left|\lambda b(t, X_t^{(1)}) + (1-\lambda) b(t, X_t^{(2)})\right| dt\right] \\
&\le \lambda\mathbb{E}\left[\int_0^{T} \left|b(t, X_t^{(1)})\right| dt\right] + (1-\lambda)\mathbb{E}\left[\int_0^{T} \left|b(t, X_t^{(2)})\right| dt\right] \\
&\le C(d, p, q, T, c_\sigma, \tilde{c}_\sigma, \|b\|_{L_p^q(T)})
\end{aligned}$$

for every $\lambda \in [0, 1]$. □

Lemma 3.1. *Let* (c1), (c3), (c4) *of Assumption 2.2 be fulfilled and* X_t *be a solution to* (5) *such that condition* (6) *holds. Then we have for every* $v, r \ge d + 1$ *and any nonnegative measurable function* $f : [0, T] \times \mathbb{R}^d \to \mathbb{R}$

$$\mathbb{E}\left[\int_0^{T} f(t, X_t)\, dt\right] \le C(T, d, p, q, v, r, c_\sigma, \tilde{c}_\sigma, \|b\|_{L_p^q(T)}) \|f\|_{L_v^r(T)}.$$

Proof. With $X_t^{(1)} = X_t$ and $\lambda = 1$ we have, with the estimate on the determinant (42) from the proof of Lemma A.11,

$$\begin{aligned}
\mathbb{E}\left[\int_0^{T\wedge\tau_R^1} f(t, X_t)\, dt\right] &= \mathbb{E}\left[\int_0^{T\wedge\tau_R^1} \left(\frac{\det(a_t^1)}{\det(a_t^1)}\right)^{\frac{1}{d+1}} f(t, X_t)\, dt\right] \\
&\le \left(\frac{2}{c_\sigma}\right)^{\frac{d}{d+1}} \mathbb{E}\left[\int_0^{T\wedge\tau_R^1} \det(a_t^1)^{\frac{1}{d+1}} f(t, X_t)\, dt\right].
\end{aligned}$$

Thus, with Lemma A.10 and Lemma A.11

$$\begin{aligned}\mathbb{E}\left[\int\limits_0^{T\wedge\tau_R^1} f(t,X_t)\,dt\right] &\le \left(\frac{2}{c_\sigma}\right)^{\frac{d}{d+1}} C(d,v,r,T)(\mathbb{B}^2+\mathbb{A})^{\frac{d}{d+1}-\frac{d}{2v}}\|f\|_{L^r_v(T)}\\ &\le C(T,d,p,q,v,r,\|b\|_{L^q_p(T)},c_\sigma,\tilde{c}_\sigma)\|f\|_{L^r_v(T)}.\end{aligned}$$

But the right-hand side is independent of R and therefore, the result follows by taking the limit $R\to\infty$. $\square$

References

[AF09] Robert A. Adams and John J. F. Fournier. *Sobolev spaces*, volume 140 of *Pure and applied mathematics series ; 140*. Acad. Press, Amsterdam [u.a.], 2. ed., digital print. edition, 2009.

[Alt16] Hans Wilhelm Alt. *Linear Functional Analysis*. Universitext. Springer London, London, 2016.

[Del78] Claude Dellacherie. *Probabilities and potential/[A]*, volume 29 of *North Holland mathematics studies ; 29*. 1978.

[DPFRV16] G. Da Prato, F. Flandoli, M. Röckner, and A. Yu. Veretennikov. Strong uniqueness for SDEs in Hilbert spaces with nonregular drift. *Ann. Probab.*, 44(3):1985–2023, 2016.

[FF11] E. Fedrizzi and F. Flandoli. Pathwise uniqueness and continuous dependence of SDEs with non-regular drift. *Stochastics*, 83(3):241–257, 2011.

[Fol99] Gerald B. Folland. *Real analysis*. Pure and applied mathematics. Wiley, New York [u.a.], 2. ed. edition, 1999.

[FZ05] Shizan Fang and Tusheng Zhang. A study of a class of stochastic differential equations with non-Lipschitzian coefficients. *Probab. Theory Related Fields*, 132(3):356–390, 2005.

[GM01] István Gyöngy and Teresa Martínez. On stochastic differential equations with locally unbounded drift. *Czechoslovak Math. J.*, 51(126)(4):763–783, 2001.

[Itô46] Kiyosi Itô. On a stochastic integral equation. *Proc. Japan Acad.*, 22(nos. 1-4):32–35, 1946.

[KR05] N. V. Krylov and M. Röckner. Strong solutions of stochastic equations with singular time dependent drift. *Probab. Theory Related Fields*, 131(2):154–196, 2005.

[Kry86] N. V. Krylov. Estimates of the maximum of the solution of a parabolic equation and estimates of the distribution of a semimartingale. *Mat. Sb. (N.S.)*, 130(172)(2):207–221, 284, 1986.

[Kry87] Nikolaj V. Krylov. *Nonlinear elliptic and parabolic equations of the second order*, volume 7 of *Mathematics and its applications : (Soviet series) ; 7*. Reidel, Dordrecht [u.a.], 1987.

[Kry99] N. V. Krylov. On Kolmogorov's equations for finite-dimensional diffusions. In *Stochastic PDE's and Kolmogorov equations in infinite dimensions (Cetraro, 1998)*, volume 1715 of *Lecture Notes in Math.*, pages 1–63. Springer, Berlin, 1999.

[Kry01] N. V. Krylov. The heat equation in $L_q((0,T), L_p)$-spaces with weights. *SIAM J. Math. Anal.*, 32(5):1117–1141, 2001.

[Kry07] N. V. Krylov. Parabolic equations with VMO coefficients in Sobolev spaces with mixed norms. *J. Funct. Anal.*, 250(2):521–558, 2007.

[Kry09] Nikolaj V. Krylov. *Controlled diffusion processes*, volume 14 of *Stochastic modelling and applied probability ; 14*. Springer, Berlin [u.a.], reprint of the 1980 ed. edition, 2009.

[KS91] Ioannis Karatzas and Steven E. Shreve. *Brownian motion and stochastic calculus*, volume 113 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, second edition, 1991.

[LR15] Wei Liu and Michael Röckner. *Stochastic partial differential equations: an introduction*. Universitext. Springer, Cham, 2015.

[RY05] Daniel Revuz and Marc Yor. *Continuous martingales and Brownian motion*. Die Grundlehren der mathematischen Wissenschaften ; 293. Springer, 2005.

[Ver78] A. Yu. Veretennikov. Strong solutions of some stochastic equations. *Uspekhi Mat. Nauk*, 33(5(203)):173–174, 1978.

[YW71] Toshio Yamada and Shinzo Watanabe. On the uniqueness of solutions of stochastic differential equations. *J. Math. Kyoto Univ.*, 11:155–167, 1971.

[Zha05] Xicheng Zhang. Strong solutions of SDES with singular drift and Sobolev diffusion coefficients. *Stochastic Process. Appl.*, 115(11):1805–1818, 2005.

[Zha11] Xicheng Zhang. Stochastic homeomorphism flows of SDEs with singular drifts and Sobolev diffusion coefficients. *Electron. J. Probab.*, 16:no. 38, 1096–1116, 2011.

[Zvo74] A. K. Zvonkin. A transformation of the phase space of a diffusion process that will remove the drift. *Mat. Sb. (N.S.)*, 93(135):129–149, 152, 1974.