On conditional densities of partially observed jump diffusions

Fabian Germ

Doctor of Philosophy
University of Edinburgh
November, 2022

Declaration

I declare that this thesis was composed by myself and that the work contained therein is my own, except where explicitly stated otherwise in the text.

(*Fabian Germ*)

Acknowledgements

First of all, I would like to thank my advisor, Professor István Gyöngy, for his continuous support, endless patience and inspiring guidance throughout the years of my PhD.

Moreover, I would like to express my sincere gratitude to Professor Nicolai Krylov, for his valuable comments and suggestions on the thesis, which greatly improved its presentation.
Thanks goes also to Professor Sotirios Sabanis, for organising my PhD defense and for his great advice on life as young mathematician.

I am no less grateful to my friends and colleagues, who made this journey a truly marvelous experience.

Na koncu bi se rad zahvalil tudi svoji družini in partnerki, saj brez njihove brezmejne spodbude in brezpogojne podpore nič od tega ne bi bilo mogoče uresničiti. Hvala za vse!

Für Viktoria.

Le temps, cher ami, le temps amène l'occasion,
l'occasion c'est la martingale de l'homme:
plus on a engagé, plus l'on gagne quand
on sait attendre.

∼ Athos, Le comte Olivier de la Fère[1]

[1]In *Les Trois Mousquetaires* by Alexandre Dumas

Lay summary

Consider a system we would like to observe to determine its current state, using some form of measurement. In many situations, the system may exhibit complex behaviour and depend on many variables. The observations we can obtain may be limited, as well as subject to errors and uncertainty, due to our imperfect measurement method. In such a case, filtering aims to give the best possible estimate of the system's state, while accounting for the flaws in the information we obtain.
To give an example, let us imagine we want to track weather parameters in a given region. In order to do so, we have to come up with a model, which we think describes their dynamics well and which usually depends on a number of variables. Let us assume that, among them, we are interested in the air pressure, the temperature and the humidity. The sensors we are given can only measure the latter two. Moreover, they cannot do so everywhere, but only in certain locations and with some error. With methods from filtering theory we can make an estimate on the weather and its other parameters, for instance the air pressure, given the flawed temperature and humidity measurements we have.
In many cases, it is desirable to know how likely a certain state of the system is, given partial information, i.e. the probability of the system having said state. Mathematically, this means we have to analyse the distribution of the state conditioned on partial observations. In our example, we could be interested in how likely the air pressure is between certain values, based on our temperature and humidity measurements. We refer to this as the conditional distribution, i.e. the distribution of the air pressure conditional on the temperature and humidity.

In our work, we consider a class of systems which has received a lot of interest in mathematics lately, referred to as jump-diffusions. They model phenomena that experience instantaneous changes in their dynamics, in such a way, that their parameters exhibit jumps in their values. We prove that, if at the start of our experiment, the conditional distribution is well-behaved in a certain sense and if our model satisfies certain assumptions, then the conditional distribution remains well-behaved at all future times as well.

Abstract

In this thesis, we study the filtering problem for a partially observed jump diffusion $(Z_t)_{t\in[0,T]} = (X_t, Y_t)_{t\in[0,T]}$ driven by Wiener processes and Poisson martingale measures, such that the signal and observation noises are correlated. We derive the filtering equations, describing the time evolution of the normalised conditional distribution $(P_t(dx))_{t\in[0,T]}$ and the unnormalised conditional distribution of the unobservable signal X_t given the observations $(Y_s)_{s\in[0,t]}$. We prove that if the coefficients satisfy linear growth and Lipschitz conditions in space, as well as some additional assumptions on the jump coefficients, then, if $\mathbb{E}|\pi_0|^p_{L_p} < \infty$ for some $p \geqslant 2$, the conditional density $\pi = (\pi_t)_{t\in[0,T]}$, where $\pi_t = dP_t/dx$, exists and is a weakly cadlag L_p-valued process. Moreover, for an integer $m \geqslant 0$ and $p \geqslant 2$, we show that if we additionally impose $m+1$ continuous and bounded spatial derivatives on the coefficients and if the initial conditional density $\mathbb{E}|\pi_0|^p_{W^m_p} < \infty$, then π is weakly cadlag as a W^m_p-valued process and strongly cadlag as a W^s_p-valued process for $s \in [0, m)$.

Contents

Chapter I

Introduction

Since its early developments in the mid-twentieth century, filtering theory has been a heavily researched area in a number of mathematical and more applied disciplines. The primary objective of stochastic filtering is to develop methods that allow us to infer properties of the state of a system, given partial, noisy or otherwise corrupted information of it. More precisely, most often we are concerned with two random dynamical systems, one modeling a (partially) unobservable *signal* $X = (X_t)_{t\in[0,T]}$ and the other representing the *observation* $Y = (Y_t)_{t\in[0,T]}$ on some time interval $[0,T]$. In most cases, the "best" estimate of X_t at time t given past observations $\{Y_s : s \leqslant t\}$ is considered to be the mean-square estimate and it is well-known that for each time t, this is given by the conditional expectation $\mathbb{E}(X_t|\{Y_s : s \leqslant t\})$. Thus, in a more general sense, the classic task of filtering theory is to calculate for a real-valued Borel function f the conditional expectation

$$\mathbb{E}(f(X_t)|\{Y_s, s \leqslant t\}) = \int_{\mathbb{R}^d} f(x)P_t(dx), \quad t \in [0,T], \tag{I.0.1}$$

and hence it is of interest to study the properties of $P_t(dx)$, the conditional distribution of X_t given $\{Y_s, s \leqslant t\}$.
This has been done for a variety of dynamical systems $Z = (X,Y)$ and to give an overview of the corresponding literature would exceed the scope of this work; instead, we refer the reader to [11] for a historical account, to [3] for an exposition of various methods and approaches, as well as to the references therein for further reading.

In this thesis, we consider a $d + d'$-dimensional stochastic process $(Z_t)_{t\in[0,T]} = (X_t, Y_t)_{t\in[0,T]}$ on a given complete filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geqslant 0}, P)$, satisfying the stochastic differential equation (SDE)

$$
\begin{aligned}
dX_t &= b(t, Z_t)dt + \sigma(t, Z_t)dW_t + \rho(t, Z_t)dV_t \\
&\quad + \int_{\mathfrak{Z}_0} \eta(t, Z_{t-}, \mathfrak{z})\, \tilde{N}_0(d\mathfrak{z}, dt) + \int_{\mathfrak{Z}_1} \xi(t, Z_{t-}, \mathfrak{z})\, \tilde{N}_1(d\mathfrak{z}, dt), \\
dY_t &= B(t, Z_t)dt + dV_t + \int_{\mathfrak{Z}_1} \mathfrak{z}\, \tilde{N}_1(d\mathfrak{z}, dt),
\end{aligned}
\tag{I.0.2}
$$

on the interval $[0, T]$ for a given $\mathcal{F}_0$-measurable initial value $Z_0 = (X_0, Y_0)$, where $(W_t, V_t)_{t \geqslant 0}$ is $d_1 + d'$-dimensional $\mathcal{F}_t$-Wiener process, and $\tilde{N}_i(d\mathfrak{z}, dt) = N_i(d\mathfrak{z}, dt) - \nu_i(d\mathfrak{z})dt$, $i = 0, 1$, are independent $\mathcal{F}_t$-Poisson martingale measures on $\mathbb{R}_+ \times \mathfrak{Z}_i$ with σ-finite characteristic measures ν_0 and ν_1 on separable measurable spaces $(\mathfrak{Z}_0, \mathcal{Z}_0)$ and $(\mathfrak{Z}_1, \mathcal{Z}_1) = (\mathbb{R}^{d'} \backslash \{0\}, \mathcal{B}(\mathbb{R}^{d'} \backslash \{0\}))$, respectively. The mappings $b = (b^i)$, $B = (B^i)$, $\sigma = (\sigma^{ij})$ and $\rho = (\rho^{il})$ are Borel functions of $(t, z) = (t, x, y) \in \mathbb{R}_+ \times \mathbb{R}^{d+d'}$, with values in $\mathbb{R}^d$, $\mathbb{R}^{d'}$, $\mathbb{R}^{d \times d_1}$ and $\mathbb{R}^{d \times d'}$, respectively, and $\eta = (\eta^i)$ and $\xi = (\xi^i)$ are $\mathbb{R}^d$-valued $\mathcal{B}(\mathbb{R}_+ \times \mathbb{R}^{d+d'}) \otimes \mathcal{Z}_0$-measurable and $\mathbb{R}^d$-valued $\mathcal{B}(\mathbb{R}_+ \times \mathbb{R}^{d+d'}) \otimes \mathcal{Z}_1$-measurable functions of $(t, z, \mathfrak{z}_0) \in \mathbb{R}_+ \times \mathbb{R}^{d+d'} \times \mathfrak{Z}_0$ and $(t, z, \mathfrak{z}_1) \in \mathbb{R}_+ \times \mathbb{R}^{d+d'} \times \mathfrak{Z}_1$, respectively.
As in (I.0.1) we consider X to be the signal process and Y the observation process, governed by (I.0.2) on the finite time interval $[0, T]$. The primary aim of this thesis is to study the filtering density $\pi = (\pi_t)_{t \in [0,T]}$, satisfying for each $t \in [0, T]$ and bounded Borel function f,

$$
\pi_t = \frac{P(X_t \in dx | \{Y_s : s \in [0, t]\})}{dx} \quad \text{and} \quad \int_{\mathbb{R}^d} f(x)\, P_t(dx) = \int_{\mathbb{R}^d} f(x) \pi_t(x)\, dx,
\tag{I.0.3}
$$

almost surely.

The novelty of this thesis is threefold. Firstly, we derive the filtering equations, the Kushner-Shiryaev equation and the Zakai equation, associated to our model (I.0.2). The former describes the time evolution of the conditional distribution $(P_t(dx))_{t \in [0,T]}$ satisfying (I.0.1). We obtain it by a change of measure method and first deriving the Zakai equation for the "unnormalised" conditional density $(\mu_t(dx))_{t \in [0,T]}$, related to P via a normalising multiplicative process. While the filtering equations for models with jumps have also been derived in the recent works [2,5,9,10,19,47,49,50], to the best of the author's knowledge, they have not previously been derived for a model as general as (I.0.2). We refer the reader to the introduction of Chapter III for a more detailed comparison of our model to SDEs considered by other authors and related results on the derivation of the filtering equations.
Secondly, we prove the existence of density valued processes associated to $(P_t)_{t \in [0,T]}$ and $(\mu_t)_{t \in [0,T]}$. We show that if the coefficients in (I.0.2) have linear growth in $z = (x, y) \in \mathbb{R}^{d+d'}$ and are Lipschitz continuous in z, uniformly in the other variables, if the measure ν_1 admits an r-th moment for some $r > 2$, if η, ξ satisfy

some additional regularity condition in terms of a Jacobian and if the initial condition satisfies $\mathbb{E}|X_0|^r < \infty$, then so long as $\mathbb{E}|\pi_0|^p < \infty$ for some $p \geqslant 2$, the conditional distribution $(P_t(dx))_{t\in[0,T]}$ admits a density process $(\pi_t)_{t\in[0,T]}$, which is weakly cadlag as L_p-valued process, satisfying (I.0.3) almost surely for each $t \in [0,T]$ and bounded Borel function f. As a special case we obtain that if $p = 2$, it is enough to assume $r = 2$. Our methods rely on some existence and regularity results from [23] as well as Itô formulas for jump processes (in L_p-spaces) from [21, 22]. These articles provide very recent results on L_p-calculus for jump processes. While there are results on L_2-valued density processes for jump-diffusions, see [6,44,50], we obtained first results for existence of L_p-valued densities for our more general jump diffusion model (I.0.2) for $p \geqslant 2$.
Thirdly and finally, we investigate the regularity (in the Sobolev sense) of $(\pi_t)_{t\in[0,T]}$ under additional regularity assumptions on the coefficients. More precisely, we show that if in addition to the assumptions described above for the L_p-case, we assume the coefficients of (I.0.2) admit $m + 1$ continuous and bounded derivatives in $x \in \mathbb{R}^d$, then for an integer $m \geqslant 0$ and $p \geqslant 2$, the filtering density π is a W^m_p-valued weakly cadlag process, so long as $\mathbb{E}|\pi|^p_{W^m_p} < \infty$, where by W^m_p we mean the space of functions with generalised derivatives in L_p, up to order m. Moreover, π is strongly cadlag as W^s_p-valued process, for $s \in [0, m)$. There have been pioneering works, see [39, 41, 51], as well as some extensions [32, 35, 36], on the regularity of filtering densities for the case of no jumps, i.e. when $\xi = \eta = 0$ in (I.0.2) and if Y contains no jumps. However, to the best of the authors knowledge, these investigations have not previously been extended to systems with Lévy noise. This thesis provides a first extension of such results on regularity (in the Sobolev sense) to jump diffusions.
Our results, both on the conditional distribution P as well as its density process π, are obtained by first studying the Zakai equation. Its linearity makes it easier to analyse in our case and we first obtain results on existence and regularity of the unnormalised conditional density $u = (u_t)_{t\in[0,T]}$, satisfying for each t, $u_t = d\mu_t/dx$ almost surely. Through multiplication with a normalising process we then get the desired results also for π.

For more details on related results, relevant literature and the methods we have employed, we refer the reader to the introductions to Chapters III, IV and V. Moreover, the reader may also consult the articles [16], [17] and [18], which provide the bases for the Chapters III, IV and V, respectively. The articles [16] and [18] are joint work with István Gyöngy, and the article [17] is joint work with Alexander Davie and István Gyöngy.

Structure of the thesis

While the reader will find a more detailed outline of the individual Chapters III, IV and V at their respective beginnings, we include a concise description of them here.

In Chapter II we first collect and review some important results on jump processes and related calculus. While it provides the reader with some background knowledge for results used in this thesis, it is meant to motivate our choice of model, in particular its form (I.0.2). Finally, we state useful results on existence, uniqueness and moment estimates for solutions to (I.0.2).

In Chapter III we derive the filtering equations associated to (I.0.2). We generalise some necessary concepts from filtering theory and, using optional projections, derive first the Zakai equation. By multiplication with a normalising process we then also obtain the Kushner-Shiryaev equation.

In Chapter IV we are concerned with the existence of an L_p-valued density process π. We first generalise some estimates originally presented in [41] and obtain a supremum estimate for the L_p-norm of the smoothed unnormalised conditional density. A limit argument directly yields our result for the case of L_2. To obtain the result for L_p, with general $p \geqslant 2$, we first prove it for the case of compactly supported coefficients and then show the general case by using a limit procedure.

In Chapter V we prove the existence of a W_p^m-valued density process. We show this by relying on results for the L_p-case from Chapter IV and first establishing existence and regularity for compactly supported coefficients. As before, a limit argument yields the desired result for general coefficients.

I.1 Notation

We conclude with some notions and notations used throughout the paper.

Spaces of continuous functions

For an integer $n \geqslant 0$ the notation $C_b^n(\mathbb{R}^d)$ means the space of real-valued bounded continuous functions on $\mathbb{R}^d$, which have bounded and continuous derivatives up to order n. (If $n = 0$, then $C_b^0(\mathbb{R}^d) = C_b(\mathbb{R}^d)$ denotes the space of real-valued bounded continuous functions on $\mathbb{R}^d$). We use the notation $C_0^\infty = C_0^\infty(\mathbb{R}^d)$ for the space of real-valued compactly supported smooth functions on $\mathbb{R}^d$.

Signed measures and related notions

We denote by $\mathbb{M} = \mathbb{M}(\mathbb{R}^d)$ the set of finite Borel measures on $\mathbb{R}^d$ and by $\mathfrak{M} = \mathfrak{M}(\mathbb{R}^d)$ the set of finite signed Borel measures on $\mathbb{R}^d$. For $\mu \in \mathfrak{M}$ we use the notation

$$\mu(\varphi) = \int_{\mathbb{R}^d} \varphi(x)\, \mu(dx)$$

for Borel functions φ on $\mathbb{R}^d$. We say that a function $\nu : \Omega \to \mathbb{M}$ is $\mathcal{G}$-measurable for a σ-algebra $\mathcal{G} \subset \mathcal{F}$, if $\nu(\varphi)$ is a $\mathcal{G}$-measurable random variable for every bounded Borel function φ on $\mathbb{R}^d$. An $\mathfrak{M}$-valued process $(\nu_t)_{t\geqslant 0}$ is said to be adapted to a filtration $(\mathcal{G}_t)_{t\in[0,T]}$ if $\nu_t(\varphi)$ is a $\mathcal{G}_t$-measurable random variable for every $t \in [0, T]$ and bounded Borel function φ on $\mathbb{R}^d$. An $\mathbb{M}$-valued stochastic process $\nu = (\nu_t)_{t\in[0,T]}$ is said to be weakly cadlag if almost surely $\nu_t(\varphi)$ is a cadlag function of t for all $\varphi \in C_b(\mathbb{R}^d)$. An $\mathfrak{M}$-valued process $(\nu_t)_{t\in[0,T]}$ is weakly cadlag, if it is the difference of two $\mathbb{M}$-valued weakly cadlag processes.

σ-algebras and filtrations

A measurable space $(\mathfrak{Z}, \mathcal{Z})$, or a measure space $(\mathfrak{Z}, \mathcal{Z}, \nu)$, is called separable if the σ-algebra $\mathcal{Z}$ is countably generated. For processes $U = (U_t)_{t\in[0,T]}$ we use the notation $\mathcal{F}_t^U$ for the P-completion of the σ-algebra generated by $\{U_s : s \leqslant t\}$. By an abuse of notation, we often write $\mathcal{F}_t^U$ when referring to the filtration $(\mathcal{F}_t^U)_{t\in[0,T]}$, whenever this is clear from the context. For σ-algebras $\mathcal{G}_i \subset \mathcal{F}$, $i = 1, 2$, the notation $\mathcal{G}_1 \vee \mathcal{G}_2$ means the P-completion of the smallest σ-algebra containing G_i for $i = 1, 2$.

Sobolev and Bessel potential spaces

For a measure space $(\mathfrak{Z}, \mathcal{Z}, \nu)$ and $p \geqslant 1$ we use the notation $L_p(\mathfrak{Z})$ for the L_p-space of $\mathbb{R}$-valued $\mathcal{Z}$-measurable processes defined on $\mathfrak{Z}$. However, if not otherwise specified, the function spaces are considered to be over $\mathbb{R}^d$. We always use without mention the summation convention, by which repeated integer valued indices imply a summation. For a multi-index $\alpha = (\alpha_1, \ldots, \alpha_d)$ of nonnegative integers $\alpha_i, i = 1, \ldots, d$, a function φ of $x = (x_1, \ldots, x_d) \in \mathbb{R}^d$ and a nonnegative integer k we use the notation

$$D^\alpha \varphi(x) = D_1^{\alpha_1} D_2^{\alpha_2} \ldots D_d^{\alpha_d} \varphi(x), \quad \text{as well as} \quad |D^k \varphi|^2 = \sum_{|\gamma|=k} |D^\gamma \varphi|^2,$$

where $D_i = \frac{\partial}{\partial x_i}$ and $|\cdot|$ denotes an appropriate norm. We also use the notation $D_{ij} = D_i D_j$. If we want to stress that the derivative is taken in a variable x, we write D_x^α. If the norm $|\cdot|$ is not clear from the context, we sometimes use appropriate subscripts, as in $|\varphi|_{L_p}$ for the $L_p(\mathbb{R}^d)$-norm of φ. For $p \geqslant 1$ and integers $m \geqslant 0$ the space of functions from L_p, whose generalized derivatives up to order m are also in L_p, is denoted by W_p^m. The norm $|f|_{W_p^m}$ of f in W_p^m is defined by

$$|f|^p_{W_p^m} := \sum_{k=0}^m \int_{\mathbb{R}^d} |D^k f(x)|^p \, dx < \infty.$$

For real-valued functions f and g defined on $\mathbb{R}^d$ the notation (f, g) means the Lebesgue integral of fg over $\mathbb{R}^d$ whenever it is well-defined. Throughout the paper we work on the finite time interval $[0, T]$, where $T > 0$ is fixed but arbitrary, as well as on a given complete probability space $(\Omega, \mathcal{F}, P)$ equipped with a filtration $(\mathcal{F}_t)_{t \geqslant 0}$ such that $\mathcal{F}_0$ contains all the P-null sets. For $p, q \geqslant 1$ and integers $m \geqslant 1$ we denote by $\mathbb{W}_p^m = L_p((\Omega, \mathcal{F}_0, P), W_p^m(\mathbb{R}^d))$ and $\mathbb{W}_{p,q}^m \subset L_p(\Omega, L_q([0, T], W_p^m(\mathbb{R}^d)))$ the set of $\mathcal{F}_0 \otimes \mathcal{B}(\mathbb{R}^d)$-measurable real-valued functions $f = f(\omega, x)$ and $\mathcal{F}_t$-optional W_p^m-valued functions $g = g_t(\omega, x)$ such that

$$|f|^p_{\mathbb{W}_p^m} := \mathbb{E}|f|^p_{W_p^m} < \infty \quad \text{and} \quad |g|^p_{\mathbb{W}_{p,q}^m} := \mathbb{E}\Big(\int_0^T |g_t|^q_{W_p^m} dt \Big)^{p/q} < \infty,$$

respectively. If $m = 0$ we set $\mathbb{L}_p = \mathbb{W}_p^0$ and $\mathbb{L}_{p,q} = \mathbb{W}_{p,q}^0$. When instead of $\mathcal{F}_0$ we consider the measurability with respect to another σ-algebra $\mathcal{G}$, we write this explicitly as $\mathbb{L}_p(\mathcal{G})$ or $\mathbb{W}_p^m(\mathcal{G})$. If $m \geqslant 0$ is not an integer and $p > 1$, then W_p^m denotes the space of real-valued functions h on $\mathbb{R}^d$ such that

$$|h|_{W_p^m} := |(1 - \Delta)^{m/2} h|_{L_p} < \infty.$$

Chapter II

Jump processes and related stochastic calculus

In this chapter we collect important notions and results on stochastic calculus related to Lévy processes as well as on stochastic differential equations driven by Lévy noise, i.e., when they are driven by both Wiener processes and Poisson random measures. We start by introducing random measures more generally and finally conclude that the jumps of a cadlag semi-martingale can by expressed in terms of Poisson random measures, yielding the semi-martingale decomposition. We then state the decompositon of a particular semi-martingale, a Lévy process, to motivate the form of stochastic differential equation considered as model in this thesis. More precisely, we outline how a Lévy process can be decomposed into a continuous drift, an integral against a Wiener process and an integral against a Poisson random measure.
Having justified the type of SDE models we investigate, we provide some useful results on existence of solutions to them, their uniqueness, as well as some moment estimates which are used in later sections.
While these results are well-known, we provide some sketches of proofs for the reader's convenience.

Throughout this chapter we consider a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geqslant 0}, P)$ such that $\mathcal{F}_0$ contains all the P-null sets and a separable measurable space $(\mathfrak{Z}, \mathcal{Z})$.

II.1 Lévy processes and random measures

Definition II.1.1. [1, p. 43] An $\mathbb{R}^d$-valued $\mathcal{F}_t$-adapted stochastic process $(X_t)_{t\geqslant 0}$ is called Lévy process if

(i) $X_0 = 0$ almost surely,

(ii) it is stochastically continuous, i.e. for all $\varepsilon > 0$ and $s, t \geqslant 0$,

$$\lim_{t \to s} P(|X_t - X_s| \geqslant \varepsilon) = 0,$$

(iii) it has stationary and independent increments, i.e. for all $t > s \geqslant 0$, the random variable $X_t - X_s$ is independent of the σ-algebra $\mathcal{F}_s$ and has the same distribution as X_{t-s}.

In the following we will refer to an $\mathcal{F}_t$-adapted Lévy process as $\mathcal{F}_t$-Lévy process.

Proposition II.1.1. *[25, Thm. 2.68] Let X be an $\mathcal{F}_t$-Lévy process. Then it has an $\mathcal{F}_t$-adapted cadlag stochastic modification.*

Henceforth we always mean the cadlag modification whenever we introduce a Lévy process. Note that then, the Lévy process is in particular a semi-martingale.

Definition II.1.2. [25, Def. 8.1] An $\mathcal{F}_t$-adapted process X is called a semi-martingale if it admits the decomposition

$$X = M + A,$$

where M is a local $\mathcal{F}_t$-martingale and A is an $\mathcal{F}_t$-adapted process with finite variation.

We say a martingale M is locally square integrable if there exists a sequence of stopping times $(\tau_n)_{n=1}^{\infty}$, such that $\tau_n \to \infty$ and for each $n \geqslant 1$, $(M_{\tau_n \wedge t})_{t \geqslant 0}$ is a square integrable martingale. For two square integrable martingales M_1 and M_2 there exists a unique predictable locally integrable increasing process $\langle M_1, M_2 \rangle = (\langle M_1, M_2 \rangle_t)_{t \geqslant 0}$, such that $M_1 M_2 - \langle M_1, M_2 \rangle$ is a locally square integrable martingale starting at 0. In this case $\langle M_1, M_2 \rangle$ is referred to as Doob-Meyer process of $M_1 M_2$. For the following theorem, we recall further the definition of the quadratic variation of two semi-martingales X, Y,

$$[X, Y] = X_0 Y_0 + \langle X^c, Y^c \rangle + \sum_{s \leqslant \cdot} \Delta X_s \Delta Y_s,$$

where X^c and Y^c denote the continuous (locally square integrable) martingale part of their semi-martingale decomposition (see [25]), $\Delta X_s := X_s - Y_{s-}$, $s \geqslant 0$ with $X_{0-} := X_0$ and ΔY is defined analogously.

Theorem II.1.2. *[25, Thm. 11.43] Let X_1 and X_2 be $\mathcal{F}_t$-Lévy processes. If their quadratic variation $[X_1, X_2] = 0$ almost surely, then they are independent.*

Theorem II.1.3. *[25, Thm. 13.44] Let X be an $\mathcal{F}_t$-Lévy process. Let $(\mathcal{F}_t^X)_{t \geqslant 0}$ be the natural filtration of X and denote by $\mathcal{N}$ the P-null sets. Then*

$$\mathcal{F}_{t-}^X \vee \mathcal{N} = \mathcal{F}_t^X \vee \mathcal{N} = \mathcal{F}_{t+}^X \vee \mathcal{N}, \quad t \geqslant 0,$$

where

$$\mathcal{F}_{t-}^X := \bigvee_{s<t} \mathcal{F}_s^X, \quad \textit{and} \quad \mathcal{F}_{t+}^X := \bigcap_{r>t} \mathcal{F}_r^X.$$

Definition II.1.3. A mapping $\mu : \Omega \times \mathcal{B}(\mathbb{R}_+) \times \mathcal{Z} \mapsto \mathbb{R}_+ \cup \{\infty\}$ is called a random measure if

(i) for all $\omega \in \Omega$, $\mu(\omega, \cdot)$ is a σ-finite measure on $\mathcal{B}(\mathbb{R}_+) \otimes \mathcal{Z}$,

(ii) for all $B \in \mathcal{B}(\mathbb{R}_+) \otimes \mathcal{Z}$, $\mu(\cdot, B)$ is a random variable on $(\Omega, \mathcal{F})$.

We set

$$(\tilde{\Omega}, \tilde{\mathcal{F}}) = (\Omega \times \mathbb{R}_+ \times \mathfrak{Z}, \mathcal{F} \otimes \mathcal{B}(\mathbb{R}_+) \otimes \mathcal{Z}).$$

Define further

$$\tilde{\mathcal{O}} = \mathcal{O} \otimes \mathcal{Z} \quad \text{and} \quad \tilde{\mathcal{P}} = \mathcal{P} \otimes \mathcal{Z},$$

referred to as optional and predictable σ-algebra on $\tilde{\Omega}$ respectively, where $\mathcal{O}$ and $\mathcal{P}$ are the usual optional and predictable σ-algebras on $\Omega \times \mathbb{R}_+$. In accordance with this, we call a measurable function f on $(\tilde{\Omega}, \tilde{\mathcal{O}})$ (resp. $(\tilde{\Omega}, \tilde{\mathcal{P}})$) optional (resp. predictable).

In the usual way, we define the stochastic integral of an optional function f on $(\tilde{\Omega}, \tilde{\mathcal{O}})$ against μ as (see [28, equ. II.1.6])

$$(f * \mu)_t = \begin{cases} \int_0^t \int_{\mathfrak{Z}} f(s, \mathfrak{z})\, \mu(d\mathfrak{z}, ds), & \text{if } \int_0^t \int_{\mathfrak{Z}} |f(s, \mathfrak{z})|\, \mu(d\mathfrak{z}, ds) < \infty \\ \infty, & \text{otherwise} \end{cases} \qquad t \geqslant 0.$$

Definition II.1.4. [28, Def. II.1.6] A random measure is called optional (resp. predictable) if for every optional (resp. predictable) function f on $(\tilde{\Omega}, \mathcal{O})$ (resp. $(\tilde{\Omega}, \mathcal{P})$) the process

$$(f * \mu)_t = \int_0^t \int_{\mathfrak{Z}} f(s, \mathfrak{z})\, \mu(d\mathfrak{z}, ds), \quad t \geqslant 0,$$

is optional (resp. predictable). Moreover, we say that μ is finite, if $\mu(\tilde{\Omega}) < \infty$. Finally, a random measure μ is $\tilde{\mathcal{P}}$-σ-finite, or predictably σ-finite, if there exists a $\tilde{\mathcal{P}}$-measurable partition $(A_n)_{n=1}^\infty$ of $\tilde{\Omega}$, such that $\mathbb{E}(\mathbb{1}_{A_n} * \mu)_\infty < \infty$ for each n.

Theorem II.1.4. *[28, Thm. II.1.8] Let μ be a $\tilde{\mathcal{P}}$-σ-finite optional random measure on $(\tilde{\Omega}, \tilde{\mathcal{F}})$. Then there exists a predictable random measure μ^p on $(\tilde{\Omega}, \tilde{\mathcal{P}})$, unique up to a P-null set, such that for every nonnegative $\tilde{\mathcal{P}}$-measurable function f we have*

$$\mathbb{E} \int_0^\infty \int_{\mathfrak{Z}} f(t, \mathfrak{z})\, \mu(\omega, d\mathfrak{z}, dt) = \mathbb{E} \int_0^\infty \int_{\mathfrak{Z}} f(t, \mathfrak{z})\, \mu^p(\omega, d\mathfrak{z}, dt).$$

In this case, μ^p is referred to as (predictable) compensator of μ.

We recall a few properties of μ and μ^p. If μ is already a predictably σ-finite random measure, then $\mu^p = \mu$ almost surely. Moreover, for any $\tilde{\mathcal{P}}$-measurable function f on $\tilde{\Omega}$, such that $|f| * \mu$ is a locally integrable increasing process, we have that $f * \mu^p$ is the compensator of $f * \mu$ and therefore, with the random measure $\tilde{\mu} := \mu - \mu^p$ we have that

$$\int_0^t \int_{\mathfrak{Z}} f(s, \mathfrak{z}) \, \mu(d\mathfrak{z}, ds) - \int_0^t \int_{\mathfrak{Z}} f(s, \mathfrak{z}) \, \mu^p(d\mathfrak{z}, ds)$$

is a local martingale (see [28, Prop. II.1.28].

The random measures in this theses will only be integer-valued random measures, as well as integrals against the latter, so that we henceforth focus our exposition thereon.

Integer-valued random measures

Definition II.1.5. [25, Def. II.1.13] We say that an optional $\tilde{\mathcal{P}}$-σ-finite random measure μ is an integer-valued random measure (on $(\tilde{\Omega}, \mathcal{O})$), if

(i) $\mu(\omega, \mathfrak{Z}, \{t\}) \in \{0, 1\}$ for all $(\omega, t) \in \Omega \times \mathbb{R}_+$,

(ii) for each $A \in \mathcal{B}(\mathbb{R}_+) \otimes \mathcal{Z}$ the random variable $\mu(\omega, A)$ takes values in $\mathbb{N} \cup \{\infty\}$.

For a stopping time τ we denote by $[\![\tau]\!]$ the graph of τ, that is $[\![\tau]\!] = \{(\omega, \tau(\omega)) \in \Omega \times \mathbb{R}_+ : \omega \in \Omega\}$.

If μ is an integer-valued random measure, then there exists a sequence of stopping times $(\tau_n)_{n=1}^{\infty}$, satisfying $[\![\tau_n]\!] \cap [\![\tau_m]\!] = \varnothing$ for $n \neq m$, and a $\mathfrak{Z}$-valued optional process $\beta = (\beta_t)_{t \geqslant 0}$ such that with $D = \bigcup_{n=1}^{\infty} [\![\tau_n]\!]$ we have

$$\mu(\omega, d\mathfrak{z}, dt) = \sum_{s \geqslant 0} \mathbb{1}_D(\omega, s) \delta_{(\beta_s(\omega), s)}(d\mathfrak{z}, ds), \tag{II.1.1}$$

with δ denoting the Dirac measure.

Example II.1.1. [28, Prop. II.1.16] Consider an $\mathcal{F}_t$-adapted cadlad $\mathbb{R}^d$-valued process $X = (X_t)_{t \geqslant 0}$ and consider the "jump-measure" μ^X associated to X, i.e. for $\Delta X_t := X_t - X_{t-}$, $t \geqslant 0$ and $X_{0-} := X_0$,

$$\mu^X(\omega, A \times (s, t)) = |\{\Delta X_r(\omega) \in A : r \in (s, t)\}|. \tag{II.1.2}$$

Then we can write μ^X as in (II.1.1) with $D := [\Delta X \neq 0] \subset \Omega \times \mathbb{R}_+$, $\beta := \Delta X$ and $\mathfrak{Z} := \mathbb{R}^d$.

If $g = g(\omega, t, x)$ is an optional real-valued function and μ an integer-valued random measure with D, $(\tau_n)_{n=1}^{\infty}$ and β as described above, then we can write, see [28, equ. 1.15],

$$\int_0^t \int_A g(s, x) \, \mu(dx, ds) = \sum_{n=1}^{\infty} \mathbb{1}_{A \times [0,t]}(\beta_n, \tau_n) g(\beta_{\tau_n}, \tau_n) = \sum_{s \leqslant t} \mathbb{1}_D(s) \mathbb{1}_A(\beta_s) g(\beta_s, s).$$

If a set A is such that $0 \notin \bar{A}$, the closure of A, then we say it is bounded below.

Proposition II.1.5. *[1, Lemma 2.3.4] Consider X and μ^X as in Example II.1.1. If $A \in \mathcal{B}(\mathbb{R}^d)$ is bounded below, then we have $\mu^X(A \times (0,t]) < \infty$ almost surely for all $t \geqslant 0$.*

Proof. Define the sequence of stopping times

$$\tau_n := \inf\{t \geqslant 0 : \mu(A \times (0,t]) \geqslant n\},$$

$n \geqslant 1$, where we observe that for $t \geqslant 0$,

$$\Delta\mu^X(A \times (0,t]) > 0 \quad \Leftrightarrow \quad \Delta X_t \in A.$$

First, note that due to the continuity of X from the right at 0, almost surely $\tau_1 > 0$. Similarly we see that $\tau_n \to \infty$ as $n \to \infty$. If not, and $\tau_n \to \tau < \infty$ on some set Ω' of positive probability, X would not have left limits at τ on Ω'. Thus for each $t \geqslant 0$ we have

$$\mu^X(A \times (0,t]) = \sum_{n=1}^{\infty} \mathbb{1}_{(0,\tau_n]}(t) < \infty, \quad \text{(a.s.)}.$$

$\square$

Though we do not use the following theorem, Theorem II.1.6, in its full generality, but only for Lévy processes appearing in later sections, we include it here to outline the role of random measures in semi-martingale decompositions. The respective special case of Theorem II.1.6, the decomposition of Lévy processes presented later in this section, motivates our choice of model in this thesis.

Theorem II.1.6. *[25, Thm. 11.25] Let $X = (X_t)_{t\geqslant 0}$ be an $\mathbb{R}^d$-valued semi-martingale, let μ^X be its jump measure and ν^X the predictable compensator of μ^X. The X can be written as*

$$X_t = X_0 + \alpha_t + X_t^c + \int_0^t \int_{|x|\leqslant 1} x\,(\mu^X - \nu^X)(dx, ds) + \int_0^t \int_{|x|>1} x\,\mu^X(dx, ds), \quad \text{(II.1.3)}$$

where $(\alpha_t)_{t\geqslant 0}$ is a predictable process with finite variation and M^c is the continuous (locally square integrable) martingale part of X with $M_0^c = \alpha_0 = 0$. Moreover, we have the properties

(i) $\nu^X(\mathbb{R}^d \times \{0\}) = \nu^X(\{0\} \times \mathbb{R}_+) = 0,$

(ii) $((|x|^2 \wedge 1) * \nu_t^X)_{t\geqslant 0}$ *is increasing with locally integrable variation,*

(iii) $\Delta\alpha_t = \int_{|x|\leqslant 1} x\,\nu^X(dx, \{t\}).$

Definition II.1.6. Consider a semi-martingale X with representation as in (II.1.3) of Theorem II.1.6. Then we call the processes α, the Doob-Meyer process $\beta = \langle X^c \rangle$ and the predictable compensator ν^X the predictable characteristics of X, or simply characteristics in short, often written as triple (α, β, ν^X).

Proposition II.1.7. *[25, Cor. 11.28] A semi-martingale with predictable characteristics (α, β, ν^X) is stochastically continuous if and only if for all $t \geqslant 0$ we have $\nu^X(\mathbb{R}^d \times \{t\}) = 0$ almost surely.*

Similarly, if we consider a cadlag process with stationary and independent increments, then one can see that the distribution of the random variables ΔX_t, with $t \geqslant 0$ does not depend on t and hence X cannot have a fixed time of discontinuity.

Theorem II.1.8. *[25, Thm. 11.36] A stochastically continuous semimartingale X is an $\mathcal{F}_t$-Lévy process if and only if its predictable characteristics (α, β, ν^X) are non-random. Then we have moreover that*

(i) α is continuous with finite variation and

(ii) β is continuous and monotone increasing, with $\beta_0 = 0$.

We recall that a random variable N is Poisson distributed with intensity λ if

$$P(N = k) = \frac{\lambda^k \exp(-k)}{k!}, \quad \text{for } k = 0, 1, 2, \ldots.$$

Analogously we say that $N = (N_t)_{t \geqslant 0}$ is a Poisson process with intensity λ if N is a Lévy process taking values in $\mathbb{N} \cup \{0\}$ and if for each t the random variable N_t has a Poisson distribution with intensity $t\lambda$.

Definition II.1.7. [30, Def. 19.1] Consider a σ-finite measure space $(H, \mathcal{H}, \tilde{\nu})$. A family of $\mathbb{N} \cup \{\infty\}$-valued random variables $(N(A))_{A \in \mathcal{F}}$ is called a Poisson random measure (on H) with characteristic measure $\tilde{\nu}$ if

(i) for each $A \in \mathcal{H}$ the random variable $N(A)$ has a Poisson distribution with intensity $\tilde{\nu}(A)$,

(ii) $N(A_1)$ and $N(A_2)$ are independent whenever $A_1 \cap A_2 = \varnothing$ and

(iii) for each $\omega \in \Omega$, $N(\omega, \cdot)$ is a measure on H.

Proposition II.1.9. *[30, Prop. 19.4] Consider a σ-finite measure space $(H, \mathcal{H}, \tilde{\nu})$. Then there exists a probability space $(\Omega', \mathcal{F}', P')$ and a random measure N on H, such that it is a Poisson random measure with characteristic measure $\tilde{\nu}$.*

We consider now the case when $(H, \mathcal{H}) = (\mathbb{R}_+ \times \mathbb{R}^d, \mathcal{B}(\mathbb{R}_+ \times \mathbb{R}^d))$ and show that the jumps of a Lévy process can be described by a Poisson random measure.

Theorem II.1.10. *Consider an $\mathbb{R}^d$-valued $\mathcal{F}_t$-Lévy process X on $\Omega \times \mathbb{R}_+$ and let N denote the measure of its jumps. Then N is a Poisson random measure on $(\mathbb{R}_+ \times (\mathbb{R}^d \backslash \{0\}), \mathcal{B}(\mathbb{R}_+) \otimes \mathcal{B}(\mathbb{R}^d \backslash \{0\}))$, such that with a σ-finite measure ν, the characteristic measure of N is given by $\tilde{\nu}(dx, dt) = \nu(dx) \otimes dt$.*

Proof. We provide a rough sketch of a proof and refer to [1, Thm. 2.2.13] for full details. For a set $A \in \mathcal{B}(\mathbb{R}_+)$ with $0 \notin \bar{A}$, $(N_t^A)_{t \geqslant 0} := (N(A \times (0, t]))_{t \geqslant 0}$ is given by the right-hand side of (II.1.2), with $s := 0$. First, one can show that N^A is also a Lévy process. Clearly $N_0^A = 0$ almost surely and it is not difficult to see that, due to the independent and stationary increments of X and its stochastic continuity, N^A shares these properties. Next we show that N^A is a Poisson process. Taking a version of X with cadlag sample paths, we know that X has at most countably many jumps of size in A, to which we assign the ordered set of stopping times

$$\tau_1 := \inf\{t \geqslant 0 : \Delta X_t \in A\}, \quad \tau_n := \inf\{t > \tau_{n-1} : \Delta X_t \in A\}, \quad n \geqslant 2.$$

Using the stationary and independent increments of X, one can show that the random variables

$$(\tau_1, \tau_2 - \tau_1, \tau_3 - \tau_2, \dots)$$

are independent and identically distributed. Further, we see immediately that $f_1(t) := P(\tau_1 > t)$ is decreasing. Moreover, $f_1(0) = 1$, as otherwise there would be a set F of positive probability such that on F we have $\Delta X_0 > 0$. By the same argument one can see that f is right-continuous at 0. Further, by due to the independent and stationary increments,

$$f_1(t + s) = P(N_s^A = 0 | N_{t+s}^A - N_s^A = 0) = f_1(t) f_1(s).$$

Hence one can deduce that there exists $\lambda > 0$ such that $f_1(t) = \exp(-\lambda t)$, and that as sum of exponentially and identically distributed random variables,

$$\tau_n = \tau_1 + (\tau_2 - \tau_1) + \cdots$$

has a gamma distribution. Finally, one can use induction to see that N^A is a Poisson process with intensity $t\lambda$. □

If the compensator $\tilde{\nu}$ of a Poisson random measure N on $(\mathbb{R}_+ \times \mathfrak{Z}, \mathcal{B}(\mathbb{R}_+) \otimes \mathcal{Z})$ admits the decomposition $\tilde{\nu}(d\mathfrak{z}, dt) = \nu(d\mathfrak{z}) \otimes dt$ for a σ-finite measure ν on $(\mathfrak{Z}, \mathcal{Z})$, then we also refer to ν as the characteristic measure of N. This will always be clear from the context. Moreover, we often write $\nu(d\mathfrak{z}) \otimes dt = \nu(d\mathfrak{z}) dt$.

Theorem II.1.11. *[25, Thm. 11.45] [1, Thm. 2.4.16] Let X be an $\mathcal{F}_t$-Lévy process, $N(ds, dx)$ its associated jump measure, with predictable compensator $ds \otimes \nu(dx)$ and compensated martingale measure $\tilde{N}(dx, dt) = N(dx, dt) - \nu(dx) dt$.*

Then it admits the representation

$$X_t = \alpha t + \int_0^t \sigma^k \, dW_s^k + \int_0^t \int_{|x|<1} x \, \tilde{N}(dx, ds) + \int_0^t \int_{|x|\geqslant 1} x \, N(dx, ds), \quad t \geqslant 0,$$

where $\alpha \in \mathbb{R}^d$*,* $\sigma = \sigma^{ij} \in \mathbb{R}^{d\times m}$ *for some* $m \geqslant 1$ *and* W *an* m*-dimensional Wiener process.*

Finally, we see that a Lévy process X exhibits a certain structure that allows us to analyse it efficiently. This decomposition motivates our choice of model (I.0.2) in investigating jump-diffusion processes, i.e., SDEs driven by Lévy processes.

We finish with some integral properties of Poisson (martingale) random measures. For details and proofs we refer the reader to [27], in particular Section II.3 therein.

Proposition II.1.12. *Let* $f = f(\omega, t, \mathfrak{z})$ *be a* $\tilde{\mathcal{P}}$*-measurable real-valued function and let moreover* $N(ds, d\mathfrak{z})$ *be a Poisson random measure on* $(\mathbb{R}_+ \times \mathfrak{Z}, \mathcal{B}(\mathbb{R}_+) \otimes \mathcal{Z})$ *with compensator* $\nu(d\mathfrak{z})dt$*. Then we have the following properties.*
(i) If for $T > 0$*,*

$$\mathbb{E} \int_0^T \int_{\mathfrak{Z}} |f(s, \mathfrak{z})|^q \, \nu(d\mathfrak{z}) ds < \infty, \tag{II.1.4}$$

with $q := 1$*, then we have*

$$\mathbb{E} \int_0^T \int_{\mathbb{R}^d} |f(s, \mathfrak{z})| \, N(d\mathfrak{z}, ds) = \mathbb{E} \int_0^T \int_{\mathfrak{Z}} |f(s, \mathfrak{z})| \, \nu(d\mathfrak{z}) ds.$$

Moreover, we know that

$$\int_0^t \int_{\mathfrak{Z}} f(s, \mathfrak{z}) \, \tilde{N}(d\mathfrak{z}, ds) := \int_0^t \int_{\mathfrak{Z}} f(s, \mathfrak{z}) \, N(d\mathfrak{z}, ds) - \int_0^t \int_{\mathfrak{Z}} f(s, \mathfrak{z}) \, \nu(d\mathfrak{z}) ds \tag{II.1.5}$$

is an $\mathcal{F}_t$*-martingale.*
(ii) If for $q := 1$ *and* $q := 2$ (II.1.4) *holds, then the right-hand side of* (II.1.5) *is a square integrable* $\mathcal{F}_t$*-martingale and the Doob-Meyer process*

$$\Big\langle \int_0^{\cdot} \int_{\mathfrak{Z}} f(s, \mathfrak{z}) \, \tilde{N}(d\mathfrak{z}, ds) \Big\rangle_t = \int_0^t \int_{\mathfrak{Z}} |f(s, \mathfrak{z})|^2 \, \nu(d\mathfrak{z}) ds, \quad t \in [0, T].$$

(iii) If (II.1.4) *holds with* $q := 2$*, then there exists a sequence of* $\tilde{\mathcal{P}}$*-measurable real-valued functions* $(f_n)_{n=1}^{\infty}$*, such that for each* f_n *we have* (II.1.4) *with* $q := 1, 2$ *and* f_n *in place of* f*,* $f_n \to f$ *as* $n \to \infty$ *for* $P \otimes dt \otimes d\mathfrak{z}$*-almost every* $(\omega, t, \mathfrak{z}) \in \Omega \times \mathbb{R}_+ \times \mathfrak{Z}$ *and*

$$\int_0^t \int_{\mathfrak{Z}} f_n(s, \mathfrak{z}) \, \tilde{N}(d\mathfrak{z}, ds), \quad t \geqslant 0, \quad n = 1, 2, \ldots,$$

is a Cauchy sequence in $\mathcal{M}_2$, the space of square integrable $\mathcal{F}_t$-martingales. Then we define its limit by

$$\int_0^t \int_{\mathfrak{Z}} f(s,\mathfrak{z})\, \tilde{N}(d\mathfrak{z}, ds), \quad t \geqslant 0. \tag{II.1.6}$$

(iv) Finally, if f is such that for a sequence of stopping times $(\rho_n)_{n=1}^{\infty}$ going to ∞, for each $n \geqslant 1$ (II.1.4) *holds for $q := 2$ and $f\mathbb{1}_{t\leqslant\rho_n}$ in place of f, then* (II.1.6) *is defined as the unique locally square integrable $\mathcal{F}_t$-martingale M, such that for each $n \geqslant 1$,*

$$M_{t\wedge\rho_n} = \int_0^t \int_{\mathfrak{Z}} \mathbb{1}_{s\leqslant\rho_n} f(s,\mathfrak{z})\, \tilde{N}(d\mathfrak{z}, ds), \quad t \in [0,T].$$

It is worth noting that in case (iii) above the equality (II.1.5) may no longer hold, as the integrals on the right-hand side thereof may not be well-defined.

II.2 Some results on Itô-Lévy processes

Consider again a complete filtered probability space $(\Omega, (\mathcal{F}_t)_{t\geqslant 0}, P)$, a separable σ-finite measure space $(\mathfrak{Z}, \mathcal{Z}, \nu)$ and the stochastic differential equation

$$dZ_t = b(t,Z_t)\,dt + \sigma(t,Z_t)\,dW_t + \int_{\mathfrak{Z}} \eta(t, Z_{t-}, \mathfrak{z})\, \tilde{N}(d\mathfrak{z}, dt), \tag{II.2.7}$$

where $b = (b^i)$, $\sigma = (\sigma^{ij})$ are $\mathcal{B}(\mathbb{R}_+)\otimes\mathcal{B}(\mathbb{R}^d)$ measurable functions, $\eta = (\eta^i)$ and is a $\mathcal{B}(\mathbb{R}_+)\otimes\mathcal{B}(\mathbb{R}^d)\otimes\mathcal{Z}$-measurable function, for $i = 1,\dots,d$, $j = 1,\dots,d_1$, W is a d_1-dimensional $\mathcal{F}_t$-Wiener process, $\tilde{N}(d\mathfrak{z}, dt) := N(d\mathfrak{z}, dt) - \nu(d\mathfrak{z})dt$ with N an $\mathcal{F}_t$-Poisson random measure with characteristic measure $\nu(d\mathfrak{z})dt$, the initial condition Z_0 is $\mathbb{R}^d$-valued and $\mathcal{F}_0$-measurable and Z_0, W and N are independent. We impose the following assumption on the coefficients.

Assumption II.2.1. There are nonnegative constants K_0, K_1 and L such that

(i) for all $t \geqslant 0$ and $z \in \mathbb{R}^d$,

$$|b(t,z)|^2 + |\sigma(t,z)|^2 + \int_{\mathfrak{Z}} |\eta(t,z,\mathfrak{z})|^2\, \nu(d\mathfrak{z}) \leqslant K_0 + K_1|z|^2,$$

(ii) for all $t \geqslant 0$ and $z_1, z_2 \in \mathbb{R}^d$

$$|b(t,z_1) - b(t,z_2)|^2 + |\sigma(t,z_1) - \sigma(t,z_2)|^2$$
$$+ \int_{\mathfrak{Z}} |\eta(t,z_1,\mathfrak{z}) - \eta(t,z_2,\mathfrak{z})|^2\, \nu(d\mathfrak{z}) \leqslant L|z_1 - z_2|^2.$$

Though the following theorem is well known, we provide a partial proof of it for the reader's convenience. For the full proof we refer the reader to [1], or to [20] for a more general exposition.

Theorem II.2.1. *Let Assumption II.2.1 hold. Then there exists a unique $\mathcal{F}_t$-adapted cadlag solution Z to* (II.2.7)*.*

Proof. We only outline the case $\mathbb{E}|Z_0|^2 < \infty$. For the other case, a truncation argument can be applied and we refer the reader to [1] for full details. We define a sequence by $Z_t^{(0)} \equiv Z_0$ and

$$Z_t^{(n+1)} = b(t, Z_t^{(n)})\, dt + \sigma(t, Z_t^{(n)})\, dW_t + \int_{\mathfrak{Z}} \eta(t, Z_{t-}^{(n)}, \mathfrak{z})\, \tilde{N}(d\mathfrak{z}, dt), \quad n = 1, 2, \dots$$

Clearly (a modification of) Z_1 is cadlag and by an inductive argument we see that this holds for $Z^{(n)}$ for all $n \geqslant 1$. First observe that by Doob's martingale inequality,

$$\mathbb{E} \sup_{0 \leqslant s \leqslant t} \left| \int_0^s \sigma(r, Z_0)\, dW_r \right|^2 \leqslant 4\mathbb{E} \int_0^t |\sigma(r, Z_0)|^2\, dr,$$

$$\mathbb{E} \sup_{0 \leqslant s \leqslant t} \left| \int_0^s \int_{\mathfrak{Z}} \eta(r, Z_0, \mathfrak{z})\, \tilde{N}(d\mathfrak{z}, dr) \right|^2 \leqslant 4\mathbb{E} \int_0^t \int_{\mathfrak{Z}} |\eta(r, Z_0, \mathfrak{z})|^2\, \nu(d\mathfrak{z}) dr$$

and hence, due to the linear growth conditions on the coefficients,

$$\begin{aligned}
\mathbb{E} \sup_{0 \leqslant s \leqslant t} |Z_t^{(1)} - Z_t^{(0)}|^2 &\leqslant 2\mathbb{E} \sup_{0 \leqslant s \leqslant t} \left| \int_0^s b(r, Z_0)|\, dr \right|^2 + 2\mathbb{E} \sup_{0 \leqslant s \leqslant t} \left| \int_0^s \sigma(r, Z_0)\, dW_r \right|^2 \\
&\quad + 2\mathbb{E} \sup_{0 \leqslant s \leqslant t} \left| \int_0^s \int_{\mathfrak{Z}} \eta(r, Z_0, \mathfrak{z})\, \tilde{N}(d\mathfrak{z}, dr) \right|^2 \\
&\leqslant 2t^2(K_0 + K_1 \mathbb{E}|Z_0|^2) + 16t(K_0 + K_1 \mathbb{E}|Z_0|^2) \leqslant N(t)t(K_0 + K_1 \mathbb{E}|Z_0|^2),
\end{aligned}$$

with $N(t) = \max\{2t, 16\}$. By induction we get that for each n, $\mathbb{E} \sup_{0 \leqslant s \leqslant t} |Z_s^{(n)}|^2 < \infty$ and using also the Lipschitz property of the coefficients, in a similar way, using also Fubini's theorem

$$\begin{aligned}
\mathbb{E} \sup_{0 \leqslant s \leqslant t} |Z_s^{(n+1)} - Z_s^{(n)}|^2 &\leqslant 2\mathbb{E} \sup_{0 \leqslant s \leqslant t} \left| \int_0^s (b(r, Z_r^{(n)}) - b(r, Z_r^{(n-1)}))\, dr \right|^2 \\
&\quad + 2\mathbb{E} \sup_{0 \leqslant s \leqslant t} \left| \int_0^s (\sigma(r, Z_r^{(n)}) - \sigma(r, Z_r^{(n-1)}))\, dW_r \right|^2 \\
&\quad + 2\mathbb{E} \sup_{0 \leqslant s \leqslant t} \left| \int_0^s \int_{\mathfrak{Z}} (\eta(r, Z_{r-}^{(n)}, \mathfrak{z}) - \eta(r, Z_{r-}^{(n-1)}, \mathfrak{z}))\, \tilde{N}(d\mathfrak{z}, dr) \right|^2 \\
&\leqslant N(t)L\mathbb{E} \int_0^t \sup_{0 \leqslant r \leqslant s} |Z_r^{(n)} - Z_r^{(n-1)}|^2\, dr = N(t)L \int_0^t \mathbb{E} \sup_{0 \leqslant r \leqslant s} |Z_r^{(n)} - Z_r^{(n-1)}|^2\, dr.
\end{aligned}$$

Inductively we then get that

$$\mathbb{E}\sup_{0\leqslant s\leqslant t}|Z_s^{(n+1)}-Z_s^{(n)}|^2 \leqslant N^{n+1}(t)L^n\frac{t^{n+1}}{(n+1)!}(K_0+K_1\mathbb{E}|Z_0|^2),$$

which converges to 0 as $n\to\infty$. Using the triangle inequality it is then easy to show that

$$\mathbb{E}|Z_t^{(n)}-Z_t^{(m)}|^2\to 0,\quad \text{for each } t\geqslant 0 \text{ as } n,m\to\infty,$$

i.e. $(Z_t^{(n)})_{n=1}^{\infty}$ is a Cauchy sequence in $L_2(\Omega)$ and we denote Z as the limit process. Indeed, using the same kind of estimates, it is not difficult to show that Z is the almost surely limit of the cadlag processes $(Z_t^{(n)})_{n=1}^{\infty}$ and hence itself cadlag. It remains to see that Z satisfies the SDE (II.2.7). To see this, one can define a process $\tilde{Z}$ as the right-hand side of (II.2.7), with Z as the limit process constructed above and use the same arguments to verify that for each t we have $\lim_{n\to\infty} Z_t^{(n)}=\tilde{Z}_t$ in mean square. By the uniqueness of the limit it then follows that $\tilde{Z}=Z$. Uniqueness of solution can be by shown by standard techniques. deriving similar estimates as above for the difference of two solutions Z^1-Z^2 and finally applying Gronwall's lemma. □

Assumption II.2.2. For a $p\geqslant 2$ and nonnegative constants K_0, K_1 let η satisfy, for all $t\geqslant 0$, $z\in\mathbb{R}^d$,

$$\int_{\mathfrak{Z}}|\eta(t,z,\mathfrak{z})|^p\,\nu(d\mathfrak{z})\leqslant K_0+K_1|z|^p.$$

The following can be found in [12, Lm. 2.2].

Theorem II.2.2. *Let Assumptions II.2.1 and II.2.2 hold for a $p\geqslant 2$ and let $\mathbb{E}|Z_0|^p<\infty$. Then the solution Z of* (II.2.7) *satisfies*

$$\mathbb{E}\sup_{0\leqslant t\leqslant T}|Z_t|^p\leqslant N(1+\mathbb{E}|Z_0|^p),$$

for a constant $N=N(d,d_1,p,K_0,K_1,T)$.

We provide a sketch of the proof.

Proof. First, we know by Theorem II.2.1 that Z is the unique $\mathcal{F}_t$-adapted cadlag process satisfying (II.2.7). Fix a $T>0$. By Itô's formula, see [22, Cor. 2.4], we have almost surely

$$\begin{aligned}|Z_t|^p &= |Z_0|^p + p\int_0^t |Z_s|^{p-2}Z_s^i\sigma_s^{ij}\,dW_s^j\\ &\quad+\tfrac{p}{2}\int_0^t\Big(2|Z_s|^{p-2}Z_s^ib^i_{\tau_n\wedge s}+(p-2)|Z_s|^{p-4}\sum_{j=1}^{d_1}|Z_s^i\sigma_s^{ij}|^2\Big)\,ds \qquad \text{(II.2.8)}\end{aligned}$$

$$
\begin{aligned}
&+ \tfrac{p}{2}\int_0^t |Z_s|^{p-2}\sum_{i=1}^{d}\sum_{j=1}^{d_1}|\sigma^{ij}_s|^2\,ds + p\int_0^t\int_{\mathfrak{Z}} |Z_{s-}|^{p-2} Z^i_{s-}\eta^i_s(\mathfrak{z})\,\tilde N(d\mathfrak{z},ds)\\
&+ \int_0^t\int_{\mathfrak{Z}} \big(|Z_{s-}+\eta_s(\mathfrak{z})|^p - |Z_{s-}|^p - p|Z_{s-}|^{p-2}Z^i_{s-}\eta^i_s(\mathfrak{z})\big)\,N(d\mathfrak{z},ds)
\end{aligned}
$$

for all $t \in [0,T]$, where we define, for $t \geqslant 0$ and $\mathfrak{z} \in \mathfrak{Z}$,

$$
b_s := b(s,Z_s), \quad \sigma_s := \sigma(s,Z_s), \quad \eta_s(\mathfrak{z}) := \eta(s,Z_{s-},\mathfrak{z}).
$$

Due to the linear growth conditions on the coefficients, by using Young's inequality, the third and fourth terms in (II.2.8) can be estimated by

$$
N + N\int_0^t |Z_s|^p\,ds \tag{II.2.9}
$$

for a constant $N = N(d,d_1,p,K_0,K_1,T)$. For the fourth term in (II.2.8), we can use Taylor's formula to rewrite its integrand as

$$
\begin{aligned}
0 \leqslant \int_0^1 (1-\theta)p\big((p-2)|Z_{s-}|^{p-4}Z^i_{s-}Z^j_{s-} + |Z_{s-}|^{p-2}\delta_{ij}\big)\eta^i_s(\mathfrak{z})\eta^j_s(\mathfrak{z})\,d\theta\\
:= A_s(\mathfrak{z})
\end{aligned}
$$

where δ_{ij} denoted the Dirac delta symbol and where we nonnegativity stems from the fact that, with $p \geqslant 2$, we have $|a+b|^p - |b|^p - p|a|^{p-2}ab \geqslant 0$. Hence, applying Proposition II.1.12 and using Assumptions II.2.1 and II.2.2 as well as Young's inequality, we have, with the stopping times $\tau_n := \inf\{t \geqslant 0 : |Z_t|^p \geqslant n\}$, for each $n \geqslant 1$,

$$
\mathbb{E}\int_0^t\int_{\mathfrak{Z}} A_{\tau_n\wedge s}(\mathfrak{z})\,N(d\mathfrak{z},ds) = \mathbb{E}\int_0^t\int_{\mathfrak{Z}} A_{\tau_n\wedge s}(\mathfrak{z})\,\nu(d\mathfrak{z})ds \leqslant N + N\mathbb{E}\int_0^t |Z_{\tau_n\wedge s}|^p\,ds,
$$

for a constant $N = N(d,d_1,p,K_0,K_1,T)$. Moreover, one can see that for each $n \geqslant 1$, the second and fourth term in (II.2.8) are local martingales, which disappear after taking the expectation, using the stopping times τ_n. Hence, from (II.2.8) we get, for (another) constant $N = N(d,d_1,p,K_0,K_1,T)$,

$$
\mathbb{E}|Z_{\tau_n\wedge t}|^p \leqslant \mathbb{E}|Z_0|^p + N + N\int_0^t \mathbb{E}|Z_{\tau_n\wedge s}|^p\,ds,
$$

so that by Gronwall's inequality and Fatou's lemma, for a constant N' depending only on d, d_1, p, K_0, K_1 and T,

$$
\sup_{t\in[0,T]} \mathbb{E}|Z_t|^p \leqslant N'\mathbb{E}|Z_0|^p.
$$

Further, for the martingale terms in (II.2.8), by Doob's and Young's inequalities,

$$\mathbb{E}\sup_{0\leqslant s\leqslant t\wedge\tau_n}\Big|\int_0^s\int_{\mathfrak{Z}}|Z_{r-}|^{p-2}Z^i_{r-}\eta^i_r(\mathfrak{z})\,\tilde{N}(d\mathfrak{z},dr)\Big|$$

$$\leqslant\mathbb{E}\Big(\int_0^{t\wedge\tau_n}\int_{\mathfrak{Z}}|Z_{s-}|^{2p-2}|\eta_s(\mathfrak{z})|^2\,\nu(d\mathfrak{z})ds\Big)^{1/2}\leqslant\mathbb{E}\Big(\int_0^t\int_{\mathfrak{Z}}|Z_{\tau_n\wedge s}|^{2p-2}|\eta_{\tau_n\wedge s}(\mathfrak{z})|^2\,\nu(d\mathfrak{z})ds\Big)^{1/2}$$

$$\leqslant N+N\mathbb{E}\Big(\int_0^t|Z_{\tau_n\wedge s}|^{2p}\,ds\Big)^{1/2}\leqslant N+N\mathbb{E}(\sup_{0\leqslant s\leqslant t}|Z_{\tau_n\wedge s}|^p)^{1/2}\Big(\int_0^t|Z_{\tau_n\wedge s}|^p\,ds\Big)^{1/2}$$

$$\leqslant N+\tfrac{1}{4}\mathbb{E}\sup_{0\leqslant s\leqslant t}|Z_{\tau_n\wedge s}|^p+4N\mathbb{E}\int_0^t|Z_{\tau_n\wedge s}|^p\,ds,$$

for a constant $N=N(p,K_0,K_1,T)$. Similarly, also for the second term in (II.2.9),

$$\mathbb{E}\sup_{0\leqslant s\leqslant t\wedge\tau_n}\Big|\int_0^s|Z_s|^{p-2}Z^i_s\sigma^{ij}_s\,dW^j_s\Big|\leqslant N+\tfrac{1}{4}\mathbb{E}\sup_{0\leqslant s\leqslant t}|Z_{\tau_n\wedge s}|^p+4N\mathbb{E}\int_0^t|Z_{\tau_n\wedge s}|^p\,ds$$

for another constant $N=N(p,K_0,K_1,T)$. Hence, taking the supremeum on the left and right-hand side of (II.2.9), and the expectation, as well as using the above estimates yields for all $n\geqslant 1$,

$$\mathbb{E}\sup_{t\leqslant\tau_n\wedge T}|Z_t|^p\leqslant N+N\mathbb{E}|Z_0|^p+\tfrac{1}{2}\mathbb{E}\sup_{t\leqslant\tau_n\wedge T}|Z_t|^p,$$

for another constant $N=N(d,d_1,p,K_0,K_1,T)$. Rearranging the above equations and using Fatou's lemma yields the desired result. $\square$

Chapter III

The filtering equations

In this chapter, based on the article [16], we derive the filtering equations for the signal-observation system (I.0.2), with its coefficients satisfying the measurability properties described in the introduction of Chapter I. More precisely, we aim to derive an equation for the conditional distribution $(P_t(dx))_{t\in[0,T]}$ of X_t given the observations $\{Y_s : s \in [0,t]\}$, such that for real-valued Borel functions f we have

$$\mathbb{E}(f(X_t)|\{Y_s, s \in [0,t]\}) = \int_{\mathbb{R}^d} f(x)\, P_t(dx), \quad \text{almost surely for } t \in [0,T].$$

As we mentioned, there has been an immense interest in the development of filtering theory due to its wide applicability in various disciplines, be they of theoretical or applied nature. A vast amount of research has been done on filtering of partially observed processes governed by stochastic differential equations driven by Wiener processes, i.e., when $\eta = \xi = 0$ in (I.0.2) and when Y contains no jumps, and a quite complete nonlinear filtering theory was built up, see for instance [11] for a historical account.

In this case it is well-known that $(P_t(dx))_{t\in[0,T]}$ satisfies a nonlinear stochastic PDE (SPDE), often called the Kushner-Shiryayev equation in filtering theory. It is also well-known that this equation can be transformed into a linear SPDE, called Zakai equation, or Duncan-Mortensen-Zakai equation for $\mu_t(dx) = \lambda_t P_t(dx)$, the unnormalised conditional distribution, where $(\lambda_t)_{t\in[0,T]}$ is a positive normalising stochastic process.

There exist several known methods of deriving the filtering equations for partially observed diffusion processes, three prominent of which are the "innovation method", the "reference measure method" and a "direct approach". The innovation method is based on "innovation process" representations, (see [43] and [15]), and the direct approach is based on suitable existence and uniqueness theorems for stochastic PDEs (see [40]). The reference probability method is employed in this paper, where we make use of the fact that by Girsanov's theorem one can introduce a new measure under which the observation σ-algebra, $\sigma(Y_s : s \leqslant t)$, is the product σ-algebra of three independent σ-algebras: the σ-algebra generated by the initial observation Y_0 and the σ-algebras generated by a Wiener process with the stochastic differential $B(t, Z_t)dt + dV_t$, and the Poisson random measure $N_1(d\mathfrak{z}, dt)$ on $\mathfrak{Z}_1 \times [0, t]$, respectively. This structure of the observation σ-algebra makes it possible to calculate conditional expectations of functions of the process Z given the observations. (See, e.g., [7] for descriptions of various methods used in filtering theory.)

Recently, also filtering for jump diffusion systems has been intensively studied, which are most often modeled as SDEs driven by Wiener processes as well as random jump measures, a classical case of which are Poisson random measures. In an early article thereon, [47], the filtering equations were derived for uncorrelated continuous observations, as well as an observation process driven only by a jump process that has no common jumps with the signal. A similar system with continuous uncorrelated observations has also been considered in [48]. A more general nonlinear system with jumps in the observation process was considered in [5]. In [2] the filtering equations for a large class of uncorrelated linear systems with jumps are derived. In [19] a very general model is considered and a representation for optional projection of the signal process is derived. However, due to the generality a number of additional assumptions are imposed on their model and equations for the filtering measures are not obtained.
In [9] and [10] the authors deal with a one-dimensional jump diffusion where observation and signal may have common jumps, by introducing a new random measure, nonzero only for observable jumps, relying on a construction in [8]. However, they impose a finiteness condition on the support of the integrand in front of the jump term, which translates to observing only finitely many jumps almost surely. In such a case, the jump measure and the associated predictable compensator, also referred to as dual predictable projection, allow for a specific decomposition, see for instance [25, Sec. XI.4]. The filtering equations have been derived for a class of jump diffusion systems [50], later generalised to include correlated Wiener process noises in [49], however, it seems to us that certain important results needed for this derivation, including Lemma 3.2 in [50], also used in [49], do not hold, for instance if one considers the case of vanishing coefficients. A model where a correlation structure between the Lévy process noises in signal and observation is described using copulas is used in [14] to derive the Zakai equation.

In this chapter we obtain the filtering equations for a jump diffusion system driven by correlated Wiener process noises, as well as correlated Poisson martingale measure noises. We impose common linear growth conditions. We do not assume any non-degeneracy conditions and allow for the number of jumps in any component of $(Z_t)_{t\geqslant 0}$ to be infinite over finite intervals. In order to obtain the equations, we generalise some results from filtering theory and in particular prove a "projection theorem" for a wide class of functions.

In Section III.1 a fairly general condition for Girsanov's transformation and our main result are presented. In Section III.2 a projection theorem covering a wide class of processes is proven, and thereby in the last section the filtering equations are derived.

Conditions and results on the existence and regularity of the filtering density are presented in the subsequent chapters of this thesis.

We conclude by asking the reader to recall the notation presented in Section I.1.

III.1 Formulation of the main results

We consider on a given finite interval $[0,T]$ a $d+d'$-dimensional stochastic process $Z = (Z_t)_{t\in[0,T]} = (X_t, Y_t)_{t\in[0,T]}$ carried by a complete probability space $(\Omega, \mathcal{F}, P)$, equipped with a filtration $(\mathcal{F}_t)_{t\geqslant 0}$ such that $\mathcal{F}_0$ contains the P-null sets of $\mathcal{F}$. We assume that Z satisfies the stochastic differential equation (I.0.2) on the interval $[0,T]$, with an $\mathcal{F}_0$-measurable initial value $Z_0 = (X_0, Y_0)$.

Besides the natural measurability conditions on the coefficients b, σ, ρ, ξ, η and B, described in the Introduction, we assume the following conditions.

Assumption III.1.1. (i) There are nonnegative constants K_0, K_1 and K_2, as well as nonnegative real-valued functions $\bar{\eta} \in L_2(\mathfrak{Z}_0, \mathcal{Z}_0, \nu_0)$ and $\bar{\xi} \in L_2(\mathfrak{Z}_1, \mathcal{Z}_1, \nu_1)$, such that

$$|b(t,z)|^2 \leqslant K_0 + K_1|z|^2, \quad |\sigma(t,z)|^2 + |\rho(t,z)|^2 + |B(t,z)|^2 \leqslant K_0 + K_2|z|^2,$$

$$|\eta(t,z,\mathfrak{z}_0)|^2 \leqslant \bar{\eta}(\mathfrak{z}_0)(K_0 + K_1|z|^2), \quad |\xi(t,z,\mathfrak{z}_1)|^2 \leqslant \bar{\xi}(\mathfrak{z}_1)(K_0 + K_1|z|^2),$$

$$\int_{\mathfrak{Z}_1} |\mathfrak{z}|^2 \, \nu_1(d\mathfrak{z}) \leqslant K_0$$

for all $z = (x,y) \in \mathbb{R}^{d+d'}$, $\mathfrak{z}_i \in \mathfrak{Z}_i$, $i = 0,1$ and $t \in [0,T]$, and we have

(ii)
$$K_1 \mathbb{E}|X_0| + K_2 \mathbb{E}|X_0|^2 < \infty. \tag{III.1.1}$$

Note that in (III.1.1) we use the convention that $0 \times \infty = 0$, i.e., if $K_2 = 0$, then the finiteness of the second moment of $|X_0|$ is not required, and if $K_1 = K_2 = 0$ then Assumption III.1.1 (ii) clearly holds.

The following moment estimate is known and can be easily proved by the help of well-known martingale inequalities.

Remark III.1.1. If Assumption III.1.1(i) holds, then for every $p \in [1,2]$ and $A \in \mathcal{F}_0$ we have

$$\mathbb{E} \sup_{t\leqslant T} \mathbf{1}_A |Z_t|^p \leqslant N(1 + \mathbb{E}\mathbf{1}_A |Z_0|^p) \tag{III.1.2}$$

with a constant N depending only on p, T, K_0, K_1, K_2 and $d+d'$.

We make also the following assumption.

Assumption III.1.2. We have $\mathbb{E}\gamma_T = 1$, where

$$\gamma_t = \exp\left(-\int_0^t B(s, X_s, Y_s)\, dV_s - \tfrac{1}{2}\int_0^t |B(s, X_s, Y_s)|^2\, ds\right), \quad t \in [0, T]. \tag{III.1.3}$$

This assumption implies that the measure Q, defined by $dQ = \gamma_T dP$ on $\mathcal{F}$, is a probability measure equivalent to P, and hence by Girsanov's theorem under Q the process

$$\tilde{V}_t = \int_0^t B(s, X_s, Y_s)\, ds + V_t, \quad t \in [0, T], \tag{III.1.4}$$

is an $\mathcal{F}_t$-Wiener process.

To describe the evolution of the conditional distribution $P_t(dx) = P(X_t \in dx | Y_s, s \leqslant t)$ for $t \in [0, T]$, we introduce the random differential operators

$$\mathcal{L}_t = a_t^{ij}(x) D_{ij} + b_t^i(x) D_i, \quad \mathcal{M}_t^k = \rho_t^{ik}(x) D_i + B_t^k(x), \quad k = 1, 2, ..., d',$$

where

$$a_t^{ij}(x) := \tfrac{1}{2}\sum_{k=1}^{d_1}(\sigma_t^{ik}\sigma_t^{jk})(x) + \tfrac{1}{2}\sum_{l=1}^{d'}(\rho_t^{il}\rho_t^{jl})(x), \quad \sigma_t^{ik}(x) := \sigma^{ik}(t, x, Y_t),$$

$$\rho_t^{il}(x) := \rho^{il}(t, x, Y_t), \quad b_t^i(x) := b^i(t, x, Y_t), \quad B_t^k(x) := B^k(t, x, Y_t)$$

for $\omega \in \Omega$, $t \in [0, T]$, $x = (x^1, ..., x^d) \in \mathbb{R}^d$, and $D_i = \partial/\partial x^i$, $D_{ij} = \partial^2/(\partial x^i \partial x^j)$ for $i, j = 1, 2..., d$. Moreover for every $t \in [0, T]$ and $\mathfrak{z} \in \mathfrak{Z}_1$ we introduce the random operators I_t^ξ and J_t^ξ defined by

$$I_t^\xi \varphi(x, \mathfrak{z}) = \varphi(x + \xi_t(x, \mathfrak{z}), \mathfrak{z}) - \varphi(x, \mathfrak{z}), \quad J_t^\xi \phi(x, \mathfrak{z}) = I_t^\xi \phi(x, \mathfrak{z}) - \sum_{i=1}^d \xi_t^i(x, \mathfrak{z}) D_i \phi(x, \mathfrak{z}) \tag{III.1.5}$$

for functions $\varphi = \varphi(x, \mathfrak{z})$ and $\phi = \phi(x, \mathfrak{z})$ of $x \in \mathbb{R}^d$ and $\mathfrak{z} \in \mathfrak{Z}_1$, and furthermore the random operators I_t^η and J_t^η, defined as I_t^ξ and J_t^ξ, respectively, with $\eta_t(x, \mathfrak{z})$ in place of $\xi_t(x, \mathfrak{z})$, where

$$\xi_t(x, \mathfrak{z}_1) := \xi(t, x, Y_{t-}, \mathfrak{z}_1), \quad \eta_t(x, \mathfrak{z}_0) := \eta(t, x, Y_{t-}, \mathfrak{z}_0)$$

for $\omega \in \Omega$, $t \in [0, T]$, $x \in \mathbb{R}^d$ and $\mathfrak{z}_i \in \mathfrak{Z}_i$ for $i = 0, 1$ ($Y_{0-} := Y_0$). Now we are in the position to formulate our main result. Recall that we denote by $(\mathcal{F}_t^Y)_{t \in [0,T]}$ the completed filtration generated by $(Y_t)_{t \in [0,T]}$.

Theorem III.1.1. *Let Assumptions III.1.1 and III.1.2 hold. Then there exist measure-valued $\mathcal{F}_t^Y$-adapted weakly cadlag processes $(P_t)_{t \in [0,T]}$ and $(\mu_t)_{t \in [0,T]}$ such that almost surely*

$$P_t(\varphi) = \mu_t(\varphi)/\mu_t(\mathbf{1}), \quad \textit{for all } t \in [0, T] \textit{ and}$$

$$P_t(\varphi) = \mathbb{E}(\varphi(X_t)|\mathcal{F}_t^Y), \quad \mu_t(\varphi) = \mathbb{E}_Q(\gamma_t^{-1}\varphi(X_t)|\mathcal{F}_t^Y) \quad (a.s.) \text{ for each } t \in [0,T],$$

for bounded Borel functions φ *on* $\mathbb{R}^d$*, and for every* $\varphi \in C_b^2(\mathbb{R}^d)$ *almost surely*

$$\begin{aligned}\mu_t(\varphi) =& \mu_0(\varphi) + \int_0^t \mu_s(\mathcal{L}_s\varphi)\,ds + \int_0^t \mu_s(\mathcal{M}_s^k\varphi)\,d\tilde{V}_s^k + \int_0^t \int_{\mathfrak{Z}_0} \mu_s(J_s^\eta\varphi)\,\nu_0(d\mathfrak{z})ds \\ &+ \int_0^t \int_{\mathfrak{Z}_1} \mu_s(J_s^\xi\varphi)\,\nu_1(d\mathfrak{z})ds + \int_0^t \int_{\mathfrak{Z}_1} \mu_{s-}(I_s^\xi\varphi)\,\tilde{N}_1(d\mathfrak{z}, ds),\end{aligned} \tag{III.1.6}$$

and

$$\begin{aligned}P_t(\varphi) =& P_0(\varphi) + \int_0^t P_s(\mathcal{L}_s\varphi)\,ds + \int_0^t \big(P_s(\mathcal{M}_s^k\varphi) - P_s(\varphi)P_s(B_s^k)\big)\,d\bar{V}_s^k \\ &+ \int_0^t \int_{\mathfrak{Z}_0} P_s(J_s^\eta\varphi)\,\nu_0(d\mathfrak{z})ds + \int_0^t \int_{\mathfrak{Z}_1} P_s(J_s^\xi\varphi)\,\nu_1(d\mathfrak{z})ds \\ &+ \int_0^t \int_{\mathfrak{Z}_1} P_{s-}(I_s^\xi\varphi)\,\tilde{N}_1(d\mathfrak{z}, ds)\end{aligned} \tag{III.1.7}$$

for all $t \in [0,T]$*, where* $(\tilde{V}_t)_{t\in[0,T]}$ *is given in* (III.1.4)*, and the process* $(\bar{V}_t)_{t\in[0,T]}$ *is defined by*

$$d\bar{V}_t = d\tilde{V}_t - P_t(B_t)\,dt = dV_t + (B_t(X_t) - P_t(B_t))\,dt, \quad \bar{V}_0 = 0.$$

Remark III.1.2. Clearly, $\bar{V} = (\bar{V}_t)_{t\in[0,T]}$ is a continuous process, starting from zero, and by the help of Lemma III.3.2 below it is easy to see that it is $\mathcal{F}_t^Y$-adapted. Moreover, it is not difficult to see that $\bar{V}$ is a martingale (under P) with respect to $(\mathcal{F}_t)_{t\geqslant 0}$, with quadratic variation process $[\bar{V}]_t = t$, $t \in [0,T]$. Hence by Lévy's theorem, $\bar{V}$ is an $\mathcal{F}_t^Y$-Wiener process. It is called the *innovation process* in the case when the observation process does not have a stochastic integral component with respect to Poisson measures, i.e., when $\nu_1 = 0$. In this case it was conjectured that $(\bar{V}_s)_{s\in[0,t]}$ together with Y_0 carry the same information as the observation $(Y_s)_{s\in[0,t]}$, i.e., that the σ-algebra generated by $(\bar{V}_s)_{s\in[0,t]}$ and Y_0 coincides with the σ-algebra generated by $(Y_s)_{s\in[0,t]}$ for every t. An affirmative result concerning this conjecture, under quite general conditions on the filtering models (but without jump components) was proved in [31] and [26]. For our filtering model we conjecture that $(\bar{V}_s)_{s\in[0,t]}$, together with Y_0 and $\{\tilde{N}((0,s]\times\Gamma) : s \in [0,t], \Gamma \in \mathcal{Z}_1\}$ carry the same information as the observation $(Y_s)_{s\in[0,t]}$, if Assumption III.1.1 holds and the coefficients of (I.0.2) satisfy an appropriate Lipschitz condition.

Remark III.1.3. For an $\mathbb{M}$-valued weakly cadlag process $(\nu_t)_{t\in[0,T]}$ (in the sense introduced in the Introduction) there is a set $\Omega' \subset \Omega$ of full probability and there is a uniquely defined (up to indistinguishability) $\mathbb{M}$-valued process $(\nu_{t-})_{t\in(0,T]}$

such that for every $\omega \in \Omega'$

$$\nu_{t-}(\varphi) = \lim_{s \uparrow t} \nu_s(\varphi) \quad \text{for all } \varphi \in C_b(\mathbb{R}^d) \text{ and } t \in (0,T], \tag{III.1.8}$$

and for each $\omega \in \Omega'$ we have $\nu_{t-} = \nu_t$, for all but at most countably many $t \in (0,T]$.

Proof. First we show that $(\nu_{t-})_{t\in[0,T]}$ defined by the right-hand side of (III.1.8) defines a measure-valued process. Since ν is weakly cadlag, there exists a set Ω' of full probability such that for all $\omega \in \Omega'$ and $\varphi \in C_0 = C_0(\mathbb{R}^d)$, the space of continuous and compactly supported functions, the function $(\nu_t(\varphi))_{t\in[0,T]}$ is cadlag. Hence we can see that, for $\omega \in \Omega'$,

$$F(\varphi)(\omega) := \lim_{r \uparrow s} \nu_r(\varphi)(\omega)$$

defines a positive linear functional on C_0. By the Riesz-Markov-Kakutani theorem, see for instance [24, Thm. D], we then know that, for each $\omega \in \Omega'$, there exists a measure, denoted by $\nu_{s-}(\omega)$, such that

$$F(\varphi)(\omega) = \nu_{s-}(\varphi)(\omega), \quad \text{for all } \omega \in \Omega' \text{ and } \varphi \in C_0(\mathbb{R}^d).$$

Finally, since for each $\omega \in \Omega'$ and $\varphi \in \Phi$, for a countable measure determining subset $\Phi \subset C_b$, we have that $\nu_{t-}(\varphi) = \nu_t(\varphi)$ for all but countably many $t \in [0,T]$, we conclude that also $\nu_{t-} = \nu_t$ for all but countably many $t \in [0,T]$. □

We will prove Theorem III.1.1 by deducing equation (III.1.7) from equation (III.1.6), which we obtain by taking, under Q, the conditional expectation of the terms in the equation for $\gamma_t^{-1}\varphi(X_t)$, given the observation $\{Y_s : s \leqslant t\}$.

There are several known conditions ensuring that Assumption III.1.2 is satisfied. For a simple proof for the well-known Novikov condition and Kazamaki condition, and their generalisations we refer to Exercise 6.8.VI in [33], [34] and [38]. These conditions, are clearly satisfied if $|B|$ is bounded, but it does not seem to be easy to reformulate them in terms of the coefficients of the system of equations (I.0.2), if $|B|$ is unbounded. Here we give a condition, which together with Assumption III.1.1(i) ensures that Assumption III.1.2 holds.

Assumption III.1.3. There is a constant K such that

$$-x^i \rho^{ik}(t,z) B^k(t,z) \leqslant K(1+|z|^2) \quad \text{for all } t \in [0,T],\ z = (x,y) \in \mathbb{R}^{d+d'}.$$

Remark III.1.4. Define the $\mathbb{R}^{(d+d')\times d'}$-valued function $\hat{\rho}$ by $\hat{\rho}^{jk} := \rho^{jk}$ for $j = 1,2,...,d$, $k = 1,2,...,d'$ and $\hat{\rho}^{jk} := 0$ for $j = d+1,...,d+d'$, $k = 1,2,...,d'$. Then Assumption III.1.3 means that the "one-sided linear growth" condition

$$zf(t,z) \leqslant K(1+|z|^2), \quad t \in [0,T],\ z \in \mathbb{R}^{d+d'},$$

holds for the $\mathbb{R}^{d+d'}$-valued function $f = -\hat{\rho}B$, where zf denotes the standard inner product of the vectors $z, f \in \mathbb{R}^{d+d'}$. Clearly, this condition is essentially weaker then the linear growth condition on f (in $z \in \mathbb{R}^{d+d'}$), which obviously holds if one of the functions ρ and B is bounded in magnitude and the other satisfies the linear growth condition in Assumption III.1.1 (i).

The following theorem provides a condition under which Assumption III.1.2 holds. A more comprehensive investigation on absolutely continuous changes of measures associated to jump diffusion processes, including a generalisation of the following theorem, can be found in [29].

Theorem III.1.2. *Let Assumptions III.1.1(i) and III.1.3 hold. Then $\mathbb{E}\gamma_T = 1$, i.e., Assumption III.1.2 holds.*

Proof. We want to prove $\mathbb{E}(\gamma_T \mathbf{1}_{|Z_0| \leqslant R}) = P(|Z_0| \leqslant R)$ for every constant $R > 0$, since by monotone convergence it implies

$$\mathbb{E}\gamma_T = \lim_{R\to\infty} \mathbb{E}(\gamma_T \mathbf{1}_{|Z_0|\leqslant R}) = \lim_{R\to\infty} P(|Z_0| \leqslant R) = 1.$$

To this end we fix a constant $R > 0$ and set $\bar{\gamma}_t := \gamma_t \mathbf{1}_{|Z_0|\leqslant R}$. By Itô's formula

$$d\bar{\gamma}_t = -\bar{\gamma}_t B(t, Z_t)\, dV_t,$$

that shows that $\bar{\gamma}$ is a local $\mathcal{F}_t$-martingale. Thus $\mathbb{E}\bar{\gamma}_{T\wedge\tau_n} = P(|Z_0| \leqslant R)$ for an increasing sequence $(\tau_n)_{n=1}^{\infty}$ of stopping times τ_n such that τ_n converges to ∞ as $n \to \infty$, and $(\gamma_{t\wedge\tau_n})_{t\in[0,T]}$ is a martingale for every n. Consequently, if we can show $\mathbb{E}\sup_{t\leqslant T} \bar{\gamma}_t < \infty$, then we can use Lebesgue's theorem on dominated convergence to get $\mathbb{E}\bar{\gamma}_T = P(|Z_0| \leqslant R)$. Define the stopping times

$$\tau_n = \inf\{t \in [0, T] : [\bar{\gamma}]_t \geqslant n\} \quad \text{for integers } n \geqslant 1,$$

where

$$[\bar{\gamma}]_t = \int_0^t \bar{\gamma}_s^2 |B(s, Z_t)|^2\, ds.$$

Then by standard estimates, using the Davis inequality, we have

$$\begin{aligned}\mathbb{E}\sup_{t\leqslant T} \bar{\gamma}_{t\wedge\tau_n} \leqslant 1 + 3\mathbb{E}[\bar{\gamma}]_{T\wedge\tau_n}^{1/2} &\leqslant 1 + 3\mathbb{E}\sup_{t\leqslant T} \bar{\gamma}_{t\wedge\tau_n}^{1/2} \Big(\int_0^{T\wedge\tau_n} \bar{\gamma}_t |B(t, Z_t)|^2\, dt \Big)^{1/2} \\ &\leqslant 1 + \tfrac{1}{2}\mathbb{E}\sup_{t\leqslant T} \bar{\gamma}_{t\wedge\tau_n} + 5\mathbb{E}\int_0^T \bar{\gamma}_t |B(t, Z_t)|^2\, dt,\end{aligned}$$

which, after we subtract $\frac{1}{2}E\sup_{t\leqslant T} \bar{\gamma}_{t\wedge\tau_n}$ and let $n \to \infty$, by Fatou's lemma gives

$$\tfrac{1}{2}\mathbb{E}\sup_{t\leqslant T} \bar{\gamma}_t \leqslant 1 + 5\mathbb{E}\int_0^T \bar{\gamma}_t |B(t, Z_t)|^2\, dt \leqslant 1 + 5\mathbb{E}\int_0^T \bar{\gamma}_t (K_0 + K_2|Z_t|^2)\, dt.$$

Since $\mathbb{E}\bar{\gamma}_t \leqslant 1$, to show that the right-hand side of the last inequality is finite we need only prove that if $K_2 \neq 0$ then

$$\sup_{t \leqslant T} \mathbb{E}\bar{\gamma}_t |Z_t|^2 < \infty. \tag{III.1.9}$$

To this end we apply Itô's formula to $U_t := \bar{\gamma}_t |Z_t|^2$ and use Assumptions III.1.1 (ii) and III.1.3 to get

$$\begin{aligned} dU_t =& \bar{\gamma}_t (2X_t b(t, Z_t) + 2Y_t B(t, Z_t) + |\sigma(t, Z_t)|^2 + |\rho(t, Z_t)|^2 + 1)\, dt \\ & - 2\bar{\gamma}_t (X_t \rho(t, Z_t) B(t, Z_t) + Y_t B_t(t, Z_t))\, dt + \bar{\gamma}_t \int_{\mathfrak{Z}_0} |\eta(t, Z_t, \mathfrak{z})|^2 \, \nu_0(d\mathfrak{z}) dt \\ & + \bar{\gamma}_t \int_{\mathfrak{Z}_1} |\xi(t, Z_t, \mathfrak{z})|^2 \, \nu_1(d\mathfrak{z}) dt + \bar{\gamma}_t \int_{\mathfrak{Z}_1} |\mathfrak{z}|^2 \, \nu_1(d\mathfrak{z}) dt + dm_t \\ \leqslant& N \bar{\gamma}_t \, dt + N U_t \, dt + dm_t \end{aligned} \tag{III.1.10}$$

with a constant N and a cadlag local martingale m starting from zero. Hence by a standard stopping time argument and Gronwall's inequality we get a constant N such that

$$\sup_{t \leqslant T} \mathbb{E} U_{t \wedge \tau_n} \leqslant N(1 + \mathbb{E}(\mathbf{1}_{|Z_0| \leqslant R} |Z_0|^2)) < \infty$$

for an increasing sequence of stopping times $\tau_n \uparrow \infty$. Letting here $n \to \infty$ by Fatou's lemma we get (III.1.9), which finishes the proof of the theorem. $\square$

III.2 Preliminaries

The following lemma is our main tool for calculating conditional expectations of Lebesgue and Itô stochastic integrals of simple processes under Q given $\mathcal{F}_t^Y$.

Lemma III.2.1. *Let X and Y be random variables such that $\mathbb{E}|X| < \infty$, $\mathbb{E}|Y| < \infty$ and $\mathbb{E}|XY| < \infty$. Let $\mathcal{G}^1$, $\mathcal{G}^2$ and $\mathcal{G}$ be σ-algebras of events such that $\mathcal{G}^1 \subset \mathcal{G}$, $\mathcal{G}^2$ is independent of $\mathcal{G}$, X is $\mathcal{G}$-measurable and Y is independent of $\mathcal{G} \vee \mathcal{G}^2$. Then almost surely*

$$\mathbb{E}(XY | \mathcal{G}^1 \vee \mathcal{G}^2) = \mathbb{E}(X | \mathcal{G}^1) \mathbb{E}Y.$$

Proof. The right-hand side of the above equation is a $\mathcal{G}^1$-measurable random variable, hence it is obviously $\mathcal{G}^1 \vee \mathcal{G}^2$-measurable. Let $\mathcal{H}$ denote the family of $G \in \mathcal{G}^1 \vee \mathcal{G}^2$ such that

$$\mathbb{E}Y \mathbb{E}(\mathbb{E}(X | \mathcal{G}^1) \mathbf{1}_G) = \mathbb{E}(XY \mathbf{1}_G).$$

Then $\mathcal{H}$ is a λ-system, and for $G = G_1 \cap G_2$, $G_i \in \mathcal{G}^i$ we have

$$\mathbb{E}Y \mathbb{E}(\mathbb{E}(X | \mathcal{G}^1) \mathbf{1}_G) = \mathbb{E}Y \mathbb{E}(\mathbb{E}(\mathbf{1}_{G_1} X | \mathcal{G}^1) \mathbf{1}_{G_2})) = \mathbb{E}Y \mathbb{E}(\mathbb{E}(\mathbf{1}_{G_1} X | \mathcal{G}^1)) \mathbb{E}\mathbf{1}_{G_2}$$

$$= \mathbb{E}Y \mathbb{E}(\mathbf{1}_{G_1} X) \mathbb{E}\mathbf{1}_{G_2} = \mathbb{E}(XY \mathbf{1}_G),$$

that shows that $\mathcal{H}$ contains the π-system $\{G_1 \cap G_2 : G_i \in \mathcal{G}^i, i = 1, 2\}$. Hence, by Dynkin's monotone class lemma $\mathcal{H} = \mathcal{G}^1 \vee \mathcal{G}^2$, which completes the proof. $\square$

To formulate a theorem on conditional expectations of Lebesgue and Itô integrals we consider a complete filtered probability space $(\Omega, \mathcal{F}, P, \mathcal{F}_t)$ carrying independent $\mathcal{F}_t$-Wiener processes $W^i = (W^i_t)_{t \geqslant 0}$ and independent $\mathcal{F}_t$-Poisson random measures $N_i = N_i(d\mathfrak{z}, dt)$ with σ-finite characteristic measures ν_i on separable measurable spaces $(\mathfrak{Z}_i, \mathcal{Z}_i)$ for $i = 0, 1$, respectively. We denote by $\mathcal{G}_t$ the P-completion of the σ-algebra generated by the events of a σ-algebra $\mathcal{Y}_0 \subset \mathcal{F}_0$ together with the random variables W^1_s and $N_1((0, s] \times \Gamma)$ for $s \leqslant t$ and $\Gamma \in \mathcal{Z}_1$ such that $\nu_1(\Gamma) < \infty$. The predictable σ-algebras on $\Omega \times [0, T]$, relative to $(\mathcal{F}_t)_{t \geqslant 0}$ and $(\mathcal{G}_t)_{t \geqslant 0}$ are denoted by $\mathcal{P}_{\mathcal{F}}$ and $\mathcal{P}_{\mathcal{G}}$, respectively. The optional σ-algebras relative to $(\mathcal{F}_t)_{t \geqslant 0}$ and $(\mathcal{G}_t)_{t \geqslant 0}$ are denoted by $\mathcal{O}_{\mathcal{F}}$ and $\mathcal{O}_{\mathcal{G}}$, respectively.

The following definition will be frequently used.

Definition III.2.1. Given a probability space $(\Omega, \mathcal{F}, P)$ and a sub-σ-algebra $\mathcal{G} \subset \mathcal{F}$, we say that a random variable f is σ-integrable (with respect to P) relative to $\mathcal{G}$, if there exists an increasing sequence $(\Omega_n)_{n=1}^{\infty}$ such that $\bigcup_n \Omega_n = \Omega$, $\Omega_n \in \mathcal{G}$ and $\mathbb{E}|f \mathbf{1}_{\Omega_n}| < \infty$ for all n.

One knows that for nonnegative random variables f and σ-algebras $\mathcal{G} \subset \mathcal{F}$ the conditional expectation $\mathbb{E}(f|\mathcal{G})$ is well-defined, and that for general random variables f, such that $\mathbb{E}(|f||\mathcal{G}) < \infty$ almost surely, the extended conditional expectation is defined as $\mathbb{E}(f^+|\mathcal{G}) - \mathbb{E}(f^-|\mathcal{G})$ on the set $\mathbb{E}(|f||\mathcal{G}) < \infty$, and it is defined to be $+\infty$ on $\mathbb{E}(|f||\mathcal{G}) = \infty$. It is not difficult to show that we have $\mathbb{E}(|f||\mathcal{G}) < \infty$ almost surely if and only if f is σ-integrable relative to $\mathcal{G}$ (see [25, Thm. 1.17] for a proof), meaning that $\mathbb{E}(f^+|\mathcal{G}) < \infty$ and $\mathbb{E}(f^-|\mathcal{G}) < \infty$ almost surely.

We consider real-valued $\mathcal{F} \otimes \mathcal{B}([0, T])$-measurable $\mathcal{F}_t$-adapted processes $f = (f_t)_{t \in [0,T]}$ and $g = (g_t)_{t \in [0,T]}$ on $\Omega \times [0, T]$, real-valued $\mathcal{F} \otimes \mathcal{B}([0, T]) \otimes \mathcal{Z}_i$-measurable functions $h^{(i)} = h^{(i)}_t(\omega, \mathfrak{z})$ of $(\omega, t, \mathfrak{z}) \in \Omega \times [0, T] \times \mathfrak{Z}_i$ for $i = 0, 1$, and a real-valued $\mathcal{F} \otimes \mathcal{B}([0, T]) \otimes \mathcal{Z}$-measurable function $h = h_t(\omega, \mathfrak{z})$ of $(\omega, t, \mathfrak{z}) \in \Omega \times [0, T] \times \mathfrak{Z}$, such that for every $t \in [0, T]$ the functions $h^{(i)}_t$ and h_t are $\mathcal{F}_t \otimes \mathcal{Z}_i$-measurable and $\mathcal{F}_t \otimes \mathcal{Z}$-measurable, respectively, for $i = 0, 1$, where $(\mathfrak{Z}, \mathcal{Z})$ is a separable measurable space, equipped with a σ-finite measure ν. Assume that almost surely

$$F := \left(\int_0^T |f_s|^2 \, ds \right)^{1/2} < \infty \quad H^{(i)} := \left(\int_0^T \int_{\mathfrak{Z}_i} |h^{(i)}_s(\mathfrak{z})|^2 \, \nu_i(d\mathfrak{z}) ds \right)^{1/2} < \infty \tag{III.2.1}$$

$$G := \int_0^T |g_s| \, ds < \infty, \quad H := \int_0^T \int_{\mathfrak{Z}} |h_s(\mathfrak{z})| \, \nu(d\mathfrak{z}) ds < \infty \tag{III.2.2}$$

for $i = 0, 1$. Then the processes

$$\alpha_t := \int_0^t g_s \, ds, \quad \delta_t := \int_0^t \int_{\mathfrak{Z}} h_s(\mathfrak{z}) \, \nu(d\mathfrak{z}) ds, \quad t \in [0, T],$$

and

$$\beta_t^{(i)} = \int_0^t f_s \, dW_s^i, \quad \delta_t^{(i)} = \int_0^t \int_{\mathfrak{Z}_i} h_s^{(i)}(\mathfrak{z}) \, \tilde{N}_i(d\mathfrak{z}, ds), \quad t \in [0, T], \tag{III.2.3}$$

are well-defined for $i = 0, 1$, and we have the following theorem.

Theorem III.2.2. *Assume the random variables F^r, G, H and $|H^{(i)}|^2$ for $i = 0, 1$, for some $r > 1$ are σ-integrable (with respect to P) relative to $\mathcal{G}_0$ and that for every $\mathcal{G}_t$-stopping time $\tau \leqslant T$, $\mathfrak{z} \in \mathfrak{Z}$ and $\mathfrak{z}_i \in \mathfrak{Z}_i$ $(i = 0, 1)$ the random variables f_τ, g_τ, $h_\tau(\mathfrak{z})$, $h_\tau^{(i)}(\mathfrak{z}_i)$ $(i = 0, 1)$ are σ-integrable relative to $\mathcal{G}_0$. Then for $t \in [0, T]$ we have*

$$\mathbb{E}(\beta_t^{(1)} | \mathcal{G}_t) = \int_0^t \hat{f}_s \, dW_s^1, \quad \mathbb{E}(\beta_t^{(0)} | \mathcal{G}_t) = 0, \tag{III.2.4}$$

$$\mathbb{E}(\alpha_t | \mathcal{G}_t) = \int_0^t \hat{g}_s \, ds, \quad \mathbb{E}(\delta_t | \mathcal{G}_t) = \int_0^t \int_{\mathfrak{Z}} \hat{h}_s(\mathfrak{z}) \, \nu(d\mathfrak{z}) ds, \tag{III.2.5}$$

$$\mathbb{E}(\delta_t^{(1)} | \mathcal{G}_t) = \int_0^t \int_{\mathfrak{Z}_1} \hat{h}_s^{(1)}(\mathfrak{z}) \, \tilde{N}_1(d\mathfrak{z}, ds), \quad \mathbb{E}(\delta_t^{(0)} | \mathcal{G}_t) = 0 \tag{III.2.6}$$

almost surely for some $\mathcal{P}_{\mathcal{G}}$-measurable functions $\hat{f}$ and $\hat{g}$ on $\Omega \times [0, T]$, a $\mathcal{P}_{\mathcal{G}} \otimes \mathcal{Z}_1$-measurable function $\hat{h}^1$ on $\Omega \times [0, T] \times \mathfrak{Z}_1$, and a $\mathcal{P}_{\mathcal{G}} \otimes \mathcal{Z}$-measurable function $\hat{h}$ on $\Omega \times [0, T] \times \mathfrak{Z}$ such that

$$\hat{f}_t = \mathbb{E}(f_t | \mathcal{G}_t), \quad \hat{g}_t = \mathbb{E}(g_t | \mathcal{G}_t) \quad \text{(a.s.) for } dt\text{-a.e. } t \in [0, T], \tag{III.2.7}$$

$$\hat{h}_t^{(1)} = \mathbb{E}(h_t^{(1)}(\mathfrak{z}) | \mathcal{G}_t) \quad \text{(a.s.) for } dt \otimes \nu_1\text{-a.e. } (t, \mathfrak{z}) \in [0, T] \times \mathfrak{Z}_1, \tag{III.2.8}$$

$$\hat{h}_t = \mathbb{E}(h_t(\mathfrak{z}) | \mathcal{G}_t) \quad \text{(a.s.) for } dt \otimes \nu\text{-a.e. } (t, \mathfrak{z}) \in [0, T] \times \mathfrak{Z}. \tag{III.2.9}$$

Proof. Since F^r is σ-integrable with respect to $\mathcal{G}_0$, there is an increasing sequence $\Omega_n \in \mathcal{G}_0$ such that $\bigcup_{n=1}^\infty \Omega_n = \Omega$ and $\mathbb{E}(\mathbf{1}_{\Omega_n} F^r) < \infty$ for every integer $n \geqslant 1$. By the definition and elementary properties of (extended) conditional expectations and stochastic integrals, we have

$$\mathbf{1}_{\Omega_n} \mathbb{E}\Big(\int_0^t f_s \, dW_s^i \Big| \mathcal{G}_t \Big) = \mathbb{E}\Big(\mathbf{1}_{\Omega_n} \int_0^t f_s \, dW_s^i \Big| \mathcal{G}_t \Big) = \mathbb{E}\Big(\int_0^t \mathbf{1}_{\Omega_n} f_s \, dW_s^i \Big| \mathcal{G}_t \Big),$$

$$\mathbf{1}_{\Omega_n} \mathbb{E}(f_t | \mathcal{G}_t) = \mathbb{E}(\mathbf{1}_{\Omega_n} f_t | \mathcal{G}_t), \quad t \in [0, T]$$

for $i = 0, 1$ and every $n \geqslant 1$. Thus, taking $\mathbf{1}_{\Omega_n} f$ in place of f, we may assume that $\mathbb{E} F^r < \infty$. Similarly, we may also assume that $\mathbb{E} G$, $\mathbb{E} H$ and $\mathbb{E}|H^{(i)}|^2$ are finite in what follows below. Assume first that f belongs to $\mathcal{H}_0$, the set of simple processes of the form

$$f_t = \sum_{i=0}^{k-1} \xi_i \mathbb{1}_{(t_i,t_{i+1}]}(t), \tag{III.2.10}$$

where $0 = t_0 \leqslant \cdots \leqslant t_k = T$ are deterministic time instants, and ξ_i is a bounded $\mathcal{F}_{t_i}$-measurable random variable for every $i = 0, 1, ..., k-1$ for an integer $k \geqslant 1$. Then we have

$$\mathbb{E}\Big(\int_0^t f_s \, dW_s^1 \Big| \mathcal{G}_t \Big) = \sum_i \mathbb{E}\big(\xi_i (W^1_{t_{i+1}\wedge t} - W^1_{t_i \wedge t}) \big| \mathcal{G}_t \big), \quad \text{for } t \in [0,T]. \tag{III.2.11}$$

For $0 \leqslant r \leqslant s \leqslant T$ define the σ-algebra

$$\mathcal{G}_{r,s} = \sigma(W^1_v - W^1_u, N_1(\Gamma \times (u,v]) : r \leqslant u \leqslant v \leqslant s, \Gamma \in \mathcal{Z}_1, \nu_1(\Gamma) < \infty).$$

Then σ-algebras $\mathcal{G}_r$ and $\mathcal{G}_{r,s}$ are independent and $\mathcal{G}_s = \mathcal{G}_r \vee \mathcal{G}_{r,s}$. Thus, using Lemma III.2.1 with $X := \xi_i$, $Y := 1$, $\mathcal{G}^1 := \mathcal{G}_{t_i}$, $\mathcal{G} := \mathcal{F}_{t_i}$ and $\mathcal{G}^2 := \mathcal{G}_{t_i,s}$ for $t_i \leqslant s \leqslant T$, we have

$$\mathbb{E}(\xi_i | \mathcal{G}_s) = \mathbb{E}(\xi_i | \mathcal{G}_{t_i}) \quad \text{for } i = 0, 1, 2, ..., k-1. \tag{III.2.12}$$

Hence for $t_i \leqslant s \leqslant t_{i+1} \leqslant t \leqslant T$,

$$\mathbb{E}(\xi_i (W^1_{t_{i+1}} - W^1_{t_i}) | \mathcal{G}_t) = \mathbb{E}(\xi_i | \mathcal{G}_t)(W^1_{t_{i+1}} - W^1_{t_i}) = \mathbb{E}(\xi_i | \mathcal{G}_s)(W^1_{t_{i+1}} - W^1_{t_i}) \tag{III.2.13}$$

and for $t_j \leqslant s \leqslant t \leqslant T$,

$$\mathbb{E}(\xi_j (W^1_t - W^1_{t_j}) | \mathcal{G}_t) = \mathbb{E}(\xi_j | \mathcal{G}_t)(W^1_t - W^1_{t_j}) = \mathbb{E}(\xi_j | \mathcal{G}_s)(W^1_t - W^1_{t_j}). \tag{III.2.14}$$

Consequently, defining $\hat{f}_s = \mathbb{E}(\xi_i | \mathcal{G}_s) = \mathbb{E}(f_s | \mathcal{G}_s)$ for $s \in (t_i, t_{i+1}]$, $i = 0, 1, ..., k-1$, the function $\hat{f}$ on $\Omega \times [0,T]$ is $\mathcal{P}_{\mathcal{G}}$-measurable, and using (III.2.11) we can see that the first equation in (III.2.4) holds. Assume now that f is $\mathcal{F} \otimes \mathcal{B}([0,T])$-measurable and $\mathcal{F}_t$-adapted such that $\mathbb{E}F^r < \infty$. Then there are sequences $(f^n)_{n=1}^\infty$ and $(\hat{f}^n)_{n=1}^\infty$ such that $f^n \in \mathcal{H}_0$, $\hat{f}^n$ is $\mathcal{P}_{\mathcal{G}}$-measurable,

$$\lim_{n\to\infty} \mathbb{E} \left(\int_0^T |f_t - f_t^n|^2 \, dt \right)^{r/2} = 0, \tag{III.2.15}$$

and almost surely

$$\mathbb{E}(I_t(f^n) | \mathcal{G}_t) := \mathbb{E}\Big(\int_0^t f_s^n \, dW_s^1 \Big| \mathcal{G}_t \Big) = \int_0^t \hat{f}_s^n \, dW_s^1 =: I_t(\hat{f}^n) \quad \text{for all } t \in [0,T], \tag{III.2.16}$$

$$\hat{f}_t^n = \mathbb{E}(f_t^n | \mathcal{G}_t) \quad \text{for } dt\text{-a.e. } t \in [0,T] \tag{III.2.17}$$

for all $n \geqslant 1$. Using the Davis inequality, Doob's inequality, Jensen's and

Burkholder's inequalities for any $r > 1$ we have

$$\mathbb{E}\left(\int_0^T |\hat{f}_t^n - \hat{f}_t^m|^2\,dt\right)^{1/2} \leqslant 3\mathbb{E}\sup_{t\leqslant T}|I_t(\hat{f}^n - \hat{f}^m)|$$

$$= 3\mathbb{E}\sup_{t\in[0,T]\cap\mathbb{Q}}|\mathbb{E}\big(I_t(f^n - f^m)\big|\mathcal{G}_t\big)| \leqslant 3\mathbb{E}\sup_{t\in[0,T]\cap\mathbb{Q}}(\mathbb{E}(\sup_{s\leqslant T}|I_s(f^n - f^m)||\mathcal{G}_t)$$

$$\leqslant 3\frac{r}{r-1}\left(\mathbb{E}\sup_{t\leqslant T}|I_t(f^n - f^m)|^r\right)^{1/r} \leqslant N\left(\mathbb{E}\left(\int_0^T |f_t^n - f_t^m|^2\,dt\right)^{r/2}\right)^{1/r},$$

where $\mathbb{Q}$ is the set of rational numbers and $N = N(r)$ is a constant, which gives

$$\lim_{n,m\to\infty}\mathbb{E}\left(\int_0^T |\hat{f}_t^n - \hat{f}_t^m|^2\,dt\right)^{1/2} = 0.$$

Thus there exists a $\mathcal{P}_{\mathcal{G}}$-measurable function $\hat{f}$ on $\Omega \times [0,T]$, such that

$$\lim_{n\to\infty}\mathbb{E}\left(\int_0^T |\hat{f}_t - \hat{f}_t^n|^2\,dt\right)^{1/2} = 0, \tag{III.2.18}$$

which implies

$$\lim_{n\to\infty}\mathbb{E}\sup_{t\in[0,T]}|I_t(\hat{f}) - I_t(\hat{f}^n)| = 0. \tag{III.2.19}$$

Using Jensen's and Davis' inequalities again we have

$$\mathbb{E}|\mathbb{E}(I_t(f)|\mathcal{G}_t) - \mathbb{E}(I_t(f^n)|\mathcal{G}_t)| \leqslant \mathbb{E}\mathbb{E}(|I_t(f - f^n)||\mathcal{G}_t)$$

$$= \mathbb{E}|I_t(f - f^n)| \leqslant 3\mathbb{E}\left(\int_0^T |f_t - f_t^n|^2\,dt\right)^{1/2} \quad \text{for every } t \in [0,T],$$

i.e., for $n \to \infty$

$$\mathbb{E}(I_t(f^n)|\mathcal{G}_t) \to \mathbb{E}(I_t(f)|\mathcal{G}_t) \quad \text{in } L_1(\Omega) \text{ for every } t \in [0,T]. \tag{III.2.20}$$

Thus letting $n \to \infty$ in equation (III.2.16), by virtue of (III.2.19) and (III.2.20) we get the first equation in (III.2.4). Clearly, (III.2.15) and (III.2.18) imply

$$\lim_{n\to\infty}\int_0^T \mathbb{E}|f_t - f_t^n| + \mathbb{E}|\hat{f}_t - \hat{f}_t^n|\,dt = 0.$$

Hence there is a subsequence $n_l \to \infty$ and a set $S \in \mathcal{B}([0,T])$ of Lebesgue measure 0 such that for $n_l \to \infty$,

$$f_t^{n_l} \to f_t \text{ and a } \hat{f}_t^{n_l} \to \hat{f}_t \text{ in } L_1(\Omega) \text{ for each } t \in [0,T]\backslash S =: S^c,$$

and taking into account (III.2.17), we can assume that S is a dt-zero set such that we also have $\hat{f}_t^{n_l} = \mathbb{E}(f_t^{n_l}|\mathcal{G}_t)$ (a.s.) for every $t \in S^c$. Thus for $n_l \to \infty$ we have $\mathbb{E}(f_t^{n_l}|\mathcal{G}_t) \to \mathbb{E}(f_t|\mathcal{G}_t)$ in $L_1(\Omega)$ for each $t \in S^c$, which gives

$$\hat{f}_t = \mathbb{E}(f_t|\mathcal{G}_t) \quad \text{almost surely for every } t \in S^c,$$

i.e., the first equation in (III.2.7) holds. To prove the second equation in (III.2.4) we note that for ξ_i from the expression (III.2.10) we have

$$\mathbb{E}(\xi_i(W^0_{t_{i+1}\wedge t} - W^0_{t_i})|\mathcal{G}_t) = \mathbb{E}(\xi_i|\mathcal{G}_{t_i})\mathbb{E}(W^0_{t_{i+1}\wedge t} - W^0_{t_i}) = 0 \quad \text{for } i = 1, 2, ..., N-1, \tag{III.2.21}$$

by using Lemma III.2.1 with $X = \xi_i$, $Y = W^0_{t_{i+1}\wedge t} - W^0_{t_i\wedge t}$ $\mathcal{G}^1 := \mathcal{G}_{t_i} \subset \mathcal{F}_{t_i} =: \mathcal{G}$ and $\mathcal{G}^2 := \mathcal{G}_{t_i,t}$ for $t_i \leqslant t$. Hence we get the second equation in (III.2.4) for f given in (III.2.10), and the general case follows by approximation as above. To prove the first equation in (III.2.5) assume that g is given by the right-hand side of (III.2.10). Then using (III.2.12) we can see that

$$\hat{g}_t := \sum_{i=0}^{k-1} \mathbb{E}(\xi_i|\mathcal{G}_{t_i})\mathbf{1}_{(t_i,t_{i+1}]}(t) = \mathbb{E}(g_t|\mathcal{G}_t), \quad t \in [0, T],$$

and that the first equation in (III.2.5) and the second equation in (III.2.7) hold. Assume now that g is an $\mathcal{F} \otimes \mathcal{B}([0, T])$-measurable F_t-adapted random process such that $\mathbb{E}G < \infty$. Then there are sequences $(g^n)_{n=1}^\infty$ and $(\hat{g}^n)_{n=1}^\infty$ such that $g^n \in \mathcal{H}_0$, $\hat{g}^n$ is $\mathcal{P}_{\mathcal{G}}$-measurable,

$$\lim_{n\to\infty} \mathbb{E} \int_0^T |g_t - g_t^n|\, dt = 0,$$

and almost surely

$$\mathbb{E}\Big(\int_0^t g_s^n\, ds \Big| \mathcal{G}_t\Big) = \int_0^t \hat{g}_s^n\, ds \quad \text{for all } t \in [0, T],$$

$$\hat{g}_t^n = \mathbb{E}(g_t^n|\mathcal{G}_t) \quad \text{for } dt\text{-a.e. } t \in [0, T].$$

Hence noting that by Tonelli's theorem and Jensen's inequality

$$\mathbb{E} \int_0^T |\hat{g}_t^n - \hat{g}_t^m|\, dt = \int_0^T \mathbb{E}|\mathbb{E}(g_t^n|\mathcal{G}_t) - \mathbb{E}(g_t^m|\mathcal{G}_t)|\, dt$$

$$\leqslant \int_0^T \mathbb{E}\mathbb{E}(|g_t^n - g_t^m||\mathcal{G}_t)\, dt = \mathbb{E} \int_0^T |g_t^n - g_t^m|\, dt,$$

and repeating previous arguments we get a $\mathcal{P}_{\mathcal{F}}$-measurable $\hat{g}$ such that the first equation in (III.2.5) and the second equation in (III.2.7) hold. To prove the second equation in (III.2.5) we assume first that

$$h_t(\mathfrak{z}) = \sum_{i=0}^{k-1} \xi_i \mathbf{1}_{(t_i,t_{i+1}]\times\Gamma_i}(t,\mathfrak{z}), \tag{III.2.22}$$

for a partition $0 \leqslant t_0 \leqslant t_1 \leqslant ... \leqslant t_k = T$ of $[0,T]$, bounded $\mathcal{F}_{t_i}$-measurable random variables ξ_i and sets $\Gamma_i \in \mathcal{Z}$, $\nu(\Gamma_i) < \infty$ for $i = 0,...,k-1$. Then

$$\mathbb{E}\Big(\int_0^t \int_{\mathfrak{Z}} h_s(\mathfrak{z})\,\nu(d\mathfrak{z})\,ds\Big|\mathcal{G}_t\Big) = \sum_{i=0}^{k-1} \mathbb{E}(\xi_i|\mathcal{G}_t)\nu(\Gamma_i)(t_{i+1}\wedge t - t_i \wedge t), \quad t \in [0,T].$$

Thus, since by virtue of (III.2.12) we have

$$\hat{h}_t(\mathfrak{z}) := \sum_{i=0}^{k-1} \mathbb{E}(\xi_i|\mathcal{G}_{t_i})\mathbf{1}_{(t_i,t_{i+1}]}(t)\mathbf{1}_{\Gamma_i}(\mathfrak{z}) = \mathbb{E}(h_t(\mathfrak{z})|\mathcal{G}_t), \quad t \in [0,T],\ \mathfrak{z} \in \mathfrak{Z},$$

for $\hat{h}$ the second equation in (III.2.5) and by definition (III.2.9) hold. Hence we can get these equations in the general case by a straightforward approximation procedure in the same way as the first equation in (III.2.5) and the second equation in (III.2.7) have been proved above.

Now we are going to prove (III.2.6). Assume first that $h^{(1)}$ is a simple function, given by the right-hand side of equation (III.2.22) with $\Gamma_i \in \mathcal{Z}_1$, $\nu_1(\Gamma_i) < \infty$, $i = 0,1,...,k-1$. Then

$$\mathbb{E}\Big(\int_0^t \int_{\mathfrak{Z}_1} h_s^{(1)}(\mathfrak{z})\,\tilde{N}_1(d\mathfrak{z},ds)\Big|\mathcal{G}_t\Big) = \sum_{i=0}^{k-1} \mathbb{E}\big(\xi_i\tilde{N}_1(\Gamma_i \times (t_i \wedge t, t_{i+1}\wedge t])\big|\mathcal{G}_t\big).$$

In the same way as equations (III.2.13) and (III.2.14) are obtained, by using (III.2.12) we get

$$\mathbb{E}\big(\xi_i\tilde{N}_1(\Gamma_i\times(t_i,t_{i+1}])\big|\mathcal{G}_t\big) = \mathbb{E}(\xi_i|\mathcal{G}_{t_i})\tilde{N}_1(\Gamma_i\times(t_i,t_{i+1}]) = \mathbb{E}(\xi_i|\mathcal{G}_s)\tilde{N}_1(\Gamma_i\times(t_i,t_{i+1}])$$

for $t_i \leqslant s \leqslant t_{i+1} \leqslant t$, and

$$\mathbb{E}\big(\xi_j\tilde{N}_1(\Gamma_j \times (t_j,t])\big|\mathcal{G}_t\big) = \mathbb{E}(\xi_j|\mathcal{G}_{t_j})\tilde{N}_1(\Gamma_j \times (t_j,t]) = \mathbb{E}(\xi_j|\mathcal{G}_s)\tilde{N}_1(\Gamma_j \times (t_j,t])$$

for $t_j \leqslant s \leqslant t \leqslant t_{j+1}$. Thus for

$$\hat{h}_t^{(1)}(\mathfrak{z}) = \sum_{i=0}^{k-1} \mathbb{E}(\xi_i|\mathcal{G}_{t_i})\mathbf{1}_{(t_i,t_{i+1}]\times\Gamma_i}(t,\mathfrak{z}) = \mathbb{E}(h_t^{(1)}(\mathfrak{z})|\mathcal{G}_t),$$

equations in (III.2.6) and (III.2.8) hold. Assume now that $h^{(1)}$ is $\mathcal{F}\otimes\mathcal{B}([0,T])\otimes\mathcal{Z}$-measurable such that for every $t \in [0,T]$ the function $h_t^{(1)}$ is $\mathcal{F}_t\otimes\mathcal{Z}_1$-measurable and $\mathbb{E}|H^{(1)}|^2 < \infty$, where $H^{(1)}$ is defined in (III.2.1). Then there exist sequences

$(h^n)_{n=1}^{\infty}$ and $(\hat{h}^n)_{n=1}^{\infty}$, such that h^n is a simple function of the form (III.2.22), $\hat{h}^n$ is a $\mathcal{P}_{\mathcal{G}} \otimes \mathcal{Z}_1$-measurable function,

$$\mathbb{E}(\tilde{I}_t(h^n)|\mathcal{G}_t) := \mathbb{E}\Big(\int_0^t \int_{\mathfrak{Z}_1} h_s^n(\mathfrak{z})\, \tilde{N}_1(d\mathfrak{z}, ds)|\mathcal{G}_t\Big) = \int_0^t \int_{\mathfrak{Z}_1} \hat{h}_s^n(\mathfrak{z})\, \tilde{N}_1(d\mathfrak{z}, ds), \tag{III.2.23}$$

$$\hat{h}_t^n(\mathfrak{z}) = \mathbb{E}(h_t^n(\mathfrak{z})|\mathcal{G}_t), \quad \text{almost surely, for } \nu_1(d\mathfrak{z}) \otimes dt\text{-a.e. } (\mathfrak{z}, t) \in \mathfrak{Z}_1 \times [0, T], \tag{III.2.24}$$

for every $n \geqslant 1$, and

$$\lim_{n\to\infty} \mathbb{E} \int_0^T \int_{\mathfrak{Z}_1} |h_t^{(1)}(\mathfrak{z}) - h_t^n(\mathfrak{z})|^2\, \nu_1(d\mathfrak{z})\, dt = 0. \tag{III.2.25}$$

Hence using Jensen's inequality we get

$$\lim_{n,m\to\infty} \mathbb{E} \int_0^T \int_{\mathfrak{Z}_1} |\hat{h}_t^n(\mathfrak{z}) - \hat{h}_t^m(\mathfrak{z})|^2\, \nu_1(d\mathfrak{z}) dt = 0,$$

which implies the existence of a $\mathcal{P}_{\mathcal{G}} \otimes \mathcal{Z}_1$-measurable function $\hat{h}^{(1)}$ such that

$$\lim_{n\to\infty} \mathbb{E} \int_0^T \int_{\mathfrak{Z}_1} |\hat{h}_t^{(1)}(\mathfrak{z}) - \hat{h}_t^n(\mathfrak{z})|^2\, \nu_1(d\mathfrak{z})\, dt = 0. \tag{III.2.26}$$

Thus letting $n \to \infty$ in (III.2.23) we obtain (III.2.6). By virtue of (III.2.25) and (III.2.26) there is a subsequence $n_l \to \infty$ and a set $A \in \mathcal{B}([0, T]) \otimes \mathcal{Z}_1$ such that $dt \otimes \nu_1(A) = 0$ and for $n_l \to \infty$

$$h_t^{n_l}(\mathfrak{z}) \to h_t^{(1)}(\mathfrak{z}) \quad \text{and} \quad \hat{h}_t^{n_l}(\mathfrak{z}) \to \hat{h}_t^{(1)}(\mathfrak{z}) \text{ in mean square}$$

for every $(t, \mathfrak{z}) \in A^c := [0, T] \times \mathfrak{Z}_1 \backslash A$. Consequently,

$$\mathbb{E}(h_t^{n_k}(\mathfrak{z})|\mathcal{G}_t) \to \mathbb{E}(h_t^{(1)}(\mathfrak{z})|\mathcal{G}_t) \quad \text{in mean square for every } (\mathfrak{z}, t) \in A^c,$$

and letting $n := n_l \to \infty$ in (III.2.24) we obtain $\mathbb{E}(h_t^{(1)}(\mathfrak{z})|\mathcal{G}_t) = \hat{h}_t^{(1)}(\mathfrak{z})$ for $(\mathfrak{z}, t) \in A^c$, which proves (III.2.8). To prove the second equation in (III.2.6) assume first that $h^{(0)}$ is a simple function of the form (III.2.22) with $\Gamma_i \in \mathcal{Z}_0$, $\nu_0(\Gamma_i) < \infty$ for $i = 0, 1, ...k-1$. Just like (III.2.21) is obtained, by Lemma III.2.1 we get

$$\mathbb{E}(\xi_i \tilde{N}_0(\Gamma_i \times (t_i \wedge t, t_{i+1} \wedge t])|\mathcal{G}_t) = \mathbb{E}(\xi_i|\mathcal{G}_{t_i})\mathbb{E}\tilde{N}_0(\Gamma_i \times (t_i \wedge t, t_{i+1} \wedge t]) = 0$$

for $i = 0, 1, ..., k-1$ and $t \in [0, T]$, that implies the second equation in (III.2.6). Hence, we obtain the second equation in (III.2.6) for $\mathcal{O}_{\mathcal{F}}$-measurable functions satisfying (III.2.1) by approximation with simple functions. $\square$

We can reformulate the above theorem by using the notion of optional projections of processes with respect to a given filtration. It is well-known (see for instance [25, Thm 5.1], [13, Thm 2.43]) that if $f = (f_t)_{t\in[0,T]}$ is a $\mathcal{B}([0,T])\otimes\mathcal{F}$-measurable process such that f_τ is σ-integrable (with respect to a probability measure P) relative to the σ-algebra $\mathcal{G}_\tau$ for every $\mathcal{G}_t$-stopping time $\tau \leqslant T$ (with respect to a P-complete filtration $(\mathcal{G}_t)_{t\in[0,T]}$), then there exists a unique (up to evanescence) $\mathcal{G}_t$-optional process ${}^o\!f = ({}^o\!f_t)_{t\in[0,T]}$ such that for every $\mathcal{G}_t$-stopping time $\tau \leqslant T$

$$\mathbb{E}(f_\tau|\mathcal{G}_\tau) = {}^o\!f_\tau \quad \text{(a.s.)}.$$

The process ${}^o\!f$ is called the optional projection of f (under P with respect to $(\mathcal{G}_t)_{t\in[0,T]}$). If f is a cadlag process such that almost surely $\sup_{t\leqslant T}|f_t| \leqslant \eta$ for some σ-integrable random variable η with respect to P relative to $\mathcal{G}_0$, then almost surely the trajectories of ${}^o\!f$ have left and right limits at every $t \in (0,T]$ and $[0,T)$, respectively, and moreover, they are also almost surely right-continuous if $(\mathcal{G}_t)_{t\in[0,T]}$ is right-continuous. Notice that for every $t \in [0,T]$ and process f, such that f_τ is σ-integrable relative to $\mathcal{G}_\tau$ for every $\mathcal{G}_t$-stopping time $\tau \leqslant T$, the extended conditional expectations $\mathbb{E}(f_t^+|\mathcal{G}_t)$ and $\mathbb{E}(f_t^-|\mathcal{G}_t)$ are almost surely equal to ${}^o(f_t^+)$ and ${}^o(f_t^-)$, respectively. Let $h = (h_t(\mathfrak{z}))$ be an $\mathcal{F}\otimes\mathcal{B}([0,T])\otimes\mathcal{Z}$-measurable function on $\Omega\times[0,T]\times\mathfrak{Z}$ such that for every $\mathcal{G}_t$-stopping time $\tau \leqslant T$ and $\mathfrak{z} \in \mathfrak{Z}$ the random variable $h_\tau(\mathfrak{z})$ is σ-integrable relative to $\mathcal{G}_\tau$. Then by the help of the Monotone Class Theorem it is not difficult to show the existence of an $\mathcal{O}_\mathcal{G}\otimes\mathcal{Z}$-measurable function, which for each fixed $\mathfrak{z} \in \mathfrak{Z}$ gives the (possibly extended) $\mathcal{O}_\mathcal{G}$-optional projection of $h(\mathfrak{z}) := (h_t(\mathfrak{z}))_{t\in[0,T]}$. We denote this function by ${}^o\!h$, and call it the (extended) $\mathcal{O}_\mathcal{G}$-optional projection of h.

Corollary III.2.3. *Assume the random variables F, $H^{(i)}$ and G, H, defined in* (III.2.1) *and* (III.2.2), *respectively, are σ-integrable relative to $\mathcal{G}_0$ for $i = 0,1$ and that for every $\mathcal{G}_t$-stopping time $\tau \leqslant T$, $\mathfrak{z} \in \mathfrak{Z}$ and $\mathfrak{z}_i \in \mathfrak{Z}_i$ $(i = 0,1)$ the random variables f_τ, g_τ, $h_\tau(\mathfrak{z})$, $h_\tau^{(i)}(\mathfrak{z}_i)$ $(i = 0,1)$ are σ-integrable relative to $\mathcal{G}_0$. Assume moreover that almost surely*

$$\int_0^T |{}^o\!f_t|^2 dt < \infty, \quad \int_0^T \int_{\mathfrak{Z}_i} |{}^o\!h_t^{(i)}(\mathfrak{z})|^2\, \nu_i(d\mathfrak{z})dt < \infty \quad \textit{for } i = 0,1, \tag{III.2.27}$$

where ${}^o\!f$ and ${}^o\!h^{(i)}$ are the (extended) $\mathcal{O}_\mathcal{G}$-optional projections of f and $h^{(i)}$, respectively. Then for every $t \in [0,T]$ equations (III.2.4), (III.2.5) *and* (III.2.6) *hold almost surely with the $\mathcal{O}_\mathcal{G}$-optional projections ${}^o\!f$, ${}^o\!g$, ${}^o\!h^{(i)}$ and ${}^o\!h$ in place of $\hat{f}$, $\hat{g}$, $\hat{h}^{(i)}$ and $\hat{h}$, respectively, for $i = 0,1$. Moreover, there is a dt-null set $T_0 \subset [0,T]$, a $dt\otimes\nu_1$-null set $B_0 \subset [0,T]\times\mathfrak{Z}_1$ and a $dt\otimes\nu$-null set $B \subset [0,T]\times\mathfrak{Z}$, such that*

(i) for each $t \in [0,T]\backslash T_0$ the random variable $|f_t| + |g_t|$ is σ-integrable relative to $\mathcal{G}_0$ and

$$\mathbb{E}(f_t|\mathcal{G}_t) = {}^o\!f_t \in \mathbb{R}, \ \textit{(a.s.)}, \quad \mathbb{E}(g_t|\mathcal{G}_t) = {}^o\!g_t \in \mathbb{R} \ \textit{(a.s.)}, \tag{III.2.28}$$

(ii) for each $(t,\mathfrak{z}) \in [0,T] \times \mathfrak{Z}_1 \backslash B_0$ *the random variable* $|h_t^{(1)}(\mathfrak{z})|$ *is* σ*-integrable relative to* $\mathcal{G}_0$ *and*

$$\mathbb{E}(h_t^{(1)}(\mathfrak{z})|\mathcal{G}_t) = {}^{o}\!h_t{}^{(1)}(\mathfrak{z}) \in \mathbb{R} \text{ (a.s.)}, \tag{III.2.29}$$

(iii) for each $(t,\mathfrak{z}) \in [0,T] \times \mathfrak{Z} \backslash B$ *the random variable* $|h_t(\mathfrak{z})|$ *is* σ*-integrable relative to* $\mathcal{G}_0$*, and*

$$\mathbb{E}(h_t(\mathfrak{z})|\mathcal{G}_t) = {}^{o}\!h_t(\mathfrak{z}) \in \mathbb{R} \text{ (a.s.)}. \tag{III.2.30}$$

Proof. Just like in the proof of the previous lemma without loss of generality we may and will assume that F, G, H and $H^{(i)}$, $i = 0, 1$, have finite expectation. Thus by Minkowski's inequality and Tonelli's theorem we have

$$\left(\int_0^T (\mathbb{E}|f_t|)^2\,dt\right)^{1/2} \leqslant \mathbb{E}\left(\int_0^T |f_t|^2\,dt\right)^{1/2} < \infty, \quad \int_0^T \mathbb{E}|g_t|\,dt = \mathbb{E}\int_0^T |g_t|\,dt < \infty,$$

$$\int_0^T \int_{\mathfrak{Z}} \mathbb{E}|h_t(\mathfrak{z})|\,\nu(d\mathfrak{z})\,dt = \mathbb{E}\int_0^T \int_{\mathfrak{Z}} |h_t(\mathfrak{z})|\,\nu(d\mathfrak{z})\,dt < \infty$$

$$\left(\int_0^T \int_{\mathfrak{Z}_1} (\mathbb{E}|h_t^{(1)}(\mathfrak{z})|)^2\,\nu_1(d\mathfrak{z})\,dt\right)^{1/2} \leqslant \mathbb{E}\left(\int_0^T \int_{\mathfrak{Z}_1} |h_t^{(1)}(\mathfrak{z})|^2 \nu_1(d\mathfrak{z})\,dt\right)^{1/2} < \infty.$$

Therefore $\mathbb{E}|f_t| + \mathbb{E}|g_t| < \infty$ for dt-almost every $t \in [0,T]$, $\mathbb{E}|h_t(\mathfrak{z})| < \infty$ for $dt\otimes\nu$-a.e. $(t,\mathfrak{z}) \in [0,T] \times \mathfrak{Z}$, and $\mathbb{E}|h_t^{(1)}(\mathfrak{z})| < \infty$ for $dt \otimes \nu_1$-a.e. $(t,\mathfrak{z}) \in [0,T] \times \mathfrak{Z}_1$, i.e., we get (III.2.28), (III.2.29) and (III.2.30). Hence due to (III.2.7) and (III.2.9) we have (III.2.5) with ${}^{o}g$ and ${}^{o}\!h$ in place of $\hat{g}$ and $\hat{h}$, respectively. We also have (III.2.4) and (III.2.6) with ${}^{o}\!f$ and ${}^{o}\!h^{(1)}$ in place of $\hat{f}$ and $\hat{h}^{(1)}$, provided F^r and $|H^{(i)}|^2$ are σ-integrable relative to $\mathcal{G}_0$ for $i = 0, 1$ for some $r > 1$. Thus it remains to prove (III.2.4) and (III.2.6) with ${}^{o}\!f$ and ${}^{o}\!h^{(1)}$ in place of $\hat{f}$ and $\hat{h}^{(1)}$, respectively, under the condition that F and $H^{(i)}$ are σ-integrable relative to $\mathcal{G}_0$ for $i = 0, 1$, and (III.2.27) holds. We show only (III.2.6) under these conditions, because (III.2.4) can be proven similarly. To this end define $h^{(1)n} = \mathbf{1}_{\mathfrak{Z}^n}(-n \vee h^{(1)} \wedge n)$ for integers $n \geqslant 1$, where $(\mathfrak{Z}^n)_{n=1}^{\infty}$ is an increasing sequence of sets $\mathfrak{Z}^n \in \mathcal{Z}_1$ such that $\bigcup_{n=1}^{\infty} \mathfrak{Z}^n = \mathfrak{Z}_1$ and $\nu_1(\mathfrak{Z}^n) < \infty$ for every $n \geqslant 1$. Then for each $t \in [0,T]$

$$\mathbb{E}(\delta_t^{(1)n}|\mathcal{G}_t) = \int_0^t \int_{\mathfrak{Z}_1} {}^{o}\!h_s^{(1)n}(\mathfrak{z})\,\tilde{N}_1(d\mathfrak{z}, ds) \quad \text{(a.s.)}, \tag{III.2.31}$$

where $\delta^{(1)n}$ is defined as $\delta^{(1)}$ in (III.2.3), but with $h^{(1)n}$ in place of $h^{(1)}$. Note that

$$|{}^{o}\!h^{(1)n}| \leqslant |{}^{o}\!h^{(1)}| \quad P \otimes dt \otimes \nu_1\text{-almost every } (\omega, t, \mathfrak{z}) \in \Omega \times [0,T] \times \mathfrak{Z}_1,$$

and for $n \to$ we have ${}^{o}\!h_s^{(1)n}(\mathfrak{z}) \to {}^{o}\!h_s^{(1)}(\mathfrak{z})$ almost surely for every $(s,\mathfrak{z}) \in [0,T] \times \mathfrak{Z}_1$

such that ${}^{o}\!h_s^{(1)}(\mathfrak{z}) \neq \infty$. Hence due to condition (III.2.27), by Lebesgue's theorem on dominated convergence we have

$$\int_0^T \int_{\mathfrak{Z}_1} |{}^{o}\!h_s^{(1)n}(\mathfrak{z}) - {}^{o}\!h_s^{(1)}(\mathfrak{z})|^2 \, \nu_1(d\mathfrak{z})dt \to 0 \text{ (a.s.) as } n \to \infty,$$

which implies

$$\int_0^t \int_{\mathfrak{Z}_1} {}^{o}\!h_s^{(1)n}(\mathfrak{z}) \, \tilde{N}_1(d\mathfrak{z}, ds) \to \int_0^t \int_{\mathfrak{Z}_1} {}^{o}\!h_s^{(1)}(\mathfrak{z}) \, \tilde{N}_1(d\mathfrak{z}, ds) \tag{III.2.32}$$

in probability, uniformly in $t \in [0, T]$. Using obvious properties of conditional expectations, by Davis' inequality and Lebesgue's theorem on dominated convergence we get

$$\begin{aligned} \lim_{n\to\infty} \mathbb{E}|\mathbb{E}(\delta_t^{(1)n}|\mathcal{G}_t) - \mathbb{E}(\delta_t^{(1)}|\mathcal{G}_t)| &\leqslant \lim_{n\to\infty} \mathbb{E}|\delta_t^{(1)n} - \delta_t^{(1)}| \\ &\leqslant 3 \lim_{n\to\infty} \mathbb{E}\left(\int_0^T \int_{\mathfrak{Z}_1} |h_s^{(1)n}(\mathfrak{z}) - h_s^{(1)}(\mathfrak{z})|^2 \, \nu_1(d\mathfrak{z})ds\right)^{1/2} = 0, \end{aligned}$$

which by virtue of (III.2.31) and (III.2.32) finishes the proof of the first equation in (III.2.6). The second equation in (III.2.6) can be obtained similarly. □

Remark III.2.1. We have that almost surely

$$\int_0^t \int_{\mathfrak{Z}_i} |{}^{o}\!h_s^{(i)}(\mathfrak{z})|^2 \, \nu_i(d\mathfrak{z}) \, ds \leqslant \int_0^t {}^{o}(|h_s^{(i)}|^2_{L_2(\mathfrak{Z}_i)}) \, ds \quad \text{(a.s.) for } i = 0, 1,$$

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

$$\int_0^T {}^{o}(|h_t^{(i)}|^2_{L_2(\mathfrak{Z}_i)}) \, dt < \infty \quad \text{(a.s.) for } i = 0, 1 \tag{III.2.33}$$

implies the assumption on $h^{(i)}$ in (III.2.27).

Proof. Let $i \in \{0, 1\}$ be fixed and let $(A_n)_{n=1}^\infty$ be an increasing sequence of sets from $\mathfrak{Z}_i$ such that $\cup_{n=1}^\infty A_n = \mathfrak{Z}_i$ and $\nu_i(A_n) < \infty$ for every $n \geqslant 1$. Set

$$h_t^{i,n}(\mathfrak{z}) := (-n) \vee (\mathbf{1}_{A_n} h_t^{(i)}(\mathfrak{z})) \wedge n.$$

Then by Jensen's inequality for the optional projections we have $|{}^{o}h_s^{i,n}(\mathfrak{z})|^2 \leqslant {}^{o}(|h_s^{i,n}(\mathfrak{z})|^2)$ for every $\mathfrak{z} \in \mathfrak{Z}_i$, and by an application of Corollary III.2.3 we obtain

$$\begin{aligned} \int_0^t \int_{\mathfrak{Z}_i} |{}^{o}h_s^{i,n}(\mathfrak{z})|^2 \, \nu_i(d\mathfrak{z})ds &\leqslant \int_0^t \int_{\mathfrak{Z}_i} {}^{o}(|h_s^{i,n}(\mathfrak{z})|^2) \, \nu_i(d\mathfrak{z})ds \\ &= \mathbb{E}\left(\int_0^t \int_{\mathfrak{Z}_i} |h_s^{i,n}(\mathfrak{z})|^2 \, \nu_i(d\mathfrak{z})ds \Big| \mathcal{G}_t\right) \end{aligned}$$

$$= \int_0^t {}^o(|h_s^{i,n}|^2_{L_2(\mathfrak{Z}_i)})\, ds \leqslant \int_0^t {}^o(|h_s^{(i)}|^2_{L_2(\mathfrak{Z}_i)})\, ds.$$

Letting here $n \to \infty$ and using the Monotone Convergence Theorem and the properties of extended optional projections on the left-hand side of the first inequality, we finish the proof of the remark. □

Let $\mathbb{P}(\mathbb{R}^d)$ be the space of of probability measures on the Borel sets of $\mathbb{R}^d$, equipped with the topology of weak convergence of measures. Recall that $C_b(\mathbb{R}^d)$ denotes the space of bounded continuous real functions on $\mathbb{R}^d$, and as before, let $(\mathfrak{Z}, \mathcal{Z})$ be a separable measurable space.

Lemma III.2.4. *Let $(\Omega, \mathcal{F}, P)$ be a complete probability space equipped with a right-continuous filtration $(\mathcal{G}_t)_{t\geqslant 0}$, $\mathcal{G}_t \subset \mathcal{F}$ for $t \geqslant 0$, such that $\mathcal{G}_0$ contains all P-zero sets of $\mathcal{F}$. Let $(X_t)_{t\geqslant 0}$ be an $\mathbb{R}^d$-valued $\mathcal{F} \otimes \mathcal{B}(\mathbb{R}_+)$-measurable cadlag process. Then the following statements hold.*

(i) There is a $\mathbb{P}(\mathbb{R}^d)$-valued weakly cadlag process $(P_t)_{t\geqslant 0}$ such that for every bounded real-valued Borel function φ on $\mathbb{R}^d$ and for each $t \geqslant 0$

$$P_t(\varphi) = \mathbb{E}(\varphi(X_t)|\mathcal{G}_t) \quad (a.s.). \tag{III.2.34}$$

(ii) Let $(P_t)_{t\geqslant 0}$ be the measure-valued process from (i). Assume $f = f(\omega, t, \mathfrak{z}, x)$ is a $\mathcal{O}_{\mathcal{G}} \otimes \mathcal{Z} \otimes \mathcal{B}(\mathbb{R}^d)$-measurable real function on $\Omega \times \mathbb{R}_+ \times \mathfrak{Z} \times \mathbb{R}^d$ such that for every finite $\mathcal{G}_t$-stopping time τ and $(x, \mathfrak{z}) \in \mathbb{R}^d \times \mathfrak{Z}$ the random variable $f_\tau(x, \mathfrak{z})$ is σ-integrable relative to $\mathcal{G}_\tau$. Define

$$P_t(f(t,\mathfrak{z})) := \begin{cases} \int_{\mathbb{R}^d} f(t,\mathfrak{z},x)\, P_t(dx), \ \textit{for } (t,\omega,\mathfrak{z}), \ \textit{if } \int_{\mathbb{R}^d} |f(t,\mathfrak{z},x)|\, P_t(dx) < \infty \\ \infty \ \textit{elsewhere.} \end{cases}$$

Then $P_t(f(t,\mathfrak{z}))$ is an $\mathcal{O}_{\mathcal{G}} \otimes \mathcal{Z}$-measurable (extended) function of $(\omega, t, \mathfrak{z})$ such that

$$\mathbb{E}(f(t,\mathfrak{z},X_t)|\mathcal{G}_t) = P_t(f(t,\mathfrak{z})) \ \ (a.s.) \quad \textit{for each } (t,\mathfrak{z}) \in \mathbb{R}_+ \times \mathfrak{Z}. \tag{III.2.35}$$

Proof. Statement (i) is shown in [52]. Thus (ii) holds if $f = g(t,\mathfrak{z})\varphi(x)$ for bounded $\mathcal{O}_{\mathcal{G}} \otimes \mathcal{Z}$-measurable functions g on $\Omega \times \mathbb{R}_+ \times \mathfrak{Z}$ and bounded Borel functions φ on $\mathbb{R}^d$. Hence by a standard monotone class argument we get (ii) under the additional assumption that f is bounded. In the general case, the set $A \subset \Omega \times \mathbb{R}_+ \times \mathfrak{Z}$ where

$$\int_{\mathbb{R}^d} |f(t,\mathfrak{z},x)| P_t(dx) = \infty$$

is in $\mathcal{O}_{\mathcal{G}} \otimes \mathcal{Z}$. Consequently, $P_t(f(t,\mathfrak{z}))$ is $\mathcal{O}_{\mathcal{G}} \otimes \mathcal{Z}$-measurable in $(\omega, t, \mathfrak{z})$. We have

$$\mathbb{E}(|f(t,\mathfrak{z},X_t)|\wedge n|\mathcal{G}_t)=\int_{\mathbb{R}^d}|f(t,\mathfrak{z},x)|\wedge n\,P_t(dx)\quad\text{(a.s.)}$$

for every integer $n\geqslant 1$. Letting here $n\to\infty$ we get

$$\mathbb{E}(|f(t,\mathfrak{z},X_t)||\mathcal{G}_t)=\int_{\mathbb{R}^d}|f(t,\mathfrak{z},x)|\,P_t(dx)\quad\text{(a.s.)},\tag{III.2.36}$$

that implies (III.2.35). Since $f(t_0,\mathfrak{z}_0,X_{t_0})$ is σ-integrable relative to $\mathcal{G}_{t_0}$, there is an increasing sequence $(\Omega_n)_{n=1}^{\infty}$ such that $\Omega_n\in\mathcal{G}_{t_0}$, $P(\cup_{n=1}^{\infty}\Omega_n)=1$, and

$$\mathbf{1}_{\Omega_n}\int_{\mathbb{R}^d}f(t_0,\mathfrak{z}_0,x)\,P_{t_0}(dx)=\mathbb{E}(\mathbf{1}_{\Omega_n}f_n(t_0,\mathfrak{z}_0,X_{t_0})|\mathcal{G}_{t_0})$$

is almost surely finite for every $n\geqslant 1$. □

Corollary III.2.5. *Let $(\Omega,\mathcal{F},P,(\mathcal{G}_t)_{t\geqslant 0})$ and $(X_t)_{t\in[0,T]}$ be a filtered probability space and a stochastic process, respectively, satisfying the conditions in Lemma III.2.4. Let $(\mathcal{F}_t)_{t\geqslant 0}$ be a filtration such that $\mathcal{G}_t\subset\mathcal{F}_t\subset\mathcal{F}$ for $t\geqslant 0$. Let Q be a probability measure on $\mathcal{F}$ such that $dQ=\gamma_T dP$ for a $\mathcal{F}_T$-measurable positive random variable γ_T. Then the following statements hold.*

(i) There is an $\mathbb{M}(\mathbb{R}^d)$-valued weakly cadlag stochastic process $(\mu_t)_{t\in[0,T]}$ such that for every bounded real-valued Borel function φ on $\mathbb{R}^d$ and for every $t\in[0,T]$

$$\mu_t(\varphi)=\mathbb{E}_Q(\gamma_T^{-1}\varphi(X_t)|\mathcal{G}_t)=\mathbb{E}_Q(\gamma_t^{-1}\varphi(X_t)|\mathcal{G}_t)\,(a.s.).\tag{III.2.37}$$

(ii) Let $f=f(\omega,t,\mathfrak{z},x)$ be a $\mathcal{O}_{\mathcal{G}}\otimes\mathcal{Z}\otimes\mathcal{B}(\mathbb{R}^d)$-measurable real function on $\Omega\times[0,T]\times\mathfrak{Z}\times\mathbb{R}^d$ such that for every finite $\mathcal{G}_t$-stopping time τ and $(x,\mathfrak{z})\in\mathbb{R}^d\times\mathfrak{Z}$ the random variable $f_\tau(x,\mathfrak{z})$ is σ-integrable relative to $\mathcal{G}_\tau$. Define

$$\mu_t(f(t,\mathfrak{z})):=\begin{cases}\int_{\mathbb{R}^d}f(t,\mathfrak{z},x)\,\mu_t(dx), & \text{for } (t,\omega,\mathfrak{z}),\ \text{if } \int_{\mathbb{R}^d}|f(t,\mathfrak{z},x)|\,\mu_t(dx)<\infty\\ \infty & \text{elsewhere.}\end{cases}$$

Then $\mu_t(f(t,\mathfrak{z}))$ is a $\mathcal{O}_{\mathcal{G}}\otimes\mathcal{Z}$-measurable function such that for each $(t,\mathfrak{z})$ we have

$$\mathbb{E}_Q(\gamma_T^{-1}f(t,\mathfrak{z},X_t)|\mathcal{G}_t)=\mathbb{E}_Q(\gamma_t^{-1}f(t,\mathfrak{z},X_t)|\mathcal{G}_t)=\mu_t(f(t,\mathfrak{z}))\quad(a.s.).\tag{III.2.38}$$

Proof. Considering $(\mathcal{F}_{t+})_{t\geqslant 0}$ in place of $(\mathcal{F}_t)_{t\geqslant 0}$ we may assume in the proof that $(\mathcal{F}_t)_{t\geqslant 0}$ is right-continuous. By Doob's theorem there is a cadlag $\mathcal{F}_t$-martingale, $(\gamma_t)_{t\in[0,T]}$, such that $\gamma_t=\mathbb{E}_P(\gamma_T|\mathcal{F}_t)$ (P-a.s) for each $t\in[0,T]$. Clearly, almost surely $\gamma_t>0$ for all $t\in[0,T]$ since

$$0=\mathbb{E}_P(\mathbf{1}_{\gamma_t=0}\gamma_t)=\mathbb{E}_P(\mathbf{1}_{\gamma_t=0}\gamma_T)$$

implies $P(\gamma_t = 0) = 0$ for every $t \in [0, T]$. Thus $(\gamma_t^{-1})_{t\in[0,T]}$ is a cadlag process, and it is an $\mathcal{F}_t$-martingale under Q, because

$$\mathbb{E}_Q(\gamma_T^{-1}|\mathcal{F}_t) = 1/\mathbb{E}_P(\gamma_T|\mathcal{F}_t) = \gamma_t^{-1} \quad \text{almost surely for } t \in [0, T].$$

Since, $\gamma = (\gamma)_{t\in[0,T]}$ is a (cadlag) $\mathcal{F}_t$-martingale under P, the set $\{\gamma_\tau\}$ for $\mathcal{F}_t$-stopping times $\tau \leqslant T$ is uniformly P-integrable, and hence one knows that ${}^o\gamma$, the $\mathcal{G}_t$-optional projection of γ under P, is a cadlag process. Due to $\gamma > 0$, we have ${}^o\gamma > 0$ (a.s.). Define $\mu_t := ({}^o\gamma_t)^{-1}P_t$ for $t \in [0, T]$, where $(P_t)_{t\in[0,T]}$ is the $\mathbb{P}(\mathbb{R}^d)$-valued $\mathcal{G}_t$-adapted cadlag process (in the topology of weak convergence of measures) by Lemma III.2.4. Hence, $(\mu_t)_{t\in[0,T]}$ is a $\mathcal{G}_t$-adapted cadlag $\mathbb{M}(\mathbb{R}^d)$-valued process, and by (III.2.34) for every bounded Borel function φ on $\mathbb{R}^d$ we have

$$\mathbb{E}_Q(\gamma_T^{-1}\varphi(X_t)|\mathcal{G}_t) = \mathbb{E}_P(\varphi(X_t)|\mathcal{G}_t)/\mathbb{E}_P(\gamma_T|\mathcal{G}_t)$$
$$= \mathbb{E}_P(\varphi(X_t)|\mathcal{G}_t)({}^o\gamma_t)^{-1} = ({}^o\gamma_t)^{-1}P_t(\varphi) = \mu_t(\varphi) \quad \text{(a.s) for each } t \in [0, T].$$

On the other hand, by well-known properties of conditional expectations

$$\mathbb{E}_Q(\gamma_T^{-1}\varphi(X_t)|\mathcal{G}_t) = \mathbb{E}_Q(\mathbb{E}_Q(\gamma_T^{-1}\varphi(X_t)|\mathcal{F}_t)|\mathcal{G}_t)$$
$$= \mathbb{E}_Q(\varphi(X_t)\mathbb{E}_Q(\gamma_T^{-1}|\mathcal{F}_t)|\mathcal{G}_t) = \mathbb{E}_Q(\gamma_t^{-1}\varphi(X_t))|\mathcal{G}_t),$$

which completes the proof of (i). To prove (ii), note that the function $\mu_t(f(t, \mathfrak{z}))$ is $\mathcal{O}_\mathcal{G} \otimes \mathcal{Z}$-measurable in $(\omega, t, \mathfrak{z})$, and by (III.2.35) for each $(t, \mathfrak{z})$ almost surely

$$\mu_t(f(t, \mathfrak{z})) = ({}^o\gamma_t)^{-1}P_t(f(t, \mathfrak{z})) = \mathbb{E}(({}^o\gamma_t)^{-1}f(t, \mathfrak{z}, X_t)|\mathcal{G}_t)$$
$$= \mathbb{E}_Q(\gamma_T^{-1}({}^o\gamma_t)^{-1}f(t, \mathfrak{z}, X_t)|\mathcal{G}_t)/\mathbb{E}_Q(\gamma_T^{-1}|\mathcal{G}_t) = \mathbb{E}_Q(\gamma_T^{-1}({}^o\gamma_t)^{-1}f(t, \mathfrak{z}, X_t)|\mathcal{G}_t){}^o\gamma_t$$
$$= \mathbb{E}_Q(\gamma_T^{-1}f(t, \mathfrak{z}, X_t)|\mathcal{G}_t) = \mathbb{E}_Q(\gamma_t^{-1}f(t, \mathfrak{z}, X_t)|\mathcal{G}_t),$$

where the last equation holds because γ^{-1} is an $\mathcal{F}_t$-martingale under Q. We finish the proof with the obvious observation that $\gamma_{t_0}^{-1}f(t_0, \mathfrak{z}_0, X_{t_0})$ is σ-integrable with respect to Q relative to $\mathcal{G}_{t_0}$ if $f(t_0, \mathfrak{z}_0, X_{t_0})$ is σ-integrable with respect to P relative to $\mathcal{G}_{t_0}$. □

III.3 Proof of Theorem III.1.1

Recall that by Assumption III.1.2 the measure Q, defined by $dQ = \gamma_T dP$ is a probability measure, equivalent to P, and by Girsanov's theorem under Q the process $(W_t, \tilde{V}_t)_{t\in[0,T]}$, where $(\tilde{V}_t)_{t\in[0,T]}$ is defined by (III.1.4), is a $d_1 + d'$-dimensional $\mathcal{F}_t$-Wiener process. Moreover, under Q the random measures $\tilde{N}_0$ and $\tilde{N}_1$ remain independent $\mathcal{F}_t$-Poisson martingale measures, with characteristic measures ν_0 and ν_1, respectively, see e.g. [45, Sec. 1.3] for a proof. Clearly,

$(\gamma_t)_{t\in[0,T]}$ is an $\mathcal{F}_t$-martingale under P. By Itô's formula

$$d\gamma_t^{-1} = \gamma_t^{-1} B_t^l(X_t)\, d\tilde{V}_t^l, \tag{III.3.1}$$

the process $\gamma^{-1} = (\gamma_t^{-1})_{t\in[0,T]}$ is an $\mathcal{F}_t$-local martingale under Q. Hence, taking into account $\mathbb{E}_Q\gamma_T^{-1} = 1$, we get that γ^{-1} is an $\mathcal{F}_t$-martingale under Q. Thus the Bayes formula for bounded Borel functions φ on $\mathbb{R}^d$ gives

$$\mathbb{E}(\varphi(X_t)|\mathcal{F}_t^Y) = \frac{\mathbb{E}_Q(\gamma_T^{-1}\varphi(X_t)|\mathcal{F}_t^Y)}{\mathbb{E}_Q(\gamma_T^{-1}|\mathcal{F}_t^Y)} = \frac{\mathbb{E}_Q(\gamma_t^{-1}\varphi(X_t)|\mathcal{F}_t^Y)}{\mathbb{E}_Q(\gamma_t^{-1}|\mathcal{F}_t^Y)} \text{ (a.s.)}, \tag{III.3.2}$$

often also referred as Kallianpur-Striebel formula in the literature. Using $\tilde{V}$ we can rewrite system (I.0.2) in the form

$$\begin{aligned} dX_t =& b(t,Z_t)\,dt + \sigma(t,Z_t)\,dW_t + \rho(t,Z_t)\,dV_t \\ &+ \int_{\mathfrak{Z}_0} \eta(t,Z_{t-},\mathfrak{z})\,\tilde{N}_0(d\mathfrak{z},dt) + \int_{\mathfrak{Z}_1} \xi(t,Z_{t-},\mathfrak{z})\,\tilde{N}_1(d\mathfrak{z},dt), \\ dY_t =& d\tilde{V}_t + \int_{\mathfrak{Z}_1} \mathfrak{z}\,\tilde{N}_1(dt,d\mathfrak{z}), \end{aligned} \tag{III.3.3}$$

which shows, in particular, that $(Y_t)_{t\in[0,T]}$ is a Lévy process under Q, and hence it is well-known that the filtration $(\mathcal{F}_t^Y)_{t\in[0,T]}$ is right-continuous. Thus we can apply Lemma III.2.4 and Corollary III.2.5 with the unobservable process $(X_t)_{t\in[0,T]}$ and the filtration $(\mathcal{G}_t)_{t\in[0,T]} = (\mathcal{F}_t^Y)_{t\in[0,T]}$ to have a $\mathbb{P}$-valued and $\mathbb{M}$-valued weakly cadlag $\mathcal{F}_t^Y$-adapted processes $P_t(dx)$ and $\mu_t(dx)$, respectively, such that for every bounded Borel function φ on $\mathbb{R}^d$ for each $t \in [0,T]$ we have

$$P_t(\varphi) = \mathbb{E}(\varphi(X_t)|\mathcal{F}_t^Y), \quad \mu_t(\varphi) = \mathbb{E}_Q(\gamma_t^{-1}\varphi(X_t)|\mathcal{F}_t^Y) \text{ (a.s.)},$$

and by (III.3.2) it follows that almost surely $P_t = \mu_t/\mu_t(\mathbf{1})$ for all $t \in [0,T]$. To get an equation for $d\mu_t(\varphi)$ for sufficiently smooth functions we calculate first the stochastic differential $d(\gamma_t^{-1}\varphi(X_t))$.

Proposition III.3.1. *Let $\varphi \in C_b^2(\mathbb{R}^d)$. Then for the stochastic differential of $\gamma_t^{-1}\varphi(X_t)$ we have*

$$\begin{aligned} d\big(\gamma_t^{-1}\varphi(X_t)\big) =& \gamma_t^{-1}\mathcal{L}_t\varphi(X_t)\,dt + \gamma_t^{-1}\mathcal{M}_t^l\varphi(X_t)\,d\tilde{V}_t^l + \gamma_t^{-1}\sigma_t^{ik}(X_t)D_i\varphi(X_t)\,dW_t^k \\ &+ \gamma_t^{-1}\int_{\mathfrak{Z}_0} I_t^\eta\varphi(X_{t-})\,\tilde{N}_0(d\mathfrak{z},dt) + \gamma_t^{-1}\int_{\mathfrak{Z}_1} I_t^\xi\varphi(X_{t-})\,\tilde{N}_1(d\mathfrak{z},dt) \\ &+ \gamma_t^{-1}\int_{\mathfrak{Z}_0} J_t^\eta\varphi(X_t)\,\nu_0(d\mathfrak{z})\,dt + \gamma_t^{-1}\int_{\mathfrak{Z}_1} J_t^\xi\varphi(X_t)\,\nu_1(d\mathfrak{z})\,dt. \end{aligned} \tag{III.3.4}$$

Proof. By Itô's formula, see for example in [1] or [27], for $\varphi \in C_b^2(\mathbb{R}^d)$ we have

$$\begin{aligned} d\varphi(X_t) = &\big(\mathcal{L}_t\varphi(X_t) - \rho_t^{il}B_t^l(X_t)D_i\varphi(X_t)\big)\, dt \\ &+ \sigma_t^{ik}(X_t)D_i\varphi(X_t)\, dW_t^k + \rho_t^{il}(X_t)D_i\varphi(X_t)\, d\tilde{V}_t^l \\ &+ \int_{\mathfrak{Z}_0} I_t^\eta\varphi(X_{t-})\, \tilde{N}_0(dt, d\mathfrak{z}) + \int_{\mathfrak{Z}_1} I_t^\xi\varphi(X_{t-})\, \tilde{N}_1(dt, d\mathfrak{z}) \\ &+ \int_{\mathfrak{Z}_0} J_t^\eta\varphi(X_{t-})\, \nu_0(d\mathfrak{z})dt + \int_{\mathfrak{Z}_1} J_t^\xi\varphi(X_{t-})\, \nu_1(d\mathfrak{z})dt, \end{aligned}$$

where we use the notations introduced before the formulation of Theorem III.1.1. Hence using (III.3.1) and the stochastic differential rule for products,

$$d(\gamma_t^{-1}\varphi(X_t)) = \gamma_t^{-1}d\varphi(X_t) + \varphi(X_{t-})\, d\gamma_t^{-1} + d\gamma_t^{-1}d\varphi(X_t),$$

where

$$d\gamma_t^{-1}d\varphi(X_t) = \gamma_t^{-1}\rho_t^{il}B_t^l(X_t)D_i\varphi(X_t)\, dt,$$

we obtain (III.3.4). □

To calculate the conditional expectation (under Q) of the terms in the equation for $\gamma_t^{-1}\varphi(X_t)$, given $\mathcal{F}_t^Y$, we describe below the structure of $\mathcal{F}_t^Y$. For each $t \geqslant 0$ we denote by $\mathcal{F}_t^{\tilde{N}}$ the P-completion of the σ-algebra generated by the random variables $N_1((0,s] \times \Gamma)$ for $s \in (0,t]$ and $\Gamma \in \mathcal{Z}_1$ such that $\nu_1(\Gamma) < \infty$.

Lemma III.3.2. *For every $t \in [0,T]$ we have*

$$\mathcal{F}_t^Y = \mathcal{F}_0^Y \vee \mathcal{F}_t^{\tilde{V}} \vee \mathcal{F}_t^{\tilde{N}_1},$$

where $\mathcal{F}_0^Y \vee \mathcal{F}_t^{\tilde{V}} \vee \mathcal{F}_t^{\tilde{N}_1}$ denotes the P-completion of the smallest σ-algebra containing $\mathcal{F}_0^Y$, $\mathcal{F}_t^{\tilde{V}}$ and $\mathcal{F}_t^{\tilde{N}_1}$.

Proof. From (III.3.3) it immediately follows that

$$\mathcal{F}_t^Y \subseteq \mathcal{F}_0^Y \vee \mathcal{F}_t^{\tilde{V}} \vee \mathcal{F}_t^{\tilde{N}_1}.$$

To prove the reversed inclusion, we claim

$$N^Y((0,t] \times A) = N_1((0,t] \times A) \quad \text{almost surely for all } t \in [0,T] \tag{III.3.5}$$

for every $A \in \mathcal{Z}_1$, where N^Y is the measure of jumps for the process Y. Clearly, $N^Y(d\mathfrak{z}, dt) = N^M(d\mathfrak{z}, dt)$, where N^M is the measure of jumps for the process

$$M_t = \int_0^t \int_{\mathfrak{Z}_1} \mathfrak{z}\, \tilde{N}_1(d\mathfrak{z}, dt), \quad t \geqslant 0,$$

i.e.,

$$N^Y((0,t] \times A) = \sum_{0<s\leqslant t} \mathbf{1}_A(\Delta Y_s) = \sum_{0<s\leqslant t} \mathbf{1}_A(\Delta M_s) \quad A \in \mathcal{Z}_1.$$

To show (III.3.5) let $A = A_0$ be a set from $\mathcal{Z}_1$ such that $\nu_1(A_0) < \infty$. Then

$$M_t^{A_0} := \int_0^t \int_{A_0} \mathfrak{z}\, \tilde{N}_1(d\mathfrak{z}, ds) = \sum_{0<s\leqslant t} p_s \mathbf{1}_{A_0}(p_s) - t \int_{A_0} \mathfrak{z}\, \nu_1(d\mathfrak{z}),$$

where $(p_t)_{t\in[0,T]}$ is the Poisson point process associated with N_1. Hence for $N^0(d\mathfrak{z}, dt)$, the measure of jumps of the process M^{A_0}, we have that almost surely

$$N^0((0,t] \times A_0) = N_1((0,t] \times A_0) \quad \text{for all } t \in [0,T]. \tag{III.3.6}$$

It is not difficult to see that $N^0((0,t] \times A_0) = N^M((0,t] \times A_0)$. Hence (III.3.5) for $A = A_0$ follows.

Since ν_1 is σ-finite, for an arbitrary $B \in \mathcal{Z}_1$ there is a sequence $(B_n)_{n=1}^\infty$ of disjoint sets $B_n \in \mathcal{Z}_1$ such that $B = \bigcup_{n=1}^\infty B_n$ and $\nu_1(B_n) < \infty$ for each $n \geqslant 1$. Thus for each integer $n \geqslant 1$ we have (III.3.5) with B_n in place of A, and summing this up over $n \geqslant 1$ and using the σ-additivity of N^Y and N_1 we obtain (III.3.5) with B in place of A. Noting that $\mathcal{F}_t^Y$ contains the σ-algebra generated by $N^Y((0,s] \times B)$ for each $s \leqslant t$ and $B \in \mathcal{Z}$, we see that $\mathcal{F}_t^Y \supset \mathcal{F}_t^{\tilde{N}_1}$. Clearly, $\mathcal{F}_t^Y \supset \mathcal{F}_0^Y$, and taking into account

$$Y_t - Y_0 - \int_0^t \int_{\mathfrak{Z}_1} \mathfrak{z}\, \tilde{N}_1(d\mathfrak{z}, ds) = \tilde{V}_t, \quad \text{for } t \in [0,T],$$

we get $\mathcal{F}_t^Y \supset \mathcal{F}_t^{\tilde{V}}$. Consequently,

$$\mathcal{F}_t^Y \supseteq \mathcal{F}_0^Y \vee \mathcal{F}_t^{\tilde{V}} \vee \mathcal{F}_t^{\tilde{N}},$$

that completes the proof. □

The above lemma is an essential tool in obtaining the filtering equations. A similar lemma in a more general setting in some directions is presented in [50] and [49] to obtain the filtering equations for the model considered in these papers. It seems to us, however, that this lemma, Lemma 3.2 in [50], used as well in [49, p.4], may not hold under the general conditions formulated in these papers, since it is not true in the simple case of vanishing coefficients in front of the random measures in the observation process. It is worth noticing that when instead of the integrand $\mathfrak{z}$ a stochastic integrand depending on $Z_t = (X_t, Y_t)$ is considered in the observation process Y, the integral of such a term against a Poisson random measure may fail to be a Lévy process, as it may not have independent increments, which is a crucial property for the filtration generated by the observation.

Now we are going to get an equation for $\mu(\varphi)$ by noting that by Proposition III.3.1 we have

$$\gamma_t^{-1}\varphi(X_t) = \varphi(X_0) + \alpha_t + \alpha_t^0 + \alpha_t^1 + \beta_t^0 + \beta_t^1 + \delta_t^0 + \delta_t^1, \quad t \in [0,T], \tag{III.3.7}$$

where

$$\alpha_t := \int_0^t \gamma_s^{-1}\mathcal{L}_s\varphi(X_s)\,ds,$$

$$\alpha_t^0 := \int_0^t\int_{\mathfrak{Z}_0} \gamma_s^{-1}J_s^\eta\varphi(X_s)\,\nu_0(d\mathfrak{z})ds, \quad \alpha_t^1 := \int_0^t\int_{\mathfrak{Z}_1} \gamma_s^{-1}J_s^\xi\varphi(X_s)\,\nu_1(d\mathfrak{z})ds,$$

$$\beta_t^0 := \int_0^t \gamma_s^{-1}\sigma_s^{ik}(X_s)D_i\varphi(X_s)\,dW_s^k, \quad \beta_t^1 := \int_0^t \gamma_s^{-1}\mathcal{M}_s^l\varphi(X_s)\,d\tilde{V}_s^l,$$

$$\delta_t^0 := \int_0^t\int_{\mathfrak{Z}_0} \gamma_s^{-1}I_s^\eta\varphi(X_{s-})\,\tilde{N}_0(d\mathfrak{z},ds), \quad \delta_t^1 := \int_0^t\int_{\mathfrak{Z}_1} \gamma_s^{-1}I_s^\xi\varphi(X_{t-})\,\tilde{N}_1(d\mathfrak{z},ds),$$

for $\varphi \in C_b^2(\mathbb{R}^d)$. We want to take the conditional expectation of both sides of equation (III.3.7) for each $t \in [0,T]$, under Q, given $\mathcal{F}_t^Y$. In order to apply Corollary III.2.3, we should verify that the random variables

$$G := \int_0^T \gamma_s^{-1}|\mathcal{L}_t\varphi(X_s)|\,ds,$$

$$G^{(0)} := \int_0^T\int_{\mathfrak{Z}_0} \gamma_s^{-1}|J_s^\eta\varphi(X_s)|\,\nu_0(d\mathfrak{z})\,ds, \quad G^{(1)} := \int_0^T\int_{\mathfrak{Z}_1} \gamma_s^{-1}|J_s^\xi\varphi(X_s)|\,\nu_1(d\mathfrak{z})\,ds,$$

$$F^{(0)} := \Big(\int_0^T \gamma_s^{-2}|\sigma_s^i(X_s)D_i\varphi(X_s)|^2\,ds\Big)^{1/2}, \quad F^{(1)} := \Big(\int_0^T \gamma_s^{-2}\sum_l|\mathcal{M}_s^l\varphi(X_s)|^2\,ds\Big)^{1/2},$$

$$H^{(0)} := \Big(\int_0^T\int_{\mathfrak{Z}_0} \gamma_s^{-2}|I_s^\eta\varphi(X_{s-})|^2\,\nu_0(d\mathfrak{z})ds\Big)^{1/2},$$

$$H^{(1)} := \Big(\int_0^T\int_{\mathfrak{Z}_1} \gamma_s^{-2}|I_s^\xi\varphi(X_{s-})|^2\,\nu_1(d\mathfrak{z})ds\Big)^{1/2}$$

are σ-integrable with respect to Q relative to $\mathcal{F}_0^Y$, and that (III.2.27) holds for ${}^Qf^{(i)}$ in place of ${}^of$, and for ${}^Qh^{(i)}$ in place of ${}^oh^{(i)}$, where ${}^Qf^{(0)}$, ${}^Qf^{(1)}$, ${}^Qh^{(0)}$ and ${}^Qh^{(1)}$ are the $\mathcal{F}_t^Y$-optional projection under Q of

$$f^{k(0)} := (\gamma_s^{-1}\sigma_s^{ik}(X_s)D_i\varphi(X_s))_{s\in[0,T]}, \quad f^{l(1)} := (\gamma_s^{-1}\mathcal{M}_s^l\varphi(X_s))_{s\in[0,T]},$$

$$h^{(0)} := (\gamma_s^{-1}I_s^\eta\varphi(X_{s-}))_{s\in[0,T]} \quad \text{and} \quad h^{(1)} := (\gamma_s^{-1}I_s^\xi\varphi(X_{s-}))_{s\in[0,T]},$$

respectively for each fixed $k = 1,2,...,d_1$ and $l = 1,...,d'$. For a fixed integer $n \geqslant 1$ let $\Omega_n = \{\omega \in \Omega : |Y_0| \leqslant n\}$. Then due to Assumption III.1.1, the martingale property of $(\gamma_t)_{t\in[0,T]}$ and (III.1.2) we have

$$\mathbb{E}_Q(\mathbf{1}_{\Omega_n}G) \leqslant N\mathbb{E}\Big(\gamma_T\int_0^T \gamma_t^{-1}(K_0 + K_1\mathbf{1}_{\Omega_n}|Z_t| + K_2\mathbf{1}_{\Omega_n}|Z_t|^2)\,dt\Big)$$

$$= N\int_0^T \mathbb{E}(\gamma_T\gamma_t^{-1}(K_0 + K_1\mathbf{1}_{\Omega_n}|Z_t| + K_2\mathbf{1}_{\Omega_n}|Z_t|^2)\,dt$$

$$= N\int_0^T \mathbb{E}(K_0 + K_1\mathbf{1}_{\Omega_n}|Z_t| + K_2\mathbf{1}_{\Omega_n}|Z_t|^2)\,dt$$

$$\leqslant N'(K_0 + K_1\mathbb{E}|X_0| + K_1\mathbb{E}(\mathbf{1}_{\Omega_n}|Y_0|) + K_2\mathbb{E}|X_0|^2 + K_2\mathbb{E}(\mathbf{1}_{\Omega_n}|Y_0|^2)) < \infty$$

with constants N and N', which shows that G is σ-integrable with respect to Q relative to $\mathcal{F}_0^Y$. Similarly, using the estimate

$$|J^\eta\varphi(X_t)| \leqslant \sup_{x\in\mathbb{R}^d}|D_{ij}\varphi(x)||\eta_t^i(X_t)||\eta_t^j(X_t)|,$$

we get

$$\mathbb{E}_Q(\mathbf{1}_{\Omega_n}G^{(0)}) = \int_0^T \mathbb{E}\int_{\mathfrak{Z}_0}\mathbf{1}_{\Omega_n}|J_s^\eta\varphi(X_s)|\,\nu_0(d\mathfrak{z})ds$$

$$\leqslant N\int_0^T \mathbb{E}\int_{\mathfrak{Z}_0}\mathbf{1}_{\Omega_n}|\eta(s,Z_s,\mathfrak{z})|^2\,\nu_0(d\mathfrak{z})\,ds \leqslant N'\int_0^T \mathbb{E}(K_0 + K_2\mathbf{1}_{\Omega_n}|Z_s|^2)\,ds < \infty$$

with constants N and N'. In the same way we get $\mathbb{E}_Q(\mathbf{1}_{\Omega_n}G^{(1)}) < \infty$. To prove that $F^{(i)}$ and $H^{(i)}$ are σ-integrable (with respect to Q) relative to F_0^Y, we claim first that

$$A_n := \mathbb{E}_Q\mathbf{1}_{\Omega_n}\sup_{t\leqslant T}\gamma_t^{-1} < \infty \quad \text{for every integer } n \geqslant 1. \tag{III.3.8}$$

To prove this we repeat a method used in proof of Theorem III.1.2. From (III.3.1) by using the Davis inequality and then Young's inequality we get

$$\mathbb{E}_Q\mathbf{1}_{\Omega_n}\sup_{t\in[0,T]}\gamma_{t\wedge\tau_k}^{-1} \leqslant 1 + 3\mathbb{E}\left(\int_0^{T\wedge\tau_k}\mathbf{1}_{\Omega_n}\gamma_t^{-2}|B(t,Z_t)|^2\,dt\right)^{1/2}$$

$$\leqslant 1 + \tfrac{1}{2}\mathbb{E}_Q\mathbf{1}_{\Omega_n}\sup_{t\in[0,T]}\gamma_{t\wedge\tau_k}^{-1} + 5\mathbb{E}\int_0^T \mathbf{1}_{\Omega_n}\gamma_t^{-1}|B(t,Z_t)|^2\,dt$$

for stopping times

$$\tau_k = \inf\{t \in [0,T] : \gamma_t^{-1} \geqslant k\}, \quad \text{for integers } k \geqslant 1.$$

Rearranging this inequality and then letting $k \to \infty$ by Fatou's lemma we obtain

$$\mathbb{E}_Q\mathbf{1}_{\Omega_n}\sup_{t\in[0,T]}\gamma_t^{-1} \leqslant 2 + 10\int_0^T \mathbb{E}_Q\mathbf{1}_{\Omega_n}\gamma_t^{-1}|B(t,Z_t)|^2\,dt.$$

Hence we get (III.3.8) by noticing that using the martingale property of γ, the estimate in (III.1.2) and $K_2\mathbb{E}|X_0|^2 < \infty$, for every $t \in [0,T]$ we have

$$\mathbb{E}_Q\mathbf{1}_{\Omega_n}\gamma_t^{-1}|B(t,Z_t)|^2 = \mathbb{E}\mathbf{1}_{\Omega_n}\gamma_T\gamma_t^{-1}|B(t,Z_t)|^2 = \mathbb{E}\mathbf{1}_{\Omega_n}|B(t,Z_t)|^2$$

$$\leqslant K_0 + K_2\mathbb{E}\mathbf{1}_{\Omega_n}|Z_t|^2 \leqslant K_0 + K_2N\mathbb{E}(1 + |X_0|^2 + \mathbf{1}_{\Omega_n}|Y_0|^2) < \infty.$$

Consequently,

$$\mathbb{E}_Q(\mathbf{1}_{\Omega_n}F^{(0)}) \leqslant \mathbb{E}_Q\left(\mathbf{1}_{\Omega_n}\sup_{s\leqslant T}\gamma_s^{-1/2}\Big(\int_0^T \mathbf{1}_{\Omega_n}\gamma_s^{-1}|\sigma_s^i(X_s)D_i\varphi(X_s)|^2\,ds\Big)^{1/2}\right) \leqslant A_n+B_n,$$

with $A_n<\infty$, and

$$B_n := \mathbb{E}_Q\int_0^T \mathbf{1}_{\Omega_n}\gamma_s^{-1}|\sigma_s^i(X_s)D_i\varphi(X_s)|^2\,ds = \int_0^T \mathbb{E}(\mathbf{1}_{\Omega_n}\gamma_T\gamma_s^{-1}|\sigma_s^i(X_s)D_i\varphi(X_s)|^2)\,ds$$

$$= \int_0^T \mathbb{E}|\mathbf{1}_{\Omega_n}\sigma_s^i(X_s)D_i\varphi(X_s)|^2\,ds \leqslant N\int_0^T \mathbb{E}(K_0+K_2\mathbf{1}_{\Omega_n}|Z_s|^2)\,ds<\infty.$$

We get $\mathbb{E}_Q(\mathbf{1}_{\Omega_n}F^{(1)})<\infty$ in the same way. Similarly, $\mathbb{E}_Q(\mathbf{1}_{\Omega_n}H^{(0)})\leqslant A_n+C_n$, with A_n given in (III.3.8) and

$$C_n := \mathbb{E}_Q\int_0^T\gamma_s^{-1}\int_{\mathfrak{Z}_0}\mathbf{1}_{\Omega_n}|I_s^\eta\varphi(X_s)|^2\,\nu_0(d\mathfrak{z})ds = \int_0^T\mathbb{E}\int_{\mathfrak{Z}_0}\mathbf{1}_{\Omega_n}|I_s^\eta\varphi(X_s)|^2\,\nu_0(d\mathfrak{z})ds$$

$$\leqslant N\int_0^T\mathbb{E}\int_{\mathfrak{Z}_0}\mathbf{1}_{\Omega_n}|\eta(s,Z_s,\mathfrak{z})|^2\,\nu_0(d\mathfrak{z})ds \leqslant N'\int_0^T\mathbb{E}(K_0+K_2\mathbf{1}_{\Omega_n}|Z_s|^2)\,ds<\infty$$

with constants N and N', where we use that by Taylor's formula we have

$$|I_s^\eta\varphi(X_s)| \leqslant \sup_{x\in\mathbb{R}^d}|D_i\varphi(x)||\eta_s^i(X_s)|.$$

In the same way we have $\mathbb{E}_Q(\mathbf{1}_{\Omega_n}H^{(1)})<\infty$. For processes $h=(h_t)_{t\in[0,T]}$ recall that ${}^{Q}h$ and ${}^{o}h$ denote the $\mathcal{F}_t^Y$-optional projections of h under Q and under P, respectively. Then using the formula ${}^{Q}h={}^{o}(\gamma h)/{}^{o}\gamma$, well-known properties of optional projections and Remark III.2.1 we have

$$|{}^{Q}h^{(0)}|^2_{L_2(\mathfrak{Z}_0)} = \frac{|{}^{o}(I^\eta\varphi(X))|^2_{L_2(\mathfrak{Z}_0)}}{({}^{o}\gamma)^2} \leqslant \frac{{}^{o}\big(|I^\eta\varphi(X))|^2_{L_2(\mathfrak{Z}_0)}\big)}{({}^{o}\gamma)^2}$$

$$\leqslant N\frac{{}^{o}(K_0+K_2|Z|^2)}{({}^{o}\gamma)^2} = N\frac{K_0}{({}^{o}\gamma)^2}+NK_2\frac{{}^{o}(|X|^2)}{({}^{o}\gamma)^2}+NK_2\frac{|Y|^2}{({}^{o}\gamma)^2}$$

with a constant N. Remember that since $\gamma=(\gamma)_{t\in[0,T]}$ is a (cadlag) $\mathcal{F}_t$-martingale under P, the set $\{\gamma_\tau\}$ for $\mathcal{F}_t$-stopping times $\tau\leqslant T$ is uniformly P-integrable and hence due to the right-continuity of $(\mathcal{F}_t^Y)_{t\in[0,T]}$, the optional projection ${}^{o}\gamma$ is a cadlag process. Moreover, due to $\gamma>0$, we have ${}^{o}\gamma>0$ (a.s.). Since by (III.1.2)

$$K_2\mathbb{E}(\sup_{t\leqslant T}\mathbf{1}_{\Omega_n}|X_t|^2)<\infty \quad \text{for every } n\geqslant 1,$$

(and $(\mathcal{F}_t^Y)_{t\in[0,T]}$ is right-continuous), the process $K_2{}^{o}(|X|^2)$ is a cadlag process.

Consequently, $K_0/|{}^o\gamma|^2$, $K_2{}^o(|X|^2)/|{}^o\gamma|^2$ and $|Y|^2/|{}^o\gamma|^2$ are cadlag processes. Hence

$$\int_0^T \frac{1}{({}^o\gamma_s)^2}\,ds + K_2\int_0^T \frac{{}^o(|X|^2)_s}{({}^o\gamma_s)^2}\,ds + K_2\int_0^T \frac{|Y_s|^2}{({}^o\gamma_s)^2}\,ds < \infty \quad \text{(a.s.)},$$

which proves

$$\int_0^T\int_{\mathfrak{Z}_i} |{}^Q h_s^{(i)}|^2\,\nu_i(d\mathfrak{z})\,ds < \infty \quad \text{(a.s.)}$$

for $i = 0$, and we get this for $i = 1$ in the same way. By the same argument we have

$$\int_0^T |{}^Q f_s^{(i)}|^2\,ds < \infty \quad \text{(a.s.)} \quad \text{for } i = 0, 1.$$

Thus we can apply Corollary III.2.3 to the processes α, α^i, β^i and δ^i (i=0,1), and then use Corollary III.2.5, to get

$$\mathbb{E}_Q(\alpha_t|\mathcal{F}_t^Y) = \int_0^t \mu_s(\mathcal{L}_s\varphi)\,ds,$$

$$\mathbb{E}_Q(\alpha_t^0|\mathcal{F}_t^Y) = \int_0^t\int_{\mathfrak{Z}_0} \mu_s(J_s^\eta\varphi)\,\nu_0(d\mathfrak{z})ds, \quad \mathbb{E}_Q(\alpha_t^1|\mathcal{F}_t^Y) = \int_0^t\int_{\mathfrak{Z}_1} \mu_s(J_s^\xi\varphi)\,\nu_1(d\mathfrak{z})ds,$$

$$\mathbb{E}_Q(\beta_t^0|\mathcal{F}_t^Y) = 0, \quad \mathbb{E}_Q(\beta_t^1|\mathcal{F}_t^Y) = \int_0^t \mu_s(\mathcal{M}_s^l\varphi)\,d\tilde{V}_s^l,$$

$$\mathbb{E}_Q(\delta_t^0|\mathcal{F}_t^Y) = 0, \quad \mathbb{E}_Q(\delta_t^1|\mathcal{F}_t^Y) = \int_0^t\int_{\mathfrak{Z}_1} \mu_s(I_s^\xi\varphi)\,\tilde{N}_1(d\mathfrak{z}, ds)$$

for $t \in [0, T]$ and $\varphi \in C_b^2(\mathbb{R}^d)$ almost surely, where $(\mu_t)_{t\in[0,T]}$ is an $\mathbb{M}(\mathbb{R}^d)$-valued $\mathcal{F}_t^Y$-adapted weakly cadlag process such that

$$\mu_t(\varphi) := \int_{\mathbb{R}^d} \varphi(x)\,\mu_t(dx) = \mathbb{E}_Q(\gamma_t^{-1}\varphi(X_t)|\mathcal{F}_t^Y) \quad \text{(a.s.)} \quad \text{for each } t \in [0, T],$$

for every bounded Borel function φ on $\mathbb{R}^d$. Using Lemma III.2.1 with random variables $X := \varphi(X_0)$, $Y := 1$ and σ-algebras $\mathcal{G}_1 := \mathcal{F}_0^Y$, $\mathcal{G} := \mathcal{F}_0$ and $\mathcal{G}_2 := \mathcal{F}_t^{\tilde{V}} \vee \mathcal{F}_t^{\tilde{N}_1}$ we get

$$\mathbb{E}_Q(\varphi(X_0)|\mathcal{F}_t^Y) = \mathbb{E}_Q(\varphi(X_0)|\mathcal{F}_0^Y) = \mu_0(\varphi) \quad \text{(a.s.)}.$$

Consequently, taking the conditional expectation of both sides of equation (III.3.7) under Q given $\mathcal{F}_t^Y$, we see that equation (III.1.6) holds for each $t \in [0, T]$ and $\varphi \in C_b^2(\mathbb{R}^d)$ almost surely, that implies that for each $\varphi \in C_b^2(\mathbb{R}^d)$ equation (III.1.6) holds almost surely for all $t \in [0, T]$, since we have cadlag processes in both sides of equation (III.1.6) for each $\varphi \in C_b^2(\mathbb{R}^d)$. To prove (III.1.7) first notice that for

$\varphi := \mathbf{1}$ equation (III.1.6) gives

$$d\mu_t(\mathbf{1}) = \mu_t(B_t^k)\, d\tilde{V}_t^k, \quad \mu_0(\mathbf{1}) = 1.$$

Since $\mu_t(\mathbf{1}) = ({}^o\gamma_t)^{-1} P_t(\mathbf{1}) = ({}^o\gamma_t)^{-1}$, $t \in [0,T]$, is a continuous process such that $\mu_t(\mathbf{1}) = \mathbb{E}_Q(\gamma_t^{-1} | \mathcal{F}_t^Y)$ (a.s.) for each $t \in [0,T]$, it is the $\mathcal{F}_t^Y$-optional projection under Q of the positive process $(\gamma_t^{-1})_{t\in[0,T]}$. Hence $\lambda_t := \mu_t(\mathbf{1})$, $t \in [0,T]$, is a positive process, and by Itô's formula

$$d\lambda_t^{-1} = -\lambda_t^{-2}\mu_t(B_t^k)\, d\tilde{V}_t^k + \lambda_t^{-3}\sum_k \mu_t^2(B_t^k)\, dt.$$

By Itô's formula for the product $P_t(\varphi) = \lambda_t^{-1}\mu_t(\varphi)$ we have

$$\begin{aligned} dP_t(\varphi) = P_t(\mathcal{L}_t\varphi)\, dt + P_t(\mathcal{M}_t^k\varphi)\, d\tilde{V}_t^k + \int_{\mathfrak{Z}_0} P_t(J_t^\eta\varphi)\, \nu_0(d\mathfrak{z})dt + \int_{\mathfrak{Z}_1} P_t(J_t^\xi\varphi)\, \nu_1(d\mathfrak{z})dt \\ + \int_{\mathfrak{Z}_1} P_t(I_t^\xi\varphi)\, \tilde{N}_1(d\mathfrak{z}, dt) + \lambda_t^{-3}\mu_t(\varphi)\sum_k \mu_t^2(B_t^k)\, dt \\ -\mu_t(\varphi)\lambda_t^{-2}\mu_t(B_t^k)\, d\tilde{V}_t^k - \lambda_t^{-2}\mu_t(B_t^k)\mu_t(\mathcal{M}_t^k\varphi)\, dt \end{aligned}$$

Hence noting that

$$\lambda_t^{-3}\mu_t(\varphi)\sum_k \mu_t^2(B_t^k) = P_t(\varphi)\sum_k P_t^2(B_t^k), \quad \mu_t(\varphi)\lambda_t^{-2}\mu_t(B_t^k) = P_t(\varphi)P_t(B_t^k)$$

$$\lambda_t^{-2}\mu_t(B_t^k)\mu_t(\mathcal{M}_t^k\varphi) = P_t(B_t^k)P_t(\mathcal{M}_t^k\varphi),$$

we obtain

$$\begin{aligned} dP_t(\varphi) = P_t(\mathcal{L}_t\varphi)\, dt + \big(P_t(\mathcal{M}_t^k\varphi) - P_t(\varphi)P_t(B_t^k)\big)\, d\tilde{V}_t^k \\ -\big(P_t(\mathcal{M}_t^k\varphi) - P_t(\varphi)P_t(B_t^k)\big)P_t(B_t^k)\, dt \\ + \int_{\mathfrak{Z}_0} P_t(J_t^\eta\varphi)\, \nu_0(d\mathfrak{z})dt + \int_{\mathfrak{Z}_1} P_t(J_t^\xi\varphi)\, \nu_1(d\mathfrak{z})dt + \int_{\mathfrak{Z}_1} P_t(I_t^\xi\varphi)\, \tilde{N}_1(d\mathfrak{z}, dt). \end{aligned}$$

Since clearly,

$$\begin{aligned} \big(P_t(\mathcal{M}_t^k\varphi) - P_t(\varphi)P_t(B_t^k)\big)\, d\tilde{V}_t^k - \big(P_t(\mathcal{M}_t^k\varphi) - P_t(\varphi)P_t(B_t^k)\big)P_t(B_t^k)\, dt \\ = \big(P_t(\mathcal{M}_t^k\varphi) - P_t(\varphi)P_t(B_t^k)\big)\, d\bar{V}_t^k \end{aligned}$$

with the process $(\bar{V}_t)_{t\in[0,T]}$, given by $d\bar{V}_t = d\tilde{V}_t - P_t(B_t)\, dt$, $\bar{V}_0 = 0$, this gives equation (III.1.7), and finishes the proof of Theorem III.1.1.

Chapter IV

The filtering density

IV.1 Introduction

In Chapter III we were interested in the equations for the evolution of the conditional distribution $P_t(dx) = P(X_t \in dx|Y_s, s \leqslant t)$ of the unobserved component X_t given the observations $(Y_s)_{s\in[0,T]}$, where $Z = (Z_t)_{t\in[0,T]} = (X_t, Y_t)_{t\in 0,T]}$ is given by the stochastic differential equation (I.0.2) and where the coefficients satisfy the measurability conditions, and dimensionality properties, stated in the introduction, Chapter I. In the present chapter we investigate the existence of the conditional density $\pi_t = dP_t/dx$ of the signal-observation system (I.0.2). More precisely, we show, under fairly general conditions, that if the conditional distribution of X_0 given Y_0 has a density π_0, such that its $L_p(\mathbb{R}^d)$-norm has a finite p-th moment, in other words $\mathbb{E}|\pi_0|^p_{L_p} < \infty$ for some $p \geqslant 2$, then X_t for every t has a conditional density π_t given $(Y_s)_{t\in[0,t]}$, which belongs also to L_p, almost surely for all t. This chapter is based on the article [17].

We do not assume any non-degeneracy conditions on σ and η, i.e., they are allowed to vanish. Thus, given the observations, there may not remain any randomness to smooth the conditional distribution $P_t(dx)$ of X_t, i.e., if the initial conditional density π_0 does not exists, then the conditional density π_t for $t > 0$ may not exist either. Therefore assuming that the initial conditional density π_0 exists, we are interested in the smoothness and growth conditions which we should require from the coefficients in order to get that π_t exists for every $t \in [0, T]$ as well.

For partially observed diffusion processes, i.e., when $\xi = \eta = 0$ and the observation process Y does not have jumps, the existence and the regularity properties of the conditional density π_t have been extensively studied in the literature. For important results under non-degeneracy conditions see, for example, [36], [39], [35], [46], and the references therein. Without any non-degeneracy assumptions, in [51] the existence of π_t is proved if $\pi_0 \in W^2_p \cap W^2_2$ for some $p \geqslant 2$, the coefficients are bounded, σ, ρ have uniformly bounded derivatives in x up to order 4, and b, B have uniformly bounded derivatives in x up to order 3. Under these conditions it is also proved that $(\pi_t)_{t\in[0,T]}$ is a weakly continuous process with values in $W_p^2 \cap W_2^2$, and that π_t has higher regularity if π_0 and the coefficients are appropriately smoother. In [41] the existence of conditional densities in $L_2(\mathbb{R}^d)$ was proved using a very nice method, deriving a priori estimates for an SPDE for the unnormalised conditional distribution smoothed with Gaussian kernels. More precisely, it was shown, without assuming differentiability conditions on the coefficients, that if they are bounded, Lipschitz continuous in space and if the initial conditional density π_0 satisfies $\mathbb{E}|\pi_0|_{L_2}^2 < \infty$, then π_t remains in L_2 for all t.

More recently also filtering densities associated to systems with jumps have been investigated, i.e. when ξ, η are not zero and the observation may also contain jump terms. However, to the best of the authors knowledge, most results treat only the case of L_2-valued densities.

Indeed, the result from [41] was also obtained with the same methods in [4] for the case when the observation is driven by an Ornstein-Uhlenbeck process independent of the signal. This smoothing approach is used again in [5] to prove uniqueness of measure-valued solutions for the Zakai equation in the case where the signal is a diffusion process, the observation contains a jump term and the coefficients are time-independent, globally Lipschitz, except for the observation drift term, which contains a time dependence, but is bounded and globally Lipschitz. The approach from [41] is extended in [44] to partially observed jump diffusions when the Wiener process in the observation process Y is independent of the Wiener process in the unobserved process, to prove, in particular, the existence of the conditional density in L_2, if the initial conditional density exists, belongs to L_2, the coefficients are bounded Lipschitz functions, the coefficients of the random measures in the unobservable process are differentiable in x and satisfy a condition in terms of their Jacobian. Another application of this method, yielding an analogous result in L_2, can be found in [6] for the case when the coefficients satisfy Lipschitz and linear growth conditions, the signal has a bounded cadlag disturbance with bounded variation, adapted to the filtration generated by Y, the observation has no jump terms and where additionally π_0 has finite third moment. In [50] and [49] the filtering equations for fairly general filtering models with partially observed jump diffusions are obtained and studied, but the existence of the conditional density (in L_2) is proved only in [50], in the special case when the equation for the unobserved process is driven by a Wiener process and an α-stable additive Lévy process, $\rho = 0$, the coefficients b and σ are bounded functions of $x \in \mathbb{R}^d$, b has bounded first order derivatives, σ has bounded derivatives up to second order and $B = B(t, x, y)$ is a bounded Lipschitz function in $z = (x, y)$.

The main theorem, Theorem IV.2.1, of the present chapter reads as follows. Assume that the coefficients b, σ, ρ, B, ξ, η and ρB are Lipschitz continuous in $z = (x, y) \in \mathbb{R}^{d+d'}$, B is bounded, b, σ, ρ, ξ and η satisfy a linear growth condition, ξ and η admit uniformly equicontinuous derivatives in $x \in \mathbb{R}^d$, $x + \xi(x)$, $x + \eta(x)$ are bijective mappings in $x \in \mathbb{R}^d$, and have a Lipschitz continuous inverse with Lipschitz constant independent of the other variables. Assume, moreover, that

$\mathbb{E}|X_0|^r < \infty$ for some $r > 2$ and that ν_1 has finite r-th moment. Under these conditions, if the initial conditional density π_0 exists for some $p \geqslant 2$, then the conditional density π_t exists and belongs to L_p for every t. Moreover, $(\pi_t)_{t\in[0,T]}$ is weakly cadlag as L_p-valued process.

To prove our main theorem we use the Itô formula from [22] and adapt an approach from [41] to estimate the L_p-norm of the smoothed unnormalised conditional distribution for even integers $p \geqslant 2$. Hence we obtain Theorem IV.2.1 for even integers $p \geqslant 2$. Then we use an interpolation theorem combined with an approximation procedure to get the main theorem for every $p \geqslant 2$.

The chapter is organised as follows. In Section IV.2 we formulate our main result. In Section IV.3 we recall important results from Chapter III together with the filtering equations obtained therein. In Section IV.4 we prove L_p estimates needed for a priori bounds for the smoothed conditional distribution. In Section IV.5 we obtain an Itô formula for the L_p-norm of the smoothed conditional distribution and prove our result for the case $p = 2$. Section IV.6 contains existence and uniqueness results for the filtering equation in L_p-spaces. Finally, in Section IV.7 we prove our main theorem.

IV.2 Formulation of the main results

We fix nonnegative constants K_0, K_1, L, K and functions $\bar{\xi} \in L_2(\mathfrak{Z}_1) = L_2(\mathfrak{Z}_1, \mathcal{Z}_1, \nu_1)$, $\bar{\eta} \in L_2(\mathfrak{Z}_0) = L_2(\mathfrak{Z}_0, \mathcal{Z}_0, \nu_0)$, used throughout the paper, and make the following assumptions.

Assumption IV.2.1. (i) For $z_j = (x_j, y_j) \in \mathbb{R}^{d+d'}$ $(j = 1, 2)$, $t \geqslant 0$ and $\mathfrak{z}_i \in \mathfrak{Z}_i$ $(i = 0, 1)$,

$$|b(t, z_1) - b(t, z_2)| + |B(t, z_1) - B(t, z_2)| + |\sigma(t, z_1) - \sigma(t, z_2)|$$

$$+|\rho(t, z_1) - \rho(t, z_2)| \leqslant L|z_1 - z_2|,$$

$$|\eta(t, z_1, \mathfrak{z}_0) - \eta(t, z_2, \mathfrak{z}_0)| \leqslant \bar{\eta}(\mathfrak{z}_0)|z_1 - z_2|,$$

$$|\xi(t, z_1, \mathfrak{z}_1) - \xi(t, z_2, \mathfrak{z}_1)| \leqslant \bar{\xi}(\mathfrak{z}_1)|z_1 - z_2|.$$

(ii) For all $z = (x, y) \in \mathbb{R}^{d+d'}$, $t \geqslant 0$ and $\mathfrak{z}_i \in \mathfrak{Z}_i$ for $i = 0, 1$ we have

$$|b(t, z)| + |\sigma(t, z)| + |\rho(t, z)| \leqslant K_0 + K_1|z|, \quad |B(t, z)| \leqslant K,$$

$$|\eta(t, z, \mathfrak{z}_0)| \leqslant \bar{\eta}(\mathfrak{z}_0)(K_0 + K_1|z|), \quad |\xi(t, z, \mathfrak{z}_1)| \leqslant \bar{\xi}(\mathfrak{z}_1)(K_0 + K_1|z|),$$

$$\int_{\mathfrak{Z}_1} |\mathfrak{z}|^2 \, \nu_1(d\mathfrak{z}) \leqslant K_0^2.$$

(iii) The initial condition $Z_0 = (X_0, Y_0)$ is an $\mathcal{F}_0$-measurable random variable with values in $\mathbb{R}^{d+d'}$.

Assumption IV.2.2. The functions $\bar{\eta} \in L_2(\mathfrak{Z}_0)$ and $\bar{\xi} \in L_2(\mathfrak{Z}_1)$ satisfy $\bar{\eta}(\mathfrak{z}_0) \leqslant K_\eta$ and $\bar{\xi}(\mathfrak{z}_1) \leqslant K_\xi$ for all $\mathfrak{z}_i \in \mathfrak{Z}_i$, $i = 0, 1$ and some nonnegative constants K_η and K_ξ.

Assumption IV.2.3. For some $r > 2$ we have $\mathbb{E}|X_0|^r < \infty$, and the measure ν_1 satisfies

$$K_r := \int_{\mathfrak{Z}_1} |\mathfrak{z}|^r \, \nu_1(d\mathfrak{z}) < \infty.$$

By Theorem II.2.1 we know that Assumption IV.2.1 ensures the existence and uniqueness of a solution $(X_t, Y_t)_{t \geqslant 0}$ to (I.0.2) for any given $\mathcal{F}_0$-measurable initial value $Z_0 = (X_0, Y_0)$, and for every $T > 0$,

$$\mathbb{E} \sup_{t \leqslant T} (|X_t|^q + |Y_t|^q) \leqslant N(1 + \mathbb{E}|X_0|^q + \mathbb{E}|Y_0|^q) \tag{IV.2.1}$$

holds for $q = 2$ with a constant N depending only on T, K_0, K, K_1, L, $|\bar{\xi}|_{L_2}$, $|\bar{\eta}|_{L_2}$ and $d + d'$. If in addition to Assumption IV.2.1 we have that Assumptions IV.2.2 and IV.2.3 hold, then by Theorem II.2.2, see also [12], we know that the moment estimate (IV.2.1) holds with $q := r$ for every $T > 0$, where now the constant N depends also on r, K_r K_ξ and K_η.

As in the previous chapter, we also need the following additional assumption.

Assumption IV.2.4. (i) The functions $f_0(t, x, y, \mathfrak{z}_0) := \eta(t, x, y, \mathfrak{z}_0)$ and $f_1(t, x, y, \mathfrak{z}_1) := \xi(t, x, y, \mathfrak{z}_1)$ are continuously differentiable in $x \in \mathbb{R}^d$ for each $(t, y, \mathfrak{z}_i) \in \mathbb{R}_+ \times \mathbb{R}^{d'} \times \mathfrak{Z}_i$, for $i = 0$ and $i = 1$, respectively, such that

$$\lim_{\varepsilon \downarrow 0} \sup_{t \in [0,T]} \sup_{\mathfrak{z} \in \mathfrak{Z}_i} \sup_{|y| \leqslant R} \sup_{|x| \leqslant R, |x'| \leqslant R, |x - x'| \leqslant \varepsilon} |D_x f_i(t, x, y, \mathfrak{z}_i) - D_x f_i(t, x', y, \mathfrak{z}_i)| = 0$$

for every $R > 0$.

(ii) There is a constant $\lambda > 0$ such that for $\theta \in [0, 1]$, $(t, y, \mathfrak{z}_i) \in \mathbb{R}_+ \times \mathbb{R}^{d'} \times \mathfrak{Z}_i$ for $i = 0, 1$ we have

$$\lambda |x_1 - x_2| \leqslant |x_1 - x_2 + \theta(f_i(t, x_1, y, \mathfrak{z}_i) - f_i(t, x_2, y, \mathfrak{z}_i))| \quad \text{for } x_1, x_2 \in \mathbb{R}^d.$$

(iii) The function $\rho B = (\rho^{ik} B^k)$ is Lipschitz in $x \in \mathbb{R}^d$, uniformly in (t, y), i.e.,

$$|(\rho B)(t, x_1, y) - (\rho B)(t, x_2, y)| \leqslant L|x_1 - x_2|,$$

for all $x_1, x_2 \in \mathbb{R}^d$ and $(t, y) \in [0, T] \times \mathbb{R}^{d'}$.

Recall that $\mathcal{F}_t^Y$ denotes the completion of the σ-algebra generated by $(Y_s)_{s \leqslant t}$. Then the main result of the paper reads as follows.

Theorem IV.2.1. *Let Assumptions IV.2.1, IV.2.2 and IV.2.4 hold. If $K_1 \neq 0$ in Assumption IV.2.1, then let additionally Assumption IV.2.3 hold. Assume the conditional density $\pi_0 = P(X_0 \in dx | \mathcal{F}_0^Y)/dx$ exists and $\mathbb{E}|\pi_0|_{L_p}^p < \infty$ for some*

$p \geqslant 2$. Then for each $t \in [0,T]$ the conditional density $P(X_t \in dx|\mathcal{F}_t^Y)/dx$ exists almost surely. Moreover, there is an L_p-valued weakly cadlag process $(\pi_t)_{t\in[0,T]}$ such that for each $t \in [0,T]$ almost surely $\pi_t = P(X_t \in dx|\mathcal{F}_t^Y)/dx$.

IV.3 The filtering equations revisited

We briefly review the main results needed from Chapter III for the reader's convenience, as well as present some new notions and auxiliary facts needed in later sections of the present chapter.

We recall, for $t \in [0,T]$, the random differential operators

$$\mathcal{L}_t = a_t^{ij}(x)D_{ij} + b_t^i(x)D_i, \quad \mathcal{M}_t^k = \rho_t^{ik}(x)D_i + B_t^k(x), \quad k = 1,2,...,d',$$

where

$$a_t^{ij}(x) := \tfrac{1}{2}\sum_{k=1}^{d_1}(\sigma_t^{ik}\sigma_t^{jk})(x) + \tfrac{1}{2}\sum_{l=1}^{d'}(\rho_t^{il}\rho_t^{jl})(x), \quad \sigma_t^{ik}(x) := \sigma^{ik}(t,x,Y_t),$$

$$\rho_t^{il}(x) := \rho^{il}(t,x,Y_t), \quad b_t^i(x) := b^i(t,x,Y_t), \quad B_t^k(x) := B^k(t,x,Y_t)$$

for $\omega \in \Omega$, $t \geqslant 0$, $x = (x^1,...,x^d) \in \mathbb{R}^d$ and $i,j = 1,2...,d$, as well as for $\mathfrak{z} \in \mathfrak{Z}_1$ the random operators I_t^ξ and J_t^ξ defined by

$$I_t^\xi\varphi(x,\mathfrak{z}) = \varphi(x+\xi_t(x,\mathfrak{z}),\mathfrak{z})-\varphi(x,\mathfrak{z}), \quad J_t^\xi\phi(x,\mathfrak{z}) = I_t^\xi\phi(x,\mathfrak{z})-\sum_{i=1}^d \xi_t^i(x,\mathfrak{z})D_i\phi(x,\mathfrak{z}) \tag{IV.3.1}$$

for functions $\varphi = \varphi(x,\mathfrak{z})$ and $\phi = \phi(x,\mathfrak{z})$ of $x \in \mathbb{R}^d$ and $\mathfrak{z} \in \mathfrak{Z}_1$, and furthermore the random operators I_t^η and J_t^η, defined as I_t^ξ and J_t^ξ, respectively, with $\eta_t(x,\mathfrak{z})$ in place of $\xi_t(x,\mathfrak{z})$, where

$$\xi_t(x,\mathfrak{z}_1) := \xi(t,x,Y_{t-},\mathfrak{z}_1), \quad \eta_t(x,\mathfrak{z}_0) := \eta(t,x,Y_{t-},\mathfrak{z}_0)$$

for $\omega \in \Omega$, $t \geqslant 0$, $x \in \mathbb{R}^d$ and $\mathfrak{z}_i \in \mathfrak{Z}_i$ for $i = 0,1$. We recall also the processes

$$\gamma_t = \exp\left(-\int_0^t B_s(X_s)\,dV_s - \tfrac{1}{2}\int_0^t |B_s(X_s)|^2\,ds\right), \quad t \in [0,T],$$

$$\tilde{V}_t = \int_0^t B_s(X_s)\,ds + V_t, \quad t \in [0,T]. \tag{IV.3.2}$$

Since by Assumption IV.2.1 (ii) B is bounded in magnitude by a constant, we know that Assumption III.1.2 holds and hence, $(\gamma_t)_{t\in[0,T]}$ is an $\mathcal{F}_t$-martingale such that, by Girsanov's theorem, the measure Q defined by $dQ = \gamma_T dP$ is a probability measure equivalent to P and under Q the process $(W_t,\tilde{V}_t)_{t\in[0,T]}$ is a $d_1 + d'$-dimensional $\mathcal{F}_t$-Wiener process.

By Theorem III.1.1 we know that if $Z = (X_t, Y_t)_{t\in[0,T]}$ satisfies equation (I.0.2), Assumption IV.2.1(ii) holds and that if $\mathbb{E}|X_0|^2 < \infty$ so long as $K_1 \neq 0$ in Assumption IV.2.1(ii), then there exist measure-valued $\mathcal{F}_t^Y$-adapted weakly cadlag processes $(P_t)_{t\in[0,T]}$ and $(\mu_t)_{t\in[0,T]}$ such that almost surely

$$P_t(\varphi) = \mu_t(\varphi)/\mu_t(\mathbf{1}), \quad \text{for all } t \in [0,T],$$

$$P_t(\varphi) = \mathbb{E}(\varphi(X_t)|\mathcal{F}_t^Y), \quad \mu_t(\varphi) = \mathbb{E}_Q(\gamma_t^{-1}\varphi(X_t)|\mathcal{F}_t^Y) \quad \text{(a.s.) for each } t \in [0,T],$$

for bounded Borel functions φ on $\mathbb{R}^d$, and for every $\varphi \in C_b^2(\mathbb{R}^d)$ almost surely

$$\begin{aligned}\mu_t(\varphi) =& \mu_0(\varphi) + \int_0^t \mu_s(\mathcal{L}_s\varphi)\,ds + \int_0^t \mu_s(\mathcal{M}_s^k\varphi)\,d\tilde{V}_s^k + \int_0^t\int_{\mathfrak{Z}_0} \mu_s(J_s^\eta\varphi)\,\nu_0(d\mathfrak{z})ds \\ &+ \int_0^t\int_{\mathfrak{Z}_1} \mu_s(J_s^\xi\varphi)\,\nu_1(d\mathfrak{z})ds + \int_0^t\int_{\mathfrak{Z}_1} \mu_{s-}(I_s^\xi\varphi)\,\tilde{N}_1(d\mathfrak{z},ds),\end{aligned} \tag{IV.3.3}$$

for all $t \in [0,T]$. Clearly, equation (IV.3.3) can be rewritten as

$$\begin{aligned}\mu_t(\varphi) =& \mu_0(\varphi) + \int_0^t \mu_s(\tilde{\mathcal{L}}_s\varphi)\,ds + \int_0^t \mu_s(\mathcal{M}_s^k\varphi)\,dV_s^k + \int_0^t\int_{\mathfrak{Z}_0} \mu_s(J_s^\eta\varphi)\,\nu_0(d\mathfrak{z})ds \\ &+ \int_0^t\int_{\mathfrak{Z}_1} \mu_s(J_s^\xi\varphi)\,\nu_1(d\mathfrak{z})ds + \int_0^t\int_{\mathfrak{Z}_1} \mu_{s-}(I_s^\xi\varphi)\,\tilde{N}_1(d\mathfrak{z},ds),\end{aligned} \tag{IV.3.4}$$

where $\tilde{\mathcal{L}}_s = \mathcal{L}_s + B_s(X_s)\mathcal{M}_s$. Moreover, if $d\mu_t/dx$ exists for $P \otimes dt$-a.e. $(\omega, t) \in \Omega \times [0,T]$, and $u = u_t(x)$ is an $\mathcal{F}_t$-adapted L_p-valued weakly cadlag process, for $p > 1$, such that almost surely $u_t = d\mu_t/dx$ for all $t \in [0,T]$, then for each $\varphi \in C_b^2(\mathbb{R}^d)$ we have that almost surely

$$\begin{aligned}(u_t,\varphi) =& (u_0,\varphi) + \int_0^t (u_s,\tilde{\mathcal{L}}_s\varphi)\,ds + \int_0^t (u_s,\mathcal{M}_s^k\varphi)\,dV_s^k + \int_0^t\int_{\mathfrak{Z}_0} (u_s,J_s^\eta\varphi)\,\nu_0(d\mathfrak{z})ds \\ &+ \int_0^t\int_{\mathfrak{Z}_1} (u_s,J_s^\xi\varphi)\,\nu_1(d\mathfrak{z})ds + \int_0^t\int_{\mathfrak{Z}_1} (u_{s-},I_s^\xi\varphi)\,\tilde{N}_1(d\mathfrak{z},ds).\end{aligned} \tag{IV.3.5}$$

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

Finally we recall from Chapter III that there exists a cadlag $\mathcal{F}_t^Y$-adapted positive process $({}^o\gamma_t)_{t\in[0,T]}$, the optional projection of $(\gamma_t)_{t\in[0,T]}$ under P with respect to $(\mathcal{F}_t^Y)_{t\in[0,T]}$, such that for every $\mathcal{F}_t^Y$-stopping time $\tau \leqslant T$ we have

$$\mathbb{E}(\gamma_\tau|\mathcal{F}_\tau^Y) = {}^o\gamma_\tau, \quad \text{almost surely.} \tag{IV.3.6}$$

Since for each $t \in [0,T]$, by known properties of conditional expectations, almost surely

$$\mu_t(\mathbf{1}) = \mathbb{E}_Q(\gamma_t^{-1}|\mathcal{F}_t^Y) = 1/\mathbb{E}(\gamma_t|\mathcal{F}_t^Y) = 1/{}^o\gamma_t,$$

we also have that for each $\varphi \in C_b^2$ almost surely $P_t(\varphi) = \mu_t(\varphi)^o\gamma_t$ for all $t \in [0, T]$.

Definition IV.3.1. An $\mathfrak{M}$-valued weakly cadlag $\mathcal{F}_t$-adapted process $(\mu_t)_{t\in[0,T]}$ is said to be an $\mathfrak{M}$-solution to the equation

$$\begin{aligned} d\mu_t =& \tilde{\mathcal{L}}_t^*\mu_t dt + \mathcal{M}_t^{k*}\mu_t\, dV_t^k + \int_{\mathfrak{Z}_0} J_t^{\eta*}\mu_t\, \nu_0(d\mathfrak{z})dt \\ &+ \int_{\mathfrak{Z}_1} J_t^{\xi*}\mu_t\, \nu_1(d\mathfrak{z})dt + \int_{\mathfrak{Z}_1} I_t^{\xi*}\mu_{t-}\, \tilde{N}_1(d\mathfrak{z}, dt) \end{aligned} \tag{IV.3.7}$$

with initial value μ_0, if for each $\varphi \in C_b^2$ almost surely equation (IV.3.4) holds for all $t \in [0, T]$. If $(\mu_t)_{t\in[0,T]}$ is an $\mathfrak{M}$-solution to equation (IV.3.7), such that it takes values in $\mathbb{M}$, then we call it a measure-valued solution.

Definition IV.3.2. Let $p \geqslant 1$ and let ψ be an L_p-valued $\mathcal{F}_0$-measurable random variable. Then we say that an L_p-valued $\mathcal{F}_t$-adapted weakly cadlag process $(u_t)_{t\in[0,T]}$ is an L_p-solution of the equation

$$\begin{aligned} du_t =& \tilde{\mathcal{L}}_t^*u_t dt + \mathcal{M}_t^{k*}u_t\, dV_t^k + \int_{\mathfrak{Z}_0} J_t^{\eta*}u_t\, \nu_0(d\mathfrak{z})dt \\ &+ \int_{\mathfrak{Z}_1} J_t^{\xi*}u_t\, \nu_1(d\mathfrak{z})dt + \int_{\mathfrak{Z}_1} I_t^{\xi*}u_{t-}\, \tilde{N}_1(d\mathfrak{z}, dt) \end{aligned} \tag{IV.3.8}$$

with initial condition ψ, if for every $\varphi \in C_0^\infty$ almost surely (IV.3.5) holds for all $t \in [0, T]$ and $u_0 = \psi$ (a.s.).

Lemma IV.3.1. *Let Assumption IV.2.1 hold, and assume also $\mathbb{E}|X_0|^2 < \infty$ if $K_1 \neq 0$ in Assumptions IV.2.1(ii). Let $(\mu_t)_{t\in[0,T]}$ be the measure-valued process from Theorem III.1.1. Then we have*

$$\mathbb{E} \sup_{t\in[0,T]} \mu_t(\mathbf{1}) \leqslant N, \tag{IV.3.9}$$

with a constant N depending only on d, K and T.

Proof. Taking $\mathbf{1}$ instead of φ in the Zakai equation (IV.3.3) yields for $(\beta_t)_{t\in[0,T]} := (B_t(X_t))_{t\in[0,T]}$,

$$\mu_t(\mathbf{1}) = 1 + \int_0^t \mu_s(\beta_s^k B_s^k)\, ds + \int_0^t \mu_s(B_s^k)\, dV_s^k. \tag{IV.3.10}$$

Since $\int_0^t \mu_s(B_s^k)\, dV_s^k$ is a martingale we can define

$$\tau_n := \inf\{t \geqslant 0 : \int_0^t \mu_s(\mathbf{1})\, ds \geqslant n\}, \quad n \geqslant 1,$$

and use $|B| \leqslant K$ to compute after taking expectations on both sides,

$$\mathbb{E}\mu_{t\wedge\tau_n}(\mathbf{1}) \leqslant 1 + dK^2 \int_0^t \mathbb{E}\mu_{s\wedge\tau_n}(\mathbf{1})\, ds.$$

Using Gronwall's inequality and Fatou's lemma, we obtain for each n,

$$\sup_{t\in[0,T]} \mathbb{E}\mu_t(\mathbf{1}) \leqslant N, \tag{IV.3.11}$$

with a constant $N = N(d, K, T)$. Moreover, due to (IV.3.10), the process $\mu_t(\mathbf{1})$ is continuous almost surely, wherefore $c_t^* := (\sup_{s\leqslant t} \mu_s(\mathbf{1}))_{t\in[0,T]}$ is locally integrable and there exists a sequence of stopping times $\rho_m \uparrow \infty$ such that $\mathbb{E}c^*_{t\wedge\rho_m} < \infty$ for all $m \geqslant 1$. Hence, by Davis' and Young's inequalities as well as (IV.3.11),

$$\mathbb{E}\sup_{t\in[0,T]} \int_0^{t\wedge\rho_m} \mu_s(B_s^k)\, dV_s^k \leqslant 3\mathbb{E}\Big(\sum_k \int_0^{T\wedge\rho_m} \mu_s(B_s^k)^2\, ds\Big)^{1/2}$$

$$\leqslant N\mathbb{E}\Big(\sup_{t\in[0,T]} \mu_{t\wedge\rho_m}(\mathbf{1}) \int_0^T \mu_s(\mathbf{1})\, ds\Big)^{1/2} \leqslant \delta\mathbb{E}\sup_{t\in[0,T]} \mu_{t\wedge\rho_m}(\mathbf{1}) + N',$$

for all $\delta > 0$, $n \geqslant 1$ and constants $N = N(d, K)$ and $N' = N'(\delta, d, K, T)$. Thus also

$$\mathbb{E}\sup_{t\in[0,T]} \mu_{t\wedge\rho_m}(\mathbf{1}) \leqslant N, \quad \text{for all } m,$$

for another constant $N = N(\delta, d, K, T)$. By Fatou's lemma we then obtain (IV.3.9) □

IV.4 L_p-estimates

Recall that $\mathbb{M} = \mathbb{M}(\mathbb{R}^d)$ denotes the set of finite measures on $\mathcal{B}(\mathbb{R}^d)$, and $\mathfrak{M} := \{\mu - \nu : \mu, \nu \in \mathbb{M}\}$. For $\nu \in \mathfrak{M}$ we use the notation $|\nu| := \nu^+ + \nu^-$ for the total variation and set $\|\nu\| = |\nu|(\mathbb{R}^d)$, where $\nu^+ \in \mathbb{M}$ and $\nu^- \in \mathbb{M}$ are the positive and negative parts of ν. For $\varepsilon > 0$ we use the notation k_ε for the Gaussian density function on $\mathbb{R}^d$ with mean 0 and covariance matrix $\varepsilon\mathbb{I}$. For linear functionals Φ, acting on a real vector space V containing $\mathcal{S} = \mathcal{S}(\mathbb{R}^d)$, the rapidly decreasing smooth functions on $\mathbb{R}^d$, the mollification $\Phi^{(\varepsilon)}$ is defined by

$$\Phi^{(\varepsilon)}(x) = \Phi(k_\varepsilon(x - \cdot)), \quad x \in \mathbb{R}^d.$$

In particular, when $\Phi = \mu$ is a (signed) measure from $\mathcal{S}^*$, the dual of $\mathcal{S}$, or $\Phi = f$ is a function from $\mathcal{S}^*$, then

$$\mu^{(\varepsilon)}(x) = \int_{\mathbb{R}^d} k_\varepsilon(x - y)\, \mu(dy), \quad f^{(\varepsilon)}(x) = \int_{\mathbb{R}^d} k_\varepsilon(x - y) f(y)\, dy, \quad x \in \mathbb{R}^d,$$

and

$$(L^*\mu)^{(\varepsilon)}(x) := \int_{\mathbb{R}^d} L_y k_\varepsilon(x-y)\mu(dy), \quad x \in \mathbb{R}^d,$$

$$(L^*f)^{(\varepsilon)}(x) := \int_{\mathbb{R}^d} (L_y k_\varepsilon(x-y)) f(y)\,dy, \quad x \in \mathbb{R}^d,$$

when L is a linear operator on V such that the integrals are well-defined for every $x \in \mathbb{R}^d$. The subscript y in L_y indicates that the operator L acts in the y-variable of the function $\bar{k}_\varepsilon(x,y) := k_\varepsilon(x-y)$. For example, if L is a differential operator of the form $a^{ij}D_{ij} + b^i D_i + c$, where a^{ij}, b^i and c are functions defined on $\mathbb{R}^d$, then

$$(L^*\mu)^{(\varepsilon)}(x) = \int_{\mathbb{R}^d} (a^{ij}(y)\tfrac{\partial^2}{\partial y^i \partial y^j} + b^i(y)\tfrac{\partial}{\partial y^i} + c(y))k_\varepsilon(x-y)\mu(dy).$$

We will often use the following well-known properties of mollifications with k_ε:

(i) $|\varphi^{(\varepsilon)}|_{L_p} \leqslant |\varphi|_{L_p}$ for $\varphi \in L_p(\mathbb{R}^d)$, $p \in [1,\infty)$;

(ii) $\mu^{(\varepsilon)}(\varphi) := \int_{\mathbb{R}^d} \mu^{(\varepsilon)}(x)\varphi(x)\,dx = \int_{\mathbb{R}^d} \varphi^{(\varepsilon)}(x)\mu(dx) =: \mu(\varphi^{(\varepsilon)})$ for $\mu \in \mathfrak{M}$ and $\varphi \in L_p(\mathbb{R}^d)$, $p \geqslant 1$;

(iii) $|\mu^{(\delta)}|_{L_p} \leqslant |\mu^{(\varepsilon)}|_{L_p}$ for $0 \leqslant \varepsilon \leqslant \delta$, $\mu \in \mathfrak{M}$ and $p \geqslant 1$. This property follows immediately from (i) and the "semigroup property" of the Gaussian kernel,

$$k_{r+s}(y-z) = \int_{\mathbb{R}^d} k_r(y-x)k_s(x-z)\,dx, \quad y,z \in \mathbb{R}^d \text{ and } r,s \in (0,\infty). \tag{IV.4.1}$$

The following generalization of (iii) is also useful: for integers $p \geqslant 2$ we have

$$\rho_\varepsilon(y) := \int_{\mathbb{R}^d} \Pi_{r=1}^p k_\varepsilon(x-y_r)\,dx = c_{p,\varepsilon} e^{-\sum_{1\leqslant r<s\leqslant p}|y_r-y_s|^2/(2\varepsilon p)}, \quad y = (y_1, ..., y_p) \in \mathbb{R}^{pd}, \tag{IV.4.2}$$

for $\varepsilon > 0$, with a constant $c_{p,\varepsilon} = c_{p,\varepsilon}(d) = p^{-d/2}(2\pi\varepsilon)^{(1-p)d/2}$. This can be seen immediately by noticing that for $x, y_k \in \mathbb{R}^d$ and $y = (y_k)_{k=1}^p \in \mathbb{R}^{pd}$ we have

$$\sum_{k=1}^p (x-y_k)^2 = p\Big(x - \sum_k y_k/p\Big)^2 + \tfrac{1}{p}\sum_{1\leqslant k<l\leqslant p}(y_k-y_l)^2.$$

Clearly, for every $r = 1,2,...,p$ and $i = 1,2,...,d$,

$$\partial_{y_r^i}\rho_\varepsilon(y) = \tfrac{1}{\varepsilon p}\sum_{s=1}^p (y_s^i - y_r^i)\rho_\varepsilon(y), \quad y = (y_1,...,y_p) \in \mathbb{R}^d, \quad y_r = (y_r^1..., y_r^d) \in \mathbb{R}^d. \tag{IV.4.3}$$

It is easy to see that

$$\sum_{r=1}^{p} \partial_{y_r^j} \rho_\varepsilon(y) = 0 \quad \text{for } y \in \mathbb{R}^{pd},\ j = 1, 2, ..., d.$$

We will often use this in the form

$$\partial_{y_r^j} \rho_\varepsilon(y) = -\sum_{s \neq r}^{p} \partial_{y_s^j} \rho_\varepsilon(y) \quad \text{for } r = 1, ..., p \text{ and } j = 1, 2, ..., d. \tag{IV.4.4}$$

In order for the left-hand side of the following L_p-estimates for $\mu \in \mathfrak{M}$ in this section to be well-defined, we require that

$$K_1 \int_{\mathbb{R}^d} |x|^2 \, |\mu|(dx) < \infty, \tag{IV.4.5}$$

where we use the formal convention that $0 \cdot \infty = 0$, i.e., if $K_1 = 0$, then condition (IV.4.5) is satisfied. The following lemma generalises a lemma from [41].

Lemma IV.4.1. *Let $p \geqslant 2$ be an integer. Let $\sigma = (\sigma^{ik})$ and $b = (b^i)$ be Borel functions on $\mathbb{R}^d$ with values in $\mathbb{R}^{d \times m}$ and $\mathbb{R}^d$, respectively, such that for some nonnegative constants K_0, K_1 and L we have*

$$|\sigma(x)| + |b(x)| \leqslant K_0 + K_1|x| \quad |\sigma(x) - \sigma(y)| \leqslant L|x - y|, \quad |b(x) - b(y)| \leqslant L|x - y| \tag{IV.4.6}$$

for all $x, y \in \mathbb{R}^d$. Set $a^{ij} = \sigma^{ik}\sigma^{jk}/2$ for $i, j = 1, 2, ..., d$. Let $\mu \in \mathfrak{M}$ such that it satisfies (IV.4.5)*. Then we have*

$$p((\mu^{(\varepsilon)})^{p-1}, ((a^{ij}D_{ij})^*\mu)^{(\varepsilon)}) + \tfrac{p(p-1)}{2}((\mu^{(\varepsilon)})^{p-2}((\sigma^{ik}D_i)^*\mu)^{(\varepsilon)}, ((\sigma^{jk}D_j)^*\mu)^{(\varepsilon)})$$
$$\leqslant NL^2 || \mu|^{(\varepsilon)}|_{L_p}^p, \tag{IV.4.7}$$

$$((\mu^{(\varepsilon)})^{p-1}, ((b^i D_i)^*\mu)^{(\varepsilon)}) \leqslant NL^2 || \mu|^{(\varepsilon)}|_{L_p}^p \tag{IV.4.8}$$

with a constant $N = N(d, p)$.

Proof. Let A and B denote the left-hand side of the inequalities (IV.4.7) and (IV.4.8), respectively. Note first that using

$$\sup_{x \in \mathbb{R}^d} \sum_{k=0}^{2} |D^k k_\varepsilon(x)| < \infty, \quad \text{and} \quad \int_{\mathbb{R}^d} (1 + |x| + K_1|x|^2) \, |\mu|(dx) < \infty, \tag{IV.4.9}$$

as well as the conditions on on σ and b, it is easy to verify that A and B are well-defined. Then by Fubini's theorem and the symmetry of the Gaussian kernel

$$A = \int_{\mathbb{R}^{(p+1)d}} f(x, y) \, \mu_p(dy) \, dx$$

with

$$f(x,y) := \Big(pa^{ij}(y_p)\partial_{y_p^i}\partial_{y_p^j} + \tfrac{p(p-1)}{2}\sigma^{ik}(y_{p-1})\sigma^{jk}(y_p)\partial_{y_{p-1}^i}\partial_{y_p^j}\Big)\,\Pi_{k=1}^p k_\varepsilon(x-y_k),$$

where $x \in \mathbb{R}^d$, $y = (y_k)_{k=1}^p \in \mathbb{R}^{dp}$, and y_k^i denotes the i-th coordinate of $y_k \in \mathbb{R}^d$ for $k = 1,...,p$, and $\mu_p(dy) := \mu^{\otimes p}(dy) = \mu(dy_1)...\mu(dy_p)$. Hence by Fubini's theorem and symmetry again

$$A = \int_{\mathbb{R}^{pd}} \Big(pa^{ij}(y_p)\partial_{y_p^i}\partial_{y_p^j} + \tfrac{p(p-1)}{2}\sigma^{ik}(y_{p-1})\sigma^{jk}(y_p)\partial_{y_{p-1}^i}\partial_{y_p^j}\Big)\,\rho_\varepsilon(y)\,\mu_p(dy) \tag{IV.4.10}$$

$$= \tfrac{1}{2}\sum_{r=1}^p \int_{\mathbb{R}^{pd}} \Big(2a^{ij}(y_r)\partial_{y_r^i}\partial_{y_r^j} + \sum_{s\neq r}\sigma^{ik}(y_r)\sigma^{jk}(y_s)\partial_{y_r^i}\partial_{y_s^j}\Big)\rho_\varepsilon(y)\,\mu_p(dy), \tag{IV.4.11}$$

where ρ_ε is given in (IV.4.2). Using here (IV.4.4) and symmetry of expressions in y_k and y_l, we obtain

$$A = -\tfrac{1}{2}\sum_{r=1}^p\sum_{s\neq r}\int_{\mathbb{R}^{pd}} \Big(2a^{ij}(y_r)\partial_{y_r^i}\partial_{y_s^j} - \sigma^{ik}(y_r)\sigma^{jk}(y_s)\partial_{y_r^i}\partial_{y_s^j}\Big)\,\rho_\varepsilon(y)\,\mu_p(dy)$$

$$= -\tfrac{1}{2}\sum_{r=1}^p\sum_{s\neq r}\int_{\mathbb{R}^{pd}} \Big((a^{ij}(y_r)+a^{ij}(y_s))\partial_{y_r^i}\partial_{y_s^j} - \sigma^{ik}(y_r)\sigma^{jk}(y_s)\partial_{y_r^i}\partial_{y_s^j}\Big)\,\rho_\varepsilon(y)\,\mu_p(dy)$$

$$= -\tfrac{1}{2}\sum_{r=1}^p\sum_{s=1}^p\int_{\mathbb{R}^{pd}} a^{ij}(y_r,y_s)\partial_{y_r^i}\partial_{y_s^j}\rho_\varepsilon(y)\,\mu_p(dy)$$

$$= -\tfrac{1}{2}\sum_{r=1}^p\sum_{s=1}^p\int_{\mathbb{R}^{pd}} a^{ij}(y_r,y_s)l_\varepsilon^{ij,rs}(y)\rho_\varepsilon(y)\,\mu_p(dy),$$

where

$$a^{ij}(x,z) = \frac{1}{2}(\sigma^{ik}(x)-\sigma^{ik}(z))(\sigma^{jk}(x)-\sigma^{jk}(z)) \quad \text{for } x,z \in \mathbb{R}^d \tag{IV.4.12}$$

and

$$l_\varepsilon^{ij,rs}(y) = \rho_\varepsilon^{-1}(y)\partial_{y_r^i}\partial_{y_s^j}\rho_\varepsilon(y) = \tfrac{1}{(p\varepsilon)^2}\sum_{k=1}^p\sum_{l=1}^p (y_k^i - y_r^i)(y_l^j - y_s^j) + \tfrac{\delta_{ij}}{p\varepsilon}.$$

Making use of the Lipschitz condition on σ and using for $q = 1,2$ that

$$\varepsilon^{-q}\sum_{s\neq r}|y_s - y_r|^{2q}\rho_\varepsilon(y) \leqslant N\rho_{2\varepsilon}(y), \quad y \in \mathbb{R}^{pd}, \quad q \geqslant 0 \tag{IV.4.13}$$

with a constant $N = N(d,p,q)$, we have

$$|\sum_{r=1}^{p}\sum_{s=1}^{p} a^{ij}(y_r,y_s) l_\varepsilon^{ij,rs}(y)| \leqslant \tfrac{NL^2}{(p\varepsilon)^2} \sum_{1\leqslant r<s\leqslant p} |y_r - y_s|^4 \rho_\varepsilon(y) + \tfrac{NL^2}{p\varepsilon} \sum_{1\leqslant r<s\leqslant p} |y_r - y_s|^2 \rho_\varepsilon(y)$$

$$\leqslant N' L^2 \rho_{2\varepsilon}(y) \quad \text{for } y \in \mathbb{R}^{pd}$$

with constants $N = N(d,p)$ and $N' = N'(d,p)$. Hence

$$A \leqslant N'L^2 \int_{\mathbb{R}^{pd}} \rho_{2\varepsilon}(y)|\mu_p|(dy) = N'L^2 \int_{\mathbb{R}^{pd}} \int_{\mathbb{R}^d} \Pi_{r=1}^p k_{2\varepsilon}(x - y_r)\, dx |\mu_p|(dy)$$

$$= N'L^2 \int_{\mathbb{R}^d} \Pi_{r=1}^p \int_{\mathbb{R}^d} k_{2\varepsilon}(x - y_r)|\mu|(dy_r)\, dx = N'L^2 || \mu|^{(2\varepsilon)}|_{L_p}^p.$$

To prove (IV.4.8) we proceed similarly. By Fubini's theorem and symmetry

$$pB = \int_{\mathbb{R}^{pd}} p b^i(y_p) \partial_{y_p^i} \rho_\varepsilon(y)\, \mu_p(dy) = \sum_{r=1}^{p} \int_{\mathbb{R}^{pd}} b^i(y_r) \partial_{y_r^i} \rho_\varepsilon(y)\, \mu_p(dy)$$

$$= -\sum_{r=1}^{p}\sum_{s\neq r} \int_{\mathbb{R}^{pd}} b^i(y_r) \partial_{y_s^i} \rho_\varepsilon(y)\, \mu_p(dy) = -\sum_{r=1}^{p}\sum_{s\neq r} \int_{\mathbb{R}^{pd}} b^i(y_s) \partial_{y_r^i} \rho_\varepsilon(y)\, \mu_p(dy).$$

Thus

$$B = \tfrac{p-1}{p} \sum_{r=1}^{p} \int_{\mathbb{R}^{pd}} b^i(y_r) \partial_{y_r^i} \rho_\varepsilon(y)\, \mu_p(dy) - \tfrac{1}{p} \sum_{r=1}^{p}\sum_{s\neq r} \int_{\mathbb{R}^{pd}} b^i(y_s) \partial_{y_r^i} \rho_\varepsilon(y)\, \mu_p(dy)$$

$$= \tfrac{1}{p} \sum_{r=1}^{p} \int_{\mathbb{R}^{pd}} \sum_{s\neq r} (b^i(y_r) - b^i(y_s)) \partial_{y_r^i} \rho_\varepsilon(y)\, \mu_p(dy)$$

$$= \tfrac{1}{\varepsilon p^2} \sum_{r=1}^{p} \int_{\mathbb{R}^{pd}} \sum_{s\neq r} (b^i(y_r) - b^i(y_s)) \sum_{l\neq r} (y_l^i - y_r^i) \rho_\varepsilon(y)\, \mu_p(dy). \tag{IV.4.14}$$

Using the Lipschitz condition on b and the inequality (IV.4.13), we obtain

$$B \leqslant \tfrac{NL}{\varepsilon} \int_{\mathbb{R}^{pd}} \sum_{s\neq r} |y_r - y_s|^2 \rho_\varepsilon(y) |\mu_p|(dy) \leqslant N'L \int_{\mathbb{R}^{pd}} \rho_{2\varepsilon}(y)|\mu_p|(dy) = N'L || \mu|^{(2\varepsilon)}|_{L_p}^p$$

with constants $N = N(p,d)$ and $N' = N(p,d)$, which completes the proof of the lemma. □

Corollary IV.4.2. *Let the conditions of Lemma IV.4.1 hold for some even $p \geqslant 2$. Then we have*

$$((\mu^{(\varepsilon)})^{p-1}, ((a^{ij}D_{ij})^*\mu)^{(\varepsilon)}) \leqslant NL^2 || \mu|^{(\varepsilon)}|_{L_p}^p - \tfrac{p-1}{2} ((\mu^{(\varepsilon)})^{p-2} ((\sigma^{ik}D_i)^*\mu)^{(\varepsilon)}, ((\sigma^{jk}D_j)^*\mu)^{(\varepsilon)})$$

$$\leqslant NL^2 ||\mu|^{(\varepsilon)}|^p_{L_p}$$

for a constant $N = N(d,p)$.

Proof. It suffices to observe that always

$$((\mu^{(\varepsilon)})^{p-2}((\sigma^{ik}D_i)^*\mu)^{(\varepsilon)}, ((\sigma^{jk}D_j)^*\mu)^{(\varepsilon)}) = \int_{\mathbb{R}^d} (\mu^{(\varepsilon)})^{p-2} \sum_{k=1}^{d} \big(((\sigma^{ik}D_i)^*\mu)^{(\varepsilon)}\big)^2 dx \geqslant 0.$$

□

Lemma IV.4.3. *Let* $p \geqslant 2$ *be an integer and let* $\sigma = (\sigma^i)$ *and* b *be Borel functions on* $\mathbb{R}^d$ *with values in* $\mathbb{R}^d$ *and* $\mathbb{R}$ *respectively. Assume furthermore that there exist constants* $K \geqslant 1$, K_0 *and* L *such that*

$$|b(x)| \leqslant K, \quad |\sigma(x)| \leqslant K_0 + K_1|x|, \quad |\sigma(x) - \sigma(y)| + |b\sigma(x) - b\sigma(y)| \leqslant L|x-y|$$

for all $x, y \in \mathbb{R}^d$. *Let* $\mu \in \mathfrak{M}$ *such that it satisfies* (IV.4.5). *Then we have*

$$\big((\mu^{(\varepsilon)})^{p-2}(b\mu)^{(\varepsilon)}, (b\mu)^{(\varepsilon)}\big) \leqslant K^2 ||\mu|^{(\varepsilon)}|^p_{L_p}, \tag{IV.4.15}$$

$$\big((\mu^{(\varepsilon)})^{p-2}, ((\sigma^i D_i)^*\mu)^{(\varepsilon)}(b\mu)^{(\varepsilon)}\big) \leqslant NKL ||\mu|^{(\varepsilon)}|^p_{L_p} \tag{IV.4.16}$$

for every $\varepsilon > 0$ *with a constant* $N = N(d,p)$.

Proof. We note again that by (IV.4.9) together with the conditions on σ and b, the left-hand sides of (IV.4.15) and (IV.4.16) are well-defined. Rewriting products of integrals as multiple integrals and using Fubini's theorem for the left-hand side of the inequality (IV.4.15) we have

$$\int_{\mathbb{R}^{dp}} b(y_r)b(y_s) \int_{\mathbb{R}^d} \Pi_{j=1}^p k_\varepsilon(x - y_j)\, dx \mu_p(dy)$$

$$\leqslant K^2 \int_{\mathbb{R}^{d(p+1)}} \Pi_{k=1}^p k_\varepsilon(x - y_k)\, dx |\mu_p|(dy) = K^2 ||\mu|^{(\varepsilon)}|^p_{L_p},$$

for any $r, s \in \{1, 2, ..., p\}$, where $y = (y_1, ..., y_p) \in \mathbb{R}^{pd}$, $y_j \in \mathbb{R}^d$ for $j = 1, 2, ..., p$, and the notation $\mu_p(dy) = \mu(dy_1)...\mu(dy_p)$ is used. This proves (IV.4.15).

Rewriting products of integrals as multiple integrals, using Fubini's theorem, interchanging the order of taking derivatives and integrals, and using equation (IV.4.2), for the left-hand side R of the inequality (IV.4.16) we have

$$R = \int_{\mathbb{R}^{dp}} b(y_k)\sigma^i(y_r)\partial_{y_r^i} \int_{\mathbb{R}^d} \Pi_{j=1}^p k_\varepsilon(x - y_j)\, dx \mu_p(dy)$$

$$= \int_{\mathbb{R}^{dp}} b(y_k)\sigma^i(y_r)\partial_{y_r^i}\rho_\varepsilon(y)\, \mu_p(dy) \tag{IV.4.17}$$

for any $r, k \in \{1, 2, .., p\}$ such that $r \neq k$. Hence

$$p(p-1)^2 R = \sum_{s=1}^{p} \sum_{r \neq s} \sum_{k \neq s} \int_{\mathbb{R}^{dp}} b(y_k)\sigma^i(y_s)\partial_{y_s^i}\rho_\varepsilon(y)\,\mu_p(dy)$$

$$= \sum_{s=1}^{p} \sum_{r \neq s} \sum_{k \neq r} \int_{\mathbb{R}^{dp}} b(y_k)\sigma^i(y_s)\partial_{y_s^i}\rho_\varepsilon(y)\,\mu_p(dy)$$

$$+ \sum_{s=1}^{p} \sum_{r \neq s} \int_{\mathbb{R}^{dp}} (b(y_r) - b(y_s))\sigma^i(y_s)\partial_{y_s^i}\rho_\varepsilon(y)\,\mu_p(dy), \tag{IV.4.18}$$

and using (IV.4.4) from (IV.4.17) we obtain

$$p(p-1)R = -\sum_{r=1}^{p} \sum_{k \neq r} \sum_{s \neq r} \int_{\mathbb{R}^{dp}} b(y_k)\sigma^i(y_r)\partial_{y_s^i}\rho_\varepsilon(y)\,\mu_p(dy)$$

$$= -\sum_{s=1}^{p} \sum_{r \neq s} \sum_{k \neq r} \int_{\mathbb{R}^{dp}} b(y_k)\sigma^i(y_r)\partial_{y_s^i}\rho_\varepsilon(y)\,\mu_p(dy) \tag{IV.4.19}$$

Adding up equations (IV.4.18) and (IV.4.19), and taking into account the equation

$$(b(y_r) - b(y_s))\sigma^i(y_s) = b(y_r)\sigma^i(y_r) - b(y_s)\sigma^i(y_s) - b(y_r)(\sigma^i(y_r) - \sigma^i(y_s))$$

we get

$$p^2(p-1)R = \sum_{s=1}^{p} \sum_{r \neq s} \sum_{k=1}^{p} \int_{\mathbb{R}^{dp}} f^i(y_k, y_s, y_r)\partial_{y_s^i}\rho_\varepsilon(y)\,\mu_p(dy)$$

$$+ \sum_{s=1}^{p} \sum_{r \neq s} \int_{\mathbb{R}^{dp}} g^i(y_r, y_s)\partial_{y_s^i}\rho_\varepsilon(y)\,\mu_p(dy) \tag{IV.4.20}$$

with functions

$$f^i(x, u, v) := b(x)(\sigma^i(u) - \sigma^i(v)), \quad g^i(u, v) := b(u)\sigma^i(u) - b(v)\sigma^i(v) \tag{IV.4.21}$$

defined for $x, u, v \in \mathbb{R}^d$ for each $i = 1, 2, ..., d$. By the boundedness of $|b|$ and the Lipschitz condition on σ and $b\sigma$ we have

$$|f^i(x, u, v)| \leqslant KL|u - v|, \quad |g^i(u, v)| \leqslant L|u - v| \quad x, u, v \in \mathbb{R}^d,\ i = 1, 2, ..., d.$$

Thus, taking into account (IV.4.3) and (IV.4.13), from (IV.4.20) we obtain

$$p^2(p-1)R \leqslant KLN \int_{\mathbb{R}^{dp}} \rho_{2\varepsilon}(y)|\mu_p|(dy) = KLN||\mu|^{(2\varepsilon)}|_{L_p}^p \leqslant NKL||\mu|^{(\varepsilon)}|_{L_p}^p$$

with a constant $N = N(d, p)$, that finishes the proof of (IV.4.16). $\square$

For $\mathbb{R}^d$-valued functions ξ on $\mathbb{R}^d$ we define the linear operators I^ξ, J^ξ and T^ξ by

$$T^\xi\varphi(x) = \varphi(x+\xi(x)), \quad I^\xi\varphi(x) := T^\xi\varphi(x) - \varphi(x),$$
$$J^\xi\psi(x) := I^\xi\psi(x) - \xi^i(x)D_i\psi(x), \quad x \in \mathbb{R}^d \tag{IV.4.22}$$

acting on functions φ and differentiable functions ψ on $\mathbb{R}^d$. If ξ depends also on some parameters, then $I^\xi\phi$ and $J^\xi\psi$ are defined for each fixed parameter as above.

Lemma IV.4.4. *Let ξ be an $\mathbb{R}^d$-valued function of $x \in \mathbb{R}^d$ such that for some constants $\lambda > 0$, K_0 and L*

$$|\xi(x) - \xi(y)| \leqslant L|x-y| \quad \textit{for all } x, y \in \mathbb{R}^d$$

and

$$\lambda|x-y| \leqslant |x-y+\theta(\xi(x)-\xi(y))| \quad \textit{for all } x,y \in \mathbb{R}^d \textit{ and } \theta \in [0,1]. \tag{IV.4.23}$$

Let $\mu \in \mathfrak{M}$ such that it satisfies (IV.4.5)*, let $p \geqslant 2$ be an integer, and for $\varepsilon > 0$ set*

$$C := \int_{\mathbb{R}^d} p(\mu^{(\varepsilon)})^{p-1}(J^{\xi*}\mu)^{(\varepsilon)} + (\mu^{(\varepsilon)} + (I^{\xi*}\mu)^{(\varepsilon)})^p - (\mu^{(\varepsilon)})^p - p(\mu^{(\varepsilon)})^{p-1}(I^{\xi*}\mu)^{(\varepsilon)}\,dx,$$

where, to ease notation, the argument $x \in \mathbb{R}^d$ is suppressed in the integrand. Then

$$|C| \leqslant N(1+L^2)L^2 ||\mu|^{(\varepsilon)}|^p_{L_p} \quad \textit{for all } \varepsilon > 0, \tag{IV.4.24}$$

with a constant $N = N(d,p,\lambda)$.

Remark IV.4.1. Notice that in the special case $p = 2$ the estimate (IV.4.24) can be rewritten as

$$2(\mu^{(\varepsilon)}, (J^{\xi*}\mu)^{(\varepsilon)}) + ((I^{\xi*}\mu)^{(\varepsilon)}, (I^{\xi*}\mu)^{(\varepsilon)}) \leqslant N(1+L^2)L^2 ||\mu|^{(\varepsilon)}|^2_{L_2} \quad \text{for all } \varepsilon > 0.$$

Proof of Lemma IV.4.4. Again we note that by (IV.4.9), together with the conditions on ξ and that by Taylor's formula

$$I^\xi k_\varepsilon(x) = \int_0^1 (D_i k_\varepsilon)(x - r - \theta\xi(r))\,d\theta\,\xi^i(r),$$

$$J^\xi k_\varepsilon(x) = \int_0^1 (1-\theta)(D_{ij}k_\varepsilon)(x - r - \theta\xi(r))\,d\theta\,\xi^i(r)\xi^j(r),$$

as well as that $\sup_{x\in\mathbb{R}^d}\sum_{k=0}^2 |D^k\rho_\varepsilon(x)| < \infty$, it is easy to verify that C is well-defined. Notice that

$$\mu^{(\varepsilon)} + (I^{\xi*}\mu)^{(\varepsilon)} = (T^{\xi*}\mu)^{(\varepsilon)}$$

and

$$p(\mu^{(\varepsilon)})^{p-1}(J^{\xi*}\mu)^{(\varepsilon)} - p(\mu^{(\varepsilon)})^{p-1}(I^{\xi*}\mu)^{(\varepsilon)} = -p(\mu^{(\varepsilon)})^{p-1}((\xi^i D_i)^*\mu)^{(\varepsilon)}.$$

Hence

$$C = \int_{\mathbb{R}^d} ((T^{\xi*}\mu)^{(\varepsilon)})^p - (\mu^{(\varepsilon)})^p - p(\mu^{(\varepsilon)})^{p-1}((\xi^i D_i)^*\mu)^{(\varepsilon)}\,dx.$$

Rewriting here the product of integrals as multiple integrals and using the product measure $\mu_p(dy) := \mu(dy_1)...\mu(dy_p)$ by Fubini's theorem we get

$$((T^{\xi*}\mu)^{(\varepsilon)})^p(x) = \int_{\mathbb{R}^{pd}} \Pi_{r=1}^p T^{\xi}_{y_r}\Pi_{r=1}^p k_\varepsilon(x-y_r)\,\mu_p(dy),$$

$$(\mu^{(\varepsilon)})^p = \int_{\mathbb{R}^{pd}} \Pi_{r=1}^p k_\varepsilon(x-y_r)\,\mu_p(dy),$$

$$\begin{aligned} p(\mu^{(\varepsilon)})^{p-1}((\xi^i D_i)^*\mu)^{(\varepsilon)} =& p\int_{\mathbb{R}^{pd}} \Pi_{r=1}^{p-1}k_\varepsilon(x-y_r)\xi^i(y_p)\partial_{y_p^i}k_\varepsilon(x-y_p)\mu_p(dy), \\ =& p\int_{\mathbb{R}^{pd}} \xi^i(y_p)\partial_{y_p^i}\Pi_{r=1}^p k_\varepsilon(x-y_r)\mu_p(dy) \\ =& \int_{\mathbb{R}^{pd}} \sum_{r=1}^p \xi^i(y_r)\partial_{y_r^i}\Pi_{r=1}^p k_\varepsilon(x-y_r)\,\mu_p(dy) \end{aligned} \quad \text{(IV.4.25)}$$

where the last equation is due to the symmetry of the function $\Pi_r^p k_\varepsilon(x-y_r)$ and the measure $\mu_p(dy)$ in $y = (y_1, ..., y_p) \in \mathbb{R}^{pd}$. Thus

$$C = \int_{\mathbb{R}^d}\int_{\mathbb{R}^{pd}} L_y \Pi_{r=1}^p k_\varepsilon(x-y_r)\,\mu_p(dy)\,dx$$

with the operator

$$L_y^\xi = \Pi_{r=1}^p T_{y_r}^\xi - \mathbb{I} - \sum_{r=1}^p \xi^i(y_r)\partial_{y_r^i}.$$

Using here Fubini's theorem then changing the order of the operator L_y^ξ and the integration against dx we have

$$\begin{aligned} C = \int_{\mathbb{R}^{pd}}\int_{\mathbb{R}^d} L_y^\xi \Pi_{r=1}^p k_\varepsilon(x-y_r)\,dx\,\mu_p(dy) &= \int_{\mathbb{R}^{pd}} L_y^\xi \int_{\mathbb{R}^d} \Pi_{r=1}^p k_\varepsilon(x-y_r)\,dx\,\mu_p(dy) \\ &= \int_{\mathbb{R}^{pd}} L_y^\xi \rho_\varepsilon(y)\,\mu_p(dy), \end{aligned} \quad \text{(IV.4.26)}$$

where, see (IV.4.2),

$$\rho_\varepsilon(y) = c_{p,\varepsilon}e^{-\sum_{1\leqslant r<s\leqslant p}|y_r-y_s|^2/(2\varepsilon p)} \quad \text{(IV.4.27)}$$

with $c_{p,\varepsilon} = p^{-d/2}(2\pi\varepsilon)^{(1-p)d/2}$.

Introduce for $\varepsilon > 0$ the function

$$\psi_\varepsilon(z) = c_{p,\varepsilon} e^{-\sum_{1\leqslant r<s\leqslant p} z_{rs}^2/(2p\varepsilon)}, \quad z = (z_{rs})_{1\leqslant r<s\leqslant p} \in \mathbb{R}^{p(p-1)d/2}.$$

Then clearly, $\rho_\varepsilon(y) = \psi_\varepsilon(\tilde{y})$ with

$$\tilde{y} := (y_{rs})_{1\leqslant r<s\leqslant p} := (y_r - y_s)_{1\leqslant r<s\leqslant p} \in \mathbb{R}^{p(p-1)d/2},$$

$\Pi_{r=1}^p T_{y_r}^\xi \rho_\varepsilon(y) = \psi_\varepsilon(\tilde{y}+\tilde{\xi}(y))$ with $\tilde{\xi}(y) = (\tilde{\xi}_{rs}(y))_{1\leqslant r<s\leqslant p}$, $\tilde{\xi}_{rs}(y) = \xi(y_r) - \xi(y_s)$, and by the chain rule

$$\sum_{r=1}^p \xi^i(y_r)\partial_{y_r^i}\rho_\varepsilon = \sum_{r=1}^p \xi^i(y_r) \sum_{1\leqslant k<l\leqslant p} (\delta_{kr} - \delta_{lr})(\partial_{z_{kl}^i}\psi_\varepsilon)(\tilde{y})$$

$$= \sum_{1\leqslant k<l\leqslant p} \sum_{r=1}^p (\delta_{kr} - \delta_{lr})\xi^i(y_r)(\partial_{z_{kl}^i}\psi_\varepsilon)(\tilde{y}) = \sum_{1\leqslant k<l\leqslant p} (\xi^i(y_k) - \xi^i(y_l))(\partial_{z_{kl}^i}\psi_\varepsilon)(\tilde{y}).$$

Consequently,

$$L_y^\xi \rho_\varepsilon(y) = \psi_\varepsilon(\tilde{y} + \tilde{\xi}(y)) - \psi_\varepsilon(\tilde{y}) - \sum_{1\leqslant k<l\leqslant p} \tilde{\xi}_{kl}^i(y)(\partial_{z_{kl}^i}\psi_\varepsilon)(\tilde{y}),$$

which by Taylor's formula gives

$$L_y^\xi \rho_\varepsilon(y) = \int_0^1 (1-\theta) \sum_{1\leqslant k<l\leqslant p} \sum_{1\leqslant r<s\leqslant p} (\partial_{z_{kl}^i}\partial_{z_{rs}^j}\psi_\varepsilon)(\tilde{y} + \theta\tilde{\xi}(y))\tilde{\xi}_{kl}^i(y)\tilde{\xi}_{rs}^j(y)\, d\theta,$$

where the summation convention is used with respect to the repeated indices $i, j = 1, 2, ..., d$. Note that

$$(\partial_{z_{kl}^i}\partial_{z_{rs}^j}\psi_\varepsilon)(\tilde{y} + \theta\tilde{\xi}(y)) = \psi_\varepsilon(\tilde{y} + \theta\tilde{\xi}(y)) l_\varepsilon^{ij,rs,kl}(\tilde{y} + \theta\tilde{\xi}(y))$$

with

$$l_\varepsilon^{ij,rs,kl}(z) := \tfrac{1}{(p\varepsilon)^2} z_{kl}^i z_{rs}^j - \tfrac{1}{p\varepsilon}\delta_{rk}\delta_{sl}\delta_{ij} \quad \text{for } z = (z_{kl})_{1\leqslant k<l\leqslant p}.$$

Due to the condition (IV.4.23) there is a constant $\kappa = \kappa(d, \lambda) > 1$ such that for $\theta \in [0, 1]$

$$\kappa^{-1}|x_1 - x_2|^2 \leqslant |x_1 - x_2 + \theta(\xi(x_1) - \xi(x_2))|^2 \quad \text{for } x_1, x_2 \in \mathbb{R}^d, \qquad \text{(IV.4.28)}$$

which implies

$$\psi_\varepsilon(\tilde{y} + \theta\tilde{\xi}(y)) \leqslant N\psi_\varepsilon(\tilde{y}/\kappa) = N\rho_{\kappa\varepsilon}(y), \quad y \in \mathbb{R}^{dp},$$

and together with the Lipschitz condition on ξ,

$$|l_{\varepsilon}^{ij,rs,kl}(\tilde{y}+\theta\tilde{\xi}(y))| \leqslant \tfrac{N}{\varepsilon^2}(1+L^2)\sum_{1\leqslant r<s\leqslant p}|y_r-y_s|^2+\tfrac{N}{\varepsilon}$$

for all $1 \leqslant r < s \leqslant p$, $1 \leqslant k < l \leqslant p$ and $i,j = 1,2,...,d$ with a constant $N = N(d,p,\lambda)$. Moreover,

$$|\tilde{\xi}_{kl}^{i}(y)\tilde{\xi}_{rs}^{j}(y)| \leqslant NL^2\sum_{1\leqslant r<s\leqslant p}|y_s-y_p|^2 \quad \text{for } y=(y_r)_{r=1}^{p}\in\mathbb{R}^{pd}$$

with a constant $N = N(d,p)$. Consequently, taking into account (IV.4.13) we have

$$\begin{aligned}|L_y^{\xi}\rho_{\varepsilon}(y)| \leqslant &\tfrac{N}{\varepsilon^2}L^2(1+L^2)\Big(\sum_{1\leqslant r<s\leqslant p}|y_r-y_s|^2\Big)^2\rho_{\kappa\varepsilon}(y)+\tfrac{N}{\varepsilon}L^2\sum_{1\leqslant r<s\leqslant p}|y_r-y_s|^2\rho_{\kappa\varepsilon}(y)\\ \leqslant &N'L^2(1+L^2)\rho_{2\kappa\varepsilon}(y) \quad \text{for } y\in\mathbb{R}^{dp}\end{aligned}$$

with constants $N = N(d,p,\lambda)$ and $N' = N'(d,p,\lambda)$. Using this we finish the proof by noting that (IV.4.26) implies

$$\begin{aligned}|C| \leqslant N'L^2(1+L^2)\int_{\mathbb{R}^{pd}}\rho_{2\kappa\varepsilon}(y)\,|\mu_p|(dy) &= N'L^2(1+L^2)||\mu|^{(2\kappa\varepsilon)}|_{L_p}^{p}\\ &\leqslant NL^2(1+L^2)||\mu|^{(\varepsilon)}|_{L_p}^{p}.\end{aligned}$$

□

Corollary IV.4.5. *Let the conditions of Lemma IV.4.4 hold. Then for every even integer $p \geqslant 2$ there is a constant $N = N(d,p,\lambda)$ such that for $\varepsilon > 0$ and $\mu \in \mathfrak{M}$ we have*

$$\int_{\mathbb{R}^d}(\mu^{(\varepsilon)})^{p-1}(x)(J^{\xi*}\mu)^{(\varepsilon)}(x)\,dx \leqslant NL^2(1+L^2)||\mu|^{(\varepsilon)}|_{L_p}^{p}. \tag{IV.4.29}$$

Proof. Notice that $|a+b|^p-|b|^p-p|a|^{p-2}ab \geqslant 0$ for $p \geqslant 2$ for any $a,b\in\mathbb{R}$ by the convexity of the function $f(a)=|a|^p$, $a\in\mathbb{R}$. Using this with $a=\mu^{(\varepsilon)}$ and $b=(I^{\xi}\mu)^{(\varepsilon)}$ we have

$$(\mu^{(\varepsilon)}+(I^{\xi*}\mu)^{(\varepsilon)})^p-(\mu^{(\varepsilon)})^p-p(\mu^{(\varepsilon)})^{p-1}(I^{\xi*}\mu)^{(\varepsilon)} \geqslant 0 \quad \text{for } x\in\mathbb{R}^d,$$

which shows that (IV.4.24) implies (IV.4.29), since

$$\int_{\mathbb{R}^d}|\mu^{(\varepsilon)}(x)|^{p-1}|(J^{\xi*}\mu)^{(\varepsilon)}(x)|\,dx<\infty.$$

□

Lemma IV.4.6. *Let the conditions of Lemma IV.4.4 hold. Let $p \geqslant 2$ be an even* *integer. Then there is a constant $N = N(d, p, \lambda)$ such that for $\varepsilon > 0$ and $\mu \in \mathfrak{M}$ we have*

$$\Big|\int_{\mathbb{R}^d} (\mu^{(\varepsilon)} + (I^{\xi *}\mu)^{(\varepsilon)})^p - (\mu^{(\varepsilon)})^p \, dx\Big| \leqslant N(1+L)L||\mu|^{(\varepsilon)}|^p_{L_p}, \tag{IV.4.30}$$

where the argument $x \in \mathbb{R}^d$ is suppressed in the integrand.

Proof. By the same arguments as in the proof of Lemma IV.4.4 we see that the left-hand side of (IV.4.30) is well-defined. Clearly,

$$D := \int_{\mathbb{R}^d} ((T^{\xi *}\mu)^{(\varepsilon)})^p - (\mu^{(\varepsilon)})^p \, dx = \int_{\mathbb{R}^d} \int_{\mathbb{R}^{pd}} M^{\xi}_y \Pi^p_{r=1} k_{\varepsilon}(x - y_r) \, \mu_p(dy) \, dx$$

with the operator

$$M^{\xi}_y = \Pi^p_{r=1} T^{\xi}_{y_r} - \mathbb{I},$$

where $\mu_p(dy) = \Pi^p_{r=1}\mu(dy_r)$, $y = (y_1, ..., y_p) \in \mathbb{R}^{pd}$. Hence by Fubini's theorem, then changing the order of the operator M^{ξ}_y and the integration against dx and by taking into account (IV.4.2) we have

$$D = \int_{\mathbb{R}^{pd}} \int_{\mathbb{R}^d} M^{\xi}_y \Pi^p_{r=1} k_{\varepsilon}(x - y_r) \, dx \, \mu_p(dy) = \int_{\mathbb{R}^{pd}} M^{\xi}_y \int_{\mathbb{R}^d} \Pi^p_{r=1} k_{\varepsilon}(x - y_r) \, dx \, \mu_p(dy)$$

$$= \int_{\mathbb{R}^{pd}} M^{\xi}_y \rho_{\varepsilon}(y) \, \mu_p(dy). \tag{IV.4.31}$$

As in the proof of Lemma IV.4.4 we introduce for $\varepsilon > 0$ the function

$$\psi_{\varepsilon}(z) = c_{p,\varepsilon} e^{-\sum_{1 \leqslant r < s \leqslant p} z^2_{rs}/(2p\varepsilon)}, \quad z = (z_{rs})_{1 \leqslant r < s \leqslant p} \in \mathbb{R}^{p(p-1)d/2},$$

such that $\rho_{\varepsilon}(y) = \psi_{\varepsilon}(\tilde{y})$ with $\tilde{y} := (y_{rs})_{1 \leqslant r < s \leqslant p} := (y_r - y_s)_{1 \leqslant r < s \leqslant p} \in \mathbb{R}^{p(p-1)d/2}$, and

$$\Pi^p_{r=1} T^{\xi}_{y_r} \rho_{\varepsilon}(y) = \psi_{\varepsilon}(\tilde{y} + \tilde{\xi}(y)) \quad \text{with } \tilde{\xi}(y) = (\tilde{\xi}_{rs}(y))_{1 \leqslant r < s \leqslant p}, \quad \tilde{\xi}_{rs}(y) = \xi(y_r) - \xi(y_s).$$

By Taylor's formula

$$M^{\xi}_y \rho_{\varepsilon}(y) = \psi_{\varepsilon}(\tilde{y} + \tilde{\xi}(y)) - \psi_{\varepsilon}(\tilde{y}) = \int_0^1 \sum_{1 \leqslant k < l \leqslant p} (\partial_{z^i_{kl}} \psi_{\varepsilon})(\tilde{y} + \theta\tilde{\xi}(y)) \tilde{\xi}^i_{kl}(y) \, d\theta,$$

where the summation convention is used with respect to the repeated indices $i = 1, 2, ..., d$. Note that

$$(\partial_{z^i_{kl}} \psi_{\varepsilon})(\tilde{y} + \theta\tilde{\xi}(y)) = \psi_{\varepsilon}(\tilde{y} + \theta\tilde{\xi}(y)) l^{kl,i}_{\varepsilon}(\tilde{y} + \theta\tilde{\xi}(y))$$

with

$$l^{kl,i}_{\varepsilon}(z) := \tfrac{1}{p\varepsilon} z^i_{kl} \quad \text{for } z = (z_{kl})_{1 \leqslant k < l \leqslant p} \in \mathbb{R}^{p(p-1)d/2}.$$

By (IV.4.28) we have

$$\psi_\varepsilon(\tilde{y}+\theta\tilde{\xi}(y)) \leqslant N\psi_\varepsilon(\tilde{y}/\kappa) = N\rho_{\kappa\varepsilon}(y), \quad y \in \mathbb{R}^{dp},$$

and due to the Lipschitz condition on ξ,

$$|l_\varepsilon^{kl,i}(\tilde{y}+\theta\tilde{\xi}(y))| \leqslant \tfrac{N}{\varepsilon}(1+L)|y_k-y_l|$$

for all $i=1,2,...,d$ with constants $\kappa(d,\lambda)>1$ and $N=N(d,p,\lambda)$. Moreover, we get

$$|\tilde{\xi}^i_{kl}(y)| \leqslant NL|y_k-y_l| \quad \text{for } y=(y_r)_{r=1}^p \in \mathbb{R}^{pd}$$

with a constant $N=N(d,p)$. Consequently, taking into account (IV.4.13) we have

$$|M_y^\xi\rho_\varepsilon(y)| \leqslant \tfrac{N}{\varepsilon}(1+L)L \sum_{1\leqslant k<l\leqslant p}|y_k-y_l|^2\rho_{\kappa\varepsilon}(y) \leqslant N'(1+L)L\rho_{2\kappa\varepsilon}(y) \quad \text{for } y\in\mathbb{R}^{dp}$$

with constants N and N' depending only on d, p and λ. Using this, from (IV.4.31) we obtain

$$|D| \leqslant N'(1+L)L\int_{\mathbb{R}^{pd}}\rho_{2\kappa\varepsilon}(y)\,|\mu_p|(dy) = N'(1+L)L||\mu|^{(2\kappa\varepsilon)}|^p_{L_p} \leqslant N'(1+L)L||\mu|^{(\varepsilon)}|^p_{L_p},$$

which completes the proof of the lemma. □

IV.5 The smoothed measures

We use the notations introduced in Section IV.4. Moreover, we ask the reader to recall the notations introduced in section I.1, in particular the notion of cadlagness employed for $\mathbb{M}$-valued, or $\mathfrak{M}$-valued processes.

We present first a version of an Itô formula, Theorem 2.1 from [21], for L_p-valued processes. To formulate it, let $\psi=(\psi(x))$, $f=(f_t(x))$, $g=(g_t^j(x))$ and $h=(h_t(x,\mathfrak{z}))$ be functions with values in $\mathbb{R}$, $\mathbb{R}$, $\mathbb{R}^m$ and $\mathbb{R}$, respectively, defined on $\Omega\times\mathbb{R}^d$, $\Omega\times H_T$, $\Omega\times H_T$ and $\Omega\times H_T\times\mathfrak{Z}$, respectively, where $H_T:=[0,T]\times\mathbb{R}^d$ and $(\mathfrak{Z},\mathcal{Z},\nu)$ is a measure space with a σ-finite measure ν and countably generated σ-algebra $\mathcal{Z}$. Assume that ψ is $\mathcal{F}_0\otimes\mathcal{B}(\mathbb{R}^d)$-measurable, f and g are $\mathcal{O}\otimes\mathcal{B}(\mathbb{R}^d)$-measurable and h is $\mathcal{P}\otimes\mathcal{B}(\mathbb{R}^d)\otimes\mathcal{Z}$-measurable, such that almost surely

$$\int_0^T|f_t(x)|\,dt<\infty, \qquad \int_0^T\sum_j|g_t^j(x)|^2\,dt<\infty, \qquad \int_0^T\int_{\mathfrak{Z}}|h_t(x,\mathfrak{z})|^2\,\nu(d\mathfrak{z})\,dt<\infty \tag{IV.5.1}$$

for each $x \in \mathbb{R}^d$, and for each bounded Borel set $\Gamma \subset \mathbb{R}^d$ almost surely

$$\int_\Gamma \int_0^T |f_t(x)| \, dt \, dx < \infty, \quad \int_\Gamma \Big(\int_0^T \sum_j |g_t^j(x)|^2 \, dt \Big)^{1/2} dx < \infty,$$

$$\int_\Gamma \Big(\int_0^T \int_{\mathfrak{Z}} \sum_j |h_t(x, \mathfrak{z})|^2 \, \nu(d\mathfrak{z}) \, dt \Big)^{1/2} dx < \infty. \tag{IV.5.2}$$

Assume, moreover, that for a number $p \in [2, \infty)$ almost surely

$$\int_{\mathbb{R}^d} |\psi(x)|^p \, dx < \infty, \quad \int_0^T \int_{\mathbb{R}^d} |f_t(x)|^p \, dx \, dt < \infty,$$

$$\int_0^T \int_{\mathbb{R}^d} \Big(\sum_j |g_t^j(x)|^2 \Big)^{p/2} dx \, dt < \infty, \tag{IV.5.3}$$

$$\int_0^T \int_{\mathbb{R}^d} \int_{\mathfrak{Z}} |h_t(x, \mathfrak{z})|^p \, \nu(d\mathfrak{z}) \, dx \, dt < \infty, \quad \int_0^T \int_{\mathbb{R}^d} \Big(\int_{\mathfrak{Z}} |h_t(x, \mathfrak{z})|^2 \, \nu(d\mathfrak{z}) \Big)^{p/2} dx \, dt < \infty.$$

Theorem IV.5.1. *Let conditions* (IV.5.1), (IV.5.2) *and* (IV.5.3) *with a number $p \geqslant 2$ hold. Assume there is an $\mathcal{O} \otimes \mathcal{B}(\mathbb{R}^d)$-measurable real-valued function v on $\Omega \times H_T$ such that almost surely*

$$\int_{\mathbb{R}^d} |v_t(x)|^p \, dx < \infty \quad \text{for all } t \in [0, T] \tag{IV.5.4}$$

and for every $x \in \mathbb{R}^d$ almost surely

$$v_t(x) = \psi(x) + \int_0^t f_s(x) \, ds + \int_0^t g_s^j(x) \, dw_s^j + \int_0^t \int_{\mathfrak{Z}} h_s(x, \mathfrak{z}) \, \tilde{\pi}(d\mathfrak{z}, ds) \tag{IV.5.5}$$

holds for all $t \in [0, T]$, where $(w_t)_{t \geqslant 0}$ is an m-dimensional $\mathcal{F}_t$-Wiener process, $\pi(d\mathfrak{z}, ds)$ is an $\mathcal{F}_t$-Poisson measure with characteristic measure ν, and $\tilde{\pi}(d\mathfrak{z}, ds) = \pi(d\mathfrak{z}, ds) - \nu(d\mathfrak{z})ds$ is the compensated martingale measure. Then $(v_t)_{t \in [0,T]}$ is an L_p-valued $\mathcal{F}_t$-adapted cadlag process, satisfying almost surely

$$\begin{aligned}
|v_t|_{L_p}^p = {} & |\psi|_{L_p}^p + p \int_0^t \int_{\mathbb{R}^d} |v_s|^{p-2} v_s g_s^j \, dx \, dw_s^j \\
& + \tfrac{p}{2} \int_0^t \int_{\mathbb{R}^d} \big(2|v_s|^{p-2} v_s f_s + (p-1)|v_s|^{p-2} \sum_j |g_s^j|^2 \big) \, dx \, ds \\
& + p \int_0^t \int_{\mathfrak{Z}} \int_{\mathbb{R}^d} |v_{s-}|^{p-2} v_{s-} h_s \, dx \, \tilde{\pi}(d\mathfrak{z}, ds) \\
& + \int_0^t \int_{\mathfrak{Z}} \int_{\mathbb{R}^d} (|v_{s-} + h_s|^p - |v_{s-}|^p - p|v_{s-}|^{p-2} v_{s-} h_s) \, dx \, \pi(d\mathfrak{z}, ds)
\end{aligned} \tag{IV.5.6}$$

for all $t \in [0,T]$, where v_{s-} means the left-hand limit in L_p at s of v.

Proof. By a truncation and stopping time argument it is not difficult to see that without loss of generality we may assume that the random variables in (IV.5.3) have finite expectation.

Our aim is to use an Itô formula from [21], stated in Theorem 2.1 therein. First we need to show that there exist "regular" versions of the stochastic integrals in (IV.5.5) such that they are measurable in $x \in \mathbb{R}^d$. Indeed, by Lemma 2.6 in [37] we know that there exists a real-valued function m on $\Omega \times [0,T] \times \mathbb{R}^d$ such that

(i) it is measurable with respect to $\mathcal{F} \otimes \mathcal{B}(\mathbb{R}_+) \otimes \mathcal{B}(\mathbb{R}^d)$,

(ii) for each $x \in \mathbb{R}^d$ the process $(m_t(x))_{t\in[0,T]}$ has continuous paths,

(iii) for each $x \in \mathbb{R}^d$ the process $(m_t(x))_{t\in[0,T]}$ is a local $\mathcal{F}_t$-martingale starting at zero

and for each $x \in \mathbb{R}^d$ we have

$$m_t(x) = \int_0^t g_s^j(x)\, dw_s^j, \quad \text{almost surely for all } t \in [0,T].$$

By Theorem 3.4 in [21] there exists a function r on $\Omega \times [0,T] \times \mathbb{R}^d$ with the properties (i) & (iii) above, such that for each $x \in \mathbb{R}^d$ the process $(r_t(x))_{t\in[0,T]}$ has cadlag paths, as well as such that for each $x \in \mathbb{R}^d$ we have

$$r_t(x) = \int_0^t \int_{\mathfrak{Z}} h_s(x,\mathfrak{z})\, \tilde{\pi}(d\mathfrak{z}, ds), \quad \text{almost surely for all } t \in [0,T].$$

From (IV.5.5) then we get that for each $\varphi \in C_0^\infty$ almost surely

$$(v_t,\varphi) = (\psi,\varphi) + \int_0^t (f_s,\varphi)\, ds + \int_0^t (g_s^j,\varphi)\, dw_s^j + \int_0^t \int_{\mathfrak{Z}} (h_s(\mathfrak{z}),\varphi)\, \tilde{\pi}(d\mathfrak{z}, ds) \quad \text{(IV.5.7)}$$

holds for all $t \in [0,T]$. This we can see if we multiply both sides of equation (IV.5.5) with φ and then, making use of our measurability conditions and the conditions (IV.5.1) and (IV.5.2), as well as the versions m and r of the stochastic integrals, we integrate over $\mathbb{R}^d$ with respect to dx and use deterministic and stochastic Fubini theorems from [37] and [21] to change the order of integrations. Due to (IV.5.7), the measurability conditions on ψ, f, g, h and v and to their integrability conditions, (IV.5.3) and (IV.5.4), by virtue of Theorem 2.1 from [21] there is an L_p-valued $\mathcal{F}_t$-adapted cadlag process $(\bar{v}_t)_{t\in[0,T]}$ such that for each $\varphi \in C_0^\infty$ almost surely (IV.5.7) holds with $\bar{v}$ in place of v, and almost surely (IV.5.6) holds with $\bar{v}$ in place of v. Moreover, for each $\varphi \in C_0^\infty$ almost surely $(v_t,\varphi) = (\bar{v}_t,\varphi)$ for all $t \in [0,T]$, which implies that almost surely $v = \bar{v}$ as L_p-valued processes, and this finishes the proof of the theorem. □

Lemma IV.5.2. *Let Assumption IV.2.1 hold. Assume* $(\mu_t)_{t\in[0,T]}$ *is an* $\mathfrak{M}$*-solution to equation* (IV.3.7)*. If* $K_1 \neq 0$ *in Assumption IV.2.1 (ii) then assume also*

$$\operatorname*{ess\,sup}_{t\in[0,T]} \int_{\mathbb{R}^d} |y|^2 \, |\mu_t|(dy) < \infty \ (a.s.). \tag{IV.5.8}$$

Then for each $x \in \mathbb{R}^d$ *and* $\varepsilon > 0$*,*

$$\begin{aligned}\mu_t^{(\varepsilon)}(x) &= \mu_0^{(\varepsilon)}(x) + \int_0^t (\tilde{\mathcal{L}}_s^* \mu_s)^{(\varepsilon)}(x)\, ds + \int_0^t (\mathcal{M}_s^{j*} \mu_s)^{(\varepsilon)}(x)\, dV_s^j \\ &+ \int_0^t \int_{\mathfrak{Z}_0} (J_s^{\eta*} \mu_s)^{(\varepsilon)}(x)\, \nu_0(d\mathfrak{z})\, ds + \int_0^t \int_{\mathfrak{Z}_1} (J_s^{\xi*} \mu_s)^{(\varepsilon)}(x)\, \nu_1(d\mathfrak{z})\, ds \\ &+ \int_0^t \int_{\mathfrak{Z}_1} (I_s^{\xi*} \mu_{s-})^{(\varepsilon)}(x)\, \tilde{N}_1(d\mathfrak{z}, ds)\end{aligned} \tag{IV.5.9}$$

holds almost surely for all $t \in [0,T]$*. Moreover, for each* $\varepsilon > 0$ *and* $p \geqslant 2$

$$\begin{aligned}|\mu_t^{(\varepsilon)}|_{L_p}^p &= |\mu_0^{(\varepsilon)}|_{L_p}^p + p\int_0^t \left(|\mu_s^{(\varepsilon)}|^{p-2}\mu_s^{(\varepsilon)}, (\mathcal{M}_s^{k*}\mu_s)^{(\varepsilon)}\right) dV_s^k \\ &+p\int_0^t \left(|\mu_s^{(\varepsilon)}|^{p-2}\mu_s^{(\varepsilon)}, (\tilde{\mathcal{L}}_s^*\mu_s)^{(\varepsilon)}\right) ds + \tfrac{p(p-1)}{2}\sum_k \int_0^t \left(|\mu_s^{(\varepsilon)}|^{p-2}, |(\mathcal{M}_s^{k*}\mu_s)^{(\varepsilon)}|^2\right) ds \\ &+p\int_0^t\int_{\mathfrak{Z}_0} \left(|\mu_s^{(\varepsilon)}|^{p-2}\mu_s^{(\varepsilon)}, (J_s^{\eta*}\mu_s)^{(\varepsilon)}\right) \nu_0(d\mathfrak{z})ds \\ &+ p\int_0^t\int_{\mathfrak{Z}_1} \left(|\mu_s^{(\varepsilon)}|^{p-2}\mu_s^{(\varepsilon)}, (J_s^{\xi*}\mu_s)^{(\varepsilon)}\right) \nu_1(d\mathfrak{z})ds \\ &+p\int_0^t\int_{\mathfrak{Z}_1} \left(|\mu_{s-}^{(\varepsilon)}|^{p-2}\mu_{s-}^{(\varepsilon)}, (I_s^{\xi*}\mu_{s-})^{(\varepsilon)}\right) \tilde{N}_1(d\mathfrak{z}, ds) \\ &+\int_0^t\int_{\mathfrak{Z}_1}\int_{\mathbb{R}^d} \left\{\left|\mu_{s-}^{(\varepsilon)} + (I_s^{\xi*}\mu_{s-})^{(\varepsilon)}\right|^p - |\mu_{s-}^{(\varepsilon)}|^p - p|\mu_{s-}^{(\varepsilon)}|^{p-2}\mu_{s-}^{(\varepsilon)}(I_s^{\xi*}\mu_{s-})^{(\varepsilon)}\right\} dx N_1(d\mathfrak{z}, ds)\end{aligned} \tag{IV.5.10}$$

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

Proof. Let $\psi \in C_0^\infty(\mathbb{R}^d)$ such that $\psi(0) = 1$, and for integers $r \geqslant 1$ define ψ_r by dilation, $\psi_r(x) = \psi(x/r)$, $x \in \mathbb{R}^d$. Then substituting $k_\varepsilon(x-\cdot)\psi_r(\cdot) \in C_0^\infty$ in place of φ in (IV.3.4), for each $x \in \mathbb{R}^d$ we get

$$\mu_t(k_\varepsilon(x-\cdot)\psi_r) = \mu_0(k_\varepsilon(x-\cdot)\psi_r) + \int_0^t \mu_s(\tilde{\mathcal{L}}_s(k_\varepsilon(x-\cdot)\psi_r))\, ds$$

$$+\int_0^t \mu_s(\mathcal{M}_s^k(k_\varepsilon(x-\cdot)\psi_r))\, dV_s^k + \int_0^t\int_{\mathfrak{Z}_0} \mu_s(J_s^\eta(k_\varepsilon(x-\cdot)\psi_r))\, \nu_0(d\mathfrak{z})ds$$

$$+\int_0^t\int_{\mathfrak{Z}_1}\mu_s(J_s^{\xi}(k_\varepsilon(x-\cdot)\psi_r))\,\nu_1(d\mathfrak{z})ds+\int_0^t\int_{\mathfrak{Z}_1}\mu_{s-}(I_s^{\xi}(k_\varepsilon(x-\cdot)\psi_r))\,\tilde{N}_1(d\mathfrak{z},ds), \tag{IV.5.11}$$

almost surely for all $t\in[0,T]$. Clearly, $\lim_{r\to\infty}k_\varepsilon(x-y)\psi_r(y)=k_\varepsilon(x-y)$ and there is a constant N, independent of r, such that

$$|k_\varepsilon(x-y)\psi_r(y)|\leqslant N\quad\text{for all }x,y\in\mathbb{R}^d.$$

Hence almost surely

$$\lim_{r\to\infty}\mu_t(k_\varepsilon(x-\cdot)\psi_r)=\mu_t(k_\varepsilon(x-\cdot))\quad\text{for each }x\in\mathbb{R}^d\text{ and }t\in[0,T]. \tag{IV.5.12}$$

It is easy to see that for every $\omega\in\Omega$, $x,y\in\mathbb{R}^d$, $s\in[0,T]$ and $\mathfrak{z}_i\in\mathfrak{Z}_i$ $(i=0,1)$ we have

$$\lim_{r\to\infty}A_s(k_\varepsilon(x-y)\psi_r(y))=A_s(k_\varepsilon(x-y)) \tag{IV.5.13}$$

with $\tilde{\mathcal{L}}$, $\mathcal{M}^k$, J^η, J^ξ and I^ξ in place of A. Clearly,

$$\sup_{x\in\mathbb{R}^d}|\psi_r(x)|=\sup_{x\in\mathbb{R}^d}|\psi(x)|<\infty,\quad\sup_{x\in\mathbb{R}^d}|D\psi_r(x)|=r^{-1}\sup_{x\in\mathbb{R}^d}|D\psi(x)|<\infty,$$

$$\sup_{x\in\mathbb{R}^d}|D^2\psi_r(x)|=r^{-2}\sup_{x\in\mathbb{R}^d}|D^2\psi(x)|<\infty,$$

and there is a constant N depending only on d and ε such that for all $x,y\in\mathbb{R}^d$

$$|k_\varepsilon(x-y)|+|Dk_\varepsilon(x-y)|+|D^2k_\varepsilon(x-y)|\leqslant N. \tag{IV.5.14}$$

Hence, due to Assumption IV.2.1, we have a constant $N=N(\varepsilon,d,K,K_0,K_1)$ such that

$$|\tilde{\mathcal{L}}_s(k_\varepsilon(x-y)\psi_r(y))|\leqslant N(K_0^2+K_1^2|y|^2+K_1^2|Y_s|^2), \tag{IV.5.15}$$

$$\sum_k|\mathcal{M}_s^k(k_\varepsilon(x-y)\psi_r(y))|^2\leqslant N(K_0^2+K_1^2|y|^2+K_1^2|Y_s|^2) \tag{IV.5.16}$$

for $x,y\in\mathbb{R}^d$, $s\in[0,T]$, $r\geqslant 1$ and $\omega\in\Omega$. Similarly, applying Taylor's formula to

$$A_s(k_\varepsilon(x-y)\psi_r(y))\quad\text{with }J^\eta,\ J^\xi\text{ and }I^\xi\text{ in place of }A,$$

we have a constant $N=N(\varepsilon,d,K_0,K_1)$ such that

$$\begin{aligned}|J_s^\eta(k_\varepsilon(x-y)\psi_r(y))|&\leqslant\sup_{v\in\mathbb{R}^d}|D_v^2(k_\varepsilon(x-v)\psi_r(v))||\eta_s(y,\mathfrak{z}_0)|^2\\&\leqslant N|\eta_s(y,\mathfrak{z}_0)|^2, &\text{(IV.5.17)}\\|J_s^\xi(k_\varepsilon(x-y)\psi_r(y))|&\leqslant\sup_{v\in\mathbb{R}^d}|D_v^2(k_\varepsilon(x-v)\psi_r(v))||\xi_s(y,\mathfrak{z}_1)|^2\\&\leqslant N|\xi_s(y,\mathfrak{z}_1)|^2 &\text{(IV.5.18)}\end{aligned}$$

and

$$|I_s^\xi(k_\varepsilon(x-y)\psi_r(y))|^2 \leqslant \sup_{v\in\mathbb{R}^d}|D_v(k_\varepsilon(x-v)\psi_r(v))|^2|\xi_s(y,\mathfrak{z}_1)|^2$$
$$\leqslant N|\xi_s(y,\mathfrak{z}_1)|^2, \tag{IV.5.19}$$

respectively, for all $x,y\in\mathbb{R}^d$, $s\in[0,T]$, $\mathfrak{z}_i\in\mathfrak{Z}_i$, $i=0,1$ and $\omega\in\Omega$. Using (IV.5.13) (with $A:=\mathcal{L}$), (IV.5.15) and (IV.5.8), by Lebesgue's theorem on dominated convergence we get for each $x\in\mathbb{R}^d$

$$\lim_{r\to\infty}\int_0^t \mu_s(\tilde{\mathcal{L}}_s(k_\varepsilon(x-\cdot)\psi_r))\,ds = \int_0^t \mu_s(\tilde{\mathcal{L}}_s k_\varepsilon(x-\cdot))\,ds \quad \text{almost surely,}$$

uniformly in $t\in[0,T]$. Using Jensen's inequality, (IV.5.13) (with $A:=\mathcal{M}$), (IV.5.16) and (IV.5.8), by Lebesgue's theorem on dominated convergence we obtain

$$\limsup_{r\to\infty}\int_0^T\sum_k|\mu_s(\mathcal{M}_s^k(k_\varepsilon(x-\cdot)\psi_r))-\mu_s(\mathcal{M}_s^k(k_\varepsilon(x-\cdot))|^2\,ds$$

$$\leqslant \operatorname*{ess\,sup}_{s\in[0,T]}\|\mu_s\|\limsup_{r\to\infty}\int_0^T\int_{\mathbb{R}^d}\sum_k|\mathcal{M}_s^k(k_\varepsilon(x-\cdot)\psi_r))-\mathcal{M}_s^k(k_\varepsilon(x-\cdot))|^2\,|\mu_s|(dy)\,ds=0$$

almost surely, which implies that for $r\to\infty$

$$\int_0^t \mu_s(\mathcal{M}_s^k(k_\varepsilon(x-\cdot)\psi_r))\,dV_s^k \to \int_0^t \mu_s(\mathcal{M}_s^k k_\varepsilon(x-\cdot))\,dV_s^k$$

in probability, for each $x\in\mathbb{R}^d$, uniformly in $t\in[0,T]$. Since by Assumption IV.2.1(ii) and (IV.5.8)

$$\int_0^T\int_{\mathbb{R}^d}\int_{\mathfrak{Z}_0}|\eta_s(y,\mathfrak{z}_0)|^2\nu_0(d\mathfrak{z}_0)\,|\mu_s|(dy)\,ds \leqslant 2K_0^2|\bar{\eta}|_{L_2}^2\int_0^T\|\mu_s\|\,ds$$

$$+2K_1^2|\bar{\eta}|_{L_2}^2\int_0^T\int_{\mathbb{R}^d}|y|^2\,|\mu_s|(dy)\,ds+2K_1^2|\bar{\eta}|_{L_2}^2\int_0^T\int_{\mathbb{R}^d}|Y_s|^2\,|\mu_s|(dy)\,ds<\infty \quad \text{(a.s.)},$$

from (IV.5.13) (with $A:=J^\eta$) and (IV.5.17) by Lebesgue's theorem on dominated convergence we get

$$\lim_{r\to\infty}\int_0^t\int_{\mathfrak{Z}_0}\mu_s(J_s^\eta(k_\varepsilon(x-\cdot)\psi_r))\,\nu_0(d\mathfrak{z})ds=\int_0^t\int_{\mathfrak{Z}_0}\mu_s(J_s^\eta k_\varepsilon(x-\cdot))\,\nu_0(d\mathfrak{z})ds \quad \text{(a.s.)},$$

uniformly in $t\in[0,T]$. In the same way we obtain this with J^ξ, ν_1 and $\mathfrak{Z}_1$ in place of J^η, ν_0 and $\mathfrak{Z}_0$, respectively. Similarly, using first Jensen's inequality and

Fubini's theorem we have

$$\limsup_{r\to\infty}\int_0^T\int_{\mathfrak{Z}_1}|\mu_s(I_s^\xi(k_\varepsilon(x-\cdot)\psi_r))-\mu_s(I_s^\xi(k_\varepsilon(x-\cdot))|^2\,\nu_1(d\mathfrak{z})\,ds$$

$$\leqslant \operatorname*{ess\,sup}_{s\in[0,T]}\|\mu_s\|\limsup_{r\to\infty}\int_0^T\int_{\mathbb{R}^d}\int_{\mathfrak{Z}_1}|I_s^\xi(k_\varepsilon(x-y)\psi_r(y)))-I_s^\xi(k_\varepsilon(x-y))|^2\,\nu_1(d\mathfrak{z})\,|\mu_s|(dy)\,ds$$
$$=0$$

almost surely, which implies that for $r\to\infty$ for each $x\in\mathbb{R}^d$ we have

$$\int_0^t\int_{\mathfrak{Z}_1}\mu_s(I_{s-}^\xi(k_\varepsilon(x-\cdot)\psi_r))\,\tilde N_1(d\mathfrak{z},ds)\to\int_0^t\int_{\mathfrak{Z}_1}\mu_s(I_{s-}^\xi(k_\varepsilon(x-\cdot)))\,\tilde N_1(d\mathfrak{z},ds)$$

in probability, uniformly in $t\in[0,T]$. Consequently, letting $r\to\infty$ in equation (IV.5.11), we obtain that (IV.5.9) holds almost surely for each $t\in[0,T]$.

To prove (IV.5.10) we are going to verify that

$$f_t(x):=(\tilde{\mathcal{L}}_t^*\mu_t)^{(\varepsilon)}(x)+\int_{\mathfrak{Z}_0}(J_t^{\eta*}\mu_s)^{(\varepsilon)}(x)\,\nu_0(d\mathfrak{z}_0)+\int_{\mathfrak{Z}_1}(J_t^{\xi*}\mu_s)^{(\varepsilon)}(x)\,\nu_1(d\mathfrak{z}_1),$$

$$g_t^j(x):=(\mathcal{M}_t^{j*}\mu_t)^{(\varepsilon)}(x),\quad h_t(x,\mathfrak{z}):=(I_t^{\xi*}\mu_{t-})^{(\varepsilon)}(x),\quad v_t(x):=\mu_t^{(\varepsilon)}(x),$$

($\omega\in\Omega$, $t\in[0,T]$, $x\in\mathbb{R}^d$, $\mathfrak{z}\in\mathfrak{Z}_1$, $j=1,2,...,d'$) satisfy the conditions of Theorem IV.5.1 with the $\mathcal{F}_t$-Wiener process $w:=V$ and $\mathcal{F}_t$-Poisson martingale measure $\tilde\pi:=\tilde N_1$, carried by the probability space $(\Omega,\mathcal{F},Q)$ equipped with the filtration $(\mathcal{F}_t)_{t\geqslant 0}$. To see that f, g, h satisfy the required measurability properties first we claim that for bounded $\mathcal{O}\otimes\mathcal{B}(\mathbb{R}^d)\otimes\mathcal{B}(\mathbb{R}^d)\otimes\mathcal{Z}_0$-measurable functions $A=A_t(x,y,\mathfrak{z})$ and bounded $\mathcal{P}\otimes\mathcal{B}(\mathbb{R}^d)\otimes\mathcal{B}(\mathbb{R}^d)\otimes\mathcal{Z}_0$-measurable $A=A_t(x,y,\mathfrak{z})$, the functions

$$\int_{\mathbb{R}^d}A_t(x,y,\mathfrak{z})\mu_t(dy)\quad\text{and}\quad\int_{\mathbb{R}^d}B_t(x,y,\mathfrak{z})\mu_{t-}(dy)\tag{IV.5.20}$$

are $\mathcal{O}\otimes\mathcal{B}(\mathbb{R}^d)\otimes\mathcal{Z}_0$- and $\mathcal{P}\otimes\mathcal{B}(\mathbb{R}^d)\otimes\mathcal{Z}_0$-measurable, in $(\omega,t,x,\mathfrak{z})\in\Omega\times[0,T]\times\mathbb{R}^d\times\mathfrak{Z}_0$, respectively. Indeed, this is obvious if $A_t(x,y,\mathfrak{z})=\alpha_t\varphi(x)\phi(y)\kappa(\mathfrak{z})$ and $B_t(x,y,\mathfrak{z})=\beta_t\varphi(x)\phi(y)\kappa(\mathfrak{z})$ with $\varphi,\phi\in C_b(\mathbb{R}^d)$, bounded $\mathcal{Z}_0$-measurable function κ on $\mathfrak{Z}_0$, and bounded $\mathcal{O}$-measurable function α and bounded $\mathcal{P}$-measurable β on $\Omega\times[0,T]$. Thus our claim follows by a standard application of the monotone class lemma for functions. Hence one can easily see that our claim remains valid if we replace the boundedness condition with the existence of the integrals in (IV.5.20). Using this and taking into account (IV.5.15) and (IV.5.16) and the estimates obtained by Taylor's formula,

$$|J_s^\eta k_\varepsilon(x-y)|\leqslant N|\eta_s(y,\mathfrak{z}_0)|^2,\quad|J_s^\xi k_\varepsilon(x-y)|\leqslant N|\xi_s(y,\mathfrak{z}_1)|^2,\tag{IV.5.21}$$

$$|I_s^\xi k_\varepsilon(x-y)|^2\leqslant N|\xi_s(y,\mathfrak{z}_1)|^2\tag{IV.5.22}$$

for $x, y \in \mathbb{R}^d$, $s \in [0,T]$, $\mathfrak{z}_i \in \mathfrak{Z}_i$ and $\omega \in \Omega$, where $N = N(\varepsilon, d)$, it is not difficult to show that $(\tilde{\mathcal{L}}_t^* \mu_t)^{(\varepsilon)}(x)$, $(\mathcal{M}^{j*}\mu_t)^{(\varepsilon)}(x)$ are $\mathcal{O} \otimes \mathcal{B}(\mathbb{R}^d)$-measurable in (ω, t), $(J_t^{\eta*}\mu_t)^{(\varepsilon)}(x)$ and $(J_t^{\xi*}\mu_t)^{(\varepsilon)}(x)$ are $\mathcal{P}\otimes\mathcal{B}(\mathbb{R}^d)\otimes\mathcal{Z}_0$- and $\mathcal{P}\otimes\mathcal{B}(\mathbb{R}^d)\otimes\mathcal{Z}_1$-measurable in $(\omega, t, x, \mathfrak{z}_0)$ and $(\omega, t, x, \mathfrak{z}_1)$, respectively, and $(I_t^{\xi*}\mu_{t-})^{(\varepsilon)}(x)$ is $\mathcal{P}\otimes\mathcal{B}(\mathbb{R}^d)\otimes\mathcal{Z}_1$-measurable in $(\omega, t, x, \mathfrak{z}_1)$. Finally, integrating $(J_t^{\eta*}\mu_t)^{(\varepsilon)}(x)$ and $(J_t^{\xi*}\mu_t)^{(\varepsilon)}(x)$ over $\mathfrak{Z}_0$ and $\mathfrak{Z}_1$, respectively, by Fubini's theorem we get that f is $\mathcal{O}\otimes\mathcal{B}(\mathbb{R}^d)$-measurable. Using the estimates (IV.5.15), (IV.5.16) together with (IV.5.21) and (IV.5.22) it is easy to see that due to $\operatorname{ess\,sup}_{t\in[0,T]} |\mu_t|(\mathbb{R}^d) < \infty$ (a.s.) and (IV.5.8) the conditions (IV.5.1), (IV.5.2) hold. By Minkowski's inequality for every $x \in \mathbb{R}^d$, $t \in [0,T]$ and $\omega \in \Omega$ we have

$$|\mu_t^{(\varepsilon)}|_{L_p}^p = \int_{\mathbb{R}^d} \Big| \int_{\mathbb{R}^d} k_\varepsilon(x-y)\mu_t(dy)\Big|^p \, dx \leqslant |k_\varepsilon|_{L_p}^p |\mu_t|^p(\mathbb{R}^d) < \infty,$$

which shows that condition (IV.5.4) holds. To complete the proof of the lemma it remains to show that almost surely

$$A := \int_0^T \int_{\mathbb{R}^d} |(\tilde{\mathcal{L}}_s^*\mu_s)^{(\varepsilon)}(x)|^p \, dxds < \infty,$$

$$B := \int_0^T \int_{\mathbb{R}^d} \Big(\sum_k |(\mathcal{M}_s^{k*}\mu_s)^{(\varepsilon)}(x)|^2\Big)^{p/2} dxds < \infty,$$

$$C_\eta := \int_0^T \int_{\mathbb{R}^d} \Big| \int_{\mathfrak{Z}_0} (J_s^{\eta*}\mu_s)^{(\varepsilon)}(x,\mathfrak{z}) \, \nu_0(d\mathfrak{z})\Big|^p dxds < \infty,$$

$$C_\xi := \int_0^T \int_{\mathbb{R}^d} \Big| \int_{\mathfrak{Z}_1} (J_s^{\xi*}\mu_s)^{(\varepsilon)}(x,\mathfrak{z}) \, \nu_1(d\mathfrak{z})\Big|^p dxds < \infty,$$

$$G := \int_0^T \int_{\mathbb{R}^d} \int_{\mathfrak{Z}_1} |(I_s^{\xi*}\mu_s)^{(\varepsilon)}(x,\mathfrak{z})|^p \, \nu_1(d\mathfrak{z}) dxds < \infty,$$

$$H := \int_0^T \int_{\mathbb{R}^d} \Big(\int_{\mathfrak{Z}_1} |(I_s^{\xi*}\mu_s)^{(\varepsilon)}(x,\mathfrak{z})|^2 \, \nu_1(d\mathfrak{z})\Big)^{p/2} dxds < \infty.$$

To this end note first that with a constant $N = N(\varepsilon, d)$

$$|k_\varepsilon(x-y)| + |Dk_\varepsilon(x-y)| + |D^2k_\varepsilon(x-y)| \leqslant Nk_{2\varepsilon}(x-y) \quad \text{for all } x, y \in \mathbb{R}^d. \tag{IV.5.23}$$

Thus, using Minkowski's inequality and Assumption IV.2.1(ii), we have a constant N, depending on ε, d, K_0, K and K_1, such that almost surely

$$A \leqslant \int_0^T \Big(\int_{\mathbb{R}^d} \Big(\int_{\mathbb{R}^d} |\tilde{\mathcal{L}}_s k_\varepsilon(x-y)|^p dx\Big)^{1/p} |\mu_s|(dy)\Big)^p ds$$

$$\leqslant N|k_{2\varepsilon}|_{L_p}^p \int_0^T \Big(\int_{\mathbb{R}^d} (K_0^2 + K_1^2|y|^2 + K_1^2|Y_s|^2)|\mu_s|(dy)\Big)^p ds.$$

Hence taking into account $\operatorname{ess\,sup}_{s\in[0,T]}\|\mu_s\| < \infty$ (a.s.), (IV.5.8) (if $K_1 \neq 0$), as well as the cadlagness of $(Y_t)_{t\in[0,T]}$, we get $A < \infty$ (a.s.). In the same way we have $B < \infty$ (a.s.). By Taylor's formula and (IV.5.23) for each $x \in \mathbb{R}^d$ we have

$$|J^\eta_y k_\varepsilon(x-y)| \leqslant \int_0^1 |D^2_y k_\varepsilon|(x-y-\theta\eta(y,\mathfrak{z}))|\eta(y,\mathfrak{z})|^2\,d\theta,$$

$$\leqslant N\int_0^1 k_{2\varepsilon}(x-y-\theta\eta(y,\mathfrak{z}))\,d\theta\,|\eta(y,\mathfrak{z})|^2,$$

for all $y \in \mathbb{R}^d$, $s \in [0,T]$, $\mathfrak{z} \in \mathfrak{Z}_0$ and $\omega \in \Omega$. Here, and often later on, the variable s is suppressed, and the subscript y in J^η_y indicates that the operator J^η acts in the variable y. Hence Minkowski's inequality gives

$$\Big(\int_{\mathbb{R}^d}|J^\eta_y k_\varepsilon(x-y)|^p\,dx\Big)^{1/p} \leqslant N|k_{2\varepsilon}|_{L_p}|\eta(y,\mathfrak{z})|^2$$

with a constant $N = N(d,\varepsilon)$. Thus by the Minkowski inequality and Fubini's theorem,

$$C_\eta \leqslant \int_0^T\left(\int_{\mathfrak{Z}_0}\int_{\mathbb{R}^d}\left(\int_{\mathbb{R}^d}|J^\eta_y k_\varepsilon(x-y)|^p\,dx\right)^{1/p}|\mu_s|(dy)\,\nu_0(d\mathfrak{z})\right)^p ds$$

$$\leqslant N^p|k_{2\varepsilon}|^p_{L_p}\int_0^T\Big(\int_{\mathfrak{Z}_0}\int_{\mathbb{R}^d}|\eta_s(y,\mathfrak{z})|^2\,|\mu_s|(dy)\,\nu_0(d\mathfrak{z})\Big)^p ds$$

$$\leqslant 2^pN^p|\bar\eta|^{2p}_{L_2}|k_{2\varepsilon}|^p_{L_p}\int_0^T\Big(\int_{\mathbb{R}^d}(K_0^2+K_1^2|y|^2+K_1^2|Y_s|^2)\,|\mu_s|(dy)\Big)^p ds < \infty \text{ (a.s.)}.$$

In the same way we get $C_\xi < \infty$ (a.s.). By Taylor's formula and (IV.5.23) for each $x \in \mathbb{R}^d$ we have

$$|I^\xi_y k_\varepsilon(x-y)| \leqslant \int_0^1 |D_y k_\varepsilon|(x-y-\theta\xi(y,\mathfrak{z}))|\xi(y,\mathfrak{z})|\,d\theta,$$

$$\leqslant N\int_0^1 k_{2\varepsilon}(x-y-\theta\xi(y,\mathfrak{z}))\,d\theta\,|\xi(y,\mathfrak{z})|, \tag{IV.5.24}$$

for all $y \in \mathbb{R}^d$, $s \in [0,T]$, $\mathfrak{z} \in \mathfrak{Z}_0$ and $\omega \in \Omega$, with a constant $N = N(d,p,\varepsilon)$. Hence similarly to above we obtain

$$G \leqslant NK_\xi^{p-2}|\bar\xi|^2_{L_2}|k_{2\varepsilon}|^p_{L_p}\int_0^T\Big(\int_{\mathbb{R}^d}(K_0+K_1|y|+K_1|Y_s|)\,|\mu_s|(dy)\Big)^p ds < \infty \text{ (a.s.)}.$$

with a constant $N = N(d,p,\varepsilon)$. By Minkowski's inequality, taking into account

(IV.5.24) and Assumption IV.2.2 we have

$$
\begin{aligned}
H &\leqslant \int_0^T \left(\int_{\mathfrak{Z}_1} \left(\int_{\mathbb{R}^d} |(I_t^{\xi *}\mu_t)^{(\varepsilon)}|^p \, dx \right)^{2/p} \nu_1(d\mathfrak{z}) \right)^{p/2} dt \\
&\leqslant \int_0^T \left(\int_{\mathfrak{Z}_1} \left(\int_{\mathbb{R}^d} \left(\int_{\mathbb{R}^d} |I_t^{\xi} k_\varepsilon(x-y)|^p \, dx \right)^{1/p} |\mu_t|(dy) \right)^2 \nu_1(d\mathfrak{z}) \right)^{p/2} dt \\
&\leqslant N|\bar{\xi}|_{L_2}^p |k_{2\varepsilon}|_{L_p}^p \int_0^T \left(\int_{\mathbb{R}^d} (K_0 + K_1|y| + K_1|Y_t|) \, |\mu_t|(dy) \right)^p dt < \infty
\end{aligned}
$$

almost surely, with a constant $N = N(d, p, \varepsilon)$. □

Lemma IV.5.3. *Let Assumption IV.2.1 hold. Assume $(u_t)_{t\in[0,T]}$ is an L_p-solution of equation* (IV.3.8) *for a given $p \geqslant 2$ such that* $\operatorname{ess\,sup}_{t\in[0,T]} |u_t|_{L_1} < \infty$ *(a.s.), and if $K_1 \neq 0$ in Assumption IV.2.1 (ii), then*

$$
\operatorname*{ess\,sup}_{t\in[0,T]} \int_{\mathbb{R}^d} |y|^2 \, |u_t|(dy) < \infty \ (a.s.). \tag{IV.5.25}
$$

Then for each $x \in \mathbb{R}^d$ and $\varepsilon > 0$,

$$
\begin{aligned}
u_t^{(\varepsilon)}(x) &= u_0^{(\varepsilon)}(x) + \int_0^t (\tilde{\mathcal{L}}_s^* u_s)^{(\varepsilon)}(x) \, ds + \int_0^t (\mathcal{M}_s^{j*} u_s)^{(\varepsilon)}(x) \, dV_s^j \\
&+ \int_0^t \int_{\mathfrak{Z}_0} (J_s^{\eta *} u_s)^{(\varepsilon)}(x) \, \nu_0(d\mathfrak{z}) \, ds + \int_0^t \int_{\mathfrak{Z}_1} (J_s^{\xi *} u_s)^{(\varepsilon)}(x) \, \nu_1(d\mathfrak{z}) \, ds \\
&+ \int_0^t \int_{\mathfrak{Z}_1} (I_s^{\eta *} u_{s-})^{(\varepsilon)}(x) \, \tilde{N}_1(d\mathfrak{z}, ds)
\end{aligned} \tag{IV.5.26}
$$

holds almost surely for all $t \in [0,T]$. Moreover, for each $\varepsilon > 0$ and $p \geqslant 2$

$$
\begin{aligned}
|u_t^{(\varepsilon)}|_{L_p}^p &= |u_0^{(\varepsilon)}|_{L_p}^p + p\int_0^t \left(|u_s^{(\varepsilon)}|^{p-2} u_s^{(\varepsilon)}, (\mathcal{M}_s^{k*} u_s)^{(\varepsilon)} \right) dV_s^k + p\int_0^t \left(|u_s^{(\varepsilon)}|^{p-2} u_s^{(\varepsilon)}, (\tilde{\mathcal{L}}_s^* u_s)^{(\varepsilon)} \right) ds \\
&+ \tfrac{p(p-1)}{2} \sum_k \int_0^t \left(|u_s^{(\varepsilon)}|^{p-2}, |(\mathcal{M}_s^{k*} u_s)^{(\varepsilon)}|^2 \right) ds + p\int_0^t \int_{\mathfrak{Z}_0} \left(|u_s^{(\varepsilon)}|^{p-2} u_s^{(\varepsilon)}, (J_s^{\eta *} u_s)^{(\varepsilon)} \right) \nu_0(d\mathfrak{z}) ds \\
&+ p\int_0^t \int_{\mathfrak{Z}_1} \left(|u_s^{(\varepsilon)}|^{p-2} u_s^{(\varepsilon)}, (J_s^{\xi *} u_s)^{(\varepsilon)} \right) \nu_1(d\mathfrak{z}) ds + p\int_0^t \int_{\mathfrak{Z}_1} \left(|u_{s-}^{(\varepsilon)}|^{p-2} u_{s-}^{(\varepsilon)}, (I^{\xi *} u_{s-})^{(\varepsilon)} \right) \tilde{N}_1(d\mathfrak{z}, ds) \\
&+ \int_0^t \int_{\mathfrak{Z}_1} \int_{\mathbb{R}^d} \left\{ |u_{s-}^{(\varepsilon)} + (I_s^{\xi *} u_{s-})^{(\varepsilon)}|^p - |u_{s-}^{(\varepsilon)}|^p - p|u_{s-}^{(\varepsilon)}|^{p-2} u_{s-}^{(\varepsilon)} (I_s^{\xi *} u_{s-})^{(\varepsilon)} \right\} dx N_1(d\mathfrak{z}, ds)
\end{aligned}
$$

(IV.5.27)

holds almost surely for all $t \in [0,T]$, where u_{s-} denotes the left limit in L_p.

Proof. Notice that equations (IV.5.26) and (IV.5.27) can be formally obtained from equations (IV.5.9) and (IV.5.10), respectively, by substituting $u_t(x)dx$ and $u_{t-}(x)dx$ in place of $\mu_t(dx)$ and $\mu_{t-}(dx)$, respectively. Note, however, that $u_t(x)dx$, defines a signed measure only for $P\otimes dt$-almost every $(\omega,t) \in \Omega\times[0,T]$. Thus this lemma does not follow directly from Lemma IV.5.2. We can copy, however, the proof of Lemma IV.5.2 by replacing $\mu_t(dx)$ and $\mu_{t-}(dx)$ with $u_t(x)dx$ and $u_{t-}(x)dx$, respectively. We need also take into account that since $(u_t)_{t\in[0,T]}$ is an L_p-valued weakly cadlag process, we have have a set Ω' of full probability such that $u_{t-}(\omega) = u_t(\omega)$ for all but countably many $t \in [0,T]$, and $\sup_{t\in[0,T]} |u_t(\omega)|_{L_p} < \infty$ for $\omega \in \Omega'$. □

Lemma IV.5.4. *Let Assumptions IV.2.1, IV.2.2 and IV.2.4 hold. Let $(\mu_t)_{t\in[0,T]}$ be a measure-valued solution to* (IV.3.7). *If $K_1 \neq 0$ in Assumption IV.2.1, then assume additionally* (IV.5.8). *Then for $\varepsilon > 0$ and even integers $p \geqslant 2$ we have*

$$\mathbb{E}\sup_{t\in[0,T]} |\mu_t^{(\varepsilon)}|_{L_p}^p \leqslant N\mathbb{E}|\mu_0^{(\varepsilon)}|_{L_p}^p \tag{IV.5.28}$$

with a constant $N = N(p,d,T,K,K_\xi,K_\eta,L,\lambda,|\bar\xi|_{L_2},|\bar\eta|_{L_2})$.

Proof. We may assume that $\mathbb{E}|\mu_0^{(\varepsilon)}|_{L_p}^p < \infty$. Define

$$\mathcal{Q}_p(b,\sigma,\rho,\beta,\mu,k_\varepsilon) = p\big(|\mu^{(\varepsilon)}|^{p-2}\mu^{(\varepsilon)},(\tilde{\mathcal{L}}^*\mu)^{(\varepsilon)}\big)+\tfrac{p(p-1)}{2}\sum_k\big(|\mu^{(\varepsilon)}|^{p-2},|(\mathcal{M}^{k*}\mu)^{(\varepsilon)}|^2\big),$$

$$\mathcal{Q}_p^{(0)}(\eta(\mathfrak{z}_0),\mu,k_\varepsilon) = p\big(|\mu^{(\varepsilon)}|^{p-2}\mu^{(\varepsilon)},(J^{\eta(\mathfrak{z}_0)*}\mu)^{(\varepsilon)}\big),$$

$$\mathcal{Q}_p^{(1)}(\xi(\mathfrak{z}_1),\mu,k_\varepsilon) = p(|\mu^{(\varepsilon)}|^{p-2}\mu^{(\varepsilon)},(J^{\xi(\mathfrak{z}_1)*}\mu)^{(\varepsilon)}), \tag{IV.5.29}$$

$$\mathcal{R}_p(\xi(\mathfrak{z}_1),\mu,k_\varepsilon) = |\mu^{(\varepsilon)}+(I^{\xi(\mathfrak{z}_1)*}\mu)^{(\varepsilon)}|_{L_p}^p-|\mu^{(\varepsilon)}|_{L_p}^p-p(|\mu^{(\varepsilon)}|^{p-2}\mu^{(\varepsilon)},(I^{\xi(\mathfrak{z}_1)*}\mu)^{(\varepsilon)}),$$

for $\mu \in \mathbb{M}$, $\beta \in \mathbb{R}^{d'}$, functions b, σ and ρ on $\mathbb{R}^d$, with values in $\mathbb{R}^d$, $\mathbb{R}^{d\times d_1}$ and $\mathbb{R}^{d\times d'}$, respectively, and $\mathbb{R}^d$-valued functions $\eta(\mathfrak{z}_0)$ and $\xi(\mathfrak{z}_1)$ on $\mathbb{R}^d$ for each $\mathfrak{z}_i \in \mathfrak{Z}_i$, $i = 0,1$, where

$$\tilde{\mathcal{L}} = \tfrac12(\sigma^{il}\sigma^{jl}+\rho^{ik}\rho^{jk})D_{ij}+\beta^l\rho^{il}D_i+\beta^lB^l, \quad \mathcal{M}^k = \rho^{ik}D_i+B^k, \quad k = 1,2,...,d'. \tag{IV.5.30}$$

Recall that (f,g) denotes the integral of the product of Lebesgue measurable functions f and g over $\mathbb{R}^d$ against the Lebesgue measure on $\mathbb{R}^d$. By Lemma IV.5.2

$$d|\mu_t^{(\varepsilon)}|_{L_p}^p = \mathcal{Q}_p(b_t,\sigma_t,\rho_t,\beta_t,\mu_t,k_\varepsilon)\,dt+\int_{\mathfrak{Z}_0}\mathcal{Q}_p^{(0)}(\eta_t(\mathfrak{z}),\mu_t,k_\varepsilon)\,\nu_0(d\mathfrak{z})\,dt$$

$$+\int_{\mathfrak{Z}_1}\mathcal{Q}_p^{(1)}(\xi_t(\mathfrak{z}),\mu_t,k_\varepsilon)\,\nu_1(d\mathfrak{z})\,dt+\int_{\mathfrak{Z}_1}\mathcal{R}_p(\xi_t(\mathfrak{z}),\mu_{t-},k_\varepsilon)\,N_1(d\mathfrak{z},dt)+d\zeta_1(t)+d\zeta_2(t), \tag{IV.5.31}$$

where $\beta_t = B_t(X_t)$ and

$$\zeta_1(t) = p\int_0^t \big(|\mu_s^{(\varepsilon)}|^{p-2}\mu_s^{(\varepsilon)}, (\mathcal{M}_s^{k*}\mu_s)^{(\varepsilon)}\big)\,dV_s^k, \tag{IV.5.32}$$

$$\zeta_2(t) = p\int_0^t\int_{\mathfrak{Z}_1} \big(|\mu_s^{(\varepsilon)}|^{p-2}\mu_s^{(\varepsilon)}, (I_s^{\xi*}\mu_s)^{(\varepsilon)}\big)\,\tilde{N}_1(d\mathfrak{z},ds) \quad t\in[0,T]$$

are local martingales under P. We write

$$\int_{\mathfrak{Z}_1} \mathcal{R}_p(\xi_t(\mathfrak{z}_1),\mu_{t-},k_\varepsilon)\,N_1(d\mathfrak{z},dt) = \int_{\mathfrak{Z}_1} \mathcal{R}_p(\xi_t(\mathfrak{z}_1),\mu_{t-},k_\varepsilon)\,\nu_1(d\mathfrak{z})dt + d\zeta_3(t) \tag{IV.5.33}$$

with

$$\zeta_3(t) = \int_0^t\int_{\mathfrak{Z}_1} \mathcal{R}_p(\xi_s(\mathfrak{z}),\mu_{s-},k_\varepsilon)\,N_1(d\mathfrak{z},ds) - \int_0^t\int_{\mathfrak{Z}_1} \mathcal{R}_p(\xi_s(\mathfrak{z}),\mu_{s-},k_\varepsilon)\,\nu_1(d\mathfrak{z})ds,$$

which we can justify if we show

$$A := \int_0^T\int_{\mathfrak{Z}_1} |\mathcal{R}_p(\xi_s(\mathfrak{z}),\mu_s,k_\varepsilon)|\,\nu_1(d\mathfrak{z})\,ds < \infty \text{ (a.s.)}. \tag{IV.5.34}$$

To this end observe that by Taylor's formula

$$0 \leqslant \mathcal{R}_p(\xi_t(\mathfrak{z}),\mu_t,k_\varepsilon) \leqslant N\int_{\mathbb{R}^d} |\mu_t^{(\varepsilon)}|^{p-2}|(I_t^{\xi(\mathfrak{z})*}\mu_t)^{(\varepsilon)}|^2 + |(I_t^{\xi(\mathfrak{z})*}\mu_t)^{(\varepsilon)}|^p\,dx \tag{IV.5.35}$$

with a constant $N = N(d,p)$. Hence

$$\int_{\mathfrak{Z}_1} \mathcal{R}_p(\xi_t(\mathfrak{z}),\mu_t,k_\varepsilon)\,\nu_1(d\mathfrak{z}) \leqslant N\int_{\mathbb{R}^d} |\mu_t^{(\varepsilon)}|^{p-2}|(I_t^{\xi(\mathfrak{z})*}\mu_t)^{(\varepsilon)}|^2_{L_2(\mathfrak{Z}_1)} + |(I_t^{\xi(\mathfrak{z})*}\mu_t)^{(\varepsilon)}|^p_{L_p(\mathfrak{Z}_1)}\,dx$$

$$\leqslant N'\big(|\mu_t^{(\varepsilon)}|^p_{L_p} + A_1(t) + A_2(t)\big)$$

with

$$A_1(t) = \int_{\mathbb{R}^d} |(I_t^{\xi(\mathfrak{z})*}\mu_t)^{(\varepsilon)}|^p_{L_2(\mathfrak{Z}_1)}\,dx, \quad A_2(t) = \int_{\mathbb{R}^d} |(I_t^{\xi(\mathfrak{z})*}\mu_t)^{(\varepsilon)}|^p_{L_p(\mathfrak{Z}_1)}\,dx \tag{IV.5.36}$$

and constants N and N' depending only on d and p. By Minkowski's inequality

$$|\mu_t^{(\varepsilon)}|^p_{L_p} = \int_{\mathbb{R}^d}\Big|\int_{\mathbb{R}^d} k_\varepsilon(x-y)\,\mu_t(dy)\Big|^p dx \leqslant \Big|\int_{\mathbb{R}^d} |k_\varepsilon|_{L_p}\,\mu_t(dy)\Big|^p \leqslant |k_\varepsilon|^p_{L_p}\,\mu_t^p(\mathbf{1}), \tag{IV.5.37}$$

$$A_1(t) = \int_{\mathbb{R}^d}\Big|\int_{\mathfrak{Z}_1}\Big|\int_{\mathbb{R}^d}(k_\varepsilon(x-y-\xi_t(y,\mathfrak{z})) - k_\varepsilon(x-y))\,\mu_t(dy)\Big|^2\nu_1(d\mathfrak{z})\Big|^{p/2} dx$$

$$\leqslant \Big|\int_{\mathfrak{Z}_1}\Big|\int_{\mathbb{R}^d}(k_\varepsilon(\cdot-y-\xi_t(y,\mathfrak{z}))-k_\varepsilon(\cdot-y))\,\mu_t(dy)\Big|^2_{L_p}\nu_1(d\mathfrak{z})\Big|^{p/2}$$

$$\leqslant \Big|\int_{\mathfrak{Z}_1}\Big|\int_{\mathbb{R}^d}|Dk_\varepsilon|_{L_p}|\xi_t(y,\mathfrak{z})|\,\mu_t(dy)\Big|^2\nu_1(d\mathfrak{z})\Big|^{p/2}$$

$$\leqslant |Dk_\varepsilon|^p_{L_p}|\bar{\xi}|^p_{L_2(\mathfrak{Z}_1)}\Big(\int_{\mathbb{R}^d}(K_0+K_1|y|+K_1|Y_t|)\,\mu_t(dy)\Big)^p, \tag{IV.5.38}$$

and similarly, using Assumption IV.2.2,

$$A_2(t)=\int_{\mathbb{R}^d}\int_{\mathfrak{Z}_1}\Big|\int_{\mathbb{R}^d}(k_\varepsilon(x-y-\xi_t(y,\mathfrak{z}))-k_\varepsilon(x-y))\,\mu_t(dy)\Big|^p\nu_1(d\mathfrak{z})dx$$

$$\leqslant \int_{\mathfrak{Z}_1}\Big|\int_{\mathbb{R}^d}\big|k_\varepsilon(\cdot-y-\xi_t(y,\mathfrak{z}))-k_\varepsilon(\cdot-y))\big|_{L_p}\mu_t(dy)\Big|^p\nu_1(d\mathfrak{z})$$

$$\leqslant K_\xi^{p-2}|Dk_\varepsilon|^p_{L_p}|\bar{\xi}|^2_{L_2(\mathfrak{Z}_1)}\Big(\int_{\mathbb{R}^d}(K_0+K_1|y|+K_1|Y_t|)\,\mu_t(dy)\Big)^p. \tag{IV.5.39}$$

By (IV.5.35)–(IV.5.39) we have a constant $N=N(K_\xi,p,d,\varepsilon,|\bar{\xi}|_{L_2(\mathfrak{Z}_1)})$ such that

$$A\leqslant N\int_0^T\mu_t^p(\mathbf{1})\,dt+N\int_0^T\Big(\int_{\mathbb{R}^d}(K_0+K_1|y|+K_1|Y_t|)\,\mu_t(dy)\Big)^p dt<\infty\,\text{(a.s.)}.$$

Next we claim that, with the operator T^ξ defined in (IV.4.22),

$$\zeta_2(t)+\zeta_3(t)=\int_0^t\int_{\mathfrak{Z}_1}|(T_s^{\xi*}\mu_s)^{(\varepsilon)}|^p_{L_p}-|\mu_s^{(\varepsilon)}|^p_{L_p}\tilde{N}_1(d\mathfrak{z},ds)=:\zeta(t)\quad\text{for } t\in[0,T]. \tag{IV.5.40}$$

To see that the stochastic integral $\zeta(t)$ is well-defined as an Itô integral note that by Lemma IV.4.6 and (IV.5.37)

$$\int_0^T\int_{\mathfrak{Z}_1}||(T_s^{\xi*}\mu_s)^{(\varepsilon)}|^p_{L_p}-|\mu_s^{(\varepsilon)}|^p_{L_p}|^2\,\nu_1(d\mathfrak{z})ds\leqslant N|\bar{\xi}|^2_{L_2(\mathfrak{Z}_1)}\int_0^T|\mu_s^{(\varepsilon)}|^{2p}_{L_p}\,ds \tag{IV.5.41}$$

$$\leqslant N|\bar{\xi}|^2_{L_2(\mathfrak{Z}_1)}|k_\varepsilon|^{2p}_{L_p}\int_0^T\mu_s^{2p}(\mathbf{1})\,ds<\infty\,\text{(a.s.)}$$

with a constant $N=N(d,p,\lambda,K_\xi)$. Since $\mathfrak{Z}_1$ is σ-finite, there is an increasing sequence $(\mathfrak{Z}_{1n})_{n=1}^\infty$, $\mathfrak{Z}_{1n}\in\mathcal{Z}_1$, such that $\nu_1(\mathfrak{Z}_{1n})<\infty$ for every n and $\cup_{n=1}^\infty\mathfrak{Z}_{1n}=\mathfrak{Z}_1$. Then it is easy to see that

$$\bar{\zeta}_{2n}(t)=p\int_0^t\int_{\mathfrak{Z}_1}\mathbf{1}_{\mathfrak{Z}_{1n}}(\mathfrak{z})\big(|\mu_s^{(\varepsilon)}|^{p-2}\mu_s^{(\varepsilon)},(I_s^{\xi*}\mu_s)^{(\varepsilon)}\big)\,N_1(d\mathfrak{z},ds),$$

$$\hat{\zeta}_{2n}(t)=p\int_0^t\int_{\mathfrak{Z}_1}\mathbf{1}_{\mathfrak{Z}_{1n}}(\mathfrak{z})\big(|\mu_s^{(\varepsilon)}|^{p-2}\mu_s^{(\varepsilon)},(I_s^{\xi*}\mu_s)^{(\varepsilon)}\big)\,\nu_1(d\mathfrak{z})ds,$$

$$\bar{\zeta}_{3n}(t) = \int_0^t \int_{\mathfrak{Z}_1} \mathbf{1}_{\mathfrak{Z}_{1n}}(\mathfrak{z}) \mathcal{R}_p(\xi_s(\mathfrak{z}), \mu_{s-}, k_\varepsilon)\, N_1(d\mathfrak{z}, ds),$$

$$\hat{\zeta}_{3n}(t) = \int_0^t \int_{\mathfrak{Z}_1} \mathbf{1}_{\mathfrak{Z}_{1n}}(\mathfrak{z}) \mathcal{R}_p(\xi_s(\mathfrak{z}), \mu_{s-}, k_\varepsilon)\, \nu_1(d\mathfrak{z}) ds$$

are well-defined, and

$$\zeta_2(t) = \lim_{n\to\infty} (\bar{\zeta}_{2n}(t) - \hat{\zeta}_{2n}(t)), \quad \zeta_3(t) = \lim_{n\to\infty} \bar{\zeta}_{3n}(t) - \lim_{n\to\infty} \hat{\zeta}_{3n}(t),$$

where the limits are understood in probability. Hence

$$\zeta_2(t) + \zeta_3(t) = \lim_{n\to\infty} \Big(\bar{\zeta}_{2n}(t) + \bar{\zeta}_{3n}(t) - \big(\hat{\zeta}_{2n}(t) + \hat{\zeta}_{3n}(t)\big)\Big)$$

$$= \lim_{n\to\infty} \Big(\int_0^t \int_{\mathfrak{Z}_1} \mathbf{1}_{\mathfrak{Z}_{1n}}(\mathfrak{z}) (|(T_s^{\xi *} \mu_s)^{(\varepsilon)}|_{L_p}^p - |\mu_s^{(\varepsilon)}|_{L_p}^p) \tilde{N}_1(d\mathfrak{z}, ds)\Big) = \zeta(t),$$

which completes the proof of (IV.5.40). Consequently, from (IV.5.31)-(IV.5.33) we have

$$d|\mu_t^{(\varepsilon)}|_{L_p}^p = \mathcal{Q}_p(b_t, \sigma_t, \rho_t, \beta_t, \mu_t, k_\varepsilon)\, dt + \int_{\mathfrak{Z}_0} \mathcal{Q}_p^{(0)}(\eta_t(\mathfrak{z}_0), \mu_t, k_\varepsilon)\, \nu_0(d\mathfrak{z})\, dt$$

$$+ \int_{\mathfrak{Z}_1} \mathcal{Q}_p^{(1)}(\xi_t(\mathfrak{z}_1), \mu_t, k_\varepsilon) + \mathcal{R}_p(\xi_t(\mathfrak{z}_1), \mu_t, k_\varepsilon)\, \nu_1(d\mathfrak{z})\, dt + d\zeta_1(t) + d\zeta(t). \quad \text{(IV.5.42)}$$

By Lemma IV.4.1, Corollary IV.4.2 and Lemma IV.4.3 we have

$$Q_p(b_s, \sigma_s, \rho_s, \beta_s, \mu_s, k_\varepsilon) \leqslant N(L^2 + K^2) |\mu_s^{(\varepsilon)}|_{L_p}^p \quad \text{(IV.5.43)}$$

with a constant $N = N(d, p)$, and by Lemma IV.4.4 and Corollary IV.4.5, using that $\bar{\xi} \leqslant K_\xi$ and $\bar{\eta} \leqslant K_\eta$, we have

$$\begin{aligned} \mathcal{Q}_p^{(0)}(\eta_s(\mathfrak{z}), \mu_s, k_\varepsilon) &\leqslant N \bar{\eta}^2(\mathfrak{z}) |\mu_s^{(\varepsilon)}|_{L_p}^p, \\ (\mathcal{Q}_p^{(1)} + \mathcal{R}_p)(\xi_s(\mathfrak{z}), \mu_s, k_\varepsilon) &\leqslant N \bar{\xi}^2(\mathfrak{z}) |\mu_s^{(\varepsilon)}|_{L_p}^p \end{aligned} \quad \text{(IV.5.44)}$$

with a constant $N = N(K_\xi, K_\eta, d, p, \lambda)$. Thus from (IV.5.42) for $c_t^\varepsilon := |\mu_t^{(\varepsilon)}|_{L_p}^p$ we obtain that almost surely

$$c_t^\varepsilon \leqslant |\mu_0^{(\varepsilon)}|_{L_p}^p + N \int_0^t c_s^\varepsilon\, ds + m_t^\varepsilon \quad \text{for all } t \in [0, T] \quad \text{(IV.5.45)}$$

with a constant $N = N(T, p, d, K, K_\xi, K_\eta, L, \lambda, |\bar{\xi}|_{L_2}, |\bar{\eta}|_{L_2})$ and the local martingale $m^\varepsilon = \zeta_1 + \zeta$. For integers $n \geqslant 1$ set $\tau_n = \bar{\tau}_n \wedge \tilde{\tau}_n$, where $(\tilde{\tau}_n)_{n=1}^\infty$ is a localising sequence of stopping times for m^ε and

$$\bar\tau_n = \bar\tau_n(\varepsilon) = \inf\Big\{t \in [0,T] : \int_0^t c^\varepsilon_s\,ds \geqslant n\Big\}.$$

Then from (IV.5.45) we get

$$\mathbb{E}c^\varepsilon_{t\wedge\tau_n} \leqslant \mathbb{E}|\mu_0^{(\varepsilon)}|^p_{L_p} + N\int_0^t \mathbb{E}c^\varepsilon_{s\wedge\tau_n}\,ds < \infty \quad \text{for } t \in [0,T] \text{ and integers } n \geqslant 1.$$

Hence by Gronwall's lemma

$$\mathbb{E}c^\varepsilon_{t\wedge\tau_n} \leqslant N\mathbb{E}|\mu_0^{(\varepsilon)}|^p_{L_p} \quad \text{for } t \in [0,T] \text{ and integers } n \geqslant 1$$

with a constant $N = N(T,p,d,K,,K_\xi,K_\eta,L,\lambda,|\bar\xi|_{L_2},|\bar\eta|_{L_2})$. Letting here $n \to \infty$, by Fatou's lemma we obtain

$$\sup_{t\in[0,T]} \mathbb{E}|\mu_t^{(\varepsilon)}|^p_{L_p} \leqslant N\mathbb{E}|\mu_0^{(\varepsilon)}|^p_{L_p}. \tag{IV.5.46}$$

Hence we follow a standard way to prove (IV.5.28). Clearly, from (IV.5.45), taking into account (IV.5.46), we have a constant $N = N(T,p,d,K,K_\xi,K_\eta,L,\lambda,|\bar\xi|_{L_2},|\bar\eta|_{L_2})$ such that

$$\mathbb{E}\sup_{t\leqslant T} c^\varepsilon_{t\wedge\tau} \leqslant N\mathbb{E}|\mu_0^{(\varepsilon)}|^p_{L_p} + \mathbb{E}\sup_{t\leqslant T}|\zeta_1(t\wedge\tau)| + \mathbb{E}\sup_{t\leqslant T}|\zeta(t\wedge\tau)| \tag{IV.5.47}$$

for every stopping time τ. By estimates in Lemmas IV.4.1 and IV.4.6 for the Doob-Meyer processes $\langle\zeta_1\rangle$ and $\langle\zeta\rangle$ of ζ_1 and ζ we have

$$\begin{aligned}\langle\zeta_1\rangle(t) &= p^2\int_0^t \big|(|\mu_s^{(\varepsilon)}|^{p-2}\mu_s^{(\varepsilon)},(\mathcal{M}_s^{k*}\mu_s)^{(\varepsilon)})\big|^2\,ds \leqslant N_1\int_0^t |\mu_s^{(\varepsilon)}|^{2p}_{L_p}\,ds < \infty,\\ \langle\zeta\rangle(t) &= \int_0^t\int_{\mathfrak{Z}_1} ||(T_s^{\xi*}\mu_s)^{(\varepsilon)}|^p_{L_p} - |\mu_s^{(\varepsilon)}|^p_{L_p}|^2\nu_1(d\mathfrak{z})ds \leqslant N_2\int_0^t |\mu_s^{(\varepsilon)}|^{2p}_{L_p}\,ds < \infty\end{aligned} \tag{IV.5.48}$$

almost surely for all $t \in [0,T]$, with constants $N_1 = N_1(d,p,L)$ and $N_2 = N_2(d,p,\lambda,K_\xi,|\bar\xi|_{L_2(\mathfrak{Z}_1)})$. Using the Davis inequality, by (IV.5.48) and (IV.5.46) we get

$$\mathbb{E}\sup_{t\leqslant T}|\zeta_1(t\wedge\tau)| + \mathbb{E}\sup_{t\leqslant T}|\zeta(t\wedge\tau)| \leqslant 3\mathbb{E}\langle\zeta_1\rangle^{1/2}(t\wedge\tau) + 3\mathbb{E}\langle\zeta\rangle^{1/2}(t\wedge\tau)$$

$$\leqslant N'\mathbb{E}\Big(\int_0^T |\mu^{(\varepsilon)}_{s\wedge\tau}|^{2p}_{L_p}\,ds\Big)^{1/2} \leqslant N'\mathbb{E}\Big(\sup_{t\leqslant T}|\mu^{(\varepsilon)}_{s\wedge\tau}|^p_{L_p}\int_0^T |\mu_s^{(\varepsilon)}|^p_{L_p}\,ds\Big)^{1/2}$$

$$\leqslant \frac{1}{2}\mathbb{E}\sup_{t\leqslant T}|\mu^{(\varepsilon)}_{t\wedge\tau}|^p_{L_p} + N''\mathbb{E}\int_0^T |\mu^{(\varepsilon)}_s|^p_{L_p}\,ds \leqslant \tfrac{1}{2}\mathbb{E}\sup_{t\leqslant T}|\mu^{(\varepsilon)}_{t\wedge\tau}|^p_{L_p} + N'''\mathbb{E}|\mu^{(\varepsilon)}_0|^p_{L_p}, \tag{IV.5.49}$$

which by (IV.5.47) gives

$$\mathbb{E}\sup_{t\leqslant T} c^{\varepsilon}_{t\wedge\tau} \leqslant N\mathbb{E}|\mu^{(\varepsilon)}_0|^p_{L_p} + \tfrac{1}{2}\mathbb{E}\sup_{t\leqslant T} c^{\varepsilon}_{t\wedge\tau}$$

with constants N, N', N'', N''' depending on T, p, d, K, K_ξ, K_η, L, λ, $|\bar\xi|_{L_2}$ and $|\bar\eta|_{L_2}$ Substituting here the stopping time

$$\rho_n = \inf\{t\in[0,T] : \langle\zeta_1\rangle(t) + \langle\zeta\rangle(t) \geqslant n\}$$

in place of τ, from (IV.5.47) by virtue of the Davis inequality we have

$$\mathbb{E}\sup_{t\leqslant T} c^{\varepsilon}_{t\wedge\rho_n} \leqslant N\mathbb{E}|\mu^{(\varepsilon)}_0|^p_{L_p} + \tfrac{1}{2}\mathbb{E}\sup_{t\leqslant T} c^{\varepsilon}_{t\wedge\rho_n} < \infty$$

for every integer $n\geqslant 1$. Hence

$$\mathbb{E}\sup_{t\leqslant T\wedge\rho_n}|\mu^{(\varepsilon)}_t|^p_{L_p} \leqslant 2N\mathbb{E}|\mu^{(\varepsilon)}_0|^p_{L_p},$$

and letting here $n\to\infty$ by Fatou's lemma we finish the proof of (IV.5.28). □

Lemma IV.5.5. *Let Assumptions IV.2.1, IV.2.2 and IV.2.4 hold. Let $(u_t)_{t\in[0,T]}$ be an L_p-solution to* (IV.3.8) *for an even integer $p\geqslant 2$ such that $\operatorname{ess\,sup}_{t\in[0,T]}|u_t|_{L_1} < \infty$ (a.s.). If $K_1\neq 0$ in Assumption IV.2.1, then assume additionally* (IV.5.8). *Then we have*

$$\mathbb{E}\sup_{t\in[0,T]}|u_t|^p_{L_p} \leqslant N\mathbb{E}|u_0|^p_{L_p} \tag{IV.5.50}$$

with a constant $N = N(p,d,T,K,K_\xi,K_\eta,L,\lambda,|\bar\xi|_{L_2},|\bar\eta|_{L_2})$.

Proof. We may assume $\mathbb{E}|u_0|^p_{L_p} < \infty$. By Lemma IV.5.3 for every $\varepsilon>0$ equation (IV.5.27) holds almost surely for all $t\in[0,T]$. Hence following the proof of Lemma IV.5.4 with $u^{(\varepsilon)}_t(x)$, $u_t(x)dx$, $u_{t-}(x)dx$, $|u_t(x)|dx$ in place of $\mu^{(\varepsilon)}_t(x)$, $\mu_t(dx)$, $\mu_{t-}(dx)dx$ and $|\mu_t|(dx)$, respectively, and taking into account that almost surely $u_t = u_{t-}$ for all but countable many $t\in[0,T]$, we obtain the counterpart of (IV.5.45),

$$\begin{aligned}|u^{(\varepsilon)}_t|^p_{L_p} &\leqslant |u^{(\varepsilon)}_0|^p_{L_p} + N\int_0^t ||u_s|^{(\varepsilon)}|^p_{L_p}\,ds + m^{\varepsilon}_t\\ &\leqslant |u_0|^p_{L_p} + N\int_0^t |u_s|^p_{L_p}\,ds + m^{\varepsilon}_t \quad \text{almost surely for all } t\in[0,T]\end{aligned} \tag{IV.5.51}$$

with a constant $N = N(T,p,d,K,K_\xi,K_\eta,L,\lambda,|\bar\xi|_{L_2},|\bar\eta|_{L_2})$ and a (cadlag) local martingale $m_t^\varepsilon = \zeta_1^\varepsilon(t) + \zeta^\varepsilon(t)$, $t \in [0,T]$, where

$$\zeta_1^\varepsilon(t) = p\int_0^t \big(|u_s^{(\varepsilon)}|^{p-2}u_s^{(\varepsilon)}, (\mathcal{M}_s^{k*}u_s)^\varepsilon\big)\, dV_s^k,$$

$$\zeta^\varepsilon(t) := \int_0^t\int_{\mathfrak{Z}_1} |(T^{\xi*}u_s)^{(\varepsilon)}|_{L_p}^p - |u_s^{(\varepsilon)}|_{L_p}^p \tilde{N}_1(d\mathfrak{z}, ds).$$

Since $(u_t)_{t\in[0,T]}$ is a weakly cadlag $\mathcal{F}_t$-adapted process, we have $\sup_{t\in[0,T]}|u_t|_{L_p} < \infty$ (a.s.), and hence

$$\int_0^t |u_s|_{L_p}^r\, ds, \quad t \in [0,T]$$

is a continuous $\mathcal{F}_t$-adapted process for every $r > 0$. For $\varepsilon > 0$ and integers $n \geqslant 1$, $k \geqslant 1$ define the stopping times $\tau_{n,k}^\varepsilon := \bar{\tau}_n \wedge \tilde{\tau}_k^\varepsilon$, where

$$\bar{\tau}_n := \inf\Big\{t \in [0,T] : \int_0^t |u_s|_{L_p}^p\, ds \geqslant n\Big\}$$

for integers $n \geqslant 1$, and $(\tilde{\tau}_k^\varepsilon)_{k=1}^\infty$ is a localizing sequence for the local martingale m^ε. Thus from (IV.5.51) for $c_t^\varepsilon := |u_t^\varepsilon|_{L_p}^p$ and $c_t := |u_t|_{L_p}^p$ we get, using that for all $\varepsilon > 0$ we have $|u^{(\varepsilon)}|_{L_p} \leqslant |u|_{L_p}$,

$$\begin{aligned}
\mathbb{E}c_{t\wedge\tau_{n,k}^\varepsilon}^\varepsilon &\leqslant \mathbb{E}c_0 + N\mathbb{E}\int_0^{t\wedge\tau_{n,k}^\varepsilon} c_s\, ds \\
&\leqslant \mathbb{E}c_0 + N\mathbb{E}\int_0^{t\wedge\bar{\tau}_n} c_s\, ds \\
&\leqslant \mathbb{E}c_0 + N\mathbb{E}\int_0^t c_{s\wedge\bar{\tau}_n}\, ds < \infty
\end{aligned}$$

for every $t \in [0,T]$. Letting here first $k \to \infty$ and then $\varepsilon \to 0$ by Fatou's lemma we obtain

$$\mathbb{E}c_{t\wedge\bar{\tau}_n} \leqslant \mathbb{E}c_0 + N\mathbb{E}\int_0^t c_{s\wedge\bar{\tau}_n}\, ds < \infty, \quad t \in [0,T],$$

which by Gronwall's lemma gives

$$\mathbb{E}c_{t\wedge\bar{\tau}_n} \leqslant e^{NT}\mathbb{E}|u_0|_{L_p}^p \quad \text{for } t \in [0,T].$$

Letting now $n \to \infty$ by Fatou's lemma we have

$$\sup_{t\in[0,T]} \mathbb{E}|u_t|_{L_p}^p \leqslant e^{NT}\mathbb{E}|u_0|_{L_p}^p. \tag{IV.5.52}$$

Hence we are going to prove (IV.5.50) in an already familiar way. Analogously to (IV.5.47), due to Lemmas IV.4.1 and IV.4.6, for the Doob-Meyer processes of ζ_1^ε and ζ^ε we have with constants $N_1 = N_1(d,p,L)$ and $N_2 =$

$N_2(d,p,\lambda,K_\xi,|\bar\xi|_{L_2(\mathfrak{Z}_1)})$,

$$
\begin{aligned}
\langle\zeta_1^\varepsilon\rangle(t) &= p^2\int_0^t \big|(|u_s^{(\varepsilon)}|^{p-2}u_s^{(\varepsilon)},(\mathcal{M}_s^{k*}u_s)^{(\varepsilon)})\big|^2\,ds\\
&\leqslant N_1\int_0^t ||u_s|^{(\varepsilon)}|_{L_p}^{2p}\,ds \leqslant N_1\int_0^t |u_s|_{L_p}^{2p}\,ds,\\
\langle\zeta^\varepsilon\rangle(t) &= \int_0^t\int_{\mathfrak{Z}_1} ||(T_s^{\xi*}u_s)^{(\varepsilon)}|_{L_p}^p - |u_s^{(\varepsilon)}|_{L_p}^p|^2\nu_1(\mathfrak{z})ds\\
&\leqslant N_2\int_0^t ||u_s|^{(\varepsilon)}|_{L_p}^{2p}\,ds \leqslant N_2\int_0^t |u_s|_{L_p}^{2p}\,ds. \qquad \text{(IV.5.53)}
\end{aligned}
$$

We define the stopping time $\rho_{n,k}^\varepsilon = \tilde\tau_k^\varepsilon \wedge \rho_n$, where

$$
\rho_n = \inf\Big\{t\in[0,T] : \int_0^t |u_s|_{L_p}^{2p}\,ds \geqslant n\Big\} \quad \text{for every integer } n\geqslant 1,
$$

and $(\tilde\tau_k^\varepsilon)_{k=1}^\infty$ denotes, as before, a localizing sequence of stopping times for m^ε. Notice that from (IV.5.51), due to (IV.5.52) and (IV.5.53), by using the Davis inequality we have

$$
\begin{aligned}
\mathbb{E}\sup_{t\leqslant T} c^\varepsilon_{t\wedge\rho^\varepsilon_{n,k}} &\leqslant N'\mathbb{E}|u_0|_{L_p}^p + \mathbb{E}\sup_{t\leqslant T}|\zeta_1(t\wedge\rho_{n,k}^\varepsilon)| + \mathbb{E}\sup_{t\leqslant T}|\zeta(t\wedge\rho^\varepsilon_{n,k})|\\
&\leqslant N\mathbb{E}|u_0|_{L_p}^p + N\mathbb{E}\Big(\int_0^{T\wedge\rho_n}|u_t|_{L_p}^{2p}\,dt\Big)^{1/2} < \infty,
\end{aligned}
$$

where N' and N are constants, depending only on p, d, T, K, K_ξ, K_η L λ, $|\bar\xi|_{L_2}$ and $|\bar\eta|_{L_2}$. Letting here first $k\to\infty$ and then $\varepsilon\to 0$ by Fatou's lemma we obtain

$$
\mathbb{E}\sup_{t\leqslant T} c_{t\wedge\rho_n} \leqslant N\mathbb{E}|u_0|_{L_p}^p + N\mathbb{E}\Big(\int_0^{T\wedge\rho_n}|u_t|_{L_p}^{2p}\,dt\Big)^{1/2} < \infty \quad \text{for every } n. \qquad \text{(IV.5.54)}
$$

Hence, in the same standard way as before, by Young's inequality and (IV.5.52) we have

$$
\begin{aligned}
\mathbb{E}\sup_{t\leqslant T} c_{t\wedge\rho_n} &\leqslant N\mathbb{E}|u_0|_{L_p}^p + N\mathbb{E}\Big(\sup_{t\leqslant T} c_{t\wedge\rho_n}\int_0^T |u_t|_{L_p}^p\,dt\Big)^{1/2}\\
&\leqslant N\mathbb{E}|u_0|_{L_p}^p + \tfrac{1}{2}\mathbb{E}\sup_{t\leqslant T} c_{t\wedge\rho_n} + \tfrac{N^2}{2}\mathbb{E}\int_0^T |u_t|_{L_p}^p\,dt\\
&\leqslant N'\mathbb{E}|u_0|_{L_p}^p + \tfrac{1}{2}\mathbb{E}\sup_{t\leqslant T} c_{t\wedge\rho_n} < \infty,
\end{aligned}
$$

with a constant $N' = N'(T,p,d,K,K_\xi,K_\eta,L,\lambda,|\bar\xi|_{L_2},|\bar\eta|_{L_2})$, which gives

$$
\mathbb{E}\sup_{t\leqslant T} c_{t\wedge\rho_n} \leqslant 2N'\mathbb{E}|u_0|_{L_p}^p.
$$

Letting here $n \to \infty$ by Fatou's lemma we finish the proof of (IV.5.50). $\square$

To formulate the next lemma let $(S, \mathcal{S})$ denote a measurable space, and let $\mathcal{H} \subset \mathcal{F} \otimes \mathcal{S}$ be a σ-algebra.

Lemma IV.5.6. *Let $\mu = (\mu_s)_{s \in S}$ be an $\mathbb{M}$-valued function on $\Omega \times S$ such that $\mu_s(\varphi)$ is an $\mathcal{H}$-measurable random variable for every bounded Borel function φ on $\mathbb{R}^d$, and $\mathbb{E}\mu_s(\mathbf{1}) < \infty$ for every $s \in S$. Let $p > 1$ and assume that for a positive sequence $\varepsilon_n \to 0$ we have*

$$\limsup_{\varepsilon_n \to 0} \mathbb{E}|\mu_s^{(\varepsilon_n)}|_{L_p}^p =: N_s^p < \infty \quad \textit{for every } s \in S.$$

Then for every $s \in S$ the density $d\mu_s/dx$ exists almost surely, and there is an $L_p(\mathbb{R}^d)$-valued $\mathcal{H}$-measurable mapping u on $\Omega \times S$ such that for each s we have $u_s = d\mu_s/dx$ (a.s.). Moreover, $\lim_{n\to\infty} |\mu_s^{(\varepsilon_n)} - u_s|_{L_p} = 0$ (a.s.) and $\mathbb{E}|u_s|_{L_p}^p \leqslant N_s^p$ for each $s \in S$.

Proof. Fix $s \in S$. Since $(\mu_s^{(\varepsilon_n)})_{n=1}^\infty$ is a bounded sequence in $\mathbb{L}_p := L_p((\Omega, \mathcal{F}, P), L_p(\mathbb{R}^d))$ from any subsequence of it one can choose a subsequence, $\mu_s^{(\varepsilon_{n'})}$, which converges weakly in $\mathbb{L}_p$ to some $\bar{u}_s \in \mathbb{L}_p$. Thus for every $\varphi \in C_0^\infty(\mathbb{R}^d)$ and $G \in \mathcal{F}$ we have

$$\mathbb{E}\int_{\mathbb{R}^d} \mu_s^{(\varepsilon_{n'})}(x) \mathbf{1}_G \varphi(x)\, dx \to \mathbb{E}\int_{\mathbb{R}^d} \bar{u}_s(x)\mathbf{1}_G\varphi(x)\, dx \quad \text{as } n' \to \infty.$$

On the other hand, since

$$\mathbb{E}\int_{\mathbb{R}^d}\int_{\mathbb{R}^d} k_{\varepsilon_n}(x-y)\mathbf{1}_G|\varphi(x)|\,\mu_s(dy)dx \leqslant |\mu_s^{(\varepsilon_n)}|_{\mathbb{L}_p}|\varphi|_{L_q} < \infty \quad \text{with } q = p/(p-1),$$

we can use Fubini's theorem, and then, due to $\mathbb{E}\mu_s(\mathbf{1}) < \infty$, we can use Lebesgue's theorem on dominated convergence to get

$$\mathbb{E}\int_{\mathbb{R}^d} \mu_s^{(\varepsilon_{n'})}(x)\mathbf{1}_G\varphi(x)\, dx = \mathbb{E}\int_{\mathbb{R}^d} \mathbf{1}_G \varphi^{(\varepsilon_{n'})}(x)\,\mu_s(dx) \to \mathbb{E}\int_{\mathbb{R}^d}\mathbf{1}_G\varphi(x)\,\mu_s(dx).$$

Consequently,

$$\mathbb{E}\mathbf{1}_G\int_{\mathbb{R}^d}\varphi(x)\,\mu_s(dx) = \mathbb{E}\mathbf{1}_G\int_{\mathbb{R}^d}\varphi(x)\bar{u}_s(x)\,dx \quad \text{for any } G \in \mathcal{F} \text{ and } \varphi \in C_0^\infty, \tag{IV.5.55}$$

which implies that $d\mu_s/dx$ almost surely exists in L_p and equals $\bar{u}_s$. Notice, that $\bar{u}_s$, as an element of $\mathbb{L}_p$, is independent of the chosen subsequences, i.e., if $\tilde{u}_s$ is the weak limit in $\mathbb{L}_p$ of some subsequence of a subsequence of $\mu_s^{(\varepsilon_n)}$, then by (IV.5.55) we have

$$\mathbb{E}\mathbf{1}_G\int_{\mathbb{R}^d}\varphi(x)\bar{u}_s(x)\,dx = \mathbb{E}\mathbf{1}_G\int_{\mathbb{R}^d}\varphi(x)\tilde{u}_s(x)\,dx \quad \text{for any } G \in \mathcal{F} \text{ and } \varphi \in C_0^\infty,$$

which means $\bar{u}_s = \tilde{u}_s$ in $\mathbb{L}_p$. Consequently, the whole sequence $\mu_s^{(\varepsilon_n)}$ converges weakly to $\bar{u}_s$ in $\mathbb{L}_p$ for every s, and for each s almost surely $\bar{u}_s = d\mu_s/dx \in L_p$. Hence $\mu_s^{(\varepsilon_n)} = \bar{u}_s^{(\varepsilon_n)} \in L_p$ (a.s.), and thus by a well-known property of mollifications, $\lim_{n\to\infty} |\mu_s^{(\varepsilon_n)} - \bar{u}_s|_{L_p} = 0$ (a.s.). Set

$$A := \{(\omega, s) \in \Omega \times S : \mu_s^{(\varepsilon_n)} \text{ is convergent in } L_p \text{ as } n \to \infty\},$$

and let u_s denote the limit of $\mathbf{1}_A \mu_s^{(\varepsilon_n)}$ in L_p. Then, since $(\mu_s^{(\varepsilon_n)})_{s\in S}$ is an L_p-valued $\mathcal{H}$-measurable function of (ω, s) for every n, the function $u = (u_s)_{s\in S}$ is also an L_p-valued $\mathcal{H}$-measurable function, and clearly, $u_s = d\mu_s/dx$ (a.s.) and $\mathbb{E}|u_s|_{L_p}^p \leqslant N_s^p$ for each s. □

Lemma IV.5.7. *Let Assumptions IV.2.1, IV.2.2 and IV.2.4 hold. Let $\mu = (\mu_t)_{t\in[0,T]}$ be a measure-valued solution to* (IV.3.7). *If $K_1 \neq 0$ in Assumption IV.2.1, then assume additionally* (IV.5.8). *Assume $u_0 = d\mu_0/dx$ exists almost surely and $\mathbb{E}|u_0|_{L_p}^p < \infty$ for some even $p \geqslant 2$. Then the following statements hold.*

(i) For each $t \in [0, T]$ the density $d\mu_t/dx$ exists almost surely, and there is an L_p-valued $\mathcal{F}_t$-adapted weakly cadlag process $(u_t)_{t\in[0,T]}$ such that almost surely $u_t = d\mu_t/dx$ for every $t \in [0, T]$ and $\mathbb{E}\sup_{t\in[0,T]} |u_t|_{L_p}^p < \infty$.

(ii) If $\mu' = (\mu'_t)_{t\in[0,T]}$ satisfies the same conditions (with the same even integer p) as μ, then for $u_t = d\mu_t/dx$ and $u'_t = d\mu'_t/dx$ we have

$$\mathbb{E}\sup_{t\in[0,T]} |u_t - u'_t|_{L_p}^p \leqslant N\mathbb{E}|u_0 - u'_0|_{L_p}^p \quad \textit{for } t \in [0, T], \tag{IV.5.56}$$

with a constant N depending only on d, p, K, K_ξ, K_η, L, λ, T, $|\bar{\eta}|_{L_2(\mathfrak{Z}_1)}$ and $|\bar{\xi}|_{L_2(\mathfrak{Z}_0)}$.

Proof. By Lemma IV.5.4 we have

$$\mathbb{E}\sup_{t\in[0,T]} |\mu_t^{(\varepsilon)}|_{L_p}^p \leqslant N\mathbb{E}|\mu_0^{(\varepsilon)}|_{L_p}^p < \infty \quad \text{for every } t \in [0, T] \text{ and } \varepsilon > 0$$

with a constant $N = N(d, p, K, T, K_\xi, K_\eta, L, \lambda, |\bar{\eta}|_{L_2(\mathfrak{Z}_1)}, |\bar{\xi}|_{L_2(\mathfrak{Z}_0)})$. Moreover, by Lemma IV.5.6, there is an L_p-valued $\mathcal{F}_t$-adapted $\mathcal{F} \otimes \mathcal{B}([0, T])$-measurable process $(\bar{u}_t)_{t\in[0,T]}$ such that $\bar{u}_t = d\mu_t/dx$ (a.s.) for every $t \in [0, T]$. To prove (i) let A be a countable dense subset of $[0, T]$, such that $T \in A$. Then

$$\mathbb{E}\sup_{t\in A} |\bar{u}_t|_{L_p}^p = \mathbb{E}\sup_{t\in A} \liminf_{n\to\infty} |\mu_t^{(\varepsilon_n)}|_{L_p}^p \leqslant \mathbb{E}\liminf_{n\to\infty} \sup_{t\in A} |\mu_t^{(\varepsilon_n)}|_{L_p}^p$$

$$\leqslant \liminf_{n\to\infty} \mathbb{E}\sup_{t\in A} |\mu_t^{(\varepsilon_n)}|_{L_p}^p \leqslant N\mathbb{E}|d\mu_0/dx|_{L_p}^p < \infty \tag{IV.5.57}$$

for a sequence $\varepsilon_n \downarrow 0$, and there is a set $\Omega' \in \mathcal{F}_0$ of full probability such that

$$\sup_{t \in A} |\bar{u}_t(\omega)|_{L_p} < \infty, \quad d\mu_t/dx = \bar{u}_t \quad \text{for every } \omega \in \Omega' \text{ and } t \in A,$$

and $\mu_t(\varphi)$ is a cadlag function in $t \in [0,T]$ for $\omega \in \Omega'$ and $\varphi \in C_0^\infty(\mathbb{R}^d)$. Hence, if $t \in [0,T]$ and $\omega \in \Omega'$, then there is a sequence $t_n = t_n(\omega) \in A$ such that $t_n \downarrow t$ and $\bar{u}_{t_n}(\omega)$ converges weakly in L_p to an element, denoted by $u_t(\omega)$. Note that since $\bar{u}_{t_n}(\omega)$ is dx-everywhere nonnegative for every n, the function $u_t(\omega)$ is also dx-almost everywhere nonnegative. Moreover, by property of a weak limit we have

$$|u_t(\omega)|_{L_p} \leqslant \liminf_{n\to\infty} |\bar{u}_{t_n}(\omega)|_{L_p} \leqslant \sup_{s \in A} |\bar{u}_s(\omega)|_{L_p},$$

which gives

$$\sup_{t \in [0,T]} |u_t(\omega)|_{L_p} \leqslant \sup_{s \in A} |\bar{u}_s(\omega)|_{L_p} < \infty \quad \text{for } \omega \in \Omega'. \tag{IV.5.58}$$

Notice that

$$(u_t(\omega), \varphi) = \lim_{n\to\infty} \mu_{t_n}(\omega, \varphi) = \mu_t(\omega, \varphi) \quad \text{for } \omega \in \Omega',\ t \in [0,T] \text{ and } \varphi \in C_0^\infty, \tag{IV.5.59}$$

which shows that $u_t(\omega)$ does not depend on the sequence t_n. In particular, for $\omega \in \Omega'$ we have $\bar{u}_t(\omega) = u_t(\omega)$ for $t \in A$. Moreover, it shows that $(u_t(\omega), \varphi)$ is a cadlag function of $t \in [0,T]$ for every $\varphi \in C_0^\infty$. Hence, due to (IV.5.58), since C_0^∞ is dense in L_q, it follows that $u_t(\omega)$ is a weakly cadlag L_p-valued function of $t \in [0,T]$ for each $\omega \in \Omega'$. Moreover, from (IV.5.59), by the monotone class lemma it follows that $u_t = d\mu_t/dx$ for every $\omega \in \Omega'$ and $t \in [0,T]$. Define $u_t(\omega) = 0$ for $\omega \notin \Omega'$ and $t \in [0,T]$. Then $(u_t)_{t\in[0,T]}$ is an L_p-valued weakly cadlag function in $t \in [0,T]$ for every $\omega \in \Omega$, and since due to (IV.5.59) almost surely $(u_t, \varphi) = \mu_t(\varphi)$ for $\varphi \in C_0^\infty$, it follows that u_t is an $\mathcal{F}_t$-measurable L_p-valued random variable for every $t \in [0,T]$. Moreover, by virtue of (IV.5.57) and (IV.5.58) we have $\mathbb{E}\sup_{t\in[0,T]} |u_t|_{L_p}^p < \infty$. To prove (ii), notice that by (i) the process $\bar{u}_t := u_t - u'_t$, $t \in [0,T]$, is an L_p-solution to equation (IV.3.8) such that $\operatorname{ess\,sup}_{t\in[0,T]} |\bar{u}_t|_{L_1} < \infty$ (a.s.). Thus we have (IV.5.56) by Lemma IV.5.5. □

Definition IV.5.1. Let $p > 1$ and let ψ be an L_p-valued $\mathcal{F}_0$-measurable random variable. Then we say that an L_p-valued $\mathcal{F}_t$-optional process $v = (v_t)_{t\in[0,T]}$ is a $\mathbb{V}_p$-solution to (IV.3.8) with initial value ψ if for each $\varphi \in C_0^\infty$

$$\begin{aligned}(v_t, \varphi) =& (\psi, \varphi) + \int_0^t (v_s, \tilde{\mathcal{L}}_s\varphi)\, ds + \int_0^t (v_s, \mathcal{M}_s^k\varphi)\, dV_s^k + \int_0^t \int_{\mathfrak{Z}_0} (v_s, J_s^\eta\varphi)\, \nu_0(d\mathfrak{z})ds \\ &+ \int_0^t \int_{\mathfrak{Z}_1} (v_s, J_s^\xi\varphi)\, \nu_1(d\mathfrak{z})ds + \int_0^t \int_{\mathfrak{Z}_1} (v_s, I_s^\xi\varphi)\, \tilde{N}_1(d\mathfrak{z}, ds)\end{aligned} \tag{IV.5.60}$$

for $P \otimes dt$-a.e. $(\omega, t) \in \Omega \times [0, T]$.

Lemma IV.5.8. *Let Assumption IV.2.1 (ii) hold. Let $(v_t)_{t\in[0,T]}$ be a $\mathbb{V}_p$-solution for a $p > 1$ such that $\operatorname{ess\,sup}_{t\in[0,T]} |v_t|_{L_p} < \infty$ (a.s.), and there is an L_p-valued random variable g such that for each $\varphi \in C_0^\infty$ equation* (IV.5.60) *for $t := T$ holds almost surely with g in place of v_T. Then there exists an L_p-solution $u = (u_t)_{t\in[0,T]}$ to equation* (IV.3.8) *such that $u_0 = \psi$ and $u = v$, $P \otimes dt$-almost everywhere.*

Proof. Let $\Phi \in C_0^\infty$ be a countable dense set in L_q for $q = p/(p-1)$. Then there is a set $\Omega' \in \Omega$ of full probability and for every $\omega \in \Omega'$ there is a set $\mathbb{T}_\omega \subset [0, T]$ of full Lebesgue measure in $[0, T]$, such that $\sup_{t\in\mathbb{T}_\omega} |v_t(\omega)|_{L_p} < \infty$ for $\omega \in \Omega'$, and for all $\varphi \in \Phi$ equation (IV.5.60) holds for all $\omega \in \Omega'$ and $t \in \mathbb{T}_\omega$. We may also assume that for each $\varphi \in \Phi$ and $\omega \in \Omega'$ equation (IV.5.60) holds for $t = T$ with g in place of v_T. Since the right-hand side of equation (IV.5.60), which we denote by $F_t(\varphi)$ for short, is almost surely a cadlag function of t, we may assume, that for $\omega \in \Omega'$ it is cadlag for all $\varphi \in \Phi$. Since $\mathbb{T}_\omega$ is dense in $[0, T]$ and $\sup_{t\in\mathbb{T}_\omega} |v_t(\omega)|_{L_p} < \infty$ for $\omega \in \Omega'$, for each $\omega \in \Omega'$ and $t \in [0, T)$ we have a sequence $t_n = t_n(\omega) \in \mathbb{T}_\omega$ such that $t_n \downarrow t$ and $v_{t_n} \to \bar{v}_t$ weakly in L_p for some element $\bar{v}_t = \bar{v}_t(\omega) \in L_p$. Hence

$$(\bar{v}_t(\omega), \varphi) = \lim_{n\to\infty} (v_{t_n}(\omega), \varphi) = \lim_{n\to\infty} F_{t_n(\omega)}(\omega, \varphi) = F_t(\omega, \varphi) \quad \text{for all } \varphi \in \Phi, \tag{IV.5.61}$$

which implies that for every sequence $t_n = t_n(\omega) \in \mathbb{T}_\omega$ such that $t_n \downarrow t$ the sequence $v_{t_n(\omega)}(\omega)$ converges weakly to $\bar{v}_t(\omega)$ in L_p. In particular, $\bar{v}_t(\omega) = v_t(\omega)$ for $\omega \in \Omega'$ and $t \in \mathbb{T}_\omega$. For $\omega \in \Omega'$ we define $u_t(\omega) := \bar{v}_t(\omega)$ for $t \in [0, T)$ and $u_T(\omega) := g(\omega)$, and for $\omega \in \Omega\backslash\Omega'$ we set $u_t(\omega) = 0$ for all $t \in [0, T]$. Then due to (IV.5.61) and that almost surely $(u_T, \varphi) = F_T(\varphi)$ for all $\varphi \in \Phi$, the process $u = (u_t)_{t\in[0,T]}$ is an L_p-valued $\mathcal{F}_t$-adapted weakly cadlag process such that almost surely (IV.5.60) holds for all $\varphi \in C_0^\infty$. Clearly, $u = v$ ($P \otimes dt$-a.e.). Thus we also have that almost surely

$$\int_0^t \int_{\mathfrak{Z}_1} (u_{s-}, I_s^\xi \varphi)\, \tilde{N}_1(d\mathfrak{z}, ds) = \int_0^t \int_{\mathfrak{Z}_1} (u_s, I_s^\xi \varphi)\, \tilde{N}_1(d\mathfrak{z}, ds)$$

for all $t \in [0, T]$ and hence u satisfies (IV.5.60), with u_s replaced by u_{s-} in the last term on the right-hand side, almost surely for all $\varphi \in C_0^\infty$ for all $t \in [0, T]$, i.e., u is an L_p solution to (IV.3.8). $\square$

IV.6 Solvability of the filtering equations in L_p-spaces

To show the solvability of the linear filtering equation (IV.3.8), the Zakai equation, with any $\mathcal{F}_0$-measurable L_p-valued initial condition, we want to apply the existence and uniqueness theorem for stochastic integro-differential equations proved in [23]. With this purpose in mind first we assume that the coefficients σ, b, ρ, B, ξ, η are smooth in $x \in \mathbb{R}^d$, and under this additional assumption we are going to determine the form of the "adjoint" operators $\tilde{\mathcal{L}}^*$, $\mathcal{M}^{k*}$, $J^{\eta*}$, $J^{\xi*}$ and $I^{\xi*}$ as operators acting directly on C_0^∞ such that

$$\int_{\mathbb{R}^d} A^*\varphi(x)\phi(x)\,dx = \int_{\mathbb{R}^d} \varphi(x)A\phi(x)\,dx \quad \text{for all } \varphi, \phi \in C_0^\infty,$$

for $\tilde{\mathcal{L}}$, $\mathcal{M}$, J^ξ, J^η and I^ξ in place of A. The form of $\tilde{\mathcal{L}}^*$ and $\mathcal{M}^{k*}$ is immediately obvious by integrating by parts. To find the form of the other operators (defined in (IV.4.22)), let $\zeta : \mathbb{R}^d \to \mathbb{R}^d$ such that

$$\tau(x) := \tau^\zeta(x) := x + \zeta(x), \quad x \in \mathbb{R}^d,$$

is a C^1-diffeomorphism on $\mathbb{R}^d$. Then observe that for φ, $\phi \in C_0^\infty$ we have

$$(\phi, T^\zeta\varphi) = \int_{\mathbb{R}^d} \phi(\tau^{-1}(x))|\det D\tau^{-1}(x)|\varphi(x)\,dx = (|\det D\tau^{-1}|\,T^{\zeta^*}\phi, \varphi)$$

with

$$\zeta^*(x) := -x + \tau^{-1}(x) = -\zeta(\tau^{-1}(x)), \quad T^{\zeta^*}\phi(x) = \phi(x + \zeta^*(x)). \tag{IV.6.62}$$

Similarly,

$$\begin{aligned}
(\phi, I^\zeta\varphi) &= \int_{\mathbb{R}^d} \big(\phi(\tau^{-1}(x))|\det D\tau^{-1}(x)| - \phi(x)\big)\varphi(x)\,dx \\
&= \int_{\mathbb{R}^d} \big(\phi(\tau^{-1}(x))|\det D\tau^{-1}(x)| - \phi(\tau^{-1}(x)) + \phi(\tau^{-1}(x)) - \phi(x)\big)\varphi(x)\,dx \\
&= (\mathfrak{c}T^{\zeta^*}\phi, \varphi) + (I^{\zeta^*}\phi, \varphi),
\end{aligned}$$

where

$$\mathfrak{c}(x) = |\det D\tau^{-1}(x)| - 1,$$

and

$$\begin{aligned}
(\phi, J^\zeta\varphi) &= (\phi, I^\zeta\varphi) - (\phi, \zeta^i D_i\varphi) \\
&= (I^{\zeta^*}\phi, \varphi) + (\mathfrak{c}T^{\zeta^*}\phi, \varphi) + (\zeta^i D_i\phi, \varphi) + ((D_i\zeta^i)\phi, \varphi) \\
&= (J^{\zeta^*}\phi, \varphi) + (\mathfrak{c}I^{\zeta^*}\phi, \varphi) + ((\mathfrak{c} + D_i\zeta^i)\phi, \varphi) + ((\zeta^{*i} + \zeta^i)D_i\phi, \varphi) \\
&= (J^{\zeta^*}\phi, \varphi) + (\mathfrak{c}I^{\zeta^*}\phi, \varphi) + ((\bar{\mathfrak{c}} + D_i\zeta^{*i} + D_i\zeta^i)\phi, \varphi) + ((\zeta^{*i} + \zeta^i)D_i\phi, \varphi),
\end{aligned}$$

where

$$\bar{\mathfrak{c}} = |\det D\tau^{-1}(x)| - 1 - D_i\zeta^{*i}.$$

Consequently, $T^{\zeta *}$, $I^{\zeta *}$ and $J^{\zeta *}$, the formal adjoint of T^{ζ}, I^{ζ} and J^{ζ}, can be written in the form

$$T^{\zeta *} = |\det D\tau^{-1}|\, T^{\zeta^*},$$

$$I^{\zeta *} = I^{\zeta^*} + \mathfrak{c}T^{\zeta^*}, \quad J^{\zeta *} = J^{\zeta^*} + \mathfrak{c}I^{\zeta^*} + (\zeta^{*i} + \zeta^i)D_i + \bar{\mathfrak{c}} + D_i(\zeta^{*i} + \zeta^i). \quad \text{(IV.6.63)}$$

Lemma IV.6.1. *Let ζ be an $\mathbb{R}^d$-valued function on $\mathbb{R}^d$ such that for an integer $m \geqslant 1$ it is continuously differentiable up to order m, and*

$$\inf_{\theta\in[0,1]} \inf_{x\in\mathbb{R}^d} |\det(\mathbb{I} + \theta D\zeta(x))| =: \lambda > 0, \quad \max_{1\leqslant k\leqslant m} \sup_{x\in\mathbb{R}^d} |D^k\zeta(x)| =: M_m < \infty. \quad \text{(IV.6.64)}$$

Then the following statements hold.

(i) The function $\tau = x + \theta\zeta(x)$, $x \in \mathbb{R}^d$, is a C^m-diffeomorphism for each $\theta \in [0,1]$, such that

$$\inf_{\theta\in[0,1]} \inf_{x\in\mathbb{R}^d} |\det D\tau^{-1}(x)| \geqslant \lambda', \quad \max_{1\leqslant k\leqslant m} \sup_{x\in\mathbb{R}^d} |D^k\tau^{-1}| \leqslant M'_m < \infty, \quad \text{(IV.6.65)}$$

with constants $\lambda' = \lambda'(d, M_1) > 0$ and $M'_m = M'_m(d, \lambda, M_m)$.

(ii) The function $\zeta^*(x) = -x + \tau^{-1}(x)$, $x \in \mathbb{R}^d$, is continuously differentiable up to order m, such that

$$\sup_{\mathbb{R}^d} |\zeta^*| = \sup_{\mathbb{R}^d} |\zeta|, \quad \text{(IV.6.66)}$$

$$\sup_{\mathbb{R}^d} |D^k\zeta^*| \leqslant M^*_m \max_{1\leqslant j\leqslant k} \sup_{\mathbb{R}^d} |D^j\zeta|, \quad \text{for } k = 1, 2, ..., m, \quad \text{(IV.6.67)}$$

$$\inf_{\theta\in[0,1]} \inf_{\mathbb{R}^d} |\det(\mathbb{I} + \theta D\zeta^*)| \geqslant \lambda' \inf_{\theta\in[0,1]} \inf_{\mathbb{R}^d} |\det(\mathbb{I} + \theta D\zeta)|, \quad \text{(IV.6.68)}$$

with a constant $M^*_m = M^*_m(d, \lambda, M_m)$ and with λ' from (IV.6.65).

(iii) For the functions $\mathfrak{c} = \det(\mathbb{I} + D\zeta^*) - 1$, $\bar{\mathfrak{c}} = \mathfrak{c} - D_i\zeta^{*i}$ and $\zeta + \zeta^*$ we have

$$\sup_{x\in\mathbb{R}^d} |D^l\mathfrak{c}(x)| \leqslant N \max_{1\leqslant j\leqslant l+1} \sup_{\mathbb{R}^d} |D^j\zeta|, \quad \sup_{x\in\mathbb{R}^d} |D^k\bar{\mathfrak{c}}(x)| \leqslant N \max_{1\leqslant j\leqslant k+1} \sup_{\mathbb{R}^d} |D^j\zeta|^2 \quad \text{(IV.6.69)}$$

$$\sup_{\mathbb{R}^d} |D^k(\zeta + \zeta^*)| \leqslant N \max_{1\leqslant j\leqslant k+1} \sup_{\mathbb{R}^d} |D^j\zeta|^2. \quad \text{(IV.6.70)}$$

for $0 \leqslant l \leqslant m-1$, $1 \leqslant k \leqslant m-1$ with a constant $N = N(d, \lambda, m, M_m)$.

Proof. To prove (i) note that (IV.6.64) implies that τ is a C^m-diffeomorphism and the estimates in (IV.6.65) are proved in [42] (see Lemma 3.3 therein). From $\tau(x) = x + \zeta(x)$, by substituting $\tau^{-1}(x)$ in place of x we obtain $\zeta^*(x) =$

$-\zeta(\tau^{-1}(x))$. Hence (IV.6.66) follows immediately, and due to the second estimates in (IV.6.65), the estimate in (IV.6.67) also follows. Notice that

$$x + \theta\zeta^*(x) = \tau^{-1}(x) + \zeta(\tau^{-1}(x)) - \theta\zeta(\tau^{-1}(x)) = \tau^{-1}(x) + (1-\theta)\zeta(\tau^{-1}(x)).$$

Hence, by the first inequality in (IV.6.65),

$$\begin{aligned}|\det(\mathbb{I} + \theta D\zeta^*)| &= |\det(\mathbb{I} + (1-\theta)D\zeta(\tau^{-1}))||\det D\tau^{-1}| \\ &\geqslant \lambda' |\det(\mathbb{I} + (1-\theta)D\zeta(\tau^{-1}))|,\end{aligned}$$

which implies (IV.6.68). To prove the inequalities in (IV.6.69) notice that for the function $F(A) = \det A$, considered as the function of the entries A^{ij} of $d \times d$ real matrices A, we have

$$\frac{\partial}{\partial A^{ij}} \det A\big|_{A=\mathbb{I}} = \delta_{ij}, \quad i, j = 1, 2, ..., d.$$

Thus

$$\frac{\partial}{\partial \theta} \det(\mathbb{I} + \theta D\zeta^*)\big|_{\theta=0} = \delta_{ij} D_j \zeta^{*i} = D_i \zeta^{*i},$$

and by Taylor's formula we get

$$\mathfrak{c} = \det(\mathbb{I} + D\zeta^*) - \det \mathbb{I} = \int_0^1 \tfrac{\partial}{\partial A^{ij}} F(\mathbb{I} + \theta D\zeta^*)\, d\theta D_i \zeta^{*j}$$

and

$$\bar{\mathfrak{c}} = \det(\mathbb{I} + D\zeta^*) - \det \mathbb{I} - D_i\zeta^{*i} = \int_0^1 (1-\theta) \tfrac{\partial^2}{\partial A^{ij} \partial A^{kl}} F(\mathbb{I} + \theta D\zeta^*)\, d\theta D_i \zeta^{*j} D_k \zeta^{*l}.$$

Hence using the estimates in (IV.6.67) we get (IV.6.69). Note that

$$\zeta + \zeta^* = \zeta(\tau^{-1} - \theta\zeta^*)\big|_{\theta=0}^{\theta=1} = \zeta^{*i} \int_0^1 (D_i\zeta)(\tau^{-1} - \theta\zeta^*)\, d\theta.$$

Hence by the second estimate in (IV.6.64) and (IV.6.67) we obtain (IV.6.70). $\square$

In this section for $\varepsilon > 0$ and functions v on $\mathbb{R}^d$ we use the notation $v^{(\varepsilon)}$ for the convolution of v with $\kappa_\varepsilon(\cdot) = \varepsilon^{-d}\kappa(\cdot/\varepsilon)$, where κ is a fixed nonnegative C_0^∞ function of unit integral such that $\kappa(x) = 0$ for $|x| \geqslant 1$ and $\kappa(-x) = \kappa(x)$ for $x \in \mathbb{R}^d$.

Lemma IV.6.2. *Let τ be an $\mathbb{R}^d$-valued function on $\mathbb{R}^d$ with uniformly continuous derivative $D\tau$ on $\mathbb{R}^d$ such that with positive constants λ and K*

$$\lambda \leqslant |\det D\tau| \quad \textit{and} \quad |D\tau| \leqslant M \quad \textit{on } \mathbb{R}^d.$$

Then

$$\tfrac{1}{2}\lambda \leqslant |\det D\tau^{(\varepsilon)}| \quad \text{on } \mathbb{R}^d$$

for $\varepsilon \in (0, \varepsilon_0)$ for $\varepsilon_0 > 0$ satisfying $\delta(\varepsilon_0) \leqslant \lambda/(2d!dM^{d-1})$, where $\delta = \delta(\varepsilon)$ is the modulus of continuity of $D\tau$.

Proof. Clearly,

$$\sup_{x\in\mathbb{R}^d} |D_j\tau^i - D_j\tau^{i(\varepsilon)}| \leqslant \delta(\varepsilon) \quad \text{for } \varepsilon > 0,\ i, j = 1, 2, ..., d.$$

Hence, for $\varepsilon > 0$,

$$\sup_{x\in\mathbb{R}^d} |\Pi_{i=1}^d D_{j_i}\tau^i - \Pi_{i=1}^d D_{j_i}\tau^{i(\varepsilon)}| \leqslant \sum_{i=1}^d M^{d-1} \sup_{\mathbb{R}^d} |D_{j_i}\tau^i - D_{j_i}\tau^{i(\varepsilon)}| \leqslant dM^{d-1}\delta(\varepsilon),$$

for every permutation $(j_1, ..., j_d)$ of $1, 2, ..., d$. Therefore

$$\sup_{x\in\mathbb{R}^d} |\det D\tau - \det D\tau^{(\varepsilon)}| \leqslant d!\, dM^{d-1}\delta(\varepsilon) \quad \text{for } \varepsilon > 0.$$

Consequently, choosing $\varepsilon_0 > 0$ such that $\delta(\varepsilon_0) \leqslant \lambda/(2d!dM^{d-1})$, for $\varepsilon \in (0, \varepsilon_0)$ we have

$$|\det D\tau^{(\varepsilon)}| \geqslant |\det D\tau| - |\det D\tau - \det D\tau^{(\varepsilon)}| \geqslant \lambda/2 \quad \text{on } \mathbb{R}^d.$$

$\square$

Corollary IV.6.3. *Let ζ be an $\mathbb{R}^d$-valued function on $\mathbb{R}^d$ such that $D\zeta$ is a uniformly continuous function on $\mathbb{R}^d$ and*

$$0 < \lambda \leqslant \inf_{\mathbb{R}^d} \det(\mathbb{I} + D\zeta), \quad \sup_{\mathbb{R}^d} |D\zeta| \leqslant M < \infty \tag{IV.6.71}$$

with some positive constants λ and M. Let $\varepsilon_0 > 0$ such that $\delta(\varepsilon_0) \leqslant \lambda/(2d!dM^{d-1})$. Then for every $\varepsilon \in (0, \varepsilon_0)$ the first inequality in (IV.6.71) *holds for $\zeta^{(\varepsilon)}$ in place of ζ with $\lambda/2$ in place of λ. Moreover,* $\sup_{\mathbb{R}^d} |D^k\zeta^{(\varepsilon)}| \leqslant M_k$ *for every integer k with a constant $M_k = M_k(d, M, \varepsilon)$, where $M_1 = M$. Hence Lemma IV.6.1 holds with $\zeta^{(\varepsilon)}$ in place of ζ, for $\varepsilon \in (0, \varepsilon_0)$ for every integer $m \geqslant 1$.*

Consider for $\varepsilon \in (0, 1)$ the equation

$$\begin{aligned} du_t^\varepsilon =& \tilde{\mathcal{L}}_t^{\varepsilon *} u_t^\varepsilon \, dt + \mathcal{M}_t^{\varepsilon k *} u_t^\varepsilon \, dV_t^k + \int_{\mathfrak{Z}_0} J_t^{\eta^\varepsilon *} u_t^\varepsilon \, \nu_0(d\mathfrak{z})dt \\ &+ \int_{\mathfrak{Z}_1} J_t^{\xi^\varepsilon *} u_t^\varepsilon \, \nu_1(d\mathfrak{z})dt + \int_{\mathfrak{Z}_1} I_t^{\xi^\varepsilon *} u_t^\varepsilon \, \tilde{N}_1(d\mathfrak{z}, dt), \quad \text{with } u_0^\varepsilon = \psi^{(\varepsilon)}, \end{aligned} \tag{IV.6.72}$$

where

$$\mathcal{M}_t^{\varepsilon k} = \rho_t^{(\varepsilon)ik} D_i + B_t^{(\varepsilon)k}, \quad k = 1, \ldots, d',$$

$$\tilde{\mathcal{L}}_t^\varepsilon = a_t^{\varepsilon,ij} D_{ij} + b_t^{(\varepsilon)i} D_i + \beta_t^k \mathcal{M}_t^{\varepsilon k}, \quad \beta_t = B(t, X_t, Y_t),$$
$$a_t^{\varepsilon,ij} := \tfrac{1}{2} \sum_k (\sigma_t^{(\varepsilon)ik} \sigma_t^{(\varepsilon)jk} + \rho_t^{(\varepsilon)ik} \rho_t^{(\varepsilon)jk}), \quad i,j = 1,2,...,d,$$

the operators $J_t^{\eta^\varepsilon}$ and $J_t^{\xi^\varepsilon}$ are defined as J_t^ξ in (IV.3.1) with $\eta_t^{(\varepsilon)}$ and $\xi_t^{(\varepsilon)}$ in place of ξ_t, and the operator $I_t^{\xi^\varepsilon}$ is defined as I_t^ξ in (IV.3.1) with $\xi_t^{(\varepsilon)}$ in place of ξ_t. (Remember that $v^{(\varepsilon)}$ denotes the convolution of functions v in $x \in \mathbb{R}^d$, with the kernel κ_ε described above.) We define the L_p-solution $(u_t^\varepsilon)_{t \in [0,T]}$ to (IV.6.72) in the sense of Definition IV.3.2. Define now for each $\omega \in \Omega$, $t \geqslant 0$ and $\mathfrak{z}_i \in \mathfrak{Z}_i$ the functions

$$\tau^{\eta_t^\varepsilon}(x) = x + \eta_t^{(\varepsilon)}(x), \quad \tau^{\xi_t^\varepsilon}(x) = x + \xi_t^{(\varepsilon)}(x), \quad x \in \mathbb{R}^d, \tag{IV.6.73}$$

where, and later on, we suppress the variables $\mathfrak{z}_i$, $i = 0, 1$.

We recall that for $p \geqslant 1$, $\mathbb{L}_p$ denotes the space of L_p-valued $\mathcal{F}_0$-measurable random variables Z such that $\mathbb{E}|Z|_{L_p}^p < \infty$, as well as that for $p, q \geqslant 1$ the notation $\mathbb{L}_{p,q}$ stands for the space of L_p-valued $\mathcal{F}_t$-optional processes $v = (v_t)_{t \in [0,T]}$ such that

$$|v|_{\mathbb{L}_{p,q}}^p := \mathbb{E} \left(\int_0^T |v_t|_{L_p}^q \, dt \right)^{p/q} < \infty.$$

Let $\mathbb{B}_0$ denote the set of those functions $\psi \in \bigcap_{p \geqslant 1} \mathbb{L}_p$ such that $\psi(x) = 0$ for $|x| \geqslant R$ for some constant $R > 0$ depending on ψ and such that $\sup_{\omega \in \Omega} \sup_{x \in \mathbb{R}^d} |\psi(x)| < \infty$. It is easy to see that $\mathbb{B}_0$ is a dense subspace of $\mathbb{L}_p$ for every $p \in [1, \infty)$.

Lemma IV.6.4. *Let Assumptions IV.2.1, IV.2.2 and IV.2.4 hold with $K_1 = 0$. Assume that the following "support condition" holds: There is some $R > 0$ such that*

$$\big(b_t(x), B_t(x), \sigma_t(x), \rho_t(x), \eta_t(x, \mathfrak{z}_0), \xi_t(x, \mathfrak{z}_1)\big) = 0 \tag{IV.6.74}$$

for $\omega \in \Omega$, $t \geqslant 0$, $\mathfrak{z}_0 \in \mathfrak{Z}_0$, $\mathfrak{z}_1 \in \mathfrak{Z}_1$ and $x \in \mathbb{R}^d$ such that $|x| \geqslant R$. Let $\psi \in \mathbb{B}_0$ such that $\psi(x) = 0$ if $|x| \geqslant R$. Then there exists an $\varepsilon_0 > 0$ and a constant $\bar{R} = \bar{R}(R, K, K_0, K_\xi, K_\eta)$ such that the following statements hold for all $m \geqslant 1$ and even integers $p \geqslant 2$.

(i) For every $\varepsilon \in (0, \varepsilon_0)$ there is an L_p-solution $u^\varepsilon = (u_t^\varepsilon)_{t \in [0,T]}$ to (IV.6.72)*, which is a W_p^m-valued weakly cadlag process. Moreover, it satisfies*

$$\mathbb{E} \sup_{t \in [0,T]} |u_t^\varepsilon|_{W_p^m}^p < \infty \quad \textit{and} \quad u_t^\varepsilon(x) = 0, \quad \textit{for } dx\textit{-a.e. } x \in \{x \in \mathbb{R}^d : |x| \geqslant \bar{R}\}, \tag{IV.6.75}$$

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

(ii) There exists a unique L_p-solution $u = (u_t)_{t \in [0,T]}$ to equation (IV.3.8) *with initial condition $u_0 = \psi$, such that almost surely $u_t(x) = 0$ for dx-almost every $x \in \{x \in \mathbb{R}^d : |x| \geqslant \bar{R}\}$ for every $t \in [0, T]$ and*

$$\mathbb{E} \sup_{t \in [0,T]} |u_t|_{L_p}^p \leqslant N \mathbb{E} |\psi|_{L_p}^p \tag{IV.6.76}$$

with a constant $N = N(d, p, T, K, K_\xi, K_\eta, L, \lambda, |\bar{\xi}|_{L_2}, |\bar{\eta}|_{L_2})$.
(iii) If $(\varepsilon_n)_{n=1}^{\infty} \subset (0, \varepsilon_0)$ *such that* $\varepsilon \to 0$ *then we have*

$$u^{\varepsilon_n} \to u \quad \textit{weakly in } \mathbb{L}_{p,q}, \textit{ for every integer } q \geqslant 1.$$

Proof. To prove (i), we look for a W_p^m-valued weakly cadlag $\mathcal{F}_t$-adapted process $(u_t^\varepsilon)_{t\in[0,T]}$ such that for each $\varphi \in C_0^\infty$ almost surely

$$\begin{aligned}(u_t^\varepsilon, \varphi) = &(\psi^{(\varepsilon)}, \varphi) + \int_0^t (\tilde{\mathcal{L}}_s^{\varepsilon *} u_s^\varepsilon, \varphi)\, ds + \int_0^t (\mathcal{M}_s^{\varepsilon k *} u_s^\varepsilon, \varphi)\, dV_s^k \\ &+ \int_0^t \int_{\mathfrak{Z}_0} (J_s^{\eta^\varepsilon *} u_s^\varepsilon, \varphi)\, \nu_0(d\mathfrak{z})\, ds + \int_0^t \int_{\mathfrak{Z}_1} (J_s^{\xi^\varepsilon *} u_s^\varepsilon, \varphi)\, \nu_1(d\mathfrak{z})\, ds \\ &+ \int_0^t \int_{\mathfrak{Z}_1} (I_s^{\xi^\varepsilon *} u_s^\varepsilon, \varphi)\, \tilde{N}_1(d\mathfrak{z}, ds),\end{aligned} \tag{IV.6.77}$$

holds for all $t \in [0, T]$, where by virtue of (IV.6.63)

$$I_s^{\xi^\varepsilon *} = I^{\xi_s^{\varepsilon *}} + \mathfrak{c}^{\xi_s^\varepsilon} T^{\xi_s^{\varepsilon *}}, \quad J_s^{\xi^\varepsilon *} = J_s^{\xi^{\varepsilon *}} + \mathfrak{c}^{\xi_s^\varepsilon} I^{\xi_s^{\varepsilon *}} + (\xi_s^{\varepsilon * i} + \xi_s^{(\varepsilon) i}) D_i + \bar{\mathfrak{c}}^{\xi_t^\varepsilon} + D_i(\xi_s^{\varepsilon * i} + \xi_s^{(\varepsilon) i}),$$

$$J_s^{\eta^\varepsilon *} = J_s^{\eta^{\varepsilon *}} + \mathfrak{c}^{\eta_t^\varepsilon} I^{\eta_s^{\varepsilon *}} + (\eta_s^{\varepsilon * i} + \eta_s^{(\varepsilon) i}) D_i + \bar{\mathfrak{c}}^{\eta_s^\varepsilon} + D_i(\eta_s^{\varepsilon * i} + \eta_s^{(\varepsilon) i}), \tag{IV.6.78}$$

with the functions

$$\eta_t^{\varepsilon *}(x) = -x + (\tau^{\eta_t^\varepsilon})^{-1}(x), \quad \xi_t^{\varepsilon *}(x) = -x + (\tau^{\xi_t^\varepsilon})^{-1}(x),$$

$$\mathfrak{c}^{\xi_t^\varepsilon}(x) = |\det D(\tau^{\xi_t^\varepsilon})^{-1}(x)| - 1, \quad \mathfrak{c}^{\eta_t^\varepsilon}(x) = |\det D(\tau^{\eta_t^\varepsilon})^{-1}(x)| - 1,$$

$$\begin{aligned}\bar{\mathfrak{c}}^{\xi_t^\varepsilon}(x) &= |\det D(\tau^{\xi_t^\varepsilon})^{-1}(x)| - 1 - D_i \xi_t^{\varepsilon * i}(x), \\ \bar{\mathfrak{c}}^{\eta_t^\varepsilon}(x) &= |\det D(\tau^{\eta_t^\varepsilon})^{-1}(x)| - 1 - D_i \eta_t^{\varepsilon * i}(x) \quad x \in \mathbb{R}^d,\end{aligned} \tag{IV.6.79}$$

and clearly,

$$\mathcal{M}_s^{\varepsilon * k} \phi = -D_i(\rho_s^{(\varepsilon) ik} \phi) + B_s^{(\varepsilon) k} \phi,$$

$$\tilde{\mathcal{L}}_s^{\varepsilon *} \phi = D_{ij}(a_s^{\varepsilon, ij} \phi) - D_i(b_s^{(\varepsilon) i} \phi) - \beta_s^k D_i(\rho_s^{(\varepsilon) ik} \phi) + \beta_s^k B_s^{(\varepsilon) k} \phi \quad \text{for } \phi \in W_p^m.$$

Note that by Assumption IV.2.1(i) together with Assumption IV.2.4(i) & (ii), for each $\omega \in \Omega$, $t \in [0, T]$ and $\mathfrak{z}_i \in \mathfrak{Z}_i$, $i = 0, 1$, the mappings

$$\tau^\eta(x) = x + \eta_t(x, \mathfrak{z}_0), \quad \text{and} \quad \tau^\xi(x) = x + \xi_t(x, \mathfrak{z}_1)$$

are biLipschitz and continuously differentiable as functions of $x \in \mathbb{R}^d$. Hence, as biLipschitz functions admit Lipschitz continuous inverses, it is easy to see that for each $\omega \in \Omega$, $t \in [0, T]$ and $\mathfrak{z}_i \in \mathfrak{Z}_i$, $i = 0, 1$,

$$\lambda' \leqslant \inf_{x \in \mathbb{R}^d} |\det D\tau^\eta(x)|, \quad \text{and} \quad \lambda' \leqslant \inf_{x \in \mathbb{R}^d} |\det D\tau^\xi(x)|$$

for some $\lambda' = \lambda'(d, \lambda, L, K_\eta, K_\xi)$. Due to Assumption IV.2.4 (i) by virtue of Corollary IV.6.3 there is $\varepsilon_0 \in (0,1)$ such that for $\varepsilon \in (0, \varepsilon_0)$ the functions $\tau^{\eta_t^\varepsilon}$ and $\tau^{\xi_t^\varepsilon}$, defined in (IV.6.73), are C^∞-diffeomorphisms on $\mathbb{R}^d$ for all $\omega \in \Omega$, $t \in [0,T]$ and $\mathfrak{z}_i$, $i = 0, 1$. Moreover, the functions defined in (IV.6.79) are infinitely differentiable functions in $x \in \mathbb{R}^d$, for all $t \in [0,T]$ and $\mathfrak{z}_i \in \mathfrak{Z}_i$, $i = 0, 1$.

Hence we can easily verify that for each $\varepsilon \in (0, \varepsilon_0)$ equation (IV.6.72) satisfies the conditions of the existence and uniqueness theorem, Theorem 2.1 in [23]. Hence (IV.6.72) has a unique L_p-solution $(u_t^\varepsilon)_{t\in[0,T]}$ which is weakly cadlag as W_p^m-valued process and satisfies the first equation in (IV.6.75), for every $m \geqslant 1$. Due to the support condition (IV.6.74) and that $|\xi| \leqslant K_0 K_\xi$, $|\eta| \leqslant K_0 K_\eta$, there is a constant $\bar{R} = \bar{R}(R, K_0, K, K_\xi, K_\eta)$ such that for $\varepsilon \in (0, \varepsilon_0)$ and $s \in [0,T]$ we have

$$\tilde{\mathcal{L}}_s^\varepsilon \varphi = \mathcal{M}_s^{\varepsilon k} \varphi = I_s^{\xi^\varepsilon} \varphi = J_s^{\xi^\varepsilon} \varphi = J_s^{\eta^\varepsilon} \varphi = 0, \quad k = 1, 2, ..., d',$$

for all $\varphi \in C_0^\infty$ such that $\varphi(x) = 0$ for $|x| \leqslant \bar{R}$. Thus from equation (IV.6.77) we get that almost surely

$$(u_t^\varepsilon, \varphi) = 0 \quad \text{for all } \varphi \in C_0^\infty \text{ such that } \varphi(x) = 0 \text{ for } |x| \leqslant \bar{R}$$

for all $t \in [0,T]$, which implies

$$u_t^\varepsilon = 0 \quad \text{for } dx\text{-almost every } x \in \{x \in \mathbb{R}^d, |x| \geqslant \bar{R}\} \text{ for all } t \in [0,T] \tag{IV.6.80}$$

for each $\varepsilon \in (0, \varepsilon_0)$. To prove (ii) and (iii), note first that

$$\sup_{t\in[0,T]} |u_t^\varepsilon|_{L_1} \leqslant \bar{R}^{d(p-1)/p} \sup_{t\in[0,T]} |u_t^\varepsilon|_{L_p} < \infty \, \text{(a.s.)}.$$

It is not difficult to see that $\sigma_t^{(\varepsilon)}$, $\rho_t^{(\varepsilon)}$, $b_t^{(\varepsilon)}$ and $B_t^{(\varepsilon)}$ are bounded and Lipschitz continuous in $x \in \mathbb{R}^d$, uniformly in $\omega \in \Omega$, $t \in [0,T]$ and $\varepsilon \in (0, \varepsilon_0)$. Moreover, for $\varepsilon \in (0, \varepsilon_0)$

$$|\eta_t^{(\varepsilon)}(x, \mathfrak{z}_0)| \leqslant K_0 \bar{\xi}(\mathfrak{z}_0), \quad |\xi_t^{(\varepsilon)}(x, \mathfrak{z}_1)| \leqslant K_0 \bar{\xi}(\mathfrak{z}_1),$$

$$|\eta_t^{(\varepsilon)}(x, \mathfrak{z}_0) - \eta_t^{(\varepsilon)}(y, \mathfrak{z}_0)| \leqslant \bar{\eta}(\mathfrak{z}_0)|x - y|, \quad |\xi_t^{(\varepsilon)}(x, \mathfrak{z}_1) - \xi_t^{(\varepsilon)}(y, \mathfrak{z}_1)| \leqslant \bar{\xi}(\mathfrak{z}_1)|x - y|,$$

for all $x, y \in \mathbb{R}^d$, $\omega \in \Omega$, $t \in [0,T]$, $\mathfrak{z}_i \in \mathfrak{Z}_i$, $i = 0, 1$. Hence by Lemma IV.5.5 for $\varepsilon \in (0, \varepsilon_0)$ we have

$$\mathbb{E}|u_T^\varepsilon|_{L_p}^p + \mathbb{E}\Big(\int_0^T |u_t^\varepsilon|_{L_p}^q \, dt\Big)^{p/q} \leqslant \mathbb{E}|u_T^\varepsilon|_{L_p}^p + T^{p/q} \mathbb{E} \sup_{t\in[0,T]} |u_t^\varepsilon|_{L_p}^p \leqslant N \mathbb{E}|\psi|_{L_p}^p \tag{IV.6.81}$$

for all $q \geqslant 1$ with a constant $N = N(d, p, T, K, K_\xi, K_\eta, R, |\bar{\eta}|_{L_2}, |\bar{\xi}|_{L_2})$. By virtue of (IV.6.81) there exists a sequence $\varepsilon_n \downarrow 0$ such that u^{ε_n} converges weakly in $\mathbb{L}_{p,q}$ to some $\bar{u} \in \mathbb{L}_{p,q}$ for every integer $q > 1$ and $u_T^{\varepsilon_n}$ converges weakly to some g in $\mathbb{L}_p(\mathcal{F}_T)$, the space of L_p-valued $\mathcal{F}_T$-measurable random variables Z with the norm $(\mathbb{E}|Z|^p_{L_p})^{1/p} < \infty$. From (IV.6.81) we get

$$\mathbb{E}|g|^p_{L_p} + |\bar{u}|^p_{\mathbb{L}_{p,q}} \leqslant N\mathbb{E}|\psi|^p_{L_p} \quad \text{for every integer } q > 1 \tag{IV.6.82}$$

with the constant N in (IV.6.81). Taking ε_n in place of ε in equation (IV.6.77) then multiplying both sides of the equation with an $\mathcal{F}_t$-optional bounded process ϕ and integrating over $\Omega \times [0,T]$ against $P \otimes dt$ we obtain

$$F(u^{\varepsilon_n}) = F(\psi^{\varepsilon_n}) + \sum_{i=1}^{5} F_i^{\varepsilon_n}(u^{\varepsilon_n}), \tag{IV.6.83}$$

where F and F_i^ε are linear functionals over $\mathbb{L}_{p,q}$, defined by

$$F(v) := \mathbb{E}\int_0^T \phi_t (v_t, \varphi)\, dt, \quad F_1^\varepsilon(v) := \mathbb{E}\int_0^T \phi_t \int_0^t (v_s, \tilde{\mathcal{L}}_s^\varepsilon \varphi)\, ds\, dt,$$

$$F_2^\varepsilon(v) := \mathbb{E}\int_0^T \phi_t \int_0^t (v_s, \mathcal{M}_s^{\varepsilon k}\varphi)\, dV_s^k\, dt,$$

$$F_3^\varepsilon(v) := \mathbb{E}\int_0^T \phi_t \int_0^t \int_{\mathfrak{Z}_0} (v_s, J_s^{\eta^\varepsilon}\varphi)\, \nu_0(d\mathfrak{z}) ds\, dt,$$

$$F_4^\varepsilon(v) := \mathbb{E}\int_0^T \phi_t \int_0^t \int_{\mathfrak{Z}_1} (v_s, J_s^{\xi^\varepsilon}\varphi)\, \nu_1(d\mathfrak{z}) ds\, dt,$$

$$F_5^\varepsilon(v) := \mathbb{E}\int_0^T \phi_t \int_0^t \int_{\mathfrak{Z}_1} (v_s, I_s^{\xi^\varepsilon}\varphi)\, \tilde{N}_1(d\mathfrak{z}, ds)\, dt$$

for a fixed $\varphi \in C_0^\infty$. Define also F_i as F_i^ε for $i = 1, 2, ..., 5$, with $\tilde{\mathcal{L}}_s$, $\mathcal{M}_s^k$, J_s^η, J_s^ξ and I_s^ξ in place of $\tilde{\mathcal{L}}_s^\varepsilon$, $\mathcal{M}_s^{\varepsilon k}$, $J_s^{\eta^\varepsilon}$, $J_s^{\xi^\varepsilon}$ and $I_s^{\xi^\varepsilon}$, respectively. It is an easy exercise to show that F and F_i and F_i^ε, $i = 1, 2, 3, 4, 5$ are continuous linear functionals on $\mathbb{L}_{p,q}$ for all $q > 1$ such that

$$\lim_{\varepsilon \downarrow 0} \sup_{|v|_{\mathbb{L}_{p,q}}=1} |F_i(v) - F_i^\varepsilon(v)| = 0 \quad \text{for every } q > 1.$$

Since u^{ε_n} converges weakly to $\bar{u}$ in $\mathbb{L}_{p,q}$, and $F_i^{\varepsilon_n}$ converges strongly to F_i in $\mathbb{L}_{p,q}^*$, the dual of $\mathbb{L}_{p,q}$, we get that $F_i^{\varepsilon_n}(u^{\varepsilon_n})$ converges to $F_i(\bar{u})$ for $i = 1, 2, 3, 4, 5$. Therefore letting $\varepsilon \downarrow 0$ in (IV.6.83) we obtain

$$\mathbb{E}\int_0^T \phi_t(\bar{u}_t, \varphi)\, dt = \mathbb{E}\int_0^T \phi_t(\psi, \varphi)\, dt + \mathbb{E}\int_0^T \phi_t \int_0^t (\bar{u}_s, \tilde{\mathcal{L}}_s\varphi)\, ds\, dt$$

$$+ \mathbb{E}\int_0^T \phi_t \int_0^t (\bar{u}_s, \mathcal{M}_s^k\varphi)\, dV_s^k\, dt + \mathbb{E}\int_0^T \phi_t \int_0^t \int_{\mathfrak{Z}_0} (\bar{u}_s, J_s^\eta\varphi)\, \nu_0(d\mathfrak{z}) ds\, dt \tag{IV.6.84}$$

$$+\mathbb{E}\int_0^T \phi_t \int_0^t \int_{\mathfrak{Z}_1} (\bar{u}_s, J_s^\xi \varphi)\, \nu_1(d\mathfrak{z})ds\, dt + \mathbb{E}\int_0^T \phi_t \int_0^t \int_{\mathfrak{Z}_1} (\bar{u}_s, I_s^\xi \varphi)\, \tilde{N}_1(d\mathfrak{z}, ds)\, dt.$$

Since this equation holds for all bounded $\mathcal{F}_t$-optional processes $\phi = (\phi_t)_{t\in[0,T]}$ and functions $\varphi \in C_0^\infty$ we conclude that $\bar{u}$ is a $\mathbb{V}_p$-solution to (IV.3.8). Letting $n \to \infty$ in equation (IV.6.77) after taking ε_n in place of ε, T in place of t, multiplying both sides of the equation with an arbitrary $\mathcal{F}_T$-measurable bounded random variable ρ and taking expectation we get

$$\mathbb{E}\rho(g,\varphi) = \mathbb{E}\rho(\psi,\varphi) + \mathbb{E}\rho\int_0^T (\bar{u}_s, \tilde{\mathcal{L}}_s\varphi)\, ds$$

$$+\mathbb{E}\rho\int_0^T (\bar{u}_s, \mathcal{M}_s^k\varphi)\, dV_s^k + \mathbb{E}\rho\int_0^t \int_{\mathfrak{Z}_0} (\bar{u}_s, J_s^\eta\varphi)\, \nu_0(d\mathfrak{z})ds$$

$$+\mathbb{E}\rho\int_0^T \int_{\mathfrak{Z}_1} (\bar{u}_s, J_s^\xi\varphi)\, \nu_1(d\mathfrak{z})ds + \mathbb{E}\rho\int_0^T \int_{\mathfrak{Z}_1} (\bar{u}_s, I_s^\xi\varphi)\, \tilde{N}_1(d\mathfrak{z}, ds),$$

which implies that almost surely

$$(g,\varphi) = (\psi,\varphi) + \int_0^T (\bar{u}_s, \tilde{\mathcal{L}}_s\varphi)\, ds + \int_0^T (\bar{u}_s, \mathcal{M}_s^k\varphi)\, dV_s^k$$

$$+\int_0^T \int_{\mathfrak{Z}_0} (\bar{u}_s, J_s^\eta\varphi)\, \nu_0(d\mathfrak{z})ds + \int_0^T \int_{\mathfrak{Z}_1} (\bar{u}_s, J_s^\xi\varphi)\, \nu_1(d\mathfrak{z})ds + \int_0^T \int_{\mathfrak{Z}_1} (\bar{u}_s, I_s^\xi\varphi)\, \tilde{N}_1(d\mathfrak{z}, ds).$$

Letting $q \to \infty$ in (IV.6.82) we get

$$\mathbb{E}\operatorname*{ess\,sup}_{t\in[0,T]} |\bar{u}_t|_{L_p}^p \leqslant N\mathbb{E}|\psi|_{L_p}^p < \infty.$$

Consequently, by virtue of Lemma IV.5.8 we get the existence of a $P \otimes dt$-modification u of $\bar{u}$, which is an L_p-solution to (IV.3.8), and hence

$$\mathbb{E}\sup_{t\in[0,T]} |u_t|_{L_p}^p \leqslant N\mathbb{E}|\psi|_{L_p}^p. \tag{IV.6.85}$$

By (IV.6.80) almost surely $u_t = 0$ for dx-almost every $x \in \{x \in \mathbb{R}^d : |x| \geqslant \bar{R}\}$, for all $t \in [0,T]$, which due to (IV.6.85) by Hölder's inequality implies

$$\mathbb{E}\sup_{t\in[0,T]} |u_t|_{L_1} \leqslant N\bar{R}^{d(p-1)/p}\mathbb{E}|\psi|_{L_p} < \infty.$$

Hence by (IV.5.50) in Lemma IV.5.5 the uniqueness of the L_p-solution follows, which completes the proof of the lemma. □

Corollary IV.6.5. *Let Assumptions IV.2.1, IV.2.2 and IV.2.4 hold with* $K_1 = 0$*. Assume, moreover, that the "support condition"* (IV.6.74) *holds for some* $R > 0$*. Then for every* $p \geqslant 2$ *there is a linear operator* $\mathbb{S}$ *defined on* $\mathbb{L}_p$ *such that* $\mathbb{S}\psi$

admits a $P \otimes dt$-modification $u = (u_t)_{t\in[0,T]}$ which is an L_p-solution to equation (IV.3.8) *for every $\psi \in \mathbb{L}_p$, with initial condition $u_0 = \psi$, and*

$$\mathbb{E} \sup_{t\in[0,T]} |u_t|_{L_p}^p \leqslant N\mathbb{E}|\psi|_{L_p}^p \tag{IV.6.86}$$

with a constant $N = N(d, p, T, K, K_\xi, K_\eta, L, \lambda, |\bar{\xi}|_{L_2}, |\bar{\eta}|_{L_2})$. Moreover, if $\psi \in \mathbb{L}_p$ such that almost surely $\psi(x) = 0$ for $|x| \geqslant R$, then there is a constant $\bar{R} = \bar{R}(R, K, K_0, K_\xi, K_\eta)$ such that almost surely $u_t(x) = 0$ for dx-a.e. $x \in \{x \in \mathbb{R}^d : |x| \geqslant \bar{R}\}$ for all $t \in [0, T]$.

Proof. If p is an even integer, then the corollary follows from Lemma IV.6.4. Assume p is not an even integer. Then let p_0 be the greatest even integer such that $p_0 \leqslant p$ and let p_1 be the smallest even integer such that $p \leqslant p_1$. By Lemma IV.6.4 there are linear operators $\mathbb{S}$ and $\mathbb{S}_T$ defined on $\mathbb{B}_0$ such that $\mathbb{S}\psi := (u_t)_{t\in[0,T]}$ is the unique L_{p_i}-solution of equation (IV.3.8) with initial condition $u_0 = \psi \in \mathbb{B}_0$ and $\mathbb{S}_T\psi = u_T$. for $i = 0, 1$. Moreover, by (IV.6.76) we have

$$|\mathbb{S}_T\psi|_{\mathbb{L}_{p_i}} + |\mathbb{S}\psi|_{\mathbb{L}_{p_i,q}} \leqslant N|\psi|_{\mathbb{L}_{p_i}} \quad \text{for } i = 0, 1$$

for every $q \in [1, \infty)$ with a constant $N = N(d, p, T, K, K_\xi, K_\eta, L, \lambda, |\bar{\xi}|_{L_2}, |\bar{\eta}|_{L_2})$. Hence by a well-known generalization of the Riesz-Thorin interpolation theorem we have

$$|\mathbb{S}_T\psi|_{\mathbb{L}_p} \leqslant N|\psi|_{\mathbb{L}_p}, \quad |\mathbb{S}\psi|_{\mathbb{L}_{p,q}} \leqslant N|\psi|_{\mathbb{L}_p} \quad \text{for every } q \in [1, \infty), \tag{IV.6.87}$$

for $\psi \in \mathbb{B}_0$ with a constant $N = N(d, p, T, K, K_\xi, K_\eta, L, \lambda, |\bar{\xi}|_{L_2}, |\bar{\eta}|_{L_2})$. Assume $\psi \in \mathbb{L}_p$. Then there is a sequence $(\psi^n)_{n=1}^\infty \subset \mathbb{B}_0$ such that $\psi^n \to \psi$ in $\mathbb{L}_p$ and $u^n = \mathbb{S}\psi^n$ has a $P \otimes dt$-modification, again denoted by $u^n = (u_t^n)_{t\in[0,T]}$ which is an L_p-solution for every n with initial condition $u_0^n = \psi^n$. In particular, for each $\varphi \in C_0^\infty$ almost surely

$$\begin{aligned}(u_t^n, \varphi) =& (\psi^n, \varphi) + \int_0^t (u_s^n, \tilde{\mathcal{L}}_s\varphi)\, ds + \int_0^t (u_s^n, \mathcal{M}_s^k\varphi)\, dV_s^k + \int_0^t \int_{\mathfrak{Z}_0} (u_s^n, J_s^\eta\varphi)\, \nu_0(d\mathfrak{z})\, ds \\ &+ \int_0^t \int_{\mathfrak{Z}_1} (u_s^n, J_s^\xi\varphi)\, \nu_1(d\mathfrak{z})\, ds + \int_0^t \int_{\mathfrak{Z}_1} (u_s^n, I_s^\xi\varphi)\, \tilde{N}_1(d\mathfrak{z}, ds),\end{aligned} \tag{IV.6.88}$$

holds for all $t \in [0, T]$. By virtue of (IV.6.87) u^n converges in $\mathbb{L}_{p,q}$ to some $\bar{u} \in \mathbb{L}_{p,q}$ for every $q > 1$, and u_T^n converges in $\mathbb{L}_p$ to some $g \in \mathbb{L}_p$. Hence, letting $n \to \infty$ in equation (IV.6.88) (after multiplying both sides of it with any bounded $\mathcal{F}_t$-optional process $\phi = (\phi_t)_{t\in[0,T]}$ and integrating it over $\Omega \times [0, T]$ against $P \otimes dt$) we can see that $\bar{u}$ is a $\mathbb{V}_p$-solution such that (IV.6.87) holds. Letting $n \to \infty$ in equation (IV.6.88) with $t := T$ (after multiplying both sides with an arbitrary

$\mathcal{F}_T$-measurable bounded random variable ρ and taking expectation) we get

$$\mathbb{E}\rho(g,\varphi) = \mathbb{E}\rho(\psi,\varphi) + \mathbb{E}\rho\int_0^T (\bar{u}_s, \tilde{\mathcal{L}}_s\varphi)\,ds + \mathbb{E}\rho\int_0^T (\bar{u}_s, \mathcal{M}_s^k\varphi)\,dV_s^k$$

$$+\mathbb{E}\rho\int_0^T\int_{\mathfrak{Z}_0} (\bar{u}_s, J_s^\eta\varphi)\,\nu_0(d\mathfrak{z})\,ds + \mathbb{E}\rho\int_0^T\int_{\mathfrak{Z}_1} (\bar{u}_s, J_s^\xi\varphi)\,\nu_1(d\mathfrak{z})\,ds$$

$$+\mathbb{E}\rho\int_0^T\int_{\mathfrak{Z}_1} (\bar{u}_s, I_s^\xi\varphi)\,\tilde{N}_1(d\mathfrak{z},ds),$$

which implies

$$(g,\varphi) = (\psi,\varphi) + \int_0^T (\bar{u}_s, \tilde{\mathcal{L}}_s\varphi)\,ds + \int_0^T (\bar{u}_s, \mathcal{M}_s^k\varphi)\,dV_s^k$$

$$+\int_0^T\int_{\mathfrak{Z}_0} (\bar{u}_s, J_s^\eta\varphi)\,\nu_0(d\mathfrak{z})\,ds + \int_0^T\int_{\mathfrak{Z}_1} (\bar{u}_s, J_s^\xi\varphi)\,\nu_1(d\mathfrak{z})\,ds$$

$$+\int_0^T\int_{\mathfrak{Z}_1} (\bar{u}_s, I_s^\xi\varphi)\,\tilde{N}_1(d\mathfrak{z},ds) \quad (a.s.).$$

Letting $q \to \infty$ in (IV.6.87) we get

$$\mathbb{E}\operatorname*{ess\,sup}_{t\in[0,T]} |\bar{u}_t|_{L_p}^p \leqslant N|\psi|_{L_p}^p.$$

Hence by virtue of Lemma IV.5.8 the process $\bar{u}$ has a $P\otimes dt$ modification $u = (u_t)_{t\in[0,T]}$ which is an L_p-solution to equation (IV.3.8) and (IV.6.86) holds. Finally the last statement of the corollary about the compact support of u can be proved in the same way as it was shown for u^ε in the proof of Lemma IV.6.4. □

IV.7 Proof of Theorem IV.2.1

To prove Theorem IV.2.1 we want to show that for $p \geqslant 2$ equation (IV.3.8) has an L_p-solution which we can identify as the unnormalised conditional density of the conditional distribution of X_t given the observation $\{Y_s : s \leqslant t\}$. To this end we need some lemmas. To formulate the first one, we recall that $\mathbb{W}_p^m$ denotes the space of W_p^m-valued $\mathcal{F}_0$-measurable random variables Z such that

$$|Z|_{\mathbb{W}_p^m}^p = \mathbb{E}|Z|_{W_p^m}^p < \infty.$$

Lemma IV.7.1. *Let (X,Y) be an $\mathcal{F}_0$-measurable $\mathbb{R}^{d+d'}$-valued random variable such that the conditional density $\pi = P(X\in dx|Y)/dx$ exists. Assume $(\Omega,\mathcal{F}_0,P)$ is "rich enough" to carry an $\mathbb{R}^d$-valued random variable ζ which is independent of (X,Y) and has a smooth probability density g supported in the unit ball centred at the origin. Then there exists a sequence of $\mathcal{F}_0$-measurable random variables*

$(X_n)_{n=1}^{\infty}$ *such that the conditional density* $\pi_n = P(X_n \in dx|Y)/dx$ *exists, almost surely* $\pi_n(x) = 0$ *for* $|x| \geqslant n+1$ *for each* n,

$$\lim_{n\to\infty} X_n = X \quad \textit{for every } \omega \in \Omega,$$

and, if $\pi \in \mathbb{W}_p^m$ *for some* $p \geqslant 1$, $m \geqslant 0$, *then* $\pi_n \in \mathbb{W}_p^m$ *for every* $n \geqslant 1$, *and*

$$\lim_{n\to\infty} |\pi_n - \pi|_{\mathbb{W}_p^m} = 0.$$

Moreover, for every $n \geqslant 1$ *we have*

$$\mathbb{E}|X_n|^q \leqslant N(1 + \mathbb{E}|X|^q) \quad \textit{for every } q \in (0, \infty)$$

with a constant N *depending only on* q.

Proof. For $\varepsilon \in (0,1)$ define

$$X_k^{\varepsilon} := X\mathbf{1}_{|X|\leqslant k} + \varepsilon\zeta \quad \text{for integers } n \geqslant 1.$$

Let g_ε denote the density function of $\varepsilon\zeta$, and let μ_k be the regular conditional distribution of $Z_k := X\mathbf{1}_{|X|\leqslant k}$ given Y. Then

$$\mu_k^{(\varepsilon)}(x) = \int_{\mathbb{R}^d} g_\varepsilon(x-y)\,\mu_k(dy) \quad \text{and} \quad \pi^{(\varepsilon)}(x) = \int_{\mathbb{R}^d} g_\varepsilon(x-y)\pi(y)dy, \quad x \in \mathbb{R}^d,$$

are the conditional density functions of X_k^{ε} and $X + \varepsilon\zeta$, given Y, respectively. Clearly, if $\pi \in \mathbb{W}_p^m$, then $\mu_k^{(\varepsilon)}$ and $\pi^{(\varepsilon)}$ belong to $\mathbb{W}_p^m$ for every k and ε. Moreover, by Fubini's theorem, for each multi-index $\alpha = (\alpha_1, \ldots, \alpha_d)$, such that $0 \leqslant |\alpha| \leqslant m$ we have

$$|D^\alpha\mu_k^{(\varepsilon)} - D^\alpha\pi^{(\varepsilon)}|_{\mathbb{L}_p}^p = \mathbb{E}\int_{\mathbb{R}^d}\Big|\int_{\mathbb{R}^d} D^\alpha g_\varepsilon(x-y)\mu_k(dy) - \int_{\mathbb{R}^d} D^\alpha g_\varepsilon(x-y)\pi(y)dy\Big|^p dx \tag{IV.7.1}$$

$$= \int_{\mathbb{R}^d}\mathbb{E}\Big|\int_{\mathbb{R}^d} D^\alpha g_\varepsilon(x-y)\mu_k(dy) - \int_{\mathbb{R}^d} D^\alpha g_\varepsilon(x-y)\pi(y)dy\Big|^p dx$$

$$= \int_{\mathbb{R}^d}\mathbb{E}\big|\mathbb{E}(D^\alpha g_\varepsilon(x-Z_k) - D^\alpha g_\varepsilon(x-X)|Y)\big|^p\,dx$$

$$\leqslant \int_{\mathbb{R}^d}\mathbb{E}\big|D^\alpha g_\varepsilon(x-Z_k) - D^\alpha g_\varepsilon(x-X)\big|^p\,dx = \mathbb{E}\int_{\mathbb{R}^d}\big|D^\alpha g_\varepsilon(x-Z_k) - D^\alpha g_\varepsilon(x-X)\big|^p\,dx,$$

where the inequality is obtained by an application of Jensen's inequality. Clearly, for every $0 \leqslant |\alpha| \leqslant m$,

$$\int_{\mathbb{R}^d}\big|D^\alpha g_\varepsilon(x-Z_k) - D^\alpha g_\varepsilon(x-X)\big|^p\,dx \leqslant 2^p|g_\varepsilon|_{W_p^m}^p < \infty \quad \text{for every } \omega \in \Omega \text{ and } k \geqslant 1.$$

Hence by Lebesgue's theorem on dominated convergence, for each $0 \leqslant |\alpha| \leqslant m$,

$$\lim_{k\to\infty} \mathbb{E}\int_{\mathbb{R}^d} \big|D^\alpha g_\varepsilon(x - Z_k) - D^\alpha g_\varepsilon(x - X)\big|^p \, dx$$

$$= \mathbb{E}\lim_{k\to\infty} \int_{\mathbb{R}^d} \big|D^\alpha g_\varepsilon(x - Z_k) - D^\alpha g_\varepsilon(x - X)\big|^p \, dx = 0.$$

Consequently, by virtue of (IV.7.1) we have $\lim_{k\to\infty} |\mu_k^{(\varepsilon)} - \pi^{(\varepsilon)}|_{\mathbb{W}_p^m} = 0$ for every $\varepsilon \in (0,1)$. Since almost surely $|\pi^{(\varepsilon)} - \pi|_{W_p^m} \to 0$ as $\varepsilon \downarrow 0$, and $|\pi^{(\varepsilon)} - \pi|_{W_p^m} \leqslant 2|\pi|_{W_p^m}$ for every $\omega \in \Omega$, we have $\lim_{\varepsilon\downarrow 0} |\pi^{(\varepsilon)} - \pi|_{\mathbb{W}_p^m} = 0$ by Lebesgue's theorem on dominated convergence. Hence there is a sequence of positive integers $k_n \uparrow \infty$ such that for $\pi_n := \mu_{k_n}^{(1/n)}$ we have $\lim_{n\to\infty} |\pi_n - \pi|_{\mathbb{W}_p^m} = 0$. Clearly, for $X_n := X_{k_n}^{\varepsilon_n}$ with $\varepsilon_n = 1/n$ we have $\lim_{n\to\infty} X_n = X$ for every $\omega \in \Omega$. Moreover, for every integer $n \geqslant 1$

$$\mathbb{E}|X_n|^q \leqslant N\big(\mathbb{E}|X\mathbf{1}_{|X|\leqslant k_n}|^q + \varepsilon_n^q \mathbb{E}|\zeta|^q\big) \leqslant N(\mathbb{E}|X|^q + 1) \quad \text{for } q \in (0,\infty)$$

with a constant $N = N(q)$, which completes the proof of the lemma. □

To formulate our next lemma let χ be a smooth function on $\mathbb{R}$ such that $\chi(r) = 1$ for $r \in [-1,1]$, $\chi(r) = 0$ for $|r| \geqslant 2$, $\chi(r) \in [0,1]$ and $\chi'(r) = \frac{d}{dr}\chi(r) \in [-2,2]$ for all $r \in \mathbb{R}$.

Lemma IV.7.2. *Let $b = (b^i)$ be an $\mathbb{R}^d$-valued function on $\mathbb{R}^m$ such that for a constant L*

$$|b(v) - b(z)| \leqslant L|v - z| \quad \text{for all } v, z \in \mathbb{R}^m. \tag{IV.7.2}$$

Then for $b_n(z) = \chi(|z|/n)b(z)$, $z \in \mathbb{R}^m$, for integers $n \geqslant 1$ we have

$$|b_n(z)| \leqslant 2nL + |b(0)|, \quad |b_n(v) - b_n(z)| \leqslant (5L + 2|b(0)|)|v - z| \quad \text{for all } v, z \in \mathbb{R}^m. \tag{IV.7.3}$$

Proof. We leave the proof as an easy exercise for the reader. □

We will truncate the coefficients ξ and η of the system' (I.0.2) by the help of the following lemma, in which for each fixed $R > 0$ and $\epsilon > 0$ we use a function κ_ϵ^R defined on $\mathbb{R}^d$ by

$$\kappa_\epsilon^R(x) = \int_{\mathbb{R}^d} \phi_\epsilon^R(x - y)k(y)\, dy, \tag{IV.7.4}$$

$$\phi_\epsilon^R(x) = \begin{cases} 1, & |x| \leqslant R + 1, \\ 1 + \epsilon \log\left(\frac{R+1}{|x|}\right), & R + 1 < |x| < (R+1)e^{1/\epsilon}, \\ 0, & |x| \geqslant (R+1)e^{1/\epsilon}, \end{cases}$$

where k is a nonnegative C^∞ mapping on $\mathbb{R}^d$ with support in $\{x \in \mathbb{R}^d : |x| \leqslant 1\}$ and unit integral. Notice that $\kappa_\epsilon^R \in C_0^\infty$ for each $R, \epsilon > 0$, such that if $x, y \in \mathbb{R}^d$

and $|y| \leqslant |x|$, then

$$|\phi_\epsilon^R(x) - \phi_\epsilon^R(y)| \leqslant \frac{\epsilon|x-y|}{\max(R, |y|)},$$

and hence

$$|\kappa_\epsilon^R(x) - \kappa_\epsilon^R(y)| \leqslant \int_{\mathbb{R}^d} |\phi_\epsilon^R(x-u) - \phi_\epsilon^R(y-u)| k(u)\, du \leqslant \frac{\epsilon|x-y|}{\max(R, |y|-1)}. \quad \text{(IV.7.5)}$$

Lemma IV.7.3. *Let $\xi : \mathbb{R}^d \mapsto \mathbb{R}^d$ be such that for a constant $L \geqslant 1$ and for every $\theta \in [0,1]$ the function $\tau_\theta(x) = x + \theta\xi(x)$ is L-biLipschitz, i.e.,*

$$L^{-1}|x-y| \leqslant |\tau_\theta(x) - \tau_\theta(y)| \leqslant L|x-y| \quad \text{(IV.7.6)}$$

for all $x, y \in \mathbb{R}^d$. Then for any $M > L$ and any $R > 0$ there is an $\epsilon = \epsilon(L, M, R, |\xi(0)|) > 0$ such that with $\kappa^R := \kappa_\epsilon^R$ the function $\xi^R := \kappa^R\xi$ vanishes for $|x| \geqslant \bar{R}$ for a constant $\bar{R} = \bar{R}(L, M, R, |\xi(0)|) > R$, $|\xi^R|$ is bounded by a constant $N = N(L, M, R, |\xi(0)|)$, and for every $\theta \in [0,1]$ the mapping

$$\tau_\theta^R(x) = x + \theta\xi^R(x), \quad x \in \mathbb{R}^d$$

is M-biLipschitz.

Proof. To show τ_θ^R is M-biLipschitz, we first note that if $x, y \in \mathbb{R}^d$ with $|x| \geqslant |y|$ then $\tau_\theta^R(x) - \tau_\theta^R(y) = A + B$ where $A = \tau_{\theta\kappa^R(x)}(x) - \tau_{\theta\kappa^R(x)}(y)$ and $B = \theta\xi(y)(\kappa^R(x) - \kappa^R(y))$. The biLipschitz hypothesis (IV.7.6), with θ replaced by $\theta\kappa^R(x)$, implies $L^{-1}|x-y| \leqslant |A| \leqslant L|x-y|$. Due to (IV.7.5) and since ξ has linear growth, we can choose a sufficiently small $\epsilon = \epsilon(L, M, R, |\xi(0)|)$ to get $|B| < (L^{-1} - M^{-1})|x-y|$ and hence

$$M^{-1}|x-y| \leqslant |\tau_\theta^R(x) - \tau_\theta^R(y)| \leqslant M|x-y|$$

as required. Finally the boundedness of $|\xi^R|$ follows from the fact that it vanishes for $|x| > Re^{1/\epsilon}$ and that ξ has linear growth. ☐

Remark IV.7.1. Note that if τ is a continuously differentiable L-biLipschitz function on $\mathbb{R}^d$ then

$$L^{-d} \leqslant |\det(D\tau(x))| \leqslant L^d \quad \text{for } x \in \mathbb{R}^d.$$

Proof. This remark must be well-known, since for $d = 1$ it is obvious, and for $d > 1$ it can be easily shown by using the singular value decomposition for the matrices $D\tau$, $D\tau^{-1}$, or by applying Hadamard's inequality to their determinants. ☐

Proof of Theorem IV.2.1. The proof is structured into three steps. First we prove the theorem for the case where $p = 2$. As second step we prove the results for all $p \geqslant 2$ for compactly supported coefficients and compactly supported initial conditional densities. The third step then involves an approximation procedure to obtain the desired results for coefficients and initial conditional densities with unbounded support.

Step I: Let Assumptions IV.2.1, IV.2.2 and IV.2.4 hold. Then by Theorem III.1.1, the process $(P_t)_{t\in[0,T]}$ of the regular conditional distribution P_t of X_t given $\mathcal{F}_t^Y$, and $\mu = (\mu_t)_{t\in[0,T]} = (P_t({}^o\gamma_t)^{-1})_{t\in[0,T]}$, the "unnormalised" (regular) conditional distribution process, are measure-valued weakly cadlag processes, and μ is a measure-valued solution to equation (IV.3.3). (Note that $({}^o\gamma_t)_{t\in[0,T]}$ is the positive normalising process which we recall in IV.3.6.) Assume that $u_0 := P(X_0 \in dx|Y_0)/dx$ exists almost surely such that $\mathbb{E}|u_0|_{L_p}^p < \infty$ for $p = 2$. In order to apply Lemma IV.5.7 if $K_1 \neq 0$, we need to verify that

$$G(\mu) = \sup_{t\in[0,T]} \int_{\mathbb{R}^d} |x|^2 \, \mu_t(dx) < \infty \quad \text{almost surely.} \tag{IV.7.7}$$

For integers $k \geqslant 1$ let $\Omega_k := [|Y_0| \leqslant k] \in \mathcal{F}_0^Y$. Then $\Omega_k \uparrow \Omega$ as $k \to \infty$. Taking $r > 2$ from Assumption IV.2.3, by Doob's inequality, and by Jensen's inequality for optional projections we get

$$\mathbb{E} \sup_{t\in[0,T]} \big(\mathbb{E}(|X_t|^2 \mathbf{1}_{\Omega_k}|\mathcal{F}_t^Y)\big)^{r/2} \leqslant \mathbb{E} \sup_{t\in[0,T]} \big(\mathbb{E}(\sup_{s\in[0,T]} |X_s|^2 \mathbf{1}_{\Omega_k}|\mathcal{F}_t^Y)\big)^{r/2}$$

$$\leqslant N\mathbb{E}\big(\mathbb{E}(\sup_{s\in[0,T]} |X_s|^2 \mathbf{1}_{\Omega_k}|\mathcal{F}_T^Y)\big)^{r/2} \leqslant N\mathbb{E} \sup_{s\in[0,T]} |X_s|^r \mathbf{1}_{\Omega_k},$$

for all k with a constant N depending only on r. Thus, by Fubini's theorem and Hölder's inequality, if $K_1 \neq 0$, for all k we have

$$G_k(\mu) := \mathbb{E} \sup_{t\in[0,T]} \int_{\mathbb{R}^d} |x|^2 \mu_t(dx) \mathbf{1}_{\Omega_k}$$

$$= \mathbb{E} \sup_{t\in[0,T]} \mathbb{E}(|X_t|^2|\mathcal{F}_t^Y)({}^o\gamma_t)^{-1}\mathbf{1}_{\Omega_k} = \mathbb{E} \sup_{t\in[0,T]} \mathbb{E}(|X_t|^2 \mathbf{1}_{\Omega_k}|\mathcal{F}_t^Y)({}^o\gamma)_t^{-1}$$

$$\leqslant \mathbb{E} \sup_{t\in[0,T]} \mathbb{E}(|X_t|^2 \mathbf{1}_{\Omega_k}|\mathcal{F}_t^Y) \sup_{t\in[0,T]} ({}^o\gamma)_t^{-1}$$

$$\leqslant \big(\mathbb{E} \sup_{t\in[0,T]} \big(\mathbb{E}(|X_t|^2 \mathbf{1}_{\Omega_k}|\mathcal{F}_t^Y)\big)^{r/2}\big)^{2/r} \big(\mathbb{E} \sup_{t\in[0,T]} ({}^o\gamma_t)^{-r'}\big)^{1/r'}$$

$$\leqslant N\big(\mathbb{E} \sup_{t\in[0,T]} |X_t|^r \mathbf{1}_{\Omega_k}\big)^{2/r},$$

where $2/r + 1/r' = 1$, $N = N(r, d, C)$ is a constant, and we use that by Jensen's inequality for optional projections and the boundedness of $|B|$

$$\mathbb{E} \sup_{t\in[0,T]} ({}^o\gamma_t)^{-r'} \leqslant \mathbb{E} \sup_{t\in[0,T]} \gamma_t^{-r'} := C < \infty \tag{IV.7.8}$$

with a constant C only depending on the bound in magnitude of $|B|$ and r.

Hence, using (IV.2.1) with $q = r$ we have

$$G_k(\mu) \leqslant N\big(1 + \mathbb{E}\sup_{t\in[0,T]}|X_t|^r\mathbf{1}_{\Omega_k}\big) \leqslant N'\big(k^r + \mathbb{E}|X_0|^r\big) < \infty,$$

for constant $N = N(r,d,C)$ and $N' = N'(d,d',r,K,K_0,K_1,K_\xi,K_\eta,T,|\bar\xi|_{L_2},|\bar\eta|_{L_2})$. Since for all $k \geqslant 1$ we have that $G_k(\mu) < \infty$ we can conclude that (IV.7.7) holds. Hence, by Lemma IV.5.7, almost surely $d\mu_t/dx$ exists, and there is an L_2-valued weakly cadlag stochastic process $(u_t)_{t\in[0,T]}$ such that almost surely $u_t = d\mu_t/dx$ for all $t \in [0,T]$ and

$$\mathbb{E}\sup_{t\in[0,T]}|u_t|^2_{L_2} \leqslant N\mathbb{E}|\pi_0|^2_{L_2} \tag{IV.7.9}$$

for every T with a constant $N = N(d,d',K,K_\xi,K_\eta,L,T,|\bar\xi|_{L_2},|\bar\eta|_{L_2},\lambda)$. Thus $\pi_t = dP_t/dx = u_t{}^o\gamma_t$, $t \in [0,T]$, is an L_2-valued weakly cadlag process, which proves Theorem IV.2.1 for $p = 2$.

Step II. Let the assumptions of Theorem IV.2.1 hold with $K_1 = 0$ in Assumption IV.2.1. Assume that $\pi_0 = P(X_0 \in dx|Y_0)/dx \in \mathbb{L}_p$ for some $p > 2$, such that almost surely $u_0(x) = 0$ for $|x| \geqslant R$ for a constant R. Assume moreover, that the support condition (IV.6.74) holds. Then by Corollary IV.6.5 there is an L_p-solution $(v_t)_{t\in[0,T]}$ to (IV.3.8) with initial condition $v_0 = \pi_0$ such that

$$\mathbb{E}\sup_{t\in[0,T]}|v_t|^p_{L_p} \leqslant N\mathbb{E}|\psi|^p_{L_p} \tag{IV.7.10}$$

with a constant $N = N(d,d',K,L,K_\xi,K_\eta,T,p,\lambda,|\bar\xi|_{L_2},|\bar\eta|_{L_2})$, and almost surely

$$v_t(x) = 0 \quad \text{for } dx\text{-a.e. } x \in \{x \in \mathbb{R}^d : |x| \geqslant \bar R\} \text{ for all } t \in [0,T]$$

with a constant $\bar R = \bar R(R,K,K_0,K_\xi,K_\eta)$. Hence $(v_t)_{t\in[0,T]}$ is also an L_2-solution to equation (IV.3.8), and clearly,

$$\sup_{t\in[0,T]}|v_t|_{L_1} \leqslant \bar R^{d(p-1)/p}\sup_{t\in[0,T]}|v_t|_{L_p} < \infty.$$

Since in particular $\mathbb{E}|\pi_0|^2_{L_2} < \infty$, by Step I there is an L_2-solution $(u_t)_{t\in[0,T]}$ to equation (IV.3.8) such that almost surely $u_t = d\mu_t/dx$ for all $t \in [0,T]$, where $\mu_t = P_t({}^o\gamma_t)^{-1}$ is the unnormalised (regular) conditional distribution of X_t given $\mathcal{F}^Y_t$. Clearly,

$$\sup_{t\in[0,T]}|u_t|_{L_1} = \sup_{t\in[0,T]}({}^o\gamma_t)^{-1} < \infty \text{ (a.s.)}.$$

Hence by virtue of (IV.5.50) in Lemma IV.5.5 we obtain $\sup_{t\in[0,T]}|u_t - v_t|_{L_2} = 0$ (a.s.), which completes the proof of Theorem IV.2.1 under the additional assumptions of Step II.

Step III. Finally, we dispense with the assumption that the coefficients and the initial condition are compactly supported, and that $K_1 = 0$ in Assumption IV.2.1. Define the functions $b^n = (b^{ni}(t,z))$, $B^n = (B^{nj}(t,z))$, $\sigma^n = (\sigma^{nij}(t,z))$,

$\eta^n = (\eta^{ni}(t, z, \mathfrak{z}_0))$ and $\xi^n = (\xi^{ni}(t, z, \mathfrak{z}_1))$ by

$$(b^n, B^n, \sigma^n, \rho^n) = (b, B, \sigma, \rho)\chi_n, \quad (\eta^n, \xi^n) = (\eta, \xi)\bar{\chi}_n$$

for every integer $n \geqslant 1$, where $\chi_n(z) = \chi(|z|/n)$ and $\bar{\chi}_n(x, y) = \kappa^n(x)\chi(|y|/n)$, with χ defined before Lemma IV.7.2 and with κ^n stemming from Lemma IV.7.3 applied to ξ and η as functions of $x \in \mathbb{R}^d$. By Lemma IV.7.2, Assumptions IV.2.1 and IV.2.2 hold for b^n, B^n, σ^n, ρ^n, η^n and ξ^n, in place of b, σ, ρ, η and ξ, respectively, with $K_1 = 0$ and with appropriate constants $K_0' = K_0'(n, K, K_0, K_1, K_\eta, K_\xi, L))$ and $L' = L'(K, K_0, K_1, L, K_\xi, K_\eta)$ in place of K_0 and L. Moreover, by Lemma IV.7.3, Assumption IV.2.4 is satisfied with a constant $\lambda' = \lambda'(K_0, K_1, K_\xi, K_\eta, \lambda)$ in place of λ. Since $\pi_0 = P(X_0 \in dx|Y_0)/dx \in \mathbb{L}_p$ for $p > 2$ by assumption (the case $p = 2$ was proved in Step I) and clearly $\pi_0 \in \mathbb{L}_1$, by Hölder's inequality we have

$$|\pi_0|_{\mathbb{L}_2} \leqslant |\pi_0|_{\mathbb{L}_1}^{1-\theta} |\pi_0|_{\mathbb{L}_p}^{\theta} < \infty \quad \text{with } \theta = \tfrac{p}{2(p-1)} \in (0, 1).$$

Thus by Lemma IV.7.1 there exists a sequence $(X_0^n)_{n=1}^\infty$ of $\mathcal{F}_0$-measurable random variables such that the conditional density $\pi_0^n = P(X_0^n \in dx|\mathcal{F}_0^Y)/dx$ exists, $\pi_0^n(x) = 0$ for $|x| \geqslant n + 1$ for every n, $\lim_{n\to\infty} X_0^n = X_0$ for every $\omega \in \Omega$,

$$\lim_{n\to\infty} |\pi_0^n - \pi_0|_{\mathbb{L}_r} = 0 \quad \text{for } r = 2, p, \tag{IV.7.11}$$

and

$$\mathbb{E}|X_0^n|^q \leqslant N(1 + \mathbb{E}|X_0|^q) \quad \text{for any } q > 0$$

with a constant $N = N(q)$. Let $(X_t^n, Y_t^n)_{t\in[0,T]}$ denote the solution of equation (I.0.2) with initial value (X_0^n, Y_0) and with b^n, σ^n, ρ^n, ξ^n, η^n and B^n in place of b, σ, ρ, ξ, η and B. Define the random fields,

$$b_t^n(x) = b^n(t, x, Y_{t-}^n), \quad \sigma_t^n(x) = \sigma^n(t, x, Y_{t-}^n),$$

$$\rho_t^n(x) = \sigma^n(t, x, Y_{t-}^n), \quad B_t^n(x) = B^n(t, x, Y_{t-}^n)$$

$$\eta_t^n(x, \mathfrak{z}_0) = \eta^n(t, x, Y_{t-}^n, \mathfrak{z}_0), \quad \xi_t^n(x, \mathfrak{z}_1) = \xi^n(t, x, Y_{t-}^n, \mathfrak{z}_1), \quad \beta_t^n = B^n(t, X_t^n, Y_{t-}^n) \tag{IV.7.12}$$

for $\omega \in \Omega$, $t \geqslant 0$, $x \in \mathbb{R}^d$, $\mathfrak{z}_i \in \mathfrak{Z}_i$, $i = 0, 1$. Consider the equation

$$\begin{aligned} du_t^n =& \tilde{\mathcal{L}}_t^{n*} u_t^n \, dt + \mathcal{M}_t^{nk*} u_t^n \, dV_t^k + \int_{\mathfrak{Z}_0} J_t^{\eta^n *} u_t^n \, \nu_0(d\mathfrak{z}) dt \\ &+ \int_{\mathfrak{Z}_1} J_t^{\xi^n *} u_t^n \, \nu_1(d\mathfrak{z}) dt + \int_{\mathfrak{Z}_1} I_t^{\xi^n *} u_t^n \, \tilde{N}_1(d\mathfrak{z}, dt), \quad \text{with } u_0^n = \pi_0^n, \end{aligned} \tag{IV.7.13}$$

where for each fixed n and $k = 1, 2, ..., d'$

$$\tilde{\mathcal{L}}_t^n := a_t^{nij} D_{ij} + b_t^{ni} D_i + \beta_t^{nk} \rho_t^{nik} D_i + \beta_t^{nk} B_t^{nk}, \quad \mathcal{M}_t^{nk} := \rho_t^{nik} D_i + B_t^{nk},$$

$$a_t^{nij} := \tfrac{1}{2}\sum_k \sigma_t^{nik}\sigma_t^{njk} + \tfrac{1}{2}\sum_k \rho_t^{nik}\rho_t^{njk}, \quad \beta_t^n := B^n(t, X_{t-}^n, Y_{t-}^n), \quad i,j = 1,2,...,d,$$

the operators $J_t^{\eta^n}$ and $J_t^{\xi^n}$ are defined as J_t^{ξ} in (IV.3.1) with η_t^n and ξ_t^n in place of η_t and ξ_t, respectively, and the operator $I_t^{\xi^n}$ is defined as I_t^{ξ} in (IV.3.1) with ξ_t^n in place of ξ_t. For each n let γ^n denote the solution to $d\gamma_t^n = -\gamma_t^n\beta_t^n\,dV_t$, $\gamma_0^n = 1$. By virtue of Step II (IV.7.13) has an L_p-solution $u^n = (u_t^n)_{t\in[0,T]}$, which is also its unique L_2-solution, i.e., for each $\varphi \in C_0^\infty$ almost surely

$$\begin{aligned}(u_t^n, \varphi) =&(\pi_0^n, \varphi) + \int_0^t (u_s^n, \tilde{\mathcal{L}}_s^n\varphi)\,ds + \int_0^t (u_s^n, \mathcal{M}_s^{nk}\varphi)\,dV_s^k + \int_0^t\int_{\mathfrak{Z}_0} (u_s^n, J_s^{\eta^n}\varphi)\,\nu_0(d\mathfrak{z})ds \\ &+ \int_0^t\int_{\mathfrak{Z}_1} (u_s^n, J_s^{\xi^n}\varphi)\,\nu_1(d\mathfrak{z})ds + \int_0^t\int_{\mathfrak{Z}_1} (u_s^n, I_s^{\xi^n}\varphi)\,\tilde{N}_1(d\mathfrak{z}, ds) \qquad \text{(IV.7.14)}\end{aligned}$$

for all $t \in [0,T]$. Moreover, almost surely $u_t^n = d\mu_t^n/dx$ for all $t \in [0,T]$, where $\mu_t^n = P_t^n({}^o\gamma_t^n)^{-1}$ is the unnormalised conditional distribution, P_t^n is the regular conditional distribution of X_t^n given $\mathcal{F}_t^{Y^n}$, and ${}^o\gamma^n$ denotes the $\mathcal{F}_t^{Y^n}$-optional projection of γ^n under P. Furthermore, for sufficiently large n we have

$$\mathbb{E}\sup_{t\in[0,T]} |u_t^n|_{L_r}^r \leqslant N|\pi_0^n|_{\mathbb{L}_r} \leqslant N|\pi_0|_{\mathbb{L}_r} \quad \text{for } r = p, 2 \qquad \text{(IV.7.15)}$$

with a constant $N = N(d, d', p, K, K_\xi, K_\eta, L, T, |\bar{\xi}|_{L_2}, |\bar{\eta}|_{L_2}, \lambda)$, which together with (IV.7.11) implies

$$\sup_{n\geqslant 1}(|u_T^n|_{\mathbb{L}_r} + |u^n|_{\mathbb{L}_{r,q}}) < \infty \quad \text{for } r = 2, p \text{ and every } q > 1.$$

Hence there exist a subsequence, denoted again by $(u^n)_{n=1}^\infty$, $\bar{u} \in \bigcap_{q=2}^\infty \mathbb{L}_{r,q}$ and $g \in \mathbb{L}_r$ for $r = 2, p$ such that

$$u^n \to \bar{u} \quad \text{weakly in } \mathbb{L}_{r,q} \text{ for } r = p, 2 \text{ and all integers } q > 1, \qquad \text{(IV.7.16)}$$

and

$$u_T^n \to g \quad \text{weakly in } \mathbb{L}_r \text{ for } r = p, 2. \qquad \text{(IV.7.17)}$$

One knows, see e.g. [20], that $(X_t^n, Y_t^n)_{t\geqslant 0}$ converges to $(X_t, Y_t)_{t\geqslant 0}$ in probability, uniformly in t in finite intervals. Hence it is not difficult to show (see Lemma 3.8 in [21]) that there is a subsequence of Y^n, denoted for simplicity also by Y^n, and there is an $\mathcal{F}_t$-adapted cadlag process $(U_t)_{t\in[0,T]}$, such that almost surely $|Y_t^n| + |Y_t| \leqslant U_t$ for every $t \in [0,T]$ and integers $n \geqslant 1$. For every integer $m \geqslant 1$ define the stopping time

$$\tau_m = \inf\{t \in [0,T] : U_t \geqslant m\}$$

To show that $\bar{u}$ is a $\mathbb{V}_r$-solution for $r = p, 2$ to (IV.3.8) with initial condition $u_0 = \pi_0$, we pass to the limit $u^n \to \bar{u}$ in equation (IV.7.14) in a similar way to that as we passed to the limit $u^{\varepsilon_n} \to \bar{u}$ in equation (IV.6.77) in the proof of Lemma IV.6.4. We fix an integer $m \geqslant 1$ and multiply both sides of (IV.7.14) with $(\phi_t \mathbf{1}_{t \leqslant \tau_m})_{t\in[0,T]}$, where $(\phi_t)_{t\in[0,T]}$ is an arbitrary bounded $\mathcal{F}_t$-optional process $\phi = (\phi_t)_{t\in[0,T]}$. Then we integrate both sides of the equation we obtained over $\Omega \times [0,T]$ against $P \otimes dt$ to get

$$F(u^n) = F(\pi_0^n) + \sum_{i=1}^{5} F_i^n(u^n), \tag{IV.7.18}$$

where F and F_i^n are linear functionals over $\mathbb{L}_{r,q}$, defined by

$$F(v) := \mathbb{E}\int_0^{T\wedge\tau_m} \phi_t (v_t, \varphi)\, dt, \quad F_1^n(v) := \mathbb{E}\int_0^{T\wedge\tau_m} \phi_t \int_0^t (v_s, \tilde{\mathcal{L}}_s^n \varphi)\, ds\, dt,$$

$$F_2^n(v) := \mathbb{E}\int_0^{T\wedge\tau_m} \phi_t \int_0^t (v_s, \mathcal{M}_s^{nk}\varphi)\, dV_s^k\, dt,$$

$$F_3^n(v) := \mathbb{E}\int_0^{T\wedge\tau_m} \phi_t \int_0^t \int_{\mathfrak{Z}_0} (v_s, J_s^{\eta^n}\varphi)\, \nu_0(d\mathfrak{z}) ds\, dt,$$

$$F_4^n(v) := \mathbb{E}\int_0^{T\wedge\tau_m} \phi_t \int_0^t \int_{\mathfrak{Z}_1} (v_s, J_s^{\xi^n}\varphi)\, \nu_1(d\mathfrak{z}) ds\, dt,$$

$$F_5^n(v) := \mathbb{E}\int_0^{T\wedge\tau_m} \phi_t \int_0^t \int_{\mathfrak{Z}_1} (v_s, I_s^{\xi^n}\varphi)\, \tilde{N}_1(d\mathfrak{z}, ds)\, dt$$

for a fixed $\varphi \in C_0^\infty$. Define also F_i as F_i^n for $i = 1, 2, ..., 5$, with $\tilde{\mathcal{L}}_s$, $\mathcal{M}_s^k$, J_s^η, J_s^ξ and I_s^ξ in place of $\tilde{\mathcal{L}}_s^n$, $\mathcal{M}_s^{nk}$, $J_s^{\eta^n}$, $J_s^{\xi^n}$ and $I_s^{\xi^n}$, respectively. It is an easy exercise to show that hat F and F_i^n, $i = 1, 2, 3, 4, 5$, are continuous linear functionals on $\mathbb{L}_{r,q}$ for $r = p, 2$ and all $q > 1$. We are going to show now that for $r = p, 2$

$$\lim_{n\to\infty} \sup_{|v|_{\mathbb{L}_{r,q}}=1} |F_i(v) - F_i^n(v)| = 0 \quad \text{for every } q > 1, \text{ for i=1,2,...,5.} \tag{IV.7.19}$$

Let $r' = r/(r-1)$, $q' = q/(q-1)$. Then for $v \in \mathbb{L}_{r,q}$ by Hölder's inequality we have

$$|F_1(v) - F_1^n(v)| \leqslant KT|v|_{\mathbb{L}_{r,q}} |(\tilde{\mathcal{L}} - \tilde{\mathcal{L}}^n)\varphi|_{\mathbb{L}_{r',q'}} \tag{IV.7.20}$$

with $K = \sup_{\omega\in\Omega} \sup_{t\in[0,T]} |\phi_t| < \infty$. Clearly, $\lim_{n\to\infty}(\tilde{\mathcal{L}}_s - \tilde{\mathcal{L}}_s^n)\varphi(x) = 0$ almost surely for all $s \in [0,T]$ and $x \in \mathbb{R}^d$, and there is a constant N independent of n and m such that

$$|(\tilde{\mathcal{L}}_s - \tilde{\mathcal{L}}_s^n)\varphi(x)| \leqslant N(1 + |x|^2 + 2m^2)\mathbf{1}_{|x|\leqslant R} \tag{IV.7.21}$$

for $\omega \in \Omega$, $s \in [0, T \wedge \tau_m]$ and $x \in \mathbb{R}^d$, where R is the diameter of the support of φ. Hence a repeated application of Lebesgue's theorem on dominated convergence gives

$$\lim_{n\to\infty} |(\tilde{\mathcal{L}} - \tilde{\mathcal{L}}^n)\varphi|_{\mathbb{L}_{r',q'}} = 0,$$

and by (IV.7.20) proves (IV.7.19) for $i = 1$. By the Davis inequality and Hölder's inequality we have

$$|F_2(v) - F_2^n(v)| \leqslant 3KT\mathbb{E}\Big(\int_0^{T\wedge\tau_m} \sum_k |(v_s, (\mathcal{M}_s^k - \mathcal{M}_s^{nk})\varphi)|^2\, ds\Big)^{1/2} \leqslant C_n^{(2)}|v|_{\mathbb{L}_{r,q}}$$

with

$$C_n^{(2)} = 3KT\Big(\mathbb{E}\Big(\int_0^{T\wedge\tau_m} \big(\sum_k |(\mathcal{M}_s^k - \mathcal{M}_s^{nk})\varphi|^2_{L_{r'}}\big)^{q/(q-2)}\, ds\Big)^{r'(q-2)/2q}\Big)^{1/r'}.$$

Clearly, $\lim_{n\to\infty}(\mathcal{M}_s^k - \mathcal{M}_s^{nk})\varphi(x) = 0$, and with a constant N independent of n and m we have

$$\sum_k |(\mathcal{M}_s^k - \mathcal{M}_s^{nk})\varphi(x)| \leqslant N(1 + |x| + 2m)\mathbf{1}_{|x|\leqslant R}$$

for all $\omega \in \Omega$, $s \in [0, T \wedge \tau_m]$ and $x \in \mathbb{R}^d$. Thus repeating the above argument we obtain (IV.7.19) for $i = 2$. By Hölder's inequality we have

$$|F_n^3(v) - F^3(v)| \leqslant KT|v|_{\mathbb{L}_{r,q}}C_n^{(3)}$$

with

$$C_n^{(3)} = \Big(\mathbb{E}\Big(\int_0^{T\wedge\tau_m} \Big|\int_{\mathfrak{Z}_0} |(J_s^\eta - J_s^{\eta^n})\varphi|_{L_{r'}}\nu_0(d\mathfrak{z})\Big|^{q'}\, ds\Big)^{r'/q'}\Big)^{1/r'},$$

where we have suppressed the variable $\mathfrak{z} \in \mathfrak{Z}_0$ in the integrand. Clearly,

$$\lim_{n\to\infty}(J^\eta - J^{\eta^n})\varphi(x) = 0 \quad \text{almost surely for all } s \in [0,T],\ x \in \mathbb{R}^d \text{ and } \mathfrak{z} \in \mathfrak{Z}_0.$$

By Taylor's formula

$$|J_s^{\eta^n}\varphi(x)| \leqslant \sup_{\theta\in[0,1]} |D^2\varphi(x + \theta\eta_s^n(x,\mathfrak{z}))||\eta_s(x,\mathfrak{z})|^2,$$

$$|J_s^{\eta}\varphi(x)| \leqslant \sup_{\theta\in[0,1]} |D^2\varphi(x + \theta\eta_s(x,\mathfrak{z}))||\eta_s(x,\mathfrak{z})|^2,$$

and by Lemma IV.7.3 with λ' from above we have

$$\lambda'|x| \leqslant |x + \theta(\eta_s^n(x,\mathfrak{z}) - \eta_s^n(0,\mathfrak{z}))| \leqslant |x + \theta\eta_s^n(x,\mathfrak{z})| + |\eta_s^n(0,\mathfrak{z})|,$$

$$\lambda'|x| \leqslant |x + \theta(\eta_s(x,\mathfrak{z}) - \eta_s(0,\mathfrak{z}))| \leqslant |x + \theta\eta_s(x,\mathfrak{z})| + |\eta_s(0,\mathfrak{z})|$$

for all $\theta \in [0,1]$, $\omega \in \Omega$, $s \in [0, T \wedge \tau_m]$. Hence, taking into account the the linear growth condition on η, see Assumption (IV.2.1) (ii), for any given $R > 0$

we have a constant $\tilde{R} = \tilde{R}(R, K_0, K_1, K_\eta, m) > R$ such that

$$|x + \theta\eta_s(x, \mathfrak{z})| \geqslant R, \quad |x + \theta\eta^n_s(x, \mathfrak{z})| \geqslant R \quad \text{for } |x| \geqslant \tilde{R},$$

for all $\theta \in [0, 1]$, $\omega \in \Omega$, $s \in [0, T \wedge \tau_m]$. Taking R such that $\varphi(x) = 0$ for $|x| \geqslant R$ we have

$$|J^{\eta^n}_s \varphi(x) - J^\eta_s \varphi(x)| \leqslant |J^{\eta^n}_s \varphi(x)| + |J^\eta_s \varphi(x)| \leqslant 2 \sup_{x \in \mathbb{R}^d} |D^2\varphi(x)| \bar{\eta}^2(\mathfrak{z}) \mathbf{1}_{|x| \leqslant \tilde{R}}$$

for $x \in \mathbb{R}$, $\omega \in \Omega$, $s \in [0, T \wedge \tau_m]$ and $\mathfrak{z} \in \mathfrak{Z}_0$. Hence by Lebesgue's theorem on dominated convergence $\lim_{n\to\infty} C^{(3)}_n = 0$ which gives (IV.7.19) for $i = 3$. We get (IV.7.19) for $i = 4$ in the same way. By the Davis inequality and Hölder's inequality we have

$$|F_5(v) - F^n_5(v)| \leqslant 3KT\mathbb{E}\Big(\int_0^{T \wedge \tau_m} \int_{\mathfrak{Z}_1} |(v_s, (I^{\xi^n}_s - I^\xi_s)\varphi)|^2 \nu_1(d\mathfrak{z})\, ds\Big)^{1/2} \leqslant C^{(5)}_n |v|_{\mathbb{L}_{r,q}}$$

with

$$C^{(5)}_n = 3KT\Big(\mathbb{E}\Big(\int_0^{T \wedge \tau_m} \Big(\int_{\mathfrak{Z}_1} |(I^{\xi^n}_s - I^\xi_s)\varphi|^2_{L_{r'}} \nu_1(d\mathfrak{z})\Big)^{q/(q-2)} ds\Big)^{r'(q-2)/2q}\Big)^{1/r'}.$$

Clearly, $\lim_{n\to\infty}(I^{\xi^n}_s - I^\xi_s)\varphi(x) = 0$ almost surely for all $s \in [0, T]$, $x \in \mathbb{R}^d$ and $\mathfrak{z} \in \mathfrak{Z}_1$. By Taylor's formula

$$|I^{\xi^n}_s \varphi(x)| \leqslant \sup_{\theta \in [0,1]} |D\varphi(x + \theta\xi^n_s(x, \mathfrak{z}))||\xi_s(x, \mathfrak{z})|,$$

$$|I^\xi_s \varphi(x)| \leqslant \sup_{\theta \in [0,1]} |D\varphi(x + \theta\xi_s(x, \mathfrak{z}))||\xi_s(x, \mathfrak{z})|.$$

Hence, using Assumptions IV.2.1, IV.2.2 and IV.2.4 in the same way as above, we get a constant $\tilde{R} = \tilde{R}(R, K_0, K_1, K_\eta, m)$ such that

$$|I^{\xi^n}_s \varphi(x) - I^\xi_s \varphi(x)| \leqslant |I^{\xi^n}_s \varphi(x)| + |I^\xi_s \varphi(x)| \leqslant 2 \sup_{x \in \mathbb{R}^d} |D\varphi(x)| \bar{\xi}(\mathfrak{z}) \mathbf{1}_{|x| \leqslant \tilde{R}}$$

for $x \in \mathbb{R}$, $\omega \in \Omega$, $s \in [0, T \wedge \tau_m]$ and $\mathfrak{z} \in \mathfrak{Z}_0$. Consequently, by Lebesgue's theorem on dominated convergence we obtain (IV.7.19) for $i = 5$, which completes the proof of (IV.7.19). Since u^n converges weakly to $\bar{u}$ in $\mathbb{L}_{r,q}$, and F^n_i converges strongly to F_i in $\mathbb{L}^*_{p,q}$, the dual of $\mathbb{L}_{p,q}$, we get that $F^n_i(u^n)$ converges to $F_i(\bar{u})$ for for $i = 1, 2, 3, 4, 5$. Therefore letting $n \to \infty$ in (IV.7.18) we obtain

$$\mathbb{E}\int_0^{T \wedge \tau_m} \phi_t(\bar{u}_t, \varphi)\, dt = \mathbb{E}\int_0^{T \wedge \tau_m} \phi_t(\psi, \varphi)\, dt + \mathbb{E}\int_0^{T \wedge \tau_m} \phi_t \int_0^t (\bar{u}_s, \tilde{\mathcal{L}}_s\varphi)\, ds\, dt$$

$$+\mathbb{E}\int_0^{T \wedge \tau_m} \phi_t \int_0^t (\bar{u}_s, \mathcal{M}^k_s\varphi)\, dV^k_s\, dt + \mathbb{E}\int_0^{T \wedge \tau_m} \phi_t \int_0^t \int_{\mathfrak{Z}_0} (\bar{u}_s, J^\eta_s\varphi)\, \nu_0(d\mathfrak{z}) ds\, dt$$

$$+\mathbb{E}\int_0^{T\wedge\tau_m}\phi_t\int_0^t\int_{\mathfrak{Z}_1}(\bar u_s, J_s^\xi\varphi)\,\nu_1(d\mathfrak{z})ds\,dt+\mathbb{E}\int_0^{T\wedge\tau_m}\phi_t\int_0^t\int_{\mathfrak{Z}_1}(\bar u_s, I_s^\xi\varphi)\,\tilde N_1(d\mathfrak{z},ds)\,dt.$$

Since this equation holds for all bounded $\mathcal{F}_t$-optional processes $\phi=(\phi_t)_{t\in[0,T]}$, we get

$$\mathbf{1}_{t\leqslant\tau_m}(\bar u_t,\varphi)=\mathbf{1}_{t\leqslant\tau_m}\Big((\psi,\varphi)+\int_0^t(\bar u_s,\tilde{\mathcal{L}}_s\varphi)\,ds+\int_0^t(\bar u_s,\mathcal{M}_s^k\varphi)\,dV_s^k\Big)$$

$$+\mathbf{1}_{t\leqslant\tau_m}\Big(\int_0^t\int_{\mathfrak{Z}_0}(\bar u_s,J_s^\eta\varphi)\,\nu_0(d\mathfrak{z})ds+\int_0^t\int_{\mathfrak{Z}_1}(\bar u_s,J_s^\xi\varphi)\,\nu_1(d\mathfrak{z})ds+\int_0^t\int_{\mathfrak{Z}_1}(\bar u_s,I_s^\xi\varphi)\,\tilde N_1(d\mathfrak{z},ds)\Big)$$

for $P\otimes dt$-almost every $(t,\omega)\in[0,T]\times\Omega$ for every $\varphi\in C_0^\infty$ and integer $m\geqslant 1$, which implies that $\bar u$ is a $\mathbb{V}_r$-solution to (IV.3.8) for $r=2,p$. In the same way as in the proof Lemma IV.6.4 we can show first that almost surely

$$\begin{aligned}\mathbf{1}_{\tau_m>T}(g,\varphi)=\mathbf{1}_{\tau_m>T}\Big((\psi,\varphi)+\int_0^T(\bar u_s,\tilde{\mathcal{L}}_s\varphi)\,ds+\int_0^T(\bar u_s,\mathcal{M}_s^k\varphi)\,dV_s^k\Big)\\+\mathbf{1}_{\tau_m>T}\Big(\int_0^T\int_{\mathfrak{Z}_0}(\bar u_s,J_s^\eta\varphi)\,\nu_0(d\mathfrak{z})ds+\int_0^T\int_{\mathfrak{Z}_1}(\bar u_s,J_s^\xi\varphi)\,\nu_1(d\mathfrak{z})ds\Big)\\+\mathbf{1}_{\tau_m>T}\int_0^T\int_{\mathfrak{Z}_1}(\bar u_s,I_s^\xi\varphi)\,\tilde N_1(d\mathfrak{z},ds)\end{aligned}\qquad\text{(IV.7.22)}$$

for every $m\geqslant 1$. Hence taking into account $P(\cup_{m=1}^\infty\{\tau_m>T\})=1$, we get that equation (IV.7.22) remains valid if we omit $\mathbf{1}_{\tau_m>T}$ everywhere in it. From (IV.7.15) we get that for all n,

$$|u^n|_{\mathbb{L}_{r,q}}\leqslant N|\pi_0^n|_{\mathbb{L}_r}\quad\text{for } r=2,p,\text{ for integers } q>1$$

with a constant $N=N(d,d',p,K,K_\xi,K_\eta,L,T,|\bar\xi|_{L_2},|\bar\eta|_{L_2},\lambda)$. Letting here $n\to\infty$ and taking into account (IV.7.11) and (IV.7.17) we obtain

$$|\bar u|_{\mathbb{L}_{r,q}}\leqslant\liminf_{n\to\infty}|u^n|_{\mathbb{L}_{r,q}}\leqslant N\lim_{n\to\infty}|\pi_0^n|_{L_r}\leqslant N|\pi_0|_{\mathbb{L}_r}\quad\text{for } r=2,p\text{ and integers } q>1.$$

Letting here $q\to\infty$ we get

$$\mathbb{E}\operatorname*{ess\,sup}_{t\in[0,T]}|\bar u_t|_{\mathbb{L}_r}^r\leqslant N^r\mathbb{E}|\pi_0|_{\mathbb{L}_r}^r,\quad\text{for } r=2,p.$$

Hence, taking into account (IV.7.22), by Lemma IV.5.8 we get a $P\otimes dt$-modification u of $\bar u$, which is an L_r-solution for $r=2,p$ to equation (IV.3.8) with initial condition $u_0=\pi_0$. As the limit of $P\otimes dt\otimes dx$-almost everywhere nonnegative functions, u is also $P\otimes dt\otimes dx$ almost everywhere nonnegative. We now show that u satisfies

$$G(u):=\sup_{t\in[0,T]}\int_{\mathbb{R}^d}|x|^2u_t(dx)<\infty\,(a.s.).\qquad\text{(IV.7.23)}$$

To show this recall that for each n and $\varphi \in C_b^2$, by Theorem III.1.1, Remark IV.3 and by what we have proven above,

$$\mu_t^n(\varphi) = P_t^n(\varphi)\mu_t^n(\mathbf{1}) = \mathbb{E}(\varphi(X_t^n)|\mathcal{F}_t^{Y^n})({}^o\gamma_t^n)^{-1},$$

where $\mu_t^n(dx) = u_t^n(x)dx$, $P_t^n(dx) = \pi_t^n(x)dx$ and ${}^o\gamma^n$ denotes the $\mathcal{F}_t^{Y^n}$-optional projection of $(\gamma_t^n)_{t\in[0,T]}$. Further, for integers $m \geqslant 1$ let again $\Omega_m := [|Y_0| \leqslant m] \in \mathcal{F}_0^Y$. Thus by Doob's inequality and Jensen's inequality for optional projections, for $r > 1$ we have, in the same way as in Step I,

$$G_m(u^n) := \mathbb{E} \sup_{t\in[0,T]} \int_{\mathbb{R}^d} |x|^2 u_t^n(x)\, dx \mathbf{1}_{\Omega_m} = \mathbb{E} \sup_{t\in[0,T]} \mathbb{E}(|X_t^n|^2|\mathcal{F}_t^{Y^n})({}^o\gamma_t^n)^{-1}\mathbf{1}_{\Omega_m}$$

$$\leqslant N\big(\mathbb{E} \sup_{t\in[0,T]} |X_t^n|^r \mathbf{1}_{\Omega_m}\big)^{2/r} \quad \text{for } t \in [0,T]$$

with a constant $N = N(r, C)$, where C is the constant from (IV.7.8), which depends only on K, r and T. Taking r from Assumption IV.2.3, by Young's inequality, (IV.2.1) for all m and n we have

$$G_m(u^n) \leqslant N\big(m^r + \mathbb{E}|X_0^n|^r)\big) \leqslant N\big(m^r + \sup_n \mathbb{E}|X_0^n|^r\big) =: N'(m) < \infty. \quad \text{(IV.7.24)}$$

By Mazur's theorem there exists a sequence of convex linear combinations $v^k = \sum_{i=1}^k c_{i,k}u^i$ converging to u (strongly) in $\mathbb{L}_{p,q}$ as $k \to \infty$. Thus there exists a subsequence, also denoted by $(v^k)_{k=1}^\infty$ which converges to u for $P \otimes dt \otimes dx$-almost every (ω, t, x). Then, by Fatou's lemma and (IV.7.24),

$$G_m(u) = \mathbb{E} \sup_{t\in[0,T]} \int_{\mathbb{R}^d} |x|^2 \liminf_{k\to\infty} v_t^k(x)\, dx \mathbf{1}_{\Omega_m} \leqslant \liminf_{k\to\infty} G_m(v^k)$$

$$= \liminf_{k\to\infty} \sum_{i=1}^k c_{k,i} G_m(u^i) \leqslant N'(m) \quad \text{for each integer } m \geqslant 1,$$

which proves (IV.7.23). Next, due to Lemma IV.3.1, using $|B^n| \leqslant |B| \leqslant K$, we have

$$\sup_{n\in\mathbb{N}} \mathbb{E} \sup_{t\in[0,T]} |u_t^n|_{L_1} \leqslant N, \quad \text{(IV.7.25)}$$

for a constant $N = N(d, K, T)$. The estimate above implies that $u^n \in \mathbb{L}_{1,q}$ for all $q \geqslant 1$. Returning to the sequence $(v^k)_{k\in\mathbb{N}} \subset \mathbb{L}_{1,q} \cap \mathbb{L}_{p,q}$ converging point-wise to u for $P \otimes dt \otimes dx$-almost every (ω, t, x), we can compute by use of Fatou's lemma

$$\mathbb{E} \operatorname*{ess\,sup}_{t\in[0,T]} |u_t|_{L_1} = \mathbb{E} \operatorname*{ess\,sup}_{t\in[0,T]} |\liminf_{k\to\infty} v_t^k|_{L_1} \leqslant \liminf_{k\to\infty} \mathbb{E} \sup_{t\in[0,T]} |v_t^k|_{L_1}$$

$$\leqslant \liminf_{k\to\infty} \sum_{i=1}^{k} c_{i,k} \mathbb{E} \sup_{t\in[0,T]} |u_t^i|_{L_1} \leqslant N, \tag{IV.7.26}$$

with the constant N from (IV.7.25). As also (IV.7.23) holds and since u is in particular an L_2-solution to (IV.3.8) we can apply Lemma IV.5.5, in particular the uniqueness of L_2-solutions (satisfying (IV.7.23) if $K_1 \neq 0$, as well as $\operatorname{ess\,sup}_{t\in[0,T]} |u_t|_{L_1} < \infty$ (a.s.)) implied by the supremum estimate (IV.5.50) of Lemma IV.5.5. Hence we see that indeed for all $t \in [0,T]$, $u_t = d\mu_t/dx$ almost surely and thus $\pi_t = u_t{}^{o}\gamma_t$, or in other words, the L_2-solution constructed above coincides with the L_2-solution from Step I. This finishes the proof. □

Chapter V

Regularity of the filtering density

V.1 Introduction

In this chapter we are interested in the regularity (in the Sobolev sense) of the filtering density $(\pi_t)_{t\in[0,T]}$ associated to the signal-observation system (I.0.2). Again we assume the measurability conditions on the coefficients, given in the introduction to Chapter I, to hold. This chapter is based on the article [18].

In Chapter IV we showed that if the coefficients of equation (I.0.2) satisfy natural linear growth and Lipschitz conditions, the drift coefficient of the observation process is bounded, $\tau^\xi = x + \theta\xi(x)$, $\tau^\eta = x + \theta\eta(x)$ are bijective mappings on $\mathbb{R}^d$, they have Lipschitz continuous inverses with a Lipschitz constant independent of $t, y, \mathfrak{z}$ and $\vartheta \in [0,1]$, and their derivatives in x are equicontinuous in x, uniformly in $t, y, \mathfrak{z}$, then for $p \geqslant 2$ the conditional density π_t exists almost surely for each t and $(\pi_t)_{t\geqslant 0}$ is a weakly cadlag L_p-valued process, whenever the initial conditional distribution P_0 has a density π_0 almost surely such that $\mathbb{E}|\pi_0|^p_{L_p} < \infty$.

For partially observed diffusion processes, i.e., when $\xi = \eta = 0$ and the observation process Y does not have jumps, the existence and the regularity properties of the conditional density π_t have been extensively studied in the literature. In [39], an early work on the regularity of the filtering density for continuous diffusions, it was shown that if the coefficients are bounded, σ, ρ admit $m+1$ bounded derivatives in $x \in \mathbb{R}^d$, b, B admit m bounded derivatives in x, the functions σ, ρ satisfy a nondegeneracy condition and $\pi_0 \in W^m_p \cap W^m_2$, then the filtering density $(\pi_t)_{t\in[0,T]}$ is weakly continuous as W^m_p-valued process, where $p \geqslant 2$ and $m \geqslant 0$. In [51] it was proven that the nondegeneracy condition can be dropped if one imposes $m+2$ bounded derivatives on σ, ρ in x, as well as $m+1$ derivatives on b, B in x, to obtain the same result under otherwise the same assumptions. The results for $m = 2$ from [39] were later strengthened in [32], for bounded coefficients, Lipschitz in space and such that σ, ρ are differentiable with respect to $x \in \mathbb{R}^d$, such that the differential is continuous in $y \in \mathbb{R}^{d'}$ and Lipschitz in x. Similarly, in [35], it was shown that if the derivatives in x of b, B satisfy a certain Lipschitz condition, σ, ρ are bounded and π_0 belongs to a certain subspace of W^1_p for $p \geqslant 2$, then π_t belongs almost surely to L_r, $r \in [1, p]$ for all time. In a later work, [36], for Lipschitz (in space), bounded coefficients, as well as under a nondegeneracy condition it was shown that if π_0 belongs to a fractional Sobolev space with integrability index $p \geqslant 2$, then π_t belongs to W^1_p for all time.

To the best of the author's knowledge, the Sobolev regularity of filtering densities associated to jump diffusion systems has not been considered yet. Our article [18], which serves as a basis for this chapter, provides a first result in this direction.

More precisely, in the present chapter we show that if the coefficients admit $m+1$ continuous bounded derivatives in $x \in \mathbb{R}^d$, have linear growth in $z = (x, y) \in \mathbb{R}^{d+d'}$, the jump coefficients are biLipschitz in x, the initial condition together with a Lévy measure have finite r-th moment for some $r > 2$, and if $\mathbb{E}|\pi_0|^p_{W^m_p} < \infty$, then for all time π_t remains in that Sobolev space W^m_p with $p \geqslant 2$ and integer $m \geqslant 1$. Moreover it is weakly cadlag as W^m_p-valued process and, if $m \geqslant 1$ and $K_1 = 0$, then it is strongly cadlag as W^s_p-valued process, for $s \in [0, m)$.

This chapter is structured as follows. Section V.2 contains the main results along with the required assumptions. In section V.3 we state some important results from Chapters III and IV which we build on. Section V.4 contains Sobolev estimates necessary to obtain a priori estimates for the smoothed filtering equations. In section V.5 we investigate some solvability properties of the Zakai equation. Section V.6 finally contains the proof of our main theorem, as well as some auxiliary results.

V.2 Formulation of the main results

We fix nonnegative constants K_0, K_1, L, K and functions $\bar{\xi} \in L_2(\mathfrak{Z}_1) = L_2(\mathfrak{Z}_1, \mathcal{Z}_1, \nu_1)$, $\bar{\eta} \in L_2(\mathfrak{Z}_0) = L_2(\mathfrak{Z}_0, \mathcal{Z}_0, \nu_0)$, used throughout the paper, and make the following assumptions.

Assumption V.2.1. (i) For $z_j = (x_j, y_j) \in \mathbb{R}^{d+d'}$ $(j = 1, 2)$, $t \geqslant 0$ and $\mathfrak{z}_i \in \mathfrak{Z}_i$ $(i = 0, 1)$,

$$|b(t, z_1) - b(t, z_2)| + |B(t, z_1) - B(t, z_2)| + |\sigma(t, z_1) - \sigma(t, z_2)|$$

$$+|\rho(t, z_1) - \rho(t, z_2)| \leqslant L|z_1 - z_2|,$$
$$|\eta(t, z_1, \mathfrak{z}_0) - \eta(t, z_2, \mathfrak{z}_0)| \leqslant \bar{\eta}(\mathfrak{z}_0)|z_1 - z_2|,$$
$$|\xi(t, z_1, \mathfrak{z}_1) - \xi(t, z_2, \mathfrak{z}_1)| \leqslant \bar{\xi}(\mathfrak{z}_1)|z_1 - z_2|.$$

(ii) For all $z = (x, y) \in \mathbb{R}^{d+d'}$, $t \geqslant 0$ and $\mathfrak{z}_i \in \mathfrak{Z}_i$ for $i = 0, 1$ we have

$$|b(t, z)|+|\sigma(t, z)|+|\rho(t, z)| \leqslant K_0+K_1|z|, \quad |B(t, z)| \leqslant K, \quad \int_{\mathfrak{Z}_1} |\mathfrak{z}|^2 \,\nu_1(d\mathfrak{z}) \leqslant K_0^2,$$

$$|\eta(t,z,\mathfrak{z}_0)| \leqslant \bar\eta(\mathfrak{z}_0)(K_0+K_1|z|), \quad |\xi(t,z,\mathfrak{z}_1)| \leqslant \bar\xi(\mathfrak{z}_1)(K_0+K_1|z|).$$

(iii) The initial condition $Z_0 = (X_0, Y_0)$ is an $\mathcal{F}_0$-measurable random variable with values in $\mathbb{R}^{d+d'}$.

Assumption V.2.2. The functions $\bar\eta \in L_2(\mathfrak{Z}_0, \mathcal{Z}_0, \nu_0)$ and $\bar\xi \in L_2(\mathfrak{Z}_1, \mathcal{Z}_1, \nu_1)$ are such that for nonnegative constants K_η and K_ξ we have $|\bar\eta(\mathfrak{z}_0)| \leqslant K_\eta$ and $|\bar\xi(\mathfrak{z}_1)| \leqslant K_\xi$ for all $\mathfrak{z}_i \in \mathfrak{Z}_i$ $(i = 0, 1)$.

Assumption V.2.3. For some $r > 2$ let $\mathbb{E}|X_0|^r < \infty$ and the measure ν_1 satisfy

$$K_r := \int_{\mathfrak{Z}_1} |\mathfrak{z}|^r \, \nu_1(d\mathfrak{z}) < \infty.$$

We recall that by Theorem II.2.1 and Theorem II.2.2, Assumption V.2.1 ensures the existence and uniqueness of a solution $(Z_t)_{t\geqslant0} = (X_t, Y_t)_{t\geqslant0}$ to (I.0.2) and for every $T > 0$,

$$\mathbb{E}\sup_{t\leqslant T}(|X_t|^q + |Y_t|^q) \leqslant N(1 + \mathbb{E}|X_0|^q + \mathbb{E}|Y_0|^q) \tag{V.2.1}$$

holds for $q = 2$ with a constant $N = N(K_0, K_1, d, d', T, L, |\bar\xi|_{L_2}, |\bar\eta|_{L_2})$ and if additionally Assumptions V.2.2 and V.2.3 hold, then the moment estimate (V.2.1) holds with $q := r$ for every $T > 0$, where now N depends also on r, K_r K_ξ and K_η.

Assumption V.2.4. (i) For a constant $\lambda > 0$ we have

$$\lambda|x - \bar x| \leqslant |x - \bar x + \theta(f_i(t,x,y,\mathfrak{z}_i) - f_i(t,\bar x,y,\mathfrak{z}_i))|$$

for all $\theta \in [0,1]$, $t \in [0,T]$, $y \in \mathbb{R}^{d'}$, $x, \bar x \in \mathbb{R}^d$, $\mathfrak{z}_i \in \mathfrak{Z}_i$, $i = 0$, and $f_0(t,x,y,\mathfrak{z}_0) = \eta(t,x,y,\mathfrak{z}_0)$, $f_1(t,x,y,\mathfrak{z}_1) = \xi(t,x,y,\mathfrak{z}_1)$.
(ii) For all $(t,y) \in \mathbb{R}_+ \times \mathbb{R}^{d'}$ and all $x_1, x_2 \in \mathbb{R}^d$,

$$|(\rho B)(t,x_1,y) - (\rho B)(t,x_2,y)| \leqslant L|x_1 - x_2|.$$

(iii) The functions $f_0(t,x,y,\mathfrak{z}) := \xi(t,x,y,\mathfrak{z})$ and $f_1(t,x,y,\mathfrak{z}) := \eta(t,x,y,\mathfrak{z})$ are continuously differentiable in $x \in \mathbb{R}^d$ for each $(t,y,\mathfrak{z}) \in \mathbb{R}_+ \times \mathbb{R}^{d'} \times \mathfrak{Z}_i$, for $i = 0$ and $i = 1$, respectively, such that

$$\lim_{\varepsilon\downarrow0} \sup_{t\in[0,T]} \sup_{\mathfrak{z}\in\mathfrak{Z}_i} \sup_{|y|\leqslant R} \sup_{|x|\leqslant R, |\bar x|\leqslant R, |x-x'|\leqslant\varepsilon} |D_x f_i(t,x,y,\mathfrak{z}) - D_x f_i(t,\bar x,y,\mathfrak{z})| = 0$$

for every $R > 0$.

Assumption V.2.5. Let $m \geqslant 0$ be an integer.

(i) The partial derivatives in $x \in \mathbb{R}^d$ of the coefficients b, B, σ, ρ, (ρB), η and ξ up to order $m+1$ are functions such that

$$\sum_{k=1}^{m+1} |D_x^k(b, B, \sigma, \rho, (\rho B))| \leqslant L \quad \text{for all } t \in [0,T],\ x \in \mathbb{R}^d,\ y \in \mathbb{R}^{d'}.$$

(ii) Moreover,

$$\sum_{k=1}^{m+1} |D_x^k \eta| \leqslant L\bar{\eta}, \quad \sum_{k=1}^{m+1} |D_x^k \xi| \leqslant L\bar{\xi},$$

for all $t \in [0,T]$, $x \in \mathbb{R}^d$, $y \in \mathbb{R}^{d'}$ and $\mathfrak{z}_i \in \mathfrak{Z}_i, i = 0, 1$.

Remark V.2.1. Note that Assumption V.2.4(i), together with Assumptions V.2.2 and V.2.1(i), implies that for a constant $c = c(\lambda, K_\xi, K_\eta)$ we have for all $\theta \in [0,1]$, $y \in \mathbb{R}^{d'}$, $t \in [0,T]$ and $\mathfrak{z}_i \in \mathfrak{Z}_i$, $i = 0, 1$,

$$c^{-1}|x - \bar{x}| \leqslant |x - \bar{x} + \theta(f_i(t, x, y, \mathfrak{z}_i) - f(t, \bar{x}, y, \mathfrak{z}_i))| \leqslant c|x - \bar{x}| \quad \text{for } x, \bar{x} \in \mathbb{R}^d,$$

with $f_0(t, x, y, \mathfrak{z}_0) := \eta(t, x, y, \mathfrak{z}_0)$ and $f_1(t, x, y, \mathfrak{z}_1) := \xi(t, x, y, \mathfrak{z}_1)$. This, together with Assumption V.2.4(iii) in particular implies that for all $\theta \in [0,1]$, $y \in \mathbb{R}^{d'}$, $t \in [0,T]$ and $\mathfrak{z}_i \in \mathfrak{Z}_i$, $i = 0, 1$ the mappings

$$\tau^\eta(x) = x + \theta\eta(t, x, y, \mathfrak{z}_0) \quad \text{and} \quad \tau^\xi(x) = x + \theta\xi(t, x, y, \mathfrak{z}_1)$$

are C^1-diffeomorphisms.

Let $\mathcal{F}_t^Y$ denote the completion of the σ-algebra generated by $(Y_s)_{s \leqslant t}$.

Theorem V.2.1. *Let Assumptions V.2.1, V.2.2, V.2.4 and V.2.5 hold. If $K_1 \neq 0$ in Assumption V.2.1, then let additionally Assumption V.2.3 hold. Assume the conditional density $\pi_0 = P(X_0 \in dx|\mathcal{F}_0^Y)/dx$ exists almost surely and for some $p \geqslant 2$ and integer $m \geqslant 0$ we have $\mathbb{E}|\pi_0|^p_{W_p^m} < \infty$. Then almost surely $P(X_t \in dx|\mathcal{F}_t^Y)/dx$ exits and belongs to W_p^m for every $t \in [0,T]$.*
Moreover, there is an W_p^m-valued weakly cadlag process $\pi = (\pi_t)_{t \in [0,T]}$ such that for each t almost surely $\pi_t = P(X_t \in dx|\mathcal{F}_t^Y)/dx$. If $K_1 = 0$ and $m \geqslant 1$, then π is strongly cadlag as W_p^s-valued process for $s \in [0,m)$.

V.3 Preliminaries

As this chapter is a direct continuation of Chapters III and IV, we ask the reader to recall their main results, which we rely on in the following sections, as well as the notions of solutions to the Zakai equation, which we introduced earlier. A summary of the aforementioned matters was provided in Section IV.3.

We also recall that that if the unnormalised conditional distribution μ_t has a density such that $u_t = d\mu_t/dx$ (a.s.) for each $t \in [0,T]$ for an L_p-valued weakly cadlag process $(u_t)_{t\in[0,T]}$ for some $p \geqslant 2$, then it satisfies for each $\varphi \in C_0^\infty$ almost surely

$$\begin{aligned}(u_t,\varphi) =&(\psi,\varphi) + \int_0^t (u_s, \tilde{\mathcal{L}}_s\varphi)\,ds + \int_0^t (u_s, \mathcal{M}_s^k\varphi)\,dV_s^k + \int_0^t\int_{\mathfrak{Z}_0}(u_s, J_s^\eta\varphi)\,\nu_0(d\mathfrak{z})ds \\ &+ \int_0^t\int_{\mathfrak{Z}_1}(u_s, J_s^\xi\varphi)\,\nu_1(d\mathfrak{z})ds + \int_0^t\int_{\mathfrak{Z}_1}(u_{s-}, I_s^\xi\varphi)\,\tilde{N}_1(d\mathfrak{z},ds), \quad t\in[0,T].\end{aligned} \tag{V.3.1}$$

for all $t \in [0,T]$, which formally, we may write as the Cauchy problem

$$\begin{aligned} du_t =&\tilde{\mathcal{L}}_t^* u_t\,dt + \mathcal{M}_t^{*k}u_t\,dV_t^k + \int_{\mathfrak{Z}_0} J_t^{\eta*}u_t\,\nu_0(d\mathfrak{z})dt \\ &+ \int_{\mathfrak{Z}_1} J_t^{\xi*}u_t\,\nu_1(d\mathfrak{z})dt + \int_{\mathfrak{Z}_1} I_t^{\xi*}u_{t-}\,\tilde{N}_1(d\mathfrak{z},dt), \qquad \text{(V.3.2)} \\ u_0 =&\psi. \end{aligned}$$

for a given ψ.

Additional to the concept of L_p-solution defined in Definition IV.3.2, we introduce the following.

Definition V.3.1. Let integers $m \geqslant 0$ and $p \geqslant 2$. Let ψ be an W_p^m-valued $\mathcal{F}_0$-measurable random variable. Then we say that a W_p^m-valued $\mathcal{F}_t$-adapted weakly cadlag process $(u_t)_{t\in[0,T]}$ is a W_p^m-solution of (V.3.2) with initial condition ψ, if for each $\varphi \in C_0^\infty$ almost surely (V.3.1) holds for every $t \in [0,T]$.

Notice that for $m = 0$, a W_p^0-solution is the same as an L_p-solution. We summarize some important results from Chapter IV

As in Chapter IV, we are interested in solutions that satisfy

$$\operatorname*{ess\,sup}_{t\in[0,T]} |u_t|_{L_1} < \infty \quad \text{and} \quad \sup_{t\in[0,T]}\int_{\mathbb{R}^d} |y|^2|u_t(y)|\,dy < \infty \quad \text{(a.s.)}. \tag{V.3.3}$$

For the following theorem we denote again by $(\mu_t)_{t\in[0,T]}$ and $(P_t)_{t\in[0,T]}$ the unnormalised and normalised conditional distribution, respectively, of X given $\mathcal{F}^Y$ from Theorem III.1.1.

Theorem V.3.1. *Let Assumptions V.2.1, V.2.2 and V.2.4 hold. If $K_1 \neq 0$, then let additionally Assumption V.2.3 hold for some $r > 2$. Assume the conditional density $\pi_0 = dP_0/dx$ exists almost surely and $\mathbb{E}|\pi_0|_{L_p}^p < \infty$ for some $p \geqslant 2$.*
(i) The unnormalized conditional density $(u_t)_{t\in[0,T]}$ exists almost surely and is an L_p-valued weakly cadlag process such that for each $t \in [0,T]$ almost surely $u_t = d\mu_t/dx$ and

$$\mathbb{E}\sup_{t\in[0,T]}|u_t|_{L_p}^p \leqslant N\mathbb{E}|\pi_0|_{L_p}^p.$$

for a constant $N = N(d, d', p, K, K_\xi, K_\eta, L, T, \lambda, |\bar{\xi}|_{L_2}, |\bar{\eta}|_{L_2})$. Moreover, u is the unique L_2-solution to (V.3.2) *satisfying the conditions in* (V.3.3).
(ii) For each $t \in [0, T]$ the conditional density dP_t/dx exists and belongs to L_p almost surely. Moreover, there is an L_p-valued weakly cadlag process $(\pi_t)_{t\in[0,T]}$ such that for each $t \in [0, T]$ almost surely $\pi_t = dP_t/dx$ and almost surely $\pi_t = u_t{}^o\gamma_t$ for all $t \in [0, T]$, where $({}^o\gamma_t)_{t\in[0,T]}$ is the optional projection of $(\gamma_t)_{t\in[0,T]}$ under P with respect to $(\mathcal{F}^Y_t)_{t\in[0,T]}$.

Proof. See Lemma IV.5.7 and Theorem IV.2.1. □

Lemma V.3.2. *Let $1 < p < \infty$ and let $(v_t)_{t\in[0,T]}$ be a weakly cadlag L_p-valued process. Assume moreover that for an $m \geqslant 0$ almost surely* $\operatorname{ess\,sup}_{t\in[0,T]} |v_t|_{W^m_p} < \infty$ *and $v_T \in W^m_p$. Then v is weakly cadlag as W^m_p-valued process.*

Proof. Let Ω' be the set of those $\omega \in \Omega$ such that $(v_t(\omega))_{t\in[0,T]}$ is weakly cadlag as an L_p-valued function, $v_T(\omega) \in W^m_p$ and $\operatorname{ess\,sup}_{t\in[0,T]} |v_t(\omega)|_{W^m_p} < \infty$. Then $P(\Omega') = 1$, and for each $\omega \in \Omega'$ there exists a dense subset $\mathbb{T}_\omega$ in $[0, T]$ such that $\sup_{t\in\mathbb{T}_\omega} |v_t(\omega)|_{W^m_p} < \infty$. If $\omega \in \Omega'$ and $t \notin \mathbb{T}_\omega$, $t \neq T$, then there exists a sequence $(t_n)_{n=1}^\infty \subset \mathbb{T}_\omega$ such that $t_n \downarrow t$. Since $\sup_{t\in\mathbb{T}_\omega} |v_t(\omega)|_{W^m_p} < \infty$ there exists a subsequence, also denoted by $(t_n)_{n=1}^\infty$, such that $v_{t_n}(\omega)$ converges weakly in W^m_p to some element $\tilde{v} \in W^m_p$. However, as v is weakly cadlag as an L_p-valued process, we know that $v_{t_n} \to v_t$ weakly in L_p as $n \to \infty$ and hence $\tilde{v} = v_t \in W^m_p$. Thus clearly also $\sup_{t\in[0,T]} |v_t(\omega)|_{W^m_p} < \infty$ if $\omega \in \Omega'$. To see that v is weakly cadlag as a W^m_p-valued process, note first that since W^m_p is a reflexive space, which is embedded continuously and densely into L_p, we have that the dual $(L_p)^* = L_q$, $q = p/(p-1)$, is embedded continuously and densely into $(W^m_p)^*$. Therefore, for each $\varepsilon > 0$ and $\phi \in (W^m_p)^*$ there is an $\phi_\varepsilon \in L_q$ such that $|\phi - \phi_\varepsilon|_{(W^m_p)^*} < \varepsilon$. Fix a $t \in [0, T)$ and a sequence $t_n \downarrow t$. Then

$$
\begin{aligned}
|(v_{t_n}, \phi) - (v_t, \phi)| &\leqslant |(v_{t_n}, \phi - \phi_\varepsilon)| + |(v_{t_n}, \phi_\varepsilon) - (v_t, \phi_\varepsilon)| + |(v_t, \phi_\varepsilon - \phi)| \\
&\leqslant 2\varepsilon \sup_{t\in[0,T]} |v_t|_{W^m_p} + |(v_{t_n}, \phi_\varepsilon) - (v_t, \phi_\varepsilon)|.
\end{aligned}
$$

Recalling that v is weakly cadlag as an L_p-valued process finishes the proof. □

In this following we show that, for $p \geqslant 2$ and integers $m \geqslant 1$, if $K_1 = 0$ in Assumption V.2.1, then a W^m_p-solution of (V.3.2) is strongly cadlag as L_p-valued process, using an Itô formula from [21], see Theorem 2.2 therein. For this purpose we first rewrite the Zakai equation (V.3.2) in the form used in [21] to apply the Itô formula proved therein. To derive the required form of the J operators, consider on a measurable σ-finite space $(\mathfrak{Z}, \mathcal{Z}, \nu)$ a function $\zeta : \mathbb{R}^d \times \mathfrak{Z} \to \mathbb{R}^d$, smooth in $x \in \mathbb{R}^d$ such that for all $\theta \in [0, 1]$, $\mathfrak{z} \in \mathfrak{Z}$ the mapping

$$\tau_{\theta,\mathfrak{z}}(x) := \tau^\zeta_{\theta,\mathfrak{z}}(x) := x + \theta\zeta(x, \mathfrak{z}), \quad x \in \mathbb{R}^d,$$

is a C^1-diffeomorphism on $\mathbb{R}^d$. Then, for smooth functions φ and ϕ on $\mathbb{R}^d$,

$$\big(\phi, \int_{\mathfrak{Z}} J^\zeta \varphi(x)\, \nu(d\mathfrak{z})\big) = \int_{\mathfrak{Z}} \int_{\mathbb{R}^d} \phi(x) J^\zeta \varphi(x)\, dx\, \nu(d\mathfrak{z})$$

$$= \int_{\mathfrak{Z}} \int_{\mathbb{R}^d} \int_0^1 \phi(x)(1-\theta)(D_{ij}\varphi)(x+\theta\zeta(x,\mathfrak{z}))\zeta^i(x,\mathfrak{z})\zeta^j(x,\mathfrak{z})\, d\theta\, dx\, \nu(d\mathfrak{z})$$

$$= \int_{\mathfrak{Z}} \int_{\mathbb{R}^d} \int_0^1 \phi(\tau^{-1}_{\theta,\mathfrak{z}}(x))(1-\theta) D_{ij}\varphi(x)\zeta^i(\tau^{-1}_{\theta,\mathfrak{z}}(x),\mathfrak{z})\zeta^j(\tau^{-1}_{\theta,\mathfrak{z}}(x),\mathfrak{z})|\det D\tau^{-1}_{\theta,\mathfrak{z}}(x)|\, d\theta\, dx\, \nu(d\mathfrak{z}).$$

Hence, integrating by parts, we get that

$$\big(\phi, \int_{\mathfrak{Z}} J^\zeta \varphi(x)\, \nu(d\mathfrak{z})\big) = (\tilde{\mathcal{J}}^{\zeta,i}\phi, D_i\varphi) + (\hat{\mathcal{J}}^{\zeta,i}\phi, D_i\varphi)$$

where

$$\tilde{\mathcal{J}}^{\zeta,i}\phi(x) = -\int_{\mathfrak{Z}} \int_0^1 (D_k\phi)(\tau^{-1}_{\theta,\mathfrak{z}}(x)) D_j(\tau^{-1}_{\theta,\mathfrak{z}}(x))^k (1-\theta) \qquad \text{(V.3.4)}$$
$$\zeta^i(\tau^{-1}_{\theta,\mathfrak{z}}(x),\mathfrak{z})\zeta^j(\tau^{-1}_{\theta,\mathfrak{z}}(x),\mathfrak{z})|\det D\tau^{-1}_{\theta,\mathfrak{z}}(x)|\, d\theta\, \nu(d\mathfrak{z}),$$

$$\hat{\mathcal{J}}^{\zeta,i}\phi(x) = -\int_{\mathfrak{Z}} \int_0^1 \phi(\tau^{-1}_{\theta,\mathfrak{z}}(x))(1-\theta)$$
$$D_j\big[\zeta^i(\tau^{-1}_{\theta,\mathfrak{z}}(x),\mathfrak{z})\zeta^j(\tau^{-1}_{\theta,\mathfrak{z}}(x),\mathfrak{z})|\det D\tau^{-1}_{\theta,\mathfrak{z}}(x)|\big]\, d\theta\, \nu(d\mathfrak{z}).$$

Moreover, we ask the reader to recall the form of $I^{\xi *}$, the adjoint operator to I^ξ, derived in (IV.6.63).

Lemma V.3.3. *Let Assumptions V.2.1, V.2.2, V.2.4 and V.2.5 hold with $K_1 = 0$ and some integer $m \geqslant 1$. Let $p \geqslant 2$ and let $u = (u_t)_{t\in[0,T]}$ be a W^m_p-solution to* (V.3.1)*. Then u is strongly cadlag as an L_p-valued process.*

Proof. We apply Theorem 2.2 in [21]. In order to do so, we rewrite equation (V.3.1) into the form used therein. For that purpose, note first that the adjoint operator to $\mathcal{M}^k$, $k = 1, \dots, d'$ for $v \in W^m_p$ is

$$\mathcal{M}^{*k}_s v(x) = -D_i(\rho^{ik}_s(x)v(x)) + B^k_s(x)v(x).$$

Similarly, with $\beta_s = B_s(X_s)$ we notice that with $\varphi \in C^\infty_0$,

$$(v, \tilde{\mathcal{L}}_s\varphi) = -(D_j(a^{ij}_s v), D_i\varphi) - (D_i(b^i_s v), \varphi) + \beta^k_s(\mathcal{M}^{*k}_s v, \varphi)$$

Since η and ξ satisfy Assumption V.2.5 with $m = 1$, we can define the operators $\tilde{\mathcal{J}}^{\xi,i}_s, \hat{\mathcal{J}}^{\xi,i}_s$ and $\tilde{\mathcal{J}}^{\eta,i}_s, \hat{\mathcal{J}}^{\eta,i}_s$, $i = 1, \dots, d$, as in (V.3.4), only with ξ_s and η_s in place of ζ, accordingly, with the C^2-diffeomorphisms on $\mathbb{R}^d$ (for $\omega \in \Omega, t \in [0,T], \theta \in$

$[0,1]$, $\mathfrak{z}_i \in \mathfrak{Z}_i$ for $i = 0, 1)$

$$\tau^{\eta}_{t,\theta,\mathfrak{z}_0}(x) := x + \theta\eta_t(x, \mathfrak{z}_0) \quad \text{and} \quad \tau^{\xi}_{t,\theta,\mathfrak{z}_1}(x) := x + \theta\xi_t(x, \mathfrak{z}_1).$$

From (IV.6.63) in Chapter IV we know that with the mappings (note that $\theta = 1$ below)

$$\xi^*_t(x, \mathfrak{z}_1) := -x + (\tau^{\xi}_{t,1,\mathfrak{z}_1})^{-1}(x) \quad \text{and} \quad \mathfrak{c}_t(x, \mathfrak{z}_1) = \det(\mathbb{I} + D\xi^*_t(x, \mathfrak{z}_1)) - 1$$

and the translation operator $T^{\xi^*}_t v(x) = v(x + \xi^*_t(x, \mathfrak{z}_1))$, the adjoint operator to I^{ξ}_t is

$$I^{\xi *}_t v(x) = I^{\xi^*}_t v(x) + \mathfrak{c}_t(x, \mathfrak{z}_1) T^{\xi^*}_t v(x), \quad v \in W^m_p. \tag{V.3.5}$$

Then, rewriting equation (V.3.1) yields that for all $\varphi \in C^\infty_0$ almost surely

$$\begin{aligned}(u_t, \varphi) = (\psi, \varphi) - \int_0^t (D_j(a^{ij}_s u_s), D_i\varphi)\, ds - \int_0^t (D_i(b^i_s u_s) + \beta^k_s \mathcal{M}^{*k}_s u_s, \varphi)\, ds \\ + \int_0^t (\mathcal{M}^{*k}_s u_s, \varphi)\, dV^k_s + \int_0^t (\tilde{\mathcal{J}}^{\eta,i}_s u_s + \hat{\mathcal{J}}^{\eta,i} u_s, D_i\varphi)\, ds \\ + \int_0^t (\tilde{\mathcal{J}}^{\xi,i}_s u_s + \hat{\mathcal{J}}^{\xi,i} u_s, D_i\varphi)\, ds + \int_0^t \int_{\mathfrak{Z}_1} (I^{\xi *}_s u_s, \varphi)\, \tilde{N}_1(d\mathfrak{z}, ds)\end{aligned}$$

for all $t \in [0, T]$. Hence, to apply Theorem 2.2 from [21], it remains to verify that almost surely, for $i = 1, \dots, d$,

$$A_D := \int_0^T \int_{\mathbb{R}^d} |D_i(a^{ij}_s u_s)|^p\, dxds < \infty, \quad A_M := \int_0^T \int_{\mathbb{R}^d} |-D_i(b^i_s u_s) + \beta^k_s \mathcal{M}^{*k}_s u_s|^p\, dxds < \infty,$$

$$A^i_\eta := \int_0^T \int_{\mathbb{R}^d} |\tilde{\mathcal{J}}^{\eta,i}_s u_s + \hat{\mathcal{J}}^{\eta,i}_s u_s|^p\, dxds < \infty, \quad A^i_\xi := \int_0^T \int_{\mathbb{R}^d} |\tilde{\mathcal{J}}^{\xi,i}_s u_s + \hat{\mathcal{J}}^{\xi,i}_s u_s|^p\, dxds < \infty,$$

$$B := \int_0^T \int_{\mathbb{R}^d} \big(\sum_k |(\mathcal{M}^{k*}_s u_s)(x)|^2\big)^{p/2}\, dxds < \infty,$$

$$G := \int_0^T \int_{\mathbb{R}^d} \int_{\mathfrak{Z}_1} |(I^{\xi *}_s u_s)(x, \mathfrak{z})|^p\, \nu_1(d\mathfrak{z}) dxds < \infty,$$

$$H := \int_0^T \int_{\mathbb{R}^d} \Big(\int_{\mathfrak{Z}_1} |(I^{\xi *}_s u_s)(x, \mathfrak{z})|^2\, \nu_1(d\mathfrak{z}) \Big)^{p/2} dxds < \infty.$$

By the boundedness of the coefficients and their derivatives up to order $m + 1$ and due to $\int_0^T |u_s|^p_{W^1_p}\, ds < \infty$ (a.s.) clearly $A_D + A_M + B < \infty$ (a.s.). Suppressing the superscript i,

$$A_\eta \leqslant \int_0^T \int_{\mathbb{R}^d} |\tilde{\mathcal{J}}^{\eta}_s u_s|^p\, dx\, ds + \int_0^T \int_{\mathbb{R}^d} |\hat{\mathcal{J}}^{\eta}_s u_s|^p\, dx\, ds = A_{\eta,1} + A_{\eta,2}.$$

Using that by Lemma IV.6.1 (i) & (ii), for all $\omega \in \Omega, t \in [0,T], \mathfrak{z} \in \mathfrak{Z}_0$ and $x \in \mathbb{R}^d$,

$$(1-\theta)|D(\tau^{\eta}_{t,\theta,\mathfrak{z}})^{-1}(x))||\eta((\tau^{\eta}_{t,\theta,\mathfrak{z}})^{-1}(x),\mathfrak{z})|^2|\det D(\tau^{\eta}_{t,\theta,\mathfrak{z}})^{-1}(x)| \leqslant N\bar{\eta}^2(\mathfrak{z})$$

for a constant $N = N(d, K_0, L, K_\eta, \lambda)$, we can apply Minkowski's inequality to compute

$$A_{\eta,1} \leqslant N' \int_0^T \Big(\int_{\mathfrak{Z}_0} \int_0^1 \Big(\int_{\mathbb{R}^d} |(Du_s)(\tau^{-1}_{s,\theta,\mathfrak{z}}(x))|^p \bar{\eta}^{2p}(\mathfrak{z})\, dx \Big)^{1/p} d\theta\, \nu_0(d\mathfrak{z}) \Big)^p ds$$

$$\leqslant N'' \int_0^T |u_s|^p_{W^1_p}\, ds < \infty \quad \text{(a.s.)},$$

for constants N' and N'' depending on $d, p, K_0, K_\eta, \lambda$ and $|\bar{\eta}|_{L_2}$. Using that by the Lemma's assumption, together with Lemma IV.6.1, for all $s \in [0,T], \theta \in [0,1], x \in \mathbb{R}^d$ and $\mathfrak{z} \in \mathfrak{Z}_0$,

$$\big|D\big[\eta^i_s(\tau^{-1}_{s,\theta,\mathfrak{z}}(x),\mathfrak{z})\eta^j_s(\tau^{-1}_{s,\theta,\mathfrak{z}}(x),\mathfrak{z})|\det D\tau^{-1}_{s,\theta,\mathfrak{z}}(x)|\big]\big| \leqslant N\bar{\eta}^2(\mathfrak{z}), \quad i,j = 1,\ldots,d,$$

for a constant $N = N(d, K_0, K_\eta, \lambda, L)$, we get $A_{\eta,2} < \infty$ in the same way, proving $A^i_\eta < \infty$ almost surely, $i = 1,\ldots,d$. Analogously also $A^i_\xi < \infty$ almost surely, $i = 1,\ldots,d$. Next, using the form (V.3.5),

$$\begin{aligned} G \leqslant& \int_0^T \int_{\mathbb{R}^d} \int_{\mathfrak{Z}_1} |(I^{\xi^*}_s u_s)(x,\mathfrak{z})|^p\, \nu_1(d\mathfrak{z})dxds \\ &+ \int_0^T \int_{\mathbb{R}^d} \int_{\mathfrak{Z}_1} |\mathfrak{c}_s(x,\mathfrak{z})(T^{\xi^*}_s u_s)(x,\mathfrak{z})|^p\, \nu_1(d\mathfrak{z})dxds \\ =:& G_1 + G_2. \end{aligned}$$

To treat the first term, we define

$$\tau^{\eta^*}_{t,\mathfrak{z}_0,\theta}(x) = x + \theta\eta^*_t(x,\mathfrak{z}_0), \quad \text{for } \omega \in \Omega, x \in \mathbb{R}^d, t \in [0,T], \mathfrak{z}_0 \in \mathfrak{Z}_0 \text{ and } \theta \in [0,1],$$

and use that by Lemma IV.6.1 (ii) almost surely

$$|\det(\tau^{\eta^*}_{s,\mathfrak{z}_0,\theta})^{-1}(x)| \leqslant |\det(\mathbb{I} + D\eta_t(x,\mathfrak{z}))|^{-1} < \infty$$

for $\omega \in \Omega, x \in \mathbb{R}^d, t \in [0,T], \mathfrak{z}_0 \in \mathfrak{Z}_0$ and $\theta \in [0,1]$. Thus, by Taylor's theorem, Minkowksi's inequality and Lemma IV.6.1,

$$G_1 \leqslant \int_0^T \Big(\int_0^1 \Big(\int_{\mathbb{R}^d} \int_{\mathfrak{Z}_1} |(Du_s)(\tau^{\xi^*}_{s,\theta,\mathfrak{z}}(x))|^p |\xi^*_s(x,\mathfrak{z})|^p\, \nu(d\mathfrak{z})dx \Big)^{1/p} d\theta \Big)^p ds$$

$$\leqslant N \int_0^T |u_s|^p_{W^1_p}\, ds < \infty$$

almost surely, for a constant $N = N(d, p, \lambda, K_0, K_\xi, |\bar{\xi}|_{L_2})$. Using that by Lemma IV.6.1 $|\mathfrak{c}_s(x, \mathfrak{z})| \leqslant N\bar{\xi}(\mathfrak{z})$ with a constant $N = N(d, L, \lambda, K_0, K_\xi)$ we get $G_2 < \infty$ (a.s.) in the same way, proving $G < \infty$ (a.s.). Using Taylor's theorem and Minkowski's inequality in a very similar way we also obtain $H < \infty$ almost surely. Therefore we can apply Theorem 2.2 in [21] to know that there exists a stochastic modification $\bar{u}$ of u, such that for $P \otimes dt$-a.e. $(\omega, t) \in \Omega \times [0, T]$ we have $\bar{u}_t = u_t$ and such that $\bar{u}$ is strongly cadlag as L_p-valued process. Finally, by Lemma V.3.2 $\bar{u}$ is also weakly cadlag as W^m_p-valued process, and therefore we have that almost surely $u_t = \bar{u}_t$ for all $t \in [0, T]$, i.e., u and $\bar{u}$ are almost surely identical. This finishes the proof. $\square$

V.4 Sobolev estimates

Here we present some estimates which are needed in the subsequent sections. We use the same notations that were introduced in Section IV.4, in particular the Gaussian density function k_ε on $\mathbb{R}^d$ with mean 0 and variance ε, as well as the function

$$\rho_\varepsilon(y) = \int_{\mathbb{R}^d} \Pi_{r=1}^p k_\varepsilon(x - y_r)\, dx = c_{p,\varepsilon} e^{-\sum_{1 \leqslant r < s \leqslant p} |y_r - y_s|^2/(2\varepsilon p)}, \quad y = (y_1, ..., y_p) \in \mathbb{R}^{pd} \tag{V.4.1}$$

for $\varepsilon > 0$, with the constant $c_{p,\varepsilon} = c_{p,\varepsilon}(d) = p^{-d/2}(2\pi\varepsilon)^{(1-p)d/2}$. We recall that ρ_ε satisfies, for every $r = 1, 2, ..., p$ and $i = 1, 2, ..., d$

$$\partial_{y^i_r} \rho_\varepsilon(y) = \tfrac{1}{\varepsilon p} \sum_{s=1}^p (y^i_s - y^i_r) \rho_\varepsilon(y), \quad y = (y_1, ..., y_p) \in \mathbb{R}^d, \quad y_r = (y^1_r ..., y^d_r) \in \mathbb{R}^d, \tag{V.4.2}$$

$$\partial_{y^j_r} \rho_\varepsilon(y) = -\sum_{s \neq r}^p \partial_{y^j_s} \rho_\varepsilon(y) \quad \text{for } r = 1, ..., p \text{ and } j = 1, 2, ..., d, \tag{V.4.3}$$

as well as, for $q = 1, 2$ with a constant $N = N(d, p, q)$

$$\varepsilon^{-q} \sum_{s \neq r} |y_s - y_r|^{2q} \rho_\varepsilon(y) \leqslant N \rho_{2\varepsilon}(y), \quad y \in \mathbb{R}^{pd}. \tag{V.4.4}$$

The case of $\alpha = 0$ in the following Lemmas in this section are proven in Section IV.4 and hence this case will be omitted in the proofs.

In the following we present estimates for $\mu \in \mathfrak{M}$ with a density $u = d\mu/dx \in W^m_p$, for $m \geqslant 0$ and $p \geqslant 2$ even. In order for the left-hand side of these estimates to be well-defined, we require that

$$K_1 \int_{\mathbb{R}^d} |x|^2 \, |u(x)| \, dx < \infty, \tag{V.4.5}$$

where we use the formal convention that $0 \cdot \infty = 0$, i.e. if $K_1 = 0$, then the second moment of $|\mu(dx)| = |u(x)|dx$ is not required to be finite.

Lemma V.4.1. *Consider integers $m \geqslant 0$ and $p \geqslant 2$ even. Let $\sigma = (\sigma^{ik})$ be a Borel function on $\mathbb{R}^d$ with values in $\mathbb{R}^{d\times k}$, such that for some nonnegative constants K_0 and L*

$$|\sigma(x)| \leqslant K_0, \quad \sum_{k=1}^{m+1} |D^k\sigma(x)| \leqslant L, \tag{V.4.6}$$

for all $x, y \in \mathbb{R}^d$. Set $a^{ij} = \sigma^{ik}\sigma^{jk}/2$ for $i, j = 1, 2, ..., d$. Let $\mu \in \mathfrak{M}$ such that it admits a density $u = d\mu/dx \in W_p^m$ which satisfies (V.4.5)*. Then for $\varepsilon > 0$ we have*

$$\begin{aligned} A^\alpha :=& p((D^\alpha\mu^{(\varepsilon)})^{p-1}, D^\alpha((a^{ij}D_{ij})^*\mu)^{(\varepsilon)}) \\ &+ \tfrac{p(p-1)}{2}((D^\alpha\mu^{(\varepsilon)})^{p-2}D^\alpha((\sigma^{ik}D_i)^*\mu)^{(\varepsilon)}, D^\alpha((\sigma^{jk}D_j)^*\mu)^{(\varepsilon)}) \leqslant NL^2|u|^p_{W_p^m} \end{aligned} \tag{V.4.7}$$

for multi-indices $\alpha = (\alpha_1, ..., \alpha_d)$ such that $0 \leqslant |\alpha| \leqslant m$, where N is a constant depending only on d, m and p.

Proof. Note first that using

$$\sup_{x\in\mathbb{R}^d} \sum_{k=0}^{m+2} |D^k k_\varepsilon(x)| < \infty, \quad \sup_{x\in\mathbb{R}^d} \sum_{k=0}^{m+2} |D^k \rho_\varepsilon(x)| < \infty, \quad \text{for all } \varepsilon > 0 \tag{V.4.8}$$

and

$$\int_{\mathbb{R}^d} (1 + |x| + |x|^2)\, |u(x)|(dx) < \infty, \tag{V.4.9}$$

as well as the conditions on σ, it is easy to verify that the left-hand side of (V.4.7) is well-defined. Changing the order of taking derivatives and integrals, then writing integer powers of integrals as iterated integrals and using

$$D_x^\alpha k_\varepsilon(x-y) = (-1)^{|\alpha|} D_y^\alpha k_\varepsilon(x-y),$$

we have

$$\begin{aligned} ((D^\alpha\mu^{(\varepsilon)}(x))^{p-1} &= \int_{\mathbb{R}^{(p-1)d}} \Pi_{r=1}^{p-1} D_x^\alpha k_\varepsilon(x-y_r)\, \mu(dy_1)...\mu(dy_{p-1}) \\ &= \int_{\mathbb{R}^{(p-1)d}} (-1)^{(p-1)|\alpha|} D_{y_1}^\alpha ... D_{y_{p-1}}^\alpha \Pi_{r=1}^{p-1} k_\varepsilon(x-y_r)\, \mu(dy_1)...\mu(dy_{p-1}), \\ D^\alpha((a^{ij}D_{ij})^*\mu)^{(\varepsilon)}(x) &= \int_{\mathbb{R}^d} a^{ij}(y_p)\partial_{y_p^i}\partial_{y_p^j} D_x^\alpha k_\varepsilon(x-y_p)\, \mu(dy_p) \\ &= \int_{\mathbb{R}^d} (-1)^{|\alpha|} a^{ij}(y_p)\partial_{y_p^i}\partial_{y_p^j} D_{y_p}^\alpha k_\varepsilon(x-y_p)\, \mu(dy_p), \end{aligned}$$

and hence for their product we get

$$(D^\alpha \mu^{(\varepsilon)})^{p-1} D^\alpha ((a^{ij} D_{ij})^* \mu)^{(\varepsilon)}(x) = \int_{\mathbb{R}^{pd}} a^{ij}(y_p) \partial_{y_p^i} \partial_{y_p^j} D_y^{p\alpha} \Pi_{r=1}^p k_\varepsilon (x - y_r) \mu_p(dy), \tag{V.4.10}$$

where $D_y^{p\alpha} := D_{y^1}^\alpha ... D_{y_p}^\alpha$ and $\mu(dy) := \mu(dy_1)...\mu(dy_p)$. Similarly,

$$(D^\alpha \mu^{(\varepsilon)})^{p-2} D^\alpha ((\sigma^{ik} D_i)^* \mu)^{(\varepsilon)} D^\alpha ((\sigma^{jk} D_j)^* \mu)^{(\varepsilon)}(x)$$

$$= \int_{\mathbb{R}^{pd}} \sigma^{ik}(y_{p-1}) \sigma^{jk}(y_p) \partial_{y_{p-1}^i} \partial_{y_p^j} D_y^{p\alpha} \Pi_{r=1}^p k_\varepsilon (x - y_r) \mu_p(dy).$$

Adding this to (V.4.10), then integrating against dx over $\mathbb{R}^d$ and using (V.4.1) we obtain

$$A = \int_{\mathbb{R}^{pd}} \left(p a^{ij}(y_p) \partial_{y_p^i} \partial_{y_p^j} + \tfrac{p(p-1)}{2} \sigma^{ik}(y_{p-1}) \sigma^{jk}(y_p) \partial_{y_{p-1}^i} \partial_{y_p^j} \right) D_y^{p\alpha} \rho_\varepsilon(y) \, \mu_p(dy).$$

Using here the symmetry of $D_y^{p\alpha} \rho_\varepsilon(y)$ and $\mu_p(dy)$ in $y \in \mathbb{R}^{dp}$ and then interchanging differential operators we get

$$A = \int_{\mathbb{R}^{pd}} \Big(\sum_{r=1}^p a^{ij}(y_r) D_y^{p\alpha} \partial_{y_r^i} \partial_{y_r^j} + \sum_{1 \leqslant r < s \leqslant p} \sigma^{ik}(y_r) \sigma^{jk}(y_s) D_y^{p\alpha} \partial_{y_r^i} \partial_{y_s^j} \Big) \rho_\varepsilon(y) \, \mu_p(dy)$$

Using

$$\partial_{y_r^j} \rho_\varepsilon(y) = - \sum_{s \neq r} \partial_{y_s^j} \rho_\varepsilon(y),$$

see (V.4.3), we have

$$\sum_{r=1}^p a^{ij}(y_r) D_y^{p\alpha} \partial_{y_r^i} \partial_{y_r^j} \rho_\varepsilon(y) = - \sum_{1 \leqslant r < s \leqslant p} (a^{ij}(y_r) + a^{ij}(y_s)) D_y^{p\alpha} \partial_{y_r^i} \partial_{y_s^j} \rho_\varepsilon(y),$$

and due to $a^{ij} = \sigma^{ik} \sigma^{jk}/2$ we have

$$-2a^{ij}(y_r, y_s) := -2(a^{ij}(y_r) + a^{ij}(y_s)) + \sigma^{ik}(y_r) \sigma^{jk}(y_s) + \sigma^{ik}(y_s) \sigma^{jk}(y_r)$$

$$= -(\sigma^{ik}(y_r) - \sigma^{ik}(y_s))(\sigma^{jk}(y_r) - \sigma^{jk}(y_s)).$$

Hence

$$A = -\tfrac{1}{2} \sum_{r \neq s} \int_{\mathbb{R}^{pd}} a^{ij}(y_r, y_s) D_y^{p\alpha} \partial_{y_r^i} \partial_{y_s^j} \rho_\varepsilon(y) \, \mu_p(dy), \tag{V.4.11}$$

that by integration by parts gives

$$= -\tfrac{1}{2} \sum_{\beta \leqslant \alpha} \sum_{\gamma \leqslant \alpha} c_\beta^\alpha c_\gamma^\alpha \int_{\mathbb{R}^{pd}} \sum_{r \neq s} a_{\beta\gamma}^{ij}(y_r, y_s) \partial_{y_r^i} \partial_{y_s^j} \rho_\varepsilon(y) u_{\bar{\beta}}(y_r) u_{\bar{\gamma}}(y_s) \Pi_{q \neq r, q \neq s} u_\alpha(y_q) \, dy, \tag{V.4.12}$$

where $a^{ij}_{\beta\gamma}(x,r) := \partial^\beta_x\partial^\gamma_r a^{ij}(x,r)$ and $u_\delta(x) := \partial^\delta u(x)$ for $x, r \in \mathbb{R}^d$, for multi-indices β, γ and δ, $\bar{\delta} := \alpha - \delta$ for multi-indices $\delta \leqslant \alpha$ (i.e. $\delta_i \leqslant \alpha_i$ for $i = 1, 2, ..., d$), $c^\alpha_\delta = \Pi^d_{i=1} c^{\alpha_i}_{\delta_i}$ with binomial coefficients c^n_k for integers $0 \leqslant k \leqslant n$,

$$u(y) := u(y_1)....u(y_p) \quad \text{for } y = (y_1, ..., y_p) \in \mathbb{R}^{dp},$$

and $dy = dy_1...dy_p$ is the Lebesgue measure on $\mathbb{R}^{pd}$. For each $\beta \leqslant \alpha$ and $\gamma \leqslant \alpha$ we are going to estimate the integrand

$$f^{\beta\gamma}(y) := \sum_{r\neq s} a^{ij}_{\beta\gamma}(y_r, y_s)\partial_{y^i_r}\partial_{y^j_s}\rho_\varepsilon(y)u_{\bar{\beta}}(y_r)u_{\bar{\gamma}}(y_s)\Pi_{q\neq r, q\neq s}u_\alpha(y_q), \quad y \in \mathbb{R}^{dp}, \quad \beta, \gamma \leqslant \alpha$$

in the integral in (V.4.12). Because of the symmetry in β and γ, we need only consider the following cases: (i) $|\beta| \geqslant 1$ and $|\gamma| \geqslant 1$, (ii) $|\beta| \geqslant 1$ and $\gamma = 0$ and (iii) $\beta = \gamma = 0$. To proceed with the calculations in each of these cases, for functions $h = h(y)$ and $g = (y)$ of $y \in \mathbb{R}^{pd}$ we will use the notations $h \sim g$ if the integral of $g - h$ against dy over $\mathbb{R}^{pd}$ is zero. In case (i) by integration by parts we have

$$f^{\beta\gamma} \sim \sum_{j=1}^4 f^{\beta\gamma}_j$$

with

$$\begin{aligned}
f^{\beta\gamma}_1 &:= \sum_{r\neq s} \partial_{y^i_r}\partial_{y^j_s} a^{ij}_{\beta\gamma}(y_r, y_s)\rho_\varepsilon(y)u_{\bar{\beta}}(y_r)u_{\bar{\gamma}}(y_s)\Pi_{q\neq r, q\neq s}u_\alpha(y_q),\\
f^{\beta\gamma}_2 &:= \sum_{r\neq s} \partial_{y^j_s} a^{ij}_{\beta\gamma}(y_r, y_s)\rho_\varepsilon(y)\partial_{y^i_r}u_{\bar{\beta}}(y_r)u_{\bar{\gamma}}(y_s)\Pi_{q\neq r, q\neq s}u_\alpha(y_q),\\
f^{\beta\gamma}_3 &:= \sum_{r\neq s} \partial_{y^i_r} a^{ij}_{\beta\gamma}(y_r, y_s)\rho_\varepsilon(y)u_{\bar{\beta}}(y_r)\partial_{y^j_s}u_{\bar{\gamma}}(y_s)\Pi_{q\neq r, q\neq s}u_\alpha(y_q),\\
f^{\beta\gamma}_4 &:= \sum_{r\neq s} a^{ij}_{\beta\gamma}(y_r, y_s)\rho_\varepsilon(y)\partial_{y^i_r}u_{\bar{\beta}}(y_r)\partial_{y^j_s}u_{\bar{\gamma}}(y_s)\Pi_{q\neq r, q\neq s}u_\alpha(y_q).
\end{aligned} \tag{V.4.13}$$

It is easy to see that for $j = 1, 2, 3, 4$

$$|f^{\beta\gamma}_j(y)| \leqslant NL^2\rho_\varepsilon(y)\sum_{|\delta|\leqslant m}|u_\delta(y_1)|...\sum_{|\delta|\leqslant m}|u_\delta(y_p)|, \quad (y_1, y_2, ..., y_p) \in \mathbb{R}^{pd}$$

with a constant $N = N(d, m, p)$. Hence in the case (i) we get

$$\begin{aligned}
\int_{\mathbb{R}^{pd}} f^{\beta\gamma}(y)\,dy \leqslant& NL^2\int_{\mathbb{R}^{pd}}\sum_{|\delta|\leqslant m}|u_\delta(y_1)|...\sum_{|\delta|\leqslant m}|u_\delta(y_p)|\rho_\varepsilon(y)\,dy\\
=& NL^2\int_{\mathbb{R}^{pd}}\int_{\mathbb{R}^d}\sum_{|\delta|\leqslant m}|u_\delta(y_1)|...\sum_{|\delta|\leqslant m}|u_\delta(y_p)|\Pi^p_{r=1}k_\varepsilon(x - y_r)\,dx\,dy
\end{aligned}$$

$$\leqslant N'L^2 \sum_{|\delta|\leqslant m} ||D^\delta u|^{(\varepsilon)}|^p_{L_p}.$$

with constants N and N' depending only on d, m and p. Integrating by parts in the case (ii) we have

$$f^{\beta 0} \sim -f_1^{\beta 0} - f_2^{\beta 0}$$

with

$$f_1^{\beta 0} = \sum_{r\neq s} \partial_{y_r^i} a^{ij}_{\beta 0}(y_r,y_s)\partial_{y_s^j}\rho_\varepsilon(y)u_{\bar\beta}(y_r)\Pi_{q\neq r}u_\alpha(y_q)$$

$$f_2^{\beta 0} = \sum_{r\neq s} a^{ij}_{\beta 0}(y_r,y_s)\partial_{y_s^j}\rho_\varepsilon(y)\partial_{y_r^i}u_{\bar\beta}(y_r)\Pi_{q\neq r}u_\alpha(y_q).$$

Clearly, for $r\neq s$ we have

$$\partial_{y_r^i} a^{ij}_{\beta 0}(y_r,y_s) = g^{\beta,j}(y_r,y_s) + h^{\beta,j}(y_r),$$

with

$$\begin{aligned} g^{\beta,j}(y_r,y_s) =& \partial_{y_r^i}\partial^\beta_{y_r}\sigma^{ik}(y_r)(\sigma^{jk}(y_r)-\sigma^{jk}(y_s)) + \partial_{y_r^i}\partial^\beta_{y_r}\sigma^{jk}(y_r)(\sigma^{ik}(y_r)-\sigma^{ik}(y_s)),\\ h^{\beta,j}(y_r) =& \sum_{1\leqslant|\delta|,\delta<\beta(i)} c_\delta^{\beta(i)}\partial^\delta_{y_r}\sigma^{ik}(y_r)\partial^{\beta(i)-\delta}_{y_r}\sigma^{jk}(y_r), \end{aligned}$$

where the multi-index $\beta(i)$ is defined by $\partial^{\beta(i)} = \partial_{y_r^i}\partial^\beta_{y_r}$. Thus

$$f_1^{\beta 0} = f_{11}^{\beta 0} + f_{12}^{\beta 0}$$

with

$$\begin{aligned} f_{11}^{\beta 0} =& \sum_{r=1}^p\sum_{s\neq r} g^{\beta,j}(y_r,y_s)\partial_{y_s^j}\rho_\varepsilon(y)u_{\bar\beta}(y_r)\Pi_{q\neq r}u_\alpha(y_q),\\ f_{12}^{\beta 0} =& \sum_{r=1}^p\sum_{s\neq r} h^{\beta,j}(y_r)\partial_{y_s^j}\rho_\varepsilon(y)u_{\bar\beta}(y_r)\Pi_{q\neq r}u_\alpha(y_q). \end{aligned} \tag{V.4.14}$$

Since

$$|g^{\beta,j}(y_r,y_s)| \leqslant NL^2|y_r-y_s| \quad j=1,2,...,p,$$

for some $N=N(d,m,p)$, taking into account (V.4.2) we have

$$|g^{\beta,j}(y_r,y_s)\partial_{y_s^j}\rho_\varepsilon(y)| \leqslant \tfrac{N}{p\varepsilon}\sum_{1\leqslant k<l\leqslant p}^{p} |y_k-y_l|^2\rho_\varepsilon(y) \leqslant N'\rho_{2\varepsilon}(y)$$

and hence

$$|f_{11}^{\beta 0}| \leqslant N'\rho_{2\varepsilon}(y)\sum_{|\delta|\leqslant m}|u_\delta(y_1)|...\sum_{|\delta|\leqslant m}|u_\delta(y_p)|$$

with a constant $N' = N'(d,m,p)$. Remembering (V.4.3) by integration by parts we obtain

$$f_{12}^{\beta 0} = -\sum_{r=1}^{p} h^{\beta,j}(y_r)\partial_{y_r^j}\rho_\varepsilon(y)u_{\bar\beta}(y_r)\Pi_{q\neq r}u_\alpha(y_q) \sim f_{121}^{\beta 0} + f_{122}^{\beta 0}$$

with

$$f_{121}^{\beta 0} = \sum_{r=1}^{p} h^{\beta,j}(y_r)\rho_\varepsilon(y)\partial_{y_r^j}u_{\bar\beta}(y_r)\Pi_{q\neq r}u_\alpha(y_q),$$

$$f_{122}^{\beta 0} = \sum_{r=1}^{p} \partial_{y_r^j}h^{\beta,j}(y_r)\rho_\varepsilon(y)u_{\bar\beta}(y_r)\Pi_{q\neq r}u_\alpha(y_q).$$

Hence noting that

$$|h^{\beta,j}(y_r)| + |\partial_{y_r^j}h^{\beta,j}(y_r)| \leqslant NL^2$$

with a constant $N = N(d,m,p)$, we get

$$|f_{121}^{\beta 0} + f_{122}^{\beta 0}| \leqslant NL^2\rho_\varepsilon(y)\sum_{|\delta|\leqslant m}|u_\delta(y_1)|\ldots\sum_{|\delta|\leqslant m}|u_\delta(y_p)|$$

Consequently, for a constant $N' = N'(d,m,p)$,

$$\begin{aligned}\int_{\mathbb{R}^{pd}} f_1^{\beta 0}(y)\,dy \leqslant & NL^2\int_{\mathbb{R}^{pd}}\sum_{|\delta|\leqslant m}|u_\delta(y_1)|\ldots\sum_{|\delta|\leqslant m}|u_\delta(y_p)|\rho_{2\varepsilon}(y)\,dy \\ \leqslant & N'L^2\sum_{|\delta|\leqslant m}||D^\delta u|^{(2\varepsilon)}|_{L_p}^p \leqslant N'L^2\sum_{|\delta|\leqslant m}||D^\delta u|^{(\varepsilon)}|_{L_p}^p. \end{aligned} \tag{V.4.15}$$

Now we are going to estimate the integral of $f_2^{\beta 0}$. If $|\beta| = 1$, then

$$|a_{\beta 0}^{ij}(y_r,y_s)| \leqslant NL^2|y_r - y_s|,$$

and taking into account (V.4.2), we get

$$|f_2^{\beta 0}| \leqslant NL^2\rho_{2\varepsilon}(y)\sum_{|\delta|\leqslant m}|u_\delta(y_1)|\ldots\sum_{|\delta|\leqslant m}|u_\delta(y_p)|$$

with $N = N(d,p,m)$ in the same way as $|f_{11}|$ is estimated. Hence, as above,

$$\int_{\mathbb{R}^{pd}} f_2^{\beta 0}(y)\,dy \leqslant NL^2\sum_{|\delta|\leqslant m}||D^\delta u|^{(2\varepsilon)}|_{L_p}^p \leqslant NL^2\sum_{|\delta|\leqslant m}||D^\delta u|^{(\varepsilon)}|_{L_p}^p. \tag{V.4.16}$$

for $|\beta| = 1$. If $|\beta| \geqslant 2$, then

$$a_{\beta 0}^{ij}(y_r,y_s) = g^{\beta,ij}(y_r,y_s) + h^{\beta,ij}(y_r)$$

with

$$\begin{aligned}
g^{\beta,ij}(y_r,y_s) =&\partial^\beta_{y_r}\sigma^{ik}(y_r)(\sigma^{jk}(y_r)-\sigma^{jk}(y_s))+\partial^\beta_{y_r}\sigma^{jk}(y_r)(\sigma^{ik}(y_r)-\sigma^{ik}(y_s))\\
h^{\beta,ij}(y_r) =& \sum_{1\leqslant|\delta|,\delta<\beta} c^\beta_\delta \partial^\delta_{y_r}\sigma^{ik}(y_r)\partial^{\beta-\delta}_{y_r}\sigma^{jk}(y_r).
\end{aligned}$$

Noticing that for a constant $N=N(d,m,p)$,

$$\sum_{i,j}|g^{\beta,ij}(y_r,y_s)|\leqslant NL^2|y_r-y_s|$$

and

$$\sum_{ij}|h^{\beta,ij}(y_r)|+\sum_{ij}|\partial_{y^j_r}h^{\beta,ij}(y_r)|\leqslant NL^2,$$

we obtain (V.4.16) for $|\beta|\geqslant 2$ in the same way as the integral of $f_1^{\beta 0}$ is estimated. It remains to consider the case (iii), i.e., to estimate the integral of f^{00}. Since

$$|a^{ij}_{00}(y_r,y_s)|\leqslant NL^2|y_r-y_s|^2$$

with a constant $N=N(d,m,p)$ and

$$\partial_{y^i_r}\partial_{y^j_s}\rho_\varepsilon(y)=\tfrac{1}{p^2\varepsilon^2}\sum_{k=1}^p\sum_{l=1}^p(y^i_k-y^i_r)(y^j_l-y^j_s)+\frac{1}{p\varepsilon}\delta_{ij},$$

we have for a constant $N'=N'(d,m,p)$,

$$\begin{aligned}
|a^{ij}_{00}(y_r,y_s)\partial_{y^i_r}\partial_{y^j_s}\rho_\varepsilon(y)| \leqslant&\tfrac{N}{\varepsilon^2}L^2\sum_{1\leqslant k<l\leqslant p}|y_k-y_l|^4\rho_\varepsilon(y)+\tfrac{N}{\varepsilon}L^2\sum_{1\leqslant k<l\leqslant p}|y_k-y_l|^2\rho_\varepsilon(y)\\
\leqslant& N'L^2\rho_{2\varepsilon}(y)\quad\text{for } y=(y_1,...,y_p)\in\mathbb{R}^{pd}.
\end{aligned}$$

Hence

$$|f^{00}(y)|\leqslant NL^2\rho_{2\varepsilon}(y)\Pi^p_{r=1}|u_\alpha(y_r)|,$$

that gives

$$\int_{\mathbb{R}^{pd}}f^{00}(y)\,dy\leqslant NL^2||D^\alpha u|^{(2\varepsilon)}|^p_{L_p}\leqslant NL^2||D^\alpha u|^{(\varepsilon)}|^p_{L_p}$$

with a constant $N=N(d,m,p)$, and we finish the proof of (V.4.7) by using $|v^{(\varepsilon)}|_{L_p}\leqslant|v|_{L_p}$ for $v\in L_p(\mathbb{R}^d)$. □

Corollary V.4.2. *Let the conditions of Lemma V.4.1 hold for integers $m\geqslant 0$ and $p\geqslant 2$ even. Then for $\varepsilon>0$ we have*

$$((D^\alpha\mu^{(\varepsilon)})^{p-1},D^\alpha((a^{ij}D_{ij})^*\mu)^{(\varepsilon)})\leqslant NL^2|u|^p_{W^m_p}$$

for multi-indices $\alpha=(\alpha_1,...,\alpha_d)$ such that $0\leqslant|\alpha|\leqslant m$, where N is a constant *depending only on d, m and p.*

Proof. As in the proof of Corollary IV.4.2, it suffices to note that

$$((D^\alpha\mu^{(\varepsilon)})^{p-2}D^\alpha((\sigma^{ik}D_i)^*\mu)^{(\varepsilon)}, D^\alpha((\sigma^{jk}D_j)^*\mu)^{(\varepsilon)})$$

$$= \int_{\mathbb{R}^d}(D^\alpha\mu^{(\varepsilon)})^{p-2}(x)\sum_{k=1}^{d}\left|D^\alpha((\sigma^{ik}D_i)^*\mu)^{(\varepsilon)}(x)\right|^2 dx \geqslant 0.$$

□

Lemma V.4.3. *Let $p \geqslant 2$ and $m \geqslant 0$ be integers, and let $\sigma = (\sigma^i)$ and b be Borel functions on $\mathbb{R}^d$ with values in $\mathbb{R}^d$ and $\mathbb{R}$ respectively. Assume the partial derivatives of σ and $b\sigma$ up to order m are functions such that there exist constants $K \geqslant L \geqslant 1$ such that*

$$\sum_{k=0}^{m+1}|D^k b(x)| \leqslant K, \quad |\sigma(x)| \leqslant K_0,$$

$$\sum_{k=1}^{m+1}|D^k\sigma(x)| + \sum_{k=1}^{m+1}|D^k(b\sigma)(x)| \leqslant L$$

for all $x, y \in \mathbb{R}^d$. Then for finite signed Borel measures μ on $\mathbb{R}$ with density $u := d\mu/dx \in W_p^m$, satisfying (V.4.5), *we have*

$$\big((D^\alpha\mu^{(\varepsilon)})^{p-2}D^\alpha(b\mu)^{(\varepsilon)}, D^\alpha(b\mu)^{(\varepsilon)}\big) \leqslant NK^2|u|^p_{W_p^m}, \tag{V.4.17}$$

$$\big((D^\alpha\mu^{(\varepsilon)})^{p-2}, D^\alpha((\sigma^iD_i)^*\mu)^{(\varepsilon)}D^\alpha(b\mu)^{(\varepsilon)}\big) \leqslant NKL|u|^p_{W_p^m} \tag{V.4.18}$$

for $\varepsilon > 0$ and multi-indices α such that $|\alpha| \leqslant m$, where N is a constant depending only on d, p, m.

Proof. First note that by (V.4.8) and (V.4.9), as well as the conditions on σ and b, the left-hand sides of (V.4.17) and (V.4.18) are well-defined. Interchanging the order of integration and the differential operator D^α, rewriting the product of integrals as multiple integral, using Fubini's theorem and the identity

$$D_x^\alpha k_\varepsilon(x-z) = (-1)^{|\alpha|}D_z^\alpha k_\varepsilon(x-z), \quad x, y \in \mathbb{R}^d,$$

as well as (V.4.1), for the left-hand side F_α of (V.4.17) we compute

$$F^\alpha = \int_{\mathbb{R}^d}\int_{\mathbb{R}^{pd}} b(y_r)b(y_s)\Pi_{j=1}^p D_x^\alpha k_\varepsilon(x-y_j)\,\mu_p(dy)\,dx$$

$$= (-1)^{p|\alpha|}\int_{\mathbb{R}^d}\int_{\mathbb{R}^{pd}} b(y_r)b(y_s)\Pi_{j=1}^p D_{y_j}^\alpha k_\varepsilon(x-y_j)\,\mu_p(dy)\,dx$$

$$= (-1)^{p|\alpha|}\int_{\mathbb{R}^{pd}} b(y_r)b(y_s)D_y^{p\alpha}\int_{\mathbb{R}^d}\Pi_{j=1}^p k_\varepsilon(x-y_j)\,dx\,\mu_p(dy)$$

$$= (-1)^{p|\alpha|} \int_{\mathbb{R}^{pd}} b(y_r) b(y_s) D_y^{p\alpha} \rho_\varepsilon(y) \Pi_{j=1}^p u(y_j)\, dy$$

for any $r, s \in \{1, 2, ..., p\}$ such that $r \neq s$, where recall that $dy = dy_1...dy_p$ and $D_y^{p\alpha} = \Pi_{j=1}^p D_{y_j}^\alpha$. Hence by integration by parts we obtain

$$F^\alpha = \sum_{\beta \leqslant \alpha} \sum_{\gamma \leqslant \alpha} c_\beta^\alpha c_\gamma^\alpha \int_{\mathbb{R}^{pd}} b_\beta(y_r) b_\gamma(y_s) \rho_\varepsilon(y) u_{\bar{\beta}}(y_r) u_{\bar{\gamma}}(y_s) \Pi_{j \neq s,r}^p u_\alpha(y_j)\, dy,$$

where $v_\delta := D^\delta v$ and $\bar{\delta} := \alpha - \delta$ for functions v on $\mathbb{R}^d$ and multi-indices $\delta \leqslant \alpha$. Using here (V.4.1) and the boundedness condition on $|b|$ and $|D^\delta b|$ we have

$$F^\alpha \leqslant NK^2 \sum_{\beta \leqslant \alpha} \sum_{\gamma \leqslant \alpha} \int_{\mathbb{R}^{pd}} \int_{\mathbb{R}^d} \Pi_{j=1}^p k_\varepsilon(x - y_j) |u_{\bar{\beta}}(y_r)| |u_{\bar{\gamma}}(y_s)| \Pi_{j \neq s,r}^p |u_\alpha(y_j)|\, dx\, dy$$

$$= NK^2 \sum_{\beta \leqslant \alpha} \int_{\mathbb{R}^d} |u_{\bar{\beta}}|^{(\varepsilon)} |u_{\bar{\gamma}}|^{(\varepsilon)} ||u_\alpha|^{(\varepsilon)}|^{p-2}\, dx \leqslant N' K^2 |u|_{W_p^m}^p$$

with constants N and N' depending only on p, d and m, where the last inequality follows by Young's inequality and the boundedness of the mollification operator in L_p. Now we are going to prove (V.4.18). By the same way as we have rewritten F^α we can rewrite the left-hand side R^α of the inequality (V.4.18) as

$$R^\alpha = \int_{\mathbb{R}^{dp}} f_{krr}(y)\, dy, \tag{V.4.19}$$

for any $r, k \in \{1, 2, .., p\}$ such that $r \neq k$, where

$$f_{krs}(y) := (-1)^{p|\alpha|} b(y_k) \sigma^i(y_r) \partial_{y_s^i} D_y^{p\alpha} \rho_\varepsilon(y)\, \Pi_{j=1}^p u(y_j), \quad y = (y_1, ..., y_p) \in \mathbb{R}^{pd}$$

for $k, r, s \in \{1, 2, ..., p\}$. As in the proof of Lemma V.4.1, for real functions f and g we write $f \sim g$ if they have the same (finite) Lebesgue integral against $dy = dy_1...dy_p$ over $\mathbb{R}^{pd}$. We write $f \preceq g$ if the integrals of f and g against dy over $\mathbb{R}^d$ are finite, and the integral of $f - g$ can be estimated by $NKL|u|_{W_p^m}$ for all $u \in W_p^m$ with a constant $N = N(d, m, p)$, independent of u. By integration by parts we have

$$f_{krr} \sim \sum_{\gamma \leqslant \alpha} \sum_{\beta \leqslant \alpha} f_{krr}^{\gamma\beta}$$

with

$$f_{krs}^{\gamma\beta}(y) := c_\gamma^\alpha c_\beta^\alpha b_\gamma(y_k) \sigma_\beta^i(y_r) \partial_{y_s^i} \rho_\varepsilon(y)\, u_{\bar{\gamma}}(y_k) u_{\bar{\beta}}(y_r) \Pi_{j \neq k,r} u_\alpha(y_j)$$

If $\beta \neq 0$ then by integration by parts (dropping $\partial_{y_r^i}$ from ρ_ε to the other terms), and using the boundedness of b, its derivatives up to order $m + 1$, and the boundedness of the derivatives of σ up to order $m + 1$, we see that $f_{krr}^{\gamma\beta} \preceq 0$ for any $k = 1, 2, ..p$, $r \neq k$ and $\gamma \leqslant \alpha$. If $\beta = 0$ and $\gamma = 0$, then f_{krr}^{00} can be estimated by an exact repetition of the proof of Lemma 4.2 in [17], by replacing

μ therein with $u_\alpha dy$, to yield $f^{00}_{krr} \preceq 0$. Consequently,

$$f_{krr} \preceq \sum_{0\neq\gamma\leqslant\alpha} f^{\gamma 0}_{krr} \quad \text{for every } k = 1, ..., p \text{ and } r \in \{1, 2, ..., p\}\backslash\{k\}.$$

Writing $f^{\gamma 0}_{krr}(y) = g^{\gamma}_{krr}(y)h^{\bar{\gamma}}_{k}(y)$, with

$$g^{\gamma}_{krs}(y) := c^{\alpha}_{\gamma} b_\gamma(y_k)\sigma^i(y_r)\partial_{y^i_s}\rho_\varepsilon(y), \quad h^{\bar{\gamma}}_{k}(y) := u_{\bar{\gamma}}(y_k)\Pi_{j\neq k}u_\alpha(y_j),$$

we get

$$p(p-1)(p-2)R^\alpha \leqslant \sum_{0\neq\gamma\leqslant\alpha}\sum_{s=1}^{p}\sum_{r\neq s}\sum_{k\neq s,r}\int_{\mathbb{R}^{dp}} g^{\gamma}_{kss}(y)h^{\bar{\gamma}}_{k}(y)\,dy + NKL|u|^p_{W^m_p}, \tag{V.4.20}$$

and by (V.4.3),

$$\begin{aligned} p(p-1)R^\alpha &\leqslant -\sum_{0\neq\gamma\leqslant\alpha}\sum_{k=1}^{p}\sum_{r\neq k}\sum_{s\neq r}\int_{\mathbb{R}^{pd}} g^{\gamma}_{krs}(y)h^{\bar{\gamma}}_{k}(y)\,dy + NKL|u|^p_{W^m_p} \\ &= -\sum_{0\neq\gamma\leqslant\alpha}\sum_{s=1}^{p}\sum_{r\neq s}\sum_{k\neq s,r}\int_{\mathbb{R}^{pd}} g^{\gamma}_{krs}(y)h^{\bar{\gamma}}_{k}(y)\,dy \\ &\quad -\sum_{0\neq\gamma\leqslant\alpha}\sum_{s=1}^{p}\sum_{r\neq s}\int_{\mathbb{R}^{pd}} g^{\gamma}_{srs}(y)h^{\bar{\gamma}}_{s}(y)\,dy + NKL|u|^p_{W^m_p} \end{aligned} \tag{V.4.21}$$

with a constant $N = N(d, m, p)$. Summing up (V.4.20) and (V.4.21) we obtain

$$\begin{aligned} c_p R^\alpha \leqslant& \sum_{0\neq\gamma\leqslant\alpha}\sum_{s=1}^{p}\sum_{r\neq s}\sum_{k\neq r,s}\int_{\mathbb{R}^{pd}} (g^{\gamma}_{kss}(y) - g^{\gamma}_{krs}(y))h^{\bar{\gamma}}_{k}(y)\,dy + NKL|u|^p_{W^m_p} \\ &-\sum_{0\neq\gamma\leqslant\alpha}\sum_{s=1}^{p}\sum_{r\neq s}\int_{\mathbb{R}^{pd}} g^{\gamma}_{srs}(y)h^{\bar{\gamma}}_{s}(y)\,dy + NKL|u|^p_{W^m_p} \end{aligned} \tag{V.4.22}$$

where $c_p = p(p-1)^2$, and

$$(g^{\gamma}_{kss}(y) - g^{\gamma}_{krs}(y))h^{\bar{\gamma}}_{k}(y) = c^{\alpha}_{\gamma} b_\gamma(y_k)(\sigma^i(y_s) - \sigma^i(y_r))\partial_{y^i_s}\rho_\varepsilon(y)u_{\bar{\gamma}}(y_k)\Pi_{j\neq k}u_\alpha(y).$$

By the boundedness of $|b_\gamma|$ and the Lipschitz condition on σ, using (V.4.4) we get

$$(g^{\gamma}_{kss} - g^{\gamma}_{krs})h^{\bar{\gamma}}_{k} \preceq 0, \quad \text{for all } 0 \neq \gamma \leqslant \alpha,\ s = 1, 2, ..., p \text{ and } r \neq s,\ k \neq r, s.$$

By integration by parts we have for the last term in (V.4.22),

$$g^{\gamma}_{srs}h^{\bar{\gamma}}_{s} \preceq 0, \quad \text{for } 0 \neq \gamma \leqslant \alpha \text{ and } s = 1, \dots, p,\ r \neq s,$$

which finishes the proof of (V.4.18). $\square$

For vectors $\xi = \xi(x) \in \mathbb{R}^d$, depending on $x \in \mathbb{R}^d$ we consider the linear operators I^ξ and J^ξ defined by

$$T^\xi\varphi(x) = \varphi(x + \xi(x)) \tag{V.4.23}$$

$$I^\xi\varphi(x) := T^\xi\varphi(x) - \varphi(x), \quad J^\xi\psi(x) := I^\xi\psi(x) - \xi(x)D_i\psi(x),$$

$x \in \mathbb{R}^d$, acting on functions φ and differentiable functions ψ on $\mathbb{R}^d$.

Lemma V.4.4. *Let $\xi = \xi(x, \mathfrak{z})$ be an $\mathbb{R}^d$-valued function of $x \in \mathbb{R}^d$ for every $\mathfrak{z} \in \mathfrak{Z}$ for a set $\mathfrak{Z}$. Assume that for an integer $m \geqslant 1$ the partial derivatives of ξ in $x \in \mathbb{R}^d$ up to order m are functions on $\mathbb{R}^d$ for each $\mathfrak{z} \in \mathfrak{Z}$, such that for a constant $\lambda > 0$, a function $\bar{\xi}$ on $\mathfrak{Z}$ and a constant $K_\xi \geqslant 0$ we have*

$$|\xi(x, \mathfrak{z})| \leqslant \bar{\xi}(\mathfrak{z}) \leqslant K_\xi,$$

$$\sum_{k=1}^{m+1} |D_x^k\xi(x, \mathfrak{z})| \leqslant \bar{\xi}(\mathfrak{z}), \quad |\det(\mathbb{I} + \theta D_x\xi(x, \mathfrak{z}))| \geqslant \lambda^{-1} \tag{V.4.24}$$

for all $x, y \in \mathbb{R}^d$, $\mathfrak{z} \in \mathfrak{Z}$ and $\theta \in [0, 1]$. Let $p \geqslant 2$ be an even integer. Then for every finite signed Borel measure μ with density $u = d\mu/dx \in W_p^m$, satisfying (V.4.5), *we have*

$$C := \int_{\mathbb{R}^d} p(D_x^\alpha\mu^{(\varepsilon)})^{p-1} D_x^\alpha(J^{\xi*}\mu)^{(\varepsilon)}\, dx$$

$$+ \int_{\mathbb{R}^d} (D_x^\alpha\mu^{(\varepsilon)} + D_x^\alpha(I^{\xi*}\mu)^{(\varepsilon)})^p - (D_x^\alpha\mu^{(\varepsilon)})^p - p(D_x^\alpha\mu^{(\varepsilon)})^{p-1} D_x^\alpha(I^{\xi*}\mu)^{(\varepsilon)}\, dx$$

$$\leqslant N\bar{\xi}^2(\mathfrak{z})|u|_{W_p^m}^p \quad \textit{for } \mathfrak{z} \in \mathfrak{Z},\ \varepsilon > 0 \tag{V.4.25}$$

for multi-indices α, $0 \leqslant |\alpha| \leqslant m$ with a constant $N = N(d, p, m, \lambda, K_\xi)$.

Proof. Again we note that by (V.4.8) & (V.4.9), together with the conditions on ξ, it is easy to verify that C is well-defined. Notice that

$$D_x^\alpha\mu^{(\varepsilon)} + D_x^\alpha(I^{\xi*}\mu)^{(\varepsilon)} = D_x^\alpha(T^{\xi*}\mu)^{(\varepsilon)}$$

and

$$p(D_x^\alpha\mu^{(\varepsilon)})^{p-1} D_x^\alpha(J^{\xi*}\mu)^{(\varepsilon)} - p(D_x^\alpha\mu^{(\varepsilon)})^{p-1} D_x^\alpha(I^{\xi*}\mu)^{(\varepsilon)}$$
$$= -p(D_x^\alpha\mu^{(\varepsilon)})^{p-1} D_x^\alpha((\xi^i D_i)^*\mu)^{(\varepsilon)},$$

Hence

$$C = \int_{\mathbb{R}^d} (D_x^\alpha(T^{\xi*}\mu)^{(\varepsilon)})^p - (D_x^\alpha\mu^{(\varepsilon)})^p - p(D_x^\alpha\mu^{(\varepsilon)})^{p-1} D_x^\alpha((\xi^i D_i)^*\mu)^{(\varepsilon)}\, dx. \tag{V.4.26}$$

First we change the order of D_x^α and the integrals and operators T_y^ξ and I_y^ξ acting in the variable $y \in \mathbb{R}^d$, then we use

$$D_x^\alpha k_\varepsilon(x-y) = (-1)^{|\alpha|} D_y^\alpha k_\varepsilon(x-y)$$

to get

$$D_x^\alpha (T^{\xi *}\mu)^{(\varepsilon)} = (-1)^{|\alpha|} \int_{\mathbb{R}^d} T_y^\xi D_y^\alpha k_\varepsilon(x-y)\,\mu(dy),$$

$$D_x^\alpha \mu^{(\varepsilon)} = (-1)^{|\alpha|} \int_{\mathbb{R}^d} D_y^\alpha k_\varepsilon(x-y)\,\mu(dy),$$

$$D_x^\alpha ((\xi^i D_i)^*\mu)^{(\varepsilon)} = (-1)^{|\alpha|} \int_{\mathbb{R}^d} \xi^i(y)\partial_{y^i} D_y^\alpha k_\varepsilon(x-y)\,\mu(dy).$$

Thus rewriting the product of integrals as multiple integrals, and using the product measure $\mu_p(dy) := \mu(dy_1)...\mu(dy_p)$ on $\mathbb{R}^{dp}$ by Fubini's theorem we get

$$\begin{aligned}(D_x^\alpha (T^{\xi *}\mu)^{(\varepsilon)})^p(x) &= \int_{\mathbb{R}^{pd}} \Pi_{r=1}^p (T_{y_r}^\xi D_{y_r}^\alpha k_\varepsilon(x-y_r))\,\mu_p(dy)\\ &= \int_{\mathbb{R}^{pd}} \Pi_{r=1}^p T_{y_r}^\xi D_y^{p\alpha} \Pi_{r=1}^p k_\varepsilon(x-y_r)\,\mu_p(dy),\\ (D_x^\alpha \mu^{(\varepsilon)})^p &= \int_{\mathbb{R}^{pd}} \Pi_{r=1}^p (D_{y_r}^\alpha k_\varepsilon(x-y_r))\,\mu_p(dy)\\ &= \int_{\mathbb{R}^{pd}} D_y^{p\alpha} \Pi_{r=1}^p k_\varepsilon(x-y_r)\,\mu_p(dy) \end{aligned} \tag{V.4.27}$$

and

$$\begin{aligned} p(D_x^\alpha \mu^{(\varepsilon)})^{p-1} D_x^\alpha ((\xi^i D_i)^*\mu)^{(\varepsilon)} &= p\int_{\mathbb{R}^{pd}} \Pi_{r=1}^{p-1}(D_{y_r}^\alpha k_\varepsilon(x-y_r))\xi^i(y_p)\partial_{y_p^i} D_{y_p}^\alpha k_\varepsilon(x-y_p)\mu_p(dy)\\ &= p\int_{\mathbb{R}^{pd}} \xi^i(y_p)\partial_{y_p^i} D_y^{p\alpha} \Pi_{r=1}^p k_\varepsilon(x-y_r)\mu_p(dy)\\ &= \int_{\mathbb{R}^{pd}} \sum_{r=1}^p \xi^i(y_r)\partial_{y_r^i} D_y^{p\alpha}\Pi_{r=1}^p k_\varepsilon(x-y_r)\,\mu_p(dy), \end{aligned} \tag{V.4.28}$$

where again

$$D_y^{p\alpha} := \Pi_{r=1}^p D_{y_r}^\alpha \quad \text{for } y = (y_1, ..., y_p) \in \mathbb{R}^{dp},$$

and the last equation is due to the symmetry of the function $\Pi_{r=1}^p D_{y_r}^\alpha k_\varepsilon(x-y_r)$ and the measure $\mu_p(dy)$ in $y = (y_1, ..., y_p) \in \mathbb{R}^{pd}$. Thus from (V.4.26) we get

$$C = \int_{\mathbb{R}^d}\int_{\mathbb{R}^{pd}} L_y^\xi D_y^{p\alpha} \Pi_{r=1}^p k_\varepsilon(x-y_r)\,\mu_p(dy)\,dx$$

with the operator

$$L^{\xi}_y = \Pi^p_{r=1} T^{\xi}_{y_r} - \mathbb{I} - \sum_{r=1}^{p} \xi^i(y_r)\partial_{y^i_r},$$

defined by

$$L^{\xi}_y\varphi(y) = \varphi(y_1+\xi(y_1), ..., y_p+\xi(y_p)) - \varphi(y) - \sum_{r=1}^{p} \xi^i(y_r)\partial_{y^i_r}\varphi(y), \quad y = (y_1, ..., y_p) \in \mathbb{R}^{pd}$$

for differentiable functions φ of $y = (y_1, ..., y_p) \in \mathbb{R}^{pd}$. Using here Fubini's theorem then changing the order of the operator $L^{\xi}_y D^{p\alpha}_y$ and the integration against dx, by virtue of (V.4.1) we have

$$C = \int_{\mathbb{R}^{pd}} L^{\xi}_y D^{p\alpha}_y \int_{\mathbb{R}^d} \Pi^p_{r=1} k_\varepsilon(x - y_r)\, dx\, \mu_p(dy) = \int_{\mathbb{R}^{pd}} L^{\xi}_y D^{p\alpha}_y \rho_\varepsilon(y)\, \mu_p(dy), \tag{V.4.29}$$

By Taylor's formula

$$L^{\xi}_y D^{p\alpha}_y \rho_\varepsilon(y) = \int_0^1 (1 - \vartheta)\xi^i(y_k)\xi^j(y_l)(\partial_{y^i_k}\partial_{y^j_l} D^{p\alpha}\rho_\varepsilon)(y + \vartheta\bar{\xi}(y))\, d\vartheta,$$

where $y = (y_1, ..., y_p) \in \mathbb{R}^{pd}$, $y_k \in \mathbb{R}^d$ for $k = 1, 2, ..., p$, and $\bar{\xi}(y) := (\xi(y_1), ..., \xi(y_p))$ for $y = (y_1, ..., y_p) \in \mathbb{R}^{dp}$. Thus by changing the order of integrals and then changing the variables y_k with $y_k + \vartheta\xi(y_k)$ for $k = 1, 2, ..., p$, from (V.4.29) we obtain

$$C = \int_0^1 (1 - \vartheta)C(\vartheta)\, d\vartheta \tag{V.4.30}$$

with

$$C(\vartheta) = \int_{\mathbb{R}^{pd}} \sum_{k=1}^{p}\sum_{l=1}^{p} \hat{\xi}^i(y_k)\hat{\xi}^j(y_l)\partial_{y^i_k}\partial_{y^j_l} D^{p\alpha}_y \rho_\varepsilon(y)\Pi^p_{r=1}\hat{u}(y_r)\, dy,$$

where, with $\tau_\vartheta(x) := x + \vartheta\xi(x)$,

$$\hat{\xi}^i(x) := \xi^i(\tau^{-1}_\vartheta(x)), \quad \hat{u}(x) = u(\tau^{-1}_\vartheta(x))|\det D\tau^{-1}_\vartheta(x)|, \quad x \in \mathbb{R}^d, \quad i = 1, 2, ..., d, \tag{V.4.31}$$

and $dy := dy_1 dy_2 ... dy_p$ denotes the Lebegue measure on $\mathbb{R}^{pd}$. Clearly,

$$C(\vartheta) = C_1(\vartheta) + C_2(\vartheta)$$

with

$$C_1(\vartheta) = \int_{\mathbb{R}^{pd}} \sum_{k=1}^{p} \hat{\xi}^i(y_k)\hat{\xi}^j(y_k)\partial_{y^i_k}\partial_{y^j_k} D^{p\alpha}_y \rho_\varepsilon(y)\Pi^p_{r=1}\hat{u}(y_r)\, dy,$$

$$C_2(\vartheta) = \int_{\mathbb{R}^{pd}} \sum_{k=1}^{p}\sum_{l\neq k} \hat{\xi}^i(y_k)\hat{\xi}^j(y_l)\partial_{y^i_k}\partial_{y^j_l} D^{p\alpha}_y \rho_\varepsilon(y)\Pi^p_{r=1}\hat{u}(y_r)\, dy.$$

Using (V.4.3) and the symmetry in y_k and y_l, we have

$$
\begin{aligned}
C_1(\vartheta) =& -\tfrac{1}{2}\int_{\mathbb{R}^{pd}}\sum_{k=1}^{p}\sum_{l\neq k}(\hat{\xi}^i(y_k)\hat{\xi}^j(y_k)+\hat{\xi}^i(y_l)\hat{\xi}^j(y_l))\partial_{y_k^i}\partial_{y_l^j}D_y^{p\alpha}\rho_\varepsilon(y)\Pi_{r=1}^{p}\hat{u}(y_r)\,dy,\\
C_2(\vartheta) =& \tfrac{1}{2}\int_{\mathbb{R}^{pd}}\sum_{k=1}^{p}\sum_{l\neq k}(\hat{\xi}^i(y_k)\hat{\xi}^j(y_l)+\hat{\xi}^i(y_l)\hat{\xi}^j(y_k))\partial_{y_k^i}\partial_{y_l^j}D_y^{p\alpha}\rho_\varepsilon(y)\Pi_{r=1}^{p}\hat{u}(y_r)\,dy.
\end{aligned}
\tag{V.4.32}
$$

Hence

$$
C(\vartheta) = \int_{\mathbb{R}^{pd}}\sum_{r=1}^{p}\sum_{s\neq r}\hat{a}^{ij}(y_r,y_s)\partial_{y_r^i}\partial_{y_s^j}D_y^{p\alpha}\rho_\varepsilon(y)\Pi_{r=1}^{p}\hat{u}(y_r)\,dy \tag{V.4.33}
$$

with

$$
\hat{a}^{ij}(y_r,y_s) = -\tfrac{1}{2}(\hat{\xi}^i(y_r)-\hat{\xi}^i(y_s))(\hat{\xi}^j(y_r)-\hat{\xi}(y_s^j))
$$

Notice that the right-hand side of equation (V.4.33) is the same as the right-hand side of (V.4.11) with $\hat{\xi}^i$ in place of $\sigma^{i\cdot}$ for each $i = 1, 2, ..., d$ and with $\hat{u}$ in place of u. It is easy to verify, see Lemma 3.3 in [42], that for a constant $N = N(d, \lambda, m, K_\xi)$ we have

$$
\sum_{k=1}^{m+1}|D_x^k(\tau_\vartheta^{-1}(x))| \leqslant N, \quad \text{for each } \vartheta \in [0,1], \mathfrak{z} \in \mathfrak{Z}, x \in \mathbb{R}^d.
$$

Thus also for each $\vartheta \in [0, 1]$,

$$
\sum_{k=1}^{m+1}|D^k\hat{\xi}(x,\mathfrak{z})| \leqslant N\bar{\xi}(\mathfrak{z}) \quad \text{for } x \in \mathbb{R}^d,\ \mathfrak{z} \in \mathfrak{Z}, \tag{V.4.34}
$$

with a constant $N = N(d, m, \lambda, K_\xi)$, i.e., for each $\vartheta \in [0, 1]$ and $\mathfrak{z} \in \mathfrak{Z}$ the function $\hat{\xi}$ of $x \in \mathbb{R}^d$ satisfies the condition (V.4.6) on σ in Lemma V.4.1, with $N\bar{\xi}(\mathfrak{z})$ in place of L. Consequently, copying the calculations which lead from equation (V.4.11) to the estimate (V.4.7) in the proof of Lemma V.4.1, we obtain

$$
C(\vartheta) \leqslant N\bar{\xi}^2(\mathfrak{z})|\hat{u}|_{W_p^m}^p \quad \text{for each } \vartheta \in [0,1],\ \mathfrak{z} \in \mathfrak{Z}
$$

with a constant $N = N(d, m, p, \lambda, K_\xi)$. Note that due to the condition (V.4.24) there is a constant $N = N(d, p, m, \lambda, K_\xi)$ such that

$$
|\hat{u}|_{W_p^m} \leqslant N|u|_{W_p^m} \quad \text{for all } \vartheta \in [0,1]. \tag{V.4.35}
$$

Hence by virtue of (V.4.30) the estimate (V.4.25) follows. □

Corollary V.4.5. *Let the conditions of Lemma V.4.4 hold. Then for every* *finite signed Borel measure μ with density $u = d\mu/dx \in W_p^m$, satisfying* (V.4.5), *we have*

$$\int_{\mathbb{R}^d} (D_x^\alpha \mu^{(\varepsilon)})^{p-1} D_x^\alpha (J^{\xi *}\mu)^{(\varepsilon)}\, dx \leqslant N\bar{\xi}^2(\mathfrak{z})|u|^p_{W_p^m} \quad \textit{for } \mathfrak{z} \in \mathfrak{Z},\ \varepsilon > 0 \qquad \text{(V.4.36)}$$

for multi-indices α, $0 \leqslant |\alpha| \leqslant m$ with a constant $N = N(d, p, m, \lambda, K_\xi)$.

Proof. By the convexity of the function $f(a) = a^p$ for even $p \geqslant 2$ we know that $(a + b)^p - a^p - pa^{p-1}b \geqslant 0$ for all $a, b \in \mathbb{R}^d$. Applying this with $a = D^\alpha u^{(\varepsilon)}$ and $b = D^\alpha(I^\xi u)^{(\varepsilon)}$ shows that (V.4.25) implies (V.4.36). □

Lemma V.4.6. *Let the conditions of Lemma V.4.4 hold. Then for every finite signed Borel measure μ with density $u = d\mu/dx \in W_p^m$, satisfying* (V.4.5), *we have*

$$\Big|\int_{\mathbb{R}^d} (D^\alpha u^{(\varepsilon)} + D^\alpha (I^{\xi *}\mu)^{(\varepsilon)})^p - (D^\alpha u^{(\varepsilon)})^p\, dx\Big| \leqslant N\bar{\xi}(\mathfrak{z})|u|^p_{W_p^m}, \qquad \text{(V.4.37)}$$

for a constant $N = N(d, p, m, \lambda, K_\xi)$ for $\mathfrak{z} \in \mathfrak{Z}$, where the argument $x \in \mathbb{R}^d$ is suppressed in the integrand.

Proof. Define

$$F := \int_{\mathbb{R}^d} (D^\alpha u^{(\varepsilon)} + D^\alpha (I^{\xi *}u)^{(\varepsilon)})^p - (D^\alpha u^{(\varepsilon)})^p\, dx$$

$$= \int_{\mathbb{R}^d} (D^\alpha (T^{\xi *}u)^{(\varepsilon)})^p - (D^\alpha u^{(\varepsilon)})^p\, dx,$$

where we use the operator T defined in (V.4.23). As in the proof of Lemma 4.5 in [17] we define the operator

$$M_y^\xi = \Pi_{i=1}^p T_{y_i}^\xi - \mathbb{I}$$

where $\mathbb{I}$ is the identity operator. Observe that using Fubini's theorem and the notation $D_y^{p\alpha} = \Pi_{r=1}^p D_{y_r}^\alpha$, $dy = dy_1 \cdots dy_p$, $y = (y_1, \ldots, y_p) \in \mathbb{R}^{pd}$,

$$F = \int_{\mathbb{R}^d}\int_{\mathbb{R}^{dp}} \Big(D_x^{p\alpha}\Pi_{i=1}^p T_{y_i}^\xi \Pi_{j=1}^p k_\varepsilon(x-y_j)\Pi_{k=1}^p u(y_k) - D_x^{p\alpha}\Pi_{j=1}^p k_\varepsilon(x-y_j)\Pi_{k=1}^p u(y_k)\Big)\, dy\, dx$$

$$= \int_{\mathbb{R}^d}\int_{\mathbb{R}^{dp}} \Big(M_y^\xi D_y^{p\alpha}\Pi_{j=1}^p k_\varepsilon(x-y_j)\Big)\Pi_{k=1}^p u(y_k)\, dy\, dx = \int_{\mathbb{R}^{dp}} \Big(M_y^\xi D_y^{p\alpha}\rho_\varepsilon(y)\Big)\Pi_{k=1}^p u(y_k)\, dy.$$

Next, note that by Taylor's formula with $\bar{\xi}(y) = (\xi(y_1), \ldots, \xi(y_p)) \in \mathbb{R}^{dp}$,

$$M_y^\xi D_y^{p\alpha}\rho_\varepsilon(y) = \sum_{k=1}^p \int_0^1 (\partial_{y_k^i} D_y^{p\alpha}\rho_\varepsilon)(y + \vartheta\bar{\xi}(y))\, d\vartheta\, \xi^i(y_k).$$

Thus, by a change of variables, Fubini's theorem and the functions defined in (V.4.31),

$$F = \sum_{k=1}^{p} \int_0^1 \int_{\mathbb{R}^{dp}} \partial_{y_k^i} D_y^{p\alpha} \rho_\varepsilon(y) \hat{\xi}^i(y_k) \Pi_{j=1}^p \hat{u}(y_j)\, dy\, d\vartheta,$$

which by integration by parts gives, with multi-indices $\beta \leqslant \alpha$, $\bar{\beta} := \alpha - \beta$ and constants c_β^α,

$$F = \sum_{\beta \leqslant \alpha} c_\beta^\alpha \sum_{k=1}^{p} \int_0^1 \int_{\mathbb{R}^{dp}} \partial_{y_k^i} \rho_\varepsilon(y) \hat{\xi}_\beta^i(y_k) \hat{u}_{\bar{\beta}}(y_k) \Pi_{j \neq k}^p \hat{u}_\alpha(y_j)\, dy\, d\vartheta$$

$$= \sum_{\beta \leqslant \alpha} c_\beta^\alpha \sum_{k=1}^{p} \int_0^1 f_k^\beta(\vartheta)\, d\vartheta,$$

where for $k = 1, \ldots, p$, $\vartheta \in [0, 1]$ and $\beta \leqslant \alpha$,

$$f_k^\beta(\vartheta) := \int_{\mathbb{R}^{dp}} \partial_{y_k^i} \rho_\varepsilon(y) \hat{\xi}_\beta^i(y_k) \hat{u}_{\bar{\beta}}(y_k) \Pi_{j \neq k}^p \hat{u}_\alpha(y_j)\, dy$$

and where $\hat{u}_\gamma(y_k) = D_{y_k}^\gamma \hat{u}(y_k)$ for $\gamma = \alpha, \bar{\beta}$. We consider two cases. In the first case, let $\bar{\beta} < \alpha$ and hence $|\beta| \geqslant 1$. Then by integration by parts, for all $k = 1, \ldots, p$ and a constant $N = N(d, p, m, \lambda)$,

$$f_k^\beta(\vartheta) = -\int_{\mathbb{R}^{dp}} \rho_\varepsilon(y) \big((\partial_{y_k^i} \hat{\xi}_\beta^i(y_k)) \hat{u}_{\bar{\beta}}(y_k) + \hat{\xi}_\beta^i(y_k)(\partial_{y_k^i} \hat{u}_{\bar{\beta}}(y_k)) \big) \Pi_{j \neq k}^p \hat{u}_\alpha(y_j)\, dy\, d\vartheta$$

$$\leqslant N \bar{\xi}(\mathfrak{z}) |u|_{W_p^m}^p,$$

where we used (V.4.34) and (V.4.35). In the second case $\bar{\beta} = \alpha$ so that $\beta = 0$ and we have

$$\sum_{k=1}^{p} f_k^0 = \sum_{k=1}^{p} \int_{\mathbb{R}^{pd}} \partial_{y_k^i} \rho_\varepsilon(y) \hat{\xi}^i(y_k) \Pi_{j=1}^p \hat{u}_\alpha(y_j)\, dy,$$

as well as by using (V.4.3) and the symmetry in s and k,

$$\sum_{k=1}^{p} f_k^0 = -\sum_{k=1}^{p} \sum_{s \neq k} \int_{\mathbb{R}^{pd}} \partial_{y_k^i} \rho_\varepsilon(y) \hat{\xi}^i(y_s) \Pi_{j=1}^p \hat{u}_\alpha(y_j)\, dy.$$

Therefore also, with a constant $N = N(d, p, m, \lambda, K_\xi)$,

$$\Big|(p-1) \sum_{k=1}^{p} f_k^0 + p \sum_{k=1}^{p} f_k^0 \Big|$$

$$= \Big| \sum_{k=1}^{p} \sum_{s \neq k} \int_{\mathbb{R}^{pd}} \partial_{y_k^i} \rho_\varepsilon(y) \big(\hat{\xi}^i(y_k) - \hat{\xi}^i(y_s) \big) \Pi_{j=1}^p \hat{u}_\alpha(y_j)\, dy \Big| \leqslant N \bar{\eta}(\mathfrak{z}) |u|_{W_p^m}^p,$$

where we used (V.4.34) together with (V.4.4), as well as (V.4.35). This proves the lemma. □

V.5 Solvability of the filtering equations in Sobolev spaces

The following two lemmas are essentially Lemma IV.5.3 in Chapter IV, where instead of $D^\alpha k_\varepsilon$ the kernel k_ε is considered. However, keeping this difference in mind, the arguments in the proofs of Lemma IV.5.3 can easily be adapted. Hence we only provide an outline and refer the reader to Chapter IV for full details.

Lemma V.5.1. *Let the Assumption V.2.1 hold. Let u be an L_p-solution of (V.3.2), $p \geqslant 2$, and assume moreover that* $\operatorname{ess\,sup}_{t\in[0,T]} |u_t|_{L_1} < \infty$. *If $K_1 \neq 0$ in Assumption V.2.1 (ii), then assume additionally*

$$\operatorname*{ess\,sup}_{t\in[0,T]} \int_{\mathbb{R}^d} |y|^2 |u_t(y)|\, dy < \infty, \quad \text{almost surely.} \tag{V.5.1}$$

Then for each $\varepsilon > 0$ and integer $m \geqslant 0$, for any multi-index $\alpha = (\alpha_1, \dots, \alpha_d)$, $|\alpha| \leqslant m$, for all $x \in \mathbb{R}^d$ almost surely

$$\begin{aligned} D^\alpha u_t^{(\varepsilon)}(x) &= D^\alpha u_0^{(\varepsilon)}(x) + \int_0^t D^\alpha(\tilde{\mathcal{L}}_s^* u_s)^{(\varepsilon)}(x)\, ds + \int_0^t D^\alpha(\mathcal{M}_s^{*k} u_s)^{(\varepsilon)}(x)\, dV_s^k \\ &+ \int_0^t \int_{\mathfrak{Z}_0} D^\alpha(J_s^{\eta *} u_s)^{(\varepsilon)}(x)\, \nu_0(d\mathfrak{z}) ds + \int_0^t \int_{\mathfrak{Z}_1} D^\alpha(J_s^{\xi *} u_s)^{(\varepsilon)}(x)\, \nu_1(d\mathfrak{z}) ds \\ &+ \int_0^t \int_{\mathfrak{Z}_1} D^\alpha(I_s^{\xi *} u_{s-})^{(\varepsilon)}(x)\, \tilde{N}_1(ds, d\mathfrak{z}), \end{aligned} \tag{V.5.2}$$

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

Proof. The case of $\alpha = 0$ is Lemma IV.5.3. The case of $\alpha \neq 0$ such that $0 < |\alpha| \leqslant m$ works exactly in the same way. We first define for a $\psi \in C_0^\infty(\mathbb{R})$ such that $\psi(0) = 1$, $\psi(r) = 0$ for $|r| \geqslant 2$, for $n \geqslant 1$, $\psi_n(x) := \psi(|x|/n) \in C_0^\infty(\mathbb{R}^d)$. Setting $\varphi_x(y) := k_{\varepsilon,\alpha}(x - y)\psi_n(y) \in C_0^\infty$ in (V.3.2), where $k_{\varepsilon,\alpha}(x - y) = D_x^\alpha k_\varepsilon(x - y)$, yields that for each $x \in \mathbb{R}^d$ almost surely

$$\begin{aligned} (u_t, k_{\varepsilon,\alpha}(x - \cdot)\psi_n) &= (u_0, k_{\varepsilon,\alpha}(x - \cdot)\psi_n) + \int_0^t \big(u_s, \tilde{\mathcal{L}}_s(k_{\varepsilon,\alpha}(x - \cdot)\psi_n)\big)\, ds \\ &+ \int_0^t \big(u_s, \mathcal{M}_s^k(k_{\varepsilon,\alpha}(x-\cdot)\psi_n)\big)\, dV_s^k + \int_0^t \int_{\mathfrak{Z}_0} \big(u_s, J_s^\eta(k_{\varepsilon,\alpha}(x-\cdot)\psi_n)\big)\, \nu_0(d\mathfrak{z}) ds \end{aligned} \tag{V.5.3}$$

$$+\int_0^t\int_{\mathfrak{Z}_1}\big(u_s, J^\xi_s(k_{\varepsilon,\alpha}(x-\cdot)\psi_n)\big)\,\nu_1(d\mathfrak{z})ds+\int_0^t\int_{\mathfrak{Z}_1}\big(u_{s-}, I^\xi_s(k_{\varepsilon,\alpha}(x-\cdot)\psi_n)\big)\,\tilde N_1(d\mathfrak{z},ds)$$

for all $t\in[0,T]$. Then we notice that

$$|k_{\varepsilon,\alpha}(x-y)|\leqslant\sum_{|\gamma|\leqslant m+2}|D^\gamma k_\varepsilon(x-y)|\leqslant Nk_{2\varepsilon}(x-y), \tag{V.5.4}$$

as well as that by Assumption for all $x,y\in\mathbb{R}^d$, $s\in[0,T]$, $\mathfrak{z}_i\in\mathfrak{Z}_i$. $i=0,1$ and $n\geqslant 0$ we have

$$\sup_{x\in\mathbb{R}^d}|D^k\psi_n|=n^{-k}\sup_{\mathbb{R}^d}|D^k\psi|<\infty,\quad \text{for } k\in\mathbb{N},.$$

$$|\tilde{\mathcal{L}}_s(k_{\varepsilon,\alpha}(x-y)\psi_n(y))|+\sum_k|\mathcal{M}^k_s(k_{\varepsilon,\alpha}(x-y)\psi_n(y))|^2\leqslant N(K_0^2+K_1^2|y|^2+K_1^2|Y_s|^2),$$

$$\begin{aligned}|J^\eta_s(k_{\varepsilon,\alpha}(x-y)\psi_n(y))|&\leqslant\sup_{v\in\mathbb{R}^d}|D^2_v(k_{\varepsilon,\alpha}(x-v)\psi_n(v))||\eta_s(y,\mathfrak{z}_0)|^2\leqslant N|\eta_s(y,\mathfrak{z}_0)|^2\\&\leqslant N\bar\eta^2(\mathfrak{z}_0)(K_0^2+K_1^2|y|^2+K_1^2|Y_s|^2),\end{aligned}$$

and

$$\begin{aligned}&|J^\xi_s(k_{\varepsilon,\alpha}(x-y)\psi_n(y))|+|I^\xi_s(k_{\varepsilon,\alpha}(x-y)\psi_n(y))|^2\\&\leqslant\sup_{v\in\mathbb{R}^d}|D^2_v(k_{\varepsilon,\alpha}(x-v)\psi_n(v))||\xi_s(y,\mathfrak{z}_1)|^2+\sup_{v\in\mathbb{R}^d}|D_v(k_{\varepsilon,\alpha}(x-v)\psi_n(v))|^2|\xi_s(y,\mathfrak{z}_1)|^2\\&\leqslant N|\xi_s(y,\mathfrak{z}_1)|^2\leqslant N\bar\xi^2(\mathfrak{z}_1)(K_0^2+K_1^2|y|^2+K_1^2|Y_s|^2),\end{aligned}$$

for a constant $N=N(\varepsilon,m,d,K_0,K_1,K,K_\xi,K_\eta)$. Using

$$\operatorname*{ess\,sup}_{t\in[0,T]}\int_{\mathbb{R}^d}(1+|y|^2+|Y_t|^2)|u_t(y)|\,dy<\infty,\quad\text{(a.s.)}$$

together with the estimates above, we can apply Lebesgue's theorem on Dominated Convergence to get that for all $x\in\mathbb{R}^d$,

$$(u_t,k_{\varepsilon,\alpha}(x-\cdot)\psi_n)\to(u_t,k_{\varepsilon,\alpha}(x-\cdot)),\quad(u_0,k_{\varepsilon,\alpha}(x-\cdot)\psi_n)\to(u_0,k_{\varepsilon,\alpha}(x-\cdot))\quad\text{and}$$

$$\int_0^t\big(u_s,\mathcal{A}_s(k_{\varepsilon,\alpha}(x-\cdot)\psi_n)\big)\,ds\to\int_0^t(u_s,\mathcal{A}_sk_{\varepsilon,\alpha}(x-\cdot))\,ds$$

as $n\to\infty$, almost surely uniformly in time, as well as that

$$\lim_{n\to\infty}\int_0^t\big(u_s,\mathcal{M}^k_s(k_{\varepsilon,\alpha}(x-\cdot)\psi_n(\cdot))\big)\,dV^k_s=\int_0^t\big(u_s,\mathcal{M}^k_sk_{\varepsilon,\alpha}(x-\cdot)\big)\,dV^k_s,$$

$$\lim_{n\to\infty}\int_0^t\int_{\mathfrak{Z}_1}(u_{s-},I^\xi_s(k_{\varepsilon,\alpha}(x-\cdot)\psi_n))\,\tilde N_1(d\mathfrak{z},ds)=\int_0^t\int_{\mathfrak{Z}_1}(u_{s-},I^\xi_sk_{\varepsilon,\alpha}(x-\cdot))\,\tilde N_1(d\mathfrak{z},ds)$$

in probability, uniformly in time. Thus, letting $n\to\infty$ in (V.5.3) it remains to note that since $\mathcal{A}$ acts in the y variable,

$$
\begin{aligned}
(u_s, \mathcal{A}_s k_{\varepsilon,\alpha}(x-\cdot)) &= \int_{\mathbb{R}^d} u_s(y)\mathcal{A}_s D_x^\alpha k_\varepsilon(x-y)\,dy \\
&= D_x^\alpha \int_{\mathbb{R}^d} u_s(y)\mathcal{A}_s k_\varepsilon(x-y)\,dy = D^\alpha(\mathcal{A}_s^* u_s)^{(\varepsilon)}(x)
\end{aligned}
$$

for all $(\omega, s, x) \in \Omega \times [0,T] \times \mathbb{R}^d$ if $\mathcal{A} = \tilde{\mathcal{L}}, \mathcal{M}^k$ or the identity, as well as for all $(\omega, s, x, \mathfrak{z}_i) \in \Omega \times [0,T] \times \mathbb{R}^d \times \mathfrak{Z}_i$ if $\mathcal{A} = J^\eta$ or $\mathcal{A} = I^\xi, J^\xi$ with $i = 0, 1$ respectively.

□

Lemma V.5.2. *Let the Assumptions V.2.1 and V.2.2 hold. Let u be an L_p-solution of* (V.3.2), *$p \geqslant 2$ and assume moreover that* $\operatorname{ess\,sup}_{t\in[0,T]} |u_t|_{L_1} < \infty$. *If $K_1 \neq 0$ in Assumption V.2.1 (ii), then assume additionally* (V.5.1). *Then for each $\varepsilon > 0$ and integer $m \geqslant 0$, for any multi-index $\alpha = (\alpha_1, \ldots, \alpha_d)$, $|\alpha| \leqslant m$, almost surely*

$$
\begin{aligned}
|D^\alpha u_t^{(\varepsilon)}|_{L_p}^p &= |D^\alpha u_0^{(\varepsilon)}|_{L_p}^p + p\int_0^t \left(|D^\alpha u_s^{(\varepsilon)}|^{p-2} D^\alpha u_s^{(\varepsilon)}, D^\alpha(\tilde{\mathcal{L}}_s^* u_s)^{(\varepsilon)}\right) ds \\
&\quad + p\int_0^t \left(|D^\alpha u_s^{(\varepsilon)}|^{p-2} D^\alpha u_s^{(\varepsilon)}, D^\alpha(\mathcal{M}_s^{k*} u_s)^{(\varepsilon)}\right) dV_s^k \\
&\quad + \tfrac{p(p-1)}{2} \sum_k \int_0^t \left(|D^\alpha u_s^{(\varepsilon)}|^{p-2}, |D^\alpha(\mathcal{M}_s^{k*} u_s)^{(\varepsilon)}|^2\right) ds \\
&\quad + p\int_0^t \int_{\mathfrak{Z}_0} \left(|D^\alpha u_s^{(\varepsilon)}|^{p-2} D^\alpha u_s^{(\varepsilon)}, D^\alpha(J_s^{\eta*} u_s)^{(\varepsilon)}\right) \nu_0(d\mathfrak{z})ds \qquad \text{(V.5.5)} \\
&\quad + p\int_0^t \int_{\mathfrak{Z}_1} \left(|D^\alpha u_s^{(\varepsilon)}|^{p-2} D^\alpha u_s^{(\varepsilon)}, D^\alpha(J_s^{\xi*} u_s)^{(\varepsilon)}\right) \nu_1(d\mathfrak{z})ds \\
&\quad + p\int_0^t \int_{\mathfrak{Z}_1} \left(|D^\alpha u_{s-}^{(\varepsilon)}|^{p-2} D^\alpha u_{s-}^{(\varepsilon)}, D^\alpha(I_s^{\xi*} u_{s-})^{(\varepsilon)}\right) \tilde{N}_1(d\mathfrak{z}, ds) \\
&\quad + \int_0^t \int_{\mathfrak{Z}_1} \int_{\mathbb{R}^d} \Big\{ |D^\alpha u_{s-}^{(\varepsilon)} + D^\alpha(I_s^{\xi*} u_{s-})^{(\varepsilon)}|^p - |D^\alpha u_{s-}^{(\varepsilon)}|^p \\
&\qquad - p|D^\alpha u_{s-}^{(\varepsilon)}|^{p-2} D^\alpha u_{s-}^{(\varepsilon)} D^\alpha(I_s^{\xi*} u_{s-})^{(\varepsilon)} \Big\}\, dx N_1(d\mathfrak{z}, ds)
\end{aligned}
$$

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

Proof. We apply the Itô formula from Chapter IV, Theorem IV.5.1, to $|D^\alpha u_t^{(\varepsilon)}|_{L_p}^p$. In order to do that, we need to verify that almost surely for each $x \in \mathbb{R}^d$ and α, such that $0 \leqslant |\alpha| \leqslant m$,

$$\int_0^T |D^\alpha(\tilde{\mathcal{L}}_s^* u_s)^{(\varepsilon)}(x)|\, ds < \infty, \quad \int_0^T \sum_k |D^\alpha(\mathcal{M}_s^{k*} u_s)^{(\varepsilon)}(x)|^2 ds < \infty,$$

$$\int_0^T \int_{\mathfrak{Z}_0} |D^\alpha(J_s^{\eta*} u_s)^{(\varepsilon)}(x)|\, \nu_0(d\mathfrak{z})\, ds < \infty, \quad \int_0^T \int_{\mathfrak{Z}_1} |D^\alpha(J_s^{\xi*} u_s)^{(\varepsilon)}(x)|\, \nu_1(d\mathfrak{z})\, ds < \infty,$$

$$\int_0^T \int_{\mathfrak{Z}_1} |D^\alpha(I_s^{\xi*} u_s)^{(\varepsilon)}(x)|^2\, \nu_1(d\mathfrak{z}) ds < \infty,$$

that for every finite set $\Gamma \in \mathcal{B}(\mathbb{R}^d)$, almost surely

$$\int_\Gamma \int_0^T |D^\alpha(\tilde{\mathcal{L}}_s^* u_s)^{(\varepsilon)}(x)|\, dxds < \infty, \quad \int_\Gamma \Big(\int_0^T \sum_k |D^\alpha(\mathcal{M}_s^{k*} u_s)^{(\varepsilon)}(x)|^2 ds \Big)^{1/2} dx < \infty,$$

$$\int_\Gamma \int_0^T \int_{\mathfrak{Z}_0} |D^\alpha(J_s^{\eta*} u_s)^{(\varepsilon)}(x)|\, \nu_0(d\mathfrak{z})\, dxds < \infty, \quad \int_\Gamma \int_0^T \int_{\mathfrak{Z}_1} |D^\alpha(J_s^{\xi*} u_s)^{(\varepsilon)}(x)|\, \nu_1(d\mathfrak{z})\, dxds < \infty,$$

$$\int_\Gamma \Big(\int_0^T \int_{\mathfrak{Z}_1} |D^\alpha(I_s^{\xi*} u_s)^{(\varepsilon)}(x)|^2\, \nu_1(d\mathfrak{z}) ds \Big)^{1/2} dx < \infty,$$

as well as that almost surely

$$A := \int_0^T \int_{\mathbb{R}^d} |D^\alpha(\tilde{\mathcal{L}}_s^* u_s)^{(\varepsilon)}(x)|^p\, dxds < \infty,$$

$$A_\eta := \int_0^T \int_{\mathbb{R}^d} \Big| \int_{\mathfrak{Z}_0} D^\alpha(J_s^{\eta*} u_s)^{(\varepsilon)}(x) \nu_0(d\mathfrak{z}) \Big|^p dxds < \infty,$$

$$A_\xi := \int_0^T \int_{\mathbb{R}^d} \Big| \int_{\mathfrak{Z}_1} D^\alpha(J_s^{\xi*} u_s)^{(\varepsilon)}(x) \nu_1(d\mathfrak{z}) \Big|^p dxds < \infty,$$

$$B := \int_0^T \int_{\mathbb{R}^d} \big(\sum_k |D^\alpha(\mathcal{M}_s^{k*} u_s)^{(\varepsilon)}(x)|^2 \big)^{p/2} dxds < \infty,$$

$$G := \int_0^T \int_{\mathbb{R}^d} \int_{\mathfrak{Z}_1} |D^\alpha(I_s^{\xi*} u_s)^{(\varepsilon)}(x, \mathfrak{z})|^p\, \nu_1(d\mathfrak{z}) dxds < \infty,$$

$$H := \int_0^T \int_{\mathbb{R}^d} \Big(\int_{\mathfrak{Z}_1} |D^\alpha(I_s^{\xi*} u_s)^{(\varepsilon)}(x, \mathfrak{z})|^2\, \nu_1(d\mathfrak{z}) \Big)^{p/2} dxds < \infty.$$

For $\alpha = 0$ the claim is Lemma IV.5.2 and the estimates can be found in the proof thereof. To prove the case where $0 < |\alpha| \leqslant m$, we note that for $\mathcal{A} = \tilde{\mathcal{L}}, \mathcal{M}, I^\xi, J^\xi, J^\eta$ we have

$$D^\alpha(\mathcal{A}^* u)^{(\varepsilon)} = \int_{\mathbb{R}^d} D_x^\alpha(\mathcal{A}_y k_\varepsilon(x-y)) u(y)\, dy = \int_{\mathbb{R}^d} (\mathcal{A}_y k_{\varepsilon,\alpha}(x-y)) u(y)\, dy,$$

for $k_{\varepsilon,\alpha}(x) = D^\alpha k_\varepsilon(x)$. Hence, a word for word repetition of the proof of Lemma 5.2 & 5.4 in [17], where we replace k_ε by $k_{\varepsilon,\alpha}$ and recall (V.5.4), yields the desired result. ☐

Lemma V.5.3. *Let Assumptions V.2.1, V.2.2, V.2.4 andV.2.5 hold with an integer $m \geqslant 0$ and let $p \geqslant 2$ be even. Let u be an W^m_p-solution to* (V.3.2)*, such that $\mathbb{E}|u_0|^p_{W^m_p} < \infty$ and almost surely $\operatorname{ess\,sup}_{t\in[0,T]} |u_t|_{L_1} < \infty$. Then*

$$\mathbb{E} \sup_{t\in[0,T]} |u_t|^p_{W^m_p} \leqslant N\mathbb{E}|u_0|^p_{W^m_p} \tag{V.5.6}$$

for a constant $N = N(m, d, p, K, K_\eta, K_\xi, L, T, \lambda, |\bar{\xi}|_{L_2(\mathfrak{Z}_1)}, |\bar{\eta}_{L_2(\mathfrak{Z}_1)})$.

Proof. For $m = 0$ the claim is Lemma IV.5.5. We proceed similarly here. For the present case, fix a multi-index α such that $0 \neq |\alpha| \leqslant m$, and define

$$\mathcal{Q}_p(\alpha, b, \sigma, \rho, \beta, u, k_\varepsilon) = p\big((D^\alpha u^{(\varepsilon)})^{p-1}, D^\alpha(\tilde{\mathcal{L}}^* u)^{(\varepsilon)}\big) \tag{V.5.7}$$
$$+ \tfrac{p(p-1)}{2} \sum_k \big((D^\alpha u^{(\varepsilon)})^{p-2}, (D^\alpha(\mathcal{M}^{k*}u)^{(\varepsilon)})^2\big), \tag{V.5.8}$$

$$\mathcal{Q}^{(0)}_p(\alpha, \eta(\mathfrak{z}_0), u, k_\varepsilon) = p\big((D^\alpha u^{(\varepsilon)})^{p-1}, D^\alpha(J^{\eta(\mathfrak{z}_0)*}u)^{(\varepsilon)}\big),$$
$$\mathcal{Q}^{(1)}_p(\alpha, \xi(\mathfrak{z}_1), u, k_\varepsilon) = p((D^\alpha u^{(\varepsilon)})^{p-1}, D^\alpha(J^{\xi(\mathfrak{z}_1)*}u)^{(\varepsilon)}), \tag{V.5.9}$$

$$\mathcal{R}_p(\alpha, \xi(\mathfrak{z}_1), u, k_\varepsilon) = |D^\alpha u^{(\varepsilon)} + D^\alpha(I^{\xi(\mathfrak{z}_1)*}u)^{(\varepsilon)}|^p_{L_p}$$
$$- |D^\alpha u^{(\varepsilon)}|^p_{L_p} - p((D^\alpha u^{(\varepsilon)})^{p-1}, D^\alpha(I^{\xi(\mathfrak{z}_1)*}u)^{(\varepsilon)}),$$

for $u \in W^m_p$, $\beta \in \mathbb{R}^{d'}$, functions b, σ and ρ on $\mathbb{R}^d$, with values in $\mathbb{R}^d$, $\mathbb{R}^{d\times d_1}$ and $\mathbb{R}^{d\times d'}$, respectively, and $\mathbb{R}^d$-valued functions $\eta(\mathfrak{z}_0)$ and $\xi(\mathfrak{z}_1)$ for each $\mathfrak{z}_i \in \mathfrak{Z}_i$, $i = 0, 1$, where $\beta_t = B_t(X_t)$,

$$\tilde{\mathcal{L}} = \tfrac{1}{2}(\sigma^{il}\sigma^{jl} + \rho^{ik}\rho^{jk})D_{ij} + \beta^l\rho^{il}D_i + \beta^l B^l, \quad \mathcal{M}^k = \rho^{ik}D_i + B^k, \quad k = 1, 2, ..., d'.$$

By Lemma V.5.2 almost surely

$$d|D^\alpha u^{(\varepsilon)}_t|^p_{L_p} = \mathcal{Q}_p(\alpha, b_t, \sigma_t, \rho_t, \beta_t, u_t, k_\varepsilon)\, dt + \int_{\mathfrak{Z}_0} \mathcal{Q}^{(0)}_p(\alpha, \eta_t(\mathfrak{z}), u_t, k_\varepsilon)\, \nu_0(d\mathfrak{z})\, dt$$
$$+ \int_{\mathfrak{Z}_1} \mathcal{Q}^{(1)}_p(\alpha, \xi_t(\mathfrak{z}), u_t, k_\varepsilon)\, \nu_1(d\mathfrak{z})\, dt + \int_{\mathfrak{Z}_1} \mathcal{R}_p(\alpha, \xi_t(\mathfrak{z}), u_{t-}, k_\varepsilon)\, N_1(d\mathfrak{z}, dt) \tag{V.5.10}$$
$$+ d\zeta_1(\alpha, t) + d\zeta_2(\alpha, t),$$

for all $t \in [0, T]$ and

$$\zeta_1(\alpha, t) = p\int_0^t \big((D^\alpha u^{(\varepsilon)}_s)^{p-1}, D^\alpha(\mathcal{M}^{k*}_s u_s)^{(\varepsilon)}\big)\, dV^k_s, \tag{V.5.11}$$

$$\zeta_2(\alpha,t)=p\int_0^t\int_{\mathfrak{Z}_1}\big((D^\alpha u_s^{(\varepsilon)})^{p-1},D^\alpha(I_s^{\xi *}u_s)^{(\varepsilon)}\big)\,\tilde N_1(d\mathfrak{z},ds)\quad t\in[0,T]$$

are local martingales under P. We write

$$\int_{\mathfrak{Z}_1}\mathcal{R}_p(\alpha,\xi_t(\mathfrak{z}_1),u_{t-},k_\varepsilon)\,N_1(d\mathfrak{z},dt)=\int_{\mathfrak{Z}_1}\mathcal{R}_p(\alpha,\xi_t(\mathfrak{z}_1),u_{t-},k_\varepsilon)\,\nu_1(d\mathfrak{z})dt+d\zeta_3(\alpha,t)\tag{V.5.12}$$

with

$$\zeta_3(\alpha,t)=\int_0^t\int_{\mathfrak{Z}_1}\mathcal{R}_p(\alpha,\xi_s(\mathfrak{z}),u_{s-},k_\varepsilon)\,N_1(d\mathfrak{z},ds)-\int_0^t\int_{\mathfrak{Z}_1}\mathcal{R}_p(\alpha,\xi_s(\mathfrak{z}),u_{s-},k_\varepsilon)\,\nu_1(d\mathfrak{z})ds,$$

which we can justify if we show

$$A:=\int_0^T\int_{\mathfrak{Z}_1}|\mathcal{R}_p(\alpha,\xi_s(\mathfrak{z}),u_s,k_\varepsilon)|\,\nu_1(d\mathfrak{z})\,ds<\infty\,\text{(a.s.)}.\tag{V.5.13}$$

To this end observe that by Taylor's formula

$$0\leqslant\mathcal{R}_p(\alpha,\xi_t(\mathfrak{z}),u_t,k_\varepsilon))\leqslant N\int_{\mathbb{R}^d}(D^\alpha u_t^{(\varepsilon)})^{p-2}(D^\alpha(I^{\xi(\mathfrak{z})*}u_t)^{(\varepsilon)})^2+(D^\alpha(I^{\xi(\mathfrak{z})*}u_t)^{(\varepsilon)})^p\,dx\tag{V.5.14}$$

with a constant $N=N(d,p)$. Hence

$$\int_{\mathfrak{Z}_1}\mathcal{R}_p(\alpha,\xi_t(\mathfrak{z}),u_t,k_\varepsilon))\,\nu_1(d\mathfrak{z})$$

$$\leqslant N\int_{\mathbb{R}^d}(D^\alpha u_t^{(\varepsilon)})^{p-2}|D^\alpha(I_t^{\xi(\mathfrak{z})*}u_t)^{(\varepsilon)}|^2_{L_2(\mathfrak{Z}_1)}+|D^\alpha(I_t^{\xi(\mathfrak{z})*}u_t)^{(\varepsilon)}|^p_{L_p(\mathfrak{Z}_1)}\,dx$$

$$\leqslant N'\big(|D^\alpha u_t^{(\varepsilon)}|^p_{L_p}+A_1(t)+A_2(t)\big)$$

with

$$A_1(t)=\int_{\mathbb{R}^d}|D^\alpha(I_t^{\xi(\mathfrak{z})*}u_t)^{(\varepsilon)}|^p_{L_2(\mathfrak{Z}_1)}\,dx,\quad A_2(t)=\int_{\mathbb{R}^d}|D^\alpha(I_t^{\xi(\mathfrak{z})*}u_t)^{(\varepsilon)}|^p_{L_p(\mathfrak{Z}_1)}\,dx\tag{V.5.15}$$

and constants N and N' depending only on d and p. By Minkowski's inequality and using again that $D_x^\alpha I^\xi k_\varepsilon(x-y)=I^\xi D_x^\alpha k_\varepsilon(x-y)$,

$$\begin{aligned}|D^\alpha u_t^{(\varepsilon)}|^p_{L_p}&=\int_{\mathbb{R}^d}\Big|\int_{\mathbb{R}^d}D_x^\alpha k_\varepsilon(x-y)\,u_t(y)\,dy\Big|^p\,dx,\\&\leqslant\Big|\int_{\mathbb{R}^d}|D^\alpha k_\varepsilon|_{L_p}\,|u_t(y)|\,dy\Big|^p\leqslant|D^\alpha k_\varepsilon|^p_{L_p}|u_t|^p_{L_1},\end{aligned}\tag{V.5.16}$$

$$A_1(t)=\int_{\mathbb{R}^d}\Big|\int_{\mathfrak{Z}_1}\Big|\int_{\mathbb{R}^d}\big((D^\alpha k_\varepsilon)(x-y-\xi_t(y,\mathfrak{z}))-(D^\alpha k_\varepsilon)(x-y)\big)\,u_t(y)\,dy\Big|^2\nu_1(d\mathfrak{z})\Big|^{p/2}dx$$

$$\leqslant \Big|\int_{\mathfrak{z}_1}\Big|\int_{\mathbb{R}^d}|(D^\alpha k_\varepsilon)(\cdot - y - \xi_t(y,\mathfrak{z})) - (D^\alpha k_\varepsilon)(\cdot - y)||u_t(y)|\,dy\Big|^2_{L_p}\nu_1(d\mathfrak{z})\Big|^{p/2}$$

$$\leqslant \Big|\int_{\mathfrak{z}_1}\Big|\int_{\mathbb{R}^d}|D^{\alpha+1}k_\varepsilon|_{L_p}\bar{\xi}(\mathfrak{z}_1)(K_0 + K_1|y| + K_1|Y_t|)|u_t(y)|\,dy\Big|^2\nu_1(d\mathfrak{z})\Big|^{p/2}$$

$$\leqslant |D^{\alpha+1}k_\varepsilon|^p_{L_p}|\bar{\xi}|^p_{L_2(\mathfrak{z}_1)}\Big(\int_{\mathbb{R}^d}(K_0 + K_1|y| + K_1|Y_t|)|u_t(y)|\,dy\Big)^p, \qquad \text{(V.5.17)}$$

where $D^{\alpha+1} = D\,D^\alpha$ and similarly, using Assumption V.2.2,

$$A_2(t) = \int_{\mathbb{R}^d}\int_{\mathfrak{z}_1}\Big|\int_{\mathbb{R}^d}\big((D^\alpha k_\varepsilon)(x-y-\xi_t(y,\mathfrak{z})) - (D^\alpha k_\varepsilon)(x-y)\big)\,u_t(y)\,dy\Big|^p\nu_1(d\mathfrak{z})dx$$

$$\leqslant \int_{\mathfrak{z}_1}\Big|\int_{\mathbb{R}^d}\big|(D^\alpha k_\varepsilon)(\cdot - y - \xi_t(y,\mathfrak{z})) - (D^\alpha k_\varepsilon)(\cdot - y))\big|_{L_p}|u_t(y)|\,dy\Big|^p\nu_1(d\mathfrak{z})$$

$$\leqslant K_\xi^{p-2}|D^{\alpha+1}k_\varepsilon|^p_{L_p}|\bar{\xi}|^2_{L_2(\mathfrak{z}_1)}\Big(\int_{\mathbb{R}^d}(K_0 + K_1|y| + K_1|Y_t|)|u_t(y)|\,dy\Big)^p. \qquad \text{(V.5.18)}$$

By (V.5.14)–(V.5.18) we have a constant $N = N(p, d, \varepsilon, |\bar{\xi}|_{L_2(\mathfrak{z}_1)}, K_\xi)$ such that

$$A \leqslant N\int_0^T |u_t|_{L_1}\,dt + N\int_0^T\Big(\int_{\mathbb{R}^d}(K_0 + K_1|y| + K_1|Y_t|)|u_t(y)|\,dy\Big)^p dt < \infty\,\text{(a.s.)}.$$

Next we claim that, with the operator T^ξ defined in (V.4.23), we have

$$\zeta_2(\alpha,t) + \zeta_3(\alpha,t)$$

$$= \int_0^t\int_{\mathfrak{z}_1}|D^\alpha(T_s^{\xi*}u_s)^{(\varepsilon)}|^p_{L_p} - |D^\alpha u_s^{(\varepsilon)}|^p_{L_p}\tilde{N}_1(d\mathfrak{z},ds) =: \zeta(\alpha,t) \qquad \text{(V.5.19)}$$

for $t \in [0,T]$. For that purpose not first that $D^\alpha u^{(\varepsilon)} + D^\alpha(I^{\xi(\mathfrak{z}_1)*}u)^{(\varepsilon)} = D^\alpha(T^{\xi*}u_s)^{(\varepsilon)}$. To see that the stochastic integral $\zeta(\alpha,t)$ is well-defined as an Itô integral note that by Lemma V.4.6,

$$\int_0^T\int_{\mathfrak{z}_1}||D^\alpha(T_s^{\xi*}u_s)^{(\varepsilon)}|^p_{L_p} - |D^\alpha u_s^{(\varepsilon)}|^p_{L_p}|^2\,\nu_1(d\mathfrak{z})ds$$

$$\leqslant N|\bar{\xi}|^2_{L_2(\mathfrak{z}_1)}\int_0^T|u_s|^{2p}_{W_p^m}\,ds < \infty\,\text{(a.s.)} \qquad \text{(V.5.20)}$$

with a constant $N = N(d,p,m,\lambda,K_\xi)$. Since $\mathfrak{z}_1$ is σ-finite, there is an increasing sequence $(\mathfrak{z}_{1n})_{n=1}^\infty$, $\mathfrak{z}_{1n} \in \mathcal{Z}_1$, such that $\nu_1(\mathfrak{z}_{1n}) < \infty$ for every n and $\cup_{n=1}^\infty\mathfrak{z}_{1n} = \mathfrak{z}_1$. Then it is easy to see that

$$\bar{\zeta}_{2n}(\alpha,t) = p\int_0^t\int_{\mathfrak{z}_1}\mathbf{1}_{\mathfrak{z}_{1n}}(\mathfrak{z})\big((D^\alpha u_s^{(\varepsilon)})^{p-1}, D^\alpha(I_s^{\xi*}u_s)^{(\varepsilon)}\big)\,N(d\mathfrak{z},ds),$$

$$\hat{\zeta}_{2n}(\alpha,t) = p\int_0^t\int_{\mathfrak{Z}_1}\mathbf{1}_{\mathfrak{Z}_{1n}}(\mathfrak{z})\big((D^\alpha u_s^{(\varepsilon)})^{p-1}, D^\alpha(I_s^{\xi*}u_s)^{(\varepsilon)}\big)\,\nu_1(d\mathfrak{z})ds,$$

$$\bar{\zeta}_{3n}(\alpha,t) = \int_0^t\int_{\mathfrak{Z}_1}\mathbf{1}_{\mathfrak{Z}_{1n}}(\mathfrak{z})\mathcal{R}_p(\alpha,\xi_s(\mathfrak{z}),u_{s-},k_\varepsilon)\,N_1(d\mathfrak{z},ds),$$

$$\hat{\zeta}_{3n}(\alpha,t) = \int_0^t\int_{\mathfrak{Z}_1}\mathbf{1}_{\mathfrak{Z}_{1n}}(\mathfrak{z})\mathcal{R}_p(\alpha,\xi_s(\mathfrak{z}),u_{s-},k_\varepsilon)\,\nu_1(d\mathfrak{z})ds$$

are well-defined, and

$$\zeta_2(\alpha,t) = \lim_{n\to\infty}(\bar{\zeta}_{2n}(\alpha,t)-\hat{\zeta}_{2n}(\alpha,t)), \quad \zeta_3(\alpha,t) = \lim_{n\to\infty}\bar{\zeta}_{3n}(\alpha,t) - \lim_{n\to\infty}\hat{\zeta}_{3n}(\alpha,t),$$

where the limits are understood in probability. Hence

$$\zeta_2(\alpha,t)+\zeta_3(\alpha,t) = \lim_{n\to\infty}\Big(\bar{\zeta}_{2n}(\alpha,t)+\bar{\zeta}_{3n}(\alpha,t) - \big(\hat{\zeta}_{2n}(t)+\hat{\zeta}_{3n}(\alpha,t)\big)\Big)$$

$$= \lim_{n\to\infty}\Big(\int_0^t\int_{\mathfrak{Z}_1}\mathbf{1}_{\mathfrak{Z}_{1n}}(\mathfrak{z})(|D^\alpha(T_s^{\xi*}u_s)^{(\varepsilon)}|_{L_p}^p - |D^\alpha u_s^{(\varepsilon)}|_{L_p}^p)\,\tilde{N}_1(d\mathfrak{z},ds)\Big) = \zeta(\alpha,t),$$

which completes the proof of (V.5.19). Consequently, from (V.5.10)-(V.5.12) we have

$$d|D^\alpha u_t^{(\varepsilon)}|_{L_p}^p = \mathcal{Q}_p(\alpha,b_t,\sigma_t,\rho_t,\beta_t,u_t,k_\varepsilon)\,dt + \int_{\mathfrak{Z}_0}\mathcal{Q}_p^{(0)}(\alpha,\eta_t(\mathfrak{z}_0),u_t,k_\varepsilon)\,\nu_0(d\mathfrak{z})\,dt$$

$$+\int_{\mathfrak{Z}_1}\mathcal{Q}_p^{(1)}(\alpha,\xi_t(\mathfrak{z}_1),u_t,k_\varepsilon)+\mathcal{R}_p(\alpha,\xi_t(\mathfrak{z}_1),u_t,k_\varepsilon)\,\nu_1(d\mathfrak{z})\,dt + d\zeta_1(\alpha,t)+d\zeta(\alpha,t). \tag{V.5.21}$$

By Lemma V.4.1, Corollary V.4.2 and Lemma V.4.3 we have

$$Q_p(\alpha,b_s,\sigma_s,\rho_s,\beta_s,u_s,k_\varepsilon) \leqslant N(L^2+K^2)|u_s|_{W_p^m}^p \tag{V.5.22}$$

with a constant $N = N(d,p,m)$, and by Lemma V.4.4 and Corollary V.4.5, using that $\bar{\xi} \leqslant K_\xi$ and $\bar{\eta} \leqslant K_\eta$,

$$\begin{aligned}\mathcal{Q}_p^{(0)}(\alpha,\eta_s(\mathfrak{z}),u_s,k_\varepsilon) &\leqslant N\bar{\eta}^2(\mathfrak{z})|u_s|_{W_p^m}^p,\\ (\mathcal{Q}_p^{(1)}+\mathcal{R}_p)(\alpha,\xi_s(\mathfrak{z}),u_s,k_\varepsilon) &\leqslant N\bar{\xi}^2(\mathfrak{z})|u_s|_{W_p^m}^p\end{aligned} \tag{V.5.23}$$

with a constant $N = N(K_\xi,K_\eta,d,p,\lambda,m)$. Thus from (V.5.21) we obtain that for all α with $|\alpha| \leqslant m$ almost surely

$$|D^\alpha u_t^{(\varepsilon)}|_{L_p}^p \leqslant |u_0^{(\varepsilon)}|_{W_p^m}^p + N\int_0^t |u_s|_{W_p^m}^p\,ds + m_t^{(\varepsilon)} \quad \text{for all } t\in[0,T]$$

with a constant $N = N(m,p,d,K,K_\xi,K_\eta,L,\lambda,|\bar{\xi}|_{L_2(\mathfrak{Z}_1)},|\bar{\eta}|_{L_2(\mathfrak{Z}_0)})$ and the local martingale $m^{(\varepsilon)}(\alpha, t) = \zeta_1(\alpha, t) + \zeta(\alpha, t)$. Summing over all $|\alpha| \leqslant m$ gives

$$|u_t^{(\varepsilon)}|_{W_p^m}^p \leqslant |u_0^{(\varepsilon)}|_{W_p^m}^p + N\int_0^t |u_s|_{W_p^m}^p \, ds + m_t^{(\varepsilon)} \quad \text{for all } t \in [0, T] \tag{V.5.24}$$

with another constant $N = N(m, p, d, K, K_\xi, K_\eta, L, \lambda, |\bar{\xi}|_{L_2(\mathfrak{Z}_1)}, |\bar{\eta}|_{L_2(\mathfrak{Z}_0)})$ and another local martingale, denoted again by $m^{(\varepsilon)}$. For integers $n, k \geqslant 1$ and $\varepsilon > 0$ set $\tau_{n,k}^\varepsilon = \bar{\tau}_n \wedge \tilde{\tau}_k^\varepsilon$, where $(\tilde{\tau}_k^\varepsilon)_{n=1}^\infty$ is a localising sequence of stopping times for $m^{(\varepsilon)}$ and

$$\bar{\tau}_n = \inf\Big\{t \in [0, T] : \int_0^t |u_s|_{W_p^m}^p \, ds \geqslant n\Big\}.$$

Then from (V.5.24), using also $|D^\alpha u^{(\varepsilon)}|_{L_p} = |(D^\alpha u)^{(\varepsilon)}|_{L_p} \leqslant |D^\alpha u|_{L_p}$ for multi-indices $\alpha \leqslant m$ and $\varepsilon > 0$ we get

$$\begin{aligned}
\mathbb{E}|u_{t\wedge\tau_{n,k}^\varepsilon}^{(\varepsilon)}|_{W_p^m}^p &\leqslant \mathbb{E}|u_0|_{W_p^m}^p + N\mathbb{E}\int_0^{t\wedge\tau_{n,k}^\varepsilon} |u_s|_{W_p^m}^p \, ds \\
&\leqslant \mathbb{E}|u_0|_{W_p^m}^p + N\mathbb{E}\int_0^{t\wedge\bar{\tau}_n} |u_s|_{W_p^m}^p \, ds \\
&\leqslant \mathbb{E}|u_0|_{W_p^m}^p + N\mathbb{E}\int_0^{t} |u_{s\wedge\bar{\tau}_n}|_{W_p^m}^p \, ds < \infty,
\end{aligned}$$

for $t \in [0, T]$ and integers $n \geqslant 1$. Applying Fatou's lemma, first for the limit $k \to \infty$ and then for the limit $\varepsilon \to 0$, followed by Grönwall's lemma gives

$$\mathbb{E}|u_{t\wedge\tau_n}|_{W_p^m}^p \leqslant N\mathbb{E}|u_0|_{W_p^m}^p \quad \text{for } t \in [0, T] \text{ and integers } n \geqslant 1$$

with a constant $N = N(m, p, d, T, K, K_\xi, K_\eta, L, \lambda, |\bar{\xi}|_{L_2}, |\bar{\eta}|_{L_2})$. Letting here $n \to \infty$, by Fatou's lemma we obtain

$$\sup_{t\in[0,T]} \mathbb{E}|u_t|_{W_p^m}^p \leqslant N\mathbb{E}|u_0|_{W_p^m}^p. \tag{V.5.25}$$

To prove (V.5.6) we define a localizing sequence of stopping times $(\tau_k^\varepsilon)_{k=1}^\infty$ for the local martingale m^ε, as well as

$$\tilde{\rho}_n = \inf\Big\{t \in [0, T] : \int_0^t |u_s|_{W_p^m}^{2p} \, ds \geqslant n\Big\}, \quad \text{and} \quad \rho_{n,k}^\varepsilon = \bar{\rho}_n \wedge \tilde{\rho}_k^\varepsilon.$$

Using the Davis inequality and Lemma V.4.3 by standard calculations for every $n \geqslant 1$ we get for each $|\alpha| \leqslant m$ for the Doob-Meyer process of ζ_1,

$$\mathbb{E}\sup_{t\leqslant T} |\zeta_1(\alpha, t \wedge \rho_{n,k}^\varepsilon)| \leqslant 3\mathbb{E}\Big(\sum_k \int_0^{T\wedge\rho_{n,k}^\varepsilon} \big((D^\alpha u_s^{(\varepsilon)})^{p-1}, D^\alpha(\mathcal{M}_s^{k*}u_s)^{(\varepsilon)}\big)^2 \, ds\Big)^{1/2} \tag{V.5.26}$$

$$\leqslant N\mathbb{E}\Big(\int_0^{T\wedge\rho^\varepsilon_{n,k}} |u_s|^{2p}_{W^m_p}\,ds\Big)^{1/2} < \infty,$$

and similarly, for each $|\alpha|\leqslant m$, the Doob-Meyer process of $\zeta(\alpha,\cdot)$ is

$$\langle\zeta(\alpha,\cdot)\rangle(t) = \int_0^t\int_{\mathfrak{Z}_1} ||D^\alpha(T^{\xi*}u_s)^{(\varepsilon)}|^p_{L_p} - |D^\alpha u^{(\varepsilon)}_s|^p_{L_p}|^2\nu_1(\mathfrak{z})ds, \quad t\in[0,T].$$

Using the Davis inequality and Lemma V.4.6,

$$\mathbb{E}\sup_{s\leqslant T}|\zeta(\alpha,s\wedge\rho^\varepsilon_{n,k})| \leqslant 3\mathbb{E}\langle\zeta(\alpha,\cdot)\rangle^{1/2}(T\wedge\rho^\varepsilon_{n,k}) \leqslant N\mathbb{E}\Big(\int_0^{T\wedge\rho^\varepsilon_{n,k}} |u_s|^{2p}_{W^m_p}\,ds\Big)^{1/2} < \infty, \tag{V.5.27}$$

with a constant $N = N(m,d,p,K,K_\xi,L,\lambda,|\bar\xi|_{L_2(\mathfrak{Z}_1)})$. Thus, due to (V.5.25) together with (V.5.26) and (V.5.27), we get from (V.5.24), with constant N depending only on m, p, d, T, K, K_ξ, K_η, L, λ, $|\bar\xi|_{L_2}$ and $|\bar\eta|_{L_2}$,

$$\mathbb{E}\sup_{t\in[0,T]}|u^{(\varepsilon)}_{t\wedge\rho^\varepsilon_{n,k}}|^p_{W^m_p} \leqslant N\mathbb{E}|u_0|^p_{W^m_p} + \sum_{|\alpha|\leqslant m}\mathbb{E}\sup_{t\leqslant T}|\zeta_1(\alpha,t\wedge\rho^\varepsilon_{n,k})| + \sum_{|\alpha|\leqslant m}\mathbb{E}\sup_{t\leqslant T}|\zeta(\alpha,t\wedge\rho^\varepsilon_{n,k})|$$

$$\leqslant N\mathbb{E}|u_0|^p_{W^m_p} + N\mathbb{E}\Big(\int_0^T |u_{s\wedge\rho_n}|^{2p}_{W^m_p}\,ds\Big)^{1/2}.$$

Letting here $k\to\infty$ and then $\varepsilon\to 0$, we obtain by Fatou's lemma with constants N' and N'' only depending on m, p, d, T, K, K_ξ, K_η, L, λ, $|\bar\xi|_{L_2(\mathfrak{Z}_1)}$ and $|\bar\eta|_{L_2(\mathfrak{Z}_0)}$,

$$\mathbb{E}\sup_{t\in[0,T]}|u_{t\wedge\rho_n}|^p_{W^m_p} \leqslant N\mathbb{E}|u_0|^p_{W^m_p} + N\mathbb{E}\Big(\int_0^T |u_{s\wedge\rho_n}|^{2p}_{W^m_p}\,ds\Big)^{1/2}$$

$$\leqslant N\mathbb{E}|u_0|^p_{W^m_p} + N\mathbb{E}\Big(\sup_{t\in[0,T]}|u_{t\wedge\rho_n}|^p_{W^m_p}\int_0^T |u_{s\wedge\rho_n}|^p_{W^m_p}\,ds\Big)^{1/2}$$

$$\leqslant N\mathbb{E}|u_0|^p_{W^m_p} + \tfrac{1}{2}\mathbb{E}\sup_{t\in[0,T]}|u_{t\wedge\rho_n}|^p_{W^m_p} + N'\mathbb{E}\int_0^T |u_{s\wedge\rho_n}|^p_{W^m_p}\,ds$$

$$\leqslant N''\mathbb{E}|u_0|^p_{W^m_p} + \tfrac{1}{2}\mathbb{E}\sup_{t\in[0,T]}|u_{t\wedge\rho_n}|^p_{W^m_p},$$

where we used Young's inequality. Thus also, we get for all n,

$$\mathbb{E}\sup_{t\in[0,T]}|u_{t\wedge\rho_n}|^p_{W^m_p} \leqslant 2N''\mathbb{E}|u_0|^p_{W^m_p}.$$

Using Fatou's lemma we get the desired result. □

For an integer $m\geqslant 0$, let $\mathbb{B}^m_0$ denote the space of those functions $\psi\in\bigcap_{p\geqslant 1}\mathbb{W}^m_p$

such that

$$\sum_{k=0}^{m}\sup_{\omega\in\Omega}\sup_{x\in\mathbb{R}^d}|D^k\psi(x)|<\infty\quad\text{and almost surely }\psi(x)=0\text{ for }|x|\geqslant R,$$

for some constant R depending on ψ. It is easy to see that $\mathbb{B}_0^m$ is a dense subspace of $\mathbb{W}_p^m$ for every $p\in[1,\infty)$. For $\varepsilon>0$ let in the following proposition $v^{(\varepsilon)}$ denote the convolution

$$v^{(\varepsilon)}(x)=\int_{\mathbb{R}^d}\chi_\varepsilon(x-y)v(y)\,dy$$

of a Borel function v on $\mathbb{R}^d$, where χ is a smooth, symmetric function of unit integral on $\mathbb{R}^d$, such that $\chi(x)=0$ for $|x|\geqslant 1$ and $\chi_\varepsilon(\cdot):=\varepsilon^{-d}\chi(\cdot/\varepsilon)$. Let

$$\mathcal{M}_t^{\varepsilon k}=\rho_t^{(\varepsilon)ik}D_i+B_t^{(\varepsilon)k},\quad k=1,\dots,d',$$

$$\tilde{\mathcal{L}}_t^{\varepsilon}=a_t^{\varepsilon,ij}D_{ij}+b_t^{(\varepsilon)i}D_i+\beta_t^k\mathcal{M}_t^{\varepsilon k},\quad \beta_t=B(t,X_t,Y_t),$$
$$a_t^{\varepsilon,ij}:=\tfrac{1}{2}\sum_k(\sigma_t^{(\varepsilon)ik}\sigma_t^{(\varepsilon)jk}+\rho_t^{(\varepsilon)ik}\rho_t^{(\varepsilon)jk}),\quad i,j=1,2,...,d$$

and let I^{ξ^ε}, J^{ξ^ε} and J^{η^ε} be defined as I^ξ, J^ξ and J^η, only with $\xi^{(\varepsilon)}$ and $\eta^{(\varepsilon)}$ instead of ξ and η, respectively.

Consider for $\varepsilon\in(0,1)$ the equation

$$\begin{aligned}du_t^\varepsilon=&\tilde{\mathcal{L}}_t^{\varepsilon*}u_t^\varepsilon\,dt+\mathcal{M}_t^{\varepsilon k*}u_t^\varepsilon\,dV_t^k+\int_{\mathfrak{Z}_0}J_t^{\eta^\varepsilon*}u_t^\varepsilon\,\nu_0(d\mathfrak{z})dt\\&+\int_{\mathfrak{Z}_1}J_t^{\xi^\varepsilon*}u_t^\varepsilon\,\nu_1(d\mathfrak{z})dt+\int_{\mathfrak{Z}_1}I_t^{\xi^\varepsilon*}u_t^\varepsilon\,\tilde{N}_1(d\mathfrak{z},dt),\quad\text{with }u_0^\varepsilon=\psi^{(\varepsilon)}.\end{aligned}\qquad\text{(V.5.28)}$$

Lemma V.5.4. *Let Assumptions V.2.1, V.2.2, V.2.4 and V.2.5 hold with $K_1=0$. Consider integers $m\geqslant 0$ and $p\geqslant 2$ even. Assume there is some $R>0$ such that*

$$\big(b_t(x),B_t(x),\sigma_t(x),\rho_t(x),\eta_t(x,\mathfrak{z}_0),\xi_t(x,\mathfrak{z}_1)\big)=0\qquad\text{(V.5.29)}$$

for $\omega\in\Omega$, $t\geqslant 0$, $\mathfrak{z}_0\in\mathfrak{Z}_0$, $\mathfrak{z}_1\in\mathfrak{Z}_1$ and $x\in\mathbb{R}^d$ such that $|x|\geqslant R$. Let $\psi\in\mathbb{B}_0$ such that $\psi(x)=0$ if $|x|\geqslant R$. Then there exists a unique W_p^m-solution $(u_t)_{t\in[0,T]}$ to equation (V.3.2) *with initial condition $u_0=\psi$. Moreover, almost surely $u_t(x)=0$ for dx-almost every $x\in\{x\in\mathbb{R}^d:|x|\geqslant\bar{R}\}$ for every $t\in[0,T]$ for a constant $\bar{R}=\bar{R}(R,K,K_0,K_\xi,K_\eta)$, and*

$$\mathbb{E}\sup_{t\in[0,T]}|u_t|_{W_p^m}^p\leqslant N\mathbb{E}|\psi|_{W_p^m}^p\qquad\text{(V.5.30)}$$

with a constant $N=N(m,d,p,T,K,K_\xi,K_\eta,L,\lambda,|\bar{\xi}|_{L_2},|\bar{\eta}|_{L_2})$.

Proof. By Lemma IV.6.4 (i) for $\varepsilon>0$ sufficiently small there exists a W_p^m-valued weakly cadlag $\mathcal{F}_t$-adapted process $(u_t^\varepsilon)_{t\in[0,T]}$, such that for each $\varphi\in C_0^\infty$ almost surely

$$(u_t^\varepsilon,\varphi) = (\psi^{(\varepsilon)},\varphi) + \int_0^t (u_s^\varepsilon, \tilde{\mathcal{L}}_s^\varepsilon\varphi)\,ds + \int_0^t (u_s^\varepsilon, \mathcal{M}_s^{\varepsilon k}\varphi)\,dV_s^k + \int_0^t\int_{\mathfrak{Z}_0} (u_s^\varepsilon, J_s^{\eta^\varepsilon}\varphi)\,\nu_0(d\mathfrak{z})\,ds$$
$$+ \int_0^t\int_{\mathfrak{Z}_1} (u_s^\varepsilon, J_s^{\xi^\varepsilon}\varphi)\,\nu_1(d\mathfrak{z})\,ds + \int_0^t\int_{\mathfrak{Z}_1} (u_s^\varepsilon, I_s^{\xi^\varepsilon}\varphi)\,\tilde N_1(d\mathfrak{z},ds), \tag{V.5.31}$$

holds for all $t\in[0,T]$. By Lemma IV.6.4 (ii), since almost surely $u_t^\varepsilon = 0$ for $|x| \geqslant \bar R$ for all $t\in[0,T]$ for a constant $\bar R = \bar R(R,K,K_0,K_\xi,K_\eta)$, we also have

$$\mathbb{E}\sup_{t\in[0,T]}|u_t^\varepsilon|_{L_1} \leqslant \bar R^{d/q}\big(\mathbb{E}\sup_{t\in[0,T]}|u_t^\varepsilon|_{L_p}^p\big)^{1/p} < \infty$$

for $q = p/(p-1)$. Next, note that the smoothed coefficients $b^{(\varepsilon)}, B^{(\varepsilon)}, \sigma^{(\varepsilon)}, \rho^{(\varepsilon)}, \xi^{(\varepsilon)}$ and $\eta^{(\varepsilon)}$ are bounded and satisfy Assumptions V.2.1, V.2.2, V.2.5 and Assumption V.2.4 (iii) with the same constants K_0, L, K_ξ and K_η, independent of ε, as well as Assumption V.2.4 (ii) with a constant $L' = L'(L,K_0,K)$. By Remark V.2.1 we have that for all $t\in[0,T], \theta\in[0,1], y\in\mathbb{R}^{d'}$ and $\mathfrak{z}_i\in\mathfrak{Z}_i$, $i=0,1$, the mappings

$$\tau^\eta_{t,\theta,\mathfrak{z}_0}(x) = x + \theta\eta_t^{(\varepsilon)}(x,\mathfrak{z}_0), \quad \text{and} \quad \tau^\xi_{t,\theta,\mathfrak{z}_1}(x) = x + \theta\xi_t^{(\varepsilon)}(x,\mathfrak{z}_1)$$

are C^1-diffeomorphisms. Moreover, by Corollary IV.6.3, we know that for ε sufficiently small we have that for all $t\in[0,T], \theta\in[0,1]$ and $\mathfrak{z}_i\in\mathfrak{Z}_i$, $i=0,1$, the mappings

$$(\tau^\eta_{t,\theta,\mathfrak{z}_0})^{(\varepsilon)} = \tau^{\eta^{(\varepsilon)}}_{t,\theta,\mathfrak{z}_0}(x) = x+\theta\eta_t^{(\varepsilon)}(x,\mathfrak{z}_0) \quad \text{and} \quad (\xi^\xi_{t,\theta,\mathfrak{z}_1})^{(\varepsilon)} = \tau^{\xi^{(\varepsilon)}}_{t,\theta,\mathfrak{z}_1}(x) = x+\theta\xi_t^{(\varepsilon)}(x,\mathfrak{z}_1)$$

are also C^1-diffeomorphisms such that

$$|\det D\tau^{\eta^{(\varepsilon)}}_{t,\theta,\mathfrak{z}_0}(x)| \geqslant \lambda' \quad \text{and} \quad |\det D\tau^{\xi^{(\varepsilon)}}_{t,\theta,\mathfrak{z}_1}(x)| \geqslant \lambda',$$

with a $\lambda' = \lambda'(\lambda,K_\xi,K_\eta,K_0)$ independent of ε. Moreover, by Remark V.2.1 we then know that Assumption V.2.4 (i) is satisfied with (another) $\lambda'' = \lambda''(\lambda,K_\xi,K_\eta,K_0)$ independent of ε. Hence by Lemma V.5.3 for each $\varepsilon>0$ also

$$\mathbb{E}|u_T^\varepsilon|^p_{W_p^m} + \mathbb{E}\Big(\int_0^T |u_t^\varepsilon|^r_{W_p^m}\,dt\Big)^{p/r} \leqslant \mathbb{E}|u_T^\varepsilon|^p_{W_p^m} + T^{p/r}\mathbb{E}\sup_{t\in[0,T]}|u_t^\varepsilon|^p_{W_p^m} \leqslant N\mathbb{E}|\psi|^p_{W_p^m} \tag{V.5.32}$$

for a constant $N = N(m,d,p,K,K_\eta,K_\xi,L,T,\lambda,|\bar\xi|_{L_2(\mathfrak{Z}_1)},|\bar\eta_{L_2(\mathfrak{Z}_1)})$ independent of ε for all integers $r\geqslant 1$. Letting $(\varepsilon_n)_{n=1}^\infty$ be the sequence from Lemma IV.6.4 (iii), we know that

$$u_T^{\varepsilon_n} \to u_T \quad \text{weakly in } \mathbb{L}_p(\mathcal{F}_T) \text{ and } \quad u^{\varepsilon_n} \to u \quad \text{weakly in } \mathbb{L}_{p,r} \text{ for integers } r\geqslant 2 \text{ as } n\to\infty$$

where u is the unique L_p-solution to (V.3.2) and, if necessary by passing to a subsequence,

$$u_T^{\varepsilon_n} \to u_T \quad \text{weakly in } \mathbb{W}_p^m(\mathcal{F}_T) \text{ and } \quad u^{\varepsilon_n} \to u \quad \text{weakly in } \mathbb{W}_{p,r}^m \text{ for integers } r \geqslant 2.$$

Letting $r \to \infty$ in (V.5.32) yields

$$\mathbb{E}|u_T|_{W_p^m}^p + \mathbb{E} \operatorname*{ess\,sup}_{t \in [0,T]} |u_t|_{W_p^m}^p < N\mathbb{E}|\psi|_{W_p^m}^p.$$

By Lemma V.3.2 u is weakly cadlag as W_p^m-valued process. Thus we can replace the essential supremum above by the supremum to obtain (V.5.30). By Lemma IV.6.4 (ii) we also have that almost surely $u_t(x) = 0$ for dx-almost every $x \in \{x \in \mathbb{R}^d : |x| \geqslant \bar{R}\}$ for every $t \in [0,T]$ for a constant $\bar{R} = \bar{R}(R, K, K_0, K_\xi, K_\eta)$. This finishes the proof. □

Corollary V.5.5. *Let Assumptions V.2.1, V.2.2, V.2.4 and V.2.5 hold with an integer $m \geqslant 0$. Assume, moreover that the support condition* (V.5.29) *holds for some $R > 0$. Then for every $p \geqslant 2$ there is a linear operator $\mathbb{S}$ defined on $\mathbb{W}_p^m$ such that $\mathbb{S}\psi$ admits a $P \otimes dt$-modification $u = (u_t)_{t \in [0,T]}$ which is a W_p^m-solution to equation* (V.3.2) *for every $\psi \in \mathbb{W}_p^m$, with initial condition $u_0 = \psi$, and*

$$\mathbb{E} \sup_{t \in [0,T]} |u_t|_{W_p^m}^p \leqslant N\mathbb{E}|\psi|_{W_p^m}^p \tag{V.5.33}$$

with a constant $N = N(m, d, p, T, K, K_\xi, K_\eta, L, \lambda, |\bar{\xi}|_{L_2}, |\bar{\eta}|_{L_2})$. Moreover, if $\psi \in \mathbb{W}_p^m$ such that almost surely $\psi(x) = 0$ for $|x| \geqslant R$, then almost surely $u_t(x) = 0$ for $|x| \geqslant \bar{R}$ for $t \in [0,T]$ for a constant $\bar{R} = \bar{R}(R, K, K_0, K_\xi, K_\eta)$.

Proof. By Corollary IV.6.5 we know that there exist linear operators $\mathbb{S}$ and $\mathbb{S}_T$ on $\mathbb{L}_p$ such that $\mathbb{S}\psi$ admits a $P \otimes dt$-modification $u = (u_t)_{t \in [0,T]}$ that is an L_p-solution to (V.3.2) such that $u_T = \mathbb{S}_T\psi$ satisfies equation (V.3.2) for each $\varphi \in C_0^\infty$ almost surely with u_T in place of u_t and $t := T$. By an abuse of notation we refer to this stochastic modification u whenever we write $\mathbb{S}\psi$ in the following. It remains to show that if $\psi \in \mathbb{W}_p^m$, then u is in particular a W_p^m-solution to (V.3.2), i.e. it is weakly cadlag as W_p^m-valued process.
If p is an even integer, then this follows from Lemma V.5.4. Assume p is not an even integer. Then let p_0 be the greatest even integer such that $p_0 \leqslant p$ and let p_1 be the smallest even integer such that $p \leqslant p_1$. By Lemma V.5.4, in particular (V.5.30), we get that

$$|\mathbb{S}_T\psi|_{\mathbb{W}_{p_i}^m(\mathcal{F}_T)} + |\mathbb{S}\psi|_{\mathbb{W}_{p_i,r}^m} \leqslant N_i|\psi|_{\mathbb{W}_{p_i}^m} \quad \text{for } i = 0, 1 \tag{V.5.34}$$

for every $r \in [1, \infty)$ and constants $N_i = N_i(m, d, p_i, T, K, K_\xi, K_\eta, L, \lambda, |\bar{\xi}|_{L_2}, |\bar{\eta}|_{L_2})$, $i = 0, 1$, independent of r. Hence, by a well-known generalization of the Riesz-Thorin interpolation theorem we also get for all $r \geqslant 1$,

$$|\mathbb{S}_T\psi|_{\mathbb{W}_p^m(\mathcal{F}_T)} + |\mathbb{S}\psi|_{\mathbb{W}_{p,r}^m} \leqslant N|\psi|_{\mathbb{W}_p^m} \quad \text{for } i = 0, 1, \tag{V.5.35}$$

for (another) constant $N = N(m, d, p, T, K, K_\xi, K_\eta, L, \lambda, |\bar{\xi}|_{L_2}, |\bar{\eta}|_{L_2})$. Consider a sequence $(\psi^n)_{n=1}^\infty \subset \mathbb{B}_0^m$ such that $\psi^n \to \psi$ in $\mathbb{W}_p^m$. For each n, $u^n = \mathbb{S}\psi^n$ is the unique $W_{p_i}^m$-solution to (V.3.2), $i = 0, 1$, with initial condition ψ^n. By virtue of (V.5.35), using that $|\psi^n - \psi|_{\mathbb{W}_p^m} \to 0$, as $n \to \infty$ we know that also

$$u^n \to u \quad \text{weakly in } \mathbb{W}_{p,r}^m \text{ for every integer } r \geqslant 2 \text{ and } \quad u_T^n \to u_T \quad \text{weakly in } \mathbb{W}_p^m(\mathcal{F}_T),$$

where $u = \mathbb{S}\psi$ is the unique L_p-solution introduced in the beginning of the proof, satisfying (V.5.35). To see that u is weakly cadlag as W_p^m-valued process, note that by letting $r \to \infty$ in (V.5.35) or $\mathbb{S}\psi = u$ and $\mathbb{S}_T\psi = u_T$ yields

$$\mathbb{E}|u_T|^p_{W_p^m(\mathcal{F}_T)} + \mathbb{E} \operatorname*{ess\,sup}_{t \in [0,T]} |u_t|^p_{W_p^m} \leqslant N\mathbb{E}|\psi|^p_{W_p^m},$$

for (another) constant $N = N(m, d, p, T, K, K_\xi, K_\eta, L, \lambda, |\bar{\xi}|_{L_2}, |\bar{\eta}|_{L_2})$. By Lemma V.3.2 we then know that u is weakly cadlag as W_p^m-valued process. Thus we can replace the essential supremum above with the supremum, to obtain (V.5.33). To prove the claim about the support of u, note that if $\psi(x) = 0$ for $|x| \geqslant R$, for a constant R, and $\psi^n \to \psi$ in $\mathbb{W}_p^m$, then for sufficiently large n we have $\psi^n(x) = 0$ for $|x| \geqslant 2R$. By Lemma V.5.4 (ii) thus also $u_t^n(x) = 0$ for dx-almost every $x \in \{x \in \mathbb{R}^d : |x| \geqslant \bar{R}\}$ for every $t \in [0, T]$ and n sufficiently large, for a constant $\bar{R} = \bar{R}(R, K, K_0, K_\xi, K_\eta)$. This is clearly preserved in the limit as $n \to \infty$. This finishes the proof. □

V.6 Proof of Theorem V.2.1

Let χ be a smooth function on $\mathbb{R}$ such that $\chi(r) = 1$ for $r \in [-1, 1]$, $\chi(r) = 0$ for $|r| \geqslant 2$, $\chi(r) \in [0, 1]$ and $\sum_{k=1}^{m+2} |d^k/(dr^k)\chi(r)| \leqslant C$ for all $r \in \mathbb{R}$ and a real nonnegative constant C. For integers $n \geqslant 1$ we define the function χ_n by $\chi_n(x) = \chi(|x|/n)$, $x \in \mathbb{R}^d$.

Lemma V.6.1. *Let $b = (b^i)$ be an $\mathbb{R}^d$-valued function on $\mathbb{R}^m$ such that for a constant $M \geqslant 0$,*

$$\sum_{k=1}^{m} |D^k b(x)| \leqslant M, \quad \text{for all } x \in \mathbb{R}^d. \tag{V.6.1}$$

Then $b_n := \chi_n b$ satisfies (V.6.1) *in place of b for a constant $M' = M'(M, C, m, |b(0)|)$ in place of M.*

Proof. A straight forward calculation yields the result. □

We recall from Chapter IV that to preserve the diffeomorphic property of the mappings

$$\tau^\eta_{t,\mathfrak{z}_0,\theta}(x) = x + \theta\eta_t(x, \mathfrak{z}_0) \quad \text{and} \quad \tau^\xi_{t,\mathfrak{z}_1,\theta}(x) = x + \theta\xi_t(x, \mathfrak{z}_1) \tag{V.6.2}$$

(for all $\omega \in \Omega$, $t \in [0,T]$, $\theta \in [0,1]$ and $\mathfrak{z}_i \in \mathfrak{Z}_i$, $i = 0,1$) as a function of $x \in \mathbb{R}^d$, when the functions ξ and η are truncated, we introduced, for each fixed $R > 0$ and $\epsilon > 0$, the function κ_ϵ^R defined on $\mathbb{R}^d$ by

$$\kappa_\epsilon^R(x) = \int_{\mathbb{R}^d} \phi_\varepsilon^R(x-y)k(y)\,dy, \tag{V.6.3}$$

$$\phi_\varepsilon^R(x) \begin{cases} 1, & |x| \leqslant R+1, \\ 1 + \epsilon \log\left(\frac{R+1}{|x|}\right), & R+1 < |x| < (R+1)e^{1/\epsilon}, \\ 0, & |x| \geqslant (R+1)e^{1/\epsilon}, \end{cases}$$

where k is a nonnegative C^∞ mapping on $\mathbb{R}^d$ with support in $\{x \in \mathbb{R}^d : |x| \leqslant 1\}$.

We summarize the results of Lemmas IV.7.1, IV.7.2, IV.7.3 and some facts from the proof of Theorem IV.2.1 in Section IV.7 in the following proposition. For that purpose, define the functions $b^n = (b^{ni}(t,z))$, $B^n = (B^{nj}(t,z))$, $\sigma^n = (\sigma^{nij}(t,z))$, $\eta^n = (\eta^{ni}(t,z,\mathfrak{z}_0))$ and $\xi^n = (\xi^{ni}(t,z,\mathfrak{z}_1))$ by

$$(b^n, B^n, \sigma^n, \rho^n) = (b, B, \sigma, \rho)\chi_n, \quad (\eta^n, \xi^n) = (\eta, \xi)\bar{\chi}_n \tag{V.6.4}$$

for every integer $n \geqslant 1$, where χ_n and $\bar{\chi}_n$ are functions on $\mathbb{R}^{d+d'}$ defined by $\chi_n(z) = \chi(|z|/n)$ and $\bar{\chi}_n(x,y) = \kappa^n(|x|/n)\chi(|y|/n)$ for $z = (x,y) \in \mathbb{R}^{d+d'}$, with χ introduced at the beginning of this section and with $\kappa^n = \kappa^n_{\varepsilon(n)}$ defined in V.6.3, such that, by the L-biLipschitzness of the mappings in (V.6.2), for all $n \geqslant 1$ the mappings

$$\tau^{\eta^n}_{t,\mathfrak{z}_0,\theta}(x) = x + \theta\eta^n_t(x,\mathfrak{z}_0) \quad \text{and} \quad \tau^{\xi^n}_{t,\mathfrak{z}_1,\theta}(x) = x + \theta\xi^n_t(x,\mathfrak{z}_1)$$

are biLipschitz (for all $\omega \in \Omega$, $t \in [0,T]$, $\theta \in [0,1]$ and $\mathfrak{z}_i \in \mathfrak{Z}_i$, $i = 0,1$).

Proposition V.6.2. *Let the conditions of Theorem IV.2.1 hold for some $p \geqslant 2$. Assume the initial conditional density $\pi_0 = P(X_0 \in dx|\mathcal{F}_0^Y)/dx$ additionally satisfies $\mathbb{E}|\pi_0|^p_{W^m_p} < \infty$ for some integer $m \geqslant 0$. Then there exist sequences*

$$(X_0^n)_{n=1}^\infty, \quad ((X_t^n, Y_t^n)_{t\in[0,T]})_{n=1}^\infty, \quad \textit{as well as} \quad (\pi_0^n)_{n=1}^\infty \quad \textit{and} \quad ((\pi_t^n)_{t\in[0,T]})_{n=1}^\infty$$

such that the following are satisfied:
(i) For each $n \geqslant 1$ the coefficients $b^n, B^n, \sigma^n, \rho^n, \xi^n$ and η^n, defined in (V.6.4)*, satisfy Assumptions V.2.1 and V.2.2 with $K_1 = 0$ and constants $K_0' = K_0'(n, L, K, K_0, K_1, K_\xi, K_\eta)$ and $L' = L'(K, K_0, K_1, L, K_\xi, K_\eta)$ in place of K_0 and L, as well as Assumption V.2.4 with $\lambda' = \lambda'(\lambda, K_0, K_1, K_\xi, K_\eta)$ in place of λ. Moreover, for each $n \geqslant 1$ they satisfy the support condition* (V.5.29) *of Lemma V.5.4 with some $R > 0$ depending only on n.*
(ii) For each $n \geqslant 1$ the random variable X_0^n is $\mathcal{F}_0$-measurable and such that

$$\lim_{n\to\infty} X_0^n = X_0, \quad \omega \in \Omega, \quad \textit{and} \quad \mathbb{E}|X_0^n|^r \leqslant N(1 + \mathbb{E}|X_0|^r),$$

for $r > 0$ with a constant $N = N(r)$ independent of n.
(iii) $Z^n_t = (X^n_t, Y^n_t)$ is the solution to the SDE (I.0.2) *with the coefficients $b^n, B^n, \sigma^n, \rho^n, \xi^n$ and η^n in place of b, B, σ, ρ, ξ and η, respectively, and with initial condition $Z^n_0 = (X^n_0, Y_0)$.*
(iv) For each $n \geqslant 1$ we have $\pi^n_0 = P(X^n_0 \in dx|\mathcal{F}^Y_0)/dx$, $\pi^n_0(x) = 0$ for $|x| \geqslant n+1$ and

$$\lim_{n\to\infty} |\pi^n_0 - \pi_0|_{\mathbb{W}^m_p} = 0.$$

(v) For each $n \geqslant 1$ there exists an L_r-solution u^n to (V.3.2)*, $r = 2, p$, with initial condition π^n_0, such that u^n is the unnormalised conditional density of X^n given Y^n, almost surely*

$$u^n_t(x) = 0 \quad \textit{for } dx\textit{-a.e. } x \in \{x \in \mathbb{R}^d : |x| \geqslant \bar{R}\} \textit{ for all } t \in [0,T]$$

with a constant $\bar{R} = \bar{R}(n, K, K_0, K_\xi, K_\eta)$ and

$$\mathbb{E} \sup_{t\in[0,T]} |u^n_t|^p_{L_p} \leqslant N\mathbb{E}|\pi^n_0|^p_{L_p} \tag{V.6.5}$$

with a constant $N = N(d, d', K, L, K_\xi, K_\eta, T, p, \lambda, |\bar{\xi}|_{L_2}, |\bar{\eta}|_{L_2})$. Moreover,

$$u^n \to u \quad \textit{weakly in } \mathbb{L}_{r,q} \textit{ for } r = p, 2 \textit{ and all integers } q > 1,$$

where u is the unnormalised conditional density of X given Y, satisfying (V.6.5) *with the same constant N and u, π_0 in place of u^n, π^n_0.*
(vi) Consequently, for each $n \geqslant 1$ and $t \in [0,T]$ we have

$$\pi^n_t = P(X^n_t \in dx|\mathcal{F}^{Y^n}_t)/dx = u^n_t(x)^o\gamma^n_t, \quad \textit{almost surely,}$$

as well as

$$\pi_t = P(X_t \in dx|\mathcal{F}^Y_t)/dx = u_t(x)^o\gamma_t, \quad \textit{almost surely,}$$

where ${}^o\gamma^n$ and ${}^o\gamma$ are cadlag positive normalising processes.

Now we are in the position to prove our main result.

Proof of Theorem V.2.1. **Step I.** Assume first that the support condition (V.5.29) holds with some $R > 0$ and that the initial conditional density π_0 is such that $\pi_0(x) = 0$ for $|x| \geqslant R$. By Corollary V.5.5 we know that there exists a W^m_p-solution $(u_t)_{t\in[0,T]}$ to (V.3.2) with initial condition π_0, satisfying

$$\mathbb{E} \sup_{t\in[0,T]} |u_t|^p_{W^m_p} \leqslant N\mathbb{E}|\pi_0|^p_{W^m_p} \tag{V.6.6}$$

with a constant $N = N(m, d, p, T, K, K_\xi, K_\eta, L, \lambda, |\bar{\xi}|_{L_2}, |\bar{\eta}|_{L_2})$. Moreover, we have $u_t = 0$ for $|x| \geqslant \bar{R}$, for a constant $\bar{R} = \bar{R}(R, K, K_0, K_1, K_\xi, K_\eta)$, and hence clearly

$$\sup_{t\in[0,T]}|u_t|_{L_1}\leqslant \bar{R}^{d/q}\sup_{t\in[0,T]}|u_t|_{L_p}\quad\text{and}\quad \sup_{t\in[0,T]}\int_{\mathbb{R}^d}|y|^2|u_t(y)|\,dy<\infty\ \text{(a.s.)},$$

with $q = p/(p-1)$. Since also $\pi_0 = P(X_0 \in dx|\mathcal{F}_0^Y)/dx \in \mathbb{L}_1$, then in particular $\pi_0 \in \mathbb{L}_2$ and hence

$$\mathbb{E}\sup_{t\in[0,T]}|u_t|_{L_2}^2 \leqslant N\mathbb{E}|\pi_0|_{L_2}^2, \tag{V.6.7}$$

with a constant $N = N(d,p,T,K,K_\xi,K_\eta,L,\lambda,|\bar{\xi}|_{L_2},|\bar{\eta}|_{L_2})$. By Lemma V.5.4 u is the unique L_2-solution and therefore by Theorem V.3.1, u is in particular the unnormalised conditional density, i.e., $u_t = d\mu_t/dx$ for all $t \in [0,T]$, almost surely, with μ the unnormalised conditional distribution from Theorem III.1.1. Thus also for each $t \in [0,T]$,

$$\pi_t = P(X_t \in dx|\mathcal{F}_t^Y)/dx = u_t{}^o\gamma_t,\quad \text{almost surely},$$

where ${}^o\gamma_t$ is the $\mathcal{F}_t^Y$-optional projection of the normalizing process γ under P introduced in Chapter III.

Step II. Finally, we dispense with the assumption that the coefficients and the initial condition are compactly supported. Define the functions $b_n, B_n, \sigma_n, \rho_n, \xi_n$ and η_n as in (V.6.4). Note that by Proposition V.6.2, as well as Lemma V.6.1, the truncated coefficients satisfy Assumptions V.2.1 and V.2.2 with $K_1 = K_2 = 0$ and constants $K_0' = K_0'(n,K,K_0,K_1,K_\xi,K_\eta)$ and $L' = L'(K,K_0,K_1,L,K_\xi,K_\eta)$ in place of K_0 and L, the coefficients $b_n, B_n, \sigma_n, \rho_n$ satisfy Assumption V.2.5 with a constant $K' = K'(m,K_0,K_1)$ in place of L, and moreover that the coefficients η_n and ξ_n satisfy Assumption V.2.5 with $K'\bar{\eta}$ and $K'\bar{\xi}$ instead of $\bar{\eta}$ and $\bar{\xi}$ respectively. Furthermore, by Lemma IV.7.3, for each $n \geqslant 1$ the coefficients η_n and ξ_n satisfy Assumption V.2.4 with a constant $\lambda' = \lambda'(\lambda,K_0,K_1,K_\eta,K_\xi)$ in place of λ. Note that K', L' and λ' do not depend on n. Moreover, for each $n \geqslant 1$ they satisfy the support condition (V.5.29) of Lemma V.5.4 for some $R = R(n) > 0$. By assumption, $\pi_0 = P(X_0 \in dx|\mathcal{F}_0^Y)/dx$ exists almost surely and $\mathbb{E}|\pi_0|_{W_p^m}^p < \infty$. Then let $(X_0^n)_{n=1}^\infty$ and $(\pi_0^n)_{n=1}^\infty \subset \mathbb{W}_p^m$ be the sequences from Proposition V.6.2 such that

$$\lim_{n\to\infty}|\pi_0^n - \pi_0|_{\mathbb{W}_p^m} = 0, \tag{V.6.8}$$

$\pi_0^n(x) = 0$ for $|x| \geqslant R(n)$ and $\pi_0^n = P(X_0^n \in dx|\mathcal{F}_0^Y)/dx$ (a.s.), where (X_0^n, Y_0) is the initial condition to the system (I.0.2), and $(R(n))_{n=1}^\infty$ is the sequence of positive numbers from the support condition for the coefficients $(\sigma^n, ..., \xi^n)$. By Step I we know that there exists a W_p^m-solution $(u_t)_{t\in[0,T]}$ to (V.3.2) with initial condition π_0^n, which is the unnormalized conditional density of $X^n = (X_t^n)_{t\in[0,T]}$ given $Y^n = (Y_t^n)_{t\in[0,T]}$, where $Z^n = (X^n, Y^n)$ is the solution to (I.0.2) with initial condition (X_0^n, Y_0). By Proposition V.6.2 (v) we know moreover that

$$u^n \to u\quad \text{weakly in } \mathbb{L}_{r,q} \text{ for } r = p, 2 \text{ and all integers } q > 1,$$

where u is the unnormalised conditional density of X given Y from Theorem V.3.1, satisfying

$$\mathbb{E}\sup_{t\in[0,T]}|u_t|_{L_2}^2 \leqslant N\mathbb{E}|\pi_0|_{L_2}^2,$$

with a constant $N = N(d,p,T,K,K_\xi,K_\eta,L,\lambda,|\bar{\xi}|_{L_2},|\bar{\eta}|_{L_2})$ independent of n. Moreover, u is an L_p-solution to (V.3.2) and by Theorem V.3.1 (i), it is the unique L_2-solution to (V.3.2). It remains to show that u is also a W_p^m-solution to (V.3.2), as well as that it is strongly cadlag as W_p^s-valued process, for $s\in[1,m)$. To prove the former, by (V.6.6) together with (V.6.8) we get that for n sufficiently large,

$$\mathbb{E}|u_T^n|_{W_p^m}^p + \mathbb{E}\left(\int_0^T |u_t^n|_{W_p^m}^r\,dt\right)^{p/r} \leqslant \mathbb{E}|u_T^n|_{W_p^m}^p + T^{p/r}\mathbb{E}\sup_{t\in[0,T]}|u_t^n|_{W_p^m}^p \leqslant 2N\mathbb{E}|\pi_0|_{W_p^m}^p. \tag{V.6.9}$$

Hence we know that

$$u_T^n \to u_T, \quad \text{weakly in } \mathbb{W}_p^m \text{ and } \quad u^n \to u \quad \text{weakly in } \mathbb{W}_{p,r}^m \text{ for any } r>1,$$

where u satisfies for all $r\geqslant 1$,

$$\mathbb{E}|u_T|_{W_p^m}^p + \mathbb{E}\left(\int_0^T |u_t|_{W_p^m}^r\,dt\right)^{p/r} \leqslant 2N\mathbb{E}|\pi_0|_{W_p^m}.$$

Letting $r\to\infty$ above yields

$$\mathbb{E}|u_T|_{W_p^m}^p + \mathbb{E}\operatorname*{ess\,sup}_{t\in[0,T]}|u_t|_{W_p^m}^p \leqslant 2N\mathbb{E}|\pi_0|_{W_p^m}.$$

By Lemma V.3.2 we then know that u is weakly cadlag as an W_p^m-valued process, i.e. it is a W_p^m-solution to (V.3.2). Clearly, by Proposition V.6.2 (vi), also for each $t\in[0,T]$

$$\pi_t(x) = P(X_t\in dx|\mathcal{F}_t^Y)/dx = u_t(x)^{o}\gamma_t, \quad \text{almost surely,}$$

with ${}^{o}\gamma$ from Theorem V.3.1. We now show that if $m\geqslant 1$ and $K_1=0$, then u is strongly cadlag as W_p^s-valued process for $s\in[0,m)$. For that purpose, recall first that by Lemma V.3.3, u is a strongly cadlag L_p-valued process, as well as weakly cadlag as an W_p^m-valued process. By interpolation we then have a constant $N=N(d,m,s,p)$ such that

$$|u_t-u_{t_n}|_{W_p^s} \leqslant N|u_t-u_{t_n}|_{W_p^m}|u_t-u_{t_n}|_{L_p} \leqslant 2N\zeta|u_t-u_{t_n}|_{L_p},$$

$$|u_{r_n}-u_{r-}|_{W_p^s} \leqslant N|u_{r_n}-u_{r-}|_{W_p^m}|u_{r_n}-u_{r-}|_{L_p} \leqslant 2N\zeta|u_{r_n}-u_{r-}|_{L_p}$$

for any $t\in[0,T)$, $r\in(0,T]$, any strictly decreasing sequences $t_n\to t$ and strictly increasing sequences $r_n\to r$ with $r_n,t_n\in(0,T)$, where u_{r-} denotes the weak limit in W^m_p of u at r from the left, and $\zeta := \sup_{t\in[0,T]} |u_t|_{W^m_p} < \infty$ (a.s.). Letting here $n \to \infty$ we finish the proof. $\square$

Bibliography

[1] David Applebaum. *Lévy processes and stochastic calculus*. Cambridge university press, 2009.

[2] David Applebaum and Stefan Blackwood. The Kalman-Bucy filter for integrable Lévy processes with infinite second moment. *Journal of Applied Probability*, 52(3):636–648, 2015.

[3] Alan Bain and Dan Crisan. *Fundamentals of stochastic filtering*, volume 3. Springer, 2009.

[4] Abhay Bhatt, Balram Rajput, and Jie Xiong. Zakai equation of nonlinear filtering with Ornstein-Uhlenbeck noise: Existence and uniqueness. *Lecture Notes in Pure and Applied Mathematics, p. 67–80, 2004.*

[5] Stefan Blackwood. *Lévy processes and filtering theory*. PhD thesis, University of Sheffield, 2014.

[6] Alessandro Calvia and Giorgio Ferrari. Nonlinear filtering of partially observed systems arising in singular stochastic optimal control. *Applied Mathematics & Optimization*, 85(2):1–43, 2022.

[7] Thomas Cass, Martin Clark, and Dan Crisan. The filtering equations revisited. In *Stochastic Analysis and Applications*, pages 129–162. Springer, 2014.

[8] Claudia Ceci. Risk minimizing hedging for a partially observed high frequency data model. *Stochastics: An International Journal of Probability and Stochastics Processes*, 78(1):13–31, 2006.

[9] Claudia Ceci and Katia Colaneri. Nonlinear filtering for jump diffusion observations. *Advances in Applied Probability*, 44(3):678–701, 2012.

[10] Claudia Ceci and Katia Colaneri. The Zakai equation of nonlinear filtering for jump-diffusion observations: existence and uniqueness. *Applied Mathematics & Optimization*, 69(1):47–82, 2014.

[11] Dan Crisan. The stochastic filtering problem: a brief historical account. *Journal of Applied Probability*, 51(A):13–22, 2014.

[12] Konstantinos Dareiotis, Chaman Kumar, and Sotirios Sabanis. On tamed Euler approximations of sdes driven by Lévy noise with applications to delay equations. *SIAM Journal on Numerical Analysis*, 54(3):1840–1872, 2016.

[13] Claude Dellacherie and Paul-André Meyer. *Probabilities and Potential B: Theory of Martingales*, volume 72. 1982.

[14] Bandhisattambige P.W. Fernando and Erika Hausenblas. Nonlinear filtering with correlated Lévy noise characterized by copulas. *Brazilian Journal of Probability and Statistics*, 32(2):374–421, 2018.

[15] Masatoshi Fujisaki, Gopinath Kallianpur, and Hiroshi Kunita. Stochastic differential equations for the non linear filtering problem. *Osaka Journal of Mathematics*, 9(1):19–40, 1972.

[16] Fabian Germ and István Gyöngy. On partially observed jump diffusions I. The filtering equations. *arXiv:2205.08286*, 2022.

[17] Fabian Germ and István Gyöngy. On partially observed jump diffusions II. The filtering density. *arXiv:2205.14534*, 2022.

[18] Fabian Germ and István Gyöngy. On partially observed jump diffusions III. Regularity of the filtering density. *arXiv:2211.07239*, 2022.

[19] Bronius Grigelionis and Remigijus Mikulevicius. Nonlinear filtering equations for stochastic processes with jumps. *The Oxford handbook of nonlinear filtering*, pages 95–128, 2011.

[20] István Gyöngy and Nicolai V. Krylov. On stochastic equations with respect to semimartingales I. *Stochastics: An International Journal of Probability and Stochastic Processes*, 4(1):1–21, 1980.

[21] István Gyöngy and Sizhou Wu. Itô's formula for jump processes in Lp-spaces. *Stochastic processes and their applications*, 131:523–552, 2021.

[22] István Gyöngy and Sizhou Wu. On Itô formulas for jump processes. *Queueing Systems*, 98(3):247–273, 2021.

[23] István Gyöngy and Sizhou Wu. On l_p-solvability of stochastic integro-differential equations. *Stochastics and Partial Differential Equations: Analysis and Computations*, 9(2):295–342, 2021.

[24] Paul R. Halmos. *Measure theory*, volume 18. Springer, 2013.

[25] Sheng-wu He, Jia-gang Wang, and Jia-an Yan. *Semimartingale theory and stochastic calculus*. Routledge, 2019.

[26] Andrew Heunis and Vladimir Lucic. On the innovations conjecture of nonlinear filtering with dependent data. *Electronic Journal of Probability*, 13:2190–2216, 2008.

[27] Nobuyuki Ikeda and Shinzo Watanabe. *Stochastic differential equations and diffusion processes*. Elsevier, 2014.

[28] Jean Jacod and Albert Shiryaev. *Limit theorems for stochastic processes*, volume 288. Springer Science & Business Media, 2013.

[29] Yuri M. Kabanov, Robert. S. Liptser, and Albert. N. Shiryayev. On absolute continuity of probability measures for Markov-Itô processes. In *Stochastic Differential Systems Filtering and Control*, pages 114–128, Berlin, Heidelberg, 1980. Springer Berlin Heidelberg.

[30] Sato Ken-Iti. *Lévy processes and infinitely divisible distributions*. Cambridge university press, 1999.

[31] Nicolai V. Krylov. On the equivalence of σ-algebras in the filtering problem of diffusion processes. *Theory of Probability & Its Applications*, 24(4):772–781, 1980.

[32] Nicolai V. Krylov. An analytic approach to SPDEs. Stochastic partial differential equations: six perspectives, 185–242. *Math. Surveys Monogr*, 64, 1999.

[33] Nicolai V. Krylov. *Introduction to the theory of random processes*, volume 43. American Mathematical Sociey, 2002.

[34] Nicolai V. Krylov. A simple proof of a result of A. Novikov. *arXiv preprint arXiv:0207013*, 2002.

[35] Nicolai V. Krylov. Kalman–Bucy filter and SPDEs with growing lower-order coefficients in W^1_p spaces without weights. *Illinois Journal of Mathematics*, 54(3):1069–1114, 2010.

[36] Nicolai V. Krylov. Filtering equations for partially observable diffusion processes with Lipschitz continuous coefficients. *The Oxford Handbook of Nonlinear Filtering*, 2011.

[37] Nicolai V. Krylov. On the Itô–Wentzell formula for distribution-valued processes and related topics. *Probability theory and related fields*, 150(1):295–319, 2011.

[38] Nicolai V. Krylov. A few comments on a result of A. Novikov and Girsanov's theorem. *Stochastics*, 91(8):1186–1189, 2019.

[39] Nicolai V. Krylov and Boris L. Rozovskiĭ. On conditional distributions of diffusion processes. *Mathematics of the USSR-Izvestiya*, 12(2):336, 1978.

[40] Nicolai V. Krylov and Aleksandar Zatezalo. A direct approach to deriving filtering equations for diffusion processes. *Applied Mathematics and Optimization*, 42(3):315–332, 2000.

[41] Thomas G. Kurtz and Jie Xiong. Particle representations for a class of nonlinear SPDEs. *Stochastic Processes and their Applications*, 83(1):103–126, 1999.

[42] Marta De León-Contreras, István Gyöngy, and Sizhou Wu. On solvability of integro-differential equations. *Potential Analysis*, 55(3):443–475, 2021.

[43] Robert S. Liptser and Albert N. Shiryaev. *Statistics of Random Processes II: Applications*, volume 2. Springer Science & Business Media, 2001.

[44] Vasileios Maroulas, Xiaoyang Pan, and Jie Xiong. Large deviations for the optimal filter of nonlinear dynamical systems driven by Lévy noise. *Stochastic Processes and their Applications*, 130(1):203–231, 2020.

[45] Bernt Øksendal and Agnès Sulem. *Stochastic Calculus with Lévy Processes*, pages 1–26. Springer International Publishing, Cham, 2019.

[46] Etienne Pardoux. Équations du filtrage non linéaire de la prédiction et du lissage. *Stochastics*, 6(3-4):193–231, 1982.

[47] Darja R. Poklukar. Nonlinear filtering for jump-diffusions. *Journal of computational and applied mathematics*, 197(2):558–567, 2006.

[48] Silvia Popa and Sivaguru S. Sritharan. Nonlinear filtering of Itô-Lévy stochastic differential equations with continuous observations. *Communications on Stochastic Analysis*, 3(3):1, 2009.

[49] Huijie Qiao. Nonlinear filtering of stochastic differential equations with correlated Lévy noises. *Stochastics*, 93(8):1156–1185, 2021.

[50] Huijie Qiao and Jinqiao Duan. Nonlinear filtering of stochastic dynamical systems with Lévy noises. *Advances in Applied Probability*, 47(3):902–918, 2015.

[51] Boris L. Rozovskii. On conditional distributions of degenerate diffusion processes. *Theory of Probability & Its Applications*, 25(1):147–151, 1980.

[52] Marc Yor. Sur les théories du filtrage et de la prédiction. *Séminaire de probabilités de Strasbourg*, 11:257–297, 1977.