Technische Universität München
Fakultät für Mathematik
Lehrstuhl für Mathematische Statistik

Tempo–Spatial Stochastic Integral Processes: Theory and Applications

Carsten Chong

Vollständiger Abdruck der von der Fakultät für Mathematik der Technischen Universität München zur Erlangung des akademischen Grades eines

Doktors der Naturwissenschaften (Dr. rer. nat.)

genehmigten Dissertation.

Vorsitzender:		Univ.-Prof. Dr. Oliver Junge
Prüfer der Dissertation:	1.	Univ.-Prof. Dr. Claudia Klüppelberg
	2.	Prof. Dr. Jean Jacod, Univ. P. et M. Curie, Paris/Frankreich
	3.	Prof. Dr. Mark Podolskij, Aarhus Univ./Dänemark

Die Dissertation wurde am 18.05.2015 bei der Technischen Universität München eingereicht und durch die Fakultät für Mathematik am 17.08.2015 angenommen.

Summary

This dissertation investigates several aspects of tempo–spatial stochastic integral processes and is outlined as follows.

In Chapter 1 we derive explicit integrability conditions for stochastic integrals taken over time and space driven by a random measure. Our main tool is a canonical decomposition of a random measure which extends the results from the purely temporal case. We show that the characteristics of this decomposition can be chosen as predictable strict random measures, and we compute the characteristics of the stochastic integral process. We apply our conditions to a variety of examples, in particular to ambit processes, which represent a rich model class.

In Chapter 2 we examine nonlinear stochastic Volterra equations in space and time driven by Lévy bases, whose solutions, if they exist, form a subclass of ambit processes as discussed in Chapter 1. Under a Lipschitz condition on the nonlinear term, we give existence and uniqueness criteria in weighted function spaces that depend on integrability properties of the kernel and the characteristics of the Lévy basis. Particular attention is devoted to equations with stationary solutions, or more generally, to equations with infinite memory, that is, where the time domain of integration starts at minus infinity. Here, in contrast to the usual case where time is positive, existence and uniqueness depend critically on the size of the kernel and the Lévy characteristics. Furthermore, once the existence of a solution is guaranteed, we analyze its asymptotic stability, that is, whether its moments remain bounded when time goes to infinity. Stability is proved whenever kernel and characteristics are small enough, or the nonlinearity of the equation exhibits a fractional growth of order strictly smaller than one. The results are applied to the stochastic heat equation for illustration.

In Chapter 3 we present two numerical schemes for the simulation of stochastic Volterra equations as in Chapter 2 when the Lévy noise is of pure-jump type. The first one is based on truncating the small jumps of the noise, while the second one

relies on series representation techniques for infinitely divisible random variables. Under reasonable assumptions, we prove for both methods L^p- and almost sure convergence of the approximations to the true solution of the Volterra equation. We give explicit convergence rates in terms of the Volterra kernel and the characteristics of the noise. A simulation study visualizes the most important path properties of the investigated processes.

In Chapter 4 we suggest three superpositions of COGARCH (supCOGARCH) volatility processes driven by Lévy processes or Lévy bases. We investigate second-order properties, jump behaviour, and prove that they exhibit Pareto-like tails. Corresponding price processes are defined and studied as well. We find that the supCOGARCH models allow for more flexible autocovariance structures than the COGARCH. Moreover, in contrast to most financial volatility models, the supCOGARCH processes do not exhibit a deterministic relationship between price and volatility jumps. Furthermore, one of the supCOGARCH volatility processes has a representation as an ambit process.

In Chapter 5 we investigate interacting particle systems with two kinds of heterogeneity: one originating from different weights of the linkages, and one concerning their asymptotic relevance when the system becomes large. Hereby, we define a partial mean field system where only the asymptotically vanishing pairs are averaged out, and prove a law of large number result with explicit bounds on the mean squared error. Furthermore, a large deviation result will be established in certain cases. The theory will be illustrated by several examples: on the one hand, we recover the classical results of chaos propagation for homogeneous systems, and, on the other hand, we demonstrate the validity of our assumptions for quite general heterogeneous networks including those arising from preferential attachment random graph models.

Zusammenfassung

Die vorliegende Dissertation untersucht verschiedene Aspekte stochastischer Integralprozesse in Raum und Zeit und hat folgenden Aufbau.

In Kapitel 1 leiten wir explizite Integrabilitätsbedingungen für stochastische Integrale über Raum und Zeit bezüglich eines Zufallsmaßes her. Das wichtigste Hilfsmittel ist dabei eine kanonische Zerlegung des Zufallsmaßes, die entsprechende Resultate aus dem rein zeitlichen Fall erweitert. Wir zeigen, dass die Charakteristiken dieser Zerlegung als vorhersagbare strikte Zufallsmaße gewählt werden können, und berechnen die Charakteristiken von stochastischen Integralprozessen. Wir wenden die gefundenen Bedingungen auf eine Auswahl an Beispielen an, insbesondere auf Ambitprozesse, welche eine umfangreiche Modellklasse darstellen.

In Kapitel 2 untersuchen wir nichtlineare stochastische Volterragleichungen, die von räumlich-zeitlichen Lévybasen getrieben werden. Deren Lösungen bilden, falls sie existieren, eine Unterklasse der in Kapitel 1 diskutierten Ambitprozesse. Unter einer Lipschitzbedingung an den nichtlinearen Term formulieren wir Existenz- und Eindeutigkeitskriterien in gewichteten Funktionenräumen, welche nur von den Integrabilitätseigenschaften des Kerns und der Charakteristiken der Lévybasis abhängen. Besondere Aufmerksamkeit wird Gleichungen mit stationären Lösungen oder noch allgemeiner Gleichungen mit unendlichem Gedächtnis gewidmet, das heißt, wo der zeitliche Bereich der Integration bei minus Unendlich beginnt. Im Gegensatz zum gewöhnlichen Fall positiver Zeit, hängen nun Existenz und Eindeutigkeit von Lösungen nicht nur von der Endlichkeit, sondern in ausschlaggebender Weise auch von der Größe des Kerns und der Charakteristiken ab. Darüber hinaus analysieren wir, sobald die Existenz einer Lösung sichergestellt ist, deren asymptotische Stabilität, was also die Frage bedeutet, ob deren Momente beschränkt bleiben, wenn die Zeit gegen Unendlich strebt. Stabilität wird gezeigt, sobald Kern und Charakteristiken genügend klein sind oder wenn die Nichtlinearität der Gleichung mit einer fraktionalen Ordnung strikt kleiner als 1 wächst. Die Ergebnisse werden anhand der

stochastischen Wärmeleitungsgleichung veranschaulicht.

In Kapitel 3 präsentieren wir zwei numerische Verfahren zur Simulation von stochastischen Volterragleichungen wie aus Kapitel 2, wenn das Lévyrauschen nur Sprünge aufweist. Das erste basiert auf dem Abschneiden kleiner Sprünge aus dem Rauschen, während sich das zweite auf Reihendarstellungen für unendlich teilbare Zufallsvariablen stützt. Unter geeigneten Bedingungen beweisen wir für beide Methoden L^p- sowie fast sichere Konvergenz der Approximationen gegen die wahre Lösung der Volterragleichung. Wir geben explizite Konvergenzraten an, welche von dem Volterrakern und den Charakteristiken der Lévybasis abhängen. Eine Simulationsstudie veranschaulicht die wichtigsten Pfadeigenschaften der untersuchten Prozesse.

In Kapitel 4 konstruieren wir drei Superpositionen von COGARCH-Volatilitätsprozessen (supCOGARCH), die von Lévyprozessen oder -basen getrieben werden. Wir untersuchen die Eigenschaften zweiter Ordnung, das Sprungverhalten und zeigen deren paretoähnliche Tails. Im Gegensatz zu den meisten Volatilitätsmodellen der Finanzmathematik, zeigen supCOGARCH-Prozesse keinen deterministischen Zusammenhang zwischen Preis- und Volatilitätssprüngen. Darüber hinaus hat einer der supCOGARCH-Volatilitätsprozesse eine Darstellung als Ambitprozess.

In Kapitel 5 untersuchen wir interagierende Teilchensysteme mit zwei Arten von Heterogenität: eine, die durch verschiedene Gewichte der Verbindungen entsteht, und eine, die deren asymptotische Relevanz beim Wachsen des Systems betrifft. Dabei definieren wir ein partielles Molekularfeldsystem, bei dem nur über asymptotisch verschwindende Paare gemittelt wird, und zeigen ein Resultat im Sinne eines Gesetzes der großen Zahl mit expliziten Schranken für den mittleren quadratischen Fehler. Außerdem wird ein Prinzip großer Abweichungen in bestimmten Fällen bewiesen. Die Theorie wird anhand einiger Beipiele verdeutlicht: zum einen gewinnen wir klassische Resultate zur Chaosausbreitung bei homogenenen Systemen wieder, und zum anderen zeigen wir die Gültigkeit unserer Annahmen bei recht allgemeinen heterogenen Netzwerken einschließlich solcher, die durch den Mechanismus eines Preferential-Attachment-Zufallsgraphen entstehen.

Acknowledgements

First of all, I want to express my gratitude to my supervisor Claudia Klüppelberg. Besides having introduced me to a fascinating research area, she has constantly supported and helped me during the past few years. I am greatly indebted to her for integrating me into the scientific community from the very beginning and for encouraging me to present my work at international conferences and to speak with top scientists about it.

Next, I take pleasure in thanking Jean Jacod for extremely useful discussions and invaluable advice over the years. His constructive comments and expertise on the subject have significantly improved the results of my research. I also thank him for his hospitality during my visit to the Université Pierre et Marie Curie in April 2015. I have benefited a lot from discussions with him and other members of the university.

Furthermore, I am also grateful to Rama Cont for enlightening discussions on systemic risk and to the Isaac Newton Institute of the University of Cambridge for its hospitality during September and October 2014. Discussions with him and other participants of the systemic risk programme have been very fruitful and stimulating.

Also, many thanks go to Ole Barndorff-Nielsen, Mark Podolskij and Almut Veraart for enjoyable stays and conferences at the University of Aarhus and the Imperial College London, to my collegues at the Chair for Mathematical Statistics, to my coauthors Anita Behme and Bohan Chen for pleasant collaborations, and to the graduate program TopMath at Technische Universität München and the Studienstiftung des deutschen Volkes for financial support.

Finally, I want to thank my parents, my sister and the rest of the family for their patience and constant encouragement during the whole time. In particular, I want to thank Julian, whose love and support are of infinite value to me.

Introduction

Itô's pioneering work on stochastic integration [79] is certainly one of the most important cornerstones in modern probability theory. On the one hand, from a mathematical point of view, it culminates in the theory of semimartingales and stochastic differential equations; on the other hand, it lays the foundation for manifold applications in mathematical finance, econometrics, physics, chemistry, engineering and other disciplines. Traditionally, and even today this is the most investigated situation, the stochastic processes under consideration are indexed by a scalar parameter t often referred to as time. However, over the last few decades, there has been a growing interest in models of stochastic processes that additionally feature a space coordinate in an infinite set E. Depending on the model at hand, this set can be physical space, so $E = \mathbb{R}^d$, a parameter space, or a countable index set, e.g. $E = \mathbb{N}$. In some situations, this spatial structure gives rise to interesting phenomena that are usually not encountered with purely temporal stochastic processes.

The present thesis aims to address several topics on such tempo–spatial stochastic processes and is organized as follows. Chapter 1, which is based on a joint paper with my supervisor Claudia Klüppelberg [44] and has been accepted as my Master's Thesis within the TopMath graduate programme at the Technische Universität München, answers the basic question of when a stochastic integral of the form

$$\int_{\mathbb{R}} \int_{E} H(t,x)\, M(\mathrm{d}t, \mathrm{d}x) \tag{1}$$

is well defined for some stochastic integrand H and random measure M. The focus here lies on *global* integrability (not just on a compact time interval) and on easily verifiable conditions. The main results are Theorems 1.4.1 and 1.4.8, where the existence of (1) is characterized in terms of H and the characteristic triplet of M (see Theorem 1.3.2). The main motivation behind this study is to clarify existence conditions for processes of the form

$$Y(t,x) = \int_{\mathbb{R}} \int_{\mathbb{R}^d} G(t,x;s,y)\sigma(s,y)\, M(\mathrm{d}s, \mathrm{d}y), \quad (t,x) \in \mathbb{R}^d, \tag{2}$$

where G is a deterministic kernel and σ a stochastic field. Such processes are called ambit processes in the context of turbulence modelling, see [18], and are related to stochastic partial differential equations as well, see the same reference and in particular Chapter 2.

In Chapter 2, which is taken from the paper [43], we analyze existence and uniqueness conditions for stochastic Volterra equations given by

$$Y(t,x) = Y_0(t,x) + \int_I \int_{\mathbb{R}^d} G(t,x;s,y)\sigma(Y(s,y))\,\Lambda(\mathrm{d}s,\mathrm{d}y) \tag{3}$$

for $(t,x) \in I \times \mathbb{R}^d$, where I is $[0,\infty)$ or $\mathbb{R}$, Y_0 some given process, σ a Lipschitz function and Λ a Lévy basis. When G is the Green's function of a partial differential operator, (3) is the mild formulation of a stochastic partial differential equation with multiplicative noise, see the initial paper [136]. This Chapter has three main contributions: for $I = [0,\infty)$ it extends, under reasonable assumptions, the work of different authors to general kernels G and noises Λ, see Theorems 2.3.1 and 2.3.5; for $I = \mathbb{R}$, which is intimately related to stationary solutions to (3), we find that existence and uniqueness of solutions no longer only depend on integrability properties of kernel and noise, but also on the size of these integrals, see Theorem 2.4.4. Finally, once a solution exists, Theorem 2.5.2 gives conditions for the boundedness of moments when time tends to infinity.

The subsequent Chapter 3 is based on joint work [41] with Bohan Chen, who has created the figures in Section 3.5 in his Master's Thesis [40], and Claudia Klüppelberg. Here we establish two numerical schemes for the simulation of (3) when Λ is of pure-jump type. One method relies on the truncation of the small jumps of Λ (see Algorithm 3.3.1), while the other one relies on series representation techniques (see Algorithm 3.4.2). For both methods we prove L^p- and a.s. convergence of the numerical approximation to the true solution in Theorems 3.3.2 and 3.4.3. As already mentioned, a simulation study is carried out in Section 3.5.

While in Chapters 1–3 the spatial component of the stochastic processes mainly refers to physical space, this changes in Chapter 4, which is from a paper [26] with Anita Behme and Claudia Klüppelberg. Here, we investigate multi-factor continuous-time stochastic volatility models so that space actually means a parameter space and all processes only evolve in time. More precisely, we construct three different superpositions of COGARCH processes, which are the continuous-time analogue [88] to the celebrated GARCH models (see e.g. [30]). Like superpositions of Ornstein-Uhlenbeck processes in [11], the supCOGARCH models are designed to

inherit the desired features of the COGARCH model (volatility clustering, heavy tails, absence of correlation for returns but positive correlation for the squared returns) and, at the same time, to overcome some of its shortcomings (exponential autocovariance function, functional relationship between volatility and price jumps [82]). For all three supCOGARCH models, we compute the moments of the volatility processes in Propositions 4.3.4, 4.3.12 and 4.3.18, prove their Pareto-like tails in Propositions 4.3.5, 4.3.13 and 4.3.19, and show dependence without correlation of the returns in Theorems 4.4.1, 4.4.2 and 4.4.3.

In the last Chapter 5, which is based on [45], we generalize McKean's mean field theory for interacting diffusions (see e.g. [131]) to heterogeneous networks. First, the interaction rates are allowed to vary for different pairs of particles, and second, some of these rates need *not* to vanish when the systems gets large. Furthermore, the driving noises in the system may have jumps and dependencies. Under reasonable assumptions we determine in Theorem 5.3.1 explicit bounds on the mean squared distance between the original system and the partial mean field system (i.e. only the rates that are asymptotically small are taken into the mean field limit). In several examples, which include homogeneous systems, sparse heterogeneous systems or networks arising from a preferential attachment algorithm, and to be discussed in Sections 5.3.1–5.3.3, we show that these bounds converge to 0 as the system grows: a law of large number holds for the partial mean field system. Under certain conditions, also a large deviation principle holds, see Theorem 5.4.1. We remark that in this Chapter the role of space is played by the underlying network structure.

We conclude this introduction with some further remarks. For reasons of brevity, references to existing literature in the overview above are kept to small number and by far not representative. For a more detailed review of the literature, we refer to the introductions within each chapter below. In addition, we made no attempt to unify notations, definitions and abbreviations among the different chapters of this thesis since a large number of symbols and variables are used, and since different abbreviations seem reasonable in different settings. Instead, the reader will find all notational conventions needed to understand the chapters in their respective introductory sections. Therefore, it is also inevitable that certain definitions or explanations are repeated in the different chapters, but this contributes to easier reading and also makes the individual chapters more self-contained.

Contents

Chapter 1:
Integrability conditions for space–time stochastic integrals: theory and applications

1.1 Introduction

Following Itô's seminal paper [79], stochastic integration theory with respect to semimartingales was brought to maturity during the 1970s and 1980s. One of the fundamental results in this area is the Bichteler-Dellacherie theorem, which shows the equivalence between the class of semimartingales and the class of finite L^0-random measures. As a consequence, semimartingales constitute the largest class of integrators that allow for stochastic integrals of predictable integrands satisfying the dominated convergence theorem. The natural analogue to semimartingale integrals in a space–time setting are integrals of the form

$$\int_{\mathbb{R}\times E} H(t,x)\, M(\mathrm{d}t, \mathrm{d}x), \tag{1.1.1}$$

where E is some space and M is an L^0-random measure on $\mathbb{R}\times E$. The construction of such integrals is discussed in [28] in its full generality, so the theory is complete from this point of view.

However, whether H is integrable with respect to M or not, depends on whether

$$\limsup_{r\to 0} \left\{ \mathbb{E}\left[\left| \int S \,\mathrm{d}M \right| \wedge 1 \right] : |S| \leq |rH|, S \text{ is a simple integrand} \right\} = 0 \tag{1.1.2}$$

or not, a property which is hard to check. Thus, the aim of this Chapter is to characterize (1.1.2) in terms of equivalent conditions, which can be verified in concrete

situations. In the purely temporal case, this subject is addressed in [22]. The result there is obtained by using the local semimartingale characteristics corresponding to a random measure. Our approach parallels this method, but it turns out that the notion of characteristics in the space–time setting is much more complex. We will show that, if M has different times of discontinuity (cf. Definition 1.3.1 below), we can associate a characteristic triplet to it consisting of strict random measures (cf. Definition 1.2.1(3)) that are jointly σ-additive in space and time. Moreover, we will determine the characteristics of stochastic integral processes, which is more involved than in the temporal case, since a concept is needed to merge space and time appropriately. Having achieved this step, integrability conditions in the same fashion as in [22, 119] can be given for space–time integrals. We will also compare our results to those of [80], [119], [136].

Applications of our theoretical results will be chosen from the class of ambit processes

$$Y(t,x) := \int_{\mathbb{R}\times\mathbb{R}^d} h(t,s;x,y)\sigma(s,y)\, M(\mathrm{d}s,\mathrm{d}y), \quad t \in \mathbb{R}, \quad x \in \mathbb{R}^d, \tag{1.1.3}$$

which have been suggested for modelling physical space–time phenomena like turbulence, see e.g. [18]. In the case, where $\sigma = 1$ and M is a Lévy basis (see Remark 1.4.4), such multiparameter integrals have already been investigated by many authors: for instance, [37, 98, 122] discuss path properties of the resulting process Y, while [60, 106] address the extremal behaviour of Y; mixing conditions are examined in [69].

As a broad model class, the applications of ambit processes go far beyond turbulence modelling. For example, [111] describes the movement of relativistic quantum particles by equations of the form (1.1.3). Moreover, solutions to stochastic partial differential equations driven by random noise are often of the form (1.1.3), cf. [18, 136] and Section 1.5.2. Furthermore, stochastic processes like forward contracts in bond and electricity markets based on a Heath-Jarrow-Morten approach also rely on a spatial structure, cf. [4, 19]. Other applications include brain imaging [84] and tumour growth [13, 83].

The concept of an ambit process has also been successfully invoked to define superpositions of stochastic processes like Ornstein-Uhlenbeck processes or, more generally, continuous-time ARMA (CARMA) processes. In these models, only integrals of deterministic integrands with respect to Lévy bases are involved, so the integration theory of [119] is sufficient. Our integrability conditions, however, allow

for a volatility modulation of the noise, which generates a greater model flexibility. Moreover, in Section 4.3.3 ambit processes are used to define superpositions of continuous-time GARCH (COGARCH) processes. In its simplest case superposition leads to multi-factor models, economically and statistically necessary extensions of the one-factor models; cf. [82]. As we shall see, the supCOGARCH model again needs the integrability criteria we have developed since for this model the volatility σ and the random measure M are not independent.

This Chapter is organized as follows. Section 1.2 introduces the notation and gives a summary on the concept of a random measure and its stochastic integration theory. Section 1.3 derives a canonical decomposition for random measures as known for semimartingales and calculates the characteristic triplet of stochastic integral processes. Section 1.4 presents integrability conditions in terms of the characteristics from Section 1.3. Section 1.5 is dedicated to examples to highlight our results.

1.2 Preliminaries

Let $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in\mathbb{R}}, P)$ be a stochastic basis satisfying the usual assumptions of completeness and right-continuity. Denote the base space by $\bar{\Omega} := \Omega\times\mathbb{R}$ and the optional (resp. predictable) σ-field on $\bar{\Omega}$ by $\mathcal{O}$ (resp. $\mathcal{P}$). Furthermore, fix some Lusin space E, equipped with its Borel σ-field $\mathcal{E}$. Using the abbreviations $\tilde{\Omega} := \Omega \times \mathbb{R} \times E$ and $\tilde{\mathcal{O}} := \mathcal{O} \otimes \mathcal{E}$ (resp. $\tilde{\mathcal{P}} := \mathcal{P} \otimes \mathcal{E}$), we call a function $H\colon \tilde{\Omega} \to \mathbb{R}$ *optional* (resp. *predictable*) if it is $\tilde{\mathcal{O}}$-measurable (resp. $\tilde{\mathcal{P}}$-measurable). We will often use the symbols $\mathcal{O}$ and $\mathcal{P}$ (resp. $\tilde{\mathcal{O}}$ and $\tilde{\mathcal{P}}$) also for the collection of optional and predictable functions from $\bar{\Omega}$ (resp. $\tilde{\Omega}$) to $\mathbb{R}$. We refer to Chapters I and II of [80] for all notions not explicitly explained.

Some further notational conventions: we write $A_t := A\cap(\Omega\times(-\infty, t])$ for $A \in \mathcal{P}$, and $\tilde{A}_t := \tilde{A}\cap(\Omega\times(-\infty, t]\times E)$ for $\tilde{A} \in \tilde{\mathcal{P}}$. $\mathcal{B}_{\mathrm{b}}(\mathbb{R}^d)$ denotes the collection of bounded Borel sets in $\mathbb{R}^d$. Next, if μ is a signed measure and X a finite variation process, we write $|\mu|$ and $|X|$ for the variation of μ and the variation process of X, respectively. Finally, we equip $L^p = L^p(\Omega, \mathcal{F}, P)$, $p \in [0, \infty)$, with the topology induced by

$$\|X\|_p := \mathbb{E}[|X|^p]^{1/p}, \quad p \geq 1, \quad \|X\|_p := \mathbb{E}[|X|^p], \quad 0 < p < 1, \quad \|X\|_0 := \mathbb{E}[|X| \wedge 1]$$

for $X \in L^p$. Among several definitions of a random measure in the literature, the following two are the most frequent ones: in essence, a random measure is either a random variable whose realizations are measures on some measurable space (e.g.

[80, 85]) or it is a σ-additive set function with values in the space L^p (e.g. [28, 91, 104, 119, 136]). Our terminology is as follows:

Definition 1.2.1 Let $(\tilde{O}_k)_{k\in\mathbb{N}}$ be a sequence of sets in $\tilde{\mathcal{P}}$ with $\tilde{O}_k \uparrow \tilde{\Omega}$. Set $\tilde{\mathcal{P}}_M := \bigcup_{k=1}^{\infty} \tilde{\mathcal{P}}|_{\tilde{O}_k}$, which is the collection of all sets $A \in \tilde{\mathcal{P}}$ such that $A \subseteq \tilde{O}_k$ for some $k \in \mathbb{N}$.

(1) An *L^p-random measure* on $\mathbb{R} \times E$ is a mapping $M\colon \tilde{\mathcal{P}}_M \to L^p$ satisfying:

   (a) $M(\emptyset) = 0$ a.s.,

   (b) For pairwise disjoint sets $(A_i)_{i\in\mathbb{N}}$ in $\tilde{\mathcal{P}}_M$ with $\bigcup_{i=1}^{\infty} A_i \in \tilde{\mathcal{P}}_M$ we have

$$M\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} M(A_i) \quad \text{in } L^p.$$

   (c) For all $A \in \tilde{\mathcal{P}}_M$ with $A \subseteq \tilde{\Omega}_t$ for some $t \in \mathbb{R}$, the random variable $M(A)$ is $\mathcal{F}_t$-measurable.

   (d) For all $A \in \tilde{\mathcal{P}}_M$, $t \in \mathbb{R}$ and $F \in \mathcal{F}_t$, we have

$$M\Big(A \cap (F \times (t,\infty) \times E)\Big) = \mathbb{1}_F M\Big(A \cap (\Omega \times (t,\infty) \times E)\Big) \quad \text{a.s.}$$

(2) If $p = 0$, we only say *random measure*; if $\tilde{O}_k$ can be chosen as $\tilde{\Omega}$ for all $k \in \mathbb{N}$, M is called a *finite* random measure; and finally, if E consists of only one point, M is called a *null-spatial* random measure.

(3) A *strict random measure* is a signed transition kernel $\mu(\omega, \mathrm{d}t, \mathrm{d}x)$ from $(\Omega, \mathcal{F})$ to $(\mathbb{R} \times E, \mathcal{B}(\mathbb{R}) \otimes \mathcal{E})$ with the following properties:

   (a) There is a strictly positive function $V \in \tilde{\mathcal{P}}$ such that

$$\int_{\mathbb{R}\times E} V(t,x)\, |\mu|(\mathrm{d}t, \mathrm{d}x) \in L^1.$$

   (b) For $\tilde{\mathcal{O}}$-measurable functions W such that W/V is bounded, the process

$$W * \mu_t := \int_{(-\infty,t]\times E} W(s,x)\, \mu(\mathrm{d}s, \mathrm{d}x), \quad t \in \mathbb{R},$$

   is optional.

Remark 1.2.2 (1) If we can choose $O_k = \Omega \times O'_k$ with $O'_k \uparrow \mathbb{R} \times E$, one popular choice for $(\mathcal{F}_t)_{t\in\mathbb{R}}$ is the *natural filtration* $(\mathcal{F}_t^M)_{t\in\mathbb{R}}$ of M which is the smallest filtration satisfying the usual assumptions such that for all $t \in \mathbb{R}$ we have $M(\Omega \times B) \in \mathcal{F}_t^M$ if $B \subseteq ((-\infty, t] \times E) \cap O'_k$ with some $k \in \mathbb{N}$.

(2) If μ is a positive transition kernel in Definition 1.2.1(3), μ is an optional $\tilde{\mathcal{P}}$-σ-finite random measure in the sense of [80, Chap. II], where also the predictable compensator of a strict random measure is defined. Obviously, a strict random measure *is* a random measure, see [28, Ex. 5 and 6]. □

Stochastic integration theory in space–time with respect to L^p-random measures is discussed in [28], see also [27]. The special case of L^2-integration theory is also discussed in [56, 136]. Let us recall the details involved: a *simple integrand* is a function $\tilde{\Omega} \to \mathbb{R}$ of the form

$$S := \sum_{i=1}^{r} a_i \mathbb{1}_{A_i}, \quad r \in \mathbb{N}, \quad a_i \in \mathbb{R}, \quad A_i \in \tilde{\mathcal{P}}_M, \tag{1.2.1}$$

for which the stochastic integral with respect to M is canonically defined as

$$\int S \,\mathrm{d}M := \sum_{i=1}^{r} a_i M(A_i). \tag{1.2.2}$$

Now consider the collection $\mathcal{S}_M^{\uparrow}$ of positive functions $\tilde{\Omega} \to \mathbb{R}$ which are the pointwise supremum of simple integrands and define the *Daniell mean* $\|\cdot\|_{M,p}^{\mathrm{D}} \colon \mathbb{R}^{\tilde{\Omega}} \to [0, \infty]$ by

- $\|K\|_{M,p}^{\mathrm{D}} := \sup_{S\in\mathcal{S}_M, |S|\le K} \left\| \int S \,\mathrm{d}M \right\|_p$, if $K \in \mathcal{S}_M^{\uparrow}$, and
- $\|H\|_{M,p}^{\mathrm{D}} := \inf_{K\in\mathcal{S}_M^{\uparrow}, |H|\le K} \|K\|_{M,p}^{\mathrm{D}}$ for arbitrary functions $H \colon \tilde{\Omega} \to \mathbb{R}$.

An arbitrary function $H \colon \tilde{\Omega} \to \mathbb{R}$ is called *integrable* with respect to M if there is a sequence of simple integrands $(S_n)_{n\in\mathbb{N}}$ such that $\|H - S_n\|_{M,p}^{\mathrm{D}} \to 0$ as $n \to \infty$. Then the *stochastic integral* of H with respect to M defined by

$$\int H \,\mathrm{d}M := \int_{\mathbb{R}\times E} H(t,x)\, M(\mathrm{d}t, \mathrm{d}x) := \lim_{n\to\infty} \int S_n \,\mathrm{d}M \tag{1.2.3}$$

exists in L^p and does not depend on the choice of $(S_n)_{n\in\mathbb{N}}$. The collection of integrable functions is denoted by $L^{1,p}(M)$ and can be characterized as follows [27, Thms. 3.4.10 and 3.2.24]:

Theorem 1.2.3. *Let $F^{1,p}(M)$ be the collection of functions H with $\|rH\|^{\mathrm{D}}_{M,p} \to 0$ as $r \to 0$. If we identify two functions coinciding up to a set whose indicator function has Daniell mean 0, then*

$$L^{1,p}(M) = \tilde{\mathcal{P}} \cap F^{1,p}(M). \tag{1.2.4}$$

Moreover, the following dominated convergence theorem holds: Let $(H_n)_{n\in\mathbb{N}}$ be a sequence in $L^{1,p}(M)$ converging pointwise to some limit H. If there exists some function $F \in F^{1,p}(M)$ with $|H_n| \leq F$ for each $n \in \mathbb{N}$, both H and H_n are integrable with $\|H - H_n\|^{\mathrm{D}}_{M,p} \to 0$ as $n \to \infty$ and

$$\int H \,\mathrm{d}M = \lim_{n\to\infty} \int H_n \,\mathrm{d}M \quad \text{in } L^p. \tag{DCT}$$

Given a predictable function $H \in \tilde{\mathcal{P}}$, we can obviously define a new random measure $H.M$ in the following way:

$$K \in L^{1,0}(H.M) :\Leftrightarrow KH \in L^{1,0}(M), \quad \int K \,\mathrm{d}(H.M) := \int KH \,\mathrm{d}M. \tag{1.2.5}$$

This indeed defines a random measure provided there exists a sequence $(\tilde{O}_k)_{k\in\mathbb{N}} \subseteq \tilde{\mathcal{P}}$ with $\tilde{O}_k \uparrow \tilde{\Omega}$ and $\mathbb{1}_{\tilde{O}_k} \in L^{1,0}(H.M)$ for all $k \in \mathbb{N}$. But this construction does not extend the class $L^{1,0}(M)$ of integrable functions with respect to M. However, as shown in [28, §3], $L^{1,p}(M)$ can indeed be extended further in the following way. Given an L^p-random measure M, fix some $\tilde{\mathcal{P}}$-measurable function H such that:

$$\text{There exists a predictable process } K\colon \bar{\Omega} \to (0,\infty) \text{ with } KH \in L^{1,p}(M). \tag{1.2.6}$$

Now set $\bar{O}_k := \{K \geq k^{-1}\}$ for $k \in \mathbb{N}$, which obviously defines predictable sets increasing to $\bar{\Omega}$, and then $\mathcal{P}_{H\cdot M} := \{A \in \mathcal{P}\colon A \subseteq \bar{O}_k \text{ for some } k \in \mathbb{N}\}$. Then we define a new null-spatial L^p-random measure by

$$H \cdot M\colon \mathcal{P}_{H\cdot M} \to L^p, (H \cdot M)(A) := \int \mathbb{1}_A H \,\mathrm{d}M.$$

The following is known from [28], see also [22, Thm. A.2]:

(1) If $H \in L^{1,p}(M)$, $H\cdot M$ is a finite L^p-random measure and $\int 1 \,\mathrm{d}(H\cdot M) = \int H \,\mathrm{d}M$.

(2) If $G\colon \bar{\Omega} \to \mathbb{R}$ is a predictable process, we have $G \in L^{1,p}(H \cdot M)$ if and only if $\|rGH\|_{M,p} \to 0$ as $r \to 0$, where for every $\tilde{\mathcal{P}}$-measurable function H we set

$$\|H\|_{M,p} := \sup_{\substack{F\colon \bar{\Omega}\to\mathbb{R} \text{ predictable},\\ |F|\leq 1, FH\in L^{1,p}(M)}} \left\| \int FH \,\mathrm{d}M \right\|_p. \tag{1.2.7}$$

In this case we have $\int G \,\mathrm{d}(H \cdot M) = \int GH \,\mathrm{d}M$.

Therefore, it is reasonable to extend the set of *integrable* functions with respect to M from $L^{1,p}(M)$ to

$$L^p(M) = \{H \in \tilde{\mathcal{P}} \colon H \text{ satisfies (1.2.6) and } \|rH\|_{M,p} \xrightarrow{r\to 0} 0\} \tag{1.2.8}$$

by setting

$$\int H \,\mathrm{d}M := (H \cdot M)(\bar{\Omega}), \quad H \in L^p(M).$$

We remark that in the null-spatial case $L^{1,0}(M) = L^0(M)$. But in general, the inclusion $L^{1,p}(M) \subseteq L^p(M)$ is strict, see [28, §3b] and Example 1.4.7 below.

Let us also remark that [48] introduces a stochastic integral for a Gaussian random measure where the integrands are allowed to be distribution-valued. It is still an open question whether it is possible to extend this to the general setting of L^p-random measures, in particular if $p < 2$; we do not pursue this direction in the present thesis.

In the sequel we will frequently use the following fact from [22, Ex. 3.1]: If M is a finite random measure, the process $(M(\tilde{\Omega}_t))_{t\in\mathbb{R}}$ has a càdlàg modification, which is then a semimartingale up to infinity with respect to to the underlying filtration (see [22, Sect. 2] for a definition). This semimartingale will be also be denoted by $M = (M_t)_{t\in\mathbb{R}}$.

1.3 Predictable characteristics of random measures

Let us introduce three important subclasses of random measures:

Definition 1.3.1 Let M be a random measure where $\tilde{O}_k = O_k \times E_k$ with $O_k \uparrow \bar{\Omega}$ and $E_k \uparrow E$. Set $\mathcal{E}_M := \bigcup_{k=1}^{\infty} \mathcal{E}|_{E_k}$.

(1) M has *different times of discontinuity* if for all $k \in \mathbb{N}$ and disjoint sets U_1, U_2 in $\mathcal{E}_M$ the semimartingales $\mathbb{1}_{O_k\times U_i} \cdot M$, $i = 1, 2$, a.s. never jump at the same time.

(2) M is called *orthogonal* if for all pairs of disjoint sets $U_1, U_2 \in \mathcal{E}_M$ and $k \in \mathbb{N}$ we have $[(\mathbb{1}_{O_k\times U_1} \cdot M)^{\mathrm{c}}, (\mathbb{1}_{O_k\times U_2} \cdot M)^{\mathrm{c}}] = 0$.

(3) M has *no fixed time of discontinuity* if for all $U \in \mathcal{E}_M$, $k \in \mathbb{N}$ and $t \in \mathbb{R}$ we have $\Delta(\mathbb{1}_{O_k\times U} \cdot M)_t = 0$ a.s.

In the next theorem we prove a canonical decomposition for random measures with different times of discontinuity generalizing the results of [80] and [22]. Without this extra assumption on the random measure, only non-explicit results such as [28, Thm. 4.21] or results for $p \geq 2$ as in [94, Thm. 1] are known. We also remark that the integrability conditions in Theorem 1.4.1 will be stated in terms of this decomposition. Some notation beforehand: we write $\mathcal{B}_0(\mathbb{R})$ for the collection of Borel sets on $\mathbb{R}$ which are bounded away from 0. Furthermore, if X is a semimartingale up to infinity, we write $\mathfrak{B}(X)$ for its first characteristic, $[X]$ for its quadratic variation, X^{c} for its continuous part (all of them starting at $-\infty$ with 0), μ^X for its jump measure and ν^X for its predictable compensator. Finally, if $U \in \mathcal{E}$, $M|_U$ denotes the random measure given by $M|_U(A) = M(A \cap (\bar{\Omega} \times U))$ for $A \in \tilde{\mathcal{P}}_M$.

Theorem 1.3.2. *Let M have different times of discontinuity.*

(1) The mappings

$$B(A) := \mathfrak{B}(\mathbb{1}_A \cdot M)_\infty, \quad M^{\mathrm{c}}(A) := (\mathbb{1}_A \cdot M)^{\mathrm{c}}_\infty, \quad A \in \tilde{\mathcal{P}}_M,$$

are random measures, the mapping

$$C(A;B) := [(\mathbb{1}_A \cdot M)^{\mathrm{c}}, (\mathbb{1}_B \cdot M)^{\mathrm{c}}]_\infty, \quad A \in \tilde{\mathcal{P}}_M,$$

is a random bimeasure (i.e. a random measure in both arguments when the other one is fixed) and the set functions

$$\mu(A,V) := \mu^{\mathbb{1}_A \cdot M}(\mathbb{R} \times V), \quad \nu(A,V) := \nu^{\mathbb{1}_A \cdot M}(\mathbb{R} \times V), \tag{1.3.1}$$

defined for $A \in \tilde{\mathcal{P}}_M$ and $V \in \mathcal{B}_0(\mathbb{R})$, can be extended to random measures on $\tilde{\mathcal{P}}_M \otimes \mathcal{B}_0(\mathbb{R})$. Moreover, (B, C, ν) can be chosen as predictable strict random (bi-)measures and form the characteristic triplet *of M.*

(2) Let $A \in \tilde{\mathcal{P}}_M$ and τ be a truncation function (i.e. a bounded function with $\tau(y) = y$ in a neighbourhood of 0). Then $\mathbb{1}_A(t,x)(y - \tau(y))$ (resp. $\mathbb{1}_A(t,x)\tau(y)$) is integrable with respect to μ (resp. $\mu - \nu$), and we have

$$\begin{aligned} M(A) = B(A) + M^{\mathrm{c}}(A) &+ \int_{\mathbb{R}\times E\times\mathbb{R}} \mathbb{1}_A(t,x)(y - \tau(y))\, \mu(\mathrm{d}t, \mathrm{d}x, \mathrm{d}y) \\ &+ \int_{\mathbb{R}\times E\times\mathbb{R}} \mathbb{1}_A(t,x)\tau(y)\, (\mu - \nu)(\mathrm{d}t, \mathrm{d}x, \mathrm{d}y). \end{aligned} \tag{1.3.2}$$

(3) There are a positive predictable strict random measure $A(\omega, \mathrm{d}t, \mathrm{d}x)$, a $\tilde{\mathcal{P}}$-measurable function $b(\omega, t, x)$ and a transition kernel $K(\omega, t, x, \mathrm{d}y)$ from $(\tilde{\Omega}, \tilde{\mathcal{P}})$ to $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ such that for a.e. $\omega \in \Omega$

$$B(\omega, \mathrm{d}t, \mathrm{d}x) = b(\omega, t, x)\, A(\omega, \mathrm{d}t, \mathrm{d}x),$$
$$\nu(\omega, \mathrm{d}t, \mathrm{d}x, \mathrm{d}y) = K(\omega, t, x, \mathrm{d}y)\, A(\omega, \mathrm{d}t, \mathrm{d}x).$$

For the proof of Theorem 1.3.2 let us recall the semimartingale topology of [59] on the space $\mathcal{SM}$ of semimartingales up to infinity, which is induced by

$$\|X\|_{\mathcal{SM}} := \sup_{|H| \le 1, H \in \mathcal{P}} \left\| \int_{-\infty}^{\infty} H_t \, \mathrm{d}X_t \right\|_0, \quad X \in \mathcal{SM}.$$

The following results are known.

Lemma 1.3.3. *(1) Let $(X^n)_{n\in\mathbb{N}} \subseteq \mathcal{SM}$ and (B^n, C^n, ν^n) denote the semimartingale characteristics of X^n. If $X^n \to 0$ in $\mathcal{SM}$, then each of the following semimartingale sequences converges to 0 in $\mathcal{SM}$ as well: B^n, $X^{\mathrm{c},n}$, C^n, $[X^n]$, $(y - \tau(y)) * \mu^n$ and $\tau(y) * (\mu^n - \nu^n)$.*

(2) If $W(\omega, t, y)$ is a positive bounded predictable function, then $W * \mu^n \to 0$ in probability if and only if $W * \nu^n \to 0$ in probability. Moreover, $W * \nu^n < \infty$ a.s. implies $W * \mu^n < \infty$ a.s.

(3) The collection of predictable finite variation processes is closed under the semimartingale topology.

For the first part of this lemma, see [22, Thm. 3.5] and [59, p. 276]. The second part is taken from [22, Lemmata 3.1 and 3.3]. The third assertion is proved in [103, Thm. IV.7].

Proof of Theorem 1.3.2. Let $k \in \mathbb{N}$ and consider the set function given by $(S, U) \mapsto B(S \times U)$ from the semiring $\mathcal{H} := \mathcal{P}|_{O_k} \times \mathcal{E}|_{E_k}$ to L^0. Obviously, it is finitely additive in each component: for fixed U, additivity in time holds by the definition of B, while for fixed S, additivity in space is due to the assumption of different times of discontinuity. By a straightforward induction argument this implies that B is also finitely additive jointly in space and time. Next, let

$$\mathcal{R}(\mathcal{H}) = \left\{ \bigcup_{n=1}^{N} C_n \colon N \in \mathbb{N}, C_n \in \mathcal{H} \text{ pairwise disjoint} \right\}$$

denote the ring generated by $\mathcal{H}$. Setting $B(\bigcup_{n=1}^N C_n) := \sum_{n=1}^N B(C_n)$ one obtains a well-defined extension of B to $\mathcal{R}(\mathcal{H})$, which is consistent with the original definition of B and still finitely additive. Furthermore, since $\mathcal{R}(\mathcal{H})$ contains $O_k \times E_k$, we can further extend B to a measure on $\sigma(\mathcal{H}) = \tilde{\mathcal{P}}|_{\tilde{O}_k}$ using [92, Thm. B.1.1]. We only have to show the implication

$$(A_n)_{n\in\mathbb{N}} \subseteq \mathcal{R}(\mathcal{H}) \quad \text{with} \quad \limsup_{n\to\infty} A_n = \emptyset \implies \lim_{n\to\infty} B(A_n) = 0 \quad \text{in } L^0. \tag{1.3.3}$$

In fact, under the assumption on the left-hand side of (1.3.3), $\mathbb{1}_{A_n} \cdot M \to 0$ in $\mathcal{SM}$:

$$\begin{aligned} \|\mathbb{1}_{A_n} \cdot M\|_{\mathcal{SM}} &= \sup_{|H|\le 1, H\in\mathcal{P}} \left\| \int H \,\mathrm{d}(\mathbb{1}_{A_n} \cdot M) \right\|_0 = \sup_{|H|\le 1, H\in\mathcal{P}} \left\| \int H \mathbb{1}_{A_n} \,\mathrm{d}M \right\|_0 \\ &\le \sup_{S\in\mathcal{S}_M, |S|\le \mathbb{1}_{A_n}} \left\| \int S \,\mathrm{d}M \right\|_0 = \|\mathbb{1}_{A_n}\|_{M,0}^{\mathrm{D}} \overset{n\to\infty}{\longrightarrow} 0 \end{aligned}$$

by (DCT) with $\mathbb{1}_{O_k\times E_k}$ as dominating function. Using Lemma 1.3.3(1), Equation (1.3.3) follows.

This extension still coincides with the definition of B in Theorem 1.3.2: From the construction given in the proof of [92, Thm. B.1.1], we know that given $A \in \tilde{\mathcal{P}}|_{\tilde{O}_k}$, there is a sequence of sets $(A_n)_{n\in\mathbb{N}}$ in $\mathcal{R}(\mathcal{H})$ with $\limsup((A \setminus A_n) \cup (A_n \setminus A)) = \emptyset$ and $B(A_n) \to B(A)$ in L^0 as $n \to \infty$. As above we obtain $\mathbb{1}_{A_n} \cdot M \to \mathbb{1}_A \cdot M$ in $\mathcal{SM}$, which implies the assertion. And of course, B is unique and $B(A)$ does not depend on the choice of $k \in \mathbb{N}$ with $A \subseteq O_k$.

Finally, we prove that B corresponds to a predictable strict random measure. By [28, Thm. 4.10] it suffices to show that for $H \in L^{1,0}(B)$ the semimartingale $H \cdot B$ is predictable and has finite variation on bounded intervals. If $H \in \mathcal{S}_M$, this follows from linearity and the fact that the first characteristic of a semimartingale up to infinity is a predictable finite variation process. In the general case choose a sequence $(S_n)_{n\in\mathbb{N}} \subseteq \mathcal{S}_M$ with $S_n \to H$ pointwise and $|S_n| \le H$ for all $n \in \mathbb{N}$. As $n \to \infty$, we have $S_n \cdot B \to H \cdot B$ in $\mathcal{SM}$ by (DCT). By Lemma 1.3.3(3) we conclude that also $H \cdot B$ is a predictable finite variation process.

For C we fix one argument and apply the same procedure to the other argument; for M^{c} we refer to [28, Thm. 4.13]. Let us proceed to μ and ν, where in both cases we first fix some $V \in \mathcal{B}_0(\mathbb{R})$ with $\inf\{|x| \colon x \in V\} \ge \epsilon > 0$ and $\epsilon < 1$. In order to apply the same construction scheme as for B, only the proof of (1.3.3) is different for μ and ν. To this end, let $(A_n)_{n\in\mathbb{N}}$ be as on the left-hand side of (1.3.3), that is, $\mathbb{1}_{A_n} \cdot M \to 0$ in $\mathcal{SM}$. Now define $\tilde{\tau}(y) = (y \wedge \epsilon) \vee (-\epsilon)$ and choose $K > 1$ such that

$|\tilde{\tau}(y)| \leq K(y^2 \wedge 1)$ for $|y| \geq \epsilon$. Then

$$\begin{aligned}\|\mu(A_n, V)\|_0 &= \left\| \frac{\mathbb{1}_V(y)}{|\tilde{\tau}(y)|} |\tilde{\tau}(y)| * \mu_\infty^{\mathbb{1}_{A_n} \cdot M} \right\|_0 \leq \epsilon^{-1} \left\| \mathbb{1}_V(y) |\tilde{\tau}(y)| * \mu_\infty^{\mathbb{1}_{A_n} \cdot M} \right\|_0 \\ &\leq K\epsilon^{-1} \left\| (y^2 \wedge 1) * \mu_\infty^{\mathbb{1}_{A_n} \cdot M} \right\|_0 \leq K\epsilon^{-1} \left\| [\mathbb{1}_{A_n} \cdot M]_\infty \right\|_0 \to 0,\end{aligned}$$

where the last step follows from Lemma 1.3.3(1). Part (2) of the same lemma yields that also $\nu(A_n, V) \to 0$ in L^0 as $n \to \infty$. Consequently, [28, Thm. 4.12] shows that $\mu(\cdot, V)$ and $\nu(\cdot, V)$ can be chosen as positive strict random measures. Observing that $\mu(A, \cdot)$ (resp. $\nu(A, \cdot)$) is clearly also a positive (and predictable) strict random measure for given $A \in \tilde{\mathcal{P}}_M$, μ (resp. ν) can be extended to a positive (and predictable) strict random measure on the product $\tilde{\mathcal{P}}_M \otimes \mathcal{B}_0(\mathbb{R})$ (see [119, Prop. 2.4]). Of course, ν is the predictable compensator of μ.

The integrability of $\mathbb{1}_A(t,x)(y - \tau(y))$ (resp. $\mathbb{1}_A(t,x)\tau(y)$) with respect to μ (resp. $\mu - \nu$) is an obvious consequence of (1.3.1) and the corresponding statements in the null-spatial case. The canonical decomposition of M follows since both sides of (1.3.2) are random measures coinciding on $\mathcal{H}$.

Finally, part (3) of Theorem 1.3.2 can be proved analogously to Proposition II.2.9 of [80]. □

Remark 1.3.4 If M is additionally orthogonal, we have $C(A; B) = C(A \cap B; A \cap B)$ for all $A, B \in \tilde{\mathcal{P}}_M$. Consequently, we may identify C with $C(A) := [(\mathbb{1}_A \cdot M)^{\mathrm{c}}]_\infty$ for $A \in \tilde{\mathcal{P}}_M$. Of course, C can then be chosen as a predictable strict random measure. □

Next we calculate the characteristics introduced in Theorem 1.3.2 in two concrete situations: first, for the random measure of a stochastic integral process, and second, for a random measure under an absolutely continuous change of measure. Although the results in both cases are comparable with the purely temporal setting, the first task turns out to be the more difficult one. Moreover, the characteristics for stochastic integral processes are of particular importance for our integrability conditions in Section 1.4.

Beforehand, we need some bimeasure theory: it is well known that bimeasures cannot be extended to measures on the product σ-field in general and that integration theory with respect to bimeasures differs from integration theory with respect to measures. Following [39], let two measurable spaces $(\Omega_i, \mathcal{F}_i)$, $i = 1, 2$, and a bimeasure $\beta \colon \mathcal{F}_1 \times \mathcal{F}_2 \to \mathbb{R}$ be given. We call a pair (f_1, f_2) of $\mathcal{F}_i$-measurable functions f_i,

$i = 1, 2$, *strictly β-integrable* if

(1) f_1 (resp. f_2) is integrable with respect to $\beta(\cdot; B)$ for all $B \in \mathcal{F}_2$ (resp. $\beta(A; \cdot)$ for all $A \in \mathcal{F}_1$),

(2) f_2 is integrable with respect to the measure $B \mapsto \int_{\Omega_1} f_1(\omega_1)\, \beta(\mathrm{d}\omega_1; B)$ and f_1 is integrable with respect to the measure $A \mapsto \int_{\Omega_2} f_2(\omega_2)\, \beta(A; \mathrm{d}\omega_2)$,

(3) for all $A \in \mathcal{F}_1$ and $B \in \mathcal{F}_2$, the following integrals are equal:

$$\int_A f_1(\omega_1) \left(\int_B f_2(\omega_2)\, \beta(\mathrm{d}\omega_1; \mathrm{d}\omega_2) \right) = \int_B f_2(\omega_2) \left(\int_A f_1(\omega_1)\, \beta(\mathrm{d}\omega_1; \mathrm{d}\omega_2) \right). \quad (1.3.4)$$

The *strict β-integral* of $(f_1; f_2)$ on $(A; B)$, denoted by $\int_{(A;B)} (f_1; f_2)\, \mathrm{d}\beta$, is then defined as the common value (1.3.4).

The next theorem determines the characteristics of stochastic integral processes, which is [80, Prop. IX.5.3] in the null-spatial case.

Theorem 1.3.5. *Let M be a random measure with different times of discontinuity and $H \in \tilde{\mathcal{P}}$ satisfy* (1.2.6) *with some $K > 0$. Then the null-spatial random measure $H \cdot M$ has characteristics $(B^{H\cdot M}, C^{H\cdot M}, \nu^{H\cdot M})$ given by*

$$\begin{aligned} B^{H\cdot M}(A) &= (H \cdot B)(A) \\ &\quad + \int_{\mathbb{R}\times E\times\mathbb{R}} \mathbb{1}_A(t)[\tau(H(t,x)y) - H(t,x)\tau(y)]\, \nu(\mathrm{d}t, \mathrm{d}x, \mathrm{d}y), \quad (1.3.5) \\ C^{H\cdot M}(A) &= \int_{\mathbb{R}} K_t^{-2}\, \mathrm{d} \left(\int_{(A_t\times E; A_t\times E)} (HK; HK)\, \mathrm{d}C \right), \quad (1.3.6) \\ W(t,y) * \nu^{H\cdot M} &= W(t, H(t,x)y) * \nu \quad (1.3.7) \end{aligned}$$

for all $A \in \mathcal{P}_{H\cdot M}$ and $\mathcal{P} \otimes \mathcal{B}(\mathbb{R})$-measurable functions W such that $W(t,y) * \nu^{H\cdot M}$ exists.

Moreover, if in addition M is orthogonal, then

$$C^{H\cdot M}(\mathrm{d}t) = \int_E H^2(t,x)\, C(\mathrm{d}t, \mathrm{d}x). \quad (1.3.8)$$

Proof. The second part of this theorem is clear as soon as we have proved the first part. Since characteristics are defined locally, we may assume that $H \in L^{1,0}(M)$. We first consider the continuous part $C^{H\cdot M}$: to this end, let $(H_n)_{n\in\mathbb{N}}$ be a sequence of simple integrands with $|H_n| \le |H|$ for all $n \in \mathbb{N}$ and $H_n \to H$ pointwise. Since for simple integrands the claim follows directly from the definition of C and the

bimeasure integral, we would like to use the (DCT) and Lemma 1.3.3(1) on the one hand and the dominated convergence theorem for bimeasure integrals (see [39, Cor. 2.9]) on the other hand to obtain the result. In order to do so, we only have to show that $(H;H)$ is strictly C-integrable, which means by the symmetry of C the following two points: first, that for all $B \in \tilde{\mathcal{P}}_M$, H is integrable with respect to the measure $A \mapsto C(A;B) = [(\mathbb{1}_A \cdot M)^{\mathrm{c}}, (\mathbb{1}_B \cdot M)^{\mathrm{c}}]_\infty$, and second, that H is integrable with respect to the measure $A \mapsto \int H(t,x)\,\mathrm{d}C(A;\mathrm{d}t,\mathrm{d}x) = [(\mathbb{1}_A \cdot M)^{\mathrm{c}}, (H \cdot M)^{\mathrm{c}}]_\infty$.

Let G be $\mathbb{1}_B$ or H. From [94], Theorem 2 and its Corollary, we know that there exists a probability measure Q equivalent to P such that M is an $L^2(Q)$-random measure with $G, H \in L^{1,2}(M;Q)$. Since the bounded sets in $L^0(P)$ are exactly the bounded sets in $L^0(Q)$, convergence in $\|\cdot\|^{\mathrm{D}}_{M,0;P}$ is equivalent to convergence in $\|\cdot\|^{\mathrm{D}}_{M,0;Q}$. Similarly, stochastic integrals and predictable quadratic covariation remain unchanged under Q (cf. [27, Prop. 3.6.20] and [80, Thm. III.3.13]). Consequently, if we write $\gamma(A) := [\mathbb{1}_A \cdot M^{\mathrm{c}}, G \cdot M^{\mathrm{c}}]_\infty$ for $A \in \tilde{\mathcal{P}}_M$, it suffices to show that

$$\sup_{S \in \mathcal{S}_M, |S| \le |rH|} \left\| \int S \,\mathrm{d}\gamma \right\|_{L^0(Q)} = \sup_{S \in \mathcal{S}_M, |S| \le |rH|} \left\| [(S \cdot M)^{\mathrm{c}}, (G \cdot M)^{\mathrm{c}}]_\infty \right\|_{L^0(Q)} \to 0 \quad \text{as } r \to 0.$$

Indeed, using Fefferman's inequality (cf. [27, Thm. 4.2.7]), we can find a constant $R > 0$, which only depends on G, such that

$$\begin{aligned} \sup_{S \in \mathcal{S}_M, |S| \le |rH|} \left\| [(S \cdot M)^{\mathrm{c}}, (G \cdot M)^{\mathrm{c}}]_\infty \right\|_{L^0(Q)} &\le R \sup_{S \in \mathcal{S}_M, |S| \le |rH|} \mathbb{E}_Q\Big[[(S \cdot M)^{\mathrm{c}}]_\infty \Big]^{1/2} \\ &= R \sup_{S \in \mathcal{S}_M, |S| \le |rH|} \|(S \cdot M)^{\mathrm{c}}_\infty\|_{L^2(Q)} = R \|rH\|^{\mathrm{D}}_{M^{\mathrm{c}},2;Q} \to 0 \end{aligned}$$

as $r \to 0$, which finishes the proof for $C^{H \cdot M}$.

For $B^{H \cdot M}$ and $\nu^{H \cdot M}$, we first take some $D \in \mathcal{P} \otimes \mathcal{B}_0(\mathbb{R})$ and claim that

$$\mathbb{1}_D(s,y) * \mu^{H \cdot M} = \mathbb{1}_D(s, H(s,x)y) * \mu. \tag{1.3.9}$$

This identity immediately extends to finite linear combinations of such indicators and thus, by (DCT), also to all functions $W(\omega,t,y)$ for which $W * \mu^{H \cdot M}$ exists. By the definition of the predictable compensator, this statement also passes to the case where μ is replaced by ν.

In order to prove (1.3.9), first observe that the jump process of the semimartingale $H \cdot M$ up to infinity is given by $\Delta(H \cdot M)_t = (H \cdot M)(\Omega \times \{t\} \times E)$. Furthermore, we can assume that D does not contain any points in $\bar{\Omega} \times \{0\}$. Hence, in the case where $H = \mathbb{1}_A$ with $A \in \tilde{\mathcal{P}}_M$, we have for all $t \in \mathbb{R}$

$$\mathbb{1}_D(s,y) * \mu_t^{H \cdot M} = \mathbb{1}_D(s,y) * \mu_t^{\mathbb{1}_A \cdot M} = \mathbb{1}_D(s,y)\mathbb{1}_A(s,x) * \mu_t = \mathbb{1}_D(s, \mathbb{1}_A(s,x)y) * \mu_t.$$

Now a similar calculation yields that (1.3.9) remains true for all $H \in \mathcal{S}_M$. Finally, let $H \in L^{1,0}(M)$. By decomposing $H = H^+ - H^-$ into its positive and negative part, we may assume that $H \geq 0$ and choose a sequence $(H_n)_{n\in\mathbb{N}}$ of simple functions with $H_n \uparrow H$ as $n \to \infty$. As we have already seen in the proof of Theorem 1.3.2, we have $\mathbb{1}_D(s,y) * \mu^{H_n \cdot M} \to \mathbb{1}_D(s,y) * \mu^{H\cdot M}$ in $\mathcal{SM}$. On the other hand, if D is of the form $R \times (a,b]$ with $R \in \mathcal{P}$ and $(a,b] \subseteq (0,\infty)$ or of the form $R \times [a,b)$ with $[a,b) \subseteq (-\infty, 0)$, then $\mathbb{1}_D(\omega,s,H_n(\omega,s,x)y) \to \mathbb{1}_D(\omega,s,H(\omega,s,x)y)$ as $n \to \infty$ for every $(\omega,s,x,y) \in \tilde{\Omega} \times \mathbb{R}$, which shows that (1.3.9) holds up to indistinguishability. For general D, use Dynkin's π-λ-lemma [29, Thm. 3.2].

Finally, we compute $B^{H\cdot M}$. The results up to now yield that for all $t \in \mathbb{R}$,

$$\begin{aligned}(H\cdot M)_t - (y-\tau(y)) * \mu^{H\cdot M}_t = {} & (H\cdot B)_t + (H \cdot M^{\mathrm{c}})_t + H(s,x)(y-\tau(y)) * \mu_t \\ & + H(s,x)\tau(y) * (\mu-\nu)_t \\ & - [H(s,x)y - \tau(H(s,x)y)] * \mu_t.\end{aligned}$$

By definition, $B^{H\cdot M}$ is the finite variation part of this special semimartingale, which exactly equals $H \cdot B + [\tau(H(t,x)y) - H(t,x)\tau(y)] * \nu$. □

Finally, we show a Girsanov-type theorem comparable to [80, Thm. III.3.24] for semimartingales. First, let us introduce some notation. We consider another probability measure P' on $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in\mathbb{R}})$ such that $P'_t := P'|_{\mathcal{F}_t}$ is absolutely continuous with respect to $P_t := P|_{\mathcal{F}_t}$ for all $t \in \mathbb{R}$. Then denote by Z the unique P-martingale such that $Z \geq 0$ identically and Z_t is a version of the Radon-Nikodym derivative $\mathrm{d}P'_t/\mathrm{d}P_t$ for all $t \in \mathbb{R}$, cf. [80, Thm. III.3.4].

Now let M be a random measure with different times of discontinuity under the probability measure P with characteristics (B,C,ν) with respect to the truncation function τ. We modify the sequence $(\tilde{O}_k)_{k\in\mathbb{N}}$ of Definition 1.2.1(1) by setting $\tilde{O}'_k := \tilde{O}_k \cap (\Omega \times (-k,k] \times E)$ for $k \in \mathbb{N}$ and $\tilde{\mathcal{P}}'_M := \bigcup_{k=1}^{\infty} \tilde{\mathcal{P}}|_{\tilde{O}'_k}$. Next, we denote the jump measure of M by μ and set $M^P_\mu(W) := \mathbb{E}_P[W * \mu_\infty]$ for all non-negative $\mathcal{F}\otimes\mathcal{B}(\mathbb{R})\otimes\mathcal{E}\otimes\mathcal{B}(\mathbb{R})$-measurable functions W. Furthermore, for every such W, there exists an M^P_μ-a.e. unique $\tilde{\mathcal{P}}\otimes\mathcal{B}(\mathbb{R})$-measurable function $M^P_\mu(W|\tilde{\mathcal{P}}\otimes\mathcal{B}(\mathbb{R}))$ such that

$$M^P_\mu(WU) = M^P_\mu(M^P_\mu(W|\tilde{\mathcal{P}}\otimes\mathcal{B}(\mathbb{R}))U) \quad \text{for all } \tilde{\mathcal{P}}\otimes\mathcal{B}(\mathbb{R})\text{-measurable } U \geq 0.$$

Finally, we set

$$Y(t,x,y) := M_\mu^P(Z/Z_-\mathbb{1}_{\{Z_->0\}}|\tilde{\mathcal{P}}\otimes\mathcal{B}(\mathbb{R}))(t,x,y), \quad t\in\mathbb{R}, \quad x\in E, \quad y\in\mathbb{R},$$
$$C^Z(A) := [(Z_-^{-1}\cdot Z)^{\mathrm{c}}, (\mathbb{1}_A\cdot M)^{\mathrm{c}}]_\infty, \quad A\in\tilde{\mathcal{P}}'_M.$$

In the last line, the stochastic integral process $Z_-^{-1}\cdot Z$ is meant to start at t_0, where $t_0\in\mathbb{R}$ is chosen such that $(\mathbb{1}_A\cdot M)^{\mathrm{c}}=0$ on $(-\infty,t_0]$. Then $C^Z(A)$ is well defined by [80, Prop. III.3.5a] and does not depend on the choice of t_0. Moreover, as in Theorem 1.3.2, one shows that C^Z can be chosen as a positive predictable strict random measure.

The following theorem extends [80, Thm. III.3.24] to the space–time framework.

Theorem 1.3.6. *Under P', M is also a random measure with different times of discontinuity (with respect to $(\tilde{O}'_k)_{k\in\mathbb{N}}$). Its P'-characteristics (B',C',ν') with respect to τ are versions of*

$$B'(\mathrm{d}t,\mathrm{d}x) := B(\mathrm{d}t,\mathrm{d}x) + C^Z(\mathrm{d}t,\mathrm{d}x) + \tau(y)(Y(t,x,y)-1)\,\nu(\mathrm{d}t,\mathrm{d}x,\mathrm{d}y),$$
$$C' := C,$$
$$\nu'(\mathrm{d}t,\mathrm{d}x,\mathrm{d}y) := Y(t,x,y)\,\nu(\mathrm{d}t,\mathrm{d}x,\mathrm{d}y).$$

Proof. Since each set in $\tilde{\mathcal{P}}'_M$ is $\mathcal{F}_t$-measurable for some $t\in\mathbb{R}$, properties (a), (b) and (d) of Definition 1.2.1(1) still hold under P'. Since (c) does not depend on the underlying probability measure, M is also a random measure under $\tilde{P}$. To show that M still has different times of discontinuity under P', it suffices to notice the following: using the notation of Definition 1.3.1, the event that $\mathbb{1}_{O_k\times U_1}\cdot M$ and $\mathbb{1}_{O_k\times U_2}\cdot M$ have a common jump in $\mathbb{R}$ is the union over $n\in\mathbb{N}$ of the events that they have a common jump in $(-\infty,n]$. Since these latter events are $\mathcal{F}_n$-measurable, their P'-probability is 0, as desired. Finally, the characteristics under P' can be derived, up to obvious changes, exactly as in [80, Thm. III.3.24].

1.4 An integrability criterion

The canonical decomposition of M in Theorem 1.3.2 together with Theorem 1.3.5 enables us to reformulate (1.2.8) in terms of conditions only depending on the characteristics of M. This result extends the null-spatial case as found in [80, Thm. III.6.30], [42, Thm. 4.5], [22, Thm. 3.2] or [92, Thm. 9.4.1]. It also generalizes the results of

[119, Thm. 2.7] to predictable integrands and also to random measures which are not necessarily Lévy bases. Our proof mimics the approach in [22, Thm. 3.2] and takes care of the additional spatial structure.

Theorem 1.4.1. *Let M be a random measure with different times of discontinuity whose characteristics with respect to some truncation function τ are given by Theorem 1.3.2. Furthermore, let $H \in \tilde{\mathcal{P}}$ satisfy* (1.2.6)*. Then $H \in L^0(M)$ if and only if each of the following conditions is satisfied a.s.:*

$$\int_{\mathbb{R}\times E} \Bigg| H(t,x)b(t,x) + \int_{\mathbb{R}} \left[\tau(H(t,x)y) - H(t,x)\tau(y)\right] K(t,x,\mathrm{d}y) \Bigg| A(\mathrm{d}t,\mathrm{d}x) < \infty, \tag{1.4.1}$$

$$\int_{\mathbb{R}} K_t^{-2}\, \mathrm{d}\left(\int_{((-\infty,t]\times E;(-\infty,t]\times E)} (HK;HK)\, \mathrm{d}C \right) < \infty, \tag{1.4.2}$$

$$\int_{\mathbb{R}\times E} \int_{\mathbb{R}} \left(1 \wedge (H(t,x)y)^2\right) K(t,x,\mathrm{d}y)\, A(\mathrm{d}t,\mathrm{d}x) < \infty. \tag{1.4.3}$$

If M is additionally orthogonal, the spaces $L^0(M)$ and $L^{1,0}(M)$ are equal and condition (1.4.2) *is equivalent to*

$$\int_{\mathbb{R}\times E} H^2(t,x)\, C(\mathrm{d}t,\mathrm{d}x) < \infty. \tag{1.4.4}$$

The following lemma is a straightforward extension of [119, Lemma 2.8]. We omit its proof.

Lemma 1.4.2. *For $t \in \mathbb{R}$, $x \in E$ and $a \in \mathbb{R}$ define*

$$U(t,x,a) := \left| ab(t,x) + \int_{\mathbb{R}} \left(\tau(ay) - a\tau(y)\right) K(t,x,\mathrm{d}y) \right|,$$
$$\tilde{U}(t,x,a) := \sup_{-1\le c\le 1} U(t,x,ca).$$

Then there exists a constant $\kappa > 0$ such that

$$\tilde{U}(t,x,a) \le U(t,x,a) + \kappa \int_{\mathbb{R}} \left(1 \wedge (ay)^2\right) K(t,x,\mathrm{d}y).$$

Proof of Theorem 1.4.1. We first prove that $H \in L^0(M)$ implies (1.4.1)–(1.4.3). Since $H \cdot M$ is a semimartingale up to infinity, $B^{H\cdot M}(\mathbb{R})$ and $C^{H\cdot M}(\mathbb{R})$ exist. Thus, Theorem 1.3.5 gives the first two conditions. For the last condition observe that $(1 \wedge y^2) * \nu_\infty^{H\cdot M} < \infty$ a.s. is equivalent to $(1 \wedge y^2) * \mu_\infty^{H\cdot M} < \infty$ a.s., which obviously holds since $H \cdot M$ is a semimartingale up to infinity. This completes the first direction of the proof.

For the converse statement, we define $\mathcal{D} := \{G \in \mathcal{P} \colon |G| \leq 1, GH \in L^{1,0}(M)\}$. By (1.2.8) we have to show that the set $\{\int GH \,\mathrm{d}M \colon G \in \mathcal{D}\}$ is bounded in L^0 (i.e. bounded in probability) whenever H satisfies (1.4.1)–(1.4.3). By Theorem 1.3.5,

$$\int GH \,\mathrm{d}M = \int GH \,\mathrm{d}M^{\mathrm{c}} + \tau(GHy) * (\mu - \nu)_\infty + (GHy - \tau(GHy)) * \mu_\infty + B^{GH\cdot M}(\mathbb{R}).$$

We consider each part of this formula separately and show that each of the sets

$$\{B^{GH\cdot M}(\mathbb{R}) \colon G \in \mathcal{D}\}, \tag{1.4.5}$$
$$\{\textstyle\int GH \,\mathrm{d}M^{\mathrm{c}} \colon G \in \mathcal{D}\}, \tag{1.4.6}$$
$$\{\tau(GHy) * (\mu - \nu)_\infty \colon G \in \mathcal{D}\}, \tag{1.4.7}$$
$$\{(GHy - \tau(GHy)) * \mu_\infty \colon G \in \mathcal{D}\} \tag{1.4.8}$$

is bounded in probability.

If $G \in \mathcal{D}$ and $\kappa > 0$ denotes the constant in Lemma 1.4.2, (1.4.1) and (1.4.3) imply

$$\begin{aligned}
&\int_{\mathbb{R}\times E} U(t,x,G_tH(t,x))\, A(\mathrm{d}t,\mathrm{d}x) \leq \int_{\mathbb{R}\times E} \tilde{U}(t,x,G_tH(t,x))\, A(\mathrm{d}t,\mathrm{d}x)\\
\leq\ &\int_{\mathbb{R}\times E} \tilde{U}(t,x,H(t,x))\, A(\mathrm{d}t,\mathrm{d}x)\\
\leq\ &\int_{\mathbb{R}\times E} U(t,x,H(t,x))\, A(\mathrm{d}t,\mathrm{d}x) + \kappa \int_{\mathbb{R}\times E}\int_{\mathbb{R}} \big(1 \wedge (H(t,x)y)\big)\, K(t,x,\mathrm{d}y)\, A(\mathrm{d}t,\mathrm{d}x)\\
<\ &\infty
\end{aligned}$$

a.s., which shows that (1.4.5) is bounded in probability.

Next consider (1.4.6) and fix some $G \in \mathcal{D}$ for a moment. Using Lenglart's in-

equality [80, Lemma I.3.30a], we have for all ϵ, $\eta > 0$

$$\begin{aligned} P\left[\left|\int GH \,\mathrm{d}M^{\mathrm{c}}\right| \geq \epsilon\right] &\leq P\left[\sup_{t\in\mathbb{R}} |(GH \cdot M^{\mathrm{c}})(\bar{\Omega}_t)| \geq \epsilon\right] \\ &= P\left[\sup_{t\in\mathbb{R}} |(GH \cdot M^{\mathrm{c}})(\bar{\Omega}_t)|^2 \geq \epsilon^2\right] \leq \frac{\eta}{\epsilon^2} + P\Big[[GH \cdot M^{\mathrm{c}}]_\infty \geq \eta\Big] \\ &= \frac{\eta}{\epsilon^2} + P\Big[G^2K^{-2} \cdot [KH \cdot M^{\mathrm{c}}]_\infty \geq \eta\Big] \leq \frac{\eta}{\epsilon^2} + P\Big[K^{-2} \cdot [KH \cdot M^{\mathrm{c}}]_\infty \geq \eta\Big]. \end{aligned}$$

Now (1.4.2) allows us to make the quantity on the left-hand side arbitrarily small, independently of $G \in \mathcal{D}$, by first choosing $\eta > 0$ and then $\epsilon > 0$ large enough.

For (1.4.7), we use the abbreviation $W(t,x,y) = \tau(G_t H(t,x)y)$. Lenglart's inequality again yields

$$\begin{aligned} P[|W * (\mu-\nu)_\infty| \geq \epsilon] &\leq P\left[\sup_{t\in\mathbb{R}} |W * (\mu-\nu)_t|^2 \geq \epsilon^2\right] \\ &\leq \frac{\eta}{\epsilon^2} + P\Big[\langle W * (\mu-\nu)\rangle_\infty \geq \eta\Big] \end{aligned} \tag{1.4.9}$$

for every ϵ, $\eta > 0$. Furthermore, by Theorem 1.3.5 and [80, Prop. II.2.17] we have

$$\langle W * (\mu-\nu)\rangle_\infty = \langle \tau(y) * (\mu^{GH\cdot M} - \nu^{GH\cdot M})\rangle_\infty \leq \tau(y)^2 * \nu_\infty,$$

which is finite by (1.4.3) yielding the boundedness of (1.4.7).

Next choose r, $\epsilon > 0$ such that $f(y) := r|y|\mathbb{1}_{\{|y|>\epsilon\}}$ satisfies $|y - \tau(y)| \leq f(y)$ for all $y \in \mathbb{R}$. Obviously, f is symmetric and increasing on $\mathbb{R}_+$ so that

$$\Big|\Big(GHy - \tau(GHy)\Big) * \mu_\infty\Big| \leq f(GHy) * \mu_\infty \leq f(Hy) * \mu_\infty.$$

Now the third condition and Lemma 1.3.3(2) imply that

$$\sum_{t\in\mathbb{R}} (1 \wedge \epsilon^2)\mathbb{1}_{\{|\Delta(H\cdot M)_t|>\epsilon\}} \leq (1 \wedge y^2) * \mu_\infty^{H\cdot M} = \Big(1 \wedge (H(t,x)y)^2\Big) * \mu_\infty < \infty$$

a.s. such that $\{|\Delta(H \cdot M)_t| > \epsilon\}$ only happens for finitely many time points. Hence

$$f(Hy) * \mu_\infty = f(y) * \mu_\infty^{H\cdot M} = r \sum_{t\in\mathbb{R}} |\Delta(H \cdot M)_t| \mathbb{1}_{\{|\Delta(H\cdot M)_t|>\epsilon\}} < \infty$$

a.s., which implies that the set in (1.4.8) is also bounded in probability.

Finally, in the case where M is also orthogonal, we show that (1.4.1),(1.4.4) and (1.4.3) imply $H \in L^{1,0}(M)$. By Theorem 1.2.3 and the fact that for predictable functions H

$$\|H\|^{\mathrm{D}}_{M,0} = \sup_{S\in\mathcal{S}_M,|S|\le|H|} \left\| \int S \,\mathrm{d}M \right\|_0 = \sup_{G\in\tilde{\mathcal{P}},|G|\le 1,GH\in L^{1,0}(M)} \left\| \int GH \,\mathrm{d}M \right\|_0,$$

we have to show that the set $\{\int GH \,\mathrm{d}M \colon G \in \mathcal{D}'\}$ is bounded in L^0, where $\mathcal{D}'$ consists of all functions $G \in \tilde{\mathcal{P}}$ with $|G| \le 1$ and $GH \in L^{1,0}(M)$. Obviously, the previously considered set $\mathcal{D}$ is a subset of $\mathcal{D}'$. Intending to verify (1.4.5)–(1.4.8) with G taken from $\mathcal{D}'$, we observe that all calculations remain valid except those for (1.4.6). For (1.4.6) we argument as follows: for all ϵ, $\eta > 0$, Lenglart's inequality implies

$$\begin{aligned} P\left[\left|\int GH \,\mathrm{d}M^{\mathrm{c}}\right| \ge \epsilon\right] &\le P\left[\sup_{t\in\mathbb{R}} |(GH\cdot M)^{\mathrm{c}}(\bar{\Omega}_t)|^2 \ge \epsilon^2\right] \\ &\le \frac{\eta}{\epsilon^2} + P\Big[[(GH\cdot M)^{\mathrm{c}}]_\infty \ge \eta\Big] = \frac{\eta}{\epsilon^2} + P\left[\int_{\mathbb{R}\times E} G^2(t,x)H^2(t,x)\, C(\mathrm{d}t,\mathrm{d}x) \ge \eta\right] \\ &\le \frac{\eta}{\epsilon^2} + P\left[\int_{\mathbb{R}\times E} H^2(t,x)\, C(\mathrm{d}t,\mathrm{d}x) \ge \eta\right]. \end{aligned}$$

This finishes the proof of Theorem 1.4.1. □

The remaining part of this section illustrates Theorem 1.4.1 by a series of remarks, examples and useful extensions.

Remark 1.4.3 If M has summable jumps, which means that each of the semimartingales $(M(\tilde{\Omega}_t \cap \tilde{O}_k))_{t\in\mathbb{R}}$, $k \in \mathbb{N}$, has summable jumps over finite intervals, it is often convenient to construct the characteristics with respect to $\tau = 0$, which is not a proper truncation function. Then one would like to use $\tau = 0$ in (1.4.1) and replace (1.4.3) by

$$\int_{\mathbb{R}\times E} \int_{\mathbb{R}} \Big(1 \wedge |H(t,x)y|\Big)\, K(t,x,\mathrm{d}y)\, A(\mathrm{d}t,\mathrm{d}x) < \infty. \tag{1.4.10}$$

We show that (1.4.1) with $\tau = 0$, (1.4.2) and (1.4.10) are together sufficient conditions for $H \in L^0(M)$. First note that we can choose $\kappa = 0$ in Lemma 1.4.2(2) since τ is identical 0 and therefore $\tilde{U} = U$. So the calculations done for (1.4.5) remain valid. Moreover, (1.4.6) does not depend on τ and the boundedness of (1.4.7) becomes trivial. For (1.4.8) observe that

$$|GHy|*\mu_\infty \le |Hy|*\mu_\infty = |y|*\mu_\infty^{H\cdot M} = |y|\mathbb{1}_{\{|y|\le 1\}}*\mu_\infty^{H\cdot M} + |y|\mathbb{1}_{\{|y|>1\}}*\mu_\infty^{H\cdot M}. \tag{1.4.11}$$

Now (1.4.10) implies by Lemma 1.3.3(2) that a.s.,

$$|y|\mathbb{1}_{\{|y|\leq 1\}} * \mu_\infty^{H\cdot M} + \mathbb{1}_{\{|y|>1\}} * \mu_\infty^{H\cdot M} < \infty.$$

As a result, on the right-hand side of (1.4.11), the first summand converges a.s. and the second one is in fact just a finite sum a.s.

The converse statement is not true, already in the null-spatial case: let $(N_t)_{t\geq 0}$ be a standard Poisson process and $\tilde{N}_t = N_t - t$, $t \geq 0$, its compensation. Set $H_t := (1+t)^{-1}$ for $t \geq 0$. Then $H \in L^0(\tilde{N})$ as one can see from (1.4.1)–(1.4.3) with the proper truncation function $\tau(y) = y\mathbb{1}_{\{|y|<1\}}$; but $\int_0^\infty H_t \,\mathrm{d}t = \infty$ violating both (1.4.1) with $\tau = 0$ and (1.4.10).

However, if M is a positive (or negative) random measure, that is, $M(A)$ is a positive (or negative) random variable for all $A \in \tilde{\mathcal{P}}_M$, then $C = 0$ necessarily and (1.4.1) with $\tau = 0$ and (1.4.10) also become necessary conditions for H belonging to $L^0(M) = L^{1,0}(M)$; cf. [28, Ex. 5, p. 7, and Thm. 4.12]. □

Next we compare our results and techniques to the standard literature.

Remark 1.4.4 (Lévy bases [119]) Lévy bases are originally called infinitely divisible independently scattered random measures in [119]. They are the space–time analogues of processes with independent increments and have attracted interest in several applications in the last few years, see Section 1.5 for some examples. The precise definition is as follows: Assume that we have $\tilde{O}_k = \Omega \times O'_k$ in the notation of Definition 1.2.1, where $(O'_k)_{k\in\mathbb{N}}$ is a sequence increasing to $\mathbb{R}\times E$. Set $\mathcal{S} := \bigcup_{k=1}^\infty \mathcal{B}(\mathbb{R}^{1+d})|_{O'_k}$. Then a *Lévy basis* Λ is a random measure on $\mathbb{R}\times E$ with the following additional properties:

(1) If $(A_n)_{n\in\mathbb{N}}$ is a sequence of pairwise disjoint sets in $\mathcal{S}$, then $(\Lambda(\Omega\times A_n))_{n\in\mathbb{N}}$ are independent random variables.

(2) For all $A \in \mathcal{S}$, $\Lambda(\Omega\times A)$ has an infinitely divisible distribution.

Note that we have altered the original definition of [119]: in order to perform stochastic integration, we need to single out one coordinate to be time and introduce a filtration based definition of the integrator Λ. For notational convenience, we will write $\Lambda(A)$ instead of $\Lambda(\Omega\times A)$ in the following. As shown in [119, Prop. 2.1 and Lemma 2.3], Λ induces a characteristic triplet (B, C, ν) with respect to some

truncation function τ via the Lévy-Khintchine formula: for all $A \in \mathcal{S}$ and $u \in \mathbb{R}$

$$\mathbb{E}[\mathrm{e}^{\mathrm{i}u\Lambda(A)}] = \exp\left(\mathrm{i}uB(A) - \frac{u^2}{2}C(A) + \int_{\mathbb{R}} (\mathrm{e}^{\mathrm{i}uy} - 1 - \mathrm{i}u\tau(y))\,\nu(A, \mathrm{d}y)\right).$$

It is natural to ask how this notion of characteristics compares with Theorem 1.3.2. Obviously, Λ is an orthogonal random measure. In order that Λ has different times of discontinuity, it suffices by independence to assume that Λ has no fixed times of discontinuity. In this case, recalling the construction in the proof of Theorem 1.3.2 and using [127, Thm. 3.2] together with [80, Thm. II.4.15], one readily sees that the two different definitions of characteristics agree in the natural filtration of Λ. In particular, the canonical decomposition of Λ determines its Lévy-Itô decomposition as derived in [113].

Consequently, the integrability criteria obtained in Theorem 1.4.1 extend the corresponding result of [119, Thm. 2.7] for deterministic functions (or, as used in [18], for integrands which are independent of Λ) to allow for predictable integrands. □

Remark 1.4.5 (Martingale measures [136]) In [136] a stochastic integration theory for predictable integrands is developed with so-called worthy martingale measures as integrators. The concept of worthiness is needed since a martingale measure in Walsh's sense does not guarantee that it is a random measure in the sense of Definition 1.2.1. What is missing is, loosely speaking, a joint σ-additivity condition in space and time; see also the example in [136, pp. 305ff.]. The worthiness of a martingale measure, i.e. the existence of a dominating (σ-additive) measure, turns it into a random measure.

In essence, the integration theory presented in [136] for worthy martingale measures is an L^2-theory similar to [56, 79], where the extension from simple to general integrands is governed by a dominating measure. The latter also determines whether a predictable function is integrable or not in terms of a square-integrability condition; see [136, p. 292]. We see the main advantages of the L^2-theory as follows: it does not require the martingale measure to have different times of discontinuity, works with fairly easy integrability conditions and produces stochastic integrals again belonging to L^2. However, many interesting integrators (e.g. stable noises) are not L^2-random measures. Moreover, even if the integrator M is an L^2-random measure, the class $L^0(M)$ is usually considerably larger than the class $L^2(M)$. Thus, in comparison to [136], it is the compensation of these two shortages of the L^2-theory that constitutes

the main advantage of our integrability conditions in Theorem 1.4.1. We will come back to this point in Section 1.5.2, where it is shown that in the study of stochastic PDEs, solutions often do not exist in the L^2-sense but in the L^0-sense. □

Remark 1.4.6 ((Compensated) strict random measures [80]) Chapters I and II of [80] are an established reference for integration theory with respect to semimartingales. Moreover, they also cover the integration theory with respect to strict random measures or compensated strict random measures as follows: if M is a strict random measure, they define stochastic integrals with respect to M path-by-path. More precisely, a measurable function $H\colon \tilde{\Omega} \to \mathbb{R}$ is pathwise integrable with respect to M if for a.e. $\omega \in \Omega$

$$\int_{\mathbb{R}\times E} |H|(\omega, t, x)\, |M|(\omega, \mathrm{d}t, \mathrm{d}x) < \infty. \tag{1.4.12}$$

If $\tilde{M} := M - M^{\mathrm{p}}$ is the compensation of an integer-valued strict random measure M, we have the following situation: let $H \in \tilde{\mathcal{P}}$ and introduce an auxiliary process by

$$\tilde{H}_t(\omega) := \int_E H(\omega, t, x)\, \tilde{M}(\omega, \{t\} \times \mathrm{d}x), \quad (\omega, t) \in \bar{\Omega}, \tag{1.4.13}$$

hereby setting $\tilde{H}_t(\omega) := +\infty$ whenever (1.4.13) diverges. Then H is integrable in the sense of [80, Def. II.1.27] if there for every $a \in \mathbb{R}$ there exists a sequence of stopping times $(T_n)_{n\in\mathbb{N}}$ with $T_n \uparrow +\infty$ a.s. and

$$\mathbb{E}\left[\left(\sum_{a\le t\le T_n} \tilde{H}_t^2\right)^{1/2}\right] < \infty. \tag{1.4.14}$$

How do these integrability conditions compare to those of Theorem 1.4.1? Obviously, pathwise integrability with respect to M does not require the integrand to be predictable. Furthermore, if H is predictable and (1.4.12) holds, then the pathwise integral coincides with the stochastic integral $H \cdot M$. Still, Theorem 1.4.1 provides a useful extension in some situations: first, there are examples $H \in L^0(M)$ which fail the condition (1.4.12) (see the example at the end of Remark 1.4.3). And second, given some specific H, it may be difficult in general to determine whether (1.4.12) holds or not (e.g., if M has no finite first moment). The characteristic triplet that is used in Theorem 1.4.1 is often easier to handle than $|M|$.

As for $\tilde{M}$ we have following situation: first, one should notice that (1.4.14) ensures integrability on *finite* intervals, whereas Theorem 1.4.1 is concerned with *global*

integrability on $\mathbb{R}$. Second, even on finite intervals, the conditions of Theorem 1.4.1 are more general than (1.4.14), see [28, Prop. 3.10]. Finally, whereas (1.4.14) involves a localizing sequence of stopping times and moment considerations, Theorem 1.4.1 relates integrability only to the integrand itself and the characteristics of $\tilde{M}$, which is often more convenient. □

In order to illustrate condition (1.4.2) in Theorem 1.4.1, we now discuss the example of a Gaussian random measure, which is white in time but coloured in space. Such random measures are often encountered as the driving noise of stochastic PDEs, see [48] and references therein.

Example 1.4.7 Let $(M(\Omega\times B))_{B\in\mathcal{B}_{\mathrm{b}}(\mathbb{R}^{1+d})}$ be a mean-zero Gaussian process whose covariance functional for $B, B' \in \mathcal{B}_{\mathrm{b}}(\mathbb{R}^{1+d})$ is given by

$$C(B;B') := \mathbb{E}[M(\Omega\times B)M(\Omega\times B')] = \int_{\mathbb{R}}\int_{B(t)\times B'(t)} f(x-x')\,\mathrm{d}(x,x')\,\mathrm{d}t, \quad (1.4.15)$$

where $B(t) := \{x\in\mathbb{R}^d\colon (t,x)\in B\}$. For the existence of such a process, it is well known [56, Thm. II.3.1] that $f\colon \mathbb{R}^d\to[0,\infty)$ must be a symmetric and nonnegative definite function for which the integral on the right-hand side of (1.4.15) exists. Under these conditions, C defines a deterministic bimeasure which is symmetric in $B, B' \in \mathcal{B}_{\mathrm{b}}(\mathbb{R}^{1+d})$.

For the further procedure let $(\mathcal{F}_t)_{t\in\mathbb{R}}$ be the natural filtration of M and set

$$M(F\times(s,t]\times U) := \mathbb{1}_F M(\Omega\times(s,t]\times U), \quad F\in\mathcal{F}_s.$$

By [28, Thm. 2.25], M can be extended to a random measure on $\mathbb{R}\times\mathbb{R}^d$ provided that

$$S_n\to 0 \text{ pointwise}, |S_n|\le|S| \implies \int S_n\,\mathrm{d}M\to 0 \quad \text{in } L^0$$

for all step functions S_n and S over sets of the form $F\times(s,t]\times U$ with $F\in\mathcal{F}_s$, $s<t$ and $U\in\mathcal{B}_{\mathrm{b}}(\mathbb{R}^d)$. Indeed, using obvious notation and observing that $\mathbb{1}_F$ is independent of $M(\Omega\times(s,t]\times U)$ for $F\in\mathcal{F}_s$ since M is white in time, we have

$$\begin{aligned}\mathbb{E}\left[\left(\int S_n\,\mathrm{d}M\right)^2\right] &= \sum_{i,j=1}^{r_n} a_i^n a_j^n \mathbb{E}[M(A_i^n)M(A_j^n)]\\ &= \sum_{i,j=1}^{r_n} a_i^n a_j^n P[F_i^n]P[F_j^n]\mathrm{Leb}((s_i^n,t_i^n]\cap(s_j^n,t_j^n])\int_{U_i^n\times U_j^n} f(x-x')\,\mathrm{d}(x,x')\\ &= \int_{(\mathbb{R}^{1+d};\mathbb{R}^{1+d})}(\tilde{S}_n,\tilde{S}_n)\,\mathrm{d}C\to 0\end{aligned}$$

by dominated convergence [39, Cor. 2.9]. Here $\tilde{S}_n$ arises from S_n by replacing a_i^n with $a_i^n P[F_i^n]$.

Having established that M is a random measure, let us derive its characteristics. Obviously, B and ν are identically 0. It is also easy to see that C is the second characteristic of M: it is clear for sets of the form $(s,t] \times U$, and extends to general sets in $\mathcal{B}_{\mathrm{b}}(\mathbb{R}^{1+d})$ by dominated convergence. Therefore, as shown in the proof of Theorem 1.3.5, $L^{1,0}(M)$ consists of those $H \in \tilde{\mathcal{P}}$ such that $(H;H)$ is strictly C-integrable, or, equivalently,

$$\int_{\mathbb{R}} \int_{\mathbb{R}^d \times \mathbb{R}^d} |H|(t,x)|H|(t,x') f(x-x') \,\mathrm{d}(x,x') \,\mathrm{d}t < \infty \quad \text{a.s.} \tag{1.4.16}$$

The class $L^0(M)$, however, is the set of all $H \in \tilde{\mathcal{P}}$ such that a.s. the inner integral in (1.4.16) is finite for a.e. $t \in \mathbb{R}$, and

$$\int_{\mathbb{R}} \int_{\mathbb{R}^d \times \mathbb{R}^d} H(t,x)H(t,x') f(x-x') \,\mathrm{d}(x,x') \,\mathrm{d}t < \infty \quad \text{a.s.} \tag{1.4.17}$$

A (deterministic) function $H \in L^0(M)$ which is not in $L^{1,0}(M)$ is, for instance, given by $H(t,x) := th(x)$ where h is chosen such that

$$\int_{\mathbb{R}^d \times \mathbb{R}^d} h(x)h(x') f(x-x') \,\mathrm{d}(x,x') = 0.$$

One important example is a fractional correlation structure in space. In this case, we have $f(x_1,\dots,x_d) = \prod_{i=1}^d |x_i|^{2H_i-2}$, where $H_i \in (1/2,1)$ is the Hurst index of the i-th coordinate. Then $L^0(M)$ can be interpreted as the extension of the class $|\Lambda_H|$ studied in [115] to several parameters and stochastic integrands. However, in [115] as well as in [21], stochastic integrals are constructed for even larger classes of integrands. These classes, denoted Λ_H or Λ_X, respectively, are obtained as limits of simple functions under L^2-norms ($\|\cdot\|_{\Lambda_H}$ and $\|\cdot\|_{\Lambda_X}$, respectively), which are defined via fractional derivatives or Fourier transforms. In particular, the stochastic integrals defined via these norms are no longer of Itô type, i.e. no dominated convergence theorem holds for these stochastic integrands. Indeed, $L^{1,0}(M)$ is the largest class of predictable integrands for which a dominated convergence theorem holds (see Theorem 1.2.3), and $L^0(M)$ is its improper extension to functions for which $H \cdot M$ is a finite measure. □

The investigation of multi-dimensional stochastic processes often involves stochastic integrals where the integrand H is a matrix-valued predictable function

and the integrator $M = (M^1, \dots, M^d)$ is a d-dimensional random measure, that is, $M^1, \dots, M^d$ are all random measures in the sense of Definition 1.2.1 with respect to the same underlying filtration and the same sequence $(\tilde{O}_k)_{k\in\mathbb{N}}$. By considering each row of H separately, we can assume for the following that H is an $\mathbb{R}^d$-valued predictable function. It is obvious that the construction of stochastic integrals requires no more techniques than those presented in Section 1.2. In fact, replacing E by E^d reduces the multivariate case to the univariate one. However, there is a difference when we want to apply the canonical decomposition as in Theorem 1.3.2 or the integrability conditions in Theorem 1.4.1: in the multi-dimensional case, it is not reasonable to assume that M^i and M^j for $i \neq j$ have different times of discontinuity. Instead, one would define d-dimensional characteristics (B, C, ν) for M, similar to [80, Chap. II] or [22, Thm. 3.1], and use these to characterize integrability.

In the next theorem we rephrase 1.4.1 for the multivariate setting. Since no novel arguments are needed, we omit its proof. We will use the product notation in a self-explanatory way: for instance, if $x, y \in \mathbb{R}^d$, xy denotes their inner product; for $A \in \tilde{\mathcal{P}}_M$, $\mathbb{1}_A \cdot M$ denotes the d-dimensional semimartingale $(\mathbb{1}_A \cdot M^1, \dots, \mathbb{1}_A \cdot M^d)$; $H \cdot M$ denotes $\sum_{i=1}^d H^i \cdot M^i$ for $H \in L^{1,0}(M)$ and is suitably extended to $H \in L^0(M)$, cf. Section 1.2. Similarly, given a matrix $\beta = (\beta^{ij})_{i,j=1}^d$ of bimeasures from $\mathcal{F}_1 \times \mathcal{F}_2 \to \mathbb{R}$ and $\mathcal{F}_i$-measurable functions $f_i = (f_i^1, \dots, f_i^d)$ for $i = 1, 2$, we define

$$\int_{(A;B)} (f_1; f_2) \, \mathrm{d}\beta := \sum_{i,j=1}^d \int_{(A;B)} (f_1^i; f_2^j) \, \mathrm{d}\beta^{ij}, \quad A \in \mathcal{F}_1, \quad B \in \mathcal{F}_2,$$

whenever the right-hand side exists.

Assume that M has different times of discontinuity, which means that $\mathbb{1}_{O_k \times U_i} \cdot M$, $i = 1, 2$, a.s. never jump at the same time for all disjoint sets $U_1, U_2 \in \mathcal{E}_M$ and $k \in \mathbb{N}$. Given a truncation function $\tau \colon \mathbb{R}^d \to \mathbb{R}^d$, define for $A, B \in \tilde{\mathcal{P}}_M$ and $V \in \mathcal{B}_0(\mathbb{R}^d)$

$$\begin{aligned} B(A) &:= \mathfrak{B}(\mathbb{1}_A \cdot M)_\infty, \quad \mu(A, V) := \mu^{\mathbb{1}_A \cdot M}(\mathbb{R}, V), \quad \nu(A, V) := \nu^{\mathbb{1}_A \cdot M}(\mathbb{R}, V) \\ M^{\mathrm{c}}(A) &:= (\mathbb{1}_A \cdot M)^{\mathrm{c}}, \quad C^{ij}(A; B) := [(\mathbb{1}_A \cdot M^i)^{\mathrm{c}}, (\mathbb{1}_B \cdot M^j)^{\mathrm{c}}]_\infty. \end{aligned} \tag{1.4.18}$$

As in Theorem 1.3.2 (B, C, ν) can be extended to predictable strict random (bi-) measures and give rise to the following canonical decomposition of M:

$$\begin{aligned} M(A) = B(A) + M^{\mathrm{c}}(A) &+ \int_{\mathbb{R}\times E\times\mathbb{R}^d} \mathbb{1}_A(t, x)(y - \tau(y)) \, \mu(\mathrm{d}t, \mathrm{d}x, \mathrm{d}y) \\ &+ \int_{\mathbb{R}\times E\times\mathbb{R}^d} \mathbb{1}_A(t, x)\tau(y) \, (\mu - \nu)(\mathrm{d}t, \mathrm{d}x, \mathrm{d}y), \quad A \in \tilde{\mathcal{P}}_M. \end{aligned} \tag{1.4.19}$$

Moreover, there exist a positive predictable strict random measure $A(\omega, \mathrm{d}t, \mathrm{d}x)$, a $\tilde{\mathcal{P}}$-measurable $\mathbb{R}^d$-valued function $b(\omega, t, x)$ and a transition kernel $K(\omega, t, x, \mathrm{d}y)$ from $(\tilde{\Omega}, \tilde{\mathcal{P}})$ to $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$ such that for all $\omega \in \Omega$,

$$B(\omega, \mathrm{d}t, \mathrm{d}x) = b(\omega, t, x)\, A(\omega, \mathrm{d}t, \mathrm{d}x), \quad \nu(\omega, \mathrm{d}t, \mathrm{d}x, \mathrm{d}y) = K(\omega, t, x, \mathrm{d}y)\, A(\omega, \mathrm{d}t, \mathrm{d}x).$$

The multi-dimensional version of Theorem 1.4.1 reads as follows.

Theorem 1.4.8. *Let M be a d-dimensional random measure with different times of discontinuity and $H\colon \tilde{\Omega} \to \mathbb{R}^d$ be a predictable function such that there exists a strictly positive predictable process $K\colon \bar{\Omega} \to \mathbb{R}$ with $HK \in L^{1,0}(M)$. Then $H \in L^0(M)$ if and only if each of the following conditions is satisfied a.s.:*

$$\int_{\mathbb{R}\times E} \left| H(t,x)b(t,x) + \int_{\mathbb{R}^d} \left[\tau(H(t,x)y) - H(t,x)\tau(y)\right] K(t,x,\mathrm{d}y) \right| A(\mathrm{d}t, \mathrm{d}x) < \infty,$$

$$\int_{\mathbb{R}} K_t^{-2} \mathrm{d} \left(\int_{((-\infty,t]\times E;(-\infty,t]\times E)} (HK; HK)\, \mathrm{d}C \right) < \infty,$$

$$\int_{\mathbb{R}\times E} \int_{\mathbb{R}^d} \left(1 \wedge |H(t,x)y|^2\right) K(t,x,\mathrm{d}y)\, A(\mathrm{d}t, \mathrm{d}x) < \infty.$$

1.5 Ambit processes

In this section we present various applications where the integrability conditions of Theorem 1.4.1 are needed. Given a filtered probability space satisfying the usual assumptions, our examples are processes of the following form:

$$Y(t,x) := \int_{\mathbb{R}\times\mathbb{R}^d} h(t,s;x,y)\, M(\mathrm{d}s, \mathrm{d}y), \quad t \in \mathbb{R}, \quad x \in \mathbb{R}^d, \tag{1.5.1}$$

where $h\colon \mathbb{R} \times \mathbb{R} \times \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$ is a deterministic measurable function and M a random measure with different times of discontinuity such that the integral in (1.5.1) exists in the sense of (1.2.3). If the characteristics of M in the sense of Theorem 1.3.2 are known, (1.5.1) exists if and only if the conditions of Theorem 1.4.1 are satisfied for each pair $(t,x) \in \mathbb{R} \times \mathbb{R}^d$. We call processes of the form (1.5.1) *ambit processes* although the original definition in [18] requires the random measure to be a volatility modulated Lévy basis, i.e. $M = \sigma.\Lambda$ where Λ is a Lévy basis and $\sigma \in \tilde{\mathcal{P}}$. As already explained in the Introduction, this class of models is relevant in many different areas of applications. In the following subsections, we discuss two applications where interesting choices for h and M will be presented.

1.5.1 Stochastic PDEs

The connection between stochastic PDEs and ambit processes is exemplified in [18] relying on the integration theory of [119] or [136]. Let U be an open subset of $\mathbb{R}\times\mathbb{R}^d$ with boundary ∂U, P a polynomial in $1+d$ variables with constant coefficients and M a random measure with different times of discontinuity. The goal is to find a solution Z to the stochastic PDE

$$P(\partial_t,\partial_1,\dots,\partial_d)Z(t,x)=\partial_t\partial_1\dots\partial_d M(t,x),\quad (t,x)\in U, \tag{1.5.2}$$

subjected to some boundary conditions on ∂U, where $\partial_t\partial_1\dots\partial_d M(t,x)$ is the formal derivative of M, its noise. We want to apply the method of Green's function to our random setting: first, we find a solution Y to (1.5.2) with vanishing boundary conditions, then we find a solution Y' to the homogeneous version of (1.5.2) which satisfies the prescribed boundary conditions, and finally we obtain a solution Z by the sum of Y and Y'. Since the problem of finding Y' is the same as in ordinary PDE theory, we concentrate on finding Y. However, since the noise of M does not exist formally, there exists no solution Y' in the strong sense. One standard approach based on [136, Sect. 3] is to interprete (1.5.2) in weak form and to define

$$Y(t,x):=\int_U G(t,s;x,y)\,M(\mathrm{d}s,\mathrm{d}y),\quad (t,x)\in U, \tag{1.5.3}$$

as a solution, where G is the Green's function for P in the domain U. Obviously, Y is then an ambit process, where the integrand is determined by the partial differential operator and the domain, and the integrator is the driving noise of the stochastic PDE. Therefore, Theorem 1.4.1 provides necessary and sufficient conditions for the existence of Y. Let us stress again that, in contrast to [119] and [136], we need no distributional assumptions on M.

Finally, we want to come back to Remark 1.4.5 and explain why the L^2-approach is too stringent for stochastic PDEs. To this end, we consider the stochastic heat equation in $\mathbb{R}^d$:

Example 1.5.1 We consider $P(t,x)=t-\sum_{i=1}^d x_i$, $U=(0,\infty)\times\mathbb{R}^d$ and $M=\sigma.\Lambda$ where σ is a predictable function and Λ a Lévy basis with characteristics $(0,\Sigma\,\mathrm{d}t\,\mathrm{d}x,\nu(\mathrm{d}\xi)\,\mathrm{d}t\,\mathrm{d}x)$, where $\Sigma\geq 0$ and ν is a symmetric Lévy measure. [136, Sect. 3] considers a similar equation with $\nu=0$. The Green's function for P and U is the heat kernel

$$G(t,s;x,y)=\frac{\exp(-|x-y|^2/(4(t-s)))}{(4\pi(t-s))^{d/2}}\mathbb{1}_{\{0<s<t\}},\quad s,t>0,\quad x,y\in\mathbb{R}^d.$$

Since for all $(t,x) \in U$ the kernel $G(t,\cdot;x,\cdot) \in L^p(U)$ if and only if $p < 1 + 2/d$, it is square-integrable only for $d = 1$. Therefore, in the L^2-approach function-valued solutions only exist for $d = 1$. However, if $\Sigma = 0$, a sufficient condition for (1.4.3) and thus the existence of (1.5.3) is

$$\int_0^t \int_{\mathbb{R}^d} |G(t,s;x,y)\sigma(s,y)|^p \,\mathrm{d}s\,\mathrm{d}y < \infty \quad \text{a.s.}, \qquad \int_{[-1,1]} |\xi|^p \,\nu(\mathrm{d}\xi) < \infty \tag{1.5.4}$$

for all $(t,x) \in U$ and some $p \in [0,2)$. For instance, if σ is stationary in U with finite p-th moment, (1.5.4) becomes

$$\int_{[-1,1]} |\xi|^p \,\nu(\mathrm{d}\xi) < \infty \quad \text{for some } p < 1 + 2/d.$$

In particular, we see that function-valued solutions exist in arbitrary dimensions, which cannot be "detected" in the L^2-framework, even for integrators which are L^2-random measures. □

The stochastic heat equation or similar equations driven by non-Gaussian noise have already been studied in a series of papers, e.g. [3, 5, 107, 108, 125], partly also extending Walsh's approach beyond the L^2-framework. Although they do not only consider the linear case (1.5.2), there are always limitations in dimension (e.g. only $d = 1$) or noise type (e.g. only stable noise without volatility modulation). Thus, in the linear case, Theorem 1.4.1 provides a unifying extension of the corresponding results in the given references.

1.5.2 Superposition of stochastic volatility models

In this subsection we give examples of ambit processes, where the spatial component in the stochastic integral has the meaning of a parameter space. First we discuss one possibility of constructing a superposition of COGARCH processes, following Section 4.3.3. The COGARCH model of [88] itself is designed as a continuous-time version of the celebrated GARCH model and is defined as follows: Let $(L_t)_{t\in\mathbb{R}}$ be a two-sided Lévy process with Lévy measure ν_L. Given $\beta, \eta > 0$ the COGARCH model (V^φ, G^φ) with parameter $\varphi \geq 0$ is given by the equations

$$\mathrm{d}G_t^\varphi = \sqrt{V_{t-}^\varphi}\,\mathrm{d}L_t, \quad G_0^\varphi = 0, \tag{1.5.5}$$

$$\mathrm{d}V_t^\varphi = (\beta - \eta V_t^\varphi)\,\mathrm{d}t + \varphi V_{t-}^\varphi \,\mathrm{d}S_t, \quad t \in \mathbb{R}, \tag{1.5.6}$$

where $S := [L]^{\mathrm{d}}$ denotes the pure-jump part of the quadratic variation of L. By [88, Thm. 3.1], (1.5.6) has a stationary solution if and only if

$$\int_{\mathbb{R}_+} \log(1+\varphi y^2)\,\nu_L(\mathrm{d}y) < \eta. \tag{1.5.7}$$

Let us denote the collection of all $\varphi \geq 0$ satisfying (1.5.7) by Φ, which by (1.5.7) must be of the form $[0, \varphi_{\max})$ with some $0 < \varphi_{\max} < \infty$. Although the COGARCH model essentially reproduces the same stylized features as the GARCH model, there are two unsatisfactory aspects:

(1) Right from the definition, the COGARCH shows a deterministic relationship between volatility and price jumps, an effect shared by many continuous-time stochastic volatility models [82]. More precisely, we have

$$\Delta V_t^{\varphi} = \varphi V_{t-}^{\varphi} (\Delta L_t)^2 = \varphi (\Delta G_t^{\varphi})^2, \quad t \in \mathbb{R}. \tag{1.5.8}$$

A realistic stochastic volatility model should allow for different scale parameters φ.

(2) The autocovariance function of the COGARCH volatility is, when existent and $\varphi > 0$, always of exponential type: $\mathrm{Cov}[V_t^{\varphi}, V_{t+h}^{\varphi}] = C\mathrm{e}^{-ah}$ for $h \geq 0$, $t \in \mathbb{R}$ and some constants $C, a > 0$. A more flexible autocovariance structure is desirable.

In Section 4, three approaches to construct superpositions of COGARCH processes (supCOGARCH) with different values of φ are suggested in order to obtain a stochastic volatility model keeping the desirable features of the COGARCH but avoiding the two disadvantages mentioned above. One of them is the following: With β and η remaining constant, take a Lévy basis Λ on $\mathbb{R} \times \Phi$ with characteristics $(b\,\mathrm{d}t\,\pi(\mathrm{d}\varphi), \Sigma\,\mathrm{d}t\,\pi(\mathrm{d}\varphi), \nu_L(\mathrm{d}y)\,\mathrm{d}t\,\pi(\mathrm{d}\varphi))$, where $b \in \mathbb{R}$, $\Sigma \geq 0$, π is a probability measure on Φ and ν_L the Lévy measure of the Lévy process given by

$$L_t := \Lambda^L((0,t] \times \Phi), \quad t \geq 0, \quad L_t := -\Lambda^L((-t,0] \times \Phi), \quad t < 0,$$

Furthermore, define another Lévy basis by $\Lambda^S(\mathrm{d}t, \mathrm{d}\varphi) := \int_{\mathbb{R}} y^2\,\mu^{\Lambda}(\mathrm{d}t, \mathrm{d}\varphi, \mathrm{d}y)$, where μ^{Λ} is the jump measure of Λ as in Theorem 1.3.2. Next define V^{φ} for each $\varphi \in \Phi$ as the COGARCH volatility process driven by L with parameter φ. Motivated by (1.5.6), the *supCOGARCH* $\bar{V}$ is now defined by the stochastic differential equation

$$\mathrm{d}\bar{V}_t = (\beta - \eta\bar{V}_t)\,\mathrm{d}t + \int_{\Phi} \varphi V_{t-}^{\varphi}\,\Lambda(\mathrm{d}t, \mathrm{d}\varphi), \quad t \in \mathbb{R}. \tag{1.5.9}$$

As shown in Proposition 4.3.15, (1.5.9) has a unique solution given by

$$\bar{V}_t = \frac{\beta}{\eta} + \int_{-\infty}^{t} \int_{\Phi} \mathrm{e}^{-\eta(t-s)} \varphi V_{s-}^{\varphi} \, \Lambda(\mathrm{d}s, \mathrm{d}\varphi), \quad t \in \mathbb{R}, \tag{1.5.10}$$

such that $\bar{V}$ is an ambit process as in (1.5.1).

Here the integrability conditions of Section 1.4 come into play. Immediately from Theorem 1.4.1 and Remark 1.4.3 we obtain the following corollary.

Corollary 1.5.2. *The supCOGARCH $\bar{V}$ as in* (1.5.10) *exists if and only if*

$$\int_{\mathbb{R}_+} \int_{\Phi} \int_{\mathbb{R}_+} 1 \wedge \left(y^2 \varphi \mathrm{e}^{-\eta s} V_s^{\varphi} \right) \nu_L(\mathrm{d}y) \, \pi(\mathrm{d}\varphi) \, \mathrm{d}s < \infty \quad a.s. \tag{1.5.11}$$

In particular, the supCOGARCH (1.5.10) provides an example where the stochastic volatility process $\sigma(s, \varphi) := \varphi V_{s-}^{\varphi}$ is *not* independent of the underlying Lévy basis Λ. So the conditions of [119, Thm. 2.7] are not applicable. For further properties of the supCOGARCH, in particular regarding its jump behaviour, autocovariance structure etc., we refer to Section 4.3.3.

Finally, let us comment on superpositions of other stochastic volatility models.

Remark 1.5.3 The usage of Ornstein-Uhlenbeck processes in stochastic volatility modelling has become popular through the Barndorff-Nielsen-Shephard model [14]. A natural extension is given by the CARMA stochastic volatility model [134], which generates a more flexible autocovariance structure. Another generalization of the BNS model is obtained via a superposition of OU processes with different memory parameters leading to the class of supOU processes [11]. This method does not only yield a more general second-order structure but can also generate long-memory processes; cf. [11, 62]. A similar technique was used in [16, 99] to construct supCARMA processes, again leading to a possible long-range dependent process.

Note that in all these models the driving noise is assumed to have stationary independent increments, which is certainly a model restriction. Therefore, [17] suggests a volatility modulation of this noise to obtain a greater model flexibility. In this way, it is possible to generate a volatility clustering effect, similar to the behaviour of the (sup)COGARCH. Without volatility modulation, supOU or supCARMA processes are defined as stochastic integrals of deterministic kernel functions with respect to a Lévy basis, so the approach of [119] is sufficient. Theorem 1.4.1 now enables us to replace Λ by a volatility modulated Lévy basis $\sigma.\Lambda$ with a possible dependence structure between σ and Λ. □

Chapter 2:
Lévy-driven Volterra equations in space and time

2.1 Introduction

In this Chapter we investigate stochastic tempo–spatial Volterra equations of the following form:

$$Y(t,x) = Y_0(t,x) + \int_I \int_{\mathbb{R}^d} G(t,x;s,y)\sigma(Y(s,y))\,\Lambda(\mathrm{d}s,\mathrm{d}y), \quad (t,x) \in I \times \mathbb{R}^d. \tag{2.1.1}$$

Here, Y_0 is a given stochastic process, I is a real time interval, G a deterministic kernel function and σ a deterministic function. Apart from Y_0, the stochasticity of (2.1.1) comes from its integrator Λ, which is an infinitely divisible independently scattered random measure, or a *Lévy basis* for short.

While the theory of deterministic Volterra equations is very well studied by now (see, for example, the monograph [74]), the literature on Volterra equations with stochastic integrators is considerably smaller. If no space is involved, [117] proves existence and uniqueness for general semimartingale integrators under differentiability assumptions on the kernel G. In the special case of Lévy-driven stochastic delay equations, the asymptotic behaviour of solutions and the existence of stationary solutions are discussed in [120]. As soon as the kernel becomes explosive, existence and uniqueness results have been found for Brownian integrators, see [46, 47, 138]. In the tempo–spatial case, singular kernels are typically encountered in the theory of stochastic PDEs, with two main approaches having become established in this context: on the one hand, there is the functional analytic approach that treats infinite-dimensional stochastic evolution equations as ordinary SDEs with irregular

coefficients driven by Hilbert or Banach space-valued Lévy processes; see, for instance, [114] for an excellent account on this subject. On the other hand, there is the random field approach that directly considers (2.1.1) as a scalar-valued equation driven by a multi-parameter Lévy noise.

Since our treatment of (2.1.1) will be within the latter framework, we review the existing literature in this field in more detail: based on the seminal work [136], which uses equations of type (2.1.1) in order to solve certain stochastic PDEs driven by Gaussian white noise, several attempts have been made to generalize Walsh's method to other noise types. One possibility is, for instance, to consider Gaussian noise that is white in time but coloured in space, which is proposed in [48]. Leaving the Gaussian world, [3, 5] study the stochastic heat equation driven by Lévy white noise. However, since both references still employ the L^2-theory of Walsh, they are confronted with the uncomfortable fact that the stochastic heat equation will have no solutions in dimensions greater than 1, cf. [136, pp. 328ff.]. This is due to the bad integrability properties of the heat kernel that plays the role of G in (2.1.1): it is square-integrable only for $d = 1$.

Therefore, the passage from the L^2- to an L^p-framework, $p \in (0, 2]$, is inevitable. The first paper that discusses Lévy-driven stochastic PDEs in an L^p-framework with $p \in [1, 2]$ is, to our best knowledge, [125]. Under the usual Lipschitz condition on σ, existence and uniqueness for (2.1.1) are proved when G is the heat kernel and Λ a homogeneous Lévy basis that is either a martingale measure or of locally finite variation. In [107, 108] a specific equation that goes beyond the results of [125] is studied: they take the non-Lipschitz coefficient $\sigma(x) = x^\beta$ with $\beta \neq 1$ and an α-stable spectrally positive Lévy basis for Λ, where $\alpha \in (0, 1)$ and $\alpha \in (1, 2)$, respectively. Finally, [9] treats the Lipschitz case with α-stable Λ where $\alpha \neq 1$. In all articles mentioned so far, the time horizon is $I = \mathbb{R}_+$.

Let us also point out that processes of the form (2.1.1) are closely related to a class of random fields that are called *ambit processes* and have found applications in physics, finance, biology among other disciplines; see [116] for a recent survey article. While for an ambit process $\sigma(Y(s, y))$ in (2.1.1) is replaced by some given random field $\sigma(s, y)$, which means that (2.1.1) actually becomes a definition, the Volterra equation (2.1.1) has to be solved first of all. Once a solution is found, it is a special type of ambit processes. For the connection between ambit processes and stochastic PDEs, we also refer to [18].

This Chapter is organized as follows: after we have provided all necessary back-

ground information in Section 2.2, we start to discuss (2.1.1) in Section 2.3 for $I = \mathbb{R}_+$. In Theorem 2.3.1 we establish existence and uniqueness conditions for (2.1.1) in L^p-spaces for $p \in (0, 2]$ under Lipschitz conditions on σ. They generalize the results mentioned in the literature review to kernels G that need not be of convolution type or related to stochastic PDEs, as well as to Lévy bases that are combinations of martingale and finite variation parts, and whose characteristics are potentially inhomogeneous in space and time. The most stringent condition in Theorem 2.3.1 is that, loosely speaking, Λ must have a moment structure that is at least as nice as its variation structure. This, for instance, a priori excludes any stable Lévy basis. An extension to such cases is provided in Theorem 2.3.5 if Λ only has finitely many large jumps on finite time intervals. Using localization methods as in [9], we are able to reduce the situation to the framework of Theorem 2.3.1 and prove existence and uniqueness of solutions this way. Beyond that, if σ has sublinear growth, we prove that they have finite L^p-moments for some $p \in (0, 2]$.

In Section 2.4, we extend the results from Section 2.3 to the case of infinite memory, which, to our knowledge, has not been considered before in the literature. More precisely, we investigate existence and uniqueness for (2.1.1) when $I = \mathbb{R}$ (Theorem 2.4.4), which turns out to be much more involved than the case $I = [0, \infty)$. First, the method of Theorem 2.3.5 will no longer work, that is, Λ is required to have a good moment structure. Second, and more importantly, an explicit size condition on G, σ and Λ comes into play, which is already a characteristic feature of deterministic Volterra equations, see Example 2.4.1. Therefore, detailed L^p-estimates for the stochastic integral in (2.1.1) are required. Furthermore, under certain conditions on Y_0, one can improve the results by using weighted L^p-spaces. If G is a kernel of convolution form and Λ is homogeneous in space and time, the stationarity of the solution is discussed in Theorem 2.4.8. Section 2.4 is round off with some results concerning the L^p-continuity of the solution Y and its continuous dependence on Y_0; see Theorem 2.4.7.

In Section 2.5 we assume that we have already found a solution to (2.1.1) that is L^p-bounded up to time T for every $T \in \mathbb{R}_+$. We want to address the question when the solution remains L^p-bounded as $T \to \infty$. An affirmative answer is given under two types of conditions (Theorem 2.5.2): first, if G, σ and Λ are small enough, a feature that we have already encountered in Theorem 2.4.4 and that is also similar to the conditions in [120] in the context of stationary solutions to stochastic delay equations; and second, if the function σ is of sublinear growth. Both condi-

tions are intrinsic for Volterra-type equations as a deterministic example shows, see Example 2.5.1.

In Sections 2.3 to 2.5, we illustrate all our results by means of the stochastic heat equation, see Examples 2.3.4, 2.3.8, 2.4.9 and 2.5.3.

Finally, Section 2.6 contains several lemmata needed for the proof of the main theorems, which is carried out in Section 2.7.

2.2 Preliminaries

We begin with a table of frequently used notations and abbreviations:

$\mathbb{R}_+$	the set $[0,\infty)$ of *positive* real numbers (*strict positivity* excludes 0);
$\bar{\mathbb{R}}$	the extended real line $\mathbb{R}\cup\{\pm\infty\}$;
$\mathbb{N}$	the set $\{1,2,\ldots\}$ of natural numbers;
I	either $I=\mathbb{R}_+$ or $I=\mathbb{R}$;
I_T	$I\cap(-\infty,T]$ for some $T\in\mathbb{R}\cup\{\infty\}$;
p^*	$p\vee 1$ for $p\in[0,\infty)$;
$\lvert z\rvert_s^r$	$\lvert z\rvert^r\mathbb{1}_{\{\lvert z\rvert>1\}}+\lvert z\rvert^s\mathbb{1}_{\{\lvert z\rvert\le 1\}}$ for $r,s,z\in\mathbb{R}$;
$\mathbb{B}$	a stochastic basis $(\Omega,\mathcal{F},\mathbb{F}=(\mathcal{F}_t)_{t\in I},\mathbb{P})$ satisfying the usual hypotheses of right-continuity and completeness that is large enough to support all random elements of this chapter;
$\tilde\Omega$	$\tilde\Omega:=\Omega\times I\times\mathbb{R}^d$ for some $d\in\mathbb{N}\cup\{0\}$ with the convention $\mathbb{R}^0:=\{1\}$;
$\tilde{\mathcal{P}}$	depending on the context, either the *tempo–spatial predictable σ-field* $\mathcal{P}\otimes\mathcal{B}(\mathbb{R}^d)$ where $\mathcal{P}$ is the usual predictable σ-field and $\mathcal{B}(\mathbb{R}^d)$ is the Borel σ-field on $\mathbb{R}^d$, or the class of *predictable* (i.e. $\tilde{\mathcal{P}}$-measurable) mappings $\tilde\Omega\to\bar{\mathbb{R}}$;
$\tilde{\mathcal{P}}_{\mathrm{b}}$	the collection of all sets $A\in\tilde{\mathcal{P}}$ such that there exists $k\in\mathbb{N}$ with $A\subseteq\Omega\times(I\cap[-k,k])\times[-k,k]^d$;
$\mathcal{B}_{\mathrm{b}}$	the collection of all bounded Borel sets in $I\times\mathbb{R}^d$;
$[\![R,S]\!]$	$\{(\omega,t)\in\Omega\times I\colon R(\omega)\le t\le S(\omega)\}$ for two $\mathbb{F}$-stopping times R,S, analogously for the other stochastic intervals;
$\lvert\mu\rvert$	the total variation measure of a signed Borel measure μ;
$x+A$	$\{x+a\colon a\in A\}$ for $x\in\mathbb{R}^d$ and $A\subseteq\mathbb{R}^d$;
A^{C}	$\mathbb{R}^d\setminus A$ for $A\subseteq\mathbb{R}^d$;
$\operatorname{diam}(A)$	$\sup\{\lvert x-y\rvert\colon x,y\in A\}$ for $A\subseteq\mathbb{R}^d$;
$(x,y]$	$\{z\in\mathbb{R}^d\colon x_i<z_i\le y_i \text{ for all } i=1,\ldots,d\}$ for $x,y\in\mathbb{R}^d$;

L^p the usual spaces $L^p(\Omega, \mathcal{F}, \mathbb{P})$ for $p \in [0, \infty)$ endowed with the topologies induced by $\|X\|_{L^p} := \mathbb{E}[|X|^p]^{1/p^*}$ for $p \in (0, \infty)$ and $\|X\|_{L^0} := \mathbb{E}[|X| \wedge 1]$ for $p = 0$

In model (2.1.1), Λ will always be a *Lévy basis* on $I \times \mathbb{R}^d$, that is, a mapping $\Lambda \colon \tilde{\mathcal{P}}_{\mathrm{b}} \to L^0$ with the following properties:

(1) $\Lambda(\emptyset) = 0$ a.s.

(2) For every sequence $(A_i)_{i \in \mathbb{N}}$ of pairwise disjoint sets in $\tilde{\mathcal{P}}_{\mathrm{b}}$ with $\bigcup_{i=1}^{\infty} A_i \in \tilde{\mathcal{P}}_{\mathrm{b}}$ we have
$$\Lambda\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \Lambda(A_i) \quad \text{in } L^0.$$

(3) For all $A \in \tilde{\mathcal{P}}_{\mathrm{b}}$ with $A \subseteq \Omega \times I_t \times \mathbb{R}^d$ for some $t \in I$, the random variable $\Lambda(A)$ is $\mathcal{F}_t$-measurable.

(4) For all $A \in \tilde{\mathcal{P}}_{\mathrm{b}}$, $t \in I$ and $\Omega_0 \in \mathcal{F}_t$, we have
$$\Lambda\Big(A \cap (\Omega_0 \times (t, \infty) \times \mathbb{R}^d)\Big) = \mathbb{1}_{\Omega_0} \Lambda\Big(A \cap (\Omega \times (t, \infty) \times \mathbb{R}^d)\Big) \quad \text{a.s.}$$

(5) If $(B_i)_{i \in \mathbb{N}}$ is a sequence of pairwise disjoint sets in $\mathcal{B}_{\mathrm{b}}$, then $(\Lambda(\Omega \times B_i))_{i \in \mathbb{N}}$ is a sequence of independent random variables. Furthermore, if $B \in \mathcal{B}_{\mathrm{b}}$ satisfies $B \subseteq (t, \infty) \times \mathbb{R}^d$ for some $t \in I$, then $\Lambda(\Omega \times B)$ is independent of $\mathcal{F}_t$.

(6) For all $B \in \mathcal{B}_{\mathrm{b}}$, $\Lambda(\Omega \times B)$ has an infinitely divisible distribution.

(7) For all $t \in I$ and $k \in \mathbb{N}$ we have $\Lambda(\Omega \times \{t\} \times [-k, k]^d) = 0$ a.s.

Just as Lévy processes are semimartingales in the purely temporal case, Lévy bases are random measures, that is, stochastic integrators in space–time. In other words, it is possible to develop an Itô stochastic integration theory for Lévy bases. Let us briefly recall this; all details can be found in [27, Chap. 3] and [28]. Starting with simple integrands $H \in \mathcal{S}$, that is, $H = \sum_{i=1}^{r} a_i \mathbb{1}_{A_i}$ with $r \in \mathbb{N}$, real numbers a_i and sets $A_i \in \tilde{\mathcal{P}}_{\mathrm{b}}$, we define the stochastic integral in the canonical way:
$$\int_I \int_{\mathbb{R}^d} H(t, x)\, \Lambda(\mathrm{d}t, \mathrm{d}x) := \sum_{i=1}^{r} a_i \Lambda(A_i).$$

Given a general predictable function $H \in \tilde{\mathcal{P}}$, we introduce the *Daniell mean*

$$\|H\|_\Lambda := \sup_{S \in \mathcal{S}, |S| \leq |H|} \left\| \int_I \int_{\mathbb{R}^d} S(t,x)\, \Lambda(\mathrm{d}t, \mathrm{d}x) \right\|_{L^0},$$

and define the class of *integrable* functions $L^0(\Lambda)$ as the closure of $\mathcal{S}$ under the Daniell mean $\|\cdot\|_\Lambda$. This is to say that $H \in \tilde{\mathcal{P}}$ is integrable with respect to Λ if and only if there exists a sequence $(S_n)_{n\in\mathbb{N}}$ of elements in $\mathcal{S}$ such that $\|H - S_n\|_\Lambda \to 0$ as $n \to \infty$. Then the *stochastic integral*

$$\int_I \int_{\mathbb{R}^d} H(t,x)\, \Lambda(\mathrm{d}t, \mathrm{d}x) := \lim_{n\to\infty} \int_I \int_{\mathbb{R}^d} S_n(t,x)\, \Lambda(\mathrm{d}t, \mathrm{d}x)$$

as a limit in probability exists and does not depend on the chosen sequence $(S_n)_{n\in\mathbb{N}}$. Moreover, defining

$$H \cdot \Lambda_t := \int_{I_t} \int_{\mathbb{R}^d} H(s,y)\, \Lambda(\mathrm{d}s, \mathrm{d}y), \quad t \in I,$$

the process $H \cdot \Lambda = (H \cdot \Lambda_t)_{t\in I}$ has a modification that is a semimartingale on I. In the case $I = \mathbb{R}$, we mean by this that $X_{-\infty} := \lim_{t\downarrow -\infty} X_t$ exists as a limit in probability, and for all bijective increasing functions $\phi\colon \mathbb{R}_+ \to [-\infty, \infty)$ the process $X^\phi := (X_{\phi(t)})_{t\in\mathbb{R}_+}$ is a usual semimartingale with respect to $(\mathcal{F}_{\phi(t)})_{t\in\mathbb{R}_+}$. For later reference, we shall mention that its quadratic variation process is defined by $[X]_t := [X^\phi]_{\phi^{-1}(t)}$ for $t \in \bar{\mathbb{R}}$. Finally, given a function $H \in \tilde{\mathcal{P}}$, one can define a new random measure $H.\Lambda$ by setting

$$\begin{aligned} K \in L^0(H.\Lambda) &:\Leftrightarrow KH \in L^0(\Lambda), \\ \int_I \int_{\mathbb{R}^d} K(t,x)\, (H.\Lambda)(\mathrm{d}t, \mathrm{d}x) &:= \int_I \int_{\mathbb{R}^d} K(t,x) H(t,x)\, \Lambda(\mathrm{d}t, \mathrm{d}x). \end{aligned} \tag{2.2.1}$$

This indeed defines a random measure $H.\Lambda$ if there exists a sequence $(A_k)_{k\in\mathbb{N}} \subseteq \tilde{\mathcal{P}}$ with $A_k \uparrow \tilde{\Omega}$ such that $\mathbb{1}_{A_k} \in L^0(H.\Lambda)$ for all $k \in \mathbb{N}$.

Every Lévy basis Λ has a canonical decomposition of the following form, see Theorem 1.3.2.

$$\begin{aligned} \Lambda(\mathrm{d}t, \mathrm{d}x) = B(\mathrm{d}t, \mathrm{d}x) + \Lambda^{\mathrm{c}}(\mathrm{d}t, \mathrm{d}x) + \int_{\mathbb{R}} z \mathbb{1}_{\{|z|\leq 1\}}\, (\mu - \nu)(\mathrm{d}t, \mathrm{d}x, \mathrm{d}z) \\ + \int_{\mathbb{R}} z \mathbb{1}_{\{|z|>1\}}\, \mu(\mathrm{d}t, \mathrm{d}x, \mathrm{d}z), \end{aligned} \tag{2.2.2}$$

where the ingredients are as follows:

(1) B is a deterministic σ-finite signed Borel measure on $I \times \mathbb{R}^d$.

(2) Λ^{c}, the continuous part of Λ in the usual sense ([28, Thm. 4.13]), is a Gaussian random measure with variance measure C, which means that it is itself a Lévy basis and $\Lambda^{\mathrm{c}}(\Omega \times B)$ has a normal distribution with mean 0 and variance $C(B)$ for every $B \in \mathcal{B}_{\mathrm{b}}$.

(3) μ is a Poisson measure on $I \times \mathbb{R}^d \times \mathbb{R}$ relative to $\mathbb{F}$ with intensity measure ν, see [80, Def. II.1.20].

Moreover, we have a representation

$$\begin{aligned} B(\mathrm{d}t, \mathrm{d}x) &= b(t,x)\,\lambda(\mathrm{d}t, \mathrm{d}x), & C(\mathrm{d}t, \mathrm{d}x) &= c(t,x)\,\lambda(\mathrm{d}t, \mathrm{d}x), \\ \nu(\mathrm{d}t, \mathrm{d}x, \mathrm{d}z) &= \pi(t,x,\mathrm{d}z)\,\lambda(\mathrm{d}t, \mathrm{d}x), \end{aligned} \tag{2.2.3}$$

with measurable functions $b\colon I \times \mathbb{R}^d \to \mathbb{R}$, $c\colon I \times \mathbb{R}^d \to \mathbb{R}_+$, a transition kernel π from $I \times \mathbb{R}^d$ to $\mathbb{R}$ such that $\pi(t,x,\cdot)$ is a Lévy measure for each (t,x), and a positive σ-finite measure λ on $I \times \mathbb{R}^d$ satisfying $\lambda(\{t\} \times \mathbb{R}^d) = 0$ for all $t \in I$.

If π satisfies

$$\int_{|z|>1} |z|\,\pi(t,x,\mathrm{d}z) < \infty, \tag{2.2.4}$$

$$\text{or} \quad \int_{|z|\le 1} |z|\,\pi(t,x,\mathrm{d}z) < \infty, \quad \text{respectively,} \tag{2.2.5}$$

for all $(t,x) \in I \times \mathbb{R}^d$, then it makes sense to introduce the *mean measure* (resp. *drift measure*)

$$B_1(\mathrm{d}t, \mathrm{d}x) := b_1(t,x)\,\lambda(\mathrm{d}t, \mathrm{d}x), \quad b_1(t,x) := b(t,x) + \int_{\mathbb{R}} z \mathbb{1}_{\{|z|>1\}}\,\pi(t,x,\mathrm{d}z), \tag{2.2.6}$$

$$B_0(\mathrm{d}t, \mathrm{d}x) := b_0(t,x)\,\lambda(\mathrm{d}t, \mathrm{d}x), \quad b_0(t,x) := b(t,x) - \int_{\mathbb{R}} z \mathbb{1}_{\{|z|\le 1\}}\,\pi(t,x,\mathrm{d}z). \tag{2.2.7}$$

If in the first case we have $b_1(t,x) = 0$ for all $(t,x) \in I \times \mathbb{R}^d$, then Λ is called a *martingale Lévy basis*, which will be denoted by $\Lambda \in \mathcal{M}$; if in the second case we have $b_0(t,x) = 0$ for all $(t,x) \in I \times \mathbb{R}^d$, then Λ is called a *Lévy basis without drift*. Next, Λ is called *symmetric* if for all $(t,x) \in I \times \mathbb{R}^d$ we have $b(t,x) = 0$ and the Lévy measure $\pi(t,x,\cdot)$ is symmetric. Furthermore, Λ is called a *homogeneous* Lévy basis if λ is the Lebesgue measure on $I \times \mathbb{R}^d$ and b, c and π do not depend on $(t,x) \in I \times \mathbb{R}^d$. In this case, a function $\phi \in \tilde{\mathcal{P}}$ is *jointly stationary with* Λ if for

arbitrary $n \in \mathbb{N}$, $(h, \eta) \in \mathbb{R} \times \mathbb{R}^d$, points $(t_1, x_1), \dots, (t_n, x_n) \in I \times \mathbb{R}^d$ and pairwise disjoint sets $B_1, \dots, B_n \in \mathcal{B}_{\mathrm{b}}$, we have

$$\begin{aligned}&(\phi(t_i, x_i), \Lambda(B_i) \colon i = 1, \dots, n, t_i + h \in I)\\ &\quad \stackrel{\mathrm{d}}{=} (\phi(t_i + h, x_i + \eta), \Lambda(B_i + (h, \eta)) \colon i = 1, \dots, n, t_i + h \in I).\end{aligned}$$

Let us come back to Equation (2.1.1). We first clarify what we mean by a solution Y to (2.1.1):

Definition 2.2.1 Equation (2.1.1) is said to have a *solution* if there exists a predictable process $Y \in \tilde{\mathcal{P}}$ such that for all $(t, x) \in I \times \mathbb{R}^d$ the stochastic integral on the right-hand side of (2.1.1) is well defined and equation (2.1.1) holds a.s. We identify two solutions Y_1 and Y_2 if for all $(t, x) \in I \times \mathbb{R}^d$ we have $Y_1(t, x) = Y_2(t, x)$ a.s. □

In order to construct solutions to (2.1.1), we introduce some spaces of stochastic processes. Let $w \colon I \times \mathbb{R}^d \to \mathbb{R}$ be a *weight function*, that is, a strictly positive measurable function. We denote by $L_I^{\infty,w}$ the Banach space of all measurable functions $f \colon I \times \mathbb{R}^d \to \mathbb{R}$ satisfying

$$\|f\|_{L_I^{\infty,w}} := \sup_{(t,x) \in I \times \mathbb{R}^d} \frac{|f(t, x)|}{w(t, x)} < \infty. \tag{2.2.8}$$

Similarly, for $p \in (0, \infty)$, $B_I^{p,w}$ is the space of all $\phi \in \tilde{\mathcal{P}}$ with

$$\|\phi\|_{B_I^{p,w}} := \sup_{(t,x) \in I \times \mathbb{R}^d} \left(\frac{\mathbb{E}[|\phi(t, x)|^p]}{w(t, x)} \right)^{1/p^*} < \infty. \tag{2.2.9}$$

If $f \in L_{I_T}^{\infty,w}$ or $\phi \in B_{I_T}^{p,w}$ for all $T \in I$, then we write $f \in L_{I,\mathrm{loc}}^{\infty,w}$ or $\phi \in B_{I,\mathrm{loc}}^{p,w}$, respectively. In the special case $w \equiv 1$, we use the notations L_I^∞, $L_{I,\mathrm{loc}}^\infty$, B_I^p and $B_{I,\mathrm{loc}}^p$.

Before we proceed to the main results of this Chapter, we recall how stochastic PDEs can be treated in the framework of (2.1.1). Let $I \subset \mathbb{R}$ be an interval, U an open subset of $\mathbb{R}^d$ with boundary ∂U and P a polynomial in $1 + d$ variables. Given some deterministic coefficient σ and some Lévy basis Λ, they give rise to the following formal equation:

$$P(\partial_t, \partial_1, \dots, \partial_d) Y(t, x) = \sigma(Y(t, x)) \dot{\Lambda}(t, x), \quad (t, x) \in I \times U, \tag{2.2.10}$$

where $\dot{\Lambda} = \partial_t \partial_1 \dots \partial_d \Lambda$ is the formal derivative of Λ, its noise. Usually, (2.2.10) is subjected to some boundary conditions on ∂U. Of course, the derivative of Λ is not

defined except in trivial cases, so a strong solution to (2.2.10) will not exist. Going back to [136] is the idea of constructing a so-called *mild solution* to (2.2.10). For this method to work, one has to assume that the operator P possesses a function-valued Green's function on $I \times U$. Then a mild solution to (2.2.10) is nothing but a solution in the sense of Definition 2.2.1 to (2.1.1), where G is the Green's function and Y_0 a term that only depends on the boundary conditions posed on ∂U.

Remark 2.2.2 While the notion of a solution as in Definition 2.2.1 is very common in the theory of stochastic PDEs, it is different to the standard notion of solutions to (ordinary) SDEs: let $I = \mathbb{R}_+$ and $d = 0$, that is, space contains only one point, and consider $G(t, 1; s, 1) = g(s)\mathbb{1}_{\{s \leq t\}}$ with some smooth function g. Then Equation (2.1.1) is equivalent to the SDE

$$\mathrm{d}Y(t) = g(t)\sigma(Y(t-))\,\Lambda(\mathrm{d}t), \quad t \geq 0, \quad Y(0) = Y_0, \tag{2.2.11}$$

where Λ is a semimartingale with independent increments. Ordinary SDE theory tells us that Equation (2.2.11) has a càdlàg solution Y that is unique up to indistinguishability. In contrast, a solution in the sense of Definition 2.2.1 would be the predictable version $Y(\cdot-)$, and uniqueness is only understood up to modifications. The reason why we have chosen this slightly different notion of a solution is that we are particularly interested in the case where G in Equation (2.1.1) has singularities. In such cases, Equation (2.1.1) permits no càdlàg solutions. □

2.3 Existence and uniqueness results on $I = \mathbb{R}_+$

The goal of this section is to provide sufficient conditions under which there exists a (unique) solution to (2.1.1) on the interval $I = \mathbb{R}_+$. It is clear that everything in this section holds analogously if we replace $I = [0, \infty)$ by $I = [a, \infty)$ with some $a \in \mathbb{R}$. As mentioned in the Introduction, the forthcoming theorem generalizes the results of [125] to potentially inhomogeneous Lévy bases and kernels different from the heat kernel. It holds under the following list of assumptions:

Assumption A Let $p \in (0, 2]$ and the predictable characteristics of Λ be given by (2.2.3). We impose the following conditions:

(1) $Y_0 \in B^p_{[0,\infty),\mathrm{loc}}$.

(2) There exists $C_{\sigma,1} \in \mathbb{R}_+$ such that $|\sigma(x) - \sigma(y)| \leq C_{\sigma,1}|x - y|$ for all $x, y \in \mathbb{R}$.

(3) $G\colon (\mathbb{R}_+ \times \mathbb{R}^d)^2 \to \mathbb{R}$ is a measurable function such that $G(t, \cdot; s, \cdot) \equiv 0$ whenever $s > t$.

(4) If $p < 2$, then Λ has no Gaussian part: $c(t, x) = 0$ for $(t, x) \in \mathbb{R}_+ \times \mathbb{R}^d$. If $p = 2$, then we assume for all $T \in \mathbb{R}_+$

$$\sup_{(t,x)\in[0,T]\times\mathbb{R}^d} \int_0^t \int_{\mathbb{R}^d} |G(t, x; s, y)|^2 c(s, y)\, \lambda(\mathrm{d}s, \mathrm{d}y) < \infty. \tag{2.3.1}$$

(5) For all $T \in \mathbb{R}_+$

$$\sup_{(t,x)\in[0,T]\times\mathbb{R}^d} \int_0^t \int_{\mathbb{R}^d} \int_{\mathbb{R}} |G(t, x; s, y)z|^p\, \nu(\mathrm{d}s, \mathrm{d}y, \mathrm{d}z) < \infty. \tag{2.3.2}$$

(6) Recall the definition of b_1 and b_0 from (2.2.6) and (2.2.7). If $p \geq 1$, assume that ν satisfies (2.2.4) and that for all $T \in \mathbb{R}$

$$\sup_{(t,x)\in[0,T]\times\mathbb{R}^d} \int_0^t \int_{\mathbb{R}^d} |G(t, x; s, y)b_1(s, y)|\, \lambda(\mathrm{d}s, \mathrm{d}y) < \infty; \tag{2.3.3}$$

if $p < 1$, assume that ν satisfies (2.2.5) and that $b_0(t, x) = 0$ for all $(t, x) \in \mathbb{R}_+ \times \mathbb{R}^d$.

(7) Define for $(t, x), (s, y) \in \mathbb{R}_+ \times \mathbb{R}^d$

$$\begin{aligned} G^A(t, x; s, y) := {} & |G(t, x; s, y)|^p \left(\int_{\mathbb{R}} |z|^p\, \pi(s, y, \mathrm{d}z) + c(s, y) \right) \\ & + |G(t, x; s, y)b_1(s, y)| \mathbb{1}_{\{p\geq 1\}}, \end{aligned}$$

and assume that for every $T \in \mathbb{R}_+$ and $\epsilon > 0$ there exists $k \in \mathbb{N}$ together with a subdivision $\mathcal{T}\colon 0 = t_0 < t_1 < \ldots < t_{k+1} = T$ such that

$$\sup_{(t,x)\in[0,T]\times\mathbb{R}^d} \sup_{i=0,\ldots,k} \int_{t_i}^{t_{i+1}} \int_{\mathbb{R}^d} G^A(t, x; s, y)\, \lambda(\mathrm{d}s, \mathrm{d}y) < \epsilon. \tag{2.3.4}$$

□

Theorem 2.3.1. *Let Assumption A be valid. Then Equation* (2.1.1) *has a unique solution in* $B^p_{[0,\infty),\mathrm{loc}}$.

The conditions of Assumption A simplify a lot if G and Λ are *quasi-stationary*, that is,

$$\begin{aligned} &|G(t,x;s,y)| \leq g(t-s,x-y), \quad \lambda(\mathrm{d}t,\mathrm{d}x) = \mathrm{d}(t,x), \quad b,c \in L^{\infty}_{[0,\infty),\mathrm{loc}}, \\ &\pi(t,x,\mathrm{d}z) \leq \pi_0(\mathrm{d}z), \end{aligned} \tag{2.3.5}$$

where $g\colon \mathbb{R}_+ \times \mathbb{R}^d \to \mathbb{R}$ is a positive measurable function.

Corollary 2.3.2. *Suppose that* (2.3.5) *holds and that Assumption A(1), (2) and (3) are given. Furthermore, assume that we have for some* $p \in (0,2]$

$$b_0 \equiv 0 \quad \textit{if} \quad p<1, \quad c \equiv 0 \quad \textit{if} \quad p<2, \quad \int_{\mathbb{R}} |z|^p \, \pi_0(\mathrm{d}z) < \infty, \tag{2.3.6}$$

and for all $T \in \mathbb{R}_+$

$$\int_0^T \int_{\mathbb{R}^d} g^p(t,x) + g(t,x) \mathbb{1}_{\{p \geq 1, \Lambda \notin \mathcal{M}\}} \, \mathrm{d}(t,x) < \infty. \tag{2.3.7}$$

Then all conditions of Assumption A are satisfied and Theorem 2.3.1 holds.

Remark 2.3.3 (1) Assumption A and Theorem 2.3.1 are special cases of Assumption C and Theorem 2.4.4, respectively, which we will discuss in Section 2.4. In fact, Theorem 2.3.1 follows if we take $I = [0,\infty)$ and $w \equiv 1$ in Theorem 2.4.4.

(2) Conditions (4), (5) and (6) in Assumption A are conditions on the joint size of G and the three characteristics of Λ, respectively. Although they are valid for many interesting examples, especially condition (5) might be too restrictive: it is violated as soon as the moment structure of Λ is worse than its variation structure, which, for instance, occurs if Λ is an α-stable Lévy basis with $\alpha \in (0,2)$; see also the last condition in (2.3.6). Theorem 2.3.5 below provides, under some additional hypotheses, an extension of Theorem 2.3.1 that includes such cases.

(3) The following observation follows from Corollary 2.3.2: in the quasi-stationary case (2.3.5), condition (7) in Assumption A is already implied by conditions (4), (5) and (6). In other words, condition (7) is a smallness assumption on the non-stationary part of G and the characteristics of Λ.

(4) As we shall see in the more general Theorem 2.4.4 in Section 2.4, it actually suffices that the left-hand side of (2.3.4) can be made smaller than some fixed

constant that does not depend on $\mathcal{T}$. Due to the previous remark, however, this fact is not that important in the case $I = [0, \infty)$ (in the case $I = \mathbb{R}$, it is!). □

Next, we apply Theorem 2.3.1 and its corollary to the stochastic heat equation. In fact, this equation will serve as our toy example and will be revisited after each main theorem: see the Examples 2.3.8, 2.4.9 and 2.5.3.

Example 2.3.4 We consider the stochastic heat equation on $\mathbb{R}_+ \times \mathbb{R}^d$, that is, (2.2.10) with P given by $P(t,x) = t - \sum_{i=1}^d x_i + a$, $a \in \mathbb{R}$, and some Lipschitz coefficient σ. The Green's function is the heat kernel

$$G_a(t,x;s,y) = g_a(t-s, x-y) = \frac{\exp\left(-\frac{|x-y|^2}{4(t-s)} - a(t-s)\right)}{(4\pi(t-s))^{d/2}} \mathbb{1}_{\{s<t\}}. \tag{2.3.8}$$

We pose an initial condition at time $t = 0$, that is, we require $Y(0,x) = y_0(x)$, where $y_0 \colon \mathbb{R}^d \to \mathbb{R}$ is some bounded continuous and, for simplicity, deterministic function. Then the correct term for Y_0 in (2.1.1) is

$$Y_0(t,x) := \int_{\mathbb{R}^d} g_a(t, x-y) y_0(y) \, \mathrm{d}y, \quad (t,x) \in \mathbb{R}_+ \times \mathbb{R}^d. \tag{2.3.9}$$

The stochastic heat equation on $I = \mathbb{R}_+$ is then given by

$$Y(t,x) = Y_0(t,x) + \int_0^t \int_{\mathbb{R}^d} g_a(t-s, x-y) \sigma(Y(s,y)) \, \Lambda(\mathrm{d}s, \mathrm{d}y). \tag{2.3.10}$$

for $(t,x) \in \mathbb{R}_+ \times \mathbb{R}^d$. Let us determine sufficient conditions for existence and uniqueness of solutions to (2.3.10): assuming that the characteristics of Λ satisfy (2.3.5), we have to check the conditions of Corollary 2.3.2: (1) and (2) of Assumption A are clear. Since

$$\int_0^T \int_{\mathbb{R}^d} g_a^p(s,y) \, \mathrm{d}(s,y) < \infty \text{ for all } T \in \mathbb{R}_+ \iff p < 1 + 2/d, \tag{2.3.11}$$

we obtain existence and uniqueness for the stochastic heat equation (2.3.10) on $I = \mathbb{R}_+$ if (2.3.6) holds with some $0 < p < 1 + 2/d$. In particular, this excludes the choice $p = 2$ and therefore the possibility of taking a non-zero Gaussian part whenever $d \geq 2$. □

As pointed out in Remark 2.3.3(2), Theorem 2.3.1 excludes any Lévy basis that has the property that for every $p \in (0, 2]$

$$\lambda\left(\left\{(t,x) \in \mathbb{R}_+ \times \mathbb{R}^d \colon \int_{\mathbb{R}} |z|^p \, \pi(t,x,\mathrm{d}z) = \infty\right\}\right) > 0. \tag{2.3.12}$$

We now discuss a possibility to circumvent this.

Assumption B Consider the following hypotheses:

(1) Assumption A(2) and (3) are valid.

(2) There exists some $q \in (0, 2]$ such that for all $n \in \mathbb{N}$ conditions (4)–(7) of Assumption A are valid when p is replaced by q and ν is replaced by

$$\nu^n(\mathrm{d}t, \mathrm{d}x, \mathrm{d}z) := \mathbb{1}_{\{|z| \le n\}}\, \nu(\mathrm{d}t, \mathrm{d}x, \mathrm{d}z).$$

Of course, b_1 is changed accordingly.

(3) For all $T \in \mathbb{R}_+$ we have $\nu\Big([0, T] \times \mathbb{R}^d \times [-1, 1]^{\mathrm{C}}\Big) < \infty$.

(4) $Y_0 \in \tilde{\mathcal{P}}$ and there are stopping times $(T_n)_{n\in\mathbb{N}}$ with $T_n \uparrow \infty$ a.s. and $Y_0 \mathbb{1}_{[\![0,T_n]\!]}$ belongs to $B^q_{[0,\infty),\mathrm{loc}}$ for all $n \in \mathbb{N}$.

(5) There exist $\gamma \in (0, 1)$ and $C_{\sigma,2} \in \mathbb{R}_+$ such that $|\sigma(x)| \le |\sigma(0)| + C_{\sigma,2}|x|^\gamma$ for all $x \in \mathbb{R}$.

(6) There exists $p \in (0, 2)$ satisfying $p < q$ and $q\gamma \le p$ such that $Y_0 \in B^p_{[0,\infty),\mathrm{loc}}$.

(7) For all $T \in \mathbb{R}_+$

$$\sup_{(t,x)\in[0,T]\times\mathbb{R}^d} \int_0^t \int_{\mathbb{R}^d} \int_{\mathbb{R}} |G(t, x; s, y)z|^p_q\, \nu(\mathrm{d}s, \mathrm{d}y, \mathrm{d}z) < \infty.$$

(8) If $p \ge 1$, (2.3.3) holds.

(9) If $p < 1$, there exist exponents $\alpha \in (-\infty, 2], \beta \in [0, \infty)$ with the following properties:

(9a) For all $(t, x) \in \mathbb{R}_+ \times \mathbb{R}^d$, $A \in [1, \infty)$ and $a \in (0, 1]$ we have

$$\left|b(t, x) - \int_{\mathbb{R}} z\mathbb{1}_{\{|z|\in(a,1]\}}\, \pi(t, x, \mathrm{d}z)\right| \le F_0(t, x)a^{1-\alpha}, \tag{2.3.13}$$

$$\left|b(t, x) + \int_{\mathbb{R}} z\mathbb{1}_{\{|z|\in(1,A]\}}\, \pi(t, x, \mathrm{d}z)\right| \le F_1(t, x)A^{1-\beta} \tag{2.3.14}$$

for some positive measurable functions $F_0, F_1 \colon \mathbb{R}_+ \times \mathbb{R}^d \to \mathbb{R}$.

(9b) For all $T \in \mathbb{R}_+$ we have

$$\sup_{(t,x)\in[0,T]\times\mathbb{R}^d} \int_0^t \int_{\mathbb{R}^d} (F_0(s,y) \vee F_1(s,y)) |G(t,x;s,y)|_\beta^\alpha \, \lambda(\mathrm{d}s, \mathrm{d}y) < \infty. \quad (2.3.15)$$

(9c) $(\alpha \vee \beta)\gamma \leq p$.

(10) The partition property (2.3.4) holds with G^B instead of G^A, where for $(t,x),(s,y) \in \mathbb{R}_+ \times \mathbb{R}^d$

$$G^B(t,x;s,y) := |G(t,x;s,y)|^2 c(s,y) + \int_{\mathbb{R}} |G(t,x;s,y)z|_q^p \, \pi(s,y,\mathrm{d}z) + \begin{cases} |G(t,x;s,y)b_1(s,y)|, & p \geq 1, \\ (F_0(s,y) \vee F_1(s,y)) |G(t,x;s,y)|_\beta^\alpha, & p < 1 \end{cases}, \quad (2.3.16)$$

□

Theorem 2.3.5. *(1) Suppose that conditions (1)–(4) of Assumption B are true. Then there exists a unique solution to Equation* (2.1.1) *among those $Y \in \tilde{\mathcal{P}}$ for which there exist stopping times $(T_n)_{n\in\mathbb{N}}$ with $T_n \uparrow \infty$ a.s. such that $Y \mathbb{1}_{[\![0,T_n]\!]}$ belongs to $B^q_{[0,\infty),\mathrm{loc}}$ for all $n \in \mathbb{N}$.*

(2) If in addition also conditions (5)–(10) of Assumption B are valid, then the solution Y from part (1) belongs to $B^p_{[0,\infty),\mathrm{loc}}$.

Remark 2.3.6 (1) Part (1) of this theorem relies on some stopping time techniques that have already been used in [9] to construct solutions to (2.1.1) driven by α-stable noise with $\alpha \neq 1$. Theorem 2.3.5 extends this result to more general Lévy bases and, more importantly, provides in part (2) conditions under which this solution belongs to the space $B^p_{[0,\infty),\mathrm{loc}}$.

(2) The smaller the growth index γ of σ is, the smaller can p be chosen and therefore, the weaker the conditions (6)–(9) of Assumption B are. For α-stable Lévy bases with $\alpha \in (0,2)$, any $\gamma \in (0,1)$ and $p < q$ will suffice.

(3) If $p < 1$, condition (9) of Assumption B looks quite technical but is actually only a very mild assumption. In the next Corollary 2.3.7 where we treat the quasi-stationary case, it is already implied by condition (6) below.

(4) Remark 2.3.3(3) holds analogously: see the next corollary.

(5) For the second condition of Assumption B, if $p \geq 1$, one has to check Assumption A(6) for different replacements of b_1 as n varies, which are usually non-zero even when $\Lambda \in \mathcal{M}$.

(6) The most stringent condition in Assumption B is (3): it requires the intensity of large jumps of Λ to decay quickly enough in space. For example, it is typically *not* enough to have $\pi(t, x, \mathrm{d}z) = \pi_0(\mathrm{d}z)$. See Corollary 2.3.7 and Example 2.3.8 for more details. □

Again we reformulate Assumption B in the quasi-stationary case:

Corollary 2.3.7. *Assume that G and Λ satisfy* (2.3.5)*, but with the stronger conditions*

$$\begin{aligned} &\pi(t, x, \mathrm{d}z) \leq \pi_1(t, x)\, \pi_0(\mathrm{d}z), \quad (t, x) \in [0, T] \times \mathbb{R}^d, \quad \pi_1 \in L^\infty_{[0,\infty),\mathrm{loc}}, \\ &\int_0^T \int_{\mathbb{R}^d} \pi_1(t, x)\, \mathrm{d}(t, x) < \infty, \quad t \in [0, T], \end{aligned} \tag{2.3.17}$$

for all $T \in \mathbb{R}_+$. Then part (1) of Theorem 2.3.5 holds if:

(1) Assumption B(1) and (4) are valid.

(2) For some $q \in (0, 2]$ conditions (2.3.6) *and* (2.3.7) *hold with p replaced by q and π_0 replaced by $\mathbb{1}_{\{|z| \leq 1\}}\, \pi_0(\mathrm{d}z)$.*

(3) If $q \geq 1$, either $\int_0^T \int_{\mathbb{R}^d} g(t, x)\, \mathrm{d}(t, x) < \infty$ for all $T \in \mathbb{R}_+$, or Λ is symmetric.

Part (2) of the same theorem holds if additionally:

(4) σ satisfies the growth condition of Assumption B(5) with $\gamma \in (0, 1)$.

(5) There exists $p \in (0, 2)$ with $p < q$ and $q\gamma \leq p$ such that $Y_0 \in B^p_{[0,\infty),\mathrm{loc}}$.

(6) $\int_{\mathbb{R}} |z|_q^p\, \pi_0(\mathrm{d}z) < \infty$ and $\int_0^T \int_{\mathbb{R}^d} |g(t, x)|_p^q\, \mathrm{d}(t, x) < \infty$ for all $T \in \mathbb{R}_+$.

For illustration purposes we go through the conditions of Theorem 2.3.5 and Corollary 2.3.7 for the stochastic heat equation.

Example 2.3.8 (Continuation of Example 2.3.4) Our aim is to extend the findings of Example 2.3.4 when Λ has bad moment properties in the sense of (2.3.12). For simplicity we assume that the characteristics of Λ are within the setting of Corollary 2.3.7, that is, they satisfy (2.3.5) and (2.3.17). As before, σ is a Lipschitz continuous function and the equation of interest is (2.3.10) with Y_0 given by (2.3.9). In view of (2.3.11), it is immediate to see that Corollary 2.3.7 yields the following conditions for part (1) of Theorem 2.3.5 to hold: there exists $q \in (0, 1 + 2/d)$ such that

$$\int_{[-1,1]} |z|^q \, \pi_0(\mathrm{d}z) < \infty, \quad c \equiv 0 \text{ if } d \geq 2, \quad b_0 \equiv 0 \text{ if } q < 1. \tag{2.3.18}$$

Furthermore, if σ has growth of order $\gamma \in (0, 1)$ and

$$\int_{|z|>1} |z|^p \, \pi_0(\mathrm{d}z) < \infty \text{ for some } p < 1 + 2/d \text{ with } p < q \text{ and } q\gamma \leq p, \tag{2.3.19}$$

then the solution Y belongs to $B^p_{[0,\infty),\mathrm{loc}}$. Indeed, this claim follows from Corollary 2.3.7 and the fact that for all $p, q \in (0, \infty)$ we have

$$\int_0^T \int_{\mathbb{R}^d} |g_a(t, x)|_p^q \, \mathrm{d}(t, x) < \infty \tag{2.3.20}$$

for all $T \in \mathbb{R}_+$ if and only if $q \in (0, 1 + 2/d)$ (p does not matter). From (2.3.19) we also see the following: the smaller the growth order γ of σ is, the fewer moments π_0 is required to have.

At last, we give some further explanation for the integrability condition on π_1 given in (2.3.17). We assume that $\pi(t, x, \mathrm{d}z) = \pi_1(t, x)\pi_0(\mathrm{d}z)$ with a Lévy measure π_0 of unbounded support. Then it is obvious to see that we cannot take $\pi_1 \equiv 1$, that is, a homogeneous noise Λ, but have to choose π_1 with sufficient decay in space. For instance, if there exists some exponent $r \in \mathbb{R}$ such that for all $T \in \mathbb{R}_+$ we have $\pi_1(t, x) \leq C_T|x|^{-r}$ for all $(t, x) \in [0, T] \times \mathbb{R}^d$ and some constant $C_T \in \mathbb{R}_+$, then we need for (2.3.17) that $r > d$, a condition that is stronger in higher dimensions. Finally, (2.3.17) is always met if π_1 is bounded and vanishes outside a compact in $\mathbb{R}^d$, which corresponds to a noise that only acts locally. In particular, this assumption is very natural if we consider the stochastic heat equation on bounded domains as, for instance, in [5, 9, 136]. □

Remark 2.3.9 Theorem 2.3.1 and 2.3.5 can actually be extended to even more general random measures than Lévy bases. Let us consider a random measure M

on $\mathbb{R}_+ \times \mathbb{R}^d$ that is defined by

$$M(\mathrm{d}t, \mathrm{d}x) = b(t,x)\,\mathrm{d}(t,x) + \rho(t,x)\,W(\mathrm{d}t, \mathrm{d}x) + \int_E \underline{\delta}(t,x,z)\,(\mathfrak{p}-\mathfrak{q})(\mathrm{d}t, \mathrm{d}x, \mathrm{d}z) + \int_E \overline{\delta}(t,x,z)\,\mathfrak{p}(\mathrm{d}t, \mathrm{d}x, \mathrm{d}z), \tag{2.3.21}$$

where $(E, \mathcal{E})$ is an arbitrary Polish space equipped with its Borel σ-field, $b, \rho \in \tilde{\mathcal{P}}$, $\delta = \underline{\delta} + \overline{\delta} = \delta \mathbb{1}_{\{|\delta| \le 1\}} + \delta \mathbb{1}_{\{|\delta| > 1\}}$ is an $\tilde{\mathcal{P}} \otimes \mathcal{E}$-measurable function, W is a Gaussian random measure with the Lebesgue measure on $I \times \mathbb{R}^d$ as variance measure, $\mathfrak{p}$ is a homogeneous Poisson random measure on $I \times \mathbb{R}^d \times E$ relative to the filtration $\mathbb{F}$ with intensity measure $\mathfrak{q}(\mathrm{d}t, \mathrm{d}x, \mathrm{d}z) = \mathrm{d}t\,\mathrm{d}x\,\lambda(\mathrm{d}z)$ where λ is a σ-finite infinite atomless measure on $(E, \mathcal{E})$. Moreover, all ingredients are such that for all $k \in \mathbb{N}$ $M(\Omega \times (I \cap (-k,k]) \times (-k,k]^d)$ is well defined. Such a measure M can be viewed as the space–time analogue of Itô semimartingales. We impose the following conditions on the coefficients (these are classical in the semimartingale setting, cf. [2, Chap. 6]): there exist positive constants $(\beta_N)_{N\in\mathbb{N}}$, a sequence of stopping times $(\tau_N)_{N\in\mathbb{N}}$ increasing to infinity a.s., and deterministic positive measurable functions $j_N(z)$ such that for all $(\omega, t, x) \in \tilde{\mathcal{P}}$ with $t \le \tau_N(\omega)$ we have

(1) $|b(\omega,t,x)|, |c(\omega,t,x)| \le \beta_N$,

(2) $|\delta(\omega,t,x,z)|^p \le j_N(z)$ and $\int_E j_N(z)\,\lambda(\mathrm{d}z) < \infty$.

Then with obvious changes to Assumptions A and B, respectively, Theorems 2.3.1 and 2.3.5 also apply to Equation (2.1.1) when driven by the random measure M as given in (2.3.21). □

2.4 Existence and uniqueness results on $I = \mathbb{R}$

While Section 2.3 deals with Equation (2.1.1) on $I = [0, \infty)$, this section investigates the case $I = \mathbb{R}$. In particular, we obtain conditions for Equation (2.1.1) to possess a stationary solution. In order to demonstrate the difference between the two cases $I = [0, \infty)$ and $I = \mathbb{R}$, we analyze the following deterministic example.

Example 2.4.1 Let $\lambda \in \mathbb{R}$ and consider the following equation:

$$v(t) = 1 + \int_{-\infty}^{t} \mathrm{e}^{-\lambda(t-s)} v(s)\,\mathrm{d}s, \quad t \in \mathbb{R}. \tag{2.4.1}$$

By standard computation one can show the following: if $\lambda \leq 0$, Equation (2.4.1) has no solution; if $\lambda > 0$ and $\lambda \neq 1$, then the solutions to (2.4.1) are

$$v(t) = c\mathrm{e}^{(1-\lambda)t} + \frac{\lambda}{\lambda - 1}, \quad c \in \mathbb{R};$$

if $\lambda = 1$, the solutions are

$$v(t) = t + c, \quad c \in \mathbb{R}.$$

We draw some important conclusions, also regarding possibilities and limitations for Equation (2.1.1) with $I = \mathbb{R}$:

(1) The reason why (2.4.1) possesses no solution for $\lambda \leq 0$ is simply the non-integrability of the kernel:

$$\int_{-\infty}^{t} \mathrm{e}^{-\lambda(t-s)} \, \mathrm{d}s = \int_{0}^{\infty} \mathrm{e}^{-\lambda s} \, \mathrm{d}s = \infty. \tag{2.4.2}$$

(2) If Equation (2.4.1) has a solution, it has uncountably many. If $\lambda \in (1, \infty)$, only one solution is in $L^{\infty}_{\mathbb{R},\mathrm{loc}}$, namely if $c = 0$. The reason for this is that the integral of the kernel given in (2.4.2) is smaller than 1. In this case the uniqueness of solutions in $L^{\infty}_{\mathbb{R},\mathrm{loc}}$ follows from Lemma 2.6.4(2). Thus, in the stochastic case of (2.1.1), we can expect existence and uniqueness of solutions in $B^{p}_{\mathbb{R},\mathrm{loc}}$ only if the quantities (2.3.1), (2.3.2) and (2.3.3) are *small enough* (not only finite) in a sense to be made precise.

(3) In contrast to the case $\lambda \in (1, \infty)$, we have for $\lambda \in (0, 1)$ that all solutions belong to $L^{\infty}_{\mathbb{R},\mathrm{loc}}$ and for $\lambda = 1$ that no solution belongs to $L^{\infty}_{\mathbb{R},\mathrm{loc}}$. Furthermore, in these cases, all solutions start with strictly negative values at $-\infty$. This is somewhat surprising given the fact that all ingredients of (2.4.1) (the exponential kernel, the constant driving force and the Lebesgue measure as integrator) are positive. This phenomenon is typical when the integral of the kernel in (2.4.2) becomes greater or equal to one: the kernel is too large to allow for a positive solution. Finally, none of the solutions can be found via a Picard iteration scheme (since the Picard iterates are always positive when the input factors are). Thus, if the kernel in (2.1.1) is too large in a certain sense, we will not be able to construct a solution in general.

(4) Under certain circumstances, however, one can make the kernel size smaller (which then implies the existence and uniqueness of solutions) by considering

Volterra equations in weighted spaces. For instance, consider the following modification of Equation (2.4.1):

$$v(t) = \mathrm{e}^{\alpha t} + \int_{-\infty}^{t} \mathrm{e}^{-\lambda(t-s)} v(s)\,\mathrm{d}s, \quad t \in \mathbb{R}, \tag{2.4.3}$$

with $\alpha, \lambda \in \mathbb{R}$ satisfying $\lambda > 0$ and $\alpha + \lambda > 1$. The family of solutions in this case is

$$v(t) = \frac{\alpha + \lambda}{\alpha + \lambda - 1} \mathrm{e}^{\alpha t} + c\mathrm{e}^{(1-\lambda)t}, \quad t \in \mathbb{R}, \quad c \in \mathbb{R}. \tag{2.4.4}$$

First note that positive solutions do exist, namely, when $c \geq 0$. Furthermore, with $w(t) := \mathrm{e}^{\alpha t}$, we have

$$\int_{-\infty}^{t} w^{-1}(t)\mathrm{e}^{-\lambda(t-s)} w(s)\,\mathrm{d}s = \int_{-\infty}^{t} \mathrm{e}^{-(\alpha+\lambda)(t-s)}\,\mathrm{d}s = (\alpha + \lambda)^{-1} < 1.$$

That is, by Lemma 2.6.4(2), there exists a unique solution to (2.4.3) in $L^{\infty,w}_{\mathbb{R},\mathrm{loc}}$, which corresponds to the case $c = 0$ in (2.4.4). Roughly speaking, this device was possible because the force function $\mathrm{e}^{\alpha t}$ is small enough at $-\infty$ (the constant function in (2.4.1) was obviously *not* small enough). This motivates us to work in the weighted spaces $B^{p,w}_{\mathbb{R},\mathrm{loc}}$ for Equation (2.1.1) on $I = \mathbb{R}$. □

We are about to formulate a set of conditions that generalizes those of Assumption A and leads to the existence and uniqueness of solutions for Equation (2.1.1) on arbitrary intervals, in particular on $I = \mathbb{R}$. In order to do so, we need the following definition.

Definition 2.4.2 Let $p \in (0, \infty)$.

(1) For $p \in (0, 1)$ we set $C_p^{\mathrm{BDG}} := 1$.

(2) For $p \in [1, \infty)$ we denote by C_p^{BDG} the smallest positive number such that for all local martingales $(M_t)_{t\in\mathbb{R}_+}$ with respect to $\mathbb{F}$ we have

$$\sup_{t\geq 0} \|M_t\|_{L^p} \leq C_p^{\mathrm{BDG}} \|[M]_\infty^{1/2}\|_{L^p}. \tag{2.4.5}$$

□

Remark 2.4.3 We make some comments on Definition 2.4.2:

(1) The Burkholder-Davis-Gundy inequality ensures the finiteness of C_p^{BDG} for $p \in [1, \infty)$. Of course, inequality (2.4.5) becomes false in general for $p < 1$; the definition above for $p \in (0, 1)$ is merely for notational convenience. Moreover, the inequality for $p \in [1, \infty)$ is usually stated with the supremum inside the L^p-norm on the left-hand side of (2.4.5). However, this may enlarge the optimal constant C_p^{BDG}.

(2) The choice $I = \mathbb{R}_+$ is unimportant: a straightforward time change argument shows that C_p^{BDG} remains optimal for any other non-trivial interval $I \subseteq \mathbb{R}$.

(3) For $p \in [1, \infty)$, the actual value of C_p^{BDG} is not known in general. We are only interested in the case $p \in [1, 2]$, for which the following results are available: $C_p^{\mathrm{BDG}} \leq \sqrt{8p}$ for $p \in (1, 2)$, $C_2^{\mathrm{BDG}} = 1$ (cf. [27, Eq. (4.2.3)]) and $C_1^{\mathrm{BDG}} = 2$ (cf. [112, Thm. 8.7]). □

Assumption C Let $0 < p \leq 2$, $I \subseteq \mathbb{R}$ be an interval and $w \colon I \times \mathbb{R}^d \to \mathbb{R}$ be a weight function. We impose the following conditions:

(1) $Y_0 \in B^{p,w}_{I,\mathrm{loc}}$.

(2) There exists $C_{\sigma,1} \in \mathbb{R}_+$ such that $|\sigma(x) - \sigma(y)| \leq C_{\sigma,1}|x - y|$ for all $x, y \in \mathbb{R}$.

(3) $G \colon (I \times \mathbb{R}^d)^2 \to \mathbb{R}$ is a measurable function such that $G(t, \cdot; s, \cdot) \equiv 0$ whenever $s > t$.

(4) If $p < 2$, then Λ has no Gaussian part: $c(t, x) = 0$ for all $(t, x) \in I \times \mathbb{R}^d$. If $p = 2$, then we assume for all $T \in I$ the finiteness of

$$\sup_{(t,x) \in I_T \times \mathbb{R}^d} \int_I \int_{\mathbb{R}^d} w^{-1}(t, x) |G(t, x; s, y)|^2 c(s, y) (w(s, y) \vee \sigma(0)) \, \lambda(\mathrm{d}s, \mathrm{d}y). \quad (2.4.6)$$

(5) For all $T \in I$ the following is finite:

$$\sup_{(t,x) \in I_T \times \mathbb{R}^d} \int_I \int_{\mathbb{R}^d} \int_{\mathbb{R}} w^{-1}(t, x) |G(t, x; s, y) z|^p (w(s, y) \vee \sigma(0)) \, \nu(\mathrm{d}s, \mathrm{d}y, \mathrm{d}z). \quad (2.4.7)$$

(6) Recall the definition of b_1 and b_0 from (2.2.6) and (2.2.7). If $p \geq 1$, assume that ν satisfies (2.2.4) and that for all $T \in I$

$$\sup_{(t,x) \in I_T \times \mathbb{R}^d} \int_I \int_{\mathbb{R}^d} w^{-1}(t, x) |G(t, x; s, y) b_1(s, y)| (w(s, y) \vee \sigma(0)) \, \lambda(\mathrm{d}s, \mathrm{d}y) \quad (2.4.8)$$

is finite; if $p < 1$, assume that ν satisfies (2.2.5) and that $b_0(t,x) = 0$ for all $(t,x) \in I \times \mathbb{R}^d$.

(7) If $p \geq 1$ and $\Lambda \notin \mathcal{M}$, assume that (6) also holds with w replaced by the constant function 1.

(8) Define for $(t,x),(s,y) \in I \times \mathbb{R}^d$

$$\begin{aligned}
\bar{G}^{C,1}(t,x;s,y) &:= (C_{\sigma,1} C_p^{\mathrm{BDG}})^p |G(t,x;s,y)|^p \left(\int_{\mathbb{R}} |z|^p \, \pi(s,y,\mathrm{d}z) + c(s,y) \right), \\
\bar{G}^{C,2}(t,x;s,y) &:= C_{\sigma,1}^p \left(\int_I \int_{\mathbb{R}^d} |G(t,x;s,y) b_1(s,y)| \, \lambda(\mathrm{d}s,\mathrm{d}y) \right)^{p-1} \\
&\qquad \times |G(t,x;s,y) b_1(s,y)| \mathbb{1}_{\{p \geq 1\}}, \\
G^{C,1}(t,x;s,y) &:= w^{-1}(t,x) \bar{G}^{C,1}(t,x;s,y) w(s,y), \\
G^{C,2}(t,x;s,y) &:= w^{-1}(t,x) \bar{G}^{C,2}(t,x;s,y) w(s,y),
\end{aligned} \tag{2.4.9}$$

and assume that for every $T \in I$ there exists $k \in \mathbb{N}$ together with a subdivision $\mathcal{T}\colon \inf I = t_0 < t_1 < \ldots < t_{k+1} = T$ such that

$$\sup_{(t,x) \in I_T \times \mathbb{R}^d} \sup_{i=0,\ldots,k} \sum_{l=1}^{2} \left(\int_{t_i}^{t_{i+1}} \int_{\mathbb{R}^d} G^{C,l}(t,x;s,y) \, \lambda(\mathrm{d}s,\mathrm{d}y) \right)^{1/p^*} < 1. \tag{2.4.10}$$

□

Theorem 2.4.4. *Under Assumption C there exists a unique solution to Equation* (2.1.1) *in* $B^{p,w}_{I,\mathrm{loc}}$.

In the quasi-stationary case, Assumption C simplifies a lot:

Corollary 2.4.5. *Let* $I = \mathbb{R}$, $w \equiv 1$ *and Assumption C(1), (2) and (3) be valid. We assume that* G *and* Λ *satisfy*

$$\begin{aligned}
&|G(t,x;s,y)| \leq g(t-s, x-y), \quad \lambda(\mathrm{d}t,\mathrm{d}x) = \mathrm{d}(t,x), \quad b, c \in L^\infty_{\mathbb{R}}, \\
&\pi(t,x,\mathrm{d}z) \leq \pi_0(\mathrm{d}z)
\end{aligned} \tag{2.4.11}$$

for all $(t,x) \in \mathbb{R} \times \mathbb{R}^d$ *and some positive measurable* $g\colon \mathbb{R}_+ \times \mathbb{R}^d \to \mathbb{R}$. *Furthermore, we suppose that for some* $p \in (0,2]$ *we have*

$$b_0 \equiv 0 \text{ if } p < 1, \quad c \equiv 0 \text{ if } p < 2, \quad \zeta_p := \int_{\mathbb{R}} |z|^p \, \pi_0(\mathrm{d}z) < \infty, \tag{2.4.12}$$

and that the following size condition is fulfilled: if $p \in (0,1)$*, then*

$$C_{\sigma,1}^p \zeta_p \int_0^\infty \int_{\mathbb{R}^d} g^p(t,x)\,\mathrm{d}(t,x) < 1, \tag{2.4.13}$$

and if $p \in [1,2]$*, then*

$$C_{\sigma,1}\left[C_p^{\mathrm{BDG}}\left((\zeta_p + \|c\|_{L^\infty_{\mathbb{R}}})\int_0^\infty \int_{\mathbb{R}^d} g^p(t,x)\,\mathrm{d}(t,x)\right)^{1/p} + \|b_1\|_{L^\infty_{\mathbb{R}}}\int_0^\infty \int_{\mathbb{R}^d} g(t,x)\,\mathrm{d}(t,x)\right] < 1. \tag{2.4.14}$$

Then all conditions of Assumption C are satisfied and Theorem 2.4.4 holds.

We write down some important observations:

Remark 2.4.6 (1) There is a fundamental difference between condition (7) of Assumption A and condition (8) of Assumption C. For instance, consider the quasi-stationary case in Corollary 2.3.2 and Corollary 2.4.5, where they reduce to (2.3.7) and either (2.4.13) or (2.4.14). While in the former case we only need certain integrability properties of the kernel, we explicitly have to care about the size of the integrals involved in the latter case, which is also the size condition we have mentioned in Example 2.4.1(2). Also notice that this is related to the fact that in the case $I = \mathbb{R}$, we typically cannot make the left-hand side of (2.4.10) as small as we want by refining the subdivision $\mathcal{T}$ since the first interval $(t_0, t_1] = (-\infty, t_1]$ always has infinite length. So whereas condition (7) of Assumption A is quite natural for $I = [0,\infty)$, the analogous condition for $I = \mathbb{R}$ would be very restrictive.

(2) By the nature of Equation (2.1.1), the size condition (8) of Assumption C is "symmetric" in G, σ and Λ.

(3) In Theorem 2.4.4 uniqueness does not hold in $\tilde{\mathcal{P}}$: see Equation (2.4.1) with $\lambda \in (1,\infty)$. □

The next theorem reports some basic properties of the solution found in Theorem 2.4.4:

Theorem 2.4.7. *Let Assumption C be valid and* Y *be the unique solution to Equation* (2.1.1) *in* $B^{p,w}_{I,\mathrm{loc}}$.

(1) For $(t,x),(\tau,\xi),(s,y) \in I \times \mathbb{R}^d$ define $\tilde{G}(t,x;\tau,\xi;s,y)$ as

$$\Bigg(|G(t,x;s,y) - G(\tau,\xi;s,y)|^p \left(\int_{\mathbb{R}} |z|^p \, \pi(s,y,\mathrm{d}z) + c(s,y)\right) + \Big|[G(t,x;s,y) - G(\tau,\xi;s,y)]b_1(s,y)\Big| \mathbb{1}_{\{p\geq 1\}}\Bigg) w(s,y). \tag{2.4.15}$$

If for all $(t,x) \in I \times \mathbb{R}^d$

$$\int_I \int_{\mathbb{R}^d} \tilde{G}(t,x;\tau,\xi;s,y) \, \lambda(\mathrm{d}s,\mathrm{d}y) \to 0 \tag{2.4.16}$$

whenever $(\tau,\xi) \to (t,x)$, then Y is an L^p-continuous process, that is,

$$\mathbb{E}[|Y(t,x) - Y(\tau,\xi)|^p] \to 0, \quad \text{whenever } (\tau,\xi) \to (t,x). \tag{2.4.17}$$

(2) Assume the case of Corollary 2.4.5 with $G(t,x;s,y) = g(t-s,x-y)$. Then (2.4.16) *and therefore the conclusion of (1) hold automatically.*

(3) Y depends continuously on Y_0. In other words, if Y and Y' are the solutions to (2.1.1) *with $Y_0, Y_0' \in B^{p,w}_{I,\mathrm{loc}}$ as force functions, respectively, then there exists a constant $C_{I,T,w} \in \mathbb{R}_+$ that may depend on I, T and w, but is independent of Y_0, Y_0' such that*

$$\|Y - Y'\|_{B^{p,w}_{I_T}} \leq C_{I,T,w} \|Y_0 - Y_0'\|_{B^{p,w}_{I_T}}. \tag{2.4.18}$$

One of our basic motivations for studying Equation (2.1.1) on $I = \mathbb{R}$ is to construct stationary solutions. We show that if G is of convolution form and Λ is homogeneous over space and time, then the stationarity of the solution in Theorem 2.4.4 follows naturally.

Theorem 2.4.8. *Assume that $G(t,x;s,y) = g(t-s,x-y)$ and that Λ is a homogeneous Lévy basis, satisfying the assumptions of Corollary 2.4.5. Furthermore, suppose that for all $(t,x) \in \mathbb{R}_+ \times \mathbb{R}^d$ we have*

$$s \downarrow t, \quad y \downarrow x \ (\text{i.e. } y_i \downarrow x_i \text{ for all } i = 1,\dots,d) \quad \Longrightarrow \quad g(s,y) \to g(t,x), \tag{2.4.19}$$

or that for all $(t,x) \in \mathbb{R}_+ \times \mathbb{R}^d$ implication (2.4.19) *holds with $\downarrow$ replaced by $\uparrow$. If Y_0 is L^p-continuous and jointly stationary with Λ, then also Y and Λ are jointly stationary.*

Example 2.4.9 (Continuation of Examples 2.3.4 and 2.3.8) While the number a in (2.3.8) did not play any role in Examples 2.3.4 and 2.3.8, this changes when we consider the stochastic heat equation on $I = \mathbb{R}$. Let $p \in (0, 1 + 2/d)$ and set $p(d) := (1-p)d/2$. Then we have the following trichotomy: for $a > 0$ we have

$$\int_0^\infty \int_{\mathbb{R}^d} g_a^p(t,x)\,\mathrm{d}(t,x) = (4\pi)^{p(d)} p^{-\frac{d}{2}} (ap)^{-1-p(d)} \Gamma(1+p(d)); \tag{2.4.20}$$

for $a = 0$ we have for $T \in \mathbb{R}_+$

$$\int_0^T \int_{\mathbb{R}^d} g_0^p(t,x)\,\mathrm{d}(t,x) = \frac{(4\pi)^{p(d)} p^{-\frac{d}{2}}}{1+p(d)} T^{1+p(d)}, \tag{2.4.21}$$

which is of polynomial growth when $T \to \infty$; finally, if $a < 0$, we have

$$\int_0^T \int_{\mathbb{R}^d} g_a^p(t,x)\,\mathrm{d}(t,x) = (4\pi)^{p(d)} p^{-\frac{d}{2}} \int_0^T \mathrm{e}^{-apt} t^{p(d)}\,\mathrm{d}t, \tag{2.4.22}$$

which grows faster than e^{-apT} as $T \to \infty$. Thus, in the latter two cases, for Theorem 2.4.4 to be applicable, the characteristics of Λ must decay fast enough at $-\infty$ to ensure the integrability conditions (4), (5) and (6) of Assumption C.

We will only focus on the case $a > 0$. Given sufficiently strong decay properties of Λ at $-\infty$, the subsequent arguments can easily be transferred to the other two cases. First, we assume that $w \equiv 1$ and that (1) and (2) of Assumption C hold. We further suppose the quasi-stationary case of (2.4.11), and that the following conditions hold:

$$p < 1 + \frac{2}{d}, \quad b_0 \equiv 0 \text{ if } p < 1, \quad c \equiv 0 \text{ if } p < 2, \quad \zeta_p := \int_{\mathbb{R}} |z|^p\,\pi_0(\mathrm{d}z) < \infty. \tag{2.4.23}$$

The only condition left is the size condition (2.4.13) for $p \in (0,1)$ and (2.4.14) for $p \in [1,2]$, respectively, before we can apply Corollary 2.4.5. By (2.4.20), they are equivalent to

$$\zeta_p C_{\sigma,1}^p (4\pi)^{p(d)} p^{-\frac{d}{2}} (ap)^{-1-p(d)} \Gamma(1+p(d)) < 1 \tag{2.4.24}$$

in the case $p \in (0,1)$, and to

$$C_{\sigma,1}\left[C_p^{\mathrm{BDG}}\left((\zeta_p + \|c\|_{L^\infty_{\mathbb{R}}})(4\pi)^{p(d)} p^{-\frac{d}{2}} (ap)^{-1-p(d)} \Gamma(1+p(d))\right)^{1/p} + \|b_1\|_{L^\infty_{\mathbb{R}}} a^{-1}\right] < 1 \tag{2.4.25}$$

in the case $p \in [1,2]$.

Finally, we would like to demonstrate how weighted spaces can be useful in Theorem 2.4.4. Let $a > 0$ and $p \in (0, 1 + 2/d)$ as before and define $w(t,x) := \mathrm{e}^{\eta t}$ with $\eta \in \mathbb{R}$ satisfying $ap + \eta > 0$. Assume that $Y_0 \in B^{p,w}_{\mathbb{R},\mathrm{loc}}$ and, if $\eta < 0$, that $\sigma(0) = 0$. Since

$$\sup_{(t,x)\in\mathbb{R}\times\mathbb{R}^d} \int_{-\infty}^{t} \int_{\mathbb{R}^d} w^{-1}(t,x) g_a^p(t-s, x-y) w(s,y)\,\mathrm{d}(s,y)$$
$$= \int_0^\infty \int_{\mathbb{R}^d} g_a^p(s,y)\mathrm{e}^{-\eta s}\,\mathrm{d}(s,y) = (4\pi)^{p(d)} p^{-\frac{d}{2}} (ap+\eta)^{-1-p(d)} \Gamma(1+p(d)),$$

we have that in the conditions (2.4.24) and (2.4.25), the term ap is now replaced by $ap + \eta$. We draw two conclusions: if Y_0 is sufficiently small at $-\infty$, meaning $Y_0 \in B^{p,w}_{\mathbb{R},\mathrm{loc}}$ for some $\eta > 0$, then the conditions (2.4.24) and (2.4.25) can be relaxed by using $ap+\eta$ instead of ap. Contrarily, if $\sigma(0) = 0$, $\eta < 0$, and the left-hand side of (2.4.24) or (2.4.25), respectively, remains smaller than 1 with $ap + \eta$ instead of ap, then one can even construct solutions with $Y_0 \in B^{p,w}_{\mathbb{R},\mathrm{loc}}$ that diverges at $-\infty$. □

2.5 Asymptotic stability

In Theorems 2.3.1, 2.3.5 and 2.4.4 we have established solutions to (2.1.1) that belong to the space $B^{p,w}_{I,\mathrm{loc}}$. In this section we will give criteria under which they even belong to the space $B^{p,w}_I$. Our primary focus is on the case where $\sup I = +\infty$, that is, we want to investigate whether solutions to (2.1.1) are asymptotically L^p-stable. Moreover, we shall replace the Lipschitz condition on σ, which was essential in Sections 2.3 and 2.4, by another growth condition, which, as we shall see, will determine the asymptotic behaviour of the solution. Of course, due to the possible non-Lipschitzianity of σ, we now have to *assume* the existence of a solution in $B^{p,w}_{I,\mathrm{loc}}$. In fact, this approach allows us to include solutions to (2.1.1) with non-Lipschitz σ which go beyond the results of the Sections 2.3 and 2.4 but are, for instance, studied in [107, 108].

Let us again start with a deterministic example that highlights the main features of the behaviour at infinity.

Example 2.5.1 Let $g \in L^1_{[0,\infty)}$, $f \in L^\infty_{[0,\infty)}$ and $v \in L^\infty_{[0,\infty),\mathrm{loc}}$ be positive functions satisfying

$$v(t) = f(t) + \int_0^t g(t-s) v^\gamma(s)\,\mathrm{d}s, \quad t \in \mathbb{R}_+, \tag{2.5.1}$$

with $\gamma \in (0,1]$. The question is under what conditions we have $v \in L^\infty_{[0,\infty)}$. It turns out that there is a fundamental difference between the cases $\gamma \in (0,1)$ and $\gamma = 1$. In the former case, we always have $v \in L^\infty_{[0,\infty)}$. In fact, if we denote the convolution on the right-hand side of (2.5.1) by $(g * v^\gamma)(t)$, iteration of (2.5.1) yields

$$v = f + g * v^\gamma = f + g * (f + g * v^\gamma)^\gamma = f + g * (f + g * (f + g * v^\gamma)^\gamma)^\gamma = \dots$$

Using Young's inequality, we obtain

$$\begin{aligned}
&\|v\|_{L^\infty_{[0,T]}} \\
&\le \|f\|_{L^\infty_{[0,\infty)}} + \|g\|_{L^1_{[0,\infty)}} \|v\|^\gamma_{L^\infty_{[0,T]}} \\
&\le \|f\|_{L^\infty_{[0,\infty)}} + \|g\|_{L^1_{[0,\infty)}} (\|f\|_{L^\infty_{[0,\infty)}} + \|g\|_{L^1_{[0,\infty)}} \|v\|^\gamma_{L^\infty_{[0,T]}})^\gamma \\
&\le \|f\|_{L^\infty_{[0,\infty)}} + \|g\|_{L^1_{[0,\infty)}} (\|f\|_{L^\infty_{[0,\infty)}} + \|g\|_{L^1_{[0,\infty)}} (\|f\|_{L^\infty_{[0,\infty)}} + \|g\|_{L^1_{[0,\infty)}} \|v\|^\gamma_{L^\infty_{[0,T]}})^\gamma)^\gamma \\
&\le \dots,
\end{aligned}$$

or, equivalently, $\|v\|_{L^\infty_{[0,T]}} \le a_n(T)$ for every $T \in [0,\infty)$ and $n \in \mathbb{N}$ where

$$a_1(T) := \|v\|_{L^\infty_{[0,T]}}, \quad a_{n+1}(T) := \|f\|_{L^\infty_{[0,\infty)}} + \|g\|_{L^1_{[0,\infty)}} (a_n(T))^\gamma.$$

By induction it can be shown that $0 \le a_n(T) \le a \vee a_1(T)$, where a is the unique solution in $(0,\infty)$ of the equation

$$a - \|f\|_{L^\infty_{[0,\infty)}} - \|g\|_{L^1_{[0,\infty)}} a^\gamma = 0.$$

Note that a does not depend on T, so we conclude that $\limsup_{n\to\infty} a_n(T) \le a$ and $\|v\|_{L^\infty_{[0,T]}} \le a$ for all $T \in [0,\infty)$. Hence we have $v \in L^\infty_{[0,\infty)}$ with $\|v\|_{L^\infty_{[0,\infty)}} \le a$.

The situation is totally different for $\gamma = 1$. Then (2.5.1) becomes

$$v(t) = f(t) + \int_0^t g(t-s)v(s)\,\mathrm{d}s, \quad t \in \mathbb{R}_+, \tag{2.5.2}$$

which is the well known *renewal equation*. If $f \in L^\infty_{[0,\infty)}$, one can show under some technical assumptions that the unique solution v to (2.5.2) exhibits the following behaviour: if $\|g\|_{L^1_{[0,\infty)}} < 1$, we have $v \in L^\infty_{[0,\infty)}$; if $\|g\|_{L^1_{[0,\infty)}} = 1$, the boundedness of v depends on whether $f \in L^1_{[0,\infty)}$ or not; if $\|g\|_{L^1_{[0,\infty)}} > 1$, then $v(t) \to \infty$ exponentially fast as $t \to \infty$. For precise statements with the required assumptions, we refer to [6, Chap. V], especially to the Theorems V.4.3 and V.7.1 and Proposition V.7.4.

In summary, whereas locally bounded solutions to (2.5.1) with $\gamma \in (0,1)$ are automatically globally bounded as soon as $f \in L^\infty_{[0,\infty)}$ and $g \in L^1_{[0,\infty)}$, the behaviour of the solution to (2.5.2) at infinity strongly depends on the *size* of $\|g\|_{L^1_{[0,\infty)}}$. For a formalization of this example see also Lemma 2.6.5 for $\gamma \in (0,1)$ and Lemma 2.6.4 for $\gamma = 1$. □

For Equation (2.1.1) the precise requirements are the following:

Assumption D Let $p \in (0,2]$, I be an interval and $w\colon I \times \mathbb{R}^d \to \mathbb{R}$ be a weight function satisfying $\sup_{(t,x)\in I\times\mathbb{R}^d} w^{-1}(t,x) < \infty$. We assume the following hypotheses:

(1) $Y_0 \in B^{p,w}_I$.

(2) $\sigma\colon \mathbb{R} \to \mathbb{R}$ satisfies $|\sigma(x)| \le |\sigma(0)| + C_{\sigma,2}|x|^\gamma$ for all $x \in \mathbb{R}$ with some $\gamma \in (0,1]$.

(3) Either $c(t,x) = 0$ for all $(t,x) \in I \times \mathbb{R}^d$, or we have $2\gamma \le p$ and

$$\sup_{(t,x)\in I\times\mathbb{R}^d} \int_I \int_{\mathbb{R}^d} w^{-1}(t,x)|G(t,x;s,y)|^2 w(s,y)c(s,y)\,\lambda(\mathrm{d}s,\mathrm{d}y) < \infty. \tag{2.5.3}$$

(4) There exists $q \in (0,2]$ with $p \le q$ and $q\gamma \le p$ such that

$$\sup_{(t,x)\in I\times\mathbb{R}^d} \int_I \int_{\mathbb{R}^d} \int_{\mathbb{R}} w^{-1}(t,x)|G(t,x;s,y)z|^p_q w(s,y)\,\nu(\mathrm{d}s,\mathrm{d}y,\mathrm{d}z) < \infty. \tag{2.5.4}$$

(5) If $p \ge 1$, then ν satisfies (2.2.4) and

$$\sup_{(t,x)\in I\times\mathbb{R}^d} \int_I \int_{\mathbb{R}^d} w^{-1}(t,x)|G(t,x;s,y)b_1(s,y)|w(s,y)\,\lambda(\mathrm{d}s,\mathrm{d}y) < \infty, \tag{2.5.5}$$

and (2.5.5) also holds with $w \equiv 1$; if $p < 1$, then there exist $\alpha \in (-\infty, 2]$ and $\beta \in [0,\infty)$ satisfying (2.3.13), (2.3.14) (with $\mathbb{R}_+$ replaced by I) and $(\alpha \vee \beta)\gamma \le p$ such that

$$\sup_{(t,x)\in I\times\mathbb{R}^d} \int_I \int_{\mathbb{R}^d} (F_0(s,y) \vee F_1(s,y))|G(t,x;s,y)|^\alpha_\beta\,\lambda(\mathrm{d}s,\mathrm{d}y) < \infty. \tag{2.5.6}$$

(6) At least one of the following three cases occurs:

(6a) We have $\gamma < 1$, $q\gamma < p$, $2\gamma < p$ if $c \not\equiv 0$ and $(\alpha \vee \beta)\gamma < p$ if $p < 1$.

(6b) We have $p \in [1,2]$, and if we define for $(t,x),(s,y) \in I \times \mathbb{R}^d$

$$\begin{aligned}
\bar{G}^{D,1}(t,x;s,y) &:= 2^{p-1} \left(\int_I \int_{\mathbb{R}^d} |G(t,x;s,y) b_1(s,y)| \, \lambda(\mathrm{d}s, \mathrm{d}y) \right)^{p-1} \\
&\quad \times |G(t,x;s,y) b_1(s,y)|, \\
\bar{G}^{D,2}(t,x;s,y) &:= 2 (C_p^{\mathrm{BDG}})^2 |G(t,x;s,y)|^2 c(s,y), \\
\bar{G}^{D,3}(t,x;s,y) &:= 2^{p-1} (C_p^{\mathrm{BDG}})^p \int_{\mathbb{R}} |G(t,x;s,y) z|^p \mathbb{1}_{\{|G(t,x;s,y)z|>1\}} \, \pi(s,y,\mathrm{d}z), \\
\bar{G}^{D,4}(t,x;s,y) &:= 2^{q-1} (C_p^{\mathrm{BDG}})^q \int_{\mathbb{R}} |G(t,x;s,y) z|^q \mathbb{1}_{\{|G(t,x;s,y)z|\leq 1\}} \, \pi(s,y,\mathrm{d}z), \\
G^{D,l}(t,x;s,y) &:= w^{-1}(t,x) G^{D,l}(t,x;s,y) w(s,y), \quad l = 1,2,3,4,
\end{aligned} \tag{2.5.7}$$

then there exists a partition of I into pairwise disjoint intervals $I_1, \dots, I_k$ such that

$$\sup_{(t,x) \in I \times \mathbb{R}^d} \sup_{j=1,\dots,k} \sum_{l=1}^4 C_{\sigma,2} \left(\int_{I_j} \int_{\mathbb{R}^d} G^{D,l}(t,x;s,y) \, \lambda(\mathrm{d}s, \mathrm{d}y) \right)^{1/p} < 1. \tag{2.5.8}$$

(6c) We have $p \in (0,1)$, and if we define for $(t,x),(s,y) \in I \times \mathbb{R}^d$

$$\begin{aligned}
G^{D,1}(t,x;s,y) &:= 2^{(\alpha \vee \beta \vee 1)-1} (F_0(s,y) \vee F_1(s,y)) |G(t,x;s,y)|_\beta^\alpha, \\
G^{D,2}(t,x;s,y) &:= 2^{p+1} |G(t,x;s,y)|^2 c(s,y), \\
G^{D,3}(t,x;s,y) &:= 2^p 2^{(q \vee 1)-1} \int_{\mathbb{R}} |G(t,x;s,y) z|_q^p \, \pi(s,y,\mathrm{d}z),
\end{aligned} \tag{2.5.9}$$

and

$$r_1 := \alpha \vee \beta, \quad r_2 := 2, \quad r_3 := 1, \tag{2.5.10}$$

then there exists a partition of I into pairwise disjoint intervals $I_1, \dots, I_k$ such that

$$\sup_{(t,x) \in I \times \mathbb{R}^d} \sup_{j=1,\dots,k} \sum_{l=1}^3 C_{\sigma,2}^{r_l} \int_{I_j} \int_{\mathbb{R}^d} G^{D,l}(t,x;s,y) \, \lambda(\mathrm{d}s, \mathrm{d}y) < 1. \tag{2.5.11}$$

□

Theorem 2.5.2. *Let Assumption D be valid. If Equation* (2.1.1) *has a solution* $Y \in B_{I,\mathrm{loc}}^{p,w}$*, it automatically also belongs to* $B_I^{p,w}$*.*

Example 2.5.3 (Continuation of Examples 2.3.4, 2.3.8 and 2.4.9) Let $I = [0, \infty)$, $a = 0$ and $w \equiv 1$. We assume that $Y \in B^p_{[0,\infty),\mathrm{loc}}$ solves

$$Y(t,x) = Y_0(t,x) + \int_0^t \int_{\mathbb{R}^d} g_0(t-s, x-y)\sigma(Y(s,y))\,\Lambda(\mathrm{d}s, \mathrm{d}y), \quad (t,x) \in \mathbb{R}_+ \times \mathbb{R}^d,$$

where Y_0 is given by (2.3.9) and σ satisfies condition (2) of Assumption D with $\gamma \in (0,1]$. We want to find conditions that guarantee $Y \in B^p_{[0,\infty)}$. Let us check the requirements of Assumption D. (1) and (2) are clear. For (3), (4) and (5), the key observation is the following: for $p, q \in (0,2]$

$$\int_0^\infty \int_{\mathbb{R}^d} |g_0(s,y)|_q^p \,\mathrm{d}(s,y) < \infty \iff p \in (0, 1+2/d) \text{ and } q \in (1+2/d, 2]. \quad (2.5.12)$$

As a consequence of the last condition, unless in trivial cases, the classical stochastic heat equation with $a = 0$ will be asymptotically unstable in dimensions 1 and 2. Only in dimensions $d \geq 3$ there is a chance for asymptotic stability. We pose the following conditions:

$$\begin{aligned} &\lambda(\mathrm{d}t, \mathrm{d}x) = \mathrm{d}(t,x), \quad \pi(t,x,\mathrm{d}z) \leq \pi_0(\mathrm{d}z), \quad p \in (0, 1+2/d), \\ &q \in (1+2/d, 2], \quad q\gamma \leq p, \quad c \equiv 0, \quad b_1 \equiv 0 \text{ if } p \geq 1, \\ &\Lambda \text{ is symmetric if } p < 1, \quad \int_{\mathbb{R}} |z|_p^q \,\pi_0(\mathrm{d}z) < \infty. \end{aligned} \quad (2.5.13)$$

We notice that $\gamma = 1$ is not possible, and that $\int_{\mathbb{R}} |z|^p \,\pi_0(\mathrm{d}z) < \infty$ is no longer sufficient, but π_0 must have a moment structure that is strictly better than its variation structure. Moreover, c must be 0; if $p \geq 1$, only $\Lambda \in \mathcal{M}$ is possible; and if $p < 1$, Λ is required to have no drift and a symmetric Lévy measure. All this is because g_0 is not L^p-integrable on $\mathbb{R}_+ \times \mathbb{R}^d$ for any $p \in (0,2]$. One readily sees that (2.5.13) implies conditions (3), (4) and (5). So if (6a) holds, we obtain $Y \in B^p_{[0,\infty)}$. In the case of (6b) or (6c), again a size condition has to be verified, which is analogous to the calculations in Example 2.4.9. We leave the details to the reader. Note that in this example we have $\gamma < 1$, and therefore (6b) or (6c) is only needed in rare situations. Finally, for $a > 0$ we refer the reader to the calculations in Example 2.4.9 again which can be re-used. In particular, one can find conditions for asymptotic stability in dimensions 1 and 2 this time. □

2.6 A series of lemmata

This section contains several lemmata that will play a crucial role in proving the main theorems in Section 2.7. First, we investigate the stochastic integral mapping in Equation (2.1.1): fix some $\phi_0 \in \tilde{\mathcal{P}}$ and define for a predictable process $\phi \in \tilde{\mathcal{P}}$ the process $J(\phi)$ by

$$J(\phi)(t,x) := \phi_0(t,x) + \int_I \int_{\mathbb{R}^d} G(t,x;s,y)\sigma(\phi(s,y))\,\Lambda(\mathrm{d}s,\mathrm{d}y) \tag{2.6.1}$$

for all $(t,x) \in I \in \mathbb{R}^d$ for which the stochastic integral exists, and set $J(\phi)(t,x) := \infty$ otherwise. The next lemma, which is of crucial importance for all main results in this Chapter, relates the moment structure of $J(\phi)$ to that of ϕ.

Lemma 2.6.1. *Let $w\colon I \times \mathbb{R}^d \to \mathbb{R}$ be a weight function.*

(1) Suppose that Assumption C holds with $p \in (0,2]$ and recall the definition of $G^{C,1}$ and $G^{C,2}$ in (2.4.9)*. Then for all $\phi \in \tilde{\mathcal{P}}$ and $(t,x) \in I \times \mathbb{R}^d$, we have*

$$\frac{\|J(\phi)(t,x)\|_{L^p}}{(w(t,x))^{1/p^*}} \le \frac{\|\phi_0(t,x)\|_{L^p}}{(w(t,x))^{1/p^*}} + \sum_{l=1}^{2} \left(\int_I \int_{\mathbb{R}^d} \frac{G^{C,l}(t,x;s,y)}{C_{\sigma,1}^p} \right.$$
$$\left. \times \left(\frac{|\sigma(0)|^{p\wedge 1} + C_{\sigma,1}^{p\wedge 1}\|\phi(s,y)\|_{L^p}}{(w(s,y))^{1/p^*}} \right)^{p^*} \lambda(\mathrm{d}s,\mathrm{d}y) \right)^{1/p^*}, \tag{2.6.2}$$

where in the case $C_{\sigma,1} = 0$ we use the convention $0/0 := 1$.

(2) Furthermore, still under Assumption C, we have for any $\phi_1, \phi_2 \in \tilde{\mathcal{P}}$ for which the right-hand side of (2.6.2) *is finite that*

$$\frac{\|J(\phi_1)(t,x) - J(\phi_2)(t,x)\|_{L^p}}{(w(t,x))^{1/p^*}} \le \sum_{l=1}^{2} \left(\int_I \int_{\mathbb{R}^d} G^{C,l}(t,x;s,y) \right.$$
$$\left. \times \left(\frac{\|\phi_1(s,y) - \phi_2(s,y)\|_{L^p}}{(w(s,y))^{1/p^*}} \right)^{p^*} \lambda(\mathrm{d}s,\mathrm{d}y) \right)^{1/p^*}. \tag{2.6.3}$$

(3) Let Assumption B or Assumption D be valid with $p \in [1,2]$. In the first case let

$I = [0, \infty)$ *and* $w \equiv 1$. *Then the following holds for all* $\phi \in \tilde{\mathcal{P}}$ *and* $(t, x) \in I \times \mathbb{R}^d$*:*

$$\begin{aligned}\frac{\|J(\phi)(t,x)\|_{L^p}}{(w(t,x))^{1/p}} &\leq \frac{\|\phi_0(t,x)\|_{L^p}}{(w(t,x))^{1/p}} + \frac{2[1+|\sigma(0)|+C_{\sigma,2}]}{(w(t,x))^{1/p}} \\ &\quad + (|\sigma(0)| + C_{\sigma,2}) \sum_{l=1}^{4} \left(\int_I \int_{\mathbb{R}^d} G^{D,l}(t,x;s,y)(w(s,y))^{-1} \, \lambda(\mathrm{d}s, \mathrm{d}y) \right)^{1/p} \\ &\quad + \sum_{l=1}^{4} C_{\sigma,2} \left(\int_I \int_{\mathbb{R}^d} \frac{G^{D,l}(t,x;s,y)}{(w(s,y))^{1-\rho}} \left(\frac{\|\phi(s,y)\|_{L^p}}{(w(s,y))^{1/p}} \right)^{p\rho} \lambda(\mathrm{d}s, \mathrm{d}y) \right)^{1/p}, \quad (2.6.4)\end{aligned}$$

where $G^{D,l}$ *is defined by* (2.5.7)*, and* ρ *can be chosen as* $\rho = (q \vee 2\mathbb{1}_{\{c \neq 0\}})\gamma/p$ *or* $\rho = 1$.

(4) Let Assumption B or Assumption D be valid with $p \in (0, 1)$. *In the first case let* $I = [0, \infty)$ *and* $w \equiv 1$. *Then for all* $\phi \in \tilde{\mathcal{P}}$ *and* $(t, x) \in I \times \mathbb{R}^d$

$$\begin{aligned}\frac{\|J(\phi)(t,x)\|_{L^p}}{w(t,x)} &\leq \frac{\|\phi_0(t,x)\|_{L^p}}{w(t,x)} + \frac{2^{p+1}+1}{w(t,x)} \\ &\quad + \sum_{l=1}^{3} (|\sigma(0)|_0^{r_l} + C_{\sigma,2}^{r_l}) \int_I \int_{\mathbb{R}^d} G^{D,l}(t,x;s,y)(w(s,y))^{-1} \, \lambda(\mathrm{d}s, \mathrm{d}y) \\ &\quad + \sum_{l=1}^{3} C_{\sigma,2}^{r_l} \int_I \int_{\mathbb{R}^d} \frac{G^{D,l}(t,x;s,y)}{(w(s,y))^{1-\rho}} \left(\frac{\|\phi(s,y)\|_{L^p}}{w(s,y)} \right)^{\rho} \lambda(\mathrm{d}s, \mathrm{d}y). \quad (2.6.5)\end{aligned}$$

where $G^{D,l}$ *and* r_l *are given by* (2.5.9) *and* (2.5.10)*, and* ρ *can be chosen as* $\rho = 1$ *or* $\rho = (q \vee 2\mathbb{1}_{\{c \neq 0\}} \vee \alpha \vee \beta)\gamma/p$.

Proof. It suffices to prove the lemma for $w \equiv 1$: the general case follows if we divide the equations (2.6.2), (2.6.3) and (2.6.4) by w^{1/p^*}. Throughout the proof, $(t, x) \in I \times \mathbb{R}^d$ is fixed, and the abbreviations $\Psi(s, y) := G(t, x; s, y)\sigma(\phi(s, y))$ and $\Phi(s, y) := G(t, x; s, y)[\sigma(\phi_1(s, y)) - \sigma(\phi_2(s, y))]$ are used. Moreover, in the numerous integrals below, we will often drop the integration variables and use the shorthand notations $\iint_t := \int_{I_t} \int_{\mathbb{R}^d}$ and $\iiint_t := \int_{I_t} \int_{\mathbb{R}^d} \int_{\mathbb{R}}$.

a) We first prove (2) when $p \geq 1$. To this end, we decompose

$$\begin{aligned}\Lambda(\mathrm{d}t, \mathrm{d}x) &= \left[\Lambda^{\mathrm{c}}(\mathrm{d}t, \mathrm{d}x) + \int_{\mathbb{R}} z \, (\mu - \nu)(\mathrm{d}t, \mathrm{d}x, \mathrm{d}z) \right] \\ &\quad + \left[B(\mathrm{d}t, \mathrm{d}x) + \int_{\mathbb{R}} z \mathbb{1}_{\{|z|>1\}} \, \nu(\mathrm{d}t, \mathrm{d}x, \mathrm{d}z) \right] \\ &=: M(\mathrm{d}t, \mathrm{d}x) + B_1(\mathrm{d}t, \mathrm{d}x), \quad (2.6.6)\end{aligned}$$

and obtain that $\|J(\phi_1)(t,x) - J(\phi_2)(t,x)\|_{L^p}$ is bounded by

$$\|J^{(1)}(\phi_1)(t,x) - J^{(1)}(\phi_2)(t,x)\|_{L^p} + \|J^{(2)}(\phi_1)(t,x) - J^{(2)}(\phi_2)(t,x)\|_{L^p},$$

where $J^{(1)}$ and $J^{(2)}$ are defined as in (2.6.1) with Λ replaced by M and B_1, respectively. For the $J^{(2)}$-part, Hölder's inequality yields

$$\begin{aligned}
&\|J^{(2)}(\phi_1)(t,x) - J^{(2)}(\phi_2)(t,x)\|_{L^p} \\
&\quad \le C_{\sigma,1} \left[\left(\iint_t |G| \,\mathrm{d}|B_1| \right)^{p-1} \iint_t |G| \mathbb{E}[|\phi_1 - \phi_2|^p] \,\mathrm{d}|B_1| \right]^{1/p} \\
&\quad = \left(\iint_t G^{C,2}(t,x;s,y) \|\phi_1(s,y) - \phi_2(s,y)\|_{L^p}^p \, \lambda(\mathrm{d}s, \mathrm{d}y) \right)^{1/p}. \qquad (2.6.7)
\end{aligned}$$

For the $J^{(1)}$-part, we assume for the moment that the process

$$N_\tau := \iint_\tau G(t,x;s,y)[\sigma(\phi_1(s,y)) - \sigma(\phi_2(s,y))] \, M(\mathrm{d}s, \mathrm{d}y) = \Phi \cdot M_\tau, \quad \tau \in I, \ (2.6.8)$$

which is well defined by assumption, is a local martingale. Then we have by Definition 2.4.2 and the assumption that $c \equiv 0$ for $p < 2$

$$\begin{aligned}
&\|J^{(1)}(\phi_1)(t,x) - J^{(1)}(\phi_2)(t,x)\|_{L^p} \\
&\le C_p^{\mathrm{BDG}} \left\| [N]_t^{1/2} \right\|_{L^p} = C_p^{\mathrm{BDG}} \left\| \left(\iiint_t |\Phi z|^2 \,\mathrm{d}\mu + \iint_t |\Phi|^2 \,\mathrm{d}C \right)^{1/2} \right\|_{L^p} \\
&\le C_p^{\mathrm{BDG}} \mathbb{E} \left[\iiint_t |\Phi z|^p \,\mathrm{d}\mu + \iint_t |\Phi|^2 \,\mathrm{d}C \right]^{1/p} = C_p^{\mathrm{BDG}} \mathbb{E} \left[\iiint_t |\Phi z|^p \,\mathrm{d}\nu + \iint_t |\Phi|^2 \,\mathrm{d}C \right]^{1/p} \\
&\le \left(\iint_t G^{C,1}(t,x;s,y) \|\phi_1(s,y) - \phi_2(s,y)\|_{L^p}^p \, \lambda(\mathrm{d}s, \mathrm{d}y) \right)^{1/p}. \qquad (2.6.9)
\end{aligned}$$

Equations (2.6.7) and (2.6.9) together imply (2.6.3) for $p \in [1,2]$. It remains to discuss whether N in (2.6.8) is a local martingale. Without loss of generality, we may assume that the right-hand side of (2.6.9) is finite; otherwise (2.6.3) becomes trivial. Let $\epsilon > 0$ and $H \in \tilde{\mathcal{P}}$ be a bounded function satisfying $|H(\omega,s,y)| \le \epsilon |\Phi(\omega,s,y)|$ pointwise for all $(\omega,s,y) \in \Omega \times I \times \mathbb{R}^d$. Then $H \cdot M$ is a martingale such that we have by the Burkholder-Davis-Gundy inequality

$$\begin{aligned}
\sup_{\tau \in I} \|H \cdot M_\tau\|_{L^p} &\le C_p^{\mathrm{BDG}} \left\| \left(\iiint_t |Hz|^2 \,\mathrm{d}\mu + \iint_t |H|^2 \,\mathrm{d}C \right)^{1/2} \right\|_{L^p} \\
&\le \epsilon C_p^{\mathrm{BDG}} \left\| \left(\iiint_t |\Phi z|^2 \,\mathrm{d}\mu + \iint_t |\Phi|^2 \,\mathrm{d}C \right)^{1/2} \right\|_{L^p}. \qquad (2.6.10)
\end{aligned}$$

The right-hand side of (2.6.10) is finite by (2.6.9). Moreover, as $\epsilon \downarrow 0$, it goes to 0 independently of H. Thus, [28, Prop. 4.9b] is applicable (the extension of this proposition to intervals I different from $I = \mathbb{R}_+$ is straightforward) and shows that N is indeed a local martingale.

b) We prove (2) when $p < 1$. By hypothesis, Λ is Lévy basis without drift. Thus,

$$\begin{aligned}\|J(\phi_1)(t,x) - J(\phi_2)(t,x)\|_{L^p} &= \left\| \iiint_t \Phi z \, \mathrm{d}\mu \right\|_{L^p} \leq \mathbb{E}\left[\iiint_t |\Phi z|^p \, \mathrm{d}\mu \right] \\ &\leq C_{\sigma,1}^p \iiint_t |Gz|^p \mathbb{E}[|\phi_1 - \phi_2|^p] \, \mathrm{d}\nu \\ &= \iint_t G^{C,1}(t,x;s,y) \|\phi_1(s,y) - \phi_2(s,y)\|_{L^p} \, \lambda(\mathrm{d}s, \mathrm{d}y),\end{aligned}$$

which is (2.6.3).

c) Because the Lipschitz condition on σ implies $|\sigma(x)| \leq |\sigma(0)| + C_{\sigma,1}|x|$ for all $x \in \mathbb{R}$, (1) can be deduced in complete analogy to a) and b).

d) We prove (3). To this end, we again consider the decomposition $\Lambda = M + B_1$ as in (2.6.6). Using Definition 2.4.2, Jensen's inequality and the hypothesis that $q\gamma \leq p$ and $2\gamma \mathbb{1}_{\{c \neq 0\}} \leq p$, we obtain

$$\begin{aligned}&\|\Psi \cdot M_t\|_{L^p} \leq C_p^{\mathrm{BDG}} \left\| \left(\iiint_t |\Psi z|^2 \, \mathrm{d}\mu + \iint_t |\Psi|^2 \, \mathrm{d}C \right)^{1/2} \right\|_{L^p} \\ &\leq C_p^{\mathrm{BDG}} \left(\mathbb{E}\left[\iiint_t |\Psi z|^q \mathbb{1}_{\{|Gz| \leq 1\}} \, \mathrm{d}\nu \right]^{p/q} \right. \\ &\quad \left. + \mathbb{E}\left[\iiint_t |\Psi z|^p \mathbb{1}_{\{|Gz|>1\}} \, \mathrm{d}\nu + \left(\iint_t |\Psi|^2 \, \mathrm{d}C \right)^{p/2} \right] \right)^{1/p} \\ &\leq C_p^{\mathrm{BDG}} \left[\left(2^{q-1} \iiint_t |Gz|^q \mathbb{1}_{\{|Gz| \leq 1\}} (|\sigma(0)|^q + C_{\sigma,2}^q \|\phi\|_{L^p}^{q\gamma}) \, \mathrm{d}\nu \right)^{1/q} \right. \\ &\quad + \left(2^{p-1} \iiint_t |Gz|^p \mathbb{1}_{\{|Gz|>1\}} (|\sigma(0)|^p + C_{\sigma,2}^p \|\phi\|_{L^p}^{p\gamma}) \, \mathrm{d}\nu \right)^{1/p} \\ &\quad \left. + \left(2 \iint_t |G|^2 (|\sigma(0)|^2 + C_{\sigma,2}^2 \|\phi\|_{L^p}^{2\gamma}) \, \mathrm{d}C \right)^{1/2} \right] \\ &\leq C_p^{\mathrm{BDG}} \left[\left(2^{q-1} \iiint_t |Gz|^q \mathbb{1}_{\{|Gz| \leq 1\}} (|\sigma(0)|^q + C_{\sigma,2}^q + C_{\sigma,2}^q \|\phi\|_{L^p}^{p\rho}) \, \mathrm{d}\nu \right)^{1/q} \right. \\ &\quad \left. + \left(2^{p-1} \iiint_t |Gz|^p \mathbb{1}_{\{|Gz|>1\}} (|\sigma(0)|^p + C_{\sigma,2}^p + C_{\sigma,2}^p \|\phi\|_{L^p}^{p\rho}) \, \mathrm{d}\nu \right)^{1/p} \right.\end{aligned}$$

$$
\begin{aligned}
&+\left(2\iint_t |G|^2(|\sigma(0)|^2+C_{\sigma,2}^2+C_{\sigma,2}^2\|\phi\|_{L^p}^{p\rho})\,\mathrm{d}C\right)^{1/2}\Bigg]\\
\leq&(|\sigma(0)|+C_{\sigma,2})\left[2+\sum_{l=2}^{4}\left(\int_I\int_{\mathbb{R}^d} G^{D,l}(t,x;s,y)\,\lambda(\mathrm{d}s,\mathrm{d}y)\right)^{1/p}\right]\\
&+2+\sum_{l=2}^{4}C_{\sigma,2}\left(\int_I\int_{\mathbb{R}^d} G^{D,l}(t,x;s,y)\|\phi(s,y)\|_{L^p}^{p\rho}\,\lambda(\mathrm{d}s,\mathrm{d}y)\right)^{1/p}. \qquad (2.6.11)
\end{aligned}
$$

Again, one can justify that $\Psi\cdot\Lambda$ is indeed a well-defined local martingale whenever the right-hand side of (2.6.4) is finite.

For the B_1-integral, another application of Hölder's inequality demonstrates

$$
\begin{aligned}
\|(\Psi\cdot B_1)_t\|_{L^p} &\leq \left[2^{p-1}\left(\iint_t |G|\,\mathrm{d}|B_1|\right)^{p-1}\iint_t |G|(|\sigma(0)|^p+C_{\sigma,2}^p\|\phi\|_{L^p}^{p\gamma})\,\mathrm{d}|B_1|\right]^{1/p}\\
&\leq \left[2^{p-1}\left(\iint_t |G|\,\mathrm{d}|B_1|\right)^{p-1}\iint_t |G|(|\sigma(0)|^p+C_{\sigma,2}^p+C_{\sigma,2}^p\|\phi\|_{L^p}^{p\rho})\,\mathrm{d}|B_1|\right]^{1/p}. \qquad (2.6.12)
\end{aligned}
$$

Equation (2.6.4) now follows from (2.6.11) and (2.6.12).

e) We consider the last part (4). In this case we directly use the canonical decomposition of $\Psi\cdot\Lambda$:

$$
\begin{aligned}
\Psi\cdot\Lambda_t &= \Psi\cdot\Lambda_t^{\mathrm{c}}+\iiint_t \Psi z\mathbb{1}_{\{|\Psi z|\leq 1\}}\,\mathrm{d}(\mu-\nu)+\iiint_t \Psi z\mathbb{1}_{\{|\Psi z|>1\}}\,\mathrm{d}\mu+B_t^{\Psi\cdot\Lambda}\\
&=: J^1+J^2+J^3+J^4,
\end{aligned}
$$

where

$$
B^{\Psi\cdot\Lambda}(\mathrm{d}t,\mathrm{d}x)=\Psi(t,x)\left[b(t,x)+\int_{\mathbb{R}} z(\mathbb{1}_{\{|\Psi(t,x)z|\leq 1\}}-\mathbb{1}_{\{|z|\leq 1\}})\,\pi(t,x,\mathrm{d}z)\right]\lambda(\mathrm{d}t,\mathrm{d}x).
$$

We begin with J^1:

$$
\begin{aligned}
\|J^1\|_{L^p} &\leq (C_1^{\mathrm{BDG}})^p\mathbb{E}\left[\left(\iint_t |\Psi|^2\,\mathrm{d}C\right)^{1/2}\right]^p \leq 2^p\left(\iint_t \mathbb{E}[|\Psi|^2]\,\mathrm{d}C\right)^{p/2}\\
&\leq 2^p\left(1+2\iint_t G^2(|\sigma(0)|^2+C_{\sigma,2}^2\|\phi\|_{L^p}^{2\gamma/p})\,\mathrm{d}C\right)\\
&\leq 2^p\left(1+2\iint_t G^2(|\sigma(0)|^2+C_{\sigma,2}^2+C_{\sigma,2}^2\|\phi\|_{L^p}^{\rho})\,\mathrm{d}C\right).
\end{aligned}
$$

For the jumps part, we obtain

$$\begin{aligned}
&\|J^2 + J^3\|_{L^p} \\
&\leq (C_1^{\mathrm{BDG}})^p \mathbb{E}\left[\left(\iiint_t |\Psi z|^2 \mathbb{1}_{\{|\Psi z|\leq 1\}} \,\mathrm{d}\mu\right)^{1/2}\right]^p + \mathbb{E}\left[\iiint_t |\Psi z|^p \mathbb{1}_{\{|\Psi z|>1\}} \,\mathrm{d}\mu\right] \\
&\leq 2^p \left(\iiint_t \mathbb{E}[|\Psi z|^q \mathbb{1}_{\{|\Psi z|\leq 1\}}] \,\mathrm{d}\nu\right)^{p/2} + \iiint_t \mathbb{E}[|\Psi z|^p \mathbb{1}_{\{|\Psi z|>1\}}] \,\mathrm{d}\nu \\
&\leq 2^p \left(1 + \iiint_t \mathbb{E}[|\Psi z|_q^p] \,\mathrm{d}\nu\right) \leq 2^p \left(1 + 2^{(q\vee 1)-1} \iiint_t |Gz|_q^p (|\sigma(0)|_0^q + C_{\sigma,2}^q \|\phi\|_{L^p}^{q\gamma/p}) \,\mathrm{d}\nu\right) \\
&\leq 2^p \left(1 + 2^{(q\vee 1)-1} \iiint_t |Gz|_q^p (|\sigma(0)|_0^q + C_{\sigma,2}^q + C_{\sigma,2}^q \|\phi\|_{L^p}^{\rho}) \,\mathrm{d}\nu\right).
\end{aligned}$$

Finally, since

$$\begin{aligned}
|J^4| \leq \iint_t |\Psi(s,y)| \Bigg| & b(s,y) \\
&+ \int_{\mathbb{R}} \left[z \mathbb{1}_{\{|z|\in(1,|\Psi(s,y)|^{-1}]\}} - z \mathbb{1}_{\{|z|\in(|\Psi(s,y)|^{-1},1]\}}\right] \pi(s,y,\mathrm{d}z) \Bigg| \lambda(\mathrm{d}s, \mathrm{d}y),
\end{aligned}$$

we deduce the following bound for J^4 from Assumption B(9) or Assumption D(4), respectively:

$$\begin{aligned}
\|J^4\|_{L^p} &\leq \mathbb{E}\left[\iint_t |\Psi| (|\Psi|^{\beta-1} F_1 \mathbb{1}_{\{|\Psi|\leq 1\}} + |\Psi|^{\alpha-1} F_0 \mathbb{1}_{\{|\Psi|>1\}}) \,\mathrm{d}\lambda\right]^p \\
&\leq \left(\iint_t (F_0 \vee F_1) \mathbb{E}[|\Psi|_\beta^\alpha] \,\mathrm{d}\lambda\right)^p \\
&\leq 1 + 2^{(\alpha\vee\beta\vee 1)-1} \iint_t (F_0 \vee F_1) |G|_\beta^\alpha \left(|\sigma(0)|_0^{\alpha\vee\beta} + C_{\sigma,2}^{\alpha\vee\beta} \|\phi\|_{L^p}^{(\alpha\vee\beta)\gamma/p}\right) \,\mathrm{d}\lambda \\
&\leq 1 + 2^{(\alpha\vee\beta\vee 1)-1} \iint_t (F_0 \vee F_1) |G|_\beta^\alpha \left(|\sigma(0)|_0^{\alpha\vee\beta} + C_{\sigma,2}^{\alpha\vee\beta} + C_{\sigma,2}^{\alpha\vee\beta} \|\phi\|_{L^p}^{\rho}\right) \,\mathrm{d}\lambda.
\end{aligned}$$

Together with the estimates for J^1, J^2 and J^3, this finishes the proof of (2.6.5). □

The next lemma allows us to take good versions of the stochastic integral process (2.6.1):

Lemma 2.6.2. *For every $\phi \in \tilde{\mathcal{P}}$ there exists a predictable modification of $J(\phi)$, that is, a $(-\infty,\infty]$-valued process $\bar{J}(\phi) \in \tilde{\mathcal{P}}$ such that for each $(t,x) \in I \times \mathbb{R}^d$ we have $J(\phi)(t,x) = \bar{J}(\phi)(t,x)$ a.s.*

Proof. The set A of all $(t,x) \in I \times \mathbb{R}^d$ for which $G(t,x;\cdot,\cdot)\sigma(\phi)$ is integrable with respect to Λ is deterministic by definition, and by Theorem 1.4.1 and Fubini's theorem also measurable. It follows that there exists a measurable modification $J^{\mathrm{m}}(\phi)$ of $J(\phi)$: set $J^{\mathrm{m}}(\phi) = \infty$ on A^{C} and use [93, Thm. 1] on A. Next, define ${}^{\mathrm{p}}J(\phi)$ as the extended predictable projection of $J^{\mathrm{m}}(\phi)$ in the sense of [80, Thm. I.2.28]. By [130, Prop. 3] we may choose ${}^{\mathrm{p}}J(\phi)(t,x)$ measurably in x. And indeed, ${}^{\mathrm{p}}J(\phi)$ is still a modification of $J(\phi)$ since for each $(t,x) \in I \times \mathbb{R}^d$ we have a.s.

$$\begin{aligned} {}^{\mathrm{p}}J(\phi)(t,x) &= \mathbb{E}[J^{\mathrm{m}}(\phi)(t,x) \,|\, \mathcal{F}_{t-}] = \int_{I_t} \int_{\mathbb{R}^d} G(t,x;s,y)\sigma(\phi(s,y))\,\Lambda(\mathrm{d}s,\mathrm{d}y) \\ &= J(\phi)(t,x). \end{aligned}$$

□

We proceed with a discretization result for stochastic integrals:

Lemma 2.6.3. *Let $I \subseteq \mathbb{R}$ be an interval and $w \equiv 1$, and assume that G, σ and Λ satisfy (2)–(6) of Assumption C. Fix some $(t,x) \in I \times \mathbb{R}^d$ and assume that $G(t,x;\cdot,\cdot)$ has the following properties: for all $(s,y) \in I_t, \times \mathbb{R}^d$ we have*

$$r \uparrow s, \quad z_i \uparrow y_i \text{ for all } i = 1,\dots,d \quad \Longrightarrow \quad G(t,x;r,z) \to G(t,x;s,y), \tag{2.6.13}$$

and for some $\epsilon > 0$ the function $G^*_\epsilon(t,x;s,y) := \sup_{r \in I, s-\epsilon < r \le s, |y-z|<\epsilon} |G(t,x;r,z)|$ satisfies

$$\begin{aligned} \int_{I_t} \int_{\mathbb{R}^d} \Bigg(& |G^*_\epsilon(t,x;s,y)|^p \left(\int_{\mathbb{R}} |z|^p \,\pi(s,y,\mathrm{d}z) + c(s,y) \right) \\ & + |G^*_\epsilon(t,x;s,y) b_1(s,y)| \mathbb{1}_{\{p \ge 1\}} \Bigg) \lambda(\mathrm{d}s,\mathrm{d}y) < \infty. \end{aligned} \tag{2.6.14}$$

Moreover, we specify discretization schemes for both time and space: first, we choose for each $N \in \mathbb{N}$ a number $k(N) \in \mathbb{N} \cup \{\infty\}$ of time points $(s_i^N)_{i=1}^{k(N)} \subseteq I_t$ such that

$$s_i^N < s_{i+1}^N, \quad \text{and} \quad s_1^N \downarrow \inf I, \quad s_{k(N)}^N \uparrow t, \quad \sup_{i=1,\dots,k(N)-1} |s_{i+1}^N - s_i^N| \downarrow 0 \quad \text{as } N \uparrow \infty;$$

and second, we fix for each $N \in \mathbb{N}$ a number $l(N) \in \mathbb{N} \cup \{\infty\}$ of non-empty pairwise disjoint hyperrectangles $(Q_j^N = (a_j^N, b_j^N])_{j=1}^{l(N)} \subseteq \mathbb{R}^d$ satisfying

$$\bigcup_{j=1}^{l(N)} Q_j^N \uparrow \mathbb{R}^d \quad \text{and} \quad \sup_{j=1,\dots,l(N)} \operatorname{diam}(Q_j^N) \downarrow 0 \quad \text{as } N \uparrow \infty.$$

(1) If $\phi \in B^p_{I,\mathrm{loc}}$ is an L^p-continuous process (cf. (2.4.17)), then the stochastic integral $J(\phi)(t,x)$ is well defined and

$$\phi_0(t,x) + \sum_{i=1}^{k(N)-1} \sum_{j=1}^{l(N)} G(t;x;s_i^N,a_j^N)\sigma(\phi(s_i^N,a_j^N))\Lambda\Big((s_i^N,s_{i+1}^N]\times Q_j^N\Big) \\ \to J(\phi)(t,x) \tag{2.6.15}$$

in L^p as $N\to\infty$.

(2) The statement of (1) remains true if we replace $\uparrow$ in (2.6.13) by $\downarrow$, and at the same time replace $G(t,x;s_i^N,a_j^N)$ by $G(t,x;s_{i+1}^N,b_j^N)$ in (2.6.15).

Proof. Part (2) is proved in the same fashion as part (1). That the stochastic integral $J(\phi)(t,x)$ exists, is a consequence of Lemma 2.6.1(1), the assumptions posed on G and Λ, and the fact that $\phi\in B^p_{I,\mathrm{loc}}$. To prove (2.6.15), let us call its left-hand side $J^N(\phi)(t,x)$. It follows that

$$J^N(\phi)(t,x) = \phi_0(t,x) + \int_{I_t}\int_{\mathbb{R}^d} H^N(t,x;s,y)\,\Lambda(\mathrm{d}s,\mathrm{d}y), \quad \text{where}$$
$$H^N(t,x;s,y) = \sum_{i=1}^{k(N)-1}\sum_{j=1}^{l(N)} G(t;x;s_i^N,a_j^N)\sigma(\phi(s_i^N,a_j^N))\mathbb{1}_{(s_i^N,s_{i+1}^N]\times Q_j^N}(s,y).$$

We notice that $H^N(t,x;s,y)=0$ if the point (s,y) does not belong to the set $A^N := (s_1^N, s_{k(N)}^N]\times\bigcup_{j=1}^{l(N)} Q_j^N$, and that for each $(s,y)\in(\inf I,t)\times\mathbb{R}^d$ we have $\mathbb{1}_{(A^N)^{\mathrm{C}}}(s,y)\to 0$ as $N\to\infty$. Now, we distinguish between two cases: first, if $p<1$, or $p\geq 1$ and $\Lambda\in\mathcal{M}$, then similar calculations as done for Lemma 2.6.1(2) lead to (set $H(t,x;s,y) := G(t,x;s,y)\sigma(\phi(s,y))$)

$$\begin{aligned}
&\mathbb{E}[|J(\phi)(t,x)-J^N(\phi)(t,x)|^p] \\
&\quad\leq (C_p^{\mathrm{BDG}})^p \int_{I_t}\int_{\mathbb{R}^d} \mathbb{E}[|H(t,x;s,y)-H^N(t,x;s,y)|^p] \\
&\qquad\times\left(\int_{\mathbb{R}}|z|^p\,\pi(s,y,\mathrm{d}z)+c(s,y)\right)\lambda(\mathrm{d}s,\mathrm{d}y) \\
&\quad\leq \int_{(A^N)^{\mathrm{C}}} \frac{G^{C,1}(t,x;s,y)}{C^p_{\sigma,1}}\mathbb{E}[|\sigma(\phi(s,y))|^p]\,\lambda(\mathrm{d}s,\mathrm{d}y) \\
&\qquad+\iint_{A^N}\frac{G^{C,1}(t,x;s,y)}{C^p_{\sigma,1}}\sum_{i,j}\mathbb{E}[|\sigma(\phi(s,y))-\sigma(\phi(s_i^N,a_j^N))|^p] \\
&\qquad\times\mathbb{1}_{(s_i^N,s_{i+1}^N]\times Q_j^N}(s,y)\,\lambda(\mathrm{d}s,\mathrm{d}y)
\end{aligned}$$

$$
\begin{aligned}
&+ (C_p^{\mathrm{BDG}})^p \iint_{A^N} \sum_{i,j} |G(t,x;s,y) - G(t,x;s_i^N,a_j^N)|^p \mathbb{E}[|\sigma(\phi(s_i^N,a_j^N))|^p] \\
&\times \mathbb{1}_{(s_i^N,s_{i+1}^N]\times Q_j^N}(s,y) \left(\int_{\mathbb{R}} |z|^p \, \pi(s,y,\mathrm{d}z) + c(s,y) \right) \lambda(\mathrm{d}s,\mathrm{d}y) \\
&=: I_1^N + I_2^N + I_3^N. \qquad (2.6.16)
\end{aligned}
$$

Since $\phi \in B^p_{I,\mathrm{loc}}$ and $G^{C,1}$ is integrable with respect to λ by hypothesis, $I_1^N \to 0$ as $N \to \infty$ by dominated convergence. Next, as a consequence of the L^p-continuity of ϕ and the refining properties of our discretization scheme, the sum within I_2^N goes to 0 pointwise for each $(s,y) \in I_t \times \mathbb{R}^d$. Moreover, this sum is majorized by $2\|\sigma(\phi)\|_{B^p_{I_t}}$ such that also $I_2^N \to 0$ as $N \to \infty$. Regarding I_3^N, we obtain as an upper bound

$$
\begin{aligned}
I_3^N \le (C_p^{\mathrm{BDG}})^p \|\sigma(\phi)\|_{B^p_{I_t}} \iint_{A^N} &\left| G(t,x;s,y) - \sum_{i,j} G(t,x;s_i^N,a_j^N) \mathbb{1}_{(s_i^N,s_{i+1}^N]\times Q_j^N}(s,y) \right|^p \\
&\times \left(\int_{\mathbb{R}} |z|^p \, \pi(s,y,\mathrm{d}z) + c(s,y) \mathbb{1}_{\{p=2\}} \right) \lambda(\mathrm{d}s,\mathrm{d}y).
\end{aligned}
$$

Because of (2.6.13), the integrand in the last line goes to 0 as $N \to \infty$, pointwise for $(s,y) \in I_t \times \mathbb{R}^d$. Moreover, it is dominated by $2G_\epsilon^*$, when ϵ is chosen according to (2.6.14) and N is large enough such that $\sup_{i=1,\dots,k(N)-1} |s_{i+1}^N - s_i^N|$ and $\sup_{j=1,\dots,l(N)} \operatorname{diam}(Q_j^N)$ are smaller than ϵ. By dominated convergence, we conclude $I_3^N \to 0$ as $N \to \infty$.

It remains to discuss the case $p \ge 1$ and $\Lambda \notin \mathcal{M}$. As in Lemma 2.6.1(2), we decompose $\Lambda = M + B_1$, where M is a martingale measure and B_1 the drift measure. For M we can apply the calculations above. For B_1 we obtain an analogous decomposition as in (2.6.16): $G^{C,1}$ is replaced by $G^{C,2}$, and instead of the Burkholder-Davis-Gundy constants, the factor

$$
\left(\int_{I_t} \int_{\mathbb{R}^d} \sum_{i,j} |G(t,x;s,y) - G(t,x;s_i^N,a_j^N)| \mathbb{1}_{(s_i^N,s_{i+1}^N]\times Q_j^N}(s,y) \, |B_1|(\mathrm{d}s,\mathrm{d}y) \right)^{p-1}
$$

appears. But this also goes to 0 as $N \to \infty$, as desired. □

The next lemma concerns the solvability of deterministic integral equations and provides a comparison result. Certainly, there is a huge literature on deterministic Volterra equations, but we did not find a reference completely satisfying our purposes. Thus, we decided to include the proof, which is also very instructive for the proofs of the main theorems below.

Lemma 2.6.4. *Let $I \subseteq \mathbb{R}$ be an interval, $p \in [1,\infty)$ and λ a positive measure on $(I \times \mathbb{R}^d, \mathcal{B}(I \times \mathbb{R}^d))$. Further suppose that for every $l \in \mathbb{N}$ we have a positive measurable function $G^{(l)}\colon (I \times \mathbb{R}^d)^2 \to \mathbb{R}$ with $G^{(l)}(t,\cdot;s,\cdot) \equiv 0$ for $s > t$. Moreover, assume that there exists $k \in \mathbb{N}$ and a partition of I into pairwise disjoint intervals $I_1, \ldots, I_k$ such that*

$$\rho := \sup_{(t,x)\in I\times\mathbb{R}^d} \sup_{j=1,\ldots,k} \sum_{l=1}^{\infty} \left(\int_{I_j} \int_{\mathbb{R}^d} G^{(l)}(t,x;s,y)\, \lambda(\mathrm{d}s,\mathrm{d}y) \right)^{1/p} < 1. \tag{2.6.17}$$

Then the following statements hold:

(1) Let $(v_n)_{n\in\mathbb{N}}$ be a sequence of positive functions in L_I^∞ satisfying

$$v_{n+1}(t,x) \le \sum_{l=1}^{\infty} \left(\int_I \int_{\mathbb{R}^d} G^{(l)}(t,x;s,y)(v_n(s,y))^p\, \lambda(\mathrm{d}s,\mathrm{d}y) \right)^{1/p} \tag{2.6.18}$$

for all $n \in \mathbb{N}$. Then $\sum_{n=1}^{\infty} \|v_n\|_{L_I^\infty}$ is finite. In particular, $v_n \to 0$ in L_I^∞.

(2) For every positive $f \in L_I^\infty$ the equation

$$v(t,x) = f(t,x) + \sum_{l=1}^{\infty} \left(\int_I \int_{\mathbb{R}^d} G^{(l)}(t,x;s,y)(v(s,y))^p\, \lambda(\mathrm{d}s,\mathrm{d}y) \right)^{1/p} \tag{2.6.19}$$

for $(t,x) \in I \times \mathbb{R}^d$ has a unique solution $v \in L_I^\infty$. Furthermore, this solution v is positive.

(3) If $\bar{v} \in L_I^\infty$ is a positive function satisfying

$$\bar{v}(t,x) \le f(t,x) + \sum_{l=1}^{\infty} \left(\int_I \int_{\mathbb{R}^d} G(t,x;s,y)(\bar{v}(s,y))^p\, \lambda(\mathrm{d}s,\mathrm{d}y) \right)^{1/p} \tag{2.6.20}$$

for $(t,x) \in I \times \mathbb{R}^d$, then we have $\bar{v}(t,x) \le v(t,x)$ for all $(t,x) \in I \times \mathbb{R}^d$. In particular, if $f \equiv 0$, then $v \equiv \bar{v} \equiv 0$.

Proof. a) We start with (1). Let $I = I_1 \cup \ldots \cup I_k$ be as in the hypothesis and suppose that the intervals I_j are arranged in increasing order (i.e. $\sup I_j = \inf I_{j+1}$). Furthermore, define for $\phi \in L_I^\infty$, $(t,x) \in I \times \mathbb{R}^d$, $l \in \mathbb{N}$ and $j = 1,\ldots,k$

$$\begin{aligned} \|\phi\|_{G^{(l)},p}(t,x) &:= \left(\int_I \int_{\mathbb{R}^d} G^{(l)}(t,x;s,y)|\phi(s,y)|^p\, \lambda(\mathrm{d}s,\mathrm{d}y) \right)^{1/p} \\ \|\phi\|_{G^{(l)},p,j}(t,x) &:= \left(\int_{I_j} \int_{\mathbb{R}^d} G^{(l)}(t,x;s,y)|\phi(s,y)|^p\, \lambda(\mathrm{d}s,\mathrm{d}y) \right)^{1/p}. \end{aligned} \tag{2.6.21}$$

Obviously, we have $\|\phi\|_{G^{(l)},p}(t,x) \le \sum_{j=1}^{k} \|\phi\|_{G^{(l)},p,j}(t,x)$ for each $(t,x) \in I \times \mathbb{R}^d$ and $l \in \mathbb{N}$. Hence, it follows from (2.6.18) that

$$v_{n+1} \le \sum_{l=1}^{\infty} \|v_n\|_{G^{(l)},p} \le \sum_{j=1}^{k} \sum_{l=1}^{\infty} \|v_n\|_{G^{(l)},p,j}, \tag{2.6.22}$$

an equation that holds pointwise for all $(t,x) \in I \times \mathbb{R}^d$ and for all $n \in \mathbb{N}$. Iterating (2.6.22) n times, together with the subadditivity of the functional $\|\cdot\|_{G^{(l)},p,j}$, yields

$$\begin{aligned} v_{n+1} &\le \sum_{j_1=1}^{k} \sum_{l_1=1}^{\infty} \|v_n\|_{G^{(l_1)},p,j_1} \le \sum_{j_1,j_2=1}^{k} \sum_{l_1=1}^{\infty} \left\| \sum_{l_2=1}^{\infty} \|v_{n-1}\|_{G^{(l_2)},p,j_2} \right\|_{G^{(l_1)},p,j_1} \le \dots \\ &\le \sum_{j_1,\dots,j_n=1}^{k} \sum_{l_1=1}^{\infty} \left\| \sum_{l_2=1}^{\infty} \right\| \dots \sum_{l_n=1}^{\infty} \|v_1\|_{G^{(l_n)},p,j_n} \dots \Big\|_{G^{(l_2)},p,j_2} \Big\|_{G^{(l_1)},p,j_1}. \end{aligned} \tag{2.6.23}$$

Observe that the Volterra property of G implies that on the right-hand side of (2.6.23), only those summands are non-zero for which $j_1 \ge \dots \ge j_n$. Since there are exactly $\binom{n+k-1}{n}$ such sequences, and $\sup_{j=1,\dots,k} \left\| \sum_{l=1}^{\infty} \|1\|_{G^{(l)},p,j} \right\|_{L_I^\infty} = \rho$, we deduce that

$$\sum_{n=1}^{\infty} \|v_n\|_{L_I^\infty} \le \|v_1\|_{L_I^\infty} \sum_{n=0}^{\infty} \binom{n+k-1}{n} \rho^n < \infty \tag{2.6.24}$$

by the ratio test and the fact that $\rho < 1$.

b) Next we prove (2) and construct a solution to (2.6.19) by Picard iteration. Define $v^0(t,x) = f(t,x)$ and for $n \in \mathbb{N}$ and $(t,x) \in I \times \mathbb{R}^d$

$$v^n(t,x) := f(t,x) + \sum_{l=1}^{\infty} \left(\int_I \int_{\mathbb{R}^d} G^{(l)}(t,x;s,y) (v^{n-1}(s,y))^p \, \lambda(\mathrm{d}s, \mathrm{d}y) \right)^{1/p}. \tag{2.6.25}$$

Since G satisfies (2.6.17), f belongs to L_I^∞, and both functions are positive, v^n is by induction again a positive function in L_I^∞. Now form the difference sequence $u^n := |v^{n+1} - v^n|$ for $n \in \mathbb{N}$, which satisfies property (2.6.18) by the reverse triangle inequality. By (1), $\sum_{n=1}^{\infty} \|u^n\|_{L_I^\infty} < \infty$, in other words, v as the limit in L_I^∞ of v^n exists. Of course, v is positive. Moreover, taking the limit on both sides of (2.6.25), we conclude that v indeed satisfies (2.6.19). The uniqueness part follows by applying part (1) to the difference of two solutions in L_I^∞.

c) For $\phi \in L_I^\infty$ set $I_f(\phi) := f + \sum_{l=1}^{\infty} \|\phi\|_{G^{(l)},p}$, which again belongs to L_I^∞. By (2.6.25), we have $v^n = I_f^{(n)}(f)$, which is the n-fold iteration $I_f(I_f(\dots I_f(f)\dots))$.

Moreover, by (2.6.20),

$$\bar{v} \le I_f(\bar{v}) \le I_f\Big(I_f(\bar{v})\Big) \le \ldots \le I_f^{(n)}(\bar{v}) \le I_f^{(n-1)}(f) + I_0^{(n)}(\bar{v}) = v^{n-1} + I_0^{(n)}(\bar{v}).$$

As shown in a), v^{n-1} converges to v uniformly on $I \times \mathbb{R}^d$. In addition, $I_0^{(n)}(\bar{v})$ is less or equal to the right-hand side of (2.6.23) when v_1 is replaced by $\bar{v}$. Thus, the considerations in a) show that $I_0^{(n)}(\bar{v}) \le \|\bar{v}\|_{L_I^\infty}\binom{n+k-1}{n}\rho^n \to 0$ as $n \to \infty$, which implies (3). □

The next lemma concerns the asymptotic behaviour of deterministic Volterra equations with a fractional nonlinearity:

Lemma 2.6.5. *Let I, p and $G^{(l)}$ be as in Lemma 2.6.4. Further suppose that $f \in L_I^\infty$ is a positive function and*

$$\theta := \sup_{(t,x)\in I\times\mathbb{R}^d} \sum_{l=1}^{\infty} \Big(\int_I \int_{\mathbb{R}^d} G^{(l)}(t,x;s,y)\, \lambda(\mathrm{d}s,\mathrm{d}y) \Big)^{1/p} < \infty.$$

Moreover, we assume that $v \in L_{I,\mathrm{loc}}^\infty$ is positive and satisfies

$$v(t,x) \le f(t,x) + \left(\sum_{l=1}^{\infty} \int_I \int_{\mathbb{R}^d} G^{(l)}(t,x;s,y)(v(s,y))^{p\gamma}\, \lambda(\mathrm{d}s,\mathrm{d}y) \right)^{1/p} \tag{2.6.26}$$

for $(t,x) \in I \times \mathbb{R}^d$ and some $\gamma \in (0,1)$. Then $v \in L_I^\infty$ with $\|v\|_{L_I^\infty} \le a$, where a is the unique strictly positive solution to the equation $a - \|f\|_{L_I^\infty} - \theta a^\gamma = 0$.

Proof. The proof is a straightforward generalization of the arguments given in Example 2.5.1. We include it for the sake of completeness. Fix $T \in I$ and recall the definition of $\|\cdot\|_{G^{(l)},p}$ and $I_f(\cdot)$ from the proof of Lemma 2.6.4. By (2.6.26), it follows that

$$\|v\|_{L_{I_T}^\infty} \le \|I_f(v^\gamma)\|_{L_{I_T}^\infty} \le I_{\|f\|_{L_I^\infty}}(\|v\|_{L_{I_T}^\infty}^\gamma).$$

By iteration of the last inequality, we deduce that $\|v\|_{L_{I_T}^\infty} \le a_n(T)$ for all $n \in \mathbb{N}$ where $a_1(T) := \|v\|_{L_{I_T}^\infty}$ and $a_{n+1}(T) = I_{\|f\|_{L_I^\infty}}((a_n(T))^\gamma) = \|f\|_{L_I^\infty} + \theta(a_n(T))^\gamma$ for $n \in \mathbb{N}$. Straightforward analysis reveals that $\limsup_{n\to\infty} a_n(T) \le a$, a number independent of T. Hence, $\|v\|_{L_I^\infty} \le a$. □

2.7 Proof of the main theorems

Proof of Theorem 2.3.1. We show that Theorem 2.3.1 is a special case of Theorem 2.4.4, or more precisely, that Assumption A is contained in Assumption C: setting $I = \mathbb{R}_+$ and $w \equiv 1$ in Assumption C, it is not hard to see that the first six conditions break down to conditions (1)–(6) of Assumption A, and that condition (7) of Assumption C becomes superfluous. The only thing to show is that (2.3.4) implies (2.4.10). To this end, fix $T \in \mathbb{R}_+$, define ϵ as

$$2^{-p^*}\left[(C_{\sigma,1}C_p^{\mathrm{BDG}})^p + C_{\sigma,1}^p\left(\sup_{(t,x)\in[0,T]\times\mathbb{R}^d}\int_0^t\int_{\mathbb{R}^d}|G(t,x;s,y)b_1(s,y)|\,\lambda(\mathrm{d}s,\mathrm{d}y)\right)^{p-1}\right]^{-1},$$

and let $\mathcal{T}$ be a subdivision of $[0,T]$ such that (2.3.4) holds. Then we have for all $(t,x)\in[0,T]\times\mathbb{R}^d$ and $i=0,\dots,k$ that

$$\begin{aligned}\sum_{l=1}^{2}&\left(\int_{t_i}^{t_{i+1}}\int_{\mathbb{R}^d}G^{C,l}(t,x;s,y)\,\lambda(\mathrm{d}s,\mathrm{d}y)\right)^{1/p^*}\\ &\le 2\left(\int_{t_i}^{t_{i+1}}\int_{\mathbb{R}^d}G^{C,1}(t,x;s,y)+G^{C,2}(t,x;s,y)\,\lambda(\mathrm{d}s,\mathrm{d}y)\right)^{1/p^*}\\ &\le \epsilon^{-1/p^*}\left(\int_{t_i}^{t_{i+1}}\int_{\mathbb{R}^d}G^{A}(t,x;s,y)\,\lambda(\mathrm{d}s,\mathrm{d}y)\right)^{1/p^*}<1,\end{aligned}$$

which is (2.4.10). □

Proof of Corollary 2.3.2. We check the conditions of Assumption A. (1), (2) and (3) are also assumed in the corollary. Regarding (4), (5) and (6), it is easy to see that because of (2.3.5), conditions (2.3.1), (2.3.2) and (2.3.3) split into separated conditions for both G and Λ, which are fulfilled thanks to (2.3.6) and (2.3.7), respectively. Only (7) if left to be verified. Let $T\in\mathbb{R}_+$ be arbitrary and define $t_n^i := i/n^2$ for $n\in\mathbb{N}$ and $i=0,\dots,Tn^2$. Then, using the notation

$$g^A := \left(\int_{\mathbb{R}}|z|^p\,\pi_0(\mathrm{d}z)+\|c\|_{L^\infty_{[0,T]}}\right)g^p + \|b_1\|_{L^\infty_{[0,T]}}g\mathbb{1}_{\{p\ge1\}}, \tag{2.7.1}$$

we have for all $(t,x)\in[0,T]\times\mathbb{R}^d$

$$\begin{aligned}\int_{t_n^i}^{t_n^{i+1}}\int_{\mathbb{R}^d}G^A(t,x;s,y)\,\mathrm{d}(s,y) &\le \int_{t_n^i}^{t_n^{i+1}}\int_{\mathbb{R}^d}g^A(t-s,x-y)\,\mathrm{d}(s,y)\\ &\le \int_{(t-t_n^{i+1})\vee0}^{(t-t_n^i)\vee0}\int_{\mathbb{R}^d}g^A(s,y)\,\mathrm{d}(s,y).\end{aligned}$$

The right-hand side becomes arbitrarily small as $n \to \infty$, uniformly for (t,x) in $[0,T] \times \mathbb{R}^d$ and $i = 0, \dots, Tn^2 - 1$. If not, there would exist some $\epsilon > 0$ as well as for each $n \in \mathbb{N}$ some $\tau_n \in [0,T]$ and $i(n) \in \{0, \dots, Tn^2 - 1\}$ such that

$$\int_{(\tau_n - t_n^{i(n)+1}) \vee 0}^{(\tau_n - t_n^{i(n)}) \vee 0} \int_{\mathbb{R}^d} g^A(s,y) \, \mathrm{d}(s,y) \geq \epsilon.$$

This, however, would contradict the dominated convergence theorem and the Borel-Cantelli lemma since $|((\tau_n - t_n^{i(n)}) \vee 0) - ((\tau_n - t_n^{i(n)+1}) \vee 0)| \leq |t_n^{i(n)+1} - t_n^{i(n)}| = 1/n^2$. Thus, Corollary 2.3.2 is proved. □

Proof of Theorem 2.3.5. a) We first prove the existence of a solution to (2.1.1). To this end, define

$$T_n := \inf\{t > 0 \colon |\Lambda(\{t\} \times \mathbb{R}^d)| > n\}, \quad n \in \mathbb{N}.$$

Assumption B(3) implies that $(T_n)_{n \in \mathbb{N}}$ is a sequence of stopping times such that we have $T_n > 0$ a.s. for each $n \in \mathbb{N}$ and $T_n \uparrow +\infty$ a.s. as $n \to \infty$. Next, we introduce for each $n \in \mathbb{N}$ a truncation of Λ in the following sense:

$$\begin{aligned} \Lambda^n(\mathrm{d}t, \mathrm{d}x) := B(\mathrm{d}t, \mathrm{d}x) + \Lambda^{\mathrm{c}}(\mathrm{d}t, \mathrm{d}x) + \int_{\mathbb{R}} z \mathbb{1}_{\{|z| \leq 1\}} \, (\mu - \nu)(\mathrm{d}t, \mathrm{d}x, \mathrm{d}z) \\ + \int_{\mathbb{R}} z \mathbb{1}_{\{1 < |z| \leq n\}} \, \mu(\mathrm{d}t, \mathrm{d}x, \mathrm{d}z). \end{aligned}$$

By Assumption B(4) we may assume without loss of generality that $Y_0 \in B^q_{[0,\infty),\mathrm{loc}}$. Consequently, thanks to Assumption B(1) and (2) and Theorem 2.3.1, Equation (2.1.1) with Λ^n as driving noise has a unique solution $Y^n \in B^q_{[0,\infty),\mathrm{loc}}$. We claim that $Y := Y^1 \mathbb{1}_{[\![0,T_1]\!]} + \sum_{n=2}^{\infty} Y^n \mathbb{1}_{]\!]T_{n-1},T_n]\!]}$ is a solution to the original equation (2.1.1) with Λ. The predictability of Y is clear. Now fix a (non-random) time $T \in \mathbb{R}_+$ and define

$$\Omega_T^n := \left\{ \omega \in \Omega \colon \sup_{(t,x) \in [0,T] \times \mathbb{R}^d} |\Lambda(\{(t,x)\})(\omega)| \in [0,n] \right\}, \quad n \in \mathbb{N}.$$

By Assumption B(3) the sequence $(\Omega_T^n)_{n \in \mathbb{N}}$ increases to Ω up to a $\mathbb{P}$-null set. Moreover, we have $\mathbb{1}_{[\![0,T_k]\!]}(t) Y^k(t,x) = \mathbb{1}_{[\![0,T_k]\!]}(t) Y^n(t,x)$ a.s. for all $n \in \mathbb{N}$ and $k = 1, \dots, n$ as a consequence of the uniqueness statement of Theorem 2.3.1 and the fact that $\mathbb{P}[T_k = t] = 0$. Now part (1) of Theorem 2.3.5 follows from the observation that for

all $(t,x) \in [0,T] \times \mathbb{R}^d$ and $n \in \mathbb{N}$ we have a.s.

$$\begin{aligned}
&\mathbb{1}_{\Omega_T^n} \int_0^t \int_{\mathbb{R}^d} G(t,x;s,y)\sigma(Y(s,y))\,\Lambda(\mathrm{d}s,\mathrm{d}y) \\
&\quad = \mathbb{1}_{\Omega_T^n} \int_0^t \int_{\mathbb{R}^d} G(t,x;s,y) \Bigg(\sigma(Y^1(s,y))\mathbb{1}_{[\![0,T_1]\!](s)} \\
&\qquad + \sum_{k=2}^n \sigma(Y^k(s,y))\mathbb{1}_{]\!]T_{k-1},T_k]\!]}(s) \Bigg) \Lambda^n(\mathrm{d}s,\mathrm{d}y) \\
&\quad = \mathbb{1}_{\Omega_T^n} \int_0^t \int_{\mathbb{R}^d} G(t,x;s,y)\sigma(Y^n(s,y))\,\Lambda^n(\mathrm{d}s,\mathrm{d}y) = \mathbb{1}_{\Omega_T^n} Y^n(t,x) = \mathbb{1}_{\Omega_T^n} Y(t,x).
\end{aligned}$$

To be utterly precise, for the transition from the second to the third line to be true, we must show that $J(\phi)$ and $J(\phi')$ as defined in (2.6.1) are modifications of each other as soon as ϕ and ϕ' are. But this follows from (2.6.3). Finally, the uniqueness statement follows from that of Theorem 2.3.1 by localization.

b) Next, we verify that the solution Y found in a) belongs to $B^p_{[0,\infty),\mathrm{loc}}$ if also (5)–(10) of Assumption B are valid. We only carry out the proof for $p \geq 1$. The case $p < 1$ can be proved in the same fashion. Let $T \in \mathbb{R}_+$ and observe from a) that Y equals Y^n on Ω_T^n. Define $v^n(t,x) := \|Y^n(t,x)\|_{L^p}$ for $(t,x) \in [0,T] \times \mathbb{R}^d$, which is always finite because $Y^n \in B^q_{[0,\infty),\mathrm{loc}}$. Moreover, if we define $G^{D,l}$ as in (2.5.7) with $w \equiv 1$, then we have for all $(t,x) \in [0,T] \times \mathbb{R}^d$ according to Lemma 2.6.1(3) with $\rho = 1$

$$\begin{aligned}
&\|Y(t,x)\mathbb{1}_{\Omega_T^n}\|_{L^p} \leq v^n(t,x) \\
&\leq f(t,x) + \sum_{l=1}^4 C_{\sigma,2} \left(\int_0^t \int_{\mathbb{R}^d} G^{D,l}(t,x;s,y)(v^n(s,y))^p\,\lambda(\mathrm{d}s,\mathrm{d}y) \right)^{1/p}, \qquad (2.7.2)
\end{aligned}$$

where f is the sum of the first three summands on the right-hand side of (2.6.4). A priori, $G^{D,l}$ may depend on n since it involves the underlying underlying Lévy measure ν^n. However, it is obvious that inequality (2.6.4) remains true if we use the original Lévy measure ν to form $G^{D,l}$: the right-hand side of (2.7.2) will only be enlarged. In this case, (2.7.2) falls into the category of Lemma 2.6.4(3). Indeed, Assumption B(10) guarantees that $f \in L^\infty_{[0,T]}$, and that the key assumption (2.6.17) is met (note that the different constants appearing in $G^{D,l}$ compared to G^B are irrelevant because G^B satisfies the partition property (2.3.4) for all $\epsilon > 0$). Thus, we have $v^n(t,x) \leq v(t,x)$ where $v \in L^\infty_{[0,T]}$ is again independent of n and is the solution

of the corresponding Volterra equation if we replace the second inequality sign in (2.7.2) by equality. Taking the limit $n \to \infty$, we conclude

$$\|Y(t,x)\|_{L^p} = \lim_{n\to\infty} \|Y(t,x)\mathbb{1}_{\Omega_T^n}\|_{L^p} \le v(t,x),$$

that is, $Y \in B^p_{[0,\infty),\mathrm{loc}}$. □

Proof of Corollary 2.3.7. a) We begin with the first statement, for which we need to verify (2) and (3) of Assumption B. That (2) holds, follows from the proof of Corollary 2.3.2, where we have shown that (2.3.5), (2.3.6) and (2.3.7) imply the validity of Assumption A(4)–(7). Notice that in the quasi-stationary case, it suffices to check Assumption B(2) only for $n = 1$ because $\int_{1<|z|\le n} |z|^q \, \pi_0(\mathrm{d}z)$ is always finite and condition (3) of Corollary 2.3.7 is in force. That (3) of Assumption B holds, is due to (2.3.17):

$$\nu\Big([0,T] \times \mathbb{R}^d \times [-1,1]^{\mathrm{C}}\Big) \le \int_0^T \int_{\mathbb{R}^d} \pi_1(t,x) \, \mathrm{d}(t,x) \pi_0(|z| > 1) < \infty.$$

b) For the second part we must prove (5)–(10) of Assumption B. (5) and (6) hold by hypothesis. Furthermore, since $p < q$ implies $|ab|_q^p \le |a|_q^p |b|_p^q$ for all $a, b \in \mathbb{R}$, we have by (2.3.5)

$$\begin{aligned}\sup_{(t,x)\in[0,T]\times\mathbb{R}^d} \int_0^t \int_{\mathbb{R}^d} \int_{\mathbb{R}} |G(t,x;s,y)z|_q^p \, \nu(\mathrm{d}s, \mathrm{d}y, \mathrm{d}z) \\ \le \|\pi_1\|_{L^\infty_{[0,T]}} \int_{\mathbb{R}} |z|_q^p \, \pi_0(\mathrm{d}z) \int_0^T \int_{\mathbb{R}^d} |g(t,x)|_p^q \, \mathrm{d}(t,x) < \infty,\end{aligned}$$

which implies (7) of Assumption B. Next, (8) is a direct consequence of condition (3) of the corollary. For (9) we choose $\alpha = q$ and $\beta = p$, which clearly satisfy (9c). For (9a) and (9b) first observe that

$$\begin{aligned}\left|b(t,x) + \int_{\mathbb{R}} z\mathbb{1}_{\{|z|\in(1,A]\}} \, \pi(t,x,\mathrm{d}z)\right| &\le \|b\|_{L^\infty_{[0,T]}} + \|\pi_1\|_{L^\infty_{[0,T]}} \int_{|z|>1} |z|^p \, \pi_0(\mathrm{d}z) A^{1-p} \\ &\le F_1 A^{1-p}\end{aligned}$$

holds for all $A \in [1,\infty)$ if $F_1 \in \mathbb{R}_+$ is chosen large enough. Second, if $q < 1$, we have $b_0 \equiv 0$ by (2.3.6), which means that

$$\begin{aligned}\left|b(t,x) - \int_{\mathbb{R}} z\mathbb{1}_{\{|z|\in(a,1]\}} \, \pi(t,x,\mathrm{d}z)\right| &= \left|\int_{\mathbb{R}} z\mathbb{1}_{\{|z|\in(0,a]\}} \, \pi(t,x,\mathrm{d}z)\right| \\ &\le \|\pi_1\|_{L^\infty_{[0,T]}} \int_{|z|\le 1} |z|^q \, \pi_0(\mathrm{d}z) a^{1-q}.\end{aligned}$$

Finally, if $q \geq 1$ we have

$$\left| b(t,x) - \int_{\mathbb{R}} z \mathbb{1}_{\{|z| \in (a,1]\}} \, \pi(t,x,\mathrm{d}z) \right| \leq \|b\|_{L^\infty_{[0,T]}} + \|\pi_1\|_{L^\infty_{[0,T]}} \int_{|z| \leq 1} |z|^q \, \pi_0(\mathrm{d}z) a^{1-q}$$
$$\leq F_0 a^{1-q}$$

for all $a \in (0,1]$ and some constant $F_0 \in \mathbb{R}_+$. Finally, condition (10) holds by the same arguments used in the proof of Corollary 2.3.2. □

Proof of Theorem 2.4.4. We base the proof on a Picard iteration scheme, which parallels the construction of a solution to (2.6.19) in Lemma 2.6.4. We define processes $Y^n \in \tilde{\mathcal{P}}$ inductively as follows: starting with $Y^0(t,x) := Y_0(t,x)$, we assume that $Y^{n-1} \in B^{p,w}_{I,\mathrm{loc}}$ has already been constructed for some $n \in \mathbb{N}$. Define for each $(t,x) \in I \times \mathbb{R}^d$

$$Y^n(t,x) := Y_0(t,x) + \int_I \int_{\mathbb{R}^d} G(t,x;s,y) \sigma(Y^{n-1}(s,y)) \, \Lambda(\mathrm{d}s, \mathrm{d}y), \tag{2.7.3}$$

hereby choosing a predictable version of Y^n, cf. Lemma 2.6.2. Let $T \in I$. Then we have by Lemma 2.6.1(1) for all $(t,x) \in I_T \times \mathbb{R}^d$

$$\frac{\|Y^n(t,x)\|_{L^p}}{(w(t,x))^{1/p^*}} \leq \frac{\|Y_0(t,x)\|_{L^p}}{(w(t,x))^{1/p^*}} + \sum_{l=1}^{2} \left(\int_I \int_{\mathbb{R}^d} \frac{G^{C,l}(t,x;s,y)}{C^p_{\sigma,1}} \right.$$
$$\left. \times \left(\frac{|\sigma(0)|^{p \wedge 1} + C^{p \wedge 1}_{\sigma,1} \|Y^{n-1}(s,y)\|_{L^p}}{(w(s,y))^{1/p^*}} \right)^{p^*} \lambda(\mathrm{d}s, \mathrm{d}y) \right)^{1/p^*},$$

which is finite by Assumption C. Thus, $Y^n \in B^{p,w}_{I,\mathrm{loc}}$ for all $n \in \mathbb{N}$. Next, Lemma 2.6.1(2) implies that $u^n := Y^n - Y^{n-1}$ satisfies

$$\frac{\|u^{n+1}(t,x)\|_{L^p}}{(w(t,x))^{1/p^*}} \leq \sum_{l=1}^{2} \left(\int_I \int_{\mathbb{R}^d} G^{C,l}(t,x;s,y) \left(\frac{\|u^n(s,y)\|_{L^p}}{(w(s,y))^{1/p^*}} \right)^{p^*} \lambda(\mathrm{d}s, \mathrm{d}y) \right)^{1/p^*} \tag{2.7.4}$$

for all $(t,x) \in I \times \mathbb{R}^d$, which is a recursive relation as in Lemma 2.6.4(1). Note that the key hypothesis (2.6.17) is fulfilled because of Assumption C(8). We conclude that $\sum_{n=1}^{\infty} \|u^n\|_{B^{p,w}_{I_T}} < \infty$, in other words, Y^n converges in $B^{p,w}_{I_T}$ to some limit Y. Applying Lemma 2.6.1(2) to $\phi_1 := Y$ and $\phi_2 := Y^{n-1}$, the convergence $Y^{n-1} \to Y$ also implies that $J(Y^{n-1}) = Y^n \to J(Y)$ in $B^{p,w}_{I_T}$, that is, Y indeed satisfies (2.1.1). The uniqueness of the solution to (2.1.1) follows if we substitute u^n in (2.7.4) by the difference of two solutions. Since $T \in I$ is arbitrary, Theorem 2.4.4 follows. □

Proof of Corollary 2.4.5. We verify Assumption C for $I = \mathbb{R}$ and $w \equiv 1$. (1), (2) and (3) hold by hypothesis; (4), (5) and (6) are consequences of (2.4.11), (2.4.12), (2.4.13) and (2.4.14). Moreover, condition (7) of Assumption C is redundant such that it remains to verify (8). To this end, define

$$\begin{aligned} g^{C,1} &:= (C_{\sigma,1} C_p^{\mathrm{BDG}})^p (\zeta_p + \|c\|_{L^\infty_{\mathbb{R}}}) g^p, \\ g^{C,2} &:= (C_{\sigma,1} \|b_1\|_{L^\infty_{\mathbb{R}}})^p \left(\int_0^\infty \int_{\mathbb{R}^d} g(t,x)\, \mathrm{d}(t,x) \right)^{p-1} g \mathbb{1}_{\{p \geq 1\}}. \end{aligned}$$

Then, for any subdivision $\mathcal{T}\colon\ -\infty = t_0 < \ldots < t_{k+1} = T$, all $(t,x) \in (-\infty, T] \times \mathbb{R}^d$ and $i = 0, \ldots, k$, we have by (2.4.13) and (2.4.14)

$$\begin{aligned} \sum_{l=1}^2 \left(\int_{t_i}^{t_{i+1}} \int_{\mathbb{R}^d} G^{C,l}(t,x;s,y)\, \mathrm{d}(s,y) \right)^{1/p^*} &\leq \sum_{l=1}^2 \left(\int_{-\infty}^{t} \int_{\mathbb{R}^d} g^{C,l}(t-s, x-y)\, \mathrm{d}(s,y) \right)^{1/p^*} \\ &= \sum_{l=1}^2 \left(\int_0^\infty \int_{\mathbb{R}^d} g^{C,l}(t,x)\, \mathrm{d}(t,x) \right)^{1/p^*} < 1. \end{aligned}$$

□

Proof of Theorem 2.4.7. a) Fix $T \in I$ and choose $(t,x), (\tau,\xi) \in I_T \times \mathbb{R}^d$. Then similar calculations as in Lemma 2.6.1(2) lead to

$$\begin{aligned} &\|Y(t,x) - Y(\tau,\xi)\|_{L^p} \\ &\quad \leq \sum_{l=1}^2 \left(\int_I \int_{\mathbb{R}^d} \tilde{G}^{(l)}(t,x;\tau,\xi;s,y) \left(\frac{\|\sigma(Y(s,y))\|_{L^p}}{(w(s,y))^{1/p^*}} \right)^{p^*} \lambda(\mathrm{d}s, \mathrm{d}y) \right)^{1/p^*} \\ &\quad \leq \|\sigma(Y)\|_{B^{p,w}_{I_T}} \sum_{l=1}^2 \left(\int_I \int_{\mathbb{R}^d} \tilde{G}^{(l)}(t,x;\tau,\xi;s,y)\, \lambda(\mathrm{d}s, \mathrm{d}y) \right)^{1/p^*}, \end{aligned}$$

where

$$\begin{aligned} \tilde{G}^{(1)}(t,x;\tau,\xi;s,y) &:= (C_p^{\mathrm{BDG}})^p |G(t,x;s,y) - G(\tau,\xi;s,y)|^p \\ &\quad \times \left(\int_{\mathbb{R}} |z|^p\, \pi(s,y,\mathrm{d}z) + c(s,y) \right) w(s,y), \\ \tilde{G}^{(2)}(t,x;\tau,\xi;s,y) &:= \left(\int_I \int_{\mathbb{R}^d} |[G(t,x;s,y) - G(\tau,\xi;s,y)] b_1(s,y)|\, \lambda(\mathrm{d}s, \mathrm{d}y) \right)^{p-1} \\ &\quad \times |[G(t,x;s,y) - G(\tau,\xi;s,y)] b_1(s,y)| w(s,y) \mathbb{1}_{\{p \geq 1\}}. \end{aligned}$$

The claim now follows from (2.4.16) because Assumption C(7) implies

$$\sup_{(t,x),(\tau,\xi)\in I_T\times\mathbb{R}^d}\left(\int_I\int_{\mathbb{R}^d}|[G(t,x;s,y)-G(\tau,\xi;s,y)]b_1(s,y)|\,\lambda(\mathrm{d}s,\mathrm{d}y)\right)^{p-1}$$
$$\leq 2\sup_{(t,x)\in I_T\times\mathbb{R}^d}\left(\int_I\int_{\mathbb{R}^d}|G(t,x;s,y)b_1(s,y)|\,\lambda(\mathrm{d}s,\mathrm{d}y)\right)^{p-1}<\infty.$$

b) In the situation of Corollary 2.4.5 with G in convolution form, we have

$$\begin{aligned}
&\int_{\mathbb{R}}\int_{\mathbb{R}^d}\tilde{G}(t,x;\tau,\xi;s,y)\,\mathrm{d}(s,y)\\
&\quad\leq(\zeta_p+\|c\|_{L^\infty_{\mathbb{R}}})\int_{\mathbb{R}}\int_{\mathbb{R}^d}\left|g(t-s,x-y)-g(\tau-s,\xi-y)\right|^p\mathrm{d}(s,y)\\
&\qquad+\|b_1\|_{L^\infty_{\mathbb{R}}}\mathbb{1}_{\{p\geq1\}}\int_{\mathbb{R}}\int_{\mathbb{R}^d}|g(t-s,x-y)-g(\tau-s,\xi-y)|\,\mathrm{d}(s,y)\\
&\quad=(\zeta_p+\|c\|_{L^\infty_{\mathbb{R}}})\int_{\mathbb{R}}\int_{\mathbb{R}^d}\left|g(s+h,y+\eta)-g(s,y)\right|^p\mathrm{d}(s,y)\\
&\qquad+\|b_1\|_{L^\infty_{\mathbb{R}}}\mathbb{1}_{\{p\geq1\}}\int_{\mathbb{R}}\int_{\mathbb{R}^d}|g(s+h,y+\eta)-g(s,y)|\,\mathrm{d}(s,y)\to0
\end{aligned}$$

because $(h,\eta)=(|t-\tau|,|x-\xi|)\to0$, cf. [65, Lemma 0.12].

c) Let $T\in I$ and define $v(t,x):=w^{-1/p^*}(t,x)\|Y(t,x)-Y'(t,x)\|_{L^p}$ as well as $v_0(t,x):=w^{-1/p^*}(t,x)\|Y_0(t,x)-Y_0'(t,x)\|_{L^p}$. Furthermore, choose $k\in\mathbb{N}$ and a partition $I_T=I_1\cup\ldots\cup I_k$ such that (2.4.10) is satisfied. Next, recall from (2.6.21) the definition of $\|\phi\|_{G^{(l)},p^*}(t,x)$ and $\|\phi\|_{G^{(l)},p^*,j}(t,x)$ for $(t,x)\in I_T\times\mathbb{R}^d$, $l=1,2$ and $j=1,\ldots,k$. From Lemma 2.6.1(2) we deduce

$$v\leq v_0+\sum_{l=1}^{2}\|v\|_{G^{(l)},p^*}\leq v_0+\sum_{j=1}^{k}\sum_{l=1}^{2}\|v\|_{G^{(l)},p^*,j}. \tag{2.7.5}$$

By the same arguments as in the proof of Lemma 2.6.4(1), iterating (2.7.5) N times produces

$$\|v\|_{L^\infty_{I_T}}\leq\|v_0\|_{L^\infty_{I_T}}\sum_{n=0}^{N-1}\binom{n+k-1}{n}\rho^n+\|v\|_{L^\infty_{I_T}}\binom{N+k-1}{N}\rho^N,$$

with $\rho<1$ being the left-hand side of (2.4.10). Letting $N\to\infty$ leads to the assertion. □

Proof of Theorem 2.4.8. It suffices to prove the case where (2.4.19) holds. Since $Y\in B^p_{\mathbb{R},\mathrm{loc}}$ is constructed as the limit of the Picard iterates Y^n in (2.7.3),

it suffices to prove that Y^n, Y_0 and Λ are jointly stationary for all $n \in \mathbb{N}$. By induction, we assume that Y^{n-1} is jointly stationary with Λ and Y_0 (that Y_0 is, holds by assumption). First, we assume that g is bounded and has compact support in $\mathbb{R}_+ \times \mathbb{R}^d$, which obviously implies that (2.6.14) holds for arbitrary $\epsilon > 0$. Moreover, Y^{n-1} is L^p-continuous because Y^0 is by hypothesis and thus also Y^{n-1} for general n by the same arguments as in the proof of Theorem 2.4.7(2). Next, we fix $(t,x),(h,\eta) \in \mathbb{R} \times \mathbb{R}^d$ and define for $N \in \mathbb{N}$ and $i = 0,\dots,N^2$ the time points $s_i^N := t - N + i/N$. Moreover, we set $Q_N := \big(0,(1/N,\dots,1/N)\big]$ and $\Gamma_N := \Big\{(i_1/N,\dots,i_d/N)\colon i_1,\dots,i_d \in \{-N^2,\dots,N^2\}\Big\}$. Lemma 2.6.3 now gives

$$
\begin{aligned}
&Y^n(t+h,x+\eta)\\
&= Y_0(t+h,x+\eta) + \int_{-\infty}^{t+h}\int_{\mathbb{R}^d} g(t+h-s,x+\eta-y)\sigma(Y^{n-1}(s,y)\,\Lambda(\mathrm{d}s,\mathrm{d}y)\\
&= Y_0(t+h,x+\eta) + \int_{-\infty}^{t}\int_{\mathbb{R}^d} g(t-s,x-y)\sigma(Y^{n-1}(s+h,y+\eta)\,\Lambda(h+\mathrm{d}s,\eta+\mathrm{d}y)\\
&= Y_0(t+h,x+\eta) + L^p\text{-}\lim_{N\to\infty}\sum_{i=0}^{N^2-1}\sum_{y_j^N\in\Gamma_N} g(t-s_i^N,x-y_j^N)\sigma(Y^{n-1}(s_i^N+h,y_j^N+\eta))\\
&\quad\times\Lambda\Big((s_i^N+h,s_{i+1}^N+h)\times(y_j^N+\eta+Q_N)\Big)\\
&\overset{\mathrm{d}}{=} Y_0(t,x) + L^p\text{-}\lim_{N\to\infty}\sum_{i=0}^{N^2-1}\sum_{y_j^N\in\Gamma_N} g(t-s_i^N,x-y_j^N)\sigma(Y^{n-1}(s_i^N,y_j^N))\\
&\quad\times\Lambda\Big((s_i^N,s_{i+1}^N)\times(y_j^N+Q_N)\Big)\\
&= Y^n(t,x).
\end{aligned}
$$

The calculation remains valid when we consider joint distributions with Y_0 and Λ, and when we extend it to n space–time points. So the theorem is proved for bounded functions g with compact support. For general functions g we notice that property (2.4.19) implies that we can write $g = \sum_{i=1}^{\infty} g_i$ where each g_i is bounded with compact support. The theorem follows since the calculation above is invariant under summation and taking limits. □

Proof of Theorem 2.5.2. Let $Y \in B^{p,w}_{I,\mathrm{loc}}$ be a solution to (2.1.1). Then we have $v \in L^\infty_{I,\mathrm{loc}}$ where v is defined by $v(t,x) := w^{-1/p^*}(t,x)\|Y(t,x)\|_{L^p}$. The claim is that v also belongs to L^∞_I. We only consider the case $p \in [1,2]$, the case $p \in (0,1)$ can be treated analogously. First, we suppose that Assumption D(6a) holds. In this case, it

follows from Lemma 2.6.1(3) that there exists some $\rho \in (0,1)$ with

$$v(t,x) \leq f(t,x) + \sum_{l=1}^{4} C_{\sigma,2} \Bigg(\int_I \int_{\mathbb{R}^d} G^{D,l}(t,x;s,y) \times (w(s,y))^{\rho-1} (v(s,y))^{p\rho} \, \lambda(\mathrm{d}s, \mathrm{d}y) \Bigg)^{1/p}, \tag{2.7.6}$$

where f denotes the sum of the first three terms on the right-hand side of (2.6.4). By hypothesis, the functions w^{-1}, $w^{-1/p}$ and $w^{\rho-1}$ are uniformly bounded on $I \times \mathbb{R}^d$, which means that f belongs to L_I^∞. Consequently, Lemma 2.6.5 together with (3), (4) and (5) of Assumption D shows that $v \in L_I^\infty$. Now suppose that Assumption D(6b) holds. Then, by replacing r in (2.7.6) by 1, the claim follows from Lemma 2.6.4(3) and assumption (2.5.8). □

Chapter 3:
Simulation of stochastic Volterra equations driven by space–time Lévy noise

3.1 Introduction

The aim of this Chapter is to investigate different simulation techniques for stochastic Volterra equations (SVEs) of the form

$$Y(t,x) = Y_0(t,x) + \int_0^t \int_{\mathbb{R}^d} G(t,x;s,y)\sigma(Y(s,y))\,\Lambda(\mathrm{d}s,\mathrm{d}y), \tag{3.1.1}$$

where $(t,x) \in \mathbb{R}_+ \times \mathbb{R}^d$, G is a deterministic kernel function, σ a Lipschitz coefficient and Λ a Lévy basis on $\mathbb{R}_+ \times \mathbb{R}^d$ of pure-jump type with no Gaussian part. In the purely temporal case where no space is involved and the kernel G is sufficiently regular on the diagonal $\{(t;s) \in \mathbb{R}_+ \times \mathbb{R}_+ \colon t = s\}$, the existence and uniqueness of the solution Y to (3.1.1) are established for general semimartingale integrators in [117]. The space–time case (3.1.1) is treated in Section 2 for quite general Lévy bases. In particular, G is allowed to be singular on the diagonal, which typically happens in the context of stochastic partial differential equations (SPDEs) where G is the Green's function of the underlying differential operator. More details on the connection between SPDEs and the SVE (3.1.1) are presented in Section 3.2, or can be found in [18, 136] and Section 2.

Since in most cases there exists no explicit solution formula for the SVE (3.1.1), it is a natural task to develop appropriate simulation algorithms. For SPDEs driven by Gaussian noise, research on this topic is rather far advanced, see e.g. [49, 75, 137]. However, for SPDEs driven by jump noises such as non-Gaussian Lévy bases, the related literature is considerably smaller, see [20] and the work of Hausenblas and

coauthors [57, 76, 77]. The case $\sigma \equiv 1$ has been treated in [40]. The contribution of this Chapter can be summarized as follows:

- We propose and analyze two approximation schemes for (3.1.1), each of which replaces the original noise by a truncated noise that only has finitely many atoms on compact subsets of $\mathbb{R}_+ \times \mathbb{R}^d$. For the first scheme, we simply cut off all jumps whose size is smaller than a constant. For the second scheme, we use series representation techniques for the noise as in [123] such that the jumps to be dropped off are chosen randomly. Both methods have already been applied successfully to the simulation of Lévy processes, cf. [7, 124].

- In the case where G originates from an SPDE, the crucial difference of our numerical schemes to the Euler or finite element methods in the references mentioned before is that we do not simulate small space–time increments of the noise but successively the true jumps of the Lévy basis, which is an easier task given that one usually only knows the underlying Lévy measure. It is important to recognize that this is only possible because the noise Λ is of pure-jump type, and contains neither a Gaussian part nor a drift. We shall point out in Section 3.6 how to relax this assumption.

The remaining article is organized as follows: Section 3.2 gives the necessary background for the SVE (3.1.1). In particular, we present sufficient conditions for the existence and uniqueness of solutions, and address the connection between (3.1.1) and SPDEs. In Section 3.3 we construct approximations to the solution Y of (3.1.1) by truncating the small jumps of the Lévy basis. We prove in Theorem 3.3.2 their L^p-convergence, and in some cases also their almost sure (a.s.) convergence to the target process Y. In Section 3.4 we approximate the driving Lévy basis using series representation methods. This leads to an algorithm that produces approximations again converging in the L^p-sense, sometimes also almost surely, to Y, see Theorem 3.4.3. In both theorems, we find explicit L^p-convergence rates that only depend on the kernel G and the characteristics of Λ. Section 3.5 presents a simulation study for the stochastic heat equation which highlights the typical path behaviour of stochastic Volterra equations. The final Section 3.6 compares the two simulation algorithms developed in this Chapter and discusses some further directions of the topic.

3.2 Preliminaries

We start with a summary of notations that will be employed in this Chapter.

$\mathbb{R}_+$	the set $[0,\infty)$ of *positive* real numbers;
$\mathbb{N}$	the natural numbers $\{1,2,\ldots\}$;
$\mathbb{B}$	a stochastic basis $(\Omega,\mathcal{F},\mathbb{F}=(\mathcal{F}_t)_{t\in\mathbb{R}_+},\mathbb{P})$ satisfying the usual hypotheses of completeness and right-continuity;
$\bar{\Omega}$,$\tilde{\Omega}$	$\bar{\Omega}:=\Omega\times\mathbb{R}_+$ and $\tilde{\Omega}:=\Omega\times\mathbb{R}_+\times\mathbb{R}^d$ where $d\in\mathbb{N}$;
$\mathcal{B}(\mathbb{R}^d)$	the Borel σ-field on $\mathbb{R}^d$;
$\tilde{\mathcal{B}}_{\rm b}$	the collection of all bounded Borel sets of $\mathbb{R}_+\times\mathbb{R}^d$;
$\mathcal{P}$	the predictable σ-field on $\mathbb{B}$ *or* the collection of all predictable processes $\bar{\Omega}\to\mathbb{R}$;
$\tilde{\mathcal{P}}$	the product $\mathcal{P}\otimes\mathcal{B}(\mathbb{R}^d)$ *or* the collection all $\mathcal{P}\otimes\mathcal{B}(\mathbb{R}^d)$-measurable processes $\tilde{\Omega}\to\mathbb{R}$;
$\tilde{\mathcal{P}}_{\rm b}$	the collection of sets in $\tilde{\mathcal{P}}$ which are a subset of $\Omega\times[0,k]\times[-k,k]^d$ for some $k\in\mathbb{N}$;
δ_a	the Dirac measure at a;
p^*	$p\vee 1$;
L^p	the space $L^p(\Omega,\mathcal{F},\mathbb{P})$, $p\in(0,\infty)$ with $\|X\|_{L^p}:=\mathbb{E}[\lvert X\rvert^p]^{1/p^*}$;
L^0	the space $L^0(\Omega,\mathcal{F},\mathbb{P})$ of all random variables on $\mathbb{B}$ endowed with the topology of convergence in probability;
$B^p_{\rm loc}$	the set of all $Y\in\tilde{\mathcal{P}}$ for which $\|Y(t,x)\|_{L^p}$ is uniformly bounded on $[0,T]\times\mathbb{R}^d$ for all $T\in\mathbb{R}_+$ ($p\in(0,\infty)$);
$A^{\rm C}$	the complement of A within the superset it belongs to (which will be clear from the context);
$A-B$	$\{x-y\colon x\in A,y\in B\}$;
$-A$	$\{-x\colon x\in A\}$;
Leb	the Lebesgue measure on $\mathbb{R}^d$ (d should be clear from the context);
$\|\cdot\|$	the Euclidean norm on $\mathbb{R}^d$;
$C, C(T)$	two generic constants in $\mathbb{R}_+$, one dependent and one independent of T, whose values are irrelevant in this Chapter and may therefore change from one place to the other

We suppose that the stochastic basis $\mathbb{B}$ supports a *Lévy basis*, that is, a mapping $\Lambda\colon\tilde{\mathcal{P}}_{\rm b}\to L^0$ with the following properties:

- $\Lambda(\emptyset) = 0$ a.s.
- For all pairwise disjoint sets $(A_i)_{i\in\mathbb{N}} \subset \tilde{\mathcal{P}}_{\mathrm{b}}$ with $\bigcup_{i=1}^{\infty} A_i \in \tilde{\mathcal{P}}_{\mathrm{b}}$ we have

$$\Lambda\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \Lambda(A_i) \quad \text{in } L^0. \tag{3.2.1}$$

- $(\Lambda(\Omega \times B_i))_{i\in\mathbb{N}}$ is a sequence of independent random variables if $(B_i)_{i\in\mathbb{N}}$ are pairwise disjoint sets in $\tilde{\mathcal{B}}_{\mathrm{b}}$.
- For every $B \in \mathcal{B}_{\mathrm{b}}$, $\Lambda(\Omega \times B)$ has an infinitely divisible distribution.
- $\Lambda(A)$ is $\mathcal{F}_t$-measurable when $A \in \tilde{\mathcal{P}}_{\mathrm{b}}$ and $A \subseteq \Omega \times [0,t] \times \mathbb{R}^d$ for $t \in \mathbb{R}_+$.
- For every $t \in \mathbb{R}_+$, $A \in \tilde{\mathcal{P}}_{\mathrm{b}}$ and $\Omega_0 \in \mathcal{F}_t$ we have a.s.

$$\Lambda(A \cap (\Omega_0 \times (t,\infty) \times \mathbb{R}^d)) = \mathbb{1}_{\Omega_0} \Lambda(A \cap (\Omega \times (t,\infty) \times \mathbb{R}^d)).$$

Just as Lévy processes are semimartingales and thus allow for an Itô integration theory, Lévy bases belong to the class of L^0-valued σ-finite random measures. Therefore, it is possible to define the stochastic integral

$$\int_{\mathbb{R}_+ \times \mathbb{R}^d} H(s,y)\, \Lambda(\mathrm{d}s, \mathrm{d}y)$$

for $H \in \tilde{\mathcal{P}}$ that are *integrable* with respect to Λ, see Section 1 for the details.

Similarly to Lévy processes, there exist two notions of characteristics for Lévy bases: one going back to [119, Prop. 2.1] that is based on the Lévy-Khintchine formula and is independent of $\mathbb{F}$, and a filtration-based one that is useful for stochastic analysis, see Theorem 1.3.2. For the whole Chapter, we will assume that both notions coincide such that Λ has a canonical decomposition under the filtration $\mathbb{F}$ of the form

$$\begin{aligned}\Lambda(\mathrm{d}t, \mathrm{d}x) = B(\mathrm{d}t, \mathrm{d}x) + \Lambda^{\mathrm{c}}(\mathrm{d}t, \mathrm{d}x) + \int_{\mathbb{R}} z \mathbb{1}_{\{|z|\le 1\}} \, (\mu - \nu)(\mathrm{d}t, \mathrm{d}x, \mathrm{d}z) \\ + \int_{\mathbb{R}} z \mathbb{1}_{\{|z|>1\}} \, \mu(\mathrm{d}t, \mathrm{d}x, \mathrm{d}z),\end{aligned}$$

where B is a σ-finite signed Borel measure on $\mathbb{R}_+ \times \mathbb{R}^d$, Λ^{c} a Lévy basis such that $\Lambda(\Omega \times B)$ is normally distributed with mean 0 and variance $C(B)$ for all $B \in \tilde{\mathcal{B}}_{\mathrm{b}}$,

and μ a Poisson measure on $\mathbb{R}_+ \times \mathbb{R}^d$ relative to $\mathbb{F}$ with intensity measure ν (cf. [80, Def. II.1.20]). There exists also a σ-finite Borel measure λ on $\mathbb{R}_+ \times \mathbb{R}^d$ such that

$$\begin{aligned} B(\mathrm{d}t, \mathrm{d}x) = b(t,x)\,\lambda(\mathrm{d}t, \mathrm{d}x), \quad C(\mathrm{d}t, \mathrm{d}x) = c(t,x)\,\lambda(\mathrm{d}t, \mathrm{d}x) \quad \text{and} \\ \nu(\mathrm{d}t, \mathrm{d}x, \mathrm{d}z) = \pi(t,x,\mathrm{d}z)\,\lambda(\mathrm{d}t, \mathrm{d}x) \end{aligned} \tag{3.2.2}$$

with two functions $b\colon \mathbb{R}_+ \times \mathbb{R}^d \to \mathbb{R}$ and $c\colon \mathbb{R}_+ \times \mathbb{R}^d \to \mathbb{R}_+$ as well as a transition kernel π from $(\mathbb{R}_+ \times \mathbb{R}^d, \mathcal{B}(\mathbb{R}_+ \times \mathbb{R}^d))$ to $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ such that $\pi(t,x,\cdot)$ is a Lévy measure for each $(t,x) \in \mathbb{R}_+ \times \mathbb{R}^d$.

We have already mentioned in the introduction that we will assume

$$C = 0 \tag{3.2.3}$$

throughout the Chapter. For simplicity we will also make two further assumptions: first, that there exist $b \in \mathbb{R}$ and a Lévy measure π such that for all $(t,x) \in \mathbb{R}_+ \times \mathbb{R}^d$ we have

$$b(t,x) = b, \quad \pi(t,x,\cdot) = \pi \quad \text{and} \quad \lambda(\mathrm{d}t, \mathrm{d}x) = \mathrm{d}(t,x); \tag{3.2.4}$$

second, that

$$\Lambda \in \mathcal{S} \cup \mathcal{V}_0, \tag{3.2.5}$$

where $\mathcal{S}$ is the collection of all *symmetric* Lévy bases and $\mathcal{V}_0$ is the class of Lévy bases with *locally finite variation and no drift*, defined by the property that

$$\int_{\mathbb{R}} |z| \mathbb{1}_{\{|z| \leq 1\}}\, \pi(\mathrm{d}z) < \infty, \quad \text{and} \quad b_0 := b - \int_{\mathbb{R}} z \mathbb{1}_{\{|z| \leq 1\}}\, \pi(\mathrm{d}z) = 0.$$

Furthermore, if Λ has a finite first moment, that is,

$$\int_{\mathbb{R}} |z| \mathbb{1}_{\{|z| > 1\}}\, \pi(\mathrm{d}z) < \infty, \tag{3.2.6}$$

we define

$$\begin{aligned} B_1(\mathrm{d}t, \mathrm{d}x) &:= b_1\, \mathrm{d}(t,x), \quad b_1 := b + \int_{\mathbb{R}} z \mathbb{1}_{\{|z| > 1\}}\, \pi(\mathrm{d}z), \\ M(\mathrm{d}t, \mathrm{d}x) &:= \Lambda(\mathrm{d}t, \mathrm{d}x) - B_1(\mathrm{d}t, \mathrm{d}x) = \int_{\mathbb{R}} z\, (\mu - \nu)(\mathrm{d}t, \mathrm{d}x, \mathrm{d}z). \end{aligned}$$

Next, let us summarize the most important facts regarding the SVE (3.1.1). All details that are not explained can be found in Section 2. First, many SPDEs of evolution type driven by Lévy noise can be written in terms of (3.1.1), where G is

the Green's function of the corresponding differential operator. Most prominently, taking G being the heat kernel in $\mathbb{R}^d$, (3.1.1) is the so-called *mild formulation* of the stochastic heat equation (with constant coefficients and multiplicative noise). Typically for parabolic equations, the heat kernel is very smooth in general but explodes on the diagonal $t = s$ and $x = y$. In fact, it is only p-fold integrable on $[0,T]\times\mathbb{R}^d$ for $p < 1+2/d$. In particular, as soon as $d \geq 2$, it is not square-integrable, and as a consequence, no solution to the stochastic heat equation in the form (3.1.1) will exist for Lévy noises with non-zero Gaussian component. This is another reason for including assumption (3.2.3) in this Chapter.

Second, let us address the existence and uniqueness problem for (3.1.1). By a *solution* to this equation we mean a predictable process $Y \in \tilde{\mathcal{P}}$ such that for all $(t,x) \in \mathbb{R}_+ \times \mathbb{R}^d$, the stochastic integral on the right-hand side of (3.1.1) is well defined and the equation itself for each $(t,x) \in [0,T]\times\mathbb{R}^d$ holds a.s. We identify two solutions as soon as they are modifications of each other. Given a number $p \in (0,2]$, the following conditions A1–A6 guarantee a unique solution to (3.1.1) in B^p_{loc} by Theorem 2.3.1:

A1. $Y_0 \in B^p_{\mathrm{loc}}$ is independent of Λ.

A2. $\sigma\colon \mathbb{R} \to \mathbb{R}$ is Lipschitz continuous, that is, there exists $C \in \mathbb{R}_+$ such that

$$|\sigma(x) - \sigma(y)| \leq C|x-y|, \quad x,y \in \mathbb{R}. \tag{3.2.7}$$

A3. $G\colon (\mathbb{R}_+ \times \mathbb{R}^d)^2 \to \mathbb{R}$ is a measurable function with $G(t,\cdot;s,\cdot) \equiv 0$ for $s > t$.

A4. Λ satisfies (3.2.2)–(3.2.5) and

$$\int_{\mathbb{R}} |z|^p \, \pi(\mathrm{d}z) < \infty. \tag{3.2.8}$$

A5. If we define for $(t,x),(s,y) \in \mathbb{R}_+ \times \mathbb{R}^d$

$$\tilde{G}(t,x;s,y) := |G(t,x;s,y)|\mathbb{1}_{\{p>1,\Lambda\notin\mathcal{S}\}} + |G(t,x;s,y)|^p, \tag{3.2.9}$$

then we have for all $T \in \mathbb{R}_+$

$$\sup_{(t,x)\in[0,T]\times\mathbb{R}^d} \int_0^T \int_{\mathbb{R}^d} \tilde{G}(t,x;s,y) \, \mathrm{d}(s,y) < \infty. \tag{3.2.10}$$

A6. For all $\varepsilon > 0$ and $T \in \mathbb{R}_+$ there exist $k \in \mathbb{N}$ and a partition $0 = t_0 < \ldots < t_k = T$ such that

$$\sup_{(t,x)\in[0,T]\times\mathbb{R}^d} \sup_{i=1,\ldots,k} \int_{t_{i-1}}^{t_i} \int_{\mathbb{R}^d} \tilde{G}(t,x;s,y)\,\mathrm{d}(s,y) < \varepsilon. \tag{3.2.11}$$

Apart from A1–A6, we will add another assumption in this Chapter:

A7. There exists a sequence $(U^N)_{N\in\mathbb{N}}$ of compact sets increasing to $\mathbb{R}^d$ such that for all $T \in \mathbb{R}_+$ and compact sets $K \subseteq \mathbb{R}^d$ we have, as $N \to \infty$,

$$\begin{aligned} r_1^N(T,K) := \sup_{(t,x)\in[0,T]\times K} \Bigg(& \int_0^t \int_{(U^N)^{\mathrm{C}}} |G(t,x;s,y)| \mathbb{1}_{\{p>1,\Lambda\notin\mathcal{S}\}}\,\mathrm{d}(s,y) \\ & + \left(\int_0^t \int_{(U^N)^{\mathrm{C}}} |G(t,x;s,y)|^p\,\mathrm{d}(s,y) \right)^{1/p^*} \Bigg) \to 0. \end{aligned} \tag{3.2.12}$$

Conditions A6 and A7 are automatically satisfied if $|G(t,x;s,y)| \le g(t-s,x-y)$ for some measurable function g and A5 holds with G replaced by g. For A6 see Remark 2.3.3(3); for A7 choose $U^N := \{x \in \mathbb{R}^d\colon \|x\| \le N\}$ such that for $p > 1$

$$\begin{aligned} & \sup_{(t,x)\in[0,T]\times K} \int_0^t \int_{(U^N)^{\mathrm{C}}} |G(t,x;s,y)|^p\,\mathrm{d}(s,y) \\ & \quad \le \sup_{(t,x)\in[0,T]\times K} \int_0^t \int_{(U^N)^{\mathrm{C}}} g^p(t-s,x-y)\,\mathrm{d}(s,y) \\ & \quad = \sup_{x\in K} \int_0^T \int_{x-(U^N)^{\mathrm{C}}} g^p(s,y)\,\mathrm{d}(s,y) \le \int_0^T \int_{K-(U^N)^{\mathrm{C}}} g^p(s,y)\,\mathrm{d}(s,y) \\ & \quad \to 0 \quad \text{as} \quad N \to \infty \end{aligned}$$

by the fact that $K - (U^N)^{\mathrm{C}} \downarrow 0$. A similar calculation applies to the case $p \in (0,1]$ and the first term in $r_1^N(T,K)$.

Example 3.2.1 We conclude this section with the stochastic heat equation in $\mathbb{R}^d$, whose mild formulation is given by the SVE (3.1.1) with

$$G(t,x;s,y) = g(t-s,x-y), \quad g(t,x) = \frac{\exp(-\|x\|^2/(4t))}{(4\pi t)^{d/2}} \mathbb{1}_{[0,t)}(s) \tag{3.2.13}$$

for $(t,x),(s,y) \in \mathbb{R}_+ \times \mathbb{R}^d$. We assume that Y_0 and σ satisfy conditions A1 and A2, respectively. Furthermore, we suppose that (3.2.2)–(3.2.5) are valid, and that

(3.2.8) holds with some $p \in (0, 1 + 2/d)$. It is straightforward to show that then A3–A6 are satisfied with the same p. Let us estimate the rate $r_1^N(T, K)$ for $T \in \mathbb{R}_+$, $K := \{\|x\| \leq R\}$ with $R \in \mathbb{N}$, and $U^N := \{\|x\| \leq N\}$. We first consider the case $p \leq 1$ or $\Lambda \in \mathcal{S}$. Since $K - (U^N)^{\mathrm{C}} = (U^{N-R})^{\mathrm{C}}$ for $N \geq R$, the calculations after A7 yield ($p(d) := 1 + (1-p)d/2$ and $\Gamma(\cdot,\cdot)$ denotes the upper incomplete gamma function)

$$\begin{aligned}(r_1^N(T,K))^{p^*} &\leq \int_0^T \int_{(U^{N-R})^{\mathrm{C}}} g^p(t,x)\,\mathrm{d}(t,x) = \int_0^T \int_{N-R}^\infty \frac{\exp(-pr^2/(4t))}{(4\pi t)^{pd/2}} r^{d-1}\,\mathrm{d}r\,\mathrm{d}t \\ &= C\int_0^T t^{p(d)-1}\Gamma\left(\frac{d}{2}, \frac{p(N-R)^2}{4t}\right)\mathrm{d}t \\ &= C(T)\Bigg((p(d))^{-1}\Gamma\left(\frac{d}{2}, \frac{p(N-R)^2}{4T}\right) - \left(\frac{p(N-R)^2}{4T}\right)^{p(d)} \\ &\quad \times \Gamma\left(\frac{d}{2} - p(d), \frac{p(N-R)^2}{4T}\right)\Bigg) \\ &\leq C(T)\exp\left(-\frac{p(N-R)^2}{4T}\right)(N-R)^{d-2}, \end{aligned} \tag{3.2.14}$$

which tends to 0 exponentially fast as $N \to \infty$. If $p > 1$ and $\Lambda \notin \mathcal{S}$, it follows from formula (3.2.12) that we need an extra summand for $r_1^N(T, K)$, namely (3.2.14) with $p = 1$. □

3.3 Truncation of small jumps

In this section we approximate equation (3.1.1) by cutting off the small jumps of Λ. To this end, we first define for each $N \in \mathbb{N}$

$$G^N(t,x;s,y) := G(t,x;s,y)\mathbb{1}_{U^N}(y), \quad (t,x),(s,y) \in \mathbb{R}_+ \times \mathbb{R}^d, \tag{3.3.1}$$

where the meaning of the sets U^N is explained in A7. Furthermore, we introduce

$$r_2^N := \left(\int_{[-\varepsilon^N,\varepsilon^N]} |z|^p\,\pi(\mathrm{d}z)\right)^{1/p^*}, \quad r_3^N := \left|\int_{[-\varepsilon^N,\varepsilon^N]} z\mathbb{1}_{\{p>1,\Lambda\notin\mathcal{S}\}}\,\pi(\mathrm{d}z)\right|, \tag{3.3.2}$$

where $(\varepsilon^N)_{N\in\mathbb{N}} \subseteq (0,1)$ satisfies $\varepsilon^N \to 0$ as $N \to \infty$. Condition A4 implies that $r_2^N, r_3^N \to 0$ as $N \to \infty$. Next, defining truncations of the Lévy basis Λ by

$$\Lambda^N(\mathrm{d}t,\mathrm{d}x) := \int_{[-\varepsilon^N,\varepsilon^N]^{\mathrm{C}}} z\,\mu(\mathrm{d}t,\mathrm{d}x,\mathrm{d}z), \tag{3.3.3}$$

our approximation scheme for the solution Y to (3.1.1) is given as:

$$Y^N(t,x) := Y_0(t,x) + \int_0^t \int_{\mathbb{R}^d} G^N(t,x;s,y)\sigma(Y^N(s,y))\,\Lambda^N(\mathrm{d}s,\mathrm{d}y) \tag{3.3.4}$$

for $(t,x) \in \mathbb{R}_+ \times \mathbb{R}^d$. Indeed, Y^N can be simulated exactly because for all $T \in \mathbb{R}_+$ the truncation Λ^N only has a finite (random) number $R^N(T)$ of jumps on $[0,T] \times U^N$, say at the space–time locations (τ_i^N, ξ_i^N) with sizes J_i^N. This implies that we have the following alternative representation of $Y^N(t,x)$ for $(t,x) \in [0,T] \times \mathbb{R}^d$:

$$Y^N(t,x) = Y_0(t,x) + \sum_{i=1}^{R^N(T)} G(t,x;\tau_i^N,\xi_i^N)\sigma(Y^N(\tau_i^N,\xi_i^N))J_i^N \mathbb{1}_{\{\tau_i^N < t\}}. \tag{3.3.5}$$

What remains to do is to simulate $Y^N(\tau_i^N,\xi_i^N)$, $i = 1,\dots,R^N(T)$, iteratively, from which the values $Y(t,x)$ for all other $(t,x) \in [0,T] \times \mathbb{R}^d$ can be computed.

The following algorithm summarizes up the simulation procedure:

Algorithm 3.3.1 Consider a finite grid $\mathcal{G} \subseteq [0,T] \times \mathbb{R}^d$. For each N proceed as follows:

(1) Draw a Poisson random variable $R^N(T)$ with parameter

$$R^N(T) := \int_0^T \int_{U^N} \int_{\mathbb{R}} \mathbb{1}_{\{z \in [-\varepsilon^N,\varepsilon^N]^{\mathrm{C}}\}}\,\nu(\mathrm{d}t,\mathrm{d}x,\mathrm{d}z) = T\mathrm{Leb}(U^N)\pi\Big([-\varepsilon^N,\varepsilon^N]^{\mathrm{C}}\Big).$$

(2) For $i = 1,\dots,R^N(T)$:

   (a) Draw a pair (τ_i^N,ξ_i^N) with uniform distribution from $[0,T] \times U^N$.

   (b) Draw J_i^N from $[-\varepsilon^N,\varepsilon^N]^{\mathrm{C}}$ with distribution $\pi/\pi\Big([-\varepsilon^N,\varepsilon^N]^{\mathrm{C}}\Big)$.

   (c) Rename $(\tau_i^N,\xi_i^N,J_i^N\colon i = 1,\dots,R^N(T))$ by sorting $(\tau_i^N\colon i = 1,\dots,R^N(T))$ in increasing order.

(3) For each $i = 1,\dots,R^N(T)$ and $(t,x) \in \mathcal{G}$ simulate $Y_0(\tau_i^N,\xi_i^N)$ and $Y_0(t,x)$.

(4) For each $i = 1,\dots,R^N(T)$ set

$$Y^N(\tau_i^N,\xi_i^N) := Y_0(\tau_i^N,\xi_i^N) + \sum_{j=1}^{i-1} G(\tau_i^N,\xi_i^N;\tau_j^N,\xi_j^N)\sigma(Y^N(\tau_j^N,\xi_j^N))J_j^N.$$

(5) For each $(t,x) \in \mathcal{G}$ define $Y^N(t,x)$ via (3.3.5). □

The next theorem determines the convergence behaviour of (3.3.4) to (3.1.1).

Theorem 3.3.2. *Grant assumptions A1–A7 under which the SVE* (3.1.1) *has a unique solution in* B^p_{loc}. *Then* Y^N *as defined in* (3.3.4) *belongs to* B^p_{loc} *for all* $N \in \mathbb{N}$, *and for all* $T \in \mathbb{R}_+$ *and compact sets* $K \subseteq \mathbb{R}^d$ *there exists a constant* $C(T) \in \mathbb{R}_+$ *independent of* N *and* K *such that*

$$\sup_{(t,x)\in[0,T]\times K} \|Y(t,x) - Y^N(t,x)\|_{L^p} \leq C(T)(r_1^N(T,K) + r_2^N + r_3^N). \tag{3.3.6}$$

Furthermore, if $\sum_{N=1}^{\infty}(r_1^N(T,K) + r_2^N + r_3^N)^{p^*} < \infty$ *is fulfilled, then we also have for all* $(t,x) \in [0,T] \times K$ *that* $Y^N(t,x) \to Y(t,x)$ *a.s. as* $N \to \infty$.

Proof. It is obvious that $|G^N| \leq |G|$ pointwise and that we have $\nu^N \leq \nu$ for the third characteristic ν^N of Λ^N. Thus, A1–A6 are still satisfied when G and ν are replaced by G^N and ν^N (if $\Lambda \in \mathcal{S}$, also $\Lambda^N \in \mathcal{S}$). So Y^N as a solution to (3.1.1) with G^N and Λ^N instead of G and Λ belongs to B^p_{loc} as well. Moreover, for all $T \in \mathbb{R}_+$ there exists $C(T) \in \mathbb{R}_+$ independent of $N \in \mathbb{N}$ such that

$$\sup_{(t,x)\in[0,T]\times\mathbb{R}^d} \|Y^N(t,x)\|_{L^p} \leq C(T), \quad N \in \mathbb{N}. \tag{3.3.7}$$

We only sketch the proof for this statement. In fact, using Lemma 2.6.1(1) it can be shown that the left-hand side of (3.3.7) satisfies an inequality of the same type as in Lemma 6.4(3) of the same paper. In particular, it is bounded by a constant $C^N(T)$ that depends on N only through $|G^N|$ and ν^N, and that this constant is only increased if we replace $|G^N|$ and ν^N by the larger $|G|$ and ν. In this way, we obtain an upper bound $C(T)$ that does not depend on N.

Next, we prove the convergence of Y^N to Y as stated in (3.3.6). We have

$$\begin{aligned} Y(t,x) - Y^N(t,x) &= \int_0^t \int_{\mathbb{R}^d} [G(t,x;s,y) - G^N(t,x;s,y)]\sigma(Y(s,y))\,\Lambda(\mathrm{d}s,\mathrm{d}y) \\ &\quad + \int_0^t \int_{\mathbb{R}^d} G^N(t,x;s,y)[\sigma(Y(s,y)) - \sigma(Y^N(s,y))]\,\Lambda(\mathrm{d}s,\mathrm{d}y) \\ &\quad + \int_0^t \int_{\mathbb{R}^d} G^N(t,x;s,y)\sigma(Y^N(s,y))\,(\Lambda - \Lambda^N)(\mathrm{d}s,\mathrm{d}y) \\ &=: I_1^N(t,x) + I_2^N(t,x) + I_3^N(t,x), \quad (t,x) \in \mathbb{R}_+ \times \mathbb{R}^d. \end{aligned} \tag{3.3.8}$$

If $p > 1$, we have

$$\begin{aligned}\|I_2^N(t,x)\|_{L^p} &\le \left\| \int_0^t \int_{\mathbb{R}^d} G^N(t,x;s,y)[\sigma(Y(s,y)) - \sigma(Y^N(s,y))]\, B_1(\mathrm{d}s,\mathrm{d}y) \right\|_{L^p} \\ &\quad + \left\| \int_0^t \int_{\mathbb{R}^d} G^N(t,x;s,y)[\sigma(Y(s,y)) - \sigma(Y^N(s,y))]\, M(\mathrm{d}s,\mathrm{d}y) \right\|_{L^p} \\ &\le C \Bigg(\left(\int_0^t \int_{\mathbb{R}^d} |G(t,x;s,y)|\, |B_1|(\mathrm{d}s,\mathrm{d}y) \right)^{p-1} \\ &\quad \times \int_0^t \int_{\mathbb{R}^d} |G(t,x;s,y)| \|Y(s,y) - Y^N(s,y)\|_{L^p}^p\, |B_1|(\mathrm{d}s,\mathrm{d}y) \Bigg)^{1/p} \\ &\quad + C \left(\int_0^t \int_{\mathbb{R}^d} |G(t,x;s,y)|^p \|Y(s,y) - Y^N(s,y)\|_{L^p}^p\, \mathrm{d}(s,y) \right)^{1/p} \end{aligned} \tag{3.3.9}$$

by (3.2.7), Hölder's inequality and the Burkholder-Davis-Gundy-inequality. If $p \in (0,1]$, we have $\Lambda \in \mathcal{V}_0$ by (3.2.5) and (3.2.8), and thus Jensen's inequality gives

$$\begin{aligned}\|I_2^N(t,x)\|_{L^p} &= \mathbb{E}\left[\left(\int_0^t \int_{\mathbb{R}^d} G^N(t,x;s,y)[\sigma(Y(s,y)) - \sigma(Y^N(s,y))] z\, \mu(\mathrm{d}s,\mathrm{d}y,\mathrm{d}z) \right)^p \right] \\ &\le \mathbb{E}\left[\int_0^t \int_{\mathbb{R}^d} |G^N(t,x;s,y)[\sigma(Y(s,y)) - \sigma(Y^N(s,y))] z|^p\, \nu(\mathrm{d}s,\mathrm{d}y,\mathrm{d}z) \right] \\ &\le C \int_0^t \int_{\mathbb{R}^d} |G(t,x;s,y)|^p \|Y(s,y) - Y^N(s,y)\|_{L^p}\, \mathrm{d}(s,y). \end{aligned} \tag{3.3.10}$$

Inserting the estimates (3.3.9) and (3.3.10) back into (3.3.8), we obtain for the function $w^N(t,x) := \|Y(t,x) - Y^N(t,x)\|_{L^p}$ that

$$\begin{aligned} w^N(t,x) \le C(T) \Bigg(& \left(\int_0^t \int_{\mathbb{R}^d} |G(t,x;s,y)| \mathbb{1}_{\{p>1, \Lambda \notin \mathcal{S}\}} (w^N(s,y))^p\, \mathrm{d}(s,y) \right)^{1/p} \\ & + \left(\int_0^t \int_{\mathbb{R}^d} |G(t,x;s,y)|^p (w^N(s,y))^{p^*}\, \mathrm{d}(s,y) \right)^{1/p^*} \Bigg) \\ & + \|I_1^N(t,x) + I_3^N(t,x)\|_{L^p}, \quad (t,x) \in [0,T] \times \mathbb{R}^d. \end{aligned}$$

By a Gronwall-type estimate, which is possible because of A5 (see the proof of Theorem 2.4.7(3) for an elaboration of an argument of this type), we conclude

$$\sup_{(t,x) \in [0,T] \times K} w^N(t,x) \le C(T) \sup_{(t,x) \in [0,T] \times K} \|I_1^N(t,x) + I_3^N(t,x)\|_{L^p}.$$

where $C(T)$ does not depend on K because of (3.2.10). For $I_1^N(t,x)$ we have for $p>1$

$$\begin{aligned}\|I_1^N(t,x)\|_{L^p} &\le \left\|\int_0^t\int_{\mathbb{R}^d}[G(t,x;s,y)-G^N(t,x;s,y)]\sigma(Y(s,y))\,B_1(\mathrm{d}s,\mathrm{d}y)\right\|_{L^p}\\ &\quad+\left\|\int_0^t\int_{\mathbb{R}^d}[G(t,x;s,y)-G^N(t,x;s,y)]\sigma(Y(s,y))\,M(\mathrm{d}s,\mathrm{d}y)\right\|_{L^p}\\ &\le C\left(1+\sup_{(t,x)\in[0,T]\times\mathbb{R}^d}\|Y(t,x)\|_{L^p}\right)\left(\int_0^t\int_{(U^N)^{\mathrm{C}}}|G(t,x;s,y)|\,|B_1|(\mathrm{d}s,\mathrm{d}y)\right.\\ &\quad\left.+\left(\int_0^t\int_{(U^N)^{\mathrm{C}}}|G(t,x;s,y)z|^p\,\nu(\mathrm{d}s,\mathrm{d}y,\mathrm{d}z)\right)^{1/p}\right)\\ &\le C(T)r_1^N(T,K),\end{aligned}\tag{3.3.11}$$

uniformly in $(t,x)\in[0,T]\times K$. In similar fashion one proves the estimate (3.3.11) for $p\in(0,1]$, perhaps with a different $C(T)$. Next, when $p>1$, (3.3.7) implies

$$\begin{aligned}\|I_3^N(t,x)\|_{L^p} &= \left\|\int_0^t\int_{\mathbb{R}^d}\int_{[-\varepsilon^N,\varepsilon^N]}G^N(t,x;s,y)\sigma(Y^N(s,y))z\,(\mu-\nu)(\mathrm{d}s,\mathrm{d}y,\mathrm{d}z)\right\|_{L^p}\\ &\quad+\left\|\int_0^t\int_{\mathbb{R}^d}\int_{[-\varepsilon^N,\varepsilon^N]}G^N(t,x;s,y)\sigma(Y^N(s,y))z\mathbb{1}_{\{\Lambda\notin\mathcal{S}\}}\,\nu(\mathrm{d}s,\mathrm{d}y,\mathrm{d}z)\right\|_{L^p}\\ &\le C(T)\left(\left(\int_0^t\int_{\mathbb{R}^d}\int_{[-\varepsilon^N,\varepsilon^N]}|G(t,x;s,y)z|^p\,\pi(\mathrm{d}z)\,\mathrm{d}(s,y)\right)^{1/p}\right.\\ &\quad\left.+\left|\int_{[-\varepsilon^N,\varepsilon^N]}z\mathbb{1}_{\{\Lambda\notin\mathcal{S}\}}\,\pi(\mathrm{d}z)\right|\int_0^t\int_{\mathbb{R}^d}|G(t,x;s,y)|\mathbb{1}_{\{\Lambda\notin\mathcal{S}\}}\,\mathrm{d}(s,y)\right)\\ &\le C(T)(r_2^N+r_3^N).\end{aligned}$$

The case $p\in(0,1]$ can be treated similarly, cf. the estimation of $I_2^N(t,x)$ above.

It remains to prove that for each $(t,x)\in[0,T]\times K$ the convergence of $Y^N(t,x)$ to $Y(t,x)$ is almost sure when $r_1^N(T,K)$, r_2^N and r_3^N are p^*-summable. To this end, choose an arbitrary sequence $(a_N)_{N\in\mathbb{N}}\subseteq(0,1)$ converging to 0 such that

$$\sum_{N=1}^{\infty}A_N<\infty\quad\text{with}\quad A_N:=\frac{(r_1^N(T,K)+r_2^N+r_3^N)^{p^*}}{a_N^p}.$$

Such a sequence always exists, see [90, Thm. 175.4], for example. So by (3.3.6) and Chebyshev's inequality we derive

$$\mathbb{P}\left[|Y(t,x)-Y^N(t,x)|\ge a_N\right]\le\frac{\|Y(t,x)-Y^N(t,x)\|_{L^p}^{p^*}}{a_N^p}\le C(T)A_N.$$

Our assertion now follows from the Borel-Cantelli lemma. □

Example 3.3.3 The rates r_2^N and r_3^N from (3.3.2) only depend on the underlying Lévy measure π. Let $p, q \in (0, 2]$ with $q < p$, and assume that $\int_{[-1,1]} |z|^q \, \pi(\mathrm{d}z) < \infty$. If $\Lambda \in \mathcal{V}_0$, assume that $q < 1$. Then

$$\begin{aligned} r_2^N &= \left(\int_{[-\varepsilon^N, \varepsilon^N]} |z|^p \, \pi(\mathrm{d}z) \right)^{1/p^*} \leq \left(\int_{[-1,1]} |z|^q \, \nu(\mathrm{d}z) (\varepsilon^N)^{p-q} \right)^{1/p^*} \\ &= \mathcal{O}\left((\varepsilon^N)^{(p-q)/p^*} \right), \\ r_3^N &= \left| \int_{[-\varepsilon^N, \varepsilon^N]} z \mathbb{1}_{\{p>1, \Lambda \notin \mathcal{S}\}} \, \pi(\mathrm{d}z) \right| \leq \mathcal{O}\left((\varepsilon^N)^{1-q} \mathbb{1}_{\{p>1, \Lambda \notin \mathcal{S}\}} \right). \end{aligned}$$

For instance, if $\varepsilon^N = 1/N^k$, then the sequence $(r_2^N)^{p^*} = \mathcal{O}(N^{-k(p-q)})$ is summable for all $k > (p-q)^{-1}$. So in order to obtain a.s. convergence of $Y^N(t,x) \to Y(t,x)$, a sufficient condition is to choose the truncation rates ε^N small enough. Similar conclusions are valid for the other two rates $r_1^N(T,K)$ and r_3^N. □

3.4 Truncation via series representations

Another approach to the simulation of (3.1.1) uses series representations for the Lévy basis. The idea, going back to [122, 123] and already applied to the simulation of Lévy processes [124], is to choose the jumps to be simulated in a random order. Instead of selecting the big jumps first and the smaller jumps later as in Section 3.3, we only choose the big jumps first more *likely*. The details are as follows: we fix a finite time horizon $T \in \mathbb{R}_+$ and, recalling A7, a partition $(Q^i)_{i \in \mathbb{N}}$ of $\mathbb{R}^d$ into pairwise disjoint compact sets such that $U^N = \bigcup_{i=1}^N Q^i$. We now assume that the jump measure μ of Λ on the strip $[0,T] \times \mathbb{R}^d \times \mathbb{R}$ can be represented in the form

$$\begin{aligned} \mu(\mathrm{d}t, \mathrm{d}x, \mathrm{d}z) &= \sum_{i=1}^{\infty} \mu_i(\mathrm{d}t, \mathrm{d}x, \mathrm{d}z), \\ \mu_i(\mathrm{d}t, \mathrm{d}x, \mathrm{d}z) &= \sum_{j=1}^{\infty} \delta_{(\tau_j^i, \xi_j^i, H(\Gamma_j^i, V_j^i))}(\mathrm{d}t, \mathrm{d}x, \mathrm{d}z) \quad \text{a.s.}, \end{aligned} \tag{3.4.1}$$

where $H \colon (0, \infty) \times \mathbb{R} \to \mathbb{R}$ is a measurable function, satisfying $H(\cdot, v) = -H(\cdot, -v)$ for all $v \in \mathbb{R}$ when $\Lambda \in \mathcal{S}$, and the random variables involved have the following properties for each $i \in \mathbb{N}$:

- $(\tau_j^i\colon j \in \mathbb{N})$ and $(\xi_j^i\colon j \in \mathbb{N})$ are i.i.d. sequences with uniform distribution on $[0,T]$ and Q^i, respectively.
- $(\Gamma_j^i\colon j \in \mathbb{N})$ is a random walk whose increments are exponentially distributed with mean $1/T$.
- $(V_j^i\colon j \in \mathbb{N})$ is an i.i.d. sequence with distribution F on $\mathbb{R}$, which we should be able to simulate from. We assume that F is symmetric when $\Lambda \in \mathcal{S}$.
- The sequences τ^i, ξ^i, Γ^i and V^i are independent from each other.
- $(\tau^i, \xi^i, \Gamma^i, V^i)$ is independent from $(\tau^k, \xi^k, \Gamma^k, V^k\colon k \neq i)$.

Because of (3.2.5), μ can always be written in the form (3.4.1) whenever the underlying stochastic basis is rich enough. We give three examples of such series representations.

Example 3.4.1 The proofs that the following choices are valid can be found in [124, Sect. 3], where also more examples are discussed.

(1) *LePage's method*: we set $F := (\delta_{-1} + \delta_1)/2$ and $H(r, \pm 1) := \pm \varrho^{\leftarrow}(r, \pm 1)$, where $\varrho^{\leftarrow}(r, \pm 1) = \inf\{x \in (0,\infty)\colon \pi(\pm[x,\infty)) < r\}$ for $r \in (0,\infty)$.

(2) *Bondesson's method*: we assume that $\pi(A) = \int_0^\infty F(A/g(t))\,\mathrm{d}t$ for $A \in \mathcal{B}(\mathbb{R}^d)$ with some non-increasing $g\colon \mathbb{R}_+ \to \mathbb{R}_+$. Then we define $H(r,v) := g(r)v$.

(3) *Thinning method*: we choose F in such a way that Q is absolutely continuous with respect to F with density q, and define $H(r,v) := v\mathbb{1}_{\{q(v)\geq r\}}$. □

Our approximation scheme is basically the same as in Section 3.3: we define G^N by (3.3.1) and Y^N by (3.3.4), with the difference that Λ^N on $[0,T]\times\mathbb{R}^d$ is now defined as

$$\begin{aligned} \Lambda^N(\mathrm{d}t,\mathrm{d}x) &:= \int_{\mathbb{R}} z\,\mu^N(\mathrm{d}t,\mathrm{d}x,\mathrm{d}z), \\ \mu^N(\mathrm{d}t,\mathrm{d}x,\mathrm{d}z) &:= \sum_{i=1}^{\infty} \sum_{j\colon \Gamma_j^i \leq N} \delta_{(\tau_j^i, \xi_j^i, H(\Gamma_j^i, V_j^i))}(\mathrm{d}t,\mathrm{d}x,\mathrm{d}z). \end{aligned} \tag{3.4.2}$$

We can therefore rewrite $Y^N(t,x)$ for $(t,x) \in [0,T]\times\mathbb{R}^d$ as

$$Y^N(t,x) = Y_0(t,x) + \sum_{i=1}^{N} \sum_{j\colon \Gamma_j^i \leq N} G(t,x;\tau_j^i,\xi_j^i)\sigma(Y^N(\tau_j^i,\xi_j^i))H(\Gamma_j^i,V_j^i)\mathbb{1}_{\{\tau_j^i<t\}}. \tag{3.4.3}$$

This yields the following simulation algorithm:

Algorithm 3.4.2 Let $\mathcal{G}$ be a finite grid in $[0,T]\times\mathbb{R}^d$ and $N\in\mathbb{N}$.

(1) For each $i=1,\dots,N$ set $j:=1$ and repeat the following:

  (a) Draw E_j^i from an exponential distribution with mean $1/T$.

  (b) Define $\Gamma_j^i:=\Gamma_{j-1}^i+E_j^i$ ($\Gamma_0^i:=0$).

  (c) If $\Gamma_j^i>N$, set $J_i:=j-1$ and leave the loop; otherwise set $j:=j+1$.

(2) For each $i=1,\dots,N$ and $j=1,\dots,J_i$ simulate independently

  (a) a pair (τ_j^i,ξ_j^i) with uniform distribution on $[0,T]\times Q^i$;

  (b) a random variable V_j^i with distribution F;

  (c) the random variable $Y_0(\tau_j^i,\xi_j^i)$.

(3) Sort the sequence $(\tau_j^i\colon i=1,\dots,N, j=1,\dots,J_i)$ in increasing order, yielding sequences $(\tau_i,\xi_i,\Gamma_i,V_i\colon i=1,\dots,\sum_{j=1}^N J_j)$. Now define

$$Y^N(\tau_i,\xi_i):=Y_0(\tau_i,\xi_i)+\sum_{j=1}^{i-1}G(\tau_i,\xi_i;\tau_j,\xi_j)\sigma(Y^N(\tau_j,\xi_j))H(\Gamma_j,V_j).$$

(4) For each $(t,x)\in\mathcal{G}$ simulate $Y_0(t,x)$ and define $Y(t,x)$ by (3.4.3). □

We can now prove a convergence theorem for Y^N to Y, similar to Theorem 3.3.2. Define

$$\begin{aligned} r_2^N&:=\left(\int_N^\infty\int_{\mathbb{R}}|H(r,v)|^p\,F(\mathrm{d}v)\,\mathrm{d}r\right)^{1/p^*},\\ r_3^N&:=\left|\int_N^\infty\int_{\mathbb{R}}H(r,v)\mathbb{1}_{\{p>1,\Lambda\notin\mathcal{S}\}}\,F(\mathrm{d}v)\,\mathrm{d}r\right|. \end{aligned}\tag{3.4.4}$$

Theorem 3.4.3. *Grant assumptions A1–A7 under which the SVE* (3.1.1) *has a unique solution in* B^p_{loc}. *Further suppose that the jump measure* μ *of* Λ *has a representation in form of* (3.4.1). *Then* Y^N *as defined in* (3.4.3) *belongs to* B^p_{loc} *for all* $N\in\mathbb{N}$, *and for all* $T\in\mathbb{R}_+$ *and compact sets* $K\subseteq\mathbb{R}^d$ *there exists a constant* $C(T)\in\mathbb{R}_+$ *independent of* N *and* K *such that*

$$\sup_{(t,x)\in[0,T]\times K}\|Y(t,x)-Y^N(t,x)\|_{L^p}\le C(T)(r_1^N(T,K)+r_2^N+r_3^N).\tag{3.4.5}$$

If $\sum_{N=1}^{\infty}(r_1^N(T,K)+r_2^N+r_3^N)^{p^*}<\infty$*, then we also have for all* $(t,x)\in[0,T]\times K$ *that* $Y^N(t,x)\to Y(t,x)$ *a.s. as* $N\to\infty$*.*

Proof. We start with some preliminaries. It follows from (3.4.1) and [124, Prop. 2.1] that we have $\nu=\bar{\nu}\circ h^{-1}$ where $\bar{\nu}(\mathrm{d}t,\mathrm{d}x,\mathrm{d}r,\mathrm{d}v)=\mathrm{d}t\,\mathrm{d}x\,\mathrm{d}r\,F(\mathrm{d}v)$ and $h(t,x,r,v)=(t,x,H(r,v))$ on $[0,T]\times\mathbb{R}^d\times\mathbb{R}$. Therefore, conditions (3.2.5) and (3.2.8) imply that

$$\int_0^\infty\int_{\mathbb{R}}|H(r,v)|^p\,F(\mathrm{d}v)\,\mathrm{d}r=\int_{\mathbb{R}}|z|^p\,\pi(\mathrm{d}z)<\infty,$$

$$\text{and}\quad\int_0^\infty\int_{\mathbb{R}}|H(r,v)|\mathbb{1}_{\{p>1,\Lambda\notin\mathcal{S}\}}\,F(\mathrm{d}v)\,\mathrm{d}r=\int_{\mathbb{R}}|z|\mathbb{1}_{\{p>1,\Lambda\notin\mathcal{S}\}}\,\pi(\mathrm{d}z)<\infty.$$

Consequently, r_2^N and r_3^N are well defined and converge to 0 when $N\to\infty$. Similarly, the compensator ν_N of the measure $\mu-\mu^N$ is given by $\nu_N(\mathrm{d}t,\mathrm{d}x,\mathrm{d}z)=\mathrm{d}t\,\mathrm{d}x\,\pi_N(\mathrm{d}z)$, where $\pi_N=(\mathrm{Leb}\otimes F)\circ H_N^{-1}$ and $H_N(r,v)=H(r,v)\mathbb{1}_{(N,\infty)}(r)$.

For the actual proof of Theorem 3.4.3 one can basically follow the proof of Theorem 3.3.2. Only the estimation of $I_3^N(t,x)$ as defined in (3.3.8) is different, which we shall carry out now. In the case of $p>1$, we again use the Burkholder-Davis-Gundy inequality and obtain for $(t,x)\in[0,T]\times\mathbb{R}^d$

$$\begin{aligned}\|I_3^N(t,x)\|_{L^p}&=\left\|\int_0^t\int_{\mathbb{R}^d}\int_{\mathbb{R}}G^N(t,x;s,y)\sigma(Y^N(s,y))z\,(\mu_N-\nu_N)(\mathrm{d}s,\mathrm{d}y,\mathrm{d}z)\right\|_{L^p}\\&\quad+\left\|\int_0^t\int_{\mathbb{R}^d}\int_{\mathbb{R}}G^N(t,x;s,y)\sigma(Y^N(s,y))z\mathbb{1}_{\{\Lambda\notin\mathcal{S}\}}\,\nu_N(\mathrm{d}s,\mathrm{d}y,\mathrm{d}z)\right\|_{L^p}\\&\le C(T)\Bigg(\left(\int_0^t\int_{\mathbb{R}^d}\int_N^\infty\int_{\mathbb{R}}|G(t,x;s,y)H(r,v)|^p\,F(\mathrm{d}v)\,\mathrm{d}r\,\mathrm{d}(s,y)\right)^{1/p}\\&\quad+\int_0^t\int_{\mathbb{R}^d}|G(t,x;s,y)|\left|\int_N^\infty\int_{\mathbb{R}}H(r,v)\mathbb{1}_{\{\Lambda\notin\mathcal{S}\}}\,F(\mathrm{d}v)\,\mathrm{d}r\right|\mathrm{d}(s,y)\Bigg)\\&\le C(T)(r_2^N+r_3^N).\end{aligned}$$

The case $p\in(0,1]$ is treated analogously. One only needs to replace $\mu_N-\nu_N$ by μ_N and estimate via Jensen's inequality. □

Example 3.4.4 (Continuation of Example 3.4.1) We calculate the rates r_2^N and r_3^N from (3.4.4) for the series representations given in Example 3.4.1. We assume that $p,q\in(0,2]$ with $q<p$ are chosen such that $\int_{[-1,1]}|z|^q\,\pi(\mathrm{d}z)<\infty$, and $q<1$ if $\Lambda\in\mathcal{V}_0$. For all three examples we use the fact that $\pi=(\mathrm{Leb}\otimes F)\circ H^{-1}$ and that $r>N$ implies $|H(r,v)|\le|H(N,v)|$ for all $v\in\mathbb{R}$.

(1) *LePage's method*: We have

$$(r_2^N)^{p^*} = \int_N^\infty \frac{|H(r,1)|^p + |H(r,-1)|^p}{2}\,\mathrm{d}r \le \frac{1}{2}\int_{[H(N,-1),H(N,1)]} |z|^p\,\pi(\mathrm{d}z)$$
$$\le \frac{1}{2}\int_{[H(1,-1),H(1,1)]} |z|^q\,\pi(\mathrm{d}z)(|H(N,-1)| \vee |H(N,1)|)^{p-q},$$

and therefore

$$r_2^N = \mathcal{O}\left((\varrho^{\leftarrow}(N,1) \vee \varrho^{\leftarrow}(N,-1))^{(p-q)/p^*}\right),$$
$$r_3^N = \mathcal{O}\left((\varrho^{\leftarrow}(N,1) \vee \varrho^{\leftarrow}(N,-1))^{1-q}\mathbb{1}_{\{p>1,\Lambda\notin\mathcal{S}\}}\right).$$

(2) *Bondesson's method*: Since $H(r,v) = g(r)v$ and g is non-increasing, we obtain

$$(r_2^N)^{p^*} = \int_N^\infty \int_{\mathbb{R}} |g(r)v|^p\,F(\mathrm{d}v)\,\mathrm{d}r \le (g(N))^{p-q}\int_0^\infty g^q(r)\,\mathrm{d}r\int_{\mathbb{R}}|v|^p\,F(\mathrm{d}v),$$

and consequently

$$r_2^N = \mathcal{O}\left(g(N)^{(p-q)/p^*}\right), \quad r_3^N = \mathcal{O}\left(g(N)^{1-q}\mathbb{1}_{\{p>1,\Lambda\notin\mathcal{S}\}}\right).$$

(3) *Thinning method*: Here we have

$$r_2^N = \left(\int_{\mathbb{R}}\int_N^{q(v)\vee N} |v|^p\,\mathrm{d}r\,F(\mathrm{d}v)\right)^{1/p^*} = \left(\int_{\mathbb{R}}|v|^p\frac{(q(v)-N)\vee 0}{q(v)}\,\pi(\mathrm{d}v)\right)^{1/p^*}$$
$$\le \left(\int_{\mathbb{R}}|z|^p\mathbb{1}_{\{q(v)\ge N\}}\,\pi(\mathrm{d}z)\right)^{1/p^*},$$
$$r_3^N \le \int_{\mathbb{R}}|z|\mathbb{1}_{\{q(v)\ge N\}}\mathbb{1}_{\{p>1,\Lambda\notin\mathcal{S}\}}\,\pi(\mathrm{d}z).$$

In most situations, there exist $(\varepsilon^N)_{N\in\mathbb{N}} \subseteq \mathbb{R}_+$ with $\varepsilon^N \to 0$ as $N\to\infty$ such that $\{q(v)\ge N\} \subseteq [-\varepsilon^N,\varepsilon^N]$. In this case, one can apply the estimates in Example 3.3.3. □

3.5 Simulation study

In this section we visualize the sample path behaviour of the stochastic heat equation from Example 3.3.3 via a simulation study, using MATLAB programs from [40]. We

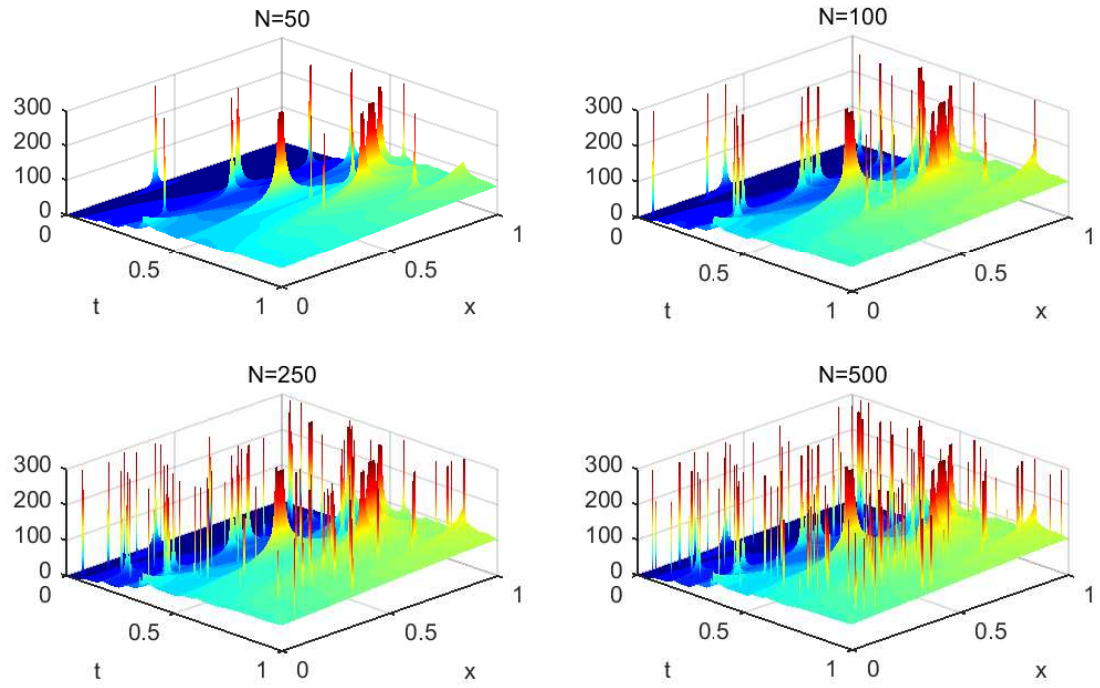

Figure 3.1: Successive approximations of Y as given in (3.5.1) via Bondesson's method in dimension 1 for $(t,x) \in [0,1] \times [0,1]$ with $N \in \{50, 100, 250, 500\}$ jumps in the region $[0,1] \times [-1,2]$

take Λ to be a Lévy basis without drift, whose Lévy measure π is that of a gamma process, i.e.

$$\pi(\mathrm{d}z) = \gamma z^{-1} \exp(-\lambda z) \mathbb{1}_{\{z>0\}} \, \mathrm{d}z$$

with two parameters $\gamma, \lambda > 0$. In the figures below their values are always $\gamma = 10$ and $\lambda = 0.1$. Furthermore, we set $Y_0 \equiv 0$ and $\sigma \equiv 1$. Especially the latter choice simplifies the subsequent discussion a lot, but none of the issues we address below relies on this assumption. Thus, the process we would like to simulate is

$$Y(t,x) = \int_0^t \int_{\mathbb{R}^d} g(t-s, x-y) \, \Lambda(\mathrm{d}s, \mathrm{d}y), \quad (t,x) \in \mathbb{R}_+ \times \mathbb{R}^d, \tag{3.5.1}$$

with g being the heat kernel given in (3.2.13). In order to understand the path properties of Y, it is important to notice that g is smooth on the whole $\mathbb{R}_+ \times \mathbb{R}^d$ except at the origin where it explodes. More precisely, for every $t \in (0, \infty)$ the function $x \mapsto g(t,x)$ is the Gaussian density with mean 0 and variance $2t$, which

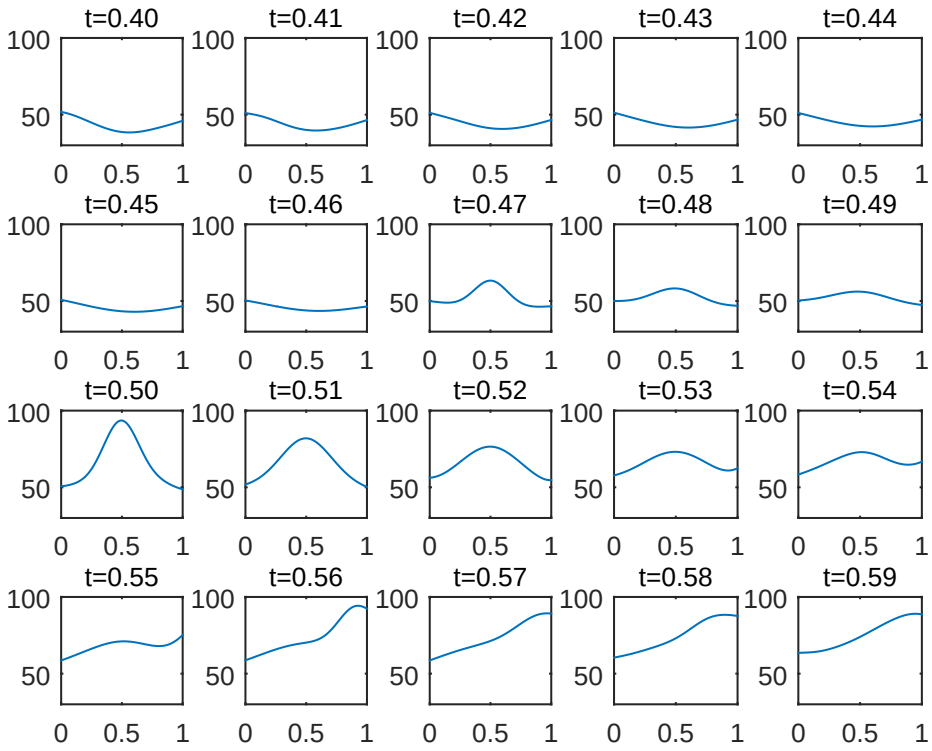

Figure 3.2: Several t-sections of the realization of Y shown in Figure 3.1 with $N = 500$

is smooth and assumes its maximum at 0. Also, for every $x \neq 0$, the function $t \mapsto g(t, x)$ is smooth (also at $t = 0$), with maximum at $t = \|x\|^2/(2d)$. However, if $x = 0$, then $g(t, 0) = (4\pi t)^{-d/2}$ has a singularity at $t = 0$.

These analytical properties have direct consequences on the sample paths of Y. When Λ is of compound Poisson type, that is, has only finitely many atoms on compact sets, it can be readily seen from (3.5.1) that the evolution of Y after a jump J at (τ, ξ) follows the shape of the heat kernel until a next jump arrives. In particular, for $x = \xi$, $Y(t, x)$ jumps to infinity at τ, and decays in t like $J(4\pi(t - \tau))^{-d/2}$ afterwards. But for every $x \neq \xi$, the evolution $t \mapsto Y(t, x)$ is *smooth* at $t = \tau$. In fact, it first starts to increase until $t = \tau + \|x - \xi\|^2/(2d)$ and then decays again. As a consequence, in space dimension 1, the space–time plot of Y shows a basically smoothly evolving path, only interrupted with slim poles at the jump locations of Λ; see the case $N = 50$ in Figure 3.1. However, when Λ has infinite activity, that is,

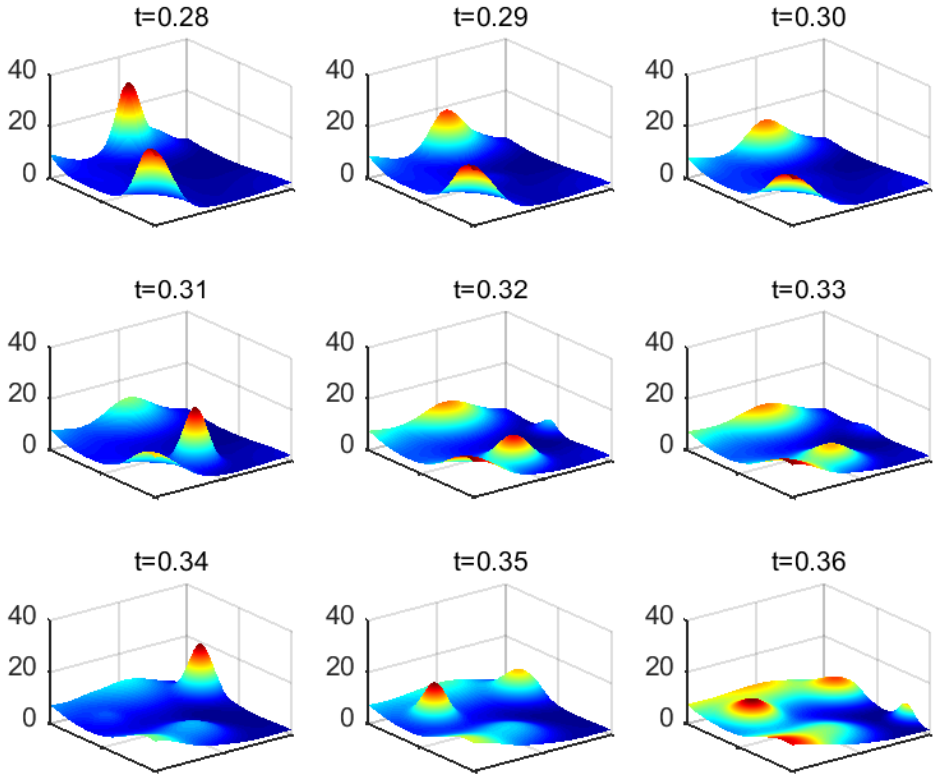

Figure 3.3: Several t-sections in the region $[-1,1]^2$ of a realization of Y in dimension 2 by Bondesson's method with $N = 500$ jumps within $[0,1] \times [-2,2]^2$

has infinitely many jumps on any non-empty open set, then it is known from [122, Thm. 4] that on any such set Y is unbounded, at least with positive probability. Therefore, the space–time plots of the approximations of Y with finitely many jumps must be treated with caution: in the limiting situation, no smooth area exists any more, but there will be a dense subset of singularities on the plane, which is in line with Figure 3.1.

Another interesting observation, however, is the following: if we consider a countable number of x- or t-sections of Y (for $x \in \mathbb{R}^d$, the x-section of Y is given by the function $t \mapsto Y(t,x)$; for $t \in \mathbb{R}_+$, the t-section of Y is the function $x \mapsto Y(t,x)$), then it is shown in [125, Sect. 2] that these are continuous with probability one. Intuitively, this is possible because a.s. the sections never hit a jump (although they are arbitrarily close). For instance, Figures 3.2 and 3.3 show t-sections of a realization of (3.5.1) in one, respectively two space dimensions. So as long as we only take

countably many "measurements", we do not observe the space–time singularities of Y but only its relatively regular sections. In theory, this also includes the x-sections of the process Y. But if we plot them for one space dimension as in Figure 3.4, one would conjecture from the simulation that they exhibit jumps in time. However, this is *not* true: the jump-like appearance of the x-sections are due to the fact that $g(\cdot, x)$ resembles a discontinuous function at $t = 0$ for small x. Of course, it follows right from the definition (3.2.13) that all x-sections of g are smooth everywhere.

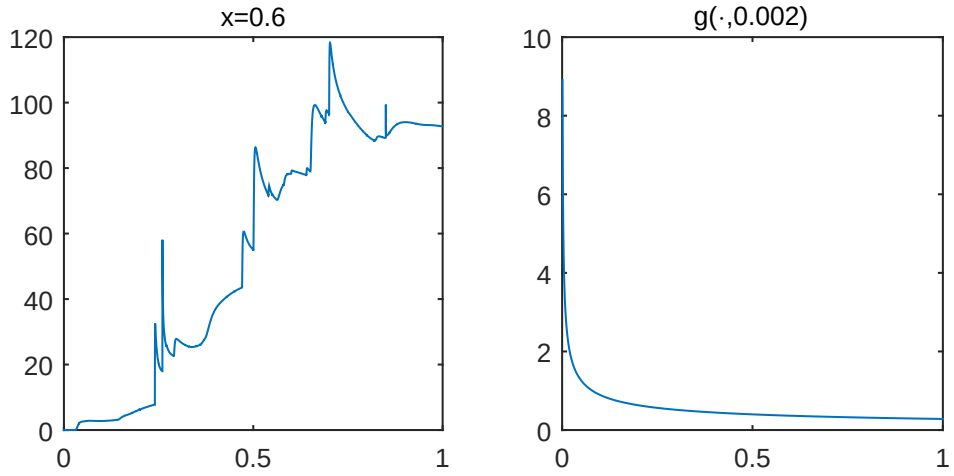

Figure 3.4: The x-section of the realization of Y as in Figure 3.1 with $N = 500$ at $x = 0.6$ and the heat kernel $g(\cdot, x)$ at $x = 0.002$

Remark 3.5.1 Throughout this Chapter we have taken the stochastic heat equation as our key example for illustration purposes. Obviously, all theory developed in Sections 3.3 and 3.4 holds under much more general assumptions, namely under A1–A7. However, the path behaviour of SVEs as examined through a simulation in this section heavily depends on the chosen kernel. While the path properties above are typical of kernels that originate from parabolic type SPDEs, or more generally, that are smooth everywhere except for an explosive singularity at the origin, a completely different picture arises when we consider bounded kernels. We refer to [18] and the references therein for examples of such kernels in different fields of application. In the case of a simple Ornstein-Uhlenbeck type kernel, a simulation study is carried out in [40, Chap. 5]. □

3.6 Conclusion and outlook

In Sections 3.3 and 3.4 we have presented two simulation algorithms for the SVE (3.1.1): Algorithms 3.3.1 and 3.4.2. In Theorems 3.3.2 and 3.4.3 we have determined the rate of convergence of the approximations Y^N to Y in the L^p-sense. If these rates are small enough, we have also proved a.s. convergence. Although the theoretical analysis of both schemes lead to quite similar results regarding their convergence behaviour, there are important differences which will decide on whether the one or the other method is preferable in concrete situations. For the first method of truncating the small jumps to work, one must be able to efficiently simulate from the truncated Lévy measure $\pi/\pi([-\varepsilon,\varepsilon]^{\mathrm{C}})$ for small ε. For the second method, which relies on series representations, the main challenge is to choose H and F in a way such that H is explicitly known *and* F can be easily simulated from. For instance, if one uses LePage's method (see Example 3.4.1), then $F = (\delta_{-1} + \delta_1)/2$ is easily simulated, but for H, which is given by the generalized inverse tails of the underlying Lévy measure, maybe no tractable expression exists.

Finally, let us comment on further generalizations of our results. Throughout this Chapter, we have assumed that the driving noise Λ is a homogeneous Lévy basis, i.e. satisfies (3.2.4). In fact, we have introduced this condition only for the sake of simplicity: with a straightforward adjustment, all results obtained in this Chapter also hold for time- and space-varying (but deterministic) characteristics. Another issue is the finite time perspective which we have taken up for our analysis. An interesting question would be under which conditions (3.1.1) has a stationary solution, and in this case, whether one can simulate from it. Sufficient conditions for the existence and uniqueness of stationary solutions to (3.1.1) are determined in Theorem 2.4.8. Under these conditions, the methods used to derive Theorems 3.3.2 and 3.4.3 can indeed be extended to the case of infinite time horizon. We leave the details to the reader at this point.

At last, also the hypothesis that Λ is of pure-jump type can be weakened. If Λ has an additional drift (including the case where Λ has locally infinite variation and is not symmetric) but still no Gaussian part, the approximations Y^N in (3.3.4) or (3.4.3) will contain a further term that is a Volterra integral with respect to the Lebesgue measure. So each time in between two simulated jumps, a deterministic Volterra equation has to be solved numerically, which boils down to a deterministic PDE in the case where G comes from an SPDE. For this subject, there exists a huge

literature, which is, of course, also very different to the stochastic case as considered above. If Λ also contains a Gaussian part, then one has to apply techniques from the papers cited in Section 3.1 and ours simultaneously. We content ourselves with referring to [139], who numerically analyzes a Volterra equation driven by a drift plus a Brownian motion. Finally, let us remark that if $p = 2$ (in particular, G must be square-integrable), it is possible for some Lévy bases to improve the results of Section 3.3 if we do not neglect the small jumps completely but approximate them via a Gaussian noise with the same variance, cf. [7] in the case of Lévy processes.

Chapter 4:
Superposition of COGARCH processes

4.1 Introduction

GARCH models have been used throughout the last decades to model returns sampled at regular intervals on stocks, currencies and other assets. They capture many of the stylized features of such data; e.g. heavy tails, volatility clustering and dependence without correlation. Also because of their interesting probabilistic properties as solutions to stochastic recurrence equations, they have attracted research by probabilists and statisticians; e.g. [68]. Various attempts have been made to capture the stylized features of financial time series using continuous-time models. The interest in continuous-time models originates in the current wide-spread availability of irregularly spaced and high-frequency data. There was a long debate, whether price and volatility fluctuations are caused by jumps or not. This question was answered convincingly in previous years by Jacod and collaborators, who developed sophisticated statistical tools to extract jumps of price and volatility out of high-frequency data (cf. [2, 81, 82] and references therein).

A prominent continuous-time stochastic volatility model was proposed by Barndorff-Nielsen and Shephard in [14], where the volatility process V and the martingale part of the logarithmic asset price G satisfy the equations

$$\begin{aligned} \mathrm{d}V_t &= -\lambda V_t \,\mathrm{d}t + \mathrm{d}L_{\lambda t}, \\ \mathrm{d}G_t &= \sqrt{V_t}\,\mathrm{d}W_t + \rho\,\mathrm{d}\widetilde{L}_{\lambda t}, \end{aligned} \tag{4.1.1}$$

where $\lambda > 0$, $\rho \leq 0$, $L = (L_t)_{t\geq 0}$ is a non-decreasing Lévy process with compensated version $\widetilde{L}$ and $W = (W_t)_{t\geq 0}$ is a standard Brownian motion independent of L. The volatility process V is taken to be the stationary solution of (4.1.1), in other words,

a stationary Lévy-driven Ornstein-Uhlenbeck (OU) process. In this model, price jumps are modelled by (scaled) upwards jumps in the volatility.

It was noticed early on that the exponential autocovariance function of the OU process may be too restrictive. Two suggestions have been made to allow for more flexibility in the autocovariance function: Barndorff-Nielsen [11] suggested to replace V by a superposition of such processes (called supOU process), which yields more flexible monotone autocovariance functions. It is defined as

$$V_t = \int_{(-\infty,t]} \int_{(0,\infty)} \mathrm{e}^{-\lambda(t-s)}\, \Lambda(\mathrm{d}s, \mathrm{d}\lambda), \quad t \in \mathbb{R}, \tag{4.1.2}$$

where Λ is an independently scattered infinitely divisible random measure, also called Lévy basis. Superpositions of CARMA processes can be defined analogously; cf. [14] and Remark 1.5.3. As shown in e.g. [62, Prop. 2.6], supOU models can also model long-range dependence for specific superposition measures.

On the other hand, [34, 134] suggested higher-order Lévy-driven CARMA models, which also allow for non-monotone autocovariance functions. The drawback of both model classes is their linearity and its consequences towards the stylized features of financial data. For instance, linear models inherit their distributions from that of the Lévy increments in a linear way. As a consequence, only when the driving Lévy process has heavy-tailed (regularly varying) increments, they model high-level volatility clusters; cf. [61, Prop. 5]. Moreover, in contrast to empirical findings (cf. [82]), these models allow only for negative price jumps coupled to the jumps in the volatility.

A continuous-time GARCH (COGARCH) model has been introduced in [88] with volatility process V and martingale part of the logarithmic asset price given by

$$\begin{aligned} \mathrm{d}V_t &= (\beta - \eta V_t)\, \mathrm{d}t + V_{t-} \varphi\, \mathrm{d}[L, L]^{\mathrm{d}}_t, \\ \mathrm{d}G_t &= \sqrt{V_{t-}}\, \mathrm{d}L_t, \end{aligned} \tag{4.1.3}$$

where $\beta, \eta, \varphi > 0$ and L is an arbitrary mean-zero Lévy process. The volatility process V is taken to be the stationary solution of (4.1.3). This model satisfies all stylized features of financial prices, exactly as the GARCH model for low frequency data. The drawback of an exponentially decreasing covariance function has been taken care of by higher-order models; cf. [35], like generalizing from OU to CARMA.

All models mentioned above have price jumps exactly at the times when the volatility jumps, since their prices are driven by the same Lévy process. Moreover,

with the exception of the supOU/supCARMA process, all jump sizes in volatility and price exhibit a fixed deterministic relationship; cf. [82]. As this is not very realistic, multi-factor models are needed. In this Chapter we want to construct such a multi-factor model, based on the COGARCH.

In contrast to the OU or CARMA models, the COGARCH model is defined as a stochastic integral with stochastic integrand. But also in this framework there is a canonical way to construct a superposition.

Starting by the fact that the ratio of volatility jumps and squared price jumps is always equal to φ in the COGARCH model, we randomize this scale parameter φ. There are various ways how to do this in a meaningful way, and we present three different possibilities, all leading to multi-factor COGARCH models. Our three models have different qualitative behaviour. For instance, the first supCOGARCH allows for jumps in the volatility, which do not necessarily lead to jumps in the price process. On the other hand, for certain choices of the distribution of the random parameter φ, the third supCOGARCH model allows for jumps in the price without having a jump in the volatility. More properties will be reported.

An interesting feature is that some of the presented new supCOGARCH volatility processes can be written in terms of a so-called *ambit process*, which has been introduced in [12] in the context of turbulence modelling. In our context the ambit process has a stochastic integrand, which is not independent of the integrator. This implies that we are no longer in the framework of [119]. Moreover, since COGARCH models are heavy-tailed, having possibly not even a second finite moment, the theory presented in [136] is also not applicable. Instead we need the concept presented in Section 1, which allows to integrate stochastic processes with respect to a Lévy basis in the generality needed for our supCOGARCH models.

This Chapter is organized as follows. In Section 4.2, we recall the COGARCH model and give a short summary of Lévy bases. In Section 4.3, we present three different superpositions of COGARCH volatility processes. For each of the three models we give necessary and sufficient conditions for strict stationarity and derive the second-order structure of the stationary process. The superpositions allow for more flexible autocorrelation structures than the COGARCH model (Propositions 4.3.4, 4.3.12 and 4.3.18). However, the stationary distributions of the supCOGARCH processes preserve the Pareto-like tails of the COGARCH process (Propositions 4.3.5, 4.3.13 and 4.3.19). Section 4.4 is devoted to the corresponding price processes and the second-order properties of their stationary increments. Again, main characteris-

tics of the COGARCH are preserved like the uncorrelated increments but positively correlated squared increments (Theorems 4.4.1, 4.4.2 and 4.4.3). Nevertheless, each of the supCOGARCH models has its specific characteristics as highlighted in Section 4.5. Furthermore, for all three models there is no longer a deterministic relationship between the jump sizes in volatility and price. Although we concentrate on the probabilistic properties of our new models, statistical issues are shortly addressed here. Finally, Section 4.6 contains the proofs of our results.

4.2 Notation and preliminaries

By the Lévy-Khintchine formula (e.g. [126, Thm. 8.1]) the *characteristic exponent* of a real-valued Lévy process $X = (X_t)_{t\geq 0}$ is given by

$$\psi_X(u) := \log \mathbb{E}\left[\mathrm{e}^{\mathrm{i}uX_1}\right] = \mathrm{i}\gamma_X u - \frac{1}{2}\sigma_X^2 u^2 + \int_{\mathbb{R}} (\mathrm{e}^{\mathrm{i}uy} - 1 - \mathrm{i}uy\mathbb{1}_{\{|y|\leq 1\}})\,\nu_X(\mathrm{d}y), \quad u \in \mathbb{R},$$

where $(\gamma_X, \sigma_X^2, \nu_X)$ is the *characteristic triplet* of X with *Lévy measure* ν_X satisfying $\nu_X(\{0\}) = 0$ and $\int_{\mathbb{R}} 1 \wedge |y|^2\,\nu_X(\mathrm{d}y) < \infty$. If additionally $\int_{|y|\leq 1} |y|\,\nu_X(\mathrm{d}y) < \infty$, we may also write the characteristic exponent in the form

$$\psi_X(u) = \mathrm{i}\gamma_X^0 u - \frac{1}{2}\sigma_X^2 u^2 + \int_{\mathbb{R}} (\mathrm{e}^{\mathrm{i}uy} - 1)\,\nu_X(\mathrm{d}y), \quad u \in \mathbb{R},$$

and call γ_X^0 the *drift* of X. This is in particular the case for subordinators, i.e. Lévy processes with increasing sample paths. We also recall that the *quadratic variation process* of the Lévy process X is given by

$$[X,X]_t := \sigma_X^2 t + [X,X]_t^{\mathrm{d}} := \sigma_X^2 t + \sum_{0<s\leq t} (\Delta X_s)^2, \quad t \geq 0,$$

where $[X,X]^{\mathrm{d}}$ is called the *pure-jump part* of $[X,X]$.

Every Lévy process $(X_t)_{t\geq 0}$ can be extended to a *two-sided Lévy process* $(X_t)_{t\in\mathbb{R}}$ by setting $X_t = -X'_{-t-}$, $t < 0$, for some i.i.d. copy X' of X. We say that $(X_t)_{t\in\mathbb{R}}$ has characteristic triplet $(\gamma_X, \sigma_X^2, \nu_X)$ if $(X_t)_{t\geq 0}$ has characteristic triplet $(\gamma_X, \sigma_X^2, \nu_X)$.

Throughout we use the notation $\mathbb{R}_+ = (0,\infty)$, $\mathbb{R}_- = (-\infty, 0)$ and $\mathbb{N}_0 = \mathbb{N} \cup \{0\}$.

4.2.1 The COGARCH model

Let $(L_t)_{t\geq 0}$ be a Lévy process with characteristic triplet $(\gamma_L, \sigma_L^2, \nu_L)$ and define

$$S_t := [L,L]_t^{\mathrm{d}} = \sum_{0<s\leq t} (\Delta L_s)^2, \quad t \geq 0. \tag{4.2.1}$$

Then $(S_t)_{t\geq 0}$ is a subordinator without drift and its Lévy measure ν_S is the image measure of ν_L under the transformation $y \mapsto y^2$. For $\eta > 0$ and $\varphi \geq 0$ define another Lévy process by

$$X_t^{\varphi} = \eta t - \sum_{0<s\leq t} \log(1 + \varphi \Delta S_s), \quad t \geq 0, \tag{4.2.2}$$

which is completely determined by S (and hence by L). Then X^{φ} has characteristic triplet $(\eta, 0, \nu_{X^{\varphi}})$, where $\nu_{X^{\varphi}}$ is the image measure of ν_S under the mapping $y \mapsto -\log(1 + \varphi y)$, and is therefore a spectrally negative Lévy process, i.e. it only has negative jumps. For $t \geq 0$ we have

$$\mathbb{E}[\mathrm{e}^{-uX_t^{\varphi}}] = \mathrm{e}^{t\Psi(u,\varphi)} \quad \text{with} \quad \Psi(u,\varphi) = -\eta u + \int_{\mathbb{R}_+} ((1+\varphi y)^u - 1)\, \nu_S(\mathrm{d}y), \tag{4.2.3}$$

where, whenever $\varphi > 0$, we have $\mathbb{E}[\mathrm{e}^{-uX_t^{\varphi}}] < \infty$ for $u > 0$ for some $t > 0$ or, equivalently, for all $t > 0$ if and only if $\mathbb{E}[S_1^u] < \infty$ [88, Lemma 4.1]. In particular, if $\mathbb{E}[S_1] < \infty$ or $\mathbb{E}[S_1^2] < \infty$, respectively, we have from [126, Ex. 25.12]

$$\Psi(1,\varphi) = \varphi \mathbb{E}[S_1] - \eta \quad \text{and} \quad \Psi(2,\varphi) = 2\varphi \mathbb{E}[S_1] + \varphi^2 \mathrm{Var}[S_1] - 2\eta. \tag{4.2.4}$$

Recall from [88] that the *COGARCH (volatility) process* driven by the Lévy process L (or the subordinator S) with parameter φ is given by

$$V_t^{\varphi} = \mathrm{e}^{-X_t^{\varphi}} \left(V_0^{\varphi} + \beta \int_{(0,t]} \mathrm{e}^{X_s^{\varphi}}\, \mathrm{d}s \right), \quad t \geq 0, \tag{4.2.5}$$

where $\beta > 0$ is a constant and V_0^{φ} is a nonnegative random variable, independent of $(S_t)_{t\geq 0}$.

Moreover, the COGARCH volatility process V^{φ} is a special case of a generalized Ornstein-Uhlenbeck process (cf. [24, 97]) and is the solution of the SDE

$$\mathrm{d}V_t^{\varphi} = (\beta - \eta V_t^{\varphi})\, \mathrm{d}t + V_{t-}^{\varphi} \varphi\, \mathrm{d}S_t = V_{t-}^{\varphi}(\varphi \mathrm{d}S_t - \eta\, \mathrm{d}t) + \beta\, \mathrm{d}t, \quad t \geq 0. \tag{4.2.6}$$

It admits the integral representation

$$V_t^{\varphi} = V_0^{\varphi} + \beta t - \eta \int_{(0,t]} V_s^{\varphi}\, \mathrm{d}s + \sum_{0<s\leq t} V_{s-}^{\varphi} \varphi \Delta S_s, \quad t \geq 0. \tag{4.2.7}$$

The corresponding *price process* or *integrated COGARCH process* is then defined as

$$G_t = \int_{(0,t]} \sqrt{V_{s-}^{\varphi}}\, \mathrm{d}L_s, \quad t \geq 0. \tag{4.2.8}$$

4.2.2 Stationary COGARCH processes

By [88, Thm. 3.1], the process defined in (4.2.5) or equivalently in (4.2.7) has a strictly stationary distribution if and only if

$$\int_{\mathbb{R}_+} \log(1+\varphi y)\,\nu_S(\mathrm{d}y) = \int_{\mathbb{R}} \log(1+\varphi y^2)\,\nu_L(\mathrm{d}y) < \eta. \tag{4.2.9}$$

In this case, the stationary distribution of the COGARCH process is given by the distribution of $V_\infty^\varphi := \beta \int_{\mathbb{R}_+} \mathrm{e}^{-X_s^\varphi}\,\mathrm{d}s$. Note that for $\varphi = 0$, the stationary COGARCH reduces to $V_t^0 = \beta/\eta$ for all $t \geq 0$.

In the sequel we denote by the set Φ_L all $\varphi \geq 0$ where (4.2.9) is satisfied. By monotone convergence, the left-hand side of (4.2.9) is continuous in φ and converges to $+\infty$ as $\varphi \to \infty$, which means that $\varphi_{\max} := \sup \Phi_L$ is finite and hence $\Phi_L = [0, \varphi_{\max})$.

Let us recall the moment structure of V^φ in the stationary case. It follows by direct computation from [25, Thm. 3.1] that, if $\kappa > 0$ is a constant, then

$$\mathbb{E}[S_1^{\max\{\kappa,1\}}] < \infty \quad \text{and} \quad \log \mathbb{E}\left[\mathrm{e}^{-\kappa X_1^\varphi}\right] = \Psi(\kappa,\varphi) < 0 \tag{4.2.10}$$

imply $\mathbb{E}[(V_0^\varphi)^\kappa] < \infty$. If (4.2.10) holds for $\kappa = 1$ or $\kappa = 2$, respectively, for every $t \geq 0$, $h \geq 0$ the first two moments of the stationary process V^φ are given by ([88, Cor. 4.1])

$$\mathbb{E}[V_t^\varphi] = -\frac{\beta}{\Psi(1,\varphi)} = \frac{\beta}{\eta - \varphi\mathbb{E}[S_1]}, \tag{4.2.11}$$

$$\mathbb{E}[(V_t^\varphi)^2] = \beta^2 \frac{2}{\Psi(1,\varphi)\Psi(2,\varphi)} \quad \text{and} \tag{4.2.12}$$

$$\begin{aligned}
\mathrm{Cov}[V_t^\varphi, V_{t+h}^\varphi] &= \mathrm{e}^{h\,\Psi(1,\varphi)}\mathrm{Var}[V_0^\varphi] \\
&= \mathrm{e}^{h\,\Psi(1,\varphi)}\beta^2 \left(\frac{2}{\Psi(1,\varphi)\Psi(2,\varphi)} - \frac{1}{\Psi(1,\varphi)^2}\right) \\
&= \mathrm{e}^{h\,(\varphi\mathbb{E}[S_1]-\eta)} \frac{\beta^2\varphi^2\mathrm{Var}[S_1]}{(\varphi\mathbb{E}[S_1]-\eta)^2(2\eta - 2\varphi\mathbb{E}[S_1] - \varphi^2\mathrm{Var}[S_1])}.
\end{aligned} \tag{4.2.13}$$

From (4.2.10) we have the clear picture that, although a stationary V^φ exists for all $\varphi \in \Phi_L = [0, \varphi_{\max})$, moments only exist on some subinterval, which shrinks with the increasing order of the moment. Moreover, it is known that no COGARCH process has moments of all orders [88, Prop. 4.3]. For later reference we set

$$\Phi_L^{(\kappa)} := [0, \varphi_{\max}^{(\kappa)}) \quad \text{with} \quad \varphi_{\max}^{(\kappa)} = \sup\{\varphi : \mathbb{E}[(V_0^\varphi)^\kappa] < \infty\}. \tag{4.2.14}$$

We have $0 < \varphi_{\max}^{(\kappa_2)} \leq \varphi_{\max}^{(\kappa_1)} < \varphi_{\max} < \infty$ whenever $0 < \kappa_1 \leq \kappa_2 < \infty$, that is, $\Phi_L^{(\kappa_2)} \subset \Phi_L^{(\kappa_1)} \subset \Phi_L$.

In [89] the tail behaviour of the COGARCH process is studied. In particular, it is shown that under rather weak assumptions the distribution of V_0^φ has Pareto-like tails [89, Thm. 6].

Regarding the price process G^φ in the stationary case, it is known from [88, Prop. 5.1] that G^φ has stationary increments that are uncorrelated on disjoint intervals while the squared increments are, under some technical assumptions, positively correlated, an effect which is typical for financial time series.

For later reference we extend the stationary COGARCH volatility process (4.2.5) to a two-sided process in the following way. For a two-sided Lévy process $(L_t)_{t\in\mathbb{R}}$ we obtain a two-sided subordinator $(S_t)_{t\in\mathbb{R}}$ by setting

$$S_t := \sum_{0<s\leq t} (\Delta L_s)^2, \quad t \geq 0 \quad \text{and} \quad S_t := -\sum_{t<s\leq 0} (\Delta L_s)^2, \quad t \leq 0. \tag{4.2.15}$$

Now we automatically obtain for every φ another two-sided Lévy process $(X_t^\varphi)_{t\in\mathbb{R}}$ given by

$$\begin{aligned} X_t^\varphi &= \eta t - \sum_{0<s\leq t} \log(1+\varphi\Delta S_s), \quad t \geq 0, \\ X_t^\varphi &= \eta t + \sum_{t<s\leq 0} \log(1+\varphi\Delta S_s), \quad t < 0. \end{aligned} \tag{4.2.16}$$

The two-sided COGARCH process $(V_t^\varphi)_{t\in\mathbb{R}}$ is then given by

$$V_t^\varphi := \beta \int_{(-\infty,t]} \mathrm{e}^{-(X_t^\varphi - X_s^\varphi)}\,\mathrm{d}s, \quad t \in \mathbb{R}, \tag{4.2.17}$$

and it is well defined for every $\varphi \in \Phi_L$. Obviously, the restriction of this process to $t \geq 0$ equals the process given in (4.2.5) with $V_0^\varphi := \beta \int_{(-\infty,0]} \mathrm{e}^{X_s^\varphi}\,\mathrm{d}s$ as starting random variable. Hence the two-sided COGARCH is always stationary with the same finite-dimensional distributions as the one-sided stationary COGARCH.

4.2.3 Lévy bases

Let $(\Omega, \mathcal{F}, \mathbb{F} = (\mathcal{F}_t)_{t\in\mathbb{R}}, \mathrm{P})$ be a filtered probability space satisfying the usual assumptions of completeness and right-continuity. Denote the space of all P-a.s. finite random variables by L^0, the optional (resp. predictable) σ-field by $\mathcal{O}$ (resp. $\mathcal{P}$)

and set $\tilde{\mathcal{P}} := \mathcal{P} \otimes \mathcal{B}(\mathbb{R}^d)$, where $\mathcal{B}(\mathbb{R}^d)$ is the Borel-σ-field on $\mathbb{R}^d$. Now let $(E_k)_{k\in\mathbb{N}}$ be a sequence of measurable subsets increasing to $\mathbb{R}^d$ and define $\tilde{\mathcal{P}}_{\mathrm{b}}$ as the collection of all $\tilde{\mathcal{P}}$-measurable subsets of $\Omega \times (-k, k] \times E_k$ for $k \in \mathbb{N}$. Similarly, set $\mathcal{B}_{\mathrm{b}} := \bigcup_{k=1}^{\infty} \mathcal{B}((-k, k] \times E_k)$.

In this set-up, we use the term Lévy basis as follows:

Definition 4.2.1 A *Lévy basis* on $\mathbb{R} \times \mathbb{R}^d$ is a mapping $\Lambda\colon \tilde{\mathcal{P}}_{\mathrm{b}} \to L^0$ satisfying:

(1) $\Lambda(\emptyset) = 0$ a.s.

(2) If $(A_n)_{n\in\mathbb{N}}$ are pairwise disjoint sets in $\tilde{\mathcal{P}}_{\mathrm{b}}$ whose union again lies in $\tilde{\mathcal{P}}_{\mathrm{b}}$, then

$$\Lambda\Big(\bigcup_{n=1}^{\infty} A_n\Big) = \sum_{n=1}^{\infty} \Lambda(A_n) \quad \text{a.s.}$$

(3) If $(B_n)_{n\in\mathbb{N}}$ are pairwise disjoint sets in $\mathcal{B}_{\mathrm{b}}$, then $(\Lambda(\Omega \times B_n))_{n\in\mathbb{N}}$ is a sequence of independent random variables with each of them having an infinitely divisible distribution.

(4) If $A \in \tilde{\mathcal{P}}_{\mathrm{b}}$ is a subset of $\Omega \times (-\infty, t] \times \mathbb{R}^d$ for some $t \in \mathbb{R}$, then $\Lambda(A)$ is $\mathcal{F}_t$-measurable.

(5) If $A \in \tilde{\mathcal{P}}_{\mathrm{b}}$, $t \in \mathbb{R}$ and $F \in \mathcal{F}_t$, then

$$\Lambda\Big(A \cap (F \times (t, \infty) \times \mathbb{R}^d)\Big) = \mathbb{1}_F \Lambda\Big(A \cap (\Omega \times (t, \infty) \times \mathbb{R}^d)\Big).$$

(6) For all $t \in \mathbb{R}$ and measurable $U \subset E_k$ for some $k \in \mathbb{N}$, we have $\Lambda(\Omega \times \{t\} \times U) = 0$ a.s.

In the following, we often write $\Lambda(B) = \Lambda(\Omega \times B)$ for a set $B \in \mathcal{B}_{\mathrm{b}}$. □

A natural choice for $\mathbb{F}$ is certainly the *augmented natural filtration* $\mathbb{G} = (\mathcal{G}_t)_{t\in\mathbb{R}}$ of the Lévy basis Λ, which means that for $t \in \mathbb{R}$, $\mathcal{G}_t$ is the completion of the σ-field generated by the collection of all $\Lambda(B)$ with $B \in \mathcal{B}_{\mathrm{b}}, B \subseteq (-\infty, t] \times \mathbb{R}^d$.

The first three points of Definition 4.2.1 are similar to the notion of infinitely divisible independently scattered random measures in [119]. Further we have added condition (6) because this ensures that Λ induces a jump measure μ^Λ by

$$\mu^\Lambda(\omega, \mathrm{d}t, \mathrm{d}x, \mathrm{d}y) := \sum_{s\in\mathbb{R}} \sum_{\xi\in\mathbb{R}^d} \mathbb{1}_{\{\Lambda(\{s\}\times\{\xi\})(\omega)\neq 0\}} \delta_{(s,\xi,\Lambda(\{s\}\times\{\xi\})(\omega))}(\mathrm{d}t, \mathrm{d}x, \mathrm{d}y), \quad (4.2.18)$$

where $\omega \in \Omega$ and δ stands for the Dirac measure. We will follow the usual convention of suppressing ω in the sequel. Thanks to (4) and (5), μ^Λ is an optional $\tilde{\mathcal{P}}$-σ-finite random measure in the sense of [80, Thm. II.1.8]. Therefore, the predictable compensator Π of μ^Λ is well defined.

In this Chapter, we will only consider Lévy bases Λ which are of the form

$$\Lambda(\mathrm{d}s, \mathrm{d}x) = \int_{\mathbb{R}} y \, \mu^\Lambda(\mathrm{d}s, \mathrm{d}x, \mathrm{d}y). \tag{4.2.19}$$

In addition, the predictable compensator of μ^Λ in the augmented natural filtration $\mathbb{G}$ will always be given by $\Pi(\mathrm{d}s, \mathrm{d}x, \mathrm{d}y) = \mathrm{d}s \, \pi(\mathrm{d}x) \, \nu(\mathrm{d}s)$, where π is some probability measure on $\mathbb{R}^d$ and ν the Lévy measure of a subordinator. In this particular case, if we write

$$W(s,x,y) * \mu_t^\Lambda := W * \mu_t^\Lambda := \begin{cases} \displaystyle\int_{(0,t]\times\mathbb{R}^d\times\mathbb{R}} W(s,x,y) \, \mu^\Lambda(\mathrm{d}s, \mathrm{d}x, \mathrm{d}y), & \text{if } t \geq 0, \\ \displaystyle\int_{(t,0]\times\mathbb{R}^d\times\mathbb{R}} W(s,x,y) \, \mu^\Lambda(\mathrm{d}s, \mathrm{d}x, \mathrm{d}y), & \text{if } t < 0, \end{cases}$$

for some $\mathcal{O} \otimes \mathcal{B}(\mathbb{R}^d) \otimes \mathcal{B}(\mathbb{R})$-measurable function W which is integrable with respect to μ^Λ (ω-wise as a Lebesgue integral), then we have

$$\mathbb{E}[W * \mu_t^\Lambda] = \mathbb{E}[W * \Pi_t] = \int_{(0,t]\times\mathbb{R}^d\times\mathbb{R}} \mathbb{E}[W(s,x,y)] \, \Pi(\mathrm{d}s, \mathrm{d}x, \mathrm{d}y), \quad t \geq 0, \tag{4.2.20}$$

for all integrable functions W (and similarly for $t < 0$), see [80, Thm. II.1.8]. Moreover, when taking stochastic integrals with respect to Λ, these can be expressed in terms of μ^Λ:

$$\int_{(0,t]\times\mathbb{R}^d} H(s,x) \, \Lambda(\mathrm{d}s, \mathrm{d}x) = \int_{(0,t]\times\mathbb{R}^d\times\mathbb{R}} H(s,x) y \, \mu^\Lambda(\mathrm{d}s, \mathrm{d}x, \mathrm{d}y), \quad t \geq 0,$$

for all H which are integrable with respect to Λ on $(0,t]$ (similarly for $t < 0$); see Section 1 for integrability conditions and further details on Lévy bases.

For later reference, we also introduce the pure-jump part of the quadratic variation measure of Λ defined as

$$[\Lambda, \Lambda]^{\mathrm{d}}(A) := \int_{\mathbb{R}\times\mathbb{R}^d\times\mathbb{R}} \mathbb{1}_A(t,x) y^2 \, \mu^\Lambda(\mathrm{d}t, \mathrm{d}x, \mathrm{d}y), \quad A \in \tilde{\mathcal{P}}_{\mathrm{b}}. \tag{4.2.21}$$

4.3 Superposition of COGARCH (supCOGARCH) processes

In the following three subsections we propose different approaches to construct a superposition of COGARCH processes. As seen in Equation (4.2.6), the parameters β and η only influence the continuous part of the COGARCH process, whereas φ scales its jump sizes. Since our goal is to find a model which shares the basic features of the COGARCH model but has a more flexible jump structure, we let β and η be fixed in the following three approaches and only allow the parameter φ to vary.

4.3.1 The supCOGARCH 1 volatility process

The obvious idea of defining a supCOGARCH process as a weighted integral of independent COGARCH processes with different parameters φ yields to consider

$$\bar{V}_t^{(1)} := \int_{[0,\infty)} V_t^{\varphi}\, \pi(\mathrm{d}\varphi), \quad t \geq 0, \tag{4.3.1}$$

for some probability measure π on $[0,\infty)$, where each COGARCH process V^{φ} is driven by $S^{\varphi} = [L^{\varphi}, L^{\varphi}]^{\mathrm{d}}$ and $(L^{\varphi})_{\varphi\in[0,\infty)}$ are i.i.d. copies of a canonical Lévy process L, which, together with $S = [L,L]^{\mathrm{d}}$, we only use for notational convenience. As a consequence, $(V^{\varphi})_{\varphi\in[0,\infty)}$ is a family of independent COGARCH processes such that the integral in (4.3.1) is only well defined if π has countable support. This leads to the *supCOGARCH 1 volatility process*

$$\bar{V}_t^{(1)} = \int_{[0,\infty)} V_t^{\varphi}\, \pi(\mathrm{d}\varphi) = \sum_{i=1}^{\infty} p_i V_t^{\varphi_i}, \quad t \geq 0, \tag{4.3.2}$$

where $\pi = \sum_{i=1}^{\infty} p_i \delta_{\varphi_i}$ for nonnegative weights $(p_i)_{i\in\mathbb{N}}$ with $\sum_{i=1}^{\infty} p_i = 1$.

To avoid degenerate cases we will assume throughout that

$$\bar{V}_0^{(1)} = \sum_{i=1}^{\infty} p_i V_0^{\varphi_i} < \infty \quad \text{a.s.} \tag{4.3.3}$$

Note that this does not automatically imply finiteness of the supCOGARCH process at all times unless we are in the stationary case (see below).

Remark 4.3.1 The supCOGARCH 1 process can also be written in terms of a Lévy basis. First, define a Lévy basis on $\mathbb{R}_+ \times [0,\infty)$ by

$$\Lambda^L((0,t]\times\{\varphi_i\}) := \sqrt{p_i} L_t^{\varphi_i}, \quad t \geq 0, \quad i \in \mathbb{N},$$

and $\Lambda^L(\mathbb{R} \times ([0,\infty) \setminus \bigcup_{i=1}^{\infty}\{\varphi_i\})) := 0$. Now with $\Lambda^S = [\Lambda^L, \Lambda^L]^{\mathrm{d}}$ being the pure-jump quadratic variation measure of Λ^L (in particular, $\Lambda^S((0,t] \times \{\varphi_i\}) = p_i S_t^{\varphi_i}$) and inserting (4.2.7) in (4.3.2), we see that

$$\begin{aligned}\bar{V}_t^{(1)} &= \sum_{i=1}^{\infty} p_i V_0^{\varphi_i} + \beta t - \eta \sum_{i=1}^{\infty} p_i \int_{(0,t]} V_s^{\varphi_i}\,\mathrm{d}s + \sum_{i=1}^{\infty} \int_{(0,t]} p_i \varphi_i V_{s-}^{\varphi_i}\,\mathrm{d}S_s^{\varphi_i} \\ &= \bar{V}_0^{(1)} + \beta t - \eta \int_{(0,t]} \bar{V}_s^{(1)}\,\mathrm{d}s + \int_{(0,t]} \int_{[0,\infty)} \varphi V_{s-}^{\varphi}\,\Lambda^S(\mathrm{d}s, \mathrm{d}\varphi), \quad t \geq 0. \end{aligned} \tag{4.3.4}$$

Note that for each $i \in \mathbb{N}$, V^{φ_i} is driven by S^{φ_i}. □

It follows directly from (4.3.4) that the jumps of the supCOGARCH 1 process are given by

$$\Delta \bar{V}_t^{(1)} = \sum_{i=1}^{\infty} p_i \Delta V_t^{\varphi_i} = \sum_{i=1}^{\infty} p_i V_{t-}^{\varphi_i} \varphi_i \Delta S_t^{\varphi_i} = \int_{[0,\infty)} \varphi V_{t-}^{\varphi}\,\Lambda^S(\{t\} \times \mathrm{d}\varphi), \quad t \geq 0. \tag{4.3.5}$$

Since the independent subordinators a.s. jump at different times, a.s. only one summand in (4.3.5) is nonzero at each jump time.

The following example for a probability measure π with two-point support will be carried through the three different supCOGARCH processes in this section to clarify their definitions.

Example 4.3.2 Let $\pi = p_1\delta_{\varphi_1} + p_2\delta_{\varphi_2}$ with $p_1 + p_2 = 1$ and $\varphi_1, \varphi_2 \in \mathbb{R}_+$. Then the supCOGARCH 1 process is the weighted sum of two independent COGARCH processes. More precisely, we have $\bar{V}_t^{(1)} = p_1 V_t^{\varphi_1} + p_2 V_t^{\varphi_2}$ for $t \geq 0$, where V^{φ_1} and V^{φ_2} are driven by *independent* copies of the canonical Lévy process L. From Figure 4.1, we clearly see that the supCOGARCH 1 process inherits both the jumps of V^{φ_1} and V^{φ_2}, scaled with p_1 or p_2, respectively. □

Stationarity and second-order properties of the supCOGARCH 1 process are given in the following three results. Proofs are postponed to Section 4.6.1.

Theorem 4.3.3. *Suppose that $\pi = \sum_{i=1}^{\infty} p_i \delta_{\varphi_i}$ is a probability measure on $[0,\infty)$, $\{L^{\varphi_i} : i \in \mathbb{N}\}$ a family of i.i.d. Lévy processes, $\{S^{\varphi_i} : i \in \mathbb{N}\}$ the corresponding family of subordinators and $\{V^{\varphi_i} : i \in \mathbb{N}\}$ the corresponding family of COGARCH processes. Assuming that* (4.3.3) *holds, a finite random variable $\bar{V}_0^{(1)}$ can be chosen such that $\bar{V}^{(1)}$ is strictly stationary if and only if*

$$\pi(\Phi_L) = 1. \tag{4.3.6}$$

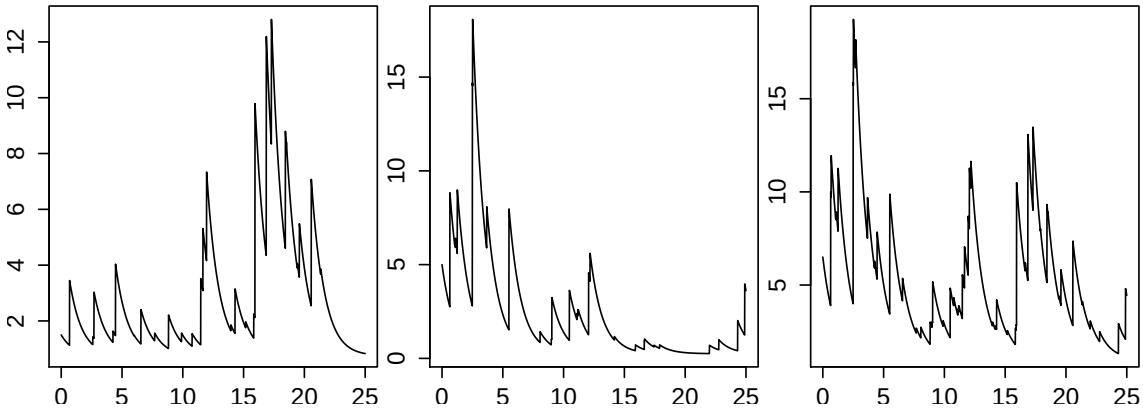

Figure 4.1: Sample paths of two independent COGARCH processes with different values for φ, scaled with the corresponding p_i, and the resulting supCOGARCH 1 process. The driving Lévy processes are independent compound Poisson processes with rate 1 and standard normal jumps. The parameters are: $\beta = 1$, $\eta = 1$, $\varphi_1 = 0.5$, $\varphi_2 = 0.95$ and $\pi = 0.75\delta_{\varphi_1} + 0.25\delta\varphi_2$, starting value is the respective mean.

In case a stationary distribution exists, it is uniquely determined by the law of

$$\bar{V}_\infty^{(1)} := \int_{\Phi_L} V_\infty^\varphi \, \pi(\mathrm{d}\varphi) = \beta \int_{\Phi_L} \int_{\mathbb{R}_+} \mathrm{e}^{-X_t^\varphi} \, \mathrm{d}t \, \pi(\mathrm{d}\varphi) = \beta \sum_{i=1}^{\infty} p_i \int_{\mathbb{R}_+} \mathrm{e}^{-X_t^{\varphi_i}} \, \mathrm{d}t. \tag{4.3.7}$$

Proposition 4.3.4. *Assume we are in the setting of Theorem 4.3.3 and let $\bar{V}^{(1)}$ be a strictly stationary solution of* (4.3.4)*. Recall the notation $\Phi_L^{(\kappa)}$ from Equation* (4.2.14)*.*

(1) Suppose that $\pi(\Phi_L^{(1)}) = 1$. Then for every $t \geq 0$,

$$\mathbb{E}[\bar{V}_t^{(1)}] = \int_{\Phi_L} \mathbb{E}[V_0^\varphi] \, \pi(\mathrm{d}\varphi) = \beta \sum_{i=1}^{\infty} \frac{p_i}{\eta - \varphi_i \mathbb{E}[S_1]}. \tag{4.3.8}$$

(2) Suppose that $\pi(\Phi_L^{(2)}) = 1$. Then for every $t \geq 0$, $h \geq 0$ we have

$$\mathrm{Var}[\bar{V}_t^{(1)}] = \sum_{i=1}^{\infty} p_i^2 \mathrm{Var}[V_0^{\varphi_i}] \quad \textit{and} \tag{4.3.9}$$

$$\mathrm{Cov}[\bar{V}_t^{(1)}, \bar{V}_{t+h}^{(1)}] = \sum_{i=1}^{\infty} p_i^2 \mathrm{Cov}[V_0^{\varphi_i}, V_h^{\varphi_i}], \tag{4.3.10}$$

with $\mathrm{Var}[V_0^{\varphi_i}]$ and $\mathrm{Cov}[V_0^{\varphi_i}, V_h^{\varphi_i}]$ as given in (4.2.12) *and* (4.2.13)*.*

Note that the quantities in (4.3.8), (4.3.9) *and* (4.3.10) *may be infinite.*

Proposition 4.3.5. *Assume we are in the setting of Theorem 4.3.3 and let* $\bar{V}^{(1)}$ *be a strictly stationary solution of* (4.3.4). *Set*

$$\bar{\varphi} := \inf\{\varphi > 0 \colon \pi((\varphi, \infty)) = 0\} \leq \varphi_{\max} < \infty$$

and assume that there exists $\bar{\kappa} > 0$ *with*

$$\mathbb{E}[S_1^{\bar{\kappa}} \log^+(S_1)] < \infty \quad \textit{and} \quad \Psi(\bar{\kappa}, \bar{\varphi}) = 0. \tag{4.3.11}$$

Then we have for $\kappa > 0$

$$\lim_{x\to\infty} x^{\kappa} \mathrm{P}[\bar{V}_0^{(1)} > x] = \begin{cases} 0 & \textit{if } \kappa < \bar{\kappa}, \\ \infty & \textit{if } \kappa > \bar{\kappa}, \end{cases}$$

while for $\kappa = \bar{\kappa}$ *there exists a constant* $C > 0$ *such that*

$$\lim_{x\to\infty} x^{\bar{\kappa}} \mathrm{P}[\bar{V}_0^{(1)} > x] = \begin{cases} C & \textit{if } \pi(\{\bar{\varphi}\}) = \bar{p} > 0, \\ 0 & \textit{if } \pi(\{\bar{\varphi}\}) = 0. \end{cases}$$

Remark 4.3.6 Recall from [89, Thm. 5] that the stationary distribution of the COGARCH V^{φ} is self-decomposable, i.e. for all $b \in (0,1)$ there exists a random variable Y_b such that $V_\infty^{\varphi} \stackrel{\mathrm{d}}{=} b(V_\infty^{\varphi})' + Y_b$ where $(V_\infty^{\varphi})'$ is an independent copy of V_∞^{φ}. Due to the fact that self-decomposability is preserved under scaling, convolution and taking limits, see e.g. [129, Prop. V.2.2], it follows directly from (4.3.7) that the stationary distribution of the supCOGARCH 1 process $\bar{V}^{(1)}$ is self-decomposable, too. □

Remark 4.3.7 Unless we are in the degenerate case $\pi = \delta_{\varphi}$ and the supCOGARCH is in fact just the COGARCH with parameter φ, the supCOGARCH process $\bar{V}^{(1)}$ is no longer a Markov process with respect to its augmented natural filtration, i.e. the smallest filtration such that $\bar{V}^{(1)}$ is adapted and which satisfies the usual hypotheses of right-continuity and completeness. But it follows directly from (4.3.4) that, letting $\mathbb{F}^{(1)} = (\mathcal{F}_t^{(1)})_{t\geq 0}$ be the augmented natural filtration of $((V_t^{\varphi_i})_{i\in\mathbb{N}})_{t\geq 0}$, we have for every measurable function $f \colon \mathbb{R}_+ \to \mathbb{R}$ and every $t \geq 0$

$$\mathbb{E}\Big[f\Big(\bar{V}_t^{(1)}\Big)\Big|\mathcal{F}_t^{(1)}\Big] = \mathbb{E}\Big[f\Big(\bar{V}_t^{(1)}\Big)\Big|(V_t^{\varphi_i})_{i\in\mathbb{N}}\Big].$$

Remark 4.3.8 In the representation $\bar{V}^{(1)} = \sum_{i=1}^{\infty} p_i V^{\varphi_i}$ a priori the φ_i do not have to be pairwise different and still the results of this section remain valid (apart from some obvious notational changes). □

4.3.2 The supCOGARCH 2 volatility process

In order to deal with uncountable superpositions, one possibility is to drop the assumption of independence, which led to the supCOGARCH 1. Hence we fix a Lévy process L, define the subordinator $(S_t)_{t\geq 0}$ by (4.2.1) and define the superposition as a weighted integral of COGARCH processes V^φ as given in (4.2.7) with different parameters φ, but all driven by the single Lévy process L, i.e. we set

$$\bar{V}_t^{(2)} := \int_{\Phi_L} V_t^\varphi \, \pi(\mathrm{d}\varphi), \quad t \geq 0,$$

for some probability measure π on the parameter space Φ_L. To ensure that $\varphi \mapsto V_t^\varphi$ is measurable at all times and in particular at time $t = 0$, we will use two-sided COGARCH processes as in (4.2.17) and define the *supCOGARCH 2 volatility process*

$$\bar{V}_t^{(2)} := \int_{\Phi_L} V_t^\varphi \, \pi(\mathrm{d}\varphi) = \beta \int_{\Phi_L} \int_{(-\infty,t]} \mathrm{e}^{-(X_t^\varphi - X_s^\varphi)} \, \mathrm{d}s \, \pi(\mathrm{d}\varphi), \quad t \in \mathbb{R}, \tag{4.3.12}$$

for $(X_t^\varphi)_{t\in\mathbb{R}}$ as given in (4.2.16). As a consequence, we have for $t \geq 0$

$$\begin{aligned}\bar{V}_t^{(2)} &= \int_{\Phi_L} V_0^\varphi \, \pi(\mathrm{d}\varphi) + \beta t - \eta \int_{\Phi_L} \int_{(0,t]} V_s^\varphi \, \mathrm{d}s \, \pi(\mathrm{d}\varphi) + \int_{\Phi_L} \int_{(0,t]} \varphi V_{s-}^\varphi \, \mathrm{d}S_s \, \pi(\mathrm{d}\varphi) \\ &= \bar{V}_0^{(2)} + \beta t - \eta \int_{(0,t]} \bar{V}_s^{(2)} \, \mathrm{d}s + \int_{(0,t]} \int_{\Phi_L} \varphi V_{s-}^\varphi \, \pi(\mathrm{d}\varphi) \, \mathrm{d}S_s. \end{aligned} \tag{4.3.13}$$

In order to ensure that (4.3.12) is finite, we always assume

$$\int_{\Phi_L} V_0^\varphi \, \pi(\mathrm{d}\varphi) < \infty. \tag{4.3.14}$$

If $\pi = \sum_{i=1}^\infty p_i \delta_{\varphi_i}$, we obviously have $\bar{V}^{(2)} = \sum_{i=1}^\infty p_i V^{\varphi_i}$ with dependent summands.

Observe that in this setting all single COGARCH processes jump at the same times and thus we have

$$\Delta \bar{V}_t^{(2)} = \int_{\Phi_L} \varphi V_{t-}^\varphi \, \pi(\mathrm{d}\varphi) \Delta S_t, \quad t \geq 0. \tag{4.3.15}$$

Example 4.3.9 (Example 4.3.2 continued) Let $\pi = p_1\delta_{\varphi_1} + p_2\delta_{\varphi_2}$ be given with $p_1 + p_2 = 1$ and $\varphi_1, \varphi_2 \in \Phi_L$. Then the supCOGARCH 2 process is the weighted sum of two COGARCH processes with parameters φ_1 and φ_2, i.e. $\bar{V}_t^{(2)} = p_1 V_t^{\varphi_1} + p_2 V_t^{\varphi_2}$. In contrast to the supCOGARCH 1 process in Example 4.3.2, V^{φ_1} and V^{φ_2} are driven by the same subordinator, say S, of the form (4.2.1). In Figure 4.2 we illustrate the

typical relationship between the original COGARCH processes and the resulting supCOGARCH 2 process. We observe that V^{φ_1}, V^{φ_2} and $\bar{V}^{(2)}$ all jump at the same times, with the jump sizes of the supCOGARCH being the weighted average jump sizes of the two COGARCH processes. □

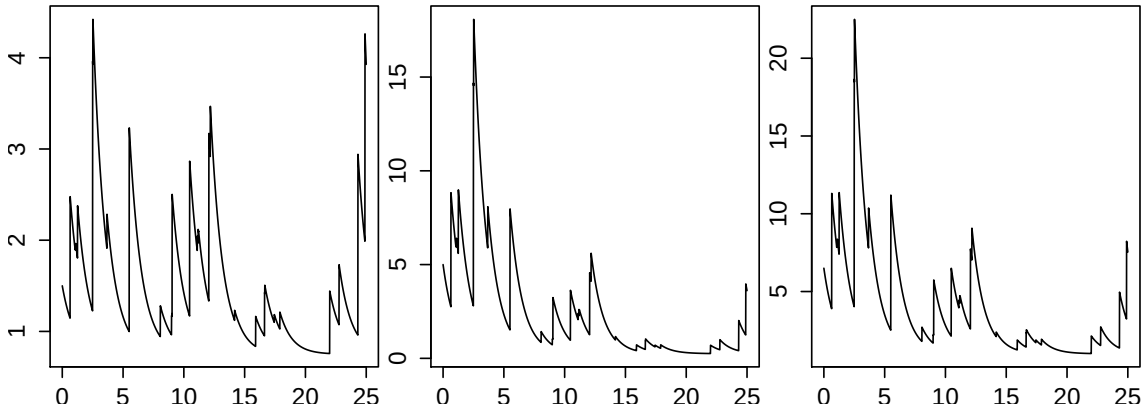

Figure 4.2: Sample paths of two COGARCH processes V^{φ_1} and V^{φ_2} with different parameters, driven by the same Lévy process L, scaled with the corresponding p_i, and the resulting supCOGARCH $\bar{V}^{(2)}$. The driving Lévy process L is a compound Poisson process with rate 1 and standard normal jumps. The parameters are the same as in Figure 4.1.

In the following we present stationarity and second-order properties of the supCOGARCH process $\bar{V}^{(2)}$. Proofs are given in Section 4.6.2.

Theorem 4.3.10. *Assume that* (4.3.14) *holds. Then* $(\bar{V}_t^{(2)})_{t\in\mathbb{R}}$ *as defined in* (4.3.12) *is strictly stationary.*

Before we can calculate the moments of the stationary supCOGARCH process $\bar{V}^{(2)}$ in Proposition 4.3.12 we need to establish covariances between single COGARCH processes with different parameters in the following proposition.

Proposition 4.3.11. *Let* $(S_t)_{t\in\mathbb{R}}$ *be a subordinator without drift, let* $\varphi, \tilde{\varphi} \in \Phi_L$ *be fixed and define the stationary two-sided COGARCH processes* $(V_t^{\varphi})_{t\in\mathbb{R}}$, $(V_t^{\tilde{\varphi}})_{t\in\mathbb{R}}$ *according to* (4.2.17). *If*

$$\mathbb{E}[S_1^2] < \infty, \quad \Psi(2,\varphi) < 0 \quad \textit{and} \quad \Psi(2,\tilde{\varphi}) < 0,$$

then $\mathbb{E}[V_t^{\varphi} V_{t+h}^{\tilde{\varphi}}] < \infty$ *for all* $t \in \mathbb{R}$ *and* $h \geq 0$. *In this case, we have for all* $t \in \mathbb{R}$ *that*

$$\mathbb{E}[V_t^{\varphi} V_t^{\tilde{\varphi}}] = \frac{\beta^2((\varphi+\tilde{\varphi})\mathbb{E}[S_1] - 2\eta)}{(\varphi\mathbb{E}[S_1]-\eta)(\tilde{\varphi}\mathbb{E}[S_1]-\eta)((\varphi+\tilde{\varphi})\mathbb{E}[S_1] + \varphi\tilde{\varphi}\mathrm{Var}[S_1] - 2\eta)}, \quad (4.3.16)$$

$$\mathrm{Cov}[V_t^{\varphi}, V_t^{\tilde{\varphi}}] = \frac{\beta^2\varphi\tilde{\varphi}\mathrm{Var}[S_1]}{(\varphi\mathbb{E}[S_1]-\eta)(\tilde{\varphi}\mathbb{E}[S_1]-\eta)(2\eta - (\varphi+\tilde{\varphi})\mathbb{E}[S_1] - \varphi\tilde{\varphi}\mathrm{Var}[S_1])}, \quad (4.3.17)$$

while for all $t \in \mathbb{R}$ *and* $h \geq 0$

$$\mathrm{Cov}[V_t^{\varphi}, V_{t+h}^{\tilde{\varphi}}] = \mathrm{e}^{h\Psi(1,\tilde{\varphi})}\mathrm{Cov}[V_0^{\varphi}, V_0^{\tilde{\varphi}}]. \quad (4.3.18)$$

Both covariances in (4.3.17) *and* (4.3.18) *are nonnegative.*

Now we can describe the covariance structure of the supCOGARCH process $\bar{V}^{(2)}$.

Proposition 4.3.12. *Let* $\bar{V}^{(2)}$ *be the strictly stationary supCOGARCH 2 process as defined in* (4.3.12). *Recall the notation* $\Phi_L^{(\kappa)}$ *from Equation* (4.2.14).

(1) Suppose that $\pi(\Phi_L^{(1)}) = 1$. *Then we have for all* $t \geq 0$

$$\mathbb{E}[\bar{V}_t^{(2)}] = \int_{\Phi_L} \mathbb{E}[V_0^{\varphi}]\,\pi(\mathrm{d}\varphi) = \beta \int_{\Phi_L} \frac{1}{\eta - \varphi\mathbb{E}[S_1]}\,\pi(\mathrm{d}\varphi). \quad (4.3.19)$$

(2) Suppose that $\pi(\Phi_L^{(2)}) = 1$. *Then for* $t \in \mathbb{R}$ *and* $h \geq 0$ *we have*

$$\mathbb{E}[(\bar{V}_t^{(2)})^2] = \int_{\Phi_L}\int_{\Phi_L} \mathbb{E}[V_0^{\varphi} V_0^{\tilde{\varphi}}]\,\pi(\mathrm{d}\varphi)\,\pi(\mathrm{d}\tilde{\varphi}), \quad (4.3.20)$$

$$\mathrm{Var}[\bar{V}_t^{(2)}] = \int_{\Phi_L}\int_{\Phi_L} \mathrm{Cov}[V_0^{\varphi}, V_0^{\tilde{\varphi}}]\,\pi(\mathrm{d}\varphi)\,\pi(\mathrm{d}\tilde{\varphi}), \quad (4.3.21)$$

$$\mathrm{Cov}[\bar{V}_t^{(2)}, \bar{V}_{t+h}^{(2)}] = \int_{\Phi_L}\int_{\Phi_L} \mathrm{Cov}[V_0^{\varphi}, V_h^{\tilde{\varphi}}]\,\pi(\mathrm{d}\varphi)\,\pi(\mathrm{d}\tilde{\varphi}), \quad (4.3.22)$$

with $\mathbb{E}[V_0^{\varphi} V_0^{\tilde{\varphi}}]$ *and* $\mathrm{Cov}[V_0^{\varphi}, V_h^{\tilde{\varphi}}]$ *as given in Proposition 4.3.11.*

Note that the quantities in (4.3.19), (4.3.20), (4.3.21) *and* (4.3.22) *may be infinite.*

The tail behaviour of $\bar{V}^{(2)}$ is similar to the tail behaviour of the supCOGARCH 1 process.

Proposition 4.3.13. *Let $\bar{V}^{(2)}$ be the strictly stationary supCOGARCH 2 process as defined in* (4.3.12). *Set $\bar{\varphi} := \inf\{\varphi > 0 \colon \pi((\varphi, \infty)) = 0\} \leq \varphi_{\max} < \infty$ and assume that there exists $\bar{\kappa} > 0$ such that* (4.3.11) *holds. Then we have for $\kappa > 0$*

$$\lim_{x\to\infty} x^{\kappa} \mathrm{P}[\bar{V}_0^{(2)} > x] = \begin{cases} 0 & \text{if } \kappa < \bar{\kappa}, \\ \infty & \text{if } \kappa > \bar{\kappa}, \end{cases}$$

while for $\kappa = \bar{\kappa}$ there exists a constant $C > 0$ such that

$$\lim_{x\to\infty} x^{\bar{\kappa}} \mathrm{P}[\bar{V}_0^{(2)} > x] = \begin{cases} C & \text{if } \pi(\{\bar{\varphi}\}) = \bar{p} > 0, \\ 0 & \text{if } \pi(\{\bar{\varphi}\}) = 0. \end{cases}$$

Remark 4.3.14 Similarly to $\bar{V}^{(1)}$, the process $\bar{V}^{(2)}$ is no Markov process with respect to its augmented natural filtration (unless in the degenerate case $\pi = \delta_{\varphi}$), but again we have a Markov property in a wide sense. More precisely, for the augmented natural filtration $\mathbb{F}^{(2)} = (\mathcal{F}_t^{(2)})_{t\geq 0}$ of $((V_t^{\varphi})_{\varphi\in\Phi_L})_{t\geq 0}$, we obtain for every measurable function $f \colon \mathbb{R}_+ \to \mathbb{R}$ and every $t \geq 0$

$$\mathbb{E}\Big[f\Big(\bar{V}_t^{(2)}\Big)\Big|\mathcal{F}_t^{(2)}\Big] = \mathbb{E}\left[f\Big(\bar{V}_t^{(2)}\Big)\Big|(V_t^{\varphi})_{\varphi\in\Phi_L}\right].$$

4.3.3 The supCOGARCH 3 volatility process

Our third superposition model invokes a Lévy basis Λ^L on $\mathbb{R} \times \Phi_L$ such that

$$L_t := \Lambda^L((0,t] \times \Phi_L), \quad t \geq 0, \quad L_t := -\Lambda^L((-t,0] \times \Phi_L), \quad t < 0,$$

exists for every $t \in \mathbb{R}$. With $\Lambda^S := [\Lambda^L, \Lambda^L]^{\mathrm{d}}$ in the sense of (4.2.21), Λ^S is of the form (4.2.19) and we assume that the predictable compensator of μ^{Λ^S} is given by $\Pi^S(\mathrm{d}t, \mathrm{d}y, \mathrm{d}\varphi) = \mathrm{d}t\, \nu_S(\mathrm{d}y)\, \pi(\mathrm{d}\varphi)$, where π is a probability measure on Φ_L and ν_S the Lévy measure of the following two-sided subordinator:

$$S_t := \Lambda^S((0,t] \times \Phi_L), \quad t \geq 0, \quad S_t := -\Lambda^S((-t,0] \times \Phi_L), \quad t < 0. \tag{4.3.23}$$

For every $\varphi \in \Phi_L$ we denote by V^{φ} the two-sided COGARCH process driven by S as in (4.2.17). The *supCOGARCH 3 volatility process* $\bar{V}^{(3)}$ is then defined by the integral equation

$$\bar{V}_t^{(3)} = \bar{V}_0^{(3)} + \beta t - \eta \int_{(0,t]} \bar{V}_s^{(3)}\, \mathrm{d}s + \int_{(0,t]} \int_{\Phi_L} \varphi V_{s-}^{\varphi}\, \Lambda^S(\mathrm{d}s, \mathrm{d}\varphi), \quad t \geq 0, \tag{4.3.24}$$

where $\bar{V}_0^{(3)}$ is some starting random variable independent of the restriction of Λ^L to $\mathbb{R}_+ \times \Phi_L$ From (4.3.24) it follows directly that

$$\Delta \bar{V}_t^{(3)} = \int_{\mathbb{R}_+ \times \Phi_L} \varphi V_{t-}^{\varphi} y \, \mu^{\Lambda^S}(\{t\}, \mathrm{d}\varphi, \mathrm{d}y), \quad t \geq 0. \tag{4.3.25}$$

We present now conditions for stationarity and calculate the second-order properties. The proofs can be found in Section 4.6.3.

Proposition 4.3.15. *The stochastic integral equation* (4.3.24) *has a unique solution given by*

$$\bar{V}_t^{(3)} = \mathrm{e}^{-\eta t} \left(\bar{V}_0^{(3)} + \beta \int_{(0,t]} \mathrm{e}^{\eta s} \, \mathrm{d}s + \int_{(0,t]} \mathrm{e}^{\eta s} \, \mathrm{d}A_s \right), \quad t \geq 0, \tag{4.3.26}$$

where

$$A_t := \int_{(0,t]} \int_{\Phi_L} \varphi V_{s-}^{\varphi} \, \Lambda^S(\mathrm{d}s, \mathrm{d}\varphi), \quad t \geq 0, \tag{4.3.27}$$

is a semimartingale with increasing sample paths, finite at every fixed $t \geq 0$.

Example 4.3.16 (Example 4.3.2 and 4.3.9 continued) Let $\pi = p_1 \delta_{\varphi_1} + p_2 \delta_{\varphi_2}$ with $p_1 + p_2 = 1$ and $\varphi_1, \varphi_2 \in \Phi_L$. As opposed to the supCOGARCH 1 process in Example 4.3.2 or the supCOGARCH 2 process in Example 4.3.9, the supCOGARCH 3 process is not the sum of two (independent or dependent) COGARCH processes. In fact, there is a subordinator S driving two COGARCH processes V^{φ_1} and V^{φ_2} and each time when S jumps, a value of φ is randomly chosen from $\{\varphi_1, \varphi_2\}$: φ takes the value φ_1 with probability p_1 and the value φ_2 with probability p_2. Now the jump size of the supCOGARCH 3 at a particular jump time of S is exactly the jump size of the COGARCH with the chosen parameter φ. If $(T_i)_{i \in \mathbb{N}}$ denote the jump times of S, we have

$$\Delta \bar{V}_{T_i}^{(3)} = \Delta V_{T_i}^{\varphi_i} = \varphi_i V_{T_i-}^{\varphi_i} \Delta S_{T_i}, \quad i \in \mathbb{N},$$

and $(\varphi_i)_{i \in \mathbb{N}}$ is an i.i.d. sequence with distribution π. Moreover, $(\varphi_i)_{i \in \mathbb{N}}$ is independent of S. This effect is illustrated in Figure 4.3. □

The next theorem establishes necessary and sufficient conditions for the existence of a stationary distribution of the supCOGARCH 3 process.

Theorem 4.3.17. *Define the supCOGARCH 3 process* $(\bar{V}_t^{(3)})_{t \geq 0}$ *by* (4.3.26). *Then a finite random variable* $\bar{V}_0^{(3)}$ *can be chosen such that* $\bar{V}^{(3)}$ *is strictly stationary if*

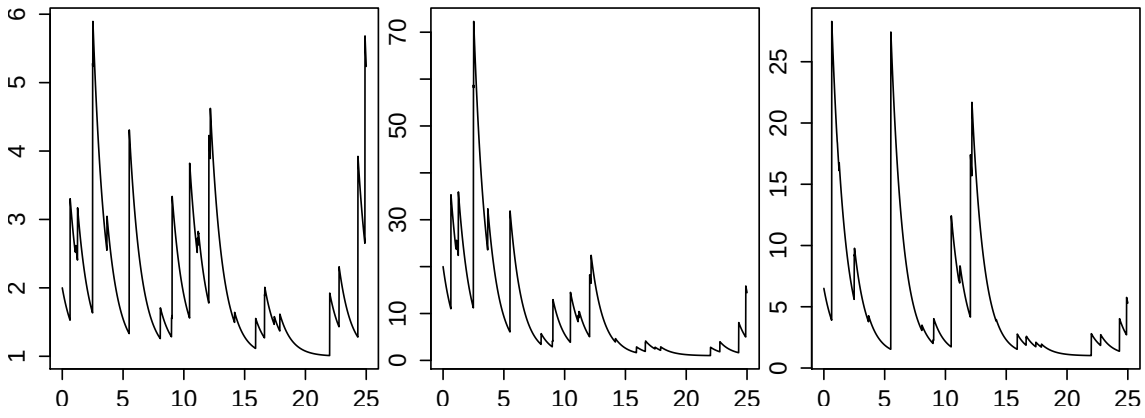

Figure 4.3: Two COGARCH processes V^{φ_1} and V^{φ_2} driven by the same Lévy process L and the resulting supCOGARCH $\bar{V}^{(3)}$. L is a compound Poisson process with rate 1 and standard normal jumps. The parameters are the same as in Figure 4.1.

and only if

$$\int_{\mathbb{R}_+}\int_{\Phi_L}\int_{\mathbb{R}_+} 1 \wedge (y\varphi V_s^{\varphi}\mathrm{e}^{-\eta s})\,\mathrm{d}s\,\pi(\mathrm{d}\varphi)\,\nu_S(\mathrm{d}y) < \infty \quad a.s. \tag{4.3.28}$$

In the case that a stationary distribution exists, it is uniquely determined by the law of $\frac{\beta}{\eta} + \int_{\mathbb{R}_+} \mathrm{e}^{-\eta s}\,\mathrm{d}A_s$. *In particular, setting* $\bar{V}_0^{(3)} := \frac{\beta}{\eta} + \int_{(-\infty,0]}\int_{\Phi_L} \mathrm{e}^{\eta s}\varphi V_{s-}^{\varphi}\,\Lambda^S(\mathrm{d}s,\mathrm{d}\varphi)$, *we obtain the two-sided stationary supCOGARCH 3 process*

$$\begin{aligned}\bar{V}_t^{(3)} &= \mathrm{e}^{-\eta t}\left(\beta\int_{(-\infty,t]}\mathrm{e}^{\eta s}\,\mathrm{d}s + \int_{(-\infty,t]}\mathrm{e}^{\eta s}\,\mathrm{d}A_s\right)\\ &= \frac{\beta}{\eta} + \int_{(-\infty,t]}\int_{\Phi_L}\mathrm{e}^{-\eta(t-s)}\varphi V_{s-}^{\varphi}\,\Lambda^S(\mathrm{d}s,\mathrm{d}\varphi)\end{aligned} \tag{4.3.29}$$

for $t \in \mathbb{R}$. *Moreover,* (4.3.28) *holds in each of the following cases:*

(1) $\pi([0,\varphi_0]) = 1$ *with some* $\varphi_0 < \varphi_{\max}$.

(2) $\pi(\Phi_L^{(\kappa)}) = 1$ *for some* $\kappa > 0$.

The second-order properties of the strictly stationary supCOGARCH 3 process are as follows.

Proposition 4.3.18. *Let* $\bar{V}^{(3)}$ *be the stationary supCOGARCH 3 process given by* (4.3.29). *Recall the notation* $\Phi_L^{(\kappa)}$ *from Equation* (4.2.14).

(1) Assume that $\pi(\Phi_L^{(1)}) = 1$. *Then for* $t \in \mathbb{R}$

$$\mathbb{E}[\bar{V}_t^{(3)}] = \int_{\Phi_L} \mathbb{E}[V_0^{\varphi}]\,\pi(\mathrm{d}\varphi) = \int_{\Phi_L} \frac{\beta}{\eta - \mathbb{E}[S_1]\varphi}\,\pi(\mathrm{d}\varphi). \tag{4.3.30}$$

(2) Assume that $\pi(\Phi_L^{(2)}) = 1$. *Then with* $\mathbb{E}[V_0^{\varphi}V_0^{\tilde{\varphi}}]$ *and* $\mathrm{Cov}[V_0^{\varphi}, V_0^{\tilde{\varphi}}]$ *as given in Proposition 4.3.11, for* $t \in \mathbb{R}$ *and* $h \geq 0$ *we have*

$$\begin{aligned}\mathbb{E}[(\bar{V}_t^{(3)})^2] = \int_{\Phi_L}\int_{\Phi_L} \Bigg(&\mathbb{E}[V_0^{\varphi}V_0^{\tilde{\varphi}}] \\ &+ \frac{\beta}{\eta}\frac{\mathrm{Var}[V_0^{\varphi}] - \mathrm{Cov}[V_0^{\varphi}, V_0^{\tilde{\varphi}}]}{\mathbb{E}[V_0^{\varphi}]}\Bigg)\,\pi(\mathrm{d}\tilde{\varphi})\,\pi(\mathrm{d}\varphi).\end{aligned} \tag{4.3.31}$$

function of $\bar{V}^{(3)}$ *is for every* $t \in \mathbb{R}$, $h \geq 0$ *given by*

$$\begin{aligned}\mathrm{Cov}[\bar{V}_t^{(3)}, \bar{V}_{t+h}^{(3)}] = \int_{\Phi_L}\int_{\Phi_L} \Bigg(&\mathrm{e}^{h\Psi(1,\varphi)}\mathrm{Cov}[V_0^{\varphi}, V_0^{\tilde{\varphi}}] \\ &+ \mathrm{e}^{-\eta h}\frac{\beta}{\eta}\frac{\mathrm{Var}[V_0^{\varphi}] - \mathrm{Cov}[V_0^{\varphi}, V_0^{\tilde{\varphi}}]}{\mathbb{E}[V_0^{\varphi}]}\Bigg)\,\pi(\mathrm{d}\tilde{\varphi})\,\pi(\mathrm{d}\varphi).\end{aligned} \tag{4.3.32}$$

Note that the quantities in (4.3.30), (4.3.31) *and* (4.3.32) *may be infinite.*

The supCOGARCH 3 process also exhibits Pareto-like tails.

Proposition 4.3.19. *Let* $\bar{V}^{(3)}$ *be the stationary supCOGARCH 3 process given by* (4.3.29). *Set* $\bar{\varphi} := \inf\{\varphi > 0 \colon \pi((\varphi, \infty)) = 0\} \leq \varphi_{\max} < \infty$ *and assume that there exists* $\bar{\kappa} > 0$ *such that* (4.3.11) *is fulfilled. Then for* $\kappa > 0$

$$\lim_{x\to\infty} x^{\kappa}\mathrm{P}[\bar{V}_0^{(3)} > x] = \begin{cases} 0 & \text{if } \kappa < \bar{\kappa}, \\ \infty & \text{if } \kappa > \bar{\kappa}, \end{cases}$$

and for $\kappa = \bar{\kappa}$ *and* $\pi(\{\bar{\varphi}\}) = 0$ *we have*

$$\lim_{x\to\infty} x^{\bar{\kappa}}\mathrm{P}[\bar{V}_0^{(3)} > x] = 0,$$

while for $\kappa = \bar{\kappa}$ *and* $\pi(\{\bar{\varphi}\}) = \bar{p} > 0$

$$0 < C_* := \liminf_{x\to\infty} x^{\bar{\kappa}}\mathrm{P}[\bar{V}_0^{(3)} > x] \leq \limsup_{x\to\infty} x^{\bar{\kappa}}\mathrm{P}[\bar{V}_0^{(3)} > x] =: C^* < \infty.$$

Remark 4.3.20 Just like $\bar{V}^{(1)}$ and $\bar{V}^{(2)}$, the process $\bar{V}^{(3)}$ is not a Markov process with respect to its augmented natural filtration (unless in the case $\pi = \delta_\varphi$), but, denoting the augmented natural filtration of $((V_t^\varphi)_{\varphi\in\Phi_L})_{t\geq 0}$ by $\mathbb{F}^{(3)} = (\mathcal{F}_t^{(3)})_{t\geq 0}$, we obtain for every measurable function $f\colon \mathbb{R}_+ \to \mathbb{R}$ and every $t \geq 0$

$$\mathbb{E}\Big[f\Big(\bar{V}_t^{(3)}\Big)\Big|(\mathcal{F}_s^{(3)})_{s\leq t}\Big] = \mathbb{E}\Big[f\Big(\bar{V}_t^{(3)}\Big)\Big|(V_t^\varphi)_{\varphi\in\Phi_L}\Big].$$

4.4 The price processes

Recall that in the COGARCH model, or its discrete-time analogue, the GARCH model (cf. [30]), the driving noises for volatility and price processes are the same (4.2.8). In this section, we suggest and investigate price processes corresponding to the supCOGARCH volatility processes. All proofs can be found in Section 4.6.4.

4.4.1 The integrated supCOGARCH 1 price process

For the supCOGARCH 1 volatility process $\bar{V}^{(1)}$ as defined in Section 4.3.1, there is no canonical choice for a price process, since a whole sequence $(L^{\varphi_i})_{i\in\mathbb{N}}$ of Lévy processes is used in its definition. Hence a priori any function of this sequence is a reasonable candidate for the driver in the price process. As a simple example we take the Lévy process L^{φ_1} as integrator; i.e. we define

$$G_t^{(1)} := \int_{(0,t]} \sqrt{\bar{V}_{s-}^{(1)}}\,\mathrm{d}L_s^{\varphi_1}, \quad t \geq 0. \tag{4.4.1}$$

It is an interesting observation that this process not only allows for common jumps of volatility and price (as it is usual in the standard COGARCH model), but also for jumps only in the volatility and not in the price process. There is evidence that this happens in real data (cf. [81]).

It is obvious from the definition that, if $(\bar{V}_t^{(1)})_{t\geq 0}$ is strictly stationary, then $(G_t^{(1)})_{t\geq 0}$ has stationary increments. Furthermore, its second-order structure is comparable to that of the integrated COGARCH process [88, Prop. 5.1].

Theorem 4.4.1. *Let $\bar{V}^{(1)} = \sum_{i=1}^\infty p_i V^{\varphi_i}$, $\varphi_i \in \Phi_L$, be a stationary supCOGARCH 1 process as defined in Section 4.3.1, where each V^{φ_i} is driven by $S^{\varphi_i} = [L^{\varphi_i}, L^{\varphi_i}]^{\mathrm{d}}$*

and $(L^{\varphi_i})_{i\in\mathbb{N}}$ *are i.i.d. copies of a Lévy processes* L *with zero mean. Define the price process* $G^{(1)}$ *by* (4.4.1) *and set*

$$\Delta^r G_t^{(1)} := G_{t+r}^{(1)} - G_t^{(1)} = \int_{(t,t+r]} \sqrt{\bar{V}_{s-}^{(1)}}\, \mathrm{d}L_s^{\varphi_1}, \quad t \geq 0, \quad r > 0.$$

Recall the notation $\Phi_L^{(\kappa)}$ *from Equation* (4.2.14) *and that the support of* π *is countable in this case.*

(1) If π *has support in* $\Phi_L^{(1/2)}$*, then* $\mathbb{E}[\Delta^r G_t^{(1)}] = 0$ *for all* $t \geq 0$ *and* $r > 0$*.*

(2) If further $\mathbb{E}[L_1^2] < \infty$ *and* π *has support in* $\Phi_L^{(1)}$*, then for* $t \in \mathbb{R}$*,* $h \geq r > 0$

$$\mathbb{E}[(\Delta^r G_t^{(1)})^2] = r\mathbb{E}[L_1^2]\mathbb{E}[\bar{V}_0^{(1)}] = r\mathbb{E}[L_1^2] \int_{\Phi_L^{(1)}} \frac{\beta}{\eta - \varphi(\mathbb{E}[L_1^2] - \sigma_L^2)}\, \pi(\mathrm{d}\varphi),$$

$$\mathrm{Cov}[\Delta^r G_t^{(1)}, \Delta^r G_{t+h}^{(1)}] = 0.$$

(3) Assume further that $\mathbb{E}[L_1^4] < \infty$*,* $\int_{\mathbb{R}} y^3\, \nu_L(\mathrm{d}y) = 0$ *and that* $\pi \neq \delta_0$ *has support in* $\Phi_L^{(2)}$*. Then for* $t \in \mathbb{R}$*,* $h \geq r > 0$

$$\begin{aligned}
&\mathrm{Cov}[(\Delta^r G_t^{(1)})^2, (\Delta^r G_{t+h}^{(1)})^2] \\
&\quad = \mathbb{E}[L_1^2] \int_{\Phi_L^{(2)}} \frac{\mathrm{e}^{h\Psi(1,\varphi)} - \mathrm{e}^{(h-r)\Psi(1,\varphi)}}{\Psi(1,\varphi)} \mathrm{Cov}[(\Delta^r G_0^{(1)})^2, V_r^{\varphi}]\, \pi(\mathrm{d}\varphi) \\
&\quad > 0.
\end{aligned}$$

4.4.2 The integrated supCOGARCH 2 price process

Let $(L_t)_{t\in\mathbb{R}}$ be a two-sided Lévy process, define the subordinator S by (4.2.15) and let $(\bar{V}_t^{(2)})_{t\in\mathbb{R}}$ be the supCOGARCH 2 process driven by S as defined in Section 4.3.2. In view of the standard definition of the integrated COGARCH price process (4.2.8) it makes sense to define the *integrated supCOGARCH 2 price process* by

$$\mathrm{d}G_t^{(2)} := \sqrt{\bar{V}_{t-}^{(2)}}\, \mathrm{d}L_t, \quad G_0^{(2)} = 0, \quad t \in \mathbb{R}. \tag{4.4.2}$$

Hence, as in the standard COGARCH model, the process $G^{(2)}$ jumps exactly when the volatility $\bar{V}^{(2)}$ jumps. Also $(G_t^{(2)})_{t\in\mathbb{R}}$ has stationary increments if $(\bar{V}_t^{(2)})_{t\in\mathbb{R}}$ is strictly stationary. The integrated supCOGARCH 2 process has the same second-order structure as the integrated supCOGARCH 1 process and, hence, as the integrated COGARCH process as shown in the following.

Theorem 4.4.2. *Suppose that the two-sided Lévy process L has expectation 0, define S by (4.2.15), the supCOGARCH volatility $\bar{V}^{(2)}$ as in Section 4.3.2 with $\pi(\Phi_L) = 1$ and the process $G^{(2)}$ by (4.4.2). Set*

$$\Delta^r G_t^{(2)} := G_{t+r}^{(2)} - G_t^{(2)} = \int_{(t,t+r]} \sqrt{\bar{V}_{s-}^{(2)}}\, \mathrm{d}L_s, \quad t \in \mathbb{R}, \quad r > 0.$$

(1) If π has support in $\Phi_L^{(1/2)}$, then $\mathbb{E}[\Delta^r G_t^{(2)}] = 0$ for all $t \in \mathbb{R}$ and $r > 0$.

(2) If further $\mathbb{E}[L_1^2] < \infty$ and π has support in $\Phi_L^{(1)}$, then for $t \in \mathbb{R}$, $h \geq r > 0$

$$\mathbb{E}[(\Delta^r G_t^{(2)})^2] = r\mathbb{E}[L_1^2]\mathbb{E}[\bar{V}_0^{(2)}] = r\mathbb{E}[L_1^2] \int_{\Phi_L^{(1)}} \frac{\beta}{\eta - \varphi(\mathbb{E}[L_1^2] - \sigma_L^2)}\, \pi(\mathrm{d}\varphi),$$

$$\mathrm{Cov}[\Delta^r G_t^{(2)}, \Delta^r G_{t+h}^{(2)}] = 0.$$

(3) Assume further that $\mathbb{E}[L_1^4] < \infty$, $\int_{\mathbb{R}} y^3\, \nu_L(\mathrm{d}y) = 0$ and $\pi \neq \delta_0$ has support in $\Phi_L^{(2)}$. Then for $t \in \mathbb{R}$, $h \geq r > 0$

$$\begin{aligned}
&\mathrm{Cov}[(\Delta^r G_t^{(2)})^2, (\Delta^r G_{t+h}^{(2)})^2] \\
&\quad = \mathbb{E}[L_1^2] \int_{\Phi_L^{(2)}} \frac{\mathrm{e}^{h\Psi(1,\varphi)} - \mathrm{e}^{(h-r)\Psi(1,\varphi)}}{\Psi(1,\varphi)} \mathrm{Cov}[(\Delta^r G_0^{(2)})^2, V_r^{\varphi}]\, \pi(\mathrm{d}\varphi) \\
&\quad > 0.
\end{aligned}$$

4.4.3 The integrated supCOGARCH 3 price process

As in the case of the supCOGARCH 2 there is a canonical choice for the driving noise in the price process of the supCOGARCH 3. With L being a Lévy process and $V^{(3)}$ the stationary supCOGARCH 3 as defined in (4.3.29), we define the *integrated supCOGARCH 3 price process* by

$$G_t^{(3)} := \int_{(0,t]} \sqrt{\bar{V}_{t-}^{(3)}}\, \mathrm{d}L_t, \quad t \geq 0. \tag{4.4.3}$$

Evidently, $G^{(3)}$ has stationary increments and, if $\pi(\{0\}) = 0$, it jumps at exactly the times when $\bar{V}^{(3)}$ jumps. However, whenever $\pi(\{0\}) > 0$, the supCOGARCH 3 model features price jumps without volatility jumps, a behaviour attested by the empirical findings of [81]. The second-order structure of $G^{(3)}$ is calculated in the following theorem.

Theorem 4.4.3. *Suppose that L is a Lévy process with expectation 0 and that $\pi(\Phi_L) = 1$. Define $V^{(3)}$ by (4.3.29) and set*

$$\Delta^r G_t^{(3)} := G_{t+r}^{(3)} - G_t^{(3)} = \int_{(t,t+r]} \sqrt{\bar{V}_{s-}^{(3)}}\, \mathrm{d}L_s, \quad t \geq 0, \quad r > 0.$$

(1) If π has support in $\Phi_L^{(1/2)}$, then $\mathbb{E}[\Delta^r G_t^{(3)}] = 0$ for all $t \geq 0$ and $r > 0$.

(2) If further $\mathbb{E}[L_1^2] < \infty$ and π has support in $\Phi_L^{(1)}$, then for $t \geq 0$ and $h \geq r > 0$

$$\mathbb{E}[(\Delta^r G_t^{(3)})^2] = r\mathbb{E}[L_1^2]\mathbb{E}[\bar{V}_0^{(3)}] = r\mathbb{E}[L_1^2] \int_{\Phi_L^{(1)}} \frac{\beta}{\eta - \varphi(\mathbb{E}[L_1^2] - \sigma_L^2)}\, \pi(\mathrm{d}\varphi),$$

$$\mathrm{Cov}[\Delta^r G_t^{(3)}, \Delta^r G_{t+h}^{(3)}] = 0.$$

(3) Assume further that $\mathbb{E}[L_1^4] < \infty$, $\int_{\mathbb{R}} y^3\, \nu_L(\mathrm{d}y) = 0$ and $\pi \neq \delta_0$ has support in $\Phi_L^{(2)}$. Then for $t \geq 0$ and $h \geq r > 0$

$$\begin{aligned}
&\mathrm{Cov}[(\Delta^r G_t^{(3)})^2, (\Delta^r G_{t+h}^{(3)})^2] \\
&= \mathbb{E}[L_1^2] \left[\frac{\mathrm{e}^{-\eta(h-r)} - \mathrm{e}^{-\eta h}}{\eta} \mathrm{Cov}[(\Delta^r G_0^{(3)})^2, \bar{V}_r^{(3)}] \right. \\
&\quad \left. + \int_{\Phi_L^{(2)}} \left(\frac{\mathrm{e}^{h\Psi(1,\varphi)} - \mathrm{e}^{(h-r)\Psi(1,\varphi)}}{\Psi(1,\varphi)} + \frac{\mathrm{e}^{-\eta h} - \mathrm{e}^{-\eta(h-r)}}{\eta} \right) \mathrm{Cov}[(\Delta^r G_0^{(3)})^2, V_r^{\varphi}]\, \pi(\mathrm{d}\varphi) \right] \\
&> 0.
\end{aligned}$$

4.5 Comparison and conclusions

This section is devoted to highlight the analogies and differences between the three supCOGARCH processes, and to compare them to the standard COGARCH process. First note that in all three models, setting $\pi = \delta_\varphi$ for $\varphi \in \Phi_L$ yields the standard COGARCH process $(V_t^\varphi)_{t\geq 0}$ as defined in (4.2.5). Hence it seems natural that some features of the COGARCH process are preserved under superpositioning. The next remark summarizes the most important properties.

Remark 4.5.1 (1) Comparing the autocovariance functions of the supCOGARCH volatility processes (cf. (4.3.10), (4.3.22) and (4.3.32)) to those of the COGARCH (cf. (4.2.13)), we find for large lags h exponential decay in all three supCOGARCH models, but allowing for more flexibility than in the COGARCH model for small and medium lags.

(2) The important property of Pareto-like tails of the stationary distribution of a COGARCH process [89, Thm. 6] persists as shown in Propositions 4.3.5, 4.3.13 and 4.3.19.

(3) Another similarity is given in the behaviour between jumps, where the COGARCH process exhibits exponential decay [89, Prop. 2]. More precisely, assuming that $\bar{V}^{(1)}$, $\bar{V}^{(2)}$ and $\bar{V}^{(3)}$ only have finitely many jumps on compact intervals, and fixing two consecutive jump times $T_j < T_{j+1}$, we obtain for $i \in \{1,2,3\}$ and $t \in (T_j, T_{j+1})$

$$\frac{\mathrm{d}}{\mathrm{d}t}\bar{V}_t^{(i)} = \beta - \eta \bar{V}_t^{(i)}, \quad \bar{V}_t^{(i)} = \frac{\beta}{\eta} + \left(\bar{V}_{T_j}^{(i)} - \frac{\beta}{\eta}\right) \mathrm{e}^{-\eta(t-T_j)}.$$

(4) An important difference between the supCOGARCH processes and the COGARCH process is the jump behaviour. This is highlighted in Corollary 4.5.2 and Example 4.5.3.

(5) In general, all supCOGARCH models have common jumps in volatility and price as it is characteristic for the COGARCH model. Additionally, the supCOGARCH 1 model also features volatility jumps without price jumps and the supCOGARCH 3, if $\pi(\{0\}) > 0$, also price jumps without volatility jumps. Moreover, if we replace L^{φ_1} in (4.4.1) by a (finite or infinite) linear combination of $(L^{\varphi_i})_{i\in\mathbb{N}}$, we can control the proportion of common volatility and price jumps to sole volatility jumps in the supCOGARCH 1 model.

(6) Our three models have different degrees of randomness in the following sense. The supCOGARCH 1 is defined via a sequence of independent Lévy processes. So by the adjustment of π there is an arbitrary degree of randomness in the model. The supCOGARCH 2 model has only one single source of randomness, namely the driving Lévy process. Finally, the supCOGARCH 3 incorporates two sources of randomness: one originating from the Lévy process $L = \Lambda^L((0,\cdot]\times\Phi_L)$ and one from the sequence $(\varphi_i)_{i\in\mathbb{N}}$ chosen at the jump times of L. □

One of the motivations for this study was the observation made in [82] that for a COGARCH process (V^φ, G^φ) there is always a deterministic relationship between volatility jumps and price jumps given by

$$q_T^\varphi := \frac{\phi(V_{T-}^\varphi, V_T^\varphi)}{\psi(G_{T-}^\varphi, G_T^\varphi)} \equiv \varphi$$

for every jump time T of the driving Lévy process and deterministic functions

$$\psi(x,y) = (y-x)^2, \quad \phi(x,y) = y - x. \tag{4.5.1}$$

From the following corollary, which is a direct consequence of the respective definitions, we see that for all three supCOGARCH models such a deterministic functional relationship between volatility and price jumps is no longer present.

Corollary 4.5.2. *Let T be a jump time of L^{φ_1} for the supCOGARCH 1, and a jump time of L for the supCOGARCH 2 and 3. Furthermore, define the numbers $\bar{\varphi} := \inf\{\varphi > 0 \colon \pi((\varphi,\infty)) = 0\}$ and $\underline{\varphi} := \sup\{\varphi > 0 \colon \pi((0,\varphi)) = 0\}$ (using the convention $\sup\emptyset := 0$, $\inf\emptyset := \infty$).*

(1) We have

$$\Delta\bar{V}_T^{(1)} = p_1\varphi_1 V_{T-}^{\varphi_1}(\Delta L_T^{\varphi_1})^2, \qquad \Delta G_T^{(1)} = \sqrt{\sum_{i=1}^{\infty} p_i V_{T-}^{\varphi_i}}\Delta L_T^{\varphi_1}, \tag{4.5.2}$$

$$\Delta\bar{V}_T^{(2)} = \int_{\Phi_L} \varphi V_{T-}^{\varphi}\,\pi(\mathrm{d}\varphi)(\Delta L_T)^2, \qquad \Delta G_T^{(2)} = \sqrt{\int_{\Phi_L} V_{T-}^{\varphi}\,\pi(\mathrm{d}\varphi)}\Delta L_T, \tag{4.5.3}$$

$$\Delta\bar{V}_T^{(3)} = \varphi_T V_{T-}^{\varphi_T}(\Delta L_T)^2, \qquad \Delta G_T^{(3)} = \sqrt{\bar{V}_{T-}^{(3)}}\Delta L_T, \tag{4.5.4}$$

where in the last line φ_T is a random variable which has distribution π and is independent of L.

(2) Define

$$q_T^{(i)} := \frac{\phi(\bar{V}_{T-}^{(i)}, \bar{V}_T^{(i)})}{\psi(G_{T-}^{(i)}, G_T^{(i)})} \tag{4.5.5}$$

for $i = 1,2,3$ with ϕ and ψ given in (4.5.1). Then we have

$$q_T^{(1)} \le \bar{\varphi} \quad \text{and} \quad \underline{\varphi} \le q_T^{(2)} \le \bar{\varphi};$$

moreover, if $\varphi_T = \bar{\varphi}$ (resp. $\varphi_T = \underline{\varphi}$), we have

$$q_j^{(3)} \ge \bar{\varphi} \quad (\text{resp. } q_j^{(3)} \le \underline{\varphi}).$$

Example 4.5.3 (Example 4.3.2, 4.3.9 and 4.3.16 continued) Let us compare the jumps in the supCOGARCH volatility processes for $\pi = p_1\delta_{\varphi_1} + p_2\delta_{\varphi_2}$ with $p_1 + p_2 = 1$ and $\varphi_1, \varphi_2 \in \Phi_L$: We see from (4.5.2) that in the supCOGARCH 1

model a squared jump of L^{φ_i} is always scaled with $p_i\varphi_i V_{t-}^{\varphi_i}$ and, hence, the parameter φ_i as well as the weight p_i take part in the scaling. In contrast, defining $S^{\varphi_i} = \Lambda^S((0,\cdot] \times \{\varphi_i\})$ for $i = 1, 2$ in the case of the supCOGARCH 3 process, each jump of $S = S^{\varphi_1} + S^{\varphi_2} = [L, L]^{\mathrm{d}}$ is scaled with $\varphi_1 V_{t-}^{\varphi_1}$ or $\varphi_2 V_{t-}^{\varphi_2}$, depending on whether S^1 or S^2 actually jumps. Here the probabilities p_i do not influence the scaling of the jump, but the intensity of the driving processes S^{φ_i}. In other words, the p_i determine the probability for the value φ_i to be chosen at a specific jump time. Finally, for the supCOGARCH 2 process, the jump size of the subordinator $S = [L, L]^{\mathrm{d}}$ is always scaled with $p_1\varphi_1 V_{t-}^{\varphi_1} + p_2\varphi_2 V_{t-}^{\varphi_2}$, so all weights and parameters are involved. □

Simulation results

To illustrate the theoretical findings above, we present simulations of the different supCOGARCH volatility processes as well as the different price processes in Figures 4.4 and 4.5 below. As Lévy process L we choose a variance gamma process arising through time changing a standard Brownian motion by an independent gamma process with mean and variance 1.

Note that we have chosen different parameters for the simulations presented in Figures 4.4 and 4.5, respectively, in order to better visualize the differences between the three volatility and the three price processes.

To illustrate the profound difference between the COGARCH and the three supCOGARCH models with reference to (4.5.1), we also compute $q^{(1)}, q^{(2)}$ and $q^{(3)}$ as defined in (4.5.5) for the jump times of the simulation in Figure 4.5. The histograms of $\log q^{(i)}$ are given in Figure 4.6. We see that both the supCOGARCH 1 and 2 exhibit a certain interval of values for $\log q^{(1)}$ and $\log q^{(2)}$. As we would expect from Corollary 4.5.2, both $\log q^{(1)}$ and $\log q^{(2)}$ are bounded from above by $\log \varphi_2$, but only $\log q^{(2)}$ is bounded from below by $\log \varphi_1$ whereas the $\log q^{(1)}$ has a relatively long tail on the negative side. Also, in general, the values of $q^{(1)}$ tend to be smaller than those of $q^{(2)}$. This is due to the fact that at a common jump time of volatility and price, the volatility jump size is the sum of two terms for the supCOGARCH 2 but only a single term for the supCOGARCH 1 (see (4.5.3) and (4.5.2)). As a result, the nominator in (4.5.5) is usually smaller for the supCOGARCH 1 than for the supCOGARCH 2. Finally, again in coincidence with Corollary 4.5.2, the supCOGARCH 3 shows two disjoint intervals for the values of $q^{(3)}$, corresponding to the two different values of φ chosen for the superposition.

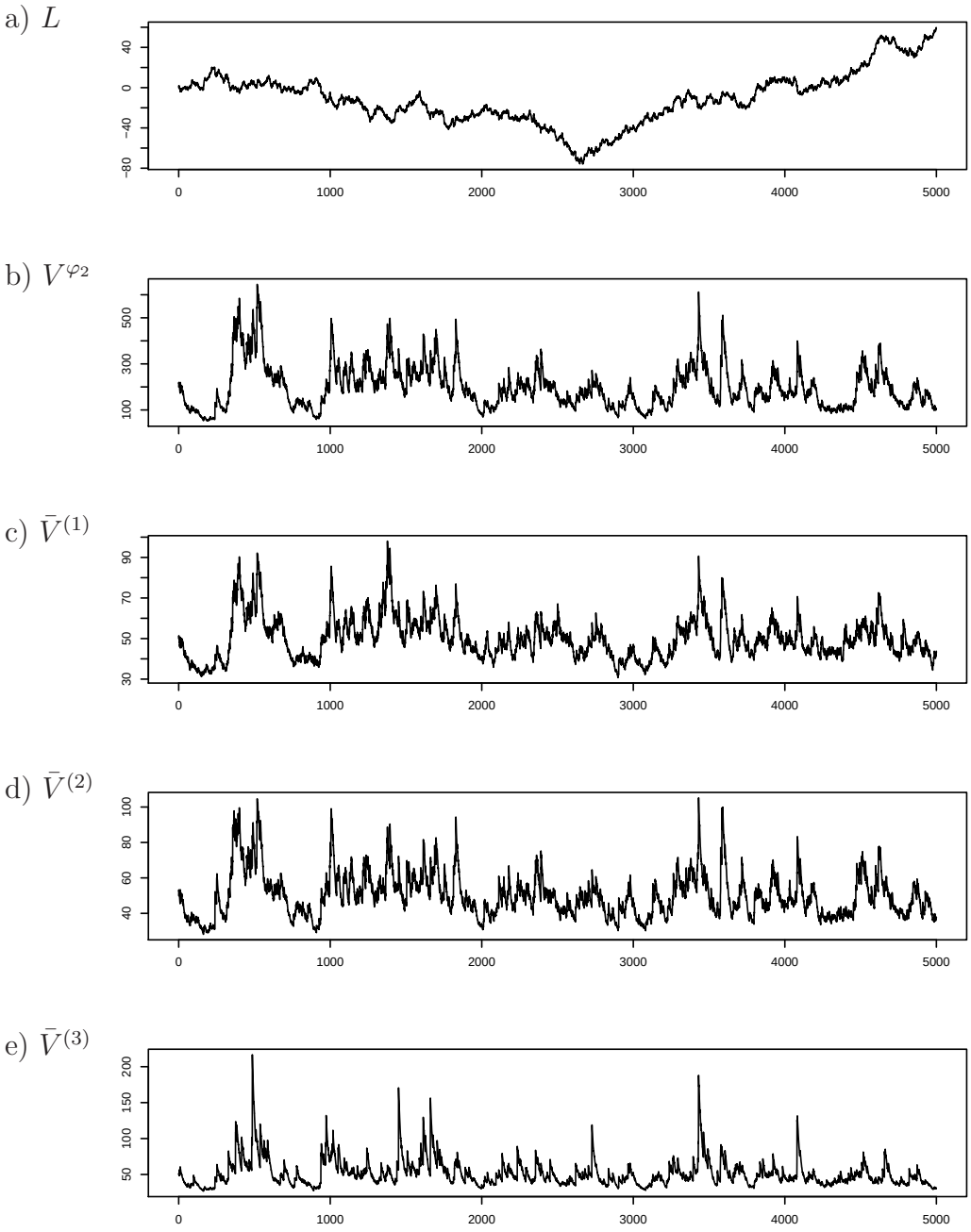

Figure 4.4: The parameters are: $\beta = 1$, $\eta = 0.05$, $\pi = 0.9\delta_{\varphi_1} + 0.1\delta_{\varphi_2}$, $\varphi_1 = 0.02$, $\varphi_2 = 0.045$, starting value is the mean; a) L is a variance gamma process with mean 0 and variance 1; b) COGARCH process driven by L with parameter φ_2; c) supCOGARCH process $\bar{V}^{(1)}$ where V^{φ_2} is driven by L and V^{φ_1} is driven by an independent copy of L; d) supCOGARCH process $\bar{V}^{(2)}$ driven by L; e) supCOGARCH process $\bar{V}^{(3)}$ driven by L.

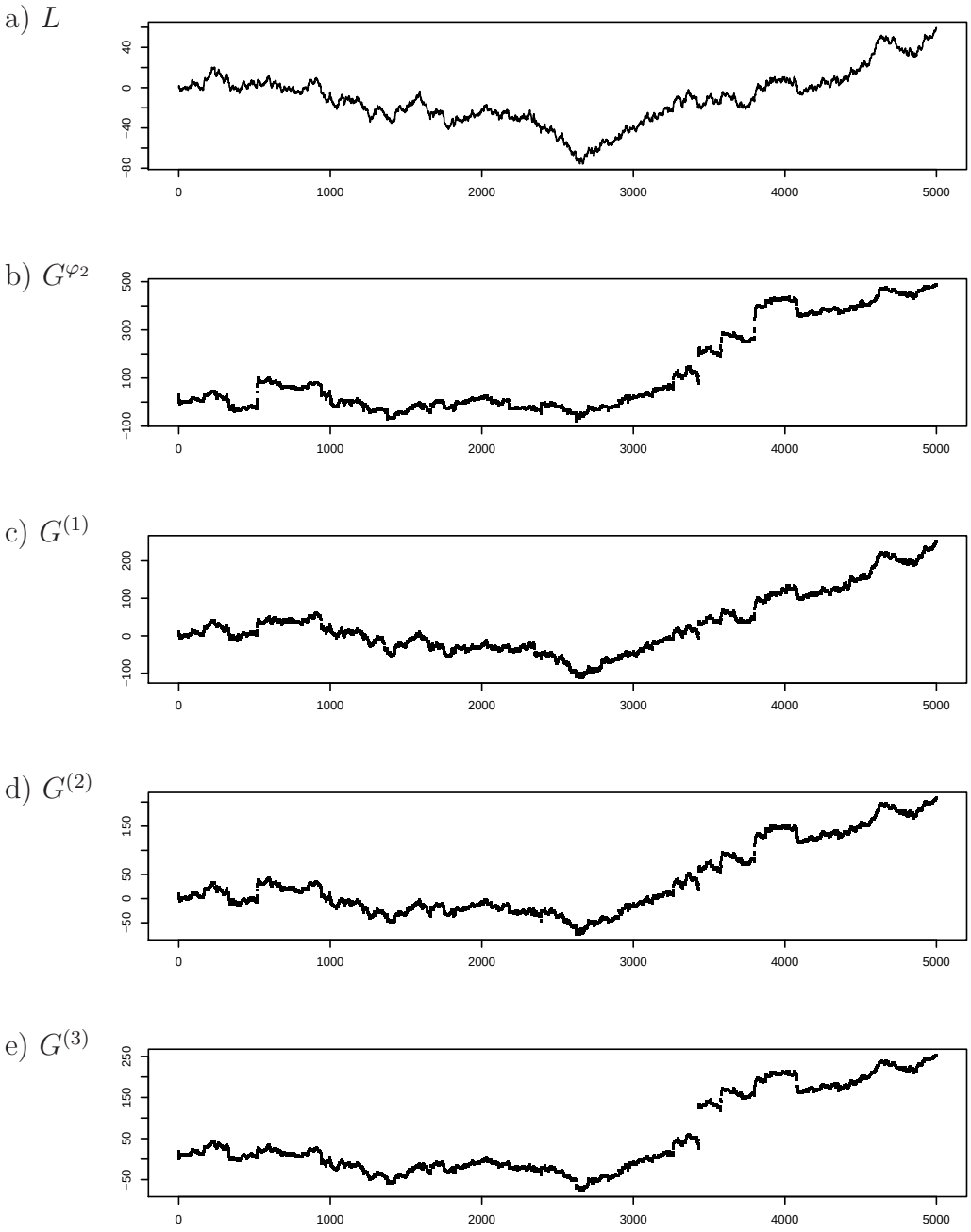

Figure 4.5: The parameters are: $\beta = 1$, $\eta = 1$, $\pi = 0.9\delta_{\varphi_1} + 0.1\delta_{\varphi_2}$, $\varphi_1 = 0.5$, $\varphi_2 = 0.995$; a) L is the same Lévy process as in Figure 4.4; b) COGARCH price process driven by L with parameter φ_2; c), d) and f) supCOGARCH price processes $G^{(1)}$, $G^{(2)}$ and $G^{(3)}$ driven by L.

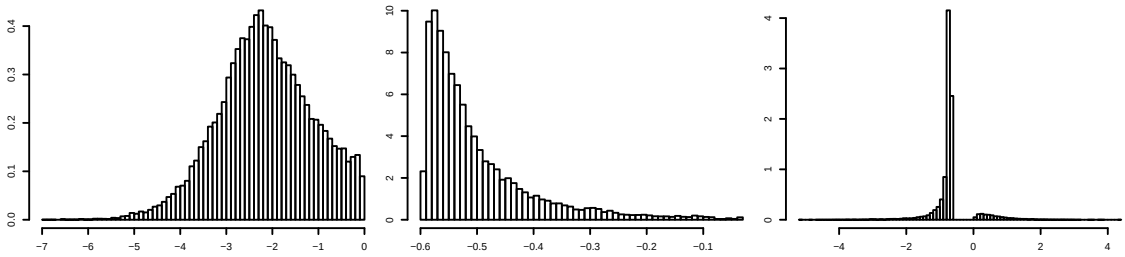

Figure 4.6: The pictures (from left to right) show the histograms for $\log(q^{(1)})$, $\log(q^{(2)})$ and $\log(q^{(3)})$.

Estimation

A thorough investigation of the statistical analysis of the supCOGARCH processes via parameter estimation goes far beyond the scope of this Chapter. Nevertheless let us shortly comment on the main task, namely the estimation of the superposition measure π, for which no standard estimation procedure is available as it is typical for multifactor models.

In the case of the supOU stochastic volatility model, several attempts have been made to infer the underlying superposition measure. For example, assuming the form $\pi = \sum_{i=1}^{K} p_i \delta_{\varphi_i}$ for some known $K \in \mathbb{N}$, in [14, 15] a least-square fit of the autocovariance function is employed. In [128] a generalized method of moments is used to estimate the supOU model under the hypothesis of a gamma distribution for π. Further, in [72, 73] a Bayesian nonparametric approach is proposed in the case that π is a discrete or continuous measure, respectively. Whether and how these approaches, or the estimation procedures for the COGARCH model mentioned in the Introduction can be adapted to the supCOGARCH case, is open.

4.6 Proofs and auxiliary results

4.6.1 Proofs for Section 4.3.1

Proof of Theorem 4.3.3. First assume that (4.3.6) holds. Then we know that each COGARCH process V^{φ_i} in the representation $\bar{V}^{(1)} = \sum_{i=1}^{\infty} p_i V^{\varphi_i}$ admits a unique stationary distribution given by the law of $V_\infty^{\varphi_i} = \beta \int_{\mathbb{R}_+} \mathrm{e}^{-X_t^{\varphi_i}} \, \mathrm{d}t$ and that by choosing $V_0^{\varphi_i} \stackrel{\mathrm{d}}{=} V_\infty^{\varphi_i}$ independently of S^{φ_i}, the corresponding COGARCH process

V^{φ_i} is strictly stationary. Thus setting $\bar{V}_0^{(1)} := \sum_{i\in\mathbb{N}} p_i V_0^{\varphi_i}$, $\bar{V}^{(1)}$ becomes strictly stationary as shown in the following.

Assume for a moment that π has finite support. Then for every $n \in \mathbb{N}$, $h > 0$ and $0 \le t_1 < t_2 < \ldots < t_n$ we can use the independence of $(V^{\varphi_i})_{i\in\mathbb{N}}$ to obtain

$$\begin{aligned}(\bar{V}_{t_1}^{(1)}, \ldots, \bar{V}_{t_n}^{(1)}) &= \left(\sum_{i=1}^m p_i V_{t_1}^{\varphi_i}, \ldots, \sum_{i=1}^m p_i V_{t_n}^{\varphi_i}\right) = \sum_{i=1}^m p_i (V_{t_1}^{\varphi_i}, \ldots, V_{t_n}^{\varphi_i}) \\ &\overset{\mathrm{d}}{=} \sum_{i=1}^m p_i (V_{t_1+h}^{\varphi_i}, \ldots, V_{t_n+h}^{\varphi_i}) = (\bar{V}_{t_1+h}^{(1)}, \ldots, \bar{V}_{t_n+h}^{(1)}).\end{aligned}$$

Due to the fact that $\sum_{i=1}^m p_i V_t^{\varphi_i}$ is strictly increasing in m, the case for π having countable support follows now by a standard monotonicity argument.

Conversely, assume that (4.3.6) is violated, i.e. there exists a φ_j with $\pi(\{\varphi_j\}) > 0$ such that V^{φ_j} has no stationary distribution. Then by [88, Thm. 3.1] $V_t^{\varphi_j}$ converges in probability to ∞ as $t \to \infty$. This yields that also $\bar{V}_t^{(1)} = p_j V_t^{\varphi_j} + \sum_{i=1, i\neq j}^\infty p_i V_t^{\varphi_i}$ converges in probability to ∞ as $t \to \infty$ since $\sum_{i=1,i\neq j}^\infty p_i V_t^{\varphi_i}$ is nonnegative. Hence $\bar{V}_t^{(1)}$ cannot be strictly stationary. □

Proof of Proposition 4.3.4. The moment conditions as well as the formulas for expectation and covariance follow directly from (4.3.7) together with the corresponding results for the COGARCH process (4.2.10), (4.2.11) and (4.2.13) observing that all appearing processes are strictly positive. □

Proof of Proposition 4.3.5. Throughout this proof we slightly change our notation as follows. Given i.i.d. subordinators $(S^i)_{i\in\mathbb{N}}$, we denote the COGARCH process driven by S^i with parameter $\varphi > 0$ by $V^{i,\varphi}$ such that we have $\bar{V}^{(1)} = \sum_{i=1}^\infty p_i V^{i,\varphi_i}$. If $\kappa < \bar{\kappa}$, then we know by the definition of Ψ in (4.2.3) and [88, Lemma 4.1(d)] that for every $\varphi \in (0, \bar{\varphi}]$ there exists a unique constant $\kappa(\varphi) > 0$ which satisfies (4.3.11) with $\bar{\kappa}$ replaced by $\kappa(\varphi)$ and such that $\kappa(\varphi)$ is strictly decreasing in φ. Moreover, as shown in [89, Thm. 6], for each $i \in \mathbb{N}$ the tail of $V^{i,\varphi}$ is asymptotically equivalent to $C(\varphi)x^{-\kappa(\varphi)}$ with some specific constant $C(\varphi) > 0$. So by [58, Lemma A3.26] we have

$$x^\kappa \mathrm{P}[\bar{V}_0^{(1)} > x] \le x^{\kappa-\bar{\kappa}} x^{\bar{\kappa}} \mathrm{P}\left[\sum_{i=1}^\infty p_i V_0^{i,\bar{\varphi}} > x\right] \to 0$$

as $x \to \infty$ for all $\kappa < \bar{\kappa}$. Conversely, if $\kappa > \kappa(\varphi_i)$ for some $i \in \mathbb{N}$, then

$$x^\kappa \mathrm{P}[\bar{V}_0^{(1)} > x] \ge x^\kappa \mathrm{P}[p_i V_0^{i,\varphi_i} > x] = x^{\kappa-\kappa(\varphi_i)} p_i^{\kappa(\varphi_i)} (x/p_i)^{\kappa(\varphi_i)} \mathrm{P}[V_0^{i,\varphi_i} > x/p_i] \to \infty.$$

Recalling that $\kappa(\varphi)$ is defined via the equation $\Psi(\kappa(\varphi),\varphi)=0$, this result is still valid for $\kappa > \bar{\kappa}$ since we have $\inf_{i\in\mathbb{N}} \kappa(\varphi_i) = \bar{\kappa}$ by the implicit function theorem.

Finally, it remains to consider the case $\kappa = \bar{\kappa}$. If $\pi(\{\bar{\varphi}\}) = 0$, then using [89, Lemma 2] and again [58, Lemma A3.26], we obtain

$$\begin{aligned} x^{\bar{\kappa}}\mathrm{P}[\bar{V}_0^{(1)} > x] &\le x^{\bar{\kappa}}\mathrm{P}\left[\sum_{\varphi_i\le\varphi} p_i V_0^{i,\varphi} + \sum_{\varphi_i>\varphi} p_i V_0^{i,\bar{\varphi}} > x\right] \sim x^{\bar{\kappa}}\mathrm{P}\left[\sum_{\varphi_i>\varphi} p_i V_0^{i,\bar{\varphi}} > x\right] \\ &\to \sum_{\varphi_i>\varphi} p_i^{\bar{\kappa}} C(\bar{\varphi}) \end{aligned}$$

as $x\to\infty$ for every $\varphi\in(0,\bar{\varphi})$. Letting $\varphi\to\bar{\varphi}$, the assertion follows. The case $\pi(\{\bar{\varphi}\}) =: \bar{p} > 0$ now follows from the results above and

$$\begin{aligned} x^{\bar{\kappa}}\mathrm{P}[\bar{V}_0^{(1)} > x] &= x^{\bar{\kappa}}\mathrm{P}\left[\sum_{\varphi_i<\bar{\varphi}} p_i V_0^{i,\varphi_i} + \bar{p} V_0^{\bar{i},\bar{\varphi}} > x\right] \\ &\le x^{\bar{\kappa}}\mathrm{P}[\bar{p} V_0^{\bar{i},\bar{\varphi}} > x(1-\epsilon)] + x^{\bar{\kappa}}\mathrm{P}\left[\sum_{\varphi_i<\varphi} p_i V_0^{i,\varphi_i} > \epsilon x\right] \to \left(\frac{\bar{p}}{1-\epsilon}\right)^{\bar{\kappa}} C(\bar{\varphi}) \end{aligned}$$

with $\bar{i}$ being the index corresponding to $\bar{\varphi}$. Letting $\epsilon\to 0$, the proof is completed by setting $C := \bar{p}^{\bar{\kappa}} C(\bar{\varphi})$. □

4.6.2 Proofs for Section 4.3.2

For the proof of Theorem 4.3.10 we need the following lemma.

Lemma 4.6.1. *Let $(S_t)_{t\in\mathbb{R}}$ be a subordinator without drift and define the double-indexed processes $(X_t^\varphi)_{t\in\mathbb{R},\varphi\in\Phi_L}$ and $(V_t^\varphi)_{t\in\mathbb{R},\varphi\in\Phi_L}$ according to* (4.2.16) *and* (4.2.17)*. Then for all $n\in\mathbb{N}$, $-\infty < t_1 < t_2 < \ldots < t_n < \infty$, $h > 0$*

$$((V_{t_1}^\varphi)_{\varphi\in\Phi_L}, (V_{t_2}^\varphi)_{\varphi\in\Phi_L}, \ldots, (V_{t_n}^\varphi)_{\varphi\in\Phi_L}) \overset{\mathrm{d}}{=} ((V_{t_1+h}^\varphi)_{\varphi\in\Phi_L}, (V_{t_2+h}^\varphi)_{\varphi\in\Phi_L}, \ldots, (V_{t_n+h}^\varphi)_{\varphi\in\Phi_L}),$$

i.e. the $\mathbb{R}^{\Phi_L}$-valued stochastic process $((V_t^\varphi)_{\varphi\in\Phi_L})_{t\in\mathbb{R}}$ is strictly stationary. In particular, every finite-dimensional process $(V_t^{\varphi_1},\ldots,V_t^{\varphi_m})_{t\in\mathbb{R}}$, $m\in\mathbb{N}$, is strictly stationary.

Proof. Imitating the proof of [88, Thm. 3.2] for the finite-dimensional process $(V_t^{\varphi_1},\ldots,V_t^{\varphi_m})_{t\in\mathbb{R}}$, $m\in\mathbb{N}$, one readily sees that

$$((V_{t_1}^{\varphi_1},\ldots,V_{t_1}^{\varphi_m}),\ldots,(V_{t_n}^{\varphi_1},\ldots,V_{t_n}^{\varphi_m})) \overset{\mathrm{d}}{=} \left((V_{t_1+h}^{\varphi_1},\ldots,V_{t_1+h}^{\varphi_m}),\ldots,(V_{t_n+h}^{\varphi_1},\ldots,V_{t_n+h}^{\varphi_m})\right)$$

holds. As stochastic processes with the same index space are equal in distribution, whenever their finite-dimensional distributions are equal (e.g. [86, Prop. 2.2]), this already yields the assertion. □

Proof of Theorem 4.3.10. The result follows from the definition of $\bar{V}^{(2)}$ and Lemma 4.6.1. □

To prove Proposition 4.3.11, another auxiliary lemma will be established.

Lemma 4.6.2. *Let $(S_t)_{t\in\mathbb{R}}$ be a subordinator without drift, let $\varphi, \tilde{\varphi} \geq 0$ be fixed and define the processes $(X_t^{\varphi})_{t\in\mathbb{R}}$ and $(X_t^{\tilde{\varphi}})_{t\in\mathbb{R}}$ according to* (4.2.16)*. Set $X_t := X_t^{\varphi} + X_t^{\tilde{\varphi}}$, $t \in \mathbb{R}$.*

(1) The process $(X_t)_{t\in\mathbb{R}}$ is a Lévy process with characteristic triplet $(2\eta, 0, \nu_X)$ where $\nu_X = \nu_S \circ T^{-1}$ for $T\colon \mathbb{R}_+ \to \mathbb{R}_-, y \mapsto -\log(1 + (\varphi + \tilde{\varphi})y + \varphi\tilde{\varphi}y^2)$.

(2) Let $\varphi, \tilde{\varphi} > 0$. Then $\mathbb{E}[\mathrm{e}^{-\kappa X_t}]$ is finite at $\kappa > 0$ for some $t > 0$, or, equivalently, for all $t > 0$, if and only if $\mathbb{E}[S_1^{2\kappa}] < \infty$. In this case we have $\mathbb{E}[\mathrm{e}^{-\kappa X_t}] = \mathrm{e}^{t h_\kappa(\varphi,\tilde{\varphi})}$, where

$$h_\kappa(\varphi, \tilde{\varphi}) = -2\eta\kappa + \int_{\mathbb{R}_+} \Big(((1 + \varphi y)(1 + \tilde{\varphi} y))^\kappa - 1\Big)\, \nu_S(\mathrm{d}y).$$

For $\kappa = 1$ we have

$$h(\varphi, \tilde{\varphi}) := h_1(\varphi, \tilde{\varphi}) = -2\eta + (\varphi + \tilde{\varphi})\mathbb{E}[S_1] + \varphi\tilde{\varphi}\mathrm{Var}[S_1]. \tag{4.6.1}$$

Proof. a) Observe that by definition

$$\begin{aligned} X_t &= 2\eta t - \sum_{0<s\leq t} \log\Big[(1 + \varphi\Delta S_s)(1 + \tilde{\varphi}\Delta S_s)\Big] \\ &= 2\eta t - \sum_{0<s\leq t} \log(1 + (\varphi + \tilde{\varphi})\Delta S_s + \varphi\tilde{\varphi}(\Delta S_s)^2) \end{aligned}$$

for $t \geq 0$, which directly yields the assertion in (1).

b) By [126, Thm. 25.17] $\mathbb{E}[\mathrm{e}^{-\kappa X_t}]$ is finite for some, or, equivalently, for every $t > 0$ if and only if

$$\begin{aligned} \int_{|y|>1} \mathrm{e}^{-\kappa y}\, \nu_X(\mathrm{d}y) = \int_{|y|>1} \mathrm{e}^{-\kappa y}\, \nu_S(T^{-1}(\mathrm{d}y)) &= \int_{y\in D^c} (1 + (\varphi + \tilde{\varphi})y + \varphi\tilde{\varphi}y^2)^\kappa\, \nu_S(\mathrm{d}y) \\ &< \infty \end{aligned}$$

where $D = \left[\frac{-(\varphi+\tilde{\varphi}) - \sqrt{(\varphi-\tilde{\varphi})^2 + 4\varphi\tilde{\varphi}\mathrm{e}}}{2\varphi\tilde{\varphi}}, \frac{-(\varphi+\tilde{\varphi}) + \sqrt{(\varphi-\tilde{\varphi})^2 + 4\varphi\tilde{\varphi}\mathrm{e}}}{2\varphi\tilde{\varphi}}\right]$. This yields (2). □

Proof of Proposition 4.3.11. Due to Lemma 4.6.1 $(V_t^{\varphi}, V_t^{\tilde{\varphi}})_{t\in\mathbb{R}}$ is strictly stationary such that it suffices to consider $t > 0$. Assume without loss of generality that $0 < \varphi \leq \tilde{\varphi}$. Then it follows from the definition of the COGARCH process that $V^{\varphi} \leq V^{\tilde{\varphi}}$. Hence $\mathbb{E}[V_t^{\varphi} V_t^{\tilde{\varphi}}] \leq \mathbb{E}[V_t^{\tilde{\varphi}} V_t^{\tilde{\varphi}}]$ and similarly $\mathbb{E}[V_t^{\varphi} V_{t+h}^{\tilde{\varphi}}] \leq \mathbb{E}[V_t^{\tilde{\varphi}} V_{t+h}^{\tilde{\varphi}}]$, which are both finite as (4.2.10) is given for $\kappa = 2$. We start with the computation of $\mathbb{E}[V_t^{\varphi} V_t^{\tilde{\varphi}}]$ and use (4.2.5) to obtain

$$\begin{aligned}\mathbb{E}[V_t^{\varphi} V_t^{\tilde{\varphi}}] &= \beta^2 \mathbb{E}\left[\int_{(0,t]} \mathrm{e}^{X_s^{\varphi} - X_t^{\varphi}} \,\mathrm{d}s \int_{(0,t]} \mathrm{e}^{X_r^{\tilde{\varphi}} - X_t^{\tilde{\varphi}}} \,\mathrm{d}r\right] + \beta \mathbb{E}[V_0^{\tilde{\varphi}}] \mathbb{E}\left[\int_{(0,t]} \mathrm{e}^{X_s^{\varphi} - X_t^{\varphi} - X_t^{\tilde{\varphi}}} \,\mathrm{d}s\right] \\ &\quad + \beta \mathbb{E}[V_0^{\varphi}] \mathbb{E}\left[\int_{(0,t]} \mathrm{e}^{X_r^{\tilde{\varphi}} - X_t^{\tilde{\varphi}} - X_t^{\varphi}} \,\mathrm{d}r\right] + \mathbb{E}[V_0^{\varphi} V_0^{\tilde{\varphi}}] \mathbb{E}[\mathrm{e}^{-X_t^{\varphi} - X_t^{\tilde{\varphi}}}] \\ &=: \beta^2 I_1 + \beta \mathbb{E}[V_0^{\tilde{\varphi}}] I_2 + \beta \mathbb{E}[V_0^{\varphi}] I_3 + \mathbb{E}[V_0^{\varphi} V_0^{\tilde{\varphi}}] I_4. \end{aligned} \tag{4.6.2}$$

Recall the Lévy process X defined in Lemma 4.6.2 and observe that the increments of X and X^{φ} on disjoint intervals are mutually independent. Thus we have by (4.2.3) and Lemma 4.6.2(2)

$$\begin{aligned} I_1 &= \mathbb{E}\left[\int_{(0,t]} \int_{(0,r]} \mathrm{e}^{X_s^{\varphi} - X_r^{\varphi} + X_r^{\varphi} - X_t^{\varphi} + X_r^{\tilde{\varphi}} - X_t^{\tilde{\varphi}}} \,\mathrm{d}s \,\mathrm{d}r\right] \\ &\quad + \mathbb{E}\left[\int_{(0,t]} \int_{(r,t]} \mathrm{e}^{X_r^{\tilde{\varphi}} - X_s^{\tilde{\varphi}} + X_s^{\tilde{\varphi}} - X_t^{\tilde{\varphi}} + X_s^{\varphi} - X_t^{\varphi}} \,\mathrm{d}s \,\mathrm{d}r\right] \\ &= \int_{(0,t]} \int_{(0,r]} \mathrm{e}^{(r-s)\Psi(1,\varphi) + (t-r)h(\varphi,\tilde{\varphi})} \,\mathrm{d}s \,\mathrm{d}r + \int_{(0,t]} \int_{(r,t]} \mathrm{e}^{(s-r)\Psi(1,\tilde{\varphi}) + (t-s)h(\varphi,\tilde{\varphi})} \,\mathrm{d}s \,\mathrm{d}r \\ &= \frac{-a\mathrm{e}^{ct} + c\mathrm{e}^{at} + a - c}{a^2 c - a c^2} + \frac{-b\mathrm{e}^{ct} + c\mathrm{e}^{bt} + b - c}{b^2 c - b c^2}, \end{aligned}$$

where $a := \Psi(1,\varphi)$, $b := \Psi(1,\tilde{\varphi})$ and $c := h(\varphi,\tilde{\varphi})$. Very similar calculations lead to

$$I_2 = \frac{\mathrm{e}^{bt} - \mathrm{e}^{ct}}{b - c}, \quad I_3 = \frac{\mathrm{e}^{at} - \mathrm{e}^{ct}}{a - c}$$

while we know from Lemma 4.6.2(2) that $I_4 = \mathrm{e}^{ct}$.

According to (4.2.11) we have $\mathbb{E}[V_0^{\varphi}] = -\beta/a$ and $\mathbb{E}[V_0^{\tilde{\varphi}}] = -\beta/b$. Furthermore, we have $\mathbb{E}[V_0^{\varphi} V_0^{\tilde{\varphi}}] = \mathbb{E}[V_t^{\varphi} V_t^{\tilde{\varphi}}]$ due to stationarity. Putting all this into (4.6.2), we obtain

$$(1 - \mathrm{e}^{ct}) \mathbb{E}[V_t^{\varphi} V_t^{\tilde{\varphi}}] = \beta^2 (1 - \mathrm{e}^{ct}) \left(\frac{1}{ac} + \frac{1}{bc}\right).$$

Since $t > 0$ we have $1 - \mathrm{e}^{ct} \neq 0$, so dividing the last equation by this term yields

$$\mathbb{E}[V_t^{\varphi} V_t^{\tilde{\varphi}}] = \frac{\beta^2}{\Psi(1,\tilde{\varphi}) h(\varphi,\tilde{\varphi})} + \frac{\beta^2}{\Psi(1,\varphi) h(\varphi,\tilde{\varphi})}$$

from which (4.3.16) and (4.3.17) follow immediately.

To obtain the formula for $\mathrm{Cov}[V_t^{\varphi}, V_{t+h}^{\tilde{\varphi}}]$ observe first that

$$V_{t+h}^{\tilde{\varphi}} = A_{t,t+h}^{\tilde{\varphi}} V_t^{\tilde{\varphi}} + B_{t,t+h}^{\tilde{\varphi}}, \tag{4.6.3}$$

where

$$A_{t,t+h}^{\tilde{\varphi}} = \mathrm{e}^{-(X_{t+h}^{\tilde{\varphi}} - X_t^{\tilde{\varphi}})} \quad \text{and} \quad B_{t,t+h}^{\tilde{\varphi}} = \beta \int_{(t,t+h]} \mathrm{e}^{-(X_{t+h}^{\tilde{\varphi}} - X_s^{\tilde{\varphi}})} \, \mathrm{d}s.$$

In particular, we see that $A_{t,t+h}^{\tilde{\varphi}}$ and $B_{t,t+h}^{\tilde{\varphi}}$ are independent of $(V_t^{\varphi}, V_t^{\tilde{\varphi}})$ such that

$$\begin{aligned} \mathbb{E}[V_t^{\varphi} V_{t+h}^{\tilde{\varphi}}] &= \mathbb{E}[V_t^{\varphi}(A_{t,t+h}^{\tilde{\varphi}} V_t^{\tilde{\varphi}} + B_{t,t+h}^{\tilde{\varphi}})] \\ &= \mathbb{E}[A_{t,t+h}^{\tilde{\varphi}}] \mathbb{E}[V_t^{\varphi} V_t^{\tilde{\varphi}}] + \mathbb{E}[V_t^{\varphi}] \mathbb{E}[B_{t,t+h}^{\tilde{\varphi}}]. \end{aligned} \tag{4.6.4}$$

Now since

$$\mathbb{E}[A_{t,t+h}^{\tilde{\varphi}}] = \mathbb{E}[\mathrm{e}^{-(X_{t+h}^{\tilde{\varphi}} - X_t^{\tilde{\varphi}})}] = \mathbb{E}[\mathrm{e}^{-X_h^{\tilde{\varphi}}}] = \mathrm{e}^{h\Psi(1,\tilde{\varphi})}$$

and

$$\begin{aligned} \mathbb{E}[B_{t,t+h}^{\tilde{\varphi}}] &= \beta \int_{(t,t+h]} \mathrm{e}^{(t+h-s)\Psi(1,\tilde{\varphi})} \, \mathrm{d}s = \frac{\beta}{\Psi(1,\tilde{\varphi})} \left(\mathrm{e}^{h\Psi(1,\tilde{\varphi})} - 1 \right) \\ &= \mathbb{E}[V_0^{\tilde{\varphi}}] \left(1 - \mathrm{e}^{h\Psi(1,\tilde{\varphi})} \right), \end{aligned}$$

Equation (4.6.4) directly yields

$$\mathbb{E}[V_t^{\varphi} V_{t+h}^{\tilde{\varphi}}] = \mathrm{e}^{h\Psi(1,\tilde{\varphi})} \mathbb{E}[V_0^{\varphi} V_0^{\tilde{\varphi}}] + \left(1 - \mathrm{e}^{h\Psi(1,\tilde{\varphi})} \right) \mathbb{E}[V_0^{\varphi}] \mathbb{E}[V_0^{\tilde{\varphi}}],$$

which gives (4.3.18). □

Proof of Proposition 4.3.12. Due to the fact that all appearing processes are nonnegative we can use Tonelli's theorem to determine the given formulas directly from the definition of $\bar{V}^{(2)}$. □

Proof of Proposition 4.3.13.. The proof is mainly the same as the proof of Proposition 4.3.5, so we only indicate the differences. For $\kappa < \bar{\kappa}$ use the estimation $\mathrm{P}[\bar{V}_0^{(2)} > x] \leq \mathrm{P}[V_0^{\bar{\varphi}} > x]$. For $\kappa > \bar{\kappa}$ and $\pi(\{\bar{\varphi}\}) = 0$, it suffices to consider $\kappa > \kappa(\varphi_i)$ after having chosen sequences $(\varphi_i)_{i\in\mathbb{N}}$ and $(\epsilon_i)_{i\in\mathbb{N}}$ with $\pi((\varphi_i - \epsilon_i, \varphi_i]) > 0$ for each $i \in \mathbb{N}$. Using $\mathrm{P}[\bar{V}_0^{(2)} > x] \geq \mathrm{P}\left[\pi((\varphi_i - \epsilon_i, \varphi_i]) V_0^{\varphi_i} > x\right]$ gives the result. Similarly, use $\mathrm{P}[\bar{V}_0^{(2)} > x] \leq \mathrm{P}\left[\pi((0,\varphi]) V_0^{\varphi} + \pi((\varphi,\bar{\varphi}]) V_0^{\bar{\varphi}} > x\right]$ for $\kappa = \bar{\kappa}$ and $\pi(\{\bar{\varphi}\}) = 0$. For $\kappa = \bar{\kappa}$ and $\pi(\{\bar{\varphi}\}) =: \bar{p} > 0$, we may use $\bar{V}_0^2 = \int_{(0,\bar{\varphi})} V_0^{\varphi} \, \pi(\mathrm{d}\varphi) + \bar{p} V_0^{\bar{\varphi}}$. Finally, the case $\kappa > \bar{\kappa}$ and $\pi(\{\bar{\varphi}\}) > 0$ follows from the preceding arguments. □

4.6.3 Proofs for Section 4.3.3

Proof of Proposition 4.3.15. By (4.2.16) and (4.2.17), the function $\varphi \mapsto V_s^\varphi$ is increasing in φ for every $s \in \mathbb{R}$. As a consequence, we have for all $t \geq 0$

$$A_t \leq \int_{(0,t]} \int_{\Phi_L} \varphi_{\max} V_s^{\varphi_{\max}} \Lambda^S(\mathrm{d}s, \mathrm{d}\varphi) = \varphi_{\max} \int_{(0,t]} V_s^{\varphi_{\max}} \, \mathrm{d}S_s < \infty.$$

Since A is by definition càdlàg, $\mathbb{G}^{(3)}$-adapted and increasing, A is a semimartingale [80, e.g. Def. I.4.21] such that uniqueness of the solution of (4.3.24) follows from [118, Thm. V.7]. It remains to show that (4.3.26) solves (4.3.24). Using integration by parts [80, Def. I.4.45] and [80, Prop. I.4.49d], we obtain

$$\begin{aligned} \mathrm{d}\bar{V}_t^{(3)} &= \left(\bar{V}_0^{(3)} + \int_{(0,t]} \mathrm{e}^{\eta s} \, \mathrm{d}A_s + \beta \int_{(0,t]} \mathrm{e}^{\eta s} \, \mathrm{d}s \right) \mathrm{d}\left(\mathrm{e}^{-\eta t}\right) + \mathrm{e}^{-\eta t}(\mathrm{e}^{\eta t} \, \mathrm{d}A_t + \beta \mathrm{e}^{\eta t} \, \mathrm{d}t) \\ &= -\eta \bar{V}_t^{(3)} \, \mathrm{d}t + \mathrm{d}A_t + \beta \, \mathrm{d}t = (\beta - \eta \bar{V}_t^{(3)}) \, \mathrm{d}t + \mathrm{d}A_t. \end{aligned}$$

□

In order to show that the supCOGARCH 3 process $\bar{V}^{(3)}$ from (4.3.24) has a stationary solution we need a series of lemmata.

Lemma 4.6.3. *Let $n, m \in \mathbb{N}$. For $h > 0$, $-\infty < t_1 < \ldots < t_{m+1} < \infty$ and $0 < \varphi_1 < \ldots < \varphi_{n+1} < \varphi_{\max}$ we have*

$$\begin{aligned} &(V_{t_j}^{\varphi_i}, \Lambda^S((t_j, t_{j+1}] \times (\varphi_i, \varphi_{i+1}]) \colon i \leq n, j \leq m) \\ &\quad \stackrel{\mathrm{d}}{=} (V_{t_j+h}^{\varphi_i}, \Lambda^S((t_j + h, t_{j+1} + h] \times (\varphi_i, \varphi_{i+1}]) \colon i \leq n, j \leq m). \end{aligned} \tag{4.6.5}$$

Proof. For $1 \leq i \leq n$ and $1 \leq j \leq m$ write $\Lambda_j^i := \Lambda^S((t_j, t_{j+1}] \times (\varphi_i, \varphi_{i+1}])$ and $\Lambda_{j,h}^i := \Lambda^S((t_j + h, t_{j+1} + h] \times (\varphi_i, \varphi_{i+1}])$ and let Z^m and Z_h^m denote the left- and right-hand side of (4.6.5), respectively. We first consider $m = 1$. On the one hand, we obtain from Lemma 4.6.1 that $(V_{t_1}^{\varphi_1}, \ldots, V_{t_1}^{\varphi_n}) \stackrel{\mathrm{d}}{=} (V_{t_1+h}^{\varphi_1}, \ldots, V_{t_1+h}^{\varphi_n})$. On the other hand, due to the independence of their single components, the vectors $(\Lambda_1^1, \ldots, \Lambda_1^n)$ and $(\Lambda_{1,h}^1, \ldots, \Lambda_{1,h}^n)$ have the same distribution. Since additionally the V-vector is independent of the Λ^S-vector, the assertion in the case $m = 1$ follows. For $m \geq 2$, using induction and the independence of Λ_m^i and Z^{m-1}, it suffices to show that the conditional distribution of $(V_{t_m}^{\varphi_i} \colon i = 1, \ldots, n)$ given Z^{m-1} does not change if shifted by h. By Markovianity (see [88, Thm. 3.2]) this distribution only depends on $(V_{t_{m-1}}^{\varphi_i}, \Lambda_{m-1}^i \colon i = 1, \ldots, n)$ such that by (4.6.3) and using the notation

there, we only need to consider the distribution of $(A^{\varphi_i}_{t_{m-1},t_m}, B^{\varphi_i}_{t_{m-1},t_m} : i = 1,\dots,n)$ given $(\Lambda^i_{m-1} : i = 1,\dots,n)$. Since the former vector is a measurable transformation of $(\Delta S_s : t_{m-1} \le s \le t_m)$, it is evident that this distribution is invariant under a shift by h, which finishes the proof. □

In connection to (4.3.27), we show a further auxiliary result. To this end define

$$\begin{aligned} A_t &:= \int_{(0,t]} \int_{\Phi_L} \varphi V^\varphi_{s-} \, \Lambda^S(\mathrm{d}s, \mathrm{d}\varphi), \quad t \ge 0, \\ A_t &:= -\int_{(t,0]} \int_{\Phi_L} \varphi V^\varphi_{s-} \, \Lambda^S(\mathrm{d}s, \mathrm{d}\varphi), \quad t < 0. \end{aligned} \tag{4.6.6}$$

Lemma 4.6.4. *The process $(A_t)_{t\in\mathbb{R}}$ defined in (4.6.6) has stationary increments, i.e. for every $n \in \mathbb{N}$, $-\infty < t_1 < \ldots < t_{n+1} < \infty$ and $h > 0$ we have*

$$(A_{t_2} - A_{t_1}, \dots, A_{t_{n+1}} - A_{t_n}) \stackrel{\mathrm{d}}{=} (A_{t_2+h} - A_{t_1+h}, \dots, A_{t_{n+1}+h} - A_{t_n+h}). \tag{4.6.7}$$

Proof. By an approximation via Riemann sums (note that $\varphi \mapsto V^\varphi_s$ is continuous in φ for all s), cf. [80, Prop. I.4.44], we may use Lemma 4.6.3 to obtain

$$\begin{aligned} &(A_{t_2} - A_{t_1}, \dots, A_{t_{n+1}} - A_{t_n}) \\ &\quad = \left(\int_{t_1}^{t_2} \int_{\Phi_L} \varphi V^\varphi_{s-} \, \Lambda^S(\mathrm{d}s, \mathrm{d}\varphi), \dots, \int_{t_n}^{t_{n+1}} \int_{\Phi_L} \varphi V^\varphi_{s-} \, \Lambda^S(\mathrm{d}s, \mathrm{d}\varphi) \right) \\ &\quad \stackrel{\mathrm{d}}{=} \left(\int_{t_1+h}^{t_2+h} \int_{\Phi_L} \varphi V^\varphi_{s-} \, \Lambda^S(\mathrm{d}s, \mathrm{d}\varphi), \dots, \int_{t_n+h}^{t_{n+1}+h} \int_{\Phi_L} \varphi V^\varphi_{s-} \, \Lambda^S(\mathrm{d}s, \mathrm{d}\varphi) \right) \\ &\quad = (A_{t_2+h} - A_{t_1+h}, \dots, A_{t_{n+1}+h} - A_{t_n+h}). \end{aligned}$$

□

Proof of Theorem 4.3.17. Since $\mathrm{e}^{-\eta t} \int_{(0,t]} \mathrm{e}^{\eta s} \, \mathrm{d}s \to \eta^{-1}$ as $t \to \infty$ the process $(V^{(3)}_t)_{t\ge 0}$ converges in distribution to a finite random variable as $t \to \infty$ if and only if

$$\begin{aligned} \mathrm{e}^{-\eta t} \int_{(0,t]} \mathrm{e}^{\eta s} \, \mathrm{d}A_s = \int_{(0,t]} \mathrm{e}^{\eta(s-t)} \, \mathrm{d}A_s = \int_{(-t,0]} \mathrm{e}^{\eta s} \, \mathrm{d}A_{s+t} &\stackrel{\mathrm{d}}{=} \int_{(-t,0]} \mathrm{e}^{\eta s} \, \mathrm{d}A_s \\ &\stackrel{\mathrm{d}}{=} \int_{(0,t]} \mathrm{e}^{-\eta s} \, \mathrm{d}A_s \end{aligned}$$

converges to a finite random variable in distribution as $t \to \infty$, where we used Lemma 4.6.4 for the distributional equalities. By monotonicity this is equivalent to

the existence of

$$\int_{\mathbb{R}_+} \mathrm{e}^{-\eta s}\,\mathrm{d}A_s = \int_{\mathbb{R}_+}\int_{\Phi_L} \mathrm{e}^{-\eta s}\varphi V_{s-}^{\varphi}\,\Lambda^S(\mathrm{d}s,\mathrm{d}\varphi)$$

in probability. As shown in Theorem 1.4.1 and the following remark, this holds if and only if (4.3.28) is valid.

Hence in case that (4.3.28) is violated, no stationary distribution can exist. On the other hand, given (4.3.28), following the above computations, the process $(V_t^{(3)})_{t\geq 0}$ converges in distribution to $\bar{V}_\infty^{(3)} := \frac{\beta}{\eta} + \int_0^\infty \mathrm{e}^{-\eta s}\,\mathrm{d}A_s$, which is thus the unique possible stationary distribution.

To show that $(V_t^{(3)})_{t\geq 0}$ is actually strictly stationary when started in a random variable $V_0^{(3)} \stackrel{\mathrm{d}}{=} \bar{V}_\infty^{(3)}$ which is independent of Λ^L on $\mathbb{R}_+ \times \Phi_L$, we set

$$\bar{V}_0^{(3)} := \frac{\beta}{\eta} + \int_{(-\infty,0]} \mathrm{e}^{\eta s}\,\mathrm{d}A_s.$$

Then using Lemma 4.6.4 we obtain for all $0 \leq t_1 < \ldots < t_n$ and $h > 0$

$$\begin{aligned}
&(\bar{V}_{t_1}^{(3)},\ldots,\bar{V}_{t_n}^{(3)})\\
&\stackrel{\mathrm{d}}{=} \left(\int_{(-\infty,t_1]} \mathrm{e}^{-\eta(t_1-s)}\,\mathrm{d}A_s + \beta\int_{(-\infty,t_1]} \mathrm{e}^{-\eta(t_1-s)}\,\mathrm{d}s,\ldots,\right.\\
&\qquad\left.\int_{(-\infty,t_n]} \mathrm{e}^{-\eta(t_n-s)}\,\mathrm{d}A_s + \beta\int_{(-\infty,t_n]} \mathrm{e}^{-\eta(t_n-s)}\,\mathrm{d}s\right)\\
&= \left(\int_{(-\infty,0]} \mathrm{e}^{\eta s}\,\mathrm{d}A_{s+t_1} + \beta\int_{\mathbb{R}_+} \mathrm{e}^{-\eta s}\,\mathrm{d}s,\ldots,\int_{(-\infty,0]} \mathrm{e}^{\eta s}\,\mathrm{d}A_{s+t_n} + \beta\int_{\mathbb{R}_+} \mathrm{e}^{-\eta s}\,\mathrm{d}s\right)\\
&\stackrel{\mathrm{d}}{=} \left(\int_{(-\infty,0]} \mathrm{e}^{\eta s}\,\mathrm{d}A_{s+t_1+h} + \beta\int_{\mathbb{R}_+} \mathrm{e}^{-\eta s}\,\mathrm{d}s,\ldots,\int_{(-\infty,0]} \mathrm{e}^{\eta s}\,\mathrm{d}A_{s+t_n+h} + \beta\int_{\mathbb{R}_+} \mathrm{e}^{-\eta s}\,\mathrm{d}s\right)\\
&\stackrel{\mathrm{d}}{=} (\bar{V}_{t_1+h}^{(3)},\ldots,\bar{V}_{t_n+h}^{(3)})
\end{aligned}$$

and hence the process $(V_t^{(3)})_{t\geq 0}$ is strictly stationary.

It remains to show that (1) and (2) imply (4.3.28). First observe from (4.2.16) and (4.2.17) that for fix s the function $\varphi \mapsto V_s^\varphi$ is increasing in φ. So if (1) holds, we have

$$\begin{aligned}
\int_{\mathbb{R}_+}\int_{\Phi_L}\int_{\mathbb{R}_+} 1 \wedge (y\varphi V_s^\varphi \mathrm{e}^{-\eta s})\,\mathrm{d}s\,\pi(\mathrm{d}\varphi)\,\nu_S(\mathrm{d}y) &\leq \int_{\mathbb{R}_+}\int_{\mathbb{R}_+} 1 \wedge (y\varphi_0 V_s^{\varphi_0}\mathrm{e}^{-\eta s})\,\mathrm{d}s\,\nu_S(\mathrm{d}y)\\
&< \infty
\end{aligned}$$

because (4.3.28) holds for $\pi = \delta_{\varphi_0}$ (in this case $\bar{V}^{(3)}$ is just the COGARCH process V^{φ_0}). Finally, (2) follows from (1) together with the fact that $\varphi_{\max}^{(\kappa)} < \varphi_{\max}$. □

For the proof of Proposition 4.3.18 we need the following Lemma.

Lemma 4.6.5. *Let $(A_t)_{t\in\mathbb{R}}$, V^φ and $\bar{V}^{(3)}$ be defined as in* (4.6.6), (4.2.17) *and* (4.3.29)*, respectively. Then, under the assumptions of Proposition 4.3.18, we have for $t \geq 0$*

$$[A,A]_t = \int_{(0,t]}\int_{\Phi_L}\int_{\mathbb{R}_+} \varphi^2 (V^\varphi_{s-})^2 y^2 \,\mu^{\Lambda^S}(\mathrm{d}s,\mathrm{d}\varphi,\mathrm{d}y) \quad \textit{and}$$
$$[\bar{V}^{(3)},V^\varphi]_t = [A,V^\varphi]_t = \varphi\int_{(0,t]}\int_{\Phi_L}\int_{\mathbb{R}_+} \tilde{\varphi} V^\varphi_{s-} V^{\tilde{\varphi}}_{s-} y^2 \,\mu^{\Lambda^S}(\mathrm{d}s,\mathrm{d}\tilde{\varphi},\mathrm{d}y),$$

with μ^{Λ^S} as defined in (4.2.18)*. For $t<0$, let the expressions on the left-hand side denote the respective quadratic (co-)variation on $(t,0]$. Then the integrals have to be computed on $(-t,0]$ instead of $(0,t]$.*

Proof. Obviously it suffices to consider $t \geq 0$. Since A is an increasing pure-jump process,

$$\begin{aligned}[A,A]_t &= \sum_{0<s\leq t} (\Delta A_s)^2 = \sum_{0<s\leq t} \left(\Delta(\varphi V^\varphi_{\cdot-} y * \mu^{\Lambda^S})_s\right)^2 \\ &= \sum_{0<s\leq t} \left(\sum_{\varphi\in\Phi_L} \varphi V^\varphi_{s-} \Lambda^S(\{s\}\times\{\varphi\})\right)^2\end{aligned}$$

Noting that for almost every ω there is at most one $\varphi\in\Phi_L$ at time s such that $\Lambda^S(\{s\}\times\{\varphi\})(\omega)\neq 0$, we obtain

$$[A,A]_t = \sum_{0<s\leq t}\sum_{\varphi\in\Phi_L} \varphi^2 (V^\varphi_{s-})^2 \Lambda^S(\{s\}\times\{\varphi\})^2,$$

as desired. Similarly,

$$[A,V^\varphi]_t = \sum_{0<s\leq t} \Delta A_s \Delta V^\varphi_s = \sum_{0<s\leq t}\left(\sum_{\tilde{\varphi}\in\Phi_L} \tilde{\varphi} V^{\tilde{\varphi}}_{s-} \Lambda^S(\{s\}\times\{\tilde{\varphi}\})\right) \varphi V^\varphi_{s-} \Delta S_s$$

according to (4.2.6). Now observe that $\Delta S_s = \Lambda^S(\{s\}\times\mathbb{R}_+) = \sum_{\varphi\in\Phi_L}\Lambda^S(\{s\}\times\{\varphi\})$ for all $s\in\mathbb{R}$, where again for almost every ω there is at most one $\varphi\in\Phi_L$ at time s with $\Lambda^S(\{s\}\times\{\varphi\})(\omega)\neq 0$. As a result,

$$[A,V^\varphi]_t = \sum_{0<s\leq t} \varphi V^\varphi_{s-} \sum_{\tilde{\varphi}\in\Phi_L} \tilde{\varphi} V^{\tilde{\varphi}}_{s-} \Lambda^S(\{s\}\times\{\tilde{\varphi}\})^2 = \varphi(\tilde{\varphi} V^\varphi_{\cdot-} V^{\tilde{\varphi}}_{\cdot-} y^2 * \mu^{\Lambda^S}_t).$$

Finally, we have $[\bar{V}^{(3)},V^\varphi] = [A,V^\varphi]$ by (4.3.24). □

Proof of Proposition 4.3.18. First observe that Theorem 4.3.17(2) ensures the existence of the given stationary version of $\bar{V}^{(3)}$ under the assumptions of the present theorem.

We set $m_1 := \int_{\mathbb{R}_+} y\,\nu_S(\mathrm{d}y) = \mathbb{E}[S_1]$, $m_2 := \int_{\mathbb{R}_+} y^2\,\nu_S(\mathrm{d}y) = \mathrm{Var}[S_1]$ and assume without loss of generality $\pi(\{0\}) = 0$. For the mean we use (4.2.11) and obtain

$$\begin{aligned}
\mathbb{E}[\bar{V}_t^{(3)}] = \mathbb{E}[\bar{V}_0^{(3)}] &= \frac{\beta}{\eta} + \mathbb{E}\left[\int_{(-\infty,0]} \mathrm{e}^{\eta s}\,\mathrm{d}A_s\right] \\
&= \frac{\beta}{\eta} + m_1 \int_{(-\infty,0]} \mathrm{e}^{\eta s}\,\mathrm{d}s \int_{\Phi_L} \varphi \mathbb{E}[V_0^{\varphi}]\,\pi(\mathrm{d}\varphi) \\
&= \frac{\beta}{\eta} - \frac{\beta}{\eta} \int_{\Phi_L} \left(1 + \frac{\eta}{m_1\varphi - \eta}\right) \pi(\mathrm{d}\varphi) \\
&= -\int_{\Phi_L} \frac{\beta}{m_1\varphi - \eta}\,\pi(\mathrm{d}\varphi) = \int_{\Phi_L} \mathbb{E}[V_0^{\varphi}]\,\pi(\mathrm{d}\varphi).
\end{aligned}$$

To compute the autocovariance function of $\bar{V}^{(3)}$ observe that for $t \geq 0$, $h \geq 0$ we have from (4.3.29)

$$\begin{aligned}
\mathrm{Cov}[\bar{V}_t^{(3)}, \bar{V}_{t+h}^{(3)}] &= \mathrm{e}^{-2\eta t}\mathrm{e}^{-\eta h}\mathbb{E}\left[\int_{(-\infty,t]} \mathrm{e}^{\eta s}\,\mathrm{d}A_s \int_{(-\infty,t+h]} \mathrm{e}^{\eta s}\,\mathrm{d}A_s\right] \\
&\quad - \mathbb{E}\left[\int_{(-\infty,t]} \mathrm{e}^{-\eta(t-s)}\,\mathrm{d}A_s\right] \mathbb{E}\left[\int_{(-\infty,t+h]} \mathrm{e}^{-\eta(t+h-s)}\,\mathrm{d}A_s\right] \\
&= \mathrm{e}^{-2\eta t}\mathrm{e}^{-\eta h}\left(\mathbb{E}\left[\left(\int_{(-\infty,t]} \mathrm{e}^{\eta s}\,\mathrm{d}A_s\right)^2\right]\right. \\
&\quad \left. + \mathbb{E}\left[\int_{(-\infty,t]} \mathrm{e}^{\eta s}\,\mathrm{d}A_s \int_t^{t+h} \mathrm{e}^{\eta s}\,\mathrm{d}A_s\right]\right) - \frac{m_1^2}{\eta^2}\left(\int_{\Phi_L} \varphi\mathbb{E}[V_0^{\varphi}]\,\pi(\mathrm{d}\varphi)\right)^2 \\
&=: \mathrm{e}^{-2\eta t}\mathrm{e}^{-\eta h}(E_1 + E_2) - \frac{m_1^2}{\eta^2}\left(\int_{\Phi_L} \varphi\mathbb{E}[V_0^{\varphi}]\,\pi(\mathrm{d}\varphi)\right)^2. \qquad (4.6.8)
\end{aligned}$$

For E_1 we can use integration by parts (see [80, Eq. I.4.45]) together with [118, Thms. II.19 and VI.29] and Lemma 4.6.5 to obtain

$$\begin{aligned}
E_1 &= 2\mathbb{E}\left[\int_{(-\infty,t]} \left(\int_{(-\infty,s)} \mathrm{e}^{\eta r}\,\mathrm{d}A_r\right) \mathrm{e}^{\eta s}\,\mathrm{d}A_s\right] + \mathbb{E}\left[\int_{(-\infty,t]} \mathrm{e}^{2\eta s}\,\mathrm{d}[A,A]_s\right] \\
&= 2m_1 \int_{\Phi_L} \int_{(-\infty,t]} \mathbb{E}\left[\left(\int_{(-\infty,s]} \mathrm{e}^{\eta r}\,\mathrm{d}A_r\right) V_s^{\varphi}\right] \mathrm{e}^{\eta s}\varphi\,\mathrm{d}s\,\pi(\mathrm{d}\varphi)
\end{aligned}$$

$$
\begin{aligned}
&+ m_2 \int_{\Phi_L} \int_{(-\infty,t]} \mathrm{e}^{2\eta s} \varphi^2 \mathbb{E}[(V_s^\varphi)^2] \,\mathrm{d}s \,\pi(\mathrm{d}\varphi) \\
&= 2m_1 \int_{\Phi_L} \int_{(-\infty,t]} g(s,\varphi) \mathrm{e}^{\eta s} \varphi \,\mathrm{d}s \,\pi(\mathrm{d}\varphi) + \frac{m_2}{2\eta} \mathrm{e}^{2\eta t} \int_{\Phi_L} \varphi^2 \mathbb{E}[(V_0^\varphi)^2] \,\pi(\mathrm{d}\varphi), \quad (4.6.9)
\end{aligned}
$$

where $g(s,\varphi) := \mathbb{E}\left[V_s^\varphi \int_{(-\infty,s]} \mathrm{e}^{\eta r} \,\mathrm{d}A_r\right]$. Then again using integration by parts, Lemmata 4.6.3 and 4.6.5 and Equations (4.2.4), (4.2.6), (4.2.20) and (4.2.11), we obtain

$$
\begin{aligned}
g(s,\varphi) &= \mathbb{E}\left[\int_{(-\infty,s]} \int_{(-\infty,r)} \mathrm{e}^{\eta u} \,\mathrm{d}A_u \,\mathrm{d}V_r^\varphi\right] + \mathbb{E}\left[\int_{(-\infty,s]} V_{r-}^\varphi \mathrm{e}^{\eta r} \,\mathrm{d}A_r\right] \\
&\quad + \mathbb{E}\left[\int_{(-\infty,s]} \mathrm{e}^{\eta r} \,\mathrm{d}[A, V^\varphi]_r\right] \\
&= \mathbb{E}\left[\int_{(-\infty,s]} \left(\int_{(-\infty,r]} \mathrm{e}^{\eta u} \,\mathrm{d}A_u\right) (\beta - \eta V_r^\varphi) \,\mathrm{d}r\right] + \mathbb{E}\left[\int_{(-\infty,s]} V_{r-}^\varphi \mathrm{e}^{\eta r} \,\mathrm{d}A_r\right] \\
&\quad + \mathbb{E}\left[\int_{(-\infty,s]} \left(\int_{(-\infty,r)} \mathrm{e}^{\eta u} \,\mathrm{d}A_u\right) \varphi V_{r-}^\varphi \,\mathrm{d}S_r\right] + \mathbb{E}\left[\int_{(-\infty,s]} \mathrm{e}^{\eta r} \,\mathrm{d}[A, V^\varphi]_r\right] \\
&= \beta m_1 \int_{(-\infty,s]} \int_{(-\infty,r]} \mathrm{e}^{\eta u} \,\mathrm{d}u \,\mathrm{d}r \int_{\Phi_L} \tilde{\varphi} \mathbb{E}[V_0^{\tilde{\varphi}}] \,\pi(\mathrm{d}\tilde{\varphi}) \\
&\quad + (m_1\varphi - \eta) \int_{(-\infty,s]} g(r,\varphi) \,\mathrm{d}r \\
&\quad + m_1 \int_{(-\infty,s]} \mathrm{e}^{\eta r} \,\mathrm{d}r \int_{\Phi_L} \tilde{\varphi} \mathbb{E}[V_0^\varphi V_0^{\tilde{\varphi}}] \,\pi(\mathrm{d}\tilde{\varphi}) \\
&\quad + m_2 \varphi \int_{(-\infty,s]} \mathrm{e}^{\eta r} \,\mathrm{d}r \int_{\Phi_L} \tilde{\varphi} \mathbb{E}[V_0^\varphi V_0^{\tilde{\varphi}}] \,\pi(\mathrm{d}\tilde{\varphi}) \\
&= \frac{\mathrm{e}^{\eta s}}{\eta} \left(\frac{m_1 \beta}{\eta} \int_{\Phi_L} \tilde{\varphi} \mathbb{E}[V_0^{\tilde{\varphi}}] \,\pi(\mathrm{d}\tilde{\varphi}) + (m_1 + m_2\varphi) \int_{\Phi_L} \tilde{\varphi} \mathbb{E}[V_0^\varphi V_0^{\tilde{\varphi}}] \,\pi(\mathrm{d}\tilde{\varphi})\right) \\
&\quad + \Psi(1,\varphi) \int_{(-\infty,s]} g(r,\varphi) \,\mathrm{d}r \\
&= \frac{\mathrm{e}^{\eta s}}{\eta} C(\varphi) + \Psi(1,\varphi) \int_{(-\infty,s]} g(r,\varphi) \,\mathrm{d}r,
\end{aligned}
$$

with

$$
\begin{aligned}
C(\varphi) &:= \int_{\Phi_L} C(\varphi,\tilde{\varphi}) \,\pi(\mathrm{d}\tilde{\varphi}), \\
C(\varphi,\tilde{\varphi}) &:= -\frac{m_1}{\eta} \Psi(1,\varphi) \tilde{\varphi} \mathbb{E}[V_0^\varphi] \mathbb{E}[V_0^{\tilde{\varphi}}] + (m_1 + m_2\varphi) \tilde{\varphi} \mathbb{E}[V_0^\varphi V_0^{\tilde{\varphi}}].
\end{aligned}
$$

Solving this integral equation yields $g(s,\varphi) = \frac{C(\varphi)\mathrm{e}^{\eta s}}{\eta - \Psi(1,\varphi)}$. Inserting this result in (4.6.9)

gives

$$E_1 = \frac{m_1}{\eta}\mathrm{e}^{2\eta t}\int_{\Phi_L}\frac{\varphi C(\varphi)}{\eta-\Psi(1,\varphi)}\,\pi(\mathrm{d}\varphi) + \frac{m_2}{2\eta}\mathrm{e}^{2\eta t}\int_{\Phi_L}\varphi^2\mathbb{E}[(V_0^\varphi)^2]\,\pi(\mathrm{d}\varphi).$$

Let us turn to E_2 and use the notation $\mathbb{G}^{(3)} = (\mathcal{G}_t^{(3)})_{t\in\mathbb{R}}$ for the augmented natural filtration of Λ^L. Now taking conditional expectation with respect to $\mathcal{G}_t^{(3)}$ and observing that V^φ, $\bar{V}^{(3)}$ as well as A are all adapted to $\mathbb{G}^{(3)}$, we obtain

$$E_2 = \mathbb{E}\left[\left(\int_{(-\infty,t]}\mathrm{e}^{\eta s}\,\mathrm{d}A_s\right)\mathbb{E}\left[\int_t^{t+h}\int_{\Phi_L}\mathrm{e}^{\eta s}\varphi V_{s-}^\varphi\,\Lambda^S(\mathrm{d}s,\mathrm{d}\varphi)\,\middle|\,\mathcal{G}_t^{(3)}\right]\right].$$

Observing that the restriction of Λ^S on $(t,t+h]$ is independent of $\mathcal{F}_t$, we have

$$E_2 = \mathbb{E}\left[\left(\int_{(-\infty,t]}\mathrm{e}^{\eta s}\,\mathrm{d}A_s\right)m_1\int_{\Phi_L}\int_{(t,t+h]}\mathrm{e}^{\eta s}\varphi\mathbb{E}[V_{s-}^\varphi|\mathcal{G}_t^{(3)}]\,\mathrm{d}s\,\pi(\mathrm{d}\varphi)\right].$$

According to [88, Eq. (4.5)] we have $\mathbb{E}[V_{s-}^\varphi|\mathcal{G}_t^{(3)}] = (V_t^\varphi - \mathbb{E}[V_0^\varphi])\mathrm{e}^{(s-t)\Psi(1,\varphi)} + \mathbb{E}[V_0^\varphi]$ for $s > t$. So we get

$$\begin{aligned}
E_2 &= m_1\mathbb{E}\Bigg[\left(\int_{(-\infty,t]}\mathrm{e}^{\eta s}\,\mathrm{d}A_s\right)\\
&\quad\times\int_{\Phi_L}\int_{(t,t+h]}\mathrm{e}^{\eta s}\varphi\Big((V_t^\varphi-\mathbb{E}[V_0^\varphi])\mathrm{e}^{(s-t)\Psi(1,\varphi)}+\mathbb{E}[V_0^\varphi]\Big)\,\mathrm{d}s\,\pi(\mathrm{d}\varphi)\Bigg]\\
&= m_1\int_{\Phi_L}\varphi\mathbb{E}\left[V_t^\varphi\int_{(-\infty,t]}\mathrm{e}^{\eta s}\,\mathrm{d}A_s\right]\int_{(t,t+h]}\mathrm{e}^{\eta s}\mathrm{e}^{(s-t)\Psi(1,\varphi)}\,\mathrm{d}s\,\pi(\mathrm{d}\varphi)\\
&\quad + m_1\mathbb{E}\left[\int_{(-\infty,t]}\mathrm{e}^{\eta s}\,\mathrm{d}A_s\right]\int_{\Phi_L}\varphi\mathbb{E}[V_0^\varphi]\int_{(t,t+h]}\mathrm{e}^{\eta s}(1-\mathrm{e}^{(s-t)\Psi(1,\varphi)})\,\mathrm{d}s\,\pi(\mathrm{d}\varphi)\\
&= \int_{\Phi_L}g(t,\varphi)\mathrm{e}^{\eta t}(\mathrm{e}^{m_1\varphi h}-1)\,\pi(\mathrm{d}\varphi) + m_1^2\int_{(-\infty,t]}\mathrm{e}^{\eta s}\,\mathrm{d}s\int_{\Phi_L}\varphi\mathbb{E}[V_0^\varphi]\,\pi(\mathrm{d}\varphi)\\
&\quad\times\int_{\Phi_L}\varphi\mathbb{E}[V_0^\varphi]\mathrm{e}^{\eta t}\left(\frac{\mathrm{e}^{\eta h}-1}{\eta}-\frac{\mathrm{e}^{m_1\varphi h}-1}{m_1\varphi}\right)\pi(\mathrm{d}\varphi)\\
&= \mathrm{e}^{2\eta t}\left(\int_{\Phi_L}\frac{C(\varphi)}{\eta-\Psi(1,\varphi)}(\mathrm{e}^{m_1\varphi h}-1)\,\pi(\mathrm{d}\varphi)+\frac{m_1^2}{\eta^2}(\mathrm{e}^{\eta h}-1)\left(\int_{\Phi_L}\varphi\mathbb{E}[V_0^\varphi]\,\pi(\mathrm{d}\varphi)\right)^2\right.\\
&\quad\left.-\frac{m_1}{\eta}\int_{\Phi_L}\varphi\mathbb{E}[V_0^\varphi]\,\pi(\mathrm{d}\varphi)\int_{\Phi_L}\mathbb{E}[V_0^\varphi](\mathrm{e}^{m_1\varphi h}-1)\,\pi(\mathrm{d}\varphi)\right).
\end{aligned}$$

Now inserting the results for E_1 and E_2 in (4.6.8), we obtain

$$
\begin{aligned}
&\operatorname{Cov}[\bar{V}_t^{(3)}, \bar{V}_{t+h}^{(3)}] \\
&\quad = \mathrm{e}^{-\eta h}\left(\frac{m_1}{\eta}\int_{\Phi_L}\frac{\varphi C(\varphi)}{\eta-\Psi(1,\varphi)}\,\pi(\mathrm{d}\varphi) + \frac{m_2}{2\eta}\int_{\Phi_L}\varphi^2\mathbb{E}[(V_0^{\varphi})^2]\,\pi(\mathrm{d}\varphi)\right) \\
&\qquad + \int_{\Phi_L}\frac{C(\varphi)}{\eta-\Psi(1,\varphi)}(\mathrm{e}^{\Psi(1,\varphi)h}-\mathrm{e}^{-\eta h})\,\pi(\mathrm{d}\varphi) - \frac{m_1^2}{\eta^2}\mathrm{e}^{-\eta h}\left(\int_{\Phi_L}\varphi\mathbb{E}[V_0^{\varphi}]\,\pi(\mathrm{d}\varphi)\right)^2 \\
&\qquad - \frac{m_1}{\eta}\int_{\Phi_L}\tilde{\varphi}\mathbb{E}[V_0^{\tilde{\varphi}}]\,\pi(\mathrm{d}\tilde{\varphi})\int_{\Phi_L}\mathbb{E}[V_0^{\varphi}](\mathrm{e}^{\Psi(1,\varphi)h}-\mathrm{e}^{-\eta h})\,\pi(\mathrm{d}\varphi) \\
&\quad = \int_{\Phi_L}\int_{\Phi_L}\left(\frac{C(\varphi,\tilde{\varphi})}{\eta-\Psi(1,\varphi)} - \frac{m_1}{\eta}\tilde{\varphi}\mathbb{E}[V_0^{\tilde{\varphi}}]\mathbb{E}[V_0^{\varphi}]\right)\mathrm{e}^{\Psi(1,\varphi)h}\,\pi(\mathrm{d}\varphi)\,\pi(\mathrm{d}\tilde{\varphi}) \\
&\qquad + \mathrm{e}^{-\eta h}\int_{\Phi_L}\int_{\Phi_L}\left(\frac{m_1\varphi C(\varphi,\tilde{\varphi})}{\eta(\eta-\Psi(1,\varphi))} + \frac{m_2\varphi^2\mathbb{E}[(V_0^{\varphi})^2]}{2\eta} - \frac{C(\varphi,\tilde{\varphi})}{\eta-\Psi(1,\varphi)}\right. \\
&\qquad \left. - \frac{m_1^2}{\eta^2}\varphi\tilde{\varphi}\mathbb{E}[V_0^{\varphi}]\mathbb{E}[V_0^{\tilde{\varphi}}] + \frac{m_1}{\eta}\tilde{\varphi}\mathbb{E}[V_0^{\varphi}]\mathbb{E}[V_0^{\tilde{\varphi}}]\right)\pi(\mathrm{d}\varphi)\,\pi(\mathrm{d}\tilde{\varphi}), \qquad (4.6.10)
\end{aligned}
$$

where using Proposition 4.3.11 together with Equations (4.2.4) and (4.6.1) gives

$$
\begin{aligned}
&\frac{C(\varphi,\tilde{\varphi})}{\eta-\Psi(1,\varphi)} - \frac{m_1}{\eta}\tilde{\varphi}\mathbb{E}[V_0^{\varphi}]\mathbb{E}[V_0^{\tilde{\varphi}}] \\
&= -\frac{m_1\Psi(1,\varphi)\tilde{\varphi}\mathbb{E}[V_0^{\varphi}]\mathbb{E}[V_0^{\tilde{\varphi}}]}{\eta(\eta-\Psi(1,\varphi))} + \frac{(m_1+m_2\varphi)\tilde{\varphi}\mathbb{E}[V_0^{\varphi}V_0^{\tilde{\varphi}}]}{\eta-\Psi(1,\varphi)} - \frac{m_1}{\eta}\tilde{\varphi}\mathbb{E}[V_0^{\varphi}]\mathbb{E}[V_0^{\tilde{\varphi}}] \\
&= \frac{(m_1+m_2\varphi)\tilde{\varphi}}{\eta-\Psi(1,\varphi)}\mathbb{E}[V_0^{\varphi}V_0^{\tilde{\varphi}}] - \frac{m_1\tilde{\varphi}}{\eta(\eta-\Psi(1,\varphi))}\mathbb{E}[V_0^{\varphi}]\mathbb{E}[V_0^{\tilde{\varphi}}](\Psi(1,\varphi)+\eta-\Psi(1,\varphi)) \\
&= \frac{\eta+\Psi(1,\tilde{\varphi})+h(\varphi,\tilde{\varphi})-\Psi(1,\varphi)-\Psi(1,\tilde{\varphi})}{\eta-\Psi(1,\varphi)}\mathbb{E}[V_0^{\varphi}V_0^{\tilde{\varphi}}] - \frac{\eta+\Psi(1,\tilde{\varphi})}{\eta-\Psi(1,\varphi)}\mathbb{E}[V_0^{\varphi}]\mathbb{E}[V_0^{\tilde{\varphi}}] \\
&= \left(1+\frac{h(\varphi,\tilde{\varphi})}{\eta-\Psi(1,\varphi)}\right)\mathbb{E}[V_0^{\varphi}V_0^{\tilde{\varphi}}] - \left(1+\frac{\Psi(1,\varphi)+\Psi(1,\tilde{\varphi})}{\eta-\Psi(1,\varphi)}\right)\mathbb{E}[V_0^{\varphi}]\mathbb{E}[V_0^{\tilde{\varphi}}] \\
&= \operatorname{Cov}[V_0^{\varphi}, V_0^{\tilde{\varphi}}], \qquad (4.6.11)
\end{aligned}
$$

while for the second part of (4.6.10) we have by Equations (4.2.4), (4.2.11) and (4.2.13)

$$\begin{aligned}
&\frac{m_1\varphi C(\varphi,\tilde\varphi)}{\eta(\eta-\Psi(1,\varphi))}+\frac{m_2\varphi^2\mathbb{E}[(V_0^\varphi)^2]}{2\eta}-\frac{C(\varphi,\tilde\varphi)}{\eta-\Psi(1,\varphi)}-\frac{m_1^2}{\eta^2}\varphi\tilde\varphi\mathbb{E}[V_0^\varphi]\mathbb{E}[V_0^{\tilde\varphi}]\\
&\qquad+\frac{m_1}{\eta}\tilde\varphi\mathbb{E}[V_0^\varphi]\mathbb{E}[V_0^{\tilde\varphi}]\\
&=\frac{\Psi(1,\varphi)C(\varphi,\tilde\varphi)}{\eta(\eta-\Psi(1,\varphi))}+\frac{\big(\Psi(2,\varphi)-2\Psi(1,\varphi)\big)}{2\eta}\mathbb{E}[(V_0^\varphi)^2]-\frac{m_1\tilde\varphi\Psi(1,\varphi)}{\eta^2}\mathbb{E}[V_0^\varphi]\mathbb{E}[V_0^{\tilde\varphi}]\\
&=:\ F_1+F_2+F_3. \qquad (4.6.12)
\end{aligned}$$

Now observe that by (4.2.12) and (4.2.11)

$$F_2=\frac{\Psi(2,\varphi)}{2\eta}\mathbb{E}[(V_0^\varphi)^2]-\frac{\Psi(1,\varphi)}{\eta}\mathbb{E}[(V_0^\varphi)^2]=-\frac{\beta}{\eta}\mathbb{E}[V_0^\varphi]+\frac{\beta}{\eta}\frac{\mathbb{E}[(V_0^\varphi)^2]}{\mathbb{E}[V_0^\varphi]},$$

while

$$\begin{aligned}
F_3&=-\frac{(\Psi(1,\tilde\varphi)+\eta)\Psi(1,\varphi)}{\eta^2}\mathbb{E}[V_0^\varphi]\mathbb{E}[V_0^{\tilde\varphi}]\\
&=-\frac{\beta^2}{\eta^2}+\frac{\Psi(1,\varphi)}{\eta}\mathrm{Cov}[V_0^\varphi,V_0^{\tilde\varphi}]-\frac{\Psi(1,\varphi)}{\eta}\mathbb{E}[V_0^\varphi V_0^{\tilde\varphi}]\\
&=-\frac{\beta^2}{\eta^2}-\frac{\beta}{\eta}\frac{\mathrm{Cov}[V_0^\varphi,V_0^{\tilde\varphi}]}{\mathbb{E}[V_0^\varphi]}+\frac{\beta}{\eta}\frac{\mathbb{E}[V_0^\varphi V_0^{\tilde\varphi}]}{\mathbb{E}[V_0^\varphi]}.
\end{aligned}$$

On the other hand we obtain by similar means

$$\begin{aligned}
\Psi(1,\varphi)C(\varphi,\tilde\varphi)&=\frac{(m_1+m_2\varphi)\tilde\varphi\beta^2(\Psi(1,\varphi)+\Psi(1,\tilde\varphi))}{h(\varphi,\tilde\varphi)\Psi(1,\tilde\varphi)}-\frac{m_1\beta^2\tilde\varphi\Psi(1,\varphi)}{\eta\Psi(1,\tilde\varphi)}\\
&=\frac{\beta^2}{\eta\Psi(1,\tilde\varphi)h(\varphi,\tilde\varphi)}\Big(\eta(m_1+m_2\varphi)\tilde\varphi(\Psi(1,\varphi)+\Psi(1,\tilde\varphi))\\
&\quad-m_1\tilde\varphi\Psi(1,\varphi)h(\varphi,\tilde\varphi)\Big)\\
&=\frac{\beta^2(\eta-\Psi(1,\varphi))}{\eta\Psi(1,\tilde\varphi)h(\varphi,\tilde\varphi)}\Big(h(\varphi,\tilde\varphi)\Psi(1,\tilde\varphi)+\eta(\Psi(1,\varphi)+\Psi(1,\tilde\varphi))\Big)
\end{aligned}$$

such that by Proposition 4.3.11

$$F_1=\frac{\beta^2}{\eta^2}-\frac{\beta}{\eta}\frac{\mathbb{E}[V_0^\varphi V_0^{\tilde\varphi}]}{\mathbb{E}[V_0^\varphi]}.$$

Finally inserting (4.6.11) and (4.6.12) with the obtained formulas for F_1, F_2 and F_3

in (4.6.10) gives

$$\operatorname{Cov}[\bar{V}_t^{(3)}, \bar{V}_{t+h}^{(3)}] = \int_{\Phi_L}\int_{\Phi_L}\Bigg(\operatorname{Cov}[V_0^{\varphi}, V_0^{\tilde{\varphi}}]\mathrm{e}^{\Psi(1,\varphi)h} + \mathrm{e}^{-\eta h}\Bigg(-\frac{\beta}{\eta}\mathbb{E}[V_0^{\varphi}] + \frac{\beta}{\eta}\frac{\mathbb{E}[(V_0^{\varphi})^2]}{\mathbb{E}[V_0^{\varphi}]} - \frac{\beta}{\eta}\frac{\operatorname{Cov}[V_0^{\varphi}, V_0^{\tilde{\varphi}}]}{\mathbb{E}[V_0^{\varphi}]}\Bigg)\Bigg)\pi(\mathrm{d}\varphi)\,\pi(\mathrm{d}\tilde{\varphi}),$$

which yields the result. □

Proof of Proposition 4.3.19. To show the assertion for $\kappa < \bar{\kappa}$ we use the fact that $\mathrm{P}[\bar{V}_0^{(3)} > x] \le \mathrm{P}[V_0^{\bar{\varphi}} > x]$ and proceed as in the proof of Proposition 4.3.5. For the other cases, observe that

$$\bar{V}_0^{(3)} \stackrel{\mathrm{d}}{=} \frac{\beta}{\eta} + \int_{\mathbb{R}_+}\int_{\Phi_L}\mathrm{e}^{-\eta t}\varphi V_{t-}^{\varphi}\,\Lambda^S(\mathrm{d}t, \mathrm{d}\varphi) = \sum_{i=1}^{\infty}\mathrm{e}^{-\eta T_i}\varphi_i V_{T_i-}^{\varphi_i}\Delta S_{T_i},$$

where $(T_i)_{i\in\mathbb{N}}$ are the jump times of S and $(\varphi_i)_{i\in\mathbb{N}}$ is an i.i.d. sequence with common distribution π which is also independent of S. We start by proving that, if I is a measurable subset of Φ_L with $\pi(I) =: p > 0$ and $\varphi \in \Phi_L$, then there are constants $0 < C_*(\varphi, p), C^*(\varphi, p) < \infty$, only dependent on I via p, with

$$\begin{aligned} C_*(\varphi, p) &= \liminf_{x\to\infty} x^{\kappa(\varphi)}\mathrm{P}\Big[\sum_{\varphi_i\in I}\mathrm{e}^{-\eta T_i}\varphi V_{T_i-}^{\varphi}\Delta S_{T_i} > x\Big] \\ &\le \limsup_{x\to\infty} x^{\kappa(\varphi)}\mathrm{P}\Big[\sum_{\varphi_i\in I}\mathrm{e}^{-\eta T_i}\varphi V_{T_i-}^{\varphi}\Delta S_{T_i} > x\Big] = C^*(\varphi, p) \end{aligned} \tag{4.6.13}$$

and moreover, if $p \to 0$, then $C_*(\varphi, p), C^*(\varphi, p) \to 0$.

We abbreviate the sum in (4.6.13) by $V(\varphi, I)$ or $V(I)$. Since the sequence $(\varphi_i)_{i\in\mathbb{N}}$ is independent of everything else, the distribution of $V(I)$ only depends on p, which means that the constants $C_*(\varphi, p) =: C_*(p)$ and $C^*(\varphi, p) =: C^*(p)$ only depend on p. Also, they are obviously decreasing in p. Hence, for the claimed convergence to 0, it suffices to show $C^*(2^{-n}) \le ((1+2^{-\kappa(\varphi)})/2)^n C(\varphi)$ for all $n \in \mathbb{N}_0$, where $C(\varphi)$ is the tail constant of V_0^{φ} as in the proof of Proposition 4.3.5. The case $n = 0$ corresponds to $V(I) \stackrel{\mathrm{d}}{=} V_0^{\varphi}$ and the statement is clear. For $n \ge 1$, find a set I' disjoint with I such that $\pi(I') = \pi(I) = 2^{-n}$ and therefore $\pi(J) = 2^{-(n-1)}$ for $J = I \cup I'$. Since

$$\begin{aligned} \mathrm{P}[V(J) > x] &= \mathrm{P}[V(I) + V(I') > x] \ge 2\mathrm{P}[V(I) > x] - \mathrm{P}[V(I) > x, V(I') > x] \\ &\ge 2\mathrm{P}[V(I) > x] - \mathrm{P}[V(J) > 2x], \end{aligned}$$

we have by induction

$$C^*(2^{-n}) = \limsup_{x\to\infty} x^{\kappa(\varphi)}\mathrm{P}[V(I) > x] \leq \frac{1+2^{-\kappa(\varphi)}}{2} C^*(2^{-(n-1)}).$$

It remains to show that $C^*(p) < \infty$ and $C_*(p) > 0$ for all $p > 0$. Again by monotonicity, the first inequality is obvious and in the second inequality we only need to consider $p = 1/n$. To this end, partition Φ_L into n disjoint sets $(I_k)_{k=1,\dots,n}$, each with $\pi(I_k) = 1/n$. Then observe that

$$\mathrm{P}[V_0^{\varphi} > x] \leq \mathrm{P}[V(I_1) > x/n \text{ or } \dots \text{ or } V(I_n) > x/n] \leq n\mathrm{P}[V(I_1) > x/n],$$

which implies

$$C_*(1/n) = \liminf_{x\to\infty} x^{\kappa(\varphi)}\mathrm{P}[V(I_1) > x] \geq Cn^{-(\kappa(\varphi)+1)} > 0.$$

Let us come back to the main line of the proof of Proposition 4.3.19. If $\varphi < \bar{\varphi}$, then we have by the above

$$\liminf_{x\to\infty} x^{\kappa}\mathrm{P}[\bar{V}_0^{(3)} > x] \geq \liminf_{x\to\infty} x^{\kappa}\mathrm{P}\Big[V(\varphi, [\varphi, \bar{\varphi}]) > x\Big] \to \infty$$

for all $\kappa > \kappa(\varphi)$ and therefore, by the same argument as in the proof of Proposition 4.3.5, for all $\kappa > \bar{\kappa}$.

Next, consider the case $\kappa = \bar{\kappa}$ and $\bar{p} = 0$. Then, again by the above and the proof of [89, Lemma 2]

$$\begin{aligned}\limsup_{x\to\infty} x^{\bar{\kappa}}\mathrm{P}[\bar{V}_0^{(3)} > x] &\leq \limsup_{x\to\infty} x^{\bar{\kappa}}\mathrm{P}\Big[V(\varphi, (0, \varphi]) + V(\bar{\varphi}, (\varphi, \bar{\varphi}]) > x\Big] \\ &= \limsup_{x\to\infty} x^{\bar{\kappa}}\mathrm{P}\Big[V(\bar{\varphi}, (\varphi, \bar{\varphi}]) > x\Big] = C^*(\bar{\varphi}, \pi((\varphi, \bar{\varphi}])),\end{aligned}$$

which converges to 0 as $\varphi \to \bar{\varphi}$. For the case $\bar{p} > 0$ first decompose

$$\bar{V}_0^{(3)} = \frac{\beta}{\eta} + \sum_{\varphi_i \neq \bar{\varphi}} \mathrm{e}^{-T_i}\varphi_i V_{T_i-}^{\varphi_i}\Delta S_{T_i} + V(\bar{\varphi}, \{\bar{\varphi}\}) =: \frac{\beta}{\eta} + Z + V(\bar{\varphi}, \{\bar{\varphi}\})$$

and observe that $\limsup_{x\to\infty} x^{\bar{\kappa}}\mathrm{P}[Z > x] = 0$ by the results so far. Reading along the lines of the proof of [89, Lemma 2], we obtain

$$\begin{aligned}\liminf_{x\to\infty} x^{\bar{\kappa}}\mathrm{P}[\bar{V}_0^{(3)} > x] &= \liminf_{x\to\infty} x^{\bar{\kappa}}\mathrm{P}[V(\bar{\varphi}, \{\bar{\varphi}\}) > x] = C_*(\bar{\varphi}, \bar{p}), \\ \limsup_{x\to\infty} x^{\bar{\kappa}}\mathrm{P}[\bar{V}_0^{(3)} > x] &= \limsup_{x\to\infty} x^{\bar{\kappa}}\mathrm{P}[V(\bar{\varphi}, \{\bar{\varphi}\}) > x] = C^*(\bar{\varphi}, \bar{p}),\end{aligned}$$

which finishes the proof. □

4.6.4 Proofs for Section 4.4

Proof of Theorem 4.4.1. First observe that the assumption that π has support in $\Phi_L^{(\kappa)}$ implies $\mathbb{E}\big[(\bar{V}_t^{(1)})^\kappa\big] < \infty$. Therefore, $\mathbb{E}[L_1^{\varphi_1}] = 0$ implies

$$\mathbb{E}[\Delta^r G_t^{(1)}] = \mathbb{E}\left[\int_{(t,t+r]} \sqrt{\bar{V}_s^{(1)}}\,\mathrm{d}L_s^{\varphi_1}\right] = 0.$$

Next assume $\mathbb{E}[L_1^2] < \infty$. Using integration by parts and the fact that $G^{(1)}$ has stationary increments, we have

$$\begin{aligned}\mathbb{E}[(\Delta^r G_t^{(1)})^2] = \mathbb{E}[(G_r^{(1)})^2] &= 2\mathbb{E}\left[\int_{(0,r]} G_{s-}^{(1)}\sqrt{\bar{V}_{s-}^{(1)}}\,\mathrm{d}L_s^{\varphi_1}\right] + \mathbb{E}\left[\int_{(0,r]} \bar{V}_{s-}^{(1)}\,\mathrm{d}[L^{\varphi_1}, L^{\varphi_1}]_s\right] \\ &= 0 + \mathrm{Var}[L_1]\mathbb{E}[\bar{V}_0^{(1)}]r,\end{aligned}$$

which, together with Proposition 4.3.12 and the relation between S and L in (4.2.1), gives the stated formula. Furthermore, for $h \geq r > 0$ we have, in view of the above computations and again using integration by parts,

$$\begin{aligned}\mathrm{Cov}[\Delta^r G_t^{(1)}, \Delta^r G_{t+h}^{(1)}] &= \mathbb{E}\left[\Delta^r G_t^{(1)} \Delta^r G_{t+h}^{(1)}\right] \\ &= \mathbb{E}\left[\int_{(0,t+h+r]} \mathbb{1}_{(t,t+r]}(s)\sqrt{\bar{V}_{s-}^{(1)}}\,\mathrm{d}L_s^{\varphi_1} \int_{(0,t+h+r]} \mathbb{1}_{(t+h,t+h+r]}(u)\sqrt{\bar{V}_{u-}^{(1)}}\,\mathrm{d}L_u^{\varphi_1}\right] \\ &= \mathbb{E}\left[\int_{(0,t+h+r]} \mathbb{1}_{(t,t+r]}(s)\mathbb{1}_{(t+h,t+h+r]}(s)\bar{V}_{s-}^{(1)}\,\mathrm{d}[L^{\varphi_1}, L^{\varphi_1}]_s\right] \\ &\quad + \mathbb{E}\left[\int_{(0,t+h+r]} \left(\int_{(0,u]} \mathbb{1}_{(t,t+r]}(s)\sqrt{\bar{V}_{s-}^{(1)}}\,\mathrm{d}L_s^{\varphi_1}\right) \mathbb{1}_{(t+h,t+h+r]}(u)\sqrt{\bar{V}_{u-}^{(1)}}\,\mathrm{d}L_u^{\varphi_1}\right] \\ &\quad + \mathbb{E}\left[\int_{(0,t+h+r]} \left(\int_{(0,u]} \mathbb{1}_{(t+h,t+h+r]}(s)\sqrt{\bar{V}_{s-}^{(1)}}\,\mathrm{d}L_s^{\varphi_1}\right) \mathbb{1}_{(t,t+r]}(u)\sqrt{\bar{V}_{u-}^{(1)}}\,\mathrm{d}L_u^{\varphi_1}\right] \\ &= 0.\end{aligned}$$

To compute the covariance of the squared increments, let $\mathbb{G}^{(1)} = (\mathcal{G}_t^{(1)})_{t\geq 0}$ denote the augmented natural filtration of $(L^{\varphi_i})_{i\in\mathbb{N}}$ and observe that

$$\begin{aligned}\mathbb{E}\left[(\Delta^r G_0^{(1)})^2(\Delta^r G_h^{(1)})^2\right] &= \mathbb{E}\left[\mathbb{E}\left[(\Delta^r G_0^{(1)})^2(\Delta^r G_h^{(1)})^2|\mathcal{G}_r^{(1)}\right]\right] \\ &= \mathbb{E}\left[(\Delta^r G_0^{(1)})^2\mathbb{E}\left[(\Delta^r G_h^{(1)})^2|\mathcal{G}_r^{(1)}\right]\right],\end{aligned}$$

where again by integration by parts

$$\begin{aligned}\mathbb{E}\left[(\Delta^r G_h^{(1)})^2|\mathcal{G}_r^{(1)}\right] &= \mathbb{E}\left[\left(\int_{(h,h+r]}\sqrt{\bar{V}_s^{(1)}}\,\mathrm{d}L_s^{\varphi_1}\right)^2\middle|\mathcal{G}_r^{(1)}\right]\\ &= 2\mathbb{E}\left[\int_{(h,h+r]}\left(\int_{(0,s]}\sqrt{\bar{V}_{u-}^{(1)}}\,\mathrm{d}L_u^{\varphi_1}\right)\sqrt{\bar{V}_{s-}^{(1)}}\mathrm{d}L_s^{\varphi_1}\middle|\mathcal{G}_r^{(1)}\right]\\ &\quad+\mathbb{E}\left[\int_{(h,h+r]}\bar{V}_{s-}^{(1)}\,\mathrm{d}[L^{\varphi_1},L^{\varphi_1}]_s\middle|\mathcal{G}_r^{(1)}\right]\\ &= 0+\mathbb{E}[L_1^2]\int_{(h,h+r]}\mathbb{E}[\bar{V}_{s-}^{(1)}|\mathcal{G}_r^{(1)}]\,\mathrm{d}s.\end{aligned}$$

Next, for $s>r$ we obtain, using the notation as in the proof of Proposition 4.3.11,

$$\begin{aligned}\mathbb{E}\left[\bar{V}_s^{(1)}|\mathcal{G}_r^{(1)}\right] &= \int_{\Phi_L^{(1)}}\mathbb{E}\left[V_s^{\varphi}|\mathcal{G}_r^{(1)}\right]\,\pi(\mathrm{d}\varphi)=\int_{\Phi_L^{(1)}}\mathbb{E}\left[(A_{r,s}^{\varphi}V_r^{\varphi}+B_{r,s}^{\varphi})|\mathcal{G}_r^{(1)}\right]\,\pi(\mathrm{d}\varphi)\\ &= \int_{\Phi_L^{(1)}}\left(\mathbb{E}[A_{r,s}^{\varphi}]V_r^{\varphi}+\mathbb{E}[B_{r,s}^{\varphi}]\right)\,\pi(\mathrm{d}\varphi)\\ &= \int_{\Phi_L^{(1)}}\left(\mathrm{e}^{(s-r)\Psi(1,\varphi)}V_r^{\varphi}+\mathbb{E}[V_0^{\varphi}]\left(1-\mathrm{e}^{(s-r)\Psi(1,\varphi)}\right)\right)\,\pi(\mathrm{d}\varphi).\end{aligned}$$

Together with the preceding computations, this yields

$$\begin{aligned}&\mathrm{Cov}[(\Delta^r G_0^{(1)})^2,(\Delta^r G_h^{(1)})^2]\\ &\quad=\mathbb{E}\left[(\Delta^r G_0^{(1)})^2\mathbb{E}[L_1^2]\int_{(h,h+r]}\mathbb{E}[\bar{V}_{s-}^{(1)}|\mathcal{G}_r^{(1)}]\,\mathrm{d}s\right]-\mathbb{E}\left[(\Delta^r G_0^{(1)})^2\right]\mathbb{E}\left[(\Delta^r G_h^{(1)})^2\right]\\ &\quad=\mathbb{E}[L_1^2]\mathbb{E}\Bigg[(\Delta^r G_0^{(1)})^2\int_{(h,h+r]}\int_{\Phi_L^{(1)}}\Big(\mathrm{e}^{(s-r)\Psi(1,\varphi)}V_r^{\varphi}\\ &\qquad+\mathbb{E}[V_0^{\varphi}]\left(1-\mathrm{e}^{(s-r)\Psi(1,\varphi)}\right)\Big)\pi(\mathrm{d}\varphi)\,\mathrm{d}s\Bigg]-\left(\mathbb{E}\left[(\Delta^r G_0^{(1)})^2\right]\right)^2\\ &\quad=\mathbb{E}[L_1^2]\mathbb{E}\Bigg[(\Delta^r G_0^{(1)})^2\int_{\Phi_L^{(1)}}\Bigg(\frac{1}{\Psi(1,\varphi)}\left(\mathrm{e}^{h\Psi(1,\varphi)}-\mathrm{e}^{(h-r)\Psi(1,\varphi)}\right)(V_r^{\varphi}-\mathbb{E}[V_0^{\varphi}])\\ &\qquad+r\mathbb{E}[V_0^{\varphi}]\Bigg)\pi(\mathrm{d}\varphi)\Bigg]-\left(\mathbb{E}\left[(\Delta^r G_0^{(1)})^2\right]\right)^2\\ &\quad=\mathbb{E}[L_1^2]\int_{\Phi_L^{(1)}}\frac{1}{\Psi(1,\varphi)}\left(\mathrm{e}^{h\Psi(1,\varphi)}-\mathrm{e}^{(h-r)\Psi(1,\varphi)}\right)\\ &\qquad\times\left(\mathbb{E}[(\Delta^r G_0^{(1)})^2V_r^{\varphi}]-\mathbb{E}[(\Delta^r G_0^{(1)})^2]\mathbb{E}[V_r^{\varphi}]\right)\,\pi(\mathrm{d}\varphi)\\ &\qquad+\mathbb{E}[(\Delta^r G_0^{(1)})^2]r\mathbb{E}[L_1^2]\int_{\Phi_L^{(1)}}\mathbb{E}[V_0^{\varphi}]\,\pi(\mathrm{d}\varphi)-\left(\mathbb{E}\left[(\Delta^r G_0^{(1)})^2\right]\right)^2\end{aligned}$$

$$= \mathbb{E}[L_1^2] \int_{\Phi_L^{(1)}} \frac{1}{\Psi(1,\varphi)} \left(\mathrm{e}^{h\Psi(1,\varphi)} - \mathrm{e}^{(h-r)\Psi(1,\varphi)} \right) \mathrm{Cov}[(\Delta^r G_0^{(1)})^2, V_r^\varphi] \, \pi(\mathrm{d}\varphi).$$

It remains to prove $\mathrm{Cov}[(\Delta^r G_0^{(1)})^2, V_r^\varphi] \geq 0$ with strict inequality if $\pi(\{\varphi\}) > 0$ in order to obtain the claimed positivity of the covariance of the squared increments. Again using integration by parts, we get

$$(\Delta^r G_0^{(1)})^2 = \left(\int_{(0,r]} \sqrt{\bar{V}_{s-}^{(1)}} \, \mathrm{d}L_s^{\varphi_1} \right)^2 = 2M_r + \int_{(0,r]} \bar{V}_{s-}^{(1)} \, \mathrm{d}[L^{\varphi_1}, L^{\varphi_1}]_s,$$

where

$$M_r := \int_{(0,r]} \sqrt{\bar{V}_{s-}^{(1)}} \left(\int_{(0,s)} \sqrt{\bar{V}_{u-}^{(1)}} \, \mathrm{d}L_u^{\varphi_1} \right) \mathrm{d}L_s^{\varphi_1}$$

satisfies $\mathbb{E}[M_r] = 0$ due to $\mathbb{E}[L_1] = 0$ and

$$\begin{aligned} \mathbb{E}[M_r V_r^\varphi] &= \mathbb{E}\left[\int_{(0,r]} M_s(\beta - \eta V_s^\varphi) \, \mathrm{d}s \right] + \mathbb{E}\left[\int_{(0,r]} M_{s-} \varphi V_{s-}^\varphi \, \mathrm{d}S_s^\varphi \right] + \mathbb{E}\left[\int_{(0,r]} V_{s-}^\varphi \, \mathrm{d}M_s \right] \\ &\quad + \mathbb{E}\Big[[V^\varphi, M]_r \Big] \\ &= \Psi(1,\varphi) \int_{(0,r]} \mathbb{E}[M_s V_s^\varphi] \, \mathrm{d}s + \mathbb{E}\Big[[V^\varphi, M]_r \Big]. \end{aligned} \tag{4.6.14}$$

Applying $\int_{\mathbb{R}} y^3 \, \nu_L(\mathrm{d}y) = 0$ and the independence of L^φ and L^{φ_1}, if $\varphi \neq \varphi_1$, we have

$$\begin{aligned} \mathbb{E}\Big[[V^\varphi, M]_r \Big] &= \varphi \mathbb{E}\left[\int_{(0,r]} V_{s-}^\varphi \sqrt{\bar{V}_{s-}^{(1)}} \left(\int_{(0,s)} \sqrt{\bar{V}_{u-}^{(1)}} \, \mathrm{d}L_u^{\varphi_1} \right) \mathrm{d}[L^{\varphi_1}, S^\varphi]_s \right] \\ &= \begin{cases} 0 & \text{if } \varphi \neq \varphi_1 \\ \varphi \displaystyle\int_{\mathbb{R}} y^3 \, \nu_L(\mathrm{d}y) \int_{(0,r]} \mathbb{E}\left[V_{s-}^\varphi \sqrt{\bar{V}_{s-}^{(1)}} \left(\int_{(0,s)} \sqrt{\bar{V}_{u-}^{(1)}} \, \mathrm{d}L_u^{\varphi_1} \right) \right] \mathrm{d}s & \text{if } \varphi = \varphi_1 \end{cases} \\ &= 0 \end{aligned} \tag{4.6.15}$$

Thus, (4.6.14) together with the fact that $\mathbb{E}[M_0 V_0^\varphi] = 0$ implies that $\mathbb{E}[M_r V_r^\varphi] = 0$ for all $r \geq 0$. As a consequence, we have

$$\begin{aligned} \mathrm{Cov}[(\Delta^r G_0^{(1)})^2, V_r^\varphi] &= \mathrm{Cov}\left[2M_r + \int_{(0,r]} \bar{V}_{s-}^{(1)} \, \mathrm{d}[L^{\varphi_1}, L^{\varphi_1}]_s, V_r^\varphi \right] \\ &= \mathbb{E}\left[V_r^\varphi \int_{(0,r]} \bar{V}_{s-}^{(1)} \, \mathrm{d}[L^{\varphi_1}, L^{\varphi_1}]_s \right] - \mathbb{E}[V_1^\varphi] \mathbb{E}\left[\int_{(0,r]} \bar{V}_{s-}^{(1)} \, \mathrm{d}[L^{\varphi_1}, L^{\varphi_1}]_s \right] \\ &= \mathbb{E}\left[V_r^\varphi \int_{(0,r]} \bar{V}_{s-}^{(1)} \, \mathrm{d}[L^{\varphi_1}, L^{\varphi_1}]_s \right] - r\mathbb{E}[L_1^2] \mathbb{E}[\bar{V}_0^{(1)}] \mathbb{E}[V_0^\varphi], \end{aligned}$$

where an application of the integration by parts formula yields

$$\begin{aligned}
f(r) &:= \mathbb{E}\left[V_r^{\varphi}\int_{(0,r]}\bar{V}_{s-}^{(1)}\,\mathrm{d}[L^{\varphi_1},L^{\varphi_1}]_s\right]\\
&= \mathbb{E}[L_1^2]\int_{(0,r]}\mathbb{E}[V_s^{\varphi}\bar{V}_s^{(1)}]\,\mathrm{d}s+\beta\int_{(0,r]}\mathbb{E}\left[\int_{(0,s]}\bar{V}_{u-}^{(1)}\,\mathrm{d}[L^{\varphi_1},L^{\varphi_1}]_u\right]\,\mathrm{d}s\\
&\quad+\Psi(1,\varphi)\int_{(0,r]}\mathbb{E}\left[V_s^{\varphi}\int_{(0,s]}\bar{V}_{u-}^{(1)}\,\mathrm{d}[L^{\varphi_1},L^{\varphi_1}]_u\right]\,\mathrm{d}s+\mathbb{E}\left[\int_{(0,r]}\bar{V}_{s-}^{(1)}\,\mathrm{d}\big[S^{\varphi_1},V^{\varphi}\big]_s\right]\\
&= \mathbb{E}[L_1^2]\mathbb{E}[V_0^{\varphi}\bar{V}_0^{(1)}]r+\beta\mathbb{E}[L_1^2]\mathbb{E}[\bar{V}_0^{(1)}]\frac{r^2}{2}+\Psi(1,\varphi)\int_{(0,r]}f(s)\,\mathrm{d}s\\
&\quad+\mathbb{1}_{\{\varphi=\varphi_1\}}\varphi\int_{\mathbb{R}}y^2\,\nu_S(\mathrm{d}y)\mathbb{E}[V_0^{\varphi}\bar{V}_0^{(1)}]r,\\
f(0) &= 0.
\end{aligned}$$

Solving this integral equation yields ($m_2 := \int_{\mathbb{R}} y^2\,\nu_S(\mathrm{d}y)$)

$$\begin{aligned}
f(r) = &\frac{(\mathbb{E}[L_1^2]+\mathbb{1}_{\{\varphi=\varphi_1\}}\varphi m_2)\mathbb{E}[V_0^{\varphi}\bar{V}_0^{(1)}]\Psi(1,\varphi)(\mathrm{e}^{\Psi(1,\varphi)r}-1)}{\Psi(1,\varphi)^2}\\
&+\frac{\beta\mathbb{E}[L_1^2]\mathbb{E}[\bar{V}_0^{(1)}](-\Psi(1,\varphi)r+\mathrm{e}^{\Psi(1,\varphi)r}-1)}{\Psi(1,\varphi)^2},
\end{aligned}$$

which by (4.2.11) yields the claimed positive correlation, since

$$\begin{aligned}
\mathrm{Cov}[(\Delta^r G_0^{(1)})^2,V_r^{\varphi}] &= f(r)-\mathbb{E}[L_1^2]\mathbb{E}[V_0^{\varphi}]\mathbb{E}[\bar{V}_0^{(1)}]r\\
&= \frac{(\mathbb{E}[L_1^2]+\mathbb{1}_{\{\varphi=\varphi_1\}}\varphi m_2)\mathbb{E}[V_0^{\varphi}\bar{V}_0^{(1)}](\mathrm{e}^{\Psi(1,\varphi)r}-1)}{\Psi(1,\varphi)}+\frac{\beta\mathbb{E}[L_1^2]\mathbb{E}[\bar{V}_0^{(1)}](\mathrm{e}^{\Psi(1,\varphi)r}-1)}{\Psi(1,\varphi)^2}\\
&= \frac{\mathrm{e}^{\Psi(1,\varphi)r}-1}{\Psi(1,\varphi)}\left(\mathbb{E}[L_1^2]\mathrm{Cov}[V_0^{\varphi},\bar{V}_0^{(1)}]+\mathbb{1}_{\{\varphi=\varphi_1\}}\varphi\int_{\mathbb{R}}y^2\,\nu_S(\mathrm{d}y)\mathbb{E}[V_0^{\varphi}\bar{V}_0^{(1)}]\right)\\
&\geq 0 \qquad (4.6.16)
\end{aligned}$$

with $\mathrm{Cov}[V_0^{\varphi},\bar{V}_0^{(1)}]=\pi(\{\varphi\})\mathrm{Var}[V_1^{\varphi}]$. □

Proof of Theorem 4.4.2. The proof works similarly to the proof of Theorem 4.4.1 with the obvious changes, when independence of the single COGARCH processes was used (e.g. (4.6.15)). Also replace $\mathbb{G}^{(1)}$ by $\mathbb{G}^{(2)}=(\mathcal{G}_t^{(2)})_{t\in\mathbb{R}}$, the augmented natural filtration of L, and notice that $\mathrm{Cov}[V_0^{\varphi},\bar{V}_0^{(2)}]=\int_{\Phi_L^{(2)}}\mathrm{Cov}[V_0^{\varphi},V_0^{\tilde{\varphi}}]\,\pi(\mathrm{d}\tilde{\varphi})>0$ by Proposition 4.3.11. □

Proof of Theorem 4.4.3. Analogously to the proof of Theorem 4.4.1, one can show that (1) and (2) hold and that for (3) we have

$$\mathbb{E}\left[(\Delta^r G_0^{(3)})^2(\Delta^r G_h^{(3)})^2\right] = \mathbb{E}[L_1^2]\mathbb{E}\left[(\Delta^r G_0^{(3)})^2 \int_{(h,h+r]} \mathbb{E}[\bar{V}_{s-}^{(3)}|\mathcal{G}_r^{(3)}]\,\mathrm{d}s\right], \qquad (4.6.17)$$

where from (4.3.29) and [88, Eq. (4.5)] we have

$$\begin{aligned}
&\mathbb{E}[\bar{V}_{s-}^{(3)}|\mathcal{G}_r^{(3)}] \\
&= \mathrm{e}^{-\eta(s-r)}\bar{V}_r^{(3)} + \beta\mathrm{e}^{-\eta s}\int_{(r,s)} \mathrm{e}^{\eta u}\mathrm{d}u + \mathbb{E}\left[\int_{(r,s)}\int_{\Phi_L^{(2)}} \mathrm{e}^{-\eta(s-u)}\varphi V_{u-}^{\varphi}\,\Lambda^S(\mathrm{d}u,\mathrm{d}\varphi)\middle|\,\mathcal{G}_r^{(3)}\right] \\
&= \mathrm{e}^{-\eta(s-r)}\bar{V}_r^{(3)} + \frac{\beta}{\eta}(1-\mathrm{e}^{-\eta(s-r)}) + \mathbb{E}[S_1]\int_{(r,s]}\int_{\Phi_L^{(2)}} \mathrm{e}^{-\eta(s-u)}\varphi\mathbb{E}[V_{u-}^{\varphi}|\mathcal{G}_r^{(3)}]\,\pi(\mathrm{d}\varphi)\,\mathrm{d}u \\
&= \mathrm{e}^{-\eta(s-r)}\bar{V}_r^{(3)} + \frac{\beta}{\eta}(1-\mathrm{e}^{-\eta(s-r)}) \\
&\quad + \mathbb{E}[S_1]\int_{(r,s]}\int_{\Phi_L^{(2)}} \mathrm{e}^{-\eta(s-u)}\varphi\Big((V_r^{\varphi}-\mathbb{E}[V_0^{\varphi}])\mathrm{e}^{(u-r)\Psi(1,\varphi)} + \mathbb{E}[V_0^{\varphi}]\Big)\,\pi(\mathrm{d}\varphi)\,\mathrm{d}u.
\end{aligned}$$

Applying (4.2.11) we obtain

$$\begin{aligned}
&\mathbb{E}[S_1]\int_{(r,s]}\int_{\Phi_L^{(2)}} \mathrm{e}^{-\eta(s-u)}\varphi\Big((V_r^{\varphi}-\mathbb{E}[V_0^{\varphi}])\mathrm{e}^{(u-r)\Psi(1,\varphi)} + \mathbb{E}[V_0^{\varphi}]\Big)\,\pi(\mathrm{d}\varphi)\,\mathrm{d}u \\
&\quad = \mathbb{E}[S_1]\int_{\Phi_L^{(2)}}\left(\frac{\varphi(V_r^{\varphi}-\mathbb{E}[V_0^{\varphi}])}{\varphi\mathbb{E}[S_1]}\left(\mathrm{e}^{\Psi(1,\varphi)(s-r)}-\mathrm{e}^{-\eta(s-r)}\right)\right. \\
&\qquad \left. + \frac{\varphi\mathbb{E}[V_0^{\varphi}]}{\eta}\left(1-\mathrm{e}^{-\eta(s-r)}\right)\right)\pi(\mathrm{d}\varphi) \\
&\quad = \int_{\Phi_L^{(2)}} \mathrm{e}^{\Psi(1,\varphi)(s-r)}(V_r^{\varphi}-\mathbb{E}[V_0^{\varphi}])\,\pi(\mathrm{d}\varphi) - \mathrm{e}^{-\eta(s-r)}\left(\int_{\Phi_L^{(2)}} V_r^{\varphi}\,\pi(\mathrm{d}\varphi) - \mathbb{E}[\bar{V}_0^{(3)}]\right) \\
&\qquad + \left(1-\mathrm{e}^{-\eta(s-r)}\right)\int_{\Phi_L^{(2)}} \frac{\mathbb{E}[S_1]\varphi}{\eta}\frac{-\beta}{\Psi(1,\varphi)}\,\pi(\mathrm{d}\varphi) \\
&\quad = \int_{\Phi_L^{(2)}} \mathrm{e}^{\Psi(1,\varphi)(s-r)}(V_r^{\varphi}-\mathbb{E}[V_0^{\varphi}])\,\pi(\mathrm{d}\varphi) - \mathrm{e}^{-\eta(s-r)}\left(\int_{\Phi_L^{(2)}} V_r^{\varphi}\,\pi(\mathrm{d}\varphi) - \mathbb{E}[\bar{V}_0^{(3)}]\right) \\
&\qquad - \frac{\beta}{\eta}\left(1-\mathrm{e}^{-\eta(s-r)}\right)\int_{\Phi_L^{(2)}}\left(1+\frac{\eta}{\Psi(1,\varphi)}\right)\pi(\mathrm{d}\varphi) \\
&\quad = \int_{\Phi_L^{(2)}} \mathrm{e}^{\Psi(1,\varphi)(s-r)}(V_r^{\varphi}-\mathbb{E}[V_0^{\varphi}])\,\pi(\mathrm{d}\varphi) - \mathrm{e}^{-\eta(s-r)}\int_{\Phi_L^{(2)}} V_r^{\varphi}\,\pi(\mathrm{d}\varphi) \\
&\qquad - \frac{\beta}{\eta}\left(1-\mathrm{e}^{-\eta(s-r)}\right) + \mathbb{E}[\bar{V}_0^{(3)}]
\end{aligned}$$

such that

$$\mathbb{E}[\bar{V}_{s-}^{(3)}|\mathcal{G}_r^{(3)}] = \mathrm{e}^{-\eta(s-r)}\left(\bar{V}_r^{(3)} - \int_{\Phi_L^{(2)}} V_r^{\varphi}\,\pi(\mathrm{d}\varphi)\right) + \int_{\Phi_L^{(2)}} \mathrm{e}^{\Psi(1,\varphi)(s-r)}(V_r^{\varphi} - \mathbb{E}[V_0^{\varphi}])\,\pi(\mathrm{d}\varphi) + \mathbb{E}[\bar{V}_0^{(3)}].$$

Inserting this into (4.6.17) yields

$$\begin{aligned}
&\mathrm{Cov}[(\Delta^r G_0^{(3)})^2, (\Delta^r G_h^{(3)})^2] \\
&= \mathbb{E}[L_1^2]\mathbb{E}\left[(\Delta^r G_0^{(3)})^2 \int_{(h,h+r]} \mathbb{E}[\bar{V}_{s-}^{(3)}|\mathcal{G}_r^{(3)}]\,\mathrm{d}s\right] - \mathbb{E}[(\Delta^r G_0^{(3)})^2]^2 \\
&= \mathbb{E}[L_1^2]\mathbb{E}\left[(\Delta^r G_0^{(3)})^2 \left(\frac{\mathrm{e}^{-\eta h} - \mathrm{e}^{-\eta(h-r)}}{-\eta}\left(\bar{V}_r^{(3)} - \int_{\Phi_L^{(2)}} V_r^{\varphi}\,\pi(\mathrm{d}\varphi)\right)\right)\right. \\
&\quad \left. + \int_{\Phi_L^{(2)}} \frac{\mathrm{e}^{\Psi(1,\varphi)h} - \mathrm{e}^{\Psi(1,\varphi)(h-r)}}{\Psi(1,\varphi)}(V_r^{\varphi} - \mathbb{E}[V_0^{\varphi}])\,\pi(\mathrm{d}\varphi)\right] \\
&= \mathbb{E}[L_1^2]\left[\frac{\mathrm{e}^{-\eta h} - \mathrm{e}^{-\eta(h-r)}}{-\eta}\mathrm{Cov}[(\Delta^r G_0^{(3)})^2, \bar{V}_r^{(3)}]\right. \\
&\quad \left. + \int_{\Phi_L^{(2)}} \left(\frac{\mathrm{e}^{\Psi(1,\varphi)h} - \mathrm{e}^{\Psi(1,\varphi)(h-r)}}{\Psi(1,\varphi)} - \frac{\mathrm{e}^{-\eta h} - \mathrm{e}^{-\eta(h-r)}}{-\eta}\right)\mathrm{Cov}[(\Delta^r G_0^{(3)})^2, V_r^{\varphi}]\,\pi(\mathrm{d}\varphi)\right].
\end{aligned}$$

Since $\Psi(1,\varphi) > -\eta$ and the function $x \mapsto (\mathrm{e}^{hx} - \mathrm{e}^{(h-r)x})/x$ is increasing in x for $x < 0$, it remains to prove $\mathrm{Cov}[(\Delta^r G_0^{(3)})^2, \bar{V}_r^{(3)}] > 0$ and $\mathrm{Cov}[(\Delta^r G_0^{(3)})^2, V_r^{\varphi}] > 0$. For the latter one, proceed as in the proof of Theorem 4.4.1 and note that $\mathrm{Cov}[V_0^{\varphi}, \bar{V}_0^{(3)}] > 0$. Indeed, using integration by parts,

$$\begin{aligned}
V_r^{\varphi}\bar{V}_r^{(3)} &= V_0^{\varphi}\bar{V}_0^{(3)} + \int_{(0,r]} \bar{V}_{s-}^{(3)}\,\mathrm{d}V_s^{\varphi} + \int_{(0,r]} V_{s-}^{\varphi}\,\mathrm{d}\bar{V}_s^{(3)} + [V^{\varphi}, \bar{V}^{(3)}]_r \\
&= V_0^{\varphi}\bar{V}_0^{(3)} + \int_{(0,r]} \bar{V}_s^{(3)}(\beta - \eta V_s^{\varphi})\,\mathrm{d}s + \int_{(0,r]} \bar{V}_{s-}^{(3)}\varphi V_{s-}^{\varphi}\,\mathrm{d}S_s + \\
&\quad + \int_{(0,r]} V_{s-}^{\varphi}(\beta - \eta\bar{V}_s^{(3)})\,\mathrm{d}s + \int_{(0,r]}\int_{\Phi_L^{(2)}} V_{s-}^{\varphi}\tilde{\varphi}V_{s-}^{\tilde{\varphi}}\,\Lambda^S(\mathrm{d}s, \mathrm{d}\tilde{\varphi}) \\
&\quad + \varphi\int_{(0,r]}\int_{\Phi_L^{(2)}}\int_{\mathbb{R}_+} V_{s-}^{\varphi}\tilde{\varphi}V_{s-}^{\tilde{\varphi}}y^2\,\mu^{\Lambda^S}(\mathrm{d}s, \mathrm{d}\tilde{\varphi}, \mathrm{d}y),
\end{aligned}$$

with $[V^{\varphi}, \bar{V}^{(3)}]_r$ as given in Lemma 4.6.5. Taking expectations, differentiating with respect to r and using the stationarity of $V^{\varphi}\bar{V}^{(3)}$, which is a consequence of Lemma 4.6.3, we find that ($m_1 := \int_{\mathbb{R}_+} y\,\nu_S(\mathrm{d}y)$ and $m_2 := \int_{\mathbb{R}_+} y^2\,\nu_S(\mathrm{d}y)$)

$$\beta(\mathbb{E}[\bar{V}_0^{(3)}] + \mathbb{E}[V_0^{\varphi}]) + (\varphi m_1 - 2\eta)\mathbb{E}[V_0^{\varphi}\bar{V}_0^{(3)}] + (m_1 + \varphi m_2)\int_{\Phi_L^{(2)}} \tilde{\varphi}\mathbb{E}[V_0^{\varphi}V_0^{\tilde{\varphi}}]\,\pi(\mathrm{d}\tilde{\varphi}) = 0,$$

which implies that

$$\begin{aligned}&\operatorname{Cov}[\bar{V}_0^{(3)}, V_0^{\varphi}]\\ &\quad= \frac{\beta(\mathbb{E}[\bar{V}_0^{(3)}] + \mathbb{E}[V_0^{\varphi}]) + (m_1 + \varphi m_2)\int_{\Phi_L^{(2)}} \tilde{\varphi}\mathbb{E}[V_0^{\varphi}V_0^{\tilde{\varphi}}]\,\pi(\mathrm{d}\tilde{\varphi})}{\eta - \Psi(1,\varphi)} - \mathbb{E}[\bar{V}_0^{(3)}]\mathbb{E}[V_0^{\varphi}].\end{aligned}$$

To show the positivity of this term, we can equally well consider its product with $\eta - \Psi(1,\varphi)$, which by (4.2.11), (4.3.30) and (4.6.1) can be simplified to

$$\begin{aligned}&\int_{\Phi_L^{(2)}}(m_1 + \varphi m_2)\tilde{\varphi}\operatorname{Cov}[V_0^{\varphi}, V_0^{\tilde{\varphi}}]\,\pi(\mathrm{d}\tilde{\varphi}) + \beta(\mathbb{E}[\bar{V}_0^{(3)}] + \mathbb{E}[V_0^{\varphi}]) + h(\varphi,\tilde{\varphi})\mathbb{E}[V_0^{\varphi}]\mathbb{E}[\bar{V}_0^{(3)}]\\ &\quad= \int_{\Phi_L^{(2)}}(m_1 + \varphi m_2)\tilde{\varphi}\operatorname{Cov}[V_0^{\varphi}, V_0^{\tilde{\varphi}}]\,\pi(\mathrm{d}\tilde{\varphi}) + \beta^2 m_2\int_{\Phi_L^{(2)}}\frac{\varphi\tilde{\varphi}}{\Psi(1,\varphi)\Psi(1,\tilde{\varphi})}\,\pi(\mathrm{d}\tilde{\varphi})\\ &\quad> 0.\end{aligned}$$

Finally, using the same methods as in the proof of Theorem 4.4.1, one can derive the following analogue of Equation (4.6.16):

$$\operatorname{Cov}[(\Delta^r G_0^{(3)})^2, \bar{V}_0^{(3)}] = g(r) - \mathbb{E}[L_1^2]\mathbb{E}[\bar{V}_0]^2 r,$$

where

$$\begin{aligned}g(r) &= \mathrm{e}^{-\eta r}\left(\int_{(0,r]}\mathrm{e}^{\eta s}\left(a + bs + \int_{\Phi_L^{(2)}} m_1\varphi f(\varphi,s)\,\pi(\mathrm{d}\varphi)\right)\mathrm{d}s\right), \quad r \geq 0,\\ a &= \mathbb{E}[L_1^2]\mathbb{E}[(\bar{V}_0^{(3)})^2] + \int_{\mathbb{R}_+} y^2\,\nu_S(\mathrm{d}y)\int_{\Phi_L^{(2)}}\varphi\mathbb{E}[V_0^{\varphi}\bar{V}_0]\,\pi(\mathrm{d}\varphi),\\ b &= \beta\mathbb{E}[L_1^2]\mathbb{E}[\bar{V}_0^{(3)}] \quad \text{and} \quad f(\varphi,r) = \mathbb{E}\left[V_r^{\varphi}\int_{(0,r]}\bar{V}_{u-}^{(3)}\,\mathrm{d}[L,L]_u\right].\end{aligned}$$

The positivity now follows from

$$\operatorname{Cov}[(\Delta^r G_0^{(3)})^2, \bar{V}_0^{(3)}] \geq \mathrm{e}^{\eta r}\int_{(0,r]}\mathrm{e}^{-\eta s}\,\mathrm{d}s\mathbb{E}[L_1^2]\mathbb{E}[(\bar{V}_0^{(3)})^2] - \mathbb{E}[L_1^2]\mathbb{E}[\bar{V}_0^{(3)}]^2 r$$

and the fact that $\mathrm{e}^{\eta r}\int_{(0,r]}\mathrm{e}^{-\eta s}\,\mathrm{d}s = (\mathrm{e}^{\eta r} - 1)/\eta > r$ for all $r > 0$. □

Chapter 5:
Partial mean field limits in heterogeneous networks

5.1 Introduction

The application of mean field theory to large systems of stochastic differential equations (SDEs) was initiated by McKean's seminal work [100, 101, 102]. In the classical case, an N-dimensional interacting particle system is governed by SDEs of the form

$$\begin{aligned} \mathrm{d}X_i^N(t) &= \frac{1}{N-1}\sum_{j\neq i}\Big(X_j^N(t) - X_i^N(t)\Big)\,\mathrm{d}t + \mathrm{d}B_i(t), \quad t \in \mathbb{R}_+, \\ X_i^N(0) &= X_i(0), \quad i = 1, \dots, N, \end{aligned} \tag{5.1.1}$$

with independent starting random variables $X_i(0)$ and independent Brownian motions B_i. As the number of particles increases, the pair dependencies in this coupled system decrease with order $1/N$ such that a law of large numbers applies (see [131, Thm. 1.4]). Defining

$$\begin{aligned} \mathrm{d}\bar{X}_i^N(t) &= \frac{1}{N-1}\sum_{j\neq i}\Big(\mathbb{E}[\bar{X}_j^N(t)] - \bar{X}_i^N(t)\Big)\,\mathrm{d}t + \mathrm{d}B_i(t), \quad t \in \mathbb{R}_+, \\ \bar{X}_i^N(0) &= X_i(0), \quad i = 1, \dots, N, \end{aligned} \tag{5.1.2}$$

there exists for every $T \in \mathbb{R}_+$ a constant $C(T) \in \mathbb{R}_+$ independent of N such that

$$\sup_{i=1,\dots,N} \mathbb{E}\left[\sup_{t\in[0,T]} |X_i^N(t) - \bar{X}_i^N(t)|^2\right]^{1/2} \leq \frac{C(T)}{\sqrt{N}}. \tag{5.1.3}$$

In other words, in a large system, the behaviour of a fixed number of particles evolving according to (5.1.1) is well described by the so-called *mean field system* (5.1.2),

where all stochastic processes are stochastically independent, a phenomenon that is called *propagation of chaos*. Thus, mean field theory provides a model simplification by reducing a many-body problem as in (5.1.1) to a one-body problem as in (5.1.2) with explicit L^2-estimates on the occurring error. Moreover, it can be shown that the empirical measure of the particles satisfies a large deviation principle as $N \to \infty$, see [50, 96]. There exists a huge literature dealing with this or related topics, and we only mention the review papers [70, 131], where one can also find further references.

The systems (5.1.1) and (5.1.2) describe statistically equal or exchangeable particles: any permutation of the indices $i \in \{1, \dots, N\}$ leads to a system with the same distribution (cf. [135]). In particle physics such an assumption is certainly reasonable and underlies many other similar models of mean field type, see for example the two treatises [132, 133] for numerous examples.

However, when mean field models are considered in other disciplines than statistical mechanics, the homogeneity assumption may not be appropriate in all situations. For instance, in [33, 78] the processes (5.1.1) are used to model the wealth of trading agents in an economy, who are typically far from being equal in their trading behaviour (there are "market makers" and others). Similarly, the stochastic Cucker-Smale model that is considered in [1, 31] describes the "flocking" phenomenon of individuals. Also here it only seems natural that one or several "leaders" may have a distinguished role, setting them apart from the remaining system. Moreover, in systemic risk modelling the particles represent financial institutions that interact with each other through mutual exposures, see [23, 66, 87] for some approaches in this direction. The different players in the banking sector vary considerably in size and importance, which is obvious by the fact that some banks have been considered as too big to fail during the financial crisis since 2007. Further fields of applications where mean field theory is used for interacting particle systems include genetic algorithms [54], neuron modelling (see [67] and references therein), epidemics modelling [95] and Monte Carlo integration [53].

Partly triggered by the examples in the previous paragraph, this research aims to investigate deviations from homogeneous systems to heterogeneous systems. First, we allow for different interaction rates between pairs (instead of $1/(N-1)$ throughout), and second, we permit the subsistence of a core–periphery structure in the mean field limit, that is, some particles may have a non-vanishing influence even when the system becomes large. Another restriction we will relax in our analysis concerns the driving noises of the interacting SDEs: instead of independence we ex-

plicitly allow for different degrees of dependence in the noise terms, even asymptotically. Until now there exists only a very small amount of literature that generalizes (5.1.1) in these directions: in [36, 87, 109, 110] the particles are divided into finitely many groups within which they are homogeneous (and the number of members in both groups must tend to infinity for the law of large numbers), and [38], where one major agent exists and propagation of chaos for the minor agents is considered conditioned on the major one. Other papers that consider general heterogeneous systems include [52], where the propagation of chaos result is *assumed*, and [63, 64, 71], where a law of large numbers for the empirical measure is proved under various conditions. Regarding the last-mentioned papers, two aspects are worthwhile to notice. First, assuming that finitely many core particles do exist in the system, their contribution to the empirical distribution becomes less and less for $N \to \infty$ although their impact may very well stay high. Thus, in this case the empirical distribution may fail to describe the behaviour of the system as a whole. Second, whereas for homogeneous systems the convergence of the empirical measure is equivalent to the existence of a mean field limit in the sense of (5.1.3) (see e.g. [131, Prop. 2.2(i)]), this is no longer true for heterogeneous systems. For core particles the left-hand side of (5.1.3) need not converge to 0 even if the empirical distribution converges, say, to a deterministic limit. For example, in the case of [38] with one core particle, an unconditional propagation of chaos result does not hold for this particle without further assumptions (even if it does for the periphery particles).

Due to the two aforementioned reasons, we will *not* work with the empirical distribution in this paper but state and prove mean field limit theorems for the particles on the process level. In Section 5.2 we start by introducing the precise interacting particle model we want to investigate. Then we define a corresponding *partial mean field model*, for which we prove a law of large numbers type result (Theorem 5.3.1) with explicit convergence rates in Section 5.3. It generalizes (5.1.3) by taking into account the different kinds of heterogeneity due to varying pair interaction rates, a distinction between important/core and less important/periphery pair relationships, and interdependencies between the driving noise terms.

The main difficulty here is to identify the correct rates that govern the distance between the original system and the mean field approximation. As we will see, a total of twelve rates is required, each expressing a connectivity property of the underlying interaction and correlation networks. In order to elucidate the meaning of each rate, we discuss three exemplary situations in detail. In Section 5.3.1, in particular

in Example 5.3.4, we show that in the quasi-homogeneous case all twelve rates typically boil down to a single rate like in (5.1.3). In Section 5.3.2, we explain why the prerequisites for Theorem 5.3.1 in the heterogeneous case are essentially sparsity assumptions on the particle network, which are satisfied for instance if this network is generated from a preferential attachment mechanism, see Section 5.3.3. In order to show the last statement, we have to derive the asymptotics of the maximal in- and out-degrees of directed preferential attachment graphs, see Lemma 5.3.8. This result may be of independent interest and generalizes that of [105] for undirected graphs.

The second main result of our paper is a large deviation principle for the difference $X^N - \bar{X}^N$, which is presented in Section 5.4 as Theorem 5.4.1. In contrast to homogeneous systems, where such a principle is proved for the empirical measure (see [50, 96]), we work on the process level again and therefore need to require the existence of all exponential moments. Furthermore, due to heterogeneity, we do not obtain an explicit formula for the large deviation rate function, but a variational representation as Fenchel-Legendre transform. The final Section 5.5 contains the proofs.

5.2 The model

Before we introduce the model we analyze in this paper, we list a number of notations that will be employed throughout the paper.

$\mathbb{R}_+$	the set $[0,\infty)$ of *positive* real numbers;
$[z]$	the largest integer smaller or equal to $z \in \mathbb{R}$;
$\mathbb{N}$	the natural numbers $\{1,2,\dots\}$;
A, x	the typical notation for a matrix $A = (A_{ij} \colon i,j \in \mathbb{N}) \in \mathbb{R}^{\mathbb{N}\times\mathbb{N}}$ and a vector $x = (x_i \colon i \in \mathbb{N})' \in \mathbb{R}^{\mathbb{N}}$, with all binary relations such as $\leq$, or operations relying on them such as the absolute value $\lvert\cdot\rvert$ or taking the supremum being understood componentwise when applied to matrices and vectors;
$(\cdot)'$	the transposition operator;
AB, Ax, e^A	matrix–matrix and matrix–vector multiplication and the matrix exponential, all defined in analogy to the finite-dimensional case, provided that the involved series converge;
$x.y$	the entrywise product $x.y = (x_i y_i \colon i \in \mathbb{N})'$ for $x, y \in \mathbb{R}^{\mathbb{N}}$;

$\lvert A\rvert_\infty, \lvert x\rvert_\infty$	$\lvert A\rvert_\infty := \sup_{i\in\mathbb{N}} \sum_{j\in\mathbb{N}} \lvert A_{ij}\rvert$ and $\lvert x\rvert_\infty := \sup_{i\in\mathbb{N}} \lvert x_i\rvert$ for $A \in \mathbb{R}^{\mathbb{N}\times\mathbb{N}}$ and $x \in \mathbb{R}^{\mathbb{N}}$;
$\lvert A\rvert_{\mathrm{d}}$	$\lvert A\rvert_{\mathrm{d}} := \sup_{i\in\mathbb{N}} \lvert A_{ii}\rvert$ for matrices A;
$A^\times$	the matrix A with all diagonal entries set to 0;
I	the identity matrix in $\mathbb{R}^{\mathbb{N}\times\mathbb{N}}$ or $\mathbb{R}^{d\times d}$ for some $d \in \mathbb{N}$;
L^p	the space $L^p(\Omega, \mathcal{F}, \mathbb{P})$, $p \in [1,\infty]$, endowed with the topology induced by $\lVert X\rVert_{L^p} := \mathbb{E}[\lvert X\rvert^p]^{1/p}$, and to be understood entrywise when applied to matrix- or vector-valued random variables;
$\mathbb{E}[X]$, $\mathrm{Var}[X]$	componentwise expectation and variance for random variables in $\mathbb{R}^{\mathbb{N}\times\mathbb{N}}$ or $\mathbb{R}^{\mathbb{N}}$;
$\mathrm{Cov}[X,Y]$, $\mathrm{Cov}[X]$	the matrices whose (ij)-th entry is $\mathrm{Cov}[X_i, Y_j]$ and $\mathrm{Cov}[X_i, X_j]$, respectively, when X and Y are random vectors;
x^*	$x^*(t) := \sup_{s\in[0,t]} \lvert x(s)\rvert$ for $t \in \mathbb{R}_+$ and functions $x\colon \mathbb{R}_+ \to \mathbb{R}$, again considered entrywise when x takes values in $\mathbb{R}^{\mathbb{N}\times\mathbb{N}}$ or $\mathbb{R}^{\mathbb{N}}$;
D_T^d, D_T^∞	the space of $\mathbb{R}^d$-valued (resp. $\mathbb{R}^{\mathbb{N}}$-valued) functions on $[0,T]$ whose coordinates are all càdlàg functions;
C_T^d, C_T^∞	elements of D_T^d and D_T^∞ where each coordinate is a continuous function;
AC_T^d, AC_T^∞	elements of D_T^d and D_T^∞ where each coordinate is an absolutely continuous function;
$\mathcal{D}_T^d$, $\mathcal{D}_T^\infty$	the σ-field on D_T^d (resp. D_T^∞) generated by the evaluation maps $\pi_t(x) = x(t)$, $x \in D_T^d$ (resp. D_T^∞), for $t \in [0,T]$;
U,J_1	the uniform topology and the Skorokhod topology on D_T^d and D_T^∞ (in the latter case they are defined via the product of the d-dimensional topologies);
M_T^d	the space of all $(\theta_1, \ldots, \theta_d)$ where each θ_i is a signed Borel measure on $[0,T]$ of finite total variation $\lvert\theta_i\rvert([0,T])$

Given a stochastic basis $(\Omega, \mathcal{F}, \mathbb{F} = (\mathcal{F}(t))_{t\in\mathbb{R}_+}, \mathbb{P})$ satisfying the usual hypotheses of completeness and right-continuity, we investigate a network described by the following *interacting particle system* (IPS):

$$\begin{aligned} \mathrm{d}X_i(t) = \sum_{j=1}^\infty a_{ij}(t)X_j(t)\,\mathrm{d}t + \sum_{j=1}^\infty \sigma_{ij}(t)X_j(t-)\,\mathrm{d}L_i(t) + \sum_{j=1}^\infty f_{ij}(t)\,\mathrm{d}B_j(t) \\ + \sum_{j=1}^\infty \rho_{ij}(t)\,\mathrm{d}M_j(t), \quad t \in \mathbb{R}_+, \quad i \in \mathbb{N}, \end{aligned} \tag{5.2.1}$$

subjected to some $\mathcal{F}(0)$-measurable $\mathbb{R}^{\mathbb{N}}$-valued initial condition $X(0)$. We will also use the more compact form

$$\mathrm{d}X(t) = a(t)X(t)\,\mathrm{d}t + \sigma(t)X(t-).\mathrm{d}L(t) + f(t)\,\mathrm{d}B(t) + \rho(t)\,\mathrm{d}M(t), \quad t \in \mathbb{R}_+, \tag{5.2.2}$$

for (5.2.1). The ingredients hereby satisfy the following conditions:

- The two measurable functions $t \mapsto a(t)$ and $t \mapsto \sigma(t)$ are decomposed into $a = a^{\mathrm{C}} + a^{\mathrm{P}}$ and $\sigma = \sigma^{\mathrm{C}} + \sigma^{\mathrm{P}}$ such that for all $T \in \mathbb{R}_+$, $i, j \in \mathbb{N}$ and $\diamond \in \{\mathrm{C}, \mathrm{P}\}$

$$A^{\diamond}_{ij}(T) := \sup_{t\in[0,T]} |a^{\diamond}_{ij}(t)| < \infty, \quad \Sigma^{\diamond}_{ij}(T) := \sup_{t\in[0,T]} |\sigma^{\diamond}_{ij}(t)| < \infty. \tag{5.2.3}$$

  We define $A_{ij}(T) := A^{\mathrm{C}}_{ij}(T) + A^{\mathrm{P}}_{ij}(T)$ and $\Sigma_{ij}(T) := \Sigma^{\mathrm{C}}_{ij}(T) + \Sigma^{\mathrm{P}}_{ij}(T)$.

- L is an $\mathbb{R}^{\mathbb{N}}$-valued $\mathbb{F}$-Lévy process (i.e. an $\mathbb{F}$-adapted Lévy process whose increments are independent of the past σ-fields in $\mathbb{F}$) with finite second moment and mean 0.

- M is an $\mathbb{R}^{\mathbb{N}}$-valued square-integrable martingale on any finite time interval, and B is an $\mathbb{R}^{\mathbb{N}}$-valued predictable process such that each coordinate process is of locally finite variation. We assume that B and the predictable quadratic variation process $\langle M, M\rangle$ have progressively measurable Lebesgue densities $b\colon \Omega \times \mathbb{R}_+ \to \mathbb{R}^{\mathbb{N}}$ and $c\colon \Omega \times \mathbb{R}_+ \to \mathbb{R}^{\mathbb{N}\times\mathbb{N}}$.

- f is the sum of two deterministic measurable functions $f^{\mathrm{C}}, f^{\mathrm{P}}\colon \mathbb{R}_+ \to \mathbb{R}^{\mathbb{N}\times\mathbb{N}}$, and ρ the sum of two predictable processes $\rho^{\mathrm{C}}, \rho^{\mathrm{P}}\colon \Omega \times \mathbb{R}_+ \to \mathbb{R}^{\mathbb{N}\times\mathbb{N}}$.

Of course, the stochastic integrals behind (5.2.2) must make sense: each single integral must be well defined *and* the infinite sums must converge in an appropriate sense. We do not go into details at this point, which can be found in the general reference [28], but only point out that a sufficient condition for the existence of the infinite-dimensional integral is the existence of the one-dimensional ones plus the summability of their L^2-norms.

Next, we shall explain the rationale behind the IPS model (5.2.2) and the specific choices for the involved processes. By the definition given in (5.2.1), the processes $(X_i\colon i \in \mathbb{N})'$ are coupled in two ways in general: first, they interact internally with each other through a drift term (determined by a) and a volatility term (determined by σ in conjunction with L); and second, they are exposed to the same external

forces (given by B and M), where f and ρ determine the level of influence these noises have on the particles. In particular, by tuning the parameters a, σ, f and ρ appropriately, one obtains a large range of possible dependence structures for the model (5.2.2).

The question this paper aims to attack is how and to which degree the complexity of the high-dimensional IPS (5.2.2) can be reduced. Of course, if all entries of the matrices a, σ, f and ρ are zero or large, there is no hope in simplifying the model. Therefore, our focus lies on particle networks, where only a small number of pairs have strong interaction, while the majority of links in the system are relatively weak. This is implemented in the decomposition of a, σ, f and ρ into a *core matrix* (superscript C) and a *periphery matrix* part (superscript P). It is important to notice that our distinction between core and periphery is not made on the basis of the particles, but on the linkages between them. This allows for greater modelling flexibility since it includes multi-tier networks in our analysis.

In the presence of non-negligible pair interactions it is natural to apply the mean field limit only to the links encoded by the periphery matrices. Therefore, we propose the following *partial mean field system* (PMFS) as an approximation to the IPS (5.2.2):

$$\begin{aligned}\mathrm{d}\bar{X}(t) &= \Big(a^{\mathrm{C}}(t)\bar{X}(t) + a^{\mathrm{P}}(t)\mathbb{E}[\bar{X}(t)]\Big)\,\mathrm{d}t + \Big(\sigma^{\mathrm{C}}(t)\bar{X}(t-) + \sigma^{\mathrm{P}}(t)\mathbb{E}[\bar{X}(t)]\Big).\mathrm{d}L(t)\\ &\quad + f^{\mathrm{C}}(t)b(t)\,\mathrm{d}t + f^{\mathrm{P}}(t)\mathbb{E}[b(t)]\,\mathrm{d}t + \rho^{\mathrm{C}}(t)\,\mathrm{d}M(t), \quad t \in \mathbb{R}_+,\\ \bar{X}(0) &= X(0). \end{aligned} \tag{5.2.4}$$

Written for each row $i \in \mathbb{N}$, this is equivalent to:

$$\begin{aligned}\mathrm{d}\bar{X}_i(t) &= \sum_{j=1}^{\infty}\Big(a^{\mathrm{C}}_{ij}(t)\bar{X}_j(t) + a^{\mathrm{P}}_{ij}(t)\mathbb{E}[\bar{X}_j(t)]\Big)\,\mathrm{d}t\\ &\quad + \sum_{j=1}^{\infty}\Big(\sigma^{\mathrm{C}}_{ij}(t)\bar{X}_j(t-) + \sigma^{\mathrm{P}}_{ij}(t)\mathbb{E}[\bar{X}_j(t)]\Big)\,\mathrm{d}L_i(t)\\ &\quad + \sum_{j=1}^{\infty}\Big(f^{\mathrm{C}}_{ij}(t)b_j(t) + f^{\mathrm{P}}_{ij}(t)\mathbb{E}[b_j(t)]\Big)\,\mathrm{d}t + \sum_{j=1}^{\infty}\rho^{\mathrm{C}}_{ij}(t)\,\mathrm{d}M_j(t), \quad t \in \mathbb{R}_+,\\ \bar{X}_i(0) &= X_i(0). \end{aligned} \tag{5.2.5}$$

It is clear that a priori there is no reason for (5.2.4) to be a good approximation for (5.2.2). Therefore, in the next section, we will give precise L^2-estimates in terms of the model coefficients for the difference between the IPS and the PMFS. Moreover,

we will determine conditions under which this difference becomes small such that we can indeed speak of a law of large numbers.

5.3 Law of large numbers

The first main result of this paper assesses the distance between the original IPS (5.2.2) and the PMFS (5.2.4). To formulate this we have to introduce some further notation. For $T \in \mathbb{R}_+$ we define

$$
\begin{aligned}
v_a(T) &:= \sup_{i\in\mathbb{N}} \sum_{j=1}^{\infty} A_{ij}(T), \quad v_{a,\mathrm{d}}(T) := \sup_{i\in\mathbb{N}} A^{\mathrm{C}}_{ii}(T), \qquad v_\sigma(T) := \sup_{i\in\mathbb{N}} \sum_{j=1}^{\infty} \Sigma_{ij}(T),\\
v_L &:= \sup_{i\in\mathbb{N}} \|L_i(1)\|_{L^2}, \qquad v_b(T) := \sup_{i\in\mathbb{N}} \sup_{t\in[0,T]} \|b_i(t)\|_{L^2}, \qquad v_X := \sup_{i\in\mathbb{N}} \|X_i(0)\|_{L^2},\\
v_f(T) &:= \sup_{i\in\mathbb{N}} \sup_{t\in[0,T]} \sum_{j=1}^{\infty} (|f^{\mathrm{C}}_{ij}(t)| + |f^{\mathrm{P}}_{ij}(t)|),\\
v_{\rho,M}(T) &:= \sup_{i\in\mathbb{N}} \sup_{t\in[0,T]} \left(\sum_{j,k=1}^{\infty} \left|\mathbb{E}[\rho^{\mathrm{C}}_{ij}(t)\rho^{\mathrm{C}}_{ik}(t)c_{jk}(t)]\right| + \left|\mathbb{E}[\rho^{\mathrm{P}}_{ij}(t)\rho^{\mathrm{P}}_{ik}(t)c_{jk}(t)]\right| \right)^{1/2}, \quad (5.3.1)
\end{aligned}
$$

and introduce the rates

$$
\begin{aligned}
r_1(T) &:= \left| A^{\mathrm{P}}(T)|\mathrm{Cov}[X(0)]|(A^{\mathrm{P}}(T))' \right|^{1/2}_{\mathrm{d}}, \quad r_2(T) := \left| \Sigma^{\mathrm{P}}(T)|\mathrm{Cov}[X(0)]|(\Sigma^{\mathrm{P}}(T))' \right|^{1/2}_{\mathrm{d}},\\
r_3(T) &:= \left| A^{\mathrm{P}}(T)|\mathrm{Cov}[L(1)]|(A^{\mathrm{P}}(T))' \right|^{1/2}_{\mathrm{d}}, \quad r_4(T) := \left| \Sigma^{\mathrm{P}}(T)|\mathrm{Cov}[L(1)]|(\Sigma^{\mathrm{P}}(T))' \right|^{1/2}_{\mathrm{d}},\\
r_5(T) &:= \sup_{t\in[0,T]} \left| f^{\mathrm{P}}(t)\mathrm{Cov}[b(t)](f^{\mathrm{P}}(t))' \right|^{1/2}_{\mathrm{d}}, \quad r_6(T) := \sup_{t\in[0,T]} \left| \mathbb{E}\left[\rho^{\mathrm{P}}(t)c(t)(\rho^{\mathrm{P}}(t))'\right] \right|^{1/2}_{\mathrm{d}},\\
r_7(T) &:= \left| A^{\mathrm{P}}(T)A^{\mathrm{C}}(T)^{\times} \right|_{\infty}, \qquad r_8(T) := \left| \Sigma^{\mathrm{P}}(T)A^{\mathrm{C}}(T)^{\times} \right|_{\infty},\\
r_9(T) &:= \sup_{s,t\in[0,T]} \left| A^{\mathrm{P}}(T)|f^{\mathrm{C}}(s)\mathrm{Cov}[b(s),b(t)](f^{\mathrm{C}}(t))'|(A^{\mathrm{P}}(T))' \right|^{1/2}_{\mathrm{d}},\\
r_{10}(T) &:= \sup_{s,t\in[0,T]} \left| \Sigma^{\mathrm{P}}(T)|f^{\mathrm{C}}(s)\mathrm{Cov}[b(s),b(t)]f^{\mathrm{C}}(t))'|(\Sigma^{\mathrm{P}}(T))' \right|^{1/2}_{\mathrm{d}},\\
r_{11}(T) &:= \sup_{t\in[0,T]} \left| A^{\mathrm{P}}(T)|\mathbb{E}[\rho^{\mathrm{C}}(t)c(t)(\rho^{\mathrm{C}}(t))']|(A^{\mathrm{P}}(T))' \right|^{1/2}_{\mathrm{d}},\\
r_{12}(T) &:= \sup_{t\in[0,T]} \left| \Sigma^{\mathrm{P}}(T)|\mathbb{E}[\rho^{\mathrm{C}}(t)c(t)(\rho^{\mathrm{C}}(t))']|(\Sigma^{\mathrm{P}}(T))' \right|^{1/2}_{\mathrm{d}}. \quad (5.3.2)
\end{aligned}
$$

Theorem 5.3.1. *Fix some $T \in \mathbb{R}_+$ and grant the general model assumptions as given in Section 5.2. Furthermore, assume that each of the numbers in* (5.3.1) *is finite. Then* (5.2.2) *and* (5.2.4) *have a pathwise unique solution X and $\bar{X}$, respectively, and there exist constants $K(T)$ and $K_\iota(T)$, $\iota = 1, \ldots, 12$, which depend on the model coefficients only through the numbers in* (5.3.1)*, such that*

$$\sup_{i\in\mathbb{N}} \left\| (X_i - \bar{X}_i)^*(T) \right\|_{L^2} \leq K(T) \sum_{\iota=1}^{12} K_\iota(T) r_\iota(T). \tag{5.3.3}$$

The proof of Theorem 5.3.1 will be given in Section 5.5. Compared to the homogeneous case of [131], we have to take care of several kinds of heterogeneous dependencies in the system: different weights on the edges, the distinction between core and periphery links, and possibly dependent driving noises. This explains why we have twelve rates in contrast to a single one in (5.1.3).

Remark 5.3.2 Our calculations furnish the following constants in (5.3.3):

$$K(T) := \sqrt{2} \exp((T^{1/2} v_a(T) + 2 v_\sigma(T) v_L)^2 T), \tag{5.3.4}$$

and

$$\begin{aligned}
K_1(T) &:= E(T)T, & K_2(T) &:= 2 v_L E(T) T^{1/2}, \\
K_3(T) &:= \frac{2}{3} E(T) v_\sigma(T) V(T) T^{3/2}, & K_4(T) &:= \sqrt{2} v_L E(T) v_\sigma(T) V(T) T, \\
K_5(T) &:= T, & K_6(T) &:= 2T^{1/2}, \\
K_7(T) &:= \frac{1}{2} E(T) V(T) T^2, & K_8(T) &:= \frac{2}{\sqrt{3}} v_L E(T) V(T) \\
K_9(T) &:= \frac{1}{2} E(T) T^2, & K_{10}(T) &:= \frac{2}{\sqrt{3}} v_L E(T) T^{3/2}, \\
K_{11}(T) &:= \frac{2}{3} E(T) T^{3/2}, & K_{12}(T) &:= \sqrt{2} v_L E(T) T,
\end{aligned}$$

where

$$\begin{aligned}
E(T) &:= \mathrm{e}^{v_{a,\mathrm{d}}(T)}, \\
V(T) &:= \sqrt{2} \mathrm{e}^{(v_a(T) T^{1/2} + 2 v_L v_\sigma(T))^2 T} \left(v_X + v_f(T) v_b(T) T + 2 v_{\rho,M}(T) T^{1/2} \right).
\end{aligned}$$

□

Remark 5.3.3 There are several possibilities to extend Theorem 5.3.1 without substantially new arguments.

(1) It is straightforward to show that Theorem 5.3.1 can be extended to the case where the interaction matrices a and σ are replaced by (still deterministic but possibly history-dependent) linear functionals.

(2) Suppose that $L = \Gamma L^0$ with some matrix $\Gamma \in \mathbb{R}^{\mathbb{N}\times\mathbb{N}}$ and some other Lévy process L^0 with finite variance and mean zero. Furthermore, $\Gamma = \Gamma^{\mathrm{C}}+\Gamma^{\mathrm{P}}$ and accordingly $L^{\mathrm{C}} = \Gamma^{\mathrm{C}}L^0$ and $L^{\mathrm{P}} = \Gamma^{\mathrm{P}}L^0$. What one would like to do when passing to the PMFS (5.2.4) is to replace L there by L^{C}. How does this affect the estimate (5.3.3) in Theorem 5.3.1? A similar analysis as for Theorem 5.3.1 reveals that an extra rate
$$r_{13} := \left|\Gamma^{\mathrm{P}}\mathrm{Cov}[\tilde{L}(1)](\Gamma^{\mathrm{P}})'\right|_{\mathrm{d}}^{1/2}$$
appears with constant $K_{13} := 2v_\sigma(T)V(T)T^{1/2}$.

(3) Two further generalizations are discussed in Remark 5.3.5 and Remark 5.3.7 below. □

It is obvious that the usefulness of Theorem 5.3.1 depends on the sizes of the rates in (5.3.2): only if they are small, the PMFS (5.2.4) is a good approximation to the IPS (5.2.2). Moreover, there are two different views on Theorem 5.3.1: first, if we assume that the underlying network of the IPS is static, it gives an upper bound on the L^2-error when the IPS is approximated by the PMFS; and second, if the interaction network (i.e. a, σ, f and ρ) is assumed to evolve according to an index $N \in \mathbb{N}$, Theorem 5.3.1 gives conditions under which the PMFS converges in the L^2-sense to the IPS when $N \to \infty$ (this happens precisely when all rates in (5.3.2) converge to 0 as $N \to \infty$, and the numbers in (5.3.1) are majorized independently of N). It is also this second point of view that is the traditional one in mean field analysis and that justifies the title "Law of large numbers" for the current section.

In the following subsections we will study three examples of dynamical networks and the corresponding conditions for the law of large numbers to hold for the PMFS.

5.3.1 Propagation of chaos

We first discuss the phenomenon of chaos propagation, and our results will particularly extend the results of [66, Sect. 17.3], [87, Cor. 4.1] and [131, Thm. 1.4] by

including inhomogeneous weights in the model. The setting is as follows:

(1) The underlying network changes with $N \in \mathbb{N}$. In particular, we will index X and $\bar{X}$, the coefficients a, σ, f and ρ as well as the rates in (5.3.2) by N.

(2) All structural assumptions in Section 5.2 hold and the numbers in (5.3.1), some of which now depend on N, are uniformly bounded in N.

(3) The core matrices $a^{N,\mathrm{C}}(t)$, $\sigma^{N,\mathrm{C}}(t)$, $f^{N,\mathrm{C}}(t)$ and $\rho^{N,\mathrm{C}}(t)$ are diagonal matrices for all times $t \in \mathbb{R}_+$.

(4) For each $N \in \mathbb{N}$, $(L_i, b_i, M_i, \rho_{ii}^{N,\mathrm{C}}, X_i^N(0) \colon i \in \mathbb{N})$ is a sequence of independent random elements (note that the noises indexed by a fixed i may depend on each other).

(5) For each $T \in \mathbb{R}_+$ the following rates converge to 0 as $N \to \infty$:

$$r_a^N(T) := \sup_{i\in\mathbb{N}} \left(\sum_{j=1}^{\infty} (A_{ij}^{N,\mathrm{P}}(T))^2 \right)^{1/2}, \quad r_\sigma^N(T) := \sup_{i\in\mathbb{N}} \left(\sum_{j=1}^{\infty} (\Sigma_{ij}^{N,\mathrm{P}}(T))^2 \right)^{1/2},$$
$$r_f^N(T) := \sup_{i\in\mathbb{N}} \sup_{t\in[0,T]} \left(\sum_{j=1}^{\infty} (f_{ij}^{N,\mathrm{P}}(t))^2 \right)^{1/2},$$
$$r_{\rho,M}^N(T) := \sup_{i\in\mathbb{N}} \sup_{t\in[0,T]} \left(\sum_{j=1}^{\infty} \mathbb{E}[(\rho_{ij}^{N,\mathrm{P}}(t))^2 c_{jj}(t)] \right)^{1/2}.$$

These hypotheses ensure that all pair dependencies between the processes X_i^N, $i \in \mathbb{N}$, vanish when $N \to \infty$. As a result, in the PMFS, the independence of the particles i at $t = 0$ propagates through all times $t > 0$: the PMFS decouples in contrast to the original IPS.

Example 5.3.4 In classical mean field theory as in the references mentioned in the introduction, the N-th network consists of exactly N particles. In other words, a_{ij}^N, σ_{ij}^N, f_{ij}^N, ρ_{ij}^N and $X_i^N(0)$ are all 0 for $i > N$ or $j > N$. Moreover, all pair interaction is assumed to be of order $1/N$, that is, we have for each $T \in \mathbb{R}_+$

$$A_{ij}^{N,\mathrm{P}}(T) = \frac{A_{ij}(T)}{N}, \quad \Sigma_{ij}^{N,\mathrm{P}}(T) = \frac{\Sigma_{ij}(T)}{N}, \quad i,j \in \mathbb{N}, \tag{5.3.5}$$

where $A_{ij}(T), \Sigma_{ij}(T) \in \mathbb{R}_+$ are uniformly bounded in $i, j \in \mathbb{N}$. Furthermore, the driving noises are supposed to be independent for different particles and to enter the PMFS completely. This means that (3) and (4) hold and that $f^{N,\mathrm{P}} = \rho^{N,\mathrm{P}} = 0$. It is easily shown that under these specifications the rates in (5) above converge to 0 as $N \to \infty$: $r^N_{\rho,M}(T)$ and $r^N_f(T)$ are simply 0, and $r^N_a(T)$ and $r^N_\sigma(T)$ are of order $1/\sqrt{N}$ as $N \to \infty$. □

We still need to show that under assumptions (1)–(5) above, all twelve rates $r^N_\iota(T)$, $\iota = 1, \ldots, 12$, converge to 0 as $N \to \infty$. Since $A^{N,\mathrm{C}}(T)$ is diagonal, we have $A^{N,\mathrm{C}}(T)^\times = 0$, and since the driving noises for different particles are independent, all covariances (or covariations) vanish outside the diagonal. Thus, we have

$$
\begin{array}{lll}
r^N_1(T) \le v_X r^N_a(T), & r^N_2(T) \le v_X r^N_\sigma(T), & r^N_3(T) \le v_L r^N_a(T), \\
r^N_4(T) \le v_L r^N_\sigma(T), & r^N_5(T) \le v_b(T) r^N_f(T), & r^N_6(T) = r^N_{\rho,M}(T), \\
r^N_7(T) = 0, & r^N_8(T) = 0, & r^N_9(T) \le v_b(T) v_f(T) r^N_a(T), \\
r^N_{10}(T) \le v_b(T) v_f(T) r^N_\sigma(T), & r^N_{11}(T) \le v_{\rho,M}(T) r^N_a(T), & r^N_{12}(T) \le v_{\rho,M}(T) r^N_\sigma(T),
\end{array}
$$

which all converge to 0 as $N \to \infty$ by hypothesis. The following remark continues Remark 5.3.3 regarding further generalizations of Theorem 5.3.1.

Remark 5.3.5 In the setting of this subsection there are actually no core relationships between different particles: every pair interaction rate tends to 0 with large N. If we even assume that there is no dependence at all originating from the noises (i.e. $f^{N,\mathrm{P}} = \rho^{N,\mathrm{P}} = 0$ above), the propagation of chaos result can easily be extended to nonlinear Lipschitz interaction terms (suitably bounded in N) instead of the matrices a^N and σ^N. As a matter of fact, the classical method of [131, Thm. 1.4] can be applied with obvious changes. □

5.3.2 Sparse interaction versus sparse correlation

The propagation of chaos result in the last subsection was based on two core hypotheses: asymptotically vanishing pair interaction rates and the independence of the particles' driving noises. The motivation for establishing Theorem 5.3.1, however, is to deal with situations where these two conditions are precisely not satisfied, that is, when the coefficients a, σ, f and ρ of (5.2.2) are decomposed into a core and a periphery part in a non-trivial way. In fact, in this subsection we discuss a typical situation where the full generality of Theorem 5.3.1 is required. Before that,

we recall that we consider networks indexed by $N \in \mathbb{N}$, and that we are interested in the cases when the rates in (5.3.2) vanish when N becomes large.

General assumptions

The following list of hypotheses describes the setting in this subsection.

(1) The statements (1) and (2) of Section 5.3.1 hold.

(2) M is an $\mathbb{R}^{\mathbb{N}}$-valued $\mathbb{F}$-Lévy process, implying that $c_{ij}(t) = \mathrm{Cov}[M_i(1), M_j(1)]t$.

(3) At stage N, the system consists of $N_0 + N$ particles with some fixed $N_0 \in \mathbb{N}$, that is, we have $a^N_{ij} = \sigma^N_{ij} = f^N_{ij} = \rho^N_{ij} = X^N_i(0) = 0$ as soon as $i > N_0 + N$ or $j > N_0 + N$.

(4) $\mathcal{C} := \{1, \dots, N_0\}$ contains the *core particles*, $\mathcal{P}^N := \{N_0 + 1, \dots, N\}$ the *periphery particles*, whose number increases with N. Correspondingly, $a^{N,\mathrm{C}}$ and $\sigma^{N,\mathrm{C}}$ (resp. $a^{N,\mathrm{P}}$ and $\sigma^{N,\mathrm{P}}$) characterize the influence of the core (resp. periphery) particles in the system. In other words, $j \in \mathcal{C}$ implies that $a^{N,\mathrm{P}}_{ij}(t) = \sigma^{N,\mathrm{P}}_{ij}(t) = 0$ for all $i \in \mathbb{N}$ and $t \in \mathbb{R}_+$, while $j \in \mathcal{P}^N$ implies $a^{N,\mathrm{C}}_{ij}(t) = \sigma^{N,\mathrm{C}}_{ij}(t) = 0$ for all $i \neq j$ and $t \in \mathbb{R}_+$. We assume that the diagonals of a^N and σ^N are completely contained in $a^{N,\mathrm{C}}$ and $\sigma^{N,\mathrm{C}}$, respectively. It follows that the partitions of a^N and σ^N can be illustrated as (omitting all zero rows and columns, and using $*$ for all potentially non-zero elements):

$$\left(\begin{array}{ccc|cccc} * & \cdots & * & 0 & \cdots & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \ddots & \vdots \\ * & \cdots & * & 0 & \cdots & \cdots & 0 \\ \hline * & \cdots & * & * & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \vdots & \vdots & \ddots & \ddots & 0 \\ * & \cdots & * & 0 & \cdots & 0 & * \end{array}\right), \qquad \left(\begin{array}{ccc|cccc} 0 & \cdots & 0 & * & \cdots & \cdots & * \\ \vdots & \ddots & \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & * & \cdots & \cdots & * \\ \hline 0 & \cdots & 0 & 0 & * & \cdots & * \\ \vdots & \ddots & \vdots & * & \ddots & \ddots & \vdots \\ \vdots & \ddots & \vdots & \vdots & \ddots & \ddots & * \\ 0 & \cdots & 0 & * & \cdots & * & 0 \end{array}\right).$$

N_0 and N mark the block boundaries; left: $a^{N,\mathrm{C}}$ and $\sigma^{N,\mathrm{C}}$, right: $a^{N,\mathrm{P}}$ and $\sigma^{N,\mathrm{P}}$.

(5) There is a finite number of *systematic noises*, namely $B_1, \dots B_{N_{00}}$ and $M_1, \dots, M_{N_{00}}$ for some fixed $N_{00} \in \mathbb{N}$ independent of N, that are important to a large part of the system, and there are *idiosyncratic noises* $B_{N_{00}+i}$

and $M_{N_{00}+i}$ that only affect the specific particle $i \in \{1,\dots,N\}$. Thus, we assume for all $i = 1,\dots,N$ and $t \in \mathbb{R}_+$ that $\rho_{ij}^{N,\mathrm{P}}(t) = f_{ij}^{N,\mathrm{P}}(t) = 0$ for $j \in \{1,\dots,N_{00}\} \cup \{N_{00}+i\}$ and $\rho_{ij}^{N,\mathrm{C}}(t) = f_{ij}^{N,\mathrm{C}}(t) = 0$ for the other values of j. Hence, f^N and ρ^N are of the form

$$\underset{f^{N,\mathrm{C}} \text{ and } \rho^{N,\mathrm{C}}}{\left(\begin{array}{ccc|cccc} & & & \multicolumn{4}{c}{\overbrace{}^{N_{00}+N}} \\ * & \cdots & * & * & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & 0 & \ddots & \ddots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \ddots & 0 \\ * & \cdots & * & 0 & \cdots & 0 & * \end{array}\right)}, \qquad \underset{f^{N,\mathrm{P}} \text{ and } \rho^{N,\mathrm{P}}}{\left(\begin{array}{ccc|cccc} & & & \multicolumn{4}{c}{\overbrace{}^{N_{00}+N}} \\ 0 & \cdots & 0 & 0 & * & \cdots & * \\ \vdots & \ddots & \vdots & * & \ddots & \ddots & \vdots \\ \vdots & \ddots & \vdots & \vdots & \ddots & \ddots & * \\ 0 & \cdots & 0 & * & \cdots & * & 0 \end{array}\right)}.$$

(6) We have for all $T \in \mathbb{R}_+$

$$A_{ij}^{N,\mathrm{P}}(T) = \frac{\phi_{ij}^N(T)}{R_A^N}, \quad \Sigma_{ij}^{N,\mathrm{P}}(T) = \frac{\psi_{ij}^N(T)}{R_\Sigma^N}, \quad i,j = 1,\dots,N, \tag{5.3.6}$$

where the rates $R_A^N, R_\Sigma^N \in \mathbb{R}_+$ satisfy

$$\frac{R_A^N}{\sqrt{N}} \to \infty, \quad \frac{R_\Sigma^N}{\sqrt{N}} \to \infty, \quad \text{as } N \to \infty, \tag{5.3.7}$$

and the numbers $\phi_{ij}^N(T), \psi_{ij}^N(T) \in \mathbb{R}_+$ satisfy

$$\phi(T) := \sup_{i,j,N \in \mathbb{N}} \phi_{ij}^N(T) < \infty, \quad \psi(T) := \sup_{i,j,N \in \mathbb{N}} \psi_{ij}^N(T) < \infty.$$

Note that we always have $\phi_{ii}^N(T) = \psi_{ii}^N(T) = 0$.

(7) For different $i, j \in \mathbb{N}$, the noises M_i and M_j as well as B_i and B_j are uncorrelated.

(8) The rates $r_f^N(T)$ and $r_{\rho,M}^N(T)$ from Section 5.3.1 converge to 0 as $N \to \infty$ for all $T \in \mathbb{R}_+$.

(9) For each $N \in \mathbb{N}$, the initial values $(X_i^N(0) \colon i \in \mathcal{P}^N)$ are mutually uncorrelated.

gets large, the influence of the core particles is non-negligible. On the contrary, there is a large number of so-called *periphery particles*, which scarcely interact with each other and are connected to the core particles with only a small amount of linkages. The number of periphery particles grows with the system and their individual influence on the system is asymptotically vanishing. Similarly, we distinguish

between two types of noises within the components of B^N and M^N: *systematic noises*, which may have a strong impact on a large number of particles even when N increases, and *idiosyncratic noises*, which mainly affect single particles and whose effect on the other particles diminish when N grows.

Conditions (4) and (5) determine the core–periphery structure of the IPS. In practice, a fixed distinction between core and periphery particles is often not possible because a large number of particles may be engaged in some strong and some weak linkages at the same time. As already pointed out, this does not affect the applicability of Theorem 5.3.1, since the concept of core and periphery refers to the linkages there. The choice of fixed core and periphery *particles* in this subsection is only a special case thereof, intended to simplify the arguments below. Next, regarding (6), one can take $R_A^N, R_\Sigma^N = N$ for concreteness, which can then be compared with Section 5.3.1. Furthermore, let us point out that assumption (7) is only for convenience (namely that f^N and ρ^N carry the whole correlation structure of the noises). Indeed, it is always possible (under our second-moment conditions) to replace any stochastic integral $\rho \cdot M$, where M is a Lévy process, with an arbitrary correlation structure by $\rho' \cdot M'$ where M' consists of mutually uncorrelated Lévy processes (of course, (8) would change accordingly). Finally, if $X^N(0)$ is independent of the driving noises, (9) can be enforced simply by switching to the conditional distribution given $X^N(0)$.

Under (1)–(9) it is easy to prove that the rates $r_1^N(T)$, $r_2^N(T)$, $r_5^N(T)$ and $r_6^N(T)$ converge to 0 when $N \to \infty$. For the latter two, this can be deduced in the same way as in Section 5.3.1 because the driving noises of different particles are uncorrelated. For the other two, we use that the starting random variables of periphery particles are assumed to be uncorrelated. Hence, we have by (5.3.7), as $N \to \infty$, that

$$r_1^N(T) = \sup_{i\in\mathbb{N}} \left(\sum_{j\in\mathcal{P}^N} (A_{ij}^{N,\mathrm{P}}(T))^2 \mathrm{Var}[X_j^N(0)] \right)^{1/2} \leq \phi(T) v_X \frac{\sqrt{N}}{R_A^N} \to 0,$$

$$r_2^N(T) = \sup_{i\in\mathbb{N}} \left(\sum_{j\in\mathcal{P}^N} (\Sigma_{ij}^{N,\mathrm{P}}(T))^2 \mathrm{Var}[X_j^N(0)] \right)^{1/2} \leq \psi(T) v_X \frac{\sqrt{N}}{R_\Sigma^N} \to 0.$$

However, the nine conditions above are in general *not* sufficient to imply the smallness of the other rates in (5.3.2). We need to add extra hypotheses.

Sparseness assumptions

For each of the remaining rates, we further examine what type of conditions are needed to make them asymptotically small. As we shall see, it is always a mixture of a sparseness condition on the interaction matrices A^N and Σ^N and a sparseness condition on the correlation matrices f^N and ρ^N.

$r_3^N(T)$ and $r_4^N(T)$: We first present a counterexample to show that we have to require further conditions. Consider the simple case where $L_i = L_1$ for all $i \in \mathbb{N}$ and that $A_{ij}^{N,\mathrm{P}}(T) = 1/R_A^N$ for all $T \in \mathbb{R}_+$ and $i, j \in \{1, \dots, N_0 + N\}$ with $i \neq j$. Then

$$r_3^N(T) = \sup_{i\in\mathbb{N}} \left(\sum_{j,k\in\mathcal{P}^N\setminus\{i\}} \left(\frac{1}{\rho_A^N}\right)^2 \mathrm{Cov}[L_j(1), L_k(1)] \right)^{1/2} = v_L \frac{N}{R_A^N},$$

which need not to converge to 0 in general. A similar calculation can be done for $r_4^N(T)$. In order to make the rates $r_3^N(T)$ and $r_4^N(T)$ small, there are basically two options: we require the interaction matrices $A^{N,\mathrm{P}}$ and $\Sigma^{N,\mathrm{P}}$ to be sparse, or we require the correlation matrix of L to be sparse. Any other possibility is a suitable combination of these two.

(10a) The noises $(L_i \colon i \in \mathcal{P}^N)$ corresponding to periphery particles only have sparse correlation (which, in particular, includes the case of mutual independence as in Section 5.3.1). More precisely, we require

$$\begin{aligned} p_L^N &:= \#\{(i,j) \in \mathcal{P}^N \times \mathcal{P}^N \colon \mathrm{Cov}[L_i(1), L_j(1)] \neq 0\}| \\ &= o\Big((R_A^N)^2 \wedge (R_\Sigma^N)^2\Big) \end{aligned} \tag{5.3.8}$$

for large N. Then

$$\begin{aligned} r_3^N(T) &= \sup_{i\in\mathbb{N}} \left(\sum_{j,k\in\mathcal{P}^N} A_{ij}^{N,\mathrm{P}}(T) A_{ik}^{N,\mathrm{P}}(T) \mathrm{Cov}[L_j(1), L_k(1)] \right)^{1/2} \\ &\leq \phi(T) v_L \frac{\sqrt{p_L^N}}{R_A^N} \to 0, \end{aligned}$$

and, similarly, $r_4^N(T) \to 0$ as $N \to \infty$.

(10b) The matrices $A^{N,\mathrm{P}}(T)$ and $\Sigma^{N,\mathrm{P}}(T)$, which describe the influence of periphery particles on the system, are only sparsely occupied, in the sense that every

particle in the system is only affected by a small number of periphery particles. In mathematical terms this condition reads as

$$p_{A,1}^N(T) := \sup_{i\in\mathbb{N}} \#\{j \in \mathcal{P}^N \colon A_{ij}^{N,\mathrm{P}}(T) \neq 0\} = o(R_A^N),$$
$$p_{\Sigma}^N(T) := \sup_{i\in\mathbb{N}} \#\{j \in \mathcal{P}^N \colon \Sigma_{ij}^{N,\mathrm{P}}(T) \neq 0\} = o(R_\Sigma^N). \tag{5.3.9}$$

In this case, we get

$$r_3^N(T) = \sup_{i\in\mathbb{N}} \left(\sum_{j,k\in\mathcal{P}^N} A_{ij}^{N,\mathrm{P}}(T) A_{ik}^{N,\mathrm{P}}(T) \mathrm{Cov}[L_j(1), L_k(1)] \right)^{1/2} \leq \phi(T) v_L \frac{p_{A,1}^N}{R_A^N} \to 0,$$

and similarly $r_4^N(T) \to 0$ as $N \to \infty$.

$r_7^N(T)$ and $r_8^N(T)$: These two rates express the connectivity between core and periphery particles. In general, they will be not become small with large N. For instance, if $A_{ij}^{N,\mathrm{C}}(T) = 1$ for all $j \in \mathcal{C}$ and $i \neq j$, and $A_{ij}^{N,\mathrm{P}} = 1/R_A^N$ for all $j \in \mathcal{P}^N$ and $i \neq j$, then

$$r_7^N(T) = \sup_{i\in\mathbb{N}} \sum_{j\in\mathcal{P}^N} \sum_{k\in\mathcal{C}} A_{ij}^{N,\mathrm{P}}(T) A_{jk}^{N,\mathrm{C}}(T) = N_0 \frac{N}{R_A^N},$$

does not necessarily converge to 0. An analogous statement holds for $r_8^N(T)$. For $r_7^N(T), r_8^N(T) \to 0$ we have to require that the lower left block of $A^{N,\mathrm{C}}$, which describes the influence of core particles on periphery particles, or the matrices $A^{N,\mathrm{P}}(T)$ and $\Sigma^{N,\mathrm{P}}(T)$, which describe the influence of periphery particles on the system, be sparse (or a combination thereof):

(11a) The influence of core on periphery particles is sparse. In other words, we suppose for the maximal number of periphery particles a single core particle interacts with through the drift:

$$p_{A,2}^N := \sup_{j\in\mathcal{C}} \#\{i \in \mathcal{P}^N \colon A_{ij}^{N,\mathrm{C}}(T) \neq 0\} = o(R_A^N \wedge R_\Sigma^N). \tag{5.3.10}$$

Then

$$r_7^N(T) = \sup_{i\in\mathbb{N}} \sum_{j\in\mathcal{P}^N} \sum_{k\in\mathcal{C}} A_{ij}^{N,\mathrm{P}}(T) A_{jk}^{N,\mathrm{C}}(T) \leq N_0 \phi(T) v_a(T) \frac{p_{A,2}^N}{R_A^N} \to 0$$

as well as $r_8^N(T) \to 0$ as $N \to \infty$.

(11b) $A^{N,\mathrm{P}}(T)$ and $\Sigma^{N,\mathrm{P}}(T)$ are sparse in the sense of (5.3.9). Then $r_7^N(T), r_8^N(T) \to 0$ follow similarly.

$r_9^N(T)$, $r_{10}^N(T)$, $r_{11}^N(T)$ and $r_{12}^N(T)$: Similar considerations as before show that these four rates do not converge to 0 in general. Instead, we again need to require some mixture of sparsely correlated driving noises and sparsely occupied matrices $A^{N,\mathrm{P}}$ and $\Sigma^{N,\mathrm{P}}$:

(12a) We assume that for all $T \in \mathbb{R}_+$

$$p_f^N(T) := \sup_{j\in\{1,\dots,N_{00}\}} \#\{i \in \mathcal{P}^N \colon f_{ij}^{N,\mathrm{C}} \not\equiv 0 \text{ on } [0,T]\} = o(R_A^N \wedge R_\Sigma^N), \quad (5.3.11)$$

$$p_\rho^N(T) := \sup_{j\in\{1,\dots,N_{00}\}} \#\{i \in \mathcal{P}^N \colon \rho_{ij}^{N,\mathrm{C}} \not\equiv 0 \text{ on } [0,T]\} = o(R_A^N \wedge R_\Sigma^N). \quad (5.3.12)$$

Then, recalling that the components of b and M are mutually uncorrelated,

$$\begin{aligned} r_9^N(T) &= \sup_{i\in\mathbb{N}} \sup_{s,t\in[0,T]} \Bigg(\sum_{j,k\in\mathcal{P}^N} \sum_{l=1}^{N_{00}} A_{ij}^{N,\mathrm{P}}(T) A_{ik}^{N,\mathrm{P}}(T) \left| f_{jl}^{N,\mathrm{C}}(s) f_{kl}^{N,\mathrm{C}}(t) \mathrm{Cov}[b_l(s), b_l(t)] \right| \\ &\qquad + \sum_{j\in\mathcal{P}^N} (A_{ij}^{N,\mathrm{P}}(T))^2 \left| f_{j(N_{00}+i)}^{N,\mathrm{C}}(s) f_{j(N_{00}+i)}^{N,\mathrm{C}}(t) \mathrm{Cov}[b_{N_{00}+i}(s), b_{N_{00}+i}(t)] \right| \Bigg)^{1/2} \\ &\le \phi(T) v_b(T) v_f(T) \frac{\sqrt{N_{00}} p_f^N(T) + \sqrt{N}}{R_A^N} \to 0, \\ r_{11}^N(T) &= \sup_{i\in\mathbb{N}} \sup_{t\in[0,T]} \Bigg(\sum_{j,k\in\mathcal{P}^N} \sum_{l=1}^{N_{00}} A_{ij}^{N,\mathrm{P}}(T) A_{ik}^{N,\mathrm{P}}(T) \left| \mathbb{E}[\rho_{jl}^{N,\mathrm{C}}(t) \rho_{kl}^{N,\mathrm{C}}(t)] \right| \mathrm{Var}[M_l(1)] \\ &\qquad + \sum_{j\in\mathcal{P}^N} (A_{ij}^{N,\mathrm{P}}(T))^2 \left| \mathbb{E}[(\rho_{j(N_{00}+i)}^{N,\mathrm{C}}(t))^2] \right| \mathrm{Var}[M_{N_{00}+i}(1)] \Bigg)^{1/2} \\ &\le \phi(T) v_{\rho,M}(T) \frac{\sqrt{N_{00}} p_\rho^N(T) + \sqrt{N}}{R_A^N} \to 0, \end{aligned}$$

and similarly $r_{10}^N(T), r_{12}^N(T) \to 0$ as $N \to \infty$.

(12b) $A^{N,2}$ and $\Sigma^{N,2}$ are sparse in the sense of (5.3.9). Then one can deduce $r_\iota^N(T) \to 0$ for $\iota = 9, 10, 11, 12$ as before.

We conclude this subsection with two remarks.

Remark 5.3.6 In the sparseness conditions (5.3.8)–(5.3.12) it is not essential that the majority of entries is exactly zero. As one can see from the definition of the rates (5.3.2), they depend continuously on the underlying matrix entries. It suffices therefore that the matrix entries are small enough in a large proportion. □

Remark 5.3.7 What can be said about Theorem 5.3.1 in the general case of nonlinear Lipschitz coefficients a^N and σ^N, apart from the special case discussed in Remark 5.3.5? In fact, a law of large numbers in the fashion of Theorem 5.3.1 can still be shown, but under more stringent conditions: namely we have to require condition (10b) above in addition, with $A^{N,\mathrm{P}}$ and $\Sigma^{N,\mathrm{P}}$ now containing the Lipschitz constants of the interaction terms. The reason is that (10b) suffices to make r_ι^N ($\iota \in \{3, 4, 7, \ldots, 12\}$) small. The remaining four rates are unrelated to a^N and σ^N and therefore not affected by their nonlinear structure. It is important to notice that conditions like (10a) and (12a) are no longer sufficient to make the corresponding rates small. The reason is that they are conditions of correlation type. Since correlation is a *linear* measure of dependence, it is not surprising that these conditions are not suitable for the nonlinear case. We do not go into the details here. □

5.3.3 Networks arising from preferential attachment

As demonstrated in the last subsection, the crucial criterion for the rates (5.3.2) in Theorem 5.3.1 to vanish asymptotically with growing network size can be described as a combination of sparse interaction and sparse correlation among the particles. Condition (5.3.9) plays a distinguished role here: when valid, it implies that eight out of twelve rates in (5.3.2) are small. Moreover, it is the key factor for a nonlinear generalization of Theorem 5.3.1 to hold or not; see Remark 5.3.7. The aim of this subsection is therefore to find algorithms for the generation of the underlying networks such that the resulting interaction matrices satisfy (5.3.9). We will assume that $a^{N,\mathrm{P}}(t) = a^{N,\mathrm{P}}$ and $\sigma^{N,\mathrm{P}}(t) = \sigma^{N,\mathrm{P}}$ are independent of $t \in \mathbb{R}_+$, such that also $A^{N,\mathrm{P}}(t)$ and $\Sigma^{N,\mathrm{P}}(t)$ as well as $p_{A,1}^N(t)$ and $p_\Sigma^N(t)$ (see (5.3.9) for their definitions) are independent of t. Furthermore, we only concentrate on $p_{A,1}^N$ as the analysis for p_Σ^N is completely analogous.

We will base the creation of the IPS network on dynamical random graph mechanisms. Since we are mainly interested in heterogeneous graphs, we will investigate the *preferential attachment* or *scale-free* random graph. Popularized by [10] as a model for the world wide web network, it is by far the most successful random

graph model in applications besides the classical homogeneous Erdős-Rényi model. There are many similar but different constructions of preferential attachment graphs; in the following, we rely on the construction of [32] for directed graphs. We remark that the random graphs to be constructed will be indexed by N, corresponding to a family of growing networks for the IPS. In particular, "time" in the random graph process must not be confused with the time t in the IPS (5.2.2); the correct view is rather that the IPS network has been built from the random graphs before time $t = 0$, and, of course, independently of all random variables in (5.2.2).

The preferential attachment algorithm works as follows: we start with a given graph $G(0) = (V, E(0))$ consisting of vertices $V = \mathbb{N}$ and edges $E(0) = \{e_1, \dots, e_\nu\}$, where $\nu \in \mathbb{N}$ and e_i stands for a directed edge between two vertices. We allow for multiple edges and loops in our graphs. Without loss of generality, we assume that the set of vertices in $G(0)$ with at least one neighbour given by $\{1, \dots, n(0)\}$ with some $n(0) \in \mathbb{N}$. Furthermore, we fix $\alpha, \beta, \gamma \in \mathbb{R}_+$ with $\alpha + \beta + \gamma = 1$ and $\alpha + \gamma > 0$ and two numbers $\delta^{\text{in}}, \delta^{\text{out}} \in \mathbb{R}_+$. For $N \in \mathbb{N}$ we construct $G(N) = (V, E(N))$ from $G(N-1)$ according to the following algorithm.

- With probability α, we create a new edge $e_{\nu+N}$ from $v = n(N-1)+1$ to a node w that is already connected in $G(N-1)$. Here w is chosen randomly from $\{1, \dots, n(N-1)\}$ according to the probability mass function

$$\frac{\mathrm{d}^{\text{in}}_{G(N-1)}(w) + \delta^{\text{in}}}{\nu + N - 1 + \delta^{\text{in}} n(N-1)}, \quad w \in \{1, \dots, n(N-1)\},$$

  where $\mathrm{d}^{\text{in}}_G(v)$ denotes the in-degree of vertex v in a graph G. Moreover, we define $n(N) := n(N-1) + 1$ and $E(N) := E(N-1) \cup \{e_{\nu+N}\}$.

- With probability β, a new edge $e_{\nu+N}$ from some vertex $v \in \{1, \dots, n(N-1)\}$ to some $w \in \{1, \dots, n(N-1)\}$ is formed (the case $v = w$ is possible). Here v and w are chosen independently according to the probability mass functions

$$\frac{\mathrm{d}^{\text{out}}_{G(N-1)}(v) + \delta^{\text{out}}}{\nu + N - 1 + \delta^{\text{out}} n(N-1)} \quad \text{and} \quad \frac{\mathrm{d}^{\text{in}}_{G(N-1)}(w) + \delta^{\text{in}}}{\nu + N - 1 + \delta^{\text{in}} n(N-1)}$$

  with $v, w \in \{1, \dots, n(N-1)\}$, respectively, where $\mathrm{d}^{\text{out}}_G(v)$ denotes the out-degree of vertex v in a graph G. Moreover, we set $n(N) := n(N-1)$ and $E(N) := E(N-1) \cup \{e_{\nu+N}\}$.

- With probability γ, a new edge $e_{\nu+N}$ from some $v \in \{1, \dots, n(N-1)\}$ to $w = n(N-1)+1$ is formed. Here v is chosen randomly according to the probability mass function

$$\frac{\mathrm{d}^{\mathrm{out}}_{G(N-1)}(v) + \delta^{\mathrm{out}}}{\nu + N - 1 + \delta^{\mathrm{out}} n(N-1)}, \quad v \in \{1, \dots, n(N-1)\}.$$

  We set $n(N) := n(N-1)+1$ and $E(N) := E(N-1) \cup \{e_{\nu+N}\}$.

Evidently, we always have $|E(N)| = \nu + N$ while the number $n(N)$ of non-isolated vertices in $G(N)$ is random in general.

The most important result for our purposes is the following one. We define

$$M^{\mathrm{in}}(N) := \max\{d^{\mathrm{in}}_{G(N)}(i) \colon i \in \mathbb{N}\}, \quad M^{\mathrm{out}}(N) := \max\{d^{\mathrm{out}}_{G(N)}(i) \colon i \in \mathbb{N}\}, \quad N \in \mathbb{N}_0,$$

as the maximal in-degree and out-degree in $G(N)$, respectively.

Lemma 5.3.8. *The maximum in-degree $M^{\mathrm{in}}(N)$ and out-degree $M^{\mathrm{out}}(N)$ of $G(N)$ satisfy the following asymptotics:*

$$c^{\mathrm{in}}(N) M^{\mathrm{in}}(N) \to \mu^{\mathrm{in}}, \quad c^{\mathrm{out}}(N) M^{\mathrm{in}}(N) \to \mu^{\mathrm{out}}, \quad N \to \infty. \tag{5.3.13}$$

Here the convergence to the random variables μ^{in} and μ^{out}, respectively, holds in the almost sure as well as in the L^p-sense for all $p \in [1, \infty)$, and $(c^{\mathrm{in}}(N))_{N\in\mathbb{N}}$ and $(c^{\mathrm{out}}(N))_{N\in\mathbb{N}}$ are sequences of random variables which can be chosen such that for every $\epsilon \in (0, \alpha+\gamma)$ we have a.s., as $N \to \infty$, that

$$c^{\mathrm{in}}(N)^{-1} = \mathcal{O}\left(N^{\frac{\alpha+\beta}{1+\delta^{\mathrm{in}}(\alpha+\gamma-\epsilon)}}\right), \quad c^{\mathrm{out}}(N)^{-1} = \mathcal{O}\left(N^{\frac{\beta+\gamma}{1+\delta^{\mathrm{out}}(\alpha+\gamma-\epsilon)}}\right). \tag{5.3.14}$$

It follows from this lemma that for every $\epsilon \in (0, \alpha+\gamma)$ we have a.s.

$$M^{\mathrm{in}}(N) = \mathcal{O}\left(N^{\frac{\alpha+\beta}{1+\delta^{\mathrm{in}}(\alpha+\gamma-\epsilon)}}\right), \quad M^{\mathrm{out}}(N) = \mathcal{O}\left(N^{\frac{\beta+\gamma}{1+\delta^{\mathrm{out}}(\alpha+\gamma-\epsilon)}}\right), \quad N \to \infty.$$

In particular, if $G(N)$ is used to model the underlying network of a^N (i.e. an edge in $G(N)$ from i to j is equivalent to $a^N_{ij} \neq 0$), we have

$$p^N_{A,1} \leq M^{\mathrm{out}}(N) = \mathcal{O}\left(N^{\frac{\beta+\gamma}{1+\delta^{\mathrm{out}}(\alpha+\gamma-\epsilon)}}\right), \quad N \to \infty.$$

In other words, the first part of condition (5.3.9) holds as soon as R_A^N, as specified through (5.3.6) and (5.3.7), increases in N at least with rate

$$N^{\frac{\beta+\gamma}{1+\delta^{\text{out}}(\alpha+\gamma-\epsilon)}} \tag{5.3.15}$$

for some small ϵ. For example, in the classical case of Example 5.3.4 where $R_A^N = N$, this is always true except in the case $\alpha = \delta^{\text{out}} = 0$, where all edges start from one of the initial nodes with probability one. We conclude that in all non-trivial situations of the preferential attachment model, the resulting networks are sparse enough for the law of large numbers implied by Theorem 5.3.1 to be in force.

5.4 Large deviations

In Theorem 5.3.1 we have established bounds on the mean squared difference between the IPS (5.2.2) and the PMFS (5.2.4). In Sections 5.3.1–5.3.3 we have given examples of dynamical networks in which these bounds converge to 0 as the network size increases. A natural question is now whether a large deviation principle holds as $N \to \infty$, which would then assure that the probability of X^N deviating strongly from $\bar{X}^N$ decreases exponentially fast in N. In the classical case of homogeneous networks, [50] is the first paper to prove a large deviation principle for the empirical measures of the processes (5.1.1). For heterogeneous networks, however, the empirical measure might no longer be a good quantity to investigate: the weight of a particle now depends on which particle's perspective is chosen. A sequence of differently weighted empirical measures seems to be more appropriate, but then their analysis becomes considerably more involved. Therefore, in this paper we take a more direct approach and study the large deviation behaviour of the difference $X^N - \bar{X}^N$ itself. In order to do so, we have to put stronger assumptions on the coefficients than in the previous sections. These are as follows.

(A1) $X^N(0)$ is deterministic for each $N \in \mathbb{N}$.

(A2) For all $N \in \mathbb{N}$ we have $\sigma^N = 0$. All other coefficients $a^{N,\mathrm{C}}$, $a^{N,\mathrm{P}}$, $\rho^{N,\mathrm{C}}$ and $\rho^{N,\mathrm{P}}$ are constant in time. $\rho^{N,\mathrm{C}}$ and $\rho^{N,\mathrm{P}}$ (resp. $a^{N,\mathrm{C}}$ and $a^{N,\mathrm{P}}$) only have $\gamma(N)$ (resp. $\Gamma(N)$) non-zero columns, where $\gamma(N)$ forms a sequence of natural numbers increasing to infinity and $\Gamma(N)$ grows at most like $\exp(\gamma(N))$.

(A3) All numbers in (5.3.1), which are indexed by N now, are bounded independently of N.

(A4) $(M_i \colon i \in \mathbb{N})$ is a sequence of independent mean-zero Lévy processes whose Brownian motion part has variance c_i and whose Lévy measure is ν_i. Moreover, there exists a real-valued mean-zero Lévy process M_0 that dominates M_i, that is, its characteristics c_0 and ν_0 satisfy $c_i \leq c_0$ and $\nu_i(A) \leq \nu_0(A)$ for all $i \in \mathbb{N}$ and Borel sets $A \subseteq \mathbb{R}$, and that has finite exponential moments of all orders: $\mathbb{E}[\mathrm{e}^{uM_0(1)}] < \infty$ for all $u \in \mathbb{R}_+$.

(A5) Assume that $G^N(t,s) := \gamma(N)\mathrm{e}^{a^N t} a^{N,\mathrm{P}} \mathrm{e}^{a^{N,\mathrm{C}} s} \rho^{N,\mathrm{C}}$, $s,t \in [0,T]$, converges uniformly to a limit $G(t,s) \in \mathbb{R}^{\mathbb{N}\times\mathbb{N}}$:

$$\sup_{i,j\in\mathbb{N}} \sup_{s,t\in[0,T]} |G_{ij}(t,s)| < \infty, \quad \sup_{i,j\in\mathbb{N}} \sup_{s,t\in[0,T]} |G^N_{ij}(t,s) - G_{ij}(t,s)| \to 0, \quad N \to \infty.$$

(A6) With $R^N(t) := \gamma(N)\mathrm{e}^{a^N t}\rho^{N,\mathrm{P}}$, $t \in [0,T]$, there exists $R(t) \in \mathbb{R}^{\mathbb{N}\times\mathbb{N}}$ such that

$$\sup_{i,j\in\mathbb{N}} \sup_{t\in[0,T]} |R_{ij}(t)| < \infty, \quad \sup_{i,j\in\mathbb{N}} \sup_{t\in[0,T]} |R^N_{ij}(t) - R_{ij}(t)| \to 0, \quad N \to \infty.$$

(A7) The following two quantities are finite:

$$q_1 := \limsup_{N\to\infty} q_1(N) := \limsup_{N\to\infty} \sup_{i\in\mathbb{N}} \gamma(N) \sum_{j=1}^{\infty} \sum_{k\neq j} |a^{N,\mathrm{P}}_{ij} a^{N,\mathrm{C}}_{jk}|,$$
$$q_2 := \limsup_{N\to\infty} q_2(N) := \limsup_{N\to\infty} \sup_{i,k\in\mathbb{N}} \gamma(N) \sum_{j=1}^{\infty} |a^{N,\mathrm{P}}_{ij} \rho^{N,\mathrm{C}}_{jk}|.$$

(A8) Define for $m \in \mathbb{N} \cup \{0\}$

$$\Psi_m(u) := \frac{1}{2} c_m u^2 + \int_{\mathbb{R}} (\mathrm{e}^{uz} - 1 - uz)\, \nu_m(\mathrm{d}z), \quad u \in \mathbb{R}_+.$$

We assume that the following holds for every $d \in \mathbb{N}$: denoting for $m \in \mathbb{N}$, $r \in [0,T]$ and $\theta \in M_T^d$

$$H_m(\theta, r) := \int_r^T \int_s^T \sum_{i=1}^d G_{im}(t-s, s-r)\, \theta_i(\mathrm{d}t)\, \mathrm{d}s + \int_r^T \sum_{i=1}^d R_{im}(t-r)\, \theta_i(\mathrm{d}t),$$

the sequence $\left(\int_0^T \Psi_m(H_m(\theta,r))\, \mathrm{d}r\right)_{m\in\mathbb{N}}$ is Cesàro summable, i.e. the following limit exists:

$$\lim_{N\to\infty} \frac{1}{\gamma(N)} \sum_{m=1}^{\gamma(N)} \int_0^T \Psi_m(H_m(\theta,r))\, \mathrm{d}r. \tag{5.4.1}$$

Theorem 5.4.1. *Let $T \in \mathbb{R}_+$. Under (A1)–(A8), the sequence $(X^N - \bar{X}^N)_{N\in\mathbb{N}}$ satisfies a large deviation principle in $(D_T^\infty, \mathcal{D}_T^\infty, J_1)$ with a good rate function $I\colon D_T^\infty \to [0,\infty]$, that is, for every $\alpha \in \mathbb{R}_+$ the set $\{x \in D_T^\infty \colon I(x) \le \alpha\}$ is compact in D_T^∞ (with respect to the J_1-topology), and for every $M \in \mathcal{D}_T^\infty$ we have*

$$\begin{aligned} -\inf_{x\in \operatorname{int} M} I(x) &\le \liminf_{N\to\infty} \frac{1}{\gamma(N)} \log \mathbb{P}[X^N - \bar{X}^N \in M] \\ &\le \limsup_{N\to\infty} \frac{1}{\gamma(N)} \log \mathbb{P}[X^N - \bar{X}^N \in M] \le -\inf_{x\in \operatorname{cl} M} I(x), \end{aligned}$$

where $\operatorname{int} M$ *and* $\operatorname{cl} M$ *denote the interior and the closure of M in (D_T^∞, J_1), respectively. Moreover, the rate function I is convex, attains its minimum* 0 *uniquely at the origin and is infinite for $x \notin AC_T^\infty$.*

Remark 5.4.2 (1) We cannot drop the requirement $\sigma^N = 0$ or condition (A4) in Theorem 5.4.1. If violated, the processes X^N and $\bar{X}^N$ will typically not have exponential moments of all order, whose existence is essential for our proof below. This kind of problem does not arise when empirical measures are considered as in [50, 96] for the homogeneous case.

(2) The Cesàro summability condition (A8) accounts for the possible inhomogeneity of the coefficients and the distribution of the noises. It holds in particular for the homogeneous case. Since a convergent series is Cesàro summable with the same limit, it also holds when we have asymptotic homogeneity (in the sense that the sequence inside the sum of (5.4.1) converges with $m \to \infty$).

(3) With $\gamma(N) = N$ and assumptions (A2)–(A4) in force, McKean's example (5.1.1) or the model considered in [87] both satisfy the assumptions of the theorem. In McKean's case our large deviation principle follows from that of [50] for the empirical measure by applying the contraction principle.

5.5 Proofs

We start with some preparatory results that are needed for the proof of Theorem 5.3.1.

Lemma 5.5.1. *Under the assumptions of Theorem 5.3.1 we have*

$$\sup_{i\in\mathbb{N}} \left\|\bar{X}_i^*(T)\right\|_{L^2} \leq V(T),$$

where $V(T)$ is given in Theorem 5.3.1.

Proof. It is a consequence of (5.2.4) and the Burkholder-Davis-Gundy inequality that

$$\begin{aligned}\left\|(\bar{X}_i)^*(t)\right\|_{L^2} \leq \left\|X_i(0)\right\|_{L^2} &+ \int_0^t \sum_{j=1}^{\infty} A_{ij}(T)\left\|(\bar{X}_j)^*(s)\right\|_{L^2} \mathrm{d}s \\ &+ 2\mathrm{Var}[L_i(1)] \left(\int_0^t \left(\sum_{j=1}^{\infty} \Sigma_{ij}(T)\left\|(\bar{X}_j)^*(s)\right\|_{L^2}\right)^2 \mathrm{d}s\right)^{1/2} \\ &+ \int_0^t \sum_{j=1}^{\infty} \left|f_{ij}(s)\right|\left\|b_j(s)\right\|_{L^2} \mathrm{d}s + 2\left(\sum_{j,k=1}^{\infty} \int_0^t \mathbb{E}\left[\rho_{ij}^{\mathrm{C}}(s)\rho_{ik}^{\mathrm{C}}(s)c_{jk}(s)\right] \mathrm{d}s\right)^{1/2}\end{aligned}$$

for all $t \in [0,T]$ and $i \in \mathbb{N}$. Therefore, if we define $w(t) := \sup_{i\in\mathbb{N}} \|(\bar{X}_i)^*(t)\|_{L^2}$, we obtain

$$\begin{aligned}w(t) &\leq v_X + v_f(T)v_b(T)T + 2v_{\rho,M}(T)T^{1/2} + v_a(T)\int_0^t w(s)\,\mathrm{d}s \\ &\quad + 2v_L v_\sigma(T)\left(\int_0^t (w(s))^2\,\mathrm{d}s\right)^{1/2} \\ &\leq v_X + v_f(T)v_b(T)T + 2v_{\rho,M}(T)T^{1/2} \\ &\quad + (v_a(T)T^{1/2} + 2v_L v_\sigma(T))\left(\int_0^t (w(s))^2\,\mathrm{d}s\right)^{1/2}.\end{aligned}$$

Now we square the last inequality, apply the basic estimate $(a+b)^2 \leq 2(a^2+b^2)$ and use Gronwall's inequality to deduce our claim, namely that

$$w(T) \leq \sqrt{2}\mathrm{e}^{(v_a(T)T^{1/2}+2v_L v_\sigma(T))^2 T}\left(v_X + v_f(T)v_b(T)T + 2v_{\rho,M}(T)T^{1/2}\right).$$

□

Lemma 5.5.2. *Assume that $(M_1(t))_{t\in\mathbb{R}_+}$ and $(M_2(t))_{t\in\mathbb{R}_+}$ are square-integrable martingales relative to the same filtration and $(H_1(t))_{t\in\mathbb{R}_+}$ and $(H_2(t))_{t\in\mathbb{R}_+}$ are predictable processes that are respectively integrable with respect to M_1 and M_2 up to some time*

$T \in \mathbb{R}_+$. Furthermore, suppose that $\sup_{t\in[0,T]} \mathbb{E}[H_1^2(t)]$ and $\sup_{t\in[0,T]} \mathbb{E}[H_2^2(t)]$ are finite. Then for all $s,t \in [0,T]$ with $s < t$ we have

$$\begin{aligned}\mathbb{E}\left[\int_0^s H_1(r)\,\mathrm{d}M_1(r)\int_0^t H_2(u)\,\mathrm{d}M_2(u)\right] &= \mathbb{E}\left[\int_0^s H_1(r)\,\mathrm{d}M_1(r)\int_0^s H_2(u)\,\mathrm{d}M_2(u)\right]\\ &= \mathbb{E}\left[\int_0^s H_1(r)H_2(r)\,\mathrm{d}\langle M_1,M_2\rangle(r)\right].\end{aligned}$$

Proof. The martingale property of M_1 and M_2 passes over to the stochastic integral processes. Thus, the second equality is a simple consequence of the classical integration-by-parts formula, see I.3.17 and I.4.45 of [80]. The first equality holds in a more general setting: for any martingales M_1 and M_2 relative to the same filtration, we have

$$\begin{aligned}\mathbb{E}[M_1(s)M_2(t)] &= \mathbb{E}\Big[\mathbb{E}[M_1(s)M_2(t)\mid\mathcal{F}(s)]\Big] = \mathbb{E}\Big[M_1(s)\mathbb{E}[M_2(t)\mid\mathcal{F}(s)]\Big]\\ &= \mathbb{E}[M_1(s)M_2(s)].\end{aligned}$$

□

Lemma 5.5.3. *Let $T \in \mathbb{R}_+$ and assume the finiteness of the numbers* (5.3.1)*. We fix some $j \in \mathbb{N}$ throughout this lemma and define for $t \in [0,T]$*

$$\begin{aligned}Y_j(t) &:= Y_j^1 + Y_j^2(t) + Y_j^3(t) + Y_j^4(t) + Y_j^5(t)\\ &:= (X_j(0) - \mathbb{E}[X_j(0)]) + \sum_{k\neq j}\int_0^t a_{jk}^{\mathrm{C}}(s)(\bar{X}_k(s) - \mathbb{E}[\bar{X}_k(s)])\,\mathrm{d}s\\ &\quad + \sum_{k=1}^{\infty}\int_0^t \left(\sigma_{jk}^{\mathrm{C}}(s)\bar{X}_k(s-) + \sigma_{jk}^{\mathrm{P}}(s)\mathbb{E}[\bar{X}_k(s)]\right)\mathrm{d}L_j(s)\\ &\quad + \sum_{k=1}^{\infty}\int_0^t f_{jk}^{\mathrm{C}}(s)(b_k(s) - \mathbb{E}[b_k(s)])\,\mathrm{d}s + \sum_{k=1}^{\infty}\int_0^t \rho_{jk}^{\mathrm{C}}(s)\,\mathrm{d}M_k(s).\end{aligned}$$

Furthermore, introduce the integrals

$$I_0^j[x](t) := x(t), \quad I_n^j[x](t) := \int_0^t a_{jj}^{\mathrm{C}}(s)I_{n-1}^j[x](s)\,\mathrm{d}s, \quad n \in \mathbb{N}, \tag{5.5.1}$$

where $x\colon [0,T] \to \mathbb{R}$ is a measurable function such that the integrals in (5.5.1) *exist for $t \in [0,T]$. Then*

$$\bar{X}_j(t) - \mathbb{E}[\bar{X}_j(t)] = \sum_{n=0}^{\infty} I_n^j[Y_j](t) = \sum_{\iota=1}^{5}\sum_{n=0}^{\infty} I_n^j[Y_j^\iota](t), \quad t \in [0,T], \tag{5.5.2}$$

where the sums converge with respect to the maximal L^2-norm $X \mapsto \|X^*(T)\|_{L^2}$.

Proof. We deduce from (5.2.4) that

$$\begin{aligned}\bar{X}_j(t) - \mathbb{E}[\bar{X}_j(t)] &= (X_j(0) - \mathbb{E}[X_j(0)]) + \sum_{k\neq j} \int_0^t a^{\mathrm{C}}_{jk}(s)(\bar{X}_k(s) - \mathbb{E}[\bar{X}_k(s)])\,\mathrm{d}s \\ &\quad + \int_0^t a^{\mathrm{C}}_{jj}(s)(\bar{X}_j(s) - \mathbb{E}[\bar{X}_j(s)])\,\mathrm{d}s \\ &\quad + \sum_{k=1}^{\infty} \int_0^t \left(\sigma^{\mathrm{C}}_{jk}(s)\bar{X}_k(s-) + \sigma^{\mathrm{P}}_{jk}(s)\mathbb{E}[\bar{X}_k(s)]\right) \mathrm{d}L_j(s) \\ &\quad + \sum_{k=1}^{\infty} \int_0^t f^{\mathrm{C}}_{jk}(s)(b_k(s) - \mathbb{E}[b_k(s)])\,\mathrm{d}s + \sum_{k=1}^{\infty} \int_0^t \rho^{\mathrm{C}}_{jk}(s)\,\mathrm{d}M_k(s) \\ &= I^j_1[\bar{X}_j - \mathbb{E}[\bar{X}_j]](t) + Y_j(t).\end{aligned}$$

Iterating this equality n times, we obtain

$$\bar{X}_j(t) - \mathbb{E}[\bar{X}_j(t)] = \sum_{\nu=0}^{n-1} I^j_\nu[Y_j](t) + I^j_n[\bar{X}_j - \mathbb{E}[\bar{X}_j]](t), \quad t \in [0,T]. \tag{5.5.3}$$

Next, observe that for any càdlàg process $(X(t))_{t\in\mathbb{R}_+}$ with $\|X^*(T)\|_{L^2} < \infty$ we have

$$\left\|(I^j_\nu[X])^*(T)\right\|_{L^2} \leq \|X^*(T)\|_{L^2} I^j_\nu[1](T) \leq \|X^*(T)\|_{L^2} \frac{(A^{\mathrm{C}}_{jj}(T))^\nu}{\nu!},$$

which is summable in ν. Therefore, recalling from Lemma 5.5.1 that both Y_j and $\bar{X}_j - \mathbb{E}[\bar{X}_j]$ have finite maximal L^2-norm, we can let $n \to \infty$ in (5.5.3) and get

$$\bar{X}_j(t) - \mathbb{E}[\bar{X}_j(t)] = \sum_{\nu=0}^{\infty} I^j_\nu[Y_j](t), \quad t \in [0,T],$$

which is the first assertion. The second part of (5.5.2) holds by linearity. □

Proof of Theorem 5.3.1. The existence and uniqueness of solutions to (5.2.2) and (5.2.4) follow from the general theory of SDEs, see [118], Theorem V.7. Since the numbers (5.3.1) are finite, there are no difficulties in dealing with infinite-dimensional systems as in our case.

It follows from (5.2.2) and (5.2.4) that the difference between X and $\bar{X}$ satisfies the SDE

$$\begin{aligned}\mathrm{d}(X(t) - \bar{X}(t)) &= \left(a(t)(X(t) - \bar{X}(t)) + a^{\mathrm{P}}(t)(\bar{X}(t) - \mathbb{E}[\bar{X}(t)])\right)\,\mathrm{d}t \\ &\quad + \left(\sigma(t)(X(t-) - \bar{X}(t-)) + \sigma^{\mathrm{P}}(t)(\bar{X}(t-) - \mathbb{E}[\bar{X}(t)])\right).\mathrm{d}L(t) \\ &\quad + f^{\mathrm{P}}(t)(b(t) - \mathbb{E}[b(t)])\,\mathrm{d}t + \rho^{\mathrm{P}}(t)\,\mathrm{d}M(t), \quad t \in \mathbb{R}_+, \\ X(0) - \bar{X}(0) &= 0.\end{aligned}$$

Thus, denoting the left-hand side of (5.3.3) by $\Delta(T)$, we obtain from the Burkholder-Davis-Gundy inequality and Jensen's inequality that

$$\begin{aligned}\Delta(T) &\le v_a(T)\int_0^T \Delta(t)\,\mathrm{d}t + 2v_\sigma(T)v_L\left(\int_0^T (\Delta(t))^2\,\mathrm{d}t\right)^{1/2}\\ &\quad + \left|\int_0^T \left\|a^{\mathrm{P}}(t)(\bar{X}(t)-\mathbb{E}[\bar{X}(t)])\right\|_{L^2}\,\mathrm{d}t\right|_\infty + \left|\left\|\left(\sigma^{\mathrm{P}}(\bar{X}-\mathbb{E}[\bar{X}])\,.\,L\right)^*(T)\right\|_{L^2}\right|_\infty\\ &\quad + T\sup_{t\in[0,T]}\left|\left\|f^{\mathrm{P}}(t)(b(t)-\mathbb{E}[b(t)])\right\|_{L^2}\right|_\infty + \left|\left\|(\rho^{\mathrm{P}}\cdot M)^*(T)\right\|_{L^2}\right|_\infty\\ &\le (T^{1/2}v_a(T)+2v_\sigma(T)v_L)\left(\int_0^T(\Delta(t))^2\,\mathrm{d}t\right)^{1/2} + \sum_{\iota=1}^4 \Delta^\iota(T), \qquad (5.5.4)\end{aligned}$$

where $\Delta^\iota(T)$ stands for the last four summands in the line before. So Gronwall's inequality produces the bound

$$\Delta(T) \le K(T)\sum_{\iota=1}^4 \Delta^\iota(T), \qquad (5.5.5)$$

where $K(T) = \sqrt{2}\exp((T^{1/2}v_a(T)+2v_\sigma(T)v_L)^2T)$. We now consider each $\Delta^\iota(T)$ separately.

For $\iota = 3$ we simply have

$$\Delta^3(T) \le T\sup_{t\in[0,T]}\sup_{i\in\mathbb{N}}\left(\sum_{j,k=1}^\infty f^{\mathrm{P}}_{ij}(t)f^{\mathrm{P}}_{ik}(t)\mathrm{Cov}[b_j(t),b_k(t)]\right)^{1/2} = Tr_5(T). \qquad (5.5.6)$$

For $\iota = 4$ another application of the Burkholder-Davis-Gundy inequality yields

$$\begin{aligned}\Delta^4(T) &\le 2\sup_{i\in\mathbb{N}}\left(\sum_{j,k=1}^\infty \mathbb{E}\left[\int_0^T \rho^{\mathrm{P}}_{ij}(t)\rho^{\mathrm{P}}_{ik}(t)\,\mathrm{d}[M_j,M_k](t)\right]\right)^{1/2}\\ &\le 2T^{1/2}\sup_{i\in\mathbb{N}}\sup_{t\in[0,T]}\left(\sum_{j,k=1}^\infty \mathbb{E}\left[\rho^{\mathrm{P}}_{ij}(t)c_{jk}(t)\rho^{\mathrm{P}}_{ik}(t)\right]\right)^{1/2} = 2T^{1/2}r_6(T). \qquad (5.5.7)\end{aligned}$$

For $\iota = 1$, we use Lemma 5.5.3 including the notations introduced there and the fact that for all stochastic processes $(X(t))_{t\in\mathbb{R}_+}$ and $(Y(t))_{t\in\mathbb{R}_+}$ with càdlàg sample paths we have

$$\sup_{r,s\in[0,T]}\left|\mathbb{E}\left[I^j_n[X](s)I^k_m[Y](r)\right]\right| \le \frac{(A^{\mathrm{C}}_{jj}(T))^n}{n!}\frac{(A^{\mathrm{C}}_{kk}(T))^m}{m!}\sup_{r,s\in[0,T]}|\mathbb{E}[X(s)Y(r)]| \qquad (5.5.8)$$

for any $j,k\in\mathbb{N}$ and $m,n\in\mathbb{N}\cup\{0\}$. In this way we obtain

$$\begin{aligned}
\Delta^1(T) &= \left|\int_0^T \left\|a^{\mathrm{P}}(t)(\bar{X}(t)-\mathbb{E}[\bar{X}(t)])\right\|_{L^2}\,\mathrm{d}t\right|_\infty \\
&= \sup_{i\in\mathbb{N}}\int_0^T \left\|\sum_{j=1}^\infty a^{\mathrm{P}}_{ij}(t)(\bar{X}_j(t)-\mathbb{E}[\bar{X}_j(t)])\right\|_{L^2}\,\mathrm{d}t \\
&\le \sum_{\iota=1}^5 \sup_{i\in\mathbb{N}}\int_0^T \left\|\sum_{j=1}^\infty a^{\mathrm{P}}_{ij}(t)\sum_{n=0}^\infty I_n^j[Y_j^\iota](t)\right\|_{L^2}\,\mathrm{d}t \\
&= \sum_{\iota=1}^5 \sup_{i\in\mathbb{N}}\int_0^T \left(\sum_{j,k=1}^\infty\sum_{n,m=0}^\infty a^{\mathrm{P}}_{ij}(t)a^{\mathrm{P}}_{ik}(t)\mathbb{E}\left[I_n^j[Y_j^\iota](t)I_m^k[Y_k^\iota](t)\right]\right)^{1/2}\mathrm{d}t \\
&\le \mathrm{e}^{|A^{\mathrm{C}}(T)|_{\mathrm{d}}}\sum_{\iota=1}^5 \sup_{i\in\mathbb{N}}\int_0^T \left(\sum_{j,k=1}^\infty A^{\mathrm{P}}_{ij}(T)A^{\mathrm{P}}_{ik}(T)\sup_{r,s\in[0,t]}\left|\mathbb{E}[Y_j^\iota(s)Y_k^\iota(r)]\right|\right)^{1/2}\mathrm{d}t \\
&=: \mathrm{e}^{|A^{\mathrm{C}}(T)|_{\mathrm{d}}}\sum_{\iota=1}^5 R_\iota(T). \qquad (5.5.9)
\end{aligned}$$

Using Lemma 5.5.1 and Lemma 5.5.2, the five terms in (5.5.9) can be estimated as follows:

$$\begin{aligned}
R_1(T) &\le T\sup_{i\in\mathbb{N}}\left(\sum_{j,k=1}^\infty A^{\mathrm{P}}_{ij}(T)A^{\mathrm{P}}_{ik}(T)|\mathrm{Cov}[X_j(0),X_k(0)]|\right)^{1/2} = Tr_1(T), \\
R_2(T) &\le \sup_{i\in\mathbb{N}}\int_0^T\sum_{j=1}^\infty A^{\mathrm{P}}_{ij}(T)\sup_{s\in[0,t]}\|Y_j^2(s)\|_{L^2}\,\mathrm{d}t \\
&\le \sup_{i\in\mathbb{N}}\int_0^T\sum_{j=1}^\infty A^{\mathrm{P}}_{ij}(T)\left(\sum_{k\neq j}\int_0^t A^{\mathrm{C}}_{jk}(s)\|\bar{X}_k(s)-\mathbb{E}[\bar{X}_k(s)]\|_{L^2}\,\mathrm{d}s\right)\mathrm{d}t \\
&\le \frac{T^2}{2}V(T)\sup_{i\in\mathbb{N}}\sum_{j=1}^\infty\sum_{k\neq j}A^{\mathrm{P}}_{ij}A^{\mathrm{C}}_{jk} = \frac{T^2}{2}V(T)r_7(T), \\
R_3(T) &\le \sup_{i\in\mathbb{N}}\int_0^T\left(\sum_{j,k=1}^\infty A^{\mathrm{P}}_{ij}(T)A^{\mathrm{P}}_{ik}(T)\sup_{s\in[0,t]}\left|\mathbb{E}\Big[\Big(\left(\sigma^{\mathrm{C}}\bar{X}+\sigma^{\mathrm{P}}\mathbb{E}[\bar{X}]\right)_j\cdot L_j\Big)(s)\right.\right. \\
&\qquad\left.\left.\times\Big(\left(\sigma^{\mathrm{C}}\bar{X}+\sigma^{\mathrm{P}}\mathbb{E}[\bar{X}]\right)_k\cdot L_k\Big)(s)\Big]\right|\right)^{1/2}\mathrm{d}t \\
&\le \sup_{i\in\mathbb{N}}\int_0^T\left(\sum_{j,k=1}^\infty A^{\mathrm{P}}_{ij}(T)A^{\mathrm{P}}_{ik}(T)\mathrm{Cov}[L_j(1),L_k(1)]\right.
\end{aligned}$$

$$
\begin{aligned}
&\times \int_0^t \mathbb{E}\Bigg[\Big|\Big(\sigma^{\mathrm{C}}(s)\bar{X}(s)+\sigma^{\mathrm{P}}(s)\mathbb{E}[\bar{X}(s)]\Big)_j\\
&\times\Big(\sigma^{\mathrm{C}}(s)\bar{X}(s)+\sigma^{\mathrm{P}}(s)\mathbb{E}[\bar{X}(s)]\Big)_k\Big|\Bigg]\,\mathrm{d}s\Bigg)^{1/2}\mathrm{d}t\\
&\le \frac{2}{3}T^{3/2}v_\sigma(T)V(T)r_3(T),\\
R_4(T) &\le \sup_{i\in\mathbb{N}}\int_0^T\Bigg(\sum_{j,k=1}^{\infty}A^{\mathrm{P}}_{ij}(T)A^{\mathrm{P}}_{ik}(T)\\
&\times\int_0^t\int_0^t\Bigg|\sum_{l,m=1}^{\infty}f^{\mathrm{C}}_{jl}(s)f^{\mathrm{C}}_{km}(r)\mathrm{Cov}[b_l(s),b_m(r)]\Bigg|\,\mathrm{d}r\,\mathrm{d}s\Bigg)^{1/2}\mathrm{d}t\\
&\le \frac{T^2}{2}\sup_{i\in\mathbb{N}}\Bigg(\sum_{j,k=1}^{\infty}A^{\mathrm{P}}_{ij}(T)A^{\mathrm{P}}_{ik}(T)\sup_{s,t\in[0,T]}\Bigg|\sum_{l,m=1}^{\infty}f^{\mathrm{C}}_{jl}(s)f^{\mathrm{C}}_{km}(t)\mathrm{Cov}[b_l(s),b_m(t)]\Bigg|\Bigg)^{1/2}\\
&= \frac{T^2}{2}r_9(T),\\
R_5(T) &\le \sup_{i\in\mathbb{N}}\int_0^T\Bigg(\sum_{j,k=1}^{\infty}A^{\mathrm{P}}_{ij}(T)A^{\mathrm{P}}_{ik}(T)\\
&\times\sup_{s\in[0,t]}\Bigg|\sum_{l,m=1}^{\infty}\mathbb{E}\left[(\rho^{\mathrm{C}}_{jl}\cdot M_l)(s)(\rho^{\mathrm{C}}_{km}\cdot M_m)(s)\right]\Bigg|\Bigg)^{1/2}\mathrm{d}t\\
&\le \sup_{i\in\mathbb{N}}\int_0^T\Bigg(\sum_{j,k=1}^{\infty}A^{\mathrm{P}}_{ij}(T)A^{\mathrm{P}}_{ik}(T)\int_0^t\Bigg|\sum_{l,m=1}^{\infty}\mathbb{E}\left[\rho^{\mathrm{C}}_{jl}(s)c_{lm}(s)\rho^{\mathrm{C}}_{km}(s)\right]\Bigg|\,\mathrm{d}s\Bigg)^{1/2}\mathrm{d}t\\
&\le \frac{2}{3}T^{3/2}r_{11}(T).
\end{aligned}
$$

The last step in the proof is the estimation of $\Delta^2(T)$. To this end, we make use of the Burkholder-Davis-Gundy inequality another time and get

$$
\begin{aligned}
\Delta^2(T) &\le \sup_{i\in\mathbb{N}}\Big\|\Big(\Big(\sigma^{\mathrm{P}}(\bar{X}-\mathbb{E}[\bar{X}])\Big)_i\cdot L_i\Big)^*(T)\Big\|_{L^2}\\
&\le 2v_L\sup_{i\in\mathbb{N}}\left(\int_0^T\mathbb{E}\left[\Big(\sigma^{\mathrm{P}}(t)(\bar{X}(t)-\mathbb{E}[\bar{X}(t)])\Big)_i^2\right]\mathrm{d}t\right)^{1/2}.
\end{aligned}
$$

The further procedure is analogous to what we have done for $\Delta^1(T)$: instead of a^{P} we have σ^{P} here. We leave the details to the reader and only state the result, which

is

$$\Delta^2(T) \leq \sum_{\iota \in \{2,4,8,10,12\}} K_\iota(T) r_\iota(T).$$

This completes the proof of Theorem 5.3.1. □

Our next goal is to prove Lemma 5.3.8 concerning the rate of growth of the maximal degree in the preferential attachment random graph as described in Section 5.3.3. For the undirected version as in [10] the corresponding result goes back to [105]. Indeed, the proof there basically works for our case as well, but there are some steps that require different arguments. Thus, we decided to include the proof to our lemma.

Proof of Lemma 5.3.8. The statement is evidently true for M^{in} when $\alpha + \beta = 0$ (resp. for M^{out} when $\beta + \gamma = 0$). In fact, for this extremal case, in every step of the random graph a new edge is created pointing to (resp. from) a new node. This means that $M^{\text{in}}(N)$ (resp. $M^{\text{out}}(N)$) remains constant for all $N \in \mathbb{N}_0$, and the claim follows with $c^{\text{in}} = 1$ (resp. $c^{\text{out}} = 1$) identically. In the other cases, we closely follow the proof of [105, Thm. 3.1]. In addition to the notation introduced in Section 5.3.3, we further define for $N \in \mathbb{N}_0$ and $\diamond \in \{\text{in}, \text{out}\}$:

$$\begin{aligned}
S^\diamond(N) &:= \nu + N + \delta^\diamond n(N), \\
X^\diamond(N,j) &:= \mathrm{d}^\diamond_{G(N)}(j) + \delta^\diamond, \quad j \in \mathbb{N}, \\
N^\diamond_j &:= \inf\{N \in \mathbb{N}_0 \colon \mathrm{d}^\diamond_{G(N)}(j) \neq 0\}, \quad j \in \mathbb{N}, \\
s^\diamond &:= \alpha \mathbb{1}_{\{\diamond = \text{in}\}} + \beta + \gamma \mathbb{1}_{\{\diamond = \text{out}\}}, \\
c^\diamond(0,k) &:= 1, \quad c^\diamond(N+1,k) := c^\diamond(N,k) \frac{S^\diamond(N)}{S^\diamond(N) + s^\diamond k}, \quad k \in \mathbb{R}_+, \\
Z^\diamond(N,j,k) &:= c^\diamond(N^\diamond_j + N, k) \binom{X^\diamond(N^\diamond_j + N, j) + k - 1}{k} \mathbb{1}_{\{N^\diamond_j < \infty\}}, \quad j \in \mathbb{N}, \quad k \in \mathbb{R}_+, \\
\mathcal{G}(N) &:= \sigma(e_{\nu+i} \colon i = 1, \ldots, N), \quad \mathcal{G}(\infty) := \sigma\left(\bigcup_{N=0}^{\infty} \mathcal{G}(N) \right).
\end{aligned}$$

Obviously, $\mathcal{G}(N)$ is the σ-field of all information up to step N in the preferential attachment algorithm, and $N^\diamond_j$ is a stopping time relative to the filtration $(\mathcal{G}(N))_{N \in \mathbb{N}}$ for every $j \in \mathbb{N}$. Analogously to [105, Thm. 2.1] one can now show that for all $k \in \mathbb{R}_+$ and $j \in \mathbb{N}$ the sequence $(Z^\diamond(N,j,k))_{N \in \mathbb{N}_0}$ is a positive martingale relative to

the filtration $(\mathcal{G}(N_j^\diamond + N))_{N\in\mathbb{N}_0}$. As a consequence, Doob's martingale convergence theorem implies that

$$Z^\diamond(N,j,k) \to \zeta^\diamond(j,k) \quad \text{a.s.} \tag{5.5.10}$$

for some random variables $\zeta^\diamond(j,k)$. The convergence in (5.5.10) also holds in L^p for all $p \in [1,\infty)$ because we have

$$Z^\diamond(N,j,k)^p \le C(k,p) Z^\diamond(N,j,kp) \quad \text{a.s.} \tag{5.5.11}$$

for some deterministic constants $C(k,p) \in \mathbb{R}_+$ independent of N and j. Indeed, on $\{N_j^\diamond < \infty\}$ we have by definition

$$\begin{aligned}\frac{Z^\diamond(N,j,k)^p}{Z^\diamond(N,j,kp)} &= \frac{c^\diamond(N_j^\diamond+N,k)^p}{c^\diamond(N_j^\diamond+N,kp)} \binom{X^\diamond(N_j^\diamond+N,j)+k-1}{k}^p \\ &\quad \times \binom{X^\diamond(N_j^\diamond+N,j)+kp-1}{kp}^{-1},\end{aligned}$$

where

$$\begin{aligned}\frac{c^\diamond(N,k)^p}{c^\diamond(N,kp)} &= \frac{c^\diamond(N-1,k)^p}{c^\diamond(N-1,kp)} \frac{S^\diamond(N)^{p-1}(S^\diamond(N)+s^\diamond kp)}{(S^\diamond(N)+s^\diamond k)^p} \le \frac{c^\diamond(N-1,k)^p}{c^\diamond(N-1,kp)} \le \cdots \\ &\le \frac{c^\diamond(0,k)^p}{c^\diamond(0,kp)} = 1,\end{aligned}$$

$$\binom{x+k-1}{k}^p \binom{x+kp-1}{kp}^{-1} = \frac{\Gamma(kp+1)}{\Gamma(k+1)^p} \frac{\Gamma(k+x)^p}{\Gamma(x)^{p-1}\Gamma(kp+x)} \overset{x\to\infty}{\longrightarrow} \frac{\Gamma(kp+1)}{\Gamma(k+1)^p},$$

which shows (5.5.11). Next, define for $N \in \mathbb{N}_0$ and $j \in \mathbb{N}$

$$\begin{aligned} m^\diamond(N,j) &:= \max\{Z^\diamond(N-N_i^\diamond,i,1)\colon i=1,\dots,j, N_i^\diamond \le N\}, \\ m^\diamond(N) &:= m^\diamond(N,n(N)), \\ \mu^\diamond(j) &:= \max\{\zeta^\diamond(i,1)\colon i=1,\dots,j\}, \\ \mu^\diamond &:= \sup\{\mu^\diamond(j)\colon j\in\mathbb{N}\},\end{aligned}$$

such that in particular the relationship $m^\diamond(N) = c^\diamond(N,1)(M^\diamond(N)+\delta^\diamond)$ holds. It is not hard to see that $(m^\diamond(N))_{N\in\mathbb{N}_0}$, as the maximum of martingale expressions, is a submartingale relative to $(\mathcal{G}(N))_{N\in\mathbb{N}_0}$. By definition the sequence $(c^\diamond(N,k))_{N\in\mathbb{N}_0}$

decreases to 0 as $N \to \infty$; more precisely, we have

$$\begin{aligned} c^\diamond(N,k) &= c^\diamond(N-1,k)\frac{S^\diamond(N-1)}{S^\diamond(N-1)+s^\diamond k} \\ &\le c^\diamond(N-1,k)\frac{\nu+N-1+\delta^\diamond(n(0)+N-1)}{\nu+N-1+\delta^\diamond(n(0)+N-1)+s^\diamond k} \\ &\le \prod_{j=0}^{N-1}\frac{(1+\delta^\diamond)j+\delta^\diamond n(0)+\nu}{(1+\delta^\diamond)j+\delta^\diamond n(0)+\nu+s^\diamond k} = \frac{\Gamma\left(N+\frac{\delta^\diamond n(0)+\nu}{1+\delta^\diamond}\right)}{\Gamma\left(N+\frac{\delta^\diamond n(0)+\nu+s^\diamond k}{1+\delta^\diamond}\right)} \sim N^{-\frac{s^\diamond k}{1+\delta^\diamond}} \end{aligned}$$

as $N \to \infty$. As a consequence, when p is large enough,

$$\begin{aligned} \mathbb{E}[m^\diamond(N)^p] &\le \mathbb{E}\left[\sum_{i=1}^{n(N)} Z^\diamond(N-N_i^\diamond,i,1)^p\right] \le C(1,p)\mathbb{E}\left[\sum_{i=1}^{n(N)} Z^\diamond(N-N_i^\diamond,i,p)\right] \\ &\le C(1,p)\sum_{i=1}^{\infty}\mathbb{E}[Z^\diamond(0,i,p)] \le C(1,p)\binom{n(0)+p+\delta^\diamond-1}{p}\sum_{i=1}^{\infty}\mathbb{E}[c^\diamond(N_i^\diamond,p)] \\ &\le C(1,p)\binom{n(0)+p+\delta^\diamond-1}{p}\left(n(0)+\sum_{i=1}^{\infty}\mathbb{E}[c^\diamond(i,p)]\right) < \infty \end{aligned} \tag{5.5.12}$$

independently of N. This implies that the submartingale $m^\diamond$ converges a.s. and in L^p for all $p \in [1,\infty)$. It follows from (5.5.12) that for $j \ge n(0)$ we have

$$\begin{aligned} &\mathbb{E}[(m^\diamond(N)-m^\diamond(N,j))^p] \\ &\quad\le \mathbb{E}\left[\sum_{i=j+1}^{n(N)} Z^\diamond(N-N_i^\diamond,i,1)^p\right] \\ &\quad\le C(1,p)\binom{n(0)+p+\delta^\diamond-1}{p}\sum_{i=j-n(0)+1}^{\infty}\mathbb{E}[c^\diamond(i,p)]. \end{aligned} \tag{5.5.13}$$

Letting $N \to \infty$, the left-hand side of (5.5.13) converges to

$$\mathbb{E}\left[\left(\lim_{N\to\infty} c^\diamond(N,1)M^\diamond(N)-\mu^\diamond(j)\right)^p\right],$$

while the right-hand side is independent of N. Now taking the limit $j \to \infty$ and again assuming that p is large, we obtain the desired result (5.3.13). Note at this point that $\mu^\diamond$ is indeed an a.s. finite random variable that belongs to L^p for all $p \in [1,\infty)$, which is proved using a similar argument as in (5.5.12).

It remains to prove (5.3.14). To this end, observe that by the law of large numbers we have $(n(N)-n(0))/N \to \alpha+\gamma$ a.s. In other words, for every $\epsilon \in (0, \alpha+\gamma)$ there exists a possibly random $\bar{N} \in \mathbb{N}$ such that for all $N \geq \bar{N}$ we have

$$\left| \frac{n(N)-n(0)}{N} - (\alpha+\gamma) \right| \leq \epsilon,$$

or, equivalently, $n(N) \in [n(0)+(\alpha+\gamma-\epsilon)N, n(0)+(\alpha+\gamma+\epsilon)N]$. Consequently, for all $k \in \mathbb{N}$ and $N \geq \bar{N}$

$$\begin{aligned}
c^\diamond(N,k) &= \prod_{i=0}^{N-1} \frac{S^\diamond(i)}{S^\diamond(i)+s^\diamond k} \\
&\geq \prod_{i=0}^{\bar{N}-1} \frac{S^\diamond(i)}{S^\diamond(i)+s^\diamond k} \prod_{i=\bar{N}}^{N-1} \frac{\nu+i+\delta^\diamond(n(0)+(\alpha+\gamma-\epsilon)i)}{\nu+i+\delta^\diamond(n(0)+(\alpha+\gamma-\epsilon)i)+s^\diamond k} \\
&= \prod_{i=0}^{\bar{N}-1} \frac{S^\diamond(i)}{S^\diamond(i)+s^\diamond k} \prod_{i=0}^{\bar{N}-1} \frac{\nu+i+\delta^\diamond(n(0)+(\alpha+\gamma-\epsilon)i)+s^\diamond k}{\nu+i+\delta^\diamond(n(0)+(\alpha+\gamma-\epsilon)i)} \\
&\quad \times \prod_{i=0}^{N-1} \frac{\nu+i+\delta^\diamond(n(0)+(\alpha+\gamma-\epsilon)i)}{\nu+i+\delta^\diamond(n(0)+(\alpha+\gamma-\epsilon)i)+s^\diamond k} \\
&= c^\diamond(\bar{N},k) \frac{\Gamma\left(\bar{N}+\frac{\delta^\diamond n(0)+\nu+s^\diamond k}{1+\delta^\diamond(\alpha+\gamma-\epsilon)}\right)\Gamma\left(N+\frac{\delta^\diamond n(0)+\nu}{1+\delta^\diamond(\alpha+\gamma-\epsilon)}\right)}{\Gamma\left(\bar{N}+\frac{\delta^\diamond n(0)+\nu}{1+\delta^\diamond(\alpha+\gamma-\epsilon)}\right)\Gamma\left(N+\frac{\delta^\diamond n(0)+\nu+s^\diamond k}{1+\delta^\diamond(\alpha+\gamma-\epsilon)}\right)} \\
&\sim c^\diamond(\bar{N},k) \frac{\Gamma\left(\bar{N}+\frac{\delta^\diamond n(0)+\nu+s^\diamond k}{1+\delta^\diamond(\alpha+\gamma-\epsilon)}\right)}{\Gamma\left(\bar{N}+\frac{\delta^\diamond n(0)+\nu}{1+\delta^\diamond(\alpha+\gamma-\epsilon)}\right)} N^{-\frac{s^\diamond k}{1+\delta^\diamond(\alpha+\gamma-\epsilon)}}, \quad N \to \infty.
\end{aligned}$$

So choosing $c^\diamond(N) := c^\diamond(N,1)$ for $N \in \mathbb{N}_0$ fulfills (5.3.14). □

Finally, we turn to the proof of the large deviation result in Section 5.4.

Proof of Theorem 5.4.1. By definition, we have for $t \in [0,T]$ that

$$X^N(t)-\bar{X}^N(t) = \int_0^t a^N(X^N-\bar{X}^N)(s)\,\mathrm{d}s + \int_0^t a^{N,\mathrm{P}}(\bar{X}^N(s)-\mathbb{E}[\bar{X}^N(s)])\,\mathrm{d}s + \rho^{N,\mathrm{P}} M(t),$$

whose solution is

$$X^N(t)-\bar{X}^N(t) = \int_0^t \mathrm{e}^{a^N(t-s)} a^{N,\mathrm{P}}(\bar{X}^N(s)-\mathbb{E}[\bar{X}^N(s)])\,\mathrm{d}s + \int_0^t \mathrm{e}^{a^N(t-s)} \rho^{N,\mathrm{P}}\,\mathrm{d}M(s).$$

In order to establish a large deviation principle, it suffices by [55, Thm. 4.6.1] to prove such a principle in $(D_T^d, \mathcal{D}_T^d, J_1)$ for the first d coordinates of the process for every $d \in \mathbb{N}$, that is, for the D_T^d-valued process

$$\begin{aligned} Y_i^N(t) &:= Y_i^{N,1}(t) + Y_i^{N,2}(t) + Y_i^{N,3}(t) \\ &:= \int_0^t \int_0^s \sum_{j,k=1}^{\infty} \mathrm{e}_{ij}^{a^N(t-s)} a_{jk}^{N,\mathrm{P}} \mathrm{e}^{a_{kk}^{N,\mathrm{C}}(s-r)} \sum_{l \neq k} a_{kl}^{N,\mathrm{C}} (\bar{X}_l^N(r) - \mathbb{E}[\bar{X}_l^N(r)]) \,\mathrm{d}r \,\mathrm{d}s \\ &\quad + \int_0^t \sum_{j,k,l=1}^{\infty} \mathrm{e}_{ij}^{a^N(t-s)} a_{jk}^{N,\mathrm{P}} \int_0^s \mathrm{e}^{a_{kk}^{N,\mathrm{C}}(s-r)} \rho_{kl}^{N,\mathrm{C}} \,\mathrm{d}M_l(r) \,\mathrm{d}s \\ &\quad + \sum_{j,k=1}^{\infty} \int_0^t \mathrm{e}_{ij}^{a^N(t-s)} \rho_{jk}^{N,\mathrm{P}} \,\mathrm{d}M_k(s), \quad i = 1, \dots, d, \quad t \in [0,T], \end{aligned} \tag{5.5.14}$$

where we have used the formula

$$\begin{aligned} \bar{X}_i^N(t) - \mathbb{E}[\bar{X}_i^N(t)] &= \int_0^t \mathrm{e}^{a_{ii}^{N,\mathrm{C}}(t-s)} \sum_{j \neq i} a_{ij}^{N,\mathrm{C}} (\bar{X}_j^N(s) - \mathbb{E}[\bar{X}_j^N(s)]) \,\mathrm{d}s \\ &\quad + \sum_{j=1}^{\infty} \int_0^t \mathrm{e}^{a_{ii}^{N,\mathrm{C}}(t-s)} \rho_{ij}^{N,\mathrm{C}} \,\mathrm{d}M_j(s), \end{aligned}$$

valid for all $i \in \mathbb{N}$ and $t \in \mathbb{R}_+$. Actually, we will even prove the large deviation principle in $(D_T^d, \mathcal{D}_T^d, U)$, which is stronger. To this end, we introduce the notation

$$\hat{x}(t) := \sum_{k=1}^{[\gamma(N)T]-1} x\left(\tfrac{k}{\gamma(N)}\right) \mathbb{1}_{\left[\frac{k}{\gamma(N)}, \frac{k+1}{\gamma(N)}\right)}(t) + x\left(\tfrac{[\gamma(N)T]}{\gamma(N)}\right) \mathbb{1}_{\left[\frac{[\gamma(N)T]}{\gamma(N)}, T\right)}(t), \quad t \in [0,T],$$

for $x \in D_T^d$. Then by [55, Thm. 4.2.13] and Lemma 5.5.4 below we can equally well show a large deviation principle for $\hat{Y}^N = \hat{Y}^{N,1} + \hat{Y}^{N,2} + \hat{Y}^{N,3}$. The same principle will then hold for Y^N. But this is proved in Lemma 5.5.6. That the rate function for $X^N - \bar{X}^N$ is convex with unique minimum 0 at 0 and can only be finite for functions in AC_T^∞, is inherited from the rate function of $\hat{Y}^N$. □

Lemma 5.5.4. *For each $d \in \mathbb{N}$ and $\iota = 1, 2, 3$, the D_T^d-valued processes $Y^{N,\iota}$ and $\hat{Y}^{N,\iota}$ are* exponentially equivalent*, that is, for all $\epsilon \in (0,1)$ we have*

$$\lim_{N \to \infty} \frac{1}{\gamma(N)} \log \mathbb{P}\left[\sup_{t \in [0,T]} \sup_{i=1,\dots,d} |Y_i^{N,\iota}(t) - \hat{Y}_i^{N,\iota}(t)| > \epsilon \right] = -\infty.$$

Proof. We start with $\iota = 1$. Writing $\hat{t} = [\gamma(N)t]/\gamma(N)$ and $\mathrm{diag}(a) := a - a^{\times}$, we obtain

$$\begin{aligned}
&\sup_{t\in[0,T]} \sup_{i=1,\dots,d} \left| Y_i^{N,1}(t) - \hat{Y}_i^{N,1}(t) \right| \\
&\quad \le \sup_{t\in[0,T]} \left| (\mathrm{e}^{a^N t} - \mathrm{e}^{a^N \hat{t}}) \left(\int_0^t \int_0^s \mathrm{e}^{-a^N s} a^{N,\mathrm{P}} \mathrm{e}^{\mathrm{diag}(a^{N,\mathrm{C}})(s-r)} (a^{N,\mathrm{C}})^{\times} \right.\right. \\
&\qquad \left.\left. \times (\bar{X}^N(r) - \mathbb{E}[\bar{X}^N(r)])\, \mathrm{d}r\, \mathrm{d}s \right) \right|_{\infty} \\
&\qquad + \sup_{t\in[0,T]} \left| \mathrm{e}^{a^N \hat{t}} \left(\int_{\hat{t}}^t \int_0^s \mathrm{e}^{-a^N s} a^{N,\mathrm{P}} \mathrm{e}^{\mathrm{diag}(a^{N,\mathrm{C}})(s-r)} (a^{N,\mathrm{C}})^{\times} \right.\right. \\
&\qquad \left.\left. \times (\bar{X}^N(r) - \mathbb{E}[\bar{X}^N(r)])\, \mathrm{d}r\, \mathrm{d}s \right) \right|_{\infty}. \qquad (5.5.15)
\end{aligned}$$

We can proceed with these two terms separately. Since $|\mathrm{e}^{a^N t} - \mathrm{e}^{a^N \hat{t}}|_{\infty} \le v_a \mathrm{e}^{v_a T}/\gamma(N)$, we have for the first term in (5.5.15)

$$\begin{aligned}
&\mathbb{P}\left[\sup_{t\in[0,T]} \left| (\mathrm{e}^{a^N t} - \mathrm{e}^{a^N \hat{t}}) \left(\int_0^t \int_0^s \mathrm{e}^{-a^N s} a^{N,\mathrm{P}} \mathrm{e}^{\mathrm{diag}(a^{N,\mathrm{C}})(s-r)} (a^{N,\mathrm{C}})^{\times} \right.\right.\right. \\
&\qquad \left.\left.\left. \times (\bar{X}^N(r) - \mathbb{E}[\bar{X}^N(r)])\, \mathrm{d}r\, \mathrm{d}s \right) \right|_{\infty} > \epsilon \right] \\
&\le \mathbb{P}\left[\sup_{t\in[0,T]} \left| a^{N,\mathrm{P}} (a^{N,\mathrm{C}})^{\times} (\bar{X}^N(t) - \mathbb{E}[\bar{X}^N(t)]) \right|_{\infty} > \frac{\epsilon\gamma(N)}{v_a (\mathrm{e}^{v_a T})^3 T^2} \right] \\
&\le \mathbb{P}\left[\sup_{t\in[0,T]} \sup_{i\in\mathbb{N}} \sum_{j=1}^{\infty} \sum_{k\neq j} \left| a_{ij}^{N,\mathrm{P}} a_{jk}^{N,\mathrm{C}} (\bar{X}_k^N(t) - \mathbb{E}[\bar{X}_k^N(t)]) \right| > \frac{\epsilon\gamma(N)}{v_a (\mathrm{e}^{v_a T})^3 T^2} \right] =: p(N).
\end{aligned}$$

We note that $\xi^N := \bar{X}^N - \mathbb{E}[\bar{X}^N]$ satisfies the integral equation

$$\xi^N(t) = \int_0^t a^{N,\mathrm{C}} \xi^N(s)\, \mathrm{d}s + \rho^{N,\mathrm{C}} M(t), \quad t \in \mathbb{R}_+.$$

Hence we have

$$(\xi^N)^*(t) \le \int_0^t |a^{N,\mathrm{C}}| (\xi^N)^*(s)\, \mathrm{d}s + (\rho^{N,\mathrm{C}} M)^*(t), \quad t \in \mathbb{R}_+,$$

or after $n \in \mathbb{N}$ iterations,

$$(\xi^N)^*(t) \leq \frac{(|a^{N,\mathrm{C}}|t)^n}{n!}(\xi^N)^*(t) + \sum_{m=0}^{n-1} \frac{(|a^{N,\mathrm{C}}|t)^m}{m!}(\rho^{N,\mathrm{C}}M)^*(t),$$
$$(\xi^N)^*(t) \leq \left(\mathrm{I} - \frac{(|a^{N,\mathrm{C}}|T)^n}{n!}\right)^{-1} \sum_{m=0}^{n-1} \frac{(|a^{N,\mathrm{C}}|t)^m}{m!}(\rho^{N,\mathrm{C}}M)^*(t),$$

where the last line holds when n is large enough such that $(v_a T)^n/n! < 1$. It is not difficult to recognize that the exact value of n only affects some constants in the subsequent arguments with no impact on the final result; we therefore assume without loss of generality that $n = 1$ (i.e. $v_a T < 1$). Then

$$\begin{aligned} p(N) &\leq \mathbb{P}\Bigg[\sup_{i\in\mathbb{N}} \sum_{j=1}^{\infty} \sum_{k\neq j} \sum_{l=1}^{\infty} |a_{ij}^{N,\mathrm{P}} a_{jk}^{N,\mathrm{C}}| (\mathrm{I} - |a^{N,\mathrm{C}}|T)_{kl}^{-1} \\ &\qquad \times \sup_{t\in[0,T]} \sum_{m=1}^{\gamma(N)} |\rho_{lm}^{N,\mathrm{C}} M_m(t)| > \frac{\epsilon\gamma(N)}{v_a(\mathrm{e}^{v_a T})^3 T^2} \Bigg] \\ &\leq \mathbb{P}\left[\frac{1}{\gamma(N)} \sup_{l\in\mathbb{N}} \sup_{t\in[0,T]} \sum_{m=1}^{\gamma(N)} |\rho_{lm}^{N,\mathrm{C}} M_m(t)| > \frac{\epsilon\gamma(N)(1-v_a T)}{q_1(N) v_a(\mathrm{e}^{v_a T})^3 T^2} \right]. \end{aligned}$$

Let $\lambda(N)$ be positive numbers to be chosen later. Using the independence of the Lévy processes M_i and Doob's maximal inequality, we arrive at

$$\begin{aligned} p(N) &\leq \exp\left(-\frac{\epsilon\lambda(N)\gamma(N)(1-v_a T)}{q_1(N) v_a(\mathrm{e}^{v_a T})^3 T^2} \right) \prod_{m=1}^{\gamma(N)} \mathbb{E}\left[\exp\left(\frac{\lambda(N)}{\gamma(N)} \sup_{l\in\mathbb{N}} |\rho_{lm}^{N,\mathrm{C}}| |M_m(T)| \right) \right] \\ &\leq \exp\left(-\frac{\epsilon\lambda(N)\gamma(N)(1-v_a T)}{q_1(N) v_a(\mathrm{e}^{v_a T})^3 T^2} \right) \\ &\quad \times \prod_{m=1}^{\gamma(N)} \left(1 + \mathbb{E}\left[\exp\left(\frac{\lambda(N)}{\gamma(N)} \sup_{l\in\mathbb{N}} |\rho_{lm}^{N,\mathrm{C}}| M_m(T) \right) \right] \right) \\ &\leq 2^{\gamma(N)} \exp\left(-\frac{\epsilon\lambda(N)\gamma(N)(1-v_a T)}{q_1(N) v_a(\mathrm{e}^{v_a T})^3 T^2} \right) \exp\left(T\gamma(N)\Psi_0\left(\frac{\lambda(N)}{\gamma(N)} \sup_{l,m\in\mathbb{N}} |\rho_{lm}^{N,\mathrm{C}}| \right) \right). \end{aligned}$$

Now define

$$\lambda(N) := \gamma(N)\Psi_0^{-1}(1) / \left(\sup_{l,m\in\mathbb{N}} |\rho_{lm}^{N,\mathrm{C}}| \right), \quad N \in \mathbb{N}.$$

Since Ψ_0 is a convex function, its inverse Ψ_0^{-1} is concave and therefore we have for large N that $\lambda(N) \geq \Psi_0^{-1}(\gamma(N)) / \left(\sup_{l,m\in\mathbb{N}} |\rho_{lm}^{N,\mathrm{C}}| \right)$, which increases to infinity with

N. With this choice of $\lambda(N)$ it follows that

$$\lim_{N\to\infty} \frac{1}{\gamma(N)} \log p(N) = -\infty,$$

which completes the proof for the first term in (5.5.15). The second term can be treated in analogous way: now the factor $\gamma(N)$ does not come from the difference $|\mathrm{e}^{a^N t} - \mathrm{e}^{a^N \hat{t}}|_\infty$, but from the domain of integration $(\hat{t}, t]$. The details are left to the reader.

For $\iota = 2$ similar methods apply. Also here we do not give the details. Instead, we sketch the proof for $\iota = 3$ where some modifications are necessary. Recalling the meaning of $\Gamma(N)$ from (A2), we have

$$\begin{aligned}
&\sup_{t\in[0,T]} \sup_{i=1,\dots,d} \left| Y_i^{N,3}(t) - \hat{Y}_i^{N,3}(t) \right| \\
&\quad \le \sup_{t\in[0,T]} \sup_{i=1,\dots,d} \left| \left((\mathrm{e}^{a^N t} - \mathrm{e}^{a^N \hat{t}}) \left(\int_0^t \mathrm{e}^{-a^N s} \rho^{N,\mathrm{P}} \, \mathrm{d}M(s) \right) \right)_i \right| \\
&\qquad + \sup_{t\in[0,T]} \sup_{i=1,\dots,d} \left| \left(\mathrm{e}^{a^N \hat{t}} \left(\int_{\hat{t}}^t \mathrm{e}^{-a^N s} \rho^{N,\mathrm{P}} \, \mathrm{d}M(s) \right) \right)_i \right| \\
&\quad \le |\mathrm{e}^{a^N t} - \mathrm{e}^{a^N \hat{t}}|_\infty \sup_{t\in[0,T]} \sup_{i=1,\dots,\Gamma(N)} \left| \left(\int_0^t \mathrm{e}^{-a^N s} \rho^{N,\mathrm{P}} \, \mathrm{d}M(s) \right)_i \right| \\
&\qquad + \mathrm{e}^{v_a T} \sup_{t\in[0,T]} \sup_{i=1,\dots,\Gamma(N)} \left| \left(\int_{\hat{t}}^t \mathrm{e}^{-a^N s} \rho^{N,\mathrm{P}} \, \mathrm{d}M(s) \right)_i \right|. \qquad (5.5.16)
\end{aligned}$$

We can again consider these two terms separately. For the first one we have

$$\begin{aligned}
&\mathbb{P}\left[|\mathrm{e}^{a^N t} - \mathrm{e}^{a^N \hat{t}}|_\infty \sup_{t\in[0,T]} \sup_{i=1,\dots,\Gamma(N)} \left| \left(\int_0^t \mathrm{e}^{-a^N s} \rho^{N,\mathrm{P}} \, \mathrm{d}M(s) \right)_i \right| > \epsilon \right] \\
&\quad \le \Gamma(N) \sup_{i=1,\dots,\Gamma(N)} \mathbb{P}\left[\sup_{t\in[0,T]} \left| \left(\int_0^t \mathrm{e}^{-a^N s} \rho^{N,\mathrm{P}} \, \mathrm{d}M(s) \right)_i \right| > \frac{\epsilon\gamma(N)}{v_a \mathrm{e}^{v_a T}} \right] \\
&\quad \le \Gamma(N) \sup_{i=1,\dots,\Gamma(N)} \exp\left(-\frac{\epsilon\lambda(N)\gamma(N)}{v_a \mathrm{e}^{v_a T}} \right) \\
&\qquad \times \prod_{k=1}^{\gamma(N)} \mathbb{E}\left[\exp\left(\lambda(N) \sup_{i\in\mathbb{N}} \left| \int_0^T \sum_{j=1}^{\infty} \mathrm{e}_{ij}^{-a^N s} \rho_{jk}^{N,\mathrm{P}} \, \mathrm{d}M_k(s) \right| \right) \right] \\
&\quad = \Gamma(N) \sup_{i=1,\dots,\Gamma(N)} \exp\left(-\frac{\epsilon\lambda(N)\gamma(N)}{v_a \mathrm{e}^{v_a T}} \right) \\
&\qquad \times \prod_{k=1}^{\gamma(N)} \mathbb{E}\left[\exp\left(\frac{\lambda(N)}{\gamma(N)} \sup_{i\in\mathbb{N}} \left| \int_0^T \sum_{j=1}^{\infty} \mathrm{e}_{ij}^{-a^N s} \gamma(N) \rho_{jk}^{N,\mathrm{P}} \, \mathrm{d}M_k(s) \right| \right) \right].
\end{aligned}$$

Now recall from [119, Thm. 2.7(iv)] that the stochastic integral in the last line has an infinitely divisible distribution. Moreover, the larger the integrand, the larger the exponential moment is. Since the integrand above is uniformly bounded in i and k by our hypotheses, the stochastic integral above can be replaced by some constant times $M_k(T)$ for the further estimation. Therefore, the remaining calculation can be completed as in the case $\iota = 1$. For the second term in (5.5.16) the reasoning is the same, except that the factor $\gamma(N)$ is now due to the domain $(\hat{t}, t]$ of the stochastic integral. Observe at this point that $M_k(t) - M_k(\hat{t})$ has the same distribution as $M_k(t - \hat{t})$ and that $|t - \hat{t}| \leq 1/\gamma(N)$. Again, we do not carry out the details. □

Lemma 5.5.5. *For each $\iota = 1, 2, 3$ the processes $(\hat{Y}^{N,\iota} \colon N \in \mathbb{N})$ form an* exponentially tight *sequence in $(D_T^d, \mathcal{D}_T^d, U)$, that is, for every $L \in \mathbb{R}_+$ there exists a compact subset K^L of D_T^d (with respect to the uniform topology U) such that*

$$\limsup_{N\to\infty} \frac{1}{\gamma(N)} \log \mathbb{P}[\hat{Y}^{N,\iota} \notin K^L] \leq -L.$$

Proof. We first consider $\iota = 1$. We will adapt the idea of [51, Lemma 4.1] to our setting. As shown in part (I) of the proof there, it suffices to show that for every $a, \epsilon \in (0, \infty)$ there exist a compact set $H \subseteq D_T^d$, some $C \in (0, \infty)$ and $n \in \mathbb{N}$ such that for all $N \geq n$

$$\mathbb{P}[d(\hat{Y}^{N,1}, H) > \epsilon] \leq C\mathrm{e}^{-\gamma(N)a}, \tag{5.5.17}$$

where $d(f, H) := \inf\{\sup_{t\in[0,T]} \sup_{i=1,\dots,d} |f_i(t) - g_i(t)| \colon g \in H\}$ for $f \in D_T^d$. In order to prove (5.5.17), we first define for $n \in \mathbb{N}$ and $A \subseteq \mathbb{R}^d$ the set $H_n(A)$ as the collection of all $f \in D_T^d$ of the form

$$f = \sum_{\kappa=1}^{[\gamma(n)T]-1} x_\kappa \mathbb{1}_{[\frac{\kappa}{\gamma(n)}, \frac{\kappa+1}{\gamma(n)})} + x_{[\gamma(n)T]} \mathbb{1}_{[\frac{\kappa}{\gamma(n)}, T]}, \quad x_1, \dots, x_{[\gamma(n)T]} \in A.$$

It follows from [51, Eq. (4.3)] that for $N \geq n$, $A \subseteq \mathbb{R}^d$ and $f \in H_N(A)$ we have

$$\begin{aligned} &d(f, H_n(A)) \\ &\leq \sup_{\kappa=0,\dots,[\gamma(n)T]-1} \sup_{\lambda\in[1, \frac{\gamma(N)}{\gamma(n)}+1)} \sup_{i=1,\dots,d} \left| f_i\left(\frac{[\frac{\gamma(N)\kappa}{\gamma(n)}] + \lambda}{\gamma(N)} \wedge T\right) - f_i\left(\frac{[\frac{\gamma(N)\kappa}{\gamma(n)}]}{\gamma(N)}\right)\right|. \end{aligned} \tag{5.5.18}$$

Next, define $K := [-1, 1]^d$. Then for every $\beta \in (0, \infty)$ and $N \geq n$ we have

$$\begin{aligned} \mathbb{P}[d(\hat{Y}^{N,1}, H_n(\beta K)) > \epsilon] &\leq \mathbb{P}[\hat{Y}^{N,1} \notin H_N(\beta K)] \\ &\quad + \mathbb{P}[\hat{Y}^{N,1} \in H_N(\beta K), d(\hat{Y}^{N,1}, H_n(\beta K)) > \epsilon]. \end{aligned} \tag{5.5.19}$$

The first probability is bounded as follows:

$$\begin{aligned}
&\mathbb{P}[\hat{Y}^{N,1} \notin H_N(\beta K)] \\
&\quad = \mathbb{P}\Bigg[\sup_{i=1,\dots,d} \sup_{\kappa=1,\dots,[\gamma(N)T]} \Bigg| \int_0^{\frac{\kappa}{\gamma(N)}} \int_0^s \sum_{j,k=1}^{\infty} \mathrm{e}_{ij}^{a^N(\kappa/\gamma(N)-s)} a_{jk}^{N,\mathrm{P}} \mathrm{e}^{a_{kk}^{N,\mathrm{C}}(s-r)} \\
&\qquad \times \sum_{l\neq k} a_{kl}^{N,\mathrm{C}} (\bar{X}_l^N(r) - \mathbb{E}[\bar{X}_l^N(r)])\,\mathrm{d}r\,\mathrm{d}s \Bigg| > \beta \Bigg] \\
&\quad \leq \mathbb{P}\left[\sup_{t\in[0,T]} \sup_{i=1,\dots,d} \sum_{j=1}^{\infty} \sum_{k\neq j} \left| a_{ij}^{N,\mathrm{P}} a_{jk}^{N,\mathrm{C}} (\bar{X}_k^N(t) - \mathbb{E}[\bar{X}_k^N(t)]) \right| > \frac{\beta}{(\mathrm{e}^{v_a T} T)^2} \right] =: p'(N).
\end{aligned}$$

By the same arguments as in Lemma 5.5.4, one obtains (again assuming $v_a T < 1$ without loss of generality)

$$p'(N) \leq 2^{\gamma(N)} \exp\left(-\frac{\beta\lambda(N)(1-v_a T)}{q_1(N)(\mathrm{e}^{v_a T} T)^2} \right) \exp\left(T\gamma(N) \Psi_0 \left(\frac{\lambda(N)}{\gamma(N)} \sup_{l,m\in\mathbb{N}} |\rho_{lm}^{N,\mathrm{C}}| \right) \right).$$

We choose $\lambda(N) := \gamma(N)$ this time. Then we can make $\log(p'(N))/\gamma(N)$ arbitrarily small uniformly for large N by varying the value of β.

For the second step of the proof of (5.5.17) we conclude from (5.5.18) that

$$\begin{aligned}
&\mathbb{P}[\hat{Y}^{N,1} \in H_N(\beta K), d(\hat{Y}^{N,1}, H_n(\beta K)) > \epsilon] \\
&\quad \leq \mathbb{P}\Bigg[\sup_{\kappa=0,\dots,[\gamma(n)T]-1} \sup_{\lambda\in\left[1,\frac{\gamma(N)}{\gamma(n)}+1\right)} \sup_{i=1,\dots,d} \\
&\qquad \left| \hat{Y}_i^{N,1} \left(\frac{[\frac{\gamma(N)\kappa}{\gamma(n)}] + \lambda}{\gamma(N)} \wedge T \right) - \hat{Y}_i^{N,1} \left(\frac{[\frac{\gamma(N)\kappa}{\gamma(n)}]}{\gamma(N)} \right) \right| > \epsilon \Bigg]. \qquad (5.5.20)
\end{aligned}$$

For the further procedure, we split the difference in the last line into two terms in the same way as in (5.5.15). We only treat the corresponding first term. As before, the other one can be estimated similarly. Introducing the notation $t_{\kappa,\lambda}^{N,n}$ (resp. $t_{\kappa}^{N,n}$) for the time point in the first (resp. second) parenthesis of (5.5.20), and observing that $0 \leq \widehat{t_{\kappa,\lambda}^{N,n}} - t_{\kappa}^{N,n} \leq 2/\gamma(N) + 1/\gamma(n)$, we obtain

$$
\begin{aligned}
&\mathbb{P}\Bigg[\sup_{\kappa=0,\dots,[\gamma(n)T]-1}\sup_{\lambda\in\left[1,\frac{\gamma(N)}{\gamma(n)}+1\right)}\sup_{i=1,\dots,d}\Bigg|\Bigg(\left(\mathrm{e}^{a^N\widehat{t^{N,n}_{\kappa,\lambda}}}-\mathrm{e}^{a^N t^{N,n}_{\kappa}}\right)\\
&\times\int_0^t\int_0^s \mathrm{e}^{-a^N s}a^{N,\mathrm{P}}\mathrm{e}^{\mathrm{diag}(a^{N,\mathrm{C}})(s-r)}(a^{N,\mathrm{C}})^{\times}(\bar{X}^N(r)-\mathbb{E}[\bar{X}^N(r)])\,\mathrm{d}r\,\mathrm{d}s\Bigg)_i\Bigg|>\epsilon\Bigg]\\
&\leq\mathbb{P}\left[\sup_{t\in[0,T]}\sup_{i=1,\dots,d}\left|\sum_{j=1}^{\infty}\sum_{k\neq j}a^{N,\mathrm{P}}_{ij}a^{N,\mathrm{C}}_{jk}(\bar{X}^N_k(t)-\mathbb{E}[\bar{X}^N_k(t)])\right|>\frac{\epsilon}{v_a\mathrm{e}^{3v_aT}T^2\left(\frac{2}{\gamma(N)}+\frac{1}{\gamma(n)}\right)}\right]\\
&\leq 2^{\gamma(N)}\exp\left(-\frac{\epsilon\lambda(N)(1-v_aT)}{q_1(N)v_a\mathrm{e}^{3v_aT}T^2\left(\frac{2}{\gamma(N)}+\frac{1}{\gamma(n)}\right)}+T\gamma(N)\Psi_0\left(\frac{\lambda(N)}{\gamma(N)}\sup_{l,m\in\mathbb{N}}|\rho^{N,\mathrm{C}}_{lm}|\right)\right),
\end{aligned}
$$

where the last line follows in similar fashion as before. With $\lambda(N):=\gamma(N)$ we can make, by taking n large enough, the logarithm of the last term divided by $\gamma(N)$ arbitrarily small for $N\geq n$. This finishes the proof for $\iota=1$. The case $\iota=2$ is analogous, while for $\iota=3$ the line of argument remains the same in principle, with slight changes to account for the discretization of Lévy processes, cf. the proofs of Lemma 5.5.4 and [51, Lemma 4.1]. □

Lemma 5.5.6. *The process $(\hat{Y}^N_i\colon i=1,\dots,d)$ satisfies a large deviation principle in $(D^d_T,\mathcal{D}^d_T,U)$ with a good convex rate function $I_d\colon D^d_T\to[0,\infty]$ such that $I_d(x)<\infty$ implies $x\in AC^d_T$. Moreover, we have $I_d(0)=0$ and this minimum is unique.*

Proof. We apply the abstract Gärtner-Ellis theorem of [51, Thms. 2.1 and 2.4] to $\hat{Y}^N$ and prove the following steps.

(1) The laws of $\hat{Y}^N$, $N\in\mathbb{N}$, are exponentially tight in $(D^d_T,\mathcal{D}^d_T,U)$.

(2) For all $\theta\in M^d_T$ the limit $\Lambda(\theta)=\lim_{N\to\infty}(\gamma(N))^{-1}\Lambda_N(\gamma(N)\theta)$ exists, where

$$\Lambda_N(\theta):=\log\mathbb{E}\left[\exp\left(\sum_{i=1}^d\int_0^T\hat{Y}^N_i(t)\,\theta_i(\mathrm{d}t)\right)\right].$$

(3) The mapping Λ is C^d_T-Gâteaux differentiable, in the sense that for all $\theta\in M^d_T$ there exists $x^\theta\in C^d_T$ such that for all $\eta\in M^d_T$

$$\delta\Lambda(\theta;\eta):=\lim_{\epsilon\to0}\frac{\Lambda(\theta+\epsilon\eta)-\Lambda(\theta)}{\epsilon}=\sum_{i=1}^d\int_0^T x^\theta_i(t)\,\eta_i(\mathrm{d}t).\tag{5.5.21}$$

Part of the claim is that the limit in (5.5.21) exists. Furthermore, we have that $\Lambda(0;\eta) = 0$ for all $\eta \in M_T^d$.

(4) We have $\{x \in D_T^d \colon \Lambda^*(x) < \infty\} \subseteq AC_T^d$, where

$$\Lambda^*(x) := \sup_{\theta \in M_T^d} \left(\sum_{i=1}^d \int_0^T x_i(t)\, \theta_i(\mathrm{d}t) - \Lambda(\theta) \right), \quad x \in D_T^d.$$

(5) For every $\alpha \in \mathbb{R}_+$ the set $\{x \in D_T^d \colon \Lambda^*(x) \le \alpha\}$ is compact in $(D_T^d, \mathcal{D}_T^d, U)$.

Part of the Gärtner-Ellis theorem is that the rate function I_d is given by Λ^*, the convex conjugate or Fenchel-Legendre transform of Λ. Since Λ is a convex function in θ satisfying (3), the conjugate Λ^{**} of Λ^* is again Λ, see [121, Thm. 12]. Thus, by the first corollary to Theorem 1 in [8], we have $I_d(0) = \Lambda^*(0) = 0$ and this minimum is unique.

Let us now prove (1)–(5) above. Part (1) has been proved in Lemma 5.5.5. For (2) we first compute Λ. For all $\theta \in M_T^d$ we have (recall that $\hat{t} := [\gamma(N)t]/\gamma(N)]$)

$$\begin{aligned}
&\Lambda_N(\gamma(N)\theta) \\
&\quad = \log \mathbb{E}\Bigg[\exp\Bigg(\sum_{i=1}^d \int_0^T \Bigg(\gamma(N) \sum_{j,k=1}^\infty \int_0^{\hat{t}} \mathrm{e}_{ij}^{a^N(\hat{t}-s)} \rho_{jk}^{N,\mathrm{P}}\, \mathrm{d}M_k(s) \\
&\qquad + \int_0^{\hat{t}} \int_0^s \gamma(N) \sum_{j,k,l,m=1}^\infty \mathrm{e}_{ij}^{a^N(\hat{t}-s)} a_{jk}^{N,\mathrm{P}} \mathrm{e}_{kl}^{a^{N,\mathrm{C}}(s-r)} \rho_{lm}^{N,\mathrm{C}}\, \mathrm{d}M_m(r)\, \mathrm{d}s \Bigg) \theta_i(\mathrm{d}t) \Bigg) \Bigg] \\
&\quad = \sum_{m=1}^{\gamma(N)} \log \mathbb{E}\Bigg[\exp\Bigg(\sum_{i=1}^d \int_0^T \Bigg(\int_0^{\hat{t}} \gamma(N) \sum_{j=1}^\infty \mathrm{e}_{ij}^{a^N(\hat{t}-s)} \rho_{jm}^{N,\mathrm{P}}\, \mathrm{d}M_m(s) \\
&\qquad + \int_0^{\hat{t}} \int_0^s \sum_{j,k,l=1}^\infty \gamma(N) \mathrm{e}_{ij}^{a^N(\hat{t}-s)} a_{jk}^{N,\mathrm{P}} \mathrm{e}_{kl}^{a^{N,\mathrm{C}}(s-r)} \rho_{lm}^{N,\mathrm{C}}\, \mathrm{d}M_m(r)\, \mathrm{d}s \Bigg) \theta_i(\mathrm{d}t) \Bigg) \Bigg] \qquad (5.5.22)
\end{aligned}$$

by the independence of the processes M_m. By a stochastic Fubini argument (see [118, Thm. IV.65]), the term within the exponential in the previous line can also be written as ($\check{s}$ denotes the smallest multiple of $\gamma(N)$ that is larger or equal to s)

$$\begin{aligned}
&\int_0^{\hat{T}} \Bigg(\int_r^{\hat{T}} \int_{\check{s}}^T \sum_{i=1}^d G_{im}^N(\hat{t}-s, s-r)\, \theta_i(\mathrm{d}t)\, \mathrm{d}s \\
&\qquad + \int_{\check{r}}^T \sum_{i=1}^d R_{im}^N(\hat{t}-r)\, \theta_i(\mathrm{d}t) \Bigg) M_m(\mathrm{d}r), \qquad (5.5.23)
\end{aligned}$$

and has an infinitely divisible distribution such that its logarithmic Laplace exponent in (5.5.22) is explicitly known. Denoting the parenthesis in (5.5.23) by $H_m^N(\theta, r)$, it is given by $\int_0^{\hat{T}} \Psi_m(H_m^N(\theta, r))\,\mathrm{d}r$. We claim that this term converges uniformly in m to $\int_0^T \Psi_m(H_m(\theta, r))\,\mathrm{d}r$. Indeed, by the dominating property of M_0, the claim follows as soon as we can prove that $H_m^N(\theta, r) \to H_m(\theta, r)$ as $N \to \infty$, uniformly in $m \in \mathbb{N}$ and $r \in [0, T]$. This in turn follows from

$$
\begin{aligned}
&|H_m(\theta, r) - H_m^N(\theta, r)| \\
&\quad \le \left| \int_{\hat{T}}^T \int_s^T \sum_{i=1}^d G_{im}(t-s, s-r)\,\theta_i(\mathrm{d}t)\,\mathrm{d}s \right| \\
&\qquad + \left| \int_r^{\hat{T}} \int_s^{\check{s}} \sum_{i=1}^d G_{im}(t-s, s-r)\,\theta_i(\mathrm{d}t)\,\mathrm{d}s \right| \\
&\qquad + \left| \int_r^{\hat{T}} \int_{\check{s}}^T \sum_{i=1}^d (G_{im}(t-s, s-r) - G_{im}^N(t-s, s-r))\,\theta_i(\mathrm{d}t)\,\mathrm{d}s \right| \\
&\qquad + \left| \int_r^{\hat{T}} \int_{\check{s}}^T \sum_{i=1}^d (G_{im}^N(t-s, s-r) - G_{im}^N(\hat{t}-s, s-r))\,\theta_i(\mathrm{d}t)\,\mathrm{d}s \right| \\
&\qquad + \left| \int_r^{\check{r}} \sum_{i=1}^d R_{im}(t-r)\,\theta_i(\mathrm{d}t) \right| + \left| \int_{\check{r}}^T \sum_{i=1}^d (R_{im}(t-r) - R_{im}^N(t-r))\,\theta_i(\mathrm{d}t) \right| \\
&\qquad + \left| \int_{\check{r}}^T \sum_{i=1}^d (R_{im}^N(t-r) - R_{im}^N(\hat{t}-r))\,\theta_i(\mathrm{d}t) \right| \\
&\quad \le d \sup_{i,m\in\mathbb{N}} \sup_{s,t\in[0,T]} |G_{im}(t,s)| \left(\frac{1}{\gamma(N)} \sup_{i=1,\dots,d} |\theta_i|([0,T]) + \sup_{i=1,\dots,d} \int_r^{\hat{T}} |\theta_i|([s,\check{s}))\,\mathrm{d}s \right) \\
&\qquad + d \sup_{i=1,\dots,d} |\theta_i|([0,T]) \Bigg(T \sup_{i,m\in\mathbb{N}} \sup_{s,t\in[0,T]} |G_{im}(t,s) - G_{im}^N(t,s)| \\
&\qquad + \frac{v_a T}{\gamma(N)} \sup_{N,i,m\in\mathbb{N}} \sup_{s,t\in[0,T]} |G_{im}^N(t,s)| + \frac{1}{\gamma(N)} \sup_{i,j\in\mathbb{N}} \sup_{t\in[0,T]} |R_{ij}(t)| \\
&\qquad + \sup_{i,j\in\mathbb{N}} \sup_{t\in[0,T]} |R_{ij}(t) - R_{ij}^N(t)| + \frac{v_a}{\gamma(N)} \sup_{N,i,j\in\mathbb{N}} \sup_{t\in[0,T]} |R_{ij}^N(t)| \Bigg),
\end{aligned}
$$

where all terms converge to 0 by hypothesis independently of m and r. For the second summand one has to notice that the integral term equals $\int_r^{\hat{T}} \int_{\hat{t}}^t 1\,\mathrm{d}s\,|\theta_i|(\mathrm{d}t)$ and thus converges to 0 uniformly in i and r with rate $1/\gamma(N)$. Since the value of Cesàro sums remains unchanged under uniform approximations, it follows from

assumption (A8) of Theorem 5.4.1 that

$$\Lambda(\theta) = \lim_{N\to\infty} \frac{1}{\gamma(N)} \sum_{m=1}^{\gamma(N)} \int_0^T \Psi_m(H_m^N(\theta, r))\,\mathrm{d}r = \lim_{N\to\infty} \frac{1}{\gamma(N)} \sum_{m=1}^{\gamma(N)} \int_0^T \Psi_m(H_m(\theta, r))\,\mathrm{d}r.$$

Next, we prove the C_T^d-Gâteaux differentiability of Λ. First, regarding the existence of $\delta\Lambda(\theta;\eta)$ in (5.5.21), we note that the mappings $M_T^d \to L^\infty(\mathbb{N}\times[0,T])$, $\theta \mapsto (H_m(\theta, r)\colon m \in \mathbb{N},\ r \in [0,T])$, are continuous linear operators and therefore Fréchet differentiable, which is stronger than Gâteaux differentiability. Together with the fact that Ψ_m is differentiable with locally bounded derivative, and $c_m \le c_0$ and $\nu_m \le \nu_0$ for all $m \in \mathbb{N}$, this implies that for every θ and η

$$\epsilon^{-1} \int_0^T \left(\Psi_m(H_m(\theta+\epsilon\eta, r)) - \Psi_m(H_m(\theta, r))\right)\,\mathrm{d}r$$

converges uniformly in $m \in \mathbb{N}$ as $\epsilon \to 0$. This in turn proves the Gâteaux differentiability of Λ. Moreover, it enables us to compute the derivative explicitly. Using the chain rule for Fréchet derivatives, we obtain

$$\begin{aligned}
\delta\Lambda(\theta;\eta) &= \lim_{N\to\infty} \sum_{m=1}^{\gamma(N)} \int_0^T \lim_{\epsilon\to 0} \frac{\Psi_m(H_m(\theta+\epsilon\eta, r)) - \Psi_m(H_m(\theta, r))}{\epsilon}\,\mathrm{d}r \\
&= \lim_{N\to\infty} \sum_{m=1}^{\gamma(N)} \int_0^T \Psi_m'(H_m(\theta, r)) H_m(\eta, r)\,\mathrm{d}r \\
&= \lim_{N\to\infty} \sum_{m=1}^{\gamma(N)} \int_0^T \Psi_m'(H_m(\theta, r)) \Bigg(\int_r^T \int_s^T \sum_{i=1}^d G_{im}(t-s, s-r)\,\eta_i(\mathrm{d}t)\,\mathrm{d}s \\
&\quad + \int_r^T \sum_{i=1}^d R_{im}(t-r)\,\eta_i(\mathrm{d}t) \Bigg)\,\mathrm{d}r \\
&= \lim_{N\to\infty} \sum_{m=1}^{\gamma(N)} \sum_{i=1}^d \int_0^T \Bigg(\int_0^t \int_0^s \Psi_m'(H_m(\theta, r)) G_{im}(t-s, s-r)\,\mathrm{d}r\,\mathrm{d}s \\
&\quad + \int_0^t R_{im}(t-r)\,\mathrm{d}r \Bigg)\,\eta_i(\mathrm{d}t) \\
&= \sum_{i=1}^d \int_0^T \int_0^t \lim_{N\to\infty} \sum_{m=1}^{\gamma(N)} \Bigg(\int_0^s G_{im}(t-s, s-r) \Psi_m'(H_m(\theta, r))\,\mathrm{d}r \\
&\quad + R_{im}(t-s) \Psi_m'(H_m(\theta, s)) \Bigg)\,\mathrm{d}s\,\eta_i(\mathrm{d}t),
\end{aligned}$$

where all interchanges of integration, summation and taking limits are justified by dominated convergence. From the last line we deduce the existence of $x^\theta \in C_T^d$ satisfying (5.5.21). Since $H_m(0,r) = 0$ and $\Psi_m'(0) = 0$, we have $\delta\Lambda(0;\eta) = 0$ identically.

Next, we demonstrate (4), namely that Λ^* only assumes finite values on the set AC_T^d, that is, $\Lambda^*(x) < \infty$ implies that for every $\epsilon \in (0,\infty)$ there exists $\delta \in (0,\infty)$ such that whenever $n \in \mathbb{N}$, $0 \le a_1 < b_1 \le \ldots \le a_n < b_n \le T$ and $\sum_{j=1}^n (b_j - a_j) < \delta$, we have $\sum_{i=1}^d \sum_{j=1}^n |x_i(b_j) - x_i(a_j)| < \epsilon$. In order to do so, we follow the strategy of proof in [51, Thm. 3.1]. We consider $\theta_i := \sum_{j=1}^n \xi_i^j (\delta_{b_j} - \delta_{a_j})$ where $\xi_i^j \in \mathbb{R}^d$ is arbitrary. Then we evidently have $\theta_i((r,T]) = \sum_{j=1}^n \xi_i^j \mathbb{1}_{[a_j,b_j)}(r)$. Using the notation $C_T := T \sup_{s,t\in[0,T]} \sup_{i,m\in\mathbb{N}} |G_{im}(t,s)| + \sup_{t\in[0,T]} \sup_{i,m\in\mathbb{N}} |R_{im}(t)|$, it follows that

$$\begin{aligned}
\Lambda(\theta) &\le \sup_{m\in\mathbb{N}} \int_0^T \Psi_m(H_m(\theta,r)\,\mathrm{d}r \le \int_0^T \Psi_0\left(C_T \sum_{i=1}^d \theta_i((r,T])\right)\,\mathrm{d}r \\
&= \int_0^T \sum_{j=1}^n \Psi_0\left(C_T \sum_{i=1}^d \xi_i^j\right) \mathbb{1}_{[a_j,b_j)}(r)\,\mathrm{d}r \le \sup_{j=1,\ldots,n} \Psi_0\left(C_T \sum_{i=1}^d |\xi_i^j|\right) \sum_{j=1}^n (b_j - a_j) \\
&=: C(T, \|\xi^1\|_1, \ldots, \|\xi^n\|_1) \sum_{j=1}^n (b_j - a_j),
\end{aligned}$$

where $\|\xi^j\|_1 := \sum_{i=1}^d |\xi_i^j|$. As a consequence, we deduce from the definition of Λ^* that for all $\tau \in (0,\infty)$ and $\|\xi^j\|_1 \le \tau$

$$\sum_{i=1}^d \sum_{j=1}^n \xi_i^j (x_i(b_j) - x_i(a_j)) \le C(T,\tau,\ldots,\tau) \sum_{j=1}^n (b_j - a_j) + \Lambda^*(x).$$

Taking ξ_i^j as the τ times the sign of $x_i(b_j) - x_i(a_j)$, it follows that

$$\sum_{i=1}^d \sum_{j=1}^n |x_i(b_j) - x_i(a_j)| \le \tau^{-1} C(T,\tau,\ldots,\tau) \sum_{j=1}^n (b_j - a_j) + \tau^{-1}\Lambda^*(x). \tag{5.5.24}$$

If $\Lambda^*(x) < \infty$, we can now choose τ first and then δ to make the left-hand side arbitrarily small.

It only remains to prove (5), the compactness of the level sets of Λ^*. By step (4) and the lower semicontinuity of Λ^*, its level sets are closed subsets of AC_T^d. Thus, the Arzelà-Ascoli theorem provides a compactness criterion. First, observe that for all $t \in [0,T]$ we have for $x \in AC_T^d$ with $\Lambda^*(x) \le \alpha$ that

$$\sum_{i=1}^d |x_i(t)| = \sup_{\theta\in\Theta_t} \sum_{i=1}^d \int_0^T x_i(t)\,\theta_i(\mathrm{d}t) \le \alpha + \sup_{\theta\in\Theta_t} \Lambda(\theta) < \infty,$$

where Θ_t is the finite collection of all θ for which each coordinate is either δ_t or $-\delta_t$. Second, for the proof of the uniform equicontinuity of the functions $x \in AC_T^d$ with $\Lambda^*(x) \leq \alpha$, we recall from (5.5.24) that

$$\sum_{i=1}^{d} |x_i(t) - x_i(s)| \leq \tau^{-1} C(T, \tau, \ldots, \tau)(t-s) + \tau^{-1}\alpha,$$

which converges to 0 independently of x when $s \uparrow t$ and $\tau \to \infty$. □

Bibliography

[1] S.M. Ahn and S.-Y. Ha. Stochastic flocking dynamics of the Cucker-Smale model with multiplicative white noises. *J. Math. Phys.*, 51(10):103301, 2010.

[2] Y. Aït-Sahalia and J. Jacod. *High-Frequency Financial Econometrics*. Princeton University Press, Princeton, 2014.

[3] S. Albeverio, J.-L. Wu, and T.-S. Zhang. Parabolic SPDEs driven by Poisson white noise. *Stoch. Process. Appl.*, 74(1):21–36, 1998.

[4] A. Andresen, F.E. Benth, S. Koekebakker, and V. Zakamulin. The CARMA interest rate model. *Int. J. Theor. Appl. Finan.*, 17(2), 2014. 27 pages.

[5] D. Applebaum and J.-L. Wu. Stochastic partial differential equations driven by Lévy space–time white noise. *Random Oper. Stoch. Equ.*, 8(3):245–259, 2000.

[6] S. Asmussen. *Applied Probability and Queues*. Springer, New York, 2nd edition, 2003.

[7] S. Asmussen and J. Rosiński. Approximations of small jumps of Lévy processes with a view towards simulation. *J. Appl. Probab.*, 38(2):482–493, 2001.

[8] E. Asplund and R.T. Rockafellar. Gradient of convex functions. *Trans. Amer. Math. Soc.*, 139:443–467, 1969.

[9] R.M. Balan. SPDEs with α-stable Lévy noise: a random field approach. *Int. J. Stoch. Anal.*, 2014, 2014. Article ID 793275, 22 pages.

[10] A.-L. Barabási and R. Albert. Emergence of scaling in random networks. *Science*, 286(5439):509–512, 1999.

[11] O.E. Barndorff-Nielsen. Superposition of Ornstein-Uhlenbeck type processes. *Theory Probab. Appl.*, 45(2):175–194, 2001.

[12] O.E. Barndorff-Nielsen and J. Schmiegel. Lévy-based spatial–temporal modelling, with applications to turbulence. *Russ. Math. Surv.*, 59(1):65–90, 2004.

[13] O.E. Barndorff-Nielsen and J. Schmiegel. Ambit processes; with applications to turbulence and tumour growth. In F.E. Benth, editor, *Stochastic Analysis and Applications*, pages 93–124. Springer, Berlin, 2007.

[14] O.E. Barndorff-Nielsen and N. Shephard. Non-Gaussian Ornstein-Uhlenbeck based models and some of their uses in financial economics. *J. R. Stat. Soc. Ser. B Stat. Methodol.*, 63(2):167–241, 2001.

[15] O.E. Barndorff-Nielsen and N. Shephard. Econometric analysis of realized volatility and its use in estimating stochastic volatility models. *J. R. Stat. Soc. Ser. B Stat. Methodol.*, 64(2):253–280, 2012.

[16] O.E. Barndorff-Nielsen and R. Stelzer. Multivariate supOU processes. *Ann. Appl. Probab.*, 21(1):140–182, 2011.

[17] O.E. Barndorff-Nielsen and A.E.D. Veraart. Stochastic volatility of volatility and variance risk premia. *J. Financ. Econom.*, 11(1):1–46, 2012.

[18] O.E. Barndorff-Nielsen, F.E. Benth, and A.E.D. Veraart. Ambit processes and stochastic partial differential equations. In G. Di Nunno and B. Øksendal, editors, *Advanced Mathematical Methods for Finance*, pages 35–74. Springer, Berlin, 2011.

[19] O.E. Barndorff-Nielsen, F.E. Benth, and A.E.D. Veraart. Modelling energy spot prices by volatility modulated Lévy-driven Volterra processes. *Bernoulli*, 19(3):803–845, 2013.

[20] A. Barth and A. Lang. Simulation of stochastic partial differential equations using finite element methods. *Stochastics*, 84(2–3):217–231, 2012.

[21] A. Basse-O'Connor, S.-E. Graversen, and J. Pedersen. Multiparameter processes with stationary increments: spectral representation and integration. *Electron. J. Probab.*, 17(74), 2012. 21 pages.

[22] A. Basse-O'Connor, S.-E. Graversen, and J. Pedersen. Stochastic integration on the real line. *Theory Probab. Appl.*, 58(2):193–215, 2014.

[23] S. Battiston, D. Delli Gatti, M. Gallegati, B. Greenwald, and J.E. Stiglitz. Liaisons dangereuses: Increasing connectivity, risk sharing, and systemic risk. *J. Econ. Dyn. Control*, 36(8):1121–1141, 2012.

[24] A. Behme. *Generalized Ornstein-Uhlenbeck Processes and Extensions*. PhD thesis, Technische Universität Braunschweig, 2011.

[25] A. Behme. Distributional properties of solutions of $\mathrm{d}V_t = V_{t-}\mathrm{d}U_t + \mathrm{d}L_t$ with Lévy noise. *Adv. Appl. Probab.*, 43(3):688–711, 2011.

[26] A. Behme, C. Chong, and C. Klüppelberg. Superposition of COGARCH processes. *Stoch. Process. Appl.*, 125(4):1426–1469, 2015.

[27] K. Bichteler. *Stochastic Integration with Jumps*. Cambridge University Press, Cambridge, 2002.

[28] K. Bichteler and J. Jacod. Random measures and stochastic integration. In G. Kallianpur, editor, *Theory and Application of Random Fields*, pages 1–18. Springer, Berlin, 1983.

[29] P. Billingsley. *Probability and Measure*. Wiley, New York, 3rd edition, 1995.

[30] T. Bollerslev, R.F. Engle, and D.B. Nelson. ARCH models. In R.F. Engle and D. McFadden, editors, *Handbook of Econometrics*, volume 4, pages 2959–3038. North-Holland, Amsterdam, 1994.

[31] F. Bolley, J.A. Cañizo, and J.A. Carrillo. Stochastic mean-field limit: non-Lipschitz forces and swarming. *Math. Model. Methods Appl. Sci.*, 21(11):2179–2210, 2011.

[32] B. Bollobás, C. Borgs, J. Chayes, and O. Riordan. Directed scale-free graphs. In *Proceedings of the Fourteenth Annual ACM-SIAM Symposium on Discrete Algorithms (Baltimore, MD, 2003)*, pages 132–139, New York, 2003. ACM.

[33] J.-P. Bouchaud and M. Mézard. Wealth condensation in a simple model of economy. *Phys. A*, 282(3–4):536–545, 2000.

[34] P.J. Brockwell. Lévy-driven CARMA processes. *Ann. Inst. Stat. Math.*, 53 (1):113–124, 2001.

[35] P.J. Brockwell, E. Chadraa, and A. Lindner. Continuous-time GARCH processes. *Ann. Appl. Probab.*, 16(2):790–826, 2006.

[36] A. Budhiraja and R. Wu. Some fluctuation results for weakly interacting multitype particle systems. Preprint under arXiv:1505.00856 [math.PR], 2015.

[37] S. Cambanis, J.P. Nolan, and J. Rosiński. On the oscillation of infinitely divisible and some other processes. *Stoch. Process. Appl.*, 35(1):87–97, 1990.

[38] R. Carmona and X. Zhu. A probabilistic approach to mean field games with major and minor players. *Ann. Appl. Probab.*, 2015. Forthcoming, preprint available on arXiv:1409.7141 [math.PR].

[39] D.K. Chang and M.M. Rao. Bimeasures and sampling theorems for weakly harmonizable processes. *Stoch. Anal. Appl.*, 1(1):21–55, 1983.

[40] B. Chen. Simulation of stochastic partial differential equations. Master's thesis, Technische Universität München, 2014. Available at https://mediatum.ub.tum.de/node?id=1225387.

[41] B. Chen, C. Chong, and C. Klüppelberg. Simulation of stochastic Volterra equations driven by space–time Lévy noise. In *A Fascinating Journey Through Probability, Statistics and Applications: In Honour of Ole E. Barndorff-Nielsen's 80th Birthday*. Springer, 2015. Forthcoming, preprint under arXiv:1501.01645 [math.PR].

[42] A.S. Cherny and A.N. Shiryaev. On stochastic integrals up to infinity and predictable criteria for integrability. In M. Émery, M. Ledoux, and M. Yor, editors, *Séminaire de Probabilités XXXVIII*, pages 49–66. Springer, Berlin, 2005.

[43] C. Chong. Lévy-driven Volterra equations in space and time. Preprint under arXiv:1407.8092 [math.PR], 2015.

[44] C. Chong and C. Klüppelberg. Integrability conditions for space–time stochastic integrals: Theory and applications. *Bernoulli*, 21(4):2190–2216, 2015.

[45] C. Chong and C. Klüppelberg. Partial mean field limit in heterogeneous networks. Preprint under arXiv:1507.01905 [math.PR], 2015.

[46] W.G. Couchran, J.-S. Lee, and J. Potthoff. Stochastic Volterra equations with singular kernels. *Stoch. Process. Appl.*, 56(2):337–349, 1995.

[47] L. Coutin and L. Decreusefond. Stochastic Volterra equations with singular kernels. In A.B. Cruzeiro and J.-C. Zambrini, editors, *Stochastic Analysis and Mathematical Physics*, pages 39–50. Birkhäuser, Boston, 2001.

[48] R.C. Dalang. Extending martingale measures stochastic integral with applications to spatially homogeneous S.P.D.E's. *Electron. J. Probab.*, 4(6), 1999. 24 pages.

[49] A.M. Davie and J.G. Gaines. Convergence of numerical schemes for the solution of parabolic stochastic partial differential equations. *Math. Comput.*, 70 (233):121–134, 2001.

[50] D.A. Dawson and J. Gärtner. Large deviations from the McKean-Vlasov limit for weakly interacting diffusions. *Stochastics*, 20(4):247–308, 1987.

[51] A. de Acosta. Large deviations for vector-valued Lévy processes. *Stoch. Process. Appl.*, 51(1):75–115, 1994.

[52] P. Degond, J.-G. Liu, and C. Ringhofer. Evolution of the distribution of wealth in an economic environment driven by local Nash equilibria. *J. Stat. Phys.*, 154(3):751–780, 2014.

[53] P. Del Moral. *Mean Field Simulation for Monte Carlo Integration*. CRC Press, Boca Raton, 2013.

[54] P. Del Moral, L. Kallel, and J. Rowe. Modelling genetics algorithms with interacting particle systems. *Rev. Mat. Teor. Apl.*, 8(2):19–77, 2001.

[55] A. Dembo and O. Zeitouni. *Large Deviations Techniques and Applications*. Springer, New York, 2nd edition, 2010.

[56] J.L. Doob. *Stochastic Processes*. Wiley, New York, 1953.

[57] T. Dunst, E. Hausenblas, and A. Prohl. Approximate Euler method for parabolic stochastic partial differential equations driven by space–time Lévy noise. *SIAM J. Numer. Anal.*, 50(6):2873–2896, 2012.

[58] P. Embrechts, C. Klüppelberg, and T. Mikosch. *Modelling Extremal Events for Insurance and Finance*. Springer, Berlin, 1997.

[59] M. Émery. Une topologie sur l'espace des semimartingales. In C. Dellacherie, P.A. Meyer, and M. Weil, editors, *Séminaire de Probabilités XIII*, pages 260–280. Springer, Berlin, 1979.

[60] V. Fasen. Extremes of Lévy driven mixed MA processes with convolution equivalent distributions. *Extremes*, 12(3):265–296, 2009.

[61] V. Fasen. Extremes of continuous-time processes. In T.G. Andersen, R.A. Davis, J.-P. Kreiss, and T. Mikosch, editors, *Handbook of Financial Time Series*, pages 653–667. Springer, Heidelberg, 2009.

[62] V. Fasen and C. Klüppelberg. Extremes of supOU processes. In F.E. Benth, G. Di Nunno, T. Lindstrøm, B. Øksendal, and T. Zhang, editors, *Stochastic Analysis and Applications: The Abel Symposium 2005*, pages 340–359. Springer, Heidelberg, 2007.

[63] W. Finnoff. Law of large numbers for a general system of stochastic differential equations with global interaction. *Stoch. Process. Appl.*, 46(1):153–182, 1993.

[64] W. Finnoff. Law of large numbers for a heterogeneous system of stochastic differential equations with strong local interaction and economic applications. *Ann. Appl. Probab.*, 4(2):494–528, 1994.

[65] G.B. Folland. *Introduction to Partial Differential Equations*. Princeton University Press, Princeton, 2nd edition, 1995.

[66] J.-P. Fouque and L.-H. Sun. Systemic risk illustrated. In J.-P. Fouque and J.A. Langsam, editors, *Handbook on Systemic Risk*, pages 444–452. Cambridge University Press, Cambridge, 2013.

[67] N. Fournier and E. Löcherbach. On a toy model of interacting neurons. *Ann. Inst. Henri Poincaré Probab. Stat.*, 2015. Forthcoming, preprint under arXiv:1410.3263 [math.PR].

[68] C. Francq and J.-M. Zakoian. *GARCH Models*. Wiley, Chichester, 2010.

[69] F. Fuchs and R. Stelzer. Mixing conditions for multivariate infinitely divisible processes with an application to mixed moving averages and the supOU stochastic volatility model. *ESAIM Probab. Stat.*, 17:455–471, 2013.

[70] J. Gärtner. On the McKean-Vlasov limit for interacting diffusions. *Math. Nachr.*, 137(1):197–248, 1988.

[71] K. Giesecke, K. Spiliopoulos, R.B. Sowers, and J. Sirignano. Large portfolio asymptotics for loss from default. *Math. Financ.*, 25(1):77–114, 2015.

[72] J.E. Griffin. Inference in infinite superpositions of non-Gaussian Ornstein-Uhlenbeck processes using Bayesian nonparametic methods. *J. Finan. Econom.*, 9(3):519–549, 2011.

[73] J.E. Griffin and M.F.J. Steel. Bayesian inference with stochastic volatility models using continuous superpositions of non-Gaussian Ornstein-Uhlenbeck processes. *Comput. Stat. Data Anal.*, 54(11):2594–2608, 2010.

[74] G. Gripenberg, S.-O. Londen, and O. Steffans. *Volterra Integral and Functional Equations*. Cambridge University Press, Cambridge, 1990.

[75] I. Gyöngy. Lattice approximations for stochastic quasi-linear parabolic partial differential equations driven by space–time white noise II. *Potential Anal.*, 11 (1):1–37, 1999.

[76] E. Hausenblas. Finite element approximation of stochastic partial differential equations driven by Poisson random measures of jump type. *SIAM J. Numer. Anal.*, 46(1):437–471, 2008.

[77] E. Hausenblas and I. Marchis. A numerical approximation of parabolic stochastic differential equations driven by a Poisson random measure. *BIT Numer. Math.*, 46(4):773–811, 2006.

[78] T. Ichinomiya. Bouchard-Mézard model on a random network. *Phys. Rev. E*, 86(3):036111, 2012.

[79] K. Itô. Stochastic integral. *Proc. Imp. Acad. Tokyo*, 20(8):519–524, 1944.

[80] J. Jacod and A.N. Shiryaev. *Limit Theorems for Stochastic Processes*. Springer, Berlin, 2nd edition, 2003.

[81] J. Jacod and V. Todorov. Do price and volatility jump together? *Ann. Appl. Probab.*, 20(4):1425–1469, 2010.

[82] J. Jacod, C. Klüppelberg, and G. Müller. Functional relationships between price and volatility jumps and its consequences for discretely observed data. *J. Appl. Probab.*, 49(4):901–914, 2012.

[83] E.B.V. Jensen, K.Ý. Jónsdóttir, J. Schmiegel, and O.E. Barndorff-Nielsen. Spatio–temporal modelling — with a view to biological growth. In B. Finkenstadt, L. Held, and V. Isham, editors, *Statistical methods for spatio-temporal systems*, pages 47–76. Chapman & Hall, Boca Raton, 2007.

[84] K.Ý. Jónsdóttir, A. Rønn-Nielsen, K. Mouridsen, and E.B.V. Jensen. Lévy-based modelling in brain imaging. *Scand. J. Stat.*, 40(3):511–529, 2013.

[85] O. Kallenberg. *Random Measures*. Akademie-Verlag, Berlin, 4th edition, 1986.

[86] O. Kallenberg. *Foundations of Modern Probability*. Springer, New York, 2nd edition, 2002.

[87] O. Kley, C. Klüppelberg, and L. Reichel. Systemic risk through contagion in a core–periphery structured banking network. Preprint under arXiv:1406.6575 [q-fin.RM], 2014.

[88] C. Klüppelberg, A. Lindner, and R. Maller. A continuous-time GARCH process driven by a Lévy process: stationarity and second-order behaviour. *J. Appl. Probab.*, 41(3):601–622, 2004.

[89] C. Klüppelberg, A. Lindner, and R. Maller. Continuous time volatility modelling: COGARCH versus Ornstein-Uhlenbeck models. In Y. Kabanov, R. Liptser, and J. Stoyanov, editors, *From Stochastic Calculus to Mathematical Finance. The Shiryaev Festschrift*, pages 393–419. Springer, Berlin, 2006.

[90] K. Knopp. *Theory and Application of Infinite Series*. Dover, New York, 1990.

[91] T.G. Kurtz and P.E. Protter. Weak convergence of stochastic integrals and differential equations II: Infinite dimensional case. In D. Talay and L. Tubaro, editors, *Probabilistic Models for Nonlinear Partial Differential Equations*, pages 197–285. Springer, Berlin, 1996.

[92] S. Kwapień and W.A. Woyczyński. *Random Series and Stochastic Integrals: Single and Multiple*. Birkhäuser, Boston, 1992.

[93] V.A. Lebedev. The Fubini theorem for stochastic integrals with respect to L^0-valued random measures depending on a parameter. *Theory Probab. Appl.*, 40(2):285–293, 1995.

[94] V.A. Lebedev. Behavior of random measures under filtration change. *Theory Probab. Appl.*, 40(4):645–652, 1995.

[95] C. Léonard. Some epidemic systems are long range interacting particle systems. In J.-P. Gabriel, C. Lefèvre, and P. Picard, editors, *Stochastic Processes in Epidemic Theory*, pages 170–183. Springer, Berlin, 1990.

[96] C. Léonard. Large deviations for long range interacting particle systems with jumps. *Ann. Inst. Henri Poincaré Probab. Stat.*, 31(2):289–323, 1995.

[97] A. Lindner and R. Maller. Lévy integrals and the stationarity of generalised Ornstein-Uhlenbeck processes. *Stoch. Process. Appl.*, 115(10):1701–1722, 2005.

[98] M.B. Marcus and J. Rosiński. Continuity and boundedness of infinitely divisible processes: a Poisson point process approach. *J. Theor. Probab.*, 18(1): 109–160, 2005.

[99] T. Marquardt and L.F. James. Generating long memory models based on CARMA processes. Technical report, Technische Universität München, 2007.

[100] H.P. McKean, Jr. A class of Markov processes associated with nonlinear parabolic equations. *Proc. Nat. Acad. Sci.*, 56(6):1907–1911, 1966.

[101] H.P. McKean, Jr. Speed of approach to equilibrium for Kac's caricature of a Maxwellian gas. *Arch. Ration. Mech. Anal.*, 21(5):343–367, 1966.

[102] H.P. McKean, Jr. Propagation of chaos for a class of non-linear parabolic equations. In *Stochastic Differential Equations (Lecture Series in Differential*

Equations, Session 7, Catholic University, 1967), pages 41–57. Air Force Office of Scientific Research, Arlington, 1967.

[103] J. Mémin. Espaces de semi martingales et changement de probabilité. *Z. Wahrscheinlichkeitstheorie verw. Geb.*, 52(1):9–39, 1980.

[104] M. Metivier and J. Pellaumail. Mesures stochastiques à valeurs dans des espaces L_0. *Z. Wahrscheinlichkeitstheorie verw. Geb.*, 40(2):101–114, 1977.

[105] T.F. Móri. The maximum degree of the Barabási-Albert random tree. *Comb. Probab. Comput.*, 14(3):339–348, 2005.

[106] M. Moser and R. Stelzer. Functional regular variation of Lévy-driven multivariate mixed moving average processes. *Extremes*, 16(3):351–382, 2013.

[107] C. Mueller. The heat equation with Lévy noise. *Stoch. Process. Appl.*, 74(1): 67–82, 1998.

[108] L. Mytnik. Stochastic partial differential equation driven by stable noise. *Probab. Theory Relat. Fields*, 123(2):157–201, 2002.

[109] M. Nagasawa and H. Tanaka. Diffusion with interactions and collisions between coloured particles and the propagation of chaos. *Probab. Theory Relat. Fields*, 74(2):161–198, 1987.

[110] M. Nagasawa and H. Tanaka. On the propagation of chaos for diffusion processes with drift coefficients not of average form. *Tokyo J. Math.*, 10(2):403–418, 1987.

[111] M. Nagasawa and H. Tanaka. Stochastic differential equations of pure-jumps in relativistic quantum theory. *Chaos Solitons Fractals*, 10(8):1265–1280, 1999.

[112] A. Osękowski. *Sharp Martingale and Semimartingale Inequalities*. Birkhäuser, Basel, 2012.

[113] J. Pedersen. The Lévy-Itô decomposition of an independently scattered random measure. Technical report, MaPhySto, University of Aarhus, 2003.

[114] S. Peszat and J. Zabczyk. *Stochastic Partial Differential Equations with Lévy Noise*. Cambridge University Press, Cambridge, 2007.

[115] V. Pipiras and M.S. Taqqu. Integration questions related to fractional Brownian motion. *Probab. Theory Relat. Fields*, 118(2):251–291, 2000.

[116] M. Podolskij. Ambit fields: Survey and new challenges. In R.H. Mena, J.C. Pardo, V. Rivero, and G.U. Bravo, editors, *XI Symposium on Probability and Stochastic Processes*, pages 241–279. Springer, Cham, 2015.

[117] P. Protter. Volterra equations driven by semimartingales. *Ann. Probab.*, 13 (2):519–530, 1985.

[118] P.E. Protter. *Stochastic Integration and Differential Equations*. Springer, Berlin, 2nd edition, 2005.

[119] B.S. Rajput and J. Rosiński. Spectral representations of infinitely divisible processes. *Probab. Theory Relat. Fields*, 82(3):451–487, 1989.

[120] M. Reiß, M. Riedle, and O. van Gaans. Delay differential equations driven by Lévy processes: Stationarity and Feller properties. *Stoch. Process. Appl.*, 116 (10):1409–1432, 2006.

[121] R.T. Rockafellar. *Conjugate Duality and Optimization*. SIAM, Philadelphia, 1974.

[122] J. Rosiński. On path properties of certain infinitely divisible processes. *Stoch. Process. Appl.*, 33(1):73–87, 1989.

[123] J. Rosiński. On series representations of infinitely divisible random vectors. *Ann. Probab.*, 18(1):405–430, 1990.

[124] J. Rosiński. Series representations of Lévy processes from the perspective of point processes. In O. E. Barndorff-Nielsen, T. Mikosch, and S. I. Resnick, editors, *Lévy Processes*, pages 401–415. Birkhäuser, Boston, 2001.

[125] E. Saint Loubert Bié. Étude d'une EDPS conduite par un bruit poissonnien. *Probab. Theory Relat. Fields*, 111(2):287–321, 1998.

[126] K.-I. Sato. *Lévy Processes and Infinitely Divisible Distributions*. Cambridge University Press, Cambridge, 1999.

[127] K.-I. Sato. Stochastic integrals in additive processes and applications to semi-Lévy processes. *Osaka J. Math.*, 41(1):211–236, 2004.

[128] R. Stelzer, T. Tosstorff, and M. Wittlinger. Moment based estimation of supOU processes and a related stochastic volatility model. *Stat. Risk Model.*, 32(1):1–24, 2015.

[129] F.W. Steutel and K. van Harn. *Infinite Divisibility of Probability Distributions on the Real Line*. Marcel Dekker, New York, 2004.

[130] C. Stricker and M. Yor. Calcul stochastique dépendant d'un paramètre. *Z. Wahrscheinlichkeitstheorie verw. Geb.*, 45(2):109–133, 1978.

[131] A.-S. Sznitman. Topics in propagation of chaos. In P.-L. Hennequin, editor, *École d'Été de Probabilités de Saint Flour XIX - 1989*, pages 165–251. Springer, Berlin, 1991.

[132] M. Talagrand. *Mean Field Models for Spin Glasses. Volume I: Basic Examples*. Springer, Berlin, 2011.

[133] M. Talagrand. *Mean Field Models for Spin Glasses. Volume II: Advanced Replica-Symmetry and Low Temperature*. Springer, Berlin, 2011.

[134] V. Todorov and G. Tauchen. Simulation methods for Lévy-driven continuous-time autoregressive moving average (CARMA) stochastic volatility models. *J. Bus. Econ. Stat.*, 24(4):455–469, 2006.

[135] J. Vaillancourt. On the existence of random McKean-Vlasov limits for triangular arrays of exchangeable diffusions. *Stoch. Anal. Appl.*, 6(4):431–446, 1988.

[136] J.B. Walsh. An introduction to stochastic partial differential equations. In P.L. Hennequin, editor, *École d'Été de Probabilités de Saint Flour XIV - 1984*, pages 265–439. Springer, Berlin, 1986.

[137] J.B. Walsh. Finite elment methods for parabolic stochastic PDE's. *Potential Anal.*, 23(1):1–43, 2005.

[138] Z. Wang. Existence and uniqueness of solutions to stochastic Volterra equations with singular kernels and non-Lipschitz coefficients. *Stat. Probab. Lett.*, 78(9):1062–1071, 2008.

[139] X. Zhang. Euler schemes and large deviations for stochastic Volterra equations with singular kernels. *J. Differ. Equ.*, 244(9):2226–2250, 2008.