Technische Universität München

Fakultät für Mathematik

Lehrstuhl für Mathematische Statistik

Spatio–temporal Processes of Stochastic Integral Type

Thomas Delerue

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: Prof. Dr. Robert König

Prüfer der Dissertation:
1. Prof. Dr. Claudia Klüppelberg
2. Prof. Dr. Carsten Chong, Columbia University, New York
3. Prof. Dr. Jean Jacod, Sorbonne University, Paris

Die Dissertation wurde am 30.06.2021 bei der Technischen Universität München eingereicht und durch die Fakultät für Mathematik am 17.08.2021 angenommen.

Summary

In this dissertation, we tackle both theoretical challenges within the field of stochastic partial differential equations as well as statistical inference problems based on high-frequency estimation for a new class of stochastic processes.

In Chapter 1, given a sequence $\dot{L}^\varepsilon$ of Lévy noises, we derive necessary and sufficient conditions in terms of their variances $\sigma^2(\varepsilon)$ such that the solution to the stochastic heat equation with noise $\sigma(\varepsilon)^{-1}\dot{L}^\varepsilon$ converges in law to the solution to the same equation with Gaussian noise. This *normal approximation* result applies to both equations with additive and multiplicative noise and hence lifts the findings of S. Asmussen and J. Rosiński [*J. Appl. Probab.* **38** (2001) 482–493] and S. Cohen and J. Rosiński [*Bernoulli* **13** (2007) 195–210] for finite-dimensional Lévy processes to the infinite-dimensional setting without making distributional assumptions on the solutions such as infinite divisibility. One important ingredient of our proof is to characterize the solution to the limit equation by a sequence of martingale problems. To this end, it is crucial to view the solution processes both as random fields and as càdlàg processes with values in a Sobolev space of negative real order.

In Chapter 2, we prove the normal approximation for the solution u^ε to the stochastic wave equation driven by the Lévy noise $\sigma(\varepsilon)^{-1}\dot{L}^\varepsilon$ and thus extend the result of Chapter 1 to the class of hyperbolic stochastic PDEs. Furthermore, u^ε is shown to have a space–time version with a càdlàg property determined by the wave kernel, and its derivative $\partial_t u^\varepsilon$ a càdlàg version when viewed as a distribution-valued process. These two path properties are essential to our proof of the normal approximation as the limit is characterized by martingale problems that necessitate both random elements. Our results apply to additive as well as to multiplicative noises.

In Chapter 3, we consider the problem of estimating volatility for high-frequency data when the observed process is the sum of a continuous Itô semimartingale and a noise process that locally behaves like fractional Brownian motion with Hurst parameter H. The resulting class of processes, which we call mixed semimartingales, generalizes the mixed fractional Brownian motion introduced by P. Cheridito [*Bernoulli* **7** (2001) 913–934] to time-dependent and stochastic volatility. Based on central limit theorems for variation functionals, we derive consistent estimators and asymptotic confidence intervals for H and the integrated volatilities of both the semimartingale and the noise part, in all cases where these quantities are identifiable. When applied to recent stock price data, we find strong empirical evidence for the presence of fractional noise, with Hurst parameters H that vary considerably over time and between assets.

Finally, Chapter 4 consists of a supplementary material that gives in detail all proofs of the theoretical results of Chapter 3.

Zusammenfassung

In dieser Dissertation gehen wir sowohl an theoretische Herausforderungen im Bereich stochastische partielle Differentialgleichungen als auch an statistische Inferenzprobleme basierend auf Hochfrequenzschätzung für eine neue Klasse von stochastischen Prozessen heran.

In Kapitel 1, gegeben eine Folge $\dot{L}^\varepsilon$ von *Lévy Noises* leiten wir notwendige und hinreichende Bedingungen für deren Varianzen $\sigma^2(\varepsilon)$ her, sodass die Lösung der stochastischen Wärmeleitungsgleichung mit Noise $\sigma(\varepsilon)^{-1}\dot{L}^\varepsilon$ in Verteilung gegen die Lösung der gleichen stochastichen PDEs mit Gauss'schem Noise konvergiert. Diese *normale Approximation* gilt für Gleichungen sowohl mit additivem als auch mit multiplikativem Noise. Sie hebt damit die Befunde von S. Asmussen und J. Rosiński [*J. Appl. Probab.* **38** (2001) 482–493] und S. Cohen and J. Rosiński [*Bernoulli* **13** (2007) 195–210] für endlich-dimensionale Lévy Prozesse auf ein unendlich-dimensionales Setting hoch, ohne Verteilungsannahmen über die Lösungen zu machen wie unendliche Teilbarkeit. Ein wichtiger Bestandteil unseres Beweises besteht darin, die Lösung der Grenzwertgleichung durch eine Folge von Martingalproblemen zu charakterisieren. Dazu ist es entscheidend, die Lösungsprozesse sowohl als Zufallsfelder als auch als càdlàg-Prozesse mit Werten in einem Sobolev-Raum negativer reeller Ordnung zu betrachten.

In Kapitel 2 weisen wir die normale Approximation für die Lösung u^ε der stochastischen Wellengleichung angetrieben durch den Lévy Noise $\sigma(\varepsilon)^{-1}\dot{L}^\varepsilon$ nach und erweitern damit die Resultate von Kapitel 1 auf die Klasse der hyperbolischen stochastischen PDEs. Außerdem wird gezeigt, dass u^ε als Zufallsfeld in Raum und Zeit eine Version mit einer càdlàg-Eigenschaft hat, welche durch den Wellenkern bestimmt wird, und seine Ableitung $\partial_t u^\varepsilon$ eine càdlàg-Version, wenn sie als distributionswertiger Prozess betrachtet wird. Diese beiden Pfadeigenschaften sind für unseren Beweis der normalen Approximation wesentlich, da der Grenzwert durch Martingalprobleme charakterisiert ist, welche beide Zufallselemente erfordern. Unsere Ergebnisse gelten sowohl für additiven als auch für multiplikativen Noise.

In Kapitel 3 betrachten wir das Problem der Volatilitätsschätzung für hochfrequente Daten, wenn der beobachtete Prozess die Summe eines kontinuierlichen Itô Semimartingals und eines Noise-Prozesses ist, welcher sich lokal wie eine fraktioniale Brownsche Bewegung mit Hurst-Parameter H verhält. Die resultierende Klasse von Prozessen, die wir *gemischte Semimartingale* nennen, verallgemeinert die von P. Cheridito [*Bernoulli* **7** (2001) 913-934] eingeführte gemischte fraktionelle Brownsche Bewegung auf zeitabhängige und stochastische Volatilität . Basierend auf zentralen Grenzwertsätzen für Variationsfunktionale leiten wir konsistente Schätzer und asymptotische Konfidenzintervalle für H und die integrierten Volatilitäten sowohl des Semimartingal- als auch des Noiseanteils in allen Fällen ab, in denen diese Größen identifizierbar sind. Bei der Anwendung auf aktuelle Aktienkursdaten finden wir starke empirische Beweise für das Vorhandensein von fraktionalem Noise.

Kapitel 4 schließlich besteht aus einem ergänzenden Material, das im Detail alle Beweise der theoretischen Ergebnisse von Kapitel 3 liefert.

Acknowledgments

I am very grateful to my supervisor Claudia Klüppelberg for giving me the opportunity to accomplish this research work at the Chair of Mathematical Statistics at the Technical University of Munich. I have enjoyed my time at the Chair very much and appreciated the open environment auspicious for fruitful work and constructive discussions with colleagues under the guidance of Claudia. She has helped me throughout and encouraged me from the very beginning to participate at many international conferences and to present my work.

I am greatly indebted in my cosupervisor Carsten Chong, for having accompanied me through all three projects that constitute this dissertation. Thanks to him, I could discover and work on many areas of stochastic analysis as well as state-of-the-art research topics in stochastic PDEs and financial statistics. I gained much knowledge but, most importantly to me, I could always benefit from his insights and aid to get through the hardest theoretical aspects of this dissertation. I thank him also for being permanently available despite the geographical distance, as well as for countless meetings toward the end of my thesis. Many thanks also to him for having invited me to EPFL as well as to Robert Dalang for his pleasant hospitality during my stay there. I am also thankful to Guoying Li, who assisted us with the financial data analysis in `R` for the last project of this thesis.

I would like to thank the Deutsche Forschungsgemeinschaft for their financial support of this thesis.

My many thanks go to all colleagues from the Chair and the other Chairs of the Department of Mathematics that I had the pleasure to meet through the years, for the friendly environment as well as for their support. I am particularly grateful to Mathias Drton for his support over the last year, to Stephan Haug for his help and guidance, and to Andrea Grant for her repeated help.

I am deeply grateful to my family, in particular my mother and brother, for their understanding, patience and help, also despite the geographical distance. My father accompanies my thoughts all the time.

Last but not least, my thoughts go to Shoko. She accompanied, supported and helped me through the toughest phase of this thesis and for that alone, my gratitude to her is boundless.

Introduction

In this thesis, we develop functional central limit theorems for the solutions to **stochastic partial differential equations** (stochastic PDEs) driven by a pure-jump Lévy space–time white noise as well as for a new model within financial statistics that we call **mixed semimartingale**, in which stock prices observed at high frequency are contaminated by a fractional microstructure noise. We further develop estimation procedures for the volatility of price processes and perform an extensive simulation study and data analysis based on high-frequency financial data.

The motivation for the first part of this thesis is the normal (or Gaussian) approximation for Lévy processes. If we consider a family $L^\varepsilon_t = \int_0^t \int_{\mathbb{R}} z\,(\mu^\varepsilon - \nu^\varepsilon)(\mathrm{d}s, \mathrm{d}z)$ of Lévy processes consisting of compensated jumps (μ^ε is a Poisson random measure on $\mathbb{R}^+ \times \mathbb{R}$, $\nu^\varepsilon(\mathrm{d}s, \mathrm{d}z) = \mathrm{d}s Q^\varepsilon(\mathrm{d}z)$ its intensity measure and $Q^\varepsilon(\mathrm{d}z)$ a Lévy measure on $\mathbb{R}$) such that the variance of the jumps is finite (that is, $\sigma^2(\varepsilon) = \int_{\mathbb{R}} z^2\, Q^\varepsilon(\mathrm{d}z) < \infty$), then it holds that

$$\sigma(\varepsilon)^{-1}\, L^\varepsilon \xrightarrow{d} W, \quad \text{as} \quad \varepsilon \to 0, \tag{I}$$

where W is a standard Brownian motion, if and only if the following analytical condition on the Lévy measure $Q^\varepsilon(\mathrm{d}z)$ holds:

$$\lim_{\varepsilon\to 0} \frac{1}{\sigma^2(\varepsilon)} \int_{|z|>\kappa\sigma(\varepsilon)} z^2\, Q^\varepsilon(\mathrm{d}z) = 0 \qquad \text{for all} \quad \kappa > 0. \tag{II}$$

In (I), the convergence takes place in the Skorokhod space $D([0,\infty), \mathbb{R})$. This is a well-known weak convergence result which has been proven in [36] in its full generality. An important special case of this **normal approximation** is the so-called small jumps approximation of Lévy processes; see [9]. In that situation, we have a single Lévy process L with a Lévy measure $Q(\mathrm{d}z)$ and we consider the processes L^ε that arise from setting $\mu^\varepsilon(\mathrm{d}s, \mathrm{d}z) = \mathbb{1}_{\{|z|\le\varepsilon\}}\mu(\mathrm{d}s, \mathrm{d}z)$ and $Q^\varepsilon(\mathrm{d}z) = \mathbb{1}_{\{|z|\le\varepsilon\}}Q(\mathrm{d}z)$, $\varepsilon > 0$. Then, provided (II) holds, the normal approximation (I) has the following interpretation: As the amplitude ε decreases, the Lévy process $\sigma(\varepsilon)^{-1}L^\varepsilon$ consisting only of normalized jumps of amplitude $\le \sigma(\varepsilon)^{-1}\varepsilon$ tends to look more and more like the Brownian motion W.

Investigating the condition (II) for prominent examples of Lévy processes, [9] obtains the following: The normal approximation neither holds for the compound Poisson process (hence, an infinite activity of jumps is necessary for (I)), nor for the gamma process. It does, however, for any stable process of index $\alpha \in (0, 2)$ and for the normal inverse Gaussian process.

From this starting point, we investigate convergence in law of stochastic PDEs for which the driving noise is L^ε. More precisely, we consider the *stochastic heat equation*

$$\partial_t u^\varepsilon(t,x) = \partial_{xx} u^\varepsilon(t,x) + f(u^\varepsilon(t,x)) \frac{\dot{L}^\varepsilon(t,x)}{\sigma(\varepsilon)}, \qquad (t,x) \in [0,T] \times [0,\pi], \tag{III}$$

and the *stochastic wave equation*

$$\partial_{tt}u^\varepsilon(t,x) = \partial_{xx}u^\varepsilon(t,x) + f(u^\varepsilon(t,x))\frac{\dot{L}^\varepsilon(t,x)}{\sigma(\varepsilon)}, \qquad (t,x)\in[0,T]\times\mathbb{R}, \tag{IV}$$

the prototypes of parabolic and hyperbolic stochastic PDEs, respectively. In these two equations, $\dot{L}^\varepsilon(t,x)$ is a pure-jump Lévy space–time white noise with a representation

$$L^\varepsilon(A) = \int_{\mathbb{R}^+\times\mathbb{R}}\int_{\mathbb{R}} \mathbb{1}_A(t,x)z\,(\mu^\varepsilon-\nu^\varepsilon)(\mathrm{d}t,\mathrm{d}x,\mathrm{d}z), \qquad A\in\mathcal{B}_b(\mathbb{R}^+\times\mathbb{R}). \tag{V}$$

The components defining L^ε are similar to the purely temporal situation: μ^ε is a Poisson random measure on $\mathbb{R}^+\times\mathbb{R}^2$ with intensity measure $\nu^\varepsilon(\mathrm{d}s,\mathrm{d}x,\mathrm{d}z) = \mathrm{d}s\mathrm{d}xQ^\varepsilon(\mathrm{d}z)$, and $Q^\varepsilon(\mathrm{d}z)$ is a Lévy measure with $\sigma^2(\varepsilon) = \int_{\mathbb{R}} z^2\,Q^\varepsilon(\mathrm{d}z) < \infty$. It is important to note that we consider equations with *multiplicative noise* $f(u^\varepsilon(t,x))\sigma(\varepsilon)^{-1}\dot{L}^\varepsilon(t,x)$, that is, the perturbation of the system considered is modeled as the product of the noise and (some function of) the solution itself.

We then raise the question: Do we have $u^\varepsilon(t,x)\xrightarrow{d} u(t,x)$ where u solves

$$\partial_t u(t,x) = \partial_{xx}u(t,x) + f(u(t,x))\dot{W}(t,x), \qquad (t,x)\in[0,T]\times[0,\pi], \tag{VI}$$

or

$$\partial_{tt}u(t,x) = \partial_{xx}u(t,x) + f(u(t,x))\dot{W}(t,x), \qquad (t,x)\in[0,T]\times\mathbb{R}, \tag{VII}$$

depending on the stochastic PDE considered? In the equations above, $\dot{W}(t,x)$ is a Gaussian space–time white noise, that is, a centered Gaussian random measure $\{W(A)\mid A\in\mathcal{B}_b(\mathbb{R}^+\times\mathbb{R})\}$ with covariance structure $\mathbb{E}[W(A)W(B)] = \mathrm{Leb}_{\mathbb{R}^+\times\mathbb{R}}(A\cap B)$ for bounded Borel sets $A,B\subseteq\mathbb{R}^+\times\mathbb{R}$. In Chapter 1 and Chapter 2, which led to the papers [32] and [43], respectively, we give a positive answer to that question.

As a first step toward our results, it is necessary to find suitable function spaces that support the solution u^ε to (III) and to (IV). In particular, we need some sort of càdlàg path regularity for u^ε because our proofs are based on the general semimartingale theory. First of all, throughout this thesis we work with the concept of mild solutions to stochastic PDEs, an approach that traces back to a seminal work by J. B. Walsh; see [99]. In our setting, this is a predictable random field u^ε satisfying for each space–time point (t,x),

$$u^\varepsilon(t,x) = \int_0^t\int_{\mathbb{R}} G_{t-s}(x,y)\frac{f(u^\varepsilon(s,y))}{\sigma(\varepsilon)}\,L^\varepsilon(\mathrm{d}s,\mathrm{d}y).$$

The kernel G is the Green's function of the heat operator on $[0,T]\times[0,\pi]$ or of the wave operator on $[0,T]\times\mathbb{R}$. Recent results guarantee the existence and uniqueness of u^ε; see [28], for instance. In the two equations considered, u^ε is locally square-integrable. Furthermore, for the heat equation, one issue is that G is singular at the origin, whence each time a jump from $L^\varepsilon(\mathrm{d}s,\mathrm{d}y)$ occurs, it creates an explosion of the solution. So u^ε cannot have any path regularity jointly in space and time. An alternative is to view u^ε as a process with only a time variable and values in a space of distributions. In doing so it was shown in[31] that the solution has a càdlàg version $\overline{u}^\varepsilon$ in the fractional Sobolev space of negative order denoted by $H_{-r}([0,\pi])$, for any $r>1/2$. Drawing upon this result, instead of u^ε only, we consider for the stochastic heat equation, the *pairs*

$$(u^\varepsilon,\overline{u}^\varepsilon)\in L^2([0,T]\times[0,\pi])\times D([0,T],H_{-r}([0,\pi])). \tag{VIII}$$

In the main result of Chapter 1, Theorem 1.2.1, we then prove that $(u^\varepsilon, \overline{u}^\varepsilon) \stackrel{d}{\longrightarrow} (u, \overline{u})$, where u is the continuous mild solution to (VI) and $\overline{u}$ its continuous version in $H_{-r}([0,\pi])$, if and only if the condition (II) is satisfied. It is remarkable that, comparing with the purely temporal case, no additional assumption on the noise $\dot{L}^\varepsilon$ is required for that result.

By contrast, the wave kernel in one space dimension has no singularity as it is the indicator function is the backward light cone. Hence, we can expect the solution to the wave equation to exhibit some path regularity in space–time. In fact, we show that it has a version $\overline{u}^\varepsilon$ with a two-dimensional càdlàg property denoted by $\preceq$ that fits the shape of the wave kernel. In addition, we also need the (distributional) time derivative $\partial_t u^\varepsilon$. We show that there exists a càdlàg process $\overline{v}^\varepsilon$ with values in a distribution space $H_{-r}(\mathbb{R})$ defined similarly as $H_{-r}([0,\pi])$ that completely characterizes $\partial_t u^\varepsilon$. We then prove for the wave equation, in Theorem 2.4.1, that the *pair*

$$(\overline{u}^\varepsilon, \overline{v}^\varepsilon) \in \Big(D_{\preceq}([0,T]\times[0,L]) \cap L^2([0,T]\times[0,L])\Big) \times D([0,T], H_{-r}(\mathbb{R})) \tag{IX}$$

converges weakly to the pair $(u, \overline{v})$, where u is the continuous mild solution to (VII) and $\overline{v}$ a continuous $H_{-r}(\mathbb{R})$-valued process that characterizes $\partial_t u$. This is the main result of Chapter 2.

We see that the functional setting is different for both equations and, consequently, many aspects in the proof of the normal approximation differ in the two chapters. However, the general idea of the proof is the same. It is important to note that because we consider stochastic PDEs with multiplicative noise, which is a major contribution of our results, the proof is much more challenging than in the case of additive noise (f constant). For instance, we cannot make use of weak convergence results for infinitely divisible distributions since we have stochastic integrands.

First, we show tightness of each random element within the function space that supports it. To that aim, we use various tightness criteria (in particular, the Aldous condition and generalization thereof to multi-indexed processes). Next, the identification step of the limit is based on the fact that we can find a characterization of mild solutions in terms of semimartingales with known characteristics. For the heat equation, we have for all test functions $\phi \in C_c^\infty((0,\pi))$,

$$\langle \overline{u}_t^\varepsilon, \phi\rangle = \int_0^t \int_0^\pi u^\varepsilon(s,x)\phi''(x)\,\mathrm{d}s\,\mathrm{d}x + \int_0^t \int_0^\pi \frac{f(u^\varepsilon(s,x))}{\sigma(\varepsilon)}\phi(x)\,L^\varepsilon(\mathrm{d}s,\mathrm{d}x). \tag{X}$$

And for the wave equation, for all $\phi_1, \phi_2 \in C_c^\infty(\mathbb{R})$,

$$\begin{aligned}\int_{\mathbb{R}} \overline{u}^\varepsilon(t,x)\phi_1(x)\,\mathrm{d}x + \langle \overline{v}_t^\varepsilon, \phi_2\rangle \\ = \int_0^t \int_{\mathbb{R}} u^\varepsilon(s,x)\phi_2''(x)\,\mathrm{d}x\mathrm{d}s + \langle \partial_t u^\varepsilon(t,\cdot), \phi_1\rangle + \int_0^t \int_{\mathbb{R}} \phi_2(x)\frac{f(u^\varepsilon(s,x))}{\sigma(\varepsilon)}\,L^\varepsilon(\mathrm{d}s,\mathrm{d}x).\end{aligned} \tag{XI}$$

By well-known results in [65], the semimartingales in (X) and (XI) are completely characterized by martingale problems involving certain complex-valued martingales $M^\varepsilon_{\mathrm{HEAT}}$ and $M^\varepsilon_{\mathrm{WAVE}}$, respectively. We then construct further processes M_{HEAT} and M_{WAVE} in such a way that *if* M_{HEAT} and M_{WAVE} are shown to be martingales as well, then the random field w (obtained from weak converging subsequences) necessarily is the mild solution to (VI) and (VII), respectively. But under the condition (II) on the Lévy measure, we can show that $M^\varepsilon_{\mathrm{HEAT}} \longrightarrow M_{\mathrm{HEAT}}$ and $M^\varepsilon_{\mathrm{WAVE}} \longrightarrow M_{\mathrm{WAVE}}$ as $\varepsilon \to 0$ in the Skorokhod topology, from which the desired martingale property of M_{HEAT} and M_{WAVE} can be inferred, hereby yielding our normal approximation results.

It is important to note the intricate interplay of the different random elements at hand in (X) and (XI): The semimartingale characteristics of the processes on the left-hand sides do not directly depend on the processes themselves but on the random field solution u^ε (and also on

$\partial_t u^\varepsilon$ for the wave equation). In particular, they are not deterministic. This is precisely the reason why we need to consider *pairs* in (VIII) and (IX).

The normal approximation of stochastic PDEs has applications, for instance, in neurophysiology: There, the stochastic cable equation is used to describe the propagation of *action potentials* within one neuron, from the site of initiation to the end of its axon, which is viewed as a thin cylinder. We then have for the potential Y^ε,

$$\partial_t Y^\varepsilon(t,x) = \partial_{xx} Y^\varepsilon(t,x) - Y^\varepsilon(t,x) + \sigma(Y^\varepsilon(t,x))\dot{L}^\varepsilon(t,x)$$

where $\dot{L}^\varepsilon(t,x)$ describes the electrical impulses arriving at time t and position x. By our results, this stochastic PDE can be approximated after appropriate normalization by

$$\partial_t Y(t,x) = \partial_{xx} Y(t,x) - Y(t,x) + \sigma(Y(t,x))\dot{W}(t,x)$$

with a Gaussian white noise $\dot{W}$, an equation far easier to simulate in practice.

We now turn to the second part of this thesis, that is, Chapter 3 and Chapter 4. Here, we introduce a new statistical model that we call **mixed semimartingale** and that we define (in a somewhat simplified version) as follows:

$$Y_t = X_t + Z_t, \quad X_t = \int_0^t a_s \,\mathrm{d}s + \int_0^t \sigma_s \,\mathrm{d}B_s, \quad Z_t = \int_0^t g(t-s)\rho_s \,\mathrm{d}W_s, \quad g(t) = K_H^{-1} t^{H-\frac{1}{2}} \quad \text{(XII)}$$

where $0 < H < 1/2$. The process $(X_t)_{t\geq 0}$ is a continuous Itō semimartingale and the process $(Z_t)_{t\geq 0}$ a continuous moving-average. σ and ρ are volatility processes (possibly dependent), B and W are independent Brownian motions. Furthermore, if $\rho \equiv 1$, one sees that Z_t is, up to a process of finite variation, the Mandelbrot–van Ness representation of *fractional Brownian motion (fBM)* with a Hurst index $H \in (0, 1/2)$. Note that if both σ and ρ in (XII) are constant, we obtain a special class of models called mixed fractional Brownian motion (mfBM) that was first introduced by P. Cheridito in [25].

The mixed semimartingale model is motivated by empirical observations in financial data. Consider the realized volatility $\widehat{V}^n_{0,t} = \sum_{i=1}^{[T/\Delta_n]} |Y_{i\Delta_n} - Y_{(i-1)\Delta_n}|^2$ where Y_t in (XII) is now an observed asset price process, Δ_n is a sampling frequency and T a finite time horizon. For each stock of the DJIA index (log of the mid-quote data) and for each trading day in 2019, we calculated $\widehat{V}^n_{0,t}$ for several sampling frequencies Δ_n. From this, we observed that $\widehat{V}^n_{0,t}$ diverges as $\Delta_n \to 0$ at a rate of Δ_n^α with $\alpha \in (-1, 0)$ in most cases. The reason for this blow-up is the so-called *market microstructure noise*. It can be viewed as the difference between the *observed* price process Y_t and the *efficient* price, which is the continuous-time process X_t in (XII) that arises as a scaling limit of prices when transactions occur more and more frequently.

We propose to model the microstructure noise by the fractional component Z_t in (XII). This is interesting for the following reasons. First, as shown by our theoretical results, under this model we have $\Delta_n^{1-2H}\widehat{V}^n_{0,T} \xrightarrow{\mathbb{P}} \int_0^T \rho_s^2 \,\mathrm{d}s$. Since $H < 1/2$, all scaling exponents between $(-1, 0)$ can be reproduced via the Hurst index. An analogous observation was made for the sample variance of price increments. Furthermore, it is a continuous-time model, whereas most existing microstructure noise models are discrete time series. In a discrete setting, as soon as the variance of the noise variables goes to 0 as $\Delta_n \to 0$ (*shrinking noise*), compatibility issues between sampling frequencies appear. On the other hand, if the variance of the noise remains constant when increasing the sampling frequency (*non-shrinking noise*, i.i.d. noise for example), then theoretical results show that $\widehat{V}^n_{0,T}$ explodes as $\Delta_n \to 0$ at a rate of Δ_n^{-1}, which was observed empirically only in few cases. Our continuous-time model also has the desirable feature that the process $t \mapsto Z_t$ is Lebesgue-measurable. Finally, it allows for serial dependence, since the increments of fBM are negative correlated (for $H < 1/2$), or dependence with the price process X_t via the noise volatility ρ.

With the statistical model (XII) at hand, we pursue a theoretical investigation of the asymptotic properties of variation functionals $\widehat{V}^n_{r,T} = \sum_{i=1}^{[T/\Delta_n]} \Delta_i^n Y \Delta_{i+r}^n Y$ where $\Delta_i^n Y = Y_{i\Delta_n} - Y_{(i-1)\Delta_n}$ is an increment and $r \in \mathbb{N}_0$, in order to develop statistical procedures for estimating parameters of interest: The integrated price volatility $\int_0^T \sigma_s^2 \,\mathrm{d}s$, the integrated noise volatility $\int_0^T \rho_s^2 \,\mathrm{d}s$ and the Hurst parameter $H \in (0, 1/2)$ of the noise. Our theoretical findings include a law of large numbers (LLN) and a central limit theorem (CLT) for these functionals: If we set $\widehat{V}^n_t = (\widehat{V}^n_{0,t}, \ldots, \widehat{V}^n_{R,t})$ with $R \in \mathbb{N}_0$, then we have for all $H \in (0, \frac{1}{2})$,

$$\Delta_n^{-\frac{1}{2}} \left\{ \Delta_n^{1-2H} \widehat{V}^n_t - \Gamma^H \int_0^t \rho_s^2 \,\mathrm{d}s - e_1 \int_0^t \sigma_s^2 \,\mathrm{d}s \, \Delta_n^{1-2H} \, \mathbb{1}_{[\frac{1}{4}, \frac{1}{2})}(H) \right\} \xrightarrow{\mathrm{st}} \mathcal{Z}_t \qquad \text{(XIII)}$$

where $\xrightarrow{\mathrm{st}}$ denotes stable convergence in law, $\mathcal{Z}_t = \int_0^t c_s \,\mathrm{d}\overline{W}_s$ with a standard Brownian motion $\overline{W} \in \mathbb{R}^{1+R}$ (independent of all previous processes) and $c_s \in \mathbb{R}^{(1+R)\times(1+R)}$ satisfies $c_s c_s^T = \mathcal{C}^H \rho_s^4$ and $\mathcal{C}^H$ further describes the covariance structure of $\mathcal{Z}_t$. The numbers $(\Gamma_r^H)_{r \geq 0}$ appearing in (XIII) are

$$\Gamma_0^H = 1, \qquad \Gamma_r^H = ((r+1)^{2H} - 2r^{2H} + (r-1)^{2H})/2, \quad r \geq 1, \qquad \Gamma^H = (\Gamma_0^H, \ldots, \Gamma_R^H)$$

and correspond exactly to the autocorrelation function of the increments of an fBM with Hurst index H. In fact, both the LLN limit and the CLT limit come from the fractional part Z_t. Setting $\sigma = 0$ would thus lead to a CLT for fBM with the same form as in [14], for instance. In the general situation, however, X_t produces an asymptotic bias term equal to the integrated price volatility at the rate of Δ_n^{1-2H} for the realized volatility and only if $H > 1/4$. This is in line with the results of [96] stating that if $H < 1/4$, the processes Y_t and Z_t have equivalent laws and, thus, the price volatility cannot be estimated consistently. Note also that the knowledge of the fluctuation process $\mathcal{Z}_t$ in (XIII) will allow us to construct feasible estimators that are asymptotically normal. Indeed, one crucial benefit of stable convergence in law is the ability to combine it with other converging random variables:

$$Z_n \xrightarrow{\mathrm{st}} Z, \quad Y_n \xrightarrow{\mathbb{P}} Y \Longrightarrow (Z_n, Y_n) \xrightarrow{\mathrm{st}} (Z, Y).$$

A central limit theorem for much more general variation functionals than $\widehat{V}^n_t$, and that are normalized, constitutes the main result of Chapter 3; see Theorem 3.2.1. It exhibits additional asymptotic bias terms with a convergence rate of $\Delta_n^{j(1-2H)}$ with $j = 1, \ldots, N(H)$ (so they are negligible in the LLN but not in the CLT due to the rate of $\Delta_n^{1/2}$) and their number $N(H) = [1/(2-4H)]$ can be arbitrarily large depending on the value of the Hurst index. This shows that a CLT for functionals of a mixed semimartingale is not just the sum of the purely semimartingale model and the purely fractional model, which are both well studied in the literature. The proof of our main CLT revolves around two main steps: First, the CLT as such based on a centering with appropriate conditional expectations and second, the convergence of the latter to the LLN limit after removing the asymptotic bias terms. Regarding the first part, it turns out that the semimartingale part X_t is asymptotically negligible and does not play a role in the shape of the CLT. So the fractional part Z_t dominates and in order to analyze it, we based our approach on martingale approximations for fractional processes and martingale CLTs as conducted in [27, 29]. We also use constantly size estimates based on the structure of the mixed semimartingale model and on important regularity assumptions that are commonly encountered in the literature: First, σ has the same regularity as Brownian motion, and second, ρ is the sum of a process with a regularity strictly greater than $1/2$ and of an Itō semimartingale.

Based on the CLT (XIII), we construct consistent, asymptotically mixed normal and, hence, bias-free estimators for the aforementioned parameters of interest. As a matter of fact, the Hurst index H must be estimated first, which is thus our first focus. Our basic estimator is

$$\widetilde{H}^n = \phi^{-1}\left(\frac{\langle a, \widehat{V}^n_T\rangle}{\langle b, \widehat{V}^n_T\rangle}\right) \qquad \text{with} \qquad \phi(H) = \frac{\langle a, \Gamma^H\rangle}{\langle b, \Gamma^H\rangle},$$

for some weight vectors $a = (a_0, \ldots, a_R)$, $b = (b_0, \ldots, b_R)$ in $\mathbb{R}^{1+R}$ (and $\langle \cdot, \cdot\rangle$ the standard inner product on $\mathbb{R}^{1+R}$). In order to deal with the bias term in (XIII), we follow two strategies that lead to two different estimators. First, we do not consider the realized volatility $\widehat{V}^n_{0,T}$ in the estimator for H. In that case, it is straightforward to find feasible estimators $\widetilde{H}^{n,0}$, $\widetilde{C}^{n,0}_T$ and $\widetilde{\Pi}^{n,0}_T$ for H, the price and noise volatility, respectively (Theorem 3.3.4). For instance, we obtain $\widetilde{\alpha}_n^{-1/2}\Delta_n^{-1/2}(\widetilde{H}^{n,0} - H) \xrightarrow{\text{st}} \mathcal{N}(0,1)$ as well as $\widetilde{\beta}_n^{-1/2}\Delta_n^{1/2-2\widetilde{H}^{n,0}}(\widetilde{C}^{n,0}_T - C_T) \xrightarrow{\text{st}} \mathcal{N}(0,1)$ if $H \in (1/4, 1/2)$ where

$$\widetilde{C}^{n,0}_T = \left\{\widehat{V}^n_{0,T} - \frac{\langle c, \widehat{V}^n_T\rangle}{\langle c, \Gamma^{\widetilde{H}^{n,0}}\rangle}\right\}\left(1 - \frac{c_0}{\langle c, \Gamma^{\widetilde{H}^{n,0}}\rangle}\right)^{-1}$$

and $c \in \mathbb{R}^{1+R}$ is another weight vector. The random variables $\widetilde{\alpha}_n$ and $\widetilde{\beta}_n$ are known normalization terms.

However, $\widetilde{H}^{n,0}$ above suffers from a serious drawback: If $H = 1/2$ (which amounts to saying there is no fractional noise), since $\widehat{V}^n_{0,T}$ is excluded from the computations, it can be shown that $\langle a, \widehat{V}^n_T\rangle / \langle b, \widehat{V}^n_T\rangle$ converges stably in law to a Cauchy distribution. So $\widetilde{H}^{n,0}$ is not consistent anymore because with positive probability, it can take values on any nonempty open interval. Therefore, it is necessary to construct an alternate estimator that performs well for high values of the true parameter H. We can achieve that by including the realized volatility in (XIII). But then, it turns out that the asymptotic bias term in (XIII) produces further bias terms: The basic CLT for $\widetilde{H}^n$ becomes

$$\Delta_n^{-\frac{1}{2}}\left\{\widetilde{H}^n - H + \sum_{j=1}^{N(H)} \Phi^n_j \left(\int_0^T \sigma_s^2 \,\mathrm{d}s\right)^j\right\} \xrightarrow{\text{st}} \mathcal{N}(0, C)$$

where $N(H) = [1/(2-4H)]$, the random variables Φ^n_j are higher order terms coming from a Taylor expansion and $C > 0$. So in order to obtain a bias-free estimator for H, we need to estimate the price volatility. But as mentioned previously, to estimate $\int_0^T \sigma_s^2 \,\mathrm{d}s$ we first need an estimate for H. This leads to a complex iterated estimation procedure for both H and the price volatility, which ultimately yields feasible estimators for the three quantities of interest (Theorem 3.3.11). It is important to note that in practice, it becomes increasingly challenging to compute the terms Φ^n_j for high values of j as they are higher order derivatives. Hence, for practical computations $N(H)$ must be fixed. To the value $N(H) = 3$ we associate the feasible estimators $\widehat{H}^{n,3}$, $\widehat{C}^{n,3}_T$ and $\widehat{\Pi}^{n,3}_T$ (this is the highest value that we compute in this dissertation and it yields the best numerical results).

With these estimators at hand, we perform a simulation study in R, hereby comparing them with estimators already existing in the literature. Regarding the Hurst index, we consider a regression estimator $\widetilde{H}^n_{\text{VS}}$ based on a volatility signature plot, the autocorrelation estimator $\widetilde{H}^n_{\text{acf}} = (1 + \log_2(\widehat{V}^n_{1,t}/\widehat{V}^n_{0,t} + 1))/2$ and a change-of-frequency estimator $\widetilde{H}^n_{\text{DMS}}$ developed in [47], which was originally constructed for the mfBM model and has an optimal convergence rate of $\Delta_n^{1/2}$. Regarding the price volatility, we consider the two-scale realized variance estimator $\widetilde{C}^n_{\mathrm{TSRV}}$ introduced in [100] as well as the preaveraging estimator $\widetilde{C}^n_{\mathrm{preave}}$ of [56]. Both are well-known noise robust estimators in the field of financial econometrics and are already implemented in the `R` package `highfrequency`.

We generate paths corresponding to 20 trading days for the (mfBM) process $Y = \sigma B + \rho B^H$ with B^H an fBM, for several values of H ranging from 0 to 1/2. Here, $H = 0$ means that instead of B^H, we generate i.i.d. standard normal variables, and $H = 1/2$ means that we set $\rho = 0$. The simulation results show that our estimators for H outperform the existing estimators in the whole range $(0, 1/2)$ in terms of root-mean-squared-error (RMSE). The best estimator in the range $H \leq 0.35$ is by far $\widetilde{H}^{n,0}$. However, for the reason explained before, $\widetilde{H}^{n,0}$ has an ever increasing standard error for higher values of H (the same goes for $\widetilde{H}^n_{\mathrm{DMS}}$ as it can be shown that it is equal to $\widetilde{H}^{n,0}$ with special weight vectors a, b). But for $H \in (0.4, 0.5)$, our estimator $\widehat{H}^{n,3}$ has a lower RMSE than $\widetilde{H}^n_{\mathrm{VS}}$, $\widetilde{H}^n_{\mathrm{acf}}$ and $\widetilde{H}^n_{\mathrm{DMS}}$. Moreover, we conjecture that computing the estimator for H with realized volatility and an increasing number of corrected bias terms (that is, greater than 3) will gradually bring down the sample bias while maintaining the standard error at a low level, ultimately yielding better estimators for H. Regarding the price volatility, it turns out that $\widetilde{C}^{n,0}_T$ performs better than $\widehat{C}^{n,3}_T$, both in terms of bias and standard error. Furthermore, $\widetilde{C}^{n,0}_T$ is better than $\widetilde{C}^n_{\mathrm{TSRV}}$ and $\widetilde{C}^n_{\mathrm{preave}}$ in the range $H \in (0.15, 0.35)$ in terms of RMSE (sometimes exhibiting an RMSE half so big as the other two estimators). However, due to larger standard errors, $\widetilde{C}^{n,0}_T$ performs worse than both standard noise-robust estimators for lower and higher values of H.

Using these findings, we conclude with an empirical analysis. Based on the log mid-quote data of the 29 companies that were constituents of the DJIA index for the whole year of 2019 and using a sampling frequency Δ_n of one second, we produce a daily estimate of H using $\widetilde{H}^{n,0}$ and $\widehat{H}^{n,3}$ as well as a daily estimate of the price volatility using $\widetilde{C}^{n,0}_T$ and $\widehat{C}^{n,3}_T$ (and similarly for the noise volatility). For each estimate of H, we also compute a 95%-confidence interval. Our results for the Hurst index exhibit values varying within the interval $(0, 1/2)$ from one trading day to the other and differing between companies. Most of the time, the estimate of H is significantly above 0 and beneath 1/2. This constitutes a strong empirical evidence for asset- and time-dependent values of H. Moreover, in many cases our estimates of the price volatility exhibit a similar behavior to the standard noise-robust estimators $\widetilde{C}^n_{\mathrm{TSRV}}$ and $\widetilde{C}^n_{\mathrm{preave}}$. For concrete examples, we refer to Section 3.5 of Chapter 3.

The theoretical results regarding our mixed semimartingale model as well as the simulation study and the empirical analysis are exposed in Chapter 3, which led to the paper [33]. For the sake of clarity, all proofs related to that chapter were put separately in Chapter 4, which constitute the supplementary material [34] of the paper [33].

To conclude, we briefly present possible directions of future research that might further develop/enhance the mixed semimartingale model. First, we can allow the semimartingale part X_t in (XII) to exhibit jumps and the observed process Y_t to be sampled at irregular (random) observation times. These two topics were analyzed extensively in [64] in the semimartingale setting. Furthermore, an important component of market microstructure noise is the *rounding error* due to the discreteness of the price. In reality, transactions are usually executed up to the nearest cent, so that prices are often unchanged over several observations. This is not quite compatible with the continuity of the fractional component Z_t. Hence, the rounding error should be included in the modeling (XII) of the observed price Y_t.

Contents

Chapter 1

Normal approximation of the solution to the stochastic heat equation with Lévy noise

1.1 Introduction

The importance of the Gaussian distribution in probability theory and its popularity in applications are manifested in the central limit theorem: The total effect of a large number of small independent contributions is approximately normal. Therefore, when physical systems governed by one or several equations are perturbed by white noise (where "white" means stationary and uncorrelated), it is frequently assumed that the noise is *Gaussian*.

For example, in his Saint-Flour lecture notes [99], J. B. Walsh discusses an application of parabolic stochastic PDEs to the modeling of neuron potentials. Subject to impulses arriving according to a marked Poisson point process (with mean 0 and atoms of size $\dot{L}(t,x)$ at time t and position x), the electrical potential $u(t,x)$ of the neuron, viewed as a thin cylinder of length, say, π, is then well described by the *stochastic cable equation*

$$\partial_t u(t,x) = \partial_{xx} u(t,x) - u(t,x) + \dot{L}(t,x), \quad (t,x) \in [0,T] \times [0,\pi], \tag{1.1.1}$$

with suitable boundary and initial conditions. Arguing that "the impulses are generally small, and there are many of them, so that in fact $\dot{L}$ is very nearly a white noise" ([99], p. 311), the author then approximates (1.1.1) by

$$\partial_t u(t,x) = \partial_{xx} u(t,x) - u(t,x) + \dot{W}(t,x), \quad (t,x) \in [0,T] \times [0,\pi], \tag{1.1.2}$$

where $\dot{W}$ is a Gaussian space–time white noise.

But of course, the central limit theorem has limitations. In the absence of finite second moments, stable limits may arise; and if there are rare but large contributions, we may have a Poisson limit. In general, any *infinitely divisible distribution* can arise as a possible limit of compound Poisson laws; see Corollary 3.8 in [91]. This leads us to the following question: If we have a sequence of noises $\dot{L}^\varepsilon$ as above where the atom sizes of $\dot{L}^\varepsilon$ converge to 0 as $\varepsilon \to 0$, and if u^ε denotes the solution to (1.1.1) with noise $\sigma(\varepsilon)^{-1}\dot{L}^\varepsilon$ (where $\sigma^2(\varepsilon)$ is the variance of $\dot{L}^\varepsilon$), do we have convergence in distribution of u^ε to the solution u of (1.1.2) with the Gaussian noise $\dot{W}$? A positive answer for this *normal approximation* is given in Theorem 7.10 in [99]: If the atoms of $\dot{L}^\varepsilon$ are locally summable and the jump measure Q^ε of $\dot{L}^\varepsilon$ satisfies

$$\frac{1}{\sigma^{2+\delta}(\varepsilon)} \int_{\mathbb{R}} |z|^{2+\delta}\, Q^\varepsilon(\mathrm{d}z) \longrightarrow 0 \quad \text{as} \quad \varepsilon \to 0 \tag{1.1.3}$$

for some $\delta > 0$, then u^ε converges in distribution to u.

The purpose of this work is to substantially generalize this result in two aspects. Given that (1.1.3) is sufficient but not necessary for $u^\varepsilon \xrightarrow{d} u$ (see Remark 1.2.3 below), our first contribution is to show that the *necessary and sufficient* condition for the normal approximation is

$$\lim_{\varepsilon\to 0} \frac{1}{\sigma^2(\varepsilon)} \int_{|z|>\kappa\sigma(\varepsilon)} z^2\, Q^\varepsilon(\mathrm{d}z) = 0 \tag{1.1.4}$$

for all $\kappa > 0$. In fact, if L^ε (resp., W) is a Lévy process (resp., Brownian motion), the same condition was found to be necessary and sufficient for $\sigma(\varepsilon)^{-1}L^\varepsilon \xrightarrow{d} W$ in [36]. Somewhat surprisingly, in the special case of *small jump approximation*, that is, where $Q^\varepsilon(\mathrm{d}z) = \mathbb{1}_{\{|z|\le\varepsilon\}}\, Q(\mathrm{d}z)$ and Q is a given Lévy measure, it was shown in [9] that condition (1.1.4) *fails* for prominent examples such as the compound Poisson or the gamma distribution. So in these cases, the small jump approximation is not true for Lévy processes and by our results, not true for stochastic PDEs, either.

Our second contribution is to consider equations with *multiplicative* noise. To our best knowledge, previous works on the normal approximation of stochastic PDEs with jumps have only considered the situation of *additive* noise; see, besides the mentioned results in [99], also [70, 98] (there is, of course, literature concerning approximation of multiplicative Gaussian white noise by *smoother* noises [13, 54], but these problems are very different in nature than the one considered here). The proofs in [70, 98, 99] (as well as those of [9, 36]) are based on characteristic functions and the Lévy–Khintchine formula for infinitely divisible distributions, which obviously do not generalize to the situation of multiplicative noise.

Instead, our approach will be to show that u^ε satisfies *martingale problems* which, assuming (1.1.4) only and not the stronger condition (1.1.3), have a limit with a unique solution. But this leads to several complications. In order to prove convergence of the associated martingales, we need some sort of uniformity in the time variable (as given, for example, by convergence in the Skorokhod topology). So taking simply the space $L^2([0,T]\times[0,\pi])$ to support the solutions u^ε and u will not be sufficient. This is why we will draw upon the results of [31] and view the solution u^ε (and also u) as a càdlàg process on $[0,T]$ with values in the Sobolev space H_{-r} for some $r > \frac{1}{2}$ (see Section 1.2.1 for a definition). In order to show tightness in that space with the Aldous criterion [7], we will use the factorization method from [39, 90] to obtain uniform bounds in time without taking moments of order higher than two. Another subtlety that arises in the analysis of the (semi-)martingales mentioned above is that their predictable characteristics are not given by a function of the former, which distinguishes our proof from the corresponding ones for (finite- or infinite-dimensional) stochastic differential equations in [48, 75, 76].

We shall also mention that the initial motivation in [9, 36] to study the normal approximation of Lévy processes comes from numerical simulation. Indeed, for stochastic PDEs as in (1.1.1) with multiplicative Lévy noise, the rate of convergence of a numerical scheme obtained by removing the small jumps of the noise is slower for noises with a high intensity of small jumps; see [24]. However, the results in [73] and [94] show that in the case of SDEs, an additional Gaussian approximation of the otherwise neglected small jumps improves the rate of convergence. We leave it to future research to examine to what extent this also holds for stochastic PDEs.

The remaining paper is organized as follows. In Section 1.2, we first describe in detail the considered equations and recall the definition of Sobolev spaces of real order in Section 1.2.1 before we state our main result, Theorem 1.2.1, in Section 1.2.2. Here we also explain the main steps and ideas behind the proof, whereas the details are given in Section 1.3.

1.2 Results

1.2.1 Preliminaries

Let $T > 0$ and consider on a filtered probability space $(\Omega, \mathcal{F}, \boldsymbol{F} = (\mathcal{F}_t)_{t\leq T}, \mathbb{P})$ that satisfies the usual conditions, for any $\varepsilon > 0$, the *stochastic heat equation* on $[0,T] \times [0,\pi]$ with Dirichlet boundary conditions:

$$\begin{cases} \partial_t u^\varepsilon(t,x) = \partial_{xx} u^\varepsilon(t,x) + f(u^\varepsilon(t,x)) \dfrac{\dot{L}^\varepsilon(t,x)}{\sigma(\varepsilon)}, & (t,x) \in [0,T] \times [0,\pi], \\ u^\varepsilon(t,0) = u^\varepsilon(t,\pi) = 0, & \text{for all } t \in [0,T], \\ u^\varepsilon(0,x) = 0, & \text{for all } x \in [0,\pi]. \end{cases} \tag{1.2.1}$$

The function $f\colon \mathbb{R} \longrightarrow \mathbb{R}$ in equation (1.2.1) describes the *multiplicative* part of the noise and will be assumed to be a Lipschitz continuous function. Concerning the driving noise $\sigma(\varepsilon)^{-1}\dot{L}^\varepsilon$, we assume that $\dot{L}^\varepsilon$ is a pure-jump Lévy space–time white noise on $[0,T] \times [0,\pi]$ given by

$$L^\varepsilon(A) = \int_0^T \int_0^\pi \int_{\mathbb{R}} \mathbb{1}_A(t,x) z \, (\mu^\varepsilon - \nu^\varepsilon)(\mathrm{d}t, \mathrm{d}x, \mathrm{d}z) \tag{1.2.2}$$

for all $A \in \mathcal{B}([0,T]\times[0,\pi])$. In this representation, μ^ε is a homogeneous Poisson random measure on $[0,T] \times [0,\pi] \times \mathbb{R}$ relative to the filtration $\boldsymbol{F}$, with intensity measure $\nu^\varepsilon = \mathrm{Leb}_{[0,T]\times[0,\pi]} \otimes Q^\varepsilon$. Here Q^ε is a Lévy measure on $\mathbb{R}$, that is, $Q^\varepsilon(\{0\}) = 0$ and $\int_{\mathbb{R}} (1 \wedge z^2) \, Q^\varepsilon(\mathrm{d}z) < \infty$. We refer to Chapter II in [65] for the definition of stochastic integrals with respect to Poisson random measures. Furthermore, we assume that for all $\varepsilon > 0$,

$$0 < \sigma^2(\varepsilon) = \int_{\mathbb{R}} z^2 \, Q^\varepsilon(\mathrm{d}z) < \infty. \tag{1.2.3}$$

Note that this integral is the variance of $L^\varepsilon([0,1] \times [0,1])$. In the special case where we have a single Poisson random measure μ having intensity measure $\nu = \mathrm{Leb}_{[0,T]\times[0,\pi]} \otimes Q$, setting

$$Q^\varepsilon(A) = \int_{|z|\leq\varepsilon} \mathbb{1}_A(z) \, Q(\mathrm{d}z), \quad A \in \mathcal{B}(\mathbb{R}), \quad \varepsilon > 0, \tag{1.2.4}$$

leads us to the case of *small jump approximation* considered in [9].

A predictable random field $u^\varepsilon = \{u^\varepsilon(t,x) \mid (t,x) \in [0,T] \times [0,\pi]\}$ is called a *mild solution* to (1.2.1) if for all $(t,x) \in [0,T] \times [0,\pi]$,

$$\begin{aligned} u^\varepsilon(t,x) &= \int_0^t \int_0^\pi G_{t-s}(x,y) \frac{f(u^\varepsilon(s,y))}{\sigma(\varepsilon)} \, L^\varepsilon(\mathrm{d}s, \mathrm{d}y) \\ &= \int_0^t \int_0^\pi \int_{\mathbb{R}} G_{t-s}(x,y) f(u^\varepsilon(s,y)) \frac{z}{\sigma(\varepsilon)} \, (\mu^\varepsilon - \nu^\varepsilon)(\mathrm{d}s, \mathrm{d}y, \mathrm{d}z) \end{aligned} \tag{1.2.5}$$

$\mathbb{P}$-almost surely, where

$$G_t(x,y) = \frac{2}{\pi} \sum_{k=1}^\infty \sin(kx) \sin(ky) e^{-k^2 t} \mathbb{1}_{\{t\geq 0\}}, \tag{1.2.6}$$

for $(t,x,y) \in [0,T] \times [0,\pi]^2$, is the Dirichlet Green's function of the heat operator on $[0,\pi]$.

The existence of a mild solution u^ε to (1.2.1) is guaranteed by Theorem 3.1 in [28] and condition (1.2.3) on the Lévy measure Q^ε and it is, up to modifications, unique among all predictable random fields satisfying

$$\sup_{(t,x)\in[0,T]\times[0,\pi]} \mathbb{E}\left[|u^\varepsilon(t,x)|^p\right] < \infty \tag{1.2.7}$$

for any $0 < p \le 2$ and $\varepsilon > 0$.

In this paper, we want to examine when the *normal approximation* holds for u^ε, that is, when u^ε can be approximated in law by the mild solution u to the same stochastic heat equation as above, but driven by a Gaussian space–time white noise on $[0,T]\times[0,\pi]$:

$$\left[\begin{array}{ll} \partial_t u(t,x) = \partial_{xx} u(t,x) + f(u(t,x))\dot{W}(t,x), & (t,x)\in[0,T]\times[0,\pi], \\ u(t,0) = u(t,\pi) = 0, & \text{for all } t\in[0,T], \\ u(0,x) = 0, & \text{for all } x\in[0,\pi]. \end{array}\right. \tag{1.2.8}$$

The driving noise $\dot{W}$ is now a centered Gaussian random field $\{W(A) \mid A \in \mathcal{B}([0,T]\times[0,\pi])\}$ with covariance structure $\mathbb{E}[W(A)W(B)] = \mathrm{Leb}_{[0,T]\times[0,\pi]}(A\cap B)$ for any measurable sets $A, B \subseteq [0,T]\times[0,\pi]$. It is well-known (see, for example, Theorem 3.2 in [99]) that, up to modifications, equation (1.2.8) has a unique mild solution u satisfying the corresponding bound in (1.2.7) for all $p > 0$.

Throughout this work, we will look at the mild solutions u^ε and u from two different points of view. First, they are random elements in the function space $L^2([0,T]\times[0,\pi])$ as the uniform bound (1.2.7) shows. But then, as mentioned in the introduction, we will need stronger path regularity in the time variable for our proofs. This is why we shall consider u^ε and u also as stochastic processes with values in an infinite dimensional space, which we will describe in the following.

Consider for any $r > 0$, the fractional Sobolev space

$$H_r([0,\pi]) = \left\{ \phi \in L^2([0,\pi]) \;\middle|\; \sum_{k=1}^\infty (1+k^2)^r \langle \phi, \phi_k\rangle^2 < \infty \right\},$$

where $\phi_k(x) = \sqrt{2/\pi}\sin(kx)$, $k\in\mathbb{N}$, form an orthonormal basis of $L^2([0,\pi])$. This is a Hilbert space with scalar product

$$\langle f, g\rangle_r = \sum_{k=1}^\infty (1+k^2)^r \langle f, \phi_k\rangle\langle g, \phi_k\rangle, \quad f, g \in H_r([0,\pi]),$$

and norm $\|\phi\|_r = \sqrt{\langle \phi, \phi\rangle_r}$ for $\phi \in H_r([0,\pi])$.

The topological dual $H_{-r}([0,\pi])$ of $H_r([0,\pi])$ is also a Hilbert space, whose dual norm $\|\cdot\|_{-r}$ can be expressed, by the Riesz representation theorem, as

$$\|\phi'\|_{-r}^2 = \sum_{k=1}^\infty \langle \phi', \phi_{r,k}\rangle^2 = \sum_{k=1}^\infty (1+k^2)^{-r}\langle \phi', \phi_k\rangle^2, \quad \phi' \in H_{-r}([0,\pi]). \tag{1.2.9}$$

(Note that if φ_1, φ_2 are elements of the same L^2-space, then $\langle \varphi_1, \varphi_2\rangle$ will always denote the standard scalar product of that space. If ϕ is an element of a Hilbert space and ϕ' an element of its topological dual, then $\langle \phi', \phi\rangle$ will always denote the dual pairing of ϕ' with ϕ.)

Coming back to the mild solution to (1.2.1), if we identify u^ε with the process $(u_t^\varepsilon)_{t\le T}$ where

$$u_t^\varepsilon \colon H_r([0,\pi]) \longrightarrow \mathbb{R}, \quad \phi \mapsto \langle u^\varepsilon(t,\cdot), \phi\rangle = \int_0^\pi u^\varepsilon(t,y)\phi(y)\,\mathrm{d}y \tag{1.2.10}$$

for all $t \leq T$, then by Theorem 2.5 in [31], u^ε has a càdlàg modification in $H_{-r}([0,\pi])$ for any $r > 1/2$, which will be denoted by $\overline{u}^\varepsilon = (\overline{u}^\varepsilon_t)_{t\leq T}$ throughout this work. Similarly, by the identification (1.2.10) and Corollary 3.4 in [99], the mild solution u to (1.2.8) has a continuous modification $\overline{u}$ in $H_{-r}([0,\pi])$ for each $r > 1/2$.

1.2.2 Main result

We now introduce the Cartesian product

$$\Omega^* = L^2([0,T] \times [0,\pi]) \times D([0,T], H_{-r}([0,\pi])), \tag{1.2.11}$$

with $r > 1/2$. Let d_1 denote the metric induced by the L^2-norm on $L^2([0,T] \times [0,\pi])$ and d_2 be the Skorokhod metric on $D([0,T], H_{-r}([0,\pi]))$. We then equip Ω^* with the product metric

$$\tau((f_1,x_1),(f_2,x_2)) = d_1(f_1,f_2) + d_2(x_1,x_2) \tag{1.2.12}$$

for any $f_1, f_2 \in L^2([0,T] \times [0,\pi])$ and $x_1, x_2 \in D([0,T], H_{-r}([0,\pi]))$. The main result of this paper is the following limit theorem.

Theorem 1.2.1. *Assume that L^ε is given by* (1.2.2) *with a variance $\sigma^2(\varepsilon)$ that satisfies* (1.2.3) *for all $\varepsilon > 0$. Let u^ε be the $L^2([0,T] \times [0,\pi])$-valued mild solution to the stochastic heat equation* (1.2.1) *driven by $\dot{L}^\varepsilon/\sigma(\varepsilon)$ and $\overline{u}^\varepsilon$ be its càdlàg modification in $H_{-r}([0,\pi])$. Similarly, let u be the $L^2([0,T] \times [0,\pi])$-valued mild solution to the stochastic heat equation* (1.2.8) *driven by $\dot{W}$ and $\overline{u}$ be its continuous modification in $H_{-r}([0,\pi])$.*

Suppose also that the Lipschitz function f in (1.2.1) *satisfies $f(0) \neq 0$. Then, as $\varepsilon \to 0$,*

$$(u^\varepsilon, \overline{u}^\varepsilon) \xrightarrow{d} (u, \overline{u}) \quad in \quad (\Omega^*, \tau) \tag{1.2.13}$$

for all $r > 1/2$ if and only if (1.1.4) *holds for all $\kappa > 0$.*

Remark 1.2.2. We can generalize Theorem 1.2.1 to nonzero initial conditions. Assume that in both equations (1.2.1) and (1.2.8) we now have $u^\varepsilon(0,x) = u(0,x) = u_0(x)$ for all $x \in [0,\pi]$, where $u_0\colon [0,\pi] \longrightarrow \mathbb{R}$ is a bounded continuous function with $u_0(0) = u_0(\pi) = 0$. Define

$$u_0(t,x) = \int_0^\pi G_t(x,y) u_0(y)\,\mathrm{d}y, \quad (t,x) \in [0,T] \times [0,\pi].$$

Then Theorem 1.2.1 can be shown in a completely analogous manner if we assume that there exists $(t_0,x_0) \in [0,T] \times [0,\pi]$ such that $f(u_0(t_0,x_0)) \neq 0$ (instead of $f(0) \neq 0$), and this assumption is only needed for showing the necessity of (1.1.4). To be more precise, since the mild solution to (1.2.8) now satisfies

$$u(t,x) = u_0(t,x) + \int_0^t \int_0^\pi G_{t-s}(x,y) f(u(s,y))\, W(\mathrm{d}s,\mathrm{d}y)$$

$\mathbb{P}$-almost surely, a similar argument as in Remark 1.3.14 shows that $\mathbb{P}(f(u(t_1,x_1)) \neq 0) > 0$ for some $(t_1,x_1) \in [0,T] \times [0,\pi]$ and hence, the expectation in (1.3.70) is nonzero.

Remark 1.2.3. Let us relate the two conditions (1.1.3) and (1.1.4) to each other. Using Hölder's and Chebyshev's inequalities, we see from the estimate

$$\begin{aligned} \frac{1}{\sigma^2(\varepsilon)} \int_{|z|>\kappa\sigma(\varepsilon)} z^2\, Q^\varepsilon(\mathrm{d}z) &\leq \frac{1}{\sigma^2(\varepsilon)} \left(\int_{\mathbb{R}} |z|^{2+\delta}\, Q^\varepsilon(\mathrm{d}z) \right)^{\frac{2}{2+\delta}} Q^\varepsilon(\{|z| > \kappa\sigma(\varepsilon)\})^{\frac{\delta}{2+\delta}} \\ &\leq \frac{1}{\kappa^\delta \sigma^{2+\delta}(\varepsilon)} \int_{\mathbb{R}} |z|^{2+\delta}\, Q^\varepsilon(\mathrm{d}z) \end{aligned}$$

that (1.1.3) implies (1.1.4).

The other implication is not true in general. For example, assume that Q^ε has density

$$q_\varepsilon(z) = \frac{1}{2|z|^2}\mathbb{1}_{\{|z|\leq\varepsilon\}} + \frac{\varepsilon^2}{2C|z|^3\log(1+|z|)^2}\mathbb{1}_{\{|z|>1\}}, \quad z\in\mathbb{R},$$

where $C = \int_1^\infty z^{-1}\log(1+z)^{-2}\,\mathrm{d}z$. Then $\int_\mathbb{R} |z|^{2+\delta}\, Q^\varepsilon(\mathrm{d}z) = \infty$ for every $\varepsilon, \delta > 0$, so condition (1.1.3) does not hold. But a direct calculation shows that $\sigma^2(\varepsilon) = \varepsilon + \varepsilon^2$. So given $\kappa > 0$, we have $1 > \kappa\sigma(\varepsilon) > \kappa\sqrt{\varepsilon} > \varepsilon$ for small values of ε, which implies (1.1.4) because

$$\lim_{\varepsilon\to 0}\frac{1}{\sigma^2(\varepsilon)}\int_{|z|>\kappa\sigma(\varepsilon)} z^2\, Q^\varepsilon(\mathrm{d}z) = \lim_{\varepsilon\to 0}\frac{1}{\varepsilon+\varepsilon^2}\int_{|z|>\varepsilon} z^2\, Q^\varepsilon(\mathrm{d}z) = \lim_{\varepsilon\to 0}\frac{\varepsilon^2}{\varepsilon+\varepsilon^2} = 0.$$

Proof of Theorem 1.2.1. We begin by showing that (1.1.4) implies the weak convergence (1.2.13). Since Ω^* is a metric space, we follow the classical scheme of first showing tightness and then uniqueness of the limiting distribution.

In Theorem 1.3.4, we show that $\{u^\varepsilon \mid \varepsilon > 0\}$ is tight in $L^2([0,T]\times[0,\pi])$ and in Theorem 1.3.7 that $\{\overline{u}^\varepsilon \mid \varepsilon > 0\}$ is tight in $D([0,T], H_{-r}([0,\pi]))$. By the subsequence principle, this immediately implies that the random elements $\{(u^\varepsilon, \overline{u}^\varepsilon) \mid \varepsilon > 0\}$ are tight in (Ω^*, τ). As it turns out, no assumptions on the Lévy noise $\dot{L}^\varepsilon$ other than the ones specified in (1.2.2) and (1.2.3) are needed for this tightness property.

As a consequence, we can apply Prokhorov's theorem, which provides for any sequence $(\varepsilon_k)_{k\in\mathbb{N}}$ with $\varepsilon_k \to 0$, a subsequence $(\varepsilon_{k_l})_{l\in\mathbb{N}}$ such that $(u^{\varepsilon_{k_l}}, \overline{u}^{\varepsilon_{k_l}})_{l\in\mathbb{N}}$ converges weakly to some distribution on (Ω^*, τ) as $l \to \infty$. For notational simplicity, we will assume without loss of generality that the whole sequence $(u^{\varepsilon_k}, \overline{u}^{\varepsilon_k})_{k\in\mathbb{N}}$ converges weakly.

Since (Ω^*, τ) is a complete separable metric space, we can further apply Skorokhod's representation theorem (see Theorem 4.30 in [69]) and obtain random elements

$$(v^k, \overline{v}^k), (v, \overline{v})\colon (\overline{\Omega}, \overline{\mathcal{F}}, \overline{\mathbb{P}}) \longrightarrow (\Omega^*, \tau), \tag{1.2.14}$$

defined on a possibly different probability space $(\overline{\Omega}, \overline{\mathcal{F}}, \overline{\mathbb{P}})$, satisfying the following properties:

$$\begin{aligned}&(v^k, \overline{v}^k) \stackrel{d}{=} (u^{\varepsilon_k}, \overline{u}^{\varepsilon_k}) \quad \text{for all} \quad k\in\mathbb{N} \quad \text{and}\\ &(v^k, \overline{v}^k)(\overline{\omega}) \longrightarrow (v, \overline{v})(\overline{\omega}) \quad \text{in} \quad (\Omega^*, \tau) \quad \text{as} \quad k\to\infty \quad \text{for all} \quad \overline{\omega}\in\overline{\Omega}.\end{aligned} \tag{1.2.15}$$

We will show that

$$(v, \overline{v}) \stackrel{d}{=} (u, \overline{u}),$$

which in turn implies (1.2.13). To do this, we first define a filtration $\overline{\boldsymbol{F}} = (\overline{\mathcal{F}}_t)_{t\leq T}$ on $\overline{\Omega}$ by setting

$$\overline{\mathcal{F}}_t = \bigcap_{u>t} \sigma\left(v(s,x), \overline{v}_s \mid 0\leq x\leq \pi,\, s\leq u\right), \quad t\leq T. \tag{1.2.16}$$

We further define for $\xi\in\mathbb{R}$, $\phi\in C_c^\infty((0,\pi))$ and $t\leq T$,

$$\overline{B}_t = \int_0^t \langle v(s,\cdot), \phi''\rangle\,\mathrm{d}s, \quad \overline{C}_t = \int_0^t\int_0^\pi f^2(v(s,x))\phi^2(x)\,\mathrm{d}s\,\mathrm{d}x, \quad \overline{A}_t = i\xi\overline{B}_t - \frac{1}{2}\xi^2\overline{C}_t \tag{1.2.17}$$

as well as

$$\overline{M}_t = e^{i\xi\langle\overline{v}_t,\phi\rangle} - \int_0^t e^{i\xi\langle\overline{v}_s,\phi\rangle}\,\overline{A}(\mathrm{d}s). \tag{1.2.18}$$

We will then show in Theorem 1.3.12 that, under assumption (1.1.4), the pair $(v, \overline{v})$ satisfies the following martingale problem. For all $\xi \in \mathbb{R}$ and $\phi \in C_c^\infty((0,\pi))$, the process $(\overline{M}_t)_{t \leq T}$ is a martingale with respect to $(\overline{\Omega}, \overline{\mathcal{F}}, \overline{\boldsymbol{F}}, \overline{\mathbb{P}})$. Note that we are able to obtain this property only because in (1.2.15), we have $\overline{\omega}$-wise convergence both in the Skorokhod topology and in $L^2([0,T] \times [0,\pi])$, which is the reason why we view the solutions to (1.2.1) and (1.2.8) as *pairs* in Ω^*.

Next, we will show in Theorem 1.3.13 that this martingale property in turn implies that there exists a Gaussian space–time white noise $\dot{\widetilde{W}}$ on $[0,T] \times [0,\pi]$, possibly defined on a filtered extension $(\widetilde{\Omega}, \widetilde{\mathcal{F}}, \widetilde{\boldsymbol{F}}, \widetilde{\mathbb{P}})$ of $(\overline{\Omega}, \overline{\mathcal{F}}, \overline{\boldsymbol{F}}, \overline{\mathbb{P}})$ such that, with probability one, the random field v is equal in $L^2([0,T] \times [0,\pi])$ to the mild solution $\widetilde{v}$ to the stochastic heat equation

$$\left[\begin{array}{ll} \partial_t \widetilde{v}(t,x) = \partial_{xx} \widetilde{v}(t,x) + f(\widetilde{v}(t,x)) \dot{\widetilde{W}}(t,x), & (t,x) \in [0,T] \times [0,\pi], \\ \widetilde{v}(t,0) = \widetilde{v}(t,\pi) = 0, & \text{for all } t \in [0,T], \\ \widetilde{v}(0,x) = 0, & \text{for all } x \in [0,\pi], \end{array}\right. \tag{1.2.19}$$

and such that $\overline{v}$ is indistinguishable from the continuous version in $H_{-r}([0,\pi])$ of $\widetilde{v}$, which concludes the first part of the proof.

For the second part, under the assumption $f(0) \neq 0$, Theorem 1.3.15 directly shows that (1.2.13) implies (1.1.4). □

1.3 Details of the proof

In the remainder of this work, the letter C will always denote a strictly positive constant whose value may change from line to line. Furthermore, by the Lipschitz continuity of the function f, there exists a positive constant K that we hold fixed from now on such that $|f(x)| \leq K|x| + |f(0)|$ for all $x \in \mathbb{R}$.

1.3.1 Tightness

We start with three lemmas that will provide uniform bounds in $\varepsilon > 0$ for the second moments of u^ε, which will be crucial for proving tightness of $\{(u^\varepsilon, \overline{u}^\varepsilon) \mid \varepsilon > 0\}$ in (Ω^*, τ).

Lemma 1.3.1. *The family $\{u^\varepsilon \mid \varepsilon > 0\}$ of mild solutions to (1.2.1) satisfies*

$$\sup_{\varepsilon > 0} \sup_{(t,x) \in [0,T] \times [0,\pi]} \mathbb{E}\left[|u^\varepsilon(t,x)|^2\right] < \infty \tag{1.3.1}$$

and this uniform bound only depends on the Lipschitz function f.

Proof. Using Itō's isometry and the definition (1.2.5) of u^ε, we have for fixed $\varepsilon > 0$ and $(t,x) \in [0,T] \times [0,\pi]$,

$$\begin{aligned} \mathbb{E}\left[|u^\varepsilon(t,x)|^2\right] &= \mathbb{E}\left[\int_0^t \int_0^\pi \int_{\mathbb{R}} G_{t-s}^2(x,y) \frac{f^2(u^\varepsilon(s,y))}{\sigma^2(\varepsilon)} z^2\, \nu^\varepsilon(\mathrm{d}s, \mathrm{d}y, \mathrm{d}z)\right] \\ &= \mathbb{E}\left[\int_0^t \int_0^\pi G_{t-s}^2(x,y) f^2(u^\varepsilon(s,y))\, \mathrm{d}s\, \mathrm{d}y\right] \left(\frac{1}{\sigma^2(\varepsilon)} \int_{\mathbb{R}} z^2\, Q^\varepsilon(\mathrm{d}z)\right) \\ &= \int_0^t \int_0^\pi G_{t-s}^2(x,y) \mathbb{E}\left[f^2(u^\varepsilon(s,y))\right] \mathrm{d}s\, \mathrm{d}y. \end{aligned}$$

Using the Lipschitz continuity of f and the elementary inequality $(a+b)^2 \leq 2a^2 + 2b^2$, we then obtain for $\varepsilon > 0$ and $(t,x) \in [0,T] \times [0,\pi]$,

$$\mathbb{E}\left[|u^\varepsilon(t,x)|^2\right] \leq C \int_0^t \int_0^\pi G_{t-s}^2(x,y)\mathbb{E}\left[|u^\varepsilon(s,y)|^2\right] \mathrm{d}s\,\mathrm{d}y + C \int_0^t \int_0^\pi G_{t-s}^2(x,y)\,\mathrm{d}s\,\mathrm{d}y. \quad (1.3.2)$$

Now in order to find a bound for $\mathbb{E}[|u^\varepsilon(t,x)|^2]$ uniformly in t, x and ε, we will use a comparison principle for deterministic Volterra equations. By (B.5) in [10], there exists a constant $C > 0$ such that $|G_t(x,y)| \leq C g_t(x-y)$ on $[0,T] \times [0,\pi]^2$, where

$$g_t(x) = \frac{1}{\sqrt{4\pi t}} \exp\left(-\frac{|x|^2}{4t}\right) \mathbb{1}_{\{t \geq 0\}} \quad (1.3.3)$$

is the heat kernel on $\mathbb{R}$. Since $\int_0^T \int_\mathbb{R} |g_t(x)|^q \,\mathrm{d}t\,\mathrm{d}x < \infty$ for all $q < 3$, we obtain

$$\sup_{(t,x)\in[0,T]\times[0,\pi]} \int_0^t \int_0^\pi G_{t-s}^2(x,y)\,\mathrm{d}s\,\mathrm{d}y \leq C \int_0^T \int_\mathbb{R} g_t^2(x)\,\mathrm{d}t\,\mathrm{d}x < \infty.$$

Recall from (1.2.7) that $\mathbb{E}[|u^\varepsilon(t,x)|^2]$ is uniformly bounded in (t,x) for fixed $\varepsilon > 0$. Therefore, by Lemma 6.4 (2) and (3) in [28], the mild solution u^ε satisfies

$$\mathbb{E}\left[|u^\varepsilon(t,x)|^2\right] \leq v(t,x)$$

for all $(t,x) \in [0,T] \times [0,\pi]$ and $\varepsilon > 0$, where v is the unique nonnegative solution of the deterministic Volterra equation

$$\begin{aligned} v(t,x) = C &\int_0^t \int_0^\pi G_{t-s}^2(x,y) v(s,y)\,\mathrm{d}s\,\mathrm{d}y \\ &+ C \int_0^t \int_0^\pi G_{t-s}^2(x,y)\,\mathrm{d}s\,\mathrm{d}y, \quad (t,x) \in [0,T]\times[0,\pi], \end{aligned}$$

and satisfies $\sup_{(t,x)\in[0,T]\times[0,\pi]} v(t,x) < \infty$. □

The next lemma gives an alternative integral representation of u^ε and is an extension of the factorization method in [39] and [90].

Lemma 1.3.2. *For $\delta \in (0, 1/4)$ define*

$$Y_\delta^\varepsilon(t,x) = \int_0^t \int_0^\pi \frac{G_{t-s}(x,y)}{(t-s)^\delta} \frac{f(u^\varepsilon(s,y))}{\sigma(\varepsilon)} L^\varepsilon(\mathrm{d}s,\mathrm{d}y), \quad (t,x) \in [0,T] \times [0,\pi].$$

We then have

$$\sup_{\varepsilon>0} \sup_{(t,x)\in[0,T]\times[0,\pi]} \mathbb{E}\left[|Y_\delta^\varepsilon(t,x)|^2\right] < \infty \quad (1.3.4)$$

and for all $(t,x) \in [0,T] \times [0,\pi]$, the representation

$$u^\varepsilon(t,x) = \frac{\sin(\delta\pi)}{\pi} \int_0^t \int_0^\pi \frac{G_{t-s}(x,y)}{(t-s)^{1-\delta}} Y_\delta^\varepsilon(s,y)\,\mathrm{d}s\,\mathrm{d}y \quad (1.3.5)$$

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

Proof. First, using Itō's isometry, the Lipschitz continuity of f and Lemma 1.3.1, we have

$$\mathbb{E}\left[|Y^\varepsilon_\delta(t,x)|^2\right] = \mathbb{E}\left[\int_0^t\int_0^\pi \frac{G^2_{t-s}(x,y)}{(t-s)^{2\delta}} f^2(u^\varepsilon(s,y))\,\mathrm{d}s\,\mathrm{d}y\right]\left(\frac{1}{\sigma^2(\varepsilon)}\int_{\mathbb{R}} z^2\,Q^\varepsilon(\mathrm{d}z)\right)$$
$$= \int_0^t\int_0^\pi \frac{G^2_{t-s}(x,y)}{(t-s)^{2\delta}}\mathbb{E}\left[f^2(u^\varepsilon(s,y))\right]\,\mathrm{d}s\,\mathrm{d}y \le C\int_0^t\int_0^\pi \frac{G^2_{t-s}(x,y)}{(t-s)^{2\delta}}\,\mathrm{d}s\,\mathrm{d}y.$$

The last integral on the right-hand side is finite if $\delta < 1/4$. Indeed, by (B.5) in [10], it can be bounded by

$$C\int_0^t\int_0^\pi \frac{1}{(t-s)^{2\delta+1}}\exp\left(-\frac{|x-y|^2}{t-s}\right)\,\mathrm{d}s\,\mathrm{d}y = C\int_0^t \frac{1}{(t-s)^{2\delta+1/2}}\,\mathrm{d}s = Ct^{-2\delta+1/2}.$$

The identity (1.3.5) follows in the same way as Lemma 5 in [90]. □

Lemma 1.3.3. *The family* $\{u^\varepsilon \mid \varepsilon > 0\}$ *of mild solutions to (1.2.1) satisfies*

$$\sup_{\varepsilon>0}\mathbb{E}\left[\left(\int_0^T\left(\int_0^\pi |u^\varepsilon(t,x)|^2\,\mathrm{d}x\right)^p\mathrm{d}t\right)^{1/p}\right] < \infty \tag{1.3.6}$$

for all $p \in (1, 4/3)$*.*

Proof. We will use the integral representation (1.3.5) of Lemma 1.3.2. Fix $\delta \in (0, 1/4)$ and $t \in [0,T]$. Using Fubini's theorem, we have

$$\int_0^\pi |u^\varepsilon(t,x)|^2\,\mathrm{d}x$$
$$= C\int_0^\pi\left(\int_0^t\int_0^\pi\int_0^t\int_0^\pi \frac{G_{t-s}(x,y)}{(t-s)^{1-\delta}}\frac{G_{t-s'}(x,y')}{(t-s')^{1-\delta}}Y^\varepsilon_\delta(s,y)Y^\varepsilon_\delta(s',y')\,\mathrm{d}y'\,\mathrm{d}s'\,\mathrm{d}y\,\mathrm{d}s\right)\mathrm{d}x.$$

By the semigroup property of the Green's function, the integral on the right-hand side is equal to

$$\int_0^t\int_0^\pi\int_0^t\int_0^\pi \frac{G_{2t-s-s'}(y,y')}{(t-s)^{1-\delta}(t-s')^{1-\delta}}Y^\varepsilon_\delta(s,y)Y^\varepsilon_\delta(s',y')\,\mathrm{d}y'\,\mathrm{d}s'\,\mathrm{d}y\,\mathrm{d}s$$
$$= 2\int_0^t\int_0^\pi\int_0^s\int_0^\pi \frac{G_{2t-s-s'}(y,y')}{(t-s)^{1-\delta}(t-s')^{1-\delta}}Y^\varepsilon_\delta(s,y)Y^\varepsilon_\delta(s',y')\,\mathrm{d}y'\,\mathrm{d}s'\,\mathrm{d}y\,\mathrm{d}s.$$

Using (B.5) in [10] and (1.3.3), we obtain

$$\int_0^\pi |u^\varepsilon(t,x)|^2\,\mathrm{d}x \le C\int_0^t\int_0^\pi\int_0^s\int_0^\pi \frac{g_{2t-s-s'}(y-y')}{(t-s)^{1-\delta}(t-s')^{1-\delta}}|Y^\varepsilon_\delta(s,y)Y^\varepsilon_\delta(s',y')|\,\mathrm{d}y'\,\mathrm{d}s'\,\mathrm{d}y\,\mathrm{d}s.$$

Now let $p \in (1, 4/3)$, take the $\|\cdot\|_{L^p([0,T])}$-norm of $t \mapsto \int_0^\pi |u^\varepsilon(t,x)|^2\,\mathrm{d}x$ and apply Minkowski's integral inequality to obtain

$$\left(\int_0^T\left(\int_0^\pi |u^\varepsilon(t,x)|^2\,\mathrm{d}x\right)^p\mathrm{d}t\right)^{1/p}$$
$$\le C\int_0^T\int_0^\pi\int_0^s\int_0^\pi\left(\int_0^T\left(\frac{g_{2t-s-s'}(y-y')\mathbb{1}_{\{s\le t\}}}{(t-s)^{1-\delta}\,(t-s')^{1-\delta}}\right)^p\mathrm{d}t\right)^{1/p}|Y^\varepsilon_\delta(s,y)Y^\varepsilon_\delta(s',y')|\,\mathrm{d}y'\,\mathrm{d}s'\,\mathrm{d}y\,\mathrm{d}s.$$

Take expectation, use the Cauchy–Schwarz inequality and (1.3.4) to further obtain

$$\begin{aligned}\mathbb{E}&\left[\left(\int_0^T\left(\int_0^\pi |u^\varepsilon(t,x)|^2\,\mathrm{d}x\right)^p\mathrm{d}t\right)^{1/p}\right]\\&\le C\int_0^T\int_0^\pi\int_0^s\int_0^\pi\left(\int_s^T\left(\frac{g_{2t-s-s'}(y-y')}{(t-s)^{1-\delta}(t-s')^{1-\delta}}\right)^p\mathrm{d}t\right)^{1/p}\mathrm{d}y'\,\mathrm{d}s'\,\mathrm{d}y\,\mathrm{d}s.\end{aligned}\tag{1.3.7}$$

Note that the right-hand side of (1.3.7) does not depend on ε anymore. We now consider the integrand

$$\int_s^T\left(\frac{g_{2t-s-s'}(y-y')}{(t-s)^{1-\delta}(t-s')^{1-\delta}}\right)^p\mathrm{d}t=\int_0^{T-s}\frac{g^p_{2t+s-s'}(y-y')}{t^{(1-\delta)p}(t+s-s')^{(1-\delta)p}}\,\mathrm{d}t.$$

For fixed $x\in\mathbb{R}$, the maximum of the function $t\mapsto g_t(x)$ is $C/|x|$ for some C that is independent of x. Let $\eta\in(0,1)$ and consider the estimate

$$g_t(x)=g_t(x)^{1-\eta}g_t(x)^\eta\le C\frac{1}{|x|^{1-\eta}}\frac{1}{t^{\eta/2}},\quad t>0,\quad x\in\mathbb{R}.$$

Since $s'\le s$, we obtain

$$\begin{aligned}&\left(\int_0^{T-s}\frac{g^p_{2t+s-s'}(y-y')}{t^{(1-\delta)p}(t+s-s')^{(1-\delta)p}}\,\mathrm{d}t\right)^{1/p}\\&\quad\le C\frac{1}{|y-y'|^{1-\eta}}\left(\int_0^T\frac{1}{t^{(1-\delta)p}(t+s-s')^{(1-\delta)p}(2t+s-s')^{p\eta/2}}\,\mathrm{d}t\right)^{1/p}\\&\quad\le C\frac{1}{|y-y'|^{1-\eta}}\left(\int_0^T\frac{1}{t^{((1-\delta)+\eta/2)p}(t+s-s')^{(1-\delta)p}}\,\mathrm{d}t\right)^{1/p}.\end{aligned}$$

Moreover, the integral $\int_0^\pi\int_0^\pi|y-y'|^{\eta-1}\,\mathrm{d}y'\,\mathrm{d}y$ is finite because $\eta>0$, and the expectation in (1.3.7) is now bounded by

$$C\int_0^T\int_0^s\left(\int_0^T\frac{1}{t^{((1-\delta)+\eta/2)p}(t+s-s')^{(1-\delta)p}}\,\mathrm{d}t\right)^{1/p}\mathrm{d}s'\,\mathrm{d}s\tag{1.3.8}$$

for any $\delta\in(0,1/4)$ and $\eta\in(0,1)$. By assumption, $3/4<1/p<1$ and $3/4<(1-\delta)+\eta/2<3/2$. Hence, we can choose δ and η such that $(1-\delta)+\eta/2<1/p$. As a consequence, by Lemma 2 of Chapter 1 in [49], the estimate

$$\int_0^T\frac{1}{t^{((1-\delta)+\eta/2)p}(t+s-s')^{(1-\delta)p}}\,\mathrm{d}t\le C(s-s')^{1-((1-\delta)+\eta/2)p-(1-\delta)p}$$

holds for $0\le s'<s\le T$. We have assumed that $((1-\delta)+\eta/2)p+(1-\delta)p>1$ since otherwise, the last integral is bounded by some constant, which immediately implies that the expectation in (1.3.7) is uniformly bounded in ε.

Therefore, we further estimate the integral in (1.3.8) by

$$C\int_0^T\int_0^s(s-s')^{1/p-2(1-\delta)-\eta/2}\,\mathrm{d}s'\,\mathrm{d}s=C\int_0^T\int_0^s r^{1/p-2(1-\delta)-\eta/2}\,\mathrm{d}r\,\mathrm{d}s,$$

which is finite because our choice of δ and η implies $2(1-\delta)+\eta/2-1/p<1$. This concludes the proof. □

We can now proceed to showing tightness.

Theorem 1.3.4. *The family $\{u^\varepsilon \mid \varepsilon > 0\}$ of mild solutions to (1.2.1) is tight in the Hilbert space $L^2([0,T]\times[0,\pi])$.*

Proof. It is easy to see that the functions

$$\psi_{ij}(t,x) = \overline{\phi}_i(t)\phi_j(x), \quad (t,x) \in [0,T]\times[0,\pi],$$

where $\overline{\phi}_i(t) = \sqrt{2/T}\sin(it\pi/T)$ and $\phi_j(x) = \sqrt{2/\pi}\sin(jx)$ for all $i,j \in \mathbb{N}$, form an orthonormal basis of $L^2([0,T]\times[0,\pi])$.

First, using the stochastic Fubini theorem (see, for example, Theorem 2.6 in [99]), we have for all $i,j\in\mathbb{N}$,

$$\begin{aligned}\langle u^\varepsilon, \psi_{ij}\rangle &= \int_0^T \int_0^\pi \left(\int_0^t \int_0^\pi G_{t-s}(x,y)\frac{f(u^\varepsilon(s,y))}{\sigma(\varepsilon)}\, L^\varepsilon(\mathrm{d}s,\mathrm{d}y)\right)\psi_{ij}(t,x)\,\mathrm{d}t\,\mathrm{d}x \\ &= \frac{1}{\sigma(\varepsilon)}\int_0^T\int_0^\pi f(u^\varepsilon(s,y))\left(\int_s^T\int_0^\pi G_{t-s}(x,y)\psi_{ij}(t,x)\,\mathrm{d}t\,\mathrm{d}x\right) L^\varepsilon(\mathrm{d}s,\mathrm{d}y).\end{aligned} \tag{1.3.9}$$

Define for all $i,j\in\mathbb{N}$,

$$H_{ij}(s,y) = \int_s^T\int_0^\pi G_{t-s}(x,y)\psi_{ij}(t,x)\,\mathrm{d}t\,\mathrm{d}x, \quad (s,y)\in[0,T]\times[0,\pi].$$

Using Fubini's theorem, the expression (1.2.6) of the Green's function G and the orthogonal properties of ϕ_j, we obtain for all $(s,y)\in[0,T]\times[0,\pi]$,

$$H_{ij}(s,y) = \int_s^T \overline{\phi}_i(t)\left(\int_0^\pi G_{t-s}(x,y)\phi_j(x)\,\mathrm{d}x\right)\mathrm{d}t = \int_s^T \overline{\phi}_i(t)\phi_j(y)e^{-j^2(t-s)}\,\mathrm{d}t. \tag{1.3.10}$$

Using the integral formula $\int(\sin ax)e^{bx}\,\mathrm{d}x = (b\sin ax - a\cos ax)\,e^{bx}/(a^2+b^2)+C$, we can further calculate

$$\begin{aligned}H_{ij}(s,y) &= \sqrt{\frac{2}{T}}\phi_j(y)e^{j^2 s}\int_s^T \sin\left(i\frac{\pi}{T}t\right)e^{-j^2 t}\,\mathrm{d}t \\ &= \sqrt{\frac{2}{T}}\phi_j(y)\frac{1}{i^2(\pi/T)^2+j^4}\left(e^{-j^2(T-s)}i\frac{\pi}{T}(-1)^{i+1} + j^2\sin\left(i\frac{\pi}{T}s\right)+i\frac{\pi}{T}\cos\left(i\frac{\pi}{T}s\right)\right) \\ &\le C\left(\frac{i}{i^2+j^4}+\frac{j^2}{i^2+j^4}\right)\le C\frac{1}{i+j^2} \quad \text{for all} \quad i,j\in\mathbb{N}.\end{aligned}$$

For the L^2-norm of H_{ij}, we then have

$$\int_0^T\int_0^\pi H_{ij}^2(s,y)\,\mathrm{d}s\,\mathrm{d}y \le C\frac{1}{i^2+j^4}$$

for all $i,j\in\mathbb{N}$. Since

$$\sum_{i,j=2}^\infty \frac{1}{i^2+j^4} \le \int_1^\infty\int_1^\infty \frac{1}{x^2+y^4}\,\mathrm{d}x\,\mathrm{d}y = \int_1^\infty \frac{\arctan y^2}{y^2}\,\mathrm{d}y < \infty,$$

we obtain from Lemma 1.3.1

$$\sum_{i,j=1}^\infty \sup_{\varepsilon>0}\mathbb{E}\left[\langle u^\varepsilon,\psi_{ij}\rangle^2\right] \le C\sum_{i,j=1}^\infty \int_0^T\int_0^\pi H_{ij}^2(s,y)\,\mathrm{d}s\,\mathrm{d}y < \infty,$$

which implies that

$$\sup_{\varepsilon>0}\mathbb{P}\left(\sum_{i,j\geq N}\langle u^\varepsilon,\psi_{ij}\rangle^2>\delta\right)\leq\frac{1}{\delta}\sum_{i,j\geq N}^{\infty}\sup_{\varepsilon>0}\mathbb{E}\left[\langle u^\varepsilon,\psi_{ij}\rangle^2\right]\longrightarrow 0\quad\text{as}\quad N\to\infty$$

for all $\delta>0$. Moreover, again by Lemma 1.3.1, we have

$$\begin{aligned}\sup_{\varepsilon>0}\mathbb{P}\left(\sum_{i,j<N}\langle u^\varepsilon,\psi_{ij}\rangle^2>\delta\right)&\leq\frac{1}{\delta}\sup_{\varepsilon>0}\mathbb{E}\left[\sum_{i,j<N}\langle u^\varepsilon,\psi_{ij}\rangle^2\right]\leq\frac{1}{\delta}\sup_{\varepsilon>0}\mathbb{E}\left[\sum_{i,j=1}^{\infty}\langle u^\varepsilon,\psi_{ij}\rangle^2\right]\\&=\frac{1}{\delta}\sup_{\varepsilon>0}\mathbb{E}\left[\int_0^T\int_0^\pi u^\varepsilon(t,x)^2\,\mathrm{d}t\,\mathrm{d}x\right]\leq\frac{C}{\delta}\longrightarrow 0\end{aligned}$$

as $\delta\to\infty$ for all $N\in\mathbb{N}$. Therefore, we can conclude from Theorem 1 in [93] that $\{u^\varepsilon\mid\varepsilon>0\}$ is tight in $L^2([0,T]\times[0,\pi])$. □

The next two propositions will imply that $\{\overline{u}^\varepsilon\mid\varepsilon>0\}$ is tight in $D([0,T],H_{-r}([0,\pi]))$.

Proposition 1.3.5. *The càdlàg processes $\{\overline{u}^\varepsilon\mid\varepsilon>0\}$ satisfy the Aldous condition: Let $(\varepsilon_n)_{n\in\mathbb{N}}$ and $(h_n)_{n\in\mathbb{N}}$ be sequences of positive numbers with $\varepsilon_n\to 0$ and $h_n\to 0$ as $n\to\infty$. For each $n\in\mathbb{N}$, further let $\tau_n\in[0,T]$ be a stopping time with respect to the filtration generated by the stochastic process $(\overline{u}^{\varepsilon_n}_t)_{t\leq T}$. Then we have for any $r>1/2$,*

$$\mathbb{E}\left[\|\overline{u}^{\varepsilon_n}_{\tau_n+h_n}-\overline{u}^{\varepsilon_n}_{\tau_n}\|^2_{-r}\right]\longrightarrow 0\quad\textit{as}\quad n\to\infty.$$

Proof. Recall the expression of the dual norm $\|\cdot\|_{-r}$ in (1.2.9). We have

$$\|\overline{u}^{\varepsilon_n}_{\tau_n+h_n}-\overline{u}^{\varepsilon_n}_{\tau_n}\|^2_{-r}=\sum_{k=1}^{\infty}(1+k^2)^{-r}\left(\langle\overline{u}^{\varepsilon_n}_{\tau_n+h_n},\phi_k\rangle-\langle\overline{u}^{\varepsilon_n}_{\tau_n},\phi_k\rangle\right)^2.\tag{1.3.11}$$

We will find a convenient semimartingale decomposition for the real-valued stochastic process $\langle\overline{u}^\varepsilon,\phi_k\rangle=(\langle\overline{u}^\varepsilon_t,\phi_k\rangle)_{t\leq T}$ for any $\varepsilon>0$ and $k\in\mathbb{N}$ that will then allow us to estimate the expectation of the terms appearing in (1.3.11).

First, proceeding as in (1.3.9) and (1.3.10), we have for all $t\leq T$,

$$\begin{aligned}\langle u^\varepsilon(t,\cdot),\phi_k\rangle&=\int_0^\pi\left(\int_0^t\int_0^\pi G_{t-s}(x,y)\frac{f(u^\varepsilon(s,y))}{\sigma(\varepsilon)}\,L^\varepsilon(\mathrm{d}s,\mathrm{d}y)\right)\phi_k(x)\,\mathrm{d}x\\&=\frac{1}{\sigma(\varepsilon)}\int_0^t\int_0^\pi f(u^\varepsilon(s,y))\phi_k(y)e^{-k^2(t-s)}\,L^\varepsilon(\mathrm{d}s,\mathrm{d}y).\end{aligned}$$

If we define

$$X^{k,\varepsilon}_t=\int_0^t\int_0^\pi f(u^\varepsilon(s,y))\phi_k(y)\,L^\varepsilon(\mathrm{d}s,\mathrm{d}y),\quad t\leq T,$$

for all $k\in\mathbb{N}$ and $\varepsilon>0$, then

$$\frac{1}{\sigma(\varepsilon)}\int_0^t\int_0^\pi f(u^\varepsilon(s,y))\phi_k(y)e^{-k^2(t-s)}\,L^\varepsilon(\mathrm{d}s,\mathrm{d}y)=\frac{1}{\sigma(\varepsilon)}\int_0^t e^{-k^2(t-s)}\,X^{k,\varepsilon}(\mathrm{d}s)$$

for all $t\leq T$, where the last term is the Itō integral of the deterministic function $s\mapsto e^{-k^2(t-s)}$ against the square-integrable martingale $X^{k,\varepsilon}$. Because the integrand is a C^∞-function, the integration by parts formula for semimartingales yields

$$\int_0^t e^{-k^2(t-s)}\,X^{k,\varepsilon}(\mathrm{d}s)=X^{k,\varepsilon}_t-\int_0^t X^{k,\varepsilon}_s k^2e^{-k^2(t-s)}\,\mathrm{d}s.$$

Altogether we obtain the semimartingale decomposition

$$\begin{aligned}\langle u^\varepsilon(t,\cdot),\phi_k\rangle &= \int_0^t \int_0^\pi \frac{f(u^\varepsilon(s,y))}{\sigma(\varepsilon)}\phi_k(y)\, L^\varepsilon(\mathrm{d}s,\mathrm{d}y) \\ &\quad - \int_0^t \left(\int_0^s \int_0^\pi \frac{f(u^\varepsilon(r,y))}{\sigma(\varepsilon)}\phi_k(y)\, L^\varepsilon(\mathrm{d}r,\mathrm{d}y) \right) k^2 e^{-k^2(t-s)}\,\mathrm{d}s\end{aligned} \tag{1.3.12}$$

for all $t \le T$, $k \in \mathbb{N}$ and $\varepsilon > 0$.

Now the process $\langle \overline{u}^\varepsilon, \phi_k\rangle$ is the càdlàg version of $(\langle u^\varepsilon(t,\cdot),\phi_k\rangle)_{t\le T}$, so we can infer that $\langle \overline{u}^\varepsilon, \phi_k\rangle$ and the right-hand side of (1.3.12) are indistinguishable since the latter is also càdlàg. Coming back to (1.3.11), we can now decompose

$$\langle \overline{u}^{\varepsilon_n}_{\tau_n+h_n}, \phi_k\rangle - \langle \overline{u}^{\varepsilon_n}_{\tau_n}, \phi_k\rangle = I_{k,n} + J^1_{k,n} + J^2_{k,n},$$

where

$$\begin{aligned} I_{k,n} &= \int_0^T \int_0^\pi \frac{f(u^{\varepsilon_n}(s,y))}{\sigma(\varepsilon_n)}\phi_k(y)\mathbb{1}_{(\tau_n,\tau_n+h_n]}(s)\, L^{\varepsilon_n}(\mathrm{d}s,\mathrm{d}y), \\ J^1_{k,n} &= \int_0^{\tau_n} \left(\int_0^s \int_0^\pi \frac{f(u^{\varepsilon_n}(r,y))}{\sigma(\varepsilon_n)}\phi_k(y)\, L^{\varepsilon_n}(\mathrm{d}r,\mathrm{d}y) \right) k^2 \left(e^{-k^2(\tau_n - s)} - e^{-k^2(\tau_n+h_n-s)} \right) \mathrm{d}s, \\ J^2_{k,n} &= -\int_{\tau_n}^{\tau_n+h_n} \left(\int_0^s \int_0^\pi \frac{f(u^{\varepsilon_n}(r,y))}{\sigma(\varepsilon_n)}\phi_k(y)\, L^{\varepsilon_n}(\mathrm{d}r,\mathrm{d}y) \right) k^2 e^{-k^2(\tau_n+h_n-s)}\,\mathrm{d}s \end{aligned}$$

for all $k, n \in \mathbb{N}$. We now gather some moment estimates for these three terms. First, for the martingale term $I_{k,n}$, we have by Itō's isometry,

$$\begin{aligned}\mathbb{E}[I^2_{k,n}] &= \mathbb{E}\left[\int_0^T \int_0^\pi \int_{\mathbb{R}} f^2(u^{\varepsilon_n}(s,y))\phi_k^2(y)\mathbb{1}_{(\tau_n,\tau_n+h_n]}(s)\frac{z^2}{\sigma^2(\varepsilon_n)}\,\mathrm{d}s\,\mathrm{d}y\, Q^{\varepsilon_n}(\mathrm{d}z)\right] \\ &= \mathbb{E}\left[\int_0^T \int_0^\pi f^2(u^{\varepsilon_n}(s,y))\phi_k^2(y)\mathbb{1}_{(\tau_n,\tau_n+h_n]}(s)\,\mathrm{d}s\,\mathrm{d}y\right].\end{aligned} \tag{1.3.13}$$

Using the Lipschitz continuity of f, we can bound the last term in (1.3.13) by

$$C\mathbb{E}\left[\int_0^T \int_0^\pi u^{\varepsilon_n}(s,y)^2 \mathbb{1}_{(\tau_n,\tau_n+h_n]}(s)\,\mathrm{d}s\,\mathrm{d}y\right] + C\mathbb{E}\left[\int_0^T \int_0^\pi \mathbb{1}_{(\tau_n,\tau_n+h_n]}(s)\,\mathrm{d}s\,\mathrm{d}y\right]$$

for any $k \in \mathbb{N}$. The second term equals $C\pi h_n$, which converges to 0 as $n \to \infty$.

For the first term, choose $p \in (1, 4/3)$. Using Hölder's inequality and Lemma 1.3.3, we obtain

$$\begin{aligned} &\mathbb{E}\left[\int_0^T \left(\int_0^\pi u^{\varepsilon_n}(s,y)^2\,\mathrm{d}y\right)\mathbb{1}_{(\tau_n,\tau_n+h_n]}(s)\,\mathrm{d}s\right] \\ &\quad \le \mathbb{E}\left[\left(\int_0^T\left(\int_0^\pi u^{\varepsilon_n}(s,y)^2\,\mathrm{d}y\right)^p \mathrm{d}s\right)^{1/p}\left(\int_0^T \mathbb{1}_{(\tau_n,\tau_n+h_n]}(s)\,\mathrm{d}s\right)^{1-1/p}\right] \\ &\quad \le h_n^{1-1/p}\sup_{\varepsilon>0}\mathbb{E}\left[\left(\int_0^T\left(\int_0^\pi u^{\varepsilon}(s,y)^2\,\mathrm{d}y\right)^p \mathrm{d}s\right)^{1/p}\right] \longrightarrow 0 \quad \text{as} \quad n\to\infty. \end{aligned}$$

Altogether, this implies $\mathbb{E}[I^2_{k,n}] \to 0$ as $n \to \infty$ for all $k \in \mathbb{N}$.

Next, we have

$$\int_0^{\tau_n} k^2 \left(e^{-k^2(\tau_n-s)} - e^{-k^2(\tau_n+h_n-s)}\right) \mathrm{d}s = \left(1-e^{-k^2h_n}\right)\int_0^{\tau_n} k^2 e^{-k^2(\tau_n-s)}\,\mathrm{d}s$$
$$= \left(1-e^{-k^2h_n}\right)\left(1-e^{-k^2\tau_n}\right) \le 1-e^{-k^2h_n},$$

and by Itō's isometry as well as Lemma 1.3.1,

$$\mathbb{E}\left[\left(\int_0^T\int_0^\pi \frac{f(u^{\varepsilon_n}(s,y))}{\sigma(\varepsilon_n)}\phi_k(y)\,L^{\varepsilon_n}(\mathrm{d}s,\mathrm{d}y)\right)^2\right] = \int_0^T\int_0^\pi \mathbb{E}\left[f^2(u^{\varepsilon_n}(s,y))\right]\phi_k^2(y)\,\mathrm{d}s\,\mathrm{d}y \le C$$

for all $k,n\in\mathbb{N}$. Therefore, by Doob's inequality, $\mathbb{E}[(J^1_{k,n})^2]$ is bounded by

$$\mathbb{E}\left[\sup_{s\le T}\left|\int_0^s\int_0^\pi \frac{f(u^{\varepsilon_n}(r,y))}{\sigma(\varepsilon_n)}\phi_k(y)\,L^{\varepsilon_n}(\mathrm{d}r,\mathrm{d}y)\right|^2\left(\int_0^{\tau_n} k^2\left(e^{-k^2(\tau_n-s)}-e^{-k^2(\tau_n+h_n-s)}\right)\mathrm{d}s\right)^2\right]$$
$$\le \left(1-e^{-k^2h_n}\right)^2\mathbb{E}\left[\sup_{s\le T}\left|\int_0^s\int_0^\pi \frac{f(u^{\varepsilon_n}(r,y))}{\sigma(\varepsilon_n)}\phi_k(y)\,L^{\varepsilon_n}(\mathrm{d}r,\mathrm{d}y)\right|^2\right]$$
$$\le C\left(1-e^{-k^2h_n}\right)^2 \longrightarrow 0 \quad \text{as} \quad n\to\infty$$

for all $k\in\mathbb{N}$.

Finally, with $\int_{\tau_n}^{\tau_n+h_n} k^2 e^{-k^2(\tau_n+h_n-s)}\,\mathrm{d}s = 1-e^{-k^2h_n}$ and similar calculations, $\mathbb{E}[(J^2_{k,n})^2]$ is bounded by

$$\mathbb{E}\left[\sup_{s\le T}\left|\int_0^s\int_0^\pi \frac{f(u^{\varepsilon_n}(r,y))}{\sigma(\varepsilon_n)}\phi_k(y)\,L^{\varepsilon_n}(\mathrm{d}r,\mathrm{d}y)\right|^2\left(\int_{\tau_n}^{\tau_n+h_n} k^2 e^{-k^2(\tau_n+h_n-s)}\,\mathrm{d}s\right)^2\right]$$
$$= \left(1-e^{-k^2h_n}\right)^2\mathbb{E}\left[\sup_{s\le T}\left|\int_0^s\int_0^\pi \frac{f(u^{\varepsilon_n}(r,y))}{\sigma(\varepsilon_n)}\phi_k(y)\,L^{\varepsilon_n}(\mathrm{d}r,\mathrm{d}y)\right|^2\right]$$
$$\le C\left(1-e^{-k^2h_n}\right)^2 \longrightarrow 0 \quad \text{as} \quad n\to\infty$$

for all $k\in\mathbb{N}$. As a consequence, recalling that $r>1/2$ and thus $\sum_{k=1}^\infty(1+k^2)^{-r}<\infty$, we obtain by (1.3.11) and dominated convergence,

$$\mathbb{E}\left[\|\overline{u}^{\varepsilon_n}_{\tau_n+h_n}-\overline{u}^{\varepsilon_n}_{\tau_n}\|^2_{-r}\right] = \sum_{k=1}^\infty(1+k^2)^{-r}\mathbb{E}\left[\left(\langle\overline{u}^{\varepsilon_n}_{\tau_n+h_n},\phi_k\rangle-\langle\overline{u}^{\varepsilon_n}_{\tau_n},\phi_k\rangle\right)^2\right]$$
$$\le 3\sum_{k=1}^\infty(1+k^2)^{-r}\left(\mathbb{E}\left[I^2_{k,n}\right]+\mathbb{E}\left[(J^1_{k,n})^2\right]+\mathbb{E}\left[(J^2_{k,n})^2\right]\right)\longrightarrow 0$$

as $n\to\infty$, which is the assertion of the proposition. □

Proposition 1.3.6. *For any fixed $t\le T$ and $r>1/2$, the random elements $\{\overline{u}^\varepsilon_t \mid \varepsilon>0\}$ are tight in $H_{-r}([0,\pi])$.*

Proof. Proceeding as in the proof of Proposition 1.3.5 and using Lemma 1.3.1, we have

$$\mathbb{E}\left[\langle\overline{u}^\varepsilon_t,\phi_k\rangle^2\right] = \int_0^t\int_0^\pi \mathbb{E}\left[f^2(u^\varepsilon(s,y))\right]\phi_k^2(y)e^{-2k^2(t-s)}\,\mathrm{d}s\,\mathrm{d}y \le C\frac{1}{k^2}\left(1-e^{-2k^2t}\right)$$

for any $k \in \mathbb{N}$ and $t \leq T$. Hence, we have for all $q < 1/2$, $t \leq T$ and $\varepsilon > 0$,

$$\mathbb{E}\left[\|\overline{u}_t^\varepsilon\|_q^2\right] = \sum_{k=1}^{\infty}(1+k^2)^q \mathbb{E}\left[\langle \overline{u}_t^\varepsilon, \phi_k\rangle^2\right] \leq C\sum_{k=1}^{\infty}(1+k^2)^q \frac{1}{k^2} < \infty,$$

and thus $\overline{u}_t^\varepsilon \in H_q([0,\pi])$ $\mathbb{P}$-almost surely.

Because the penultimate term in the inequality above does not depend on ε, by Markov's inequality, we can further deduce

$$\lim_{\delta\to\infty}\sup_{\varepsilon>0}\mathbb{P}\left(\|\overline{u}_t^\varepsilon\|_q > \delta\right) = 0$$

for all $q < 1/2$ and $t \leq T$. Since the embeddings

$$H_q([0,\pi]) \hookrightarrow L^2([0,\pi]) \hookrightarrow H_{-r}([0,\pi])$$

are compact for $0 < q < 1/2 < r$ by Theorem 4.58 in [44], it follows that $\{\overline{u}_t^\varepsilon \mid \varepsilon > 0\}$ is tight in $H_{-r}([0,\pi])$ for any fixed $t \leq T$ and $r > 1/2$. □

Theorem 1.3.7. *For any $r > 1/2$, the càdlàg modifications $\{\overline{u}^\varepsilon \mid \varepsilon > 0\}$ are tight in the Skorokhod space $D([0,T], H_{-r}([0,\pi]))$.*

Proof. By Theorem 6.8 in [99], this is a direct consequence of Propositions 1.3.5 and 1.3.6. □

1.3.2 Characterization of the limit

After proving tightness in Section 1.3.1, our next goal is to characterize the limit distribution of weakly converging subsequences. Following the outline of the proof of Theorem 1.2.1, the first step is to show that under condition (1.1.4) on the Lévy measure Q^ε, the process $\overline{M}$ in (1.2.18) is a martingale with respect to the filtration $\overline{\boldsymbol{F}}$ defined in (1.2.16). In order to achieve this result, which is Theorem 1.3.12 below, we prove that the pairs $(u^\varepsilon, \overline{u}^\varepsilon)$ satisfy related martingale problems (Theorem 1.3.8) and that these "converge" as $\varepsilon \to 0$ (Theorem 1.3.9).

Recall that for all test functions $\phi \in C_c^\infty((0,\pi))$ and fixed $t \leq T$,

$$\int_0^\pi u^\varepsilon(t,x)\phi(x)\,\mathrm{d}x = \int_0^t\int_0^\pi u^\varepsilon(s,x)\phi''(x)\,\mathrm{d}s\,\mathrm{d}x + \int_0^t\int_0^\pi \frac{f(u^\varepsilon(s,x))}{\sigma(\varepsilon)}\phi(x)\,L^\varepsilon(\mathrm{d}s,\mathrm{d}x) \tag{1.3.14}$$

$\mathbb{P}$-almost surely. This follows, in a similar way to Theorem 3.2 in [99], from the fact that in our situation, we may apply the stochastic Fubini theorem; see, for example, Theorem 2.6 in [99].

Theorem 1.3.8. *For each $\varepsilon > 0$, the pair $(u^\varepsilon, \overline{u}^\varepsilon)$ where $u^\varepsilon \in L^2([0,T]\times[0,\pi])$ is the mild solution to the stochastic heat equation (1.2.1) and $\overline{u}^\varepsilon$ is its càdlàg modification in $H_{-r}([0,\pi])$, with $r > 1/2$, satisfies the following martingale problem. For all $\xi \in \mathbb{R}$ and $\phi \in C_c^\infty((0,\pi))$, the complex-valued stochastic process*

$$\begin{aligned} M_t^\varepsilon = {} & e^{i\xi\langle\overline{u}_t^\varepsilon,\phi\rangle} - i\xi\int_0^t e^{i\xi\langle\overline{u}_s^\varepsilon,\phi\rangle}\langle u^\varepsilon(s,\cdot),\phi''\rangle\,\mathrm{d}s \\ & - \int_0^t\int_0^\pi\int_{\mathbb{R}} e^{i\xi\langle\overline{u}_s^\varepsilon,\phi\rangle}\left(e^{i\xi\frac{f(u^\varepsilon(s,x))}{\sigma(\varepsilon)}\phi(x)z} - 1 - i\xi\frac{f(u^\varepsilon(s,x))}{\sigma(\varepsilon)}\phi(x)z\right)\mathrm{d}s\,\mathrm{d}x\,Q^\varepsilon(\mathrm{d}z), \end{aligned} \tag{1.3.15}$$

with $t \leq T$, is a square-integrable $\boldsymbol{F}$-martingale with the uniform bound

$$\sup_{\varepsilon>0}\sup_{t\leq T}\mathbb{E}\left[|M_t^\varepsilon|^2\right] < \infty. \tag{1.3.16}$$

Proof. First, since $\overline{u}^\varepsilon$ is the càdlàg version of u^ε, the stochastic process $\langle \overline{u}^\varepsilon, \phi\rangle$ is indistinguishable from the right-hand side of (1.3.14) for any $\phi \in C_c^\infty((0,\pi))$. This directly implies that $\langle \overline{u}^\varepsilon, \phi\rangle$ is an $\boldsymbol{F}$-semimartingale without continuous martingale part. Furthermore, one can easily verify that the $\boldsymbol{F}$-compensator of the jump measure μ^ε_ϕ of $\langle \overline{u}^\varepsilon, \phi\rangle$ (on $[0,T]\times\mathbb{R}$) is given by

$$\nu^\varepsilon_\phi(A) = \int_0^T \int_0^\pi \int_{\mathbb{R}} \mathbb{1}_A\left(t, \frac{f(u^\varepsilon(t,x))}{\sigma(\varepsilon)}\phi(x)z\right) \mathrm{d}t\,\mathrm{d}x\,Q^\varepsilon(\mathrm{d}z), \quad A \in \mathcal{B}([0,T]\times\mathbb{R}). \tag{1.3.17}$$

As a consequence, using Itō's formula (see, for example, Theorem I.4.57 in [65]), (1.3.14) and the fact that $u^\varepsilon(0,x) = 0$, we have

$$\begin{aligned} e^{i\xi\langle \overline{u}^\varepsilon_t, \phi\rangle} &= 1 + i\xi \int_0^t e^{i\xi\langle \overline{u}^\varepsilon_s, \phi\rangle} \langle u^\varepsilon(s,\cdot), \phi''\rangle\,\mathrm{d}s + i\xi \int_0^t \int_0^\pi e^{i\xi\langle \overline{u}^\varepsilon_{s-}, \phi\rangle} \frac{f(u^\varepsilon(s,x))}{\sigma(\varepsilon)}\phi(x)\,L^\varepsilon(\mathrm{d}s,\mathrm{d}x) \\ &\quad + \int_0^t \int_{\mathbb{R}} e^{i\xi\langle \overline{u}^\varepsilon_{s-}, \phi\rangle} (e^{i\xi x} - 1 - i\xi x)\,\mu^\varepsilon_\phi(\mathrm{d}s,\mathrm{d}x), \end{aligned}$$

and therefore, by (1.3.15),

$$\begin{aligned} M^\varepsilon_t &= 1 + i\xi \int_0^t \int_0^\pi e^{i\xi\langle \overline{u}^\varepsilon_{s-}, \phi\rangle} \frac{f(u^\varepsilon(s,x))}{\sigma(\varepsilon)}\phi(x)\,L^\varepsilon(\mathrm{d}s,\mathrm{d}x) \\ &\quad + \int_0^t \int_{\mathbb{R}} e^{i\xi\langle \overline{u}^\varepsilon_{s-}, \phi\rangle} (e^{i\xi x} - 1 - i\xi x)\,(\mu^\varepsilon_\phi - \nu^\varepsilon_\phi)(\mathrm{d}s,\mathrm{d}x) \end{aligned} \tag{1.3.18}$$

for all $t \le T$. The two integral processes on the right-hand side of (1.3.18) are square-integrable $\boldsymbol{F}$-martingales (for the second, this is implied by the elementary inequalities $(\cos(x) - 1)^2 \le x^2$ and $(\sin(x) - x)^2 \le 4x^2$ for all $x \in \mathbb{R}$, together with Lemma 1.3.1) and hence, this is also the case for M^ε. By Itō's isometry, we further obtain

$$\mathbb{E}\left[\left|\int_0^t \int_{\mathbb{R}} e^{i\xi\langle \overline{u}^\varepsilon_s, \phi\rangle} \frac{f(u^\varepsilon(s,x))}{\sigma(\varepsilon)}\phi(x)\,L^\varepsilon(\mathrm{d}s,\mathrm{d}x)\right|^2\right] = \mathbb{E}\left[\int_0^t \int_0^\pi f^2(u^\varepsilon(s,x))\phi^2(x)\,\mathrm{d}s\,\mathrm{d}x\right]$$

as well as

$$\begin{aligned} &\mathbb{E}\left[\left|\int_0^t \int_{\mathbb{R}} e^{i\xi\langle \overline{u}^\varepsilon_s, \phi\rangle} (e^{i\xi x} - 1 - i\xi x)\,(\mu^\varepsilon_\phi - \nu^\varepsilon_\phi)(\mathrm{d}s,\mathrm{d}x)\right|^2\right] \\ &\qquad = \mathbb{E}\left[\int_0^t \int_{\mathbb{R}} |e^{i\xi x} - 1 - i\xi x|^2\,\nu^\varepsilon_\phi(\mathrm{d}s,\mathrm{d}x)\right] \\ &\qquad = \mathbb{E}\left[\int_0^t \int_{\mathbb{R}} \left((\cos(\xi x) - 1)^2 + (\sin(\xi x) - \xi x)^2\right)\,\nu^\varepsilon_\phi(\mathrm{d}s,\mathrm{d}x)\right] \end{aligned}$$

for all $t \le T$. We estimate the last expectation, using the elementary inequalities given above as well as the definition of ν^ε_ϕ, by

$$\begin{aligned} 5\xi^2\mathbb{E}\left[\int_0^t \int_{\mathbb{R}} x^2\,\nu^\varepsilon_\phi(\mathrm{d}s,\mathrm{d}x)\right] &= 5\xi^2\mathbb{E}\left[\int_0^t \int_0^\pi \frac{f^2(u^\varepsilon(s,x))}{\sigma^2(\varepsilon)}\phi^2(x)z^2\,\mathrm{d}s\,\mathrm{d}x\,Q^\varepsilon(\mathrm{d}z)\right] \\ &= 5\xi^2\mathbb{E}\left[\int_0^t \int_0^\pi f^2(u^\varepsilon(s,x))\phi^2(x)\,\mathrm{d}s\,\mathrm{d}x\right], \quad t \le T. \end{aligned}$$

Altogether, we obtain (1.3.16) from (1.3.18), the Lipschitz continuity of f and Lemma 1.3.1. □

We now switch to the probability space $(\overline{\Omega}, \overline{\mathcal{F}}, \overline{\boldsymbol{F}}, \overline{\mathbb{P}})$ from the Skorokhod construction in (1.2.14) and define the process $\overline{M}^k$ in the same way as M^ε in (1.3.15), but with $(u^\varepsilon, \overline{u}^\varepsilon)$ and Q^ε replaced by $(v^k, \overline{v}^k)$ in (1.2.15) and Q^{ε_k}, respectively.

Theorem 1.3.9. *Under* (1.1.4)*, we have pointwise on* $\overline{\Omega}$ *for any* $\xi \in \mathbb{R}$ *and* $\phi \in C_c^\infty((0,\pi))$*,*

$$\overline{M}^k \longrightarrow \overline{M} \quad \textit{as} \quad k \to \infty$$

in the Skorokhod space $D([0,T], \mathbb{C})$*, where* $\overline{M}$ *is the process in* (1.2.18)*.*

In order to prove this result, we first rewrite $\overline{M}^k$ in a more convenient form. For any fixed $k \in \mathbb{N}$, $\xi \in \mathbb{R}$ and $\phi \in C_c^\infty((0,\pi))$, let

$$\begin{aligned} \overline{\nu}^k(A) &= \int_0^T \int_0^\pi \int_{\mathbb{R}} \mathbb{1}_A \left(t, \frac{f(v^k(t,x))}{\sigma(\varepsilon_k)} \phi(x) z \right) \mathrm{d}t \, \mathrm{d}x \, Q^{\varepsilon_k}(\mathrm{d}z), \\ \overline{B}_t^k &= \int_0^t \langle v^k(s,\cdot), \phi'' \rangle \, \mathrm{d}s - \int_0^t \int_{\mathbb{R}} x \mathbb{1}_{\{|x|>1\}} \, \overline{\nu}^k(\mathrm{d}s, \mathrm{d}x), \\ \overline{A}_t^k &= i\xi \overline{B}_t^k + \int_0^t \int_{\mathbb{R}} \left(e^{i\xi x} - 1 - i\xi x \mathbb{1}_{\{|x|\le 1\}} \right) \overline{\nu}^k(\mathrm{d}s, \mathrm{d}x) \end{aligned} \tag{1.3.19}$$

for all $A \in \mathcal{B}([0,T] \times \mathbb{R})$ and $t \le T$. Note that $(\overline{B}^k, 0, \overline{\nu}^k)$ are the predictable characteristics of the $\overline{\boldsymbol{F}}$-semimartingale $\langle \overline{v}^k, \phi \rangle$ and that they are functions of the random field v^k and not of $\langle \overline{v}^k, \phi \rangle$ itself (which is another reason why we have adopted a dual view on the solutions to (1.2.1) and (1.2.8) as elements of Ω^*). The process $\overline{M}^k$ introduced above can thus be written as

$$\overline{M}_t^k = e^{i\xi \langle \overline{v}_t^k, \phi \rangle} - \int_0^t e^{i\xi \langle \overline{v}_s^k, \phi \rangle} \, \overline{A}^k(\mathrm{d}s), \quad t \le T. \tag{1.3.20}$$

Define the truncation functions

$$\varrho_h \colon \mathbb{R} \longrightarrow \mathbb{R}, \quad x \mapsto x \mathbb{1}_{\{|x| \le h\}}, \quad h > 0. \tag{1.3.21}$$

The key idea of the proof of Theorem 1.3.9 is to see that for fixed $\overline{\omega} \in \overline{\Omega}$, $t \le T$ and $\phi \in C_c^\infty((0,\pi))$, the function $\overline{A}_t^k$ in (1.3.19) (resp., $\overline{A}_t$ in (1.2.17)) is the Lévy exponent of the infinitely divisible distribution η_k (resp., η) with characteristics $(\overline{B}_t^k, 0, \overline{\nu}^k([0,t] \times \mathrm{d}x))$ (resp., $(\overline{B}_t, \overline{C}_t, 0)$) with respect to ϱ_1. Then we can make use of the following result, which is the only place in this work where (1.1.4) will actually be needed.

Theorem 1.3.10. *If* (1.1.4) *holds, then for any* $\phi \in C_c^\infty((0,\pi))$*,* $t \in [0,T]$ *and* $\overline{\omega} \in \overline{\Omega}$*, we have*

$$\eta_k \xrightarrow{w} \eta \quad \textit{as} \quad k \to \infty.$$

Proof. As in the proof of Theorem 2.2 in [36] (see also Theorem 2.1 in [9]), it suffices to show that

$$\begin{aligned} &(i) \quad \int_{|x|\le h} x^2 \, \overline{\nu}^k([0,t] \times \mathrm{d}x) \longrightarrow \overline{C}_t, \\ &(ii) \quad \overline{B}_t^k \longrightarrow \overline{B}_t, \\ &(iii) \quad \overline{\nu}^k\left([0,t] \times \{|x| > 1\}\right) \longrightarrow 0 \end{aligned} \tag{1.3.22}$$

as $k \to \infty$ for all $h > 0$. Starting with (i), we have

$$\int_{|x|\le h} x^2\, \overline{\nu}^k([0,t]\times \mathrm{d}x) \\ = \int_0^t \int_0^\pi f^2(v^k(s,x))\phi^2(x) \frac{1}{\sigma^2(\varepsilon_k)} \int_{\mathbb{R}} z^2 \mathbb{1}_{\left\{|z|\le \left(h/|f(v^k(s,x))\phi(x)|\right)\sigma(\varepsilon_k)\right\}}\, Q^{\varepsilon_k}(\mathrm{d}z)\,\mathrm{d}s\,\mathrm{d}x.$$

We can ignore all points in the domain of integration where $|f(v^k(s,x))\phi(x)| = 0$. So if we let

$$I_h^k(s,x) = \frac{1}{\sigma^2(\varepsilon_k)} \int_{\mathbb{R}} z^2 \mathbb{1}_{\left\{|z|\le \left(h/|f(v^k(s,x))\phi(x)|\right)\sigma(\varepsilon_k)\right\}}\, Q^{\varepsilon_k}(\mathrm{d}z) \quad \text{and} \\ \Sigma_h^k(s,x) = 1 - I_h^k(s,x) = \frac{1}{\sigma^2(\varepsilon_k)} \int_{\mathbb{R}} z^2 \mathbb{1}_{\left\{|z|> \left(h/|f(v^k(s,x))\phi(x)|\right)\sigma(\varepsilon_k)\right\}}\, Q^{\varepsilon_k}(\mathrm{d}z) \tag{1.3.23}$$

for any $(s,x) \in [0,T]\times[0,\pi]$, $h > 0$ and $k \in \mathbb{N}$, then, using the triangle inequality and the fact that $0 < I_h^k(s,x) \le 1$, we obtain

$$\left| \int_{|x|\le h} x^2\, \overline{\nu}^k([0,t]\times \mathrm{d}x) - \overline{C}_t \right| \\ \le \left| \int_0^t \int_0^\pi \left(f^2(v^k(s,x)) - f^2(v(s,x)) \right) \phi^2(x) I_h^k(s,x)\,\mathrm{d}s\,\mathrm{d}x \right| \\ + \left| \int_0^t \int_0^\pi f^2(v(s,x))\phi^2(x) \left(I_h^k(s,x) - 1 \right) \mathrm{d}s\,\mathrm{d}x \right| \\ \le \int_0^t \int_0^\pi \left| f^2(v^k(s,x)) - f^2(v(s,x)) \right| \phi^2(x)\,\mathrm{d}s\,\mathrm{d}x + \int_0^t \int_0^\pi f^2(v(s,x))\phi^2(x)\Sigma_h^k(s,x)\,\mathrm{d}s\,\mathrm{d}x. \tag{1.3.24}$$

By the Lipschitz continuity of f and Hölder's inequality, we have for the first integral on the right-hand side of (1.3.24),

$$\int_0^t \int_0^\pi \left| f^2(v^k(s,x)) - f^2(v(s,x)) \right| \phi^2(x)\,\mathrm{d}s\,\mathrm{d}x \\ \le C \int_0^t \int_0^\pi \left| f(v^k(s,x)) - f(v(s,x)) \right| \left| f(v^k(s,x)) + f(v(s,x)) \right| \mathrm{d}s\,\mathrm{d}x \\ \le C \left(\int_0^t \int_0^\pi \left(v^k(s,x) - v(s,x) \right)^2 \mathrm{d}s\,\mathrm{d}x \right)^{1/2} \left(\int_0^t \int_0^\pi \left(f(v^k(s,x)) + f(v(s,x)) \right)^2 \mathrm{d}s\,\mathrm{d}x \right)^{1/2}. \tag{1.3.25}$$

By (1.2.15), $v^k \longrightarrow v$ in $L^2([0,T]\times[0,\pi])$ pointwise on $\overline{\Omega}$. Hence, the sequence $(v^k)_{k\in\mathbb{N}}$ is bounded in $L^2([0,T]\times[0,\pi])$ and we have

$$\int_0^t \int_0^\pi \left(f(v^k(s,x)) + f(v(s,x)) \right)^2 \mathrm{d}s\,\mathrm{d}x \\ \le C \left(1 + \sup_{k\in\mathbb{N}} \int_0^t \int_0^\pi v^k(s,x)^2\,\mathrm{d}s\,\mathrm{d}x + \int_0^t \int_0^\pi v(s,x)^2\,\mathrm{d}s\,\mathrm{d}x \right) < \infty, \tag{1.3.26}$$

which implies $\int_0^t \int_0^\pi |f^2(v^k(s,x)) - f^2(v(s,x))|\phi^2(x)\,\mathrm{d}s\,\mathrm{d}x \longrightarrow 0$ as $k \to \infty$.

The second integral in (1.3.24) is more difficult to handle. We decompose it into $I_1^{k,n} + I_2^{k,n,M} + I_3^{k,n,M}$, where

$$\begin{aligned} I_1^{k,n} &= \int_0^t \int_0^\pi f^2(v(s,x))\Sigma_h^k(s,x)\mathbb{1}_{\{|v^k(s,x)|\le n\}}\,\mathrm{d}s\,\mathrm{d}x, \\ I_2^{k,n,M} &= \int_0^t \int_0^\pi f^2(v(s,x))\Sigma_h^k(s,x)\mathbb{1}_{\{|v^k(s,x)|>n\}}\mathbb{1}_{\{|f(v(s,x))|\le M\}}\,\mathrm{d}s\,\mathrm{d}x, \\ I_3^{k,n,M} &= \int_0^t \int_0^\pi f^2(v(s,x))\Sigma_h^k(s,x)\mathbb{1}_{\{|v^k(s,x)|>n\}}\mathbb{1}_{\{|f(v(s,x))|>M\}}\,\mathrm{d}s\,\mathrm{d}x \end{aligned}$$

for all $k, n, M \in \mathbb{N}$.

Again we will study each of these three integrals separately. On the set $\{|v^k(s,x)| \le n\}$, we have $|f(v^k(s,x))\phi(x)| \le (Kn + |f(0)|)\|\phi\|_\infty$ and therefore

$$\mathbb{1}_{\{|z|>\left(h/|f(v^k(s,x))\phi(x)|\right)\sigma(\varepsilon_k)\}} \le \mathbb{1}_{\{|z|>\left(h/(Kn+|f(0)|)\|\phi\|_\infty\right)\sigma(\varepsilon_k)\}}.$$

Thus,

$$I_1^{k,n} \le \int_0^t \int_0^\pi f^2(v(s,x))\frac{1}{\sigma^2(\varepsilon_k)}\int_{\mathbb{R}} z^2 \mathbb{1}_{\{|z|>\left(h/(Kn+|f(0)|)\|\phi\|_\infty\right)\sigma(\varepsilon_k)\}}\,Q^{\varepsilon_k}(\mathrm{d}z)\,\mathrm{d}s\,\mathrm{d}x.$$

Because the term $h/(Kn + |f(0)|)\|\phi\|_\infty$ does not depend on k, we can use condition (1.1.4), whence

$$\frac{1}{\sigma(\varepsilon_k)}\int_{\mathbb{R}} z^2 \mathbb{1}_{\{|z|>\left(h/(Kn+|f(0)|)\|\phi\|_\infty\right)\sigma(\varepsilon_k)\}}\,Q^{\varepsilon_k}(\mathrm{d}z) \longrightarrow 0 \quad \text{as} \quad k\to\infty$$

for all $n \in \mathbb{N}$ and $h > 0$. Since $v \in L^2([0,T]\times[0,\pi])$, we obtain by dominated convergence that $I_1^{k,n} \longrightarrow 0$ as $k \to \infty$ for all $n \in \mathbb{N}$.

Next, we have by Chebyshev's inequality,

$$I_2^{k,n,M} \le M^2 \int_0^t \int_0^\pi \mathbb{1}_{\{|v^k(s,x)|>n\}}\,\mathrm{d}s\,\mathrm{d}x \le \frac{M^2}{n^2}\sup_{k\in\mathbb{N}}\int_0^t \int_0^\pi v^k(s,x)^2\,\mathrm{d}s\,\mathrm{d}x,$$

which tends to 0 as $n \to \infty$, uniformly in k.

Finally, we have by dominated convergence,

$$I_3^{k,n,M} \le \int_0^t \int_0^\pi f^2(v(s,x))\mathbb{1}_{\{|f(v(s,x))|>M\}}\,\mathrm{d}s\,\mathrm{d}x \longrightarrow 0 \quad \text{as} \quad M\to\infty,$$

uniformly in n and k. Altogether, we have just shown that the left-hand side of (1.3.24) converges to 0 as $k \to \infty$, which is condition (i) in (1.3.22).

The two other conditions will follow from our last calculations. Indeed, we have

$$|\overline{B}_t^k - \overline{B}_t| \le \int_0^t \left|\langle v^k(s,\cdot),\phi''\rangle - \langle v(s,\cdot),\phi''\rangle\right|\mathrm{d}s + \left|\int_0^t \int_{\mathbb{R}} x\mathbb{1}_{\{|x|>1\}}\,\overline{\nu}^k(\mathrm{d}s,\mathrm{d}x)\right|,$$

where the first term vanishes because

$$\begin{aligned} \int_0^t \left|\langle v_s^k,\phi''\rangle - \langle v_s,\phi''\rangle\right|\mathrm{d}s &\le \int_0^t \int_0^\pi \left|v^k(s,x) - v(s,x)\right||\phi''(x)|\,\mathrm{d}s\,\mathrm{d}x \\ &\le C\left(\int_0^t \int_0^\pi \left(v^k(s,x)-v(s,x)\right)^2\mathrm{d}s\,\mathrm{d}x\right)^{1/2} \longrightarrow 0 \end{aligned} \tag{1.3.27}$$

as $k \to \infty$. Furthermore,

$$\begin{aligned}\int_0^t \int_{\mathbb{R}} x \mathbb{1}_{\{|x|>1\}} \,\overline{\nu}^k(\mathrm{d}s, \mathrm{d}x) &\leq \int_0^t \int_{\mathbb{R}} x^2 \mathbb{1}_{\{|x|>1\}} \,\overline{\nu}^k(\mathrm{d}s, \mathrm{d}x) \\ &= \int_0^t \int_0^\pi f^2(v^k(s,x))\phi^2(x)\Sigma_1^k(s,x)\,\mathrm{d}s\,\mathrm{d}x \\ &\leq \int_0^t \int_0^\pi \left|f^2(v^k(s,x)) - f^2(v(s,x))\right|\phi^2(x)\,\mathrm{d}s\,\mathrm{d}x \\ &\quad + \int_0^t \int_0^\pi f^2(v(s,x))\phi^2(x)\Sigma_1^k(s,x)\,\mathrm{d}s\,\mathrm{d}x.\end{aligned} \tag{1.3.28}$$

These integrals are exactly the same as in the last line of (1.3.24), so we obtain $|\overline{B}_t^k - \overline{B}_t| \longrightarrow 0$, which is condition (ii). From this, condition (iii) immediately follows since

$$\overline{\nu}^k\left([0,t] \times \{|x|>1\}\right) = \int_0^t \int_{\mathbb{R}} \mathbb{1}_{\{|x|>1\}} \,\overline{\nu}^k(\mathrm{d}s, \mathrm{d}x) \leq \int_0^t \int_{\mathbb{R}} x \mathbb{1}_{\{|x|>1\}} \,\overline{\nu}^k(\mathrm{d}s, \mathrm{d}x). \tag{1.3.29}$$

□

The following technical lemma is a direct consequence of Theorem 1.3.10 and will be crucial for proving Theorem 1.3.9 afterwards.

If $t \mapsto A_t$ is a function of locally finite variation, we denote by $\mathrm{Var}(A)_t$ the total variation of the function A on the interval $[0,t]$. If A is complex-valued, we have $\mathrm{Var}(A) = \mathrm{Var}(\mathrm{Re}\,A) + \mathrm{Var}(\mathrm{Im}\,A)$.

Lemma 1.3.11. *If* (1.1.4) *holds, then we have pointwise on* $\overline{\Omega}$,

$$\overline{A}_t^k \longrightarrow \overline{A}_t \quad \textit{and} \quad \mathrm{Var}(\overline{A}^k - \overline{A})_t \longrightarrow 0 \quad \textit{as} \quad k \to \infty$$

for any $\xi \in \mathbb{R}$, $\phi \in C_c^\infty((0,\pi))$ *and* $t \leq T$, *where the processes* $\overline{A}^k$ *and* $\overline{A}$ *are defined in* (1.3.19) *and* (1.2.17), *respectively.*

Proof. For fixed $\phi \in C_c^\infty((0,\pi))$, $t \in [0,T]$ and $\overline{\omega} \in \overline{\Omega}$, the infinitely divisible distributions η_k and η, defined before Theorem 1.3.10, have Lévy exponents $\overline{A}_t^k$ and $\overline{A}_t$, respectively. By that theorem, $\eta_k \xrightarrow{w} \eta$ as $k \to \infty$. This immediately implies the first claim of the proposition (see, for example, Equation VII.2.6 in [65]).

For the second claim, we will need the truncation function

$$\vartheta(x) = \begin{cases} -1, & x < -1, \\ x, & |x| \leq 1, \\ 1, & x > 1. \end{cases}$$

The main difference between the function $\varrho_1(x) = x\mathbb{1}_{\{|x|\leq 1\}}$, used so far, and ϑ is that the latter is continuous. Since this property will be needed for technical reasons, we replace ϱ_1 by ϑ in the expression of $\overline{A}_t^k$ in (1.3.19) and thus obtain

$$\begin{aligned}\overline{A}_t^k &= i\xi\left(\int_0^t \langle v^k(s,\cdot), \phi''\rangle\,\mathrm{d}s - \int_0^t \int_{\mathbb{R}} (x - \vartheta(x))\,\overline{\nu}^k(\mathrm{d}s, \mathrm{d}x)\right) \\ &\quad + \int_0^t \int_{\mathbb{R}} \left(e^{i\xi x} - 1 - i\xi\vartheta(x)\right)\overline{\nu}^k(\mathrm{d}s, \mathrm{d}x).\end{aligned} \tag{1.3.30}$$

With (1.3.30) and (1.2.17), we then calculate

$$\begin{aligned}\operatorname{Re}(\overline{A}_t^k - \overline{A}_t) &= \frac{1}{2}\xi^2 \int_0^t \int_0^\pi f^2(v(s,x))\phi^2(x)\,\mathrm{d}s\,\mathrm{d}x + \int_{\mathbb{R}} (\cos(\xi x) - 1)\,\overline{\nu}^k([0,t]\times \mathrm{d}x),\\ \operatorname{Im}(\overline{A}_t^k - \overline{A}_t) &= \xi\left(\int_0^t \langle v^k(s,\cdot),\phi''\rangle\,\mathrm{d}s - \int_0^t \langle v(s,\cdot),\phi''\rangle\,\mathrm{d}s - \int_{\mathbb{R}} (x-\vartheta(x))\,\overline{\nu}^k([0,t]\times\mathrm{d}x)\right)\\ &\quad + \int_{\mathbb{R}} (\sin(\xi x) - \xi\vartheta(x))\,\overline{\nu}^k([0,t]\times\mathrm{d}x).\end{aligned} \tag{1.3.31}$$

Consequently,

$$\begin{aligned}\operatorname{Var}(\operatorname{Re}(\overline{A}^k - \overline{A}))_t \le \frac{1}{2}\xi^2 \int_0^t \int_0^\pi &\left| f^2(v(s,x))\phi^2(x) - \int_{\mathbb{R}} \vartheta^2\left(\frac{f(v^k(s,x))}{\sigma(\varepsilon_k)}\phi(x) z\right) Q^{\varepsilon_k}(\mathrm{d}z)\right| \mathrm{d}x\,\mathrm{d}s\\ &+ \int_{\mathbb{R}} \left|\cos(\xi x) - 1 + \frac{1}{2}\xi^2\vartheta^2(x)\right| \overline{\nu}^k([0,t]\times\mathrm{d}x).\end{aligned} \tag{1.3.32}$$

We will show that the two integrals above converge to 0 as $k\to\infty$.

The function $|\cos(\xi x) - 1 + \frac{1}{2}\xi^2\vartheta^2(x)|$ is bounded, continuous (because ϑ is) and $o(x^2)$ as $x\to 0$. Hence, by condition $[\delta_{1,3}]$ of Theorem VII.2.9 in [65], we can infer

$$\int_{\mathbb{R}} \left|\cos(\xi x) - 1 + \frac{1}{2}\xi^2\vartheta^2(x)\right| \overline{\nu}^k([0,t]\times\mathrm{d}x) \longrightarrow 0 \quad \text{as} \quad k\to\infty$$

from Theorem 1.3.10.

Consider now the first integral in (1.3.32), and notice that

$$\vartheta^2(x) = x^2 \mathbb{1}_{\{|x|\le 1\}} + \mathbb{1}_{\{|x|>1\}} \quad \text{for all} \quad x\in\mathbb{R}.$$

Using the triangle inequality, we can therefore estimate this integral by

$$\begin{aligned}\int_0^t\int_0^\pi &\left| f^2(v(s,x))\phi^2(x) - \int_{\mathbb{R}} \Big(f(v^k(s,x))\phi(x)z/\sigma(\varepsilon_k)\Big)^2 \mathbb{1}_{\{|f(v^k(s,x))\phi(x)z|/\sigma(\varepsilon_k)\le 1\}}\, Q^{\varepsilon_k}(\mathrm{d}z)\right| \mathrm{d}x\,\mathrm{d}s\\ &+ \int_{\mathbb{R}} \mathbb{1}_{\{|x|>1\}}\,\overline{\nu}^k([0,t]\times\mathrm{d}x).\end{aligned}$$

The second integral above converges to 0 as shown in (1.3.29), while the same holds for the first integral by (1.3.24) (set $h=1$). Together with (1.3.32), we conclude that $\operatorname{Var}(\operatorname{Re}(\overline{A}^k - \overline{A}))_t \longrightarrow 0$ as $k\to\infty$.

It remains to show that $\operatorname{Var}(\operatorname{Im}(\overline{A}^k - \overline{A}))_t \longrightarrow 0$ as $k\to\infty$, which will be done in a similar manner as before. From (1.3.31), we have

$$\begin{aligned}\operatorname{Var}(\operatorname{Im}(\overline{A}^k - \overline{A}))_t \le |\xi| \int_0^t &\Big|\langle v^k(s,\cdot),\phi''\rangle - \langle v(s,\cdot),\phi''\rangle\Big|\,\mathrm{d}s + |\xi|\int_{\mathbb{R}} |x-\vartheta(x)|\,\overline{\nu}^k([0,t]\times\mathrm{d}x)\\ &+ \int_{\mathbb{R}} \Big|\sin(\xi x) - \xi\vartheta(x)\Big|\,\overline{\nu}^k([0,t]\times\mathrm{d}x).\end{aligned}$$

The first integral on the right-hand side above converges to 0 by (1.3.27). Furthermore, since

$$\int_{\mathbb{R}} |x-\vartheta(x)|\,\overline{\nu}^k([0,t]\times\mathrm{d}x) \le 2\int_{\mathbb{R}} |x|^2 \mathbb{1}_{\{x>1\}}\,\overline{\nu}^k([0,t]\times\mathrm{d}x),$$

also the second integral vanishes by (1.3.28). Finally, the function $\left|\sin(\xi x) - \xi\vartheta(x)\right|$ is bounded, continuous and $o(x^2)$ as $x \to 0$. Hence, we can again apply Theorem VII.2.9 in [65] in order to obtain

$$\int_{\mathbb{R}} \left|\sin(\xi x) - \xi\vartheta(x)\right| \overline{\nu}^k([0,t] \times \mathrm{d}x) \longrightarrow 0 \quad \text{as} \quad k \to \infty.$$

We conclude that $\mathrm{Var}(\mathrm{Im}(\overline{A}^k - \overline{A}))_t \longrightarrow 0$, and altogether $\mathrm{Var}(\overline{A}^k - \overline{A})_t \longrightarrow 0$ as $k \to \infty$. □

Proof of Theorem 1.3.9. According to Proposition VI.1.23 in [65], because the function $t \mapsto \int_0^t \exp(i\xi\langle\overline{v}_s, \phi\rangle)\, \overline{A}(\mathrm{d}s)$ is continuous, $\overline{M}^k$ converges to $\overline{M}$ in the Skorokhod topology for fixed $\overline{\omega} \in \overline{\Omega}$ if

$$e^{i\xi\langle\overline{v}^k,\phi\rangle} \longrightarrow e^{i\xi\langle\overline{v},\phi\rangle} \quad \text{and} \quad \int_0^{\cdot} e^{i\xi\langle\overline{v}^k_s,\phi\rangle}\, \overline{A}^k(\mathrm{d}s) \longrightarrow \int_0^{\cdot} e^{i\xi\langle\overline{v}_s,\phi\rangle}\, \overline{A}(\mathrm{d}s)$$

in $D([0,T],\mathbb{C})$ as $k \to \infty$.

Using the definition of the Skorokhod topology, we can easily infer from the convergence of $(\overline{v}^k)_{k\in\mathbb{N}}$ to $\overline{v}$ in $D([0,T], H_{-r}([0,\pi]))$ given in (1.2.15) that

$$\begin{aligned} \langle\overline{v}^k, \phi\rangle &\longrightarrow \langle\overline{v}, \phi\rangle \quad \text{in} \quad D([0,T],\mathbb{R}) \quad \text{and} \\ e^{i\xi\langle\overline{v}^k,\phi\rangle} &\longrightarrow e^{i\xi\langle\overline{v},\phi\rangle} \quad \text{in} \quad D([0,T],\mathbb{C}) \end{aligned}$$

as $k \to \infty$ for all $\phi \in H_r([0,\pi])$.

Next, we have

$$\begin{aligned} &\sup_{t\le T} \left| \int_0^t e^{i\xi\langle\overline{v}^k_s,\phi\rangle}\, \overline{A}^k(\mathrm{d}s) - \int_0^t e^{i\xi\langle\overline{v}_s,\phi\rangle}\, \overline{A}(\mathrm{d}s) \right| \\ &\quad \le \sup_{t\le T} \left| \int_0^t e^{i\xi\langle\overline{v}^k_s,\phi\rangle}\, (\overline{A}^k - \overline{A})(\mathrm{d}s) \right| + \sup_{t\le T} \left| \int_0^t \left(e^{i\xi\langle\overline{v}^k_s,\phi\rangle} - e^{i\xi\langle\overline{v}_s,\phi\rangle} \right) \overline{A}(\mathrm{d}s) \right| \\ &\quad \le \mathrm{Var}(\overline{A}^k - \overline{A})_T + \int_0^T |e^{i\xi\langle\overline{v}^k_s,\phi\rangle} - e^{i\xi\langle\overline{v}_s,\phi\rangle}|\, \mathrm{Var}(\overline{A})(\mathrm{d}s). \end{aligned}$$

Lemma 1.3.11 then immediately gives us $\mathrm{Var}(\overline{A}^k - \overline{A})_T \to 0$ as $k \to \infty$. In addition, the Skorokhod convergence of $e^{i\xi\langle\overline{v}^k,\phi\rangle}$ towards $e^{i\xi\langle\overline{v},\phi\rangle}$ implies $e^{i\xi\langle\overline{v}^k_t,\phi\rangle} \longrightarrow e^{i\xi\langle\overline{v}_t,\phi\rangle}$ for all continuity points of $e^{i\xi\langle\overline{v},\phi\rangle}$; see, for example, VI.2.3 of [65]. Since a càdlàg function has at most countably many discontinuities, we have $e^{i\xi\langle\overline{v}^k_t,\phi\rangle} \longrightarrow e^{i\xi\langle\overline{v}_t,\phi\rangle}$ for almost all $t \in [0,T]$. So dominated convergence implies that also the last term of the previous display converges to 0 as $k \to \infty$. □

We have now gathered all the intermediate results needed for the following theorem.

Theorem 1.3.12. *If* (1.1.4) *holds, then* $(v, \overline{v})$ *in* (1.2.15) *satisfies the following martingale problem. For all* $\xi \in \mathbb{R}$ *and* $\phi \in C_c^\infty((0,\pi))$, *the process* $(\overline{M}_t)_{t\le T}$ *defined in* (1.2.18) *is a martingale with respect to the filtration* $\overline{\boldsymbol{F}}$ *in* (1.2.16).

Furthermore, v *has an* $\overline{\boldsymbol{F}}$*-predictable modification and*

$$\operatorname*{ess\,sup}_{(t,x)\in[0,T]\times[0,\pi]} \mathbb{E}\left[|v(t,x)|^2\right] < \infty. \tag{1.3.33}$$

Finally, for almost all $t \in [0,T]$, $\overline{v}_t = \langle v(t,\cdot), \cdot\rangle$ *as well as* $\overline{v}_0 = 0$ *holds with probability one.*

Proof. By Theorem 1.3.8, for any $\xi \in \mathbb{R}$, $\phi \in C_c^\infty((0,\pi))$ and $k \in \mathbb{N}$, the process M^{ε_k} defined in (1.3.15) is a square-integrable $\boldsymbol{F}$-martingale.

Now define for each $k \in \mathbb{N}$, the filtration $\overline{\boldsymbol{F}}^k = (\overline{\mathcal{F}}_t^k)_{t \leq T}$ on $\overline{\Omega}$ by setting

$$\overline{\mathcal{F}}_t^k = \bigcap_{u>t} \sigma\left(v^k(s,x), \overline{v}_s^k \mid 0 \leq x \leq \pi,\, s \leq u\right), \quad t \leq T. \tag{1.3.34}$$

Since $\overline{v}^k$ and v^k in (1.2.15) are adapted to the filtration $\overline{\boldsymbol{F}}^k$, the same holds for $\overline{M}^k$ from (1.3.20). Similarly, $\overline{v}$, v and $\overline{M}$ are $\overline{\boldsymbol{F}}$-adapted. Since $\overline{M}^k$ has the same distribution as M^{ε_k} by (1.2.15), standard arguments now show that $\overline{M}^k$ is an $\overline{\boldsymbol{F}}^k$-martingale for all $\xi \in \mathbb{R}$, $\phi \in C_c^\infty((0,\pi))$ and $k \in \mathbb{N}$. This is the martingale problem satisfied by the pair $(v^k, \overline{v}^k)$.

Using Theorem 1.3.9, we have

$$\overline{M}^k(\overline{\omega}) \longrightarrow \overline{M}(\overline{\omega}) \quad \text{in} \quad D([0,T],\mathbb{C}) \tag{1.3.35}$$

as $k \to \infty$ for all $\overline{\omega} \in \overline{\Omega}$. This implies $\overline{M}_t^k(\overline{\omega}) \longrightarrow \overline{M}_t(\overline{\omega})$ almost everywhere on $[0,T]$ for all $\overline{\omega} \in \overline{\Omega}$. Furthermore,

$$\mathbb{E}\left[|\overline{M}_t^k|^2\right] = \mathbb{E}\left[|M_t^{\varepsilon_k}|^2\right] < \infty \tag{1.3.36}$$

uniformly in $k \in \mathbb{N}$ and $t \leq T$ by Theorem 1.3.8.

Moreover, the convergence in (1.2.15) implies convergence in measure (with respect to the Lebesgue measure on $[0,T] \times [0,\pi]$) of $v^k(\overline{\omega})$ towards $v(\overline{\omega})$ for all $\overline{\omega} \in \overline{\Omega}$. Hence, we have by dominated convergence,

$$\overline{\mathbb{P}} \otimes \mathrm{Leb}_{[0,T]\times[0,\pi]}\left(|v^k - v| \geq \varepsilon\right) = \mathbb{E}\left[\int_0^T \int_0^\pi \mathbb{1}_{\{|v^k(\overline{\omega},t,x) - v(\overline{\omega},t,x)| \geq \varepsilon\}}\, \mathrm{d}t\, \mathrm{d}x\right] \longrightarrow 0$$

as $k \to \infty$, and thus, v^k converges to v in $\overline{\mathbb{P}} \otimes \mathrm{Leb}_{[0,T]\times[0,\pi]}$-measure. Therefore, there exists a subsequence $(k_l)_{l \in \mathbb{N}}$ such that

$$v^{k_l} \longrightarrow v \quad \overline{\mathbb{P}} \otimes \mathrm{Leb}_{[0,T]\times[0,\pi]}\text{-almost everywhere} \quad \text{as} \quad l \to \infty, \tag{1.3.37}$$

and we will assume without loss of generality that (1.3.37) holds for the whole sequence. In turn, this implies $v^k \longrightarrow v$ $\overline{\mathbb{P}}$-almost surely as $k \to \infty$ for almost all $(t,x) \in [0,T] \times [0,\pi]$. Using Fatou's lemma, we obtain

$$\mathbb{E}\left[|v(t,x)|^2\right] \leq \liminf_{k\to\infty} \mathbb{E}\left[|v^k(t,x)|^2\right] \quad \mathrm{Leb}_{[0,T]\times[0,\pi]}\text{-almost everywhere.} \tag{1.3.38}$$

Furthermore,

$$v^k(t,x) \stackrel{d}{=} u^{\varepsilon_k}(t,x) \quad \mathrm{Leb}_{[0,T]\times[0,\pi]}\text{-almost everywhere,} \tag{1.3.39}$$

for all $k \in \mathbb{N}$, so (1.3.33) follows from Lemma 1.3.1. (In order to show (1.3.39), consider for $\alpha > 0$ the mollified random fields $J_\alpha v^k$ and $J_\alpha u^{\varepsilon_k}$ on $[0,T] \times [0,\pi]$, defined exactly as in (1.8) of Chapter 10 in [49]. Then (1.2.15) implies

$$(J_\alpha v^k)(t,x) \stackrel{d}{=} (J_\alpha u^{\varepsilon_k})(t,x) \tag{1.3.40}$$

for all $(t,x) \in [0,T] \times [0,\pi]$, $\alpha > 0$ and $k \in \mathbb{N}$. In addition, using Lemma 3 of Chapter 10 in [49], we have

$$J_\alpha v^k(\overline{\omega}) \longrightarrow v^k(\overline{\omega}) \quad \text{and} \quad J_\alpha u^{\varepsilon_k}(\omega) \longrightarrow u^{\varepsilon_k}(\omega) \quad \text{in} \quad L^2([0,T] \times [0,\pi]) \quad \text{as} \quad \alpha \to 0,$$

for all $k \in \mathbb{N}$, $\overline{\omega} \in \overline{\Omega}$ and $\omega \in \Omega$. As a consequence, we can find a sequence $(\alpha_l)_{l\in\mathbb{N}}$ converging to 0 such that

$$J_{\alpha_l} v^k(\overline{\omega}) \longrightarrow v^k(\overline{\omega}) \quad \text{and} \quad J_{\alpha_l} u^{\varepsilon_k}(\omega) \longrightarrow u^{\varepsilon_k}(\omega) \quad \mathrm{Leb}_{[0,T]\times[0,\pi]}\text{-almost everywhere} \tag{1.3.41}$$

as $l \to \infty$ for all $k \in \mathbb{N}$, $\overline{\omega} \in \overline{\Omega}$ and $\omega \in \Omega$. So (1.3.39) follows from (1.3.40) and (1.3.41).)

Finally, since we have $\overline{v}^k \longrightarrow \overline{v}$ in $D([0,T], H_{-r}([0,\pi]))$ for all $\overline{\omega} \in \overline{\Omega}$, we also infer that $\overline{v}^k_t \longrightarrow \overline{v}_t$ $\overline{\mathbb{P}}$-almost surely in $H_{-r}([0,\pi])$ for almost all $t \le T$.

Choose continuous bounded functions $\overline{h}\colon H_{-r}([0,\pi])^M \to \mathbb{R}$ and $h\colon \mathbb{R}^N \to \mathbb{R}$ with $M, N \in \mathbb{N}$. Using the $\overline{\boldsymbol{F}}^k$-martingale property of $\overline{M}^k$, (1.3.36) as well as Vitali convergence theorem, we now obtain the following: For almost all $0 \le s < t \le T$, $s_1, \ldots, s_M \le s$, $r_1, \ldots, r_N \le s$ and $x_1, \ldots, x_N \in [0,\pi]$, we have as $k \to \infty$,

$$\begin{aligned} 0 &= \mathbb{E}\left[\left(\overline{M}^k_t - \overline{M}^k_s\right) \overline{h}(\overline{v}^k_{s_1}, \ldots, \overline{v}^k_{s_M}) h(v^k(r_1,x_1), \ldots, v^k(r_N,x_N))\right] \\ &\longrightarrow \mathbb{E}\left[\left(\overline{M}_t - \overline{M}_s\right) \overline{h}(\overline{v}_{s_1}, \ldots, \overline{v}_{s_M}) h(v(r_1,x_1), \ldots, v(r_N,x_N))\right]. \end{aligned} \tag{1.3.42}$$

Next, u^ε is L^2-continuous by Theorem 4.7 in [28] and Lemma B.1 in [10]. This and (1.3.39) imply that v^k is also L^2-continuous. Since v^k satisfies the uniform bound (1.3.33) and with Vitali convergence theorem, we can infer that the random field v is L^1-continuous. Therefore, by a density argument, we can further deduce that the right-hand side of (1.3.42) is 0 for even all choices $0 \le s < t \le T$, $s_i, r_j \le s$ and $x_j \in [0,\pi]$ with $i = 1, \ldots, M$, $j = 1, \ldots, N$. Hence, again by standard arguments, we can deduce that $\overline{M}$ is an $\overline{\boldsymbol{F}}$-martingale for any $\xi \in \mathbb{R}$ and $\phi \in C_c^\infty((0,\pi))$.

Finally, by a straightforward extension of Proposition 3.21 in [83] to two-parameter processes and since v is stochastically continuous, v has a predictable modification.

The last statement of the theorem is easy and we leave the details to the reader. □

We can now finish the proof of the weak convergence (1.2.13). Indeed, the martingale problem stated in Theorem 1.3.12 and satisfied by $(v, \overline{v})$ in (1.2.15) will allow us to identify uniquely the distribution of $(v, \overline{v})$ (from now on we may and will assume that v is predictable).

Note that the next theorem holds independently of all our previous results.

Theorem 1.3.13. *On a filtered probability space $(\overline{\Omega}, \overline{\mathcal{F}}, \overline{\boldsymbol{F}}, \overline{\mathbb{P}})$, let $v = \{v(t,x) \mid (t,x) \in [0,T] \times [0,\pi]\}$ be an $\overline{\boldsymbol{F}}$-predictable random field and $\overline{v}$ an $\overline{\boldsymbol{F}}$-adapted càdlàg process in $H_{-r}([0,\pi])$, with $r > 1/2$. Assume that for almost all $t \in [0,T]$, $\overline{v}_t = \langle v(t,\cdot), \cdot\rangle$ as well as $\overline{v}_0 = 0$ holds $\overline{\mathbb{P}}$-almost surely and that*

$$\operatorname*{ess\,sup}_{(t,x)\in[0,T]\times[0,\pi]} \mathbb{E}\left[|v(t,x)|^2\right] < \infty. \tag{1.3.43}$$

In addition, assume that the pair $(v, \overline{v})$ satisfies the following martingale problem. For all $\xi \in \mathbb{R}$ and $\phi \in C_c^\infty((0,\pi))$, the process $(\overline{M}_t)_{t\le T}$ defined via (1.2.17) and (1.2.18) is a local $\overline{\boldsymbol{F}}$-martingale.

Then there exists a Gaussian space–time white noise $\widetilde{W}$ on $[0,T] \times [0,\pi]$, possibly defined on a filtered extension $(\widetilde{\Omega}, \widetilde{\mathcal{F}}, \widetilde{\boldsymbol{F}}, \widetilde{\mathbb{P}})$ of $(\overline{\Omega}, \overline{\mathcal{F}}, \overline{\boldsymbol{F}}, \overline{\mathbb{P}})$ such that, with probability one, v is equal in $L^2([0,T] \times [0,\pi])$ to the mild solution to the stochastic heat equation (1.2.8) with noise $\dot{\widetilde{W}}$. Furthermore, $\overline{v}$ is indistinguishable from the modification of the latter that is continuous in $H_{-r}([0,\pi])$.

Proof. The proof is inspired by Lemma 2.4 in [74]. First, Theorem II.2.42 in [65] shows that for any $\phi \in C_c^\infty((0,\pi))$, the stochastic process $\langle \overline{v}, \phi\rangle$ is an $\overline{\boldsymbol{F}}$-semimartingale with first and second characteristic given by

$$t \mapsto \int_0^t \langle v(s,\cdot), \phi''\rangle \, \mathrm{d}s \quad \text{and} \quad t \mapsto \int_0^t \int_0^\pi f^2(v(s,x))\phi^2(x) \, \mathrm{d}s \, \mathrm{d}x,$$

respectively. Furthermore, the third characteristic of $\langle \overline{v}, \phi\rangle$ equals 0, which implies that $\langle \overline{v}, \phi\rangle$ is continuous. As $\overline{v}_0 = 0$ $\overline{\mathbb{P}}$-almost surely, its canonical decomposition is

$$\langle \overline{v}, \phi\rangle = \int_0^{\cdot} \langle v(s,\cdot), \phi''\rangle \, \mathrm{d}s + \langle \overline{v}, \phi\rangle^c,$$

where $\langle \overline{v}, \phi\rangle^c$ denotes the continuous martingale part of $\langle \overline{v}, \phi\rangle$. Since

$$\mathbb{E}\left[\int_0^T \int_0^\pi f^2(v(s,x))\phi^2(x) \, \mathrm{d}s \, \mathrm{d}x\right] \le C\left(\operatorname*{ess\,sup}_{(s,x)\in[0,T]\times[0,\pi]} \mathbb{E}\left[|v(s,x)|^2\right] + 1\right),$$

which is finite by assumption, the quadratic variation of $\langle \overline{v}, \phi\rangle^c$ is integrable, so

$$M_t(\phi) = \langle \overline{v}_t, \phi\rangle - \int_0^t \langle v(s,\cdot), \phi''\rangle \, \mathrm{d}s, \quad t \le T, \tag{1.3.44}$$

is a continuous square-integrable $\overline{\boldsymbol{F}}$-martingale with quadratic variation process

$$t \mapsto \int_0^t \int_0^\pi f^2(v(s,x))\phi^2(x) \, \mathrm{d}s \, \mathrm{d}x, \tag{1.3.45}$$

for all $\phi \in C_c^\infty((0,\pi))$. The specifications (1.3.44) and (1.3.45) define an orthogonal martingale measure $\{M_t(A),\, t \in [0,T],\, A \in \mathcal{B}([0,\pi])\}$ relative to $(\overline{\Omega}, \overline{\mathcal{F}}, \overline{\boldsymbol{F}}, \overline{\mathbb{P}})$, in the sense of Chapter 2 in [99], with covariation measure

$$Q_M(A \times B \times [s,t]) = \int_s^t \int_{A\cap B} f^2(v(r,x)) \, \mathrm{d}r \, \mathrm{d}x \tag{1.3.46}$$

for all $A, B \in \mathcal{B}([0,\pi])$.

Now let $(\Omega', \mathcal{F}', \boldsymbol{F}', \mathbb{P}')$ be another filtered probability space on which a Gaussian space–time white noise W' on $[0,T]\times[0,\pi]$ is defined. Set

$$\widetilde{\Omega} = \overline{\Omega} \times \Omega', \quad \widetilde{\mathcal{F}} = \overline{\mathcal{F}} \otimes \mathcal{F}', \quad \widetilde{\mathcal{F}}_t = \bigcap_{s>t} \overline{\mathcal{F}}_s \otimes \mathcal{F}'_s, \quad \widetilde{\mathbb{P}} = \overline{\mathbb{P}} \otimes \mathbb{P}',$$

and extend the random measures M and W' as well as the random elements $\overline{v}$ and v to $\widetilde{\Omega}$ in the standard way so that on $(\widetilde{\Omega}, \widetilde{\mathcal{F}}, \widetilde{\boldsymbol{F}}, \widetilde{\mathbb{P}})$, W' is independent of $(\overline{v}, v)$ and thus of M. In addition, on this extension, M is still an orthogonal martingale measure satisfying (1.3.44) and (1.3.46) by Lemma II.7.3 in [65]. Define

$$\begin{aligned}\widetilde{W}_t(\phi) = &\int_0^t \int_0^\pi \frac{1}{f(v(s,x))} \mathbb{1}_{\{f^2(v(s,x))\neq 0\}} \phi(x) \, M(\mathrm{d}s, \mathrm{d}x) \\ &+ \int_0^t \int_0^\pi \mathbb{1}_{\{f^2(v(s,x)) = 0\}} \phi(x) \, W'(\mathrm{d}s, \mathrm{d}x)\end{aligned}$$

for all $t \le T$ and $\phi \in C_c^\infty((0,\pi))$. As before, this defines a martingale measure $\{\widetilde{W}_t(A),\, t \in [0,T],\, A \in \mathcal{B}([0,\pi])\}$ relative to $(\widetilde{\Omega}, \widetilde{\mathcal{F}}, \widetilde{\boldsymbol{F}}, \widetilde{\mathbb{P}})$.

Since M and W' are independent, we have from (1.3.46),

$$\begin{aligned}Q_{\widetilde{W}}(A\times B\times[s,t]) &= \int_s^t\int_{A\cap B}\frac{1}{f^2(v(r,x))}\mathbb{1}_{\{f^2(v(r,x))\neq 0\}}f^2(v(r,x))\,\mathrm{d}r\,\mathrm{d}x\\ &\quad+\int_s^t\int_{A\cap B}\mathbb{1}_{\{f^2(v(r,x))=0\}}\,\mathrm{d}r\,\mathrm{d}x\\ &=\int_s^t\int_{A\cap B}\mathrm{d}r\,\mathrm{d}x\end{aligned}$$

for all $A, B\in\mathcal{B}([0,\pi])$. Therefore, it follows from Proposition 2.1 in [99] that $\widetilde{W}$ is orthogonal and from Proposition 2.10 in [99] that the martingale measure $\widetilde{W}$ is a Gaussian space–time white noise on $[0,T]\times[0,\pi]$ with respect to $(\widetilde{\Omega},\widetilde{\mathcal{F}},\widetilde{\boldsymbol{F}},\widetilde{\mathbb{P}})$. Moreover, we have

$$\begin{aligned}\int_0^t\int_0^\pi f(v(s,x))\phi(x)\,\widetilde{W}(\mathrm{d}s,\mathrm{d}x) &= \int_0^t\int_0^\pi f(v(s,x))\frac{1}{f(v(s,x))}\mathbb{1}_{\{f^2(v(s,x))\neq 0\}}\phi(x)\,M(\mathrm{d}s,\mathrm{d}x)\\ &\quad+\int_0^t\int_0^\pi f(v(s,x))\mathbb{1}_{\{f^2(v(s,x))=0\}}\phi(x)\,W'(\mathrm{d}s,\mathrm{d}x)\\ &=\int_0^t\int_0^\pi \mathbb{1}_{\{f^2(v(s,x))\neq 0\}}\phi(x)\,M(\mathrm{d}s,\mathrm{d}x).\end{aligned}\tag{1.3.47}$$

Since, by (1.3.46),

$$\begin{aligned}&\mathbb{E}\left[\left(\int_0^T\int_0^\pi\mathbb{1}_{\{f^2(v(s,x))\neq 0\}}\phi(x)\,M(\mathrm{d}s,\mathrm{d}x)-\int_0^T\int_0^\pi\phi(x)\,M(\mathrm{d}s,\mathrm{d}x)\right)^2\right]\\ &\quad=\mathbb{E}\left[\left(\int_0^T\int_0^\pi\mathbb{1}_{\{f^2(v(s,x))=0\}}\phi(x)\,M(\mathrm{d}s,\mathrm{d}x)\right)^2\right]\\ &\quad=\mathbb{E}\left[\int_0^T\int_0^\pi\int_0^\pi\phi(x)\mathbb{1}_{\{f^2(v(s,x))=0\}}\phi(y)\mathbb{1}_{\{f^2(v(s,y))=0\}}\,Q_M(\mathrm{d}s,\mathrm{d}x,\mathrm{d}y)\right]\\ &\quad=\mathbb{E}\left[\int_0^T\int_0^\pi\phi^2(x)\mathbb{1}_{\{f^2(v(s,x))=0\}}f^2(v(s,x))\,\mathrm{d}x\,\mathrm{d}s\right]=0,\end{aligned}$$

the $\widetilde{\boldsymbol{F}}$-martingales $t\mapsto\int_0^t\int_0^\pi\mathbb{1}_{\{f^2(v(s,x))\neq 0\}}\phi(x)\,M(\mathrm{d}s,\mathrm{d}x)$ and $t\mapsto\int_0^t\int_0^\pi\phi(x)\,M(\mathrm{d}s,\mathrm{d}x)$ are indistinguishable. This implies, together with (1.3.44) and (1.3.47), that we have for any $\phi\in C_c^\infty((0,\pi))$,

$$\int_0^t\int_0^\pi f(v(s,x))\phi(x)\,\widetilde{W}(\mathrm{d}s,\mathrm{d}x)=M_t(\phi)=\langle\overline{v}_t,\phi\rangle-\int_0^t\langle v(s,\cdot),\phi''\rangle\,\mathrm{d}s,\quad t\le T,\tag{1.3.48}$$

$\widetilde{\mathbb{P}}$-almost surely. By assumption, the equality in (1.3.48) holds also $\widetilde{\mathbb{P}}$-almost surely for almost all $t\le T$ if we replace $\langle\overline{v}_t,\phi\rangle$ with $\langle v(t,\cdot),\phi\rangle$. This and the assumption (1.3.43) imply, by the proof of Theorem 3.2 in [99], that we have

$$v(t,x)=\int_0^t\int_0^\pi G_{t-s}(x,y)f(v(s,y))\,\widetilde{W}(\mathrm{d}s,\mathrm{d}y)\quad\widetilde{\mathbb{P}}\text{-almost surely}\tag{1.3.49}$$

for almost all $(t,x)\in[0,T]\times[0,\pi]$, i.e., v satisfies the mild formulation of (1.2.19) almost everywhere. Now let $\widetilde{v}$ be a mild solution to (1.2.19). Again by Theorem 3.2 in [99] and its proof, we can infer that $\widetilde{\mathbb{P}}$-almost surely, v and $\widetilde{v}$ are equal almost everywhere and hence, in $L^2([0,T]\times[0,\pi])$.

Finally, let $\widehat{v}$ be the continuous modification in $H_{-r}([0,\pi])$ of $\widetilde{v}$, which we obtain from Corollary 3.4 in [99]. By (1.3.49), $\widehat{v}_t = \langle v(t,\cdot),\cdot\rangle = \overline{v}_t$ $\widetilde{\mathbb{P}}$-almost surely for almost all $t \le T$, and therefore, because $\widehat{v}$ is continuous and $\overline{v}$ càdlàg, these two processes are indistinguishable. □

1.3.3 Necessity of the condition (1.1.4)

Remark 1.3.14. Suppose that the Lipschitz function f satisfies $f(0)\neq 0$. Then there must be $(t_1,x_1)\in[0,T]\times[0,\pi]$ such that $\mathbb{P}(f(u(t_1,x_1))\neq 0)>0$, where u is the mild solution to (1.2.8). Indeed, if we had $f(u(t,x))=0$ $\mathbb{P}$-almost surely for all (t,x), it would imply $u=0$ everywhere on $[0,T]\times[0,\pi]$ by equation (1.2.8). This in turn would imply $f(0)=0$, which contradicts the assumption.

Theorem 1.3.15. *Assume that $f(0)\neq 0$. In the setting of Theorem 1.2.1, if* (1.2.13) *holds, then we have* (1.1.4) *for all $\kappa>0$.*

Proof. If (1.2.13) holds, we can use Skorokhod's representation theorem as in the first part of the proof of Theorem 1.2.1 and obtain for any sequence $(\varepsilon_k)_{k\in\mathbb{N}}$ converging to 0, random elements

$$(v^k,\overline{v}^k),(v,\overline{v})\colon(\overline{\Omega},\overline{\mathcal{F}},\overline{\mathbb{P}})\longrightarrow(\Omega^*,\tau)$$

on a probability space $(\overline{\Omega},\overline{\mathcal{F}},\overline{\mathbb{P}})$ possibly different from $(\Omega,\mathcal{F},\mathbb{P})$ that satisfy (1.2.15). Of course, we now have

$$(v,\overline{v})\stackrel{d}{=}(u,\overline{u}). \tag{1.3.50}$$

Consider the same filtrations $\overline{\boldsymbol{F}}=(\overline{\mathcal{F}}_t)_{t\le T}$ and $\overline{\boldsymbol{F}}^k=(\overline{\mathcal{F}}^k_t)_{t\le T}$ on $\overline{\Omega}$ as defined in (1.2.16) and (1.3.34), respectively. For fixed $\phi\in C_c^\infty((0,\pi))$, define the processes

$$\begin{aligned}\overline{X}^k_t&=\langle\overline{v}^k_t,\phi\rangle-\int_0^t\int_0^\pi v^k(s,x)\phi''(x)\,\mathrm{d}s\,\mathrm{d}x,\\ \overline{X}_t&=\langle\overline{v}_t,\phi\rangle-\int_0^t\int_0^\pi v(s,x)\phi''(x)\,\mathrm{d}s\,\mathrm{d}x\end{aligned} \tag{1.3.51}$$

for all $k\in\mathbb{N}$ and $t\le T$. It is straightforward to infer from (1.2.15) that pointwise on $\overline{\Omega}$,

$$\overline{X}^k\longrightarrow\overline{X}\quad\text{in}\quad D([0,T],\mathbb{R})\quad\text{as}\quad k\to\infty. \tag{1.3.52}$$

Furthermore, by (1.2.15), (1.3.50) and (1.3.51), $\overline{X}^k$ and $\overline{X}$ have the same distribution as the square-integrable $\boldsymbol{F}$-martingales

$$t\mapsto\int_0^t\int_0^\pi\frac{f(u^{\varepsilon_k}(s,x))}{\sigma(\varepsilon_k)}\phi(x)\,L^{\varepsilon_k}(\mathrm{d}s,\mathrm{d}x)\quad\text{and}\quad t\mapsto\int_0^t\int_0^\pi f(u(s,x))\phi(x)\,W(\mathrm{d}s,\mathrm{d}x),$$

respectively, and therefore, by standard arguments, we can deduce that $\overline{X}^k$, resp. $\overline{X}$, is an $\overline{\boldsymbol{F}}^k$-martingale, resp. $\overline{\boldsymbol{F}}$-martingale, and that $\overline{X}$ is continuous.

Recall the truncation function ϱ_h introduced in (1.3.21). Using Theorem II.2.21 in [65], we can further infer that the semimartingale characteristics of $\overline{X}^k$ and $\overline{X}$ with respect to $\overline{\boldsymbol{F}}^k$ and $\overline{\boldsymbol{F}}$, respectively, and relative to ϱ_h for a fixed but arbitrary $h>0$, are given by $(\overline{B}^{k,h},0,\overline{\nu}^k)$ and $(0,\overline{C},0)$, respectively, where $\overline{\nu}^k$ is defined as in (1.3.19), $\overline{C}$ is defined as in (1.2.17) and

$$\overline{B}^{k,h}_t=-\int_0^t\int_{\mathbb{R}}x\mathbb{1}_{\{|x|>h\}}\,\overline{\nu}^k(\mathrm{d}s,\mathrm{d}x),\quad t\le T. \tag{1.3.53}$$

Define $\overline{X}^k(\varrho_h) = \overline{X}^k - \sum_{s\leq \cdot} \Delta\overline{X}^k_s \mathbb{1}_{\{|\Delta\overline{X}^k_s|>h\}}$ for all $k \in \mathbb{N}$. Then we have, by definition of the first characteristic,

$$\overline{X}^k(\varrho_h) = \overline{M}^{k,h} + \overline{B}^{k,h}, \tag{1.3.54}$$

where $\overline{M}^{k,h}$ is a local $\overline{\boldsymbol{F}}$-martingale.

Now since $\overline{X}$ is continuous, Proposition VI.2.7 in [65] and (1.3.52) imply that $\overline{\omega}$-wise,

$$\overline{X}^k(\varrho_h) \longrightarrow \overline{X} \quad \text{in} \quad D([0,T],\mathbb{R}) \quad \text{as} \quad k \to \infty. \tag{1.3.55}$$

We also have

$$\overline{\nu}^k([0,t] \times \{|x| > a\}) \xrightarrow{\overline{\mathbb{P}}} 0 \quad \text{as} \quad k \to \infty \tag{1.3.56}$$

for any $t \leq T$ and $a > 0$ by Proposition VI.3.26 and Lemma VI.4.22 in [65]. Therefore, there exists a subsequence of $(\overline{\nu}^k([0,T] \times \{|x| > h\}))_{k\in\mathbb{N}}$ converging $\overline{\mathbb{P}}$-almost surely to 0. For the sake of clarity, assume without loss of generality that this holds for the whole sequence. Applying the Cauchy–Schwarz inequality to $\overline{B}^{k,h}$ in (1.3.53), we further deduce that

$$\begin{aligned}\sup_{t\leq T} |\overline{B}^{k,h}_t|^2 &\leq \left(\int_0^T \int_0^\pi \int_{\mathbb{R}} \frac{f^2(v^k(t,x))}{\sigma^2(\varepsilon_k)} \phi^2(x) z^2 \,\mathrm{d}t\,\mathrm{d}x\, Q^{\varepsilon_k}(\mathrm{d}z) \right) \overline{\nu}^k([0,T] \times \{|x| > h\}) \\ &\leq C\overline{\nu}^k([0,T] \times \{|x| > h\}) \left(1 + \sup_{k\in\mathbb{N}} \int_0^T \int_0^\pi v^k(t,x)^2 \,\mathrm{d}t\,\mathrm{d}x \right)\end{aligned}$$

and the last term converges $\overline{\mathbb{P}}$-almost surely to 0 (note that the supremum is finite because $v^k \longrightarrow v$ in $L^2([0,T] \times [0,\pi])$). This implies

$$\overline{B}^{k,h} \longrightarrow 0 \quad \text{in} \quad D([0,T],\mathbb{R}) \quad \text{as} \quad k \to \infty \tag{1.3.57}$$

$\overline{\mathbb{P}}$-almost surely. Using Proposition VI.1.23 in [65], (1.3.54), (1.3.55) and (1.3.57), we obtain

$$\overline{M}^{k,h} \longrightarrow \overline{X} \quad \text{in} \quad D([0,T],\mathbb{R}) \quad \text{as} \quad k \to \infty$$

as well as

$$(\overline{M}^{k,h}, -2\overline{M}^{k,h}, (\overline{M}^{k,h})^2) \longrightarrow (\overline{X}, -2\overline{X}, \overline{X}^2) \quad \text{in} \quad D([0,T],\mathbb{R}^3) \quad \text{as} \quad k \to \infty \tag{1.3.58}$$

$\overline{\mathbb{P}}$-almost surely. Since the jumps of $\overline{M}^{k,h}$ are uniformly bounded by h, we can apply Proposition VI.6.13 in [65] on the sequence $(\overline{M}^{k,h})_{k\in\mathbb{N}}$ and then Theorem VI.6.22 (c) in [65] on the processes in (1.3.58) in order to obtain

$$\left(\overline{M}^{k,h}, -2\overline{M}^{k,h}, (\overline{M}^{k,h})^2, -2\int_0^\cdot \overline{M}^{k,h}_s\, \overline{M}^{k,h}(\mathrm{d}s) \right) \xrightarrow{\overline{\mathbb{P}}} \left(\overline{X}, -2\overline{X}, \overline{X}^2, -2\int_0^\cdot \overline{X}_s\, \overline{X}(\mathrm{d}s) \right)$$

in $D([0,T],\mathbb{R}^4)$ as $k \to \infty$. By definition of the quadratic variation, we can therefore deduce that

$$(\overline{M}^{k,h}, [\overline{M}^{k,h}, \overline{M}^{k,h}]) \xrightarrow{\overline{\mathbb{P}}} (\overline{X}, \overline{C}) \quad \text{in} \quad D([0,T],\mathbb{R}^2) \quad \text{as} \quad k \to \infty. \tag{1.3.59}$$

Denoting by $\overline{\mu}^k$ the jump measure of $\overline{X}^k$, we have, since $\overline{B}^{k,h}$ is continuous,

$$[\overline{M}^{k,h}, \overline{M}^{k,h}]_t = \int_0^t \int_{\mathbb{R}} x^2 \mathbb{1}_{\{|x|\leq h\}}\, \overline{\mu}^k(\mathrm{d}s,\mathrm{d}x), \quad t \leq T. \tag{1.3.60}$$

Now denote for any $k \in \mathbb{N}$,

$$\begin{aligned} \widetilde{C}_t^k &= \int_0^t \int_{\mathbb{R}} x^2 \mathbb{1}_{\{|x|\le h\}}\, \overline{\nu}^k(\mathrm{d}s,\mathrm{d}x) \quad \text{and} \\ \overline{Y}_t^k &= [\overline{M}^{k,h}, \overline{M}^{k,h}]_t - \widetilde{C}_t^k = \int_0^t \int_{\mathbb{R}} x^2 \mathbb{1}_{\{|x|\le h\}}\, (\overline{\mu}^k - \overline{\nu}^k)(\mathrm{d}s,\mathrm{d}x). \end{aligned} \tag{1.3.61}$$

Then $\overline{Y}^k$ is a square-integrable $\overline{\boldsymbol{F}}^k$-martingale with $|\Delta \overline{Y}^k| \le h$, and for any bounded stopping time T, we have, by the optional stopping theorem, $\mathbb{E}[(\overline{Y}_T^k)^2] \le \mathbb{E}[[\overline{Y}^k, \overline{Y}^k]_T]$. Therefore, by Lenglart's inequality (see Lemma I.3.30 in [65]), we obtain for all $\delta > 0$ and $\eta > 0$,

$$\begin{aligned} \mathbb{P}\left(\sup_{s\le t} |\overline{Y}_s^k|^2 \ge \delta\right) &\le \frac{1}{\delta}\left(\eta + \mathbb{E}\left[\sup_{s\le t} \Delta[\overline{Y}^k, \overline{Y}^k]_s\right]\right) + \mathbb{P}([\overline{Y}^k, \overline{Y}^k]_t \ge \eta) \\ &\le 2\frac{\eta}{\delta} + \left(\frac{h}{\delta} + 1\right) \mathbb{P}([\overline{Y}^k, \overline{Y}^k]_t \ge \eta). \end{aligned} \tag{1.3.62}$$

By (1.3.60), we have

$$[\overline{Y}^k, \overline{Y}^k]_t = \int_0^t \int_{\mathbb{R}} x^4 \mathbb{1}_{\{|x|\le h\}}\, \overline{\mu}^k(\mathrm{d}s,\mathrm{d}x) \le \left(\sup_{s\le t} |\Delta[\overline{M}^{k,h}, \overline{M}^{k,h}]_s|\right) [\overline{M}^{k,h}, \overline{M}^{k,h}]_t.$$

Moreover, because $[\overline{M}^{k,h}, \overline{M}^{k,h}]_t \xrightarrow{\overline{\mathbb{P}}} \overline{C}_t$ by (1.3.59) and $\sup_{s\le t} |\Delta[\overline{M}^{k,h}, \overline{M}^{k,h}]_s| \xrightarrow{\overline{\mathbb{P}}} 0$ by Proposition VI.3.26 (iii) in [65], we deduce from the inequality above that $[\overline{Y}^k, \overline{Y}^k]_t \xrightarrow{\overline{\mathbb{P}}} 0$ and, by (1.3.62), that

$$\sup_{s\le t} |\overline{Y}_s^k| \xrightarrow{\overline{\mathbb{P}}} 0 \quad \text{as} \quad k \to \infty \tag{1.3.63}$$

for all $t \le T$. Finally, combine (1.3.59), (1.3.61) and (1.3.63) to see that

$$\widetilde{C}_t^k = \int_0^t \int_{\mathbb{R}} x^2 \mathbb{1}_{\{|x|\le h\}}\, \overline{\nu}^k(\mathrm{d}s,\mathrm{d}x) \xrightarrow{\overline{\mathbb{P}}} \overline{C}_t \quad \text{as} \quad k \to \infty \tag{1.3.64}$$

for all $t \le T$ and $h > 0$. Taking a subsequence if necessary, we will from now on assume that the convergence in (1.3.64) holds even $\overline{\mathbb{P}}$-almost surely.

Recall now the definition of $\Sigma_h^k(s,x)$ in (1.3.23) and that, because $v^k \longrightarrow v$ in $L^2([0,T] \times [0,\pi])$, we have $\int_0^t \int_0^\pi |f^2(v^k(s,x)) - f^2(v(s,x))| \phi^2(x)\, \mathrm{d}s\, \mathrm{d}x \longrightarrow 0$ as $k \to \infty$; see the calculations in (1.3.25) and (1.3.26). Together with (1.3.64) this implies $\overline{\mathbb{P}}$-almost surely,

$$\int_0^T \int_0^\pi f^2(v(s,x)) \phi^2(x) \Sigma_h^k(s,x)\, \mathrm{d}s\, \mathrm{d}x \longrightarrow 0 \quad \text{as} \quad k \to \infty \tag{1.3.65}$$

for all $h > 0$ and $\phi \in C_c^\infty((0,\pi))$ by a similar calculation as in (1.3.24) (note that the first inequality there becomes an equality if $|\cdot|$ is replaced by $(\cdot)$ throughout).

Now on the set $\{|f(v^k(s,x))\phi(x)| \ge \delta\}$, where $\delta > 0$, we have

$$\mathbb{1}_{\left\{|z| \ge \left(h/|f(v^k(s,x))\phi(x)|\right)\sigma(\varepsilon_k)\right\}} \ge \mathbb{1}_{\{|z|\ge (h/\delta)\sigma(\varepsilon_k)\}},$$

and thus

$$\Sigma_h^k(s,x) \ge \frac{1}{\sigma^2(\varepsilon_k)} \int_{\mathbb{R}} z^2 \mathbb{1}_{\{|z|\ge(h/\delta)\sigma(\varepsilon_k)\}}\, Q^{\varepsilon_k}(\mathrm{d}z).$$

Therefore, as a consequence of (1.3.65), we obtain $\overline{\mathbb{P}}$-almost surely,

$$\begin{aligned}\frac{1}{\sigma^2(\varepsilon_k)}\int_{\mathbb{R}} z^2 \mathbb{1}_{\{|z|\geq (h/\delta)\sigma(\varepsilon_k)\}}\, Q^{\varepsilon_k}(\mathrm{d}z)\\ \times \int_0^T\int_0^\pi f^2(v(s,x))\phi^2(x)\mathbb{1}_{\{|f(v^k(s,x))\phi(x)|\geq\delta,\,|f(v(s,x))\phi(x)|>\delta\}}\,\mathrm{d}s\,\mathrm{d}x \longrightarrow 0\end{aligned}\tag{1.3.66}$$

as $k\to\infty$ for all $h>0$, $\delta>0$ and $\phi\in C_c^\infty((0,\pi))$.

We have seen in (1.3.37) that we can assume (perhaps for a subsequence) that

$$v^k \longrightarrow v \quad \text{as} \quad k\to\infty \quad \overline{\mathbb{P}}\otimes \mathrm{Leb}_{[0,T]\times[0,\pi]}\text{-almost everywhere,}$$

which implies, by dominated convergence and continuity of f,

$$\begin{aligned}\mathbb{E}\left[\int_0^T\int_0^\pi f^2(v(s,x))\phi^2(x)\mathbb{1}_{\{|f(v^k(s,x))\phi(x)|\geq\delta,\,|f(v(s,x))\phi(x)|>\delta\}}\,\mathrm{d}s\,\mathrm{d}x\right]\\ \longrightarrow \mathbb{E}\left[\int_0^T\int_0^\pi f^2(v(s,x))\phi^2(x)\mathbb{1}_{\{|f(v(s,x))\phi(x)|>\delta\}}\,\mathrm{d}s\,\mathrm{d}x\right] \quad \text{as} \quad k\to\infty.\end{aligned}\tag{1.3.67}$$

So from (1.2.15), (1.3.66) and (1.3.67), we deduce that

$$\begin{aligned}\frac{1}{\sigma^2(\varepsilon_k)}\int_{\mathbb{R}} z^2 \mathbb{1}_{\{|z|\geq (h/\delta)\sigma(\varepsilon_k)\}}\, Q^{\varepsilon_k}(\mathrm{d}z)\\ \times \mathbb{E}\left[\int_0^T\int_0^\pi f^2(u(s,x))\phi^2(x)\mathbb{1}_{\{|f(u(s,x))\phi(x)|>\delta\}}\,\mathrm{d}s\,\mathrm{d}x\right] \longrightarrow 0 \quad \text{as} \quad k\to\infty\end{aligned}\tag{1.3.68}$$

for all $h>0$, $\delta>0$ and $\phi\in C_c^\infty((0,\pi))$. Moreover,

$$\begin{aligned}&\left(\frac{1}{\sigma^2(\varepsilon_k)}\int_{\mathbb{R}} z^2 \mathbb{1}_{\{|z|\geq (h/\delta)\sigma(\varepsilon_k)\}}\, Q^{\varepsilon_k}(\mathrm{d}z)\right)\mathbb{E}\left[\int_0^T\int_0^\pi f^2(u(s,x))\phi^2(x)\,\mathrm{d}s\,\mathrm{d}x\right]\\ &\quad\leq \left(\frac{1}{\sigma^2(\varepsilon_k)}\int_{\mathbb{R}} z^2 \mathbb{1}_{\{|z|\geq (h/\delta)\sigma(\varepsilon_k)\}}\, Q^{\varepsilon_k}(\mathrm{d}z)\right)\\ &\qquad\times \mathbb{E}\left[\int_0^T\int_0^\pi f^2(u(s,x))\phi^2(x)\mathbb{1}_{\{|f(u(s,x))\phi(x)|>\delta\}}\,\mathrm{d}s\,\mathrm{d}x\right] + T\pi\delta^2.\end{aligned}\tag{1.3.69}$$

So if we choose $h=\kappa\delta$ with $\kappa>0$ arbitrary, then by (1.3.68), the first term on the right-hand side of (1.3.69) converges to 0 as $k\to\infty$ for all $\kappa>0$ and $\delta>0$. The second term does not depend on k nor h and converges to 0 as $\delta\to 0$. This implies

$$\left(\frac{1}{\sigma^2(\varepsilon_k)}\int_{\mathbb{R}} z^2 \mathbb{1}_{\{|z|\geq \kappa\sigma(\varepsilon_k)\}}\, Q^{\varepsilon_k}(\mathrm{d}z)\right)\mathbb{E}\left[\int_0^T\int_0^\pi f^2(u(s,x))\phi^2(x)\,\mathrm{d}s\,\mathrm{d}x\right] \longrightarrow 0 \tag{1.3.70}$$

as $k\to\infty$ for all $\kappa>0$. Since $f(0)\neq 0$, there exists $(t_1,x_1)\in[0,T]\times[0,\pi]$ such that $\mathbb{E}[f^2(u(t_1,x_1))]>0$ by Remark 1.3.14. Moreover, the mild solution u is continuous in $L^2(\Omega,\mathcal{F},\mathbb{P})$, which follows from the proof of Corollary 3.4 in [99]. We can thus infer that the expectation in (1.3.70) is not 0 and we obtain

$$\frac{1}{\sigma^2(\varepsilon_k)}\int_{\mathbb{R}} z^2 \mathbb{1}_{\{|z|\geq \kappa\sigma(\varepsilon_k)\}}\, Q^{\varepsilon_k}(\mathrm{d}z) \longrightarrow 0 \quad \text{as} \quad k\to\infty$$

for all $\kappa>0$, which is exactly (1.1.4). □

Chapter 2

Normal approximation of the solution to the stochastic wave equation with Lévy noise

2.1 Introduction

The wave equation is the prototype of an hyperbolic PDE, widely used e.g. in acoustics and signal processing ([46]). In the literature of stochastic PDEs, the corresponding equation with random perturbation has been extensively studied, especially when the driving noise is Gaussian: See e.g. [99] for the case of a space–time white noise and Chapter 2 of [40] for a noise that is white in time but spatially correlated. In this paper, we consider the stochastic wave equation

$$\partial_{tt}u(t,x) = \partial_{xx}u(t,x) + f(u(t,x))\dot{L}(t,x), \quad (t,x) \in \mathbb{R}^+ \times \mathbb{R}, \tag{2.1.1}$$

where $\dot{L}$ is a Lévy space–time white noise. We investigate the *normal approximation* on $[0,T] \times [0,L]$ of solutions to (2.1.1) when $\dot{L}$ has a finite second moment and no Gaussian component, that is, when can they be approximated in law on compact domains by the solution to

$$\partial_{tt}u(t,x) = \partial_{xx}u(t,x) + f(u(t,x))\dot{W}(t,x), \quad (t,x) \in \mathbb{R}^+ \times \mathbb{R}, \tag{2.1.2}$$

with a Gaussian space–time white noise $\dot{W}$. The purpose of this work is to show that the *necessary and sufficient* condition for this functional convergence is

$$\lim_{\varepsilon \to 0} \frac{1}{\sigma^2(\varepsilon)} \int_{|z|>\kappa\sigma(\varepsilon)} z^2 \, Q^\varepsilon(\mathrm{d}z) = 0 \tag{2.1.3}$$

where $\sigma^2(\varepsilon)$ is the variance of a homogeneous Lévy noise $\dot{L}^\varepsilon$ with Lévy measure Q^ε, for all $\kappa > 0$.

Intuitively, to make such an approximation plausible, $\dot{L}^\varepsilon$ should be close to $\dot{W}$ in distribution. For Lévy processes having jumps decreasing in size to 0, this was made rigorous in [9] and more generally in [36], where condition (2.1.3) was first introduced. The passage to an infinite-dimensional setting has been addressed in such generality, to our best knowledge, only in the case of *parabolic* stochastic PDEs like the stochastic heat equation, see [32] and references therein. We substantially generalize these results to the category of hyperbolic stochastic PDEs. Our second contribution is that we consider throughout equations with *multiplicative noise*. As an application, in the situation of *small jumps approximation* of [9], if the impulses of the noise $\dot{L}^\varepsilon$ decrease too fast to 0 (such as for a gamma noise), then the corresponding stochastic PDE will not admit a normal approximation, but it will, for instance, if $\dot{L}^\varepsilon$ is α-stable for any $\alpha \in (0,2)$.

The strategy of proof of our main result, Theorem 2.4.1, is identical to [32]: We show tightness of the solutions to (2.1.1) via a generalization of the Aldous criterion [59] and then identify uniquely the limit. Here classical methods relying on the Lévy–Khintchine formula do not apply due to the multiplicative noise, so we resort to *martingale problems* that correspond to the solutions and whose associated martingales *converge* under condition (2.1.3) to a limit that we link to the solution to (2.1.2). However, their predictable characteristics do not depend on the martingale process itself, which makes other well-established techniques, see e.g. Chapter IX of [65], inapplicable as well. Instead, we prove convergence by hand, so to say, in Skorokhod's representation. To this end, we show that all solution processes considered belong to a suitable Skorokhod space (and to an L^2-space as well), a fact we will extensively utilize because, in our setting, convergence in Skorokhod topology preserves the martingale property.

The *random field solution* u to (2.1.1) will exhibit a càdlàg property *jointly* in space and time, as we show in Theorem 2.3.2, that is directly linked to the shape of the wave kernel. Now it turns out that to show our normal approximation result, we will need to investigate two different processes simultaneously: u and its time derivative $\partial_t u$, because both appear in the weak formulation of the stochastic wave equation that we consider in this work and, hence, in the aforementioned martingale problems, see Section 2.3.2 for details. This is a substantial difference with [32] where two different representations of the *same* process needed to be adopted due to the singularities of the heat kernel. We show in Theorem 2.3.5 that we can view $\partial_t u$ as a càdlàg process taking values on a space of distributions constructed via Hermite expansions. We also mention that u becomes a *strong martingale* after an appropriate change of coordinate system, a crucial property that we will use to show tightness and the path properties above in place of the factorization method from [39, 90] (applied in [32]).

As in [9, 32, 36], our motivation comes from numerical simulation: An additional normal approximation of the small jumps of the noise in (2.1.1) might improve the rate of convergence of numerical schemes, as suggested by the results in [73] for SDEs and in [24] for SPDEs.

This paper is organized as follows. In Section 2.2, we describe in detail equations (2.1.1) and (2.1.2). In Section 2.3, we introduce all function spaces needed and show existence of the random elements that will be studied in Section 2.4, which contains our main result as well as the main ideas of its proof. The details as well as the proofs for Section 2.3 are postponed to Section 2.5.

2.2 Preliminaries

Consider on a filtered probability space $(\Omega, \mathcal{F}, \boldsymbol{F} = (\mathcal{F}_t)_{t\leq T}, \mathbb{P})$ that satisfies the usual conditions, for any $\varepsilon > 0$, the *stochastic wave equation* on $\mathbb{R}^+ \times \mathbb{R}$ with vanishing initial conditions:

$$\left[\begin{array}{ll} \partial_{tt} u^\varepsilon(t,x) = \partial_{xx} u^\varepsilon(t,x) + f(u^\varepsilon(t,x)) \dfrac{\dot{L}^\varepsilon(t,x)}{\sigma(\varepsilon)}, & (t,x) \in \mathbb{R}^+ \times \mathbb{R}, \\ u^\varepsilon(0,x) = \partial_t u^\varepsilon(0,x) = 0, & \text{for all} \quad x \in \mathbb{R}, \end{array}\right. \tag{2.2.1}$$

where $\dot{L}^\varepsilon(t,x)$ is a pure-jump Lévy space–time white noise on $\mathbb{R}^+ \times \mathbb{R}$ given by

$$L^\varepsilon(A) = \int_{\mathbb{R}^+ \times \mathbb{R}} \int_{\mathbb{R}} \mathbb{1}_A(t,x) z \, (\mu^\varepsilon - \nu^\varepsilon)(\mathrm{d}t, \mathrm{d}x, \mathrm{d}z) \tag{2.2.2}$$

for all bounded Borel sets $A \in \mathcal{B}_b(\mathbb{R}^+ \times \mathbb{R})$. In this representation, μ^ε is a homogeneous Poisson random measure on $(\mathbb{R}^+ \times \mathbb{R}) \times \mathbb{R}$ relative to the filtration $\boldsymbol{F}$, with intensity measure $\nu^\varepsilon = \mathrm{Leb}_{\mathbb{R}^+ \times \mathbb{R}} \otimes Q^\varepsilon$. Here Q^ε is a Lévy measure on $\mathbb{R}$, that is, $Q^\varepsilon(\{0\}) = 0$ and $\int_{\mathbb{R}} (1 \wedge z^2) \, Q^\varepsilon(\mathrm{d}z) < \infty$, see e.g. Chapter II in [65] for the definition of stochastic integrals with respect to Poisson random measures. Furthermore, we assume that for all $\varepsilon > 0$,

$$0 < \sigma^2(\varepsilon) = \int_{\mathbb{R}} z^2 \, Q^\varepsilon(\mathrm{d}z) < \infty, \tag{2.2.3}$$

which is the variance of $L^\varepsilon([0,1] \times [0,1])$. The special case

$$Q^\varepsilon(A) = \int_{|z|\le\varepsilon} \mathbb{1}_A(z)\, Q(\mathrm{d}z), \quad A \in \mathcal{B}(\mathbb{R}), \quad \varepsilon > 0,$$

for a single Poisson random measure μ having intensity measure $\nu = \mathrm{Leb}_{\mathbb{R}^+\times\mathbb{R}} \otimes Q$, corresponds to the small jump approximation in [9].

The function $f\colon \mathbb{R} \longrightarrow \mathbb{R}$ in equation (2.2.1) will be assumed to be Lipschitz continuous throughout this work.

We are interested in the notion of *mild solution* to (2.2.1). It is defined as an $\boldsymbol{F}$-predictable random field $u^\varepsilon = \{u^\varepsilon(t,x) \mid (t,x) \in \mathbb{R}^+ \times \mathbb{R}\}$ satisfying for all $(t,x) \in \mathbb{R}^+ \times \mathbb{R}$,

$$\begin{aligned} u^\varepsilon(t,x) &= \int_0^t \int_{\mathbb{R}} G_{t-s}(x,y) \frac{f(u^\varepsilon(s,y))}{\sigma(\varepsilon)} \, L^\varepsilon(\mathrm{d}s,\mathrm{d}y) \\ &= \int_{[0,t]\times\mathbb{R}} \int_{\mathbb{R}} G_{t-s}(x,y) f(u^\varepsilon(s,y)) \frac{z}{\sigma(\varepsilon)} \, (\mu^\varepsilon - \nu^\varepsilon)(\mathrm{d}s,\mathrm{d}y,\mathrm{d}z) \quad \mathbb{P}\text{-almost surely.} \end{aligned} \tag{2.2.4}$$

In this equation, G denotes the Green's function of the wave operator $\partial_{tt} - \partial_{xx}$ and has the following expression:

$$G_{t-s}(x,y) = G(t,x;s,y) = \frac{1}{2} \mathbb{1}_{A^+(t,x)}(s,y) \tag{2.2.5}$$

for any $(t,x,s,y) \in (\mathbb{R}^+ \times \mathbb{R})^2$, where

$$A^+(t,x) = \Big\{(s,y) \in \mathbb{R}^+ \times \mathbb{R} \mid |y-x| \le t-s\Big\} \tag{2.2.6}$$

denotes the backward light cone with apex (t,x) restricted to $\mathbb{R}^+ \times \mathbb{R}$. In particular, G is bounded and not differentiable. By Theorem 3.1 in [28], there exists a unique mild solution u^ε to (2.2.1) satisfying

$$\sup_{(t,x)\in[0,T]\times\mathbb{R}} \mathbb{E}\left[|u^\varepsilon(t,x)|^p\right] < \infty \tag{2.2.7}$$

for all $T > 0$, $0 < p \le 2$ and $\varepsilon > 0$. Indeed, from (2.2.3) and

$$\sup_{(t,x)\in[0,T]\times\mathbb{R}} \int_0^t \int_{\mathbb{R}} G(t,x;s,y)^p \, \mathrm{d}s \, \mathrm{d}y = (1/2)^p \, T^2 < \infty \tag{2.2.8}$$

for all $T > 0$ and $p > 0$, Assumption A in [28] is easily seen to be satisfied for $p = 2$.

We will investigate the *normal approximation* of u^ε and for this, we also consider the solution to the same stochastic PDE as above, but now driven by a Gaussian space–time white noise:

$$\left[\begin{array}{ll} \partial_{tt} u(t,x) = \partial_{xx} u(t,x) + f(u(t,x))\dot{W}(t,x), & (t,x) \in \mathbb{R}^+ \times \mathbb{R}, \\ u(0,x) = \partial_t u(0,x) = 0, & \text{for all} \quad x \in \mathbb{R}. \end{array}\right. \tag{2.2.9}$$

The driving noise $\dot{W}$ in (2.2.9) is a centered Gaussian random field $\{W(A) \mid A \in \mathcal{B}_b(\mathbb{R}^+ \times \mathbb{R})\}$ with covariance structure $\mathbb{E}[W(A)W(B)] = \mathrm{Leb}_{\mathbb{R}^+\times\mathbb{R}}(A \cap B)$ for bounded Borel sets $A, B \subseteq \mathbb{R}^+ \times \mathbb{R}$. As is well-known (see e.g. Exercise 3.7 of Chapter 3 in [99]), equation (2.2.9) has a (unique) continuous mild solution u that satisfies the corresponding bound in (2.2.7) for all $p > 0$.

2.3 Functional setting

In this work, the letter C will always denote a strictly positive constant whose value may change from line to line. Note that $|f(x)| \leq C|x| + |f(0)|$ for all $x \in \mathbb{R}$ by the Lipschitz continuity of f.

If φ_1, φ_2 are elements of the same L^2-space, we will always use the notation $\langle \varphi_1, \varphi_2 \rangle$ for the standard scalar product of that space and $\|\cdot\|$ for the induced norm. If ϕ is an element of a topological vector space and ϕ' an element of its topological dual, then $\langle \phi', \phi \rangle$ will always denote the dual pairing of ϕ' with ϕ.

2.3.1 Path property of mild solutions

Consider the partial order $\preceq$ on $\mathbb{R}^2$:

$$(\tilde{t}, \tilde{x}) \preceq (t, x) :\Leftrightarrow \tilde{t} \leq t \quad \text{and} \quad |\tilde{x} - x| \leq t - \tilde{t} \tag{2.3.1}$$

introduced in Section 5 of [84]. We define a space–time càdlàg property corresponding to $\preceq$.

Definition 2.3.1. A function $\phi\colon M \to \mathbb{R}$ with $M \subseteq \mathbb{R}^2$ is called $\preceq$-*càdlàg* if for every $(t, x) \in M$,

(1) $\lim_{\substack{(\tilde{t},\tilde{x}) \to (t,x) \\ (\tilde{t},\tilde{x}) \succeq (t,x)}} \phi(\tilde{t}, \tilde{x}) = \phi(t, x)$,

(2) The limits from the *flanks*, that is,

$$\lim_{\substack{(\tilde{t},\tilde{x}) \to (t,x) \\ \tilde{t}<t,\, |\tilde{x}-x|<t-\tilde{t}}} \phi(\tilde{t}, \tilde{x}), \qquad \lim_{\substack{(\tilde{t},\tilde{x}) \to (t,x) \\ \tilde{x}>x,\, x-\tilde{x}\leq \tilde{t}-t<\tilde{x}-x}} \phi(\tilde{t}, \tilde{x}) \quad \text{and} \quad \lim_{\substack{(\tilde{t},\tilde{x}) \to (t,x) \\ \tilde{x}<x,\, \tilde{x}-x\leq \tilde{t}-t<x-\tilde{x}}} \phi(\tilde{t}, \tilde{x}) \quad \text{all exist.}$$

We further denote the space of all $\preceq$-càdlàg functions on M by $D_{\preceq}(M)$.

We have the following result.

Theorem 2.3.2. *For any $\varepsilon > 0$, let u^ε be a mild solution to the stochastic wave equation* (2.2.1) *with noise $\sigma^{-1}(\varepsilon)\dot{L}^\varepsilon$. Then u^ε has a modification $\overline{u}^\varepsilon$ in $D_{\preceq}(\mathbb{R}^+ \times \mathbb{R})$.*

We will investigate the functional convergence of the $\preceq$-càdlàg version $\overline{u}^\varepsilon$ of Theorem 2.3.2 towards u and to this end, we need a suitable Skorokhod topology for $\preceq$-càdlàg functions.

Consider the order-preserving change of basis in $\mathbb{R}^2$ obtained by rotating the standard basis vectors clockwise by 45 degrees

$$H\colon (\mathbb{R}^2, \preceq) \longrightarrow (\mathbb{R}^2, \leq), \qquad \begin{pmatrix} t \\ x \end{pmatrix} \mapsto \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} t \\ x \end{pmatrix} \tag{2.3.2}$$

as well as a shifting in $\mathbb{R}^2$ by $u_0 = (-3/2, 1/2)$ composed with H and then rescaled

$$J\colon (\mathbb{R}^2, \preceq) \longrightarrow (\mathbb{R}^2, \leq), \quad u \mapsto \frac{\sqrt{2}}{3} H(u - u_0). \tag{2.3.3}$$

We set $u^* = (3/2, 1/2)$ and use the following notation for closed rectangles with respect to $\preceq$:

$$[(\tilde{t}, \tilde{x}), (t, x)]_{\preceq} = \Big\{ (s, y) \in \mathbb{R}^2 \mid (\tilde{t}, \tilde{x}) \preceq (s, y) \preceq (t, x) \Big\} \quad \text{for} \quad (\tilde{t}, \tilde{x}) \preceq (t, x). \tag{2.3.4}$$

We then have $[0,1]^2 \subsetneq [u_0, u^*]_{\preceq}$ and J builds a bijection between $[u_0, u^*]_{\preceq}$ and $[0,1]^2$. This particular choice of the vectors u_0 and u^* is for simplicity only, it guarantees that the processes we consider in the proofs of Theorem 2.3.2 (and Theorem 2.5.1) vanish on the axes, a technical requirement of strong martingales often seen in the literature.

Now let $D([0,1]^2)$ be the usual Skorokhod space of càdlàg functions on $[0,1]^2$ with respect to the partial order $\leq$ where $(\tilde{t},\tilde{x}) \leq (t,x)$ if and only if $\tilde{t} \leq t$ and $\tilde{x} \leq x$, see e.g. Section 2 in [58] for a definition. Consider the well-defined bijective transformation

$$\Phi : D([0,1]^2) \longrightarrow D_{\preceq}([u_0,u^*]_{\preceq}), \quad x \mapsto x \circ J. \tag{2.3.5}$$

We now draw upon the results of [92] on general Skorokhod spaces to obtain the following.

Lemma 2.3.3. *There exists a Skorokhod metric, that will be denoted by τ throughout this work, that makes $D_{\preceq}([0,1]^2)$ and $D_{\preceq}([u_0,u^*]_{\preceq})$ complete and separable metric spaces and with respect to which the composition*

$$D([0,1]^2) \xrightarrow{\Phi} D_{\preceq}([u_0,u^*]_{\preceq}) \overset{\iota}{\hookrightarrow} D_{\preceq}([0,1]^2), \tag{2.3.6}$$

with Φ as in (2.3.5) *and ι the restriction map, is* continuous. *Furthermore, $\preceq$-càdlàg functions are continuous except on at most countably many lines and bounded. If $x_n \xrightarrow{\tau} x$ with $x_n, x \in D_{\preceq}([0,1]^2)$, then $x_n(u) \longrightarrow x(u)$ at all continuity points u of x.*

An immediate consequence of Lemma 2.3.3 is that *tightness* of probability measures in $D([0,1]^2)$, for which there exist criteria in the literature, implies tightness of the transformed measures (according to (2.3.6)) in $D_{\preceq}([0,1]^2)$. This will be of crucial importance for the proof of Theorem 2.5.1.

A straightforward extension of Lemma 2.3.3, in the proof of which a definition of τ is given, yields a Skorokhod topology on $D_{\preceq}([0,T] \times I)$ with $T > 0$ and $I \subseteq \mathbb{R}$ a finite closed interval.

2.3.2 Weak formulation

The martingale problem approach mentioned in the introduction relies on a suitable weak formulation of the stochastic wave equation on $\mathbb{R}^+ \times \mathbb{R}$ that we formally compute from (2.2.1) in this section. It is inspired by Section 13.1 (together with Definition 9.11) of [83].

Let $\phi_1, \phi_2 \in C_c^\infty(\mathbb{R})$ and $\varepsilon > 0$. Take the scalar multiplication of both sides of (2.2.1) with ϕ_2 and integrate over $[0,t] \times \mathbb{R}$. Use the initial condition of $\partial_t u^\varepsilon$ as well as partial integration twice to obtain

$$\int_{\mathbb{R}} \partial_t u^\varepsilon(t,x)\phi_2(x)\,\mathrm{d}x = \int_{[0,t]\times\mathbb{R}} u^\varepsilon(s,x)\phi_2''(x)\,\mathrm{d}s\,\mathrm{d}x + \int_0^t \int_{\mathbb{R}} \phi_2(x) f(u^\varepsilon(s,x)) \frac{\dot{L}^\varepsilon(s,x)}{\sigma(\varepsilon)}\,\mathrm{d}s\,\mathrm{d}x. \tag{2.3.7}$$

Now add the equation $\int_{\mathbb{R}} u^\varepsilon(t,x)\phi_1(x)\,\mathrm{d}x = \int_{[0,t]\times\mathbb{R}} \partial_t u^\varepsilon(s,x)\phi_1(x)\,\mathrm{d}s\,\mathrm{d}x$, which readily follows from the initial condition of u^ε, to (2.3.7) in order to obtain for all $t \geq 0$,

$$\begin{aligned}&\int_{\mathbb{R}} u^\varepsilon(t,x)\phi_1(x)\,\mathrm{d}x + \int_{\mathbb{R}} \partial_t u^\varepsilon(t,x)\phi_2(x)\,\mathrm{d}x \\ &\quad = \int_0^t \left(\int_{\mathbb{R}} u^\varepsilon(s,x)\phi_2''(x)\,\mathrm{d}x + \int_{\mathbb{R}} \partial_t u^\varepsilon(s,x)\phi_1(x)\,\mathrm{d}x \right) \mathrm{d}s + \int_0^t \int_{\mathbb{R}} \phi_2(x) f(u^\varepsilon(s,x)) \frac{\dot{L}^\varepsilon(s,x)}{\sigma(\varepsilon)}\,\mathrm{d}s\,\mathrm{d}x.\end{aligned} \tag{2.3.8}$$

It turns out that using equation (2.3.7) alone is enough to prove the necessity of (2.1.3) for $\overline{u}^\varepsilon \xrightarrow{d} u$, but not its sufficiency. For the latter, it is really equation (2.3.8) that will be needed in because it yields an equivalence of weak and mild solution to (2.2.1) (and analogously for (2.2.9)).

Note that in [99], page 309, the author develops a different weak formulation for (2.2.9). However, it does not yield an equality of stochastic processes by fixed test function (because the latter must satisfy a condition that depends on the current time point), which is required for martingale problems.

Because u^ε is locally integrable on $\mathbb{R}^+ \times \mathbb{R}$ by (2.2.7), it is a distribution on $\mathbb{R}^+ \times \mathbb{R}$ and so, $\partial_t u^\varepsilon = \partial u^\varepsilon / \partial t$ in (2.3.8) will be the time derivative of u^ε in the sense of distributions. The aim of the next section is to find a convenient representation of $\partial u^\varepsilon/\partial t$ by means of a distribution-valued càdlàg process, that we can insert into equation (2.3.8) and thereby use for showing our normal approximation result.

2.3.3 Distributional time derivative and path property

For simplicity, we write in this paper $\delta_{y\pm(t-s)}(\mathrm{d}x) = \delta_{y+(t-s)}(\mathrm{d}x) + \delta_{y-(t-s)}(\mathrm{d}x)$ and we use this notation for functions as well. Let $\Psi \in C_c^\infty(\mathbb{R}^+ \times \mathbb{R})$ and fix $(s,y) \in \mathbb{R}^+ \times \mathbb{R}$. Straightforward computations yield for the Green's function G,

$$\begin{aligned}\int_{\mathbb{R}^+\times\mathbb{R}} & G(t,x;s,y)\partial_t\Psi(t,x)\,\mathrm{d}t\,\mathrm{d}x = \frac{1}{2}\int_{\mathbb{R}}\int_{\mathbb{R}^+} \mathbb{1}_{\{|y-x|\leq t-s\}}\partial_t\Psi(t,x)\,\mathrm{d}t\,\mathrm{d}x \\ &= -\frac{1}{2}\int_{\mathbb{R}} \Psi(s+|y-x|,x)\,\mathrm{d}x = -\frac{1}{2}\left(\int_s^\infty \Psi(t,y+(t-s))\,\mathrm{d}t + \int_s^\infty \Psi(t,y-(t-s))\,\mathrm{d}t\right) \\ &= -\frac{1}{2}\int_0^\infty\int_{\mathbb{R}} \Psi(t,x)\,\delta_{y\pm(t-s)}(\mathrm{d}x)\mathbb{1}_{\{s\leq t\}}\,\mathrm{d}t.\end{aligned}$$

Hence, the distributional time derivative of $G(\cdot,\cdot;s,y)$ on $\mathbb{R}^+ \times \mathbb{R}$ is a measure on $\mathbb{R}^+ \times \mathbb{R}$ that we will henceforth denote by $\partial G/\partial t\,(\mathrm{d}t,\mathrm{d}x;s,y)$ and such that

$$\frac{\partial G}{\partial t}(\mathrm{d}t,\mathrm{d}x;s,y) = \frac{1}{2}\delta_{y\pm(t-s)}(\mathrm{d}x)\mathbb{1}_{\{s\leq t\}}\,\mathrm{d}t = \frac{\mathrm{d}G}{\mathrm{d}x}(t,\mathrm{d}x;s,y)\mathbb{1}_{\{s\leq t\}}\,\mathrm{d}t \tag{2.3.9}$$

where $\mathrm{d}G/\mathrm{d}x\,(t,\mathrm{d}x;s,y)$ denotes the distributional derivative of $G(t,\cdot;s,y)$ for fixed t,s,y, which is readily seen to be equal to $(1/2)\delta_{y\pm(t-s)}(\mathrm{d}x)$ whenever $t \geq s$ and to 0 otherwise.

Using the expression (2.2.4) of u^ε, the stochastic Fubini theorem (see, for example, Theorem 2.6 in [99]) and (2.3.9), we further have

$$\begin{aligned}\int_{\mathbb{R}^+\times\mathbb{R}} & u^\varepsilon(t,x)\partial_t\Psi(t,x)\,\mathrm{d}t\,\mathrm{d}x \\ &= \int_{\mathbb{R}^+\times\mathbb{R}}\left(\int_{\mathbb{R}^+\times\mathbb{R}} G(t,x;s,y)\partial_t\Psi(t,x)\,\mathrm{d}t\,\mathrm{d}x\right)\frac{f(u^\varepsilon(s,y))}{\sigma(\varepsilon)}\,L^\varepsilon(\mathrm{d}s,\mathrm{d}y) \\ &= -\int_{\mathbb{R}^+\times\mathbb{R}}\left(\int_0^\infty\int_{\mathbb{R}} \Psi(t,x)\,\frac{\mathrm{d}G}{\mathrm{d}x}(t,\mathrm{d}x;s,y)\mathbb{1}_{\{s\leq t\}}\,\mathrm{d}t\right)\frac{f(u^\varepsilon(s,y))}{\sigma(\varepsilon)}\,L^\varepsilon(\mathrm{d}s,\mathrm{d}y) \\ &= -\int_0^\infty\left(\int_0^t\int_{\mathbb{R}}\left(\int_{\mathbb{R}}\Psi(t,x)\,\frac{\mathrm{d}G}{\mathrm{d}x}(t,\mathrm{d}x;s,y)\right)\frac{f(u^\varepsilon(s,y))}{\sigma(\varepsilon)}\,L^\varepsilon(\mathrm{d}s,\mathrm{d}y)\right)\mathrm{d}t \quad \mathbb{P}\text{-a.s.}\end{aligned} \tag{2.3.10}$$

Recall now that the Schwartz space $\mathcal{S}(\mathbb{R})$ consists of all $C^\infty(\mathbb{R})$-functions with rapid decrease, see e.g. Definition 4.1 in [44]. It has a natural topology induced by the seminorms $\sup_{x\in\mathbb{R}}|x^j\phi^{(k)}(x)|$ with $j,k \in \mathbb{N}$ and $\phi \in \mathcal{S}(\mathbb{R})$. We define for each $\varepsilon > 0$, an $\mathcal{S}'(\mathbb{R})$-valued stochastic process

$$\begin{aligned} v^\varepsilon : \mathbb{R}^+ &\longrightarrow \mathcal{S}'(\mathbb{R}) \\ t &\mapsto \left[\phi \mapsto \int_0^t\int_{\mathbb{R}}\left(\int_{\mathbb{R}}\phi(x)\,\frac{\mathrm{d}G}{\mathrm{d}x}(t,\mathrm{d}x;s,y)\right)\frac{f(u^\varepsilon(s,y))}{\sigma(\varepsilon)}\,L^\varepsilon(\mathrm{d}s,\mathrm{d}y)\right].\end{aligned} \tag{2.3.11}$$

This stochastic integral is well-defined because by (2.2.7), the Lipschitz continuity of f and (2.3.9),

$$\begin{aligned}
&\mathbb{E}\left[\int_0^t \int_{\mathbb{R}^2} \left(\int_{\mathbb{R}} \phi(x)\, \frac{\mathrm{d}G}{\mathrm{d}x}(t, \mathrm{d}x; s, y)\right)^2 \frac{f^2(u^\varepsilon(s,y))}{\sigma^2(\varepsilon)} z^2\, \nu^\varepsilon(\mathrm{d}s, \mathrm{d}y, \mathrm{d}z)\right] \\
&\qquad = \frac{1}{4}\int_0^t \int_{\mathbb{R}} \phi^2(y \pm (t-s))\mathbb{E}\left[f^2(u^\varepsilon(s,y))\right] \mathrm{d}s\, \mathrm{d}y \left(\frac{1}{\sigma^2(\varepsilon)} \int_{\mathbb{R}} z^2\, Q^\varepsilon(\mathrm{d}z)\right) \\
&\qquad \le C\int_0^t \int_{\mathbb{R}} \phi^2(y \pm (t-s))\, \mathrm{d}s\, \mathrm{d}y \le C \int_{\mathbb{R}} \phi^2(x)\, \mathrm{d}x < \infty \quad \text{for all} \quad \phi \in \mathcal{S}(\mathbb{R}) \quad \text{and} \quad t \ge 0.
\end{aligned} \tag{2.3.12}$$

Combining the definition of distributions and of v^ε, and with (2.3.10), we obtain the following representation for $\partial u^\varepsilon/\partial t$:

$$\left\langle \frac{\partial u^\varepsilon}{\partial t}, \Psi \right\rangle = \int_{\mathbb{R}^+} \langle v_t^\varepsilon, \Psi(t, \cdot)\rangle\, \mathrm{d}t \quad \text{for all} \quad \Psi \in C_c^\infty(\mathbb{R}^+ \times \mathbb{R}). \tag{2.3.13}$$

Furthermore, using v^ε, we can now derive mathematically a weak formulation of (2.2.1) corresponding to equation (2.3.8), see Proposition 2.5.6.

Actually, v_t^ε is not yet a random distribution: We only have for all Schwartz functions ϕ_1, ϕ_2 and scalars α_1, α_2, $\langle v_t^\varepsilon, \alpha_1\phi_1 + \alpha_2\phi_2\rangle = \alpha_1\langle v_t^\varepsilon, \phi_1\rangle + \alpha_2\langle v_t^\varepsilon, \phi_2\rangle$ $\mathbb{P}$-almost surely, so v_t^ε is not a linear functional but rather a random linear functional as defined in [99] on page 332. However, we can show that the random field $\{\langle v_t^\varepsilon, \phi\rangle \mid \phi \in \mathcal{S}(\mathbb{R})\}$ has a version with values in $\mathcal{S}'(\mathbb{R})$. For this, we first recall a few facts on $\mathcal{S}(\mathbb{R})$. For $q \in \mathbb{N}$, let h_q denote the qth Hermite function

$$h_q(x) = \frac{(-1)^q}{(2^q q!\sqrt{\pi})^{1/2}} e^{x^2/2} \frac{\mathrm{d}^q}{\mathrm{d}x^q} e^{-x^2}, \quad x \in \mathbb{R}.$$

As is well-known, $h_q \in \mathcal{S}(\mathbb{R})$ and a possible orthonormal basis of $L^2(\mathbb{R})$ is given by $\{h_q \mid q \in \mathbb{N}\}$. Define now for each $r \ge 0$, the function space

$$H_r(\mathbb{R}) = \left\{\phi \in L^2(\mathbb{R}) \;\middle|\; \sum_{q=0}^\infty (1+2q)^r \langle \phi, h_q\rangle^2 < \infty\right\}. \tag{2.3.14}$$

Note that this is *not* the fractional Sobolev space on $\mathbb{R}$ of order r which is usually defined via the Fourier transform. It is a Hilbert space whose topology is induced by the norm $\|\phi\|_r = \sqrt{\langle \phi, \phi\rangle_r}$ with the scalar product

$$\langle \phi, \varphi\rangle_r = \sum_{q=0}^\infty (1+2q)^r \langle \phi, h_q\rangle\langle \varphi, h_q\rangle \quad \text{for all} \quad \phi, \varphi \in H_r(\mathbb{R}). \tag{2.3.15}$$

We denote the topological dual of $H_r(\mathbb{R})$ by $H_{-r}(\mathbb{R})$ with dual norm $\|\cdot\|_{-r}$. For each $r \ge 0$ and $q \in \mathbb{N}$, consider the continuous and linear functional

$$e_{q,-r} \colon H_r(\mathbb{R}) \longrightarrow \mathbb{R}, \quad \phi \mapsto (1+2q)^{-r/2}\langle \phi, h_q\rangle_r. \tag{2.3.16}$$

By the Riesz representation theorem, the duality $\langle \phi', e_{q,-r}\rangle_{-r} = (1+2q)^{-r/2}\langle \phi', h_q\rangle$ holds, the set $\{e_{q,-r} \mid q \in \mathbb{N}\}$ forms an orthonormal basis of $H_{-r}(\mathbb{R})$ and for all $\phi' \in H_{-r}(\mathbb{R})$, we have

$$\phi' = \sum_{q=0}^\infty (1+2q)^{-r/2}\langle \phi', h_q\rangle e_{q,-r} \quad \text{in} \quad H_{-r}(\mathbb{R}) \quad \text{and} \quad \|\phi'\|_{-r}^2 = \sum_{q=0}^\infty (1+2q)^{-r}\langle \phi', h_q\rangle^2. \tag{2.3.17}$$

By Example 2 in Chapter 4 of [99], $\mathcal{S}(\mathbb{R})$ is then a *nuclear space*, see e.g. pages 330–332 of that chapter for a definition. In particular, $\mathcal{S}(\mathbb{R}) \subseteq H_r(\mathbb{R})$ for all $r \geq 0$ and the injection $(\mathcal{S}(\mathbb{R}), \|\cdot\|_s) \hookrightarrow (\mathcal{S}(\mathbb{R}), \|\cdot\|_r)$ is a Hilbert-Schmidt operator if $r < s+1$ (and not $r < s+1/2$ as indicated in that example, which is a typo). We then obtain the following regularization.

Proposition 2.3.4. *For any $r > 1$, $\varepsilon > 0$ and $t \geq 0$, the random field $\{\langle v_t^\varepsilon, \phi\rangle \mid \phi \in \mathcal{S}(\mathbb{R})\}$ has a version which is in $H_{-r}(\mathbb{R})$ and hence, in $\mathcal{S}'(\mathbb{R})$ as well.*

Proof. This is a direct application of Theorem 4.1 in [99]: By the calculation in (2.3.12),

$$\mathbb{E}\left[\left|\langle v_t^\varepsilon, \phi\rangle - \langle v_t^\varepsilon, \varphi\rangle\right|^2\right] \leq C \int_{\mathbb{R}} |\phi(x) - \varphi(x)|^2 \,\mathrm{d}x \quad \text{for all} \quad \phi, \varphi \in \mathcal{S}(\mathbb{R}),$$

so v_t^ε is continuous in probability in the norm $\|\cdot\|_r$ for any $r \geq 0$. □

As a consequence, we may and will assume from now on that $v_t^\varepsilon \in H_{-r}(\mathbb{R})$ for arbitrary $r > 1$. We then obtain the following path property.

Theorem 2.3.5. *For any $r > 2$ and $\varepsilon > 0$, the process v^ε introduced in (2.3.11) has a version $\overline{v}^\varepsilon$ in $D(\mathbb{R}^+, H_{-r}(\mathbb{R}))$, the Skorokhod space of $H_{-r}(\mathbb{R})$-valued càdlàg functions on $\mathbb{R}^+$.*

In the remainder of this paper, we will work with the càdlàg process $\overline{v}^\varepsilon$ obtained in Theorem 2.3.5 instead of $\partial u^\varepsilon/\partial t$. Here we point out that even though we will investigate in Section 2.4 convergence in distribution on *finite* intervals, for technical reasons only (e.g. to avoid tedious calculations related to the boundaries of the interval), we chose in this work a space of distributions on the whole of $\mathbb{R}$.

Finally, we follow the same scheme for the continuous mild solution u to (2.2.9) and by the proofs of Proposition 2.3.4 and Theorem 2.3.5, for any $r > 2$, there exists a unique continuous process $\overline{v}$ with values in $H_{-r}(\mathbb{R})$ such that for all $\phi \in \mathcal{S}(\mathbb{R})$ and $t \geq 0$,

$$\begin{aligned} \langle \overline{v}_t, \phi\rangle &= \int_0^t \int_{\mathbb{R}} \left(\int_{\mathbb{R}} \phi(x) \frac{\mathrm{d}G}{\mathrm{d}x}(t, \mathrm{d}x; s, y)\right) f(u(s,y))\, W(\mathrm{d}s, \mathrm{d}y) \\ &= \frac{1}{2} \int_0^t \int_{\mathbb{R}} \phi(y \pm (t-s)) f(u(s,y))\, W(\mathrm{d}s, \mathrm{d}y) \quad \mathbb{P}\text{-almost surely.} \end{aligned} \tag{2.3.18}$$

2.4 Main result

In this section, we fix $T > 0$ as well as $L > 0$. Consider the Cartesian space

$$\Omega^\dagger = \Big(D_{\preceq}([0,T] \times [0,L]) \cap L^2([0,T] \times [0,L])\Big) \times D([0,T], \mathbb{R}).$$

Let ϱ be defined as the sum of the metrics τ and d_1, where τ is the Skorokhod metric on $D_{\preceq}([0,T] \times [0,L])$, see Lemma 2.3.3, and d_1 the metric induced by the standard L^2-norm on $L^2([0,T] \times [0,L])$. Let also $\tau^\dagger$ denote the usual Skorokhod metric on $D([0,T], \mathbb{R})$. We equip $\Omega^\dagger$ with the product metric

$$\chi^\dagger((f_1, g_1), (f_2, g_2)) = \varrho(f_1, f_2) + \tau^\dagger(g_1, g_2) = (\tau(f_1, f_2) + d_1(f_1, f_2)) + \tau^\dagger(g_1, g_2) \tag{2.4.1}$$

for all $(f_1, g_1), (f_2, g_2) \in \Omega^\dagger$. The main result of this paper is the following limit theorem.

Theorem 2.4.1. *Let L^ε be as in* (2.2.2) *with variance $\sigma^2(\varepsilon)$ as in* (2.2.3) *for all $\varepsilon > 0$. Further let u^ε be a mild solution to the stochastic wave equation* (2.2.1) *driven by $\sigma^{-1}(\varepsilon)\dot{L}^\varepsilon$, $\overline{u}^\varepsilon$ its $\preceq$-càdlàg version given by Theorem 2.3.2 and $\overline{v}^\varepsilon$ the càdlàg $H_{-r}(\mathbb{R})$-valued process obtained in Theorem 2.3.5 for an arbitrary fixed $r > 2$.*

In addition, let u be the continuous mild solution to the stochastic wave equation (2.2.9) *driven by $\dot{W}$ and $\overline{v}$ the continuous $H_{-r}(\mathbb{R})$-valued process satisfying* (2.3.18)*.*

Suppose the Lipschitz function f satisfies $f(0) \neq 0$. We then have

$$(\overline{u}^\varepsilon, \langle \overline{v}^\varepsilon, \phi\rangle) \xrightarrow{d} (u, \langle \overline{v}, \phi\rangle) \quad \textit{in} \quad \left(\Omega^\dagger, \chi^\dagger\right) \quad \textit{as} \quad \varepsilon \to 0 \quad \textit{for all} \quad \phi \in C_c^\infty((0, L)) \tag{2.4.2}$$

if and only if condition (2.1.3) *holds for each $\kappa > 0$.*

Remark 2.4.2. The weak convergence of $(\langle \overline{u}^\varepsilon(t,\cdot), \phi\rangle, \langle \overline{v}_t^\varepsilon, \phi\rangle)_{t\leq T}$ for all $\phi \in C_c^\infty((0, L))$, is needed to show the necessity of (2.1.3). That is why the distribution-valued processes $\overline{v}^\varepsilon$ and $\overline{v}$ were added to the actual normal approximation of $\overline{u}^\varepsilon$ by u.

Proof of Theorem 2.4.1. In a first part, we show that (2.1.3) implies (2.4.2). For any fixed $\phi \in C_c^\infty((0, L))$, we will show convergence in distribution of subsequences of $\{(\overline{u}^\varepsilon, \langle \overline{v}^\varepsilon, \phi\rangle) \mid \varepsilon > 0\}$ toward the limit distribution $(u, \langle \overline{v}, \phi\rangle)$. To this end, we need to consider $\overline{u}^\varepsilon$ on the larger domain $[0, T] \times [-T, L + T]$. This is due to the Green's function G (recall (2.2.5)): The value of $u(t, x)$ for any $(t, x) \in [0, T] \times [0, L]$ depends on values taken by u on $\{(s, x) \in \mathbb{R}^+ \times \mathbb{R} \mid 0 \leq s \leq t,\ \mathrm{dist}(x, [0, L]) \leq t - s\} \subseteq [0, T] \times [-T, L + T]$. This larger domain will be necessary for the proof of Theorem 2.5.7. Since we also need to work with $\overline{v}^\varepsilon$, in order to prove (2.4.2), we shall work with the second Cartesian space

$$\begin{aligned}\Omega^* = &\left(D_\preceq\left([0, T] \times [-T, L + T]\right) \cap L^2\left([0, T] \times [-T, L + T]\right)\right)\\ &\times \left(D([0, T], H_{-r}(\mathbb{R})) \cap L^2([0, T], H_{-r}(\mathbb{R}))\right).\end{aligned} \tag{2.4.3}$$

Let ρ be the sum of the usual Skorokhod metric τ^* on $D([0, T], H_{-r}(\mathbb{R}))$ and of the standard L^2-metric d_2 on $L^2([0, T], H_{-r}(\mathbb{R}))$. We then equip Ω^* with the product metric

$$\chi^*((f_1, g_1), (f_2, g_2)) = \varrho(f_1, f_2) + \rho(g_1, g_2) = (\tau(f_1, f_2) + d_1(f_1, f_2)) + (\tau^*(g_1, g_2) + d_2(g_1, g_2))$$

for all $(f_1, g_1), (f_2, g_2) \in \Omega^*$.

By Theorem 2.5.1 and Theorem 2.5.2, $\overline{u}^\varepsilon$ is tight both in $D_\preceq([0, T] \times [-T, L + T])$ and in $L^2([0, T] \times [-T, L + T])$. This readily implies that $\overline{u}^\varepsilon$ is also tight in $(D_\preceq([0, T] \times [-T, L + T]) \cap L^2([0, T] \times [-T, L + T]), \varrho)$ (it is easy to see that if K_1 and K_2 are compact sets, one in each function space, then $K_1 \cap K_2$ is compact in the intersection space considered). Analogously, $\overline{v}^\varepsilon$ is tight in $(D([0, T], H_{-r}(\mathbb{R})) \cap L^2([0, T], H_{-r}(\mathbb{R})), \rho)$ as a consequence of Theorem 2.5.3 and Corollary 2.5.5. Since the product of compact spaces is compact, we draw the crucial conclusion that $\{(\overline{u}^\varepsilon, \overline{v}^\varepsilon) \mid \varepsilon > 0\}$ is tight in (Ω^*, χ^*). Note that no assumptions other than (2.2.2) and (2.2.3) on the Lévy noise are needed for this result.

Subsequently, apply Prokhorov's theorem and let without loss of generality $(\varepsilon_k)_{k\in\mathbb{N}}$ be a sequence with $\varepsilon_k \to 0$ such that $(\overline{u}^{\varepsilon_k}, \overline{v}^{\varepsilon_k})_{k\in\mathbb{N}}$ converges weakly to some distribution on (Ω^*, χ^*) as $k \to \infty$. Then we may further apply Skorokhod's representation theorem, see e.g. Section 1 in [68], and obtain random elements

$$(w^k, \theta^k), (w, \theta) \colon (\overline{\Omega}, \overline{\mathcal{F}}, \overline{\mathbb{P}}) \longrightarrow (\Omega^*, \chi^*), \tag{2.4.4}$$

defined on a common probability space $(\overline{\Omega}, \overline{\mathcal{F}}, \overline{\mathbb{P}})$, and satisfying the following properties:

$$\begin{aligned} &(w^k, \theta^k) \overset{d}{=} (\overline{u}^{\varepsilon_k}, \overline{v}^{\varepsilon_k}) \quad \text{for all} \quad k \in \mathbb{N} \quad \text{and} \\ &(w^k, \theta^k)(\overline{\omega}) \longrightarrow (w, \theta)(\overline{\omega}) \quad \text{in} \quad (\Omega^*, \chi^*) \quad \text{as} \quad k \to \infty \quad \text{for all} \quad \overline{\omega} \in \overline{\Omega}. \end{aligned} \tag{2.4.5}$$

We will show in the following that for any $\phi \in C_c^\infty((0, L))$,

$$(w, \langle \theta, \phi \rangle) \overset{d}{=} (u, \langle \overline{v}, \phi \rangle) \quad \text{in} \quad (\Omega^\dagger, \chi^\dagger), \tag{2.4.6}$$

which together with (2.4.5) implies

$$(\overline{u}^{\varepsilon_k}, \langle \overline{v}^{\varepsilon_k}, \phi \rangle) \xrightarrow{d} (u, \langle \overline{v}, \phi \rangle) \quad \text{in} \quad \left(\Omega^\dagger, \chi^\dagger\right) \quad \text{as} \quad k \to \infty$$

by the continuous mapping theorem, and altogether, (2.4.2). In this identification step of the distribution of $(w, \langle \theta, \phi \rangle)$, we will refer to the parts of [32] that are identical.

First, define a filtration $\overline{\boldsymbol{F}} = (\overline{\mathcal{F}}_t)_{t \le T}$ on $(\overline{\Omega}, \overline{\mathcal{F}})$:

$$\overline{\mathcal{F}}_t = \bigcap_{u \ge t} \sigma\left(w(s, x), \theta_s \mid s \le u,\ -T \le x \le L + T\right) \vee \mathcal{N}^{\overline{\mathbb{P}}}, \quad 0 \le t \le T, \tag{2.4.7}$$

where $\mathcal{N}^{\overline{\mathbb{P}}}$ is the set of all $\overline{\mathbb{P}}$-null sets of $\overline{\mathcal{F}}$ (we assume that $\overline{\mathcal{F}}$ is $\overline{\mathbb{P}}$-complete), as well as

$$\overline{B}_t = \int_0^t \left(\int_{\mathbb{R}} w(s, x) \phi_2''(x) \, \mathrm{d}x + \langle \theta_s, \phi_1 \rangle \right) \mathrm{d}s \quad \text{and} \quad \overline{C}_t = \int_0^t \int_{\mathbb{R}} \phi_2^2(x) f^2(w(s, x)) \, \mathrm{d}s \, \mathrm{d}x \tag{2.4.8}$$

for $\phi_1, \phi_2 \in C_c^\infty((-T, L + T))$ and $t \le T$.

Assume now for the time being that the pair (w, θ) satisfies the following martingale problem: The complex-valued càdlàg process

$$\overline{M}_t = e^{i\xi(\langle w(t,\cdot), \phi_1 \rangle + \langle \theta_t, \phi_2 \rangle)} - \int_0^t e^{i\xi(\langle w(s,\cdot), \phi_1 \rangle + \langle \theta_s, \phi_2 \rangle)} \, \overline{A}(\mathrm{d}s) \quad \text{with} \quad \overline{A}_t = i\xi \overline{B}_t - \frac{1}{2}\xi^2 \overline{C}_t, \quad t \le T, \tag{2.4.9}$$

is a martingale with respect to $(\overline{\Omega}, \overline{\mathcal{F}}, \overline{\boldsymbol{F}}, \overline{\mathbb{P}})$ for all $\xi \in \mathbb{R}$ and $\phi_1, \phi_2 \in C_c^\infty((-T, L + T))$. (Note that because w is $\preceq$-càdlàg, the process $(\langle w(t, \cdot), \phi_1 \rangle)_{t \le T}$ is càdlàg, and that by a limit argument, w, θ as well as $\overline{M}$ are $\overline{\boldsymbol{F}}$-adapted.) Using (2.4.5), we also have

$$\operatorname*{ess\,sup}_{(t,x) \in [0,T] \times [-T, L+T]} \mathbb{E}\left[|w(t, x)|^2 \right] < \infty \quad \text{and} \quad \text{for all} \quad x \in \mathbb{R}, \quad w(0, x) = \theta_0 = 0 \quad \overline{\mathbb{P}}\text{-a.s.} \tag{2.4.10}$$

Indeed, the Skorokhod convergence of w^k implies for almost all $(t, x) \in [0, T] \times [-T, L + T]$, $w^k(t, x) \longrightarrow w(t, x)$ $\overline{\mathbb{P}}$-almost surely and because the projection maps $\pi_{(t,x)} : D_{\preceq}([0, T] \times [-T, L + T]) \longrightarrow \mathbb{R}$, $f \mapsto f(t, x)$ are measurable, we also have $w^k(t, x) \overset{d}{=} \overline{u}^{\varepsilon_k}(t, x)$ for all (t, x). Now the random fields $\{u^\varepsilon \mid \varepsilon > 0\}$ satisfy the uniform bound

$$\sup_{\varepsilon > 0} \sup_{(t,x) \in [0,T] \times \mathbb{R}} \mathbb{E}\left[|u^\varepsilon(t, x)|^2 \right] < \infty, \tag{2.4.11}$$

which only depends on the Lipschitz function f. With (2.2.8) and (2.2.7), the proof of (2.4.11) goes as Lemma 3.1 in [32]. (Note that (2.4.11) is also crucial for proving the existence of $\overline{u}^\varepsilon$ and tightness of $\{(\overline{u}^\varepsilon, \overline{v}^\varepsilon) \mid \varepsilon > 0\}$ in (Ω^*, χ^*).) Apply then Fatou's lemma to obtain (2.4.10).

Assumption (2.4.9) enables us, together with (2.4.10), to show that there exists a Gaussian space–time white noise $\widetilde{W}$ on $[0,T]\times[-T,L+T]$, possibly defined on a complete stochastic basis $(\widetilde{\Omega},\widetilde{\mathcal{F}},\widetilde{\boldsymbol{F}},\widetilde{\mathbb{P}})$ extending $(\overline{\Omega},\overline{\mathcal{F}},\overline{\boldsymbol{F}},\overline{\mathbb{P}})$ such that for all $\phi_1,\phi_2\in C_c^\infty((-T,L+T))$,

$$\int_{\mathbb{R}} w(t,x)\phi_1(x)\,\mathrm{d}x+\langle\theta_t,\phi_2\rangle = \int_0^t\Big(\int_{\mathbb{R}} w(s,x)\phi_2''(x)\,\mathrm{d}x+\langle\theta_s,\phi_1\rangle\Big)\,\mathrm{d}s+\int_0^t\int_{\mathbb{R}}\phi_2(x)f(w_-(s,x))\,\widetilde{W}(\mathrm{d}s,\mathrm{d}x)\quad\forall t\le T \tag{2.4.12}$$

holds $\widetilde{\mathbb{P}}$-almost surely, where we have set $w_-(s,x)=\lim_{r\to s,\,r<s}w(r,x)$. As a consequence, apply Theorem 2.5.7 to deduce that w is on $[0,T]\times[0,L]$ the continuous mild solution to the stochastic wave equation

$$\begin{cases}\partial_{tt}w(t,x)=\partial_{xx}w(t,x)+f(w(t,x))\dot{\widetilde{W}}(t,x), & (t,x)\in\mathbb{R}^+\times\mathbb{R},\\ w(0,x)=\partial_t w(0,x)=0, & \text{for all}\quad x\in\mathbb{R},\end{cases} \tag{2.4.13}$$

of which (2.4.12) is the weak formulation (on $[0,T]\times[-T,L+T]$), and that θ satisfies for all $\phi\in C_c^\infty((0,L))$ and $t\le T$,

$$\langle\theta_t,\phi\rangle=\int_0^t\int_{\mathbb{R}}\Big(\int_{\mathbb{R}}\phi(x)\frac{\mathrm{d}G}{\mathrm{d}x}(t,\mathrm{d}x;s,y)\Big)\,f(w(s,y))\,\widetilde{W}(\mathrm{d}s,\mathrm{d}y)\quad\widetilde{\mathbb{P}}\text{-almost surely.} \tag{2.4.14}$$

Recalling (2.3.18), we then infer that (2.4.6) holds.

Now we show how to obtain (2.4.12). Note that $\overline{C}$ in (2.4.8) only depends on ϕ_2, so fix $\phi_1\in C_c^\infty((-T,L+T))$ and apply Theorem II.2.42 of [65] on the process $(\overline{M}_t)_{t\le T}$ in (2.4.9) to first see that $(\overline{B},\overline{C},0)$, with $\overline{B}$ in (2.4.8), are the predictable characteristics of the $\overline{\boldsymbol{F}}$-semimartingale $(\langle w(t,\cdot),\phi_1\rangle+\langle\theta_t,\phi_2\rangle)_{t\le T}$. Observe that they do not directly depend on the process itself, but on w and θ, which also explains why we are working on the space (Ω^*,χ^*). Consequently,

$$M_t(\phi_2)=\int_{\mathbb{R}} w(t,x)\phi_1(x)\,\mathrm{d}x+\langle\theta_t,\phi_2\rangle-\overline{B}_t,\quad t\le T,$$

is a continuous square-integrable $\overline{\boldsymbol{F}}$-martingale with quadratic variation process $\overline{C}$ for all $\phi_2\in C_c^\infty((-T,L+T))$. This induces, relative to $(\overline{\Omega},\overline{\mathcal{F}},\overline{\boldsymbol{F}},\overline{\mathbb{P}})$, an orthogonal martingale measure $\{M_t(A),\,t\in[0,T],\,A\in\mathcal{B}([-T,L+T])\}$ (see Chapter 2 in [99] for a definition) with covariation measure $Q_M(A\times B\times[s,t])=\int_s^t\int_{A\cap B}f^2(w(r,x))\,\mathrm{d}r\,\mathrm{d}x$ for all $A,B\in\mathcal{B}([-T,L+T])$. Now use the proof of Theorem 3.13 in [32] to define, possibly on a complete filtered extension $(\widetilde{\Omega},\widetilde{\mathcal{F}},\widetilde{\boldsymbol{F}},\widetilde{\mathbb{P}})$,

$$\widetilde{W}_t(\phi_2)=\int_0^t\int_{\mathbb{R}}\mathbb{1}_{\{f^2(w_-(s,x))\neq0\}}\frac{\phi_2(x)}{f(w_-(s,x))}\,M(\mathrm{d}s,\mathrm{d}x)+\int_0^t\int_{\mathbb{R}}\mathbb{1}_{\{f^2(w_-(s,x))=0\}}\phi_2(x)\,W'(\mathrm{d}s,\mathrm{d}x) \tag{2.4.15}$$

where W' is a Gaussian white noise on $[0,T]\times[-T,L+T]$ independent of M, for all $t\le T$ and $\phi_2\in C_c^\infty((-T,L+T))$, and to further deduce that (2.4.15) defines a Gaussian white noise $\widetilde{W}$ on $[0,T]\times[-T,L+T]$ with respect to $(\widetilde{\Omega},\widetilde{\mathcal{F}},\widetilde{\boldsymbol{F}},\widetilde{\mathbb{P}})$ such that for all $\phi_1,\phi_2\in C_c^\infty((-T,L+T))$, (2.4.12) holds $\widetilde{\mathbb{P}}$-almost surely.

Of course, it remains to show that $\overline{M}$ in (2.4.9) is an $\overline{\boldsymbol{F}}$-martingale. To this end, consider first on Ω the càdlàg process $(\langle\overline{u}^\varepsilon(t,\cdot),\phi_1\rangle+\langle\overline{v}_t^\varepsilon,\phi_2\rangle)_{t\le T}$ with $\phi_1,\phi_2\in C_c^\infty((-T,L+T))$ and $\varepsilon>0$. By Proposition 2.5.6, it is indistinguishable from the right-hand side of (2.5.27). Replicate the proof of Theorem 3.8 in [32] and use (2.4.11) to see that for each $\varepsilon>0$, the pair $(\overline{u}^\varepsilon,\overline{v}^\varepsilon)$

satisfies the following martingale problem: For all $\xi \in \mathbb{R}$ and $\phi_1, \phi_2 \in C_c^\infty((-T, L+T))$, the complex-valued process

$$\begin{aligned} M_t^\varepsilon = e^{i\xi(\langle \overline{u}^\varepsilon(t,\cdot),\phi_1\rangle + \langle \overline{v}_t^\varepsilon,\phi_2\rangle)} - i\xi \int_0^t e^{i\xi(\langle \overline{u}^\varepsilon(s,\cdot),\phi_1\rangle + \langle \overline{v}_s^\varepsilon,\phi_2\rangle)} \left(\langle \overline{u}^\varepsilon(s,\cdot), \phi_2''\rangle + \langle \overline{v}_s^\varepsilon, \phi_1\rangle\right) \mathrm{d}s \\ + \int_0^t \int_{\mathbb{R}^2} e^{i\xi(\langle \overline{u}^\varepsilon(s,\cdot),\phi_1\rangle + \langle \overline{v}_s^\varepsilon,\phi_2\rangle)} \left(e^{i\xi \frac{f(\overline{u}^\varepsilon(s,x))}{\sigma(\varepsilon)}\phi_2(x)z} - 1 - i\xi \frac{f(\overline{u}^\varepsilon(s,x))}{\sigma(\varepsilon)}\phi_2(x)z \right) \mathrm{d}s\,\mathrm{d}x\,Q^\varepsilon(\mathrm{d}z) \end{aligned} \tag{2.4.16}$$

is a square-integrable $\boldsymbol{F}$-martingale satisfying $\sup_{\varepsilon>0}\sup_{t\le T}\mathbb{E}[|M_t^\varepsilon|^2] < \infty$.

Define now on $\overline{\Omega}$ the filtration $\overline{\boldsymbol{F}}^k = (\overline{\mathcal{F}}_t^k)_{t\le T}$ with

$$\overline{\mathcal{F}}_t^k = \bigcap_{u\ge t} \sigma\left(w^k(s,x), \theta_s^k \mid s \le u,\ -T \le x \le L+T\right) \vee \mathcal{N}^{\overline{\mathbb{P}}}, \quad 0 \le t \le T, \tag{2.4.17}$$

for each $k \in \mathbb{N}$, as well as the $\overline{\boldsymbol{F}}^k$-adapted process $(\overline{M}_t^k)_{t\le T}$ in the same way as M^ε in (2.4.16), but with $(\overline{u}^\varepsilon, \overline{v}^\varepsilon)$ and Q^ε replaced by (w^k, θ^k) of (2.4.5) and Q^{ε_k}, respectively. Furthermore, because $\overline{M}^k$ has the same distribution as M^{ε_k} by (2.4.5), by standard arguments, $\overline{M}^k$ is a square-integrable $\overline{\boldsymbol{F}}^k$-martingale satisfying $\sup_{k\in\mathbb{N}}\sup_{t\le T}\mathbb{E}[|\overline{M}_t^k|^2] < \infty$ for all $\xi \in \mathbb{R}$ and $\phi_1, \phi_2 \in C_c^\infty((-T, L+T))$. This is the martingale problem satisfied by the pair (w^k, θ^k). For any fixed $\xi \in \mathbb{R}$ and $\phi_1, \phi_2 \in C_c^\infty((-T, L+T))$, we can infer that $\overline{M}$ is an $\overline{\boldsymbol{F}}$-martingale, again by standard arguments, if we have:

$$\text{for almost all} \quad t \le T, \quad \overline{M}_t^k \longrightarrow \overline{M}_t \quad \text{as} \quad k \to \infty \quad \overline{\mathbb{P}}\text{-almost surely.} \tag{2.4.18}$$

Indeed, with this convergence result and the càdlàg properties of $\overline{M}$, θ and w, we can infer that

$$\mathbb{E}\left[\left(\overline{M}_t - \overline{M}_s\right) \overline{h}(\theta_{s_1}, \dots, \theta_{s_M}) h(w(r_1, x_1), \dots, w(r_N, x_N))\right] = 0 \tag{2.4.19}$$

for all continuous bounded functions $\overline{h} \colon H_{-r}(\mathbb{R})^M \to \mathbb{R}$, $h \colon \mathbb{R}^N \to \mathbb{R}$ with $M, N \in \mathbb{N}$ and all $0 \le s < t \le T$, $s_i, r_j \le s$ and $x_j \in [-T, L+T]$ with $i = 1, \dots, M$, $j = 1, \dots, N$.

Now in order to show (2.4.18), which is the final step, first set for each $k \in \mathbb{N}$,

$$\begin{aligned} \overline{\nu}^k(A) &= \int_0^T \int_{\mathbb{R}^2} \mathbb{1}_A\left(t, \frac{f(w^k(t,x))}{\sigma(\varepsilon_k)}\phi_2(x)z\right) \mathrm{d}t\,\mathrm{d}x\,Q^{\varepsilon_k}(\mathrm{d}z), \\ \overline{B}_t^k &= \int_0^t \left(\langle w^k(s,\cdot), \phi_2''\rangle + \langle \theta_s^k, \phi_1\rangle\right) \mathrm{d}s - \int_0^t \int_{\mathbb{R}} x \mathbb{1}_{\{|x|>1\}}\, \overline{\nu}^k(\mathrm{d}s, \mathrm{d}x), \\ \overline{A}_t^k &= i\xi \overline{B}_t^k + \int_0^t \int_{\mathbb{R}} \left(e^{i\xi x} - 1 - i\xi x \mathbb{1}_{\{|x|\le 1\}}\right) \overline{\nu}^k(\mathrm{d}s, \mathrm{d}x) \end{aligned} \tag{2.4.20}$$

for all $A \in \mathcal{B}([0,T]\times\mathbb{R})$ and $t \le T$, so that $\overline{M}^k$ can be written as

$$\overline{M}_t^k = e^{i\xi\left(\langle w^k(t,\cdot),\phi_1\rangle + \langle \theta_t^k,\phi_2\rangle\right)} - \int_0^t e^{i\xi\left(\langle w^k(s,\cdot),\phi_1\rangle + \langle \theta_s^k,\phi_2\rangle\right)}\, \overline{A}^k(\mathrm{d}s), \quad t \le T. \tag{2.4.21}$$

Now replicate the proofs of Theorem 3.9, Theorem 3.10 and Lemma 3.11 in [32] to see that the assumption (2.1.3) on the Lévy measure Q^ε and (2.4.5) imply

$$\sup_{t\le T}\left| \int_0^t e^{i\xi\left(\langle w^k(s,\cdot),\phi_1\rangle + \langle \theta_s^k,\phi_2\rangle\right)}\, \overline{A}^k(\mathrm{d}s) - \int_0^t e^{i\xi(\langle w(s,\cdot),\phi_1\rangle + \langle \theta_s,\phi_2\rangle)}\, \overline{A}(\mathrm{d}s) \right| \longrightarrow 0 \quad \text{as} \quad k \to \infty \tag{2.4.22}$$

pointwise on $\overline{\Omega}$. Note that it is the only place in our proof where (2.1.3) is actually needed. For the proofs of the aforementioned theorems to actually hold here, we need the extra convergence $\int_0^t \langle \theta_s^k, \phi_1 \rangle \,\mathrm{d}s \longrightarrow \int_0^t \langle \theta_s, \phi_1 \rangle \,\mathrm{d}s$. But this readily follows from $\theta^k \longrightarrow \theta$ in $L^2([0,T], H_{-r}(\mathbb{R}))$ in (2.4.5). Recalling the expression of $\overline{M}$ in (2.4.9), resp. of $\overline{M}^k$ in (2.4.21), it is now easy to see that (2.4.18) follows from (2.4.22), the Skorokhod convergence of θ^k and the L^2-convergence of w^k.

For the second part of the proof, assume (2.4.2) and fix a sequence $(\varepsilon_k)_{k\in\mathbb{N}}$ converging to 0. We apply again Skorokhod's representation theorem and obtain for each $\phi \in C_c^\infty((0,L))$, random elements

$$(w^k, \vartheta^{k,\phi}), (w, \vartheta^\phi) \colon (\overline{\Omega}, \overline{\mathcal{F}}, \overline{\mathbb{P}}) \longrightarrow (\Omega^\dagger, \chi^\dagger)$$

on a probability space $(\overline{\Omega}, \overline{\mathcal{F}}, \overline{\mathbb{P}})$ possibly different from $(\Omega, \mathcal{F}, \mathbb{P})$ but that does *not* depend on ϕ, satisfying

$$\begin{aligned} &(w^k, \vartheta^{k,\phi}) \overset{d}{=} (\overline{u}^{\varepsilon_k}, \langle \overline{v}^{\varepsilon_k}, \phi \rangle) \quad \text{for all} \quad k \in \mathbb{N}, \quad (w, \vartheta^\phi) \overset{d}{=} (u, \langle \overline{v}, \phi \rangle) \quad \text{and} \\ &(w^k, \vartheta^{k,\phi})(\overline{\omega}) \longrightarrow (w, \vartheta^\phi)(\overline{\omega}) \quad \text{in} \quad (\Omega^\dagger, \chi^\dagger) \quad \text{as} \quad k \to \infty \quad \text{for all} \quad \overline{\omega} \in \overline{\Omega}. \end{aligned} \tag{2.4.23}$$

We then define filtrations $\overline{\boldsymbol{F}}^k = (\overline{\mathcal{F}}_t^k)_{t\le T}$ and $\overline{\boldsymbol{F}} = (\overline{\mathcal{F}}_t)_{t\le T}$ on $(\overline{\Omega}, \overline{\mathcal{F}}, \overline{\mathbb{P}})$ by setting

$$\begin{aligned} \overline{\mathcal{F}}_t^k &= \bigcap_{u\ge t} \sigma\left(w^k(s,x), \vartheta_s^{k,\phi} \mid s \le u,\, 0 \le x \le L,\, \phi \in C_c^\infty((0,L))\right) \vee \mathcal{N}^{\overline{\mathbb{P}}}, \\ \overline{\mathcal{F}}_t &= \bigcap_{u\ge t} \sigma\left(w(s,x), \vartheta_s^{\phi} \mid s \le u,\, 0 \le x \le L,\, \phi \in C_c^\infty((0,L))\right) \vee \mathcal{N}^{\overline{\mathbb{P}}}, \quad 0 \le t \le T, \end{aligned} \tag{2.4.24}$$

as well as for arbitrary fixed $\phi \in C_c^\infty((0,L))$, the càdlàg processes

$$\begin{aligned} \overline{X}_t^k &= \vartheta_t^{k,\phi} - \int_0^t \int_0^L w^k(s,x)\phi''(x)\,\mathrm{d}x\,\mathrm{d}s, \\ \overline{X}_t &= \vartheta_t^{\phi} - \int_0^t \int_0^L w(s,x)\phi''(x)\,\mathrm{d}x\,\mathrm{d}s \end{aligned} \tag{2.4.25}$$

for all $k \in \mathbb{N}$ and $t \le T$. Since $\overline{v}$ is continuous, this is also the case for the real-valued process ϑ^ϕ by (2.4.23), hence $\overline{X}$ is continuous. Now (2.4.23) readily implies

$$\overline{X}^k \longrightarrow \overline{X} \quad \text{in} \quad D([0,T], \mathbb{R}) \quad \text{as} \quad k \to \infty$$

pointwise on $\overline{\Omega}$. Furthermore, by (the proof of) Proposition 2.5.6, (2.4.23) and (2.4.25), the processes $\overline{X}^k$ and $\overline{X}$ have the same distribution as the square-integrable $\boldsymbol{F}$-martingales

$$t \mapsto \int_0^t \int_0^L \phi(x) \frac{f(u^{\varepsilon_k}(s,x))}{\sigma(\varepsilon_k)} \, L^{\varepsilon_k}(\mathrm{d}s, \mathrm{d}x) \quad \text{and} \quad t \mapsto \int_0^t \int_0^L \phi(x) f(u(s,x))\, W(\mathrm{d}s, \mathrm{d}x),$$

respectively. By standard arguments, we can thus deduce that $\overline{X}^k$, resp. $\overline{X}$, is an $\overline{\boldsymbol{F}}^k$-martingale, resp. $\overline{\boldsymbol{F}}$-martingale.

Consider the truncation functions

$$\varrho_h \colon \mathbb{R} \longrightarrow \mathbb{R}, \quad x \mapsto x\mathbb{1}_{\{|x|\le h\}}, \quad h > 0,$$

and apply Theorem II.2.21 in [65] to see that the semimartingale characteristics of $\overline{X}^k$ and $\overline{X}$, relative to ϱ_h for a fixed but arbitrary $h>0$, are given by $(\overline{B}^{k,h}, 0, \overline{\nu}^k)$ and $(0, \overline{C}, 0)$, respectively, where $\overline{\nu}^k$ is defined as in (2.4.20) and $\overline{C}$ as in (2.4.8) (with ϕ_2 replaced by ϕ), and

$$\overline{B}_t^{k,h} = -\int_0^t \int_{\mathbb{R}} x \mathbb{1}_{\{|x|>h\}}\, \overline{\nu}^k(\mathrm{d}s, \mathrm{d}x), \quad t \leq T.$$

The remainder of the proof now goes exactly as the proof of Theorem 3.15 in [32] (where the remaining assumption $f(0) \neq 0$ of the theorem is then needed). □

2.5 Proofs

2.5.1 Proofs for Section 2.3 and for tightness

We begin by showing that each u^ε has a $\preceq$-càdlàg version.

Proof of Theorem 2.3.2. Fix $\varepsilon > 0$ for the whole proof. For each $n \in \mathbb{N}$, we introduce a truncated Lévy space–time white noise $\dot{L}^{\varepsilon,n}$ on $\mathbb{R}^+ \times \mathbb{R}$ by setting

$$L^{\varepsilon,n}(A) = \int_{\mathbb{R}^+\times\mathbb{R}} \int_{\mathbb{R}} \mathbb{1}_A(t,x) \mathbb{1}_{\{|x|\leq n\}}\, z \mathbb{1}_{\{|z|>1/n\}}\, (\mu^\varepsilon - \nu^\varepsilon)(\mathrm{d}t, \mathrm{d}x, \mathrm{d}z) \tag{2.5.1}$$

for all $A \in \mathcal{B}_b(\mathbb{R}^+ \times \mathbb{R})$. Now let $u^{\varepsilon,n}$ be a mild solution to the stochastic wave equation (2.2.1) when $\sigma^{-1}(\varepsilon)\dot{L}^\varepsilon$ is replaced by $\sigma^{-1}(\varepsilon)\dot{L}^{\varepsilon,n}$. Because $\dot{L}^{\varepsilon,n}$ generates on $[0,t]\times\mathbb{R}$ finitely many jumps only, we can write for all $(t,x) \in \mathbb{R}^+ \times \mathbb{R}$ and $n \in \mathbb{N}$,

$$\begin{aligned} u^{\varepsilon,n}(t,x) = \frac{1}{\sigma(\varepsilon)} \sum_{k=1}^\infty G_{t-T_k}(x, X_k) f(u^{\varepsilon,n}(T_k, X_k)) \mathbb{1}_{\{|X_k|\leq n\}}\, Z_k \mathbb{1}_{\{|Z_k|>1/n\}} \\ - \frac{\int_{|z|>1/n} z\, Q^\varepsilon(\mathrm{d}z)}{\sigma(\varepsilon)} \int_0^t \int_{|y|\leq n} G_{t-s}(x,y) f(u^{\varepsilon,n}(s,y))\, \mathrm{d}s\, \mathrm{d}y \quad \mathbb{P}\text{-almost surely} \end{aligned} \tag{2.5.2}$$

where the T_k indicate the jump times of μ^ε and X_k (resp. Z_k) the space locations (resp. amplitudes) of the jumps of μ^ε.

The random field on the right-hand side of (2.5.2) is a $\preceq$-càdlàg version of $u^{\varepsilon,n}$. Indeed, through the reformulation of the Green's function

$$G_{t-s}(x,y) = \frac{1}{2}\mathbb{1}_{A^-(s,y)}(t,x), \quad (t,x,s,y) \in (\mathbb{R}^+\times\mathbb{R})^2,$$

where

$$A^-(s,y) = \Big\{(t,x) \in \mathbb{R}^+ \times \mathbb{R} \mid |y-x| \leq t-s\Big\} \tag{2.5.3}$$

denotes the forward light cone with apex (s,y), one sees that $(t,x) \mapsto G_{t-T_k}(x, X_k)$ is already $\preceq$-càdlàg and, hence, the finite sum as well as the integral in (2.5.2) are $\mathbb{P}$-almost surely $\preceq$-càdlàg.

We will show that $u^{\varepsilon,n}$ converges uniformly on compact sets of $\mathbb{R}^+\times\mathbb{R}$ in probability to u^ε as $n \to \infty$. For this, assume first without loss of generality using Theorem 2 in Chapter 3, § 2 in [52], that u^ε and $u^{\varepsilon,n}$ are separable random fields. The first step is to obtain the maximal inequality

$$\begin{aligned} \mathbb{E}\left[\sup_{(s,y)\in[(\tilde{t},\tilde{x}),(t,x)]_\preceq} \left|u^\varepsilon(s,y) - u^{\varepsilon,n}(s,y)\right|^2\right] \leq \mathbb{E}\left[\sup_{(s,y)\in A^+(t,x)} \left|u^\varepsilon(s,y) - u^{\varepsilon,n}(s,y)\right|^2\right] \\ \leq \sup_{(s,y)\in A^+(t,x)} \mathbb{E}\left[\left|u^\varepsilon(s,y) - u^{\varepsilon,n}(s,y)\right|^2\right] = \mathbb{E}\left[\left|u^\varepsilon(t,x) - u^{\varepsilon,n}(t,x)\right|^2\right] \end{aligned} \tag{2.5.4}$$

for all $(\tilde{t},\tilde{x}) \preceq (t,x)$ in $\mathbb{R}^+ \times \mathbb{R}$. Choose for simplicity $(\tilde{t},\tilde{x}) = 0$ as well as $x = 0$ and fix $t > 0$. Recall the change of coordinates H introduced in (2.3.2) and define $K(u) = H(u - u_0)$ on $\mathbb{R}^2$ with $u_0 = (-t, 0)$. Then K builds a bijection of $[u_0, u^*]_\preceq$ onto $[0, \sqrt{2}t]^2$ with $u^* = (0, t)$. Define also a two-parameter filtration $\boldsymbol{F}^\varepsilon$ on $\mathbb{R}^2$ with respect to the partial order $\preceq$ by setting

$$\mathcal{F}^\varepsilon_{(s,y)} = \bigcap_{(s,y)\preceq(\tilde{s},\tilde{y})} \sigma\left(L^\varepsilon(A) \mid A \in \mathcal{B}(A^+(\tilde{s},\tilde{y}))\right) \vee \mathcal{N}^\mathbb{P} \quad \text{for} \quad (s,y) \in \mathbb{R}^+ \times \mathbb{R}, \tag{2.5.5}$$

with $\mathcal{N}^\mathbb{P}$ the set of all $\mathbb{P}$ null-sets of $\mathcal{F}$ (and $\mathcal{F}^\varepsilon_{(s,y)} = \{\emptyset, \Omega\}$ for all (s,y) with $s < 0$). We further define $\widetilde{u}^\varepsilon(v_1,v_2) = u^\varepsilon(K^{-1}(v_1,v_2))$ for all $v = (v_1,v_2) \in [0,\sqrt{2}t]^2$ (extending u^ε to 0 whenever $v_1 + v_2 < \sqrt{2}t$) as well as a filtration $\widetilde{\boldsymbol{F}}^\varepsilon$ on $[0,\sqrt{2}t]^2$ with respect to $\leq$ by $\widetilde{\mathcal{F}}^\varepsilon_{(v_1,v_2)} = \mathcal{F}^\varepsilon_{K^{-1}(v_1,v_2)}$. With the stochastic integration theory of Cairoli and Walsh in [23], we now show that $\widetilde{u}^\varepsilon$ is a two-parameter *strong martingale* with respect to $\widetilde{\boldsymbol{F}}^\varepsilon$, see e.g. page 115 there for a definition.

Consider on $[0,\sqrt{2}t]^2$ the two-parameter process

$$\widetilde{L}^\varepsilon(v_1,v_2) = \begin{cases} L^\varepsilon\left(A^+(K^{-1}(v_1,v_2))\right), & \text{if} \quad v_1 + v_2 \geq \sqrt{2}t, \\ 0, & \text{otherwise.} \end{cases}$$

By the properties of the Lévy noise L^ε, $\widetilde{L}^\varepsilon$ is a Lévy sheet as well as an $\widetilde{\boldsymbol{F}}^\varepsilon$-strong martingale (the latter follows exactly as in the proof of Lemma 6.2 in [84]) and $\widetilde{\boldsymbol{F}}^\varepsilon$ satisfies the commuting condition $F4$ of [23], see pp. 113–114. Choose the filtration $\boldsymbol{F}$ on $[0,t]$ to be $\mathcal{F}_r = \bigcap_{t\geq s\geq r} \sigma\left(L^\varepsilon(A) \mid A \in \mathcal{B}(A^+(s,0))\right) \vee \mathcal{N}^\mathbb{P}$ for all $0 \leq r \leq t$ (note that on $A^+(s,0)$ the mild solution u^ε depends on the values of L^ε on $A^+(s,0)$ only). Then $\widetilde{u}^\varepsilon$ is a valid integrand (see also page 121 of [23]) and $\widetilde{L}^\varepsilon$ a valid integrator for Theorem 2.2 in [23] to apply, whence

$$\int_0^{v_1}\int_0^{v_2} f(\widetilde{u}^\varepsilon(z_1,z_2))\,\widetilde{L}^\varepsilon(\mathrm{d}z_1,\mathrm{d}z_2) = \int_{\mathbb{R}^+\times\mathbb{R}} \mathbb{1}_{A^+(K^{-1}(v_1,v_2))} f(u^\varepsilon(s,y))\,L^\varepsilon(\mathrm{d}s,\mathrm{d}y) = \widetilde{u}^\varepsilon(v_1,v_2)$$

is an $\widetilde{\boldsymbol{F}}^\varepsilon$-strong martingale on $[0,\sqrt{2}t]^2$. Analogously, $\widetilde{u}^{\varepsilon,n} = u^{\varepsilon,n} \circ K^{-1}$ defines an $\widetilde{\boldsymbol{F}}^\varepsilon$-strong martingale on $[0,\sqrt{2}t]^2$ for each $n \in \mathbb{N}$. As a consequence, apply Cairoli's strong maximal inequality, see e.g. Corollary 2.3.1 of Chapter 7 in [72] (note that u^ε and $u^{\varepsilon,n}$ are also *orthomartingales* by Proposition 1.1 in [97] and L^2-continuous by Theorem 4.7 in [28]) to obtain

$$\begin{aligned} \mathbb{E}\left[\sup_{(v_1,v_2)\in K([0,(t,0)]_\preceq)} \left|\widetilde{u}^\varepsilon(v_1,v_2) - \widetilde{u}^{\varepsilon,n}(v_1,v_2)\right|^2\right] &\leq \mathbb{E}\left[\sup_{(v_1,v_2)\in[0,\sqrt{2}t]^2} \left|\widetilde{u}^\varepsilon(v_1,v_2) - \widetilde{u}^{\varepsilon,n}(v_1,v_2)\right|^2\right] \\ \leq \sup_{(v_1,v_2)\in[0,\sqrt{2}t]^2} \mathbb{E}\left[\left|\widetilde{u}^\varepsilon(v_1,v_2) - \widetilde{u}^{\varepsilon,n}(v_1,v_2)\right|^2\right] &= \mathbb{E}\left[\left|\widetilde{u}^\varepsilon(\sqrt{2}t,\sqrt{2}t) - \widetilde{u}^{\varepsilon,n}(\sqrt{2}t,\sqrt{2}t)\right|^2\right]. \end{aligned}$$

By bijectivity, the terms in these inequalities agree exactly with the corresponding ones in (2.5.4).

In a second step, we show that

$$u^{\varepsilon,n}(t,x) \longrightarrow u^\varepsilon(t,x) \quad \text{in} \quad L^2(\Omega,\mathcal{F},\mathbb{P}) \quad \text{as} \quad n \to \infty \quad \text{for all} \quad (t,x) \in \mathbb{R}^+ \times \mathbb{R}. \tag{2.5.6}$$

Write

$$\begin{aligned} u^\varepsilon(t,x) - u^{\varepsilon,n}(t,x) &= \int_0^t \int_\mathbb{R} G_{t-s}(x,y) \frac{f(u^\varepsilon(s,y)) - f(u^{\varepsilon,n}(s,y))}{\sigma(\varepsilon)}\,L^\varepsilon(\mathrm{d}s,\mathrm{d}y) \\ &\quad + \int_0^t \int_\mathbb{R} G_{t-s}(x,y) \frac{f(u^{\varepsilon,n}(s,y))}{\sigma(\varepsilon)} (L^\varepsilon - L^{\varepsilon,n})(\mathrm{d}s,\mathrm{d}y) =: I_{\varepsilon,n}(t,x) + J_{\varepsilon,n}(t,x). \end{aligned}$$

Fix $T > 0$. Using Itō's isometry and the Lipschitz continuity of f, we estimate

$$\mathbb{E}\left[I_{\varepsilon,n}(t,x)^2\right] \leq C \int_0^t \int_{\mathbb{R}} \mathbb{1}_{A^+(t,x)}(s,y)\mathbb{E}\left[\left|u^\varepsilon(s,y) - u^{\varepsilon,n}(s,y)\right|^2\right] \mathrm{d}s\,\mathrm{d}y \tag{2.5.7}$$

as well as

$$\mathbb{E}\left[J_{\varepsilon,n}(t,x)^2\right] \leq C\left(1 + \sup_{(s,y)\in[0,T]\times\mathbb{R}} \sup_{\varepsilon>0,n\in\mathbb{N}} \mathbb{E}\left[|u^{\varepsilon,n}(s,y)|^2\right]\right) \\ \times \sigma^{-2}(\varepsilon) \int_0^t \int_{\mathbb{R}^2} \mathbb{1}_{A^+(t,x)}(s,y) z^2 \left(1 - \mathbb{1}_{\{|y|\leq n, |z|>1/n\}}\right)^2 \mathrm{d}s\,\mathrm{d}y\, Q^\varepsilon(\mathrm{d}z) \tag{2.5.8}$$

for all $(t,x) \in [0,T]\times\mathbb{R}$ and $n \in \mathbb{N}$. Since the uniform bound (2.4.11) also applies to all $u^{\varepsilon,n}$ and noting that $1 - \mathbb{1}_{\{|y|\leq n,|z|>1/n\}} = \mathbb{1}_{\{|y|>n\}} + \mathbb{1}_{\{|y|\leq n,|z|\leq 1/n\}}$ pointwise on $\mathbb{R}^2$, the right-hand side of (2.5.8) can further be estimated by C times the function

$$f_{\varepsilon,n}(t,x) = \int_0^t \int_{\mathbb{R}} \mathbb{1}_{A^+(t,x)}(s,y)\mathbb{1}_{\{|y|>n\}}\,\mathrm{d}s\,\mathrm{d}y \\ + \sigma^{-2}(\varepsilon)\int_{\mathbb{R}} z^2 \mathbb{1}_{\{|z|\leq 1/n\}}\, Q^\varepsilon(\mathrm{d}z) \int_0^t \int_{\mathbb{R}} \mathbb{1}_{A^+(t,x)}(s,y)\,\mathrm{d}s\,\mathrm{d}y, \quad (t,x)\in[0,T]\times\mathbb{R},$$

which together with (2.5.7) yields:

$$\mathbb{E}\left[\left|u^\varepsilon(t,x) - u^{\varepsilon,n}(t,x)\right|^2\right] \leq C\int_0^t \int_{\mathbb{R}} \mathbb{1}_{A^+(t,x)}(s,y)\mathbb{E}\left[\left|u^\varepsilon(s,y) - u^{\varepsilon,n}(s,y)\right|^2\right] \mathrm{d}s\,\mathrm{d}y + C f_{\varepsilon,n}(t,x) \tag{2.5.9}$$

for all $(t,x)\in[0,T]\times\mathbb{R}$ and $n\in\mathbb{N}$.

Set $v_{\varepsilon,n}(t,x) = \mathbb{E}\left[|u^\varepsilon(t,x) - u^{\varepsilon,n}(t,x)|^2\right]$ and hold from now on C in (2.5.9) fixed. Let $t_1 > 0$ such that $t_1^2 < 2/C$ and set $t_k = kt_1$ with $k\in\mathbb{N}$. We now show by induction that for all $k\in\mathbb{N}$, $v_{\varepsilon,n}(t,x) \longrightarrow 0$ as $n\to\infty$ for any $(t,x)\in[0,t_k\wedge T]\times\mathbb{R}$, which altogether implies (2.5.6). First, (2.2.5), (2.5.9) and dominated convergence yield for $t\leq t_1$,

$$\sup_{(s,y)\preceq(t,x)} v_{\varepsilon,n}(s,y) \leq \frac{C}{1 - Ct_1^2/2} f_{\varepsilon,n}(t,x) \longrightarrow 0 \quad \text{as} \quad n\to\infty. \tag{2.5.10}$$

Next, let $k\geq 2$ and assume $t_k < t \leq t_{k+1} \leq T$. We have

$$\int_0^t \int_{\mathbb{R}} \mathbb{1}_{A^+(t,x)}(s,y) v_{\varepsilon,n}(s,y)\,\mathrm{d}s\,\mathrm{d}y \leq \int_0^{t_k}\int_{\mathbb{R}} \mathbb{1}_{A^+(t,x)}(s,y)v_{\varepsilon,n}(s,y)\,\mathrm{d}s\,\mathrm{d}y \\ + \sup_{\substack{(\tilde{t},\tilde{x})\preceq(t,x)\\ t_k<\tilde{t}}} v_{\varepsilon,n}(\tilde{t},\tilde{x}) \int_{t_k}^t \int_{\mathbb{R}} \mathbb{1}_{A^+(t,x)}(s,y)\,\mathrm{d}s\,\mathrm{d}y.$$

Combine this inequality with (2.5.9), note that $\int_{t_k}^t \int_{\mathbb{R}} \mathbb{1}_{A^+(t,x)}(s,y)\,\mathrm{d}s\,\mathrm{d}y = (t-t_k)^2 \leq t_1^2$ and use similar calculations as for (2.5.10) to conclude that

$$\sup_{\substack{(\tilde{t},\tilde{x})\preceq(t,x)\\ t_k<\tilde{t}}} v_{\varepsilon,n}(\tilde{t},\tilde{x}) \leq \frac{C}{1-Ct_1^2/2}\left(\int_0^{t_k}\int_{\mathbb{R}} \mathbb{1}_{A^+(t,x)}(s,y)v_{\varepsilon,n}(s,y)\,\mathrm{d}s\,\mathrm{d}y + f_{\varepsilon,n}(t,x)\right) \longrightarrow 0$$

as $n\to\infty$ by induction hypothesis and dominated convergence.

We infer, using (2.5.4) and (2.5.6), that $u^{\varepsilon,n} - u^\varepsilon$ converges uniformly on compacts in probability to 0 as $n\to\infty$ for any $\varepsilon > 0$ and therefore, by standard arguments, the existence of a $\preceq$-càdlàg version $\overline{u}^\varepsilon$ of u^ε on $\mathbb{R}^+\times\mathbb{R}$. □

We now turn to the Skorokhod topology for $\preceq$-càdlàg functions.

Proof of Lemma 2.3.3. We first recall a few facts on the usual Skorokhod topology on $D([0,1]^2)$ that can all be found in Section 5 of [92]. It is induced by the Skorokhod metric

$$\delta'(x,y) = \inf_{\lambda\in\Lambda_s\times\Lambda_s}\left(\sup_{v\in[0,1]^2}\Big|x(v)-y(\lambda(v))\Big| \vee \|\lambda\|_s\right), \quad x,y\in D([0,1]^2), \tag{2.5.11}$$

where Λ_s is the set of all homeomorphisms of $[0,1]$ onto itself which have 0 as a fixed point, $\Lambda_s\times\Lambda_s$ the set of all homeomorphisms λ of the form

$$\lambda\colon [0,1]^2 \longrightarrow [0,1]^2, \quad v=(v_1,v_2)\mapsto(\lambda_1(v_1),\lambda_2(v_2))$$

with $\lambda_1,\lambda_2\in\Lambda_s$, and $\|\lambda\|_s = \sup_{0\le p\le 1}(\max_{i=1,2}|\lambda_i(p)-p|)$ for $\lambda\in\Lambda_s\times\Lambda_s$. There exists a Skorokhod metric δ that is equivalent to δ' and makes $D([0,1]^2)$ a complete and separable metric space.

Now recall (2.3.3), (2.3.5) and give $D_{\preceq}([0,1]^2)$ the topology induced by the metric

$$\tau'(x,y) = \delta'(\Phi^{-1}(x),\Phi^{-1}(y)), \quad x,y\in D_{\preceq}([u_0,u^*]_{\preceq}).$$

This is a Skorokhod distance in the sense of [92], see (3.14) of Section 3. Indeed, consider the group of homeomorphisms from $[u_0,u^*]_{\preceq}$ onto itself $\Theta_s := \{J^{-1}\circ\lambda\circ J \mid \lambda\in\Lambda_s\times\Lambda_s\}$ equipped with the induced norm $\|J^{-1}\circ\lambda\circ J\|_s := \|\lambda\|_s$, see Section 3 in [92]. Then we can rewrite

$$\tau'(x,y) = \inf_{\theta\in\Theta_s}\left(\sup_{u\in[u_0,u^*]_{\preceq}}\Big|x(u)-y(\theta(u))\Big| \vee \|\theta\|_s\right), \quad x,y\in D_{\preceq}([u_0,u^*]_{\preceq}).$$

Defining τ on $D_{\preceq}([u_0,u^*]_{\preceq})$ analogously to τ', but with δ instead of δ', yields an equivalent Skorokhod metric to τ' that makes $D_{\preceq}([u_0,u^*]_{\preceq})$ a complete and separable metric space.

Definition 2.3.1 of $D_{\preceq}([u_0,u^*]_{\preceq})$ coincides exactly with the construction (3.15) in Section 3 of [92] of the Skorokhod space on the set $[u_0,u^*]_{\preceq}$ relative to the group Θ_s (in order to see this, consider all preimages under J of the partitions used in (5.5) and (5.6) of that paper to construct $D([0,1]^2)$, define the Skorokhod space and use Theorem 5.1 in [92]).

At last, use the exact same procedure to obtain a Skorokhod topology on $D_{\preceq}([0,1]^2)$ as well as Skorokhod metrics, denoted by the same letters as before. The map Φ of (2.3.5) is now a homeomorphic transformation between $D(J([0,1]^2))$ and $D_{\preceq}([0,1]^2)$. For the definition of the Skorokhod metric on $D(J([0,1]^2))$, we now use the subgroup $\Gamma_s = \{\lambda\in\Lambda_s\times\Lambda_s \mid \lambda(J([0,1]^2)) = J([0,1]^2)\}$ equipped with the norm $\|\lambda\|_{s'} = \sup_{(v_1,v_2)\in J([0,1]^2)}(\max_{i=1,2}|\lambda_i(v_i)-v_i|)$. As a consequence, it is easy to see that the restriction map $\iota : D_{\preceq}([u_0,u^*]_{\preceq}) \hookrightarrow D_{\preceq}([0,1]^2)$ is continuous.

The remaining assertions readily follow from Section 3 and 5 of [92]. □

Next, we proceed to show that the $\preceq$-càdlàg version $\overline{u}^\varepsilon$ is tight.

Theorem 2.5.1. *The random fields $\{\overline{u}^\varepsilon \mid \varepsilon>0\}$ where $\overline{u}^\varepsilon$ is the $\preceq$-càdlàg version of u^ε obtained in Theorem 2.3.2, are* tight *in the Skorokhod space $D_{\preceq}([0,T]\times I)$ for any $T>0$ and finite closed interval $I\subseteq\mathbb{R}$.*

Proof. Without loss of generality, assume that $T=1$ and $I=[0,1]$ and recall the transformation J in (2.3.3). Set $u^\varepsilon(t,x)=\overline{u}^\varepsilon(t,x)=0$ whenever $t<0$. By Lemma 2.3.3, it suffices to show that the random elements $\{\overline{u}^\varepsilon\circ J^{-1} \mid \varepsilon>0\}$ are tight in $D([0,1]^2)$.

By (2.4.11), all random variables $\overline{u}^\varepsilon \circ J^{-1}(v_1, v_2)$ are tight. Furthermore, by the same arguments as in the proof of Theorem 2.3.2, the processes $u^\varepsilon \circ J^{-1}$ and $\overline{u}^\varepsilon \circ J^{-1}$ are strong martingales in $[0,1]^2$ with respect to the push-forward of filtration (2.5.5) through J for each $\varepsilon > 0$.

We will apply a generalization of Aldous condition for tightness to strong martingales. First of all, fix $\varepsilon > 0$ and note that if τ is a natural 1-stopping time for $\overline{u}^\varepsilon \circ J^{-1}$ with $\tau \in [0,1]$, see page 112 in [59] for a definition, then the processes $(\overline{u}^\varepsilon \circ J^{-1}(\tau, v))_{0 \le v \le 1}$ and $(u^\varepsilon \circ J^{-1}(\tau, v))_{0 \le v \le 1}$ are versions of one another. To see this, approximate τ from above by a sequence $(\tau_n)_{n \in \mathbb{N}}$ of natural 1-stopping times taking on finitely many values only. Then for all $0 \le v \le 1$, $\overline{u}^\varepsilon \circ J^{-1}(\tau_n, v) = u^\varepsilon \circ J^{-1}(\tau_n, v)$ $\mathbb{P}$-almost surely. Now since $\overline{u}^\varepsilon \circ J^{-1}$ is càdlàg and $(\tau, v) \le (\tau_n, v)$, by dominated convergence, $\overline{u}^\varepsilon \circ J^{-1}(\tau_n, v) \longrightarrow \overline{u}^\varepsilon \circ J^{-1}(\tau, v)$ in $L^1(\Omega, \mathcal{F}, \mathbb{P})$ as $n \to \infty$ for any $0 \le v \le 1$. Finally, by Itō's isometry, the Lipschitz continuity of f and with $\mathbb{1}_\emptyset \equiv 0$,

$$\begin{aligned} &\mathbb{E}\left[\left(u^\varepsilon \circ J^{-1}(\tau_n, v) - u^\varepsilon \circ J^{-1}(\tau, v)\right)^2\right] \\ &\quad \le C\mathbb{E}\left[\int_{\mathbb{R}^+ \times \mathbb{R}} \mathbb{1}_{A^+(J^{-1}(\tau_n,v)) \setminus A^+(J^{-1}(\tau,v))}(s,y)\left(\left|u^\varepsilon(s,y)\right|^2 + 1\right) \mathrm{d}s\,\mathrm{d}y\right] \end{aligned} \tag{2.5.12}$$

for all $0 \le v \le 1$ and $n \in \mathbb{N}$. We infer $u^\varepsilon \circ J^{-1}(\tau_n, v) \longrightarrow u^\varepsilon \circ J^{-1}(\tau, v)$ in $L^2(\Omega, \mathcal{F}, \mathbb{P})$ as $n \to \infty$ for any $0 \le v \le 1$, again by dominated convergence ($\mathbb{1}_{A^+(J^{-1}(\tau_n,v)) \setminus A^+(J^{-1}(\tau,v))} \longrightarrow 0$ pointwise on $\Omega \times \mathbb{R}^+ \times \mathbb{R}$ as $n \to \infty$ and the integrand above may be approximated by the integrable function $\mathbb{1}_{A^+(u^*)}(|u^\varepsilon|^2 + 1)$ with $u^* = (3/2, 1/2)$).

Now we assume that each u^ε is separable and let $(\varepsilon_n)_{n \in \mathbb{N}}$, $(h_n)_{n \in \mathbb{N}}$ be sequences of positive numbers with $\varepsilon_n \longrightarrow 0$ and $h_n \longrightarrow 0$ as $n \to \infty$. Let also $(T_n)_{n \in \mathbb{N}}$ be a sequence of natural 1-stopping times for $\overline{u}^{\varepsilon_n} \circ J^{-1}$ with $T_n \in [0,1]$. As for (2.5.12) and using (2.4.11), we obtain

$$\begin{aligned} &\mathbb{E}\left[\left|\overline{u}^{\varepsilon_n} \circ J^{-1}(T_n + h_n, v) - \overline{u}^{\varepsilon_n} \circ J^{-1}(T_n, v)\right|^2\right] = \mathbb{E}\left[\left|u^{\varepsilon_n} \circ J^{-1}(T_n + h_n, v) - u^{\varepsilon_n} \circ J^{-1}(T_n, v)\right|^2\right] \\ &\quad \le C\mathbb{E}\left[\left(\sup_{(s,y) \in A^+(u^*)} \left|u^{\varepsilon_n}(s,y)\right|^2 + 1\right) \int_{\mathbb{R}^+ \times \mathbb{R}} \mathbb{1}_{A^+(J^{-1}(T_n + h_n, v)) \setminus A^+(J^{-1}(T_n, v))}(s,y)\,\mathrm{d}s\,\mathrm{d}y\right] \\ &\quad \le Ch_n\left(\mathbb{E}\left[\sup_{(s,y) \in A^+(u^*)} \left|u^{\varepsilon_n}(s,y)\right|^2\right] + 1\right) \le Ch_n\left(\sup_{(s,y) \in A^+(u^*)} \mathbb{E}\left[\left|u^{\varepsilon_n}(s,y)\right|^2\right] + 1\right) \le Ch_n, \end{aligned} \tag{2.5.13}$$

which goes to 0 as $n \to \infty$ for all $0 \le v \le 1$. We used the inverse mapping of H to see that whenever $T_n + v \ge 1$, the surface integral inside the third expectation in (2.5.13) equals

$$\frac{9}{2}\left(T_n + h_n + v - \frac{1}{\sqrt{2}}\right)^2 - \frac{9}{2}\left(T_n + v - \frac{1}{\sqrt{2}}\right)^2 = 9h_n\left(T_n + \frac{h_n}{2} + v - \frac{1}{\sqrt{2}}\right) (\le Ch_n).$$

Plus, the maximal inequality on the last line of (2.5.13) is a consequence of (2.5.4). Analogously, if $(T_n)_{n \in \mathbb{N}}$ is a sequence of natural 2-stopping times for $\overline{u}^{\varepsilon_n} \circ J^{-1}$ with $T_n \in [0,1]$,

$$\sup_{0 \le v \le 1} \mathbb{E}\left[\left|\overline{u}^{\varepsilon_n} \circ J^{-1}(v, T_n + h_n) - \overline{u}^{\varepsilon_n} \circ J^{-1}(v, T_n)\right|^2\right] \le Ch_n \longrightarrow 0 \quad \text{as} \quad n \to \infty.$$

Therefore, the random fields $\overline{u}^\varepsilon \circ J^{-1}$ satisfy all conditions for Theorem 4.I in [59] to apply. □

We now state a tightness result for u^ε and $\overline{u}^\varepsilon$ in L^2-space.

Theorem 2.5.2. *The family* $\{u^\varepsilon \mid \varepsilon > 0\}$ *of mild solutions to* (2.2.1) *is tight in the Hilbert space* $L^2([0,T] \times I)$ *for any* $T > 0$ *and finite interval* $I \subseteq \mathbb{R}$.

Proof. Let $\{\Psi_k \mid k \in \mathbb{N}\}$ be a countable orthonormal basis of $L^2([0,T] \times I)$. By the stochastic Fubini theorem (see e.g. Theorem 2.6 in [99]), for all $\varepsilon > 0$ and $k \in \mathbb{N}$,

$$\begin{aligned}\langle u^\varepsilon, \Psi_k\rangle &= \int_0^T \int_I \left(\int_0^T \int_{\mathbb{R}} G(t,x;s,y)\frac{f(u^\varepsilon(s,y))}{\sigma(\varepsilon)}\, L^\varepsilon(\mathrm{d}s,\mathrm{d}y)\right) \Psi_k(t,x)\,\mathrm{d}t\,\mathrm{d}x \\ &= \int_0^T \int_{\mathbb{R}} \left(\int_0^T \int_I G(t,x;s,y)\Psi_k(t,x)\,\mathrm{d}t\,\mathrm{d}x\right) \frac{f(u^\varepsilon(s,y))}{\sigma(\varepsilon)}\, L^\varepsilon(\mathrm{d}s,\mathrm{d}y) \quad \mathbb{P}\text{-almost surely.}\end{aligned}$$

Using (2.4.11), Parseval's identity and Fubini's theorem, we infer

$$\begin{aligned}\sum_{k=0}^\infty \sup_{\varepsilon>0} \mathbb{E}\left[\langle u^\varepsilon, \Psi_k\rangle^2\right] &\le C \int_0^T \int_{\mathbb{R}} \left(\sum_{k=0}^\infty \langle G(\cdot,\cdot;s,y),\Psi_k\rangle^2\right) \mathrm{d}s\,\mathrm{d}y \\ &= \int_0^T \int_{\mathbb{R}} \left(\int_0^T \int_I G^2(t,x;s,y)\,\mathrm{d}t\,\mathrm{d}x\right)\mathrm{d}s\,\mathrm{d}y = \int_0^T \int_I \left(\int_0^T \int_{\mathbb{R}} G^2(t,x;s,y)\,\mathrm{d}s\,\mathrm{d}y\right)\mathrm{d}t\,\mathrm{d}x,\end{aligned}$$

which is finite since the last inner integral equals $t^2/4$. This implies by Markov's inequality,

$$\sup_{\varepsilon>0} \mathbb{P}\left(\sum_{k=N}^\infty \langle u^\varepsilon, \Psi_k\rangle^2 > \delta\right) \le \frac{1}{\delta}\sum_{k=N}^\infty \sup_{\varepsilon>0}\mathbb{E}\left[\langle u^\varepsilon, \Psi_k\rangle^2\right] \longrightarrow 0 \quad \text{as} \quad N\to\infty$$

for all $\delta > 0$ as well as

$$\sup_{\varepsilon>0} \mathbb{P}\left(\sum_{k=0}^N \langle u^\varepsilon, \Psi_k\rangle^2 > \delta\right) \le \frac{1}{\delta}\sum_{k=0}^\infty \sup_{\varepsilon>0}\mathbb{E}\left[\langle u^\varepsilon, \Psi_k\rangle^2\right] \longrightarrow 0 \quad \text{as} \quad \delta\to\infty$$

for all $N \in \mathbb{N}$. So we can apply Theorem 1 in [93] and conclude the proof. □

We turn to the $H_{-r}(\mathbb{R})$-valued process v^ε in (2.3.11) and first show that it has a càdlàg version.

Proof of Theorem 2.3.5. The proof relies on the Hilbert space structure of $H_{-r}(\mathbb{R})$. First, we show that for any $\phi \in \mathcal{S}(\mathbb{R})$ and $\varepsilon > 0$, the real-valued process $(\langle v_t^\varepsilon, \phi\rangle)_{t\ge 0}$ has a càdlàg modification. Use (2.3.9), (2.3.11) and the fundamental theorem of calculus to rewrite for all $t \ge 0$,

$$\begin{aligned}\langle v_t^\varepsilon, \phi\rangle &= \int_0^t \int_{\mathbb{R}} \phi(y)\frac{f(u^\varepsilon(s,y))}{\sigma(\varepsilon)}\, L^\varepsilon(\mathrm{d}s,\mathrm{d}y) \\ &\quad + \frac{1}{2}\int_0^t \int_{\mathbb{R}} \left(\int_s^t \phi'(y+(r-s))\,\mathrm{d}r - \int_s^t \phi'(y-(r-s))\,\mathrm{d}r\right) \frac{f(u^\varepsilon(s,y))}{\sigma(\varepsilon)}\, L^\varepsilon(\mathrm{d}s,\mathrm{d}y)\end{aligned}$$

$\mathbb{P}$-almost surely, and then the stochastic Fubini theorem on the last double integral to obtain

$$\langle v_t^\varepsilon, \phi\rangle = \int_0^t \int_{\mathbb{R}} \phi(y)\frac{f(u^\varepsilon(s,y))}{\sigma(\varepsilon)}\, L^\varepsilon(\mathrm{d}s,\mathrm{d}y) + \frac{1}{2}\int_0^t J_r^\varepsilon(\phi)\,\mathrm{d}r \quad \mathbb{P}\text{-almost surely} \tag{2.5.14}$$

where we have set

$$J_r^\varepsilon(\phi) = \int_0^r \int_{\mathbb{R}} \left(\phi'(y+(r-s)) - \phi'(y-(r-s))\right) \frac{f(u^\varepsilon(s,y))}{\sigma(\varepsilon)}\, L^\varepsilon(\mathrm{d}s,\mathrm{d}y), \quad r \ge 0.$$

The semimartingale on the right-hand side of (2.5.14), that we will denote by $X^\varepsilon(\phi)$, is càdlàg.

Next, fix an arbitrary $T > 0$ and $\varepsilon > 0$. Doob's inequality, Itō's isometry and (2.2.7) yield

$$\mathbb{E}\left[\sup_{t\le T}\left|X_t^\varepsilon(\phi)\right|^2\right] \le C\int_0^T\int_{\mathbb{R}}\phi^2(y)\,\mathrm{d}s\,\mathrm{d}y + C\int_0^T\left(\int_0^r\int_{\mathbb{R}}\phi'(y\pm(r-s))^2\,\mathrm{d}s\,\mathrm{d}y\right)\mathrm{d}r \tag{2.5.15}$$

for all $\phi \in \mathcal{S}(\mathbb{R})$. Now the Hermite functions h_q satisfy the recursion relation $h_q'(x) = \sqrt{q/2}h_{q-1}(x) - \sqrt{(q+1)/2}h_{q+1}(x)$ for all $q \in \mathbb{N}$ and $x \in \mathbb{R}$, from which we obtain by orthogonality,

$$\int_{\mathbb{R}} h_q'(x)^2\,\mathrm{d}x = \frac{q}{2}\int_{\mathbb{R}} h_{q-1}^2(x)\mathrm{d}x + \frac{q+1}{2}\int_{\mathbb{R}} h_{q+1}^2(x)\mathrm{d}x = q + \frac{1}{2}. \tag{2.5.16}$$

We carry forward the estimation in (2.5.15) for $\phi = h_q$, whence

$$\mathbb{E}\left[\sup_{t\le T}\left|X_t^\varepsilon(h_q)\right|^2\right] \le C\left(1 + \int_{\mathbb{R}} h_q'(x)^2\,\mathrm{d}x\right) \le C(1+2q) \quad \text{for all} \quad q \in \mathbb{N}. \tag{2.5.17}$$

Fix $r > 2$. It is easy to see that for each $N \in \mathbb{N}$, the $H_{-r}(\mathbb{R})$-valued process

$$\sum_{q=0}^{N}(1+2q)^{-r/2}X^\varepsilon(h_q)e_{q,-r} \tag{2.5.18}$$

with $e_{q,-r}$ as in (2.3.16), is càdlàg. Recall the Fourier expansion (2.3.17) and use (2.5.17) to obtain

$$\begin{aligned}\mathbb{E}\left[\sup_{t\le T}\left\|\sum_{q=N+1}^{M}(1+2q)^{-r/2}X_t^\varepsilon(h_q)e_{q,-r}\right\|_{-r}^2\right] &\le \sum_{q=N+1}^{M}(1+2q)^{-r}\mathbb{E}\left[\sup_{t\le T}\left|X_t^\varepsilon(h_q)\right|^2\right] \\ &\le C\sum_{q=N+1}^{M}(1+2q)^{-r+1} \longrightarrow 0\end{aligned} \tag{2.5.19}$$

as $N, M \to \infty$ since $r > 2$. Consequently, standard arguments show that there exists a process $\overline{v}^\varepsilon \in D([0,T], H_{-r}(\mathbb{R}))$ such that $\mathbb{P}$-almost surely,

$$\overline{v}_t^\varepsilon = \sum_{q=0}^{\infty}(1+2q)^{-r/2}X_t^\varepsilon(h_q)e_{q,-r} \quad \text{in} \quad H_{-r}(\mathbb{R}) \quad \text{for all} \quad t \le T. \tag{2.5.20}$$

By (2.5.14), this process is a version of $(v_t^\varepsilon)_{t\le T}$ in $H_{-r}(\mathbb{R})$.

□

Next, we show tightness of the càdlàg version $\overline{v}^\varepsilon$.

Theorem 2.5.3. *The family of processes $\{\overline{v}^\varepsilon \mid \varepsilon > 0\}$ where $\overline{v}^\varepsilon$ is the càdlàg version of v^ε obtained in Theorem 2.3.5, is tight in the Skorokhod space $D([0,T], H_{-r}(\mathbb{R}))$ for any $r > 2$ and $T > 0$.*

Proof. We first check that $\{\overline{v}^\varepsilon \mid \varepsilon > 0\}$ satisfies the Aldous condition for tightness. To this end, let $(\varepsilon_n)_{n\in\mathbb{N}}$ and $(h_n)_{n\in\mathbb{N}}$ be sequences of positive numbers with $\varepsilon_n \to 0$ and $h_n \to 0$ as $n \to \infty$. In addition, for each $n \in \mathbb{N}$, let $\tau_n \in [0,T]$ be a stopping time with respect to the filtration generated by the process $(\overline{v}_t^{\varepsilon_n})_{t\le T}$. We will show

$$\mathbb{E}\left[\|\overline{v}_{\tau_n+h_n}^{\varepsilon_n} - \overline{v}_{\tau_n}^{\varepsilon_n}\|_{-r}^2\right] \longrightarrow 0 \quad \text{as} \quad n \to \infty. \tag{2.5.21}$$

Recall the series representation (2.5.20) of $\overline{v}^{\varepsilon_n}$, where $X^{\varepsilon_n}(h_q)$ is the right-hand side of (2.5.14) with $\phi = h_q$. We have for each $q \in \mathbb{N}$,

$$\begin{aligned} X^{\varepsilon_n}_{\tau_n+h_n}(h_q) - X^{\varepsilon_n}_{\tau_n}(h_q) &= \int_{\tau_n}^{\tau_n+h_n} \int_{\mathbb{R}} h_q(y) \frac{f(u^{\varepsilon_n}(s,y))}{\sigma(\varepsilon_n)}\, L^{\varepsilon_n}(\mathrm{d}s,\mathrm{d}y) + \frac{1}{2}\int_{\tau_n}^{\tau_n+h_n} J^{\varepsilon_n}_r(h_q)\,\mathrm{d}r \\ &=: I_{q,n} + J_{q,n}. \end{aligned} \tag{2.5.22}$$

We estimate the second moment of each of these two terms. For the first one, by Itō's isometry and the Lipschitz continuity of f,

$$\begin{aligned} \mathbb{E}\left[I^2_{q,n}\right] &= \mathbb{E}\left[\int_0^T \int_{\mathbb{R}} h_q^2(y) \mathbb{1}_{(\tau_n,\tau_n+h_n]}(s) f^2(u^{\varepsilon_n}(s,y))\,\mathrm{d}s\,\mathrm{d}y\right] \\ &\le C\mathbb{E}\left[\int_0^T \int_{\mathbb{R}} h_q^2(y) \mathbb{1}_{(\tau_n,\tau_n+h_n]}(s) |u^{\varepsilon_n}(s,y)|^2\,\mathrm{d}s\,\mathrm{d}y\right] + C\mathbb{E}\left[\int_0^T \mathbb{1}_{(\tau_n,\tau_n+h_n]}(s)\,\mathrm{d}s\right] \\ &= C\int_{\mathbb{R}} h_q^2(y)\mathbb{E}\left[\int_0^T \mathbb{1}_{(\tau_n,\tau_n+h_n]}(s)|u^{\varepsilon_n}(s,y)|^2\,\mathrm{d}s\right]\mathrm{d}y + Ch_n. \end{aligned} \tag{2.5.23}$$

Furthermore, by the maximal inequality (2.5.4) (assuming separability),

$$\mathbb{E}\left[\sup_{(s,y)\in[(0,x),(T,x)]_{\preceq}} \left|u^{\varepsilon}(s,y)\right|^2\right] \le \mathbb{E}\left[\left|u^{\varepsilon}(T,x)\right|^2\right] \quad \text{for all} \quad x\in\mathbb{R} \quad \text{and} \quad \varepsilon>0. \tag{2.5.24}$$

Hence, the remaining integral on the right-hand side of (2.5.23) can further be estimated by

$$\begin{aligned} &\int_{\mathbb{R}} h_q^2(y)\mathbb{E}\left[\sup_{(s,z)\in[(0,y),(T,y)]_{\preceq}} \left|u^{\varepsilon_n}(s,z)\right|^2 \int_0^T \mathbb{1}_{(\tau_n,\tau_n+h_n]}(s)\,\mathrm{d}s\right]\mathrm{d}y \\ &\quad = h_n\int_{\mathbb{R}} h_q^2(y)\mathbb{E}\left[\sup_{(s,z)\in[(0,y),(T,y)]_{\preceq}} \left|u^{\varepsilon_n}(s,z)\right|^2\right]\mathrm{d}y \le h_n \int_{\mathbb{R}} h_q^2(y)\mathbb{E}\left[\left|u^{\varepsilon_n}(T,y)\right|^2\right]\mathrm{d}y \\ &\quad \le Ch_n\int_{\mathbb{R}} h_q^2(y)\mathrm{d}y = Ch_n \quad \text{for all} \quad q,n\in\mathbb{N}, \end{aligned} \tag{2.5.25}$$

where (2.4.11) was used for the last inequality. Note a significant difference here with the stochastic heat equation addressed in [32]: The mild solution to that equation is not a multiparameter martingale, so instead of maximal inequalities as (2.5.24), the factorization method from [39, 90] was used to prove the Aldous condition, see in particular Lemma 3.3 and (3.13) in [32].

Next, by the same calculations as in (2.5.15) (but using (2.4.11) instead of (2.2.7)) and (2.5.17), $\sup_{\varepsilon>0}\mathbb{E}\left[|J^{\varepsilon}_r(h_q)|^2\right] \le C(1+2q)$ for all $q\in\mathbb{N}$ and $r\le T$, so using Hölder's inequality,

$$\mathbb{E}\left[J^2_{q,n}\right] = \frac{1}{4}\mathbb{E}\left[\left(\int_0^T \mathbb{1}_{(\tau_n,\tau_n+h_n]} J^{\varepsilon_n}_r(h_q)\,\mathrm{d}r\right)^2\right] \le C\mathbb{E}\left[h_n\int_0^T |J^{\varepsilon_n}_r(h_q)|^2\,\mathrm{d}r\right] \le Ch_n(1+2q)$$

for all $n,q\in\mathbb{N}$. Combine this with (2.5.22), (2.5.23) and (2.5.25) to obtain altogether

$$\begin{aligned} \mathbb{E}\left[\|\overline{v}^{\varepsilon_n}_{\tau_n+h_n} - \overline{v}^{\varepsilon_n}_{\tau_n}\|^2_{-r}\right] &= \sum_{q=0}^{\infty}(1+2q)^{-r}\mathbb{E}\left[\left(X^{\varepsilon_n}_{\tau_n+h_n}(h_q) - X^{\varepsilon_n}_{\tau_n}(h_q)\right)^2\right] \\ &\le 2\sum_{q=0}^{\infty}(1+2q)^{-r}\left(\mathbb{E}\left[I^2_{q,n}\right] + \mathbb{E}\left[J^2_{q,n}\right]\right) \le h_n C\sum_{q=0}^{\infty}(1+2q)^{-r+1} \longrightarrow 0 \end{aligned}$$

as $n \to \infty$ since $r > 2$, which is (2.5.21).

In addition, $\sum_{q=0}^{\infty}(1+2q)^{-r}\sup_{\varepsilon>0}\mathbb{E}[X_t^\varepsilon(h_q)^2] < \infty$ by (2.4.11), (2.5.17) and since $r > 2$, so we can readily deduce, as in the proof of Theorem 2.5.2, that the random elements $\{\overline{v}_t^\varepsilon \mid \varepsilon > 0\}$ are tight in $H_{-r}(\mathbb{R})$ for any fixed $t \le T$.

The claim of the theorem now directly follows from Theorem 6.8 in [99]. □

We end this section with a tightness result for v^ε and $\overline{v}^\varepsilon$ in L^2-space.

Theorem 2.5.4. *The distribution-valued processes $\{v^\varepsilon \mid \varepsilon > 0\}$ with v^ε as in* (2.3.11)*, are tight in the Hilbert space $L^2([0,T], H_{-r}(\mathbb{R}))$ for each $r > 1$ and $T > 0$.*

Proof. First, each v^ε is an element of $L^2([0,T], H_{-r}(\mathbb{R}))$ as is seen from

$$\mathbb{E}\left[\int_0^T \|v_t^\varepsilon\|_{-r}^2 \,\mathrm{d}t\right] = \sum_{q=0}^{\infty}(1+2q)^{-r}\int_0^T \mathbb{E}\left[\langle v_t^\varepsilon, h_q\rangle^2\right]\mathrm{d}t \le C\sum_{q=0}^{\infty}(1+2q)^{-r} < \infty$$

which follows from (2.3.12) and $r > 1$.

The scalar product in $L^2([0,T], H_{-r}(\mathbb{R}))$ is given by $\langle f, g\rangle = \int_0^T \langle f_t, g_t\rangle_{-r}\,\mathrm{d}t$ and it is easy to see that an orthonormal basis is formed by $\{\phi_i e_{q,-r} \mid i, q \in \mathbb{N}\}$ with $\phi_i(t) = \sqrt{2/T}\sin(it\pi/T)$ and $e_{q,-r}$ as in (2.3.16). By (2.3.11) and the stochastic Fubini theorem, for all $i, q \in \mathbb{N}$ and $\varepsilon > 0$,

$$\int_0^T \langle v_t^\varepsilon, h_q\rangle\phi_i(t)\,\mathrm{d}t = \frac{1}{2}\int_0^T\int_{\mathbb{R}}\left(\int_s^T h_q(y \pm (t-s))\phi_i(t)\,\mathrm{d}t\right)\frac{f(u^\varepsilon(s,y))}{\sigma(\varepsilon)}\,L^\varepsilon(\mathrm{d}s,\mathrm{d}y) \quad \mathbb{P}\text{-a.s.}$$

Therefore, by duality, Itō's isometry and (2.4.11), we have

$$\mathbb{E}\left[\langle v^\varepsilon, \phi_i e_{q,-r}\rangle^2\right] \le C(1+2q)^{-r}\int_0^T\int_{\mathbb{R}}\left(\int_s^T h_q(y \pm (t-s))\phi_i(t)\,\mathrm{d}t\right)^2 \mathrm{d}s\,\mathrm{d}y.$$

Using Parseval's identity relative to the orthonormal basis of $L^2([0,T])$, we obtain altogether

$$\begin{aligned}\sum_{i,q=0}^{\infty}\sup_{\varepsilon>0}\mathbb{E}\left[\langle v^\varepsilon, \phi_i e_{q,-r}\rangle^2\right] &\le C\sum_{q=0}^{\infty}(1+2q)^{-r}\int_0^T\int_{\mathbb{R}}\sum_{i=0}^{\infty}\left(\int_s^T h_q(y \pm (t-s))\phi_i(t)\,\mathrm{d}t\right)^2\mathrm{d}s\,\mathrm{d}y \\ &= C\sum_{q=0}^{\infty}(1+2q)^{-r}\int_0^T\int_{\mathbb{R}}\int_s^T h_q^2(y \pm (t-s))\,\mathrm{d}t\,\mathrm{d}y\,\mathrm{d}s \le C\sum_{q=0}^{\infty}(1+2q)^{-r}\end{aligned} \tag{2.5.26}$$

which is finite since $r > 1$. We can now conclude analogously to the proof of Theorem 2.3.5. □

Corollary 2.5.5. *The distribution-valued processes $\{\overline{v}^\varepsilon \mid \varepsilon > 0\}$ where $\overline{v}^\varepsilon$ is the càdlàg version of v^ε in Theorem 2.3.5, are tight in $L^2([0,T], H_{-r}(\mathbb{R}))$ for any $r > 2$ and $T > 0$.*

Proof. Since $\overline{v}_t^\varepsilon$ lives in $H_{-r}(\mathbb{R})$ for $r > 2$ only, this is a direct consequence of Theorem 2.5.4. □

2.5.2 Proofs for Section 2.4

We first show that a mild solution to (2.2.1) satisfies equation (2.3.8) with $\partial_t u^\varepsilon$ replaced by v^ε.

Proposition 2.5.6. *Let u^ε be a mild solution to* (2.2.1) *and v^ε the process defined in* (2.3.11). *For each $\varepsilon > 0$, the pair $(u^\varepsilon, v^\varepsilon)$ satisfies the following weak formulation of the stochastic wave equation on $\mathbb{R}^+ \times \mathbb{R}$: For any $\phi_1, \phi_2 \in C_c^\infty(\mathbb{R})$ and $t \geq 0$, we have*

$$\begin{aligned}\int_{\mathbb{R}} u^\varepsilon(t,x)\phi_1(x)\,\mathrm{d}x + \langle v_t^\varepsilon, \phi_2\rangle \\ = \int_0^t \left(\int_{\mathbb{R}} u^\varepsilon(s,x)\phi_2''(x)\,\mathrm{d}x + \langle v_s^\varepsilon, \phi_1\rangle\right)\mathrm{d}s + \int_0^t\int_{\mathbb{R}} \phi_2(x)\frac{f(u^\varepsilon(s,x))}{\sigma(\varepsilon)}\,L^\varepsilon(\mathrm{d}s,\mathrm{d}x) \quad \mathbb{P}\text{-a.s.}\end{aligned} \tag{2.5.27}$$

Proof. Recall first (2.3.9) and apply the stochastic Fubini theorem in order to obtain

$$\begin{aligned}\int_0^t\int_{\mathbb{R}} u^\varepsilon(s,x)\phi_2''(x)\,\mathrm{d}x\,\mathrm{d}s &= \int_0^t\int_{\mathbb{R}}\left(\int_0^t\int_{\mathbb{R}} G(s,x;r,y)\phi_2''(x)\,\mathrm{d}s\,\mathrm{d}x\right)\frac{f(u^\varepsilon(r,y))}{\sigma(\varepsilon)}\,L^\varepsilon(\mathrm{d}r,\mathrm{d}y) \quad \text{and} \\ \int_0^t \langle v_s^\varepsilon, \phi_1\rangle\,\mathrm{d}s &= \int_0^t\int_{\mathbb{R}}\left(\int_0^t\int_{\mathbb{R}} \phi_1(x)\frac{\mathrm{d}G}{\mathrm{d}x}(s,\mathrm{d}x;r,y)\,\mathrm{d}s\right)\frac{f(u^\varepsilon(r,y))}{\sigma(\varepsilon)}\,L^\varepsilon(\mathrm{d}r,\mathrm{d}y).\end{aligned} \tag{2.5.28}$$

We further calculate for both inner integrals in (2.5.28) and fixed $0 \leq r \leq t$ and $y \in \mathbb{R}$,

$$\int_0^t\int_{\mathbb{R}} G(s,x;r,y)\phi_2''(x)\,\mathrm{d}x\,\mathrm{d}s = \frac{1}{2}\phi_2(y \pm (t-r)) - \phi_2(y) = \int_{\mathbb{R}} \phi_2(x)\frac{\mathrm{d}G}{\mathrm{d}x}(t,\mathrm{d}x;r,y) - \phi_2(y) \quad \text{and}$$

$$\int_0^t\int_{\mathbb{R}} \phi_1(x)\frac{\mathrm{d}G}{\mathrm{d}x}(s,\mathrm{d}x;r,y)\,\mathrm{d}s = \int_{\mathbb{R}} G(t,z;r,y)\phi_1(z)\,\mathrm{d}z.$$

Now insert the last integral accordingly in (2.5.28) and apply again the stochastic Fubini theorem. □

The next theorem is a converse of Proposition 2.5.6 in the following sense: If a random field on $[0,T]\times[-T,L+T]$ satisfies (together with an auxiliary distribution-valued process) the weak formulation of the stochastic wave equation (on $\mathbb{R}^+\times\mathbb{R}$) driven by Gaussian noise "restricted" to $[0,T]\times[-T,L+T]$, then it is a mild solution to (2.2.9) on $[0,T]\times[0,L]$.

Theorem 2.5.7. *On a complete stochastic basis $(\widetilde{\Omega}, \widetilde{\mathcal{F}}, \widetilde{\boldsymbol{F}}, \widetilde{\mathbb{P}})$, let $\widetilde{W}$ be a Gaussian space–time white noise on $[0,T]\times[-T,L+T]$ for some $T > 0$ and $L > 0$. Assume we have a $\preceq$-càdlàg random field $w = \{w(t,x) \mid (t,x) \in [0,T]\times[-T,L+T]\}$ satisfying*

$$\operatorname*{ess\,sup}_{(t,x)\in[0,T]\times[-T,L+T]} \mathbb{E}\left[\left|w(t,x)\right|^2\right] < \infty \tag{2.5.29}$$

and a $H_{-r}(\mathbb{R})$-valued càdlàg process $(\theta_t)_{t\leq T}$ for some $r > 2$. Assume for any $x \in [-T, L+T]$, $w(0,x) = \theta_0 = 0$ $\widetilde{\mathbb{P}}$-a.s. and that for all $\phi_1, \phi_2 \in C_c^\infty((-T,L+T))$, the pair (w,θ) satisfies (2.4.12) *with probability one. Then w is on $[0,T]\times[0,L]$ the continuous mild solution to the stochastic wave equation* (2.4.13) *driven by $\dot{\widetilde{W}}$, and θ satisfies* (2.4.14) *for all $t \leq T$ and $\phi \in C_c^\infty((0,L))$.*

For the proof of this theorem, we need the following technical lemma.

Lemma 2.5.8. *Let $T > 0$ and $I \subseteq \mathbb{R}$ be a finite* open *interval. The tensor product $C^\infty([0,T]) \otimes C_c^\infty(I)$ is dense in $C_c^\infty([0,T]\times I)$ with respect to each norm $\sum_{|\alpha|\leq N}\|\cdot\|_{\infty,\alpha}$ with $N \in \mathbb{N}$ and*

$$\|f\|_{\infty,\alpha} = \|f^{(\alpha)}\|_\infty = \sup_{(t,x)\in[0,T]\times I}\left|f^{(\alpha)}(t,x)\right| \tag{2.5.30}$$

and $f^{(\alpha)} = \partial_t^{\alpha_1}\partial_x^{\alpha_2} f$ with multi-index $\alpha = (\alpha_1,\alpha_2) \in \mathbb{N}^2$.

Proof. Assume for simplicity $I = (0,1)$, fix $f \in C_c^\infty([0,T]\times I)$ and a compact set $A \subseteq I$ such that $\operatorname{supp} f \subseteq [0,T]\times A$. Furthermore, let b be a $C_c^\infty(\mathbb{R})$-function such that $0 \le b \le 1$, $b \equiv 1$ on A and $A \subseteq \operatorname{supp} b \subsetneq I$. Set $K = \operatorname{supp} b$.

The set of all polynomials on $[0,T]\times K$ is dense in $C^\infty([0,T]\times K)$ with respect to each norm $\sum_{|\alpha|\le N} \|\cdot\|_{\infty,\alpha}$ (with the obvious restriction of domain of definition). We prove this by induction on the differentiation order N: If $N = 0$, it is a direct consequence of the Stone–Weierstrass theorem and if the claim holds for $N-1$, choose $g \in C^\infty([0,T]\times K)$ and write $g(t,x) = \int_0^t \partial_t g(s,x)\,\mathrm{d}s + \int_a^x \partial_x g(0,y)\,\mathrm{d}y + g(a,0)$ (assuming $K=[a,b]$ for simplicity). By assumption, we can find polynomials A_n, resp. B_n, that converge in $\sum_{|\alpha|\le N-1}\|\cdot\|_{\infty,\alpha}$ to $\partial_t g$, resp. $\partial_x g$. Then the polynomial $C_n(t,x) = \int_0^t \partial_t A_n(s,x)\,\mathrm{d}s + \int_a^x B_n(0,y)\,\mathrm{d}y + g(a,0)$ converges to g in $\sum_{|\alpha|\le N}\|\cdot\|_{\infty,\alpha}$.

Now fix $N \in \mathbb{N}$ and choose a sequence of polynomials $P_n(t,x) = \sum_{i,j=1}^{N_n} \alpha_{i,n}\beta_{j,n} t^i x^j$ with $\alpha_{i,n}, \beta_{j,n} \in \mathbb{R}$ and $N_n \in \mathbb{N}$ such that $P_n^{(\alpha)}$ converges to $f^{(\alpha)}$ uniformly on $[0,T]\times K$ for all $|\alpha| \le N$. We can write $b(x)P_n(t,x) = \sum_{i,j=0}^{N_n} \alpha_{i,n}\beta_{j,n} t^i (b(x)x^j)$, so each bP_n lies in $C^\infty([0,T]) \otimes C_c^\infty(I)$ since $b \in C_c^\infty(I)$. We now make the following calculations on $[0,T]\times I$. By the Leibniz rule, $\|(f-bP_n)^{(0,k)}\|_\infty \le \|f^{(0,k)} - bP_n^{(0,k)}\|_\infty + C\sum_{l=1}^k \|b^{(l)} P_n^{(l,k-l)}\|_\infty$ for any $k \in \mathbb{N}$. The first term on the right-hand side can further be estimated by

$$\sup_{(t,x)\in[0,T]\times A} \left| f^{(0,k)}(t,x) - P_n^{(0,k)}(t,x) \right| + \sup_{(t,x)\in[0,T]\times(K\setminus A)} \left| P_n^{(0,k)}(t,x) \right|,$$

which goes to 0 for each $k \le N$ by assumption on P_n. On the other hand,

$$\sum_{l=1}^k \|b^{(l)} P_n^{(l,k-l)}\|_\infty \le \left(\max_{l=1,\dots,k} \|b^{(l)}\|_\infty \right) \sum_{l=1}^k \sup_{(t,x)\in[0,T]\times(K\setminus A)} \left| P_n^{(l,k-l)}(t,x) \right|,$$

which also goes to zero for all $k \le N$. Hence, $\|(f-bP_n)^{(0,k)}\|_\infty \longrightarrow 0$ as $n \to \infty$ and the same holds for multi-indices $(k,0)$ with $k \le N$ as b is time independent. This concludes the proof. □

Proof of Theorem 2.5.7. The proof is inspired by Theorem 9.15 in [83]. The key idea is that we can extend (2.4.12) to test functions with a space *and* a time variable. To be precise, for any $\psi_1, \psi_2 \in C_c^\infty([0,T]\times(-T, L+T))$, we will show that $\widetilde{\mathbb{P}}$-almost surely,

$$\begin{aligned} &\int_{\mathbb{R}} w(t,x)\psi_1(t,x)\,\mathrm{d}x + \langle \theta_t, \psi_2(t,\cdot)\rangle \\ &\quad = \int_0^t \int_{\mathbb{R}} w(s,x) \left(\frac{\partial \psi_1}{\partial t}(s,x) + \frac{\partial^2 \psi_2}{\partial x^2}(s,x) \right) \mathrm{d}s\,\mathrm{d}x + \int_0^t \left\langle \theta_s, \psi_1(s,\cdot) + \frac{\partial \psi_2}{\partial t}(s,\cdot) \right\rangle \mathrm{d}s \\ &\qquad + \int_0^t \int_{\mathbb{R}} \psi_2(s,x) f(w_-(s,x))\, \widetilde{W}(\mathrm{d}s,\mathrm{d}x) \quad \text{for all} \quad t \le T. \end{aligned} \tag{2.5.31}$$

First, we show (2.5.31) for special functions

$$\Psi_i(t,x) = \varphi(t)\phi_i(x) \quad \text{with} \quad \varphi \in C^\infty([0,T]) \quad \text{and} \quad \phi_i \in C_c^\infty((-T, L+T)). \tag{2.5.32}$$

Using the integration by parts formula for càdlàg functions of Proposition 9.16 in [83] and taking into account the initial conditions of w and θ, we compute for all $t \le T$,

$$\begin{aligned} &\int_{\mathbb{R}} w(t,x)\psi_1(t,x)\,\mathrm{d}x + \langle \theta_t, \psi_2(t,\cdot)\rangle = \varphi(t) \left(\int_{\mathbb{R}} w(t,x)\phi_1(x)\,\mathrm{d}x + \langle \theta_t, \phi_2\rangle \right) \\ &\quad = \int_0^t \varphi'(s) \left(\int_{\mathbb{R}} w(s,x)\phi_1(x)\,\mathrm{d}x + \langle \theta_s, \phi_2\rangle \right) \mathrm{d}s + \int_0^t \varphi(s)\,\mathrm{d}\left(\int_{\mathbb{R}} w(s,x)\phi_1(x)\,\mathrm{d}x + \langle \theta_s, \phi_2\rangle \right). \end{aligned} \tag{2.5.33}$$

Now the last integral process in (2.5.33) is indistinguishable from the process

$$t \mapsto \int_0^t \varphi(s) \left(\int_{\mathbb{R}} w(s,x)\phi_2''(x)\,\mathrm{d}x + \langle \theta_s, \phi_1 \rangle \right) \mathrm{d}s + \int_0^t \int_{\mathbb{R}} \varphi(s)\phi_2(x) f(w_-(s,x))\, \widetilde{W}(\mathrm{d}s, \mathrm{d}x), \tag{2.5.34}$$

since its integrator equals the right-hand side of (2.4.12) by assumption. Inserting (2.5.34) into (2.5.33) and recombining the functions ψ_i as well as their derivatives yields exactly (2.5.31).

Next, we prove (2.5.31) for general $\psi_i \in C_c^\infty([0,T]\times(-T,L+T))$ by a density argument. Let $N_0 \in \mathbb{N}$ to be determined later in the proof. Using Lemma 2.5.8, choose sequences $(\psi_i^n)_{n\in\mathbb{N}} \in C^\infty([0,T])\otimes C_c^\infty((-T,L+T))$ such that ψ_i^n converges to ψ_i in $\sum_{|\alpha|\le N_0} \|\cdot\|_{\infty,\alpha}$ with each $\|\cdot\|_{\infty,\alpha}$ as in (2.5.30). This implies uniform convergence in $[0,T]$ of each of the corresponding terms in (2.5.31) as we show in the following. (Note that by linearity, (2.5.31) readily holds for linear combinations of special functions (2.5.32).)

Since w is $\preceq$-càdlàg, and θ is càdlàg in $H_{-r}(\mathbb{R})$, both processes are bounded and therefore,

$$\sup_{t\le T} \left| \int_{\mathbb{R}} w(t,x)\psi_1(t,x)\,\mathrm{d}x - \int_{\mathbb{R}} w(t,x)\psi_1^n(t,x)\,\mathrm{d}x \right| \le C\|w\|_\infty \|\psi_1 - \psi_1^n\|_\infty \longrightarrow 0 \quad \text{as} \quad n\to\infty$$

as well as

$$\sup_{t\le T} \left| \langle \theta_t, \psi_2(t,\cdot)\rangle - \langle \theta_t, \psi_2^n(t,\cdot)\rangle \right| \le \left(\sup_{t\le T} \|\theta_t\|_{-r} \right) \sup_{t\le T} \|\psi_2(t,\cdot) - \psi_2^n(t,\cdot)\|_r < \infty. \tag{2.5.35}$$

We now show that for any $r\ge 0$, $\psi_i^n \longrightarrow \psi_i$ in all $\|\cdot\|_{\infty,\alpha}$ with $|\alpha|\le N_0$ and sufficiently large N_0 implies $\psi_i^n \longrightarrow \psi_i$ and $\partial_t\psi_2^n \longrightarrow \partial_t\psi_2$ in $\sup_{t\le T}\|\cdot\|_r$ (thus convergence to 0 of all terms in (2.5.35)). For this, we use the well-known differential equation satisfied by the Hermite functions

$$h_q''(x) + (1+2q-x^2)h_q(x) = 0 \quad \text{for} \quad x\in\mathbb{R} \quad \text{and} \quad q\in\mathbb{N}. \tag{2.5.36}$$

Let $q_0\in\mathbb{N}$ be such that $\sqrt{1+2q} > L+T$ for all $q\ge q_0$. Then $1/|x^2-(1+2q)| \le 1/((1+2q)-(L+T)^2)$ on $[-T,L+T]$ for all $q\ge q_0$. Let $\phi\in C_c^\infty((-T,L+T))$. Insert (2.5.36) into $\langle\phi,h_q\rangle$ and use integration by parts twice, repeat k times this procedure, apply then Hölder's inequality and the elementary inequality above to see that for all $q\ge q_0$ and $k\in\mathbb{N}$,

$$\langle \phi, h_q\rangle^2 \le \frac{C}{((1+2q)-(L+T)^2)^{2k}} \left(\sum_{l=0}^{2k} \|\phi^{(l)}\|_\infty^2 \right) \int_{-T}^{L+T} P_{4k}(x)\,\mathrm{d}x \tag{2.5.37}$$

with a polynomial P_{4k} of degree $4k$ (the remaining details of these calculations are left to the reader). Now choose N_0 such that $r-N_0<-3$ and infer from (2.5.37) that

$$\begin{aligned} &\|\partial_t\psi_i^n(t,\cdot) - \partial_t\psi_i(t,\cdot)\|_r^2 \\ &\quad\le C\left(\sum_{q=0}^{q_0} (1+2q)^r + \sum_{q=q_0+1}^{\infty} \frac{(1+2q)^r}{((1+2q)-(L+T)^2)^{N_0-2}} \right) \sum_{l=0}^{N_0-2} \|\partial_t\partial_x^l(\psi_i^n - \psi_i)\|_\infty^2 \end{aligned} \tag{2.5.38}$$

for all $t\le T$ and $n\in\mathbb{N}$. The series in (2.5.38) is finite since $r-N_0<-3$ and the last sum converges to 0 as $n\to\infty$ by assumption on (ψ_i^n), which proves the desired convergences.

Finally, by Doob's inequality, Itō's isometry, the Lipschitz continuity of f and (2.5.29),

$$\begin{aligned} &\mathbb{E}\left[\sup_{t\le T} \left| \int_0^t \int_{\mathbb{R}} \psi_2(s,x) f(w_-(s,x))\,\widetilde{W}(\mathrm{d}s,\mathrm{d}x) - \int_0^t\int_{\mathbb{R}} \psi_2^n(s,x) f(w_-(s,x))\,\widetilde{W}(\mathrm{d}s,\mathrm{d}x) \right| \right] \\ &\qquad \le \int_0^T \int_{\mathbb{R}} (\psi_2(s,x) - \psi_2^n(s,x))^2 \mathbb{E}\left[f(w_-(s,x))^2 \right] \mathrm{d}s\,\mathrm{d}x \le C\|\psi_2 - \psi_2^n\|_\infty^2 \longrightarrow 0 \quad \text{as} \quad n\to\infty. \end{aligned}$$

Analogous arguments for the remaining terms of (2.5.31) finishes the density argument.

We now choose two particular functions to be inserted in (2.5.31). Fix $\phi \in C_c^\infty((0,L))$ as well as $t \leq T$, and define

$$\psi_1(s,y) = \frac{1}{2}\phi(y \pm (t-s)) \quad \text{and} \quad \psi_2(s,y) = \frac{1}{2}\int_{y-(t-s)}^{y+(t-s)} \phi(x)\,\mathrm{d}x \tag{2.5.39}$$

with $(s,y) \in [0,T]\times(-T,L+T)$. Then $\psi_1, \psi_2 \in C_c^\infty([0,T]\times(-T,L+T))$. In addition, $\psi_1(t,y) = \phi(y)$ and $\psi_2(t,y) = 0$ for all $y \in \mathbb{R}$, and straightforward calculus yields

$$\psi_2^{(1,0)}(s,y) = -\psi_1(s,y) \quad \text{and} \quad \psi_2^{(2,0)}(s,y) = \psi_2^{(0,2)}(s,y). \tag{2.5.40}$$

The freedom we have to choose two different functions in (2.5.39) is another reason why we considered the weak formulation (2.3.8) of the stochastic wave equation in this work: By (2.5.40), the first two integrals on the right-hand side of (2.5.31) *vanish*, and (2.5.31) yields at time point t

$$\int_{\mathbb{R}} w(t,x)\phi(x)\,\mathrm{d}x = \int_0^t \int_{\mathbb{R}} \psi_2(s,y) f(w_-(s,y))\,\widetilde{W}(\mathrm{d}s,\mathrm{d}y) \quad \widetilde{\mathbb{P}}\text{-almost surely,}$$

which, recalling (2.3.9) and using the stochastic Fubini theorem, has the equivalent form

$$\int_{\mathbb{R}} \left(w(t,x) - \int_0^t \int_{\mathbb{R}} G(t,x;s,y) f(w_-(s,y))\,\widetilde{W}(\mathrm{d}s,\mathrm{d}y) \right) \phi(x)\,\mathrm{d}x = 0 \quad \widetilde{\mathbb{P}}\text{-almost surely,} \tag{2.5.41}$$

and this holds for all $\phi \in C_c^\infty((0,L))$ and $t \leq T$. We can now infer the first claim of the theorem. Denote by $Z_t(x)$ the random field in parenthesis in (2.5.41) with $(t,x) \in [0,T]\times[0,L]$. It is easy to see that (2.5.29) implies $Z_t \in L^2([0,L])$ for all $t \leq T$. For any $\epsilon > 0$ and $t \leq T$, consider the mollified random field $J_\epsilon(Z_t)$ on $[0,L]$ defined exactly as in (1.8) of Chapter 10 in [49]. By Lemma 3 of that chapter, $J_\epsilon(Z_t) \longrightarrow Z_t$ in $L^2([0,L])$ as $\epsilon \to 0$. Consequently, we can choose a sequence $(\epsilon_l)_{l\in\mathbb{N}}$ converging to 0 such that $\widetilde{\omega}$-wise,

$$J_{\epsilon_l}(Z_t) \longrightarrow Z_t \quad \mathrm{Leb}_{[0,L]}\text{-almost everywhere} \quad \text{as} \quad l \to \infty. \tag{2.5.42}$$

Now for any fixed $y \in (0,L)$, the support of the function $\rho((y-\cdot)/\epsilon_l)/\epsilon_l^2$ used to mollify Z_t will be contained in $(0,L)$ if ϵ_l is sufficiently small and, hence, (2.5.41) applies to $J_{\epsilon_l}(Z_t)(y)$ for those ϵ_l with ϕ being chosen as $\rho((y-\cdot)/\epsilon_l)/\epsilon_l^2$. Combining this with (2.5.42) has the following outcome: For all $t \leq T$ and almost all $y \in (0,L)$, $Z_t(y) = 0$ $\widetilde{\mathbb{P}}$-almost surely. We deduce that w_- satisfies the mild formulation of (2.4.13) almost everywhere on $[0,T]\times[0,L]$.

Let $\widetilde{u}$ be the continuous mild solution to (2.4.13) on $(\widetilde{\Omega}, \widetilde{\mathcal{F}}, \widetilde{\boldsymbol{F}}, \widetilde{\mathbb{P}})$. We have

$$\mathbb{E}\left[\left|\widetilde{u}(t,x) - w_-(t,x)\right|^2\right] \leq \int_0^T \int_{\mathbb{R}} G^2(t,x;s,y)\mathbb{E}\left[\left|\widetilde{u}(s,y) - w_-(s,y)\right|^2\right]\mathrm{d}s\,\mathrm{d}y \quad \text{a.e.}$$

and by Lemma 6.4 (3) in [28], $\mathbb{E}\left[|\widetilde{u}(t,x) - w_-(t,x)|^2\right] = 0$ for almost all $(t,x) \in [0,T]\times[0,L]$. It follows that $\widetilde{\mathbb{P}}$-almost surely, w and $\widetilde{u}$ agree almost everywhere on $[0,T]\times[0,L]$ and therefore, since w is $\preceq$-càdlàg, they are indistinguishable and w is actually continuous on $[0,T]\times[0,L]$. Finally, by the usual computations, we obtain for all $(t,x) \in [0,T]\times[0,L]$,

$$w(t,x) = \widetilde{u}(t,x) = \int_0^t \int_{\mathbb{R}} G_{t-s}(x,y) f(\widetilde{u}(s,y))\,\widetilde{W}(\mathrm{d}s,\mathrm{d}y) = \int_0^t \int_{\mathbb{R}} G_{t-s}(x,y) f(w(s,y))\,\widetilde{W}(\mathrm{d}s,\mathrm{d}y)$$

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

For the second claim of the theorem, we define new functions ψ_1, ψ_2 by

$$\psi_1(s,y) = \frac{1}{2}\left(\phi'(y+(t-s)) - \phi'(y-(t-s))\right) \quad \text{and} \quad \psi_2(s,y) = \frac{1}{2}\phi(y \pm (t-s))$$

with fixed $\phi \in C_c^\infty((0,L))$ and $t \leq T$, for all $(s,y) \in [0,T] \times (-T, L+T)$. Again we have $\psi_1, \psi_2 \in C_c^\infty([0,T] \times (-T, L+T))$. Since $w_- = w$, by straightforward calculus and again with (2.3.9), inserting ψ_1, ψ_2 into (2.5.31) yields at time point t exactly (2.4.14). This concludes the proof. $\square$

Chapter 3

Mixed semimartingales: Volatility estimation in the presence of fractional noise

3.1 Introduction

Over the last two decades, a large amount of work has been dedicated to the problem of estimating volatility for a continuous Itô semimartingale X based on high-frequency observations. Motivated by financial applications, sophisticated methods have been employed to construct volatility estimators that are robust to, for example, jumps, irregular observation schemes and/or the presence of market microstructure noise ([3, 64]). Concerning the last point, the type of noises considered in the literature is usually one of the following two (or a combination thereof): rounding errors due to the discreteness of prices ([41, 60, 77, 88, 89]) or additive noise due to data errors, informational asymmetries, transaction costs etc. In the latter case, one assumes that observations at high frequency are only available for

$$Y_t = X_t + Z_t, \tag{3.1.1}$$

where $(X_t)_{t\geq 0}$ is the efficient price process and $(Z_t)_{t\geq 0}$ is a noise process, both of which are unobservable. A common approach in the literature is to model $(Z_t)_{t\geq 0}$ at the observation times (say, $i\Delta_n$ for $i = 1, \ldots, [T/\Delta_n]$ where Δ_n is a small step size and $T > 0$ is a finite time horizon) as

$$Z_{i\Delta_n} = \varepsilon_i^n, \tag{3.1.2}$$

where for each n, $(\varepsilon_i^n)_{i=1}^{[T/\Delta_n]}$ is a discrete time series. Examples for ε_i^n include i.i.d. noise [11, 15, 16, 35, 85, 87, 100], AR- or ARCH-type noise [5, 95], and nonparametric variants thereof [56, 61, 62, 63, 67, 78, 79]. We also refer to [12, 51] for overviews of and comparisons between the estimators developed in the mentioned works.

In these *non-shrinking noise models* where the (conditional) variance of ε_i^n is bounded away from 0 as $\Delta_n \to 0$, it is known that under mild assumptions on ε_i^n, the *realized variance (RV)* of the noisy process Y, defined as $\widehat{V}_{0,T}^n = \sum_{i=1}^{[T/\Delta_n]} (Y_{i\Delta_n} - Y_{(i-1)\Delta_n})^2$, diverges at a rate of Δ_n^{-1} as $\Delta_n \to 0$. While this behavior was empirically confirmed in previous studies[1], examining quotes

[1] For example, [5] found that the average RV estimator of the 30 Dow Jones Industrial Average (DJIA) stocks during the last ten trading days of April 2004 exploded as $\Delta_n \to 0$, with a slope coefficient of 1.02 on a log–log scale.

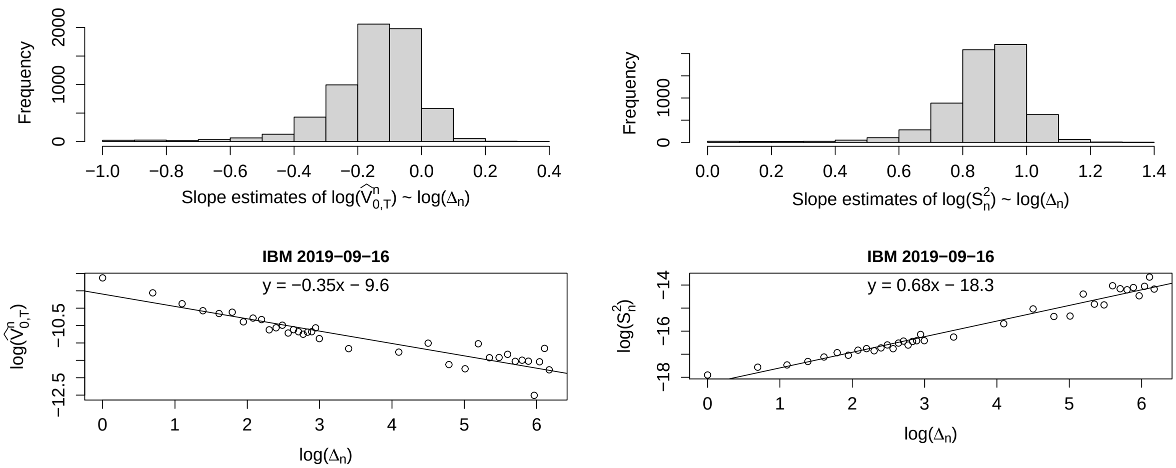


Figure 3.1: Top row: Histograms of slope estimates obtained in a linear regression of $\log \widehat{V}^n_{0,T}$ (left) and $\log S^2_n$ (right) on $\log \Delta_n$ (where Δ_n expressed in seconds). Both $\widehat{V}^n_{0,T}$ and S^2_n are computed for the logarithmic mid-quote data of 29 DJIA stocks on all trading days in 2019. Each data point corresponds to one company and day. Bottom row: One particular asset and day (IBM on September 16, 2019) including least-square estimates.

data for the individual DJIA stocks in 2019, we found a substantial deviation of the regression coefficient of $\log \widehat{V}^n_{0,T}$ on $\log \Delta_n$ from the theoretical value of -1 implied by non-shrinking noise models. In fact, for over 90% of the examined assets and days, the estimated slope coefficient was larger than -0.3; cf. the left histogram depicted in Figure 3.1. This indicates that while the RV estimator explodes as $\Delta_n \to 0$, in general, the rate of explosion need not be Δ_n^{-1} and, in particular, may be asset- and/or time-dependent. Similarly, if we compute the sample variance S^2_n of increments at different frequencies Δ_n for the same data set, again in over 90% of the cases, the slope estimate in a regression of $\log S^2_n$ on $\log \Delta_n$ was larger than 0.6 (see the right histogram in Figure 3.1), whereas non-shrinking noise models would predict stabilization of the variance for small Δ_n (i.e., zero slope).

Shrinking noise models, in which the variance of the noise decreases as $\Delta_n \to 0$, are much less studied in the literature. In [3, Chapter 7], the noise variables ε^n_i are allowed to have vanishing variances as $\Delta_n \to 0$, but conditionally on the filtration generated by ingredients of X, they are assumed to be independent of each other for different values of i. In [71], the noise model is (essentially) $\Delta_n^{\alpha/2}$ times i.i.d. variables, where $\alpha \in [0, \frac{1}{2})$, which can only explain scaling exponents of $\log \widehat{V}^n_{0,T}$ (resp., $\log S^2_n$) in the range $[-1, -\frac{1}{2})$ (resp., $[0, \frac{1}{2})$). In [6], noises with a variance proportional to Δ_n^γ with $\gamma \geq \frac{3}{2}$ are considered, but this cannot explain the slope distribution in the top-right panel of Figure 3.1. Moreover, in these three works, the noise is white and admits no serial dependence, conditionally on the price process. In particular, the increments of the observed price process are asymptotically uncorrelated at lags 2 or higher. However, in line with the results of [5], we find a significant second-order autocorrelation coefficient for the majority of companies and days in our sample; see Figure 3.2.

A considerably more general treatment of shrinking noises *with* (conditional) serial dependence is found in the recent paper [38], where the authors essentially assume that

$$Z_{i\Delta_n} = \Delta_n^\gamma \rho_{i\Delta_n} \varepsilon^n_i, \tag{3.1.3}$$

where ρ is a noise volatility process, ε^n_i is a discrete-time infinite moving-average process and γ is

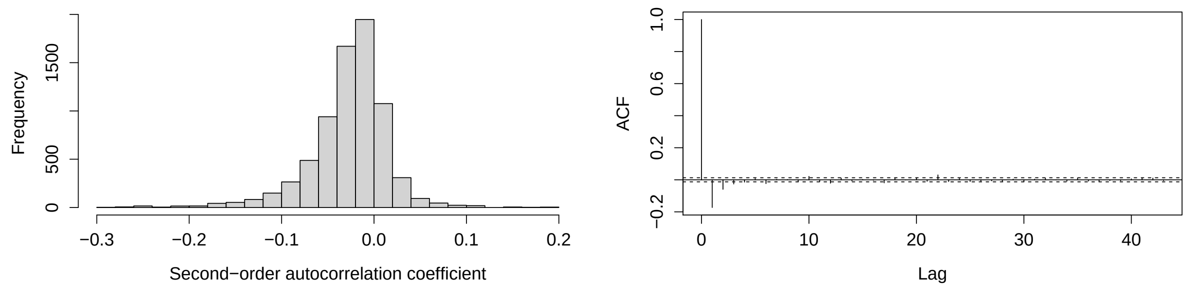


Figure 3.2: Left: Histogram of second-order autocorrelation coefficients of 1s increments for 29 DJIA stocks and all trading days in 2019. Right: Autocorrelation function for one particular asset and day (IBM on September 16, 2019).

parameter that determines the speed at which the noise variance shrinks. As noted in [3, Chapter 7.1], a typical issue of discrete-time noise processes with shrinking variance is compatibility between different frequencies: in (3.1.3), if $i\Delta_n = j\Delta_m$, we typically do not have $Z_{i\Delta_n} = Z_{j\Delta_m}$ because the variances are of order Δ_n^γ and Δ_m^γ, respectively.[2] Let us also mention that [38, Assumption 5] essentially requires the autocovariances of the noise at lag r decay faster than r^{-3}. In another recent paper [79], a decay rate of more than r^{-2} is needed. As we explain below, this, in fact, cannot hold for a large class of noise models as soon as a natural compatibility condition between frequencies is assumed.

In view of the previous observations, we will introduce a noise model that

(a) captures market microstructure noise in continuous time without compatibility issues between different sampling frequencies;

(b) reproduces scaling exponents of $\log \widehat{V}^n_{0,T}$ as a function of $\log \Delta_n$ in the full range of $(-1, 0)$;

(c) reproduces decay rates of S_n^2 in the full range of $(0, 1)$; and

(d) retains desirable features of existing noise models such as serial dependence, dependence on the price process, cross-sectional dependence between assets and diurnal heteroscedasticity.

3.1.1 Model

On a filtered probability space $(\Omega, \mathcal{F}, \mathbb{F} = (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$ satisfying the usual conditions, we assume that the latent efficient price process X is a d-dimensional continuous Itô semimartingale

$$X_t = X_0 + \int_0^t a_s \,\mathrm{d}s + \int_0^t \sigma_s \,\mathrm{d}B_s, \qquad t \geq 0, \tag{3.1.4}$$

where B is a standard $\mathbb{F}$-Brownian motion in $\mathbb{R}^d$ and a and σ are $\mathbb{F}$-adapted locally bounded $\mathbb{R}^d$- and $\mathbb{R}^{d\times d}$-valued processes, respectively. Our assumption on the noise is as follows:

Assumption (Z). The process $(Z_t)_{t\geq 0}$ is given by

$$Z_t = Z_0 + \int_0^t g(t-s)\rho_s \,\mathrm{d}W_s, \qquad t \geq 0, \tag{3.1.5}$$

[2] In the white noise case, to ensure compatibility, $(Z_t)_{t\geq 0}$ must be a collection of independent variables. But then a quantity like $\frac{1}{T}\int_0^T Z_t \,\mathrm{d}t$, the average observational error on $[0, T]$, though meaningful from a practical point of view, cannot be defined anymore because $t \mapsto Z_t$ is non-measurable.

where W is another d-dimensional standard $\mathbb{F}$-Brownian motion, independent of B, and $(\rho_t)_{t\geq 0}$ is an $\mathbb{F}$-adapted locally bounded $\mathbb{R}^{d\times d}$-valued process. The kernel $g\colon (0,\infty)\to\mathbb{R}$ is of the form

$$g(t) = K_H^{-1} t^{H-\frac{1}{2}} + g_0(t) \tag{3.1.6}$$

where $H \in (0, \frac{1}{2})$,

$$K_H = \sqrt{\frac{1}{2H} + \int_1^\infty \left(r^{H-\frac{1}{2}} - (r-1)^{H-\frac{1}{2}}\right)^2 \mathrm{d}r} = \frac{\sqrt{2H\sin(\pi H)\Gamma(2H)}}{\Gamma(H+\frac{1}{2})} \tag{3.1.7}$$

is a normalizing constant (the second identity can be found in [81, Theorem 1.3.1]) and the function $g_0\colon [0,\infty)\to\mathbb{R}$ is smooth (including at $t=0$) with $g_0(0)=0$.

In other words, we assume that the noise process is a continuous-time moving average process (with ρ representing diurnal features of the noise) for which the kernel behaves as $t^{H-1/2}$ as $t\to 0$. If $H\in(0,\frac{1}{2})$, then our results show that $\widehat{V}^n_{0,T}$ diverges at a rate of Δ_n^{2H-1}. Hence, the exponent covers the whole interval $(-1,0)$. Moreover, the increments of Z over a time interval of length Δ_n have variances of order Δ_n^{2H}, so the exponent can be any value in $(0,1)$. If $H=\frac{1}{2}$, then Z itself is a semimartingale and there is no way to discern Z from the efficient price process X; if $H>\frac{1}{2}$, then Z is smoother than X and $\widehat{V}^n_{0,T}$ converges to $\int_0^t \sigma_s^2\,\mathrm{d}s$, as in the noise-free case. This is why we exclude the case $H\geq\frac{1}{2}$ in Assumption (Z). Note that X and Z may be dependent through ρ.

In the special case where $g_0\equiv 0$ and $\rho_s\equiv\rho$ is a constant, Z is—up to a term of finite variation—simply a multiple of a *fractional Brownian motion (fBM)*. If further $X_t=\sigma B_t$ with constant volatility σ, then the resulting observed process $Y_t=\sigma B_t+\rho Z_t$ is a *mixed fractional Brownian motion (mfBM)* as introduced by [25]. Our model for Y, as the sum of X in (3.1.4) and Z in (3.1.5), can be viewed as a nonparametric generalization of mfBM that allows for stochastic volatility in both its Brownian and its fractional component. We do keep the parameter H, though, which we call the *Hurst index* in analogy with fBM. This is why we refer to our model for the observation process

$$Y_t = X_t + Z_t = X_0 + \int_0^t a_s\,\mathrm{d}s + \int_0^t \sigma_s\,\mathrm{d}B_s + Z_0 + \int_0^t g(t-s)\rho_s\,\mathrm{d}W_s, \qquad t\geq 0, \tag{3.1.8}$$

as a *mixed semimartingale*.

To our best knowledge, the mixed semimartingale hypothesis also leads to the first continuous-time microstructure noise model in which $t\mapsto Z_t$ is a measurable process and, at the same time, the divergence of the RV estimator as $\Delta_n\to 0$ can be explained.[3] The condition (3.1.6) on the kernel g implies that Z locally resembles an fBM with Hurst index H. At first sight, this might seem restrictive, but actually it is not: Suppose that $(Z'_t)_{t\geq 0}$ is a measurable stationary noise process. Under mild assumptions, Z' admits a continuous-time moving average representation $Z'_t=\int_{-\infty}^t g(t-s)\rho_s\,\mathrm{d}B_s$ for some $g\in L^2((0,\infty))$ and stationary ρ; see, for example, [45, Chapter XII, Theorem 5.3]. If we ignore ρ (i.e., take $\rho\equiv 1$) and assume that the autocorrelation functions (ACFs) of the increments of Z' are compatible between frequencies, that is, $\Gamma_r^{(n)}=\mathrm{Corr}(Z'_{i\Delta_n}-Z'_{(i-1)\Delta_n}, Z'_{(i+r)\Delta_n}-Z'_{(i+r-1)\Delta_n})$ converges to some Γ_r for all $r\geq 0$ as

[3] In [4, 55], continuous-time noise models are considered for which the noise increments are of the same order as the efficient price increments. Thus, the RV estimator does not explode at high frequency in these models. In [55], a moving-average noise model with finite dependence is briefly mentioned in Section 4.1, but only unbiasedness of an associated estimator is shown subsequently.

$n \to \infty$, then, as explained in [14, Remark 7], we necessarily have $\Gamma_r = \Gamma_r^H$ for some $H \in (0,1)$ where

$$\Gamma_0^H = 1 \quad \text{and} \quad \Gamma_r^H = \frac{1}{2}\left((r+1)^{2H} - 2r^{2H} + (r-1)^{2H}\right), \quad r \geq 1. \tag{3.1.9}$$

This is exactly the ACF of the increments of fractional Brownian motion and, if $H \in (0, \frac{1}{2})$ and $\rho \equiv 1$, the limiting ACF of the increments of Z in (3.1.5) as $\Delta_n \to 0$. We conclude that Z from (3.1.5) captures the behavior of a large class of "compatible" stationary noise processes in the high-frequency limit. Since $r \mapsto \Gamma_r^H$ decays like r^{2H-2} as $r \to \infty$, which is slower than r^{-2} for any $H \in (0,1)$, the assumptions imposed on the serial dependence of the noise in [38, 62, 79] are not satisfied.

Note that fractional Brownian motion and other fractional models were also considered as asset price models in the mathematical finance literature, often in the context of long-range dependence; see the seminal work of [80] but also [18, 19, 20], for example. In those works, it is typically the behavior of the kernel g at $t = \infty$ that is of primary interest, as this determines whether the resulting process has short or long memory. Our concern, by contrast, is the behavior of this kernel around $t = 0$, which governs the local regularity, or *roughness*, of the high-frequency increments of Z. In fact, on a finite time interval $[0,T]$, there is no way to distinguish between short- and long-range dependence. Therefore, the Hurst parameter H should really be viewed as a measure of roughness in this work.

Let us finally mention that mixed semimartingale models are in line with no-arbitrage concepts in mathematical finance. Clearly, as non-semimartingales, they admit arbitrage in the FLVR sense [42]. However, as shown in [26, 53], mixed fractional Brownian motion does not admit arbitrage in the presence of transaction costs, which are exactly one of the market inefficiencies that microstructure noise models are supposed to capture. Also, in the mixed semimartingale model, the efficient price process is still assumed to follow a continuous semimartingale; it is only the noise process that exhibits fractional and hence non-semimartingale behavior.

3.1.2 Identifiability

Before we describe the main results in next Section, let us comment on the identifiability of the involved parameters and processes. The following result, due to [96] (see also [22, 25]), puts a constraint on estimating volatility in a mixed semimartingale model:

Proposition 3.1.1. *Assume that Y is an mfBM, that is, $Y = X + Z$ where $X = \sigma B$ and $Z = \rho B^H$ and $\rho, \sigma \in (0,\infty)$, B is a Brownian motion and B^H is an independent fBM with Hurst parameter $H \in (0, \frac{1}{2})$. For any $T > 0$, the laws of $(Y_t)_{t\in[0,T]}$ and $(Z_t)_{t\in[0,T]}$ are mutually equivalent if $H \in (0, \frac{1}{4})$ and mutually singular if $H \in [\frac{1}{4}, \frac{1}{2})$.*

In other words, if $H \in (0, \frac{1}{4})$, there is no way to consistently estimate σ in the presence of Z on a finite time interval. Against this background, we will first establish a central limit theorem (CLT) for variation functionals of mixed semimartingales in Section 3.2 and then use this CLT in Section 3.3 to derive consistent and asymptotically mixed normal estimators for H, $\int_0^T \sigma_s^2 \, \mathrm{d}s$ and $\int_0^T \rho_s^2 \, \mathrm{d}s$ if $H \in (\frac{1}{4}, \frac{1}{2})$ and for H and $\int_0^T \rho_s^2 \, \mathrm{d}s$ if $H \in (0, \frac{1}{4})$. Sections 3.4 and 3.5 contain a simulation study and an empirical application of our results, respectively. Section 3.6 concludes.

3.2 Central limit theorem for variation functionals

As with most estimators in high-frequency statistics, ours are based on limit theorems for power variations and related functionals. For semimartingales, this is well studied topic by now; see

[3, 64] for in-depth treatments of this subject. For fractional Brownian motion or moving-average processes as in (3.1.5), the theory is similarly well understood; see [14, 21, 37, 50] and the references therein. Surprisingly, it turns out that the mixed case is more complicated than the "union" of the purely semimartingale and the purely fractional case. A first instance of this are additional bias terms that already appear in the central limit theorem for power variations.

3.2.1 The result

Given $L, M \in \mathbb{N}$ and a test function $f\colon \mathbb{R}^{d\times L} \to \mathbb{R}^M$, our goal is to establish a CLT for *normalized variation functionals* of the form

$$V_f^n(Y,t) = \Delta_n \sum_{i=1}^{[t/\Delta_n]-L+1} f\left(\frac{\underline{\Delta}_i^n Y}{\Delta_n^H}\right),$$

where

$$\Delta_i^n Y = Y_{i\Delta_n} - Y_{(i-1)\Delta_n} \in \mathbb{R}^d, \qquad \underline{\Delta}_i^n Y = (\Delta_i^n Y, \Delta_{i+1}^n Y, \dots, \Delta_{i+L-1}^n Y) \in \mathbb{R}^{d\times L}. \tag{3.2.1}$$

This will be achieved under the following set of assumptions:

Assumption (CLT). Let $\|\cdot\|$ denote the Euclidean norm (in $\mathbb{R}^n$ if applied to vectors and in $\mathbb{R}^{nm}$ if applied to a matrix in $\mathbb{R}^{n\times m}$). We assume that the observation process Y is given by the sum of X from (3.1.4) and Z from (3.1.5) with the following specifications:

(1) The function $f\colon \mathbb{R}^{d\times L} \to \mathbb{R}^M$ is even and infinitely differentiable. Moreover, all its derivatives (including f itself) have at most polynomial growth.

(2) The process a is d-dimensional, locally bounded and $\mathbb{F}$-adapted.

(3) The volatility process σ is an $\mathbb{F}$-adapted locally bounded $\mathbb{R}^{d\times d}$-valued process. Moreover, for every $T > 0$, there is $K_1 \in (0,\infty)$ such that for all $s,t \in [0,T]$,

$$\mathbb{E}\Big[1 \wedge \|\sigma_t - \sigma_s\|\Big] \le K_1 |t-s|^{\frac{1}{2}}. \tag{3.2.2}$$

(4) The noise volatility process ρ takes the form

$$\rho_t = \rho_t^{(0)} + \int_0^t \widetilde{b}_s \,\mathrm{d}s + \int_0^t \widetilde{\rho}_s \,\mathrm{d}\widetilde{W}_s, \qquad t \ge 0, \tag{3.2.3}$$

where

(a) $\rho^{(0)}$ is an $\mathbb{F}$-adapted locally bounded $\mathbb{R}^{d\times d}$-valued process such that for all $T > 0$,

$$\mathbb{E}\Big[1 \wedge \|\rho_t^{(0)} - \rho_s^{(0)}\|\Big] \le K_2 |t-s|^{\gamma}, \qquad s,t \in [0,T], \tag{3.2.4}$$

for some $\gamma \in (\frac{1}{2}, 1]$ and $K_2 \in (0,\infty)$;

(b) $\widetilde{b}$ is $d \times d$-dimensional, locally bounded and $\mathbb{F}$-adapted;

(c) $\widetilde{\rho}$ is an $\mathbb{F}$-adapted locally bounded $\mathbb{R}^{d\times d\times d}$-valued process (for example, the (ij)th component of the stochastic integral in (3.2.3) equals $\sum_{k=1}^d \int_0^t \widetilde{\rho}_s^{ijk} \,\mathrm{d}\widetilde{W}_s^k$) such that for all $T > 0$ there are $\varepsilon > 0$ and $K_3 \in (0,\infty)$ with

$$\mathbb{E}\Big[1 \wedge \|\widetilde{\rho}_t - \widetilde{\rho}_s\|\Big] \le K_3 |t-s|^{\varepsilon}, \qquad s,t \in [0,T]. \tag{3.2.5}$$

(d) $\widetilde{W}$ is a d-dimensional $\mathbb{F}$-Brownian motion that is jointly Gaussian with (B, W).

(5) The kernel g takes the form (3.1.6) with $H \in (0, \frac{1}{2})$ and some $g_0 \in C^\infty([0,\infty))$ with $g_0(0) = 0$.

To describe the CLT for $V^n_f(Y,t)$, we need some more notation. Define μ_f as the $\mathbb{R}^M$-valued function that maps $v = (v_{k\ell,k'\ell'}) \in (\mathbb{R}^{d\times L})^2$ to $\mathbb{E}[f(\mathcal{Z})]$ where $\mathcal{Z} \in (\mathbb{R}^{d\times L})^2$ follows a multivariate normal distribution with mean 0 and $\mathrm{Cov}(\mathcal{Z}_{k\ell}, \mathcal{Z}_{k'\ell'}) = v_{k\ell,k'\ell'}$. Note that μ_f is infinitely differentiable because f is. Furthermore, if $\mathcal{Z}' \in (\mathbb{R}^{d\times L})^2$ is such that $\mathcal{Z}$ and $\mathcal{Z}'$ are jointly Gaussian with mean 0, covariances $\mathrm{Cov}(\mathcal{Z}_{k\ell}, \mathcal{Z}_{k'\ell'}) = \mathrm{Cov}(\mathcal{Z}'_{k\ell}, \mathcal{Z}'_{k'\ell'}) = v_{k\ell,k'\ell'}$ and cross-covariances $\mathrm{Cov}(\mathcal{Z}_{k\ell}, \mathcal{Z}'_{k'\ell'}) = q_{k\ell,k'\ell'}$, we define

$$\gamma_{f_{m_1},f_{m_2}}(v,q) = \mathrm{Cov}\left(f_{m_1}(\mathcal{Z}), f_{m_2}(\mathcal{Z}')\right), \qquad m_1, m_2 = 1,\dots,M.$$

We further introduce a multi-index notation adapted to the definition of μ_f. For $\chi = (\chi_{k\ell,k'\ell'}) \in \mathbb{N}_0^{(d\times L)\times(d\times L)}$ and v as above, we let

$$|\chi| = \sum_{k,k'=1}^{d} \sum_{\ell,\ell'=1}^{L} \chi_{k\ell,k'\ell'}, \quad \chi! = \prod_{k,k'=1}^{d} \prod_{\ell,\ell'=1}^{L} \chi_{k\ell,k'\ell'}, \quad v^\chi = \prod_{k,k'=1}^{d} \prod_{\ell,\ell'=1}^{L} {v_{k\ell,k'\ell'}}^{\chi_{k\ell,k'\ell'}}$$

and

$$\partial^\chi \mu_f = \frac{\partial^{|\chi|}\mu_f}{\partial v_{11,11}^{\chi_{11,11}} \cdots \partial v_{dL,dL}^{\chi_{dL,dL}}}.$$

Finally, recalling (3.1.9), we define for all $k,k' \in \{1,\dots,d\}$, $\ell,\ell' \in \{1,\dots,L\}$ and $r \in \mathbb{N}_0$,

$$\pi_r(s)_{k\ell,k'\ell'} = (\rho_s\rho_s^T)_{kk'}\Gamma^H_{|\ell-\ell'+r|}, \qquad c(s)_{k\ell,k'\ell'} = (\sigma_s\sigma_s^T)_{kk'}\mathbb{1}_{\{\ell=\ell'\}}$$

and, as a special case,

$$\pi(s) = \pi_0(s). \tag{3.2.6}$$

The following CLT is our first main result. We use $\overset{\mathrm{st}}{\Longrightarrow}$ (resp., $\overset{L^1}{\Longrightarrow}$) to denote functional stable convergence in law (resp., convergence in L^1) in the space of càdlàg functions $[0,\infty) \to \mathbb{R}$ equipped with the local uniform topology. In the special case where Y follows the parametric model of an mfBM and the test function is $f(x) = x^2$, the CLT was already obtained by [47].[4]

Theorem 3.2.1. *Grant Assumption (CLT) and let $N(H) = [1/(2-4H)]$. Then we have that*

$$\Delta_n^{-\frac{1}{2}}\left\{V^n_f(Y,t) - \int_0^t \mu_f(\pi(s))\,\mathrm{d}s - \sum_{j=1}^{N(H)} \Delta_n^{j(1-2H)} \sum_{|\chi|=j} \frac{1}{\chi!}\int_0^t \partial^\chi\mu_f\left(\pi(s)\right)c(s)^\chi\,\mathrm{d}s\right\} \overset{\mathrm{st}}{\Longrightarrow} \mathcal{Z}, \tag{3.2.7}$$

where $\mathcal{Z} = (\mathcal{Z}_t)_{t\geq 0}$ is an $\mathbb{R}^M$-valued continuous process defined on a very good filtered extension $(\overline{\Omega}, \overline{\mathcal{F}}, (\overline{\mathcal{F}}_t)_{t\geq0}, \overline{\mathbb{P}})$ of the original probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq0}, \mathbb{P})$ which, conditionally on the σ-field $\mathcal{F}$, is a centered Gaussian process with independent increments and such that the covariance function $\mathcal{C}^{m_1m_2}_t = \overline{\mathbb{E}}[\mathcal{Z}^{m_1}_t\mathcal{Z}^{m_2}_t \mid \mathcal{F}]$, for $m_1, m_2 = 1,\dots,M$, is given by

$$\mathcal{C}^{m_1m_2}_t = \int_0^t \left\{\gamma_{f_{m_1},f_{m_2}}(\pi(s),\pi(s)) + \sum_{r=1}^{\infty}\left(\gamma_{f_{m_1},f_{m_2}} + \gamma_{f_{m_2},f_{m_1}}\right)(\pi(s),\pi_r(s))\right\}\mathrm{d}s. \tag{3.2.8}$$

[4]The authors also considered $f(x) = x^4$ but only obtained a law of large numbers without a CLT.

Let us make a few comments on the assumptions and the structure of this CLT.

Remark 3.2.2. From the proofs, one can see that it suffices to require f be $2(N(H)+1)$-times continuously differentiable with derivatives of at most polynomial growth. A decomposition as in (3.2.3) is standard for CLTs in high-frequency statistics. But here we need it for ρ (instead of σ) as the noise process dominates the efficient price process in the limit $\Delta_n \to 0$. Condition (3.2.2) on σ is satisfied if, for example, σ is itself a continuous Itô semimartingale. Finally, it should be possible to allow for an additional slowly varying (at 0) function as a factor in front of $t^{H-1/2}$ in (3.1.6) (cf. [14, 37]), but to simplify the exposition, we do not pursue this extra bit of generality.

Remark 3.2.3. Both the law of large numbers (LLN) limit

$$V_f(Y,t) = \int_0^t \mu_f(\pi(s))\,\mathrm{d}s \tag{3.2.9}$$

and the fluctuation process $\mathcal{Z}$ originate from the moving-average process Z. In other words, if $\sigma \equiv 0$ (i.e., in the pure fractional case), we would have (3.2.7) without the $\sum_{j=1}^{N(H)}$-expression; cf. [14, 37]. Even if $\sigma \not\equiv 0$, in the case where $H < \frac{1}{4}$, no additional terms are present because $N(H) = 0$. This is in line with Proposition 3.1.1, which states that it is impossible to consistently estimate $C_t = \int_0^t \sigma_s^2\,\mathrm{d}s$ if $H < \frac{1}{4}$. If $H \in (\frac{1}{4}, \frac{1}{2})$, the "mixed" terms in the $\sum_{j=1}^{N(H)}$-expression will allow us to estimate C_t.

Remark 3.2.4. Let us consider the special case where $d = 1$ and $f(x) = x^{2p}$ for some $p \in \mathbb{N}$. Then (3.2.7) reads

$$\Delta_n^{-\frac{1}{2}}\left\{V_f^n(Y,t) - V_f(Y,t) - \sum_{j=1}^{N(H)} \Delta_n^{j(1-2H)}\mu_{2p}\binom{p}{j}\int_0^t \rho_s^{2p-2j}\sigma_s^{2j}\,\mathrm{d}s\right\} \stackrel{\text{st}}{\Longrightarrow} \mathcal{Z},$$

where μ_{2p} is the moment of order $2p$ of a standard normal variable. Typically, one is interested in estimating only one of the terms in the sum $\sum_{j=1}^{N(H)}$ at a time (e.g., $\int_0^t \sigma_s^{2p}\,\mathrm{d}s$ corresponding to $j = p$). All other terms (e.g., $j \neq p$) have to be considered as higher-order bias terms in this case. The appearance of (potentially many, if $N(H)$ is large) bias terms for test functions as simple as powers of even order neither happens in the pure semimartingale nor in the pure fractional setting.

Remark 3.2.5. The following values for H are special:

$$\mathcal{H} = \left\{\frac{1}{2} - \frac{1}{4n} : n \geq 1\right\} = \left\{\frac{1}{4}, \frac{3}{8}, \frac{5}{12}, \frac{7}{16}, \ldots\right\}. \tag{3.2.10}$$

Indeed, if $H \in \mathcal{H}$, then $N(H) = 1/(2-4H)$. In particular, the term in (3.2.7) that corresponds to $j = N(H)$ is exactly of order $\Delta_n^{1/2}$. So in this case, (3.2.7) can also be viewed as convergence to a non-central mixed normal distribution.

3.2.2 Overview of the proof

In the following, we describe the main steps of the proof of Theorem 3.2.1 and defer the details to the supplementary material in [34]. By a standard localization argument (cf. [64, Lemma 4.4.9]), we may and will assume a strengthened version of Assumption (CLT):

Assumption (CLT'). In addition to Assumption (CLT), there is a constant $C > 0$ such that

$$\sup_{(\omega,t)\in\Omega\times[0,\infty)} \Big\{\|a_t(\omega)\| + \|\sigma_t(\omega)\| + \|\rho_t(\omega)\| + \|\rho^{(0)}_t(\omega)\| + \|\widetilde{b}_t(\omega)\| + \|\widetilde{\rho}_t(\omega)\|\Big\} < C. \qquad (3.2.11)$$

Moreover, for every $p > 0$, there is $C_p > 0$ such that for all $s, t > 0$,

$$\mathbb{E}[\|\sigma_t - \sigma_s\|^p]^{\frac{1}{p}} \leq C_p|t-s|^{\frac{1}{2}}, \quad \mathbb{E}[\|\rho^{(0)}_t - \rho^{(0)}_s\|^p]^{\frac{1}{p}} \leq C_p|t-s|^{\gamma}, \quad \mathbb{E}[\|\widetilde{\rho}_t - \widetilde{\rho}_s\|^p]^{\frac{1}{p}} \leq C_p|t-s|^{\varepsilon}. \qquad (3.2.12)$$

Proof of Theorem 3.2.1. Except for (3.2.17) below, we may and will assume that $M = 1$. Recalling the decomposition (3.1.6), since g_0 is smooth with $g_0(0) = 0$, we can use the stochastic Fubini theorem (see [86, Chapter IV, Theorem 65]) to write

$$\int_0^t g_0(t-r)\rho_r \,\mathrm{d}W_r = \int_0^t \left(\int_r^t g_0'(s-r)\,\mathrm{d}s\right)\rho_r \,\mathrm{d}W_r = \int_0^t \left(\int_0^s g_0'(s-r)\rho_r \,\mathrm{d}W_r\right)\mathrm{d}s.$$

This is a finite variation process and can be incorporated in the drift process in (3.1.8). So without loss of generality, we may assume $g_0 \equiv 0$ and $g(t) = K_H^{-1} t^{H-1/2}$ in the following. Then

$$Y_t = A_t + M_t + Z_t, \quad A_t = \int_0^t a_s \,\mathrm{d}s, \quad M_t = \int_0^t \sigma_s \,\mathrm{d}B_s,$$

and we have $\underline{\Delta}^n_i Y = \underline{\Delta}^n_i A + \underline{\Delta}^n_i M + \underline{\Delta}^n_i Z$ in the notation of (3.2.1). Writing $g(t) = 0$ for $t \leq 0$, we also define for all $s, t \geq 0$ and $i, n \in \mathbb{N}$,

$$\Delta^n_i g(s) = g(i\Delta_n - s) - g((i-1)\Delta_n - s), \qquad \underline{\Delta}^n_i g(s) = (\Delta^n_i g(s), \ldots, \Delta^n_{i+L-1} g(s)), \quad (3.2.13)$$

such that, in matrix notation,

$$\underline{\Delta}^n_i Z = \left(\int_0^\infty \Delta^n_i g(s)\rho_s \,\mathrm{d}W_s, , \ldots, \int_0^\infty \Delta^n_{i+L-1} g(s)\rho_s \,\mathrm{d}W_s\right) = \int_0^\infty \rho_s \,\mathrm{d}W_s \,\underline{\Delta}^n_i g(s).$$

As with most CLTs in high-frequency statistics (cf. [64, Chapters 5.2 and 5.3]), our proof is divided into two parts: a CLT based on centering with appropriate conditional expectations and the convergence of the latter to the LLN limit (3.2.9) after removing the asymptotic bias terms given by the sum over j in (3.2.7). The second part showcases how the mixed setting is different from both the pure fractional and pure semimartingale framework (in which no asymptotic bias terms are present). The first part is dominated by the fractional component in the sense that the contribution from the semimartingale part is asymptotically negligible. This, in fact, is rather surprising: if H is close to $\frac{1}{2}$, X is only slightly smoother than Z, so one would expect X to have a non-negligible contribution even if we center by conditional expectations. However, this is not the case and is related to the special structure of semimartingales. To deal with the fractional part, we could, in principle, use the methods employed by [14, 37], which involve Malliavin calculus and fractional calculus. However, one important hypothesis in both works is that ρ, the volatility of the noise, be strictly smoother than a continuous semimartingale. In order to accommodate the possibility of modeling ρ by a semimartingale (as in [3, 38, 62, 79], for example), we shall use an approach developed in [29, 27], which relies on martingale approximations for fractional processes and martingale CLTs. As the method was developed in the context of stochastic partial differential equations, we will give most details in the supplement for the reader's convenience.

The first step in our proof is to shrink the domain of integration for each $\underline{\Delta}^n_i Z$. Let

$$\frac{1}{4(1-H)} < \theta < \frac{1}{2}, \qquad (3.2.14)$$

which is always possible for $H \in (0, \frac{1}{2})$, and set $\theta_n = [\Delta_n^{-\theta}]$. Further define

$$\underline{\Delta}_i^n Y^{\mathrm{tr}} = \underline{\Delta}_i^n A + \underline{\Delta}_i^n M + \xi_i^n, \qquad \xi_i^n = \int_{(i-\theta_n)\Delta_n}^{(i+L-1)\Delta_n} \rho_s \, \mathrm{d}W_s \, \underline{\Delta}_i^n g(s). \tag{3.2.15}$$

Lemma 3.2.6. *If θ is chosen according to* (3.2.14)*, then*

$$\Delta_n^{-\frac{1}{2}} \left(V_f^n(Y,t) - \Delta_n \sum_{i=\theta_n+1}^{[t/\Delta_n]-L+1} f\left(\frac{\underline{\Delta}_i^n Y^{\mathrm{tr}}}{\Delta_n^H} \right) \right) \stackrel{L^1}{\Longrightarrow} 0.$$

The sum involving the truncated increments can be further decomposed into three parts:

$$\Delta_n^{\frac{1}{2}} \sum_{i=\theta_n+1}^{[t/\Delta_n]-L+1} f\left(\frac{\underline{\Delta}_i^n Y^{\mathrm{tr}}}{\Delta_n^H} \right) = V^n(t) + U^n(t) + \Delta_n^{\frac{1}{2}} \sum_{i=\theta_n+1}^{[t/\Delta_n]-L+1} \mathbb{E}\left[f\left(\frac{\underline{\Delta}_i^n Y^{\mathrm{tr}}}{\Delta_n^H} \right) \,\middle|\, \mathcal{F}_{i-\theta_n}^n \right], \tag{3.2.16}$$

where

$$V^n(t) = \sum_{i=\theta_n+1}^{[t/\Delta_n]-L+1} \Xi_i^n, \qquad \Xi_i^n = \Delta_n^{\frac{1}{2}} \left(f\left(\frac{\xi_i^n}{\Delta_n^H} \right) - \mathbb{E}\left[f\left(\frac{\xi_i^n}{\Delta_n^H} \right) \,\middle|\, \mathcal{F}_{i-\theta_n}^n \right] \right),$$

$$U^n(t) = \Delta_n^{\frac{1}{2}} \sum_{i=\theta_n+1}^{[t/\Delta_n]-L+1} \left\{ f\left(\frac{\underline{\Delta}_i^n Y^{\mathrm{tr}}}{\Delta_n^H} \right) - f\left(\frac{\xi_i^n}{\Delta_n^H} \right) - \mathbb{E}\left[f\left(\frac{\underline{\Delta}_i^n Y^{\mathrm{tr}}}{\Delta_n^H} \right) - f\left(\frac{\xi_i^n}{\Delta_n^H} \right) \,\middle|\, \mathcal{F}_{i-\theta_n}^n \right] \right\}.$$

Lemma 3.2.7. *For all $H < \frac{1}{2}$, we have that $U^n \stackrel{L^1}{\Longrightarrow} 0$.*

In other words, in the limit $\Delta_n \to 0$, the impact of the semimartingale component is negligible, except for its contributions to the conditional expectations in (3.2.16). As we mentioned above, this is somewhat surprising: It is true that L^2-norm of the semimartingale increment $\underline{\Delta}_i^n A + \underline{\Delta}_i^n M$, divided by Δ_n^H, converges to 0. But the rate $\Delta_n^{1/2-H}$ at which this takes place can be arbitrarily slow if H is close to $\frac{1}{2}$. So Lemma 3.2.7 implies that there is a big gain in convergence rate if one considers the sum of the centered differences $f(\underline{\Delta}_i^n Y^{\mathrm{tr}}/\Delta_n^H) - f(\xi_i^n/\Delta_n^H)$. In the proof, we will need for the first time that f has at least $2(N(H)+1)$ continuous derivatives.

The process V^n only contains the fractional part and is responsible for the limit $\mathcal{Z}$ in (3.2.7). For the sake of brevity, we borrow a result from [29]: For each $m \in \mathbb{N}$, consider the sums

$$V^{n,m,1}(t) = \sum_{j=1}^{J^{n,m}(t)} V_j^{n,m}, \qquad V_j^{n,m} = \sum_{k=1}^{m\theta_n} \Xi_{(j-1)((m+1)\theta_n+L-1)+k}^n,$$

$$V^{n,m,2}(t) = \sum_{j=1}^{J^{n,m}(t)} \sum_{k=1}^{\theta_n+L-1} \Xi_{(j-1)((m+1)\theta_n+L-1)+m\theta_n+k}^n,$$

$$V^{n,m,3}(t) = \sum_{j=((m+1)\theta_n+L-1)J^{n,m}(t)+1}^{[t/\Delta_n]-L+1} \Xi_j^n,$$

where $J^{n,m}(t) = [([t/\Delta_n] - L + 1)/((m+1)\theta_n + L - 1)]$. We then have $V^n(t) = \sum_{i=1}^3 V^{n,m,i}(t)$. This is very similar to the decomposition in [29, Section 3.2, p. 1161]. With only minimal changes (cf. Lemma 3.9, Lemma 3.10 and Proposition 3.11 in [29]), we infer that $V^n(t) \stackrel{\mathrm{st}}{\Longrightarrow} \mathcal{Z}$ and, hence,

$$\Delta_n^{-\frac{1}{2}} \left\{ \Delta_n \sum_{i=1}^{[t/\Delta_n]-L+1} f\left(\frac{\underline{\Delta}_i^n Y}{\Delta_n^H} \right) - \Delta_n \sum_{i=\theta_n+1}^{[t/\Delta_n]-L+1} \mathbb{E}\left[f\left(\frac{\underline{\Delta}_i^n Y^{\mathrm{tr}}}{\Delta_n^H} \right) \,\middle|\, \mathcal{F}_{i-\theta_n}^n \right] \right\} \stackrel{\mathrm{st}}{\Longrightarrow} \mathcal{Z}, \tag{3.2.17}$$

where $\mathcal{Z}$ is exactly as in (3.2.7). Therefore, in order to complete the proof of Theorem 3.2.1, it remains to show that (recall $N(H) = [1/(2-4H)]$)

$$\Delta_n^{-\frac{1}{2}}\Bigg\{\Delta_n \sum_{i=\theta_n+1}^{[t/\Delta_n]-L+1} \mathbb{E}\left[f\left(\frac{\underline{\Delta}_i^n Y^{\mathrm{tr}}}{\Delta_n^H}\right) \,\middle|\, \mathcal{F}_{i-\theta_n}^n\right] - \int_0^t \mu_f(\pi(s))\,\mathrm{d}s - \sum_{j=1}^{N(H)} \Delta_n^{j(1-2H)} \sum_{|\chi|=j} \frac{1}{\chi!}\int_0^t \partial^\chi \mu_f\left(\pi(s)\right) c(s)^\chi\,\mathrm{d}s\Bigg\} \overset{L^1}{\Longrightarrow} 0.$$

To that end, we will discretize ("freeze") the volatility processes σ and ρ in $\underline{\Delta}_i^n Y^{\mathrm{tr}}$. The proof is technical and will be divided into further smaller steps in the supplement.

Lemma 3.2.8. *Assuming* (3.2.14)*, we have that*

$$\Delta_n^{\frac{1}{2}} \sum_{i=\theta_n+1}^{[t/\Delta_n]-L+1} \left\{\mathbb{E}\left[f\left(\frac{\underline{\Delta}_i^n Y^{\mathrm{tr}}}{\Delta_n^H}\right) \,\middle|\, \mathcal{F}_{i-\theta_n}^n\right] - \mu_f(\Upsilon^{n,i})\right\} \overset{L^1}{\Longrightarrow} 0,$$

where $\Upsilon^{n,i} \in (\mathbb{R}^{d\times L})^2$ *is defined by*

$$(\Upsilon^{n,i})_{k\ell,k'\ell'} = c((i-1)\Delta_n)_{k\ell,k'\ell'}\,\Delta_n^{1-2H} + \left(\rho_{(i-1)\Delta_n}\rho_{(i-1)\Delta_n}^T\right)_{kk'} \int_{(i-\theta_n)\Delta_n}^{(i+L-1)\Delta_n} \frac{\Delta_{i+\ell-1}^n g(s)\Delta_{i+\ell'-1}^n g(s)}{\Delta_n^{2H}}\,\mathrm{d}s. \tag{3.2.18}$$

The last part of the proof consists of evaluating

$$\Delta_n^{\frac{1}{2}} \sum_{i=\theta_n+1}^{[t/\Delta_n]-L+1} \mu_f(\Upsilon^{n,i}).$$

This is the place where the asymptotic bias terms arise and which is different from the pure (semimartingale or fractional) cases. Roughly speaking, the additional terms are due to fact that in the LLN limit (3.2.9), there is a contribution of magnitude $\Delta_n^{1-2H}c(s)$ coming from the semimartingale part that is negligible on first order but not at a rate of $\sqrt{\Delta_n}$. Expanding $\mu_f(\Upsilon^{n,i})$ in a Taylor sum around the point $\pi((i-1)\Delta_n)$ up to order $N(H)$, we obtain

$$\mu_f(\Upsilon^{n,i}) = \mu_f(\pi((i-1)\Delta_n)) + \sum_{j=1}^{N(H)} \sum_{|\chi|=j} \frac{1}{\chi!}\partial^\chi \mu_f\left(\pi((i-1)\Delta_n)\right)\left(\Upsilon^{n,i} - \pi((i-1)\Delta_n)\right)^\chi + \sum_{|\chi|=N(H)+1} \frac{1}{\chi!}\partial^\chi \mu_f(\upsilon_i^n)(\Upsilon^{n,i} - \pi((i-1)\Delta_n))^\chi,$$

where υ_i^n is a point between $\Upsilon^{n,i}$ and $\pi((i-1)\Delta_n)$. The next lemma shows two things: first, the term of order $N(H)+1$ is negligible, and second, for $j = 1,\ldots,N(H)$, we may replace $\Upsilon^{n,i} - \pi((i-1)\Delta_n)$ by $\Delta_n^{1-2H}c((i-1)\Delta_n)$.

Lemma 3.2.9. *Recall that* $N(H) = [1/(2-4H)]$*. We have that* $\mathbb{X}_1^n \overset{L^1}{\Longrightarrow} 0$ *and* $\mathbb{X}_2^n \overset{L^1}{\Longrightarrow} 0$*, where*

$$\mathbb{X}_1^n(t) = \Delta_n^{\frac{1}{2}} \sum_{i=\theta_n+1}^{[t/\Delta_n-L+1} \sum_{j=1}^{N(H)} \sum_{|\chi|=j} \frac{1}{\chi!}\partial^\chi \mu_f(\pi((i-1)\Delta_n)) \times \left\{(\Upsilon^{n,i} - \pi((i-1)\Delta_n))^\chi - \Delta_n^{j(1-2H)}c((i-1)\Delta_n)^\chi\right\}, \tag{3.2.19}$$

$$\mathbb{X}_2^n(t) = \Delta_n^{\frac{1}{2}} \sum_{i=\theta_n+1}^{[t/\Delta_n]-L+1} \sum_{|\chi|=N(H)+1} \frac{1}{\chi!}\partial^\chi \mu_f(\upsilon_i^n)(\Upsilon^{n,i} - \pi((i-1)\Delta_n))^\chi.$$

In a final step, we remove the discretization of σ and ρ.

Lemma 3.2.10. *If θ is chosen according to* (3.2.14)*, then*

$$\Delta_n^{-\frac{1}{2}}\left\{\Delta_n \sum_{i=\lambda_n+1}^{[t/\Delta_n]-L+1} \mu_f(\pi((i-1)\Delta_n)) - \int_0^t \mu_f(\pi(s))\,\mathrm{d}s\right\} \xrightarrow{L^1} 0 \tag{3.2.20}$$

and

$$\begin{aligned}\Delta_n^{-\frac{1}{2}} &\left\{\Delta_n \sum_{i=\theta_n+1}^{[t/\Delta_n]-L+1} \sum_{j=1}^{N(H)} \Delta_n^{j(1-2H)} \sum_{|\chi|=j} \frac{1}{\chi!}\partial^\chi \mu_f\left(\pi((i-1)\Delta_n)\right) c((i-1)\Delta_n)^\chi \right.\\ &\left. -\int_0^t \sum_{j=1}^{N(H)} \sum_{|\chi|=j} \frac{1}{\chi!}\partial^\chi \mu_f\left(\pi(s)\right) \Delta_n^{j(1-2H)} c(s)^\chi\,\mathrm{d}s\right\} \xrightarrow{L^1} 0.\end{aligned} \tag{3.2.21}$$

By the properties of stable convergence in law (see, for example, [64, Equation (2.2.5)]), the CLT in (3.2.7) follows by combining Lemmas 3.2.6–3.2.10. □

3.3 Estimating Hurst index and integrated volatility

In this section, we assume $d = 1$ for simplicity. We will develop an estimation procedure for H and the integrated volatilities of the efficient price (if $H > \frac{1}{4}$) and of the noise process, that is, for

$$C_t = \int_0^t \sigma_s^2\,\mathrm{d}s, \qquad \Pi_t = \int_0^t \rho_s^2\,\mathrm{d}s.$$

To avoid additional bias terms (cf. Remark 3.2.4), we use quadratic functionals only, that is, we consider

$$f_r(x) = x_1 x_{r+1}, \quad x = (x_1, \ldots, x_{r+1}) \in \mathbb{R}^{r+1}, \quad r \in \mathbb{N}_0,$$

and the associated variation functionals

$$V_{r,t}^n = V_{f_r}^n(Y,t) = \Delta_n \sum_{k=1}^{[t/\Delta_n]-r} \frac{\Delta_k^n Y \Delta_{k+r}^n Y}{\Delta_n^{2H}}.$$

Note that $V_{r,t}^n$ is not a statistic as it depends on the unknown Hurst parameter H. Therefore, we introduce $\widehat{V}_t^n = (\widehat{V}_{0,t}^n, \ldots, \widehat{V}_{R,t}^n)$, a non-normalized version of $V_{r,t}^n$ that is a statistic:

$$\widehat{V}_{r,t}^n = \widehat{V}_{f_r}^n(Y,t) = \sum_{k=1}^{[t/\Delta_n]-r} \Delta_k^n Y \Delta_{k+r}^n Y, \qquad r \in \mathbb{N}_0.$$

Clearly, $\Delta_n^{1-2H}\widehat{V}_{r,t}^n = V_{r,t}^n$, so our main CLT (Theorem 3.2.1) immediately yields:

Corollary 3.3.1. *Let $\widehat{V}_t^n = (\widehat{V}_{0,t}^n, \ldots, \widehat{V}_{R,t}^n)$ for a fixed but arbitrary $R \in \mathbb{N}_0$. For all $H \in (0, \frac{1}{2})$,*

$$\Delta_n^{-\frac{1}{2}}\left\{\Delta_n^{1-2H}\widehat{V}_t^n - \Gamma^H \int_0^t \rho_s^2\,\mathrm{d}s - e_1 \int_0^t \sigma_s^2\,\mathrm{d}s\, \Delta_n^{1-2H}\, \mathbb{1}_{[\frac{1}{4},\frac{1}{2})}(H)\right\} \xrightarrow{\mathrm{st}} \mathcal{Z}, \tag{3.3.1}$$

where $\Gamma^H = (\Gamma_0^H, \ldots, \Gamma_R^H)$, $e_1 = (1, 0, \ldots, 0) \in \mathbb{R}^{1+R}$ and $\mathcal{Z}$ is as in (3.2.7)*. The covariance process $\mathcal{C}^H(t) = (\mathcal{C}_{ij}^H(t))_{i,j=0,\ldots,R}$ in* (3.2.8) *is given by*

$$\begin{aligned}\mathcal{C}_{ij}^H(t) &= \mathcal{C}_{ij}^H \int_0^t \rho_s^4\,\mathrm{d}s,\\ \mathcal{C}_{ij}^H &= \Gamma_{|i-j|}^H + \Gamma_i^H\Gamma_j^H + \sum_{r=1}^\infty \left(\Gamma_r^H\Gamma_{|i-j+r|}^H + \Gamma_{|r-j|}^H\Gamma_{i+r}^H + \Gamma_r^H\Gamma_{|j-i+r|}^H + \Gamma_{|r-i|}^H\Gamma_{j+r}^H\right).\end{aligned} \tag{3.3.2}$$

As we can see, if $H \in (\frac{1}{4}, \frac{1}{2})$, only RV ($r = 0$) contains information about $C_t = \int_0^t \sigma_s^2 \, \mathrm{d}s$. But to first order, $V_{0,t}^n = \Delta_n^{1-2H} \widehat{V}_{0,t}^n$ estimates $\Pi_t = \int_0^t \rho_s^2 \, \mathrm{d}s$, the integrated noise volatility. In order to obtain C_t, our strategy is to use $\widehat{V}_{r,t}^n$ for $r \geq 1$ to remove the first-order limit of $\widehat{V}_{0,t}^n$. But here is a problem: both the factor Δ_n^{1-2H} and the vector Γ^H contain the unknown Hurst index H, so we need to estimate H first.

The most obvious estimator for H is the one obtained from regressing $\log \Delta_n$ on $\log \widehat{V}_{0,t}^n$. This is, in fact, what we did in Figure 3.1 and what *volatility signature plots* (i.e., plots of $\widehat{V}_{0,t}^n$ as a function of Δ_n; see [5, 8]) are based on. We also refer to [89] for a more general but related concept. However, as noted by [47, Remark 3.1], already in an mfBM model, this regression based estimator only has a logarithmic rate of convergence. Indeed, as our simulation study in Section 3.4 shows, this estimator systematically overestimates H unless H is very close to 0 or $\frac{1}{2}$. In the pure fractional case, better and sometimes even rate-optimal estimators are given by the so-called change-of-frequency estimators or autocorrelation estimators (see [14, 37] and the references therein as well as [29, 47]). Both types of estimators extract information about H by considering the ratio of (different combinations of) $\widehat{V}_{r,t}^n$ for different values of r. For example, the simplest autocorrelation estimator is

$$\widetilde{H}_{\mathrm{acf}}^n = \frac{1}{2} \left[1 + \log_2 \left(\frac{\widehat{V}_{1,t}^n}{\widehat{V}_{0,t}^n} + 1 \right) \right], \tag{3.3.3}$$

which is based on the fact that $\widehat{V}_{1,t}^n / \widehat{V}_{0,t}^n = V_{1,t}^n / V_{0,t}^n \overset{P}{\to} \Gamma_1^H = 2^{2H-1} - 1$. While this estimator converges faster than logarithmically, the convergence rate is still not optimal because of the bias term that appears in (3.3.1) when $r = 0$. The first rate-optimal estimator for H in the case of mfBM was constructed in [47, Theorem 3.2] by using a variant of (3.3.3) that cancels out the contribution from $\widehat{V}_{0,t}^n$. However, this estimator suffers from a potentially large constant in the asymptotic variance. For example, in [47, Remark 3.2], the authors do not recommend using it in practice even though it has a better convergence rate than the regression-based estimator. This point is further confirmed in the simulations in Section 3.4.

To do better, our strategy is to use linear combinations of $\widehat{V}_{r,t}^n$ for multiple values of r. To this end, we choose two weight vectors $a = a(R) = (a_0, \ldots, a_R)$ and $b = b(R) = (b_0, \ldots, b_R)$ in $\mathbb{R}^{1+R}$ and consider the statistic

$$\widetilde{H}^n = \phi^{-1} \left(\frac{\langle a, \widehat{V}_t^n \rangle}{\langle b, \widehat{V}_t^n \rangle} \right) \qquad \text{with} \qquad \phi(H) = \frac{\langle a, \Gamma^H \rangle}{\langle b, \Gamma^H \rangle}, \tag{3.3.4}$$

where $\langle \cdot, \cdot \rangle$ denotes the standard inner product on $\mathbb{R}^{1+R}$ and a and b are assumed to be such that ϕ is invertible. The further analysis is now dependent on whether $H \in (0, \frac{1}{4})$ or $H \in (\frac{1}{4}, \frac{1}{2})$ and, in the latter case, whether $a_0 = b_0 = 0$ or at least one of a_0 and b_0 is not zero.

3.3.1 Estimation without quadratic variation or if $H \in (0, \frac{1}{4})$

If $a_0 = b_0 = 0$, we exclude quadratic variation from our estimation procedure for H. This has the advantage that the term $e_1 \int_0^t \sigma_s^2 \, \mathrm{d}s \, \Delta_n^{1-2H}$ in (3.3.1), which is only nonzero for $r = 0$, disappears. The same holds true (even for $r = 0$) if $H < \frac{1}{4}$: there is no asymptotic drift term in (3.3.1).

Theorem 3.3.2. *Assume that* $H \in (0, \frac{1}{2})$ *and choose* $R \in \mathbb{N}$ *and* $a, b \in \mathbb{R}^{1+R}$ *such that* ϕ *from* (3.3.4) *is invertible. If* $H \in (\frac{1}{4}, \frac{1}{2})$*, further assume that* $a_0 = b_0 = 0$*.*

(1) The estimator $\widetilde{H}^n$ *introduced in* (3.3.4) *satisfies*

$$\Delta_n^{-\frac{1}{2}}(\widetilde{H}^n - H) \xrightarrow{\text{st}} \mathcal{N}\left(0, \mathrm{Var}_{H,0}\frac{\int_0^t \rho_s^4\,\mathrm{d}s}{(\int_0^t \rho_s^2\,\mathrm{d}s)^2}\right), \tag{3.3.5}$$

where $\mathcal{Z}$ *is the same as in* (3.3.1) *and*

$$\mathrm{Var}_{H,0} = \mathrm{Var}_{H,0}(R,a,b,H) = \left(\frac{(\phi^{-1})'(\phi(H))}{\langle b, \Gamma^H\rangle}\right)^2 \{a^T - \phi(H)b^T\}\mathcal{C}^H\{a - \phi(H)b\}. \tag{3.3.6}$$

(2) If $H \in (\frac{1}{4}, \frac{1}{2})$*, choose* $c \in \mathbb{R}^{1+R}$ *and define*

$$\widehat{C}_t^n = \left\{\widehat{V}_{0,t}^n - \frac{\langle c, \widehat{V}_t^n\rangle}{\langle c, \Gamma^{\widetilde{H}^n}\rangle}\right\}\left(1 - \frac{c_0}{\langle c, \Gamma^{\widetilde{H}^n}\rangle}\right)^{-1}. \tag{3.3.7}$$

Then

$$\Delta_n^{\frac{1}{2}-2H}\left\{\widehat{C}_t^n - C_t\right\} \xrightarrow{\text{st}} \mathcal{N}\left(0, \mathrm{Var}_C \int_0^t \rho_s^4\,\mathrm{d}s\right), \tag{3.3.8}$$

where

$$\begin{aligned} \mathrm{Var}_C &= \mathrm{Var}_C(R,a,b,c,H) = u^T\mathcal{C}^H u,\\ u &= \left(e_1 - \frac{c}{\langle c, \Gamma^H\rangle} + \frac{\langle c, \partial_H\Gamma^H\rangle}{\langle c, \Gamma^H\rangle}\frac{(\phi^{-1})'(\phi(H))}{\langle b, \Gamma^H\rangle}(a - \phi(H)b)\right)\left(1 - \frac{c_0}{\langle c, \Gamma^H\rangle}\right)^{-1} \end{aligned} \tag{3.3.9}$$

and the vector $\partial_H\Gamma^H = (\partial_H\Gamma_0^H, \dots, \partial_H\Gamma_R^H)$ *is given by*

$$\partial_H\Gamma_0^H = 0, \quad \partial_H\Gamma_r^H = \log(r+1)(r+1)^{2H} - 2\log(r)r^{2H} + \log(r-1)(r-1)^{2H}, \quad r \geq 1. \tag{3.3.10}$$

(3) The estimator

$$\widehat{\Pi}_t^n = \Delta_n^{1-2\widetilde{H}^n}\frac{\langle a, \widehat{V}_t^n\rangle}{\langle a, \Gamma^{\widetilde{H}^n}\rangle} \tag{3.3.11}$$

satisfies

$$\frac{\Delta_n^{-\frac{1}{2}}}{|\log\Delta_n|}(\widehat{\Pi}_t^n - \Pi_t) \xrightarrow{\text{st}} \mathcal{N}\left(0, 4\mathrm{Var}_{H,0}\int_0^t \rho_s^4\,\mathrm{d}s\right). \tag{3.3.12}$$

Remark 3.3.3. To construct the estimator $\widehat{C}_t^n$, we allow for the possibility of choosing a new weight vector c. Therefore, a and b should be thought of as weights that one can choose to, for example, minimize $\mathrm{Var}_{H,0}(R,a,b,H)$, while c can then be chosen to minimize $\mathrm{Var}_C(R,a,b,c,H)$. Alternatively, one may decide to choose a, b and c to minimize $\mathrm{Var}_C(R,a,b,c,H)$ directly (if $H > \frac{1}{4}$).

In order to obtain feasible CLTs, we replace the unknown quantities in $\mathrm{Var}_{H,0}$ and Var_C by consistent estimators thereof. To this end, consider $f(x) = x^4$ and

$$Q_t^n = V_f^n(Y,t) = \Delta_n\sum_{i=1}^{[t/\Delta_n]}\left|\frac{\Delta_i^n Y}{\Delta_n^H}\right|^4, \qquad \widehat{Q}_t^n = \sum_{i=1}^{[t/\Delta_n]}(\Delta_i^n Y)^4.$$

By Theorem 3.2.1, we have the LLN

$$Q_t^n \xrightarrow{L^1} 3\int_0^t \rho_s^4\,\mathrm{d}s. \tag{3.3.13}$$

Therefore, the following theorem is a direct consequence of Theorem 3.3.2 and well-known properties of stable convergence in law (see [64, Equation (2.2.5)]).

Theorem 3.3.4. *Grant the assumptions of Theorem 3.3.2. For* (3.3.15) *below, further assume that* $H \in (\frac{1}{4}, \frac{1}{2})$. *Then*

$$\Delta_n^{-\frac{1}{2}}(\widetilde{H}^n - H)\sqrt{\frac{3\Delta_n(\widehat{V}^n_{0,t})^2}{\mathrm{Var}_{H,0}(R,a,b,\widetilde{H}^n)\widehat{Q}^n_t}} \xrightarrow{\mathrm{st}} \mathcal{N}(0,1), \tag{3.3.14}$$

$$\Delta_n^{\frac{1}{2}-2\widetilde{H}^n}(\widehat{C}^n_t - C_t)\sqrt{\frac{3\Delta_n^{4\widetilde{H}^n-1}}{\mathrm{Var}_C(R,a,b,c,\widetilde{H}^n)\widehat{Q}^n_t}} \xrightarrow{\mathrm{st}} \mathcal{N}(0,1), \tag{3.3.15}$$

$$\frac{\Delta_n^{-\frac{1}{2}}}{|\log\Delta_n|}(\widehat{\Pi}^n_t - \Pi_t)\sqrt{\frac{3\Delta_n^{4\widetilde{H}^n-1}}{4\mathrm{Var}_{H,0}(R,a,b,\widetilde{H}^n)\widehat{Q}^n_t}} \xrightarrow{\mathrm{st}} \mathcal{N}(0,1). \tag{3.3.16}$$

3.3.2 Estimation with quadratic variation if $H \in (\frac{1}{4}, \frac{1}{2})$

The estimators based on weight vectors a and b with $a_0 = b_0 = 0$ were easy to construct but suffer from a serious shortcoming: Let us consider the simple case where $Y = \sigma B$ for some constant $\sigma > 0$ (i.e., there is no noise). Then, by standard CLTs for Brownian motion, the ratio $\langle a, \widehat{V}^n_t\rangle / \langle b, \widehat{V}^n_t\rangle$ converges stably in law to the ratio Z_1/Z_2 of two centered (possibly correlated) normals that are independent of B. In particular, because Z_1/Z_2 has a density supported on $\mathbb{R}$, the asymptotic probability that the estimator $\widetilde{H}^n$ from (3.3.4) falls into any nonempty open subinterval of $(0,1)$ is nonzero. Therefore, based on $\widetilde{H}^n$ only, it is impossible to tell whether there is evidence for the presence of fractional noise or whether an estimate produced by $\widetilde{H}^n$ is simply the result of chance!

To solve this problem, we have to include lag 0 in our estimation of H. If $H \in (\frac{1}{4}, \frac{1}{2})$, this significantly complicates the estimation procedure: By the discussion at the beginning of Section 3.3, in order to estimate C_t, we need to estimate H first. At the same time, as Corollary 3.3.1 shows, using $\widehat{V}^n_{0,t}$ to estimate H induces an asymptotic bias term coming from the $\int_0^t \sigma_s^2\,\mathrm{d}s$ term, which can only be corrected with an estimator of C_t. In other words, in order to estimate C_t, we need to first estimate H, but in order to (precisely) estimate H, we need to estimate C_t. Resolving this circular dependence necessitates a complex iterated estimation procedure for H and C_t that we describe in the following. In particular, as $H \uparrow \frac{1}{2}$, even though there is only one intermediate limit $e_1 C_t \Delta_n^{1-2H}$ between the LLN limit $\Gamma^H \Pi_t$ and the CLT, we obtain an increasing number of higher-order bias terms as a result of the interdependence between the H- and the C_t-estimators.

A consistent but not asymptotically normal estimator of H

To simplify the exposition, we assume that at least one of a_0 and b_0 is zero. By symmetry, we shall consider the case where

$$a_0 \neq 0, \qquad b_0 = 0. \tag{3.3.17}$$

Also, again to simplify the argument and because this is not really a severe restriction from a statistical point of view, we shall assume that the true value of H satisfies

$$H \in (\tfrac{1}{4}, \tfrac{1}{2}) \setminus \mathcal{H}, \tag{3.3.18}$$

where $\mathcal{H}$ is the set from (3.2.10).

Proposition 3.3.5. *Let $H \in (\frac{1}{4}, \frac{1}{2}) \setminus \mathcal{H}$ and suppose that $a, b \in \mathbb{R}^{1+R}$ satisfy* (3.3.17) *and are such that ϕ from* (3.3.4) *is invertible. Recalling that $N(H) = [1/(2-4H)]$, we further define*

$$\Phi_j^n = \Phi_j^n(R, a, b, \widehat{V}_t^n, \widetilde{H}^n) = \frac{(-1)^j}{j!}(\phi^{-1})^{(j)}(\phi(\widetilde{H}^n))\frac{a_0^j}{\langle b, \widehat{V}_t^n\rangle^j}, \qquad j = 1, \dots, N(H). \tag{3.3.19}$$

Then $\widetilde{H}^n$, as defined in (3.3.4)*, satisfies*

$$\Delta_n^{-\frac{1}{2}}\left\{\widetilde{H}^n - H + \sum_{j=1}^{N(H)} \Phi_j^n C_t^j\right\} \xrightarrow{\text{st}} \mathcal{N}\left(0, \mathrm{Var}_{H,0}\frac{\int_0^t \rho_s^4\,\mathrm{d}s}{(\int_0^t \rho_s^2\,\mathrm{d}s)^2}\right), \tag{3.3.20}$$

where $\mathrm{Var}_{H,0}$ *is defined in* (3.3.6)*.*

For each j, the term Φ_j^n is of order $\Delta_n^{j(1-2H)}$. As a result, while $\widetilde{H}^n$ is consistent for H, it is affected by many higher-order asymptotic bias terms that depend on C_t. So our next goal is to find consistent estimators of C_t that we can use to correct H.

A consistent but not asymptotically normal estimator of C_t

With a first estimator of H at hand, we can now construct an estimator of C_t by removing the first-order limit of $\widehat{V}_{0,t}^n$, hereby replacing H by $\widetilde{H}^n$ throughout. Doing so, we have to employ an estimator of Π_t, the integrated noise volatility. To avoid even more higher-order bias terms, we need one with convergence rate $\sqrt{\Delta_n}$. One possibility is to use the estimator from Theorem 3.3.2, constructed from an additional pair of weights a^0 and b^0 with $a_0^0 = b_0^0 = 0$. Note that even in the case where $\rho = 0$, the estimator $\widehat{\Pi}_t^n$ from Theorem 3.3.2, in contrast to $\widetilde{H}^n$, converges to the desired limit (i.e., 0) in probability.

Proposition 3.3.6. *In addition to $a, b \in \mathbb{R}^{1+R}$ satisfying* (3.3.17)*, choose $a^0, b^0 \in \mathbb{R}^{1+R}$ with $a_0^0 = b_0^0 = 0$ and let*

$$\widehat{P}_t^n = \frac{\langle a^0, \widehat{V}_t^n\rangle}{\langle a^0, \Gamma^{\widetilde{H}^{n,0}}\rangle}, \qquad \widetilde{H}^{n,0} = \varphi^{-1}\left(\frac{\langle a^0, \widehat{V}_t^n\rangle}{\langle b^0, \widehat{V}_t^n\rangle}\right). \tag{3.3.21}$$

Further define

$$\widetilde{C}_t^{n,1} = \left\{\widehat{V}_{0,t}^n - \frac{\langle a, \widehat{V}_t^n\rangle}{\langle a, \Gamma^{\widetilde{H}^n}\rangle}\right\}\Theta(\widehat{V}_t^n, \widetilde{H}^n, \widetilde{H}^{n,0})^{-1} \tag{3.3.22}$$

where

$$\Theta(\widehat{V}_t^n, \widetilde{H}^n, \widetilde{H}^{n,0}) = \Theta(R, a, b, a^0, b^0, \widehat{V}_t^n, \widetilde{H}^n, \widetilde{H}^{n,0}) = 1 - \frac{a_0}{\langle a, \Gamma^{\widetilde{H}^n}\rangle} + \frac{\widehat{P}_t^n}{\langle b, \widehat{V}_t^n\rangle}\frac{a_0\psi'(\phi(\widetilde{H}^n))}{\langle a, \Gamma^{\widetilde{H}^n}\rangle} \tag{3.3.23}$$

and

$$\psi(y) = \langle a, \Gamma^{\phi^{-1}(y)}\rangle, \qquad y \in \mathbb{R}. \tag{3.3.24}$$

Then, under the assumptions made in Proposition 3.3.5,

$$\Delta_n^{\frac{1}{2}-2H}\left\{\widetilde{C}_t^{n,1} - C_t + \sum_{j=2}^{N(H)} \Psi_j^n C_t^j\right\} \xrightarrow{\text{st}} \mathcal{N}\left(0, \mathrm{Var}_{C,1}\int_0^t \rho_s^4\,\mathrm{d}s\right), \tag{3.3.25}$$

where

$$\begin{aligned}\Psi_j^n &= \Psi_j^n(R,a,b,a^0,b^0,\widehat{V}_t^n,\widetilde{H}^n,\widetilde{H}^{n,0}) \\ &= \frac{(-1)^j}{j!}\psi^{(j)}(\phi(\widetilde{H}^n))\frac{a_0^j}{\langle a,\Gamma^{\widetilde{H}^n}\rangle}\frac{\widehat{P}_t^n}{\langle b,\widehat{V}_t^n\rangle^j}\Theta(\widehat{V}_t^n,\widetilde{H}^n,\widetilde{H}^{n,0})^{-1}\end{aligned} \tag{3.3.26}$$

for $j=2,\dots,N(H)$ *and*

$$\begin{aligned}&\mathrm{Var}_{C,1} = \mathrm{Var}_{C,1}(R,a,b,H) = u_1^T\mathcal{C}^H u_1, \qquad (3.3.27)\\ &u_1 = \left(e_1 - \frac{a}{\langle a,\Gamma^H\rangle} + \frac{\psi'(\phi(H))}{\langle a,\Gamma^H\rangle\langle b,\Gamma^H\rangle}(a-\phi(H)b)\right)\left(1-\frac{a_0}{\langle a,\Gamma^H\rangle}+\frac{\psi'(\phi(H))}{\langle a,\Gamma^H\rangle\langle b,\Gamma^H\rangle}a_0\right)^{-1},\end{aligned}$$

and $\mathcal{C}^H$ *is the matrix in* (3.3.2).

Note that Ψ_j^n is of magnitude $\Delta_n^{(j-1)(1-2H)}$. Thus, just as for the initial estimator of H, the estimator $\widetilde{C}_t^{n,1}$ is consistent but has higher-order bias terms.

The first asymptotically normal estimators of H and C_t

What is different between the two initial estimators of H and C_t is that in (3.3.25) the bias terms only hinge on C_t, the quantity that $\widetilde{C}_t^{n,1}$ is supposed to estimate in the first place. Therefore, we can set up an iteration procedure and use $\widetilde{C}_t^{n,1}$ to correct itself.

Proposition 3.3.7. *Recall that* $N(H)=[1/(2-4H)]$ *and define*

$$\widetilde{C}_t^{n,\ell+1} = \widetilde{C}_t^{n,1} + \sum_{j=2}^{\ell+1}\Psi_j^n\left(\widetilde{C}_t^{n,\ell-j+2}\right)^j, \qquad \ell\geq 0, \tag{3.3.28}$$

and

$$\widehat{C}_t^{n,1} = \widetilde{C}_t^{n,N(\widetilde{H}^n)}. \tag{3.3.29}$$

Then we have that

$$\Delta_n^{\frac{1}{2}-2H}\left\{\widetilde{C}_t^{n,1} - C_t + \sum_{j=2}^{N(\widetilde{H}^n)}\Psi_j^n\left(\widetilde{C}_t^{n,N(\widetilde{H}^n)-j+1}\right)^j\right\} = \Delta_n^{\frac{1}{2}-2H}(\widehat{C}_t^{n,1}-C_t) \xrightarrow{\mathrm{st}} \mathcal{N}\left(0,\mathrm{Var}_{C,1}\int_0^t\rho_s^4\,\mathrm{d}s\right)$$

with the same $\mathrm{Var}_{C,1}$ *as in* (3.3.27).

The corrected estimator $\widehat{C}_t^{n,1}$ is our first consistent and asymptotically mixed normal estimator for C_t and has a convergence rate of $\Delta_n^{2H-1/2}$. According to work in progress by F. Mies[5], this rate is optimal. With a bias-free estimator of C_t at hand, we can now proceed to correcting the initial estimator $\widetilde{H}^n$ of H.

Proposition 3.3.8. *Recall* $\widetilde{H}^n$ *in* (3.3.4) *and define*

$$\widehat{H}_1^n = \widetilde{H}^n + \sum_{j=1}^{N(\widetilde{H}^n)}\Phi_j^n\left(\widehat{C}_t^{n,1}\right)^j \tag{3.3.30}$$

[5]Private communication.

with Φ_j^n *as in* (3.3.19)*. We have the central limit theorem*

$$\Delta_n^{-\frac{1}{2}}(\widehat{H}_1^n - H) \xrightarrow{\text{st}} \mathcal{N}\left(0, \operatorname{Var}_{H,1}\frac{\int_0^t \rho_s^4\,\mathrm{d}s}{(\int_0^t \rho_s^2\,\mathrm{d}s)^2}\right),$$

where

$$\operatorname{Var}_{H,1} = \operatorname{Var}_{H,1}(R,a,b,H) = w_1^T\mathcal{C}^H w_1, \qquad w_1 = \frac{(\phi^{-1})'(\phi(H))}{\langle b, \Gamma^H\rangle}\{a - \phi(H)b - a_0u_1\}$$

and the vector u_1 *is exactly as in* (3.3.27) *and the matrix* $\mathcal{C}^H$ *as in* (3.3.2)*.*

A multi-step algorithm

Even though $\widehat{C}_t^{n,1}$ and $\widehat{H}_1^n$ from Propositions 3.3.7 and 3.3.8 are rate-optimal and asymptotically bias-free estimators of C_t and H, respectively, we can still do better: The estimator $\widehat{C}_t^{n,1}$ is based on the initial estimator $\widetilde{C}_t^{n,1}$ from (3.3.22), which in turn is based on the initial estimator $\widetilde{H}^n$ of H. Now that we have a better estimator of H, namely $\widehat{H}_1^n$, the idea is to use $\widehat{H}_1^n$ to construct an updated estimator, say, $\widetilde{C}_t^{n,2}$, of C_t. And with this updated estimator of C_t, we next update $\widehat{H}_1^n$ to, say, $\widehat{H}_2^n$, which we can then use again to update $\widetilde{C}_t^{n,2}$, and so on. A related approach was used in [78].

Proposition 3.3.9. *For* $k = 2,\dots,m$ *where* $m \geq 2$ *is an integer, we define iteratively*

$$\widehat{C}_t^{n,k} = \left\{\widehat{V}_{0,t}^n - \frac{\langle a, \widehat{V}_t^n\rangle}{\langle a, \Gamma^{\widehat{H}_{k-1}^n}\rangle}\right\}\left(1 - \frac{a_0}{\langle a, \Gamma^{\widehat{H}_{k-1}^n}\rangle}\right)^{-1} \tag{3.3.31}$$

and

$$\widehat{H}_k^n = \widetilde{H}^n + \sum_{j=1}^{N(\widehat{H}_{k-1}^n)} \Phi_j^n\left(\widehat{C}_t^{n,k}\right)^j. \tag{3.3.32}$$

Then

$$\Delta_n^{-\frac{1}{2}}(\widehat{H}_k^n - H) \xrightarrow{\text{st}} \mathcal{N}\left(0, \operatorname{Var}_{H,k}\frac{\int_0^t \rho_s^4\,\mathrm{d}s}{(\int_0^t \rho_s^2\,\mathrm{d}s)^2}\right), \tag{3.3.33}$$

$$\Delta_n^{\frac{1}{2}-2H}(\widehat{C}_t^{n,k} - C_t) \xrightarrow{\text{st}} \mathcal{N}\left(0, \operatorname{Var}_{C,k}\int_0^t \rho_s^4\,\mathrm{d}s\right), \tag{3.3.34}$$

where, for each $k = 2,\dots,m$*,*

$$\operatorname{Var}_{H,k} = \operatorname{Var}_{H,k}(R,a,b,H) = w_k^T\mathcal{C}^H w_k, \qquad \operatorname{Var}_{C,k} = \operatorname{Var}_{C,k}(R,a,b,H) = u_k^T\mathcal{C}^H u_k,$$

and

$$u_k = \left(e_1 - \frac{a}{\langle a, \Gamma^H\rangle} + \frac{\langle a, \partial_H\Gamma^H\rangle}{\langle a, \Gamma^H\rangle}w_{k-1}\right)\left(1 - \frac{a_0}{\langle a, \Gamma^H\rangle}\right)^{-1},$$

$$w_k = \frac{(\phi^{-1})'(\phi(H))}{\langle b, \Gamma^H\rangle}\{a - \phi(H)b - a_0u_k\}.$$

Our final estimator of H is

$$\widehat{H}^n = \widehat{H}^n_m. \tag{3.3.35}$$

For later references, let us define

$$\mathrm{Var}_H = \mathrm{Var}_H(R,a,b,H) = \mathrm{Var}_{H,m}(R,a,b,H). \tag{3.3.36}$$

The next theorem exhibits our final estimators for C_t and Π_t

Theorem 3.3.10. *Choose $c \in \mathbb{R}^{1+R}$ and define*

$$\widehat{C}^n_t = \left\{ \widehat{V}^n_{0,t} - \frac{\langle c, \widehat{V}^n_t \rangle}{\langle c, \Gamma^{\widehat{H}^n} \rangle} \right\} \left(1 - \frac{c_0}{\langle c, \Gamma^{\widehat{H}^n} \rangle} \right)^{-1} \tag{3.3.37}$$

and

$$\widehat{\Pi}^n_t = \left\{ \frac{\langle a, \widehat{V}^n_t \rangle}{\langle a, \Gamma^{\widehat{H}^n} \rangle} - \frac{a_0}{\langle a, \Gamma^{\widehat{H}^n} \rangle} \widehat{C}^n_t \right\} \Delta_n^{1-2\widehat{H}^n}. \tag{3.3.38}$$

Then

$$\Delta_n^{\frac{1}{2}-2H} (\widehat{C}^n_t - C_t) \xrightarrow{\mathrm{st}} \mathcal{N}\left(0, \mathrm{Var}_C \int_0^t \rho_s^4 \,\mathrm{d}s\right), \tag{3.3.39}$$

$$\frac{\Delta_n^{-\frac{1}{2}}}{|\log \Delta_n|} (\widehat{\Pi}^n_t - \Pi_t) \xrightarrow{\mathrm{st}} \mathcal{N}\left(0, 4\mathrm{Var}_H \int_0^t \rho_s^4 \,\mathrm{d}s\right), \tag{3.3.40}$$

where

$$\begin{aligned} \mathrm{Var}_C &= \mathrm{Var}_C(R,a,b,c,H) = u^T \mathcal{C}^H u, \\ u &= \left(e_1 - \frac{c}{\langle c, \Gamma^H \rangle} + \frac{\langle c, \partial_H \Gamma^H \rangle}{\langle c, \Gamma^H \rangle} w_m \right) \left(1 - \frac{c_0}{\langle c, \Gamma^H \rangle} \right)^{-1}. \end{aligned} \tag{3.3.41}$$

Feasible CLTs

Finally, we replace unknown parameters that appear in Var_H and Var_C by consistent estimators to obtain feasible CLTs. The following theorem is a direct consequence of (3.3.13), Proposition 3.3.9 and Theorem 3.3.10.

Theorem 3.3.11. *Assume that $H \in (\frac{1}{4}, \frac{1}{2})$. Choose $R \geq 0$, $m \geq 2$ and $a, b, c \in \mathbb{R}^{1+R}$ such that* (3.3.17) *is satisfied and ϕ from* (3.3.4) *is invertible. Further choose $a^0, b^0 \in \mathbb{R}^{1+R}$ as in Proposition 3.3.6. Then*

$$\Delta_n^{-\frac{1}{2}} (\widehat{H}^n - H) \sqrt{\frac{3\Delta_n (\widehat{V}^n_{0,t})^2}{\mathrm{Var}_H(R,a,b,\widehat{H}^n)\widehat{Q}^n_t}} \xrightarrow{\mathrm{st}} \mathcal{N}(0,1), \tag{3.3.42}$$

$$\Delta_n^{\frac{1}{2}-2\widehat{H}^n} (\widehat{C}^n_t - C_t) \sqrt{\frac{3\Delta_n^{4\widehat{H}^n-1}}{\mathrm{Var}_C(R,a,b,c,\widehat{H}^n)\widehat{Q}^n_t}} \xrightarrow{\mathrm{st}} \mathcal{N}(0,1), \tag{3.3.43}$$

$$\frac{\Delta_n^{-\frac{1}{2}}}{|\log \Delta_n|} (\widehat{\Pi}^n_t - \Pi_t) \sqrt{\frac{3\Delta_n^{4\widehat{H}^n-1}}{4\mathrm{Var}_H(R,a,b,\widehat{H}^n)\widehat{Q}^n_t}} \xrightarrow{\mathrm{st}} \mathcal{N}(0,1). \tag{3.3.44}$$

3.4 Simulation study

All results reported in this Section are based on 5,000 simulations from the mfBM

$$Y_t = X_t + Z_t = \sigma B_t + \rho B_t^H, \qquad t \in [0,T],$$

where $\sigma = 0.01$, $\rho = 0.001$, B and B^H are independent and $T = 1$ or $T = 20$ trading days, each consisting of 6.5 hours or $n = 23{,}400$ seconds. Accordingly, we choose $\Delta_n = 1/n = 1/23{,}400$. The values of H will be taken from the set

$$H \in \{0.05, 0.1, 0.15, 0.2, 0.25, 0.275, 0.3, 0.325, 0.35, 0.4, 0.45\}. \tag{3.4.1}$$

We additionally consider the cases "$H = 0.5$", which means $\rho = 0$, and "$H = 0$", which means that $(B_t^0)_{t\in[0,T]}$ is a collection of i.i.d. standard normal noise variables.[6]

3.4.1 Choosing the tuning parameters

We fix the number of iterations in the multi-step algorithm of Section 3.3.2 at $m = 50$. In fact, for an overwhelming majority of estimates obtained in the simulation and the empirical analysis of Section 3.5, a precision of 10^{-5} was attained after fewer than 50 steps. We further make the choice $R = 60$, which corresponds to considering quadratic variations with time lags up to one minute. In order to tune the remaining parameters, we want to choose the vectors $a, b, c \in \mathbb{R}^{1+R}$ in such a way that $(a, b, c) \mapsto \mathrm{Var}_C(R, a, b, c, H)$ is as small as possible. Due to the complexity of how Var_C depends on a, b and c, we were not able to find (and doubt there is) an analytical expression for the minimizers. In addition, Var_C depends on H, which is unknown. Pretending we knew H for the moment and $H \in (\frac{1}{4}, \frac{1}{2})$, in order to resolve the first issue, we choose[7]

$$a = c = \frac{\Gamma^H - \langle \Gamma^H, b\rangle b}{\|\Gamma^H - \langle \Gamma^H, b\rangle b\|}, \qquad b = \frac{\partial_H \Gamma^H}{\|\partial_H \Gamma^H\|} \tag{3.4.2}$$

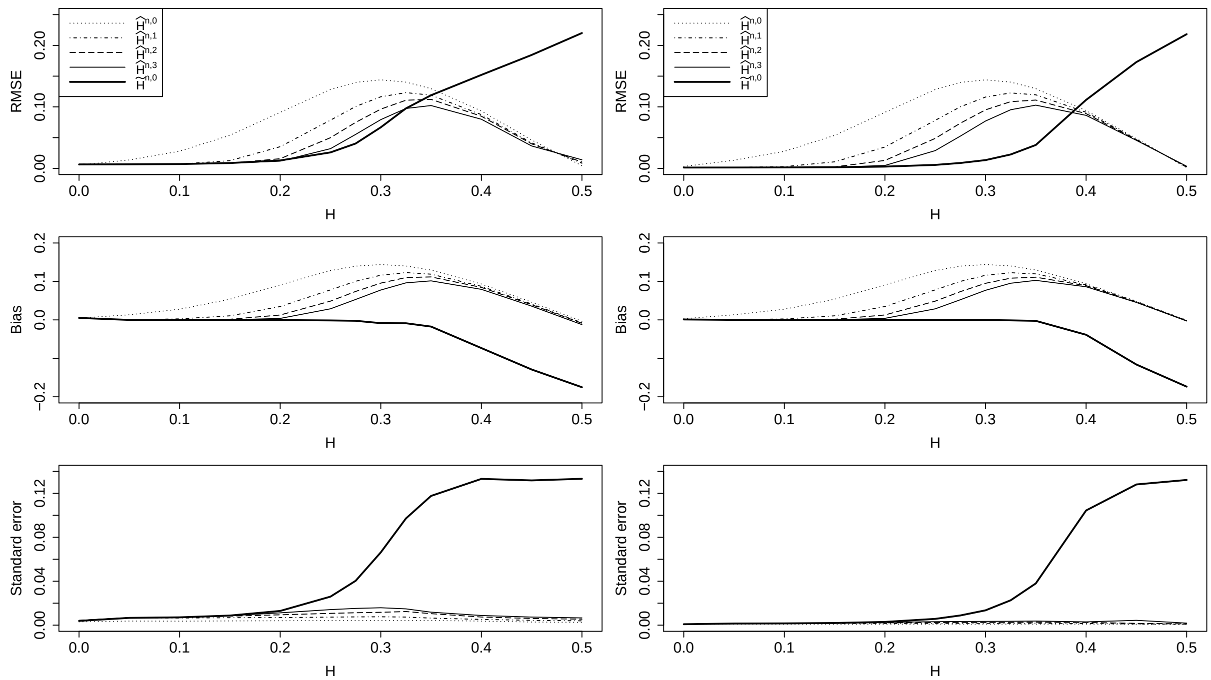
as initial values and run the R function `fminsearch()` from the package `pracma` to find (local) minimizers $a(H)$, $b(H)$ and $c(H)$ of $(a, b, c) \mapsto \mathrm{Var}_C(R, a, b, c, H)$ from (3.3.41). Similarly, we obtain $a^0(H)$, $b^0(H)$ and $c^0(H)$ as minimizers of $(a, b, c) \mapsto \mathrm{Var}_C(R, a, b, c, H)$ from (3.3.9) by taking the same initial weights b and c from (3.4.2) for b^0 and c^0 and the vector obtained by substituting 0 for the first component of a from (3.4.2) for a^0. As H is unknown in this process, one could, in principle, plug in a consistent estimator of H (e.g., $\widehat{H}^n$, computed for some initial choice of a, b and c), determine the minimizing vectors, use them to construct an update of $\widehat{H}^n$, and repeat this procedure. However, such an adaptive scheme of constructing $\widehat{H}^n$ makes the weight vectors dependent on the latest estimator of H and therefore changes its asymptotic variance in every step. Unfortunately, we see no way of keeping track of those changes, in particular because we do not know the precise form of how $a(H)$, $b(H)$ and $c(H)$ depend on H. Instead, we take the minimizers at $H_0 = 0.35$, that is, we let

$$a^0 = a^0(0.35), \qquad b^0 = b^0(0.35), \qquad c^0 = c^0(0.35), \tag{3.4.3}$$

$$a = a(0.35), \qquad b = b(0.35), \qquad c = c(0.35). \tag{3.4.4}$$

[6] Caution: The law of B^H does not tend to that of i.i.d. noise as $H \to 0$. We formally set $H = 0$ only because the variance of i.i.d. noise is non-shrinking, which is what the order Δ_n^{2H} of fBM increments yields when H is set to 0.

[7] This is a heuristic choice: with these vectors, $\langle c, \partial_H \Gamma^H\rangle = 0$ in (3.3.41) and $\langle u, \Gamma^H\rangle = 0$. Consequently, if $\mathcal{C}_{0,1}^H$ and $\mathcal{C}_{0,2}^H$ denote the two zeroth-order terms in (3.3.2), then $u^T \mathcal{C}_{0,2}^H u = 0$ and, in $u^T \mathcal{C}_{0,1}^H u = \sum_{i,j=0}^R u_i u_j (\mathcal{C}_{0,1}^H)_{ij}$, the part of the sum where $i = j$ is 0.



(a) $T = 1$ (b) $T = 20$

Figure 3.3: RMSE, bias and standard error of $\widehat{H}^{n,i}$, $i = 0, 1, 2, 3$ and $\widetilde{H}^{n,0}$.

The resulting variances $\mathrm{Var}_C(R, a^0(H_0), b^0(H_0), c^0(H_0), H)$ and $\mathrm{Var}_C(R, a(H_0), b(H_0), c(H_0), H)$ at other values of H turn out to be reasonably close to the H-dependent minimal values given by $\mathrm{Var}_C(R, a^0(H), b^0(H), c^0(H), H)$ and $\mathrm{Var}_C(R, a(H), b(H), c(H), H)$, respectively (no more than 2.1% larger in the former case; no more than 8.1% larger in the latter case for all H in (3.4.1) except for $H = 0.45$, where the variance based on (3.4.4) is 2.6 times larger).

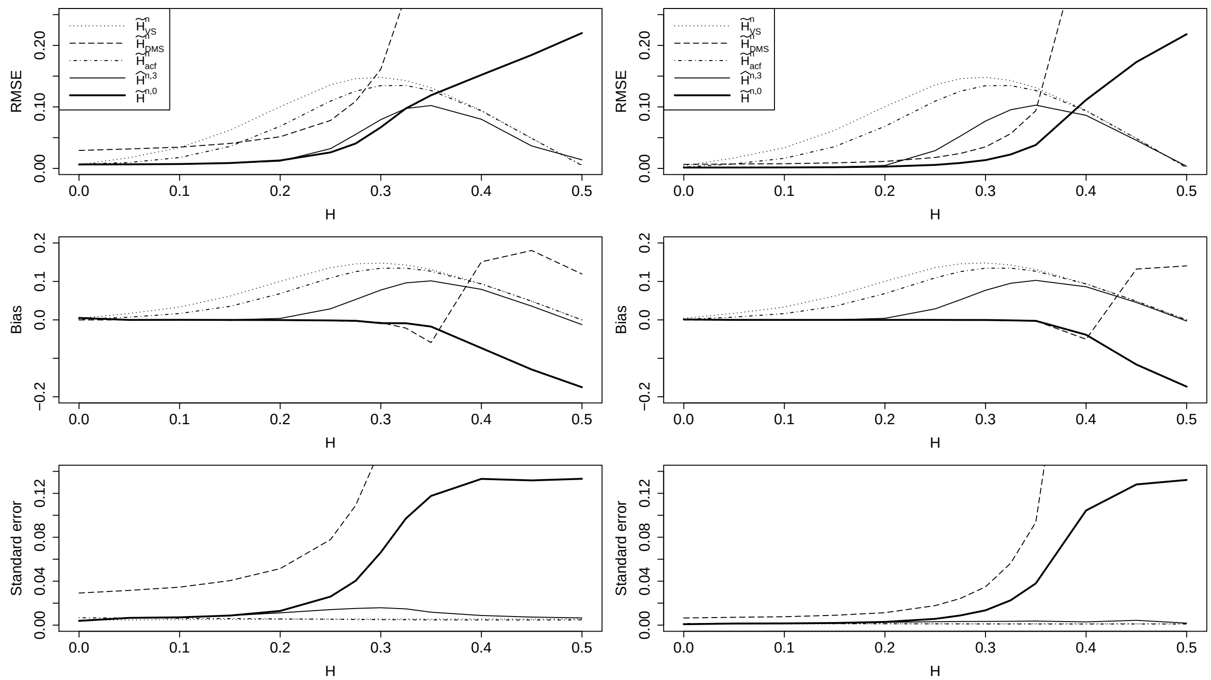
3.4.2 Performance of estimators and comparison with existing estimators

For H, we first compare our estimator $\widetilde{H}^{n,0} = \widetilde{H}^n$ from (3.3.4), constructed with a^0 and b^0 from (3.4.3), with four variants of $\widehat{H}^n$ from (3.3.35), denoted by $\widehat{H}^{n,i}$ for $i = 0, 1, 2, 3$. For each i, $\widehat{H}^{n,i}$ is defined in the same way as $\widehat{H}^n$ in (3.3.35) [constructed with a and b from (3.4.4) and, in step (3.3.22), with a^0 and b^0 from (3.4.3)] except that $N(\widetilde{H}^n)$ in (3.3.29) and (3.3.30) and $N(\widehat{H}^n_{k-1})$ in (3.3.32) are replaced by the fixed number i. In particular, if n is large, then with high probability,

$$\widehat{H}^n = \begin{cases} \widehat{H}^{n,0} & \text{if } H \in (0, 0.25), \\ \widehat{H}^{n,1} & \text{if } H \in (0.25, 0.375), \\ \widehat{H}^{n,2} & \text{if } H \in (0.375, 0.41\overline{6}), \\ \widehat{H}^{n,3} & \text{if } H \in (0.41\overline{6}, 0.4375). \end{cases} \tag{3.4.5}$$

We do not include four or more correction terms as it becomes increasingly intractable to compute higher-order derivatives of composite functions like ϕ^{-1} and ψ in (3.3.19) or (3.3.26) by hand.

As Figure 3.3 shows, $\widehat{H}^{n,3}$ has a lower root-mean-square error (RMSE) than $\widehat{H}^{n,i}$ for any $i = 0, 1, 2$, although, according to the theory in Section 3.3.2, it suffices for asymptotic normality to only include i corrections and consider $\widehat{H}^{n,i}$ in the ranges of H specified in (3.4.5). In fact,



(a) $T = 1$ (b) $T = 20$

Figure 3.4: RMSE, bias and standard error of $\widehat{H}^{n,3}$, $\widetilde{H}^{n,0}$, $\widetilde{H}^n_{\mathrm{VS}}$, $\widetilde{H}^n_{\mathrm{DMS}}$ and $\widetilde{H}^n_{\mathrm{acf}}$.

taking $i = 3$ decreases the finite sample bias considerably compared to $i = 0, 1, 2$, while increasing the standard error only moderately, so that, in total, $i = 3$ is best in terms of RMSE among the four variants in (3.4.5). Moreover, if $T = 1$ (resp., $T = 20$), $\widetilde{H}^{n,0}$ is superior to $\widehat{H}^{n,3}$ in terms of RMSE if $H \leq 0.3$ (resp., $H \leq 0.35$) and inferior to $\widehat{H}^{n,3}$ if $H \geq 0.325$ (resp., $H \geq 0.4$). The better/worse performance of $\widetilde{H}^{n,0}$ on the respective range of H is due to a considerably smaller bias / larger standard error. Also, taking $T = 20$ instead of $T = 1$ significantly reduces the RMSE of $\widetilde{H}^{n,0}$ for $H \leq 0.35$, but the RMSE of $\widehat{H}^{n,i}$ is largely unaffected. Extrapolating from these results, we conjecture that including an even higher number of correction terms in (3.4.5) would gradually bring down the finite sample bias while keeping the standard error low. As we mentioned, computing higher-order correction terms is computationally challenging, so examining this conjecture is beyond the scope of the current paper.

In Figure 3.4, we further compare $\widehat{H}^{n,3}$ and $\widetilde{H}^{n,0}$ with

- the regression estimator $\widetilde{H}^n_{\mathrm{VS}}$ based on a volatility signature plot, that is,

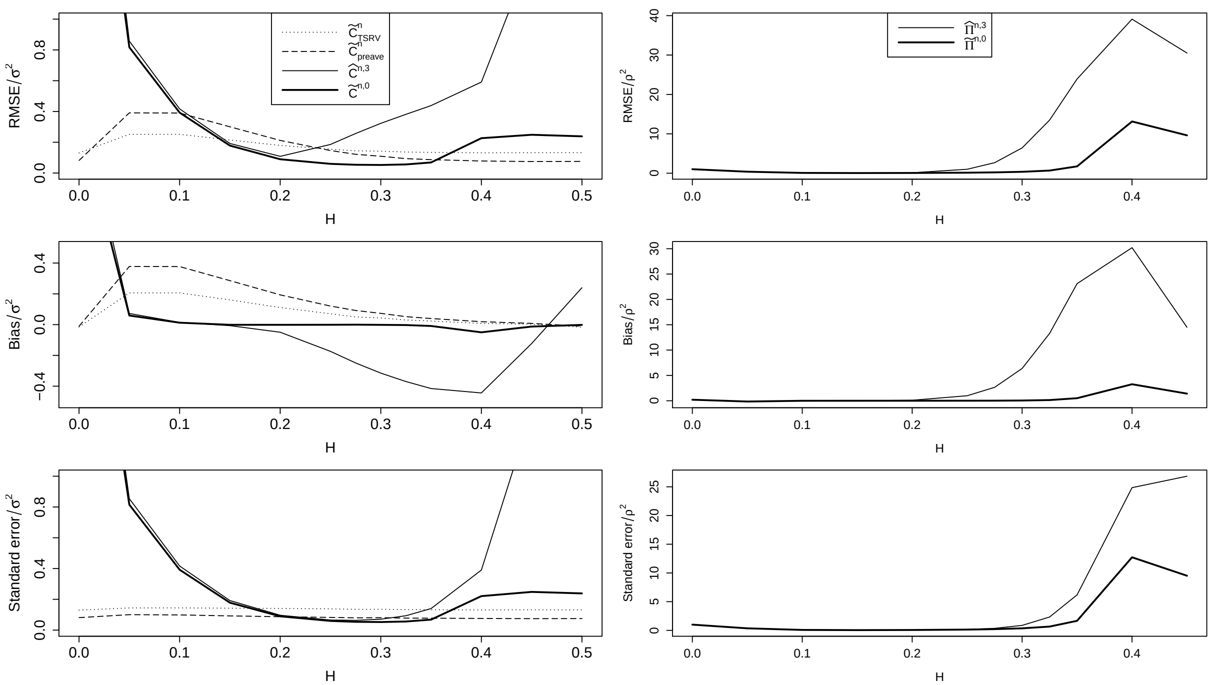
$$\widetilde{H}^n_{\mathrm{VS}} = \tfrac{1}{2}(\widetilde{\beta}^n_{\mathrm{VS}} + 1),$$

  where $\widetilde{\beta}^n_{\mathrm{VS}}$ is the slope estimate in a linear regression of $\log \widehat{V}^{n/i}_{0,t}$ on $\log i$ for $i = 1, \ldots, 10$;

- the rate-optimal estimator from [47, Theorem 3.2] given by ($\log_{2+} x = \log_2 x$ if $x > 0$ and $\log_{2+} x = 0$ otherwise)

$$\widetilde{H}^n_{\mathrm{DMS}} = \frac{1}{2}\left(\log_{2+} \frac{\widehat{V}^{n/4}_{0,t} - \widehat{V}^{n/2}_{0,t}}{\widehat{V}^{n/2}_{0,t} - \widehat{V}^{n}_{0,t}} + 1 \right);$$

- the autocorrelation estimator $\widetilde{H}^n_{\mathrm{acf}}$ from (3.3.3) (cf. [47, Proposition 3.1]).



(a) Estimators for volatility (b) Estimators for noise volatility

Figure 3.5: RMSE, bias and standard error of (a) $\widehat{C}^{n,3}$, $\widetilde{C}^{n,0}$, $\widetilde{C}^n_{\text{TSRV}}$ and $\widetilde{C}^n_{\text{preave}}$, and (b) of $\widehat{\Pi}^{n,3}$ and $\widetilde{\Pi}^{n,0}$. Negative volatility estimates were replaced by 0 in the evaluation.

Comparing the values for RMSE for $T = 1$, we see that $\widehat{H}^{n,3}$ performs better than $\widetilde{H}^n_{\text{VS}}$, $\widetilde{H}^n_{\text{DMS}}$ and $\widetilde{H}^n_{\text{acf}}$ for all values of H except when there is no noise ("$H = 0.5$"). When $T = 20$, $\widehat{H}^{n,3}$ is still better than $\widetilde{H}^n_{\text{VS}}$ and $\widetilde{H}^n_{\text{acf}}$ but is not as good as $\widetilde{H}^n_{\text{DMS}}$ when $H \in \{0.25, \dots, 0.35\}$. When $T = 20$, the best estimator in the range $H \leq 0.35$ is $\widetilde{H}^{n,0}$; but if $H \geq 0.4$, just like $\widetilde{H}^n_{\text{DMS}}$ which does not use quadratic variation, either, the RMSE of $\widetilde{H}^{n,0}$ becomes large.

Finally, we study the performance of our volatility estimators. To this end, we implement $\widetilde{C}^{n,0} = \widehat{C}^n_{20} - \widehat{C}^n_{19}$ and $\widetilde{\Pi}^{n,0} = \widehat{\Pi}^n_{20} - \widehat{\Pi}^n_{19}$ from (3.3.7) and (3.3.11) on the last simulated day, using the estimator $\widetilde{H}^{n,0} = \widetilde{H}^n$ from (3.3.4) that is based on the whole simulated period of 20 days. Similarly, we consider $\widehat{C}^{n,3} = \widehat{C}^n_{20} - \widehat{C}^n_{19}$ and $\widehat{\Pi}^{n,3} = \widehat{\Pi}^n_{20} - \widehat{\Pi}^n_{19}$ from (3.3.37) and (3.3.38), using, instead of $\widehat{H}^n$, the estimator $\widehat{H}^{n,3}$ from above computed again based on the whole period of 20 simulated days. We further compare $\widetilde{C}^{n,0}$ and $\widehat{C}^{n,3}$ with the *two-scale realized variance* estimator $\widetilde{C}^n_{\text{TSRV}}$ of [100] and the pre-averaging estimator $\widetilde{C}^n_{\text{preave}}$ of [56], as implemented by the functions `rTSCov()` and `rMRC()` in the R package `highfrequency`.

From Figure 3.5, we find that $\widetilde{C}^{n,0}$ (resp., $\widetilde{\Pi}^{n,0}$) performs better than $\widehat{C}^{n,3}$ (resp., $\widehat{\Pi}^{n,3}$), both in terms of bias and standard error. Furthermore, in terms of RMSE, $\widetilde{C}^{n,0}$ outperforms both $\widetilde{C}^n_{\text{TSRV}}$ and $\widetilde{C}^n_{\text{preave}}$ for $H \in \{0.15, \dots, 0.35\}$, but is inferior to the latter two when H is small ($H \in \{0, 0.05, 0.1\}$) or large ($H \in \{0.4, 0.45, 0.5\}$). Interestingly, the poorer performance of $\widetilde{C}^n_{\text{TSRV}}$ and $\widetilde{C}^n_{\text{preave}}$ in the range $H \in \{0.15, \dots, 0.35\}$ is due to larger biases, while $\widetilde{C}^{n,0}$ performs worse for $H \in \{0, 0.05, 0.1, 0.4, 0.45, 0.5\}$ due to larger standard errors.

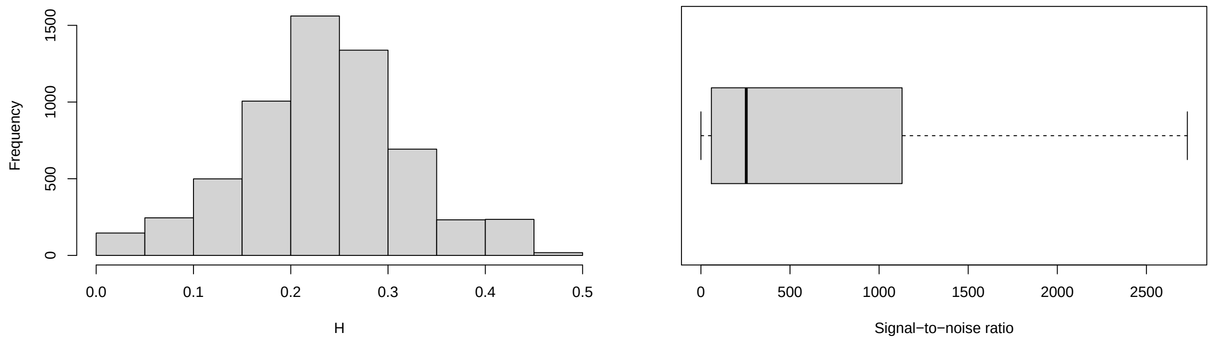


Figure 3.6: Histogram of estimates for H and boxplot of signal–to–noise ratios. Each data point corresponds to one company and day. Signal–to–noise ratios where the volatility or the noise volatility estimate is negative are omitted. Outliers in the boxplot are not shown.

3.5 Empirical analysis

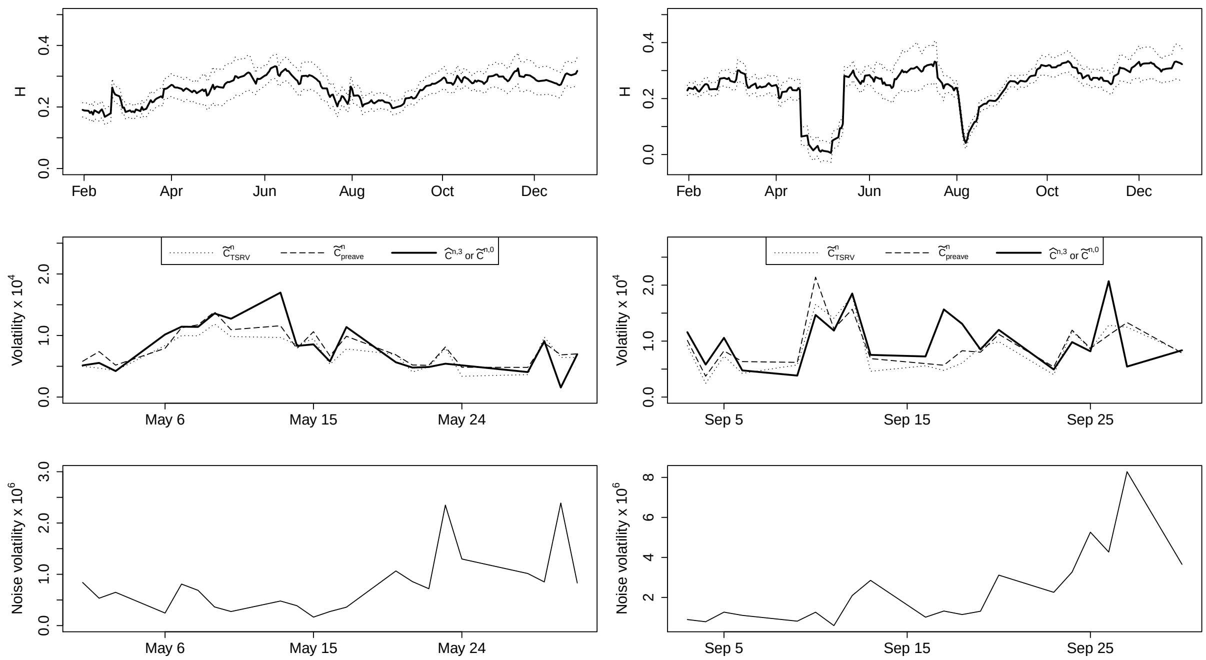
We apply the estimators from Theorems 3.3.2 and 3.3.11 to (logarithmic) mid-quote data for each of the 29 stocks that were constituents of the DJIA index for the whole year of 2019. The data source is the TAQ database. For each trading day in 2019, we collect all quotes[8] on the NYSE or NASDAQ from 9:00 am until 4:00 pm Eastern Time and preprocess them using the `quotesCleanup()` function from the R package `highfrequency`. We sample in calendar time every second.

To reduce the variability of the resulting estimates, we calculate, for each trading day from January 31 to December 31, the estimators $\widehat{H}^{n,3}$ and $\widetilde{H}^{n,0}$ based on the previous 20 trading days. Afterwards, based on the insights from the simulation study of Section 3.4, we calculate an estimate of H using $\widehat{H}^{n,3}$ if its asymptotic 95%-confidence interval contains 0.5 or is a subset of $(0.4, 0.5)$; otherwise, we report the estimate produced by $\widetilde{H}^{n,0}$. Correspondingly, we either take $\widehat{C}^{n,3}$ or $\widetilde{C}^{n,0}$ (resp., $\widehat{\Pi}^{n,3}$ or $\widetilde{\Pi}^{n,0}$) to estimate the daily integrated volatility (resp., noise volatility). Figure 3.6 shows the empirical distribution of the daily estimators of H and a boxplot of the daily signal–to–noise ratios (i.e., of $\widehat{C}^{n,3}/\widehat{\Pi}^{n,3}$ or $\widetilde{C}^{n,0}/\widetilde{\Pi}^{n,0}$). The top row of Figure 3.7 shows the daily H-estimates for the two individual stocks, American Express (AXP) and IBM, including 95%-confidence intervals. In the second and third row, we show the daily volatility and noise volatility estimates for the month of May (AXP) and September (IBM). In these months, the respective H-estimates are all above 0.25.

3.6 Discussion

In this paper, we introduced mixed semimartingales as a natural class of microstructure noise models for high-frequency financial data. Defined as a sum of a continuous semimartingale and a continuous-time moving average process that locally resembles fractional Brownian motion, mixed semimartingales are nonparametric extensions of the mixed fractional Brownian motion of [25] that include (possibly dependent) stochastic volatility and noise volatility. From a modeling point of view, mixed semimartingales can capture microstructure noise in continuous time and,

[8] We also examined transaction data for the same stocks in 2019. But sampled at five seconds, many daily estimators for H were not significantly different from $\frac{1}{2}$ at the 5% level. This is in line with other research (see, for example, [6]) demonstrating that transaction data of DJIA stocks have become less and less noisy over the recent years.



(a) AXP 2019 (b) IBM 2019

Figure 3.7: Top row: Estimates of H with asymptotic 95%-confidence intervals. Middle and bottom row: Volatility and noise volatility estimates for May 2019 (AXP) and September 2019 (IBM).

at the same time, allow for varying rates at which the variance of price increments shrinks or at which realized variance explodes as the sampling frequency increases.

Based on central limit theorems for variation functionals, we constructed consistent and asymptotically (mixed) normal estimators for the Hurst parameter, the integrated volatility and the integrated noise volatility—in all cases where these quantities are identifiable. Our analysis showed profound differences to the pure semimartingale or fractional process setting: the interplay between the two gives rise to a possibly large number of intermediate limits (or higher-order bias terms), whose removal is intricate and necessitates an iteration procedure. In a simulation study, we found that our estimators of the Hurst index H outperform existing ones in the literature, for all considered values of H. Our volatility estimators perform best in the region $H \in (0.15, 0.35)$ and are superior to standard noise-robust volatility estimators in this range in terms of RMSE. For smaller and higher values of H, even though one of our volatility estimators shows almost no bias, they suffer from large standard errors. How to tackle these problems and perhaps combine them with existing techniques such as pre-averaging remain open problems. Also, in this first paper, we did not examine the effect of, for example, jumps [1, 2, 66] or irregular observation times [17, 57, 62] on our estimators. While our estimators of H and noise volatility from Theorem 3.3.4 might not be affected by jumps too much, as they do not use quadratic variation, certainly the volatility estimators and all estimators from Theorem 3.3.11 are. We leave it to future research to develop estimators that are fully robust to jumps and asynchronous sampling.

We further applied our estimators to 2019 quote data of DJIA stocks. We found strong empirical evidence for asset- and time-dependent values of H, which further underpins the need to model the roughness of noise in a flexible manner. Allowing for time-dependent and/or stochastic Hurst parameters ("multifractality") is yet another challenge to be explored in the future.

Chapter 4

Supplement to: "Mixed semimartingales: Volatility estimation in the presence of fractional noise"

We gather some frequently used moment estimates in Section 4.1 before moving in Section 4.2 to the details of proof of the main result of Chapter 3, which is Theorem 3.2.1. Section 4.3 contains the proofs of the results announced in Section 3.3 of Chapter 3 and Section 4.4 lists some estimates on fractional kernels needed in the proofs.

We use the notation from Chapter 3. In addition, we write $A \lesssim B$ if there is a constant C that is independent of any quantity of interest such that $A \leq CB$.

4.1 Size estimates

In the following, we repeatedly make use of so-called *standard size estimates* (cf. [30, Appendix D]). Under the strengthened hypotheses of Assumption (CLT'), consider for fixed indices $j, k \in \{1, \dots, d\}$ and $\ell \in \{1, \dots, L\}$ an expression like

$$S_n(t) = \Delta_n^{\frac{1}{2}} \sum_{i=\theta_n+1}^{[t/\Delta_n]} h(\zeta_i^n) \Bigg(\frac{\Delta_{i+\ell-1}^n A^k}{\Delta_n^H} + \frac{1}{\Delta_n^H} \int_{(i+\ell-2)\Delta_n}^{(i+\ell-1)\Delta_n} \left(\sigma_s^{kj} - \sigma_{(i-\theta_n'')\Delta_n}^{kj} \right) \mathrm{d}B_s^j + \int_0^\infty \frac{\Delta_{i+\ell-1}^n g(s)}{\Delta_n^H} \left(\rho_s^{kj} - \rho_{(i-\theta_n)\Delta_n}^{kj} \right) \mathbb{1}_{((i-\theta_n)\Delta_n, (i-\theta_n')\Delta_n)}(s) \,\mathrm{d}W_s^j \Bigg), \tag{4.1.1}$$

where $\theta_n = [\Delta_n^{-\theta}]$, $\theta_n' = [\Delta_n^{-\theta'}]$, $\theta_n'' = [\Delta_n^{-\theta''}]$ and $-\infty \leq \theta', \theta'' < \theta \leq \infty$. In addition, h is a function such that $|h(x)| \lesssim 1 + \|x\|^p$ for some $p > 1$, and ζ_i^n are random variables with

$$\sup_{n\in\mathbb{N}} \sup_{i=1,\dots,[T/\Delta_n]} \mathbb{E}[\|\zeta_i^n\|^p] < \infty.$$

For any $q \geq 1$, because a is uniformly bounded by (3.2.11), Minkowski's integral inequality yields

$$\mathbb{E}\left[\left\| \frac{\Delta_{i+\ell-1}^n A}{\Delta_n^H} \right\|^q \right]^{\frac{1}{q}} \leq \frac{1}{\Delta_n^H} \int_{(i+\ell-2)\Delta_n}^{(i+\ell-1)\Delta_n} \mathbb{E}\left[\|a_s\|^q \right]^{\frac{1}{q}} \,\mathrm{d}s \lesssim \Delta_n^{1-H}. \tag{4.1.2}$$

Similarly, using the Burkholder–Davis–Gundy (BDG) inequality and (3.2.11) and (3.2.12), we obtain

$$\mathbb{E}\left[\left|\frac{1}{\Delta_n^H}\int_{(i+\ell-2)\Delta_n}^{(i+\ell-1)\Delta_n}\left(\sigma_s^{kj}-\sigma_{(i-\theta_n'')\Delta_n}^{kj}\right)\mathrm{d}B_s^j\right|^q\right]^{\frac{1}{q}} \lesssim (\theta_n''\Delta_n)^{\frac{1}{2}}\Delta_n^{\frac{1}{2}-H}. \tag{4.1.3}$$

Combining (3.2.11) and (3.2.12) with Lemma 4.4.1, we deduce that

$$\begin{aligned}&\mathbb{E}\left[\left|\int_0^\infty \frac{\Delta_{i+\ell-1}^n g(s)}{\Delta_n^H}\left(\rho_s^{kj}-\rho_{(i-\theta_n)\Delta_n}^{kj}\right)\mathbb{1}_{((i-\theta_n)\Delta_n,(i-\theta_n')\Delta_n)}(s)\,\mathrm{d}W_s^j\right|^q\right]^{\frac{1}{q}}\\ &\quad\lesssim(\theta_n\Delta_n)^{\frac{1}{2}}\left(\frac{1}{\Delta_n^{2H}}\int_0^{(i-\theta_n')\Delta_n}\Delta_{i+\ell-1}^n g(s)^2\,\mathrm{d}s\right)^{\frac{1}{2}}\lesssim(\theta_n\Delta_n)^{\frac{1}{2}}\,\Delta_n^{\theta'(1-H)}.\end{aligned} \tag{4.1.4}$$

Finally, using Hölder's inequality to separate $h(\zeta_i^n)$ from the subsequent expression in (4.1.1), we have shown that

$$\begin{aligned}\mathbb{E}\left[\sup_{t\le T}\left|S_n(t)\right|\right]&\lesssim\Delta_n^{\frac{1}{2}}\sum_{i=\theta_n+1}^{[T/\Delta_n]}\left\{\Delta_n^{1-H}+\Delta_n^{1-H}(\theta_n'')^{\frac{1}{2}}+(\theta_n\Delta_n)^{\frac{1}{2}}\,\Delta_n^{\theta'(1-H)}\right\}\\ &\lesssim\Delta_n^{\frac{1}{2}-H}+\Delta_n^{\frac{1}{2}-H-\frac{\theta''}{2}}+\Delta_n^{\theta'(1-H)-\theta}.\end{aligned} \tag{4.1.5}$$

The upshot of this example is that the absolute moments of sums and products of more or less complicated expressions can always be bounded term by term: for example, in (4.1.1), the terms

$$\sum_{i=\theta_n+1}^{[t/\Delta_n]},\qquad h(\zeta_i^n),\qquad \Delta_{i+\ell-1}^n A^k,\qquad \int_{(i+\ell-2)\Delta_n}^{(i+\ell-1)\Delta_n}(\cdots)\,\mathrm{d}B_s^j,\qquad \sigma_s^{kj}-\sigma_{(i-\theta_n'')\Delta_n}^{kj},$$
$$\int_0^{(i-\theta_n')\Delta_n}\frac{\Delta_{i+\ell-1}^n g(s)}{\Delta_n^H}(\cdots)\,\mathrm{d}W_s^j,\qquad \rho_s^{kj}-\rho_{(i-\theta_n)\Delta_n}^{kj}$$

have *sizes* (i.e., their L^q-moments, for any q, can be uniformly bounded by a constant times)

$$\Delta_n^{-1},\qquad 1,\qquad \Delta_n,\qquad \sqrt{\Delta_n},\qquad (\theta_n''\Delta_n)^{\frac{1}{2}},\qquad \Delta_n^{\theta'(1-H)},\qquad (\theta_n\Delta_n)^{\frac{1}{2}},$$

respectively. The final estimate (4.1.5) is then obtained by combining these bounds. Clearly, size estimates can be estimated to variants of (4.1.1), too, for example, when the stochastic integral in (4.1.1) is squared, when we have products of integrals, when $S_n(t)$ is matrix-valued, etc.

Even though size estimates are optimal in general, better estimates may be available in specific cases. One such case occurs when sums have a martingale structure. To illustrate this, let $\mathcal{F}_i^n=\mathcal{F}_{i\Delta_n}$ and consider

$$S_n'(t)=\Delta_n^{\frac{1}{2}}\sum_{i=1}^{[t/\Delta_n]-L+1}\varpi_i^n$$

with random variables ϖ_i^n that are $\mathcal{F}_i^n$-measurable and satisfy $\mathbb{E}[\varpi_i^n\mid\mathcal{F}_{i-\theta_n'''}^n]=0$, where $\theta_n'''=[\Delta_n^{-\theta'''}]$ for some $0<\theta'''<1$. Suppose that $\mathbb{E}[|\varpi_i^n|^2]^{1/2}\lesssim\Delta_n^{\varpi}$ uniformly in i and n for some $\varpi>0$. Writing

$$S_n'(t)=\sum_{j=1}^{\theta_n'''}S_{n,j}'(t),\qquad S_{n,j}'(t)=\Delta_n^{\frac{1}{2}}\sum_{k=1}^{([t/\Delta_n]-L+1)/\theta_n'''}\varpi_{j+(k-1)\theta_n'''}^n,$$

we observe that each $S'_{n,j}$ is a martingale in t (albeit relative to different filtrations), so the BDG inequality (applied to $S'_{n,j}$) and Minkowski's inequality (applied to the sum over j) yields

$$\mathbb{E}\left[\sup_{t\leq T}|S'_n(t)|\right] \lesssim (\theta'''_n)^{\frac{1}{2}}\Delta_n^{\varpi}. \tag{4.1.6}$$

Very often, ϖ_i^n will actually only be $\mathcal{F}^n_{i+L-1}$-measurable. However, a shift by L increments will not change the value of the above estimate. Following [27, Section 4], we refer to (4.1.6) as a *martingale size estimate*.

4.2 Details for the proof of Theorem 3.2.1

This section is devoted to proving the lemmas that appear in Section 3.2.2. Assumption (CLT') is in force throughout.

Proof of Lemma 3.2.6. By the calculations in (4.1.2), (4.1.3) and (4.1.4), we have $\mathbb{E}[\|\underline{\Delta}_i^n Y/\Delta_n^H\|^p]^{1/p} \lesssim 1$ for all $p\geq 1$. As f grows at most polynomially, we see that $\mathbb{E}[|f(\underline{\Delta}_i^n Y/\Delta_n^H)|]$ is of size 1. Hence, $\mathbb{E}[|\Delta_n^{1/2}\sum_{i=1}^{\theta_n} f(\underline{\Delta}_i^n Y/\Delta_n^H)|] \lesssim \Delta_n^{1/2-\theta}$, which implies

$$\Delta_n^{\frac{1}{2}}\sum_{i=1}^{\theta_n} f\left(\frac{\underline{\Delta}_i^n Y}{\Delta_n^H}\right) \xrightarrow{L^1} 0$$

since $\theta < \frac{1}{2}$ by (3.2.14). As a result, omitting the first θ_n terms in the definition of $V_f^n(Y,t)$ does no harm asymptotically. Next, defining

$$\Lambda_i^n = f\left(\frac{\underline{\Delta}_i^n Y}{\Delta_n^H}\right) - f\left(\frac{\underline{\Delta}_i^n Y^{\mathrm{tr}}}{\Delta_n^H}\right), \qquad \overline{\Lambda}_i^n = \Lambda_i^n - \mathbb{E}[\Lambda_i^n \mid \mathcal{F}^n_{i-\theta_n}], \tag{4.2.1}$$

we consider the decomposition

$$\begin{aligned}
&\Delta_n^{\frac{1}{2}}\sum_{i=\theta_n+1}^{[t/\Delta_n]-L+1}\left\{f\left(\frac{\underline{\Delta}_i^n Y}{\Delta_n^H}\right) - f\left(\frac{\underline{\Delta}_i^n Y^{\mathrm{tr}}}{\Delta_n^H}\right)\right\}\\
&\quad = \Delta_n^{\frac{1}{2}}\sum_{i=\theta_n+1}^{[t/\Delta_n]-L+1}\overline{\Lambda}_i^n + \Delta_n^{\frac{1}{2}}\sum_{i=\theta_n+1}^{[t/\Delta_n]-L+1}\mathbb{E}[\Lambda_i^n|\mathcal{F}^n_{i-\theta_n}].
\end{aligned} \tag{4.2.2}$$

By our choice (3.2.14) of θ and since $H<\frac{1}{2}$, the lemma is proved once

$$\begin{aligned}
\mathbb{E}\left[\sup_{t\leq T}\left|\Delta_n^{\frac{1}{2}}\sum_{i=\theta_n+1}^{[t/\Delta_n]-L+1}\overline{\Lambda}_i^n\right|\right] &\lesssim \Delta_n^{\theta(\frac{1}{2}-H)},\\
\mathbb{E}\left[\sup_{t\leq T}\left|\Delta_n^{\frac{1}{2}}\sum_{i=\theta_n+1}^{[t/\Delta_n]-L+1}\mathbb{E}[\Lambda_i^n|\mathcal{F}^n_{i-\theta_n}]\right|\right] &\lesssim \Delta_n^{\theta(\frac{1}{2}-H)} + \Delta_n^{2\theta(1-H)-\frac{1}{2}}
\end{aligned} \tag{4.2.3}$$

are established. To this end, let

$$\lambda_i^n = \frac{\underline{\Delta}_i^n Y - \underline{\Delta}_i^n Y^{\mathrm{tr}}}{\Delta_n^H} = \int_0^{(i-\theta_n)\Delta_n}\rho_s\,\mathrm{d}W_s\,\frac{\underline{\Delta}_i^n g(s)}{\Delta_n^H}.$$

By Assumption (CLT), we have $|f(z) - f(z')| \lesssim (1 + \|z\|^{p-1} + \|z'\|^{p-1})\|z - z'\|$. In addition, $\mathbb{E}[|f(\underline{\Delta}_i^n Y/\Delta_n^H)|]$ is of size 1, so

$$\mathbb{E}[(\overline{\Lambda}_i^n)^2] \lesssim \mathbb{E}[(\Lambda_i^n)^2] \lesssim \mathbb{E}[\|\lambda_i^n\|^2] \lesssim \Delta_n^{2\theta(1-H)},$$

where we used (4.1.4) for the last estimation. By construction, $\overline{\Lambda}_i^n$ is $\mathcal{F}_{i+L-1}^n$-measurable and has conditional expectation 0 given $\mathcal{F}_{i-\theta_n}^n$. Therefore, we can further use an estimate of the kind (4.1.6) to derive

$$\mathbb{E}\left[\sup_{t\leq T}\left|\Delta_n^{\frac{1}{2}}\sum_{i=\theta_n+1}^{[t/\Delta_n]-L+1}\overline{\Lambda}_i^n\right|\right] \lesssim \sqrt{\theta_n}\Delta_n^{\theta(1-H)} \lesssim \Delta_n^{\theta(1-H)-\frac{1}{2}\theta} = \Delta_n^{\theta(\frac{1}{2}-H)},$$

which is the first property in (4.2.3).

Next, we define the following random matrices

$$\psi_i^n = \sigma_{(i-\theta_n)\Delta_n}\underline{\Delta}_i^n B + \int_{(i-\theta_n)\Delta_n}^{(i+L-1)\Delta_n} \rho_{(i-\theta_n)\Delta_n}\,\mathrm{d}W_s\,\underline{\Delta}_i^n g(s).$$

Since f is smooth by Assumption (CLT), we can apply Taylor's theorem twice to obtain the following decomposition:

$$\begin{aligned}\Lambda_i^n &= \sum_{|\chi|=1}\partial^\chi f\left(\frac{\underline{\Delta}_i^n Y^{\mathrm{tr}}}{\Delta_n^H}\right)(\lambda_i^n)^\chi + \sum_{|\chi|=2}\frac{1}{\chi!}\partial^\chi(\eta_i^n)(\lambda_i^n)^\chi \\ &= \sum_{|\chi|=1}\partial^\chi f\left(\frac{\psi_i^n}{\Delta_n^H}\right)(\lambda_i^n)^\chi + \sum_{|\chi|,|\chi'|=1}\partial^{\chi+\chi'}f(\widetilde{\eta}_i^n)\left(\frac{\underline{\Delta}_i^n Y^{\mathrm{tr}} - \psi_i^n}{\Delta_n^H}\right)^{\chi'}(\lambda_i^n)^\chi + \sum_{|\chi|=2}\frac{1}{\chi!}\partial^\chi(\eta_i^n)(\lambda_i^n)^\chi \\ &=: \Lambda_i^{n,1} + \Lambda_i^{n,2} + \Lambda_i^{n,3},\end{aligned}$$

where $\chi, \chi' \in \mathbb{N}_0^{d\times L}$ are multi-indices and η_i^n (resp., $\widetilde{\eta}_i^n$) is a point on the line between $\underline{\Delta}_i^n Y/\Delta_n^H$ and $\underline{\Delta}_i^n Y^{\mathrm{tr}}/\Delta_n^H$ (resp., $\underline{\Delta}_i^n Y^{\mathrm{tr}}/\Delta_n^H$ and ψ_i^n/Δ_n^H). Accordingly, we split

$$\Delta_n^{\frac{1}{2}}\sum_{i=\theta_n+1}^{[t/\Delta_n]-L+1}\mathbb{E}[\Lambda_i^n \mid \mathcal{F}_{i-\theta_n}^n] = \sum_{j=1}^{3}\mathbb{L}_j^n(t), \qquad \mathbb{L}_j^n(t) = \Delta_n^{\frac{1}{2}}\sum_{i=\theta_n+1}^{[t/\Delta_n]-L+1}\mathbb{E}[\Lambda_i^{n,j} \mid \mathcal{F}_{i-\theta_n}^n].$$

As λ_i^n is $\mathcal{F}_{i-\theta_n}^n$-measurable,

$$\mathbb{E}\left[\partial^\chi f\left(\frac{\psi_i^n}{\Delta_n^H}\right)(\lambda_i^n)^\chi \;\middle|\; \mathcal{F}_{i-\theta_n}^n\right] = (\lambda_i^n)^\chi\,\mathbb{E}\left[\partial^\chi f\left(\frac{\psi_i^n}{\Delta_n^H}\right) \;\middle|\; \mathcal{F}_{i-\theta_n}^n\right] = 0$$

because ψ_i^n is centered normal given $\mathcal{F}_{i-\theta_n}^n$ and f has odd partial derivatives of first orders (since f is even). It follows that $\mathbb{L}_1^n(t) = 0$.

Writing $\mathbb{1}_i^n(s) = (\mathbb{1}_{((i-1)\Delta_n, i\Delta_n)}(s), \ldots, \mathbb{1}_{((i+L-2)\Delta_n,(i+L-1)\Delta_n)}(s))$, we further have that

$$\underline{\Delta}_i^n Y^{\mathrm{tr}} - \psi_i^n = \underline{\Delta}_i^n A + \int_0^t (\sigma_s - \sigma_{(i-\theta_n)\Delta_n})\,\mathrm{d}B_s\,\mathbb{1}_i^n(s) + \int_{(i-\theta_n)\Delta_n}^{(i+L-1)\Delta_n}(\rho_s - \rho_{(i-\theta_n)\Delta_n})\,\mathrm{d}W_s\,\underline{\Delta}_i^n g(s).$$

By a standard size estimate, it follows that

$$\begin{aligned}\mathbb{E}\left[\sup_{t\leq T}|\mathbb{L}_2^n(t)|\right] &\lesssim (\Delta_n^{\frac{1}{2}}\Delta_n^{-1})\left(\Delta_n^{1-H} + \theta_n^{\frac{1}{2}}\Delta_n^{1-H} + (\theta_n\Delta_n)^{\frac{1}{2}}\right)\Delta_n^{\theta(1-H)} \\ &\lesssim \Delta_n^{-\frac{1}{2}}\Delta_n^{\theta(1-H)}(\theta_n\Delta_n)^{\frac{1}{2}} = \Delta_n^{\theta\left(\frac{1}{2}-H\right)}, \\ \mathbb{E}\left[\sup_{t\leq T}|\mathbb{L}_3^n(t)|\right] &\lesssim \Delta_n^{-\frac{1}{2}}\left(\Delta_n^{\theta(1-H)}\right)^2 = \Delta_n^{2\theta(1-H)-\frac{1}{2}},\end{aligned}$$

proving the second property in (4.2.3) and thus the lemma. □

Proof of Lemma 3.2.7. Recall the definition of ξ_i^n in (3.2.15) and let

$$\xi_i^{n,\mathrm{dis}} = \int_{(i-\theta_n)\Delta_n}^{(i+L-1)\Delta_n} \rho_{(i-\theta_n)\Delta_n}\,\mathrm{d}W_s\,\underline{\Delta}_i^n g(s).$$

In a first step, we shall show that U^n can be approximated by

$$\overline{U}^n(t) = \Delta_n^{\frac{1}{2}} \sum_{i=\theta_n+1}^{[t/\Delta_n]-L+1} \Biggl\{ f\left(\frac{\sigma_{(i-1)\Delta_n}\underline{\Delta}_i^n B + \xi_i^{n,\mathrm{dis}}}{\Delta_n^H}\right) - f\left(\frac{\xi_i^{n,\mathrm{dis}}}{\Delta_n^H}\right) \\ - \mathbb{E}\left[f\left(\frac{\sigma_{(i-1)\Delta_n}\underline{\Delta}_i^n B + \xi_i^{n,\mathrm{dis}}}{\Delta_n^H}\right) - f\left(\frac{\xi_i^{n,\mathrm{dis}}}{\Delta_n^H}\right) \,\Big|\, \mathcal{F}_{i-\theta_n}^n \right] \Biggr\}.$$

By (3.2.12) and a size estimate as in (4.1.4), the difference $\xi_i^n - \xi_i^{n,\mathrm{dis}}$ is of size $(\theta_n\Delta_n)^{1/2}$. Together with (4.1.2) and (4.1.3), we further have that $\underline{\Delta}_i^n Y^{\mathrm{tr}} - \sigma_{(i-1)\Delta_n}\underline{\Delta}_i^n B - \xi_i^{n,\mathrm{dis}}$ is of size $\Delta_n + \sqrt{\Delta_n} + (\theta_n\Delta_n)^{1/2}$. By the mean-value theorem, these size bounds imply that

$$\mathbb{E}\left[\left| f\left(\frac{\underline{\Delta}_i^n Y^{\mathrm{tr}}}{\Delta_n^H}\right) - f\left(\frac{\sigma_{(i-1)\Delta_n}\underline{\Delta}_i^n B + \xi_i^{n,\mathrm{dis}}}{\Delta_n^H}\right) \right|^p + \left| f\left(\frac{\xi_i^n}{\Delta_n^H}\right) - f\left(\frac{\xi_i^{n,\mathrm{dis}}}{\Delta_n^H}\right) \right|^p \right]^{1/p} \lesssim (\theta_n\Delta_n)^{\frac{1}{2}}$$

for any $p > 0$. Moreover, the ith term in the definition of $\overline{U}^n(t)$ is $\mathcal{F}_{i+L-1}^n$-measurable with zero mean conditionally on $\mathcal{F}_{i-\theta_n}^n$. Therefore, employing a martingale size estimate as in (4.1.6), we obtain

$$\mathbb{E}\left[\sup_{t\leq T} \left| U^n(t) - \overline{U}^n(t) \right| \right] \lesssim \sqrt{\theta_n}(\theta_n\Delta_n)^{\frac{1}{2}} \leq \Delta_n^{\frac{1}{2}-\theta},$$

which converges to 0 by (3.2.14).

Next, because B and W are independent, we can apply Itô's formula with $\xi_i^{n,\mathrm{dis}}$ as starting point and write

$$\begin{aligned} &f\left(\frac{\sigma_{(i-1)\Delta_n}\underline{\Delta}_i^n B + \xi_i^{n,\mathrm{dis}}}{\Delta_n^H}\right) - f\left(\frac{\xi_i^{n,\mathrm{dis}}}{\Delta_n^H}\right) \\ &\quad = \Delta_n^{-H} \sum_{j,k=1}^{d} \sum_{\ell=1}^{L} \int_{(i+\ell-2)\Delta_n}^{(i+\ell-1)\Delta_n} \frac{\partial}{\partial z_{k\ell}} f\left(\frac{\underline{\Delta}Y_i^{n,\mathrm{dis}}(s)}{\Delta_n^H}\right) \sigma_{(i-1)\Delta_n}^{kj}\,\mathrm{d}B_s^j \\ &\quad\quad + \frac{1}{2}\Delta_n^{-2H} \sum_{k,k'=1}^{d} \sum_{\ell=1}^{L} \int_{(i+\ell-2)\Delta_n}^{(i+\ell-1)\Delta_n} \frac{\partial^2}{\partial z_{k\ell}\partial z_{k'\ell}} f\left(\frac{\underline{\Delta}Y_i^{n,\mathrm{dis}}(s)}{\Delta_n^H}\right) (\sigma\sigma^T)_{(i-1)\Delta_n}^{kk'}\,\mathrm{d}s, \end{aligned} \tag{4.2.4}$$

where $\underline{\Delta}Y_i^{n,\mathrm{dis}}(s) = \int_{(i-1)\Delta_n}^{s} \sigma_{(i-1)\Delta_n}\,\mathrm{d}B_r\,\mathbb{1}_i^n(r) + \xi_i^{n,\mathrm{dis}}$. Clearly, the stochastic integral is $\mathcal{F}_{i+L-1}^n$-measurable and conditionally centered given $\mathcal{F}_{i-1}^n$. Therefore, by a martingale size estimate, its contribution to $\overline{U}^n(t)$ is of magnitude $\Delta_n^{1/2-H}$, which is negligible because $H < \frac{1}{2}$. For the Lebesgue integral, we need to apply Itô's formula again and write

$$\begin{aligned} &\frac{\partial^2}{\partial z_{k\ell}\partial z_{k'\ell}} f\left(\frac{\underline{\Delta}Y_i^{n,\mathrm{dis}}(s)}{\Delta_n^H}\right) = \frac{\partial^2}{\partial z_{k\ell}\partial z_{k'\ell}} f\left(\frac{\xi_i^{n,\mathrm{dis}}}{\Delta_n^H}\right) \\ &\quad + \Delta_n^{-H} \sum_{j_2,k_2=1}^{d} \sum_{\ell_2=1}^{L} \int_{(i+\ell_2-2)\Delta_n}^{s\wedge(i+\ell_2-1)\Delta_n} \frac{\partial^3}{\partial z_{k\ell}\partial z_{k'\ell}\partial z_{k_2\ell_2}} f\left(\frac{\underline{\Delta}Y_i^{n,\mathrm{dis}}(r)}{\Delta_n^H}\right) \sigma_{(i-1)\Delta_n}^{k_2 j_2}\,\mathrm{d}B_r^{j_2} \\ &\quad + \frac{1}{2}\Delta_n^{-2H} \sum_{k_2,k_2'=1}^{d} \sum_{\ell_2=1}^{L} \int_{(i+\ell_2-2)\Delta_n}^{s\wedge(i+\ell_2-1)\Delta_n} \frac{\partial^4}{\partial z_{k\ell}\partial z_{k\ell}\partial z_{k_2\ell_2}\partial z_{k_2'\ell_2}} f\left(\frac{\underline{\Delta}Y_i^{n,\mathrm{dis}}(r)}{\Delta_n^H}\right) (\sigma\sigma^T)_{(i-1)\Delta_n}^{k_2 k_2'}\,\mathrm{d}r. \end{aligned}$$

By the same reason as before, the stochastic integral [even after we plug it into the drift in (4.2.4)] is $\mathcal{F}^n_{i+L-1}$-measurable with zero $\mathcal{F}^n_{i-1}$-conditional mean and therefore negligible. The Lebesgue integral is essentially of the form as the one from (4.2.4). Because f is smooth, we can repeat this procedure as often as we want. What is important, is that we gain a net factor of Δ_n^{1-2H} in each step (we have Δ_n^{-2H} times a Lebesgue integral over an interval of length at most Δ_n). After N applications of Itô's formula, the final drift term yields a contribution of size $\sqrt{\theta_n}\Delta_n^{N(1-2H)}$ to $\overline{U}^n(t)$. As $\theta < \frac{1}{2}$, it suffices to take $N = N(H)+1$ to make this convergent to 0. □

Proof of Lemma 3.2.8. We begin by discretizing ρ on a finer scale and let

$$\Theta_i^n = \int_{(i-\theta_n)\Delta_n}^{(i+L-1)\Delta_n} \left(\sum_{k=1}^{Q} \rho_{(i-\theta_n^{(q-1)})\Delta_n} \mathbb{1}_{((i-\theta_n^{(q-1)})\Delta_n,(i-\theta_n^{(q)})\Delta_n)}(s) \right) \mathrm{d}W_s \, \underline{\Delta}_i^n g(s), \tag{4.2.5}$$

where $\theta_n^{(q)} = [\Delta_n^{-\theta^{(q)}}]$ for $q = 0,\dots,Q-1$, $\theta_n^{(Q)} = -(L-1)$ and the numbers $\theta^{(q)}$, $q = 0,\dots,Q-1$ for some $Q \in \mathbb{N}$, are chosen such that $\theta = \theta^{(0)} > \theta^{(q)} > \cdots > \theta^{(Q-1)} > \theta^{(Q)} = 0$ and

$$\theta^{(q)} > \frac{\gamma}{1-H}\theta^{(q-1)} - \frac{\gamma - \frac{1}{2}}{1-H}, \qquad q = 1,\dots,Q, \tag{4.2.6}$$

where γ describes the regularity of the volatility process $\rho^{(0)}$ in (3.2.12). Because $H < \frac{1}{2}$ and we can make γ arbitrarily close to $\frac{1}{2}$ if we want, there is no loss of generality to assume that $\gamma/(1-H) < 1$. In this case, the fact that a choice as in (4.2.6) is possible can be verified by solving the associated linear recurrence equation. Defining

$$\underline{\Delta}_i^n Y^{\mathrm{dis}} = \sigma_{(i-1)\Delta_n} \underline{\Delta}_i^n B + \Theta_i^n,$$

we will show in Lemma 4.2.1 below that

$$\Delta_n^{\frac{1}{2}} \sum_{i=\theta_n+1}^{[t/\Delta_n]-L+1} \left\{ \mathbb{E}\left[f\left(\frac{\underline{\Delta}_i^n Y^{\mathrm{tr}}}{\Delta_n^H} \right) \,\Big|\, \mathcal{F}^n_{i-\theta_n} \right] - \mathbb{E}\left[f\left(\frac{\underline{\Delta}_i^n Y^{\mathrm{dis}}}{\Delta_n^H} \right) \,\Big|\, \mathcal{F}^n_{i-\theta_n} \right] \right\} \xrightarrow{L^1} 0. \tag{4.2.7}$$

Next, we define another matrix $\Upsilon_i^{n,0} \in (\mathbb{R}^{d\times L})^2$ by

$$(\Upsilon_i^{n,0})_{k\ell,k'\ell'} = c((i-1)\Delta_n)\Delta_n^{1-2H} + \sum_{q=1}^{Q} \left(\rho_{(i-\theta_n^{(q-1)})\Delta_n} \rho^T_{(i-\theta_n^{(q-1)})\Delta_n} \right)_{kk'} \times \int_{(i-\theta_n)\Delta_n}^{(i+L-1)\Delta_n} \frac{\Delta^n_{i+\ell-1} g(s) \Delta^n_{i+\ell'-1} g(s)}{\Delta_n^{2H}} \mathbb{1}_{((i-\theta_n^{(q-1)})\Delta_n,(i-\theta_n^{(q)})\Delta_n)}(s) \, \mathrm{d}s. \tag{4.2.8}$$

Observe that this is the covariance matrix of $\underline{\Delta}_i^n Y^{\mathrm{tr}}/\Delta_n^H$ if all discretized values of c and ρ are deterministic. Also notice that the only difference to Υ_i^n are the discretization points for ρ. The next step is to show the convergence

$$\Delta_n^{\frac{1}{2}} \sum_{i=\theta_n+1}^{[t/\Delta_n]-L+1} \left\{ \mathbb{E}\left[f\left(\frac{\underline{\Delta}_i^n Y^{\mathrm{dis}}}{\Delta_n^H} \right) \,\Big|\, \mathcal{F}^n_{i-\theta_n} \right] - \mu_f\left(\mathbb{E}\left[\Upsilon_i^{n,0} \mid \mathcal{F}^n_{i-\theta_n} \right] \right) \right\} \xrightarrow{L^1} 0, \tag{4.2.9}$$

where μ_f is the mapping defined after Assumption (CLT). This will be achieved through successive conditioning in Lemma 4.2.2. Finally, as we show in Lemma 4.2.3, we have

$$\mathbb{E}\left[\sup_{t\leq T}\left|\Delta_n^{\frac{1}{2}}\sum_{i=\theta_n+1}^{[t/\Delta_n]-L+1}\left\{\mu_f\left(\mathbb{E}[\Upsilon_i^{n,0}|\mathcal{F}_{i-\theta_n}^n]\right)-\mu_f(\Upsilon_i^{n,0})\right\}\right|\right]\to 0 \tag{4.2.10}$$

and

$$\mathbb{E}\left[\sup_{t\leq T}\left|\Delta_n^{\frac{1}{2}}\sum_{i=\theta_n+1}^{[t/\Delta_n]-L+1}\left\{\mu_f(\Upsilon_i^{n,0})-\mu_f(\Upsilon^{n,i})\right\}\right]\to 0, \tag{4.2.11}$$

which completes the proof of the current lemma. □

Lemma 4.2.1. *The convergence* (4.2.7) *holds true.*

Proof. Denote the left-hand side by $\mathbb{Q}^n(t)$. By Taylor's theorem, we can write $\mathbb{Q}^n(t) = \mathbb{Q}_1^n(t) + \mathbb{Q}_2^n(t)$ with

$$\mathbb{Q}_1^n(t) = \Delta_n^{\frac{1}{2}}\sum_{i=\theta_n+1}^{[t/\Delta_n]-L+1}\sum_{|\chi|=1}\mathbb{E}\left[\partial^\chi f\left(\frac{\underline{\Delta}_i^n Y^{\mathrm{dis}}}{\Delta_n^H}\right)(\kappa_i^n)^\chi \;\Big|\; \mathcal{F}_{i-\theta_n}^n\right],$$

$$\mathbb{Q}_2^n(t) = \Delta_n^{\frac{1}{2}}\sum_{i=\theta_n+1}^{[t/\Delta_n]-L+1}\sum_{|\chi|=2}\frac{1}{\chi!}\mathbb{E}[\partial^\chi f(\overline{\kappa}_i^n)(\kappa_i^n)^\chi \mid \mathcal{F}_{i-\theta_n}^n],$$

where $\kappa_i^n = (\underline{\Delta}_i^n Y^{\mathrm{tr}} - \underline{\Delta}_i^n Y^{\mathrm{dis}})/\Delta_n^H$ and $\overline{\kappa}_i^n$ is some point on the line between $\underline{\Delta}_i^n Y^{\mathrm{tr}}/\Delta_n^H$ and $\underline{\Delta}_i^n Y^{\mathrm{dis}}/\Delta_n^H$. By definition,

$$\begin{aligned}(\kappa_i^n)_{k\ell} = {} & \frac{\Delta_{i+\ell-1}^n A^k}{\Delta_n^H} + \frac{1}{\Delta_n^H}\int_{(i+\ell-2)\Delta_n}^{(i+\ell-1)\Delta_n}\sum_{\ell'=1}^{d'}\left(\sigma_s^{k\ell'}-\sigma_{(i-1)\Delta_n}^{k\ell'}\right)\mathrm{d}B_s^{\ell'} \\ & + \sum_{q=1}^{Q}\int_{(i-\theta_n)\Delta_n}^{(i+L-1)\Delta_n}\frac{\Delta_{i+\ell-1}^n g(s)}{\Delta_n^H}\mathbb{1}_{((i-\theta_n^{(q-1)})\Delta_n,((i-\theta_n^{(q)})\Delta_n)}(s)\sum_{\ell'=1}^{d'}\left(\rho_s^{k\ell'}-\rho_{(i-\theta_n^{(q-1)})\Delta_n}^{k\ell'}\right)\mathrm{d}W_s^{\ell'}.\end{aligned} \tag{4.2.12}$$

Using Hölder's inequality, the estimates (4.1.2), (4.1.3) and (4.1.4) and polynomial growth assumption on $\partial^\chi f$, we see that $\underline{\Delta}_i^n Y^{\mathrm{dis}}/\Delta_n^H$ is of size one and, therefore, that

$$\mathbb{E}\left[\sup_{t\leq T}|\mathbb{Q}_2^n(t)|\right] \lesssim \Delta_n^{-\frac{1}{2}}\left(\Delta_n^{2(1-H)}+\Delta_n^{2(1-H)}+\sum_{q=1}^{Q}\Delta_n^{(1-\theta^{(q-1)})+2\theta^{(q)}(1-H)}\right)\to 0 \tag{4.2.13}$$

as $n\to\infty$ since $0<\theta^{(q)}<\frac{1}{2}$.

Next, we further split $\mathbb{Q}_1^n(t) = \mathbb{Q}_{11}^n(t)+\mathbb{Q}_{12}^n(t)+\mathbb{Q}_{13}^n(t)$ into three terms according to the decomposition (4.2.12). Using again the estimates (4.1.2) and (4.1.3), we see that both $\mathbb{Q}_{11}^n(t)$ and $\mathbb{Q}_{12}^n(t)$ are of size $\Delta_n^{-1/2+(1-H)} = \Delta_n^{1/2-H}$. We now tackle the term $\mathbb{Q}_{13}^n(t)$, which requires a more careful analysis. Here we need assumption (3.2.3) on the noise volatility process ρ. First, since the process $t\mapsto\int_0^t\widetilde{b}_s\,\mathrm{d}s$ satisfies a better regularity condition than (3.2.12), we may incorporate the drift term in $\rho^{(0)}$ for the remainder of the proof. Then we further decompose $\mathbb{Q}_{13}^n(t)$ into $\mathbb{R}_1^n(t)+\mathbb{R}_2^n(t)$ where $\mathbb{R}_1^n(t)$ and $\mathbb{R}_2^n(t)$ correspond to taking only $\rho^{(0)}$ and $t\mapsto\int_0^t\widetilde{\rho}_s\,\mathrm{d}\widetilde{W}_s$ instead of ρ, respectively. Using the estimate (4.1.4) and taking into account the regularity condition of $\rho^{(0)}$, we find that $\mathbb{R}_1^n(t)$ is of size

$$\sum_{q=1}^{Q} \Delta_n^{-\frac{1}{2}+\gamma(1-\theta^{(q-1)})+\theta^{(q)}(1-H)}, \tag{4.2.14}$$

which goes to 0 as $\theta^{(q)}$, $q = 0, \dots, Q$, are chosen according to the recursion formula (4.2.6).

Concerning $\mathbb{R}_2^n(t)$, we fix $\chi \in \mathbb{N}_0^{d\times L}$ such that $|\chi| = 1$ and $\chi_{k\ell} = 1$ for some k and ℓ. We then resolve to a final decomposition:

$$\Delta_n^{\frac{1}{2}} \sum_{i=\theta_n+1}^{[t/\Delta_n]-L+1} \mathbb{E}\Bigg[\partial^\chi f\left(\frac{\underline{\Delta}_i^n Y^{\mathrm{dis}}}{\Delta_n^H}\right) \sum_{q=1}^{Q} \int_{(i-\theta_n)\Delta_n}^{(i+L-1)\Delta_n} \frac{\Delta_{i+\ell-1}^n g(s)}{\Delta_n^H} \mathbb{1}_{((i-\theta_n^{(q-1)})\Delta_n,(i-\theta_n^{(q)})\Delta_n)}(s)$$
$$\times \sum_{\ell',\ell''=1}^{d} \left(\int_{(i-\theta_n^{(q-1)})\Delta_n}^{s} \widetilde{\rho}_r^{k,\ell',\ell''} \,\mathrm{d}\widetilde{W}_r^{\ell''}\right) \mathrm{d}W_s^{\ell'} \,\Bigg|\, \mathcal{F}_{i-\theta_n}^n\Bigg] = \mathbb{R}_{21}^{n,\chi}(t) + \mathbb{R}_{22}^{n,\chi}(t) + \mathbb{R}_{23}^{n,\chi}(t)$$

where

$$\mathbb{R}_{21}^{n,\chi}(t) = \Delta_n^{\frac{1}{2}} \sum_{i=\theta_n+1}^{[t/\Delta_n]-L+1} \sum_{\ell',\ell''=1}^{d} \mathbb{E}\Bigg[\partial^\chi f\left(\frac{\underline{\Delta}_i^n Y^{\mathrm{dis}}}{\Delta_n^H}\right) \sum_{q=1}^{Q} \int_{(i-\theta_n^{(q-1)})\Delta_n}^{(i-\theta_n^{(q)})\Delta_n} \frac{\Delta_{i+\ell-1}^n g(s)}{\Delta_n^H}$$
$$\times \int_{(i-\theta_n^{(q-1)})\Delta_n}^{s} \left(\widetilde{\rho}_r^{k,\ell',\ell''} - \widetilde{\rho}_{(i-\theta_n^{(q-1)})\Delta_n}^{k,\ell',\ell''}\right) \mathrm{d}\widetilde{W}_r^{\ell''} \,\mathrm{d}W_s^{\ell'} \,\Bigg|\, \mathcal{F}_{i-\theta_n}^n\Bigg],$$

$$\mathbb{R}_{22}^{n,\chi}(t) = \Delta_n^{\frac{1}{2}} \sum_{i=\theta_n+1}^{[t/\Delta_n]-L+1} \sum_{q=1}^{Q} \sum_{\ell',\ell''=1}^{d} \mathbb{E}\Bigg[\left\{\partial^\chi f\left(\frac{\underline{\Delta}_i^n Y^{\mathrm{dis}}}{\Delta_n^H}\right) - \partial^\chi f\left(\frac{\underline{\Delta}_i^n Y^{\mathrm{dis},q}}{\Delta_n^H}\right)\right\}$$
$$\times \int_{(i-\theta_n^{(q-1)})\Delta_n}^{(i-\theta_n^{(q)})\Delta_n} \frac{\Delta_{i+\ell-1}^n g(s)}{\Delta_n^H} \int_{(i-\theta_n^{(q-1)})\Delta_n}^{s} \widetilde{\rho}_{(i-\theta_n^{(q-1)})\Delta_n}^{k,\ell',\ell''} \,\mathrm{d}\widetilde{W}_r^{\ell''} \,\mathrm{d}W_s^{\ell'} \,\Bigg|\, \mathcal{F}_{i-\theta_n}^n\Bigg],$$

$$\mathbb{R}_{23}^{n,\chi}(t) = \Delta_n^{\frac{1}{2}} \sum_{i=\theta_n+1}^{[t/\Delta_n]-L+1} \sum_{q=1}^{Q} \sum_{\ell',\ell''=1}^{d} \mathbb{E}\Bigg[\partial^\chi f\left(\frac{\underline{\Delta}_i^n Y^{\mathrm{dis},q}}{\Delta_n^H}\right) \int_{(i-\theta_n^{(q-1)})\Delta_n}^{(i-\theta_n^{(q)})\Delta_n} \frac{\Delta_{i+\ell-1}^n g(s)}{\Delta_n^H}$$
$$\times \int_{(i-\theta_n^{(q-1)})\Delta_n}^{s} \widetilde{\rho}_{(i-\theta_n^{(q-1)})\Delta_n}^{k,\ell',\ell''} \,\mathrm{d}\widetilde{W}_r^{\ell''} \,\mathrm{d}W_s^{\ell'} \,\Bigg|\, \mathcal{F}_{i-\theta_n}^n\Bigg]$$

and

$$\underline{\Delta}_i^n Y^{\mathrm{dis},q} = \int_{(i-\theta_n^{(q-1)})\Delta_n}^{(i+L-1)\Delta_n} \rho_{(i-\theta_n^{(q-1)})\Delta_n} \,\mathrm{d}W_s \,\underline{\Delta}_i^n g(s).$$

We now use the BDG and Minkowski's integral inequality alternatingly to obtain for any $p \geq 2$,

$$\mathbb{E}\Bigg[\Bigg|\int_{(i-\theta_n)\Delta_n}^{(i+L-1)\Delta_n} \frac{\Delta_{i+\ell-1}^n g(s)}{\Delta_n^H} \mathbb{1}_{((i-\theta_n^{(q-1)})\Delta_n,(i-\theta_n^{(q)})\Delta_n)}(s)$$
$$\times \left(\int_{(i-\theta_n^{(q-1)})\Delta_n}^{s} \left(\widetilde{\rho}_r^{k,\ell',\ell''} - \widetilde{\rho}_{(i-\theta_n^{(q-1)})\Delta_n}^{k,\ell',\ell''}\right) \mathrm{d}\widetilde{W}_r^{\ell''}\right) \mathrm{d}W_s^{\ell'}\Bigg|^p\Bigg]^{\frac{1}{p}}$$

$$\lesssim \Bigg(\int_{(i-\theta_n)\Delta_n}^{(i+L-1)\Delta_n} \frac{\Delta_{i+\ell-1}^n g(s)^2}{\Delta_n^{2H}} \mathbb{1}_{((i-\theta_n^{(q-1)})\Delta_n,(i-\theta_n^{(q)})\Delta_n)}(s)$$

$$\times \mathbb{E}\left[\left|\int_{(i-\theta_n^{(q-1)})\Delta_n}^{s} \left(\widetilde{\rho}_r^{k,\ell',\ell''} - \widetilde{\rho}_{(i-\theta_n^{(q-1)})\Delta_n}^{k,\ell',\ell''}\right) \mathrm{d}\widetilde{W}_r^{\ell''}\right|^p\right]^{\frac{2}{p}} \mathrm{d}s\Bigg)^{\frac{1}{2}}$$

$$\lesssim \Bigg(\int_{(i-\theta_n)\Delta_n}^{(i+L-1)\Delta_n} \frac{\Delta_{i+\ell-1}^n g(s)^2}{\Delta_n^{2H}} \mathbb{1}_{((i-\theta_n^{(q-1)})\Delta_n,(i-\theta_n^{(q)})\Delta_n)}(s)$$

$$\times \left(\int_{(i-\theta_n^{(q-1)})\Delta_n}^{s} \mathbb{E}\left[\left|\widetilde{\rho}_r^{k,\ell',\ell''} - \widetilde{\rho}_{(i-\theta_n^{(q-1)})\Delta_n}^{k,\ell',\ell''}\right|^p\right]^{\frac{2}{p}} \mathrm{d}r\right) \mathrm{d}s\Bigg)^{\frac{1}{2}}$$

$$\lesssim (\theta_n^{(q-1)}\Delta_n)^{\frac{1}{2}(1+2\varepsilon')} \left(\int_{(i-\theta_n^{(q-1)})\Delta_n}^{(i-\theta_n^{(q)})\Delta_n} \frac{\Delta_{i+\ell-1}^n g(s)^2}{\Delta_n^{2H}} \,\mathrm{d}s\right)^{\frac{1}{2}} \lesssim \Delta_n^{(\frac{1}{2}+\varepsilon')(1-\theta^{(q-1)})+\theta^{(q)}(1-H)},$$

where ε' is as in (3.2.5). We thus infer that $\mathbb{R}_{21}^{n,\chi}(t)$ is of size

$$\sum_{q=1}^{Q} \Delta_n^{-\frac{1}{2}+(\frac{1}{2}+\varepsilon')(1-\theta^{(q-1)})+\theta^{(q)}(1-H)}.$$

This is almost the same as (4.2.14); the only difference is that γ is replaced by $\frac{1}{2}+\varepsilon'$. Since we can assume without loss of generality that $\frac{1}{2}+\varepsilon' < \gamma$, the formula (4.2.6) implies that we have $-\frac{1}{2}+(\frac{1}{2}+\varepsilon')(1-\theta^{(q-1)})+\theta^{(q)}(1-H) > 0$ for all $q=1,\dots,Q$, which means that $\mathbb{R}_{21}^{n,\chi}(t)$ is asymptotically negligible.

Next, using Lemma 4.4.1 (3) and a similar estimate to the previous display, we see that $(\Theta_i^n - \underline{\Delta}_i^n Y^{\mathrm{dis},q})/\Delta_n^H$ is of size $\Delta_n^{\theta^{(q-1)}(1-H)} + \Delta_n^{(1-\theta^{(q-1)})/2}$. Hence, with the two estimates (4.1.2) and (4.1.3) at hand, we deduce that $\mathbb{R}_{22}^{n,\chi}(t)$ is of size

$$\sum_{q=1}^{Q} \Delta_n^{-\frac{1}{2}}(\Delta_n^{\frac{1}{2}-H} + \Delta_n^{\theta^{(q-1)}(1-H)} + \Delta_n^{\frac{1}{2}(1-\theta^{(q-1)})})\Delta_n^{\theta^{(q)}(1-H)+\frac{1}{2}(1-\theta^{(q-1)})}$$

$$\leq \sum_{q=1}^{Q}\left(\Delta_n^{\frac{1}{2}-H-(\gamma-\frac{1}{2})(1-\theta^{(q-1)})} + \Delta_n^{(\gamma+\frac{1}{2}-H)\theta^{(q-1)}-(\gamma-\frac{1}{2})} + \Delta_n^{\theta^{(q)}(1-H)+(\frac{1}{2}-\theta^{(q-1)})}\right).$$

The last term clearly goes to 0 because $\theta^{(q-1)} \leq \theta < \frac{1}{2}$ by (3.2.14). Without loss of generality, we can assume that $\gamma > \frac{1}{2}$ is sufficiently close to $\frac{1}{2}$ such that the first term is negligible as well. With this particular value, we then make sure that

$$\frac{\gamma-\frac{1}{2}}{\gamma+\frac{1}{2}-H} < \theta^{(Q-1)} < \frac{\gamma-\frac{1}{2}}{\gamma},$$

which, on the one hand, is in line with (4.2.6) and, on the other hand, guarantees that the second term in the preceding display tends to 0 for all $q=1,\dots,Q$.

Finally, to compute $\mathbb{R}_{23}^{n,\chi}(t)$, we first condition on $\mathcal{F}_{i-\theta_n^{(q-1)}}^n$. Because f is even and $\underline{\Delta}_i^n Y^{\mathrm{dis},q}/\Delta_n^H$ has a centered normal distribution given $\mathcal{F}_{i-\theta_n^{(q-1)}}^n$, if follows that $\partial^\chi f(\Theta_i^{n,q}/\Delta_n^H)$ is an element of the direct sum of all odd-order Wiener chaoses. At the same time, the double stochastic integrals in $\mathbb{R}_{23}^{n,\chi}(t)$ belongs to the second Wiener chaos; see [82, Proposition 1.1.4]. Since Wiener chaoses are mutually orthogonal, we obtain $\mathbb{R}_{23}^{n,\chi}(t)=0$. Because this reasoning is valid for all multi-indices with $|\chi|=1$, we have shown that $\mathbb{R}_2^n(t)$ is asymptotically negligible. □

Lemma 4.2.2. *The convergence* (4.2.9) *holds true.*

Proof. For $r = 0, \dots, Q$ (where Q is as in Lemma 4.2.1), define

$$\mathbb{Y}_i^{n,r} = \int_{(i-\theta_n)\Delta_n}^{(i+L-1)\Delta_n} \left(\sum_{q=1}^{r} \rho_{(i-\theta_n^{(q-1)})\Delta_n} \mathbb{1}_{((i-\theta_n^{(q-1)})\Delta_n,(i-\theta_n^{(q)})\Delta_n)}(s) \right) \mathrm{d}W_s \, \frac{\underline{\Delta}_i^n g(s)}{\Delta_n^H},$$

$$\begin{aligned}\Upsilon_i^{n,r} &= c((i-1)\Delta_n)\Delta_n^{1-2H} + \sum_{q=r+1}^{Q} \rho_{(i-\theta_n^{(q-1)})\Delta_n} \rho^T_{(i-\theta_n^{(q-1)})\Delta_n} \\ &\quad \times \int_{(i-\theta_n)\Delta_n}^{(i+L-1)\Delta_n} \frac{\underline{\Delta}_i^n g(s)^T \underline{\Delta}_i^n g(s)}{\Delta_n^{2H}} \mathbb{1}_{((i-\theta_n^{(q-1)})\Delta_n,(i-\theta_n^{(q)})\Delta_n)}(s) \, \mathrm{d}s\end{aligned}$$

Note that $\mathbb{Y}_i^{n,r} \in \mathbb{R}^{d\times L}$, $\Upsilon_i^{n,r} \in \mathbb{R}^{(d\times L)\times(d\times L)}$ and that $\mathbb{Y}_i^{n,Q} = \Theta_i^n/\Delta_n^H$ from (4.2.5). In order to show (4.2.9), we need the following approximation result for each $r = 1, \dots, Q-1$:

$$\Delta_n^{\frac{1}{2}} \sum_{i=\theta_n+1}^{[t/\Delta_n]-L+1} \mathbb{E}\left[\mu_{f(\mathbb{Y}_i^{n,r}+\cdot)} \left(\mathbb{E}[\Upsilon_i^{n,r} \mid \mathcal{F}_{i-\theta_n^{(r)}}] \right) - \mu_{f(\mathbb{Y}_i^{n,r}+\cdot)} \left(\mathbb{E}[\Upsilon_i^{n,r} \mid \mathcal{F}_{i-\theta_n^{(r-1)}}] \right) \,\middle|\, \mathcal{F}^n_{i-\theta_n} \right] \overset{L^1}{\Longrightarrow} 0. \tag{4.2.15}$$

Let us proceed with the proof of (4.2.9), taking the previous statement for granted. Defining

$$\overline{\mathbb{Y}}_i^{n} = \int_{(i-\theta_n)\Delta_n}^{(i-1)\Delta_n} \left(\sum_{q=1}^{Q} \rho_{(i-\theta_n^{(q-1)})\Delta_n} \mathbb{1}_{((i-\theta_n^{(q-1)})\Delta_n,(i-\theta_n^{(q)})\Delta_n)}(s) \right) \mathrm{d}W_s \, \frac{\underline{\Delta}_i^n g(s)}{\Delta_n^H},$$

we can use the tower property of conditional expectation to derive

$$\begin{aligned}
\mathbb{E}\left[f\left(\frac{\underline{\Delta}_i^n Y^{\mathrm{dis}}}{\Delta_n^H} \right) \,\middle|\, \mathcal{F}_{i-\theta_n} \right] &= \mathbb{E}\left[\mathbb{E}\left[f\left(\frac{\underline{\Delta}_i^n Y^{\mathrm{dis}}}{\Delta_n^H} \right) \,\middle|\, \mathcal{F}_{i-1} \right] \,\middle|\, \mathcal{F}_{i-\theta_n} \right] \\
&= \mathbb{E}\Bigg[\mu_{f(\overline{\mathbb{Y}}_i^n+\cdot)} \Bigg(c((i-1)\Delta_n)\Delta_n^{1-2H} \\
&\qquad + \rho_{(i-\theta_n^{(Q-1)})\Delta_n} \rho^T_{(i-\theta_n^{(Q-1)})\Delta_n} \int_{(i-1)\Delta_n}^{(I+L-1)\Delta_n} \frac{\underline{\Delta}_i^n g(s)^T \underline{\Delta}_i^n g(s)}{\Delta_n^{2H}} \, \mathrm{d}s \Bigg) \,\Bigg|\, \mathcal{F}_{i-\theta_n} \Bigg] \\
&= \mathbb{E}\Bigg[\mathbb{E}\Bigg[\mu_{f(\overline{\mathbb{Y}}_i^n+\cdot)} \Bigg(c((i-1)\Delta_n)\Delta_n^{1-2H} \\
&\qquad + \rho_{(i-\theta_n^{(Q-1)})\Delta_n} \rho^T_{(i-\theta_n^{(Q-1)})\Delta_n} \int_{(i-1)\Delta_n}^{(i+L-1)\Delta_n} \frac{\underline{\Delta}_i^n g(s)^T \underline{\Delta}_i^n g(s)}{\Delta_n^{2H}} \, \mathrm{d}s \Bigg) \,\Bigg|\, \mathcal{F}_{i-\theta_n^{(Q-1)}} \Bigg] \,\Bigg|\, \mathcal{F}_{i-\theta_n} \Bigg] \\
&= \mathbb{E}\left[\mu_{f(\mathbb{Y}_i^{n,Q-1}+\cdot)} \left(\Upsilon_i^{n,Q-1} \right) \,\middle|\, \mathcal{F}_{i-\theta_n} \right].
\end{aligned}$$

Thanks to (4.2.15) and in view of (4.2.9), we can replace $\Upsilon_i^{n,Q-1} = \mathbb{E}[\Upsilon_i^{n,Q-1} \mid \mathcal{F}_{i-\theta_n^{(Q-1)}}]$ in the last line by $\mathbb{E}[\Upsilon_i^{n,Q-1} \mid \mathcal{F}_{i-\theta_n^{(Q-2)}}]$, because the error resulting from this approximation is asymptotically negligible. We can then further compute

$$\begin{aligned}
&\mathbb{E}\left[\mathbb{E}\left[\mu_{f(\mathbb{Y}_i^{n,Q-1}+\cdot)} \left(\mathbb{E}\left[\Upsilon_i^{n,Q-1} \mid \mathcal{F}_{i-\theta_n^{(Q-2)}} \right] \right) \,\middle|\, \mathcal{F}_{i-\theta_n^{(Q-2)}} \right] \,\middle|\, \mathcal{F}_{i-\theta_n} \right] \\
&\qquad = \mathbb{E}\left[\mu_{f(\mathbb{Y}_i^{n,Q-2}+\cdot)} \left(\mathbb{E}\left[\Upsilon_i^{n,Q-2} \mid \mathcal{F}_{i-\theta_n^{(Q-2)}} \right] \right) \,\middle|\, \mathcal{F}_{i-\theta_n} \right] \\
&\qquad = \mathbb{E}\left[\mathbb{E}\left[\mu_{f(\mathbb{Y}_i^{n,Q-2}+\cdot)} \left(\mathbb{E}\left[\Upsilon_i^{n,Q-2} \mid \mathcal{F}_{i-\theta_n^{(Q-2)}} \right] \right) \,\middle|\, \mathcal{F}_{i-\theta_n^{(Q-3)}} \right] \,\middle|\, \mathcal{F}_{i-\theta_n} \right].
\end{aligned} \tag{4.2.16}$$

Again by (4.2.15), we may replace $\mathbb{E}[\Upsilon_i^{n,Q-2} \mid \mathcal{F}_{i-\theta_n^{(Q-2)}}]$ by $\mathbb{E}[\Upsilon_i^{n,Q-2} \mid \mathcal{F}_{i-\theta_n^{(Q-3)}}]$ in (4.2.16). Repeating this procedure Q times, we obtain at the end $\mu_{f(\mathbb{Y}_i^{n,0}+\cdot)}(\mathbb{E}[\Upsilon_i^{n,0} \mid \mathcal{F}_{i-\theta_n^{(0)}}]) = \mu_f(\mathbb{E}[\Upsilon_i^{n,0} \mid \mathcal{F}_{i-\theta_n}])$, which shows (4.2.9).

So it remains to prove (4.2.15). For the function $(u,v) \mapsto \mu_{f(u+\cdot)}(v)$, we use $\partial^{\chi'}$ to denote differentiation with respect to u (where $\chi' \in \mathbb{N}_0^{d\times L}$) and $\partial^{\chi''}$ to denote differentiation with respect to v (where $\chi'' \in \mathbb{N}_0^{(d\times L)\times(d\times L)}$). A Taylor expansion of $\mu_{f(\mathbb{Y}_i^{n,r}+\cdot)}(\mathbb{E}[\Upsilon_i^{n,r} \mid \mathcal{F}_{i-\theta_n^{(r)}}])$ around the point $(\mathbb{Y}_i^{n,r}, \mathbb{E}[\Upsilon_i^{n,r} \mid \mathcal{F}_{i-\theta_n^{(r-1)}}])$ decomposes the difference inside $\mathbb{E}[\cdot \mid \mathcal{F}_{i-\theta_n}^n]$ in (4.2.15) into

$$\begin{aligned}
&\Delta_n^{\frac{1}{2}} \sum_{i=\theta_n+1}^{[t/\Delta_n]-L+1} \sum_{|\chi''|=1} \mathbb{E}\Bigg[\partial^{\chi''}\mu_{f(\mathbb{Y}_i^{n,r}+\cdot)}(\mathbb{E}[\Upsilon_i^{n,r} \mid \mathcal{F}_{i-\theta_n^{(r-1)}}]) \\
&\qquad \times \left\{\mathbb{E}[\Upsilon_i^{n,r} \mid \mathcal{F}_{i-\theta_n^{(r)}}] - \mathbb{E}[\Upsilon_i^{n,r} \mid \mathcal{F}_{i-\theta_n^{(r-1)}}]\right\}^{\chi''} \Bigg| \mathcal{F}_{i-\theta_n}^n\Bigg] \\
&+ \Delta_n^{\frac{1}{2}} \sum_{i=\theta_n+1}^{[t/\Delta_n]-L+1} \sum_{|\chi''|=2} \frac{1}{\chi''!}\mathbb{E}\Bigg[\partial^{\chi''}\mu_{f(\mathbb{Y}_i^{n,r}+\cdot)}(\overline{v}_i^n) \\
&\qquad \times \left\{\mathbb{E}[\Upsilon_i^{n,r} \mid \mathcal{F}_{i-\theta_n^{(r)}}] - \mathbb{E}[\Upsilon_i^{n,r} \mid \mathcal{F}_{i-\theta_n^{(r-1)}}]\right\}^{\chi''} \Bigg| \mathcal{F}_{i-\theta_n}^n\Bigg]
\end{aligned} \tag{4.2.17}$$

with some $\overline{v}_i^n$ between $\mathbb{E}[\Upsilon_i^{n,r} \mid \mathcal{F}_{i-\theta_n^{(r)}}]$ and $\mathbb{E}[\Upsilon_i^{n,r} \mid \mathcal{F}_{i-\theta_n^{(r-1)}}]$. Write

$$\begin{aligned}
&\mathbb{E}\left[\rho_{(i-\theta_n^{(q-1)})\Delta_n}\rho^T_{(i-\theta_n^{(q-1)})\Delta_n} \,\middle|\, \mathcal{F}^n_{i-\theta_n^{(r)}}\right] - \mathbb{E}\left[\rho_{(i-\theta_n^{(q-1)})\Delta_n}\rho^T_{(i-\theta_n^{(q-1)})\Delta_n} \,\middle|\, \mathcal{F}^n_{i-\theta_n^{(r-1)}}\right] \\
&\quad = \mathbb{E}\left[\rho_{(i-\theta_n^{(q-1)})\Delta_n}\rho^T_{(i-\theta_n^{(q-1)})\Delta_n} - \rho_{(i-\theta_n^{(r-1)})\Delta_n}\rho^T_{(i-\theta_n^{(r-1)})\Delta_n} \,\middle|\, \mathcal{F}^n_{i-\theta_n^{(r)}}\right] \\
&\qquad - \mathbb{E}\left[\rho_{(i-\theta_n^{(q-1)})\Delta_n}\rho^T_{(i-\theta_n^{(q-1)})\Delta_n} - \rho_{(i-\theta_n^{(r-1)})\Delta_n}\rho^T_{(i-\theta_n^{(r-1)})\Delta_n} \,\middle|\, \mathcal{F}^n_{i-\theta_n^{(r-1)}}\right],
\end{aligned} \tag{4.2.18}$$

and note that, because of (3.2.11), (3.2.12) and the identity

$$xy - x_0y_0 = y_0(x-x_0) + x_0(y-y_0) + (x-x_0)(y-y_0), \tag{4.2.19}$$

the two conditional expectations on the right-hand side of (4.2.18) are both of size $(\theta_n^{(r-1)}\Delta_n)^{1/2}$. The same holds true if we replace $\rho_{(i-\theta_n^{(q-1)})\Delta_n}$ by $\sigma_{(i-1)\Delta_n}$. Therefore, we deduce that

$$\mathbb{E}\left[\left\|\mathbb{E}[\Upsilon_i^{n,r} \mid \mathcal{F}_{i-\theta_n^{(r)}}] - \mathbb{E}[\Upsilon_i^{n,r} \mid \mathcal{F}_{i-\theta_n^{(r-1)}}]\right\|^p\right]^{\frac{1}{p}} \lesssim (\theta_n^{(r-1)}\Delta_n)^{\frac{1}{2}}. \tag{4.2.20}$$

As a consequence, the second expression in (4.2.17) is of size $\Delta_n^{-1/2}((\theta_n^{(r-1)}\Delta_n)^{1/2})^2 = \Delta_n^{1/2-\theta^{(r-1)}}$ which goes to 0 as $n\to\infty$ since all numbers $\theta^{(r)}$ are chosen to be smaller than $\frac{1}{2}$; see (4.2.6).

Next, we expand $\partial^{\chi}\mu_{f(\mathbb{Y}_i^{n,r}+\cdot)}(\mathbb{E}[\Upsilon_i^{n,r} \mid \mathcal{F}_{i-\theta_n^{(r-1)}}])$ around $(0, \mathbb{E}[\Upsilon_i^{n,r} \mid \mathcal{F}_{i-\theta_n^{(r-1)}}])$ and write the first expression in (4.2.17) as $\mathbb{S}_1^n(t)+\mathbb{S}_2^n(t)+\mathbb{S}_3^n(t)$, where

$$\mathbb{S}_1^n(t)=\Delta_n^{\frac{1}{2}}\sum_{i=\theta_n+1}^{[t/\Delta_n]-L+1}\sum_{|\chi''|=1}\mathbb{E}\Bigg[\partial^{\chi''}\mu_f(\mathbb{E}[\Upsilon_i^{n,r}\mid\mathcal{F}_{i-\theta_n^{(r-1)}}])\times\left\{\mathbb{E}[\Upsilon_i^{n,r}\mid\mathcal{F}_{i-\theta_n^{(r)}}]-\mathbb{E}[\Upsilon_i^{n,r}\mid\mathcal{F}_{i-\theta_n^{(r-1)}}]\right\}^{\chi''}\Bigg|\,\mathcal{F}^n_{i-\theta_n}\Bigg],$$

$$\mathbb{S}_2^n(t)=\Delta_n^{\frac{1}{2}}\sum_{i=\theta_n+1}^{[t/\Delta_n]-L+1}\sum_{|\chi'|=1,\,|\chi''|=1}\mathbb{E}\Bigg[\partial^{\chi'}\partial^{\chi''}\mu_f(\mathbb{E}[\Upsilon_i^{n,r}\mid\mathcal{F}_{i-\theta_n^{(r-1)}}])\times(\mathbb{Y}_i^{n,r})^{\chi'}\left\{\mathbb{E}[\Upsilon_i^{n,r}\mid\mathcal{F}_{i-\theta_n^{(r)}}]-\mathbb{E}[\Upsilon_i^{n,r}\mid\mathcal{F}_{i-\theta_n^{(r-1)}}]\right\}^{\chi''}\Bigg|\,\mathcal{F}^n_{i-\theta_n}\Bigg],$$

$$\mathbb{S}_3^n(t)=\Delta_n^{\frac{1}{2}}\sum_{i=\theta_n+1}^{[t/\Delta_n]-L+1}\sum_{|\chi'|=2,\,|\chi''|=1}\frac{1}{\chi'!}\mathbb{E}\Bigg[\partial^{\chi'}\partial^{\chi''}\mu_{f(\varsigma_i^n+\cdot)}(\mathbb{E}[\Upsilon_i^{n,r}\mid\mathcal{F}_{i-\theta_n^{(r-1)}}])\times(\mathbb{Y}_i^{n,r})^{\chi'}\left\{\mathbb{E}[\Upsilon_i^{n,r}\mid\mathcal{F}_{i-\theta_n^{(r)}}]-\mathbb{E}[\Upsilon_i^{n,r}\mid\mathcal{F}_{i-\theta_n^{(r-1)}}]\right\}^{\chi''}\Bigg|\,\mathcal{F}^n_{i-\theta_n}\Bigg],$$

and ς_i^n is a point between 0 and $\mathbb{Y}_i^{n,r}$. Observe that $\partial^{\chi''}\mu_f(\mathbb{E}[\Upsilon_i^{n,r}\mid\mathcal{F}_{i-\theta_n^{(r-1)}}])$ is $\mathcal{F}_{i-\theta_n^{(r-1)}}$-measurable and that the $\mathcal{F}^n_{i-\theta_n^{(r-1)}}$-conditional expectation of $\mathbb{E}[\Upsilon_i^{n,r}\mid\mathcal{F}_{i-\theta_n^{(r)}}]-\mathbb{E}[\Upsilon_i^{n,r}\mid\mathcal{F}_{i-\theta_n^{(r-1)}}]$ is 0. Hence,

$$\mathbb{E}\left[\partial^{\chi''}\mu_f(\mathbb{E}[\Upsilon_i^{n,r}\mid\mathcal{F}_{i-\theta_n^{(r-1)}}])\left\{\mathbb{E}[\Upsilon_i^{n,r}\mid\mathcal{F}_{i-\theta_n^{(r)}}]-\mathbb{E}[\Upsilon_i^{n,r}\mid\mathcal{F}_{i-\theta_n^{(r-1)}}]\right\}^{\chi''}\Bigg|\,\mathcal{F}^n_{i-\theta_n}\right]=0$$

and it follows that $\mathbb{S}_1^n(t)$ vanishes.

By [30, Equation (D.46)], given $|\chi'|=|\chi''|=1$, we can find $\alpha,\beta,\gamma\in\{1,\dots,d\}\times\{1,\dots,L\}$ such that

$$\partial^{\chi'}\partial^{\chi''}\mu_{f(u+\cdot)}(v)=\frac{\partial\mu_{f(u+\cdot)}}{\partial u_\gamma\partial v_{\alpha,\beta}}(v)=\frac{1}{2^{\mathbb{1}_{\{\alpha=\beta\}}}}\mu_{\partial_{\alpha\beta\gamma}f(u+\cdot)}(v).$$

If $u=0$, since f has odd third derivatives, we have that $\mu_{\partial_{\alpha\beta\gamma}f}(v)=0$. Therefore, the $\partial^{\chi'}\partial^{\chi''}\mu_f$-expression in $\mathbb{S}_2^n(t)$ is equal to 0 and $\mathbb{S}_2^n(t)$ vanishes as well. Finally, we use the generalized Hölder inequality as well as the estimates (4.2.20) and (4.1.4) to see that

$$\begin{aligned}\mathbb{E}\left[\sup_{t\le T}|\mathbb{S}_3^n(t)|\right]&\lesssim\Delta_n^{\frac{1}{2}}\sum_{i=\theta_n+1}^{[T/\Delta_n]-L+1}\mathbb{E}\left[\|\mathbb{Y}_i^{n,r}\|^4\right]^{\frac{1}{2}}\mathbb{E}\left[\left\|\mathbb{E}[\Upsilon_i^{n,r}\mid\mathcal{F}_{i-\theta_n^{(r)}}]-\mathbb{E}[\Upsilon_i^{n,r}\mid\mathcal{F}_{i-\theta_n^{(r-1)}}]\right\|^4\right]^{\frac{1}{4}}\\&\lesssim\Delta_n^{-\frac{1}{2}}\Delta_n^{2\theta^{(r)}(1-H)}(\theta_n^{(r-1)}\Delta_n)^{\frac{1}{2}}.\end{aligned}$$

This converges to 0 as $n\to\infty$ if $2\theta^{(r)}(1-H)-\frac{1}{2}\theta^{(r-1)}>0$ for all $r=1,\dots,Q-1$, which is equivalent to $\theta^{(r)}>\frac{1}{4(1-H)}\theta^{(r-1)}$. Because $\frac{1}{4(1-H)}<1$, this condition means that $\theta^{(r)}$ must not decrease to 0 too fast. By adding more intermediate θ's between $\theta^{(0)}$ and $\theta^{(Q-1)}$ if necessary, which does no harm to (4.2.6), we can make sure that this is satisfied. □

Lemma 4.2.3. *The convergences* (4.2.10) *and* (4.2.11) *hold true.*

Proof. We perform a Taylor expansion of $\mu_f(\Upsilon_i^{n,0})$ around $\mathbb{E}[\Upsilon_i^{n,0} \mid \mathcal{F}_{i-\theta_n}^n]$ and we write

$$\begin{aligned}\mu_f(\Upsilon_i^{n,0}) - \mu_f(\mathbb{E}[\Upsilon_i^{n,0} \mid \mathcal{F}_{i-\theta_n}^n]) &= \sum_{|\chi|=1} \partial^\chi \mu_f\left(\mathbb{E}[\Upsilon_i^{n,0} \mid \mathcal{F}_{i-\theta_n}^n]\right)\left(\Upsilon_i^{n,0} - \mathbb{E}[\Upsilon_i^{n,0} \mid \mathcal{F}_{i-\theta_n}^n]\right)^\chi \\ &\quad + \sum_{|\chi|=2} \frac{1}{\chi!}\partial^\chi \mu_f(\widetilde{\upsilon}_i^n)\left(\Upsilon_i^{n,0} - \mathbb{E}[\Upsilon_i^{n,0} \mid \mathcal{F}_{i-\theta_n}^n]\right)^\chi\end{aligned} \tag{4.2.21}$$

with some $\widetilde{\upsilon}_i^n$ on the line between $\Upsilon_i^{n,0}$ and $\mathbb{E}[\Upsilon_i^{n,0} \mid \mathcal{F}_{i-\theta_n}^n]$. The expression $\Upsilon_i^{n,0} - \mathbb{E}[\Upsilon_i^{n,0} \mid \mathcal{F}_{i-\theta_n}^n]$ contains the difference

$$\rho_{(i-\theta_n^{(q-1)})\Delta_n}\rho^T_{(i-\theta_n^{(q-1)})\Delta_n} - \mathbb{E}\left[\rho_{(i-\theta_n^{(q-1)})\Delta_n}\rho^T_{(i-\theta_n^{(q-1)})\Delta_n} \,\middle|\, \mathcal{F}_{i-\theta_n}^n\right]$$

and a similar one with $\rho_{(i-\theta_n^{(q-1)})\Delta_n}$ replaced by $\sigma_{(i-1)\Delta_n}$. Inserting $\rho\rho^T$ or $\sigma\sigma^T$ at $(i-\theta_n)\Delta_n$ artificially [cf. (4.2.18)], we can use (4.2.19) and the assumptions (3.2.11) and (3.2.12) on both ρ and σ to find that the term in the display above is of size at most $(\theta_n\Delta_n)^{1/2}$. This immediately leads to the bound $\mathbb{E}[\|\Upsilon_i^{n,0} - \mathbb{E}[\Upsilon_i^{n,0} \mid \mathcal{F}_{i-\theta_n}^n]\|^2]^{1/2} \lesssim (\theta_n\Delta_n)^{1/2}$, which in turn shows that the second-order term in (4.2.21) is $o_{\mathbb{P}}(\sqrt{\Delta_n})$ by (3.2.14). Therefore, in (4.2.10), it remains to consider

$$\sqrt{\Delta_n}\sum_{i=\theta_n+1}^{[t/\Delta_n]-L+1}\sum_{|\chi|=1}\partial^\chi\mu_f\left(\mathbb{E}[\Upsilon_i^{n,0} \mid \mathcal{F}_{i-\theta_n}^n]\right)\left(\Upsilon_i^{n,0} - \mathbb{E}[\Upsilon_i^{n,0} \mid \mathcal{F}_{i-\theta_n}^n]\right)^\chi.$$

For each i, the $\sum_{|\chi|=1}$-expression is $\mathcal{F}_i^n$-measurable and has a vanishing conditional expectation given $\mathcal{F}_{i-\theta_n}^n$. We can therefore use a martingale size estimate of the type (4.1.6) to show that

$$\mathbb{E}\left[\sup_{t\leq T}\left|\sqrt{\Delta_n}\sum_{i=\theta_n+1}^{[t/\Delta_n]-L+1}\sum_{|\chi|=1}\partial^\chi\mu_f\left(\mathbb{E}[\Upsilon_i^{n,0} \mid \mathcal{F}_{i-\theta_n}^n]\right)\left(\Upsilon_i^{n,0} - \mathbb{E}[\Upsilon_i^{n,0} \mid \mathcal{F}_{i-\theta_n}^n]\right)^\chi\right|\right] \lesssim \sqrt{\theta_n}(\theta_n\Delta_n)^{\frac{1}{2}},$$

which tends to 0 by (3.2.14). This proves (4.2.10).

For (4.2.11), recall $\Upsilon^{n,i}$ in (3.2.18) and define $\Delta_i^n\Upsilon = \Upsilon^{n,i} - \Upsilon_i^{n,0}$. Then

$$\begin{aligned}(\Delta_i^n\Upsilon)_{k\ell,k'\ell'} &= \sum_{q=1}^{Q}\left(\rho_{(i-1)\Delta_n}\rho^T_{(i-1)\Delta_n} - \rho_{(i-\theta_n^{(q-1)})\Delta_n}\rho^T_{(i-\theta_n^{(q-1)})\Delta_n}\right)_{kk'} \\ &\quad \times \int_{(i-\theta_n)\Delta_n}^{(i+L-1)\Delta_n}\frac{\Delta_{i+\ell-1}^n g(s)\Delta_{i+\ell'-1}^n g(s)}{\Delta_n^{2H}}\mathbb{1}_{((i-\theta_n^{(q-1)})\Delta_n,(i-\theta_n^{(q)})\Delta_n)}(s)\,\mathrm{d}s\end{aligned}$$

for all $k,k'=1,\ldots,d$ and $\ell,\ell'=1,\ldots,L$. By Taylor's theorem,

$$\begin{aligned}&\Delta_n^{\frac{1}{2}}\sum_{i=\theta_n+1}^{[t/\Delta_n]-L+1}\left\{\mu_f(\Upsilon^{n,i}) - \mu_f(\Upsilon_i^{n,0})\right\} \\ &= \Delta_n^{\frac{1}{2}}\sum_{i=\theta_n+1}^{[t/\Delta_n]-L+1}\sum_{|\chi|=1}\partial^\chi\mu_f(\Upsilon^{n,i})(\Delta_i^n\Upsilon)^\chi + \Delta_n^{\frac{1}{2}}\sum_{i=\theta_n+1}^{[t/\Delta_n]-L+1}\sum_{|\chi|=2}\frac{1}{\chi!}\partial^\chi\mu_f(\widehat{\upsilon}_i^n)(\Delta_i^n\Upsilon)^\chi,\end{aligned} \tag{4.2.22}$$

where $\widehat{\upsilon}_i^n$ is some point between $\Upsilon^{n,i}$ and $\Upsilon_i^{n,0}$. Hölder's inequality together with the identity (4.2.19) as well as the moment and regularity assumptions on ρ show that the last sum in the above display is of size

$$\Delta_n^{-\frac{1}{2}}\sum_{q=1}^{Q}(\theta_n^{(q-1)}\Delta_n)\Delta_n^{4\theta^{(q)}(1-H)},$$

which goes to 0 as $n \to \infty$ [compare with (4.2.13)]. Next, recall the decomposition (3.2.3) of the noise volatility process ρ. As before, we incorporate the drift $t \mapsto \int_0^t \widetilde{b}_s \, \mathrm{d}s$ into $\rho^{(0)}$ so that $\rho = \rho^{(0)} + \rho^{(1)}$ with $\rho^{(1)}_t = \int_0^t \widetilde{\rho}_s \, \mathrm{d}\widetilde{W}_s$. By (4.2.19),

$$\begin{aligned}
&\rho^{k\ell}_{(i-1)\Delta_n}\rho^{k'\ell}_{(i-1)\Delta_n} - \rho^{k\ell}_{(i-\theta_n^{(q-1)})\Delta_n}\rho^{k'\ell}_{(i-\theta_n^{(q-1)})\Delta_n} \\
&\quad = \left(\rho^{k\ell}_{(i-\theta_n^{(q-1)})\Delta_n}\left\{\rho^{(0),k'\ell}_{(i-1)\Delta_n} - \rho^{(0),k'\ell}_{(i-\theta_n^{(q-1)})\Delta_n}\right\} + \rho^{k'\ell}_{(i-\theta_n^{(q-1)})\Delta_n}\left\{\rho^{(0),k\ell}_{(i-1)\Delta_n} - \rho^{(0),k\ell}_{(i-\theta_n^{(q-1)})\Delta_n}\right\}\right) \\
&\qquad + \left(\rho^{k\ell}_{(i-\theta_n^{(q-1)})\Delta_n}\left\{\rho^{(1),k'\ell}_{(i-1)\Delta_n} - \rho^{(1),k'\ell}_{(i-\theta_n^{(q-1)})\Delta_n}\right\} + \rho^{k'\ell}_{(i-\theta_n^{(q-1)})\Delta_n}\left\{\rho^{(1),k\ell}_{(i-1)\Delta_n} - \rho^{(1),k\ell}_{(i-\theta_n^{(q-1)})\Delta_n}\right\}\right) \\
&\qquad + \left(\rho^{k\ell}_{(i-1)\Delta_n} - \rho^{k\ell}_{(i-\theta_n^{(q-1)})\Delta_n}\right)\left(\rho^{k'\ell}_{(i-1)\Delta_n} - \rho^{k'\ell}_{(i-\theta_n^{(q-1)})\Delta_n}\right).
\end{aligned}$$

The remaining term $\Delta_n^{1/2}\sum_{i=\theta_n+1}^{[t/\Delta_n]-L+1}\sum_{|\chi|=1}\partial^\chi \mu_f(\Upsilon^{n,i})(\Delta_i^n\Upsilon)^\chi$ in (4.2.22) can thus be written as $\mathbb{T}_1^n(t)+\mathbb{T}_2^n(t)+\mathbb{T}_3^n(t)$ according to this decomposition. By Hölder's inequality and the moment and regularity assumptions on ρ, we see that $\mathbb{T}_3^n(t)$ is of size at most

$$\Delta_n^{-\frac{1}{2}}\sum_{q=1}^{Q}(\theta_n^{(q-1)}\Delta_n)\Delta_n^{2\theta^{(q)}(1-H)}, \tag{4.2.23}$$

which goes to 0 as $n \to \infty$ as we saw in (4.2.13). Similarly, thanks to the regularity property (3.2.12) of $\rho^{(0)}$, we further obtain

$$\mathbb{E}\left[\sup_{t\le T}\left|\mathbb{T}_1^n(t)\right|\right] \lesssim \Delta_n^{-\frac{1}{2}}\sum_{q=1}^{Q}(\theta_n^{(q-1)}\Delta_n)^\gamma\Delta_n^{2\theta^{(q)}(1-H)},$$

and this also goes to 0 as $n \to \infty$ by our choice (4.2.6) of the numbers $\theta_n^{(q-1)}$. Finally,

$$\begin{aligned}
\mathbb{T}_2^n(t) = \Delta_n^{\frac{1}{2}}&\sum_{i=\theta_n+1}^{[t/\Delta_n]-L+1}\sum_{q=1}^{Q}\sum_{|\chi|=1}\partial^\chi\mu_f(\Upsilon^{n,i}) \\
&\times\left\{\pi_{q-1}^{n,i}\int_{(i-\theta_n)\Delta_n}^{(i+L-1)\Delta_n}\frac{\underline{\Delta}_i^n g(s)^T\underline{\Delta}_i^n g(s)}{\Delta_n^{2H}}\mathbb{1}_{((i-\theta_n^{(q-1)})\Delta_n,(i-\theta_n^{(q)})\Delta_n)}(s)\,\mathrm{d}s\right\}^\chi,
\end{aligned}$$

where

$$\pi_q^{n,i} = \rho_{(i-\theta_n^{(q)})\Delta_n}\left(\rho^{(1)}_{(i-1)\Delta_n} - \rho^{(1)}_{(i-\theta_n^{(q)})\Delta_n}\right)^T + \left(\rho^{(1)}_{(i-1)\Delta_n} - \rho^{(1)}_{(i-\theta_n^{(q)})\Delta_n}\right)\rho^T_{(i-\theta_n^{(q)})\Delta_n}.$$

Let $\widetilde{\mathbb{T}}_2^n(t)$ be defined in the same way as $\mathbb{T}_2^n(t)$ except that in the previous display, $\Upsilon^{n,i}$ is replaced by $\widetilde{\Upsilon}^{n,i}_{q-1}$, obtained from $\Upsilon^{n,i}$ by substituting $(i-\theta_n^{(q-1)})\Delta_n$ for $(i-1)\Delta_n$ everywhere. By the generalized Hölder inequality and the regularity assumptions on ρ and σ, the difference

$\mathbb{T}_2^n(t) - \widetilde{\mathbb{T}}_2^n(t)$ is of the same size (4.2.23) as $\mathbb{T}_2^n(t)$ and hence asymptotically negligible. Next,

$$\begin{aligned}\widetilde{\mathbb{T}}_2^n(t) &= \sum_{q=1}^{Q} \Delta_n^{\frac{1}{2}} \sum_{i=\theta_n+1}^{[t/\Delta_n]-L+1} \sum_{|\chi|=1} \partial^\chi \mu_f(\widetilde{\Upsilon}^{n,i}_{q-1}) \bigg\{ \left(\pi^{n,i}_{q-1} - \mathbb{E}\left[\pi^{n,i}_{q-1} \,\middle|\, \mathcal{F}^n_{i-\theta_n^{(q-1)}}\right]\right) \\ &\qquad \times \int_{(i-\theta_n)\Delta_n}^{(i+L-1)\Delta_n} \frac{\underline{\Delta}_i^n g(s)^T \underline{\Delta}_i^n g(s)}{\Delta_n^{2H}} \mathbb{1}_{((i-\theta_n^{(q-1)})\Delta_n,(i-\theta_n^{(q)})\Delta_n)}(s)\,\mathrm{d}s \bigg\}^\chi \\ &\quad + \Delta_n^{\frac{1}{2}} \sum_{i=\theta_n+1}^{[t/\Delta_n]-L+1} \sum_{q=1}^{Q} \sum_{|\chi|=1} \partial^\chi \mu_f(\widetilde{\Upsilon}^{n,i}_{q-1}) \bigg\{ \mathbb{E}\left[\pi^{n,i}_{q-1} \,\middle|\, \mathcal{F}^n_{i-\theta_n^{(q-1)}}\right] \\ &\qquad \times \int_{(i-\theta_n)\Delta_n}^{(i+L-1)\Delta_n} \frac{\underline{\Delta}_i^n g(s)^T \underline{\Delta}_i^n g(s)}{\Delta_n^{2H}} \mathbb{1}_{((i-\theta_n^{(q-1)})\Delta_n,(i-\theta_n^{(q)})\Delta_n)}(s)\,\mathrm{d}s \bigg\}^\chi .\end{aligned} \tag{4.2.24}$$

For fixed q, the first term on the right-hand side of (4.2.24) is a sum where the ith summand is $\mathcal{F}^n_{i+L-1}$-measurable and has, by construction, a zero $\mathcal{F}^n_{i-\theta_n^{(q-1)}}$-conditional mean. By a martingale size estimate of the type (4.1.6), that first term is therefore of size

$$\sum_{q=1}^{Q} \sqrt{\theta_n^{(q-1)}} (\theta_n^{(q-1)}\Delta_n)^{\frac{1}{2}} \Delta_n^{2\theta^{(q)}(1-H)} = \sum_{q=1}^{Q} \Delta_n^{\frac{1}{2}-\theta_n^{(q-1)}+2\theta^{(q)}(1-H)} \to 0$$

as $n \to \infty$ since all $\theta_n^{(q)} < \frac{1}{2}$. Since stochastic integrals have mean 0,

$$\mathbb{E}\left[(\rho^{(1),k\ell}_{(i-1)\Delta_n} - \rho^{(1),k\ell}_{(i-\theta_n^{(q-1)})\Delta_n}) \,\middle|\, \mathcal{F}^n_{i-\theta_n^{(q-1)}}\right] = \sum_{m=1}^{d} \mathbb{E}\left[\int_{(i-\theta_n^{(q-1)})\Delta_n}^{(i-1)\Delta_n} \widetilde{\rho}^{k\ell m}_s \,\mathrm{d}\widetilde{W}^m_s \,\middle|\, \mathcal{F}^n_{i-\theta_n^{(q-1)}}\right] = 0,$$

which means that, in fact,

$$\begin{aligned}\mathbb{E}\left[\pi^{n,i}_{q-1} \,\middle|\, \mathcal{F}^n_{i-\theta_n^{(q-1)}}\right] &= \rho_{(i-\theta_n^{(q-1)})\Delta_n} \mathbb{E}\left[(\rho^{(1)}_{(i-1)\Delta_n} - \rho^{(1)}_{(i-\theta_n^{(q-1)})\Delta_n})^T \,\middle|\, \mathcal{F}^n_{i-\theta_n^{(q-1)}}\right] \\ &\quad + \mathbb{E}\left[(\rho^{(1)}_{(i-1)\Delta_n} - \rho^{(1)}_{(i-\theta_n^{(q-1)})\Delta_n}) \,\middle|\, \mathcal{F}^n_{i-\theta_n^{(q-1)}}\right] \rho^T_{(i-\theta_n^{(q-1)})\Delta_n} \\ &= 0.\end{aligned}$$

Therefore, $\mathbb{T}_2^n(t)$ is asymptotically negligible and the proof of (4.2.11) is complete. □

Proof of Lemma 3.2.9 . Recall the expressions $\mathbb{X}_1^n(t)$ and $\mathbb{X}_2^n(t)$ defined in (3.2.19). For a given multi-index $\chi \in \mathbb{N}_0^{(d\times L)\times(d\times L)}$, let

$$Q_\chi(x) = x^\chi, \qquad x \in \mathbb{R}^{(d\times L)\times(d\times L)}, \tag{4.2.25}$$

which is a polynomial of degree $|\chi|$. By Taylor's theorem,

$$\begin{aligned}\mathbb{X}_1^n(t) &= \Delta_n^{\frac{1}{2}} \sum_{i=\theta_n+1}^{[t/\Delta_n]-L+1} \sum_{j=1}^{N(H)} \sum_{|\chi|=j} \frac{1}{\chi!} \partial^\chi \mu_f(\pi((i-1)\Delta_n)) \sum_{k=1}^{j} \sum_{|\chi'|=k} \frac{\Delta_n^{(j-k)(1-2H)}}{\chi'!} \\ &\quad \times \partial^{\chi'} Q_\chi(c((i-1)\Delta_n)) \left\{\Upsilon^{n,i} - \pi((i-1)\Delta_n) - \Delta_n^{1-2H} c((i-1)\Delta_n)\right\}^{\chi'} .\end{aligned} \tag{4.2.26}$$

The key term in (4.2.26) is the expression in braces and we have [recall (3.2.6) and (3.1.9)]

$$\begin{aligned}
&\Upsilon^{n,i} - \pi((i-1)\Delta_n) - \Delta_n^{1-2H} c((i-1)\Delta_n) \\
&\quad = \rho_{(i-1)\Delta_n}\rho_{(i-1)\Delta_n}^T \left\{ \int_{(i-\theta_n)\Delta_n}^{(i+L-1)\Delta_n} \frac{\underline{\Delta}_i^n g(s)^T \underline{\Delta}_i^n g(s)}{\Delta_n^{2H}}\,\mathrm{d}s - \left(\Gamma^H_{|\ell-\ell'|}\right)_{\ell,\ell'=1}^{L,L} \right\} \\
&\quad = -\rho_{(i-1)\Delta_n}\rho_{(i-1)\Delta_n}^T \int_{-\infty}^{(i-\theta_n)\Delta_n} \frac{\underline{\Delta}_i^n g(s)^T \underline{\Delta}_i^n g(s)}{\Delta_n^{2H}}\,\mathrm{d}s,
\end{aligned} \tag{4.2.27}$$

because $\Gamma^H_{|\ell-\ell'|} = \Delta_n^{-2H} \int_{-\infty}^{\infty} \Delta_{i+\ell}^n g(s) \Delta_{i+\ell'}^n g(s)\,\mathrm{d}s$ by (4.4.2). The last integral is $\lesssim \Delta_n^{2\theta(1-H)}$ by Lemma 4.4.1 (3). Consequently, if we apply the generalized Hölder inequality to (4.2.26) and take into account the moment conditions (3.2.11) on ρ and σ, we obtain

$$\mathbb{E}\left[\sup_{t\leq T} |\mathbb{X}_1^n(t)|\right] \lesssim \Delta_n^{-\frac{1}{2}} \sum_{j=1}^{N(H)} \sum_{k=1}^{j} \Delta_n^{(j-k)(1-2H)} \Delta_n^{k2\theta(1-H)} \lesssim \Delta_n^{-\frac{1}{2}+2\theta(1-H)} \to 0$$

by (3.2.14). Using (4.2.27) as well as (3.2.11), we further see that the magnitude of $\Upsilon^{n,i} - \pi((i-1)\Delta_n)$ is $\lesssim \Delta_n^{1-2H} + \Delta_n^{2\theta(1-H)}$. Thus, again by the generalized Hölder inequality, we deduce that

$$\mathbb{E}\left[\sup_{t\leq T} |\mathbb{X}_2^n(t)|\right] \lesssim \Delta_n^{-\frac{1}{2}} \left\{\Delta_n^{(N(H)+1)(1-2H)} + \Delta_n^{(N(H)+1)2\theta(1-H)}\right\} \to 0$$

by the definition of $N(H)$. □

Proof of Lemma 3.2.10. The first convergence (3.2.20) can be shown analogously to [64, Equation (5.3.24)] and is omitted. For (3.2.21), we write the left-hand side as $\sum_{j=1}^{N(H)} \mathbb{Z}_j^n(t) - \overline{\mathbb{Z}}^n(t)$ where

$$\begin{aligned}
\mathbb{Z}_j^n(t) = \Delta_n^{-\frac{1}{2}+j(1-2H)} \sum_{i=\theta_n+1}^{[t/\Delta_n]-L+1} \sum_{|\chi|=j} \frac{1}{\chi!} \int_{(i-1)\Delta_n}^{i\Delta_n} &\Big\{ \partial^\chi \mu_f\left(\pi((i-1)\Delta_n)\right) c((i-1)\Delta_n)^\chi \\
&- \partial^\chi \mu_f\left(\pi(s)\right) c(s)^\chi \Big\}\,\mathrm{d}s,
\end{aligned}$$

$$\overline{\mathbb{Z}}^n(t) = \Delta_n^{-\frac{1}{2}} \int_0^t \left(\mathbb{1}_{\{0\leq s\leq \theta_n\Delta_n\}} + \mathbb{1}_{\{([t/\Delta_n]+L-1)\Delta_n\leq s\leq t\}}\right) \sum_{j=1}^{N(H)} \sum_{|\chi|=j} \frac{1}{\chi!} \partial^\chi \mu_f\left(\pi(s)\right) \Delta_n^{j(1-2H)} c(s)^\chi\,\mathrm{d}s.$$

Using the moment assumptions on σ and ρ, since $t - ([t/\Delta_n] - L + 1)\Delta_n \leq L\Delta_n$, we readily see that

$$\mathbb{E}\left[\sup_{t\leq T} |\overline{\mathbb{Z}}^n(t)|\right] \lesssim \Delta_n^{-\frac{1}{2}}(\theta_n\Delta_n + L\Delta_n) \lesssim \Delta_n^{\frac{1}{2}-\theta} + \Delta_n^{\frac{1}{2}} \to 0.$$

Let $j = 1, \ldots, N(H)$ (in particular, everything in the following can be skipped if $H < \frac{1}{4}$) and consider, for $\chi \in \mathbb{N}_0^{(d\times L)\times(d\times L)}$, again the polynomial Q_χ introduced in (4.2.25). Using the mean-value theorem, we can write

$$\begin{aligned}
\mathbb{Z}_j^n(t) = \Delta_n^{-\frac{1}{2}+j(1-2H)} \sum_{i=\theta_n+1}^{[t/\Delta_n]-L+1} \sum_{|\chi|=j} \frac{1}{\chi!} \int_{(i-1)\Delta_n}^{i\Delta_n} \sum_{|\chi_1+\chi_2|=1} \partial^{\chi+\chi_1} \mu_f(\zeta_{n,i}^1) \partial^{\chi_2} Q_\chi(\zeta_{n,i}^2) \\
\times \{\pi((i-1)\Delta_n) - \pi(s)\}^{\chi_1} \{c((i-1)\Delta_n) - c(s)\}^{\chi_2}\,\mathrm{d}s
\end{aligned}$$

for some matrices $\zeta^1_{n,i}$ and $\zeta^2_{n,i}$. By the generalized Hölder inequality as well as the moment and regularity assumptions on σ and ρ, we deduce that

$$\mathbb{E}\left[\sup_{t\leq T}|\mathbb{Z}^n_j(t)|\right] \lesssim \Delta_n^{-\frac{1}{2}+j(1-2H)}\Delta_n^{-1}\Delta_n\Delta_n^{\frac{1}{2}} = \Delta_n^{j(1-2H)} \to 0$$

for any $H < \frac{1}{2}$. This concludes the proof of the lemma. □

4.3 Proofs for Section 3.3

Proof of Theorem 3.3.2. Since φ is invertible, we can write

$$\begin{aligned} H &= \varphi^{-1}\left(\frac{\langle a, \Gamma^H\rangle \Pi_t}{\langle b, \Gamma^H\rangle \Pi_t}\right) = G(\langle a, \Gamma^H\rangle \Pi_t, \langle b, \Gamma^H\rangle \Pi_t), \\ \widetilde{H}^n &= G(\langle a, \widehat{V}^n_t\rangle, \langle b, \widehat{V}^n_t\rangle) = G(\langle a, V^n_t\rangle, \langle b, V^n_t\rangle), \qquad G(x,y) = \varphi^{-1}\left(\frac{x}{y}\right). \end{aligned} \tag{4.3.1}$$

As G is infinitely differentiable on $\mathbb{R}\times(\mathbb{R}\setminus\{0\})$, we can expand $\widetilde{H}^n$ in a Taylor sum around $(\langle a, \Gamma^H\rangle \Pi_t, \langle b, \Gamma^H\rangle \Pi_t)$ and obtain

$$\begin{aligned} \widetilde{H}^n - H &= \sum_{|\chi|=1} \partial^\chi G(\langle a, \Gamma^H\rangle \Pi_t, \langle b, \Gamma^H\rangle \Pi_t)\left(\langle a, V^n_t\rangle - \langle a, \Gamma^H\rangle \Pi_t, \langle b, V^n_t\rangle - \langle b, \Gamma^H\rangle \Pi_t\right)^\chi + \mathbb{H}^n, \\ \mathbb{H}^n &= \sum_{|\chi|=2} \frac{\partial^\chi G(\alpha^n)}{\chi!}\left(\langle a, V^n_t\rangle - \langle a, \Gamma^H\rangle \Pi_t, \langle b, V^n_t\rangle - \langle b, \Gamma^H\rangle \Pi_t\right)^\chi, \end{aligned} \tag{4.3.2}$$

where $\chi \in \mathbb{N}_0^2$ and α^n is a random vector between $(\langle a, V^n_t\rangle, \langle b, V^n_t\rangle)$ and $(\langle a, \Gamma^H\rangle \Pi_t, \langle b, \Gamma^H\rangle \Pi_t)$. By straightforward computations,

$$\partial^{(1,0)}G(x,y) = (\varphi^{-1})'\left(\frac{x}{y}\right)\frac{1}{y} \quad \text{and} \quad \partial^{(0,1)}G(x,y) = -(\varphi^{-1})'\left(\frac{x}{y}\right)\frac{x}{y^2}. \tag{4.3.3}$$

Therefore, (4.3.2) becomes

$$\begin{aligned} \widetilde{H}^n - H &= (\varphi^{-1})'\left(\frac{\langle a, \Gamma^H\rangle}{\langle b, \Gamma^H\rangle}\right)\frac{1}{\langle b, \Gamma^H\rangle \Pi_t}\left(\langle a, V^n_t\rangle - \langle a, \Gamma^H\rangle \Pi_t\right) \\ &\quad - (\varphi^{-1})'\left(\frac{\langle a, \Gamma^H\rangle}{\langle b, \Gamma^H\rangle}\right)\frac{\langle a, \Gamma^H\rangle \Pi_t}{(\langle b, \Gamma^H\rangle \Pi_t)^2}\left(\langle b, V^n_t\rangle - \langle b, \Gamma^H\rangle \Pi_t\right) + \mathbb{H}^n \\ &= \frac{(\varphi^{-1})'(\varphi(H))}{\langle b, \Gamma^H\rangle \Pi_t}\left\{\left(\langle a, V^n_t\rangle - \langle a, \Gamma^H\rangle \Pi_t\right) - \varphi(H)\left(\langle b, V^n_t\rangle - \langle b, \Gamma^H\rangle \Pi_t\right)\right\} + \mathbb{H}^n. \end{aligned} \tag{4.3.4}$$

Because $H \in (0, \frac{1}{4})$ or $a_0 = b_0 = 0$, the first expression on the right-hand side of (4.3.4) can further be written as

$$\frac{(\varphi^{-1})'(\varphi(H))}{\langle b, \Gamma^H\rangle \Pi_t}\left\{a^T - \varphi(H)b^T\right\}\left\{V^n_t - \Gamma^H\int_0^t \rho_s^2\,\mathrm{d}s - e_1\int_0^t \sigma_s^2\,\mathrm{d}s\,\Delta_n^{1-2H}\,\mathbb{1}_{[\frac{1}{4},\frac{1}{2})}(H)\right\},$$

Moreover, by Corollary 3.3.1, the term $\mathbb{H}^n$ is of magnitude Δ_n and hence,

$$\begin{aligned}\Delta_n^{-\frac{1}{2}}(\widetilde{H}^n - H) &= \frac{(\varphi^{-1})'(\varphi(H))}{\langle b, \Gamma^H\rangle \Pi_t}\left\{a^T - \varphi(H)b^T\right\} \\ &\quad \times \Delta_n^{-\frac{1}{2}}\left\{V_t^n - \Gamma^H\int_0^t \rho_s^2\,\mathrm{d}s - e_1\int_0^t \sigma_s^2\,\mathrm{d}s\,\Delta_n^{1-2H}\,\mathbb{1}_{[\frac{1}{4},\frac{1}{2})}(H)\right\} + \Delta_n^{-\frac{1}{2}}\mathbb{H}^n \\ &\xrightarrow{\text{st}} \frac{(\varphi^{-1})'(\varphi(H))}{\langle b, \Gamma^H\rangle \Pi_t}\left\{a^T - \varphi(H)b^T\right\}\mathcal{Z}_t \sim \mathcal{N}\left(0, \mathrm{Var}_{H,0}\frac{\int_0^t \rho_s^4\,\mathrm{d}s}{(\int_0^t \rho_s^2\,\mathrm{d}s)^2}\right),\end{aligned}$$

which proves (3.3.5).

We now turn to the convergence stated in (3.3.7) when $H > \frac{1}{4}$. We decompose

$$\begin{aligned}V_{0,t}^n - \frac{\langle c, V_t^n\rangle}{\langle c, \Gamma^{\widetilde{H}^n}\rangle} &= \{V_{0,t}^n - \Pi_t\} - \frac{1}{\langle c, \Gamma^{\widetilde{H}^n}\rangle}\left\{\langle c, V_t^n\rangle - \langle c, \Gamma^H\rangle\Pi_t\right\} + \frac{\Pi_t}{\langle c, \Gamma^{\widetilde{H}^n}\rangle}\left\{\langle c, \Gamma^{\widetilde{H}^n}\rangle - \langle c, \Gamma^H\rangle\right\} \\ &= \{V_{0,t}^n - \Pi_t\} - \frac{1}{\langle c, \Gamma^{\widetilde{H}^n}\rangle}c^T\{V_t^n - \Gamma^H\Pi_t\} + \Pi_t\frac{\langle c, \partial_H\Gamma^{\widetilde{H}^n}\rangle}{\langle c, \Gamma^{\widetilde{H}^n}\rangle}\{\widetilde{H}^n - H\} + \mathbb{V}^n, \\ \mathbb{V}^n &= \frac{1}{2}\Pi_t\frac{\langle c, \partial_{HH}\Gamma^{\beta^n}\rangle}{\langle c, \Gamma^{\widetilde{H}^n}\rangle}\{\widetilde{H}^n - H\}^2,\end{aligned} \tag{4.3.5}$$

where $\partial_{HH}\Gamma^H$ is the second derivative of $H \mapsto (\Gamma_0^H, \ldots, \Gamma_R^H)$ evaluated at H and β^n is somewhere between $\widetilde{H}^n$ and H. Since $c_0 \neq 0$, the first two terms in the second line of (4.3.5) are of magnitude Δ_n^{1-2H}, while the third is of magnitude $\Delta_n^{1/2}$ by our first result (3.3.5). Finally, $\mathbb{V}^n$ is of magnitude Δ_n, so using Corollary 3.3.1, we deduce

$$\Delta_n^{2H-1}\left\{V_{0,t}^n - \frac{\langle c, V_t^n\rangle}{\langle c, \Gamma^{\widetilde{H}^n}\rangle}\right\} \xrightarrow{P} C_t - \frac{1}{\langle c, \Gamma^H\rangle}\langle c, e_1\rangle C_t = \left(1 - \frac{1}{\langle c, \Gamma^H\rangle}c_0\right)C_t. \tag{4.3.6}$$

Reusing the Taylor expansion of $\widetilde{H}^n$ from (4.3.4) and recalling that $a_0 = b_0 = 0$, we further have that

$$\begin{aligned}&\Delta_n^{-\frac{1}{2}}\left\{V_{0,t}^n - \frac{\langle c, V_t^n\rangle}{\langle c, \Gamma^{\widetilde{H}^n}\rangle} - \left(1 - \frac{1}{\langle c, \Gamma^{\widetilde{H}^n}\rangle}c_0\right)C_t\,\Delta_n^{1-2H}\right\} \\ &= \Delta_n^{-\frac{1}{2}}\{V_{0,t}^n - \Pi_t - C_t\,\Delta_n^{1-2H}\} - \frac{\Delta_n^{-\frac{1}{2}}c^T\{V_t^n - \Gamma^H\Pi_t - e_1C_t\,\Delta_n^{1-2H}\}}{\langle c, \Gamma^{\widetilde{H}^n}\rangle} \\ &\quad + \Pi_t\frac{\langle c, \partial_H\Gamma^{\widetilde{H}^n}\rangle}{\langle c, \Gamma^{\widetilde{H}^n}\rangle}\Delta_n^{-\frac{1}{2}}\{\widetilde{H}^n - H\} + \Delta_n^{-\frac{1}{2}}\mathbb{V}^n \\ &= \left(e_1^T - \frac{c^T}{\langle c, \Gamma^{\widetilde{H}^n}\rangle} + \Pi_t\frac{\langle c, \partial_H\Gamma^{\widetilde{H}^n}\rangle}{\langle c, \Gamma^{\widetilde{H}^n}\rangle}\frac{(\varphi^{-1})'(\varphi(H))}{\langle b, \Gamma^H\rangle\Pi_t}\left\{a^T - \varphi(H)b^T\right\}\right) \\ &\qquad \times \Delta_n^{-\frac{1}{2}}\left\{V_t^n - \Gamma^H\Pi_t - e_1C_t\,\Delta_n^{1-2H}\right\} + \Delta_n^{-\frac{1}{2}}\left(\Pi_t\frac{\langle c, \partial_H\Gamma^{\widetilde{H}^n}\rangle}{\langle c, \Gamma^{\widetilde{H}^n}\rangle}\mathbb{H}^n + \mathbb{V}^n\right) \\ &\xrightarrow{\text{st}} \left(e_1^T - \frac{c^T}{\langle c, \Gamma^H\rangle} + \frac{\langle c, \partial_H\Gamma^H\rangle}{\langle c, \Gamma^H\rangle}\frac{(\varphi^{-1})'(\varphi(H))}{\langle b, \Gamma^H\rangle}\left\{a^T - \varphi(H)b^T\right\}\right)\mathcal{Z}_t.\end{aligned}$$

It remains to normalize the left-hand side of (4.3.6) in order to obtain (3.3.7):

$$\Delta_n^{-\frac{1}{2}+(1-2H)} \left\{ \left(\widehat{V}_{0,t}^n - \frac{\langle c, \widehat{V}_t^n \rangle}{\langle c, \Gamma^{\widetilde{H}^n} \rangle} \right) \left(1 - \frac{1}{\langle c, \Gamma^{\widetilde{H}^n} \rangle} c_0 \right)^{-1} - C_t \right\}$$
$$\xrightarrow{\text{st}} \left(1 - \frac{1}{\langle c, \Gamma^H \rangle} c_0 \right)^{-1} \left(e_1^T - \frac{c^T}{\langle c, \Gamma^H \rangle} + \frac{\langle c, \partial_H \Gamma^H \rangle}{\langle c, \Gamma^H \rangle} \frac{(\varphi^{-1})'(\varphi(H))}{\langle b, \Gamma^H \rangle} \left\{ a^T - \varphi(H) b^T \right\} \right) \mathcal{Z}_t$$
$$\sim \mathcal{N}\left(0, \mathrm{Var}_C \int_0^t \rho_s^4 \,\mathrm{d}s \right).$$

Finally, we tackle (3.3.12). We use the mean-value theorem to decompose

$$\begin{aligned} &\Delta_n^{-\frac{1}{2}} \left(\frac{\langle a, V_t^n \rangle}{\langle a, \Gamma^{\widetilde{H}^n} \rangle} - \Pi_t \right) \\ &\quad = \Delta_n^{-\frac{1}{2}} \frac{1}{\langle a, \Gamma^{\widetilde{H}^n} \rangle} \left\{ \langle a, V_t^n \rangle - \langle a, \Gamma^H \rangle \Pi_t \right\} - \frac{\Pi_t}{\langle a, \Gamma^{\widetilde{H}^n} \rangle} \Delta_n^{-\frac{1}{2}} \left\{ \langle a, \Gamma^{\widetilde{H}^n} \rangle - \langle a, \Gamma^H \rangle \right\} \\ &\quad = \frac{a^T}{\langle a, \Gamma^{\widetilde{H}^n} \rangle} \Delta_n^{-\frac{1}{2}} \left\{ V_t^n - \Gamma^H \Pi_t \right\} - \frac{\Pi_t}{\langle a, \Gamma^{\widetilde{H}^n} \rangle} \langle a, \partial_H \Gamma^{\widetilde{\beta}^n} \rangle \Delta_n^{-\frac{1}{2}} \left\{ \widetilde{H}^n - H \right\} \end{aligned} \tag{4.3.7}$$

where $\widetilde{\beta}^n$ is between $\widetilde{H}^n$ and H and therefore satisfies $\widetilde{\beta}^n \xrightarrow{P} H$. As before, because $\frac{1}{4} < H < \frac{1}{2}$ or $a_0 = b_0 = 0$, we prefer to write

$$\frac{a^T}{\langle a, \Gamma^{\widetilde{H}^n} \rangle} \Delta_n^{-\frac{1}{2}} \left\{ V_t^n - \Gamma^H \Pi_t \right\} = \frac{a^T}{\langle a, \Gamma^{\widetilde{H}^n} \rangle} \Delta_n^{-\frac{1}{2}} \left\{ V_t^n - \Gamma^H \Pi_t - C_t \, \Delta_n^{1-2H} \, \mathbb{1}_{[\frac{1}{4}, \frac{1}{2})}(H) \right\}.$$

Using Corollary 3.3.1 and our first result (3.3.5), we infer that $\Delta_n^{-1/2}(\langle a, V_t^n \rangle / \langle a, \Gamma^{\widetilde{H}^n} \rangle - \Pi_t)$ converges stably in distribution. Applying again the mean-value theorem, this time on the function $H \mapsto \Delta_n^{-2H}$, and recalling the identity $\Delta_n^{1-2H} \widehat{V}_{r,t}^n = V_{r,t}^n$, we further obtain

$$\begin{aligned} \Delta_n^{-\frac{1}{2}} (\widehat{\Pi}_t^n - \Pi_t) &= \Delta_n^{-\frac{1}{2}} \left(\frac{\langle a, V_t^n \rangle}{\langle a, \Gamma^{\widetilde{H}^n} \rangle} - \Pi_t \right) + \Delta_n^{-\frac{1}{2}} \frac{\langle a, \widehat{V}_t^n \rangle}{\langle a, \Gamma^{\widetilde{H}^n} \rangle} \left\{ \Delta_n^{1-2\widetilde{H}^n} - \Delta_n^{1-2H} \right\} \\ &= \Delta_n^{-\frac{1}{2}} \left(\frac{\langle a, V_t^n \rangle}{\langle a, \Gamma^{\widetilde{H}^n} \rangle} - \Pi_t \right) - 2 \frac{\langle a, \widehat{V}_t^n \rangle}{\langle a, \Gamma^{\widetilde{H}^n} \rangle} \Delta_n^{1-2H} \log(\Delta_n) \Delta_n^{2(H - \overline{\beta}^n)} \Delta_n^{-\frac{1}{2}} \left\{ \widetilde{H}^n - H \right\} \end{aligned}$$

and $\overline{\beta}^n$ is again some point between $\widetilde{H}^n$ and H. Observe that by (3.3.5), $\overline{\beta}^n$ converges to H at a rate of $\Delta_n^{1/2}$. Therefore, $\Delta_n^{2(H-\overline{\beta}^n)} \to 1$ as $n \to \infty$. Normalizing by $\log(\Delta_n)$, we conclude from (3.3.5) that

$$\begin{aligned} \frac{\Delta_n^{-\frac{1}{2}}}{\log(\Delta_n)} (\widehat{\Pi}_t^n - \Pi_t) &= \frac{\Delta_n^{-\frac{1}{2}}}{\log(\Delta_n)} \left(\frac{\langle a, V_t^n \rangle}{\langle a, \Gamma^{\widetilde{H}^n} \rangle} - \Pi_t \right) - 2 \frac{\langle a, V_t^n \rangle}{\langle a, \Gamma^{\widetilde{H}^n} \rangle} \Delta_n^{2(H-\overline{\beta}^n)} \Delta_n^{-\frac{1}{2}} \left\{ \widetilde{H}^n - H \right\} \\ &\xrightarrow{\text{st}} -2\Pi_t \mathcal{N}\left(0, \mathrm{Var}_{H,0} \frac{\int_0^t \rho_s^4 \,\mathrm{d}s}{(\int_0^t \rho_s^2 \,\mathrm{d}s)^2} \right) \sim \mathcal{N}\left(0, 4\mathrm{Var}_{H,0} \int_0^t \rho_s^4 \,\mathrm{d}s \right). \end{aligned}$$

□

Proof of Proposition 3.3.5. Starting from (4.3.1), we expand

$$\Delta_n^{-\frac{1}{2}}\left(\widetilde{H}^n - H\right) = -\sum_{j=1}^{N(H)}\sum_{|\chi|=j}\frac{\partial^\chi G(\langle a, V_t^n\rangle, \langle b, V_t^n\rangle)}{\chi!}(-1)^j$$
$$\times \Delta_n^{-\frac{1}{2}}\left(\langle a, V_t^n\rangle - \langle a, \Gamma^H\rangle\Pi_t, \langle b, V_t^n\rangle - \langle b, \Gamma^H\rangle\Pi_t\right)^\chi - \mathbb{I}^n,$$
$$\mathbb{I}^n = \sum_{|\chi|=N(H)+1}\frac{\partial^\chi G(\overline{\alpha}^n)}{\chi!}(-1)^{|\chi|}\Delta_n^{-\frac{1}{2}}\left(\langle a, V_t^n\rangle - \langle a, \Gamma^H\rangle\Pi_t, \langle b, V_t^n\rangle - \langle b, \Gamma^H\rangle\Pi_t\right)^\chi, \tag{4.3.8}$$

where $\chi \in \mathbb{N}_0^2$ and $\overline{\alpha}^n$ is a point between $(\langle a, \Gamma^H\rangle\Pi_t, \langle b, \Gamma^H\rangle\Pi_t)$ and $(\langle a, V_t^n\rangle, \langle b, V_t^n\rangle)$. In contrast to the proof of (3.3.5), we expanded $\widetilde{H}^n$ around $(\langle a, \Gamma^H\rangle\Pi_t, \langle b, \Gamma^H\rangle\Pi_t)$ and not $(\langle a, V_t^n\rangle, \langle b, V_t^n\rangle)$. We consider the terms where $\chi = (j, 0)$ for some $j = 1, \ldots, N(H)$ or $\chi = (0, 1)$ separately. In the first case, $\partial^\chi G$ takes a simple form, namely

$$\partial^\chi G(x, y) = (\varphi^{-1})^{(j)}\left(\frac{x}{y}\right)\frac{1}{y^j}, \quad \chi = (j, 0), \quad j \geq 1;$$

in the second case, $\partial^\chi G(x, y)$ was computed in (4.3.3). With that in mind, and recalling (3.3.19), we have that

$$\Delta_n^{-\frac{1}{2}}\left(\widetilde{H}^n - H + \sum_{j=1}^{N(H)}\Phi_j^n C_t^j\right)$$
$$= (\phi^{-1})'(\phi(\widetilde{H}^n))\frac{1}{\langle b, V_t^n\rangle}\left\{a^T - \varphi(\widetilde{H}^n)b^T\right\}\Delta_n^{-\frac{1}{2}}\left\{V_t^n - \Gamma^H\Pi_t - e_1 C_t\,\Delta_n^{1-2H}\right\}$$
$$+ \sum_{j=2}^{N(H)}\frac{(-1)^{j+1}}{j!}(\phi^{-1})^{(j)}(\phi(\widetilde{H}^n))\frac{1}{\langle b, V_t^n\rangle^j}\Delta_n^{-\frac{1}{2}}\left\{\left(\langle a, V_t^n\rangle - \langle a, \Gamma^H\rangle\Pi_t\right)^j - a_0^j C_t^j\Delta_n^{j(1-2H)}\right\}$$
$$- \sum_{j=2}^{N(H)}\sum_{\chi\neq(j,0)}\frac{\partial^\chi G(\langle a, V_t^n\rangle, \langle b, V_t^n\rangle)}{\chi!}(-1)^j\Delta_n^{-\frac{1}{2}}\left(\langle a, V_t^n\rangle - \langle a, \Gamma^H\rangle\Pi_t, \langle b, V_t^n\rangle - \langle b, \Gamma^H\rangle\Pi_t\right)^\chi$$
$$- \mathbb{I}^n. \tag{4.3.9}$$

By the mean-value theorem and Corollary 3.3.1, one can easily see that $\Delta_n^{-1/2}\{(\langle a, V_t^n\rangle - \langle a, \Gamma^H\rangle\Pi_t)^j - a_0^j C_t^j\Delta_n^{j(1-2H)}\}$ is of magnitude $\Delta_n^{(j-1)/2}$. Thus, the second term on the right-hand side of (4.3.9) is asymptotically negligible. And so are the third term in (4.3.9) and $\mathbb{I}^n$: For any $\chi = (j-i, i) \in \mathbb{N}_0^2$, Corollary 3.3.1 and assumption (3.3.17) imply that $\Delta_n^{-1/2}(\langle a, V_t^n\rangle - \langle a, \Gamma^H\rangle\Pi_t, \langle b, V_t^n\rangle - \langle b, \Gamma^H\rangle\Pi_t)^\chi$ is of magnitude $\Delta_n^{(j-i)(1-2H)+i/2-1/2}$ and therefore asymptotically negligible as soon as $i \geq 1$ and $j-i \geq 1$. Similarly, $\mathbb{I}^n$ is of magnitude at most $\Delta_n^{(N(H)+1)(1-2H)-1/2}$, which goes to 0 by the definition of $N(H)$. Altogether, we obtain by Corollary 3.3.1,

$$\Delta_n^{-\frac{1}{2}}\left(\widetilde{H}^n - H + \sum_{j=1}^{N(H)}\Phi_j^n C_t^j\right)$$
$$\xrightarrow{\text{st}} (\phi^{-1})'(\phi(H))\frac{1}{\langle b, \Gamma^H\rangle\Pi_t}\left\{a^T - \varphi(H)b^T\right\}\mathcal{Z}_t \sim \mathcal{N}\left(0, \mathrm{Var}_{H,0}\frac{\int_0^t \rho_s^4\,\mathrm{d}s}{(\int_0^t \rho_s^2\,\mathrm{d}s)^2}\right),$$

which concludes the proof. □

Proof of Proposition 3.3.6. We start similarly to the proof of (3.3.8) and decompose

$$V_{0,t}^n - \frac{\langle a, V_t^n\rangle}{\langle a, \Gamma^{\widetilde{H}^n}\rangle} = \{V_{0,t}^n - \Pi_t\} - \frac{1}{\langle a, \Gamma^{\widetilde{H}^n}\rangle}\left\{\langle a, V_t^n\rangle - \langle a, \Gamma^H\rangle\Pi_t\right\} + \frac{\Pi_t}{\langle a, \Gamma^{\widetilde{H}^n}\rangle}\left\{\langle a, \Gamma^{\widetilde{H}^n}\rangle - \langle a, \Gamma^H\rangle\right\}. \tag{4.3.10}$$

We further analyze the last term in the above display and write

$$\begin{aligned}\langle a, \Gamma^{\widetilde{H}^n}\rangle &= K(\langle a, \widehat{V}_t^n\rangle, \langle b, \widehat{V}_t^n\rangle) = K(\langle a, V_t^n\rangle, \langle b, V_t^n\rangle),\\ \langle a, \Gamma^H\rangle &= K(\langle a, \Gamma^H\rangle, \langle b, \Gamma^H\rangle) = K(\langle a, \Gamma^H\rangle\Pi_t, \langle b, \Gamma^H\rangle\Pi_t),\end{aligned}$$

where $K(x,y) = \psi(x/y)$ and ψ is the function from (3.3.24). The following derivatives will be needed in the course of the proof:

$$\begin{aligned}\partial^{(j,0)}K(x,y) &= \psi^{(j)}\left(\frac{x}{y}\right)\frac{1}{y^j}, \qquad j \geq 1,\\ \partial^{(0,1)}K(x,y) &= -\psi'\left(\frac{x}{y}\right)\frac{x}{y^2}.\end{aligned}$$

We now expand $\langle a, \Gamma^H\rangle$ in a Taylor sum around the point $(\langle a, V_t^n\rangle, \langle b, V_t^n\rangle)$ up to order $N(H)$, singling out the two first-order derivatives as well as the derivatives $\partial^{(j,0)}$:

$$\begin{aligned}&\langle a, \Gamma^{\widetilde{H}^n}\rangle - \langle a, \Gamma^H\rangle\\ &\quad= \psi'(\phi(\widetilde{H}^n))\frac{1}{\langle b, V_t^n\rangle}\left(\left\{\langle a, V_t^n\rangle - \langle a, \Gamma^H\rangle\Pi_t\right\} - \phi(\widetilde{H}^n)\left\{\langle b, V_t^n\rangle - \langle b, \Gamma^H\rangle\Pi_t\right\}\right)\\ &\qquad - \sum_{j=2}^{N(H)}\frac{(-1)^j}{j!}\psi^{(j)}(\phi(\widetilde{H}^n))\frac{1}{\langle b, V_t^n\rangle^j}\left\{\langle a, V_t^n\rangle - \langle a, \Gamma^H\rangle\Pi_t\right\}^j - \mathbb{J}^n,\end{aligned} \tag{4.3.11}$$

where

$$\begin{aligned}\mathbb{J}^n = &\sum_{j=2}^{N(H)}\sum_{\chi\neq(j,0)}\frac{\partial^\chi K(\langle a, V_t^n\rangle, \langle b, V_t^n\rangle)}{\chi!}(-1)^j\Delta_n^{-\frac{1}{2}}\left(\langle a, V_t^n\rangle - \langle a, \Gamma^H\rangle\Pi_t, \langle b, V_t^n\rangle - \langle b, \Gamma^H\rangle\Pi_t\right)^\chi\\ &+ \sum_{|\chi|=N(H)+1}\frac{\partial^\chi K(\widetilde{\alpha}^n)}{\chi!}(-1)^{|\chi|}\Delta_n^{-\frac{1}{2}}\left(\langle a, V_t^n\rangle - \langle a, \Gamma^H\rangle\Pi_t, \langle b, V_t^n\rangle - \langle b, \Gamma^H\rangle\Pi_t\right)^\chi\end{aligned}$$

and $\widetilde{\alpha}^n$ is between $(\langle a, \Gamma^H\rangle\Pi_t, \langle b, \Gamma^H\rangle\Pi_t)$ and $(\langle a, V_t^n\rangle, \langle b, V_t^n\rangle)$. Using (4.3.10) for the first and (4.3.11) for the second equality, we find that

$$\begin{aligned}&\Delta_n^{-\frac{1}{2}}\Bigg(\left\{V_{0,t}^n - \frac{\langle a, V_t^n\rangle}{\langle a, \Gamma^{\widetilde{H}^n}\rangle}\right\} - \left(1 - \frac{a_0}{\langle a, \Gamma^{\widetilde{H}^n}\rangle} + \frac{\Pi_t}{\langle a, \Gamma^{\widetilde{H}^n}\rangle}\frac{a_0\psi'(\phi(\widetilde{H}^n))}{\langle b, V_t^n\rangle}\right)C_t\,\Delta_n^{1-2H}\\ &\qquad + \frac{\Pi_t}{\langle a, \Gamma^{\widetilde{H}^n}\rangle}\sum_{j=2}^{N(H)}\frac{(-1)^j}{j!}\psi^{(j)}(\phi(\widetilde{H}^n))\frac{a_0^j}{\langle b, V_t^n\rangle^j}C_t^j\,\Delta_n^{j(1-2H)}\Bigg)\\ &\quad= \Delta_n^{-\frac{1}{2}}\left\{V_{0,t}^n - \Pi_t - C_t\Delta_n^{1-2H}\right\} - \frac{1}{\langle a, \Gamma^{\widetilde{H}^n}\rangle}\Delta_n^{-\frac{1}{2}}\left\{\langle a, V_t^n\rangle - \langle a, \Gamma^H\rangle\Pi_t - a_0C_t\,\Delta_n^{1-2H}\right\}\end{aligned}$$

$$
\begin{aligned}
&+ \frac{\Pi_t}{\langle a, \Gamma^{\widetilde{H}^n}\rangle} \Delta_n^{-\frac{1}{2}} \left\{ \langle a, \Gamma^{\widetilde{H}^n}\rangle - \langle a, \Gamma^H\rangle + \sum_{j=1}^{N(H)} \frac{(-1)^j}{j!} \psi^{(j)}(\phi(\widetilde{H}^n)) \frac{a_0^j}{\langle b, V_t^n\rangle^j} C_t^j \, \Delta_n^{j(1-2H)} \right\} \\
&= \left\{ e_1^T - \frac{a^T}{\langle a, \Gamma^{\widetilde{H}^n}\rangle} + \frac{\Pi_t}{\langle a, \Gamma^{\widetilde{H}^n}\rangle} \frac{\psi'(\phi(\widetilde{H}^n))}{\langle b, V_t^n\rangle} (a^T - \phi(\widetilde{H}^n) b^T) \right\} \Delta_n^{-\frac{1}{2}} \left\{ V_t^n - \Gamma^H \Pi_t - e_1 C_t \Delta_n^{1-2H} \right\} \\
&- \frac{\Pi_t}{\langle a, \Gamma^{\widetilde{H}^n}\rangle} \sum_{j=2}^{N(H)} \frac{(-1)^j}{j!} \psi^{(j)}(\phi(\widetilde{H}^n)) \frac{1}{\langle b, V_t^n\rangle^j} \Delta_n^{-\frac{1}{2}} \left\{ \left\{ \langle a, V_t^n\rangle - \langle a, \Gamma^H\rangle \Pi_t \right\}^j - a_0^j C_t^j \Delta_n^{j(1-2H)} \right\} \\
&- \frac{\Pi_t}{\langle a, \Gamma^{\widetilde{H}^n}\rangle} \Delta_n^{-\frac{1}{2}} \mathbb{J}^n. \qquad (4.3.12)
\end{aligned}
$$

For the exact same reasons as explained after (4.3.8), the term involving $\mathbb{J}^n$ is asymptotically negligible: $(\langle a, V_t^n\rangle - \langle a, \Gamma^H\rangle \Pi_t, \langle b, V_t^n\rangle - \langle b, \Gamma^H\rangle \Pi_t)^\chi$ is of magnitude $\Delta_n^{(j-i)(1-2H)+i/2} \le \Delta_n^{3/2-2H}$ if $|\chi| = 2, \ldots, N(H)$ and $\chi \ne (j, 0)$, and it is of magnitude $\le \Delta_n^{(N(H)+1)(1-2H)}$ if $|\chi| = N(H)+1$; in both cases, the exponent is strictly bigger than $\frac{1}{2}$. Moreover, by Corollary 3.3.1,

$$
\Delta_n^{-j(1-2H)} \left(\langle a, V_t^n\rangle - \langle a, \Gamma^H\rangle \Pi_t \right)^j \xrightarrow{\mathbb{P}} a_0^j C_t^j,
$$

which implies that the second term on the right-hand side of (4.3.12) is of magnitude $\Delta_n^{(j-1)1/2}$ for $j = 2, \ldots, N(H)$. Thus, by Corollary 3.3.1, the left-hand side of (4.3.12) converges stably in law to

$$
\mathcal{Z}_t' = \left\{ e_1^T - \frac{a^T}{\langle a, \Gamma^H\rangle} + \frac{\Pi_t}{\langle a, \Gamma^H\rangle} \frac{\psi'(\phi(H))}{\langle b, \Gamma^H\rangle \Pi_t} (a^T - \phi(H) b^T) \right\} \mathcal{Z}_t. \qquad (4.3.13)
$$

Next, we replace Π_t in the first two lines of (4.3.12) by $\Delta_n^{1-2H} \widehat{P}_t^n$, where $\widehat{P}_t^n$ was introduced in (3.3.21). The resulting difference is given by

$$
\frac{\Delta_n^{-\frac{1}{2}} \{ \Delta_n^{1-2H} \widehat{P}_t^n - \Pi_t \}}{\langle a, \Gamma^{\widetilde{H}^n}\rangle} \sum_{j=1}^{N(H)} \frac{(-1)^j}{j!} \psi^{(j)}(\phi(\widetilde{H}^n)) \frac{a_0^j}{\langle b, V_t^n\rangle^j} C_t^j \, \Delta_n^{j(1-2H)}. \qquad (4.3.14)
$$

By the proof of Theorem 3.3.2 [see (4.3.7) in particular], the term $\Delta_n^{-1/2} \{ \Delta_n^{1-2H} \widehat{P}_t^n - \Pi_t \}$ converges stably in distribution. As a consequence, the expression in the previous display converges to 0 in probability as $n \to \infty$. By (3.3.23), (3.3.26) and (4.3.13), it follows that

$$
\begin{aligned}
&\Delta_n^{-\frac{1}{2}+(1-2H)} \left(\left\{ V_{0,t}^n - \frac{\langle a, V_t^n\rangle}{\langle a, \Gamma^{\widetilde{H}^n}\rangle} \right\} \Delta_n^{2H-1} \Theta(\widehat{V}_t^n, \widetilde{H}^n, \widetilde{H}^{n,0})^{-1} - C_t + \sum_{j=2}^{N(H)} \Psi_j^n C_t^j \right) \\
&= \Theta(\widehat{V}_t^n, \widetilde{H}^n, \widetilde{H}^{n,0})^{-1} \Delta_n^{-\frac{1}{2}} \left(\left\{ V_{0,t}^n - \frac{\langle a, V_t^n\rangle}{\langle a, \Gamma^{\widetilde{H}^n}\rangle} \right\} - \Theta(\widehat{V}_t^n, \widetilde{H}^n, \widetilde{H}^{n,0}) C_t \, \Delta_n^{1-2H} \right. \\
&\quad \left. + \frac{\Delta_n^{1-2H} \widehat{P}_t^n}{\langle a, \Gamma^{\widetilde{H}^n}\rangle} \sum_{j=2}^{N(H)} \frac{(-1)^j}{j!} \psi^{(j)}(\phi(\widetilde{H}^n) \frac{a_0^j}{\langle b, V_t^n\rangle^j} C_t^j \, \Delta_n^{j(1-2H)} \right) \qquad (4.3.15) \\
&\xrightarrow{\text{st}} \left(1 - \frac{a_0}{\langle a, \Gamma^H\rangle} + \frac{\psi'(\phi(H))}{\langle a, \Gamma^H\rangle \langle b, \Gamma^H\rangle} a_0 \right)^{-1} \mathcal{Z}_t' \sim \mathcal{N}\left(0, \mathrm{Var}_{C,1} \int_0^t \rho_s^4 \, \mathrm{d}s \right).
\end{aligned}
$$

The CLT stated in (3.3.25) is proved. □

Proof of Proposition 3.3.7. We first prove by induction that for $\ell = 0, \dots, N(H)-2$, the difference $\widetilde{C}_t^{n,\ell+1} - C_t$ converges in probability with a convergence rate of $\Delta_n^{(1+\ell)(1-2H)}$. If $\ell = 0$, then $\widetilde{C}_t^{n,\ell+1} = \widetilde{C}_t^{n,1}$, so by (4.3.15),

$$\Delta_n^{2H-1}\left(\widetilde{C}_t^{n,1} - C_t\right) \xrightarrow{\mathbb{P}} -\frac{1}{2}\psi^{(2)}(\phi(H))\frac{a_0^2}{\langle a, \Gamma^H\rangle}\frac{1}{\langle b, \Gamma^H\rangle^2}\left(1 - \frac{a_0}{\langle a, \Gamma^H\rangle} + \frac{\psi'(\phi(H))}{\langle a, \Gamma^H\rangle\langle b, \Gamma^H\rangle}a_0\right)^{-1} C_t^2.$$

Suppose now that $\widetilde{C}_t^{n,\ell+1} - C_t$ converges at a rate of $\Delta_n^{(1+\ell)(1-2H)}$ for $\ell = 0, \dots, \ell'-1$. Decomposing

$$\widetilde{C}_t^{n,\ell'+1} - C_t = \left\{\widetilde{C}_t^{n,1} - C_t + \sum_{j=2}^{\ell'+1}\Psi_j^n C_t^j\right\} + \sum_{j=2}^{\ell'+1}\Psi_j^n\left\{(\widetilde{C}_t^{n,\ell'-j+2})^j - C_t^j\right\}, \tag{4.3.16}$$

we note that the first term on the right-hand side converges at a rate of $\Delta_n^{(\ell'+1)(1-2H)}$ by (4.3.15). The second term can be rewritten as

$$\sum_{j=2}^{\ell'+1}\Psi_j^n\left\{(\widetilde{C}_t^{n,\ell'-j+2})^j - C_t^j\right\} = \sum_{j=2}^{\ell'+1}\sum_{m=1}^{j}\frac{j!}{(j-m)!}C_t^{j-m}\left\{\Psi_j^n(\widetilde{C}_t^{n,\ell'-j+2} - C_t)^m\right\}. \tag{4.3.17}$$

By assumption, $\widetilde{C}_t^{n,(\ell'-j+1)+1} - C_t$ is of magnitude $\Delta_n^{(\ell'-j+2)(1-2H)}$. Furthermore, by the definition (3.3.26) of Ψ_j^n, the product $\Psi_j^n\Delta_n^{(1-j)(1-2H)}$ converges in the probability. Thus, Ψ_j^n is of magnitude $\Delta_n^{(j-1)(1-2H)}$ and we conclude that $\Psi_j^n(\widetilde{C}_t^{n,\ell'-j+2} - C_t)^m$ is of magnitude $\Delta_n^{(j-1+m(\ell'-j+2))(1-2H)} \leq \Delta_n^{(\ell'+1)(1-2H)}$. Altogether, we have shown that $\widetilde{C}_t^{n,\ell'+1} - C_t$ is of magnitude $\Delta_n^{(\ell'+1)(1-2H)}$.

We can now complete the proof of the proposition. By an analogous decomposition to (4.3.16) with $\ell' = N(H)-1$,

$$\widetilde{C}_t^{n,N(H)} - C_t = \left\{\widetilde{C}_t^{n,1} - C_t + \sum_{j=2}^{N(H)}\Psi_j^n C_t^j\right\} + \sum_{j=2}^{N(H)}\Psi_j^n\left\{(\widetilde{C}_t^{n,N(H)-j+1})^j - C_t^j\right\}. \tag{4.3.18}$$

We know that $\widetilde{C}_t^{n,N(H)-j+1} - C_t$ is of magnitude $\Delta_n^{(N(H)-j+1)(1-2H)}$. Therefore, proceeding exactly as in (4.3.17), we see that the right-hand side of (4.3.18) times $\Delta_n^{1/2-2H}$ is of magnitude $\Delta_n^{(N(H)+1)(1-2H)-1/2}$ which goes to 0 as $n \to \infty$ since the exponent is positive by the definition of $N(H)$. Therefore, $\Delta_n^{1/2-2H}\{\widetilde{C}_t^{n,N(H)} - C_t\}$ converges stably to the same distribution as $\Delta_n^{1/2-2H}\{\widetilde{C}_t^{n,1} - C_t\}$ does.

Finally,

$$\Delta_n^{-\frac{1}{2}}\{\widetilde{C}_t^{n,N(\widetilde{H}^n)} - C_t\} = \Delta_n^{-\frac{1}{2}}\{\widetilde{C}_t^{n,N(H)} - C_t\} + \Delta_n^{-\frac{1}{2}}\{\widetilde{C}_t^{n,N(\widetilde{H}^n)} - \widetilde{C}_t^{n,N(H)}\}.$$

Since $\widetilde{H}^n$ is a consistent estimator for H and $H \notin \mathcal{H}$, for small enough $\varepsilon > 0$ (such that the event $\{|\widetilde{H}^n - H| \leq \varepsilon\} \subseteq \{N(\widetilde{H}^n) = N(H)\}$), we have

$$\mathbb{P}(\Delta_n^{-\frac{1}{2}}|\widetilde{C}_t^{n,N(\widetilde{H}^n)} - \widetilde{C}_t^{n,N(H)}| > \varepsilon) \leq \mathbb{P}(|\widetilde{H}^n - H| > \varepsilon) \to 0 \tag{4.3.19}$$

as $n \to \infty$. Thus, the CLT of $\{\widetilde{C}_t^{n,N(H)} - C_t\}$ is not affected when $N(H)$ is replaced by $N(\widetilde{H}^n)$. □

Proof of Proposition 3.3.8. We first decompose

$$
\begin{aligned}
&\Delta_n^{-\frac{1}{2}}\left\{\widetilde{H}^n - H + \sum_{j=1}^{N(H)} \Phi_j^n \left(\widehat{C}_t^{n,1}\right)^j\right\} \\
&\quad = \Delta_n^{-\frac{1}{2}}\left\{\widetilde{H}^n - H + \sum_{j=1}^{N(H)} \Phi_j^n C_t^j\right\} + \Phi_1^n \Delta_n^{-\frac{1}{2}}\left\{\widetilde{C}_t^{n,N(H)} - C_t\right\} \\
&\qquad + \sum_{j=2}^{N(H)} \Phi_j^n \Delta_n^{-\frac{1}{2}}\left\{(\widetilde{C}_t^{n,N(H)})^j - C_t^j\right\} + \sum_{j=1}^{N(H)} \Phi_j^n \Delta_n^{-\frac{1}{2}}\left\{\left(\widetilde{C}_t^{n,N(\widetilde{H}^n)}\right)^j - \left(\widetilde{C}_t^{n,N(H)}\right)^j\right\} \\
&\quad = \Delta_n^{-\frac{1}{2}}\left\{\widetilde{H}^n - H + \sum_{j=1}^{N(H)} \Phi_j^n C_t^j\right\} + \Phi_1^n \Delta_n^{-\frac{1}{2}}\left\{\widetilde{C}_t^{n,1} - C_t + \sum_{k=2}^{N(H)} \Psi_k^n C_t^k\right\} + \mathbb{I}_1^n,
\end{aligned}
\tag{4.3.20}
$$

where

$$
\begin{aligned}
\mathbb{I}_1^n = {} & \sum_{k=2}^{N(H)} \Phi_1^n \Psi_k^n \Delta_n^{-\frac{1}{2}}\left\{(\widetilde{C}_t^{n,N(H)-k+1})^k - C_t^k\right\} + \sum_{j=2}^{N(H)} \Phi_j^n \Delta_n^{-\frac{1}{2}}\left\{(\widetilde{C}_t^{n,N(H)})^j - C_t^j\right\} \\
& + \sum_{j=1}^{N(H)} \Phi_j^n \Delta_n^{-\frac{1}{2}}\left\{\left(\widetilde{C}_t^{n,N(\widetilde{H}^n)}\right)^j - \left(\widetilde{C}_t^{n,N(H)}\right)^j\right\}.
\end{aligned}
$$

By the proof of Proposition 3.3.7 and the mean-value theorem, $(\widetilde{C}_t^{n,N(H)-k+1})^k - C_t^k$ is of size $\Delta_n^{(N(H)-k+1)(1-2H)}$ and $(\widetilde{C}_t^{n,N(H)})^j - C_t^j$ is of size $\Delta_n^{2H-1/2}$. Furthermore, recalling the definition (3.3.19) of Φ_j^n, we see that $\Phi_j^n \Delta_n^{-j(1-2H)}$ converges in probability. Hence, $\Phi_j^n\{(\widetilde{C}_t^{n,N(H)})^j - C_t^j\}$ is of size $\Delta_n^{1/2+(j-1)(1-2H)}$. Also, Ψ_k^n is of size $\Delta_n^{(k-1)(1-2H)}$, so $\Phi_1^n \Psi_k^n \Delta_n^{-1/2}\{(\widetilde{C}_t^{n,N(H)-k+1})^k - C_t^k\}$ of size $\Delta_n^{(N(H)+1)(1-2H)-1/2}$. Recall also that $\Delta_n^{-1/2}\{\widetilde{C}_t^{n,N(\widetilde{H}^n)} - \widetilde{C}_t^{n,N(H)}\}$ is asymptotically negligible by the last part of the proof of Proposition 3.3.7. Altogether, we obtain that $\mathbb{I}_1^n$ is asymptotically negligible.

Now recall the precise definition (3.3.22) of $\widetilde{C}_t^{n,1}$ as well as that of $\Theta(\widehat{V}_t^n, \widetilde{H}^n, \widetilde{H}^{n,0})$ and Ψ_k^n

given in (3.3.23) and (3.3.26), respectively. With those definitions at hand, we can decompose

$$
\begin{aligned}
&\Delta_n^{\frac{1}{2}-2H}\left\{\widetilde{C}_t^{n,1}-C_t+\sum_{k=2}^{N(H)}\Psi_k^n C_t^k\right\}\\
&\quad=\Delta_n^{\frac{1}{2}-2H}\Theta(\widehat{V}_t^n,\widetilde{H}^n,\widetilde{H}^{n,0})^{-1}\Bigg\{\left\{V_{0,t}^n-\frac{\langle a,V_t^n\rangle}{\langle a,\Gamma^{\widetilde{H}^n}\rangle}\right\}\Delta_n^{2H-1}-\Theta(\widehat{V}_t^n,\widetilde{H}^n,\widetilde{H}^{n,0})C_t\\
&\qquad+\frac{\Delta_n^{1-2H}\widehat{P}_t^n}{\langle a,\Gamma^{\widetilde{H}^n}\rangle}\sum_{j=2}^{N(H)}\frac{(-1)^j}{j!}\psi^{(j)}(\phi(\widetilde{H}^n))\frac{a_0^j}{\langle b,V_t^n\rangle^j}C_t^j\,\Delta_n^{(j-1)(1-2H)}\Bigg\}\\
&\quad=\Delta_n^{-\frac{1}{2}}\Theta(\widehat{V}_t^n,\widetilde{H}^n,\widetilde{H}^{n,0})^{-1}\Bigg\{\left\{V_{0,t}^n-\frac{\langle a,V_t^n\rangle}{\langle a,\Gamma^{\widetilde{H}^n}\rangle}\right\}\\
&\qquad-\left(1-\frac{a_0}{\langle a,\Gamma^{\widetilde{H}^n}\rangle}+\frac{\Pi_t}{\langle a,\Gamma^{\widetilde{H}^n}\rangle}\frac{a_0\psi'(\phi(\widetilde{H}^n))}{\langle b,V_t^n\rangle}\right)C_t\Delta_n^{1-2H}\\
&\qquad+\frac{\Pi_t}{\langle a,\Gamma^{\widetilde{H}^n}\rangle}\sum_{j=2}^{N(H)}\frac{(-1)^j}{j!}\psi^{(j)}(\phi(\widetilde{H}^n))\frac{a_0^j}{\langle b,V_t^n\rangle^j}C_t^j\,\Delta_n^{j(1-2H)}\Bigg\}\\
&\qquad+\Theta(\widehat{V}_t^n,\widetilde{H}^n,\widetilde{H}^{n,0})^{-1}\frac{\Delta_n^{-\frac{1}{2}}\{\Delta_n^{1-2H}\widehat{P}_t^n-\Pi_t\}}{\langle a,\Gamma^{\widetilde{H}^n}\rangle}\sum_{j=1}^{N(H)}\frac{(-1)^j}{j!}\psi^{(j)}(\phi(\widetilde{H}^n))\frac{a_0^j}{\langle b,V_t^n\rangle^j}C_t^j\,\Delta_n^{j(1-2H)}.
\end{aligned}
\tag{4.3.21}
$$

The last term in the above display is asymptotically negligible as already seen in the discussion following (4.3.14), while the first term on the right-hand side of (4.3.21) was analyzed in the (4.3.12). Combining this with the Taylor expansion (4.3.9) of $\widetilde{H}^n$, we can continue the computations started in (4.3.20). We have

$$
\begin{aligned}
&\Delta_n^{-\frac{1}{2}}\left\{\widetilde{H}^n-H+\sum_{j=1}^{N(H)}\Phi_j^n\left(\widehat{C}_t^{n,1}\right)^j\right\}\\
&\quad=(\phi^{-1})'(\phi(\widetilde{H}^n))\frac{1}{\langle b,V_t^n\rangle}\left\{a^T-\varphi(\widetilde{H}^n)b^T\right\}\Delta_n^{-\frac{1}{2}}\left\{V_t^n-\Gamma^H\Pi_t-e_1C_t\,\Delta_n^{1-2H}\right\}\\
&\qquad+(\Phi_1^n\Delta_n^{2H-1})\Theta(\widehat{V}_t^n,\widetilde{H}^n,\widetilde{H}^{n,0})^{-1}\left\{e_1^T-\frac{a^T}{\langle a,\Gamma^{\widetilde{H}^n}\rangle}+\frac{\Pi_t}{\langle a,\Gamma^{\widetilde{H}^n}\rangle}\frac{\psi'(\phi(\widetilde{H}^n))}{\langle b,V_t^n\rangle}(a^T-\phi(\widetilde{H}^n)b^T)\right\}\\
&\qquad\times\Delta_n^{-\frac{1}{2}}\left\{V_t^n-\Gamma^H\Pi_t-e_1C_t\Delta_n^{1-2H}\right\}+\widehat{\mathbb{I}}_1^n\\
&\quad=w_1(\widetilde{H}^n,\widetilde{H}^{n,0},V_t^n)\Delta_n^{-\frac{1}{2}}\left\{V_t^n-\Gamma^H\Pi_t-e_1C_t\Delta_n^{1-2H}\right\}+\widehat{\mathbb{I}}_1^n,
\end{aligned}
\tag{4.3.22}
$$

where

$$
\begin{aligned}
w_1(\widetilde{H}^n,\widetilde{H}^{n,0},V_t^n)&=\frac{(\phi^{-1})'(\phi(\widetilde{H}^n))}{\langle b,V_t^n\rangle}\left\{a^T-\phi(\widetilde{H}^n)b^T-a_0u_1(\widetilde{H}^n,\widetilde{H}^{n,0},V_t^n)\right\},\\
u_1(\widetilde{H}^n,\widetilde{H}^{n,0},V_t^n)&=\frac{1}{\Theta(\widehat{V}_t^n,\widetilde{H}^n,\widetilde{H}^{n,0})}\left(e_1^T-\frac{a^T}{\langle a,\Gamma^{\widetilde{H}^n}\rangle}+\frac{\Pi_t\psi'(\phi(\widetilde{H}^n))}{\langle b,V_t^n\rangle\langle a,\Gamma^{\widetilde{H}^n}\rangle}(a^T-\phi(\widetilde{H}^n)b^T)\right).
\end{aligned}
\tag{4.3.23}
$$

In $\widehat{\mathbb{I}}_1^n$, we have incorporated the last three terms on the right-hand side of (4.3.9), the last two terms on the right-hand side of (4.3.12), the last expression in (4.3.21) as well as $\mathbb{I}_1^n$ from

(4.3.20). By the discussions following these equations, we know that $\widehat{\mathbb{I}}_1^n$ is asymptotically negligible. Therefore, we obtain

$$\Delta_n^{-\frac{1}{2}} \left\{ \widetilde{H}^n - H + \sum_{j=1}^{N(H)} \Phi_j^n \left(\widehat{C}_t^{n,1} \right)^j \right\} \xrightarrow{\text{st}} \frac{w_1^T}{\Pi_t} \mathcal{Z}_t \sim \mathcal{N} \left(0, \mathrm{Var}_{H,1} \frac{\int_0^t \rho_s^4 \,\mathrm{d}s}{(\int_0^t \rho_s^2 \,\mathrm{d}s)^2} \right).$$

To conclude, it remains to observe that this CLT is not affected when $N(H)$ is replaced by $N(\widetilde{H}^n)$ because $H \notin \mathcal{H}$; cf. the argument used to show (4.3.19). $\square$

Proof of Proposition 3.3.9. For $k = 2, \ldots, m$, define

$$\begin{aligned} u_k(\widehat{H}_{k-1}^n, \widetilde{H}^n, V_t^n) &= \left\{ e_1^T - \frac{a^T}{\langle a, \Gamma^{\widehat{H}_{k-1}^n} \rangle} + \frac{\Pi_t}{\langle a, \Gamma^{\widehat{H}_{k-1}^n} \rangle} \langle a, \partial_H \Gamma^H \rangle w_{k-1}(\widehat{H}_{k-2}^n, \widetilde{H}^n, V_t^n) \right\} \\ &\quad \times \left(1 - \frac{a_0}{\langle a, \Gamma^{\widehat{H}_{k-1}^n} \rangle} \right)^{-1}, \\ w_k(\widehat{H}_{k-1}^n, \widetilde{H}^n, V_t^n) &= \frac{(\phi^{-1})'(\phi(\widetilde{H}^n))}{\langle b, V_t^n \rangle} \left\{ a^T - \phi(\widetilde{H}^n) b^T - a_0 u_k(\widehat{H}_{k-1}^n, \widetilde{H}^n, V_t^n) \right\}. \end{aligned}$$

In the definition of $u_2(\widehat{H}_{k-1}^n, \widetilde{H}^n, V_t^n)$, the term $w_1(\widehat{H}_{k-1}^n, \widetilde{H}^n, V_t^n)$ is replaced by $w_1(\widetilde{H}^n, \widetilde{H}^{n,0}, V_t^n)$ from (4.3.23). By induction over k, we are going to show for all $k = 1, \ldots, m$ that

$$\Delta_n^{-\frac{1}{2}} (\widehat{H}_k^n - H) = w_k(\widehat{H}_{k-1}^n, \widetilde{H}^n, V_t^n) \Delta_n^{-\frac{1}{2}} \left\{ V_t^n - \Gamma^H \Pi_t - e_1 C_t \Delta_n^{1-2H} \right\} + \widehat{\mathbb{I}}_k^n \tag{4.3.24}$$

for some asymptotically negligible expression $\widehat{\mathbb{I}}_k^n$ and

$$u_k(\widehat{H}_{k-1}^n, \widetilde{H}^n, V_t^n) \xrightarrow{P} u_k^T, \qquad w_k(\widehat{H}_{k-1}^n, \widetilde{H}^n, V_t^n) \xrightarrow{P} \frac{w_k^T}{\Pi_t}, \tag{4.3.25}$$

where, for $k = 1$, we take the expressions in (4.3.23) instead. Since (4.3.24) was already shown in (4.3.22) and (4.3.25) is obvious for $k = 1$, we may consider $k \geq 2$ now and assume (4.3.24) and (4.3.25) for $k - 1$. In particular,

$$\Delta_n^{-\frac{1}{2}} \left\{ \widehat{H}_{k-1}^n - H \right\} \xrightarrow{\text{st}} \frac{w_{k-1}^T}{\Pi_t} \mathcal{Z}_t \sim \mathcal{N} \left(0, \mathrm{Var}_{H,k-1} \frac{\int_0^t \rho_s^4 \,\mathrm{d}s}{(\int_0^t \rho_s^2 \,\mathrm{d}s)^2} \right).$$

It is straightforward to see that

$$u_k(\widehat{H}_{k-1}^n, \widetilde{H}^n, V_t^n) \xrightarrow{P} u_k^T \qquad \text{and} \qquad w_k(\widehat{H}_{k-1}^n, \widetilde{H}^n, V_t^n) \xrightarrow{P} \frac{w_k^T}{\Pi_t}, \tag{4.3.26}$$

so we can proceed to showing (4.3.24) for k. Expanding $\langle a, \Gamma^{\widehat{H}_{k-1}^n} \rangle$ around H and using the induction hypothesis, we can find β_{k-1}^n between $\widehat{H}_{k-1}^n$ and H such that

$$\begin{aligned}
&\Delta_n^{-\frac{1}{2}}\left\{V_{0,t}^n-\frac{\langle a,V_t^n\rangle}{\langle a,\Gamma^{\widehat{H}_{k-1}^n}\rangle}-\left(1-\frac{a_0}{\langle a,\Gamma^{\widehat{H}_{k-1}^n}\rangle}\right)C_t\Delta_n^{1-2H}\right\}\\
&\quad=\Delta_n^{-\frac{1}{2}}\left\{V_{0,t}^n-\Pi_t-C_t\Delta_n^{1-2H}\right\}-\frac{1}{\langle a,\Gamma^{\widehat{H}_{k-1}^n}\rangle}\Delta_n^{\frac{1}{2}}\left\{\langle a,V_t^n\rangle-\langle a,\Gamma^H\rangle\Pi_t-a_0C_t\Delta_n^{1-2H}\right\}\\
&\qquad+\frac{\Pi_t}{\langle a,\Gamma^{\widehat{H}_{k-1}^n}\rangle}\Delta_n^{-\frac{1}{2}}\left\{\langle a,\Gamma^{\widehat{H}_{k-1}^n}\rangle-\langle a,\Gamma^H\rangle\right\}\\
&\quad=\Delta_n^{-\frac{1}{2}}\left\{V_{0,t}^n-\Pi_t-C_t\Delta_n^{1-2H}\right\}-\frac{1}{\langle a,\Gamma^{\widehat{H}_{k-1}^n}\rangle}\Delta_n^{\frac{1}{2}}\left\{\langle a,V_t^n\rangle-\langle a,\Gamma^H\rangle\Pi_t-a_0C_t\Delta_n^{1-2H}\right\}\\
&\qquad+\frac{\Pi_t}{\langle a,\Gamma^{\widehat{H}_{k-1}^n}\rangle}\langle a,\partial_H\Gamma^H\rangle\left\{w_{k-1}(\widehat{H}_{k-2}^n,\widetilde{H}^n,V_t^n)\Delta_n^{-\frac{1}{2}}\left\{V_t^n-\Gamma^H\Pi_t-e_1C_t\Delta_n^{1-2H}\right\}+\widehat{\mathbb{I}}_{k-1}^n\right\}\\
&\qquad+\frac{1}{2!}\frac{\Pi_t}{\langle a,\Gamma^{\widehat{H}_{k-1}^n}\rangle}\langle a,\partial_{HH}\Gamma^{\bar{\xi}^n}\rangle\Delta_n^{-\frac{1}{2}}\left\{\widehat{H}_{k-1}^n-H\right\}^2\\
&\quad=u_k(\widehat{H}_{k-1}^n,\widetilde{H}^n,V_t^n)\left(1-\frac{a_0}{\langle a,\Gamma^{\widehat{H}_{k-1}^n}\rangle}\right)\Delta_n^{-\frac{1}{2}}\left\{V_t^n-\Gamma^H\Pi_t-e_1C_t\Delta_n^{1-2H}\right\}+\mathbb{J}_k^n,
\end{aligned}\tag{4.3.27}$$

where $\mathbb{J}_k^n$ is given by

$$\frac{\Pi_t}{\langle a,\Gamma^{\widehat{H}_{k-1}^n}\rangle}\langle a,\partial_H\Gamma^H\rangle\widehat{\mathbb{I}}_{k-1}^n+\frac{1}{2!}\frac{\Pi_t}{\langle a,\Gamma^{\widehat{H}_{k-1}^n}\rangle}\langle a,\partial_{HH}\Gamma^{\beta_{k-1}^n}\rangle\Delta_n^{-\frac{1}{2}}\left\{\widehat{H}_{k-1}^n-H\right\}^2.$$

Because $\widehat{I}_{k-1}^n$ is asymptotically negligible by induction hypothesis and $\widehat{H}_{k-1}^n-H$ is of size $\Delta_n^{1/2}$, we see that $\mathbb{J}_k^n\xrightarrow{P}0$. Recall the definition of $\widehat{C}_t^{n,k}$ given in (3.3.31). From (4.3.27) and (4.3.26), we infer that

$$\Delta_n^{\frac{1}{2}-2H}(\widehat{C}_t^{n,k}-C_t)\xrightarrow{\text{st}}u_k^T\mathcal{Z}_t\sim\mathcal{N}\left(0,\mathrm{Var}_{C,k}\int_0^t\rho_s^4\,\mathrm{d}s\right),$$

which (3.3.34). Now recall the definitions (3.3.31) and (3.3.32). Using (4.3.9) and the formula $\Phi_1^n=-\Delta_n^{1-2H}(\phi^{-1})(\phi(\widetilde{H}^n))a_0/\langle b,V_t^n\rangle$ for the second equality and (4.3.27) for the third, we obtain

$$\begin{aligned}
\Delta_n^{-\frac{1}{2}}\left\{\widehat{H}_k^n-H\right\}&=\Delta_n^{-\frac{1}{2}}\left\{\widetilde{H}^n-H+\sum_{j=1}^{N(\widehat{H}_{k-1}^n)}\Phi_j^nC_t^j\right\}+\Phi_1^n\Delta_n^{-\frac{1}{2}}\left\{\widehat{C}_t^{n,k}-C_t\right\}\\
&\quad+\sum_{j=2}^{N(\widehat{H}_{k-1}^n)}\Phi_j^n\Delta_n^{-\frac{1}{2}}\left\{\left(\widehat{C}_t^{n,k}\right)^j-C_t^j\right\}\\
&=\frac{(\phi^{-1})'(\phi(\widetilde{H}^n))}{\langle b,V_t^n\rangle}\left\{a^T-\phi(\widetilde{H}^n)b^T-a_0u_k(\widehat{H}_{k-1}^n,\widetilde{H}^n,V_t^n)\right\}\\
&\quad\times\Delta_n^{-\frac{1}{2}}\left\{V_t^n-\Gamma^H\Pi_t-e_1C_t\Delta_n^{1-2H}\right\}+\widehat{\mathbb{I}}_k^n\\
&=w_k(\widehat{H}_{k-1}^n,\widetilde{H}^n,V_t^n)\Delta_n^{-\frac{1}{2}}\left\{V_t^n-\Gamma^H\Pi_t-e_1C_t\Delta_n^{1-2H}\right\}+\widehat{\mathbb{I}}_k^n.
\end{aligned}\tag{4.3.28}$$

In the last line, $\widehat{\mathbb{I}}_k^n$ contains the last three terms on the right-hand side of (4.3.9) as well as

$$\mathbb{J}_k^n\left(1-\frac{a_0}{\langle a,\Gamma^{\widehat{H}_{k-1}^n}\rangle}\right)^{-1}+\sum_{j=2}^{N(H)}\Phi_j^n\Delta_n^{-\frac{1}{2}}\left\{\left(\widehat{C}_t^{n,k}\right)^j-C_t^j\right\}+\Delta_n^{-\frac{1}{2}}\sum_{j=N(H)\wedge N(\widehat{H}_{k-1}^n)}^{N(H)\vee N(\widehat{H}_{k-1}^n)}\Phi_j^n\left(\widehat{C}_t^{n,k}\right)^j.$$

The term $\Phi_j^n\Delta_n^{-1/2}\{(\widehat{C}_t^{n,k})^j-C_t^j\}$ is of size $\Delta_n^{(j-1)(1-2H)}$ because Φ_j^n is of size $\Delta_n^{j(1-2H)}$. Also, the last sum goes to 0 in probability by a similar argument to (4.3.19). Therefore, $\widehat{\mathbb{I}}_k^n$ is asymptotically negligible. This together with (4.3.28) implies (4.3.24) and our induction argument is complete. From (4.3.24), we immediately obtain (3.3.33). □

Proof of Theorem 3.3.10. The proof of (3.3.39) is completely analogous to that for (3.3.34) in Theorem 3.3.9. while the proof of (3.3.40) follows the same steps as that for (3.3.12) in Theorem 3.3.2. □

4.4 Estimates for fractional kernels

Here we gather some useful results about the kernel $g(t)=K_H^{-1}t^{H-1/2}$ introduced in (3.1.6) [we consider the case $g_0\equiv 0$ here].

Lemma 4.4.1. *Recall the notation* $\underline{\Delta}_i^n g$ *introduced in* (3.2.13). *Also recall the constant* K_H *in* (3.1.7) *and the numbers* $(\Gamma_r^H)_{r\geq 0}$ *in* (3.1.9).

(1) For any $k,n\in\mathbb{N}$,

$$\int_0^\infty \Delta_k^n g(t)^2\,\mathrm{d}t=K_H^{-2}\left\{\frac{1}{2H}+\int_1^k\left(r^{H-\frac{1}{2}}-(r-1)^{H-\frac{1}{2}}\right)^2\mathrm{d}r\right\}\Delta_n^{2H}\leq\Delta_n^{2H}. \tag{4.4.1}$$

(2) For any $k,\ell,n\in\mathbb{N}$ *with* $k<\ell$,

$$\begin{aligned}\int_0^\infty \Delta_k^n g(t)\Delta_\ell^n g(t)\,\mathrm{d}t&=\Delta_n^{2H}K_H^{-2}\int_0^k\left(r^{H-\frac{1}{2}}-(r-1)_+^{H-\frac{1}{2}}\right)\\&\qquad\times\left((r+\ell-k))^{H-\frac{1}{2}}-(r+(\ell-k)-1)^{H-\frac{1}{2}}\right)\mathrm{d}r,\end{aligned} \tag{4.4.2}$$

$$\int_{-\infty}^\infty \Delta_k^n g(t)\Delta_\ell^n g(t)\,\mathrm{d}t=\Delta_n^{2H}\Gamma_{\ell-k}^H\lesssim\Delta_n^{2H}\overline{\Gamma}_{\ell-k}^H$$

where

$$\overline{\Gamma}_r^H=\frac{1}{(r-1)^{2(1-H)}},\quad r\geq 2,\qquad \overline{\Gamma}_1^H=\Gamma_1^H. \tag{4.4.3}$$

(3) For any $\theta\in(0,1)$, *setting* $\theta_n=[\Delta_n^{-\theta}]$, *we have for any* $i>\theta_n$ *and* $r\in\mathbb{N}$,

$$\int_{-\infty}^{(i-\theta_n)\Delta_n}\Delta_i^n g(s)\Delta_{i+r}^n g(s)\,\mathrm{d}s\lesssim\Delta_n^{2H}\Delta_n^{2\theta(1-H)}. \tag{4.4.4}$$

Proof. Let $k\leq\ell$. By direct calculation,

$$\begin{aligned}&\int_0^\infty \Delta_k^n g(t)\Delta_\ell^n g(t)\,\mathrm{d}t\\&=K_H^{-2}\int_0^{k\Delta_n}\left(s^{H-\frac{1}{2}}-(s-\Delta_n)_+^{H-\frac{1}{2}}\right)\left((s+(\ell-k)\Delta_n)^{H-\frac{1}{2}}-(s+(\ell-k)\Delta_n-\Delta_n)_+^{H-\frac{1}{2}}\right)\mathrm{d}s\\&=\Delta_n^{2H}K_H^{-2}\int_0^k\left(r^{H-\frac{1}{2}}-(r-1)_+^{H-\frac{1}{2}}\right)\left((r+(\ell-k))^{H-\frac{1}{2}}-(r+(\ell-k)-1)_+^{H-\frac{1}{2}}\right)\mathrm{d}r,\end{aligned}$$

which is the first equality in (4.4.2). If further $k = \ell$, then

$$\int_0^\infty \Delta_k^n g(t)^2 \,\mathrm{d}t = \Delta_n^{2H} K_H^{-2} \left(\int_0^1 r^{2H-1} \,\mathrm{d}r + \int_1^k \left(r^{H-\frac{1}{2}} - (r-1)^{H-\frac{1}{2}} \right)^2 \mathrm{d}r \right),$$

which shows (4.4.1).

Now let $(B^H)_{t\geq 0}$ be a fractional Brownian motion with Hurst index H. Then B^H has the Mandelbrot–van Ness representation

$$B_t^H = K_H^{-1} \int_{\mathbb{R}} \left((t-s)_+^{H-\frac{1}{2}} - (-s)_+^{H-\frac{1}{2}} \right) \mathrm{d}\overline{B}_s, \quad t \geq 0,$$

where $\overline{B}$ is a two-sided standard Brownian motion. Then

$$\Delta_i^n B^H = \int_{\mathbb{R}} \Delta_i^n g(s) \,\mathrm{d}\overline{B}_s$$

for any i. Therefore, by well-known properties of fractional Brownian motion,

$$\begin{aligned}
\int_{-\infty}^{\infty} \Delta_k^n g(s) \Delta_\ell^n g(s) \,\mathrm{d}s &= \mathbb{E}\left[\Delta_k^n B^H \Delta_\ell^n B^H \right] \\
&= \mathbb{E}\left[B_{\Delta_n}^H \Delta_{\ell-k+1}^n B^H \right] = \mathbb{E}\left[B_{\Delta_n}^H B_{(\ell-k+1)\Delta_n}^H \right] - \mathbb{E}\left[B_{\Delta_n}^H B_{(\ell-k)\Delta_n}^H \right] \\
&= \frac{1}{2} \Big\{ \Delta_n^{2H} + ((\ell-k+1)\Delta_n)^{2H} - ((\ell-k)\Delta_n)^{2H} \\
&\qquad - \Delta_n^{2H} - ((\ell-k)\Delta_n)^{2H} + ((\ell-k-1)\Delta_n)^{2H} \Big\} \\
&= \Delta_n^{2H} \Gamma_{\ell-k}^H,
\end{aligned}$$

which is the second equality in (4.4.2). Next, use the mean value theorem twice on Γ_r^H in order to obtain for all $r \geq 2$,

$$\begin{aligned}
\Gamma_r^H &= \frac{1}{2} \left(\left\{ (r+1)^{2H} - r^{2H} \right\} - \left\{ r^{2H} - (r-1)^{2H} \right\} \right) \leq \frac{1}{2} (2H) \left((r+1)^{2H-1} - (r-1)^{2H-1} \right) \\
&\leq H(2H-1)(r-1)^{2H-2},
\end{aligned}$$

which completes the proof of (4.4.3). Finally,

$$\begin{aligned}
&\int_{-\infty}^{(i-\theta_n)\Delta_n} \Delta_i^n g(s) \Delta_{i+r}^n g(s) \,\mathrm{d}s \\
&\quad = \Delta_n^{2H} K_H^{-2} \int_{\theta_n}^{\infty} \left(t^{H-\frac{1}{2}} - (t-1)^{H-\frac{1}{2}} \right) \left((t+r)^{H-\frac{1}{2}} - (t+r-1)^{H-\frac{1}{2}} \right) \mathrm{d}t \\
&\quad \lesssim \Delta_n^{2H} \int_{\theta_n}^{\infty} \left(t^{H-\frac{1}{2}} - (t-1)^{H-\frac{1}{2}} \right)^2 \mathrm{d}t \lesssim \Delta_n^{2H} \int_{\theta_n}^{\infty} (t-1)^{2H-3} \,\mathrm{d}t \lesssim \Delta_n^{2H} \Delta_n^{\theta(2-2H)},
\end{aligned}$$

which yields (4.4.4). □

Bibliography

[1] Y. Aït-Sahalia and J. Jacod. Estimating the degree of activity of jumps in high frequency data. *Ann. Statist.*, 37(5A):2202–2244, 2009.

[2] Y. Aït-Sahalia and J. Jacod. Testing for jumps in a discretely observed process. *Ann. Statist.*, 37(1):184–222, 2009.

[3] Y. Aït-Sahalia and J. Jacod. *High-Frequency Financial Econometrics*. Princeton University Press, Princeton, 2014.

[4] Y. Aït-Sahalia, P. A. Mykland, and L. Zhang. How often to sample a continuous-time process in the presence of market microstructure noise. *Rev. Financ. Stud.*, 18(2):351–416, 2005.

[5] Y. Aït-Sahalia, P. A. Mykland, and L. Zhang. Ultra high frequency volatility estimation with dependent microstructure noise. *J. Econometrics*, 160(1):160–175, 2011.

[6] Y. Aït-Sahalia and D. Xiu. A Hausman test for the presence of market microstructure noise in high frequency data. *J. Econometrics*, 211(1):176–205, 2019.

[7] D. Aldous. Stopping times and tightness. *Ann. Probab.*, 6(2):335–340, 1978.

[8] T. Andersen, T. Bollerslev, F. Diebold, and P. Labys. Great realizations. *Risk*, 13:105–108, 2000.

[9] 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.

[10] V. Bally, A. Millet, and M. Sanz-Solé. Approximation and support theorem in Hölder norm for parabolic stochastic partial differential equations. *Ann. Probab.*, 23(1):178–222, 1995.

[11] F. M. Bandi and J. R. Russell. Separating microstructure noise from volatility. *J. Financ. Econ.*, 79(3):655–692, 2006.

[12] F. M. Bandi, J. R. Russell, and C. Yang. Realized volatility forecasting and option pricing. *J. Econometrics*, 147(1):34–46, 2008.

[13] X. Bardina, M. Jolis, and L. Quer-Sardanyons. Weak convergence for the stochastic heat equation driven by Gaussian white noise. *Electron. J. Probab.*, 15(1):1267–1295, 2010.

[14] O. Barndorff-Nielsen, J. Corcuera, and M. Podolskij. Multipower variation for Brownian semistationary processes. *Bernoulli*, 17(4):1159–1194, 2011.

[15] O. E. Barndorff-Nielsen, P. R. Hansen, A. Lunde, and N. Shephard. Designing realized kernels to measure the ex post variation of equity prices in the presence of noise. *Econometrica*, 76(6):1481–1536, 2008.

[16] O. E. Barndorff-Nielsen, P. R. Hansen, A. Lunde, and N. Shephard. Subsampling realised kernels. *J. Econometrics*, 160(1):204–219, 2011.

[17] O. E. Barndorff-Nielsen and N. Shephard. Power variation and time change. *Teor. Veroyatn. Primen.*, 50(1):115–130, 2005.

[18] E. Bayraktar, H. V. Poor, and K. R. Sircar. Estimating the fractal dimension of the S&P 500 index using wavelet analysis. *Int. J. Theor. Appl. Finance*, 7(5):615–643, 2004.

[19] C. Bender, T. Sottinen, and E. Valkeila. Fractional processes as models in stochastic finance. In *Advanced mathematical methods for finance*, pages 75–103. Springer, Heidelberg, 2011.

[20] S. Bianchi and A. Pianese. Time-varying Hurst-Hölder exponents and the dynamics of (in)efficiency in stock markets. *Chaos Solitons Fractals*, 109:64–75, 2018.

[21] A. Brouste and M. Fukasawa. Local asymptotic normality property for fractional Gaussian noise under high-frequency observations. *Ann. Statist.*, 46(5):2045–2061, 2018.

[22] C. Cai, P. Chigansky, and M. Kleptsyna. Mixed Gaussian processes: a filtering approach. *Ann. Probab.*, 44(4):3032–3075, 2016.

[23] R. Cairoli and J. B. Walsh. Stochastic integrals in the plane. *Acta Math.*, 134:111–183, 1975.

[24] B. Chen, C. Chong, and C. Klüppelberg. Simulation of stochastic Volterra equations driven by space–time Lévy noise. In M. Podolskij, R. Stelzer, S. Thorbjørnsen, and A. Veraart, editors, *The Fascination of Probability, Statistics and their Applications*, pages 209–229. Springer, Cham, 2016.

[25] P. Cheridito. Mixed fractional Brownian motion. *Bernoulli*, 7(6):913–934, 2001.

[26] A. Cherny. Brownian moving averages have conditional full support. *Ann. Appl. Probab.*, 18(5):1825–1830, 2008.

[27] C. Chong. High-frequency analysis of parabolic stochastic PDEs with multiplicative noise. Submitted, 2020. Preprint under arXiv:1908.04145.

[28] C. Chong. Lévy-driven Volterra equations in space and time. *J. Theor. Probab.*, 30(3):1014–1058, 2017.

[29] C. Chong. High-frequency analysis of parabolic stochastic PDEs. *Ann. Statist.*, 48(2):1143–1167, 2020.

[30] C. Chong. Supplement to “High-frequency analysis of parabolic stochastic PDEs”. 60 pages, 2020.

[31] C. Chong, R. C. Dalang, and T. Humeau. Path properties of the solution to the stochastic heat equation with Lévy noise. *Stoch. Partial Differ. Equ. Anal. Comput.*, 7(1):123–168, 2019.

[32] C. Chong and T. Delerue. Normal approximation of the solution to the stochastic heat equation with Lévy noise. *Stoch. Partial Differ. Equ. Anal. Comput.*, 8(2):362–401, 2020.

[33] C. Chong, T. Delerue, and G. Li. Mixed semimartingales: Volatility estimation in the presence of fractional noise. Submitted, 2021. Preprint under arXiv:2106.16149.

[34] C. Chong, T. Delerue, and G. Li. Supplement to "Mixed semimartingales: Volatility estimation in the presence of fractional noise". Submitted, 2021. Preprint under arXiv:2106.16149.

[35] K. Christensen, M. Podolskij, and M. Vetter. Bias-correcting the realized range-based variance in the presence of market microstructure noise. *Finance Stoch.*, 13(2):239–268, 2009.

[36] S. Cohen and J. Rosiński. Gaussian approximation of multivariate Lévy processes with applications to simulation of tempered stable processes. *Bernoulli*, 13(1):195–210, 2007.

[37] J. M. Corcuera, E. Hedevang, M. S. Pakkanen, and M. Podolskij. Asymptotic theory for Brownian semi-stationary processes with application to turbulence. *Stochastic Process. Appl.*, 123(7):2552–2574, 2013.

[38] R. Da and D. Xiu. When moving-average models meet high-frequency data: Uniform inference on volatility. Chicago Booth Research Paper No. 17-27, Fama-Miller Working Paper, available at SSRN: http://dx.doi.org/10.2139/ssrn.3043834, 2019.

[39] G. Da Prato, S. Kwapien, and J. Zabczyk. Regularity of solutions of linear stochastic equations in Hilbert spaces. *Stochastics*, 23(1):1–23, 1987.

[40] R. Dalang, D. Khoshnevisan, C. Mueller, D. Nualart, and Y. Xiao. *A Minicourse on Stochastic Partial Differential Equations*. Springer, Berlin, 2009.

[41] S. Delattre and J. Jacod. A central limit theorem for normalized functions of the increments of a diffusion process, in the presence of round-off errors. *Bernoulli*, 3(1):1–28, 1997.

[42] F. Delbaen and W. Schachermayer. A general version of the fundamental theorem of asset pricing. *Math. Ann.*, 300(3):463–520, 1994.

[43] T. Delerue. Normal approximation of the solution to the stochastic wave equation with Lévy noise. Submitted, 2021. Preprint under arXiv:1911.01811.

[44] F. Demengel and G. Demengel. *Functional Spaces for the Theory of Elliptic Partial Differential Equations*. Springer, London, 2012.

[45] J. L. Doob. *Stochastic Processes*. John Wiley & Sons, Inc., New York; Chapman & Hall, Limited, London, 1953.

[46] A. A. Douglas. *Underwater Acoustic Signal Processing: Modeling, Detection, and Estimation*. Springer, Cham, 2019.

[47] M. Dozzi, Y. Mishura, and G. Shevchenko. Asymptotic behavior of mixed power variations and statistical estimation in mixed models. *Stat. Inference Stoch. Process.*, 18(2):151–175, 2015.

[48] N. Fournier. Simulation and approximation of Lévy-driven stochastic differential equations. *ESAIM Probab. Stat.*, 15(1):233–248, 2011.

[49] A. Friedman. *Partial Differential Equations of Parabolic Type.* Prentice-Hall, Englewood Cliffs, NJ, 1964.

[50] M. Fukasawa and T. Takabatake. Asymptotically efficient estimators for self-similar stationary Gaussian noises under high frequency observations. *Bernoulli*, 25(3):1870–1900, 2019.

[51] J. Gatheral and R. C. A. Oomen. Zero-intelligence realized variance estimation. *Finance Stoch.*, 14(2):249–283, 2010.

[52] I. I. Gikhman and A. V. Skorokhod. *The Theory of Stochastic Processes I.* Springer, New-York, 1974.

[53] P. Guasoni, M. Rásonyi, and W. Schachermayer. Consistent price systems and face-lifting pricing under transaction costs. *Ann. Appl. Probab.*, 18(2):491–520, 2008.

[54] M. Hairer and É. Pardoux. A Wong-Zakai theorem for stochastic PDEs. *J. Math. Soc. Japan*, 67(4):1551–1604, 2015.

[55] P. R. Hansen and A. Lunde. Realized variance and market microstructure noise. *J. Bus. Econom. Statist.*, 24(2):127–218, 2006. With comments and a rejoinder by the authors.

[56] N. Hautsch and M. Podolskij. Preaveraging-based estimation of quadratic variation in the presence of noise and jumps: theory, implementation, and empirical evidence. *J. Bus. Econom. Statist.*, 31(2):165–183, 2013.

[57] T. Hayashi and N. Yoshida. Asymptotic normality of a covariance estimator for nonsynchronously observed diffusion processes. *Ann. Inst. Statist. Math.*, 60(2):367–406, 2008.

[58] B. G. Ivanoff. The function space $D([0,\infty)^q, E)$. *Canad. J. Statist.*, 8(2):179–181, 1980.

[59] B. G. Ivanoff. Stopping times and tightness for multiparameter martingales. *Statist. Probab. Lett.*, 28(2):111–114, 1996.

[60] J. Jacod. La variation quadratique du brownien en présence d'erreurs d'arrondi. *Astérisque*, 236:155–161, 1996. Hommage à P. A. Meyer et J. Neveu.

[61] J. Jacod, Y. Li, P. A. Mykland, M. Podolskij, and M. Vetter. Microstructure noise in the continuous case: the pre-averaging approach. *Stochastic Process. Appl.*, 119(7):2249–2276, 2009.

[62] J. Jacod, Y. Li, and X. Zheng. Statistical properties of microstructure noise. *Econometrica*, 85(4):1133–1174, 2017.

[63] J. Jacod, M. Podolskij, and M. Vetter. Limit theorems for moving averages of discretized processes plus noise. *Ann. Statist.*, 38(3):1478–1545, 2010.

[64] J. Jacod and P. Protter. *Discretization of Processes*, volume 67 of *Stochastic Modelling and Applied Probability.* Springer, Heidelberg, 2012.

[65] J. Jacod and A. N. Shiryaev. *Limit Theorems for Stochastic Processes.* Springer, Berlin, 2nd edition, 2003.

[66] J. Jacod and V. Todorov. Efficient estimation of integrated volatility in presence of infinite variation jumps. *Ann. Statist.*, 42(3):1029–1069, 2014.

[67] J. Jacod and V. Todorov. Limit theorems for integrated local empirical characteristic exponents from noisy high-frequency data with application to volatility and jump activity estimation. *Ann. Appl. Probab.*, 28(1):511–576, 2018.

[68] A. Jakubowski. The almost sure Skorokhod representation for subsequences in nonmetric spaces. *Theory Probab. Appl.*, 42(1):167–174, 1997.

[69] O. Kallenberg. *Foundations of Modern Probability.* Springer, New York, 2nd edition, 2002.

[70] G. Kallianpur and R. Wolpert. Infinite dimensional stochastic differential equation models for spatially distributed neurons. *Appl. Math. Optim.*, 12(1):125–172, 1984.

[71] I. Kalnina and O. Linton. Estimating quadratic variation consistently in the presence of endogenous and diurnal measurement error. *J. Econometrics*, 147(1):47–59, 2008.

[72] D. Khoshnevisan. *Multiparameter Processes.* Springer, New York, 2002.

[73] A. Kohatsu-Higa and P. Tankov. Jump-adapted discretization schemes for Lévy-driven SDEs. *Stochastic Process. Appl.*, 120(1):2258–2285, 2010.

[74] N. Konno and T. Shiga. Stochastic partial differential equations for some measure-valued diffusions. *Probab. Theory Related Fields*, 79(2):201–225, 1988.

[75] T. G. Kurtz and P. E. Protter. Weak limit theorems for stochastic integrals and stochastic differential equations. *Ann. Probab.*, 19(3):1035–1070, 1991.

[76] 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.

[77] Y. Li and P. A. Mykland. Are volatility estimators robust with respect to modeling assumptions? *Bernoulli*, 13(3):601–622, 2007.

[78] Z. M. Li, R. J. A. Laeven, and M. H. Vellekoop. Dependent microstructure noise and integrated volatility estimation from high-frequency data. *J. Econometrics*, 215(2):536–558, 2020.

[79] Z. M. Li and O. Linton. A ReMeDI for microstructure noise. *Econometrica*, 2021. forthcoming.

[80] B. B. Mandelbrot. *Fractals and Scaling in Finance.* Selected Works of Benoit B. Mandelbrot. Springer-Verlag, New York, 1997. Discontinuity, concentration, risk, Selecta Volume E, With a foreword by R. E. Gomory.

[81] Y. S. Mishura. *Stochastic Calculus for Fractional Brownian Motion and Related Processes*, volume 1929 of *Lecture Notes in Mathematics.* Springer-Verlag, Berlin, 2008.

[82] D. Nualart. *The Malliavin calculus and related topics.* Probability and its Applications (New York). Springer-Verlag, Berlin, second edition, 2006.

[83] S. Peszat and J. Zabczyk. *Stochastic Partial Differential Equations with Lévy Noise.* Cambridge University Press, Cambridge, 2007.

[84] V. S. Pham and C. Chong. Volterra-type Ornstein-Uhlenbeck processes in space and time. *Stoch. Process. Appl.*, 128(9):3082–3117, 2018.

[85] M. Podolskij and M. Vetter. Estimation of volatility functionals in the simultaneous presence of microstructure noise and jumps. *Bernoulli*, 15(3):634–658, 2009.

[86] P. E. Protter. *Stochastic Integration and Differential Equations*, volume 21 of *Stochastic Modelling and Applied Probability.* Springer-Verlag, Berlin, 2005. Second edition. Version 2.1, Corrected third printing.

[87] M. Reiß. Asymptotic equivalence for inference on the volatility from noisy observations. *Ann. Statist.*, 39(2):772–802, 2011.

[88] M. Rosenbaum. Integrated volatility and round-off error. *Bernoulli*, 15(3):687–720, 2009.

[89] M. Rosenbaum. A new microstructure noise index. *Quant. Finance*, 11(6):883–899, 2011.

[90] M. Sanz-Solé and P.-A. Vuillermot. Equivalence and Hölder-Sobolev regularity of solutions for a class of non-autonomous stochastic partial differential equations. *Ann. Inst. Henri Poincaré Probab. Stat.*, 39(4):703–742, 2003.

[91] K.-I. Sato. *Lévy Processes and Infinitely Divisible Distributions.* Cambridge University Press, Cambridge, 1999.

[92] M. L. Straf. Weak convergence of stochastic processes with several parameters. In L. M. L. Cam, J. Neyman, and E. L. Scott, editors, *Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability, Volume 2: Probability Theory*, pages 187–221. Univ. California Press, Berkeley, 1972.

[93] C. Suquet. Tightness in Schauder decomposable Banach spaces. In N. N. Uraltseva, editor, *Proceedings of the St. Petersburg Mathematical Society Volume V*, pages 201–224. American Mathematical Society, Providence, RI, 1999.

[94] H. Tanaka and A. Kohatsu-Higa. An operator approach for Markov chain weak approximations with an application to infinite activity Lévy driven SDEs. *Ann. Appl. Probab.*, 19(3):1026–1062, 2009.

[95] H. Taniai, T. Usami, N. Suto, and M. Taniguchi. Asymptotics of realized volatility with non-Gaussian ARCH(∞) microstructure noise. *J. Financ. Econom.*, 10(4):617–636, 2012.

[96] H. van Zanten. When is a linear combination of independent fBm's equivalent to a single fBm? *Stochastic Process. Appl.*, 117(1):57–70, 2007.

[97] J. B. Walsh. Convergence and regularity of multiparameter strong martingales. *Probab. Theory Relat. Fields*, 46(2):177–192, 1979.

[98] J. B. Walsh. A stochastic model of neural response. *Adv. Appl. Probab.*, 13(2):231–281, 1981.

[99] J. B. Walsh. An introduction to stochastic partial differential equations. In P. Hennequin, editor, *École d'Été de Probabilités de Saint Flour XIV - 1984*, pages 265–439. Springer, Berlin, 1986.

[100] L. Zhang, P. A. Mykland, and Y. Aït-Sahalia. A tale of two time scales: determining integrated volatility with noisy high-frequency data. *J. Amer. Statist. Assoc.*, 100(472):1394–1411, 2005.