



# Martingale-driven integrals and singular SPDEs

P. Grazieschi<sup>1</sup> · K. Matetski<sup>2</sup> · H. Weber<sup>3</sup>

Received: 20 July 2023 / Revised: 1 July 2024 / Accepted: 27 July 2024 /

Published online: 12 August 2024

© The Author(s), under exclusive licence to Springer-Verlag GmbH Germany, part of Springer Nature 2024

## Abstract

We consider multiple stochastic integrals with respect to càdlàg martingales, which approximate a cylindrical Wiener process. We define a chaos expansion, analogous to the case of multiple Wiener stochastic integrals, for these integrals and use it to show moment bounds. Key tools include an iteration of the Burkholder–Davis–Gundy inequality and a multi-scale decomposition similar to the one developed in Hairer and Quastel (Forum Math Pi 6:e3, 2018). Our method can be combined with the recently developed discretisation framework for regularity structures (Hairer and Matetski in Ann Probab 46(3):1651–1709, 2018, Erhard and Hairer in Ann Inst Henri Poincaré Probab Stat 55(4):2209–2248, 2019) to prove convergence of interacting particle systems to singular stochastic PDEs. A companion article (Grazieschiet al. in The dynamical Ising–Kac model in 3D converges to  $\Phi_3^4$ , 2023. arXiv:2303.10242) applies the results of this paper to prove convergence of a rescaled Glauber dynamics for the three-dimensional Ising–Kac model near criticality to the  $\Phi_3^4$  dynamics on a torus.

**Mathematics Subject Classification** 60H15 · 60G44

## Contents

1	Introduction	1064
2	Integrals with respect to càdlàg martingales	1067
3	Moment bounds for iterated integrals	1076
4	Kernels given by generalised convolutions	1088
5	Application to a discrete martingale model	1107
	References	1119

✉ K. Matetski  
matetski@msu.edu

P. Grazieschi  
p.grazieschi@bath.ac.uk

H. Weber  
hendrik.weber@uni-muenster.de

<sup>1</sup> University of Bath, Bath, UK

<sup>2</sup> Michigan State University, East Lansing, USA

<sup>3</sup> University of Münster, Münster, Germany

## 1 Introduction

We consider a class of càdlàg martingales which approximate a cylindrical Wiener process over a  $d$ -dimensional spatial domain, i.e. integrated-in-time space-time white noise. We develop a theory of iterated integrals with respect to these martingales and derive moment bounds.

Our results serve as a technical tool for proving the convergence of Interacting Particle Systems (IPSSs) to solutions of non-linear stochastic partial differential equations (SPDEs). The limiting SPDEs are usually of the form

$$\mathcal{L}u = F(u, \nabla u) + \sigma(u)\xi, \quad (1.1)$$

where  $\mathcal{L}$  is a linear parabolic operator (e.g.  $\mathcal{L} = \partial_t - \Delta$ ),  $\xi$  is an irregular random noise (e.g. a Gaussian white noise) and  $F$  and  $\sigma$  are local non-linearities. There are by now a number of convergence results of this type. These include 1 + 1-dimensional surface growth models rescaling to the KPZ equation

$$\partial_t h - \partial_x^2 h = -(\partial_x h)^2 + \xi, \quad (1.2)$$

e.g. [3, 8, 9, 12, 19], long range (Kac) spin models rescaling to  $\varphi^{2n}$  dynamics

$$\partial_t \varphi - \Delta \varphi = -\varphi^{2n-1} + \xi, \quad (1.3)$$

in one [5, 16] and two dimensions [20, 32, 40, 42] as well as diffusions in random environment rescaling to the parabolic Anderson model / multiplicative stochastic heat equation

$$\partial_t u - \Delta u = u\xi, \quad (1.4)$$

[14, 39]. Ultimately, the motivating goal of the theory developed in this article is to show the convergence of the Ising–Kac model to the  $\varphi^4$  dynamics in three dimensions, and this is accomplished in our companion article [21].

A common feature of all of these limiting results is that particle systems are simultaneously rescaled (i.e. observed on large scales) while a certain parameter is changed. The specific nature of this parameter depends on the model under consideration; examples are the strength of the *weak* asymmetry in exclusion processes approximating the KPZ equation [3], or the range of the interaction in Kac-models [20]. The typical strategy is to tune down the effect of the “non-linearity” as one moves to larger scales.

This procedure is necessary to obtain convergence to one of the SPDEs (1.2), (1.3), (1.4) and reflects the fact that the SPDEs are themselves not scale-invariant. The fact that a relatively small class of SPDEs arises as scaling limit of this type for a relatively large number of particle systems sharing just a few key characteristics is sometimes referred to as weak universality.

A key technical challenge in deriving such scaling results is the low regularity of the solutions of the limiting Eqs. (1.2), (1.3), (1.4): the noise term  $\xi$  is typically very irregular, leading to irregular solutions which in turn lead to difficulties in dealing

with the non-linearities. This problem does not appear in more common Gaussian fluctuation limits [35]—while the solutions of the limiting equations there are typically also irregular, this is less problematic due to the absence of a non-linear term. Good theories for non-linear SPDEs and their renormalisation have only been developed over the last years, including Hairer’s theory of regularity structures [26], the theory of paracontrolled distributions put forward by Gubinelli, Imkeller and Perkowski [17] and more recently theories of weak solutions for specific equations, in particular the KPZ equation [18, 19, 23, 24].

The theory of regularity structures and the theory of paracontrolled distributions both build on a two-step approach: first, the construction of approximate solutions building a local expansion (the *model* in the jargon of [26]) which relies on probabilistic tools, in particular Gaussian analysis and explicit calculations of covariance functions, and second, analytic techniques (in particular regularity estimates and commutator estimates) for dealing with the remainder. The weak solution theories developed in [18, 19, 23, 24] use a very different approach and make explicit use of the invariant Gaussian measure to give a direct characterisation of the generator of the dynamics.

In principle, both approaches can be used to study scaling limits. In situations, where a simple invariant measure for an interacting particle system is given, the weak solution approach has proved highly efficient, see e.g. [2, 19, 25, 33]. The approach which consists of mimicking the theory of regularity structure / paracontrolled distributions has also been implemented in a few examples, in particular [14, 22, 39, 40]. Still, implementing this programme for “interesting” limiting equations remains a challenging enterprise: for the second, deterministic, step of the analysis a systematic theory has been developed in [13, 29], but the first probabilistic part remains challenging, because the number of terms in this perturbative expansion (the “trees”) can become prohibitively large when looking at interesting equations. For the continuum there is by now a very systematic treatment for the trees (see [6, 30, 31, 37]). The aim of this paper is to develop a—at least somewhat—systematic approach to bound these trees for approximations of white noise. A particular focus is on the jump martingales that typically arise in the analysis of IPSs.

On a technical level: the noise approximations we deal with are of bounded variation, but discontinuous because of the jumps. Therefore, the non-linear functionals that make up the model can rigorously be written in terms of integrals with respect to product measure in the underlying noise. We then decompose these integrals according to “diagonals” or “contractions”. This is in the spirit of the Wiener chaos decomposition, however many more terms than in the Gaussian case arise, as in the latter only diagonals where precisely two coordinates coincide, make a non-vanishing contribution. In our noise approximations, many more “diagonals” appear, and we aim to show that their impact vanishes as space-time white noise is approached. Our main technical tool is an iteration of a Burkholder–Davis–Gundy (BDG) type inequality. For our purpose, the most convenient form is in terms of the predictable quadratic variation with an error term that depends on the size of jumps, as was used previously in [40, Lemma 4.1]. The advantage is that under our assumptions (which are motivated by the analysis of the Ising–Kac model [21]), an explicit and optimal bound on the predictable quadratic variation are available. The error term does not matter too much in the Ising–Kac application, because the size of individual jumps is suppressed by the

smoothing from the Kac-potential. Another key assumption we need to make, is that the magnitude of the jumps of the martingales is fixed by a deterministic constant. This allows to rewrite contractions of an odd number of variables in terms of a martingale and ultimately permits to prove that in the Kac-Ising application these contractions vanish in the limit, even though they are integrated against a very singular kernel.

## 1.1 Structure of the article

In Sect. 2 we define multiple stochastic integrals with respect to càdlàg square integrable martingales. In Sect. 3.1 we derive moment bounds on stochastic integrals with respect to only one variable, while moment bounds on multiple integrals are obtained in Sect. 3.2. Section 3.2.1 is devoted to renormalised stochastic integrals and their moment bounds. In Sect. 4 we analyse stochastic integrals with kernels given by generalised convolutions, which are typical objects in the theory of regularity structures. As an example, we apply the result in Sect. 5 to a discrete approximation of the  $\Phi_3^4$  equation.

## 1.2 Notation

We use the standard notation  $\mathbf{N} = \{1, 2, 3, \dots\}$  for the set of natural numbers,  $\mathbf{N}_0$  for  $\mathbf{N} \cup \{0\}$  and the set  $\mathbf{R}_+ := [0, \infty)$  for the time variables. For  $n \in \mathbf{N}$  we define  $[\![n]\!] := \{1, \dots, n\}$ . We use  $\mathbb{1}_A$  for the indicator function of the set  $A$ .

For  $\Lambda_0$  being either  $(\mathbf{R}/\mathbf{Z})^d$  or  $\mathbf{R}^d$  we use the standard notation  $\mathcal{D}'(\Lambda_0)$  for the space of distributions on  $\Lambda_0$ . For  $n \geq 0$ , the space  $\mathcal{C}^n(\Lambda_0)$  contains all  $n$ -times continuously differentiable functions on  $\Lambda_0$ , and we write  $\mathcal{C}(\Lambda_0)$  for this space when  $n = 0$ . The Skorokhod space of càdlàg functions on  $[0, T]$  with values in  $\mathcal{D}'(\Lambda_0)$  is denoted by  $D([0, T], \mathcal{D}'(\Lambda_0))$ .

Given a random variable  $X$  and some  $p \geq 1$ , we use the following shorthand notation for the stochastic  $L^p$  norm

$$\mathbf{E}_p X := \mathbf{E}[|X|^p]^{1/p}. \quad (1.5)$$

In estimates we often use “ $\lesssim$ ”, which means that the bound “ $\leq$ ” holds up to a constant which is independent of the quantities relevant in our statements, which will be always clear from the context. If we want to indicate dependence of the proportionality constant on some parameters  $\alpha, \beta, \dots$ , we write “ $\lesssim_{\alpha, \beta, \dots}$ ”.

Finally, let  $\Lambda_\varepsilon = (\varepsilon \mathbf{Z}/\mathbf{Z})^d$  be a discrete torus with mesh size  $\varepsilon > 0$ . For  $T > 0$ ,  $p \geq 1$  and for a function  $F : [0, T] \times \Lambda_\varepsilon \rightarrow \mathbf{R}$ , we define

$$\|F\|_{L_\varepsilon^p} := \left( \varepsilon^d \sum_{x \in \Lambda_\varepsilon} \int_0^T |F(r, x)|^p dr \right)^{\frac{1}{p}}, \quad (1.6)$$

that is, we take the  $L^p$  norm in time and the  $l^p$  norm in space with a weight  $\varepsilon^d$  on the points of the lattice. This and several other norms in the article depend on the parameter  $T$ , but we omit this dependence from our notation.

## 2 Integrals with respect to càdlàg martingales

### 2.1 Properties of càdlàg martingales

Following [34, Ch. I.4], we recall some properties of martingales which are used in the article. Let  $(M_t)_{t \geq 0}$  and  $(N_t)_{t \geq 0}$  be two càdlàg square-integrable martingales on the same filtered probability space. Their *predictable quadratic covariation*  $\langle M, N \rangle_t$  is the unique adapted process with bounded total variation, such that  $M_t N_t - \langle M, N \rangle_t$  is a martingale. The *quadratic covariation*  $[M, N]_t$  is defined by

$$[M, N]_t := M_t N_t - M_0 N_0 - \int_0^t M_s \mathrm{d}N_s - \int_0^t N_s \mathrm{d}M_s, \quad (2.1)$$

where  $M_{s-} := \lim_{r \uparrow s} M_r$  is the left limit of  $M$  at time  $s$ . Another way to define these quadratic covariations is the following: if  $0 = t_0 \leq \dots \leq t_n = t$  is a partition with diameter  $\max_i (t_{i+1} - t_i)$  tending to zero as  $n \rightarrow \infty$ , then  $[M, N]_t$  is equal to the limit in probability of the sums  $\sum_{i=0}^{n-1} (M_{t_{i+1}} - M_{t_i})(N_{t_{i+1}} - N_{t_i})$  as  $n \rightarrow \infty$  (see [34, Thm. I.4.47]), and  $\langle M, N \rangle_t$  is the probability limit of the sums  $\sum_{i=0}^{n-1} \mathbf{E}[(M_{t_{i+1}} - M_{t_i})(N_{t_{i+1}} - N_{t_i}) | \mathcal{F}_{t_i}]$ , where  $(\mathcal{F}_t)_{t \geq 0}$  is the underlying filtration [34, Prop. I.4.50]. The difference of the two bracket processes  $[M, N]_t - \langle M, N \rangle_t$  is always a càdlàg martingale [34, Prop. I.4.50]. In the case  $M = N$ , it will be convenient to use the shorthands  $[M]_t = [M, M]_t$  and  $\langle M \rangle_t = \langle M, M \rangle_t$ .

We will use the Burkholder–Davis–Gundy inequality in the following form, which is obtained by approximating  $M$  by discrete-time martingales and applying the discrete-time Burkholder–Davis–Gundy inequality [28].

**Proposition 2.1** *Let  $(M_t)_{t \in [0, T]}$  be a càdlàg square integrable martingale. Then, for any  $p \geq 1$  there exists a constant  $C > 0$  depending on  $p$  such that*

$$\mathbf{E} \left[ \sup_{t \in [0, T]} |M_t|^p \right]^{\frac{1}{p}} \leq C \left( \mathbf{E} \left[ \langle M \rangle_t^{p/2} \right]^{\frac{1}{p}} + \mathbf{E} \left[ \sup_{t \in [0, T]} |\Delta_t M|^p \right]^{\frac{1}{p}} \right), \quad (2.2)$$

where  $\Delta_t M := M_t - M_{t-}$  is a jump at time  $t$ .

### 2.2 Assumptions on martingales

Let  $d \in \mathbf{N}$  and let  $\Lambda_0$  be a  $d$ -dimensional torus  $(\mathbf{R}/\mathbf{Z})^d$ . For  $\varepsilon > 0$ , let  $\Lambda_\varepsilon$  be a discretisation of  $\Lambda_0$  with mesh size  $\varepsilon$ , i.e.  $\Lambda_\varepsilon$  is a  $d$ -dimensional discrete torus  $(\varepsilon\mathbf{Z}/\mathbf{Z})^d$  (in this case we need  $\varepsilon^{-1}$  to be integer). The moment bounds for stochastic integrals, which we prove in the following sections, depend on the Lebesgue measure

of the domain  $\Lambda_0$ , which is bounded. Let  $\mathcal{D}_\varepsilon := \mathbf{R}_+ \times \Lambda_\varepsilon$  be a discretised space-time domain, and let  $\mathcal{D}_{\varepsilon,t} := [0, t) \times \Lambda_\varepsilon$  be the space-time domain with time horizon  $t > 0$ .

For a function  $u_\varepsilon$  on the domain  $\mathcal{D}_\varepsilon$ , we introduce its natural extensions to the space of distributions

$$(\iota_\varepsilon u_\varepsilon)(\varphi) := \sum_{x \in \Lambda_\varepsilon} \varepsilon^d \int_{\mathbf{R}_+} u_\varepsilon(t, x) \varphi(t, x) dt, \quad (\iota_\varepsilon u_\varepsilon)(t, \psi) := \sum_{x \in \Lambda_\varepsilon} \varepsilon^d u_\varepsilon(t, x) \psi(x), \quad (2.3)$$

where  $\varphi : \mathbf{R}_+ \times \Lambda_0 \rightarrow \mathbf{R}$  and  $\psi : \Lambda_0 \rightarrow \mathbf{R}$  are smooth compactly supported functions.

Let  $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbf{P})$  be a filtered probability space, which satisfies the “usual conditions” (i.e. completeness and right-continuity [34, Def. I.1.3]). We then introduce a family of càdlàg martingales  $(\mathbb{M}_\varepsilon(t, x))_{t \geq 0}$ , indexed by points  $x \in \Lambda_\varepsilon$ . Let  $\mathbb{M}_\varepsilon(t-, x) = \lim_{\delta \rightarrow 0+} \mathbb{M}_\varepsilon(t - \delta, x)$  be the left-limit of  $\mathbb{M}_\varepsilon(x)$  at time  $t$  and let  $\Delta_t \mathbb{M}_\varepsilon(x) := \mathbb{M}_\varepsilon(t, x) - \mathbb{M}_\varepsilon(t-, x)$  denote the jump at time  $t$ . We make the following assumption on these martingales.

**Assumption 1** For  $\varepsilon > 0$ , we assume that  $(\mathbb{M}_\varepsilon(t, x))_{t \geq 0}$  are càdlàg square-integrable martingales with the following properties.

1. The predictable quadratic covariation  $\langle \mathbb{M}_\varepsilon(x), \mathbb{M}_\varepsilon(y) \rangle_t$  vanishes whenever  $x \neq y$ , and

$$\langle \mathbb{M}_\varepsilon(x) \rangle_t = \varepsilon^{-d} \int_0^t \mathbf{C}_\varepsilon(s, x) ds, \quad (2.4)$$

where  $(s, x) \mapsto \mathbf{C}_\varepsilon(s, x)$  is a progressively measurable stochastic process satisfying  $|\mathbf{C}_\varepsilon(s, x)| \lesssim 1$  a.s. uniformly in  $s$  and  $x$ . The proportionality constant in this bound is non-random.

2. Two martingales almost surely never jump simultaneously, i.e. for any  $T > 0$

$$\mathbf{P}(\Delta_t \mathbb{M}_\varepsilon(x) \Delta_t \mathbb{M}_\varepsilon(y) = 0, \forall x \neq y, \forall t \in [0, T]) = 1.$$

3. There exist  $\mathbf{k} > -\frac{d}{2}$  and a non-random value  $c > 0$  such that for all  $x \in \Lambda_\varepsilon$  and  $t \geq 0$  the following holds: if  $\Delta_t \mathbb{M}_\varepsilon(x) \neq 0$ , then  $|\Delta_t \mathbb{M}_\varepsilon(x)| = c \varepsilon^{\mathbf{k}}$  a.s.
4. The martingale  $\mathbb{M}_\varepsilon(s, x)$  follows a dynamics which can be expressed in the form

$$\mathbb{M}_\varepsilon(t, x) = J_\varepsilon(t, x) - \varepsilon^{-\mathbf{k}-d} \int_0^t C_\varepsilon(s, x) ds, \quad (2.5)$$

where  $t \mapsto J_\varepsilon(t, x)$  is a pure jump process (i.e.  $J_\varepsilon(t, x) = \sum_{0 \leq s \leq t} \Delta_s \mathbb{M}_\varepsilon(x)$ ) and where  $t \mapsto C_\varepsilon(t, x)$  is a progressively measurable process such that  $|C_\varepsilon(t, x)| \lesssim 1$  a.s. uniformly in  $x$  and  $t$ . The proportionality constant in the last estimate is non-random.

**Remark 2.2** Assumption 1(1.) implies that in the case  $\mathbf{C}_\varepsilon(s, x) = 1$  the quadratic variation of the martingales approximates the quadratic variation of a cylindrical Wiener process, see also the following Lemma 2.6. Assumption 1(2.) is satisfied in many applications, e.g. when jumps are sub-sampled from independent Poisson processes. Assumption 1(3.) implies that the size of an individual jump is smaller than the size of  $\mathbb{M}_\varepsilon(t, x)$  for bounded  $t$ , which is of order  $\varepsilon^{-d/2}$  by Assumption 1(1.). We show in Lemma 2.4 below that Assumptions 1(1.) and 1(3.) combined imply that jumps happen with frequency  $\varepsilon^{d+2k}$ .

We will use the following martingales (see Sect. 2.1)

$$\bar{\mathbb{M}}_\varepsilon(t, x) := \varepsilon^{-k} (\mathbb{M}_\varepsilon(x))_t - \langle \mathbb{M}_\varepsilon(x) \rangle_t. \quad (2.6)$$

The multiplier  $\varepsilon^{-k}$  in (2.6) is chosen to have the following.

**Lemma 2.3** *The martingales  $\bar{\mathbb{M}}_\varepsilon(t, x)$  satisfy Assumption 1 with the same value of  $k$  (but with the constant  $c^2$  in place of  $c$ ), with  $\mathbf{C}_\varepsilon$  replaced by  $c^2 \mathbf{C}_\varepsilon$  in (2.4) and with  $C_\varepsilon$  replaced by  $\mathbf{C}_\varepsilon$  in (2.5).*

**Proof** Assumptions 1(1.) and (4.) yield  $[\mathbb{M}_\varepsilon(x)]_t = \sum_{0 \leq s \leq t} (\Delta_s \mathbb{M}_\varepsilon(x))^2$  and

$$\bar{\mathbb{M}}_\varepsilon(t, x) = \varepsilon^{-k} \sum_{0 \leq s \leq t} (\Delta_s \mathbb{M}_\varepsilon(x))^2 - \varepsilon^{-k-d} \int_0^t \mathbf{C}_\varepsilon(s, x) ds.$$

This identity gives Assumptions 1(4.) for  $\bar{\mathbb{M}}_\varepsilon$  with  $C_\varepsilon$  replaced by  $\mathbf{C}_\varepsilon$ . The second term is a function of finite variation with respect to the variable  $t$ , and it does not contribute to the predictable quadratic covariation. Then, using Assumptions 1(3.), we get

$$[\bar{\mathbb{M}}_\varepsilon(x)]_t = \varepsilon^{-2k} \sum_{0 \leq s \leq t} (\Delta_s \mathbb{M}_\varepsilon(x))^4 = c^2 \sum_{0 \leq s \leq t} (\Delta_s \mathbb{M}_\varepsilon(x))^2 = c^2 [\mathbb{M}_\varepsilon(x)]_t$$

almost surely. Similarly, we get  $[\bar{\mathbb{M}}_\varepsilon(x), \bar{\mathbb{M}}_\varepsilon(y)]_t = 0$  almost surely for  $x \neq y$ . Since  $[\mathbb{M}_\varepsilon(x)]_t - \langle \mathbb{M}_\varepsilon(x) \rangle_t$  is a martingale, predictable quadratic variation of  $\bar{\mathbb{M}}_\varepsilon(x)$  has to be  $c^2 \langle \mathbb{M}_\varepsilon(x) \rangle_t$ , which yields Assumptions 1(1.) for  $\bar{\mathbb{M}}_\varepsilon$  with  $\mathbf{C}_\varepsilon$  replaced by  $c^2 \mathbf{C}_\varepsilon$ .

The rest of Assumptions 1 for  $\bar{\mathbb{M}}_\varepsilon$  follows readily from these identities and properties of the martingales  $\mathbb{M}_\varepsilon$ .  $\square$

For  $x \in \Lambda_\varepsilon$  and for a bounded set  $A \subseteq \mathbf{R}_+$ , we define

$$\mathbf{n}_A^\varepsilon(x) := \#\{t \in A : \Delta_t \mathbb{M}_\varepsilon(x) \neq 0\} \quad (2.7)$$

to be the number of jumps of the martingale  $\mathbb{M}_\varepsilon(x)$  in  $A$ . We are going to show that Assumption 1 implies moment bounds for the number of jumps.

**Lemma 2.4** *For any  $[a, b] \subseteq \mathbf{R}_+$  and any  $p \geq 1$  the number of jumps satisfies*

$$\sup_{\varepsilon \in (0, 1]} \sup_{x \in \Lambda_\varepsilon} \mathbf{E}_p |\mathbf{n}_{[a, b]}^\varepsilon(x)| < \infty, \quad (2.8)$$

locally uniformly in  $a, b$ .

**Proof** Assumptions 1(3.) and (4.) yield  $[\mathbb{M}_\varepsilon(x)]_t = \sum_{0 \leq s \leq t} (\Delta_s \mathbb{M}_\varepsilon(x))^2 = c^2 \varepsilon^{2k} \mathbb{m}_{[0,t]}^\varepsilon(x)$ . Then for any  $p \geq 1$  we have  $\mathbf{E}_p |\mathbb{m}_{[0,t]}^\varepsilon(x)| = c^{-2} \varepsilon^{-2k} \mathbf{E}_p [\mathbb{M}_\varepsilon(x)]_t$ . Using the martingales (2.6) and applying Minkowski's inequality, we get furthermore

$$\mathbf{E}_p |\mathbb{m}_{[0,t]}^\varepsilon(x)| \leq c^{-2} \varepsilon^{-2k} \mathbf{E}_p \langle \mathbb{M}_\varepsilon(x) \rangle_t + c^{-2} \varepsilon^{-2k} \mathbf{E}_p |\overline{\mathbb{M}}_\varepsilon(t, x)|.$$

The first term is bounded using Assumptions 1(1.) as  $\varepsilon^{-2k} \mathbf{E}_p \langle \mathbb{M}_\varepsilon(x) \rangle_t \lesssim \varepsilon^{-d-2k} t$ , while to bound the second term we apply the Burkholder–Davis–Gundy inequality (2.2):

$$\varepsilon^{-k} \mathbf{E}_p |\overline{\mathbb{M}}_\varepsilon(t, x)| \lesssim \varepsilon^{-k} \mathbf{E}_p \langle \overline{\mathbb{M}}_\varepsilon(x) \rangle_t^{1/2} + \varepsilon^{-k} \mathbf{E}_p \sup_{s \in [0,t]} |\Delta_s \overline{\mathbb{M}}_\varepsilon(x)|,$$

where the proportionality constant depends only on  $p$ . Lemma 2.3 allows to bound the preceding expression by a constant multiple of  $\varepsilon^{-k-d/2} t^{1/2} + 1$ . Hence, we have

$$\mathbf{E}_p |\mathbb{m}_{[0,t]}^\varepsilon(x)| \lesssim \varepsilon^{-d-2k} t + \varepsilon^{-k-d/2} t^{1/2} + 1.$$

Since the proportionality constant in this bound is independent of  $t$ , we can replace  $t$  by  $\varepsilon^{d+2k} t$  to get  $\mathbf{E}_p |\mathbb{m}_{[0,\varepsilon^{d+2k} t]}^\varepsilon(x)| \lesssim t + 1$  from which the required bound follows.  $\square$

**Lemma 2.5** *Let  $\|\cdot\|_{TV([0,T])}$  be the total variation norm on the interval  $[0, T]$ . Then*

$$\sup_{\varepsilon \in (0,1]} \sup_{x \in \Lambda_\varepsilon} \varepsilon^{k+d} \mathbf{E}_p \|\mathbb{M}_\varepsilon(x)\|_{TV([0,T])} < \infty$$

for any  $T > 0$  and  $p \geq 1$ .

**Proof** Assumption 1(4.) yields

$$\begin{aligned} \|\mathbb{M}_\varepsilon(x)\|_{TV([0,T])} &\leq \sum_{0 \leq s \leq T} |\Delta_s \mathbb{M}_\varepsilon(x)| + \varepsilon^{-k-d} \int_0^T |\mathbb{C}_\varepsilon(s, x)| ds \\ &\lesssim \varepsilon^k \mathbb{m}_{[0,T]}^\varepsilon(x) + T \varepsilon^{-k-d} \end{aligned}$$

a.s. uniformly in  $x$ , where we used the properties  $|\Delta_s \mathbb{M}_\varepsilon(x)| \lesssim \varepsilon^k$  and  $|\mathbb{C}_\varepsilon(s, x)| \lesssim 1$ . Then the required bound follows from Lemma 2.4.  $\square$

The next result shows that martingales satisfying Assumption 1 weakly converge to a cylindrical Wiener process [11].

**Lemma 2.6** *Let martingales  $(\mathbb{M}_\varepsilon(t, x))_{t \geq 0}$ , with  $x \in \Lambda_\varepsilon$  and either  $\Lambda_\varepsilon = (\varepsilon \mathbf{Z}/\mathbf{Z})^d$  or  $\Lambda_\varepsilon = \varepsilon \mathbf{Z}^d$ , satisfy Assumption 1 (except possibly Assumption 1(4.)) and let  $\mathbb{M}_\varepsilon(0, x) =$*

0. For every continuous, compactly supported function  $\varphi : \Lambda_0 \rightarrow \mathbf{R}$ , for every fixed  $t > 0$  and some constant  $\sigma > 0$ , let the following limit hold in distribution

$$\lim_{\varepsilon \rightarrow 0} \int_0^t (\iota_\varepsilon \mathbf{C}_\varepsilon)(s, \varphi) ds = \sigma t \int_{\Lambda_0} \varphi(x) dx. \quad (2.9)$$

Then the martingales  $(\mathbb{M}_\varepsilon(t, x))_{t \geq 0}$  weakly converge in the Skorokhod topology  $D(\mathbf{R}_+, \mathcal{D}'(\Lambda_0))$  to a cylindrical Wiener process on  $L^2(\Lambda_0)$  with variance  $\sigma$ .

**Proof** Let us take a continuous and compactly supported function  $\varphi$  and consider the martingales  $t \mapsto (\iota_\varepsilon \mathbb{M}_\varepsilon)(t, \varphi)$ . To prove tightness of the laws of these martingales, we will use a version of the Aldous' criterion [4, Corollary 16.11] (see [1] for the original result).

By Assumption 1(2.),(3.), for any  $T > 0$ , the jumps of this martingale can be a.s. bounded by  $\sup_{0 \leq t \leq T} |\Delta_t(\iota_\varepsilon \mathbb{M}_\varepsilon)(\varphi)| \lesssim \varepsilon^{d+\mathbf{k}} \|\varphi\|_{L^\infty}$ , which vanishes as  $\varepsilon \rightarrow 0$ . Furthermore, for fixed  $t \geq 0$  the process  $\delta \mapsto \mathbb{M}_\varepsilon(t + \delta, x) - \mathbb{M}_\varepsilon(t, x)$  is a martingale with respect to the filtration  $(\mathcal{F}_{t+\delta})_{\delta \geq 0}$  with the predictable quadratic covariation  $\varepsilon^{-d} \int_t^{t+\delta} \mathbf{C}_\varepsilon(s, x) ds$ . Hence, for any stopping time  $\tau \in [0, T]$  and for any  $\delta \in (0, 1]$  we apply the Burkholder–Davis–Gundy inequality (2.2) to get

$$\begin{aligned} & \mathbf{E}|(\iota_\varepsilon \mathbb{M}_\varepsilon)(\tau + \delta, \varphi) - (\iota_\varepsilon \mathbb{M}_\varepsilon)(\tau, \varphi)| \\ & \leq \mathbf{E} \left[ \sup_{0 \leq t \leq T} \mathbf{E} \left[ |(\iota_\varepsilon \mathbb{M}_\varepsilon)(t + \delta, \varphi) - (\iota_\varepsilon \mathbb{M}_\varepsilon)(t, \varphi)| \middle| \mathcal{F}_t \right] \right] \\ & \lesssim \mathbf{E} \left[ \sup_{0 \leq t \leq T} \mathbf{E} \left[ \int_t^{t+\delta} (\iota_\varepsilon \mathbf{C}_\varepsilon)(s, \varphi^2) ds \middle| \mathcal{F}_t \right]^{\frac{1}{2}} \right] + \mathbf{E} \left[ \sup_{0 \leq t \leq T+\delta} \mathbf{E} \left[ |\Delta_t(\iota_\varepsilon \mathbb{M}_\varepsilon)(\varphi)| \middle| \mathcal{F}_t \right] \right]. \end{aligned}$$

Using Assumption 1(1.), the first term is bounded by a constant proportional to  $\delta^{1/2}$ , while the second term is bounded by a constant proportional to  $\varepsilon^{d+\mathbf{k}}$ . Hence, for any  $\alpha > 0$  the Markov inequality yields

$$\mathbf{P}(|(\iota_\varepsilon \mathbb{M}_\varepsilon)(\tau + \delta, \varphi) - (\iota_\varepsilon \mathbb{M}_\varepsilon)(\tau, \varphi)| \geq \alpha) \leq C \frac{1}{\alpha} (\delta^{1/2} + \varepsilon^{d+\mathbf{k}}),$$

and the assumptions of [4, Corollary 16.11] are satisfied. This gives tightness of the stochastic processes  $t \mapsto (\iota_\varepsilon \mathbb{M}_\varepsilon)(t, \varphi)$  in  $D([0, T], \mathbf{R})$ , and moreover every limiting point is in  $\mathcal{C}([0, T], \mathbf{R})$ . From [34, Corollary IX.1.19] we conclude that every limiting point is a martingale. Finally, [38] yields tightness of  $t \mapsto \iota_\varepsilon \mathbb{M}_\varepsilon(t)$  in  $D([0, T], \mathcal{D}'(\Lambda_0))$ . Convergence (2.9) and Assumption 1(1.) imply the limit in distribution

$$\lim_{\varepsilon \rightarrow 0} \langle (\iota_\varepsilon \mathbb{M}_\varepsilon)(\varphi) \rangle_t = \sigma t \|\varphi\|_{L^2}^2,$$

combining which with the Lévy characterization theorem we conclude that the limit of  $\mathbb{M}_\varepsilon$  in the Skorokhod topology  $D([0, T], \mathcal{D}'(\Lambda_0))$  is a cylindrical Wiener process with variance  $\sigma$ .  $\square$

## 2.3 Iterated integrals with respect to martingales

Let  $\mathbf{M}_\varepsilon$  be the random measure (recall that the paths of  $\mathbb{M}_\varepsilon$  are almost surely of bounded variation) over  $\mathcal{D}_\varepsilon$  such that, for any  $F : \mathcal{D}_\varepsilon \rightarrow \mathbf{R}$  which is continuous in the time variable,

$$\int_{\mathcal{D}_\varepsilon} F(z) d\mathbf{M}_\varepsilon(z) = \sum_{x \in \Lambda_\varepsilon} \varepsilon^d \int_{s=0}^{\infty} F(s, x) d\mathbb{M}_\varepsilon(s, x). \quad (2.10)$$

Atoms of  $\mathbf{M}_\varepsilon$  correspond to jumps of the martingales; in fact, by Assumption 1, the magnitude of the jumps of the martingales is deterministic in absolute value and equal to  $c\varepsilon^k$ ; as such, given that different martingales never jump simultaneously, the absolute value of atoms of  $\mathbf{M}_\varepsilon$  is always equal to  $c\varepsilon^{d+k}$  and this quantity—see again Assumption 1(3)—goes to zero as  $\varepsilon \rightarrow 0$ .

Let  $n \in \mathbf{N}$  and let  $\mathbf{M}_\varepsilon^n$  be the product measure on  $\mathcal{D}_\varepsilon^n$ . We want to analyse integrals of the form

$$\int_{\mathcal{D}_\varepsilon^n} F d\mathbf{M}_\varepsilon^n := \int_{\mathcal{D}_\varepsilon^n} F(z_1, \dots, z_n) \prod_{i=1}^n d\mathbf{M}_\varepsilon(z_i). \quad (2.11)$$

Here and throughout this section  $F : \mathcal{D}_\varepsilon^n \rightarrow \mathbf{R}$  is a function of  $n$  space-time variables, which is continuous in all time variables.

### 2.3.1 Contractions and orderings

The following is motivated by the analysis of  $n$ -fold iterated integrals against space-time white noise, which are conventionally defined as limits of Riemann sums that cut out diagonals (see e.g. [41, Section 1.1.2], [36, Section 9], or [10, Appendix A]): we call a *contraction* on  $\llbracket n \rrbracket$  any equivalence relation on  $\llbracket n \rrbracket$ , and its equivalent classes are called *components*. We use the symbol  $\mathfrak{C}(n)$  to denote the set of contractions and the symbol  $\mathfrak{C}_m(n)$  to denote contractions on  $\llbracket n \rrbracket$  with  $m$  components. For  $a, b \in \llbracket n \rrbracket$  and  $\gamma \in \mathfrak{C}(n)$ , we use the notation  $a \sim_\gamma b$  to indicate that  $a$  and  $b$  belong to the same component, and we denote by  $[a]_\gamma$  the component containing  $a$ .

For  $\gamma \in \mathfrak{C}(n)$  we define

$$\tilde{\mathcal{D}}_\gamma := \left\{ (z_1, \dots, z_n) \in \mathcal{D}_\varepsilon^n : z_i = z_j \text{ if and only if } i \sim_\gamma j \right\}.$$

The sets  $\tilde{\mathcal{D}}_\gamma$  form a partition of  $\mathcal{D}_\varepsilon^n = \bigsqcup_{\gamma \in \mathfrak{C}(n)} \tilde{\mathcal{D}}_\gamma$ , so that we can write the integral in (2.11) as

$$\int_{\mathcal{D}_\varepsilon^n} F d\mathbf{M}_\varepsilon^n = \sum_{\gamma \in \mathfrak{C}(n)} \int_{\tilde{\mathcal{D}}_\gamma} F d\mathbf{M}_\varepsilon^n. \quad (2.12)$$

The next lemma shows that under the measure  $\mathbf{M}_\varepsilon^n$  we can disregard all the points in  $\mathcal{D}_\varepsilon^n$  which have different space components but the same time components. To this

end, for  $i \neq j \in \llbracket n \rrbracket$  and for  $T > 0$  we define

$$\mathcal{C}_{i,j}(T) := \{(z_1, \dots, z_n) \in \mathcal{D}_\varepsilon^n : z_i = (s_i, x_i), z_j = (s_j, x_j) \text{ with } s_i = s_j \in [0, T), x_i \neq x_j\},$$

as well as

$$\mathcal{C} := \bigcup_{T \in \mathbb{N}} \bigcup_{i \neq j \in \llbracket n \rrbracket} \mathcal{C}_{i,j}(T).$$

**Lemma 2.7** *Let  $n \in \mathbb{N}$ ,  $n \geq 2$ . Then  $\mathbf{M}_\varepsilon^n(\mathcal{C}) = 0$  almost surely.*

**Proof** It suffices to show that for all  $T > 0$  and all  $i \neq j \in \llbracket n \rrbracket$  we have  $\mathbf{M}_\varepsilon^n(\mathcal{C}_{i,j}(T)) = 0$ . We have

$$\mathbf{M}_\varepsilon^n(\mathcal{C}_{i,j}(T)) = \mathbf{M}_\varepsilon^2\left(\left\{((s, x_1), (s, x_2)) : x_1 \neq x_2 \text{ and } s \in [0, T)\right\}\right) \mathbf{M}_\varepsilon^{n-2}(\mathcal{D}_{\varepsilon, T}^{n-2}).$$

The quantity  $\mathbf{M}_\varepsilon^{n-2}(\mathcal{D}_{\varepsilon, T}^{n-2})$  is almost surely finite. Recall that the measure  $\mathbf{M}_\varepsilon^n$  is defined in terms of the martingales  $\mathbb{M}_\varepsilon$ , see (2.10), and that the martingales  $\mathbb{M}_\varepsilon$  are given by a sum of a jump part and an absolutely continuous part, see Assumption 1(4.). The absolutely continuous part does not contribute to the diagonal considered here, and we get

$$\begin{aligned} & \mathbf{M}_\varepsilon^2\left(\left\{((s, x_1), (s, x_2)) : x_1 \neq x_2 \text{ and } s \in [0, T)\right\}\right) \\ & \leq \sum_{\substack{x_1, x_2 \in \Lambda_\varepsilon \\ x_1 \neq x_2}} \varepsilon^{2d} \sum_{0 \leq s < T} |\Delta_s \mathbb{M}_\varepsilon(x_1) \Delta_s \mathbb{M}_\varepsilon(x_2)|. \end{aligned}$$

By Assumption 1(2.) the last expression is 0 almost surely.  $\square$

Combining (2.12) with Lemma 2.7, we obtain

$$\int_{\mathcal{D}_\varepsilon^n} F \, d\mathbf{M}_\varepsilon^n = \sum_{\gamma \in \mathfrak{C}(n)} \int_{\mathcal{D}_\gamma} F \, d\mathbf{M}_\varepsilon^n, \quad (2.13)$$

where  $\mathcal{D}_\gamma := \tilde{\mathcal{D}}_\gamma \setminus \mathcal{C}$ .

Lemma 2.7 ensures that we can decompose the integral further according to the order between the different components with respect to the time argument. For a given  $\gamma \in \mathfrak{C}_m(n)$ , we denote by  $\Sigma_\gamma$  the set of bijections from  $\llbracket m \rrbracket$  to the components of  $\gamma$ . We interpret  $\sigma \in \Sigma_\gamma$  as an ordering of the components and write  $[i]_\gamma \prec_\sigma [j]_\gamma$  if  $\sigma^{-1}([i]_\gamma) < \sigma^{-1}([j]_\gamma)$ . Given an ordering  $\sigma \in \Sigma_\gamma$  over the components of a contraction  $\gamma \in \mathfrak{C}(n)$ , we define the sets

$$\mathcal{D}_{\gamma, \sigma} := \left\{((s_1, x_1), \dots, (s_n, x_n)) \in \mathcal{D}_\gamma : s_i < s_j \text{ whenever } [i]_\gamma \prec_\sigma [j]_\gamma\right\}.$$

Now (2.13) becomes

$$\int_{\mathcal{D}_\varepsilon^n} F \, d\mathbf{M}_\varepsilon^n = \sum_{\gamma \in \mathfrak{C}(n)} \sum_{\sigma \in \Sigma_\gamma} \int_{\mathcal{D}_{\gamma, \sigma}} F \, d\mathbf{M}_\varepsilon^n. \quad (2.14)$$

### 2.3.2 A recursive representation

We introduce the notation

$$\mathcal{I}_{\gamma, \sigma}^\varepsilon F := \int_{\mathcal{D}_{\gamma, \sigma}} F \, d\mathbf{M}_\varepsilon^n, \quad (2.15)$$

which allows to write (2.14) as

$$\int_{\mathcal{D}_\varepsilon^n} F \, d\mathbf{M}_\varepsilon^n = \sum_{\gamma \in \mathfrak{C}(n)} \sum_{\sigma \in \Sigma_\gamma} \mathcal{I}_{\gamma, \sigma}^\varepsilon F. \quad (2.16)$$

We will call this expression *a chaos expansion* by analogy with the Wiener chaos expansion for multiple Wiener integrals. Usually, we will work with stochastic integrals on a finite time interval  $[0, t]$ , in which case the preceding two identities become

$$(\mathcal{I}_{\gamma, \sigma}^\varepsilon F)_t := \int_{\mathcal{D}_{\gamma, \sigma} \cap \mathcal{D}_{\varepsilon, t}^n} F \, d\mathbf{M}_\varepsilon^n \quad (2.17)$$

and

$$\int_{\mathcal{D}_\varepsilon^n \cap \mathcal{D}_{\varepsilon, t}^n} F \, d\mathbf{M}_\varepsilon^n = \sum_{\gamma \in \mathfrak{C}(n)} \sum_{\sigma \in \Sigma_\gamma} (\mathcal{I}_{\gamma, \sigma}^\varepsilon F)_t.$$

The aim of the next sections is to derive an estimate on moments of  $(\mathcal{I}_{\gamma, \sigma}^\varepsilon F)_t$  and this will be done recursively. To this end we introduce some more notation: first, for given  $\gamma \in \mathfrak{C}_m(n)$  and  $\sigma \in \Sigma_\gamma$  we define the function  $F^{\gamma, \sigma} : \mathcal{D}_\varepsilon^m \rightarrow \mathbf{R}$  by

$$F^{\gamma, \sigma}(z_1, \dots, z_m) := F(\bar{z}_1, \dots, \bar{z}_n), \quad (2.18)$$

where  $\bar{z} = (\bar{z}_1, \dots, \bar{z}_n) \in \mathcal{D}_\varepsilon^n$  is defined by

$$\bar{z}_i = z_j \iff i \text{ is a member of the } j\text{-th equivalence class according to } \sigma.$$

Furthermore, for any  $n$  we define the measure  $\mathbf{M}_{\varepsilon, \text{diag}}^n$  on  $\mathcal{D}_\varepsilon$  as

$$\mathbf{M}_{\varepsilon, \text{diag}}^n = \iota^* \mathbf{M}_\varepsilon^n |_{\mathcal{D}_{\varepsilon, \text{diag}}^n}, \quad (2.19)$$

where  $\mathcal{D}_{\varepsilon, \text{diag}}^n$  denotes the *full diagonal*  $\mathcal{D}_{\varepsilon, \text{diag}}^n := \{(z_1, z_2, \dots, z_n) \in \mathcal{D}_\varepsilon^n : z_1 = z_2 = \dots = z_n\}$ ,  $\mathbf{M}_\varepsilon^n |_{\mathcal{D}_{\varepsilon, \text{diag}}^n}$  is the restriction of the product measure  $\mathbf{M}_\varepsilon^n$  to  $\mathcal{D}_{\varepsilon, \text{diag}}^n$

and  $\iota^*$  denotes the image measure under the identification  $\iota: \mathcal{D}_{\varepsilon, \text{diag}}^n \rightarrow \mathcal{D}_\varepsilon$  given by  $\iota(z, z, \dots, z) = z$ . With this notations in place, the following recursive formula follows immediately.

**Lemma 2.8** *For any  $F \in \mathcal{D}_\varepsilon^n \rightarrow \mathbf{R}$  which is continuous in all time variables, for any contraction  $\gamma \in \mathfrak{C}_m(n)$  and any ordering  $\sigma \in \Sigma_\gamma$  we have*

$$(\mathcal{I}_{\gamma, \sigma}^\varepsilon F)_t = \int_{s_m < t} \int_{s_{m-1} < s_m} \dots \int_{s_1 < s_2} F^{\gamma, \sigma}(z_1, \dots, z_m) d\mathbf{M}_{\varepsilon, \text{diag}}^{|\gamma_{\sigma(1)}|}(z_1) \dots d\mathbf{M}_{\varepsilon, \text{diag}}^{|\gamma_{\sigma(m)}|}(z_m). \quad (2.20)$$

The subscript  $s_m < t$  in the first integral of (2.20) is used as a shorthand for  $\{z_m = (s_m, x_m) \in \mathcal{D}_\varepsilon : s_m < t\}$ , and similarly the subscripts  $s_i < s_{i+1}$  mean that the corresponding integrals are taken over  $\{z_i = (s_i, x_i) \in \mathcal{D}_\varepsilon : s_i < s_{i+1}\}$ .

## 2.4 Analysis of the measure on the diagonals

We analyse further the measures  $\mathbf{M}_{\varepsilon, \text{diag}}^n$  defined in (2.19). By definition, we have for  $F: \mathcal{D}_\varepsilon \rightarrow \mathbf{R}$

$$\int_{\mathcal{D}_\varepsilon} F d\mathbf{M}_{\varepsilon, \text{diag}}^n = \int_{\mathcal{D}_{\varepsilon, \text{diag}}^n} F(z_1) d\mathbf{M}_\varepsilon^n(z_1, z_2, \dots, z_n).$$

In this formula we may allow  $F$  to be random which does not affect our computations.

For  $n \geq 2$  Assumption 1(4.) implies that that only the jump parts of the martingales produce non-trivial contributions. We get that

$$\int_{s < t} F(s, x) d\mathbf{M}_{\varepsilon, \text{diag}}^n(s, x) = \sum_{x \in \Lambda_\varepsilon} \varepsilon^{nd} \sum_{0 \leq s < t} F(s, x) (\Delta_s \mathbb{M}_\varepsilon(x))^n \quad (2.21)$$

In the case  $n = 2$  we use the bracket processes of the martingales (see Sect. 2.1) to write (2.21) as

$$\begin{aligned} \int_{s < t} F(s, x) d\mathbf{M}_{\varepsilon, \text{diag}}^2(s, x) &= \sum_{x \in \Lambda_\varepsilon} \varepsilon^{2d} \int_{s < t} F(s, x) d[\mathbb{M}_\varepsilon(x)]_s \\ &= \sum_{x \in \Lambda_\varepsilon} \varepsilon^{2d} \int_{s < t} F(s, x) d\langle \mathbb{M}_\varepsilon(x) \rangle_s \\ &\quad + \sum_{x \in \Lambda_\varepsilon} \varepsilon^{2d} \int_{s < t} F(s, x) d([\mathbb{M}_\varepsilon(x)]_s - \langle \mathbb{M}_\varepsilon(x) \rangle_s) \\ &= \sum_{x \in \Lambda_\varepsilon} \varepsilon^d \int_{s < t} F(s, x) \mathbf{C}_\varepsilon(s, x) ds \\ &\quad + \sum_{x \in \Lambda_\varepsilon} \varepsilon^{2d+k} \int_{s < t} F(s, x) d\bar{\mathbb{M}}_\varepsilon(s, x), \end{aligned} \quad (2.22)$$

where in the last equality we have used Assumption 1(1.) in the first term and the definition of the martingales (2.6) in the second one.

To bound (2.21) for  $n \geq 3$  we make crucial use of the Assumption 1(3.) that guarantees that the jumps  $|\Delta_t \mathbb{M}(x)|$  are of fixed size  $c\varepsilon^k$ . Then for  $n$  odd we get

$$\begin{aligned} \int_{s < t} F(s, x) d\mathbb{M}_{\varepsilon, \text{diag}}^n(s, x) &= \sum_{x \in \Lambda_\varepsilon} \varepsilon^{nd} \sum_{0 \leq s < t} F(s, x) (c\varepsilon^k)^{n-1} \Delta_s \mathbb{M}_\varepsilon(x) \\ &= c^{n-1} \varepsilon^{(d+k)(n-1)} \sum_{x \in \Lambda_\varepsilon} \varepsilon^d \int_{0 \leq s < t} F(s, x) d\mathbb{M}_\varepsilon(s, x) \\ &\quad - c^{n-1} \varepsilon^{(d+k)(n-2)} \sum_{x \in \Lambda_\varepsilon} \varepsilon^d \int_{0 \leq s < t} F(s, x) C_\varepsilon(s, x) ds, \end{aligned} \quad (2.23)$$

where in the last identity we made use of Assumption 1(4.). Similarly, when  $n$  is even we get

$$\begin{aligned} \int_{s < t} F(s, x) d\mathbb{M}_{\varepsilon, \text{diag}}^n(s, x) &= (c\varepsilon^k)^{n-2} \sum_{x \in \Lambda_\varepsilon} \varepsilon^{nd} \sum_{0 \leq s < t} F(s, x) d[\mathbb{M}_\varepsilon(x)]_s \\ &= c^{n-2} \varepsilon^{(d+k)(n-1)} \sum_{x \in \Lambda_\varepsilon} \varepsilon^d \int_{0 \leq s < t} F(s, x) d\overline{\mathbb{M}}_\varepsilon(s, x) \\ &\quad + c^{n-2} \varepsilon^{(d+k)(n-2)} \sum_{x \in \Lambda_\varepsilon} \varepsilon^d \int_{0 \leq s < t} F(s, x) C_\varepsilon(s, x) ds, \end{aligned} \quad (2.24)$$

where we made use of the martingale (2.6) and Assumption 1(1.). Remarkably, the equations (2.23) and (2.24) are of exactly the same structure and consequently, even and odd contractions can be bounded in the same way.

### 3 Moment bounds for iterated integrals

We aim to estimate integrals (2.20). This is done recursively and here we perform the recursive step by deriving an estimate on

$$\int_{\mathcal{D}_\varepsilon} G_t(t, x) d\mathbb{M}_{\varepsilon, \text{diag}}^{n_0}(t, x), \quad (3.1)$$

where

$$G_t(s, x) = \int_{s_m < t} \int_{s_{m-1} < s_m} \cdots \int_{s_1 < s_2} F(s, x; z_1, \dots, z_m) d\mathbb{M}_{\varepsilon, \text{diag}}^{n_1}(z_1) \cdots d\mathbb{M}_{\varepsilon, \text{diag}}^{n_m}(z_m) \quad (3.2)$$

and  $n_0, n_1, \dots, n_m \geq 1$ . Throughout this section we make the assumption that  $F : \mathcal{D}_\varepsilon^{m+1} \rightarrow \mathbf{R}$  is a deterministic function that is  $\mathcal{C}^1$  in each time variable. We note that the domain of integration in (3.2) guarantees that for any fixed  $s$  the function  $t \mapsto G_t(s, x)$  is predictable. In the following Sect. 3.1 we bound the simple integral (3.1) in terms of  $G$  and  $F$ . The resulting estimate is then used in a recursive argument to bound the full iterated integral  $\mathcal{I}_{\gamma, \sigma}^\varepsilon F$  in the subsequent Sect. 3.2.

To control the function  $F$  we will use the following norm

$$\|F\|_{\mathcal{C}_s^1(L_\varepsilon^\infty)} := \|F\|_{L_\varepsilon^\infty} + \|\partial_s F\|_{L_\varepsilon^\infty}, \quad (3.3)$$

where the subscript  $s$  refers to the variable with respect to which the  $\mathcal{C}^1$ -norm is computed.

### 3.1 Simple integrals

We will need the following result.

**Lemma 3.1** *Let  $f : I \rightarrow \mathbf{R}$  be a  $\mathcal{C}^1$  function on a interval  $I \subseteq \mathbf{R}$  with length  $|I| > 0$ . Then for any  $p \geq 1$*

$$\sup_{t \in I} |f(t)|^p \leq \frac{1}{|I|} \int_I |f(r)|^p dr + p \left( \int_I |f(r)|^p dr \right)^{\frac{p-1}{p}} \left( \int_I |f'(r)|^p dr \right)^{\frac{1}{p}}. \quad (3.4)$$

**Proof** For any fixed  $t, t_0 \in I$  we can write  $f(t)^p = f(t_0)^p + \int_{t_0}^t p f(r)^{p-1} f'(r) dr$ . Taking absolute values, then using the Hölder inequality and finally taking the supremum over  $t$ , we deduce that

$$\sup_{t \in I} |f(t)|^p \leq |f(t_0)|^p + p \left( \int_I |f(r)|^p dr \right)^{\frac{p-1}{p}} \left( \int_I |f'(r)|^p dr \right)^{\frac{1}{p}}.$$

We conclude by averaging the variable  $t_0$  over the interval  $I$ .  $\square$

The following proposition provides moment bounds for a simple stochastic integral.

**Proposition 3.2** *Let  $G_t : \mathcal{D}_\varepsilon \rightarrow \mathbf{R}$  be a possibly random function, such that the function  $t \mapsto G_t(t, x)$  is predictable. Then for any  $p \geq 2$  and  $T \in [0, 1]$  we have*

$$\begin{aligned} \mathbf{E}_p \sup_{t \in [0, T]} \left| \int_{\mathcal{D}_{\varepsilon, t}} G_s(s, x) d\mathbf{M}_\varepsilon(s, x) \right| \\ \lesssim_p \left( \varepsilon^d \sum_{y \in \Lambda_\varepsilon} \int_0^T (\mathbf{E}_p G_r(r, y))^2 dr \right)^{\frac{1}{2}} + \varepsilon^{d+k} \mathbf{E}_p \sup_{(s, x) \in \mathcal{D}_{\varepsilon, T}} |G_s(s, x)|. \end{aligned} \quad (3.5)$$

If  $G_t(s, x)$  is of the form (3.2), then we have

$$\mathbf{E}_p \sup_{(s,x) \in \mathcal{D}_{\varepsilon,T}} |G_s(s, x)| \lesssim \varepsilon^{-(d+(d+k)n)\frac{1}{p}} \left\| \mathbf{E}_p \sup_{t \in [0, T]} |G_t| \right\|_{L_\varepsilon^\infty}^{\frac{p-1}{p}} \|F\|_{\mathcal{C}_s^1(L_\varepsilon^\infty)}^{\frac{1}{p}}, \quad (3.6)$$

where  $n_1 + \dots + n_m = n$  and where we use the norm (3.3).

**Remark 3.3** In our application,  $\partial_s F$  will be badly behaved and in general blow up as negative power in  $\varepsilon$ , similarly to the other exploding terms. However, as  $p$  can and will be chosen arbitrarily large, all can be absorbed in a small pre-factor  $\varepsilon^{d+k}$  and therefore these error terms are all harmless.

**Proof of Proposition 3.2** Using the Burkholder–Davis–Gundy inequality (Proposition 2.1) and Assumption 1, we get

$$\mathbf{E}_p \sup_{t \in [0, T]} \left| \int_{\mathcal{D}_{\varepsilon,t}} G_s(s, x) d\mathbf{M}_\varepsilon(s, x) \right| \quad (3.7)$$

$$\lesssim_p \mathbf{E} \left[ \left( \sum_{y \in \Lambda_\varepsilon} \varepsilon^{2d} \int_{r=0}^T G_r(r, y)^2 d\langle \mathbb{M}_\varepsilon(y) \rangle_r \right)^{\frac{p}{2}} \right]^{\frac{1}{p}} + \mathbf{E} \left[ \sup_{t \in [0, T]} |G_t(t, x) \Delta_t \mathbb{M}_\varepsilon(x)|^p \right]^{\frac{1}{p}}$$

$$\lesssim_p \mathbf{E} \left[ \left( \sum_{y \in \Lambda_\varepsilon} \varepsilon^d \int_0^T G_r(r, y)^2 dr \right)^{\frac{p}{2}} \right]^{\frac{1}{p}} + \varepsilon^{d+k} \mathbf{E} \left[ \sup_{(s,x) \in \mathcal{D}_{\varepsilon,T}} |G_s(s, x)|^p \right]^{\frac{1}{p}}, \quad (3.8)$$

where the proportionality constant in the last bound comes from Assumption 1(1.) and (3.). The first term on the right hand side of (3.8) can be controlled by the first term in (3.5) by an application of Minkowski's inequality. This yields the required bound (3.5).

Now we will prove (3.6). First, the supremum over the lattice points is replaced by a sum at the expense of a small negative power of  $\varepsilon$  and second, we use  $\sup_{s \in [0, T]} |G_s(s, x)| \leq \sup_{s, t \in [0, T]} |G_t(s, x)|$  to arrive at

$$\mathbf{E} \left[ \sup_{(s,x) \in \mathcal{D}_{\varepsilon,T}} |G_s(s, x)|^p \right]^{\frac{1}{p}} \leq \varepsilon^{-\frac{d}{p}} \left( \varepsilon^d \sum_{y \in \Lambda_\varepsilon} \mathbf{E} \left[ \sup_{s, t \in [0, T]} |G_t(s, y)|^p \right] \right)^{\frac{1}{p}}. \quad (3.9)$$

We note that this bound holds because  $\Lambda_\varepsilon$  is a finite lattice. We apply first (3.4) to the supremum in the variable  $s$  and then Hölder's inequality to bound the right-hand side of (3.9) by a constant times

$$\begin{aligned} & \varepsilon^{-\frac{d}{p}} \left( \varepsilon^d \sum_{y \in \Lambda_\varepsilon} \mathbf{E} \sup_{t \in [0, T]} \left[ \frac{1}{T} \int_0^T |G_t(s, y)|^p ds \right. \right. \\ & \quad \left. \left. + \left( \int_0^T |G_t(s, y)|^p ds \right)^{\frac{p-1}{p}} \left( \int_0^T |\partial_s G_t(s, y)|^p ds \right)^{\frac{1}{p}} \right] \right)^{\frac{1}{p}} \end{aligned}$$

$$\begin{aligned}
&\lesssim \varepsilon^{-\frac{d}{p}} \left( \varepsilon^d \sum_{y \in \Lambda_\varepsilon} \frac{1}{T} \int_0^T \mathbf{E} \left[ \sup_{t \in [0, T]} |G_t(s, y)|^p \right] ds \right. \\
&\quad \left. + \left( \varepsilon^d \sum_{y \in \Lambda_\varepsilon} \int_0^T \mathbf{E} \left[ \sup_{t \in [0, T]} |G_t(s, y)|^p \right] ds \right)^{\frac{p-1}{p}} \right. \\
&\quad \left. \left( \varepsilon^d \sum_{y \in \Lambda_\varepsilon} \int_0^T \mathbf{E} \left[ \sup_{t \in [0, T]} |\partial_s G_t(s, y)|^p \right] ds \right)^{\frac{1}{p}} \right)^{\frac{1}{p}}. \tag{3.10}
\end{aligned}$$

To bound the norms of the function  $G_t$ , we observe that

$$\varepsilon^d \sum_{y \in \Lambda_\varepsilon} \int_0^T \mathbf{E} \left[ \sup_{t \in [0, T]} |G_t(s, y)|^p \right] ds \lesssim T \sup_{(s, y) \in \mathcal{D}_{\varepsilon, T}} \mathbf{E} \left[ \sup_{t \in [0, T]} |G_t(s, y)|^p \right],$$

where we used that the grid  $\Lambda_\varepsilon$  is finite. Then (3.10) is estimated by a constant multiple of

$$\begin{aligned}
&\varepsilon^{-\frac{d}{p}} \left( \left\| \mathbf{E}_p \sup_{t \in [0, T]} |G_t| \right\|_{L_\varepsilon^\infty}^p + \left\| \mathbf{E}_p \sup_{t \in [0, T]} |G_t| \right\|_{L_\varepsilon^\infty}^{p-1} \right. \\
&\quad \left. \left( \varepsilon^d \sum_{y \in \Lambda_\varepsilon} \int_0^T \mathbf{E} \left[ \sup_{t \in [0, T]} |\partial_s G_t(s, y)|^p \right] ds \right)^{\frac{1}{p}} \right)^{\frac{1}{p}}, \tag{3.11}
\end{aligned}$$

where we used our assumption  $T \leq 1$ .

Now, we will bound the expectation containing  $\partial_s G_t$  in (3.11). The definition (3.2) yields

$$|\partial_s G_t(s, y)| \lesssim \|\partial_s F\|_{L_\varepsilon^\infty} \sum_{x_1, \dots, x_m \in \Lambda_\varepsilon} \varepsilon^{md} \prod_{i=1}^m \|\mathbf{M}_{\varepsilon, \text{diag}}^{n_i}(x_i)\|_{\text{TV}([0, T])},$$

and hence

$$\mathbf{E}_p \sup_{t \in [0, T]} |\partial_s G_t(s, y)| \lesssim \|\partial_s F\|_{L_\varepsilon^\infty} \sup_{x_1, \dots, x_m \in \Lambda_\varepsilon} \mathbf{E}_p \prod_{i=1}^m \|\mathbf{M}_{\varepsilon, \text{diag}}^{n_i}(x_i)\|_{\text{TV}([0, T])}.$$

Hölder's inequality allows to bound the preceding expression by

$$\mathbf{E}_p \sup_{t \in [0, T]} |\partial_s G_t(s, y)| \lesssim \|\partial_s F\|_{L_\varepsilon^\infty} \sup_{x_1, \dots, x_m \in \Lambda_\varepsilon} \prod_{i=1}^m \mathbf{E}_{mp} \|\mathbf{M}_{\varepsilon, \text{diag}}^{n_i}(x_i)\|_{\text{TV}([0, T])}.$$

Identities (2.23)/(2.24) and Lemma 2.5 yield

$$\sup_{x_i \in \Lambda_\varepsilon} \mathbf{E}_{mp} \|\mathbf{M}_{\varepsilon, \text{diag}}^{n_i}(x_i)\|_{\text{TV}([0, T])} \lesssim \varepsilon^{-(\mathbf{k}+d)n_i},$$

and we get

$$\mathbf{E}_p \sup_{t \in [0, T]} |\partial_s G_t(s, y)| \lesssim \|\partial_s F\|_{L_\varepsilon^\infty} \prod_{i=1}^m \varepsilon^{-(\mathbf{k}+d)n_i} \lesssim \|\partial_s F\|_{L_\varepsilon^\infty} \varepsilon^{-(\mathbf{k}+d)n}, \quad (3.12)$$

where  $n_1 + \dots + n_m = n$ .

Using (3.12) in (3.11), we conclude that (3.9) is estimated by a constant multiple of

$$\begin{aligned} & \varepsilon^{-\frac{d}{p}} \left( \left\| \mathbf{E}_p \sup_{t \in [0, T]} |G_t| \right\|_{L_\varepsilon^\infty}^p + \varepsilon^{-(d+\mathbf{k})n} \|\partial_s F\|_{L_\varepsilon^\infty} \left\| \mathbf{E}_p \sup_{t \in [0, T]} |G_t| \right\|_{L_\varepsilon^\infty}^{p-1} \right)^{\frac{1}{p}} \\ &= \varepsilon^{-\frac{d}{p}} \left\| \mathbf{E}_p \sup_{t \in [0, T]} |G_t| \right\|_{L_\varepsilon^\infty}^{\frac{p-1}{p}} \left( \left\| \mathbf{E}_p \sup_{t \in [0, T]} |G_t| \right\|_{L_\varepsilon^\infty} + \varepsilon^{-(d+\mathbf{k})n} \|\partial_s F\|_{L_\varepsilon^\infty} \right)^{\frac{1}{p}}, \end{aligned} \quad (3.13)$$

where we used our assumption  $T \leq 1$ . Similarly to (3.12) we get

$$\mathbf{E}_p \sup_{t \in [0, T]} |G_t(s, y)| \lesssim \|F\|_{L_\varepsilon^\infty} \prod_{i=1}^m \varepsilon^{-(\mathbf{k}+d)n_i} \lesssim \|F\|_{L_\varepsilon^\infty} \varepsilon^{-(\mathbf{k}+d)n},$$

and (3.13) is estimated by a constant multiple of

$$\varepsilon^{-\frac{d}{p} - (d+\mathbf{k})\frac{n}{p}} \left\| \mathbf{E}_p \sup_{t \in [0, T]} |G_t| \right\|_{L_\varepsilon^\infty}^{\frac{p-1}{p}} \|F\|_{\mathcal{C}_s^1(L_\varepsilon^\infty)}^{\frac{1}{p}},$$

where we used the norm (3.3). This gives the required bound (3.6).  $\square$

In connection with the representation developed in Sect. 2.3.2 we get the following.

**Proposition 3.4** *Let  $G_t$  be as in (3.5) and let  $n_0 \geq 2$ . Then for any  $p \geq 2$  and  $T \in [0, 1]$  we have*

$$\begin{aligned} & \mathbf{E}_p \sup_{t \in [0, T]} \left| \int_{\mathcal{D}_{\varepsilon, t}} G_s(s, x) d\mathbf{M}_{\varepsilon, \text{diag}}^{n_0}(s, x) \right| \\ & \lesssim_p \varepsilon^{(d+\mathbf{k})(n_0-1)} \left( \varepsilon^d \sum_{y \in \Lambda_\varepsilon} \int_0^T (\mathbf{E}_p G_r(r, y))^2 dr \right)^{\frac{1}{2}} \\ & \quad + \varepsilon^{(d+\mathbf{k})(n_0-2)} \left\| \mathbf{E}_p \sup_{t \in [0, T]} |G_t| \right\|_{L_\varepsilon^1} + \varepsilon^{(d+\mathbf{k})n_0} \mathbf{E}_p \sup_{(s, x) \in \mathcal{D}_{\varepsilon, T}} |G_s(s, x)|. \end{aligned} \quad (3.14)$$

If  $G_t(s, x)$  is of the form (3.2), then the last expectation is bounded by (3.6).

**Proof** Using (2.23)/(2.24) and Assumption 1 we get

$$\begin{aligned} \left| \int_{\mathcal{D}_{\varepsilon,t}} G_s(s, x) d\mathbf{M}_{\varepsilon, \text{diag}}^{n_0}(s, x) \right| &\lesssim \varepsilon^{(d+k)(n_0-1)} \left| \int_{\mathcal{D}_{\varepsilon,t}} G_s(s, x) d\mathbb{N}_\varepsilon(s, x) \right| \quad (3.15) \\ &\quad + \varepsilon^{(d+k)(n_0-2)} \sum_{x \in \Lambda_\varepsilon} \varepsilon^d \int_{0 \leq s \leq t} |G_s(s, x)| ds \end{aligned}$$

a.s., where the martingale  $\mathbb{N}_\varepsilon$  is either  $\mathbf{M}_\varepsilon$  or  $\overline{\mathbf{M}}_\varepsilon$ , depending on whether  $n_0$  is odd or even. In particular, the martingale satisfies Assumption 1. For the first term in (3.15) we use Proposition 3.2 to get

$$\begin{aligned} \mathbf{E}_p \sup_{t \in [0, T]} \left| \int_{\mathcal{D}_{\varepsilon,t}} G_s(s, x) d\mathbb{N}_\varepsilon(s, x) \right| &\lesssim \left( \varepsilon^d \sum_{y \in \Lambda_\varepsilon} \int_0^T (\mathbf{E}_p G_r(r, y))^2 dr \right)^{\frac{1}{2}} \\ &\quad + \varepsilon^{d+k} \mathbf{E}_p \sup_{(s,x) \in \mathcal{D}_{\varepsilon,T}} |G_s(s, x)|, \end{aligned}$$

and for the second term in (3.15) we have

$$\mathbf{E}_p \left| \sum_{x \in \Lambda_\varepsilon} \varepsilon^d \int_{0 \leq s \leq t} |G_s(s, x)| ds \right| \lesssim \left\| \mathbf{E}_p \sup_{t \in [0, T]} |G_t| \right\|_{L_\varepsilon^1}.$$

This gives the required bound (3.14).  $\square$

### 3.2 Bounds for general iterated integrals

Let  $n, m \in \mathbf{N}$  with  $m \leq n$  be fixed. For a contraction  $\gamma \in \mathfrak{C}_m(n)$  and a permutation  $\sigma \in \Sigma_\gamma$ , we want to prove moment bounds for general multiple iterated integrals (2.17). For this, we will define a norm on the function  $F^{\gamma, \sigma}$  from (2.18).

For  $m \geq 1$  we denote by  $\mathcal{P}_m$  the set of all functions  $\mathbf{p} : \llbracket m \rrbracket \rightarrow \{1, 2, \infty\}$ . Then for  $\mathbf{p} \in \mathcal{P}_1$  we set  $\|F^{\gamma, \sigma}\|_{L_\varepsilon^\mathbf{p}} = \|F^{\gamma, \sigma}\|_{L_\varepsilon^{\mathbf{p}(1)}}$ , and for  $m \geq 2$  and  $\mathbf{p} \in \mathcal{P}_m$  we define the norm recursively

$$\|F^{\gamma, \sigma}\|_{L_\varepsilon^\mathbf{p}} := \left\| \left\| (F^{\gamma, \sigma})^{(z_m)} \mathbb{1}_{s_m > s_{m-1}} \right\|_{L_\varepsilon^{\mathbf{p} \restriction \llbracket m-1 \rrbracket}} \right\|_{L_\varepsilon^{\mathbf{p}(m)}}, \quad (3.16)$$

where  $\mathbf{p} \restriction \llbracket m-1 \rrbracket$  is the restriction of  $\mathbf{p}$  to  $\llbracket m-1 \rrbracket$ , the function  $(F^{\gamma, \sigma})^{(z_m)} : \mathcal{D}_\varepsilon^{m-1} \rightarrow \mathbf{R}$  is defined as  $(F^{\gamma, \sigma})^{(z_m)}(z_1, \dots, z_{m-1}) = F^{\gamma, \sigma}(z_1, \dots, z_{m-1}, z_m)$ , and the outer norm in (3.16) is computed with respect to the variable  $z_m$ . The indicator  $\mathbb{1}_{s_m > s_{m-1}}$ , with the convention  $s_0 = 0$ , is needed to respect the domain of integration in (2.20).

For  $\mathbf{p} \in \mathcal{P}_m$  we will also use the standard notation  $\mathbf{p}^{-1} : \{1, 2, \infty\} \rightarrow 2^{\llbracket m \rrbracket}$  for the inverse function. For any  $\gamma \in \mathfrak{C}_m(n)$ , we define the set of non-contracted variables

$$\Gamma_1(\gamma) := \{i \in \llbracket m \rrbracket : |\gamma_i| = 1\}. \quad (3.17)$$

The following is our main result for the general iterated integrals.

**Theorem 3.5** *Let martingales  $(\mathbb{M}_\varepsilon(t, x))_{t \geq 0}$  satisfy Assumption 1,  $\gamma \in \mathfrak{C}_m(n)$ ,  $\sigma \in \Sigma_\gamma$  be any ordering for the components of  $\gamma$ , and let  $F: \mathcal{D}_\varepsilon^n \rightarrow \mathbf{R}$  be  $\mathcal{C}^1$  in each time variable. Then for every  $p \geq 2$  and  $T \in [0, 1]$*

$$E_p \left[ \sup_{t \in [0, T]} |(\mathcal{I}_{\gamma, \sigma}^\varepsilon F)_t| \right] \quad (3.18)$$

$$\lesssim_p \sum_{\substack{\mathbf{p} \in \mathcal{P}_m: \\ \mathbf{p}^{-1}(1) \cap \Gamma_1(\gamma) = \emptyset}} \varepsilon^{\alpha_\gamma(\mathbf{p})} \|F^{\gamma, \sigma}\|_{L_\varepsilon^{\mathbf{p}}}^{\beta_{\gamma, \mathbf{p}}(\mathbf{p})} \left( \prod_{\substack{i \in \mathbf{p}^{-1}(\infty) \setminus \Gamma_1(\gamma): \\ i \geq 2}} \varepsilon^{-\kappa_{\gamma, i}(\mathbf{p})} \|F^{\gamma, \sigma}\|_{\mathcal{C}_{s_i}^1(L_\varepsilon^\infty)}^{\beta_{\gamma^{\geq i}, \mathbf{p}}(\mathbf{p}^{\geq i})} \right)^{\frac{1}{p}},$$

for some constants  $\kappa_{\gamma, i}(\mathbf{p}) > 0$ , where the function  $F^{\gamma, \sigma}$  is defined in (2.18), the powers are

$$\alpha_\gamma(\mathbf{p}) := (d + \mathbf{k}) \left( \sum_{i=1}^m |\gamma_i| - 2|\mathbf{p}^{-1}(1)| - |\mathbf{p}^{-1}(2)| \right), \quad (3.19)$$

$$\beta_{\gamma, \mathbf{p}}(\mathbf{p}) := \prod_{\substack{i \in \mathbf{p}^{-1}(\infty) \setminus \Gamma_1(\gamma): \\ i \geq 2}} \left( \frac{p-1}{p} \right), \quad (3.20)$$

the contraction  $\gamma^{\geq i}$  has components  $\gamma_1^{\geq i}, \dots, \gamma_{m-i+1}^{\geq i}$  such that  $\gamma_j^{\geq i} = \gamma_{i+j-1}$ , and the function  $\mathbf{p}^{\geq i} \in \mathcal{P}_{m-i+1}$  is defined as  $\mathbf{p}^{\geq i}(j) = \mathbf{p}(i+j-1)$ .

**Remark 3.6** Precise values of the constants  $\kappa_{\gamma, i}(\mathbf{p})$  in (3.18) will not be important to us, although they may be obtained from the proof of Theorem 3.5. We show below that for  $p$  sufficiently large the divergent factors  $\varepsilon^{-\kappa_{\gamma, i}(\mathbf{p})/p}$  are compensated by the multiplier  $\varepsilon^{\alpha_\gamma(\mathbf{p})}$ .

**Remark 3.7** One can see that for any function  $\mathbf{p}$  in (3.18) satisfying  $\mathbf{p}^{-1}(\infty) \neq \emptyset$  we have  $\alpha_\gamma(\mathbf{p}) \geq d + \mathbf{k}$ . Indeed, we can estimate  $\sum_{i=1}^m |\gamma_i| \geq |\Gamma_1(\gamma)| + 2(m - |\Gamma_1(\gamma)|) = 2m - |\Gamma_1(\gamma)|$ , because the sum over  $|\Gamma_1(\gamma)|$  components  $\gamma_i$  of cardinality 1 is exactly  $|\Gamma_1(\gamma)|$  and the sum over the other  $m - |\Gamma_1(\gamma)|$  component is at least  $2(m - |\Gamma_1(\gamma)|)$ . On the other hand, the assumptions on  $\mathbf{p}$  yield  $2|\mathbf{p}^{-1}(1)| + |\mathbf{p}^{-1}(2)| < m + |\mathbf{p}^{-1}(1)| \leq 2m - |\Gamma_1(\gamma)|$ . Thus, we conclude that  $\alpha_\gamma(\mathbf{p}) \geq d + \mathbf{k}$ .

Similarly, we have  $\alpha_\gamma(\mathbf{p}) \geq d + \mathbf{k}$  if the contraction  $\gamma$  has a component  $\gamma_i$  such that  $|\gamma_i| \geq 3$ . Repeating the preceding computations, we get  $\sum_{i=1}^m |\gamma_i| > 2m - |\Gamma_1(\gamma)|$  and  $2|\mathbf{p}^{-1}(1)| + |\mathbf{p}^{-1}(2)| \leq 2m - |\Gamma_1(\gamma)|$ , which yield the required estimate.

Finally, one can see that we have  $\alpha_\gamma(\mathbf{p}) \geq d + \mathbf{k}$  if there is a component  $\gamma_i$  such that  $i \in \mathbf{p}^{-1}(2)$  and  $|\gamma_i| \geq 2$ .

Hence, if we have a sufficiently good control on the norms  $\|F^{\gamma, \sigma}\|_{L_\varepsilon^{\mathbf{p}}}$  and the product in (3.18) over  $i \in \mathbf{p}^{-1}(\infty)$  is bounded by  $C\varepsilon^{-\kappa}$  for some  $\kappa > 0$ , we can take  $p$  sufficiently large such that  $\varepsilon^{\alpha_\gamma(\mathbf{p})} \varepsilon^{-\kappa/p}$  vanishes as  $\varepsilon \rightarrow 0$ . So that we expect that the only non-vanishing terms in the limit  $\varepsilon \rightarrow 0$  are those with correspond to contractions

$\gamma$  with components of cardinalities at most 2 and functions  $\mathbf{p}$  satisfying  $\mathbf{p}^{-1}(\infty) = \emptyset$  and  $\mathbf{p}^{-1}(2) = \Gamma_1(\gamma)$ . In these cases we have  $\alpha_\gamma(\mathbf{p}) = 0$  and  $\beta_{\gamma,p}(\mathbf{p}) = 1$ , and the estimate (3.18) is similar to the bound for a multiple Wiener integral.

**Remark 3.8** The bound (3.18) is homogeneous with respect to  $F$ , i.e. multiplication of  $F$  by a constant  $\lambda > 0$  is equivalent to multiplication of the two sides of (3.18) by  $\lambda$ . While the homogeneity of the left-hand side is trivial, seeing it for the right-hand side is more complicated. Let us consider the term in the sum in (3.18) corresponding to a function  $\mathbf{p}$ .

If  $\mathbf{p}^{-1}(\infty) \setminus \Gamma_1(\gamma) = \emptyset$  or  $\mathbf{p}^{-1}(\infty) \setminus \Gamma_1(\gamma) = \{1\}$ , then  $\beta_{\gamma,p}(\mathbf{p}) = 1$  and the product in the parentheses in (3.18) equals 1. The respective term in the sum in (3.18) equals  $\varepsilon^{\alpha_\gamma(\mathbf{p})} \|F^{\gamma,\sigma}\|_{L_\varepsilon^\mathbf{p}}$  and is homogeneous with respect to  $F$ .

Let us now look at the case when the set  $\{i \in \mathbf{p}^{-1}(\infty) \setminus \Gamma_1(\gamma) : i \geq 2\}$  is non-empty, and let  $N$  be the magnitude of this set. Then the  $\lambda$ -multiplier corresponding to this term in the sum in (3.18) equals

$$\lambda^{\beta_{\gamma,p}(\mathbf{p})} \left( \prod_{\substack{i \in \mathbf{p}^{-1}(\infty) \setminus \Gamma_1(\gamma): \\ i \geq 2}} \lambda^{\beta_{\gamma^{\geq i},p}(\mathbf{p}^{\geq i})} \right)^{\frac{1}{p}}.$$

Furthermore, from the definition (3.20) we conclude that the power of  $\lambda$  may be written as

$$\left( \frac{p-1}{p} \right)^N + \frac{1}{p} \sum_{i=1}^N \left( \frac{p-1}{p} \right)^{N-i} = 1.$$

Hence, this term in the sum in (3.18) is homogeneous with respect to  $F$ .

**Proof of Theorem 3.5** We prove this theorem by induction over the number  $m$  of components in  $\gamma$ .

The base of induction is  $m = 1$ , in which case  $\gamma$  has only one component  $\gamma_1$  such that  $|\gamma_1| = n$ . If  $n = 1$  then the required bound (3.18) is given in Proposition 3.2. (In this case, only two functions  $\mathbf{p}$  contribute to the sum in (3.18):  $\mathbf{p}(1) = 2$  and  $\mathbf{p}(1) = \infty$ , which correspond to  $\alpha_\gamma(\mathbf{p}) = 0$  and  $\alpha_\gamma(\mathbf{p}) = d + \mathbf{k}$  respectively. In both cases  $\beta_{\gamma,p}(\mathbf{p}) = 1$ . Then the two terms on the right-hand side of (3.18) coincide with the two terms in (3.5).) If  $n \geq 2$  then the bound (3.18) is provided by Proposition 3.4. (The three functions  $\mathbf{p}$  contributing to the sum in (3.18) are  $\mathbf{p}(1) = 1$ ,  $\mathbf{p}(1) = 2$  and  $\mathbf{p}(1) = \infty$ , which correspond to  $\alpha_\gamma(\mathbf{p}) = (d + \mathbf{k})(n-2)$ ,  $\alpha_\gamma(\mathbf{p}) = (d + \mathbf{k})(n-1)$  and  $\alpha_\gamma(\mathbf{p}) = (d + \mathbf{k})n$  respectively. In all cases  $\beta_{\gamma,p}(\mathbf{p}) = 1$ .)

To make an inductive step, we assume that (3.18) holds for all contractions having  $m$  components and we will prove it for a contraction  $\gamma \in \mathfrak{C}_{m+1}(n)$ . Lemma 2.8 yields

$$(\mathcal{I}_{\gamma,\sigma}^\varepsilon F)_t = \int_{s_{m+1} < t} \left( \mathcal{I}_{\gamma,\bar{\sigma}}^\varepsilon F^{(z_{m+1})} \right)_{s_{m+1}-} \mathrm{d}\mathbf{M}_{\varepsilon,\mathrm{diag}}^{|\gamma_{m+1}|}(z_{m+1}), \quad (3.21)$$

where  $\bar{\gamma}$  is the contraction obtained from  $\gamma$  by removing the  $(m+1)^{\text{th}}$  component,  $\bar{\sigma}$  is obtained by restricting  $\sigma$  to  $\llbracket m \rrbracket$ , and the function  $F^{(z_{m+1})}$  is obtained from  $F$  by setting all the variables, whose labels are in  $(m+1)$ -st equivalence class according to  $\sigma$ , to  $z_{m+1}$ . If  $|\gamma_{m+1}| = 1$ , we bound (3.21) using Proposition 3.2:

$$\begin{aligned} \mathbf{E}_p \left[ \sup_{t \in [0, T]} |(\mathcal{I}_{\gamma, \sigma}^\varepsilon F)_t| \right] &\lesssim_p \left\| \mathbf{E}_p \sup_{t \in [0, T]} \left| \left( \mathcal{I}_{\bar{\gamma}, \bar{\sigma}}^\varepsilon F^{(z_{m+1})} \right)_t \right| \right\|_{L_\varepsilon^2} \\ &\quad + \varepsilon^{d+\mathbf{k}} \left\| \mathbf{E}_p \sup_{t \in [0, T]} \left| \left( \mathcal{I}_{\bar{\gamma}, \bar{\sigma}}^\varepsilon F^{(z_{m+1})} \right)_t \right| \right\|_{L_\varepsilon^\infty}, \end{aligned} \quad (3.22)$$

and if  $|\gamma_{m+1}| \geq 2$ , we use Proposition 3.4 and the estimate (3.6):

$$\begin{aligned} \mathbf{E}_p \left[ \sup_{t \in [0, T]} |(\mathcal{I}_{\gamma, \sigma}^\varepsilon F)_t| \right] &\lesssim_p \varepsilon^{(d+\mathbf{k})(|\gamma_{m+1}|-1)} \left\| \mathbf{E}_p \sup_{t \in [0, T]} \left| \left( \mathcal{I}_{\bar{\gamma}, \bar{\sigma}}^\varepsilon F^{(z_{m+1})} \right)_t \right| \right\|_{L_\varepsilon^2} \\ &\quad + \varepsilon^{(d+\mathbf{k})(|\gamma_{m+1}|-2)} \left\| \mathbf{E}_p \sup_{t \in [0, T]} \left| \left( \mathcal{I}_{\bar{\gamma}, \bar{\sigma}}^\varepsilon F^{(z_{m+1})} \right)_t \right| \right\|_{L_\varepsilon^1} \\ &\quad + \varepsilon^{(d+\mathbf{k})|\gamma_{m+1}|} \varepsilon^{-(d+(d+\mathbf{k})(n-|\gamma_{m+1}|)) \frac{1}{p}} \left\| \mathbf{E}_p \right. \\ &\quad \left. \sup_{t \in [0, T]} \left| \left( \mathcal{I}_{\bar{\gamma}, \bar{\sigma}}^\varepsilon F^{(z_{m+1})} \right)_t \right| \right\|_{L_\varepsilon^\infty}^{\frac{p-1}{p}} \|F^{\gamma, \sigma}\|_{\mathcal{C}_{s_{m+1}}^1(L_\varepsilon^\infty)}^{\frac{1}{p}}. \end{aligned} \quad (3.23)$$

The function inside the expectations is itself an iterated integral of the function  $F^{(z_{m+1})}$  with the contraction  $\bar{\gamma}$  having  $m$  components. We can use the induction hypothesis and the simple bound  $\|(F^{\bar{\gamma}, \bar{\sigma}})^{(z_{m+1})}\|_{\mathcal{C}_{s_i}^1(L_\varepsilon^\infty)} \leq \|F^{\gamma, \sigma}\|_{\mathcal{C}_{s_i}^1(L_\varepsilon^\infty)}$  to get moment bounds for the expectation:

$$\mathbf{E}_p \left[ \sup_{t \in [0, T]} \left| \left( \mathcal{I}_{\bar{\gamma}, \bar{\sigma}}^\varepsilon F^{(z_{m+1})} \right)_t \right| \right] \lesssim_p \sum_{\substack{\bar{\mathbf{p}} \in \mathcal{P}_m: \\ \bar{\mathbf{p}}^{-1}(1) \cap \Gamma_1(\bar{\gamma}) = \emptyset}} \varepsilon^{\alpha_{\bar{\gamma}}(\bar{\mathbf{p}})} \|(F^{\bar{\gamma}, \bar{\sigma}})^{(z_{m+1})}\|_{L_\varepsilon^{\bar{\mathbf{p}}}}^{\beta_{\bar{\gamma}, p}(\bar{\mathbf{p}})} E_p(\bar{\gamma}, \bar{\mathbf{p}}), \quad (3.24)$$

where we denoted

$$E_p(\bar{\gamma}, \bar{\mathbf{p}}) := \left( \prod_{\substack{i \in \bar{\mathbf{p}}^{-1}(\infty) \setminus \Gamma_1(\bar{\gamma}): \\ i \geq 2}} \varepsilon^{-\kappa_{\bar{\gamma}, i}(\bar{\mathbf{p}})} \|F^{\gamma, \sigma}\|_{\mathcal{C}_{s_i}^1(L_\varepsilon^\infty)}^{\beta_{\bar{\gamma}, p}(\bar{\mathbf{p}})} \right)^{\frac{1}{p}}, \quad (3.25)$$

for some constants  $\kappa_{\bar{\gamma}, i}(\bar{\mathbf{p}}) > 0$ . Then we use the preceding bound in (3.22) to get

$$\mathbf{E}_p \left[ \sup_{t \in [0, T]} |(\mathcal{I}_{\gamma, \sigma}^\varepsilon F)_t| \right] \lesssim_p \sum_{\substack{\bar{\mathbf{p}} \in \mathcal{P}_m: \\ \bar{\mathbf{p}}^{-1}(1) \cap \Gamma_1(\bar{\gamma}) = \emptyset}} \varepsilon^{\alpha_{\bar{\gamma}}(\bar{\mathbf{p}})} \left\| (F^{\bar{\gamma}, \bar{\sigma}})^{(z_{m+1})} \right\|_{L_\varepsilon^{\bar{\mathbf{p}}}}^{\beta_{\bar{\gamma}, p}(\bar{\mathbf{p}})} E_p(\bar{\gamma}, \bar{\mathbf{p}}) \quad (3.26)$$

$$+ \sum_{\substack{\bar{\mathbf{p}} \in \mathcal{P}_m: \\ \bar{\mathbf{p}}^{-1}(1) \cap \Gamma_1(\bar{\gamma}) = \emptyset}} \varepsilon^{\alpha_{\bar{\gamma}}(\bar{\mathbf{p}}) + d + \mathbf{k}} \left\| (F^{\bar{\gamma}, \bar{\sigma}})^{(z_{m+1})} \right\|_{L_{\varepsilon}^{\bar{\mathbf{p}}}}^{\beta_{\bar{\gamma}, p}(\bar{\mathbf{p}})} E_p(\bar{\gamma}, \bar{\mathbf{p}}).$$

We have  $\left\| (F^{\bar{\gamma}, \bar{\sigma}})^{(z_{m+1})} \right\|_{L_{\varepsilon}^{\bar{\mathbf{p}}}}^{\beta_{\bar{\gamma}, p}(\bar{\mathbf{p}})} \leq \left\| (F^{\bar{\gamma}, \bar{\sigma}})^{(z_{m+1})} \right\|_{L_{\varepsilon}^{\bar{\mathbf{p}}}}^{\beta_{\bar{\gamma}, p}(\bar{\mathbf{p}})}$ . Moreover, we have  $\beta_{\bar{\gamma}, p}(\bar{\mathbf{p}}) \leq 1$  and Jensen's inequality yields  $\left\| (F^{\bar{\gamma}, \bar{\sigma}})^{(z_{m+1})} \right\|_{L_{\varepsilon}^{\bar{\mathbf{p}}}}^{\beta_{\bar{\gamma}, p}(\bar{\mathbf{p}})} \lesssim \left\| (F^{\bar{\gamma}, \bar{\sigma}})^{(z_{m+1})} \right\|_{L_{\varepsilon}^{\bar{\mathbf{p}}}}^{\beta_{\bar{\gamma}, p}(\bar{\mathbf{p}})}$ . We introduce new functions  $\mathbf{p} \in \mathcal{P}_{m+1}$ , such that  $\mathbf{p}(i) = \bar{\mathbf{p}}(i)$  for  $i \in \llbracket m \rrbracket$ , and  $\mathbf{p}(m+1) = 2$  and  $\mathbf{p}(m+1) = \infty$  in the two sums in (3.26) respectively. Then (3.19) and  $|\gamma_{m+1}| = 1$  imply that the powers of  $\varepsilon$  in (3.26) are exactly  $\alpha_{\gamma}(\mathbf{p})$ . Furthermore, (3.20) yields  $\beta_{\bar{\gamma}, p}(\bar{\mathbf{p}}) = \beta_{\gamma, p}(\mathbf{p})$  and (3.25) yields  $E_p(\bar{\gamma}, \bar{\mathbf{p}}) = E_p(\gamma, \mathbf{p})$ . Hence, recalling the definition (3.16) we get the required bound (3.18) for  $m+1$ .

Now, we use the bound (3.24) in (3.23) and get

$$\begin{aligned} & \mathbf{E}_p \left[ \sup_{t \in [0, T]} |(\mathcal{I}_{\gamma, \sigma}^{\varepsilon} F)_t| \right] \\ & \lesssim_p \sum_{\substack{\bar{\mathbf{p}} \in \mathcal{P}_m: \\ \bar{\mathbf{p}}^{-1}(1) \cap \Gamma_1(\bar{\gamma}) = \emptyset}} \varepsilon^{\alpha_{\bar{\gamma}}(\bar{\mathbf{p}}) + (d + \mathbf{k})(|\gamma_{m+1}| - 1)} \left\| (F^{\bar{\gamma}, \bar{\sigma}})^{(z_{m+1})} \right\|_{L_{\varepsilon}^{\bar{\mathbf{p}}}}^{\beta_{\bar{\gamma}, p}(\bar{\mathbf{p}})} E_p(\bar{\gamma}, \bar{\mathbf{p}}) \\ & + \sum_{\substack{\bar{\mathbf{p}} \in \mathcal{P}_m: \\ \bar{\mathbf{p}}^{-1}(1) \cap \Gamma_1(\bar{\gamma}) = \emptyset}} \varepsilon^{\alpha_{\bar{\gamma}}(\bar{\mathbf{p}}) + (d + \mathbf{k})(|\gamma_{m+1}| - 2)} \left\| (F^{\bar{\gamma}, \bar{\sigma}})^{(z_{m+1})} \right\|_{L_{\varepsilon}^{\bar{\mathbf{p}}}}^{\beta_{\bar{\gamma}, p}(\bar{\mathbf{p}})} E_p(\bar{\gamma}, \bar{\mathbf{p}}) \quad (3.27) \\ & + \sum_{\substack{\bar{\mathbf{p}} \in \mathcal{P}_m: \\ \bar{\mathbf{p}}^{-1}(1) \cap \Gamma_1(\bar{\gamma}) = \emptyset}} \varepsilon^{\alpha_{\bar{\gamma}}(\bar{\mathbf{p}}) \frac{p-1}{p} + (d + \mathbf{k})|\gamma_{m+1}| - (d + (d + \mathbf{k})(n - |\gamma_{m+1}|)) \frac{1}{p}} \left\| (F^{\bar{\gamma}, \bar{\sigma}})^{(z_{m+1})} \right\|_{L_{\varepsilon}^{\bar{\mathbf{p}}}}^{\beta_{\bar{\gamma}, p}(\bar{\mathbf{p}})} \\ & \times E_p(\bar{\gamma}, \bar{\mathbf{p}})^{\frac{p-1}{p}} \|F^{\gamma, \sigma}\|_{\mathcal{C}_{s_{m+1}}^1(L_{\varepsilon}^{\infty})}^{\frac{1}{p}}, \end{aligned}$$

where in the last line we used subadditivity of the function  $x \mapsto x^{\frac{p-1}{p}}$  for  $x \geq 0$ . As above, we estimate the norms by moving the power  $\beta_{\bar{\gamma}, p}(\bar{\mathbf{p}})$  to the outer norms. Furthermore, we introduce functions  $\mathbf{p} \in \mathcal{P}_{m+1}$ , such that  $\mathbf{p}(i) = \bar{\mathbf{p}}(i)$  for  $i \in \llbracket m \rrbracket$ , and  $\mathbf{p}(m+1) = 2$ ,  $\mathbf{p}(m+1) = 1$  and  $\mathbf{p}(m+1) = \infty$  in the three sums respectively. Then the powers of  $\varepsilon$  in the first and second sums in (3.27) equal  $\alpha_{\gamma}(\mathbf{p})$ . Moreover, we have  $\beta_{\bar{\gamma}, p}(\bar{\mathbf{p}}) = \beta_{\gamma, p}(\mathbf{p})$  and  $E_p(\bar{\gamma}, \bar{\mathbf{p}}) = E_p(\gamma, \mathbf{p})$  in these sums. The last sum in (3.27) is more complicated. The power of  $\varepsilon$  equals  $\alpha_{\gamma}(\mathbf{p}) - \kappa_{\gamma, m+1}(\mathbf{p}) \frac{1}{p}$  with  $\kappa_{\gamma, m+1}(\mathbf{p}) = \alpha_{\bar{\gamma}}(\bar{\mathbf{p}}) + d + (d + \mathbf{k})(n - |\gamma_{m+1}|) > 0$ . Furthermore,  $\beta_{\bar{\gamma}, p}(\bar{\mathbf{p}})^{\frac{p-1}{p}} = \beta_{\gamma, p}(\mathbf{p})$ . Setting  $\kappa_{\gamma, i}(\mathbf{p}) = \kappa_{\bar{\gamma}, i}(\bar{\mathbf{p}})^{\frac{p-1}{p}}$  for  $i \leq m$ , we get

$$\begin{aligned}
& \varepsilon^{-\kappa_{\gamma, m+1}(\mathbf{p})} \frac{1}{p} E_p(\bar{\gamma}, \bar{\mathbf{p}}) \frac{p-1}{p} \|F^{\gamma, \sigma}\|_{\mathcal{C}_{s_{m+1}}^1(L_\varepsilon^\infty)}^{\frac{1}{p}} \\
&= \left( \varepsilon^{-\kappa_{\gamma, m+1}(\mathbf{p})} \|F^{\gamma, \sigma}\|_{\mathcal{C}_{s_{m+1}}^1(L_\varepsilon^\infty)} \left( \prod_{\substack{i \in \bar{\mathbf{p}}^{-1}(\infty) \setminus \Gamma_1(\bar{\gamma}): \\ i \geq 2}} \varepsilon^{-\kappa_{\bar{\gamma}, i}(\bar{\mathbf{p}})} \|F^{\gamma, \sigma}\|_{\mathcal{C}_{s_i}^1(L_\varepsilon^\infty)}^{\beta_{\bar{\gamma}, i, p}(\bar{\mathbf{p}}_{\geq i})} \right)^{\frac{p-1}{p}} \right)^{\frac{1}{p}} \\
&= \left( \prod_{\substack{i \in \bar{\mathbf{p}}^{-1}(\infty) \setminus \Gamma_1(\gamma): \\ i \geq 2}} \varepsilon^{-\kappa_{\gamma, i}(\mathbf{p})} \|F^{\gamma, \sigma}\|_{\mathcal{C}_{s_i}^1(L_\varepsilon^\infty)}^{\beta_{\gamma, i, p}(\mathbf{p}_{\geq i})} \right)^{\frac{1}{p}},
\end{aligned}$$

where we used the identities  $\beta_{\gamma_{\geq m+1}, p}(\mathbf{p}_{\geq m+1}) = 1$  and  $\beta_{\gamma_{\geq i}, p}(\mathbf{p}_{\geq i}) = \beta_{\bar{\gamma}_{\geq i}, p}(\bar{\mathbf{p}}_{\geq i}) \frac{p-1}{p}$  which follow from the definitions. The preceding expression has the form (3.25) for the contraction  $\gamma$  and the function  $\mathbf{p}$ . Hence, recalling the definition (3.16), we get from (3.27) the required bound (3.18) for  $m+1$ .  $\square$

### 3.2.1 Renormalised iterated integrals

In the theory of regularity structures [26], there is usually the need to renormalise stochastic objects. Introducing renormalised integrals against martingales is the goal of this section.

Let  $(\mathbb{M}_\varepsilon(t, x))_{t \geq 0}$  be martingales satisfying Assumption 1. Let a function  $F : \mathcal{D}_\varepsilon^n \rightarrow \mathbf{R}$  be as in (2.17), where the contraction  $\gamma$  has only one component  $\gamma_1$ , such that  $|\gamma_1| = n$  is even, and let the permutation  $\sigma$  be trivial. We define the integral

$$(\mathcal{I}_{\gamma, \sigma}^{\varepsilon, \nabla} F)_t := \varepsilon^{(d+k)(n-2)} \int_{\mathcal{D}_{\varepsilon, t}} F^{\gamma, \sigma}(z) \mathbf{C}_\varepsilon(z) dz, \quad (3.28)$$

where we use the function  $F^{\gamma, \sigma}$  defined in (2.18). In this expression we integrate the contracted variable with respect to the bracket process (2.4) of the martingale. Furthermore, we define the renormalised integral

$$(\mathcal{I}_{\gamma, \sigma}^{\varepsilon, \diamond} F)_t := (\mathcal{I}_{\gamma, \sigma}^{\varepsilon} F)_t - (\mathcal{I}_{\gamma, \sigma}^{\varepsilon, \nabla} F)_t = \varepsilon^{(d+k)(n-1)} \int_{\mathcal{D}_{\varepsilon, t}} F^{\gamma, \sigma}(s, x) d\bar{\mathbb{M}}_\varepsilon(s, x), \quad (3.29)$$

where the last equality follows from (2.24), with the martingale  $\bar{\mathbb{M}}_\varepsilon$  defined in (2.6). As we will see in our application in Sect. 5, we will consider the situation when the integral  $\mathcal{I}_{\gamma, \sigma}^{\varepsilon}$  diverges as  $\varepsilon \rightarrow 0$ , and in order to control the latter we need to consider its renormalisation  $\mathcal{I}_{\gamma, \sigma}^{\varepsilon, \diamond}$  instead. If the noise was Gaussian, then the renormalising term  $\mathcal{I}_{\gamma, \sigma}^{\varepsilon, \nabla}$  would be deterministic. In our case, it is however a stochastic process.

In general, let  $\gamma \in \mathfrak{C}_m(n)$  with  $1 \leq m \leq n$ , and let  $\sigma \in \Sigma_\gamma$  be a permutation. Moreover, let us label components of  $\gamma$  using  $\mathbf{L} \in \{\nabla, \diamond, \text{nil}\}^{\llbracket m \rrbracket}$ , such that the label  $\mathbf{L}(i)$  assigned to a component shows with respect to which process the variable is

integrated. For  $m = 1$  we set

$$(\mathcal{I}_{\gamma, \sigma}^{\varepsilon, \mathbb{L}} F)_t := \begin{cases} (\mathcal{I}_{\gamma_1, \sigma}^{\varepsilon, \nabla} F)_t & \text{if } \mathbb{L}(1) = \nabla, \\ (\mathcal{I}_{\gamma_1, \sigma}^{\varepsilon, \diamond} F)_t & \text{if } \mathbb{L}(1) = \diamond, \\ (\mathcal{I}_{\gamma_1, \sigma}^{\varepsilon} F)_t & \text{if } \mathbb{L}(1) = \text{nil}. \end{cases}$$

Since we defined the integrals (3.28) and (3.29) only for even  $n$ , we will always assume that  $\mathbb{L}(i) = \text{nil}$  for any  $i$  such that  $|\gamma(i)|$  is odd. Then for  $m \geq 2$  we define recursively

$$(\mathcal{I}_{\gamma, \sigma}^{\varepsilon, \mathbb{L}} F)_t := \begin{cases} \varepsilon^{(d+k)(|\gamma_m|-2)} \int_{\mathcal{D}_{\varepsilon, t}} \left( \mathcal{I}_{\bar{\gamma}, \bar{\sigma}}^{\varepsilon, \mathbb{L} \restriction [m-1]} F(z_m) \right)_{s_m-} \mathbf{C}_\varepsilon(z_m) dz_m & \text{if } \mathbb{L}(m) = \nabla, \\ \varepsilon^{(d+k)(|\gamma_m|-1)} \int_{\mathcal{D}_{\varepsilon, t}} \left( \mathcal{I}_{\bar{\gamma}, \bar{\sigma}}^{\varepsilon, \mathbb{L} \restriction [m-1]} F(z_m) \right)_{s_m-} d\bar{\mathbb{M}}_\varepsilon(z_m) & \text{if } \mathbb{L}(m) = \diamond, \\ \int_{\mathcal{D}_{\varepsilon, t}} \left( \mathcal{I}_{\bar{\gamma}, \bar{\sigma}}^{\varepsilon, \mathbb{L} \restriction [m-1]} F(z_m) \right)_{s_m-} d\mathbf{M}_{\varepsilon, \text{diag}}^{|\gamma_m|}(z_m) & \text{if } \mathbb{L}(m) = \text{nil}, \end{cases} \quad (3.30)$$

where the function  $F^{(z_m)}$  is obtained from  $F$  by setting the values of the variables in  $\{\sigma(i) : i \in \gamma_m\}$  to  $z_m$ , where  $\bar{\gamma}$  is the contraction that removes the  $m^{\text{th}}$  component of  $\gamma$ , where the labeling  $\mathbb{L}$  is restricted to the indices in  $[\![m-1]\!]$ , and where  $\bar{\sigma}$  is the restriction of  $\sigma$  to  $[\![m-1]\!]$ .

For a labeling  $\mathbb{L}$  it will be convenient to define the sets

$$\mathbb{L}^{-1}(\nabla) := \{i : \mathbb{L}(i) = \nabla\}, \quad \mathbb{L}^{-1}(\diamond) := \{i : \mathbb{L}(i) = \diamond\},$$

which contain the indices of the components labeled by “ $\nabla$ ” and “ $\diamond$ ” respectively. Similarly to (3.17) we define the set

$$\Gamma(\gamma) := \Gamma_1(\gamma) \cup \mathbb{L}^{-1}(\diamond). \quad (3.31)$$

of variables integrated with respect to martingales.

The following result is an analogue of Theorem 3.5 for the renormalised integrals.

**Theorem 3.9** *In the setting of Theorem 3.5, let  $\mathbb{L}$  be a labeling of the contraction  $\gamma \in \mathfrak{C}_m(n)$ . Then for every  $p \geq 2$  and  $T \in [0, 1]$*

$$\begin{aligned} & E_p \left[ \sup_{t \in [0, T]} \left| (\mathcal{I}_{\gamma, \sigma}^{\varepsilon, \mathbb{L}} F)_t \right| \right] \\ & \lesssim_p \sum_{\substack{\mathbf{p} \in \mathcal{P}_m: \\ \mathbf{p}^{-1}(1) \cap \Gamma(\gamma) = \emptyset, \\ \mathbb{L}^{-1}(\nabla) \subseteq \mathbf{p}^{-1}(1)}} \varepsilon^{\alpha_\gamma(\mathbf{p})} \|F^{\gamma, \sigma}\|_{L_\varepsilon^p}^{\beta_{\gamma, p}(\mathbf{p})} \left( \prod_{\substack{i \in \mathbf{p}^{-1}(\infty) \setminus \Gamma_1(\gamma): \\ i \geq 2}} \varepsilon^{-\kappa_{\gamma, i}(\mathbf{p})} \|F^{\gamma, \sigma}\|_{\mathcal{C}_{s_i}^1(L_\varepsilon^\infty)}^{\beta_{\gamma \geq i, p}(\mathbf{p}^{\geq i})} \right)^{\frac{1}{p}}, \end{aligned} \quad (3.32)$$

for some constants  $\kappa_{\gamma, i}(\mathbf{p}) > 0$ , where  $\alpha_\gamma(\mathbf{p})$  and  $\beta_{\gamma, p}(\mathbf{p})$  are defined in (3.19) and (3.20).

**Proof** The proof is analogous to the proof of Theorem 3.5, where we use the recursive definition (3.30) and the fact that the martingales  $\bar{\mathbb{M}}_\varepsilon$  satisfy Assumption 1. The restriction  $\mathbb{L}^{-1}(\nabla) \subseteq \mathbf{p}^{-1}(1)$  in the sum in (3.32) follows from the uniform bound on the integral with the label  $\nabla$ .  $\square$

**Remark 3.10** The same argument as in Remark 3.7 implies that as  $\varepsilon \rightarrow 0$  the non-vanishing expectations (3.32) are those with contractions  $\gamma$  having components of cardinalities at most 2 and functions  $\mathbf{p}$  satisfying  $\mathbf{p}^{-1}(\infty) = \emptyset$  and  $\mathbf{p}^{-1}(2) = \Gamma(\gamma)$ .

## 4 Kernels given by generalised convolutions

In this section we prove moment bounds for the iterated integrals (2.11), when the function  $F$  is given by convolutions of singular kernels (similar to the one introduced in [30, Appendix A]). This type of kernels appears in canonical lifts of random noises in the theory of regularity structures. However, the result presented in this section is different from [30] because of two reasons: first, our noise is non-Gaussian, and second, we prove bounds on the stochastic integrals rather than on deterministic objects which appear after Wick contractions of Gaussian noises. Moment bounds for stochastic integrals driven by a general stationary non-Gaussian noise were proved in [7]. In the latter work, the authors generalised the framework of [30, Appendix A] which allowed them to deal with more general contractions of noises. In our setting, we need to use Theorem 3.9, which requires estimating more complicated norms of the functions, in contrast to the  $L^2$  norms when the noise is Gaussian. Since we adjust the ideas of [30, Appendix A] to our framework, we equip our results and definitions with references to their analogues from this article.

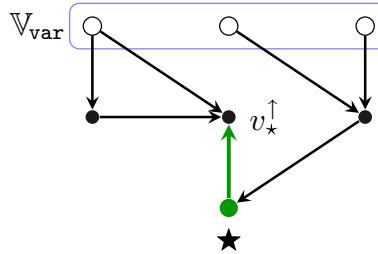
We will work in the space  $\mathbf{R}^{d+1}$  with the parabolic scaling  $\mathfrak{s} = (2, 1, \dots, 1)$ , where the first coordinate is time and the other  $d$  coordinates are spatial. We denote  $|\mathfrak{s}| := 2+d$ , and  $\|z\|_{\mathfrak{s}} := |t|^{1/2} + |x|$  for any  $z = (t, x) \in \mathbf{R}^{d+1}$ , and  $\lambda^{\mathfrak{s}} z := (\lambda^2 t, \lambda x)$ . For a multi-index  $k = (k_0, \dots, k_d) \in \mathbf{N}_0^{d+1}$  we define  $|k|_{\mathfrak{s}} := 2k_0 + \sum_{i=1}^d k_i$ . Then we denote by  $\mathcal{C}_{\mathfrak{s}}^r$  the space of function on  $\mathbf{R}^{d+1}$  with bounded mixed derivatives of the scaled order not exceeding  $r$ .

It will be convenient to consider processes  $\mathbb{M}_\varepsilon(t, x)$  defined on the whole time line  $\mathbf{R}$ . For this, we denote by  $\tilde{\mathbb{M}}_\varepsilon(t, x)$  an independent copy of  $\mathbb{M}_\varepsilon(t, x)$  and define

$$\mathbb{M}_\varepsilon(t, x) := \begin{cases} \mathbb{M}_\varepsilon(t, x) & \text{for } t \geq 0, \\ \tilde{\mathbb{M}}_\varepsilon(-t, x) & \text{for } t < 0, \end{cases} \quad (4.1)$$

for all  $t \in \mathbf{R}$ . Then the stochastic integral (2.10) can be naturally extended as

$$\begin{aligned} \int_{\mathbf{R} \times \Lambda_\varepsilon} F(z) d\mathbb{M}_\varepsilon(z) &= \varepsilon^d \sum_{x \in \Lambda_\varepsilon} \int_{s=0}^{\infty} F(s, x) d\mathbb{M}_\varepsilon(s, x) \\ &\quad + \varepsilon^d \sum_{x \in \Lambda_\varepsilon} \int_{s=0}^{\infty} F(-s, x) d\tilde{\mathbb{M}}_\varepsilon(s, x). \end{aligned} \quad (4.2)$$



**Fig. 1** An example of the graph  $\mathbb{G}$ , where the green edge connects the distinguished vertices  $\star$  and  $v_*^{\uparrow}$ . The white vertices are in  $\mathbb{V}_{\text{var}}$  and have only outgoing edges. The distinguished vertex  $\star$  has an incoming edge, but it cannot come from  $v_*^{\uparrow}$  (colour figure online)

The multiple integrals developed in Sect. 2.3 can be then naturally extended to whole  $\mathbf{R}$  in the time variable. With a little ambiguity we will use the notation as in Sect. 2.3 for the integral defined with respect to  $\mathbf{M}_\varepsilon$  on  $\mathbf{R} \times \Lambda_\varepsilon$ .

Following the idea of [30, Appendix A], it will be convenient to describe generalised convolutions using labelled graphs. More precisely, we consider a finite directed *graph*  $\mathbb{G} = (\mathbb{V}, \mathbb{E})$  with a set of vertices  $\mathbb{V}$  and with edges  $e \in \mathbb{E}$  labelled by pairs  $(a_e, r_e) \in \mathbf{R}_+ \times \mathbf{Z}$ . We assume that the graph is *weakly connected* and *loopless*, i.e. every vertex has either an outgoing or incoming edge, and there are no edges from a vertex to itself. We require  $\mathbb{G}$  to contain a distinguished vertex  $\star \in \mathbb{V}$ , connected by an outgoing edge with exactly one other vertex, denoted by  $v_*^{\uparrow} \in \mathbb{V} \setminus \{\star\}$ . We also allow  $\star$  to have incoming edges, which by the loopless assumption above cannot come from  $v_*^{\uparrow}$ . Finally, we assume that the graph contains a set  $\mathbb{V}_{\text{var}}$  of distinguished vertices, which can be empty and which satisfies  $\star \notin \mathbb{V}_{\text{var}}$ , and if it is non-empty, then it has only outgoing edges (“var” stands for “variables” because these vertices correspond to the variables integrated in the stochastic integral). This implies that there are no edges connecting two vertices from  $\mathbb{V}_{\text{var}}$ . In Fig. 1 we provide an example of such graph  $\mathbb{G}$ , where we omit labels and use various decorations for nodes and edges.

We define the set  $\mathbb{V}_* := \mathbb{V} \setminus \{\star\}$  and for a directed edge  $e \in \mathbb{E}$  we write  $e_+$  and  $e_-$  for the two vertices such that  $e = (e_-, e_+)$  is directed from  $e_-$  to  $e_+$ . We make the following assumption on the labels of the edges.

**Assumption 2** The described graph  $\mathbb{G}$  has the following properties:

1. every edge  $e$  containing  $\star$  has  $r_e = 0$ ;
2. the edge  $e = (\star, v_*^{\uparrow})$  has the label  $(a_e, r_e) = (0, 0)$ ;
3. at most one edge with  $r_e > 0$  may be incident to the same vertex;
4. if there are two vertices  $e_-$  and  $e_+$  such that the edge  $e = (e_-, e_+)$  has  $r_e < 0$ , then  $e_-$  and  $e_+$  have no other incident edge.

Let  $|\mathbb{V}_{\text{var}}| = n \geq 1$ . Then we label the elements  $\mathbb{V}_{\text{var}}$  by  $1, \dots, n$ , which gives a bijection between  $\mathbb{V}_{\text{var}}$  and  $[\![n]\!]$ . Using the notation of Sect. 2.3.1, we write  $\mathfrak{C}(\mathbb{V}_{\text{var}})$  for the set  $\mathfrak{C}(n)$  of all contractions on  $\mathbb{V}_{\text{var}}$ . For a graph  $\mathbb{G} = (\mathbb{V}, \mathbb{E})$  and a contraction  $\gamma \in \mathfrak{C}(\mathbb{V}_{\text{var}})$  we define the multigraph (i.e. two vertices are allowed to be connected by multiple edges)  $\mathbb{G}_\gamma = (\tilde{\mathbb{V}}, \tilde{\mathbb{E}})$ , with labels  $(\tilde{a}_e, \tilde{r}_e)$ , in the following way: the set

of vertices  $\tilde{\mathbb{V}} \subseteq \mathbb{V}$  is obtained from  $\mathbb{V}$  by identifying those vertices from  $\mathbb{V}_{\text{var}}$  which belong to the same component in  $\gamma$ . We denote this ‘‘identification’’ by a surjective map  $i_\gamma : \mathbb{V} \rightarrow \tilde{\mathbb{V}}$ . In particular,  $i_\gamma$  maps the vertices from  $\mathbb{V} \setminus \mathbb{V}_{\text{var}}$  (which includes  $\star$ ) to themselves. We define  $\tilde{\mathbb{V}}_{\text{var}}$  to be the image of  $\mathbb{V}_{\text{var}}$  under the map  $i_\gamma$ . Then we define the set of edges  $\tilde{\mathbb{E}}$  on  $\tilde{\mathbb{V}}$  to contain  $\tilde{e} = (i_\gamma(e_-), i_\gamma(e_+))$  for all  $e = (e_-, e_+) \in \mathbb{E}$ , with the label  $(\hat{a}_e, r_{\tilde{e}}) = (a_e, r_e)$ . In what follows, we call  $\mathbb{G}_\gamma = (\tilde{\mathbb{V}}, \tilde{\mathbb{E}})$  the *contracted (multi)graph* corresponding to  $\mathbb{G}$  and  $\gamma$ . To consider renormalised integrals, we will use a labeling  $L$  of the components of the contraction  $\gamma$ , defined as in Sect. 3.2.1. We by analogy with (3.31) we define the set of vertices

$$\Gamma(\gamma) := \{v \in \tilde{\mathbb{V}}_{\text{var}} : |i_\gamma^{-1}(v)| = 1\} \cup L^{-1}(\diamond), \quad (4.3)$$

which correspond to the variables integrated with respect to martingales. Throughout this section we will use the shorthand  $\Gamma = \Gamma(\gamma)$  because the contraction  $\gamma$  will be always fixed.

A special case is  $\mathbb{V}_{\text{var}} = \emptyset$ , in which all the definitions in the previous paragraph make sense for the identity contraction  $\gamma$ , and the contracted graph  $\mathbb{G}_\gamma$  coincides with the original one  $\mathbb{G}$ .

It will be useful to define a *simple* (containing no multiedges) graph  $(\hat{\mathbb{V}}, \hat{\mathbb{E}})$ , such that  $\hat{\mathbb{V}} = \tilde{\mathbb{V}}$  and the unique edge  $e \in \hat{\mathbb{E}}$  from  $e_-$  to  $e_+$  is obtained by contracting all edges  $\tilde{e}$  from  $e_-$  to  $e_+$  in  $\tilde{\mathbb{E}}$ , with the label  $(\hat{a}_e, r_e)$  of  $e$  being the sum of the labels of all such parallel edges  $\tilde{e}$ . It follows from Assumption 2 that if there is more than one edge connecting  $e_-$  to  $e_+$  in  $\mathbb{G}_\gamma$ , then the value  $r_e$  associated to the contracted edge is either 0 (if all these edges  $\tilde{e} \in \tilde{\mathbb{E}}$  have  $r_{\tilde{e}} = 0$ ), or coincides with the only value  $r_{\tilde{e}} > 0$ , for  $\tilde{e} \in \tilde{\mathbb{E}}$  connecting  $e_-$  to  $e_+$ . We can have  $r_e < 0$  only if there is a unique edge  $\tilde{e}$  from  $e_-$  to  $e_+$  with  $r_{\tilde{e}} < 0$ .

For a subset  $\bar{\mathbb{V}} \subseteq \mathbb{V}$  we define the outgoing edges  $\mathbb{E}^\uparrow(\bar{\mathbb{V}}) := \{e \in \mathbb{E} : e_- \in \bar{\mathbb{V}}\}$ , incoming edges  $\mathbb{E}^\downarrow(\bar{\mathbb{V}}) := \{e \in \mathbb{E} : e_+ \in \bar{\mathbb{V}}\}$ , internal edges  $\mathbb{E}_0(\bar{\mathbb{V}}) := \{e \in \mathbb{E} : e_\pm \in \bar{\mathbb{V}}\}$ , and incident edges  $\mathbb{E}(\bar{\mathbb{V}}) := \{e \in \mathbb{E} : e_- \in \bar{\mathbb{V}} \text{ or } e_+ \in \bar{\mathbb{V}}\}$ . If  $\bar{\mathbb{V}} = \mathbb{V}$ , we simply write  $\mathbb{E}^\uparrow$ ,  $\mathbb{E}^\downarrow$ , etc. Furthermore, we define the sets  $\mathbb{E}_+(\bar{\mathbb{V}}) := \{e \in \mathbb{E}(\bar{\mathbb{V}}) : r_e > 0\}$ ,  $\mathbb{E}_-(\bar{\mathbb{V}}) := \{e \in \mathbb{E}(\bar{\mathbb{V}}) : r_e < 0\}$ ,  $\mathbb{E}_+^\uparrow := \mathbb{E}_+ \cap \mathbb{E}^\uparrow$  and  $\mathbb{E}_+^\downarrow := \mathbb{E}_+ \cap \mathbb{E}^\downarrow$ . These sets, defined for the edges  $\tilde{\mathbb{E}}$  and  $\hat{\mathbb{E}}$ , will have the respective decorations.

Then we require the contracted graph to satisfy the following assumption, which we state for the simple graph  $(\hat{\mathbb{V}}, \hat{\mathbb{E}})$  defined above.

**Assumption 3** The graph  $\mathbb{G} = (\mathbb{V}, \mathbb{E})$  and the contraction  $\gamma \in \mathcal{C}(\mathbb{V}_{\text{var}})$  are such that the graph  $(\hat{\mathbb{V}}, \hat{\mathbb{E}})$ , defined above, has the following properties:

1. for any edge  $e \in \hat{\mathbb{E}}$  one has  $\hat{a}_e + (r_e \wedge 0) < |\mathfrak{s}|$ ;
2. for every subset  $\bar{\mathbb{V}} \subseteq \hat{\mathbb{V}}_\star$  of cardinality at least 3 one has

$$\sum_{e \in \hat{\mathbb{E}}_0(\bar{\mathbb{V}})} \hat{a}_e < \left(2|\bar{\mathbb{V}}| - |\bar{\mathbb{V}} \cap \Gamma| - 1 - \mathbb{1}_{\bar{\mathbb{V}} \cap \Gamma = \emptyset}\right) \frac{|\mathfrak{s}|}{2};$$

3. for every subset  $\bar{\mathbb{V}} \subseteq \hat{\mathbb{V}}$  containing  $\star$  of cardinality at least 2 one has

$$\sum_{e \in \hat{\mathbb{E}}_0(\bar{\mathbb{V}})} \hat{a}_e + \sum_{e \in \hat{\mathbb{E}}_+^\uparrow(\bar{\mathbb{V}})} (\hat{a}_e + r_e - 1) - \sum_{e \in \hat{\mathbb{E}}_+^\downarrow(\bar{\mathbb{V}})} r_e < (2|\bar{\mathbb{V}}| - |\bar{\mathbb{V}} \cap \Gamma|) \frac{|\mathfrak{s}|}{2};$$

4. for every non-empty subset  $\bar{\mathbb{V}} \subseteq \hat{\mathbb{V}}_\star \setminus \{v_\star^\uparrow\}$  one has

$$\sum_{e \in \hat{\mathbb{E}}(\bar{\mathbb{V}}) \setminus \hat{\mathbb{E}}_+^\downarrow(\bar{\mathbb{V}})} \hat{a}_e + \sum_{e \in \hat{\mathbb{E}}_+^\uparrow(\bar{\mathbb{V}})} r_e - \sum_{e \in \hat{\mathbb{E}}_+^\downarrow(\bar{\mathbb{V}})} (r_e - 1) > (2|\bar{\mathbb{V}}| - |\bar{\mathbb{V}} \cap \Gamma|) \frac{|\mathfrak{s}|}{2}.$$

**Remark 4.1** Assumption 3 coincides with Assumption 3.17 in [7] on an “elementary graph”, where the set of “external vertices” (see Definition 3.13 in [7]) is given in our case by the set  $\Gamma$ .

#### 4.1 Kernels associated to the graph

Given a graph  $\mathbb{G} = (\mathbb{V}, \mathbb{E})$  as above, to each edge we associate a kernel and each vertex corresponds to a variable in the domain  $\mathbf{R}^{d+1}$ . Then, for  $e \in \mathbb{E}$ , the values  $a_e$  will describe the order of singularity of the kernel associated to the edge  $e$ . The value  $r_e$  will describe the order of renormalisation of this kernel. For every vertex  $v \notin \mathbb{V}_{\text{var}} \cup \{\star\}$  we assume to be given a measure  $\mu_v^\varepsilon$  on  $\mathbf{R}^{d+1}$  of the form

$$\mu_v^\varepsilon(dz) = \varepsilon^d \sum_{y \in \varepsilon \mathbf{Z}^d} \delta(y - x) dt dx, \quad (4.4)$$

where  $z = (t, x)$  with  $t \in \mathbf{R}$  and  $x \in \mathbf{R}^d$ , and where  $\delta$  is the Dirac delta function on  $\mathbf{R}^d$ . This measure counts the points in the space lattice and is the Lebesgue measure in time. Notice that, as  $\varepsilon \rightarrow 0$ , the measure  $\mu_v^\varepsilon$  converges in the weak-\* topology to the Lebesgue measure on  $\mathbf{R}^{d+1}$ .

For each edge of the graph we associate a kernel with the following properties.

**Assumption 4** For every  $e \in \mathbb{E}$  we consider a smooth<sup>1</sup> kernel  $K_e^\varepsilon: \mathbf{R}^{d+1} \rightarrow \mathbf{R}$ , which can be written as  $K_e^\varepsilon(z) = \sum_{n=0}^N K_e^{\varepsilon, n}(z)$  for  $N = -[\log_2 \mathfrak{e}]$  and for some  $\mathfrak{e} \in [\varepsilon, 1]$ , where the smooth functions  $\{K_e^{\varepsilon, n}\}_{0 \leq n \leq N}$  have the following properties:

1. the function  $K_e^{\varepsilon, n}(z)$  is supported in  $C_1 2^{-n} \leq \|z\|_{\mathfrak{s}} \leq C_2 2^{-n}$  for some  $0 < C_1 < C_2$ ;
2. for any  $q \geq 0$  and for some  $C > 0$ , independent of  $\varepsilon$  and  $\mathfrak{e}$ , one has

$$|D^k K_e^{\varepsilon, n}(z)| \leq C 2^{n(a_e + |k|_{\mathfrak{s}})}, \quad (4.5)$$

uniformly in  $z$ ,  $|k|_{\mathfrak{s}} \leq q$  and  $0 \leq n \leq N$ ;

<sup>1</sup> In all our applications it is sufficient to have kernels sufficiently many times differentiable. For example, we can take them to be in  $\mathcal{C}_{\mathfrak{s}}^q$  for  $q = \sum_{e \in \mathbb{E}} (|r_e| + 2)$ .

3. if  $r_e < 0$ , then for all  $0 \leq n \leq N$  and  $|k|_s < |r_e|$  one has

$$\int_{\mathbf{R}^{d+1}} z^k K_e^{\varepsilon, n}(z) \mu_{e+}^{\varepsilon}(dz) = 0. \quad (4.6)$$

The necessity to introduce a new parameter  $\varepsilon$  can be seen in our application in Sect. 5, where the mesh size of the grid is  $\varepsilon$  and the interaction range is defined on the scale  $\varepsilon^{\frac{3}{4}} \geq \varepsilon$ .

We see from (4.5) that the value  $a_e$  characterizes the order of singularity of the kernel. Moreover, the value  $r_e$ , assigned to an edge  $e \in \mathbb{E}$ , describes a renormalisation of the singularity, which for positive and negative values are defined in different ways in the following section.

**Lemma 4.2** *If Assumption 4 is satisfied, then for any  $q \geq 0$ , the following quantity is bounded uniformly in  $\varepsilon \in [\varepsilon, 1]$  and  $\varepsilon \in (0, 1]$*

$$\|K_e^{\varepsilon}\|_{a_e; q}^{(\varepsilon)} := \sup_{z \in \mathbf{R}^{d+1}} \sup_{|k|_s < q} (|z|_s + \varepsilon)^{a_e + |k|_s} |D^k K_e^{\varepsilon}(z)|. \quad (4.7)$$

*The reverse statement is also true, i.e. if for a kernel  $K_e^{\varepsilon}$  the quantity (4.7) is bounded uniformly, then it has all the properties listen in Assumption 4.*

**Proof** The bound (4.7) is a direct consequence of Assumption 4(1)–(2). The second part of the lemma follows by repeating the proof of [29, Lemma 5.4].  $\square$

#### 4.1.1 Renormalisation

If  $r_e \neq 0$ , then the kernel corresponding to the edge  $e$  requires renormalisation. For positive and negative values of  $r_e$  the renormalisation is defined different. For  $r_e > 0$  the renormalisation of the smooth kernel is required to get a sufficiently fast decay of the kernel at the origin. In the case  $r_e < 0$  the renormalisation is required to make the kernel, with a very strong singularity at the origin, integrable.

In the case  $r_e > 0$ , we define the renormalised kernel

$$\widehat{K}_e^{\varepsilon}(z_{e-}, z_{e+}) := K_e^{\varepsilon}(z_{e+} - z_{e-}) - \sum_{|k|_s < r_e} \frac{z_{e+}^k}{k!} D^k K_e^{\varepsilon}(-z_{e-}), \quad (4.8)$$

where the sum runs over all multi-indices  $k \in \mathbf{N}_0^{d+1}$  such that  $|k|_s < r_e$ . In the case  $r_e = 0$  we simply define  $\widehat{K}_e^{\varepsilon}(z_{e-}, z_{e+}) := K_e^{\varepsilon}(z_{e+} - z_{e-})$ . The positive renormalisation (4.8) allows to define kernels, which have sufficiently fast polynomial decay at the diagonal  $z_{e-} = z_{e+}$ . This is the case when  $K_e^{\varepsilon}$  is smooth with uniformly bounded derivatives.

If  $r_e < 0$ , then for a smooth and compactly supported function  $\varphi$  on  $\mathbf{R}^{d+1} \times \mathbf{R}^{d+1}$  we define the expansion

$$(T_{r_e} \varphi)(z_{e-}, z_{e+}) := \varphi(z_{e-}, z_{e+}) - \sum_{|k|_s < |r_e|} \frac{(z_{e+} - z_{e-})^k}{k!} D_2^k \varphi(z_{e-}, z_{e+}), \quad (4.9)$$

where the sum runs over all multi-indices  $k \in \mathbf{N}_0^{d+1}$  satisfying  $|k|_{\mathfrak{s}} < |r_e|$ , and where  $D_2^k$  is the multi-derivative in the second argument. Furthermore, we associate to  $K_e^\varepsilon$  the distribution

$$\widehat{K}_e^\varepsilon(\varphi) := \int_{\mathbf{R}^{d+1}} \int_{\mathbf{R}^{d+1}} K_e^\varepsilon(z_{e+} - z_{e-})(T_{r_e} \varphi)(z_{e-}, z_{e+}) \mu_{e-}^\varepsilon(dz_{e-}) \mu_{e+}^\varepsilon(dz_{e+}), \quad (4.10)$$

which is obtained from  $K_e^\varepsilon$  by subtracting delta-functions and their derivatives. Expression (4.10) is just another way to write the integral

$$\int_{\mathbf{R}^{d+1}} \int_{\mathbf{R}^{d+1}} K_e^\varepsilon(z_{e+} - z_{e-}) \varphi(z_{e-}, z_{e+}) \mu_{e-}^\varepsilon(dz_{e-}) \mu_{e+}^\varepsilon(dz_{e+}),$$

since  $\int_{\mathbf{R}^{d+1}} K_e^\varepsilon(z) z^k \mu_{e+}^\varepsilon(dz) = 0$  and the measure  $\mu_{e+}^\varepsilon$  is translation invariant.

**Example 1** In all of the applications that we have in mind, we deal with labels  $r_e$  taking values  $+1$ ,  $0$  or  $-1$ . In Sect. 5, for example, we have  $r_e = 1$  only for the tree in Sect. 5.4; we use negative renormalisation with  $r_e = -1$  only for the tree which is dealt with in (5.16). All the other edges in the trees of Sect. 5 always have  $r_e = 0$ .

Clearly, when  $r_e = 0$ , we have no transformation to do on the kernels. When  $r_e = 1$ , on the other hand, we have

$$\widehat{K}_e^\varepsilon(z_{e-}, z_{e+}) := K_e^\varepsilon(z_{e+} - z_{e-}) - K_e^\varepsilon(-z_{e-}),$$

while when  $r_e = -1$ , we get

$$\widehat{K}_e^\varepsilon(z_{e-}, z_{e+}) := K_e^\varepsilon(z_{e+} - z_{e-}) - \left( \int_{\mathbf{R}^{d+1}} K_e^\varepsilon(z_{e+} - z_{e-}) \mu_{e+}^\varepsilon(dz_{e+}) \right) \delta(z_{e+} - z_{e-}),$$

where  $\delta$  is the Dirac delta-function. Observe that positive renormalisation corresponds to subtracting the value of the kernel itself at the “base” point  $z_{e-}$ , while negative renormalisation means removing singularities at the base point.

#### 4.1.2 A generalised convolution

Let us fix a graph  $\mathbb{G} = (\mathbb{V}, \mathbb{E})$  as described above. Then for a smooth and compactly supported function  $\varphi : \mathbf{R}^{d+1} \rightarrow \mathbf{R}$ , for  $z = (t, x), \bar{z} = (\bar{t}, \bar{x}) \in \mathbf{R}^{d+1}$  and for  $\lambda \in (0, 1]$  we define its rescaling and recentering

$$\varphi_z^\lambda(z) := \lambda^{-|\mathfrak{s}|} \varphi(\lambda^{-2}(t - \bar{t}), \lambda^{-1}(x - \bar{x})). \quad (4.11)$$

For fixed  $\mathbf{z}^{\text{var}} \in (\mathbf{R}^{d+1})^{\mathbb{V}_{\text{var}}}$  we define the product measure on  $\mathbf{z} \in (\mathbf{R}^{d+1})^{\mathbb{V}_*}$

$$\mu_{\mathbb{V}_*, \mathbf{z}^{\text{var}}}^\varepsilon(d\mathbf{z}) := \left( \prod_{v \in \mathbb{V}_* \setminus \mathbb{V}_{\text{var}}} \mu_v^\varepsilon(dz_v) \right) \left( \prod_{w \in \mathbb{V}_{\text{var}}} \delta(z_w - z_w^{\text{var}}) dz_w \right), \quad (4.12)$$

where again  $\delta$  is the Dirac delta-function, and where  $\mathbf{z} = (z_v \in \mathbf{R}^{d+1} : v \in \mathbb{V}_\star)$  and  $\mathbf{z}^{\text{var}} = (z_v^{\text{var}} \in \mathbf{R}^{d+1} : v \in \mathbb{V}_{\text{var}})$ . In other words, the variable  $z_v$ , corresponding to the vertex  $v \in \mathbb{V}_\star \setminus \mathbb{V}_{\text{var}}$ , is integrated with respect to the measures  $\mu_v^\varepsilon$ , and the variables corresponding to the vertices in  $\mathbb{V}_{\text{var}}$  are fixed to be equal to  $\mathbf{z}^{\text{var}}$ . These are the variables which we want to integrate with respect to martingales. Then we define the *generalised convolution*

$$\mathcal{K}_{\mathbb{G}}^{\lambda, \varepsilon}(\mathbf{z}^{\text{var}}) := \int_{(\mathbf{R}^{d+1})^{\mathbb{V}_\star}} \left( \prod_{e \in \mathbb{E}} \widehat{K}_e^\varepsilon(z_{e-}, z_{e+}) \right) \varphi_0^\lambda(z_{v_\star^\uparrow}) \mu_{\mathbb{V}_\star, \mathbf{z}^{\text{var}}}^\varepsilon(d\mathbf{z}). \quad (4.13)$$

Since the kernels  $K_e^\varepsilon$  are smooth, our assumptions on the graph guarantee that the generalised convolution (4.13) is well-defined.

We fix any order of the elements in  $\mathbb{V}_{\text{var}}$  (which respectively fixed the order of the variables in  $\mathcal{K}_{\mathbb{G}}^{\lambda, \varepsilon}(\mathbf{z}^{\text{var}})$ ) and we define

$$(\mathcal{I}_\gamma^{\varepsilon, \mathbb{L}} \mathcal{K}_{\mathbb{G}}^{\lambda, \varepsilon})_t := \sum_{\sigma \in \Sigma_\gamma} (\mathcal{I}_{\gamma, \sigma}^{\varepsilon, \mathbb{L}} \mathcal{K}_{\mathbb{G}}^{\lambda, \varepsilon})_t, \quad (4.14)$$

where the stochastic integral  $\mathcal{I}_{\gamma, \sigma}^{\varepsilon, \mathbb{L}}$  is defined in Sect. 3.2.1 with respect to the fixed order of the variables. The following is our main result of this section.

**Theorem 4.3** *Let  $\mathbb{G} = (\mathbb{V}, \mathbb{E})$  be a graph with labels  $\{a_e, r_e\}_{e \in \mathbb{E}}$  satisfying Assumption 2, let  $\gamma \in \mathfrak{C}_m(\mathbb{V}_{\text{var}})$ , with  $1 \leq m \leq |\mathbb{V}_{\text{var}}|$ , be a contraction with a labeling  $\mathbb{L}$  such that Assumption 3 is satisfied. Let the measures be defined as in (4.4) and let the kernels satisfy Assumption 4. Let furthermore  $\mathcal{I}_\gamma^{\varepsilon, \mathbb{L}}$  be a stochastic integral with respect to càdlàg martingales satisfying Assumption 1, let the set  $\Gamma$  be defined in (4.3), and let*

$$\nu_\gamma := |\mathfrak{s}| |\widehat{\mathbb{V}}_\star \setminus \{\widehat{v}_\star^\uparrow\}| - \frac{|\mathfrak{s}|}{2} |\Gamma| - \sum_{e \in \widehat{\mathbb{E}}} \widehat{a}_e < 0. \quad (4.15)$$

*Then for any  $p \geq 2$  and  $\theta > 0$  there is a constant  $C$  for which the following bound holds*

$$\begin{aligned} \left( E \left[ \sup_{t \in \mathbb{R}_+} |(\mathcal{I}_\gamma^{\varepsilon, \mathbb{L}} \mathcal{K}_{\mathbb{G}}^{\lambda, \varepsilon})_t|^p \right] \right)^{\frac{1}{p}} &\leq C \lambda^{\nu_\gamma} \sum_{\substack{\mathbf{p} \in \mathcal{P}_m : \mathbf{p}^{-1}(1) \cap \Gamma = \emptyset, \\ \mathbb{L}^{-1}(\nabla) \subseteq \mathbf{p}^{-1}(1), \mathbf{p}^{-1}(\infty) = \emptyset}} \varepsilon^{\alpha_\gamma(\mathbf{p})} \mathfrak{e}^{-\delta_\gamma(\mathbf{p})} \\ &+ C \lambda^{\nu_\gamma} \sum_{\substack{\mathbf{p} \in \mathcal{P}_m : \mathbf{p}^{-1}(1) \cap \Gamma = \emptyset, \\ \mathbb{L}^{-1}(\nabla) \subseteq \mathbf{p}^{-1}(1), \mathbf{p}^{-1}(\infty) \neq \emptyset}} \varepsilon^{\alpha_\gamma(\mathbf{p}) - \theta} \mathfrak{e}^{-\delta_\gamma(\mathbf{p})} \end{aligned} \quad (4.16)$$

uniformly in  $\lambda \in [\epsilon, 1]$ ,  $\epsilon \in [\varepsilon, 1]$  and  $\varepsilon \in (0, 1]$ , where the set of functions  $\mathcal{P}_m$  is defined in Sect. 3.2, the constant  $\alpha_\gamma(\mathbf{p})$  is defined in (3.19), and

$$\delta_\gamma(\mathbf{p}) := \frac{|\mathfrak{s}|}{2} \left( 2|\mathbf{p}^{-1}(\infty) \setminus \Gamma| + |\mathbf{p}^{-1}(\infty) \cap \Gamma| + |\mathbf{p}^{-1}(2) \setminus \Gamma| \right). \quad (4.17)$$

We prove this theorem in Sect. 4.4, and before that we need to get some preliminary results.

**Remark 4.4** From the proof of Theorem 4.3 we can see that there exists a value  $q \geq 0$  and a compact set  $\mathfrak{K} \subseteq \mathbf{R}^{d+1}$ , such that the constant  $C$  in (4.16) is proportional to

$$\left( \prod_{e \in \mathbb{E}} \|K_e^\varepsilon\|_{a_e; q}^{(\epsilon)} \right) \left( \prod_{v \notin \mathbb{V}_{\text{var}} \cup \{\star\}} \|\mu_v^\varepsilon\|_{\text{TV}(\mathfrak{K})} \right),$$

which by our assumptions is bounded uniformly in  $\varepsilon$  and  $\epsilon$ . For example, we can take a very rough value  $q = \sum_{e \in \mathbb{E}} (|r_e| + 2)$ .

If we would like to consider a recentered test function  $\varphi_{\bar{z}}^\lambda(z)$ , we need to shift respectively all the variables in the generalised convolution:

$$\mathcal{K}_{\mathbb{G}, \bar{z}}^{\lambda, \varepsilon}(\mathbf{z}^{\text{var}}) := \int_{(\mathbf{R}^{d+1})^{\mathbb{V}_\star}} \left( \prod_{e \in \mathbb{E}} \widehat{K}_e^\varepsilon(z_{e_-} - \bar{z}, z_{e_+} - \bar{z}) \right) \varphi_{\bar{z}}^\lambda(z_{v_\star^\uparrow}) \mu_{\mathbb{V}_\star, \mathbf{z}^{\text{var}}}^\varepsilon(d\mathbf{z}). \quad (4.18)$$

Then the following result can be proved as Theorem 4.3, by changing the value of the variable  $z_\star$  from 0 to  $\bar{z}$ . Uniformity in  $\bar{z}$  holds, because the norms of the kernels (4.30) are independent of this variable.

**Corollary 4.5** *Under the assumptions of Theorem 4.3, the bound (4.16) holds for the multiple integral  $\mathcal{I}_\gamma^{\varepsilon, L} \mathcal{K}_{\mathbb{G}, \bar{z}}^{\lambda, \varepsilon}$ , locally uniformly in  $\bar{z}$ .*

Applying Minkowski inequality, we get from the definition (4.14) the bound

$$\mathbf{E}_p \sup_{t \in \mathbf{R}_+} |(\mathcal{I}_\gamma^{\varepsilon, L} \mathcal{K}_{\mathbb{G}}^{\lambda, \varepsilon})_t| \leq \sum_{\sigma \in \Sigma_\gamma} \mathbf{E}_p \sup_{t \in \mathbf{R}_+} |(\mathcal{I}_{\gamma, \sigma}^{\varepsilon, L} \mathcal{K}_{\mathbb{G}}^{\lambda, \varepsilon})_t|. \quad (4.19)$$

In the rest of the section we are going to prove the bound (4.16) for the integral  $\mathcal{I}_{\gamma, \sigma}^{\varepsilon, L}$  with a fixed  $\sigma$ . One can see from the proof, that this bound is independent of the order of the variables (although the order plays a role in some intermediate results like Lemma 4.7), and the same bound (4.16) holds for every integral  $\mathcal{I}_{\gamma, \sigma}^{\varepsilon, L}$  in (4.19).

## 4.2 Multiscale decomposition of the generalised convolution

Our aim is to write the kernels  $K_e^\varepsilon$  in the generalised convolution (4.13) as sums of localised functions. For the edge  $(\star, v_\star^\uparrow)$ , we view the test function  $\varphi_0^\lambda(z_{v_\star^\uparrow})$  in (4.13) as a

new kernel  $K_{(\star, v_\star^\uparrow)}(z_{v_\star^\uparrow})$ , supported on  $\|z_{v_\star^\uparrow}\|_{\mathfrak{s}} \lesssim \lambda$  and satisfying  $\|K_{(\star, v_\star^\uparrow)}\|_{0; q} \lesssim \lambda^{-|\mathfrak{s}|}$  (recall that this edge has the labels  $a_e = r_e = 0$  in the graph).

Our next aim is to decompose the kernels in (4.13) into sums of localised functions. To this end, for  $e \in \mathbb{E}$  with  $r_e > 0$ , we take any smooth functions  $\psi^{(\varepsilon, n)} : \mathbf{R}^{d+1} \rightarrow \mathbf{R}$ , such that  $\psi^{(\varepsilon, n)}(z)$  is supported in  $C_1 2^{-n} \leq \|z\|_{\mathfrak{s}} \leq C_2 2^{-n}$  (where  $C_1, C_2$  are from Assumption 4), scales as  $2^{-n}$  and satisfies  $\sum_{n=0}^N \psi^{(\varepsilon, n)}(z) = 1$  for all  $z$ . Let us denote for convenience  $\mathbf{N}_{\leq N} := \{0, 1, \dots, N\}$ . Then for  $r_e > 0$  and  $\mathbf{n} = (k, p, m) \in \mathbf{N}_{\leq N}^3$  we set

$$\widehat{K}_e^{\varepsilon, \mathbf{n}}(z, \bar{z}) := \psi^{(\varepsilon, k)}(\bar{z} - z) \psi^{(\varepsilon, p)}(z) \psi^{(\varepsilon, m)}(\bar{z}) \widehat{K}_e^{\varepsilon}(z, \bar{z}), \quad (4.20)$$

where the kernel  $\widehat{K}_e^{\varepsilon}$  has been defined in (4.8). For  $\mathbf{n} \in \mathbf{N}_{\leq N}^3$  and  $e \in \mathbb{E}$  such that  $r_e \leq 0$ , we define the function

$$\widehat{K}_e^{\varepsilon, \mathbf{n}}(z, \bar{z}) := \begin{cases} K_e^{\varepsilon, k}(\bar{z} - z), & \text{if } \mathbf{n} = (k, 0, 0), \ 0 \leq k \leq N, \\ 0, & \text{otherwise,} \end{cases}$$

where we made use of the expansion of the kernel from Assumption 4.

For  $\lambda \in (0, 1]$  we define the set  $\mathcal{N}_\lambda^{\mathfrak{e}}$  of functions  $\mathbf{n} : \mathbb{E} \rightarrow \mathbf{N}_{\leq N}^3$  satisfying  $2^{-|\mathbf{n}_{(\star, v_\star^\uparrow)}|} \leq \lambda \vee \mathfrak{e}$ , with  $\mathbf{n}_{(\star, v_\star^\uparrow)}$  being the evaluation of the function  $\mathbf{n}$  on the edge  $(\star, v_\star^\uparrow)$ . Then for a function  $\mathbf{n} \in \mathcal{N}_\lambda^{\mathfrak{e}}$  and a point  $\mathbf{z} = (z_v : v \in \mathbb{V}_\star)$ , we define

$$\widehat{K}^{\varepsilon, \mathbf{n}}(\mathbf{z}) := \prod_{e \in \mathbb{E}} \widehat{K}_e^{\varepsilon, \mathbf{n}_e}(z_{e_-}, z_{e_+}), \quad (4.21)$$

where  $z_\star = 0$ . Since the functions  $\psi^{(\varepsilon, n)}$  sum up to 1 and since we consider the test function  $\varphi_0^\lambda$  as a kernel, one can rewrite the generalised convolution (4.13) as

$$\mathcal{K}_{\mathbb{G}}^{\lambda, \varepsilon}(\mathbf{z}^{\text{var}}) := \sum_{\mathbf{n} \in \mathcal{N}_\lambda^{\mathfrak{e}}} \mathcal{K}_{\mathbb{G}}^{\varepsilon, \mathbf{n}}(\mathbf{z}^{\text{var}}), \quad \mathcal{K}_{\mathbb{G}}^{\varepsilon, \mathbf{n}}(\mathbf{z}^{\text{var}}) := \int_{(\mathbf{R}^{d+1})^{\mathbb{V}_\star}} \widehat{K}^{\varepsilon, \mathbf{n}}(\mathbf{z}) \mu_{\mathbb{V}_\star, \mathbf{z}^{\text{var}}}^{\varepsilon} d\mathbf{z}. \quad (4.22)$$

Since we are interested in estimating the integrals  $\mathcal{I}_{\gamma, \sigma}^{\varepsilon, \mathbb{L}} \mathcal{K}_{\mathbb{G}}^{\lambda, \varepsilon}$ , we can exploit the fact that the integration variables  $z_v$  in the kernel (4.22), for vertices  $v$  belonging to the same component of  $\gamma$ , are equal. More precisely, we define the set  $\mathcal{N}_{\lambda, \gamma}^{\mathfrak{e}}$  in the same way as  $\mathcal{N}_\lambda^{\mathfrak{e}}$ , but using the contracted graph  $\mathbb{G}_\gamma = (\tilde{\mathbb{V}}, \tilde{\mathbb{E}})$ . Then for a function  $\mathbf{n} \in \mathcal{N}_{\lambda, \gamma}^{\mathfrak{e}}$  and a point  $\mathbf{z} = (z_v : v \in \tilde{\mathbb{V}}_\star)$ , we define the kernel  $\widehat{K}^{\varepsilon, \mathbf{n}}(\mathbf{z})$  as in (4.21), but with the product over  $\tilde{\mathbb{E}}$ . Furthermore, we define the measure  $\mu_{\tilde{\mathbb{V}}_\star, \mathbf{z}^{\text{var}}}^{\varepsilon}$  on  $(\mathbf{R}^{d+1})^{\tilde{\mathbb{V}}_\star}$  by

$$\mu_{\tilde{\mathbb{V}}_\star, \mathbf{z}^{\text{var}}}^{\varepsilon}(d\mathbf{z}) := \left( \prod_{v \in \tilde{\mathbb{V}}_\star \setminus \tilde{\mathbb{V}}_{\text{var}}} \mu_v^{\varepsilon}(dz_v) \right) \left( \prod_{w \in \tilde{\mathbb{V}}_{\text{var}}} \delta_{z_w - z_{w*}^{\text{var}}} dz_w \right), \quad (4.23)$$

where  $w_*$  is the first element (with respect to a chosen order of vertices) in  $i_\gamma^{-1}(w)$ , and the map  $i_\gamma$  has been introduced in the beginning of this section. In other words, this measure identifies the variables in  $\tilde{\mathbb{V}}_{\text{var}}$  which correspond to the same component of  $\gamma$ . Then we define the kernel

$$K_{\mathbb{G}_\gamma}^{\varepsilon, \mathbf{n}}(\mathbf{z}^{\text{var}}) := \int_{(\mathbf{R}^{d+1})^{\tilde{\mathbb{V}}_*}} \widehat{K}^{\varepsilon, \mathbf{n}}(\mathbf{z}) \mu_{\tilde{\mathbb{V}}_*, \mathbf{z}^{\text{var}}}^\varepsilon(d\mathbf{z}), \quad (4.24)$$

and write the multiple stochastic integral as  $\mathcal{I}_{\gamma, \sigma}^{\varepsilon, \mathbb{L}} \mathcal{K}_{\mathbb{G}}^{\lambda, \varepsilon} = \sum_{\mathbf{n} \in \mathcal{N}_{\lambda, \gamma}^\varepsilon} \mathcal{I}_{\gamma, \sigma}^{\varepsilon, \mathbb{L}} \mathcal{K}_{\mathbb{G}_\gamma}^{\varepsilon, \mathbf{n}}$ . Using this expansion and applying Minkowski's inequality, we obtain the bound

$$\mathbf{E}_p \sup_{t \in \mathbf{R}_+} |(\mathcal{I}_{\gamma, \sigma}^{\varepsilon, \mathbb{L}} \mathcal{K}_{\mathbb{G}}^{\lambda, \varepsilon})_t| \leq \sum_{\mathbf{n} \in \mathcal{N}_{\lambda, \gamma}^\varepsilon} \mathbf{E}_p \sup_{t \in \mathbf{R}_+} |(\mathcal{I}_{\gamma, \sigma}^{\varepsilon, \mathbb{L}} \mathcal{K}_{\mathbb{G}_\gamma}^{\varepsilon, \mathbf{n}})_t|. \quad (4.25)$$

Bounding a multiple integral of the generalised convolution boils down to bounding integrals in (4.25) and summing over the functions  $\mathbf{n} \in \mathcal{N}_{\lambda, \gamma}^\varepsilon$ . This is what we do in the next sections, where, following the idea of [30, Appendix A.2], we use a multiscale clustering in the sum over  $\mathbf{n}$ .

### 4.3 Bounds on iterated integrals

We associate to every point  $\mathbf{z} \in (\mathbf{R}^{d+1})^{\tilde{\mathbb{V}}}$  a rooted labelled binary tree  $(T, \ell)$ , such that  $\|z_v - z_w\|_{\mathfrak{s}} \sim 2^{-\ell_{v \wedge w}}$  and  $\ell_{v \wedge w} \in \mathbf{N}_{\leq N}$ , where  $v \wedge w$  is the closest common ancestor of  $v$  and  $w$ . Moreover, the labels  $\ell$  satisfy  $\ell_v \geq \ell_\omega$  whenever  $v \geq \omega$ , where  $v \geq \omega$  means that  $\omega$  belongs to the shortest path from  $v$  to the root of the tree  $T$ . See [30, Appendix A.2] for construction of such tree and also for the terminology which we are going to use. Given a set of vertices  $\tilde{\mathbb{V}}$ , we denote by  $\mathbb{T}^\varepsilon(\tilde{\mathbb{V}})$  the set of rooted labelled binary trees  $(T, \ell)$  as above, which have  $\tilde{\mathbb{V}}$  as their set of leaves. Denote furthermore by  $\mathbb{T}_\lambda^\varepsilon(\tilde{\mathbb{V}})$  the subset of those labelled trees in  $\mathbb{T}^\varepsilon(\tilde{\mathbb{V}})$  with the property that  $2^{-\ell_{* \wedge v_*}} \leq \lambda$ .

Our next aim is to write summation in (4.25) over such labelled trees  $(T, \ell)$  and then over those functions  $\mathbf{n}$  which are close in some sense to the labeling  $\ell$ . To this end, for the constant,<sup>2</sup>  $c := (\log_2 |\tilde{\mathbb{V}}| + |\log_2 C_1|) \vee |\log_2 C_2|$ , where the constants  $C_1, C_2$  are from Assumption 4, we define the set  $\mathcal{N}_\gamma^\varepsilon(T, \ell)$  consisting of all functions  $\mathbf{n}: \tilde{\mathbb{E}} \rightarrow \mathbf{N}_{\leq N}^3$  such that

1. for every edge  $e = (v, w)$  with  $r_e \leq 0$ , one has  $\mathbf{n}_e = (k, 0, 0) \in \mathbf{N}_{\leq N}^3$  with  $|k - \ell_{v \wedge w}| \leq c$ ,
2. for every edge  $e = (v, w)$  with  $r_e > 0$ , one has  $\mathbf{n}_e = (k, p, m) \in \mathbf{N}_{\leq N}^3$  with  $|k - \ell_{v \wedge w}| \leq c$ ,  $|p - \ell_{v \wedge *} \leq c$ , and  $|m - \ell_{w \wedge *} \leq c$ .

<sup>2</sup> Our value of  $c$  is different from the analogous value in [30, Definition A.8] because the kernels  $K_e^{\varepsilon, n}$  from Assumption 4 have a different support. The need to define  $c$  in this way can be seen from the proof of [30, Lemma A.9].

Then we have the following analogue of [30, Lemma A.9], which is proved in exactly the same way.

**Lemma 4.6** *Let us fix a point  $z^{\text{var}} \in (\mathbf{R}^{d+1})^{\mathbb{V}_{\text{var}}}$ . Let  $\mathbf{n}: \tilde{\mathbb{E}} \rightarrow \mathbf{N}_{\leq N}^3$  be such that the kernel  $\mathcal{K}_{\mathbb{G}_{\gamma}}^{\varepsilon, \mathbf{n}}(z^{\text{var}})$  defined in (4.24) does not vanish. Then there exists a labelled tree  $(T, \ell) \in \mathbb{T}_{\lambda}^{\mathbf{e}}(\tilde{\mathbb{V}})$  such that  $\mathbf{n} \in \mathcal{N}_{\gamma}^{\mathbf{e}}(T, \ell)$ .*

Using this result, the right-hand side of (4.25) can be estimated as

$$\mathbf{E}_p \sup_{t \in \mathbf{R}_+} |(\mathcal{I}_{\gamma, \sigma}^{\varepsilon, \mathbf{L}} \mathcal{K}_{\mathbb{G}}^{\lambda, \varepsilon})_t| \leq \sum_{(T, \ell) \in \mathbb{T}_{\lambda}^{\mathbf{e}}(\tilde{\mathbb{V}})} \sum_{\mathbf{n} \in \mathcal{N}_{\gamma}^{\mathbf{e}}(T, \ell)} \mathbf{E}_p \sup_{t \in \mathbf{R}_+} |(\mathcal{I}_{\gamma, \sigma}^{\varepsilon, \mathbf{L}} \mathcal{K}_{\mathbb{G}_{\gamma}}^{\varepsilon, \mathbf{n}})_t|. \quad (4.26)$$

We will now modify the kernels in (4.21) in the same way how it was done in [30, Appendix A.5]. Let  $A^- \subseteq \tilde{\mathbb{E}}$  contain those edges  $e = (e_-, e_+)$  which have the label  $r_e < 0$ , and for which any two vertices  $\{u, v\}$  satisfying  $u \wedge v = e_- \wedge e_+$  coincide with  $\{e_-, e_+\}$ . Then we can factorize (4.21) as

$$\widehat{K}^{\varepsilon, \mathbf{n}}(\mathbf{z}) = \widehat{G}^{\mathbf{e}, \mathbf{n}}(\mathbf{z}) \left( \prod_{e \in A^-} \widehat{K}_e^{\varepsilon, \mathbf{n}_e}(z_{e_-}, z_{e_+}) \right), \quad \widehat{G}^{\mathbf{e}, \mathbf{n}}(\mathbf{z}) := \prod_{e \notin A^-} \widehat{K}_e^{\varepsilon, \mathbf{n}_e}(z_{e_-}, z_{e_+}). \quad (4.27)$$

For  $e = (e_-, e_+)$  and  $r > 0$  we define the operator  $\mathcal{Y}_e^r$  acting on sufficiently smooth functions  $V: (\mathbf{R}^{d+1})^{\tilde{\mathbb{V}}} \rightarrow \mathbf{R}$  as

$$(\mathcal{Y}_e^r V)(\mathbf{z}) := V(\mathbf{z}) - \sum_{|k|_{\mathfrak{s}} < r} \frac{(z_{e_+} - z_{e_-})^k}{k!} (D_{e_+}^k V)(P_e(\mathbf{z})),$$

where  $D_{e_+}^k$  is a derivative with respect to  $z_{e_+}$  and where  $(P_e(\mathbf{z}))_v = z_v$  if  $v \neq e_+$  and  $(P_e(\mathbf{z}))_v = z_{e_-}$  if  $v = e_+$ . Furthermore, writing  $A^- = \{e^{(1)}, \dots, e^{(k)}\}$  for some  $k \geq 0$ , we define the kernel

$$\tilde{K}^{\varepsilon, \mathbf{n}}(\mathbf{z}) := \left( \mathcal{Y}_{e^{(k)}}^{r_{e^{(k)}}} \cdots \mathcal{Y}_{e^{(1)}}^{r_{e^{(1)}}} \widehat{G}^{\mathbf{e}, \mathbf{n}}(\mathbf{z}) \right) \left( \prod_{e \in A^-} \widehat{K}_e^{\varepsilon, \mathbf{n}_e}(z_{e_-}, z_{e_+}) \right). \quad (4.28)$$

Then for every  $\mathbf{z}^{\text{var}}$  we have

$$\int_{(\mathbf{R}^{d+1})^{\tilde{\mathbb{V}}_*}} \widehat{K}^{\varepsilon, \mathbf{n}}(\mathbf{z}) \mu_{\tilde{\mathbb{V}}_*, \mathbf{z}^{\text{var}}}^{\varepsilon} (d\mathbf{z}) = \int_{(\mathbf{R}^{d+1})^{\tilde{\mathbb{V}}_*}} \tilde{K}^{\varepsilon, \mathbf{n}}(\mathbf{z}) \mu_{\tilde{\mathbb{V}}_*, \mathbf{z}^{\text{var}}}^{\varepsilon} (d\mathbf{z}), \quad (4.29)$$

which is just a reformulation of the argument below [30, Equation A.26] in our context. Then (4.26) can be written as

$$\mathbf{E}_p \sup_{t \in \mathbf{R}_+} |(\mathcal{I}_{\gamma, \sigma}^{\varepsilon, \mathbf{L}} \mathcal{K}_{\mathbb{G}}^{\lambda, \varepsilon})_t| \leq \sum_{(T, \ell) \in \mathbb{T}_{\lambda}^{\mathbf{e}}(\tilde{\mathbb{V}})} \sum_{\mathbf{n} \in \mathcal{N}_{\gamma}^{\mathbf{e}}(T, \ell)} \mathbf{E}_p \sup_{t \in \mathbf{R}_+} |(\mathcal{I}_{\gamma, \sigma}^{\varepsilon, \mathbf{L}} \tilde{K}_{\mathbb{G}_{\gamma}}^{\varepsilon, \mathbf{n}})_t|, \quad (4.30)$$

with the new kernels

$$\tilde{\mathcal{K}}_{\mathbb{G}_\gamma}^{\varepsilon, \mathbf{n}}(\mathbf{z}^{\text{var}}) := \int_{(\mathbf{R}^{d+1})^{\tilde{\mathbb{V}}_\star}} \tilde{K}^{\varepsilon, \mathbf{n}}(\mathbf{z}) \mu_{\tilde{\mathbb{V}}_\star, \mathbf{z}^{\text{var}}}^\varepsilon(d\mathbf{z}). \quad (4.31)$$

Using the notation (2.18), we denote with  $(\tilde{\mathcal{K}}_{\mathbb{G}_\gamma}^{\varepsilon, \mathbf{n}})^{\gamma, \sigma} : (\mathbf{R}^{d+1})^{\tilde{\mathbb{V}}_\star} \rightarrow \mathbf{R}$  the kernel which is obtained from  $\tilde{\mathcal{K}}_{\mathbb{G}_\gamma}^{\varepsilon, \mathbf{n}}$  by making all the variables from the same component of  $\gamma$  equal. We would like to apply Theorem 3.9 to bound the stochastic integrals in (4.30). For this, we need to estimate the norms (3.16) of the kernel  $(\tilde{\mathcal{K}}_{\mathbb{G}_\gamma}^{\varepsilon, \mathbf{n}})^{\gamma, \sigma}(\mathbf{z}^{\text{var}})$ , which is what we are going to do now.

Let  $T^\circ$  denote the set of interior nodes of the tree  $T$ . Then for  $e \in \widehat{E}$  let us define the function  $\eta_e : T^\circ \rightarrow \mathbf{R}$  by

$$\begin{aligned} \eta_e(v) := & -\widehat{a}_e \mathbb{1}_{e_\uparrow}(v) + r_e (\mathbb{1}_{e_+ \wedge \star}(v) - \mathbb{1}_{e_\uparrow}(v)) \mathbb{1}_{r_e > 0, e_+ \wedge \star > e_\uparrow} \\ & + (1 - r_e - \widehat{a}_e) (\mathbb{1}_{e_- \wedge \star}(v) - \mathbb{1}_{e_\uparrow}(v)) \mathbb{1}_{r_e > 0, e_- \wedge \star > e_\uparrow}, \end{aligned}$$

where  $\mathbb{1}_v(w) := \mathbb{1}_{v=w}$  and  $e_\uparrow := e_- \wedge e_+ \in T^\circ$  for an edge  $e = (e_-, e_+) \in \widehat{E}$ . The function  $\eta_e$  coincides with the one defined in [30, Equation A.20] and is used to bound the generalised convolution without taking into account negative renormalisation. To consider negative renormalisation we define by analogy with [30, Equation A.27] a modified function

$$\tilde{\eta}(v) := |\mathfrak{s}| + \sum_{e \in \widehat{E}} \tilde{\eta}_e(v), \quad \tilde{\eta}_e(v) := \eta_e(v) - r_e \mathbb{1}_{e \in A^-} (\mathbb{1}_{e_\uparrow}(v) - \mathbb{1}_{e_{\uparrow\uparrow}}(v)), \quad (4.32)$$

where the interior node  $e_{\uparrow\uparrow} \in T^\circ$  is of the form  $w \wedge e_-$  with  $w \notin e$  which is the furthest from the root.

From the fixed order  $\sigma$  of the variables in (4.14) we obtain an order of the vertices in  $\tilde{\mathbb{V}}_\star$ . Then we write  $\tilde{\mathbb{V}}_\star = (v_1, \dots, v_m)$  according to this order. For every  $v \in \tilde{\mathbb{V}}_\star$  we denote by  $v^\rightarrow$  the element following after  $v$  with respect to this order, and in case when there is no following element we define  $v^\rightarrow = \star$ . For  $A \subseteq \tilde{\mathbb{V}}_\star$ , let  $T_A$  be the subtree of  $T$  containing all the leaves  $v$  and  $v^\rightarrow$ , for  $v \in A \sqcup \{\star\}$ , and all the inner nodes  $v \wedge v^\rightarrow$  for such  $v$ . Let  $T_A^\circ$  contain the inner nodes of  $T_A$ .

Then we have the following bound on the norms (3.16) of the kernels  $(\tilde{\mathcal{K}}_{\mathbb{G}_\gamma}^{\varepsilon, \mathbf{n}})^{\gamma, \sigma}$ .

**Lemma 4.7** *In the setting of Theorem 4.3, let  $\mathbf{p}$  be one of the functions in the sum in (4.16). Then for any labeled tree  $(T, \ell) \in \mathbb{T}_\lambda^\mathbf{e}(\tilde{\mathbb{V}})$  there is a constant  $C$  such that for every  $\mathbf{n} \in \mathcal{N}_\gamma^\mathbf{e}(T, \ell)$  one has the bound*

$$\|(\tilde{\mathcal{K}}_{\mathbb{G}_\gamma}^{\varepsilon, \mathbf{n}})^{\gamma, \sigma}\|_{L_\varepsilon^p} \leq C \mathfrak{e}^{-\delta_\gamma(\mathbf{p})} \left( \prod_{v \in T^\circ} 2^{-\ell_v \tilde{\eta}(v)} \right) \left( \prod_{v \in T_\Gamma^\circ} 2^{\ell_v |\mathfrak{s}|} \right)^{\frac{1}{2}}, \quad (4.33)$$

where we use the norm (3.16) and the constant  $\delta_\gamma(\mathbf{p})$  defined in (4.17).

**Proof** We are going to prove a more general result; namely, we will prove a bound on the norm of the kernel  $(\tilde{\mathcal{K}}_{\mathbb{G}_\gamma}^{\varepsilon, \mathbf{n}})^{\gamma, \sigma}$  some of whose variables are fixed. For this, we take  $0 \leq M < m$  and the set  $D = \{v_{M+1}, \dots, v_m\} \subseteq \tilde{\mathbb{V}}_{\text{var}}$  of vertices and we will fix the values of the variables corresponding to these vertices. More precisely, for  $\mathbf{z}_D \in (\mathbf{R}^{d+1})^D$  we write  $(\tilde{\mathcal{K}}_{\mathbb{G}_\gamma}^{\varepsilon, \mathbf{n}})^{\gamma, \sigma}|_{\mathbf{z}_D}$  for the function from  $(\mathbf{R}^{d+1})^{\tilde{\mathbb{V}}_{\text{var}} \setminus D}$  to  $\mathbf{R}$ , which is obtained from  $(\tilde{\mathcal{K}}_{\mathbb{G}_\gamma}^{\varepsilon, \mathbf{n}})^{\gamma, \sigma}(\mathbf{z}^{\text{var}})$  by fixing the values of the variables  $z_v^{\text{var}}$  with  $v \in D$ . We extend this definition for  $D = \emptyset$  (which corresponds to  $M = m$ ) by  $(\tilde{\mathcal{K}}_{\mathbb{G}_\gamma}^{\varepsilon, \mathbf{n}})^{\gamma, \sigma}|_{\mathbf{z}_\emptyset} = (\tilde{\mathcal{K}}_{\mathbb{G}_\gamma}^{\varepsilon, \mathbf{n}})^{\gamma, \sigma}$ .

For  $1 \leq M \leq m$  and for a function  $\mathbf{p} \in \mathcal{P}_M$ , we are going to prove the bound

$$\|(\tilde{\mathcal{K}}_{\mathbb{G}_\gamma}^{\varepsilon, \mathbf{n}})^{\gamma, \sigma}|_{\mathbf{z}_D}\|_{L_\varepsilon^{\mathbf{p}}} \leq C \left( \prod_{v \in T^\circ} 2^{-\ell_v \tilde{\eta}(v)} \right) \left( \prod_{v \in T_{D \cup \mathbf{p}^{-1}(\infty)}^\circ} 2^{\ell_v |\mathfrak{s}|} \right) \left( \prod_{v \in T_{\mathbf{p}^{-1}(2)}^\circ} 2^{\ell_v |\mathfrak{s}|} \right)^{\frac{1}{2}} \quad (4.34)$$

uniformly in  $\mathbf{z}_D \in (\mathbf{R}^{d+1})^D$ . Moreover, we will show that for  $M = 0$  (in which case  $D = \tilde{\mathbb{V}}_{\text{var}}$ ) the same bound holds for the absolute value of  $(\tilde{\mathcal{K}}_{\mathbb{G}_\gamma}^{\varepsilon, \mathbf{n}})^{\gamma, \sigma}|_{\mathbf{z}_D}$ .

We can see that the bound (4.33) follows from (4.34) in the particular case  $M = m$  corresponding to  $D = \emptyset$ . To see it, we note that (4.34) simplifies to

$$\|(\tilde{\mathcal{K}}_{\mathbb{G}_\gamma}^{\varepsilon, \mathbf{n}})^{\gamma, \sigma}\|_{L_\varepsilon^{\mathbf{p}}} \leq C \left( \prod_{v \in T^\circ} 2^{-\ell_v \tilde{\eta}(v)} \right) \left( \prod_{v \in T_{\mathbf{p}^{-1}(\infty)}^\circ} 2^{\ell_v |\mathfrak{s}|} \right) \left( \prod_{v \in T_{\mathbf{p}^{-1}(2)}^\circ} 2^{\ell_v |\mathfrak{s}|} \right)^{\frac{1}{2}}. \quad (4.35)$$

Our next goal is to replace the product over  $v \in T_{\mathbf{p}^{-1}(2)}^\circ$  by the product over  $v \in T_\Gamma^\circ$ . We do it by noting that  $\tilde{\mathbb{V}}_{\text{var}} = \mathbf{p}^{-1}(1) \sqcup \mathbf{p}^{-1}(2) \sqcup \mathbf{p}^{-1}(\infty)$  and using simple operations on the sets. Namely, we have  $\mathbf{p}^{-1}(2) = (\Gamma \sqcup (\mathbf{p}^{-1}(2) \setminus \Gamma)) \setminus (\mathbf{p}^{-1}(\infty) \cap \Gamma)$ , where we used the assumption  $\mathbf{p}^{-1}(1) \cap \Gamma = \emptyset$  in (4.16). Then we write the products over  $v \in T_{\mathbf{p}^{-1}(2)}^\circ$  as

$$\prod_{v \in T_{\mathbf{p}^{-1}(2)}^\circ} 2^{\ell_v |\mathfrak{s}|} = \left( \prod_{v \in T_\Gamma^\circ} 2^{\ell_v |\mathfrak{s}|} \right) \left( \prod_{v \in T_{\mathbf{p}^{-1}(2)}^\circ \setminus T_\Gamma^\circ} 2^{\ell_v |\mathfrak{s}|} \right) \left( \prod_{v \in T_{\mathbf{p}^{-1}(\infty)}^\circ \cap T_\Gamma^\circ} 2^{-\ell_v |\mathfrak{s}|} \right).$$

Hence, the product on the right-hand side of (4.35) equals

$$\begin{aligned} & \left( \prod_{v \in T^\circ} 2^{-\ell_v \tilde{\eta}(v)} \right) \left( \prod_{v \in T_\Gamma^\circ} 2^{\ell_v |\mathfrak{s}|} \right)^{\frac{1}{2}} \left( \prod_{v \in T_{\mathbf{p}^{-1}(\infty)}^\circ} 2^{\ell_v |\mathfrak{s}|} \right) \left( \prod_{v \in T_{\mathbf{p}^{-1}(2)}^\circ \setminus T_\Gamma^\circ} 2^{\ell_v |\mathfrak{s}|} \right)^{\frac{1}{2}} \\ & \times \left( \prod_{v \in T_{\mathbf{p}^{-1}(\infty)}^\circ \cap T_\Gamma^\circ} 2^{-\ell_v |\mathfrak{s}|} \right)^{\frac{1}{2}} \end{aligned}$$

$$\begin{aligned}
&= \left( \prod_{v \in T^\circ} 2^{-\ell_v \tilde{\eta}(v)} \right) \left( \prod_{v \in T_\Gamma^\circ} 2^{\ell_v |\mathfrak{s}|} \right)^{\frac{1}{2}} \left( \prod_{v \in T_\Gamma^\circ \setminus \mathbf{p}^{-1}(\infty) \setminus T_\Gamma^\circ} 2^{\ell_v |\mathfrak{s}|} \right) \left( \prod_{v \in T_\Gamma^\circ \cap \mathbf{p}^{-1}(\infty)} 2^{\ell_v |\mathfrak{s}|} \right)^{\frac{1}{2}} \\
&\times \left( \prod_{v \in T_{\mathbf{p}^{-1}(2)}^\circ \setminus T_\Gamma^\circ} 2^{\ell_v |\mathfrak{s}|} \right)^{\frac{1}{2}}.
\end{aligned}$$

According to our definition of the labels  $\ell$  in Sect. 4.3 we have  $2^{-\ell_v} \geq \epsilon$ , and we can bound the preceding expression by the right-hand side of (4.33).

Now, we turn to the proof of (4.34). From [30, Lemma A.16] we conclude that the kernel (4.28) satisfies

$$\sup_{\mathbf{z} \in (\mathbf{R}^{d+1})^{\tilde{\mathbb{V}}_\star}} |\tilde{K}^{\epsilon, \mathbf{n}}(\mathbf{z})| \lesssim \prod_{v \in T^\circ} 2^{-\ell_v(\tilde{\eta}(v) - |\mathfrak{s}|)}, \quad (4.36)$$

uniformly over all  $\mathbf{n} \in \mathcal{N}_\gamma^{\mathfrak{e}}(T, \ell)$ . We will use this estimate to bound the norms (3.16).

Let us first consider the case  $M = 0$  corresponding to  $D = \tilde{\mathbb{V}}_{\text{var}}$ . We have  $(\tilde{K}_{\mathbb{G}_\gamma}^{\epsilon, \mathbf{n}})^{\gamma, \sigma}|_{\mathbf{z}_D} = (\tilde{K}_{\mathbb{G}_\gamma}^{\epsilon, \mathbf{n}})^{\gamma, \sigma}(\mathbf{z}_D)$ , and we are going to bound it absolutely. From (4.23) and (4.31) we get

$$(\tilde{K}_{\mathbb{G}_\gamma}^{\epsilon, \mathbf{n}})^{\gamma, \sigma}|_{\mathbf{z}_D} = \int_{(\mathbf{R}^{d+1})^{\tilde{\mathbb{V}}_\star \setminus \tilde{\mathbb{V}}_{\text{var}}}} (\tilde{K}_{\mathbb{G}_\gamma}^{\epsilon, \mathbf{n}})^{\gamma, \sigma}(\mathbf{z}) \Big|_{\mathbf{z}_{\text{var}} = \mathbf{z}_D} \prod_{v \in \tilde{\mathbb{V}}_\star \setminus \tilde{\mathbb{V}}_{\text{var}}} \mu_v^\epsilon(dz_v). \quad (4.37)$$

We write  $\mathbf{z} = (\mathbf{z}_{\tilde{\mathbb{V}}_\star \setminus \tilde{\mathbb{V}}_{\text{var}}}, \mathbf{z}^{\text{var}})$ , where  $\mathbf{z}_{\tilde{\mathbb{V}}_\star \setminus \tilde{\mathbb{V}}_{\text{var}}}$  contains the variables  $z_v$  with  $v \in \tilde{\mathbb{V}}_\star \setminus \tilde{\mathbb{V}}_{\text{var}}$ . The definition of the kernel and properties of the measures  $\mu_v^\epsilon$  allow to bound the preceding expression by a constant times

$$|\mathcal{A}_\gamma^\epsilon|_{|\tilde{\mathbb{V}}_\star \setminus \tilde{\mathbb{V}}_{\text{var}}|} \sup_{\mathbf{z}_{\tilde{\mathbb{V}}_\star \setminus \tilde{\mathbb{V}}_{\text{var}}} \in (\mathbf{R}^{d+1})^{\tilde{\mathbb{V}}_\star \setminus \tilde{\mathbb{V}}_{\text{var}}}} |\tilde{K}^{\epsilon, \mathbf{n}}(\mathbf{z}_{\tilde{\mathbb{V}}_\star \setminus \tilde{\mathbb{V}}_{\text{var}}}, \mathbf{z}_D)|,$$

where we write  $|\cdot|_\alpha$  for the  $(\alpha|\mathfrak{s}|)$ -dimensional Lebesgue measure, and the set  $\mathcal{A}_\gamma^\epsilon$  contains all points  $\{z_v : v \in \tilde{\mathbb{V}}_\star \setminus \tilde{\mathbb{V}}_{\text{var}}\}$ , satisfying the conditions

$$\begin{aligned}
\|z_v - z_w\|_{\mathfrak{s}} &\leq C' 2^{-\ell_{v \wedge w}} \quad \text{for } v, w \in \tilde{\mathbb{V}}_\star \setminus \tilde{\mathbb{V}}_{\text{var}}, \\
\|z_v - z_w^{\text{var}}\|_{\mathfrak{s}} &\leq C' 2^{-\ell_{v \wedge w}} \quad \text{for } v \in \tilde{\mathbb{V}}_\star \setminus \tilde{\mathbb{V}}_{\text{var}}, w \in \tilde{\mathbb{V}}_{\text{var}}.
\end{aligned}$$

Here, we use the fact that  $2^{-\ell_{v \wedge w}} \geq \epsilon$ , which is a consequence of the assumption  $(T, \ell) \in \mathbb{T}_\lambda^{\mathfrak{e}}(\tilde{\mathbb{V}})$ . For an interior node  $v \in T^\circ$ , let us choose  $v_\pm \in \tilde{\mathbb{V}}$  to be such that  $v_- \wedge v_+ = v$  and there is an edge from  $v_-$  to  $v_+$ . Then the collection of edges  $\{(v_-, v_+) : v \in T^\circ\}$  forms a spanning tree of  $\tilde{\mathbb{V}}$ , and  $\mathcal{A}_\gamma^\epsilon$  is a subset of

$$\left\{ \mathbf{z}_{\tilde{\mathbb{V}}_\star \setminus \tilde{\mathbb{V}}_{\text{var}}} \in (\mathbf{R}^{d+1})^{\tilde{\mathbb{V}}_\star \setminus \tilde{\mathbb{V}}_{\text{var}}} : \|\mathbf{z}_{v_+} - \mathbf{z}_{v_-}\|_{\mathfrak{s}} \leq C' 2^{-\ell_v} \quad \forall v \in T^\circ, v_\pm \notin \tilde{\mathbb{V}}_{\text{var}}, \right. \\ \left. \|\mathbf{z}_{v_+} - \mathbf{z}_{v_-}^{\text{var}}\|_{\mathfrak{s}} \leq C' 2^{-\ell_v} \quad \forall v \in T^\circ, v_- \in \tilde{\mathbb{V}}_{\text{var}} \right\},$$

where  $z_\star = 0$ . Here, we used the property that the vertices in  $\tilde{\mathbb{V}}_{\text{var}}$  have only outgoing edges. Next, we compute the Lebesgue measure of this set. We integrate out the variables  $z_v$  one by one, for  $v \notin \tilde{\mathbb{V}}_{\text{var}}$ , which gives an expression of order  $\prod_{v \in T^\circ \setminus T_{\tilde{\mathbb{V}}_{\text{var}}}^\circ} 2^{-\ell_v |\mathfrak{s}|}$ . Hence,

$$|\mathcal{A}_\gamma^\varepsilon|_{|\tilde{\mathbb{V}}_\star \setminus \tilde{\mathbb{V}}_{\text{var}}|} \lesssim \prod_{v \in T^\circ \setminus T_{\tilde{\mathbb{V}}_{\text{var}}}^\circ} 2^{-\ell_v |\mathfrak{s}|},$$

combining which with the estimate on the kernel (4.36) we get

$$\begin{aligned} \left| (\tilde{\mathcal{K}}_{\mathbb{G}_\gamma}^{\varepsilon, \mathbf{n}})^{\gamma, \sigma} \Big|_{\mathbf{z}_D} \right| &\lesssim \left( \prod_{v \in T^\circ \setminus T_{\tilde{\mathbb{V}}_{\text{var}}}^\circ} 2^{-\ell_v |\mathfrak{s}|} \right) \left( \prod_{v \in T^\circ} 2^{-\ell_v (\tilde{\eta}(v) - |\mathfrak{s}|)} \right) \\ &\lesssim \left( \prod_{v \in T^\circ} 2^{-\ell_v \tilde{\eta}(v)} \right) \left( \prod_{v \in T_{\tilde{\mathbb{V}}_{\text{var}}}^\circ} 2^{\ell_v |\mathfrak{s}|} \right). \end{aligned} \quad (4.38)$$

Recalling that  $D = \tilde{\mathbb{V}}_{\text{var}}$ , this is exactly the right-hand side of (4.34) with  $\mathbf{p}^{-1}(\infty) = \mathbf{p}^{-1}(2) = \emptyset$ .

Now we proceed with the proof of (4.34) by induction over  $M = 1, \dots, m$ . For  $\mathbf{p} \in \mathcal{P}_M$  with  $M \geq 2$ , let  $\bar{\mathbf{p}}$  be the restriction of  $\mathbf{p}$  to  $\{1, \dots, M-1\}$ . Let us furthermore define the function with respect to the variable  $z_{v_M}$  corresponding to the vertex  $v_M$ :

$$F(z_{v_M}) := \left\| (\tilde{\mathcal{K}}_{\mathbb{G}_\gamma}^{\varepsilon, \mathbf{n}})^{\gamma, \sigma} \Big|_{\mathbf{z}_{D \cup \{v_M\}}} \right\|_{L_\varepsilon^{\bar{\mathbf{p}}}} \quad (4.39)$$

if  $M \geq 2$  and  $F(z_{v_M}) := (\tilde{\mathcal{K}}_{\mathbb{G}_\gamma}^{\varepsilon, \mathbf{n}})^{\gamma, \sigma} \Big|_{\mathbf{z}_{D \cup \{v_M\}}}$  if  $M = 1$ . Then we use the definition (3.16) to write

$$\left\| (\tilde{\mathcal{K}}_{\mathbb{G}_\gamma}^{\varepsilon, \mathbf{n}})^{\gamma} \Big|_{\mathbf{z}_D} \right\|_{L_\varepsilon^{\mathbf{p}}} \leq \|F\|_{L_\varepsilon^{\mathbf{p}(M)}}. \quad (4.40)$$

We got an inequality because we omitted the indicator functions in the definition (3.16), which corresponds to increasing the domain of integration of the function. We need to bound the norm on the right-hand side of (4.40), for what we consider all possible values of  $\mathbf{p}(M)$  one-by-one.

If  $\mathbf{p}(M) = 2$ , then the definition (1.6) yields

$$\|F\|_{L_\varepsilon^2} = \left( \varepsilon^d \sum_{x \in \Lambda_\varepsilon} \int_0^\infty F(r, x)^2 dr \right)^{\frac{1}{2}}. \quad (4.41)$$

The norms (3.16) are defined on the time interval  $[0, T]$ , which means that the integral in (4.41) with respect to  $r$  should be on  $[0, T]$ . Since the kernel  $(\hat{\mathcal{K}}_{\mathbb{G}_\gamma}^{\varepsilon, \mathbf{n}})^{\gamma, \sigma}$  is compactly supported, we can take  $T$  big enough so that the integrals can be written on  $[0, \infty)$ . We use this convention in all formulas below.

Recalling the definitions of the kernel (4.31) and the function  $\mathbf{n}$  in Sect. 4.3, we conclude that the function  $F(z_{v_M})$  is supported on  $\|z_{v_M} - z_{v_M^\rightarrow}\|_{\mathfrak{s}} \lesssim 2^{-\ell_{v_M \wedge v_M^\rightarrow}}$ . We recall that  $v_M^\rightarrow = v_{M+1}$  if  $M < m$  and  $v_m^\rightarrow = \star$ . Then the norm (4.41) can be bounded as

$$\|F\|_{L_\varepsilon^2} \lesssim \left( 2^{-\ell_{v_M \wedge v_M^\rightarrow} |\mathfrak{s}|} \sup_{\|z_{v_M} - z_{v_M^\rightarrow}\|_{\mathfrak{s}} \lesssim 2^{-\ell_{v_M \wedge v_M^\rightarrow}}} F(z_{v_M})^2 \right)^{\frac{1}{2}}. \quad (4.42)$$

If  $M = 1$ , i.e. the set  $D \sqcup \{v_1\}$  in (4.39) equals  $\tilde{\mathbb{V}}_{\text{var}}$ , then the function  $F(z_{v_1})$  satisfies the bound (4.38). Then the expression (4.42) is bounded by a constant times

$$\begin{aligned} & 2^{-\ell_{v_1 \wedge v_1^\rightarrow} |\mathfrak{s}|/2} \left( \prod_{v \in T^\circ} 2^{-\ell_v \tilde{\eta}(v)} \right) \left( \prod_{v \in T_{\tilde{\mathbb{V}}_{\text{var}}}^\circ} 2^{\ell_v |\mathfrak{s}|} \right) \\ &= \left( \prod_{v \in T^\circ} 2^{-\ell_v \tilde{\eta}(v)} \right) \left( \prod_{v \in T_{\tilde{\mathbb{V}}_{\text{var}}}^\circ \setminus \{v_1 \wedge v_1^\rightarrow\}} 2^{\ell_v |\mathfrak{s}|} \right) 2^{\ell_{v_1 \wedge v_1^\rightarrow} |\mathfrak{s}|/2}. \end{aligned} \quad (4.43)$$

Since  $T_{\tilde{\mathbb{V}}_{\text{var}}}^\circ \setminus \{v_1 \wedge v_1^\rightarrow\} = T_{\tilde{\mathbb{V}}_{\text{var}} \setminus \{v_1\}}^\circ$ , this is exactly (4.34) with  $\mathbf{p}^{-1}(2) = \{1\}$ ,  $\mathbf{p}^{-1}(1) = \mathbf{p}^{-1}(\infty) = \emptyset$  and  $D = \tilde{\mathbb{V}}_{\text{var}} \setminus \{v_1\}$ .

If  $M \geq 2$ , i.e. the set  $D \sqcup \{v_M\}$  in (4.39) is a strict subset of  $\tilde{\mathbb{V}}_{\text{var}}$ , then by the induction hypothesis the function  $F(z_{v_M})$  satisfies the bound (4.34) with the function  $\tilde{\mathbf{p}}$  and the set  $D \sqcup \{v_M\}$ . Then the expression (4.42) is bounded by a constant multiple of

$$2^{-\ell_{v_M \wedge v_M^\rightarrow} |\mathfrak{s}|/2} \left( \prod_{v \in T^\circ} 2^{-\ell_v \tilde{\eta}(v)} \right) \left( \prod_{v \in T_{D \sqcup \{v_M\} \sqcup \tilde{\mathbf{p}}^{-1}(\infty)}^\circ} 2^{\ell_v |\mathfrak{s}|} \right) \left( \prod_{v \in T_{\tilde{\mathbf{p}}^{-1}(2)}^\circ} 2^{\ell_v |\mathfrak{s}|} \right)^{\frac{1}{2}}. \quad (4.44)$$

Since  $T_{D \sqcup \{v_M\} \sqcup \tilde{\mathbf{p}}^{-1}(\infty)}^\circ = T_{D \sqcup \tilde{\mathbf{p}}^{-1}(\infty)}^\circ \sqcup \{v_M \wedge v_M^\rightarrow\}$  and  $T_{\tilde{\mathbf{p}}^{-1}(2)}^\circ = T_{\mathbf{p}^{-1}(2)}^\circ \setminus \{v_M \wedge v_M^\rightarrow\}$ , this gives the required expression (4.34).

Now we consider the case  $\mathbf{p}(M) = 1$  in (4.40). Similarly to (4.42) we get

$$\|F\|_{L_\varepsilon^1} \lesssim 2^{-\ell_{v_M \wedge v_M^\rightarrow} |\mathfrak{s}|} \sup_{\|z_{v_M} - z_{v_M^\rightarrow}\|_{\mathfrak{s}} \lesssim 2^{-\ell_{v_M \wedge v_M^\rightarrow}}} |F(z_{v_M})|. \quad (4.45)$$

If  $M = 1$ , then by analogy with (4.43) we bound the preceding expression by a constant times

$$2^{-\ell_{v_1 \wedge v_1^\rightarrow} |\mathfrak{s}|} \left( \prod_{v \in T^\circ} 2^{-\ell_v \tilde{\eta}(v)} \right) \left( \prod_{v \in T_{\tilde{\mathbb{V}}_{\text{var}}}^\circ} 2^{\ell_v |\mathfrak{s}|} \right) = \left( \prod_{v \in T^\circ} 2^{-\ell_v \tilde{\eta}(v)} \right) \left( \prod_{v \in T_{\tilde{\mathbb{V}}_{\text{var}}}^\circ \setminus \{v_1 \wedge v_1^\rightarrow\}} 2^{\ell_v |\mathfrak{s}|} \right),$$

which is exactly (4.34) with  $D = \tilde{\mathbb{V}}_{\text{var}} \setminus \{v_1\}$ . In the case  $M \geq 2$  we use the induction hypothesis, and by analogy with (4.44), we bound (4.45) by a constant times

$$2^{-\ell_{v_M \wedge v_M^\rightarrow} |\mathfrak{s}|} \left( \prod_{v \in T^\circ} 2^{-\ell_v \tilde{\eta}(v)} \right) \left( \prod_{v \in T_{D \cup \mathbf{p}^{-1}(\infty) \cup \{v_M\}}^\circ} 2^{\ell_v |\mathfrak{s}|} \right) \left( \prod_{v \in T_{\mathbf{p}^{-1}(2)}^\circ} 2^{\ell_v |\mathfrak{s}|} \right)^{\frac{1}{2}},$$

which is the required bound (4.34).

Finally, we consider the case  $\mathbf{p}(M) = \infty$  in (4.40). Similarly to (4.42) we can bound

$$\|F\|_{L_\varepsilon^p} \lesssim \sup_{\|z_{v_M} - z_{v_M^\rightarrow}\|_{\mathfrak{s}} \lesssim 2^{-\ell_{v_M \wedge v_M^\rightarrow}}} |F(z_{v_M})|. \quad (4.46)$$

In the case  $M = 1$ , we use (4.38) to bound this expression by a constant multiple of

$$\left( \prod_{v \in T^\circ} 2^{-\ell_v \tilde{\eta}(v)} \right) \left( \prod_{v \in T_{\tilde{\mathbb{V}}_{\text{var}}}^\circ} 2^{\ell_v |\mathfrak{s}|} \right),$$

which is the required bound (4.34) with  $\mathbf{p}^{-1}(1) = \mathbf{p}^{-1}(2) = \emptyset$ ,  $\mathbf{p}^{-1}(\infty) = \{v_1\}$  and  $D = \tilde{\mathbb{V}}_{\text{var}} \setminus \{v_1\}$ . If  $M \geq 2$ , we use the induction hypothesis and similarly to (4.44) we bound the expression (4.46) by a constant times

$$\left( \prod_{v \in T^\circ} 2^{-\ell_v \tilde{\eta}(v)} \right) \left( \prod_{v \in T_{D \cup \mathbf{p}^{-1}(\infty) \cup \{v_M\}}^\circ} 2^{\ell_v |\mathfrak{s}|} \right) \left( \prod_{v \in T_{\mathbf{p}^{-1}(2)}^\circ} 2^{\ell_v |\mathfrak{s}|} \right)^{\frac{1}{2}},$$

which is exactly (4.34).  $\square$

Since the product in (4.34) is different from the one in [30, Lemma A.10], we need to have an analogous result in our context. For this we define the function  $\widehat{\eta} : T^\circ \rightarrow \mathbf{R}$  by  $\widehat{\eta}(v) = \tilde{\eta}(v)$  if  $v \notin T_\Gamma^\circ$ , and  $\widehat{\eta}(v) = \tilde{\eta}(v) - |\mathfrak{s}|/2$  if  $v \in T_\Gamma^\circ$ , where we use the set  $T_\Gamma^\circ$  introduced above Lemma 4.7.

**Lemma 4.8** *In the setting of Theorem 4.3, the function  $\widehat{\eta}$  satisfies assumptions of [30, Lemma A.10], and*

$$|\widehat{\eta}| := \sum_{v \in T^\circ} \widehat{\eta}(v) = |\mathfrak{s}| |\tilde{\mathbb{V}}_\star \setminus \{v_\star^\uparrow\}| - \sum_{e \in \widehat{\mathbb{E}}} \widehat{a}_e - \frac{|\mathfrak{s}|}{2} |\Gamma|. \quad (4.47)$$

**Proof** Using the definition of the function  $\widehat{\eta}$ , the assumptions of [30, Lemma A.10] follow at once if we prove the following two properties

1. For every  $v \in T^\circ$  one has  $\sum_{v \geq v} \tilde{\eta}(v) > \frac{|\mathfrak{s}|}{2} \#\{v \in T_\Gamma^\circ : v \geq v\}$ .
2. For every  $v \in T^\circ$  such that  $v \leq v_\star$  one has  $\sum_{v \not\geq v} \tilde{\eta}(v) < \frac{|\mathfrak{s}|}{2} \#\{v \in T_\Gamma^\circ : v \not\geq v\}$ , provided that this sum contains at least one term, where  $v_\star$  is a fixed distinguished inner node.

These bounds can be shown by repeating the proof of [30, Lemma A.19] and using Assumption 3. To compute (4.47) we use  $|T_\Gamma^\circ| = |\Gamma|$ .  $\square$

**Lemma 4.9** *Let the function  $\tilde{\eta}$  be defined in (4.32). In the setting of Theorem 4.3 the following bound holds uniformly over  $\lambda \in [\mathfrak{e}, 1]$ :*

$$\sum_{\ell \in \mathcal{N}_\lambda^\mathfrak{e}(T^\circ)} \left( \prod_{v \in T^\circ} 2^{-\ell_v \tilde{\eta}(v)} \right) \left( \prod_{v \in T_\Gamma^\circ} 2^{\ell_v |\mathfrak{s}|/2} \right) \lesssim \lambda^{|\widehat{\eta}|}, \quad (4.48)$$

where  $|\widehat{\eta}|$  is computed in (4.47).

**Proof** Using the function  $\widehat{\eta}$ , defined above Lemma 4.8, we can write the left-hand side of (4.48) as

$$\sum_{\ell \in \mathcal{N}_\lambda^\mathfrak{e}(T^\circ)} \prod_{v \in T^\circ} 2^{-\ell_v \widehat{\eta}_v}.$$

Then Lemma 4.8 implies that the function  $\widehat{\eta}$  satisfies the assumptions of [30, Lemma A.10], and [30, Equation A.29] allows to bound the left-hand side of (4.48) by  $\lambda^{|\widehat{\eta}|}$ , where  $|\widehat{\eta}|$  is computed in (4.47).  $\square$

#### 4.4 Proof of Theorem 4.3

We use formulas (4.19) and (4.30), and apply Theorem 3.9 to each term:

$$\begin{aligned} & \mathbf{E}_p \sup_{t \in \mathbf{R}_+} |(\mathcal{I}_\gamma^{\varepsilon, \mathbb{L}} \mathcal{K}_{\mathbb{G}}^{\lambda, \varepsilon})_t| \\ & \lesssim_p \sum_{\sigma \in \Sigma_\gamma} \sum_{\mathbf{p} \in \mathcal{P}_n : \substack{\mathbf{p}^{-1}(1) \cap \Gamma(\gamma) = \emptyset, \\ \mathbf{L}^{-1}(\nabla) \subseteq \mathbf{p}^{-1}(1)}} \sum_{(T, \ell) \in \mathbb{T}_\lambda^\mathfrak{e}(\tilde{\mathbb{V}})} \sum_{\mathbf{n} \in \mathcal{N}_\gamma^\mathfrak{e}(T, \ell)} \varepsilon^{\alpha_\gamma(\mathbf{p})} \|(\tilde{\mathcal{K}}_{\mathbb{G}_\gamma}^{\varepsilon, \mathbf{n}})^{\gamma, \sigma}\|_{L_\varepsilon^{\mathbf{p}}}^{\beta_{\gamma, p}(\mathbf{p})} \\ & \quad \times \left( \prod_{\substack{i \in \mathbf{p}^{-1}(\infty) \setminus \Gamma_1(\gamma) : \\ i \geq 2}} \varepsilon^{-\kappa_{\gamma, i}(\mathbf{p})} \|(\tilde{\mathcal{K}}_{\mathbb{G}_\gamma}^{\varepsilon, \mathbf{n}})^{\gamma, \sigma}\|_{\mathcal{C}_{s_i}^1(L_\varepsilon^\infty)}^{\beta_{\gamma \geq i, p}(\mathbf{p}^{\geq i})} \right)^{\frac{1}{p}}, \end{aligned} \quad (4.49)$$

for some constants  $\kappa_{\gamma, i}(\mathbf{p}) > 0$ . Next, we are going to bound the terms in the sum in (4.49) for different functions  $\mathbf{p}$ .

If the function  $\mathbf{p}$  satisfies  $\mathbf{p}^{-1}(\infty) = \emptyset$ , then we have  $\beta_{\gamma, p}(\mathbf{p}) = 1$  and the product in the parentheses in (4.49) equals 1. Then the inner double sum in (4.49) simplifies to

$$\varepsilon^{\alpha_\gamma(\mathbf{p})} \sum_{(T, \ell) \in \mathbb{T}_\lambda^\epsilon(\tilde{\mathbb{V}})} \sum_{\mathbf{n} \in \mathcal{N}_\gamma^\epsilon(T, \ell)} \|(\tilde{\mathcal{K}}_{\mathbb{G}_\gamma}^{\epsilon, \mathbf{n}})^{\gamma, \sigma}\|_{L_\epsilon^p}.$$

Using Lemma 4.7, we bound this expression by a constant multiple of

$$\varepsilon^{\alpha_\gamma(\mathbf{p})} \epsilon^{-\delta_\gamma(\mathbf{p})} \sum_{(T, \ell) \in \mathbb{T}_\lambda^\epsilon(\tilde{\mathbb{V}})} \sum_{\mathbf{n} \in \mathcal{N}_\gamma^\epsilon(T, \ell)} \left( \prod_{v \in T^\circ} 2^{-\ell_v \tilde{\eta}(v)} \right) \left( \prod_{v \in T_\Gamma^\circ} 2^{\ell_v |\mathbf{s}|} \right)^{\frac{1}{2}}.$$

Lemma 4.9 allows to bound this expression by a constant times  $\varepsilon^{\alpha_\gamma(\mathbf{p})} \epsilon^{-\delta_\gamma(\mathbf{p})} \lambda^{\nu_\gamma}$ , with  $\nu_\gamma$  defined in (4.15).

Let  $\mathbf{p}^{-1}(\infty) \neq \emptyset$ . We can bound the norm  $\|(\tilde{\mathcal{K}}_{\mathbb{G}_\gamma}^{\epsilon, \mathbf{n}})^{\gamma, \sigma}\|_{\mathcal{C}_s^1(L_\epsilon^\infty)}$  defined in (3.3). More precisely, Lemma 4.7 yields

$$\|\partial_{s_i}^k (\tilde{\mathcal{K}}_{\mathbb{G}_\gamma}^{\epsilon, \mathbf{n}})^{\gamma, \sigma}\|_{L_\epsilon^\infty} \lesssim \epsilon^{-\delta_\gamma(\infty) - 2k} \left( \prod_{v \in T^\circ} 2^{-\ell_v \tilde{\eta}(v)} \right) \left( \prod_{v \in T_\Gamma^\circ} 2^{\ell_v |\mathbf{s}|} \right)^{\frac{1}{2}}, \quad (4.50)$$

for  $k = 0$  and  $k = 1$ , where we write  $\delta_\gamma(\infty)$  for the constant (4.17) defined via the function  $\mathbf{p} \equiv \infty$ , and where the derivative  $\partial_{s_i}$  gives the multiplier  $\epsilon^{-2}$  (as follows from the scaling properties of the kernels). The norm (4.50) can be brutally bounded by a negative power of  $\epsilon \gtrsim \varepsilon$ , and hence the whole expression in the parentheses in (4.49) can be bounded by  $\varepsilon^{-\kappa}$  for some  $\kappa \geq 0$ . Hence, the inner double sum in (4.49) is estimated by

$$\sum_{(T, \ell) \in \mathbb{T}_\lambda^\epsilon(\tilde{\mathbb{V}})} \sum_{\mathbf{n} \in \mathcal{N}_\gamma^\epsilon(T, \ell)} \varepsilon^{\alpha_\gamma(\mathbf{p}) - \frac{\kappa}{p}} \|(\tilde{\mathcal{K}}_{\mathbb{G}_\gamma}^{\epsilon, \mathbf{n}})^{\gamma, \sigma}\|_{L_\epsilon^p}^{\beta_{\gamma, p}(\mathbf{p})}.$$

Since  $\beta_{\gamma, p}(\mathbf{p}) \leq 1$  (see (3.20)), we use Jensen's inequality to estimate this expression by a constant times

$$\varepsilon^{\alpha_\gamma(\mathbf{p}) - \frac{\kappa}{p}} \left( \sum_{(T, \ell) \in \mathbb{T}_\lambda^\epsilon(\tilde{\mathbb{V}})} \sum_{\mathbf{n} \in \mathcal{N}_\gamma^\epsilon(T, \ell)} \|(\tilde{\mathcal{K}}_{\mathbb{G}_\gamma}^{\epsilon, \mathbf{n}})^{\gamma, \sigma}\|_{L_\epsilon^p} \right)^{\beta_{\gamma, p}(\mathbf{p})}.$$

We use Lemmas 4.7 and 4.9 to bound the double sum by  $\epsilon^{-\delta_\gamma(\mathbf{p})} \lambda^{\nu_\gamma}$ , and we bound the preceding expression by

$$\varepsilon^{\alpha_\gamma(\mathbf{p}) - \frac{\kappa}{p}} \left( \epsilon^{-\delta_\gamma(\mathbf{p})} \lambda^{\nu_\gamma} \right)^{\beta_{\gamma, p}(\mathbf{p})}.$$

The expression in the parentheses is smaller than one (recall that we assumed  $\nu_\gamma < 0$ ), and this expression can be estimated by  $\varepsilon^{\alpha_\gamma(\mathbf{p}) - \frac{\kappa}{p}} e^{-\delta_\gamma(\mathbf{p})\lambda\nu_\gamma}$ . Taking  $p$  sufficiently large, we get the required bound (4.16).

## 5 Application to a discrete martingale model

This section is a showcase of the theory we developed in this paper. We introduce a family of martingales indexed by points of the lattice  $\Lambda_\varepsilon$ ; we then build trees as iterated integrals against the martingales themselves; and we finally apply our theory to prove uniform bounds.

The martingales in this section are chosen to resemble those that appear in our companion paper [21], where we prove convergence of the dynamical Ising–Kac model to  $\Phi_3^4$  and this proof of convergence is what motivated the development of the theory here in the first place. We have therefore chosen to present a family of martingales which is both simpler and similar to the one found in the Ising–Kac model. In this way, we aim to give the reader a concrete and easy example of how the theory above can be applied.

For the proof of convergence in our companion paper [21], we use the theory of regularity structures [26] (see also [15, 27]), together with the discretisation framework by [13]. We prefer not to reintroduce all the concepts developed in these articles. Generally speaking, however, the theory of regularity structures is used as a solution theory for the (continuous)  $\Phi_3^4$  equation, while the discretisation framework by [13] gives us a solution theory for the discrete Ising–Kac model which preserves the formalism of regularity structures; [13] also supplies us with some convergence tools, while our theory develops the missing tool for convergence of models, namely uniform boundedness in the scaling parameter.

As follows from [26], the regularity structure for the  $\Phi_3^4$  equation has a basis which is convenient to write as formal expressions, which are written using the symbols  $\Xi$ ,  $\mathcal{I}$  and  $X_i$ ,  $i = 0, \dots, 3$ . Here, the symbol  $\Xi$  corresponds to the driving noise of the equation,  $\mathcal{I}$  corresponds to the convolution map with respect to the heat kernel, and  $X_i$  are the time-space variables. For example, the expression  $\Psi := \mathcal{I}(\Xi)$  corresponds to the convolution of the heat kernel with the driving noise. The first several basis elements of the regularity structure are  $\Xi$ ,  $\Psi$ ,  $\Psi^2$ ,  $\Psi^3$ ,  $\Psi^2 X_i$ ,  $\mathcal{I}(\Psi^3)\Psi$ ,  $\mathcal{I}(\Psi^3)\Psi^2$ ,  $\mathcal{I}(\Psi^3)\Psi^3$ . In this section we will prove moment bounds for a discrete model acting only on the elements  $\Xi$ ,  $\Psi$ ,  $\Psi^2$ ,  $\mathcal{I}(\Psi^3)\Psi^2$ , which we believe are the most interesting. We refer the reader to our companion paper [21] for a full description of the regularity structure for the Ising–Kac model.

A *model* is a pair of linear maps  $(\Pi, \Gamma)$  on a regularity structure, which map the basis elements into functions/distributions. These maps are required to have certain algebraic and analytic properties which can be found in [26]. In this section we will consider a discrete model (in the sense of [13]) and, more precisely, only a discretisation of the map  $\Pi$ . For this, we need to make some definitions.

For any  $\alpha \in (0, 1)$ , we define  $\varepsilon := \varepsilon^\alpha$  and the function  $\psi_\varepsilon : \mathbf{R}^3 \rightarrow \mathbf{R}_+$  by  $\psi_\varepsilon(x) := \varepsilon^{-3}\psi(\varepsilon^{-1}x)$ , with  $\psi$  being a smooth function, supported in the ball centered

at the origin and of radius 1, and satisfying  $\int_{\mathbf{R}^3} \psi(x) dx = 1$ . While any  $\alpha \in (0, 1)$  works for our purposes, we choose  $\alpha = 3/4$  to be in the setting of [21]. Then for  $t \geq 0$  and  $x \in \Lambda_\varepsilon := (\varepsilon \mathbf{Z}/\mathbf{Z})^3$  we define the martingale

$$\mathcal{M}_\varepsilon(t, x) = \frac{1}{\sqrt{2}} \varepsilon^{\frac{5}{2}} \sum_{y \in \Lambda_\varepsilon} \psi_\varepsilon(x - y) \left( \mathcal{P}_{\varepsilon^{-2}t}(\varepsilon^{-1}y) - \tilde{\mathcal{P}}_{\varepsilon^{-2}t}(\varepsilon^{-1}y) \right), \quad (5.1)$$

where  $\mathcal{P}_t(x)$  and  $\tilde{\mathcal{P}}_t(x)$  are independent Poisson processes of intensities 1. We extend these martingales periodically to  $x \in \varepsilon \mathbf{Z}^3$  and we extend them to  $\mathbf{R}$  in time as in (4.1). We denote the new space-time domain by  $D_\varepsilon := \mathbf{R} \times \varepsilon \mathbf{Z}^3$ .

As mentioned above, we want to make the family of martingales (5.1) as similar as possible to the family of martingales of the Ising–Kac interaction system; and, by the choice of  $\alpha = \frac{3}{4}$ , the two families of martingales have the same limiting behaviour as  $\varepsilon \rightarrow 0$ . We refer the reader to the [21] for a more detailed explanation.

Using these martingales, we are going to define a discretisation of the map  $\Pi$ , which we denote by  $\widehat{\Pi}^\varepsilon$ . As we mentioned above, we will bound this map only for the four elements  $\Xi$ ,  $\Psi$ ,  $\Psi^2$  and  $\mathcal{I}(\Psi^3)\Psi^2$  of the regularity structure, and a complete analysis of the map in a similar context is performed in [21]. For every fixed  $z \in D_\varepsilon$  the action of this map on the element  $\Xi$  is defined as

$$(\widehat{\Pi}_z^\varepsilon \Xi)(\bar{z}) = d\mathcal{M}_\varepsilon(\bar{z}), \quad (5.2)$$

which means that for every test function  $\varphi : \mathbf{R}^4 \rightarrow \mathbf{R}$  we have

$$\iota_\varepsilon(\widehat{\Pi}_z^\varepsilon \Xi)(\varphi) = \int_{D_\varepsilon} \varphi(\bar{z}) d\mathcal{M}_\varepsilon(\bar{z}),$$

where we used the extension (2.3), the expression on the left-hand side means the duality pairing of the distribution  $\iota_\varepsilon(\widehat{\Pi}_z^\varepsilon \Xi)$  with the test function  $\varphi$ , and where the integral with respect to the martingale is defined as in (4.2).

Let  $P(t, x) := \frac{1}{(4\pi t)^{3/2}} e^{-|x|^2/(4t)}$  be the heat kernel on  $\mathbf{R}^3$ . In order to integrate it with respect to the martingales  $\mathcal{M}_\varepsilon$ , we need to remove the singularity of  $P$  at the origin. For this, we will convolve  $P$  with a smooth function. More precisely, let us take any smooth function  $\tilde{\psi} : \mathbf{R}^4 \rightarrow \mathbf{R}$ , supported in the unit ball with the center at the origin and which satisfies  $\int_{\mathbf{R}^4} \tilde{\psi}(z) dz = 1$ . Let us set  $\tilde{\psi}_\varepsilon(t, x) := \varepsilon^{-5} \tilde{\psi}(\varepsilon^{-2}t, \varepsilon^{-1}x)$ . Then we define a smoothed heat kernel  $P^\varepsilon := P * \tilde{\psi}_\varepsilon$ , where the convolution is over  $\mathbf{R}^4$ . As follows from [26, Lemma 7.7], we can write  $P^\varepsilon = K^\varepsilon + R^\varepsilon$ , where  $K^\varepsilon$  is a compactly supported singular part of the kernel (i.e.  $K^\varepsilon(0)$  diverges as  $\varepsilon \rightarrow 0$ ) and  $R^\varepsilon$  is smooth. Then we set

$$(\widehat{\Pi}_z^\varepsilon \Psi)(\bar{z}) = \int_{D_\varepsilon} K^\varepsilon(\bar{z} - \tilde{z}) d\mathcal{M}_\varepsilon(\tilde{z}), \quad (5.3)$$

where we recall that  $\Psi = \mathcal{I}(\Xi)$ .

We will also use the kernel  $K^\epsilon := K^\epsilon \star_\epsilon \psi_\epsilon$ , where  $\star_\epsilon$  is the convolution on  $\epsilon\mathbf{Z}^3$ . Then we set

$$(\widehat{\Pi}_z^\epsilon \Psi^2)(\bar{z}) = (\widehat{\Pi}_z^\epsilon \Psi)(\bar{z})^2 - C_1^\epsilon, \quad (5.4)$$

with the renormalisation constant

$$C_1^\epsilon = \int_{D_\epsilon} K^\epsilon(z)^2 dz. \quad (5.5)$$

It is not difficult to see that  $\widehat{\Pi}_z^\epsilon \Psi$  converges to a distribution as  $\epsilon \rightarrow 0$ . This implies that the product  $(\widehat{\Pi}_z^\epsilon \Psi^2)(\bar{z})^2$  diverges in the limit, and in order to have a non-trivial limit we need to renormalise the product by subtracting the divergent constant  $C_1^\epsilon$ . The precise formula for this constant will be explained in Sect. 5.3 below.

Finally, for the element  $\mathcal{I}(\Psi^3)\Psi^2$  we set

$$\begin{aligned} (\widehat{\Pi}_z^\epsilon \mathcal{I}(\Psi^3)\Psi^2)(\bar{z}) &= (\widehat{\Pi}_z^\epsilon \Psi)(\bar{z})^2 \int_{D_\epsilon} (K^\epsilon(\bar{z} - z) - K^\epsilon(z - \bar{z})) (\widehat{\Pi}_z^\epsilon \Psi)(\bar{z})^3 dz \\ &\quad - 3C_2^\epsilon (\widehat{\Pi}_z^\epsilon \Psi)(\bar{z}), \end{aligned} \quad (5.6)$$

where the new renormalisation constant is

$$C_2^\epsilon = 2 \int_{D_\epsilon} \int_{D_\epsilon} \int_{D_\epsilon} K^\epsilon(z_1) K^\epsilon(z_1 - z_3) K^\epsilon(z_2) K^\epsilon(z_2 - z_3) K^\epsilon(z_3) dz_1 dz_2 dz_3. \quad (5.7)$$

Again, we need to subtract the renormalisation constant to have non-divergent moment bounds for the function. The formula for the renormalisation constant is explained in Sect. 5.4 below.

For a fixed  $\kappa > 0$ , we assign to these four basis elements a *homogeneity*  $|\cdot|$  as

$$|\Xi| = -\frac{5}{2} - \kappa, \quad |\Psi| = -\frac{1}{2} - \kappa, \quad |\Psi^2| = -1 - 2\kappa, \quad |\mathcal{I}(\Psi^3)\Psi^2| = -\frac{1}{2} - 5\kappa.$$

Let  $\tau$  be one of these elements. We are going to prove that for some  $\bar{\kappa} > 0$ , any  $p \geq 2$  and any test function  $\varphi : \mathbf{R}^4 \rightarrow \mathbf{R}$  the following bound holds:

$$\mathbf{E}_p |\iota_\epsilon(\widehat{\Pi}_z^\epsilon \tau)(\varphi_z^\lambda)| \leq C \lambda^{|\tau| + \bar{\kappa}}, \quad (5.8)$$

uniformly in  $z \in D_\epsilon$ ,  $\lambda \in [\epsilon, 1]$  and  $\epsilon \in (0, 1]$ , where we use the extension (2.3) and a recentered and rescaled test function (4.11). The constant  $C$  in this bound may depend on  $p$ .

For every element  $\tau \in \{\Xi, \Psi, \Psi^2, \mathcal{I}(\Psi^3)\Psi^2\}$ , we use (2.14) to write

$$(\widehat{\Pi}_z^\epsilon \tau)(\varphi_z^\lambda) = \sum_{\gamma \in \mathfrak{C}(n)} \sum_{\sigma \in \Sigma_\gamma} \int_{D_\epsilon} \varphi_z^\lambda(\bar{z}) \left( \int_{D_{\gamma, \sigma}} F_{\bar{z}}(z_1, \dots, z_n) d\mathcal{M}_\epsilon^n(z_1, \dots, z_n) \right) d\bar{z},$$

where  $n$  is the number of instances of  $\Xi$  in the definition of  $\tau$  (recall that  $\Psi = \mathcal{I}(\Xi)$ ), the measure  $\mathcal{M}_\varepsilon^n$  is the product measure built from  $\mathcal{M}_\varepsilon$ , and some function  $F$ . In order to bound such terms, we need to use martingales satisfying Assumption 1. For this, we use the definition (5.1) and rewrite the previous expression in the form

$$(\widehat{\Pi}_z^\varepsilon \tau)(\varphi_z^\lambda) = \sum_{\gamma \in \mathfrak{C}(n)} \sum_{\sigma \in \Sigma_\gamma} \int_{D_\varepsilon} \varphi_z^\lambda(\bar{z}) \\ \times \left( \int_{\mathcal{D}_{\gamma, \sigma}} \left( (F \star_\varepsilon^n \psi_\varepsilon)_{\bar{z}}(z_1, \dots, z_n) \right) d\mathbf{M}_\varepsilon^n(z_1, \dots, z_n) \right) d\bar{z}, \quad (5.9)$$

where now the measure  $\mathbf{M}_\varepsilon^n$  is built as in (2.13) using the family of martingales

$$\mathbb{M}_\varepsilon(t, x) := \frac{1}{\sqrt{2}} \varepsilon^{-\frac{1}{2}} \left( \mathcal{P}_{\varepsilon^{-2}t}(\varepsilon^{-1}x) - \widetilde{\mathcal{P}}_{\varepsilon^{-2}t}(\varepsilon^{-1}x) \right) \quad (5.10)$$

and where  $F \star_\varepsilon^n \psi_\varepsilon$  is the discrete convolution of  $F$  against the function  $\psi_\varepsilon$  in each of the variables of  $F$ . In particular,  $\mathbb{M}_\varepsilon$  are càdlàg martingales, satisfying Assumption 1 with  $\mathbf{k} = -\frac{1}{2}$  and  $\mathbf{C}_\varepsilon \equiv \mathbf{C}_\varepsilon \equiv 1$ . We note that we can replace the Poisson processes in (5.1) and in (5.10) by their compensated versions, because the integrals of their intensities cancel each other. Furthermore, (2.15) allows to write (5.9) as a chaos expansion

$$(\widehat{\Pi}_z^\varepsilon \tau)(\varphi_z^\lambda) = \sum_{\gamma \in \mathfrak{C}(n)} \sum_{\sigma \in \Sigma_\gamma} \int_{D_\varepsilon} \varphi_z^\lambda(\bar{z}) (\mathcal{I}_{\gamma, \sigma}^\varepsilon (F \star_\varepsilon^n \psi_\varepsilon)_{\bar{z}}) d\bar{z}.$$

The functions  $\varphi_z^\lambda$  and  $F \star_\varepsilon^n \psi_\varepsilon$  may be written in this expression as a suitable kernel  $\mathcal{K}_{\mathbb{G}, z}^{\lambda, \varepsilon}$  in the sense of (4.18). Moreover, due to the renormalisation involved in the definition of  $\widehat{\Pi}_z^\varepsilon \tau$  we will typically write this expression in terms of renormalised stochastic integrals (3.30) with suitable labelings  $\mathbb{L}_\gamma$ , which may be different for different contractions  $\gamma$ . Hence, we will write the preceding expression as

$$(\widehat{\Pi}_z^\varepsilon \tau)(\varphi_z^\lambda) = \sum_{\gamma \in \mathfrak{C}(n)} \sum_{\sigma \in \Sigma_\gamma} \mathcal{I}_{\gamma, \sigma}^{\varepsilon, \mathbb{L}_\gamma} \mathcal{K}_{\mathbb{G}, z}^{\lambda, \varepsilon}.$$

Finally, we use (4.14) to write it as

$$(\widehat{\Pi}_z^\varepsilon \tau)(\varphi_z^\lambda) = \sum_{\gamma \in \mathfrak{C}(n)} \mathcal{I}_\gamma^{\varepsilon, \mathbb{L}_\gamma} \mathcal{K}_{\mathbb{G}, z}^{\lambda, \varepsilon},$$

and we are going to bound each term in this sum using Corollary 4.5.

Additionally, it is convenient to use graphical notation to represent the kernels  $\mathcal{K}_{\mathbb{G}, z}^{\lambda, \varepsilon}$ . In the graphical notation, nodes represent variables and arrows represent kernels. The vertex “●” labelled with  $z$  represents the basis point  $z \in D_\varepsilon$ . The arrow “→” represents a test function  $\varphi_z^\lambda$ . The arrow “→” represents either the discrete kernel  $K^\varepsilon$  or  $K^\varepsilon$ , and we will write two labels  $(a_e, r_e)$  on this arrow, which correspond to the

labels on graphs as described in Sect. 4. More precisely, since the kernels  $K^\varepsilon$  and  $K^\epsilon$  satisfy the bound (4.7) with  $a_\varepsilon = 3$  (this follows from [26, Lemma 7.7]), Lemma 4.2 implies that the kernels  $K^\varepsilon$  and  $K^\epsilon$  have all the properties from Assumption 4 with the values  $a_\varepsilon = 3$  and  $r_\varepsilon = 0$ . Hence, we will depict this kernel by “ $\xrightarrow{-3.0}$ ”. Whenever a contracted variable  $z_i$  is integrated with respect to the measure  $\mathbb{M}_\varepsilon^n$  with  $n \geq 2$ , we denote it by a node “ $\bullet$ ”. Moreover, the variable integrated with respect to  $\mathbb{M}_\varepsilon$  will be denoted by “ $\circlearrowleft$ ”. By the node “ $\bullet$ ” we denote a variable integrated out in  $D_\varepsilon$ .

Using this notation, we will now prove the bounds (5.8) for each of the four elements  $\tau$ .

## 5.1 The element $\tau = \Xi$

The definitions (5.2), (5.1) and (5.10) yield

$$\iota_\varepsilon(\widehat{\Pi}_z^\varepsilon \Xi)(\varphi_z^\lambda) = \int_{D_\varepsilon} (\varphi_z^\lambda \star_\varepsilon \psi_\epsilon)(\bar{z}) d\mathbb{M}_\varepsilon(\bar{z}).$$

This expression is not in the scope of Theorem 4.3, but we can bound it using Proposition 3.2 since the right-hand side is a first-order stochastic integral. Recalling that we extended the martingale  $\mathbb{M}_\varepsilon$  to all times in  $\mathbf{R}$  by (4.1), we write the preceding expression explicitly as

$$\sum_{\bar{x} \in \varepsilon \mathbf{Z}^3} \varepsilon^3 \int_{\bar{t}=0}^{\infty} (\varphi_z^\lambda \star_\varepsilon \psi_\epsilon)(\bar{t}, \bar{x}) d\mathbb{M}_\varepsilon(\bar{t}, \bar{x}) + \sum_{\bar{x} \in \varepsilon \mathbf{Z}^3} \varepsilon^3 \int_{\bar{t}=0}^{\infty} (\varphi_z^\lambda \star_\varepsilon \psi_\epsilon)(-\bar{t}, \bar{x}) d\widetilde{\mathbb{M}}_\varepsilon(\bar{t}, \bar{x}),$$

where  $\widetilde{\mathbb{M}}_\varepsilon$  is an independent copy of  $\mathbb{M}_\varepsilon$  (see (4.1)). We bound these two integrals using Proposition 3.2, recalling the definition (2.10), and get

$$\mathbf{E}_p |\iota_\varepsilon(\widehat{\Pi}_z^\varepsilon \Xi)(\varphi_z^\lambda)| \lesssim \left( \varepsilon^3 \sum_{\bar{x} \in \varepsilon \mathbf{Z}^3} \int_{\mathbf{R}} (\varphi_z^\lambda \star_\varepsilon \psi_\epsilon)(\bar{t}, \bar{x})^2 d\bar{t} \right)^{\frac{1}{2}} + \varepsilon^{\frac{5}{2}} \sup_{(\bar{t}, \bar{x}) \in D_\varepsilon} |(\varphi_z^\lambda \star_\varepsilon \psi_\epsilon)(\bar{t}, \bar{x})|, \quad (5.11)$$

where we used  $\mathbf{k} = -\frac{1}{2}$ . The function  $(\varphi_z^\lambda \star_\varepsilon \psi_\epsilon)(\bar{t}, \bar{x})$  is supported on  $(\bar{t}, \bar{x}) \in D_\varepsilon$  satisfying  $|\bar{t} - t| \lesssim \lambda$  and  $|\bar{x} - x| \lesssim \lambda \vee \epsilon$ . Moreover, this function is rescaled by  $\lambda$  in time and  $\lambda \vee \epsilon$  in space. Then the first term on the right-hand side of (5.11) is bounded by a constant times  $\lambda^{-\frac{5}{2}}$  for  $\lambda \in [\epsilon, 1]$ . The second term in (5.11) is bounded by a constant times  $\varepsilon^{\frac{5}{2}} \lambda^{-5} \leq \lambda^{-\frac{5}{2}}$  for  $\lambda \in [\epsilon, 1]$ . Hence, we obtained the required bound (5.8) with  $|\Xi| = -\frac{5}{2} - \kappa$  for any  $\kappa > 0$ .

## 5.2 The element $\tau = \Psi$

Using (5.3), (5.1) and (5.10), we can represent the map  $\widehat{\Pi}_z^\varepsilon \tau$  diagrammatically as

$$\iota_\varepsilon(\widehat{\Pi}_z^\varepsilon \tau)(\varphi_z^\lambda) = \textcolor{blue}{\circlearrowleft} \xrightarrow{-3.0} \bullet \xleftarrow{z} \bullet.$$

This diagram is the stochastic integral  $\mathcal{I}^\varepsilon(F)$ , where the kernel  $F$  is in this case the generalised convolution  $\mathcal{K}_{\mathbb{G}, z}^{\lambda, \varepsilon}$ , as in (4.18), given by

$$\mathcal{K}_{\mathbb{G}, z}^{\lambda, \varepsilon}(z^{\text{var}}) = \int_{D_\varepsilon} \varphi_z^\lambda(\bar{z}) K^\varepsilon(\bar{z} - z^{\text{var}}) d\bar{z}.$$

One can check that Assumption 3 is satisfied for this diagram, because it contains only one edge with the labels  $a = 3$  and  $r = 0$ ,  $\Gamma = \widehat{\mathbb{V}}_{\text{var}} = \{1\}$ , the contraction  $\gamma$  is trivial with only one component  $\gamma_1 = \{1\}$ , and  $|\widehat{\mathbb{V}}_* \setminus \{v_*^\uparrow\}| = 1$ . The space-time scaling is  $\mathfrak{s} = (2, 1, 1, 1)$ , so that  $|\mathfrak{s}| = 5$  and the value of the constant  $v_\gamma$  in (4.15) is  $-\frac{1}{2}$ , where we used  $|\Gamma| = 1$ .

Then we can compute the quantities appearing in (4.16). There is only one function  $\mathbf{p} \in \mathcal{P}_1$  satisfying the conditions of the first sum in (4.16), this is the function  $\mathbf{p}(1) = 2$ . Using this function, the first term on the right-hand side of (4.16) is  $C\lambda^{-\frac{1}{2}}$ , because  $\alpha_\gamma(\mathbf{p}) = \delta_\gamma(\mathbf{p}) = 0$  as follows from (3.19) and (4.17).

The only function  $\mathbf{p} \in \mathcal{P}_1$  satisfying the conditions of the second sum in (4.16) is the function  $\mathbf{p}(1) = \infty$ . Then the second term on the right-hand side of (4.16) is  $C\lambda^{-\frac{1}{2}}\varepsilon^{\frac{5}{2}-\theta}e^{-\frac{5}{2}}$  for any  $\theta > 0$ , because  $\alpha_\gamma(\mathbf{p}) = \frac{5}{2}$  and  $\delta_\gamma(\mathbf{p}) = \frac{5}{2}$  as follows from (3.19) and (4.17) respectively.

Hence, applying Corollary 4.5 and recalling that  $|\tau| = -\frac{1}{2} - \kappa$ , one obtains the bound

$$\left( \mathbf{E} |\iota_\varepsilon(\widehat{\Pi}_z^\varepsilon \tau)(\varphi_z^\lambda)|^p \right)^{\frac{1}{p}} \lesssim \lambda^{-\frac{1}{2}} (1 + \varepsilon^{\frac{5}{2}-\theta} e^{-\frac{5}{2}}) \lesssim \lambda^{-\frac{1}{2}},$$

for any  $\theta > 0$  small enough. As such, we immediately get (5.8) for the element  $\tau$ .

In what follows we always have  $|\mathfrak{s}| = 5$  and we prefer not to recall it every time.

### 5.3 The element $\tau = \Psi^2$

Taking into account the renormalisation in (5.4), the map  $\widehat{\Pi}_z^\varepsilon \tau$  can be represented by the diagrams

$$\iota_\varepsilon(\widehat{\Pi}_z^\varepsilon \tau)(\varphi_z^\lambda) = \begin{array}{c} \text{Diagram 1: } \text{Two blue circles at } z \text{ with } 3.0 \text{ labels, connected by a green arrow pointing up.} \\ \text{Diagram 2: } \text{Two blue circles at } z \text{ with } 3.0 \text{ labels, connected by a green arrow pointing up, with a blue circle at } z \text{ with } 3.0 \text{ label.} \\ \text{Diagram 3: } \text{A green arrow pointing up from } z \text{ to a black dot at } z. \end{array} - C_1^\varepsilon, \quad (5.12)$$

where the renormalisation constant is given in (5.5). We denote the first diagram by  $\iota_\varepsilon(\widehat{\Pi}_z^{\varepsilon, 1} \tau)(\varphi_z^\lambda)$  and the difference of the other two terms by  $\iota_\varepsilon(\widehat{\Pi}_z^{\varepsilon, 2} \tau)(\varphi_z^\lambda)$ .

Let us start with the first diagram in (5.12). Assumption 3 is satisfied for it with a trivial contraction  $\gamma$  having two components  $\gamma_1 = \{1\}$  and  $\gamma_2 = \{2\}$ , and the sets  $\Gamma = \widehat{\mathbb{V}}_{\text{var}} = \{1, 2\}$ . Furthermore, we have  $|\widehat{\mathbb{V}}_* \setminus \{v_*^\uparrow\}| = 2$  and the value of the constant  $v_\gamma$  in (4.15) is  $-1$ .

Now, we can compute the quantities appearing in the bound (4.16). Only the function  $\mathbf{p} \in \mathcal{P}_2$  with  $\mathbf{p}(1) = \mathbf{p}(2) = 2$  contributes to the first sum in (4.16). Then, according to (3.19) and (4.17), we have  $\alpha_\gamma(\mathbf{p}) = \delta_\gamma(\mathbf{p}) = 0$  and the first term on the right-hand side of (4.16) equals  $C\lambda^{-1}$ .

There are three functions  $\mathbf{p} \in \mathcal{P}_2$  contributing to the second sum in (4.16) which we denote by  $\mathbf{p}_1$ ,  $\mathbf{p}_2$  and  $\mathbf{p}_3$ . These functions are given by  $\mathbf{p}_1(1) = 2$ ,  $\mathbf{p}_1(2) = \infty$ ,  $\mathbf{p}_2(1) = \infty$ ,  $\mathbf{p}_2(2) = 2$ , and  $\mathbf{p}_3(1) = \mathbf{p}_3(2) = \infty$ . According to (3.19) and (4.17), we have  $\alpha_\gamma(\mathbf{p}_1) = \alpha_\gamma(\mathbf{p}_2) = \delta_\gamma(\mathbf{p}_1) = \delta_\gamma(\mathbf{p}_2) = \frac{5}{2}$  and  $\alpha_\gamma(\mathbf{p}_3) = \delta_\gamma(\mathbf{p}_3) = 5$ . Then the second term on the right-hand side equals  $C\lambda^{-1}(2\varepsilon^{\frac{5}{2}-\theta}e^{-\frac{5}{2}} + \varepsilon^{5-\theta}e^{-5})$  for any  $\theta > 0$ . Hence, applying Corollary 4.5 to this diagram, we get the bound

$$\left( \mathbf{E} |\iota_\varepsilon(\widehat{\Pi}_z^{\varepsilon,1}\tau)(\varphi_z^\lambda)|^p \right)^{\frac{1}{p}} \lesssim \lambda^{-1} (1 + \varepsilon^{\frac{5}{2}-\theta}e^{-\frac{5}{2}} + \varepsilon^{5-\theta}e^{-5}) \lesssim \lambda^{-1},$$

for any  $\theta > 0$  small enough. Recalling that  $|\tau| = -1 - 2\kappa$ , we get the bound (5.8).

The second diagram in (5.12) does not satisfy Assumption 3 and it needs to be considered as a renormalised integral in the sense of (3.29). For this, let us denote by “ $\diamond$ ” the variable which is integrated with respect to the martingales  $t \mapsto (\varepsilon^{\frac{5}{2}}\bar{\mathbb{M}}_\varepsilon(t, x))_{x \in \Lambda_\varepsilon}$ , where  $\bar{\mathbb{M}}_\varepsilon(t, x) := \varepsilon^{\frac{1}{2}}([\mathbb{M}_\varepsilon(x)]_t - \langle \mathbb{M}_\varepsilon(x) \rangle_t)$ . We use the multiplier  $\varepsilon^{\frac{1}{2}}$  to be consistent with our definition (2.6), and the multiplier  $\varepsilon^{d+k} = \varepsilon^{\frac{5}{2}}$  equals the  $\varepsilon$ -factor in the definition (3.29) with  $n = 2$ . Then the identity (2.22) and our definition of the renormalisation constant (5.5) allow to write

$$\iota_\varepsilon(\widehat{\Pi}_z^{\varepsilon,2}\tau)(\varphi_z^\lambda) = \begin{array}{c} \diamond \\ \nearrow 3,0 \quad \searrow 3,0 \\ \bullet \\ \uparrow \\ z \end{array}.$$

This expression equals to a renormalised integral (3.29) with  $n = 2$ . More generally, it is equal to an integral (3.30) with the contraction  $\gamma$  having one component  $\gamma_1 = \{1, 2\}$  which is labeled with  $\diamond$ . However, this diagram still does not satisfy Assumption 3, because the kernels have very strong singularities. To resolve this problem, we notice that multiplication of a kernel by a positive power of  $\varepsilon$  decreases the order of singularity in (4.7). Hence, for any  $0 < a < 3$  we can write the preceding expression as

$$\varepsilon^{2(a-3)} \begin{array}{c} \diamond \\ \nearrow a,0 \quad \searrow a,0 \\ \bullet \\ \uparrow \\ z \end{array}, \quad (5.13)$$

where we multiplied each kernel by  $\varepsilon^{3-a}$ . For  $\frac{5}{4} < a < \frac{5}{2}$ , Assumption 3 is satisfied with the sets  $\Gamma = \widehat{\mathbb{V}}_{\text{var}} = \{1\}$ . Furthermore, we have  $|\widehat{\mathbb{V}}_\star \setminus \{\nu_\star^\uparrow\}| = 1$  and the value of the constant  $\nu_\gamma$  in (4.15) is  $\frac{5}{2} - 2a < 0$ .

To apply the bound (4.16) to the expression in the parentheses in (5.13), we need to compute the involved quantities. There is one function  $\mathbf{p} \in \mathcal{P}_1$  contributing to the first sum on the right-hand side of (4.16), and this function is  $\mathbf{p}(1) = 2$ . Using (3.19) and (4.17), we compute  $\alpha_\gamma(\mathbf{p}) = \frac{5}{2}$  and  $\delta_\gamma(\mathbf{p}) = 0$ , and the first term on the right-hand side of (4.16) equals  $C\lambda^{\frac{5}{2}-2a}\varepsilon^{\frac{5}{2}}$ .

The only function  $\mathbf{p} \in \mathcal{P}_1$  contributing to the second sum on the right-hand side of (4.16) is  $\mathbf{p}(1) = \infty$ . Using (3.19) and (4.17), we compute  $\alpha_\gamma(\mathbf{p}) = 5$  and  $\delta_\gamma(\mathbf{p}) = \frac{5}{2}$ , and the second term on the right-hand side of (4.16) equals  $C\lambda^{\frac{5}{2}-2a}\varepsilon^{5-\theta}\mathfrak{e}^{-\frac{5}{2}}$  for any  $\theta > 0$ .

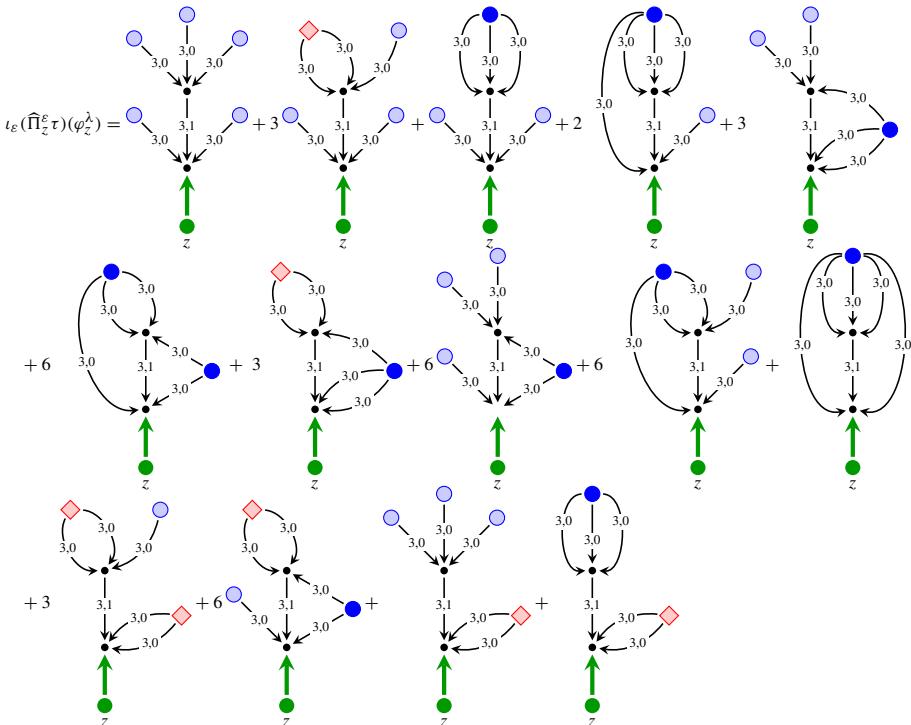
Hence, Corollary 4.5 yields

$$\left(\mathbf{E} \left| \iota_{\ell}(\widehat{\Pi}_z^{\varepsilon,2} \tau)(\varphi_z^\lambda) \right|^p\right)^{\frac{1}{p}} \lesssim \varepsilon^{2(a-3)} \lambda^{\frac{5}{2}-2a} \left( \varepsilon^{\frac{5}{2}} + \varepsilon^{5-\theta} \mathfrak{e}^{-\frac{5}{2}} \right) \lesssim \varepsilon^{2a-\frac{7}{2}-\theta} \lambda^{\frac{5}{2}-2a},$$

for any  $\theta > 0$ . If we take  $a > \frac{7}{4} + \theta$ , then the last expression becomes  $\varepsilon^\theta \lambda^{-1-2\theta}$ . Recalling that  $|\tau| = -1 - 2\kappa$  and taking  $\theta$  small enough, we get the bound (5.8). Moreover, this quantity vanishes as  $\varepsilon \rightarrow 0$ .

## 5.4 The element $\tau = \mathcal{I}(\Psi^3)\Psi^2$

Using the definition (5.6) and the renormalised integrals as in Sect. 5.3, the diagrams for the map  $\hat{\Pi}_z^\varepsilon \tau$  are the following:





All these diagrams, except the tenth and the last two, can be bounded by a direct application of Corollary 4.5.

As a demonstration, let us bound the third diagram, which we denote by  $\iota_{\varepsilon}(\widehat{\Pi}_z^{\varepsilon,3}\tau)(\varphi_z^{\lambda})$ . Let us label the vertices  $\mathbb{V}_{\text{var}}$  of the first diagram in (5.14) with  $1, \dots, 5$ , going in the clockwise direction. Then the third diagram corresponds to the contraction  $\gamma$  with three components  $\gamma_1 = \{1\}$ ,  $\gamma_2 = \{2, 3, 4\}$  and  $\gamma_3 = \{5\}$ . After performing contraction, we get three vertices in  $\widehat{\mathbb{V}}_{\text{var}}$  which we label with 1, 2 and 3. Then Assumption 3 is satisfied with  $\Gamma = \{1, 3\}$ . Furthermore, we have  $|\widehat{\mathbb{V}}_{\star} \setminus \{v_{\star}^{\uparrow}\}| = 4$  and the value of the constant  $\nu_{\gamma}$  in (4.15) is  $-3$ .

Now, we will compute the quantities which appear in the bounds (4.16). There is only one function  $\mathbf{p} \in \mathcal{P}_3$  contributing to the first sum in (4.16), which is the function  $\mathbf{p}(1) = 2$ ,  $\mathbf{p}(2) = 1$  and  $\mathbf{p}(3) = 2$ . Then (3.19) and (4.17) yield  $\alpha_{\gamma}(\mathbf{p}) = \frac{5}{2}$  and  $\delta_{\gamma}(\mathbf{p}) = 0$ , and the first term on the right-hand side of (4.16) equals  $C\lambda^{-3}\varepsilon^{\frac{5}{2}}$ . Since  $\lambda \geq \varepsilon$ , we can estimate it by  $C\lambda^{-\frac{1}{2}-\theta}\varepsilon^{\theta}$  for any  $\theta \in (0, \frac{5}{2})$ .

There are four functions  $\mathbf{p} \in \mathcal{P}_3$  contributing to the second sum in (4.16), which we denote by  $\mathbf{p}_1$ ,  $\mathbf{p}_2$ ,  $\mathbf{p}_3$  and  $\mathbf{p}_4$ , and which are given by  $\mathbf{p}_1(1) = 2$ ,  $\mathbf{p}_1(2) = \infty$ ,  $\mathbf{p}_1(3) = 2$ ,  $\mathbf{p}_2(1) = \infty$ ,  $\mathbf{p}_2(2) = \infty$ ,  $\mathbf{p}_2(3) = 2$ ,  $\mathbf{p}_3(1) = 2$ ,  $\mathbf{p}_3(2) = \infty$ ,  $\mathbf{p}_3(3) = \infty$ ,  $\mathbf{p}_4(1) = \infty$ ,  $\mathbf{p}_4(2) = \infty$  and  $\mathbf{p}_4(3) = \infty$ . We can compute by (3.19) and (4.17) the quantities  $\alpha_{\gamma}(\mathbf{p}_1) = \frac{15}{2}$ ,  $\delta_{\gamma}(\mathbf{p}_1) = 5$ ,  $\alpha_{\gamma}(\mathbf{p}_2) = \alpha_{\gamma}(\mathbf{p}_3) = 10$ ,  $\delta_{\gamma}(\mathbf{p}_2) = \delta_{\gamma}(\mathbf{p}_3) = \frac{15}{2}$ ,  $\alpha_{\gamma}(\mathbf{p}_4) = \frac{25}{2}$  and  $\delta_{\gamma}(\mathbf{p}_4) = 10$ . Then the second term on the right-hand side of (4.16) equals

$$C\lambda^{-3}\left(\varepsilon^{\frac{15}{2}-\theta}\mathfrak{e}^{-5} + 2\varepsilon^{10-\theta}\mathfrak{e}^{-\frac{15}{2}} + \varepsilon^{\frac{25}{2}-\theta}\mathfrak{e}^{-10}\right),$$

for any  $\theta > 0$ . Recalling that  $\lambda \geq \mathfrak{e} \geq \varepsilon$  and taking  $\theta$  small enough, we bound this expression by  $4C\lambda^{-\frac{1}{2}-2\theta}\varepsilon^{\theta}$ .

Hence, applying Corollary 4.5, we get

$$\left(\mathbf{E}|\iota_{\varepsilon}(\widehat{\Pi}_z^{\varepsilon,3}\tau)(\varphi_z^{\lambda})|^p\right)^{\frac{1}{p}} \lesssim \lambda^{-\frac{1}{2}-2\theta}\varepsilon^{\theta},$$

for all  $\theta > 0$  small enough. Taking a suitable value of  $\theta$ , we get the required bound (5.8) with  $|\tau| = -\frac{1}{2} - 5\kappa$ .

Now, we will bound the tenth diagram (the one that contracts all leaves), which we denote by  $\iota_{\varepsilon}(\widehat{\Pi}_z^{\varepsilon,10}\tau)(\varphi_z^{\lambda})$ . For this we write the product measure  $\mathbf{M}_{\varepsilon}^5$ , defined in (2.11), as

$$\mathbf{M}_\varepsilon^5([0, t] \times A) = \varepsilon^{15} \sum_{x \in A} \sum_{0 \leq s \leq t} (\Delta_s \mathbb{M}_\varepsilon(x))^5 = \varepsilon^{13} \sum_{x \in A} \sum_{0 \leq s \leq t} \Delta_s \mathbb{M}_\varepsilon(x),$$

where  $\gamma \in \mathfrak{C}_1(\mathbb{V}_{\text{var}})$  with  $|\mathbb{V}_{\text{var}}| = |\gamma_1| = 5$  and  $[0, t] \times A \subseteq \mathcal{D}_\gamma$ , where the set  $\mathcal{D}_\gamma$  is defined in (2.13), i.e. it is the diagonal of  $\mathcal{D}_\varepsilon^5$  on which all variables take the same values. The preceding expression equals  $\varepsilon^{13} \sum_{x \in A} \mathbb{M}_\varepsilon(t, x)$  because  $\mathbb{M}_\varepsilon$  is a piece-wise constant process. Then we can write

$$\begin{array}{c} \text{Diagram 1:} \\ \text{A circular diagram with a blue dot at the top. Inside, there are two nodes connected by a vertical line. The left node has a green arrow pointing up to a green circle labeled } z. \text{ The right node has a green arrow pointing down to a green circle labeled } z. \text{ The top arc has a blue dot at the top. The left arc has labels } 3,0, 3,0, 3,0 \text{ and } 3,1 \text{ on the left. The right arc has labels } 3,0, 3,0, 3,0 \text{ and } 3,1 \text{ on the right.} \end{array} = \varepsilon^{10} \begin{array}{c} \text{Diagram 2:} \\ \text{A circular diagram with a blue dot at the top. Inside, there are two nodes connected by a vertical line. The left node has a green arrow pointing up to a green circle labeled } z. \text{ The right node has a green arrow pointing down to a green circle labeled } z. \text{ The top arc has a blue dot at the top. The left arc has a blue dot at the top with a downward arrow labeled } 9,0. \text{ The right arc has labels } 3,0, 3,0, 3,0 \text{ and } 3,1 \text{ on the right.} \end{array},$$

where the stochastic integral is with respect to the martingale  $\mathbb{M}_\varepsilon$ . We can now view this diagram to correspond to the contraction  $\gamma$  with one component  $\gamma_1 = \{1\}$ . Taking powers of  $\varepsilon$  to improve the singularities of the kernels, similarly to how we did it in (5.13), we get

$$\varepsilon^{2a+b-5} \begin{array}{c} \text{Diagram 3:} \\ \text{A circular diagram with a blue dot at the top. Inside, there are two nodes connected by a vertical line. The left node has a green arrow pointing up to a green circle labeled } z. \text{ The right node has a green arrow pointing down to a green circle labeled } z. \text{ The top arc has a blue dot at the top. The left arc has a blue dot at the top with a downward arrow labeled } b,0. \text{ The right arc has labels } a,0, a,0, a,0 \text{ and } 3,1 \text{ on the right.} \end{array}, \quad (5.15)$$

for  $0 < a < 3$  and  $0 < b < 9$ . If  $a$  and  $b$  satisfy  $b < 5$  and  $\frac{7}{2} < 2a + b < 7$ , then Assumption 3 is satisfied with the sets  $\Gamma = \widehat{\mathbb{V}}_{\text{var}} = \{1\}$ . Furthermore, we have  $|\widehat{\mathbb{V}}_* \setminus \{u_*^\uparrow\}| = 2$  and the value of the constant  $v_\gamma$  in (4.15) equals  $\frac{9}{2} - 2a - b$ , which is strictly negative if  $2a + b > \frac{9}{2}$ .

Now, we are going to apply the bound (4.16) to the diagram (5.15), together with the multiplier  $\varepsilon^{2a+b-5}$ . There is one function  $\mathbf{p} \in \mathcal{P}_1$ , given by  $\mathbf{p}(1) = 2$ , contributing to the first sum on the right-hand side of (4.16). Using (3.19) and (4.17), we compute  $\alpha_\gamma(\mathbf{p}) = \delta_\gamma(\mathbf{p}) = 0$ , and the first term on the right-hand side of (4.16) equals  $C \varepsilon^{2a+b-5} \lambda^{\frac{9}{2}-2a-b}$ .

The only function  $\mathbf{p} \in \mathcal{P}_1$  contributing to the second sum on the right-hand side of (4.16) is  $\mathbf{p}(1) = \infty$ . Using (3.19) and (4.17), we compute  $\alpha_\gamma(\mathbf{p}) = \delta_\gamma(\mathbf{p}) = \frac{5}{2}$ , and the second term on the right-hand side of (4.16) equals  $C \varepsilon^{2a+b-5} \lambda^{\frac{9}{2}-2a-b} \varepsilon^{\frac{5}{2}-\theta} \varepsilon^{-\frac{5}{2}} \leq C \varepsilon^{2a+b-5-\theta} \lambda^{\frac{9}{2}-2a-b}$  for any  $\theta > 0$ .

Then Corollary 4.5 yields

$$\left( \mathbf{E} |\iota_\varepsilon(\widehat{\Pi}_z^{\varepsilon, 10} \tau)(\varphi_z^\lambda)|^p \right)^{\frac{1}{p}} \lesssim \varepsilon^{2a+b-5-\theta} \lambda^{\frac{9}{2}-2a-b}.$$

If we take  $\theta > 0$  sufficiently small and  $2a + b = 5 + 2\theta$ , then the last expression turns to  $\varepsilon^\theta \lambda^{-\frac{1}{2}-2\theta}$ . Recalling that  $|\tau| = -\frac{1}{2} - 5\kappa$  and taking  $\theta$  small enough, we get the bound (5.8), and this quantity vanishes as  $\varepsilon \rightarrow 0$ .

We now turn our attention to the last two terms in (5.14), which are also the most interesting ones. Using the renormalised kernel (4.8), we can write

$$2 \begin{array}{c} \text{blue dot} \\ \text{blue dot} \\ \text{blue dot} \end{array} - C_2^\varepsilon = \left( 2 \begin{array}{c} \text{blue dot} \\ \text{blue dot} \\ \text{blue dot} \end{array} - C_2^\varepsilon \right) - 2 \begin{array}{c} \text{blue dot} \\ \text{blue dot} \\ \text{blue dot} \end{array}. \quad (5.16)$$

The last diagram can be further decomposed as

$$\begin{array}{c} \text{blue dot} \\ \text{blue dot} \\ \text{blue dot} \end{array} = \begin{array}{c} \text{red diamond} \\ \text{red diamond} \\ \text{red diamond} \end{array} + 2 \begin{array}{c} \text{blue dot} \\ \text{blue dot} \\ \text{blue dot} \end{array} + \begin{array}{c} \text{blue dot} \\ \text{blue dot} \\ \text{blue dot} \end{array} \quad (5.17)$$

and applying Corollary 4.5 with  $\nu_\gamma = -\frac{11}{2}$ , the  $p$ -th moment of each term can be bounded by a constant times  $\varepsilon^\theta \lambda^{-\frac{1}{2}-2\theta}$  for any  $\theta > 0$ . To obtain this bound, we need to perform computations similar to what we did above.

The expression in the parentheses in (5.16) can be written as

$$2 \begin{array}{c} \text{blue dot} \\ \text{blue dot} \\ \text{blue dot} \end{array} - C_2^\varepsilon = 2 \begin{array}{c} \text{red diamond} \\ \text{red diamond} \\ \text{red diamond} \end{array} + 4 \begin{array}{c} \text{blue dot} \\ \text{blue dot} \\ \text{blue dot} \end{array} + \left( 2 \begin{array}{c} \text{blue dot} \\ \text{blue dot} \\ \text{blue dot} \end{array} - C_2^\varepsilon \right), \quad (5.17)$$

The first two diagrams above can again be bounded using Corollary 4.5 by a constant multiple of  $\varepsilon^\theta \lambda^{-\frac{1}{2}-2\theta}$  for  $\theta > 0$ .

The last expression in the parentheses in (5.17) needs more attention. Let us define a new kernel

$$\mathcal{G}_\varepsilon(z_1, z_2) := \bullet \leftarrow^{3,0} \bullet \downarrow^{3,0} \bullet \leftarrow^{3,0} \bullet \uparrow^{3,0} \bullet, \quad z_1 \quad z_2$$

which is in fact a function of the difference of the arguments, i.e.  $\mathcal{G}_\varepsilon(z_1, z_2) = \mathcal{G}_\varepsilon(z_2 - z_1)$ . As follows from the order of the singularity of the kernel  $K^\varepsilon$  and [29, Lemma 7.3], the function  $\mathcal{G}_\varepsilon$  satisfies  $|D^k \mathcal{G}_\varepsilon(z)| \lesssim (\|z\|_s + \varepsilon)^{-5-|k|_s}$  for all multiindices  $k$  with  $|k|_s$  large enough. Then we conclude from Lemma 4.2 that the function  $\mathcal{G}_\varepsilon$  has all the properties listed in Assumption 4 with the values  $a_e = 5$  and  $r_e = 0$ . We denote the kernel  $\mathcal{G}_\varepsilon$  by an edge “ $\rightsquigarrow^{5,0}$ ”. Then the first diagram in the parentheses in (5.17) can be represented as

$$\bullet \rightsquigarrow^{5,0} \bullet \rightsquigarrow^{5,0} \bullet \leftarrow^{3,0} \bullet,$$

and one can see that this diagram does not satisfy Assumption 3(1.) (recall that  $|s| = 5$ ). To resolve this problem, we need to use a negative renormalisation (in the sense of Sect. 4.1.1) of the kernel  $\mathcal{G}_\varepsilon$ . More precisely, for any smooth function  $\eta : \mathbf{R}^4 \times \mathbf{R}^4 \rightarrow \mathbf{R}$ , we define its negative renormalisation as

$$(\mathcal{R}_\varepsilon \mathcal{G}_\varepsilon)(\eta) := \int_{D_\varepsilon} \int_{D_\varepsilon} \mathcal{G}_\varepsilon(z_1, z_2) (\eta(z_1, z_2) - \eta(z_1, z_1)) dz_1 dz_2,$$

and we graphically depict  $\mathcal{R}_\varepsilon \mathcal{G}_\varepsilon$  as “ $\rightsquigarrow^{5,-1}$ ”, where the label “ $-1$ ” refers to the order of renormalisation. Since the renormalisation constant (5.7) can be represented as

$$C_2^\varepsilon = 2 \bullet \leftarrow^{3,0} \bullet \downarrow^{3,0} \bullet \leftarrow^{3,0} \bullet \uparrow^{3,0} \bullet, \quad 0$$

the expression in the parentheses in (5.17) equals

$$2 \bullet \rightsquigarrow^{5,-1} \bullet \leftarrow^{3,0} \bullet,$$

This diagram satisfies Assumption 3, and using Corollary 4.5 we bound its moments by a constant multiple of  $\varepsilon^\theta \lambda^{-\frac{1}{2}-2\theta}$  for  $\theta > 0$ .

By taking  $\theta$  in the preceding bounds sufficiently small, we obtain the required bound (5.8) on the expression (5.16).

**Acknowledgements** PG was supported by a scholarship from the EPSRC Centre for Doctoral Training in Statistical Applied Mathematics at Bath (SAMBa), under the project EP/L015684/1.

**Author Contributions** All authors were involved in writing all parts of the manuscript.

**Funding** KM was partially supported by NSF grant DMS-2321493. HW was supported by the Royal Society through the University Research Fellowship UF140187, by the Leverhulme Trust through a Philip Leverhulme Prize and by the European Union (ERC, GE4SPDE, 101045082). HW acknowledges funding by the Deutsche Forschungsgemeinschaft under Germany's Excellence Strategy EXC 2044 390685587, Mathematics Münster: Dynamics–Geometry–Structure. PG and HW thank the Isaac Newton Institute for Mathematical Sciences for hospitality during the programme *Scaling limits, rough paths, quantum field theory*, which was supported by EPSRC Grant No. EP/R014604/1.

## References

1. Aldous, D.: Stopping times and tightness. *Ann. Probab.* **6**(2), 335–340 (1978)
2. Bernardin, C., Funaki, T., Sethuraman, S.: Derivation of coupled KPZ–Burgers equation from multi-species zero-range processes. *Ann. Appl. Probab.* **31**(4), 1966–2017 (2021)
3. Bertini, L., Giacomin, G.: Stochastic Burgers and KPZ equations from particle systems. *Commun. Math. Phys.* **183**(3), 571–607 (1997)
4. Billingsley, P.: Convergence of Probability Measures. Wiley Series in Probability and Statistics: Probability and Statistics, 2nd edn. Wiley, New York (1999)
5. Bertini, L., Presutti, E., Rüdiger, B., Saada, E.: Dynamical fluctuations at the critical point: convergence to a nonlinear stochastic PDE. *Teor. Veroyatnost. i Primenen.* **38**(4), 689–741 (1993)
6. Chandra, A., Hairer, M.: An analytic BPHZ theorem for regularity structures. [arXiv:1612.08138](https://arxiv.org/abs/1612.08138) (2016)
7. Chandra, A., Shen, H.: Moment bounds for SPDEs with non-Gaussian fields and application to the Wong–Zakai problem. *Electron. J. Probab.* **22**, Paper No. 68, 32 (2017)
8. Corwin, I., Shen, H., Tsai, L.-C.: ASEP( $q, j$ ) converges to the KPZ equation. *Ann. Inst. Henri Poincaré Probab. Stat.* **54**(2), 995–1012 (2018)
9. Corwin, I., Tsai, L.-C.: KPZ equation limit of higher-spin exclusion processes. *Ann. Probab.* **45**(3), 1771–1798 (2017)
10. Chandra, A., Weber, H.: Stochastic PDEs, regularity structures, and interacting particle systems. *Ann. Fac. Sci. Toulouse Math.* (6) **26**(4), 847–909 (2017)
11. Da Prato, G., Zabczyk, J.: Stochastic Equations in Infinite Dimensions, vol. 44 of Encyclopedia of Mathematics and Its Applications. Cambridge University Press (1992)
12. Dembo, A., Tsai, L.-C.: Weakly asymmetric non-simple exclusion process and the Kardar–Parisi–Zhang equation. *Commun. Math. Phys.* **341**(1), 219–261 (2016)
13. Erhard, D., Hairer, M.: Discretisation of regularity structures. *Ann. Inst. Henri Poincaré Probab. Stat.* **55**(4), 2209–2248 (2019)
14. Erhard, D., Hairer, M.: A scaling limit of the parabolic Anderson model with exclusion interaction (2021). [arXiv:2103.13479](https://arxiv.org/abs/2103.13479)
15. Friz, P.K., Hairer, M.: A Course on Rough Paths. Universitext, 2nd edn. Springer, Cham (2020)
16. Fritz, J., Rüdiger, B.: Time dependent critical fluctuations of a one-dimensional local mean field model. *Probab. Theory Relat. Fields* **103**(3), 381–407 (1995)
17. Gubinelli, M., Imkeller, P., Perkowski, N.: Paracontrolled distributions and singular PDEs. *Forum Math. Pi* **3**, e6, 75 (2015)
18. Gubinelli, M., Jara, M.: Regularization by noise and stochastic Burgers equations. *Stoch. Partial Differ. Equ. Anal. Comput.* **1**(2), 325–350 (2013)
19. Gonçalves, P., Jara, M.: Nonlinear fluctuations of weakly asymmetric interacting particle systems. *Arch. Ration. Mech. Anal.* **212**(2), 597–644 (2014)
20. Giacomin, G., Lebowitz, J.L., Presutti, E.: Deterministic and stochastic hydrodynamic equations arising from simple microscopic model systems. In: Stochastic Partial Differential Equations: Six Perspectives, vol. 64 of Mathematical Surveys and Monographs, pp. 107–152. American Mathematical Society, Providence, RI (1999)
21. Grazieschi, P., Matetski, K., Weber, H.: The dynamical Ising–Kac model in 3D converges to  $\Phi_3^4$ . [arXiv:2303.10242](https://arxiv.org/abs/2303.10242) (2023)

22. Gubinelli, M., Perkowski, N.: KPZ reloaded. *Commun. Math. Phys.* **349**(1), 165–269 (2017)
23. Gubinelli, M., Perkowski, N.: Energy solutions of KPZ are unique. *J. Am. Math. Soc.* **31**(2), 427–471 (2018)
24. Gubinelli, M., Perkowski, N.: The infinitesimal generator of the stochastic Burgers equation. *Probab. Theory Relat. Fields* **178**(3–4), 1067–1124 (2020)
25. Gonçalves, P., Perkowski, N., Simon, M.: Derivation of the stochastic Burgers equation with Dirichlet boundary conditions from the WASEP. *Ann. H. Lebesgue* **3**, 87–167 (2020)
26. Hairer, M.: A theory of regularity structures. *Invent. Math.* **198**(2), 269–504 (2014)
27. Hairer, M.: Introduction to regularity structures. *Braz. J. Probab. Stat.* **29**(2), 175–210 (2015)
28. Hall, P., Heyde, C.C.: *Martingale Limit Theory and Its Application*. Academic Press Inc. [Harcourt Brace Jovanovich, Publishers], New York-London (1980) (Probability and Mathematical Statistics)
29. Hairer, M., Matetski, K.: Discretisations of rough stochastic PDEs. *Ann. Probab.* **46**(3), 1651–1709 (2018)
30. Hairer, M., Quastel, J.: A class of growth models rescaling to KPZ. *Forum Math. Pi* **6**, e3, 112 (2018)
31. Hairer, M., Steele, R.: The BPHZ theorem for regularity structures via the spectral gap inequality (2023). [arXiv:2301.10081](https://arxiv.org/abs/2301.10081)
32. Iberti, M.: Convergence of Glauber dynamic on Ising-like models with Kac interaction to  $\varphi_2^{2n}$  (2017). [arXiv:1708.00948](https://arxiv.org/abs/1708.00948)
33. Jara, M., MorenoFlores, G.R.: Scaling of the Sasamoto–Spohn model in equilibrium. *Electron. Commun. Probab.* **24**, Paper No. 3, 12 (2019)
34. Jacod, J., Shiryaev, A.N.: *Limit Theorems for Stochastic Processes*. Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 288, 2nd edn. Springer, Berlin (2003)
35. Kipnis, C., Landim, C.: *Scaling Limits of Interacting Particle Systems*. Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 320. Springer, Berlin (1999)
36. Kuo, H.: *Introduction to Stochastic Integration*. Universitext. Springer, New York (2005)
37. Linares, P., Otto, F., Tempelman, M., Tsatsoulis, P.: A diagram-free approach to the stochastic estimates in regularity structures. [arXiv:2112.10739](https://arxiv.org/abs/2112.10739) (2021)
38. Mitoma, I.: Tightness of probabilities on  $C([0, 1]; \mathcal{S}')$  and  $D([0, 1]; \mathcal{S}')$ . *Ann. Probab.* **11**(4), 989–999 (1983)
39. Martin, J., Perkowski, N.: Paracontrolled distributions on Bravais lattices and weak universality of the 2d parabolic Anderson model. *Ann. Inst. Henri Poincaré Probab. Stat.* **55**(4), 2058–2110 (2019)
40. Mourrat, J.-C., Weber, H.: Convergence of the two-dimensional dynamic Ising–Kac model to  $\Phi_2^4$ . *Commun. Pure Appl. Math.* **70**(4), 717–812 (2017)
41. Nualart, D.: *The Malliavin Calculus and Related Topics*. Probability and Its Applications. Springer, Berlin (2006)
42. Shen, H., Weber, H.: Glauber dynamics of 2D Kac–Blume–Capel model and their stochastic PDE limits. *J. Funct. Anal.* **275**(6), 1321–1367 (2018)

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.