The Annals of Probability
2021, Vol. 49, No. 2, 944–996
https://doi.org/10.1214/20-AOP1465

PROPAGATION OF CHAOS FOR MEAN FIELD ROUGH DIFFERENTIAL EQUATIONS

BY ISMAËL BAILLEUL[1], RÉMI CATELLIER[2,*] AND FRANÇOIS DELARUE[2,†]

[1]*Université Rennes 1, CNRS, IRMAR, ismael.bailleul@univ-rennes1.fr*
[2]*Université Côte d'Azur, CNRS, LJAD, *remi.catellier@unice.fr; †francois.delarue@unice.fr*

We address propagation of chaos for large systems of rough differential equations associated with random rough differential equations of mean field type

$$dX_t = \mathrm{V}(X_t, \mathcal{L}(X_t))\,dt + \mathrm{F}(X_t, \mathcal{L}(X_t))\,dW_t,$$

where W is a random rough path and $\mathcal{L}(X_t)$ is the law of X_t. We prove propagation of chaos, and provide also an explicit optimal convergence rate. The analysis is based upon the tools we developed in our companion paper (*Electron. J. Probab.* **25** (2020) 21) for solving mean field rough differential equations and in particular upon a corresponding version of the Itô-Lyons continuity theorem. The rate of convergence is obtained by a coupling argument developed first by Sznitman for particle systems with Brownian inputs.

CONTENTS

1. Introduction. The study of mean field stochastic dynamics and interacting diffusions / Markov processes finds its roots in Kac's simplified approach to kinetic theory [36] and McKean's work [41] on nonlinear parabolic equations. It provides the description of evolutions $(\mu_t)_{t\geq 0}$ in the space of probability measures under the form of a pathspace random

Received June 2019; revised June 2020.
MSC2020 subject classifications. 60H10, 60G99.
Key words and phrases. Random rough differential equations, particle system, mean field interaction, propagation of chaos, convergence rate.

dynamics

$$\begin{aligned} dX_t(\omega) &= \mathrm{V}\big(X_t(\omega), \mu_t\big)\,dt + \mathrm{F}\big(X_t(\omega), \mu_t\big)\,dW_t(\omega), \\ \mu_t &:= \mathcal{L}(X_t), \end{aligned} \tag{1.1}$$

where $\mathcal{L}(A)$ stands for the law of a random variable A over a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ containing ω and relates it to the empirical behaviour of large systems of interacting dynamics. The main emphasis of subsequent works has been on proving propagation of chaos and other limit theorems, and giving stochastic representations of solutions to nonlinear parabolic equations under more and more general settings; see, for instance, [10, 11, 23, 24, 31, 35, 42, 47, 48]. Classical stochastic calculus makes sense of equation (1.1) only when the process W is a semimartingale under $\mathbb{P}$, for some filtration, and the integrand is predictable. However, this setting happens to be too restrictive in a number of situations, especially when the diffusivity is random. This prompted several authors to address equation (1.1) by means of rough paths theory. Indeed, one may understand rough paths theory as a natural framework for providing probabilistic models of interacting populations, beyond the realm of Itô calculus. Cass and Lyons [18] did the first study of mean field random rough differential equations and proved the well-posed character of equation (1.1), and propagation of chaos for an associated system of interacting particles, under the crucial assumption that there is no mean field interaction in the diffusivity, that is, $\mathrm{F}(x, \mu) = \mathrm{F}(x)$, and that the drift depends linearly on the mean field interaction. Bailleul extended partly these results in [4] by proving well-posedness of the mean field rough differential equation (1.1) in the case where the drift depends nonlinearly on the interaction term and the diffusivity is still independent of the interaction, and by proving an existence result when the diffusivity depends on the interaction. Another breakthrough came with our earlier arXiv deposit [6], in which we explained how to handle the case when F truly depends on the interaction term by making a systematic use of Lions' approach to differential calculus on Wasserstein space. To make the content more accessible, we eventually decided to split [6] into two parts: While the current work is mainly inspired from the second half of [6], our companion article [5] corresponds to the first half of [6]; Therein, we address the well-posedness of the mean field rough equation (1.1) for a genuinely nonlinear F.

In fact, as explained in [5], the general case may be easily reduced to the study of the simpler equation

$$dX_t(\omega) = \mathrm{F}\big(X_t(\omega), \mathcal{L}(X_t)\big)\,dW_t(\omega), \tag{1.2}$$

which is precisely the version we address in this paper. To make it clear, the purpose of the present article is to prove that, under suitable assumptions, the solution of (1.2) coincides with the limit (in a convenient sense), as n tends to ∞, of the n-particle system

$$X_t^i(\omega) = X_0^i(\omega) + \int_0^t \mathrm{F}\left(X_s^i(\omega), \frac{1}{n}\sum_{j=1}^n \delta_{X_s^j(\omega)}\right) dW_s^i(\omega), \quad t \geq 0, \tag{1.3}$$

for $1 \leq i \leq n$, where $(X_0^i(\cdot), W^i(\cdot))_{1\leq i\leq n}$ is a collection of independent and identically distributed variables with the same distribution as $(X_0(\cdot), W(\cdot))$, the first component being regarded as a random variable with values in $\mathbb{R}^d$ and the second one as a random variable with values in the space of continuous functions. Of course, equation (1.3) must be understood as a rough differential equation driven by the signal $(W^1(\omega), \ldots, W^n(\omega))$ with $(X^1(\omega), \ldots, X^n(\omega))$ as output. As it is well known, this requires to lift $(W^1(\omega), \ldots, W^n(\omega))$ into an enhanced rough path $\boldsymbol{W}^{(n)}(\omega)$ and henceforth to define the various iterated integrals. Asking the paths $W(\omega)$, $\omega \in \Omega$, to have a finite p-variation for $2 \leq p < 3$, this prompts us to assume that, instead of $((X_0^1(\cdot), W^1(\cdot)), \ldots, (X_0^n, W^n(\cdot)))$, we have in fact n independent

copies $(X_0^i(\cdot), W^i(\cdot), \mathbb{W}^i(\cdot))_{1\le i\le n}$ of the triple $(X_0(\cdot), W(\cdot), \mathbb{W}(\cdot))$, where $\mathbb{W}(\omega)$ is the iterated integral of $W(\omega)$ and $\mathbb{W}^i(\omega)$ is the iterated integral of $W^i(\omega)$. Of course, it is also needed to define the iterated integrals of $W^j(\omega)$ with respect to $W^i(\omega)$, for $j \neq i$. Not only we assume below that such iterated integrals do indeed exist, but we make the additional assumption that there is a measurable map $\mathcal{I}$ giving $\mathbb{W}^{i,j}(\omega)$ from $W^i(\omega)$ and $W^j(\omega)$, that is

$$\mathbb{W}^{i,j}(\omega) = \mathcal{I}\big(W^i(\omega), W^j(\omega)\big), \quad i \neq j. \tag{1.4}$$

In words, (1.4) says that there exists a measurable way to construct the iterated integral of two independent copies of the signal in the limiting equation (1.2). Hence, (1.4) should be really regarded as an intrinsic property of (1.2) and not as a specific feature of the particle system (1.3).

More generally, it is in fact a key point in the subsequent analysis to draw a parallel between the underlying rough path used to give a meaning to (1.3) and the notion of *extended*[1] rough set-up used in [5] to address (1.2). We provide a reminder of the latter notion in Section 2. Basically, it says that, in order to solve (1.2), we must not only lift, for a given $\omega \in \Omega$, the trajectory $W(\omega)$ into an enhanced rough path $(W(\omega), \mathbb{W}(\omega))$, but we must in fact lift the whole trajectory $(W(\omega), W(\cdot))$, the second component being seen as a path with values in some $\mathbb{L}^q(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R}^m)$ space, where m is the dimension of the signal. Then, we call *extended* rough path set-up the enhancement of $(W(\omega), W(\cdot))$.

The striking fact of our analysis is then based upon an observation noticed first by Tanaka in his seminal work [49] on limit theorems for mean field type diffusions, and used crucially by Cass and Lyons in their seminal work [18]. We refer to it as Tanaka's trick. It says that, for a given $\omega \in \Omega$, the particle system (1.3) itself may be interpreted as a mean field equation, but with respect to the empirical measure of the driving noise. Adapted to the rough paths theory, it says that, for any fixed $\omega \in \Omega$, the path

$$\boldsymbol{W}^{(n)}(\omega) = \big(\big(W^i(\omega)\big)_{1\le i\le n}, \big(\mathbb{W}^{i,j}(\omega)\big)_{1\le i,j\le n}\big) =: \big(W^{(n)}(\omega), \mathbb{W}^{(n)}(\omega)\big),$$

which underpins the rough structure used to solve (1.3), may be seen as an extended rough set-up on its own—below, we just say a rough set-up) but defined on the finite probability space

$$\left(\{1,\dots,n\}, \mathcal{P}\big(\{1,\dots,n\}\big), \frac{1}{n}\sum_{i=1}^n \delta_i\right),$$

where $\mathcal{P}(\{1,\dots,n\})$ denotes the collection of subsets of $\{1,\dots,n\}$, instead of the former probability space $(\Omega, \mathcal{F}, \mathbb{P})$. We call this set-up the *empirical rough set-up*, and we make its construction entirely clear in the sequel of the paper. For sure, given the iterated integrals of the signal $(W^1(\omega), \dots, W^n(\omega))$, the rough integral (1.3) should be interpreted in the usual sense, as given by standard Lyons' rough paths theory. In short, this requires to expand locally the integrand in (1.3), which in turns requires to have a convenient notion of derivative with respect to the measure argument. In this regard, a crucial fact in [5] is to use Lions' approach [12, 14, 38] to differential calculus on the space $\mathcal{P}_2(\mathbb{R}^d)$ of probability measures on $\mathbb{R}^d$ with a finite square moment, the so-called d-dimensional Wasserstein space, d denoting here and throughout the dimension of the output in (1.2) (we refer to Section 2.5 for a longer discussion on the connection with other notions of derivatives on the Wasserstein space). The core of our analysis in Section 4 is that, whenever Wasserstein derivatives on $\mathcal{P}_2(\mathbb{R}^d)$ are projected, through empirical measures, into classical derivatives on $(\mathbb{R}^d)^n$, as it is needed

[1] In fact, the term *extended* does not appear in [5], but it is here of a convenient use to distinguish from the standard rough set-up used to solve the particle system (1.3).

to differentiate the integrand in (1.3), the resulting solution for (1.3), as given by standard rough paths theory, coincides with the solution obtained by interpreting (1.3) as a mean field rough equation driven by the aforementioned empirical rough set-up—see Section 3 for reminders on solvability results for mean field rough equations. In this way, the convergence of solutions of (1.3) to solutions of (1.2) as n tends to ∞ is reduced to a form of continuity of the solutions to mean field rough differential equations with respect to the underlying rough set-up. We called the latter *continuity of the Itô-Lyons solution map*, see Theorem 5.4 of our companion work [5]. Our first main result, Theorem 4.3, shows that, for a sufficiently large class of input signals, propagation of chaos is in fact a consequence of the *continuity of the Itô-Lyons solution map* for mean field rough differential equations. At this stage, it is worth mentioning that it is precisely in the requirements of the *continuity of the Itô-Lyons map* that the structure condition (1.4) about the cross-iterated integrals comes in. In [5], a *rough set-up* that satisfies (1.4) is said to be *strong*.

While the proofs of both our first main result and the underlying *continuity property of the Itô-Lyons solution map* are mostly based on compactness arguments, our second main result is to elucidate, under slightly stronger assumptions the convergence rate in the propagation of chaos; see Theorem 5.1 in Section 5. The strategy is directly inspired from original Sznitman's coupling argument for mean field systems driven by Brownian signals, see [47]. Although the proof is much more involved than in the Brownian setting, we recover the same rate of convergence: It coincides with the rate of convergence (in Wasserstein metric) of the empirical measure of an n-sample of (sufficiently integrable) i.i.d. variables to their common distribution. In particular, the speed decays with the dimension.

As in [5], our analysis holds for continuous rough paths whose p-variation, for some $p \in [2, 3)$, is finite and has sub-exponential tails and for which the so-called *local accumulated variation*—that counts the increments of the signal of a given size over a bounded interval—has super-exponential tails, see [5], Theorem 1.1. Among others, our results apply to continuous centred Gaussian signals defined over some time interval $[0, T]$ that have independent components and whose covariance function has finite ρ-two dimensional variation, for some $\rho \in [1, 3/2)$.

For the sake of completeness, it is worth adding that there has been a number of works recently using probabilistic tools to investigate mean field problems, or connecting probabilistic and analytical approaches. We emphasize Barbu and Röckner's study [8] of the solvability of McKean-Vlasov equations with local interactions by means of the superposition principle, and Coghi and Gess' work [21] on stochastic nonlinear nonlocal Fokker–Planck equations that arise in the mean field limit of weakly interacting diffusions with a common noise. Even though it is purely analytic, Jabin and Wang's work [33] on quantitative estimates of propagation of chaos for particle systems with a singular interaction kernel should be read at the light of a probabilistic intuition. The literature at the intersection of rough paths theory and mean field dynamics is sparser. Cass and Lyons launched the subject in [18] by proving a well-posedness result for a mean field rough differential equation with no interaction in the diffusivity and a linear interaction in the drift, and a propagation of chaos result. The well-posedness result was generalised by Bailleul under a different set of assumptions in [4] for mean field rough differential equations with no interaction in the diffusivity and a nonlinear interaction in the drift. The setting happens to be much simpler for equations with an additive noise, as no rough paths are needed and propagation of chaos, central limit theorem and large deviation results can be proved, as shown by Coghi, Deuschel, Friz and Maurelli in [19]. Coghi and Nilssen developed in [22] a variant of Bailleul and Riedel's approach of rough flows [7] to study rough nonlocal mean field type Fokker–Planck equations subjected to a rough common noise (meaning that, in [22], the process $(\mu_t)_{0\le t\le T}$ in (1.1) becomes random under the action of an additional noise that is precisely assumed to be a rough path;

obviously, there is no common noise in our model) but to an idiosyncratic Brownian noise (to make it clear, the idiosyncratic noise should be regarded as W in (1.1); of course, a substantial difference is that, in our framework, W may not be a Brownian motion). Last, Cass, dos Reis and Salkeld [16] used rough paths theory to investigate in depth support theorems for solutions of McKean Vlasov equations driven by a Brownian motion.

The present work leaves wide open the question of refining the strong law of large numbers given by the propagation of chaos result stated in Theorem 4.3. A central limit theorem for the fluctuations of the empirical measure of the particle system is expected to hold under reasonable conditions on the common law of the rough drivers. Large and moderate deviation results would also be most welcome. In a different direction, it would be interesting to investigate the propagation of chaos phenomenon for systems of interacting rough dynamics subject to a common noise. Very interesting things happen in the Itô setting in relation with mean field games [13, 37]. Also, one would get a more realistic model of natural phenomena by working with systems of particles driven by nonindependent noises. Individuals with close initial conditions could have drivers strongly correlated while individuals started far apart could have (almost-)independent drivers. Limit mean field dynamics are likely to be different from the results obtained here—see [20] for a result in this direction in the Itô setting. We invite the reader to make her/his own mind about these problems.

The paper is organized as follows. We recall in Section 2 the construction of a rough set-up, as introduced in [5]. We provide in Section 3 a sketchy presentation of related solvability results for equation (1.2), including a review of the main assumptions that we need on the diffusivity F. Convergence of the particle system (1.3) is established in Section 4. The convergence rate is addressed in Section 5, under additional regularity assumptions on F and integrability assumptions on the signal. Proofs of some technical results are given in Appendixes A.1 and A.2.

Notations. We gather here a number of notations that will be used throughout the text.

- We set $\mathcal{S}_2 := \{(s,t) \in [0,\infty)^2 : s \le t\}$, and $\mathcal{S}_2^T := \{(s,t) \in [0,T]^2 : s \le t\}$.
- We denote by $(\Omega, \mathcal{F}, \mathbb{P})$ an atomless Polish probability space, $\mathcal{F}$ standing for the completion of the Borel σ-field under $\mathbb{P}$, and denote by $\langle\cdot\rangle$ the expectation operator, by $\langle\cdot\rangle_r$, for $r \in [1,+\infty]$, the $\mathbb{L}^r$-norm on $(\Omega, \mathcal{F}, \mathbb{P})$ and by $\langle\langle\cdot\rangle\rangle$ and $\langle\langle\cdot\rangle\rangle_r$ the expectation operator and the $\mathbb{L}^r$-norm on $(\Omega^2, \mathcal{F}^{\otimes 2}, \mathbb{P}^{\otimes 2})$. When r is finite, $\mathbb{L}^r(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R})$ is separable as Ω is Polish.
- As for processes $X_\bullet = (X_t)_{t\in I}$, defined on a time interval I, we often write X for $X_\bullet$.

2. From probabilistic rough structures to rough integrals.

2.1. *Overview on probabilistic rough structures*. We here provide a brief reminder of the content of Section 2 in [5]. We refer the reader to the paper for a complete review. Throughout the section, we work on a finite time horizon $[0,T]$, for a given $T > 0$.

The first level of the rough path structure used to give a meaning to (1.2) is defined as an ω-indexed pair of paths

$$\big(W_t(\omega), W_t(\cdot)\big)_{0\le t\le T}, \tag{2.1}$$

where $(W_t(\cdot))_{0\le t\le T}$ is a collection of q-integrable $\mathbb{R}^m$-valued random variables on $(\Omega, \mathcal{F}, \mathbb{P})$, which we regard as a deterministic $\mathbb{L}^q(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R}^m)$-valued path, for some exponent $q \ge 8$, and $(W_t(\omega))_{0\le t\le T}$ stands for the realizations of these random variables along the outcome $\omega \in \Omega$; so the pair (2.1) takes values in $\mathbb{R}^m \times \mathbb{L}^q(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R}^m)$. The second level has the form

of an ω-dependent two-index path with values in $(\mathbb{R}^m \times \mathbb{L}^q(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R}^m))^{\otimes 2}$ and is encoded in matrix form as

$$\begin{pmatrix} \mathbb{W}_{s,t}(\omega) & \mathbb{W}^{\perp\!\!\!\perp}_{s,t}(\omega, \cdot) \\ \mathbb{W}^{\perp\!\!\!\perp}_{s,t}(\cdot, \omega) & \mathbb{W}^{\perp\!\!\!\perp}_{s,t}(\cdot, \cdot) \end{pmatrix}_{0 \le s \le t \le T}, \tag{2.2}$$

where:

- $\mathbb{W}_{s,t}(\omega)$ is in $(\mathbb{R}^m)^{\otimes 2} \simeq \mathbb{R}^{m \times m}$,
- $\mathbb{W}^{\perp\!\!\!\perp}_{s,t}(\omega, \cdot)$ is in $\mathbb{R}^m \otimes \mathbb{L}^q(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R}^m) \simeq \mathbb{L}^q(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R}^{m \times m})$,
- $\mathbb{W}^{\perp\!\!\!\perp}_{s,t}(\cdot, \omega)$ is in $\mathbb{L}^q(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R}^m) \otimes \mathbb{R}^m \simeq \mathbb{L}^q(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R}^{m \times m})$,
- $\mathbb{W}^{\perp\!\!\!\perp}_{s,t}(\cdot, \cdot)$ is in $\mathbb{L}^q(\Omega^{\otimes 2}, \mathcal{F}^{\otimes 2}, \mathbb{P}^{\otimes 2}; \mathbb{R}^{m \times m})$, the realizations of which read in the form $\Omega^2 \ni (\omega, \omega') \mapsto \mathbb{W}^{\perp\!\!\!\perp}_{s,t}(\omega, \omega') \in \mathbb{R}^{m \times m}$ and the two sections of which are precisely given by $\mathbb{W}^{\perp\!\!\!\perp}_{s,t}(\omega, \cdot) : \Omega \ni \omega' \mapsto \mathbb{W}^{\perp\!\!\!\perp}_{s,t}(\omega, \omega')$, and $\mathbb{W}^{\perp\!\!\!\perp}_{s,t}(\cdot, \omega) \ni \omega' \mapsto \mathbb{W}^{\perp\!\!\!\perp}_{s,t}(\omega', \omega)$, for $\omega \in \Omega$.

A convenient form of Chen's relations is required, for any $\omega \in \Omega$,

$$\begin{aligned} \mathbb{W}_{r,t}(\omega) &= \mathbb{W}_{r,s}(\omega) + \mathbb{W}_{s,t}(\omega) + W_{r,s}(\omega) \otimes W_{s,t}(\omega), \\ \mathbb{W}^{\perp\!\!\!\perp}_{r,t}(\cdot, \omega) &= \mathbb{W}^{\perp\!\!\!\perp}_{r,s}(\cdot, \omega) + \mathbb{W}^{\perp\!\!\!\perp}_{s,t}(\cdot, \omega) + W_{r,s}(\cdot) \otimes W_{s,t}(\omega), \\ \mathbb{W}^{\perp\!\!\!\perp}_{r,t}(\omega, \cdot) &= \mathbb{W}^{\perp\!\!\!\perp}_{r,s}(\omega, \cdot) + \mathbb{W}^{\perp\!\!\!\perp}_{s,t}(\omega, \cdot) + W_{r,s}(\omega) \otimes W_{s,t}(\cdot), \\ \mathbb{W}^{\perp\!\!\!\perp}_{r,t}(\cdot, \cdot) &= \mathbb{W}^{\perp\!\!\!\perp}_{r,s}(\cdot, \cdot) + \mathbb{W}^{\perp\!\!\!\perp}_{s,t}(\cdot, \cdot) + W_{r,s}(\cdot) \otimes W_{s,t}(\cdot), \end{aligned} \tag{2.3}$$

for any $0 \le r \le s \le t \le T$, with notation $f_{r,s} := f_s - f_r$, for a function f from $[0, \infty)$ into a vector space. In (2.3), we denoted by $X(\cdot) \otimes Y(\cdot)$, for any two X and Y in $\mathbb{L}^q(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R}^m)$, the random variable $(\omega, \omega') \mapsto (X_i(\omega) Y_j(\omega'))_{1 \le i,j \le m}$ defined on Ω^2. It is in $\mathbb{L}^q(\Omega^2, \mathcal{F}^{\otimes 2}, \mathbb{P}^{\otimes 2}; \mathbb{R}^{m \times m})$. The notation $\perp\!\!\!\perp$ in $\mathbb{W}^{\perp\!\!\!\perp}$ is used to indicate that $\mathbb{W}^{\perp\!\!\!\perp}_{s,t}(\cdot, \cdot)$ should be thought of as the random variable

$$(\omega, \omega') \mapsto \int_s^t \big(W_r(\omega) - W_s(\omega)\big) \otimes dW_r(\omega').$$

Since $\Omega^2 \ni (\omega, \omega') \mapsto (W_t(\omega))_{t \ge 0}$ and $\Omega^2 \ni (\omega, \omega') \mapsto (W_t(\omega'))_{t \ge 0}$ are independent under $\mathbb{P}^{\otimes 2}$, we then understand $\mathbb{W}^{\perp\!\!\!\perp}_{s,t}$ as an iterated integral for two independent copies of the noise. We refer to Examples 2.1 and 2.2 in [5]. In the end, we denote by $\boldsymbol{W}(\omega)$ the so-called rough set-up specified by the ω-dependent collection of maps given by (2.1) and (2.2).

2.2. *Regularity of the rough set-up.* Following [5], we use the notion of p-variation to handle the regularity of the various trajectories in hand. Throughout, the exponent p is taken in the interval $[2, 3)$. For a continuous function $\mathbb{G}$ from the simplex $\mathcal{S}_2^T$ into some $\mathbb{R}^\ell$, we set, for any $p' \ge 1$,

$$\|\mathbb{G}\|^{p'}_{[0,T],p'-\mathrm{v}} := \sup_{0=t_0<t_1\cdots<t_n=T} \sum_{i=1}^n |\mathbb{G}_{t_{i-1},t_i}|^{p'},$$

and define for any function g from $[0, T]$ into $\mathbb{R}^\ell$, $\|g\|^p_{[0,T],p-\mathrm{v}} := \|\mathbb{G}\|^p_{[0,T],p-\mathrm{v}}$ as the p-variation seminorm of its associated two index function $\mathbb{G}_{s,t} := g_t - g_s$. Similarly, for a random variable $\mathbb{G}(\cdot)$ on Ω with values in $\mathcal{C}(\mathcal{S}_2^T; \mathbb{R}^\ell)$, and $p' \ge 1$, we define its p'-variation in $\mathbb{L}^q$ as

$$\big\langle \mathbb{G}(\cdot) \big\rangle^{p'}_{q;[0,T],p'-\mathrm{v}} := \sup_{0=t_0<t_1\cdots<t_n=T} \sum_{i=1}^n \big\langle \mathbb{G}_{t_{i-1},t_i}(\cdot) \big\rangle^{p'}_q, \tag{2.4}$$

and define for a random variable $G(\cdot)$ on Ω, with values in $\mathcal{C}([0,T];\mathbb{R}^\ell)$,

$$\langle G(\cdot)\rangle^{p'}_{q;[0,T],p'-\mathrm{v}} := \langle \mathbb{G}(\cdot)\rangle^{p'}_{q;[0,T],p'-\mathrm{v}},$$

as the p'-variation seminorm in $\mathbb{L}^q$ of its associated two-index function $\mathcal{S}_2^T \ni (s,t) \mapsto \mathbb{G}_{s,t}(\cdot) = G_t(\cdot) - G_s(\cdot)$. Lastly, for a random variable $\mathbb{G}(\cdot,\cdot)$ from $(\Omega^2, \mathcal{F}^{\otimes 2})$ into $\mathcal{C}(\mathcal{S}_2^T;\mathbb{R}^\ell)$, we set

$$\big\langle\!\big\langle \mathbb{G}(\cdot,\cdot)\big\rangle\!\big\rangle^{p'}_{q;[0,T],p'-\mathrm{v}} := \sup_{0=t_0<t_1\cdots<t_n=T} \sum_{i=1}^n \big\langle\!\big\langle \mathbb{G}_{t_{i-1},t_i}(\cdot,\cdot)\big\rangle\!\big\rangle^{p'}_q. \tag{2.5}$$

Given these definitions, we require from the rough set-up $\boldsymbol{W}$ that:

- For any $\omega\in\Omega$, the path $W(\omega)$ is in the space $\mathcal{C}([0,T];\mathbb{R}^m)$, and the map $W:\Omega\ni\omega\mapsto W(\omega)\in\mathcal{C}([0,T];\mathbb{R}^m)$ is Borel-measurable and q-integrable.
- For any $\omega\in\Omega$, the two-index path $\mathbb{W}(\omega)$ is in $\mathcal{C}(\mathcal{S}_2^T;\mathbb{R}^{m\times m})$, and the map $\mathbb{W}:\Omega\ni\omega\mapsto \mathbb{W}(\omega)\in\mathcal{C}(\mathcal{S}_2^T;\mathbb{R}^{m\times m})$ is Borel-measurable and q-integrable.
- For any $(\omega,\omega')\in\Omega^2$, the two-index path $\mathbb{W}^{\perp}(\omega,\omega')$ is an element of $\mathcal{C}(\mathcal{S}_2^T;\mathbb{R}^{m\times m})$, and the map $\mathbb{W}^{\perp}:\Omega^2\ni(\omega,\omega')\mapsto \mathbb{W}^{\perp}(\omega,\omega')\in\mathcal{C}(\mathcal{S}_2^T;\mathbb{R}^{m\times m})$ is Borel-measurable and q-integrable.

Moreover, we may set, for some fixed $p\in[2,3)$ and for all $0\le s\le t\le T$ and $\omega\in\Omega$,

$$\begin{aligned} v(s,t,\omega) := {} & \big\|W(\omega)\big\|^p_{[s,t],p-\mathrm{v}} + \langle W(\cdot)\rangle^p_{q;[s,t],p-\mathrm{v}} \\ & + \big\|\mathbb{W}(\omega)\big\|^{p/2}_{[s,t],p/2-\mathrm{v}} + \langle \mathbb{W}^{\perp}(\omega,\cdot)\rangle^{p/2}_{q;[s,t],p/2-\mathrm{v}} \\ & + \langle \mathbb{W}^{\perp}(\cdot,\omega)\rangle^{p/2}_{q;[s,t],p/2-\mathrm{v}} + \big\langle\!\big\langle \mathbb{W}^{\perp}(\cdot,\cdot)\big\rangle\!\big\rangle^{p/2}_{q;[s,t],p/2-\mathrm{v}}, \end{aligned} \tag{2.6}$$

and we assume that, for any positive finite time T and any $\omega\in\Omega$, the quantity $v(0,T,\omega)$ is finite. Importantly, $\omega\mapsto (v(s,t,\omega))_{(s,t)\in\mathcal{S}_2^T}$ is a random variable with values in $\mathcal{C}(\mathcal{S}_2^T;\mathbb{R}_+)$ and is super-additive, namely, for any $0\le r\le s\le t\le T$, and $\omega\in\Omega$,

$$v(r,t,\omega)\ge v(r,s,\omega)+v(s,t,\omega).$$

We then assume $\langle v(0,T,\cdot)\rangle_q<\infty$, which implies, by Lebesgue's dominated convergence theorem, that the function $\mathcal{S}_2^T\ni(s,t)\mapsto\langle v(s,t,\cdot)\rangle_q$ is continuous. We assume that it is of bounded variation on $[0,T]$, that is,

$$\big\langle v(\cdot)\big\rangle_{q;[s,t],1-\mathrm{v}} := \sup_{0\le t_1<\cdots<t_K\le T} \sum_{i=1}^K \big\langle v(t_{i-1},t_i,\cdot)\big\rangle_q < \infty. \tag{2.7}$$

We then call a control any family of random variables $(\omega\mapsto w(s,t,\omega))_{(s,t)\in\mathcal{S}_2^T}$ that is jointly continuous in (s,t) and that satisfies,

$$w(s,t,\omega)\ge v(s,t,\omega)+\big\langle v(\cdot)\big\rangle_{q;[s,t],1-\mathrm{v}}, \tag{2.8}$$

together with

$$\begin{aligned} & \big\langle w(s,t,\cdot)\big\rangle_q \le 2w(s,t,\omega), \\ & w(r,t,\omega)\ge w(r,s,\omega)+w(s,t,\omega), \quad r\le s\le t. \end{aligned} \tag{2.9}$$

A typical choice to get (2.8) and (2.9) is to choose

$$w(s,t,\omega) := v(s,t,\omega)+\big\langle v(\cdot)\big\rangle_{q;[s,t],1-\mathrm{v}}. \tag{2.10}$$

2.3. *Controlled trajectories*. With a rough set-up at hands on a given finite time interval $[0,T]$, we define an associated notion of controlled path and rough integral in the spirit of Gubinelli [32]. Again, we refer to [5] for details, see Definition 3.1 therein.

DEFINITION 2.1. An ω-dependent continuous $\mathbb{R}^d$-valued path $(X_t(\omega))_{0\le t\le T}$ is called an ω-controlled path on $[0,T]$ if its increments can be decomposed as

$$X_{s,t}(\omega)=\delta_x X_s(\omega)W_{s,t}(\omega)+\mathbb{E}\big[\delta_\mu X_s(\omega,\cdot)W_{s,t}(\cdot)\big]+R^X_{s,t}(\omega), \tag{2.11}$$

where $(\delta_x X_t(\omega))_{0\le t\le T}$ belongs to $\mathcal{C}([0,T];\mathbb{R}^{d\times m})$ and $(\delta_\mu X_t(\omega,\cdot))_{0\le t\le T}$ to $\mathcal{C}([0,T];\mathbb{L}^{4/3}(\Omega,\mathcal{F},\mathbb{P};\mathbb{R}^{d\times m}))$, $(R^X_{s,t}(\omega))_{s,t\in\mathcal{S}_2^T}$ is in the space $\mathcal{C}(\mathcal{S}_2^T;\mathbb{R}^d)$, and

$$\begin{aligned}\vertiii{X(\omega)}_{\star,[0,T],w,p}&:=|X_0(\omega)|+|\delta_x X_0(\omega)|+\langle\delta_\mu X_0(\omega,\cdot)\rangle_{4/3}\\&\quad+\vertiii{X(\omega)}_{[0,T],w,p}<\infty,\end{aligned}$$

where

$$\begin{aligned}\vertiii{X(\omega)}_{[0,T],w,p}&:=\|X(\omega)\|_{[0,T],w,p}+\|\delta_x X(\omega)\|_{[0,T],w,p}\\&\quad+\langle\delta_\mu X(\omega,\cdot)\rangle_{[0,T],w,p,4/3}+\|R^X(\omega)\|_{[0,T],w,p/2},\end{aligned}$$

with

$$\|X(\omega)\|_{[0,T],w,p}:=\sup_{\varnothing\neq(s,t)\subset[0,T]}\frac{|X_{s,t}(\omega)|}{w(s,t,\omega)^{1/p}},\quad\text{and similarly for }\delta_x X$$

$$\langle\delta_\mu X(\omega,\cdot)\rangle_{[0,T],w,p,4/3}:=\sup_{\varnothing\neq(s,t)\subset[0,T]}\frac{\langle\delta_\mu X_{s,t}(\omega,\cdot)\rangle_{4/3}}{w(s,t,\omega)^{1/p}},$$

$$\|R^X(\omega)\|_{[0,T],w,p/2}:=\sup_{\varnothing\neq(s,t)\subset[0,T]}\frac{|R^X_{s,t}(\omega)|}{w(s,t,\omega)^{2/p}}.$$

We call $\delta_x X(\omega)$ and $\delta_\mu X(\omega,\cdot)$ in (2.11) the derivatives of $X(\omega)$.

We then define the notion of random controlled trajectory, which consists of a collection of ω-controlled trajectories indexed by the elements of Ω.

DEFINITION 2.2. A family of ω-controlled paths $(X(\omega))_{\omega\in\Omega}$ such that X, $\delta_x X$, $\delta_\mu X$ and R^X are measurable from Ω into $\mathcal{C}([0,T];\mathbb{R}^d)$, $\mathcal{C}([0,T];\mathbb{R}^{d\times m})$, $\mathcal{C}([0,T];\mathbb{L}^{4/3}(\Omega,\mathcal{F},\mathbb{P};\mathbb{R}^{d\times m}))$ and $\mathcal{C}(\mathcal{S}_2^T;\mathbb{R}^d)$, and satisfy

$$\langle X_0(\cdot)\rangle_2+\langle\vertiii{X(\cdot)}_{[0,T],w,p}\rangle_8<\infty \tag{2.12}$$

is called a random controlled path on $[0,T]$.

It is proven in [5], Lemma 3.3, that a random controlled trajectory induces a continuous path $t\mapsto X_t(\cdot)$ from $[0,T]$ to $\mathbb{L}^2(\Omega,\mathcal{F},\mathbb{P};\mathbb{R}^d)$. On another matter, the reader may observe that we do not require any integrability property on the initial conditions $\delta_x X_0$ and $\delta_\mu X_0$ of the two derivative processes. In fact, it must be understood that, when dealing with solutions to (1.2), both $\delta_x X_0$ and $\delta_\mu X_0$ are automatically prescribed: $\delta_x X_0(\omega)=\mathrm{F}(X_0(\omega),\mathcal{L}(X_0))$ and $\delta_\mu X_0(\omega,\cdot)=0$, see the forthcoming Theorem 2.3 for more details. Simple bounds for both $\delta_x X_0$ and $\delta_\mu X_0$ then easily follow. Lastly, as it is explained in our companion paper [5], see Definition 3.1 therein, the values of $4/3$ and 8 in Definition 2.2 are somewhat arbitrary. In particular, the analysis could be managed with other exponents provided that a certain trade-off between the two of them (here $4/3$ and 8) still holds. To make it clear, $4/3$ should be here regarded as the conjugate exponent of 4; and the reader may easily guess that 8 shows up in the computations when squaring some fourth moments.

2.4. *Rough integral.* As for the construction, of the rough integral, we recall the following statement from [5], Theorem 3.4.

THEOREM 2.3. *There exists a universal constant c_0 and, for any $\omega \in \Omega$, there exists a continuous linear map*

$$\big(X_t(\omega)\big)_{0\le t\le T} \mapsto \left(\int_s^t X_{s,u}(\omega)\otimes d\boldsymbol{W}_u(\omega)\right)_{(s,t)\in\mathcal{S}_2^T}$$

from the space of ω-controlled trajectories equipped with the norm $|||\cdot|||_{\star,[0,T],p}$, onto the space of continuous functions from $\mathcal{S}_2^T$ into $\mathbb{R}^d\otimes\mathbb{R}^m$ with finite norm $\|\cdot\|_{[0,T],w,p/2}$, with w in the latter norm being evaluated along the realization ω, that satisfies for any $0\le r\le s\le t\le T$ the identity

$$\begin{aligned}&\int_r^t X_{r,u}(\omega)\otimes d\boldsymbol{W}_u(\omega)\\&\quad=\int_r^s X_{r,u}(\omega)\otimes d\boldsymbol{W}_u(\omega)+\int_s^t X_{s,u}(\omega)\otimes d\boldsymbol{W}_u(\omega)+X_{r,s}(\omega)\otimes W_{s,t}(\omega),\end{aligned}$$

together with the estimate

$$\begin{aligned}&\left|\int_s^t X_{s,u}(\omega)\otimes d\boldsymbol{W}_u(\omega)-\delta_x X_s(\omega)\mathbb{W}_{s,t}(\omega)-\mathbb{E}\big[\delta_\mu X_s(\omega,\cdot)\mathbb{W}_{s,t}^{\perp}(\cdot,\omega)\big]\right|\\&\qquad\le c_0|||X(\omega)|||_{[0,T],w,p}\,w(s,t,\omega)^{3/p}.\end{aligned}\tag{2.13}$$

Above, $\delta_x X_s(\omega)\mathbb{W}_{s,t}(\omega)$ is the product of a $d\times m$ matrix and an $m\times m$ matrix, so it gives back a $d\times m$ matrix, with components $(\delta_x X_s(\omega)\mathbb{W}_{s,t}(\omega))_{i,j}=\sum_{k=1}^m(\delta_x X_s^i(\omega))_k\times(\mathbb{W}_{s,t}(\omega))_{k,j}$, for $i\in\{1,\dots,d\}$ and $j\in\{1,\dots,m\}$, and similarly for $\mathbb{E}[\delta_\mu X_s(\omega,\cdot)\mathbb{W}_{s,t}^{\perp}(\cdot,\omega)]$. As usual, the above construction allows us to define an additive process setting

$$\int_s^t X_u(\omega)\otimes d\boldsymbol{W}_u(\omega):=\int_s^t X_{s,u}(\omega)\otimes d\boldsymbol{W}_u(\omega)+X_s(\omega)\otimes W_{s,t}(\omega),$$

for $0\le s\le t\le T$. We can thus consider the integral process $(\int_0^t X_s(\omega)\otimes d\boldsymbol{W}_s(\omega))_{0\le t\le T}$ as an ω-controlled trajectory with values in $\mathbb{R}^{d\times m}$, with

$$\left(\delta_x\left[\int_0^\cdot X_s(\omega)\otimes d\boldsymbol{W}_s(\omega)\right]_t\right)_{(i,j),k}=\big(X_t(\omega)\big)_i\delta_{j,k},$$

for $i\in\{1,\dots,d\}$ and $j,k\in\{1,\dots,m\}$, where $\delta_{j,k}$ stands for the usual Kronecker symbol, and with null μ-derivative.

When the trajectory $X(\omega)$ takes in values in $\mathbb{R}^d\otimes\mathbb{R}^m$ rather than $\mathbb{R}^d$, the integral $\int_0^t X_s(\omega)\otimes d\boldsymbol{W}_s(\omega)$ belongs to $\mathbb{R}^d\otimes\mathbb{R}^m\otimes\mathbb{R}^m$. We then set for $i\in\{1,\dots,d\}$

$$\left(\int_0^t X_s(\omega)\,d\boldsymbol{W}_s(\omega)\right)_i:=\sum_{j=1}^m\left(\int_0^t X_s(\omega)\otimes d\boldsymbol{W}_s(\omega)\right)_{i,j,j},$$

and consider $\int_0^t X_s(\omega)\,d\boldsymbol{W}_s(\omega)$ as an element of $\mathbb{R}^d$.

2.5. *Stability of controlled paths under nonlinear maps.* A key fact in [5] is to use regularity properties of functions defined on Wasserstein space through a lifting procedure to an L^2 space standing above the probability space. We refer the reader to Lions' lectures [38], to the lecture notes [12] of Cardaliaguet or to Carmona and Delarue's monograph [14], Chapter 5, for basics on the subject. Among others, the latter addresses the connection with other

forms of derivatives on the Wasserstein space, such as flat or intrinsic derivatives. Flat derivatives are inherited from the linear structure of the space of (hence nonnormalized) signed measures and are also called, depending on the definition, extrinsic or convexity extrinsic derivatives, see for instance [44, 45]. Intrinsic derivatives were described by [3, 39, 50] after the pioneering works [34, 43] on gradient flows on the space of probability measures; a fine comparison with Lions' approach is carried out in the recent work [30]. The typical example for a lift in L^2 is to consider an ordinary differential equation driven by some vector field b starting from some random variable; it provides a lift of the solution to a Fokker–Planck equation together with a systematic way to identify derivatives of functionals of a measure argument in the direction b. To the best of our knowledge, this latter idea goes back to the earlier works [1, 2].

- Recall $(\Omega, \mathcal{F}, \mathbb{P})$ stands for an atomless probability space, with Ω a Polish space and $\mathcal{F}$ its Borel σ-algebra. Fix a finite dimensional space $E = \mathbb{R}^k$ and denote, for $r \geq 1$, by $\mathbb{L}^r := \mathbb{L}^r(\Omega, \mathcal{F}, \mathbb{P}; E)$ the space of E-valued random variables on Ω with finite r moment. We equip the space $\mathcal{P}_r(E) := \{\mathcal{L}(Z); Z \in \mathbb{L}^r\}$ with the r-Wasserstein distance

$$\mathbf{d}_r(\mu_1, \mu_2) := \inf\big\{\|Z_1 - Z_2\|_r; \mathcal{L}(Z_1) = \mu_1, \mathcal{L}(Z_2) = \mu_2\big\}. \tag{2.14}$$

  When $r = 2$, an $\mathbb{R}^k$-valued function u defined on $\mathcal{P}_2(E)$ is canonically extended to $\mathbb{L}^2$ by setting, for any $Z \in \mathbb{L}^2$, $U(Z) := u(\mathcal{L}(Z))$.
- The function u is then said to be differentiable at $\mu \in \mathcal{P}_2(E)$ if its canonical lift is Fréchet differentiable at some point Z such that $\mathcal{L}(Z) = \mu$; we denote by $\nabla_Z U \in (\mathbb{L}^2)^k$ the gradient of U at Z. The function U is then differentiable at any other point $Z' \in \mathbb{L}^2$ such that $\mathcal{L}(Z') = \mu$, and the laws of $\nabla_Z U$ and $\nabla_{Z'} U$ are equal, for any such Z'.
- The function u is said to be of class C^1 if its canonical lift is of class C^1. If u is of class C^1 on $\mathcal{P}_2(E)$, then $\nabla_Z U$ is $\sigma(Z)$-measurable and given by an $\mathcal{L}(Z)$-dependent function Du from E to E^k such that

$$\nabla_Z U = (Du)(Z). \tag{2.15}$$

  In order to emphasize the fact that Du depends upon $\mathcal{L}(Z)$, we shall write $D_\mu u(\mathcal{L}(Z))(\cdot)$ instead of $Du(\cdot)$. Importantly, this representation is independent of the choice of the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and can be easily transported from one probability space to another.

Throughout the paper, we regard the function F in (1.2) as a map from $\mathbb{R}^d \times \mathbb{L}^2(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R}^d)$ into the space $L(\mathbb{R}^m, \mathbb{R}^d) \cong \mathbb{R}^d \otimes \mathbb{R}^m$ of linear mappings from $\mathbb{R}^m$ to $\mathbb{R}^d$. Intuitively, we identify the coefficient driving equation (1.2) with its lift $\widehat{\mathrm{F}}$. Following [5], Section 3.3, we require F to satisfy the following regularity assumptions.

REGULARITY ASSUMPTIONS 1. Assume that F is continuously differentiable in the joint variable (x, Z), that $\partial_x F$ is also continuously differentiable in (x, Z) and that there is some positive finite constant Λ such that $|\mathrm{F}(x, \mu)|$, $|\partial_x \mathrm{F}(x, \mu)|$, $|\partial_x^2 \mathrm{F}(x, \mu)|$, $\|\nabla_Z \mathrm{F}(x, Z)\|_2$ and $\|\partial_x \nabla_Z \mathrm{F}(x, Z)\|_2$ are bounded by Λ, for any $x \in \mathbb{R}^d$, $\mu \in \mathcal{P}_2(\mathbb{R}^d)$ and $Z \in \mathbb{L}^2(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R}^d)$. Assume moreover that, for any $x \in \mathbb{R}^d$, the mapping $Z \mapsto \nabla_Z \mathrm{F}(x, Z)$ is a Λ-Lipschitz function of $Z \in \mathbb{L}^2(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R}^d)$.

We recall below that, for an ω-controlled path $X(\omega)$ and for an $\mathbb{R}^d$-valued random controlled path $Y(\cdot)$, $\mathrm{F}(X(\omega), Y(\cdot)) := (\mathrm{F}(X_t(\omega), Y_t(\cdot)))_{0 \leq t \leq T}$ may be also expanded in the form of an ω-controlled trajectory. As explained in [5], (3.8), it suffices for our purpose to provide the form of the expansion when $\delta_\mu X(\omega) \equiv 0$ and $\delta_\mu Y(\cdot) \equiv 0$.

PROPOSITION 2.4. *Let $X(\omega)$ be an ω-controlled path and $Y(\cdot)$ be an $\mathbb{R}^d$-valued random controlled path. Assume that $\delta_\mu X(\omega) \equiv 0$ and $\delta_\mu Y(\cdot) \equiv 0$ and that $\sup_{0\le t\le T}(|\delta_x X_t(\omega)| \vee \langle \delta_x Y_t(\cdot)\rangle_\infty) < \infty$. Then, $F(X(\omega), Y(\cdot))$ is an ω-controlled path with*

$$\delta_x\big(F\big(X(\omega), Y(\cdot)\big)\big)_t = \partial_x F\big(X_t(\omega), Y_t(\cdot)\big)\delta_x X_t(\omega),$$

which is understood as $(\sum_{\ell=1}^d \partial_{x_\ell} F^{i,j}(X_t(\omega), Y_t(\cdot))(\delta_x X_t^\ell(\omega))_k)_{i,j,k}$, with $i,k \in \{1,\dots,d\}$ and $j \in \{1,\dots,m\}$, and (with a similar interpretation for the product)

$$\delta_\mu\big(F\big(X(\omega), Y(\cdot)\big)\big)_t = D_\mu \mathrm{F}\big(X_t(\omega), \mathcal{L}(Y_t)\big)\big(X_t(\cdot)\big)\delta_x Y_t(\cdot).$$

2.6. *Local accumulation.* In order to proceed with the analysis of (1.3), we make use of the notion of local accumulation. Following [5], we define it as follows. Given a nondecreasing[2] continuous positive valued function ϖ on $\mathcal{S}_2$, a nonnegative parameter s and a positive threshold α, we define inductively a sequence of times setting $\tau_0(s,\alpha) := s$, and

$$\tau_{n+1}^\varpi(s,\alpha) := \inf\big\{u \ge \tau_n^\varpi(s,\alpha) : \varpi\big(\tau_n^\varpi(s,\alpha), u\big) \ge \alpha\big\}, \tag{2.16}$$

with the understanding that $\inf\varnothing := +\infty$. For $t \ge s$, set

$$N_\varpi\big([s,t],\alpha\big) := \sup\big\{n \in \mathbb{N} : \tau_n^\varpi(s,\alpha) \le t\big\}. \tag{2.17}$$

We call N_ϖ the local accumulation of ϖ (of size α if we specify the value of the threshold): $N_\varpi([s,t],\alpha)$ is the largest number of disjoint open sub-intervals (a,b) of $[s,t]$ on which $\varpi(a,b)$ is greater than or equal to α. When $\varpi(s,t) = w(s,t,\omega)^{1/p}$ with w a control satisfying (2.8) and (2.9) and when the framework makes it clear, we just write $N([s,t],\omega,\alpha)$ for $N_\varpi([s,t],\alpha)$. Similarly, we also write $\tau_n(s,\omega,\alpha)$ for $\tau_n^\varpi(s,\alpha)$ when $\varpi(s,t) = w(s,t,\omega)$. We will also use the convenient notation

$$\tau_n^\varpi(s,t,\alpha) := \tau_n^\varpi(s,\alpha) \wedge t.$$

3. Analysis of the mean field rough differential equation.

3.1. *Solving the equation.* The following notion of solution to (1.2) is taken from [5], Definition 4.1.

DEFINITION 3.1. Let W together with its enhancement $\boldsymbol{W}$ satisfy the assumption of Section 2.2 on a finite interval $[0,T]$. A solution to (1.2) on the time interval $[0,T]$, with initial condition $X_0(\cdot) \in \mathbb{L}^2(\Omega,\mathcal{F},\mathbb{P};\mathbb{R}^d)$, is a random controlled path $X(\cdot)$ such that for $\mathbb{P}$-a.e. ω the paths $X(\omega)$ and $X_0(\omega) + \int_0^\cdot \mathrm{F}(X_s(\omega), X_s(\cdot))\, d\boldsymbol{W}_s(\omega)$ coincide.

We formulate here the regularity assumptions on $\mathrm{F}(x,\mu)$ used in [5], in addition to Regularity assumptions 1, to show the well-posed character of Equation (1.2). Below, we denote by $(\widetilde{\Omega},\widetilde{\mathcal{F}},\widetilde{\mathbb{P}})$ a copy of $(\Omega,\mathcal{F},\mathbb{P})$, and given a random variable Z on $(\Omega,\mathcal{F},\mathbb{P})$, write $\widetilde{Z}$ for its copy on $(\widetilde{\Omega},\widetilde{\mathcal{F}},\widetilde{\mathbb{P}})$.

REGULARITY ASSUMPTIONS 2.

- The function $\partial_x \mathrm{F}$ is differentiable in (x,μ).
- For each $(x,\mu) \in \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d)$, there exists a version of $D_\mu \mathrm{F}(x,\mu)(\cdot) \in L^2_\mu(\mathbb{R}^d;\mathbb{R}^d\otimes\mathbb{R}^m)$ such that the map $(x,\mu,z) \mapsto D_\mu \mathrm{F}(x,\mu)(z)$ from $\mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d) \times \mathbb{R}^d$ to $\mathbb{R}^d \otimes \mathbb{R}^m \otimes \mathbb{R}^d$ is of class C^1, the derivative in the direction μ being understood as before.

[2] In the sense that $\varpi(a,b) \ge \varpi(a',b')$ if $(a',b') \subset (a,b)$.

- The function $(x, Z) \mapsto \partial_x^2 \mathrm{F}(x, \mathcal{L}(Z))$ from $\mathbb{R}^d \times \mathbb{L}^2(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R}^d)$ to $\mathbb{R}^d \otimes \mathbb{R}^m \otimes \mathbb{R}^d \otimes \mathbb{R}^d$ is bounded by Λ and Λ-Lipschitz continuous.
- The two derivative functions $(x, Z) \mapsto \partial_x D_\mu \mathrm{F}(x, \mathcal{L}(Z))(Z(\cdot))$ (which is the same as $(x, Z) \mapsto D_\mu \partial_x \mathrm{F}(x, \mathcal{L}(Z))(Z(\cdot))$ by Schwarz' theorem) and $(x, Z) \mapsto \partial_z D_\mu \mathrm{F}(x, \mathcal{L}(Z))(Z(\cdot))$ are bounded by Λ and Λ-Lipschitz continuous from $\mathbb{R}^d \times \mathbb{L}^2(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R}^d)$ to $\mathbb{L}^2(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R}^d \otimes \mathbb{R}^m \otimes \mathbb{R}^d \otimes \mathbb{R}^d)$.
- For each $\mu \in \mathcal{P}_2(\mathbb{R}^d)$, we denote by $D_\mu^2 \mathrm{F}(x, \mu)(z, \cdot)$, the derivative of $D_\mu \mathrm{F}(x, \mu)(z)$ with respect to μ—which is indeed given by a function. For $z' \in \mathbb{R}^d$, $D_\mu^2 \mathrm{F}(x, \mu)(z, z')$ is an element of $\mathbb{R}^d \otimes \mathbb{R}^m \otimes \mathbb{R}^d \otimes \mathbb{R}^d$. We assume that
$$(x, Z) \mapsto D_\mu^2 \mathrm{F}\big(x, \mathcal{L}(Z)\big)\big(Z(\cdot), \widetilde{Z}(\cdot)\big),$$
from $\mathbb{R}^d \times \mathbb{L}^2(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R}^d)$ to $\mathbb{L}^2(\Omega \times \widetilde{\Omega}, \mathcal{F} \otimes \widetilde{\mathcal{F}}, \mathbb{P} \otimes \widetilde{\mathbb{P}}; \mathbb{R}^d \otimes \mathbb{R}^m \otimes \mathbb{R}^d \otimes \mathbb{R}^d)$, is bounded by Λ and Λ-Lipschitz continuous.

The two functions $\mathrm{F}(x, \mu) = \int f(x, y)\mu(dy)$ and $\mathrm{F}(x, \mu) = g(x, \int y\mu(dy))$, for functions $f, g \in C_b^3$ (meaning that f and g are bounded and have bounded derivatives of order 1, 2 and 3), satisfy the Regularity assumptions 1 and 2. The following property is taken from [5], Propositions 4.3 and (4.21).

PROPOSITION 3.2. *Let* F *satisfy Regularity assumptions 1 and 2 and* w *be a control satisfying (2.8) and (2.9). Consider two* ω*-controlled paths* $X(\omega)$ *and* $X'(\omega)$ *with possibly different initial conditions* $(X_0(\omega), \delta_x X_0(\omega))$ *and* $(X_0'(\omega), \delta_x X_0'(\omega))$*, defined on a time interval* $[0, T]$*, together with two random controlled paths* $Y(\cdot)$ *and* $Y'(\cdot)$*, with possibly different initial conditions* $(Y_0(\omega), \delta_x Y_0(\omega))$ *and* $(Y_0'(\omega), \delta_x Y_0'(\omega))$*, all of them satisfying* $\delta_\mu X(\omega) \equiv \delta_\mu X'(\omega) \equiv 0$ *and* $\delta_\mu Y(\cdot) \equiv \delta_\mu Y'(\cdot) \equiv 0$ *together with*

$$\big|\delta_x X(\omega)\big| \vee \big|\delta_x X'(\omega)\big| \vee \big\langle \delta_x Y(\cdot)\big\rangle_\infty \vee \big\langle \delta_x Y'(\cdot)\big\rangle_\infty \le \Lambda, \tag{3.1}$$

and the size estimates

$$\big\langle \|\!| Y(\cdot) |\!\|_{[0,T],w,p}\big\rangle_8^2 \le L_0, \qquad \big\langle \|\!| Y'(\cdot) |\!\|_{[0,T],w,p}\big\rangle_8^2 \le L_0, \tag{3.2}$$

$$\|\!| X(\omega) |\!\|_{[t_i^0, t_{i+1}^0],w,p}^2 \le L_0, \qquad \|\!| X'(\omega) |\!\|_{[t_i^0, t_{i+1}^0],w,p}^2 \le L_0, \tag{3.3}$$

for $i \in \{0, \dots, N^0\}$*, for some* $L_0 \ge 1$*, and* $N^0 = N([0, T], \omega, 1/(4L_0))$ *given by (2.17), and for the sequence* $(t_i^0 = \tau_i(0, T, \omega, 1/(4L_0)))_{i=0,\dots,N^0+1}$ *given by (2.16).*

Then, we can find a constant γ *depending on* L_0 *and* Λ *such that, for any partition* $(t_i)_{i=0,\dots,N}$ *included in* $(t_i^0)_{i=0,\dots,N^0}$ *and satisfying* $w(t_i, t_{i+1}, \omega)^{1/p} \le 1/(4L)$ *for some* $L \ge L_0$*, we have*

$$\begin{aligned}
&\Big\|\!\Big| \int_{t_i}^{\cdot} \big(F\big(X_r(\omega), Y_r(\cdot)\big) - F\big(X_r'(\omega), Y_r'(\cdot)\big)\big)\, d\boldsymbol{W}_r(\omega) \Big|\!\Big\|_{[t_i, t_{i+1}],w,p} \\
&\quad \le \gamma\big(|\Delta X_0(\omega)| + |\delta_x \Delta X_0(\omega)|\big) + \big\langle \Delta Y_0(\cdot)\big\rangle_4 + \big\langle \delta_x \Delta Y_0(\cdot)\big\rangle_4 \\
&\qquad + \gamma\, w(0, t_i, \omega)^{1/p}\big(\|\!| \Delta X(\omega) |\!\|_{[0,t_i],w,p} + \big\langle \|\!| \Delta Y(\cdot) |\!\|_{[0,T],w,p}\big\rangle_8\big) \\
&\qquad + \frac{\gamma}{4L}\big(\|\!| \Delta X(\omega) |\!\|_{[t_i, t_{i+1}],w,p} + \big\langle \|\!| \Delta Y(\cdot) |\!\|_{[0,T],w,p}\big\rangle_8\big),
\end{aligned} \tag{3.4}$$

where $\Delta X_t(\omega) := X_t(\omega) - X_t'(\omega)$, $\Delta Y_t(\cdot) := Y_t(\cdot) - Y_t'(\cdot)$, $t \in [0, T]$.

In [5], Proposition 3.2 is used to prove the following existence and uniqueness result, see Theorems 1.1 and 4.4 therein, to which we add the final estimate in the statement.

THEOREM 3.3. *Let* F *satisfy Regularity Assumptions* 1 *and* 2 *and let* w *be a control satisfying* (2.8) *and* (2.9). *Assume there exists a positive time horizon* T *such that the random variables* $w(0,T,\cdot)$ *and* $(N([0,T],\cdot,\alpha))_{\alpha>0}$ *have sub and super exponential tails respectively, in the sense that*

$$\begin{aligned}&\mathbb{P}\big(w(0,T,\cdot)\geq t\big)\leq c_1\exp\big(-t^{\varepsilon_1}\big),\\&\mathbb{P}\big(N\big([0,T],\cdot,\alpha\big)\geq t\big)\leq c_2(\alpha)\exp\big(-t^{1+\varepsilon_2(\alpha)}\big),\end{aligned}\tag{3.5}$$

for some positive constants c_1 *and* ε_1, *and possibly* α*-dependent positive constants* $c_2(\alpha)$ *and* $\varepsilon_2(\alpha)$. *Then, for any* d*-dimensional square-integrable random variable* X_0, *the mean field rough differential equation* (1.2) *has a unique solution defined on the whole interval* $[0,T]$. *Moreover, there exist four positive real numbers* γ_0, L_0, L *and* η_0 (*with* $\gamma_0,\eta_0>1$), *only depending on* Λ *and* T, *such that, for any subinterval* $[S_1,S_2]\subset[0,T]$ *for which*

$$\big\langle N\big([S_1,S_2],\cdot,1/(4L_0)\big)\big\rangle_8\leq 1,$$

and

$$\big\langle\big[\gamma\big(1+w(0,T,\cdot)^{1/p}\big)\big]^{N([S_1,S_2],\cdot,1/(4L))}\big\rangle_{32}\leq\eta_0,$$

it holds, for any $\omega\in\Omega$

$$\big|\!\big|\!\big|X(\omega)\big|\!\big|\!\big|_{[S_1,S_2],w,p}\leq\big[C\big(1+w(0,T,\omega)^{1/p}\big)\big]^{2N([0,T],\omega,1/(4L))},$$

for a constant C *depending only on* Λ *and* T.

PROOF. We just address the derivation of the last inequality since the latter is not given in [5], Theorem 4.4. The key point is to sum over $n\geq 1$ in [5], (4.30), replacing $[0,S]$ therein by $[S_1,S_2]$, which is indeed licit provided that $\langle N([S_1,S_2],\cdot,1/(4L_0))\rangle_8\leq 1$, see, for instance, [5], (4.23), and $\langle[\gamma_0(1+w(0,T,\cdot)^{1/p})]^{N([S_1,S_2],\cdot,1/(4L))}\rangle_{32}\leq\eta_0$ for η_0 small enough, see [5], (4.29). □

3.2. *Strong rough set-ups and continuity of the Itô-lyons solution map*. Uniqueness in law of the solutions to (1.2) is proven in [5], Theorem 5.3, under the additional assumption that the set-up satisfies the following definition.

DEFINITION 3.4. A rough set-up is called strong if there exists a measurable mapping $\mathcal{I}$ from $\mathcal{C}([0,T];\mathbb{R}^m)^2$ into $\mathcal{C}(\mathcal{S}_2^T;\mathbb{R}^m\otimes\mathbb{R}^m)$ such that

$$\mathbb{P}^{\otimes 2}\big(\big\{(\omega,\omega')\in\Omega^2:\mathbb{W}^{\perp}(\omega,\omega')=\mathcal{I}\big(W(\omega),W(\omega')\big)\big\}\big)=1.\tag{3.6}$$

For our prospect, the following *continuity theorem* is of crucial interest; see[3] [5], Theorem 5.4.

THEOREM 3.5. *Let* F *satisfy the same assumptions as in* Theorem 3.3. *Given a time interval* $[0,T]$ *and a sequence of probability spaces* $(\Omega_n,\mathcal{F}_n,\mathbb{P}_n)$, *indexed by* $n\in\mathbb{N}$, *let, for any* n, $X_0^n(\cdot):=(X_0^n(\omega_n))_{\omega_n\in\Omega_n}$ *be an* $\mathbb{R}^d$*-valued square-integrable initial condition and*

$$\boldsymbol{W}^n(\cdot):=\big(W^n(\omega_n),\mathbb{W}^n(\omega_n),\mathbb{W}^{n,\perp}\big(\omega_n,\omega_n'\big)\big)_{\omega_n,\omega_n'\in\Omega_n}$$

[3] Here, we feel useful to say a word about the proof of Theorem [5], Theorem 5.4. The proof of Step 2b therein is a bit short. The reader may indeed wonder why K therein may be chosen independently of n. In fact, it suffices to observe, with the same notations as therein, that we can render $\mathbb{P}(N^n([S_j,S_{j+1}],\cdot,1/(4L_0))\geq 1)$ as small as needed, uniformly in n. This follows from the fact that $\mathbb{P}(N^n([S_j,S_{j+1}],\cdot,1/(4L_0))\geq 1)=\mathbb{P}(w^n(S_j,S_{j+1},\cdot)\geq 1/(4L_0))\leq 4L_0\langle w^n(S_j,S_{j+1},\cdot)\rangle$. By the second item in the assumption of Theorem 3.5, the last term is less than $C(S_j-S_{j+1})$, for C independent of n.

be an m-dimensional rough set-up with corresponding control w^n, as given by (2.10), *and local accumulated variation N^n, for fixed values of $p \in [2,3)$ and $q > 8$. Assume that*:

- *the collection $(\mathbb{P}_n \circ (|X_0^n(\cdot)|^2)^{-1})_{n\geq 0}$ is uniformly integrable*;
- *for positive constants ε_1, c_1 and $(\varepsilon_2(\alpha), c_2(\alpha))_{\alpha>0}$, the tail assumption* (3.5) *hold for w^n and N^n, for all $n \geq 0$*;
- *associating v^n with each $\boldsymbol{W}^n(\cdot)$ as in* (2.6), *the functions $(\mathcal{S}_2^T \ni (s,t) \mapsto \langle v^n(s,t,\cdot)\rangle_{2q})_{n\geq 0}$ are uniformly Lipschitz continuous.*

Assume also that:

- *there exist, on another probability space $(\Omega, \mathcal{F}, \mathbb{P})$, a square integrable initial condition $X_0(\cdot)$ with values in $\mathbb{R}^d$ and a strong rough set-up*

$$\boldsymbol{W}(\cdot) := \big(W(\omega), \mathbb{W}(\omega), \mathbb{W}^{\perp}(\omega,\omega')\big)_{\omega,\omega'\in\Omega}$$

with values in $\mathbb{R}^m$, such that the law under the probability measure $\mathbb{P}_n^{\otimes 2}$ of the random variable

$$\Omega_n^2 \ni \big(\omega_n, \omega_n'\big) \mapsto \big(X_0^n(\omega_n), W^n(\omega_n), \mathbb{W}_n(\omega_n), \mathbb{W}_n^{\perp}(\omega_n,\omega_n')\big),$$

seen as a random variable with values in the space $\mathbb{R}^d \times \mathcal{C}([0,T];\mathbb{R}^m) \times \{\mathcal{C}(\mathcal{S}_2^T;\mathbb{R}^m \otimes \mathbb{R}^m)\}^2$, converges in the weak sense to the law of

$$\Omega^2 \ni \big(\omega,\omega'\big) \mapsto \big(X_0(\omega), W(\omega), \mathbb{W}(\omega_n), \mathbb{W}^{\perp}(\omega,\omega')\big).$$

Then, $\boldsymbol{W}(\cdot)$ satisfies the requirements of Theorem 3.3 *for some $p' \in (p,3)$ and $q' \in [8,q)$, with control w therein given by* (2.10). *Moreover, if $X^n(\cdot)$, resp. $X(\cdot)$, is the solution of the mean field rough differential equation driven by $\boldsymbol{W}^n(\cdot)$, resp. $\boldsymbol{W}(\cdot)$, then $X^n(\cdot)$ converges in law to $X(\cdot)$ on $\mathcal{C}([0,T];\mathbb{R}^d)$.*

4. Particle system and propagation of chaos. We now have all the ingredients to write down the limiting mean field rough differential equation (1.2) as the limit of a system of particles driven by rough signals (1.3).

4.1. *Empirical rough set-up.* We recall the framework used to address (1.3). The initial conditions $(X_0^i(\cdot))_{1\leq i\leq n}$ are $\mathbb{R}^d$-valued variables with the same distribution as X_0 (in the statement of Theorem 3.3) and the enhanced signals $(W^i(\cdot), \mathbb{W}^i(\cdot))_{1\leq i\leq n}$ are $\mathbb{R}^m \oplus \mathbb{R}^m \otimes \mathbb{R}^m$-valued with the same distribution as $(W(\cdot), \mathbb{W}(\cdot))$ on the space of continuous functions. Moreover, the variables $(X_0^i(\cdot), W^i(\cdot), \mathbb{W}^i(\cdot))_{1\leq i\leq n}$ are independent and identically distributed. All of them are constructed on a single probability space, still denoted by $(\Omega, \mathcal{F}, \mathbb{P})$. Assuming the rough set-up in Theorem 3.3 to be strong, see Definition 3.4, we let

$$\mathbb{W}^{i,j}(\omega) = \mathcal{I}\big(W^i(\omega), W^j(\omega)\big), \quad i \neq j, 1 \leq i, j \leq n.$$

Obviously, equation (1.3) must be understood as a rough differential equation driven by an $(n \times m)$-dimensional signal $(W^1(\omega), \ldots, W^n(\omega))$, and with $(X^1(\omega), \ldots, X^n(\omega))$ as $(n \times d)$-dimensional output. Our first task is to prove that (1.3) may be also understood as a mean field rough differential equation on a suitable rough set-up and that the two interpretations coincide. If we require $\mathbb{P}^{\otimes 2}(\{(\omega,\omega') : \|\mathbb{W}^{\perp}(\omega,\omega')\|_{[0,T],p/2-\mathrm{v}} < \infty\}) = 1$ in Definition 3.4, then it is pretty clear that, for almost every $\omega \in \Omega$,

$$\boldsymbol{W}^{(n)}(\omega) = \big(\big(W^i(\omega)\big)_{1\leq i\leq n}, \big(\mathbb{W}^{i,j}(\omega)\big)_{1\leq i,j\leq n}\big) =: \big(W^{(n)}(\omega), \mathbb{W}^{(n)}(\omega)\big),$$

is a rough path of finite p-variation, with the convention that $\mathbb{W}^{i,i}(\omega) = \mathbb{W}^i(\omega)$, for $i \in \{1,\ldots,n\}$. As explained in [5], Proposition 2.3, we may change the definition of

$((W^i(\omega))_{1\le i\le n}, (\mathbb{W}^{i,j}(\omega))_{1\le i,j\le n})$ on a $\mathbb{P}$-null set so that $\boldsymbol{W}^{(n)}(\omega)$ is in fact a rough path for any $\omega\in\Omega$.

As mentioned in the Introduction, the striking fact of the analysis was first introduced by Tanaka in [49] and used by Cass and Lyons in their seminal work [18]. The quantity $\mathbb{W}^{(n)}(\omega)$ may be seen as a rough set-up defined on a finite probability space for any fixed $\omega\in\Omega$; we call it the **empirical rough set-up**. To make it clear, observe that, throughout Section 2, the rough structure is supported by the probability space $(\Omega,\mathcal{F},\mathbb{P})$ itself. Here, ω is fixed, and we see the probability space as

$$\left(\{1,\dots,n\}, \mathcal{P}(\{1,\dots,n\}), \frac{1}{n}\sum_{i=1}^n \delta_i\right), \tag{4.1}$$

where $\mathcal{P}(\{1,\dots,n\})$ denotes the collection of subsets of $\{1,\dots,n\}$. The reader may object that such a probability space is not atomless whilst we explicitly assumed $(\Omega,\mathcal{F},\mathbb{P})$ to be atomless in the Introduction (see also [5], Section 2); actually, the reader must realize that, in [5], the atomless property is just used to guarantee that, for any probability measure μ on a given Polish space S, the probability space $(\Omega,\mathcal{F},\mathbb{P})$ carries an S-valued random variable with μ as distribution. In other words, we could instead say that, in [5], the probability space $(\Omega,\mathcal{F},\mathbb{P})$ has to be rich enough, which is indeed guaranteed under the assumptions used therein. So, it is here not a hindrance that $\{1,\dots,n\}$ is finite: We must restrict ourselves to random variables taking at most n-values, but this is exactly what we need for our purposes since all the relevant probability distributions showing up from the particle system are n-empirical distributions.

Hence, in order to draw a parallel with (2.2), the role played by $\omega\in\Omega$ is here played by $i\in\{1,\dots,n\}$ and the matrix (2.2) must read

$$\begin{pmatrix} \mathbb{W}^{i,i}_{s,t}(\omega) & \mathbb{W}^{i,\bullet}_{s,t}(\omega) \\ \mathbb{W}^{\bullet,i}_{s,t}(\omega) & \mathbb{W}^{\bullet,\bullet}_{s,t}(\omega)\end{pmatrix}_{0\le s\le t\le T}, \tag{4.2}$$

where $\mathbb{W}^{i,\bullet}_{s,t}(\omega)$ is seen as $\{1,\dots,n\}\ni j\mapsto \mathbb{W}^{i,j}_{s,t}(\omega)$, $\mathbb{W}^{\bullet,i}_{s,t}(\omega)$ as $\{1,\dots,n\}\ni j\mapsto \mathbb{W}^{j,i}_{s,t}(\omega)$ and $\mathbb{W}^{\bullet,\bullet}_{s,t}(\omega)$ as $\{1,\dots,n\}\ni(i,j)\mapsto \mathbb{W}^{i,j}_{s,t}(\omega)$.

In the same spirit, the variation function v in (2.6) is (we put a subscript p in the variation function below to emphasize the dependence upon p)

$$\begin{aligned} v_p^{i,n}(s,t,\omega) :=& \|W^i(\omega)\|^p_{[s,t],p-\mathrm{v}} + {}^{(n)}\big(W^\bullet(\omega)\big)^p_{q;[s,t],p-\mathrm{v}} \\ &+ \|\mathbb{W}^i(\omega)\|^{p/2}_{[s,t],p/2-\mathrm{v}} + {}^{(n)}\big(\mathbb{W}^{i,\bullet}(\omega)\big)^{p/2}_{q;[s,t],p/2-\mathrm{v}} \\ &+ {}^{(n)}\big(\mathbb{W}^{\bullet,i}(\omega)\big)^{p/2}_{q;[s,t],p/2-\mathrm{v}} + {}^{(n)}\big(\!\big(\mathbb{W}^{\bullet,\bullet}(\omega)\big)\!\big)^{p/2}_{q;[s,t],p/2-\mathrm{v}}, \end{aligned} \tag{4.3}$$

where we used the notations

$${}^{(n)}\big(X^\bullet\big)_q = \left(\frac{1}{n}\sum_{j=1}^n |X^j|^q\right)^{1/q}, \quad {}^{(n)}\big(\!\big(X^{\bullet,\bullet}\big)\!\big)_q = \left(\frac{1}{n^2}\sum_{j,k=1}^n |X^{j,k}|^q\right)^{1/q},$$

the corresponding p-variation being defined as in (2.4) and (2.5). Obviously, $v_p^{i,n}(0,T,\omega)$ is almost surely finite. Hence, in order to check that $\boldsymbol{W}^{(n)}(\omega)$ defines a rough set-up, it remains to check that it satisfies (2.7). To do so, we strengthen the assumptions on the signal and assume that, for the same parameter q as in Section 2, it holds

$$\begin{aligned} &\mathbb{E}\big[\|W(\cdot)\|^{pq}_{[0,T],(1/p)-\mathrm{H}} + \|\mathbb{W}(\cdot)\|^{pq/2}_{[0,T],(2/p)-\mathrm{H}}\big] \\ &\quad + \mathbb{E}^{\otimes 2}\big[\|\mathbb{W}^{\mathbb{L}}(\cdot,\cdot)\|^{pq/2}_{[0,T],(2/p)-\mathrm{H}}\big] < \infty, \end{aligned} \tag{4.4}$$

where

$$\|W(\omega)\|_{[s,t],(1/p)-\mathrm{H}} = \sup_{\varnothing\neq(s',t')\subset[s,t]} \frac{|W_{t'}(\omega)-W_{s'}(\omega)|}{|t'-s'|^{1/p}},$$
$$\|\mathbb{W}(\omega)\|_{[s,t],(2/p)-\mathrm{H}} = \sup_{\varnothing\neq(s',t')\subset[s,t]} \frac{|\mathbb{W}_{s',t'}(\omega)|}{|t'-s'|^{2/p}},$$

and similarly for $\|\mathbb{W}^{\perp}(\omega,\omega')\|_{[s,t],(2/p)-\mathrm{H}}$, stand for the standard Hölder seminorms of the rough path. Then, back to (4.3), we can find a universal positive constant c such that

$$\begin{aligned} v_p^{i,n}(s,t,\omega) \le c\{&\|W^i(\omega)\|^p_{[s,t],(1/p)-\mathrm{H}} + \|\mathbb{W}^i(\omega)\|^{p/2}_{[s,t],(2/p)-\mathrm{H}} \\ &+ {}^{(n)}\big(\|W^{\bullet}(\omega)\|^p_{[s,t],(1/p)-\mathrm{H}}\big)_q + {}^{(n)}\big(\|\mathbb{W}^{i,\bullet}(\omega)\|^{p/2}_{[s,t],(2/p)-\mathrm{H}}\big)_q \\ &+ {}^{(n)}\big(\|\mathbb{W}^{\bullet,i}(\omega)\|^{p/2}_{[s,t],(2/p)-\mathrm{H}}\big)_q \\ &+ {}^{(n)}\big(\big(\|\mathbb{W}^{\bullet,\bullet}(\omega)\|^{p/2}_{[s,t],(2/p)-\mathrm{H}}\big)\big)_q\}(t-s). \end{aligned} \tag{4.5}$$

Taking the mean over $i\in\{1,\dots,n\}$ and invoking the law of large numbers (see Lemma A.5 in Appendix A.3 for a version of the law of large numbers with second order interactions), we deduce that, for almost every $\omega\in\Omega$,

$$\begin{aligned} &\limsup_{n\ge1}\sup_{0\le s<t\le T}\frac{{}^{(n)}(v_p^{\bullet,n}(s,t,\omega))_q}{t-s} \\ &\quad\le c\big(\langle\|W(\cdot)\|^{pq}_{[0,T],(1/p)-\mathrm{H}}+\|\mathbb{W}(\cdot)\|^{pq/2}_{[0,T],(2/p)-\mathrm{H}}\rangle^{1/q} \\ &\qquad+\langle\!\langle\|\mathbb{W}^{\perp}(\cdot,\cdot)\|^{pq/2}_{[0,T],(2/p)-\mathrm{H}}\rangle\!\rangle^{1/q}\big), \end{aligned} \tag{4.6}$$

for a new value of the constant c. Observe that, in order to derive (4.6), the law of large numbers can be directly applied to each of the first three terms in the right-hand side of (4.5), since each of them lead to empirical means over terms of the form $\mathcal{J}(W^i(\omega))$, for a suitable functional $\mathcal{J}$ (which has nothing to do with the mapping $\mathcal{I}$ used in Definition 3.4). Differently, the last three terms in (4.5) require a modicum of care as they lead to empirical means of the form

$$\frac{1}{n^2}\sum_{j,k=1,j\neq k}^{n}\|\mathcal{I}(W^j(\omega),W^k(\omega))\|^{pq/2}_{[0,T],(2/p)-\mathrm{H}}+\frac{1}{n^2}\sum_{j=1}^{n}\|\mathbb{W}^j(\omega)\|^{pq/2}_{[0,T],(2/p)-\mathrm{H}},$$

with $\mathcal{I}$ as in (3.6). Still, if the summands in the two sums are integrable, the limit is $\langle\|\mathcal{I}(W^1(\cdot),W^2(\cdot))\|^{pq/2}_{[0,T],(2/p)-\mathrm{H}}\rangle$, see once again Lemma A.5 in Appendix A.3. Hence, (4.6). Now, the fact that the right-hand side of (4.6) is finite guarantees that the 1-variation in the mean in (2.7) is uniformly controlled in $n\ge1$, the mean in (2.7) being understood as the mean on the probability space $(\{1,\dots,n\},\mathcal{P}(\{1,\dots,n\}),\frac{1}{n}\sum_{i=1}^n\delta_i)$. Here are two examples under which (4.5) holds true.

EXAMPLE 4.1. Assume that the regularity index q used in (2.6) satisfies the inequality $q>1/(1-p/3)$, and that, for some constant $K_T\ge0$, $\langle v(s,t,\cdot)\rangle_q\le K_T(t-s)$ for $(s,t)\in\mathcal{S}_2^T$. Then, we get the bounds

$$\langle|(W_t-W_s)(\cdot)|^{pq}\rangle\le K_T^q|t-s|^q,$$
$$\langle|\mathbb{W}_{s,t}(\cdot)|^{pq/2}\rangle\le K_T^q|t-s|^q,\qquad\langle\!\langle|\mathbb{W}^{\perp}_{s,t}(\cdot,\cdot)|^{pq/2}\rangle\!\rangle\le K_T^q|t-s|^q.$$

By Kolmogorov's criterion for rough paths, see Theorem 3.1 in [28], we deduce that W has paths that are $1/p' := (1-1/q)/p > 1/3$-Hölder continuous. Similarly, $\mathbb{W}$ and $\mathbb{W}^{\perp}$ have paths that are $2p' = 2(1-1/q)/p > 2/3$-Hölder continuous and (4.4) holds true with p' instead of p. So, the empirical rough set-up satisfies the required conditions provided we replace p by p'.

EXAMPLE 4.2. Assume that $W := (W^1, \dots, W^m)$ is a tuple of independent and centred continuous Gaussian processes, defined on $[0,T]$, for which there exist an exponent $\varrho \in [1, 3/2)$ and a constant K such that, for any subinterval $[s,t] \subset [0,T]$ and any $k = 1, \dots, m$, it holds

$$\sup \sum_{i,j} |\mathbb{E}[(W^k_{t_{i+1}} - W^k_{t_i})(W^k_{s_{j+1}} - W^k_{s_j})]|^{\rho} \le K|t-s|, \tag{4.7}$$

the sup being over divisions $(t_i)_i$ and $(s_j)_j$ of $[s,t]$. Then, $\|W(\cdot)\|_{[0,T],(1/p)-\mathrm{H}}$ has Gaussian tail and $\|\mathbb{W}(\cdot)\|_{[0,T],(2/p)-\mathrm{H}}$ and $\|\mathbb{W}^{\perp}(\cdot,\cdot)\|_{[0,T],(2/p)-\mathrm{H}}$ have exponential tails, for any $p \in (2\varrho, 3)$; see Theorem 11.9 in [28].

Now that we have defined the empirical rough set-up, we must make clear the meaning given to the rough differential equation (1.2) in Definition 3.1 when the rough set-up therein is precisely the empirical rough set-up. We call the corresponding rough differential equation the *empirical rough differential equation*.

For a given $\omega \in \Omega$, the probability space that carries the empirical rough-set up is given by (4.1). Despite the fact it is not atomless, whilst $(\Omega, \mathcal{F}, \mathbb{P})$ is, Theorem 3.3 applies and guarantees existence and uniqueness of a solution to the empirical rough differential equation. In this regard, observe that the square integrability requirement on the initial condition here writes $\frac{1}{n}\sum_{i=1}^n |X_0^i(\omega)|^2 < \infty$, which is indeed satisfied for ω in a full event. The solution reads in the form of a n-tuple $X^{(n)}(\omega) = (X^i(\omega))_{1\le i\le n}$ in $\mathcal{C}([0,T];\mathbb{R}^d)^n$. The coefficient driving the equation for $X^i(\omega)$ reads

$$\mathrm{F}\big(X^i_t(\omega), X^{\theta_n(\cdot)}_t(\omega)\big), \quad t \in [0,T],$$

where $\theta_n(\cdot) : \{1,\dots,n\} \ni j \mapsto j$ is the canonical random variable on $\{1,\dots,n\}$. Here the dot in the notation $X^{\theta_n(\cdot)}_t(\omega)$ refers to the current element in $\{1,\dots,n\}$. With this notation, the law of $X^{\theta_n(\cdot)}_t(\omega)$ (on $\{1,\dots,n\}$) must be understood as the empirical distribution $\mu^n_t(\omega)$. Moreover, each $X^i(\omega)$ is controlled, in standard Gubinelli's sense, by the enhanced rough path $(W^i(\omega), \mathbb{W}^i(\omega))$ (the remainder in the expansion being controlled by $v^{i,n}$). In particular, $X^i(\omega)$ may be seen as an i-controlled path on the empirical rough set-up: If we use $\delta^{(n)}_x$ and $\delta^{(n)}_\mu$ as symbols for the Gubinelli derivatives in Definition 2.1 but on the empirical rough set-up, then $\delta^{(n)}_x X^i(\omega)$ identifies with the standard Gubinelli derivative in the expansion of $X^i(\omega)$ along the variations of $(W^i(\omega), \mathbb{W}^i(\omega))$ and $\delta^{(n)}_\mu X^{\cdot}(\omega) \equiv 0$.

The key fact in our analysis lies in the interpretation of the two derivatives

$$\delta^{(n)}_x\big[\mathrm{F}\big(X^i(\omega), X^{\theta_n(\cdot)}(\omega)\big)\big] \quad \text{and} \quad \delta^{(n)}_\mu\big[\mathrm{F}\big(X^i(\omega), X^{\theta_n(\cdot)}(\omega)\big)\big]$$

in Proposition 2.4. First, it is elementary to check that

$$\begin{aligned}\delta^{(n)}_x\big(\mathrm{F}\big(X^i(\omega), X^{\theta_n(\cdot)}(\omega)\big)\big)_t &= \partial_x \mathrm{F}\big(X^i_t(\omega), X^{\theta_n(\cdot)}_t(\omega)\big)\delta^{(n)}_x X^i_t(\omega)\\ &= \partial_x \mathrm{F}\big(X^i_t(\omega), \mu^n_t(\omega)\big)\delta^{(n)}_x X^i_t(\omega).\end{aligned} \tag{4.8}$$

More interestingly, we have

$$\delta^{(n)}_\mu\big(\mathrm{F}\big(X^i(\omega), X^{\theta_n(\cdot)}(\omega)\big)\big)_t = D_\mu \mathrm{F}\big(X^i_t(\omega), \mu^n_t(\omega)\big)\big(X^{\theta_n(\cdot)}_t(\omega)\big)\delta^{(n)}_x X^{\theta_n(\cdot)}_t(\omega), \tag{4.9}$$

both the left- and the right-hand sides being seen as random variables on $\{1,\dots,n\}$. The realizations of the random variable in the right-hand side may be computed by replacing the symbol $\cdot$ by $j \in \{1,\dots,n\}$.

So, applying (2.13) with $\mathrm{F}(X^i(\omega), \mu^n(\omega))$ as integrand, the third term on the first line of (2.13) here reads

$$\frac{1}{n}\sum_{j=1}^{n} D_\mu \mathrm{F}\big(X_t^i(\omega), \mu_t^n\big)\big(X_t^j(\omega)\big)\delta_x^{(n)} X_t^j(\omega)\mathbb{W}_t^{j,i}(\omega).$$

This shows that the integral $\int_0^t \mathrm{F}(X_s^i(\omega), X_s^{\theta_n(\cdot)}(\omega))d\boldsymbol{W}_s^{(n)}(\omega)$, as defined by Theorem 2.3, is the limit of the compensated Riemann sums

$$\begin{aligned}
&\sum_{k=0}^{K-1}\Bigg(\mathrm{F}\big(X_{t_k}^i(\omega), X_{t_k}^{\theta_n(\cdot)}(\omega)\big)W_{t_k,t_{k+1}}^i(\omega)\\
&\quad+\partial_x\mathrm{F}\big(X_{t_k}^i(\omega), X_{t_k}^{\theta_n(\cdot)}(\omega)\big)\mathrm{F}\big(X_{t_k}^i(\omega), X_{t_k}^{\theta_n(\cdot)}(\omega)\big)\mathbb{W}_{t_k,t_{k+1}}^i(\omega)\\
&\quad+\frac{1}{n}\sum_{j=1}^{n} D_\mu\mathrm{F}\big(X_{t_k}^i(\omega), \mu_t^n(\omega)\big)\big(X_{t_k}^j(\omega)\big)\mathrm{F}\big(X_{t_k}^j(\omega), X_{t_k}^{\theta_n(\cdot)}(\omega)\big)\mathbb{W}_{t_k,t_{k+1}}^{j,i}(\omega)\Bigg),
\end{aligned}\tag{4.10}$$

as the mesh of the dissection $0 = t_0 < \cdots < t_K = t$ tends to 0.[4] This allows to compare the latter quantity with (1.3) if we interpret the integral with respect to $W^i(\omega)$ therein as a rough integral with respect to the enhanced setting above $(W^1(\omega),\dots,W^n(\omega))$, and consider the leading coefficient $\mathrm{F}(X_t^i(\omega), \mu_t^n(\omega))$ as a *standard* Euclidean function of the tuple $X_t^{(n)}(\omega) = (X_t^1(\omega),\dots,X_t^n(\omega))$. Indeed, under the standing Regularity assumptions 1 and 2, the function

$$f^i : (\mathbb{R}^d)^n \ni (x^1,\dots,x^n) \mapsto \mathrm{F}\Bigg(x^i, \frac{1}{n}\sum_{k=1}^{n}\delta_{x^k}\Bigg)$$

is $\mathcal{C}^2$ with Lipschitz derivatives and

$$\partial_{x^j} f^i(x^1,\dots,x^n) = \delta_{i,j}\partial_x\mathrm{F}\Bigg(x^i, \frac{1}{n}\sum_{k=1}^{n}\delta_{x^k}\Bigg) + \frac{1}{n}D_\mu\mathrm{F}\Bigg(x^i, \frac{1}{n}\sum_{k=1}^{n}\delta_{x^k}\Bigg)(x^j),$$

with $\delta_{i,j} = 1$ if $i = j$ and 0 otherwise, see Chapter 5 in [14]. Therefore, (1.3) is uniquely solvable in the classical sense and the above formulas for the derivatives show that the rough integral therein may be approximated by the same Riemann sum as in (4.10). Namely, (1.3) may be rewritten as

$$\begin{aligned}
&\sum_{k=0}^{K-1}\Bigg(f^i\big(X_{t_k}^1(\omega),\dots,X_{t_k}^n(\omega)\big)W_{t_k,t_{k+1}}^i(\omega)\\
&\quad+\sum_{j=1}^{n}\partial_{x_j} f^i\big(X_{t_k}^1(\omega),\dots,X_{t_k}^n(\omega)\big)\mathbb{W}_{t_k,t_{k+1}}^{j,i}(\omega)\Bigg).
\end{aligned}$$

This proves that the solution to (1.3), when the latter is seen as a rough differential equation driven by the enhanced setting above $(W^1(\omega),\dots,W^n(\omega))$, coincides with the solution of the empirical version of (1.2), when the latter is understood as a mean field rough differential equation driven by the empirical rough set up.

[4] In the second line, $\partial_x\mathrm{F}(X_s^i(\omega), X_s^{\theta_n(\cdot)}(\omega))(\mathrm{F}(X_s^i(\omega), X_s^{\theta_n(\cdot)}(\omega))\mathbb{W}_{s,t}^i(\omega))$ is understood as $(\sum_{\ell=1}^d\sum_{j,k=1}^m \partial_{x_\ell}\mathrm{F}^{\iota,j}(X_s^i(\omega), X_s^{\theta_n(\cdot)}(\omega))(\mathrm{F}^{\ell,k}(X_s^i(\omega), X_s^{\theta_n(\cdot)}(\cdot)(\omega))(\mathbb{W}_{s,t}^i)^{k,j}(\omega)))_{\iota=1,\dots,d}$ and similarly for the term on the third line.

4.2. *Propagation of chaos.* We now have all the ingredients to prove that the empirical measure of the solution to the particle system (1.3) converges, in some sense, to the solution of the rough mean field equation (1.2), when the rough set-up therein is interpreted as originally explained in Section 2. This is what we call *propagation of chaos*. The statement takes the following form.

THEOREM 4.3. *We make the following assumptions.*

(a) *Let* F *satisfy Regularity assumptions 1 and 2.*

(b) *Let w be a control satisfying (2.8) and (2.9) for the same parameters $p \in [2,3)$ and $q \geq 8$ as in Section 2. Assume that, for a given positive time horizon T, the random variables $w(0,T,\cdot)$ and $(N([0,T],\cdot,\alpha))_{\alpha>0}$, see (2.17), have sub and super exponential tails, see (3.5).*

(c) *Assume that the rough set-up $\boldsymbol{W}$ is strong.*

(d) *Assume also that there exists a positive constant ε_1 such that*

$$\begin{aligned}\mathbb{E}\big[\exp\big(\|W(\cdot)\|^{\varepsilon_1}_{[0,T],(1/p)-\mathrm{H}}\big)\big] + \mathbb{E}\big[\exp\big(\|\mathbb{W}(\cdot)\|^{\varepsilon_1/2}_{[0,T],(2/p)-\mathrm{H}}\big)\big] \\ + \mathbb{E}^{\otimes 2}\big[\exp\big(\|\mathbb{W}^{\mathbb{L}}(\cdot,\cdot)\|^{\varepsilon_1/2}_{[0,T],(2/p)-\mathrm{H}}\big)\big] < \infty.\end{aligned} \tag{4.11}$$

Then, for almost every $\omega \in \Omega$,

$$\frac{1}{n}\sum_{i=1}^{n} \delta_{X^{i,(n)}(\omega)} \to \mathcal{L}\big(X(\cdot)\big), \tag{4.12}$$

where $X^{(n)}(\omega) = (X^{i,(n)}(\omega))_{i=1,\ldots,n}$ is the solution to (1.3) and $X(\cdot)$ is the solution to (1.2), the convergence being the convergence in law on $\mathcal{C}([0,T];\mathbb{R}^d)$. Moreover, for any fixed $k \geq 1$, the law of $(X^{1,(n)}(\cdot),\ldots,X^{k,(n)}(\cdot))$ converges to $\mathcal{L}(X(\cdot))^{\otimes k}$.

REMARK 4.4. Before we prove the above statement, we feel useful to make the following comments:

- In item (b) of the assumption, we can always assume that w is in fact given by the *natural* control (2.10).
- The argument used below in Step 3 of the proof would show that item (d) implies the part related to w in item (b) at least when w is chosen as in (2.10). Despite this form of redundancy, we feel better to keep the current formulation (of the assumptions) as it is consistent with the rest of the text.
- Following [5], Theorem 2.4, the above assumptions hold true for Gaussian rough paths subject to the classical conditions of Friz-Victoir [27], see Example 4.2, and the related Example 2.2 in [5].

PROOF. The key tool for passing to the limit is the continuity Theorem 3.5, but with p therein replaced by some $p' \in (p,3)$. The main difficulty is in controlling the accumulated local variation of the empirical rough set-up. To make the notations clear, we write $X_0^{i,(n)}$ for X^i, $W^{i,(n)}$ for W^i, $\mathbb{W}^{i,(n)}$ for $\mathbb{W}^i$ and $\mathbb{W}^{i,j,(n)}$ for $\mathbb{W}^{i,j}$.

Step 1. As a starting point, we want to prove that, for almost every $\omega \in \Omega$, for any $\alpha > 0$, there exists a constant $\varepsilon_2 > 0$ such that, for all $n \geq 1$,

$$\sup_{n\geq 1} \frac{1}{n}\sum_{i=1}^{n} \exp\big(N^{i,n}(0,T,\omega,\alpha)^{1+\varepsilon_2}\big) < \infty, \tag{4.13}$$

where $N^{i,n}(0,T,\omega,\alpha)$ is defined as the local accumulation

$$N^{i,n}\big([0,T],\omega,\alpha\big) := N_{\varpi}\big([0,T],\alpha\big), \tag{4.14}$$

when $\varpi(s,t) = v_{p'}^{i,n}(s,t,\omega)^{1/p'}$, see (2.17) and (4.3). Following (A.1) in the Appendix (see also the longer discussion in the introduction of the appendix in [5]), it suffices to prove (4.13) when ϖ in the definition of $N^{i,n}$ is equal to each of the terms in the right-hand side of (4.3).

When $\varpi(s,t) = \|W^i(\omega)\|_{[s,t],p'-\mathrm{v}}$ or $\varpi(s,t) = \|\mathbb{W}^i(\omega)\|^{1/2}_{[s,t],p'/2-\mathrm{v}}$, the resulting variables $(N^{i,n}([0,T],\omega,\alpha))_{i=1,\dots,n}$ in (4.14) are independent and identically distributed, their common law being independent of n. Then, (4.13) follows from assumption (b) in the statement and from the law of large numbers (using the well-known fact that the p'-variation is less than or equal to the p-variation if $p' > p$).

If $\varpi(s,t) = {}^{(n)}(W^\bullet(\omega))_{q;[s,t],p'-\mathrm{v}}$ or $\varpi(s,t) = {}^{(n)}((\mathbb{W}^{\bullet,\bullet}(\omega)))^{1/2}_{q;[s,t],p'/2-\mathrm{v}}$, the resulting variables $(N^{i,n}([0,T],\omega,\alpha))_{i=1,\dots,n}$ in (4.14) only depend on n. We may denote them by $N^n([0,T],\omega,\alpha)$. Then, it suffices to prove that, for any $\alpha > 0$, $\limsup_{n\to\infty} N^n([0,T],\omega,\alpha)$ is almost surely finite. By (4.5), we may easily control $N^n([0,T],\omega,\alpha)$ from above by noticing that

$$\begin{aligned}&\alpha^p N^n\big([0,T],\omega,\alpha\big)\\&\quad\le c\big({}^{(n)}\big(\|W^\bullet(\omega)\|^p_{[0,T],(1/p)-\mathrm{H}}\big)_q + {}^{(n)}\big(\big(\|\mathbb{W}^{\bullet,\bullet}(\omega)\|^{p/2}_{[0,T],(2/p)-\mathrm{H}}\big)_q\big),\end{aligned}$$

for a constant c that is independent of n and ω. Proceeding as in (4.6), the result follows again from the law of large of numbers and from assumption (b).

In fact, the most difficult cases are $\varpi(s,t) = {}^{(n)}(\mathbb{W}^{i,\bullet}(\omega))^{1/2}_{q;[s,t],p'/2-\mathrm{v}}$ or $\varpi(s,t) = {}^{(n)}(\mathbb{W}^{\bullet,i}(\omega))^{1/2}_{q;[s,t],p'/2-\mathrm{v}}$. By symmetry, it suffices to treat the first one. And, by changing in an obvious manner the parameter α, we may just focus on $\varpi(s,t) = {}^{(n)}(\mathbb{W}^{i,\bullet}(\omega))^q_{q;[s,t],p'/2-\mathrm{v}}$. Then,

$$^{(n)}\big(\mathbb{W}^{i,\bullet}_{s,t}(\omega)\big)^q_q = \frac{1}{n}\sum_{j=1}^n\big(|\mathbb{W}^{i,j}_{s,t}(\omega)|^q - \big\langle\mathbb{W}^{i,\perp}_{s,t}(\omega,\cdot)\big\rangle^q_q\big) + \big\langle\mathbb{W}^{i,\perp}_{s,t}(\omega,\cdot)\big\rangle^q_q. \tag{4.15}$$

Now, Rosenthal's inequality (see [46]) together with (4.11) say that, for any $a \ge 2$ and any $i \in \{1,\dots,n\}$,

$$\int_\Omega\left|\frac{1}{n}\sum_{j=1}^n\big(|\mathbb{W}^{i,j}_{s,t}(\omega)|^q - \big\langle\mathbb{W}^{i,\perp}_{s,t}(\omega,\cdot)\big\rangle^q_q\big)\right|^a d\mathbb{P}(\omega) \le C_a n^{-a/2}|t-s|^{2aq/p},$$

for a constant C_a depending on a and on the upper bound for the left-hand side in (4.11), but independent of i, n and (s,t). Letting $(t_k^{(n)} = kT/n)_{k=0,\dots,n}$ and allowing the constant C_a to vary from line to line, we deduce that

$$\sum_{1\le k<\ell\le n}\sum_{i=1}^n\int_\Omega\left|\frac{1}{n}\sum_{j=1}^n\big(|\mathbb{W}^{i,j}_{t_k^{(n)},t_\ell^{(n)}}(\omega)|^q - \big\langle\mathbb{W}^{i,\perp}_{t_k^{(n)},t_\ell^{(n)}}(\omega,\cdot)\big\rangle^q_q\big)\right|^a d\mathbb{P}(\omega) \le C_a n^{3-a/2}.$$

We deduce from Markov inequality that, for any $a \ge 2$ and any $n \ge 1$,

$$\begin{aligned}&\mathbb{P}\left(\max_{1\le i\le n}\max_{1\le k<\ell\le n}\left|\frac{1}{n}\sum_{j=1}^n\big(|\mathbb{W}^{i,j}_{t_k^{(n)},t_\ell^{(n)}}(\omega)|^q - \big\langle\mathbb{W}^{i,\perp}_{t_k^{(n)},t_\ell^{(n)}}(\omega,\cdot)\big\rangle^q_q\big)\right| \ge n^{-1/4}\right)\\&\quad\le C_a n^{3-a/4}.\end{aligned} \tag{4.16}$$

By item (d) in the statement, we also have, for any $n \ge 1$, $a \ge 2$ and $\delta > 0$,

$$\begin{aligned}&\mathbb{P}\Big(\max_{1\le i,j\le n}\big(\|W^i\|^2_{[0,T],(1/p)-\mathrm{H}} + \|\mathbb{W}^{i,j}\|_{[0,T],(2/p)-\mathrm{H}}\big) \ge n^{\delta/2}\Big)\\&\quad\le C_a n^{2-a\delta/2},\end{aligned} \tag{4.17}$$

which implies, at least for n large enough (below, we absorb the constant T that appears in the length of the increments by changing $n^{\delta/2}$ into n^{δ}, which is indeed possible since n is large, it being understood that the lower threshold for n only depends on δ and T)

$$\mathbb{P}\Big(\max_{1\le i,j\le n}\sup_{|s-t|\le T/n}|\mathbb{W}^{i,j}_{s,t}(\omega)|\ge n^{\delta-2/p}\Big)\le C_a n^{2-a\delta/2},$$

$$\mathbb{P}\Big(\max_{1\le i,j\le n}\sup_{\min(|s-t|,|s'-t'|)\le T/n}\big(|W^i_{s,t}(\omega)||W^j_{s',t'}(\omega)|\big)\ge n^{\delta-1/p}\Big)\le C_a n^{2-a\delta/2}. \tag{4.18}$$

Similarly, we have (for the same ranges of values for a, δ and n)

$$\mathbb{P}\Big(\max_{1\le i\le n}\big\|\langle\mathbb{W}^{i,\perp\!\!\!\perp}(\omega,\cdot)\rangle_q\big\|_{[0,T],(2/p)-\mathrm{H}}\ge n^{\delta/2}\Big)\le C_a n^{1-a\delta/2}, \tag{4.19}$$

which implies

$$\mathbb{P}\Big(\max_{1\le i\le n}\sup_{|s-t|\le T/n}\langle\mathbb{W}^{i,\perp\!\!\!\perp}_{s,t}(\omega,\cdot)\rangle_q\ge n^{\delta-2/p}\Big)\le C_a n^{1-a\delta/2},$$

$$\mathbb{P}\Big(\max_{1\le i\le n}\sup_{\min(|s-t|,|s'-t'|)\le T/n}\big(|W^i_{s,t}(\omega)|\langle W_{s',t'}(\cdot)\rangle_q\big)\ge n^{\delta-1/p}\Big)\le C_a n^{1-a\delta/2}. \tag{4.20}$$

Using Chen's relations (2.3) to write $\mathbb{W}^{i,j}_{s,t}:=-\mathbb{W}^{i,j}_{\{s\},s}+\mathbb{W}^{i,j}_{\{s\},\{t\}}+\mathbb{W}^{i,j}_{\{t\},t}+W^i_{s,\{t\}}\otimes W^j_{\{t\},t}-W^i_{\{s\},s}\otimes W^j_{s,\{t\}}$ (with $\{s\}:=\lfloor ns/T\rfloor T/n$), we can find a constant c_q only depending on q (but the value of which is allowed to change from line to line) such that, for any $(s,t)\in\mathcal{S}_2^T$,

$$\big||\mathbb{W}^{i,j}_{s,t}(\omega)|^q-|\mathbb{W}^{i,j}_{\{s\},\{t\}}(\omega)|^q\big|\le c_q\,\Xi^{i,j}\sup_{(s',t')\in\mathcal{S}_2^T}|\mathbb{W}^{i,j}_{s',t'}(\omega)|^{q-1}$$

with

$$\Xi^{i,j}:=\sup_{|s'-t'|\le T/n}|\mathbb{W}^{i,j}_{s',t'}(\omega)|+\sup_{\min(|s'-t'|,|s''-t''|)\le T/n}\big(|W^i_{s',t'}(\omega)||W^j_{s'',t''}(\omega)|\big),$$

from which, together with (4.17) and (4.18), we deduce that

$$\mathbb{P}\Big(\max_{1\le i,j\le n}\sup_{(s,t)\in\mathcal{S}_2^T}\big||\mathbb{W}^{i,j}_{s,t}(\omega)|^q-|\mathbb{W}^{i,j}_{\{s\},\{t\}}(\omega)|^q\big|\ge c_q n^{\delta q-1/p}\Big)\le C_a n^{2-a\delta/2}.$$

Proceeding similarly with $\langle\mathbb{W}^{i,\perp\!\!\!\perp}_{s,t}(\omega,\cdot)\rangle_q^q$ and inserting (4.16), we finally obtain, for $a\ge 2$, $\delta>0$ and n large enough (in terms of T and δ only) and for a possibly new value of c_q,

$$\mathbb{P}\left(\max_{1\le i\le n}\sup_{(s,t)\in\mathcal{S}_2^T}\left|\frac{1}{n}\sum_{j=1}^n\big(|\mathbb{W}^{i,j}_{s,t}(\omega)|^q-\langle\mathbb{W}^{i,\perp\!\!\!\perp}_{s,t}(\omega,\cdot)\rangle_q^q\big)\right|\ge c_q\big(n^{-1/4}+n^{\delta q-1/p}\big)\right)\le C_a\big(n^{3-a/4}+n^{2-a\delta/2}\big). \tag{4.21}$$

Meanwhile, we also have, for $\mathbb{P}$-almost every $\omega\in\Omega$,

$$\left|\frac{1}{n}\sum_{j=1}^n\big(|\mathbb{W}^{i,j}_{s,t}(\omega)|^q-\langle\mathbb{W}^{i,\perp\!\!\!\perp}_{s,t}(\omega,\cdot)\rangle_q^q\big)\right|\le\max_{1\le i,j\le n}\big(\|\mathbb{W}^{i,j}(\omega)\|^q_{[0,T],(2/p)-\mathrm{H}}+\|\langle\mathbb{W}^{i,\perp\!\!\!\perp}(\omega,\cdot)\rangle_q\|^q_{[0,T],(2/p)-\mathrm{H}}\big)(t-s)^{2q/p}.$$

By (4.17) and (4.19), we deduce that (again for the same ranges of values for a, δ and n)

$$\mathbb{P}\left(\max_{1\le i\le n}\sup_{(s,t)\in\mathcal{S}_2^T}\left|\frac{1}{n}\sum_{j=1}^{n}\left(\left|\mathbb{W}_{s,t}^{i,j}(\omega)\right|^q-\left\langle\mathbb{W}_{s,t}^{i,\perp\!\!\!\perp}(\omega,\cdot)\right\rangle_q^q\right)\right|\ge n^{\delta q}(t-s)^{2q/p}\right)\le C_a n^{2-a\delta/2}. \tag{4.22}$$

Taking the power $1-p/p'$ in the expression underpinning the event in (4.21) and, similarly, the power p/p' in the expression underpinning the event in (4.22), cross-multiplying both expressions, we get, for n large enough and for $\delta\in(0,1/4)$,

$$\begin{aligned}&\mathbb{P}\Bigg(\max_{1\le i\le n}\sup_{(s,t)\in\mathcal{S}_2^T}\left|\frac{1}{n}\sum_{j=1}^{n}\left(\left|\mathbb{W}_{s,t}^{i,j}(\omega)\right|^q-\left\langle\mathbb{W}_{s,t}^{i,\perp\!\!\!\perp}(\omega,\cdot)\right\rangle_q^q\right)\right|\\&\qquad\ge c_{p,p',q}\left(n^{-(p'-p)/(4p')+\delta qp/p'}+n^{\delta q-(p'-p)/(pp')}\right)(t-s)^{2q/p'}\Bigg)\\&\le C_a\left(n^{3-a/4}+n^{2-a\delta/2}\right),\end{aligned}$$

where $c_{p,p',q}$ only depends on p, p' and q. Choosing δ small enough in terms of p, p' and q, we deduce that for n large enough,

$$\begin{aligned}&\mathbb{P}\left(\max_{1\le i\le n}\sup_{(s,t)\in\mathcal{S}_2^T}\left|\frac{1}{n}\sum_{j=1}^{n}\left(\left|\mathbb{W}_{s,t}^{i,j}(\omega)\right|^q-\left\langle\mathbb{W}_{s,t}^{i,\perp\!\!\!\perp}(\omega,\cdot)\right\rangle_q^q\right)\right|\ge c_{p,p',q}(t-s)^{2q/p'}\right)\\&\le C_a\left(n^{3-a/4}+n^{2-a\delta/2}\right).\end{aligned}$$

Back to (4.15), we deduce that, for δ as in the previous inequality (assuming also without any loss of generality that $\delta\in(0,1/2)$), a large enough (in terms of δ, but δ depending on p, p' and q) and then n large enough (in terms of δ and T) and also for a new value of $c_{p,p',q}$,

$$\begin{aligned}&\mathbb{P}\big(\forall i\in\{1,\dots,n\},\forall(s,t)\in\mathcal{S}_2^T,\ {}^{(n)}\big(\mathbb{W}^{i,\bullet}(\omega)\big)_{q;[s,t],p'/2-\mathrm{v}}\\&\qquad\le c_{p,p',q}(t-s)^{2/p'}+\left\langle\mathbb{W}^{i,\perp\!\!\!\perp}(\omega,\cdot)\right\rangle_{q;[s,t],p'/2-\mathrm{v}}\big)\ge 1-C_a n^{-a\delta/4}.\end{aligned} \tag{4.23}$$

By Borel-Cantelli lemma, we deduce that, for $\mathbb{P}$-almost every $\omega\in\Omega$, for n large enough, for any $i\in\{1,\dots,n\}$ and any $(s,t)\in\mathcal{S}_2^T$,

$${}^{(n)}\big(\mathbb{W}^{i,\bullet}(\omega)\big)_{q;[s,t],p'/2-\mathrm{v}}\le c_{p,p',q}(t-s)^{2/p'}+\left\langle\mathbb{W}^{i,\perp\!\!\!\perp}(\omega,\cdot)\right\rangle_{q;[s,t],p'/2-\mathrm{v}}.$$

Since the variables $(\mathbb{W}^{i,\perp\!\!\!\perp})_{i\ge1}$ are independent, the local accumulation associated with the second term in the right-hand side may be handled like the local accumulation associated to $\varpi(s,t)=\|W^i(\omega)\|_{[s,t],p'-\mathrm{v}}$. The local accumulation associated with the first term is easily handled.

Step 2. Now, from the law of large numbers (see Lemma A.5 for the law of large numbers with second order interaction terms) and from [9], Theorem 2.3 and Problem 3.1, we deduce that there exists a full subset $E\subset\Omega$ (the definition of which may vary from line to line in the rest of the proof as long as $\mathbb{P}(E)$ remains equal to 1) such that, for any $\omega\in E$,

$$\pi_n(\omega)=\left(\frac{1}{n^2}\sum_{i,j=1}^{n}\delta_{(X_0^{i,(n)}(\omega),W^{i,(n)}(\omega),\mathbb{W}^{i,(n)}(\omega),\mathbb{W}^{i,j,(n)}(\omega))}\right)_{n\ge1}$$

converges in the weak sense to $(X_0(\cdot),W(\cdot),\mathbb{W}(\cdot),\mathbb{W}^{\perp\!\!\!\perp}(\cdot,\cdot))$ on the space $\mathbb{R}^d\times\mathcal{C}([0,T];\mathbb{R}^m)\times\{\mathcal{C}(\mathcal{S}_2^T;\mathbb{R}^m\otimes\mathbb{R}^m)\}^2$.

Step 3. Back to the statement of Theorem 3.5, the first item in the statement is a consequence of the law of large numbers. As for the fourth item, it follows directly from the previous step. In order to check the check the second and third items, we now have a look at $v_{p'}^{i,n}(s,t,\omega)$ in (4.3). Following (4.6), we already know that

$$\limsup_{n\geq 1} \sup_{0\leq s<t\leq T} \frac{{}^{(n)}(v_{p'}^{\bullet,n}(s,t,\omega))_{2q}}{t-s} < \infty,$$

which proves the third item in the statement of Theorem 3.5. We end up with the proof of the second item. Following (4.5), there exists a constant c' such that, for any $\varepsilon > 0$, the quantity

$$\sup_{n\geq 1} {}^{(n)}\big(\exp\big([v_{p'}^{\bullet,n}(0,T,\omega)]^{\varepsilon}\big)\big)_1 \tag{4.24}$$

is finite if (notice that we could work below with $1/p'$ instead of $1/p$–Hölder norms, but $1/p$ obviously suffices and is in fact more adapted to the assumption (4.11))

$$\begin{aligned} &\sup_{n\geq 1} \frac{1}{n}\sum_{i=1}^{n} \exp\big(c'\|W^i(\omega)\|_{[0,T],(1/p)-\mathrm{H}}^{p'\varepsilon} + c'\|\mathbb{W}^i(\omega)\|_{[0,T],(2/p)-\mathrm{H}}^{p'\varepsilon/2}\big) < \infty, \\ &\sup_{n\geq 1} \frac{1}{n}\sum_{i=1}^{n} \exp\big(c'^{(n)}\big(\|\mathbb{W}^{\bullet,i}(\omega)\|_{[0,T],(2/p)-\mathrm{H}}^{p'/2}\big)_q^{\varepsilon}\big) < \infty, \end{aligned} \tag{4.25}$$

and similarly on the second line of (4.25) with $\mathbb{W}^{\bullet,i}(\omega)$ replaced by $\mathbb{W}^{i,\bullet}(\omega)$ or $\mathbb{W}^{\bullet,\bullet}(\omega)$ (the last two terms appearing on the last line of (4.5)). By the law of large numbers, the first line holds true on a full event if $p'\varepsilon < \varepsilon_1$. As for the second one, we use the following trick. Notice that the function

$$(0,+\infty) \ni x \mapsto \exp\big(x^{\varepsilon/q}\big), \tag{4.26}$$

is convex on $[A_\varepsilon, \infty)$, for some $A_\varepsilon > 0$. Therefore, Jensen's inequality says that, in order to check the second line in (4.25), it suffices to prove that

$$\sup_{n\geq 1} \frac{1}{n^2}\sum_{i,j=1}^{n} \exp\big[\big(A_\varepsilon^{\varepsilon/q} \vee \|\mathbb{W}^{i,j}(\omega)\|_{[0,T],(2/p)-\mathrm{H}}^{p'\varepsilon/2}\big)\big] < \infty, \tag{4.27}$$

and similarly for the two terms appearing on the last line of (4.5). Obviously, under the standing assumption, the latter holds true with probability 1 provided $p'\varepsilon < \varepsilon_1$. This proves (4.24). In the statement of Theorem 3.5, this proves the condition related to the tails of w^n by a standard application of Markov inequality.

The bound on the local accumulation in the second item of Theorem 3.5 follows from the first step of the proof.

Step 4. By Theorem 3.5, we get (4.12) on a set of full measure. By Proposition 2.2 in [47], we deduce that, for any fixed $k \geq 1$, the law of $(X^{1,(n)}, \ldots, X^{k,(n)})$ converges to $\mathcal{L}(X(\cdot))^{\otimes k}$. □

REMARK 4.5. Recently, the authors in [19] obtained a quantified propagation of chaos result for mean field stochastic equations with additive noise

$$dx_t = b\big(x_t, \mathcal{L}(x_t)\big)\,dt + dw_t, \quad x_0 = \zeta, \tag{4.28}$$

for a random path $w \in \mathcal{C}([0,T],\mathbb{R}^d)$ subject to mild integrability condition, and random initial condition ζ. There is no need of rough paths theory to make sense of this equation and solve it by elementary means, under proper regularity assumptions on the drift b. Its distribution is even a Lipschitz function of the distribution of (ζ, w), in p-Wasserstein metric.

Using Tanaka's trick, this continuity result entails a propagation of chaos result. The global Lipscthiz continuity of the solution map $\mathcal{L}(w,\zeta)\mapsto\mathcal{L}(x)$ ensures in particular a quantitative convergence rate for the particle system no greater than the corresponding convergence rate for the sample empirical mean of the driving noises, which is optimal. We get back such a sharp estimate in the present, much more complicated, setting in the next section. Note that the global Lipschitz bound satisfied by the natural map Φ giving the solution to equation (4.28) as a fixed point of Φ actually allows to deal with reflected dynamics, as the bounded variation part needed for the reflection happens to be a Lipschitz function of the nonreflected path, in Skorokhod formulation of the problem. We do not have such a strong continuity result for our solution map; see Theorem 3.5. See also the previous work [25] of the authors.

5. Rate of convergence. The goal of this section is to elucidate the rate of convergence in the convergence result stated in Theorem 4.3.

The analysis is based upon a variation of Sznitman's original coupling argument, see [47]. To make its principle clear, we recall that, on the space $(\Omega,\mathcal{F},\mathbb{P})$, the triples $(X_0^1(\cdot),W^1(\cdot),\mathbb{W}^1(\cdot)),\dots,(X_0^n(\cdot),W^n(\cdot),\mathbb{W}^n(\cdot))$ are n independent copies of the original triple $(X_0(\cdot),W(\cdot),\mathbb{W}(\cdot))$. For each $i\in\{1,\dots,n\}$, the pair $(W^i(\cdot),\mathbb{W}^i(\cdot))$ is completed into a rough set-up

$$\begin{aligned}\overline{\boldsymbol{W}}^i(\cdot)&:=\big(W^i(\cdot),\mathbb{W}^i(\cdot),\mathbb{W}^{i,\perp}(\cdot,\cdot)\big),\\ \mathbb{W}^{i,\perp}(\omega,\omega')&=\mathcal{I}\big(W^i(\omega),W^i(\omega')\big),\quad(\omega,\omega')\in\Omega^2.\end{aligned}\tag{5.1}$$

Here we put a bar on the symbol $\overline{\boldsymbol{W}}^i$ in order to distinguish it from the finite-dimensional rough set-up $\boldsymbol{W}^{(n)}(\omega)$ that lies above $(W^1(\omega),\dots,W^n(\omega))$. In comparison, the second-order level of $\boldsymbol{W}^{(n)}$ is made of $(\mathbb{W}^i)_{1\le i\le n}$ and of $(\mathbb{W}^{i,j}=\mathcal{I}(W^i,W^j))_{1\le i\ne j\le n}$, see (4.2). To make the notations more homogeneous, we sometimes write $\mathbb{W}^{i,i}(\omega)$ for $\mathbb{W}^i(\omega)$.

With each $(X_0^i(\cdot),\overline{\boldsymbol{W}}^i(\cdot))$, we associate the corresponding solution $\overline{X}^i(\cdot)$ to the mean field equation (1.2). The 5-tuples

$$\Omega\ni\omega\mapsto\big(X_0^i(\omega),W^i(\omega),\mathbb{W}^i(\omega),\mathbb{W}^{i,\perp}(\cdot,\omega),\overline{X}^i(\omega)\big)_{1\le i\le n}$$

are independent and identically distributed, $\Omega\ni\omega\mapsto(\mathbb{W}_t^{i,\perp}(\cdot,\omega))_{0\le t\le T}$ being regarded as a process with values in $\mathbb{L}^q(\Omega,\mathcal{F},\mathbb{P};\mathbb{R}^d)$. Recalling that $X^{(n)}(\omega)=(X^{1,(n)}(\omega),\dots,X^{n,(n)}(\omega))$ is the solution to (1.3), we then let

$$\mu_t^n(\omega)=\frac{1}{n}\sum_{i=1}^n\delta_{X_t^{i,(n)}(\omega)},\qquad\overline{\mu}_t^n(\omega)=\frac{1}{n}\sum_{i=1}^n\delta_{\overline{X}_t^i(\omega)},\quad t\in[0,T],\omega\in\Omega.\tag{5.2}$$

Here is now the main result. Note the use of the $\mathbf{d}_1$-distance (see (2.14)) in the assumption required from F in the statement below, $\mathbf{d}_1$-continuity being stronger than $\mathbf{d}_2$-continuity.

THEOREM 5.1. *We make the following assumptions.*

(a) *Assumptions* (a)–(d) *in the statement of Theorem 4.3 are satisfied for the same parameters* $p\in[2,3)$ *and* $q\ge 8$ *as in Section 2, for some control* w *satisfying* (2.8) *and* (2.9) *and some time horizon* $T>0$.

(b) *The first and second derivatives of* F, $(x,\mu)\mapsto\partial_x\mathrm{F}(x,\mu)$, $(x,\mu,z)\mapsto(D_\mu\mathrm{F}(x,\mu)(z),\partial_xD_\mu\mathrm{F}(x,\mu)(z))$, *and* $(x,\mu,z,z')\mapsto D_\mu^2\mathrm{F}(x,\mu)(z,z')$, *are bounded on the whole space and are Lipschitz continuous with respect to all the variables, the Lipschitz property in the direction* μ *being understood with respect to* $\mathbf{d}_1$.

(c) *Last, for any $\alpha > 0$, there exists a constant $\varepsilon_2 > 0$ such that, for some $p' \in [p, 3)$, and any random variables $\tau, \tau' : \Omega \to [0, T]$, with $\mathbb{P}(\tau < \tau') = 1$, we have*

$$\sup_{n\geq 1} \sup_{1\leq i\leq n} \mathbb{E}\Big[\exp\Big[\Big(\frac{\widehat{N}^{i,n}([\tau,\tau'],\omega,\alpha)}{\sqrt{\tau'-\tau}}\Big)^{1+\varepsilon_2}\Big]\Big] < \infty, \tag{5.3}$$

where $\widehat{N}^{i,n}([\tau,\tau'],\omega,\alpha)$ is defined as the accumulation $N_{\varpi}([\tau,\tau'],\alpha)$ when $\varpi = (\widehat{w}^{i,n}_{p'}(\omega))^{\frac{1}{p'}}$ with

$$\begin{aligned}
\widehat{w}^{i,n}_{p'}(s,t,\omega) &:= \big(w^{i,n}_{p'} + \widehat{v}^{i,n}_{p'}\big)(s,t,\omega) + {}^{(n)}(\widehat{v}^{\bullet,n}_{p'}(\omega))_{q;[s,t],1-\mathrm{v}} + (t-s),\\
w^{i,n}_{p'}(s,t,\omega) &:= v^{i,n}_{p'}(s,t,\omega) + {}^{(n)}(v^{\bullet,n}_{p'}(\omega))_{q;[s,t],1-\mathrm{v}},\\
\widehat{v}^{i,n}_{p'}(s,t,\omega) &:= \big\langle \mathbb{W}^{i,\perp}(\omega,\cdot)\big\rangle^{p'/2}_{q;[s,t],p'/2-\mathrm{v}} + \big\langle \mathbb{W}^{i,\perp}(\cdot,\omega)\big\rangle^{p'/2}_{q;[s,t],p'/2-\mathrm{v}}.
\end{aligned} \tag{5.4}$$

Then, for any $r \geq 1$, there exists an exponent $q(r) \geq 8$ such that, if $X_0(\cdot)$ is in $\mathbb{L}^{q(r)}$, then, for any $n \geq 1$,

$$\sup_{1\leq i\leq n} \mathbb{E}\Big[\sup_{0\leq t\leq T} |\overline{X}^i_t - X^{i,(n)}_t|^r\Big]^{1/r} + \mathbb{E}\Big[\sup_{0\leq t\leq T} \mathbf{d}_1(\mu^n_t, \overline{\mu}^n_t)^r\Big]^{1/r} \leq C\varsigma_n, \tag{5.5}$$

for a constant C independent of n, and $\varsigma_n = n^{-1/2}$ if $d = 1$, $\varsigma_n = n^{-1/2}\ln(1+n)$ if $d = 2$ and $\varsigma_n = n^{-1/d}$ if $d \geq 3$.

REMARK 5.2. Let us make a few remarks on this statement before embarking on its proof.

- We refer to [14], Chapter 5, for examples of a function F satisfying item *(b)* in the assumptions of the statement. Importantly, we recall that a function $G : \mathcal{P}_2(\mathbb{R}^d) \ni \mu \mapsto G(\mu) \in \mathbb{R}$, whose derivative $D_\mu G : \mathcal{P}_2(\mathbb{R}^d) \times \mathbb{R}^d \ni (\mu, z) \mapsto D_\mu G(\mu)(z) \in \mathbb{R}^d$ is uniformly bounded on the whole $\mathcal{P}_2(\mathbb{R}^d) \times \mathbb{R}^d$, is Lipschitz continuous with respect to the $\mathbf{d}_1$-Wasserstein distance. In particular, under the assumptions of the statement, F itself is Lipschitz continuous on $\mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d)$, the Lipschitz property in the direction μ being understood with respect to $\mathbf{d}_1$.
- Obviously, condition (5.3) depends on p'. We let the reader check that if (5.3) holds for some $p' \in [p, 3)$, then it holds for any other $p'' \in [p', 3)$.
- By inspecting the proof of Theorem 5.1, we could make explicit the value of $q(r)$ (in the condition $X_0(\cdot) \in \mathbb{L}^{q(r)}$), but we feel that it would not be so useful.
- The convergence rate ς_n in (5.5) corresponds to the usual rate for the convergence in the 1-Wasserstein distance of an empirical sample of independent, identically distributed, random variables toward the limiting common distribution; see [26] together with Lemma A.4.
- Theorem 5.1 applies when W is a continuous centred Gaussian process defined over $[0, T]$ as in Example 4.2.

PROOF. Observe that, for each $i \in \{1, \dots, n\}$ and any $\omega \in \Omega$, we can define the integral process $(\int_0^t \mathrm{F}(\overline{X}^i_s(\omega), \overline{\mu}^n_s(\omega))\, d\boldsymbol{W}^{i,(n)}_s(\omega))_{0\leq t\leq T}$ using usual rough paths theory, where the label i in the notation $\boldsymbol{W}^{i,(n)}(\omega)$ is here to indicate that the integral only involves $(W^i(\omega), (\mathbb{W}^{j,i}(\omega))_{1\leq j\leq n})$. Equivalently, $\boldsymbol{W}^{i,(n)}(\omega)$ must be seen as $(W^i(\omega), (\mathbb{W}^{j,i}(\omega))_{1\leq j\leq n})$. The fact that the integral may be defined with respect to $(W^i(\omega), (\mathbb{W}^{j,i}(\omega))_{1\leq j\leq n})$ follows from the fact that $\overline{X}^j(\omega)$, for each $j \in \{1, \dots, n\}$ and each $\omega \in \Omega$, is controlled by the variations of the sole $W^j(\omega)$. So, whenever we expand locally $\mathrm{F}(\overline{X}^i_s(\omega), \overline{\mu}^n_s(\omega))$, we let appear increments of $\overline{X}^i_s(\omega)$, which are controlled by increments of $W^i(\omega)$, and also increments of

$\overline{X}^j_s(\omega)$, for $j \neq i$, which are controlled by increments of $W^j(\omega)$: At the end of the day, it suffices to have the iterated integrals $(\mathbb{W}^{j,i}(\omega))_{1\le j\le n}$ to define the integral $(\int_0^t \mathrm{F}(\overline{X}^i_s(\omega), \overline{\mu}^n_s(\omega))\, d\boldsymbol{W}^{i,(n)}_s(\omega))_{0\le t\le T}$.

Step 1. The first step is to compare

$$\int_0^t \mathrm{F}\big(\overline{X}^i_s(\omega), \mathcal{L}(X_s)\big)\, d\overline{\boldsymbol{W}}^i_s(\omega) \quad \text{and} \quad \int_0^t \mathrm{F}\big(\overline{X}^i_s(\omega), \overline{\mu}^n_s(\omega)\big)\, d\boldsymbol{W}^{i,(n)}_s(\omega), \tag{5.6}$$

for $t \in [0, T]$. What makes the proof nontrivial is the fact that the rough set-ups used in the first and the second integrals are not the same. So, in order to compare the two of them, we need to come back to the original constructions of the two integrals. To simplify notations, and for $0 \le t \le T$, set

$$\overline{F}^i_t(\omega) := \mathrm{F}\big(\overline{X}^i_t(\omega), \mathcal{L}(X_t)\big), \qquad F^{i,n}_t(\omega) := \mathrm{F}\big(\overline{X}^i_t(\omega), \overline{\mu}^n_t(\omega)\big). \tag{5.7}$$

For sure, $(\overline{F}^i_t(\omega))_{0\le t\le T}$ is ω-controlled by $\overline{\boldsymbol{W}}^i(\omega)$, see Definition 2.1 and Proposition 2.4, and the collection indexed by $\omega \in \Omega$ is a random path controlled by $\overline{\boldsymbol{W}}^i$, see Definition 2.2 for a reminder. The corresponding Gubinelli derivatives are denoted by $(\delta_x \overline{F}^i_t(\omega), \delta_\mu \overline{F}^i_t(\omega, \cdot))_{0\le t\le T}$, see again Proposition 2.4. Similarly, $(F^{i,n}_t(\omega))_{0\le t\le T}$ is controlled by $\boldsymbol{W}^{i,(n)}(\omega)$ and $\boldsymbol{W}^{\bullet,(n)}(\omega)$ and Gubinelli derivatives are encoded in the form of a collection $(\delta_x F^{i,n}_t(\omega), (\delta_\mu F^{i,j,n}_t(\omega))_{1\le j\le n})_{0\le t\le T}$, see ((4.8)–(4.9)). To make it clear, set

$$\begin{aligned} \delta_x \overline{F}^i_t(\omega) &:= \partial_x \mathrm{F}\big(\overline{X}^i_t(\omega), \mathcal{L}(X_t)\big)\mathrm{F}\big(\overline{X}^i_t(\omega), \mathcal{L}(X_t)\big), \\ \delta_\mu \overline{F}^i_t(\omega, \cdot) &:= D_\mu \mathrm{F}\big(\overline{X}^i_t(\omega), \mathcal{L}(X_t)\big)\big(\overline{X}^i_t(\cdot)\big)\mathrm{F}\big(\overline{X}^i_t(\cdot), \mathcal{L}(X_t)\big), \end{aligned} \tag{5.8}$$

where $X(\cdot)$ is the solution to (1.2) with $\boldsymbol{W}(\cdot) = (W(\cdot), \mathbb{W}(\cdot), \mathbb{W}^{\perp}(\cdot, \cdot))$. We also let

$$\begin{aligned} \delta_x F^{i,n}_t(\omega) &:= \partial_x \mathrm{F}\big(\overline{X}^i_t(\omega), \overline{\mu}^n_t(\omega)\big)\mathrm{F}\big(\overline{X}^i_t(\omega), \overline{\mu}^n_t(\omega)\big), \\ \delta_\mu F^{i,j,n}_t(\omega) &:= D_\mu \mathrm{F}\big(\overline{X}^i_t(\omega), \overline{\mu}^n_t(\omega)\big)\big(\overline{X}^j_t(\omega)\big)\mathrm{F}\big(\overline{X}^j_t(\omega), \overline{\mu}^n_t(\omega)\big). \end{aligned} \tag{5.9}$$

For a subdivision $\Delta = \{s = t_0 < t_1 < \cdots < t_K = t\}$, set

$$\begin{aligned} \overline{\mathcal{I}}^{i,\Delta}_{s,t}(\omega) &:= \sum_{k=0}^{K-1} \big\{ \overline{F}^i_{t_k}(\omega) W^i_{t_k,t_{k+1}}(\omega) + \delta_x \overline{F}^i_{t_k}(\omega) \mathbb{W}^i_{t_k,t_{k+1}}(\omega) \\ &\qquad + \mathbb{E}\big[\delta_\mu \overline{F}^i_{t_k}(\omega, \cdot) \mathbb{W}^{i,\perp}_{t_k,t_{k+1}}(\cdot, \omega)\big]\big\}, \\ \mathcal{I}^{i,n,\Delta}_{s,t}(\omega) &:= \sum_{k=0}^{K-1} \Big\{ F^{i,n}_{t_k}(\omega) W^i_{t_k,t_{k+1}}(\omega) + \delta_x F^{i,n}_{t_k}(\omega) \mathbb{W}^i_{t_k,t_{k+1}}(\omega) \\ &\qquad + \frac{1}{n} \sum_{j=1}^n \delta_\mu F^{i,j,n}_{t_k}(\omega) \mathbb{W}^{j,i}_{t_k,t_{k+1}}(\omega) \Big\}. \end{aligned} \tag{5.10}$$

The two integrals in (5.6) should be understood as the respective limits of the two Riemann sums right above as K tends to ∞. In the sequel, we denote the summand in the first sum by $\overline{\mathcal{I}}^{i,\partial}_{\{t_k,t_{k+1}\}}(\omega)$ and the summand in the second sum by $\mathcal{I}^{i,n,\partial}_{\{t_k,t_{k+1}\}}(\omega)$. By Lemma A.3 proved in Appendix A.2, we can find, for any $\varrho \ge 8$, an exponent $\varrho' \ge q$, independent of n, K and Δ, such that, whenever $X_0(\cdot) \in \mathbb{L}^{\varrho'}$, it holds, for a constant C, also independent of n, K and Δ but depending on $\langle X_0(\cdot)\rangle_{\varrho'}$, and then for any $k \in \{1, \ldots, K-1\}$ (provided $K \ge 2$),

$$\big\langle \big\{\mathcal{I}^{i,n,\Delta}_{s,t}(\cdot) - \mathcal{I}^{i,n,\Delta'}_{s,t}(\cdot)\big\} - \big\{\overline{\mathcal{I}}^{i,\Delta}_{s,t}(\cdot) - \overline{\mathcal{I}}^{i,\Delta'}_{s,t}(\cdot)\big\}\big\rangle_{\varrho} \le C\varsigma_n \big\| w^+(t_{k-1}, t_{k+1}, \cdot, \cdot)\big\|^{3/p}_{\varrho'},$$

where $\Delta' := \Delta \setminus \{t_k\}$ and $w^+(s,t,\omega,\omega') := w(s,t,\omega) + \|\mathbb{W}^{\perp}(\omega,\omega')\|^{p/2}_{[s,t],p/2-\mathrm{v}}$, ($w$ being here given by (2.10)). In the rest of the proof, it is implicitly understood that the condition $X_0(\cdot) \in \mathbb{L}^{\varrho'}$ is satisfied for ϱ' large enough and that the constant C is allowed to depend on $\langle X_0(\cdot)\rangle_{\varrho'}$.

Formulating (4.5) and (4.6) but for the limit (instead of empirical) rough set-up, we know that the right-hand side in the above inequality is less than

$$C\varsigma_n\big[\big\langle \|W(\cdot)\|_{[0,T],(1/p)-\mathrm{H}}\big\rangle_{p\varrho'} + \big\langle \|\mathbb{W}(\cdot)\|_{[0,T],(2/p)-\mathrm{H}}\big\rangle^{1/2}_{p\varrho'} + \big\langle\!\big\langle \|\mathbb{W}^{\perp}(\cdot,\cdot)\|_{[0,T],(2/p)-\mathrm{H}}\big\rangle\!\big\rangle^{1/2}_{p\varrho'}\big]^3 (t_{k+1}-t_{k-1})^{3/p},$$

but by assumption all the expectations are finite. Now we can choose k such that $|t_{k+1} - t_{k-1}| \le 4|t-s|/K$ (if not, it means that $4(t-s)(K-1)/K < \sum_{k=1}^{K-1} |t_{k+1} - t_{k-1}| = \sum_{k=1}^{K-1}(t_{k+1} - t_k + t_k - t_{k-1}) = 2(t-s) - (t_K - t_{K-1} + t_1 - t_0) \le 2(t-s)$, which is a contradiction since $K \ge 2$). We get

$$\big\langle \{\mathcal{I}^{i,n,\Delta}_{s,t}(\cdot) - \mathcal{I}^{i,n,\Delta'}_{s,t}(\cdot)\} - \{\overline{\mathcal{I}}^{i,\Delta}_{s,t}(\cdot) - \overline{\mathcal{I}}^{i,\Delta'}_{s,t}(\cdot)\}\big\rangle_{\varrho} \le C\varsigma_n \left(\frac{t-s}{K}\right)^{3/p},$$

the constant C being allowed to increase from line to line as long as it remains independent of n, K, Δ and Δ'. Letting $t^{(1)} = t_k$ and applying iteratively the above bound to a sequence of meshes of the form $\Delta \setminus \{t^{(1)}\}$, $\Delta \setminus \{t^{(1)}, t^{(2)}\}, \ldots$, and then letting K tend to ∞, we deduce that

$$\begin{aligned} &\Big\langle \int_s^t F^{i,n}_r(\cdot)\, d\boldsymbol{W}^{i,(n)}_r(\cdot) - \int_s^t \overline{F}^i_r(\cdot)\, d\overline{\boldsymbol{W}}^i_r(\cdot) - \{\mathcal{I}^{i,n,\partial}_{\{s,t\}} - \overline{\mathcal{I}}^{i,\partial}_{\{s,t\}}\}\Big\rangle_{\varrho} \\ &\quad \le C\varsigma_n (t-s)^{3/p}. \end{aligned} \tag{5.11}$$

By Lemma A.3, we also have $\langle \mathcal{I}^{i,n,\partial}_{\{s,t\}} - \overline{\mathcal{I}}^{i,\partial}_{\{s,t\}}\rangle_{\varrho} \le C\varsigma_n (t-s)^{1/p}$, from which we deduce that

$$\Big\langle \int_s^t F^{i,n}_r(\cdot)\, d\boldsymbol{W}^{i,(n)}_r(\cdot) - \int_s^t \overline{F}^i_r(\cdot)\, d\overline{\boldsymbol{W}}^i_r(\cdot)\Big\rangle_{\varrho} \le C\varsigma_n (t-s)^{1/p}.$$

Similarly, Lemma A.3 says that $\langle [F^{i,n}(\cdot) - \overline{F}^i(\cdot)]_{s,t}\rangle_{\varrho} \le C\varsigma_n (t-s)^{1/p}$, and, noting that

$$\begin{aligned} R^{\int F^{i,n} d\boldsymbol{W}^{i,(n)}}_{s,t}(\omega) &= \int_s^t F^{i,n}_r(\omega)\, d\boldsymbol{W}^{i,(n)}_r(\omega) - \mathcal{I}^{i,n,\partial}_{s,t}(\omega) \\ &\quad + \delta_x F^{i,n}_s(\omega)\mathbb{W}^i_{s,t}(\omega) + \frac{1}{n}\sum_{j=1}^n \delta_\mu F^{i,j,n}_s(\omega)\mathbb{W}^{j,i}_{s,t}(\omega), \\ R^{\int \overline{F}^i d\overline{\boldsymbol{W}}^i}_{s,t}(\omega) &= \int_s^t \overline{F}^i_r(\omega) d\overline{\boldsymbol{W}}^i_r(\omega) - \overline{\mathcal{I}}^{i,\partial}_{s,t}(\omega) \\ &\quad + \delta_x \overline{F}^i_s(\omega)\mathbb{W}^i_{s,t}(\omega) + \mathbb{E}\big[\delta_\mu \overline{F}^i_s(\omega,\cdot)\mathbb{W}^{i,\perp}_{s,t}(\cdot,\omega)\big], \end{aligned}$$

we deduce in a similar manner, using (5.11) and Lemma A.3 once again, that

$$\big\langle R^{\int F^{i,n} d\boldsymbol{W}^{i,(n)}}_{s,t}(\cdot) - R^{\int \overline{F}^i d\overline{\boldsymbol{W}}^i}_{s,t}(\cdot)\big\rangle_{\varrho} \le C\varsigma_n (t-s)^{2/p}.$$

So, fixing $i \in \{1,\ldots,n\}$, choosing ϱ large enough and applying a suitable version of Kolmogorov's theorem (see, for instance, Theorem 3.1 in [28]), we can find $p' \in (p,3)$, which

may be assumed to satisfy item (c) in the assumption (see, for instance, the second bullet point in Remark 5.2) and the value of which is fixed until the end of the proof, such that

$$\begin{aligned}\left|\int_s^t F_r^{i,n}(\omega)\,d\boldsymbol{W}_r^{i,(n)} - \int_s^t \overline{F}^i_r(\omega)\,d\overline{\boldsymbol{W}}^i_r(\omega)\right| &\le \theta^{i,n}(\omega)(t-s)^{1/p'},\\ \left|\big[F^{i,n}(\omega) - \overline{F}^i(\omega)\big]_{s,t}\right| &\le \theta^{i,n}(\omega)(t-s)^{1/p'},\\ \left|R^{\int F^{i,n}\,d\boldsymbol{W}^{i,(n)}}_{s,t}(\omega) - R^{\int \overline{F}^i\,d\overline{\boldsymbol{W}}^i}_{s,t}(\omega)\right| &\le \theta^{i,n}(\omega)(t-s)^{2/p'},\end{aligned} \tag{5.12}$$

with $\langle\theta^{i,n}(\cdot)\rangle_{\varrho} \le C\varsigma_n$, for a new value of the constant C.

Observe now that the empirical control associated with our empirical rough set-up and with the exponent p' reads (compare with (2.10))

$$w^{i,n}_{p'}(s,t,\omega) := v^{i,n}_{p'}(s,t,\omega) + {}^{(n)}(v^{\bullet,n}_{p'}(\omega))_{q;[s,t],1-\mathrm{v}},$$

where we used the same notation as in (4.3). In fact, there is no loss of generality in changing the definition of $w^{i,n}_{p'}$ into

$$w^{i,n}_{p'}(s,t,\omega) := v^{i,n}_{p'}(s,t,\omega) + {}^{(n)}(v^{\bullet,n}_{p'}(\omega))_{q;[s,t],1-\mathrm{v}} + (t-s), \tag{5.13}$$

which permits to replace $(t-s)^{1/p'}$ by $w^{i,n}_{p'}(s,t,\omega)^{1/p'}$ in the inequalities (5.12). Hence,

$$\left|\!\left|\!\left|\int_0^{\cdot} F_r^{i,n}(\omega)\,d\boldsymbol{W}_r^{i,(n)} - \int_0^{\cdot} \overline{F}^i_r(\omega)\,d\overline{\boldsymbol{W}}^i_r(\omega)\right|\!\right|\!\right|_{[0,T],w^{i,n}_{p'},p'} \le \theta^{i,n}(\omega).$$

Step 2. We now make use of Proposition 3.2 to compare

$$\int_0^t \mathrm{F}\big(X^{i,(n)}_s(\omega),\mu^n_s(\omega)\big)\,d\boldsymbol{W}^{i,(n)}_s(\omega) \quad \text{and} \quad \int_0^t \mathrm{F}\big(\overline{X}^i_s(\omega),\overline{\mu}^n_s(\omega)\big)\,d\boldsymbol{W}^{i,(n)}_s(\omega),$$

see (5.2). To simplify the notations, we just write X^i for $X^{i,(n)}$ and $\boldsymbol{W}^i$ for $\boldsymbol{W}^{i,(n)}$. We then apply Proposition 3.2 with

$$\big(X(\omega),Y(\cdot)\big) = \big(X^i(\omega),X^{\bullet}(\omega)\big), \qquad \big(X'(\omega),Y'(\cdot)\big) = \big(\overline{X}^i(\omega),\overline{X}^{\bullet}(\omega)\big), \tag{5.14}$$

the underlying set-up being understood as the empirical rough set-up for a given realization ω (in particular, the various assumptions on the moments in Proposition 3.2 must be checked under the empirical distribution). The difficulty here is that the variations of these two solutions are controlled by two different functionals w, see (2.11); in short, there is the control associated with the empirical set-up and the control associated with the theoretical one. This is the rationale for introducing $\widehat{w}^{i,n}_{p'}$ in (5.4). Obviously, $\widehat{w}^{i,n}_{p'}(\cdot,\cdot,\omega)$ is not the *natural* control functional associated with $\boldsymbol{W}^i(\omega)$, but it is greater than $w^{i,n}_{p'}(s,t,\omega)$ and it satisfies

$${}^{(n)}(\widehat{w}^{\bullet,n}_{p'}(s,t,\omega))_q \le 2\widehat{w}^{i,n}_{p'}(s,t,\omega), \tag{5.15}$$

which suffices to apply Proposition 3.2, see also [5], Proposition 4.3, with $w^{i,n}_{p'}(s,t,\omega)$ replaced by $\widehat{w}^{i,n}_{p'}(s,t,\omega)$. The resulting seminorm that must be used to control the difference $(X(\omega)-X'(\omega),Y(\cdot)-Y'(\cdot)) = (X^i(\omega)-\overline{X}^i(\omega),X^{\bullet}(\omega)-\overline{X}^{\bullet}(\omega))$ on a given interval $[s,t]$ is $|\!|\!|\cdot|\!|\!|_{[s,t],\widehat{w}^{i,n},p'}$ (we feel easier not to put the index p' in $\widehat{w}^{i,n}$ but it is implicitly understood). We use the corresponding local accumulation, which we denote by $\widehat{N}^{i,n}([0,T],\omega,\alpha)$.

We first check that the pair $(X(\omega),Y(\cdot)) = (X^i(\omega),X^{\bullet}(\omega))$ in (5.14) satisfies the assumptions of Proposition 3.2, assuming that T is upper bounded by 1 (which may seem rather restrictive but which is in fact consistent with what we do in the sequel since we require, at

least for a while, T to be small enough). By construction of the processes $(X^j(\omega))_{j=1,\dots,n}$ as the solution of the empirical rough equation, we know from Theorem 3.3 and from [5], Proposition 4.2, (which guarantees a form of stability in the iterative construction of the solution) that we can find three deterministic constants L_0, η_0 and γ_0, all strictly greater than 1 and only depending on Λ (we may forget the dependence upon T since $T \le 1$; in particular the three constants are independent of n and, for sure, of the index i as well), such that $X(\omega) = X^i(\omega)$ satisfies (3.3) with $w = \widehat{w}^{i,n}_{p'}$ (it being understood that the points $(t^0_\ell = \tau_\ell(0, T, \omega, 1/(4L_0)))_{\ell=0,\dots,N^0+1}$ in the statement of Proposition 3.2 are constructed with respect to $\widehat{w}^{i,n}_{p'}$) and $Y(\cdot) = X^\bullet(\omega)$ satisfies condition (3.2) with respect to ${}^{(n)}(\cdot)_8$, provided that T satisfies

$$ {}^{(n)}\big(\widehat{N}^{\bullet,n}([0,T],\omega,1/(4L_0))\big)_8 \le 1, \tag{5.16}$$

$$ {}^{(n)}\Big(\big[\gamma_0\big(1+\widehat{w}^{\bullet,n}_{p'}(0,T,\cdot)^{1/p'}\big)\big]^{\widehat{N}^{\bullet,n}([0,T],\cdot,1/(4L_0))}\Big)_{32} \le \eta_0. \tag{5.17}$$

We now check that the pair $(X'(\omega), Y'(\cdot)) = (\overline{X}^i(\omega), \overline{X}^\bullet(\omega))$ in (5.14) also satisfies the assumptions of Proposition 3.2. In fact, using the Hölder regularity of the paths, see (4.5) for a similar use, and using the additional $t-s$ in the definition (5.4), $\widehat{w}^{i,n}_{p'}$ dominates (up to a multiplicative constant) the control $\overline{w}^i$ associated to $\overline{\boldsymbol{W}}^i$ (and p') through (2.10) (see (5.1) for the definition of $\overline{\boldsymbol{W}}^i$; in short, the variations of $W^i(\omega)$ and $\mathbb{W}^i(\omega)$ are already included in $v^{i,n}_{p'}(\omega)$, the variations of $\mathbb{W}^{i,\perp}(\omega,\cdot)$ and $\mathbb{W}^{i,\perp}(\cdot,\omega)$ are precisely included in the definition of $\widehat{v}^{i,n}_{p'}(\omega)$ and, using the Hölder regularity of the paths, the variations of $W^i(\cdot)$ and $\mathbb{W}^{i,\perp}(\cdot,\cdot)$ (in $\mathbb{L}^q$) are dominated by the additional $t-s$). Moreover, we have (observe that, since we work here with a copy of the theoretical rough set-up, we use theoretical instead of empirical moments to check the various properties of a rough set-up)

$$\big\langle \widehat{w}^{i,n}_{p'}(s,t,\cdot)\big\rangle_q \le C(t-s) \le C\,\widehat{w}^{i,n}_{p'}(s,t,\omega),$$

for a constant C independent of i, n, s and t. Although $C \ge 2$ (compare with (2.8)), this permits to use $\widehat{w}^{i,n}_{p'}(s,t,\cdot)$ as control functional when working with the rough set-up $\overline{\boldsymbol{W}}^i$ and, in particular, when invoking the solvability Theorem 3.3—the proof would be the same. This is an important point: the path $\overline{X}^i(\omega)$, defined right after (5.1), is the solution of a mean-field rough equation driven by a signal that is controlled by $\widehat{w}^{i,n}_{p'}(\cdot)$. Hence, $X'(\omega) = \overline{X}^i(\omega)$ in (5.14) satisfies the second bound in (3.3) with $w = \widehat{w}^{i,n}_{p'}$, provided (5.16) and (5.17) hold true but under the theoretical (instead of empirical) moments of order 8 and 32, respectively and T therein is deterministic. Of course, by exchangeability, the latter is in fact true automatically as soon as (5.16) and (5.17) themselves are satisfied (it suffices to take power 8 in (5.16) and power 32 in (5.17) and then to take the theoretical expectation). By the same argument, $\overline{X}^i(\cdot)$ satisfies the second condition in (3.2) with respect to $\langle\cdot\rangle_8$; in turn, this, together with (5.16), make it possible to apply the first line in [5], Proposition 4.2, (4.4), for each $i \in \{1,\dots,n\}$ and to deduce that $Y'(\cdot) = \overline{X}^\bullet(\omega)$ satisfies the second condition (3.2) but with respect to ${}^{(n)}(\cdot)_8$. Due to the assumption in [5], Proposition 4.2, this may require to work with a larger value of the threshold L_0 in the statement of Proposition 3.2, but, as made clear in the statement of [5], Proposition 4.2, itself, this is always possible. Then, by Proposition 3.2, we obtain, for a given $L \ge L_0$,

$$\begin{aligned}
&\Big\|\int_{t_k}^{\cdot} \mathrm{F}\big(X^i_r(\omega),\mu^n(\omega)\big)\,d\boldsymbol{W}^i_r(\omega) - \int_{t_k}^{\cdot} \mathrm{F}\big(\overline{X}^i_r(\omega),\overline{\mu}^n(\omega)\big)\,d\boldsymbol{W}^i_r(\omega)\Big\|_{[t_k,t_{k+1}],\widehat{w}^{i,n},p'} \\
&\quad\le \gamma\,\widehat{w}^{i,n}_{p'}(0,t_k,\omega)^{1/p'}\big(\big\|(X^i-\overline{X}^i)(\omega)\big\|_{[0,t_k],\widehat{w}^{i,n},p'}
\end{aligned}$$

$$
\begin{aligned}
&+ {}^{(n)}\big(\|\!|(X^\bullet - \overline{X}^\bullet)(\omega)\|\!|_{[0,T],\widehat{w}^{\bullet,n},p'}\big)_8\big)\\
&+ \frac{\gamma}{4L}\big(\|\!|(X^i - \overline{X}^i)(\omega)\|\!|_{[t_k,t_{k+1}],\widehat{w}^{i,n},p'}\\
&+ {}^{(n)}\big(\|\!|(X^\bullet - \overline{X}^\bullet)(\omega)\|\!|_{[t_k,t_{k+1}],\widehat{w}^{\bullet,n},p'}\big)_8\big),
\end{aligned}
$$

where $\widehat{w}^{i,n}_{p'}(t_k,t_{k+1},\omega)^{1/p'} \le 1/(4L)$ and for $k \le 2\widehat{N}^{i,n}([0,T],\omega,1/(4L))$ (since the sequence $(t_i)_i$ must refine the sequence $(t^0_j)_j$, we may assume that the collection $(t_i)_i$ counts $2\widehat{N}^{i,n}([0,T],\omega,1/(4L))+2$ points, including $t_0=0$) and where γ depends on L_0 and Λ. The point now is to insert the conclusion of the first step (replacing for free $w^{i,n}_{p'}$ by $\widehat{w}^{i,n}_{p'}$ therein). We get

$$
\begin{aligned}
\|\!|(X^i - \overline{X}^i)(\omega)\|\!|_{[t_k,t_{k+1}],\widehat{w}^{i,n},p'} \le{}& \gamma\,\widehat{w}^{i,n}_{p'}(0,t_k,\omega)^{1/p'}\big(\|\!|(X^i - \overline{X}^i)(\omega)\|\!|_{[0,t_k],\widehat{w}^{i,n},p'}\\
&+ {}^{(n)}\big(\|\!|(X^\bullet - \overline{X}^\bullet)(\omega)\|\!|_{[0,T],\widehat{w}^{\bullet,n},p'}\big)_8\big) + \theta^{i,n}(\omega)\\
&+ \frac{\gamma}{4L}\big(\|\!|(X^i - \overline{X}^i)(\omega)\|\!|_{[t_k,t_{k+1}],\widehat{w}^{i,n},p'}\\
&+ {}^{(n)}\big(\|\!|(X^\bullet - \overline{X}^\bullet)(\omega)\|\!|_{[t_k,t_{k+1}],\widehat{w}^{\bullet,n},p'}\big)_8\big).
\end{aligned}
$$

If $\gamma/(4L) \le 1/2$, we get

$$
\begin{aligned}
&\|\!|(X^i - \overline{X}^i)(\omega)\|\!|_{[t_k,t_{k+1}],\widehat{w}^{i,n},p'}\\
&\quad\le 2\gamma\left(\frac{1}{L} + \widehat{w}^{i,n}_{p'}(0,t_k,\omega)^{1/p'}\right)\big(\|\!|(X^i - \overline{X}^i)(\omega)\|\!|_{[0,t_k],\widehat{w}^{i,n},p'}\\
&\qquad + {}^{(n)}\big(\|\!|(X^\bullet - \overline{X}^\bullet)(\omega)\|\!|_{[0,T],\widehat{w}^{\bullet,n},p'}\big)_8\big) + 2\theta^{i,n}(\omega).
\end{aligned}
\tag{5.18}
$$

We deduce that there exists a constant $c \ge 1$, possibly depending on L_0 but independent of n and L and T and whose value may increase from line to line, such that (see, for instance, [5], footnote 5, for the concatenation of two intervals)

$$
\begin{aligned}
&\|\!|(X^i - \overline{X}^i)(\omega)\|\!|_{[0,t_{k+1}],\widehat{w}^{i,n},p'}\\
&\quad\le c\big(\|\!|(X^i - \overline{X}^i)(\omega)\|\!|_{[0,t_k],\widehat{w}^{i,n},p'} + \|\!|(X^i - \overline{X}^i)(\omega)\|\!|_{[t_k,t_{k+1}],\widehat{w}^{i,n},p'}\big)\\
&\quad\le c\big(1+\zeta^{i,n}_T(\omega)\big)\|\!|(X^i - \overline{X}^i)(\omega)\|\!|_{[0,t_k],\widehat{w}^{i,n},p'}\\
&\qquad + c\zeta^{i,n}_T(\omega)^{(n)}\big(\|\!|(X^\bullet - \overline{X}^\bullet)(\omega)\|\!|_{[0,T],\widehat{w}^{\bullet,n},p'}\big)_8 + c\theta^{i,n}(\omega),
\end{aligned}
$$

with $\zeta^{i,n}_T(\omega) := \frac{1}{L} + \widehat{w}^{i,n}_{p'}(0,T,\omega)^{1/p'}$. So, by induction,

$$
\begin{aligned}
&\|\!|(X^i - \overline{X}^i)(\omega)\|\!|_{[0,t_{k+1}],\widehat{w}^{i,n},p'}\\
&\quad\le c\left(\sum_{\ell=0}^{k}\big[c\big(1+\zeta^{i,n}_T(\omega)\big)\big]^\ell\right)\times\big(\zeta^{i,n}_T(\omega)^{(n)}\big(\|\!|(X^\bullet - \overline{X}^\bullet)(\omega)\|\!|_{[0,T],\widehat{w}^{\bullet,n},p'}\big)_8 + \theta^{i,n}(\omega)\big).
\end{aligned}
$$

In the end,

$$
\begin{aligned}
&\|\!|(X^i - \overline{X}^i)(\omega)\|\!|_{[0,T],\widehat{w}^{i,n},p'}\\
&\quad\le c\big[c\big(1+\zeta^{i,n}_T(\omega)\big)\big]^{2\widehat{N}^{i,n}([0,T],\omega,1/(4L))+1}\\
&\qquad\times\big(\zeta^{i,n}_T(\omega)^{(n)}\big(\|\!|(X^\bullet - \overline{X}^\bullet)(\omega)\|\!|_{[0,T],\widehat{w}^{\bullet,n},p'}\big)_8 + \theta^{i,n}(\omega)\big).
\end{aligned}
\tag{5.19}
$$

Hence, using the shorten notation $\widehat{N}_T^{i,n}(\omega)$ for $\widehat{N}^{i,n}([0,T],\omega,1/(4L))$ and recalling that $c \geq 1$, we obtain

$$\begin{aligned}
&{}^{(n)}\big(\|\!|\!|(X^\bullet - \overline{X}^\bullet)(\omega)\|\!|\!|_{[0,T],\widehat{w}^{\bullet,n},p'}\big)_8 \\
&\leq {}^{(n)}\big(\big[c^2\big(1+\zeta_T^{\bullet,n}(\omega)\big)\big]^{2\widehat{N}_T^{\bullet,n}(\omega)+1}\zeta_T^{\bullet,n}(\omega)\big)_8 \\
&\quad \times {}^{(n)}\big(\|\!|\!|(X^\bullet - \overline{X}^\bullet)(\omega)\|\!|\!|_{[0,T],\widehat{w}^{\bullet,n},p'}\big)_8 \\
&\quad + {}^{(n)}\big(\big[c^2\big(1+\zeta_T^{\bullet,n}(\omega)\big)\big]^{2\widehat{N}_T^{\bullet,n}(\omega)+1}\theta^{\bullet,n}(\omega)\big)_8.
\end{aligned} \tag{5.20}$$

Step 3. The key quantity of interest in (5.20) is the multiplicative factor in the second line, which we denote by

$$\Psi_T^n(\omega) := {}^{(n)}\big(\big[c^2\big(1+\zeta_T^{\bullet,n}(\omega)\big)\big]^{2\widehat{N}_T^{\bullet,n}(\omega)+1}\zeta_T^{\bullet,n}(\omega)\big)_8.$$

In particular, letting

$$\Theta_T^n(\omega) := {}^{(n)}\big(\big[c^2\big(1+\zeta_T^{\bullet,n}(\omega)\big)\big]^{2\widehat{N}_T^{\bullet,n}(\omega)+1}\theta^{\bullet,n}(\omega)\big)_8,$$

we rewrite (5.20) in the form

$$\begin{aligned}
&{}^{(n)}\big(\|\!|\!|(X^\bullet - \overline{X}^\bullet)(\omega)\|\!|\!|_{[0,T],\widehat{w}^{\bullet,n},p'}\big)_8 \\
&\leq \Psi_T^n(\omega)\,{}^{(n)}\big(\|\!|\!|(X^\bullet - \overline{X}^\bullet)(\omega)\|\!|\!|_{[0,T],\widehat{w}^{\bullet,n},p'}\big)_8 + \Theta_T^n(\omega).
\end{aligned} \tag{5.21}$$

Here comes the key point. The variable ω being frozen, assume that we are given a deterministic time horizon T small enough and a deterministic $L \geq L_0$ large enough, such that $\gamma/(4L) \leq 1/2$, $\Psi_T^n(\omega) \leq 1/2$, and (5.16) and (5.17) hold true. Then,

$${}^{(n)}\big(\|\!|\!|(X^\bullet - \overline{X}^\bullet)(\omega)\|\!|\!|_{[0,T],\widehat{w}^{\bullet,n},p'}\big)_8 \leq 2\Theta_T^n(\omega).$$

The above inequality sounds really close to the desired result, but it is on a small interval $[0,T]$ only. The purpose is thus to iterate it in order to cover any given time interval. The explicit construction of T and L being postponed to Step 5.

Step 4. In order to iterate in a proper way, we change our notation. While we keep the notation T for the deterministic global time horizon given in the statement (without any further requirement that $T \leq 1$), we use the letter τ instead of T in the previous analysis. Put differently, τ will stand for a deterministic time horizon less than 1 such that Ψ_τ is small enough. And then, we let $\tau_0 = \tau$ and consider a deterministic dissection $0 = \tau_0 < \tau_1 < \cdots < \tau_M = T$ of the interval $[0,T]$ into M subintervals. *The goal of this step is to explain how the error on the interval* $[\tau_\ell, \tau_{\ell+1}]$, $\ell = 1,\ldots,M-1$ *can be controlled in terms of the error on the preceding interval* $[0,\tau_\ell]$*, the construction of the dissection being achieved in Step 5.*

To do so, we need to revisit the statement of Proposition 3.2. Assume indeed that we have a bound for

$$\mathcal{E}_{\tau_\ell}^{i,n}(\omega) := \big(1+\widehat{w}_{p'}^{i,n}(0,T,\omega)^{1/p'}\big)\|\!|\!|(X^i - \overline{X}^i)(\omega)\|\!|\!|_{[0,\tau_\ell],\widehat{w}^{i,n},p'},$$

for some $\ell \leq M$. Then, in order to duplicate the previous two steps, we must consider a new dissection $\tau_\ell = t_0 < t_1 < \cdots < t_K = \tau_{\ell+1}$ of the interval $[\tau_\ell,\tau_{\ell+1}]$ with the property that $K = 2\widehat{N}^{i,n}([\tau_\ell,\tau_{\ell+1}],\omega,1/(4L)) + 1$ and that $\widehat{w}_{p'}^{i,n}(t_k,t_{k+1},\omega) \leq 1/(4L)$ if $k < K$. The key point is to apply the first inequality in (3.4) on $[t_k,t_{k+1}]$ with $(X(\omega),Y(\cdot)) = (X^i(\omega),X^i(\cdot))$

and $(X'(\omega), Y'(\cdot)) = (\overline{X}^i(\omega), \overline{X}^i(\cdot))$, but with τ_ℓ instead of 0 as initial time. Upper bounding the second line in (3.4) by $\mathcal{E}^{i,n}_{\tau_\ell}(\omega) + {}^{(n)}(\mathcal{E}^{\bullet,n}_{\tau_\ell}(\omega))_8$, we obtain

$$\begin{aligned}
&\Big\|\int_{t_k}^{\cdot} \mathrm{F}\big(X^i_r(\omega), \mu^n_r(\omega)\big)\, d\boldsymbol{W}^i_r(\omega) - \int_{t_k}^{\cdot} \mathrm{F}\big(\overline{X}^i_r(\omega), \overline{\mu}^n_r(\omega)\big)\, d\boldsymbol{W}^i_r(\omega)\Big\|_{[t_k,t_{k+1}],\widehat{w}^{i,n},p'} \\
&\quad \le \gamma\, \widehat{w}^{i,n}_{p'}(\tau_\ell,\tau_{\ell+1},\omega)^{1/p'}\big\{\big\|(X^i-\overline{X}^i)(\omega)\big\|_{[\tau_\ell,t_k],\widehat{w}^{i,n},p'} \\
&\qquad + {}^{(n)}\big(\big\|(X^\bullet-\overline{X}^\bullet)(\omega)\big\|_{[\tau_\ell,\tau_{\ell+1}],\widehat{w}^{\bullet,n},p'}\big)_8\big\} \\
&\qquad + \frac{\gamma}{4L}\big\{\big\|(X^i-\overline{X}^i)(\omega)\big\|_{[t_k,t_{k+1}],\widehat{w}^{i,n},p'} \\
&\qquad + {}^{(n)}\big(\big\|(X^\bullet-\overline{X}^\bullet)(\omega)\big\|_{[t_k,t_{k+1}],\widehat{w}^{\bullet,n},p'}\big)_8\big\} \\
&\qquad + \gamma\big[\mathcal{E}^{i,n}_{\tau_\ell}(\omega) + {}^{(n)}(\mathcal{E}^{\bullet,n}_{\tau_\ell}(\omega))_8\big],
\end{aligned}$$

provided the analogues of (3.2) and (3.3) hold true. We may argue as in the second step to check the latter two: They are consequences of Theorem 3.3, if $\tau_{\ell+1}-\tau_\ell \le 1$ and the analogues of (5.16) and (5.17) hold true, namely

$$ {}^{(n)}\big(\widehat{N}^{\bullet,n}([\tau_\ell,\tau_{\ell+1}],\omega,1/(4L_0))\big)_8 \le 1, \tag{5.22}$$

$$ {}^{(n)}\big(\big[\gamma_0\big(1+\widehat{w}^{\bullet,n}_{p'}(\tau_\ell,\tau_{\ell+1},\cdot)^{1/p'}\big)\big]^{\widehat{N}^{\bullet,n}([\tau_\ell,\tau_{\ell+1}],\cdot,1/(4L_0))}\big)_{32} \le \eta_0. \tag{5.23}$$

Then, proceeding as in the second step,

$$\begin{aligned}
&\big\|(X^i-\overline{X}^i)(\omega)\big\|_{[t_k,t_{k+1}],\widehat{w}^{i,n},p'} \\
&\quad \le c\bigg(\frac{1}{L} + \widehat{w}^{i,n}_{p'}(\tau_\ell,\tau_{\ell+1},\omega)^{1/p'}\bigg)\big\{\big\|(X^i-\overline{X}^i)(\omega)\big\|_{[\tau_\ell,t_k],\widehat{w}^{i,n},p'} \\
&\qquad + {}^{(n)}\big(\big\|(X^\bullet-\overline{X}^\bullet)(\omega)\big\|_{[\tau_\ell,\tau_{\ell+1}],\widehat{w}^{\bullet,n},p'}\big)_8\big\} \\
&\qquad + c\big\{\mathcal{E}^{i,n}_{\tau_\ell}(\omega) + {}^{(n)}(\mathcal{E}^{\bullet,n}_{\tau_\ell}(\omega))_8 + \theta^{i,n}(\omega)\big\}.
\end{aligned}$$

In the end, we are in the same situation as in Step 2, but with new $\zeta^{i,n}_T$ and $\widehat{N}^{i,n}_T$. Here, we let (pay attention that, to be consistent with the notations $\zeta^{i,n}_T$ and $\widehat{N}^{i,n}_T$, we should use $[\tau_\ell,\tau_{\ell+1}]$ instead of ℓ as subscript below, but, for simplicity, we prefer to use ℓ only)

$$\zeta^{i,n}_\ell(\omega) := \frac{1}{L} + \widehat{w}^{i,n}_{p'}(\tau_\ell,\tau_{\ell+1},\omega)^{1/p'}, \qquad \widehat{N}^{i,n}_\ell(\omega) := \widehat{N}^{i,n}\bigg([\tau_\ell,\tau_{\ell+1}],\omega,\frac{1}{4L}\bigg).$$

Following (5.19), we obtain

$$\begin{aligned}
&\big\|(X^i-\overline{X}^i)(\omega)\big\|_{[\tau_\ell,\tau_{\ell+1}],\widehat{w}^{i,n},p'} \\
&\quad \le c\big[c\big(1+\zeta^{i,n}_\ell(\omega)\big)\big]^{2\widehat{N}^{i,n}_\ell(\omega)+1} \\
&\qquad \times \big\{\zeta^{i,n}_\ell(\omega)\,{}^{(n)}\big(\big\|(X^\bullet-\overline{X}^\bullet)(\omega)\big\|_{[\tau_\ell,\tau_{\ell+1}],\widehat{w}^{\bullet,n},p'}\big)_8 \\
&\qquad + \theta^{i,n}(\omega) + \mathcal{E}^{i,n}_{\tau_\ell} + {}^{(n)}\big(\mathcal{E}^{\bullet,n}_{\tau_\ell}(\omega)\big)_8\big\}.
\end{aligned} \tag{5.24}$$

Hence,

$$\begin{aligned}
&{}^{(n)}\big(\big\|(X^\bullet-\overline{X}^\bullet)(\omega)\big\|_{[\tau_\ell,\tau_{\ell+1}],\widehat{w}^{\bullet,n},p'}\big)_8 \\
&\quad \le \Psi^n_\ell(\omega) \times {}^{(n)}\big(\big\|(X^\bullet-\overline{X}^\bullet)(\omega)\big\|_{[\tau_\ell,\tau_{\ell+1}],\widehat{w}^{\bullet,n},p'}\big)_8 + \Theta^n_\ell(\omega),
\end{aligned}$$

with

$$\Psi_\ell^n(\omega) := {}^{(n)}\big(\big[c^2\big(1+\zeta_\ell^{\bullet,n}(\omega)\big)\big]^{2\widehat{N}_\ell^{\bullet,n}(\omega)+1}\zeta_\ell^{\bullet,n}(\omega)\big)_8,$$

$$\Theta_\ell^n(\omega) := {}^{(n)}\big(\big[c^2\big(1+\zeta_\ell^{\bullet,n}(\omega)\big)\big]^{2\widehat{N}_\ell^{\bullet,n}(\omega)+1}\big(\theta^{\bullet,n}(\omega)+\mathcal{E}_{\tau_\ell}^{\bullet,n}(\omega)+{}^{(n)}(\mathcal{E}_{\tau_\ell}^{\bullet,n}(\omega))_8\big)\big)_8.$$

Following (5.21), if we can choose L large enough and then $\tau_{\ell+1}-\tau_\ell$ small enough such that $\Psi_\ell^n(\omega)\le 1/2$, then we get

$${}^{(n)}\big(\|\!|(X^\bullet-\overline{X}^\bullet)(\omega)|\!\|_{[\tau_\ell,\tau_{\ell+1}],\widehat{w}^{\bullet,n},p'}\big)_8 \le 2\Theta_\ell^n(\omega).$$

Eventually, returning to (5.24) and modifying the value of the constant c, we deduce

$$\begin{aligned}&\|\!|(X^i-\overline{X}^i)(\omega)|\!\|_{[\tau_\ell,\tau_{\ell+1}],\widehat{w}^{i,n},p'}\\&\quad\le c\big[c\big(1+\zeta_\ell^{i,n}(\omega)\big)\big]^{2\widehat{N}_\ell^{i,n}(\omega)+1}\\&\qquad\times\big(\zeta_\ell^{i,n}(\omega)\Theta_\ell^n(\omega)+\theta^{i,n}(\omega)+\mathcal{E}_{\tau_\ell}^{i,n}+{}^{(n)}\big(\mathcal{E}_{\tau_\ell}^{\bullet,n}(\omega)\big)_8\big),\end{aligned}$$

and then

$$\mathcal{E}_{\tau_{\ell+1}}^{i,n}(\omega)\le\kappa_\ell^{i,n}(\omega)\big(\zeta_\ell^{i,n}(\omega)\Theta_\ell^n(\omega)+\theta^{i,n}(\omega)+\mathcal{E}_{\tau_\ell}^{i,n}(\omega)+{}^{(n)}\big(\mathcal{E}_{\tau_\ell}^{\bullet,n}(\omega)\big)_8\big),$$

with

$$\kappa_\ell^{i,n}(\omega) := c^2\big(1+\widehat{w}_{p'}^{i,n}(0,T,\omega)^{1/p'}\big)\big[c^2\big(1+\zeta_\ell^{i,n}(\omega)\big)\big]^{2\widehat{N}_\ell^{i,n}(\omega)+1}, \tag{5.25}$$

using the fact that $c\ge 1$. By induction, we get the following global bound:

$$\mathcal{E}_{\tau_{\ell+1}}^{i,n}(\omega)\le\sum_{k=0}^{\ell}\mathcal{K}_{k,\ell}^{i,n}(\omega)\big[\zeta_k^{i,n}(\omega)\Theta_k^n(\omega)+\theta^{i,n}(\omega)+{}^{(n)}\big(\mathcal{E}_{\tau_k}^{\bullet,n}(\omega)\big)_8\big], \tag{5.26}$$

with

$$\mathcal{K}_{k,\ell}^{i,n}(\omega) := \prod_{j=k}^{\ell}\kappa_j^{i,n}(\omega). \tag{5.27}$$

We deduce that for any $r>8$, we can find a constant $q(r)$ (which has nothing to do with q in the assumption) such that

$$\begin{aligned}{}^{(n)}\big(\mathcal{E}_{\tau_{\ell+1}}^{\bullet,n}(\omega)\big)_r \le \sum_{k=0}^{\ell}\big\{&{}^{(n)}\big(\mathcal{K}_{k,\ell}^{\bullet,n}\big)_{q(r)}\times\big(1+{}^{(n)}\big(\widehat{w}^{\bullet,n}(0,T,\omega)^{1/p'}\big)_{q(r)}\big)\\&\times\big(1+{}^{(n)}\big(\big[c^2\big(1+\zeta_k^{\bullet,n}(\omega)\big)\big]^{2\widehat{N}_k^{\bullet,n}(\omega)+1}\big)_{q(r)}\big)\\&\times\big({}^{(n)}\big(\theta^{\bullet,n}(\omega)\big)_{q(r)}+{}^{(n)}\big(\mathcal{E}_{\tau_k}^{\bullet,n}(\omega)\big)_r\big)\big\}.\end{aligned}$$

Using the fact that

$${}^{(n)}\big(\mathcal{K}_{k,\ell}^{\bullet,n}\big)_{q(r)}\ge\max\big(1,{}^{(n)}\big(\widehat{w}^{\bullet,n}(0,T,\omega)^{1/p'}\big)_{q(r)},{}^{(n)}\big(\big[c^2\big(1+\zeta_k^{\bullet,n}(\omega)\big)\big]^{2\widehat{N}_k^{\bullet,n}(\omega)+1}\big)_{q(r)}\big),$$

we obtain a bound of the form $a_{\ell+1}\le\sum_{k=0}^{\ell}g_{k,\ell}(b+a_k)$, with

$$a_\ell := {}^{(n)}\big(\mathcal{E}_{\tau_\ell}^{\bullet,n}(\omega)\big)_r,\qquad g_{k,\ell} := 4\times\big({}^{(n)}\big(\mathcal{K}_{k,\ell}^{\bullet,n}\big)_{q(r)}\big)^3,\qquad b := {}^{(n)}\big(\theta^{\bullet,n}(\omega)\big)_{q(r)}.$$

Hence,

$$a_\ell\le b\sum_{j=1}^{\ell}\sum_{0\le k_1\le\cdots\le k_j\le k_{j+1}=\ell}\prod_{h=1}^{j}g_{k_h,k_{h+1}}. \tag{5.28}$$

Back to (5.27), we will use below the bound

$$\begin{aligned}
&\mathcal{K}^{i,n}_{k,\ell}(\omega)\\
&\le c^{2(\ell+1-k)}\prod_{j=k}^{\ell}\{(1+\widehat{w}^{i,n}(0,T,\omega)^{1/p'})[c^2(1+\zeta^{i,n}_j(\omega))]^{2\widehat{N}^{i,n}_j(\omega)+1}\}\\
&\le c^{4(\ell+1-k)+4\widehat{N}^{i,n}_{k,\ell+1}(\omega)}(1+\widehat{w}^{i,n}(0,T,\omega)^{1/p'})^{\ell+1-k+2\widehat{N}^{i,n}_{k,\ell+1}(\omega)},
\end{aligned}\tag{5.29}$$

with the shortened notation $\widehat{N}^{i,n}_{k,\ell}(\omega) := \widehat{N}^{i,n}([\tau_k,\tau_\ell],\omega,1/(4L))$, where we have used the fact that $\sum_{j=k}^{\ell}\widehat{N}^{i,n}_j(\omega)\le\widehat{N}^{i,n}_{k,\ell+1}(\omega)$.

Step 5. We now recall that the parameter L and the sequence of deterministic times $0=\tau_0<\tau_1<\cdots<\tau_M$ must satisfy $\gamma/(4L)\le 1/2$, (5.22), (5.23) and $\Psi^n_\ell(\omega)\le 1/2$, together with $\tau_{\ell+1}-\tau_\ell\le 1$, for $\ell\in\{0,\dots,M-1\}$.

In order to proceed, we let

$$\begin{aligned}
M_1 = {}&\langle\|W(\cdot)\|^{p'}_{[0,T],(1/p')-\mathrm{H}}\rangle_q+\langle\|\mathbb{W}(\cdot)\|^{p'/2}_{[0,T],(2/p')-\mathrm{H}}\rangle_q\\
&+\langle\|\mathbb{W}^{\perp\!\!\!\perp}(\cdot,\cdot)\|^{p'/2}_{[0,T],(2/p')-\mathrm{H}}\rangle_q,
\end{aligned}$$

and we consider the events

$$\begin{aligned}
A^n_1 = {}&\{\omega\in\Omega:{}^{(n)}(\|W^\bullet(\omega)\|^{p'}_{[0,T],(1/p')-\mathrm{H}})_q+{}^{(n)}(\|\mathbb{W}^\bullet(\omega)\|^{p'/2}_{[0,T],(2/p')-\mathrm{H}})_q\\
&+{}^{(n)}((\|\mathbb{W}^{\bullet,\bullet}(\omega)\|^{p'/2}_{[0,T],(2/p')-\mathrm{H}}))_q\le M_1+1\},\\
A^n_2 = {}&\{\omega\in\Omega:\forall i\in\{1,\dots,n\},\forall(s,t)\in\mathcal{S}^T_2,\\
&{}^{(n)}(\mathbb{W}^{i,\bullet}(\omega))_{q;[s,t],p'/2-\mathrm{v}}+{}^{(n)}(\mathbb{W}^{\bullet,i}(\omega))_{q;[s,t],p'/2-\mathrm{v}}\\
&\le 2c_{p,p',q}(t-s)^{2/p'}+\langle\mathbb{W}^{i,\perp\!\!\!\perp}(\omega,\cdot)\rangle_{q;[s,t],p'/2-\mathrm{v}}+\langle\mathbb{W}^{i,\perp\!\!\!\perp}(\cdot,\omega)\rangle_{q;[s,t],p'/2-\mathrm{v}}\},
\end{aligned}$$

with $c_{p,p',q}$, as in (4.23). On the event $A^n_1\cap A^n_2$, we have (compare with (4.3) and (4.5)) for $s,t\in[0,T]^2$, $s<t$,

$$\begin{aligned}
v^{i,n}_{p'}(s,t,\omega)\le{}&\|W^i(\omega)\|^{p'}_{[s,t],p'-\mathrm{v}}+\|\mathbb{W}^i(\omega)\|^{p'/2}_{[s,t],p'/2-\mathrm{v}}\\
&+3^{p'/2-1}(\langle\mathbb{W}^{i,\perp\!\!\!\perp}(\omega,\cdot)\rangle^{p'/2}_{q;[s,t],p'/2-\mathrm{v}}+\langle\mathbb{W}^{i,\perp\!\!\!\perp}(\cdot,\omega)\rangle^{p'/2}_{q;[s,t],p'/2-\mathrm{v}})\\
&+(c'_{p,p',q}+M_1)(t-s),
\end{aligned}$$

for a new constant $c'_{p,p',q}$ only depending on p, p' and q. Therefore, introducing the new event

$$\begin{aligned}
A^n_3 = {}&\{{}^{(n)}(\langle\|\mathbb{W}^{\bullet,\perp\!\!\!\perp}(\omega,\cdot)\|^{p'/2}_{[0,T],(2/p')-\mathrm{H}}\rangle_q)_q\\
&+{}^{(n)}(\langle\|\mathbb{W}^{\bullet,\perp\!\!\!\perp}(\cdot,\omega)\|^{p'/2}_{[0,T],(2/p')-\mathrm{H}}\rangle_q)_q\le 2(1+M_1)\},
\end{aligned}$$

we get, on $A^n_1\cap A^n_2\cap A^n_3$,

$${}^{(n)}(v^{\bullet,n}_{p'}(\omega))_{q;[s,t],1-\mathrm{v}}\le c'_{p,p',q,M_1}(t-s),$$

for a new constant c'_{p,p',q,M_1} only depending on p, p', q and M_1.

Recall now the definition of $\widehat{v}^{i,n}_{p'}$ in (5.4). We have

$$\widehat{v}^{i,n}_{p'}(s,t,\omega) = \langle \mathbb{W}^{i,\perp}(\omega,\cdot)\rangle^{p'/2}_{q;[s,t],p'/2-\mathrm{v}} + \langle \mathbb{W}^{i,\perp}(\cdot,\omega)\rangle^{p'/2}_{q;[s,t],p'/2-\mathrm{v}}$$
$$\le \big[\big\langle \|\mathbb{W}^{i,\perp}(\omega,\cdot)\|^{p'/2}_{[0,T],(2/p')-\mathrm{H}}\big\rangle_q + \big\langle \|\mathbb{W}^{i,\perp}(\cdot,\omega)\|^{p'/2}_{[0,T],(2/p')-\mathrm{H}}\big\rangle_q\big](t-s).$$

And then,

$$^{(n)}\big(\widehat{v}^{\bullet,n}_{p'}(\omega)\big)_{q;[s,t],1-\mathrm{v}} \le \big({}^{(n)}\big(\big\langle \|\mathbb{W}^{\bullet,\perp}(\omega,\cdot)\|^{p'/2}_{[0,T],(2/p')-\mathrm{H}}\big\rangle_q\big)_q$$
$$+ {}^{(n)}\big(\big\langle \|\mathbb{W}^{\bullet,\perp}(\cdot,\omega)\|^{p'/2}_{[0,T],(2/p')-\mathrm{H}}\big\rangle_q\big)_q\big)(t-s).$$

Therefore, on the event $A^n_1\cap A^n_2\cap A^n_3$, we have

$$^{(n)}\big(\widehat{v}^{\bullet,n}_{p'}(\omega)\big)_{q;[s,t],1-\mathrm{v}} \le 2(1+M_1)(t-s).$$

Using the same notations as in (5.4), we end-up with

$$\widehat{w}^{i,n}_{p'}(s,t,\omega) \le \widetilde{w}^{i,n}_{p'}(s,t,\omega), \tag{5.30}$$

for $\omega\in A^n_1\cap A^n_2\cap A^n_3$, where we let (using the fact that $3^{p'/2-1}\le\sqrt{3}\le 2$)

$$\widetilde{w}^{i,n}_{p'}(s,t,\omega) := \|W^i(\omega)\|^{p'}_{[s,t],p'-\mathrm{v}} + \|\mathbb{W}^i(\omega)\|^{p'/2}_{[s,t],p'/2-\mathrm{v}} + C_{p,p',q,M_1}(t-s)$$
$$+ 3\langle \mathbb{W}^{i,\perp}(\omega,\cdot)\rangle^{p'/2}_{q;[s,t],p'/2-\mathrm{v}} + 3\langle \mathbb{W}^{i,\perp}(\cdot,\omega)\rangle^{p'/2}_{q;[s,t],p'/2-\mathrm{v}},$$

with $C_{p,p',q,M_1} := c'_{p,p',q,M_1} + c'_{p,p',q} + 3M_1 + 3$. Using the notation (2.17), we also let $\widetilde{N}^{i,n}([\tau,\tau'],\omega,\alpha) := N_{\varpi}([\tau,\tau'],\alpha)$, with $\varpi := (\widetilde{w}^{i,n}_{p'}(\omega))^{1/p'}$. By (5.30),

$$\widehat{N}^{i,n}\big([\tau,\tau'],\omega,\alpha\big) \le \widetilde{N}^{i,n}\big([\tau,\tau'],\omega,\alpha\big), \tag{5.31}$$

for $\omega\in A^n_1\cap A^n_2\cap A^n_3$. The good point here is that the variables $(\widetilde{w}^{i,n}_{p'})_{1\le i\le n}$ are independent whilst the variables $(\widehat{w}^{i,n}_{p'})_{1\le i\le n}$ are not. Similarly, whenever τ and τ' are deterministic, the variables $(\widetilde{N}^{i,n}([\tau,\tau'],\cdot,\alpha))_{1\le i\le n}$ are independent. Moreover, it is not difficult to see that

$$\widetilde{w}^{i,n}_{p'}(s,t,\omega) \le 3\widehat{w}^{i,n}_{p'}(s,t,\omega) + C_{p,p',q,M_1}(t-s), \tag{5.32}$$

from which we deduce, see for instance (A.1), that, for any $\alpha>0$,

$$\widetilde{N}^{i,n}\big([\tau,\tau'],\omega,\alpha\big) \le \widehat{N}^{i,n}\Big([\tau,\tau'],\omega,\frac{\alpha}{2\cdot 3^{1/p'}}\Big) + C_\alpha(\tau'-\tau), \tag{5.33}$$

for a constant C_α only depending on α and on p, p', q and M_1 (we feel easier not to indicate the dependence on p, p', q and M_1). In particular, we can easily replace $\widehat{N}^{i,n}$ by $\widetilde{N}^{i,n}$ in the third item of the assumption of the statement. Moreover, by (4.11), we deduce that each $\widetilde{w}^{i,n}_{p'}$ satisfies the first bound in (3.5), uniformly in i and n.

We now claim that we can choose L and the sequence $0=\tau_0<\cdots<\tau_M=T$, independently of n, such that

$$\big\langle \widetilde{N}^{1,n}\big([\tau_\ell,\tau_{\ell+1}],\cdot,1/(4L_0)\big)\big\rangle_8 \le \frac{1}{2}, \tag{5.34}$$

$$\big\langle \big[\gamma_0\big(1+\widetilde{w}^{1,n}_{p'}(0,T,\cdot)^{1/p'}\big)\big]^{\widetilde{N}^{1,n}([\tau_\ell,\tau_{\ell+1}],\cdot,1/(4L_0))}\big\rangle_{32} \le \frac{1+\eta_0}{2}, \tag{5.35}$$

$$\Big\langle \big[c^2\big(1+\widetilde{w}^{1,n}_{p'}(0,T,\cdot)^{1/p'}\big)\big]^{2\widetilde{N}^{1,n}([\tau_\ell,\tau_{\ell+1}],\cdot,1/(4L_0))+1}$$
$$\times\Big(\frac{1}{L}+\widetilde{w}^{1,n}_{p'}(\tau_\ell,\tau_{\ell+1},\cdot)^{1/p'}\Big)\Big\rangle_8 \le \frac{1}{4}, \tag{5.36}$$

for γ_0, η_0 and c as in Steps 2 and 4 (in particular, c is independent of n and L). We then make use of the third item in the assumption of the statement, see (5.3); basically, it says that, in all the three constraints, we should normalize $\widetilde{N}^{1,n}([\tau_\ell,\tau_{\ell+1}],\cdot,1/(4L_0))$ by the root of $\tau_{\ell+1}-\tau_\ell$. This indeed makes sense since inequality (5.32) insures that $\widehat{N}^{1,n}([\tau_\ell,\tau_{\ell+1}],\cdot,1/(4L_0))$ satisfies a similar estimate as $\widehat{N}^{1,n}([\tau_\ell,\tau_{\ell+1}],\cdot,1/(4L_0))$ in (5.3). As for (5.34), (5.3) says that the left-hand side can be bounded by $C\sqrt{\tau_{\ell+1}-\tau_\ell}$, for a constant C independent of n, which makes it possible to choose $\tau_{\ell+1}-\tau_\ell$ small enough (independently of n) such that the constraint (5.34) is indeed satisfied. Recalling that $\tau_{\ell+1}-\tau_\ell$ is less than 1 and invoking Cauchy-Schwarz inequality, the left-hand side in (5.35) can be bounded by $C^{\sqrt{\tau_{\ell+1}-\tau_\ell}}$; the way the constant C here shows up is made clear in (5.41) below, with δ'_ℓ therein being here understood as $\tau_{\ell+1}-\tau_\ell$ and the function f being lower bounded by the identity. Importantly, the application of (5.41) is made possible by the upper bounds we have for $\widetilde{w}^{1,n}_{p'}$ and $\widetilde{N}^{1,n}$ in terms of $\widehat{w}^{1,n}_{p'}$ and $\widehat{N}^{1,n}$, see (5.32) and (5.33). As for (5.36), it may be bounded, using the same argument together with an additional Cauchy-Schwarz argument, by a product of the form $C^{\sqrt{\tau_{\ell+1}-\tau_\ell}}\times(\frac{1}{L}+\langle\widetilde{w}^{1,n}_{p'}(\tau_\ell,\tau_{\ell+1},\cdot)^{1/p'}\rangle_{16})$. The first factor $C^{\sqrt{\tau_{\ell+1}-\tau_\ell}}$ can be made smaller than 2 by choosing $\tau_{\ell+1}-\tau_\ell$ small enough. Then, we can take $L\geq 16$ so that $C^{\sqrt{\tau_{\ell+1}-\tau_\ell}}/L$ is less than $1/8$. It then remains to decrease $\tau_{\ell+1}-\tau_\ell$ if necessary to render $\langle\widetilde{w}^{1,n}_{p'}(\tau_\ell,\tau_{\ell+1},\cdot)^{1/p'}\rangle_{16}$ less than $1/16$; this is possible by using the analogue of (4.5) but for $\widehat{w}^{i,n}_{p'}$, as it says that $\widetilde{w}^{1,n}_{p'}(\tau_\ell,\tau_{\ell+1},\cdot)^{1/p'}$ scales like $(\tau_{\ell+1}-\tau_\ell)^{1/p'}$, the (random) scaling factor having moments of any order that are bounded independently of n. Importantly, this discussion says that the number M of intervals in the dissection can be chosen independently of n. For sure, the index 1 in the left-hand side in the three constraints (5.34), (5.35) and (5.36) can be replaced by any $i\in\{1,\dots,n\}$. We then consider the family of events

$$
\begin{aligned}
A^{\ell,n}_4 &= A^{\ell,n}_{4,1}\cap A^{\ell,n}_{4,2}\cap A^{\ell,n}_{4,3},\quad \ell=0,\dots,M-1,\\
A^{\ell,n}_{4,1} &= \big\{{}^{(n)}\big(\widetilde{N}^{\bullet,n}\big([\tau_\ell,\tau_{\ell+1}],\cdot,1/(4L_0)\big)\big)_8\leq 1\big\},\\
A^{\ell,n}_{4,2} &= \big\{{}^{(n)}\big([\gamma_0(1+\widetilde{w}^{\bullet,n}_{p'}(0,T,\cdot)^{1/p'})]^{\widetilde{N}^{\bullet,n}([\tau_\ell,\tau_{\ell+1}],\cdot,1/(4L_0))}\big)_{32}\leq 1+\eta_0\big\},\\
A^{\ell,n}_{4,3} &= \Big\{{}^{(n)}\Big([c^2(1+\widetilde{w}^{\bullet,n}_{p'}(0,T,\cdot)^{1/p'})]^{2\widetilde{N}^{\bullet,n}([\tau_\ell,\tau_{\ell+1}],\cdot,1/(4L_0))+1}\\
&\qquad\times\Big(\frac{1}{L}+\widetilde{w}^{\bullet,n}_{p'}(\tau_\ell,\tau_{\ell+1},\cdot)^{1/p'}\Big)\Big)_{32}\leq\frac{1}{2}\Big\}.
\end{aligned}
$$

By (5.30) and (5.31), on $A^n_1\cap A^n_2\cap A^n_3\cap(\bigcap_{\ell=0}^{M-1}A^{\ell,n}_4)$, the upper bounds (5.22), (5.23) and $\Psi^n_\ell\leq 1/2$ are satisfied and then the conclusion of the fourth step holds true.

Following (5.25), this prompts us to set:

$$
\begin{aligned}
\widetilde{\kappa}^{i,n}_\ell(\omega) &:= c^2\big(1+\widetilde{w}^{i,n}_{p'}(0,T,\omega)^{1/p'}\big)\\
&\qquad\times\Big[c^2\Big(1+\frac{1}{L}+\widetilde{w}^{i,n}_{p'}(\tau_\ell,\tau_{\ell+1},\omega)^{1/p'}\Big)\Big]^{2\widetilde{N}^{i,n}_\ell(\omega)+1},
\end{aligned}
$$

and then $\widetilde{\mathcal{K}}^{i,n}_{k,\ell}(\omega):=\prod_{j=k}^{\ell}\widetilde{\kappa}^{i,n}_j(\omega)$. Returning to the conclusion of the fourth step, we get, for $\omega\in A^n:=A^n_1\cap A^n_2\cap A^n_3\cap(\bigcap_{\ell=0}^{M-1}A^{\ell,n}_4)$,

$$
\begin{aligned}
{}^{(n)}\big(\mathcal{E}^{\bullet,n}_{\tau_\ell}(\omega)\big)_r &\leq {}^{(n)}\big(\theta^{\bullet,n}(\omega)\big)_{q(r)}\sum_{j=1}^{\ell}\sum_{0\leq k_1\leq\cdots\leq k_j\leq k_{j+1}=\ell}\prod_{h=1}^{j}4\times\big({}^{(n)}\big(\widetilde{\mathcal{K}}^{\bullet,n}_{k_h,k_{h+1}}(\omega)\big)_{q(r)}\big)^3\\
&\leq \ell 2^{2\ell+1}\times 4^\ell\times{}^{(n)}\big(\theta^{\bullet,n}(\omega)\big)_{q(r)}\times\big({}^{(n)}\big(\widetilde{\mathcal{K}}^{\bullet,n}_{0,\ell}(\omega)\big)_{q(r)}\big)^{3\ell}.
\end{aligned}
$$

The key fact here is that $\widetilde{\mathcal{K}}^{i,n}_{0,M}(\omega)$, for any $i \in \{1,\dots,n\}$, has finite moments of any order, independently of i and n. The proof follows from (5.29), from (4.11) and from the third item in the assumption of the statement of Theorem 5.1, the last two properties implying that $(1+\widetilde{w}^{i,n}_{p'}(\tau_\ell,\tau_{\ell+1},\omega)^{1/p'})^{\widetilde{N}^{i,n}_\ell(\omega)}$ has finite moments of any order, independently of i and n, see for instance (5.43) below. Hence, for a constant C, independent of n but possibly depending on M, we get (using an obvious exchangeability argument)

$$\big\langle \mathbf{1}_{A^n}(\cdot)^{(n)}\big(\mathcal{E}^{\bullet,n}_T(\cdot)\big)_r\big\rangle_r \le C\big\langle {}^{(n)}\big(\theta^{\bullet,n}(\cdot)\big)_{q(r)}\big\rangle_{2r} \le C\big\langle \theta^{1,n}(\cdot)\big\rangle_{q(r)},$$

where we took, without any loss of generality, $q(r)\ge 2r$. Taking $\varrho = q(r)$ in (5.12), we get that, for a constant C independent of n, but depending on r,

$$\sup_{1\le i\le n}\big\langle \mathbf{1}_{A^n}(\cdot)\big|(X^i-\overline{X}^i)(\cdot)\big|\big\rangle_r \le C\varsigma_n. \tag{5.37}$$

Step 6. From the law of large of numbers and from (4.11), we claim that $\mathbb{P}((A^n_1)^{\complement})$ decays faster than any n^{-s}, for $s>0$. The first step of the proof is to notice that

$$\begin{aligned}
\mathbb{P}\big((A^n_1)^{\complement}\big) &\le \mathbb{P}\Big(\omega : {}^{(n)}\big(\|W^\bullet(\omega)\|^{p'}_{[0,T],(1/p')-\mathrm{H}}\big)_q - \big\langle \|W(\cdot)\|^{p'}_{[0,T],(1/p')-\mathrm{H}}\big\rangle_q \ge \frac{1}{3}\Big)\\
&\quad + \mathbb{P}\Big(\omega : {}^{(n)}\big(\|\mathbb{W}^\bullet(\omega)\|^{p'/2}_{[0,T],(2/p')-\mathrm{H}}\big)_q - \big\langle \|\mathbb{W}(\cdot)\|^{p'/2}_{[0,T],(2/p')-\mathrm{H}}\big\rangle_q \ge \frac{1}{3}\Big)\\
&\quad + \mathbb{P}\Big(\omega : {}^{(n)}\big(\big(\|\mathbb{W}^{\bullet,\bullet}(\omega)\|^{p'/2}_{[0,T],(2/p')-\mathrm{H}}\big)\big)_q - \big\langle \|\mathbb{W}^{\perp}(\cdot,\cdot)\|^{p'/2}_{[0,T],(2/p')-\mathrm{H}}\big\rangle_q \ge \frac{1}{3}\Big)\\
&=: \pi^n_{1,1}+\pi^n_{1,2}+\pi^n_{1,3}.
\end{aligned}$$

Since the most difficult term is the last one, we just explain how to handle it. The other two terms may be treated in the same way. Since $q\ge 1$, we first observe that

$$\begin{aligned}
\pi^n_{1,3} &\le \mathbb{P}\Big(\omega : {}^{(n)}\big(\big(\|\mathbb{W}^{\bullet,\bullet}(\omega)\|^{p'/2}_{[0,T],(2/p')-\mathrm{H}}\big)\big)^q_q \ge \big\langle \|\mathbb{W}^{\perp}(\cdot,\cdot)\|^{p'/2}_{[0,T],(2/p')-\mathrm{H}}\big\rangle^q_q + \frac{1}{3^q}\Big)\\
&\le \mathbb{P}\Big(\omega : \frac{1}{n^2}\sum_{i\ne j}\big(\|\mathbb{W}^{i,j}(\omega)\|^{qp'/2}_{[0,T],(2/p')-\mathrm{H}} - \big\langle \|\mathbb{W}^{\perp}(\cdot,\cdot)\|^{qp'/2}_{[0,T],(2/p')-\mathrm{H}}\big\rangle_1\big) \ge \frac{1}{3^{q+1}}\Big)\\
&\quad + \mathbb{P}\Big(\omega : \frac{1}{n^2}\sum_{i=1}^n \|\mathbb{W}^{i}(\omega)\|^{qp'/2}_{[0,T],(2/p')-\mathrm{H}} \ge \frac{1}{3^{q+1}}\Big).
\end{aligned}$$

By (4.11), the last term in the right-hand side is easily handled. As for the first one, Markov's inequality yields, for any $s>1$,

$$\begin{aligned}
&\mathbb{P}\Big(\omega : \frac{1}{n^2}\sum_{j\ne i}\big(\|\mathbb{W}^{i,j}(\omega)\|^{qp'/2}_{[0,T],(2/p')-\mathrm{H}} - \big\langle \|\mathbb{W}^{\perp}(\cdot,\cdot)\|^{qp'/2}_{[0,T],(2/p')-\mathrm{H}}\big\rangle_1\big) \ge \frac{1}{3^{q+1}}\Big)\\
&\le \frac{3^{s(q+1)}}{n^{s+1}}\sum_{i=1}^n \mathbb{E}\Big[\Big|\sum_{j:j\ne i}\big(\|\mathbb{W}^{i,j}(\cdot)\|^{qp'/2}_{[0,T],(2/p')-\mathrm{H}} - \big\langle \|\mathbb{W}^{\perp}(\cdot,\cdot)\|^{qp'/2}_{[0,T],(2/p')-\mathrm{H}}\big\rangle_1\big)\Big|^s\Big].
\end{aligned}$$

By (4.11) again and by Rosenthal's inequality, see [46], we deduce that the right-hand side is less than $Cn^{-s/2}$, for a constant C independent of n. This completes the proof of our claim.

The same result holds for $\mathbb{P}((A^n_3)^{\complement})$. Also, since $(\widetilde{N}^{i,n}([\tau_\ell,\tau_{\ell+1}],\cdot,\alpha))_{1\le i\le n}$, are independent for any $\ell=0,\dots,M-1$ and have finite moments of any order that can be bounded independently of n, we also have that $\mathbb{P}((A^{\ell,n}_{4,1})^{\complement})$ decays faster than any n^{-s}, for any

$\ell = 0, \dots, M-1$. Invoking once again (5.42), the same holds for $\mathbb{P}((A^{\ell,n}_{4,2})^{\complement})$ and $\mathbb{P}((A^{\ell,n}_{4,3})^{\complement})$ and hence for $\mathbb{P}((A^{\ell,n}_{4})^{\complement})$. Since M is finite, $\mathbb{P}((A^{n}_{4})^{\complement})$ also decays faster than any n^{-s}.

We finally check that the same is true for A^n_2. In fact, this is a consequence of (4.23), choosing therein δ first in terms of p, p' and q, and then a in terms of δ such that $a\delta/4 = s$. In the end, C_a in (4.23) depends on p, p', q and s.

All in all, back to the definition of A^n at the end of the fifth step, see (5.37), we deduce that, for any $s > 0$, $\mathbb{P}((A^n)^{\complement}) \le Cn^{-s}$. Therefore, in order to conclude, it suffices to prove that, for any $r \ge 1$, we can choose $q(r) \ge 8$ such that, if $X_0(\cdot)$ is in $\mathbb{L}^{q(r)}$, then

$$\sup_{1\le i\le n} \mathbb{E}\Big[\sup_{0\le t\le T} |X^i_t|^r\Big] \le C(r), \tag{5.38}$$

for a constant $C(r)$ depending on r and on $\langle X_0(\cdot)\rangle_{q(r)}$ but independent of n. Obviously, (5.38) implies the same inequality but for $\overline{X}^i$ by (say for $i = 1$) letting n tend to ∞ and then invoking Theorem 4.3.

The proof of (5.38) relies on the final estimate in the statement of Theorem 3.3. To make it clear, we consider a new random dissection $0 = \tau_0 < \tau_1 < \cdots < \tau_M = T$ of $[0, T]$ (for simplicity, we use the same notation as in the previous step, but the new dissection has in fact nothing to do with the first one; in particular, it is random) such that

$$\begin{aligned} &\Gamma^{(n)}_1(\omega, [\tau_\ell, \tau_{\ell+1}]) := {}^{(n)}\big(\widehat{N}^{\bullet,n}([\tau_\ell, \tau_{\ell+1}], \omega, 1/(4L_0))\big)_8 \le 1, \\ &\Gamma^{(n)}_2(\omega, [\tau_\ell, \tau_{\ell+1}]) \\ &\quad := {}^{(n)}\big([\gamma_0^2(1 + \widehat{w}^{\bullet,n}(0, T, \omega)^{1/p'})]^{\widehat{N}^{\bullet,n}([\tau_\ell,\tau_{\ell+1}],\omega,1/(4L))}\big)_{32} \le \eta_0, \end{aligned} \tag{5.39}$$

for the same constants as in the statement of Theorem 3.3. We deduce from Theorem 3.3 that there exists a constant C (independent of n) such that, for any $i \in \{1, \dots, n\}$ and $\ell \in \{0, \dots, M-1\}$,

$$\|\!|X^i(\omega)|\!\|_{[\tau_\ell,\tau_{\ell+1}],\widehat{w}^{i,n},p'} \le \big[C\big(1 + \widehat{w}^{i,n}_{p'}(0, T, \omega)^{1/p'}\big)\big]^{2\widehat{N}^{i,n}([0,T],\omega,1/(4L))}. \tag{5.40}$$

Observe now that, for any $i \in \{1, \dots, n\}$,

$$\begin{aligned} \sup_{0\le t\le T} |X^i_t - X^i_0| &\le \sum_{\ell=0}^{M-1} \big(\|\!|X^i(\omega)|\!\|_{[\tau_\ell,\tau_{\ell+1}],\widehat{w}^{i,n},p'} \widehat{w}^{i,n}_{p'}(\tau_\ell, \tau_{\ell+1}, \omega)^{1/p'}\big) \\ &\le M\big[C\big(1 + \widehat{w}^{i,n}_{p'}(0, T, \omega)^{1/p'}\big)\big]^{2\widehat{N}^{i,n}([0,T],\omega,1/(4L))+1}. \end{aligned}$$

The second factor in the right-hand side has finite moments of any order, see (5.43) below, replacing therein $\widehat{N}^{i,n}_\ell/\sqrt{\delta_\ell}$ by $N([0, T], \omega, 1/(4L))/\sqrt{T}$. Moreover, we prove below that M has sub-exponential tails, that is, $\mathbb{P}(M > a) \le c\exp(-a^\varepsilon)$, for $c, \varepsilon > 0$. This suffices to prove (5.38).

We now prove that $(\tau_\ell)_{\ell=0,\dots,M}$ in (5.39) may be constructed in such a way that M has indeed sub-exponential tails. Obviously, see for instance (A.1), it suffices to construct, for each constraint in (5.39), a subdivision $(\tau_\ell)_{\ell=0,\dots,M}$ of $[0, T]$, for which the corresponding constraint in (5.39) (and only this one) holds true and the number of points M has sub-exponential tails.

We start with the second constraint in (5.39). By induction, we define the sequence $(\tau'_\ell)_{\ell=0,\dots,M'}$, letting $\tau'_0 := 0$ and $\tau'_{\ell+1} := \inf\{t \ge \tau'_\ell : \Gamma^{(n)}_2(\omega, [\tau'_\ell, t]) \ge \eta_0\} \wedge T$[5] (we recall

[5]The reader may compare with (2.16), paying attention to the fact that, here, $t \mapsto \Gamma^{(n)}_2(\omega, [\tau'_\ell, t])$ is not continuous but just right upper semicontinuous, namely $\lim_{\varepsilon\searrow 0, \varepsilon>0} \Gamma^{(n)}_2(\omega, [\tau'_\ell, t + \varepsilon]) \le \Gamma^{(n)}_2(\omega, [\tau'_\ell, t])$ for any

that $\eta_0 > 1$), with $M' := \inf_{\ell\in\mathbb{N}}\{\ell \in \mathbb{N} : \tau'_\ell = T\}$. We claim that we can choose $M = 2M'$. Indeed, since the counter $\widehat{N}^{i,n}$ appearing in (5.39) is the local accumulation of a continuous function on $\mathcal{S}_2^T$, there exists $\delta_L > 0$ such that, for any $t \in [0, T]$ and any $i \in \{1, \dots, n\}$, $\widehat{N}^{i,n}([t, (t + \delta_L) \wedge T], \omega, 1/(4L)) = 0$. (Of course, δ_L depends on n and ω, but this is not a problem in the rest of the proof.) Then, for any point $t \in [\tau'_\ell, \tau'_{\ell+1})$, we have, by definition of $\tau'_{\ell+1}$, $\Gamma_2^{(n)}(\omega, [\tau'_\ell, t]) < \eta$. Moreover, if $|\tau'_{\ell+1} - t| \le \delta_L$, then $\Gamma_2^{(n)}(\omega, [t, \tau'_{\ell+1}]) = 1 \le \eta$. Therefore, we may choose $\tau_{2\ell} = \tau'_\ell$ for $\ell \in \{0, \dots, M\}$ and then $|\tau_{2\ell+1} - \tau'_{\ell+1}| \le \delta_L$. The sequence $(\tau_\ell)_{\ell=0,\dots,2M}$ satisfies the second constraint in (5.39).

We now prove that M' has sub-exponential tails (which implies that $M = 2M'$ also has sub-exponential tails). Letting $\delta'_\ell := \tau'_{\ell+1} - \tau'_\ell$, for any $\ell \in \mathbb{N}$, we have, for any $A > 1$ (recalling $\gamma_0, \eta_0 > 1$),

$$\begin{aligned}\pi' &:= \mathbb{P}\left(\delta'_\ell < \frac{1}{A}, \ell < M' - 1\right)\\ &\le \mathbb{P}\big({}^{(n)}\big([\gamma^2(1 + \widehat{w}^{\bullet,n}_{p'}(0, T, \omega)^{1/p'})]^{\widehat{N}^{\bullet,n}_\ell(\omega)/\sqrt{\delta'_\ell}}\big)^{1/\sqrt{A}}_{32} \ge \eta_0\big)\\ &= \mathbb{P}\big({}^{(n)}\big([\gamma^2(1 + \widehat{w}^{\bullet,n}_{p'}(0, T, \omega)^{1/p'})]^{\widehat{N}^{\bullet,n}_\ell(\omega)/\sqrt{\delta'_\ell}}\big)_{32} \ge \eta_0^{\sqrt{A}}\big),\end{aligned}$$

with the shorten notation $\widehat{N}^{\bullet,n}_\ell(\omega) = \widehat{N}^{\bullet,n}([\tau'_\ell, \tau'_{\ell+1}], \omega, 1/(4L))$; in the second line, we used the fact that ${}^{(n)}([\gamma^2(1 + \widehat{w}^{\bullet,n}_{p'}(0, T, \omega)^{1/p'})]^{\widehat{N}^{\bullet,n}_\ell(\omega)})_{32} \ge \eta_0$, see footnote (5). We now introduce the function $f(x) = \exp(\ln(x)^{1+\varepsilon})$, $x > 1$; it is nondecreasing on $[1, \infty)$ and convex on $[e, \infty)$. By Markov inequality,

$$\begin{aligned}\pi' &\le e^{-(\ln[\eta_0^{32\sqrt{A}}])^{1+\varepsilon}}\mathbb{E}\left[f\left(\frac{1}{n}\sum_{i=1}^n e\big[\gamma^2(1 + \widehat{w}^{i,n}_{p'}(0, T, \cdot)^{1/p'})\big]^{32\widehat{N}^{i,n}_\ell/\sqrt{\delta'_\ell}}\right)\right]\\ &\le e^{-(\ln[\eta_0^{32\sqrt{A}}])^{1+\varepsilon}}\frac{1}{n}\sum_{i=1}^n \mathbb{E}\big[f\big(e\big[\gamma^2(1 + \widehat{w}^{i,n}_{p'}(0, T, \cdot)^{1/p'})\big]^{32\widehat{N}^{i,n}_\ell/\sqrt{\delta'_\ell}}\big)\big],\end{aligned}$$

with $e = \exp(1)$. We prove in (5.42) below that, for ε small enough (independently of n),

$$\sup_{i=1,\dots,n}\mathbb{E}\big[f\big(e\big[\gamma^2(1 + \widehat{w}^{i,n}_{p'}(0, T, \cdot)^{1/p'})\big]^{32\widehat{N}^{i,n}_\ell/\sqrt{\delta'_\ell}}\big)\big] \le C, \tag{5.41}$$

for C independent of n. As a result, $\pi' \le C\exp(-(32\ln(\eta_0))^{1+\varepsilon}A^{(1+\varepsilon)/2})$, and then,

$$\begin{aligned}\mathbb{P}(M' > \ell + 1) &= \mathbb{P}\big(\delta'_1 + \cdots + \delta'_\ell < T, \ell + 1 < M'\big)\\ &\le \sum_{i=1}^\ell \mathbb{P}\left(\delta'_i < \frac{T}{\ell}, i + 1 < M'\right)\\ &\le C\ell e^{-(32\ln(\eta_0))^{1+\varepsilon}(\ell/T)^{(1+\varepsilon)/2}},\end{aligned}$$

which shows that M' has sub-exponential tails.

We now check what happens when handling the first constraint in (5.39). We may define M' as before, that is $M' := \inf_{\ell\in\mathbb{N}}\{\ell \in \mathbb{N} : \tau'_\ell = T\}$ with $\tau'_0 := 0$ and $\tau'_{\ell+1} := \inf\{t \ge \tau'_\ell :$

$t \in [\tau'_\ell, T]$. In order to check the latter, it suffices to prove that $t \mapsto \widehat{N}^{i,n}([\tau'_\ell, t], \omega, 1/(4L))$ is also right upper semicontinuous, for any $i \in \{1, \dots, n\}$. Assume indeed that, for an index $i \in \{1, \dots, n\}$, for a time $t \ge \tau'_\ell$ and for an integer $\ell \ge 0$, it holds that $\widehat{N}^{i,n}([\tau'_\ell, t + \varepsilon], \omega, 1/(4L)) \ge \ell$ for any $\varepsilon > 0$. Then, for any $\varepsilon > 0$, we can find $\ell + 1$ reals $\tau'_\ell =: t_0^{(\varepsilon)} < t_1^{(\varepsilon)} < \cdots < t_\ell^{(\varepsilon)} \le t + \varepsilon$ such that $\widehat{w}^{i,n}_{p'}(t_j^{(\varepsilon)}, t_{j+1}^{(\varepsilon)}, \omega)^{1/p'} \ge 1/(4L)$. By an obvious compactness argument and by continuity of $\widehat{w}^{i,n}_{p'}$, we deduce that there exists $\ell + 1$ points $\tau'_\ell =: t_0^{(0)} < t_1^{(0)} < \cdots < t_\ell^{(0)} \le t$ such that $\widehat{w}^{i,n}_{p'}(t_j^{(0)}, t_{j+1}^{(0)}, \omega)^{1/p'} \ge 1/(4L)$, which in turn implies that $\widehat{N}^{i,n}([\tau'_\ell, t], \omega, 1/(4L)) \ge \ell$.

$\Gamma_1^{(n)}(\omega,[\tau'_\ell,t])\geq 1\}\wedge T$. Then, we can repeat the same proof as above by using the fact that

$$\left\{\delta'_\ell<\frac{1}{A},\ell<M'-1\right\}\subset\left\{{}^{(n)}\left(\frac{\widehat{N}^{\bullet,n}([\tau'_\ell,\tau'_{\ell+1}],\omega,1/(4L_0))}{\sqrt{\delta'_\ell}}\right)_8\geq\sqrt{A}\right\}$$

and by recalling that

$${}^{(n)}\left(\frac{\widehat{N}^{\bullet,n}([\tau'_\ell,\tau'_{\ell+1}],\omega,1/(4L_0))}{\sqrt{\delta'_\ell}}\right)_8$$

has Weibull tails with shape parameter strictly greater than 1,[6] uniformly in the choice of the dissection $0=\tau_0<\cdots<\tau_{M'}=T$, which follows from the third item in the assumption of Theorem 5.1 together with the convexity of the function $[0,+\infty)\ni x\mapsto\exp(x^{1+\varepsilon})$, for $\varepsilon\geq 0$. This permits to provide an upper bound for $\mathbb{P}(\delta'_\ell<1/A,\ell<M'-1)$ and then to deduce as before that M' has sub-exponential tails.

It now remains to prove (5.41). By (4.26) and (4.27), we can find a real $\varepsilon_1>0$, independent of n, such that $\sup_{i=1,\dots,n}\mathbb{E}[\exp(\widehat{w}^{i,n}_{p'}(0,T,\cdot)^{\varepsilon_1})]\leq C$, for C independent of n. Hence, combining with the third item in the assumption of the statement, we get, for any $n\geq 1$, $i\in\{1,\dots,n\}$, $a>1$ and $K>0$,

$$\begin{aligned}&\mathbb{P}\big((1+\widehat{w}^{i,n}_{p'}(0,T,\cdot)^{1/p'}\big)^{\widehat{N}^{i,n}_\ell/\sqrt{\delta'_\ell}}\geq a\big)\\&\leq\mathbb{P}\left(\frac{\widehat{N}^{i,n}_\ell}{\sqrt{\delta'_\ell}}\geq K\right)+\mathbb{P}\big(1+\widehat{w}^{i,n}_{p'}(0,T,\cdot)^{1/p'}\geq a^{1/K}\big)\\&\leq ce^{-K^{1+\varepsilon_2}}+ce^{-a^{\varepsilon_1 p'/K}},\end{aligned}\tag{5.42}$$

for a new constant c independent of n and i. Choosing $K=(\ln a)^{1/(1+\varepsilon_2/2)}$, we deduce that there exist a constant $c>1$ and an exponent $\varepsilon>0$ such that, for any $a>0$,

$$\mathbb{P}\big((1+\widehat{w}^{i,n}_{p'}(0,T,\cdot)^{1/p'}\big)^{\widehat{N}^{i,n}_\ell/\sqrt{\delta'_\ell}}\geq a\big)\leq ce^{-c^{-1}\ln(a)^{1+2\varepsilon}},\tag{5.43}$$

from which we obtain (5.41). □

APPENDIX: INTEGRABILITY AND AUXILIARY ESTIMATES

We prove in this appendix auxiliary results that we left aside in the body of the text to keep focused on the main problems at hand. In Appendix A.1, we show that assumption (c) in Theorem 5.1 holds true for interacting particle system driven by Gaussian rough paths satisfying Example 4.2, see Remark 5.2. Appendix A.2 is dedicated to proving a crucial moment estimate for some quantity of interest in Step 1 of the proof of Theorem 5.1. *This is where the convergence rate ς_n appears*, see for instance (5.11). In the last Appendix A.3, we elaborate on the versions of law of large numbers used in the text.

A.1. Gaussian case. Remark 5.2 asserts that the assumptions of Theorem 4.3 are satisfied in the Gaussian framework specified in Example 4.2. Since the derivation of (4.11) is already justified in the latter example, we only prove here that we can control the empirical local accumulation as in the requirement (c) of Theorem 5.1 with $p'=p$ therein.

[6]Recall that a positive random variable A has a Weibull tail with shape parameter $2/\varrho$ if $A^{1/\rho}$ has a Gaussian tail.

Following the proof of [5], Theorem 2.4, we may focus on the local accumulation of each of the various terms in (5.4). To make it clear, we have the following property: For a given threshold $\alpha > 0$ and for any two continuous functions $v_1 : \mathcal{S}_2^T \to \mathbb{R}_+$ and $v_2 : \mathcal{S}_2^T \to \mathbb{R}_+$, set $N_i(\alpha) := N_{v_i}([0,T],\alpha)$, for $1 \le i \le 2$, and $N(\alpha) := N_{v_1+v_2}([0,T],\alpha)$, see (2.17) for the original definition, then

$$\max\left(N_1\left(\frac{\alpha}{2}\right), N_2\left(\frac{\alpha}{2}\right)\right) \ge N(\alpha). \tag{A.1}$$

Throughout the proof, we choose Ω as the space $\mathcal{W} = \mathcal{C}([0,T];\mathbb{R}^d)$, equipped with the law $\mathbb{P}$ of the Gaussian process addressed in Example 4.2. We call $\mathcal{H}$ the corresponding Cameron-Martin space and we regard $(\mathcal{W},\mathcal{H},\mathbb{P})$ as an abstract Wiener space. We then regard $(W^1,\dots,W^n)$ as the canonical process on Ω^n equipped with the product measure $\mathbb{P}^{\otimes n}$. We recall from [28], Theorem 10.4, that the processes $(\mathbb{W}^i)_{1\le i\le n}$ and $(\mathbb{W}^{i,j})_{1\le i,j\le n}$ may be regarded as random variables on Ω^n. We first perform the proof when $[\tau,\tau']$ in the requirement (c) of Theorem 5.1 is the interval $[0,T]$ itself; we explain in the last step of the proof why this may be generalized to any (possibly random) subinterval $[\tau,\tau']$ of $[0,T]$.

Step 1. The first step is to consider, for a given $\alpha > 0$, the accumulation $\widetilde{N}^i([0,T],\omega,\alpha)$ associated with $\|W^i(\omega)\|^p_{[s,t],p-\mathrm{v}} + \|\mathbb{W}^i(\omega)\|^{p/2}_{[s,t],p/2-\mathrm{v}}$, see (4.3), namely

$$\widetilde{N}^i\big([0,T],\omega,\alpha\big) := N_{\varpi}\big([0,T],\alpha\big),$$

when

$$\varpi(s,t)^p = \big\|W^i(\omega)\big\|^p_{[s,t],p-\mathrm{v}} + \big\|\mathbb{W}^i(\omega)\big\|^{p/2}_{[s,t],p/2-\mathrm{v}},$$

but this follows from [5], Theorem 2.4, and from an obvious exchangeability argument. The term $\widehat{v}^{i,n}_{p'}(s,t,\omega)$ in (5.4) is handled in the same way.

Step 2. We now focus on the local accumulation of the fourth and fifth terms in (4.3). For simplicity, we just explain what happens for the fourth term. The fifth term may be handled in the same way.

We use the same notation as in Section 4.1 and proceed as in the proof of [5], Theorem 2.4. The Gaussian process $(W^1,\dots,W^n)$ has $(\mathcal{W}^n,\mathcal{H}^{\oplus n},\mathbb{P}^{\otimes n})$ as abstract Wiener space. For $\omega = (\omega_i)_{i=1}^n \in \Omega^n$ and for $\boldsymbol{h} = \bigoplus_{i=1}^n h_i \in \mathcal{H}^{\oplus n}$, we let

$$T_{\boldsymbol{h}}\boldsymbol{W}^{(n)}(\omega) = T_{\bigoplus_{i=1}^n h_i}\boldsymbol{W}^{(n)}(\omega)$$

for the translated rough path along $\boldsymbol{h}$ (see [28], (11.5)). By [28], Lemma 11.4, and by Young's inequality, with probability 1 under $\mathbb{P}^{\otimes n}$, for all $\boldsymbol{h} \in \mathcal{H}^{\oplus n}$,

$$\begin{aligned}\big\|\mathbb{W}^{i,j}(\omega)\big\|^{p/2}_{[s,t],(p/2)-\mathrm{v}} \le c\big(&\big\|(T_{\boldsymbol{h}}\mathbb{W})^{i,j}(\omega)\big\|^{p/2}_{[s,t],(p/2)-\mathrm{v}} + \big\|(T_{\boldsymbol{h}}W)^i(\omega)\big\|^p_{[s,t],p-\mathrm{v}}\\ &+ \big\|(T_{\boldsymbol{h}}W)^j(\omega)\big\|^p_{[s,t],p-\mathrm{v}} + \|h_i\|^p_{[s,t],\varrho-\mathrm{v}} + \|h_j\|^p_{[s,t],\varrho-\mathrm{v}}\big).\end{aligned}$$

Importantly, the constant c is independent of n. Below, it is allowed to increase from line to line as long as it remains independent of n. So,

$$\begin{aligned}&{}^{(n)}\big(\big\|\mathbb{W}^{i,\bullet}(\omega)\big\|^{p/2}_{[s,t],(p/2)-\mathrm{v}}\big)_q\\ &\le c\big\{{}^{(n)}\big(\big\|(T_{\boldsymbol{h}}\mathbb{W})^{i,\bullet}(\omega)\big\|^{p/2}_{[s,t],(p/2)-\mathrm{v}}\big)_q + {}^{(n)}\big(\big\|(T_{\boldsymbol{h}}W)^{\bullet}(\omega)\big\|^p_{[s,t],p-\mathrm{v}}\big)_q\\ &\quad + \big\|(T_{\boldsymbol{h}}W)^i(\omega)\big\|^p_{[s,t],p-\mathrm{v}} + \|h_i\|^p_{[s,t],\varrho-\mathrm{v}} + {}^{(n)}\big(\|h_\bullet\|^p_{[s,t],\varrho-\mathrm{v}}\big)_q\big\}\\ &\le c\big\{{}^{(n)}\big(\big[\!\big[(T_{\boldsymbol{h}}\boldsymbol{W})^{i,\bullet}(\omega)\big]\!\big]^p_{[0,T],(1/p)-\mathrm{H}}\big)_q(t-s) + \|h_i\|^p_{[s,t],\varrho-\mathrm{v}}\\ &\quad + {}^{(n)}\big(\|h_\bullet\|^p_{[s,t],\varrho-\mathrm{v}}\big)_q\big\},\end{aligned} \tag{A.2}$$

where for any $i, j \in \{1, \dots, n\}$, we let

$$\|\!|\boldsymbol{W}^{i,j}(\omega)|\!\|_{[s,t],(1/p)-\mathrm{H}} := \|(W^i, W^j)(\omega)\|_{[s,t],(1/p)-\mathrm{H}} + \sqrt{\|\mathbb{W}^{i,j}(\omega)\|_{[s,t],(2/p)-\mathrm{H}}},$$

and similarly for $\|\!|(T_{\boldsymbol{h}}\boldsymbol{W})^{i,j}(\omega)|\!\|_{[0,T],(1/p)-\mathrm{H}}$.

The tricky term in (A.2) is the last one on the last line. The key point is to notice that, for a given $\varepsilon \in (0, 2-\rho)$,

$$\begin{aligned}
{}^{(n)}\big(\|h_\bullet\|^p_{[s,t],\varrho-\mathrm{v}}\big)_q &= \left[\frac{1}{n}\sum_{j=1}^n \|h_j\|^{pq}_{[s,t],\varrho-\mathrm{v}}\right]^{1/q} \\
&= n^{-1/q}\left\{\left[\sum_{j=1}^n \|h_j\|^{pq}_{[s,t],\varrho-\mathrm{v}}\right]^{(2-\varepsilon)/(pq)}\right\}^{p/(2-\varepsilon)} \\
&\le n^{-1/q}\left[\sum_{j=1}^n \|h_j\|^{2-\varepsilon}_{[s,t],\varrho-\mathrm{v}}\right]^{p/(2-\varepsilon)},
\end{aligned}$$

where we used the fact that $2-\varepsilon < pq$. Observe in particular that, whenever $\sum_{j=1}^n \|h_j\|^{2-\varepsilon}_{[s,t],\varrho-\mathrm{v}} \le n^{(2-\varepsilon)/(pq)}$, it holds

$$\begin{aligned}
{}^{(n)}\big(\|h_\bullet\|^p_{[s,t],\varrho-\mathrm{v}}\big)_q &\le n^{-1/q}\left[\sum_{j=1}^n \|h_j\|^{2-\varepsilon}_{[s,t],\varrho-\mathrm{v}}\right]^{p/(2-\varepsilon)} \\
&\le n^{-1/q}\big(n^{(2-\varepsilon)/(pq)}\big)^{p/(2-\varepsilon)-1}\sum_{j=1}^n \|h_j\|^{2-\varepsilon}_{[s,t],\varrho-\mathrm{v}} \\
&= n^{-(2-\varepsilon)/(pq)}\sum_{j=1}^n \|h_j\|^{2-\varepsilon}_{[s,t],\varrho-\mathrm{v}},
\end{aligned}$$

where, in the second line, we used the fact that $p/(2-\varepsilon) > 1$. Returning to (A.2), we deduce that, whenever $\|h_i\|_{[s,t],\varrho-\mathrm{v}} \le 1$ and $\sum_{j=1}^n \|h_j\|^{2-\varepsilon}_{[s,t],\varrho-\mathrm{v}} \le n^{(2-\varepsilon)/(pq)}$,

$$\begin{aligned}
{}^{(n)}\big(\|\mathbb{W}^{i,\bullet}(\omega)\|^{p/2}_{[s,t],(p/2)-\mathrm{v}}\big)_q \le c\Bigg\{ & {}^{(n)}\big(\|\!|(T_{\boldsymbol{h}}\boldsymbol{W})^{i,\bullet}(\omega)|\!\|^p_{[0,T],(1/p)-\mathrm{H}}\big)_q(t-s) \\
& + \|h_i\|^{2-\varepsilon}_{[s,t],\varrho-\mathrm{v}} + n^{-(2-\varepsilon)/(pq)}\sum_{j=1}^n \|h_j\|^{2-\varepsilon}_{[s,t],\varrho-\mathrm{v}}\Bigg\}.
\end{aligned} \tag{A.3}$$

When the left-hand side is less than or equal to α^p, we can modify the constant c in such a way that the inequality remains true when $\|h_i\|_{[s,t],\varrho-\mathrm{v}} \ge 1$ or $\sum_{j=1}^n \|h_j\|^{2-\varepsilon}_{[s,t],\varrho-\mathrm{v}} \ge n^{(2-\varepsilon)/(pq)}$. Noticing that $2-\varepsilon > \rho$, (A.3) remains true with ${}^{(n)}(\mathbb{W}^{i,\bullet}(\omega))^{p/2}_{q;[s,t],(p/2)-\mathrm{v}}$ in the left-hand side.

Define now $N^{i,n,\perp}([0,T],\omega,\alpha) := N_\varpi([0,T],\alpha)$, when

$$\varpi(s,t)^p = {}^{(n)}\big(\mathbb{W}^{i,\bullet}(\omega)\big)^{p/2}_{q;[s,t],(p/2)-\mathrm{v}}.$$

Then, (A.3) (together with $2-\varepsilon > \rho$) yields

$$\begin{aligned}
N^{i,n,\perp}\big([0,T],\omega,\alpha\big)\alpha^p \le c\Bigg\{ & {}^{(n)}\big(\|\!|(T_{\boldsymbol{h}}\boldsymbol{W})^{i,\bullet}(\omega)|\!\|^p_{[0,T],(1/p)-\mathrm{H}}\big)_q T \\
& + \|h_i\|^{2-\varepsilon}_{[0,T],\varrho-\mathrm{v}} + n^{-(2-\varepsilon)/(pq)}\sum_{j=1}^n \|h_j\|^{2-\varepsilon}_{[0,T],\varrho-\mathrm{v}}\Bigg\}
\end{aligned}$$

$$\leq c\Biggl\{ {}^{(n)}\bigl(\llbracket (T_{\boldsymbol{h}}\boldsymbol{W})^{i,\bullet}(\omega)\rrbracket^{p}_{[0,T],(1/p)-\mathrm{H}}\bigr)_q T + \|h_i\|^{2-\varepsilon}_{[0,T],\varrho-\mathrm{v}} + n^{-(2-\varepsilon)/(pq)+\varepsilon/2}\Biggl[\sum_{j=1}^{n}\|h_j\|^2_{[0,T],\varrho-\mathrm{v}}\Biggr]^{(2-\varepsilon)/2}\Biggr\},$$

where we applied Hölder's inequality to handle the last term. By choosing ε small enough such that $(2-\varepsilon)/(pq)-\varepsilon/2>0$ and by applying Proposition 11.2 in [28], we get, for a possibly new value of the constant c,

$$\begin{aligned} &N^{i,n,\perp}([0,T],\omega,\alpha)\alpha^p \\ &\quad\leq c\bigl\{ {}^{(n)}\bigl(\llbracket (T_{\boldsymbol{h}}\boldsymbol{W})^{i,\bullet}(\omega)\rrbracket^{p}_{[0,T],(1/p)-\mathrm{H}}\bigr)_q T + \|\boldsymbol{h}\|^{2-\varepsilon}_{\mathcal{H}^{\oplus n}} T^{(2-\varepsilon)/(2\rho)}\bigr\}, \end{aligned} \tag{A.4}$$

with $\|\boldsymbol{h}\|^2_{\mathcal{H}^{\oplus n}}=\sum_{i=1}^n\|h_i\|^2_{\mathcal{H}}$. We then notice that $(2-\varepsilon)/(2\varrho)>1/2$ since $2-\varepsilon>\varrho$. We deduce that $T^{(2-\varepsilon)/(2\varrho)}\leq cT^{1/2}$ for a possibly new value of the constant c. We then apply Theorems 11.5 and 11.7 in [28] but on the space $(\mathcal{W}^{\otimes n},\mathcal{H}^{\oplus n},\mathbb{P}^{\otimes n})$. Importantly, we observe that

$$\mathbb{E}\bigl[{}^{(n)}\bigl(\llbracket \boldsymbol{W}^{i,\bullet}(\omega)\rrbracket^{p}_{[0,T],(1/p)-\mathrm{H}}\bigr)_q\bigr]$$

is bounded by a constant c, independent of i and n, which proves that $N^{i,n,\perp}([0,T],\cdot,\alpha)/\sqrt{T}$ has a Weibull distribution with shape parameter $2/(2-\varepsilon)$, independently of n.

Step 3. We now turn to the local accumulation of the sixth term in (4.3). Taking the norm ${}^{(n)}(\cdot)_q$ in (A.2), we get, with probability 1 under $\mathbb{P}^{\otimes n}$, for all $\boldsymbol{h}\in\mathcal{H}^{\oplus n}$,

$$\begin{aligned} &{}^{(n)}\bigl(\bigl(\|\mathbb{W}^{\bullet,\bullet}(\omega)\|^{p/2}_{[s,t],(p/2)-\mathrm{v}}\bigr)\bigr)_q \\ &\quad\leq c\bigl\{ {}^{(n)}\bigl(\bigl(\llbracket (T_{\boldsymbol{h}}\boldsymbol{W})^{\bullet,\bullet}(\omega)\rrbracket^{p}_{[0,T],(1/p)-\mathrm{H}}\bigr)\bigr)_q(t-s) + {}^{(n)}\bigl(\|h_\bullet\|^p_{[s,t],\varrho-\mathrm{v}}\bigr)_q\bigr\}. \end{aligned}$$

Following the proof of (A.3), we deduce that

$$\begin{aligned} {}^{(n)}\bigl(\bigl(\mathbb{W}^{\bullet,\bullet}(\omega)\bigr)\bigr)^{p/2}_{q;[s,t],p/2-\mathrm{v}} \leq c\Biggl\{ &{}^{(n)}\bigl(\bigl(\llbracket (T_{\boldsymbol{h}}\boldsymbol{W})^{\bullet,\bullet}(\omega)\rrbracket^{p}_{[0,T],(1/p)-\mathrm{H}}\bigr)\bigr)_q(t-s) \\ &+ n^{-(2-\varepsilon)/(pq)}\sum_{j=1}^{n}\|h_j\|^{2-\varepsilon}_{[s,t],\varrho-\mathrm{v}}\Biggr\}, \end{aligned}$$

at least when the left-hand side is less than or equal to α^p. Importantly, there is no need to distinguish the coordinate i of $\boldsymbol{h}$ from the other coordinates $j\neq i$ since the coefficient in front of any $\|h_j\|_{[s,t],\varrho-\mathrm{v}}$, $j=1,\ldots,n$, has the same power decay as n tends to ∞. So, the context is simpler than in the previous step and we may conclude in the same way.

Local accumulations associated to the second term in (4.3) and to $(s,t)\mapsto {}^{(n)}(v^{\bullet,n}_{p'}(\omega))_{q;[s,t],1-\mathrm{v}}$ and $(s,t)\mapsto {}^{(n)}(\widehat{v}^{\bullet,n}_{p'}(\omega))_{q;[s,t],1-\mathrm{v}}$ in (5.4) are handled in the same way. (As for the latter one, the reader may refer to the proof of [5], Theorem 2.4.)

Step 4. The proof has been here achieved on the interval $[0,T]$. Importantly, the fact that T is deterministic does not play any role in the proof. It is in particular quite easy to see that the interval $[0,T]$ can be replaced by any (random) sub-interval $[\tau,\tau']\subset[0,T]$ as in the requirement (c) of Theorem 5.1.

A.2. An auxiliary estimate. We prove in this appendix some auxiliary estimates that were used in Step 1 of the proof of Theorem 5.1. This is where the convergence rate ς_n in Theorem 5.1 appears. Recall we set $\varsigma_n=n^{-1/2}$ if $d=1$, and $\varsigma_n=n^{-1/2}\ln(1+n)$, if $d=2$, and $\varsigma_n=n^{-1/d}$, if $d\geq 3$. Recall also definitions (5.7), (5.8), (5.9) and (5.10).

LEMMA A.3. *Fix $\varrho \geq 8$. There exists an exponent ϱ' such that, whenever $X_0(\cdot) \in \mathbb{L}^{\varrho'}$, we can find another constant C, depending on $\langle X_0(\cdot)\rangle_{\varrho'}$ and satisfying, for any integers $1 \leq i \leq n$ and any $0 \leq r \leq s \leq t \leq T$,*

$$\begin{aligned}
&\big\langle \big[F^{i,n}(\cdot) - \overline{F}^{i}(\cdot)\big]_{s,t}\big\rangle_{\varrho} \leq C\varsigma_n \big\langle\!\big\langle w^+(s,t,\cdot,\cdot)\big\rangle\!\big\rangle_{\varrho'}^{1/p},\\
&\big\langle \mathcal{I}^{i,n,\partial}_{\{s,t\}}(\cdot) - \overline{\mathcal{I}}^{i,\partial}_{\{s,t\}}(\cdot)\big\rangle_{\varrho} \leq C\varsigma_n \big(\big\langle\!\big\langle w^+(s,t,\cdot,\cdot)\big\rangle\!\big\rangle_{\varrho'}^{1/p} + \big\langle\!\big\langle w^+(s,t,\cdot,\cdot)\big\rangle\!\big\rangle_{\varrho'}^{2/p}\big),\\
&\bigg(\int_\Omega \bigg|\frac{1}{n}\sum_{j=1}^{n} \delta_\mu F^{i,j,n}_s(\omega)\mathbb{W}^{j,i}_{s,t}(\omega) - \mathbb{E}\big[\delta_\mu \overline{F}^{i}_s(\omega,\cdot)\mathbb{W}^{i,\perp}_{s,t}(\cdot,\omega)\big]\bigg|^{\rho} d\mathbb{P}(\omega)\bigg)^{1/\rho}\\
&\quad + \big\langle \big(\delta_x F^{i,n}_s(\cdot) - \delta_x \overline{F}^{i}_s(\cdot)\big)\mathbb{W}^{i}_{s,t}(\cdot)\big\rangle_{\rho} \leq C\varsigma_n \big\langle\!\big\langle w^+(s,t,\cdot,\cdot)\big\rangle\!\big\rangle_{\varrho'}^{2/p},\\
&\big\langle \big\{\mathcal{I}^{i,n,\partial}_{\{r,s\}}(\cdot) + \mathcal{I}^{i,n,\partial}_{\{s,t\}}(\cdot) - \mathcal{I}^{i,n,\partial}_{\{r,t\}}(\cdot)\big\} - \big\{\overline{\mathcal{I}}^{i,\partial}_{\{r,s\}}(\cdot) + \overline{\mathcal{I}}^{i,\partial}_{\{s,t\}}(\cdot) - \overline{\mathcal{I}}^{i,\partial}_{\{r,t\}}(\cdot)\big\}\big\rangle_{\varrho}\\
&\quad \leq C\varsigma_n \big\langle\!\big\langle w^+(r,t,\cdot,\cdot)\big\rangle\!\big\rangle_{\varrho'}^{3/p},
\end{aligned}$$

where $w^+(r,t,\omega,\omega') := w(r,t,\omega) + \|\mathbb{W}^{\perp}(\omega,\omega')\|^{p/2}_{[r,t],\,p/2-\mathrm{v}}$, with w as in (2.10) for the same parameters p and q as therein.

The reason the appearance of the quantity w^+ instead of w, in the above upper bounds, will appear at the beginning of Step 2 in the proof.

PROOF. We directly prove the last inequality in the statement; the first three inequalities follow from similar computations. Throughout the proof, we use the following notations. For each $i \in \{1,\dots,n\}$, we call $\overline{w}^i$ the control associated with $\overline{\boldsymbol{W}}^i(\cdot)$ through identity (2.10). For $j \in \{1,\dots,n\}$, we also let

$$\overline{w}^{i,j}(s,t,\omega) := \big\|\mathbb{W}^{i,j}(\omega)\big\|^{p}_{[s,t],\,p-\mathrm{v}}.$$

We make in the course of the proof an intense use of Lemma A.4 below, giving the convergence rate of the empirical measure of a sample of independent and identically distributed random variables towards their common law. In this regard, a key fact is that the theoretical distribution driving the empirical one must be sufficiently integrable. By a variant of (5.38), we already know that, for any $\rho \geq 1$, there exists $\rho' \geq 8$ such that $\sup_{0\leq t\leq T}|X_t(\cdot)|$ is in $\mathbb{L}^{\rho}$ as soon as $X_0(\cdot)$ is in $\mathbb{L}^{\rho'}$. The proof of this variant is in fact simpler than the proof of (5.38) itself, since we can directly invoke Theorem 3.3 instead of (5.40), noticing that the analogue of M in (5.40) then becomes deterministic, see for instance footnote (3). Importantly, the same holds true with $|\!|\!|X(\cdot)|\!|\!|_{[0,T],w,p}$: it belongs to $\mathbb{L}^{\rho}$ if $X_0(\cdot)$ is in $\mathbb{L}^{\rho'}$, for a well-chosen ρ'. The proof also follows from Theorem 3.3, by concatenating a deterministic finite number of intervals of the form $[S_1,S_2]$, see [5], footnote (5), for some details about concatenation. We then compute

$$\begin{aligned}
&\big\{\mathcal{I}^{i,n,\partial}_{\{r,s\}}(\omega) + \mathcal{I}^{i,n,\partial}_{\{s,t\}}(\omega) - \mathcal{I}^{i,n,\partial}_{\{r,t\}}(\omega)\big\} - \big\{\overline{\mathcal{I}}^{i,\partial}_{\{r,s\}}(\omega) + \overline{\mathcal{I}}^{i,\partial}_{\{s,t\}}(\omega) - \overline{\mathcal{I}}^{i,\partial}_{\{r,t\}}(\omega)\big\}\\
&\quad = \big(R^{F^{i,n}}_{r,s}(\omega) - R^{\overline{F}^{i}}_{r,s}(\omega)\big)W^{i}_{s,t}(\omega) + \big(\delta_x F^{i,n}_{r,s}(\omega) - \delta_x \overline{F}^{i}_{r,s}(\omega)\big)\mathbb{W}^{i}_{s,t}(\omega)\\
&\qquad + \bigg(\frac{1}{n}\sum_{j=1}^{n} \delta_\mu F^{i,j,n}_{r,s}(\omega)\mathbb{W}^{j,i}_{s,t}(\omega) - \mathbb{E}\big[\delta_\mu \overline{F}^{i}_{r,s}(\omega,\cdot)\mathbb{W}^{i,\perp}_{s,t}(\cdot,\omega)\big]\bigg),
\end{aligned}$$

where

$$
\begin{aligned}
R^{F^{i,n}}_{r,s}(\omega) &:= F^{i,n}_s(\omega) - F^{i,n}_r(\omega) - \delta_x F^{i,n}_r(\omega) W^i_{r,s}(\omega) \\
&\quad - \frac{1}{n}\sum_{j=1}^n \delta_\mu F^{i,j,n}_r(\omega) W^j_{r,s}(\omega), \\
R^{\overline{F}^i}_{r,s}(\omega) &:= \overline{F}^i_s(\omega) - \overline{F}^i_r(\omega) - \delta_x \overline{F}^i_r(\omega) W^i_{r,s}(\omega) \\
&\quad - \mathbb{E}\big[\delta_\mu \overline{F}^i_r(\omega,\cdot) W^i_{r,s}(\cdot)\big].
\end{aligned}
\tag{A.5}
$$

Following (5.8) and (5.9), we define differentiable functions G_x and G_μ of their arguments setting

$$
\begin{aligned}
\delta_x F^{i,n}_t(\omega) &=: G_x\big(\overline{X}^i_t(\omega), \overline{\mu}^n_t(\omega)\big), \qquad \delta_x \overline{F}^i_t(\omega) =: G_x\big(\overline{X}^i_t(\omega), \mathcal{L}(X_t)\big), \\
\delta_\mu F^{i,j,n}_t(\omega) &=: G_\mu\big(\overline{X}^i_t(\omega), \overline{\mu}^n_t(\omega)\big)\big(\overline{X}^j_t(\omega)\big), \\
\delta_\mu \overline{F}^i_t(\omega,\cdot) &=: G_\mu\big(\overline{X}^i_t(\omega), \mathcal{L}(X_t)\big)\big(\overline{X}^i_t(\cdot)\big).
\end{aligned}
$$

Finally, we can write the whole difference in the form

$$
\begin{aligned}
&\big\{\mathcal{I}^\partial_{\{r,s\}}(\omega) + \mathcal{I}^\partial_{\{s,t\}}(\omega) - \mathcal{I}^\partial_{\{r,t\}}(\omega)\big\} - \big\{\overline{\mathcal{I}}^\partial_{\{r,s\}}(\omega) + \overline{\mathcal{I}}^\partial_{\{s,t\}}(\omega) - \overline{\mathcal{I}}^\partial_{\{r,t\}}(\omega)\big\} \\
&\quad = \big(R^{F^{i,n}}_{r,s}(\omega) - R^{\overline{F}^i}_{r,s}(\omega)\big) W^i_{s,t}(\omega) \\
&\qquad + \big[G_x\big(\overline{X}^i(\omega), \overline{\mu}^n(\omega)\big) - G_x\big(\overline{X}^i(\omega), \mathcal{L}(X)\big)\big]_{r,s} \mathbb{W}^i_{s,t}(\omega) \\
&\qquad + \frac{1}{n}\sum_{j=1}^n \big[G_\mu\big(\overline{X}^i(\omega), \overline{\mu}^n(\omega)\big)\big(\overline{X}^j(\omega)\big) - G_\mu\big(\overline{X}^i(\omega), \mathcal{L}(X)\big)\big(\overline{X}^j(\omega)\big)\big]_{r,s} \mathbb{W}^{j,i}_{s,t}(\omega) \\
&\qquad + \frac{1}{n}\sum_{j=1}^n \big[G_\mu\big(\overline{X}^i(\omega), \mathcal{L}(X)\big)\big(\overline{X}^j(\omega)\big)\big]_{r,s} \mathbb{W}^{j,i}_{s,t}(\omega) - \mathbb{E}\big[\delta_\mu \overline{F}^i_{r,s}(\omega,\cdot) \mathbb{W}^{i,\perp}_{s,t}(\cdot,\omega)\big].
\end{aligned}
\tag{A.6}
$$

A key fact is that G_x and G_μ are Lipschitz continuous in all the entries, the Lipschitz property in μ being understood with respect to $\mathbf{d}_1$. Moreover, similar to F itself, they are jointly continuously differentiable in all the arguments and the derivatives are Lipschitz continuous, the Lipschitz property in μ being again understood with respect to $\mathbf{d}_1$.

Step 1. Observe that

$$
\begin{aligned}
&\big[G_x\big(\overline{X}^i(\omega), \overline{\mu}^n(\omega)\big)\big]_{r,s} \\
&\quad = \int_0^1 \partial_x G_x\big(\overline{X}^{i,(\lambda)}_{r;(r,s)}(\omega), \overline{\mu}^{n,\lambda}_{r;(r,s)}(\omega)\big)\overline{X}^i_{r,s}(\omega)\, d\lambda \\
&\qquad + \frac{1}{n}\sum_{j=1}^n \int_0^1 D_\mu G_x\big(\overline{X}^{i,(\lambda)}_{r;(r,s)}(\omega), \overline{\mu}^{n,\lambda}_{r;(r,s)}(\omega)\big)\big(\overline{X}^{j,(\lambda)}_{r;(r,s)}(\omega)\big)\overline{X}^j_{r,s}(\omega)\, d\lambda \\
&\quad = \int_0^1 \partial_x G_x\big(\overline{X}^{i,(\lambda)}_{r;(r,s)}(\omega), \overline{\mu}^{n,\lambda}_{r;(r,s)}(\omega)\big)\overline{X}^i_{r,s}(\omega)\, d\lambda \\
&\qquad + \int_{\mathbb{R}^{2d}} \bigg[\int_0^1 D_\mu G_x\big(\overline{X}^{i,(\lambda)}_{r;(r,s)}(\omega), \overline{\mu}^{n,\lambda}_{r;(r,s)}(\omega)\big)(y) z\, d\lambda\bigg] d\overline{\nu}^{n,\lambda}_{r;(r,s)}(\omega; y, z)
\end{aligned}
\tag{A.7}
$$

where

$$
\overline{\mu}^{n,(\lambda)}_{r;(r,s)}(\omega) := \frac{1}{n}\sum_{j=1}^n \delta_{\overline{X}^{j,(\lambda)}_{r;(r,s)}(\omega)}, \qquad \overline{\nu}^{n,(\lambda)}_{s;(s,t)}(\omega) := \frac{1}{n}\sum_{j=1}^n \delta_{(\overline{X}^{j,(\lambda)}_{r;(r,s)}(\omega), \overline{X}^j_{r,s}(\omega))},
$$

with

$$\overline{X}^{j,(\lambda)}_{r;(r,s)}(\omega) := \overline{X}^{j}_{r}(\omega) + \lambda \overline{X}^{j}_{r,s}(\omega).$$

Proceeding similarly with $[G_x(\overline{X}^i(\omega), \mathcal{L}(X))]_{r,s}$, we get

$$\begin{aligned}
&\big[G_x\big(\overline{X}^i(\omega), \overline{\mu}^n(\omega)\big) - G_x\big(\overline{X}^i(\omega), \mathcal{L}(X)\big)\big]_{r,s} \\
&\quad= \int_0^1 \big[\partial_x G_x\big(\overline{X}^{i,(\lambda)}_{r;(r,s)}(\omega), \overline{\mu}^{n,(\lambda)}_{r;(r,s)}(\omega)\big) \\
&\qquad - \partial_x G_x\big(\overline{X}^{i,(\lambda)}_{r;(r,s)}(\omega), \mathcal{L}\big(X^{(\lambda)}_{r;(r,s)}\big)\big)\big]\overline{X}^i_{r,s}(\omega)\,d\lambda \\
&\qquad + \int_{\mathbb{R}^{2d}}\bigg[\int_0^1 D_\mu G_x\big(\overline{X}^{i,(\lambda)}_{r;(r,s)}(\omega), \overline{\mu}^{n,(\lambda)}_{r;(r,s)}(\omega)\big)(y)z\,d\lambda\bigg] d\overline{v}^{n,(\lambda)}_{r;(r,s)}(\omega; y, z) \\
&\qquad - \int_{\mathbb{R}^{2d}}\bigg[\int_0^1 D_\mu G_x\big(\overline{X}^{i,(\lambda)}_{r;(r,s)}(\omega), \mathcal{L}\big(X^{(\lambda)}_{r;(r,s)}\big)\big)(y)z\,d\lambda\bigg] d\mathcal{L}\big(X^{(\lambda)}_{r;(r,s)}, X_{r,s}\big)(y, z),
\end{aligned}$$

where, as before, $X^{(\lambda)}_{r;(r,s)}(\omega) = X_r(\omega) + \lambda X_{r,s}(\omega)$. Splitting the last two terms in the above expansion into

$$\begin{aligned}
&\int_{\mathbb{R}^{2d}}\bigg[\int_0^1 D_\mu G_x\big(\overline{X}^{i,(\lambda)}_{r;(r,s)}(\omega), \overline{\mu}^{n,(\lambda)}_{r;(r,s)}(\omega)\big)(y)z\,d\lambda\bigg] d\overline{v}^{n,(\lambda)}_{r;(r,s)}(\omega; y, z) \\
&\quad - \int_{\mathbb{R}^{2d}}\bigg[\int_0^1 D_\mu G_x\big(\overline{X}^{i,(\lambda)}_{r;(r,s)}(\omega), \mathcal{L}\big(X^{(\lambda)}_{r;(r,s)}\big)\big)(y)z\,d\lambda\bigg] d\overline{v}^{n,(\lambda)}_{r;(r,s)}(\omega; y, z) \\
&\quad + \int_{\mathbb{R}^{2d}}\bigg[\int_0^1 D_\mu G_x\big(\overline{X}^{i,(\lambda)}_{r;(r,s)}(\omega), \mathcal{L}\big(X^{(\lambda)}_{r;(r,s)}\big)\big)(y)z\,d\lambda\bigg] d\overline{v}^{n,(\lambda)}_{r;(r,s)}(\omega; y, z) \\
&\quad - \int_{\mathbb{R}^{2d}}\bigg[\int_0^1 D_\mu G_x\big(\overline{X}^{i,(\lambda)}_{r;(r,s)}(\omega), \mathcal{L}\big(X^{(\lambda)}_{r;(r,s)}\big)\big)(y)z\,d\lambda\bigg] d\mathcal{L}\big(X^{(\lambda)}_{r;(r,s)}, X_{r,s}\big)(y, z),
\end{aligned}$$

we get

$$\begin{aligned}
&\big|\big[G_x\big(\overline{X}^i(\omega), \overline{\mu}^n(\omega)\big) - G_x\big(\overline{X}^i(\omega), \mathcal{L}(X)\big)\big]_{r,s}\big| \\
&\quad\le c\int_0^1 \mathbf{d}_1\big(\overline{\mu}^{n,(\lambda)}_{r;(r,s)}(\omega), \mathcal{L}\big(X^{(\lambda)}_{r;(r,s)}\big)\big)\,d\lambda \\
&\qquad\times\left(\|\!|\overline{X}^i(\omega)|\!\|_{[0,T],\overline{w}^i,p}\,\overline{w}^i(r,s,\omega)^{1/p} + \frac{1}{n}\sum_{k=1}^n \|\!|\overline{X}^k(\omega)|\!\|_{[0,T],\overline{w}^k,p}\,\overline{w}^k(r,s,\omega)^{1/p}\right) \\
&\qquad + c\big|\mathcal{S}^{i,n}_{r,s}\big(\omega, |\overline{X}^\bullet_{r,s}(\omega)|\big)\big|,
\end{aligned}$$

where $\mathcal{S}^{i,n}_{r,s}(\omega, |\overline{X}^\bullet_{r,s}(\omega)|)$ is the n-empirical mean of n variables that are dominated by $(|\overline{X}^j_{r,s}(\omega)| + \langle X_{r,s}(\cdot)\rangle_1)_{j=1,\dots,n}$ and $n-1$ of which are conditionally centred and conditionally independent given the realization of the path $(\overline{X}^i, W^i, \mathbb{W}^i)$. Allowing the value of the constant c to increase from line to line, we obtain

$$\begin{aligned}
&\big|\big[G_x\big(\overline{X}^i(\omega), \overline{\mu}^n(\omega)\big) - G_x\big(\overline{X}^i(\omega), \mathcal{L}(X)\big)\big]_{r,s}\mathbb{W}^i_{s,t}(\omega)\big| \\
&\quad\le c\int_0^1 \mathbf{d}_1\big(\overline{\mu}^{n,(\lambda)}_{r;(r,s)}(\omega), \mathcal{L}\big(X^{(\lambda)}_{r;(r,s)}\big)\big)\,d\lambda \\
&\qquad\times\left[\|\!|\overline{X}^i(\omega)|\!\|_{[0,T],\overline{w}^i,p} + \left(\frac{1}{n}\sum_{k=1}^n \|\!|\overline{X}^k(\omega)|\!\|^2_{[0,T],\overline{w}^k,p}\right)^{1/2}\right]
\end{aligned}$$

$$\times \left[\overline{w}^i(r,t,\omega)^{3/p} + \frac{1}{n}\sum_{k=1}^n \overline{w}^k(r,t,\omega)^{3/p} \right]$$
$$+ c|\mathcal{S}^{i,n}_{r,s}(\omega, |\overline{X}^\bullet_{r,s}(\omega)|)|\overline{w}^i(r,t,\omega)^{2/p}.$$

In order to conclude for the second term in the right-hand side of (A.6), it suffices to recall from Rosenthal's inequality (applied under the conditional probability given the realization of the path $(\overline{X}^i, W^i, \mathbb{W}^i)$) that

$$\begin{aligned}\langle \mathcal{S}^{i,n}_{r,s}(\cdot, |\overline{X}^\bullet_{r,s}(\cdot)|)\rangle_{3\varrho/2} &\le cn^{-1/2}\langle |||X(\cdot)|||_{[0,T],w,p} w(r,s,\cdot)^{1/p}\rangle_{3\varrho/2} \\ &\le cn^{-1/2}\langle |||X(\cdot)|||_{[0,T],w,p}\rangle_{\chi\varrho}\langle w(r,t,\cdot)\rangle^{1/p}_{\chi\varrho},\end{aligned}$$

where $\chi \ge 1$ is a universal constant whose value may change from line to line (as long as it remains universal). If ρ is large enough, we deduce from Lemma A.4 that

$$\begin{aligned}&\langle [G_x(\overline{X}^i(\cdot), \overline{\mu}^n(\cdot)) - G_x(\overline{X}^i(\cdot), \mathcal{L}(X))]_{r,s}\mathbb{W}^i_{s,t}(\cdot)\rangle_\varrho \\ &\le c\left(\int_0^1 \langle \mathbf{d}_1(\overline{\mu}^{n,(\lambda)}_{r;(r,s)}(\cdot), \mathcal{L}(X^{(\lambda)}_{r;(r,s)}))\rangle_{\chi\varrho}\, d\lambda\right)\langle |||X(\cdot)|||_{[0,T],w,p}\rangle_{\chi\varrho} \\ &\quad \times \langle w(r,t,\cdot)\rangle^{3/p}_{\chi\varrho} \\ &\quad + cn^{-1/2}\langle |||X(\cdot)|||_{[0,T],w,p}\rangle_{\chi\varrho}\langle w(r,t,\cdot)\rangle^{3/p}_{\chi\varrho} \\ &\le c\varsigma_n\left(1 + \left\langle \sup_{0\le u\le T}|X_u(\cdot)|\right\rangle_{\chi\varrho}\right)\langle |||X(\cdot)|||_{[0,T],w,p}\rangle_{\chi\varrho}\langle\!\langle w^+(r,t,\cdot,\cdot)\rangle\!\rangle^{3/p}_{\chi\varrho}.\end{aligned}$$

Step 2. By the same argument, we have

$$\begin{aligned}&|[G_\mu(\overline{X}^i(\omega), \overline{\mu}^n(\omega))(\overline{X}^j(\omega)) - G_\mu(\overline{X}^i(\omega), \mathcal{L}(X))(\overline{X}^j(\omega))]_{r,s}\mathbb{W}^{j,i}_{s,t}(\omega)| \\ &\le c\left(\int_0^1 \mathbf{d}_1(\overline{\mu}^{n,(\lambda)}_{r;(r,s)}(\omega), \mathcal{L}(X^{(\lambda)}_{r;(r,s)}))\, d\lambda\right)\overline{w}^{j,i}(s,t,\omega)^{2/p} \\ &\quad \times \left[|||\overline{X}^i(\omega)|||_{[0,T],\overline{w}^i,p} + |||\overline{X}^j(\omega)|||_{[0,T],\overline{w}^j,p} + \left(\frac{1}{n}\sum_{k=1}^n |||\overline{X}^k(\omega)|||^2_{[0,T],\overline{w}^k,p}\right)^{1/2}\right] \\ &\quad \times \left[\overline{w}^i(r,s,\omega)^{1/p} + \overline{w}^j(r,s,\omega)^{1/p} + \left(\frac{1}{n}\sum_{k=1}^n \overline{w}^k(r,s,\omega)^{2/p}\right)^{1/2}\right] \\ &\quad + c|\mathcal{S}^{i,j,n}_{r,s}(\omega, |\overline{X}^\bullet_{r,s}(\omega)|)|\overline{w}^{j,i}(s,t,\omega)^{2/p},\end{aligned}$$

where

$$\begin{aligned}\langle \mathcal{S}^{i,j,n}_{r,s}(\cdot, |\overline{X}^\bullet_{r,s}(\cdot)|)\rangle_{3\varrho/2} &\le cn^{-1/2}\langle |||X|||_{[0,T],w,p} w(r,s,\cdot)^{1/p}\rangle_{3\varrho/2} \\ &\le cn^{-1/2}\langle |||X|||_{[0,T],w,p}\rangle_{\chi\varrho}\langle w(r,t,\cdot)\rangle^{1/p}_{\chi\varrho}.\end{aligned}$$

Observing that $\langle \overline{w}^{j,i}(s,t,\cdot)^{2/p}\rangle_{\chi\varrho} \le \langle\!\langle w^+(r,t,\cdot,\cdot)\rangle\!\rangle^{2/p}_{\chi\varrho}$—this is the rationale for introducing w^+, and taking expectation, we get

$$\begin{aligned}&\langle [G_\mu(\overline{X}^i(\cdot), \overline{\mu}^n(\cdot))(\overline{X}^j(\cdot)) - G_\mu(\overline{X}^i(\cdot), \mathcal{L}(X))(\overline{X}^j(\cdot))]_{r,s}\mathbb{W}^{j,i}_{s,t}(\omega)\rangle_\varrho \\ &\le c\left(\int_0^1 \langle \mathbf{d}_1(\overline{\mu}^{n,(\lambda)}_{r;(r,s)}(\cdot), \mathcal{L}(X^{(\lambda)}_{r;(r,s)}))\rangle_{\chi\varrho}\, d\lambda\right)\langle |||X(\cdot)|||_{[0,T],w,p}\rangle_{\chi\varrho}\end{aligned}$$

$$\times \langle\!\langle w^+(r,t,\cdot,\cdot)\rangle\!\rangle_{\chi\varrho}^{3/p}$$
$$+ cn^{-1/2}\langle |\!|\!| X(\cdot) |\!|\!|_{[0,T],w,p}\rangle_{\chi\varrho}\langle\!\langle w^+(r,t,\cdot,\cdot)\rangle\!\rangle_{\chi\varrho}^{3/p}.$$

Taking the mean over j, we obtain as upper bound for the third term in the right-hand side of (A.6) the quantity

$$\Big\langle \frac{1}{n}\sum_{j=1}^{n}\big[G_\mu\big(\overline{X}^i(\cdot),\overline{\mu}^n(\cdot)\big)\big(\overline{X}^j(\cdot)\big) - G_\mu\big(\overline{X}^i(\cdot),\mathcal{L}(X)\big)\big(\overline{X}^j(\cdot)\big)\big]_{r,s}\mathbb{W}_{s,t}^{j,i}(\omega)\Big\rangle_{\varrho}$$
$$\le c\Big(\int_0^1 \langle \mathbf{d}_1\big(\overline{\mu}^{n,(\lambda)}_{r;(r,s)}(\cdot),\mathcal{L}\big(X^{(\lambda)}_{r;(r,s)}\big)\big)\rangle_{\chi\varrho}\,d\lambda\Big)\langle |\!|\!| X(\cdot) |\!|\!|_{[0,T],w,p}\rangle_{\chi\varrho}$$
$$\times \langle\!\langle w^+(r,t,\cdot,\cdot)\rangle\!\rangle_{\chi\varrho}^{3/p}$$
$$+ cn^{-1/2}\langle |\!|\!| X(\cdot) |\!|\!|_{[0,T],w,p}\rangle_{\chi\varrho}\langle\!\langle w^+(r,t,\cdot,\cdot)\rangle\!\rangle_{\chi\varrho}^{3/p}.$$

By Lemma A.4, we get the same bound as in the first step.

Step 3. We now turn to the last term in the right-hand side of (A.6). It reads as the empirical mean of n random variables, $n-1$ of which are conditionally centred and conditionally independent given the realization of the paths $(\overline{X}^i, W^i, \mathbb{W}^i)$, namely

$$\frac{1}{n}\sum_{j=1}^{n}\big[G_\mu\big(\overline{X}^i(\omega),\mathcal{L}(X)\big)\big(\overline{X}^j(\omega)\big)\big]_{r,s}\mathbb{W}_{s,t}^{j,i}(\omega) - \mathbb{E}\big[\delta_\mu \overline{F}^i_{r,s}(\omega,\cdot)\mathbb{W}^{i,\perp}_{s,t}(\cdot,\omega)\big].$$

We can handle the above term by invoking Rosenthal's inequality once again (in a conditional form). To do so, it suffices to compute the $\mathbb{L}^\varrho$ norm of $[G_\mu(\overline{X}^i(\omega),\mathcal{L}(X))(\overline{X}^j(\omega))]_{r,s}\times \mathbb{W}_{s,t}^{j,i}(\omega)$. Doing as before (see (A.7)), it is less than $c\langle |\!|\!|\overline{X}(\cdot)|\!|\!|_{[0,T],w,p}\rangle_{\chi\varrho}\langle\!\langle w^+(r,t,\cdot,\cdot)\rangle\!\rangle_{\chi\varrho}^{3/p}$. So,

$$\Big\langle \frac{1}{n}\sum_{j=1}^{n}\big[G_\mu\big(\overline{X}^i(\omega),\mathcal{L}(X)\big)\big(\overline{X}^j(\omega)\big)\big]_{r,s}\mathbb{W}_{s,t}^{j,i}(\omega) - \mathbb{E}\big[\delta_\mu \overline{F}^i_{r,s}(\omega,\cdot)\mathbb{W}^{i,\perp}_{s,t}(\cdot,\omega)\big]\Big\rangle_{\varrho}$$
$$\le cn^{-1/2}\langle |\!|\!| X(\cdot) |\!|\!|_{[0,T],w,p}\rangle_{\chi\varrho}\langle\!\langle w^+(r,t,\cdot,\cdot)\rangle\!\rangle_{\chi\varrho}^{3/p},$$

which suffices to conclude.

Step 4. We now handle the remainders in (A.6). By expanding (A.5) and by using similar notations for the remainders in the expansion of each $(\overline{X}^j)_{j=1,\dots,n}$, we have (see for instance the proof of [5], Proposition 3.5, and in particular [5], (3.9))

$$\begin{aligned}
R^{F^{i,n}}_{r,s}(\omega) &= \partial_x \mathrm{F}\big(\overline{X}^i_r(\omega),\overline{\mu}^n_r(\omega)\big)R^{\overline{X}^i}_{r,s}(\omega)\\
&\quad + \frac{1}{n}\sum_{j=1}^{n} D_\mu \mathrm{F}\big(\overline{X}^i_r(\omega),\overline{\mu}^n_r(\omega)\big)\big(\overline{X}^j_r(\omega)\big)R^{\overline{X}^j}_{r,s}(\omega)\\
&\quad + \int_0^1 \big[\partial_x \mathrm{F}\big(\overline{X}^{i,(\lambda)}_{r;(r,s)}(\omega),\overline{\mu}^{n,(\lambda)}_{r;(r,s)}(\omega)\big) - \partial_x \mathrm{F}\big(\overline{X}^i_r(\omega),\overline{\mu}^n_r(\omega)\big)\big]\overline{X}^i_{r,s}(\omega)\,d\lambda\\
&\quad + \frac{1}{n}\sum_{j=1}^{n}\int_0^1 \big[D_\mu \mathrm{F}\big(\overline{X}^{i,(\lambda)}_{r;(r,s)}(\omega),\overline{\mu}^{n,(\lambda)}_{r;(r,s)}\big)\big(\overline{X}^{j,(\lambda)}_{r;(r,s)}(\omega)\big)\\
&\quad - D_\mu \mathrm{F}\big(\overline{X}^i_r(\omega),\overline{\mu}^n_r\big)\big(\overline{X}^j_r(\omega)\big)\big]\overline{X}^j_{r,s}(\omega)\,d\lambda.
\end{aligned} \tag{A.8}$$

Expanding $R^{\overline{F}^i}_{r,s}(\omega)$ in a similar way, we have to investigate four difference terms in order to estimate the difference $R^{F^{i,n}}_{r,s}(\omega) - R^{\overline{F}^i}_{r,s}(\omega)$. The first difference term corresponds to the first term in the right-hand side of (A.8)

$$\begin{aligned}&\big|\big[\partial_x \mathrm{F}\big(\overline{X}^i_r(\omega), \overline{\mu}^n_r(\omega)\big) - \partial_x \mathrm{F}\big(\overline{X}^i_r(\omega), \mathcal{L}(X_r)\big)\big] R^{\overline{X}^i}_{r,s}(\omega)\big| \\ &\quad \le c\mathbf{d}_1\big(\overline{\mu}^n_r(\omega), \mathcal{L}(X_r)\big) |\!|\!| \overline{X}^i(\cdot) |\!|\!|_{[0,T],\overline{w}^i,p} \overline{w}^i(r,s,\omega)^{2/p}.\end{aligned}$$

Then, we must recall that, in the first line of the right-hand side in (A.6), the difference $R^{F^{i,n}}_{r,s}(\omega) - R^{\overline{F}^i}_{r,s}(\omega)$ is multiplied by $W^i_{s,t}(\omega)$, which is less than $\overline{w}^i(s,t,\omega)^{1/p}$. In other words, we must multiply both sides in the above inequality by $\overline{w}^i(r,t,\omega)^{1/p}$. By Cauchy Schwarz inequality, the $\mathbb{L}^\varrho$ norm of the resulting bound is less than

$$c\big\langle \mathbf{d}_1\big(\overline{\mu}^n_r(\cdot), \mathcal{L}(X_r)\big)\big\rangle_{\chi\varrho} \big\langle |\!|\!| X(\cdot) |\!|\!|_{[0,T],w,p}\big\rangle_{\chi\varrho} \big\langle w(r,t,\cdot)\big\rangle^{3/p}_{\chi\varrho}.$$

The second difference term that we have to handle corresponds to the second term in the right-hand side of (A.8). With an obvious definition for $R^X(\cdot)$, it reads

$$\begin{aligned}&\bigg| \frac{1}{n}\sum_{j=1}^n D_\mu \mathrm{F}\big(\overline{X}^i_r(\omega), \overline{\mu}^n_r(\omega)\big)\big(\overline{X}^j_r(\omega)\big) R^{\overline{X}^j}_{r,s}(\omega) \\ &\qquad - \big\langle D_\mu \mathrm{F}\big(\overline{X}^i_r(\omega), \mathcal{L}(X_r)\big)\big(X_r(\cdot)\big) R^X_{r,s}(\cdot)\big\rangle \bigg|.\end{aligned}$$

Proceeding exactly as in the first step, the latter is bounded by

$$c\mathbf{d}_1\big(\overline{\mu}^n_r(\omega), \mathcal{L}(X_r)\big)\bigg(\frac{1}{n}\sum_{j=1}^n |R^{\overline{X}^j}_{r,s}(\omega)|\bigg) + c\big|\mathcal{S}^{i,n}_{r,s}\big(\omega, |R^{\overline{X}^\bullet}_{r,s}(\omega)|\big)\big|,$$

where $\mathcal{S}^{i,n}_{r,s}(\omega, |R^{\overline{X}^\bullet}_{r,s}(\omega)|)$ is the n-empirical mean of n variables that are dominated by $(|R^{\overline{X}^j}_{r,s}(\omega)| + \langle R^X_{r,s}(\cdot)\rangle_1)_{j=1,\dots,n}$ and $n-1$ of which are conditionally centred and conditionally independent given the realization of the path $(\overline{X}^i, W^i, \mathbb{W}^i)$. Hence, the $\mathbb{L}^\varrho$ norm of the right-hand side, after multiplication as before by $\overline{w}^i(s,t,\omega)^{1/p}$, is less than

$$c\big(\big\langle \mathbf{d}_1\big(\overline{\mu}^n_r(\cdot), \mathcal{L}(X_r)\big)\big\rangle_{\chi\varrho} + n^{-1/2}\big)\big\langle |\!|\!| X(\cdot) |\!|\!|_{[0,T],w,p}\big\rangle_{\chi\varrho}\big\langle w(r,t,\cdot)\big\rangle^{3/p}_{\chi\varrho}.$$

As for the third term in the right-hand side of (A.8), it fits, up to the additional factor $\overline{X}^i_{r,s}(\omega)$ and for each value of λ, the term studied in the first step. So we get as an upper bound for its $\mathbb{L}^\varrho$ norm, after multiplication by $\overline{w}^i(s,t,\omega)^{1/p}$, the quantity

$$\begin{aligned}&c\bigg(\int_0^1\int_0^1 \big\langle \mathbf{d}_1\big(\overline{\mu}^{n,(\lambda\lambda')}_{r;(r,s)}(\cdot), \mathcal{L}\big(X^{(\lambda\lambda')}_{r;(r,s)}\big)\big)\big\rangle_{\chi\varrho}\, d\lambda\, d\lambda'\bigg) \\ &\quad \times \big\langle |\!|\!| X(\cdot) |\!|\!|_{[0,T],w,p}\big\rangle^2_{\chi\varrho}\big\langle w(r,t,\cdot)\big\rangle^{3/p}_{\chi\varrho} \\ &\quad + cn^{-1/2}\big\langle |\!|\!| X(\cdot) |\!|\!|_{[0,T],w,p}\big\rangle^2_{\chi\varrho}\big\langle w(r,t,\cdot)\big\rangle^{3/p}_{\chi\varrho}.\end{aligned}$$

Following Step 2, we get exactly a similar bound for the fourth term in the right-hand side of (A.8). □

A.3. About law of large numbers.

LEMMA A.4. *There exists a real $\rho_d \geq 1$ such that, for any $\rho \geq \rho_d$ and any probability measure μ on $\mathbb{R}^d$ satisfying $M_\rho(\mu) := (\int_{\mathbb{R}^d} |x|^\rho \mu(dx))^{1/\rho} < \infty$, it holds*

$$\mathbb{E}\big[\mathbf{d}_1(\mu^n(\cdot), \mu)^{\rho/3}\big]^{3/\rho} \leq c_{\rho,d} M_\rho(\mu) \varsigma_n,$$

for a constant $c_{\rho,d}$, only depending on ρ and d, where $\mu^n(\cdot)$ is the empirical distribution of n independent identically distributed random variables and with ς_n as in the introduction of Section A.2.

PROOF. Without any loss of generality, we can assume that $M_\rho(\mu) = 1$, see the argument in [14], Chapter 5, Theorem 5.8. Then Theorem 2 in [26] gives us the following results. For $d \geq 3$, we have (for $\rho \geq 2$)

$$\mathbb{P}\big(\mathbf{d}_1(\mu^n(\cdot), \mu) \geq A\varsigma_n\big) \leq C \exp\big(-cn\varsigma_n^d A^d\big) + Cn(nA\varsigma_n)^{-\rho/2},$$

in which case the result easily follows. When $d = 1$, we have

$$\mathbb{P}\big(\mathbf{d}_1(\mu^n(\cdot), \mu) \geq A\varsigma_n\big) \leq C \exp\big(-cn\varsigma_n^2 A^2\big) + Cn(nA\varsigma_n)^{-\rho/2},$$

and the result follows as well by our choice of ς_n. Finally, when $d = 2$,

$$\mathbb{P}\big(\mathbf{d}_1(\mu^n(\cdot), \mu) \geq A\varsigma_n\big) \leq C \exp\bigg(-\frac{cn\varsigma_n^2 A^2}{(\ln(2 + A^{-1}\varsigma_n^{-1}))^2}\bigg) + Cn(nA\varsigma_n)^{-\rho/2}.$$

Assuming without any loss of generality that $A \geq 1$, we have

$$\ln\big(2 + A^{-1}\varsigma_n^{-1}\big) \leq \ln\big(2 + \varsigma_n^{-1}\big) = \ln(1 + 2\varsigma_n) - \ln(\varsigma_n),$$

which is less than $-2\ln(\varsigma_n)$ for n large enough. Given our choice of ς_n, we have $-\ln(\varsigma_n) = \ln(n)/2 - \ln(\ln(1+n))$, which is less than $\ln(n)/2$. Hence, modifying the value of the constant c, we get, for $A \geq 1$ and for n large enough, independently of the value of A, we get the bound

$$\mathbb{P}\big(\mathbf{d}_1(\mu^n(\cdot), \mu) \geq A\varsigma_n\big) \leq C \exp\bigg(-\frac{cA^2 \ln(1+n)^2}{\ln(n)^2}\bigg) + Cn(nA\varsigma_n)^{-\rho/2},$$

which suffices to complete the proof. □

LEMMA A.5. *Let $(X_n)_{n\geq 1}$ be a collection of independent and identically distributed random variables with values in a Polish space S and let f be a real-valued Borel function on S^2 such that $\mathbb{E}[|f(X_1, X_2)|]$ and $\mathbb{E}[|f(X_1, X_1)|]$ are both finite. Then, with probability 1,*

$$\lim_{n\to\infty} \frac{1}{n^2} \sum_{i,j=1}^n f(X_i, X_j) = \mathbb{E}\big[f(X_1, X_2)\big].$$

PROOF. By the standard version of the law of large numbers, it suffices to prove that, with probability 1,

$$\lim_{n\to\infty} \frac{1}{n^2} \sum_{i,j=1, i\neq j}^n f(X_i, X_j) = \mathbb{E}\big[f(X_1, X_2)\big].$$

Letting $S_n = \sum_{i,j=1,i\neq j}^n f(X_i, X_j)$, for $n \geq 1$, we then define the σ-field $\mathcal{G}_n = \sigma(S_k, k \geq n)$. By independence of the variables $(X_k)_{k\geq 1}$, we have, for any $(i, j) \in \{1, \ldots, n\}^2$ with

$i \neq j$, $\mathbb{E}[f(X_i, X_j)|\mathcal{G}_n] = \mathbb{E}[f(X_i, X_j)|S_n]$. By exchangeability, this is also equal to $\mathbb{E}[f(X_1, X_2)|S_n]$. We get

$$\mathbb{E}\big[f(X_1, X_2)|\mathcal{G}_n\big] = \frac{1}{n^2 - n} \sum_{i,j=1, i\neq j}^{n} \mathbb{E}\big[f(X_i, X_j)|S_n\big] = \frac{S_n}{n^2 - n}.$$

By Lévy's downward theorem and by Kolmogorov's zero-one law, the left-hand side converges almost surely to $\mathbb{E}[f(X_1, X_2)]$. □

Acknowledgments. I. Bailleul thanks the Centre Henri Lebesgue ANR-11-LABX-0020-01 for its stimulating mathematical research programs, and the U.B.O. for their hospitality, part of this work was written there. Partial support from the ANR-16-CE40-0020-01. F. Delarue thanks the Institut Universitaire de France.

REFERENCES

[1] ALBEVERIO, S., KONDRATIEV, Y. G. and RÖCKNER, M. (1996). Differential geometry of Poisson spaces. *C. R. Acad. Sci. Paris Sér. I Math.* **323** 1129–1134. MR1423438

[2] ALBEVERIO, S., KONDRATIEV, YU. G. and RÖCKNER, M. (1998). Analysis and geometry on configuration spaces. *J. Funct. Anal.* **154** 444–500. MR1612725 https://doi.org/10.1006/jfan.1997.3183

[3] AMBROSIO, L., GIGLI, N. and SAVARÉ, G. (2005). *Gradient Flows in Metric Spaces and in the Space of Probability Measures. Lectures in Mathematics ETH Zürich*. Birkhäuser, Basel. MR2129498

[4] BAILLEUL, I. (2015). Flows driven by rough paths. *Rev. Mat. Iberoam.* **31** 901–934. MR3420480 https://doi.org/10.4171/RMI/858

[5] BAILLEUL, I., CATELLIER, R. and DELARUE, F. (2020). Solving mean field rough differential equations. *Electron. J. Probab.* **25** Paper No. 21. MR4073682 https://doi.org/10.1214/19-ejp409

[6] BAILLEUL, I., CATELLIER, R. and DELARUE, F. (2020). Solving mean field rough differential equations. *Electron. J. Probab.* **25** Paper No. 21, 51. MR4073682 https://doi.org/10.1214/19-ejp409

[7] BAILLEUL, I. and RIEDEL, S. (2019). Rough flows. *J. Math. Soc. Japan* **71** 915–978. MR3984248 https://doi.org/10.2969/jmsj/80108010

[8] BARBU, V. and RÖCKNER, M. (2020). From nonlinear Fokker–Planck equations to solutions of distribution dependent SDE. *Ann. Probab.* **48** 1902–1920. MR4124528 https://doi.org/10.1214/19-AOP1410

[9] BILLINGSLEY, P. (1999). *Convergence of Probability Measures*, 2nd ed. *Wiley Series in Probability and Statistics*: *Probability and Statistics*. Wiley, New York. MR1700749 https://doi.org/10.1002/9780470316962

[10] BUDHIRAJA, A., DUPUIS, P. and FISCHER, M. (2012). Large deviation properties of weakly interacting processes via weak convergence methods. *Ann. Probab.* **40** 74–102. MR2917767 https://doi.org/10.1214/10-AOP616

[11] BUDHIRAJA, A. and WU, R. (2017). Moderate deviation principles for weakly interacting particle systems. *Probab. Theory Related Fields* **168** 721–771. MR3663630 https://doi.org/10.1007/s00440-016-0723-3

[12] CARDALIAGUET, P. (2013). Notes on mean field games. https://www.ceremade.dauphine.fr/cardaliaguet/MFG20130420.pdf.

[13] CARMONA, R. and DELARUE, F. (2015). Forward-backward stochastic differential equations and controlled McKean–Vlasov dynamics. *Ann. Probab.* **43** 2647–2700. MR3395471 https://doi.org/10.1214/14-AOP946

[14] CARMONA, R. and DELARUE, F. (2018). *Probabilistic Theory of Mean Field Games with Applications. I*: *Mean Field FBSDEs, Control, and Games*. *Probability Theory and Stochastic Modelling* **83**. Springer, Cham. MR3752669

[15] CARMONA, R. and DELARUE, F. (2018). *Probabilistic Theory of Mean Field Games with Applications. II*: *Mean Field Games with Common Noise and Master Equations*. *Probability Theory and Stochastic Modelling* **84**. Springer, Cham. MR3753660

[16] CASS, T., DOS REIS, G. and SLAKELD, W. (2019). The support of McKean Vlasov equations driven by Brownian motion. Available at arXiv:1911.01992.

[17] CASS, T., LITTERER, C. and LYONS, T. (2013). Integrability and tail estimates for Gaussian rough differential equations. *Ann. Probab.* **41** 3026–3050. MR3112937 https://doi.org/10.1214/12-AOP821

[18] CASS, T. and LYONS, T. (2015). Evolving communities with individual preferences. *Proc. Lond. Math. Soc.* (3) **110** 83–107. MR3299600 https://doi.org/10.1112/plms/pdu040

[19] COGHI, M., DEUSCHEL, J.-D., FRIZ, P. and MAURELLI, M. (2018). Pathwise McKean–Vlasov theory. Available at arXiv:1812.11773.
[20] COGHI, M. and FLANDOLI, F. (2016). Propagation of chaos for interacting particles subject to environmental noise. *Ann. Appl. Probab.* **26** 1407–1442. MR3513594 https://doi.org/10.1214/15-AAP1120
[21] COGHI, M. and GESS, B. (2019). Stochastic nonlinear Fokker–Planck equations. *Nonlinear Anal.* **187** 259–278. MR3954095 https://doi.org/10.1016/j.na.2019.05.003
[22] COGHI, M. and NILSSEN, T. (2019). Rough nonlocal diffusions. Available at arXiv:1905.07270.
[23] DAWSON, D. and VAILLANCOURT, J. (1995). Stochastic McKean–Vlasov equations. *NoDEA Nonlinear Differential Equations Appl.* **2** 199–229. MR1328577 https://doi.org/10.1007/BF01295311
[24] DAWSON, D. A. and GÄRTNER, J. (1987). Large deviations from the McKean–Vlasov limit for weakly interacting diffusions. *Stochastics* **20** 247–308. MR0885876 https://doi.org/10.1080/17442508708833446
[25] DEUSCHEL, J.-D., FRIZ, P. K., MAURELLI, M. and SLOWIK, M. (2018). The enhanced Sanov theorem and propagation of chaos. *Stochastic Process. Appl.* **128** 2228–2269. MR3804792 https://doi.org/10.1016/j.spa.2017.09.010
[26] FOURNIER, N. and GUILLIN, A. (2015). On the rate of convergence in Wasserstein distance of the empirical measure. *Probab. Theory Related Fields* **162** 707–738. MR3383341 https://doi.org/10.1007/s00440-014-0583-7
[27] FRIZ, P. and VICTOIR, N. (2010). Differential equations driven by Gaussian signals. *Ann. Inst. Henri Poincaré Probab. Stat.* **46** 369–413. MR2667703 https://doi.org/10.1214/09-AIHP202
[28] FRIZ, P. K. and HAIRER, M. (2014). *A Course on Rough Paths*: *With an Introduction to Regularity Structures*. *Universitext*. Springer, Cham. MR3289027 https://doi.org/10.1007/978-3-319-08332-2
[29] FRIZ, P. K. and VICTOIR, N. B. (2010). *Multidimensional Stochastic Processes as Rough Paths*: *Theory and Applications*. *Cambridge Studies in Advanced Mathematics* **120**. Cambridge Univ. Press, Cambridge. MR2604669 https://doi.org/10.1017/CBO9780511845079
[30] GANGBO, W. and TUDORASCU, A. (2019). On differentiability in the Wasserstein space and well-posedness for Hamilton-Jacobi equations. *J. Math. Pures Appl.* (9) **125** 119–174. MR3944201 https://doi.org/10.1016/j.matpur.2018.09.003
[31] GÄRTNER, J. (1988). On the McKean-Vlasov limit for interacting diffusions. *Math. Nachr.* **137** 197–248. MR0968996 https://doi.org/10.1002/mana.19881370116
[32] GUBINELLI, M. (2004). Controlling rough paths. *J. Funct. Anal.* **216** 86–140. MR2091358 https://doi.org/10.1016/j.jfa.2004.01.002
[33] JABIN, P.-E. and WANG, Z. (2018). Quantitative estimates of propagation of chaos for stochastic systems with $W^{-1,\infty}$ kernels. *Invent. Math.* **214** 523–591. MR3858403 https://doi.org/10.1007/s00222-018-0808-y
[34] JORDAN, R., KINDERLEHRER, D. and OTTO, F. (1998). The variational formulation of the Fokker–Planck equation. *SIAM J. Math. Anal.* **29** 1–17. MR1617171 https://doi.org/10.1137/S0036141096303359
[35] JOURDAIN, B. and MÉLÉARD, S. (1998). Propagation of chaos and fluctuations for a moderate model with smooth initial data. *Ann. Inst. Henri Poincaré Probab. Stat.* **34** 727–766. MR1653393 https://doi.org/10.1016/S0246-0203(99)80002-8
[36] KAC, M. (1956). Foundations of kinetic theory. In *Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability*, 1954–1955, *Vol. III* 171–197. Univ. California Press, Berkeley and Los Angeles. MR0084985
[37] KOLOKOLTSOV, V. N. and TROEVA, M. (2019). On mean field games with common noise and McKean-Vlasov SPDEs. *Stoch. Anal. Appl.* **37** 522–549. MR3954570 https://doi.org/10.1080/07362994.2019.1592690
[38] LIONS, P.-L. Théorie des jeux à champs moyen et applications. Lectures at the Collège de France, 2007–2008. Available at http://www.college-de-france.fr/default/EN/all/equ_der/cours_et_seminaires.htm.
[39] LOTT, J. (2008). Some geometric calculations on Wasserstein space. *Comm. Math. Phys.* **277** 423–437. MR2358290 https://doi.org/10.1007/s00220-007-0367-3
[40] LYONS, T. J. (1998). Differential equations driven by rough signals. *Rev. Mat. Iberoam.* **14** 215–310. MR1654527 https://doi.org/10.4171/RMI/240
[41] MCKEAN, H. P. JR. (1966). A class of Markov processes associated with nonlinear parabolic equations. *Proc. Natl. Acad. Sci. USA* **56** 1907–1911. MR0221595 https://doi.org/10.1073/pnas.56.6.1907
[42] MÉLÉARD, S. (1996). Asymptotic behaviour of some interacting particle systems; McKean–Vlasov and Boltzmann models. In *Probabilistic Models for Nonlinear Partial Differential Equations* (*Montecatini Terme*, 1995). *Lecture Notes in Math.* **1627** 42–95. Springer, Berlin. MR1431299 https://doi.org/10.1007/BFb0093177
[43] OTTO, F. (2001). The geometry of dissipative evolution equations: The porous medium equation. *Comm. Partial Differential Equations* **26** 101–174. MR1842429 https://doi.org/10.1081/PDE-100002243

[44] OVERBECK, L., RÖCKNER, M. and SCHMULAND, B. (1995). An analytic approach to Fleming–Viot processes with interactive selection. *Ann. Probab.* **23** 1–36. MR1330758

[45] REN, P. and WANG, F.-Y. (2019). Derivative formulas in measure on Riemannian manifold. Available at arXiv:1908.03711.

[46] ROSENTHAL, H. P. (1972). On the span in L^p of sequences of independent random variables. II. In *Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability* (*Univ. California, Berkeley, Calif.*, 1970/1971), *Vol. II: Probability Theory* 149–167. MR0440354

[47] SZNITMAN, A.-S. (1991). Topics in propagation of chaos. In *École D'Été de Probabilités de Saint-Flour XIX—*1989. *Lecture Notes in Math.* **1464** 165–251. Springer, Berlin. MR1108185 https://doi.org/10.1007/BFb0085169

[48] TANAKA, H. (1978/79). Probabilistic treatment of the Boltzmann equation of Maxwellian molecules. *Z. Wahrsch. Verw. Gebiete* **46** 67–105. MR0512334 https://doi.org/10.1007/BF00535689

[49] TANAKA, H. (1984). Limit theorems for certain diffusion processes with interaction. In *Stochastic Analysis* (*Katata/Kyoto*, 1982). *North-Holland Math. Library* **32** 469–488. North-Holland, Amsterdam. MR0780770 https://doi.org/10.1016/S0924-6509(08)70405-7

[50] VILLANI, C. (2003). *Topics in Optimal Transportation*. *Graduate Studies in Mathematics* **58**. Amer. Math. Soc., Providence, RI. MR1964483 https://doi.org/10.1090/gsm/058

[51] WU, C. and ZHANG, J. (2017). An elementary proof for the structure of derivatives in probability measures. Available at arXiv:1705.08046.