A law of large numbers for interacting diffusions via a mild formulation*

Florian Bechtold[†]   Fabio Coppini[‡]

Abstract

Consider a system of n weakly interacting particles driven by independent Brownian motions. In many instances, it is well known that the empirical measure converges to the solution of a partial differential equation, usually called McKean-Vlasov or Fokker-Planck equation, as n tends to infinity. We propose a relatively new approach to show this convergence by directly studying the stochastic partial differential equation that the empirical measure satisfies for each fixed n. Under a suitable control on the noise term, which appears due to the finiteness of the system, we are able to prove that the stochastic perturbation goes to zero, showing that the limiting measure is a solution to the classical McKean-Vlasov equation. In contrast with known results, we do not require any independence or finite moment assumption on the initial condition, but the only weak convergence.

The evolution of the empirical measure is studied in a suitable class of Hilbert spaces where the noise term is controlled using two distinct but complementary techniques: rough paths theory and maximal inequalities for self-normalized processes.

1 Introduction

The theory of weakly interacting particle systems has received great attention in the last fifty years. On the one hand, its mathematical tractability has allowed to obtain a deep understanding of the behavior of the empirical measure for such systems: law of large numbers [32, 5], fluctuations and central limit theorems [36, 12], large deviations [11, 19] and propagation of chaos properties [35] are by now established. On the other hand, the theory of weakly interacting particles enters in several areas of applied

---

†LPSM, Sorbonne Université, France. E-mail: fbechtold@lpsm.paris
‡LPSM, Université de Paris, France. E-mail: fcoppini@lpsm.paris

mathematics such as mean-field games or finance models [4], making it an area of active research.

Depending on the context of application, several results are available. The class of mean-field systems under the name of weakly interacting particles is rather large and models may substantially vary from one another depending on the regularity of the coefficients or the noise. This richness in models is reflected in a variety of different techniques implemented in their study (see e.g. [5, 32, 35] for three very different approaches).

If one focuses on models where the interaction function is regular enough, e.g. bounded and globally Lipschitz, one of the aspects that has not been completely investigated so far, concerns the initial condition. To the authors' knowledge, most of known results require a finite moment condition in order to prove tightness properties of the general sequence (e.g. [19]) or to apply a fixed-point argument in a suitable topological space (e.g. [5]). The only exceptions are given by [35, 36], although they require independent and identically distributed (IID) initial conditions. We want to point out that existence of a solution to the limiting system, a non-linear partial differential equation (PDE) known as Fokker-Planck or McKean-Vlasov equation, does not require any finite moment condition on the initial measure, see e.g. [35, Theorem 1.1]. Furthermore, whenever the particle system is deterministic, there is no need to assume independence (or any finite moment) for this same convergence, see, e.g., [10, 31].

We present a result in the spirit of the law of large numbers, without requiring any assumption on the initial conditions but the convergence of the associated empirical measure. Our main idea consists in exploiting a mild formulation associated to the stochastic partial differential equation satisfied by the empirical measure for a fixed (finite!) population. The main difficulty is giving a meaning to the noise term appearing in such formulation: exploiting the regularizing properties of the semigroup generated by the Laplacian in two different ways, using rough paths theory and maximal inequalities for self normalized processes respectively, we are able to adequately control it. By taking the limit for the size of the population which tends to infinity, the stochastic term vanishes and the limiting measure satisfies the well-known McKean-Vlasov equation.

1.1 Organization

The paper is organized as follows. In the rest of this section we present the model, known results and introduce the set-up in which the evolution of the empirical measure is studied along with notation used.

In Section 2 we give the definition of our notion of solution as well as a corresponding uniqueness statement. The law of large numbers, Theorem 2.3, is presented right after; the section ends with a discussion, a comparison with the existing literature and the strategy of the proof.

The noise perturbation mentioned in the introduction is tackled in Section 3 where rough paths techniques and maximal inequalities for self-normalized processes are exploited. The proof of Theorem 2.3 is given at the end of this section.

Appendix A recalls general properties of analytic semigroups; Appendix B provides an extension of Gubinelli's theory for rough integration to our setting.

1.2 The model and known results

Consider $\left(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geqslant 0}, \mathbb{P}\right)$ a filtered probability space, the filtration satisfying the usual conditions. Fix $d \in \mathbb{N}$, let $(B^i)_{i \in \mathbb{N}}$ be a sequence of IID $\mathbb{R}^d$-valued Brownian motions adapted to the filtration $(\mathcal{F}_t)_{t \geqslant 0}$.

Fix $n \in \mathbb{N}$ and $T > 0$ a finite time horizon. Let $\Gamma : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}^d$ be a bounded

Lipschitz function, and $(x^{i,n})_{1 \leqslant i \leqslant n}$ the unique strong solution to

$$
\begin{cases}
\mathrm{d}x_t^{i,n} = \frac{1}{n}\sum_{j=1}^n \Gamma(x_t^{i,n}, x_t^{j,n})\mathrm{d}t + \mathrm{d}B_t^i, \\
x_0^{i,n} = x_0^i,
\end{cases}
\tag{1.1}
$$

for $t \in [0, T]$ and $i = 1, \ldots, n$. The initial conditions are denoted by the sequence $(x_0^i)_{i \in \mathbb{N}} \subset \mathbb{R}^d$, whenever they are random they are taken independent of the Brownian motions. Existence and uniqueness for (1.1) is a classical result; see, e.g., [34].

The main quantity of interest in system (1.1) is the empirical measure $\nu^n = (\nu_t^n)_{t \in [0,T]}$, a random variable with values on the space of probability measures. It is defined for $t \in [0, T]$ by

$$
\nu_t^n := \frac{1}{n}\sum_{j=1}^n \delta_{x_t^{j,n}}.
\tag{1.2}
$$

Observe that ν^n is apriori a probability measure on the continuous trajectories with values in $\mathbb{R}^d$, i.e. $\nu^n \in \mathcal{P}(C([0,T], \mathbb{R}^d))$, however in many instances we rather consider its projection $(\nu_t^n)_{t \in [0,T]} \in C([0,T], \mathcal{P}(\mathbb{R}^d))$ as continuous function over the probability measures on $\mathbb{R}^d$. This last object does not carry the information of the time dependencies between time marginals, but is in our case more suitable when studying (1.1) in the limit for n which tends to infinity.

1.2.1 Known results

Fix an initial probability measure $\nu_0 \in \mathcal{P}(\mathbb{R}^d)$. Whenever $(x_0^i)_{i \in \mathbb{N}}$ are taken to be IID random variables sampled from ν_0, or ν_0 has a finite moment and ν_0^n weakly converges to it, it is well known (e.g. [35, Theorem 1.4] and [5, Theorem 3.1]) that ν^n converges (in a precise sense depending on the setting) to the solution of the following PDE

$$
\begin{cases}
\partial_t \nu_t = \frac{1}{2}\Delta \nu_t - \mathrm{div}[\nu_t(\Gamma * \nu_t)], \\
\nu|_{t=0} = \nu_0,
\end{cases}
\tag{1.3}
$$

for $t \in [0, T]$ and where $*$ denotes the integration with respect to the second argument, i.e. for $\mu \in \mathcal{P}(\mathbb{R}^d)$

$$
(\Gamma * \mu)(x) = \int_{\mathbb{R}^d} \Gamma(x, y)\,\mu(\mathrm{d}y), \quad x \in \mathbb{R}^d.
$$

Equation (1.3) is usually called McKean-Vlasov or non-linear Fokker-Planck equation.

Remark 1.1. Observe that requiring IID initial conditions is not an innocent assumption as they are, in particular, exchangeable, see [35, §I.2] for more on this perspective. From an applied viewpoint, independence is often a hypothesis that we do not want to assume, see e.g. [9, Example II].

A solution to (1.3) is linked to the following non-linear process:

$$
\begin{cases}
x_t = x_0 + \int_0^t \int_{\mathbb{R}^d} \Gamma(x_s, y)\,\nu_s(\mathrm{d}y)\mathrm{d}s + B_t, \\
\nu_t = Law(x_t),
\end{cases}
\tag{1.4}
$$

where B is a Brownian motion independent of $(B^i)_{i \in \mathbb{N}}$ and x_0. It is well-known that $\nu = (\nu_t)_{t \in [0,T]}$ is a solution to (1.3) if and only if the non-linear process $(x_t)_{t \in [0,T]}$ in (1.4) exists and is such that $Law(x_t) = \nu_t$ for every $t \in [0, T]$.

We have to following theorem.

Theorem 1.2 ([35, Theorem 1.1])**.** *Suppose* Γ *is bounded and Lipschitz and* x_0 *is a random variable with law* $\nu_0 \in \mathcal{P}(\mathbb{R}^d)$. *Then, system* (1.4) *has a unique solution* $(x_t)_{t \in [0,T]}$.

Moreover, if $\nu = (\nu_t)_{t \in [0,T]}$ *is the law of* $(x_t)_{t \in [0,T]}$, *then* $\nu \in C([0,T], \mathcal{P}(\mathbb{R}^d))$ *and it solves the McKean-Vlasov equation* (1.3) *in the weak sense.*

1.3 Set-up and notations

Let $W^{m,p} = W^{m,p}(\mathbb{R}^d)$ be the standard Sobolev space with $m \in \mathbb{N}$ and $p \in [1, \infty)$. Classical results as [1, Theorem 4.12] assure that

$$W_0^{m,p}(\mathbb{R}^d) = W^{m,p}(\mathbb{R}^d) \subset C_b(\mathbb{R}^d) \quad \text{whenever } mp > d, \tag{1.5}$$

where $C_b(\mathbb{R}^d)$ is the space of continuous bounded functions on $\mathbb{R}^d$. The space $W_0^{m,p}(\mathbb{R}^d)$ is the closure of $C_0^\infty(\mathbb{R}^d)$, i.e., the space of smooth functions with compact support, with respect to the norm

$$\|\varphi\|_{W^{m,p}} := \left(\sum_{0 \leqslant |\alpha| \leqslant m} \int_{\mathbb{R}^d} |\partial^\alpha \varphi(x)|^p \, \mathrm{d}x \right)^p, \quad \varphi \in C_0^\infty(\mathbb{R}^d),$$

where $\alpha = (\alpha_1, \ldots, \alpha_d)$ with $|\alpha| = \alpha_1 + \cdots + \alpha_d$ and $\partial^\alpha = (\partial_{x_1})^{\alpha_1} (\partial_{x_2})^{\alpha_2} \ldots (\partial_{x_d})^{\alpha_d}$.

Fix $p = 2$ and $m > d/2$, we consider the Hilbert space $H^m := W^{m,2}(\mathbb{R}^d)$, with norm denoted by $\|\cdot\|_m$ and its dual space $H^{-m} := (H^m)^*$ with the standard dual norm defined by $\|\mu\|_{-m} := \sup_{\|h\|_m \leqslant 1} \langle \mu, h \rangle_{-m,m}$. The action of H^{-m} on H^m is denoted by $\langle \cdot, \cdot \rangle_{-m,m}$. By duality, if follows from (1.5) that

$$\mathcal{P}(\mathbb{R}^d) \subset C_b(\mathbb{R}^d)^* \subset H^{-m}.$$

We denote by $(\cdot, \cdot)_m$ the scalar product in H^m and by $\langle \cdot, \cdot \rangle$ the natural action of a probability measure on test functions, i.e., for $\nu \in \mathcal{P}(\mathbb{R}^d)$ and a smooth function h, we write $\langle \nu, h \rangle = \int_{\mathbb{R}^d} h(x) \nu(\mathrm{d}x)$. We often abuse of notation denoting the density of a probability measure by the probability measure itself.

Let $\nu \in \mathcal{P}(\mathbb{R}^d)$, and thus $\nu \in H^{-m}$, and let $\widetilde{\nu} \in H^m$ be its Riesz representative, then we have for any $h \in H^m$

$$\langle \nu, h \rangle = \nu(h) = (\widetilde{\nu}, h)_m = \langle \nu, h \rangle_{-m,m}$$

and therefore

$$|\langle \nu, h \rangle| \leqslant \|\nu\|_{-m} \|h\|_m.$$

In particular

$$\sup_{\|h\|_m \leqslant 1} \langle \nu, h \rangle = \|\nu\|_{-m}.$$

If $(\mu^n)_{n \in \mathbb{N}}$ is a sequence of probability measures which weakly converges to some $\mu \in \mathcal{P}(\mathbb{R}^d)$, we use the notation $\mu^n \rightharpoonup \mu$. For weak convergence and weak-*-convergence of a sequence $(x_n)_n \subset X$ to some $x \in X$, X being a Banach space, we use the standard notations $x_n \rightharpoonup x$ and $x_n \overset{*}{\rightharpoonup} x$ respectively.

As introduced in [28], we will use $\|\cdot\|_{-m}$ as distance between probability measures and our results will be expressed with respect to this topology.

The various constants in the paper will always be denoted by C or C_α to emphasize the dependence on some parameter α. Their value may change from line to line.

2 Main result

Before stating the main result, we give the definition of weak-mild solutions to (1.3) in the Hilbert space H^m. We denote by $S = (S_t)_{t \in [0,T]}$ the analytic semigroup generated by the Laplacian operator $\frac{\Delta}{2}$ on H^m. We refer to Appendix A for general properties of S.

Definition 2.1 (m-weak-mild solutions to McKean-Vlasov PDEs). *Let ν_0 be an element in H^{-m}. We call $\nu \in L^\infty([0,T], H^{-m})$ an m-weak-mild solution to the problem (1.3), if for every $h \in H^m$ and $t \in [0,T]$, it holds*

$$\langle \nu_t, h \rangle_{-m,m} = \langle \nu_0, S_t h \rangle_{-m,m} + \int_0^t \langle \nu_s, (\nabla S_{t-s} h)(\Gamma * \nu_s) \rangle_{-m,m} \mathrm{d}s. \tag{2.1}$$

If Γ is sufficiently regular, uniqueness can be readily established by using a classical argument. This is illustrated in the next proposition.

Proposition 2.2 (Uniqueness)**.** *Suppose that* $\Gamma(\cdot_x, \cdot_y) \in H_y^m W_x^{m,\infty}$, *i.e.,*

$$\|\Gamma(\cdot_x, \cdot_y)\|_{H_y^m W_x^{m,\infty}} = \max_{|\beta| \leqslant m} \left\| \sum_{|\alpha| \leqslant m} \int_{\mathbb{R}^d} \left(\partial_x^\beta \partial_y^\alpha \Gamma(x,y) \right)^2 \mathrm{d}y \right\|_{L_x^\infty} < \infty. \qquad (2.2)$$

Then, m-*weak mild solutions to* (1.3) *are unique.*

Proof. Suppose $\nu, \rho \in L^\infty([0,T], H^{-m})$ are two m-weak mild solutions. Then, taking the difference between the two equations (2.1), one obtains that for every $h \in H^m$

$$\langle \nu_t - \rho_t, h \rangle_{-m,m} = \int_0^t \langle \nu_s - \rho_s, (\nabla S_{t-s} h)(\Gamma * \nu_s) \rangle_{-m,m} \mathrm{d}s$$
$$+ \int_0^t \langle \rho_s, (\nabla S_{t-s} h)(\Gamma * (\nu_s - \rho_s)) \rangle_{-m,m} \mathrm{d}s.$$

In particular,

$$\|\nu_t - \rho_t\|_{-m} \leqslant \int_0^t \|\nu_s - \rho_s\|_{-m} \sup_{\|h\|_m \leqslant 1} \|(\nabla S_{t-s} h)(\Gamma * \nu_s)\|_m \, \mathrm{d}s$$
$$+ \int_0^t \|\rho_s\|_{-m} \sup_{\|h\|_m \leqslant 1} \|(\nabla S_{t-s} h)(\Gamma * (\nu_s - \rho_s))\|_m \, \mathrm{d}s.$$

Observe that, for $\mu \in H^{-m}$ it holds that

$$\|(\nabla S_{t-s} h)(\Gamma * \mu)\|_m \leqslant \|\nabla S_{t-s} h\|_m \|\Gamma * \mu\|_{W^{m,\infty}}$$
$$\leqslant \frac{C}{\sqrt{t-s}} \|h\|_m \|\mu\|_{-m} \|\Gamma(\cdot_x, \cdot_y)\|_{H_y^m W_x^{m,\infty}}, \qquad (2.3)$$

where we have used the properties of the semigroup (see Lemma A.1). Using the continuous embedding of $\mathcal{P}(\mathbb{R}^d)$ into H^{-m}, we conclude that there exists a (new) constant $C > 0$:

$$\|\nu_t - \rho_t\|_{-m} \leqslant C \|\Gamma(\cdot_x, \cdot_y)\|_{H_y^m W_x^{m,\infty}} \int_0^t \frac{1}{\sqrt{t-s}} \|\nu_s - \rho_s\|_{-m} \, \mathrm{d}s.$$

A Gronwall-like lemma yields the proof. $\qquad \square$

We are ready to state the main result.

Theorem 2.3. *Assume* $m > d/2 + 3$ *and* $\Gamma(\cdot_x, \cdot_y) \in H_y^m W_x^{m,\infty}$. *If* $\nu_0 \in H^{-m}$, *then there exists* $\nu \in L^\infty([0,T], H^{-m})$, *unique* m-*weak-mild solution to* (2.1). *Suppose that the initial empirical measure associated to the particle system* (1.1) *is such that*

$$\nu_0^n \rightharpoonup \nu_0 \qquad \text{in } H^{-m}$$

in probability. Then, the empirical measure ν^n *of* (1.1) *satisfies*

$$\nu^n \overset{*}{\rightharpoonup} \nu \qquad \text{in } L^\infty([0,T], H^{-m})$$

in probability.

Moreover, if $\nu_0 \in \mathcal{P}(\mathbb{R}^d)$, *then* ν *is the unique weak solution of the McKean-Vlasov equation* (1.3) *and, in particular,* $\nu \in C([0,T], \mathcal{P}(\mathbb{R}^d))$.

2.1 Discussion

Theorem 2.3 shows a law of large numbers in $L^\infty([0, T], H^{-m})$ by directly studying the evolution of the empirical measure. Contrary to most of the existing proofs in the literature, it does not establish any trajectorial estimates on system (1.1) and does not invoke propagation of chaos techniques, as, e.g., in [29, 35]. This allows to deal with very general initial data: the weak convergence of $(\nu_0^n)_{n \in \mathbb{N}}$ in H^{-m} – which is implied by the weak convergence in $\mathcal{P}(\mathbb{R}^d)$ – suffices.

Working in H^{-m} for $m > d/2$ assures a bound on $\|\nu\|_{-m}$ which is uniform in $\nu \in \mathcal{P}(\mathbb{R}^d)$ thanks to the continuous embedding of $\mathcal{P}(\mathbb{R}^d)$ in H^{-m} and the duality properties of probability measures, see Lemma A.3. By exploiting the equation satisfied by ν^n, we are able to establish a compactness property for $(\nu^n)_{n \in \mathbb{N}}$, usually hard to obtain in $\mathcal{P}(\mathbb{R}^d)$, and which represents our main tool for obtaining the existence both of the limit solution and of a convergent subsequence.

Weak-mild solutions make sense for any $m > d/2$, yet we have to require the stronger condition $m > d/2 + 3$ in order to give a pathwise meaning to the stochastic term present in the dynamics. This implies that Γ is C^3. In this last case, it is already known that a weak solution to the McKean-Vlasov equation (1.3) exists for any initial probability measure ν_0. Since weak solutions are weak-mild solutions, as we will show in the sequel, a byproduct of our main result is the uniqueness of (weak) solutions to equation (1.3).

The particle system (1.1) represents an interaction setting where no transport is present in the dynamics. We have decided not to include other terms so as to keep the underlying ideas and techniques as clear as possible. However, all our arguments readily extend to the more general case of interacting particles given by

$$\mathrm{d}x_t^{i,n} = F(x_t^{i,n})\mathrm{d}t + \frac{1}{n}\sum_{j=1}^n \Gamma(x_t^{i,n}, x_t^{j,n})\mathrm{d}t + \mathrm{d}B_t^i, \tag{2.4}$$

provided that $F \in H^m$.

Finally, we point out that the need of rather high regularity in Γ (and F) is an intrinsic requirement of rough paths theory and not of the particular class of models we are working with. In particular, proving Theorem 2.3 independently of rough paths arguments would likely yield less restrictive regularity constraints on Γ. On the other hand, rough paths theory allows to give a pathwise definition of the stochastic partial differential equation satisfied by the empirical measure. Such viewpoint appears to be new in the literature. Finally, we observe that the proposed strategy represents an application to the algebraic integration with respect to semigroups, presented in [18], that can be interesting on its own.

2.2 Comparison with the existing literature

Proving a law of large numbers by directly studying the empirical measure and not the single trajectories is the classical approach in the deterministic setting [31, 10], i.e., when no Brownian motions are acting on system (1.1). In the case of interacting diffusions, the idea of studying the equation satisfied by the empirical measure for a fixed n, comes from the two articles [3, 26] and the recent [6], where a weak-mild formulation is derived and carefully studied. Contrary to our case, in [3, 6, 26] the particles live in the one dimensional torus which considerably simplifies the analysis; we refer to Remark 3.5.

A Hilbertian approach for particle systems has already been discussed in [12], where it is used to study the fluctuations of the empirical measure around the McKean-Vlasov limit. However, [12] does not make use of the theory of semigroups but instead requires strong hypothesis on the initial conditions which have to be IID and with finite $(4d + 1)$-

moment (see [12, §3]). The evolution of the empirical measure (1.2) is then studied in weighted Hilbert spaces (or, more precisely, in spaces of Bessel potentials) so as to fully exploit the properties of mass concentration given by the condition on the moments. Observe that we are not able to present a fluctuation result, given the lack of a suitable uniform estimate on the noise term.

Studying the action of an analytic semigroup in the evolution of an interacting particle system has been recently proposed in similar settings; we refer to [13, 14] and references therein. This method is referred to as the semigroup approach. We want to stress that the cited works deal with smooth mollified empirical measures and work in a weaker topology (with respect to the time variable) than the one expressed in Theorem 2.3.

The strategies developed in [30, 29, 23], and further applied in the case of mean-field games in [4], study the evolution of the joint law of system (1.1) and take a more abstract viewpoint. In particular, they study the system dynamics at the level of the flows and not directly addressing the empirical measure.

Finally, observe that under a suitable change of the time-scale, the n-dependent SPDE satisfied by the empirical measure (1.2) is the mild formulation of the Dean-Kawasaki equation [25, Theorem 1] and [24].

2.3 Strategy of the proof

Using Itô's formula, we derive an equation satisfied by ν^n for every fixed $n \in \mathbb{N}$, which turns out to be the McKean-Vlasov PDE perturbed by some noise w^n, see Lemma 3.1. This equation makes sense in $L^\infty([0,T], H^{-m})$ and in this space we study the convergence of $(\nu^n)_{n \in \mathbb{N}}$.

The main challenge towards the proof of Theorem 2.3 is giving a meaning to w^n and suitably controlling it. In Lemma 3.2, we first give a pathwise definition of this term through rough paths theory, referring to Appendix B for a suitable theory of rough integration in our setting. This in turn will allow to show that $(\nu^n)_{n \in \mathbb{N}}$ is uniformly bounded in $L^\infty([0,T], H^{-m})$ and to extract a weak-* converging subsequence, see Lemma 3.8.

To show that a converging subsequence satisfies the weak-mild formulation (2.1) in the limit, as shown in Lemma 3.9, we need a further step: the pointwise estimate of $w^n(h)$, for a fixed $h \in H^m$. Using a suitable decomposition of the semigroup and a maximal inequality for self-normalized processes, we are able to prove that $w^n(h)$ converges to zero in probability as n diverges, see Lemma 3.6. If on the one hand the rough paths bound cannot take advantage of the statistical independence of the Brownian motions and thus, cannot be improved in n; on the other hand the probability estimate does not suffice to define w^n as an element of $L^\infty([0,T], H^{-m})$. We refer to Subsection 3.2 and Remark 3.7 for more on this aspect.

The uniqueness of weak-mild solution, Proposition 2.2, is the last ingredient to obtain that any convergent subsequence of $(\nu^n)_{n \in \mathbb{N}}$ admits a further subsequence that converges $\mathbb{P}$-a.s. to the same ν satisfying equation (2.1). This is equivalent to the weak-* convergence in probability to the weak-mild solution ν.

3 Proofs

We start by giving the n-dependent stochastic equation satisfied by the empirical measure for each $n \in \mathbb{N}$. We then move to the control on the noise term and, finally, the proof of Theorem 2.3.

3.1 A weak-mild formulation satisfied by the empirical measure

Recall that $(S_t)_{t \in [0,T]}$ denotes the semigroup generated by $\frac{\Delta}{2}$ on H^m.

Lemma 3.1. *Assume $m > d/2 + 2$. The empirical measure (1.2) associated to the particle systems (1.1) satisfies for every $h \in H^m$ and $t \in [0, T]$*

$$\langle \nu_t^n, h \rangle_{-m,m} = \langle \nu_0^n, S_t h \rangle_{-m,m} + \int_0^t \langle \nu_s^n, (\nabla S_{t-s} h)(\Gamma * \nu_s^n) \rangle_{-m,m} \mathrm{d}s + w_t^n(h), \quad \mathbb{P}\text{-a.s.,} \quad (3.1)$$

where

$$w_t^n(h) = \frac{1}{n} \sum_{j=1}^n \int_0^t \left[\nabla S_{t-s} h \right] (x_s^{j,n}) \cdot \mathrm{d}B_s^j. \qquad (3.2)$$

Proof. Fix $t \in [0, T]$ and $h \in H^m$, by (1.5) h is $C^2(\mathbb{R}^d)$. For $s < t$, applying Itô's formula onto the test function $\varphi(x, s) = (S_{t-s} h)(x)$, we obtain

$$h(x_t^{i,n}) = (S_t h)(x_0^{i,n}) + \frac{1}{n} \sum_{j=1}^n \int_0^t (\nabla S_{t-s} h)(x_s^{i,n}) \Gamma(x_s^{i,n}, x_s^{j,n}) \mathrm{d}s$$

$$+ \int_0^t (\nabla S_{t-s} h)(x_s^{j,n}) \cdot \mathrm{d}B_s^j.$$

Summing over all particles and dividing by $1/n$, the claim is proved modulo well-posedness of the noise term w^n which is presented in the following subsection. $\qquad \square$

3.2 Controlling the noise term

The aim of this subsection is to control the noise term w^n appearing in the weak-mild formulation (3.1) for the empirical measure. We start by giving a pathwise definition of the integral (3.2), i.e. for any $\omega \in A \subset \Omega$ where $\mathbb{P}(A) = 1$ and any $h \in H^m$ we define

$$w_t^n(h)(\omega) = \left(\frac{1}{n} \sum_{j=1}^n \int_0^t \left[\nabla S_{t-s} h \right] (x_s^{j,n}) \cdot \mathrm{d}B_s^j \right)(\omega),$$

which in turn allows to define w^n as an element of $L^\infty([0, T], H^{-m})$, via an inequality of the form

$$\sup_{\|h\|_m = 1} |w_t^n(h)(\omega)| \leqslant C_T(\omega)$$

for $\omega \in A$, see Lemma 3.2. For this purpose, we extend Gubinelli's theory for rough integration (see [17] and [18, §3 and 4]) to our setting, see Appendix B for notations and precise results on this extension.

A probabilistic estimate is then given, exploiting the independence of the Brownian motions; Lemma 3.6 shows that

$$\mathbb{E} \left[\sup_{t \in [0, T]} |w_t^n(h)|^2 \right] \leqslant \frac{C}{n} \|h\|_m^2, \quad h \in H^m.$$

This estimate will allow us to prove the convergence of (3.1) to (2.1) for every fixed $h \in H^m$, see Lemma 3.9.

3.2.1 Pathwise definition via rough paths theory for semigroup functionals

We start by observing that the noise term $w_t^n(h)$ in (3.2) is neither a stochastic convolution that could be treated using a maximal inequality in Hilbert spaces (e.g. [7, §6.4] and [2] in the context of an unbounded diffusion operator), nor a classical controlled rough path integral (e.g. [15]) as the integrand depends on the upper integration limit.

We combine the strategies in [18, 17] so to define $w_t^n(h)$ in a pathwise sense. Note that our setting is different from [18], where an infinite dimensional theory à la Da Prato-Zabczyk is constructed, while we are interested in finite dimensional stochastic integrals over functionals of such objects. Our construction is nonetheless similar to [18]: we fix the Itô-rough path lift associated to Brownian motion and extend the algebraic integration in [18] to our setting of semigroup functionals. This extension is presented in detail in Appendix B, where the main ingredient, the Sewing lemma, is proven. Before stating Lemma 3.2, we present in a heuristic fashion the main ideas towards a rough path construction of (3.2).

Note that it suffices to define integrals of the form

$$\int_s^t \nabla S_{t-u} f(x_u) \cdot \mathrm{d}B_u \qquad (3.3)$$

in a pathwise sense for a class of sufficiently regular functions f and where $(x_u)_u$ is an $\mathbb{R}^d$-valued process controlled by the Brownian motion $(B_u)_u$, such that

$$x_t - x_s = B_t - B_s + O(|t - s|), \quad \text{for } s, t \in [0, T], \ \mathbb{P}\text{-a.s..} \qquad (3.4)$$

Recall that in the classical setting of rough paths theory, one has for $s \leqslant t$

$$\int_s^t f(x_u) \mathrm{d}B_u = f(x_s) B_{ts} + (D_x f)(x_u) \mathbb{B}_{ts} + R_{ts}$$

where we have used the notation $B_{ts} := B_t - B_s$ as well as

$$\mathbb{B}_{ts} := \int_s^t B_{us} \otimes \mathrm{d}B_u, \quad t \geqslant s \in [0, T].$$

In particular, $A_{ts} := f(x_s) B_{ts} + (D_x f)(x_u) \mathbb{B}_{ts}$ is a germ and, thanks to (3.4), $R_{ts} = o(|t-s|)$ is a remainder in the terminology of [17]. In the same spirit of [17], we rewrite the left hand side of this expression as

$$\int_s^t f(x_u) \mathrm{d}B_u = [\delta I]_{ts} = I_t - I_s$$

where

$$I_t = \int_0^t f(x_u) \mathrm{d}B_u.$$

We are thus left with

$$[\delta I]_{ts} = A_{ts} + R_{ts}. \qquad (3.5)$$

Recall that Gubinelli's Sewing Lemma formulates precise conditions under which a given germ A gives rise to a unique remainder term $R_{ts} = o(|t - s|)$ and such that I can be obtained as

$$I_t := \lim_{|\mathcal{P}[0,t]| \downarrow 0} \sum_{[u,v] \in \mathcal{P}[0,t]} A_{vu}.$$

If one tries to follow a similar approach for the quantity of interest (3.3), a canonical candidate for local approximations to (3.3) would be

$$\int_s^t (\nabla S_{t-u} f)(x_u) \mathrm{d}B_u = (\nabla S_{t-s} f)(x_s) B_{ts} + (D \nabla S_{t-s} f)(x_s) \mathbb{B}_{ts} + R_{ts}.$$

However, notice that if we were to set

$$I_t(f) := \int_0^t (\nabla S_{t-u} f)(x_u) \mathrm{d}B_u$$

then, we would obtain

$$
\begin{aligned}
[\delta I(f)]_{ts} &= I_t(f) - I_s(f) \\
&= \int_0^t (\nabla S_{t-u}f)(x_u)\mathrm{d}B_u + \int_0^s (\nabla S_{s-u}(S_{t-s} - \mathrm{Id})f)(x_u)\mathrm{d}B_u - \int_0^s (\nabla S_{t-u}f)(x_u)\mathrm{d}B_u \\
&= \int_s^t (\nabla S_{t-u}f)(x_u)\mathrm{d}B_u + I_s((S_{t-s} - \mathrm{Id})f) \\
&\neq \int_s^t (\nabla S_{t-u}f)(x_u)\mathrm{d}B_u,
\end{aligned}
$$

in contrast to the above setting, meaning the standard approach of [17] fails. If one defines, following Gubinelli and Tindel [18, p.16], the operator ϕ via

$$
[\phi I(f)]_{ts} = I_s((S_{t-s} - \mathrm{Id})f)
$$

as well as the operator $\hat{\delta}$ via

$$
[\hat{\delta} I(f)]_{ts} = [\delta I(f)]_{ts} - [\phi I(f)]_{ts},
$$

the desired relationship is recovered, indeed

$$
\begin{aligned}
[\hat{\delta} I(f)]_{ts} &= \int_s^t (\nabla S_{t-u}f)(x_u)\mathrm{d}B_u \\
&= (\nabla S_{t-s}f)(x_s)B_{ts} + (D\nabla S_{t-s}f)(x_s)\mathbb{B}_{ts} + R_{ts}.
\end{aligned}
$$

The idea is hence to change the cochain complex in [17] and to consider a perturbed version of it associated to the operator $\hat{\delta}$, this is done in Lemma B.1. Lemma B.2 proves a Sewing Lemma in this modified setting, which in turn allows to construct the above remainder R_{ts}. The germ will therefore be

$$
[Af]_{ts} = (\nabla S_{t-s}f)(x_s)B_{ts} + (D\nabla S_{t-s}f)(x_s)\mathbb{B}_{ts}.
$$

For $0 = t_0 < \cdots < t_{n+1} = t$, note that due to

$$
\begin{aligned}
I_t(f) &= \int_0^t (\nabla S_{t-u}f)(x_u)\mathrm{d}B_u \\
&= \sum_{k=0}^n \int_{t_k}^{t_{k+1}} (\nabla S_{t-u}f)(x_u)\mathrm{d}B_u \\
&= \sum_{k=0}^n \int_{t_k}^{t_{k+1}} (\nabla S_{t_{k+1}-u}(S_{t-t_{k+1}}f))(x_u)\mathrm{d}B_u \\
&= \sum_{k=0}^n [A(S_{t-t_{k+1}}f)]_{t_{k+1}t_k} + \sum_{k=0}^n R_{t_{k+1}t_k},
\end{aligned}
$$

the correct way of sewing together the germs is given by

$$
I_t(f) = \lim_{n \to \infty} \sum_{k=0}^n [A(S_{t-t_{k+1}}f)]_{t_{k+1}t_k},
$$

which is reflected in equation (B.3) in Corollary B.3. In particular, note that this Corollary comes with the stability estimate (B.2) which allows to eventually deduce the first crucial estimate (3.6) on the noise term, as shown in the next Lemma.

Lemma 3.2. *Suppose $m > d/2 + 3$. For every $\alpha \in (1/3, 1/2)$, there exists a positive random constant $C = C_\alpha$ that is finite $\mathbb{P}$-a.s.(and of finite moments for all orders) such that $\mathbb{P}$-a.s.*

$$|w_t^n(h)| \leqslant C_\alpha (1+t)^{3\alpha} \|h\|_m \tag{3.6}$$

for any $t \geqslant 0$ and $h \in H^m$.

Proof. We follow the notations of Appendix B. Fix $\alpha \in (1/3, 1/2)$ and recall that $(B, \mathbb{B})$ is the Itô rough path lift, with

$$\mathbb{B}_{ts} := \int_s^t B_{us} \otimes \mathrm{d}B_u, \quad s \leqslant t \in [0, T],$$

where $B_{us} := B_u - B_s$. Note that the above stochastic integral is understood in the Itô sense.

We use Lemma B.3 to define the Itô integral (3.3). This in turn will imply the well-posedness of $w_t^n(h)$ with the choice $f = h$, $x = x^{i,n}$ and $B = B^i$ for $i = 1, \dots, n$. Indeed, $x^{i,n}$ is controlled by B^i (recall (1.1) and the fact that Γ is bounded), and thus

$$|w_t^n(h)| \leqslant \frac{1}{n} \sum_{i=1}^n \left| \int_0^t (\nabla S_{t-s} h)(x_s^{i,n}) \mathrm{d}B_s^i \right| \leqslant C_\alpha (1+t)^{3\alpha} \|h\|_m.$$

Define the operator A acting on $f \in H^m$ into $C(\Delta_2, \mathbb{R})$ via

$$[Af]_{ts} := (\nabla S_{t-s} f)(x_s) \cdot B_{ts} + (D_x \nabla S_{t-s} f)(x_s) \cdot \mathbb{B}_{ts},$$

where D_x denotes the Jacobian in $\mathbb{R}^d$ and $\cdot$ the scalar product between tensors of the same dimension. In the sequel, we adopt the following shorter notation

$$[Af]_{ts} := \nabla S_{ts} f_s B_{ts} + D_x \nabla S_{ts} f_s \mathbb{B}_{ts}.$$

As in classical rough paths theory $[Af]_{ts}$ is not a 1-increment (i.e. a difference as B_{ts}) but a continuous function of the two variables s and t. In particular $A \in D_2$, i.e. A is a linear operator from the Banach space H^m to C_2.

One can actually prove that $A \in D_2^\alpha$: for $0 \leqslant s \leqslant t \leqslant T$ and $f \in H^m$

$$|[Af]_{ts}| \leqslant \|\nabla S_{ts} f\|_\infty |B_{ts}| + \|D_x \nabla S_{ts} f\|_\infty |\mathbb{B}_{ts}| \leqslant C_\alpha \|f\|_m |t - s|^\alpha,$$

where $C_\alpha = C_\alpha(\omega)$ depends on the α-Hölder norm of $B(\omega)$ and $\mathbb{B}(\omega)$ and we have used the properties of S, see Lemma A.2. Note in particular that $C_\alpha < \infty$, $\mathbb{P}$-a.s. and that C_α has finite moments of all orders.

Recall the definition of $\hat{\delta}$ (Lemma B.1), in order to apply Lemma B.2 and Corollary B.3 we need to show that $\hat{\delta} A \in D_3^{1+}$. Let $f \in H^m$ and $s < u < t$, one has

$$[\hat{\delta} Af]_{tus} = [\delta Af]_{tus} - [\phi Af]_{tus} = [Af]_{ts} - [Af]_{tu} - [Af]_{us} - [A(S_{t\cdot} - \mathrm{Id})f]_{us}$$
$$= [Af]_{ts} - [Af]_{tu} - [AS_{t\cdot} f]_{us}.$$

Observe that thanks to the properties of the semigroup

$$[AS_{t\cdot} f]_{us} = \nabla S_{us} S_{tu} f_s B_{us} + D_x \nabla S_{us} S_{tu} f_s \mathbb{B}_{us} = \nabla S_{ts} f_s B_{us} + D_x \nabla S_{ts} f_s \mathbb{B}_{us}.$$

In particular, using Chen's relation

$$\mathbb{B}_{ts} = \mathbb{B}_{us} + \mathbb{B}_{tu} + B_{tu} \otimes B_{us}$$

we obtain

$$[\hat{\delta}Af]_{tus} = \nabla S_{ts}f_s\, B_{tu} - \nabla S_{tu}f_u\, B_{tu} + D_x\nabla S_{ts}f_s\,(\mathbb{B}_{ts} - \mathbb{B}_{us}) - D_x\nabla S_{tu}f_u\, \mathbb{B}_{tu}$$
$$= (\nabla S_{ts}f_s - \nabla S_{tu}f_u)\, B_{tu} + D_x(\nabla S_{ts}f_s - \nabla S_{tu}f_u)\, \mathbb{B}_{tu} + D_x S_{ts}f_s\, B_{tu} \otimes B_{us}.$$

We rewrite everything as the sum of four terms

$$[\hat{\delta}Af]_{tus} = \nabla(S_{ts} - S_{tu})f_u\, B_{tu} + D_x\nabla(S_{ts} - S_{tu})f_u\, \mathbb{B}_{tu}$$
$$+ (D_x\nabla S_{ts}f_s - D_x\nabla S_{ts}f_u)\, \mathbb{B}_{tu} + (\nabla S_{ts}f_s - \nabla S_{ts}f_u + D_x\nabla S_{ts}f_s B_{us})\, B_{tu}$$
$$=: A^1 + A^2 + A^3 + A^4.$$

For A^1 we obtain

$$|\nabla(S_{ts} - S_{tu})f_u\, B_{tu}| \leqslant \|\nabla(S_{ts} - S_{tu})f_u\|_\infty\, |B_{tu}| \leqslant C_\alpha\, \|f\|_m\, |t - u|^\alpha\, |u - s|,$$

where $C_\alpha = C_\alpha(\omega)$ depends on the α-Hölder norm of $B(\omega)$ and we have used the properties of S, see Lemma A.2. Note in particular that $C_\alpha < \infty$, $\mathbb{P}$-a.s. and that C_α has finite moments of all orders. Similarly, for A^2 (with a different C_α)

$$|D_x\nabla(S_{ts} - S_{tu})f_u\, \mathbb{B}_{tu}| \leqslant C_\alpha\, \|f\|_m\, |t - u|^{2\alpha}\, |u - s|^{1/2}.$$

Observe now that, since $f \in C_b^3$, the function $D_x\nabla S_{ts}f$ is Lipschitz uniformly in s and t, from which we extract that

$$|(D_x\nabla S_{ts}f_s - D_x\nabla S_{ts}f_u)\, \mathbb{B}_{tu}| \leqslant C_\alpha\, \|f\|_m\, |x_s - x_u|\, |t - u|^{2\alpha}$$
$$\leqslant C_\alpha\, \|f\|_m\, |t - u|^{2\alpha}\, |u - s|^\alpha.$$

Using (3.4), we recognize in A^4 the Taylor expansion of $\nabla S_{ts}f$ around x_s, i.e.

$$|\nabla S_{ts}f_u - \nabla S_{ts}f_u - D_x\nabla S_{ts}f_s\, B_{us}|$$
$$\leqslant |\nabla S_{ts}f_u - \nabla S_{ts}f_u - D_x\nabla S_{ts}f_s\, x_{us}| + c\, |D_x\nabla S_{tu}f_s|\, |u - s|$$
$$\leqslant c\, \|f\|_m\, |x_{us}|^2 + c\, \|f\|_m\, |u - s| \leqslant c\, \|f\|_m\, |u - s|^{2\alpha}.$$

We conclude that

$$\left|A^4\right| \leqslant C_\alpha\, \|f\|_m\, |t - u|^\alpha\, |u - s|^{2\alpha}.$$

Putting the four estimates together, we have just shown $\hat{\delta}A \in D_3^{1+}$ and, in particular, that

$$\left\|\hat{\delta}A\right\|_{D_3^{3\alpha}} \leqslant C_\alpha$$

for some C_α which is finite $\mathbb{P}$-a.s. and admits moments of all orders. By Corollary B.3, we know that there exists $I \in D_1$ such that

$$[\hat{\delta}If]_{ts} = \lim_{|\mathcal{P}^n[s,t]| \to 0} \sum_{[u,v] \in \mathcal{P}^n[s,t]} [AS_{t\cdot}f]_{vu}$$

is well defined. For $0 \leqslant s \leqslant t \leqslant T$, se set

$$\int_s^t \nabla S_{t-u}f(x_u) \cdot \mathrm{d}B_u := [\hat{\delta}If]_{ts}.$$

Again Corollary B.3 assures that there exists a (new) constant C_α, depending on the norm of A in D_2^α and the norm of $\hat{\delta}A$ in $D_3^{3\alpha}$, such that

$$\left|\int_0^t \nabla S_{t-u}f(x_u) \cdot \mathrm{d}B_u\right| \leqslant C_\alpha\, \|f\|_m\, (1 + t)^{3\alpha}.$$

The proof is concluded. $\qquad\square$

3.2.2 Controlling $w_t^n(h)$ **via a maximal inequality for self-normalized processes**

The aim of this subsection is to give a probabilistic bound on

$$w_t^n(h) = \frac{1}{n} \sum_{j=1}^{n} \int_0^t [\nabla S_{t-s} h] (x_s^{j,n}) \mathrm{d}B_s^j$$

by exploiting the independence of the Brownian motions (we have removed the product symbol $\cdot$ for the sake of notation).

Observe that if $w_t^n(h)$ didn't involve a convolution with the semigroup S, $w_t^n(h)$ would be a standard martingale and classical estimates like the Burkholder-Davis-Gundy inequality could be used to establish the desired bound. While the convolution with the semigroup S destroys the martingale property, $w_t^n(h)$ is still closely related to maximal inequalities for self-normalized martingales for which the following fine estimate due to Graversen and Peskir [16] is available.

Lemma 3.3 ([16, Corollary 2.8] and [22, Corollary 2.4])**.** *Let* $(M_t)_{t \in [0,T]}$ *be a continuous local martingale. There exists a universal constant* C *such that*

$$\mathbb{E}\left[\sup_{t \in [0,\tau]} \frac{|M_t|^2}{1 + \langle M \rangle_t} \right] \leqslant C \, \mathbb{E}\left[\log(1 + \log(1 + \langle M \rangle_\tau)) \right]$$

for every stopping time $\tau \leqslant T$.

Observe that this result is a consequence of more general bounds on self-normalized processes of the form $X_t = A_t / B_t$ (e.g. [8]), where in this case $A_t = M_t$ is a martingale and $B_t^2 - 1 = \langle M \rangle_t$ its quadratic variation.

Let us illustrate in the following example how this interpretation can be used to directly obtain a bound on

$$v_t = \frac{1}{n} \sum_{j=1}^{n} \int_0^t e^{-a(t-s)} \mathrm{d}B_s^j, \quad a > 0,$$

which could be seen as a most simple toy model for $w_t^n(h)$.

Example 3.4. Let $(B^j)_{j \leqslant n}$ be independent Brownian motions on a common filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F})_t, \mathbb{P})$. For $a > 0$, let $(X^j)_{j \leqslant n}$ be the following associated familiy of Ornstein Uhlenbeck processes:

$$X_t^j := \int_0^t e^{-a(t-s)} \mathrm{d}B_s^j, \quad t \in [0, T]$$

and consider the quantity

$$v_t := \frac{1}{n} \sum_{j=1}^{n} X_t^j.$$

We remark that we may rewrite

$$\sum_{j=1}^{n} X_t^j = \sqrt{\frac{n}{2a}} e^{-at} \left(\sum_{j=1}^{n} \sqrt{\frac{2a}{n}} \int_0^t e^{as} \mathrm{d}B_s^j \right) =: \sqrt{\frac{n}{2a}} e^{-at} M_t.$$

Notice that M is a martingale of quadratic variation

$$\langle M \rangle_t = (e^{2at} - 1)$$

and therefore, by Lemma 3.3, we conclude that

$$
\begin{aligned}
\mathbb{E}\left[\sup_{t\in[0,T]}|v_t|^2\right] &= \frac{1}{2na}\mathbb{E}\left[\sup_{t\in[0,T]}|e^{-at}M_t|^2\right] \\
&= \frac{1}{2na}\mathbb{E}\left[\sup_{t\in[0,T]}\frac{|M_t|^2}{1+\langle M\rangle_t}\right] \\
&\leqslant C\frac{1}{2na}\log\left(1+2aT\right).
\end{aligned}
$$

Note that we crucially exploited the splitting $e^{-a(t-s)} = e^{-at}e^{as}$, which is not available in the semigroup setting we are concerned with. Intending to employ such a step suggests to pass by a functional calculus for the semigroup, which we briefly discuss next.

Recall that an analytic semigroup is a bounded linear operator that can be expressed by means of a Dunford integral (e.g. [20, 27] and Appendix A). The integral representation of S is given for every $t\in[0,T]$ by

$$
S_t = \frac{1}{2\pi i}\int_{\gamma_{r,\eta}} e^{t\lambda}R(\lambda,\tfrac{\Delta}{2})\mathrm{d}\lambda, \tag{3.7}
$$

where $R(\lambda,\tfrac{\Delta}{2}) = (\lambda\mathrm{Id}-\tfrac{\Delta}{2})^{-1}$ denotes the resolvent of $\tfrac{\Delta}{2}$ and where, for $r>0$ and $\eta\in(\pi/2,\pi)$, $\gamma_{r,\eta}$ is the curve $\{\lambda\in\mathbb{C}:|\arg\lambda|=\eta,|\lambda|\geqslant r\}\cup\{\lambda\in\mathbb{C}:|\arg\lambda|\leqslant\eta,|\lambda|=r\}$, oriented counterclockwise.

Plugging (3.7) into the expression of $w_t^n(h)$ yields

$$
w_t^n(h) = \frac{1}{2\pi in}\sum_{j=1}^n\int_0^t\int_{\gamma_{r,\eta}} e^{(t-s)\lambda}\left[\nabla R(\lambda,\tfrac{\Delta}{2})h\right](x_s^{j,n})\mathrm{d}\lambda\mathrm{d}B_s^j,
$$

splitting the complex integral into three real integrals parametrizing $\gamma_{r,\eta}$, and then using stochastic Fubini, one is left with expressions similar to

$$
\frac{1}{2\pi in}\sum_{j=1}^n\int_0^t e^{(t-s)\rho e^{i\eta}}\left[\nabla R(\rho e^{i\eta},\tfrac{\Delta}{2})h\right](x_s^{j,n})e^{i\eta}\mathrm{d}B_s^j, \quad \rho>r,
$$

which remind us of 1-dimensional self-normalized martingale for every ρ, similar to the process $(v_t)_t$ considered in Example 3.4.

It remains to establish a suitable bound on the expression

$$
\left[\nabla R(\rho e^{i\eta},\tfrac{\Delta}{2})h\right](x_s^{j,n})
$$

and to ensure that this bound is integrable for $\rho\in(r,\infty)$, see Lemma A.4.

Putting all the above considerations together with care, one obtains a maximal inequality for $w_t^n(h)$ that we present in Lemma 3.6.

Remark 3.5. A similar control has already been used in [6, Lemma 3.3], see also [3, §3.1] and [26, §4] for an estimate using the Rodemich-Garsia-Rumsey lemma. However, in all these cases the particles are living in the one dimensional torus, making the (still highly technical) noise analysis considerably simpler due to the decomposition in Fourier series.

Lemma 3.6. *Assume $m>d/2$. There exists a constant $C\geqslant 1$, independent of n and $h\in H^m$, such that for every $h\in H^m$*

$$
\mathbb{E}\left[\sup_{t\in[0,T]}|w_t^n(h)|^2\right]\leqslant\frac{C}{n}\|h\|_m^2. \tag{3.8}
$$

Law of large numbers via mild formulation

Proof. Let $h \in H^m$ and $\gamma_{r,\eta}$ be the curve in (3.7) with $\eta \in (\pi/2, \pi)$ and $r > 0$. Since the real values of η and r are not crucial for the proof, we may suppose $r > 1$. Using the decomposition of S we obtain:

$$
\begin{aligned}
w_t^n(h) &= \frac{1}{n} \sum_{j=1}^n \int_0^t \left[\nabla S_{t-s} h\right](x_s^{j,n}) \mathrm{d}B_s^j \\
&= \frac{1}{2\pi i n} \sum_{j=1}^n \int_0^t \left[\nabla \int_{\gamma_{r,\eta}} e^{(t-s)\lambda} R(\lambda, \tfrac{\Delta}{2}) h \, \mathrm{d}\lambda\right](x_s^{j,n}) \mathrm{d}B_s^j \\
&= \frac{1}{2\pi i n} \sum_{j=1}^n \int_0^t \int_{\gamma_{r,\eta}} e^{(t-s)\lambda} \left[\nabla R(\lambda, \tfrac{\Delta}{2}) h\right](x_s^{j,n}) \mathrm{d}\lambda \, \mathrm{d}B_s^j \\
&= Z_t^1(h) + Z_t^2(h) + Z_t^3(h),
\end{aligned}
$$

where in the third step we have used that ∇ is a closed linear operator on $D(\tfrac{\Delta}{2})$ and with

$$
Z_t^1(h) := \frac{1}{2\pi i n} \sum_{j=1}^n \int_0^t \int_r^\infty e^{(t-s)\rho e^{i\eta}} \left[\nabla R(\rho e^{i\eta}, \tfrac{\Delta}{2})h\right](x_s^{j,n}) e^{i\eta} \mathrm{d}\rho \, \mathrm{d}B_s^j,
$$

$$
Z_t^2(h) := \frac{1}{2\pi i n} \sum_{j=1}^n \int_0^t \int_{-\eta}^\eta e^{(t-s)r e^{i\alpha}} \left[\nabla R(r e^{i\alpha}, \tfrac{\Delta}{2})h\right](x_s^{j,n}) i r e^{i\alpha} \mathrm{d}\alpha \, \mathrm{d}B_s^j, \tag{3.9}
$$

$$
Z_t^3(h) := -\frac{1}{2\pi i n} \sum_{j=1}^n \int_0^t \int_r^\infty e^{(t-s)\rho e^{-i\eta}} \left[\nabla R(\rho e^{-i\eta}, \tfrac{\Delta}{2})h\right](x_s^{j,n}) e^{-i\eta} \mathrm{d}\rho \, \mathrm{d}B_s^j.
$$

Using the classical estimate $(a+b+c)^2 \leqslant 3(a^2+b^2+c^2)$, it follows that

$$
|w_t^n(h)|^2 \leqslant 3\left[\left|Z_t^1(h)\right|^2 + \left|Z_t^2(h)\right|^2 + \left|Z_t^3(h)\right|^2\right].
$$

We focus on $Z_t^1(h)$, but similar estimates for $Z_t^2(h)$ and $Z_t^2(h)$ follow in exactly the same way.

Fix $\epsilon > 0$ small, the stochastic Fubini theorem (e.g. [7, §4.5]) and Cauchy-Schwartz inequality imply that

$$
\begin{aligned}
\left|Z_t^1(h)\right|^2 &= \left|\int_r^\infty \left[\int_0^t \frac{\rho^{\frac{1+\epsilon}{2}}}{2\pi i n} \sum_{j=1}^n e^{(t-s)\rho e^{i\eta}} \left[\nabla R(\rho e^{i\eta}, \tfrac{\Delta}{2})h\right](x_s^{j,n}) e^{i\eta} \mathrm{d}B_s^j\right] \frac{\mathrm{d}\rho}{\rho^{\frac{1+\epsilon}{2}}}\right|^2 \\
&\leqslant C \int_r^\infty \left|\int_0^t \frac{1}{n} \sum_{j=1}^n e^{(t-s)\rho e^{i\eta}} \left[\nabla R(\rho e^{i\eta}, \tfrac{\Delta}{2})h\right](x_s^{j,n}) e^{i\eta} \mathrm{d}B_s^j\right|^2 \rho^{1+\epsilon} \mathrm{d}\rho \\
&= \frac{C}{n^2} \int_r^\infty e^{-2t\rho(-\cos\eta)} \underbrace{\left|\int_0^t \sum_{j=1}^n e^{-s\rho e^{i\eta}} \left[\nabla R(\rho e^{i\eta}, \tfrac{\Delta}{2})h\right](x_s^{j,n}) \mathrm{d}B_s^j\right|^2}_{=:M_t} \rho^{1+\epsilon} \mathrm{d}\rho
\end{aligned}
$$

where $C = \frac{1}{4\pi^2} \int_r^\infty \frac{\mathrm{d}\rho}{\rho^{1+\epsilon}}$.

We introduce the continuous martingale $X_\cdot^{\epsilon,\rho}(h)$ defined for $t \geqslant 0$ by

$$
X_t^{\epsilon,\rho}(h) := \rho^{1/2+\epsilon} \sqrt{\frac{-2\rho\cos\eta}{\|h\|_m^2 \, n}} M_t,
$$

so to obtain

$$\left|Z_t^1(h)\right|^2 \leqslant \frac{C}{-2n\cos\eta} \|h\|_m^2 \int_r^\infty e^{2t\rho\cos\eta} \left|X_t^{\epsilon,\rho}(h)\right|^2 \frac{\mathrm{d}\rho}{\rho^{1+\epsilon}}$$

$$\leqslant \frac{C}{n} \|h\|_m^2 \int_r^\infty e^{2t\rho\cos\eta} \left|X_t^{\epsilon,\rho}(h)\right|^2 \frac{\mathrm{d}\rho}{\rho^{1+\epsilon}}.$$

where we absorbed the factor $(-2\cos\eta)^{-1}$ in the unessential constant C.

We compute the quadratic variation of $X_t^{\epsilon,\rho}(h)$:

$$\langle X^{\epsilon,\rho}(h)\rangle_t = \rho^{1+2\epsilon} \frac{(-2\rho\cos\eta)}{\|h\|_m^2 \, n} \sum_{j=1}^n \int_0^t e^{-2s\rho\cos\eta} \left[\nabla R(\rho e^{i\eta}, \tfrac{\Delta}{2})h\right]^2 (x_s^{j,n})\mathrm{d}s.$$

Lemma A.3 assures that for every ϵ such that $0 < 2\epsilon < (m-d/2)\wedge 1$, $\mathcal{P}(\mathbb{R}^d)$ is continuously embedded in $H^{-m+2\epsilon}$, in particular

$$\left| \left[\nabla R(\rho e^{i\eta}, \tfrac{\Delta}{2})h\right](x_s^{j,n})\right| = \left|\langle\delta_{x_s^{j,n}}, \nabla R(\rho e^{i\eta}, \tfrac{\Delta}{2})h\rangle_{-m,m}\right|$$

$$= \left|\langle\delta_{x_s^{j,n}}, \nabla R(\rho e^{i\eta}, \tfrac{\Delta}{2})h\rangle_{-m+2\epsilon, m-2\epsilon}\right| \tag{3.10}$$

$$\leqslant \left\|\delta_{x_s^{j,n}}\right\|_{-m+2\epsilon} \left\|\nabla R(\rho e^{i\eta}, \tfrac{\Delta}{2})h\right\|_{m-2\epsilon} \leqslant C\frac{\|h\|_m}{\rho^{1/2+\epsilon}},$$

where we have exploited the properties of the resolvent operator R, see Lemma A.4.

Thus, the quadratic variation of $X_t^{\epsilon,\rho}(h)$ is bounded $\mathbb{P}$-a.s. by

$$\langle X^{\epsilon,\rho}(h)\rangle_t \leqslant C(-2\rho\cos\eta) \int_0^t e^{-2s\rho\cos\eta}\mathrm{d}s = C\left(e^{-2t\rho\cos\eta} - 1\right). \tag{3.11}$$

Observe then

$$\mathbb{E}\left[\sup_{t\in[0,T]} \left|Z_t^1(h)\right|^2\right] \leqslant \frac{C}{n} \|h\|_m^2 \int_r^\infty \mathbb{E}\left[\sup_{t\in[0,T]} e^{2t\rho\cos\eta} \left|X_t^{\epsilon,\rho}(h)\right|^2\right] \frac{\mathrm{d}\rho}{\rho^{1+\epsilon}}$$

$$\leqslant \frac{C}{n} \|h\|_m^2 \int_r^\infty \mathbb{E}\left[\sup_{t\in[0,T]} \frac{\left|X_t^{\epsilon,\rho}(h)\right|^2}{1+\langle X^{\epsilon,\rho}(h)\rangle_t} \left(\sup_{t\in[0,T]} \frac{1+\langle X^{\epsilon,\rho}(h)\rangle_t}{e^{-2t\rho\cos\eta}}\right)\right] \frac{\mathrm{d}\rho}{\rho^{1+\epsilon}}.$$

The term $\sup_{t\in[0,T]} \frac{1+\langle X^{\epsilon,\rho}(h)\rangle_t}{e^{-2t\rho\cos\eta}}$ is bounded using (3.11) by a constant, wherefore we are left with

$$\mathbb{E}\left[\sup_{t\in[0,T]} \left|Z_t^1(h)\right|^2\right] \leqslant \frac{C}{n} \|h\|_m^2 \int_r^\infty \mathbb{E}\left[\sup_{t\in[0,T]} \frac{\left|X_t^{\epsilon,\rho}(h)\right|^2}{1+\langle X^{\epsilon,\rho}(h)\rangle_t}\right] \frac{\mathrm{d}\rho}{\rho^{1+\epsilon}}.$$

We now invoke Lemma 3.3, which in conjunction with (3.11) allows to deduce that

$$\mathbb{E}\left[\sup_{t\in[0,T]} \frac{\left|X_t^{\epsilon,\rho}(h)\right|^2}{1+\langle X^{\epsilon,\rho}(h)\rangle_t}\right] \leqslant C\,\mathbb{E}\left[\log\left(1 + \log\left(1 + \langle X^{\epsilon,\rho}(h)\rangle_T\right)\right)\right]$$

$$\leqslant C\mathbb{E}\left[\log\left(1 - 2T\rho\cos\eta + \log(C)\right)\right]$$

where in the last inequality, we have bounded the constant C appearing in (3.11) by $\max\{1, C\}$. Further modifying C accordingly, we are thus left with

$$\mathbb{E}\left[\sup_{t\in[0,T]} \left|Z_t^1(h)\right|^2\right] \leqslant \frac{C}{n} \|h\|_m^2 \int_r^\infty \log(1 - 2T\rho\cos\eta + \log(C))\frac{\mathrm{d}\rho}{\rho^{1+\epsilon}} \leqslant \frac{C}{n} \|h\|_m^2.$$

Concerning $Z_t^3(h)$, computations are the same if one replaces η by $-\eta$. Concerning $Z_t^2(h)$, computations are easier since there is no a priori diverging integral to deal with and we omit the proof. The overall bound on $w_t^n(h)$ is thus obtained by summing the three estimates and choosing the constant C accordingly. □

Remark 3.7. Note that Lemma 3.6 implies by Jensen's inequality the following bound

$$\mathbb{E}\left[\sup_{t\in[0,T]} |w_t^n(h)|\right] \leqslant \frac{C}{\sqrt{n}} \|h\|_m,$$

which is sharper in n with respect to (3.6), but in a weaker topology. One could ask if it is possible to establish a similar $O(1/\sqrt{n})$ bound for

$$\mathbb{E}\left[\sup_{t\in[0,T]} \|w_t^n\|_{-m}\right] = \mathbb{E}\left[\sup_{t\in[0,T]} \sup_{\|h\|_m \leqslant 1} |w_t^n(h)|\right].$$

Such a bound cannot be obtained by rough paths theory and a full probabilistic proof, which takes the independence between the Brownian motions into account, is desirable. To the authors' knowledge, this has been established only in the case of interacting oscillators; we refer to the noise term analysis in [3, 6, 26].

3.3 Proving Theorem 2.3

The proof of Theorem 2.3 consists in two steps: using the pathwise bound on w^n, Lemma 3.8 shows that we can extract from $(\nu^n)_{n\in\mathbb{N}}$ a weak-*-convergence subsequence; then, by exploiting the probability bound on $w_t^n(h)$ for a fixed $h \in H^m$, we identify through Lemma 3.9 the limit with a solution to (2.1).

3.3.1 Extraction of a weak-*-convergent subsequence

The main result of this subsection is given by the next lemma.

Lemma 3.8. *The sequence $(\nu^n)_{n\in\mathbb{N}}$ is uniformly bounded in $L^\infty([0,T], H^{-m})$ and thus admits a subsequence that converges weak-* to some $\nu \in H^{-m}$, $\mathbb{P}$-a.s..*

Proof. It suffices to show that $(\nu^n)_{n\in\mathbb{N}}$ is uniformly bounded in $L^\infty([0,T], H^{-m})$ $\mathbb{P}$-a.s., an application of Banach-Alaoglu yields the existence of a convergent subsequence.

Exploiting the mild formulation in Lemma 3.1 and the bound on $w_t^n(h)$ in Lemma 3.2 for some $\alpha \in (1/3, 1/2)$, one obtains that

$$\|\nu_t^n\|_{-m} \leqslant \|\nu_0^n\|_{-m} + \int_0^t \|\nu_s^n\|_{-m} \sup_{\|h\|_m \leqslant 1} \|(\nabla S_{t-s}h)(\Gamma * \nu_s^n)\|_m \, \mathrm{d}s + \|w_t^n\|_{-m}$$

$$\leqslant \|\nu_0^n\|_{-m} + \int_0^t \frac{C}{\sqrt{t-s}} \|\nu_s^n\|_{-m} \, \mathrm{d}s + C_\alpha (1+t)^{3\alpha},$$

where we have exploited the properties of the semigroup and the bound already used in (2.3). A Gronwall-like argument implies the existence of a constant a independent of n and T such that

$$\sup_{t\in[0,T]} \|\nu_t^n\|_{-m} \leqslant 2\left(\|\nu_0^n\|_{-m} + C_\alpha (1+T)^{3\alpha}\right) \sqrt{T} e^{a\sqrt{T}}.$$

In particular, using Lemma A.3, we conclude

$$\sup_{n\in\mathbb{N}} \sup_{t\in[0,T]} \|\nu_t^n\|_{-m} \leqslant C_{\alpha,T}. \qquad \square$$

We move to the identification of the limit $\nu \in L^\infty([0,T], H^{-m})$.

3.3.2 The limit coincides with an m-weak-mild solution

We prove that any possible limit of $(\nu^n)_{n \in \mathbb{N}}$ is a weak-mild solution (2.1). Given the uniqueness of (2.1), this implies the weak-* convergence in $L^\infty([0,T], H^{-m})$ of $(\nu^n)_{n \in \mathbb{N}}$ to the element ν given in Lemma 3.8.

Lemma 3.9. *Let $(\nu^n)_{n \in \mathbb{N}}$ be converging weak-* to some $\bar{\nu} \in L^\infty([0,T], H^{-m})$ $\mathbb{P}$-a.s. along a subsequence that we denote by $(\nu^{n_k})_{k \in \mathbb{N}}$. Then $\bar{\nu}$ satisfies (2.1), i.e.*

$$\langle \bar{\nu}_t, h \rangle_{-m,m} = \langle \bar{\nu}_0, S_t h \rangle_{-m,m} + \int_0^t \langle \bar{\nu}_s, (\nabla S_{t-s} h)(\Gamma * \bar{\nu}_s) \rangle_{-m,m} \mathrm{d}s,$$

meaning $\bar{\nu}$ is an m-weak-mild solution to (1.3).

Proof. Recall that for every n, ν^n solves the mild formulation (3.1), i.e. for $t \in [0,T]$

$$\langle \nu_t^n, h \rangle_{-m,m} = \langle \nu_0^n, S_t h \rangle_{-m,m} + \int_0^t \langle \nu_s^n, (\nabla S_{t-s} h)(\Gamma * \nu_s^n) \rangle_{-m,m} \mathrm{d}s + w_t^n(h).$$

By hypothesis we have that for every $t \in [0,T]$ and $h \in H^m$

$$\lim_{k \to \infty} \langle \nu_t^{n_k}, h \rangle_{-m,m} = \langle \bar{\nu}_t, h \rangle_{-m,m}, \quad \mathbb{P}\text{-a.s.}.$$

In particular, this is true for $(\nu_0^{n_k})_k$ since $S_t h \in H^m$. Furthermore, Lemma 3.6 implies that

$$\lim_{k \to \infty} w_t^{n_k}(h) = 0, \quad \text{in } \mathbb{P}\text{-probability}$$

and thus in particular the convergence holds $\mathbb{P}$-a.s. along a sub-subsequence $(n_{k_j})_j$. Thus, it remains to show that $\mathbb{P}$-a.s.

$$\lim_{j \to \infty} \int_0^t \langle \nu_s^{n_{k_j}}, (\nabla S_{t-s} h)(\Gamma * \nu_s^{n_{k_j}}) \rangle_{-m,m} \mathrm{d}s = \int_0^t \langle \bar{\nu}_s, (\nabla S_{t-s} h)(\Gamma * \bar{\nu}_s) \rangle_{-m,m} \mathrm{d}s. \quad (3.12)$$

For better readability and lighter notation, we will not distinguish between n and n_{k_j} in the following, understanding that we continue to work on the sub-subsequence. Consider then

$$\langle \bar{\nu}_s^n, (\nabla S_{t-s} h)(\Gamma * \bar{\nu}_s^n) \rangle_{-m,m} - \langle \bar{\nu}_s, (\nabla S_{t-s} h)(\Gamma * \bar{\nu}_s) \rangle_{-m,m}$$
$$= \langle \bar{\nu}_s^n - \bar{\nu}_s, (\nabla S_{t-s} h)(\Gamma * \bar{\nu}_s) \rangle_{-m,m} + \langle \bar{\nu}_s^n, (\nabla S_{t-s} h)(\Gamma * (\bar{\nu}_s^n - \bar{\nu}_s)) \rangle_{-m,m}.$$

Using again (2.3), it is easy to see that $\mathbb{P}$-a.s.

$$\mathbf{1}_{[0,t]}(s)(\nabla S_{t-s} h)(\Gamma * \nu_s) \in L^1([0,T], H^m)$$

wherefore it is indeed an element of the predual to $L^\infty([0,T], H^{-m})$ and thus by weak-* convergence in this space, we have

$$\int_0^T \langle \nu_s^n - \nu_s, \mathbf{1}_{[0,t]}(s)(\nabla S_{t-s} h)(\Gamma * \nu_s) \rangle \mathrm{d}s = \int_0^t \langle \nu_s^n - \nu_s, (\nabla S_{t-s} h)(\Gamma * \nu_s) \rangle \mathrm{d}s \to 0.$$

For the second term, note that

$$\langle \bar{\nu}_s^n, (\nabla S_{t-s} h)(\Gamma * (\bar{\nu}_s^n - \bar{\nu}_s)) \rangle_{-m,m}$$
$$= \langle \bar{\nu}_s^n - \bar{\nu}_s, \langle \bar{\nu}_s^n(\mathrm{d}x), (\nabla S_{t-s} h)(x)(\Gamma(x, \cdot)) \rangle_{-m,m} \rangle_{-m,m}$$

and that the function

$$y \mapsto \langle \bar{\nu}_s^n(\mathrm{d}x), (\nabla S_{t-s} h)(x)(\Gamma(x, y)) \rangle_{-m,m}$$
$$= \frac{1}{n} \sum_{j=1}^n (\nabla S_{t-s} h(x_s^{j,n})) \cdot (\Gamma(x_s^{j,n}, y))$$

is in H^m for every $s \in [0, t]$, since $\Gamma(x, \cdot) \in H^m$ for a.e. $x \in \mathbb{R}^d$, recall (2.2). Namely,

$$\left\| \frac{1}{n} \sum_{j=1}^{n} (\nabla S_{t-s} h(x_s^{j,n})) \cdot (\Gamma(x_s^{j,n}, \cdot)) \right\|_m \leqslant \|\nabla S_{t-s} h\|_\infty \|\Gamma\|_{H_y^m L_x^\infty}$$
$$\leqslant C \frac{\|h\|_m}{\sqrt{t-s}} \|\Gamma\|_{H_y^m W_x^{m,\infty}}.$$

We conclude that

$$\mathbf{1}_{[0,t]} \langle \bar{\nu}_s^n(\mathrm{d}x), (\nabla S_{t-s}h)(x)(\Gamma(x, \cdot)) \rangle_{-m,m} \in L^1([0,T], H^m)$$

and in particular

$$\int_0^T \langle \bar{\nu}_s^n - \bar{\nu}_s, \langle \bar{\nu}_s^n(\mathrm{d}x), \mathbf{1}_{[0,t]}(s)(\nabla S_{t-s}h)(x) \cdot (\Gamma(x, \cdot)) \rangle_{-m,m} \rangle_{-m,m} \mathrm{d}s \to 0.$$

This establishes (3.12).

Overall, we have thus shown that any subsequence of $(\nu^n)_n$ converges along some further subsequence $\mathbb{P}$-a.s. weak-* in $L^\infty([0,T], H^{-m})$, the limit $\bar{\nu}$ satisfying for every $h \in H^m$ the equation

$$\langle \bar{\nu}_t, h \rangle_{-m,m} = \langle \bar{\nu}_0, S_t h \rangle_{-m,m} + \int_0^t \langle \bar{\nu}_s, (\nabla S_{t-s}h)(\Gamma * \bar{\nu}_s) \rangle_{-m,m} \mathrm{d}s,$$

meaning that $\bar{\nu}$ is indeed an m-weak-mild solution. $\qquad\square$

3.3.3 Proof of Theorem 2.3

In order to show that $\nu^n \overset{*}{\rightharpoonup} \nu$ in $L^\infty([0,T], H^{-m})$ in probability, we show that any subsequence $(\nu^{n_k})_k$ admits a further subsequence that converges $\mathbb{P}$-a.s. in weak-* topology of $L^\infty([0,T], H^{-m})$ to ν.

Let $(\nu^{n_k})_k$ be hence a subsequence. By assumption of the Theorem, Lemmas 3.6 and 3.8, we find a further subsequence $(\nu^{n_{k_j}})_j$, along which

$$
\begin{aligned}
w_t^{n_{k_j}}(h) &\to 0 \quad \forall h \in H^m \\
\langle \nu_0^{n_{k_j}}, h \rangle &\to \langle \nu_0, h \rangle \quad \forall h \in H^m \\
\nu^{n_{k_j}} &\overset{*}{\rightharpoonup} \bar{\nu} \quad \text{in } L^\infty([0,T], H^{-m})
\end{aligned}
\tag{3.13}
$$

$\mathbb{P}$-a.s., where the limit $\bar{\nu} \in L^\infty([0,T], H^{-m})$ may apriori depend on the subsequence chosen. Notice however that due to Lemma 3.9, any such limit is a m-weak-mild solution to (1.3). By the uniqueness result of Proposition 2.2, we conclude that the limit $\bar{\nu} = \nu$ must be the same for any subsequence chosen.

The first part of the Theorem is proved. Note that apriori, our limit ν is only a distribution in H^{-m} at each fixed timepoint.

Suppose $\nu_0 \in \mathcal{P}(\mathbb{R}^d)$. In order to show that ν_t is actually a probability measure for each $t \in [0,T]$, we observe that a weak solution $\mu \in C([0,T], \mathcal{P}(\mathbb{R}^d))$ to (1.3) (which exists due to Theorem 1.2) is a weak-mild solution (2.1).

Indeed, let $\mu = (\mu_t)_{t \in [0,T]} \in C([0,T], \mathcal{P}(\mathbb{R}^d))$ be a weak solution to (1.3). As done in Lemma 3.1, one can show that for every $f \in C_0^\infty$ and $t \in [0,T]$

$$\langle \mu_t, f \rangle_{-m,m} = \langle \mu_0, S_t f \rangle_{-m,m} + \int_0^t \langle \mu_s, (\nabla S_{t-s}f) \cdot (\Gamma * \mu_s) \rangle_{-m,m} \mathrm{d}s \tag{3.14}$$

holds. Note that by standard approximation, (3.14) holds also for $f \in H^m \subset C_b^3$, meaning that μ is indeed a weak-mild solution. By the uniqueness statement of Proposition 2.2 we conclude $\mu = \nu$ and thus in particular $\nu \in C([0,T], \mathcal{P}(\mathbb{R}^d))$. This concludes the second part and thus the entire proof of the Theorem.

A Hilbert spaces and semigroups

A.1 The Laplacian semigroup

The following definitions are taken from [20, 27]. For the sake of notation, we focus on Δ, the standard Laplacian on $L^2(\mathbb{R}^d)$, instead of $\frac{\Delta}{2}$. We can consider the part of Δ on (the complexification e.g. [27, Appendix A] of) H^m:

$$\Delta : \mathcal{D}(\Delta) \subset H^m \longrightarrow H^m.$$

It is not difficult to see that $H^{m+2} \subset \mathcal{D}(\Delta)$, where the inclusion is dense, and that Δ is a sectorial operator with spectrum given by $(-\infty, 0]$. In particular, it generates an analytic strongly continuous semigroup denoted for all $t \geqslant 0$ by S_t; recall that $S_0 := \mathrm{Id}$ is the identity operator.

We represent S for $t \in [0, T]$ as the following Dunford integral

$$S_t = \frac{1}{2\pi i} \int_{\gamma_{r,\eta}} e^{t\lambda} R(\lambda, \Delta) \mathrm{d}\lambda,$$

where $R(\lambda, \Delta) = (\lambda \mathrm{Id} - \Delta)^{-1}$ denotes the resolvent of Δ and where, for $r > 0$ and $\eta \in (\pi/2, \pi)$, $\gamma_{r,\eta}$ is the curve $\{\lambda \in \mathbb{C} : |\arg \lambda| = \eta, |\lambda| \geqslant r\} \cup \{\lambda \in \mathbb{C} : |\arg \lambda| \leqslant \eta, |\lambda| = r\}$, oriented counterclockwise.

Observe that $\gamma_{r,\eta}$ is contained in the resolvent set of Δ, i.e. $\gamma_{r,\eta} \subset \rho(\Delta)$, and that, for all regular values $\lambda \in \rho(\Delta)$, $R(\lambda, \Delta)$ is a bounded linear operator on H^m.

When computing the semigroup against a function h through (3.7), we use the following decomposition into three real integrals:

$$
\begin{aligned}
S_t h &= \frac{1}{2\pi i} \int_{\gamma_{r,\eta}} e^{t\lambda} R(\lambda, \Delta) h \mathrm{d}\lambda = \frac{1}{2\pi i} \Bigg[\int_r^\infty e^{t\rho e^{i\eta}} R(\rho e^{i\eta}, \Delta) e^{i\eta} \mathrm{d}\rho \\
&\quad + \int_{-\eta}^{\eta} e^{tr(\cos\alpha + i\sin\alpha)} R(re^{i\alpha}, \Delta) i r e^{i\alpha} \mathrm{d}\alpha - \int_r^\infty e^{t\rho e^{-i\eta}} R(\rho e^{-i\eta}, \Delta) e^{-i\eta} \mathrm{d}\rho \Bigg].
\end{aligned}
\tag{A.1}
$$

The section ends with some estimates basic concerning the regularity of S.

Lemma A.1. *Let $(S_t)_{t \leqslant T}$ be the heat semigroup acting on H^m. Then for $\alpha \geqslant 0$ and $t > 0$ one has*

$$\|(-\Delta)^\alpha S_t h\|_m \leqslant C \frac{1}{t^\alpha} \|h\|_m.$$

In particular, we also have

$$\|\nabla S_t h\|_m \leqslant C \frac{1}{\sqrt{t}} \|h\|_m$$

Proof. Refer to [33, Chapter II, Theorem 6.13] and the proof of [21, Proposition 4.3] for example. □

Lemma A.2. *Assume $m > d/2 + 3$. Let $(S_t)_{t \leqslant T}$ be the heat semigroup acting on H^m. For $f \in H^m$, it holds that*

$$
\begin{aligned}
\|\nabla (S_t - \mathbf{Id}) f\|_\infty &\leqslant \sqrt{t} \, \|D^2 f\|_\infty \leqslant C \sqrt{t} \, \|f\|_m, \\
\|\nabla (S_t - \mathbf{Id}) f\|_\infty &\leqslant \frac{1}{2} t \, \|D^3 f\|_\infty \leqslant \frac{C}{2} t \, \|f\|_m,
\end{aligned}
$$

where D^2 is the Hessian and D^3 the tensor with third-order derivatives. In particular

$$
\begin{aligned}
\|\nabla (S_t - S_s) f\|_\infty &\leqslant \sqrt{|t-s|} \, \|D^2 f\|_\infty \leqslant C \sqrt{|t-s|} \, \|f\|_m, \\
\|\nabla (S_t - S_s) f\|_\infty &\leqslant \frac{1}{2} |t-s| \, \|D^3 f\|_\infty \leqslant \frac{C}{2} |t-s| \, \|f\|_m.
\end{aligned}
$$

Proof. We calculate explicitly

$$
\begin{aligned}
|\nabla(S_t - \mathrm{Id})f(x)| &= \left| \int_{\mathbb{R}^d} \frac{1}{(2\pi t)^{d/2}} e^{-\frac{|y|^2}{2t}} (\nabla f(x-y) - \nabla f(x)) \mathrm{d}y \right| \\
&= \left| \int_{\mathbb{R}^d} \frac{1}{(2\pi t)^{d/2}} e^{-\frac{|y|^2}{2t}} \left(\nabla f(x-y) - \nabla f(x) + (-y)^T (D\nabla f)(x) \right) \mathrm{d}y \right| \\
&\leqslant \frac{1}{2} \left\| D^3 f \right\|_\infty \int_{\mathbb{R}^d} \frac{1}{(2\pi t)^{d/2}} e^{-\frac{|y|^2}{2t}} |y|^2 \mathrm{d}y \\
&= \frac{1}{2} \left\| D^3 f \right\|_\infty t
\end{aligned}
$$

where we exploited the asymmetry of the first Taylor component. The first statement follows from a similar consideration, considering an order one Taylor expansion of $\nabla f(x-y)$ around x instead of an order two Taylor expansion. The proof follows by Sobolev's embeddings. $\qquad\square$

A.2 The Hilbert space H^s

It is useful to give an explicit definition of H^m through the Fourier transform (e.g. [1, 7.62]). Let $s > 0$, define $(H^s, \|\cdot\|_s)$ by

$$
\begin{aligned}
H^s &= \left\{ u \in L^2(\mathbb{R}^d) \;:\; \int_{\mathbb{R}^d} (1 + |\xi|^2)^s \left| \mathcal{F}(u)(\xi) \right|^2 \mathrm{d}\xi < \infty \right\}, \\
\|u\|_s^2 &= \int_{\mathbb{R}^d} (1 + |\xi|^2)^s \left| \mathcal{F}(u)(\xi) \right|^2 \mathrm{d}\xi.
\end{aligned}
\tag{A.2}
$$

Whenever s is an integer, it is well known that this definition coincides with the standard definition of the Sobolev space $W^{s,2}(\mathbb{R}^d)$.

The next lemma extends the embedding (1.5) to H^s and its relationship with the space of probability measures.

Lemma A.3. *For all $s > d/2$, one has the following continuous embedding*

$$
H^s \subset C_b(\mathbb{R}^d).
\tag{A.3}
$$

Moreover, there exists $C > 0$ (depending on s only) such that

$$
\sup_{\mu \in \mathcal{P}(\mathbb{R}^d)} \|\mu\|_{-s} \leqslant C.
\tag{A.4}
$$

Proof. The continuous embedding (A.3) is a consequence of the embedding of Besov spaces into the space of continuous bounded functions (e.g. [1, Theorem 7.34]) and the fact that they coincide with H_s for a particular choice of the indices.

Turning to (A.4), let $\mu \in \mathcal{P}(\mathbb{R}^d)$, then

$$
\|\mu\|_{-s} = \sup_{h \in H^s} \frac{\langle \mu, h \rangle_{-s,s}}{\|h\|_s} = \sup_{h \in H^s} \frac{\langle \mu, h \rangle}{\|h\|_s} \leqslant \sup_{h \in H^s} \frac{\|h\|_\infty}{\|h\|_s} \leqslant C,
$$

where C is the norm of the identity operator between H^s and $C_b(\mathbb{R}^d)$. $\qquad\square$

A.3 Fractional operators on H^s

We have the following lemma.

Lemma A.4. *Let $\lambda = \rho e^{i\eta} \in \rho(\Delta)$ and suppose $\rho > 1$. There exists a positive constant $C = C_\eta$ such that for every $\epsilon \in (0, 1/2)$*

$$
\left\| \nabla R(\rho e^{i\eta}, \Delta) h \right\|_{m-2\epsilon}^2 \leqslant C_\eta \frac{\|h\|_m^2}{\rho^{1+2\epsilon}}, \quad h \in H^m.
\tag{A.5}
$$

Proof. Exploiting the Fourier multipliers associated to ∇ and R, one obtains

$$\left\|\nabla R(\rho e^{i\eta}, \Delta) h\right\|_{m-2\epsilon}^2 = \int_{\mathbb{R}^d} (1+|\xi|^2)^{m-2\epsilon} \left|\mathcal{F}(\nabla R(\rho e^{i\eta}, \Delta) h)(\xi)\right|^2 \mathrm{d}\xi$$

$$\leqslant \int_{\mathbb{R}^d} (1+|\xi|^2)^{m-2\epsilon} |\mathcal{F}(h)(\xi)|^2 \left|\frac{\xi}{\rho e^{i\eta} + |\xi|^2}\right|^2 \mathrm{d}\xi$$

$$\leqslant \int_{\mathbb{R}^d} (1+|\xi|^2)^m |\mathcal{F}(h)(\xi)|^2 \frac{|\xi|^2}{|\rho e^{i\eta} + |\xi|^2|^2} \frac{1}{(1+|\xi|^2)^{2\epsilon}} \mathrm{d}\xi.$$

Since we are assuming $\rho > 1$, we have that

$$\frac{|\xi|^2}{|\rho e^{i\eta} + |\xi|^2|^2} \frac{1}{(1+|\xi|^2)^{2\epsilon}} \leqslant \frac{(\rho + |\xi|^2)^{1-2\epsilon}}{|\rho e^{i\eta} + |\xi|^2|^2} \leqslant C_\eta \frac{1}{(\rho + |\xi|^2)^{1+2\epsilon}} \leqslant C_\eta \frac{1}{\rho^{1+2\epsilon}},$$

with $C_\eta = (\sup_{x \geqslant 0} (1+x)/|e^{i\eta} + x|)^2 < \infty$ since $\eta \neq \pi$. $\square$

B Rough integration associated to semigroup functionals

We mostly follow [18] and use very similar notations. Let $k \geqslant 1$ and Δ_k be the k-dimensional simplex given by

$$\Delta_k = \{t_1, \ldots, t_k \in [0, T] : T \geqslant t_1 \geqslant t_2 \geqslant \ldots \geqslant t_k \geqslant 0\}.$$

Let $C_k = C(\Delta_k; \mathbb{R})$ and W a Banach space with a strongly continuous semigroup $(S_t)_{t \in [0,T]}$ acting on it. Define D_k as the space of linear operators from W to C_k. Furthermore, let $D_* = \bigcup_{k \geqslant 1} D_k$ and define the following operators on D_*:

$$\delta : D_k \to D_{k+1}, \quad \phi : D_k \to D_{k+1}, \qquad k \geqslant 1.$$

For $A \in D_k$ and $f \in W$, they are defined as

$$[\delta A f]_{t_1 \ldots t_{k+1}} = \sum_{i=1}^{k+1} (-1)^{i+1} [A f]_{t_1 \ldots \not{t_i} \ldots t_{k+1}},$$

$$[\phi A f]_{t_1 \ldots t_{k+1}} = [A(S_{t_1 t_2} - \mathrm{Id}) f]_{t_2 \ldots t_{k+1}},$$

where $\not{t_i}$ means that the argument t_i is omitted and $S_{t_1 t_2}$ stands for $S_{t_1 - t_2}$.

We are ready for the first lemma.

Lemma B.1. *Let $\hat{\delta} := \delta - \phi$. Then $(D_*, \hat{\delta})$ is an acyclic cochain complex. In particular*

$$\mathrm{Ker}\, \hat{\delta}|_{D_{k+1}} = \mathrm{Im}\, \hat{\delta}|_{D_k}, \quad \text{for any } k \geqslant 1.$$

Proof. The proof mimics [18, Proposition 3.1]. We only mention that for proving $\mathrm{Ker}\, \hat{\delta}|_{D_{k+1}} \subset \mathrm{Im}\, \hat{\delta}|_{D_k}$, a possible choice for $A \in \mathrm{Ker}\, \hat{\delta}|_{D_{k+1}}$ is given by $B \in D_k$ defined as

$$[Bf]_{t_1 \ldots t_k} = (-1)^{k+1} [Af]_{t_1 \ldots t_k 0}, \quad f \in W. \qquad \square$$

We now introduce some analytical assumptions on the previous function spaces. We start with a Hölder-like norm on C_k for $k = 2, 3$. For $\mu > 0$ and $g \in C_2$, define

$$\|g\|_\mu := \sup_{t,s \in \Delta_2} \frac{|g_{ts}|}{|t-s|^\mu},$$

and consequently

$$C_2^\mu := \left\{ g \in C_2 \,;\, \|g\|_\mu < \infty \right\}.$$

For $\gamma, \rho > 0$ and $g \in C_3$, define

$$\|g\|_{\gamma,\rho} := \sup_{t,u,s \in \Delta_3} \frac{|g_{tus}|}{|t-u|^{\mu}\,|u-s|^{\rho}}$$

and

$$\|g\|_{\mu} := \inf\left\{\sum_i \|g_i\|_{\rho_i, \mu-\rho_i}\;;\; g = \sum_i g_i\,,\, 0 < \rho_i < \mu\right\},$$

where the infimum is taken on all sequences $(g_i) \subset C_3$ such that $g = \sum_i g_i$ and $\rho_i \in (0, \mu)$. Again, $\|\cdot\|_{\mu}$ defines a norm on C_3 and we denote the induced subspace by

$$C_3^{\mu} := \left\{g \in C_3\,;\, \|g\|_{\mu} < \infty\right\}.$$

With these definitions in mind, let

$$D_k^{\mu} := L(W, C_k^{\mu}), \quad D_k^{1+} := \bigcup_{\mu > 1} D_k^{\mu}, \qquad k = 2, 3.$$

The space $L(W, C_k^{\mu})$ is the space of linear bounded operators from W to C_k^{μ} equipped with its corresponding operator norm, i.e.

$$\|A\|_{D_k^{\mu}} := \sup_{\|f\|_W \leqslant 1} \|Af\|_{\mu}, \quad f \in D_k^{\mu}.$$

The main tool for constructing the pathwise integral associated to semigroup functionals is given by the next lemma and the following corollary. We use the notation $\hat{\delta}(D_k) := \operatorname{Im} \hat{\delta}|_{D_k}$ for $k \geqslant 1$.

Lemma B.2 (Sewing)**.** *There exists a unique linear operator*

$$\Lambda : D_3^{1+} \cap \hat{\delta}(D_2) \to D_2^{1+}$$

such that

$$\hat{\delta}\Lambda = \operatorname{Id}|_{D_3^{1+} \cap \hat{\delta}(D_2)}.$$

Moreover, if $\eta > 1$ then Λ is a continuous operator from $D_3^{\eta} \cap \hat{\delta}(D_2)$ to D_2^{η}, i.e. there exists a constant $C = C_{\eta} > 0$ such that

$$\|\Lambda A\|_{\eta} \leqslant C_{\eta} \|A\|_{\eta}, \quad A \in D_3^{\eta} \cap \hat{\delta}(D_2). \tag{B.1}$$

Proof. Concerning uniqueness, let $\widetilde{\Lambda}$ be another map satisfying the conditions stated in the Lemma. Then for $A \in D_3^{1+} \cap \hat{\delta}D_2$ we have

$$\hat{\delta}(\widetilde{\Lambda}A - \Lambda A) = A - A = 0$$

hence $Q := \widetilde{\Lambda}A - \Lambda A \in \operatorname{Ker}(\hat{\delta}) \cap D_2$. By Lemma B.1 there exists $q \in D_1$ such that $Q = \hat{\delta}q$. Note that for any partition $\mathcal{P}^n([s,t]) = (t_i)_{0 \leqslant i \leqslant n+1}$ of the interval $[s,t] \subset [0,T]$ such that $t_0 = s$ and $t_{n+1} = t$, we have the following telescopic sum expansion

$$\sum_{i=0}^{n} [\hat{\delta}q S_{tt_{i+1}}f]_{t_{i+1}t_i} = \sum_{i=0}^{n} [q S_{tt_{i+1}}f]_{t_{i+1}} - [q S_{tt_{i+1}}]_{t_i} - [q S_{tt_{i+1}}(S_{t_{i+1}t_i} - \operatorname{Id})f]_{t_i}$$

$$= \sum_{i=0}^{n} [q S_{tt_{i+1}}f]_{t_{i+1}} - [q S_{tt_i}f]_{t_i}$$

$$= [qf]_t - [q S_{ts}f]_s$$

$$= [\hat{\delta}qf]_{ts}$$

for any $f \in W$. We conclude

$$[Qf]_{ts} = [\hat{\delta}qf]_{ts} = \sum_{i=0}^{n}[\hat{\delta}qS_{tt_{i+1}}f]_{t_{i+1}t_i} = \sum_{i=0}^{n}[QS_{tt_{i+1}}f]_{t_{i+1}t_i}.$$

Letting $\mathcal{P}^n([s,t])$ be for example be the dyadic partition, one obtains for $Q \in D_2^{\gamma}$, with $\gamma > 1$, the estimate

$$|[QS_{tt_{i+1}}f]_{t_{i+1}t_i}| \leqslant 2^{-n\gamma}\left\|QS_{tt_{i+1}}f\right\|_{\gamma}|t-s|^{\gamma} \leqslant 2^{-n\gamma}\|Q\|_{D_2^{\gamma}}\|f\|_W|t-s|^{\gamma}$$

where we exploited that S is a contraction semigroup. Returning to the telescope sum, we obtain

$$|[Qf]_{ts}| \leqslant 2^{n(1-\gamma)}\|Q\|_{D_2^{\gamma}}\|f\|_W|t-s|^{\gamma}.$$

By passing to the limit for n which tends to infinity, we conclude that for any $f \in W$ and any $[s,t] \subset [0,T]$

$$[Qf]_{ts} = 0$$

yielding $Q = 0$, i.e. $\Lambda A = \widetilde{\Lambda}A$ for any $A \in D_3^{1+} \cap \hat{\delta}(D_2)$, concluding uniqueness.

Towards existence, let $A \in D_3^{1+} \cap \hat{\delta}(D_2)$, i.e. there exist a $B \in D_2$ and $\eta > 1$ such that $\hat{\delta}B = A \in D_3^{\eta}$. Let $(r_k^n)_{0 \leqslant k \leqslant 2^n}$ be the dyadic partition of $[s,t]$. We set, following [18]

$$M^n : D_3^{1+} \cap \hat{\delta}(D_2) \to D_2^{1+}$$

$$\hat{\delta}B \mapsto M^n\hat{\delta}B$$

where

$$[(M^n\hat{\delta}B)f]_{ts} := [B(f)]_{ts} - \sum_{k=0}^{2^n-1}[B(S_{tr_{k+1}^n}f)]_{r_{k+1}^n r_k^n}.$$

Note in particular that $[M^n\hat{\delta}f]_{ts} = 0$. We show that $(M^n\hat{\delta}B)_n$ is Cauchy in D_2^{η}. Note that

$$[(M^n\hat{\delta}B)f]_{ts} - [(M^{n+1}\hat{\delta}B)f]_{ts}$$

$$= \sum_{k=0}^{2^n-1}[B(S_{tr_{2k+2}^n}f)]_{r_{2k+2}^n r_{2k}^n} - [B(S_{tr_{2k+2}^n}f)]_{r_{2k+2}^n r_{2k+1}^n} - [B(S_{tr_{2k+1}^n}f)]_{r_{2k+1}^n r_{2k}^n}$$

$$= \sum_{k=0}^{2^n-1}[\delta B(S_{tr_{2k+2}^n}f)]_{r_{2k+2}^n r_{2k+1}^n r_{2k}^n} - \sum_{k=0}^{2^n-1}[\phi B(S_{tr_{2k+2}^n}f)]_{r_{2k+2}^n r_{2k+1}^n r_{2k}^n}$$

$$= \sum_{k=0}^{2^n-1}[\hat{\delta}B(S_{tr_{2k+2}^n}f)]_{r_{2k+2}^n r_{2k+1}^n r_{2k}^n} \leqslant (t-s)^{\eta}\sum_{k=0}^{2^n-1}\left\|\hat{\delta}B(S_{tr_{2k+2}^n}f)\right\|_{\eta}2^{-n\eta}$$

$$\leqslant (t-s)^{\eta}\left\|\hat{\delta}B\right\|_{D_3^{\eta}}\|f\|_W 2^{-n(\eta-1)}$$

From which we deduce that $(M^n\hat{\delta}B)_{n\in\mathbb{N}}$ is a Cauchy sequence in D_2^{η}, indeed

$$\left\|M^n\hat{\delta}B - M^{n+1}\hat{\delta}B\right\|_{D_2^{\eta}} \leqslant \left\|\hat{\delta}B\right\|_{D_3^{\eta}}2^{-n(\eta-1)}.$$

Let $\Lambda\hat{\delta}B \in D_2^{\eta}$ be its limit. By a telescope argument

$$\left\|M^n\hat{\delta}B\right\|_{D_2^{\eta}} = \left\|\sum_{k=0}^{n-1}M^k\hat{\delta}B - M^{k+1}\hat{\delta}B\right\|_{D_2^{\eta}} \leqslant \sum_{k=0}^{n-1}2^{-k(\eta-1)}\left\|\hat{\delta}B\right\|_{D_3^{\eta}} \leqslant C_{\eta}\left\|\hat{\delta}B\right\|_{D_3^{\eta}}$$

from which we obtain (B.1) using to weak-*-lower semicontinuity of the norm. One can prove that the limit does not depend on the particular sequence, see [18, Proposition 2.3].

Finally, let $u = 2^m$ for some $m \in 0, \ldots, n$ and note that

$$
\begin{aligned}
&[(\hat{\delta} M^n \hat{\delta} B) f]_{tus} \\
&= [(M^n \hat{\delta} B) f]_{ts} - [(M^n \hat{\delta} B) f]_{tu} - [(M^n \hat{\delta} B) f]_{us} - [(M^n \hat{\delta} B (S_{tu} - \mathrm{Id})) f]_{us} \\
&= [Bf]_{ts} - [Bf]_{tu} - [B S_{tu} f]_{us} + \sum_{k=0}^{2^n - 1} [B(S_{t r_{k+1}^n} f)]_{r_{k+1}^n r_k^n} \\
&\quad - \sum_{k=2^m}^{2^n - 1} [B(S_{t r_{k+1}^n} f)]_{r_{k+1}^n r_k^n} - \sum_{k=0}^{2^m - 1} [B(S_{u r_{k+1}^n} S_{tu} f)]_{r_{k+1}^n r_k^n} \\
&= [Bf]_{ts} - [Bf]_{tu} - [B S_{tu} f]_{us} \\
&= [\hat{\delta} B f]_{sut}
\end{aligned}
$$

from which we recover, in the limit $n \to \infty$, that $\hat{\delta}\Lambda = \mathrm{Id}|_{D_3^{1+} \cap \hat{\delta}(D_2)}$. $\qquad\square$

Corollary B.3. *Suppose that $A \in D_2$ is such that $\hat{\delta} A \in D_3^{1+}$. Then there exists $I \in D_1$ such that*

$$
\hat{\delta} I = \left(\mathrm{Id} - \Lambda \hat{\delta} \right) A,
$$

i.e. for every $f \in W$ and $(t,s) \in \Delta_2$, $\left[\hat{\delta} I f \right]_{ts} = [Af]_{ts} - \left[\Lambda \hat{\delta} A f \right]_{ts}$. In particular, if $A \in D_2^\mu$ with $\mu > 0$ and $\hat{\delta} A \in D_3^\eta$ with $\eta > 1$, then for every $f \in W$

$$
\left| [\hat{\delta} I f]_{ts} \right| \leqslant \left(\|A\|_{D_2^\mu} (t-s)^\mu + \left\| \hat{\delta} A \right\|_{D_3^\eta} (t-s)^\eta \right) \|f\|_W. \tag{B.2}
$$

Finally

$$
\left[\hat{\delta} I f \right]_{ts} = \lim_{|\mathcal{P}[s,t]| \downarrow 0} \sum_{[v,u] \in \mathcal{P}[s,t]} [A S_{tu} f]_{uv}, \tag{B.3}
$$

where the limit is over any partition of $[s,t]$ whose mesh tends to zero.

Proof. The proof is an easy application of the Sewing Lemma and the properties of $(D_*, \hat{\delta})$. Indeed, observe that $\hat{\delta}(Af - \Lambda \hat{\delta} A f) = 0$ for any $f \in W$, which means that $A - \Lambda \hat{\delta} A \in \mathrm{Ker}\, \hat{\delta}|_{D_2}$, and thus there exists $I \in D_1$ such that $\hat{\delta} I = A - \Lambda \hat{\delta} A$.

The estimate (B.2) follows from (B.1). Concerning (B.3), observe that for a partition $|\mathcal{P}[s,t]|$, using the properties of $\hat{\delta}$, one obtains

$$
\begin{aligned}
\sum_{[v,u] \in \mathcal{P}[s,t]} [A S_{tu} f]_{uv} &= \sum_{[v,u] \in \mathcal{P}[s,t]} \left[\hat{\delta} I S_{tu} f \right]_{uv} + \left[\Lambda \hat{\delta} A S_{tu} f \right]_{uv} \\
&= \left[\hat{\delta} I f \right]_{ts} + \sum_{[v,u] \in \mathcal{P}[s,t]} \left[\Lambda \hat{\delta} A S_{tu} f \right]_{uv}.
\end{aligned}
$$

By taking the limit for the mesh which tends to zero and using the fact that $\Lambda \hat{\delta} A \in D_2^{1+}$, the last sum converges to zero. $\qquad\square$

References

[1] R. A. Adams and J. J. F. Fournier. *Sobolev spaces*, volume 140 of *Pure and Applied Mathematics (Amsterdam)*. Elsevier/Academic Press, Amsterdam, second edition, 2003. MR-2424078

[2] F. Bechtold. Strong solutions of semilinear SPDEs with unbounded diffusion. arXiv:2001.08848, 2020.

[3] L. Bertini, G. Giacomin, and C. Poquet. Synchronization and random long time dynamics for mean-field plane rotators. *Probab. Theory Rel. Fields*, 160(3-4):593–653, 2014. MR-3278917

[4] P. Cardaliaguet, F. Delarue, J.-M. Lasry, and P.-L. Lions. *The master equation and the convergence problem in mean field games*, volume 201 of *Annals of Mathematics Studies*. Princeton University Press, Princeton, NJ, 2019. MR-3967062

[5] M. Coghi, J.-D. Deuschel, P. Friz, and M. Maurelli. Pathwise McKean-Vlasov Theory with Additive Noise. arXiv:1812.11773, 2018. MR-4149531

[6] F. Coppini. Long time dynamics for interacting oscillators on graphs. arXiv:1908.01520, 2019.

[7] G. Da Prato and J. Zabczyk. *Stochastic equations in infinite dimensions*, volume 152 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, Cambridge, second edition, 2014. MR-3236753

[8] V. H. de la Pena, M. J. Klass, and T. L. Lai. Self-normalized processes: exponential inequalities, moment bounds and iterated logarithm laws. *Ann. Probab.*, 32(3A):1902–1933, 2004. MR-2073181

[9] S. Delattre, G. Giacomin, and E. Luçon. A Note on Dynamical Models on Random Graphs and Fokker-Planck Equations. *J. Stat. Phys.*, 165(4):785–798, 2016. MR-3568168

[10] R. L. Dobrušin. Vlasov equations. *Funktsional. Anal. i Prilozhen.*, 13(2):48–58, 96, 1979. MR-0541637

[11] J. Feng and T. G. Kurtz. *Large Deviations for Stochastic Processes*, volume 131 of *Mathematical Surveys and Monographs*. American Mathematical Society, 2006. MR-2260560

[12] B. n. Fernandez and S. Méléard. A Hilbertian approach for fluctuations on the McKean-Vlasov model. *Stoch. Proc. Appl.*, 71(1):33–53, 1997. MR-1480638

[13] F. Flandoli, M. Leimbach, and C. Olivera. Uniform convergence of proliferating particles to the FKPP equation. *J. Math. Anal. Appl.*, 473(1):27–52, 2019. MR-3912811

[14] F. Flandoli, C. Olivera, and M. Simon. Uniform approximation of $2d$ Navier-Stokes equation by stochastic interacting particle systems. arXiv:2004.00458, 2020. MR-4169261

[15] P. K. Friz and M. Hairer. *A course on rough paths*. Universitext. Springer, Cham, 2014. MR-3289027

[16] S. E. Graversen and G. Peskir. Maximal inequalities for the Ornstein-Uhlenbeck process. *Proc. Amer. Math. Soc.*, 128(10):3035–3041, 2000. MR-1664394

[17] M. Gubinelli. Controlling rough paths. *J. Func. Anal.*, 216(1):86–140, 2004. MR-2091358

[18] M. Gubinelli and S. Tindel. Rough evolution equations. *Ann. Probab.*, 38(1):1–75, 01 2010. MR-2599193

[19] J. Gärtner. On the McKean-Vlasov Limit for Interacting Diffusions. *Math. Nachr.*, 137(1):197–248, 1988. MR-0968996

[20] D. Henry. *Geometric theory of semilinear parabolic equations*, volume 840 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin-New York, 1981. MR-0610244

[21] M. Hofmanová. Strong solutions of semilinear stochastic partial differential equations. *Nonlinear Differential Equations and Applications NoDEA*, 20(3):757–778, 2013. MR-3057153

[22] C. Jia and G. Zhao. Moderate maximal inequalities for the Ornstein-Uhlenbeck process. arXiv:1711.00902, 2017.

[23] V. N. Kolokoltsov. *Markov processes, semigroups, and generators*. Number 38 in De Gruyter studies in mathematics. De Gruyter, Berlin, 2011. MR-2780345

[24] V. Konarovskyi, T. Lehmann, and M. von Renesse. Dean-Kawasaki dynamics: ill-posedness vs. triviality. *Electron. Commun. Probab.*, 24(8), 2019. MR-3916340

[25] V. Konarovskyi, T. Lehmann, and M. von Renesse. On Dean-Kawasaki dynamics with smooth drift potential. *J. Stat. Phys.*, 178(3):666–681, 2020. MR-4059955

[26] E. Luçon and C. Poquet. Long time dynamics and disorder-induced traveling waves in the stochastic Kuramoto model. *Ann. Inst. Henri Poincaré Probab. Stat.*, 53(3):1196–1240, 2017. MR-3689966

[27] A. Lunardi. *Analytic semigroups and optimal regularity in parabolic problems*, volume 16 of *Progress in Nonlinear Differential Equations and their Applications*. Birkhäuser Verlag, Basel, 1995. MR-1329547

[28] M. Métivier. *Semimartingales*, volume 2 of *de Gruyter Studies in Mathematics*. Walter de Gruyter & Co., Berlin-New York, 1982. MR-0688144

[29] S. Mischler and C. Mouhot. Kac's Program in Kinetic Theory. *Inventiones Mathematicae*, 193(1):1–147, 2013. MR-3069113

[30] S. Mischler, C. Mouhot, and B. Wennberg. A new approach to quantitative propagation of chaos for drift, diffusion and jump processes. *Probability Theory and Related Fields*, 161(1-2):1–59, 2015. MR-3304746

[31] H. Neunzert. An introduction to the nonlinear Boltzmann-Vlasov equation. In *Kinetic theories and the Boltzmann equation (Montecatini, 1981)*, volume 1048 of *Lecture Notes in Mathematics*, pages 60–110. Springer, Berlin, 1984. MR-0740721

[32] K. Oelschläger. A Martingale Approach to the Law of Large Numbers for Weakly Interacting Stochastic Processes. *Ann. Probab.*, 12(2):458–479, 1984. MR-0735849

[33] A. Pazy. *Semigroups of linear operators and applications to partial differential equations*. Springer, New York, 1983. OCLC: 835128129. MR-0710486

[34] D. W. Stroock and S. R. S. Varadhan. *Multidimensional diffusion processes*, volume 233 of *Grundlehren der mathematischen Wissenschaften*. Springer, Berlin, New York, 2006. MR-0532498

[35] A.-S. Sznitman. Topics in propagation of chaos. In P.-L. Hennequin, editor, *Ecole d'Eté de Probabilités de Saint-Flour XIX — 1989*, pages 165–251. Springer, Berlin, Heidelberg, 1991. MR-1108185

[36] H. Tanaka. Limit Theorems for Certain Diffusion Processes with Interaction. In K. Itô, editor, *North-Holland Mathematical Library*, volume 32 of *Stochastic Analysis*, pages 469–488. Elsevier, 1984. MR-0780770

Electronic Journal of Probability
Electronic Communications in Probability

Advantages of publishing in EJP-ECP

- Very high standards

- Free for authors, free for readers

- Quick publication (no backlog)

- Secure publication (LOCKSS[1])

- Easy interface (EJMS[2])

Economical model of EJP-ECP

- Non profit, sponsored by IMS[3], BS[4] , ProjectEuclid[5]

- Purely electronic

Help keep the journal free and vigorous

- Donate to the IMS open access fund[6] (click here to donate!)

- Submit your best articles to EJP-ECP

- Choose EJP-ECP over for-profit journals

---

[1]LOCKSS: Lots of Copies Keep Stuff Safe http://www.lockss.org/
[2]EJMS: Electronic Journal Management System http://www.vtex.lt/en/ejms.html
[3]IMS: Institute of Mathematical Statistics http://www.imstat.org/
[4]BS: Bernoulli Society http://www.bernoulli-society.org/
[5]Project Euclid: https://projecteuclid.org/
[6]IMS Open Access Fund: http://www.imstat.org/publications/open.htm