QUASILINEAR SPDES VIA ROUGH PATHS

FELIX OTTO AND HENDRIK WEBER

ABSTRACT. We are interested in (uniformly) parabolic PDEs with a nonlinear dependance of the leading-order coefficients, driven by a rough right hand side. For simplicity, we consider a space-time periodic setting with a single spatial variable:

$$\partial_2 u - P(a(u)\partial_1^2 u - \sigma(u)f) = 0$$

where P is the projection on mean-zero functions, and f is a distribution and only controlled in the low regularity norm of $C^{\alpha-2}$ for $\alpha > \frac{2}{3}$ on the parabolic Hölder scale. The example we have in mind is a random forcing f and our assumptions allow, for example, for an f which is white in the time variable x_2 and only mildly coloured in the space variable x_1; any spatial covariance operator $(1 + |\partial_1|)^{-\lambda_1}$ with $\lambda_1 > \frac{1}{3}$ is admissible.

On the deterministic side we obtain a C^α-estimate for u, assuming that we control products of the form $v\partial_1^2 v$ and vf with v solving the constant-coefficient equation $\partial_2 v - a_0\partial_1^2 v = f$. As a consequence, we obtain existence, uniqueness and stability with respect to $(f, vf, v\partial_1^2 v)$ of small space-time periodic solutions for small data. We then demonstrate how the required products can be bounded in the case of a random forcing f using stochastic arguments.

For this we extend the treatment of the singular product $\sigma(u)f$ via a space-time version of Gubinelli's notion of controlled rough paths to the product $a(u)\partial_1^2 u$, which has the same degree of singularity but is more nonlinear since the solution u appears in both factors. The PDE ingredient mimics the (kernel-free) Krylov-Safanov approach to ordinary Schauder theory.

1. INTRODUCTION

We are interested in the parabolic PDE

$$(1) \qquad \partial_2 u - P(a(u)\partial_1^2 u - \sigma(u)f) = 0$$

for a rough driver f. The coefficients a, σ are assumed to be regular and uniformly elliptic, see (20) below for precise assumptions, and P is the projection on mean-zero functions. For the right hand side f we only assume control on the low regularity norm of $C^{\alpha-2}$ in the parabolic Hölder scale for $\alpha \in (\frac{2}{3}, 1)$ (see (19) for a precise statement). The optimal control on u one could aim to obtain under these assumption is in the C^α norm but in this regularity class there is no classical functional analytic definition of the singular products $a(u)\partial_1^2 u$ and $\sigma(u)f$. In this article we assume that we have an "off-line" interpretation for

several products such as $v\partial_2^2 v$, vf (see (111)), where v solves the constant coefficient equation $\partial_2 v - a_0 \partial_1^2 v = f$ and show that these bounds allow to control u. We are ultimately interested in a stochastic forcing f and in this case the required control of products can be obtained using explicit moment calculations to capture stochastic cancelations.

Our method is similar in spirit to Lyons' rough path theory [14, 13, 15]. This theory is based on the observation that the analysis of stochastic integrals

$$(2) \qquad \int_0^t u(s)dv(s)$$

for irregular v, such as Brownian motion or even lower regularity stochastic processes, can be conducted efficiently by splitting it into a stochastic and a deterministic step. In the stochastic step the integral (2) is defined for a single well-chosen function $\bar{u}$, e.g. v itself. In the case where $v = \bar{u}$ is a (multidimensional) Brownian motion there is a one-parameter family of canonical definitions for these integrals, with the Itô and the Stratonovich notions being the most prominent ones. Information on this single integral suffices to give a subordinate sense to integrals for a whole class of functions u with similar small-scale behaviour. This line of thought is expressed precisely in Gubinelli's notion of a controlled path [4, Definition 1]. There, a function u in the usual Hölder space C^α, $\alpha \in (\frac{1}{3}, \frac{1}{2})$, is said to be controlled by $\bar{u} \in C^\alpha$ if there exists a third function $\sigma \in C^\alpha$ such that for all $s, t \in \mathbb{R}$

$$(3) \qquad |u(t) - u(s) - \sigma(s)(\bar{u}(t) - \bar{u}(s))| \lesssim |t - s|^{2\alpha}.$$

Loosely speaking, this means that the *increments* $u(t) - u(s)$ of the function u can be approximated by those of $\bar{u}$, provided the latter are locally modulated by the amplitudes σ. In [4, Theorem 1] it is then shown that this assumption, together with a bound of the form

$$(4) \qquad \left| \int_s^t \bar{u}(r)dv(r) - \bar{u}(s)(v(t) - v(s)) \right| \lesssim |t - s|^{2\alpha},$$

suffices to define the integral $\int u(r)dv(r)$ and to obtain the bound

$$\left| \int_s^t u(r)dv(r) - u(s)(v(t) - v(s)) - \sigma(s) \int_s^t (\bar{u}(r) - \bar{u}(s))dv(r) \right|$$
$$(5) \qquad \lesssim |t - s|^{3\alpha}.$$

The construction of the integrals (4) for the specific function $\bar{u}$ can be accomplished under a less restrictive set of assumptions than required for the classical Itô theory. In many applications this construction can be carried out using Gaussian calculus without making reference to an underlying martingale structure. The construction makes very little use of the linear order of time and lends itself well to extensions to higher dimensional index sets.

This last point was the starting point for Hairer's work on singular stochastic PDE – the observation that the variable t in the rough path theory could represent "space" rather than "time" was the key insight that allowed to define stochastic PDEs with non-linearities of Burgers type [6] and the KPZ equation [7]. The notion of controlled path was also the starting point for his definition of regularity structures [8] which permits to treat semilinear stochastic PDE with an extremely irregular right hand side, possibly involving a renormalisation procedure. Parallel to that, Gubinelli, Imkeller and Perkowski put forward a notion of paracontrolled rough paths [5], a Fourier-analytic variant of (3) which has also been used to treat singular stochastic PDE.

In this article we propose yet another higher-dimensional generalisation of the notion of controlled path, see Definition 1 below, and use it to provide a solution and stability theory for (1). This definition is an immediate generalisation of Gubinelli's definition (3) and also closely related to Hairer's notion [8, Definition 3.1] of a modelled distribution in a certain regularity structure. However, the definition comes with a twist because the quasilinear nature of (1) forces us to allow the realisation of the *model*, $v(\cdot, a_0)$ in our notation, to depend on a parameter a_0, which ultimately corresponds to the variable diffusion coefficient $a(u)$. In our theory the "off-line products" vf and $v\partial_1^2 v$ play the role of the "off-line integral" $\int \bar{u} dv$ above and the regularity assumption (4) is translated into a control on the commutators

$$[v, (\cdot)_T] \diamond \{\partial_1^2 v, f\} := v(\{\partial_1^2 v, f\})_T - (v \diamond \{\partial_1^2 v, f\})_T,$$

where $(\cdot)_T$ denotes the convolution with a smooth kernel at scale T (see (17) and the discussion that follows it) and where we use the notation $\diamond$ to indicate that products are not classically defined and that their interpretations have to be specified. Furthermore, here and below we use the abbreviated notation $[v, (\cdot)_T] \diamond \{\partial_1^2 v, f\}$ when we speak about $[v, (\cdot)_T] \diamond \partial_1^2 v$ and $[v, (\cdot)_T] \diamond f$ simultaneously. Based on these assumptions we derive bounds in the spirit of (5) on the singular products $a(u) \diamond \partial_1^2 u$ and $\sigma(u) \diamond f$ (see Lemma 2 and 4) which can also be seen as a (simpler) variant of Hairer's Reconstruction Theorem [8, Theorem 3.10]. We want to point out that our method completely avoids the use of wavelet analysis which features prominently in Hairer's proof of the Reconstruction Theorem. On the PDE side, in Lemma 5, we obtain an optimal regularity result on solution u of (1) based on a control of the commutators $[a, (\cdot)_T] \diamond \partial_1^2 u$ and $[\sigma, (\cdot)_T] \diamond f$. This result is similar in spirit to Hairer's Integration Theorem [8, Theorem 5.12]. Our proof mimics the Krylov-Safanov approach to Schauder theory [12] and therefore does not make reference to a parabolic heat kernel. The main deterministic results, Proposition 1 and Theorem 2 combine these ingredients to obtain existence and uniqueness results for the linear version of (1) (i.e. a and σ do not depend on u, Proposition 1) and for

(1) under a small data assumption (Theorem 2). We want to point out that the deterministic analysis does not depend on the assumption of a $1+1$ dimensional space and would go through completely unchanged if $\partial_2 - a(u)\partial_1$ were replaced by a uniformly parabolic operator over $\mathbb{R}^n \times \mathbb{R}$.

On the stochastic side, we consider a class of stationary Gaussian distributions f of class $C^{\alpha-2}$. This class includes, for example, the case where f is "white" in the time-like variable x_2 and has covariance operator $(1 + |\partial_1|)^{-\lambda_1}$ for $\lambda_1 > \frac{1}{3}$ in the x_1 variable, or the case where the noise is constant in the time-like variable x_2 and has covariance operator $(1 + |\partial_1|)^{-\lambda_1}$ for $\lambda_1 > -\frac{5}{3}$ for the x_1 variable (see the end of Section 3 for a more detailed discussion of admissible f). For such f we construct the generalized products $v \diamond \partial_1^2 v$ and $v \diamond f$ as limits of renormalized smooth approximations: More precisely, let ψ' be an arbitrary Schwartz function with $\int \psi' = 1$ and set $\psi_\varepsilon(x_1, x_2) = \frac{1}{\varepsilon^{\frac{3}{4}}} \psi_1\big(\frac{x_1}{\varepsilon^{\frac{1}{4}}}, \frac{x_2}{\varepsilon^{\frac{1}{2}}}\big)$. Then we set $f_\varepsilon = f * \psi'_\varepsilon$, let v_ε solve $\partial_2 v_\varepsilon - a_0 \partial_1^2 v_\varepsilon = f_\varepsilon$ and construct $v \diamond f$ and $v \diamond \partial_1^2 v$ as

$$v(\cdot, a_0) \diamond f := \lim_{\varepsilon \to 0} \big(v_\varepsilon(\cdot, a_0) f_\varepsilon - \big\langle v_\varepsilon(\cdot, a_0) f_\varepsilon \big\rangle \big),$$

$$v(\cdot, a_0) \diamond \partial_1^2 v(\cdot, a_0') := \lim_{\varepsilon \to 0} \big(v_\varepsilon(\cdot, a_0) \partial_1^2 v_\varepsilon(\cdot, a_0')$$

$$(6) \hspace{5cm} - \big\langle v_\varepsilon(\cdot, a_0) \partial_1^2 v_\varepsilon(\cdot, a_0') \big\rangle \big),$$

see Proposition 2 below. In many of the examples we consider, the expectations of the regularized products $\mathfrak{c}^{(1)}(\varepsilon, a_0) = \big\langle v_\varepsilon(\cdot, a_0) f_\varepsilon \big\rangle$ and $\mathfrak{c}^{(2)}(\varepsilon, a_0, a_0') = \big\langle v_\varepsilon(\cdot, a_0) \partial_1^2 v_\varepsilon(\cdot, a_0') \big\rangle$ diverge as ε goes to zero (the precise form of these constants is given in Lemma 7). The renormalization procedure can be avoided if f satisfies the additional stronger regularity assumption (144) which holds, for example, if f is "white" in x_1 and "trace-class" in x_2.

Finally, the construction of these renormalized products and the deterministic well-posedness theory can be combined to the following main theorem:

Theorem 1. *Let the coefficients a, σ satisfy the regularity assumptions (20).*
Let f be a centered, one-periodic, stationary Gaussian random distribution satisfying the regularity assumption (129) and let $f_\varepsilon = f * \psi'_\varepsilon$ be as described above. Denote by $v(\cdot, a_0)$ (resp. $v_\varepsilon(\cdot, a_0)$) the one-periodic mean-free solutions of $(\partial_2 - a_0 \partial_1^2)v = Pf$ (resp. $(\partial_2 - a_0 \partial_1^2)v_\varepsilon = Pf_\varepsilon$).

(i) The renormalized commutators $[v, (\cdot)_T] \diamond f$ and $[v, (\cdot)_T] \diamond \partial_1^2 v$ defined via the limits (6) exist and satisfy the bounds

$$(7) \qquad \left\langle \left(\sup_{T \leq 1} (T^{\frac{1}{4}})^{2-\alpha'} \| f_T \| \right)^p \right\rangle^{\frac{1}{p}} \lesssim 1,$$

$$(8) \qquad \left\langle \left(\sup_{T \leq 1} \sup_{a_0 \in [\lambda, 1]} (T^{\frac{1}{4}})^{2-2\alpha'} \| [v(\cdot, a_0), (\cdot)_T] \diamond f \| \right)^p \right\rangle^{\frac{1}{p}} \lesssim 1,$$

$$(9) \quad \left\langle \left(\sup_{T \leq 1} \sup_{a_0 \in [\lambda, 1]} \sup_{a_0' \in [\lambda, 1]} (T^{\frac{1}{4}})^{2-2\alpha'} \| [v(\cdot, a_0), (\cdot)_T] \diamond \partial_1^2 v(\cdot, a_0') \| \right)^p \right\rangle^{\frac{1}{p}} \lesssim 1,$$

as well as analogous bounds if v is replaced by its derivatives with respect to a_0, a_0'. The convergences of the renormalized products take place almost surely and in every stochastic L^p space with respect to the norm $\sup_{\lambda < a_0 \leq 1} \sup_{\lambda < a_0' \leq 1} \| \cdot \|_{C^{\alpha-2}}$. Here $\| f \| = \sup_{x \in \mathbb{R}^2} |f(x)|$ denotes the supremum norm and a norm for $C^{\alpha-2}$ is defined in (19) below.

(ii) There exists a random constant $\eta > 0$ satisfying

$$(10) \qquad \langle \eta^{-p} \rangle^{\frac{1}{p}} \lesssim 1$$

for all $p < \infty$, such that ηf as well as the commutators $\eta^2 [v(\cdot, a_0), (\cdot)_T] \diamond \{f, \partial_1^2 v\}$ derived from ηf satisfy the smallness Assumptions (110) and (111). Therefore, there exists a unique mean-free function u with the properties

$$u \text{ is modelled after } v \text{ according to } a(u) \text{ and } \sigma(u)$$
$$(11) \qquad \qquad \qquad \text{in the sense of Definition 1,}$$
$$(12) \qquad \partial_2 u - P(a(u) \diamond \partial_1^2 u + \sigma(u) \diamond \eta f) = 0 \quad \text{distributionally,}$$

under the smallness condition

$$(13) \qquad [u]_\alpha \ll 1.$$

(iii) Let $\eta > 0$ be as in part (ii). The regularised noise terms ηf_ε as well as the corresponding renormalized commutators

$$\eta^2 [v_\varepsilon(\cdot, a_0), (\cdot)_T] f_\varepsilon(\cdot, a_0') - \eta^2 \mathfrak{c}^{(1)}(\varepsilon, a_0)$$
$$\eta^2 [v_\varepsilon(\cdot, a_0), (\cdot)_T] \partial_1^2 v_\varepsilon(\cdot, a_0') - \eta^2 \mathfrak{c}^{(2)}(\varepsilon, a_0, a_0')$$

almost surely satisfy the smallness assumptions (110) and (111) uniformly in ε. Denote by u_ε the unique solutions of (11) – (114) with ηf replaced by ηf_ε. Then these u_ε converge to u almost surely with respect to the C^α norm. Furthermore, u_ε is a classical solution to the renormalized PDE

$$\partial_2 u_\varepsilon - P(a(u) \partial_1^2 u_\varepsilon - a'(u_\varepsilon) \sigma(u_\varepsilon)^2 \mathfrak{c}^{(2)}(\cdot, a(u_\varepsilon), a(u_\varepsilon))$$
$$(14) \qquad + \sigma(u_\varepsilon) f_\varepsilon - \sigma'(u_\varepsilon) \sigma(u_\varepsilon) \mathfrak{c}^{(1)}(\cdot, a(u_\varepsilon))) = 0.$$

(iv) If f satisfies the additional assumption (144), *then the above statements hold true without renormalization, i.e. setting* $\mathfrak{c}^{(1)} = \mathfrak{c}^{(2)} = 0$.

PROOF OF THEOREM 1.
The bound (7) is proved as (130) in Lemma 6, and the bounds (8) and (9) as well as the convergence almost surely and in every stochastic L^p space are proved in Proposition 2. For (ii) we set

$$\eta^{-1} \gg \sup_{a_0, a_0' \in [\lambda, 1]} \sup_{\varepsilon \in [0,1]} \sup_{T \leq 1} (T^{\frac{1}{4}})^{2-2\alpha}$$

$$\|\{1, \frac{\partial}{\partial a_0'}, \frac{\partial^2}{\partial a_0'^2}\}\{1, \frac{\partial}{\partial a_0}, \frac{\partial^2}{\partial a_0^2}\}[v_\varepsilon, (\cdot)_T] \diamond \{f_\varepsilon, \partial_1^2 v_\varepsilon\}\|,$$

then ηf_ε, $\eta^2[v_\varepsilon, (\cdot)_T] f_\varepsilon$ etc. satisfy the smallness condition (110) and (111) uniformly in ε. The bound (10) is a consequence of Proposition 2 and the conclusion of (ii) is contained in part (i) of Theorem 2. For part (iii), we have already seen that ηf_ε, $\eta^2[v_\varepsilon, (\cdot)_T] f_\varepsilon$ etc. satisfy the smallness assumptions (110) and (111) uniformly in ε, and the convergence of the u_ε to u follows from a combination of Lemma 6, Proposition 2 and Theorem 2 (ii). The form of (14) follows from Corollary 3. Finally, part (iv) follows in the same way only replacing Proposition 2 by Corollary 5.

We finally mention that one week before posting this second version of our result, the article [2] was posted on the arXiv. In this article Furlan and Gubinelli study the equation

$$\partial_t u - a(u)\Delta u = \xi,$$

where $u = u(t, x)$ for x taking values in the two-dimensional torus, and $\xi = \xi(x)$ is a white noise over the two-dimensional torus, which is constant in the time variable t. This noise term ξ is of class C^{-1-} and therefore essentially behaves like our term f. They also define a notion of solution and prove short time existence and uniqueness of solutions for the initial value problem, as well as convergence for renormalized approximations similar to (14). Similar to the approach we present here, they locally approximate the solutions u by a family of solutions to constant coefficient problems. Their approach then deviates from ours and they implement their theory in the framework of paracontrolled distributions.

2. DETERMINISTIC ANALYSIS

2.1. **Setup.** METRIC. The parabolic operator $\partial_2 - a_0 \partial_1^2$ and its mapping properties on the scale of Hölder spaces (i.e. Schauder theory) imposes its intrinsic (Carnot-Carathéodory) metric, which is given by

$$(15) \qquad d(x, y) = |x_1 - y_1| + \sqrt{|x_2 - y_2|},$$

see for instance [12, Section 8.5]. The Hölder semi norm $[\cdot]_\alpha$ is defined based on (15):

$$(16) \qquad [u]_\alpha := \sup_{x \neq y} \frac{|u(x) - u(y)|}{d^\alpha(x, y)}.$$

CONVOLUTION. In order to define negative norms of distributions in the intrinsic way, it is convenient to have a family $\{(\cdot)_T\}_{T>0}$ of mollification operators $(\cdot)_T$ consistent with the relative scaling $(x_1, x_2) = (\ell \hat{x}_1, \ell^2 \hat{x}_2)$ of the two variables dictated by (15). It will turn out to be extremely convenient to have in addition the semi-group property

$$(17) \qquad (\cdot)_T (\cdot)_t = (\cdot)_{T+t}.$$

All is achieved by convolution with the semi-group $\exp(-T(\partial_1^4 - \partial_2^2))$ of the elliptic operator $\partial_1^4 - \partial_2^2$, which is the simplest positive operator displaying the same relative scaling between the variables as $\partial_2 - \partial_1^2$ and being symmetric in x_2 next to x_1. We note that the corresponding convolution kernel ψ_T is easily characterized by its Fourier transform $\hat{\psi}_T(k) = \exp(-T(k_1^4 + k_2^2))$; since the latter is a Schwartz function, also ψ_T is a Schwartz function. The only two (minor) inconveniences are that 1) the x_1-scale is played by $T^{\frac{1}{4}}$ (in line with (15) the x_2-scale is played by $T^{\frac{1}{2}}$) since we have $\psi_T(x_1, x_2) = \frac{1}{T^{\frac{3}{4}}} \psi_1(\frac{x_1}{T^{\frac{1}{4}}}, \frac{x_2}{T^{\frac{1}{2}}})$ and that 2) ψ_1 (and thus ψ_T) does not have a sign. The only properties of the kernel we need are moments of derivatives:

$$(18) \qquad \begin{array}{rcl} \int dy |\partial_1^k \psi_T(x-y)| d^\alpha(x, y) & \lesssim & (T^{\frac{1}{4}})^{-k+\alpha} \quad \text{and} \\ \int dy |\partial_2^k \psi_T(x-y)| d^\alpha(x, y) & \lesssim & (T^{\frac{1}{4}})^{-2k+\alpha} \end{array}$$

for all orders of derivative $k = 0, 1, \cdots$ and moment exponents $\alpha \geq 0$, as well as the fact that $\int \psi(x) x_1 dx = 0$. Estimates (18) follow immediately from the scaling and the fact that ψ_1 is a Schwartz function. In Lemma 11 we show however, that our main regularity assumption (19) on f as well as the bounds on the commutators do not depend on the specific choice of Schwartz kernel ψ. In particular, the statements ultimately do not depend on the semi-group property although this property plays an important part in the proofs.

FINITE DOMAIN. We mimic a finite domain by imposing periodicity in both directions; w.l.o.g. we may set this scale equal to one. We will typically measure the size of the distribution f by the expression

$$(19) \qquad \sup_{T \leq 1} (T^{\frac{1}{4}})^{2-\alpha} \|f_T\|,$$

where the restriction $T \leq 1$ reflects the period unity. With Lemma 9, cf Step 1, we have that this expression agrees with the standard definition of the norm of $C^{\alpha-2}$.

STANDING ASSUMPTIONS ON THE NONLINEARITIES. There exists a constant $\lambda > 0$ such that

$$
(20) \qquad
\begin{aligned}
a \in [\lambda, \tfrac{1}{\lambda}], & \quad \|a'\|, \|a''\|, \|a'''\| \leq \tfrac{1}{\lambda}, \\
\sigma \in [-1, 1], & \quad \|\sigma'\|, \|\sigma''\|, \|\sigma'''\| \leq \tfrac{1}{\lambda}.
\end{aligned}
$$

We express the bound on the various norms of a and σ by the ellipticity contrast λ in order to have a single constant that measures the quality of the data. Note that the assumption $\sigma \in [-1, 1]$ is only seemingly stronger than $\|\sigma\| \lesssim \tfrac{1}{\lambda}$ since that constant can always be absorbed into the rhs f in the equation. These fairly high regularity assumptions intervene in the proof of Lemma 1, they could be slightly weakened in the sense of [4, Proposition 4], at the expense of a more complicated notation. Here and in the entire deterministic section $\lesssim$ means $\leq C$ with a constant C only depending on λ and the exponent α.

2.2. **Definitions and results.** The following central definition is a straightforward generalization of Gubinelli's definition [4, Definition 1] of a "controlled path", a generalization from the time variable x_2 to multiple variables x, and to a "model" $(v_1, \cdots, v_I)$ (in the language of Hairer [8]) that here may depend on an additional parameter a_0. It states that the *increments* $u(y) - u(x)$ of the function u can be approximated by those of several functions v_i, if the latter are locally modulated by the amplitudes σ_i and the functions a_i that locally determine the value of the parameter a_0. The functions σ_i can therefore be interpreted as "derivatives" of u wrt v_i. The increments of the linear function x_1 also have to be included because of $\alpha > \tfrac{1}{2}$. In fact, since $2\alpha > 1$, given the model $(v_1, \cdots, v_I)$ (as modulated by the functions a_i), the "derivatives" $(\sigma_1, \cdots, \sigma_I)$ and ν determine u up to a constant. In our situation, we expect u and $(v_1, \cdots, v_I)$ to be Hölder continuous with exponent not (much) larger than α, so that imposing closeness of the increments to order 2α contains valuable additional information.

Definition 1. *Let $\tfrac{1}{2} < \alpha < 1$ and $I \in \mathbb{N}$. We say that a function u is modelled after the functions $(v_1, \cdots, v_I)$ of (x, a_0) according to the functions $(a_1, \cdots, a_I)$ and $(\sigma_1, \cdots, \sigma_I)$ provided there exists a function ν (which because of $2\alpha > 1$ is easily seen to be unique) such that*

$$
M := \sup_{x \neq y} \frac{1}{d^{2\alpha}(y, x)}
$$

$$
(21) \quad |u(y) - u(x) - \sigma_i(x)(v_i(y, a_i(x)) - v_i(x, a_i(x))) - \nu(x)(y - x)_1|
$$

is finite. Here and in the sequel we use Einstein's convention of summation over repeated indices.

Note that imposing (21) also for distant points x and y is consistent with periodicity despite the non-periodic term $(y - x)_1$ since by $\alpha \geq \tfrac{1}{2}$ the latter is dominated by $d^{2\alpha}(x, y)$ for $d(x, y) \geq 1$. Note also that (21) is reminiscent of a Hölder norm: In case of $(\sigma_1, \cdots, \sigma_I) = 0$,

the finiteness of (21) implies that u is continuously differentiable in x_1 and that $\nu(x) = \partial_1 u(x)$ so that M turns into the parabolic $C^{2\alpha}$-norm of u. In this spirit, Step 1 in the proof of Lemma 2 shows that the modelledness constant M in (21) controls the $(2\alpha - 1)$-Hölder norm of ν, provided $x \mapsto \sigma_i(x)v_i(\cdot, a_i(x))$ is α-Hölder continuous with values in C^α. In addition, in the presence of periodicity, M also controls the α-Hölder norm of u and the supremum norm of ν, which are of lower order, cf Step 2 in the proof of Lemma 2.

The following lemma shows that the notion of modelledness in Definition 1 is well-behaved under sufficiently smooth nonlinear *pointwise* transformation; it will be used in the proof of Theorem 2. It is essentially identical to [4, Proposition 4], which in turn is a consequence of Taylor's formula; because of the minor modifications due to the presence of a more general model, we reproduce the proof.

Lemma 1. *i) Suppose that u is modelled after v according to a and σ with constant M. Let the function b be twice differentiable. Then $b(u)$ is modelled after v according to a and $\mu := b'(u)\sigma$ with constant $\tilde{M}$ estimated by*

$$\text{(22)} \qquad \tilde{M} + [b(u)]_\alpha \leq (\|b'\| + \|b''\|[u]_\alpha)(M + [u]_\alpha),$$

$$\text{(23)} \qquad [\mu]_\alpha + \|\mu\| \leq (\|b'\| + \|b''\|[u]_\alpha)([\sigma]_\alpha + \|\sigma\|).$$

ii) Suppose that for $i = 0, 1$, u_i is modelled after v_i according to a_i and σ_i with constant M_i. Suppose further that $u_1 - u_0$ is modelled after (v_1, v_0) according to (a_1, a_0) and $(\sigma_1, -\sigma_0)$ with constant δM. Let the function b be three times differentiable. Then $b(u_1) - b(u_0)$ is modelled after (v_1, v_0) according to (a_1, a_0) and $(\mu_1 := b'(u_1)\sigma_1, -\mu_0 := -b'(u_0)\sigma_0)$ with constant $\delta\tilde{M}$ estimated by

$$\delta\tilde{M} + [b(u_1) - b(u_0)]_\alpha + \|b(u_1) - b(u_0)\|$$
$$\leq (\|b'\| + \|b''\|(\max_i M_i + \max_i[u_i]_\alpha) + \|b'''\|(\max_i[u_i])^2)$$
$$\text{(24)} \qquad \times (\delta M + [u_1 - u_0]_\alpha + \|u_1 - u_0\|),$$
$$[\mu_1 - \mu_0]_\alpha + \|\mu_1 - \mu_0\|$$
$$\leq (\|b'\| + \|b''\|\max_i[u_i]_\alpha)([\sigma_1 - \sigma_0]_\alpha + \|\sigma_1 - \sigma_0\|)$$
$$+ (\|b''\| + \|b'''\|\max_i[u_i]_\alpha)\max_i([\sigma_i]_\alpha + \|\sigma_i\|)$$
$$\text{(25)} \qquad \times ([u_1 - u_0]_\alpha + \|u_1 - u_0\|).$$

As discussed in the introduction, the main challenge in solving stochastic ordinary differential equations is to give a sense to integrals of the form (2). In the spirit of Hairer [8] we interpret this problem as giving a meaning to the *product* $u\partial_t v$, which does not have a canonical functional analytic definition because both u and v are only Hölder continuous in the time variable t of exponent less than $\frac{1}{2}$, because they

behave like Brownian motion. In view of the parabolic scaling, we encounter the same difficulty when giving a distributional sense to $b \diamond \partial_1^2 u$ when b and u are only Hölder continuous of exponent $\alpha < 1$ (from now we use the non-standard notation $b \diamond \partial_1^2 u$ instead of $b \, \partial_1^2 u$ to indicate that the definition of this product is non-standard).

As discussed in the introduction a main insight of Lyons' theory of rough paths, was the observation that such products can be defined provided u is controlled by $\bar{u}$ and the off-line product $\bar{u} \partial_t v$ satisfies the bound (4), which can be rewritten as $\int_s^t (\bar{u}(r) - \bar{u}(s)) \diamond \partial_r v(r)$ $= \bar{u}(s) \int_s^t \partial_r v(r) - \int_s^t \bar{u} \diamond \partial_r v =: -([\bar{u}, \int^t] \diamond f)(s)$, that is, the expression on both sides of (4) amount to a commutator $[\bar{u}, \int^t]$ of multiplication with $\bar{u}$ and integration, applied to a distribution $\partial_r v$. In our multi-dimensional framework, we replace integration $\frac{1}{t-s} \int_s^t$ by (smooth) averaging:

$$(26) \qquad\qquad [v, (\cdot)_T] \diamond f := v f_T - (v \diamond f)_T.$$

In our set up, the role of the crucial "algebraic relationship" [4, (24)] from rough path theory is played by the following straightforward consequence of the semi-group property (17)

$$(27) \qquad\quad [v, (\cdot)_{t+T}] \diamond f - ([v, (\cdot)_T] \diamond f)_t = [v, (\cdot)_t] f_T,$$

cf (263) in the proof of Lemma 2. Note that it is (only the control of) $[v, (\cdot)_T] \diamond f$ that relates the distribution $v \diamond f$ to the function v and the distribution f.

For our quasilinear SPDE, we need to give a sense to the *two* singular products $\sigma(u) \diamond f$ and $a(u) \diamond \partial_1^2 u$, so in particular to products of the form $u \diamond f$ and $b \diamond \partial_1^2 u$, where u and b behave like the solution v of $(\partial_2 - a_0 \partial_1^2) v = f$. Hence we will need the two off-line products $v \diamond f$ and $v \diamond \partial_1^2 v$. For simplicity, we split the argument into Lemma 3 dealing with the first and Lemma 4 with the second factor in the singular products. We will use Lemma 3, or rather Corollary 1, in order to pass from the definition of $v \diamond f$ and $v \diamond \partial_1^2 v$ to the definition of $u \diamond f$ and $b \diamond \partial_1^2 v$, respectively (since the distribution $\partial_1^2 v$ plays a role very similar to f, the lemma and the corollary are formulated in the notation of the former case). We will then use Lemma 4 to pass from $b \diamond \partial_1^2 v$ to $b \diamond \partial_1^2 u$.

Lemmas 2, 3, and 4, reveal a clear hierarchy of norms and measures of size:

- Functions u are measured in terms of the Hölder semi-norm $[u]_\alpha$ (the supremum norm $\|\sigma\|$ of a function σ only intervenes in scaling-wise suboptimal estimates like (60) that rely on the periodicity or the constraint $T \leq 1$ providing a large-scale cut-off, otherwise just as part of the product $\|\sigma\| [a]_\alpha$ with the Hölder norm of a),

- distributions are measured in the $C^{\alpha-2}$-norm $\sup_{T\leq 1} (T^{\frac{1}{4}})^{2-\alpha}$ $\|f_T\|$, see Step 1 in the proof of Lemma 9 for this equivalence of norms,
- commutators $[u, (\cdot)_T] \diamond f$ are measured on level $2\alpha - 2 < 0$ via $\sup_{T\leq 1} (T^{\frac{1}{4}})^{2-2\alpha} \|[u, (\cdot)_T] \diamond f\|$, and
- differences $[u, (\cdot)_T] \diamond f - [v, (\cdot)_T] \diamond f$ of commutators, like in case of the rough path expression (5) divided by $(t-s)$, are measured on level $3\alpha - 2 > 0$ via $\sup_{T\leq 1} (T^{\frac{1}{4}})^{2-3\alpha} \|[u, (\cdot)_T] \diamond f - [v, (\cdot)_T] \diamond f\|$, see (32) of Lemma 2.

Equipped with this dictionary, Lemmas 3 and 4 can be seen to be very close to [4, Theorem 1]; in particular, (32) in Lemma 2 is very close to (28) in [4, Corollary 3]. The major difference is the multi-dimensional extension through (26). A minor difference coming from the parabolic nature is the appearance of the commutator $[x_1, (\cdot)_T]f$, which however is regular, cf Lemma 10. A further minor difference arises from the a_0-dependence of the model v and the related appearance of the function a, which necessitates control of $\frac{\partial}{\partial a_0}$-derivatives of the functions and the commutators and manifests itself via the evaluation operator E. However, this minor difference can be embedded into the more general form of the upcoming Lemma 2.

Lemma 2. *Let* $\frac{2}{3} < \alpha < 1$. *Suppose we have a family of functions* $\{v(\cdot, x)\}_x$ *of class* C^α, *parameterized by points* x, *a distribution* f, *and a family of distributions* $\{v(\cdot, x) \diamond f\}_x$, *both of class* $C^{\alpha-2}$, *satisfying*

$$[v(\cdot, x) - v(\cdot, x')]_\alpha \leq N d^\alpha(x, x'), \tag{28}$$

$$\sup_{T\leq 1}(T^{\frac{1}{4}})^{2-\alpha}\|f_T\| \leq N_1, \tag{29}$$

$$\sup_{T\leq 1}(T^{\frac{1}{4}})^{2-2\alpha}\|[v(\cdot, x), (\cdot)_T] \diamond f$$
$$-[v(\cdot, x'), (\cdot)_T] \diamond f\| \leq N N_1 d^\alpha(x, x') \tag{30}$$

for all pairs of points x, x' *for some constants* N, N_1. *Suppose we are given a function* u *such that*

$$|(u(y) - u(x)) - (v(y, x) - v(x, x))$$
$$-\nu(x)(y - x)_1| \leq M d^{2\alpha}(y, x) \tag{31}$$

for all pairs of points y, x *for some constant* M *and some function* ν. *Then there exists a unique distribution* $u \diamond f$ *such that*

$$\sup_{T\leq 1}(T^{\frac{1}{4}})^{2-3\alpha}\|[u, (\cdot)_T] \diamond f$$
$$- E[v, (\cdot)_T] \diamond f - \nu[x_1, (\cdot)_T]f\| \lesssim (M + N)N_1, \tag{32}$$

where E *stands for the evaluation of the continuous function* $(x, y) \mapsto ([v(\cdot, x), (\cdot)_T] \diamond f)(y)$ *at* $y = x$.

If moreover all functions and distributions are 1-periodic and we use the constant N to also estimate the lower-order expressions

$$(33) \qquad\qquad\qquad [v(\cdot, x)]_\alpha \leq N,$$

$$(34) \qquad\qquad \sup_{T \leq 1}(T^{\frac{1}{4}})^{2-2\alpha}\|[v(\cdot, x), (\cdot)_T] \diamond f\| \leq NN_1,$$

for all points x then also

$$(35) \qquad\qquad \sup_{T \leq 1}(T^{\frac{1}{4}})^{2-2\alpha}\|[u, (\cdot)_T] \diamond f\| \leq (M + N)N_1.$$

Equipped with Lemma 2, the upcoming Lemma 3 is more of a corollary that specifies the form of the model. The general form of Lemma 2 is in particular convenient for part ii) of Lemma 3, where the Lipschitz continuity of the product $\sigma \diamond f$ in terms of the off-line product $v \diamond f$ and the modulating property (both constant and modulating functions) is established.

Lemma 3. *Let $\frac{2}{3} < \alpha < 1$. Suppose we are given a distribution f with*

$$(36) \qquad\qquad \sup_{T \leq 1}(T^{\frac{1}{4}})^{2-\alpha}\|f_T\| \leq N_1$$

for some constant N_1.
i) We consider a family of functions $\{v(\cdot, a_0)\}_{a_0}$ and a family of distributions $\{v(\cdot, a_0) \diamond f\}_{a_0}$ satisfying

$$(37) \qquad\qquad\qquad \sup_{a_0}[\{1, \frac{\partial}{\partial a_0}\}v]_\alpha \leq N_0,$$

$$(38) \qquad \sup_{T \leq 1}(T^{\frac{1}{4}})^{2-2\alpha} \sup_{a_0}\|\{1, \frac{\partial}{\partial a_0}\}[v, (\cdot)_T] \diamond f\| \leq N_1 N_0$$

for some constant N_0. We are given a function u modelled after v according to the α-Hölder functions a and σ with constant M and ν as in (21). Then there exists a unique distribution $u \diamond f$ such that

$$(39) \qquad \lim_{T \downarrow 0}\|[u, (\cdot)_T] \diamond f - \sigma E[v, (\cdot)_T] \diamond f - \nu[x_1, (\cdot)_T]f\| = 0,$$

where E evaluates a function of (x, a_0) at $(x, a(x))$. Furthermore, in case of

$$(40) \qquad\qquad [\sigma]_\alpha \leq 1, \quad [a]_\alpha \leq 1 \quad and \quad \|\sigma\| \leq 1$$

and when all functions are 1-periodic we have the sub-optimal estimate

$$(41) \qquad\qquad \sup_{T \leq 1}(T^{\frac{1}{4}})^{2-2\alpha}\|[u, (\cdot)_T] \diamond f\| \lesssim N_1(M + N_0).$$

ii) We consider two families of functions $\{v_i(\cdot, a_0)\}_{a_0}$, $i = 0, 1$, and two families of distributions $\{v_i(\cdot, a_0) \diamond f\}_{a_0}$ satisfying

$$(42) \qquad \sup_{a_0}[\{1, \frac{\partial}{\partial a_0}, \frac{\partial^2}{\partial a_0^2}\}v_i(\cdot, a_0)]_\alpha \leq N_0,$$

$$(43) \qquad \sup_{a_0}[\{1, \frac{\partial}{\partial a_0}\}(v_1 - v_0)(\cdot, a_0)]_\alpha \leq \delta N_0,$$

$$(44) \qquad \sup_{T \leq 1}(T^{\frac{1}{4}})^{2-2\alpha} \sup_{a_0} \|\{1, \frac{\partial}{\partial a_0}, \frac{\partial^2}{\partial a_0^2}\}[v_i, (\cdot)_T] \diamond f\| \leq N_1 N_0,$$

(45)

$$\sup_{T \leq 1}(T^{\frac{1}{4}})^{2-2\alpha} \sup_{a_0} \|\{1, \frac{\partial}{\partial a_0}\}([v_1, (\cdot)_T] \diamond f - [v_0, (\cdot)_T] \diamond f)\| \leq N_1 \delta N_0$$

for some constants N_0 and δN_0. Suppose the function δu is modelled after (v_1, v_0) according to the α-Hölder functions (a_1, a_0) and $(\sigma_1, -\sigma_0)$ with δM and $\delta \nu$ in analogy to (21). Then there exists a unique distribution $\delta u \diamond f$ such that

$$\lim_{T \downarrow 0} \|[\delta u, (\cdot)_T] \diamond f$$

$$(46) \qquad - (\sigma_1 E_1[v_1, (\cdot)_T] \diamond f - \sigma_0 E_0[v_0, (\cdot)_T] \diamond f) - \delta \nu[x_1, (\cdot)_T] f\| = 0,$$

where E_i denotes the operator that evaluates a function of (x, a_0) at $a_0 = a_i(x)$. Furthermore, in case of (40) and when all functions are 1-periodic we have the sub-optimal estimate

$$\sup_{T \leq 1}(T^{\frac{1}{4}})^{2-2\alpha}\|[\delta u, (\cdot)_T] \diamond f\| \lesssim N_1\big(\delta M$$

$$(47) \qquad + N_0([\sigma_1 - \sigma_0]_\alpha + \|\sigma_1 - \sigma_0\| + [a_1 - a_0]_\alpha + \|a_1 - a_0\|) + \delta N_0\big).$$

For the use in the proof of Theorem 2 it is very convenient to bring Lemma 3 into the form of Corollary 1. The difference between part i) and part iii) of the corollary on the one hand and part i) and part ii), respectively, of the lemma on the other hand is that the corollary allows for a distribution f that depends on an additional parameter a_0' and establishes estimates on the a_0'-derivatives. Part ii) of the corollary extends part i) of the lemma to two distributions f_1 and f_0.

Corollary 1. *i) Let $\{v(\cdot, a_0)\}_{a_0}$ be a family of functions and let*

$$\{f(\cdot, a_0')\}_{a_0'}, \quad \{v(\cdot, a_0) \diamond f(\cdot, a_0')\}_{a_0, a_0'}$$

be two families of distributions satisfying (37) and

$$(48) \qquad \sup_{T \leq 1}(T^{\frac{1}{4}})^{2-\alpha} \sup_{a_0'} \|\{1, \frac{\partial}{\partial a_0'}, \frac{\partial^2}{\partial a_0'^2}\}f_T\| \leq N_1,$$

$$(49) \qquad \sup_{T \leq 1}(T^{\frac{1}{4}})^{2-2\alpha} \sup_{a_0, a_0'} \|\{1, \frac{\partial}{\partial a_0}\}\{1, \frac{\partial}{\partial a_0'}, \frac{\partial^2}{\partial a_0'^2}\}[v, (\cdot)_T] \diamond f\| \leq N_1 N_0$$

for some constant N_1. If u is modelled after v according to a and σ, satisfying (40), with constant M we have

$$(50) \quad \sup_{T \leq 1}(T^{\frac{1}{4}})^{2-2\alpha} \sup_{a_0'} \|\{1, \frac{\partial}{\partial a_0'}, \frac{\partial^2}{\partial a_0'^2}\}[u, (\cdot)_T] \diamond f\| \lesssim N_1(M + N_0).$$

ii) Let $\{v(\cdot, a_0)\}_{a_0}$, $\{f_j(\cdot, a_0')\}_{a_0'}$, and $\{v(\cdot, a_0) \diamond f_j(\cdot, a_0')\}_{a_0, a_0'}$, $j = 0, 1$, be as in i) and suppose in addition

$$(51) \quad \sup_{T \leq 1}(T^{\frac{1}{4}})^{2-\alpha} \sup_{a_0'} \|\{1, \frac{\partial}{\partial a_0'}\}(f_{1T} - f_{0T})\| \leq \delta N_1,$$

$$\sup_{T \leq 1}(T^{\frac{1}{4}})^{2-2\alpha} \sup_{a_0, a_0'} \|\{1, \frac{\partial}{\partial a_0}\}\{1, \frac{\partial}{\partial a_0'}\}([v, (\cdot)_T] \diamond f_1$$

$$(52) \quad - [v, (\cdot)_T] \diamond f_0)\| \leq \delta N_1 N_0$$

for some constant δN_1. Then for u as in i) we have

$$\sup_{T \leq 1}(T^{\frac{1}{4}})^{2-2\alpha} \sup_{a_0'} \|\{1, \frac{\partial}{\partial a_0'}\}([u, (\cdot)_T] \diamond f_1$$

$$(53) \quad - [u, (\cdot)_T] \diamond f_0)\| \lesssim \delta N_1(M + N_0).$$

iii) Let the two families of functions $\{v_i(\cdot, a_0)\}_{a_0}$, $i = 0, 1$, and the three families of distributions $\{f(\cdot, a_0')\}_{a_0'}$, $\{v_i(\cdot, a_0) \diamond f(\cdot, a_0')\}_{a_0, a_0'}$ be as in i) and satisfy in addition (42), (43). Suppose we have in addition

$$(54) \quad \sup_{T \leq 1}(T^{\frac{1}{4}})^{2-2\alpha} \sup_{a_0, a_0'} \|\{1, \frac{\partial}{\partial a_0}, \frac{\partial^2}{\partial a_0^2}\}\{1, \frac{\partial}{\partial a_0'}\}[v_i, (\cdot)_T] \diamond f\| \leq N_1 N_0,$$

$$\sup_{T \leq 1}(T^{\frac{1}{4}})^{2-2\alpha} \sup_{a_0, a_0'} \|\{1, \frac{\partial}{\partial a_0}\}\{1, \frac{\partial}{\partial a_0'}\}([v_1, (\cdot)_T] \diamond f$$

$$(55) \quad - [v_0, (\cdot)_T] \diamond f)\| \leq N_1 \delta N_0.$$

Let u_i be two functions like in part i) and suppose that $u_1 - u_0$ is modelled after (v_1, v_0) according to (a_1, a_0) and $(\sigma_1, -\sigma_0)$ with constant δM. Then we have

$$\sup_{T \leq 1}(T^{\frac{1}{4}})^{2-2\alpha} \sup_{a_0'} \|\{1, \frac{\partial}{\partial a_0'}\}([u_1, (\cdot)_T] \diamond f - [u_0, (\cdot)_T] \diamond f)\| \lesssim N_1 \big(\delta M$$

$$(56) \quad + N_0([\sigma_1 - \sigma_0]_\alpha + \|\sigma_1 - \sigma_0\| + [a_1 - a_0]_\alpha + \|a_1 - a_0\|) + \delta N_0\big).$$

We now turn to Lemma 4 that deals with the second factor in $a \diamond \partial_1^2 u$. The reason why we consider several functions $v_1, \cdots, v_I$ in Lemma 4 instead of a single one for our scalar PDE is that this seems necessary when establishing the contraction property for Proposition 1; because of the a_0-dependence, it turns out that we need not just $I = 2$ but in fact $I = 3$, cf Corollary 2.

Lemma 4. *Let $\frac{2}{3} < \alpha < 1$. We are given a function b, I families of functions $\{v_1(\cdot, a_0), \cdots, v_I(\cdot, a_0)\}_{a_0}$, and I families of distributions $\{b \diamond \partial_1^2 v_1(\cdot, a_0), \cdots, b \diamond \partial_1^2 v_I(\cdot, a_0)\}_{a_0}$ with*

$$
(57) \qquad \sup_{a_0}[\{1, \frac{\partial}{\partial a_0}\}v_i]_\alpha \leq N_i,
$$

$$
(58) \qquad \sup_{T \leq 1}(T^{\frac{1}{4}})^{2-2\alpha} \sup_{a_0} \|\{1, \frac{\partial}{\partial a_0}\}[b, (\cdot)_T] \diamond \partial_1^2 v_i\| \leq N_0 N_i
$$

for some constants $N_0, \cdots, N_I$. Let the function u be modelled after $(v_1, \cdots, v_I)$ according to the α-Hölder functions a and $(\sigma_1, \cdots, \sigma_I)$ with constant M, cf Definition 1. Then there exists a unique distribution $b \diamond \partial_1^2 u$ such that on the level of the commutators

$$
(59) \qquad \lim_{T \downarrow 0} \|[b, (\cdot)_T] \diamond \partial_1^2 u - \sigma_i E[b, (\cdot)_T] \diamond \partial_1^2 v_i\| = 0,
$$

where E denotes the operator that evaluates a function in two variables (x, a_0) at $(x, a(x))$. Moreover, provided $[a]_\alpha \leq 1$, we have the suboptimal estimate

$$
(60) \qquad \sup_{T \leq 1}(T^{\frac{1}{4}})^{2-2\alpha}\|[b, (\cdot)_T] \diamond \partial_1^2 u\| \lesssim [b]_\alpha M + N_0 N_i([\sigma_i]_\alpha + \|\sigma_i\|).
$$

The following lemma is the only place where we use the PDE. It might be seen as an extension of Schauder theory in the sense that it compares, on the level of $C^{2\alpha}$, the solution u of a variable-coefficient equation $\partial_2 u - a \diamond \partial_1^2 u = \sigma \diamond f$ to the solutions of the corresponding constant-coefficient equation (62), by saying that u is modelled after v according to a and σ. To this purpose we apply $(\cdot)_T$ to the equation and rearrange to

$$
\partial_2 u_T - P(a \partial_1^2 u_T - \sigma f_T) = -P\big([a, (\cdot)_T] \diamond \partial_1^2 u + [\sigma, (\cdot)_T] \diamond f\big).
$$

Since the previous lemmas estimate the commutators on the rhs, we will right away assume that the lhs is estimated accordingly, cf (63). Working with the commutator of multiplication with a coefficient a and convolution is reminiscent of the DiPerna-Lions theory, which however deals with a transport instead of a parabolic equations with a rough coefficient, that is $\partial_2 u - a\partial_1 u$ instead of $\partial_2 u - a\partial_1^2 u$. In our proof, we follow the approach to classical Schauder theory of Krylov & Safanov, see [12], in particular Section 8.6. This approach avoids the use of kernels.

Lemma 5. *Let $\frac{1}{2} < \alpha < 1$ and suppose all functions and distributions are periodic. Were are given I families of distributions $\{f_1(\cdot, a_0), \cdots, f_I(\cdots, a_0)\}_{a_0}$ with*

$$
(61) \qquad \sup_{T \leq 1}(T^{\frac{1}{4}})^{2-\alpha} \sup_{a_0} \|\{1, \frac{\partial}{\partial a_0}\}f_{iT}\| \leq N_i
$$

for some constants $N_1, \cdots, N_I$. For $a_0 \in [\lambda, \frac{1}{\lambda}]$ we denote by $v_i(\cdot, a_0)$ the function of vanishing mean solving

$$(62) \qquad (\partial_2 - a_0 \partial_1^2) v_i(\cdot, a_0) = P f_i(\cdot, a_0) \quad \text{distributionally.}$$

We are also given a function u, modelled after $(v_1, \cdots, v_I)$ according to some functions $a \in [\lambda, \frac{1}{\lambda}]$ and $(\sigma_1, \cdots, \sigma_I)$. We assume that u approximately satisfies the PDE $\partial_2 u - P a \partial_1^2 u = P \sigma_i E f_i$ in the sense of

$$(63) \qquad \sup_{T \leq 1} (T^{\frac{1}{4}})^{2-2\alpha} \| \partial_2 u_T - P(a \partial_1^2 u_T + \sigma_i E f_{iT}) \| \leq N^2$$

for some constant N, where E is defined as in Lemma 4. Then we have for the modelling and the Hölder constant of u

$$(64) \qquad M \lesssim N^2 + [a]_\alpha M + N_i([\sigma_i]_\alpha + \|\sigma_i\|[a]_\alpha),$$

$$(65) \qquad [u]_\alpha \lesssim M + N_i \|\sigma_i\|.$$

In Corollary 2, we will combine Lemma 4 on the product $a \diamond \partial_1^2 u$ and Lemma 5 to obtain an *a priori* estimate on the modelling and Hölder constants. The use of the "infinitesimal" part ii) of this corollary will be explained in the discussion of Proposition 1.

Corollary 2. *Let $\frac{2}{3} < \alpha < 1$. i) Suppose we are given two functions σ and a, two distributions f and $\sigma \diamond f$, and a family of distributions $\{a \diamond \partial_1^2 v(\cdot, a_0)\}_{a_0}$ with*

$$(66) \qquad\qquad\qquad [\sigma]_\alpha + [a]_\alpha \leq N,$$

$$(67) \qquad\qquad\qquad \sup_{T \leq 1} (T^{\frac{1}{4}})^{2-\alpha} \|f_T\| \leq N_0,$$

$$(68) \qquad\qquad\qquad \sup_{T \leq 1} (T^{\frac{1}{4}})^{2-2\alpha} \|[\sigma, (\cdot)_T] \diamond f\| \leq N N_0,$$

$$(69) \quad \sup_{T \leq 1} (T^{\frac{1}{4}})^{2-2\alpha} \sup_{a_0} \|\{1, \frac{\partial}{\partial a_0}, \frac{\partial^2}{\partial a_0^2}\}[a, (\cdot)_T] \diamond \partial_1^2 v\| \leq N N_0,$$

where $v(\cdot, a_0)$ denotes the mean-free solution of

$$(70) \qquad\qquad\qquad (\partial_2 - a_0 \partial_1^2) v(\cdot, a_0) = P f,$$

and satisfying the constraints

$$(71) \qquad \sigma \in [-1, 1], \ a \in [\lambda, \frac{1}{\lambda}], \ [\sigma]_\alpha \leq 1, \ [a]_\alpha \ll 1.$$

Then if a function u is modelled after v according to a and σ with

$$(72) \qquad\qquad\qquad \partial_2 u - P(a \diamond \partial_1^2 u + \sigma \diamond f) = 0$$

we have for the modelling and Hölder constants

$$(73) \qquad\qquad\qquad M \lesssim N_0 N,$$

$$(74) \qquad\qquad\qquad [u]_\alpha \lesssim N_0(N + 1).$$

ii) In addition, suppose we are given two functions $\delta\sigma$ and δa, three distributions δf, $\sigma \diamond \delta f$, and $\delta\sigma \diamond f$, and two families of distributions $\{a \diamond \partial_1^2 \delta v(\cdot, a_0)\}_{a_0}$ and $\{\delta a \diamond \partial_1^2 v(\cdot, a_0)\}_{a_0}$ with

$$[\delta\sigma]_\alpha + \|\delta\sigma\| + [\delta a]_\alpha + \|\delta a\| \leq \delta N, \tag{75}$$

$$\sup_{T \leq 1}(T^{\frac{1}{4}})^{2-\alpha}\|\delta f_T\| \leq \delta N_0, \tag{76}$$

$$\sup_{T \leq 1}(T^{\frac{1}{4}})^{2-2\alpha}\|[\sigma, (\cdot)_T] \diamond \delta f\| \leq N \delta N_0, \tag{77}$$

$$\sup_{T \leq 1}(T^{\frac{1}{4}})^{2-2\alpha}\|[\delta\sigma, (\cdot)_T] \diamond f\| \leq \delta N N_0, \tag{78}$$

$$\sup_{T \leq 1}(T^{\frac{1}{4}})^{2-2\alpha} \sup_{a_0} \|\{1, \frac{\partial}{\partial a_0}\}[a, (\cdot)_T] \diamond \partial_1^2 \delta v\| \leq N \delta N_0, \tag{79}$$

$$\sup_{T \leq 1}(T^{\frac{1}{4}})^{2-2\alpha} \sup_{a_0} \|\{1, \frac{\partial}{\partial a_0}\}[\delta a, (\cdot)_T] \diamond \partial_1^2 v\| \leq \delta N N_0 \tag{80}$$

for some constants $\delta N_0, \delta N$ and where δv is the mean-free solution of

$$(\partial_2 - a_0\partial_1^2)\delta v = P\delta f. \tag{81}$$

Then if a function δu is modelled after $(v, \frac{\partial v}{\partial a_0}, \delta v)$ according to a and $(\delta\sigma, \sigma\delta a, \sigma)$ with

$$\partial_2\delta u - P(a \diamond \partial_1^2 \delta u + \delta a \diamond \partial_1^2 u + \sigma \diamond \delta f + \delta\sigma \diamond f) = 0 \tag{82}$$

then we have for the modelling and Hölder constants

$$\delta M \lesssim N_0 \delta N + \delta N_0 N \quad \textit{provided } N \leq 1, \tag{83}$$

$$[\delta u]_\alpha \lesssim N_0 \delta N + \delta N_0 \quad \textit{provided } N \leq 1. \tag{84}$$

The following Proposition 1 may be seen as the main contribution of this paper. It establishes a solution theory for the linear equation $\partial_2 u - P(a \diamond \partial_1^2 u + \sigma \diamond f) = 0$ for given driver f (a distribution) and functions σ and a. Because of the roughness of f, it does not only require an definition of $\sigma \diamond f$ but also of $a \diamond \partial_1^2 v$, where $v(\cdot, a_0)$ solves $\partial_2 v - a_0\partial_1^2 v = Pf$, so that when u is modelled after v according to a and σ, also $a \diamond \partial_1^2 u$ may be given a sense by Lemma 4. The most subtle point is to establish Lipschitz continuity of u in the data $(a, a \diamond \partial_1^2 v)$. This involves considering differences of solutions and quantifying

$$u_1 - u_0 \text{ is modelled after } (v_1, v_0)$$

$$\text{according to } (a_1, a_0) \text{ and } (\sigma_1, -\sigma_0). \tag{85}$$

When quantifying differences of solutions, variable coefficients require a somewhat different strategy compared to constant coefficients, as we shall explain now. The modelledness (85) has to come from the PDE, that is, Lemma 5. The naive approach is to consider the difference

of the PDE for two given pairs of data (σ_i, a_i, f_i), $i = 0, 1$, (plus the products), and to rearrange as follows

$$\partial_2(u_1 - u_0) - P(a_0 \diamond \partial_1^2 u_1 - a_0 \diamond \partial_1^2 u_0)$$

$$(86) \qquad = P\big(\sigma_1 \diamond f_1 - \sigma_0 \diamond f_0 + (a_1 \diamond \partial_1^2 u_1 - a_0 \diamond \partial_1^2 u_1)\big),$$

which already means breaking the permutation symmetry in $i = 0, 1$ and therefore does not bode well. By the modelledness of u_1 we expect that for the purpose of Lemma 5, we may replace u_1 by v_1 on the rhs of (86), leading to

$$\partial_2(u_1 - u_0) - P(a_0 \diamond \partial_1^2 u_1 - a_0 \diamond \partial_1^2 u_0)$$

$$(87) \qquad \approx P\big(\sigma_1 \diamond f_1 - \sigma_0 \diamond f_0 + \sigma_1(E_1 a_1 \diamond \partial_1^2 v_1 - E_1 a_0 \diamond \partial_1^2 v_1)\big).$$

In view of Lemma 5 and the discussion preceding it, this suggests that we obtain

$$(88) \qquad u_1 - u_0 \text{ is modelled after } (v_1, v_0, (\partial_2 - a_0 \partial_1^2)^{-1} P E_1 \partial_1^2 v)$$
$$\text{according to } a_0 \text{ and } (\sigma_1, -\sigma_0, \sigma_1(a_1 - a_0)),$$

which is *not* the desired (85) unless $a_1 = a_0$. Instead, our strategy will be to construct a *curve* $\{u_s\}_{s \in [0,1]}$ interpolating between u_0 and u_1. For this, we interpolate the data linearly, that is, $f_s := s f_1 + (1 - s) f_0$, $\sigma_s := s \sigma_1 + (1 - s) \sigma_0$, and $a_s := s a_1 + (1 - s) a_0$, and solve

$$(89) \qquad \partial_2 u_s - P(a_s \diamond \partial_1^2 u_s + \sigma_s \diamond f_s) = 0.$$

Provided we interpolate the products bi-linearly, that is,

$$(90) \quad \sigma_s \diamond f_s := s^2 \sigma_1 \diamond f_1 + s(1 - s)(\sigma_1 \diamond f_0 + \sigma_0 \diamond f_1) + (1 - s)^2 \sigma_0 \diamond f_0$$

and the same definition for $a_s \diamond \partial_1^2 v_s$, Leibniz' rule for $\sigma_s \diamond f_s$ holds, and we expect it to hold for $a_s \diamond \partial_1^2 u_s$ so that differentiation of (89) gives

$$\partial_2 \partial_s u - P(a_s \diamond \partial_1^2 \partial_s u) = P(\partial_s a \diamond \partial_1^2 u_s + \partial_s \sigma \diamond f_s + \sigma_s \diamond \partial_s f),$$

which in view of (89) we approximate by

$$\partial_2 \partial_s u - P(a_s \diamond \partial_1^2 \partial_s u) \approx P(\sigma_s E_s \partial_s a \diamond \partial_1^2 v_s + \partial_s \sigma \diamond f_s + \sigma_s \diamond \partial_s f),$$

with $v_s = s v_1 + (1 - s) v_0$. It is this form that motivates the part ii) of Corollary 2. Noting that $(\partial_2 - a_0 \partial_1^2) \frac{\partial v_s}{\partial a_0} = \partial_1^2 v_s$ we obtain

$$(91) \qquad \partial_s u \text{ is modelled after } (v_s, \frac{\partial v_s}{\partial a_0}, \partial_s v)$$
$$\text{according to } a_s \text{ and } (\partial_s \sigma, \sigma_s \partial_s a, \sigma_s),$$

which should be compared with (88). Using Leibniz' rule once more, but this time in the classical form of

$$\frac{\partial}{\partial s}\big(\sigma_s(x) v_s(y, a_s(x))\big) = (\partial_s \sigma)(x) v_s(y, a_s(x))$$

$$+ (\sigma_s \partial_s a)(x) \frac{\partial v_s}{\partial a_0}(y, a_s(x)) + \sigma_s(x) \partial_s v(y, a_s(x)),$$

and integrating (91) in $s \in [0,1]$ yields the desired (85). We note that this strategy differs from [4] even in case when a is constant: When passing from the modelledness of $u_1 - u_0$ to the modelledness of $\sigma(u_1) - \sigma(u_0)$, the argument in [4, Proposition 4] uses the linear interpolation $u_s = su_1 + (1-s)u_0$ (as we do in Lemma 1), which implicitly amounts to the interpolation $\sigma_s \diamond f_s = s\sigma_1 \diamond f_1 + (1-s)\sigma_0 \diamond f_0$, as opposed to (90).

Proposition 1. *Let* $\frac{2}{3} < \alpha < 1$. *i) Suppose we are given two functions* σ *and* a, *two distributions* f *and* $\sigma \diamond f$, *and a family of distributions* $\{a \diamond \partial_1^2 v(\cdot, a_0)\}_{a_0}$ *related by*

$$(92) \qquad\qquad\qquad [\sigma]_\alpha + [a]_\alpha \le N,$$

$$(93) \qquad\qquad\qquad \sup_{T \le 1}(T^{\frac{1}{4}})^{2-\alpha}\|f_T\| \le N_0,$$

$$(94) \qquad\qquad\qquad \sup_{T \le 1}(T^{\frac{1}{4}})^{2-2\alpha}\|[\sigma, (\cdot)_T] \diamond f\| \le N_0 N,$$

$$(95) \qquad \sup_{T \le 1}(T^{\frac{1}{4}})^{2-2\alpha} \sup_{a_0}\|\{1, \frac{\partial}{\partial a_0}, \frac{\partial^2}{\partial a_0^2}\}[a, (\cdot)_T] \diamond \partial_1^2 v\| \le N_0 N,$$

for some constants N_0, N, *where* $v(\cdot, a_0)$ *denotes the mean-free solution of* $(\partial_2 - a_0\partial_1^2)v(\cdot, a_0) = Pf$. *Let the functions* σ *and* a *satisfy*

$$(96) \qquad\qquad \sigma \in [-1,1], \ a \in [\lambda, \frac{1}{\lambda}], \ [\sigma]_\alpha \le 1, \ [a]_\alpha \ll 1.$$

Then there exists a unique mean-free function u *modelled after* v *according to* a *and* σ *and such that*

$$(97) \qquad\qquad \partial_2 u - P(a \diamond \partial_1^2 u + \sigma \diamond f) = 0.$$

The modelling and Hölder constants are estimated as follows

$$(98) \qquad\qquad\qquad M \lesssim N_0 N,$$

$$(99) \qquad\qquad\qquad [u]_\alpha \lesssim N_0(N+1).$$

ii) Suppose we are given four functions σ_i *and* a_i, $i = 0,1$, *six distributions* f_i *and* $\sigma_i \diamond f_j$, $j = 0,1$, *and four families of distributions* $\{a_i \diamond \partial_1^2 v_j(\cdot, a_0)\}_{a_0}$, *where* $v_i(\cdot, a_0)$ *is the mean-free solution of* $(\partial_2 - a_0\partial_1^2)v_i = Pf_i$, *satisfying the assumption (92), (93), (94), and (95), the two latter with cross terms, that is,*

$$(100) \qquad\qquad \sup_{T \le 1}(T^{\frac{1}{4}})^{2-2\alpha}\|[\sigma_i, (\cdot)_T] \diamond f_j\| \le N_0 N,$$

$$(101) \qquad \sup_{T \le 1}(T^{\frac{1}{4}})^{2-2\alpha} \sup_{a_0}\|\{1, \frac{\partial}{\partial a_0}, \frac{\partial^2}{\partial a_0^2}\}[a_i, (\cdot)_T] \diamond \partial_1^2 v_j\| \le N_0 N$$

and (96). We suppose in addition that

$$(102) \qquad [\sigma_1 - \sigma_0]_\alpha + \|\sigma_1 - \sigma_0\| + [a_1 - a_0]_\alpha + \|a_1 - a_0\| \le \delta N,$$

$$(103) \qquad \sup_{T \le 1}(T^{\frac{1}{4}})^{2-\alpha}\|f_{1T} - f_{0T}\| \le \delta N_0,$$

$$(104) \qquad \sup_{T \le 1}(T^{\frac{1}{4}})^{2-2\alpha}\|[\sigma_i, (\cdot)_T] \diamond f_1 - [\sigma_i, (\cdot)_T] \diamond f_0\| \le N\delta N_0,$$

$$(105) \qquad \sup_{T \le 1}(T^{\frac{1}{4}})^{2-2\alpha}\|[\sigma_1, (\cdot)_T] \diamond f_j - [\sigma_1, (\cdot)_T] \diamond f_j\| \le \delta N N_0,$$

$$\sup_{T \le 1}(T^{\frac{1}{4}})^{2-2\alpha} \sup_{a_0} \|\{1, \frac{\partial}{\partial a_0}\}\big([a_i, (\cdot)_T] \diamond \partial_1^2 v_1$$
$$(106) \qquad\qquad\qquad\qquad -[a_i, (\cdot)_T] \diamond \partial_1^2 v_0\big)\| \le N\delta N_0,$$

$$\sup_{T \le 1}(T^{\frac{1}{4}})^{2-2\alpha} \sup_{a_0} \|\{1, \frac{\partial}{\partial a_0}\}\big([a_1, (\cdot)_T] \diamond \partial_1^2 v_j$$
$$(107) \qquad\qquad\qquad\qquad -[a_0, (\cdot)_T] \diamond \partial_1^2 v_j\big)\| \le \delta N N_0$$

for some constants δN_0, δN. Let u_i denote the corresponding solutions ensured by part i). Then $u_1 - u_0$ is modelled after (v_1, v_0) according to (a_1, a_0) and $(\sigma_1, -\sigma_0)$ with modelling constant and Hölder norm estimated as follows

$$(108) \qquad \delta M \lesssim N_0 \delta N + \delta N_0 N,$$

$$(109) \qquad [u_1 - u_0]_\alpha + \|u_1 - u_0\| \lesssim N_0 \delta N + \delta N_0$$
$$\text{both provided } N \le 1.$$

We now proceed to Theorem 2, the main deterministic result of this paper. It can be seen as a PDE version of the ODE result in [4, Section 5]. Part i) of the theorem provides existence and uniqueness by a contraction mapping argument, corresponding to [4, Proposition 7]; part ii) provides continuity of the fixed point in the model, the analogue of the Lyons' sense of continuity for the Itô map and corresponding to [4, Proposition 8].

Theorem 2. *i) Suppose we are given a distribution f satisfying*

$$(110) \qquad \sup_{T \le 1}(T^{\frac{1}{4}})^{2-\alpha}\|f_T\| \le N_0$$

for some constant $N_0 \ll 1$; denote by $v(\cdot, a_0)$ the mean-free solution of $(\partial_2 - a_0 \partial_1^2)v = Pf$. Suppose further that we are given a one-parameter family of distributions $v(\cdot, a_0') \diamond f$ and a two-parameter family of distributions $v(\cdot, a_0') \diamond \partial_1^2 v(\cdot, a_0)$ satisfying

$$\sup_{T \le 1}(T^{\frac{1}{4}})^{2-2\alpha}\|\{1, \frac{\partial}{\partial a_0'}, \frac{\partial^2}{\partial a_0'^2}\}\{1, \frac{\partial}{\partial a_0}, \frac{\partial^2}{\partial a_0^2}\}$$
$$(111) \qquad\qquad [v, (\cdot)_T] \diamond \{f, \partial_1^2 v\}\| \le N_0^2$$

(in fact, we do not need the highest cross-derivative $\frac{\partial^2}{\partial a_0'^2}\frac{\partial^2}{\partial a_0^2}$). Then there exists a unique mean-free function u with the properties

(112) u is modelled after v according to $a(u)$ and $\sigma(u)$,

(113) $\partial_2 u - P(a(u)\diamond\partial_1^2 u + \sigma(u)\diamond f) = 0$ distributionally,

under the smallness condition

(114) $[u]_\alpha \ll 1.$

This unique u satisfies the estimate

(115) $[u]_\alpha + \|u\| \lesssim N_0$ and $M \lesssim N_0^2,$

where M denotes the modelling constant in (112).

ii) Now suppose we have two distributions f_j, $j = 0, 1$, with

(116) $\sup_{T\leq 1}(T^{\frac{1}{4}})^{2-\alpha}\|f_{jT}\| \leq N_0;$

denote by $v_j(\cdot, a_0)$ the mean-free solution of $(\partial_2 - a_0\partial_1^2)v_j = Pf_j$. Suppose further that for $i = 0, 1$ we are given four one-parameter families of distributions $v_i(\cdot, a_0')\diamond f_j$ and four two-parameter families of distributions $v_i(\cdot, a_0')\diamond\partial_1^2 v_j(\cdot, a_0)$ satisfying the analogue of (111) including cross-terms

$$\sup_{T\leq 1}(T^{\frac{1}{4}})^{2-2\alpha}\|\{1, \frac{\partial}{\partial a_0'}, \frac{\partial^2}{\partial a_0'^2}\}\{1, \frac{\partial}{\partial a_0}, \frac{\partial^2}{\partial a_0^2}\}$$

(117) $[v_i, (\cdot)_T]\diamond\{f_j, \partial_1^2 v_j\}\| \leq N_0^2.$

We measure the distance of f_1 to f_0 in terms of a constant δN_0 with

(118) $\sup_{T\leq 1}(T^{\frac{1}{4}})^{2-\alpha}\|(f_1 - f_0)_T\| \leq \delta N_0,$

$$\sup_{T\leq 1}(T^{\frac{1}{4}})^{2-2\alpha}\|\{1, \frac{\partial}{\partial a_0'}\}\{1, \frac{\partial}{\partial a_0}\}\big([v_i, (\cdot)_T]\diamond\{f_1, \partial_1^2 v_1\}$$

(119) $-[v_i, (\cdot)_T]\diamond\{f_0, \partial_1^2 v_0\}\big)\| \leq N_0\delta N_0,$

$$\sup_{T\leq 1}(T^{\frac{1}{4}})^{2-2\alpha}\|\{1, \frac{\partial}{\partial a_0'}\}\{1, \frac{\partial}{\partial a_0}\}\big([v_1, (\cdot)_T]\diamond\{f_j, \partial_1^2 v_j\}$$

(120) $-[v_0, (\cdot)_T]\diamond\{f_j, \partial_1^2 v_j\}\big)\| \leq N_0\delta N_0.$

If u_i, $i = 0, 1$, denote the corresponding solutions of (112)&(114)&(113) we have

(121) $[u_1 - u_0]_\alpha + \|u_1 - u_0\| \lesssim \delta N_0.$

Moreover, $u_1 - u_0$ is modelled after (v_1, v_0) according to $(a(u_1), a(u_0))$ and $(\sigma(u_1), -\sigma(u_0))$ with modelling constant δM estimated by

(122) $\delta M \lesssim N_0\delta N_0.$

It remains to establish a link between the solution theory presented in Theorem 2 and the classical solution theory in the case where f is smooth, e.g. $f \in C^\beta$ for any $0 < \beta < 1$. In this case by classical Schauder theory $\sup_{a_0}[\{\partial_1^2, \partial_2\}v(\cdot, a_0)]_\beta \lesssim [f]_\beta$ and in particular there is the classical choice for the products

$$(123) \qquad v(\cdot, a_0') \diamond \{f, \partial_1^2 v(\cdot, a_0)\} := v(\cdot, a_0')\{f, \partial_1^2 v(\cdot, a_0)\}.$$

In the language of Hairer [8, Sec. 8.2], this corresponds to the *canonical model* built from a smooth noise term. The only assumption on the products $v(\cdot, a_0') \diamond \{f, \partial_1^2 v(\cdot, a_0)\}$ entering the definition of the singular products are the regularity bounds (111) expressed in terms of commutators and they are easily seen to be satisfied in this case. For example we have

$$\|[v, (\cdot)_T]f\| = \sup_x \left| \int \psi_T(x-y)(v(x) - v(y))f(y)dy \right|$$

$$(124) \qquad\qquad \lesssim T^{\frac{1}{4}}\|\{\partial_1, \partial_2\}v\|\|f\|,$$

which is much more than needed. However, the canonical definition (123) is by no means the only possible choice of product. In fact, as (111) is the only requirement on $v(\cdot, a_0') \diamond \{f, \partial_1^2 v(\cdot, a_0)\}$ we can set for example

$$v(\cdot, a_0') \diamond f := v(\cdot, a_0')f + g^{(1)},$$

$$(125) \qquad v(\cdot, a_0') \diamond \partial_1^2 v(\cdot, a_0) := v(\cdot, a_0')\partial_1^2 v(\cdot, a_0) + g^{(2)}$$

for a one-parameter family of distributions $g^{(1)}$ indexed by a_0 and a two-parameter family $g^{(2)}$ indexed by a_0, a_0'. For this choice of "products" $\diamond$ the commutators turn into

$$[v, (\cdot)_T] \diamond \{f, \partial_1^2 v\} = [v, (\cdot)_T]\{f, \partial_1^2 v\} - (\{g^1, g^{(2)}\})_T$$

so that (111) reduces to the regularity assumption

$$\|\{1, \frac{\partial}{\partial a_0'}, \frac{\partial^2}{\partial a_0'^2}\}(g^{(1)})_T\|, \|\{1, \frac{\partial}{\partial a_0'}, \frac{\partial^2}{\partial a_0'^2}\}\{1, \frac{\partial}{\partial a_0}, \frac{\partial^2}{\partial a_0^2}\}(g^{(2)})_T\|$$

$$(126) \qquad\qquad\qquad\qquad\qquad \lesssim (T^{\frac{1}{4}})^{2\alpha - 2}.$$

This mild assumption leaves a lot of freedom to choose $g^{(i)}$ (any distribution of order $2\alpha - 2$ that is smooth in the parameter would do) but we are mostly interested in the case where they are constant in x depending only on a_0 and a_0'. The following corollary provides a link between solutions of (113) and classical solutions in the case where the the products $\diamond$ are defined by (125).

Corollary 3. *Let f be a function in C^β for some $0 < \beta < 1$ and let the products $v(\cdot, a_0') \diamond \{f, \partial_1^2 v(\cdot, a_0)\}$ be defined by (125) for $g^{(1)}, g^{(2)}$ which are of class C^β in x and smooth in a_0, a_0'. Then the following are equivalent:*

i) u *is modelled after* v *according to* $a(u)$ *and* $\sigma(u)$ *and solves*
$\partial_2 u - P(a(u) \diamond \partial_1^2 u + \sigma(u) \diamond f) = 0$ *distributionally.*
ii) u *is of class* $C^{\beta+2}$ *and it is a classical solution of*

$$\partial_2 u - P\big(a(u)\partial_1^2 u + a'(u)\sigma(u)^2 g^{(2)}(\cdot, a(u), a(u))$$
$$+\sigma(u)f + \sigma'(u)\sigma(u)g^{(1)}(\cdot, a(u))\big) = 0.$$

3. Stochastic bounds

We now present the stochastic bounds which are necessary as input into our deterministic theory. We consider a random distribution f, construct (renormalized) commutators, and show that the bounds (19) and (111) hold for these objects. The calculations in this section are inspired by a similar reasoning (in a more complicated situation) in [11, Sec. 5], [8, Sec. 10]; for the reader's convenience we provide self-contained proofs.

Let f be a Gaussian centered distribution which is 1-periodic in both the x_1 and the x_2 direction. Such a distribution is most conveniently represented in terms of its Fourier series development given by

$$(127) \qquad f(x) = \sum_{k \in (2\pi\mathbb{Z})^2} \sqrt{\hat{C}(k)} e^{ik \cdot x} Z_k,$$

which converges in a suitable topology on distributions. Here the Z_k are complex-valued centered Gaussians which are independent except for the symmetry constraint $Z_k = \bar{Z}_{-k}$ and satisfy $\langle Z_k Z_{-\ell} \rangle = \delta_{k,\ell}$. The coefficients $\sqrt{\hat{C}}$ are assumed to be real-valued, non-negative, and symmetric $\sqrt{\hat{C}(k)} = \sqrt{\hat{C}(-k)}$. This notation is chosen because in the case where realisations from f are (say smooth) functions the coefficients in (127) do coincide with the square root of the Fourier transform of the covariance function as we now demonstrate. If, using the stationarity of f, we define the covariance function

$$\langle f(x)f(x')\rangle = C(x - x')$$

for $x = (x_1, x_2)$, $x' = (x'_1, x'_2)$, then stationarity also implies that

$$C(x) = \langle f(x)f(0)\rangle = \langle f(0)f(-x)\rangle = C(-x).$$

Hence the (discrete) Fourier transform

$$\hat{C}(k) = \int_{[0,1)^2} e^{-ik \cdot x} C(x) dx = \int_{[0,1)^2} \cos(k \cdot x) C(x) dx, \qquad k \in (2\pi\mathbb{Z})^2$$

is real valued and symmetric. For $k, \ell \in (2\pi\mathbb{Z})^2$ we have

$$\langle \hat{f}(k)\hat{f}(-\ell) \rangle = \int_{[0,1)^2} \int_{[0,1)^2} e^{-ik\cdot x} e^{i\ell\cdot x'} \langle f(x)f(x') \rangle \, dx dx'$$

$$= \int_{[0,1)^2} \int_{[0,1)^2} e^{-i(k-\ell)\cdot x} e^{-i\ell\cdot(x-x')} C(x-x') \, dx dx'$$

$$(128) \qquad = \delta_{k,\ell} \hat{C}(k),$$

which implies in particular that $\hat{C}$ is non-negative, and exactly corresponds to (127).

The construction of non-linear functionals of f involves regularisation. For this, let ψ' be an arbitrary Schwartz function with $\int_{\mathbb{R}^2} \psi' = 1$. As in the deterministic part we define the rescaling $\psi'_\varepsilon(x_1, x_2) = \frac{1}{\varepsilon^{\frac{3}{4}}} \psi'(\frac{x_1}{\varepsilon^{\frac{1}{4}}}, \frac{x_2}{\varepsilon^{\frac{1}{2}}})$ and define $f_\varepsilon = f * \psi'_\varepsilon$. Of course, $\psi' = \psi_1$ for ψ_1 as in the deterministic analysis constitutes an admissible choice, but in the following analysis of stochastic moments the semi-group property for ψ' is not needed and we therefore do not need to restrict ourselves to this particular choice.

Throughout this section we assume that $\hat{C}(0) = 0$, i.e. f has vanishing average. Our quantitative assumptions on the regularity of f are expressed in terms of $\hat{C}$: We assume that there exist $\lambda_1, \lambda_2 \in \mathbb{R}$ and $\alpha \in (0,1)$ such that

$$(129) \quad \hat{C}(k) \leq \frac{1}{(1+|k_1|)^{\lambda_1}(\sqrt{1+|k_2|})^{\lambda_2}}, \qquad k = (k_1, k_2) \in (2\pi\mathbb{Z})^2,$$

$$\lambda_1 + \lambda_2 = -1 + 2\alpha \qquad \lambda_1, \frac{\lambda_2}{2} < 1.$$

The second condition, may seem confusing, because larger values of λ, corresponding to more smoothness for f, should help our theory. The point is here, that decay in one of the directions beyond summability cannot compensate for a lack of decay in the other direction. In order to use the bounds presented in Lemma 6 and Proposition 2 as input for the deterministic theory in Section 2 we need $\alpha > \frac{2}{3}$, but this condition does not play a role in the proof of these bounds. The following lemma shows that assumption (129) corresponds to the regularity assumption (19) on f.

Lemma 6. *Let f be a stationary centered Gaussian distribution given by (127) and for $\varepsilon > 0$ set $f_\varepsilon = f * \psi'_\varepsilon$. If the assumption (129) holds then we have for any $p < \infty$ and $\alpha' < \alpha$*

$$(130) \qquad \left\langle \left(\sup_{T \leq 1} (T^{\frac{1}{4}})^{2-\alpha'} \|f_T\| \right)^p \right\rangle^{\frac{1}{p}} \lesssim 1.$$

If additionally $0 \leq \kappa \leq 4$, then

(131) $$\left\langle \left(\sup_{0 < \varepsilon \leq 1} \sup_{T \leq 1} (T^{\frac{1}{4}})^{2-\alpha'+\kappa} (\varepsilon^{\frac{1}{4}})^{-\kappa} \|(f_\varepsilon)_T - f_T\| \right)^p \right\rangle^{\frac{1}{p}} \lesssim 1.$$

Here and in the proof the implicit constant in $\lesssim$ depends only on p and α'.

Because of $(f_\varepsilon)_T = f_T * \psi'_\varepsilon$ and because the operators $\psi'_\varepsilon *$ are bounded with respect to $\|\cdot\|$ uniformly in ε the bound (130) immediately implies a bound which holds uniformly in the regularisation ε

$$\left\langle \left(\sup_{\varepsilon \leq 1} \sup_{T \leq 1} (T^{\frac{1}{4}})^{2-\alpha'} \|(f_\varepsilon)_T\| \right)^p \right\rangle^{\frac{1}{p}} \lesssim 1.$$

For $a_0 \in [\lambda, 1]$ let $G(\cdot, a_0)$ be the (periodic) Green function of $(\partial_2 - a_0 \partial_1^2)$, where the heat operator is endowed with periodic and zero average time-space boundary conditions. Its (discrete) Fourier transform is given by

(132) $$\hat{G}(k, a_0) = \begin{cases} \frac{1}{a_0 k_1^2 - i k_2} = \frac{a_0 k_1^2 + i k_2}{a_0^2 k_1^4 + k_2^2} & \text{for } k \in (2\pi\mathbb{Z})^2 \setminus \{0\} \\ 0 & \text{for } k = 0. \end{cases}$$

With these notations in place, the periodic zero-mean solutions of $(\partial_2 - a_0 \partial_1^2) v(\cdot, a_0) = f$ and $(\partial_2 - a_0 \partial_1^2) v_\varepsilon(\cdot, a_0) = f_\varepsilon$ are characterized by their discrete Fourier transforms $\hat{v}(k, a_0) = \hat{G}(k, a_0) \hat{f}_\varepsilon(k)$, $\hat{v}_\varepsilon(k, a_0) = \hat{G}(k, a_0) \hat{f}(k)$ for $k \in (2\pi\mathbb{Z})^2$.

We aim at giving a meaning to the products $v(\cdot, a_0) \diamond f$, $v(\cdot, a_0) \diamond \partial_1^2 v(\cdot, a'_0)$ and obtaining bounds for the families of commutators $[v(\cdot, a_0), (\cdot)_T] \diamond f$, $[v(\cdot, a_0), (\cdot)_T] \diamond \partial_1^2 v(\cdot, a'_0)$ derived from them. The quantity $\partial_1^2 v(\cdot, a_0)$ is obtained from f through a regularity-preserving transformation, as can be expressed in terms of the Fourier transform

$$\widehat{\partial_1^2 v}(k, a_0) = \frac{k_1^2}{a_0 k_1^2 - i k_2} \hat{f}(k)$$

and noting that $\frac{k_1^2}{a_0 k_1^2 - i k_2}$ is a bounded symbol (see also Lemma 9). Therefore, the proofs for $v(\cdot, a_0) \diamond f$ and $v(\cdot, a_0) \diamond \partial_1^2 v(\cdot, a'_0)$ are essentially identical. The list of commutators needed for the deterministic analysis also includes various derivatives with respect to a_0 and a'_0, but these derivatives do not change the regularity either. For example we have for any $n \geq 1$

(133) $$\frac{\partial^n}{\partial a_0^n} \hat{v}(k, a_0) = \frac{\partial^n}{\partial a_0^n} \hat{G}(k, a_0) \hat{f}(k) = \frac{(-1)^n n! k_1^{2n}}{(a_0 k_1^2 - i k_2)^n} \hat{v}(k, a_0),$$

and for every n the symbol $\frac{(-1)^n n! k_1^{2n}}{(a_0 k_1^2 - i k_2)^n}$ is also bounded.

As the regularities of $v(\cdot, a_0)$, f, $\partial_1^2 v(\cdot, a_0)$ are not sufficient to give a deterministic functional analytic interpretation to these products,

we proceed by approximation and study the convergence of $v_\varepsilon(\cdot, a_0) f_\varepsilon$, $v_\varepsilon(\cdot, a_0)\partial_1^2 v_\varepsilon(\cdot, a_0')$ as ε goes to zero by bounding stochastic moments. As a first step, in the following lemma we calculate the expectations of $v_\varepsilon(\cdot, a_0) f_\varepsilon$ and $v_\varepsilon(\cdot, a_0)\partial_1^2 v_\varepsilon(\cdot, a_0')$, which by stationarity do not depend on the point $x \in [0, 1)^2$ they are evaluated at.

Lemma 7. *For $\varepsilon > 0$ we have*

$$(134) \quad \mathfrak{c}^{(1)}(\varepsilon, a_0) := \langle v_\varepsilon(\cdot, a_0) f_\varepsilon \rangle = \sum_{k \in (2\pi\mathbb{Z})^2 \setminus 0} \frac{a_0 k_1^2}{a_0^2 k_1^4 + k_2^2} \hat{C}(k) |(\psi_\varepsilon')^2(k)|,$$

$$\mathfrak{c}^{(2)}(\varepsilon, a_0, a_0') := \langle v_\varepsilon(\cdot, a_0)\partial_1^2 v_\varepsilon(\cdot, a_0') \rangle$$

$$(135) \qquad = \sum_{k \in (2\pi\mathbb{Z})^2 \setminus 0} \frac{(-a_0 a_0' k_1^4 + k_2^2) k_1^2}{(a_0^2 k_1^4 + k_2^2)((a_0')^2 k_1^4 + k_2^2)} \hat{C}(k) |(\psi_\varepsilon')^2(k)|.$$

The regularity assumption (129) does not imply that the constants $\mathfrak{c}^{(1)}(\varepsilon, a_0)$ and $\mathfrak{c}^{(2)}(\varepsilon, a_0, a_0')$ converge to a finite limit as ε tends to zero, although there are interesting cases in which they do converge. This is discussed below, but for the moment we study the convergence of the renormalized products

$$v_\varepsilon(\cdot, a_0) \diamond f_\varepsilon := v_\varepsilon(\cdot, a_0) f_\varepsilon - \mathfrak{c}^{(1)}(\varepsilon, a_0),$$

$$(136) \qquad v_\varepsilon(\cdot, a_0) \diamond \partial_1^2 v_\varepsilon(\cdot, a_0') := v_\varepsilon(\cdot, a_0)\partial_1^2 v_\varepsilon(\cdot, a_0') - \mathfrak{c}^{(2)}(\varepsilon, a_0, a_0'),$$

as well as the corresponding commutators,

$$[v_\varepsilon(\cdot, a_0), (\cdot)_T] \diamond f_\varepsilon = v_\varepsilon(\cdot, a_0)(f_\varepsilon)_T - (v_\varepsilon(\cdot, a_0) \diamond f_\varepsilon)_T,$$

$$[v_\varepsilon(\cdot, a_0), (\cdot)_T] \diamond \partial_1^2 v_\varepsilon(\cdot, a_0')$$

$$(137) \qquad = v_\varepsilon(\cdot, a_0)(\partial_1^2 v_\varepsilon(\cdot, a_0'))_T - (v_\varepsilon(\cdot, a_0) \diamond \partial_1^2 v_\varepsilon(\cdot, a_0'))_T.$$

Observe that while the singular products appearing in this expression, $v_\varepsilon(\cdot, a_0) f_\varepsilon$ and $v_\varepsilon(\cdot, a_0)\partial_1^2 v_\varepsilon(\cdot, a_0')$, are renormalized by subtracting the expectation, the products $v_\varepsilon(\cdot, a_0)(f_\varepsilon)_T$ and $v_\varepsilon(\cdot, a_0)(\partial_1^2 v_\varepsilon(\cdot, a_0'))_T$ are not changed. In particular, unlike the renormalized products in (136) the renormalized commutators in (137) do not have vanishing expectation.

The key result of this section is the following proposition which shows the convergence of the renormalized products and provides a control for stochastic moments of the renormalized commutators in (137) as well as their derivatives with respect to a_0, a_0'.

Proposition 2. *Let f be a stationary centered Gaussian distribution given by (127) and assume that (129) is satisfied. For $\varepsilon > 0$ set $f_\varepsilon = f * \psi_\varepsilon'$. and $\hat{v}_\varepsilon(\cdot, a_0') = \hat{G}(\cdot, a_0')\hat{f}_\varepsilon$ for $G(\cdot, a_0')$ as in (132).*
i) For any $n, n' \geq 0$ the random distributions $\frac{\partial^n}{\partial a_0^n} \frac{\partial^{n'}}{\partial (a_0')^{n'}} v_\varepsilon(\cdot, a_0) \diamond \{f_\varepsilon, \partial_1^2 v_\varepsilon(\cdot, a_0')\}$ converge as $\varepsilon \to 0$. This convergence takes place almost

surely uniformly over a_0, a_0' and with respect to any $C^{\alpha'-2}$ norm for $\alpha' < \alpha$. We denote the limits by $\frac{\partial^n}{\partial a_0^n} \frac{\partial^{n'}}{\partial (a_0')^{n'}} v(\cdot, a_0) \diamond \{f, \partial_1^2 v(\cdot, a_0')\}$.

ii) We have the estimates

$$\Big\langle \Big(\sup_{a_0,a_0'\in[\lambda,1]} \sup_{\varepsilon\in[0,1]} \sup_{T\le 1}(T^{\frac14})^{2-2\alpha'} \|\frac{\partial^n}{\partial a_0^n} \frac{\partial^{n'}}{\partial (a_0')^{n'}}[v_\varepsilon(\cdot,a_0),(\cdot)_T]$$

(138)
$$\diamond\{f_\varepsilon, \partial_1^2 v_\varepsilon(\cdot,a_0')\}\|\Big)^p\Big\rangle^{\frac1p} \lesssim 1,$$

where $\lesssim$ means up to a constant that may depend on n, n', α' and p, as well as

$$\Big\langle \Big(\sup_{a_0,a_0'\in[\lambda,1]} \sup_{\varepsilon\in(0,1]} \varepsilon^{-\frac{\kappa}{4}} \sup_{T\le 1}(T^{\frac14})^{2-2\alpha'+\kappa}$$

$$\Big\| \frac{\partial^n}{\partial a_0^n} \frac{\partial^{n'}}{\partial (a_0')^{n'}}\Big([v_\varepsilon(\cdot,a_0),(\cdot)_T]\diamond\{f_\varepsilon, \partial_1^2 v_\varepsilon(\cdot,a_0')\}$$

(139)
$$- [v(\cdot,a_0),(\cdot)_T]\diamond\{f, \partial_1^2 v(\cdot,a_0')\}\Big)\Big\|\Big)^p\Big\rangle^{\frac1p} \lesssim 1,$$

where $\lesssim$ means up to a constant depending only on n, n', α', κ and p.

Proposition 2 follows from the following estimate on the second moments of commutators.

Lemma 8. *Let f be defined by (127) for $\hat{C}$ satisfying (129), and let G be the Greens function defined in (132) for some $a_0 \in [\lambda, 1]$. Then for Fourier multipliers $\hat{M}_1, \hat{M}_2$ satisfying $\hat{M}_i(k) = \overline{\hat{M}_i(-k)}$ and*

(140) $\quad |\hat{M}_1(k)| \lesssim |k_1^4 + k_2^2|^{\frac{\kappa_1}{4}}, \quad |\hat{M}_2(k)| \lesssim |k_1^4 + k_2^2|^{\frac{\kappa_2}{4}}, \quad k \in (2\pi\mathbb{Z})^2,$

for $0 \le \kappa_1, \kappa_2 \ll 1$ (where $\ll$ depends on λ_1, λ_2), let f' and v' be defined through $\hat{f}' = \hat{M}_1\hat{f}$ and $\hat{v}' = \hat{M}_2\hat{G}\hat{f}$. We make the qualitative assumption that f' and v' are smooth and set $v'\diamond f' = v'f' - \langle v'f'\rangle$ and $[v',(\cdot)_T]\diamond f' = v'(f')_T - (v'\diamond f')_T$. Then

(141) $\qquad \langle([v',(\cdot)_T]\diamond f')^2\rangle^{\frac12} \lesssim (T^{\frac14})^{2\alpha-2-\kappa_1-\kappa_2}.$

Here the implicit constant depends on λ_1, λ_2 in (129), the implicit constants in (140), κ_1, κ_2 as well as the ellipticity contrast λ (but not on the smoothness assumption on f', v').

In the proof of Proposition 2 this Lemma is used in the form of the following immediate corollary:

Corollary 4. *For $\varepsilon > 0$ and $a_0, a_0' \in [\lambda, 1]$ let $f_\varepsilon = \psi_\varepsilon' * f$ and $\hat{v}_\varepsilon(\cdot, a_0') = \hat{G}(\cdot, a_0')\hat{f}_\varepsilon$ for $G(\cdot, a_0')$ as in (132). Furthermore, let $[v_\varepsilon(\cdot, a_0), (\cdot)_T] \diamond\{f_\varepsilon, \partial_1^2 v_\varepsilon(\cdot, a_0')\}$ be defined as in (137).*

Then for $n, n' \geq 0$ we have

$$\left\langle \left(\left[\frac{\partial}{\partial a_0^n} v_\varepsilon(\cdot, a_0), (\cdot)_T \right] \diamond \{ f_\varepsilon, \frac{\partial}{\partial (a_0')^{n'}} \partial_1^2 v_\varepsilon(\cdot, a_0') \} \right)^2 \right\rangle^{\frac{1}{2}}$$

$$\tag{142} \lesssim (T^{\frac{1}{4}})^{2\alpha - 2}.$$

Furthermore, we have for $0 \leq \kappa \ll 1$ ($\ll$ may depend on λ_1, λ_2)

$$\left\langle \left(\frac{\partial}{\partial \varepsilon} \left(\left[\frac{\partial}{\partial a_0^n} v_\varepsilon(\cdot, a_0), (\cdot)_T \right] \diamond \{ f_\varepsilon, \frac{\partial}{\partial (a_0')^{n'}} \partial_1^2 v_\varepsilon \} \right) \right)^2 \right\rangle^{\frac{1}{2}}$$

$$\tag{143} \lesssim \frac{(T^{\frac{1}{4}})^{2\alpha - 2 - \kappa}}{\varepsilon^{1 - \frac{\kappa}{4}}}.$$

Finally, we come back to the products and commutators without renormalization. According to Lemma 7 the constants $\mathfrak{c}^{(1)}(\varepsilon, a_0)$ converge to a non-trivial limit if and only if

$$\tag{144} \sum_{k \in (2\pi\mathbb{Z})^2 \setminus 0} \frac{k_1^2}{k_1^4 + k_2^2} \hat{C}(k) < \infty.$$

Furthermore, given that the ratio of the kernels appearing in (134) and (135)

$$\lambda \leq \left| \frac{-a_0' k_1^4 + a_0^{-1} k_2^2}{(a_0')^2 k_1^4 + k_2^2} \right| \leq \lambda^{-3}$$

is bounded away from 0 and ∞ the convergence of the $\mathfrak{c}^{(2)}(\varepsilon, a_0, a_0')$ as ε goes to zero is also equivalent to (144). The condition (144) also implies the convergence for arbitrary derivatives of $\mathfrak{c}^{(1)}, \mathfrak{c}^{(2)}$ with respect to a_0, a_0'. For example, recalling (133) and the fact that the term $\frac{a_0 k_1^2}{a_0^2 k_1^4 + k_2^2}$ is nothing but the real part $\Re$ of $\hat{G}(k, a_0)$ we can write

$$\frac{\partial^n}{\partial a_0^n} \mathfrak{c}^{(1)}(\varepsilon, a_0) = \sum_{k \in (2\pi\mathbb{Z})^2 \setminus 0} \Re \left(\frac{\partial^n}{\partial a_0^n} \hat{G}(k, a_0) \right) \hat{C}(k) |(\psi_\varepsilon')^2(k)|$$

$$= \sum_{k \in (2\pi\mathbb{Z})^2 \setminus 0} \Re \left(\frac{(-1)^n n! k_1^{2n}}{(a_0 k_1^2 - ik_2)^n} \hat{G}(k, a_0) \right) \hat{C}(k) |(\psi_\varepsilon')^2(k)|.$$

Given that for any $n \geq 1$ the absolute value of the quantity under the real part $\Re$ is $\lesssim \frac{k_1^2}{k_1^4 + k_2^2}$ the convergence as $\varepsilon \to 0$ under (144) follows. A similar argument works for $\mathfrak{c}^{(2)}$. We summarise this discussion in the following corollary.

Corollary 5. *Assume that both (129) and (144) hold. Then the statements of Proposition 2 remain true if all of the renormalized products are replaced by products without renormalisation.*

The limits which exist under the assumptions of this corollary will be denoted by $[v(\cdot, a_0), (\cdot)_T] f$, $[v(\cdot, a_0), (\cdot)_T] \partial_1^2 v(\cdot, a_0)$ etc..

We finish this section by comparing the assumptions (129) and (144) in particular cases. First consider the case

$$(145) \qquad \hat{C}(k) = (1 + |k_1|)^{-\lambda_1} (1 + |k_2|)^{-\frac{\lambda_2}{2}}.$$

For this choice of $\hat{C}$ the regularity assumption (129) is equivalent to

$$(146) \quad \lambda_1 + \lambda_2 \geq -1 + 2\alpha, \qquad \lambda_1 > -3 + 2\alpha, \quad \text{and} \quad \lambda_2 > -2 + 2\alpha.$$

(note that equality is not necessary in the first condition, because in the case of strict inequality, one can find $\lambda_1' \leq \lambda_1$ and $\lambda_2' \leq \lambda_2$ that satisfy (129) with equality. However, $\lambda_1 \leq -3 + 2\alpha$ or $\lambda_2 \leq -2 + 2\alpha$ can never be compensated without violating the second condition in (129)) The condition (144) on the other hand is equivalent to

$$(147) \qquad \lambda_1 + \lambda_2 > 1 \qquad \lambda_1 > -1, \quad \text{and} \quad \lambda_2 > -2.$$

For any $\alpha \in (0, 1)$ the first requirements in (146) is weaker than the corresponding assumptions in (147). An interesting case in which both assumptions are satisfied and for which our theory can therefore be applied without renormalisation is the case where $\lambda_1 > 1$ and $\lambda_2 = 0$; this corresponds to the case of noise which is white in the time-like variable x_2 but "trace-class" in x_1. However, if we are willing to accept renormalisation, the regularity requirement in the x_1 direction reduces to $\lambda_1 > \frac{1}{3}$ (recall that the deterministic analysis is applicable if $\alpha > \frac{2}{3}$). Another interesting case is the covariance

$$\hat{C}(k_1, k_2) = \delta_{k_2,0} (1 + |k_1|)^{-\lambda_1},$$

which corresponds to a noise term which only depends on the space-like x_1 variable. The parabolic equations with constant diffusion coefficients driven by such a noise term has recently been studied as *Parabolic Anderson model* in two and three spatial dimensions [5, 10, 9, 1]. Our theory applies without renormalisation for all $\lambda_1 > -1$, which covers in particular the case of one-dimensional spatial white noise, $\lambda_1 = 0$. If we admit renormalisation we can go all the way to $\lambda_1 > -\frac{5}{3}$ by choosing $\lambda_2 < 2$ and $\alpha > \frac{2}{3}$ as close to 2 and $\frac{2}{3}$, respectively, as we please. This covers the case $\lambda_1 = -1$ for which the noise f has the same scaling behaviour as spatial white noise in two dimensions (both are distributions of regularity C^{-1-}) but it does not cover the case $\lambda_1 = -2$ for which the noise scales like spatial white noise in three dimensions.

4. Proofs for the deterministic analysis

Proof of Theorem 2.

We write for abbreviation $[\cdot] = [\cdot]_\alpha$. We consider the map defined through

$$(148) \quad (\bar{u}, \bar{a}, \bar{\sigma}) \mapsto (\sigma := \sigma(\bar{u}), a := a(\bar{u}), \sigma \diamond f, a \diamond \partial_1^2 v) \mapsto (u, a, \sigma),$$

where u is the solution provided by Proposition 1, the map of which we seek to characterize the fixed point.

Step 1. Pointwise nonlinear transformation, application of Lemma 1. We work under the assumptions of part ii) of the theorem on the distributions f_j and the off-line products $v_i \diamond f_j$, $v_i \diamond \partial_1^2 v_j$. Suppose we are given two triplets $(\bar{u}_i, \bar{a}_i, \bar{\sigma}_i)$, $i = 0, 1$, of functions satisfying the constraints

$$(149) \quad \bar{\sigma}_i \in [-1, 1], \ \bar{a}_i \in [\lambda, \tfrac{1}{\lambda}], \quad [\bar{\sigma}_i], [\bar{a}_i] \le 1.$$

We measure the size of $\{(\bar{u}_i, \bar{a}_i, \bar{\sigma}_i)\}_i$ and their distance through

$$(150) \quad \bar{M} := \max_i (M_{\bar{u}_i} + [\bar{u}_i]) + N_0,$$

$$\delta\bar{M} := M_{\bar{u}_1 - \bar{u}_0} + [\bar{u}_1 - \bar{u}_0] + \|\bar{u}_1 - \bar{u}_0\|$$

$$(151) \quad + N_0([\bar{\sigma}_1 - \bar{\sigma}_0] + \|\bar{\sigma}_1 - \bar{\sigma}_0\| + [\bar{a}_1 - \bar{a}_0] + \|\bar{a}_1 - \bar{a}_0\|) + \delta N_0,$$

where $M_{\bar{u}_i}$ denotes the constant in the modelledness of $\bar{u}_i$ after v_i according to $\bar{a}_i$ and $\bar{\sigma}_i$, and where $M_{\bar{u}_1 - \bar{u}_0}$ denotes the constant in the modelledness of $\bar{u}_1 - \bar{u}_0$ after (v_1, v_0) according to $(\bar{a}_1, \bar{a}_0)$ and $(\bar{\sigma}_1, -\bar{\sigma}_0)$.

We now consider $\sigma_i := \sigma(\bar{u}_i)$ and $a_i := a(\bar{u}_i)$. We claim

$$(152) \quad \sigma_i \in [-1, 1], \ a_i \in [\lambda, \tfrac{1}{\lambda}], \ [\sigma_i], [a_i] \le 1 \quad \text{provided } \max_i [\bar{u}_i] \ll 1,$$

$$(153) \quad \tilde{M} \lesssim \bar{M} \quad \text{provided } \max_i [\bar{u}_i] \le 1,$$

$$(154) \quad \delta\tilde{M} \lesssim \delta\bar{M} \quad \text{provided } \bar{M} \le 1,$$

where we define in analogy with (150) and (151):

$$(155) \quad \tilde{M} := \max_i (M_{\sigma_i} + [\sigma_i] + M_{a_i} + [a_i]) + N_0,$$

$$\delta\tilde{M} := M_{\sigma_1 - \sigma_0} + [\sigma_1 - \sigma_0] + \|\sigma_1 - \sigma_0\|$$

$$+ N_0([\omega_1 - \omega_0] + \|\omega_1 - \omega_0\| + [\bar{a}_1 - \bar{a}_0] + \|\bar{a}_1 - \bar{a}_0\|)$$

$$+ M_{a_1 - a_0} + [a_1 - a_0] + \|a_1 - a_0\|$$

$$(156) \quad + N_0([\mu_1 - \mu_0] + \|\mu_1 - \mu_0\| + [\bar{a}_1 - \bar{a}_0] + \|\bar{a}_1 - \bar{a}_0\|) + \delta N_0,$$

with the understanding that σ_i is modelled after v_i according to $\bar{a}_i$ and $\omega_i := \sigma'(\bar{u}_i)\bar{\sigma}_i$ and constant M_{σ_i}; that a_i is modelled after v_i according to $\bar{a}_i$ and $\mu_i := a'(\bar{u}_i)\bar{\sigma}_i$ and constant M_{a_i}; that $\sigma_1 - \sigma_0$ is modelled after (v_1, v_0) according to $(\bar{a}_1, \bar{a}_0)$ and $(\omega_1, -\omega_0)$ and a constant we

name $M_{\sigma_1-\sigma_0}$; and that $a_1 - a_0$ is modelled after (v_1, v_0) according to $(\bar{a}_1, \bar{a}_0)$ and $(\mu_1, -\mu_0)$ and a constant we name $M_{a_1-a_0}$.

It is obvious from (20) that (149) turns into (152) under the assumption $\max_i[\bar{u}_i] \ll 1$. Estimate (153) follows from part i) of Lemma 1 with u replaced by $\bar{u}_i$ and the generic nonlinearity b replaced by σ and by a, respectively, (using our assumptions (20)). More precisely, (153) follows from (22) by $[\bar{u}_i] \leq 1$. We now turn to (154), which by definitions (151) of $\delta\bar{M}$ and (156) of $\delta\tilde{M}$ and because of $N_0 \leq 1$ we may split into the four statements

$$M_{\sigma_1-\sigma_0} + [\sigma_1 - \sigma_0] + \|\sigma_1 - \sigma_0\| \lesssim M_{\bar{u}_1-\bar{u}_0} + [\bar{u}_1 - \bar{u}_0] + \|\bar{u}_1 - \bar{u}_0\|,$$
$$[\omega_1 - \omega_0] + \|\omega_1 - \omega_0\| \lesssim [\bar{\sigma}_1 - \bar{\sigma}_0] + \|\bar{\sigma}_1 - \bar{\sigma}_0\|$$
$$+ [\bar{u}_1 - \bar{u}_0] + \|\bar{u}_1 - \bar{u}_0\|,$$
$$M_{a_1-a_0} + [a_1 - a_0] + \|a_1 - a_0\| \lesssim M_{\bar{u}_1-\bar{u}_0} + [\bar{u}_1 - \bar{u}_0] + \|\bar{u}_1 - \bar{u}_0\|,$$
$$[\mu_1 - \mu_0] + \|\mu_1 - \mu_0\| \lesssim [\bar{\sigma}_1 - \bar{\sigma}_0] + \|\bar{\sigma}_1 - \bar{\sigma}_0\|$$
$$+ [\bar{u}_1 - \bar{u}_0] + \|\bar{u}_1 - \bar{u}_0\|,$$

all provided $\max_i(M_{\bar{u}_i} + [\bar{u}_i]) \leq 1$,

where we also used the definition (150) of $\bar{M}$. This is a consequence of part ii) of Lemma 1 with $(\bar{u}_i, \bar{\sigma}_i, \bar{a}_i)$ playing the role of (u_i, σ_i, a_i). The first two estimates follow from replacing the generic nonlinearity b by σ, the last two estimates from replacing it by a. The first and the third estimate are a consequence of (24), the second and fourth one of (25), in which we use (152). It is on all four we use our full assumptions (20) on the nonlinearities σ and a.

STEP 2. Using the off-line products, application of Corollary 1. We claim that under the hypothesis of part ii) of the theorem on the distributions f_j and the off-line products $v_i \diamond f_j$ and $v_i \diamond \partial_1^2 v_j$ we have the commutator estimates

(157) $$\sup_{T \leq 1}(T^{\frac{1}{4}})^{2-2\alpha}\|[\sigma_i, (\cdot)_T] \diamond f_j\| \lesssim N_0\tilde{M},$$

(158) $$\sup_{T \leq 1}(T^{\frac{1}{4}})^{2-2\alpha}\|[\sigma_i, (\cdot)_T] \diamond f_1 - [\sigma_i, (\cdot)_T] \diamond f_0\| \lesssim \delta N_0\tilde{M},$$

(159) $$\sup_{T \leq 1}(T^{\frac{1}{4}})^{2-2\alpha}\|[\sigma_1, (\cdot)_T] \diamond f_j - [\sigma_0, (\cdot)_T] \diamond f_j\| \lesssim N_0\delta\tilde{M},$$

(160) $$\sup_{T \leq 1}(T^{\frac{1}{4}})^{2-2\alpha}\|\{1, \frac{\partial}{\partial a_0}, \frac{\partial^2}{\partial a_0^2}\}[a_i, (\cdot)_T] \diamond \partial_1^2 v_j\| \lesssim N_0\tilde{M},$$

and

$$\sup_{T \le 1}(T^{\frac{1}{4}})^{2-2\alpha}\|\{1, \frac{\partial}{\partial a_0}\}\big([a_i, (\cdot)_T] \diamond \partial_1^2 v_1$$

(161)
$$- [a_i, (\cdot)_T] \diamond \partial_1^2 v_0\big)\| \lesssim \delta N_0 \tilde{M},$$

$$\sup_{T \le 1}(T^{\frac{1}{4}})^{2-2\alpha}\|\{1, \frac{\partial}{\partial a_0}\}\big([a_1, (\cdot)_T] \diamond \partial_1^2 v_j$$

(162)
$$- [a_0, (\cdot)_T] \diamond \partial_1^2 v_j\big)\| \lesssim N_0 \delta \tilde{M}.$$

This is an application of Corollary 1 with $(N_1, \delta N_1) = (N_0, \delta N_0)$. Estimate (157) is an application of Corollary 1 i) with u replaced by σ_i; the hypotheses (48) and (49) are contained in the theorem's assumptions (110) and (111) (note that f does not depend on an extra parameter a_0'). The output (50) turns into (157) since by definition (155), $M_{\sigma_i} + N_0 \le \tilde{M}$. Estimate (158) is an application of Corollary 1 ii) still applied with u replaced by σ_i; the hypotheses (51) and (52) are contained in the theorem's assumptions (118) and (119). The output (53) turns into (158) as in the previous application. Estimate (159) is an application of Corollary 1 iii) now applied with u_i replaced by σ_i (and thus (σ_i, a_i) replaced by $(\omega_i, \bar{a}_i)$); the hypotheses (54) and (55) are contained in the theorem's assumptions (117) and (120). The output (56) turns into (159), since by definition (156) we have

$$M_{\sigma_1-\sigma_0}$$
$$+ N_0([\omega_1 - \omega_0] + \|\omega_1 - \omega_0\| + [\bar{a}_1 - \bar{a}_0] + \|\bar{a}_1 - \bar{a}_0\|) + \delta N_0 \le \delta \tilde{M}.$$

The arguments for (160), (161), and (162) follow the same lines of those for (48), (158), and (159), respectively. The only difference is that in all instances, the distribution f_j is replaced by the family of distributions $\partial_1^2 v_j(\cdot, a_0)$ (and a_i plays the role of u in Corollary 1). Hence the hypotheses (48) and (51) in Corollary 1 turn into

$$\sup_{T \le 1}(T^{\frac{1}{4}})^{2-\alpha}\|\{1, \frac{\partial}{\partial a_0}, \frac{\partial^2}{\partial a_0^2}\}\partial_1^2 v_{jT}\| \lesssim N_0,$$

$$\sup_{T \le 1}(T^{\frac{1}{4}})^{2-\alpha}\|\{1, \frac{\partial}{\partial a_0}\}\partial_1^2(v_{1T} - v_{0T})\| \lesssim \delta N_0.$$

This follows from Step 1 in the proof of Corollary 2 via (18).

STEP 3. Application of Proposition 1. We claim that under the hypothesis of part ii) of the theorem regarding the distributions f_j and

the off-line products $v_i \diamond f_j$ and $v_i \diamond \partial_1^2 v_j$

$$(163) \qquad M \lesssim N_0(\tilde{M} + 1) \qquad\qquad \text{provided } \max_i[\bar{u}_i] \ll 1,$$

$$(164) \quad \max_i M_{u_i} \lesssim N_0\tilde{M} \qquad\qquad \text{provided } \max_i[\bar{u}_i] \ll 1,$$

$$(165) \qquad \delta M \lesssim N_0\delta\tilde{M} + \delta N_0 \qquad \text{provided in addition } \tilde{M} \lesssim 1,$$

$$(166) \quad M_{u_1-u_0} \lesssim N_0\delta\tilde{M} + \delta N_0\tilde{M} \qquad \text{provided in addition } \tilde{M} \lesssim 1,$$

where we define in consistency with (150) and (151)

$$(167) \qquad M := \max_i(M_{u_i} + [u_i]) + N_0,$$

$$\delta M := M_{u_1-u_0} + [u_1 - u_0] + \|u_1 - u_0\| + N_0\big([\sigma_1 - \sigma_0]$$
$$(168) \qquad + \|\sigma_1 - \sigma_0\| + [a_1 - a_0] + \|a_1 - a_0\|\big) + \delta N_0.$$

Indeed, (163) and (164) are an application of part i) of Proposition 1: The hypothesis (92) of the proposition is build into the definition (155) of $\tilde{M}$, so that $\tilde{M}$ here plays the role of N in the proposition. The hypothesis (93) is identical with the theorem's assumption (116), hypothesis (96) was established in (152), hypotheses (94) and (95) are contained in (157) and (160) of Step 2 which is consistent with $\tilde{M}$ playing the role of N there. The combination of (98) and (99) amounts to (163) by definition (167) of M. Estimate (98) by itself amounts to (164).

Estimate (165) in turn is a consequence of part ii) of Proposition 1: Hypothesis (102) of the proposition is build into the definition (156) of $\delta\tilde{M}$, so that $\delta\tilde{M}$ here plays the role of δN in the proposition. Hypotheses (100) and (101) are identical with (157) and (160) of Step 2. Hypothesis (103) is identical with our assumption (118), hypotheses (104), (105), (106), and (107) are identical with (158), (159), (161), and (162) in Step 2. The outcome (108) of the proposition turns into (166). The latter trivially for $\tilde{M} \lesssim 1$ implies

$$M_{u_1-u_0} \lesssim N_0\delta\tilde{M} + \delta N_0,$$

whereas the outcome (109) of the proposition assumes the form

$$[u_1 - u_0] + \|u_1 - u_0\| \lesssim N_0\delta\tilde{M} + \delta N_0.$$

By definition (156) of $\delta\tilde{M}$ we have

$$[\sigma_1 - \sigma_0] + \|\sigma_1 - \sigma_0\| + [a_1 - a_0] + \|a_1 - a_0\| \le \delta\tilde{M}.$$

The combination of the last three statement yields (165) in view of definition (168).

STEP 4. Under the assumptions of part ii) of the theorem on the distributions f_j and the off-line products $v_i \diamond f_j$ and $v_i \diamond \partial_1^2 v_j$, Step 1

and Step 3 obviously combine to

(169) $\qquad M \lesssim N_0(\bar{M}+1) \qquad\qquad\qquad$ provided $\max_i[\bar{u}_i] \ll 1,$

(170) $\quad \max_i M_{u_i} \lesssim N_0\bar{M} \qquad\qquad\qquad$ provided $\max_i[\bar{u}_i] \ll 1,$

(171) $\qquad \delta M \lesssim N_0\delta\bar{M} + \delta N_0 \qquad$ provided in addition $\bar{M} \leq 1,$

(172) $\qquad M_{u_1-u_0} \lesssim N_0\delta\bar{M} + \delta N_0\bar{M} \qquad$ provided in addition $\bar{M} \leq 1.$

STEP 5. Contraction mapping argument. We work under the assumptions of part ii) of the theorem on the distributions f_j and the off-line products $v_i \diamond f_j$, $v_i \diamond \partial_1^2 v_j$. In this step, we specify to the case of a single model $f_1 = f_0 =: f$ with the corresponding constant-coefficient solution v; this means that we may set $\delta N_0 = 0$.

We consider the space of all triplets $(\bar{u}, \bar{a}, \bar{\sigma})$, where $\bar{u}$ is modelled after v according to $\bar{a}$ and $\bar{\sigma}$, which fulfill the constraints (149), and which satisfy

(173) $$\bar{M} \leq N,$$

cf. (150), for some constant N to be fixed presently. We apply Step 4 to $(f_i, \bar{a}_i, \bar{\sigma}_i) = (f, \bar{a}, \bar{\sigma})$. From (173) and the definition (150) of $\bar{M}$ we learn that the proviso of (169) is fulfilled provided the constant N is sufficiently small, which we now fix accordingly. We thus learn from (169), which by (173) assumes the form of $M \lesssim N_0$, that the map defined through (148) sends the set defined through (173) into itself, provided $N_0 \ll 1$.

For two triplets (u_i, a_i, σ_i) as above we first note that

$$d\big((u_1, a_1, \sigma_1), (u_0, a_0, \sigma_0)\big) := M_{u_1-u_0} + [u_1 - u_0] + \|u_1 - u_0\|$$

(174) $\qquad + N_0([\sigma_1 - \sigma_0] + \|\sigma_1 - \sigma_0\| + [a_1 - a_0] + \|a_1 - a_0\|)$

defines a distance function. Indeed, that also the modelledness constant $M_{u_1-u_0}$ satisfies a triangle inequality in (u_i, a_i, σ_i) can be seen by rewriting the definition (21) as

$$\sup_{x,R} \frac{1}{R^{2\alpha}} \inf_{\ell} \sup_{y:d(x,y)\leq R} |u_1(y) - \sigma_1(x)v(u, a_1(x))$$
$$- (u_0(y) - \sigma_0(x)v(y, a_0(x))) - \ell(y)|$$

where ℓ runs over all linear functionals of the form $ay_1 + b$. We now apply Step 4 to the case of $(f_i, \bar{a}_i, \bar{\sigma}_i) = (f, \bar{a}_i, \bar{\sigma}_i)$. From (173) we learn that the proviso of (171) is fulfilled; because of $\delta N_0 = 0$, (171) assumes the form $\delta M \lesssim N_0\delta\bar{M}$. By definitions (151) and (168) of $\delta\bar{M}$ and δM, combined with $\delta N_0 = 0$, this turns into

$$d\big((u_1, a_1, \sigma_1), (u_0, a_0, \sigma_0)\big) \lesssim N_0 d\big((\bar{u}_1, \bar{a}_1, \bar{\sigma}_1), (\bar{u}_0, \bar{a}_0, \bar{\sigma}_0)\big).$$

Hence the map (148) is a contraction for $N_0 \ll 1$. We further note that the space of above triplets (u, a, σ) endowed with the distance function

(174) is *complete*; and that the subset defined through the constraints (149) and (173) is *closed*. Hence by the contraction mapping principle the map (148) admits a unique fixed point on the set defined through (149) and (173).

STEP 6. Conclusion on part i) of the theorem. Let u now be as in part i) of the theorem. We note that the assumptions of part i) on the distribution f and the off-line products $v \diamond f$, $v \diamond \partial_1^2 v$ turn into the assumptions of part ii) with $\delta N_0 = 0$. We claim that $(u, a(u), \sigma(u)) =: (u, a, \sigma)$ is a fixed point of the map (148), which is obvious, and which lies in the set defined through the constraints (149) and (173), and therefore is unique. Indeed, in view of $[a] \leq \|a'\|[u] \leq 1$, $[\sigma] \leq \|\sigma'\|[u] \leq 1$ by (20) and (114), the constraints (149) are satisfied. The constraint (173) would be an immediate consequence of the stronger statement (115) (provided N_0 is sufficiently small). We thus turn to this a priori estimate (115). We apply Step 4 to $(f_i, \bar{a}_i, \bar{\sigma}_i) = (f, a(u), \sigma(u))$. Since we are dealing with fixed points, we have $\bar{M} = M$. By the theorem's assumption $[u] \ll 1$, the provisos of (169) and (170) are satisfied so that because of $N_0 \ll 1$, their application yields

(175) $$M \lesssim N_0 \quad \text{and thus} \quad M_u \lesssim N_0^2.$$

By definition (163) this turns into (115).

STEP 7. Conclusion on part ii) of the theorem. Let u_i, $i = 0, 1$, now be as in part ii) of theorem. By Step 6, the two triplets $(u_i, a(u_i), \sigma(u_i)) =: (u_i, a_i, \sigma_i)$ satisfy the constraints (149) and (173) and each triplet is a fixed point of "its own" map (148) (which depends on i through the model f_i). We apply Step 4 to $(f_i, \bar{a}_i, \bar{\sigma}_i) = (f_i, a(u_i), \sigma(u_i))$. Since we are dealing with fixed points, we have $\bar{M} = M$ and $\delta\bar{M} = \delta M$. By the a priori estimate (115) and $N_0 \ll 1$, the two provisos of Step 4 are satisfied. Because of $N_0 \ll 1$, (171) and (172) turn into

$$\delta M \lesssim \delta N_0 \quad \text{and then} \quad M_{u_1 - u_0} \lesssim N_0 \delta N_0,$$

where we used (175). By definition (168) of δM, this turns into (121) and (122).

PROOF OF PROPOSITION 1.
We write for abbreviation $[\cdot] = [\cdot]_\alpha$. When a function v depends on a_0 next to x, we continue to write $\|v\|$ when we mean $\sup_{a_0} \|v(\cdot, a_0)\|$ and $[v]$ for $\sup_{a_0}[v(\cdot, a_0)]$. When we speak of a function u, we automatically mean that it is Hölder continuous with exponent α, that is, $[u] < \infty$; when we speak of a distribution f, we imply that it is of order $\alpha - 2$ in the sense of $\sup_{T \leq 1}(T^{\frac{1}{4}})^{2-\alpha}\|f_T\| < \infty$. When a distribution depends on the additional parameter a_0, we imply that the above bound is uniform in a_0.

STEP 1. Uniqueness. Under the assumptions of part i) of the proposition we claim that there is at most one mean-free u modelled after

v according to a and σ satisfying the equation (97). Indeed, let u' be another function with these properties; we trivially have by Definition 1 that $u - u'$ is modelled after v according to a and to 0 playing the role of σ. We now apply Lemma 4 with b replaced by a. We apply it three times, namely to u, to u', and to $u - u'$. We obtain from these three versions of (59) and the triangle inequality that

$$\lim_{T\downarrow 0} \|[a, (\cdot)_T] \diamond \partial_1^2 u - [a, (\cdot)_T] \diamond \partial_1^2 u' - [a, (\cdot)_T] \diamond \partial_1^2 (u - u')\| = 0$$

and thus $\lim_{T\downarrow 0} \|(a\diamond\partial_1^2 u - a\diamond\partial_1^2 u' - a\diamond\partial_1^2(u - u'))_T\| = 0$ so that $a\diamond\partial_1^2 u - a\diamond\partial_1^2 u' = a\diamond\partial_1^2(u - u')$. Hence we obtain from taking the difference of the equations:

$$(176) \qquad \partial_2(u - u') - Pa\diamond\partial_1^2(u - u') = 0.$$

We may also say that $u - u'$ is modelled after 0 playing the role of v and 0 playing the role of σ; we call δM the corresponding modelling constant. Hence we may apply Corollary 2 i) with $f = 0$ and thus $N_0 = 0$. We apply it with u replaced by $u - u'$ (and thus M by δM), which we may thanks to (176). In this context, the output (73) of Corollary 2 assumes the form $\delta M = 0$. Since $u - u'$ is periodic, we first infer $\delta v = 0$ and then $u - u' = const$. Since $u - u'$ has vanishing average, we obtain as desired $u - u' = 0$.

STEP 2. A special regularization. Under the assumptions of Lemma 4 and for $\tau > 0$ and $i = 1, \cdots, I$ we consider the convolution $v_{i\tau}$ of v_i and *define*

$$(177) \qquad a\diamond\partial_1^2 v_{i\tau} := (a\diamond\partial_1^2 v_i)_\tau.$$

Then, we claim that for any function u of class $C^{\alpha+2}$, which is modelled after $(v_{1\tau}, \cdots, v_{I\tau})$ according to a and $(\sigma_1, \cdots, \sigma_I)$, we have

$$(178) \qquad a\diamond\partial_1^2 u = a\partial_1^2 u - \sigma_i E[a, (\cdot)_\tau]\diamond\partial_1^2 v_i.$$

Indeed, by Lemma 4 (with b replaced by a) we understand the distribution $a\diamond\partial_1^2 u$ as defined by

$$(179) \qquad \lim_{T\downarrow 0} \|[a, (\cdot)_T]\diamond\partial_1^2 u - \sigma_i E[a, (\cdot)_T]\diamond\partial_1^2 v_{i\tau}\| = 0.$$

We note that (177) implies

$$(180) \qquad [a, (\cdot)_T]\diamond\partial_1^2 v_{i\tau} = [a, (\cdot)_{T+\tau}]\diamond\partial_1^2 v_i,$$

which ensures that $[a, (\cdot)_T]\diamond\partial_1^2 v_{i\tau} \to [a, (\cdot)_\tau]\diamond\partial_1^2 v_i$ as $T \downarrow 0$ uniformly in x for fixed a_0. Thanks to the bound on the $\frac{\partial}{\partial a_0}$-derivative in (58), this convergence is even uniform in (x, a_0), so that (179) turns into

$$\lim_{T\downarrow 0} \|[a, (\cdot)_T]\diamond\partial_1^2 u - \sigma_i E[a, (\cdot)_\tau]\diamond\partial_1^2 v_i\| = 0.$$

Since u is of class $C^{\alpha+2}$, this further simplifies to

$$\lim_{T\downarrow 0}\|a\partial_1^2 u - (a\diamond\partial_1^2 u)_T - \sigma_i E[a,(\cdot)_\tau]\diamond\partial_1^2 v_i\| = 0,$$

from which we learn that the distribution $a\diamond\partial_1^2 u$ is actually the function given by (178).

STEP 3. Existence in the regularized case. Under the assumptions of part i) of this proposition and in line with Step 2, for $\tau > 0$ we consider the mollification f_τ of f, so that v_τ satisfies $(\partial_2 - a_0\partial_1^2)v_\tau = Pf_\tau$, and complement definition (177) (without the index i) by

$$(181) \qquad \sigma\diamond f_\tau := (\sigma\diamond f)_\tau.$$

Then we claim that there exists a mean-free u^τ of class $C^{\alpha+2}$ modelled after v_τ according to a and σ such that

$$(182) \qquad \partial_2 u^\tau - P(a\diamond\partial_1^2 u^\tau + \sigma\diamond f_\tau) = 0 \quad \text{distributionally,}$$

and at the same time

$$(183) \quad \partial_2 u^\tau - P(a\partial_1^2 u^\tau - \sigma E[a,(\cdot)_\tau]\diamond\partial_1^2 v + (\sigma\diamond f)_\tau) = 0 \quad \text{classically.}$$

We first turn to the existence of (183) and start by noting that the rhs $-\sigma E[a,(\cdot)_\tau]\diamond\partial_1^2 v + (\sigma\diamond f)_\tau$ in (183) is of class C^α. Leveraging upon $[a]\ll 1$ we rewrite the equation as $\partial_2 u^\tau - a_0\partial_1^2 u^\tau = P((a-a_0)\partial_1^2 u$ $-\sigma E[a,(\cdot)_\tau]\diamond\partial_1^2 v + (\sigma\diamond f)_\tau)$ for $a_0 = a(0)$. Using the invertibility of the constant-coefficient operator $\partial_2 - a_0\partial_1^2$ on periodic mean-free functions, and equipped with the corresponding Schauder estimates, see for instance [12, Theorem 8.6.1] lifted to the torus, we see that a solution of class $C^{\alpha+2}$ exists, using a contraction mapping argument based on $\|a-a_0\|\ll 1$. Since both u^τ and $v_\tau(\cdot,a_0)$ are in particular of class $C^{\alpha+1}$, u is modelled after v_τ according to — in fact any — a and σ. By Step 2 and definition (181) we see that (183) may be rewritten as (182).

STEP 4. Basic construction. We now work under the assumptions of part ii) of the proposition. We interpolate the functions σ_i, a_i, and v_i as well as the distribution f_i linearly:

$$(184) \qquad \sigma_s := s\sigma_1 + (1-s)\sigma_0 \quad \text{and same for } a, f, \text{ and } v.$$

We note that this preserves (96). We interpolate the products bilinearly

$$\sigma_s\diamond f_s := s^2\sigma_1\diamond f_1 + s(1-s)\sigma_1\diamond f_0$$
$$+ (1-s)s\sigma_0\diamond f_1 + (1-s)^2\sigma_0\diamond f_0,$$
$$\partial_s\sigma\diamond f_s := s\sigma_1\diamond f_1 + (1-s)\sigma_1\diamond f_0 - s\sigma_0\diamond f_1 - (1-s)\sigma_0\diamond f_0,$$
$$\sigma_s\diamond\partial_s f := s\sigma_1\diamond f_1 - s\sigma_1\diamond f_0 + (1-s)\sigma_0\diamond f_1 - (1-s)\sigma_0\diamond f_0,$$
$$(185) \qquad \text{and same for } a_s\diamond\partial_1^2 v_s, \ \partial_s a\diamond\partial_1^2 v_s \text{ and } a_s\diamond\partial_1^2\partial_s v.$$

Thanks to the estimate (101), which is preserved under bilinear interpolation, the family of distributions $\{a_s \diamond \partial_1^2 v_s(\cdot, a_0)\}_{a_0}$ is continuously differentiable in a_0 so that we may define

$$(186) \qquad a_s \diamond \partial_1^2 \frac{\partial v_s}{\partial a_0}(\cdot, a_0) := \frac{\partial}{\partial a_0} a_s \diamond \partial_1^2 v_s(\cdot, a_0).$$

For given $0 < \tau \leq 1$, we define the singular products with the regularized distributions as in Step 2, namely

$$\sigma_s \diamond f_{s\tau} := (\sigma_s \diamond f_s)_\tau \quad \text{and same for } \partial_s \sigma \diamond f_{s\tau}, \ \sigma_s \diamond \partial_s f_\tau,$$

$$(187) \qquad a_s \diamond \partial_1^2 v_{s\tau}, \ \partial_s a \diamond \partial_1^2 v_{s\tau}, \ a_s \diamond \partial_1^2 \partial_s v_\tau, \ a_s \diamond \partial_1^2 \frac{\partial v_{s\tau}}{\partial a_0}.$$

We claim that there exists a curve u_s^τ of mean-free functions continuously differentiable in s wrt to the class $C^{\alpha+2}$ such that

$$(188) \qquad u_s^\tau \text{ is modelled after } v_{s\tau} \text{ according to } a_s \text{ and } \sigma_s$$

and satisfies

$$(189) \qquad \partial_2 u_s^\tau - P(a_s \diamond \partial_1^2 u_s^\tau + \sigma_s \diamond f_{s\tau}) = 0 \quad \text{distributionally.}$$

Furthermore, we claim that

$$\partial_s u^\tau \text{ is modelled after } (v_{s\tau}, \frac{\partial v_{s\tau}}{\partial a_0}, \partial_s v_\tau)$$

$$(190) \qquad \text{according to } a_s \text{ and } (\partial_s \sigma, \sigma_s \partial_s a, \sigma_s)$$

and satisfies

$$(191) \quad \partial_2 \partial_s u^\tau - P(a_s \diamond \partial_1^2 \partial_s u^\tau + \partial_s a \diamond \partial_1^2 u_s^\tau + \sigma_s \diamond \partial_s f_\tau + \partial_s \sigma \diamond f_{s\tau}) = 0$$

distributionally.

By Steps 3 and 1 and our definitions of $\sigma_s \diamond f_{s\tau}$ and $a_s \diamond \partial_1^2 v_{s\tau}$ by convolution, cf (187), there exists a unique mean-free u_s^τ of class $C^{\alpha+2}$ such that (188) and (189) hold. Furthermore by Step 2 u_s^τ is characterized as the classical solution of

$$(192) \qquad \partial_2 u_s^\tau - P(a_s \partial_1^2 u_s^\tau - \sigma_s E_s[a_s, (\cdot)_\tau] \diamond \partial_1^2 v_s + (\sigma_s \diamond f_s)_\tau) = 0.$$

In preparation of taking the s-derivative of (192) we note that the definition (185) of $\sigma_s \diamond f_s$ and $a_s \diamond \partial_1^2 v_s$ by (bi-)linear interpolation ensures that Leibniz' rule holds:

$$(193) \qquad \partial_s(\sigma_s \diamond f_s) = \partial_s \sigma \diamond f_s + \sigma_s \diamond \partial_s f,$$

$$(194) \qquad \partial_s(a_s \diamond \partial_1^2 v_s) = \partial_s a \diamond \partial_1^2 v_s + a_s \diamond \partial_1^2 \partial_s v.$$

We recall that E_s denotes the evaluation operator that evaluates a function of (x, a_0) at $(x, a_s(x))$; with the obvious commutation rule

$[\partial_s, E_s] = (\partial_s a) E_s \frac{\partial}{\partial a_0}$ we obtain from (194) and (186)

$$\partial_s(E_s a_s \diamond \partial_1^2 v_s)$$

$$= E_s \partial_s a \diamond \partial_1^2 v_s + \partial_s a E_s a_s \diamond \partial_1^2 \frac{\partial v_s}{\partial a_0} + E_s a_s \diamond \partial_1^2 \partial_s v,$$

which in conjunction with the classical differentiation rules extends to the commutator:

$$\partial_s(E_s[a_s, (\cdot)_\tau] \diamond \partial_1^2 v_s) = E_s[\partial_s a, (\cdot)_\tau] \diamond \partial_1^2 v_s$$

(195)
$$+ \partial_s a E_s[a_s, (\cdot)_\tau] \diamond \partial_1^2 \frac{\partial v_s}{\partial a_0} + E_s[a_s, (\cdot)_\tau] \diamond \partial_1^2 \partial_s v.$$

Equipped with (193) and (195) we learn from (192) that u_s^τ is differentiable in s with values in the class $C^{\alpha+2}$ and

$$\partial_2 \partial_s u^\tau - P\big(a_s \partial_1^2 \partial_s u^\tau + \partial_s a \partial_1^2 u_s^\tau - \sigma_s E_s[\partial_s a, (\cdot)_\tau] \diamond \partial_1^2 v_s$$

$$- \partial_s \sigma E_s[a_s, (\cdot)_\tau] \diamond \partial_1^2 v_s - \sigma_s \partial_s a E_s[a_s, (\cdot)_\tau] \diamond \partial_1^2 \frac{\partial v_s}{\partial a_0}$$

(196)
$$- \sigma_s E_s[a_s, (\cdot)_\tau] \diamond \partial_1^2 \partial_s v + (\partial_s \sigma \diamond f_s)_\tau + (\sigma_s \diamond \partial_s f)_\tau\big) = 0.$$

Moreover, like in Step 3, (190) holds automatically because of the regularity of $\partial_s u^\tau$ and of $(v_{s\tau}, \frac{\partial v_{s\tau}}{\partial a_0}, \partial_s v_\tau)$. In view of the definition (187) of $\partial_s a \diamond \partial_1^2 v_{s\tau}$ we have by Step 2 applied to u_s^τ modelled according to (188)

$$\partial_s a \diamond \partial_1^2 u_s^\tau = \partial_s a \partial_1^2 u_s^\tau - \sigma_s E_s[\partial_s a, (\cdot)_\tau] \diamond \partial_1^2 v_s.$$

In view of the similar definition of $a_s \diamond \partial_1^2 \partial_s v$, $a_s \diamond \partial_1^2 \frac{\partial v_{s\tau}}{\partial a_0}$, and $a_s \diamond \partial_1^2 \partial_s v_\tau$ we have by Step 2 applied to $\partial_s u^\tau$ modelled according to (190)

$$a_s \diamond \partial_1^2 \partial_s u^\tau = a_s \partial_1^2 \partial_s u^\tau - \partial_s \sigma E_s[a_s, (\cdot)_\tau] \diamond \partial_1^2 v_s$$

$$- \sigma_s \partial_s a E_s[a_s, (\cdot)_\tau] \diamond \partial_1^2 \frac{\partial v_s}{\partial a_0} - \sigma_s E_s[a_s, (\cdot)_\tau] \diamond \partial_1^2 \partial_s v.$$

Plugging these two formulas and the definition (187) of $\partial_s \sigma \diamond f_\tau$ and $\sigma_s \diamond \partial_s f_\tau$ into (196), we obtain (191).

STEP 5. We now work under the assumptions of part ii) of the proposition. We claim

(197)
$$[\sigma_s] + [a_s] \leq N,$$

(198)
$$\sup_{T \leq 1}(T^{\frac{1}{4}})^{2-\alpha} \|(f_{s\tau})_T\| \lesssim N_0,$$

(199)
$$\sup_{T \leq 1}(T^{\frac{1}{4}})^{2-2\alpha} \|[\sigma_s, (\cdot)_T] \diamond (f_{s\tau})_T\| \lesssim N N_0,$$

(200)
$$\sup_{T \leq 1}(T^{\frac{1}{4}})^{2-2\alpha} \sup_{a_0} \|\{1, \frac{\partial}{\partial a_0}, \frac{\partial^2}{\partial a_0^2}\}[a_s, (\cdot)_T] \diamond \partial_1^2 v_{s\tau}\| \lesssim N N_0$$

and

$$(201) \qquad\qquad [\partial_s \sigma] + \|\partial_s \sigma\| + [\partial_s a] + \|\partial_s a\| \leq \delta N,$$

$$(202) \qquad\qquad \sup_{T \leq 1}(T^{\frac{1}{4}})^{2-\alpha}\|\partial_s (f_\tau)_T\| \leq \delta N_0,$$

$$(203) \qquad\qquad \sup_{T \leq 1}(T^{\frac{1}{4}})^{2-2\alpha}\|[\sigma_s, (\cdot)_T] \diamond \partial_s f_\tau\| \lesssim N \delta N_0,$$

$$(204) \qquad\qquad \sup_{T \leq 1}(T^{\frac{1}{4}})^{2-2\alpha}\|[\partial_s \sigma, (\cdot)_T] \diamond f_{s\tau}\| \lesssim \delta N N_0,$$

$$(205) \qquad \sup_{T \leq 1}(T^{\frac{1}{4}})^{2-2\alpha} \sup_{a_0} \|\{1, \frac{\partial}{\partial a_0}\}[a_s, (\cdot)_T] \diamond \partial_1^2 \partial_s v_\tau\| \lesssim N \delta N_0,$$

$$(206) \qquad \sup_{T \leq 1}(T^{\frac{1}{4}})^{2-2\alpha} \sup_{a_0} \|\{1, \frac{\partial}{\partial a_0}\}[\partial_s a, (\cdot)_T] \diamond \partial_1^2 v_{s\tau}\| \lesssim \delta N N_0.$$

Indeed, (197) and (201) are immediate from (92) (with i) and (102), respectively, by the linear interpolation (184). For $\tau = 0$ the remaining estimates, even with $\lesssim$ replaced by $\leq$, follow from the linear and bilinear interpolations (184) and (185) from the assumptions of this proposition: inequality (198) from (93) (with i), (199) from (100), (200) from (101). Still for $\tau = 0$, the five estimates (202), (203), (204), (205), and (206), are direct consequences of (103), (104), (105), (106), and (107), respectively.

It remains to pass from $\tau = 0$ to $0 < \tau \leq 1$ in the eight estimates of this step, based on our definition (187) of singular products. This is done with help of the next step.

STEP 6. Let the (generic) function u and the (generic) distributions f and $u \diamond f$ be such that

$$(207) \qquad [u] \leq N_0, \quad \sup_{T \leq 1}(T^{\frac{1}{4}})^{2-\alpha}\|f_T\| \leq N_1 \quad \text{and}$$

$$(208) \qquad \sup_{T \leq 1}(T^{\frac{1}{4}})^{2-2\alpha}\|[u, (\cdot)_T] \diamond f\| \leq N_0 N_1$$

for some constants N_0 and N_1. Then we claim that for $\tau \leq 1$ the distributions f_τ and $u \diamond f_\tau := (u \diamond f)_\tau$ satisfy the same estimates:

$$(209)$$
$$\sup_{T \leq 1}(T^{\frac{1}{4}})^{2-\alpha}\|(f_\tau)_T\| \lesssim N_1, \quad \sup_{T \leq 1}(T^{\frac{1}{4}})^{2-2\alpha}\|[u, (\cdot)_T] \diamond f_\tau\| \lesssim N_0 N_1.$$

Indeed, by definition of $u \diamond f_\tau$ we have like for (180) by the semi-group property

$$[u, (\cdot)_T] \diamond f_\tau = [u, (\cdot)_{T+\tau}] \diamond f,$$

so that (209) follows automatically provided we can show that (207) & (208) extend from the range of $T \leq 1$ to the range $T \leq 2$ in form of

$$(210) \quad \sup_{T \leq 1} (T^{\frac{1}{4}})^{2-\alpha} \|f_{2T}\| \lesssim N_1, \quad \sup_{T \leq 1} (T^{\frac{1}{4}})^{2-2\alpha} \|[u, (\cdot)_{2T}] \diamond f\| \lesssim N_0 N_1.$$

For this, we appeal to the semi-group property giving us

$$f_{2T} = (f_T)_T \quad \text{and} \quad [u, (\cdot)_{2T}] \diamond f = ([u, (\cdot)_T] \diamond f)_T + [u, (\cdot)_T] f_T,$$

so that by the boundedness of $(\cdot)_T$ in $\| \cdot \|$ indeed (208) entails (210), appealing to (268) and using in addition that by (207)

$$\|[u, (\cdot)_T] f_T\| \lesssim N_0 (T^{\frac{1}{4}})^{\alpha} \|f_T\| \lesssim N_0 N_1 (T^{\frac{1}{4}})^{2\alpha - 2}.$$

STEP 7. Application of Corollary 2. We claim for the modelling and Hölder constants of u_s^τ and $\partial_s u^\tau$:

$$(211) \qquad\qquad M_s^\tau \lesssim N_0 N,$$

$$(212) \qquad\qquad [u_s^\tau] \lesssim N_0(N + 1),$$

$$(213) \qquad\qquad \delta M_s^\tau \lesssim N_0 \delta N + \delta N_0 N \quad \text{provided } N \leq 1,$$

$$(214) \qquad\qquad [\partial_s u^\tau] \lesssim N_0 \delta N + \delta N_0 \quad \text{provided } N \leq 1.$$

Indeed, for estimates (211) and (212) we apply Corollary 2 i) with $(f, v, \sigma, a, \sigma \diamond f, a \diamond \partial_1^2 v, u)$ replaced by $(f_{s\tau}, v_{s\tau}, \sigma_s, a_s, \sigma_s \diamond f_{s\tau}, a_s \diamond \partial_1^2 v_{s\tau}, u_s^\tau)$ (where it is clear that linear interpolation and convolution preserves the relation between $f_{s\tau}$ and $v_{s\tau}$ through the constant coefficient equation). As already remarked in Step 4 the constraints (96) turn into (71) under the linear interpolation (184). The hypotheses (66), (67), (68), and (69) were established in Step 5, cf (197), (198), (199), and (200), respectively. Hypothesis (72) and the modelledness are clear by construction, cf (189) and (188) in Step 4. The outputs (73) and (74) assume the form (211) and (212).

For the remaining estimates (213) and (214), we apply Corollary 2 ii) with $(\delta f, \delta v, \delta \sigma, \delta a, \sigma \diamond \delta f, \delta \sigma \diamond f, a \diamond \partial_1^2 \delta v, \delta a \diamond \partial_1^2 v, \delta u)$ replaced by $(\partial_s f_\tau, \partial_s v_\tau, \partial_s \sigma, \partial_s a, \sigma_s \diamond \partial_s f_\tau, \partial_s \sigma \diamond f_{s\tau}, a_s \diamond \partial_1^2 \partial_s v_\tau, \partial_s a \diamond \partial_1^2 v_{s\tau}, \partial_s u^\tau)$. The six hypotheses (75)–(80) were established in Step 5, cf (201)– (206). Hypothesis (82) and the corresponding modelledness are clear by construction, cf (191) and (190) in Step 4. The outputs (83) and (84) assume the form of (213) and (214).

STEP 8. Integration. We claim that $u_1^\tau - u_0^\tau$ is modelled after (v_1^τ, v_0^τ) according to (a_1, a_0) and $(\sigma_1, -\sigma_0)$ with the modelling constant and Hölder constant estimated as follows

$$(215) \qquad\qquad \delta M^\tau \lesssim N_0 \delta N + \delta N_0 N \quad \text{provided } N \leq 1,$$

$$(216) \qquad\qquad [u_1^\tau - u_0^\tau] \lesssim N_0 \delta N + \delta N_0 \quad \text{provided } N \leq 1.$$

Indeed, the Hölder estimate (216) is obvious from (214) by integration in $s \in [0,1]$. The estimate on the modelling constant relies on the differentiation rule

$$\frac{\partial}{\partial s}\big(u_s^\tau(y) - \sigma_s(x)v_{s\tau}(y, a_s(x))\big) = \partial_s u^\tau(y) - (\partial_s\sigma)(x)v_{s\tau}(y, a_s(x))$$

$$- (\sigma_s\partial_s a)(x)\frac{\partial v_{s\tau}}{\partial a_0}(y, a_s(x)) - \sigma_s(x)\partial_s v_\tau(y, a_s(x)),$$

and on defining $\nu := \int_0^1 \nu_s ds$, where ν belongs to $u_1^\tau - u_0^\tau$ and ν_s to $\partial_s u^\tau$ in the sense of Definition 1. This provides the link between (213) and (215) by integration.

STEP 9. Passage to limit. We claim that we may pass to the limit $\tau \downarrow 0$ in (211) and (212) with $s = 0, 1$, recovering (98) and (99) in part i) of this proposition, and in (215) and (216), recovering (108) and (109) in part ii) of the proposition. Clearly, from the uniform-in-τ estimate (212) (in conjunction with the vanishing mean of u_i^τ which provides the same bound on the supremum norm) we learn by Arzelà-Ascoli that there exists a subsequence $\tau \downarrow 0$ (unchanged notation) and a continuous mean-free function u_i to which u_i^τ converges uniformly. Hence we may pass to the limit in the Hölder estimates (212) and (216). Since also the convolution $v_{i\tau}$ converges to v_i uniformly, we may pass to the limit in the estimates (211) and (215) of the modelling constants. By uniqueness, cf Step 1, it thus remains to argue that u_i solves (97) (with (f, σ, a) replaced by (f_i, σ_i, a_i)). In order to pass from (189) to (97) it remains to establish the distributional convergences

$$(217) \qquad\qquad \sigma_i \diamond f_{i\tau} \rightharpoonup \sigma_i \diamond f_i,$$

$$(218) \qquad\qquad a_i \diamond \partial_1^2 u_i^\tau \rightharpoonup a_i \diamond \partial_1^2 u_i.$$

The convergence (217) is build-in by the definition (187) through convolution. One of the ingredients for the convergence (218) is the analogue of (217)

$$a_i \diamond \partial_1^2 v_{i\tau}(\cdot, a_0) \rightharpoonup a_i \diamond \partial_1^2 v_i(\cdot, a_0),$$

which in conjunction with the pointwise convergence of $v_{i\tau}$ extends to the commutator

$$[a_i, (\cdot)_T] \diamond \partial_1^2 v_{i\tau}(\cdot, a_0) \to [a_i, (\cdot)_T] \diamond \partial_1^2 v_i(\cdot, a_0).$$

Since $\|\frac{\partial}{\partial a_0}[a_i, (\cdot)_T] \diamond \partial_1^2 v_{i\tau}(\cdot, a_0)\|$ is uniformly bounded, cf (101) and (187) in conjunction with a formula of type (180), we even have

$$[a_i, (\cdot)_T] \diamond \partial_1^2 v_{i\tau}(\cdot, a_0) \to [a_i, (\cdot)_T] \diamond \partial_1^2 v_i(\cdot, a_0) \quad \text{uniformly in } a_0,$$

so that

$$\sigma_i E_i[a_i, (\cdot)_T] \diamond \partial_1^2 v_{i\tau} \to \sigma_i E_i[a_i, (\cdot)_T] \diamond \partial_1^2 v_i.$$

In order to relate this to (218) we appeal to the modelledness of u_i wrt to v_i according to a_i and σ_i which by (59) in Lemma 4 yields

$$\lim_{T\downarrow 0} \|[a_i, (\cdot)_T] \diamond \partial_1^2 u_i - \sigma_i E_i[a_i, (\cdot)_T] \diamond \partial_1^2 v_i\| = 0.$$

Likewise, the uniform modelledness of u_i^τ, cf (213), in conjunction with the uniform commutator bounds (95) and the uniform bounds on $v_{i\tau}$, we have, again by (59) in Lemma 4, the uniform convergence

$$\lim_{T\downarrow 0} \sup_\tau \|[a_i, (\cdot)_T] \diamond \partial_1^2 u_i^\tau - \sigma_i E_i[a_i, (\cdot)_T] \diamond \partial_1^2 v_{i\tau}\| = 0.$$

The combination of the three last statements implies

$$\lim_{T\downarrow 0} \limsup_{\tau\downarrow 0} \|[a_i, (\cdot)_T] \diamond \partial_1^2 u_i^\tau - [a_i, (\cdot)_T] \diamond \partial_1^2 u_i\| = 0,$$

which by the convergence of u_i^τ yields

(219) $$\lim_{T\downarrow 0} \limsup_{\tau\downarrow 0} \|(a_i \diamond \partial_1^2 u_i^\tau - a_i \diamond \partial_1^2 u_i)_T\| = 0.$$

Now the next step shows that this implies (218).

STEP 10. Suppose that the sequence $\{f_n\}_{\tau\downarrow 0}$ of uniformly bounded distributions satisfies

$$\lim_{T\downarrow 0} \limsup_{n\uparrow\infty} \|f_{nT}\| = 0.$$

We claim that this implies distributional convergence:

$$f_n \rightharpoonup 0.$$

Indeed, we have for fixed $T > 0$ and any $\tau \leq T$ that $\|f_{nT}\| \lesssim \|f_{n\tau}\|$ and therefore $\limsup_{n\uparrow\infty} \|f_{nT}\| \lesssim \limsup_{n\uparrow 0} \|f_{n\tau}\|$ and $\limsup_{n\uparrow 0} \|f_{nT}\| \lesssim \lim_{\tau\downarrow 0} \limsup_{n\uparrow 0} \|f_{n\tau}\|$. The latter is equal to zero by assumption. Hence we have $f_{nT} \to 0$ for every $T > 0$, which yields the claim by the uniform boundedness of f_n in the sense of $\sup_{T\leq 1}(T^{\frac{1}{4}})^{2-\alpha}\|f_T\|$, and then also in the more classical $C^{\alpha-2}$-norm, cf (336) in Step 1 of Lemma 9.

PROOF OF COROLLARY 2.
We write $[\cdot]$ for $[\cdot]_\alpha$.
STEP 1. Application of Lemma 9. We claim

(220) $$\sup_{a_0}[\{1, \frac{\partial}{\partial a_0}, \frac{\partial^2}{\partial a_0^2}\}v] \lesssim N_0,$$

(221) $$\sup_{a_0}[\{1, \frac{\partial}{\partial a_0}\}\delta v] \lesssim \delta N_0.$$

The estimate (220) is based on the two identities following from differentiating (70) twice wrt a_0

(222) $$(\partial_2 - a_0 \partial_1^2)(v, \frac{\partial v}{\partial a_0}, \frac{\partial^2 v}{\partial a_0^2}) = (Pf, \partial_1^2 v, 2\partial_1^2 \frac{\partial v}{\partial a_0}).$$

We now see that (220) follows by an iterated application of Lemma 9: From (67) we first obtain the bound on v by Lemma 9, then the bound on $\partial_1^2 v_T$ by (18), then via (222) the bound on $\frac{\partial v}{\partial a_0}$ by Lemma 9, then the bound on $\partial_1^2 \frac{\partial v_T}{\partial a_0}$ by (18), then via (222) finally the bound on $\frac{\partial^2 v}{\partial a_0^2}$ by Lemma 9. The argument for (221) is identical, just with (f, v) replaced by $(\delta f, \delta v)$, cf (81), and starting from (76) instead of (67) and thus with N_0 replaced by δN_0.

STEP 2. Application of Lemma 4. We claim that

$$(223) \qquad \sup_{T \leq 1}(T^{\frac{1}{4}})^{2-2\alpha}\|[a, (\cdot)_T] \diamond \partial_1^2 u\| \lesssim [a]M + N N_0,$$

$$(224) \qquad \sup_{T \leq 1}(T^{\frac{1}{4}})^{2-2\alpha}\|[\delta a, (\cdot)_T] \diamond \partial_1^2 u\| \lesssim [\delta a]M + \delta N N_0,$$

$$(225) \qquad \sup_{T \leq 1}(T^{\frac{1}{4}})^{2-2\alpha}\|[a, (\cdot)_T] \diamond \partial_1^2 \delta u\| \lesssim [a]\delta M + N(N_0 \delta N + \delta N_0).$$

Here comes the argument: Estimate (223) follows from Lemma 4 with b replaced by a, $I = 1$ and $v_{i=1} = v$, so that the hypothesis (57) is satisfied by (220) in Step 1 with N_0 playing the role of $N_{i=1}$. Hypothesis (58) is satisfied by our assumption (69) with N playing the role of N_0. In view of (71), the outcome (60) of Lemma 4 turns into (223).

Estimate (224) follows from applying Lemma 4 with b replaced by δa, still $I = 1$, $v_{i=1} = v$, and N_0 playing the role of $N_{i=1}$. Hypothesis (58) is satisfied by our assumption (80) with δN playing the role of N_0. In view of (71), the outcome (60) of Lemma 4 turns into (224).

Finally, estimate (225) follows from applying Lemma 4 with b again replaced by a, but this time $I = 3$ and $(v_1, v_2, v_3) = (v, \frac{\partial v}{\partial a_0}, \delta v)$. We learn from Step 1 that hypothesis (57) is satisfied with $(N_1, N_2, N_3) = (N_0, N_0, \delta N_0)$. We now turn to the hypothesis (58): For $i = 1$ it is contained in our assumption (69) with N playing the role of N_0. In preparation of checking hypothesis (58) for $i = 2$ we note that our assumption (69) implies in particular that the family of distributions $\{a \diamond \partial_1^2 v(\cdot, a_0)\}_{a_0}$ is continuously differentiable in a_0. This allows us to *define* the family of distributions $\{a \diamond \partial_1^2 \frac{\partial v}{\partial a_0}(\cdot, a_0)\}_{a_0}$ via

$$a \diamond \partial_1^2 \frac{\partial v}{\partial a_0} := \frac{\partial}{\partial a_0} a \diamond \partial_1^2 v,$$

which extends to the commutator:

$$(226) \qquad [a, (\cdot)_T] \diamond \partial_1^2 \frac{\partial v}{\partial a_0} = \frac{\partial}{\partial a_0}[a, (\cdot)_T] \diamond \partial_1^2 v.$$

Hence the hypothesis (58) for $i = 2$ is also satisfied by (69) (here we use it up to $\frac{\partial^2}{\partial a_0^2}$). Hypothesis (58) for $i = 3$ is identical to our assumption (79). We apply Lemma 4 with δu playing the role of u; the triple

$(\delta\sigma, \sigma\delta a, \sigma)$ then plays the role of $(\sigma_1, \sigma_2, \sigma_3)$ and δM that of M. The outcome (60) of Lemma 4 assumes the form

$$\sup_{T\leq 1}(T^{\frac{1}{4}})^{2-2\alpha}\|[a,(\cdot)_T]\diamond\partial_1^2\delta u\|$$

$$\lesssim [a]\delta M + N\big(N_0([\delta\sigma] + \|\delta\sigma\| + [\sigma\delta a] + \|\sigma\delta a\|)$$

(227) $$+ \delta N_0([\sigma] + \|\sigma\|)\big).$$

We note that by (71) and (75) we have

$$N_0([\delta\sigma] + \|\delta\sigma\| + [\sigma\delta a] + \|\sigma\delta a\|) + \delta N_0([\sigma] + \|\sigma\|)$$
$$\lesssim N_0([\delta\sigma] + \|\delta\sigma\| + [\delta a] + \|\delta a\|) + \delta N_0$$
$$\lesssim N_0\delta N + \delta N_0,$$

so that (227) yields (225).

STEP 3. Commutator estimates. We claim

(228) $$\sup_{T\leq 1}(T^{\frac{1}{4}})^{2-2\alpha}\|\partial_2 u_T - P(a\partial_1^2 u_T + \sigma f_T)\| \lesssim [a]M + NN_0,$$

$$\sup_{T\leq 1}(T^{\frac{1}{4}})^{2-2\alpha}\|\partial_2\delta u_T - P(a\partial_1^2\delta u_T + \sigma\delta a E\partial_1^2 v_T + \sigma\delta f_T + \delta\sigma f_T)\|$$

(229) $$\lesssim [a]\delta M + ([\delta a] + \|\delta a\|)M + N(N_0\delta N + \delta N_0) + \delta N N_0.$$

Indeed, we apply $(\cdot)_T$ to (72) and rearrange terms:

(230) $$\partial_2 u_T - P(a\partial_1^2 u_T + \sigma f_T) = -P([a,(\cdot)_T]\diamond\partial_1^2 u + [\sigma,(\cdot)_T]\diamond f).$$

Similarly, we apply $(\cdot)_T$ to (82) and rearrange terms:

$$\partial_2\delta u_T - P(a\partial_1^2\delta u_T + \sigma\delta a E\partial_1^2 v_T + \sigma\delta f_T + \delta\sigma f_T)$$
$$= -P\Big(-\delta a(\partial_1^2 u_T - \sigma E\partial_1^2 v_T)$$
$$+ [a,(\cdot)_T]\diamond\partial_1^2\delta u + [\delta a,(\cdot)_T]\diamond\partial_1^2 u$$

(231) $$+ [\sigma,(\cdot)_T]\diamond\delta f + [\delta\sigma,(\cdot)_T]\diamond f\Big).$$

By assumption (68) and by (223) in Step 2 we obtain estimate (228) from identity (230). By assumptions (77) and (78) and by (224) and (225) from Step 2 and from writing

$$(\partial_1^2 u_T - \sigma E\partial_1^2 v_T)(x) = \int dy\,\partial_1^2\psi_T(x-y)$$
$$\times \big((u(y)-u(x)) - \sigma(x)(v(y,a(x))-v(x,a(x))) - \nu(x)(y-x)_1\big),$$

which entails with help of (18)

$$\sup_{T\leq 1}(T^{\frac{1}{4}})^{2-2\alpha}\|\delta a(\partial_1^2 u_T - \sigma E\partial_1^2 v_T)\| \lesssim \|\delta a\|M,$$

we obtain (229) from (231).

STEP 4. Application of Lemma 5 and conclusion. We first apply Lemma 5 with $I = 1$ and f playing the role of $f_{i=1}$ (which does not

depend on a_0). The hypothesis (61) is ensured by our assumption (67) with N_0 playing the role of $N_{i=1}$. The hypothesis (63) is settled through (228) in Step 3 with N^2 given by $[a]M + NN_0$. Hence the two outputs (64) and (65) of Lemma 5 take the form of

$$(232) \qquad M \lesssim [a]M + NN_0 + N_0([\sigma] + \|\sigma\|[a]),$$

$$(233) \qquad [u] \lesssim M + N_0\|\sigma\|.$$

The smallness of $[a]$ and the boundedness of $\|\sigma\|$, cf (71), imply that (232) simplifies to $M \lesssim NN_0 + N_0([\sigma] + [a])$, which by (66) means (73). Inserting (73) into (233) and using once more $\|\sigma\| \leq 1$ yields (74).

We now apply Lemma 5 with $I = 3$ and $(f, \partial_1^2 v, \delta f)$ playing the role of (f_1, f_2, f_3). In view of assumption (70), of (222) in Step 1, and of assumption (81), the triplet $(v, \frac{\partial v}{\partial a_0}, \delta v)$ plays the role of (v_1, v_2, v_3); by (220) & (221) in Step 1, it also satisfies the estimates (61) with $(N_1, N_2, N_3) = (N_0, N_0, \delta N_0)$. We apply Lemma 5 to δu playing the role of u, $(\delta\sigma, \sigma\delta a, \sigma)$ playing the role of $(\sigma_1, \sigma_2, \sigma_3)$, and δM playing the role of M. The hypothesis (63) is settled through Step 3 with N^2 estimated by the rhs of (229). Hence the two outputs (64) and (65) of Lemma 5 take the form

$$\delta M \lesssim \text{expression on rhs of (229)} + [a]\delta M$$
$$\quad + N_0([\delta\sigma] + \|\delta\sigma\|[a] + [\sigma\delta a] + \|\sigma\delta a\|[a]) + \delta N_0([\sigma] + \|\sigma\|[a]),$$
$$[\delta u] \lesssim \delta M + N_0(\|\delta\sigma\| + \|\sigma\delta a\|) + \delta N_0\|\sigma\|.$$

Making use of the constraints (71) on σ and a, in particular to absorb $[a]\delta M$ into the lhs, this simplifies to

$$\delta M \lesssim ([\delta a] + \|\delta a\|)M + N(N_0\delta N + \delta N_0) + \delta NN_0$$
$$\quad + N_0([\delta\sigma] + \|\delta\sigma\| + [\delta a] + \|\delta a\|) + \delta N_0([\sigma] + [a]),$$
$$[\delta u] \lesssim \delta M + N_0(\|\delta\sigma\| + \|\delta a\|) + \delta N_0.$$

Making use of (66) and (75), this reduces to

$$(234) \qquad \delta M \lesssim M\delta N + N(N_0\delta N + \delta N_0) + N_0\delta N,$$

$$(235) \qquad [\delta u] \lesssim \delta M + N_0\delta N + \delta N_0.$$

Making use of the estimate (73) on M we just established, (234) implies

$$\delta M \lesssim N(N_0\delta N + \delta N_0) + N_0\delta N.$$

Clearly, this estimate implies the desired (83). Plugging (83) into (235) yields the desired (84).

PROOF OF LEMMA 5.

All functions are 1-period if not stated otherwise.

STEP 1. Estimate of v_i and $\frac{\partial v_i}{\partial a_0}$. We claim

$$(236) \qquad \sup_{a_0}[\{v_i, \frac{\partial v_i}{\partial a_0}\}(\cdot, a_0)]_\alpha \lesssim N_i.$$

This follows immediately from assumption (61) on f_i and the definition (62) of v_i via Lemma 9 and the argument of Step 1 of Corollary 2.

STEP 2. Freezing-in the coefficients. We claim that we have for all points x_0

$$(237) \qquad (\partial_2 - a(x_0)\partial_1^2)\big(u_T - \sigma_i(x_0)v_{iT}(\cdot, a(x_0))\big) = Pg_{x_0}^T,$$

where the function $g_{x_0}^T$ is estimated as follows

$$(238) \qquad |g_{x_0}^T(x)| \lesssim \tilde{N}^2\big((T^{\frac{1}{4}})^{2\alpha-2} + (T^{\frac{1}{4}})^{\alpha-2}d^\alpha(x, x_0)\big) \quad \text{for } T \leq 1$$

with the abbreviation

$$(239) \qquad \tilde{N}^2 := N^2 + [a]_\alpha[u]_\alpha + N_i([\sigma_i]_\alpha + \|\sigma_i\|[a]_\alpha).$$

Indeed, making use of $P^2 = P$ we write

$$(240) \qquad (\partial_2 - a(x_0)\partial_1^2)u_T = P(\sigma_i(x_0)f_{iT}(\cdot, a(x_0)) + g_{x_0}^T)$$

with $g_{x_0}^T$ defined through

$$(241) \qquad \begin{aligned} g_{x_0}^T := {} & \partial_2 u_T - P(a\partial_1^2 u_T + \sigma_i Ef_{iT}) + (a - a(x_0))\partial_1^2 u_T \\ & + (\sigma_i - \sigma_i(x_0))Ef_{iT} + \sigma_i(x_0)(Ef_{iT} - f_{iT}(\cdot, a(x_0))). \end{aligned}$$

By definition (62) of $v_i(\cdot, a_0)$, to which we apply $(\cdot)_T$, which we evaluate for $a_0 = a(x_0)$, and which we contract with $\sigma_i(x_0)$ we have

$$(242) \qquad (\partial_2 - a(x_0)\partial_1^2)\sigma_i(x_0)v_{iT}(\cdot, a(x_0)) = P\sigma_i(x_0)f_{iT}(\cdot, a(x_0)).$$

From the combination of (240) and (242) we obtain (237), so that it remains to estimate $g_{x_0}^T$, cf (241). Making use of the assumption (63) we obtain

$$\begin{aligned} |g_{x_0}^T(x)| \leq {} & N^2(T^{\frac{1}{4}})^{2\alpha-2} + d^\alpha(x, x_0)\big([a]_\alpha\|\partial_1^2 u_T\| \\ & + [\sigma_i]_\alpha \sup_{a_0}\|f_{iT}\| + \|\sigma_i\|[a]_\alpha \sup_{a_0}\|(\frac{\partial f_i}{\partial a_0})_T\|\big), \end{aligned}$$

so that by (18) and by assumption (61)

$$\begin{aligned} |g_{x_0}^T(x)| \lesssim {} & N^2(T^{\frac{1}{4}})^{2\alpha-2} \\ & + (T^{\frac{1}{4}})^{\alpha-2}d^\alpha(x, x_0)\big([a]_\alpha[u]_\alpha + N_i([\sigma_i]_\alpha + \|\sigma_i\|[a]_\alpha)\big) \end{aligned}$$

which can be consolidated into the estimate (238).

STEP 3. PDE estimate. Under the outcome of Step 2, we have for all points x_0 and radii $R \ll L$

$$\frac{1}{R^{2\alpha}} \inf_{\ell} \|u_T - \sigma_i(x_0)v_{iT}(\cdot, a(x_0)) - \ell\|_{B_R(x_0)}$$

$$\lesssim (\frac{R}{L})^{2(1-\alpha)} \frac{1}{L^{2\alpha}} \inf_{\ell} \|u_T - \sigma_i(x_0)v_{iT}(\cdot, a(x_0)) - \ell\|_{B_L(x_0)}$$

$$(243) \qquad + \tilde{N}^2 \big(\frac{L^2}{R^{2\alpha}(T^{\frac{1}{4}})^{2-2\alpha}} + \frac{L^{2+\alpha}}{R^{2\alpha}(T^{\frac{1}{4}})^{2-\alpha}} \big),$$

where ℓ runs over all functions spanned by 1 and x_1 and $\| \cdot \|_{B_R(x_0)}$ denotes the supremum norm restricted to the ball $B_R(x_0)$ in the intrinsic metric (15) with center x_0 and radius R. This step mimics the heart of the kernel-free approach of Krylov & Safanov to the classical Schauder theory, see [12, Theorem 8.6.1]. Here comes the argument: Wlog we restrict to $x_0 = 0$ and write $B_R = B_R(0)$ and $\| \cdot \|_R := \| \cdot \|_{B_R}$. Let $w_>$ be the (non-periodic) solution of

$$(\partial_2 - a(0)\partial_1^2)w_> = I(B_L)g_0^T,$$

so that in view of (237), where we write $Pg_0^T = g_0^T + c$ with $c = -\int_{[0,1)^2} g_0^T$, the function

$$(244) \qquad w_< := u_T - \sigma_i(0)v_{iT}(\cdot, a(0)) - w_>$$

satisfies

$$(245) \qquad (\partial_2 - a(0)\partial_1^2)w_< = c \quad \text{in } B_L$$

with the constant c given by $c := -\int_{[0,1)^2} g_0^T$. By standard estimates for the heat equation we have

$$(246) \qquad \|w_>\| \lesssim L^2 \|g_0^T\|_L,$$

$$(247) \qquad \|\{\partial_1^2, \partial_2\}w_<\|_{\frac{L}{2}} \lesssim L^{-2}\|w_< - \ell_L\|_L$$

for any function $\ell_L \in \text{span}\{1, x_1\}$. The interior estimate (247) is slightly non-standard because of the non-vanishing rhs c but can be easily reduced to the case of $c = 0$: First of all, replacing w by $w - \ell_L$ in (245) and (247) we may reduce to the case of $\ell_L = 0$. Testing (245) with a cut-off function for B_L that is smooth on scale L we learn that $|c| \lesssim L^{-2}\|w_<\|_L$. We then may replace w by $w + cx_2$ which reduces the further estimate to the standard case of $c = 0$. We refer to [12, Theorem 8.4.4] for an elementary argument for (247) in case of $c = 0$ only relying on the maximum's principle via Bernstein's argument. We refer to [12, Exercise 8.4.8] for the statement (246) via the representation through the heat kernel. Since by construction, cf (244), we have $u_T - \sigma_i(0)v_{iT}(\cdot, a(0)) = w_< + w_>$ we obtain by the triangle inequality

for a suitably chosen $\ell_R \in \text{span}\{1, x_1\}$

$$\|u_T - \sigma_i(0)v_{iT}(\cdot, a(0)) - \ell_R\|_R$$
$$\leq \|w_< - \ell_R\|_R + \|w_>\|_R \lesssim R^2\|\{\partial_1^2, \partial_2\}w_<\|_R + \|w_>\|_R.$$

Inserting (247) for $R \ll L$, and another application of the triangle inequality this yield

$$\|u_T - \sigma_i(0)v_{iT}(\cdot, a(0)) - \ell_R\|_R$$
$$\lesssim L^{-2}R^2\|w_< - \ell_L\|_L + \|w_>\|_R$$
$$\leq L^{-2}R^2\|u_T - \sigma_i(0)v_{iT}(\cdot, a(0)) - \ell_L\|_L + 2\|w_>\|.$$

Inserting (246) & (238) this yields

$$\inf_\ell \|u_T - \sigma_i(0)v_{iT}(\cdot, a(0)) - \ell\|_R$$
$$\lesssim L^{-2}R^2 \inf_\ell \|u_T - \sigma_i(0)v_{iT}(\cdot, a(0)) - \ell\|_L$$
$$(248) \qquad + \tilde{N}^2L^2((T^{\frac{1}{4}})^{2\alpha-2} + L^\alpha(T^{\frac{1}{4}})^{\alpha-2}),$$

where we recall that ℓ runs over $\text{span}\{1, x_1\}$. Dividing by $R^{2\alpha}$ gives (243).

STEP 4. Equivalence of norms. We claim that the modelling constant M of u is estimated by the expression appearing in Step 3:

$$(249) \qquad\qquad M \lesssim M',$$

where we have set for abbreviation

$$(250) \qquad M' := \sup_{x_0} \sup_{R \leq 4} R^{-2\alpha} \inf_\ell \|u - \sigma_i(x_0)v_i(\cdot, a(x_0)) - \ell\|_{B_R(x_0)}$$

and where the maximal radius 4 is chosen such that a ball of half of that radius covers the periodic cell $[0, 1]^2$. In fact, also the reverse estimate holds, highlighting once more that the modulation function ν in the definition of modelledness (Definition 1) plays a small role compared to σ_i. The equivalence of (249) and (250) on the level of standard Hölder spaces is the starting point for the approach to Schauder theory by Krylov and Safanov, see [12, Theorem 8.5.2]. We first argue that the ℓ in (250) may be chosen to be independent of R, that is,

$$(251) \qquad \sup_{x_0} \inf_\ell \sup_{R \leq 4} R^{-2\alpha}\|u - \sigma_i(x_0)v_i(\cdot, a(x_0)) - \ell\|_{B_R(x_0)} \lesssim M'.$$

Indeed, fix x_0, say $x_0 = 0$, and let $\ell_R = \nu_R x_1 + c_R$ be (near) optimal in (250), then we have by definition of M' and by the triangle inequality $R^{-2\alpha}\|\ell_{2R} - \ell_R\|_R \lesssim M'$. This implies $R^{1-2\alpha}|\nu_{2R} - \nu_R| + R^{-2\alpha}|c_{2R} - c_R| \lesssim M'$, which thanks to $\alpha > \frac{1}{2}$ yields by telescoping $R^{1-2\alpha}|\nu_R - \nu_{R'}| + R^{-2\alpha}|c_R - c_{R'}| \lesssim M'$ for all $R' \leq R$ and thus the existence of $\nu, c \in \mathbb{R}$ such that $R^{1-2\alpha}|\nu_R - \nu| + R^{-2\alpha}|c_R - c| \lesssim M'$, so that $\ell := \nu x_1 + c$ satisfies

$$(252) \qquad\qquad R^{-2\alpha}\|\ell_R - \ell\|_R \lesssim M'.$$

Hence we may pass from (250) to (251) by the triangle inequality.

It is clear from (251) that necessarily for any x_0, say $x_0 = 0$, the optimal ℓ must be of the form $\ell(x) = u(0) - \sigma_i(0)v_i(0, a(0)) - \nu(0)x_1$. This establishes the main part of (249), namely the modelledness (21) for any "base" point x and any y of distance at most 4. Since $B_4(x)$ covers a periodic cell, we may use (21) for $y = x + (1,0)$ so that by periodicity of $y \mapsto (u(y) - u(x)) - \sigma_i(x)(v_i(y, a(x)) - v_i(x, a(x)))$ we extract $|\nu(x)| \lesssim M'$. Since $\alpha \geq \frac{1}{2}$, this implies that $|\nu(x)(x - y)_1| \lesssim M'd^{2\alpha}(x, y)$ for all $y \notin B_4(x)$. Hence once again by periodicity of $y \mapsto (u(y) - u(x)) - \sigma_i(x)(v_i(y, a(x)) - v_i(x, a(x)))$, (21) holds also for $y \notin B_4(x)$.

STEP 5. Modelledness implies approximation property. We claim that for any mollification parameter $0 < T \leq 1$, radius L, and point x_0 we have

$$\frac{1}{(T^{\frac{1}{4}})^{2\alpha}}\|(u_T - u) - \sigma_i(x_0)(v_{iT} - v_i)(\cdot, a(x_0))\|_{B_L(x_0)}$$

(253) $$\lesssim M + \tilde{N}^2(\frac{L}{T^{\frac{1}{4}}})^\alpha,$$

where we recall the definition (239) of $\tilde{N}$. Wlog we restrict ourselves to $x_0 = 0$, write $v_i(y, x) = v_i(y, a(x))$, and note that the first moment of ψ_T vanishes

$$(u_T - u)(x) - \sigma_i(0)(v_{iT} - v_i)(x, 0)$$
$$= \int dy\psi_T(x - y)\big((u(y) - u(x)) - \sigma_i(0)(v_i(y, 0) - v_i(x, 0))$$
$$- \nu(x)(y - x)_1\big).$$

We split the rhs into three terms:

$$(u_T - u)(x) - \sigma_i(0)(v_{iT} - v_i)(x, 0)$$
$$= \int dy\psi_T(x - y)\big((u(x) - u(y)) - \sigma_i(x)(v_i(y, x) - v_i(x, x))$$
$$- \nu(x)(x - y)_1\big)$$
$$+ \int dy\psi_T(x - y)(\sigma_i(x) - \sigma_i(0))(v_i(y, 0) - v_i(x, 0))$$
$$+ \int dy\psi_T(x - y)\sigma_i(x)\big((v_i(y, x) - v_i(y, 0)) - (v_i(x, x) - v_i(x, 0))\big).$$

For the first rhs term we appeal to the modelledness assumption (21), which implies that the integrand is estimated by $|\psi_T(x-y)|\, M\, d^{2\alpha}(x, y)$. Hence by (18) the integral is estimated by $M\,(T^{\frac{1}{4}})^{2\alpha}$. The integrand of the second rhs term is estimated by $|\psi_T(x - y)|\,[\sigma_i]_\alpha\,d^\alpha(x, 0)\,[v_i(\cdot, 0)]_\alpha$ $d^\alpha(x, y)$ so that by (18) and (236) the integral is controlled by $\lesssim [\sigma_i]_\alpha$

$d^\alpha(x,0)\ N_i\ (T^{\frac{1}{4}})^\alpha$; since $x \in B_L(0)$ it is controlled by $\lesssim [\sigma_i]_\alpha\ L^\alpha\ N_i$ $(T^{\frac{1}{4}})^\alpha$. Using the identity (and dropping the index i)

$$(v(y, a(x)) - v(y, a(0))) - (v(x, a(x)) - v(x, a(0))) = (a(x) - a(0))$$

$$\times \int_0^1 ds\big(\frac{\partial v}{\partial a_0}(y, sa(x) + (1-s)a(0)) - \frac{\partial v}{\partial a_0}(x, sa(x) + (1-s)a(0)))\big),$$

we see that the integrand of the third rhs term is estimated by $\|\sigma_i\|$ $d^\alpha(x,y)\ [a]_\alpha\ d^\alpha(x,0)\ \sup_{a_0}[\frac{\partial v_i}{\partial a_0}(\cdot, a_0)]_\alpha$; hence in view of (236) the third term itself is estimated by $\|\sigma_i\|\ N_i\ (T^{\frac{1}{4}})^\alpha\ [a]_\alpha\ L^\alpha$. Collecting these estimates we obtain for $x \in B_L(0)$

$$|(u_T - u)(x) - \sigma_i(0)(v_{iT} - v_i)(x, 0)|$$

$$\lesssim M(T^{\frac{1}{4}})^{2\alpha} + N_i([\sigma_i]_\alpha + \|\sigma_i\|[a]_\alpha)L^\alpha(T^{\frac{1}{4}})^\alpha.$$

In view of the definition (239) of $\tilde{N}^2$, this yields (253).

STEP 6. Estimate of M. We claim that

$$(254) \qquad\qquad M \lesssim \tilde{N}^2.$$

Indeed, we now may buckle and to this purpose rewrite (243) from Step 3 with help of the triangle inequality as

$$\frac{1}{R^{2\alpha}} \inf_\ell \|u - \sigma_i(x_0)v_i(\cdot, a(x_0)) - \ell\|_{B_R(x_0)}$$

$$\lesssim (\frac{R}{L})^{2-2\alpha}\frac{1}{L^{2\alpha}} \inf_\ell \|u - \sigma_i(x_0)v_i(\cdot, a(x_0)) - \ell\|_{B_L(x_0)}$$

$$+ \tilde{N}^2\big(\frac{L^2}{R^{2\alpha}(T^{\frac{1}{4}})^{2-2\alpha}} + \frac{L^{2+\alpha}}{R^{2\alpha}(T^{\frac{1}{4}})^{2-\alpha}}\big)$$

$$+ (\frac{T^{\frac{1}{4}}}{R})^{2\alpha}\frac{1}{(T^{\frac{1}{4}})^{2\alpha}}\|(u_T - u) - \sigma_i(x_0)(v_{iT} - v_i)(\cdot, a(x_0))\|_{B_L(x_0)}.$$

We now insert (253) from Step 5 to obtain

$$\frac{1}{R^{2\alpha}} \inf_\ell \|u - \sigma_i(x_0)v_i(\cdot, a(x_0)) - \ell\|_{B_R(x_0)}$$

$$\lesssim (\frac{R}{L})^{2-2\alpha}M + \tilde{N}^2\big(\frac{L^2}{R^{2\alpha}(T^{\frac{1}{4}})^{2-2\alpha}} + \frac{L^{2+\alpha}}{R^{2\alpha}(T^{\frac{1}{4}})^{2-\alpha}}\big)$$

$$(255) \qquad + (\frac{T^{\frac{1}{4}}}{R})^{2\alpha}M + \tilde{N}^2\frac{L^\alpha(T^{\frac{1}{4}})^\alpha}{R^{2\alpha}}.$$

Here we have used that

$$\sup_{x_0} \sup_L \frac{1}{L^{2\alpha}} \inf_\ell \|u - \sigma_i(x_0)v_i(\cdot, a(x_0)) - \ell\|_{B_L(x_0)} \lesssim M$$

by the definition of the modelling constant M with $\ell_{x_0}(x) = u(x_0)$ $-\sigma_i(x_0)v_i(x_0, a(x_0))\ -\nu(x_0)(x - x_0)_1$. Relating the length scales $T^{\frac{1}{4}}$ and L to the given $R \leq 4$ in (255) via $T^{\frac{1}{4}} = \epsilon R$ (so that in particular

as required $T \leq 1$ since we think of $\epsilon \ll 1$) and $L = \epsilon^{-1}R$, taking the supremum over $R \leq 4$ and x_0 yields by definition (250) of M'

$$M' \lesssim (\epsilon^{2-2\alpha} + \epsilon^{2\alpha})M + (\epsilon^{2\alpha-4} + \epsilon^{-4} + 1)\tilde{N}^2.$$

By (249) in Step 4, this implies

$$M \lesssim (\epsilon^{2-2\alpha} + \epsilon^{2\alpha})M + \epsilon^{-4}\tilde{N}^2.$$

Since $0 < \alpha < 1$, we may choose ϵ sufficiently small such that the first rhs term may be absorbed into the lhs yielding the desired estimate $M \lesssim \tilde{N}^2$ (note that $M < \infty$ is part of our assumption).

STEP 7. Conclusion. Clearly, (64) and (65) immediately follow from the combination of

$$M \lesssim N^2 + [a]_\alpha[u]_\alpha + N_i([\sigma_i]_\alpha + \|\sigma_i\|[a]_\alpha),$$
$$[u]_\alpha \lesssim M + N_i\|\sigma_i\|.$$

The first estimate is identical to (254) in Step 6 into which we plug the definition (239) of $\tilde{N}$. The second estimate is an application of Step 2 in the proof of Lemma 2 with $v(y,x) := \sigma_i(x)v_i(y, a_i(x))$, so that the hypothesis (33) holds with N replaced by $\|\sigma_i\|N_i$, cf (236) in Step 1.

PROOF OF LEMMA 2.
We write for abbreviation $[\cdot] := [\cdot]_\alpha$.
STEP 1. We claim

(256) $[\nu]_{2\alpha-1} \lesssim M + N.$

Indeed, introducing $\ell_x(y) := \nu(x)y_1$ we see that (31) can be rewritten as

$$|(u - v(\cdot, x) - \ell_x)(y) - (u - v(\cdot, x) - \ell_x)(x)| \leq Md^{2\alpha}(y,x),$$

so that we obtain by the triangle inequality

$$|(u - v(\cdot, x) - \ell_x)(y) - (u - v(\cdot, x) - \ell_x)(y')|$$
(257) $$\leq M(d^{2\alpha}(y,x) + d^{2\alpha}(y',x)).$$

In combination with (28) this yields by the triangle inequality

$$|(u - v(\cdot, x') - \ell_x)(y) - (u - v(\cdot, x') - \ell_x)(y')|$$
$$\leq M(d^{2\alpha}(y,x) + d^{2\alpha}(y',x)) + Nd^\alpha(x,x')d^\alpha(y,y').$$

We now take the difference of this with (257) with x replaced by x' to obtain, once more by the triangle inequality,

$$|(\ell_x - \ell_{x'})(y) - (\ell_x - \ell_{x'})(y')| \leq M\big(d^{2\alpha}(y,x) + d^{2\alpha}(y',x)$$
$$+ d^{2\alpha}(y,x') + d^{2\alpha}(y',x')\big) + Nd^\alpha(x,x')d^\alpha(y,y').$$

By definition of ℓ and with the choice of $y = x$ and $y' = x + (R, 0)$, this assumes the form

$$|\nu(x) - \nu(x')|R$$
$$\leq M(R^{2\alpha} + d^{2\alpha}(x, x') + (R + d(x, x'))^{2\alpha}) + N d^{\alpha}(x, x') R^{\alpha}.$$

With the choice of $R = d(x, x')$ this turns into

$$|\nu(x) - \nu(x')|d(x, x') \lesssim (M + N)d^{2\alpha}(x, x'),$$

which amounts to the desired (256).

STEP 2. We claim

$$(258) \qquad\qquad [u] + \|\nu\| \lesssim M + N.$$

By the triangle inequality on (31) we obtain for all pairs of points $|\nu(x)(x - y)_1| \leq |u(x) - u(y)| + [v(\cdot, x)]d^{\alpha}(y, x) + M d^{2\alpha}(x, y)$. Choosing $y = x + (1, 0)$, appealing to the 1-periodicity of u, taking the supremum over x, and appealing to (33), this turns into the ν-part of (258):

$$(259) \qquad\qquad \|\nu\| \lesssim M + N.$$

We now consider pairs of points (x, y) with $d(x, y) \leq 2$. By the triangle inequality from (31) we get

$$\frac{1}{d^{\alpha}(x, y)}|u(x) - u(y)| \lesssim M + N + \|\nu\|.$$

By 1-periodicity, this extends to all pairs so that

$$[u] \lesssim M + N + \|\nu\|.$$

Inserting (259) into this yields the u-part of (258).

STEP 3. Dyadic decomposition. For $\tau < T$ (with T a dyadic multiple of τ) we claim that

$$\left(uf_T - E[v, (\cdot)_T] \diamond f - \nu[x_1, (\cdot)_T]f\right)$$
$$- \left(uf_\tau - E[v, (\cdot)_\tau] \diamond f - \nu[x_1, (\cdot)_\tau]f\right)_{T - \tau}$$
$$= \sum_{\tau \leq t < T} \left(\left([u, (\cdot)_t] - E[v, (\cdot)_t] - \nu[x_1, (\cdot)_t]\right)f_t\right.$$
$$(260) \qquad \left. - [\nu, (\cdot)_t][x_1, (\cdot)_t]f - [E, (\cdot)_t][v, (\cdot)_t] \diamond f\right)_{T - 2t},$$

where the sum runs over the dyadic "times" $t = \frac{T}{2}, \frac{T}{4}, \cdots, \tau$. By telescoping based on the semi-group property (17) this reduces to

$$\left(uf_{2t} - E[v, (\cdot)_{2t}] \diamond f - \nu[x_1, (\cdot)_{2t}]f\right)$$
$$- \left(uf_t - E[v, (\cdot)_t] \diamond f - \nu[x_1, (\cdot)_t]f\right)_t$$
$$= \left([u, (\cdot)_t] - E[v, (\cdot)_t] - \nu[x_1, (\cdot)_t]\right)f_t$$
$$- [\nu, (\cdot)_t][x_1, (\cdot)_t]f - [E, (\cdot)_t][v, (\cdot)_t] \diamond f,$$

and splits into the three statements

$$(261) \qquad u f_{2t} - (u f_t)_t = [u, (\cdot)_t] f_t,$$

$$\nu[x_1, (\cdot)_{2t}] f - (\nu[x_1, (\cdot)_t] f)_t = \nu[x_1, (\cdot)_t] f_t + [\nu, (\cdot)_t][x_1, (\cdot)_t] f,$$

$$E[v, (\cdot)_{2t}] \diamond f - (E[v, (\cdot)_t] \diamond f)_t = E[v, (\cdot)_t] f_t + [E, (\cdot)_t][v, (\cdot)_t] \diamond f.$$

Plugging in the definition of the commutator $[\nu, (\cdot)_t]$, the middle statement reduces to

$$(262) \qquad [x_1, (\cdot)_{2t}] f - ([x_1, (\cdot)_t] f)_t = [x_1, (\cdot)_t] f_t.$$

By the definition of the commutator $[E, (\cdot)_t]$, the last statement reduces to

$$(263) \qquad [v, (\cdot)_{2t}] \diamond f - ([v, (\cdot)_t] \diamond f)_t = [v, (\cdot)_t] f_t,$$

which by definition of $[v, (\cdot)_T] \diamond f$ splits into

$$(264) \qquad v f_{2t} - (v f_t)_t = [v, (\cdot)_t] f_t \quad \text{and} \quad (v \diamond f)_{2t} - ((v \diamond f)_t)_t = 0.$$

Now identities (261), (262), and (264) follow immediately from the semi-group property.

STEP 4. For $\tau < T \le 1$ (with T a dyadic multiple of τ) we claim the estimate

$$\big\| \big(u f_T - E[v, (\cdot)_T] \diamond f - \nu[x_1, (\cdot)_T] f \big)$$
$$- \big(u f_\tau - E[v, (\cdot)_\tau] \diamond f - \nu[x_1, (\cdot)_\tau] f \big)_{T-\tau} \big\|$$
$$(265) \qquad \lesssim (M + N) N_1 (T^{\frac{1}{4}})^{3\alpha - 2}.$$

Indeed, by the dyadic representation (260), the triangle inequality in $\| \cdot \|$ and the fact that $(\cdot)_{T-2t}$ is bounded in that norm, cf (18), it is enough to show that the rhs term of (260) under the parenthesis is estimated by $(M + N) N_1 (t^{\frac{1}{4}})^{3\alpha - 2}$ for all $t \le 1$; here we crucially use that by assumption $3\alpha - 2 > 0$ for the convergence of the geometric series. Using Step 1 to control $[\nu]_{2\alpha - 1}$ in (266) by $M + N$, this estimate splits into

$$\big\| \big([u, (\cdot)_t] - E[v, (\cdot)_t] - \nu[x_1, (\cdot)_t] \big) f_t \big\| \lesssim M N_1 (t^{\frac{1}{4}})^{3\alpha - 2},$$

$$(266) \qquad \big\| [\nu, (\cdot)_t][x_1, (\cdot)_t] f \big\| \lesssim [\nu]_{2\alpha - 1} N (t^{\frac{1}{4}})^{3\alpha - 2},$$

$$(267) \qquad \big\| [E, (\cdot)_t][v, (\cdot)_t] \diamond f \big\| \lesssim N N_1 (t^{\frac{1}{4}})^{3\alpha - 2}.$$

Appealing to our assumptions (29) & (30) and to Lemma 10, these three estimates reduce to

$$\big\| \big([u, (\cdot)_t] - E[v, (\cdot)_t] - \nu[x_1, (\cdot)_t] \big) \tilde{f} \big\| \lesssim M \| \tilde{f} \| (t^{\frac{1}{4}})^{2\alpha},$$

$$(268) \qquad \big\| [\nu, (\cdot)_t] \tilde{f} \big\| \lesssim [\nu]_\beta \| \tilde{f} \| (t^{\frac{1}{4}})^{\beta},$$

$$(269) \qquad \big\| [E, (\cdot)_t] \tilde{v} \big\| \lesssim \sup_{x,x'} \frac{1}{d^\alpha(x, x')} \| \tilde{v}(\cdot, x) - \tilde{v}(\cdot, x') \| (t^{\frac{1}{4}})^\alpha,$$

where $\tilde{f} = \tilde{f}(y)$ plays the role of f_t or $[x_1, (\cdot)_t]f$, and $\tilde{v} = \tilde{v}(x, y)$ plays the role of $([v(\cdot, x), (\cdot)_t] \diamond f)(y)$, but now can be, like ν, generic functions; similarly, β plays the role of $2\alpha - 1$ but could be any exponent in $[0, 1]$. Using the definition of E, we may rewrite these estimates more explicitly as

$$| \int dy\, \psi_t(x - y)\Big(\big(u(x) - u(y)\big) - \big(v(x, x) - v(y, x)\big)$$
$$- \nu(x)(x - y)_1\Big)\tilde{f}(y)| \lesssim M\|\tilde{f}\|(t^{\frac{1}{4}})^{2\alpha},$$

$$| \int dy\, \psi_t(x - y)\big(\nu(x) - \nu(y)\big)\tilde{f}(y)| \lesssim [\nu]_\beta\|\tilde{f}\|(t^{\frac{1}{4}})^{\beta},$$

$$| \int dy\, \psi_t(x - y)\big(\tilde{v}(y, x) - \tilde{v}(y, y)\big)|$$
$$\lesssim \sup_{x,x'} \frac{1}{d^\alpha(x, x')}\|\tilde{v}(\cdot, x) - \tilde{v}(\cdot, x')\|(t^{\frac{1}{4}})^{\alpha}.$$

All three estimates rely on the moment bounds (18), the first estimate is then an immediate consequence of (31) and the two last ones tautological.

STEP 5. For

$$F^\tau := uf_\tau - E[v, (\cdot)_\tau] \diamond f - \nu[x_1, (\cdot)_\tau]f$$

we claim the estimates

(270) $$\sup_{T \leq 1}(T^{\frac{1}{4}})^{2-2\alpha}\|uf_T - F^\tau_{T-\tau}\| \lesssim (M + N)N_1,$$

(271) $$\sup_{T \leq 1}(T^{\frac{1}{4}})^{2-\alpha}\|F^\tau_T\| \lesssim (M + N + \|u\|)N_1.$$

Indeed, (270) follows from (265) in Step 4 via the triangle inequality and

$$\sup_{T \leq 1}(T^{\frac{1}{4}})^{2-2\alpha}\|E[v, (\cdot)_T] \diamond f\| \overset{(34)}{\leq} NN_1,$$

$$\sup_{T \leq 1}(T^{\frac{1}{4}})^{2-2\alpha}\|\nu[x_1, (\cdot)_T]f\| \overset{(258),(342)}{\leq} (M + N)N_1,$$

the latter being a consequence of Step 2 and Lemma 10; here, we make extensively use of $T \leq 1$. Estimate (271) in turn follows from (270) via $\|F^\tau_T\| = \|(F^\tau_{T-\tau})_\tau\| \lesssim \|F^\tau_{T-\tau}\|$ (cf (17) and (18)) by the triangle inequality and (29), again making use of $T \leq 1$.

STEP 6. Conclusion. Indeed by (271) in Step 5, the sequence $\{F^\tau\}_{\tau \downarrow 0}$ of functions is bounded as distributions in $C^{\alpha-2}$. By standard weak compactness based on the equivalence of norms from Step 1 in the proof of Lemma 9, there exists a subsequence $\tau_n \downarrow 0$ and a distribution we give the name of $u \diamond f$ such that $F^{\tau_n} \rightharpoonup u \diamond f$. By standard lower semi-continuity, we may pass to the limit in (270) in Step 5 to obtain

(35). Likewise, we may pass to the limit in (265) in Step 4 to obtain (32).

PROOF OF LEMMA 4.

The proof follows the lines of Steps 3 through 6 of the proof of Lemma 2.

STEP 1. For $\tau < T$ (with T a dyadic multiple of τ) we claim the formula

$$\left(b\partial_1^2 u_T - \sigma_i E[b,(\cdot)_T]\diamond\partial_1^2 v_i\right) - \left(b\partial_1^2 u_\tau - \sigma_i E[b,(\cdot)_\tau]\diamond\partial_1^2 v_i\right)_{T-\tau}$$

$$= \sum_{\tau\le t<T}\Big(\big([b,(\cdot)_t]\partial_1^2 u_t - \sigma_i E[b,(\cdot)_t]\partial_1^2 v_{it}\big)$$

$$(272)\qquad - [\sigma_i,(\cdot)_t]E[b,(\cdot)_t]\diamond\partial_1^2 v_i - \sigma_i[E,(\cdot)_t][b,(\cdot)_t]\diamond\partial_1^2 v_i\Big)_{T-2t},$$

where the sum runs over $t = \frac{T}{2},\frac{T}{4},\ldots,\tau$. By telescoping based on the semi-group property the formula reduces to

$$\left(b\partial_1^2 u_{2t} - \sigma_i E[b,(\cdot)_{2t}]\diamond\partial_1^2 v_i\right) - \left(b\partial_1^2 u_t - \sigma_i E[b,(\cdot)_t]\diamond\partial_1^2 v_i\right)_t$$

$$= \big([b,(\cdot)_t]\partial_1^2 u_t - \sigma_i E[b,(\cdot)_t]\partial_1^2 v_{it}\big)$$

$$+ [\sigma_i,(\cdot)_t]E[b,(\cdot)_t]\diamond\partial_1^2 v_i + \sigma_i[E,(\cdot)_t][b,(\cdot)_t]\diamond\partial_1^2 v_i,$$

and splits into the two statements

$$(273)\qquad b\partial_1^2 u_{2t} - (b\partial_1^2 u_t)_t = [b,(\cdot)_t]\partial_1^2 u_t,$$

$$\sigma_i E[b,(\cdot)_{2t}]\diamond\partial_1^2 v_i - (\sigma_i E[b,(\cdot)_t]\diamond\partial_1^2 v_i)_t = \sigma_i E[b,(\cdot)_t]\partial_1^2 v_{it}$$

$$+ [\sigma_i,(\cdot)_t]E[b,(\cdot)_t]\diamond\partial_1^2 v_i + \sigma_i[E,(\cdot)_t][b,(\cdot)_t]\diamond\partial_1^2 v_i.$$

By definition of the commutator $[\sigma_i,(\cdot)_t]$, the last statement reduces to

$$E[b,(\cdot)_{2t}]\diamond\partial_1^2 v_i - (E[b,(\cdot)_t]\diamond\partial_1^2 v_i)_t$$

$$= E[b,(\cdot)_t]\partial_1^2 v_{it} + [E,(\cdot)_t][b,(\cdot)_t]\diamond\partial_1^2 v_i,$$

and by the definition of $[E,(\cdot)_t]$ further to

$$(274)\qquad [b,(\cdot)_{2t}]\diamond\partial_1^2 v_i - ([b,(\cdot)_t]\diamond\partial_1^2 v_i)_t = [b,(\cdot)_t]\partial_1^2 v_{it}.$$

Now (273) and (274) are consequences of the semi group property.

STEP 2. We claim the estimate

$$\|\left(b\partial_1^2 u_T - \sigma_i E[b,(\cdot)_T]\diamond\partial_1^2 v_i\right) - \left(b\partial_1^2 u_\tau - \sigma_i E[b,(\cdot)_\tau]\diamond\partial_1^2 v_i\right)_{T-\tau}\|$$

$$\lesssim \left([b]_\alpha M + N_0 N_i([\sigma_i]_\alpha + \|\sigma_i\|[a]_\alpha)\right)(T^{\frac14})^{3\alpha-2}.$$

In view of (272) this estimate splits into

$$(275)\qquad \|[b,(\cdot)_t]\partial_1^2 u_t - \sigma_i E[b,(\cdot)_t]\partial_1^2 v_{it}\| \lesssim [b]_\alpha M(t^{\frac14})^{3\alpha-2},$$

$$(276)\qquad \|[\sigma_i,(\cdot)_t]E[b,(\cdot)_t]\diamond\partial_1^2 v_i\| \lesssim N_0 N_i[\sigma_i]_\alpha(t^{\frac14})^{3\alpha-2},$$

$$(277)\qquad \|[E,(\cdot)_t][b,(\cdot)_t]\diamond\partial_1^2 v_i\| \lesssim N_0 N_i[a]_\alpha(t^{\frac14})^{3\alpha-2}.$$

Estimate (276) follows from (268) (with σ_i playing the role of ν, $E[b, (\cdot)_T] \diamond \partial_1^2 v_i$ playing the role of $\tilde{f}$, and β playing the role of α) and our assumption (58) (without $\frac{\partial}{\partial a_0}$). Estimate (277) from (269) (with $[b, (\cdot)_t] \diamond \partial_1^2 v_i$ playing the role of $\tilde{v}$ and our assumptions (57) and (58) (with $\frac{\partial}{\partial a_0}$),

$$\frac{1}{d^\alpha(x, x')} \| ([b, (\cdot)_t] \diamond \partial_1^2 v_i)(\cdot, a(x)) - ([b, (\cdot)_t] \diamond \partial_1^2 v_i)(\cdot, a(x')) \|$$

$$\leq [a]_\alpha \sup_{a_0} \| \frac{\partial}{\partial a_0} [b, (\cdot)_t] \diamond \partial_1^2 v_i \| \leq [a]_\alpha N_0 N_i (t^{\frac{1}{4}})^{2\alpha - 2}.$$

For (275) we write

$$\big([b, (\cdot)_t] \partial_1^2 u_t - \sigma_i E[b, (\cdot)_t] \partial_1^2 v_{it} \big)(x)$$

$$(278) \quad = \int dy \psi_t(x - y)(b(x) - b(y)) \big(\partial_1^2 u_t(y) - \sigma_i(x) \partial_1^2 v_{it}(y, a(x)) \big)$$

and

$$\partial_1^2 u_t(y) - \sigma_i(x) \partial_1^2 v_{it}(y, a(x)) = \int dz \partial_1^2 \psi_t(y - z) \times$$

$$\big(u(z) - u(x) - \sigma_i(x)(v_i(z, a(x)) - v_i(x, a(x))) - \nu(x)(z - x)_1 \big).$$

Hence by the modelledness assumption of u, the triangle inequality $d(z, x) \leq d(z, y) + d(y, x)$, and (18) we obtain

$$|\partial_1^2 u_t(y) - \sigma_i(x) \partial_1^2 v_{it}(y, a(x))| \lesssim M((t^{\frac{1}{4}})^{2\alpha - 2} + (t^{\frac{1}{4}})^{-2} d^{2\alpha}(y, x)).$$

Plugging this into (278), we obtain using (18) once more

$$\big| [b, (\cdot)_t] \partial_1^2 u_t - \sigma_i E[b, (\cdot)_t] \partial_1^2 v_{it} \big|(x) \lesssim [b]_\alpha M(t^{\frac{1}{4}})^{3\alpha - 2},$$

as desired.

The further two steps are as Steps 5 and 6 in Lemma 2.

PROOF OF LEMMA 3.
We write $[\cdot]$ for $[\cdot]_\alpha$.
STEP 1. Suppose that $\{v(\cdot, a_0)\}_{a_0}$ and $\{v_i(\cdot, a_0)\}_{a_0}$, $i = 0, 1$, are three families of functions and $|\cdot|$ a semi-norm on functions of x (like $[\cdot]$) such that

$$(279) \qquad \sup_{a_0} |\{1, \frac{\partial}{\partial a_0}\} v(\cdot, a_0)| \leq N_0,$$

$$(280) \qquad \sup_{a_0} |\{1, \frac{\partial}{\partial a_0}, \frac{\partial^2}{\partial a_0^2}\} v_i(\cdot, a_0)| \leq N_0,$$

$$(281) \qquad \sup_{a_0} |\{1, \frac{\partial}{\partial a_0}\} (v_1 - v_0)(\cdot, a_0)| \leq \delta N_0$$

for some constants $N_0, \delta N_0$; the reason for this more general framework is useful because we shall also apply it with $v(\cdot, a_0)$ replaced by $[v(\cdot, a_0), (\cdot)_T] \diamond f$, the supremums norm $\|\cdot\|$ playing the role of $|\cdot|$, and

with $(N_0, \delta N_0)$ replaced by $(N_1 N_0 (T^{\frac{1}{4}})^{2\alpha-2}, N_1 \delta N_0 (T^{\frac{1}{4}})^{2\alpha-2})$. We claim that this entails

$$(282) \quad |\sigma(x)v(\cdot, a(x))| \leq N_0 \|\sigma\|,$$

$$(283) \quad |\sigma(x)v(\cdot, a(x)) - \sigma(x')v(\cdot, a(x'))| \leq N_0([\sigma] + \|\sigma\|[a])d^\alpha(x, x'),$$

$$|\sigma_1(x)v_1(\cdot, a_1(x)) - \sigma_0(x)v_0(\cdot, a_0(x))|$$

$$(284) \quad \leq N_0(\|\sigma_1 - \sigma_0\| + \max_i \|\sigma_i\|\|a_1 - a_0\|) + \delta N_0 \max_i \|\sigma_i\|,$$

$$\big|(\sigma_1(x)v_1(\cdot, a_1(x)) - \sigma_0(x)v_0(\cdot, a_0(x)))$$

$$- (\sigma_1(x')v_1(\cdot, a_1(x')) - \sigma_0(x')v_0(\cdot, a_0(x')))\big|$$

$$\leq \Big(N_0 \max_{i,j} \big([\sigma_1 - \sigma_0] + \|\sigma_i\|[a_1 - a_0] + [\sigma_i]\|a_1 - a_0\|$$

$$+ \|\sigma_1 - \sigma_0\|[a_i] + \|\sigma_i\|[a_j]\|a_1 - a_0\|\big)$$

$$(285) \quad + \delta N_0 \max_i \big([\sigma_i] + \|\sigma_i\|[a_i]\big)\Big)d^\alpha(x, x').$$

Estimate (282) follows immediately from (279). We treat (283), (284), and (285) along the same lines, which is a bit of an overkill for (283) and (284). We start with the two elementary, and purposefully symmetric, formulas

$$(286) \quad \sigma v - \sigma'v' = \frac{1}{2}(\sigma - \sigma')(v + v') + \frac{1}{2}(\sigma + \sigma')(v - v')$$

and

$$(\sigma_1 v_1 - \sigma_0 v_0) - (\sigma_1' v_1' - \sigma_0' v_0')$$

$$= \frac{1}{4}((\sigma_1 - \sigma_0) - (\sigma_1' - \sigma_0'))(v_1 + v_1' + v_0 + v_0')$$

$$+ \frac{1}{4}((\sigma_1 + \sigma_1' + \sigma_0 + \sigma_0'))((v_1 - v_0) - (v_1' - v_0'))$$

$$+ \frac{1}{4}((\sigma_1 - \sigma_1') + (\sigma_0 - \sigma_0'))((v_1 - v_0) + (v_1' - v_0'))$$

$$(287) \quad + \frac{1}{4}((\sigma_1 - \sigma_0) + (\sigma_1' - \sigma_0'))((v_1 - v_1') + (v_0 - v_0')).$$

We use the first formula twice. The first application is for $\sigma = \sigma(x)$ and $\sigma' = \sigma(x')$, $v = v(\cdot, a(x))$, and $v' = v(\cdot, a(x'))$ to obtain using the triangle inequality

$$|\sigma(x)v(\cdot, a(x)) - \sigma(x')v'(\cdot, a(x'))|$$

$$\leq [\sigma]d^\alpha(x, x') \sup_{a_0} |v(\cdot, a_0)| + \|\sigma\| \sup_{a_0} |\frac{\partial v}{\partial a_0}(\cdot, a_0)|[a]d^\alpha(x, x').$$

In view of the assumption (279) this yields (283). The second application is for $\sigma = \sigma_1(x)$ and $\sigma' = \sigma_0(x)$, $v = v_1(\cdot, a_1(x))$, and

$v' = v_0(\cdot, a_0(x))$. We obtain the inequality

$$|\sigma_1(x)v_1(\cdot, a_1(x)) - \sigma_0(x)v_0(\cdot, a_0(x))|$$

$$\leq \|\sigma_1 - \sigma_0\| \max_i \sup_{a_0} |v_i(\cdot, a_0)|$$

(288)
$$+ \max_i \|\sigma_i\| |v_1(\cdot, a_1(x)) - v_0(\cdot, a_0(x))|.$$

In view of the assumption (280), the first rhs term is estimated as desired. For the second rhs term we interpolate linearly in the sense of $v_s := sv_1 + (1 - s)v_0$ and $a_s := sa_1 + (1 - s)a_0$, to the effect of

$$v_1(\cdot, a_1(x)) - v_0(\cdot, a_0(x))$$

(289)
$$= \int_0^1 ds \big((v_1 - v_0)(\cdot, a_s(x)) + \frac{\partial v_s}{\partial a_0}(\cdot, a_s(x))(a_1 - a_0)(x),$$

from which we learn

$$|v_1(\cdot, a_1(x)) - v_0(\cdot, a_0(x))|$$

(290)
$$\leq \sup_{a_0} |v_1 - v_0| + \max_i \sup_{a_0} |\frac{\partial v_i}{\partial a_0}| \|a_1 - a_0\|.$$

Inserting this into (288) and in view of the assumption (280)&(281) we obtain the remaining part of (284).

We use the second formula (287) for $\sigma_i = \sigma_i(x)$, $\sigma'_i = \sigma_i(x')$, $v_i = v_i(\cdot, a_i(x))$, and $v'_i = v_i(\cdot, a_i(x'))$ to obtain

$$|\big(\sigma_1(x)v_1(\cdot, a_1(x)) - \sigma_0(x)v_0(\cdot, a_0(x))\big)$$

$$- \big(\sigma_1(x')v_1(\cdot, a_1(x')) - \sigma_0(x')v_0(\cdot, a_0(x'))\big)|$$

$$\leq [\sigma_1 - \sigma_0]d^\alpha(x, x') \max_i \sup_{a_0} |v_i(\cdot, a_0)|$$

$$+ \max_i \|\sigma_i\| |(v_1(\cdot, a_1(x)) - v_0(\cdot, a_0(x))) - (v_1(\cdot, a_1(x')) - v_0(\cdot, a_0(x')))|$$

$$+ \max_i [\sigma_i]d^\alpha(x, x') \sup_y |v_1(\cdot, a_1(y)) - v_0(\cdot, a_0(y))|$$

$$+ \|\sigma_1 - \sigma_0\| \max_i \sup_{a_0} |\frac{\partial v_i}{\partial a_0}(\cdot, a_0)|[a_i]d^\alpha(x, x').$$

In order to deduce (285) from this inequality, in view of (290) and of our assumption (280) & (281), it remains to show for the second rhs terms

$$|(v_1(\cdot, a_1(x)) - v_0(\cdot, a_0(x))) - (v_1(\cdot, a_1(x')) - v_0(\cdot, a_0(x')))|$$

$$\leq \sup_{a_0} |\frac{\partial}{\partial a_0}(v_1 - v_0)(\cdot, a_0)| \max_i [a_i]d^\alpha(x, x')$$

$$+ \max_i \sup_{a_0} |\frac{\partial^2 v_i}{\partial a_0^2}(\cdot, a_0)| \max_j [a_j]d^\alpha(x, x')\|a_1 - a_0\|$$

(291)
$$+ \max_i \sup_{a_0} |\frac{\partial v_i}{\partial a_0}(\cdot, a_0)|[a_1 - a_0]d^\alpha(x, x').$$

We appeal again to the outcome (289) of the linear interpolation, which immediately yields the first rhs term (291) from the first rhs term in (289). For the second rhs term in (291), we appeal once more to formula (286) (applied to $\sigma = \frac{\partial v_s}{\partial a_0}(\cdot, a_s(x))$, $\sigma' = \frac{\partial v_s}{\partial a_0}(\cdot, a_s(x'))$, $v = (a_1 - a_0)(x)$, and $v' = (a_1 - a_0)(x')$).

STEP 2. Argument for part i) of the lemma. We apply Lemma 2 to the family of functions $v(\cdot, x) := \sigma(x)v(\cdot, a(x))$ and the family of distributions $v(\cdot, x) \diamond f := \sigma(x)v(\cdot, a(x)) \diamond f$, both parameterized by x. We claim that the hypotheses (28) and (30) of Lemma 2 are satisfied with

$$(292) \qquad N = N_0([\sigma] + \|\sigma\|[a]).$$

We also claim that in addition the hypotheses (33) and (34) of Lemma 2 are satisfied provided N is enlarged to

$$(293) \qquad N = N_0([\sigma] + \|\sigma\| + \|\sigma\|[a]) \lesssim N_0 \quad \text{for } [\sigma], [a], \|\sigma\| \leq 1.$$

Indeed, for (28) this follows from (283) of Step 1 with the Hölder semi-norm $[\cdot]$ playing the role of $|\cdot|$. In the same vein, hypothesis (33) follows from (282). The relevant hypothesis (279) of Step 1 coincides with the assumption (37) of this lemma. For (30) this follows again from (283) but this time with $[v(\cdot, a_0), (\cdot)_T] \diamond f$ playing the role of $v(\cdot, a_0)$, the supremum norm $\|\cdot\|$ playing the role of $|\cdot|$, and with $N_1 N_0 (T^{\frac{1}{4}})^{2\alpha-2}$, $N_1 \delta N_0 (T^{\frac{1}{4}})^{2\alpha-2}$ playing the role of $(N_0, \delta N_0)$. Likewise, hypothesis (34) follows from (282). The relevant hypothesis (279) of Step 1 coincide with the assumptions (38) of this lemma. With the definition (292), the output (32) of Lemma 2 quantifies the claim (39) of this lemma. Likewise, with definition (293), the output (35) of Lemma 2 turns into the claim (41) of this lemma.

STEP 3. Argument for part ii) of the lemma. We apply Lemma 2 to the family of functions $v(\cdot, x) := \sigma_1(x)v_1(\cdot, a_1(x)) - \sigma_0(x)v_0(\cdot, a_0(x))$ and the family of distributions $v(\cdot, x) \diamond f := \sigma_1(x) v_1(\cdot, a_1(x)) \diamond f - \sigma_0(x) v_0(\cdot, a_0(x)) \diamond f$, both parameterized by x. We claim that the hypotheses (28) and (30) of Lemma 2 are satisfied with

$$N = N_0 \max_{i,j} \left([\sigma_1 - \sigma_0] + \|\sigma_i\|[a_1 - a_0] + [\sigma_i]\|a_1 - a_0\|\right.$$
$$+ \|\sigma_1 - \sigma_0\|[a_i] + \|\sigma_i\|[a_j]\|a_1 - a_0\|\right)$$
$$(294) \qquad + \delta N_0 \max_{i,j} \left([\sigma_i] + \|\sigma_i\|[a_j]\right).$$

We also claim that in addition the hypotheses (33) and (34) of Lemma 2 are satisfied provided N is enlarged to

$$
\begin{aligned}
N = N_0 \max_{i,j} \big(& [\sigma_1 - \sigma_0] + \|\sigma_1 - \sigma_0\| + \|\sigma_i\|[a_1 - a_0] \\
& + \|\sigma_i\|\|a_1 - a_0\| + [\sigma_i]\|a_1 - a_0\| + \|\sigma_1 - \sigma_0\|[a_i] \\
& + \|\sigma_i\|[a_j]\|a_1 - a_0\|\big) + \delta N_0 \max_{i,j} \big([\sigma_i] + \|\sigma_i\| + \|\sigma_i\|[a_j]\big) \\
& \lesssim N_0\big([\sigma_1 - \sigma_0] + \|\sigma_1 - \sigma_0\| + [a_1 - a_0] + \|a_1 - a_0\| + \delta N_0
\end{aligned}
$$

(295) \qquad for $[\sigma], [a], \|\sigma\| \leq 1$.

Indeed, for (28) this follows from (285) of Step 1 with the Hölder semi-norm $[\cdot]$ playing the role of $|\cdot|$. In the same vein, hypothesis (33) follows from (284) in Step 1. The relevant hypotheses (280) and (281) of Step 1 coincide with the assumptions (42) and (43) of this lemma. For (30) this follows from Step 1 with $[v(\cdot, a_0), (\cdot)_T] \diamond f$ playing the role of $v(\cdot, a_0)$, the supremum norm $\|\cdot\|$ playing the role of $|\cdot|$, and with $N_1 N_0 (T^{\frac{1}{4}})^{2\alpha-2}, N_1 \delta N_0 (T^{\frac{1}{4}})^{2\alpha-2}$ playing the role of $(N_0, \delta N_0)$. Likewise, hypothesis (34) follows from (284). The relevant hypotheses (280) and (281) of Step 1 coincide with the assumptions (44), and (45) of this lemma. With the definition (294), the output (32) of Lemma 2, when applied to $(\delta u, \delta \nu, \delta M)$ playing the role of (u, ν, M), coincides with the claim (46) of this lemma. Likewise, with definition (295), the output (35) of Lemma 2 turns into the claim (47) of this lemma.

PROOF OF COROLLARY 1.
STEP 1. Generalization of part i) of Lemma 3. Let f_j, $j = 1, \cdots, J$, be finitely many distributions satisfying (36) and in addition

(296) $$\sup_{T \leq 1}(T^{\frac{1}{4}})^{2-\alpha}\|\sum_{j=1}^{J} c_j f_{1T}\| \leq \delta N_1$$

for some scalars $\{c_j\}_{j=1,\cdots,J}$. Suppose that next to (38) we have

(297) $$\sup_{T \leq 1}(T^{\frac{1}{4}})^{2-2\alpha}\|\{1, \frac{\partial}{\partial a_0}\}\sum_{j=1}^{J} c_j [v, (\cdot)_T] \diamond f_j\| \leq \delta N_1 N_0.$$

Then we claim in the situation of Lemma 3 i)

(298) $$\sup_{T \leq 1}(T^{\frac{1}{4}})^{2-2\alpha}\|\sum_{j=1}^{J} c_j [u, (\cdot)_T] \diamond f_j\| \lesssim \delta N_1 (M + N_0).$$

Indeed, in view of (296) & (297) we may apply part i) of Lemma 3 to $(f, v \diamond f)$ replaced by $\big(\sum_{j=1}^{J} c_j f_j, v \diamond (\sum_{j=1}^{J} c_j f_j) := \sum_{j=1}^{J} c_j v \diamond f_j\big)$ (and thus N_1 replaced by δN_1), yielding the existence of a distribution we

name $u \diamond (\sum_{j=1}^{J} c_j f_j)$ with

$$(299) \qquad \sup_{T \le 1} (T^{\frac{1}{4}})^{2-2\alpha} \| [u, (\cdot)_T] \diamond (\sum_{j=1}^{J} c_j f_j) \| \lesssim \delta N_1 (M + N_0).$$

It follows from the unique characterization of the product with u by the products with (v, x_1) through (39) in conjunction with our definition $v \diamond (\sum_{j=1}^{J} c_j f_j) = \sum_{j=1}^{J} c_j v \diamond f_j$ and the linearity of the regular product in form of $x_1 (\sum_{j=1}^{J} c_j f_j) = \sum_{j=1}^{J} c_j x_1 f_j$ that we have $u \diamond (\sum_{j=1}^{J} c_j f_j) = \sum_{j=1}^{J} c_j u \diamond f_j$. Hence (299) turns into (298).

STEP 2. Conclusion for part i) and ii). The only new element in part i) are the terms involving $\frac{\partial}{\partial a_0'}$ and $\frac{\partial^2}{\partial (a_0')^2}$ in (48), (49), and (50). For $\frac{\partial}{\partial a_0'}$, this follows from Step 1 by the choice of $\sum_{j=1}^{J} c_j f_j = f(\cdot, a_0^+) - f(\cdot, a_0^-)$ with two arbitrary parameters a_0^+, a_0^-, in which case $\delta N_1 = N_1 |a_0^+ - a_0^-|$. Likewise for $\frac{\partial^2}{(\partial a_0')^2}$, this follows from Step 1 by the choice of $\sum_{j=1}^{J} c_j f_j = (f(\cdot, a_0^{++}) - f(\cdot, a_0^{+-})) - (f(\cdot, a_0^{-+}) - f(\cdot, a_0^{--}))$ (and thus $J = 4$) with four arbitrary parameters $a_0^{++}, a_0^{+-}, a_0^{-+}, a_0^{--}$ with $a_0^{++} - a_0^{+-} = a_0^{-+} - a_0^{--}$; in this case $\delta N_1 = N_1 |(a_0^{++} - a_0^{+-}) - (a_0^{-+} - a_0^{--})|$.

Part ii) without the term involving $\frac{\partial}{\partial a_0'}$ in assumptions and conclusions follows from Step 1 by the choice of $\sum_{j=1}^{J} c_j f_j = f_1 - f_0$. Part ii) with the term involving $\frac{\partial}{\partial a_0'}$ follows by the choice of $\sum_{j=1}^{J} c_j f_j = (f_1 - f_0)(\cdot, a_0') - (f_1 - f_0)(\cdot, a_0'')$ (hence $J = 4$).

STEP 3. Conclusion for part iii). From the unique characterization of $u_i \diamond f$ through (39) and $(u_1 - u_0) \diamond f$ through (46), and the fact that by uniqueness of ν we have $\delta \nu = \nu_1 - \nu_0$, it follows $(u_1 - u_0) \diamond f = u_1 \diamond f - u_0 \diamond f$. Modulo this identity, the only new element in assumptions and conclusions of part iii) of this corollary over part ii) of Lemma 3 is the appearance of $\frac{\partial}{\partial a_0'}$. The argument proceeds by establishing the linearity property analogous to Step 1, now on the level of part ii) of Lemma 3, and applying it to $\sum_{j=1}^{J} c_j f_j = f(\cdot, a_0') - f(\cdot, a_0'')$, cf Step 2.

PROOF OF LEMMA 1.
We write for abbreviation $[\cdot] := [\cdot]_\alpha$.

STEP 1. Taylor's formulas. We start from two levels of Taylor's formula for the function b of u:

$$b(u') - b(u) = \int_0^1 dr \, b'(ru' + (1-r)u)(u' - u),$$

$$b(u') - b(u) - b'(u)(u' - u)$$
$$= \int_0^1 dr (1-r) b''(ru' + (1-r)u)(u' - u)^2.$$

Substituting u by $su_1 + (1-s)u_0$ and u' by $su_1' + (1-s)u_0'$, taking the derivative in s and integrating over $s \in [0,1]$ we obtain

$$(b(u_1') - b(u_0')) - (b(u_1) - b(u_0))$$
$$= \int_0^1 ds \int_0^1 dr \Big(b'(\cdot)\big((u_1' - u_1) - (u_0' - u_0)\big) + b''(\cdot)$$
$$\times \big(r(u_1' - u_0') + (1-r)(u_1 - u_0)\big)\big(s(u_1' - u_1) + (1-s)(u_0' - u_0)\big)\Big),$$
$$(b(u_1') - b(u_0')) - (b(u_1) - b(u_0))$$
$$\quad - (b'(u_1)(u_1' - u_1) - b'(u_0)(u_0' - u_0))$$
$$= \int_0^1 ds \int_0^1 dr(1-r)\Big(b''(\cdot)2\big(s(u_1' - u_1) + (1-s)(u_0' - u_0)\big)$$
$$\times \big((u_1' - u_1) - (u_0' - u_0)\big) + b'''(\cdot)$$
$$\times \big(r(u_1' - u_0') + (1-r)(u_1 - u_0)\big)\big(s(u_1' - u_1) + (1-s)(u_0' - u_0)\big)^2\Big),$$

where the argument of b'' and of b''' is given by $s(ru_1' + (1-r)u_1) + (1-s)(ru_0' + (1-r)u_0)$.

STEP 2. Inequalities. We use the formulas from Step 1 in terms of the inequalities

$$(300) \qquad\qquad\qquad |b(u') - b(u)| \leq \|b'\| |u' - u|,$$

$$(301) \qquad |b(u') - b(u) - b'(u)(u' - u)| \leq \frac{1}{2}\|b''\|(u' - u)^2,$$

and

$$\Big| b(u_1') - b(u_0')) - (b(u_1) - b(u_0)\Big|$$
$$\leq \|b'\| \big|(u_1' - u_1) - (u_0' - u_0)\big|$$
$$(302) \qquad + \|b''\| \max\{|u_1' - u_0'|, |u_1 - u_0|\} \max_i |u_i' - u_i|,$$
$$\Big|(b(u_1') - b(u_0')) - (b(u_1) - b(u_0))$$
$$\quad - (b'(u_1)(u_1' - u_1) - b'(u_0)(u_0' - u_0))\Big|$$
$$\leq \|b''\| \max_i |u_i' - u_i| |(u_1' - u_1) - (u_0' - u_0)|$$
$$(303) \qquad + \frac{1}{2}\|b'''\| \max\{|u_1' - u_0'|, |u_1 - u_0|\} \max_i |u_i' - u_i|^2.$$

By smuggling in a term $b'(u)(v' - v)$, we obtain from (301) by the triangle inequality

$$|b(u') - b(u) - b'(u)(v' - v)|$$
$$(304) \qquad \leq \|b'\| |(u' - u) - (v' - v)| + \|b''\|(u' - u)^2.$$

Likewise, we smuggle in the term $b'(u_1)(v'_1 - v_1) - b'(u_0)(v'_0 - v_0)$ into (303) and use (286) in form of

$$
\begin{aligned}
&\big(b'(u_1)(u'_1 - u_1) - b'(u_0)(u'_0 - u_0)\big) \\
&\quad - \big(b'(u_1)(v'_1 - v_1) - b'(u_0)(v'_0 - v_0)\big) \\
&= \frac{1}{2}(b'(u_1) + b'(u_0)) \\
&\quad \times \big(((u'_1 - u_1) - (u'_0 - u_0)) - ((v'_1 - v_1) - (v'_0 - v_0))\big) \\
&\quad + (b'(u_1) - b'(u_0)) \\
&\quad \times \frac{1}{2}\big(((u'_1 - u_1) - (v'_1 - v_1)) + ((u'_0 - u_0) - (v'_0 - v_0))\big),
\end{aligned}
$$

to obtain

$$
\begin{aligned}
&\big|\big(b(u'_1) - b(u'_0)\big) - (b(u_1) - b(u_0)) \\
&\quad - \big(b'(u_1)(v'_1 - v_1) - b'(u_0)(v'_0 - v_0)\big)\big| \\
&\leq \|b'\| \big|\big((u'_1 - u_1) - (u'_0 - u_0)\big) - \big((v'_1 - v_1) - (v'_0 - v_0)\big)\big| \\
&\quad + \|b''\| |u_1 - u_0| \max_i |(u'_i - u_i) - (v'_i - v_i)| \\
&\quad + \|b''\| \max_i |u'_i - u_i| |(u'_1 - u_1) - (u'_0 - u_0)| \\
(305) \quad &\quad + \frac{1}{2}\|b'''\| \max\{|u'_1 - u'_0|, |u_1 - u_0|\} \max_i |u'_i - u_i|^2.
\end{aligned}
$$

STEP 3. Application of inequalities. We apply (300) from Step 2 to $u = u(x),\ u' = u(y)$, which yields

$$(306) \qquad\qquad\qquad [b(u)] \leq \|b'\|[u]$$

and (304) to $v = \sigma(x)v(x, a(x)) + \nu(x)x_1,\ v' = \sigma(x)v(y, a(x)) + \nu(x)y_1$, which yields

$$
\begin{aligned}
&|b(u)(y) - b(u)(x) - (b'(u)\sigma)(x)(v(y, a(x)) - v(x, a(x))) \\
&\quad - (b'(u)\nu)(x)(y - x)_1| \leq (\|b'\|M + \|b''\|[u]^2)d^{2\alpha}(y, x),
\end{aligned}
$$

from which we deduce

$$\tilde{M} \leq \|b'\|M + \|b''\|[u]^2.$$

We combine the latter with (306) to (22).

Likewise, we apply (302) to $u_i = u_i(x),\ u'_i = u_i(y)$, which yields

$$(307) \qquad [b(u_1) - b(u_0)] \leq \|b'\|[u_1 - u_0] + \|b''\|\|u_1 - u_0\| \max_i [u_i],$$

and (305) to $v_i = \sigma_i(x)v_i(x, a_i(x)) + \nu_i(x)x_1, \quad v_i' = \sigma_i(x)v_i(y, a_i(x)) + \nu_i(x)y_1$, which entails

$$\begin{aligned}
\Big| &\big((b(u_1) - b(u_0))(y) - (b(u_1) - b(u_0))(x)\big) \\
&- \big((b'(u_1)\sigma_1)(x)(v_1(y, a_1(x)) - v_1(x, a_1(x)))\big) \\
&- (b'(u_0)\sigma_0)(x)(v_0(y, a_0(x)) - v_0(x, a_0(x)))\big) \\
&- (b'(u_1)\nu_1 - b'(u_0)\nu_0)(x)(y - x)_1 \Big| \\
&\leq \big(\|b'\|\delta M + \|b''\|\|u_1 - u_0\| \max_i M_i + \|b''\| \max_i[u_i][u_1 - u_0] \\
&\quad + \frac{1}{2}\|b'''\|\|u_1 - u_0\| \max_i[u_i]^2\big) d^{2\alpha}(y, x).
\end{aligned}$$

The latter implies

$$\begin{aligned}
\delta\tilde{M} &\leq \|b'\|\delta M + \|b''\|\|u_1 - u_0\| \max_i M_i + \|b''\| \max_i[u_i][u_1 - u_0] \\
&\quad + \frac{1}{2}\|b'''\|\|u_1 - u_0\| \max_i[u_i]^2.
\end{aligned}$$

In combination with (307) and $\|b(u_1) - b(u_0)\| \leq \|b'\|\|u_1 - u_0\|$ this yields (24).

STEP 4. Estimate of the modulating functions. Indeed, estimate (23) is straightforward. We will now argue for (25). We appeal to identity (286) (with (σ, σ', v, v') replaced by $(b(u_1), b(u_0), \sigma_1, \sigma_0)$) for the estimate

$$\begin{aligned}
&\|b'(u_1)\sigma_1 - b'(u_0)\sigma_0\| \\
(308) \qquad &\leq \|b''\|\|u_1 - u_0\| \max_i\|\sigma_i\| + \|b'\|\|\sigma_1 - \sigma_0\|.
\end{aligned}$$

We appeal to identity (287) $(\sigma_1, \sigma_1', v_1, v_1')$ replaced by $(b(u_1(y)), b(u_0(y)), \sigma_1(y), \sigma_0(y))$ and $(\sigma_0, \sigma_0', v_0, v_0')$ by $(b(u_1(x)), b(u_0(x)), \sigma_1(x), \sigma_0(x))$ to obtain

$$\begin{aligned}
&[b'(u_1)\sigma_1 - b'(u_0)\sigma_0] \\
&\leq [b'(u_1) - b'(u_0)] \max_i\|\sigma_i\| + \max_i\|b'(u_i)\|[\sigma_1 - \sigma_0] \\
&\quad + \max_i[b'(u_i)]\|\sigma_1 - \sigma_0\| + \|b'(u_1) - b'(u_0)\| \max_i[\sigma_i].
\end{aligned}$$

The first rhs term can be estimated by (307) with b'' playing the role of b' so that

$$\begin{aligned}
&[b'(u_1)\sigma_1 - b'(u_0)\sigma_0] \\
&\leq (\|b''\|[u_1 - u_0] + \|b'''\|\|u_1 - u_0\| \max_i[u_i]) \max_i\|\sigma_i\| \\
&\quad + \|b'\|[\sigma_1 - \sigma_0] + \|b''\| \max_i[u_i]\|\sigma_1 - \sigma_0\| + \|b''\|\|u_1 - u_0\| \max_i[\sigma_i].
\end{aligned}$$

Estimate (25) follows from this and (308).

PROOF OF COROLLARY 3.

STEP 1. Proof of (i) $\Rightarrow$ (ii). As v is a $C^{\beta+2}$ function the assumption that u is modelled after v according to $a(u)$, $\sigma(u)$ implies that u is of class $C^{2\alpha}$, in particular $\partial_1 u$ is a function of class $C^{2\alpha-1}$ (of course, as we will see below, u is actually of class $C^{\beta+2}$ but we do not have this information to our disposal yet). Together with the regularity assumption on f this implies that there is a classical interpretation of the products $\sigma(u)f$ and $a(u)\partial_1^2 u$ the latter as a distribution. In fact, this is obvious for $\sigma(u)f$ and for $a(u)\partial_1^2 u$ we can set, for example,

$$(309) \qquad a(u)\partial_1^2 u := \partial_1(a(u)\partial_1 u) - \partial_1 a(u)\partial_1 u.$$

The claim then follows from standard parabolic regularity theory as soon as we have established that

$$(310) \qquad \sigma(u)\diamond f = \sigma(u)f + \sigma'(u)\sigma(u)g^{(1)}(\cdot, a(u))$$

$$(311) \qquad a(u)\diamond \partial_1^2 u = a(u)\partial_1^2 u + a'(u)\sigma^2(u)g^{(2)}(\cdot, a(u), a(u)).$$

We first argue that (310) holds. To see this, first by Lemma 1 $\sigma(u)$ is modelled after v according to $a(u)$ and $\sigma'(u)\sigma(u)$. Then, Lemma 3 characterizes $\sigma(u)\diamond f$ as the unique distribution for which

$$(312) \quad \lim_{T\downarrow 0} \|[\sigma(u), (\cdot)_T]\diamond f - \sigma'(u)\sigma(u)E[v, (\cdot)_T]\diamond f - \nu[x_1, (\cdot)]f\| = 0.$$

By the C^β regularity of f as well as the $C^{2\alpha}$ regularity of $\sigma(u)$ one sees immediately that each of the commutators in this expression goes to zero if $\diamond$ is replaced by the classical product

$$\|[\sigma(u), (\cdot)_T]f\|, \ \|\sigma'(u)\sigma(u)E[v, (\cdot)_T]f\|, \ \|\nu[x, (\cdot)]f\| \to 0$$

for $T \to 0$. Hence (312) turns into

$$\lim_{T\downarrow 0} \|\sigma(u)f - (\sigma(u)\diamond f)_T - \sigma'(u)\sigma(u)g_T^{(1)}(\cdot, a(u))\| = 0.$$

Since, $g(\cdot, a_0) \in C^\beta$ by assumption, this yields (310). In the same way, one can see that for any a_0' we have

$$(313) \qquad a(u)\diamond \partial_1^2 v(\cdot, a_0') = a(u)\partial_1^2 v(\cdot, a_0') + a'(u)\sigma(u)g^{(2)}(\cdot, a(u), a_0')$$

(the classical definition of $a(u)\partial_1^2 v(\cdot, a_0')$ poses no problem because v is of class $C^{\beta+2}$).

It remains to upgrade (313) to (311), i.e. the second factor $\partial_1^2 v$ in (313) should be replaced by $\partial_1^2 u$. To this end we make the ansatz

$$(314) \qquad a(u)\diamond \partial_1^2 u = a(u)\partial_1^2 u + a'(u)\sigma^2(u)g^{(2)}(\cdot, a(u), a(u)) + B,$$

and aim to show that $B = 0$. Recalling once more that u is modelled after v according to $a(u)$, $\sigma(u)$ we invoke Lemma 4 and plug in our

ansatz (314) to obtain

$$\lim_{T\downarrow 0}\|[a(u),(\cdot)_T]\partial_1^2 u - (a'(u)\sigma^2(u)g^{(2)}(\cdot,a(u),a(u)))_T + (B)_T$$

$$(315) \qquad - \sigma(u)E[a(u),(\cdot)_T]\diamond\partial_1^2 v\| = 0.$$

Plugging (313) into (315) we obtain

$$\lim_{T\downarrow 0}\|[a(u),(\cdot)_T]\partial_1^2 u - (a'(u)\sigma^2(u)g^{(2)}(\cdot,a(u),a(u)))_T + (B)_T$$

$$(316) \quad - \sigma(u)E[a(u),(\cdot)_T]\partial_1^2 v - a'(u)\sigma^2(u)E(g^{(2)}(\cdot,a(u),a'_0))_T\| = 0.$$

Now according to our regularity assumptions we have both

$$\|(a'(u)\sigma^2(u)g^{(2)}(\cdot,a(u),a(u)))_T - a'(u)\sigma^2(u)E(g^{(2)}(\cdot,a(u),a'_0))_T\| \to 0$$

$$\|\sigma(u)E[a(u),(\cdot)_T]\partial_1^2 v\| \to 0,$$

for $T \to 0$, which reduces (316) to

$$\lim_{T\downarrow 0}\|[a(u),(\cdot)_T]\partial_1^2 u - B_T\| = 0,$$

where we recall that the classical commutator is defined based on (309). Now, according to its definition (309) we have $[a(u),(\cdot)_T] \rightharpoonup 0$, which characterizes B as 0.

STEP 2. Proof of (ii) $\Rightarrow$ (i). If u as well as all the $v(\cdot,a_0)$ are of class $C^{\beta+2}$, then u is automatically modelled after v according to $a(u)$ and $\sigma(u)$. Thus we can conclude from Step 1 that (310) and (311) hold which in turn implies that u solves $\partial_2 u - P(a(u)\diamond\partial_1^2 u + \sigma(u)\diamond f) = 0$ distributionally.

5. PROOFS OF THE STOCHASTIC BOUNDS

PROOF OF LEMMA 6.
STEP 1. Proof of (130). Assumption (129) and the stationarity and periodicity of f imply that for $T \leq 1$

$$\langle f_T^2(0)\rangle$$

$$= \left\langle \int_{[0,1)^2} f_T^2\, dx \right\rangle = \sum_{k\in(2\pi\mathbb{Z})^2} \hat\psi_T(k)^2 \hat C(k)$$

$$\leq (T^{\frac{1}{4}})^{-3+\lambda_1+\lambda_2} \sum_{k\in(2\pi\mathbb{Z})^2\setminus\{0\}} (T^{\frac{1}{4}})^3 \frac{e^{-2|T^{\frac{1}{4}}k_1|^4-2|T^{\frac{1}{2}}k_2|^2}}{(T^{\frac{1}{4}}(1+|k_1|))^{\lambda_1}(T^{\frac{1}{2}}(1+|k_2|))^{\frac{\lambda_2}{2}}}$$

$$\lesssim (T^{\frac{1}{4}})^{2\alpha-4}.$$

In the last estimate we have used that the sum in the second line is a Riemann sum approximation to the integral $\int e^{-2\hat k_1^4-2\hat k_2^2}|\hat k_1|^{-\lambda_1}|\hat k_2|^{-\frac{\lambda_2}{2}}\, d\hat k$ which converges due to $\lambda_1, \frac{\lambda_2}{2} < 1$.

The fact that f_T is Gaussian and stationary implies that we have $\langle |f_T(x)|^p \rangle \lesssim \langle f_T(0)^2 \rangle^{\frac{p}{2}}$ uniformly over x, which permits to write

$$\Big\langle \int_{[0,1)^2} |f_T|^p dx \Big\rangle^{\frac{1}{p}} \lesssim \langle f_T^2(0) \rangle^{\frac{1}{2}} \lesssim (T^{\frac{1}{4}})^{\alpha-2}.$$

In order to upgrade this L^p bound to an L^∞ bound under the expectation we observe that by the semi-group property (17) we have $f_T = (f_{T/2})_{T/2}$ such that Young's inequality implies

$$\|f_T\| \lesssim \|f_{T/2}\|_{L^p} \|\psi_{T/2,\mathrm{per}}\|_{L^{p'}}$$

where as before $\|\cdot\|$ refers to the supremums norm over $\mathbb{R}^2$ (or equivalently $[0,1)^2$ by periodicity) and $\|\cdot\|_{L^p}$ refers to the L^p norm over $[0,1)^2$, $p' = \frac{p}{p-1}$ is the dual exponent of p and $\psi_{T,\mathrm{per}}(x) = \sum_{k \in \mathbb{Z}^2} \psi_T(x+k)$ is the periodisation of ψ_T. By observing that for for small T the difference $\big| \|\psi_{T,\mathrm{per}}\|_{L^{p'}} - \big(\int_{\mathbb{R}^2} \psi_T^{p'} dx \big)^{\frac{1}{p'}} \big|$ stays bounded and scaling we get $\|\psi_{T,\mathrm{per}}\|_{L^{p'}} \lesssim (T^{\frac{1}{4}})^{\frac{3}{p}}$ such that finally

$$\langle \|f_T\|^p \rangle^{\frac{1}{p}} \lesssim (T^{\frac{1}{4}})^{-\frac{3}{p}} \langle \|f_T\|_{L^p}^p \rangle^{\frac{1}{p}} \lesssim (T^{\frac{1}{4}})^{\alpha-2-\frac{3}{p}}.$$

To also accommodate for the supremum over the scales T we first note that $\|f_{T+t}\| \lesssim \|f_T\|$ implies

$$\sup_{T \le 1} (T^{\frac{1}{4}})^{2-\alpha'} \|f_T\| \lesssim \sup_{T \le 1, \mathrm{dyadic}} (T^{\frac{1}{4}})^{2-\alpha'} \|f_T\|,$$

where the subscript dyadic means that this supremum is only taken over T of the form $T = 2^{-k}$ for $k \ge 0$. Then we write

$$\Big\langle \Big(\sup_{T \le 1, \mathrm{dyadic}} (T^{\frac{1}{4}})^{2-\alpha'} \|f_T\| \Big)^p \Big\rangle \le \sum_{T \le 1, \mathrm{dyadic}} (T^{\frac{1}{4}})^{p(2-\alpha')} \langle \|f_T\|^p \rangle$$

$$\lesssim \sum_{T \le 1, \mathrm{dyadic}} (T^{\frac{1}{4}})^{p(2-\alpha')} (T^{\frac{1}{4}})^{p(\alpha-2)-3},$$

which converges as soon as $p > \frac{4}{\alpha-\alpha'}$ and establishes (130) for large p. The bound for smaller p can be derived from the bound for large p and Jensen's inequality.

STEP 2. Proof of (131). The bound on the ε-differences follows from (130) as soon as we have established the deterministic bound

$$(317) \qquad \|(f_\varepsilon)_T - f_T\| \lesssim \min \big\{ \frac{\varepsilon}{T}, 1 \big\} \|f_{T/2}\|.$$

Given the bound

$$(318) \qquad \|(f_\varepsilon)_T - f_T\| \le \|(f_\varepsilon)_T\| + \|f_T\| \lesssim \|f_{T/2}\|$$

which follows from the triangle inequality, the semi-group property in the form $(f_\varepsilon)_T = (f_{T/2})_{T/2} * \psi'_\varepsilon$ and $f_T = (f_{T/2})_{T/2}$ as well as the fact that the operators $(\cdot)_T$ and $\psi'_\varepsilon *$ are bounded with respect to $\|\cdot\|$

uniformly in T and ε, it suffices to establish (317) for $\varepsilon \leq \frac{1}{4}T$. We then write

$$\|(f_\varepsilon)_T - f_T\| = \|(\psi_{T/2} * \psi'_\varepsilon - \psi_{T/2}) * f_{T/2}\|$$
$$\leq \int_{\mathbb{R}^2} |\psi_{T/2} * \psi'_\varepsilon - \psi_{T/2}|dx \, \|f_{T/2}\|,$$

and have thereby reduced (317) (and hence (131)) to establishing that

$$\int_{\mathbb{R}^2} |\psi_{T/2} * \psi'_\varepsilon - \psi_{T/2}|dx \lesssim \frac{\varepsilon}{T} \qquad \text{for } \varepsilon \leq \frac{T}{4}.$$

By scaling (recalling that $\psi_T(x_1, x_2) = T^{-\frac{3}{4}}\psi_1(T^{-\frac{1}{4}}x_1, T^{-\frac{1}{2}}x_2)$), it suffices to show this bound for $\frac{T}{2} = 1$, in which case it turns into

$$\int_{\mathbb{R}^2} |\psi_1 * \ell^{-3}\psi'(\ell^{-1}x_1, \ell^{-2}x_2) - \psi_1(x_1, x_2)| \lesssim \ell$$

for $0 \leq \ell \leq (\frac{1}{2})^{\frac{1}{4}}$, which is immediate for a Schwartz kernels ψ_1 and ψ'.

PROOF OF LEMMA 7. By stationarity and (128) we may write

$$\langle v_\varepsilon(0, a_0)f_\varepsilon(0)\rangle$$
$$= \left\langle \int_{[0,1)^2} v_\varepsilon(x, a_0)f_\varepsilon(x)dx \right\rangle = \sum_{k \in (2\pi\mathbb{Z})^2} \langle \hat{v}_\varepsilon(k, a_0)\hat{f}_\varepsilon(-k)\rangle$$
$$= \sum_{k \in (2\pi\mathbb{Z})^2} \hat{G}(k, a_0)\langle f_\varepsilon(k)f_\varepsilon(-k)\rangle = \sum_{k \in (2\pi\mathbb{Z})^2} \hat{G}(k, a_0)\hat{C}(k)|(\hat{\psi}'_\varepsilon)^2(k)|.$$

As the left hand side of this expression is real valued, the imaginary part of the sum of the right hand side also has to vanish. As both $\hat{C}$ and $\hat{\psi}_\varepsilon$ are real valued this means that we can replace $\hat{G}(\cdot, a_0)$ by its real part (given in (132)) thereby yielding (134).

The same calculation yields

$$\langle v_\varepsilon(0, a_0)\partial_1^2 v(0, a'_0)\rangle = \sum_{k \in (2\pi\mathbb{Z})^2} \hat{G}(k, a_0)(-k_1^2)\hat{G}(-k, a'_0)\langle f_\varepsilon(k)f_\varepsilon(-k)\rangle$$
$$= \sum_{k \in (2\pi\mathbb{Z})^2} \hat{G}(k, a_0)(-k_1^2)\hat{G}(k, a'_0)\hat{C}(k)|(\hat{\psi}'_\varepsilon)^2(k)|,$$

and after calculating the real part of $\hat{G}(k, a_0)(-k_1^2)\hat{G}(k, a'_0)$ we arrive at (135).

PROOF OF LEMMA 8.
STEP 1. Bound on the expectation. We start the derivation of (141) by bounding $\langle[v', (\cdot)_T] \diamond f'\rangle$. By definition $\langle v' \diamond f'\rangle = 0$. Furthermore,

by stationarity and (127) and (129) we have

$$
\left| \langle v'(f')_T \rangle \right| = \left| \sum_k \langle \hat{v}'(-k)\hat{\psi}_T(k)\hat{f}'(k) \rangle \right|
$$

$$
= \left| \sum_k (\hat{M}_2 \hat{G})(-k)\hat{\psi}_T(k)\hat{M}_1(k)\hat{C}(k) \right|
$$

$$
\leq (T^{\frac{1}{4}})^{-3+2+\lambda_1+\lambda_2-\kappa_1-\kappa_2}
$$

$$
\times \sum_k (T^{\frac{1}{4}})^3 \frac{\hat{\psi}_T(k)}{|T^{\frac{1}{4}}k_1|^2 + |T^{\frac{1}{2}}k_2|} \frac{(|T^{\frac{1}{4}}k_1|^4 + |T^{\frac{1}{2}}k_2|^2)^{\frac{\kappa_1+\kappa_2}{4}}}{(T^{\frac{1}{4}}(1+|k_1|))^{\lambda_1}(T^{\frac{1}{2}}(1+|k_2|))^{\frac{\lambda_2}{2}}}
$$

$$
\lesssim (T^{\frac{1}{4}})^{2\alpha-2-\kappa_1-\kappa_2},
$$

where the sum is taken over $(2\pi\mathbb{Z})^2$. In the last step we have used the fact that the Riemann sum in the third line approximates the integral $\int \hat{\psi}_1(\hat{k}) \frac{(|\hat{k}_1|^4+|\hat{k}_2|^2)^{\frac{\kappa_1+\kappa_2}{4}}}{|\hat{k}_1|^2+|\hat{k}_2|} \frac{1}{|\hat{k}_1|^{\lambda_1}|\hat{k}_2|^{\lambda_2/2}} d\hat{k}$. This integral converges because the singularities near the axes $\hat{k}_1 = 0$ and $\hat{k}_2 = 0$ are integrable because of $\lambda_1, \frac{\lambda_2}{2} < 1$ and the singularity near the origin is integrable due to $-2+\lambda_1+\lambda_2 = -3+2\alpha > -3$, where we appeal to parabolic dimension counting (alternatively, one may split the integral into $|x_1| \leq \sqrt{|x_2|}$ and its complement). This establishes that the expectation satisfies the bound (141).

STEP 2. Bound on the variance. For the variances we seek the bound

$$
\left| \langle ([v', (\cdot)_T] \diamond f')^2 \rangle - \langle v'(f')_T \rangle^2 \right|^{\frac{1}{2}} \lesssim (T^{\frac{1}{4}})^{2\alpha-2-\kappa_1-\kappa_2},
$$

which by definition of $\diamond$ can be expressed equivalently without the renormalisation as

$$
(319) \qquad \left| \langle ([v', (\cdot)_T]f')^2 \rangle - \langle [v', (\cdot)_T]f' \rangle^2 \right|^{\frac{1}{2}} \lesssim (T^{\frac{1}{4}})^{2\alpha-2-\kappa_1-\kappa_2}.
$$

To derive the estimate in the form (319) we write using stationarity once more

$$
\langle ([v', (\cdot)_T]f')^2 \rangle = \left\langle \int_{[0,1)^2} ([v', (\cdot)_T]f')^2 dx \right\rangle = \sum_{k \in (2\pi\mathbb{Z})^2} \left\langle |\widehat{[v', (\cdot)_T]f'}(k)|^2 \right\rangle.
$$

The expression appearing in the last expectation can be evaluated according to its definition

$$
\widehat{[v', (\cdot)_T]f'}(k)
$$

$$
= \sum_{\ell \in (2\pi\mathbb{Z})^2} (\hat{\psi}_T(\ell) - \hat{\psi}_T(k))\hat{v}'(k-\ell)\hat{f}'(\ell)
$$

$$
(320) \qquad = \sum_{\ell \in (2\pi\mathbb{Z})^2} (\hat{\psi}_T(\ell) - \hat{\psi}_T(k))(\hat{M}_2\hat{G})(k-\ell)\hat{f}(k-\ell)\hat{M}_1(\ell)\hat{f}(\ell),
$$

which permits to write

$$
\begin{aligned}
&\langle ([v',(\cdot)_T]f')^2\rangle \\
&= \sum_k \sum_\ell \sum_{\ell'} (\hat\psi_T(\ell) - \hat\psi_T(k))(\hat\psi_T(-\ell') - \hat\psi_T(-k)) \\
&\quad \times (\hat M_2\hat G)(k-\ell)(\hat M_2\hat G)(-(k-\ell'))\hat M_1(\ell)\hat M_1(-\ell') \\
&\quad \times \langle \hat f(k-\ell)\hat f(\ell)\hat f(-(k-\ell'))\hat f(-\ell')\rangle,
\end{aligned}
$$

(321)

where all sums are taken over $(2\pi\mathbb{Z})^2$. We now use (127) and the Gaussian identity

(322) $\quad \langle \hat f(k-\ell)\hat f(\ell)\hat f(-(k-\ell'))\hat f(-\ell')\rangle$
$$
= \delta_{k,0}\hat C(\ell)\hat C(\ell') + \delta_{\ell,\ell'}\hat C(k-\ell)\hat C(\ell) + \delta_{k-\ell,\ell'}\hat C(k-\ell)\hat C(\ell),
$$

and bound the three terms resulting from plugging this identity into (321) one by one. The first term coincides with the square of the expectation

$$
\begin{aligned}
&\sum_\ell \sum_{\ell'} (\hat\psi_T(\ell) - \hat\psi_T(0))(\hat\psi_T(-\ell') - \hat\psi_T(0)) \\
&\quad (\hat M_2\hat G)(-\ell)(\hat M_2\hat G)(\ell')(\hat M_1\hat C)(\ell)(\hat M_1\hat C)(-\ell') \\
&= \Big(\sum_\ell (\hat\psi_T(\ell) - \hat\psi_T(0))(\hat M_2\hat G)(-\ell)(\hat M_1\hat C)(\ell) \Big)^2 \\
&= \langle [v',(\cdot)_T]f'\rangle^2
\end{aligned}
$$

so that the required bound (319) follows as soon as we can bound the remaining two terms. The term originating from the third contribution on the right hand side of (322) can be absorbed into the second term using the Cauchy-Schwarz inequality. Indeed, we may write

$$
\begin{aligned}
&\sum_k \sum_\ell (\hat\psi_T(\ell) - \hat\psi_T(k))(\hat\psi_T(-(k-\ell)) - \hat\psi_T(-k)) \\
&\quad \times (\hat M_2\hat G)(k-\ell)(\hat M_2\hat G)(-\ell)(\hat M_1\hat C)(\ell)(\hat M_1\hat C)(-(k-\ell)) \\
&\leq \Big(\sum_{k,\ell} |\hat\psi_T(\ell) - \hat\psi_T(k)|^2 |(\hat M_2\hat G)(k-\ell)|^2 |\hat M_1(\ell)|^2 \hat C(-(k-\ell))\hat C(\ell) \Big)^{\frac12} \\
&\quad \times \Big(\sum_{k,\ell} |\hat\psi_T(-(k-\ell)) - \hat\psi_T(-k)|^2 |(\hat M_2\hat G)(-\ell)|^2 \\
&\qquad\qquad \times |\hat M_1(-(k-\ell))|^2 \hat C(-(k-\ell))\hat C(\ell) \Big)^{\frac12},
\end{aligned}
$$

and the second factor on the right hand side can be seen to coincide with the first one by performing the change of variables $k' = -k$ and $\ell' = \ell - k$ and the symmetry $\hat C(k) = \hat C(-k)$. Hence, it only remains to bound the term coming from the second contribution on the right

hand side of (322). We use the assumptions (129) and (140) to write

$$\sum_{k,\ell} |\hat{\psi}_T(\ell) - \hat{\psi}_T(k)|^2 |(\hat{M}_2\hat{G})(k-\ell)|^2 |\hat{M}_1(\ell)|^2 \hat{C}(\ell)\hat{C}(k-\ell)$$

$$\lesssim \sum_{k \neq \ell} \frac{|\hat{\psi}_T(\ell) - \hat{\psi}_T(k)|^2}{(|(k-\ell)_1^2 + |(k-\ell)_2|)^2} \frac{(|\ell_1|^4 + |\ell_2|^2)^{\frac{\kappa_1}{2}}}{(1+|\ell_1|)^{\lambda_1}(1+|\ell_2|)^{\frac{\lambda_2}{2}}}$$

$$\times \frac{(|(k-\ell)_1|^4 + |(k-\ell)_2|^2)^{\frac{\kappa_2}{2}}}{(1+|(k-\ell)_1|)^{\lambda_1}(1+|(k-\ell)_2|)^{\frac{\lambda_2}{2}}}$$

$$= (T^{\frac{1}{4}})^{4\alpha-4-2\kappa_1-2\kappa_2} \sum_{k \neq \ell} (T^{\frac{1}{4}})^6 \Big(\frac{\hat{\psi}_T(\ell) - \hat{\psi}_T(k)}{|T^{\frac{1}{4}}(k-\ell)_1|^2 + T^{\frac{1}{2}}|(k-\ell)_2|}\Big)^2$$

$$\times \frac{(|T^{\frac{1}{4}}\ell_1|^4 + |T^{\frac{1}{2}}\ell_2|^2)^{\frac{\kappa_1}{2}}}{(T^{\frac{1}{4}}(1+|\ell_1|))^{\lambda_1}(T^{\frac{1}{2}}(1+|\ell_2|))^{\frac{\lambda_2}{2}}}$$

$$(323) \qquad \times \frac{(|T^{\frac{1}{4}}(k-\ell)_1|^4 + |T^{\frac{1}{2}}(k-\ell)_2|^2)^{\frac{\kappa_2}{2}}}{(T^{\frac{1}{4}}(1+|(k-\ell)_1|))^{\lambda_1}(T^{\frac{1}{2}}(1+|(k-\ell)_2|))^{\frac{\lambda_2}{2}}}.$$

STEP 3. Bound on an integral. In order to show that the expression (323) is bounded by $\lesssim (T^{\frac{1}{4}})^{4\alpha-4-2\kappa_1-2\kappa_2}$ which in turn establishes (319), it remains to show that the integral which is approximated by the Riemann sum in the last line converges. To this end we write this integral as

$$\iint \Big(\frac{\hat{\psi}_1(\hat{\ell}) - \hat{\psi}_1(\hat{k})}{(\hat{k}-\hat{\ell})_1^2 + |\hat{k}_2 - \hat{\ell}_2|}\Big)^2 \frac{(|\hat{\ell}_1|^4 + |\hat{\ell}_2|^2)^{\frac{\kappa_1}{2}}}{|\hat{\ell}_1|^{\lambda_1}|\hat{\ell}_2|^{\frac{\lambda_2}{2}}}$$

$$\times \frac{(|(\hat{k}-\hat{\ell})_1|^4 + |(\hat{k}-\hat{\ell})_2|^2)^{\frac{\kappa_2}{2}}}{|\hat{k}_1 - \hat{\ell}_1|^{\lambda_1}|\hat{k}_2 - \hat{\ell}_2|^{\frac{\lambda_2}{2}}} d\hat{\ell}d\hat{k}$$

$$= \iint \Big(\frac{\hat{\psi}_1(\hat{\ell}) - \hat{\psi}_1(\hat{\ell}+h)}{h_1^2 + |h_2|}\Big)^2 \frac{(|\hat{\ell}_1|^4 + |\hat{\ell}_2|^2)^{\frac{\kappa_1}{2}}}{|\hat{\ell}_1|^{\lambda_1}|\hat{\ell}_2|^{\frac{\lambda_2}{2}}} \frac{(|h_1|^4 + |h_2|^2)^{\frac{\kappa_2}{2}}}{|h_1|^{\lambda_1}|h_2|^{\frac{\lambda_2}{2}}} d\hat{\ell}dh.$$

We treat the inner $d\hat{\ell}$ integral first. For $|h_1|+|h_2| \leq 2$ we use the bound $|\hat{\psi}_1(\hat{\ell}) - \hat{\psi}_1(\hat{\ell}+h)| \lesssim (|h_1| + |h_2|) \int_0^1 |\nabla\hat{\psi}_1(\hat{\ell}+\theta h)|d\theta$ which yields using Jensen's inequality:

$$\int (\hat{\psi}_1(\hat{\ell}) - \hat{\psi}_1(\hat{\ell}+h))^2 \frac{(|\hat{\ell}_1|^4 + |\hat{\ell}_2|^2)^{\frac{\kappa_1}{2}}}{|\hat{\ell}_1|^{\lambda_1}|\hat{\ell}_2|^{\frac{\lambda_2}{2}}} d\hat{\ell}$$

$$\lesssim (|h_1| + |h_2|)^2 \sup_{\theta \in [0,1]} \int |\nabla\hat{\psi}_1(\hat{\ell}+\theta h)|^2 \frac{(|\hat{\ell}_1|^4 + |\hat{\ell}_2|^2)^{\frac{\kappa_1}{2}}}{|\hat{\ell}_1|^{\lambda_1}|\hat{\ell}_2|^{\frac{\lambda_2}{2}}} d\hat{\ell}$$

$$(324) \qquad \lesssim (|h_1| + |h_2|)^2.$$

The integral converges because according to our assumption (129) we have $\lambda_1, \frac{\lambda_2}{2} < 1$ which together with the exponential decay of $\nabla\hat{\psi}_1$

implies that the integrals over $\{\hat{\ell}: |\hat{\ell}_2| \geq 1\}$ and $\{\hat{\ell}: |\hat{\ell}_1| \geq 1\}$ are finite and that the integral over $\{|\hat{\ell}_1|, |\hat{\ell}_2| \leq 1\}$ is finite as well. For $|h_1| + |h_2| > 2$ we use the estimate $(\hat{\psi}_1(\hat{\ell}) - \hat{\psi}_1(\hat{\ell} + h))^2 \lesssim \hat{\psi}_1^2(\hat{\ell}) + \hat{\psi}_1^2(\hat{\ell} + h)$ and obtain

$$\int (\hat{\psi}_1(\hat{\ell}) - \hat{\psi}_1(\hat{\ell} + h))^2 \frac{(|\hat{\ell}_1|^4 + |\hat{\ell}_2|^2)^{\frac{\kappa_1}{2}}}{|\hat{\ell}_1|^{\lambda_1} |\hat{\ell}_2|^{\frac{\lambda_2}{2}}} d\hat{\ell} \lesssim 1 + (|h_1|^4 + |h_2|^2)^{\frac{\kappa_1}{2}}.$$

It remains to treat the outer dh-integral

$$(325) \qquad \int \left(\frac{\min\{1, |h_1| + |h_2|\}}{h_1^2 + |h_2|} \right)^2 \frac{(|h_1|^4 + |h_2|^2)^{\frac{\kappa_1 + \kappa_2}{2}}}{|h_1|^{\lambda_1} |h_2|^{\frac{\lambda_2}{2}}} dh.$$

The integrability of this expression outside of $\{h: |h_1|, |h_2| \leq 1\}$ can be seen using the decay of $(h_1^2 + |h_2|)^{-2}$ for large $|h_1| + |h_2|$ (note that our assumption (129) allows for negative $\lambda_1 > -3 + 2\alpha$ or $\lambda_2 > -2 + 2\alpha$, but this potential growth at infinity as well as the growth of the numerator is still compensated by the decay of $(h_1^2 + |h_2|)^{-2}$ because of $\kappa_1 + \kappa_2 \ll 1$) and the integrability of $|h_1|^{-\lambda_1}$, and $|h_2|^{-\frac{\lambda_2}{2}}$ for small $|h_1|$ and $|h_2|$ due to $\lambda_1, \frac{\lambda_2}{2} < 1$. Near the origin we drop the numerator $(|h_1|^4 + |h_2|^2)^{\frac{\kappa_1 + \kappa_2}{2}}$ and split into $\mathcal{A}_1 = \{h: h_1^2 \leq |h_2| \leq 1\}$ and $\mathcal{A}_2 = \{h: |h_2| \leq h_1^2 \leq 1\}$. We furthermore bound brutally $|h_1| + |h_2| \leq |h_1| + \sqrt{|h_2|}$ which yields the bound

$$\int_{\mathcal{A}_1} \left(\frac{|h_1| + |h_2|}{h_1^2 + |h_2|} \right)^2 \frac{1}{|h_1|^{\lambda_1} |h_2|^{\frac{\lambda_2}{2}}} dh$$

$$\lesssim \int_0^1 \int_0^{\sqrt{|h_2|}} \frac{1}{|h_1|^2 + |h_2|} \frac{1}{|h_1|^{\lambda_1} |h_2|^{\frac{\lambda_2}{2}}} dh_1 dh_2$$

$$\lesssim \int_0^1 \frac{1}{|h_2|^{1 + \frac{\lambda_2}{2}}} \int_0^{\sqrt{|h_2|}} \frac{1}{|h_1|^{\lambda_1}} dh_1 dh_2$$

$$\lesssim \int_0^1 \frac{1}{|h_2|^{\frac{1 + \lambda_1 + \lambda_2}{2}}} dh_2 < \infty.$$

Here we have made use of (129) and more specifically of $\lambda_1 < 1$ as well as $\frac{1}{2}(1 + \lambda_1 + \lambda_2) = \alpha < 1$. The integral over $\mathcal{A}_2$ can be bounded in a similar way concluding the proof of (141).

PROOF OF COROLLARY 4. The estimate (142) follows immediately from (141) either with f_ε in the role of f' (i.e. $\hat{M}_1 = \hat{\psi}_\varepsilon'$) or $\left(\frac{\partial}{\partial a_0'} \right)^{n'} \partial_1^2 v_\varepsilon(\cdot, a_0')$ in the role of f' which amounts to

$$\hat{M}_1(k) = \frac{(-1)^n n! k_1^{2n}}{(a_0' k_1^2 - ik_2)^n} \frac{-k_1^2}{a_0' k_1^2 - ik_2} \hat{\psi}_\varepsilon'(k)$$

and with $\left(\frac{\partial}{\partial a_0}\right)^n v_\varepsilon(\cdot, a_0)$ in the role of v' i.e.

$$\hat{M}_2(k) = \frac{(-1)^n n! k_1^{2n}}{(a_0' k_1^2 - i k_2)^n} \hat{\psi}_\varepsilon'(k).$$

For the derivatives wrt ε we observe that the product rule applies in the form

$$\frac{\partial}{\partial \varepsilon}\left(\left[\frac{\partial}{\partial a_0^n} v_\varepsilon(\cdot, a_0), (\cdot)_T\right] \diamond \{f_\varepsilon, \frac{\partial}{\partial (a_0')^{n'}} \partial_1^2 v_\varepsilon\}\right)$$

$$= \left[\frac{\partial}{\partial \varepsilon}\frac{\partial}{\partial a_0^n} v_\varepsilon(\cdot, a_0), (\cdot)_T\right] \diamond \{f_\varepsilon, \frac{\partial}{\partial (a_0')^{n'}} \partial_1^2 v_\varepsilon\}$$

$$+ \left[\frac{\partial}{\partial a_0^n} v_\varepsilon(\cdot, a_0), (\cdot)_T\right] \diamond \frac{\partial}{\partial \varepsilon}\{f_\varepsilon, \frac{\partial}{\partial (a_0')^{n'}} \partial_1^2 v_\varepsilon\}.$$

We then apply (142) to each of the terms on the rhs separately. The multipliers $\hat{M}_1$, $\hat{M}_2$ are the same as above only with $\hat{\psi}_\varepsilon'$ replaced by $\frac{\partial}{\partial \varepsilon}\psi_\varepsilon' \lesssim (k_1^4 + k_2^2)^{\frac{\kappa}{4}} \varepsilon^{1-\frac{\kappa}{4}}$ in M_2 for the first term and in M_1 for the second term.

Proof of Proposition 2.
Step 1. Bound on the supremum over x and T. Our first claim is that uniformly over $a_0, a_0' \in [\lambda, 1]$, $\varepsilon \in (0, 1]$ for some $\kappa \ll 1$ and for all $n, n' \geq 1$ we have

$$\Big\langle \Big(\sup_{T \leq 1} (T^{\frac{1}{4}})^{2-2\alpha'} \|\frac{\partial^n}{\partial a_0^n} \frac{\partial^{n'}}{\partial (a_0')^{n'}} [v_\varepsilon(\cdot, a_0), (\cdot)_T]$$

(326) $$\diamond \{f_\varepsilon, \partial_1^2 v_\varepsilon(\cdot, a_0')\}\| \Big)^p \Big\rangle^{\frac{1}{p}} \lesssim 1.$$

$$\Big\langle \Big(\sup_{T \leq 1} (T^{\frac{1}{4}})^{2-2\alpha'+\kappa} \|\frac{\partial}{\partial \varepsilon} \frac{\partial^n}{\partial a_0^n} \frac{\partial^{n'}}{\partial (a_0')^{n'}} [v_\varepsilon(\cdot, a_0), (\cdot)_T]$$

(327) $$\diamond \{f_\varepsilon, \partial_1^2 v_\varepsilon(\cdot, a_0')\}\| \Big)^p \Big\rangle^{\frac{1}{p}} \lesssim \varepsilon^{\frac{\kappa}{4}-1}.$$

To keep the notation concise, for the moment we restrict ourselves to the bound for $[v_\varepsilon, (\cdot)_T] \diamond \partial_1^2 v$ without the derivatives wrt a_0, a_0', ε. The general case of (326) follows in the identical way and so does (327) if in the proof (142) is replaced by (143). To simplify the notation further we drop the subscript ε as well as the dependence on a_0, a_0' for the moment.

First of all $[v, (\cdot)_T] \diamond \partial_1^2 v$ is a random variable in the second Wiener chaos over the Gaussian field f such that by equivalence of moments for random variables in the second Wiener chaos and by stationarity, the bound (142) can be upgraded to

(328) $$\langle |[v, (\cdot)_T] \diamond \partial_1^2 v|^p \rangle^{\frac{1}{p}} \lesssim (T^{\frac{1}{4}})^{2\alpha-2}$$

for any $p < \infty$ (the implicit constant then depends on p). We now aim to upgrade this L^p bound to an L^∞ bound over x. At the same time, we want to show that the supremum over all $T \leq 1$ can be reduced to a supremum over all dyadic T, i.e. over T of the form $T = 2^{-k}$ for a $k \geq 0$. For any given $T \leq 1$ there is a unique a dyadic $T' \leq \frac{1}{2}$ such that $T = 2T' + t$ for with $2T' \leq T < 4T'$ and we refer to this choice when we write T' in the sequel.

We make use of the commutator identity

$$(329) \qquad [v, (\cdot)_T] \diamond \partial_1^2 v = \big([v, (\cdot)_{T'}] \diamond \partial_1^2 v\big)_{T'+t} + [v, (\cdot)_{T'+t}](\partial_1^2 v)_{T'}.$$

The second term on the right hand side of (329) can be bounded directly by making the convolution with $\psi_{T'+t}$ explicit and using Lemma 9 and (18) to get

$$\big([v, (\cdot)_{T'+t}](\partial_1^2 v)_{T'}\big)(x)$$

$$= \int (v(x) - v(y))\psi_{T'+t}(x - y)(\partial_1^2 v)_{T'}(y)dy$$

$$\lesssim [v]_{\alpha'}\|(\partial_1^2 v)_{T'}\| \int d^{\alpha'}(x, y)|\psi_{T'+t}(x - y)|dy$$

$$\lesssim ((T')^{\frac{1}{4}})^{\alpha'-2}((T' + t)^{\frac{1}{4}})^{\alpha'}[v]_{\alpha'}(((T')^{\frac{1}{4}})^{2-\alpha'}\|(\partial_1^2 v)_T\|)$$

$$\lesssim (T^{\frac{1}{4}})^{2\alpha'-2}\big(\sup_{T\leq 1}(T^{\frac{1}{4}})^{2-\alpha'}\|f_T\|\big)^2.$$

(Note that here we have used Lemma 9 to bound the semi-norm $[v]_\alpha$ by the $C^{\alpha-2}$ norm of f. The same argument applies to $\frac{\partial}{\partial a_0^n}v(\cdot, a_0)$ using an identity like (222) and iterating Lemma 9. Similarly the $C^{\alpha-2}$ norm of $\frac{\partial}{\partial a_0^n}\partial_1^2 v(\cdot, a_0)$ is controlled by the $C^{\alpha-2}$ norm of f.) Taking the sup over x and T and then the p-th moment in the expectation we get from Lemma 6

$$\big\langle\big(\sup_{T\leq 1}(T^{\frac{1}{4}})^{2-2\alpha'}\|[v, (\cdot)_{T'+t}](\partial_1^2 v)_{T'}\|\big)^p\big\rangle^{\frac{1}{p}} \lesssim \big\langle\big(\sup_{T\leq 1}(T^{\frac{1}{4}})^{2(2-\alpha')}\|f_T\|^2\big)^p\big\rangle^{\frac{1}{p}}$$

$$\lesssim 1.$$

To bound the first term on the right hand side of (329) we use Young's inequality (on the torus) in the form

$$\big\|\big([v, (\cdot)_{T'}] \diamond \partial_1^2 v\big)_{T'+t}\big\| \lesssim \|[v, (\cdot)_{T'}] \diamond \partial_1^2 v\|_{L^p}\|\psi_{T'+t,\mathrm{per}}\|_{L^{p'}}$$

where L^p denotes the L^p norm on the torus $[0, 1)^2$, $p' = \frac{p}{p-1}$ is the dual exponent of p and $\psi_{T,\mathrm{per}}(x) = \sum_{k\in\mathbb{Z}^2}\psi_T(x + k)$ is the periodisation of ψ_T. By observing that for for small T the difference $\big|\|\psi_{T,\mathrm{per}}\|_{L^{p'}} - \big(\int_{\mathbb{R}^2}\psi_T^{p'}dx\big)^{\frac{1}{p'}}\big|$ stays bounded and scaling we get $\|\psi_{T,\mathrm{per}}\|_{L^{p'}} \lesssim (T^{\frac{1}{4}})^{-\frac{3}{p}}$ such that finally

$$\big\|\big([v, (\cdot)_{T'}] \diamond \partial_1^2 v\big)_{T'+t}\big\| \lesssim ((T' + t)^{\frac{1}{4}})^{-\frac{3}{p}}\|[v, (\cdot)_{T'}] \diamond \partial_1^2 v\|_{L^p}.$$

Dropping the t and taking the supremum over T we get for any p

$$\left(\sup_{T\leq 1}(T^{\frac{1}{4}})^{2-2\alpha'}\|\big([v,(\cdot)_{T'}]\diamond\partial_1^2 v\big)_{T'+t}\|\right)^p$$
$$\lesssim \sum_{T'\leq\frac{1}{2},\text{dyadic}}((T')^{\frac{1}{4}})^{p(2-2\alpha')}((T')^{\frac{1}{4}})^{-3}\|[v,(\cdot)_{T'}]\diamond\partial_1^2 v\|_{L^p}^p.$$

Finally, we take the expectation of this estimat and use (328) and the stationarity to get

$$\left\langle\left(\sup_{T\leq 1}(T^{\frac{1}{4}})^{2-2\alpha'}\|\big([v,(\cdot)_{T'}]\diamond\partial_1^2 v\big)_{T'+t}\|\right)^p\right\rangle$$
$$\lesssim \sum_{T'\leq\frac{1}{2},\text{dyadic}}((T')^{\frac{1}{4}})^{p(2-2\alpha')}((T')^{\frac{1}{4}})^{-3}\left\langle\|[v,(\cdot)_{T'}]\diamond\partial_1^2 v\|_{L^p}^p\right\rangle$$
$$\lesssim \sum_{T\leq\frac{1}{2},\text{dyadic}}(T^{\frac{1}{4}})^{p(2\alpha-2\alpha')}(T^{\frac{1}{4}})^{-3}.$$

To obtain (326) for fixed $\alpha' < \alpha$ and p we now apply this estimate for an exponent $p' > \max\{\frac{3}{2(\alpha-\alpha')}, p\}$, sum the resulting geometric series and use Jensen's inequality.

STEP 2. Introducing the supremum over a_0, a_0'. In the following steps we use the notation

$$(330)\quad A(T, a_0, a_0', \varepsilon) = \frac{\partial^n}{\partial a_0^n}\frac{\partial^{n'}}{\partial(a_0')^{n'}}[v_\varepsilon(\cdot, a_0), (\cdot)_T]\diamond\{f_\varepsilon, \partial_1^2 v_\varepsilon(\cdot, a_0')\}.$$

In this step aim to show that for each $\varepsilon\in(0,1]$ and $\kappa\ll 1$

$$(331)\quad \left\langle\left(\sup_{a_0,a_0'\in[\lambda,1]}\sup_{T\leq 1}(T^{\frac{1}{4}})^{2-2\alpha'}\|A(T, a_0, a_0', \varepsilon)\|\right)^p\right\rangle^{\frac{1}{p}}\lesssim 1,$$

$$(332)\quad \left\langle\left(\sup_{a_0,a_0'\in[\lambda,1]}\sup_{T\leq 1}(T^{\frac{1}{4}})^{2-2\alpha'}\|\frac{\partial}{\partial\varepsilon}A(T, a_0, a_0', \varepsilon)\|\right)^p\right\rangle^{\frac{1}{p}}\lesssim \varepsilon^{\frac{\kappa}{4}-1}.$$

For (331) we use the Sobolev inequality

$$\sup_{a_0\in[\lambda,1]}\sup_{a_0'\in[\lambda,1]}\|A\|^p\lesssim \int_\lambda^1\int_\lambda^1\|A\|^p da_0\, da_0' + \int_\lambda^1\int_\lambda^1\|\frac{\partial}{\partial a_0}A\|^p da_0\, da_0'$$
$$+ \int_\lambda^1\int_\lambda^1\|\frac{\partial}{\partial a_0'}A\|^p da_0\, da_0'$$

which holds for $p > 2$. Taking the supremum over T, then the expectation and invoking Fubini's theorem and (326) yields

$$\Big\langle \Big(\sup_{a_0 \in [\lambda,1]} \sup_{a_0' \in [\lambda,1]} \sup_{T \leq 1} (T^{\frac{1}{4}})^{2-2\alpha'} \|A\| \Big)^p \Big\rangle$$

$$\lesssim \int_\lambda^1 \int_\lambda^1 \Big\langle \Big(\sup_{T \leq 1} (T^{\frac{1}{4}})^{2-2\alpha'} \|A\| \Big)^p \Big\rangle da_0 \, da_0'$$

$$+ \int_\lambda^1 \int_\lambda^1 \Big\langle \Big(\sup_{T \leq 1} (T^{\frac{1}{4}})^{2-2\alpha'} \|\frac{\partial}{\partial a_0} A\| \Big)^p \Big\rangle da_0 \, da_0'$$

$$+ \int_\lambda^1 \int_\lambda^1 \Big\langle \Big(\sup_{T \leq 1} (T^{\frac{1}{4}})^{2-2\alpha'} \|\frac{\partial}{\partial a_0'} A\| \Big)^p \Big\rangle da_0 \, da_0' \lesssim 1,$$

so (331) follows. For (332) we repeat the same calculation with A replaced by $\frac{\partial}{\partial \varepsilon} A$ and (326) replaced by (327).

STEP 3. Bounding the supremum over ε. Let $A(T, a_0, a_0', \varepsilon)$ be defined as in (330) above. In this step we upgrade (331) and (332) to

$$(333) \qquad \Big\langle \Big(\sup_{\varepsilon \in (0,1]} \sup_{a_0, a_0' \in [\lambda,1]} \sup_{T \leq 1} (T^{\frac{1}{4}})^{2-2\alpha'} \|A(T, a_0, a_0', \varepsilon)\| \Big)^p \Big\rangle^{\frac{1}{p}} \lesssim 1$$

valid for $\alpha' < \alpha$. We start with the elementary inequality

$$\sup_{0 < \varepsilon \leq 1} |A(\varepsilon)| \leq \int_0^1 |A(\varepsilon)| d\varepsilon + \int_0^1 |\frac{\partial}{\partial \varepsilon} A(\varepsilon)| d\varepsilon.$$

We now multiply with $(T^{\frac{1}{4}})^{2-\alpha'+\kappa}$ for some $\alpha' < \alpha$ and $\kappa \ll 1$, take the supremum over T, a_0, a_0' of this estimate, take p-th moments and invoke Minkowski's inequality to arrive at

$$\Big\langle \Big(\sup_{\varepsilon \in (0,1]} \sup_{a_0 \in [\lambda,1]} \sup_{a_0' \in [\lambda,1]} \sup_{T \leq 1} (T^{\frac{1}{4}})^{2-2\alpha'+\kappa} \|A\| \Big)^p \Big\rangle^{\frac{1}{p}}$$

$$\lesssim \int_0^1 \Big\langle \Big(\sup_{a_0 \in [\lambda,1]} \sup_{a_0' \in [\lambda,1]} \sup_{T \leq 1} (T^{\frac{1}{4}})^{2-2\alpha'+\kappa} \|A\| \Big)^p \Big\rangle^{\frac{1}{p}} d\varepsilon$$

$$+ \int_0^1 \Big\langle \Big(\sup_{a_0 \in [\lambda,1]} \sup_{a_0' \in [\lambda,1]} \sup_{T \leq 1} (T^{\frac{1}{4}})^{2-2\alpha'+\kappa} \|\frac{\partial}{\partial \varepsilon} A\| \Big)^p \Big\rangle^{\frac{1}{p}} d\varepsilon$$

$$\lesssim \int_0^1 (1 + \varepsilon^{\frac{\kappa}{4}-1}) d\varepsilon \lesssim 1.$$

Now (333) follows by relabelling $-2\alpha' + \kappa$ as $-2\alpha'$.

STEP 4. Bounding ε differences. We claim that for $\kappa \ll 1$ and all $p < \infty$ and $\alpha' < \alpha$

$$\Big\langle \Big(\sup_{a_0, a_0' \in [\lambda,1]} \sup_{\varepsilon_1 \neq \varepsilon_2 \in (0,1]} \sup_{T \leq 1} (T^{\frac{1}{4}})^{2-2\alpha'+\kappa} |\varepsilon_2 - \varepsilon_1|^{-\frac{\kappa}{4}}$$

$$(334) \qquad\qquad \times \|A(T, a_0, a_0', \varepsilon_1) - (T, a_0, a_0', \varepsilon_1)\| \Big)^p \Big\rangle^{\frac{1}{p}} \lesssim 1.$$

We start the argument with Hölder's inequality in the form

$$|A(\varepsilon_2) - A(\varepsilon_1)| = \Big| \int_{\varepsilon_1}^{\varepsilon_2} \frac{\partial}{\partial \varepsilon} A(\varepsilon) d\varepsilon \Big|$$

$$\le |\varepsilon_2 - \varepsilon_1|^{\frac{\kappa}{4}} \Big(\int_0^1 \big| \frac{\partial}{\partial \varepsilon} A(\varepsilon) \big|^{\frac{1}{1-\frac{\kappa}{4}}} d\varepsilon \Big)^{1-\frac{\kappa}{4}}.$$

Now, we multiply this estimate with $(T^{\frac{1}{4}})^{2-2\alpha+\kappa+\bar{\kappa}}$ for another $\bar{\kappa} \ll 1$, take the supremum over T, a_0 and a_0', then p-th moments, and finally invoke Minkowski's inequality and (327) to get

$$\Big\langle \Big(\sup_{a_0,a_0'\in[\lambda,1]} \sup_{\varepsilon_1\neq\varepsilon_2\in(0,1]} \sup_{T\le 1} (T^{\frac{1}{4}})^{2-2\alpha'+\kappa+\bar{\kappa}} |\varepsilon_2 - \varepsilon_1|^{-\frac{\bar{\kappa}}{4}}$$

$$\times \|A(T,a_0,a_0',\varepsilon_2) - (T,a_0,a_0',\varepsilon_1)\| \Big)^p \Big\rangle^{\frac{1}{p}}$$

$$\lesssim \Big(\int_0^1 \Big\langle \Big(\sup_{a_0,a_0'\in[\lambda,1]} \sup_{T\le 1} (T^{\frac{1}{4}})^{2-2\alpha'+\kappa+\bar{\kappa}} \|\frac{\partial}{\partial\varepsilon}A\|^p \Big) \Big\rangle^{\frac{1}{p}(\frac{1}{1-\frac{\kappa}{4}})} d\varepsilon \Big)^{1-\frac{\kappa}{4}}$$

$$\lesssim \int_0^1 \varepsilon^{(\frac{\kappa+\kappa}{4}-1)(\frac{1}{1-\frac{\kappa}{4}})} d\varepsilon \lesssim 1,$$

so (334) follows by relabelling $-2\alpha' + \kappa$ as $-2\alpha'$.

STEP 5. Conclusion. To shorten notation, we only treat the product $v_\varepsilon \diamond f_\varepsilon$. Writing

$$(v_\varepsilon \diamond f_\varepsilon)_T = v_\varepsilon(f_\varepsilon)_T - [v_\varepsilon, (\cdot)_T]f_\varepsilon,$$

and invoking (130) and (131) for the first and (334) for the second term imply that $v_\varepsilon \diamond f_\varepsilon$ converges almost surely with respect to the $C^{\alpha-2}$ norm to a limit $v \diamond f$. Furthermore, the estimates (333) and (334) remain true if the supremum over $\varepsilon \in (0,1]$ is extended to include the limit as $\varepsilon \to 0$.

6. APPENDIX

Lemma 9. *The (mean-free) solution of (62) satisfies the estimate*

$$(335) \qquad \sup_{a_0}[v(\cdot,a_0)]_\alpha \lesssim \sup_{T\le 1}(T^{\frac{1}{4}})^{2-\alpha}\|f_T\|.$$

PROOF OF LEMMA 9.
All functions are 1-period if not stated otherwise.
STEP 1. Reduction. We claim that it is enough to show

$$\sup_{T\le 1}(T^{\frac{1}{4}})^{2-\alpha}\|f_T\|$$

$$(336) \qquad \sim \inf \Big\{ [f_1]_\alpha + [f_2]_\alpha + |c| \Big| f = \partial_1^2 f_1 + \partial_2 f_2 + c \Big\},$$

where the infimum is over all triplets (f_1, f_2, c) of two functions and a constant. Incidentally, the equivalence confirms that the lhs indeed

defines the (parabolic) $C^{\alpha-2}$-norm. Indeed, let the decomposition $f = \partial_1^2 f_1 + \partial_2 f_2 + c$ be near-optimal in the rhs of (336), that is,

$$(337) \qquad [f_1]_\alpha + [f_2]_\alpha \lesssim \sup_{T \leq 1} (T^{\frac{1}{4}})^{2-\alpha} \|f_T\|.$$

By uniqueness of the mean-free solution of (62) this induces $v(\cdot, a_0) = \partial_1^2 v_1 + \partial_2 v_2$ where v_i, $i = 1, 2$, denote the mean-free solutions of

$$(\partial_2 - a_0 \partial_1^2) v_1 = f_1 \quad \text{and} \quad (\partial_2 - a_0 \partial_1^2) v_2 = f_2.$$

By classical $C^{\alpha+2}$-Schauder theory [12, Theorem 8.6.1] we have $[\partial_1^2 v_i]_\alpha + [\partial_2 v_i]_\alpha \lesssim [f_i]_\alpha$, so that (335) follows from (337).

STEP 2. For the solution of

$$(338) \qquad (\partial_2 - \partial_1^2) v = Pf$$

we claim

$$(339) \qquad \|v_T - v\| \lesssim N_0 \max\{(T^{\frac{1}{4}})^\alpha, (T^{\frac{1}{4}})^2\} \quad \text{for all } T > 0,$$

where we have set for abbreviation $N_0 := \sup_{T \leq 1} (T^{\frac{1}{4}})^{2-\alpha} \|f_T\|$. We start by noting that the definition of N_0 may be extended to the control of $T \geq 1$ by the semi-group property (17) in form of $f_T = (f_1)_{T-1}$ and (18) in form of $\|f_T\| \lesssim \|f_1\|$. We thus have

$$(340) \qquad \|f_T\| \lesssim N_0 \max\{T^{\alpha-2}, 1\}.$$

By approximation through (standard) convolution, which preserves (338) and does increase N_0, we may assume that f and v are smooth. By definition of the convolution $(\cdot)_t$ we have

$$\partial_t v_t = -(\partial_1^4 - \partial_2^2) v_t = (-\partial_1^2 - \partial_2)(\partial_2 - \partial_1^2) v_t \stackrel{(338)}{=} (-\partial_1^2 - \partial_2) P f_t$$
$$\stackrel{(17)}{=} (-\partial_1^2 - \partial_2)(f_{\frac{t}{2}})_{\frac{t}{2}}.$$

Hence we obtain by (18) for all $T \leq 1$

$$\|\partial_t v_t\| \lesssim (t^{\frac{1}{4}})^{-2} \|f_{\frac{t}{2}}\| \stackrel{(340)}{\lesssim} N_0 \max\{(t^{\frac{1}{4}})^{\alpha-4}, (t^{\frac{1}{4}})^{-2}\}.$$

Integrating over $t \in (0, T)$ we obtain (339) by the triangle inequality.

STEP 3. For v defined through (338) we have

$$(341) \qquad [v]_\alpha \lesssim N_0$$

As in Step 2 we may assume that f and v are smooth so that $[v]_\alpha$ is finite. Because of periodicity, it is sufficient to probe Hölder continuity for pairs (x, y) of points with $d(y, x) \leq 4$. For any $T > 0$ we have the identity

$$v(y) - v(x) = (v_T - v)(y) - (v_T - v)(x)$$
$$- \int_0^1 \partial_1 v_T (sy + (1-s)x)(y - x)_1 + \partial_2 v_T (sy + (1-s)x)(y - x)_2 \, ds,$$

from which we obtain the inequality

$$|v(y) - v(x)| \leq 2\|v_T - v\| + \|\partial_1 v_T\|d(y, x) + \|\partial_2 v_T\|d^2(y, x).$$

From Step 2 and (18) we obtain the estimate

$$|v(y) - v(x)|$$
$$\lesssim N_0 \max\{(T^{\frac{1}{4}})^\alpha, (T^{\frac{1}{4}})^2\} + [v]_\alpha\big((T^{\frac{1}{4}})^{\alpha-1}d(y, x) + (T^{\frac{1}{4}})^{\alpha-2}d^2(y, x)\big).$$

With the ansatz $T^{\frac{1}{4}} = \frac{1}{\epsilon}d(y, x)$ for some $\epsilon \leq 1$ and making use of $d(y, x) \lesssim 1$ we obtain

$$|v(y) - v(x)| \lesssim \big(\epsilon^{-2}N_0 + [v]_\alpha(\epsilon^{1-\alpha} + \epsilon^{2-\alpha})\big)d^\alpha(y, x).$$

Fixing an ϵ sufficiently small to absorb the last rhs term into the lhs we infer (341).

STEP 4. We finally establish the equivalence of norms (336). The direction $\lesssim$ follows immediately from (18). The direction $\gtrsim$ follows from Step 3 with $f_1 = v$, $f_2 = -v$, and $c = \int_{[0,1)^2} f$.

Lemma 10.

$$(342) \qquad \sup_{T \leq 1}(T^{\frac{1}{4}})^{1-\alpha}\|[x_1, (\cdot)_T]f\| \lesssim \sup_{T \leq 1}(T^{\frac{1}{4}})^{2-\alpha}\|f_T\|$$

PROOF OF LEMMA 10.
Introducing the kernel $\tilde{\psi}_T(x) := x_1\psi_T(x)$ we start by claiming the representation

$$(343) \qquad\qquad [x_1, (\cdot)_T]f = 2\tilde{\psi}_{\frac{T}{2}} * f_{\frac{T}{2}}.$$

Indeed, by definition of the commutator and $\tilde{\psi}_T$ we have $[x_1, (\cdot)_T]f = \tilde{\psi}_T * f$, so that the above representation follows from the formula

$$(344) \qquad\qquad \tilde{\psi}_T = 2\tilde{\psi}_{\frac{T}{2}} * \psi_{\frac{T}{2}}.$$

The argument for (344) relies on the fact that convolution is commutative in form of $\tilde{\psi}_{\frac{T}{2}} * \psi_{\frac{T}{2}} = \psi_{\frac{T}{2}} * \tilde{\psi}_{\frac{T}{2}}$, which spelled out means $\int dy(x_1 - y_1)\psi_{\frac{T}{2}}(x - y)\psi_{\frac{T}{2}}(y) = \int dy\psi_{\frac{T}{2}}(x - y)y_1\psi_{\frac{T}{2}}(y)$, and thus implies $2\int dy(x_1 - y_1)\psi_{\frac{T}{2}}(x - y)\psi_{\frac{T}{2}}(y) = x_1\int dy\psi_{\frac{T}{2}}(x - y)\psi_{\frac{T}{2}}(y)$, that is $2(\tilde{\psi}_{\frac{T}{2}} * \psi_{\frac{T}{2}})(x) = x_1(\psi_{\frac{T}{2}} * \psi_{\frac{T}{2}})(x)$. Together with the semi-group property (17) in form of $\psi_{\frac{T}{2}} * \psi_{\frac{T}{2}} = \psi_T$ this yields (344).

From the representation (343) we obtain the estimate

$$\|[x_1, (\cdot)_T]f\| \leq 2\int dx|x_1\psi_{\frac{T}{2}}(x)|\|f_{\frac{T}{2}}\| \overset{(18)}{\lesssim} T^{\frac{1}{4}}\|f_{\frac{T}{2}}\|,$$

which yields the desired (342).

Lemma 11. *Let ψ and ψ' be Schwartz functions over $\mathbb{R}^2$ with $\int \psi = \int \psi' = 1$. For $T > 0$ set*

(345)
$$\psi_T(x_1, x_2) = T^{-\frac{3}{4}}\psi\Big(\frac{x_1}{T^{\frac{1}{4}}}, \frac{x_2}{T^{\frac{1}{2}}}\Big), \qquad \psi'_T(x_1, x_2) = T^{-\frac{3}{4}}\psi'\Big(\frac{x_1}{T^{\frac{1}{4}}}, \frac{x_2}{T^{\frac{1}{2}}}\Big).$$

(i) For an arbitrary Schwartz distribution $f \in \mathcal{S}'(\mathbb{R}^2)$ set

(346)
$$(f)_T = f * \psi_T \qquad and \qquad (f)'_T = f * \psi'_T.$$

Then for any $\gamma < 0$ we have

(347)
$$\sup_{T \leq 1}(T^{\frac{1}{4}})^{-\gamma}\|(f)_T\| \lesssim \sup_{T \leq 1}(T^{\frac{1}{4}})^{-\gamma}\|(f)'_T\|,$$

where $\lesssim$ only refers to ψ, ψ' and γ.

(ii) Let $\alpha > 0$ and $\gamma < 0$. Let u be function of class C^α and f a distribution of class C^γ. Furthermore, let $u \diamond f$ be an arbitrary distribution of class C^γ and define the generalised commutators $[u, (\cdot)] \diamond f = u(f)_T - (u \diamond f)_T$ and $[u, (\cdot)'] \diamond f = u(f)'_T - (u \diamond f)'_T$. Then for $\bar{\gamma} = \gamma + \alpha$ we have

$$\sup_{T \leq 1}(T^{\frac{1}{4}})^{-\bar{\gamma}}\|[u, (\cdot)_T]f\| \lesssim \sup_{T \leq 1}(T^{\frac{1}{4}})^{-\bar{\gamma}}\|[u, (\cdot)'_T]f\|$$

(348)
$$+ [u]_\alpha \sup_{T \leq 1}(T^{\frac{1}{4}})^{-\gamma}\|(f)'_T\|,$$

where $\lesssim$ depends on α, γ as well as ψ and ψ'.

PROOF OF LEMMA 11.

STEP 1. The proof relies on a variant of a construction from [3] which we recall in this step. For the reader's convenience we give self-contained proofs of the identities in Step 4 below. First of all, for any $p > 0$ there exists a Schwartz function ω^0 such that $\varphi' = \omega^0 * \psi'$ satisfies

(349)
$$\int x^n \varphi'(x)dx = \begin{cases} 1 & \text{for } \alpha = 0 \\ 0 & \text{for } 0 < \|n\|_{\text{par}} < p, \end{cases}$$

where for $n = (n_1, n_2)$ and $x = (x_1, x_2)$ we use the parabolic norm $\|n\|_{\text{par}} = |n_1| + 2|n_2|$ and $x^n = x_1^{n_1} x_2^{n_2}$. Furthermore, it is shown that for any p and any φ' satisfying (349) as well as $\theta \ll 1$ (depending on φ, ψ, p), the function ψ can be represented as

(350)
$$\psi = \sum_{k=0}^{\infty} \omega^{(k)} * \varphi'_{\theta^k},$$

where φ'_{θ^k} is the rescaled version of φ' defined as in (345) for $T = \theta^k$, and the $\omega^{(k)}$ are Schwartz functions satisfying

(351)
$$\int |\omega^{(k)}| \lesssim (C_0 \theta^{\frac{p}{4}})^k,$$

where $C_0 = C_0(\varphi', \psi, p)$. The convergence of the sum in (350) holds in $L^1(\mathbb{R}^2)$. Additionally, we will make use of the bounds

$$(352) \qquad \int d^\alpha(0, x)|\omega^{(k)}(x)|dx \lesssim (C_0\theta^{\frac{p}{4}})^k.$$

We summarize this as $\psi = \sum_{k=0}^\infty \omega^{(k)} * \omega_{\theta^k}^0 * \psi'_{\theta^k}$, which can be rescaled as

$$(353) \qquad \psi_T = \sum_{k=0}^\infty \omega_T^{(k)} * \omega_{\theta^k T}^0 * \psi'_{\theta^k T},$$

where as before the index T expresses that a function is rescaled by T as in (345).

STEP 2. Equipped with these results we now proceed to prove (347). Wlog we assume that f is smooth. We set $\sup_{T \leq 1} \|(f)'_T\|(T^{\frac{1}{4}})^{-\gamma} = N_0$ and write

$$\|(f)_T\| = \|f * \sum_{k=0}^\infty \psi'_{\theta^k T} * (\omega_T^{(k)} * \omega_{\theta^k T}^0)\|$$

$$= \|\sum_{k=0}^\infty (f)'_{\theta^k T} * (\omega_T^{(k)} * \omega_{\theta^k T}^0)\| \leq \sum_{k=0}^\infty \|(f)'_{\theta^k T}\| \int |\omega_T^{(k)}| \int |\omega_{\theta^k T}^0|$$

$$\lesssim N_0 \sum_{k=0}^\infty (\theta^{\frac{k}{4}}T^{\frac{1}{4}})^\gamma (C_0\theta^{\frac{p}{4}})^k \int |\omega^0|,$$

where in the last line, we have used (351). Then (347) follows by choosing first $p > |\gamma|$ and then $\theta^{\frac{1}{4}} \leq \frac{1}{2C_0}$ and then summing the geometric series over k.

STEP 3. As in the previous step we assume wlog that u, f, $u \diamond f$ are smooth (however, we do not assume that $u \diamond f = uf$). We set $\sup_{T \leq 1} \|[u, (\cdot)'_T] \diamond f\| (T^{\frac{1}{4}})^{-\bar\gamma} = N_1$ and as in the previous step $\sup_{T \leq 1} \|(f)'_T\|(T^{\frac{1}{4}})^{-\gamma} = N_0$. As in the proof of (347) we make use of the representation (353) of ψ_T to write

$$[u, (\cdot)_T] \diamond f = \sum_{k=0}^\infty [u, \omega_T^{(k)} * \omega_{\theta^k T}^0 * \psi'_{\theta^k T}*] \diamond f.$$

We apply the commutator relation $[A, BC] = [A, B]C + B[A, C]$ twice, to rewrite each term in this sum as

$$[u, \omega_T^{(k)} * \omega_{\theta^k T}^0 * \varphi_{\theta^k T}*] \diamond f$$

$$= [u, \omega_T^{(k)} * \omega_{\theta^k T}^0*](f)'_{\theta^k T} + \omega_T^{(k)} * \omega_{\theta^k T}^0 * [u, (\cdot)'_{\theta^k T}] \diamond f$$

$$= [u, \omega_T^{(k)}*](\omega_{\theta^k T}^0 * (f)'_{\theta^k T}) + \omega_T^{(k)} * ([u, \omega_{\theta^k T}^0*](f)'_{\theta^k T})$$

$$(354) \qquad + \omega_T^{(k)} * \omega_{\theta^k T}^0 * [u, (\cdot)'_{\theta^k T}] \diamond f.$$

We bound the terms on the rhs of (354) one by one, starting with the last. This expression can be directly bounded using Young's inequality

$$\|(\omega_T^{(k)} * \omega_{\theta^k T}^0) * [u, (\cdot)'_{\theta^k}] \diamond f\| \leq \int |\omega_T^{(k)}| \int |\omega_{\theta^k T}^0| \, (\theta^{\frac{k}{4}} T^{\frac{1}{4}})^{\bar{\gamma}} N_1$$
$$= \int |\omega^{(k)}| \int |\omega^0| \, (\theta^{\frac{k}{4}} T^{\frac{1}{4}})^{\bar{\gamma}} N_1.$$

Therefore, the sum in k over this term is controlled by invoking (351) for p large enough, then choosing θ small enough, resulting with a geometric series as above.

By Young's inequality, the second term on the rhs of (354) is bounded

$$\|\omega_T^{(k)} * ([u, \omega_{\theta^k T}^0 *](f)'_{\theta^k T})\| \leq \int |\omega^{(k)}| dx \, \|[u, \omega_{\theta^k T}^0 *](f)'_{\theta^k T}\|.$$

According to (351) the first factor on the rhs is bounded by $\lesssim (C_0 \theta^{\frac{p}{4}})^k$, while the second factor can be bounded as

$$\|[u, \omega_{\theta^k T}^0 *](f)'_{\theta^k T}\|$$
$$= \sup_x \left| \int (u(x) - u(y)) \omega_{\theta^k T}^0 (y - x)(f)'_{\theta^k T}(y) dy \right|$$
$$\leq [u]_\alpha N_0 (\theta^{\frac{k}{4}} T^{\frac{1}{4}})^\gamma \sup_x \int d^\alpha(x, y) |\omega_{\theta^k T}^0 (y - x)| dy$$
$$\leq [u]_\alpha N_0 (\theta^{\frac{k}{4}} T^{\frac{1}{4}})^\gamma ((\theta^k T)^{\frac{1}{4}})^\alpha \int d^\alpha(0, z) |\omega^0(z)| dz,$$

so that summing these terms over k also yields the required bound.

It remains to bound the first term on the rhs of (354) and for this we write

$$\|[u, \omega_T^{(k)} *](\omega_{\theta^k T}^0 * (f)'_{\theta^k T})\|$$
$$\leq \sup_x \int |u(y) - u(x)| \, |\omega_T^{(k)}(y - x)| dy \left(\int |\omega_{\theta^k T}^0| \right) \|(f)'_{\theta^k T}\|$$
$$\leq [u]_\alpha \sup_x \int d^\alpha(x, y) |\omega_T^{(k)}(y - x)| dy \left(\int |\omega^0| \right) N_0 (\theta^{\frac{k}{4}} T^{\frac{1}{4}})^\gamma$$
$$= [u]_\alpha (T^{\frac{1}{4}})^\alpha \int d^\alpha(0, z) |\omega^{(k)}(z)| dz \left(\int |\omega^0| \right) N_0 (\theta^{\frac{k}{4}} T^{\frac{1}{4}})^\gamma.$$

The first integral on the rhs is bounded $\lesssim (C_0 \theta^{\frac{p}{4}})^k$ in (352), so that finally (348) follows once more by choosing p large enough and θ small enough and summing over k.

STEP 4. It remains to give the argument for (349), (350) and (352) following [3]. The construction of ω^0 is based on the identity

$$A_{n,m} := \int x^n \partial^m \psi'(x) dx$$

$$= \begin{cases} 0 & \text{if } \|n\|_{\text{par}} \leq \|m\|_{\text{par}}, \, n \neq m \\ (-1)^{|m_1|+|m_2|} m_1! m_2! & \text{for } m = n \end{cases}.$$

This trigonal structure implies that for any fixed p the linear map $(a_m)_{\|m\|_{\text{par}}<p} \mapsto (\sum_{\|m\|_{\text{par}}<p} A_{n,m} a_m)_{\|n\|_{\text{par}}<p}$ is invertible. Furthermore, for each n, m the numbers $A^r_{n,m} := \int x^n \partial^m (\psi'_r * \psi')(x) dx$ converge to $A_{n,m}$ as $r \to 0$ and for $r > 0$ small enough the linear map associated to $(A^r_{n,m})_{\|n\|_{\text{par}}, \|m\|_{\text{par}} \leq p}$ is still invertible. This implies in particular the existence of coefficients (a_m) such that

$$\sum_{\|m\|_{\text{par}}<p} A^r_{n,m} a_m = \sum_{\|m\|_{\text{par}}<p} a_m \int x^n \partial^m (\psi'_r * \psi')(x) dx$$

$$= \begin{cases} 1 & \text{if } n = 0 \\ 0 & \text{else} \end{cases}.$$

The identity (349) thus follows for $\omega^0 = \sum_{\|m\|_{\text{par}}<p} a_m \partial^m \psi'_r$.

The key ingredient for the proof of (350) and (352) are the following estimates (355)–(358). We claim that for an arbitrary Schwartz function ω and any multi-index $m = (m_1, m_2)$ with $\|m\|_{\text{par}} \leq p+1$ we have for any $T > 0$

(355)

$$\int |\partial^m(\omega - \varphi'_T * \omega)| \leq C_0 \int |\partial^m \omega|,$$

$$\int d(0,x)^\alpha |\partial^m(\omega - \varphi'_T * \omega)| dx$$

(356)
$$\leq C_0 \left(\int d(0,x)^\alpha |\partial^m \omega| dx + (T^{\frac{1}{4}})^\alpha \int |\partial^m \omega| dx \right).$$

Furthermore, for $T \leq 1$

(357)
$$\int |\omega - \varphi'_T * \omega| \leq C_0 (T^{\frac{1}{4}})^p \sum_{\|m\|_{\text{par}}=p,p+1} \int |\partial^m \omega|$$

$$\int d(0,x)^\alpha |\omega - \varphi'_T * \omega|$$

(358) $$\leq C_0 (T^{\frac{1}{4}})^p \sum_{\|m\|_{\text{par}}=p,p+1} \left(\int d(0,x)^\alpha |\partial^m \omega| + (T^{\frac{1}{4}})^\alpha \int |\partial^m \omega| \right),$$

where we have $C_0 = C_0(\varphi')$ in (355) – (358). The estimates (357) and (358) rely on the Assumption (349) that φ' integrates to zero against

monomials of degree $0 < \|n\|_{\mathrm{par}} < p$. Once these bounds are established, the representation (350) follows if we define the $\omega^{(k)}$ recursively by

$$\omega^{(0)} = \psi \qquad \text{and } \omega^{(k+1)} = \omega^{(k)} - \varphi'_{\theta^k} * \omega^{(k)}$$

for a $\theta > 0$ small enough. Indeed, iterating (355) and (356) yields

$$\sum_{\|m\|_{\mathrm{par}}=p,p+1} \int (1 + d(0,x)^\alpha)|\partial^m \omega^{(k)}| dx$$

$$\leq (2C_0)^k \sum_{\|m\|_{\mathrm{par}}=p,p+1} \int (1 + d(0,x)^\alpha)|\partial^m \psi| dx,$$

which can then be plugged into (357) and (358) to yield

$$\int (1 + d(0,x)^\alpha)|\omega^{(k+1)}| dx$$

$$\leq (2C_0)^{k+1}(\theta^{\frac{k}{4}})^p \sum_{\|m\|_{\mathrm{par}}=p,p+1} \int (1 + d(0,x)^\alpha)|\partial^m \psi| dx,$$

which in turn yields (351) and (352). The representation then follows by observing

$$\psi = \omega^{(0)} = \omega^{(0)} * \varphi' + \omega^{(1)} = \omega^{(0)} * \varphi' + \omega^{(1)} * \varphi'_\theta + \omega^{(2)} = \ldots.$$

which together with (351) implies that the convergence holds in L^1.

The bounds (355) and (357) are provided in the discussion following Equation (295) in [3] (up to the parabolic scaling which can be included in the same way as in the following argument). Here we only present the proofs for (356) and (358) which follow along similar lines. First of all, in order to bound $\int d(0,x)^\alpha|\partial^m \omega - \varphi'_T * \partial^m \omega| dx$ we make use of the triangle inequality in the form $|\partial^m \omega - \varphi'_T * \partial^m \omega| \leq |\partial^m \omega| + |\varphi'_T * \partial^m \omega|$. The integral resulting from the first term then already has the desired form. For the second term, we write $|\varphi'_T * \partial^m \omega(x)| \leq \int |\varphi'_T(x-y)\partial^m \omega(y)| dy$ and use the triangle inequality once more, this time in the form $d(0,x)^\alpha \leq d(0,x-y)^\alpha + d(0,y)^\alpha$. Hence, it remains to bound the two integrals

$$\int\int d(0,x-y)^\alpha|\varphi'_T(x-y)| \, |\partial^m \omega(y)| dx dy$$

$$\leq \int d(0,z)^\alpha|\varphi'_T(z)| dz \int |\partial^m \omega(y)| dy,$$

$$= (T^{\frac{1}{4}})^\alpha \int d(0,\hat{z})^\alpha|\varphi'(\hat{z})| d\hat{z} \int |\partial^m \omega(y)| dy,$$

$$\int\int d(0,y)^\alpha |\varphi'_T(x-y)| \, |\partial^m \omega(y)| dx dy$$

$$\leq \int |\varphi'_T(x)| dx \int d(0,y)^\alpha |\partial^m \omega(y)| dy,$$

and estimate (356) follows.

To obtain (358), similar to [3] we obtain the pointwise bound

$$|\varphi'_T * \omega - \omega|(x)$$

$$(359) \quad \leq 2 \sum_{\|m\|_{\text{par}}=p,p+1} \int_0^1 \int d(0,z)^{\|m\|_{\text{par}}} |\varphi'_T(-z)| \, |\partial^m \omega(x+sz)| dz ds.$$

We recall their argument (adjusted to the case of parabolic scaling): First, according to (349) φ' integrates non-constant monomials of (parabolic) degree $< p$ to zero which permits us to write $(\varphi'_T * \omega - \omega)(x) = \int \left(\omega(x+z) - \sum_{\|m\|_{\text{par}}<p} \frac{1}{m_1! m_2!} \partial^m \omega(x) z^m \right) \varphi'_T(-z) dz$. At this point we seek to apply Taylor's formula, but unlike [3] we need an anisotropic version of the error term. In order to formulate this we define for $m = (m_1, m_2)$

$$T^m = \frac{\partial^m \omega(x) z^m}{(m_1+m_2)!} \qquad E^m = \int_0^1 \frac{(1-s)^{(m_1+m_2-1)}}{(m_1+m_2-1)!} z^m \partial^m \omega(x+sz) ds,$$

and observe the elementary identities $\omega(x+z) - \omega(x) = E^{(1,0)} + E^{(0,1)}$ as well as $E^m = T^m + E^{(m_1+1,m_2)} + E^{(m_1,m_2+1)}$ which permit to recursively obtain

$$\omega(x+z) - \sum_{\|m\|_{\text{par}}<p} \frac{1}{m_1! m_2!} \partial^m \omega(x) z^m$$

$$= \sum_{\|m\|_{\text{par}}=p} \binom{m_1+m_2}{m_1} E^{(m_1,m_2)} + \sum_{\|m\|_{\text{par}}=p-1} \binom{m_1+m_2}{m_1} E^{(m_1,m_2+1)}$$

$$\leq \sum_{\|m\|_{\text{par}}=p,p+1} \binom{m_1+m_2}{m_1} E^{(m_1,m_2)}.$$

Then bounding $|z^m| \leq d(0,z)^{\|m\|_{\text{par}}}$ and observing that the combinatorial pre-factor satisfies $\frac{1}{(m_1+m_2-1)!} \binom{m_1+m_2}{m_1} \leq 2$ and dropping $(1-s)^{m_1+m_2-1} \leq 1$ the claimed expression (359) follows.

To bound $\int d(0,x)^\alpha |\varphi'_T * \omega(x) - \omega(x)| dx$ we then use the triangle inequality in the form $d(0,x)^\alpha \leq d(0,sz)^\alpha + d(0,x+sz)^\alpha$ which prompts

to bound the two integrals

$$\int\int_0^1\int d(0,sz)^\alpha d(0,z)^{\|m\|_{\mathrm{par}}}|\varphi_T'(-z)|\,|\partial^m\omega(x+sz)|dzdsdx$$

$$\leq\left(\int d(0,z)^{\|m\|_{\mathrm{par}}}d(0,z)^\alpha|\varphi_T'(-z)|dz\right)\left(\int|\partial^m\omega(x)|dx\right),$$

$$\int\int_0^1\int d(0,x+sz)^\alpha d(0,z)^{\|m\|_{\mathrm{par}}}|\varphi_T'(-z)|\,|\partial^m\omega(x+sz)|dzdsdx$$

$$=\left(\int d(0,z)^{\|m\|_{\mathrm{par}}}|\varphi_T'(-z)|\right)\left(\int d(0,x)^\alpha|\partial^m\omega(x)|dx\right),$$

both of which are bounded as claimed in (358).

REFERENCES

[1] I. Bailleul and F. Bernicot. Heat semigroup and singular PDEs. *J. Funct. Anal.*, 270(9):3344–3452, 2016.

[2] M. Furlan and M. Gubinelli. Paracontrolled quasilinear SPDEs. *arXiv preprint arXiv:1610.07886*, 2016.

[3] A. Gloria and F. Otto. The corrector in stochastic homogenization: optimal rates, stochastic integrability, and fluctuations. *arXiv preprint arXiv:1510.08290*, 2015.

[4] M. Gubinelli. Controlling rough paths. *J. Funct. Anal.*, 216(1):86–140, 2004.

[5] M. Gubinelli, P. Imkeller, and N. Perkowski. Paracontrolled distributions and singular PDEs. *Forum Math. Pi*, 3:e6, 75, 2015.

[6] M. Hairer. Rough stochastic PDEs. *Comm. Pure Appl. Math.*, 64(11):1547–1585, 2011.

[7] M. Hairer. Solving the KPZ equation. *Ann. of Math. (2)*, 178(2):559–664, 2013.

[8] M. Hairer. A theory of regularity structures. *Invent. Math.*, 198(2):269–504, 2014.

[9] M. Hairer and C. Labbé. Multiplicative stochastic heat equations on the whole space. *arXiv preprint arXiv:1504.07162*, 2015.

[10] M. Hairer and C. Labbé. A simple construction of the continuum parabolic Anderson model on $\mathbf{R}^2$. *Electron. Commun. Probab.*, 20:no. 43, 11, 2015.

[11] M. Hairer and É. Pardoux. A Wong-Zakai theorem for stochastic PDEs. *J. Math. Soc. Japan*, 67(4):1551–1604, 2015.

[12] N. V. Krylov. *Lectures on elliptic and parabolic equations in Hölder spaces*, volume 12 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 1996.

[13] T. Lyons and Z. Qian. *System control and rough paths*. Oxford Mathematical Monographs. Oxford University Press, Oxford, 2002. Oxford Science Publications.

[14] T. J. Lyons. Differential equations driven by rough signals. *Rev. Mat. Iberoamericana*, 14(2):215–310, 1998.

[15] T. J. Lyons, M. Caruana, and T. Lévy. *Differential equations driven by rough paths*, volume 1908 of *Lecture Notes in Mathematics*. Lectures from the 34th Summer School on Probability Theory held in Saint-Flour, July 6–24, 2004.