Arch. Rational Mech. Anal. 232 (2019) 873–950
Digital Object Identifier (DOI) https://doi.org/10.1007/s00205-018-01335-z

Quasilinear SPDEs via Rough Paths

Felix Otto & Hendrik Weber

Communicated by L. Saint-Raymond

Abstract

We are interested in (uniformly) parabolic PDEs with a nonlinear dependence 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 which is 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. In fact, we develop a theory for the linear equation $\partial_t u - P(a\partial_1^2 u + \sigma f) = 0$ with rough but given coefficient fields a and σ and then apply a fixed point argument. The PDE ingredient mimics the (kernel-free) Safonov approach to ordinary Schauder theory.

Contents

1. Introduction

We are interested in the parabolic PDE

$$\partial_2 u - P\big(a(u)\partial_1^2 u + \sigma(u) f\big) = 0 \tag{1.1}$$

for a rough driver f. The non-linearities a, σ are assumed to be regular and uniformly elliptic, see (1.9) below for precise assumptions. In order to avoid difficulties related to initial and boundary values we adopt a more elliptic point of view and seek solutions which are periodic in *both* the space-like coordinate x_1 and the time-like coordinate x_2. This is the reason for the non-standard labelling of coordinates and the presence of the operator P-the projection onto 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 (2.5) 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 the products $v\partial_1^2 v$, vf (see (3.82)), where $v(\cdot, a_0)$ is the mean-free and space-time periodic solution to the constant coefficient equation

$$\partial_2 v(\cdot, a_0) - a_0 \partial_1^2 v(\cdot, a_0) = Pf \qquad \text{distributionally,} \tag{1.2}$$

and we show that these bounds allow us 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 [15–17]. This theory is based on the observation that the analysis of stochastic integrals

$$\int_0^t u(s)\, \mathrm{d}v(s) \tag{1.3}$$

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 (1.3) is defined for a single well-chosen function $\bar{u}$, for example v itself. In the case where $\bar{u} = v$ 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 [6, 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 $s, t \in \mathbb{R}$

$$|u(t) - u(s) - \sigma(s)(\bar{u}(t) - \bar{u}(s))| \lesssim |t - s|^{2\alpha}. \tag{1.4}$$

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 [6, Theorem 1] it is then shown that this assumption, together with an "off-line" bound of the form

$$\left| \int_s^t \bar{u}(r)\, \mathrm{d}v(r) - \bar{u}(s)(v(t) - v(s)) \right| \lesssim |t - s|^{2\alpha}, \tag{1.5}$$

suffices to define the integral $\int u(r)\, \mathrm{d}v(r)$ and to obtain the bound

$$\left| \int_s^t u(r)\, \mathrm{d}v(r) - u(s)(v(t) - v(s)) - \sigma(s) \int_s^t (\bar{u}(r) - \bar{u}(s))\, \mathrm{d}v(r) \right|$$
$$\lesssim |t - s|^{3\alpha}. \tag{1.6}$$

The construction of the integrals (1.5) 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 no use of the linear ordering 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 him to define stochastic PDEs with non-linearities of Burgers type [8] and the KPZ equation [9]. The notion of a controlled path was also the starting point for his definition of regularity structures [10] which permits treatment of semilinear stochastic PDE with an extremely irregular right hand side, possibly involving a renormalization procedure. Parallel to that, Gubinelli, Imkeller and Perkowski put forward a notion of paracontrolled distributions [7], a Fourier-analytic variant of (1.4) which has also been used to treat singular stochastic PDE.

In this article we propose yet another higher-dimensional generalization of the notion of controlled path, see Definition 3.1 below, and use it to provide a solution and stability theory for (1.1). This definition is an immediate generalization of Gubinelli's definition (1.4) and also closely related to Hairer's notion [10, Definition

3.1] of a modelled distribution in a certain regularity structure. However, the definition comes with a twist because the quasilinear nature of (1.1) forces us to allow the realization 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 (1.5) 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 (2.3) 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.[1] 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 (1.6) on the singular products $a(u)\diamond\partial_1^2 u$ and $\sigma(u)\diamond f$ (see Lemmas 3.3 and 3.5) which can also be seen as a variant of Hairer's Reconstruction Theorem [10, Theorem 3.10] in a simpler situation. 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 3.6, we obtain an optimal regularity result on solutions u of (1.1) 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 [10, Theorem 5.12]. Our proof mimics Safonov's approach to Schauder theory (as popularized in the monograph [14]) and therefore does not make reference to a parabolic heat kernel. These ingredients are combined in Proposition 3.8, to obtain a robust existence and uniqueness theory for the linear version of (1.1) (that is a and σ do not depend on u) including stability in the input data, and in Theorem 3.9 these results are used to develop a small data theory for the non-linear problem (1.1). 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 $\partial_{n+1} - \sum_{i,j=1}^n a^{ij}(u)\partial_i\partial_j$ 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 4 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 φ be an arbitrary Schwartz function with $\int \varphi = 1$ and for $\varepsilon \in (0, 1]$ set

[1] In the literature $\diamond$ is sometimes used to denote the Wick product of two random variables. Our products need not be Wick products.

$$\varphi_\varepsilon(x_1, x_2) := \frac{1}{\varepsilon^{\frac{3}{4}}}\varphi\Big(\frac{x_1}{\varepsilon^{\frac{1}{4}}}, \frac{x_2}{\varepsilon^{\frac{1}{2}}}\Big), \qquad f_\varepsilon := f * \varphi_\varepsilon, \qquad v_\varepsilon(\cdot, a_0) := v(\cdot, a_0) * \varphi_\varepsilon \tag{1.7}$$

and construct the $C^{\alpha-2}$ distributions $v \diamond f$ and $v \diamond \partial_1^2 v$ as

$$\begin{aligned} v(\cdot, a_0) \diamond f &:= \lim_{\varepsilon\to 0} \big(v_\varepsilon(\cdot, a_0) f_\varepsilon - \langle v_\varepsilon(\cdot, a_0) f_\varepsilon\rangle\big), \\ v(\cdot, a_0) \diamond \partial_1^2 v(\cdot, a_0') &:= \lim_{\varepsilon\to 0} \big(v_\varepsilon(\cdot, a_0)\partial_1^2 v_\varepsilon(\cdot, a_0') - \langle v_\varepsilon(\cdot, a_0)\partial_1^2 v_\varepsilon(\cdot, a_0')\rangle\big), \end{aligned} \tag{1.8}$$

where we use angled brackets $\langle\cdot\rangle$ for the expectation of a random variable, see Proposition 4.2 below. (We use the non-standard scaling in ε for consistency spatial scaling given by the convolution with Ψ_T, see Section 2 below.)

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

Theorem 1.1. *Let the non-linearities a, σ be smooth and uniformly elliptic in the sense that*

$$\begin{aligned} a \in [\lambda, \tfrac{1}{\lambda}], \quad \|a'\|, \|a''\|, \|a'''\| \leqq \tfrac{1}{\lambda}, \\ \sigma \in [-1, 1], \quad \|\sigma'\|, \|\sigma''\|, \|\sigma'''\| \leqq \tfrac{1}{\lambda}, \end{aligned} \tag{1.9}$$

where $\lambda > 0$ is some fixed constant and $\|\cdot\|$ denotes the supremum norm. Let f be a space-time periodic random Schwartz distribution, which is stationary, centered and Gaussian, and which satisfies the regularity assumption (4.2) for $\frac{2}{3} < \alpha' < 1$ and let α satisfy $\frac{2}{3} < \alpha < \alpha'$. Let f_ε be as in (1.7).

For any noise amplitude $\eta > 0$ we consider the following regularized and renormalized version of (1.1)

$$\begin{aligned} \partial_2 u_\varepsilon - P\big(a(u_\varepsilon)\partial_1^2 u_\varepsilon - a'(u_\varepsilon)\sigma^2(u_\varepsilon)\eta^2 g_2(\varepsilon, a(u_\varepsilon), a(u_\varepsilon)) \\ + \sigma(u_\varepsilon)\eta f_\varepsilon - \sigma'(u_\varepsilon)\sigma(u_\varepsilon)\eta^2 g_1(\varepsilon, a(u_\varepsilon))\big) = 0, \end{aligned} \tag{1.10}$$

where

$$\begin{aligned} g_1(\varepsilon, a_0) &:= \langle v_\varepsilon(\cdot, a_0) f_\varepsilon\rangle, \\ g_2(\varepsilon, a_0, a_0') &:= \langle v_\varepsilon(\cdot, a_0)\partial_1^2 v_\varepsilon(\cdot, a_0')\rangle, \end{aligned} \tag{1.11}$$

and where $v_\varepsilon(\cdot, a_0)$ is defined in (1.7).

There exists a random $\eta_0 > 0$ and a deterministic constant $\delta = \delta(\lambda, \alpha) \in (0, 1]$ such that almost surely for any $\eta \leqq \eta_0$ and for any $0 < \varepsilon \leqq 1$ there exists a unique space-time periodic smooth random function u_ε which satisfies (1.10) and which is small in the sense $[u_\varepsilon]_\alpha \leqq \delta$, where $[u_\varepsilon]_\alpha$ refers to the parabolic Hölder semi-norm, defined in (2.2). Furthermore η is not too small in the sense that

$$\langle \eta_0^{-p}\rangle^{\frac{1}{p}} < \infty \qquad \text{for all } p < \infty. \tag{1.12}$$

Almost surely, for any fixed $\eta \leqq \eta_0$ the solutions u_ε converge as $\varepsilon \downarrow 0$ to a limit u. This convergence takes place uniformly and with respect to $[\cdot]_\alpha$. The limit u does not depend on the choice of mollifying kernel φ although g_1 and g_2 do.

The small amplitude η appears here because of our choice to work with space-time periodic solutions rather than treating the initial value problem (space-time periodic here means that functions/distributions are periodic of fixed period which without loss of generality we set to be 1, both in the space-like coordinate x_1 and the time-like coordinate x_2). In initial value problems it is common to show the "local" existence and uniqueness of solutions, that is the existence and uniqueness on some small time interval (the length of which is random if there are random terms in the equation). The small amplitude η plays the role of this small time interval here. The smallness assumption $[u_\varepsilon]_\alpha \leqq \delta$ also appears because of the periodic space-time boundary conditions and is needed to ensure uniqueness of solutions. The following theorem gives a characterization of the limit u obtained in Theorem 1.1:

Theorem 1.2. *Under the assumptions of Theorem 1.1, u is almost surely the unique mean-free space-time periodic function with the properties*

$$u \text{ is modelled after } v \text{ according to } a(u) \text{ and } \sigma(u) \text{ (in the sense of Definition 3.1)}, \tag{1.13}$$

$$\partial_2 u - P\big[a(u) \diamond \partial_1^2 u + \sigma(u) \diamond \eta f\big] = 0 \quad \text{distributionally}, \tag{1.14}$$

satisfying

$$[u]_\alpha \leqq \delta. \tag{1.15}$$

We stress that the definition of the non-standard products $a(u)\diamond\partial_1^2 u$ and $\sigma(u)\diamond\eta f$ in (1.14) (see Corollary 3.7 and Lemma 3.5) relies on the "modelledness" of u as well as the definition of the renormalized products (1.8).

We finally mention that shortly before posting the second version of our result, the article [4] was posted on the arXiv. In this article Furlan and Gubinelli study the equation

$$\partial_t u - a(u)\Delta u = \xi, \tag{1.16}$$

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 (1.10). Following the approach we present here, they locally approximate the solutions u by a family of solutions to constant coefficient problems. Their approach then proceeds in the framework of paracontrolled distributions. Yet another approach by BAILLEUL et al. [2] was put forward shortly after posting our second version. They deal with the system

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

where ξ is again a two dimensional white noise and they also obtain a short-time existence and stability result for renormalized solutions. Their method is easier than ours or Furlan and Gubinelli's as they only need a single random function, namely $X = (-\Delta)^{-1}\xi$ to locally describe u. However, this makes strong use of the fact that the noise only depends on the space variable and it would also not work if the operator $a(u)\Delta$ were replaced by the more general uniformly elliptic operator $a_{ij}(u)\partial_i\partial_j u$.

2. Setup

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

$$d(x, y) := |x_1 - y_1| + \sqrt{|x_2 - y_2|}, \tag{2.1}$$

see for instance [14, Section 8.5]. The Hölder semi-norm $[\cdot]_\alpha$ is defined based on (2.1):

$$[u]_\alpha := \sup_{x \neq y} \frac{|u(x) - u(y)|}{d^\alpha(x, y)}. \tag{2.2}$$

In order to define negative norms of distributions in an intrinsic way, cf. (2.5) below, 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 (2.1). It will turn out to be extremely convenient to have in addition the semi-group property

$$(\cdot)_T \circ (\cdot)_t = (\cdot)_{T+t}. \tag{2.3}$$

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 and 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 (2.1) 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:

$$\int dy |\partial_1^k \partial_2^\ell \psi_T(x - y)| d^\alpha(x, y) \leqq C(k, \ell, \alpha)(T^{\frac{1}{4}})^{-k-2\ell+\alpha}, \tag{2.4}$$

for all orders of derivative $k, \ell = 0, 1, \ldots$ and moment exponents $\alpha \geqq 0$, as well as the fact that $\int \psi(x)x_1 \, dx = 0$. Estimates (2.4) follow immediately from the scaling and the fact that ψ_1 is a Schwartz function. In Lemma A.3 we show however that our main regularity assumption (2.5) 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. We will typically measure the size of the distribution f by the expression

$$\|f\|_{\alpha-2} := \sup_{T \leqq 1} (T^{\frac{1}{4}})^{2-\alpha} \|f_T\|, \tag{2.5}$$

where the restriction $T \leqq 1$ reflects the period unity. By Lemma A.1, cf. Step 1, this expression agrees with the standard definition of the norm of $C^{\alpha-2}$.

Here and throughout the entire deterministic analysis presented in Sections 3, 5 and Appendix 8 $\lesssim$ means $\leqq C$ with a constant C only depending on λ and the exponent α. In the derivation of the stochastic bounds in Sections 4 and 6 the implicit constant may depend on additional parameters which are specified there. Similarly, we write $\ll 1$ for $\leqq \delta$ for $\delta = \delta(\lambda, \alpha)$ small enough.

3. Deterministic Analysis

We start with the following central definition which is a straightforward generalization of Gubinelli's definition [6, Definition 1] of a "controlled path", a generalization from the time variable x_2 to multiple variables x, and to a "model" $(v_1, \ldots, v_I)$ (in the language of Hairer [10]) 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 with respect to v_i. The increments of the linear function x_1 also have to be included because of $\alpha > \frac{1}{2}$. In fact, since $2\alpha > 1$, given the model $(v_1, \ldots, v_I)$ (as modulated by the functions a_i), the "derivatives" $(\sigma_1, \ldots, \sigma_I)$ and ν determine u up to a constant. In our situation, we expect u and $(v_1, \ldots, v_I)$ 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 3.1. Let $\frac{1}{2} < \alpha < 1$ and $I \in \mathbb{N}$. We say that a function u is modelled after the functions $(v_1, \ldots, v_I)$ of (x, a_0) according to the functions $(a_1, \ldots, a_I)$ and $(\sigma_1, \ldots, \sigma_I)$ provided there exists a function ν (which because of $2\alpha > 1$ is easily seen to be unique) such that

$$M := \sup_{x \neq y} \frac{1}{d^{2\alpha}(y, x)} |u(y) - u(x) - \sigma_i(x)(v_i(y, a_i(x)) - v_i(x, a_i(x))) - \nu(x)(y - x)_1| \tag{3.1}$$

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

Note that imposing (3.1) also for distant points x and y is consistent with periodicity despite the non-periodic term $(y - x)_1$ since by $\alpha \geqq \frac{1}{2}$ the latter is dominated by $d^{2\alpha}(x, y)$ for $d(x, y) \geqq 1$. Note also that (3.1) is reminiscent of a Hölder norm: In case of $(\sigma_1, \ldots, \sigma_I) = 0$, the finiteness of (3.1) 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 3.3 shows that the modelledness constant M in (3.1) 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 3.3.

The following lemma shows that the notion of modelledness in Definition 3.1 is well-behaved under sufficiently smooth nonlinear *pointwise* transformation; it will be used in the proof of Theorem 3.9 (it is essentially identical to [6, Proposition 4], which in turn is a consequence of Taylor's formula; and we omit the proof):

Lemma 3.2. *Let $\frac{1}{2} < \alpha < 1$. Then*
i) Suppose that $u \in C^\alpha$ 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

$$\tilde{M} + [b(u)]_\alpha \leqq (\|b'\| + \|b''\|[u]_\alpha)(M + [u]_\alpha), \tag{3.2}$$

$$[\mu]_\alpha + \|\mu\| \leqq (\|b'\| + \|b''\|[u]_\alpha)([\sigma]_\alpha + \|\sigma\|). \tag{3.3}$$

ii) Suppose that for $i = 0, 1$, $u_i \in C^\alpha$ 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

$$\begin{aligned}
&\delta\tilde{M} + [b(u_1) - b(u_0)]_\alpha + \|b(u_1) - b(u_0)\| \\
&\quad \leqq \Big(\|b'\| + \|b''\|(\max_i M_i + \max_i [u_i]_\alpha) + \|b'''\|(\max_i [u_i]_\alpha)^2\Big) \\
&\qquad \times (\delta M + [u_1 - u_0]_\alpha + \|u_1 - u_0\|),
\end{aligned} \tag{3.4}$$

$$\begin{aligned}
&[\mu_1 - \mu_0]_\alpha + \|\mu_1 - \mu_0\| \\
&\quad \leqq (\|b'\| + \|b''\| \max_i [u_i]_\alpha)([\sigma_1 - \sigma_0]_\alpha + \|\sigma_1 - \sigma_0\|) \\
&\qquad + (\|b''\| + \|b'''\| \max_i [u_i]_\alpha) \max_i ([\sigma_i]_\alpha + \|\sigma_i\|)([u_1 - u_0]_\alpha + \|u_1 - u_0\|).
\end{aligned} \tag{3.5}$$

As discussed in the introduction, the main challenge in solving stochastic ordinary differential equations is to give a sense to integrals of the form (1.3). In the spirit of Hairer [10] 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 in a Brownian fashion. 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 (1.5), which can be rewritten as $\int_s^t (\bar{u}(r) - \bar{u}(s)) \diamond \partial_r v(r) = -\bar{u}(s) \int_s^t \partial_r v(r) - \int_s^t \bar{u} \diamond \partial_r v =: -([\bar{u}, \int^t] \diamond \partial_r v)(s)$, that is, the expression on both sides of (1.5) 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

$$[v, (\cdot)_T] \diamond f := v f_T - (v \diamond f)_T. \tag{3.6}$$

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. In our set up, the role of the crucial "algebraic relationship" [6, (24)] from rough path theory is played by the following straightforward consequence of the semi-group property (2.3)

$$[v, (\cdot)_{t+T}] \diamond f - ([v, (\cdot)_T] \diamond f)_t = [v, (\cdot)_t] f_T, \tag{3.7}$$

cf. (5.115) in the proof of Lemma 3.3. We also stress that the bound of order $(T^{\frac{1}{4}})^{2\alpha-2}$ on the commutator (3.6) we impose below, is equivalent to the condition on the "model" imposed in [13] in the framework of regularity structures. In fact, there in [13, Equation (3.9)] the condition (in their notation)

$$|(\Pi_z)(\tau)(\varphi_z^\lambda)| \lesssim \lambda^{|\tau|}$$

is assumed for all "stochastic basis elements" τ. Specialized to $\tau = \mathcal{I}(\Xi)\Xi$ (still in their notation) and following the definition of the "canonical admissible model" (see [13, Section 3.3]) this condition translates to our notation as

$$\|[v, *\varphi^\lambda] \diamond f\| = \sup_{x_0} \left| \int (v \diamond f(x) - v(x_0) f(x)) \varphi^\lambda(x - x_0)\, \mathrm{d}x \right| \lesssim \lambda^{|\tau|},$$

where φ^λ is a (parabolically) scaled test-function. This only differs from our assumption in our specific choice of regularising kernel ψ.

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 v defined by (1.2). 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 3.3 and Corollary 3.4 dealing with the first and Lemma 3.5 with the second factor in the singular products. We will use Corollary 3.4, 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 3.5 to pass from $b \diamond \partial_1^2 v$ to $b \diamond \partial_1^2 u$.

These upcoming statements 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 (3.37) that rely on the periodicity or the constraint $T \leqq 1$ 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 $\|f\|_{\alpha-2}$ (defined in (2.5));
- commutators $[u, (\cdot)_T]\diamond f$ are measured on level $2\alpha - 2 < 0$ via

$$\|[u, (\cdot)]\diamond f\|_{2\alpha-2} := \sup_{T\leqq 1} (T^{\frac{1}{4}})^{2-2\alpha} \|[u, (\cdot)_T]\diamond f\|; \tag{3.8}$$

- differences $[u, (\cdot)_T]\diamond f - [v, (\cdot)_T]\diamond f$ of commutators, like in case of the rough path expression (1.6) divided by $(t - s)$, are measured on level $3\alpha - 2 > 0$ via

$$\|[u, (\cdot)]\diamond f - [v, (\cdot)]\diamond f\|_{3\alpha-2} := \sup_{T\leqq 1} (T^{\frac{1}{4}})^{2-3\alpha} \|[u, (\cdot)_T]\diamond f - [v, (\cdot)_T]\diamond f\|, \tag{3.9}$$

see (3.14) of Lemma 3.3.

Equipped with this dictionary, Corollary 3.4 and Lemma 3.5 can be seen to be very close to [6, Theorem 1]; in particular, (3.14) in Lemma 3.3 is very close to (28) in [6, Corollary 3]. The major difference is the multi-dimensional extension through (3.6). 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 A.2. 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, these minor differences can be embedded into the more general form of the upcoming Lemma, 3.3.

Lemma 3.3. *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*

$$[v(\cdot, x) - v(\cdot, x')]_\alpha \leqq N d^\alpha(x, x'), \tag{3.10}$$

$$\|f\|_{\alpha-2} \leqq N_1, \tag{3.11}$$

$$\|[v(\cdot, x), (\cdot)]\diamond f - [v(\cdot, x'), (\cdot)]\diamond f\|_{2\alpha-2} \leqq N N_1 d^\alpha(x, x') \tag{3.12}$$

for all pairs of points x, x' and for some constants N, N_1.[2] *Suppose we are given a function u such that*

$$|(u(y) - u(x)) - (v(y, x) - v(x, x)) - \nu(x)(y - x)_1| \leqq M d^{2\alpha}(y, x) \tag{3.13}$$

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

$$\|[u, (\cdot)]\diamond f - E_{\text{diag}}[v, (\cdot)]\diamond f - \nu[x_1, (\cdot)]f\|_{3\alpha-2} \lesssim (M + N)N_1, \tag{3.14}$$

[2] In (3.12) the $2\alpha - 2$ semi-norm of the difference of commutators is defined as (3.9) with $3\alpha - 2$ replaced by $2\alpha - 2$.

where E_{diag} stands for the evaluation of the continuous function $(x,y) \mapsto ([v(\cdot,x),(\cdot)_T] \diamond f)(y)$ on the diagonal $y = x$.

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

$$[v(\cdot,x)]_\alpha \leqq N, \tag{3.15}$$

$$\|[v(\cdot,x),(\cdot)] \diamond f\|_{2\alpha-2} \leqq NN_1 \tag{3.16}$$

for all points x, then we also have that

$$\|[u,(\cdot)] \diamond f\|_{2\alpha-2} \leqq (M+N)N_1. \tag{3.17}$$

Equipped with Lemma 3.3, the upcoming corollary specifies the form of the model. The general form of Lemma 3.3 is in particular convenient for part iii), 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.

To shorten some of the formulas, from now on we add some more indices to the (semi-) norms referring to parameter derivatives with respect to a_0 and a_0': If $|\cdot|$ is a semi-norm and if u depends on a parameter a_0 we write

$$|u|_n := \sup_{a_0 \in [\lambda, \frac{1}{\lambda}]} \max_{i=0,\ldots,n} \left| \left(\frac{\partial}{\partial a_0}\right)^i u(a_0) \right|, \tag{3.18}$$

and if u depends on two parameters a_0 and a_0' we write

$$|u|_{n,m} := \sup_{a_0, a_0' \in [\lambda, \frac{1}{\lambda}]} \max_{i=0,\ldots,n} \max_{j=0,\ldots,m} \left| \left(\frac{\partial}{\partial a_0}\right)^i \left(\frac{\partial}{\partial a_0}\right)^j u(a_0, a_0') \right|. \tag{3.19}$$

Corollary 3.4. *i) Let $\{v(\cdot,a_0)\}_{a_0}$ be a family of functions and let $\{f(\cdot,a_0')\}_{a_0'}$, $\{v(\cdot,a_0) \diamond f(\cdot,a_0')\}_{a_0,a_0'}$ be two families of distributions satisfying*

$$[v]_{\alpha,2} \leqq N_0, \tag{3.20}$$

$$\|f\|_{\alpha-2,2} \leqq N_1, \tag{3.21}$$

$$\|[v,(\cdot)] \diamond f\|_{2\alpha-2,1,2} \leqq N_1 N_0 \tag{3.22}$$

for some constants N_0 and N_1. If u is modelled after v according to the α-Hölder functions a and σ with constant M and ν as in (3.1), then there exists a unique family of distributions $\{u \diamond f\}_{a_0,a_0'}$ such that

$$\lim_{T \downarrow 0} \left\| [u,(\cdot)_T] \diamond f - \sigma E[v,(\cdot)_T] \diamond f - \nu [x_1,(\cdot)_T] f \right\| = 0, \tag{3.23}$$

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

$$[\sigma]_\alpha \leqq 1, \ [a]_\alpha \leqq 1 \quad \textit{and} \quad \|\sigma\| \leqq 1, \tag{3.24}$$

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

$$\left\| [u,(\cdot)] \diamond f \right\|_{2\alpha-2,2} \lesssim N_1(M+N_0). \tag{3.25}$$

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,*

$$\|f_1 - f_0\|_{\alpha-2,1} \leqq \delta N_1, \tag{3.26}$$

$$\|[v, (\cdot)] \diamond f_1 - [v, (\cdot)] \diamond f_0\|_{2\alpha-2,1,1} \leqq \delta N_1 N_0 \tag{3.27}$$

for some constant δN_1. *Then for* u *and* $u \diamond f_i$ *as in i) we have*

$$\|[u, (\cdot)] \diamond f_1 - [u, (\cdot)] \diamond f_0\|_{2\alpha-2,1} \lesssim \delta N_1 (M + N_0). \tag{3.28}$$

iii) Let the two families of functions $\{v_i(\cdot, a_0)\}_{a_0}$, $i = 0, 1$, *and the three families of distributions* $\{f(\cdot, a_0')\}_{a_0'}$, $\{v_i(\cdot, a_0) \diamond f(\cdot, a_0')\}_{a_0, a_0'}$ *satisfy* (3.21) *and, in addition,*

$$[v_i]_{\alpha,2} \leqq N_0, \tag{3.29}$$

$$[v_1 - v_0]_{\alpha,1} \leqq \delta N_0, \tag{3.30}$$

$$\|[v_i, (\cdot)] \diamond f\|_{2\alpha-2,2,1} \leqq N_1 N_0, \tag{3.31}$$

$$\|[v_1, (\cdot)] \diamond f - [v_0, (\cdot)] \diamond f\|_{2\alpha-2,1,1} \leqq N_1 \delta N_0. \tag{3.32}$$

Let u_i *be two functions like in part i) and let* $u_i \diamond f$ *be as constructed there. 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*

$$\|[u_1, (\cdot)] \diamond f - [u_0, (\cdot)] \diamond f\|_{2\alpha-2,1} \lesssim N_1 \big(\delta M + N_0([\sigma_1 - \sigma_0]_\alpha + \|\sigma_1 - \sigma_0\| + [a_1 - a_0]_\alpha + \|a_1 - a_0\|) + \delta N_0\big). \tag{3.33}$$

We now turn to Lemma 3.5 that deals with the second factor in $a \diamond \partial_1^2 u$. The reason why we consider several functions $v_1, \ldots, v_I$ in Lemma 3.5 instead of a single one for our scalar PDE is that this seems necessary when establishing the contraction property for Proposition 3.8; because of the a_0-dependence, it turns out that we need not just $I = 2$ but in fact $I = 3$, cf. Corollary 3.7.

Lemma 3.5. *Let* $\frac{2}{3} < \alpha < 1$ *and* $I \in \mathbb{N}$. *We are given a function* b, I *families of functions* $\{v_1(\cdot, a_0), \ldots, v_I(\cdot, a_0)\}_{a_0}$, *and* I *families of distributions* $\{b \diamond \partial_1^2 v_1(\cdot, a_0), \ldots, b \diamond \partial_1^2 v_I(\cdot, a_0)\}_{a_0}$ *with*

$$[v_i]_{\alpha,1} \leqq N_i, \tag{3.34}$$

$$\|[b, (\cdot)] \diamond \partial_1^2 v_i\|_{2\alpha-2,1} \leqq N_0 N_i \tag{3.35}$$

for some constants $N_0, \ldots, N_I$. *Let the function* u *be modelled after* $(v_1, \ldots, v_I)$ *according to the* α*-Hölder functions* a *and* $(\sigma_1, \ldots, \sigma_I)$ *with constant* M, *cf. Definition* 3.1. *Then there exists a unique distribution* $b \diamond \partial_1^2 u$ *such that*

$$\lim_{T \downarrow 0} \|[b, (\cdot)_T] \diamond \partial_1^2 u - \sigma_i E[b, (\cdot)_T] \diamond \partial_1^2 v_i\| = 0, \tag{3.36}$$

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

$$\|[b, (\cdot)] \diamond \partial_1^2 u\|_{2\alpha-2} \lesssim [b]_\alpha M + N_0 N_i ([\sigma_i]_\alpha + \|\sigma_i\|). \tag{3.37}$$

The next 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 (3.39), 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 right hand side, we will right away assume that the left hand side is estimated accordingly, cf. (3.40). 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 equation 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 Safonov, [14], in particular Section 8.6. This approach avoids the use of kernels.

Lemma 3.6. *Let $\frac{1}{2} < \alpha < 1$ and suppose all functions and distributions are periodic. We are given I families of distributions $\{f_1(\cdot, a_0), \ldots, f_I(\cdot, a_0)\}_{a_0}$ with*

$$\|f_i\|_{\alpha-2,1} \leqq N_i \tag{3.38}$$

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

$$(\partial_2 - a_0\partial_1^2) v_i(\cdot, a_0) = P f_i(\cdot, a_0) \quad \textit{distributionally}. \tag{3.39}$$

We are also given a function u, modelled after $(v_1, \ldots, v_I)$ according to some functions $a \in [\lambda, \frac{1}{\lambda}]$ and $(\sigma_1, \ldots, \sigma_I)$. We assume that

$$\sup_{T \leqq 1} (T^{\frac{1}{4}})^{2-2\alpha} \|\partial_2 u_T - P(a\partial_1^2 u_T + \sigma_i E f_{iT})\| \leqq N^2 \tag{3.40}$$

for some constant N, where E is defined in Lemma 3.5. Then we have for the modelling and the Hölder constant of u

$$M \lesssim N^2 + [a]_\alpha M + N_i([\sigma_i]_\alpha + \|\sigma_i\|[a]_\alpha), \tag{3.41}$$

$$[u]_\alpha \lesssim M + N_i\|\sigma_i\|. \tag{3.42}$$

In the upcoming Corollary 3.7, we combine Corollary 3.4 on the product $\sigma \diamond f$, Lemma 3.5 on the product $a \diamond \partial_1^2 u$ and Lemma 3.6 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 3.8.

Corollary 3.7. *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

$$[\sigma]_\alpha + [a]_\alpha \leqq N, \tag{3.43}$$

$$\|f\|_{\alpha-2} \leqq N_0, \tag{3.44}$$
$$\|[\sigma,(\cdot)]\diamond f\|_{2\alpha-2} \leqq NN_0, \tag{3.45}$$
$$\|[a,(\cdot)]\diamond \partial_1^2 v\|_{2\alpha-2,2} \leqq NN_0, \tag{3.46}$$

for some constants N_0 and N, where $v(\cdot, a_0)$ denotes the mean-free solution of (1.2), and satisfying the constraints

$$\sigma \in [-1,1],\ a \in \left[\lambda, \frac{1}{\lambda}\right],\ [\sigma]_\alpha \leqq 1,\ [a]_\alpha \ll 1. \tag{3.47}$$

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

$$\partial_2 u - P(a\diamond \partial_1^2 u + \sigma \diamond f) = 0 \quad \text{distributionally}, \tag{3.48}$$

we have for the modelling and Hölder constants that

$$M \lesssim N_0 N, \tag{3.49}$$
$$[u]_\alpha \lesssim N_0(N+1). \tag{3.50}$$

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

$$[\delta\sigma]_\alpha + \|\delta\sigma\| + [\delta a]_\alpha + \|\delta a\| \leqq \delta N, \tag{3.51}$$
$$\|\delta f\|_{\alpha-2} \leqq \delta N_0, \tag{3.52}$$
$$\|[\sigma,(\cdot)]\diamond \delta f\|_{2\alpha-2} \leqq N\delta N_0, \tag{3.53}$$
$$\|[\delta\sigma,(\cdot)]\diamond f\|_{2\alpha-2} \leqq \delta N N_0, \tag{3.54}$$
$$\|[a,(\cdot)]\diamond \partial_1^2 \delta v\|_{2\alpha-2,1} \leqq N\delta N_0, \tag{3.55}$$
$$\|[\delta a,(\cdot)]\diamond \partial_1^2 v\|_{2\alpha-2,1} \leqq \delta N N_0 \tag{3.56}$$

for some constants δN_0, δN and where $\delta v(\cdot, a_0)$ is the mean-free solution of

$$(\partial_2 - a_0\partial_1^2)\delta v(\cdot, a_0) = P\delta f \quad \text{distributionally}. \tag{3.57}$$

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

$$\partial_2 \delta u - P(a\diamond \partial_1^2 \delta u + \delta a \diamond \partial_1^2 u + \sigma \diamond \delta f + \delta\sigma \diamond f) = 0 \tag{3.58}$$

then we have for the modelling and Hölder constants that

$$\delta M \lesssim N_0 \delta N + \delta N_0 N \quad \text{provided } N \leqq 1, \tag{3.59}$$
$$[\delta u]_\alpha \lesssim N_0\delta N + \delta N_0 \quad \text{provided } N \leqq 1. \tag{3.60}$$

The next Proposition, 3.8, 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 given coefficients σ and a. Because of the roughness of f, it does not only require a definition of $\sigma \diamond f$ but also of $a \diamond \partial_1^2 v$, 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 3.5. 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

$$\begin{aligned} &u_1 - u_0 \text{ is modelled after } (v_1, v_0) \\ &\text{according to } (a_1, a_0) \text{ and } (\sigma_1, -\sigma_0). \end{aligned} \tag{3.61}$$

When quantifying differences of solutions, variable coefficients require a somewhat different strategy compared to constant coefficients, as we shall explain now. The modelledness (3.61) has to come from the PDE, that is, Lemma 3.6. 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:

$$\begin{aligned} &\partial_2(u_1 - u_0) - P(a_0 \diamond \partial_1^2 u_1 - a_0 \diamond \partial_1^2 u_0) \\ &\quad = P\big(\sigma_1 \diamond f_1 - \sigma_0 \diamond f_0 + (a_1 \diamond \partial_1^2 u_1 - a_0 \diamond \partial_1^2 u_1)\big), \end{aligned} \tag{3.62}$$

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 3.6, we may replace u_1 by v_1 on the right hand side of (3.62), leading to

$$\begin{aligned} &\partial_2(u_1 - u_0) - P(a_0 \diamond \partial_1^2 u_1 - a_0 \diamond \partial_1^2 u_0) \\ &\quad \approx P\big(\sigma_1 \diamond f_1 - \sigma_0 \diamond f_0 + \sigma_1 (E_1 a_1 \diamond \partial_1^2 v_1 - E_1 a_0 \diamond \partial_1^2 v_1)\big). \end{aligned} \tag{3.63}$$

In view of Lemma 3.6 and the discussion preceding it, this suggests that we obtain

$$\begin{aligned} &u_1 - u_0 \text{ is modelled after } (v_1, v_0, (\partial_2 - a_0 \partial_1^2)^{-1} P E_1 \partial_1^2 v) \\ &\quad \text{according to } a_0 \text{ and } (\sigma_1, -\sigma_0, \sigma_1(a_1 - a_0)), \end{aligned} \tag{3.64}$$

which is *not* the desired (3.61) 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

$$\partial_2 u_s - P(a_s \diamond \partial_1^2 u_s + \sigma_s \diamond f_s) = 0. \tag{3.65}$$

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

$$\sigma_s \diamond f_s := s^2 \sigma_1 \diamond f_1 + s(1 - s)(\sigma_1 \diamond f_0 + \sigma_0 \diamond f_1) + (1 - s)^2 \sigma_0 \diamond f_0, \tag{3.66}$$

and use the same definition for $a_s \diamond \partial_1^2 v_s$, Leibniz's 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 (3.65) gives

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

where we write $\partial_s u$ as short hand for $\partial_s u_s$ and the same for a, σ and f. In view of (3.65) we approximate the right hand side 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 part ii) of Corollary 3.7. Noting that $(\partial_2 - a_0 \partial_1^2) \frac{\partial v_s}{\partial a_0} = \partial_1^2 v_s$ we obtain

$$\begin{aligned} &\partial_s u \text{ is modelled after } \left(v_s, \frac{\partial v_s}{\partial a_0}, \partial_s v \right) \\ &\text{according to } a_s \text{ and } (\partial_s \sigma, \sigma_s \partial_s a, \sigma_s), \end{aligned} \tag{3.67}$$

which compares favorably to (3.64).[3] Using Leibniz's rule once more, but this time in the classical form of

$$\begin{aligned} \frac{\partial}{\partial s} \big(\sigma_s(x) v_s(y, a_s(x)) \big) &= (\partial_s \sigma)(x) v_s(y, a_s(x)) \\ &+ (\sigma_s \partial_s a)(x) \frac{\partial v_s}{\partial a_0}(y, a_s(x)) + \sigma_s(x) \partial_s v(y, a_s(x)), \end{aligned}$$

and integrating (3.67) in $s \in [0, 1]$, yields the desired (3.61). We note that this strategy differs from [6] 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 [6, Proposition 4] uses the linear interpolation $u_s = s u_1 + (1-s) u_0$ (as we do in Lemma 3.2), 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 (3.66).

Proposition 3.8. *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 $\{a \diamond \partial_1^2 v(\cdot, a_0)\}_{a_0}$ satisfying (3.43) –(3.47)*. Then there exists a unique mean-free function u modelled after v according to a and σ and such that*

$$\partial_2 u - P(a \diamond \partial_1^2 u + \sigma \diamond f) = 0. \tag{3.68}$$

The modelling and Hölder constants are estimated as follows:

$$M \lesssim N_0 N, \tag{3.69}$$

$$[u]_\alpha \lesssim N_0 (N+1). \tag{3.70}$$

ii) Suppose we are given functions σ_i and a_i, $i = 0, 1$, distributions f_i and $\sigma_i \diamond f_j$, $j = 0, 1$, and 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 (1.2) *corresponding to f_i, satisfying the assumption* (3.43)– (3.46) *with cross terms, that is,*

$$\| [\sigma_i, (\cdot)] \diamond f_j \|_{2\alpha - 2} \leqq N_0 N, \tag{3.71}$$

[3] Here we use the symbol a_0 with two different meanings: as the concrete coefficient field a_0 and as an abstract parameter in v_s. It will always be clear from the context which of these interpretations is meant.

$$\|[a_i, (\cdot)] \diamond \partial_1^2 v_j\|_{2\alpha-2,2} \leqq N_0 N \tag{3.72}$$

and (3.47)*. We measure the distance of* (f_1, σ_1, a_1) *to* (f_0, σ_0, a_0) *in terms of the constants* δN_0 *and* δN *with*

$$[\sigma_1 - \sigma_0]_\alpha + \|\sigma_1 - \sigma_0\| + [a_1 - a_0]_\alpha + \|a_1 - a_0\| \leqq \delta N, \tag{3.73}$$

$$\|f_1 - f_0\|_{\alpha-2} \leqq \delta N_0, \tag{3.74}$$

$$\|[\sigma_i, (\cdot)] \diamond f_1 - [\sigma_i, (\cdot)] \diamond f_0\|_{2\alpha-2} \leqq N \delta N_0, \tag{3.75}$$

$$\|[\sigma_1, (\cdot)] \diamond f_j - [\sigma_0, (\cdot)] \diamond f_j\|_{2\alpha-2} \leqq \delta N N_0, \tag{3.76}$$

$$\|[a_i, (\cdot)] \diamond \partial_1^2 v_1 - [a_i, (\cdot)] \diamond \partial_1^2 v_0\|_{2\alpha-2,1} \leqq N \delta N_0, \tag{3.77}$$

$$\|[a_1, (\cdot)] \diamond \partial_1^2 v_j - [a_0, (\cdot)] \diamond \partial_1^2 v_j\|_{2\alpha-2,1} \leqq \delta N N_0. \tag{3.78}$$

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:*

$$\delta M \lesssim N_0 \delta N + \delta N_0 N, \tag{3.79}$$

$$[u_1 - u_0]_\alpha + \|u_1 - u_0\| \lesssim N_0 \delta N + \delta N_0 \quad \textit{both provided } N \leqq 1. \tag{3.80}$$

We now proceed to Theorem 3.9, the main deterministic result of this paper; it can be seen as a PDE version of the ODE result in [6, Section 5]. Part i) of the theorem provides the existence and uniqueness by a contraction mapping argument, corresponding to [6, 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 [6, Proposition 8].

Theorem 3.9. *Let* $\frac{2}{3} < \alpha < 1$ *and let the non-linearities satisfy* (1.9)*. Then*

i) Suppose we are given a distribution f *satisfying*

$$\|f\|_{\alpha-2} \leqq N_0 \tag{3.81}$$

for some constant $N_0 \ll 1$*; denote by* $v(\cdot, a_0)$ *the space-time periodic and mean-free solution of* (1.2)*. Suppose further that we are given a one-parameter family of distributions* $\{v(\cdot, a_0') \diamond f\}_{a_0'}$ *and a two-parameter family of distributions* $\{v(\cdot, a_0') \diamond \partial_1^2 v(\cdot, a_0)\}_{a_0, a_0'}$ *satisfying*

$$\|[v, (\cdot)] \diamond f\|_{2\alpha-2,2}, \ \|[v, (\cdot)] \diamond \partial_1^2 v\|_{2\alpha-2,2,2} \leqq N_0^2. \tag{3.82}$$

(In fact, we do not need the highest cross-derivative $\frac{\partial^2}{\partial a_0'^2} \frac{\partial^2}{\partial a_0^2} [v, (\cdot)_T] \diamond \partial_1^2 v$*). Then there exists a unique mean-free function u with the properties*

$$u \textit{ is modelled after } v \textit{ according to } a(u) \textit{ and } \sigma(u), \tag{3.83}$$

$$\partial_2 u - P(a(u) \diamond \partial_1^2 u + \sigma(u) \diamond f) = 0 \quad \textit{distributionally}, \tag{3.84}$$

under the smallness condition

$$[u]_\alpha \ll 1. \tag{3.85}$$

This unique u satisfies the estimate

$$[u]_\alpha + \|u\| \lesssim N_0 \quad \text{and} \quad M \lesssim N_0^2, \tag{3.86}$$

where M denotes the modelling constant in (3.83).

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

$$\|f_j\|_{\alpha-2} \leqq N_0; \tag{3.87}$$

and let $v_j(\cdot, a_0)$ be the corresponding solutions of (1.2). Suppose further that for $i = 0, 1$ we are given four one-parameter families of distributions $\{v_i(\cdot, a_0') \diamond f_j\}_{a_0'}$ and four two-parameter families of distributions $\{v_i(\cdot, a_0') \diamond \partial_1^2 v_j(\cdot, a_0)\}_{a_0, a_0'}$ satisfying the analogue of (3.82) including the cross-terms

$$\|[v_i, (\cdot)] \diamond f_j\|_{2\alpha-2,2}, \ \ \|[v_i, (\cdot)] \diamond \partial_1^2 v_j\|_{2\alpha-2,2,2} \leqq N_0^2. \tag{3.88}$$

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

$$\|f_1 - f_0\|_{\alpha-2} \leqq \delta N_0, \tag{3.89}$$

$$\|[v_i, (\cdot)] \diamond \{f_1, \partial_1^2 v_1\} - [v_i, (\cdot)] \diamond \{f_0, \partial_1^2 v_0\}\|_{2\alpha-2,1,1} \leqq N_0 \delta N_0, \tag{3.90}$$

$$\|[v_1, (\cdot)] \diamond \{f_j, \partial_1^2 v_j\} - [v_0, (\cdot)] \diamond \{f_j, \partial_1^2 v_j\}\|_{2\alpha-2,1,1} \leqq N_0 \delta N_0. \tag{3.91}$$

If u_i, $i = 0, 1$, denote the corresponding solutions of (3.83)–(3.85) we have

$$[u_1 - u_0]_\alpha + \|u_1 - u_0\| \lesssim \delta N_0. \tag{3.92}$$

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

$$\delta M \lesssim N_0 \delta N_0. \tag{3.93}$$

It remains to establish a link between the solution theory presented in Theorem 3.9 and the classical solution theory in the case where f is smooth, for example $f \in C^\beta$ 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 $v(\cdot, a_0') \diamond \{f, \partial_1^2 v(\cdot, a_0)\} = v(\cdot, a_0')\{f, \partial_1^2 v(\cdot, a_0)\}$. In the language of Hairer [10, 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 (3.82) expressed in terms of commutators and they are easily seen to be satisfied in this case. For example we have

$$\|[v, (\cdot)_T] f\| = \sup_x \left| \int \psi_T(x-y)(v(x) - v(y)) f(y) \, dy \right| \overset{(2.4)}{\lesssim} T^{\frac{1}{4}} \|\{\partial_1, \partial_2\} v\| \|f\|, \tag{3.94}$$

which is much more than needed. However, the canonical definition is by no means the only possible choice of product. In fact, as (3.82) is the only requirement on $v(\cdot, a_0') \diamond \{f, \partial_1^2 v(\cdot, a_0)\}$ we can set for example

$$v(\cdot, a_0') \diamond \{f, \partial_1^2 v(\cdot, a_0)\} := v(\cdot, a_0')\{f, \partial_1^2 v(\cdot, a_0)\} + \{g_1, g_2\} \tag{3.95}$$

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 (3.82) reduces to the regularity assumption

$$\|g_1\|_{2\alpha-2,2}, \|g_2\|_{2\alpha-2,2,2} < \infty. \tag{3.96}$$

The following corollary provides a link between solutions of (3.84) and classical solutions in the case where the products $\diamond$ are defined by (3.95):

Corollary 3.10. *Let f be a space-time periodic 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 (3.95) for g_1, g_2 which satisfy (3.96). Then for a periodic mean-free function u 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 a classical solution of

$$\partial_2 u - P\big(a(u)\partial_1^2 u + a'(u)\sigma^2(u) g_2(\cdot, a(u), a(u)) \\ + \sigma(u) f + \sigma'(u)\sigma(u) g_1(\cdot, a(u))\big) = 0.$$

4. 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 (3.81) and (3.82) hold for these objects. The calculations in this section are inspired by a similar reasoning (in a more complicated situation) in [13, Sec. 5], [10, Sec. 10]; for the reader's convenience we provide self-contained proofs.

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

$$f(x) = \sum_{k \in (2\pi\mathbb{Z})^2} \sqrt{\hat{C}(k)} e^{ik\cdot x} Z_k, \tag{4.1}$$

which converges in a suitable topology on distributions. The Z_k are complex-valued centered Gaussians which are independent except for the symmetry constraint $Z_k = \bar{Z}_{-k}$ and satisfy $\langle Z_k Z_{-\ell} \rangle = \delta_{k,\ell}$, where as in the introduction we use angled brackets $\langle \cdot \rangle$ to denote the expectation of a random variable. 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 realizations from f are (say smooth) functions the coefficients in (4.1) coincide with the square root of the Fourier series of the covariance function.

Throughout this section we assume that $\hat{C}(0) = 0$, that is 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 $\alpha' \in (\frac{1}{4}, 1)$ such that

$$\hat{C}(k) \leqq \frac{1}{(1+|k_1|)^{\lambda_1}(\sqrt{1+|k_2|})^{\lambda_2}}, \qquad k = (k_1, k_2) \in (2\pi\mathbb{Z})^2,$$
$$\lambda_1 + \lambda_2 = -1 + 2\alpha' \qquad \lambda_1, \frac{\lambda_2}{2} < 1. \tag{4.2}$$

The second condition may be confusing, because larger values of λ_i, corresponding to more smoothness for f, should help our theory. The point here is that decay in one of the directions beyond summability cannot compensate for a lack of decay in the other direction. The upcoming Lemma 4.1 shows that assumption (4.2) corresponds to the regularity assumption (3.81) on f. In order to use the bounds presented in Lemma 4.1 and Proposition 4.2 as input for the deterministic theory in Section 3 we only need the case where $\alpha' > \frac{2}{3}$ but the construction presented in this section works under the weaker assumption $\alpha' > \frac{1}{4}$ without additional difficulty.

As in the introduction, we fix an arbitrary Schwartz function φ with $\int_{\mathbb{R}^2} \varphi = 1$ and define the rescaling φ_ε and the regularized noise f_ε as in (1.7). Of course, $\varphi = \psi_1$ for ψ_1 as in the deterministic analysis is an admissible choice, but in the following analysis of stochastic moments the semi-group property for φ is not needed, and we therefore do not need to restrict ourselves to this particular choice.

Lemma 4.1. *Let f be given by (4.1) satisfying (4.2) for some $\alpha' < 1$ and let f_ε be as in (1.7). Then we have for any $p < \infty$ and $\alpha < \alpha'$*

$$\left\langle \sup_{\varepsilon \in [0,1]} \| f_\varepsilon \|^p_{\alpha-2} \right\rangle^{\frac{1}{p}} \lesssim 1, \tag{4.3}$$

where we use the convention $f_0 := f$. If additionally $0 \leqq \kappa \leqq 1$, then

$$\left\langle \left(\sup_{\varepsilon \in (0,1]} (\varepsilon^{\frac{1}{4}})^{-\kappa} \| f_\varepsilon - f \|_{\alpha-2-\kappa} \right)^p \right\rangle^{\frac{1}{p}} \lesssim 1. \tag{4.4}$$

Here and in the proof the implicit constant in $\lesssim$ depends only on the λ_i, p, α as well as our choice of regularising kernel φ.

As before let $v(\cdot, a_0)$ denote the space-time periodic and mean-free solution to (1.2). We aim at giving a meaning to the products $v(\cdot, a_0) \diamond f$ and $v(\cdot, a_0) \diamond \partial_1^2 v(\cdot, a_0')$, and obtaining bounds for the families of commutators derived from them. The regularities of $v(\cdot, a_0)$, f and $\partial_1^2 v(\cdot, a_0)$ are not sufficient to give a deterministic interpretation to these products, and we therefore seek a probabilistic argument to show the convergence of regularized products. We define $v_\varepsilon(\cdot, a_0)$ as in (1.7) 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. In general, under assumption (4.2), these regularized products do not converge as the regularization is removed, but convergence can be enforced by subtracting their expectation. Therefore, we define the renormalized products

$$\begin{aligned} v_\varepsilon(\cdot,a_0)\diamond f_\varepsilon &:= v_\varepsilon(\cdot,a_0) f_\varepsilon - g_1(\varepsilon,a_0),\\ v_\varepsilon(\cdot,a_0)\diamond \partial_1^2 v_\varepsilon(\cdot,a_0') &:= v_\varepsilon(\cdot,a_0)\partial_1^2 v_\varepsilon(\cdot,a_0') - g_2(\varepsilon,a_0,a_0'), \end{aligned} \tag{4.5}$$

where as in (1.11) we set $g_1(\varepsilon,a_0) = \langle v_\varepsilon(\cdot,a_0) f_\varepsilon\rangle$ and $g_2(\varepsilon,a_0,a_0') = \langle v_\varepsilon(\cdot,a_0)\partial_1^2 v_\varepsilon(\cdot,a_0')\rangle$.

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 as well as their derivatives with respect to a_0, a_0':

Proposition 4.2. *Let f be a stationary centered Gaussian distribution given by (4.1) satisfying (4.2) for some $\frac{1}{4} < \alpha' < 1$, let $v(\cdot,a_0')$ be the space-time periodic mean-free solution of (1.2) and let f_ε and $v_\varepsilon(\cdot,a_0')$ be as in (1.7). Then we have that*

i) For any $n, m \geqq 0$ the random distributions $\left(\frac{\partial}{\partial a_0}\right)^n \left(\frac{\partial}{\partial a_0'}\right)^m v_\varepsilon(\cdot,a_0)\diamond\{f_\varepsilon, \partial_1^2 v_\varepsilon(\cdot,a_0')\}$ 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 $\left(\frac{\partial}{\partial a_0}\right)^n \left(\frac{\partial}{\partial a_0'}\right)^m v(\cdot,a_0)\diamond\{f, \partial_1^2 v(\cdot,a_0')\}$.

ii) For all $p<\infty$ we have the estimates

$$\begin{aligned} &\Big\langle \sup_{\varepsilon_0,\varepsilon_1\in[0,1]} \|[v_{\varepsilon_0},(\cdot)]\diamond f_{\varepsilon_1}\|_{2\alpha-2,n}^p\Big\rangle^{\frac{1}{p}} \lesssim 1,\\ &\Big\langle \sup_{\varepsilon_0,\varepsilon_1\in[0,1]} \|[v_{\varepsilon_0},(\cdot)]\diamond \partial_1^2 v_{\varepsilon_1}\|_{2\alpha-2,n,m}^p\Big\rangle^{\frac{1}{p}} \lesssim 1, \end{aligned} \tag{4.6}$$

as well as for $0<\kappa\ll 1$ (where $\ll$ depends only on λ_1,λ_2)

$$\begin{aligned} &\Big\langle\Big(\sup_{\varepsilon\in(0,1]} (\varepsilon^{\frac14})^{-\kappa}\big\|[v_\varepsilon,(\cdot)]\diamond f_\varepsilon - [v,(\cdot)]\diamond f\big\|_{2\alpha-2-\kappa,n}\Big)^p\Big\rangle^{\frac1p} \lesssim 1,\\ &\Big\langle\Big(\sup_{\varepsilon\in(0,1]} (\varepsilon^{\frac14})^{-\kappa}\big\|[v_\varepsilon,(\cdot)]\diamond \partial_1^2 v_\varepsilon - [v,(\cdot)]\diamond \partial_1^2 v\big\|_{2\alpha-2-\kappa,n,m}\Big)^p\Big\rangle^{\frac1p} \lesssim 1, \end{aligned} \tag{4.7}$$

where here and in the proof $\lesssim$ means up to a constant depending only on n, m, the λ_i, α, κ, p, the ellipticity contrast λ as well as the specific choice of regularising kernel φ. In both estimates the subscripts n, m in the norms refer to parameter derivatives with respect to a_0, a_0' as in (3.18) and (3.19).

Proposition 4.2 is a consequence of the following estimate on the second moments of commutators:

Lemma 4.3. *Let f and $v(\cdot,a_0)$ be as in Proposition 4.2. Let $\hat{M}_1, \hat{M}_2$ be Fourier multipliers satisfying*

$$\hat{M}_i(k) = \overline{\hat{M}_i(-k)} \quad \text{and} \quad |\hat{M}_i(k)| \leqq (k_1^4+k_2^2)^{\frac{\kappa_i}{4}}, \quad k\in(2\pi\mathbb{Z})^2, i=1,2 \tag{4.8}$$

for $0 \leqq \kappa_1, \kappa_2 \ll 1$ *(where* $\ll$ *depends only on* λ_1, λ_2*). Let* f' *and* $v'(\cdot, a_0)$ *be defined through their Fourier series*

$$\hat{f}' = \hat{M}_1 \hat{f} \quad \text{and} \quad \hat{v}'(\cdot, a_0) = \hat{M}_2 \hat{v}(\cdot, a_0).$$

We make the qualitative assumption that f' *and* $v'(\cdot, a_0)$ *are smooth and set*

$$v'(\cdot, a_0) \diamond f' := v'(\cdot, a_0) f' - \langle v'(\cdot, a_0) f' \rangle.$$

Then for all $a_0 \in [\lambda, \frac{1}{\lambda}]$

$$\langle ([v'(\cdot, a_0), (\cdot)_T] \diamond f')^2 \rangle^{\frac{1}{2}} \lesssim (T^{\frac{1}{4}})^{2\alpha' - 2 - \kappa_1 - \kappa_2}. \tag{4.9}$$

Here and in the proof the implicit constant depends on $\lambda_1, \lambda_2, \kappa_1, \kappa_2$ *as well as the ellipticity contrast* λ *(but not on the qualitative smoothness assumption on* f', v'*).*

In the proof of Proposition 4.2 this lemma is used in the form of the following immediate corollary:

Corollary 4.4. *Let* f, f_ε, v *and* v_ε *be as in Proposition 4.2. Then for* $n, m \geqq 0$ *we have*

$$\Big\langle \Big(\Big[\Big(\frac{\partial}{\partial a_0} \Big)^n v_{\varepsilon_0}(\cdot, a_0), (\cdot)_T \Big] \diamond \Big\{ f_{\varepsilon_1}, \Big(\frac{\partial}{\partial a_0'} \Big)^m \partial_1^2 v_{\varepsilon_1}(\cdot, a_0') \Big\} \Big)^2 \Big\rangle^{\frac{1}{2}} \lesssim (T^{\frac{1}{4}})^{2\alpha' - 2}. \tag{4.10}$$

Furthermore, we have for $0 \leqq \kappa \ll 1$ *(*$\ll$ *depends only on* λ_1, λ_2*) and for* $i = 0, 1$

$$\Big\langle \Big(\varepsilon_i \frac{\partial}{\partial \varepsilon_i} \Big(\Big[\Big(\frac{\partial}{\partial a_0} \Big)^n v_{\varepsilon_0}(\cdot, a_0), (\cdot)_T \Big] \diamond \Big\{ f_{\varepsilon_1}, \Big(\frac{\partial}{\partial a_0'} \Big)^m \partial_1^2 v_{\varepsilon_1} \Big\} \Big) \Big)^2 \Big\rangle^{\frac{1}{2}}$$
$$\lesssim (T^{\frac{1}{4}})^{2\alpha' - 2 - \kappa} (\varepsilon_i^{\frac{1}{4}})^\kappa. \tag{4.11}$$

Here and in the proof the implicit constant depends on λ_1, λ_2, κ *the ellipticity contrast* λ, n, m *as well as the specific choice of regularising kernel* φ.

Finally, the following lemma deals with the behaviour of the expectations g_1, g_2 as the regularization is removed:

Lemma 4.5. *i) For* $\varepsilon > 0$ *we have*

$$g_1(\varepsilon, a_0) = \sum_{k \in (2\pi\mathbb{Z})^2 \setminus \{0\}} \frac{a_0 k_1^2}{a_0^2 k_1^4 + k_2^2} \hat{C}(k) |\hat{\varphi}_\varepsilon(k)|^2, \tag{4.12}$$

$$g_2(\varepsilon, a_0, a_0') = \sum_{k \in (2\pi\mathbb{Z})^2 \setminus \{0\}} \frac{(-a_0 a_0' k_1^4 + k_2^2) k_1^2}{(a_0^2 k_1^4 + k_2^2)((a_0')^2 k_1^4 + k_2^2)} \hat{C}(k) |\hat{\varphi}_\varepsilon(k)|^2. \tag{4.13}$$

ii) The expectation $g_1(\varepsilon, a_0)$ *converges to a finite limit as* $\varepsilon \to 0$ *if and only if*

$$\sum_{k \in (2\pi\mathbb{Z})^2 \setminus \{0\}} \frac{k_1^2}{k_1^4 + k_2^2} \hat{C}(k) < \infty. \tag{4.14}$$

If (4.14) holds, then $g_2(\varepsilon, a_0, a_0')$ *as well as all parameter derivatives* $(\frac{\partial}{\partial a_0})^n$ $g_1(\varepsilon, a_0)$ *and* $(\frac{\partial}{\partial a_0})^n$ $(\frac{\partial}{\partial a_0'})^m$ $g_2(\varepsilon, a_0, a_0')$ *for* $n, m \geqq 0$ *converge as well.*

In particular we immediately get

Corollary 4.6. *Assume that both* (4.2) *and* (4.14) *hold. Then the statements of Proposition* 4.2 *remain true if all of the renormalized products are replaced by products without renormalization.*

We finish this section by discussing the assumptions (4.2) and (4.14) in particular cases. First consider the case

$$\hat{C}(k) = \frac{1}{(1+|k_1|)^{\lambda_1}(\sqrt{1+|k_2|})^{\lambda_2}}. \tag{4.15}$$

For this choice of $\hat{C}$ the regularity assumption (4.2) is equivalent to

$$\lambda_1 + \lambda_2 \geqq -1 + 2\alpha', \qquad \lambda_1 > -3 + 2\alpha', \quad \text{and} \quad \lambda_2 > -2 + 2\alpha'. \tag{4.16}$$

Note that equality is not necessary in the first condition, because in the case of strict inequality, one can find $\lambda_1' \leqq \lambda_1$ and $\lambda_2' \leqq \lambda_2$ that satisfy (4.2) with equality. The condition (4.14) on the other hand is equivalent to

$$\lambda_1 + \lambda_2 > 1 \qquad \lambda_1 > -1, \quad \text{and} \quad \lambda_2 > -2. \tag{4.17}$$

An interesting case in which both assumptions are satisfied and for which our theory can therefore be applied without renormalization 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 renormalization, 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

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

which corresponds to the choice $\lambda_2 = \infty$ in (4.15) and yields 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 [1,7,11,12]. Our theory applies without renormalization for all $\lambda_1 > -1$, which covers in particular the case of one-dimensional spatial white noise, $\lambda_1 = 0$. If we admit renormalization we can go all the way to $\lambda_1 > -\frac{5}{3}$. 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.

5. Proofs for the Deterministic Analysis

5.1. Proof of Theorem 3.9

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

$$(\bar{u}, \bar{a}, \bar{\sigma}) \mapsto (\sigma := \sigma(\bar{u}), a := a(\bar{u}), \sigma \diamond f, a \diamond \partial_1^2 v) \mapsto (u, a, \sigma), \tag{5.1}$$

where u is the solution provided by Proposition 3.8. This is the map of which we seek to characterize the fixed point. Note that the right hand side depends on $\bar{a}$ and $\bar{\sigma}$ via the definition of the products $\sigma \diamond f$ and $a \diamond \partial_1^2$.

STEP 1. Pointwise nonlinear transformation, application of Lemma 3.2. 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

$$\bar{\sigma}_i \in [-1, 1],\ \bar{a}_i \in \left[\lambda, \frac{1}{\lambda}\right], \quad [\bar{\sigma}_i], [\bar{a}_i] \leqq 1. \tag{5.2}$$

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

$$\bar{M} := \max_i (M_{\bar{u}_i} + [\bar{u}_i]) + N_0, \tag{5.3}$$

$$\begin{aligned} \delta\bar{M} := {} & M_{\bar{u}_1 - \bar{u}_0} + [\bar{u}_1 - \bar{u}_0] + \|\bar{u}_1 - \bar{u}_0\| \\ & + N_0([\bar{\sigma}_1 - \bar{\sigma}_0] + \|\bar{\sigma}_1 - \bar{\sigma}_0\| + [\bar{a}_1 - \bar{a}_0] + \|\bar{a}_1 - \bar{a}_0\|) + \delta N_0, \end{aligned} \tag{5.4}$$

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

$$\sigma_i \in [-1, 1],\ a_i \in [\lambda, \frac{1}{\lambda}],\ [\sigma_i], [a_i] \ll 1 \quad \text{provided } \max_i [\bar{u}_i] \ll 1, \tag{5.5}$$

$$\tilde{M} \lesssim \bar{M} \quad \text{provided } \max_i [\bar{u}_i] \leqq 1, \tag{5.6}$$

$$\delta\tilde{M} \lesssim \delta\bar{M} \quad \text{provided } \bar{M} \leqq 1, \tag{5.7}$$

where we define, in analogy with (5.3) and (5.4),

$$\tilde{M} := \max_i (M_{\sigma_i} + [\sigma_i] + M_{a_i} + [a_i]) + N_0, \tag{5.8}$$

$$\begin{aligned} \delta\tilde{M} := {} & M_{\sigma_1 - \sigma_0} + [\sigma_1 - \sigma_0] + \|\sigma_1 - \sigma_0\| \\ & + N_0\big([\omega_1 - \omega_0] + \|\omega_1 - \omega_0\| + [\bar{a}_1 - \bar{a}_0] + \|\bar{a}_1 - \bar{a}_0\|\big) \\ & + M_{a_1 - a_0} + [a_1 - a_0] + \|a_1 - a_0\| \\ & + N_0\big([\mu_1 - \mu_0] + \|\mu_1 - \mu_0\| + [\bar{a}_1 - \bar{a}_0] + \|\bar{a}_1 - \bar{a}_0\|\big) + \delta N_0, \end{aligned} \tag{5.9}$$

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 (1.9) that we have (5.5) under the assumption $\max_i [\bar{u}_i] \ll 1$. Estimate (5.6) follows from part i) of Lemma 3.2 with u replaced by $\bar{u}_i$ and the generic nonlinearity b replaced by σ and by a, respectively, (using our assumptions (1.9)). More precisely, (5.6) follows from (3.2) by $[\bar{u}_i] \leqq 1$. We now turn to (5.7), which by definitions (5.4) of $\delta\bar{M}$ and (5.9) of $\delta\tilde{M}$ and because of $N_0 \leqq 1$ we may split into the four statements

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

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

where we also used the definition (5.3) of $\bar{M}$. This is a consequence of part ii) of Lemma 3.2 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 (3.4), the second and fourth one of (3.5), in which we use (5.5). It is on all four we use our full assumptions (1.9) on the nonlinearities σ and a.

STEP 2. Using the off-line products, application of Corollary 3.4. 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

$$\|[\sigma_i, (\cdot)] \diamond f_j\|_{2\alpha-2} \lesssim N_0\tilde{M}, \tag{5.10}$$

$$\|[\sigma_i, (\cdot)] \diamond f_1 - [\sigma_i, (\cdot)] \diamond f_0\|_{2\alpha-2} \lesssim \delta N_0\tilde{M}, \tag{5.11}$$

$$\|[\sigma_1, (\cdot)] \diamond f_j - [\sigma_0, (\cdot)] \diamond f_j\|_{2\alpha-2} \lesssim N_0\delta\tilde{M}, \tag{5.12}$$

$$\|[a_i, (\cdot)] \diamond \partial_1^2 v_j\|_{2\alpha-2,2} \lesssim N_0\tilde{M}, \tag{5.13}$$

$$\|[a_i, (\cdot)] \diamond \partial_1^2 v_1 - [a_i, (\cdot)] \diamond \partial_1^2 v_0\|_{2\alpha-2,1} \lesssim \delta N_0\tilde{M}, \tag{5.14}$$

$$\|[a_1, (\cdot)] \diamond \partial_1^2 v_j - [a_0, (\cdot)] \diamond \partial_1^2 v_j\|_{2\alpha-2,1} \lesssim N_0\delta\tilde{M}. \tag{5.15}$$

This is an application of Corollary 3.4 with $(N_1, \delta N_1) = (N_0, \delta N_0)$. Estimate (5.10) is an application of Corollary 3.4 i) with u replaced by σ_i; the hypotheses (3.21) and (3.22) are contained in the theorem's assumptions (3.81) and (3.82) (note that f does not depend on an extra parameter a_0'). The output (3.25) turns into (5.10) since by definition (5.8), $M_{\sigma_i} + N_0 \leqq \tilde{M}$. Estimate (5.11) is an application of Corollary 3.4 ii) still applied with u replaced by σ_i; the hypotheses (3.26) and (3.27) are contained in the theorem's assumptions (3.89) and (3.90). The output (3.28) turns into (5.11) as in the previous application. Estimate (5.12) is an application of Corollary 3.4 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 (3.31) and (3.32) are contained in the theorem's assumptions (3.88) and (3.91). The output (3.33) turns into (5.12), since, by definition (5.9), we have

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

The arguments for (5.13), (5.14) and (5.15) follow the same lines of those for (5.10), (5.11) and (5.12), 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 3.4). Hence the hypotheses (3.21) and (3.26) in Corollary 3.4 turn into

$$\|\partial_1^2 v_j\|_{\alpha-2,2} \lesssim N_0, \qquad \|\partial_1^2(v_1 - v_0)\|_{\alpha-2,1} \lesssim \delta N_0.$$

This follows from Step 1 in the proof of Corollary 3.7 via (2.4).

STEP 3. Application of Proposition 3.8. 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 $v_i \diamond \partial_1^2 v_j$

$$M \lesssim N_0(\tilde{M} + 1) \qquad \text{provided } \max_i [\bar{u}_i] \ll 1, \tag{5.16}$$

$$\max_i M_{u_i} \lesssim N_0 \tilde{M} \qquad \text{provided } \max_i [\bar{u}_i] \ll 1, \tag{5.17}$$

$$\delta M \lesssim N_0 \delta\tilde{M} + \delta N_0 \qquad \text{provided in addition } \tilde{M} \lesssim 1, \tag{5.18}$$

$$M_{u_1 - u_0} \lesssim N_0 \delta\tilde{M} + \delta N_0 \tilde{M} \qquad \text{provided in addition } \tilde{M} \lesssim 1, \tag{5.19}$$

where we define consistently with (5.3) and (5.4)

$$M := \max_i (M_{u_i} + [u_i]) + N_0, \tag{5.20}$$

$$\begin{aligned} \delta M := {} & M_{u_1-u_0} + [u_1 - u_0] + \|u_1 - u_0\| \\ & + N_0\Big([\sigma_1 - \sigma_0] + \|\sigma_1 - \sigma_0\| + [a_1 - a_0] + \|a_1 - a_0\|\Big) + \delta N_0. \end{aligned} \tag{5.21}$$

Indeed, (5.16) and (5.17) are an application of part i) of Proposition 3.8: The hypothesis (3.43) of the proposition is built into the definition (5.8) of $\tilde{M}$, so that $\tilde{M}$ here plays the role of N in the proposition. The hypothesis (3.44) is identical to the theorem's assumption (3.87), hypothesis (3.47) was established in (5.5), hypotheses (3.45) and (3.46) are contained in (5.10) and (5.13) of Step 2 which is consistent with $\tilde{M}$ playing the role of N there. The combination of (3.69) and (3.70) amounts to (5.16) by definition (5.20) of M. Estimate (3.69) by itself amounts to (5.17).

Estimate (5.18) in turn is a consequence of part ii) of Proposition 3.8: Hypothesis (3.73) of the proposition is build into the definition (5.9) of $\delta\tilde{M}$, so that $\delta\tilde{M}$ here plays the role of δN in the proposition. Hypotheses (3.71) and (3.72) are identical to (5.10) and (5.13) of Step 2. Hypothesis (3.74) is identical to our assumption (3.89), hypotheses (3.75), (3.76), (3.77), and (3.78) are identical to (5.11), (5.12), (5.14), and (5.15) in Step 2. The outcome (3.79) of the proposition turns into (5.19). 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 (3.80) 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 (5.9) of $\delta\tilde{M}$ we have

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

The combination of the last three statement yields (5.18) in view of definition (5.21).

STEP 4. Still 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 $v_i \diamond \partial_1^2 v_j$, estimates (5.6) and (5.7) in Step 1 and Step 3 obviously combine to

$$M \lesssim N_0(\bar{M} + 1) \qquad \text{provided } \max_i [\bar{u}_i] \ll 1, \tag{5.22}$$

$$\max_i M_{u_i} \lesssim N_0 \bar{M} \qquad \text{provided } \max_i [\bar{u}_i] \ll 1, \tag{5.23}$$

$$\delta M \lesssim N_0 \delta\bar{M} + \delta N_0 \qquad \text{provided in addition } \bar{M} \leqq 1, \tag{5.24}$$

$$M_{u_1 - u_0} \lesssim N_0 \delta\bar{M} + \delta N_0 \bar{M} \qquad \text{provided in addition } \bar{M} \leqq 1. \tag{5.25}$$

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 (5.2), and which satisfy

$$\bar{M} \leqq N, \tag{5.26}$$

cf. (5.3), 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 (5.26) and the definition (5.3) of $\bar{M}$ we learn that the proviso of (5.22) is fulfilled provided the constant N is sufficiently small, which we now fix accordingly. We thus learn from (5.22), which by (5.26) assumes the form of $M \lesssim N_0$, that the map defined through (5.1) sends the set defined through (5.26) into itself, provided $N_0 \ll 1$.

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

$$\begin{aligned} d\big((u_1, a_1, \sigma_1), (u_0, a_0, \sigma_0)\big) &:= M_{u_1 - u_0} + [u_1 - u_0] + \|u_1 - u_0\| \\ &\quad + N_0([\sigma_1 - \sigma_0] + \|\sigma_1 - \sigma_0\| + [a_1 - a_0] + \|a_1 - a_0\|) \end{aligned} \tag{5.27}$$

defines a distance function. Indeed, that also the modelledness constant $M_{u_1 - u_0}$ satisfies a triange inequality in (u_i, a_i, σ_i) can be seen by rewriting the definition (3.1) as

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

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 (5.26) we learn that the proviso of

(5.24) is fulfilled; because of $\delta N_0 = 0$, (5.24) assumes the form $\delta M \lesssim N_0 \delta \bar{M}$. By definitions (5.4) and (5.21) 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 (5.1) 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 (5.27) is *complete*; and that the subset defined through the constraints (5.2) and (5.26) is *closed*. Hence by the contraction mapping principle the map (5.1) admits a unique fixed point on the set defined through (5.2) and (5.26).

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 (5.1), which is obvious, that lies in the set defined through the constraints (5.2) and (5.26), and therefore is unique. Indeed, in view of $[a] \leqq \|a'\|[u] \leqq 1$, $[\sigma] \leqq \|\sigma'\|[u] \leqq 1$ by (1.9) and (3.85), the constraints (5.2) are satisfied. The constraint (5.26) will be an immediate consequence of the stronger statement (3.86) (provided N_0 is sufficiently small). We thus turn to this a priori estimate (3.86) and 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 (5.22) and (5.23) are satisfied so that because of $N_0 \ll 1$, their application yields

$$M \lesssim N_0 \quad \text{and thus} \quad M_u \lesssim N_0^2. \tag{5.28}$$

By definition (5.20) and the vanishing mean of u, this turns into (3.86).

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 (5.2) and (5.26) and each triplet is a fixed point of "its own" map (5.1) (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 (3.86) and $N_0 \ll 1$, the two provisos of Step 4 are satisfied. Because of $N_0 \ll 1$, (5.24) and (5.25) 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 (5.28). By definition (5.21) of δM, this turns into (3.92) and (3.93).

5.2. *Proof of Proposition 3.8*

We continue to abbreviate $[\cdot] = [\cdot]_\alpha$. When a function v depends on a_0 and 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 $\|f\|_{\alpha-2} < \infty$. 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 (3.68). Indeed, let u' be another function with these properties; we trivially have by Definition 3.1 that $u-u'$ is modelled after v according to a and to 0 playing the role of σ. We now apply Lemma 3.5 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 (3.36) and the triangle inequality that

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

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,

$$\partial_2(u-u')-Pa\diamond\partial_1^2(u-u')=0. \tag{5.29}$$

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 3.7 i) with $f=0$ and thus $N_0=0$. We apply it with u replaced by $u-u'$, which we may thanks to (5.29). In this context, the output (3.50) of Corollary 3.7 assumes the form $[u-u']=0$. Since $u-u'$ has vanishing average, we obtain as desired $u-u'=0$.

STEP 2. A special regularization. Under the assumptions of Lemma 3.5 and for $\tau>0$ and $i=1,\dots,I$ we consider the convolution $v_{i\tau}$ of v_i and *define*

$$a\diamond\partial_1^2 v_{i\tau}:=(a\diamond\partial_1^2 v_i)_\tau. \tag{5.30}$$

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

$$a\diamond\partial_1^2 u=a\partial_1^2 u-\sigma_i E[a,(\cdot)_\tau]\diamond\partial_1^2 v_i. \tag{5.31}$$

Indeed, by Lemma 3.5 (with b replaced by a) we understand the distribution $a\diamond\partial_1^2 u$ as defined by

$$\lim_{T\downarrow 0}\|[a,(\cdot)_T]\diamond\partial_1^2 u-\sigma_i E[a,(\cdot)_T]\diamond\partial_1^2 v_{i\tau}\|=0. \tag{5.32}$$

We note that (5.30) implies by the semi-group property:

$$[a,(\cdot)_T]\diamond\partial_1^2 v_{i\tau}=[a,(\cdot)_{T+\tau}]\diamond\partial_1^2 v_i, \tag{5.33}$$

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 (3.35), this convergence is even uniform in (x,a_0), so that (5.32) 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 (5.31).

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 (5.30) (without the index i) by

$$\sigma \diamond f_\tau := (\sigma \diamond f)_\tau. \tag{5.34}$$

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

$$\partial_2 u^\tau - P(a \diamond \partial_1^2 u^\tau + \sigma \diamond f_\tau) = 0 \quad \text{distributionally,} \tag{5.35}$$

and at the same time

$$\partial_2 u^\tau - P(a\partial_1^2 u^\tau - \sigma E[a, (\cdot)_\tau] \diamond \partial_1^2 v + (\sigma \diamond f)_\tau) = 0 \quad \text{classically.} \tag{5.36}$$

We first turn to the existence of (5.36) and start by noting that the right hand side $-\sigma E[a, (\cdot)_\tau] \diamond \partial_1^2 v + (\sigma \diamond f)_\tau$ in (5.36) 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 [14, 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 (5.34) we see that (5.36) may be rewritten as (5.35).

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:

$$\sigma_s := s\sigma_1 + (1 - s)\sigma_0 \quad \text{and the same for } a,\ f, \text{ and } v. \tag{5.37}$$

We note that this preserves (3.47). We interpolate the products bi-linearly:

$$\begin{aligned}
\sigma_s \diamond f_s &:= s^2\sigma_1 \diamond f_1 + s(1 - s)\sigma_1 \diamond f_0 \\
&\qquad + (1 - s)s\sigma_0 \diamond f_1 + (1 - s)^2\sigma_0 \diamond f_0, \\
\partial_s\sigma \diamond f_s &:= s\sigma_1 \diamond f_1 + (1 - s)\sigma_1 \diamond f_0 - s\sigma_0 \diamond f_1 - (1 - s)\sigma_0 \diamond f_0, \\
\sigma_s \diamond \partial_s f &:= s\sigma_1 \diamond f_1 - s\sigma_1 \diamond f_0 + (1 - s)\sigma_0 \diamond f_1 - (1 - s)\sigma_0 \diamond f_0, \\
&\text{and the same for } a_s \diamond \partial_1^2 v_s,\ \ \partial_s a \diamond \partial_1^2 v_s \text{ and } a_s \diamond \partial_1^2 \partial_s v,
\end{aligned} \tag{5.38}$$

where here and below we use the convention that ∂_s only acts on the object directly following it (with argument suppressed), that is for example $\partial_s\sigma \diamond f_s = (\partial_s\sigma_s) \diamond f_s$.

Thanks to the estimate (3.72), 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

$$a_s \diamond \partial_1^2 \frac{\partial v_s}{\partial a_0}(\cdot, a_0) := \frac{\partial}{\partial a_0} a_s \diamond \partial_1^2 v_s(\cdot, a_0). \tag{5.39}$$

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

$$\sigma_s \diamond f_{s\tau} := (\sigma_s \diamond f_s)_\tau \quad \text{and the same for } \partial_s\sigma \diamond f_{s\tau},\ \sigma_s \diamond \partial_s f_\tau,$$
$$a_s \diamond \partial_1^2 v_{s\tau},\ \partial_s a \diamond \partial_1^2 v_{s\tau},\ a_s \diamond \partial_1^2 \partial_s v_\tau,\ a_s \diamond \partial_1^2 \frac{\partial v_{s\tau}}{\partial a_0}. \tag{5.40}$$

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

$$u_s^\tau \text{ is modelled after } v_{s\tau} \text{ according to } a_s \text{ and } \sigma_s, \tag{5.41}$$
$$\partial_2 u_s^\tau - P(a_s \diamond \partial_1^2 u_s^\tau + \sigma_s \diamond f_{s\tau}) = 0 \quad \text{distributionally.} \tag{5.42}$$

Furthermore, we claim that

$$\partial_s u^\tau \text{ is modelled after } \left(v_{s\tau}, \frac{\partial v_{s\tau}}{\partial a_0}, \partial_s v_\tau\right) \text{ according to } a_s \text{ and } (\partial_s\sigma, \sigma_s\partial_s a, \sigma_s), \tag{5.43}$$

$$\partial_2\partial_s u^\tau - P\left(a_s \diamond \partial_1^2 \partial_s u^\tau + \partial_s a \diamond \partial_1^2 u_s^\tau + \sigma_s \diamond \partial_s f_\tau + \partial_s \sigma \diamond f_{s\tau}\right) = 0 \tag{5.44}$$

distributionally. Note that (5.44) is what we get from formally applying ∂_s to (5.42).

Here is the argument: 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. (5.40), there exists a unique mean-free u_s^τ of class $C^{\alpha+2}$ such that (5.41) and (5.42) hold. Furthermore by Step 2 u_s^τ is characterized as the classical solution of

$$\partial_2 u_s^\tau - P\left(a_s \partial_1^2 u_s^\tau - \sigma_s E_s[a_s, (\cdot)_\tau] \diamond \partial_1^2 v_s + (\sigma_s \diamond f_s)_\tau\right) = 0. \tag{5.45}$$

In preparation for taking the s-derivative of (5.45) we note that the definition (5.38) of $\sigma_s \diamond f_s$ and $a_s \diamond \partial_1^2 v_s$ by (bi-)linear interpolation ensures that Leibniz's rule holds:

$$\partial_s(\sigma_s \diamond f_s) = \partial_s\sigma \diamond f_s + \sigma_s \diamond \partial_s f, \tag{5.46}$$
$$\partial_s(a_s \diamond \partial_1^2 v_s) = \partial_s a \diamond \partial_1^2 v_s + a_s \diamond \partial_1^2 \partial_s v. \tag{5.47}$$

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 (5.47) and (5.39)

$$\begin{aligned}&\partial_s(E_s a_s \diamond \partial_1^2 v_s)\\ &\quad = E_s \partial_s a \diamond \partial_1^2 v_s + \partial_s a E_s a_s \diamond \partial_1^2 \frac{\partial v_s}{\partial a_0} + E_s a_s \diamond \partial_1^2 \partial_s v,\end{aligned}$$

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

$$\begin{aligned}\partial_s(E_s[a_s, (\cdot)_\tau] \diamond \partial_1^2 v_s) &= E_s[\partial_s a, (\cdot)_\tau] \diamond \partial_1^2 v_s\\ &\quad + \partial_s a E_s[a_s, (\cdot)_\tau] \diamond \partial_1^2 \frac{\partial v_s}{\partial a_0} + E_s[a_s, (\cdot)_\tau] \diamond \partial_1^2 \partial_s v.\end{aligned} \tag{5.48}$$

Equipped with (5.46), (5.47) and (5.48), we learn from (5.45) by the argument of Step 3 that u_s^τ is differentiable in s with values in the class $C^{\alpha+2}$ and

$$\partial_2\partial_s u^\tau - P\Big(a_s\partial_1^2\partial_s u^\tau + \partial_s a\partial_1^2 u_s^\tau - \sigma_s E_s[\partial_s a, (\cdot)_\tau]\diamond\partial_1^2 v_s - \partial_s\sigma E_s[a_s, (\cdot)_\tau]\diamond\partial_1^2 v_s - \sigma_s\partial_s a E_s[a_s, (\cdot)_\tau]\diamond\partial_1^2\frac{\partial v_s}{\partial a_0} - \sigma_s E_s[a_s, (\cdot)_\tau]\diamond\partial_1^2\partial_s v + (\partial_s\sigma\diamond f_s)_\tau + (\sigma_s\diamond\partial_s f)_\tau\Big) = 0. \tag{5.49}$$

Moreover, like in Step 3, (5.43) 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 (5.40) of $\partial_s a\diamond\partial_1^2 v_{s\tau}$ we have by Step 2 applied to u_s^τ modelled according to (5.41)

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

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 (5.43)

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

Plugging these two formulas and the definition (5.40) of $\partial_s\sigma\diamond f_\tau$ and $\sigma_s\diamond\partial_s f_\tau$ into (5.49), we obtain (5.44).

Step 5. We still work under the assumptions of part ii) of the proposition. We claim

$$[\sigma_s] + [a_s] \leqq N, \tag{5.50}$$

$$\|f_{s\tau}\|_{\alpha-2} \lesssim N_0, \tag{5.51}$$

$$\|[\sigma_s, (\cdot)_T]\diamond f_{s\tau}\|_{2\alpha-2} \lesssim NN_0, \tag{5.52}$$

$$\|[a_s, (\cdot)]\diamond\partial_1^2 v_{s\tau}\|_{2\alpha-2,2} \lesssim NN_0, \tag{5.53}$$

and on the corresponding estimates on the infinitesimal level

$$[\partial_s\sigma_s] + \|\partial_s\sigma_s\| + [\partial_s a] + \|\partial_s a\| \leqq \delta N, \tag{5.54}$$

$$\|\partial_s f_\tau\|_{\alpha-2} \leqq \delta N_0, \tag{5.55}$$

$$\|[\sigma_s, (\cdot)]\diamond\partial_s f_\tau\|_{2\alpha-2} \lesssim N\delta N_0, \tag{5.56}$$

$$\|[\partial_s\sigma, (\cdot)]\diamond f_{s\tau}\|_{2\alpha-2} \lesssim \delta NN_0, \tag{5.57}$$

$$\|[a_s, (\cdot)]\diamond\partial_1^2\partial_s v_\tau\|_{\alpha-2,1} \lesssim N\delta N_0, \tag{5.58}$$

$$\|[\partial_s a_s, (\cdot)]\diamond\partial_1^2 v_{s\tau}\|_{2\alpha-2,1} \lesssim \delta NN_0. \tag{5.59}$$

Indeed, (5.50) and (5.54) are immediate from our assumptions (3.43) (with i) and (3.73), respectively, by the linear interpolation (5.37). For $\tau = 0$ the remaining estimates, even with $\lesssim$ replaced by $\leqq$, follow from the linear and bilinear interpolations (5.37) and (5.38) from the assumptions of this proposition: inequality (5.51) from (3.44) (with i), (5.52) from (3.71), (5.53) from (3.72). Still, for $\tau = 0$, the five estimates-(5.55), (5.56), (5.57), (5.58) and (5.59) are direct consequences of (3.74), (3.75), (3.76), (3.77) and (3.78), respectively.

It remains to pass from $\tau = 0$ to $0 < \tau \leqq 1$ in the eight estimates of this step, based on our definition (5.40) 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

$$[u] \leqq N_0, \quad \|f\|_{\alpha-2} \leqq N_1 \quad \text{and} \quad \|[u, (\cdot)] \diamond f\|_{2\alpha-2} \leqq N_0 N_1 \tag{5.60}$$

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

$$\|f_\tau\|_{\alpha-2} \lesssim N_1 \quad \text{and} \quad \|[u, (\cdot)] \diamond f_\tau\|_{2\alpha-2} \lesssim N_0 N_1. \tag{5.61}$$

Indeed, by the definition of $u \diamond f_\tau$, we have, as for (5.33),

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

so that (5.61) follows automatically provided we can show that (5.60) extend from the range of $T \leqq 1$ to the range $T \leqq 2$ in form of

$$\sup_{T \leqq 1} (T^{\frac{1}{4}})^{2-\alpha} \|f_{2T}\| \lesssim N_1, \quad \sup_{T \leqq 1} (T^{\frac{1}{4}})^{2-2\alpha} \|[u, (\cdot)_{2T}] \diamond f\| \lesssim N_0 N_1. \tag{5.62}$$

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 the last item in (5.60) entails (5.62), appealing to (5.120) and using, in addition, that by the first items in (5.60),

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

STEP 7. Application of Corollary 3.7. We claim for the modelling and Hölder constants of u_s^τ and $\partial_s u^\tau$ that

$$M_s^\tau \lesssim N_0 N, \tag{5.63}$$

$$[u_s^\tau] \lesssim N_0 (N + 1), \tag{5.64}$$

$$\delta M_s^\tau \lesssim N_0 \delta N + \delta N_0 N \quad \text{provided } N \leqq 1, \tag{5.65}$$

$$[\partial_s u^\tau] \lesssim N_0 \delta N + \delta N_0 \quad \text{provided } N \leqq 1. \tag{5.66}$$

Indeed, for estimates (5.63) and (5.64) we apply Corollary 3.7 i) with (f, v, σ, a, $\sigma \diamond f$, $a \diamond \partial_1^2 v$, u) replaced by ($f_{s\tau}$, $v_{s\tau}$, σ_s, a_s, $\sigma_s \diamond f_{s\tau}$, $a_s \diamond \partial_1^2 v_{s\tau}$, u_s^τ) (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 linear interpolation (5.37) preserves (3.47). The hypotheses (3.43), (3.44), (3.45) and (3.46) were established in Step 5, cf. (5.50), (5.51), (5.52) and (5.53), respectively. Hypothesis (3.48) and the modelledness are clear by construction, cf.

(5.42) and (5.41) in Step 4. The outputs (3.49) and (3.50) assume the form (5.63) and (5.64).

For the remaining estimates (5.65) and (5.66), we apply Corollary 3.7 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 $(\partial_s f_\tau, \partial_s v_\tau, \partial\sigma, \partial_s a, \sigma_s\diamond\partial_s f_\tau, \partial_s\sigma\diamond f_{s\tau}, a_s\diamond\partial_1^2\partial_s v_\tau, \partial_s a\diamond\partial_1^2 v_{s\tau}, \partial_s u^\tau)$. The six hypotheses (3.51)–(3.56) were established in Step 5, cf. (5.54)–(5.59). Hypothesis (3.58) and the corresponding modelledness are clear by construction, cf. (5.44) and (5.43) in Step 4. The outputs (3.59) and (3.60) assume the form of (5.65) and (5.66).

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:

$$\delta M^\tau \lesssim N_0\delta N + \delta N_0 N \quad \text{provided } N \leqq 1, \tag{5.67}$$

$$[u_1^\tau - u_0^\tau] \lesssim N_0\delta N + \delta N_0 \quad \text{provided } N \leqq 1. \tag{5.68}$$

Indeed, the Hölder estimate (5.68) is obvious from (5.66) by integration in $s \in [0,1]$. The estimate on the modelling constant relies on the differentiation rule

$$\begin{aligned}\frac{\partial}{\partial s}\big(u_s^\tau(y) - \sigma_s(x)v_{s\tau}(y, a_s(x))\big) &= \partial_s u^\tau(y) - (\partial_s\sigma)(x)v_{s\tau}(y, a_s(x)) \\ &\quad - (\sigma_s\partial_s a)(x)\frac{\partial v_{s\tau}}{\partial a_0}(y, a_s(x)) - \sigma_s(x)\partial_s v_\tau(y, a_s(x)),\end{aligned}$$

and on defining $\nu := \int_0^1 \nu_s \,\mathrm{d}s$, where ν belongs to $u_1^\tau - u_0^\tau$ and ν_s to $\partial_s u^\tau$ in the sense of Definition 3.1. This provides the link between (5.65) and (5.67) by integration.

Step 9. Passage to limit. We claim that we may pass to the limit $\tau \downarrow 0$ in (5.63) and (5.64) with $s = 0, 1$, recovering (3.69) and (3.70) in part i) of this proposition, and in (5.67) and (5.68), recovering (3.79) and (3.80) in part ii) of the proposition. Clearly, from the uniform-in-τ estimate (5.64) (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 (5.64) and (5.68). Since we also have that the convolution $v_{i\tau}$ converges to v_i uniformly, we may pass to the limit in the estimates (5.63) and (5.67) of the modelling constants. By uniqueness, cf. Step 1, it thus remains to argue that u_i solves (3.68) (with (f, σ, a) replaced by (f_i, σ_i, a_i)). In order to pass from (5.42) to (3.68) it remains to establish the distributional convergences

$$\sigma_i \diamond f_{i\tau} \rightharpoonup \sigma_i \diamond f_i, \tag{5.69}$$

$$a_i \diamond \partial_1^2 u_i^\tau \rightharpoonup a_i \diamond \partial_1^2 u_i. \tag{5.70}$$

The convergence (5.69) is built-in by the definition (5.40) through convolution. One of the ingredients for the convergence (5.70) is the analogue of (5.69)

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

which in conjunction with the uniform 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 $\sup_{a_0} \|\frac{\partial}{\partial a_0}[a_i, (\cdot)_T] \diamond \partial_1^2 v_{i\tau}(\cdot, a_0)\|$ is uniformly bounded, cf. (3.72) and (5.40) in conjunction with a formula of type (5.33), 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 (5.70) we appeal to the modelledness of u_i with respect to v_i according to a_i and σ_i which by (3.36) in Lemma 3.5 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. (5.63), in conjunction with the uniform commutator bounds (3.46) and the uniform bounds on $v_{i\tau}$, we have, again by (3.36) in Lemma 3.5, 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

$$\lim_{T\downarrow 0} \limsup_{\tau\downarrow 0} \|(a_i \diamond \partial_1^2 u_i^\tau - a_i \diamond \partial_1^2 u_i)_T\| = 0. \tag{5.71}$$

Now the next step shows that this implies (5.70).

STEP 10. Let a sequence of distributions $\{f_n\}_{n\uparrow\infty}$ be bounded wrt $\|\cdot\|_{\alpha-2}$; then we claim

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

Indeed, we have for fixed $T > 0$ and any $\tau \leqq T$ that $\|f_{nT}\| \lesssim \|f_{n\tau}\|$ and therefore $\limsup_{n\uparrow\infty} \|f_{nT}\| \lesssim \limsup_{n\uparrow\infty} \|f_{n\tau}\|$ and $\limsup_{n\uparrow\infty} \|f_{nT}\| \lesssim \lim_{\tau\downarrow 0} \limsup_{n\uparrow\infty} \|f_{n\tau}\|$. 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 boundedness of f_n wrt $\|\cdot\|_{\alpha-2}$, and then also in the more classical $C^{\alpha-2}$-norm, cf. (A.2) in Step 1 of Lemma A.1.

5.3. Proof of Corollary 3.7

We write $[\cdot]$ for $[\cdot]_\alpha$.

STEP 1. Application of Lemma A.1. We claim

$$[v]_2 \lesssim N_0, \tag{5.72}$$

$$[\delta v]_1 \lesssim \delta N_0, \tag{5.73}$$

where we recall the notational convention (3.18) for the a_0-derivatives. The estimate (5.72) is based on the identities following from differentiating (1.2) twice with respect to a_0:

$$(\partial_2 - a_0\partial_1^2)\left\{v, \frac{\partial v}{\partial a_0}, \frac{\partial^2 v}{\partial a_0^2}\right\} = \left\{Pf, \partial_1^2 v, 2\partial_1^2 \frac{\partial v}{\partial a_0}\right\}. \tag{5.74}$$

We now see that (5.72) follows by an iterated application of Lemma A.1. From (3.44) we first obtain the bound on $[v]$ by Lemma A.1, then the bound on $\|\partial_1^2 v\|_{\alpha-2}$ by (2.4), then via (5.74) the bound on $[\frac{\partial v}{\partial a_0}]$ by Lemma A.1, then the bound on $\|\partial_1^2 \frac{\partial v}{\partial a_0}\|_{\alpha-2}$ by (2.4), then via (5.74) finally the bound on $[\frac{\partial^2 v}{\partial a_0^2}]$ by Lemma A.1. The argument for (5.73) is identical, just with (f, v) replaced by $(\delta f, \delta v)$, cf. (3.57), and starting from (3.52) instead of (3.44) and thus with N_0 replaced by δN_0.

STEP 2. Application of Lemma 3.5. We claim that

$$\|[a, (\cdot)] \diamond \partial_1^2 u\|_{2\alpha-2} \lesssim [a]M + NN_0, \tag{5.75}$$

$$\|[\delta a, (\cdot)] \diamond \partial_1^2 u\|_{2\alpha-2} \lesssim [\delta a]M + \delta NN_0, \tag{5.76}$$

$$\|[a, (\cdot)] \diamond \partial_1^2 \delta u\|_{2\alpha-2} \lesssim [a]\delta M + N(N_0\delta N + \delta N_0). \tag{5.77}$$

The argument is as follows: estimate (5.75) follows from Lemma 3.5 with b replaced by a, $I = 1$ and $v_{i=1} = v$, so that the hypothesis (3.34) is satisfied by (5.72) in Step 1 with N_0 playing the role of $N_{i=1}$. Hypothesis (3.35) is satisfied by our assumption (3.46) with N playing the role of N_0. In view of (3.47), the outcome (3.37) of Lemma 3.5 turns into (5.75).

Estimate (5.76) follows from applying Lemma 3.5 with b replaced by δa, still $I = 1$, $v_{i=1} = v$, and N_0 playing the role of $N_{i=1}$. Hypothesis (3.35) is satisfied by our assumption (3.56) with δN playing the role of N_0. In view of (3.47), the outcome (3.37) of Lemma 3.5 turns into (5.76).

Finally, estimate (5.77) follows from applying Lemma 3.5 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 (3.34) is satisfied with $(N_1, N_2, N_3) = (N_0, N_0, \delta N_0)$. We now turn to the hypothesis (3.35): For $i = 1$ it is contained in our assumption (3.46) with N playing the role of N_0. In preparation of checking hypothesis (3.35) for $i = 2$ we note that our assumption (3.46) 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

$$[a,(\cdot)_T]\diamond\partial_1^2\frac{\partial v}{\partial a_0}=\frac{\partial}{\partial a_0}[a,(\cdot)_T]\diamond\partial_1^2 v. \tag{5.78}$$

Hence the hypothesis (3.35) for $i=2$ is also satisfied by (3.46) (here we use it up to $\frac{\partial^2}{\partial a_0^2}$). Hypothesis (3.35) for $i=3$ is identical to our assumption (3.55). We apply Lemma 3.5 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 (3.37) of Lemma 3.5 assumes the form

$$\begin{aligned}&\|[a,(\cdot)]\diamond\partial_1^2\delta u\|_{2\alpha-2}\\&\quad\lesssim[a]\delta M+N\big(N_0([\delta\sigma]+\|\delta\sigma\|+[\sigma\delta a]+\|\sigma\delta a\|)+\delta N_0([\sigma]+\|\sigma\|)\big).\end{aligned} \tag{5.79}$$

We note that by (3.47) and (3.51) we have

$$\begin{aligned}&N_0([\delta\sigma]+\|\delta\sigma\|+[\sigma\delta a]+\|\sigma\delta a\|)+\delta N_0([\sigma]+\|\sigma\|)\\&\quad\lesssim N_0([\delta\sigma]+\|\delta\sigma\|+[\delta a]+\|\delta a\|)+\delta N_0\lesssim N_0\delta N+\delta N_0,\end{aligned}$$

so that (5.79) yields (5.77).

STEP 3. Commutator estimates. We claim

$$\sup_{T\leqq 1}(T^{\frac14})^{2-2\alpha}\|\partial_2 u_T-P(a\partial_1^2 u_T+\sigma f_T)\|\lesssim[a]M+NN_0, \tag{5.80}$$

$$\begin{aligned}&\sup_{T\leqq 1}(T^{\frac14})^{2-2\alpha}\|\partial_2\delta u_T-P(a\partial_1^2\delta u_T+\sigma\delta aE\partial_1^2 v_T+\sigma\delta f_T+\delta\sigma f_T)\|\\&\quad\lesssim[a]\delta M+([\delta a]+\|\delta a\|)M+N(N_0\delta N+\delta N_0)+\delta NN_0.\end{aligned} \tag{5.81}$$

Indeed, we apply $(\cdot)_T$ to (3.48) and rearrange terms to get

$$\partial_2 u_T-P(a\partial_1^2 u_T+\sigma f_T)=-P([a,(\cdot)_T]\diamond\partial_1^2 u+[\sigma,(\cdot)_T]\diamond f). \tag{5.82}$$

Similarly, we apply $(\cdot)_T$ to (3.58) and rearrange terms to obtain

$$\begin{aligned}&\partial_2\delta u_T-P(a\partial_1^2\delta u_T+\sigma\delta aE\partial_1^2 v_T+\sigma\delta f_T+\delta\sigma f_T)\\&\quad=-P\Big(-\delta a(\partial_1^2 u_T-\sigma E\partial_1^2 v_T)+[a,(\cdot)_T]\diamond\partial_1^2\delta u+[\delta a,(\cdot)_T]\diamond\partial_1^2 u\\&\qquad+[\sigma,(\cdot)_T]\diamond\delta f+[\delta\sigma,(\cdot)_T]\diamond f\Big).\end{aligned} \tag{5.83}$$

By assumption (3.45) and by (5.75) in Step 2 we obtain estimate (5.80) from identity (5.82). By assumptions (3.53) and (3.54) and by (5.76) and (5.77) from Step 2 and from writing

$$\begin{aligned}&(\partial_1^2 u_T-\sigma E\partial_1^2 v_T)(x)=\int dy\,\partial_1^2\psi_T(x-y)\\&\quad\times\big((u(y)-u(x))-\sigma(x)(v(y,a(x))-v(x,a(x)))-\nu(x)(y-x)_1\big),\end{aligned}$$

which entails, with help of (2.4) and (3.1), that

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

and we obtain (5.81) from formula (5.83).

STEP 4. Application of Lemma 3.6 and conclusion. We first apply Lemma 3.6 with $I = 1$ and f playing the role of $f_{i=1}$ (which does not depend on a_0). The hypothesis (3.38) is ensured by our assumption (3.44) with N_0 playing the role of $N_{i=1}$. The hypothesis (3.40) is settled through (5.80) in Step 3 with N^2 given by $[a]M + NN_0$. Hence the two outputs (3.41) and (3.42) of Lemma 3.6 take the form of

$$M \lesssim [a]M + NN_0 + N_0([\sigma] + \|\sigma\|[a]), \tag{5.84}$$

$$[u] \lesssim M + N_0\|\sigma\|. \tag{5.85}$$

The smallness of $[a]$ and the boundedness of $\|\sigma\|$, cf. (3.47), imply that (5.84) simplifies to $M \lesssim NN_0 + N_0([\sigma] + [a])$, which by (3.43) means (3.49). Inserting (3.49) into (5.85) and using once more $\|\sigma\| \leqq 1$ yields (3.50).

We now apply Lemma 3.6 with $I = 3$ and $(f, \partial_1^2 v, \delta f)$ playing the role of (f_1, f_2, f_3); by assumptions (3.44), (3.52) and by (5.72), this triplet satisfies (3.38) with $(N_1, N_2, N_3) = (N_0, N_0, \delta N_0)$. In view of (5.74) in Step 1, and of assumption (3.57), the triplet $(v, \frac{\partial v}{\partial a_0}, \delta v)$ plays the role of (v_1, v_2, v_3) in the sense of (3.39). We apply Lemma 3.6 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 (3.40) is settled through Step 3 with N^2 estimated by the right hand side of (5.81). Hence the two outputs (3.41) and (3.42) of Lemma 3.6 take the form

$$\begin{aligned}\delta M &\lesssim \text{expression on right hand side of (5.81)} + [a]\delta M\\ &\quad + N_0([\delta\sigma] + \|\delta\sigma\|[a] + [\sigma\delta a] + \|\sigma\delta a\|[a]) + \delta N_0([\sigma] + \|\sigma\|[a]),\\ [\delta u] &\lesssim \delta M + N_0(\|\delta\sigma\| + \|\sigma\delta a\|) + \delta N_0\|\sigma\|.\end{aligned}$$

Making use of the constraints (3.47) on σ and a, in particular to absorb $[a]\delta M$ into the lhs, this simplifies to

$$\begin{aligned}\delta M &\lesssim ([\delta a] + \|\delta a\|)M + N(N_0\delta N + \delta N_0) + \delta N N_0\\ &\quad + N_0([\delta\sigma] + \|\delta\sigma\| + [\delta a] + \|\delta a\|) + \delta N_0([\sigma] + [a]),\\ [\delta u] &\lesssim \delta M + N_0(\|\delta\sigma\| + \|\delta a\|) + \delta N_0.\end{aligned}$$

Inserting (3.43) and (3.51), this reduces to

$$\delta M \lesssim M\delta N + N(N_0\delta N + \delta N_0) + N_0\delta N, \tag{5.86}$$

$$[\delta u] \lesssim \delta M + N_0\delta N + \delta N_0. \tag{5.87}$$

Making use of the estimate (3.49) on M we just established, (5.86) implies

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

Clearly, this estimate implies the desired (3.59). Plugging (3.59) into (5.87) yields the desired (3.60).

5.4. *Proof of Lemma 3.6*

All functions are periodic if not stated otherwise.

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

$$[v_i]_{\alpha,1} \lesssim N_i, \tag{5.88}$$

where we recall the abbreviation (3.18). This follows immediately from assumption (3.38) on f_i and the definition (3.39) of v_i via Lemma A.1 and the argument of Step 1 of Corollary 3.7.

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

$$(\partial_2 - a(x_0)\partial_1^2)\big(u_T - \sigma_i(x_0)v_{iT}(\cdot, a(x_0))\big) = Pg_{x_0}^T, \tag{5.89}$$

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

$$|g_{x_0}^T(x)| \lesssim \tilde{N}^2 \left((T^{\frac14})^{2\alpha-2} + (T^{\frac14})^{\alpha-2} d^\alpha(x, x_0)\right) \quad \text{for } T \leqq 1, \tag{5.90}$$

with the abbreviation

$$\tilde{N}^2 := N^2 + [a]_\alpha [u]_\alpha + N_i([\sigma_i]_\alpha + \|\sigma_i\|[a]_\alpha). \tag{5.91}$$

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

$$(\partial_2 - a(x_0)\partial_1^2)u_T = P\left(\sigma_i(x_0) f_{iT}(\cdot, a(x_0)) + g_{x_0}^T\right), \tag{5.92}$$

with $g_{x_0}^T$ defined through

$$\begin{aligned} g_{x_0}^T &:= \partial_2 u_T - P(a\partial_1^2 u_T + \sigma_i E f_{iT}) + (a - a(x_0))\partial_1^2 u_T \\ &\quad + (\sigma_i - \sigma_i(x_0))E f_{iT} + \sigma_i(x_0)(E f_{iT} - f_{iT}(\cdot, a(x_0))). \end{aligned} \tag{5.93}$$

By definition (3.39) 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 obtain

$$(\partial_2 - a(x_0)\partial_1^2)\sigma_i(x_0)v_{iT}(\cdot, a(x_0)) = P\sigma_i(x_0) f_{iT}(\cdot, a(x_0)). \tag{5.94}$$

From the combination of (5.92) and (5.94) we obtain (5.89), so that it remains to estimate $g_{x_0}^T$. Making use of the assumption (3.40), we obtain from (5.93) that

$$\begin{aligned} |g_{x_0}^T(x)| \leqq N^2(T^{\frac14})^{2\alpha-2} + d^\alpha(x, x_0) \Bigg(&[a]_\alpha \|\partial_1^2 u_T\| \\ &+ [\sigma_i]_\alpha \sup_{a_0} \|f_{iT}\| + \|\sigma_i\|[a]_\alpha \sup_{a_0} \|(\frac{\partial f_i}{\partial a_0})_T\| \Bigg), \end{aligned}$$

so that by (2.4) and by assumption (3.38),

$$|g_{x_0}^T(x)| \lesssim N^2(T^{\frac14})^{2\alpha-2} + (T^{\frac14})^{\alpha-2} d^\alpha(x, x_0)\big([a]_\alpha [u]_\alpha + N_i([\sigma_i]_\alpha + \|\sigma_i\|[a]_\alpha)\big),$$

which can be consolidated into the estimate (5.90).

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

$$\begin{aligned} &\frac{1}{R^{2\alpha}} \inf_{\ell} \|u_T - \sigma_i(x_0) v_{iT}(\cdot, a(x_0)) - \ell\|_{B_R(x_0)} \\ &\quad \lesssim \left(\frac{R}{L}\right)^{2(1-\alpha)} \frac{1}{L^{2\alpha}} \inf_{\ell} \|u_T - \sigma_i(x_0) v_{iT}(\cdot, a(x_0)) - \ell\|_{B_L(x_0)} \\ &\qquad + \tilde{N}^2 \left(\frac{L^2}{R^{2\alpha}(T^{\frac{1}{4}})^{2-2\alpha}} + \frac{L^{2+\alpha}}{R^{2\alpha}(T^{\frac{1}{4}})^{2-\alpha}} \right), \end{aligned} \tag{5.95}$$

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 (2.1) with center x_0 and radius R. This step mimics the heart of the kernel-free approach of Safonov to the classical Schauder theory, see [14, Theorem 8.6.1]. The argument for this is as follows 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

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

where $I(B_L)$ denotes the indicator function of the set B_L. Hence in view of (5.89), where we write $P g_0^T = g_0^T + c$ with $c = -\int_{[0,1)^2} g_0^T$, the function

$$w_< := u_T - \sigma_i(0) v_{iT}(\cdot, a(0)) - w_> \tag{5.96}$$

satisfies

$$(\partial_2 - a(0)\partial_1^2) w_< = c \quad \text{in } B_L. \tag{5.97}$$

By standard estimates for the heat equation we have

$$\|w_>\| \lesssim L^2 \|g_0^T\|_L, \tag{5.98}$$

$$\|\{\partial_1^2, \partial_2\} w_<\|_{\frac{L}{2}} \lesssim L^{-2} \|w_< - \ell_L\|_L \tag{5.99}$$

for any function $\ell_L \in \operatorname{span}\{1, x_1\}$. The interior estimate (5.99) is slightly non-standard because of the non-vanishing right hand side c but can be easily reduced to the case of $c = 0$. First of all, replacing w by $w - \ell_L$ in (5.97) and (5.99) we may reduce to the case of $\ell_L = 0$. Testing (5.97) 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 $w + c x_2$ which reduces the further estimate to the standard case of $c = 0$. We refer to [14, Theorem 8.4.4] for an elementary argument for (5.99) in case of $c = 0$ only relying on the maximum principle via Bernstein's argument. We refer to [14, Exercise 8.4.8] for the statement (5.98) via the representation through the heat kernel. Since by construction, cf. (5.96), we have $u_T - \sigma_i(0) v_{iT}(\cdot, a(0)) = w_< + w_>$ we obtain by the triangle inequality for a suitably chosen $\ell_R \in \operatorname{span}\{1, x_1\}$

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

Inserting (5.99) for $R \ll L$, and by another application of the triangle inequality this yields

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

Inserting (5.98) and (5.90), this gives

$$\begin{aligned}
&\inf_{\ell} \|u_T - \sigma_i(0)v_{iT}(\cdot, a(0)) - \ell\|_R \\
&\quad \lesssim L^{-2}R^2 \inf_{\ell} \|u_T - \sigma_i(0)v_{iT}(\cdot, a(0)) - \ell\|_L \\
&\quad \quad + \tilde{N}^2L^2\left((T^{\frac{1}{4}})^{2\alpha-2} + L^{\alpha}(T^{\frac{1}{4}})^{\alpha-2}\right),
\end{aligned} \tag{5.100}$$

where we recall that ℓ runs over span$\{1, x_1\}$. Dividing by $R^{2\alpha}$ gives (5.95).

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

$$M \lesssim M', \tag{5.101}$$

where we have set for abbreviation

$$M' := \sup_{x_0} \sup_{R \leqq 1} R^{-2\alpha} \inf_{\ell} \|u - \sigma_i(x_0)v_i(\cdot, a(x_0)) - \ell\|_{B_R(x_0)} \tag{5.102}$$

and where the maximal radius 1 is chosen such that a ball of that covers a periodic cell. In fact, also the reverse estimate holds, highlighting once more that the modulation function ν in the definition of modelledness (Definition 3.1) plays a small role compared to σ_i. The equivalence of (5.101) and (5.102) on the level of standard Hölder spaces is the starting point for the approach to Schauder theory by Safonov, see [14, Theorem 8.5.2]. We first argue that the ℓ in (5.102) may be chosen to be independent of R, that is,

$$\sup_{x_0} \inf_{\ell} \sup_{R \leqq 1} R^{-2\alpha}\|u - \sigma_i(x_0)v_i(\cdot, a(x_0)) - \ell\|_{B_R(x_0)} \lesssim M'. \tag{5.103}$$

Indeed, fix x_0, say $x_0 = 0$, and let $\ell_R = \nu_R x_1 + c_R$ be (near) optimal in (5.102), 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 $R' \leqq R$ 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

$$R^{-2\alpha}\|\ell_R - \ell\|_R \lesssim M'. \tag{5.104}$$

Hence we may pass from (5.102) to (5.103) by the triangle inequality.

It is clear from (5.103) 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 (5.101), namely the modelledness (3.1) for any "base" point x and any y of distance at most 1. Since $B_1(x)$ covers a periodic cell, 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 $\alpha \geqq \frac{1}{2}$, this implies that $|\nu(x)(x-y)_1| \lesssim M' d^{2\alpha}(x, y)$ for all $y \notin B_1(x)$. Hence once again by periodicity of $y \mapsto (u(y) - u(x)) - \sigma_i(x)(v_i(y, a(x)) - v_i(x, a(x)))$, (3.1) holds also for $y \notin B_1(x)$.

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

$$\frac{1}{(T^{\frac{1}{4}})^{2\alpha}} \|(u_T - u) - \sigma_i(x_0)(v_{iT} - v_i)(\cdot, a(x_0))\|_{B_L(x_0)} \lesssim M + \tilde{N}^2 \left(\frac{L}{T^{\frac{1}{4}}}\right)^{\alpha}. \tag{5.105}$$

Wlog we consider $x_0 = 0$ and recall that the first moment of ψ_T vanishes, so that

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

We split the right hand side into three terms:

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

For the first right-hand-side term we appeal to the modelledness assumption (3.1), which implies that the integrand is estimated by $|\psi_T(x-y)|\ M\ d^{2\alpha}(x, y)$. Hence by (2.4) 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 (2.4) and (5.88) 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)

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

we see that the integrand of the third right-hand-side term is estimated by $|\psi_T(x-y)| \|\sigma_i\|\ d^\alpha(x, y)\ [a]_\alpha\ \sup_{a_0}[\frac{\partial v_i}{\partial a_0}(\cdot, a_0)]_\alpha\ d^\alpha(x, 0)$; hence in view of (5.88) 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 (5.91) of $\tilde{N}^2$, this yields (5.105).

STEP 6. Estimate of M. We claim that

$$M \lesssim \tilde{N}^2. \tag{5.106}$$

Indeed, we can now close the argument and to this purpose rewrite (5.95) from Step 3 with help of the triangle inequality as

$$\begin{aligned}
&\frac{1}{R^{2\alpha}} \inf_\ell \|u - \sigma_i(x_0)v_i(\cdot, a(x_0)) - \ell\|_{B_R(x_0)} \\
&\lesssim \left(\frac{R}{L}\right)^{2-2\alpha} \frac{1}{L^{2\alpha}} \inf_\ell \|u - \sigma_i(x_0)v_i(\cdot, a(x_0)) - \ell\|_{B_L(x_0)} \\
&\quad + \tilde{N}^2 \left(\frac{L^2}{R^{2\alpha}(T^{\frac{1}{4}})^{2-2\alpha}} + \frac{L^{2+\alpha}}{R^{2\alpha}(T^{\frac{1}{4}})^{2-\alpha}}\right) \\
&\quad + \left(\frac{T^{\frac{1}{4}}}{R}\right)^{2\alpha} \frac{1}{(T^{\frac{1}{4}})^{2\alpha}} \|(u_T - u) - \sigma_i(x_0)(v_{iT} - v_i)(\cdot, a(x_0))\|_{B_L(x_0)}.
\end{aligned}$$

We now insert (5.105) from Step 5 to obtain

$$\begin{aligned}
&\frac{1}{R^{2\alpha}} \inf_\ell \|u - \sigma_i(x_0)v_i(\cdot, a(x_0)) - \ell\|_{B_R(x_0)} \\
&\lesssim \left(\frac{R}{L}\right)^{2-2\alpha} M + \tilde{N}^2 \left(\frac{L^2}{R^{2\alpha}(T^{\frac{1}{4}})^{2-2\alpha}} + \frac{L^{2+\alpha}}{R^{2\alpha}(T^{\frac{1}{4}})^{2-\alpha}}\right) \\
&\quad + \left(\frac{T^{\frac{1}{4}}}{R}\right)^{2\alpha} M + \tilde{N}^2 \frac{L^\alpha(T^{\frac{1}{4}})^\alpha}{R^{2\alpha}}.
\end{aligned} \tag{5.107}$$

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 $R \leqq 1$ in (5.107) via $T^{\frac{1}{4}} = \varepsilon R$ (so that in particular as required $T \leqq 1$ since we think of $\varepsilon \ll 1$) and $L = \varepsilon^{-1} R$, taking the supremum over $R \leqq 1$ and x_0 yields by definition (5.102) of M'

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

By (5.101) in Step 4, this implies

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

Since $0 < \alpha < 1$, we may choose ε sufficiently small such that the first right-hand-side 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, (3.41) and (3.42) immediately follow from the combination of

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

The first estimate is identical to (5.106) in Step 6 into which we plug the definition (5.91) of $\tilde{N}$. The second estimate is an application of Step 2 in the proof of Lemma 3.3 with $v(y,x) := \sigma_i(x) v_i(y, a_i(x))$, so that the hypothesis (3.15) holds with N replaced by $\|\sigma_i\| N_i$, cf. (5.88) in Step 1.

5.5. *Proof of Lemma 3.3*

We write for abbreviation $[\cdot] := [\cdot]_\alpha$ and $E := E_{\text{diag}}$.

STEP 1. We claim

$$[\nu]_{2\alpha-1} \lesssim M + N. \tag{5.108}$$

Indeed, introducing $\ell_x(y) := \nu(x) y_1$ we see that (3.13) can be rewritten as

$$|(u - v(\cdot, x) - \ell_x)(y) - (u - v(\cdot, x) - \ell_x)(x)| \leqq M d^{2\alpha}(y, x),$$

so that we obtain by the triangle inequality

$$|(u - v(\cdot, x) - \ell_x)(y) - (u - v(\cdot, x) - \ell_x)(y')| \leqq M(d^{2\alpha}(y, x) + d^{2\alpha}(y', x)). \tag{5.109}$$

In combination with (3.10) this yields by the triangle inequality

$$\begin{aligned} &|(u - v(\cdot, x') - \ell_x)(y) - (u - v(\cdot, x') - \ell_x)(y')| \\ &\quad \leqq M(d^{2\alpha}(y, x) + d^{2\alpha}(y', x)) + N d^\alpha(x, x') d^\alpha(y, y'). \end{aligned}$$

We now take the difference of this with (5.109) with x replaced by x' to obtain, once more by the triangle inequality,

$$\begin{aligned} &|(\ell_x - \ell_{x'})(y) - (\ell_x - \ell_{x'})(y')| \\ &\quad \leqq M\big(d^{2\alpha}(y, x) + d^{2\alpha}(y', x) + d^{2\alpha}(y, x') + d^{2\alpha}(y', x')\big) \\ &\qquad + N d^\alpha(x, x') d^\alpha(y, y'). \end{aligned}$$

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

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

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 (5.108).

STEP 2. Under our additional assumption (3.15) we claim

$$[u] + \|\nu\| \lesssim M + N. \tag{5.110}$$

By the triangle inequality on (3.13) we obtain for all pairs of points $|\nu(x)(x-y)_1| \leqq |u(x)-u(y)| + [v(\cdot,x)]d^{\alpha}(y,x) + Md^{2\alpha}(x,y)$. Choosing $y = x + (1,0)$, appealing to the space-time periodicity of u, taking the supremum over x, and appealing to (3.15), this turns into the ν-part of (5.110):

$$\|\nu\| \lesssim M + N. \tag{5.111}$$

We now consider pairs of points (x,y) with $d(x,y) \leqq 1$. By the triangle inequality from (3.13) we get

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

By space-time periodicity, this extends to all pairs so that

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

Inserting (5.111) into this yields the u-part of (5.110).

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

$$\begin{aligned}&\big(uf_T - E[v,(\cdot)_T]\diamond f - \nu[x_1,(\cdot)_T]f\big) - \big(uf_\tau - E[v,(\cdot)_\tau]\diamond f - \nu[x_1,(\cdot)_\tau]f\big)_{T-\tau}\\ &\quad = \sum_{\tau \leqq t < T}\Big(\big([u,(\cdot)_t] - E[v,(\cdot)_t] - \nu[x_1,(\cdot)_t]\big)f_t\\ &\qquad - [\nu,(\cdot)_t][x_1,(\cdot)_t]f - [E,(\cdot)_t][v,(\cdot)_t]\diamond f\Big)_{T-2t},\end{aligned} \tag{5.112}$$

where the sum runs over the dyadic "times" $t = \frac{T}{2}, \frac{T}{4}, \ldots, \tau$. By telescoping based on the semi-group property (2.3) this reduces to

$$\begin{aligned}&\big(uf_{2t} - E[v,(\cdot)_{2t}]\diamond f - \nu[x_1,(\cdot)_{2t}]f\big) - \big(uf_t - E[v,(\cdot)_t]\diamond f - \nu[x_1,(\cdot)_t]f\big)_t\\ &\quad = \big([u,(\cdot)_t] - E[v,(\cdot)_t] - \nu[x_1,(\cdot)_t]\big)f_t - [\nu,(\cdot)_t][x_1,(\cdot)_t]f\\ &\qquad - [E,(\cdot)_t][v,(\cdot)_t]\diamond f,\end{aligned}$$

and splits into the three statements

$$\begin{aligned} uf_{2t} - (uf_t)_t &= [u,(\cdot)_t]f_t,\\ \nu[x_1,(\cdot)_{2t}]f - (\nu[x_1,(\cdot)_t]f)_t &= \nu[x_1,(\cdot)_t]f_t + [\nu,(\cdot)_t][x_1,(\cdot)_t]f,\\ E[v,(\cdot)_{2t}]\diamond f - (E[v,(\cdot)_t]\diamond f)_t &= E[v,(\cdot)_t]f_t + [E,(\cdot)_t][v,(\cdot)_t]\diamond f.\end{aligned} \tag{5.113}$$

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

$$[x_1,(\cdot)_{2t}]f - ([x_1,(\cdot)_t]f)_t = [x_1,(\cdot)_t]f_t. \tag{5.114}$$

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

$$[v, (\cdot)_{2t}] \diamond f - ([v, (\cdot)_t] \diamond f)_t = [v, (\cdot)_t] f_t, \tag{5.115}$$

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

$$v f_{2t} - (v f_t)_t = [v, (\cdot)_t] f_t \quad \text{and} \quad (v \diamond f)_{2t} - ((v \diamond f)_t)_t = 0. \tag{5.116}$$

Now identities (5.113), (5.114), and (5.116) follow immediately from the semigroup property.

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

$$\begin{aligned} &\|\big(u f_T - E[v, (\cdot)_T] \diamond f - \nu [x_1, (\cdot)_T] f\big) \\ &\quad - \big(u f_\tau - E[v, (\cdot)_\tau] \diamond f - \nu [x_1, (\cdot)_\tau] f\big)_{T-\tau}\| \\ &\lesssim (M + N) N_1 (T^{\frac{1}{4}})^{3\alpha - 2}. \end{aligned} \tag{5.117}$$

Indeed, by the dyadic representation (5.112), the triangle inequality in $\|\cdot\|$ and the fact that $(\cdot)_{T-2t}$ is bounded in that norm, cf. (2.4), it is enough to show that the right-hand-side term of (5.112) under the parenthesis is estimated by $(M+N)N_1$ $(t^{\frac{1}{4}})^{3\alpha-2}$ for all $t \leqq 1$; 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 (5.118) by $M + N$, this estimate splits into

$$\begin{aligned} \|\big([u, (\cdot)_t] - E[v, (\cdot)_t] - \nu[x_1, (\cdot)_t]\big) f_t\| &\lesssim M N_1 (t^{\frac{1}{4}})^{3\alpha-2}, \\ \|[\nu, (\cdot)_t][x_1, (\cdot)_t] f\| &\lesssim [\nu]_{2\alpha-1} N (t^{\frac{1}{4}})^{3\alpha-2}, \end{aligned} \tag{5.118}$$

$$\|[E, (\cdot)_t][v, (\cdot)_t] \diamond f\| \lesssim N N_1 (t^{\frac{1}{4}})^{3\alpha-2}. \tag{5.119}$$

Appealing to our assumptions (3.11) and (3.12) and to Lemma A.2, these three estimates reduce to

$$\begin{aligned} &\|\big([u, (\cdot)_t] - E[v, (\cdot)_t] - \nu[x_1, (\cdot)_t]\big) \tilde f\| \lesssim M \|\tilde f\| (t^{\frac{1}{4}})^{2\alpha}, \\ &\|[\nu, (\cdot)_t] \tilde f\| \lesssim [\nu]_\beta \|\tilde f\| (t^{\frac{1}{4}})^{\beta}, \end{aligned} \tag{5.120}$$

$$\|[E, (\cdot)_t] \tilde v\| \lesssim \sup_{x, x'} \frac{1}{d^\alpha(x, x')} \|\tilde v(\cdot, x) - \tilde v(\cdot, x')\| (t^{\frac{1}{4}})^{\alpha}, \tag{5.121}$$

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

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

$$\Big|\int \mathrm{d}y\psi_t(x-y)\big(\nu(x)-\nu(y)\big)\tilde f(y)\Big| \lesssim [\nu]_\beta\|\tilde f\|(t^{\frac14})^\beta,$$
$$\Big|\int \mathrm{d}y\psi_t(x-y)\big(\tilde v(y,x)-\tilde v(y,y)\big)\Big| \lesssim \sup_{x,x'}\frac{1}{d^\alpha(x,x')}\|\tilde v(\cdot,x)-\tilde v(\cdot,x')\|(t^{\frac14})^\alpha.$$

All three estimates rely on the moment bounds (2.4), the first estimate is then an immediate consequence of (3.13) and the two last ones tautological.

STEP 5. For

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

and under our addditional assumptions (3.15) and (3.16), we claim the estimates

$$\sup_{T\leqq 1}(T^{\frac14})^{2-2\alpha}\|uf_T - F^\tau_{T-\tau}\| \lesssim (M+N)N_1, \quad \|F^\tau\|_{\alpha-2}\lesssim (M+N+\|u\|)N_1. \tag{5.122}$$

Indeed, the first item in (5.122) follows from (5.117) in Step 4 via the triangle inequality and

$$\|E[v,(\cdot)]\diamond f\|_{2\alpha-2} \overset{(3.16)}{\leqq} NN_1, \quad \|\nu[x_1,(\cdot)]f\|_{2\alpha-2}\lesssim (M+N)N_1,$$

the latter being a consequence of (5.110) in Step 2, (A.9) in Lemma A.2, and our assumption (3.11); here, we make extensively use of $T \leqq 1$. The second item in (5.122) in turn follows from (5.122) via $\|F^\tau_T\| = \|(F^\tau_{T-\tau})_\tau\| \lesssim \|F^\tau_{T-\tau}\|$ (cf. (2.3) and (2.4)) by the triangle inequality, (5.110), and (3.11), again making use of $T\leqq 1$.

STEP 6. Conclusion. By the second item in (5.122) in Step 5, the sequence $\{F^\tau\}_{\tau\downarrow 0}$ is bounded wrt $\|\cdot\|_{\alpha-2}$. By standard weak compactness based on the equivalence of norms from Step 1 in the proof of Lemma A.1, 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 (5.122) in Step 5 to obtain (3.17). Likewise, we may pass to the limit in (5.117) in Step 4 to obtain (3.14). Note that our additional assumptions (3.15) and (3.16) were only qualitatively used in deriving (3.14) by ensuring the above boundedness of $\{F^\tau\}_{\tau\downarrow 0}$.

5.6. Proof of Lemma 3.5

The proof follows the lines of Steps 3 through 6 of the proof of Lemma 3.3.

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

$$\begin{aligned}&\big(b\partial_1^2 u_T - \sigma_i E[b,(\cdot)_T]\diamond\partial_1^2 v_i\big) - \big(b\partial_1^2 u_\tau - \sigma_i E[b,(\cdot)_\tau]\diamond\partial_1^2 v_i\big)_{T-\tau}\\ &\quad=\sum_{\tau\leqq t<T}\Big(\big([b,(\cdot)_t]\partial_1^2 u_t - \sigma_i E[b,(\cdot)_t]\partial_1^2 v_{it}\big)\\ &\qquad - [\sigma_i,(\cdot)_t]E[b,(\cdot)_t]\diamond\partial_1^2 v_i - \sigma_i[E,(\cdot)_t][b,(\cdot)_t]\diamond\partial_1^2 v_i\Big)_{T-2t},\end{aligned} \tag{5.123}$$

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

$$\begin{aligned}&\big(b\partial_1^2 u_{2t} - \sigma_i E[b,(\cdot)_{2t}]\diamond\partial_1^2 v_i\big) - \big(b\partial_1^2 u_t - \sigma_i E[b,(\cdot)_t]\diamond\partial_1^2 v_i\big)_t\\&\quad = \big([b,(\cdot)_t]\partial_1^2 u_t - \sigma_i E[b,(\cdot)_t]\partial_1^2 v_{it}\big)\\&\qquad - [\sigma_i,(\cdot)_t]E[b,(\cdot)_t]\diamond\partial_1^2 v_i - \sigma_i[E,(\cdot)_t][b,(\cdot)_t]\diamond\partial_1^2 v_i,\end{aligned}$$

and splits into the two statements

$$\begin{aligned}&b\partial_1^2 u_{2t} - (b\partial_1^2 u_t)_t = [b,(\cdot)_t]\partial_1^2 u_t,\\&\sigma_i E[b,(\cdot)_{2t}]\diamond\partial_1^2 v_i - (\sigma_i E[b,(\cdot)_t]\diamond\partial_1^2 v_i)_t = \sigma_i E[b,(\cdot)_t]\partial_1^2 v_{it}\\&\quad + [\sigma_i,(\cdot)_t]E[b,(\cdot)_t]\diamond\partial_1^2 v_i + \sigma_i[E,(\cdot)_t][b,(\cdot)_t]\diamond\partial_1^2 v_i.\end{aligned}\tag{5.124}$$

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

$$[b,(\cdot)_{2t}]\diamond\partial_1^2 v_i - ([b,(\cdot)_t]\diamond\partial_1^2 v_i)_t = [b,(\cdot)_t]\partial_1^2 v_{it}.\tag{5.125}$$

Now (5.124) and (5.125) are consequences of the semi-group property.
Step 2. We claim the estimate

$$\begin{aligned}&\|\big(b\partial_1^2 u_T - \sigma_i E[b,(\cdot)_T]\diamond\partial_1^2 v_i\big) - \big(b\partial_1^2 u_\tau - \sigma_i E[b,(\cdot)_\tau]\diamond\partial_1^2 v_i\big)_{T-\tau}\|\\&\lesssim \big([b]_\alpha M + N_0 N_i([\sigma_i]_\alpha + \|\sigma_i\|[a]_\alpha)\big)(T^{\frac14})^{3\alpha-2}.\end{aligned}$$

In view of (5.123), this estimate splits into

$$\|[b,(\cdot)_t]\partial_1^2 u_t - \sigma_i E[b,(\cdot)_t]\partial_1^2 v_{it}\| \lesssim [b]_\alpha M (t^{\frac14})^{3\alpha-2},\tag{5.126}$$

$$\|[\sigma_i,(\cdot)_t]E[b,(\cdot)_t]\diamond\partial_1^2 v_i\| \lesssim N_0 N_i [\sigma_i]_\alpha (t^{\frac14})^{3\alpha-2},\tag{5.127}$$

$$\|[E,(\cdot)_t][b,(\cdot)_t]\diamond\partial_1^2 v_i\| \lesssim N_0 N_i [a]_\alpha (t^{\frac14})^{3\alpha-2}.\tag{5.128}$$

Estimate (5.127) follows from (5.120) (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 α playing the role of β) and our assumption (3.35) (without $\frac{\partial}{\partial a_0}$). Estimate (5.128) from (5.121) (with $[b,(\cdot)_t]\diamond\partial_1^2 v_i$ playing the role of $\tilde v$) and our assumptions (3.34) and (3.35) (with $\frac{\partial}{\partial a_0}$) give

$$\begin{aligned}&\frac{1}{d^\alpha(x,x')}\|([b,(\cdot)_t]\diamond\partial_1^2 v_i)(\cdot,a(x)) - ([b,(\cdot)_t]\diamond\partial_1^2 v_i)(\cdot,a(x'))\|\\&\quad \leqq [a]_\alpha \sup_{a_0}\|\frac{\partial}{\partial a_0}[b,(\cdot)_t]\diamond\partial_1^2 v_i\| \leqq [a]_\alpha N_0 N_i (t^{\frac14})^{2\alpha-2}.\end{aligned}$$

For (5.126) we write

$$\begin{aligned}&\big([b,(\cdot)_t]\partial_1^2 u_t-\sigma_i E[b,(\cdot)_t]\partial_1^2 v_{it}\big)(x)\\&\quad=\int \mathrm{d}y\,\psi_t(x-y)(b(x)-b(y))\big(\partial_1^2 u_t(y)-\sigma_i(x)\partial_1^2 v_{it}(y,a(x))\big)\end{aligned}\tag{5.129}$$

and

$$\partial_1^2 u_t(y)-\sigma_i(x)\partial_1^2 v_{it}(y,a(x))=\int \mathrm{d}z\,\partial_1^2\psi_t(y-z)\times$$
$$\big(u(z)-u(x)-\sigma_i(x)(v_i(z,a(x))-v_i(x,a(x)))-\nu(x)(z-x)_1\big).$$

Hence by the modelledness assumption of u, the triangle inequality $d(z,x)\leqq d(z,y)+d(y,x)$, and (2.4), we obtain

$$|\partial_1^2 u_t(y)-\sigma_i(x)\partial_1^2 v_{it}(y,a(x))|\lesssim M((t^{\frac14})^{2\alpha-2}+(t^{\frac14})^{-2}d^{2\alpha}(y,x)).$$

Plugging this into (5.129), we obtain, using (2.4) once more,

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

as desired.

The further two steps are as Steps 5 and 6 in Lemma 3.3.

5.7. Proof of Corollary 3.4

This is a corollary to Lemma 3.3 in the sense that we specify the families $\{v(\cdot,x)\}_x$ and $\{v(\cdot,x)\diamond f\}_x$ there to be given by $\{\sigma(x)v(\cdot,a(x))\}_x$ and $\{\sigma(x)\, v(\cdot,a(x))\diamond f\}_x$, respectively. Step 4 provides the necessary translations of the continuity and boundedness assumptions. In addition, for part i) of this corollary, we need to deal with (up to second) derivatives in the parameter a_0', which on the level of Lemma 3.3 is taken care of in Step 3. For part ii), next to the parameter derivatives, we need to deal with differences in f, which is tackled in Step 2. Finally, for part iii), again next to parameter derivatives, we are confronted with differences in v, which is taken care of in Step 1. We write $[\cdot]$ for $[\cdot]_\alpha$.

STEP 1. Differences in v in Lemma 3.3. Suppose we are given *two* families of functions $\{v_i(\cdot,x)\}_x$, $i=0,1$, and *two* families of distributions $\{v_i(\cdot,x)\diamond f\}_x$ both satisfying (3.10) and (3.12) and (3.15) and (3.16), and satisfying the analogue for the difference, which with the abbreviations $\delta v:=v_1-v_0$, $\delta v(\cdot,x)\diamond f:=v_1(\cdot,x)\diamond f-v_0(\cdot,x)\diamond f$ can be written as

$$[\delta v(\cdot,x)]\leqq \delta N,\tag{5.130}$$
$$[\delta v(\cdot,x)-\delta v(\cdot,x')]\leqq \delta N d^\alpha(x,x'),\tag{5.131}$$
$$\|[\delta v(\cdot,x),(\cdot)]\diamond f\|_{2\alpha-2}\leqq \delta N N_1,\tag{5.132}$$
$$\|[\delta v(\cdot,x),(\cdot)]\diamond f-[\delta v(\cdot,x'),(\cdot)]\diamond f\|_{2\alpha-2}\leqq \delta N N_1 d^\alpha(x,x')\tag{5.133}$$

for some constant δN. Suppose further we are given two functions u_i both satisfying (3.13) and their difference $\delta u := u_1 - u_0$ satisfying the analogue statement for some constant δM and function $\delta \nu$:

$$|(\delta u(y) - \delta u(x)) - (\delta v(y,x) - \delta v(x,x)) - \delta \nu(x)(y-x)_1| \leqq \delta M d^{2\alpha}(y,x). \tag{5.134}$$

We claim that (3.17) holds in form of

$$\|[u_1,(\cdot)]\diamond f - [u_0,(\cdot)]\diamond f\|_{2\alpha-2} \lesssim (\delta M + \delta N)N_1. \tag{5.135}$$

Indeed, we start by applying Lemma 3.3 with (u, v, M, N) replaced by $(\delta u, \delta v, \delta M, \delta N)$. There exists δu such that (3.17) takes the form

$$\|[\delta u,(\cdot)]\diamond f\|_{2\alpha-2} \lesssim (\delta M + \delta N)N_1. \tag{5.136}$$

Note that (3.13) holds for (u, v, ν) replaced by $(\delta u, \delta v, \delta \nu)$, (u_1, v_1, ν_1) and (u_0, v_0, ν_0). Because of the definition $(\delta u, \delta v) = (u_1 - u_0, v_1 - v_0)$ we thus obtain from the triangle inequality that $|\delta \nu - (\nu_1 - \nu_0)|(x)|(y-x)_1| \leqq (\delta M + 2M)d^{2\alpha}(y,x)$, which for $y \to x$ yields $\delta \nu = \nu_1 - \nu_0$. Note that (3.14) holds with (u, v, M, N) replaced by $(\delta u, \delta v, \delta M, \delta N)$, (u_1, v_1, M, N) and (u_0, v_0, M, N). Because of $(\delta u, \delta v, \delta v \diamond f, \delta \nu) = (u_1 - u_0, v_1 - v_0, v_1 \diamond f - v_0 \diamond f, \nu_1 - \nu_0)$ we obtain from the triangle inequality in $\|\cdot\|_{3\alpha-2}$ that $\lim_{T\downarrow 0} \|(\delta u \diamond f - (u \diamond f_1 - u \diamond f_0))_T\| = 0$ and thus $\delta u \diamond f = u_1 \diamond f - u_0 \diamond f$. Therefore (5.136) turns into (5.135).

Step 2. Differences in f in Lemma 3.3. Suppose we are given *two* distributions f_j, $j = 0, 1$, and *two* families of distributions $\{v(\cdot, x) \diamond f_j\}_x$ both satisfying (3.11) and (3.12) and (3.16), and satisfying the analogue for the difference, which introducing the abbreviations $\delta f := f_1 - f_0$ and $v(\cdot,x)\diamond \delta f := v(\cdot,x)\diamond f_1 - v(\cdot,x)\diamond f_0$, we may rewrite as

$$\|\delta f\|_{\alpha-2} \leqq \delta N_1, \tag{5.137}$$

$$\|[v(\cdot,x),(\cdot)]\diamond \delta f\|_{2\alpha-2} \leqq N\delta N_1, \tag{5.138}$$

$$\|[v(\cdot,x),(\cdot)]\diamond \delta f - [v(\cdot,x'),(\cdot)]\diamond \delta f\|_{2\alpha-2} \leqq N\delta N_1 d^{\alpha}(x,x') \tag{5.139}$$

for some constant δN_1. Then we claim the analogue of (3.17), namely

$$\|[u,(\cdot)]\diamond f_1 - [u,(\cdot)]\diamond f_0\|_{2\alpha-2} \lesssim (M+N)\delta N_1. \tag{5.140}$$

Indeed, from (5.137)–(5.139) together with the remaining assumptions of Lemma 3.3 we learn from the latter that there exists a distribution we call $u \diamond \delta f$ such that (3.14) holds with (f, N_1) replaced by $(\delta f, \delta N_1)$. Since it also holds with (f_j, N_1), we obtain from the triangle inequality and the above definition of $v(\cdot,x)\diamond \delta f$ that $\lim_{T\downarrow 0}\|(u\diamond \delta f - (u\diamond f_1 - u\diamond f_0))_T\| = 0$, which gives $u\diamond \delta f = u\diamond f_1 - u\diamond f_0$ and thus (3.17), still with (f, N_1) replaced by $(\delta f, \delta N_1)$, turns into (5.140).

Step 3. C^m-dependence of $(f, v\diamond f)$ on a parameter $a_0' \in [\lambda, \frac{1}{\lambda}]$ in Lemma 3.3. Suppose $f \in C^m_{a_0'}(C^{\alpha-2})$ and that $[v(\cdot,x),(\cdot)]\diamond f$ is of class $C^m_{a_0'}(C^{2\alpha-2})$ uniformly in x, see below for the precise meaning. We claim that this is preserved: $[u,(\cdot)]\diamond f$ is of class $C^m_{a_0'}(C^{2\alpha-2})$. Moreover, if (3.11), (3.12) and (3.16) are strengthened to

$$\|f\|_{\alpha-2,m} \leqq N_1, \tag{5.141}$$

$$\|[v(\cdot,x),(\cdot)]\diamond f\|_{2\alpha-2,m} \leqq NN_1, \tag{5.142}$$
$$\|[v(\cdot,x),(\cdot)]\diamond f-[v(\cdot,x'),(\cdot)]\diamond f\|_{2\alpha-2,m} \leqq NN_1 d^{\alpha}(x,x'), \tag{5.143}$$

cf. (3.18) and (3.19), then (3.17) improves likewise:

$$\|[u,(\cdot)]\diamond f\|_{2\alpha-2,m} \lesssim (M+N)N_1. \tag{5.144}$$

In virtue of Lemma 3.3 and fixing a_0' and j, we may associate to $\big((\frac{\partial}{\partial a_0'})^j f, \{(\frac{\partial}{\partial a_0'})^j (v(\cdot,x)\diamond f)\}_x\big)$ a distribution we call $u\diamond(\frac{\partial}{\partial a_0'})^j f$. Under the assumptions of Lemma 3.3 enhanced by (5.141)–(5.143), (3.17) turns into

$$\|[u,(\cdot)]\diamond\left(\frac{\partial}{\partial a_0'}\right)^j f\|_{2\alpha-2} \lesssim (M+N)N_1. \tag{5.145}$$

It is convenient to abbreviate by $\mathcal{R}f(\tilde a_0',a_0') := f(\tilde a_0') - \sum_{j=0}^m (\tilde a_0' - a_0')^j (\frac{\partial}{\partial a_0'})^j f(a_0')$ Taylor's remainder for a generic (Banach space-valued) function f of a_0'. Our C^m-assumption on the input includes $\lim_{\tilde a_0'\to a_0'} \|\mathcal{R}f(\tilde a_0',a_0')\|_{\alpha-2} = 0$ and $\lim_{\tilde a_0'\to a_0'} \sup_x \|\mathcal{R}([v(x,\cdot),(\cdot)]\diamond f)(\tilde a_0',a_0')\|_{2\alpha-2} = 0$. From the latter we learn that $u\diamond f \in C^m_{a_0'}(C^{\alpha-2})$ with $(\frac{\partial}{\partial a_0'})^j\big(u\diamond f\big) = u\diamond(\frac{\partial}{\partial a_0'})^j f$, so that in particular (5.145) turns into (5.144). From the former we therefore learn that $\lim_{\tilde a_0'\to a_0'} \|\mathcal{R}([u,(\cdot)]\diamond f)(\tilde a_0',a_0')\|_{2\alpha-2} = 0$, so that the $C^m_{a_0'}(C^{2\alpha-2})$ property is transmitted.

STEP 4. Some algebra. 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

$$|v|_1 \leqq N_0, \tag{5.146}$$
$$|v_i|_2 \leqq N_0, \tag{5.147}$$
$$|v_1 - v_0|_1 \leqq \delta N_0 \tag{5.148}$$

for some constants N_0, δN_0 (here as in (3.18) the subscripts in $|\cdot|_1$ and $|\cdot|_2$ refer to the number of parameter derivatives with respect to a_0). The reason for this more general framework is useful because in Step 5 we apply it with $v(\cdot,a_0)$ replaced by $[v(\cdot,a_0),(\cdot)_T]\diamond f$ and with the supremum norm $N_1^{-1}(T^{\frac14})^{2-2\alpha}\|\cdot\|$ playing the role of $|\cdot|$. We claim that this entails

$$|\sigma(x)v(\cdot,a(x))| \leqq N_0\|\sigma\|, \tag{5.149}$$
$$|\sigma(x)v(\cdot,a(x)) - \sigma(x')v(\cdot,a(x'))| \leqq N_0([\sigma]+\|\sigma\|[a])d^{\alpha}(x,x'), \tag{5.150}$$
$$\begin{aligned}&|\sigma_1(x)v_1(\cdot,a_1(x)) - \sigma_0(x)v_0(\cdot,a_0(x))|\\ &\quad\leqq N_0(\|\sigma_1-\sigma_0\| + \max_i \|\sigma_i\|\|a_1-a_0\|) + \delta N_0 \max_i \|\sigma_i\|,\end{aligned} \tag{5.151}$$
$$\begin{aligned}&|\big(\sigma_1(x)v_1(\cdot,a_1(x)) - \sigma_0(x)v_0(\cdot,a_0(x))\big)\\ &\qquad - \big(\sigma_1(x')v_1(\cdot,a_1(x')) - \sigma_0(x')v_0(\cdot,a_0(x'))\big)|\\ &\quad\leqq \Big(N_0 \max_{i,j}\big([\sigma_1-\sigma_0] + \|\sigma_i\|[a_1-a_0] + [\sigma_i]\|a_1-a_0\|\end{aligned}$$

$$+ \|\sigma_1 - \sigma_0\|[a_i] + \|\sigma_i\|[a_j]\|a_1 - a_0\|\big)$$
$$+ \delta N_0 \max_i \big([\sigma_i] + \|\sigma_i\|[a_i]\big)\Big) d^\alpha(x, x'). \tag{5.152}$$

Estimate (5.149) follows immediately from (5.146). We treat (5.150), (5.151) and (5.152) along the same lines, which is a bit of an overkill for (5.150) and (5.151). We start with the two elementary, and purposefully symmetric, formulas:

$$\sigma v - \sigma' v' = \frac{1}{2}(\sigma - \sigma')(v + v') + \frac{1}{2}(\sigma + \sigma')(v - v'), \tag{5.153}$$

$$\begin{aligned}
&(\sigma_1 v_1 - \sigma_0 v_0) - (\sigma_1' v_1' - \sigma_0' v_0') \\
&\quad = \frac{1}{4}((\sigma_1 - \sigma_0) - (\sigma_1' - \sigma_0'))(v_1 + v_1' + v_0 + v_0') \\
&\qquad + \frac{1}{4}((\sigma_1 + \sigma_1' + \sigma_0 + \sigma_0'))((v_1 - v_0) - (v_1' - v_0')) \\
&\qquad + \frac{1}{4}((\sigma_1 - \sigma_1') + (\sigma_0 - \sigma_0'))((v_1 - v_0) + (v_1' - v_0')) \\
&\qquad + \frac{1}{4}((\sigma_1 - \sigma_0) + (\sigma_1' - \sigma_0'))((v_1 - v_1') + (v_0 - v_0')).
\end{aligned} \tag{5.154}$$

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

$$\begin{aligned}
&|\sigma(x) v(\cdot, a(x)) - \sigma(x') v'(\cdot, a(x'))| \\
&\quad \leqq [\sigma] d^\alpha(x, x') \sup_{a_0} |v(\cdot, a_0)| + \|\sigma\| \sup_{a_0} |\frac{\partial v}{\partial a_0}(\cdot, a_0)| [a] d^\alpha(x, x').
\end{aligned}$$

In view of the assumption (5.146) this yields (5.150). 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

$$\begin{aligned}
&|\sigma_1(x) v_1(\cdot, a_1(x)) - \sigma_0(x) v_0(\cdot, a_0(x))| \\
&\quad \leqq \|\sigma_1 - \sigma_0\| \max_i \sup_{a_0} |v_i(\cdot, a_0)| + \max_i \|\sigma_i\| |v_1(\cdot, a_1(x)) - v_0(\cdot, a_0(x))|.
\end{aligned} \tag{5.155}$$

In view of the assumption (5.147), the first right-hand-side term is estimated as desired. For the second rhs term we interpolate linearly in the sense of $v_s := s v_1 + (1 - s) v_0$ and $a_s := s a_1 + (1 - s) a_0$, to the effect of

$$\begin{aligned}
&v_1(\cdot, a_1(x)) - v_0(\cdot, a_0(x)) \\
&\quad = \int_0^1 \mathrm{d}s \big((v_1 - v_0)(\cdot, a_s(x)) + \frac{\partial v_s}{\partial a_0}(\cdot, a_s(x))(a_1 - a_0)(x),
\end{aligned} \tag{5.156}$$

from which we learn that

$$|v_1(\cdot, a_1(x)) - v_0(\cdot, a_0(x))| \leqq \sup_{a_0} |v_1 - v_0| + \max_i \sup_{a_0} |\frac{\partial v_i}{\partial a_0}| \|a_1 - a_0\|. \tag{5.157}$$

Inserting this into (5.155) and in view of the assumption (5.147) and (5.148) we obtain the remaining part of (5.151).

We use the second formula (5.154) 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

$$\begin{aligned}
&|\big(\sigma_1(x)v_1(\cdot, a_1(x)) - \sigma_0(x)v_0(\cdot, a_0(x))\big) - \big(\sigma_1(x')v_1(\cdot, a_1(x')) \\
&\quad - \sigma_0(x')v_0(\cdot, a_0(x'))\big)| \\
&\quad \leqq [\sigma_1 - \sigma_0]d^\alpha(x, x') \max_i \sup_{a_0} |v_i(\cdot, a_0)| \\
&\quad + \max_i \|\sigma_i\| |(v_1(\cdot, a_1(x)) - v_0(\cdot, a_0(x))) - (v_1(\cdot, a_1(x')) - v_0(\cdot, a_0(x')))| \\
&\quad + \max_i [\sigma_i] d^\alpha(x, x') \sup_y |v_1(\cdot, a_1(y)) - v_0(\cdot, a_0(y))| \\
&\quad + \|\sigma_1 - \sigma_0\| \max_i \sup_{a_0} \left|\frac{\partial v_i}{\partial a_0}(\cdot, a_0)\right| [a_i] d^\alpha(x, x').
\end{aligned}$$

In order to deduce (5.152) from this inequality, in view of (5.157) and of our assumption from (5.147) and (5.148), it remains to show for the second right-hand-side terms that

$$\begin{aligned}
&|(v_1(\cdot, a_1(x)) - v_0(\cdot, a_0(x))) - (v_1(\cdot, a_1(x')) - v_0(\cdot, a_0(x')))| \\
&\quad \leqq \sup_{a_0} \left|\frac{\partial}{\partial a_0}(v_1 - v_0)(\cdot, a_0)\right| \max_i [a_i] d^\alpha(x, x') \\
&\quad + \max_i \sup_{a_0} \left|\frac{\partial^2 v_i}{\partial a_0^2}(\cdot, a_0)\right| \max_j [a_j] d^\alpha(x, x') \|a_1 - a_0\| \\
&\quad + \max_i \sup_{a_0} \left|\frac{\partial v_i}{\partial a_0}(\cdot, a_0)\right| [a_1 - a_0] d^\alpha(x, x'). \qquad (5.158)
\end{aligned}$$

We appeal again to the outcome (5.156) of the linear interpolation, which immediately yields the first right-hand-side term (5.158) from the first right-hand-side term in (5.156). For the second right-hand-side term in (5.158), we appeal once more to formula (5.153) (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 5. Conclusion. We start with part i) of this corollary; we apply Lemma 3.3, in form of Step 3 with $m = 2$, to the families given by distributions $\{f(\cdot, a_0')\}_{a_0'}$, the functions $\{\sigma(x)v(\cdot, a(x))\}_x$, and the products $\{\sigma(x)\, v(\cdot, a(x)) \diamond f(\cdot, a_0')\}_{x, a_0'}$. To this end we verify the hypotheses; hypothesis (5.141) on the distribution f is identical to the corollary's hypothesis (3.21). We now turn to those on the function v, namely (3.15) and (3.10). Using (3.24), these follow, with N_0 playing the role of N, from (5.149) and (5.150) of Step 4 provided the generic semi-norm $|\cdot|$ there is chosen to be $[\cdot]$. The relevant hypothesis (5.146) of Step 4 is identical to the corollary's hypothesis (3.20). We last turn to the hypothesis on the product $v \diamond f$, that is, (5.142) and (5.143); to this end, we fix a convolution parameter $T \in (0, 1]$, an order of differentiation $j = 0, 1, 2$ and the parameter a_0'. These hypotheses follow again from (5.149) and (5.150) of Step 4, this time with $\{(\frac{\partial}{\partial a_0'})^j [v(\cdot, a_0), (\cdot)_T] \diamond f(\cdot, a_0')\}_{a_0}$

playing the role of $\{v(\cdot, a_0)\}_{a_0}$ and the norm $N_1^{-1}(T^{\frac{1}{4}})^{2-2\alpha}\|\cdot\|$ replacing $|\cdot|$. The relevant hypothesis (5.146) then holds by the corollary's hypothesis (3.22); the outputs (5.149) and (5.150) indeed turn into (5.142) and (5.143), still with N_0 playing the role of N. Finally, the outcome (5.144) of Step 3 turns into the desired (3.25).

We now turn to part ii) of this corollary. Again, we apply Lemma 3.3, this time in form of Step 2, upgraded by Step 3 with $m = 1$ in the sense that the expressions $(\|\cdot\|_{\alpha-2}, \|\cdot\|_{2\alpha-2})$ are replaced by $(\|\cdot\|_{\alpha-2,1}, \|\cdot\|_{2\alpha-2,1})$. The argument follows the lines of the one for part i); when it comes to the product $(\frac{\partial}{\partial a_0'})^m(v \diamond f_1 - v \diamond f_0)$, for fixed $m = 0, 1$ and parameter a_0', the presence of an a_0-derivative in the corollary's hypothesis (3.27) feeds into Step 4's hypothesis (5.146) with the semi-norm $|\cdot|$ $= (\delta N_1)^{-1}(T^{\frac{1}{4}})^{2-2\alpha}\|\cdot\|$. Step 4's output, (5.149) and (5.150), provides Step 2's input: (5.138) and (5.139). Step 2's output (5.140) is identical to the corollary's claim (3.28).

We finally turn to part iii) of this corollary. For a final last time, we apply Lemma 3.3, now in the form of Step 1, upgraded in terms of differentiability in the parameter a_0' by Step 3 with $m = 1$. We apply Step 1 to the families given by distributions $\{f(\cdot, a_0')\}_{a_0'}$, the functions $\{\sigma_i(x)v_i(\cdot, a_i(x))\}_x$, and the products $\{\sigma_i(x)\, v(\cdot, a_i(x)) \diamond f(\cdot, a_0')\}_{x,a_0'}$. We start with the hypotheses (5.130) and (5.131) on the difference of the functions and apply Step 4 to $|\cdot| = [\cdot]$. The relevant input (5.147) and (5.148) of that step is provided by the corollary's assumptions (3.29) and (3.30). In view of (3.24), the output (5.151) and (5.152) of Step 4 turns into the hypotheses (5.130) and (5.131) with $\delta N := N_0([\sigma_1 - \sigma_0] + \|\sigma_1 - \sigma_0\| + [a_1 - a_0] + \|a_1 - a_0\|) + \delta N_0$. We now turn to the hypotheses (5.132) and (5.133) on the difference of the products and apply Step 4 to $\{(\frac{\partial}{\partial a_0'})^j v_i(\cdot, a_0) \diamond f(\cdot, a_0')\}_{a_0}, i = 0, 1,$ playing the role of $\{v_i(\cdot, a_0)\}_{a_0}$, and with $|\cdot| = N_1^{-1}(T^{\frac{1}{4}})^{2-2\alpha}\|\cdot\|$ for fixed T, $j = 0, 1$ and a_0'. The relevant input (5.147) and (5.148) of that step is provided by the corollary's assumptions (3.31) and (3.32). The output (5.151) and (5.152) of Step 4 turns into the hypotheses (5.132) and (5.133) with the above definition of δN. Finally, we note that the modelledness assumption of our corollary assumes the form (5.134). The output (5.135) of Step 1 turns into the desired (3.33).

5.8. *Proof of Corollary 3.10*

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 at 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,

$$a(u)\partial_1^2 u := \partial_1(a(u)\partial_1 u) - \partial_1 a(u)\partial_1 u. \qquad (5.159)$$

The claim then follows from standard parabolic regularity theory as soon as we have established that

$$\sigma(u)\diamond f = \sigma(u)f + \sigma'(u)\sigma(u)g_1(\cdot, a(u)) \tag{5.160}$$

$$a(u)\diamond \partial_1^2 u = a(u)\partial_1^2 u + a'(u)\sigma^2(u)g_2(\cdot, a(u), a(u)). \tag{5.161}$$

We first argue that (5.160) holds. To see this, first by Lemma 3.2 $\sigma(u)$ is modelled after v according to $a(u)$ and $\sigma'(u)\sigma(u)$. Then, Corollary 3.4 characterizes $\sigma(u)\diamond f$ as the unique distribution for which

$$\lim_{T\downarrow 0} \|[\sigma(u), (\cdot)_T]\diamond f - \sigma'(u)\sigma(u)E[v, (\cdot)_T]\diamond f - \nu[x_1, (\cdot)]f\| = 0. \tag{5.162}$$

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 (5.162) turns into

$$\lim_{T\downarrow 0} \|\sigma(u)f - (\sigma(u)\diamond f)_T - \sigma'(u)\sigma(u)g_{1,T}(\cdot, a(u))\| = 0.$$

Since, $g(\cdot, a_0) \in C^\beta$ by assumption, this yields (5.160). In the same way, one can see that for any a_0' we have that

$$a(u)\diamond \partial_1^2 v(\cdot, a_0') = a(u)\partial_1^2 v(\cdot, a_0') + a'(u)\sigma(u)g_2(\cdot, a(u), a_0') \tag{5.163}$$

(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 (5.163) to (5.161), that is the second factor $\partial_1^2 v$ in (5.163) should be replaced by $\partial_1^2 u$. To this end we make the ansatz

$$a(u)\diamond \partial_1^2 u = a(u)\partial_1^2 u + a'(u)\sigma^2(u)g_2(\cdot, a(u), a(u)) + B, \tag{5.164}$$

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 3.5 and plug in our ansatz (5.164) to obtain

$$\begin{aligned}\lim_{T\downarrow 0} \|&[a(u), (\cdot)_T]\partial_1^2 u - (a'(u)\sigma^2(u)g_2(\cdot, a(u), a(u)))_T + (B)_T \\ &- \sigma(u)E[a(u), (\cdot)_T]\diamond \partial_1^2 v\| = 0.\end{aligned} \tag{5.165}$$

Plugging (5.163) into (5.165) we obtain

$$\begin{aligned}\lim_{T\downarrow 0} \|&[a(u), (\cdot)_T]\partial_1^2 u - (a'(u)\sigma^2(u)g_2(\cdot, a(u), a(u)))_T + (B)_T \\ &- \sigma(u)E[a(u), (\cdot)_T]\partial_1^2 v - a'(u)\sigma^2(u)E(g_2(\cdot, a(u), a_0'))_T\| = 0.\end{aligned} \tag{5.166}$$

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 (5.166) 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 (5.159). Now, according to its definition (5.159) 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 (5.160) and (5.161) 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.

6. Proofs of the Stochastic Bounds

6.1. Proof of Lemma 4.1

Step 1. Proof of (4.3). By the stationarity of $f_T = f * \psi_T$ we have for $T \leqq 1$ that

$$\langle f_T^2(0)\rangle = \left\langle \int_{[0,1)^2} f_T^2 \,\mathrm{d}x \right\rangle \overset{(4.1)}{=} \sum_{k\in(2\pi\mathbb{Z})^2} \hat{\psi}_T^2(k)\hat{C}(k)$$
$$\overset{(4.3)}{\leqq} (T^{\frac14})^{-3+\lambda_1+\lambda_2} \sum_{k\in(2\pi\mathbb{Z})^2\setminus\{0\}} (T^{\frac14})^3 \frac{e^{-2(T^{\frac14}k_1)^4-2(T^{\frac12}k_2)^2}}{(T^{\frac14}(1+|k_1|))^{\lambda_1}(T^{\frac12}(1+|k_2|))^{\frac{\lambda_2}{2}}}$$
$$\lesssim (T^{\frac14})^{2\alpha'-4}.$$

In the last estimate we have used that for $T \downarrow 0$ the sum in the third line is a Riemann sum approximation of the integral $\int e^{-2\hat{k}_1^4-2\hat{k}_2^2}|\hat{k}_1|^{-\lambda_1}|\hat{k}_2|^{-\frac{\lambda_2}{2}}\,\mathrm{d}\hat{k}$ which converges due to $\lambda_1, \frac{\lambda_2}{2} < 1$.

The fact that f_T is Gaussian and stationary implies that we have $\langle |f_T(x)|^p\rangle \lesssim \langle f_T^2(0)\rangle^{\frac{p}{2}}$, which permits us to write

$$\left\langle \int_{[0,1)^2} |f_T|^p \,\mathrm{d}x \right\rangle^{\frac1p} \lesssim \langle f_T^2(0)\rangle^{\frac12} \lesssim (T^{\frac14})^{\alpha'-2}.$$

In order to upgrade this L^p bound to an L^∞ bound under the expectation we observe that by the semi-group property (2.3) we have $f_T = (f_{T/2})_{T/2}$ such that Hölder'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 $\psi_{T,\mathrm{per}}(x) = \sum_{z\in\mathbb{Z}^2} \psi_T(x+z)$ is the periodization of ψ_T. By observing that for small T the difference $\big|\|\psi_{T,\mathrm{per}}\|_{L^{p'}} - \big(\int_{\mathbb{R}^2} |\psi_T|^{p'}\,\mathrm{d}x\big)^{\frac{1}{p'}}\big|$ stays bounded, and scaling we get $\|\psi_{T,\mathrm{per}}\|_{L^{p'}} \lesssim (T^{\frac14})^{-\frac{3}{p}}$ such that, finally,

$$\big\langle \|f_T\|^p\big\rangle^{\frac1p} \lesssim (T^{\frac14})^{-\frac3p}\big\langle \|f_T\|_{L^p}^p\big\rangle^{\frac1p} \lesssim (T^{\frac14})^{\alpha'-2-\frac3p}.$$

To also accommodate for the supremum over the scales T we first note that $\|f_{T+t}\| \lesssim \|f_T\|$ implies

$$\|f\|_{\alpha-2} = \sup_{T\leqq 1}(T^{\frac14})^{2-\alpha}\|f_T\| \lesssim \sup_{T\leqq 1,\mathrm{dyadic}}(T^{\frac14})^{2-\alpha}\|f_T\|,$$

where the subscript dyadic means that this supremum is only taken over dyadic T. Then we write

$$\begin{aligned}\Big\langle\Big(\sup_{T\leqq 1,\mathrm{dyadic}}(T^{\frac14})^{2-\alpha}\|f_T\|\Big)^p\Big\rangle &\leqq \sum_{T\leqq 1,\mathrm{dyadic}}(T^{\frac14})^{p(2-\alpha)}\big\langle\|f_T\|^p\big\rangle\\ &\lesssim \sum_{T\leqq 1,\mathrm{dyadic}}(T^{\frac14})^{p(2-\alpha)}(T^{\frac14})^{p(\alpha'-2)-3},\end{aligned}$$

which converges as soon as $p > \frac{3}{\alpha'-\alpha}$ and thus establishes $\big\langle\|f\|_{\alpha-2}^p\big\rangle^{\frac1p} \lesssim 1$ for large p. The same bound for smaller p can be derived from the bound for large p and Jensen's inequality. Finally, because of $(f_\varepsilon)_T = \varphi_\varepsilon * f_T$ and because the operators $\varphi_\varepsilon *$ are bounded with respect to $\|\cdot\|$ uniformly in ε, the bound holds uniformly in the regularization leading to the desired estimate (4.3).

STEP 2. Proof of (4.4). The bound on the ε-differences follows from (4.3) as soon as we have established the deterministic bound

$$\|(f_\varepsilon)_T - f_T\| \lesssim \min\left\{\left(\frac{\varepsilon}{T}\right)^{\frac14}, 1\right\}\|f_{T/2}\|, \tag{6.1}$$

which by the semi-group property reduces to

$$\|(f_\varepsilon)_T - f_T\| \lesssim \min\left\{\left(\frac{\varepsilon}{T}\right)^{\frac14}, 1\right\}\|f\|.$$

Since $(\cdot)_T$ and $\varphi_\varepsilon *$ are bounded with respect to $\|\cdot\|$, it suffices to consider $\varepsilon \leqq T$. We then write

$$\|(f_\varepsilon)_T - f_T\| = \|(\psi_T * \varphi_\varepsilon - \psi_T) * f\| \leqq \int_{\mathbb{R}^2}|\psi_T * \varphi_\varepsilon - \psi_T|\,\mathrm{d}x\;\|f\|,$$

and have thereby reduced (6.1) (and hence (4.4)) to establishing that

$$\int_{\mathbb{R}^2}|\psi_T * \varphi_\varepsilon - \psi_T|\,\mathrm{d}x \lesssim \left(\frac{\varepsilon}{T}\right)^{\frac14} \qquad \text{for } \varepsilon \leqq T.$$

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 $T = 1$, in which case it turns into

$$\int_{\mathbb{R}^2} |\psi_1 * \varphi_\varepsilon - \psi_1| \lesssim \varepsilon^{\frac{1}{4}} \qquad \text{for } \varepsilon \leqq 1,$$

which is immediate for Schwartz kernels ψ_1, φ and in view of the definition (1.7) of φ_ε.

6.2. *Proof of Lemma 4.3*

For $a_0 \in [\lambda, \frac{1}{\lambda}]$ 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 Fourier series is given by

$$\hat{G}(k, a_0) = \begin{cases} \frac{1}{a_0 k_1^2 - ik_2} = \frac{a_0 k_1^2 + ik_2}{a_0^2 k_1^4 + k_2^2} & \text{for } k \in (2\pi\mathbb{Z})^2 \backslash \{0\}, \\ 0 & \text{for } k = 0. \end{cases} \tag{6.2}$$

With this notation in place, $v(\cdot, a_0)$ is characterized by its discrete Fourier transforms $\hat{v}(k, a_0) = \hat{G}(k, a_0)\hat{f}(k)$. Throughout the proof the parameter dependence on a_0 only appears in $\hat{G}(k, a_0)$ for which only the bound

$$|\hat{G}(k, a_0)| \lesssim \frac{1}{k_1^2 + |k_2|} \tag{6.3}$$

is used. We thus suppress the a_0-dependence in all expressions.

Step 1. Bound on the expectation. We claim that

$$\langle [v', (\cdot)_T] \diamond f' \rangle \lesssim (T^{\frac{1}{4}})^{2\alpha' - 2 - \kappa_1 - \kappa_2}. \tag{6.4}$$

By stationarity $\langle (v' \diamond f')_T \rangle = \langle v' \diamond f' \rangle = 0$. Furthermore, by stationarity and (4.1) and (4.2) we have

$$\begin{aligned} \left|\langle v' f'_T \rangle\right| &= \left|\sum_k \langle \hat{v}'(-k)\hat{\psi}_T(k)\hat{f}'(k)\rangle\right| = \left|\sum_k (\hat{M}_2\hat{G})(-k)\hat{\psi}_T(k)\hat{M}_1(k)\hat{C}(k)\right| \\ &\overset{(4.8),(6.3)}{\leqq} (T^{\frac{1}{4}})^{-3+2+\lambda_1+\lambda_2-\kappa_1-\kappa_2} \\ &\quad \times \sum_k (T^{\frac{1}{4}})^3 \frac{\hat{\psi}_T(k)}{(T^{\frac{1}{4}}k_1)^2 + |T^{\frac{1}{2}}k_2|} \frac{((T^{\frac{1}{4}}k_1)^4 + (T^{\frac{1}{2}}k_2)^2)^{\frac{\kappa_1+\kappa_2}{4}}}{(T^{\frac{1}{4}}(1+|k_1|))^{\lambda_1}(T^{\frac{1}{2}}(1+|k_2|))^{\frac{\lambda_2}{2}}} \\ &\lesssim (T^{\frac{1}{4}})^{2\alpha' - 2 - \kappa_1 - \kappa_2}, \end{aligned}$$

where the sum is taken over $(2\pi\mathbb{Z})^2 \backslash \{0\}$. 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 on 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 $2+\lambda_1+\lambda_2 = 1+2\alpha' < 3$, where we appeal to the fact that the parabolic dimension is 3 (alternatively, one may split the integral into $|x_1| \leqq \sqrt{|x_2|}$ and its complement). This establishes (6.4).

STEP 2. Preparation for 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 the definition of $\diamond$ can be expressed equivalently without the renormalization as

$$\left|\langle([v',(\cdot)_T]f')^2\rangle - \langle [v',(\cdot)_T]f'\rangle^2\right|^{\frac{1}{2}} \lesssim (T^{\frac{1}{4}})^{2\alpha'-2-\kappa_1-\kappa_2}. \tag{6.5}$$

To derive the estimate in the form (6.5) we write using once more stationarity

$$\langle([v',(\cdot)_T]f')^2\rangle = \left\langle \int_{[0,1)^2} ([v',(\cdot)_T]f')^2\,\mathrm{d}x \right\rangle = \sum_{k\in(2\pi\mathbb{Z})^2} \left\langle |\widehat{[v',(\cdot)_T]f'}(k)|^2 \right\rangle.$$

The expression appearing in the last expectation can be evaluated according to its definition

$$\begin{aligned}\widehat{[v',(\cdot)_T]f'}(k) &= \sum_{\ell\in(2\pi\mathbb{Z})^2} (\hat\psi_T(\ell)-\hat\psi_T(k))\hat{v}'(k-\ell)\hat{f}'(\ell) \\ &= \sum_{\ell\in(2\pi\mathbb{Z})^2} (\hat\psi_T(\ell)-\hat\psi_T(k))(\hat{M}_2\hat{G})(k-\ell)\hat{f}(k-\ell)\hat{M}_1(\ell)\hat{f}(\ell),\end{aligned} \tag{6.6}$$

which permits as to write

$$\begin{aligned}\langle([v',(\cdot)_T]f')^2\rangle = \sum_k\sum_\ell\sum_{\ell'} &(\hat\psi_T(\ell)-\hat\psi_T(k))(\hat\psi_T(-\ell')-\hat\psi_T(-k)) \\ &\times (\hat{M}_2\hat{G})(k-\ell)(\hat{M}_2\hat{G})(-(k-\ell'))\hat{M}_1(\ell)\hat{M}_1(-\ell') \\ &\times \left\langle \hat{f}(k-\ell)\hat{f}(\ell)\hat{f}(-(k-\ell'))\hat{f}(-\ell')\right\rangle,\end{aligned} \tag{6.7}$$

where all sums are taken over $(2\pi\mathbb{Z})^2$. We now use (4.1) and the Gaussian identity

$$\begin{aligned}&\langle \hat{f}(k-\ell)\hat{f}(\ell)\hat{f}(-(k-\ell'))\hat{f}(-\ell')\rangle \\ &\quad = \delta_{k,0}\hat{C}(\ell)\hat{C}(\ell') + \delta_{\ell,\ell'}\hat{C}(k-\ell)\hat{C}(\ell) + \delta_{k-\ell,\ell'}\hat{C}(k-\ell)\hat{C}(\ell).\end{aligned} \tag{6.8}$$

Plugging this identity into (6.7) results in three terms which we bound one by one. The first term coincides with the square of the expectation (which is subtracted on the left hand side of (6.5))

$$\sum_{\ell}\sum_{\ell'}(\hat{\psi}_T(\ell)-\hat{\psi}_T(0))(\hat{\psi}_T(-\ell')-\hat{\psi}_T(0))$$
$$(\hat{M}_2\hat{G})(-\ell)(\hat{M}_2\hat{G})(\ell')(\hat{M}_1\hat{C})(\ell)(\hat{M}_1\hat{C})(-\ell')$$
$$=\Big(\sum_{\ell}(\hat{\psi}_T(\ell)-\hat{\psi}_T(0))(\hat{M}_2\hat{G})(-\ell)(\hat{M}_1\hat{C})(\ell)\Big)^2=\langle[v',(\cdot)_T]f'\rangle^2,$$

so that the required bound (6.5) follows as soon as we can bound the remaining two terms. The term originating from the third contribution on the right hand side of (6.8) can be absorbed into the second term $\sum_{k,\ell}(\hat{\psi}_T(\ell)-\hat{\psi}_T(k))^2|(\hat{M}_2\hat{G})(k-\ell)|^2|\hat{M}_1(\ell)|^2\hat{C}(\ell)\hat{C}(k-\ell)$ using the Cauchy–Schwarz inequality. Indeed, we may write

$$\sum_{k}\sum_{\ell}(\hat{\psi}_T(\ell)-\hat{\psi}_T(k))(\hat{\psi}_T(-(k-\ell))-\hat{\psi}_T(-k))$$
$$\times(\hat{M}_2\hat{G})(k-\ell)(\hat{M}_2\hat{G})(-\ell)(\hat{M}_1\hat{C})(\ell)(\hat{M}_1\hat{C})(-(k-\ell))$$
$$\leqq\Big(\sum_{k,\ell}(\hat{\psi}_T(\ell)-\hat{\psi}_T(k))^2|(\hat{M}_2\hat{G})(k-\ell)|^2|\hat{M}_1(\ell)|^2\hat{C}(-(k-\ell))\hat{C}(\ell)\Big)^{\frac{1}{2}}$$
$$\times\Big(\sum_{k,\ell}(\hat{\psi}_T(-(k-\ell))-\hat{\psi}_T(-k))^2|(\hat{M}_2\hat{G})(-\ell)|^2$$
$$\times|\hat{M}_1(-(k-\ell))|^2\hat{C}(-(k-\ell))\hat{C}(\ell)\Big)^{\frac{1}{2}},$$

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 (6.8). We use the assumptions (4.2) and (4.8) to bound this term as follows:

$$\sum_{k,\ell}(\hat{\psi}_T(\ell)-\hat{\psi}_T(k))^2|(\hat{M}_2\hat{G})(k-\ell)|^2|\hat{M}_1(\ell)|^2\hat{C}(\ell)\hat{C}(k-\ell)$$
$$\overset{(4.8)(6.3)}{\lesssim}\sum_{k\neq\ell}\frac{(\hat{\psi}_T(\ell)-\hat{\psi}_T(k))^2}{(|(k-\ell)_1^2+|(k-\ell)_2|)^2}\frac{(\ell_1^4+\ell_2^2)^{\frac{\kappa_1}{2}}}{(1+|\ell_1|)^{\lambda_1}(1+|\ell_2|)^{\frac{\lambda_2}{2}}}$$
$$\times\frac{((k-\ell)_1^4+(k-\ell)_2^2)^{\frac{\kappa_2}{2}}}{(1+|(k-\ell)_1)|)^{\lambda_1}(1+|(k-\ell)_2|)^{\frac{\lambda_2}{2}}}$$
$$=(T^{\frac{1}{4}})^{4\alpha'-4-2\kappa_1-2\kappa_2}\sum_{k\neq\ell}\big(T^{\frac{1}{4}}\big)^6\left(\frac{\hat{\psi}_T(\ell)-\hat{\psi}_T(k)}{(T^{\frac{1}{4}}(k-\ell)_1)^2+|T^{\frac{1}{2}}(k-\ell)_2|}\right)^2$$
$$\times\frac{((T^{\frac{1}{4}}\ell_1)^4+(T^{\frac{1}{2}}\ell_2)^2)^{\frac{\kappa_1}{2}}}{(T^{\frac{1}{4}}(1+|\ell_1|))^{\lambda_1}(T^{\frac{1}{2}}(1+|\ell_2|))^{\frac{\lambda_2}{2}}}$$
$$\times\frac{((T^{\frac{1}{4}}(k-\ell)_1)^4+(T^{\frac{1}{2}}(k-\ell)_2)^2)^{\frac{\kappa_2}{2}}}{(T^{\frac{1}{4}}(1+|(k-\ell)_1|))^{\lambda_1}(T^{\frac{1}{2}}(1+|(k-\ell)_2|))^{\frac{\lambda_2}{2}}}.\tag{6.9}$$

STEP 3. Bound on an integral. In order to show that the expression (6.9) is bounded by $\lesssim (T^{\frac{1}{4}})^{4\alpha'-4-2\kappa_1-2\kappa_2}$, which in turn establishes (6.5), it remains to show the convergence of the integral which is approximated by the Riemann sum in the last lines:

$$\iint \Big(\frac{\hat\psi_1(\hat\ell)-\hat\psi_1(\hat k)}{(\hat k-\hat\ell)_1^2+|\hat k_2-\hat\ell_2|}\Big)^2 \bar C_1(\hat\ell)\bar C_2(\hat k-\hat\ell)d\hat\ell\,\mathrm{d}\hat k, \tag{6.10}$$

where momentarily we use the short-hand

$$\bar C_i(\hat\ell) := \frac{(\hat\ell_1^4+\hat\ell_2^2)^{\frac{\kappa_i}{2}}}{|\hat\ell_1|^{\lambda_1}|\hat\ell_2|^{\frac{\lambda_2}{2}}} \qquad i=1,2.$$

As a first step we deal with the integral near the diagonal, where $|(\hat\ell-\hat k)_1|+|(\hat\ell-\hat k)_2|\leqq 1$. For these values the change of variables $\hat h=\hat k-\hat\ell$ is useful. We furthermore make use of the bound $|\hat\psi_1(\hat\ell)-\hat\psi_1(\hat\ell+\hat h)|\lesssim(|\hat h_1|+|\hat h_2|)\int_0^1|\nabla\hat\psi_1(\hat\ell+\theta\hat h)|\,\mathrm{d}\theta$ and brutally bound $(|\hat h_1|+|\hat h_2|)\leqq\sqrt{\hat h_1^2+|\hat h_2|}$ so that we need to address the convergence of

$$\iint_{|\hat h_1|+|\hat h_2|\leqq 1}\frac{\max_{\theta\in[0,1]}|\nabla\hat\psi_1|^2(\ell+\theta\hat h)}{\hat h_1^2+|\hat h_2|}\bar C_1(\hat\ell)\bar C_2(\hat h)\,\mathrm{d}\hat h\,\mathrm{d}\hat\ell.$$

For the $\mathrm{d}\hat h$ integral over a finite volume it suffices to assert that the singularities near the axes $\hat h_1=0$ and $\hat h_2=0$ are integrable due to $\lambda_1,\frac{\lambda_2}{2}<1$ and that the singularity near the origin is integrable because by assumption (4.2) $2+\lambda_1+\lambda_2=1+2\alpha'$ which is less than the parabolic dimension 3. The singularities for the $d\hat\ell$ integral are only better behaved and the convergence of the integral for $|\hat\ell|\to\infty$ is guaranteed by the exponential decay of $\nabla\hat\psi_1$.

We now discuss the convergence of (6.10) for $|(\hat\ell-\hat k)_1|+|(\hat\ell-\hat k)_2|>1$. For these values we write $(\hat\psi_1(\hat\ell)-\hat\psi_1(\hat k))^2\lesssim\hat\psi_1^2(\hat\ell)+\hat\psi_1^2(\hat k)$ and treat the resulting integrals separately. For the integral coming from $\hat\psi_1^2(\hat\ell)$ we use the same change of variables $\hat h=\hat k-\hat\ell$ which leads us to consider the integral

$$\iint_{|\hat h_1|+|\hat h_2|>1}\frac{\hat\psi_1^2(\hat\ell)}{(\hat h_1^2+|\hat h_2|)^2}\bar C_1(\hat\ell)\bar C_2(\hat h)\,\mathrm{d}\hat\ell\,\mathrm{d}\hat h.$$

As above, the $d\hat\ell$ integral converges because the singularities of $\bar C_1(\ell)$ near the axes $\ell_1=0$ and $\ell_2=0$ as well as the singularity near the origin are integrable and because of the exponential decay of $\hat\psi_1$ at infinity. The singularities of the $d\hat h$ integral near the axes are also integrable and its convergence for $|\hat h|\to\infty$ is guaranteed by the fact that $4+\lambda_1+\lambda_2=3+2\alpha'$ which is larger than the parabolic dimension 3 and by $\kappa_2\ll 1$.

It remains to treat the integral coming from $\hat\psi_1^2(\hat k)$:

$$\iint_{|(\ell-k)_1|+|(\ell-k)_2|>1}\frac{\hat\psi_1^2(\hat k)}{((\hat k-\hat\ell)_1^2+|\hat k_2-\hat\ell_2|)^2}\bar C_1(\hat\ell)\bar C_2(\hat k-\hat\ell)\,\mathrm{d}\hat\ell\,\mathrm{d}\hat k.$$

It is here that our assumption $\alpha > \frac{1}{4}$ becomes relevant to assure the convergence of the $d\hat{\ell}$ integral. We get

$$\int_{|(\hat{\ell}-\hat{k})_1|+|(\hat{\ell}-\hat{k})_2|>1} \frac{1}{((\hat{k}-\hat{\ell})_1^2+|\hat{k}_2-\hat{\ell}_2|)^2} \bar{C}_1(\hat{\ell})\bar{C}_2(\hat{k}-\hat{\ell})\,\mathrm{d}\hat{\ell},$$

which converges for $|\hat{\ell}| \to \infty$ because of $4+2(\lambda_1+\lambda_2) = 2+4\alpha'$ which is larger than the parabolic dimension 3 due to $\alpha' > \frac{1}{4}$ and because $\kappa_1, \kappa_2 \ll 1$. The convergence of the resulting $d\hat{k}$ integral near the origin is guaranteed by $2(\lambda_1+\lambda_2)-3 = 4\alpha'-5 < 3$ and for $|\hat{k}| \to \infty$ by the exponential decay of $\hat{\psi}_1^2(\hat{k})$.

6.3. Proof of Corollary 4.4

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

$$\widehat{\partial_1^2 v}(k, a_0) = \frac{k_1^2}{a_0 k_1^2 - ik_2}\hat{f}(k).$$

Derivatives with respect to a_0 and a_0' do not change the regularity either as can be seen from

$$\left(\frac{\partial}{\partial a_0}\right)^n \hat{G}(k, a_0) = \frac{(-1)^n n! k_1^{2n}}{(a_0 k_1^2 - ik_2)^n}\hat{G}(k, a_0) \qquad n \geqq 1, \tag{6.11}$$

and for every n the symbol $\frac{(-1)^n n! k_1^{2n}}{(a_0 k_1^2 - ik_2)^n}$ is also bounded. Therefore, the estimate (4.10) follows immediately from (4.9) either with f_ε in the role of f' (that is $\hat{M}_1 = \hat{\varphi}_{\varepsilon_1}$) or $\left(\frac{\partial}{\partial a_0'}\right)^m \partial_1^2 v_\varepsilon(\cdot, a_0')$ in the role of f' which amounts to

$$\hat{M}_1(k) = \frac{(-1)^m m! k_1^{2m}}{(a_0 k_1^2 - ik_2)^m}\frac{-k_1^2}{a_0' k_1^2 - ik_2}\hat{\varphi}_{\varepsilon_1}(k),$$

and with $\left(\frac{\partial}{\partial a_0}\right)^n v_{\varepsilon_0}(\cdot, a_0)$ in the role of v', that is

$$\hat{M}_2(k) = \frac{(-1)^n n! k_1^{2n}}{(a_0' k_1^2 - ik_2)^n}\hat{\varphi}_{\varepsilon_0}(k).$$

For the derivatives with respect to ε_i the multipliers $\hat{M}_1$, $\hat{M}_2$ are the same as above only with $\hat{\varphi}_{\varepsilon_i}$ replaced by $|\varepsilon_i \frac{\partial}{\partial \varepsilon_i}\hat{\varphi}_{\varepsilon_i}| \lesssim \left((k_1^4+k_2^2)\varepsilon\right)^{\frac{\kappa_i}{4}}$ in $\hat{M}_2$ for $i = 0$ and in $\hat{M}_1$ if $i = 1$.

6.4. Proof of Proposition 4.2

Step 1. Bound on the supremum over x and T. Our first claim is that for all $a_0, a_0' \in [\lambda, \frac{1}{\lambda}]$, $\varepsilon_0, \varepsilon_1 \in (0, 1]$, for $\kappa \ll 1$ and for all $n, m \geqq 1$ and $i = 0, 1$ we have

$$\Big\langle\Big(\sup_{T\leqq 1}(T^{\frac14})^{2-2\alpha}\Big\|\Big(\frac{\partial}{\partial a_0}\Big)^n\Big(\frac{\partial}{\partial a_0'}\Big)^m[v_{\varepsilon_0}(\cdot,a_0),(\cdot)_T] \diamond\{f_{\varepsilon_1},\partial_1^2 v_{\varepsilon_1}(\cdot,a_0')\}\Big\|\Big)^p\Big\rangle^{\frac1p}\lesssim 1, \tag{6.12}$$

$$\Big\langle\Big(\sup_{T\leqq 1}(T^{\frac14})^{2-2\alpha+\kappa}\Big\|\varepsilon_i\frac{\partial}{\partial\varepsilon_i}\Big(\frac{\partial}{\partial a_0}\Big)^n\Big(\frac{\partial}{\partial a_0'}\Big)^m[v_{\varepsilon_0}(\cdot,a_0),(\cdot)_T] \diamond\{f_{\varepsilon_1},\partial_1^2 v_{\varepsilon_1}(\cdot,a_0')\}\Big\|\Big)^p\Big\rangle^{\frac1p}\lesssim \varepsilon_i^{\frac{\kappa}{4}} \quad\text{for all } p<\infty. \tag{6.13}$$

To keep the notation concise, for the moment we restrict ourselves to the bound for $[v_{\varepsilon_0},(\cdot)_T]\diamond\partial_1^2 v_{\varepsilon_1}$ without the derivatives with respect to a_0, a_0', ε_i. The general case of (6.12) follows in the identical way and so does (6.13) if in the proof (4.10) is replaced by (4.11). To simplify the notation further we drop the subscript ε_i 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 (see for example [19, Chapter 1], [3, Section 1.6], or [18, Section 3]) for random variables in the second Wiener chaos and by stationarity, the bound (4.10) can be upgraded to

$$\langle|[v,(\cdot)_T]\diamond\partial_1^2 v|^p\rangle^{\frac1p}\lesssim (T^{\frac14})^{2\alpha'-2} \quad\text{for all } p<\infty. \tag{6.14}$$

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 $T\leqq 1$ can be reduced to a supremum over all dyadic T. For any given $T\leqq 1$ there is a unique a dyadic $T'\leqq\frac12$ such that $T=2T'+t$ with $2T'\leqq T<4T'$ and we refer to this choice when we write T' in the sequel.

We make use of the commutator identity (3.7) in the form of

$$[v,(\cdot)_T]\diamond\partial_1^2 v=\big([v,(\cdot)_{T'}]\diamond\partial_1^2 v\big)_{T'+t}+[v,(\cdot)_{T'+t}](\partial_1^2 v)_{T'}. \tag{6.15}$$

The second term on the right hand side of (6.15) can be bounded directly by making the convolution with $\psi_{T'+t}$ explicit

$$\begin{aligned}
&\Big|\big([v,(\cdot)_{T'+t}](\partial_1^2 v)_{T'}\big)(x)\Big|\\
&=\Big|\int (v(x)-v(y))\psi_{T'+t}(x-y)(\partial_1^2 v)_{T'}(y)\,dy\Big|\\
&\leqq [v]_\alpha\|(\partial_1^2 v)_{T'}\|\int d^\alpha(x,y)|\psi_{T'+t}(x-y)|\,dy\\
&\overset{(2.4)}{\lesssim}(T'^{\frac14})^{\alpha-2}((T'+t)^{\frac14})^\alpha[v]_\alpha^2\overset{(A.1)}{\lesssim}(T^{\frac14})^{2\alpha-2}\|f\|_{\alpha-2}^2.
\end{aligned}$$

Derivatives with respect to a_0, a_0' can be dealt with as in Step 1 of the proof of Corollary 3.7.

Taking the sup over x and T and then the p-th moment in the expectation we get, from Lemma 4.1,

$$\Big\langle\Big(\sup_{T\leqq 1}(T^{\frac14})^{2-2\alpha}\|[v,(\cdot)_{T'+t}](\partial_1^2 v)_{T'}\|\Big)^p\Big\rangle^{\frac1p}\lesssim\big\langle\|f\|_{\alpha-2}^{2p}\big\rangle^{\frac1p}\lesssim 1.$$

To bound the first term on the right hand side of (6.15) we use Young's inequality (on the torus) in the form

$$\|\big([v,(\cdot)_{T'}]\diamond\partial_1^2 v\big)_{T'+t}\| \lesssim \|[v,(\cdot)_{T'}]\diamond\partial_1^2 v\|_{L^p}\|\psi_{T'+t,\mathrm{per}}\|_{L^{p'}},$$

where we use the notation of Step 1 in the proof of Lemma 4.1, resulting in

$$\|\big([v,(\cdot)_{T'}]\diamond\partial_1^2 v\big)_{T'+t}\| \lesssim ((T'+t)^{\frac{1}{4}})^{-\frac{3}{p}}\|[v,(\cdot)_{T'}]\diamond\partial_1^2 v\|_{L^p}.$$

Taking the supremum over T we get for any p

$$\begin{aligned}&\Big(\sup_{T\leqq 1}(T^{\frac{1}{4}})^{2-2\alpha}\|\big([v,(\cdot)_{T'}]\diamond\partial_1^2 v\big)_{T'+t}\|\Big)^p\\ &\quad\lesssim \sum_{T'\leqq\frac{1}{2},\mathrm{dyadic}}(T'^{\frac{1}{4}})^{p(2-2\alpha)}((T')^{\frac{1}{4}})^{-3}\|[v,(\cdot)_{T'}]\diamond\partial_1^2 v\|_{L^p}^p.\end{aligned}$$

Finally, we take the expectation of this estimate and use (6.14) and the stationarity to get

$$\begin{aligned}&\Big\langle\Big(\sup_{T\leqq 1}(T^{\frac{1}{4}})^{2-2\alpha}\|\big([v,(\cdot)_{T'}]\diamond\partial_1^2 v\big)_{T'+t}\|\Big)^p\Big\rangle\\ &\quad\lesssim \sum_{T'\leqq\frac{1}{2},\mathrm{dyadic}}(T'^{\frac{1}{4}})^{p(2-2\alpha)}(T'^{\frac{1}{4}})^{-3}\big\langle\|[v,(\cdot)_{T'}]\diamond\partial_1^2 v\|_{L^p}^p\big\rangle\\ &\quad\lesssim \sum_{T'\leqq\frac{1}{2},\mathrm{dyadic}}(T'^{\frac{1}{4}})^{p(2\alpha'-2\alpha)}(T'^{\frac{1}{4}})^{-3}.\end{aligned}$$

Estimate (6.12) for $p>\frac{3}{2(\alpha'-\alpha)}$ then follows by summing this geometric series. The same bound for smaller p can be derived from the bound for large p and Jensen's inequality.

STEP 2. Bounding the supremum over a_0, a_0'. In the following steps we use the abbreviation

$$A(\cdot,T,a_0,a_0',\varepsilon_0,\varepsilon_1)=\Big(\frac{\partial}{\partial a_0}\Big)^n\Big(\frac{\partial}{\partial a_0'}\Big)^m[v_{\varepsilon_0}(\cdot,a_0),(\cdot)_T]\diamond\{f_{\varepsilon_1},\partial_1^2 v_{\varepsilon_1}(\cdot,a_0')\}. \tag{6.16}$$

In this step we show that for $\varepsilon_0,\varepsilon_1\in(0,1]$ and $\kappa\ll 1$

$$\Big\langle\Big(\sup_{a_0,a_0'\in[\lambda,\frac{1}{\lambda}]}\sup_{T\leqq 1}(T^{\frac{1}{4}})^{2-2\alpha}\|A\|\Big)^p\Big\rangle^{\frac{1}{p}}\lesssim 1, \tag{6.17}$$

$$\Big\langle\Big(\sup_{a_0,a_0'\in[\lambda,\frac{1}{\lambda}]}\sup_{T\leqq 1}(T^{\frac{1}{4}})^{2-2\alpha}\Big\|\varepsilon_i\frac{\partial}{\partial\varepsilon_i}A\Big\|\Big)^p\Big\rangle^{\frac{1}{p}}\lesssim \varepsilon_i^{\frac{\kappa}{4}}\qquad\text{for all } p<\infty. \tag{6.18}$$

For (6.17) we use the Sobolev inequality

$$\sup_{a_0,a_0'\in[\lambda,\frac{1}{\lambda}]}\|A\|^p\lesssim\int_{[\lambda,\frac{1}{\lambda}]}\int_{[\lambda,\frac{1}{\lambda}]}\Big\|\Big\{1,\frac{\partial}{\partial a_0},\frac{\partial}{\partial a_0'}\Big\}A\Big\|^p\,\mathrm{d}a_0\,\mathrm{d}a_0',$$

which holds for $p > 2$. Taking the supremum over T, then the expectation and invoking Fubini's theorem and (6.12) yields

$$\Big\langle\Big(\sup_{a_0,a_0'\in[\lambda,\frac{1}{\lambda}]}\sup_{T\leqq 1}(T^{\frac14})^{2-2\alpha}\|A\|\Big)^p\Big\rangle$$
$$\lesssim \int_{[\lambda,\frac{1}{\lambda}]}\int_{[\lambda,\frac{1}{\lambda}]}\Big\langle\Big(\sup_{T\leqq 1}(T^{\frac14})^{2-2\alpha}\Big\|\Big\{1,\frac{\partial}{\partial a_0},\frac{\partial}{\partial a_0'}\Big\}A\Big\|\Big)^p\Big\rangle \mathrm{d}a_0\,\mathrm{d}a_0' \lesssim 1,$$

so (6.17) follows. For (6.18) we repeat the same calculation with A replaced by $\varepsilon_i\frac{\partial}{\partial\varepsilon_i}A$ and (6.12) replaced by (6.13).

STEP 3. Bounding the supremum over ε_i. Let A be defined as in (6.16) above. In this step we upgrade (6.17)and (6.18) to

$$\Big\langle\Big(\sup_{\varepsilon_0,\varepsilon_1\in(0,1]}\sup_{a_0,a_0'\in[\lambda,\frac{1}{\lambda}]}\sup_{T\leqq 1}(T^{\frac14})^{2-2\alpha}\|A\|\Big)^p\Big\rangle^{\frac1p} \lesssim 1, \tag{6.19}$$

valid for $\alpha < \alpha'$. As in the previous step, we use the Sobolev inequality

$$\sup_{\varepsilon_0,\varepsilon_1\in(0,1]}|A(\varepsilon)|^p \lesssim \int_{[0,1]}\int_{[0,1]}\Big|\Big\{1,\frac{\partial}{\partial\varepsilon_0},\frac{\partial}{\partial\varepsilon_1}\Big\}A(\varepsilon)\Big|^p\,\mathrm{d}\varepsilon_0\,\mathrm{d}\varepsilon_1,$$

valid for $p > 2$. We now multiply with $(T^{\frac14})^{2-\alpha+\kappa}$ for some $\alpha < \alpha'$ and $0 < \kappa \ll 1$, take the supremum over x, T, a_0, a_0' of this estimate and finally take the expectation to arrive at

$$\Big\langle\Big(\sup_{\varepsilon_0,\varepsilon_1\in(0,1]}\sup_{a_0,a_0'\in[\lambda,\frac{1}{\lambda}]}\sup_{T\leqq 1}(T^{\frac14})^{2-2\alpha+\kappa}\|A\|\Big)^p\Big\rangle^{\frac1p}$$
$$\lesssim \int_{[0,1]}\int_{[0,1]}\Big\langle\Big(\sup_{a_0,a_0'\in[\lambda,\frac{1}{\lambda}]}\sup_{T\leqq 1}(T^{\frac14})^{2-2\alpha+\kappa}\Big\|\Big\{1,\frac{\partial}{\partial\varepsilon_0},\frac{\partial}{\partial\varepsilon_1}\Big\}A\Big\|\Big)^p\Big\rangle^{\frac1p}\mathrm{d}\varepsilon_0\,\mathrm{d}\varepsilon_1$$
$$\overset{(6.12),(6.13)}{\lesssim} \int_{[0,1]}\int_{[0,1]}\big\{1,\varepsilon_0^{\frac{\kappa}{4}-1},\varepsilon_1^{\frac{\kappa}{4}-1}\big\}\,\mathrm{d}\varepsilon_0\,\mathrm{d}\varepsilon_1 \lesssim 1.$$

Now (6.19) follows by relabelling $-2\alpha+\kappa$ as -2α.

STEP 4. Bounding ε differences. In this step we only consider the diagonal where $\varepsilon_0 = \varepsilon_1 = \varepsilon$ in A defined in (6.16) and simply write $A(\varepsilon)$ instead of $A(\varepsilon,\varepsilon)$. Note that with this notation

$$\varepsilon\frac{\partial}{\partial\varepsilon}A(\varepsilon) = \varepsilon_0\frac{\partial}{\partial\varepsilon_0}A(\varepsilon_0,\varepsilon_1)\Big|_{\varepsilon_0=\varepsilon_1=\varepsilon} + \frac{\partial}{\partial\varepsilon_1}A(\varepsilon_0,\varepsilon_1)\Big|_{\varepsilon_0=\varepsilon_1=\varepsilon}.$$

We claim that for $\kappa \ll 1$ and all $p < \infty$ and $\alpha < \alpha'$

$$\bigg\langle\bigg(\sup_{a_0,a_0'\in[\lambda,\frac{1}{\lambda}]}\sup_{\varepsilon_0\neq\varepsilon_1\in(0,1]}\sup_{T\leqq 1}(T^{\frac14})^{2-2\alpha+\kappa}|\varepsilon_1-\varepsilon_0|^{-\frac{\kappa}{4}}\|A(\varepsilon_1)-A(\varepsilon_0)\|\bigg)^p\bigg\rangle^{\frac1p} \lesssim 1. \tag{6.20}$$

We start the argument with Sobolev's inequality

$$\sup_{\varepsilon_0\neq\varepsilon_1\in(0,1]} \frac{|A(\varepsilon)-A(\bar\varepsilon)|}{|\varepsilon_1-\varepsilon_0|^{\frac{\kappa}{4}}} \leqq \Big(\int_0^1 \big|\frac{\partial}{\partial\varepsilon}A(\varepsilon)\big|^{\frac{1}{1-\frac{\kappa}{4}}}\,\mathrm{d}\varepsilon\Big)^{1-\frac{\kappa}{4}}.$$

Now, we multiply this estimate with $(T^{\frac14})^{2-2\alpha+\kappa+\bar\kappa}$ for another $0<\bar\kappa\ll 1$, take the supremum over x, T, a_0 and a_0', then p-th moments, and finally invoke Minkowski's inequality (for $p>\frac{1}{1-\frac{\kappa}{4}}$) and (6.13) to get

$$\begin{aligned}
&\Big\langle\Big(\sup_{a_0,a_0'\in[\lambda,\frac1\lambda]}\ \sup_{\varepsilon_0\neq\varepsilon_1\in(0,1]}\ \sup_{T\leqq 1}(T^{\frac14})^{2-2\alpha+\kappa+\bar\kappa}|\varepsilon_1-\varepsilon_0|^{-\frac{\kappa}{4}}\|A(\varepsilon_1)-A(\varepsilon_0)\|\Big)^p\Big\rangle^{\frac1p}\\
&\lesssim \Big(\int_0^1\Big\langle\Big(\sup_{a_0,a_0'\in[\lambda,\frac1\lambda]}\ \sup_{T\leqq 1}(T^{\frac14})^{2-2\alpha+\kappa+\bar\kappa}\Big\|\frac{\partial}{\partial\varepsilon}A\Big\|^p\Big)\Big\rangle^{\frac1p\frac{1}{1-\frac{\kappa}{4}}}\mathrm{d}\varepsilon\Big)^{1-\frac{\kappa}{4}}\\
&\lesssim \int_0^1 \varepsilon^{(\frac{\kappa+\bar\kappa}{4}-1)\frac{1}{1-\frac{\kappa}{4}}}\,\mathrm{d}\varepsilon\lesssim 1,
\end{aligned}$$

so (6.20) follows by relabelling $-2\alpha+\bar\kappa$ as -2α.

STEP 5. Conclusion. To shorten notation, we only treat the product $v_\varepsilon\diamond f_\varepsilon$. Writing

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

and invoking (4.3) and (4.4) for the first and (6.20) for the second term, implies 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 (6.19) and (6.20) remain true if the supremum over $\varepsilon\in(0,1]$ is extended to include the limit as $\varepsilon\to 0$.

6.5. *Proof of Lemma 4.5*

STEP 1. Proof of (i). By stationarity and (4.1) we may write

$$\begin{aligned}
g_1(\varepsilon,a_0) &\overset{(1.11)}{=} \langle v_\varepsilon(0,a_0)f_\varepsilon(0)\rangle\\
&= \Big\langle\int_{[0,1)^2} v_\varepsilon(x,a_0)f_\varepsilon(x)\,\mathrm{d}x\Big\rangle = \sum_{k\in(2\pi\mathbb{Z})^2}\langle\hat v_\varepsilon(k,a_0)\hat f_\varepsilon(-k)\rangle\\
&= \sum_{k\in(2\pi\mathbb{Z})^2}\hat G(k,a_0)\langle f_\varepsilon(k)f_\varepsilon(-k)\rangle = \sum_{k\in(2\pi\mathbb{Z})^2}\hat G(k,a_0)\hat C(k)|\hat\varphi_\varepsilon(k)|^2,
\end{aligned}$$

where $\hat G$ denotes the Fourier transform of the Greens function introduced in (6.2) above. 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 $\hat C$ is real valued this means that we can replace $\hat G(\cdot,a_0)$ by its real part (given in (6.2)) thereby yielding (4.12).

The same calculation yields

$$g_2(\varepsilon, a_0, a_0') \overset{(1.11)}{=} \langle v_\varepsilon(0, a_0)\partial_1^2 v_\varepsilon(0, a_0')\rangle$$
$$= \sum_{k\in(2\pi\mathbb{Z})^2} \hat{G}(k, a_0)(-k_1^2)\hat{G}(k, a_0')\hat{C}(k)|\hat{\varphi}_\varepsilon(k)|^2,$$

and after calculating the real part of $\hat{G}(k, a_0)(-k_1^2)\hat{G}(k, a_0')$ we arrive at (4.13).

STEP 2. Proof of (ii). By the condition $a_0 \in [\lambda, \frac{1}{\lambda}]$ we immediately see from (4.12) that convergence of $g_1(\varepsilon, a_0)$ is equivalent to (4.14). Furthermore, given that the ratio of the kernels appearing in (4.12) and (4.13) is bounded as

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

(4.14) also implies the convergence of the $g_2(\varepsilon, a_0, a_0')$ as ε goes to zero. The condition (4.14) also implies the convergence for arbitrary derivatives of g_1, g_2 with respect to a_0, a_0'. For example, recalling the fact that the term $\frac{a_0k_1^2}{a_0^2k_1^4+k_2^2}$ is nothing but the real part $\Re$ of $\hat{G}(k, a_0)$ we can write

$$\left(\frac{\partial}{\partial a_0}\right)^n g_1(\varepsilon, a_0) = \sum_{k\in(2\pi\mathbb{Z})^2\setminus\{0\}} \Re\Big(\Big(\frac{\partial}{\partial a_0}\Big)^n \hat{G}(k, a_0)\Big)\hat{C}(k)|\hat{\varphi}_\varepsilon(k)|^2$$
$$\overset{(6.11)}{=} \sum_{k\in(2\pi\mathbb{Z})^2\setminus\{0\}} \Re\left(\frac{(-1)^n n! k_1^{2n}}{(a_0k_1^2 - ik_2)^n}\hat{G}(k, a_0)\right)\hat{C}(k)|\hat{\varphi}_\varepsilon(k)|^2.$$

Given that for any $n \geqq 1$ the modulus 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 (4.14) follows. A similar argument works for g_2.

7. Proofs of Theorems 1.1 and 1.2

According to Lemma 4.1 under assumption (4.2), we have

$$\Big\langle \sup_{\varepsilon\in[0,1]} \|f\|_{\alpha-2}^p\Big\rangle^{\frac{1}{p}} < \infty$$

for any $\alpha < \alpha'$ and $p < \infty$ and we have almost surely and in every stochastic L^p space that $\|f - f_\varepsilon\|_{\alpha-2} \to 0$. Under the same assumption according to Proposition 4.2 the renormalized products $v_\varepsilon(\cdot, a_0) \diamond f_\varepsilon$ and $v_\varepsilon(\cdot, a_0) \diamond \partial_1^2 v_\varepsilon(\cdot, a_0')$ defined in (4.5) converge to limits denoted by $v \diamond f$ and $v \diamond \partial_1^2 v$ as ε goes to zero in the sense that almost surely the quantities

$$\big\|[v_\varepsilon, (\cdot)] \diamond f_\varepsilon - [v, (\cdot)] \diamond f\big\|_{2\alpha-2,2}, \quad \big\|[v_\varepsilon, (\cdot)] \diamond \partial_1^2 v_\varepsilon - [v, (\cdot)] \diamond \partial_1^2 v\big\|_{2\alpha-2,2,2}$$

converge to zero. Furthermore, we have the moment bounds

$$\Big\langle \sup_{\varepsilon_0,\varepsilon_1\in[0,1]} \|[v_{\varepsilon_0},(\cdot)]\diamond f_{\varepsilon_1}\|^p_{2\alpha-2,2}\Big\rangle^{\frac{1}{p}}, \quad \Big\langle \sup_{\varepsilon_0,\varepsilon_1\in[0,1]} \|[v_{\varepsilon_0},(\cdot)]\diamond \partial_1^2 v_{\varepsilon_1}\|^p_{2\alpha-2,2,2}\Big\rangle^{\frac{1}{p}} < \infty \tag{7.1}$$

for all $p < \infty$.

Let $N_0 \ll 1$ be so small that Theorem 3.9 holds and set

$$\eta_0^{-1} = \frac{1}{N_0} \sup_{\varepsilon,\varepsilon_0,\varepsilon_1\in[0,1]} \max\Big\{\|f_\varepsilon\|_{\alpha-2}, \|[v_{\varepsilon_0},(\cdot)]\diamond f_{\varepsilon_1}\|^{\frac{1}{2}}_{2\alpha-2,2,2}, \|[v_{\varepsilon_0},(\cdot)]\diamond \partial_1^2 v_{\varepsilon_1}\|^{\frac{1}{2}}_{2\alpha-2,2,2}\Big\}.$$

Then the moment bound (1.12) holds, and for all $\eta \leqq \eta_0$ the functions/distributions ηf_ε, $\eta^2[v_\varepsilon,(\cdot)]f_\varepsilon$ and $\eta^2[v_\varepsilon,(\cdot)]\partial_1^2 v_\varepsilon$ satisfy the smallness condition (3.81), (3.82), and (3.88) uniformly in $\varepsilon \in [0,1]$. Thus Theorem 3.9 part (i) yields the existence and uniqueness of a solution u to (1.13), (1.14), (1.15), as well as solutions u_ε to the corresponding regularized problems with δ in Theorem 1.1 being the implicit constant in (3.85). By Corollary 3.10 the regularized problem takes the form of (1.10) and part (ii) of Theorem 3.9, more precisely estimate (3.92), yields the convergence to zero of $\|u-u_\varepsilon\| + [u-u_\varepsilon]_\alpha$.

Acknowledgements. Open access funding provided by Max Planck Society. HW is supported by the Royal Society through the University Research Fellowship UF140187.

Compliance with Ethical Standards

Conflict of interest The authors declare that they have no conflict of interest.

Appendix A. Some additional lemmas

Lemma A.1. *The (mean-free) solution of (3.39) satisfies the estimate*

$$\sup_{a_0}[v(\cdot,a_0)]_\alpha \lesssim \|f\|_{\alpha-2}. \tag{A.1}$$

Proof of Lemma A.1

All functions are space-time period if not stated otherwise.
Step 1. Reduction. We claim that it is enough to show

$$\sup_{T\leqq 1}(T^{\frac{1}{4}})^{2-\alpha}\|f_T\| \sim \inf\Big\{[f_1]_\alpha + [f_2]_\alpha + |c| \,\Big|\, f = \partial_1^2 f_1 + \partial_2 f_2 + c\Big\}, \tag{A.2}$$

where the infimum is over all triplets (f_1, f_2, c) of two functions and a constant. Incidentally, the equivalence confirms that the left hand side indeed defines the (parabolic) $C^{\alpha-2}$-norm.

Let the decomposition $f=\partial_1^2 f_1+\partial_2 f_2+c$ be near-optimal in the right hand side of (A.2), that is,

$$[f_1]_\alpha+[f_2]_\alpha \leqq 2\sup_{T\leqq 1}\|f_T\|. \tag{A.3}$$

By the uniqueness of the mean-free solution of (3.39) 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 $(\partial_2-a_0\partial_1^2)v_i=f_i$. By classical $C^{\alpha+2}$-Schauder theory [14, Theorem 8.6.1] we have $[\partial_1^2 v_i]_\alpha+[\partial_2 v_i]_\alpha\lesssim[f_i]_\alpha$, so that (A.1) follows from (A.3).

STEP 2. For the solution of

$$(\partial_2-\partial_1^2)v=Pf \tag{A.4}$$

we claim

$$\|v_T-v\|\lesssim N_0\max\left\{(T^{\frac14})^\alpha,(T^{\frac14})^2\right\}\quad\text{for all } T>0, \tag{A.5}$$

where we have set for abbreviation

$$N_0:=\sup_{T\leqq 1}(T^{\frac14})^{2-\alpha}\|f_T\|. \tag{A.6}$$

We start by noting that the definition of N_0 may be extended to the control of $T\geqq 1$ by the semi-group property (2.3) in form of $f_T=(f_1)_{T-1}$ and (2.4) in form of $\|f_T\|\lesssim\|f_1\|$. We thus have

$$\|f_T\|\lesssim N_0\max\{T^{\alpha-2},1\}. \tag{A.7}$$

By approximation through (standard) convolution, which preserves (A.4) and does increase N_0, we may assume that f and v are smooth. By definition of the convolution $(\cdot)_t$ we have

$$\begin{aligned}\partial_t v_t&=-(\partial_1^4-\partial_2^2)v_t=(-\partial_1^2-\partial_2)(\partial_2-\partial_1^2)v_t\overset{(A.4)}{=}(-\partial_1^2-\partial_2)Pf_t\\&\overset{(2.3)}{=}(-\partial_1^2-\partial_2)(f_{\frac t2})_{\frac t2}.\end{aligned}$$

Hence we obtain by (2.4) for all $T\leqq 1$

$$\|\partial_t v_t\|\lesssim(t^{\frac14})^{-2}\|f_{\frac t2}\|\overset{(A.7)}{\lesssim}N_0\max\left\{(t^{\frac14})^{\alpha-4},(t^{\frac14})^{-2}\right\}.$$

Integrating over $t\in(0,T)$ we obtain (A.5) by the triangle inequality.

STEP 3. For v defined through (A.4) we have

$$[v]_\alpha\lesssim N_0, \tag{A.8}$$

where N_0 is as in (A.6). 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 $d(y,x)\leqq 4$. For any $T>0$ we have the identity

$$\begin{aligned}v(y)-v(x)&=(v_T-v)(y)-(v_T-v)(x)\\&\quad-\int_0^1\partial_1 v_T(sy+(1-s)x)(y-x)_1+\partial_2 v_T(sy+(1-s)x)(y-x)_2\,ds,\end{aligned}$$

from which we obtain the inequality

$$|v(y)-v(x)|\leqq 2\|v_T-v\|+\|\partial_1 v_T\|d(y,x)+\|\partial_2 v_T\|d^2(y,x).$$

From Step 2 and (2.4) we obtain the estimate

$$|v(y)-v(x)| \lesssim N_0 \max\{(T^{\frac14})^\alpha, (T^{\frac14})^2\} + [v]_\alpha\big((T^{\frac14})^{\alpha-1}d(y,x) + (T^{\frac14})^{\alpha-2}d^2(y,x)\big).$$

With the ansatz $T^{\frac14} = \frac{1}{\varepsilon}d(y,x)$ for some $\varepsilon \leqq 1$ and making use of $d(y,x) \leqq 1$ we obtain

$$|v(y)-v(x)| \lesssim \big(\varepsilon^{-2}N_0 + [v]_\alpha(\varepsilon^{1-\alpha}+\varepsilon^{2-\alpha})\big)d^\alpha(y,x).$$

Fixing an ε sufficiently small to absorb the last right-hand-side term into the left hand side we infer (A.8).

Step 4. We finally establish the equivalence of norms (A.2). The direction $\lesssim$ follows immediately from (2.4). The direction $\gtrsim$ follows from Step 3 with $f_1 = v$, $f_2 = -v$, and $c = \int_{[0,1)^2} f$.

Lemma A.2.

$$\|[x_1,(\cdot)]f\|_{\alpha-1} \lesssim \|f\|_{\alpha-2}. \tag{A.9}$$

Proof of Lemma A.2

Introducing the kernel $\tilde\psi_T(x) := x_1\psi_T(x)$ we start by claiming the representation

$$[x_1,(\cdot)_T]f = 2\tilde\psi_{\frac T2} * f_{\frac T2}. \tag{A.10}$$

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

$$\tilde\psi_T = 2\tilde\psi_{\frac T2} * \psi_{\frac T2}. \tag{A.11}$$

The argument for (A.11) relies on the fact that convolution is commutative in form of $\tilde\psi_{\frac T2} * \psi_{\frac T2} = \psi_{\frac T2} * \tilde\psi_{\frac T2}$, which spelled out means $\int dy(x_1-y_1)\psi_{\frac T2}(x-y)\psi_{\frac T2}(y) = \int dy\psi_{\frac T2}(x-y)y_1\psi_{\frac T2}(y)$, and thus implies $2\int dy(x_1-y_1)\psi_{\frac T2}(x-y)\psi_{\frac T2}(y) = x_1\int dy\psi_{\frac T2}(x-y)\psi_{\frac T2}(y)$, that is $2(\tilde\psi_{\frac T2} * \psi_{\frac T2})(x) = x_1(\psi_{\frac T2} * \psi_{\frac T2})(x)$. Together with the semi-group property (2.3) in form of $\psi_{\frac T2} * \psi_{\frac T2} = \psi_T$ this yields (A.11).

From the representation (A.10) we obtain the estimate

$$\|[x_1,(\cdot)_T]f\| \leqq 2\int dx|x_1\psi_{\frac T2}(x)|\|f_{\frac T2}\| \overset{(2.4)}{\lesssim} T^{\frac14}\|f_{\frac T2}\|,$$

which yields the desired (A.9).

The following lemma shows that the definitions (2.5), (3.8) and (3.9) are independent of the choice of convolution kernel:

Lemma A.3. *Let ψ and ψ' be Schwartz functions over $\mathbb{R}^2$ with $\int\psi = \int\psi' = 1$. For $T > 0$ define*

$$\psi_T(x_1,x_2) = T^{-\frac34}\psi\Big(\frac{x_1}{T^{\frac14}}, \frac{x_2}{T^{\frac12}}\Big), \qquad \psi'_T(x_1,x_2) = T^{-\frac34}\psi'\Big(\frac{x_1}{T^{\frac14}}, \frac{x_2}{T^{\frac12}}\Big), \tag{A.12}$$

and for an arbitrary Schwartz distribution $f \in \mathcal{S}'(\mathbb{R}^2)$ set

$$(f)_T = f * \psi_T \quad \text{and} \quad (f)'_T = f * \psi'_T. \tag{A.13}$$

Then
i) For any $\gamma < 0$ we have

$$\sup_{T \leqq 1} (T^{\frac{1}{4}})^{-\gamma} \|(f)_T\| \lesssim \sup_{T \leqq 1} (T^{\frac{1}{4}})^{-\gamma} \|(f)'_T\|, \tag{A.14}$$

where $\lesssim$ only refers to ψ, ψ' and γ.
ii) Let $\alpha > 0$ and $\gamma < 0$. Let u be a 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 generalized commutators $[u, (\cdot)_T] \diamond f := u(f)_T - (u \diamond f)_T$ and $[u, (\cdot)'_T] \diamond f := u(f)'_T - (u \diamond f)'_T$. Then for $\bar{\gamma} = \gamma + \alpha$ we have

$$\begin{aligned} \sup_{T \leqq 1} (T^{\frac{1}{4}})^{-\bar{\gamma}} \|[u, (\cdot)_T] \diamond f\| &\lesssim \sup_{T \leqq 1} (T^{\frac{1}{4}})^{-\bar{\gamma}} \|[u, (\cdot)'_T] \diamond f\| \\ &\quad + [u]_\alpha \sup_{T \leqq 1} (T^{\frac{1}{4}})^{-\gamma} \|(f)'_T\|, \end{aligned} \tag{A.15}$$

where $\lesssim$ depends on α, γ as well as ψ and ψ'.

Proof of Lemma A.3

STEP 1. The proof relies on a variant of a construction from [5] 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

$$\int x^n \varphi'(x)\, dx = \begin{cases} 1 & \text{for } \alpha = 0 \\ 0 & \text{for } 0 < \|n\|_{\text{par}} < p, \end{cases} \tag{A.16}$$

where for $n = (n_1, n_2)$ and $x = (x_1, x_2)$ we write $x^n = x_1^{n_1} x_2^{n_2}$ and use the parabolic norm $\|n\|_{\text{par}} = |n_1| + 2|n_2|$. Furthermore, it is shown that for any p and any φ' satisfying (A.16) as well as $\theta \ll 1$ (depending on φ, ψ, p), that the function ψ can be represented as

$$\psi = \sum_{k=0}^{\infty} \omega^{(k)} * \varphi'_{\theta^k}, \tag{A.17}$$

where φ'_{θ^k} is the rescaled version of φ' defined as in (A.12) for $T = \theta^k$, and the $\omega^{(k)}$ are Schwartz functions satisfying

$$\int |\omega^{(k)}| \lesssim (C_0 \theta^{\frac{p}{4}})^k, \tag{A.18}$$

where $C_0 = C_0(\varphi', \psi, p)$. The convergence of the sum in (A.17) holds in $L^1(\mathbb{R}^2)$. Additionally, we will make use of the bounds

$$\int d^\alpha(0, x) |\omega^{(k)}(x)|\, dx \lesssim (C_0 \theta^{\frac{p}{4}})^k. \tag{A.19}$$

We summarize this as $\psi = \sum_{k=0}^{\infty} \omega^{(k)} * \omega^0_{\theta^k} * \psi'_{\theta^k}$, which can be rescaled as

$$\psi_T = \sum_{k=0}^{\infty} \omega_T^{(k)} * \omega^0_{\theta^k T} * \psi'_{\theta^k T}, \tag{A.20}$$

where as before the index T expresses that a function is rescaled by T as in (A.12).

STEP 2. Equipped with these results we now proceed to prove (A.14). Set $N_0 := \sup_{T \leqq 1} (T^{\frac{1}{4}})^{-\gamma} \|(f)'_T\|$ and write

$$\begin{aligned} \|(f)_T\| &\overset{(A.20)}{=} \|\sum_{k=0}^{\infty} (\omega_T^{(k)} * \omega^0_{\theta^k T}) * (f)'_{\theta^k T}\| \leqq \sum_{k=0}^{\infty} \int |\omega_T^{(k)}| \int |\omega^0_{\theta^k T}| \|(f)'_{\theta^k T}\| \\ &\overset{(A.18)}{\lesssim} N_0 \sum_{k=0}^{\infty} (\theta^{\frac{k}{4}} T^{\frac{1}{4}})^{\gamma} (C_0 \theta^{\frac{p}{4}})^k \int |\omega^0|. \end{aligned}$$

Then (A.14) follows by choosing first $p > |\gamma|$ and then $\theta^{\frac{1}{4}} \leqq \frac{1}{2C_0}$ and then summing the geometric series over k.

STEP 3. We set $N_1 := \sup_{T \leqq 1} (T^{\frac{1}{4}})^{-\bar{\gamma}} \|[u, (\cdot)'_T] \diamond f\|$ and $N'_0 := \sup_{T \leqq 1} (T^{\frac{1}{4}})^{-\gamma} \|(f)'_T\|$ as before. Again, we make use of the representation (A.20) of ψ_T to write

$$[u, (\cdot)_T] \diamond f = \sum_{k=0}^{\infty} [u, \omega_T^{(k)} * \omega^0_{\theta^k T} * \psi'_{\theta^k T} *] \diamond f.$$

We apply the commutator relation $[A, BC] = [A, B]C + B[A, C]$ twice, to rewrite each term in this sum as

$$\begin{aligned} &[u, \omega_T^{(k)} * \omega^0_{\theta^k T} * \varphi_{\theta^k T} *] \diamond f \\ &= [u, \omega_T^{(k)} * \omega^0_{\theta^k T} *](f)'_{\theta^k T} + \omega_T^{(k)} * \omega^0_{\theta^k T} * [u, (\cdot)'_{\theta^k T}] \diamond f \\ &= [u, \omega_T^{(k)} *](\omega^0_{\theta^k T} * (f)'_{\theta^k T}) + \omega_T^{(k)} * ([u, \omega^0_{\theta^k T} *](f)'_{\theta^k T}) \\ &\quad + \omega_T^{(k)} * \omega^0_{\theta^k T} * [u, (\cdot)'_{\theta^k T}] \diamond f. \end{aligned} \tag{A.21}$$

Note that only the last commutator on the rhs requires the definition of $u \diamond f$ and all the other commutators are defined classically. We bound the terms on the right hand side of (A.21) one by one, starting with the last. This expression can be directly bounded

$$\|(\omega_T^{(k)} * \omega^0_{\theta^k T}) * [u, (\cdot)'_{\theta^k T}] \diamond f\| = \int |\omega^{(k)}| \int |\omega^0| (\theta^{\frac{k}{4}} T^{\frac{1}{4}})^{\bar{\gamma}} N_1.$$

Therefore, the sum in k over this term is controlled by invoking (A.18) for p large enough, then choosing θ small enough, resulting with a geometric series as in Step 2.

By Young's inequality, the second term on the right hand side of (A.21) is bounded as follows:

$$\|\omega_T^{(k)} * ([u, \omega^0_{\theta^k T} *](f)'_{\theta^k T})\| \leqq \int |\omega^{(k)}| \; \|[u, \omega^0_{\theta^k T} *](f)'_{\theta^k T}\|.$$

According to (A.18) 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

$$\begin{aligned}
&\|[u,\omega^0_{\theta^k T}*](f)'_{\theta^k T}\| \\
&\quad = \sup_x \Big| \int (u(x)-u(y))\omega^0_{\theta^k T}(y-x)(f)'_{\theta^k T}(y)\,\mathrm{d}y \Big| \\
&\quad \leqq [u]_\alpha N_0(\theta^{\frac{k}{4}}T^{\frac14})^\gamma \sup_x \int d^\alpha(x,y)|\omega^0_{\theta^k T}(y-x)|\,\mathrm{d}y \\
&\quad = [u]_\alpha N_0(\theta^{\frac{k}{4}}T^{\frac14})^{\bar\gamma} \int d^\alpha(0,z)|\omega^0(z)|\,\mathrm{d}z,
\end{aligned}$$

so that summing these terms over k also yields the required bound as above.

It remains to bound the first term on the rhs of (A.21) and for this we write

$$\begin{aligned}
&\|[u,\omega^{(k)}_T *](\omega^0_{\theta^k T} * (f)'_{\theta^k T})\| \\
&\quad \leqq \sup_x \int |u(y)-u(x)|\,|\omega^{(k)}_T(y-x)|\,\mathrm{d}y \Big(\int |\omega^0_{\theta^k T}|\Big)\|(f)'_{\theta^k T}\| \\
&\quad \leqq [u]_\alpha (T^{\frac14})^\alpha N_0(\theta^{\frac{k}{4}}T^{\frac14})^\gamma \int d^\alpha(0,z)|\omega^{(k)}(z)|\,\mathrm{d}z \Big(\int|\omega^0|\Big).
\end{aligned}$$

The first integral on the rhs is bounded $\lesssim (C_0\theta^{\frac{p}{4}})^k$ in (A.19), so that finally (A.15) 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 (A.16), (A.17) and (A.19) following [5]. The construction of ω^0 is based on the identity

$$\begin{aligned}
A_{n,m} &:= \int x^n \partial^m \psi'(x)\,\mathrm{d}x \\
&= \begin{cases} 0 & \text{if } \|n\|_{\mathrm{par}} \leqq \|m\|_{\mathrm{par}},\ n \neq m \\ (-1)^{|m_1|+|m_2|} m_1! m_2! & \text{if } n = m \end{cases}.
\end{aligned}$$

This trigonal structure implies that for any fixed p the linear map

$$(a_m)_{\|m\|_{\mathrm{par}}<p} \mapsto \Bigg(\sum_{\|m\|_{\mathrm{par}}<p} A_{n,m} a_m\Bigg)_{\|n\|_{\mathrm{par}}<p}$$

is invertible. Furthermore, for each n, m the numbers $A^r_{n,m} := \int x^n \partial^m (\psi'_r * \psi')(x)\,\mathrm{d}x$ converge to $A_{n,m}$ as $r \to 0$ and for $r > 0$ small enough the linear map associated to $(A^r_{n,m})_{\|n\|_{\mathrm{par}}, \|m\|_{\mathrm{par}} < p}$ is still invertible. This implies in particular the existence of coefficients (a_m) such that

$$\sum_{\|m\|_{\mathrm{par}} < p} A^r_{n,m} a_m = \sum_{\|m\|_{\mathrm{par}} < p} a_m \int x^n \partial^m (\psi'_r * \psi')(x)\,\mathrm{d}x = \left\{ \begin{matrix} 1 & \text{if } n = 0 \\ 0 & \text{else} \end{matrix} \right\}.$$

The identity (A.16) thus follows for $\omega^0 = \sum_{\|m\|_{\mathrm{par}} < p} a_m \partial^m \psi'_r$.
The key ingredient for the proof of (A.17) and (A.19) are the following estimates (A.22)–(A.25). We claim that for an arbitrary Schwartz function ω and any multi-index $m = (m_1, m_2)$ with $\|m\|_{\mathrm{par}} \leqq p + 1$ we have for any $T > 0$

$$\int |\partial^m (\omega - \varphi'_T * \omega)| \leqq C_0 \int |\partial^m \omega|, \tag{A.22}$$

$$\int d^\alpha(0, x) |\partial^m (\omega - \varphi'_T * \omega)|\,\mathrm{d}x \leqq C_0 \Big(\int d^\alpha(0, x) |\partial^m \omega|\,\mathrm{d}x + (T^{\frac{1}{4}})^\alpha \int |\partial^m \omega|\,\mathrm{d}x \Big). \tag{A.23}$$

Furthermore, for $T \leqq 1$,

$$\int |\omega - \varphi'_T * \omega| \leqq C_0 (T^{\frac{1}{4}})^p \sum_{\|m\|_{\mathrm{par}} = p, p+1} \int |\partial^m \omega| \tag{A.24}$$

$$\int d^\alpha(0, x) |\omega - \varphi'_T * \omega| \leqq C_0 (T^{\frac{1}{4}})^p \sum_{\|m\|_{\mathrm{par}} = p, p+1} \Big(\int d^\alpha(0, x) |\partial^m \omega| + (T^{\frac{1}{4}})^\alpha \int |\partial^m \omega| \Big), \tag{A.25}$$

where we have $C_0 = C_0(p, \varphi')$ in (A.22)–(A.25). The estimates (A.24) and (A.25) rely on the Assumption (A.16) that φ' integrates to zero against monomials of degree $0 < \|n\|_{\mathrm{par}} < p$. Once these bounds are established, the representation (A.17) 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 (A.22) and (A.23) yields

$$\sum_{\|m\|_{\mathrm{par}} = p, p+1} \int (1 + d^\alpha(0, x)) |\partial^m \omega^{(k)}|\,\mathrm{d}x \leqq (2C_0)^k \sum_{\|m\|_{\mathrm{par}} = p, p+1} \int (1 + d^\alpha(0, x)) |\partial^m \psi|\,\mathrm{d}x,$$

which can then be plugged into (A.24) and (A.25) to yield

$$\int (1 + d^\alpha(0, x)) |\omega^{(k+1)}|\,\mathrm{d}x$$

$$\leqq (2C_0)^{k+1}(\theta^{\frac{k}{4}})^p \sum_{\|m\|_{\mathrm{par}}=p,p+1} \int (1 + d^\alpha(0,x))|\partial^m \psi| \,\mathrm{d}x,$$

which in turn yields (A.18) and (A.19). The representation then follows by observing that

$$\psi = \omega^{(0)} = \omega^{(0)} * \varphi' + \omega^{(1)} = \omega^{(0)} * \varphi' + \omega^{(1)} * \varphi'_\theta + \omega^{(2)} = \dots,$$

which, together with (A.18), implies that the convergence holds in L^1.

The bounds (A.22) and (A.24) are provided in the discussion following equation (295) in [5] (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 (A.23) and (A.25) which follow along similar lines. First of all, in order to bound $\int d^\alpha(0,x)|\partial^m \omega - \varphi'_T * \partial^m \omega| \,\mathrm{d}x$ we make use of the triangle inequality in the form $|\partial^m \omega - \varphi'_T * \partial^m \omega| \leqq |\partial^m \omega| + |\varphi'_T * \partial^m \omega|$. The integral resulting from the first term then already has the desired form. For the second term, we write $|\varphi'_T * \partial^m \omega(x)| \leqq \int |\varphi'_T(x-y)\partial^m \omega(y)| \,\mathrm{d}y$ and use the triangle inequality once more, this time in the form $d^\alpha(0,x) \leqq d^\alpha(0,x-y) + d^\alpha(0,y)$. It hence remains to bound the two integrals

$$\begin{aligned} &\int\int d^\alpha(0,x-y)|\varphi'_T(x-y)|\,|\partial^m \omega(y)| \,\mathrm{d}x\,\mathrm{d}y \\ &\quad = (T^{\frac{1}{4}})^\alpha \int d^\alpha(0,\hat z)|\varphi'(\hat z)| \,\mathrm{d}\hat z \int |\partial^m \omega(y)| \,\mathrm{d}y, \\ &\int\int d^\alpha(0,y)|\varphi'_T(x-y)|\,|\partial^m \omega(y)| \,\mathrm{d}x\,\mathrm{d}y \\ &\quad \leqq \int |\varphi'_T(z)| \,\mathrm{d}z \int d^\alpha(0,y)|\partial^m \omega(y)| \,\mathrm{d}y, \end{aligned}$$

and estimate (A.23) follows.

To obtain (A.25), similar to [5] we obtain the pointwise bound

$$\begin{aligned} &|\varphi'_T * \omega - \omega|(x) \\ &\quad \leqq 2 \sum_{\|m\|_{\mathrm{par}}=p,p+1} \int_0^1 \int d^{\|m\|_{\mathrm{par}}}(0,z)|\varphi'_T(-z)|\,|\partial^m \omega(x+sz)| \,\mathrm{d}z\,\mathrm{d}s. \end{aligned} \tag{A.26}$$

We recall the argument from [5] (adjusted to the case of parabolic scaling): First, according to (A.16) φ' integrates non-constant monomials of (parabolic) degree $< p$ to zero which permits us to write $(\varphi'_T * \omega - \omega)(x) = \int \left(\omega(x+z) - \sum_{\|m\|_{\mathrm{par}}<p} \frac{1}{m_1! m_2!}\partial^m \omega(x) z^m\right) \varphi'_T(-z) \,\mathrm{d}z$. At this point we seek to apply Taylor's formula, but unlike [5] we need an anisotropic version of the error term. In order to formulate this we define, for $m = (m_1, m_2)$,

$$F^m := \frac{\partial^m \omega(x) z^m}{(m_1+m_2)!} \qquad E^m := \int_0^1 \frac{(1-s)^{m_1+m_2-1}}{(m_1+m_2-1)!} z^m \partial^m \omega(x+sz) \,\mathrm{d}s,$$

and observe the elementary identities $\omega(x+z)-\omega(x) = E^{(1,0)} + E^{(0,1)}$ as well as $E^m = F^m + E^{(m_1+1,m_2)} + E^{(m_1,m_2+1)}$ which permit its to recursively obtain

$$\begin{aligned}
&\Big|\omega(x+z) - \sum_{\|m\|_{\mathrm{par}}<p} \frac{1}{m_1!m_2!}\partial^m \omega(x) z^m\Big| \\
&= \Big|\sum_{\|m\|_{\mathrm{par}}=p} \binom{m_1+m_2}{m_1} E^{(m_1,m_2)} + \sum_{\|m\|_{\mathrm{par}}=p-1} \binom{m_1+m_2}{m_1} E^{(m_1,m_2+1)}\Big| \\
&\leqq \sum_{\|m\|_{\mathrm{par}}=p,p+1} \binom{m_1+m_2}{m_1} \big|E^{(m_1,m_2)}\big|.
\end{aligned}$$

Then bounding $|z^m| \leqq d^{\|m\|_{\mathrm{par}}}(0,z)$ and observing that the combinatorial pre-factor satisfies $\frac{1}{(m_1+m_2-1)!}\binom{m_1+m_2}{m_1} \leqq 2$ and dropping $(1-s)^{m_1+m_2-1} \leqq 1$ the claimed inequality (A.26) follows.
To bound $\int d^\alpha(0,x)|\varphi'_T * \omega - \omega|(x)\,\mathrm{d}x$ we then use the triangle inequality in the form $d^\alpha(0,x) \leqq d^\alpha(0,z) + d^\alpha(0,x+sz)$, which prompts as to bound the two integrals

$$\begin{aligned}
&\int\int_0^1\int d^{\alpha+\|m\|_{\mathrm{par}}}(0,z)|\varphi'_T(-z)|\,|\partial^m\omega(x+sz)|\,\mathrm{d}z\,\mathrm{d}s\,\mathrm{d}x \\
&\quad= \Big(\int d^{\alpha+\|m\|_{\mathrm{par}}}(0,z)|\varphi'_T(-z)|\,\mathrm{d}z\Big)\Big(\int |\partial^m\omega(x)|\,\mathrm{d}x\Big), \\
&\int\int_0^1\int d^{\alpha}(0,x+sz) d^{\|m\|_{\mathrm{par}}}(0,z)|\varphi'_T(-z)|\,|\partial^m\omega(x+sz)|\,\mathrm{d}z\,\mathrm{d}s\,\mathrm{d}x \\
&\quad= \Big(\int d^{\|m\|_{\mathrm{par}}}(0,z)|\varphi'_T(-z)|\,\mathrm{d}z\Big)\Big(\int d^\alpha(0,y)|\partial^m\omega(y)|\,\mathrm{d}y\Big),
\end{aligned}$$

both of which are bounded as claimed in (A.25).

References

1. Bailleul, I., Bernicot, F.: Heat semigroup and singular PDEs. *J. Funct. Anal.* **270**(9), 3344–3452 (2016)
2. Bailleul, I., Debussche, A., Hofmanova, M.: Quasilinear generalized parabolic Anderson model equation. arXiv prerpint arXiv:1610.06726 (2016)
3. Bogachev, V.: *Gaussian Measures*, vol. 62. American Mathematical Society, Providence, 1998
4. Furlan, M., Gubinelli, M.: Paracontrolled quasilinear SPDEs. arXiv preprint arXiv:1610.07886 (2016)
5. Gloria, A., Otto, F.: The corrector in stochastic homogenization: optimal rates, stochastic integrability, and fluctuations. arXiv preprint arXiv:1510.08290 (2015)
6. Gubinelli, M.: Controlling rough paths. *J. Funct. Anal.* **216**(1), 86–140 (2004)
7. Gubinelli, M., Imkeller, P., Perkowski, N.: Paracontrolled distributions and singular PDEs. *Forum Math. Pi* **3**, e6, 75 (2015)
8. Hairer, M.: Rough stochastic PDEs. *Commun. Pure Appl. Math.* **64**(11), 1547–1585 (2011)
9. Hairer, M.: Solving the KPZ equation. *Ann. Math. (2)* **178**(2), 559–664 (2013)
10. Hairer, M.: A theory of regularity structures. *Invent. Math.* **198**(2), 269–504 (2014)
11. Hairer, M., Labbé, C.: Multiplicative stochastic heat equations on the whole space. arXiv preprint arXiv:1504.07162 (2015)

12. Hairer, M., Labbé, C.: A simple construction of the continuum parabolic Anderson model on $\mathbf{R}^2$. *Electron. Commun. Probab.* **20**(43), 11 (2015)
13. Hairer, M., Pardoux, É.: A Wong–Zakai theorem for stochastic PDEs. *J. Math. Soc. Jpn.* **67**(4), 1551–1604 (2015)
14. Krylov, N.V.: *Lectures on Elliptic and Parabolic Equations in Hölder Spaces*, volume 12 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 1996
15. Lyons, T., Qian, Z.: *System Control and Rough Paths*. Oxford Mathematical Monographs. Oxford University Press, Oxford, 2002. Oxford Science Publications.
16. Lyons, T. J.: Differential equations driven by rough signals. *Rev. Mat. Iberoam.* **14**(2), 215–310 (1998)
17. Lyons, T.J., Caruana, M., Lévy, T.: *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
18. Mourrat, J.C., Weber, H., Xu, W.: Construction of Φ^4_3 diagrams for pedestrians. arXiv preprint arXiv:1610.08897 (2016)
19. Nualart, D.: *The Malliavin Calculus and Related Topics*, vol. 1995. Springer, Berlin, 2006

Felix Otto
Max Planck Institute for Mathematics in the Sciences,
04103 Leipzig,
Germany.
e-mail: otto@mis.mpg.de

and

Hendrik Weber
University of Bath,
Bath,
UK.
e-mail: h.weber@bath.ac.uk

(*Received May 10, 2017 / Revised May 10, 2017 / Accepted November 8, 2018*)
Published online November 30, 2018