KPZ reloaded

Massimiliano Gubinelli
CEREMADE & UMR 7534 CNRS
Université Paris–Dauphine
Institut Universitaire de France
and
University of Bonn
Hausdorff Center for Mathematics
& Institute for Applied Mathematics
gubinelli@iam.uni-bonn.de

Nicolas Perkowski
Humboldt–Universität zu Berlin
Institut für Mathematik
perkowsk@math.hu-berlin.de

September 22, 2016

Abstract

We analyze the one-dimensional periodic Kardar–Parisi–Zhang equation in the language of paracontrolled distributions, giving an alternative viewpoint on the seminal results of Hairer.

Apart from deriving a basic existence and uniqueness result for paracontrolled solutions to the KPZ equation we perform a thorough study of some related problems. We rigorously prove the links between KPZ equation, stochastic Burgers equation, and (linear) stochastic heat equation and also the existence of solutions starting from quite irregular initial conditions. We also show that there is a natural approximation scheme for the nonlinearity in the stochastic Burgers equation.

Interpreting the KPZ equation as the value function of an optimal control problem, we give a pathwise proof for the global existence of solutions and thus for the strict positivity of solutions to the stochastic heat equation.

Moreover, we study Sasamoto-Spohn type discretizations of the stochastic Burgers equation and show that their limit solves the continuous Burgers equation possibly with an additional linear transport term. As an application, we give a proof of the invariance of the white noise for the stochastic Burgers equation which does not rely on the Cole–Hopf transform.

Contents

1 Introduction

The Kardar–Parisi–Zhang (KPZ) equation is the stochastic partial differential equation (SPDE)

$$\mathscr{L}h(t,x) = (\mathrm{D}h(t,x))^2 + \xi(t,x), \qquad x \in \mathbb{R}, \quad t \geqslant 0, \tag{1}$$

where $h : \mathbb{R}_+ \times \mathbb{R} \to \mathbb{R}$ is a real valued function on the real line, $\mathscr{L} = \partial_t - \Delta$ denotes the heat operator, $\mathrm{D} = \partial/\partial x$ and $\partial_t = \partial/\partial t$ are the spatial respectively temporal derivatives and ξ is a space-time white noise: the centered Gaussian space-time random distribution with covariance

$$\mathbb{E}[\xi(t,x)\xi(s,y)] = \delta(t-s)\delta(x-y), \qquad t,s \geqslant 0, \quad x,y \in \mathbb{R}.$$

The KPZ equation was introduced by Kardar, Parisi, and Zhang [KPZ86] as an SPDE model describing the large scale fluctuations of a growing interface represented by a *height field* h. Based on non-rigorous perturbative renormalization group arguments they predicted that upon a particular "1-2-3" rescaling and centering, the height field h (or at least its finite dimensional distributions) must converge to a scale invariant random field H_{kpz} (the *KPZ fixed point*) obtained as

$$H_{\mathrm{kpz}}(t,x) = \lim_{\lambda \to \infty} \lambda h(\lambda^3 t, \lambda^2 x) - c(\lambda)t. \tag{2}$$

According to the general renormalization group (RG) understanding of dynamic critical phenomena a large class of similar interface growth mechanisms must feature the same large scale statistical behavior, namely their height fields $\tilde{h}(t,x)$ converge, upon rescaling (and maybe centering and a suitable Galilean transformation to get rid of a uniform spatial drift) to the same KPZ fixed point H_{kpz}.

Proving any kind of rigorous results about the convergence stated in (2) (called sometimes the *strong KPZ universality conjecture*) is a wide open problem for any continuous SPDE modeling interface growth and in particular for the KPZ equation; see however [SS10, ACQ11] for recent breakthroughs. On the other side there has been a tremendous amount of progress in understanding the large scale behavior of certain discrete probabilistic models of growth phenomena belonging to the same universality class, mainly thanks to a special feature of these models called *stochastic integrability*. For further details see the excellent recent surveys [Cor12, Qua14, QS15].

A weaker form of universality comes from looking at the KPZ equation as a *mesoscopic model* of a special class of growth phenomena. Indeed by the same theoretical physics RG picture it is expected that the KPZ equation is itself a universal description of the fluctuations of weakly asymmetric growth models, a prediction that is commonly referred to as the *weak KPZ universality conjecture*. In this case the microscopic model possesses a parameter which drives the system out of equilibrium and controls the speed of growth. It is then possible to keep the nonlinear contributions of the same size as the diffusion and noise effects by tuning the parameter and at the same time rescaling the height field diffusively. It is then expected that the random field so obtained satisfies the KPZ equation (1).

For a long time, the main difficulty in addressing these beautiful problems and obtaining any kind of invariance principle for space–time random fields describing random growth has been the elusiveness of the KPZ equation itself. The main difficulty in making sense of the equation stems from the fact that for any fixed time $t \geqslant 0$ the dependence of the solution $h(t, x)$ on the space variable $x \in \mathbb{T}$ cannot be better than Brownian. That is, if we measure spatial regularity in the scale of Hölder spaces $\mathscr{C}^\gamma$, we expect that $h(t, \cdot) \in \mathscr{C}^\gamma$ for any $\gamma < 1/2$ but not better, in particular certainly the quadratic term $(\mathrm{D}h(t, x))^2$ is not well defined.

The first rigorous results about KPZ is due to Bertini and Giacomin [BG97], who show that the height function of the weakly asymmetric simple exclusion process (WASEP) converges under appropriate rescaling to a continuous random field h which they characterize using the Cole–Hopf transform. Namely they show that the random field $w = e^h$ is the solution of a particular Itô SPDE: the stochastic heat equation (SHE)

$$\mathrm{d}w(t, x) = \Delta w(t, x)\mathrm{d}t + w(t, x)\mathrm{d}B_t(x), \tag{3}$$

where $B_t(x) = \int_0^t \xi(s, x)\mathrm{d}s$ is a cylindrical Brownian motion in $L^2(\mathbb{R})$. The regularity of the solution w does not allow to determine the intrinsic dynamics of h, but the convergence result shows that any candidate notion of solution to the KPZ equation should have the property that e^h solves the SHE (3).

The key tool in Bertini and Giacomin's proof is Gärtner's microscopic Cole–Hopf transform [Gär88] which allows them to prove the convergence of a transformed version of the WASEP to the stochastic heat equation. But the microscopic Cole–Hopf transform only works for this specific model (or at least only for a few special models, see the recent works [DT13, CT15] for extensions of the Bertini-Giacomin result) and cannot be used to prove universality.

Another notion of solution for the KPZ equation has been introduced by Gonçalves and Jara [GJ14] who proved tightness for a quite general class of rescaled weakly asymmetric exclusion processes and provided a description of the dynamics of the limiting random fields by showing that they must be *energy solutions* of the KPZ equation. In particular, they showed that (after a subtraction) the quadratic term in (1) would make sense only as a space-time distribution and not better. This notion of *energy solution* has been subsequently interpreted by Gubinelli and Jara in the language of controlled paths in [GJ13]. The key observation is that the fast time decorrelation of the noise provides a mechanism to make sense of the quadratic term. Unfortunately, energy solutions are very weak and allow only for few a priori estimates and currently there is no hint that they are unique[1].

A related problem is studied in the the recent work of Alberts, Khanin and Quastel [AKQ14] where a universality result is rigorously shown for discrete random directed polymers with weak noise, which converge under rescaling to the continuum random directed polymer whose partition function is given by the stochastic heat equation.

[1]Shortly after submitting the present article we were actually able to prove the uniqueness of stationary energy solutions, see [GP15b].

This was the state of the art until 2011, when Hairer [Hai13] established a well-posedness result for the periodic KPZ equation using three key ingredients: a partial series expansion of the solution, an explicit control of various stochastic terms, and a fixed point argument involving rough path theory. As a final outcome he was able to show very explicitly that solutions h to equation (1) are limits for $\varepsilon \to 0$ of the approximate solutions h_ε of the equation

$$\partial_t h_\varepsilon(t,x) = \Delta h_\varepsilon(t,x) + (\mathrm{D}h_\varepsilon(t,x))^2 - C_\varepsilon + \xi_\varepsilon(t,x), \qquad x \in \mathbb{T}, \quad t \geqslant 0, \tag{4}$$

where $\mathbb{T} = \mathbb{R}/(2\pi\mathbb{Z})$ and Δ is the Laplacian with periodic boundary conditions on $\mathbb{T}$, ξ_ε is a Gaussian process obtained by regularizing ξ by convoluting it with a smooth kernel ρ, that is $\xi_\varepsilon = \rho_\varepsilon * \xi$ with $\rho_\varepsilon(x) = \varepsilon^{-1}\rho(x/\varepsilon)$, and where the constant C_ε has to be chosen in such a way that it diverges to infinity as $\varepsilon \to 0$. While it had been known since Bertini and Giacomin's work [BG97] that the solutions to (4) converge to a unique limit, the key point of Hairer's construction is that using rough path theory he is able to give an *intrinsic* meaning to the product $(\mathrm{D}h(t,x))^2$ and to obtain good bounds for that operation on suitable function spaces, which ultimately allowed him to solve the equation.

The solution of the KPZ equation was one of the stepping stones in the development of Hairer's regularity structures [Hai14], and now it is of course also possible to solve the equation using regularity structures rather than rough paths; see for example [FH14].

Using regularity structures, Hairer and Quastel [HQ15] recently proved the universality of the KPZ equation in the following sense: they consider a regularized version ξ_ε of the white noise and an arbitrary even polynomial P and show that under the right rescaling the solution h_ε to

$$\partial_t h_\varepsilon = \Delta h_\varepsilon + \sqrt{\varepsilon} P(\mathrm{D}h_\varepsilon) + \xi_\varepsilon \tag{5}$$

converges to the KPZ equation with (possibly) a non–universal constant in front of the non–linearity. Let us also mention the work [HS15], where a similar problem is studied except that $P(x) = x^2$ but ξ_ε is not necessarily Gaussian.

In the same period Hairer was developing his theory, together with P. Imkeller we proposed the notion of *paracontrolled distributions* [GIP15, GIP16, GP15a] as a tool to give a meaning to and solve a large class of singular SPDEs. The class of models which can be analyzed using this method includes the generalized Parabolic Anderson model in 2 dimensions and the stochastic quantization equation in 3 dimensions and other interesting singular SPDEs [CC13, CF14, ZZ15a, ZZ14, BB16, CC15, BBF15, PT16].

After Hairer's breakthrough, another set of tools which should allow to solve the KPZ equation was developed by Kupiainen [Kup14], who builds on renormalization group techniques to make sense of the three dimensional stochastic quantization equation ϕ^4_3.

The main aim of the present paper is to describe the paracontrolled approach to the KPZ equation. While the analysis of the KPZ equation itself is quite simple and shorter versions of the present paper circulated informally among people interested in paracontrolled equations, we kept delaying a public version in order to use the full versatility of paracontrolled analysis to explore various questions around KPZ and related equations and more importantly to find simple and direct arguments for most of the proofs.

Indeed, despite paracontrolled calculus being currently less powerful than the full fledged regularity structure theory, we believe that it is lightweight enough to be effective in exploring various questions related to the qualitative behaviour of solutions and in particular in bridging the gap between stochastic and deterministic PDE theories.

We start in Section 2 by recalling some basic results from paracontrolled distributions, which we then apply in Section 3 to solve the conservative Burgers equation

$$\mathscr{L}u = \mathrm{D}u^2 + \mathrm{D}\xi$$

driven by an irregular signal ξ in a pathwise manner. Here our main result is Theorem 3.5 which gives the local-in-time existence and uniqueness of paracontrolled solutions. In Section 4, Theorems 4.2 and 4.5 we indicate how to adapt the arguments to also solve the KPZ equation and the rough heat equation pathwise, and we show that the formal links between these equations ($u = \mathrm{D}h$, $w = e^h$) can indeed be justified rigorously, also for a driving signal that only has the regularity of the white noise and not just for the equations driven by mollified signals.

We then show in Section 5, specifically in Corollary 5.2, that by working in a space of paracontrolled distributions one can find a natural interpretation for the nonlinearity in the stochastic Burgers equation, an observation which was not yet made in the works based on rough paths or regularity structures, where the focus is more on the continuous dependence of the solution map on the data rather than on the declaration of the nonlinear operations appearing in the equation.

Section 6 extends our previous results to more irregular initial conditions. In the previous sections we only worked with initial conditions for the KPZ equation that are relatively smooth perturbations of the Brownian bridge. But in Theorem 6.13 we show that by allowing for a singularity at 0 we can start the KPZ equation in any $\mathscr{C}^\beta$ function if $\beta > 0$. For the linear heat equation we show in Theorem 6.15 that we can take $w_0 \in B^{-\gamma}_{p,\infty}$ whenever $\gamma < 1/2$ and $p \in [1, \infty]$; in particular we can take w_0 as the Dirac delta because $\delta \in B^0_{1,\infty}$, and even initial conditions with worse regularity are possible.

In Section 7 we develop yet another approach to the KPZ equation. We show in Theorem 7.3 that its solution is given by the value function of the stochastic control problem

$$h(t, x) = \sup_v \mathbb{E}_x \left[\bar{h}(\gamma^v_t) + \int_0^t \left(\xi(t - s, \gamma^v_s) - \infty - \frac{1}{4}|v_s|^2 \right) \mathrm{d}s \right], \tag{6}$$

where under $\mathbb{E}_x$ we have

$$\gamma^v_s = x + \int_0^s v_r \mathrm{d}r + \sqrt{2} W_s$$

with a standard Brownian motion W. Such a representation has already proved very powerful in the case of a spatially smooth noise ξ, see for example [EKMS00, BCK14], but of course it is not obvious how to even make sense of it when ξ is a space-time white noise. Based on paracontrolled distributions and the techniques of [DD16, CC15] we can give a rigorous interpretation for the optimization problem and show that the identity (6) is indeed true. Immediate consequences are a comparison principle for the KPZ equation and a pathwise global existence result. Recall that in [Hai13] global existence could only be shown by identifying the rough path solution with the (stochastic) Cole–Hopf solution, which means that the null set where global existence fails may depend on the initial condition. In Corollary 7.5 we show that this is not the case, and that for any ω for which $\xi(\omega)$ can be enhanced to a *KPZ-enhancement* (see Definition 4.1) and for any initial condition $h_0 \in \mathscr{C}^\beta$ with $\beta > 0$ there exists a unique global in time paracontrolled solution $h(\omega)$ to the equation

$$\mathscr{L} h(t,x,\omega) = (\mathrm{D} h(t,x,\omega))^2 - \infty + \xi(t,x,\omega), \qquad h(0,\omega) = h_0.$$

and that the L^∞ norm of h is controlled only in terms of the KPZ-enhancement of the noise and the L^∞ norm of the initial condition.

A surprising byproduct of these estimates is a positivity result for the solution of the SHE which is independent of the precise details of the noise. This is at odds with currently available proofs of positivity [Mue91, MF14, CK14], starting from the original proof of Mueller [Mue91], which all use heavily the Gaussian nature of the noise and proceed via rather indirect arguments. Ours is a direct PDE argument showing that the ultimate reason for positivity does not lie in the law of the noise but somehow in its space-time regularity (more precisely in the regularity of its enhancement).

Section 8 is devoted to the study of Sasamoto-Spohn [SS09] type discretizations of the conservative stochastic Burgers equation. We consider a lattice model $u_N\colon [0,\infty)\times\mathbb{T}_N \to \mathbb{R}$ (where $\mathbb{T}_N = (2\pi\mathbb{Z}/N)/(2\pi\mathbb{Z})$), defined by an SDE

$$\begin{aligned} \mathrm{d}u_N(t,x) &= \Delta_N u_N(t,x)\mathrm{d}t + (\mathrm{D}_N B_N(u_N(t),u_N(t)))(x)\mathrm{d}t + \mathrm{d}(\mathrm{D}_N\sqrt{\frac{2\pi}{N}}W_N(t,x)) \\ u_N(0,x) &= u_0^N(x), \end{aligned}$$

where Δ_N, D_N are approximations of Laplacian and spatial derivative, respectively, B_N is a bilinear form approximating the pointwise product, $(W_N(t,x))_{t\in\mathbb{R}_+,x\in\mathbb{T}_N}$ is an N–dimensional standard Brownian motion, and u_0^N is independent of W_N. We show in Theorem 8.2 that if (u_0^N) converges in distribution in $\mathscr{C}^{-\beta}$ for some $\beta < 1$ to u_0, then (u_N) converges weakly to the paracontrolled solution u of

$$\mathscr{L} u = \mathrm{D}u^2 + c\mathrm{D}u + \mathrm{D}\xi, \qquad u(0) = u_0, \tag{7}$$

where $c \in \mathbb{R}$ is a constant depending on the specific choice of Δ_N, D_N, and B_N. If Δ_N and D_N are the usual discrete Laplacian and discrete gradient and B_N is the pointwise product, then $c = 1/2$. However, if we replace B_N by the bilinear form introduced in [KS91], then $c = 0$. It was observed before in the works of Hairer, Maas and Weber [HM12, HMW14] that such a "spatial Stratonovich corrector" can arise in the limit of discretizations of Hairer's [Hai11] generalized Burgers equation $\mathscr{L} u = g(u)\partial_x u + \xi$, and indeed one of the motivations for the work [HMW14] was that it would provide a first step towards the understanding of discretizations of the KPZ equation. To carry out this program we simplify many of the arguments in [HMW14], replacing rough paths by paracontrolled distributions. A key difference to [HMW14] is that for us the "paracontrolled derivative" is not constant, which introduces tremendous technical difficulties that were absent in all discretizations of singular PDEs which were studied so far, for example [ZZ14, MW14][2]. We overcome these difficulties by introducing a certain random operator which we bound using stochastic computations.

As an application of our convergence result, we show that the distribution of $m + 2^{-1/2}\eta$, where η is a space white noise and $m \in \mathbb{R}$, is invariant under the evolution of the conservative stochastic Burgers equation. While this is well known (see [BG97] and also the recent work [FQ14] for an elegant proof in the more complicated setting of the non-periodic KPZ equation), ours seems to be the first proof which does not rely on the Cole–Hopf transform.

In Section 9, Theorems 9.1 and 9.3 we construct the enhanced white noise which we needed in our pathwise analysis. We try to build a link with the Feynman diagrams from quantum field theory. The required bounds are shown by reducing the computation to a few basic integrals that can be controlled by a simple recursive algorithm. In Section 10 (Theorem 10.1) we indicate how to adapt these calculations to also obtain the convergence result for the enhanced data in the lattice models. We also calculate the explicit form of the correction constant c appearing in (7).

Acknowledgments. We are very grateful to Khalil Chouk, who pointed out that by working in more general Besov spaces than $B_{\infty,\infty}$ one can start the rough heat equation in the Dirac delta. We would also like to thank Rongchan and Xiangchan Zhu, who found a mistake in a previous version of the paper.

The main part of the work was carried out while N.P. was employed by Université Paris Dauphine and supported by the Fondation Sciences Mathématiques de Paris (FSMP) and by a public grant overseen by the French National Research Agency (ANR) as part of the "Investissements d'Avenir" program (reference: ANR-10-LABX-0098).

2 Paracontrolled calculus

Paracontrolled calculus and the relevant estimates which will be needed in our study of the KPZ (and related) equations have been introduced in [GIP15]. In this section we will recall the notations and the basic results of paracontrolled calculus without proofs. For more details on Besov spaces, Littlewood–Paley theory, and Bony's paraproduct the reader can refer to the nice recent monograph [BCD11].

2.1 Notation and conventions

Throughout the paper, we use the notation $a \lesssim b$ if there exists a constant $c > 0$, independent of the variables under consideration, such that $a \leqslant c \cdot b$, and we write $a \simeq b$ if $a \lesssim b$ and $b \lesssim a$. If we want to emphasize the dependence of c on the variable x, then we write $a(x) \lesssim_x b(x)$. For index variables i and j of Littlewood-Paley decompositions

[2] Now also [HM15] are able to deal with problems where the paracontrolled derivative is not constant.

(see below) we write $i \lesssim j$ if there exists $N \in \mathbb{N}$, independent of j, such that $i \leqslant j + N$ (or in other words if $2^i \lesssim 2^j$), and we write $i \sim j$ if $i \lesssim j$ and $j \lesssim i$.

An *annulus* is a set of the form $\mathscr{A} = \{x \in \mathbb{R}^d : a \leqslant |x| \leqslant b\}$ for some $0 < a < b$. A *ball* is a set of the form $\mathscr{B} = \{x \in \mathbb{R}^d : |x| \leqslant b\}$. $\mathbb{T} = \mathbb{R}/(2\pi\mathbb{Z})$ denotes the torus.

If f is a map from $A \subset \mathbb{R}$ to the linear space Y, then we write $f_{s,t} = f(t) - f(s)$. For $f \in L^p(\mathbb{T})$ we write $\|f(x)\|^p_{L^p_x(\mathbb{T})} := \fint_{\mathbb{T}} |f(x)|^p \mathrm{d}x$.

Given a Banach space X with norm $\|\cdot\|_X$ and $T > 0$, we write $C_T X = C([0,T], X)$ for the space of continuous maps from $[0,T]$ to X, equipped with the supremum norm $\|\cdot\|_{C_T X}$, and we set $CX = C(\mathbb{R}_+, X)$, equipped with the topology of uniform convergence on compacts. Similarly $C(\mathbb{R}, X)$ will always be equipped with the locally uniform topology. For $\alpha \in (0,1)$ we also define $C^\alpha_T X$ as the space of α-Hölder continuous functions from $[0,T]$ to X, endowed with the seminorm $\|f\|_{C^\alpha_T X} = \sup_{0 \leqslant s < t \leqslant T} \|f_{s,t}\|_X / |t-s|^\alpha$, and we write $C^\alpha_{\mathrm{loc}} X$ for the space of locally α-Hölder continuous functions from $\mathbb{R}_+$ to X.

The space of distributions on the torus is denoted by $\mathscr{D}'(\mathbb{T})$ or $\mathscr{D}'$. The Fourier transform is defined with the normalization

$$\mathscr{F} u(k) = \hat{u}(k) = \int_{\mathbb{T}} e^{-ikx} u(x) \mathrm{d}x, \qquad k \in \mathbb{Z},$$

so that the inverse Fourier transform is given by $\mathscr{F}^{-1} v(x) = (2\pi)^{-1} \sum_k e^{ikx} v(k)$. We denote Fourier multipliers by $\varphi(\mathrm{D}) u = \mathscr{F}^{-1}(\varphi \mathscr{F} u)$ whenever the right hand side is well defined.

Throughout the paper, (χ, ρ) will denote a dyadic partition of unity such that $\mathrm{supp}(\rho(2^{-i}\cdot)) \cap \mathrm{supp}(\rho(2^{-j}\cdot)) = \emptyset$ for $|i-j| > 1$. The family of operators $(\Delta_j)_{j \geqslant -1}$ will denote the Littlewood-Paley projections associated to this partition of unity, that is $\Delta_{-1} u = \mathscr{F}^{-1}(\chi \mathscr{F} u)$ and $\Delta_j = \mathscr{F}^{-1}\big(\rho(2^{-j}\cdot)\mathscr{F} u\big)$ for $j \geqslant 0$. We also use the notation $S_j = \sum_{i<j} \Delta_i$. We write $\rho_j = \rho(2^{-j}\cdot)$ for $j \geqslant 0$ and $\rho_{-1} = \chi$, and $\chi_j = \chi(2^{-j}\cdot)$ for $j \geqslant 0$ and $\chi_j = 0$ for $j < 0$. We define $\psi_{\prec}(k,\ell) = \sum_j \chi_{j-1}(k)\rho_j(\ell)$ and $\psi_{\circ}(k,\ell) = \sum_{|i-j| \leqslant 1} \rho_i(k)\rho_j(\ell)$. The Hölder-Besov space $B^\alpha_{\infty,\infty}(\mathbb{T},\mathbb{R})$ for $\alpha \in \mathbb{R}$ will be denoted by $\mathscr{C}^\alpha$ and equipped with the norm

$$\|f\|_\alpha = \|f\|_{B^\alpha_{\infty,\infty}} = \sup_{i \geqslant -1} (2^{i\alpha} \|\Delta_i f\|_{L^\infty(\mathbb{T})}).$$

If f is in $\mathscr{C}^{\alpha-\varepsilon}$ for all $\varepsilon > 0$, then we write $f \in \mathscr{C}^{\alpha-}$. For $\alpha \in (0,2)$, we also define the space $\mathscr{L}^\alpha_T = C^{\alpha/2}_T L^\infty \cap C_T \mathscr{C}^\alpha$, equipped with the norm

$$\|f\|_{\mathscr{L}^\alpha_T} = \max\left\{ \|f\|_{C^{\alpha/2}_T L^\infty}, \|f\|_{C_T \mathscr{C}^\alpha} \right\}.$$

The notation is chosen to be reminiscent of $\mathscr{L} = \partial_t - \Delta$, by which we will always denote the heat operator with periodic boundary conditions on $\mathbb{T}$. We also write $\mathscr{L}^\alpha = C^{\alpha/2}_{\mathrm{loc}} L^\infty \cap C \mathscr{C}^\alpha$.

2.2 Bony–Meyer paraproducts

Paraproducts are bilinear operations introduced by J. M. Bony [Bon81, Mey81] in order to linearize a class of nonlinear PDE problems. Paraproducts allow to describe functions which "look like" some given reference functions and to perform detailed computations on their singular behavior. They also appear naturally in the analysis of the product of two Besov distributions. In terms of Littlewood–Paley blocks, the product fg of two distributions f and g can be (at least formally) decomposed as

$$fg = \sum_{j \geqslant -1} \sum_{i \geqslant -1} \Delta_i f \Delta_j g = f \prec g + f \succ g + f \circ g.$$

Here $f \prec g$ is the part of the double sum with $i < j-1$, and $f \succ g$ is the part with $i > j+1$, and $f \circ g$ is the "diagonal" part, where $|i-j| \leqslant 1$. More precisely, we define

$$f \prec g = g \succ f = \sum_{j \geqslant -1} \sum_{i=-1}^{j-2} \Delta_i f \Delta_j g \qquad \text{and} \qquad f \circ g = \sum_{|i-j| \leqslant 1} \Delta_i f \Delta_j g.$$

This decomposition behaves nicely with respect to Littlewood–Paley theory. Of course, the decomposition depends on the dyadic partition of unity used to define the blocks $\{\Delta_j\}_{j \geqslant -1}$, and also on the particular choice of the set of pairs (i,j) which contribute to the diagonal part. Our choice of taking all (i,j) with $|i-j| \leqslant 1$ into the diagonal part corresponds to a dyadic partition of unity which satisfies $\operatorname{supp}(\rho(2^{-i}\cdot)) \cap \operatorname{supp}(\rho(2^{-j}\cdot)) = \emptyset$ for $|i-j| > 1$. This implies that every term $S_{j-1} f \Delta_j g$ in the series $f \prec g = \sum_j S_{j-1} f \Delta_j g$ has a Fourier transform which is supported in an annulus $2^j \mathscr{A}$, and of course the same holds true for $f \succ g$. On the other side, the terms in the diagonal part $f \circ g$ have Fourier transforms which are supported in balls. We call $f \prec g$ and $f \succ g$ *paraproducts*, and $f \circ g$ the *resonant* term.

Bony's crucial observation is that $f \prec g$ (and thus $f \succ g$) is always a well-defined distribution. In particular, if $\alpha > 0$ and $\beta \in \mathbb{R}$, then $(f,g) \mapsto f \prec g$ is a bounded bilinear operator from $\mathscr{C}^\alpha \times \mathscr{C}^\beta$ to $\mathscr{C}^\beta$. Heuristically, $f \prec g$ behaves at large frequencies like g (and thus retains the same regularity), and f provides only a modulation of g at larger scales. The only difficulty in defining fg for arbitrary distributions lies in handling the resonant term $f \circ g$. The basic result about these bilinear operations is given by the following estimates.

Lemma 2.1 (Paraproduct estimates)**.** *For any $\beta \in \mathbb{R}$ we have*

$$\|f \prec g\|_\beta \lesssim_\beta \|f\|_{L^\infty} \|g\|_\beta, \tag{8}$$

and for $\alpha < 0$ furthermore

$$\|f \prec g\|_{\alpha+\beta} \lesssim_{\alpha,\beta} \|f\|_\alpha \|g\|_\beta. \tag{9}$$

For $\alpha + \beta > 0$ we have

$$\|f \circ g\|_{\alpha+\beta} \lesssim_{\alpha,\beta} \|f\|_\alpha \|g\|_\beta. \tag{10}$$

A natural corollary is that the product fg of two elements $f \in \mathscr{C}^\alpha$ and $g \in \mathscr{C}^\beta$ is well defined as soon as $\alpha + \beta > 0$, and that it belongs to $\mathscr{C}^\gamma$, where $\gamma = \min\{\alpha, \beta, \alpha+\beta\}$.

2.3 Product of paracontrolled distributions and other useful results

The key result of paracontrolled calculus is the fact that we are able to multiply paracontrolled distributions, extending Bony's results beyond the case $\alpha+\beta > 0$. We present here a simplified version which is adapted to our needs. Let us start with the following meta-definition:

Definition 2.2. Let $\beta > 0$ and $\alpha \in \mathbb{R}$. A distribution $f \in \mathscr{C}^\alpha$ is called paracontrolled by $u \in \mathscr{C}^\alpha$ if there exists $f' \in \mathscr{C}^\beta$ such that $f^\sharp = f - f' \prec u \in \mathscr{C}^{\alpha+\beta}$.

Of course in general the derivative f' is not uniquely determined by f and u, so more correctly we should say that (f, f') is paracontrolled by u.

Theorem 2.3. *Let $\alpha, \beta \in (1/3, 1/2)$. Let $u \in \mathscr{C}^\alpha$, $v \in \mathscr{C}^{\alpha-1}$, and let (f, f') be paracontrolled by u and (g, g') be paracontrolled by v. Assume that $u \circ v \in \mathscr{C}^{2\alpha-1}$ is given as limit of $(u_n \circ v_n)$ in $\mathscr{C}^{2\alpha-1}$, where (u_n) and (v_n) are sequences of smooth functions that converge to u in $\mathscr{C}^\alpha$ and to v in $\mathscr{C}^{\alpha-1}$ respectively. Then fg is well defined and satisfies*

$$\|fg - f \prec g\|_{2\alpha-1} \lesssim (\|f'\|_\beta \|u\|_\alpha + \|f^\sharp\|_{\alpha+\beta})(\|g'\|_\beta \|v\|_{\alpha-1} + \|g^\sharp\|_{\alpha+\beta-1}) + \|f'g'\|_\beta \|u \circ v\|_{2\alpha-1}.$$

Furthermore, the product is locally Lipschitz continuous: Let $\tilde{u} \in \mathscr{C}^\alpha$, $\tilde{v} \in \mathscr{C}^{\alpha-1}$ with $\tilde{u} \circ \tilde{v} \in \mathscr{C}^{2\alpha-1}$ and let $(\tilde{f}, \tilde{f}')$ be paracontrolled by $\tilde{u}$ and $(\tilde{g}, \tilde{g}')$ be paracontrolled by $\tilde{v}$. Assume that $M > 0$ is an upper bound for the norms of all distributions under consideration. Then

$$\begin{aligned} &\|(fg - f \prec g) - (\tilde{f}\tilde{g} - \tilde{f} \prec \tilde{g})\|_{2\alpha-1} \\ &\quad \lesssim (1 + M^3)\Big[\|f' - \tilde{f}'\|_\beta + \|g' - \tilde{g}'\|_\beta + \|u - \tilde{u}\|_\alpha + \|v - \tilde{v}\|_{\alpha-1} \\ &\qquad + \|f^\sharp - \tilde{f}^\sharp\|_{\alpha+\beta} + \|g^\sharp - \tilde{g}^\sharp\|_{\alpha+\beta-1} + \|u \circ v - \tilde{u} \circ \tilde{v}\|_{2\alpha-1}\Big]. \end{aligned}$$

If $f' = \tilde{f}' = 1$ or $g' = \tilde{g}' = 1$, then M^3 can be replaced by M^2.

Remark 2.4. *The proof is based on the simple result of* [GIP15]*, Lemma 2.4, that the commutator*

$$C(f, g, h) := (f \prec g) \circ h - f(g \circ h)$$

is a bounded trilinear operator from $\mathscr{C}^\beta \times \mathscr{C}^\alpha \times \mathscr{C}^{\alpha-1}$ to $\mathscr{C}^{2\alpha+\beta-1}$.

We will write $f \cdot g$ instead of fg whenever we want to stress the fact that we are considering the product of paracontrolled distributions. For some computations we will also need a paralinearization result which allows us to control nonlinear functions of the unknown in terms of a paraproduct.

Lemma 2.5 (Bony–Meyer paralinearization theorem, [GIP15], Lemma 2.6)**.** *Let $\alpha \in (0, 1)$ and let $F \in C^2$. There exists a locally bounded map $R_F : \mathscr{C}^\alpha \to \mathscr{C}^{2\alpha}$ such that*

$$F(f) = F'(f) \prec f + R_F(f) \tag{11}$$

for all $f \in \mathscr{C}^\alpha$. If $F \in C^3$, then R_F is locally Lipschitz continuous.

We will also need the following lemma which allows to replace the paraproduct by a pointwise product in certain situations:

Lemma 2.6 (An associativity result, [Bon81], Theorem 2.3)**.** *Let $\alpha > 0$, $\beta \in \mathbb{R}$, and let $f, g \in \mathscr{C}^\alpha$, and $h \in \mathscr{C}^\beta$. Then*

$$\|f \prec (g \prec h) - (fg) \prec h\|_{\alpha+\beta} \lesssim \|f\|_\alpha \|g\|_\alpha \|h\|_\beta.$$

When dealing with paraproducts in the context of parabolic equations it would be natural to introduce parabolic Besov spaces and related paraproducts. But to keep a simpler setting, we choose to work with space–time distributions belonging to the scale of spaces $(C\mathscr{C}^\alpha)_{\alpha\in\mathbb{R}}$. To do so efficiently, we will use a modified paraproduct which introduces some smoothing in the time variable that is tuned to the parabolic scaling. Let therefore $\varphi \in C^\infty(\mathbb{R}, \mathbb{R}_+)$ be nonnegative with compact support contained in $\mathbb{R}_+$ and with total mass 1, and define for all $i \geqslant -1$ the operator

$$Q_i : C\mathscr{C}^\beta \to C\mathscr{C}^\beta, \qquad Q_i f(t) = \int_{\mathbb{R}} 2^{-2i} \varphi(2^{2i}(t-s)) f(s \vee 0) \mathrm{d}s.$$

We will often apply Q_i and other operators on $C\mathscr{C}^\beta$ to functions $f \in C_T\mathscr{C}^\beta$ which we then simply extend from $[0, T]$ to $\mathbb{R}_+$ by considering $f(\cdot \wedge T)$. With the help of Q_i, we define a modified paraproduct

$$f \prec\!\!\prec g = \sum_i (Q_i S_{i-1} f) \Delta_i g$$

for $f, g \in C(\mathbb{R}_+, \mathscr{D}'(\mathbb{T}))$. It is easy to see that for $f \prec\!\!\prec g$ we have essentially the same estimates as for the pointwise paraproduct $f \prec g$, only that we have to bound f uniformly in time. More precisely:

Lemma 2.7. *For any $\beta \in \mathbb{R}$ we have*

$$\|f \prec\!\!\prec g(t)\|_\beta \lesssim \|f\|_{C_t L^\infty} \|g(t)\|_\beta, \tag{12}$$

for all $t > 0$, and for $\alpha < 0$ furthermore

$$\|f \prec\!\!\prec g(t)\|_{\alpha+\beta} \lesssim \|f\|_{C_t\mathscr{C}^\alpha} \|g(t)\|_\beta. \tag{13}$$

See Lemma 6.4 below for a more general version and a proof. We will also need the following two commutation results.

Lemma 2.8. *Let $\alpha \in (0, 2)$ and $\beta \in \mathbb{R}$. Then*

$$\|(f \prec\!\!\prec g - f \prec g)(t)\|_{\alpha+\beta} \lesssim \|f\|_{\mathscr{L}_t^\alpha} \|g(t)\|_\beta$$

for all $t \geqslant 0$. If $\alpha \in (0, 1)$, then we also have

$$\|(\mathscr{L}(f \prec\!\!\prec g) - f \prec\!\!\prec (\mathscr{L} g))(t)\|_{\alpha+\beta-2} \lesssim \|f\|_{\mathscr{L}_t^\alpha} \|g(t)\|_\beta.$$

More general versions of these bounds are shown in Lemma 6.5 below. As a consequence, every distribution on $\mathbb{R}_+ \times \mathbb{T}$ which is paracontrolled in terms of the modified paraproduct is also paracontrolled using the original paraproduct, at least if the derivative is in $\mathscr{L}^\alpha$ and not just in $C\mathscr{C}^\alpha$.

Moreover, we introduce the linear operator $I: C(\mathbb{R}_+, \mathscr{D}'(\mathbb{T})) \to C(\mathbb{R}_+, \mathscr{D}'(\mathbb{T}))$

$$If(t) = \int_0^t P_{t-s} f(s) \mathrm{d}s,$$

where $P_t = e^{t\Delta}$, $t \geqslant 0$ is the heat semigroup generated by the periodic Laplacian. Some standard estimates for I are summarized in the following lemma.

Lemma 2.9 (Schauder estimates,[GIP15], Lemmas A.7-A.9)**.** *For $\alpha \in (0,2)$ we have*

$$\|If\|_{\mathscr{L}_T^\alpha} \lesssim \|f\|_{C_T\mathscr{C}^{\alpha-2}} \tag{14}$$

for all $T > 0$, as well as

$$\|s \mapsto P_s u_0\|_{\mathscr{L}_T^\alpha} \lesssim \|u_0\|_\alpha. \tag{15}$$

Combining the commutator estimate above with the Schauder estimates, we are able to control the modified paraproduct in $\mathscr{L}_T^\alpha$ spaces rather than just $C_T\mathscr{C}^\alpha$ spaces:

Lemma 2.10. *Let $\alpha \in (0,2)$, $\delta > 0$ and let $f \in \mathscr{L}_T^\delta$, $g \in C_T\mathscr{C}^\alpha$, and $\mathscr{L} g \in C_T\mathscr{C}^{\alpha-2}$. Then*

$$\|f \prec\!\!\prec g\|_{\mathscr{L}_T^\alpha} \lesssim \|f\|_{\mathscr{L}_T^\delta}\left(\|g\|_{C_T\mathscr{C}^\alpha} + \|\mathscr{L} g\|_{C_T\mathscr{C}^{\alpha-2}}\right).$$

Proof. We apply the Schauder estimates for $\mathscr{L}$ to obtain

$$\begin{aligned}\|f \prec\!\!\prec g\|_{\mathscr{L}_T^\alpha} &\lesssim \|f \prec\!\!\prec g(0)\|_{\mathscr{C}^\alpha} + \|\mathscr{L}(f \prec\!\!\prec g)\|_{C_T\mathscr{C}^{\alpha-2}} \\ &\leqslant \|\mathscr{L}(f \prec\!\!\prec g) - f \prec\!\!\prec \mathscr{L} g\|_{C_T\mathscr{C}^{\alpha+\delta-2}} + \|f \prec\!\!\prec (\mathscr{L} g)\|_{C_T\mathscr{C}^{\alpha-2}} \\ &\lesssim \|f\|_{\mathscr{L}_T^\delta}(\|g\|_{C_T\mathscr{C}^\alpha} + \|\mathscr{L} g\|_{C_T\mathscr{C}^{\alpha-2}}).\end{aligned}$$

□

It is possible to replace $\|f\|_{\mathscr{L}_T^\delta}$ on the right hand side by $\|f\|_{C_T L^\infty}$, see Lemma 2.16 of [Fur14]. But we will only apply Lemma 2.10 in the form stated above, so that we do not need this improvement.

At the end of this section, let us observe that in $(\mathscr{L}_T^\alpha)_\alpha$ spaces we can gain a small scaling factor by passing to a larger space.

Lemma 2.11. *Let $\alpha \in (0,2)$, $T > 0$, and let $f \in \mathscr{L}_T^\alpha$. Then for all $\delta \in (0,\alpha]$ we have*

$$\|f\|_{\mathscr{L}_T^\delta} \lesssim \|f(0)\|_\delta + T^{(\alpha-\delta)/2}\|f\|_{\mathscr{L}_T^\alpha}.$$

Proof. It suffices to observe that by interpolation $\|f\|_{C_T^{(\alpha-\delta)/2}\mathscr{C}^\delta} \lesssim \|f\|_{\mathscr{L}_T^\alpha}$. □

3 Stochastic Burgers equation

3.1 The strategy

Instead of dealing directly with the KPZ equation we find it convenient to consider its derivative $u = \mathrm{D}h$ which solves the Stochastic Burgers equation (SBE)

$$\mathscr{L}u = \mathrm{D}u^2 + \mathrm{D}\xi. \tag{16}$$

In a later section we will show that the two equations are essentially equivalent. The strategy we adopt is to consider a regularized version for this equation, where the noise ξ has been replaced by a noise ξ_ε that is smooth in space. Our aim is then to show that the solution u_ε of the equation

$$\mathscr{L}u_\varepsilon = \mathrm{D}u_\varepsilon^2 + \mathrm{D}\xi_\varepsilon$$

converges in the space of distributions to a limit u as $\varepsilon \to 0$. Moreover, we want to describe a space of distributions where the nonlinear term $\mathrm{D}u^2$ is well defined, and we want to show that u is the unique element in this space which satisfies $\mathscr{L}u = \mathrm{D}u^2 + \mathrm{D}\xi$ and has the right initial condition.

That it is non-trivial to study the limiting behavior of the sequence (u_ε) can be understood from the following considerations. First, the limiting solution u cannot be expected to behave better (in terms of spatial regularity) than the solution X of the linear equation

$$\mathscr{L}X = \mathrm{D}\xi$$

(for example with zero initial condition at time 0). It is well known that, almost surely, $X \in C\mathscr{C}^{-1/2-}$ but not better, at least on the scale of $(C\mathscr{C}^\alpha)_\alpha$ spaces. In that case, the term u^2 featured on the right hand side of (16) is not well defined since the product $(f, g) \mapsto fg$ is a continuous map from $C\mathscr{C}^\alpha \times C\mathscr{C}^\beta$ to $C\mathscr{C}^{\alpha\wedge\beta}$ only if $\alpha + \beta > 0$, a condition which is violated here. Thus, we cannot hope to directly control the limit of $(u_\varepsilon^2)_\varepsilon$ as $\varepsilon \to 0$.

What raises some hope to have a well–defined limit is the observation that if X_ε is the solution of the linear equation with regularized noise ξ_ε, then $(\mathrm{D}X_\varepsilon^2)_\varepsilon$ converges to a well defined space–time distribution $\mathrm{D}X^2$ which belongs to $\mathscr{L}C\mathscr{C}^{0-}$ (the space of all distributions of the form $\mathscr{L}v$ for some $v \in C\mathscr{C}^{0-}$). This suggests that there are stochastic cancellations going into $\mathrm{D}X^2$ due to the correlation structure of the white noise, cancellations which result in a well behaved limit. Taking these cancellations properly into account in the full solution u is the main aim of the strategy we will implement below.

In order to prove the convergence and related results it will be convenient to take a more general point of view on the problem and to study the *solution map* $\Phi_{\mathrm{rbe}}\colon (\theta, v) \mapsto u$, mapping well-behaved (for example smooth, but here it is enough to consider elements of $C(\mathbb{R}_+, C^1)$ and C^1 respectively) functions θ and initial conditions v to the classical solution of the evolution problem

$$\mathscr{L}u = \mathrm{D}u^2 + \mathrm{D}\theta, \qquad u(0) = v, \tag{17}$$

which we will call the Rough Burgers Equation (RBE).

To recover the original problem, our aim is to understand the properties of Φ when θ belongs to spaces of distributions of the same regularity as the space–time white noise, that is for $\theta \in C\mathscr{C}^{-1/2-}$. In order to do so we will introduce an n-tuple of distributions $\mathbb{X}(\theta)$ constructed recursively from the data θ of the problem and living in a complete metric space $\mathcal{X}_{\mathrm{rbe}}$ and show that for every bounded ball $\mathscr{B}$ around 0 in $\mathcal{X}_{\mathrm{rbe}}$ and every bounded ball $\tilde{\mathscr{B}}$ in $\mathscr{C}^{-1+}$ there exist $T > 0$ and a continuous map $\Psi_{\mathrm{rbe}} : \mathscr{B} \times \tilde{\mathscr{B}} \to C_T\mathscr{C}^{-1-}$ satisfying $\Phi_{\mathrm{rbe}}(\theta, v) = \Psi_{\mathrm{rbe}}(\mathbb{X}(\theta), v)|_{[0,T]}$ for all $(\mathbb{X}(\theta), v) \in \mathscr{B} \times \tilde{\mathscr{B}}$. This will allow us to obtain the local in time convergence of $u_\varepsilon = \Phi_{\mathrm{rbe}}(\xi_\varepsilon, \varphi(\varepsilon \mathrm{D})u_0)$ by showing that the sequence $(\mathbb{X}(\xi_\varepsilon))_\varepsilon$ converges as $\varepsilon \to 0$ to a well–defined element of $\mathcal{X}_{\mathrm{rbe}}$; here φ is a compactly supported, infinitely smooth function with $\varphi(0) = 1$ and we recall that $\varphi(\varepsilon\mathrm{D})u = \mathscr{F}^{-1}(\varphi(\varepsilon\cdot)\mathscr{F}u)$. That last step will be accomplished via direct probabilistic estimates with respect to the law of the white noise ξ. The described approach thus allows us to decouple the initial problem into two parts:

1. a functional analytic part related to the study of particular spaces of distributions that are suitable to analyze the structure of solutions to equation (17);
2. a probabilistic part related to the construction of the *RBE-enhancement* $\mathbb{X}(\xi) \in \mathcal{X}_{\mathrm{rbe}}$ associated to the white noise ξ.

3.2 Structure of the solution

In this discussion we consider the case of zero initial condition and we formally expand the solution u to (16) in certain multilinear functionals constructed from the noise, just as in [Hai13]. Let us first define a linear map $J(f)$ for each smooth f on $\mathbb{R}_+ \times \mathbb{T}$ as the solution to $\mathscr{L}J(f) = \mathrm{D}f$ with initial condition $J(f)(0) = 0$. Explicitly, we have

$$J(f)(t) = I(\mathrm{D}f)(t) = \int_0^t P_{t-s}\mathrm{D}(f(s))\mathrm{d}s,$$

where $(P_t)_{t\geqslant 0}$ is the semi–group generated by the periodic Laplacian on $\mathbb{T}$. Using the estimates on I given in Lemma 2.9, we see that J is a bounded linear operator from $C_T\mathscr{C}^\alpha$ to $C_T\mathscr{C}^{\alpha+1}$ for all $\alpha \in \mathbb{R}$ and $T > 0$.

Let us now expand the solution to the rough Burgers equation (17) around the solution $X = J(\theta)$ to the linear equation $\mathscr{L}X = \mathrm{D}\theta$. Setting $u = X + u^{\geqslant 1}$, we have

$$\mathscr{L}u^{\geqslant 1} = \mathrm{D}(u^2) = \mathrm{D}(X^2) + 2\mathrm{D}(Xu^{\geqslant 1}) + \mathrm{D}((u^{\geqslant 1})^2).$$

We can proceed by performing a further change of variables in order to remove the term $\mathrm{D}(X^2)$ from the equation by setting

$$u = X + J(X^2) + u^{\geqslant 2}. \tag{18}$$

Now $u^{\geqslant 2}$ satisfies

$$\begin{aligned}\mathscr{L}u^{\geqslant 2} \;=\;& 2\mathrm{D}(XJ(X^2)) + \mathrm{D}(J(X^2)J(X^2)) \\ &+ 2\mathrm{D}(Xu^{\geqslant 2}) + 2\mathrm{D}(J(X^2)u^{\geqslant 2}) + \mathrm{D}((u^{\geqslant 2})^2).\end{aligned} \tag{19}$$

We can imagine to make a similar change of variables to get rid of the term $2\mathrm{D}(XJ(X^2))$. As we proceed in this inductive expansion, we generate a certain number of explicit terms, obtained via various combinations of the function X and of the bilinear map

$$B(f,g) = J(fg) = I(\mathrm{D}(fg)).$$

Since we will have to deal explicitly with at least some of these terms, it is convenient to represent them with a compact notation involving binary trees. A binary tree $\tau \in \mathcal{T}$ is either the root $\bullet$ or the combination of two smaller binary trees $\tau = (\tau_1\tau_2)$, where the two edges of the root of τ are attached to τ_1 and τ_2 respectively. The natural grading $\bar{d}\colon \mathcal{T} \to \mathbb{N}$ is given by $\bar{d}(\bullet) = 0$ and $\bar{d}((\tau_1\tau_2)) = 1 + \bar{d}(\tau_1) + \bar{d}(\tau_2)$. However, for our purposes it is more convenient to work with the degree $d(\tau)$, which we define as the number of leaves (that is nodes without children) of τ. By induction we easily see that $d = \bar{d} + 1$. Then we define recursively a map $X : \mathcal{T} \to C(\mathbb{R}_+, \mathscr{D}')$ by

$$X^{\bullet} = X, \qquad X^{(\tau_1\tau_2)} = B(X^{\tau_1}, X^{\tau_2}),$$

giving

$$X^{\mathsf{V}} = B(X,X), \quad X^{\treeVdot} = B(X, X^{\mathsf{V}}), \quad X^{\treeVdotdot} = B(X, X^{\treeVdot}), \quad X^{\treeVV} = B(X^{\mathsf{V}}, X^{\mathsf{V}})$$

and so on, where

$$(\bullet\bullet) = \mathsf{V}, \quad (\mathsf{V}\bullet) = \treeVdot, \quad (\bullet\treeVdot) = \treeVdotdot, \quad (\mathsf{V}\mathsf{V}) = \treeVV, \quad \ldots$$

In this notation the expansion (18)-(19) reads

$$u = X + X^{\mathsf{V}} + u^{\geqslant 2}, \tag{20}$$

$$u^{\geqslant 2} = 2X^{\treeVdot} + X^{\treeVV} + 2B(X, u^{\geqslant 2}) + 2B(X^{\mathsf{V}}, u^{\geqslant 2}) + B(u^{\geqslant 2}, u^{\geqslant 2}). \tag{21}$$

Remark 3.1. *We observe that formally the solution u of equation (17) can be expanded as an infinite sum of terms labelled by binary trees:*

$$u = \sum_{\tau \in \mathcal{T}} c(\tau) X^{\tau},$$

where $c(\tau)$ is a combinatorial factor counting the number of planar trees which are isomorphic (as graphs) to τ. For example $c(\bullet) = 1$, $c(\mathsf{V}) = 1$, $c(\treeVdot) = 2$, $c(\treeVdotdot) = 4$, $c(\treeVV) = 1$ and in general $c(\tau) = \sum_{\tau_1,\tau_2 \in \mathcal{T}} \mathbb{1}_{(\tau_1\tau_2)=\tau} c(\tau_1)c(\tau_2)$. Alternatively, truncating the summation at trees of degree at most n and setting

$$u = \sum_{\tau \in \mathcal{T}, d(\tau) < n} c(\tau) X^{\tau} + u^{\geqslant n},$$

we obtain a remainder $u^{\geqslant n}$ that satisfies the equation

$$u^{\geqslant n} = \sum_{\substack{\tau_1,\tau_2 \,:\, d(\tau_1) < n, d(\tau_2) < n \\ d((\tau_1\tau_2)) \geqslant n}} c(\tau_1)c(\tau_2) X^{(\tau_1\tau_2)} + \sum_{\tau : d(\tau) < n} c(\tau) B(X^{\tau}, u^{\geqslant n}) + B(u^{\geqslant n}, u^{\geqslant n}).$$

Our aim is to control the truncated expansion under the natural regularity assumptions for the white noise case. These regularities turn out to be

$$X \in C\mathscr{C}^{-1/2-}, \quad X^{\vee} \in C\mathscr{C}^{0-}, \quad X^{\vee\!\!\vee}, X^{\vee\!\!\vee\!\!\vee} \in C\mathscr{C}^{1/2-}, \quad X^{\vee\vee} \in C\mathscr{C}^{1-}. \tag{22}$$

Moreover, for any $f \in C\mathscr{C}^{\alpha}$ with $\alpha > 1/2$ we have that $B(X, f) \in C\mathscr{C}^{1/2-}$: indeed, the bilinear map B satisfies

$$\|B(f,g)\|_{C_T\mathscr{C}^{\delta}} \lesssim \|f\|_{C_T\mathscr{C}^{\alpha}} \|g\|_{C_T\mathscr{C}^{\beta}} \tag{23}$$

for any $T > 0$, where $\delta = \min(\alpha, \beta, \alpha + \beta) + 1$, but only if $\alpha + \beta > 0$. Equation (21) implies that $u^{\geqslant 2}$ has at least regularity $(1/2)-$. In this case $B(u^{\geqslant 2}, u^{\geqslant 2})$ is well defined and in $C\mathscr{C}^{3/2-}$, but $B(X, u^{\geqslant 2})$ is not well defined because the sum of the regularities of X and of $u^{\geqslant 2}$ just fails to be positive. To make sense of $B(X, u^{\geqslant 2})$, we continue the expansion. Setting $u^{\geqslant 2} = 2X^{\vee\!\!\vee} + u^{\geqslant 3}$, we obtain from (21) that $u^{\geqslant 3}$ solves

$$u^{\geqslant 3} = 4X^{\vee\!\!\vee\!\!\vee} + X^{\vee\vee} + 2B(X, u^{\geqslant 3}) + 2B(X^{\vee}, u^{\geqslant 2}) + B(u^{\geqslant 2}, u^{\geqslant 2}).$$

At this stage, if we assume that $u^{\geqslant 3}, u^{\geqslant 2} \in C\mathscr{C}^{1/2-}$, all the terms but $4X^{\vee\!\!\vee\!\!\vee} + 2B(X, u^{\geqslant 3})$ are of regularity $C\mathscr{C}^{1-}$: indeed

$$B(X^{\vee}, u^{\geqslant 2}) \in C\mathscr{C}^{1-}, \quad B(u^{\geqslant 2}, u^{\geqslant 2}) \in C\mathscr{C}^{3/2-},$$

and we already assumed that $X^{\vee\vee} \in C\mathscr{C}^{1-}$. This observation implies that $u^{\geqslant 3}$ has the form

$$u^{\geqslant 3} = 2B(X, u^{\geqslant 3}) + 4X^{\vee\!\!\vee\!\!\vee} + C\mathscr{C}^{1-},$$

by which we mean $u^{\geqslant 3} - 2B(X, u^{\geqslant 3}) - 4X^{\vee\!\!\vee\!\!\vee} \in C\mathscr{C}^{1-}$. Of course, the problem still lies in obtaining an expression for $B(X, u^{\geqslant 3})$ which is not well defined since our a priori estimates are insufficient to handle it. This problem leads us to finding a suitable paracontrolled structure to overcome the lack of regularity of $u^{\geqslant 3}$.

3.3 Paracontrolled ansatz

Here we discuss formally our approach to solving the rough Burgers equation (17). The rigorous calculations will be performed in the following section.

Inspired by the partial tree series expansion of u, we set up a paracontrolled ansatz of the form

$$u = X + X^{\vee} + 2X^{\vee\!\!\vee} + u^{Q}, \qquad u^{Q} = u' \prec\!\!\prec Q + u^{\sharp}, \tag{24}$$

where the functions u', Q and $u^{\sharp}$ are for the moment arbitrary and later will be chosen so that u solves the RBE (17). The idea is that the term $u' \prec\!\!\prec Q$ takes care of the less regular contributions to u of all the terms of the form $B(X, f)$. In this sense, we assume that $Q \in C\mathscr{C}^{1/2-}$ and that $u' \in C\mathscr{C}^{1/2-} \cap C^{1/4-}L^{\infty}$. The remaining term $u^{\sharp}$ will contain the more regular part of the solution, more precisely we assume that $u^\sharp \in C\mathscr{C}^{1-}$. Any such u is called paracontrolled.

For paracontrolled distributions u, the nonlinear term takes the form

$$\mathrm{D}u^2 = \mathrm{D}(XX+2XX^{\vee}+X^{\vee}X^{\vee}+4XX^{\curlyvee})+2\mathrm{D}(Xu^Q)+2\mathrm{D}(X^{\vee}(u^Q+2X^{\curlyvee}))+\mathrm{D}((u^Q+2X^{\curlyvee})^2).$$

Of all the term involved, the only one which is problematic is $\mathrm{D}(Xu^Q)$, since by our assumptions $2\mathrm{D}(X^{\vee}(u^Q + X^{\curlyvee})) + \mathrm{D}((u^Q + X^{\curlyvee})^2)$ are well defined and

$$\mathrm{D}(XX + 2XX^{\vee} + X^{\vee}X^{\vee} + 4XX^{\curlyvee}) = \mathscr{L}(X^{\vee} + 2X^{\curlyvee} + X^{\vee\vee} + 4X^{\curlyvee\!\curlyvee})$$

are given through an explicit construction and do not depend on u^Q. Using the paracontrolled structure of u, the term $\mathrm{D}(Xu^Q)$ is defined as $\mathrm{D}(X \cdot u^Q)$ provided that $Q \circ X \in C\mathscr{C}^{0-}$ is given. Under this assumption we easily deduce from Theorem 2.3 that

$$\mathrm{D}(Xu^Q) - u^Q \prec \mathrm{D}X \in C\mathscr{C}^{-1-}.$$

In other words, for paracontrolled distributions the nonlinear operation on the right hand side of (17) is well defined as a bounded operator.

Next, we should specify how to choose Q and which form u' will take for the solution u of (17). To do so, we derive an equation for $u^\sharp$:

$$\mathscr{L}u^\sharp = \mathscr{L}(u - X - X^{\vee} - 2X^{\curlyvee} - u' \prec\!\!\prec Q) = 4\mathrm{D}(XX^{\curlyvee}) + 2u^Q \prec \mathrm{D}X - u' \prec\!\!\prec \mathscr{L}Q + C\mathscr{C}^{-1-},$$

where we used the commutator estimate Lemma 2.8 to replace $\mathscr{L}(u' \prec\!\!\prec Q)$ by $u' \prec\!\!\prec \mathscr{L}Q$. If we now assume that $X^{\curlyvee} \circ X \in C\mathscr{C}^{0-}$, then we obtain

$$\mathscr{L}u^\sharp = (2u^Q + 4X^{\curlyvee}) \prec\!\!\prec \mathrm{D}X - u' \prec\!\!\prec \mathscr{L}Q + C\mathscr{C}^{-1-},$$

which shows that we should choose Q so that $\mathscr{L}Q = \mathrm{D}X$ and $u' = 2u^Q + 4X^{\curlyvee}$. In the next section we derive a priori estimates for the equation driven by regular data which allow us to pass to the limit where θ can be replaced by the white noise ξ.

3.4 A priori estimates on small time intervals

We are now ready to turn these formal discussions into rigorous mathematics and to solve Burgers equation on small time intervals. For simplicity, we first treat the case of initial conditions that are smooth perturbations of the white noise. In Section 6 we indicate how to adapt the arguments to treat more general initial conditions.

Throughout this section we fix $\alpha \in (1/3, 1/2)$. First, let us specify which distributions will play the role of enhanced data. We will usually consider regularized versions of the white noise which leave the temporal variable untouched and convolute the spatial variable against a smooth kernel. So in the following definition we are careful to only require a spatially smooth X, but not a smooth θ.

Definition 3.2. (RBE-enhancement) Let

$$\mathcal{X}_{\mathrm{rbe}} \subseteq C\mathscr{C}^{\alpha-1} \times C\mathscr{C}^{2\alpha-1} \times \mathscr{L}^{\alpha} \times \mathscr{L}^{2\alpha} \times \mathscr{L}^{2\alpha} \times C\mathscr{C}^{2\alpha-1}$$

be the closure of the image of the map $\Theta_{\mathrm{rbe}} : \mathscr{L}C(\mathbb{R}, C^{\infty}(\mathbb{T})) \to \mathcal{X}_{\mathrm{rbe}}$ given by

$$\Theta_{\mathrm{rbe}}(\theta) = (X(\theta), X^{\vee}(\theta), X^{\vee\!\!\!\vee}(\theta), X^{\vee\!\!\!\vee\!\!\!\vee}(\theta), X^{\vee\vee}(\theta), Q(\theta) \circ X(\theta)), \tag{25}$$

where

$$\begin{array}{rcl} \mathscr{L}X(\theta) & = & \mathrm{D}\theta, \\ \mathscr{L}X^{\vee}(\theta) & = & \mathrm{D}(X(\theta)^2), \\ \mathscr{L}X^{\vee\!\!\!\vee}(\theta) & = & \mathrm{D}(X(\theta)X^{\vee}(\theta)), \\ \mathscr{L}X^{\vee\!\!\!\vee\!\!\!\vee}(\theta) & = & \mathrm{D}(X^{\vee\!\!\!\vee}(\theta) \circ X(\theta)), \\ \mathscr{L}X^{\vee\vee}(\theta) & = & \mathrm{D}(X^{\vee}(\theta)X^{\vee}(\theta)), \\ \mathscr{L}Q(\theta) & = & \mathrm{D}X(\theta), \end{array} \tag{26}$$

all with zero initial conditions except $X(\theta)$ for which we choose the "stationary" initial condition

$$X(\theta)(0) = \int_{-\infty}^{0} P_{-s}\mathrm{D}\theta(s)\mathrm{d}s.$$

We call $\Theta_{\mathrm{rbe}}(\theta)$ the *RBE-enhancement* of the driving distribution θ. For $T > 0$ we define $\mathcal{X}_{\mathrm{rbe}}(T) = \mathcal{X}_{\mathrm{rbe}}|_{[0,T]}$ and we write $\|\mathbb{X}\|_{\mathcal{X}_{\mathrm{rbe}}(T)}$ for the norm of $\mathbb{X}$ in the Banach space $C_T\mathscr{C}^{\alpha-1} \times C_T\mathscr{C}^{2\alpha-1} \times \mathscr{L}_T^{\alpha} \times \mathscr{L}_T^{2\alpha} \times \mathscr{L}_T^{2\alpha} \times C_T\mathscr{C}^{2\alpha-1}$. Moreover, we set $d_{\mathcal{X}_{\mathrm{rbe}}(T)}(\mathbb{X}, \tilde{\mathbb{X}}) = \|\mathbb{X} - \tilde{\mathbb{X}}\|_{\mathcal{X}_{\mathrm{rbe}}(T)}$.

For every RBE-enhancement we have an associated space of paracontrolled distributions.

Definition 3.3. Let $\mathbb{X} \in \mathcal{X}_{\mathrm{rbe}}$ and $\delta > 1 - 2\alpha$. We define the space $\mathscr{D}^{\delta}_{\mathrm{rbe}} = \mathscr{D}^{\delta}_{\mathrm{rbe},\mathbb{X}}$ of distributions paracontrolled by $\mathbb{X}$ as the set of all $(u, u', u^{\sharp}) \in C\mathscr{C}^{\alpha-1} \times \mathscr{L}^{\delta} \times \mathscr{L}^{\alpha+\delta}$ such that

$$u = X + X^{\vee} + 2X^{\vee\!\!\!\vee} + u' \prec\!\!\prec Q + u^{\sharp}.$$

For $\delta = \alpha$ we will usually write $\mathscr{D}_{\mathrm{rbe}} = \mathscr{D}^{\alpha}_{\mathrm{rbe}}$. For every $T > 0$ we set $\mathscr{D}^{\delta}_{\mathrm{rbe}}(T) = \mathscr{D}^{\delta}_{\mathrm{rbe}}|_{[0,T]}$, and we define

$$\|u\|_{\mathscr{D}^{\delta}_{\mathrm{rbe}}(T)} = \|u'\|_{\mathscr{L}^{\delta}_T} + \|u^{\sharp}\|_{\mathscr{L}^{\alpha+\delta}_T}.$$

We will often use the notation $u^Q = u' \prec\!\!\prec Q + u^{\sharp}$. Since u is a function of u' and $u^{\sharp}$, we will also write $(u', u^{\sharp}) \in \mathscr{D}_{\mathrm{rbe}}$, and if there is no ambiguity then we also abuse notation by writing $u \in \mathscr{D}_{\mathrm{rbe}}$. We call u' the *derivative* and $u^{\sharp}$ the *remainder*.

Strictly speaking it is not necessary to assume $u^{\sharp} \in \mathscr{L}^{\alpha+\delta}$, the weaker condition $u^{\sharp} \in C\mathscr{C}^{\alpha+\delta}$ would suffice to define $\mathrm{D}u^2$. But this stronger assumption will simplify the notation and the uniqueness argument below, and it will always be satisfied by the paracontrolled solution to Burgers equation.

Our a priori estimate below will bound $\|u\|_{\mathscr{D}_{\mathrm{rbe}}(T)}$ in terms of the weaker norm $\|u\|_{\mathscr{D}^{\delta}_{\mathrm{rbe}}(T)}$ for suitable $\delta < \alpha$. This will allow us to obtain a scaling factor T^{ε} on small time intervals $[0,T]$, from where we can then derive the locally Lipschitz continuous dependence of the solution u on the data $(\mathbb{X}, u_0)$. Throughout this section we will work under the following assumption:

Assumption (T,M)**.** *Assume that* $\theta, \tilde{\theta} \in \mathscr{L}C(\mathbb{R}, C^{\infty}(\mathbb{T}))$ *and* $u_0, \tilde{u}_0 \in \mathscr{C}^{2\alpha}$*, and that* u *is the unique global in time solution to the Burgers equation*

$$\mathscr{L}u = \mathrm{D}u^2 + \mathrm{D}\xi, \qquad u(0) = X(0) + u_0. \tag{27}$$

We define $\mathbb{X} = \Theta_{\mathrm{rbe}}(\theta)$*,* $u^{Q} = u - X - X^{\vee} - 2X^{\vee_{\!\vee}}$*, and* $u' = 2u^{Q} + 4X^{\vee_{\!\vee}}$*, and we set* $u^{\sharp} = u^{Q} - u' \prec\!\!\prec Q$*. Similarly we define* $\tilde{\mathbb{X}}, \tilde{u}, \tilde{u}^{Q}, \tilde{u}', \tilde{u}^{\sharp}$*. Finally we assume that* $T, M > 0$ *are such that*

$$\max\{\|u_0\|_{2\alpha}, \|\tilde{u}_0\|_{2\alpha}, \|\mathbb{X}\|_{\mathcal{X}_{\mathrm{rbe}}(T)}, \|\tilde{\mathbb{X}}\|_{\mathcal{X}_{\mathrm{rbe}}(T)}\} \leqslant M.$$

Lemma 3.4. *Make Assumption (T,M) and let* $\delta > 1 - 2\alpha$*. Then*

$$\|\mathscr{L}u^{\sharp} - \mathscr{L}X^{\vee\vee} - 4\mathscr{L}X^{\vee_{\!\vee}\vee}\|_{C_T\mathscr{C}^{2\alpha-2}} + \|u'\|_{\mathscr{L}^{\alpha}_T} \lesssim (1 + M^2)(1 + \|u\|^2_{\mathscr{D}^{\delta}_{\mathrm{rbe}}(T)}). \tag{28}$$

If further also $\|u\|_{\mathscr{D}_{\mathrm{rbe},\mathbb{X}}(T)}, \|\tilde{u}\|_{\mathscr{D}_{\mathrm{rbe},\tilde{\mathbb{X}}}(T)} \leqslant M$*, then*

$$\begin{aligned}\|(\mathscr{L}u^{\sharp} - \mathscr{L}X^{\vee\vee} - 4\mathscr{L}X^{\vee_{\!\vee}\vee}) - (\mathscr{L}\tilde{u}^{\sharp} - \mathscr{L}\tilde{X}^{\vee\vee} - 4\mathscr{L}\tilde{X}^{\vee_{\!\vee}\vee})\|_{C_T\mathscr{C}^{2\alpha-2}} + \|u' - \tilde{u}'\|_{\mathscr{L}^{\alpha}_T} \\ \lesssim M^2(d_{\mathcal{X}_{\mathrm{rbe}}(T)}(\mathbb{X}, \tilde{\mathbb{X}}) + (\|u' - \tilde{u}'\|_{\mathscr{L}^{\delta}_T} + \|u^{\sharp} - \tilde{u}^{\sharp}\|_{\mathscr{L}^{2\delta}_T})).\end{aligned} \tag{29}$$

Proof. Let us decompose $\mathscr{L}u^{\sharp}$ into a sum of bounded multilinear operators:

$$\begin{aligned}\mathscr{L}u^{\sharp} &= \mathrm{D}u^2 - \mathrm{D}X^2 - 2\mathrm{D}X^{\vee}X - \mathscr{L}(u' \prec\!\!\prec Q) \\ &= \mathscr{L}X^{\vee\vee} + 2\mathrm{D}[u^{Q}X - u^{Q} \prec X] + 2[\mathrm{D}(u^{Q} \prec X) - u^{Q} \prec \mathrm{D}X] \\ &\quad + 4\mathscr{L}X^{\vee_{\!\vee}\vee} + 4\mathrm{D}[X^{\vee_{\!\vee}} \succ X] + 4[\mathrm{D}(X^{\vee_{\!\vee}} \prec X) - X^{\vee_{\!\vee}} \prec \mathrm{D}X] + [u' \prec \mathrm{D}X - u' \prec\!\!\prec \mathrm{D}X] \\ &\quad + \mathrm{D}[2X^{\vee}(2X^{\vee_{\!\vee}} + u^{Q}) + (2X^{\vee_{\!\vee}} + u^{Q})^2] - \left[\mathscr{L}(u' \prec\!\!\prec Q) - u' \prec\!\!\prec (\mathscr{L}Q)\right],\end{aligned}$$

where we used that $u' = 2u^{Q} + 4X^{\vee_{\!\vee}}$. From here it is straightforward to show that

$$\begin{aligned}\|\mathscr{L}u^{\sharp} - \mathscr{L}X^{\vee\vee} - 4\mathscr{L}X^{\vee_{\!\vee}\vee}\|_{C_T\mathscr{C}^{2\alpha-2}} &\lesssim \|\mathbb{X}\|_{\mathcal{X}_{\mathrm{rbe}}(T)}(\|u^{Q}\|_{C_T\mathscr{C}^{\alpha}} + \|u'\|_{\mathscr{L}^{\alpha}_T}) + \|u^{Q}\|^2_{C_T L^{\infty}} \\ &\quad + (\|\mathbb{X}\|_{\mathcal{X}_{\mathrm{rbe}}(T)} + \|\mathbb{X}\|^2_{\mathcal{X}_{\mathrm{rbe}}(T)})(1 + \|u^{\sharp}\|_{C_T\mathscr{C}^{\alpha+\delta}} + \|u'\|_{C_T\mathscr{C}^{\delta}}).\end{aligned} \tag{30}$$

Now $\|u^{Q}\|_{C_T\mathscr{C}^{\alpha}} \lesssim \|u'\|_{\mathscr{L}^{\delta}_T}\|\mathbb{X}\|_{\mathcal{X}_{\mathrm{rbe}}(T)} + \|u^{\sharp}\|_{C_T\mathscr{C}^{\alpha}}$ by Lemma 2.10 (observe that $\|Q\|_{\mathscr{L}^{\alpha}_T} \lesssim \|\mathbb{X}\|_{\mathcal{X}_{\mathrm{rbe}}(T)}$ by the Schauder estimates, Lemma 2.9), and therefore the estimate (28) follows as soon as we can bound $(1 + \|\mathbb{X}\|_{\mathcal{X}_{\mathrm{rbe}}(T)})\|u'\|_{\mathscr{L}^{\alpha}_T}$ by the right hand side of (28). For this purpose it suffices to use the explicit form $u' = 2u^{Q} + 4X^{\vee_{\!\vee}}$ and then $u^{Q} = u' \prec\!\!\prec Q + u^{\sharp}$ in combination with Lemma 2.10.

The same arguments combined with the multilinearity of our operators also yield (29). □

The main result of this section now follows:

Theorem 3.5. *For every $(\mathbb{X}, u_0) \in \mathcal{X}_{\mathrm{rbe}} \times \mathscr{C}^{2\alpha}$ there exists $T^* \in (0, \infty]$ such that for all $T < T^*$ there is a unique solution $(u, u', u^\sharp) \in \mathscr{D}_{\mathrm{rbe}}(T)$ to the rough Burgers equation* (27) *on the interval $[0, T]$. Moreover, we can choose*

$$T^* = \sup\{t \geqslant 0 : \|u\|_{\mathscr{D}_{\mathrm{rbe}}(t)} < \infty\}.$$

Proof. Let $M = 2\max\{\|u_0\|_{2\alpha}, \|\mathbb{X}\|_{\mathcal{X}_{\mathrm{rbe}}(1)}\}$. Consider a sequence $(\theta_n, u_0^n) \subset \mathscr{L}C(\mathbb{R}, C^\infty) \times \mathscr{C}^{2\alpha}$ such that $\mathbb{X}_n = \Theta_{\mathrm{rbe}}(\theta_n)$ approximates $\mathbb{X}|_{[0,T]}$ in $\mathcal{X}_{\mathrm{rbe}}(T)$ for all $T > 0$ and (u_0^n) approximates u_0 in $\mathscr{C}^{2\alpha}$. Without loss of generality assume that $\|u_0^n\|_{2\alpha} \vee \|\mathbb{X}_n\|_{\mathcal{X}_{\mathrm{rbe}}(1)} \leqslant M$ for all n. Denote by u_n the solution to Burgers equation driven by $\mathbb{X}_n$ and started in u_0^n and define u_n^Q, u_n', $u_n^\sharp$ as in Assumption (T, M). In particular, for all n we are in the setting of Assumption $(1, M)$.

For fixed n, the remainder has the initial condition $u_n^\sharp(0) = u_0^n \in \mathscr{C}^{2\alpha}$, so that Lemma 3.4 and the Schauder estimates (Lemma 2.9) yield

$$\|u_n\|_{\mathscr{D}_{\mathrm{rbe}}(\tau)} \lesssim (1 + M^2)(1 + \|u_n\|^2_{\mathscr{D}^\delta_{\mathrm{rbe}}(\tau)})$$

for all $\tau \in (0, 1]$. Therefore, we can apply Lemma 2.11 to gain a small scaling factor:

$$\|u_n\|_{\mathscr{D}_{\mathrm{rbe}}(\tau)} \lesssim (1 + M^2)(1 + \tau^{\alpha-\delta}\|u_n\|^2_{\mathscr{D}_{\mathrm{rbe}}(\tau)}). \tag{31}$$

Now $\|u_n\|_{\mathscr{D}_{\mathrm{rbe}}(\tau)} \leqslant \|u_n\|_{\mathscr{D}_{\mathrm{rbe}}(1)}$ for all $\tau \in (0, 1]$, and therefore there exists a universal constant $C > 0$ such that $\|u_n\|_{\mathscr{D}_{\mathrm{rbe}}(\tau)} \leqslant C(1 + M^2)$ for all sufficiently small $\tau \leqslant \tau^*$. However, a priori τ^* may depend on stronger norms of θ_n and not just on $\|\mathbb{X}_n\|_{\mathcal{X}_{\mathrm{rbe}}(1)}$. To see that we may choose τ^* only depending on M, it suffices to note that $\tau \mapsto \|u_n\|_{\mathscr{D}_{\mathrm{rbe}}(\tau)}$ is a continuous function (which follows from the fact that $u_n^\sharp$ and u_n' are actually more regular than $\mathscr{L}^{2\alpha}$ and $\mathscr{L}^\alpha$ respectively), to select the smallest τ^* with $\|u_n\|_{\mathscr{D}_{\mathrm{rbe}}(\tau^*)} = C(1 + M^2)$, and to plug $\|u_n\|_{\mathscr{D}_{\mathrm{rbe}}(\tau^*)}$ into (31).

In other words $\|u_n\|_{\mathscr{D}_{\mathrm{rbe}}(\tau)} \leqslant C(1 + M^2)$ for all n, so that the second statement of Lemma 3.4 in conjunction with the Schauder estimates for the Laplacian shows that

$$\|u_m^\sharp - u_n^\sharp\|_{\mathscr{L}^{2\alpha}_\tau} + \|u_m' - u_n'\|_{\mathscr{L}^\alpha_\tau} \lesssim (1 + M^2)(d_{\mathcal{X}_{\mathrm{rbe}}(\tau)}(\mathbb{X}_m, \mathbb{X}_n) + \|u_0^m - u_0^n\|_{2\alpha})$$

for all m, n (possibly after further decreasing $\tau = \tau(M)$). So if

$$\mathscr{B} = \left\{(\Theta_{\mathrm{rbe}}(\theta), v) : \theta \in \mathscr{L}C(\mathbb{R}, C^\infty(\mathbb{T})), v \in \mathscr{C}^{2\alpha}, \|\Theta_{\mathrm{rbe}}(\theta)\|_{\mathcal{X}_{\mathrm{rbe}}(1)}, \|v\|_{2\alpha} \leqslant M\right\},$$

and if $\Psi : \mathscr{B} \to \mathscr{L}^\alpha_{\tau(M)} \times \mathscr{L}^{2\alpha}_{\tau(M)}$ is the map which assigns to (θ, v) the local solution $(u', u^\sharp)$ of (27), then Ψ has a unique continuous extension from $\mathscr{B}$ to its closure in $\mathcal{X}_{\mathrm{rbe}}(1) \times \mathscr{C}^{2\alpha}$, which is just the ball $\mathscr{B}_M(\mathcal{X}_{\mathrm{rbe}}(1) \times \mathscr{C}^{2\alpha})$ of size M in this space. By construction, $u = (u', u^\sharp) = \Psi(\mathbb{X}, u_0)$ solves Burgers equation driven by $\mathbb{X}$ and with initial condition u_0 if we interpret the product $\mathrm{D}u^2$ using the paracontrolled structure of u. To see that u is the unique element in $\mathscr{D}_{\mathrm{rbe},\mathbb{X}}(\tau(M))$ which solves the equation, it suffices to repeat the proof of estimate (29).

It remains to iterate this argument in order to obtain solutions until the blow up time T^*. Assume that we constructed a paracontrolled solution u on $[0,\tau]$ for some $\tau > 0$ and consider the time interval $[\tau, \tau+1]$. Denote the solution on this interval by $\tilde{u}$, and write $\tilde{\mathbb{X}}(t) = \mathbb{X}(t+\tau)$. The initial condition for $\tilde{u}$ is $\tilde{u}(0) = u(\tau) \in \mathscr{C}^\alpha$, so that we are not in the same setting as before. But in fact we only needed a smooth initial condition to obtain an initial condition of regularity 2α for the remainder, and now we have

$$\begin{aligned}\tilde{u}^\sharp(0) &= \tilde{u}(0) - \tilde{X}(0) - \tilde{X}^{\vee}(0) - 2\tilde{X}^{\vee_{\!\vee}}(0) - \tilde{u}' \prec\!\!\prec \tilde{Q}(0) \\ &= u(\tau) - X(0) - X^{\vee}(\tau) - 2X^{\vee_{\!\vee}}(\tau) - u'(\tau) \prec Q(\tau).\end{aligned}$$

According to Lemma 2.8 we have $u'(\tau) \prec Q(\tau) - u' \prec\!\!\prec Q(\tau) \in \mathscr{C}^{2\alpha}$, so that the paracontrolled structure of u on the time interval $[0,\tau]$ shows that the initial condition for $\tilde{u}^\sharp$ is in $\mathscr{C}^{2\alpha}$. As a consequence, there exists $\tilde{M} > 0$ which is a function of $\|\mathbb{X}\|_{\mathcal{X}_{\mathrm{rbe}}(\tau+1)}$, $\|u\|_{\mathscr{D}_{\mathrm{rbe}}(\tau)}$ such that the unique paracontrolled solution $\tilde{u}$ can be constructed on $[0, \tau(\tilde{M})]$. We then extend u and u' from $[0,\tau]$ to $[0, \tau + \tau(\tilde{M})]$ by setting $u(t) = \tilde{u}(t-\tau)$ for $t \geqslant \tau$ and similarly for u'. Then by definition $u' \in \mathscr{L}^\alpha_{\tau+\tau(\tilde{M})}$. For the remainder we know from our construction that on the interval $[\tau, \tau+\tau(\tilde{M})]$ we have $u^Q - u' \prec\!\!\prec_\tau Q \in \mathscr{L}^{2\alpha}$, where

$$u' \prec\!\!\prec_\tau Q(t) = \sum_j \int_{-\infty}^t 2^{2j} \varphi(2^{2j}(t-s)) S_{j-1} u'(s \vee \tau) \mathrm{d}s \Delta_j Q(t).$$

It therefore suffices to show that $(u' \prec\!\!\prec_\tau Q - u' \prec\!\!\prec Q)|_{[\tau,\tau+\tau(\tilde{M})]} \in \mathscr{L}^{2\alpha}$. But we already showed that $(u' \prec\!\!\prec_\tau Q - u' \prec\!\!\prec Q)(\tau) \in \mathscr{C}^{2\alpha}$, and for $t \in [\tau, \tau+\tau(\tilde{M})]$ we get

$$\begin{aligned}\mathscr{L}(u' \prec\!\!\prec_\tau Q - u' \prec\!\!\prec Q)(t) &= (u' \prec\!\!\prec_\tau \mathscr{L}Q - u' \prec\!\!\prec \mathscr{L}Q)(t) + \mathscr{C}^{2\alpha} \\ &= (u' \prec\!\!\prec_\tau \mathrm{D}X - u' \prec\!\!\prec \mathrm{D}X)(t) + \mathscr{C}^{2\alpha}.\end{aligned}$$

Now Lemma 2.8 shows that $u' \prec\!\!\prec \mathrm{D}X - u' \prec \mathrm{D}X \in C_{\tau+\tau(\tilde{M})}\mathscr{C}^{2\alpha-2}$, and by the same arguments it follows that also $(u' \prec\!\!\prec_\tau \mathrm{D}X - u' \prec \mathrm{D}X \in \mathscr{C}^{2\alpha-2})|_{[\tau,\tau+\tau(\tilde{M})]} \in C([\tau, \tau + \tau(\tilde{M})], \mathscr{C}^{2\alpha})$, so that the $\mathscr{L}^{2\alpha}$ regularity of $u^\sharp$ on $[0, \tau+\tau(\tilde{M})]$ follows from the Schauder estimates for the heat flow, Lemma 2.9.

Reversing these differences between the different paraproducts, we also obtain the uniqueness of u on $[0, \tau + \tilde{\tau}(M)]$. □

Remark 3.6. *There is a subtlety concerning the uniqueness statement, similarly as for controlled rough paths. The uniqueness is to be interpreted in the sense that there exists (locally in time) a unique fixed point* $(u, u', u^\sharp)$ *in* $\mathscr{D}_{\mathrm{rbe}}$ *for the map* $\mathscr{D}_{\mathrm{rbe}} \ni (v, v', v^\sharp) \mapsto (w, w', w^\sharp) \in \mathscr{D}_{\mathrm{rbe}}$*, where*

$$w = P.(u_0) + I(\mathrm{D}v^2) + X, \qquad w' = 2v^Q + 4X^{\vee_{\!\vee}}, \qquad w^\sharp = w - X - X^{\vee} - 2X^{\vee_{\!\vee}} - w' \prec\!\!\prec Q.$$

For general $\mathbb{X} \in \mathcal{X}_{\mathrm{rbe}}$ *it is not true that there is a unique* $u \in C_T\mathscr{C}^{\alpha-1}$ *for which there exist* $u', u^\sharp$ *with* $(u, u', u^\sharp) \in \mathscr{D}_{\mathrm{rbe},\mathbb{X}}(T)$ *and such that* $u = P.(u_0) + I(\mathrm{D}u^2) + X$*, where the* *square is calculated using u' and $u^\sharp$. The problem is that u' and $u^\sharp$ are in general not uniquely determined by u and $\mathbb{X}$ (see however* [HP13, FS13] *for conditions that guarantee uniqueness of the derivative in the case of controlled rough paths). Nonetheless, we will see in Section 5 below that this stronger notion of uniqueness always holds as long as $\mathbb{X}$ is "generated from convolution smoothings", and in particular when $\mathbb{X}$ is constructed from the white noise.*

Moreover, without additional conditions on $\mathbb{X} \in \mathcal{X}_{\mathrm{rbe}}$ there always exists a unique $u \in C_T\mathscr{C}^{\alpha-1}$ so that if (θ_n) is a sequence in $\mathscr{L}C(\mathbb{R}, C^\infty)$ for which $(\Theta_{\mathrm{rbe}}(\theta_n))$ converges to $\mathbb{X}$, then the classical solutions u_n to Burgers equation driven by θ_n converge to u.

The solution constructed in Theorem 3.5 depends continuously on the data $\mathbb{X}$ and on the initial condition u_0. To formulate this, we have to be a little careful because for now we cannot exclude the possibility that the blow up time T^* is finite (although we will see in Section 7 below that in fact $T^* = \infty$ for all $\mathbb{X} \in \mathcal{X}_{\mathrm{rbe}}$, $u_0 \in \mathscr{C}^{2\alpha}$). We define the distance $d_{\mathcal{X}_{\mathrm{rbe}}}$ as follows:

$$d_{\mathcal{X}_{\mathrm{rbe}}}(\mathbb{X}, \tilde{\mathbb{X}}) := \sum_{m=1}^{\infty} 2^{-m}(1 \wedge d_{\mathcal{X}_{\mathrm{rbe}(m)}}(\mathbb{X}, \tilde{\mathbb{X}})). \tag{32}$$

Theorem 3.7. *Let $(\mathbb{X}_n)_n, \mathbb{X} \in \mathcal{X}_{\mathrm{rbe}}$ be such that $\lim_n d_{\mathcal{X}_{\mathrm{rbe}}}(X_n, X) = 0$ and let $(u_0^n), u_0 \in \mathscr{C}^{2\alpha}$ be such that $\lim_n \|u_0^n - u_0\|_{2\alpha} = 0$. Denote the solution to the rough Burgers equation (27) driven by $\mathbb{X}_n$ and started in $X_n(0) + u_0^n$ by u_n, and write u for the solution driven by $\mathbb{X}$ and started in $X(0) + u_0$. Then there exists a sequence of times T_n with $\lim_{n\to\infty} T_n = T^* := \sup\{t \geqslant 0 : \|u\|_{\mathscr{D}_{\mathrm{rbe}}(t)} < \infty\}$ such that*

$$\lim_{n\to\infty} \|u_n - u\|_{C_{\tau_n}\mathscr{C}^\alpha} = 0.$$

We say that the solution depends "continuously" on $(\mathbb{X}, u_0)$.

Proof. Let $m \in \mathbb{N}$ and define $M = \max\{\|\mathbb{X}\|_{\mathcal{X}_{\mathrm{rbe}(m)}}, \|\mathbb{X}_n\|_{\mathcal{X}_{\mathrm{rbe}(m)}}, \|u_0\|_{2\alpha}, \|u_0^n\|_{2\alpha}\}$ as well as

$$\tau_m^n := \inf\{t \geqslant 0 : \|u_n\|_{\mathscr{D}_{\mathrm{rbe}}(t)} \vee \|u\|_{\mathscr{D}_{\mathrm{rbe}}(t)} \geqslant m\} \wedge m.$$

Using the same arguments as in the proof of Theorem 3.5, we see that there exists a constant $C(m, M) > 0$, increasing in m and M, such that

$$\|u_n - u\|_{C_{\tau_m^n}\mathscr{C}^\alpha} \leqslant C(m, M)(d_{\mathcal{X}_{\mathrm{rbe}}}(\mathbb{X}_n, \mathbb{X}) + \|u_0^n - u_0\|_{2\alpha}).$$

Now the larger n is, the smaller the second factor on the right hand side is and thus the larger we may choose m. In other words there exists a sequence $(m(n))_{n\in\mathbb{N}}$ with $\lim_n m(n) = \infty$ and $\lim_n \|u_n - u\|_{C_{\tau^n_{m(n)}}\mathscr{C}^\alpha} = 0$. It remains to show that $\lim_n \tau^n_{m(n)} = T^*$. Clearly the limit is bounded from above by T^*, so assume that it is strictly smaller. Then there exists $T < T^*$ such that for all large n we have $\tau^n_{m(n)} < T$, and by definition of $\tau^n_{m(n)}$ and T^* this means that $\|u_n\|_{\mathscr{D}_{\mathrm{rbe}}(T)} > m(n)$ for all large n. But then we can use that $\lim_n m(n) = \infty$ and apply once more the same arguments as in the proof of Theorem 3.5 to obtain $\|u\|_{\mathscr{D}_{\mathrm{rbe}}(T)} = \infty$, a contradiction to $T < T^*$. □

Remark 3.8. *We say that the solution depends "continuously" on the data because we did not define a topology on the space where the solution lives. There are canonical definitions for the distance on the space of exploding paths* $f: \mathbb{R}_+ \to Z \cup \{\zeta\}$ *where* $(Z, \|\cdot\|)$ *is a normed space and* ζ *is a cemetery state; we can for example set* $\tau_m(z) = \inf\{t \geqslant 0 : \|z(t)\| = m\} \wedge m$ *and* $y^\tau(t) = y(t \wedge \tau)$ *and start with the pseudodistances* $d_m(z, \tilde{z}) = \|z^{\tau^m(z)} - \tilde{z}^{\tau^m(\tilde{z})}\|$ *which can be combined into a metric as in* (32)*). However, for the solution to Burgers equation we are not able to prove convergence in any useful distance on the space of exploding paths which is why we state the convergence result in the same form as in* [Hai14].

Remark 3.9. *The analysis in this section is very similar and of course equivalent to the one in* [Hai13]*, where the solution of the KPZ equation is expanded in a number of multilinear functionals* Y^τ *constructed from the white noise (see also the next section where we study the KPZ equation). The key difference is that in* [Hai13] *the product* $u^Q X$ *is constructed using the controlled rough path integral of* [Gub04] *by first noting that there exists* Y *with* $\mathrm{D}Y = X$ *and then setting* $u^Q X := \mathrm{D} \int_{-\pi}^{\cdot} u^Q(y) \mathrm{d}Y(y)$*. Working directly with the product simplifies some of the arguments.*

Of course it is also possible to solve the stochastic Burgers equation using regularity structures (see [Hai14, FH14]*), where roughly speaking every term* X^τ *is identified with a basis vector for an abstract vector space in which the solution to Burgers equation can be expanded. Regularity structures are very powerful and allow to deal with more general equations than the ones that can be treated using paracontrolled distributions. On the other hand they are built completely from scratch and sometimes need sophisticated arguments to carry over results from classical analysis, while paracontrolled distributions are a relatively lightweight addition to the available PDE theory.*

4 KPZ and Rough Heat Equation

We discuss the relations of the RBE with the KPZ equation and the multiplicative heat equation in the paracontrolled setting.

4.1 The KPZ equation

For the KPZ equation

$$\mathscr{L}h = (\mathrm{D}h)^2 + \xi, \qquad h(0) = h_0,$$

one can proceed, at least formally, as for the SBE equation by introducing a series of driving terms $(Y^\tau)_{\tau \in \mathcal{T}}$ defined recursively by

$$\mathscr{L}Y^\bullet = \xi, \qquad \mathscr{L}Y^{(\tau_1 \tau_2)} = (\mathrm{D}Y^{\tau_1} \mathrm{D}Y^{\tau_2})$$

and such that $\mathrm{D}Y^\tau = X^\tau$ for all $\tau \in \mathcal{T}$; we will usually write $Y = Y^\bullet$. However, now we have to be more careful because it will turn out that for the white noise ξ some of the terms Y^τ can only be constructed after a suitable renormalization. Let us consider for example Y^{V}. If $\varphi: \mathbb{R} \to \mathbb{R}$ is a smooth compactly supported function with

$\varphi(0)=1$ and $\xi_\varepsilon=\varphi(\varepsilon\mathrm{D})\xi$, then as $\varepsilon\to 0$ we only obtain the convergence of $Y_\varepsilon^{\vee}$ after subtracting suitable diverging constants (which are deterministic): there exist $(c_\varepsilon)_{\varepsilon>0}$, with $\lim_{\varepsilon\to 0} c_\varepsilon=\infty$, such that $(Y_\varepsilon^{\vee}(t)-c_\varepsilon t)_{t\geqslant 0}$ converges to a limit $Y^{\vee}$. We stress the fact that while c_ε depends on the specific choice of φ, the limit $Y^{\vee}$ does not. To make the constant c_ε appear in the equation, we should solve for

$$\mathscr{L}h_\varepsilon=(\mathrm{D}h_\varepsilon)^2-c_\varepsilon+\xi_\varepsilon, \qquad h_\varepsilon(0)=\varphi(\varepsilon\mathrm{D})h_0.$$

In that sense, the limiting object h will actually solve

$$\mathscr{L}h=(\mathrm{D}h)^{\diamond 2}+\xi, \qquad h(0)=h_0, \tag{33}$$

where $(\mathrm{D}h)^{\diamond 2}=(\mathrm{D}h)^2-\infty$ denotes a renormalized product. The reason why we did not see this renormalization appear in Burgers equation is that there we differentiated after the multiplication, considering for example $X_\varepsilon^{\vee}(t)=\mathrm{D}(Y_\varepsilon^{\vee}(t)-c_\varepsilon t)=\mathrm{D}Y_\varepsilon^{\vee}(t)$. There are two other terms which need to be renormalized, $Y^{\vee^{\!\vee}_{\circ}}$ and $Y^{\overset{\vee\vee}{\vee}}$. But on the level of the equation these renormalizations will cancel exactly, so that it suffices to introduce the renormalization of $Y_\varepsilon^{\vee}$ by subtracting c_ε. We still consider a fixed $\alpha\in(1/3,1/2)$.

Definition 4.1. (KPZ-enhancement) Let

$$\mathcal{Y}_{\mathrm{kpz}}\subseteq\mathscr{L}^{\alpha}\times\mathscr{L}^{2\alpha}\times\mathscr{L}^{\alpha+1}\times\mathscr{L}^{2\alpha+1}\times\mathscr{L}^{2\alpha+1}\times C\mathscr{C}^{2\alpha-1}$$

be the closure of the image of the map

$$\Theta_{\mathrm{kpz}}:\mathscr{L}C^{\alpha/2}_{\mathrm{loc}}(\mathbb{R},C^\infty(\mathbb{T}))\times\mathbb{R}\times\mathbb{R}\to\mathcal{Y}_{\mathrm{kpz}},$$

given by

$$\Theta_{\mathrm{kpz}}(\theta,c^{\vee},c^{\overset{\vee\vee}{\vee}})=(Y(\theta),Y^{\vee}(\theta),Y^{\vee^{\!\vee}}(\theta),Y^{\vee^{\!\vee}_{\circ}}(\theta),Y^{\overset{\vee\vee}{\vee}}(\theta),\mathrm{D}P(\theta)\circ\mathrm{D}Y(\theta)), \tag{34}$$

where

$$\begin{array}{rcl}
\mathscr{L}Y(\theta) & = & \theta,\\
\mathscr{L}Y^{\vee}(\theta) & = & (\mathrm{D}Y(\theta))^2-c^{\vee},\\
\mathscr{L}Y^{\vee^{\!\vee}}(\theta) & = & \mathrm{D}Y(\theta)\mathrm{D}Y^{\vee}(\theta),\\
\mathscr{L}Y^{\vee^{\!\vee}_{\circ}}(\theta) & = & \mathrm{D}Y^{\vee^{\!\vee}}\circ\mathrm{D}Y+c^{\overset{\vee\vee}{\vee}}/4\\
\mathscr{L}Y^{\overset{\vee\vee}{\vee}}(\theta) & = & \mathrm{D}Y^{\vee}(\theta)\mathrm{D}Y^{\vee}(\theta)-c^{\overset{\vee\vee}{\vee}},\\
\mathscr{L}P(\theta) & = & \mathrm{D}Y(\theta),
\end{array} \tag{35}$$

all with zero initial conditions except $Y(\theta)$ for which we choose the "stationary" initial condition

$$Y(\theta)(0)=\int_{-\infty}^{0}P_{-s}\Pi_{\neq 0}\theta(s)\mathrm{d}s.$$

Here and in the following $\Pi_{\neq 0}$ denotes the projection on the (spatial) Fourier modes $\neq 0$. We call $\Theta_{\mathrm{kpz}}(\theta,c^{\vee},c^{\overset{\vee\vee}{\vee}})$ the *renormalized KPZ-enhancement* of the driving distribution θ. For $T>0$ we define $\mathcal{Y}_{\mathrm{kpz}}(T)=\mathcal{Y}_{\mathrm{kpz}}|_{[0,T]}$ and we write $\|\mathbb{Y}\|_{\mathcal{Y}_{\mathrm{kpz}}(T)}$ for the norm of $\mathbb{Y}$ in the Banach space $\mathscr{L}^{\alpha}_T\times\mathscr{L}^{2\alpha}_T\times\mathscr{L}^{\alpha+1}_T\times\mathscr{L}^{2\alpha+1}_T\times\mathscr{L}^{2\alpha+1}_T\times C_T\mathscr{C}^{2\alpha-1}$. Moreover, we define the distance $d_{\mathcal{Y}_{\mathrm{kpz}}(T)}(\mathbb{Y},\tilde{\mathbb{Y}})=\|\mathbb{Y}-\tilde{\mathbb{Y}}\|_{\mathcal{Y}_{\mathrm{kpz}}(T)}$.

Observe the link between the renormalizations for $Y^{\vee\vee}$ and $Y^{\overset{\vee}{\curlyvee}}$. It is chosen so that $c^{\vee\vee}$ will never appear in the equation. It is known from the work of Hairer [Hai13], see also Section 9.6, that in the case where θ is a suitable approximation of the space-time white noise it is possible to find such renormalizations for $Y^{\vee\vee}$ and $Y^{\overset{\vee}{\curlyvee}}$. Also note that for every $\mathbb{Y} \in \mathcal{Y}_{\mathrm{kpz}}$ there exists an associated $\mathbb{X} \in \mathcal{X}_{\mathrm{rbe}}$, obtained by differentiating the first five entries of $\mathbb{Y}$ and keeping the sixth entry fixed. Abusing notation, we write $\mathbb{X} = \mathrm{D}\mathbb{Y}$.

The paracontrolled ansatz for KPZ is $h \in \mathscr{D}_{\mathrm{kpz}} = \mathscr{D}_{\mathrm{kpz},\mathbb{Y}}$ if

$$h = Y + Y^{\vee} + 2Y^{\curlyvee} + h^P, \qquad h^P = h' \prec\!\!\prec P + h^\sharp, \tag{36}$$

where $h' \in \mathscr{L}^{\alpha}$, and $h^\sharp \in \mathscr{L}^{2\alpha+1}$. For $h \in \mathscr{D}_{\mathrm{kpz}}$ it now follows from the same arguments as in the case of Burgers equation that the term

$$(\mathrm{D}h)^{\diamond 2} = (\mathrm{D}h)^2 - c^{\vee}$$

is well defined as a bounded multilinear operator. Now we obtain the local existence and uniqueness of paracontrolled solutions to the KPZ equation (33) exactly as for the Burgers equation, at least as long as the initial condition is of the form $Y(0) + h(0)$ with $h(0) \in \mathscr{C}^{2\alpha+1}$.

Moreover, let $\mathbb{X} = \mathrm{D}\mathbb{Y} \in \mathcal{X}_{\mathrm{rbe}}$, and let $h \in \mathscr{D}_{\mathrm{kpz},\mathbb{Y}}$. Then $\mathrm{D}h = X + X^{\vee} + 2X^{\curlyvee} + \mathrm{D}h^P$, where

$$\mathrm{D}h^P = (\mathrm{D}h') \prec\!\!\prec P + h' \prec\!\!\prec \mathrm{D}P + \mathrm{D}h^\sharp = h' \prec\!\!\prec Q + (\mathrm{D}h)^\sharp$$

with $(\mathrm{D}h)^\sharp \in \mathscr{L}^{2\alpha-\varepsilon}$ for all $\varepsilon > 0$. This follows from the fact that $\mathscr{L}^{2\alpha+1} \subseteq C^{(2\alpha-\varepsilon)/2}_{\mathrm{loc}}\mathscr{C}^{1+\varepsilon}$. In other words $\mathrm{D}h \in \mathscr{D}^{\delta}_{\mathrm{rbe},\mathbb{X}}$ for all $\delta < \alpha$ (using an interpolation argument one can actually show that $(\mathrm{D}h)^\sharp \in \mathscr{L}^{2\alpha}$, but we will not need this). And if $h \in \mathscr{D}_{\mathrm{kpz}}$ solves the KPZ equation with initial condition h_0, then $u = \mathrm{D}h \in \mathscr{D}^{\delta}_{\mathrm{rbe}}$ solves

$$\mathscr{L}u = \mathrm{D}\mathscr{L}h = \mathrm{D}((\mathrm{D}h)^{\diamond 2} + \theta) = \mathrm{D}u^2 + \mathrm{D}\theta,$$

where we used that $\mathrm{D}c^{\vee} = 0$ (or more rigorously that $\mathrm{D}Y^{\vee} = X^{\vee}$). So u solves the Burgers equation with initial condition $\mathrm{D}h_0$.

Conversely, let $u \in \mathscr{D}_{\mathrm{rbe}}$. It is easy to see that $u = \mathrm{D}\tilde{h} + f$, where $\tilde{h} \in \mathscr{D}_{\mathrm{kpz}}$ and $f \in C\mathscr{C}^{2\alpha}$. Therefore, the renormalized product $u^{\diamond 2} = u^2 - c^{\vee}$ is well defined. Note that $u^{\diamond 2}$ does not depend on the decomposition $u = \mathrm{D}\tilde{h} + f$. The linear equation

$$\mathscr{L}h = u^{\diamond 2} + \theta, \qquad h(0) = Y(0) + h_0,$$

has a unique global in time solution h for all $h_0 \in \mathscr{C}^{2\alpha+1}$. Setting $h^P = h - Y - Y^{\vee} - 2Y^{\curlyvee}$, we get

$$\mathscr{L}h^P = \mathscr{L}(h - Y - Y^{\vee} - 2Y^{\curlyvee}) = u^{\diamond 2} - ((\mathrm{D}Y)^2 - c^{\vee}) - 2\mathrm{D}Y\mathrm{D}Y^{\vee}.$$

Recalling that $u^{\diamond 2} = (\mathrm{D}\tilde{h})^{\diamond 2} + 2f\mathrm{D}\tilde{h} + f^2 = ((\mathrm{D}Y)^2 - c^{\vee}) + 2\mathrm{D}Y\mathrm{D}Y^{\vee} + \mathscr{L}(\mathscr{L}^{\alpha+1})$, we deduce that $h^P \in \mathscr{L}^{\alpha+1}$. Furthermore, if we set $h' = 2u^Q + 4X^{\curlyvee}$ and expand $u^{\diamond 2}$, then

$$\mathscr{L}(h^P - h' \prec\!\!\prec P) = u^{\diamond 2} - ((\mathrm{D}Y)^2 - c^{\vee}) - 2\mathrm{D}Y\mathrm{D}Y^{\vee}$$

$$
\begin{aligned}
&- \big(\mathscr{L}(h' \prec\!\!\prec P) - h' \prec\!\!\prec \mathscr{L}P\big) + h' \prec\!\!\prec \mathscr{L}P \\
&= (h' \prec X - h' \prec\!\!\prec X) + \mathscr{L}(\mathscr{L}^{2\alpha+1}).
\end{aligned}
$$

Using Lemma 2.8, we get that $h^\sharp \in \mathscr{L}^{2\alpha+1}$ and therefore $h \in \mathscr{D}_{\mathrm{kpz}}$. By construction, $\mathrm{D}h$ solves the equation $\mathscr{L}(\mathrm{D}h) = \mathrm{D}u^2 + \mathrm{D}\theta$ with initial condition $\mathrm{D}h(0) = X(0) + \mathrm{D}h_0$. So if u solves the Burgers equation with initial condition $X(0)+\mathrm{D}h_0$, then $\mathscr{L}(\mathrm{D}h-u) = 0$ and $(\mathrm{D}h - u)(0) = 0$, and therefore $\mathrm{D}h = u$. As a consequence, h solves the KPZ equation with initial condition $Y(0) + h_0$.

Theorem 4.2. *For every $(\mathbb{Y}, h_0) \in \mathcal{X}_{\mathrm{kpz}} \times \mathscr{C}^{2\alpha+1}$ there exists $T^* \in (0,\infty]$ such that for all $T < T^*$ there is a unique solution $(h, h', h^\sharp) \in \mathscr{D}_{\mathrm{kpz}}(T)$ to the KPZ equation*

$$
\mathscr{L}h = (\mathrm{D}h)^{\diamond 2} + \theta, \qquad h(0) = Y(0) + h_0, \tag{37}
$$

on the interval $[0,T]$. The map that sends $(\mathbb{Y}, h_0) \in \mathcal{Y}_{\mathrm{kpz}} \times \mathscr{C}^{2\alpha+1}$ to the solution $h \in \mathscr{D}_{\mathrm{kpz},\mathbb{Y}}$ is "continuous" in the sense of Theorem 3.7 and we can choose

$$
T^* = \sup\{t \geqslant 0 : \|h\|_{\mathscr{D}_{\mathrm{kpz}}(t)} < \infty\}.
$$

If $\mathbb{X} = \mathrm{D}\mathbb{Y}$, then $\mathbb{X} \in \mathcal{X}_{\mathrm{rbe}}$ and $u = \mathrm{D}h \in \mathscr{D}_{\mathrm{rbe}}(T)$ solves the rough Burgers equation with data $\mathbb{X}$ and initial condition $u(0) = X(0) + \mathrm{D}h_0$. Conversely, given the solution u to Burgers equation driven by $\mathbb{X}$ and started in $X(0) + \mathrm{D}h_0$, the solution $h \in \mathscr{D}_{\mathrm{kpz}}(T)$ to the linear PDE $\mathscr{L}h = u^{\diamond 2} + \theta$, $h(0) = Y(0) + h_0$, solves the KPZ equation driven by $\mathbb{Y}$ and with initial condition $Y(0) + h_0$.

4.2 The Rough Heat Equation

Let us now solve

$$
\mathscr{L}w = w \diamond \xi, \qquad w(0) = w_0, \tag{38}
$$

where $\diamond$ denotes again a suitably renormalized product and where ξ is the space-time white noise. We replace for the moment ξ by some $\theta \in C(\mathbb{R}, C^\infty(\mathbb{T}))$ and let $\mathbb{Y} = \Theta_{\mathrm{kpz}}(\theta, c^{\vee}, c^{\vee\vee})$ for some $c^{\vee}, c^{\vee\vee} \in \mathbb{R}$ be as in Definition 4.1. Then we know from the Cole–Hopf transform that the unique classical solution to the heat equation $\mathscr{L}w = w(\theta - c^{\vee})$, $h(0) = e^{h_0}$ is given by $w = \exp(h)$, where h solves the KPZ equation with initial condition h_0. Indeed, we have

$$
\mathscr{L}w = \exp(h)\mathscr{L}h - \exp(h)(\mathrm{D}h)^2 = \exp(h)((\mathrm{D}h)^2 - c^{\vee} + \theta) - \exp(h)(\mathrm{D}h)^2 = w(\theta - c^{\vee}).
$$

Extending the Cole–Hopf transform to the paracontrolled setting, we suspect that the solution will be of the form $w = e^{Y + Y^{\vee} + 2Y^{\curlyvee} + h^P}$ where $h^P = h' \prec\!\!\prec P + h^\sharp$ with $h' \in \mathscr{L}^\alpha$ and $h^\sharp \in \mathscr{L}^{2\alpha+1}$, or in other words

$$
w = e^{Y + Y^{\vee} + 2Y^{\curlyvee}} w^P, \qquad w^P = w' \prec\!\!\prec P + w^\sharp,
$$

with $w' \in \mathscr{L}^\alpha$ and $w^\sharp \in \mathscr{L}^{2\alpha+1}$. So we call w paracontrolled and write $w \in \mathscr{D}_{\mathrm{rhe},\mathbb{Y}} = \mathscr{D}_{\mathrm{rhe}}$ if it is of this form.

Our first objective is to make sense of the term $w \diamond \theta$ for $w \in \mathscr{D}_{\mathrm{rhe}}$. In the renormalized smooth case $\mathbb{Y} = \Theta_{\mathrm{kpz}}(\theta, c^{\vee}, c^{\vee\!\vee})$ for $\theta \in C(\mathbb{R}_+, C^{\infty}(\mathbb{T}))$ and $c^{\vee}, c^{\vee\!\vee} \in \mathbb{R}$ we have for $w \in \mathscr{D}_{\mathrm{rhe}}$

$$\mathscr{L}w = e^{Y+Y^{\vee}+2Y^{\curlyvee}} \Big[\mathscr{L}w^P + \mathscr{L}(Y + Y^{\vee} + 2Y^{\curlyvee})w^P - (\mathrm{D}Y + \mathrm{D}Y^{\vee} + 2\mathrm{D}Y^{\curlyvee})^2 w^P\Big] \\ - 2e^{Y+Y^{\vee}+2Y^{\curlyvee}}\mathrm{D}(Y + Y^{\vee} + 2Y^{\curlyvee})\mathrm{D}w^P.$$

Taking into account that $\mathscr{L}Y = \theta$, $\mathscr{L}Y^{\vee} = (\mathrm{D}Y)^2 - c^{\vee}$, and $\mathscr{L}Y^{\curlyvee} = \mathrm{D}Y\mathrm{D}Y^{\vee}$, we get

$$w(\theta - c^{\vee}) = \mathscr{L}w - e^{Y+Y^{\vee}+2Y^{\curlyvee}}\Big[- [4(\mathscr{L}Y^{\curlyvee^{\vee}} + \mathrm{D}Y^{\curlyvee} \prec \mathrm{D}Y + \mathrm{D}Y^{\curlyvee} \succ \mathrm{D}Y) + \mathscr{L}Y^{\vee\!\vee}]w^P \\ + [4\mathrm{D}Y^{\vee}\mathrm{D}Y^{\curlyvee} + (2\mathrm{D}Y^{\curlyvee})^2]w^P \\ + \mathscr{L}w^P - 2\mathrm{D}(Y + Y^{\vee} + 2Y^{\curlyvee})\mathrm{D}w^P\Big]. \qquad (39)$$

At this point we can simply take the right hand side as the *definition* of $w \diamond \theta = w(\theta - c^{\vee})$. For smooth θ and $\mathbb{Y} = \Theta_{\mathrm{rbe}}(\theta, c^{\vee}, c^{\vee\!\vee})$, we have just seen that $w \diamond \theta$ is nothing but the renormalized pointwise product $w(\theta - c^{\vee})$. Moreover, the operation $w \diamond \theta$ is continuous:

Lemma 4.3. *The distribution $w \diamond \theta$ is well defined for all $\mathbb{Y} = \Theta_{\mathrm{rbe}}(\theta, c^{\vee}, c^{\vee\!\vee})$ with $\theta \in \mathscr{L}C^{\alpha/2}_{\mathrm{loc}}(\mathbb{R}, C^{\infty}(\mathbb{T}))$, $c^{\vee}, c^{\vee\!\vee} \in \mathbb{R}$, $w' \in \mathscr{L}^{\alpha}$ and $w^{\sharp} \in \mathscr{L}^{2\alpha}$. Moreover, $w \diamond \theta$ depends continuously on $(\mathbb{Y}, w', w^{\sharp}) \in \mathcal{Y}_{\mathrm{kpz}} \times \mathscr{L}^{\alpha} \times \mathscr{L}^{2\alpha}$.*

Proof. Consider (39). Inside the big square bracket on the right hand side, there are three terms which are not immediately well defined: $\mathrm{D}Y\mathrm{D}w^P$, $\mathscr{L}Y^{\curlyvee^{\vee}}w^P$, and $\mathscr{L}Y^{\vee\!\vee}w^P$. For the first one we can use the fact that $w^P = w' \prec\!\!\prec P + w^{\sharp}$ and that $\mathbb{Y}$ "contains" $\mathrm{D}P \circ \mathrm{D}Y$, just as we did when solving the KPZ equation. The product $\Delta Y^{\curlyvee^{\vee}}w^P$ is well defined because $\Delta Y^{\curlyvee^{\vee}} \in C\mathscr{C}^{2\alpha-1}$ and $w^P \in C\mathscr{C}^{\alpha}$. The product $\partial_t Y^{\curlyvee^{\vee}}w^P$ can be defined using Young integration: we have $Y^{\curlyvee^{\vee}} \in C^{\alpha+1/2}_{\mathrm{loc}}L^{\infty}$ and $w^P \in C^{\alpha/2}_{\mathrm{loc}}L^{\infty}$ and $3\alpha/2 + 1/2 > 1$. The same arguments show that $\mathscr{L}Y^{\vee\!\vee}w^P$ is well defined.

So taking into account that $w^P = w' \prec\!\!\prec P + w^{\sharp}$ and using Lemma 2.8 as well as Bony's associativity result Lemma 2.6, we see that the term in the square brackets is of the form

$$\partial_t(w^{\sharp} + Y^{\curlyvee^{\vee}} + Y^{\vee\!\vee}) + (w' - 4w^P\mathrm{D}Y^{\curlyvee} - 2\mathrm{D}w^P) \prec \mathrm{D}Y + C\mathscr{C}^{2\alpha-1}.$$

The second term is paracontrolled by $\mathrm{D}Y$, and since $e^{Y+Y^{\vee}+2Y^{\curlyvee}}$ is paracontrolled by Y, the product between the two is well defined as long as $Y \circ \mathrm{D}Y$ is given. But of course $Y \circ \mathrm{D}Y = \mathrm{D}(Y \circ Y)/2$. Finally, the product $e^{Y+Y^{\vee}+2Y^{\curlyvee}}\partial_t(w^{\sharp} + Y^{\curlyvee^{\vee}} + Y^{\vee\!\vee})$ can be defined using Young integration as described above. $\square$

Besides it agreeing with the renormalized pointwise product in the smooth case and being continuous in suitable paracontrolled spaces, there is another indication that our definition of $w \diamond \theta$ is useful, despite its unusual appearance.

Proposition 4.4. *Let ξ be a space-time white noise on $\mathbb{R} \times \mathbb{T}$, and let $(\mathcal{F}_t)_{t \geqslant 0}$ be its natural filtration,*

$$\mathcal{F}_t^{\circ} = \sigma\left(\xi(\mathbb{1}_{(s_1,s_2]}\psi) : -\infty < s_1 < s_2 \leqslant t, \psi \in C^{\infty}(\mathbb{T},\mathbb{R})\right), \qquad t \geqslant 0,$$

and $(\mathcal{F}_t)_{t\geqslant 0}$ is its right-continuous completion. Let $\mathbb{Y} \in \mathcal{Y}_{\mathrm{kpz}}$ be constructed from ξ as described in Theorem 9.3, and let $(w_t)_{t\geqslant 0}$ be an adapted process with values in $\mathscr{L}^{\alpha+1}$, such that almost surely $w \in \mathscr{D}_{\mathrm{kpz},\mathbb{Y}}$ with adapted w' and $w^{\sharp}$. Let $\psi : \mathbb{R}_+ \times \mathbb{T} \to \mathbb{R}$ be a smooth test function of compact support. Then

$$\int_{\mathbb{R}_+\times\mathbb{T}} \psi(s,x)(w \diamond \xi)(s,x)\mathrm{d}x\mathrm{d}s = \int_{\mathbb{R}_+\times\mathbb{T}} (\psi(s,x)w(s,x))\xi(s,x)\mathrm{d}x\mathrm{d}s,$$

where the left hand side is to be interpreted using our definition of $w \diamond \xi$, and the right hand side denotes the Itô integral.

Proof. Let φ be a smooth compactly supported even function with $\varphi(0) = 1$ and set $\xi_\varepsilon = \varphi(\varepsilon \mathrm{D})\xi$. Theorem 9.3 states that for $c_\varepsilon^{\vee} = \frac{1}{4\pi\varepsilon}\int_{\mathbb{R}} \varphi^2(x)\mathrm{d}x$ and an appropriate choice of $c_\varepsilon^{\vee\!\vee}$, we have that $\Theta_{\mathrm{kpz}}(\xi_\varepsilon, c_\varepsilon^{\vee}, c_\varepsilon^{\vee\!\vee})$ converges in probability to $\mathbb{Y}$ as $\varepsilon \to 0$. Moreover, by the stochastic Fubini theorem ([DPZ14], Theorem 4.33) and the continuity properties of the Itô integral,

$$\int_{\mathbb{R}_+\times\mathbb{T}} (\psi(s,x)w(s,x))\xi(s,x)\mathrm{d}x\mathrm{d}s = \lim_{\varepsilon\to 0}\int_{\mathbb{R}_+\times\mathbb{T}} (\psi(s,x)w_\varepsilon(s,x))\xi_\varepsilon(s,x)\mathrm{d}x\mathrm{d}s,$$

where we set

$$w_\varepsilon = e^{Y_\varepsilon + Y_\varepsilon^{\vee} + 2Y_\varepsilon^{\curlyvee}}(w' \prec\!\!\prec P_\varepsilon + w^{\sharp}),$$

and where the Y_ε^{τ} and P_ε are constructed from φ, $c_\varepsilon^{\vee}$, and $c_\varepsilon^{\vee\!\vee}$ as in Definition 4.1. But by definition of $\mathbb{Y}$ and the continuity properties of our product operator, also $w_\varepsilon \diamond \xi_\varepsilon$ converges in probability to $w \diamond \xi$ in the sense of distributions on $\mathbb{R}_+ \times \mathbb{T}$. It therefore suffices to show that

$$\int_0^t (w_\varepsilon \diamond \xi_\varepsilon)(s,x)\mathrm{d}s = \int_0^t w_\varepsilon(s,x)\xi_\varepsilon(s,x)\mathrm{d}s = \int_0^t w_\varepsilon(s,x)\mathrm{d}_s B_\varepsilon(s,x)$$

for all $(t,x) \in \mathbb{R}_+ \times \mathbb{T}$, where we wrote $B_\varepsilon(t,x) = \int_0^t \xi_\varepsilon(s,x)\mathrm{d}s$ which, as a function of t, is a Brownian motion with covariance $(2\pi)^{-1}\sum_k \varphi(\varepsilon k)^2 = 2c_\varepsilon^{\vee} + o(1)$ for every $x \in \mathbb{T}$. The first term on the right hand side of (39) is $\mathscr{L}w_\varepsilon$. Let us therefore apply Itô's formula: Writing $w^{P_\varepsilon} = w' \prec\!\!\prec P_\varepsilon + w^{\sharp}$, we have for fixed $x \in \mathbb{T}$

$$\begin{aligned}
\mathrm{d}_t w_\varepsilon - \Delta w_\varepsilon \mathrm{d}t = e^{Y_\varepsilon + Y_\varepsilon^{\vee} + 2Y_\varepsilon^{\curlyvee}}\Big[&\mathrm{d}_t w^{P_\varepsilon} - \Delta w^{P_\varepsilon}\mathrm{d}t \\
&+ w^{P_\varepsilon}(\mathrm{d}_t(Y_\varepsilon + Y_\varepsilon^{\vee} + 2Y_\varepsilon^{\curlyvee}) - \Delta(Y_\varepsilon + Y_\varepsilon^{\vee} + 2Y_\varepsilon^{\curlyvee})\mathrm{d}t) \\
&- (\mathrm{D}Y_\varepsilon + \mathrm{D}Y_\varepsilon^{\vee} + 2\mathrm{D}Y_\varepsilon^{\curlyvee})^2 w^{P_\varepsilon}\mathrm{d}t - 2\mathrm{D}(Y_\varepsilon + Y_\varepsilon^{\vee} + 2Y_\varepsilon^{\curlyvee})\mathrm{D}w^{P_\varepsilon}\mathrm{d}t\Big] \\
+ \frac{1}{2}w_\varepsilon \mathrm{d}\langle Y_\varepsilon(\cdot,x)\rangle_t&,
\end{aligned}$$

where we used that all terms except Y_ε have zero quadratic variation. But now $\mathrm{d}\langle Y_\varepsilon(\cdot, x)\rangle_t = \mathrm{d}\langle B_\varepsilon(\cdot, x)\rangle_t = (2c_\varepsilon^{\vee} + o(1))\mathrm{d}t$, and therefore the Itô correction term cancels in the limit with the term $-w_\varepsilon c_\varepsilon^{\vee}\mathrm{d}t$ that we get from $w_\varepsilon(\mathrm{d}_t Y_\varepsilon^{\vee} - \Delta Y_\varepsilon^{\vee}\mathrm{d}t - (\mathrm{D}Y_\varepsilon)^2\mathrm{d}t)$. So if we compare the expression we obtained with the right hand side of (39), we see that

$$w_\varepsilon \diamond \xi_\varepsilon \mathrm{d}t = w_\varepsilon(\mathrm{d}_t Y_\varepsilon - \Delta Y_\varepsilon \mathrm{d}t) + o(1) = w_\varepsilon \mathrm{d}_t B_\varepsilon + o(1),$$

and the proof is complete. □

With our definition (39) of $w \diamond \theta$, the function $w \in \mathscr{D}_{\mathrm{rhe}}$ solves the rough heat equation (38) if and only if

$$e^{Y+Y^{\vee}+2Y^{\curlyvee}}\Big[- [4(\mathscr{L}Y^{\bigvee} + \mathrm{D}Y^{\curlyvee} \prec \mathrm{D}Y + \mathrm{D}Y^{\curlyvee} \succ \mathrm{D}Y) + \mathscr{L}Y^{\vee\vee} + 4\mathrm{D}Y^{\vee}\mathrm{D}Y^{\curlyvee} + (2\mathrm{D}Y^{\curlyvee})^2]w^P + \mathscr{L}w^P - 2\mathrm{D}(Y + Y^{\vee} + 2Y^{\curlyvee})\mathrm{D}w^P\Big] = 0.$$

Since $e^{Y+Y^{\vee}+2Y^{\curlyvee}}$ is a strictly positive continuous function, this is only possible if the term in the square brackets is constantly equal to zero, that is if

$$\mathscr{L}w^P = [4(\mathscr{L}Y^{\bigvee} + \mathrm{D}Y^{\curlyvee} \prec \mathrm{D}Y + \mathrm{D}Y^{\curlyvee} \succ \mathrm{D}Y) + \mathscr{L}Y^{\vee\vee} + 4\mathrm{D}Y^{\vee}\mathrm{D}Y^{\curlyvee} + (2\mathrm{D}Y^{\curlyvee})^2]w^p + 2\mathrm{D}(Y + Y^{\vee} + 2Y^{\curlyvee})\mathrm{D}w^P. \tag{40}$$

We deduce the following equation for $w^\sharp$:

$$\mathscr{L}w^\sharp = (2\mathrm{D}w^P + 4w^P\mathrm{D}Y^{\curlyvee} - w') \prec \mathrm{D}Y + (4\mathscr{L}Y^{\bigvee} + \mathscr{L}Y^{\vee\vee})w^P + \mathscr{L}\mathscr{L}^{2\alpha+1},$$

where we used again Bony's associativity result Lemma 2.6. Thus, we should choose $w' = 2\mathrm{D}w^P + 4\mathrm{D}Y^{\curlyvee}w^P$. Now we can use similar arguments as for the Burgers and KPZ equations to obtain the existence and uniqueness of solutions to the rough heat equation for every $\mathbb{Y} \in \mathcal{Y}_{\mathrm{kpz}}$ with initial condition $e^{Y_0}w_0$ with $w_0 \in \mathscr{C}^{2\alpha}$. Since the equation is linear, now we even obtain the global existence and uniqueness of solutions. Moreover, as limit of nonnegative functions w is nonnegative whenever the initial condition w_0 is nonnegative.

Theorem 4.5. *For all $(\mathbb{Y}, w_0) \in \mathcal{X}_{\mathrm{kpz}} \times \mathscr{C}^{2\alpha+1}$ and all $T > 0$ there is a unique solution $(w, w', w^\sharp) \in \mathscr{D}_{\mathrm{kpz}}(T)$ to the rough heat equation*

$$\mathscr{L}w = w \diamond \theta, \qquad w(0) = e^{Y(0)}w_0, \tag{41}$$

on the interval $[0, T]$. The solution is nonnegative whenever w_0 is nonnegative. If $\mathbb{Y}$ is sampled from the space-time white noise ξ as described in Theorem 9.3, then w is almost surely equal to the Itô solution of (41).

4.3 Relation between KPZ and Rough Heat Equation

From KPZ to heat equation. Let now $h \in \mathscr{D}_{\mathrm{kpz}}$ and set $w = e^h$. Then $w = e^{Y+Y^{\vee}+2Y^{\curlyvee}} w^P$ with $w^P = e^{h^P}$. To see that $w \in \mathscr{D}_{\mathrm{rhe}}$, we need the following lemma:

Lemma 4.6. *Let $F \in C^2(\mathbb{R},\mathbb{R})$, $h' \in \mathscr{L}^{\alpha}$, $P \in \mathscr{L}^{\alpha+1}$ with $\mathscr{L}P \in C\mathscr{C}^{\alpha-1}$, and $h^\sharp \in \mathscr{L}^{2\alpha+1}$. Write $h^P = h' \prec\!\!\prec P + h^\sharp \in \mathscr{L}^{\alpha+1}$. Then we have for all $\varepsilon > 0$*

$$F(h^P) - (F'(h^P)h') \prec\!\!\prec P \in \mathscr{L}^{2\alpha+1-\varepsilon}.$$

Proof. It is easy to see that the difference is in $C\mathscr{C}^{2\alpha+1}$, the difficulty is to get the right temporal regularity. If we apply the heat operator to the difference and use Lemma 2.8 together with Lemma 2.6, we get

$$\begin{aligned}\mathscr{L}(F(h^P) - (F'(h^P)h') \prec\!\!\prec P) &= F'(h^P)\mathscr{L}(h' \prec\!\!\prec P + h^\sharp) - F''(h^P)(\mathrm{D}h^P)^2 \\ &\qquad - (F'(h^P)h') \prec\!\!\prec \mathscr{L}P + C\mathscr{C}^{2\alpha-1} \\ &= F'(h^P)\partial_t h^\sharp + C\mathscr{C}^{2\alpha-1}.\end{aligned}$$

So an application of the Schauder estimates for $\mathscr{L}$ shows that

$$(F(h^P) - (F'(h^P)h') \prec\!\!\prec P - \int_0^{\cdot} P_{\cdot-s}\{F'(h^P(s))\partial_s h^\sharp(s)\}\mathrm{d}s \in \mathscr{L}^{2\alpha+1}.$$

Since $F'(h^P) \in C^{(\alpha+1)/2}L^\infty$ and $h^\sharp \in C^{(2\alpha+1)/2}L^\infty$, the Young integral

$$I(t) = \int_0^t F'(h^P(s))\mathrm{d}h^\sharp(s)$$

is well defined and in $C^{(2\alpha+1)/2}L^\infty$. We can therefore apply Theorem 1 of [GLT06] to see that $\int_0^{\cdot} P_{\cdot-s}\mathrm{d}I(s) \in C^{(2\alpha+1-\varepsilon)/2}L^\infty$ for all $\varepsilon > 0$, and the proof is complete. □

As a consequence of this lemma we have $e^h \in \mathscr{D}^{\delta}_{\mathrm{rhe}}$ for every $h \in \mathscr{D}_{\mathrm{kpz}}$ and $\delta < \alpha$, with derivative $e^{h^P}h' \in \mathscr{L}^{\alpha}$.

Lemma 4.7. *Let $\mathbb{Y} \in \mathcal{Y}_{\mathrm{kpz}}$, $T > 0$, and let $h \in \mathscr{D}_{\mathrm{kpz},\mathbb{Y}}$ solve the KPZ equation on $[0,T]$ with initial condition $h(0) = h_0$, and set $w = e^h \in \mathscr{D}^{\alpha-\varepsilon}_{\mathrm{rhe},\mathbb{Y}}$ with derivative $e^{h^P}h'$. Then w solves the rough heat equation on $[0,T]$ with initial condition $w(0) = e^{h_0}$.*

Proof. Recall that

$$\mathscr{L}h^P = \mathscr{L}(h - Y - Y^{\vee} - 2Y^{\curlyvee}) = (\mathrm{D}h)^{\diamond 2} - ((\mathrm{D}Y)^2 - c^{\vee}) - 2\mathrm{D}Y\mathrm{D}Y^{\vee}.$$

From the chain rule for the heat operator we obtain

$$\mathscr{L}w^P = w^P\mathscr{L}h^P - w^P(\mathrm{D}h^P)^2 = w^P((\mathrm{D}h)^{\diamond 2} - ((\mathrm{D}Y)^2 - c^{\vee}) - 2\mathrm{D}Y\mathrm{D}Y^{\vee} - (\mathrm{D}h^P)^2).$$

But now

$$
\begin{aligned}
&(\mathrm{D}h)^{\diamond 2} - ((\mathrm{D}Y)^2 - c^{\vee}) - 2\mathrm{D}Y\mathrm{D}Y^{\vee} - (\mathrm{D}h^P)^2 \\
&\qquad = 4(\mathscr{L}Y^{\Psi} + \mathrm{D}Y^{\curlyvee} \prec \mathrm{D}Y + \mathrm{D}Y^{\curlyvee} \succ Y) + Y^{\vee\vee} + 4\mathrm{D}Y^{\curlyvee}\mathrm{D}Y^{\vee} + (2\mathrm{D}Y^{\curlyvee})^2 \\
&\qquad\quad + 2\mathrm{D}(Y + Y^{\vee} + 2Y^{\curlyvee})\mathrm{D}h^P,
\end{aligned}
$$

and $\mathrm{D}h^P = \mathrm{D}(\log w^P) = \mathrm{D}w^P/w^P$, from where we deduce that

$$
\begin{aligned}
\mathscr{L}w^P &= [4(\mathscr{L}Y^{\Psi} + \mathrm{D}Y^{\curlyvee} \prec \mathrm{D}Y + \mathrm{D}Y^{\curlyvee} \succ Y) + Y^{\vee\vee} + 4\mathrm{D}Y^{\curlyvee}\mathrm{D}Y^{\vee} + (2\mathrm{D}Y^{\curlyvee})^2]w^P \\
&\quad + 2\mathrm{D}(Y + Y^{\vee} + 2Y^{\curlyvee})\mathrm{D}w^P.
\end{aligned}
$$

In other words, w^P satisfies (40) and w solves the rough heat equation. $\square$

From the heat equation to KPZ. Conversely, let $w = e^{Y+Y^{\vee}+2Y^{\curlyvee}} w^P \in \mathscr{D}_{\mathrm{rhe}}$ be strictly positive and define $h = \log w$. Then $h = Y + Y^{\vee} + 2Y^{\curlyvee} + \log w^P$, and Lemma 4.6 yields $\log w^P = (w'/w^P) \prec\!\!\prec P + \mathscr{L}^{2\alpha+1-\varepsilon}$ for all $\varepsilon > 0$, in other words h is paracontrolled with derivative w'/w^P.

Lemma 4.8. *Let $\mathbb{Y} \in \mathcal{Y}_{\mathrm{kpz}}$, $T > 0$, and let $w \in \mathscr{D}_{\mathrm{rhe},\mathbb{Y}}$ be a strictly positive solution to the rough heat equation with initial condition $w(0) = w_0 > 0$ on the time interval $[0, T]$. Set $h = \log w \in \mathscr{D}^{\alpha-\varepsilon}_{\mathrm{rhe},\mathbb{Y}}$ with derivative w'/w^P. Then h solves the KPZ equation on $[0, T]$ with initial condition $h_0 = \log w_0$.*

Proof. We have

$$
\begin{aligned}
\mathscr{L}h^P &= \mathscr{L}(\log w^P) = \frac{1}{w^P}\mathscr{L}w^P + \frac{1}{(w^P)^2}(\mathrm{D}w^P)^2 \\
&= \frac{1}{w^P}[4(\mathscr{L}Y^{\Psi} + \mathrm{D}Y^{\curlyvee} \prec \mathrm{D}Y + \mathrm{D}Y^{\curlyvee} \succ \mathrm{D}Y) + \mathscr{L}Y^{\vee\vee} + 4\mathrm{D}Y^{\vee}\mathrm{D}Y^{\curlyvee} + (2\mathrm{D}Y^{\curlyvee})^2]w^p \\
&\quad + \frac{1}{w^P}2\mathrm{D}(Y + Y^{\vee} + 2Y^{\curlyvee})\mathrm{D}w^P + (\mathrm{D}h^P)^2,
\end{aligned}
$$

where for the last term we used that $\mathrm{D}w^P = w^P\mathrm{D}h^P$. Using this relation once more, we arrive at

$$
\begin{aligned}
\mathscr{L}h^P &= 4(\mathscr{L}Y^{\Psi} + \mathrm{D}Y^{\curlyvee} \prec \mathrm{D}Y + \mathrm{D}Y^{\curlyvee} \succ \mathrm{D}Y) + \mathscr{L}Y^{\vee\vee} + 4\mathrm{D}Y^{\vee}\mathrm{D}Y^{\curlyvee} + (2\mathrm{D}Y^{\curlyvee})^2 \\
&\quad + 2\mathrm{D}(Y + Y^{\vee} + 2Y^{\curlyvee})\mathrm{D}h^P + (\mathrm{D}h^P)^2,
\end{aligned}
$$

or in other words $\mathscr{L}h = (\mathrm{D}h)^{\diamond 2} + \theta$. $\square$

Let us summarize our observations:

Theorem 4.9. *Let $\mathbb{Y}, w_0, h$ be as in Theorem 4.5 and set*

$$T^* = \inf\{t \geqslant 0 : \min_{x \in \mathbb{T}} w(t, x) = 0\}.$$

Then for all $T < T^*$ *the function* $\log w|_{[0,T]} \in \mathscr{D}_{\mathrm{kpz}}(T)$ *solves the KPZ equation* (37) *driven by* $\mathbb{Y}$ *and started in* $Y(0) + \log w_0$. *Conversely, if* $h \in \mathscr{D}_{\mathrm{kpz}}(T)$ *solves the KPZ equation driven by* $\mathbb{Y}$ *and started in* $Y(0) + h_0$*, then* $w = \exp(h)$ *solves* (41) *with* $w_0 = \exp(h_0)$.

As an immediate consequence we obtain a better characterization of the blow up time for the solution to the KPZ equation:

Corollary 4.10. *In the context of Theorem 4.2, the explosion time* T^* *of the paracontrolled norm of the solution* h *to the KPZ equation is given by*

$$T^* = \inf\{t \geqslant 0 : \min_{x \in \mathbb{T}} \exp(h(t,x)) = 0\} = \sup\{t \geqslant 0 : \|h(t)\|_{L^\infty} < \infty\}.$$

5 Interpretation of the nonlinearity

The purpose of this section is to give a more natural interpretation of the nonlinearity $\mathrm{D}u^2$ that appears in the formulation of Burgers equation. We will show that if ξ is a space-time white noise, then u is the only stochastic process for which there exists u' with $(u, u') \in \mathscr{D}_{\mathrm{rbe}}$ almost surely and such that almost surely

$$\partial_t u = \Delta u + \lim_{\varepsilon \to 0} \mathrm{D}(\varphi(\varepsilon \mathrm{D}) u)^2 + \mathrm{D}\xi, \qquad u(0) = u_0,$$

whenever φ is an even smooth function of compact support with $\varphi(0) = 1$, where the convergence holds in probability in $\mathscr{D}'(\mathbb{R}_+ \times \mathbb{T})$.

This is a simple consequence of the following theorem.

Theorem 5.1. *Let* $\mathbb{X} \in \mathcal{X}_{\mathrm{rbe}}$ *and* $\varphi \in C^\infty(\mathbb{R})$ *with compact support and* $\varphi(0) = 0$. *Define for* $\varepsilon > 0$ *the mollification* $X^\tau_\varepsilon = \varphi(\varepsilon \mathrm{D}) X^\tau$, $Q_\varepsilon = \varphi(\varepsilon \mathrm{D}) Q$*, and then*

$$\mathbb{X}_\varepsilon = (X_\varepsilon, I(\mathrm{D}[(X_\varepsilon)^2]), I(\mathrm{D}[X^{\vee}_\varepsilon X_\varepsilon]), I(\mathrm{D}[X^{\curlyvee}_\varepsilon \circ X_\varepsilon]), I(\mathrm{D}[(X^{\vee}_\varepsilon)^2]), Q_\varepsilon \circ X_\varepsilon).$$

Let $T > 0$ *and assume that* $\|\mathbb{X}_\varepsilon - \mathbb{X}\|_{\mathcal{X}_{\mathrm{rbe}}(T)}$ *converges to 0 as* $\varepsilon \to 0$. *Let* $u \in \mathscr{D}_{\mathrm{rbe}}(T)$. *Then*

$$\lim_{\varepsilon \to 0} I(\mathrm{D}[(\varphi(\varepsilon \mathrm{D}) u)^2]) = I(\mathrm{D}u^2), \tag{42}$$

where

$$\begin{aligned} I(\mathrm{D}u^2) = {} & X^{\vee} + 2X^{\curlyvee} + X^{\vee\vee} + 4X^{\curlyvee\!\!\curlyvee} \\ & + I(2\mathrm{D}(u^Q X) + 2\mathrm{D}(X^{\vee}(u^Q + 2X^{\curlyvee})) + \mathrm{D}((u^Q + 2X^{\curlyvee})^2)) \end{aligned} \tag{43}$$

and

$$u^Q X = u \prec X + u \succ X + u^\sharp \circ X + ((u' \prec\!\!\prec Q) - u' \prec Q) \circ X + C(u', Q, X) + u'(Q \circ X) \tag{44}$$

denotes the paracontrolled square and where the convergence takes place in $C([0,T], \mathscr{D}')$.

Proof. We have $\varphi(\varepsilon \mathrm{D})u = (X_\varepsilon + X_\varepsilon^{\vee} + 2X_\varepsilon^{\curlyvee} + \varphi(\varepsilon \mathrm{D})u^Q)$, and our assumptions are chosen exactly so that the convergence of all terms of $I(\mathrm{D}[(\varphi(\varepsilon \mathrm{D})u)^2])$ is trivial, except that of $I(\mathrm{D}[(\varphi(\varepsilon \mathrm{D})(u' \prec\!\!\prec Q)) \circ X_\varepsilon])$ to $I(\mathrm{D}[(u' \prec\!\!\prec Q) \circ X])$, where the second resonant product is interpreted in the paracontrolled sense. Since $f \mapsto I(\mathrm{D}f)$ is a continuous operation on $C([0,T], \mathscr{D}')$, it suffices to show that $(\varphi(\varepsilon \mathrm{D})(u' \prec\!\!\prec Q)) \circ X_\varepsilon$ converges in $C([0,T], \mathscr{D}')$ to $(u' \prec\!\!\prec Q) \circ X$. We decompose

$$(\varphi(\varepsilon \mathrm{D})u^Q) \circ X_\varepsilon = (\varphi(\varepsilon \mathrm{D})[(u' \prec\!\!\prec Q) - (u' \prec Q)]) \circ X_\varepsilon + (\varphi(\varepsilon \mathrm{D})(u' \prec Q)) \circ X_\varepsilon$$

and use the continuity properties of the resonant term in combination with Lemma 2.8 to conclude that the first term on the right hand side converges to its "without ε counterpart". It remains to treat $(\varphi(\varepsilon \mathrm{D})(u' \prec Q)) \circ X_\varepsilon$. But now Lemma 5.3.20 of [Per14] states that

$$\|\varphi(\varepsilon \mathrm{D})(u' \prec Q) - u' \prec Q_\varepsilon\|_{C_T \mathscr{C}^{2\alpha-\delta}} \lesssim \varepsilon^\delta \|u'\|_{C_T \mathscr{C}^\alpha} \|Q\|_{C_T \mathscr{C}^\alpha}$$

whenever $\delta \leqslant 1$. If we choose $\delta > 0$ small enough so that $3\alpha - \delta > 1$, this allows us to replace $(\varphi(\varepsilon \mathrm{D})(u' \prec Q)) \circ X_\varepsilon$ by $(u' \prec Q_\varepsilon) \circ X_\varepsilon$. Then we get

$$(u' \prec Q_\varepsilon) \circ X_\varepsilon = C(u', Q_\varepsilon, X_\varepsilon) + u'(Q_\varepsilon \circ X_\varepsilon),$$

and since C is continuous and by assumption $Q_\varepsilon \circ X_\varepsilon$ converges to $Q \circ X$ in $C_T \mathscr{C}^{2\alpha-1}$, this completes the proof. □

Corollary 5.2. *Let ξ be a space-time white noise on $\mathbb{R} \times \mathbb{T}$, let $\mathbb{X} \in \mathcal{X}_{\mathrm{rbe}}$ be constructed from ξ as described in Theorem 9.1, and let $u_0 \in \mathscr{C}^{2\alpha}$ almost surely. Assume that u is a stochastic process for which there exists u' and $u^\sharp$ with $(u, u', u^\sharp) \in \mathscr{D}_{\mathrm{rbe}}$ almost surely and such that almost surely*

$$\partial_t u = \Delta u + \lim_{\varepsilon \to 0} \mathrm{D}(\varphi(\varepsilon \mathrm{D})u)^2 + \mathrm{D}\xi, \qquad u(0) = u_0 + Y(0), \tag{45}$$

whenever φ is an even smooth function of compact support with $\varphi(0) = 1$, where the convergence holds in probability in $\mathscr{D}'(\mathbb{R}_+ \times \mathbb{T})$. Then u is almost surely equal to the unique paracontrolled solution to the rough Burgers equation driven by $\mathbb{X}$ and started in $u_0 + Y(0)$.

Proof. Theorem 5.1 and the convergence result of Theorem 9.1 imply that almost surely

$$\mathscr{L}u = \mathrm{D}u^2 + \mathrm{D}\xi, \qquad u(0) = u_0 + Y(0),$$

where the square $\mathrm{D}u^2$ is interpreted as in (43), (44) using the paracontrolled structure of $(u, u', u^\sharp)$. Therefore, there exists a unique $\tilde{u}^\sharp$ such that also $(u, 2u^Q + 4X^{\curlyvee}, \tilde{u}^\sharp)$ is almost surely paracontrolled, where $u^Q = u - X - X^{\vee} - 2X^{\curlyvee}$. However, another application of Theorem 5.1 shows that $\mathrm{D}u^2$ can be computed without explicit reference to the paracontrolled derivative and remainder of u and thus it agrees for $(u, u', u^\sharp)$ and for $(u, 2u^Q + 4X^{\curlyvee}, \tilde{u}^\sharp)$, and therefore $(u, 2u^Q + 4X^{\curlyvee}, \tilde{u}^\sharp)$ is the paracontrolled solution of the equation. □

Remark 5.3. *The conclusion of Corollary 5.2 should be read more precisely as follows: There exists a unique $\tilde{u}^\sharp$ such that with $u^Q = u - X - X^{\vee} - 2X^{\curlyvee}$ the tuple $(u, 2u^Q + 4X^{\curlyvee}, \tilde{u}^\sharp)$ is almost surely equal to the unique paracontrolled solution to the rough Burgers equation driven by $\mathbb{X}$ and started in $u_0 + Y(0)$.*

Remark 5.4. *There have been several efforts to find formulations of the Burgers equation based on interpretations like* (45)*, see for example* [Ass02, GJ14, Ass13, CO14, GJ13]*. Corollary 5.2 shows that any solution satisfying* (45) *is unique provided that it is additionally paracontrolled.*

Remark 5.5. *For simplicity we formulated the result under the assumption $u(0) = Y(0) + u_0$ for some relatively regular perturbation u_0. Using the techniques of Section 6 it is possible to extend this to general $u(0) \in \mathscr{C}^{-1+\varepsilon}$.*

6 Singular initial conditions

In this section we extend the results of Section 3 to allow for more general initial conditions. For that purpose we adapt the definition of paracontrolled distributions. To allow for initial conditions that are not necessarily regular perturbations of $X(0)$, we will introduce weighted norms which permit a possible singularity of the paracontrolled norm at 0. Moreover, as was kindly pointed out to us by Khalil Chouk, in the case of the rough heat equation it is very convenient to work on

$$\mathscr{C}_p^\alpha = B_{p,\infty}^\alpha$$

for $\alpha \in \mathbb{R}$, rather than restricting ourselves to $p = \infty$. For example the Dirac delta has regularity $\mathscr{C}_p^{1/p-1}$ on $\mathbb{T}$, so by working in $\mathscr{C}_1^\alpha$ spaces we will be able to start the rough heat equation from $(-\Delta)^\beta \delta_0$ for any $\beta < 1/4$. In fact for the same reasons and in the setting of regularity structures an approach based on Besov spaces with finite integrability index was developed recently in [HL15].

Recall that the space $\mathscr{C}_p^\alpha$ is defined as

$$\mathscr{C}_p^\alpha = \{f \in \mathscr{D}'(\mathbb{T}) : \|f\|_{\mathscr{C}_p^\alpha} := \sup_{j \geqslant -1} 2^{j\alpha} \|\Delta_j f\|_{L^p} < \infty\}$$

For $p \in [1, \infty]$, $\gamma \geqslant 0$ and $T > 0$ we define $\mathcal{M}_T^\gamma L^p = \{v \colon [0,T] \to \mathscr{D}'(\mathbb{T}) : \|v\|_{\mathcal{M}_T^\gamma L^p} < \infty\}$, where

$$\|v\|_{\mathcal{M}_T^\gamma L^p} = \sup_{t \in [0,T]} \{\|t^\gamma v(t)\|_{L^p}\}.$$

In particular $\|v\|_{\mathcal{M}_T^\gamma L^p} = \|v\|_{C_T L^p}$ for $\gamma = 0$. If further $\alpha \in (0,2)$ we define the norm

$$\|f\|_{\mathscr{L}_p^{\gamma,\alpha}(T)} = \max\{\|t \mapsto t^\gamma f(t)\|_{C_T^{\alpha/2} L^p}, \|f\|_{\mathcal{M}_T^\gamma \mathscr{C}_p^\alpha}\}$$

and the space $\mathscr{L}_p^{\gamma,\alpha}(T) = \{f \colon [0,T] \to \mathscr{D}' : \|f\|_{\mathscr{L}_p^{\gamma,\alpha}(T)} < \infty\}$ as well as

$$\mathscr{L}_p^{\gamma,\alpha} = \{f \colon \mathbb{R}_+ \to \mathscr{D}' : f|_{[0,T]} \in \mathscr{L}_p^{\gamma,\alpha}(T) \text{ for all } T > 0\}.$$

In particular, we have $\mathscr{L}_\infty^{0,\alpha}(T) = \mathscr{L}_T^\alpha$.

6.1 Preliminary estimates

Here we translate the results of Section 2.3 to the setting of $\mathscr{L}_p^{\gamma,\alpha}$ spaces. The estimates involving only one fixed time remain essentially unchanged, but we carefully have to revisit every estimate that involves the modified paraproduct $\prec\!\!\!\prec$, as well as the Schauder estimates for the Laplacian.

Let us start with the paraproduct estimates:

Lemma 6.1. *For any $\beta \in \mathbb{R}$ and $p \in [1,\infty]$ we have*

$$\|f \prec g\|_{\mathscr{C}_p^\beta} \lesssim_\beta \min\{\|f\|_{L^\infty}\|g\|_{\mathscr{C}_p^\beta}, \|f\|_{L^p}\|g\|_\beta\},$$

and for $\alpha < 0$ furthermore

$$\|f \prec g\|_{\mathscr{C}_p^{\alpha+\beta}} \lesssim_{\alpha,\beta} \min\{\|f\|_{\mathscr{C}_p^\alpha}\|g\|_\beta, \|f\|_\alpha\|g\|_{\mathscr{C}_p^\beta}\}.$$

For $\alpha + \beta > 0$ we have

$$\|f \circ g\|_{\mathscr{C}_p^{\alpha+\beta}} \lesssim_{\alpha,\beta} \min\{\|f\|_{\mathscr{C}_p^\alpha}\|g\|_\beta, \|f\|_\alpha\|g\|_{\mathscr{C}_p^\beta}\}.$$

To carry over the multiplication theorem of paracontrolled distributions to our setting, we first adapt the meta-definition of a paracontrolled distribution:

Definition 6.2. Let $\beta > 0$, $\alpha \in \mathbb{R}$, and $p \in [1,\infty]$. A distribution $f \in \mathscr{C}_p^\alpha$ is called paracontrolled by $u \in \mathscr{C}^\alpha$ and we write $f \in \mathscr{D}_p^\beta(u)$, if there exists $f' \in \mathscr{C}_p^\beta$ such that $f^\sharp = f - f' \prec u \in \mathscr{C}_p^{\alpha+\beta}$.

Let us stress that $u \in \mathscr{C}^\alpha$ is not a typo, we do not weaken the integrability assumptions on the reference distribution u.

Theorem 6.3. *Let $\alpha, \beta \in (1/3, 1/2)$. Let $u \in \mathscr{C}^\alpha$, $v \in \mathscr{C}^{\alpha-1}$, and let $(f, f') \in \mathscr{D}_\infty^\beta(u)$ and $(g, g') \in \mathscr{D}_p^\beta(v)$. Assume that $u \circ v \in \mathscr{C}^{2\alpha-1}$ is given as limit of $(u_n \circ v_n)$ in $\mathscr{C}^{2\alpha-1}$, where (u_n) and (v_n) are sequences of smooth functions that converge to u in $\mathscr{C}^\alpha$ and to v in $\mathscr{C}^{\alpha-1}$ respectively. Then fg is well defined and satisfies*

$$\|fg - f \prec g\|_{\mathscr{C}_p^{2\alpha-1}} \lesssim (\|f'\|_\beta\|u\|_\alpha + \|f^\sharp\|_{\alpha+\beta})(\|g'\|_{\mathscr{C}_p^\beta}\|v\|_{\alpha-1} + \|g^\sharp\|_{\mathscr{C}_p^{\alpha+\beta-1}}) + \|f'g'\|_{\mathscr{C}_p^\beta}\|u \circ v\|_{2\alpha-1}.$$

Furthermore, the product is locally Lipschitz continuous: Let $\tilde{u} \in \mathscr{C}^\alpha$, $\tilde{v} \in \mathscr{C}^{\alpha-1}$ with $\tilde{u} \circ \tilde{v} \in \mathscr{C}^{2\alpha-1}$ and let $(\tilde{f}, \tilde{f}') \in \mathscr{D}_\infty^\alpha(\tilde{u})$ and $(\tilde{g}, \tilde{g}') \in \mathscr{D}_p^\alpha(\tilde{v})$. Assume that $M > 0$ is an upper bound for the norms of all distributions under consideration. Then

$$\begin{aligned}\|(fg - f \prec g) - (\tilde{f}\tilde{g} - \tilde{f} \prec \tilde{g})\|_{\mathscr{C}_p^{2\alpha-1}} \lesssim (1 + M^3)\Big[&\|f' - \tilde{f}'\|_\beta + \|g' - \tilde{g}'\|_{\mathscr{C}_p^\beta} + \|u - \tilde{u}\|_\alpha \\ &+ \|v - \tilde{v}\|_{\alpha-1} + \|f^\sharp - \tilde{f}^\sharp\|_{\alpha+\beta} \\ &+ \|g^\sharp - \tilde{g}^\sharp\|_{\mathscr{C}_p^{\alpha+\beta-1}}\|u \circ v - \tilde{u} \circ \tilde{v}\|_{2\alpha-1}\Big].\end{aligned}$$

If $f' = \tilde{f}' = 1$ or $g' = \tilde{g}' = 1$, then M^3 can be replaced by M^2.

In this setting the proof is a straightforward adaptation of the arguments in [GIP15], see also [AC15]. For an extension to much more general Besov spaces see [PT16].

Next we get to the modified paraproduct, which we recall was defined as

$$f \prec\!\!\prec g = \sum_i (Q_i S_{i-1} f)\Delta_i g \qquad \text{with} \qquad Q_i f(t) = \int_{\mathbb{R}} 2^{-2i}\varphi(2^{2i}(t-s)) f(s \vee 0) \mathrm{d}s.$$

Since for $f \in \mathcal{M}_T^\gamma L^p$ we have in general $f(0) \notin L^p$, this definition might lead to problems. Recalling that $f \in \mathscr{L}_\infty^{\gamma,\alpha}(T)$ if and only if $t \mapsto t^\gamma f(t) \in \mathscr{L}_T^\gamma$, it seems reasonable to replace f by $f(t)\mathbb{1}_{t>0}$, so that $\int_{-\infty}^0 \varphi(2^{2i}(t-s))f(s\vee 0)\mathrm{d}s = 0$. So in what follows we shall always silently perform this replacement when evaluating $\prec\!\!\prec$ on elements of $\mathscr{L}_p^{\gamma,\alpha}(T)$ or $\mathcal{M}_T^\gamma L^p$.

Lemma 6.4. *For any $\beta \in \mathbb{R}$, $p \in [1, \infty]$, and $\gamma \in [0, 1)$ we have*

$$t^\gamma \| f \prec\!\!\prec g(t)\|_{\mathscr{C}_p^\beta} \lesssim \min\{\|f\|_{\mathcal{M}_t^\gamma L^p}\|g(t)\|_\beta, \|f\|_{\mathcal{M}_t^\gamma L^\infty}\|g(t)\|_{\mathscr{C}_p^\beta}\} \tag{46}$$

for all $t > 0$, and for $\alpha < 0$ furthermore

$$t^\gamma \| f \prec\!\!\prec g(t)\|_{\mathscr{C}_p^{\alpha+\beta}} \lesssim \min\{\|f\|_{\mathcal{M}_t^\gamma \mathscr{C}_p^\alpha}\|g(t)\|_\beta, \|f\|_{\mathcal{M}_t^\gamma \mathscr{C}^\alpha}\|g(t)\|_{\mathscr{C}_p^\beta}\}. \tag{47}$$

Lemma 6.5. *Let $\alpha \in (0, 2)$, $\beta \in \mathbb{R}$, $p \in [1, \infty]$, and $\gamma \in [0, 1)$. Then*

$$t^\gamma \|(f \prec\!\!\prec g - f \prec g)(t)\|_{\mathscr{C}_p^{\alpha+\beta}} \lesssim \|f\|_{\mathscr{L}_p^{\gamma,\alpha}(t)}\|g(t)\|_\beta$$

for all $t > 0$, as well as

$$t^\gamma \left\|\left(\mathscr{L}(f \prec\!\!\prec g) - f \prec\!\!\prec (\mathscr{L}g)\right)(t)\right\|_{\mathscr{C}_p^{\alpha+\beta-2}} \lesssim \|f\|_{\mathscr{L}_p^{\gamma,\alpha}(t)}\|g(t)\|_\beta.$$

These two lemmas are not very difficult to show, but at least the proof for the second one is slightly technical; the proofs can be found in Appendix A.

Recall the definition of the operator $If(t) = \int_0^t P_{t-s}f(s)\mathrm{d}s$. In the singular case we can adapt the Schauder estimates for I as follows:

Lemma 6.6 (Schauder estimates)**.** *Let $\alpha \in (0, 2)$, $p \in [1, \infty]$, and $\gamma \in [0, 1)$. Then*

$$\|If\|_{\mathscr{L}_p^{\gamma,\alpha}(T)} \lesssim \|f\|_{\mathcal{M}_T^\gamma \mathscr{C}_p^{\alpha-2}} \tag{48}$$

for all $T > 0$. If further $\beta \in [-\alpha, 2-\alpha)$, then

$$\|s \mapsto P_s u_0\|_{\mathscr{L}_p^{(\beta+\alpha)/2,\alpha}(T)} \lesssim \|u_0\|_{\mathscr{C}_p^{-\beta}}. \tag{49}$$

For all $\alpha \in \mathbb{R}$, $\gamma \in [0, 1)$, and $T > 0$ we have

$$\|If\|_{\mathcal{M}_T^\gamma \mathscr{C}_p^\alpha} \lesssim \|f\|_{\mathcal{M}_T^\gamma \mathscr{C}_p^{\alpha-2}}. \tag{50}$$

To a large extent the proof is contained in [GIP15]. We indicate in Appendix A how to adapt the arguments therein to obtain the estimates above.

Just as in the "non-singular" setting, the Schauder estimates allow us to bound $f \prec\!\!\prec g$ in the parabolic space $\mathscr{L}_p^{\gamma,\alpha}(T)$:

Lemma 6.7. *Let $\alpha \in (0,2)$, $p \in [1,\infty]$, $\gamma \in [0,1)$ and $\delta > 0$. Let $f \in \mathscr{L}_p^{\gamma,\delta}(T)$, $g \in C_T\mathscr{C}^\alpha$, and $\mathscr{L} g \in C_T\mathscr{C}^{\alpha-2}$. Then*

$$\| f \prec\!\!\prec g \|_{\mathscr{L}_p^{\gamma,\alpha}(T)} \lesssim \| f \|_{\mathscr{L}_p^{\gamma,\delta}(T)} (\| g \|_{C_T\mathscr{C}^\alpha} + \| \mathscr{L} g \|_{C_T\mathscr{C}^{\alpha-2}}).$$

The proof is completely analogous to the one of Lemma 2.10. Finally, we will need a lemma which allows us to pass between different $\mathscr{L}_p^{\gamma,\alpha}$ spaces, the proof of which can be also found in Appendix A.

Lemma 6.8. *Let $\alpha \in (0,2)$, $\gamma \in (0,1)$, $p \in [1,\infty]$, $T > 0$, and let $f \in \mathscr{L}_T^{\gamma,\alpha}$. Then*

$$\| f \|_{\mathscr{L}_p^{\gamma-\varepsilon/2,\alpha-\varepsilon}(T)} \lesssim \| f \|_{\mathscr{L}_p^{\gamma,\alpha}(T)}.$$

for all $\varepsilon \in [0, \alpha \wedge 2\gamma)$.

6.2 Burgers and KPZ equation with singular initial conditions

Let us indicate how to modify our arguments to solve Burgers equation with initial condition $u_0 \in \mathscr{C}^{-\beta}$ for arbitrary $\beta < 1$. Throughout this section we fix $\alpha \in (1/3, 1/2)$ and $\beta \in (1-\alpha, 2\alpha)$. For $\varepsilon \geqslant 0$ we write

$$\gamma_\varepsilon = \frac{\beta + \varepsilon}{2}.$$

Definition 6.9. Let $\mathbb{X} \in \mathcal{X}_{\mathrm{rbe}}$ and $\delta > 1 - 2\alpha$. We define the space $\mathscr{D}_{\mathrm{rbe}}^{\mathrm{sing},\delta} = \mathscr{D}_{\mathrm{rbe},\mathbb{X}}^{\mathrm{sing},\delta}$ of distributions paracontrolled by $\mathbb{X}$ as the set of all $(u, u', u^\sharp) \in C\mathscr{C}^\beta \times \mathscr{L}_\infty^{\gamma_\delta,\delta} \times \mathscr{L}_\infty^{\gamma_{\alpha+\delta},(\alpha+\delta)}$ such that

$$u = X + X^{\vee} + 2X^{\curlyvee} + u' \prec\!\!\prec Q + u^\sharp.$$

For $\delta = \alpha$ we will usually write $\mathscr{D}_{\mathrm{rbe}}^{\mathrm{sing}} = \mathscr{D}_{\mathrm{rbe}}^{\mathrm{sing},\alpha}$. For $T > 0$ we set $\mathscr{D}_{\mathrm{rbe}}^{\mathrm{sing},\delta}(T) = \mathscr{D}_{\mathrm{rbe}}^{\mathrm{sing},\delta}|_{[0,T]}$, and we define

$$\| u \|_{\mathscr{D}_{\mathrm{rbe}}^{\mathrm{sing},\delta}(T)} = \| u' \|_{\mathscr{L}_\infty^{\gamma_\delta,\delta}(T)} + \| u^\sharp \|_{\mathscr{L}_\infty^{\gamma_{\alpha+\delta},(\alpha+\delta)}(T)}.$$

We will often use the notation $u^Q = u' \prec\!\!\prec Q + u^\sharp$, and we will also write $(u', u^\sharp) \in \mathscr{D}_{\mathrm{rbe}}^{\mathrm{sing}}$ or $u \in \mathscr{D}_{\mathrm{rbe}}^{\mathrm{sing}}$. We call u' the *derivative* and $u^\sharp$ the *remainder*.

The sing in $\mathscr{D}_{\mathrm{rbe}}^{\mathrm{sing}}$ stands for “singular”, by which we mean that the paracontrolled norm of $u(t)$ is allowed to blow up as t approaches 0. Throughout this section we work under the following assumption:

Assumption (T,M)**.** *Assume that $\theta, \tilde{\theta} \in \mathscr{L}C(\mathbb{R}, C^\infty(\mathbb{T}))$ and $u_0, \tilde{u}_0 \in \mathscr{C}^{-\beta}$, and that u is the unique global in time solution to the Burgers equation*

$$\mathscr{L} u = \mathrm{D} u^2 + \mathrm{D}\theta, \qquad u(0) = u_0. \tag{51}$$

We define $\mathbb{X} = \Theta_{\mathrm{rbe}}(\theta)$, $u^Q = u - X - X^{\vee} - 2X^{\curlyvee}$, and $u' = 2u^Q + 4X^{\curlyvee}$, and we set $u^\sharp = u^Q - u' \prec\!\!\prec Q$. Similarly we define $\tilde{\mathbb{X}}, \tilde{u}, \tilde{u}^Q, \tilde{u}', \tilde{u}^\sharp$. Finally we assume that $T, M > 0$ are such that

$$\max \{ \| u_0 \|_{-\beta}, \| \tilde{u}_0 \|_{-\beta}, \| \mathbb{X} \|_{\mathcal{X}_{\mathrm{rbe}}(T)}, \| \tilde{\mathbb{X}} \|_{\mathcal{X}_{\mathrm{rbe}}(T)} \} \leqslant M.$$

From the paracontrolled ansatz we derive the following equation for $u^\sharp$:

$$\mathscr{L}u^\sharp = \mathrm{D}u^2 - \mathrm{D}X^2 - 2\mathrm{D}X^{\vee}X - \mathscr{L}(u' \prec\!\!\prec Q), \qquad u^\sharp(0) = u_0 - X(0) - X^{\vee}(0) - 2X^{\vee\!\!\!\vee}(0),$$

where for the initial condition we used that $u' \prec\!\!\prec Q(0) = 0$. Using similar arguments as for Lemma 3.4, we deduce the following result.

Lemma 6.10. *Under Assumption (T,M) we have*

$$\|(\mathscr{L}u^\sharp - \mathscr{L}X^{\vee\vee} - 4\mathscr{L}X^{\vee\!\!\!\vee\!\!\!\vee})(t)\|_{2\alpha-2} \lesssim t^{-\gamma_\beta}(1+M^2)(1+\|u\|^2_{\mathscr{D}^{\mathrm{sing}}_{\mathrm{rbe}}(T)}).$$

for all $t \in (0,T]$. If further also $\|u\|_{\mathscr{D}^{\mathrm{sing}}_{\mathrm{rbe},\mathbb{X}}(T)}, \|\tilde{u}\|_{\mathscr{D}^{\mathrm{sing}}_{\mathrm{rbe},\tilde{\mathbb{X}}}(T)} \leqslant M$, then

$$\begin{aligned} &\|((\mathscr{L}u^\sharp - \mathscr{L}X^{\vee\vee} - 4\mathscr{L}X^{\vee\!\!\!\vee\!\!\!\vee}) - (\mathscr{L}\tilde{u}^\sharp - \mathscr{L}\tilde{X}^{\vee\vee} - 4\mathscr{L}\tilde{X}^{\vee\!\!\!\vee\!\!\!\vee}))(t)\|_{2\alpha-2} \\ &\qquad \lesssim t^{-\gamma_\beta}(1+M^2)(d_{\mathcal{X}_{\mathrm{rbe}}(T)}(\mathbb{X},\tilde{\mathbb{X}}) + \|u'-\tilde{u}'\|_{\mathscr{L}^{\gamma_\alpha,\alpha}_\infty(T)} + \|u^\sharp - \tilde{u}^\sharp\|_{\mathscr{L}^{\gamma_{2\alpha},2\alpha}_\infty(T)}). \end{aligned} \tag{52}$$

Sketch of Proof. We use the same decomposition for $(\mathscr{L}u^\sharp - \mathscr{L}X^{\vee\vee} - \mathscr{L}X^{\vee\!\!\!\vee\!\!\!\vee})(t)$ as in the proof of Lemma 3.4. The most tricky term to bound is

$$\|\mathrm{D}(u^Q(t))^2\|_{2\alpha-2} \lesssim \|(u^Q(t))^2\|_{2\alpha-1} \lesssim \|(u^Q(t))\|^2_{L^\infty} \lesssim (t^{-\gamma_0}\|u^Q\|_{\mathscr{M}^{\gamma_0}_T L^\infty})^2.$$

Now $2\gamma_0 = \gamma_\beta$ and Lemma 6.4 and Lemma 6.8 yield

$$\begin{aligned} \|u^Q\|_{\mathscr{M}^{\gamma_0}_T L^\infty} &\leqslant \|u' \prec\!\!\prec Q\|_{\mathscr{M}^{\gamma_0}_T L^\infty} + \|u^\sharp\|_{\mathscr{M}^{\gamma_0}_T L^\infty} \lesssim \|u'\|_{\mathscr{M}^{\gamma_0}_T L^\infty}\|Q\|_{C_T\mathscr{C}^\alpha} + \|u^\sharp\|_{\mathscr{L}^{\gamma_{2\alpha},2\alpha}_\infty(T)} \\ &\lesssim (\|u'\|_{\mathscr{L}^{\gamma_\alpha,\alpha}_\infty(T)} + \|u^\sharp\|_{\mathscr{L}^{\gamma_{2\alpha},2\alpha}_\infty(T)})(1+\|\mathbb{X}\|_{\mathcal{X}_{\mathrm{rbe}}(T)}). \end{aligned}$$

Every other term in the decomposition of $(\mathscr{L}u^\sharp - \mathscr{L}X^{\vee\vee} - 4\mathscr{L}X^{\vee\!\!\!\vee\!\!\!\vee})(t)$ that does not explicitly depend on $u^\sharp$ can be estimated with a factor $t^{-\gamma_\alpha}$ and since $2\gamma_0 = \gamma_\beta$ and $\alpha < \beta$ we get $t^{-\gamma_\alpha} \lesssim_T t^{-2\gamma_0}$ for all $t \in [0,T]$. The last remaining term is then

$$\|\mathrm{D}(u^\sharp \circ X(t))\|_{\mathscr{C}^{\alpha+\beta-2}} \lesssim t^{-\gamma_\beta}\|u^\sharp\|_{\mathscr{L}^{\gamma_\beta,\beta}_\infty(T)}\|X(t)\|_{\alpha-1} \lesssim t^{-\gamma_\beta}\|u^\sharp\|_{\mathscr{L}^{\gamma_{2\alpha},2\alpha}_\infty(T)}\|\mathbb{X}\|_{\mathcal{X}_{\mathrm{rbe}}(T)},$$

where we used that $\alpha + \beta > 1$ and that $2\alpha > \beta$. □

Corollary 6.11. *Under Assumption (T,M), we have*

$$\|u\|_{\mathscr{D}^{\mathrm{sing}}_{\mathrm{rbe}}(T)} \lesssim M + T^{\gamma_{2\alpha}-\gamma_\beta}(1+M^2)(1+\|u\|^2_{\mathscr{D}^{\mathrm{sing}}_{\mathrm{rbe}}(T)})$$

If further also $\|u\|_{\mathscr{D}^{\exp}_{\mathrm{rbe},\mathbb{X}}(T)}, \|\tilde{u}\|_{\mathscr{D}^{\exp}_{\mathrm{rbe},\tilde{\mathbb{X}}}(T)} \leqslant M$, then

$$\begin{aligned} \|u^\sharp - \tilde{u}^\sharp\|_{\mathscr{L}^{\gamma_{2\alpha},2\alpha}_\infty(T)} + \|u'-\tilde{u}'\|_{\mathscr{L}^{\gamma_\alpha,\alpha}_\infty(T)} &\lesssim \|u_0 - \tilde{u}_0\|_{-\beta} + d_{\mathcal{X}_{\mathrm{rbe}}(T)}(\mathbb{X},\tilde{\mathbb{X}}) \\ &\quad + T^{\gamma_{2\alpha}-\gamma_\beta}(1+M^2)(d_{\mathcal{X}_{\mathrm{rbe}}(T)}(\mathbb{X},\tilde{\mathbb{X}}) + \|u'-\tilde{u}'\|_{\mathscr{L}^{\gamma_\alpha,\alpha}_\infty(T)} + \|u^\sharp - \tilde{u}^\sharp\|_{\mathscr{L}^{\gamma_{2\alpha},2\alpha}_\infty(T)}). \end{aligned}$$

Proof. Lemma 6.10 and Lemma 6.6 yield

$$\|u^\sharp\|_{\mathscr{L}_\infty^{\gamma_{2\alpha},2\alpha}(T)} \lesssim \|u_0\|_{-\beta} + \|\mathbb{X}\|_{\mathcal{X}_{\mathrm{rbe}}(T)} + T^{\gamma_{2\alpha}-\gamma_\beta}(1+M^2)(1+\|u\|^2_{\mathscr{D}^{\mathrm{sing}}_{\mathrm{rbe}}(T)})$$

and similarly we derive the bound for $u^\sharp - \tilde{u}^\sharp$. It remains to control u' and $u' - \tilde{u}'$. We have

$$\begin{aligned}\|u'\|_{\mathscr{L}_\infty^{\gamma_\alpha,\alpha}(T)} &\leqslant 2\|u^Q\|_{\mathscr{L}_\infty^{\gamma_\alpha,\alpha}(T)} + 4\|X^{\curlyvee}\|_{\mathscr{L}_\infty^{\gamma_\alpha,\alpha}(T)} \\ &\lesssim \|u' \lll Q\|_{\mathscr{L}_\infty^{\gamma_\alpha,\alpha}(T)} + \|u^\sharp\|_{\mathscr{L}_\infty^{\gamma_{2\alpha},2\alpha}(T)} + \|\mathbb{X}\|_{\mathcal{X}_{\mathrm{rbe}}(T)},\end{aligned}$$

where we used Lemma 6.8 to bound $\|u^\sharp\|_{\mathscr{L}_\infty^{\gamma_\alpha,\alpha}(T)} \lesssim \|u^\sharp\|_{\mathscr{L}_\infty^{\gamma_{2\alpha},2\alpha}(T)}$. Now Lemma 6.7 and then once more Lemma 6.8 yield

$$\begin{aligned}\|u' \lll Q\|_{\mathscr{L}_\infty^{\gamma_\alpha,\alpha}(T)} &\lesssim T^{\gamma_\alpha-\gamma_{\beta-\alpha}}\|u' \lll Q\|_{\mathscr{L}_\infty^{\gamma_{\beta-\alpha},\alpha}(T)} \\ &\lesssim T^{\gamma_\alpha-\gamma_{\beta-\alpha}}\|u'\|_{\mathscr{L}_\infty^{\gamma_{\beta-\alpha},\beta-\alpha}(T)}\|\mathbb{X}\|_{\mathcal{X}_{\mathrm{rbe}}(T)} \\ &\lesssim T^{\gamma_{2\alpha}-\gamma_\beta}\|u'\|_{\mathscr{L}_\infty^{\gamma_\alpha,\alpha}(T)}\|\mathbb{X}\|_{\mathcal{X}_{\mathrm{rbe}}(T)}.\end{aligned}$$

Combining this with our estimate for $\|u^\sharp\|_{\mathscr{L}_\infty^{\gamma_{2\alpha},2\alpha}(T)}$, we obtain the bound for $\|u'\|_{\mathscr{L}_\infty^{\gamma_\alpha,\alpha}(T)}$. The same arguments allow us to treat $\|u' - \tilde{u}'\|_{\mathscr{L}_\infty^{\gamma_\alpha,\alpha}(T)}$. □

The main result of this section now immediately follows. Before we state it, let us recall that the distance $d_{\mathcal{X}_{\mathrm{rbe}}}$ was defined in (32).

Theorem 6.12. *For every $(\mathbb{X}, u_0) \in \mathcal{X}_{\mathrm{rbe}} \times \mathscr{C}^{-\beta}$ there exists $T^* \in (0,\infty]$ such that for all $T < T^*$ there is a unique solution $(u, u', u^\sharp) \in \mathscr{D}^{\mathrm{sing}}_{\mathrm{rbe}}(T)$ to the rough Burgers equation (51) on the interval $[0,T]$. The map that sends $(\mathbb{X}, u_0) \in \mathcal{X}_{\mathrm{rbe}} \times \mathscr{C}^{-\beta}$ to the solution $(u, u', u^\sharp) \in \mathscr{D}^{\mathrm{sing}}_{\mathrm{rbe}}$ is "continuous" in the sense of Theorem 3.7, and we can choose*

$$T^* = \sup\{t \geqslant 0 : \|u\|_{C_t\mathscr{C}^{-\beta}} < \infty\}. \tag{53}$$

Proof. The proof is essentially the same as the one for Theorem 3.5. The only difference is in the iteration argument. Assume that we constructed a paracontrolled solution u on $[0,\tau]$ for some $\tau > 0$. Then $\tau^{\gamma_{2\alpha}}(u^Q(\tau) - u' \lll Q(\tau)) \in \mathscr{C}^{2\alpha}$ by assumption, and Lemma 6.5 shows that

$$\tau^{\gamma_{2\alpha}}(u' \lll Q(\tau) - u'(\tau) \prec Q(\tau)) \in \mathscr{C}^{2\alpha}. \tag{54}$$

The initial condition for $u^\sharp$ in the iteration on $[\tau, \tau+1]$ is given by $u^Q(\tau) - u'(\tau) \prec Q(\tau)$, which as we just saw is in $\mathscr{C}^{2\alpha}$, and therefore we can use the results of Section 3.4 to construct a paracontrolled solution $(\tilde{u}, \tilde{u}', \tilde{u}^\sharp) \in C_{\tilde{\tau}}\mathscr{C}^{\alpha-1} \times \mathscr{L}^\alpha_{\tilde{\tau}} \times \mathscr{L}^{2\alpha}_{\tilde{\tau}}$ on the time interval $[0,\tilde{\tau}]$ for some $\tilde{\tau} > 0$. Extend now u and u' from $[0,\tau]$ to $[0,\tau+\tilde{\tau}]$ by setting $u(t) = \tilde{u}(t-\tau)$ for $t \geqslant \tau$ and similarly for u'. Since on $[\tau, \tau+\tilde{\tau}]$ the function $t \mapsto t^{\gamma_\alpha}$ is infinitely differentiable, we obtain that $(u, u') \in C_{\tau+\tilde{\tau}}\mathscr{C}^{\beta} \times \mathscr{L}^{\gamma_\alpha;\alpha}_{\tau+\tilde{\tau}}$. Moreover, $u^Q - u' \prec\!\!\prec_\tau Q \in \mathscr{L}^{2\alpha}$ on the interval $[\tau, \tau+\tilde{\tau}]$, where

$$u' \prec\!\!\prec_\tau Q(t) = \sum_j \int_{-\infty}^{t} 2^{2j}\varphi(2^{2j}(t-s))S_{j-1}u'(s\vee\tau)\mathrm{d}s\Delta_j Q(t).$$

Using the smoothness of $t \mapsto t^{\gamma_{2\alpha}}$ on $[\tau, \tau+\tilde{\tau}]$, it therefore suffices to show that $(u' \prec\!\!\prec_\tau Q - u' \prec\!\!\prec Q)|_{[\tau,\tau+\tilde{\tau}]} \in \mathscr{L}^{2\alpha}$. But we already saw in equation (54) above that $(u' \prec\!\!\prec_\tau Q - u' \prec\!\!\prec Q)(\tau) \in \mathscr{C}^{2\alpha}$, and for $t \in [\tau, \tau+\tilde{\tau}]$ we get from the second estimate of Lemma 6.5 that

$$\mathscr{L}(u' \prec\!\!\prec_\tau Q - u' \prec\!\!\prec Q)(t) = (u' \prec\!\!\prec_\tau \mathrm{D}X - u' \prec\!\!\prec \mathrm{D}X)(t) + t^{-\gamma_\alpha}\mathscr{C}^{2\alpha-2}.$$

On $[\tau, \tau+\tilde{\tau}]$ the factor $t^{-\gamma_\alpha}$ poses no problem and therefore it suffices to control the difference between the two modified paraproducts. Now the first estimate of Lemma 6.5 shows that $(u' \prec\!\!\prec \mathrm{D}X - u' \prec \mathrm{D}X)|_{[\tau,\tau+\tilde{\tau}]} \in C\mathscr{C}^{2\alpha-2}$, and from Lemma 2.8 it follows that also $(u' \prec\!\!\prec_\tau \mathrm{D}X - u' \prec \mathrm{D}X)|_{[\tau,\tau+\tilde{\tau}]} \in C\mathscr{C}^{2\alpha-2}$, so that the $\mathscr{L}^{2\alpha}$ regularity of $u^\sharp|_{[\tau,\tau+\tilde{\tau}]}$ follows from the Schauder estimates for the heat flow, Lemma 2.9. Uniqueness and continuous dependence on the data can be handled along the same lines.

However, so far the construction only works up to time $\tilde{T}^* = \sup\{t \geqslant 0 : \|u\|_{\mathscr{D}^{\mathrm{sing}}_{\mathrm{rbe}}(t)} < \infty\}$, and it remains to show that $\tilde{T}^* = T^*$ for T^* defined in (55). Assume therefore that $\|u\|_{C_{\tilde{T}^*}\mathscr{C}^{-\beta}} < \infty$. Then we can solve the equation starting in $u(\tilde{T}^* - \varepsilon)$ for some very small ε for which we can perform the Picard iteration on an interval $[\tilde{T}^* - \varepsilon, \tilde{T}^* + \varepsilon]$, and in particular the solution $(u, u', u^\sharp)$ restricted to the time interval $[\tilde{T}^* - \varepsilon/2, \tilde{T}^* + \varepsilon/2]$ is in $C([\tilde{T}^* - \varepsilon/2, \tilde{T}^* + \varepsilon/2], \mathscr{C}^{\alpha-1}) \times \mathscr{L}^{\alpha}([\tilde{T}^* - \varepsilon/2, \tilde{T}^* + \varepsilon/2]) \times \mathscr{L}^{2\alpha}([\tilde{T}^* - \varepsilon/2, \tilde{T}^* + \varepsilon/2])$ (with the natural interpretation for the $\mathscr{L}^{\delta}([a,b])$ spaces). Using the uniqueness of the solution we get a contradiction to $\|u\|_{\mathscr{D}^{\mathrm{sing}}_{\mathrm{rbe}}(\tilde{T}^*)} = \infty$, and the proof is complete. □

For the KPZ equation and $Y \in \mathcal{Y}_{\mathrm{kpz}}$ we introduce analogous spaces $\mathscr{D}^{\mathrm{sing}}_{\mathrm{kpz},\mathbb{Y}}$ with distance $d^{\mathrm{sing}}_{\mathrm{kpz}}(h, \tilde{h})$ and obtain the analogous result:

Theorem 6.13. *For every $(\mathbb{Y}, h_0) \in \mathcal{Y}_{\mathrm{kpz}} \times \mathscr{C}^{1-\beta}$ there exists $T^* \in (0, \infty]$ such that for all $T < T^*$ there is a unique solution $(h, h', h^\sharp) \in \mathscr{D}^{\mathrm{sing}}_{\mathrm{kpz}}(T)$ to the rough KPZ equation*

$$\mathscr{L}h = (\mathrm{D}h)^{\diamond 2} + \theta, \qquad h(0) = h_0$$

on the interval $[0, T]$. The map that sends $(\mathbb{Y}, h_0) \in \mathcal{X}_{\mathrm{rbe}} \times \mathscr{C}^{1-\beta}$ to the solution $(h, h', h^\sharp) \in \mathscr{D}^{\mathrm{sing}}_{\mathrm{rbe}}$ is "continuous" in the sense of Theorem 3.7, and we can choose

$$T^* = \sup\{t \geqslant 0 : \|u\|_{C_t\mathscr{C}^{1-\beta}} < \infty\}. \tag{55}$$

6.3 Heat equation with singular initial conditions

The rough heat equation is linear, and therefore it sometimes turns out to be advantageous to solve it in $\mathscr{C}^{\alpha}_p$ spaces for general p in order to allow for more general initial conditions. More precisely, we will be able to handle initial conditions $w_0 \in \mathscr{C}_p^{-\beta}$ for arbitrary $\beta < 1/2$ and $p \in [1, \infty]$. Taking $p = 1$, this allows us for example to start in $(-\Delta)^\gamma \delta$ for $\gamma < 1/4$, where δ denotes the Dirac delta.

We fix $p \in [1, \infty]$, $\alpha \in (1/3, 1/2)$, and $\beta \in (0, \alpha)$, and we want to solve the paracontrolled equation

$$\mathscr{L} w = w \diamond \xi, \qquad w(0) = w_0,$$

for $w_0 \in \mathscr{C}^{-\beta}$. For $\varepsilon \geqslant 0$ let us write

$$\gamma_\varepsilon = \frac{\beta + \varepsilon}{2}.$$

Definition 6.14. Let $\mathbb{Y} \in \mathcal{Y}_{\mathrm{kpz}}$ and $\delta \in (1 - 2\alpha, 1 - \alpha - \beta)$. We define the space $\mathscr{D}_{\mathrm{rhe}}^{\mathrm{sing},\delta} = \mathscr{D}_{\mathrm{rhe},\mathbb{Y}}^{\mathrm{sing},\delta}$ of distributions paracontrolled by $\mathbb{Y}$ as the set of all $(u, u', u^\sharp) \in C\mathscr{C}_p^\beta \times \mathscr{L}_p^{\gamma_\delta,\delta} \times \mathscr{L}_p^{\gamma_{1+\alpha+\delta},(1+\alpha+\delta)}$ such that

$$u = e^{Y + Y^{\vee} + 2Y^{\curlyvee}} (w' \prec\!\!\prec P + w^\sharp).$$

For $T > 0$ we set $\mathscr{D}_{\mathrm{rhe}}^{\mathrm{sing},\delta}(T) = \mathscr{D}_{\mathrm{rhe}}^{\mathrm{sing},\delta}|_{[0,T]}$, and we define

$$\|w\|_{\mathscr{D}_{\mathrm{rhe}}^{\mathrm{sing},\delta}(T)} = \|w'\|_{\mathscr{L}_p^{\gamma_\delta,\delta}(T)} + \|w^\sharp\|_{\mathscr{L}_p^{\gamma_{1+\alpha+\delta},(1+\alpha+\delta)}(T)}.$$

We will often write $w^P = w' \prec\!\!\prec P + w^\sharp$.

Recall from (39) the definition of the renormalized product $w \diamond \xi$ for $\mathbb{Y} = \Theta_{\mathrm{kpz}}(\xi, c^{\vee}, c^{\vee\vee})$:

$$\begin{aligned} w \diamond \xi = \mathscr{L} w - e^{Y + Y^{\vee} + 2Y^{\curlyvee}} \Big[& - [4(\mathscr{L} Y^{\Ydown} + \mathrm{D}Y^{\curlyvee} \prec \mathrm{D}Y + \mathrm{D}Y^{\curlyvee} \succ \mathrm{D}Y) + \mathscr{L} Y^{\vee\vee}] w^p \\ & + [4\mathrm{D}Y^{\vee}\mathrm{D}Y^{\curlyvee} + (2\mathrm{D}Y^{\curlyvee})^2] w^P \\ & + \mathscr{L} w^P - 2\mathrm{D}(Y + Y^{\vee} + 2Y^{\curlyvee})\mathrm{D}w^P \Big]. \end{aligned}$$

It is then a simple exercise to apply the results of Section 6.1 to see that $w \diamond \xi$ depends continuously on $(\mathbb{Y}, w) \in \mathcal{Y}_{kpz} \times \mathscr{D}_{\mathrm{rhe},\mathbb{Y}}^{\mathrm{sing},\delta}$. Moreover, $w \in \mathscr{D}_{\mathrm{rhe},\mathbb{Y}}^{\mathrm{sing},\delta}$ solves

$$\mathscr{L} w = w \diamond \xi, \qquad w(0) = w_0,$$

if and only if

$$\begin{aligned} \mathscr{L} w^P = {} & [4(\mathscr{L} Y^{\Ydown} + \mathrm{D}Y^{\curlyvee} \prec \mathrm{D}Y + \mathrm{D}Y^{\curlyvee} \succ \mathrm{D}Y) + \mathscr{L} Y^{\vee\vee} + 4\mathrm{D}Y^{\vee}\mathrm{D}Y^{\curlyvee} + (2\mathrm{D}Y^{\curlyvee})^2] w^p \\ & + 2\mathrm{D}(Y + Y^{\vee} + 2Y^{\curlyvee})\mathrm{D}w^P, \end{aligned}$$

$w^P(0) = w_0 e^{-Y(0) - Y^{\vee}(0) - 2Y^{\curlyvee}(0)}$. Since $w_0 \in \mathscr{C}_p^{-\beta}$ and $e^{-Y(0) - Y^{\vee}(0) - 2Y^{\curlyvee}(0)} \in \mathscr{C}^\alpha$ with $\alpha > \beta$, the latter product is well defined and in $\mathscr{C}_p^{-\beta}$.

We define the distance $d_{\mathrm{rhe}}^{\mathrm{sing}}$ analogously to the case of the KPZ or Burgers equation. Solutions to the rough heat equation can now be constructed using the same arguments as in the previous section, so that we end up with the following result:

Theorem 6.15. *Let* $\delta \in (1-2\alpha, 1-\alpha-\beta)$*. For all* $(\mathbb{Y}, w_0, T) \in \mathcal{Y}_{\mathrm{kpz}} \times \mathscr{C}_p^{-\beta} \times [0,\infty)$ *there is a unique solution* $(w, w', w^\sharp) \in \mathscr{D}^{\mathrm{sing},\delta}_{\mathrm{rhe},\mathbb{Y}}(T)$ *to the rough heat equation* (51) *on the interval* $[0,T]$*. The solution depends "continuously" on* $(\mathbb{Y}, w_0)$ *in the sense of Theorem 3.7. Moreover, if there exist* $(w_0^n) \subset C^\infty$ *with* $\|w_0 - w_0^n\|_{-\beta} \to 0$ *as* $n\to\infty$ *and such that* $w_0^n \geqslant 0$ *for all* n*, then* $w(t,x) \geqslant 0$ *for all* (t,x)*.*

7 Variational representation and global existence of solutions

7.1 KPZ as Hamilton-Jacobi-Bellman equation

Here we show that for every $\mathbb{Y} \in \mathcal{Y}_{\mathrm{rbe}}$ and every $h_0 \in \mathscr{C}^\beta$ for $\beta > 0$ there are global in time solutions to the KPZ equation. The idea is to interpret the solution as value function of an optimal control problem, and to "guess" an expansion of the optimal control. This can be made rigorous in the case of smooth data and allows us to obtain a priori bounds which show that no explosion occurs as we let the smooth data converge to $\mathbb{Y}$.

Let h solve the KPZ equation

$$\mathscr{L}h = (\mathrm{D}h)^2 - c^{\mathsf{V}} + \theta, \qquad h(0) = \bar{h}, \tag{56}$$

for $\theta \in \mathscr{L}C^{\alpha/2}_{\mathrm{loc}}(\mathbb{R}, C^\infty)$, $c^{\mathsf{V}} \in \mathbb{R}$, and $\bar{h} \in C^\infty$. Then by the Cole–Hopf transform we have $h = \log w$, where

$$\mathscr{L}w = w(\theta - c^{\mathsf{V}}), \qquad w(0) = e^{\bar{h}}. \tag{57}$$

We specified in Section 4.2 how to interpret this equation. But as long as Y (see Definition 4.1) is in $C(\mathbb{R}_+, C^\infty)$, we have the following simpler characterization. The function $w\colon \mathbb{R}_+ \times \mathbb{T} \to \mathbb{R}$ solves (57) in the sense of Section 4.2 if and only if $w = e^Y w^1$, where w^1 is a classical solution to

$$\mathscr{L}w^1 = (|\mathrm{D}Y|^2 - c^{\mathsf{V}})w^1 + 2\mathrm{D}Y\mathrm{D}w^1, \qquad w^1(0) = e^{\bar{h} - Y(0)}.$$

For the rest of this section it will be convenient to reverse time. So fix $T > 0$ and define $\overleftarrow{\varphi} = \varphi(T-t)$ for all appropriate φ. Then

$$(\partial_t + \Delta)\overleftarrow{h} = -((\mathrm{D}\overleftarrow{h})^2 - c^{\mathsf{V}}) - \overleftarrow{\theta}, \qquad \overleftarrow{h}(T) = \bar{h},$$

and $\overleftarrow{w} = e^{\overleftarrow{Y}}\overleftarrow{w}^1$ with

$$(\partial_t + \Delta + 2\mathrm{D}\overleftarrow{Y}\mathrm{D})\overleftarrow{w}^1 = -(|\mathrm{D}\overleftarrow{Y}|^2 - c^{\mathsf{V}})\overleftarrow{w}^1, \qquad \overleftarrow{w}^1(T) = e^{\bar{h} - \overleftarrow{Y}(T)}.$$

The Feynman–Kac formula ([KS88], Theorem 5.7.6) shows that for $t \in [0,T]$ we have

$$\overleftarrow{w}^1(t,x) = \mathbb{E}_{t,x}\Big[e^{\bar{h}(\gamma_T) - \overleftarrow{Y}(T,\gamma_T) + \int_t^T (|\mathrm{D}\overleftarrow{Y}|^2(s,\gamma_s) - c^{\mathsf{V}})\mathrm{d}s}\Big], \tag{58}$$

where under $\mathbb{P}_{t,x}$ the process γ solves

$$\gamma_s = \int_t^s 2\mathrm{D}\overleftarrow{Y}(r,\gamma_r)\mathrm{d}r + B_s, \quad s \geqslant t,$$

with a Brownian motion B with variance 2 (that is $\mathrm{d}\langle B\rangle_s = 2\mathrm{d}s$), started in $B_t = x$. An application of Girsanov's theorem gives

$$\begin{aligned}\overleftarrow{w}(t,x) &= e^{\overleftarrow{Y}(t,x)}\mathbb{E}_{t,x}\Big[e^{\bar{h}(B_T)-\overleftarrow{Y}(T,B_T)+\int_t^T(|\mathrm{D}\overleftarrow{Y}|^2(s,B_s)-c^{\mathsf{V}})\mathrm{d}s}e^{\int_t^T \mathrm{D}\overleftarrow{Y}(s,B_s)\mathrm{d}B_s-\int_t^T|\mathrm{D}\overleftarrow{Y}|^2(s,B_s)\mathrm{d}s}\Big] \\ &= \mathbb{E}_{t,x}\Big[e^{\bar{h}(B_T)-(\overleftarrow{Y}(T,B_T)-\overleftarrow{Y}(t,x)-\int_t^T \mathrm{D}\overleftarrow{Y}(s,B_s)\mathrm{d}B_s)-c^{\mathsf{V}}(T-t)}\Big].\end{aligned}$$

If we formally apply Itô's formula, we get

$$\overleftarrow{Y}(T,B_T)-\overleftarrow{Y}(t,x)-\int_t^T \mathrm{D}\overleftarrow{Y}(s,B_s)\mathrm{d}B_s = \int_t^T(\partial_s+\Delta)\overleftarrow{Y}(s,B_s)\mathrm{d}s = -\int_t^T \overleftarrow{\theta}(s,B_s)\mathrm{d}s,$$

and we simply take the left hand side as the definition of the right hand side so that we can write

$$\overleftarrow{w}(t,x) = \mathbb{E}_{t,x}[e^{\bar{h}(B_T)-\int_t^T(\overleftarrow{\theta}(s,B_s)-c^{\mathsf{V}})\mathrm{d}s}].$$

We will need the following generalization of the Boué–Dupuis [BD98] formula which has been recently established by Üstünel:

Theorem 7.1 ([Üst14], Theorem 7)**.** *Let $B\colon[0,T]\to\mathbb{R}^d$ be a Brownian motion with variance σ^2 and let F be a measurable functional on $C([0,T];\mathbb{R}^d)$ such that $F(B)\in L^2$ and $e^{-F(B)}\in L^2$. Then*

$$-\log\mathbb{E}[e^{-F(B)}] = \inf_v \mathbb{E}\Big[F\Big(B+\int_0^{\cdot} v_s\mathrm{d}s\Big)+\frac{1}{2\sigma^2}\int_0^T|v_s|^2\mathrm{d}s\Big],$$

where the infimum runs over all processes v that are progressively measurable with respect to $(\mathcal{F}_t)$, the augmented filtration generated by B, and that are such that $\omega\mapsto v_s(\omega)$ is $\mathcal{F}_s$–measurable for Lebesgue–almost all $s\in[0,T]$.

Corollary 7.2. *Let h solve* (56)*, let $T>0$, and let B be a Brownian motion with variance $\langle B\rangle_t = 2t$, started in $B_0=0$. Then*

$$\begin{aligned}h(T,x) = \log\overleftarrow{w}(0,x) &= \log\mathbb{E}\Big[\exp\Big(\bar{h}(x+B_T)+\int_0^T(\overleftarrow{\theta}(s,x+B_s)-c^{\mathsf{V}})\mathrm{d}s\Big)\Big] \\ &= -\inf_v\mathbb{E}\left[-\bar{h}(\gamma_T^v)-\int_0^T(\overleftarrow{\theta}(s,\gamma_s^v)-c^{\mathsf{V}})\mathrm{d}s+\frac{1}{4}\int_0^T|v_s|^2\mathrm{d}s\right] \\ &= \sup_v\mathbb{E}\left[\bar{h}(\gamma_T^v)+\int_0^T\Big(\overleftarrow{\theta}(s,\gamma_s^v)-c^{\mathsf{V}}-\frac{1}{4}|v_s|^2\Big)\mathrm{d}s\right], \qquad (59)\end{aligned}$$

where we wrote $\gamma_t^v = x+B_t+\int_0^t v_s\mathrm{d}s$ and the supremum runs over the same processes v as in Theorem 7.1.

Define now for any γ of the form $\mathrm{d}\gamma_t = v_t \mathrm{d}t + \mathrm{d}B_t$ the payoff functional

$$
\begin{aligned}
\Phi(\gamma, v) &= \bar{h}(\gamma_T) + \int_0^T \Big(\overleftarrow{\theta}(s, \gamma_s) - c^{\vee} - \frac{1}{4}|v_s|^2 \Big) \mathrm{d}s \\
&:= \bar{h}(\gamma_T) - \overleftarrow{Y}(T, \gamma_T) + \overleftarrow{Y}(0, x) + \int_0^T \mathrm{D}\overleftarrow{Y}(s, \gamma_s)\mathrm{d}\gamma_s - \int_0^T (c^{\vee} + \frac{1}{4}|v_s|^2)\mathrm{d}s,
\end{aligned}
$$

so that $h(T, x) = \sup_v \mathbb{E}[\Phi(\gamma^v, v)]$. Plugging in $\mathrm{d}\gamma_t^v = v_t \mathrm{d}t + \mathrm{d}B_t$, we get

$$
\begin{aligned}
\Phi(\gamma^v, v) &= \bar{h}(\gamma_T^v) - \overleftarrow{Y}(T, \gamma_T^v) + \overleftarrow{Y}(0, x) - \int_0^T \Big(- v\mathrm{D}\overleftarrow{Y} + c^{\vee} + \frac{1}{4}|v|^2 \Big)(s, \gamma_s^v)\mathrm{d}s + \mathrm{mart.} \\
&= \bar{h}(\gamma_T^v) - \overleftarrow{Y}(T, \gamma_T^v) + \overleftarrow{Y}(0, x) - \int_0^T \Big(- |\mathrm{D}\overleftarrow{Y}|^2 + c^{\vee} + \frac{1}{4}|v - 2\mathrm{D}\overleftarrow{Y}|^2 \Big)(s, \gamma_s^v)\mathrm{d}s \\
&\quad + \mathrm{mart.},
\end{aligned}
$$

where we write "mart." for an arbitrary martingale term whose expectation vanishes under $\mathbb{E}$. Now change the optimization variable to $v_t^1 = v_t - 2\mathrm{D}\overleftarrow{Y}(t, \gamma_t^v)$, so that

$$
\gamma_t^v = x + B_t + \int_0^t (v^1 + 2\mathrm{D}\overleftarrow{Y})(s, \gamma_s^v)\mathrm{d}s,
$$

and the payoff becomes

$$
\Phi(\gamma^v, v) = \bar{h}(\gamma_T^v) - \overleftarrow{Y}(T, \gamma_T^v) + \overleftarrow{Y}(0, x) + \int_0^T \Big(|\mathrm{D}\overleftarrow{Y}|^2 - c^{\vee} - \frac{1}{4}|v^1|^2 \Big)(s, \gamma_s^v)\mathrm{d}s + \mathrm{mart}.
$$

In the following denote $X^i = \mathrm{D}Y^i$. We can iterate the process by considering $\overleftarrow{Y}^{\vee}$, where $Y^{\vee}$ is as in Definition 4.1 (i.e. $\overleftarrow{Y}^{\vee}$ solves $(\partial_t + \Delta)\overleftarrow{Y}^{\vee} = -(|\overleftarrow{X}|^2 - c^{\vee})$ with terminal condition $\overleftarrow{Y}^{\vee}(T) = 0$), which allows us to represent the payoff function as

$$
\begin{aligned}
\Phi(\gamma^v, v) &= \bar{h}(\gamma_T^v) - \overleftarrow{Y}(T, \gamma_T^v) + \overleftarrow{Y}(0, x) + \overleftarrow{Y}^{\vee}(0, x) \\
&\quad + \int_0^T \Big(v^1 \overleftarrow{X}^{\vee} + 2\overleftarrow{X}\overleftarrow{X}^{\vee} - \frac{1}{4}|v^1|^2 \Big)(s, \gamma_s^v)\mathrm{d}s + \mathrm{mart}. \\
&= \bar{h}(\gamma_T^v) - \overleftarrow{Y}(T, \gamma_T^v) + \overleftarrow{Y}(0, x) + \overleftarrow{Y}^{\vee}(0, x) \\
&\quad + \int_0^T \Big(|\overleftarrow{X}^{\vee}|^2 + 2\overleftarrow{X}\overleftarrow{X}^{\vee} - \frac{1}{4}|v^1 - 2\overleftarrow{X}^{\vee}|^2 \Big)(s, \gamma_s^v)\mathrm{d}s + \mathrm{mart}.
\end{aligned}
$$

We change the control strategy to $v_t^2 = v_t^1 - 2\overleftarrow{X}^{\vee}(t, \gamma_t^v)$, so that the dynamics of γ^v read $\mathrm{d}\gamma_t^v = \mathrm{d}B_t + (v^2 + 2\overleftarrow{X} + 2\overleftarrow{X}^{\vee})(t, \gamma_t^v)\mathrm{d}t$. Now let

$$
\mathscr{L}Y^R = |X^{\vee}|^2 + 2XX^{\vee} + 2(X + X^{\vee})\mathrm{D}Y^R, \qquad Y^R(0) = 0. \tag{60}
$$

This is a linear paracontrolled equation whose solution is of the form $Y^R = 2Y^{\curlyvee} + Y' \prec\!\!\prec P + Y^\sharp$ and depends continuously on $\mathbb{Y}$. Moreover, another application of Itô's formula yields with $X^R = \mathrm{D}Y^R$

$$\Phi(\gamma^v, v) = \bar{h}(\gamma_T^v) - \overleftarrow{Y}(T, \gamma_T^v) + \overleftarrow{Y}(0,x) + \overleftarrow{Y}^{\vee}(0,x) + \overleftarrow{Y}^{R}(0,x) + \int_0^T \Big(v^2 \overleftarrow{X}^R - \frac{1}{4}|v^2|^2\Big)(s, \gamma_s^v)\mathrm{d}s + \mathrm{mart}\,.$$

Let us summarize the result of our calculation:

Theorem 7.3. *Let* $\theta \in \mathscr{L}C^{\alpha/2}_{\mathrm{loc}}(\mathbb{R}, C^\infty)$ *and* $c^{\vee}, c^{\vee\!\vee} \in \mathbb{R}$, *and let* $\mathbb{Y} = \Theta_{\mathrm{kpz}}(\theta, c^{\vee}, c^{\vee\!\vee}) \in \mathcal{Y}_{\mathrm{kpz}}$ *(see Definition 4.1). Let* h *solve the KPZ equation* (56) *driven by* $\mathbb{Y}$, *started in* $\bar{h} \in L^\infty$, *let* Y^R *solve* (60) *and* $X^R = \mathrm{D}Y^R$, *and let* $T > 0$. *Then*

$$(h - Y - Y^{\vee} - Y^R)(T,x) = \sup_v \mathbb{E}\Big[\bar{h}(\zeta_T^v) - Y(0, \zeta_T^v) + \int_0^T \Big(|\overleftarrow{X}^R|^2 - \frac{1}{4}|v - 2\overleftarrow{X}^R|^2\Big)(s, \zeta_s^v)\mathrm{d}s\Big], \quad (61)$$

where

$$\zeta_t^v = x + \int_0^t (2\overleftarrow{X} + 2\overleftarrow{X}^{\vee} + v)(s, \zeta_s^v)\mathrm{d}s + B_t \quad (62)$$

and the supremum is taken over the same v *as in Theorem 7.1.*

This representation is very useful for deriving a priori bounds on h.

Corollary 7.4. *In the setting of Theorem 7.3, for all* $T > 0$ *there exists a constant* $C > 0$ *depending only on* T *and* $\|\mathbb{Y}\|_{\mathcal{Y}_{\mathrm{kpz}}(T)}$, *such that*

$$\|h\|_{C_T L^\infty} \leqslant C(1 + \|\bar{h}\|_{L^\infty}).$$

In particular, if $(\Theta_{\mathrm{kpz}}(\theta_n, c_n^{\vee}, c_n^{\vee\!\vee}))_n$ *is a converging sequence in* $\mathcal{Y}_{\mathrm{kpz}}$ *and every* θ_n *is as above, then the corresponding solutions* (h_n) *stay uniformly bounded in* $C_T L^\infty$ *for all* $T > 0$.

Proof. First, we have $-\frac{1}{4}|v - 2\overleftarrow{X}^R|^2(s, \zeta_s^v) \leqslant 0$, independently of v, and therefore

$$\sup_v \mathbb{E}\Big[\bar{h}(\zeta_T^v) - Y(0, \zeta_T^v) + \int_0^T \Big(|\overleftarrow{X}^R|^2 - \frac{1}{4}|v - 2\overleftarrow{X}^R|^2\Big)(s, \zeta_s^v)\mathrm{d}s\Big] \leqslant \|\bar{h}\|_{L^\infty} + \|Y(0)\|_{L^\infty} + \|X^R\|^2_{L^2_T L^\infty}. \quad (63)$$

Moreover, choosing the specific control $v(t) \equiv 0$ we get

$$\sup_v \mathbb{E}\Big[\bar{h}(\zeta_T^v) - Y(0, \zeta_T^v) + \int_0^T \Big(|\overleftarrow{X}^R|^2 - \frac{1}{4}|v - 2\overleftarrow{X}^R|^2\Big)(s, \zeta_s^v)\mathrm{d}s\Big] \geqslant -\|\bar{h}\|_{L^\infty} - \|Y(0)\|_{L^\infty}.$$

In combination with Theorem 7.3 and (63), this yields

$$\|h(T)\|_{L^\infty} \leqslant \|(Y + Y^{\mathsf{V}} + Y^R)(T)\|_{L^\infty} + \|Y(0)\|_{L^\infty} + \|\bar{h}\|_{L^\infty} + \|X^R\|^2_{L^2_T L^\infty}, \qquad (64)$$

which gives us a bound on the solution of the KPZ equation which is linear in terms of the data $\mathbb{Y} \in \mathcal{Y}_{\mathrm{kpz}}$, and quadratic in terms of the solution Y^R to a linear paracontrolled equation that can in turn be bounded by the data $\mathbb{Y}$. □

An immediate consequence is the global in time existence of solutions to the KPZ equation:

Corollary 7.5. *In Theorem 4.2, Theorem 4.9 and Theorem 6.13 we have $T^* = \infty$. If in Theorem 3.5 and Theorem 6.12 the initial condition is $u_0 = \mathrm{D}h_0$ for some $h_0 \in \mathscr{C}^{2\alpha+1}$ (respectively $h_0 \in \mathscr{C}^{-\beta+1}$), then also here $T^* = \infty$.*

Proof. To see that $T^* = \infty$ in Theorem 4.2, Theorem 4.9 and in Theorem 3.5 under the condition $u_0 = \mathrm{D}h_0$, it suffices to combine Theorem 7.3 with Corollary 4.10. As for Theorem 6.12, recall from the proof of this theorem that in order to extend a solution $u \in \mathscr{D}^{\mathrm{sing}}_{\mathrm{rbe}}(T)$ from $[0, T]$ to $[0, T']$ with $T' > T$, it suffices to solve Burgers equation on $[0, T' - T]$ with the initial condition $u^Q(T) - u'(T) \prec Q(T) \in \mathscr{C}^{2\alpha}$ for the remainder. Moreover, if $u_0 = \mathrm{D}h_0$, then also $u(T) = \mathrm{D}h(T)$, so that the first part of the theorem tells us that u can be extended from $[0, T]$ to $[0, \infty)$. The same line of reasoning works also for Theorem 6.13. □

Remark 7.6. *The condition $u_0 = \mathrm{D}h_0$ is equivalent to $\int_{\mathbb{T}} u_0(x)\mathrm{d}x = 0$. In case the integral is equal to $c \in \mathbb{R} \setminus \{0\}$ we can consider $\tilde{u} = u - c$ which solves*

$$\mathscr{L}\tilde{u} = \mathrm{D}\tilde{u}^2 + 2c\mathrm{D}\tilde{u} + \mathrm{D}\theta, \qquad \tilde{u}(0) = \mathrm{D}h_0$$

for some h_0. This is a paracontrolled equation which we can solve up to some explosion time, and we have $\tilde{u} = \mathrm{D}\tilde{h}$ for the solution $\tilde{h}$ to

$$\mathscr{L}\tilde{h} = |\mathrm{D}\tilde{h}|^{\diamond 2} + 2c\mathrm{D}\tilde{h} + \mathrm{D}\theta, \qquad \tilde{h}(0) = h_0.$$

The Cole–Hopf transform then shows that $\tilde{h} = \log \tilde{w}$, where

$$\mathscr{L}\tilde{w} = 2c\mathrm{D}w + w \diamond \xi, \qquad \tilde{w}(0) = e^{h_0},$$

and based on these observations we could perform the same analysis as above to show that the explosion time of $\tilde{u}$ (and thus of u) is infinite. We would only have to replace the Brownian motion B by the process $(B_t + 2ct)_{t \geqslant 0}$ which corresponds to the generator $\Delta + 2c\nabla$.

Another simple consequence of (61) is a quantitative comparison result for the KPZ equation.

Lemma 7.7 ("Comparison principle")**.** *In the setting of Theorem 6.13 let $\bar{h}_1, \bar{h}_2 \in \mathscr{C}^{1-\beta}$, and let h_i solve*

$$\mathscr{L} h_i = |\mathrm{D} h_i|^{\diamond 2} + \theta, \qquad h_i(0) = \bar{h}_i, \qquad i = 1, 2.$$

Then

$$h_1(t,x) + \inf_x (\bar{h}_2(x) - \bar{h}_1(x)) \leqslant h_2(t,x) \leqslant h_1(t,x) + \sup_x (\bar{h}_2(x) - \bar{h}_1(x)) \qquad (65)$$

for all (t,x). In particular, $\|h_1 - h_2\|_{C_T L^\infty} \leqslant \|\bar{h}_1 - \bar{h}_2\|_{L^\infty}$ for all $T > 0$.

Proof. Consider regular data $(\mathbb{Y}^n, \bar{h}_1^n, \bar{h}_2^n)$ that converges to $(\mathbb{Y}, \bar{h}_1, \bar{h}_2)$ in $\mathcal{Y}_{\mathrm{kpz}} \times \mathscr{C}^{1-\beta} \times \mathscr{C}^{1-\beta}$, and denote the corresponding solutions by h_1^n and h_2^n respectively. For every n the representation (61) and the decomposition $\bar{h}_n^2 = \bar{h}_n^1 + (\bar{h}_n^2 - \bar{h}_n^1)$ gives

$$h_1^n(t,x) + \inf_x (\bar{h}_2^n(x) - \bar{h}_1^n(x)) \leqslant h_2^n(t,x) \leqslant h_1^n(t,x) + \sup_x (\bar{h}_2^n(x) - \bar{h}_1^n(x)).$$

Letting n tend to infinity, we get (65). □

Remark 7.8. *Corollary 7.5 gives a pathwise proof for the strict positivity of solutions to the rough heat equation started in strictly positive initial data. The classical proof for the stochastic heat equation is due to Mueller* [Mue91]*, whose arguments are rather involved; see also the recent works* [MF14, CK14]*. Compared to these, our proof has the advantage that it does not use the structure of the white noise at all, so that it is applicable in a wide range of scenarios. The disadvantage is that we need to start in strictly positive data, whereas Mueller's result allows to start in nonpositive, nonzero data.*

However, after the completion of this work it was shown by Cannizzaro, Friz and Gassiat in [CFG15] *that even a strong maximum principle holds for the rough heat equation. That is, for all nonpositive, nonzero initial conditions w_0 the solution to the rough heat equation w satisfies $w(t,x) > 0$ for all $t > 0$ and all $x \in \mathbb{T}$, and also their proof does not rely on the structure of the white noise. But one significant advantage of Corollary 7.4 is that it comes with good qualitative bounds on the solution h of the KPZ equation, while the result of* [CFG15] *is rather qualitative.*

7.2 Partial Girsanov transform and random directed polymers

To formulate the optimization problem (61) for non-smooth elements $\mathbb{Y}$ of $\mathcal{Y}_{\mathrm{kpz}}$, we first need to make sense of the diffusion equation (62) for ζ^v in that case. This can be done with the techniques of [DD16] or [CC15], which we will apply with a slight twist. Both [DD16] and [CC15] allow to make sense of diffusions with singular drifts; [DD16] is based on rough path integrals and applies in a one-dimensional setting while [CC15] use paracontrolled distributions to generalize [DD16] to higher dimensions. Here we could of course apply both, but since [CC15] is formulated in the language of paracontrolled distributions it relates more directly to the concepts developed in this paper.

Let us start by formally deriving the dynamics of the coordinate process under the random directed polymer measure. This is the measure given by

$$\mathrm{d}\mathbb{Q}_{T,x} = \exp\Big(\int_0^T (\overleftarrow{\xi}(t, B_t) - \infty)\mathrm{d}t\Big)\mathrm{d}\mathbb{P}_x,$$

where ξ is a space-time white noise (and thus $\overleftarrow{\xi}$ as well), and under $\mathbb{P}_x$ the process B is a Brownian motion started in x and with variance 2. The term $-\infty T$ is chosen so that $\mathbb{Q}_{T,x}$ has total mass 1. If now h solves the KPZ equation with $h(0) = 0$ and if $\overleftarrow{h}(s) = h(T - s)$, then we can write

$$\begin{aligned} 0 &= \overleftarrow{h}(0, x) + \int_0^T (\partial_t + \Delta)\overleftarrow{h}(t, B_t)\mathrm{d}t + \int_0^T \mathrm{D}\overleftarrow{h}(t, B_t)\mathrm{d}B_t \\ &= \overleftarrow{h}(0, x) + \int_0^T (-\overleftarrow{\xi}(t, B_t) + \infty)\mathrm{d}t + \int_0^T \mathrm{D}\overleftarrow{h}(t, B_t)\mathrm{d}B_t - \int_0^T |\mathrm{D}\overleftarrow{h}(t, B_t)|^2\mathrm{d}t, \end{aligned}$$

which shows that under $\mathbb{Q}_{T,x}$ the process B solves the SDE

$$B_t = x + \int_0^T 2\mathrm{D}\overleftarrow{h}(t, B_t)\mathrm{d}t + \mathrm{d}W_t,$$

where W is a $\mathbb{Q}_{T,x}$-Brownian motion with variance 2.

Of course, a priori such a diffusion equation does not make any sense because $\mathrm{D}\overleftarrow{h}(t, \cdot)$ is a distribution and not a function. But let $\overleftarrow{Y}$, $\overleftarrow{Y}^{\mathsf{V}}$, and $\overleftarrow{Y}^R$ be as in Section 7.1. Then we can rewrite

$$\begin{aligned} \int_0^T (\overleftarrow{\xi}(t, B_t) - \infty)\mathrm{d}t = {} & Y(0, x) + Y^{\mathsf{V}}(0, x) + Y^R(0, x) - Y(T, B_T) \\ & + \int_0^T (\overleftarrow{X} + \overleftarrow{X}^{\mathsf{V}})(t, B_t)\mathrm{d}B_t - \int_0^T |(\overleftarrow{X} + \overleftarrow{X}^{\mathsf{V}})(t, B_t)|^2\mathrm{d}t \\ & + \int_0^T \overleftarrow{X}^R(t, B_t)(\mathrm{d}B_t - 2(\overleftarrow{X} + \overleftarrow{X}^{\mathsf{V}})(t, B_t)\mathrm{d}t), \end{aligned}$$

and the only terms that will be distributions in the limit are X and X^{V}. So let us define

$$Z = (X + X^{\mathsf{V}}) \tag{66}$$

and set

$$\mathrm{d}\mathbb{P}^Z_{T,x} = \exp\Big(\int_0^T \overleftarrow{Z}(t, B_t)\mathrm{d}B_t - \int_0^T |\overleftarrow{Z}(t, B_t)|^2\mathrm{d}t\Big)\mathrm{d}\mathbb{P}.$$

Then under $\mathbb{P}^Z_{T,x}$ the coordinate process B solves

$$B_t = x + \int_0^t 2\overleftarrow{Z}(s, B_s)\mathrm{d}s + \tilde{B}_t,$$

where $\tilde{B}$ is a (variance 2) Brownian motion under $\mathbb{P}_{T,x}$. The advantage of this splitting of the Radon-Nikodym density is that now we isolated the singular part of the measure, and $\mathbb{Q}_{T,x}$ is absolutely continuous with respect to $\mathbb{P}^Z_{T,x}$, with

$$\mathrm{d}\mathbb{Q}_{T,x} = \frac{\exp(-Y(T,B_T) + \int_0^T \overleftarrow{X}^R(s,B_s)\mathrm{d}\tilde{B}_s)}{\mathbb{E}_{\mathbb{P}^Z_{T,x}}[\exp(-Y(T,B_T) + \int_0^T \overleftarrow{X}^R(s,B_s)\mathrm{d}\tilde{B}_s)]}\mathrm{d}\mathbb{P}^Z_{T,x}.$$

The density on the right hand side is strictly positive and in $L^p(\mathbb{Q}_{T,x})$ for all $p \in [1,\infty)$, even for general $\mathbb{Y} \in \mathcal{Y}_{\mathrm{kpz}}$ (not necessarily smooth). It remains to construct the measure $\mathbb{P}^Z_{T,x}$ for general $\mathbb{Y} \in \mathcal{Y}_{\mathrm{kpz}}$. In general $\mathbb{P}^Z_{T,x}$ will be singular with respect to the Wiener measure.

Theorem 7.9. *Let $\mathbb{Y} \in \mathcal{Y}_{\mathrm{kpz}}$ and $T > 0$. Consider for given $\varphi_T \in \mathscr{C}^{\alpha+1}$ and $f \in C([0,T],L^\infty)$ the solution φ to the paracontrolled equation*

$$(\partial_t + \Delta)\varphi = -2\overleftarrow{Z}\mathrm{D}\varphi + f, \qquad \varphi(T) = \varphi_T.$$

Then for every $x \in \mathbb{T}$ there exists a unique probability measure $\mathbb{P}^Z_{T,x}$ on $\Omega = C([0,T],\mathbb{T})$, such that $\mathbb{P}^Z_{T,x}(\gamma_0 = x) = 1$ and for all φ as above the process

$$\varphi(t,\gamma_t) - \int_0^t f(s,\gamma_s)\mathrm{d}s, \qquad t \in [0,T],$$

is a square integrable martingale. Here γ denotes the coordinate process on Ω. Moreover, assume that $\mathbb{Y}_n = \Theta_{\mathrm{kpz}}(\theta_n, c_n^{\vee}, c_n^{\vee\vee})$ for a sequence $\theta_n \in \mathscr{L}C^{\alpha/2}_{\mathrm{loc}}(\mathbb{R},C^\infty)$ and $c_n^{\vee}, c_n^{\vee\vee} \in \mathbb{R}$ is such that $(\mathbb{Y}_n)$ converges to $\mathbb{Y}$ in $\mathcal{Y}_{\mathrm{kpz}}$. Let B be a variance 2 Brownian motion and consider the process

$$\gamma^n_t = x + \int_0^t 2\overleftarrow{Z}_n(s,\gamma^n_s)\mathrm{d}s + B_t$$

with $Z_n = X_n + X_n^{\vee} = \mathrm{D}(Y_n + Y_n^{\vee})$, which is well defined because Z_n is Lipschitz continuous, uniformly in $t \in [0,T]$. Then there exists a Brownian motion $\tilde{B}$ on Ω with respect to the measure $\mathbb{P}^Z_{T,x}$ such that (γ_n, B) converges weakly to $(\gamma, \tilde{B})$. Finally, γ is a time-inhomogeneous strong Markov process under the family of measures $(\mathbb{P}^Z_{T,x})_{x\in\mathbb{T}}$.

Proof. In the paper [CC15] diffusions with singular drift are constructed as solutions to singular martingale problems. Using paracontrolled distributions it is shown that if $\overleftarrow{Z} \in C_T\mathscr{C}^{\alpha-1}$ and $\overleftarrow{Q^Z} \circ \overleftarrow{Z} \in C_T\mathscr{C}^{2\alpha-1}$ is given for the solution $\overleftarrow{Q^Z}$ to $(\partial_t + \Delta)\overleftarrow{Q^Z} = -\overleftarrow{\mathrm{D}Z}$, $\overleftarrow{Q^Z}(0) = 0$ (i.e. $\overleftarrow{Z}$ is a "ground drift" as defined in [CC15]), then there exists a unique martingale solution γ to

$$\gamma_t = x + \int_0^t 2\overleftarrow{Z}(s,\gamma_s)\mathrm{d}s + B_t,$$

and if $(\overleftarrow{Z_n}, \overleftarrow{Q^{Z_n}} \circ \overleftarrow{Z_n}) \to (\overleftarrow{Z}, \overleftarrow{Q^Z} \circ \overleftarrow{Z})$ for $n \to \infty$, then γ is the limit of the γ^n solving the same equation with $\overleftarrow{Z^n}$ instead of Z. See Theorem 2.6 and the discussion in Section 3 of [CC15].

It remains to construct $\overleftarrow{Q^Z} \circ \overleftarrow{Z}$. But $\overleftarrow{Q^Z} = \overleftarrow{Q} + \overleftarrow{Q^{\mathsf{V}}}$, where $(\partial_t + \Delta)\overleftarrow{Q^{\mathsf{V}}} = -\overleftarrow{\mathrm{D}X^{\mathsf{V}}}$. Since $\overleftarrow{X^{\mathsf{V}}} \in C_T\mathscr{C}^{2\alpha-1}$ and thus $\overleftarrow{Q^{\mathsf{V}}} \in C_T\mathscr{C}^{2\alpha}$, the only problematic term in the definition of $\overleftarrow{Q^Z} \circ \overleftarrow{Z}$ is $\overleftarrow{Q} \circ \overleftarrow{X} = \overleftarrow{Q \circ X}$. And since $Q \circ X$ is "contained" in $\mathbb{Y}$, the proof is complete. □

Remark 7.10. *We expect that for* $0 \le s < t \le T$ *under* $(\mathbb{P}^Z_{T,x})_{x\in\mathbb{T}}$ *the transition function*

$$T_{s,t}g(x) = \mathbb{E}_{\mathbb{P}^Z_T}[g(\gamma_t)|\gamma_s = x]$$

has a density $p_{s,t}(x,y)$ *which is jointly continuous in* (s,t,x,y)*. For that purpose we would have to solve the generator equation*

$$(\partial_r + \Delta + \overleftarrow{Z}\mathrm{D})\varphi = f,\ r \in [0,t), \qquad \varphi(t) = \varphi_t \tag{67}$$

for general $t \in [0,T]$*. Taking* $f \equiv 0$*, we would have* $T_{s,t}\varphi_t(x) = \varphi(s,x)$*. Then we could enlarge the class of terminal conditions* φ^t*, and using the techniques of Section 6.3 it should be possible to take* φ_t *as the Dirac delta in an arbitrary point* $y \in \mathbb{T}$*. But then the transition density* $p_{s,t}(x,y)$ *of* $T_{s,t}$ *must be given by* $\varphi(s,x)$*, where* φ *solves* (67) *with terminal condition* $\delta(y)$*.*

We also expect that the density is strictly positive: In [CFG15] *it is shown that if* u *solves a linear paracontrolled equation and* $u(0) \in C(\mathbb{T})$ *is a continuous, nonnegative and non-zero function, then* $u(t,x) > 0$ *for all* $t > 0$*,* $x \in \mathbb{T}$*. Combining this with the smoothing effect of the linear equation* (67)*, which takes as the terminal condition at time* t *the Dirac delta and returns a nonnegative non-zero continuous function for all sufficiently large* $r < t$*, the strict positivity of the density will follow.*

Remark 7.11. *In the setting of Theorem 7.9, we can construct the "full random directed polymer measure"* $\mathbb{Q}_{T,x}$ *by setting*

$$\frac{\mathrm{d}\mathbb{Q}_{T,x}}{\mathrm{d}\mathbb{P}^Z_{T,x}} = \frac{\exp(-Y(T,\gamma_T) + \int_0^T \overleftarrow{X}^R(s,\gamma_s)\mathrm{d}\tilde{B}_s)}{\mathbb{E}_{\mathbb{P}^Z_{T,x}}[\exp(-Y(T,\gamma_T) + \int_0^T \overleftarrow{X}^R(s,\gamma_s)\mathrm{d}\tilde{B}_s)]},$$

which is now a perfectly well defined expression.

Similarly we can also construct the measure $\mathbb{Q}_{T,(t,x)}$ *under which the coordinate* γ *process formally solves*

$$\gamma_s = x + \int_t^s 2\mathrm{D}\overleftarrow{h}(r,\gamma_r)\mathrm{d}r + (B_s - B_t), \qquad s \in [t,T].$$

With respect to $\mathbb{Q}_{T,(t,x)}$ *we have the following Feynman-Kac representation for the solution* w *to the paracontrolled equation heat equation driven by* $\mathbb{Y}$ *and started in* $w(0) = \bar{w}$*:*

$$w(t,x) = \mathbb{E}_{\mathbb{Q}_{T,(t,x)}}[\bar{w}(\gamma_T)]\tilde{w}(t,x),$$

where $\tilde{w}(t,x)$ *solves the same equation as* w *but started in* $\tilde{w}(0) \equiv 1$*.*

7.3 Rigorous control problem

The control problem that we formulated in Section 7.1 worked only for "smooth" elements of $\mathcal{Y}_{\mathrm{kpz}}$, that is for $\mathbb{Y} = \Theta_{\mathrm{kpz}}(\theta, c^{\vee}, c^{\vee\!\vee})$ with $\theta \in \mathscr{L}C^{\alpha/2}_{\mathrm{loc}}(\mathbb{R}, C^{\infty})$ and $c^{\vee}, c^{\vee\!\vee} \in \mathbb{R}$. Let us use the construction of Section 7.2 to make it rigorous also for general $\mathbb{Y} \in \mathcal{Y}_{\mathrm{kpz}}$.

For that purpose fix $T > 0$ and set $\Omega_T = C([0,T], \mathbb{T})$, equipped with the canonical filtration $(\mathcal{F}_t)_{t\in[0,T]}$. Recall that a function $v \colon [0,T] \times \Omega_T \to \mathbb{R}$ is called *progressively measurable* if for all $t \in [0,T]$ the map $[0,t] \times \Omega_T \ni (s,\omega) \mapsto v(s,\omega)$ is $\mathscr{B}[0,t] \otimes \mathcal{F}_t$–measurable.

From now on we also fix $\mathbb{Y} \in \mathcal{Y}_{\mathrm{kpz}}$ and define Z as in (66).

Definition 7.12. Let v be a progressively measurable process and let $(\tilde{\Omega}, \tilde{\mathcal{F}}, (\tilde{\mathcal{F}}_t)_{t\in[0,T]}, \mathbb{P})$ be a filtered probability space. A stochastic process on this space is called a *martingale solution* to the equation

$$\gamma_t = x + \int_0^t (2\overleftarrow{Z}(s,\gamma_s) + v(s,\gamma))\mathrm{d}s + B_t \tag{68}$$

if $\mathbb{P}(\gamma_0 = x) = 1$ and whenever $\varphi_T \in \mathscr{C}^{\alpha+1}$ and $f \in C([0,T], L^{\infty})$ and φ solves the paracontrolled equation

$$(\partial_t + \Delta + 2\overleftarrow{Z}\mathrm{D})\varphi = f, \qquad \varphi(T) = \varphi_T.$$

on $[0,T]$, then

$$\varphi(t,\gamma_t) - \int_0^t (f(s,\gamma_s) + \mathrm{D}\varphi(s,\gamma_s)v(s,\gamma))\mathrm{d}s, \qquad t \in [0,T],$$

is a martingale.

We write $\mathfrak{pm}$ for the set of progressively measurable processes, and for a given $v \in \mathfrak{pm}$ and $x \in \mathbb{T}$ we write $\mathfrak{M}(v,x)$ for the collection of all martingale solutions to (68). Note that here we explicitly allow the probability space to vary for different martingale solutions, and also that we do not make any claim about the existence or uniqueness (in law) of martingale solutions.

Theorem 7.13. *Let $\bar{h} - Y(0) \in \mathscr{C}^{2\alpha+1}$ and let h be a paracontrolled solution to the KPZ equation with initial condition $\bar{h}$. Then*

$$\begin{aligned}(h - Y - Y^{\vee} - Y^R)(T,x) \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad (69)\\ = \sup_{v\in\mathfrak{pm}} \sup_{\gamma\in\mathfrak{M}(v,x)} \mathbb{E}\Big[\bar{h}(\gamma_T) - Y(0,\gamma_T) + \int_0^T \Big(|\overleftarrow{X}^R|^2 - \frac{1}{4}|v - 2\overleftarrow{X}^R|^2\Big)(s,\gamma)\mathrm{d}s\Big]\end{aligned}$$

and the optimal v is

$$v(t,\gamma) = 2\mathrm{D}(\overleftarrow{h} - \overleftarrow{X} - \overleftarrow{X}^{\vee})(t,\gamma_t) = 2(\overleftarrow{X}^R + \mathrm{D}\overleftarrow{h}^R)(t,\gamma_t),$$

where h^R solves the paracontrolled equation

$$\mathscr{L}h^R = |X^R|^2+2(X+X^{\mathsf{V}}+X^R)\mathrm{D}h^R+|\mathrm{D}h^R|^2, \qquad h^R(0)=\bar{h}-Y(0). \tag{70}$$

For this v and all $x\in\mathbb{T}$, the set $\mathfrak{M}(v,x)$ is non-empty and every $\gamma\in\mathfrak{M}(v,x)$ has the same law.

Proof. Clearly h solves the KPZ equation driven by $\mathbb{Y}$ and started in $\bar{h}$ if and only if $h^R = h - X - X^{\mathsf{V}} - X^R$ solves (70). Moreover, the paracontrolled structure of h^R is

$$h^R = 2(\mathrm{D}h^R) \prec\!\!\prec P + h^{R,\sharp}$$

with $h^{R,\sharp}\in\mathscr{L}^{2\alpha+1}$, and in particular $h^R\in C\mathscr{C}^{\alpha+1}$. Reversing time, we get

$$(\partial_t+\Delta+2\overleftarrow{Z}\mathrm{D})\overleftarrow{h}^R = -|\overleftarrow{X}^R|^2-2\overleftarrow{X}^R\mathrm{D}\overleftarrow{h}^R-|\mathrm{D}\overleftarrow{h}^R|^2, \qquad \overleftarrow{h}^R(T)=\bar{h}-Y(0).$$

Let now $v\in\mathfrak{pm}$ and let γ be a martingale solution to

$$\gamma_t = x+\int_0^t (2\overleftarrow{Z}(s,\gamma_s)+v(s,\gamma))\mathrm{d}s + B_t.$$

Since $\bar{h}-Y(0)\in\mathscr{C}^{2\alpha+1}$, we can take $\overleftarrow{h}^R$ as test function in the martingale problem and get

$$\begin{aligned}\bar{h}(\gamma_T)-Y(0,\gamma_T) = \overleftarrow{h}^R(0,x) &+ \int_0^T (-|\overleftarrow{X}^R|^2-2\overleftarrow{X}^R\mathrm{D}\overleftarrow{h}^R-|\mathrm{D}\overleftarrow{h}^R|^2)(s,\gamma_s)\mathrm{d}s \\ &+ \int_0^T \mathrm{D}\overleftarrow{h}^R(s,\gamma_s)v(s,\gamma)\mathrm{d}s + \text{mart.},\end{aligned}$$

so writing $\tilde{v}(s,\gamma)=v(s,\gamma)-2\overleftarrow{X}^R(s,\gamma_s)-2\mathrm{D}\overleftarrow{h}^R(s,\gamma_s)$ we obtain

$$\begin{aligned}&\mathbb{E}\Big[\bar{h}(\gamma_T)-Y(0,\gamma_T)+\int_0^T\Big(|\overleftarrow{X}^R|^2-\frac{1}{4}|v-2\overleftarrow{X}^R|^2\Big)(s,\gamma)\mathrm{d}s\Big]-\overleftarrow{h}^R(0,x)\\ &\quad= \mathbb{E}\Big[\int_0^T(-|\overleftarrow{X}^R|^2-2\overleftarrow{X}^R\mathrm{D}\overleftarrow{h}^R-|\mathrm{D}\overleftarrow{h}^R|^2+\mathrm{D}\overleftarrow{h}^R(2\overleftarrow{X}^R+2\mathrm{D}\overleftarrow{h}^R+\tilde{v}))(s,\gamma)\mathrm{d}s\Big]\\ &\qquad+\mathbb{E}\Big[\int_0^T\Big(|\overleftarrow{X}^R|^2-\frac{1}{4}|\tilde{v}+2\mathrm{D}\overleftarrow{h}^R|^2\Big)(s,\gamma_s)\mathrm{d}s\Big]\\ &\quad= \mathbb{E}\Big[\int_0^T-\frac{1}{4}|\tilde{v}(s,\gamma)|^2\mathrm{d}s\Big].\end{aligned}$$

This shows that

$$\sup_{v\in\mathfrak{pm}}\sup_{\gamma\in\mathfrak{M}(v,x)}\mathbb{E}\Big[\bar{h}(\gamma_T)-Y(0,\gamma_T)+\int_0^T\Big(|\overleftarrow{X}^R|^2-\frac{1}{4}|v-2\overleftarrow{X}^R|^2\Big)(s,\gamma)\mathrm{d}s\Big]\leqslant h^R(T,x).$$

On the other side, taking $v = 2\overleftarrow{X}^R + 2\mathrm{D}\overleftarrow{h}^R$ we obtain a ground drift in the terminology of [CC15], and therefore for all $x \in \mathbb{T}$ there exists a $\gamma \in \mathfrak{M}(v, x)$ and its law is unique. For any such γ we obtain

$$\mathbb{E}\Big[\bar{h}(\gamma_T) - Y(0, \gamma_T) + \int_0^T \Big(|\overleftarrow{X}^R|^2 - \frac{1}{4}|v - 2\overleftarrow{X}^R|^2\Big)(s, \gamma)\mathrm{d}s\Big] = h^R(T, x),$$

and thus the proof is complete. □

Remark 7.14. *We only needed $\bar{h} - Y(0) \in \mathscr{C}^{2\alpha+1}$ to apply the results of* [CC15]*. By using similar arguments as in Section 6 it will be possible to weaken the assumptions on the ground drift (allowing a possible singularity at 0) and on the terminal condition in the martingale problem in* [CC15]*, and then the variational representation of Theorem 7.13 will extend to $\bar{h} \in \mathscr{C}^{1-\beta}$ for $\beta < 1$ as in Theorem 6.13.*

Actually, the extension to $\bar{h} \in \mathscr{C}^{\alpha}$ is immediate because we can simply start Y^R in $\bar{h} - Y(0)$ (which still gives us a ground drift $\overleftarrow{Z} + 2\overleftarrow{X}^R + 2\mathrm{D}\overleftarrow{h}^R$, where h^R is now started in 0), and change the control problem to

$$\sup_{v \in \mathfrak{pm}} \sup_{\gamma \in \mathfrak{M}(v,x)} \mathbb{E}\Big[\int_0^T \Big(|\overleftarrow{X}^R|^2 - \frac{1}{4}|v - 2\overleftarrow{X}^R|^2\Big)(s, \gamma)\mathrm{d}s\Big].$$

Similarly, if we start Y^R in $-Y(0)$ the control problem has the more appealing form

$$\sup_{v \in \mathfrak{pm}} \sup_{\gamma \in \mathfrak{M}(v,x)} \mathbb{E}\Big[\bar{h}(\gamma_T) + \int_0^T \Big(|\overleftarrow{X}^R|^2 - \frac{1}{4}|v - 2\overleftarrow{X}^R|^2\Big)(s, \gamma)\mathrm{d}s\Big].$$

8 Convergence of Sasamoto-Spohn lattice models

In this section we consider the weak universality conjecture in the context of weakly asymmetric interface models $\varphi_N : \mathbb{R}_+ \times \mathbb{Z}_N \to \mathbb{R}$ (where $\mathbb{Z}_N = \mathbb{Z}/(N\mathbb{Z})$) with

$$\begin{aligned} \mathrm{d}\varphi_N(t, x) &= \Delta_{\mathbb{Z}_N}\varphi_N(t, x)\mathrm{d}t + \sqrt{\varepsilon}(B_{\mathbb{Z}_N}(\mathrm{D}_{\mathbb{Z}_N}\varphi_N(t), \mathrm{D}_{\mathbb{Z}_N}\varphi_N(t)))(x)\mathrm{d}t + \mathrm{d}W_N(t, x), \\ \varphi_N(0, x) &= \varphi_0^N(x), \end{aligned} \tag{71}$$

where $\Delta_{\mathbb{Z}_N}$ and $\mathrm{D}_{\mathbb{Z}_N}$ are discrete versions of Laplacian and spatial derivative respectively, $B_{\mathbb{Z}_N}$ is a bilinear form taking the role of the pointwise product, $(W_N(t, x))_{t \in \mathbb{R}_+, x \in \mathbb{Z}_N}$ is an N–dimensional standard Brownian motion, and φ_0^N is independent of W_N. We assume throughout this section that

$$\varepsilon = \frac{2\pi}{N}.$$

Equation (71) is a generalization of the Sasamoto-Spohn discretization of the KPZ equation, see Remark 8.3 below. To simplify things (eliminating the need to introduce renormalization constants), let us look at the flux $\mathrm{D}_{\mathbb{Z}_N}\varphi_N$. Assume that there exists $\beta < 1$ for which $(x \mapsto (\mathrm{D}_{\mathbb{Z}_N}\varphi_0^N)(x/\varepsilon))_N$ converges weakly to 0 in $\mathscr{C}^{-\beta}$ with rate of convergence

$\varepsilon^{1/2}$. Then $((t,x) \mapsto \mathrm{D}_{\mathbb{Z}_N} \varphi_N(t/\varepsilon^2, x/\varepsilon))$ converges to 0 (this will be a consequence of our analysis below), and we can study the fluctuations defined by

$$u_N(t,x) = \varepsilon^{-1/2} \mathrm{D}_{\mathbb{Z}_N} \varphi_N(t/\varepsilon^2, x/\varepsilon).$$

This is a stochastic process on $\mathbb{R}_+ \times \mathbb{T}_N$ with $\mathbb{T}_N = (\varepsilon\mathbb{Z})/(2\pi\mathbb{Z})$ which solves the SDE

$$\begin{aligned} \mathrm{d}u_N(t,x) &= \Delta_N u_N(t,x)\mathrm{d}t + (\mathrm{D}_N B_N(u_N(t), u_N(t)))(x)\mathrm{d}t + \mathrm{d}(\mathrm{D}_N \varepsilon^{-1/2} W_N(t,x)) \\ u_N(0) &= u_0^N. \end{aligned} \tag{72}$$

where Δ_N, D_N, and B_N are approximations of Laplacian, spatial derivative, and pointwise product respectively, $\partial_t \mathrm{D}_N \varepsilon^{-1/2} W_N$ converges to $\mathrm{D}\xi$, where ξ is a space-time white noise and $u_0^N(x) = \mathrm{D}_N \varepsilon^{1/2} \varphi_0^N(x/\varepsilon)$. We show that if $\mathrm{D}_N \varepsilon^{1/2} \varphi_0^N(x/\varepsilon)$ converges in distribution in $\mathscr{C}^{-\beta}$, then (u_N) converges in distribution to the solution of a modified Burgers equation involving a sort of Itô-Stratonovich corrector.

Another way of reading our result is that (72) is a lattice discretization of the Burgers equation and we show that it might converge to a different equation in the limit, depending on how we choose Δ_N, D_N, and B_N.

There are two problems that we have to deal with before we can study the convergence. First, it is not obvious whether u_N blows up in finite time, because the equation contains a quadratic nonlinearity. Let therefore ζ be a cemetery state and define the space

$$C_N = \{\varphi \colon \mathbb{R}_+ \to \mathbb{R}^{\mathbb{T}_N} \cup \{\zeta\}, \varphi \text{ is continuous on } [0, \tau_\zeta(\varphi)) \text{ and } \varphi(\tau_\zeta(\varphi) + t) = \zeta, t \geqslant 0\},$$

where

$$\tau_\zeta(\varphi) = \inf\{t \geqslant 0 : \varphi(t) = \zeta\} \quad \text{and for } c > 0 \quad \tau_c(\varphi) = \inf\{t \geqslant 0 : \|\varphi(t)\|_{L^\infty(\mathbb{T}_N)} \geqslant c\},$$

with $\|\zeta\|_{L^\infty(\mathbb{T}_N)} = \infty$. Then a stochastic process u_N with values in C_N is a solution to (72) if $\tau_\zeta(u_N) = \sup_{c>0} \tau_c(u_N)$ and $u_N|_{[0,\tau_\zeta(u_N))}$ solves (72) on $[0, \tau_\zeta(u_N))$. It is a classical result that there exists a unique solution in that sense (which is adapted to the filtration generated by u_0^N and W_N, but we will not need this).

The next problem that we face is that $u_N(t)$ is only defined on the grid $\mathbb{T}_N$ and not on the entire torus $\mathbb{T}$. Since we will only obtain convergence in a space of distributions and not of continuous functions, some care has to be exercised when choosing an extension of u_N to $\mathbb{T}$. For $\delta > 0$ one can easily define sequences of smooth functions (f_N) and (g_N) on $\mathbb{T}$ such that f_N and g_N agree in the lattice points $\mathbb{T}_N$, but both converge to different limits in $\mathscr{C}^{-\delta}$. Here we will work with a particularly convenient extension of (u_N) that can be constructed using discrete Fourier transforms [HM12, HMW14]. Since it will simplify the notation, we make the following assumption from now on:

$$N \text{ is odd.}$$

Of course, our results do not depend on that assumption and we only make it for convenience. We define for $\varphi\colon \mathbb{T}_N \to \mathbb{C}$ the discrete Fourier transform

$$\mathscr{F}_N\varphi(k) = \varepsilon \sum_{|\ell|<N/2} \varphi(\varepsilon\ell)e^{-i\varepsilon\ell k}, \qquad k \in \mathbb{Z}_N,$$

(for even N we would have to adapt the domain of summation), and then

$$\mathcal{E}_N\varphi(x) = (2\pi)^{-1} \sum_{|k|<N/2} \mathscr{F}_N\varphi(k)e^{ikx}, \qquad x \in \mathbb{T}.$$

Then $\mathcal{E}_N\varphi(x) = \varphi(x)$ for all $x \in \mathbb{T}_N$, and by construction $\mathcal{E}_N\varphi$ is a smooth function with Fourier transform $\mathscr{F}\mathcal{E}_N\varphi(k) = \mathscr{F}_N\varphi(k)\mathbb{1}_{|k|<N/2}$. If φ is real valued, then so is $\mathcal{E}_N\varphi$.

Remark 8.1. *A quite generic method of extending a function $\varphi\colon \mathbb{T}_N \to \mathbb{R}$ to $\mathbb{T}$ is as follows: write*

$$\psi = \sum_{x\in\mathbb{T}_N} \varphi(x)\varepsilon\delta_x,$$

let $\eta_N \in \mathscr{D}'(\mathbb{T})$, and define $\bar{\varphi}(x) = \psi * \eta_N$, which yields for all $k \in \mathbb{Z}$

$$\mathscr{F}\bar{\varphi}(k) = \mathscr{F}\psi(k)\mathscr{F}\eta_N(k) = \mathscr{F}_N\varphi(k)\mathscr{F}\eta_N(k).$$

Now assume that (φ_N) is a sequence of functions on $\mathbb{T}_N$ such that $(\mathcal{E}_N\varphi_N)$ converges in $\mathscr{D}'(\mathbb{T})$ to some limit φ_∞. Convergence in $\mathscr{D}'(\mathbb{T})$ is equivalent to the convergence of all Fourier modes, together with a uniform polynomial bound on their growth, and thus we get $\lim_N \mathscr{F}_N\varphi_N(k) = \mathscr{F}\varphi_\infty(k)$ for all $k \in \mathbb{Z}$. So if (η_N) converges to δ_0 in $\mathscr{D}'(\mathbb{T})$, then $(\bar{\varphi}_N)$ converges in $\mathscr{D}'(\mathbb{T})$ to φ_∞. Typical examples for interpolations that can be constructed in this way are the “Dirac delta extension” (take $\eta_N = \delta_0$ for all N), the “piecewise constant extension” (take $\eta_N = \varepsilon^{-1}\mathbb{1}_{[0,\varepsilon)}$), or the “piecewise linear extension” (take $\eta_N(x) = \varepsilon^{-1}((\varepsilon^{-1}x+1)\mathbb{1}_{[-\varepsilon,0]}(x) + (1-\varepsilon^{-1}x)\mathbb{1}_{(0,\varepsilon]}(x))$). In particular, our convergence result below also implies the convergence of all these interpolations.

For $\varphi \in C_N$ we then get $\mathcal{E}_N\varphi\colon \mathbb{R}_+ \to C^\infty(\mathbb{T}) \cup \{\zeta\}$, and for $T > 0$ and $\beta \in \mathbb{R}$ we have

$$\mathcal{E}_N\varphi|_{[0,T]} \in C_T^*\mathscr{C}^\beta := C_T\mathscr{C}^\beta \cup \{\infty\},$$

where we identify all trajectories that take the value ζ and assign them the symbol ∞. On $C_T^*\mathscr{C}^\beta$ we consider the metric introduced by Hairer and Matetski in [HM15],

$$\begin{aligned} d^*_{T,\beta}(\psi,\infty) &:= d^*_{T,\beta}(\infty,\psi) := (1 + \|\psi\|_{C_T\mathscr{C}^\beta})^{-1} \text{ for } \psi \neq \infty, \\ d^*_{T,\beta}(\psi,\psi') &:= \min\{\|\psi-\psi'\|_{C_T\mathscr{C}^\beta}, d^*_{T,\beta}(\psi,\infty) + d^*_{T,\beta}(\psi',\infty)\}. \end{aligned} \tag{73}$$

We will need some assumptions on the operators Δ_N, D_N, B_N: Let

$$\Delta_N\varphi(x) = \varepsilon^{-2}\int_{\mathbb{Z}} \varphi(x+\varepsilon y)\pi(\mathrm{d}y), \qquad \mathrm{D}_N\varphi(x) = \varepsilon^{-1}\int_{\mathbb{Z}} \varphi(x+\varepsilon y)\nu(\mathrm{d}y), \tag{74}$$

$$B_N(\varphi,\psi)(x) = \int_{\mathbb{Z}^2} \varphi(x+\varepsilon y)\psi(x+\varepsilon z)\mu(\mathrm{d}y,\mathrm{d}z),$$

where π, ν, and μ are finite signed measures on $\mathbb{Z}$ and a probability measure on $\mathbb{Z}^2$, respectively. We define

$$f(x) = \frac{\int_{\mathbb{Z}} e^{ixy}\pi(\mathrm{d}y)}{-x^2}, \quad g(x) = \frac{\int_{\mathbb{Z}} e^{ixy}\nu(\mathrm{d}y)}{ix}, \quad h(x_1,x_2) = \int_{\mathbb{Z}^2} e^{i(x_1z_1+x_2z_2)}\mu(\mathrm{d}z_1,\mathrm{d}z_2),$$

and make the following assumptions on the measures π, ν, μ:

($\mathbf{H}_f$) The finite signed measure π on $\mathbb{Z}$ is symmetric, has total mass zero, finite fourth moment, and satisfies $\int_{\mathbb{Z}} y^2\pi(\mathrm{d}y) = 2$. Moreover, there exists $c_f > 0$ such that $f(x) > c_f$ for all $x \in [-\pi,\pi]$.

($\mathbf{H}_g$) The finite signed measure ν on $\mathbb{Z}$ has total mass zero, finite second moment, and satisfies $\int_{\mathbb{Z}} y\nu(\mathrm{d}y) = 1$.

($\mathbf{H}_h$) The probability measure μ has a finite first moment on $\mathbb{Z}^2$ and satisfies $\mu(A\times B) = \mu(B\times A)$ for all $A, B \subseteq \mathbb{Z}$.

The constant c_f in (H_f) exists for example if $\pi = \sum_{k\geqslant 1} p_k(\delta_k + \delta_{-k}) - c\delta_0$ where $p_k \geqslant 0$, $\sum_{k\geqslant 1} 2p_k = c$ and $p_1 > 0$.

Theorem 8.2. *Make assumption (H_f), (H_g), (H_h) and let $\beta \in (-1,-1/2)$. Consider for all $N \in \mathbb{N}$ an N-dimensional standard Brownian motion $(W_N(t,x))_{t\geqslant 0, x\in\mathbb{T}_N}$ and an independent random variable $(u_0^N(x))_{x\in\mathbb{T}_N}$ and denote the solution to (72) by u_N. If there exists a random variable u_0 such that $\mathcal{E}_N u_0^N$ converges to u_0 in distribution in $\mathscr{C}^\beta$, then for all $T > 0$ the sequence $(\mathcal{E}_N u_N|_{[0,T]})_N$ converges in distribution with respect to the metric $d^*_{T,\beta}$ to u, the unique paracontrolled solution of*

$$\mathscr{L}u = \mathrm{D}u^2 + 4c\mathrm{D}u + \mathrm{D}\xi, \qquad u(0) = u_0, \tag{75}$$

where ξ is a space-time white noise which is independent of u_0, and where

$$c = -\frac{1}{4\pi}\int_0^\pi \frac{\mathrm{Im}(g(x)\bar{h}(x))}{x}\frac{h(x,-x)|g(x)|^2}{|f(x)|^2}\mathrm{d}x \in \mathbb{R}.$$

Proof. We have

$$\mathscr{F}_N B_N(\varphi,\psi)(k) = (2\pi)^{-1}\sum_{\ell\in\mathbb{Z}_N} \mathscr{F}_N\varphi(\ell)\mathscr{F}_N\psi(k-\ell)\int_{\mathbb{Z}^2} e^{i(\varepsilon\ell y+\varepsilon(k-\ell)z)}\mu(\mathrm{d}y,\mathrm{d}z),$$

from where we deduce that $\mathcal{E}_N B_N(\varphi,\psi) = \Pi_N B_N(\mathcal{E}_N\varphi,\mathcal{E}_N\psi)$ for all $\varphi,\psi \in \mathbb{R}^{\mathbb{T}_N}$, with

$$\Pi_N\varphi(x) = (2\pi)^{-1}\sum_k e^{ik^N x}\mathscr{F}\varphi(k),$$

where $k^N = \arg\min\{|\ell|: \ell \in \mathbb{Z}, \ell = k + jN \text{ for some } j \in \mathbb{Z}\} \in (-N/2, N/2)$. Moreover, we obtain for $|k|< N/2$

$$\begin{aligned}
&\mathbb{E}[\varepsilon^{-1/2}\mathscr{F}\mathcal{E}_N W_N(t,k_1)\varepsilon^{-1/2}\mathscr{F}\mathcal{E}_N W_N(t,k_2)] \\
&\qquad = \varepsilon \sum_{|\ell_1|,|\ell_2|<N/2} \mathbb{E}[W_N(t,\varepsilon\ell_1)W_N(t,\varepsilon\ell_2)]e^{-i\varepsilon(\ell_1 k_1+\ell_2 k_2)} \\
&\qquad = \varepsilon \sum_{|\ell|<N/2} t e^{-i\ell(k_1+k_2)\varepsilon} = t2\pi\delta_{k_1+k_2=0},
\end{aligned}$$

which shows that $\partial_t \varepsilon^{-1/2}\mathcal{E}_N W_N(t)$ has the same distribution as $\mathcal{P}_N\xi$, where ξ is a space-time white noise and $\mathcal{P}_N\varphi = \mathbb{1}_{(-N/2,N/2)}(\mathrm{D}) = \mathscr{F}^{-1}(\mathbb{1}_{(-N/2,N/2)}\mathscr{F}\varphi)$ is a Fourier cutoff operator.

We now place ourselves on a probability space where such a space-time white noise is given and where $(\tilde{u}_0^N)$ is a sequence of random variables with values in $\mathscr{D}'(\mathbb{T})$, which is independent of ξ, such that $\tilde{u}_0^N$ has the same distribution as $\mathcal{E}_N u_0^N$ for all N, and such that u_0^N converges to a random variable u_0 in probability in $\mathscr{C}^\beta$ (which is then also independent of ξ). Then u_N has the same distribution as $\tilde{u}_N$, the solution to

$$(\partial_t - \Delta_N)\tilde{u}_N(t,x) = \mathrm{D}_N \Pi_N B_N(\tilde{u}_N, \tilde{u}_N)(t,x) + \mathrm{D}_N \mathcal{P}_N \xi(t,x), \qquad \tilde{u}_N(0) = \tilde{u}_0^N,$$

and therefore it suffices to study $\tilde{u}_N$. The pathwise analysis of this equation is carried out in Section 8.2 below, and the convergence result assuming convergence of the data $(\mathbb{X}_N(\xi))$ and the random operators (A_N) is formulated in Theorem 8.17. The convergence of $(\mathbb{X}_N(\xi))$ and (A_N) in L^p spaces is shown in Theorem 10.1 and Theorem 10.4, respectively. Thus, we get that $(\tilde{u}_N)$ converges in probability with respect to $d^*_{T,\beta}$ to u, the solution of (75). □

Remark 8.3. *If we take B_N as the pointwise product for all N, and Δ_N as the discrete Laplacian, $\Delta_N f(x) = \varepsilon^{-2}(f(x+\varepsilon)+f(x-\varepsilon)-2f(x))$ and $\mathrm{D}_N f(x) = \varepsilon^{-1}(f(x)-f(x-\varepsilon))$, then we get $c = 1/8$, so the additional term in equation (75) is $1/2\mathrm{D}u$.*

However, if we take the same Δ_N and D_N but replace the pointwise product by

$$B_N(\varphi,\psi)(x) = \frac{1}{2(\kappa+\lambda)}(\kappa\varphi(x)\psi(x)+\lambda(\varphi(x)\psi(x+\varepsilon)+\varphi(x+\varepsilon)\psi(x))+\kappa\varphi(x+\varepsilon)\psi(x+\varepsilon))$$

for some $\kappa, \lambda \in [0,\infty)$ with $\kappa+\lambda > 0$, then one can check that $c = 0$. Here the Sasamoto–Spohn discretization [ZK65, KS91, LS98, SS09] *corresponds to $\kappa = 1$, $\lambda = 1/2$, and in that case one furthermore has*

$$\langle \varphi, \mathrm{D}_N B_N(\varphi,\varphi)\rangle_{\mathbb{T}_N} = \sum_{x\in\mathbb{T}_N} \varphi(x)\mathrm{D}_N B_N(\varphi,\varphi)(x) = 0,$$

which entails that already for fixed N there is no blow up in the system, i.e. u_N is well defined for all times. Moreover, now we can explicitly write down a family of stationary measures for u_N: For all $m \in \mathbb{R}$, the evolution of u_N is invariant under

$$\mu_m^\varepsilon(\mathrm{d}x) = \prod_{j=0}^{N-1} \frac{\exp(-\varepsilon x_j^2 + m x_j)}{Z_m^\varepsilon}\mathrm{d}x_j,$$

where Z_m^ε is a constant normalizing the mass of μ_m^ε to 1; see [SS09] *or simply verify that the $L^2(\mathbb{T})$–adjoint of the generator of u_N applied to the density of μ_m^ε equals 0 and then use Echeverría's criterion to obtain the invariance of μ_m^ε from its infinitesimal invariance* [Ech82]*. If $u_N(0) \sim \mu_m^\varepsilon$, then for all $t \geqslant 0$ the vector $(u_N(t,x))_{x\in\mathbb{T}_N}$ consists of independent Gaussian random variables with variance $\varepsilon^{-1}/2$ and mean m. Therefore, for all $t > 0$ the $\mathscr{D}'$-valued random variable $(\mathcal{E}_N u_N(t,\cdot))$ converges in distribution to a space white noise with mean m and variance $1/2$. It is also straightforward to verify that $\sup_N \mathbb{E}[\|\mathcal{E}_N u_N(0,\cdot)\|_{B^\alpha_{p,p}}^p] < \infty$ whenever $\alpha < -1/2$, and then the Besov embedding theorem shows that the convergence actually takes place in distribution in $\mathscr{C}^\beta$, for β as required in Theorem 8.2. But if $\mathcal{E}_N u_N$ is a stationary process for all N, then any limit in distribution must be stationary as well, and this shows that the white noise with mean m and variance $1/2$ is an invariant distribution for the stochastic Burgers equation. This is of course well known, see for example* [BG97] *or* [FQ14]*. But to the best of our knowledge, ours is the first proof which does not rely on the Cole–Hopf transform. See also* [ZZ15b, HM15] *for a similar-in-spirit proof of the invariance of the ϕ^4_3 measure for the ϕ^4_3 equation.*

We now take

$$\mathscr{L}_N u_N(t,x) = (\partial_t - \Delta_N) u_N(t,x) = \mathrm{D}_N \Pi_N B_N(u_N,u_N)(t,x) + \mathrm{D}_N \mathcal{P}_N \xi(t,x) \qquad (76)$$

as the starting point for our analysis, where we assume that $u_N(0) = \mathcal{P}_N u_N(0)$. Recall that

$$\Pi_N \varphi(x) = (2\pi)^{-1} \sum_k e^{ik^N x} \mathscr{F}\varphi(k)$$

with

$$k^N = \arg\min\{|\ell| \colon \ell \in \mathbb{Z}, \ell = k + jN \text{ for some } j \in \mathbb{Z}\} \in (-N/2, N/2),$$

and that

$$\mathcal{P}_N \varphi = \mathbb{1}_{(-N/2,N/2)}(\mathrm{D}) = \mathscr{F}^{-1}(\mathbb{1}_{(-N/2,N/2)} \mathscr{F}\varphi).$$

The operators Δ_N, D_N, and B_N are given in terms of finite signed measures π, ν, μ as described in (74).

Lemma 8.4. *Let π be a finite signed measure on $\mathbb{R}$ that satisfies (H_f). Then the function*

$$f(x) = -\frac{\int_{\mathbb{R}} e^{ixy} \pi(\mathrm{d}y)}{x^2} = \int_{\mathbb{R}} \frac{1-\cos(xy)}{(xy)^2} y^2 \pi(\mathrm{d}y)$$

is in C_b^2 and such that $f(0) = 1$.

Proof. The function $\varphi(x) = (1-\cos(x))/x^2 = 2\sin^2(x/2)/x^2$ is nonnegative, bounded by $1/2$, and satisfies $\varphi(0) = 1/2$. Therefore, f is bounded and $f(0) = 1/2 \int_{\mathbb{R}} y^2 \pi(\mathrm{d}y) = 1$. Furthermore, it is easy to check that $\varphi \in C_b^2$, and thus

$$|f'(x)| \leqslant \int_{\mathbb{R}} |\varphi'(xy)||y|^3 |\pi|(\mathrm{d}y) \lesssim \int_{\mathbb{R}} |y|^3 |\pi|(\mathrm{d}y), \qquad |f''(x)| \lesssim \int_{\mathbb{R}} |y|^4 |\pi|(\mathrm{d}y).$$

As π has a finite fourth moment, this shows that $f \in C_b^2$. □

Lemma 8.5. *Let ν be a finite signed measure on $\mathbb{R}$ that satisfies (H_g). Then the function*

$$g(x) = \frac{\int_{\mathbb{R}} e^{ixy}\nu(\mathrm{d}y)}{ix} = \int_{\mathbb{R}} \frac{e^{ixy}-1}{ixy} y\nu(\mathrm{d}y)$$

is in C_b^1 and such that $g(0) = 1$.

Proof. It suffices to observe that the function $\varphi(x) = (e^{ix} - 1)/(ix)$ is in C_b^1 and satisfies $\varphi(0) = 1$, and then to copy the proof of Lemma 8.4. □

The next lemma is a simple and well known statement about characteristic functions of probability measures.

Lemma 8.6. *Let μ be a probability measure on $\mathbb{R}^2$ that satisfies (H_h). Then the function*

$$h(x, y) = \int_{\mathbb{R}^2} e^{i(xz_1+yz_2)}\mu(\mathrm{d}z_1, \mathrm{d}z_2)$$

is in C_b^1 and such that $h(0, 0) = 1$.

Our general strategy is to find a paracontrolled structure for (76) and then to follow the same steps as in the continuous setting. To do so, we need to translate all steps of our continuous analysis to the discrete setting.

8.1 Preliminary estimates

Fourier cutoff. The cutoff operator $\mathcal{P}_N$ is not a bounded operator on $\mathscr{C}^\alpha$ spaces (at least not uniformly in N) and will lead to a small loss of regularity.

Lemma 8.7. *We have*

$$\|\mathcal{P}_N\varphi\|_{L^\infty} \lesssim \log N \|\varphi\|_{L^\infty},$$

and in particular we get for all $\delta > 0$

$$\|\mathcal{P}_N\varphi - \varphi\|_{\alpha-\delta} \lesssim N^{-\delta} \log N \|\varphi\|_\alpha \lesssim \|\varphi\|_\alpha.$$

Proof. We have

$$\|\mathcal{P}_N\varphi\|_{L^\infty(\mathbb{T})} \lesssim \left\|\mathscr{F}^{-1}\mathbb{1}_{(-N/2,N/2)}\right\|_{L^1(\mathbb{T})} \|\varphi\|_{L^\infty(\mathbb{T})},$$

and using that N is odd we get

$$\mathscr{F}^{-1}\mathbb{1}_{(-N/2,N/2)}(x) = (2\pi)^{-1} \sum_{|k|<N/2} e^{ikx} = (2\pi)^{-1} \frac{\cos(x(N-1)/2) - \cos(x(N+1)/2)}{1 - \cos(x)}.$$

Now

$$\left|\frac{\cos(x(N-1)/2) - \cos(x(N+1)/2)}{1 - \cos(x)}\right| \lesssim \min\left\{\frac{Nx^2}{x^2}, \frac{|x|}{x^2}\right\},$$

and therefore

$$\int_{-\pi}^{\pi} \left|\mathscr{F}^{-1}\mathbb{1}_{(-N/2,N/2)}(x)\right| \mathrm{d}x \lesssim \log N.$$

To obtain the bound for $\mathcal{P}_N\varphi - \varphi$ it suffices to note that $\mathcal{P}_N$ acts trivially on Δ_j (either as identity or as zero) unless $2^j \simeq N$. □

Lemma 8.8. *Let $\alpha \geqslant 0$ and $\varphi \in \mathscr{C}^\alpha$. Then for any $\delta \geqslant 0$*

$$\|\Pi_N\varphi - \varphi\|_{\alpha-\delta} \lesssim N^{-\delta}\log N \|\varphi\|_\alpha.$$

If $\mathrm{supp}(\mathscr{F}\varphi) \subset [-cN, cN]$ *for some $c \in (0,1)$, then this inequality extends to general $\alpha \in \mathbb{R}$.*

Proof. We already know that $\|\mathcal{P}_N\varphi - \varphi\|_{\alpha-\delta} \lesssim N^{-\delta}\log N\|\varphi\|_\alpha$. So since $\Pi_N\varphi = \mathcal{P}_N((1 + e^{iN\cdot} + e^{-iN\cdot})\varphi)$ we get for $\alpha \geqslant 0$

$$\begin{aligned}
\|\Delta_q\mathcal{P}_N((e^{-iN\cdot} + e^{iN\cdot})\varphi)\|_{L^\infty} &\leqslant \mathbb{1}_{2^q \lesssim N}\|\mathcal{P}_N((e^{-iN\cdot} + e^{iN\cdot})\varphi)\|_{L^\infty} \\
&\lesssim \sum_{j:2^j\simeq N} \mathbb{1}_{2^q\lesssim N}\log(N)\|(e^{-iN\cdot} + e^{iN\cdot})\Delta_j\varphi\|_{L^\infty} \\
&\lesssim \sum_{j:2^j\simeq N} \mathbb{1}_{2^q\lesssim N}\log(N)2^{-j\alpha}\|\varphi\|_\alpha \\
&\lesssim \mathbb{1}_{2^q\lesssim N}\log(N)N^{-\alpha}\|\varphi\|_\alpha \lesssim 2^{-q(\alpha-\delta)}\log(N)N^{-\delta}\|\varphi\|_\alpha.
\end{aligned}$$

If also $\mathrm{supp}\,(\mathscr{F}u) \subset [-cN, cN]$, then the spectrum of $(e^{iN\cdot} + e^{-iN\cdot})u$ is contained in an annulus $N\mathscr{A}$, and therefore we can replace the indicator function $\mathbb{1}_{2^q\lesssim N}$ by $\mathbb{1}_{2^q\simeq N}$ in the calculation above, from where the claim follows. □

Remark 8.9. *There exists $c \in (0,1)$, independent of N, such that if* $\mathrm{supp}\,(\mathscr{F}\psi) \subset [-N/2, N/2]$*, then* $\mathrm{supp}\,(\mathscr{F}(\varphi \prec \psi)) \subset [-cN, cN]$*. This means that we can always bound $\Pi_N(\varphi \prec \psi) - \varphi \prec \psi$ using Lemma 8.8 even if the paraproduct has negative regularity. On the other side the best statement we can make about the resonant product is that if φ and ψ are both spectrally supported in $[-N/2, N/2]$, then* $\mathrm{supp}\,(\mathscr{F}(\varphi \circ \psi)) \subset [-N, N]$*. A simple consequence is that if $\alpha + \beta > 0$, $\varphi \in \mathscr{C}^\alpha$, $\psi \in \mathscr{C}^\beta$, and* $\mathrm{supp}(\mathscr{F}\varphi) \cup \mathrm{supp}(\mathscr{F}\psi) \subset [-N/2, N/2]$*, then*

$$\|\Pi_N(\varphi\psi) - \varphi\psi\|_{\alpha\wedge\beta-\delta} \lesssim N^{-\delta}\log(N)\|\varphi\|_\alpha\|\psi\|_\beta.$$

Estimates for the bilinear form. Let us define paraproduct and resonant term with respect to B_N:

$$B_N(\varphi \prec \psi) = \sum_j B_N(S_{j-1}\varphi, \Delta_j\psi), \qquad B_N(\varphi \succ \psi) = \sum_j B_N(\Delta_j\varphi, S_{j-1}\psi),$$

$$B_N(f \circ g) = \sum_{|i-j|\leqslant 1} B_N(\Delta_i f, \Delta_j g).$$

We have the same estimates as for the usual product:

Lemma 8.10. *Let μ satisfy (H_h). For any $\beta \in \mathbb{R}$ and $\delta \in [0,1]$ we have*

$$\|B_N(\varphi \prec \psi) - \varphi \prec \psi\|_{\beta-\delta} \lesssim N^{-\delta}\|\varphi\|_{L^\infty}\|\psi\|_\beta,$$

and for $\alpha < 0$ furthermore

$$\|B_N(\varphi \prec \psi) - \varphi \prec \psi\|_{\alpha+\beta-\delta} \lesssim N^{-\delta}\|\varphi\|_\alpha\|\psi\|_\beta.$$

For $\alpha + \beta - \delta > 0$ we have

$$\|B_N(\varphi \circ \psi) - \varphi \circ \psi\|_{\alpha+\beta-\delta} \lesssim N^{-\delta}\|\varphi\|_\alpha\|\psi\|_\beta.$$

Proof. It suffices to note that $\mathscr{F}B_N(\Delta_i\varphi, \Delta_j\psi)$ and $\mathscr{F}(\Delta_i\varphi\Delta_j\psi)$ have the same support and that $\|B_N(\Delta_i\varphi, \Delta_j\psi)\|_{L^\infty} \leqslant \|\Delta_i\varphi\|_{L^\infty}\|\Delta_j\psi\|_{L^\infty}$, whereas

$$\|B_N(\Delta_i\varphi, \Delta_j\psi) - \Delta_i\varphi\Delta_j\psi\|_{L^\infty} \lesssim N^{-1}(2^i + 2^j)\|\Delta_i\varphi\|_{L^\infty}\|\Delta_j\psi\|_{L^\infty}.$$

□

To invoke our commutator estimate, we have to pass from $B_N(\cdot \prec \cdot)$ to the usual paraproduct, which can be done using the following commutator lemma. Here we write

$$\overline{B}_N\varphi = \overline{B}_N(\varphi) = B_N(\varphi, 1) = B_N(1, \varphi).$$

Lemma 8.11. *Let μ satisfy (H_h). Let $\alpha < 1$, $\beta \in \mathbb{R}$, and $\varphi \in \mathscr{C}^\alpha$, $\mathcal{P}_N\psi \in \mathscr{C}^\beta$. Then*

$$\left\|B_N(\varphi \prec \mathcal{P}_N\psi) - \varphi \prec \overline{B}_N(\mathcal{P}_N\psi)\right\|_{\alpha+\beta} \lesssim \|\varphi\|_\alpha\|\mathcal{P}_N\psi\|_\beta.$$

Proof. By spectral support properties it suffices to control

$$\begin{aligned}
&\left|\left(B_N(S_{j-1}\varphi, \Delta_j\mathcal{P}_N\psi) - S_{j-1}\varphi\overline{B}_N(\Delta_j\mathcal{P}_N\psi)\right)(x)\right| \\
&\qquad = \left|\int_{\mathbb{R}^2}(S_{j-1}\varphi(x+\varepsilon y) - S_{j-1}\varphi(x))\Delta_j\mathcal{P}_N\psi(x+\varepsilon z)\mu(\mathrm{d}y, \mathrm{d}z)\right| \\
&\qquad \lesssim \varepsilon 2^{j(1-\alpha-\beta)}\|\varphi\|_\alpha\|\mathcal{P}_N\psi\|_\beta.
\end{aligned}$$

But now $\Delta_j\mathcal{P}_N \equiv 0$ unless $2^j \lesssim N$, and therefore we may estimate $\varepsilon 2^j \lesssim 1$. □

Estimates for the discrete Laplacian.

Lemma 8.12. *Let π satisfy (H_f). Let $\alpha < 1$, $\beta \in \mathbb{R}$, and let $\varphi \in \mathscr{C}^\alpha$ and $\psi \in \mathscr{C}^\beta$ both have spectral support in $(-N/2, N/2)$. Then for all $\delta \in [0,1]$ and $N \in \mathbb{N}$*

$$\|\Delta_N\psi - \Delta\psi\|_{\beta-2-\delta} \lesssim N^{-\delta}\|\psi\|_\beta \qquad \text{and for } \delta > 0 \text{ also}$$

$$\|\Delta_N\Pi_N B_N(\varphi \prec \psi) - \Pi_N B_N(\varphi \prec \Delta_N\psi)\|_{\alpha+\beta-2-\delta} \lesssim \|\varphi\|_\alpha\|\psi\|_\beta.$$

Proof. As π has zero mass, zero first moment, and finite second moment, we get

$$\begin{aligned}
|\Delta_j \Delta_N \psi(x)| &= \left| \varepsilon^{-2} \int_{\mathbb{R}} \Delta_j \psi(x+\varepsilon y) \pi(\mathrm{d}y) \right| \\
&= \left| \varepsilon^{-2} \int_{\mathbb{R}} (\Delta_j \psi(x+\varepsilon y) - \Delta_j \psi(x) - \mathrm{D}\Delta_j \psi(x) \varepsilon y) \pi(\mathrm{d}y) \right| \\
&\leqslant \|\mathrm{D}^2 \Delta_j \psi\|_{L^\infty} \int_{\mathbb{R}} y^2 |\pi|(\mathrm{d}y) \lesssim 2^{j(2-\beta)} \|\psi\|_\beta.
\end{aligned}$$

On the other hand we use that $\int_{\mathbb{Z}} y^2 \pi(\mathrm{d}y) = 2$ to obtain

$$\begin{aligned}
|\Delta_j(\Delta_N \psi - \Delta \psi)(x)| &= \left| \varepsilon^{-2} \int_{\mathbb{R}} (\Delta_j \psi(x+\varepsilon y) - \frac{1}{2} \Delta \Delta_j \psi(x) \varepsilon^2 y^2) \pi(\mathrm{d}y) \right| \\
&\leqslant \varepsilon \|\mathrm{D}^3 \Delta_j \psi\|_{L^\infty} \int_{\mathbb{R}} |y|^3 |\pi|(\mathrm{d}y) \lesssim 2^{-j(\beta-3)} \varepsilon \|\psi\|_\beta,
\end{aligned}$$

and thus by interpolation $\|\Delta_N \psi - \Delta \psi\|_{\beta-2-\delta} \lesssim \varepsilon^\delta \|\psi\|_\beta$.

For our second claim we use that φ and ψ have spectral support in $(-N/2, N/2)$, which means that there are unique lattice functions $\tilde{\varphi}$ and $\tilde{\psi}$ such that $\varphi = \mathcal{E}_N \tilde{\varphi}$ and $\psi = \mathcal{E}_N \tilde{\psi}$. For these lattice functions we get

$$\Delta_N \Pi_N B_N(\varphi, \psi) = \Delta_N \Pi_N B_N(\mathcal{E}_N \tilde{\varphi}, \mathcal{E}_N \tilde{\psi}) = \Delta_N \mathcal{E}_N B_N(\tilde{\varphi}, \tilde{\psi}) = \mathcal{E}_N \Delta_N B_N(\tilde{\varphi}, \tilde{\psi}),$$

and a direct computation shows that

$$\Delta_N B_N(\tilde{\varphi}, \tilde{\psi}) = B_N(\Delta_N \tilde{\varphi}, \tilde{\psi}) + B_N(\tilde{\varphi}, \Delta_N \tilde{\psi}) + \varepsilon^{-2} \int B_N(\tau_{-\varepsilon y} \tilde{\varphi} - \tilde{\varphi}, \tau_{-\varepsilon y} \tilde{\psi} - \tilde{\psi}) \pi(\mathrm{d}y),$$

from where we obtain

$$\begin{aligned}
\Delta_N \Pi_N B_N(\varphi, \psi) = {} & \Pi_N B_N(\Delta_N \varphi, \psi) + \Pi_N B_N(\varphi, \Delta_N \psi) \\
& + \varepsilon^{-2} \int \Pi_N B_N(\tau_{-\varepsilon y} \varphi - \varphi, \tau_{-\varepsilon y} \psi - \psi) \pi(\mathrm{d}y).
\end{aligned}$$

so also for the paraproduct

$$\begin{aligned}
\Delta_N \Pi_N B_N(\varphi \prec \psi) = {} & \Pi_N B_N(\varphi \prec \Delta_N \psi) + \Pi_N B_N(\Delta_N \varphi \prec \psi) \\
& + \varepsilon^{-2} \int \Pi_N B_N((\tau_{-\varepsilon y} \varphi - \varphi) \prec (\tau_{-\varepsilon y} \psi - \psi)) \pi(\mathrm{d}y).
\end{aligned}$$

Now $\|\Pi_N B_N(\Delta_N \varphi \prec \psi)\|_{\alpha+\beta-2-\delta} \lesssim \|\Delta_N \varphi\|_{\alpha-2} \|\psi\|_\beta$ by Lemma 8.10 and Remark 8.9, and $\|\Delta_N \varphi\|_{\alpha-2} \lesssim \|\varphi\|_\alpha$ by our first bound. For the remainder term we use once more Remark 8.9 to get

$$\left\| \varepsilon^{-2} \int \Pi_N B_N((\tau_{-\varepsilon y} \varphi - \varphi) \prec (\tau_{-\varepsilon y} \psi - \psi)) \pi(\mathrm{d}y) \right\|_{\alpha+\beta-2-\delta}$$

$$\lesssim \varepsilon^{-2} \int \| B_N((\tau_{-\varepsilon y}\varphi - \varphi) \prec (\tau_{-\varepsilon y}\psi - \psi)) \|_{\alpha+\beta-2} \pi(\mathrm{d}y).$$

Now we apply Lemma 8.10 to bound

$$\begin{aligned} \| B_N((\tau_{-\varepsilon y}\varphi - \varphi) \prec (\tau_{-\varepsilon y}\psi - \psi)) \|_{\alpha+\beta-2} &\lesssim \| \tau_{-\varepsilon y}\varphi - \varphi \|_{\alpha-1} \| \tau_{-\varepsilon y}\psi - \psi \|_{\beta-1} \\ &\lesssim (\varepsilon |y|)^2 \| \varphi \|_\alpha \| \psi \|_\beta, \end{aligned}$$

from where the claimed bound readily follows. □

While the semigroup generated by the discrete Laplacian Δ_N does not have good regularizing properties, we will only apply it to functions with spectral support contained in $[-N/2, N/2]$, where it has the same regularizing effect as the heat flow. It is here that we need the assumption that $f(x) \geqslant c_f > 0$ for $x \in [-\pi, \pi]$.

Lemma 8.13. *Assume that π satisfies (H_f). Let $\alpha \in \mathbb{R}$, $\beta \geqslant 0$, and let $\varphi \in \mathscr{D}'$ with* $\operatorname{supp}(\mathscr{F}\varphi) \subset [-N/2, N/2]$. *Then we have for all $T > 0$ uniformly in $t \in (0, T]$*

$$\| e^{t\Delta_N} \varphi \|_{\alpha+\beta} \lesssim t^{-\beta/2} \| \varphi \|_\alpha. \tag{77}$$

If $\alpha < 0$, then we also have

$$\| e^{t\Delta_N} \varphi \|_{L^\infty} \lesssim t^{\alpha/2} \| \varphi \|_\alpha$$

Proof. Let χ be a compactly supported smooth function with $\chi \equiv 1$ on $[-\pi, \pi]$ and such that $f(x) \geqslant c_f/2$ for all $x \in \operatorname{supp}(\chi)$. Then $e^{t\Delta_N}\varphi = e^{-tf(\varepsilon \mathrm{D})\mathrm{D}^2} \chi(\varepsilon \mathrm{D}) \varphi =: \psi_\varepsilon(\mathrm{D})\varphi$. According to Lemma 2.2 in [BCD11], it suffices to show that

$$\max_{k=0,1,2} \sup_{x \in \mathbb{R}} |\mathrm{D}^k \psi_\varepsilon(x)| t^{\beta/2} |x|^{\beta+k} \leqslant C < \infty,$$

uniformly in $\varepsilon \in (0, 1]$. For ψ_ε itself we have

$$|\psi_\varepsilon(x)| t^{\beta/2} |x|^\beta \lesssim e^{-\frac{c_f}{2} |\sqrt{t} x|^2} |\sqrt{t} x|^\beta \lesssim 1.$$

To calculate the derivatives, note that

$$\begin{aligned} (e^{-\varphi(x)} \rho(x))' &= e^{-\varphi(x)} [-\varphi'(x)\rho(x) + \rho'(x)], \\ (e^{-\varphi(x)} \rho(x))'' &= e^{-\varphi(x)} [\varphi'(x)^2 \rho(x) - 2\varphi'(x)\rho'(x) - \varphi''(x)\rho(x) + \rho''(x)]. \end{aligned}$$

In our case we set $\varphi_\varepsilon(x) = -tf(\varepsilon x)x^2$ and $\rho_\varepsilon(x) = \chi(\varepsilon x)$, and obtain

$$\varphi'_\varepsilon(x) = -tx(f'(\varepsilon x)\varepsilon x + 2f(\varepsilon x)), \qquad \varphi''_\varepsilon(x) = -t(f''(\varepsilon x)(\varepsilon x)^2 + 4f'(\varepsilon x)\varepsilon x + 2f(\varepsilon x)).$$

Since $|\varepsilon x \chi(\varepsilon x)| \lesssim 1$ and similarly for $(\varepsilon x)^2 \chi(\varepsilon x)$ and $\varepsilon x \chi'(\varepsilon x)$, we get

$$|x| |\varphi'_\varepsilon(x) \rho_\varepsilon(x)| \lesssim |\sqrt{t} x|^2 \mathbb{1}_{\chi(\varepsilon x) \neq 0} \qquad \text{and} \qquad |x| |\rho'_\varepsilon(x)| \lesssim \mathbb{1}_{\chi(\varepsilon x) \neq 0},$$

from where we deduce that

$$|\mathrm{D}_x \psi_\varepsilon(x)| t^{\beta/2} |x|^{\beta+1} \lesssim e^{-\frac{c_f}{2} (\sqrt{t} x)^2} \left(|\sqrt{t} x|^{2+\beta} + |\sqrt{t} x|^\beta \right) \lesssim 1.$$

Similar arguments show that also $|\mathrm{D}_x^2 \psi_\varepsilon(x)| t^{\beta/2} |x|^{\beta+2} \lesssim 1$, which concludes the proof of (77). The L^∞ estimate now follows from an interpolation argument. □

Corollary 8.14. *Let π satisfy (H_f), let $\alpha \in (0,2)$, and let $\varphi \in \mathscr{C}^\alpha$ with spectral support in $[-N/2, N/2]$. Then*

$$\|(e^{t\Delta_N} - \mathrm{id})\varphi\|_{L^\infty} \lesssim t^{\alpha/2}\|\varphi\|_\alpha.$$

Proof. By definition of $e^{t\Delta_N}$ we have

$$\|(e^{t\Delta_N} - \mathrm{id})\varphi\|_{L^\infty} \leqslant \int_0^t \|e^{s\Delta_N}\Delta_N\varphi\|_{L^\infty}\mathrm{d}s \lesssim \int_0^t s^{-(2-\alpha)/2}\|\Delta_N\varphi\|_{\alpha-2}\mathrm{d}s \lesssim t^{\alpha/2}\|\varphi\|_\alpha,$$

where we used Lemma 8.13 and Lemma 8.12 in the second step. □

Combining Lemma 8.13 and Corollary 8.14, we can apply the same arguments as in the continuous setting to derive analogous Schauder estimates for $(e^{t\Delta_N})$ as in Lemma 2.9 or Lemma 6.6 – of course always restricted to elements of $\mathscr{S}'$ that are spectrally supported in $[-N/2, N/2]$.

Estimate for the discrete derivative.

Lemma 8.15. *Let ν satisfy (H_g). Let $\alpha \in (0,1)$, $\beta \in \mathbb{R}$, and let $\varphi \in \mathscr{C}^\alpha$ and $\psi \in \mathscr{C}^\beta$. Then for all $\delta \in [0,1]$ and $N \in \mathbb{N}$*

$$\|\mathrm{D}_N\psi - \mathrm{D}\psi\|_{\beta-1-\delta} \lesssim N^{-\delta}\|\psi\|_\beta \qquad \text{and for } \delta > 0 \text{ also}$$

$$\|\mathrm{D}_N\Pi_N B_N(\varphi \prec \psi) - \Pi_N B_N(\varphi \prec \mathrm{D}_N\psi)\|_{\alpha+\beta-1-\delta} \lesssim \|\varphi\|_\alpha\|\psi\|_\beta.$$

The proof is the same as the one of Lemma 8.12, and we omit it.

Lemma 8.16. *Let μ satisfy (H_h). Let $\alpha+\beta+\gamma > 0$, $\beta+\gamma < 0$, assume that $\alpha \in (0,1)$, and let $\varphi \in \mathscr{C}^\alpha$, $\mathcal{P}_N\psi \in \mathscr{C}^\beta$, $\chi \in \mathscr{C}^\gamma$. Define the operator*

$$A_N^{\psi,\chi}(\varphi) := \int (\Pi_N(\Pi_N(\varphi \prec \tau_{-\varepsilon y}\mathcal{P}_N\psi)) \circ \tau_{-\varepsilon z}\chi) - \mathcal{P}_N((\varphi \prec \tau_{-\varepsilon y}\mathcal{P}_N\psi) \circ \tau_{-\varepsilon z}\chi))\mu(\mathrm{d}y, \mathrm{d}z).$$

Then for all $\delta \in [0, \alpha+\beta+\gamma)$

$$\begin{aligned}&\|\Pi_N B_N(\Pi_N(\varphi \prec \mathcal{P}_N\psi) \circ \chi) - \mathcal{P}_N C(\varphi, \mathcal{P}_N\psi, \chi) - \mathcal{P}_N(\varphi B_N(\mathcal{P}_N\psi \circ \chi))\|_{\beta+\gamma} \\ &\qquad \lesssim N^{-\delta}\log(N)^2\|\varphi\|_\alpha\|\mathcal{P}_N\psi\|_\beta\|\gamma\|_\gamma + \|A_N^{\psi,\chi}\|_{L(\mathscr{C}^\alpha, \mathscr{C}^{\beta+\gamma})}\|\varphi\|_\alpha,\end{aligned}$$

where C is the commutator of Remark 2.4 and $L(U,V)$ denotes the space of bounded operators between the Banach spaces U and V, equipped with the operator norm.

Proof. We decompose the difference as follows:

$$\begin{aligned}&\|\Pi_N B_N(\Pi_N(\varphi \prec \mathcal{P}_N\psi) \circ \chi) - \mathcal{P}_N C(\varphi, \mathcal{P}_N\psi, \chi) - \mathcal{P}_N(\varphi B_N(\mathcal{P}_N\psi \circ \chi))\|_{\beta+\gamma} \\ &\quad \leqslant \left\|\Pi_N\Big(B_N(\Pi_N(\varphi \prec \mathcal{P}_N\psi) \circ \chi) - \int (\Pi_N(\varphi \prec \tau_{-\varepsilon y}\mathcal{P}_N\psi)) \circ \tau_{-\varepsilon z}\chi\mu(\mathrm{d}y, \mathrm{d}z)\Big)\right\|_{\beta+\gamma} \\ &\qquad + \|A_N^{\psi,\chi}(\varphi)\|_{\beta+\gamma-\delta} + \int \|\mathcal{P}_N(C(\varphi, \tau_{-\varepsilon y}\mathcal{P}_N\psi, \tau_{-\varepsilon z}\chi) - C(\varphi, \mathcal{P}_N\psi, \chi))\|_{\beta+\gamma}\mu(\mathrm{d}y, \mathrm{d}z).\end{aligned}$$

For the first term on the right hand side the same arguments as in the proof of Lemma 8.11 show that for all $\delta \in [0,1]$ with $\alpha + \beta + \gamma - \delta > 0$

$$\left\| \Pi_N \Big(B_N(\Pi_N(\varphi \prec \mathcal{P}_N \psi)) \circ \chi) - \int (\Pi_N(\varphi \prec \tau_{-\varepsilon y} \mathcal{P}_N \psi)) \circ \tau_{-\varepsilon z} \chi) \mu(\mathrm{d}y, \mathrm{d}z) \Big) \right\|_{\alpha+\beta+\gamma-\delta} \lesssim N^{-\delta} \log(N)^2 \|\varphi\|_\alpha \|\mathcal{P}_N \psi\|_\beta \|\chi\|_\gamma$$

(the factor $\log(N)^2$ is due to the operator Π_N which appears twice, see Lemma 8.8 and Remark 8.9). The second term is trivial to bound, and for the last term we simply use that

$$\|\tau_{-\varepsilon y} u - u\|_{\kappa-\delta} \lesssim N^{-\delta} |y|^\delta \|u\|_\kappa$$

whenever $\kappa \in \mathbb{R}$, $u \in \mathscr{C}^\kappa$, and $\delta \in [0,1]$. □

8.2 Paracontrolled analysis of the discrete equation

We now have all tools at hand that are required to describe the paracontrolled structure of the solution u_N to equation (76) which as we recall is given by

$$\mathscr{L}_N u_N = \mathrm{D}_N \Pi_N B_N(u_N, u_N) + \mathrm{D}_N \mathcal{P}_N \xi, \qquad u_N(0) = \mathcal{P}_N u_0^N.$$

We set

$$\mathbb{X}_N(\xi) = (X_N(\xi), X_N^{\vee}(\xi), X_N^{\curlyvee}(\xi), X_N^{\curlyvee\circ}(\xi), X_N^{\vee\vee}(\xi), B_N(Q_N \circ X_N)(\xi)),$$

where

$$\begin{aligned} \mathscr{L} X_N(\xi) &= \mathrm{D}_N \mathcal{P}_N \xi, \\ \mathscr{L} X_N^{\vee}(\xi) &= \mathrm{D}_N \Pi_N B_N(X_N(\xi), X_N(\xi)), \\ \mathscr{L} X_N^{\curlyvee}(\xi) &= \mathrm{D}_N \Pi_N B_N(X_N(\xi), X_N^{\vee}(\xi)), \\ \mathscr{L} X_N^{\curlyvee\circ}(\xi) &= \mathrm{D}_N \Pi_N B_N(X_N^{\curlyvee}(\xi) \circ X_N(\xi)), \\ \mathscr{L} X_N^{\vee\vee}(\xi) &= \mathrm{D}_N \Pi_N B_N(X_N^{\vee}(\xi), X_N^{\vee}(\xi)), \\ \mathscr{L} Q_N(\xi) &= \mathrm{D}_N B_N(X_N(\xi), 1), \end{aligned} \tag{78}$$

all with zero initial conditions except $X_N(\xi)$ for which we choose the "stationary" initial condition

$$X_N(\xi)(0) = \int_{-\infty}^0 e^{-s f_\varepsilon |\cdot|^2}(\mathrm{D}) \mathrm{D}_N \mathcal{P}_N \xi(s) \mathrm{d}s.$$

As in Section 6.2 we fix $\alpha \in (1/3, 1/2)$ and $\beta \in (1-\alpha, 2\alpha)$.

Theorem 8.17. *Let $(\mathbb{X}_N)$ be as in (88) and assume that the sequence is uniformly bounded in $C_T \mathscr{C}^{\alpha-1} \times C_T \mathscr{C}^{2\alpha-1} \times \mathscr{L}_T^\alpha \times \mathscr{L}_T^{2\alpha} \times \mathscr{L}_T^{2\alpha} \times C_T \mathscr{C}^{2\alpha-1}$ for all $T > 0$, and converges to*

$$(X, X^{\vee}, X^{\curlyvee} + 2cQ, X^{\curlyvee\circ} + cQ^{\vee} + 2cQ^{Q \circ X}, X^{\vee\vee}, Q \circ X + c)$$

in $C_T\mathscr{C}^{\alpha-1} \times C_T\mathscr{C}^{2\alpha-1} \times \mathscr{L}_T^{\alpha} \times \mathscr{L}_T^{2\alpha} \times \mathscr{L}_T^{2\alpha} \times C([\delta, T], \mathscr{C}^{2\alpha-1})$ *for all* $0 < \delta < T$*, where* $c \in \mathbb{R}$*,*

$$\mathbb{X} = (X, X^{\vee}, X^{\curlyvee}, X^{\overset{\vee}{\curlyvee}}, X^{\vee\!\curlyvee}, Q \circ X) \in \mathbb{X}_{\mathrm{rbe}},$$

and

$$\mathscr{L}Q^{\vee} = \mathrm{D}X^{\vee}, \qquad \mathscr{L}Q^{Q \circ X} = \mathrm{D}(Q \circ X),$$

both with 0 *initial condition. Assume also that the operator* $A_N = A_N^{Q_N, X_N}$ *in Lemma 8.16 converges to 0 in* $C_T(L(\mathscr{C}^{\bar\alpha}, \mathscr{C}^{2\bar\alpha-1}))$ *for all* $T > 0$*,* $\bar\alpha \in (1/3, \alpha)$*. Finally, let* $(\mathcal{P}_N u_0^N), u_0 \in \mathscr{C}^{-\beta}$ *and assume that* $\lim_N \|\mathcal{P}_N u_0^N - u_0\|_{-\beta} = 0$*.*

Let (u_N) *be the solution to* (76) *and let* $d^*_{T,-\beta}$ *be the metric defined in* (73)*. Then for all* $T > 0$ *we have* $\lim_N d^*_{T,-\beta}(u_N, u) = 0$*, where* $u \in \mathscr{D}^{\exp}_{\mathrm{rbe},\mathbb{X}}$ *(see Definition 6.9) is the unique solution to*

$$\mathscr{L}u = \mathrm{D}u^2 + 4c\mathrm{D}u + \mathrm{D}\xi, \qquad u(0) = u_0. \tag{79}$$

Remark 8.18. *According to Remark 7.6, equation* (79) *has a unique solution in* $\mathscr{D}^{\exp}_{\mathrm{rbe},\mathbb{X}}$ *which does not blow up. In particular we get* $\lim_N T_N^* = \infty$ *for the blow-up time* $T_N^* := \sup\{t \geqslant 0 : \|u_N\|_{C_t\mathscr{C}^{-\beta}} < \infty\}$ *of* u_N*, even if for fixed* N *we cannot guarantee that* u_N *stays finite for all times.*

Proof. Throughout the proof we fix $\bar\alpha \in (1/3, \alpha)$ and we define $\gamma_\delta = (\beta + \delta)/2$ whenever $\delta \geqslant 0$. We would like to perform a paracontrolled analysis of the equation, working in spaces modeled after the $\mathscr{D}^{\exp}_{\mathrm{rbe}}$ of Definition 6.9. For that purpose, we decompose the nonlinearity as follows:

$$\begin{aligned}
\mathrm{D}_N \Pi_N B_N(u_N, u_N) = {} & \mathscr{L}_N(X_N^{\vee} + 2X_N^{\curlyvee} + 4X_N^{\overset{\vee}{\curlyvee}} + X_N^{\vee\!\curlyvee}) + 4\mathrm{D}_N\Pi_N B_N(X_N^{\curlyvee} \prec X_N) \\
& + 4\mathrm{D}_N\Pi_N B_N(X_N^{\curlyvee} \succ X_N) + 2\mathrm{D}_N\Pi_N B_N(u_N^Q, X_N) \\
& + 2\mathrm{D}_N\Pi_N B_N(X_N^{\vee}, 2X_N^{\curlyvee} + u_N^Q) \\
& + \mathrm{D}_N\Pi_N B_N(2X_N^{\curlyvee} + u_N^Q, 2X_N^{\curlyvee} + u_N^Q).
\end{aligned}$$

The term $\mathrm{D}_N\Pi_N B_N(u_N^Q, X_N)$ can be further decomposed as

$$\begin{aligned}
\mathrm{D}_N\Pi_N B_N(u_N^Q, X_N) = {} & \mathrm{D}_N\Pi_N B_N(u_N^Q \prec X_N) + \mathrm{D}_N\Pi_N B_N(u_N^Q \succ X_N) \\
& + \mathrm{D}_N\Pi_N B_N(u_N^Q \circ X_N),
\end{aligned}$$

and the critical term is of course $\mathrm{D}_N\Pi_N B_N(u_N^Q \circ X_N)$. Using Lemma 8.15, Lemma 8.11, Lemma 8.7, and Remark 8.9 we have for all $T > 0$

$$\begin{aligned}
& \|\mathrm{D}_N\Pi_N B_N((2u_N^Q + 4X_N^{\curlyvee}) \prec X_N) - \Pi_N((2u_N^Q + 4X_N^{\curlyvee}) \prec\!\!\prec \overline{B}_N(\mathrm{D}_N X_N))\|_{\mathscr{M}_T^{\gamma_{\bar\alpha}}\mathscr{C}^{2\bar\alpha-2}} \\
& \quad \lesssim (\|u_N^Q\|_{\mathscr{L}_\infty^{\gamma_{\bar\alpha},\bar\alpha}(T)} + \|X_N^{\curlyvee}\|_{\mathscr{L}_T^{\alpha}})\|X_N\|_{\alpha-1}.
\end{aligned}$$

However, the Fourier cutoff operator Π_N does not commute with the paraproduct (at least not allowing for bounds that are uniform in N), and in particular u_N^Q is not paracontrolled. Rather, we have

$$u_N^Q = \Pi_N(u_N' \prec\!\!\prec Q_N) + u_N^\sharp$$

with

$$u_N' = 2u_N^Q + 4X_N^{\mathsf{Y}} \in \mathscr{L}_\infty^{\gamma_{\bar\alpha},\bar\alpha}(T), \qquad u_N^\sharp \in \mathscr{L}_\infty^{\gamma_{2\bar\alpha},2\bar\alpha}(T).$$

This means that we need an additional ingredient beyond the paracontrolled tools in order to control the term $\mathrm{D}_N\Pi_N B_N(\Pi_N(u_N' \prec Q_N)\circ X_N)$, and it is here that we need our assumption on the operator A_N. Under this assumption we can apply Lemma 8.16 to write

$$\mathrm{D}_N\Pi_N B_N(\Pi_N(u_N' \prec Q_N)\circ X_N) = R_N + \mathcal{P}_N C(u_N', Q_N, X_N) + \mathcal{P}_N(u_N' B_N(Q_N\circ X_N)), \tag{80}$$

where R_N is a term that converges to zero in $\mathcal{M}_T^{\gamma_{\bar\alpha}}\mathscr{C}^{2\bar\alpha-2}$ if u_N' stays uniformly bounded in $\mathscr{L}_\infty^{\gamma_{\bar\alpha},\bar\alpha}(T)$. Denote now

$$(\tilde X, \tilde X^{\mathsf{V}}, \tilde X^{\mathsf{Y}}, \tilde X^{\mathsf{YY}}, \tilde X^{\mathsf{W}}, \tilde\eta, \tilde Q) = \lim_N (X_N, X_N^{\mathsf{V}}, X_N^{\mathsf{Y}}, X_N^{\mathsf{YY}}, X_N^{\mathsf{W}}, B_N(Q_N\circ X_N), Q_N).$$

Based on the above representation of the nonlinearity, it is not difficult to repeat the arguments of Section 6.2 in order to show that $d^*_{T,-\beta}(u_N, u)$ converges to 0, where

$$u = \tilde X + \tilde X^{\mathsf{V}} + 2\tilde X^{\mathsf{Y}} + \tilde u^Q, \qquad \tilde u^Q = \tilde u' \prec\!\!\prec \tilde Q + \tilde u^\sharp,$$

with $(\tilde u^Q, \tilde u', \tilde u^\sharp) = \lim_N(u_N^Q, u_N', u_N^\sharp)$, where $u(0) = u_0$ and

$$\begin{aligned}\mathscr{L}\tilde u^Q = {}& \mathscr{L}\tilde X^{\mathsf{W}} + 4\mathscr{L}\tilde X^{\mathsf{YY}} + 2\mathrm{D}((2\tilde X^{\mathsf{Y}} + \tilde u^Q)\prec \tilde X) + 2\mathrm{D}((2\tilde X^{\mathsf{Y}} + \tilde u^Q)\succ \tilde X) \\ & + 2\mathrm{D}(u^\sharp\circ\tilde X) + 2\mathrm{D}((\tilde u' \prec\!\!\prec \tilde Q - \tilde u'\prec\tilde Q)\circ\tilde X) + 2\mathrm{D}C(\tilde u', \tilde Q, \tilde X) + 2\mathrm{D}(\tilde u'\tilde\eta) \\ & + 2\mathrm{D}(\tilde X^{\mathsf{V}}(2\tilde X^{\mathsf{Y}} + \tilde u^Q)) + \mathrm{D}(2\tilde X^{\mathsf{Y}} + \tilde u^Q)^2. \end{aligned} \tag{81}$$

We also have

$$\tilde u' = 2\tilde u^Q + 4\tilde X^{\mathsf{Y}}.$$

The fact that $(B_N(Q_N\circ X_N))$ converges not uniformly but only uniformly on intervals $[\delta, T]$ for $\delta > 0$ poses no problem because the sequence is uniformly bounded, so that given $\kappa > 0$ we can fix a small $\delta > 0$ with $\sup_N \|u_N(\delta) - \mathcal{P}_N u_0^N\|_{-\beta} < \kappa$, and then use the uniform convergence of the data on $[\delta, T]$ and the convergence of $(\mathcal{P}_N u_0^N)$ to u_0.

Now observe that

$$u = X + X^{\mathsf{V}} + 2X^{\mathsf{Y}} + u^Q$$

with $u^Q - \tilde u^Q = 2\tilde X^{\mathsf{Y}} - 2X^{\mathsf{Y}} = 4cQ$, and moreover

$$\tilde u' = 2\tilde u^Q + 4X^{\mathsf{Y}} + 8cQ = 2u^Q + 4X^{\mathsf{Y}}.$$

Plugging this as well as the specific form of the $\tilde{X}^{\tau}$ into (81), we get

$$
\begin{aligned}
\mathscr{L}u^Q = {} & 4c\mathrm{D}X + \mathscr{L}X^{\vee\vee} + 4\mathscr{L}X^{\curlyvee'} + 4c\mathrm{D}X^{\vee} + 8c\mathrm{D}(Q \circ X) + 2\mathrm{D}((2X^{\curlyvee} + u^Q) \prec X) \\
& + 2\mathrm{D}((2X^{\curlyvee} + u^Q) \succ X) + 2\mathrm{D}(u^{\sharp} \circ X) + 2\mathrm{D}((\tilde{u}' \prec\!\!\prec Q - \tilde{u}' \prec Q) \circ X) \\
& + 2\mathrm{D}C(\tilde{u}', Q, X) + 2\mathrm{D}(\tilde{u}'(Q \circ X + c)) + 2\mathrm{D}(X^{\vee}(2X^{\curlyvee} + u^Q)) + \mathrm{D}(2X^{\curlyvee} + u^Q)^2.
\end{aligned}
$$

Since

$$
\begin{aligned}
8c\mathrm{D}(Q \circ X) + 2\mathrm{D}(u^{\sharp} \circ X) + 2\mathrm{D}((\tilde{u}' \prec\!\!\prec Q - \tilde{u}' \prec Q) \circ X) + 2\mathrm{D}C(\tilde{u}', Q, X) + 2\mathrm{D}(\tilde{u}'(Q \circ X)) \\
= 8c\mathrm{D}(Q \circ X) + 2\mathrm{D}(\tilde{u}^Q \circ X) = 2\mathrm{D}(u^Q \circ X),
\end{aligned}
$$

we end up with

$$
\begin{aligned}
\mathscr{L}u^Q &= \mathrm{D}u^2 - \mathscr{L}X - \mathscr{L}X^{\vee} - 2\mathscr{L}X^{\curlyvee} + 4c\mathrm{D}X + 4c\mathrm{D}X^{\vee} + 2c\mathrm{D}\tilde{u}' \\
&= \mathrm{D}u^2 - \mathscr{L}X - \mathscr{L}X^{\vee} - 2\mathscr{L}X^{\curlyvee} + 4c\mathrm{D}u,
\end{aligned}
$$

which completes the proof. □

It remains to study the convergence of the discrete stochastic data, which will be done in Section 10 below.

9 The stochastic driving terms

In this section we study the random fields

$$
X, X^{\vee}, X^{\curlyvee}, X^{\curlyvee'}, X^{\vee\vee}, Q \circ X
$$

which appear in the definition of $\mathbb{X} \in \mathcal{X}_{\mathrm{rbe}}$. Our main results are the following two theorems, whose proofs will cover the next subsections.

Theorem 9.1. *Let ξ be a space-time white noise on $\mathbb{R} \times \mathbb{T}$ and define*

$$
\begin{aligned}
\mathscr{L}X &= \mathrm{D}\xi, \\
\mathscr{L}X^{\vee} &= \mathrm{D}(X^2), \\
\mathscr{L}X^{\curlyvee} &= \mathrm{D}(XX^{\vee}), \\
\mathscr{L}X^{\curlyvee'} &= \mathrm{D}(X^{\curlyvee} \circ X), \\
\mathscr{L}X^{\vee\vee} &= \mathrm{D}(X^{\vee}X^{\vee}), \\
\mathscr{L}Q &= \mathrm{D}X,
\end{aligned}
$$

all with zero initial condition except X for which we choose the stationary initial condition

$$
X(0) = \int_{-\infty}^{0} P_{-s}\mathrm{D}\xi(s)\mathrm{d}s.
$$

Then almost surely $\mathbb{X} = (X, X^{\vee}, X^{\curlyvee}, X^{\overset{\curlyvee}{\vee}}, X^{\vee\vee}, Q \circ X) \in \mathcal{X}_{\mathrm{rbe}}$. *If* $\varphi: \mathbb{R} \to \mathbb{R}$ *is a measurable, bounded, even function of compact support, such that* $\varphi(0) = 1$ *and* φ *is continuous in a neighborhood of* 0, *and if*

$$\xi_\varepsilon = \mathscr{F}^{-1}(\varphi(\varepsilon \cdot) \mathscr{F} \xi) = \varphi(\varepsilon \mathrm{D}) \xi$$

(here $\mathscr{F}$ *denotes the spatial Fourier transform) and* $\mathbb{X}_\varepsilon = \Theta_{\mathrm{rbe}}(\xi_\varepsilon)$, *then for all* $T, p > 0$

$$\lim_{\varepsilon \to 0} \mathbb{E}[\| \mathbb{X} - \mathbb{X}_\varepsilon \|^p_{\mathcal{X}_{\mathrm{rbe}}(T)}] = 0.$$

Similarly, if $\tilde{\mathbb{X}}_\varepsilon = (\tilde{X}_\varepsilon, \tilde{X}^{\vee}_\varepsilon, \tilde{X}^{\curlyvee}_\varepsilon, \tilde{X}^{\overset{\curlyvee}{\vee}}_\varepsilon, \tilde{X}^{\vee\vee}_\varepsilon, \tilde{Q}_\varepsilon \circ \tilde{X}_\varepsilon)$ *for* $\tilde{X}_\varepsilon = \varphi(\varepsilon \mathrm{D}) X$ *and*

$$\begin{aligned} \mathscr{L} \tilde{X}^{\vee}_\varepsilon &= \mathrm{D}((\tilde{X}_\varepsilon)^2), \\ \mathscr{L} \tilde{X}^{\curlyvee}_\varepsilon &= \mathrm{D}(\tilde{X}_\varepsilon \varphi(\varepsilon \mathrm{D}) X^{\vee}), \\ \mathscr{L} \tilde{X}^{\overset{\curlyvee}{\vee}}_\varepsilon &= \mathrm{D}(\varphi(\varepsilon \mathrm{D}) X^{\curlyvee} \circ \tilde{X}_\varepsilon), \\ \mathscr{L} \tilde{X}^{\vee\vee}_\varepsilon &= \mathrm{D}(\varphi(\varepsilon \mathrm{D}) X^{\vee} \varphi(\varepsilon \mathrm{D}) X^{\vee}), \\ \mathscr{L} \tilde{Q}_\varepsilon &= \mathrm{D} X_\varepsilon, \end{aligned}$$

all with zero initial conditions, then for all $T, p > 0$

$$\lim_{\varepsilon \to 0} \mathbb{E}[\| \mathbb{X} - \tilde{\mathbb{X}}_\varepsilon \|^p_{\mathcal{X}_{\mathrm{rbe}}(T)}] = 0.$$

Remark 9.2. *The theorem would be easier to formulate if we assumed* φ *to be continuous, of compact support, and with* $\varphi(0) = 1$. *The reason why we chose the complicated formulation above is that we do not want to exclude the function* $\varphi(x) = \mathbb{1}_{[-1,1]}(x)$.

Theorem 9.3. *Let* ξ *be a space-time white noise on* $\mathbb{R} \times \mathbb{T}$. *Then there exists an element* $\mathbb{Y} \in \mathcal{Y}_{\mathrm{kpz}}$ *such that for every even, compactly supported function* $\varphi \in C^1(\mathbb{R}, \mathbb{R})$ *with* $\varphi(0) = 1$ *there are diverging constants* $c^{\vee}_\varepsilon$ *and* $c^{\vee\vee}_\varepsilon$ *for which*

$$\lim_{\varepsilon \to 0} \mathbb{E}[\| \mathbb{Y} - \mathbb{Y}_\varepsilon \|^p_{\mathcal{Y}_{\mathrm{kpz}}(T)}] = 0$$

for all $T, p > 0$, *where* $\mathbb{Y}_\varepsilon = \Theta_{\mathrm{kpz}}(\varphi(\varepsilon \mathrm{D}) \xi, c^{\vee}_\varepsilon, c^{\vee\vee}_\varepsilon)$. *Moreover, we have*

$$c^{\vee}_\varepsilon = \frac{1}{4\pi\varepsilon} \int_{\mathbb{R}} \varphi^2(x) \mathrm{d}x.$$

Remark 9.4. *We will only worry about the construction of* $\mathbb{X}$ *and* $\mathbb{Y}$. *The convergence result then follows easily from the dominated convergence theorem, because since* φ *is an even function all the symmetries in the kernels that we will use below also hold for the kernels corresponding to* $\mathbb{X}_\varepsilon$, $\tilde{\mathbb{X}}_\varepsilon$, $\mathbb{Y}_\varepsilon$.

9.1 Kernels

We can represent the white noise in terms of its spatial Fourier transform. More precisely, let $E = \mathbb{Z} \setminus \{0\}$ and let $\tilde{W}$ be a complex valued centered Gaussian process on $\mathbb{R} \times E$, such that $\tilde{W}(s,k)^* = \tilde{W}(s,-k)$ and with the covariance

$$\mathbb{E}\left[\int_{\mathbb{R}\times E} f(\eta)\tilde{W}(\mathrm{d}\eta)\int_{\mathbb{R}\times E} g(\eta')\tilde{W}(\mathrm{d}\eta')\right] = (2\pi)^{-1}\int_{\mathbb{R}\times E} g(\eta_1)f(\eta_{-1})\mathrm{d}\eta_1,$$

where $\eta_a = (s_a, k_a)$, $s_{-a} = s_a$, $k_{-a} = -k_a$ and the measure $\mathrm{d}\eta_a = \mathrm{d}s_a\mathrm{d}k_a$ is the product of the Lebesgue measure $\mathrm{d}s_a$ on $\mathbb{R}$ and of the counting measure $\mathrm{d}k_a$ on E. The functions f, g are complex valued and in $L^2(\mathbb{R}\times E)$. Then the process $\tilde{\xi}(\varphi) = \tilde{W}(\mathscr{F}\varphi)$, where $\varphi \in L^2(\mathbb{R}\times\mathbb{T})$ and $\mathscr{F}$ denotes the spatial Fourier transform, is a white noise on $L^2_0(\mathbb{R}\times\mathbb{T})$, the space of all L^2 functions φ with $\int_{\mathbb{T}}\varphi(x,y)\mathrm{d}y = 0$ for almost all x.

Convention: To eliminate many constants of the type $(2\pi)^p$ in the following calculations, let us rather work with $\sqrt{2\pi}\tilde{W}$, which we denote by the same symbol $\tilde{W}$. Of course all qualitative results that we prove for this transformed noise stay true for the original noise, and we only have to pay attention in Theorem 9.3 to get the constant $c_\varepsilon^{\mathsf{V}}$ right.

The process X then has the following representation as an integral

$$X(t,x) = \int_{\mathbb{R}\times E} e^{ikx}H_{t-s}(-k)\tilde{W}(\mathrm{d}\eta),$$

where $\eta = (s,k) \in \mathbb{R}\times E$ and

$$h_t(k) = e^{-k^2 t}\mathbb{1}_{t\geqslant 0}, \qquad H_t(k) = ikh_t(k).$$

This means that $H_{t-s}(-k) = -H_{t-s}(k)$, and it will simplify the notation if we work with $W = -\tilde{W}$ and $\xi = -\tilde{\xi}$, which of course have the same distribution as $\tilde{W}$ and $\tilde{\xi}$ and for which

$$X(t,x) = \int_{\mathbb{R}\times E} e^{ikx}H_{t-s}(k)W(\mathrm{d}\eta).$$

The space Fourier transform $\hat{X}(t,k) = \hat{X}_t(k)$ of $X(t,\cdot)$ reads

$$\hat{X}(t,k) = \int_{\mathbb{R}} 2\pi H_{t-s}(k)W_k(\mathrm{d}s),$$

where $W_{k'}(\mathrm{d}s) = \int_E \delta_{k,k'}W(\mathrm{d}s\mathrm{d}k)$ is just a countable family of complex time white noises satisfying $W_k(\mathrm{d}s)^* = W_{-k}(\mathrm{d}s)$ and $\mathbb{E}[W_{k'}(\mathrm{d}s)W_k(\mathrm{d}s')] = \delta_{k,-k'}\delta(s-s')\mathrm{d}s\mathrm{d}s'$.

Note that if $s \leqslant t$

$$\int_{\mathbb{R}} H_{s-\sigma}(k)H_{t-\sigma}(-k)\mathrm{d}\sigma = \frac{e^{-k^2|t-s|}}{2}, \tag{82}$$

from where we read the covariance of X:

$$\mathbb{E}[X(t,x)X(s,y)] = \mathbb{E}\left[\int_{\mathbb{R}\times E} e^{ik_1x}H_{t-s_1}(k_1)W(\mathrm{d}\eta_1)\int_{\mathbb{R}\times E} e^{ik_2y}H_{s-s_2}(k_2)W(\mathrm{d}\eta_2)\right]$$

$$= \int_E \mathrm{d}k_1 e^{ik_1(x-y)} \int_{\mathbb{R}} H_{t-s_1}(k_1) H_{s-s_1}(-k_1) \mathrm{d}s_1$$
$$= \int_E e^{ik_1(x-y)} \frac{e^{-k_1^2|t-s|}}{2} \mathrm{d}k_1 = \frac{1}{2} p_{|t-s|}(x-y),$$

where p_t is the kernel of the heat semigroup: $P_t f = p_t * f = \int_{\mathbb{T}} p_t(\cdot - y) f(y) \mathrm{d}y$. In Fourier space we have

$$\mathbb{E}[\hat{X}_t(k) \hat{X}_s(k')] = \delta_{k+k'=0} \frac{e^{-k^2|t-s|}}{2}$$

as expected.

These notations and preliminary results will be useful below in relation to the representation of elements in the chaos of W and the related Gaussian computations. Recall that $X^{\bullet} = X$ and $X^{(\tau_1\tau_2)} = B(X^{\tau_1}, X^{\tau_2})$. Then

$$X^{\tau}(t,x) = \int_{(\mathbb{R}\times E)^n} G^{\tau}(t,x,\eta_\tau) \prod_{i=1}^{n} W(\mathrm{d}\eta_i),$$

where $n = d(\tau)$, $\eta_\tau = \eta_{1\cdots n} = (\eta_1, \ldots, \eta_n) \in (\mathbb{R} \times E)^n$ and $\mathrm{d}\eta_\tau = \mathrm{d}\eta_{1\cdots n} = \mathrm{d}\eta_1 \cdots \mathrm{d}\eta_n$. Here we mean that each of the X^τ is a polynomial in the Gaussian variables $W(\mathrm{d}\eta_i)$, and in the next section we study how these polynomials decompose into the chaoses of W. For the moment we are interested in the analysis of the kernels G^τ involved in this representation. These kernels are defined recursively by

$$G^{\bullet}(t,x,\eta) = e^{ikx} H_{t-s}(k),$$

and then

$$G^{(\tau_1\tau_2)}(t,x,\eta_{(\tau_1\tau_2)}) = B(G^{\tau_1}(\cdot,\cdot,\eta_{\tau_1}), G^{\tau_2}(\cdot,\cdot,\eta_{\tau_2}))(t,x)$$
$$= \int_0^t \mathrm{d}\sigma \mathrm{D} P_{t-\sigma}(G^{\tau_1}(\sigma,\cdot,\eta_{\tau_1}) G^{\tau_2}(\sigma,\cdot,\eta_{\tau_2}))(x).$$

In the first few cases this gives

$$G^{\mathsf{V}}(t,x,\eta_{12}) = \int_0^t \mathrm{d}\sigma \mathrm{D} P_{t-\sigma}(G^{\bullet}(\sigma,\cdot,\eta_1) G^{\bullet}(\sigma,\cdot,\eta_2))(x)$$
$$= e^{ik_{[12]}x} \int_0^t \mathrm{d}\sigma H_{t-\sigma}(k_{[12]}) H_{\sigma-s_1}(k_1) H_{\sigma-s_2}(k_2),$$

where we set $k_{[1\cdots n]} = k_1 + \cdots + k_n$, and

$$G^{\mathsf{V\!\!\!\!\!\!\!V}}(t,x,\eta_{123}) = \int_0^t \mathrm{d}\sigma \mathrm{D} P_{t-\sigma}(G^{\mathsf{V}}(\sigma,\cdot,\eta_{12}) G^{\bullet}(\sigma,\cdot,\eta_3))(x)$$
$$= e^{ik_{[123]}x} \int_0^t \mathrm{d}\sigma H_{t-\sigma}(k_{[123]}) \left(\int_0^\sigma \mathrm{d}\sigma' H_{\sigma-\sigma'}(k_{[12]}) H_{\sigma'-s_1}(k_1) H_{\sigma'-s_2}(k_2) \right) H_{\sigma-s_3}(k_3).$$

In both cases, the kernel has the factorized form

$$G^{\tau}(t,x,\eta_{\tau}) = e^{ik_{[\tau]}x} H^{\tau}(t,\eta_{\tau}), \tag{83}$$

where we further denote $k_{[\tau]} = k_{[1\cdots n]} = k_1 + \cdots + k_n$, and this factorization holds for all G^{τ}. In fact, it is easy to show inductively that

$$G^{(\tau_1\tau_2)}(t,x,\eta_{(\tau_1\tau_2)}) = e^{ik_{[\tau]}x} \int_0^t \mathrm{d}\sigma H_{t-\sigma}(k_{[\tau]}) H^{\tau_1}(\sigma,\eta_{\tau_1}) H^{\tau_2}(\sigma,\eta_{\tau_2}),$$

from where we read

$$\begin{aligned} G^{\curlyvee}(t,x,\eta_{1234}) = e^{ik_{[1234]}x} &\int_0^t \mathrm{d}\sigma H_{t-\sigma}(k_{[1234]}) \int_0^{\sigma} \mathrm{d}\sigma' H_{\sigma-\sigma'}(k_{[123]}) \times \\ &\times \Big(\int_0^{\sigma'} \mathrm{d}\sigma'' H_{\sigma'-\sigma''}(k_{[12]}) H_{\sigma''-s_1}(k_1) H_{\sigma''-s_2}(k_2) \Big) H_{\sigma'-s_3}(k_3) H_{\sigma-s_4}(k_4) \end{aligned}$$

and thus

$$\begin{aligned} G^{\curlyvee_\circ}(t,x,\eta_{1234}) = e^{ik_{[1234]}x} &\int_0^t \mathrm{d}\sigma H_{t-\sigma}(k_{[1234]}) \psi_{\circ}(k_{[123]},k_4) \int_0^{\sigma} \mathrm{d}\sigma' H_{\sigma-\sigma'}(k_{[123]}) \times \\ &\times \Big(\int_0^{\sigma'} \mathrm{d}\sigma'' H_{\sigma'-\sigma''}(k_{[12]}) H_{\sigma''-s_1}(k_1) H_{\sigma''-s_2}(k_2) \Big) H_{\sigma'-s_3}(k_3) H_{\sigma-s_4}(k_4), \end{aligned}$$

where we recall that $\psi_{\circ}(k,\ell) = \sum_{|i-j|\leqslant 1} \rho_i(k)\rho_j(\ell)$. Similarly, we have

$$\begin{aligned} G^{\curlyvee\curlyvee}(t,x,\eta_{1234}) = e^{ik_{[1234]}x} &\int_0^t \mathrm{d}\sigma \int_0^{\sigma} \mathrm{d}\sigma' \int_0^{\sigma} \mathrm{d}\sigma'' H_{t-\sigma}(k_{[1234]}) H_{\sigma-\sigma'}(k_{[12]}) H_{\sigma-\sigma''}(k_{[34]}) \times \\ &\times H_{\sigma'-s_1}(k_1) H_{\sigma'-s_2}(k_2) H_{\sigma''-s_3}(k_3) H_{\sigma''-s_4}(k_4). \end{aligned}$$

9.2 Chaos decomposition and diagrammatic representation

The representation

$$X^{\tau}(t,x) = \int_{(\mathbb{R}\times E)^n} G^{\tau}(t,x,\eta_{\tau}) \prod_{i=1}^{n} W(\mathrm{d}\eta_i)$$

is not very useful for the analysis of the properties of the random fields X. It is more meaningful to separate the components in different chaoses. Denote by

$$\int_{(\mathbb{R}\times E)^n} f(\eta_{1\cdots n}) W(\mathrm{d}\eta_{1\cdots n})$$

a generic element of the n-th chaos of the white noise W on $(\mathbb{R}\times E)$. We find convenient not to symmetrize the kernels in the chaos decomposition. If we follow this convention then we should recall that the variance of the chaos elements will be given by

$$\mathbb{E}\left[\left| \int_{(\mathbb{R}\times E)^n} f(\eta_{1\cdots n}) W(\mathrm{d}\eta_{1\cdots n}) \right|^2 \right] = \sum_{\sigma\in\mathcal{S}_n} \int_{(\mathbb{R}\times E)^n} f(\eta_{1\cdots n}) f(\eta_{(-\sigma(1))\cdots(-\sigma(n))}) \mathrm{d}\eta_{1\cdots n}$$

where the sum runs over all the permutations $\mathcal{S}_n$ of $\{1, \dots, n\}$ and where we introduced the notation $\eta_{-1} = (s_1, -k_1)$ to describe the contraction of the Gaussian variables. By the Cauchy-Schwarz inequality

$$\mathbb{E}\left[\left|\int_{(\mathbb{R}\times E)^n} f(\eta_{1\cdots n}) W(\mathrm{d}\eta_{1\cdots n})\right|^2\right] \leqslant n! \int_{(\mathbb{R}\times E)^n} |f(\eta_{1\cdots n})|^2 \mathrm{d}\eta_{1\cdots n},$$

so that for the purpose of bounding the variance of the chaoses it is enough to bound the L^2 norm of the unsymmetrized kernels. The general formula for the chaos decomposition of a polynomial

$$\int_{(\mathbb{R}\times E)^n} f(\eta_{1\cdots n}) \prod_{i=1}^{n} W(\mathrm{d}\eta_i)$$

is given by

$$\int_{(\mathbb{R}\times E)^n} f(\eta_{1\cdots n}) \prod_{i=1}^{n} W(\mathrm{d}\eta_i) = \sum_{\ell=0}^{n} \int_{(\mathbb{R}\times E)^k} f_\ell(\eta_{1\cdots \ell}) W(\mathrm{d}\eta_{1\dots\ell})$$

with $f_\ell(\eta_{1\cdots\ell}) = 0$ if $n - \ell$ is odd. To give the expression for f_ℓ with $n - \ell$ even, let us introduce some notation: We write $\mathcal{S}(\ell, n)$ for the shuffle permutations of $\{1, \dots, n\}$ which leave the order of the first ℓ and the last $n - \ell$ terms intact. We also write

$$\eta \sqcup \upsilon = \eta_{1\dots\ell} \sqcup \upsilon_{1\dots m} = (\eta_1, \dots, \eta_\ell, \upsilon_1, \dots, \upsilon_m)$$

for the concatenation of η and υ. So if $n - \ell = 2m$ for some m, then

$$f_\ell(\eta_{1\dots\ell}) = 2^{-m} \sum_{\sigma\in\mathcal{S}(\ell,n)} \sum_{\substack{\tau\in\mathcal{S}(m,2m),\\ \tilde{\tau}\in\mathcal{S}(m)}} \int_{(\mathbb{R}\times E)^m} f(\sigma[\eta_{1\dots\ell}\sqcup\tau(\eta_{(\ell+1)\dots(\ell+m)}\sqcup\tilde{\tau}(\eta_{-(\ell+1)\dots-(\ell+m)}))])\mathrm{d}\eta_{(\ell+1)\dots(\ell+m)},$$

where $\sigma(\eta_{1\dots n}) = \eta_{\sigma(1)\dots\sigma(n)}$. This just means that we sum over all ways in which $2m$ of the $2m + \ell$ variables can be paired and integrated out. For example, we have

$$W(\mathrm{d}\eta_1)W(\mathrm{d}\eta_2) = W(\mathrm{d}\eta_{12}) + \delta(\eta_1 + \eta_{-2})\mathrm{d}\eta_1\mathrm{d}\eta_2.$$

We will denote by G_ℓ^τ the kernel of the ℓ-th chaos arising from the decomposition of X^τ:

$$X^\tau(t, x) = \sum_{\ell=0}^{n} \int_{(\mathbb{R}\times E)^\ell} G_\ell^\tau(t, x, \eta_{1\cdots\ell}) W(\mathrm{d}\eta_{1\cdots\ell}).$$

Terms X^τ of odd degree have zero mean by construction while the terms of even degree have zero mean due to the fact that if $d(\tau) = 2n$ we have

$$\mathbb{E}[X^\tau(t, x)] = 2^{-n} \sum_{\substack{\tau\in\mathcal{S}(n,2n)\\ \tilde{\tau}\in\mathcal{S}_n}} \int_{(\mathbb{R}\times E)^n} G^\tau(t, x, \tau(\eta_{1\cdots n} \sqcup \tilde{\tau}(\eta_{(-1)\cdots(-n)})))\mathrm{d}\eta_{1\dots n}.$$

But now $k_{[1\cdots n(-\tilde{\tau}(1))\cdots(-\tilde{\tau}(n))]} = k_1 + \cdots + k_n - k_{\tilde{\tau}(1)} \cdots - k_{\tilde{\tau}(n)} = 0$ and we always have $G^{\tau}(t, x, \eta_{1\cdots 2n}) \propto k_{[1\cdots(2n)]}$, which implies that

$$G^{\tau}(t, x, \tau(\eta_{1\cdots n} \sqcup \tilde{\tau}(\eta_{(-1)\cdots(-n)}))) = 0.$$

This is a special simplification of considering the stochastic Burgers equation instead of the KPZ equation. Later we will study the kernel functions for the KPZ equation to understand some subtle cancellations which appear in the terms belonging to the 0-th chaos.

Applying these considerations to the first nontrivial case given by $X^{\vee}$, we obtain

$$X^{\vee}(t, x) = \int_{(\mathbb{R}\times E)^2} G^{\vee}(t, x, \eta_{12}) W(\mathrm{d}\eta_1 \mathrm{d}\eta_2) + G_0^{\vee}(t, x)$$

with

$$G_0^{\vee}(t, x) = \int_{(\mathbb{R}\times E)^2} G^{\vee}(t, x, \eta_{1(-1)}) \mathrm{d}\eta_1.$$

But as already remarked

$$G^{\vee}(t, x, \eta_{1(-1)}) = e^{ik_{[1(-1)]}x} \int_0^t H_{t-\sigma}(k_{[1(-1)]}) H_{\sigma-s_1}(k_1) H_{\sigma-s_2}(k_{-1}) \mathrm{d}\sigma = 0$$

since $H_{t-\sigma}(0) = 0$. Consider the next term

$$X^{\curlyvee}(t, x) = \int_{(\mathbb{R}\times E)^3} G^{\curlyvee}(t, x, \eta_{123}) W(\mathrm{d}\eta_1 \mathrm{d}\eta_2 \mathrm{d}\eta_3) + \int_{\mathbb{R}\times E} G_1^{\curlyvee}(t, x, \eta_1) W(\mathrm{d}\eta_1).$$

In this case we have three possible contractions contributing to $G_1^{\curlyvee}$, which result in

$$G_1^{\curlyvee}(t, x, \eta_1) = \int_{\mathbb{R}\times E} (G^{\curlyvee}(t, x, \eta_{12(-2)}) + G^{\curlyvee}(t, x, \eta_{21(-2)}) + G^{\curlyvee}(t, x, \eta_{2(-2)1})) \mathrm{d}\eta_2.$$

But note that $G^{\curlyvee}(t, x, \eta_{2(-2)1}) = 0$ since, as above, this kernel is proportional to $k_{2(-2)} = 0$. Moreover, by symmetry $G^{\curlyvee}(t, x, \eta_{12(-2)}) = G^{\curlyvee}(t, x, \eta_{21(-2)})$ and we remain with

$$G_1^{\curlyvee}(t, x, \eta_1) = 2G^{\triangledown}(t, x, \eta_1) = 2\int_{\mathbb{R}\times E} G^{\curlyvee}(t, x, \eta_{12(-2)}) \mathrm{d}\eta_2,$$

where we introduced the intuitive notation $G^{\triangledown}(t, x, \eta_1)$ which is useful to keep track graphically of the Wick contractions of the kernels G^{τ} by representing them as arcs between leaves of the binary tree.

Now an easy computation gives

$$G^{\triangledown}(t, x, \eta_1) = e^{ik_1 x} \int_0^t \mathrm{d}\sigma \int_0^{\sigma} \mathrm{d}\sigma' H_{t-\sigma}(k_1) H_{\sigma'-s_1}(k_1) V^{\triangledown}(\sigma - \sigma', k_1),$$

where

$$V^{\tau_1}(\sigma, k_1) = \int H_\sigma(k_{[12]}) H_{\sigma - s_2}(k_2) H_{-s_2}(k_{-2}) \mathrm{d}\eta_2 = \int \mathrm{d}k_2 H_\sigma(k_{[12]}) \frac{e^{-|\sigma| k_2^2}}{2}.$$

We call the functions V_n^τ vertex functions. They are useful to compare the behavior of different kernels.

By similar arguments we establish the decomposition for the last two terms, that is

$$X^{\tau_2}(t,x) = \int_{(\mathbb{R}\times E)^3} G^{\tau_2}(t,x,\eta_{1234}) W(\mathrm{d}\eta_{1234}) + \int_{(\mathbb{R}\times E)^2} G_2^{\tau_2}(t,x,\eta_{12}) W(\mathrm{d}\eta_{12})$$

and

$$X^{\tau_3}(t,x) = \int_{(\mathbb{R}\times E)^3} G^{\tau_3}(t,x,\eta_{1234}) W(\mathrm{d}\eta_{1234}) + \int_{(\mathbb{R}\times E)^2} G_2^{\tau_3}(t,x,\eta_{12}) W(\mathrm{d}\eta_{12})$$

with

$$\begin{aligned} G_2^{\tau_2}(t,x,\eta_{12}) &= \int_{\mathbb{R}\times E} (G^{\tau_2}(t,x,\eta_{123(-3)}) + 2G^{\tau_2}(t,x,\eta_{13(-3)2}) + 2G^{\tau_2}(t,x,\eta_{132(-3)}))\mathrm{d}\eta_3 \\ &= (G^{\tau_4}(t,x,\eta_{12}) + 2G^{\tau_5}(t,x,\eta_{12}) + 2G^{\tau_6}(t,x,\eta_{12})) \end{aligned}$$

and

$$G_2^{\tau_3}(t,x,\eta_{12}) = 4\int_{\mathbb{R}\times E} G^{\tau_3}(t,x,\eta_{13(-3)2})\mathrm{d}\eta_3 = 4G^{\tau_7}(t,x,\eta_{12}).$$

Here the contributions associated to $G^{\tau_4}(t,x,\eta_{12})$ and $G^{\tau_5}(t,x,\eta_{12})$ are "reducible" since they can be conveniently factorized as follows:

$$\begin{aligned} G^{\tau_4}(t,x,\eta_{12}) &= \int_{\mathbb{R}\times E} G^{\tau_2}(t,x,\eta_{123(-3)})\mathrm{d}\eta_3 \\ &= e^{ik_{[12]}x} \int_0^t \mathrm{d}\sigma \int_0^\sigma \mathrm{d}\sigma' \int_0^{\sigma'} \mathrm{d}\sigma'' H_{t-\sigma}(k_{[12]}) H_{\sigma'-\sigma''}(k_{[12]}) H_{\sigma''-s_1}(k_1) \\ &\qquad \times H_{\sigma''-s_2}(k_2) V^{\tau_8}(\sigma-\sigma', k_{[12]}) \end{aligned}$$

with

$$\begin{aligned} V^{\tau_8}(\sigma,k_1) &= \int \psi_\circ(k_{[12]}, k_{-2}) H_\sigma(k_{[12]}) H_{\sigma-s_2}(k_2) H_{-s_2}(k_{-2})\mathrm{d}\eta_2 \\ &= \int \mathrm{d}k_2 \psi_\circ(k_{[12]}, k_{-2}) H_\sigma(k_{[12]}) \frac{e^{-|\sigma|k_2^2}}{2}. \end{aligned}$$

Also,

$$G^{\tau_5}(t,x,\eta_{12}) = \int_{\mathbb{R}\times E} G^{\tau_9}(t,x,\eta_{13(-3)2})\mathrm{d}\eta_3$$

$$
\begin{aligned}
&= e^{ik_{[12]}x}\psi_\circ(k_1,k_2)\int_0^t \mathrm{d}\sigma H_{t-\sigma}(k_{[12]})H_{\sigma-s_2}(k_2)\int_0^\sigma \mathrm{d}\sigma' \int_0^{\sigma'} \mathrm{d}\sigma'' H_{\sigma-\sigma'}(k_1) \\
&\qquad\qquad \times H_{\sigma''-s_1}(k_1)V^{\vee}(\sigma'-\sigma'',k_1) \\
&= e^{ik_{[12]}x}\psi_\circ(k_1,k_2)\int_0^t \mathrm{d}\sigma H_{t-\sigma}(k_{[12]})H_{\sigma-s_2}(k_2)e^{-ik_1x}G^{\vee}(\sigma,x,\eta_1).
\end{aligned}
$$

On the other side, the term $G^{\tau_1}(t,x,\eta_{12})$ cannot be reduced to a form involving $V^{\vee}$ or V^{τ_2}, and instead we have for it

$$
\begin{aligned}
G^{\tau_1}(t,x,\eta_{12}) &= \int_{\mathbb{R}\times E} G^{\tau_3}(t,x,\eta_{132(-3)})\mathrm{d}\eta_3 \\
&= e^{ik_{[12]}x}\int_0^t \mathrm{d}\sigma \int_0^\sigma \mathrm{d}\sigma' \int_0^{\sigma'} \mathrm{d}\sigma'' H_{t-\sigma}(k_{[12]})H_{\sigma'-s_2}(k_2)H_{\sigma''-s_1}(k_1)\times \\
&\qquad\qquad \times V^{\tau_1}(\sigma-\sigma',\sigma-\sigma'',k_{12})
\end{aligned}
$$

with

$$
\begin{aligned}
V^{\tau_1}(\sigma-\sigma',\sigma-\sigma'',k_{12}) &= \int_{\mathbb{R}\times E} \mathrm{d}\eta_3\psi_\circ(k_{[132]},k_{-3})H_{\sigma-s_3}(k_{-3})H_{\sigma-\sigma'}(k_{[132]})\times \\
&\qquad\qquad \times H_{\sigma'-\sigma''}(k_{[13]})H_{\sigma''-s_3}(k_3) \\
&= \int_E \mathrm{d}k_3\psi_\circ(k_{[123]},k_{-3})H_{\sigma-\sigma'}(k_{[123]})H_{\sigma'-\sigma''}(k_{[13]})\frac{e^{-k_3^2|\sigma-\sigma''|}}{2}.
\end{aligned}
$$

Similarly we have for G^{τ_4}

$$
\begin{aligned}
G^{\tau_4}(t,x,\eta_{12}) &= \int_{\mathbb{R}\times E} G^{\tau_5}(t,x,\eta_{13(-3)2})\mathrm{d}\eta_3 \\
&= e^{ik_{[12]}x}\int_0^t \mathrm{d}\sigma \int_0^\sigma \mathrm{d}\sigma' \int_0^{\sigma} \mathrm{d}\sigma'' H_{t-\sigma}(k_{[12]})H_{\sigma'-s_1}(k_1)H_{\sigma''-s_2}(k_2)\times \\
&\qquad\qquad \times V^{\tau_4}(\sigma-\sigma',\sigma-\sigma'',k_{12})
\end{aligned}
$$

with a vertex function which we choose to write in symmetrized form:

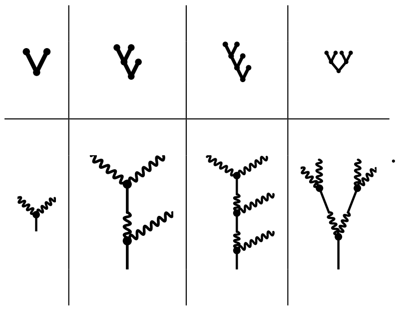
$$
\begin{aligned}
V^{\tau_4}(\sigma-\sigma',\sigma-\sigma'',k_{12}) &= \frac{1}{2}\int_{\mathbb{R}\times E} \mathrm{d}\eta_3 H_{\sigma-\sigma'}(k_{[13]})H_{\sigma-\sigma''}(k_{[2(-3)]})H_{\sigma'-s_3}(k_3)H_{\sigma''-s_3}(k_{-3}) \\
&\quad + \frac{1}{2}\int_{\mathbb{R}\times E} \mathrm{d}\eta_3 H_{\sigma-\sigma''}(k_{[13]})H_{\sigma-\sigma'}(k_{[2(-3)]})H_{\sigma'-s_3}(k_3)H_{\sigma''-s_3}(k_{-3}) \\
&= \frac{1}{2}\int_E \mathrm{d}k_3 H_{\sigma-\sigma'}(k_{[13]})H_{\sigma-\sigma''}(k_{[2(-3)]})\frac{e^{-k_3^2|\sigma'-\sigma''|}}{2} \\
&\quad + \frac{1}{2}\int_E \mathrm{d}k_3 H_{\sigma-\sigma''}(k_{[13]})H_{\sigma-\sigma'}(k_{[2(-3)]})\frac{e^{-k_3^2|\sigma'-\sigma''|}}{2}.
\end{aligned}
$$

9.3 Feynman diagrams

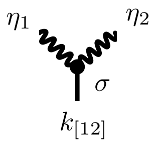
While here we only need to analyze few of the random fields X^τ, we think it useful to give a general perspective on their structure and in particular on the structure of the kernels G^τ and on the Wick contractions. Doing this will build a link with the theoretical physics methods and with quantum field theory (QFT), in particular with the Martin–Siggia–Rose response field formalism which has been used in the QFT analysis of the stochastic Burgers equations since the seminal work of Forster, Nelson, and Stephen [FNS77].

The explicit form of the kernels G^τ can be described in terms of Feynman diagrams and the associated rules. To each kernel G^τ we can associate a graph which is isomorphic to the tree τ and this graph can be mapped with Feynman rules to the explicit functional form of G^τ. The algorithm goes as follows: consider τ as a graph, where each edge and each internal vertex (i.e. not a leaf) are drawn as

To the trees  we associate, respectively, the diagrams

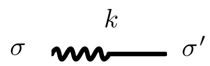


These diagrams correspond to kernels via the following rules: each internal vertex comes with a time integration and a factor (ik),



$$\longrightarrow \qquad (ik_{[12]}) \int_{\mathbb{R}_+} \mathrm{d}\sigma.$$

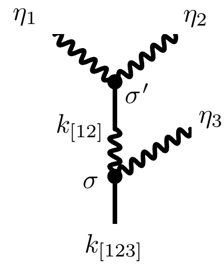
Each external wiggly line is associated to a variable η_i and a factor $H_{\sigma - s_i}(k_i)$, where σ is the integration variable of the internal vertex to which the line is attached. Each response line is associated to a factor $h_{\sigma-\sigma'}(k)$, where k is the moment carried by the line and σ, σ' are the time labels of the vertices to which it is attached:



$$\longrightarrow \qquad h_{\sigma-\sigma'}(k).$$

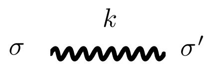
Note that these lines carry information about the casual propagation. Finally, the outgoing line always carries a factor $h_{t-\sigma}(k)$, where k is the outgoing momentum and is σ

the time label of the vertex to which the line is attached. For example:



$$\longrightarrow \quad \int_{\mathbb{R}^2_+} \mathrm{d}\sigma\mathrm{d}\sigma'(ik_{[123]})h_{t-\sigma}(k_{[123]})(ik_{[12]})h_{\sigma-\sigma'}(k_{[12]})\times \times H_{\sigma'-s_1}(k_1)H_{\sigma'-s_2}(k_2)H_{\sigma-s_3}(k_3).$$

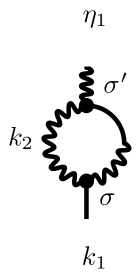
Once given a diagram, the associated Wick contractions are obtained by all possible pairings of the wiggly lines. To each of these pairings we associate the corresponding correlation function of the Ornstein-Uhlenbeck process and an integration over the momentum variable carried by the line:



$$\longrightarrow \quad \int_E \mathrm{d}k \frac{e^{-k^2|\sigma-\sigma'|}}{2}.$$

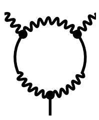
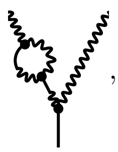
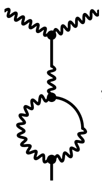
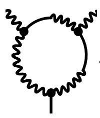
So for example we have

$$G_1(t,x,\eta_1) = 2\times$$

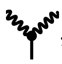


$$= 2\int_{\mathbb{R}^2_+} \mathrm{d}\sigma\mathrm{d}\sigma' H_{t-\sigma}(k_1)\int_E \mathrm{d}k_2 \frac{e^{-k_2^2|\sigma-\sigma'|}}{2} H_{\sigma-\sigma'}(k_{[1(-2)]})H_{\sigma'-s_1}(k_1),$$

which coincides with the expression obtained previously. Note that we also have to take into account the multiplicities of the different ways in which each graph can be obtained. The contractions arising from G and G result in the following set of diagrams:

$G_2 = 4G = 4\times$  and $G =$ , $G =$ , $G =$ .

The diagrammatic representation makes pictorially evident what we already have remarked with explicit computations: G and G are formed by the union of two graphs:

 and ,

while the kernel G cannot be decomposed in such a way and it has a shape very similar to that of G.

9.4 Bounds on the vertex functions

Up to now we have obtained explicit expressions for the kernels appearing in the Wick contractions. These kernels feature vertex functions. Let us start the analysis by bounding them. Consider for example the first non-trivial one,

$$G^{\vee}(t,x,\eta_1)=\int_0^t \mathrm{d}\sigma \int_0^\sigma \mathrm{d}\sigma' H_{t-\sigma}(k_1)H_{\sigma'-s_1}(k_1)V^{\vee}(\sigma-\sigma',k_1).$$

We have

$$|G^{\vee}(t,x,\eta_1)|\leqslant \int_0^t \mathrm{d}\sigma \int_0^\sigma \mathrm{d}\sigma' M_{t-\sigma}(k_1)M_{\sigma'-s_1}(k_1)|V^{\vee}(\sigma-\sigma',k_1)|,$$

where $M_t(k)=|H_t(k)|=|k|\exp(-k^2t)\mathbb{1}_{t\geqslant 0}$. Since the integrand is positive, we can extend the domain of integration to obtain an upper bound:

$$|G^{\vee}(t,x,\eta_1)|\leqslant \int_{\mathbb{R}} \mathrm{d}\sigma \int_{\mathbb{R}} \mathrm{d}\sigma' M_{t-\sigma}(k_1)M_{\sigma'-s_1}(k_1)|V^{\vee}(\sigma-\sigma',k_1)|=Z(t,x,\eta_1).$$

The quantity $Z(t,x,\eta_1)$ is given by a multiple convolution, and since we are interested in L^2 bounds of $G^{\vee}(t,x,\eta_1)$, we can pass to Fourier variables in time to decouple the various dependencies:

$$\int_{\mathbb{R}\times E}|G^{\vee}(t,x,\eta_1)|^2\mathrm{d}\eta_1 \leqslant \int_{\mathbb{R}\times E}|Z(t,x,\eta_1)|^2\mathrm{d}\eta_1=(2\pi)^{-1}\int_{\mathbb{R}\times E}|\hat{Z}(t,x,\theta_1)|^2\mathrm{d}\theta_1,$$

where $\theta_1=(\omega_1,k_1)$ and

$$\hat{Z}(t,x,\theta_1)=\int_{\mathbb{R}} e^{-i\omega_1 s_1}Z(t,x,\eta_1)\mathrm{d}s_1=\hat{M}(-\theta_1)\hat{M}(-\theta_1)\mathcal{V}^{\vee}(-\omega_1,k_1)$$

with

$$\mathcal{V}^{\vee}(\omega_1,k_1)=\int_{\mathbb{R}} e^{-i\omega_1\sigma}|V^{\vee}(\sigma,k_1)|\mathrm{d}\sigma.$$

Then

$$\begin{aligned}\int_{\mathbb{R}\times E}|\hat{Z}(t,x,\theta_1)|^2\mathrm{d}\theta_1 &\leqslant \int_{\mathbb{R}\times E}|\hat{M}(\theta_1)\hat{M}(\theta_1)|^2|\mathcal{V}^{\vee}(\omega_1,k_1)|^2\mathrm{d}\theta_1\\ &\leqslant \int_{\mathbb{R}\times E}|Q(\theta_1)Q(\theta_1)|^2\Big(\int_{\mathbb{R}}|V^{\vee}(\sigma,k_1)|\mathrm{d}\sigma\Big)^2\mathrm{d}\theta_1,\end{aligned}$$

where $Q(\theta)=\hat{H}(\theta)$ is the time Fourier transform of H:

$$Q(\theta)=\frac{ik}{i\omega+k^2}.$$

This computation hints to the fact that the relevant norm on the vertex functions is given by the supremum norm in the wave vectors of the L^1 norm in the time variable. Summing up, we have in this case

$$\int_{\mathbb{R}\times E}|G^{\vee}(t,x,\eta_1)|^2\mathrm{d}\eta_1 \leqslant \int_{\mathbb{R}\times E}|Q(\theta_1)Q(\theta_1)|^2\|V^{\vee}(\sigma,k_1)\|^2_{L^1_\sigma}\mathrm{d}\theta_1.$$

For the other contraction kernels we can proceed similarly. Consider first

$$
\begin{aligned}
G^{\text{[tree]}}(t,x,\eta_{12}) &= \int_{\mathbb{R}\times E} G^{\text{[tree]}}(t,x,\eta_{123(-3)})\mathrm{d}\eta_3 \\
&= e^{ik_{[12]}x}\int_0^t \mathrm{d}\sigma \int_0^{\sigma}\mathrm{d}\sigma' \int_0^{\sigma'} \mathrm{d}\sigma'' H_{t-\sigma}(k_{[12]})H_{\sigma'-\sigma''}(k_{[12]})H_{\sigma''-s_1}(k_1)\times \\
&\qquad \times H_{\sigma''-s_2}(k_2)V^{\text{[tree]}}(\sigma-\sigma',k_{[12]}),
\end{aligned}
$$

for which we have

$$
\int_{(\mathbb{R}\times E)^2} |G^{\text{[tree]}}(t,x,\eta_{12})|^2 \mathrm{d}\eta_{12} \leqslant \int_{(\mathbb{R}\times E)^2} |Q(\theta_{[12]})Q(\theta_{[12]})Q(\theta_1)Q(\theta_2)|^2 \|V^{\text{[tree]}}(\sigma,k_{[12]})\|^2_{L^1_\sigma}\mathrm{d}\theta_{12}.
$$

The next term is

$$
G^{\text{[tree]}}(t,x,\eta_{12}) = e^{ik_{[12]}x}\psi_\circ(k_1,k_2)\int_0^t \mathrm{d}\sigma H_{t-\sigma}(k_{[12]})H_{\sigma-s_2}(k_2)e^{-ik_1x}G^{\text{[tree]}}(\sigma,x,\eta_1),
$$

for which

$$
\int_{(\mathbb{R}\times E)^2} |G^{\text{[tree]}}(t,x,\eta_{12})|^2 \mathrm{d}\eta_{12} \leqslant \int_{(\mathbb{R}\times E)^2} |\psi_\circ(k_1,k_2)Q(\theta_{[12]})Q(\theta_1)Q(\theta_1)Q(\theta_2)|^2 \|V^{\text{[tree]}}(\sigma,k_1)\|^2_{L^1_\sigma}\mathrm{d}\theta_{12}.
$$

And finally

$$
\begin{aligned}
G^{\text{[tree]}}(t,x,\eta_{12}) &= e^{ik_{[12]}x}\int_0^t \mathrm{d}\sigma \int_0^{\sigma'}\mathrm{d}\sigma' \int_0^{\sigma'} \mathrm{d}\sigma'' H_{t-\sigma}(k_{[12]})H_{\sigma'-s_2}(k_2)H_{\sigma''-s_1}(k_1)\times \\
&\qquad \times V^{\text{[tree]}}(\sigma-\sigma',\sigma-\sigma'',k_{12}),
\end{aligned}
$$

for which

$$
\int_{(\mathbb{R}\times E)^2} |G^{\text{[tree]}}(t,x,\eta_{12})|^2 \mathrm{d}\eta_{12} \leqslant \int_{(\mathbb{R}\times E)^2} |Q(\theta_{[12]})Q(\theta_1)Q(\theta_2)|^2 \|V^{\text{[tree]}}(\sigma,\sigma',k_{12})\|^2_{L^1_\sigma L^1_{\sigma'}}\mathrm{d}\theta_{12}.
$$

Similarly, we get

$$
\int_{(\mathbb{R}\times E)^2} |G^{\text{[tree]}}(t,x,\eta_{12})|^2 \mathrm{d}\eta_{12} \leqslant \int_{(\mathbb{R}\times E)^2} |Q(\theta_{[12]})Q(\theta_1)Q(\theta_2)|^2 \|V^{\text{[tree]}}(\sigma,\sigma',k_{12})\|^2_{L^1_\sigma L^1_{\sigma'}}\mathrm{d}\theta_{12}.
$$

For these computations to be useful it remains to obtain explicit bounds for the norms of the vertex functions.

Lemma 9.5. *For any $\varepsilon > 0$ we have*

$$
\int_{\mathbb{R}} \mathrm{d}\sigma |V^{\text{[tree]}}(\sigma,k_1)| + \int_{\mathbb{R}} \mathrm{d}\sigma |V^{\text{[tree]}}(\sigma,k_1)| \lesssim |k_1|^{\varepsilon}.
$$

Proof. Some care has to be exercised since a too bold bounding would fail to give a finite result. Indeed, a direct estimation would lead to

$$\int_{\mathbb{R}} \mathrm{d}\sigma \left| V^{\curlyvee}(\sigma, k_1) \right| \leqslant \int_E \mathrm{d}k_2 \int_0^\infty \mathrm{d}\sigma |k_{[12]}| e^{-\sigma(k_2^2 + k_{[12]}^2)} \leqslant \int_E \mathrm{d}k_2 \frac{|k_{[12]}|}{k_2^2 + k_{[12]}^2} = +\infty,$$

due to the logarithmic divergence at infinity (recall that $\int_E$ stands for $\sum_{\mathbb{Z}\setminus\{0\}}$). To overcome this problem, observe that

$$V^{\curlyvee}(\sigma, 0) = 2 \int_E \mathrm{d}k_2 H_\sigma(k_2) \frac{e^{-|\sigma| k_2^2}}{2} = 0$$

since the integrand is an odd function of k_2. So we can write instead

$$V^{\curlyvee}(\sigma, k_1) = 2 \int_E \mathrm{d}k_2 [H_\sigma(k_{[12]}) - H_\sigma(k_2)] \frac{e^{-|\sigma| k_2^2}}{2},$$

and at this point it is easy to verify that

$$\int_{\mathbb{R}} \mathrm{d}\sigma |V^{\curlyvee}(\sigma, k_1)| \lesssim |k_1|^\varepsilon$$

for arbitrarily small $\varepsilon > 0$. Indeed,

$$\int_{\mathbb{R}} \mathrm{d}\sigma |V^{\curlyvee}(\sigma, k_1)| \lesssim \int_E \mathrm{d}k_2 \int_0^\infty \mathrm{d}\sigma |H_\sigma(k_{[12]}) - H_\sigma(k_2)| e^{-\sigma k_2^2}$$

and a first order Taylor expansion gives

$$|H_\sigma(k_{[12]}) - H_\sigma(k_2)| \lesssim |k_1| \int_0^1 \mathrm{d}\tau e^{-c(k_2 + \tau k_1)^2 \sigma}.$$

Therefore,

$$\begin{aligned}
\int_{\mathbb{R}} \mathrm{d}\sigma |V^{\curlyvee}(\sigma, k_1)| &\lesssim |k_1| \int_0^1 \mathrm{d}\tau \int_E \mathrm{d}k_2 \int_0^\infty \mathrm{d}\sigma e^{-c(k_2 + \tau k_1)^2 \sigma - k_2^2 \sigma} \\
&\lesssim |k_1| \int_0^1 \mathrm{d}\tau \int_E \frac{\mathrm{d}k_2}{(k_2 + \tau k_1)^2 + k_2^2},
\end{aligned}$$

but now

$$\int_E \frac{\mathrm{d}k_2}{(k_2 + \tau k_1)^2 + k_2^2} \lesssim \int_E \frac{\mathrm{d}k_2}{k_2^2} \lesssim 1.$$

On the other side the sum over E is bounded by the corresponding integral over $\mathbb{R}$, so a change of variables gives

$$\int_E \frac{\mathrm{d}k_2}{(k_2 + \tau k_1)^2 + k_2^2} \lesssim \int_{\mathbb{R}} \frac{\mathrm{d}k_2}{(k_2 + \tau k_1)^2 + k_2^2} \lesssim \frac{1}{\tau |k_1|} \int_{\mathbb{R}} \frac{\mathrm{d}k_2}{(k_2 + 1)^2 + k_2^2} \lesssim \frac{1}{\tau |k_1|}.$$

By interpolating these two bounds, we obtain

$$\int_{\mathbb{R}} \mathrm{d}\sigma |V^{\curlyvee}(\sigma, k_1)| \lesssim |k_1|^{\varepsilon} \int_0^1 \frac{\mathrm{d}\tau}{\tau^{1-\varepsilon}} \lesssim |k_1|^{\varepsilon}$$

for arbitrarily small $\varepsilon > 0$. The same arguments also give the bound for $\int_{\mathbb{R}} \mathrm{d}\sigma |V^{\curlyvee}(\sigma, k_1)|$. □

Lemma 9.6. *For any $\varepsilon > 0$ we have*

$$\int_{\mathbb{R}^2} \Big| V^{\curlyvee}(\sigma, \sigma', k_{12}) \Big| \,\mathrm{d}\sigma \mathrm{d}\sigma' \lesssim |k_{[12]}|^{-1+\varepsilon}.$$

Proof. We have

$$\begin{aligned}
\int_{\mathbb{R}^2} \Big| V^{\curlyvee}(\sigma, \sigma', k_{12}) \Big| \,\mathrm{d}\sigma \mathrm{d}\sigma' &\lesssim \int_{\mathbb{R}^2} \left| \int_E \mathrm{d}k_3 \psi_\circ(k_{[123]}, k_{-3}) H_\sigma(k_{[132]}) H_{\sigma'}(k_{[13]}) e^{-k_3^2(\sigma'+\sigma)} \right| \mathrm{d}\sigma \mathrm{d}\sigma' \\
&\lesssim \int_E \mathrm{d}k_3 \frac{|k_{[123]}||k_{[13]}|}{(k_3^2 + k_{[123]}^2)(k_3^2 + k_{[13]}^2)} \\
&\lesssim \int_E \mathrm{d}k_3 \frac{1}{(k_3^2 + k_{[123]}^2)^{1/2}(k_3^2 + k_{[13]}^2)^{1/2}} \\
&\lesssim \int_E \mathrm{d}k_3 \frac{1}{(k_3^2 + k_{[123]}^2)^{1/2}|k_3|} \lesssim |k_{[12]}|^{-1+\varepsilon}
\end{aligned}$$

for arbitrarily small $\varepsilon > 0$. □

Lemma 9.7. *For any $\varepsilon > 0$ we have*

$$\int_{\mathbb{R}^2} \Big| V^{\curlyvee}(\sigma, \sigma', k_{12}) \Big| \,\mathrm{d}\sigma \mathrm{d}\sigma' \lesssim |k_{[12]}|^{-1+\varepsilon}.$$

Proof. We can estimate

$$\begin{aligned}
&\int_{\mathbb{R}^2} \Big| V^{\curlyvee}(\sigma, \sigma', k_{12}) \Big| \,\mathrm{d}\sigma \mathrm{d}\sigma' \\
&\quad\lesssim \int_{\mathbb{R}^2} \left| \int_E \mathrm{d}k_3 H_\sigma(k_{[13]}) H_{\sigma'}(k_{[(-3)2]}) e^{-k_3^2|\sigma'-\sigma|} \right| \mathrm{d}\sigma \mathrm{d}\sigma' \\
&\quad\lesssim \int_0^\infty \mathrm{d}\sigma \int_0^\infty \mathrm{d}\sigma' \int_E \mathrm{d}k_3 |H_\sigma(k_{[13]}) H_{\sigma+\sigma'}(k_{[(-3)2]})| e^{-k_3^2\sigma'} \\
&\qquad + \int_0^\infty \mathrm{d}\sigma \int_0^\infty \mathrm{d}\sigma' \int_E \mathrm{d}k_3 |H_{\sigma+\sigma'}(k_{[13]}) H_\sigma(k_{[(-3)2]})| e^{-k_3^2\sigma'} \\
&\quad\lesssim \int_E \mathrm{d}k_3 \frac{|k_{[13]}||k_{[(-3)2]}|}{(k_3^2 + k_{[(-3)2]}^2)(k_{[13]}^2 + k_{[(-3)2]}^2)} + \int_E \mathrm{d}k_3 \frac{|k_{[13]}||k_{[(-3)2]}|}{(k_3^2 + k_{[13]}^2)(k_{[(-3)2]}^2 + k_{[13]}^2)} \\
&\quad\lesssim \int_E \mathrm{d}k_3 \frac{1}{(k_3^2 + k_{[(-3)2]}^2)^{1/2}(k_{[13]}^2 + k_{[(-3)2]}^2)^{1/2}}
\end{aligned}$$

$$
\begin{aligned}
&+ \int_E \mathrm{d}k_3 \frac{1}{(k_3^2 + k_{[13]}^2)^{1/2}(k_{[(-3)2]}^2 + k_{[13]}^2)^{1/2}} \\
&\lesssim \int_E \mathrm{d}k_3 \frac{1}{|k_3|(k_{[13]}^2 + k_{[(-3)2]}^2)^{1/2}} \lesssim |k_{[12]}|^{-1+\varepsilon}
\end{aligned}
$$

whenever $\varepsilon > 0$. □

9.5 Regularity of the driving terms

In this section we will determine the Besov regularity of the random fields X^τ. Below we will derive estimates of the form

$$
\sup_{q \geqslant 0, x \in \mathbb{T}, s,t \geqslant 0} 2^{2\gamma_1(\tau) q} |t-s|^{-\gamma_2(\tau)} \mathbb{E}[(\Delta_q X^\tau(t,x) - \Delta_q X^\tau(s,x))^2] \lesssim 1 \tag{84}
$$

for any $\gamma_2(\tau) \in [0,2]$ with $\gamma_1(\tau) + \gamma_2(\tau) = \gamma(\tau)$. Each $\Delta_q X^\tau(t,x)$ is a random variable with a finite chaos decomposition, so Gaussian hypercontractivity implies that

$$
\sup_{q \geqslant 0, x \in \mathbb{T}, s,t \geqslant 0} 2^{p\gamma_1(\tau) q} |t-s|^{-p\gamma_2(\tau)/2} \mathbb{E}[(\Delta_q X^\tau(t,x) - \Delta_q X^\tau(s,x))^p] \lesssim 1
$$

for any $p \geqslant 2$, and then

$$
\sup_{q \geqslant 0, s,t \geqslant 0} 2^{p\gamma_1(\tau) q} |t-s|^{-p\gamma_2(\tau)/2} \mathbb{E}[\|\Delta_q X^\tau(t,\cdot) - \Delta_q X^\tau(s,\cdot)\|_{L^p(\mathbb{T})}^p] \lesssim 1.
$$

From here we derive that for all $\varepsilon > 0$

$$
\begin{aligned}
&\sup_{s,t \geqslant 0} |t-s|^{-p\gamma_2(\tau)/2} \mathbb{E}[\|X^\tau(t,\cdot) - X^\tau(s,\cdot)\|_{B_{p,p}^{\gamma_1(\tau)-\varepsilon}}^p] \\
&\qquad = \sup_{s,t \geqslant 0} |t-s|^{-p\gamma_2(\tau)/2} \mathbb{E}\Big[\sum_{q \geqslant -1} 2^{(\gamma_1(\tau)-\varepsilon) q} \|\Delta_q X^\tau(t,\cdot) - \Delta_q X^\tau(s,\cdot)\|_{L^p(\mathbb{T})}^p\Big] \lesssim 1,
\end{aligned}
$$

and by the Besov embedding theorem we get for $p \geqslant 1/\varepsilon$

$$
\mathbb{E}[\|X^\tau(t,\cdot) - X^\tau(s,\cdot)\|_{B_{\infty,\infty}^{\gamma_1(\tau)-2\varepsilon}}^p] \lesssim \mathbb{E}[\|X^\tau(t,\cdot) - X^\tau(s,\cdot)\|_{B_{p,p}^{\gamma_1(\tau)-\varepsilon}}^p].
$$

Thus an application of Kolmogorov's continuity criterion gives

$$
\mathbb{E}[\|X^\tau\|_{C_T^{\gamma_2(\tau)/2-\varepsilon} \mathscr{C}^{\gamma_1(\tau)-\varepsilon}}^p] \lesssim 1
$$

whenever $\varepsilon > 0$, $\gamma_1(\tau) + \gamma_2(\tau) \leqslant \gamma(\tau)$, and $p, T > 0$. This argument reduces the regularity problem to second moment estimations.

Let us start by analyzing $Q \circ X$, whose kernel is given by

$$
G^{Q \circ X}(t,x,\eta_{12}) = e^{ik_{[12]}x} \psi_\circ(k_1,k_2) H_{t-s_1}(k_1) \int_0^t \mathrm{d}\sigma H_{t-\sigma}(k_2) H_{\sigma-s_2}(k_2).
$$

Now note that, by symmetry under the change of variables $k_1 \to -k_1$ we have

$$G_0^{Q \circ X}(t, x) = \int_{(\mathbb{R}\times E)} G^{Q \circ X}(t, x, \eta_{1(-1)}) \mathrm{d}\eta_1 = 0$$

since

$$G^{Q \circ X}(t, x, \eta_{1(-1)}) = \psi_\circ(k_1, -k_1) H_{t-s_1}(k_1) \int_0^t \mathrm{d}\sigma H_{t-\sigma}(-k_1) H_{\sigma-s_2}(-k_1)$$

and $H_{t-\sigma}(-k_1) = -H_{t-\sigma}(k_1)$. So only the second chaos is involved in the following computation:

$$\begin{aligned}
&\mathbb{E}[(\Delta_q(Q \circ X)(t,x))^2] \\
&\quad\lesssim \int_{E^2} \mathrm{d}k_{12} \rho_q(k_{[12]})^2 \psi_\circ(k_1,k_2)^2 \int_0^t \int_0^t \mathrm{d}\sigma \mathrm{d}\sigma' H_{t-\sigma}(k_2) H_{t-\sigma'}(k_2) \frac{e^{-k_2^2|\sigma-\sigma'|}}{2} \\
&\quad\lesssim \int_{E^2} \mathrm{d}k_{12} \rho_q(k_{[12]})^2 \psi_\circ(k_1,k_2)^2 |k_2|^2 \int_0^t \int_0^t \mathrm{d}\sigma \mathrm{d}\sigma' e^{-k_2^2[(t-\sigma)+(t-\sigma')+|\sigma-\sigma'|]} \\
&\quad\lesssim \int_{E^2} \mathrm{d}k_{12} \rho_q(k_{[12]})^2 \psi_\circ(k_1,k_2)^2 |k_2|^{-2} \lesssim \int_{E^2} \mathrm{d}k \mathrm{d}k' \rho_q(k)^2 \psi_\circ(k-k',k')^2 |k'|^{-2} \\
&\quad\lesssim 2^q \sum_{i \gtrsim q} 2^{i-2i} \lesssim 1.
\end{aligned}$$

Similarly we see that $\mathbb{E}[(\Delta_q(Q \circ X)(t,x) - \Delta_q(Q \circ X)(s,x))^2] \lesssim 2^{q\kappa}|t-s|^{\kappa/2}$ whenever $\kappa \in [0,2]$, from where we get the required temporal regularity.

Next, let us get to the X^τ. We would like to reduce the estimation of the time difference in (84) to an estimate at fixed times. For that purpose observe that every X^τ solves a parabolic equation started in 0 (except for $\tau = \bullet$ which is easy to treat by hand). Thus we get for $0 \leqslant s \leqslant t$

$$\begin{aligned}
X^\tau(t,\cdot) - X^\tau(s,\cdot) &= \int_0^t P_{t-r} \mathscr{L} X^\tau(r) \mathrm{d}r - \int_0^s P_{s-r} \mathscr{L} X^\tau(r) \mathrm{d}r \\
&= (P_{t-s} - 1) X^\tau(s) + \int_s^t P_{t-r} \mathscr{L} X^\tau(r) \mathrm{d}r. \qquad (85)
\end{aligned}$$

We estimate the first term by

$$|\Delta_q(P_{t-s} - 1) X^\tau(s,x)| \lesssim |e^{c(t-s)2^{2q}} - 1| |\Delta_q X^\tau(s,x)| \lesssim |t-s|^{\gamma_2(\tau)/2} 2^{q\gamma_2(\tau)} |\Delta_q X^\tau(s,x)|$$

whenever $\gamma_2(\tau) \in [0,2]$. The first estimate may appear rather formal, but it is not difficult to prove it rigorously, see for example Lemma 2.4 in [BCD11]. So if we can show that

$$\mathbb{E}[|\Delta_q X^\tau(s,x)|^2] \lesssim 2^{-q\gamma(\tau)}, \qquad (86)$$

for $\gamma(\tau) \leqslant 2$, then the estimate (84) follows for the first term on the right hand side of (85). For the second term in (85), we will see below that when estimating $\mathbb{E}[|\Delta_q X^\tau(t,x)|^2]$ we can extend the domain of integration of $\int_0^t P_{t-r}\Delta_q u^\tau(r)\mathrm{d}r$ until $-\infty$, and this gives us a factor 2^{-2q}. If instead we integrate only over the time interval $[s,t]$, then we get an additional factor $1-e^{c(t-s)2^{2q}}$, which we can then treat as above. It will therefore suffice to prove bounds of the form $\mathbb{E}[|\Delta_q X^\tau(t,x)|^2] \lesssim 2^{-2q\gamma(\tau)}$.

Next we treat the X^τ. As we have seen in the case of the vertex functions it will be convenient to pass to Fourier variables. In doing so we will establish uniform bounds for the kernel functions $(G^\tau)_\tau$ in terms of their stationary versions $(\Gamma^\tau)_\tau$: that is the kernels which govern the statistics of the random fields $X^\tau(t,\cdot)$ when $t \to +\infty$.

Let recursively define $\Gamma^\bullet = G^\bullet$ and

$$\Gamma^{(\tau_1\tau_2)}(t,x,\eta_{(\tau_1\tau_2)}) = \int_{-\infty}^t \mathrm{d}s\partial_x P_{t-s}(\Gamma^{\tau_1}(s,\cdot,\eta_{\tau_1}),\Gamma^{\tau_2}(s,\cdot,\eta_{\tau_2}))(x).$$

Like the G kernels they have the factorized form $\Gamma^\tau(t,x,\eta_\tau) = e^{ik_{[\tau]}x}\gamma_t^\tau(\eta_\tau)$. The advantage of the kernels Γ is that their Fourier transform Q in the time variables $(s_1,\ldots,s_n)$ is very simple. Letting $\theta_i = (\omega_i, k_i)$ we have

$$Q^\tau(t,x,\theta_\tau) = \int_{(\mathbb{R}\times E)^n} \mathrm{d}s_\tau e^{i\omega_\tau\cdot s_\tau}\Gamma^\tau(t,x,\eta_\tau) = e^{ik_{[\tau]}x+i\omega_{[\tau]}t}q^\tau(\theta_\tau),$$

where

$$q^\bullet(\theta) = \frac{ik}{i\omega+k^2}, \qquad q^{(\tau_1\tau_2)}(\theta_{(\tau_1\tau_2)}) = q^\bullet(\theta_{[(\tau_1\tau_2)]})q^{\tau_1}(\theta_{\tau_1})q^{\tau_2}(\theta_{\tau_2}).$$

The kernels for the Γ terms bound the corresponding kernels for the G terms:

$$|G^{(\tau_1\tau_2)}(t,x,\eta_{(\tau_1\tau_2)})| = |H^\tau(t,\eta_\tau)| \leqslant |\gamma_t^\tau(\eta_\tau)| = |\Gamma^{(\tau_1\tau_2)}(t,x,\eta_{(\tau_1\tau_2)})|.$$

This is true for $\tau = \bullet$ and by induction

$$\begin{aligned}
|G^{(\tau_1\tau_2)}(t,x,\eta_{(\tau_1\tau_2)})| &= |H^{(\tau_1\tau_2)}(t,\eta_{(\tau_1\tau_2)})| = |k_{[(\tau_1\tau_2)]}|\int_0^t h_{t-s}(k_{[(\tau_1\tau_2)]})|H^{\tau_1}(s,\eta_{\tau_1})||H^{\tau_2}(s,\eta_{\tau_2})|\mathrm{d}s \\
&\leqslant |k_{[(\tau_1\tau_2)]}|\int_{-\infty}^t h_{t-s}(k_{[(\tau_1\tau_2)]})|g_s^{\tau_1}(\eta_{\tau_1})||g_s^{\tau_2}(\eta_{\tau_2})|\mathrm{d}s \\
&\leqslant |k_{[(\tau_1\tau_2)]}|\int_{-\infty}^t h_{t-s}(k_{[(\tau_1\tau_2)]})|\gamma_s^{\tau_1}(\eta_{\tau_1})||\gamma_s^{\tau_2}(\eta_{\tau_2})|\mathrm{d}s \\
&= |\gamma_t^{(\tau_1\tau_2)}(\eta_{(\tau_1\tau_2)})| = |\Gamma^{(\tau_1\tau_2)}(t,x,\eta_{(\tau_1\tau_2)})|.
\end{aligned}$$

These bounds and the computation of the Fourier transform imply the following estimation for the L^2 norm of the kernels G uniformly in $t \geqslant 0$:

$$\int_{\mathbb{R}^n}|G^\tau(t,x,\eta_\tau)|^2\mathrm{d}s_\tau \leqslant \int_{\mathbb{R}^n}|\Gamma^\tau(t,x,\eta_\tau)|^2\mathrm{d}s_\tau \leqslant \int_{\mathbb{R}^n}|Q^\tau(t,x,\theta_\tau)|^2\mathrm{d}\omega_\tau,$$

that is

$$\int_{\mathbb{R}^n} |G^\tau(t, x, \eta_\tau)|^2 \mathrm{d}s_\tau \leqslant \int_{\mathbb{R}^n} |q^\tau(\theta_\tau)|^2 \mathrm{d}\omega_\tau. \tag{87}$$

This observation simplifies many computations of the moments of the X^τ's and gives estimates that are uniform in $t \geqslant 0$. Actually it shows that the statistics of the X^τ's are bounded by the statistics in the stationary state.

Now, note first that due to the factorization (83) we have that

$$\Delta_q G^\tau(t, x, \eta_\tau) = \rho_q(k_{[\tau]}) G^\tau(t, x, \eta_\tau),$$

so the Littlewood-Paley blocks of X^τ have the expression

$$\Delta_q X^\tau(t, x) = \int_{(\mathbb{R}\times E)^n} \rho_q(k_{[\tau]}) G^\tau(t, x, \eta_\tau) \prod_{i=1}^n W(\mathrm{d}\eta_i)$$

which we rewrite in terms of the chaos expansion as

$$\Delta_q X^\tau(t, x) = \sum_{\ell=0}^{d(\tau)} \int_{(\mathbb{R}\times E)^n} \rho_q(k_{[\tau]}) G_\ell^\tau(t, x, \eta_{1\cdots\ell}) W(\mathrm{d}\eta_{1\cdots\ell}).$$

By the orthogonality of the different chaoses and because of the bound (87) we have

$$\begin{aligned}\mathbb{E}[(\Delta_q X^\tau(t, x))^2] &\lesssim \sum_{\ell=0}^{d(\tau)} \int_{(\mathbb{R}\times E)^n} \rho_q(k_{[\tau]}) |G_\ell^\tau(t, x, \eta_{1\cdots\ell})|^2 \mathrm{d}\eta_{1\cdots\ell} \\ &\lesssim \sum_{\ell=0}^{d(\tau)} \int_{(\mathbb{R}\times E)^n} \rho_q(k_{[\tau]}) |q_\ell^\tau(\eta_{1\cdots\ell})|^2 \mathrm{d}\eta_{1\cdots\ell}.\end{aligned}$$

By proceeding recursively from the leaves to the root and using the bounds on the vertex functions that we already proved, the problem of the estimation of the above integrals is reduced to estimate at each step an integral of the form

$$\int_{\mathbb{R}\times E} |\theta|^{-\alpha} |\theta' - \theta|^{-\beta} \mathrm{d}\theta,$$

where we have a joining of two leaves into a vertex, each leave carrying a factor proportional either to $|\theta|^{-\alpha}$ or $|\theta|^{-\beta}$ with $\alpha, \beta \geqslant 2$ and where the length $|\theta|$ of the θ variables is conveniently defined as

$$|\theta| = |\omega|^{1/2} + |k|,$$

so that the estimate $|q(\theta)| \sim |\theta|^{-1}$ holds for $q = q^\bullet$.

Lemma 9.8. *For this basic integral we have the estimate*

$$\int_{\mathbb{R}\times E} |\theta|^{-\alpha} |\theta' - \theta|^{-\beta} \mathrm{d}\theta \lesssim |\theta'|^{-\rho},$$

where $\rho = \alpha + \beta - 3$ if $\alpha, \beta < 3$ and $\alpha + \beta > 3$, and where $\rho = \alpha - \varepsilon$ for an arbitrarily small $\varepsilon > 0$ if $\beta \geqslant 3$ and $\alpha \in (0, \beta]$. If $\beta \geqslant 3$ and $\alpha \in (0, \beta]$, then

$$\int_{\mathbb{R}\times E} \psi_\circ(k, k'-k)^2 |\theta|^{-\alpha} |\theta' - \theta|^{-\beta} \mathrm{d}\theta \lesssim |\theta'|^{-\alpha+\varepsilon} |k'|^{3-\beta}.$$

Proof. Let $\ell \in \mathbb{N}$ such that $|\theta'| \sim 2^\ell$:

$$\int_{\mathbb{R}\times E} |\theta|^{-\alpha} |\theta' - \theta|^{-\beta} \mathrm{d}\theta \lesssim \sum_{i,j \geqslant 0} 2^{-\alpha i - \beta j} \int_{\mathbb{R}\times E} \mathbb{1}_{|\theta'| \sim 2^\ell, |\theta| \sim 2^i, |\theta' - \theta| \sim 2^j} \mathrm{d}\theta.$$

Then there are three possibilities, either $\ell \lesssim i \sim j$ or $i \lesssim j \sim \ell$ or $j \lesssim i \sim \ell$. In the first case we bound

$$\int_{\mathbb{R}\times E} \mathbb{1}_{|\theta'| \sim 2^\ell, |\theta| \sim 2^i, |\theta' - \theta| \sim 2^j} \mathrm{d}\theta \lesssim \int_{\mathbb{R}\times E} \mathbb{1}_{|\theta| \sim 2^i} \mathrm{d}\theta \lesssim 2^{3i},$$

and in the second one

$$\int_{\mathbb{R}\times E} \mathbb{1}_{|\theta'| \sim 2^\ell, |\theta| \sim 2^i, |\theta' - \theta| \sim 2^j} \mathrm{d}\theta \lesssim \int_{\mathbb{R}\times E} \mathbb{1}_{|\theta| \sim 2^i} \mathrm{d}\theta \lesssim 2^{3i},$$

and similarly in the third case

$$\int_{\mathbb{R}\times E} \mathbb{1}_{|\theta'| \sim 2^\ell, |\theta| \sim 2^i, |\theta' - \theta| \sim 2^j} \mathrm{d}\theta \lesssim \int_{\mathbb{R}\times E} \mathbb{1}_{|\theta| \sim 2^j} \mathrm{d}\theta \lesssim 2^{3j}.$$

So if $\alpha + \beta > 3$ we have

$$\begin{aligned} \int_{\mathbb{R}\times E} |\theta|^{-\alpha} |\theta' - \theta|^{-\beta} \mathrm{d}\theta &\lesssim \sum_{\ell \lesssim i \sim j} 2^{-\alpha i - \beta j + 3i} + \sum_{i \lesssim j \sim \ell} 2^{-\alpha i - \beta j + 3i} + \sum_{j \lesssim i \sim \ell} 2^{-\alpha i - \beta j + 3j} \\ &\lesssim 2^{-\ell(\alpha + \beta - 3)} + 2^{-\beta \ell + (3-\alpha)_+ \ell} + 2^{-\alpha \ell + (3-\beta)_+ \ell} \lesssim 2^{-\rho \ell}, \end{aligned}$$

where ρ can be chosen as announced and where we understand that $(\delta)_+ = \varepsilon$ if $\delta = 0$.

Let us get to the estimate for the integral with $\psi_\circ$. Let $\ell' \leqslant \ell$ be such that $|k'| \sim 2^{\ell'}$ and write

$$\begin{aligned} &\int_{\mathbb{R}\times E} |\psi_\circ(k, k'-k)| |\theta|^{-\alpha} |\theta' - \theta|^{-\beta} \mathrm{d}\theta \\ &\qquad \lesssim \sum_{i,j \geqslant 0} \sum_{i' \leqslant i, j' \leqslant j} 2^{-\alpha i - \beta j} \int_{\mathbb{R}\times E} \mathbb{1}_{|k| \sim |k'-k|} \mathbb{1}_{|\theta| \sim 2^i, |\theta' - \theta| \sim 2^j} \mathbb{1}_{|k| \sim 2^{i'}, |k - k'| \sim 2^{j'}} \mathrm{d}\theta \\ &\qquad \lesssim \sum_{i,j \geqslant 0} \sum_{i' \lesssim i \wedge j} 2^{-\alpha i - \beta j} \int_{\mathbb{R}\times E} \mathbb{1}_{i' \gtrsim \ell} \mathbb{1}_{|\theta| \sim 2^i, |\theta' - \theta| \sim 2^j} \mathbb{1}_{|k| \sim 2^{i'}} \mathrm{d}\theta. \end{aligned}$$

Now we have to consider again the three possibilities $\ell \lesssim i \sim j$ or $i \lesssim j \sim \ell$ or $j \lesssim i \sim \ell$. In the first and second case we bound the integral on the right hand side by $2^{2i+i'}$, and in the third case by $2^{2j+i'}$. Then we end up with

$$\int_{\mathbb{R}\times E} |\psi_\circ(k, k'-k)| |\theta|^{-\alpha} |\theta' - \theta|^{-\beta} \mathrm{d}\theta$$

$$\lesssim \sum_{\substack{\ell \lesssim i \sim j \\ \ell' \lesssim i' \lesssim i}} 2^{-\alpha i - \beta j + 2i + i'} + \sum_{\substack{i \lesssim j \sim \ell \\ \ell' \lesssim i' \lesssim i}} 2^{-\alpha i - \beta j + 2i + i'} + \sum_{\substack{j \lesssim i \sim \ell \\ \ell' \lesssim i' \lesssim j}} 2^{-\alpha i - \beta j + 2j + i'}$$

$$\lesssim 2^{-\ell(\alpha+\beta-3)} + 2^{-\ell\beta + (3-\alpha)_+ \ell + (3-\alpha)_- \ell'} + 2^{-\ell\alpha + (3-\beta)_+ \ell + (3-\beta)_- \ell'},$$

with the same convention for $(\delta)_+$ as above. □

Lemma 9.9. *Let $\alpha > 2$. Then*

$$\int_{\mathbb{R}} \mathrm{d}\omega (|\omega|^{1/2} + |k|)^{-\alpha} \lesssim |k|^{2-\alpha}.$$

Proof. We have

$$\int_{\mathbb{R}} \mathrm{d}\omega (|\omega|^{1/2} + |k|)^{-\alpha} = |k|^{-\alpha} \int_{\mathbb{R}} \mathrm{d}\omega (|\omega |k|^{-2}|^{1/2} + 1)^{-\alpha} = |k|^{2-\alpha} \int_{\mathbb{R}} \mathrm{d}\omega (|\omega|^{1/2} + 1)^{-\alpha},$$

and if $\alpha > 2$ the integral on the right hand side is finite. □

Consider now $X^{\vee}$. Combining the bound (87) with Lemma 9.8 and Lemma 9.9, we get

$$\begin{aligned}
\mathbb{E}[(\Delta_q X^{\vee}(t,x))^2] &\lesssim \int_{(\mathbb{R}\times E)^2} \rho_q(k_{[12]})^2 |q(\theta_{[12]}) q(\theta_1) q(\theta_2)|^2 \mathrm{d}\theta_{12} \\
&= \int_{(\mathbb{R}\times E)^2} \mathrm{d}\theta_{[12]} \mathrm{d}\theta_2 \rho_q(k_{[12]})^2 |q(\theta_{[12]}) q(\theta_{[12]} - \theta_2) q(\theta_2)|^2 \\
&\lesssim \int_{(\mathbb{R}\times E)^2} \mathrm{d}\theta_{[12]} \mathrm{d}\theta_2 \rho_q(k_{[12]})^2 |q(\theta_{[12]})|^2 |\theta_{[12]} - \theta_2|^{-2} |\theta_2|^{-2} \\
&\lesssim \int_{\mathbb{R}\times E} \mathrm{d}\theta_{[12]} \rho_q(k_{[12]})^2 |q(\theta_{[12]})|^2 |\theta_{[12]}|^{-1} \lesssim \int_E \mathrm{d}k_{[12]} \frac{\rho_q(k_{[12]})^2}{|k_{[12]}|^{-1}} \lesssim 1.
\end{aligned}$$

As far as $X^{\curlyvee}$ is concerned, we have

$$\begin{aligned}
\mathbb{E}[(\Delta_q X^{\curlyvee}(t,x))^2] \lesssim{} & \int_{(\mathbb{R}\times E)^3} \rho_q(k_{[123]})^2 \left| G^{\curlyvee}(t,x,\eta_{123}) \right|^2 \mathrm{d}\eta_{123} \\
& + \int_{\mathbb{R}\times E} \rho_q(k_1)^2 \left| G^{\curlyvee\!\!\!\circ}(t,x,\eta_1) \right|^2 \mathrm{d}\eta_1 \\
\lesssim{} & \int_{(\mathbb{R}\times E)^3} \rho_q(k_{[123]})^2 |q(\theta_{[123]}) q(\theta_3) q(\theta_{[12]}) q(\theta_1) q(\theta_2)|^2 \mathrm{d}\theta_{123} \\
& + \int_{\mathbb{R}\times E} \rho_q(k_1)^2 |q(\theta_1) q(\theta_1)|^2 \left(\int_{\mathbb{R}} |V^{\curlyvee\!\!\!\circ}(\sigma,k_1)| \mathrm{d}\sigma \right)^2 \mathrm{d}\theta_1.
\end{aligned}$$

For the contraction term we already know that $\int_{\mathbb{R}} |V^{\curlyvee\!\!\!\circ}(\sigma,k_1)| \mathrm{d}\sigma \lesssim |k_1|^{\varepsilon} \lesssim |\theta_1|^{\varepsilon}$, so

$$\int_{\mathbb{R}\times E} \rho_q(k_1)^2 |q(\theta_1) q(\theta_1)|^2 \Big(\int_{\mathbb{R}} |V^{\curlyvee\!\!\!\circ}(\sigma,k_1)| \mathrm{d}\sigma \Big)^2 \mathrm{d}\theta_1 \lesssim \int_{\mathbb{R}\times E} \rho_q(k_1)^2 |\theta_1|^{-4+2\varepsilon} \mathrm{d}\theta_1$$

$$\lesssim \int_E \rho_q(k_1)^2|k_1|^{-2+2\varepsilon}\mathrm{d}k_1 \lesssim 2^{(2\varepsilon-1)q}.$$

The contribution of the third chaos can be estimated by

$$\begin{aligned}\int_{(\mathbb{R}\times E)^3} \rho_q(k_{[123]})^2|q(\theta_{[123]})q(\theta_3)q(\theta_{[12]})q(\theta_1)q(\theta_2)|^2\mathrm{d}\theta_{123}\\ \lesssim \int_{(\mathbb{R}\times E)^2} \rho_q(k_{[12]})^2|q(\theta_{[12]})q(\theta_1)q(\theta_2)|^2|\theta_2|^{-1}\mathrm{d}\theta_{12}\\ \lesssim \int_{\mathbb{R}\times E} \rho_q(k_1)^2|q(\theta_1)|^2|\theta_1|^{-2+\varepsilon}\mathrm{d}\theta_1 \lesssim 2^{(\varepsilon-1)q},\end{aligned}$$

which is enough to conclude that

$$\mathbb{E}[(\Delta_q X^{\curlyvee}(t,x))^2] \lesssim 2^{-q(1-\varepsilon)}$$

for arbitrarily small $\varepsilon > 0$.

The next term is

$$\begin{aligned}&\mathbb{E}[(\Delta_q X^{\curlyvee\!\!\circ}(t,x))^2]\\ &\quad\lesssim \int_{(\mathbb{R}\times E)^4} \rho_q(k_{[1234]})^2|\psi_\circ(k_{[123]},k_4)q(\theta_{[1234]})q(\theta_4)q(\theta_{[123]})q(\theta_3)q(\theta_{[12]})q(\theta_1)q(\theta_2)|^2\mathrm{d}\theta_{1234}\\ &\qquad+\int_{(\mathbb{R}\times E)^2} \rho_q(k_{[12]})^2|G_2^{\curlyvee\!\!\circ}(t,x,\eta_{12})|^2\mathrm{d}\eta_{12}.\end{aligned}$$

By proceeding inductively we bound

$$\begin{aligned}&\int_{(\mathbb{R}\times E)^4} \rho_q(k_{[1234]})^2|\psi_\circ(k_{[234]},k_1)q(\theta_{[1234]})q(\theta_1)q(\theta_{[234]})q(\theta_2)q(\theta_{[34]})q(\theta_3)q(\theta_4)|^2\mathrm{d}\theta_{1234}\\ &\quad\lesssim \int_{(\mathbb{R}\times E)^3} \rho_q(k_{[123]})^2|\psi_\circ(k_{[23]},k_1)q(\theta_{[123]})q(\theta_1)q(\theta_{[23]})q(\theta_2)q(\theta_3)|^2|\theta_3|^{-1}\mathrm{d}\theta_{123}\\ &\quad\lesssim \int_{(\mathbb{R}\times E)^2} \rho_q(k_{[12]})^2|\psi_\circ(k_2,k_1)q(\theta_{[12]})q(\theta_1)q(\theta_2)|^2|\theta_2|^{-2+\varepsilon}\mathrm{d}\theta_{12}\\ &\quad\lesssim \int_{\mathbb{R}\times E} \rho_q(k_1)^2|q(\theta_1)|^2|\theta_1|^{-2+\varepsilon}|k_1|^{-1+\varepsilon}\mathrm{d}\theta_1 \lesssim \int_E \rho_q(k_1)^2|k_1|^{-2+\varepsilon}|k_1|^{-1+\varepsilon}\mathrm{d}k_1 \lesssim 2^{q(-2+2\varepsilon)}.\end{aligned}$$

For the contractions we have

$$\begin{aligned}&\int_{(\mathbb{R}\times E)^2} \rho_q(k_{[12]})^2|G^{\curlyvee\!\!\circ}(t,x,\eta_{12})|^2\mathrm{d}\eta_{12}\\ &\quad\leqslant \int_{(\mathbb{R}\times E)^2} \rho_q(k_{[12]})^2|q(\theta_{[12]})q(\theta_{[12]})q(\theta_1)q(\theta_2)|^2\|V^{\circ}(\sigma,k_{[12]})\|^2_{L^1_\sigma}\mathrm{d}\theta_{12}\\ &\quad\leqslant \int_{(\mathbb{R}\times E)^2} \rho_q(k_1)^2|q(\theta_1)|^4|\theta_1-\theta_2|^{-2}|\theta_2|^{-2}|k_1|^{2\varepsilon}\mathrm{d}\theta_{12}\end{aligned}$$

$$\lesssim \int_{\mathbb{R}\times E} \rho_q(k_1)^2 |\theta_1|^{-4} |\theta_1|^{-1} |k_1|^{2\varepsilon} \mathrm{d}\theta_1 \lesssim \int_E \rho_q(k_1)^2 |k_1|^{-3+2\varepsilon} \mathrm{d}k_1 \lesssim 2^{q(-2+2\varepsilon)}.$$

The following term is

$$\begin{aligned}
&\int_{(\mathbb{R}\times E)^2} \rho_q(k_{[12]})^2 |G^{\text{[tree]}}(t,x,\eta_{12})|^2 \mathrm{d}\eta_{12} \\
&\qquad \leqslant \int_{(\mathbb{R}\times E)^2} \rho_q(k_{[12]})^2 \psi_\circ(k_1,k_2)^2 |q(\theta_{[12]}) q(\theta_1) q(\theta_1) q(\theta_2)|^2 \|V^{\text{[tree]}}(\sigma, k_{[12]})\|^2_{L^1_\sigma} \mathrm{d}\theta_{12} \\
&\qquad \leqslant \int_{(\mathbb{R}\times E)^2} \rho_q(k_1)^2 \psi_\circ(k_1-k_2,k_2)^2 |q(\theta_1)|^2 |\theta_1-\theta_2|^{-4} |\theta_2|^{-2} |k_1|^{2\varepsilon} \mathrm{d}\theta_{12} \\
&\qquad \lesssim \int_{\mathbb{R}\times E} \rho_q(k_1)^2 |q(\theta_1)|^2 |\theta_1|^{-2+\varepsilon} |k_1|^{-1+2\varepsilon} \mathrm{d}\theta_1 \lesssim \int_{\mathbb{R}\times E} \rho_q(k_1)^2 |k_1|^{-3+3\varepsilon} \mathrm{d}k_1 \lesssim 2^{q(-2+3\varepsilon)}.
\end{aligned}$$

Next, we have

$$\begin{aligned}
&\int_{(\mathbb{R}\times E)^2} \rho_q(k_{[12]})^2 |G^{\text{[tree]}}(t,x,\eta_{12})|^2 \mathrm{d}\eta_{12} \\
&\qquad \leqslant \int_{(\mathbb{R}\times E)^2} \rho_q(k_{[12]})^2 |q(\theta_{[12]}) q(\theta_1) q(\theta_2)|^2 \|V^{\text{[tree]}}(\sigma,\sigma',k_{12})\|^2_{L^1_\sigma L^1_{\sigma'}} \mathrm{d}\theta_{12} \\
&\qquad \lesssim \int_{(\mathbb{R}\times E)^2} \rho_q(k_1)^2 |q(\theta_1)|^2 |\theta_1-\theta_2|^{-2} |\theta_2|^{-2} |k_1|^{-2+2\varepsilon} \mathrm{d}\theta_{12} \\
&\qquad \lesssim \int_{\mathbb{R}\times E} \rho_q(k_1)^2 |\theta_1|^{-3} |k_1|^{-2+2\varepsilon} \mathrm{d}\theta_1 \lesssim \int_E \rho_q(k_1)^2 |k_1|^{-3+2\varepsilon} \mathrm{d}k_1 \lesssim 2^{q(-2+2\varepsilon)},
\end{aligned}$$

and therefore $\mathbb{E}[(\Delta_q X^{\text{[tree]}}(t,x))^2] \lesssim 2^{-q(2-3\varepsilon)}$.

The last term is then

$$\begin{aligned}
&\mathbb{E}[(\Delta_q X^{\text{[tree]}}(t,x))^2] \\
&\quad \lesssim \int_{(\mathbb{R}\times E)^4} \rho_q(k_{[1234]})^2 |q(\theta_{[1234]}) q(\theta_{[12]}) q(\theta_1) q(\theta_2) q(\theta_{[34]}) q(\theta_3) q(\theta_4)|^2 \mathrm{d}\theta_{1234} \\
&\qquad + \int_{(\mathbb{R}\times E)^2} \rho_q(k_{[12]})^2 |G_2^{\text{[tree]}}(t,x,\eta_{12})|^2 \mathrm{d}\eta_{12}.
\end{aligned}$$

The first term on the right hand side can be bounded as follows:

$$\begin{aligned}
&\int_{(\mathbb{R}\times E)^4} \rho_q(k_{[1234]})^2 |q(\theta_{[1234]}) q(\theta_{[12]}) q(\theta_1) q(\theta_2) q(\theta_{[34]}) q(\theta_3) q(\theta_4)|^2 \mathrm{d}\theta_{1234} \\
&\qquad\qquad \lesssim \int_{(\mathbb{R}\times E)^4} \rho_q(k_{[12]})^2 |q(\theta_{[12]}) q(\theta_1) q(\theta_2)|^2 |\theta_1|^{-1} |\theta_2|^{-1} \mathrm{d}\theta_{12} \\
&\qquad\qquad \lesssim \int_{(\mathbb{R}\times E)^4} \rho_q(k_1)^2 |q(\theta_1)|^2 |\theta_1|^{-3+\varepsilon} \mathrm{d}\theta_1 \lesssim 2^{(\varepsilon-2)q}.
\end{aligned}$$

On the other side, the contraction term is given by

$$\int_{(\mathbb{R}\times E)^2} \rho_q(k_{[12]})^2 |G_2^{\vee\vee}(t,x,\eta_{12})|^2 \mathrm{d}\eta_{12}$$
$$\lesssim \int_{(\mathbb{R}\times E)^2} \rho_q(k_{[12]})^2 |q(\theta_{[12]})q(\theta_1)q(\theta_2)|^2 \|V^{\vee\vee}(\sigma,\sigma',k_{12})\|^2_{L^1_\sigma L^1_{\sigma'}} \mathrm{d}\theta_{12}$$
$$\lesssim \int_{(\mathbb{R}\times E)^2} \rho_q(k_1)^2 |q(\theta_1)|^2 |\theta_1-\theta_2|^{-2}|\theta_2|^{-2}|k_1|^{-2+2\varepsilon}\mathrm{d}\theta_{12}$$
$$\lesssim \int_{\mathbb{R}\times E} \rho_q(k_1)^2 |\theta_1|^{-3}|k_1|^{-2+2\varepsilon}\mathrm{d}\theta_1 \lesssim \int_E \rho_q(k_1)^2|k_1|^{-3+2\varepsilon}\mathrm{d}k_1 \lesssim 2^{q(-2+2\varepsilon)},$$

so we can conclude that $\mathbb{E}[(\Delta_q X^{\vee\vee}(t,x))^2] \lesssim 2^{-q(2-2\varepsilon)}$.

9.6 Divergences in the KPZ equation

The data we still need to control for the KPZ equation is

$$Y, Y^{\vee}, Y^{\vee^{\bullet}}, Y^{\vee^{\vee}}, Y^{\vee\vee}$$

since $Q \circ X$ was already dealt with. The kernels for the chaos decomposition of these random fields are given by

$$\tilde{G}^{\bullet}_t(k) = \mathbb{1}_{t\geqslant 0} e^{-tk^2}, \qquad \tilde{G}^{(\tau_1,\tau_2)}(t,x,\eta_\tau) = \int_0^t \mathrm{d}\sigma P_{t-\sigma}(G^{\tau_1}(\sigma,\cdot,\eta_{\tau_1})G^{\tau_2}(\sigma,\cdot,\eta_{\tau_2}))(x),$$

so they enjoy similar estimates as the kernels G^τ and therefore all the chaos components different from the 0-th are under control. The only difference is the missing derivative which in the case of the X^τ is responsible for the fact that the constant component in the chaos expansion vanishes. The 0-th component it given by

$$tc^\tau = \mathbb{E}[Y^\tau(t,x)].$$

Some of these expectations happen to be infinite which will force us to renormalize Y^τ by subtracting its mean.

For $Y^{\vee}_\varepsilon$ we have

$$\mathbb{E}[Y^{\vee}_\varepsilon(t,x)] = \int_0^t \mathbb{E}[P_{t-\sigma}(X_\varepsilon(\sigma,\cdot)^2)(x)]\mathrm{d}\sigma$$
$$= \int_0^t \mathrm{d}\sigma \int_{\mathbb{R}\times E} \mathrm{d}\eta_1 e^{ik_{1(-1)}x} e^{-(t-\sigma)k^2_{1(-1)}} \varphi(\varepsilon k_1) H_{\sigma-s_1}(k_1)\varphi(\varepsilon k_{-1})H_{\sigma-s_1}(k_{-1})$$
$$= \int_0^t \mathrm{d}\sigma \int_{\mathbb{R}\times E} \mathrm{d}\eta_1 \varphi^2(\varepsilon k_1) H_{\sigma-s_1}(k_1)H_{\sigma-s_1}(k_{-1}) = \frac{t}{2}\int_E \mathrm{d}k_1 \varphi^2(\varepsilon k_1),$$

and since $\varphi \in C^1$ we get

$$\frac{1}{\varepsilon}\Big(\int_{\mathbb{R}} \varphi^2(x)\mathrm{d}x - \varepsilon \int_E \mathrm{d}k_1 \varphi^2(\varepsilon k_1)\Big) = \frac{1}{\varepsilon}\Big(\sum_k \int_{\varepsilon k}^{\varepsilon(k+1)} (\varphi^2(x) - \varphi^2(\varepsilon k))\mathrm{d}x\Big)$$
$$= \frac{1}{\varepsilon}\sum_k \int_0^1 \int_0^\varepsilon \mathrm{D}(\varphi^2)(\varepsilon k + \lambda x)x\mathrm{d}x\mathrm{d}\lambda,$$

which converges to $\int_{\mathbb{R}} \mathrm{D}\varphi^2(x)\mathrm{d}x = 0$ as $\varepsilon \to 0$. Now recall that here we are dealing with $(2\pi)^{1/2}$ times the white noise, so that the constant $c_\varepsilon^{\vee}$ of Theorem 9.3 can be chosen as

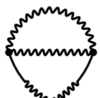
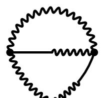
$$c_\varepsilon^{\vee} = \frac{1}{4\pi\varepsilon}\int_{\mathbb{R}} \varphi^2(x)\mathrm{d}x.$$

The next term is $Y^{\vee\!\vee}$, which has mean zero since it belongs to the odd chaoses and thus $c^{\vee\!\vee} = 0$.

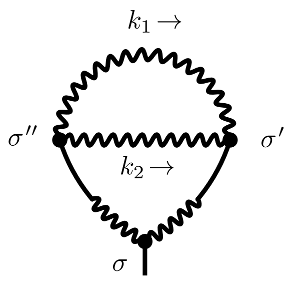
What we want to show now is that a special symmetry of the equation induces cancellations which are responsible for the fact that while $c^{\vee\vee}$ and $c^{\vee\!\vee\!\vee}$ are separately divergent, the particular combination

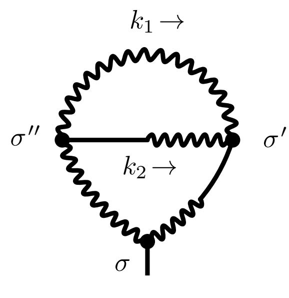
$$c^{\vee\vee} + 4c^{\vee\!\vee\!\vee}$$

is actually finite. If this is true, then we can renormalize $Y^{\vee\!\vee\!\vee}$ and $Y^{\vee\vee}$ as announced in Theorem 9.3. In terms of Feynman diagrams (which have the same translation into kernels as for Burgers equation, except that for the outgoing line the factor $ik_{[\tau]}$ is suppressed), this quantity is given by

$$I = c^{\vee\vee} + 4c^{\vee\!\vee\!\vee} = 2 \times$$  $$+ 8 \times$$  ,

because all other contractions vanish since they involve contractions of the topmost leaves (e.g. $\tilde{G}^{\vee\!\vee\!\vee}(t, x, \eta_{1(-1)2(-2)}) = 0$). Moreover, writing explicitly the remaining two contributions and fixing the integration variables according to the following picture

$$I = 2 \times$$ 

$$+ 8 \times$$ 

,

we get

$$I = 2\int \mathrm{d}k_1\mathrm{d}k_2 \int_0^t \mathrm{d}\sigma \int_0^\sigma \mathrm{d}\sigma' \int_0^\sigma \mathrm{d}\sigma''(ik_{[12]})(-ik_{[12]})e^{-|\sigma'-\sigma''|(k_1^2+k_2^2)-(\sigma-\sigma')k_{[12]}^2-(\sigma-\sigma'')k_{[12]}^2}$$
$$+ 8\int \mathrm{d}k_1\mathrm{d}k_2 \int_0^t \mathrm{d}\sigma \int_0^\sigma \mathrm{d}\sigma' \int_0^{\sigma'} \mathrm{d}\sigma''(ik_{[12]})(ik_2)e^{-(\sigma'-\sigma'')(k_1^2+k_2^2)-(\sigma-\sigma')k_{[12]}^2-(\sigma-\sigma'')k_{[12]}^2}.$$

Now note that the second term can be symmetrized over σ', σ'' to get

$$I = 2\int \mathrm{d}k_1\mathrm{d}k_2 \int_0^t \mathrm{d}\sigma \int_0^\sigma \mathrm{d}\sigma' \int_0^\sigma \mathrm{d}\sigma''(ik_{[12]})(-ik_{[12]})e^{-|\sigma'-\sigma''|(k_1^2+k_2^2)-(\sigma-\sigma')k_{[12]}^2-(\sigma-\sigma'')k_{[12]}^2}$$
$$+ 4\int \mathrm{d}k_1\mathrm{d}k_2 \int_0^t \mathrm{d}\sigma \int_0^\sigma \mathrm{d}\sigma' \int_0^{\sigma} \mathrm{d}\sigma''(ik_{[12]})(ik_2)e^{-|\sigma'-\sigma''|(k_1^2+k_2^2)-(\sigma-\sigma')k_{[12]}^2-(\sigma-\sigma'')k_{[12]}^2}.$$

At this point the second term can still be symmetrized over k_2, k_1 to finally get

$$I = 2\int \mathrm{d}k_1\mathrm{d}k_2 \int_0^t \mathrm{d}\sigma \int_0^\sigma \mathrm{d}\sigma' \int_0^\sigma \mathrm{d}\sigma''(ik_{[12]})(-ik_{[12]})e^{-|\sigma'-\sigma''|(k_1^2+k_2^2)-(\sigma-\sigma')k_{[12]}^2-(\sigma-\sigma'')k_{[12]}^2}$$
$$+ 2\int \mathrm{d}k_1\mathrm{d}k_2 \int_0^t \mathrm{d}\sigma \int_0^\sigma \mathrm{d}\sigma' \int_0^{\sigma} \mathrm{d}\sigma''(ik_{[12]})(ik_{[12]})e^{-|\sigma'-\sigma''|(k_1^2+k_2^2)-(\sigma-\sigma')k_{[12]}^2-(\sigma-\sigma'')k_{[12]}^2}$$

and conclude that $I = 0$. For the moment this computation is only formal since we did not take properly into account the regularization. Introducing the regularization φ on the noise, the integral to be considered is

$$I_\varepsilon = 2\int \mathrm{d}k_1\mathrm{d}k_2\varphi^2(\varepsilon k_1)\varphi^2(\varepsilon k_2) \int_0^t \mathrm{d}\sigma \int_0^\sigma \mathrm{d}\sigma' \int_0^\sigma \mathrm{d}\sigma'' \times$$
$$\times (ik_{[12]})(-ik_{[12]})e^{-|\sigma'-\sigma''|(k_1^2+k_2^2)-(\sigma-\sigma')k_{[12]}^2-(\sigma-\sigma'')k_{[12]}^2}$$
$$+ 8\int \mathrm{d}k_1\mathrm{d}k_2\varphi^2(\varepsilon k_1)\varphi^2(\varepsilon k_{[12]}) \int_0^t \mathrm{d}\sigma \int_0^\sigma \mathrm{d}\sigma' \int_0^{\sigma'} \mathrm{d}\sigma''(ik_{[12]})(ik_2)\times$$
$$\times e^{-(\sigma'-\sigma'')(k_1^2+k_2^2)-(\sigma-\sigma')k_{[12]}^2-(\sigma-\sigma'')k_{[12]}^2}.$$

Here the symmetrization of σ' and σ'' can still be performed, giving

$$I_\varepsilon = 2\int_{E^2} \mathrm{d}k_1\mathrm{d}k_2 \int_0^t \mathrm{d}\sigma \int_0^\sigma \mathrm{d}\sigma' \int_0^\sigma \mathrm{d}\sigma''(ik_{[12]})[\varphi^2(\varepsilon k_1)\varphi^2(\varepsilon k_2)(-ik_{[12]}) + 2\varphi^2(\varepsilon k_1)\varphi^2(\varepsilon k_{[12]})(ik_2)]\times$$
$$\times e^{-|\sigma'-\sigma''|(k_1^2+k_2^2)-(\sigma-\sigma')k_{[12]}^2-(\sigma-\sigma'')k_{[12]}^2},$$

which is equivalent to

$$I_\varepsilon = 4\int_{E^2} \mathrm{d}k_1\mathrm{d}k_2 \int_0^t \mathrm{d}\sigma \int_0^\sigma \mathrm{d}\sigma' \int_0^\sigma \mathrm{d}\sigma''(ik_{[12]})\varphi^2(\varepsilon k_1)(ik_2)[\varphi^2(\varepsilon k_{[12]}) - \varphi^2(\varepsilon k_2)]\times$$
$$\times e^{-|\sigma'-\sigma''|(k_1^2+k_2^2)-\sigma' k_{[12]}^2-\sigma'' k_{[12]}^2}.$$

Now perform the change of variables $\sigma' \to \varepsilon^2\sigma'$, $\sigma'' \to \varepsilon^2\sigma''$ to obtain

$$I_\varepsilon = 4\varepsilon^2 \int_{E^2} \mathrm{d}k_1 \mathrm{d}k_2 \int_0^t \mathrm{d}\sigma \int_0^{\sigma/\varepsilon^2} \mathrm{d}\sigma' \int_0^{\sigma/\varepsilon^2} \mathrm{d}\sigma'' (i\varepsilon k_{[12]})\varphi^2(\varepsilon k_1)(i\varepsilon k_2)[\varphi^2(\varepsilon k_{[12]}) - \varphi^2(\varepsilon k_2)] \times \\ \times e^{-|\sigma'-\sigma''|(\varepsilon^2 k_1^2 + \varepsilon^2 k_2^2) - (\sigma'+\sigma'')\varepsilon^2 k_{[12]}^2}.$$

By taking the limit $\varepsilon \to 0$, the two sums over k_1 and k_2 become integrals:

$$\begin{aligned}
\lim_{\varepsilon\to 0} I_\varepsilon &= 4t \int_0^\infty \mathrm{d}\sigma' \int_0^\infty \mathrm{d}\sigma'' \int_{\mathbb{R}^2} \mathrm{d}k_1 \mathrm{d}k_2 (ik_{[12]})\varphi^2(k_1)(ik_2)[\varphi^2(k_{[12]}) - \varphi^2(k_2)] \times \\
&\qquad \times e^{-|\sigma'-\sigma''|(k_1^2+k_2^2) - (\sigma'+\sigma'')k_{[12]}^2} \\
&= 4t \int_{\mathbb{R}^2} \mathrm{d}k_1 \mathrm{d}k_2 (ik_{[12]})\varphi^2(k_1)(ik_2)[\varphi^2(k_{[12]}) - \varphi^2(k_2)] \frac{1}{k_{[12]}^2} \frac{1}{k_1^2 + k_2^2 + k_{[12]}^2} \\
&= 2t \int_{\mathbb{R}^2} \mathrm{d}k_1 \mathrm{d}k_2 (ik_{[12]})[2\varphi^2(k_1)\varphi^2(k_{[12]})(ik_2) - \varphi^2(k_1)(ik_{[12]})\varphi^2(k_2)] \frac{1}{k_{[12]}^2} \frac{1}{k_1^2 + k_2^2 + k_{[12]}^2} \\
&= 2t \int_{\mathbb{R}^2} \mathrm{d}k_1 \mathrm{d}k_2 (ik_{[12]}) \frac{\varphi^2(k_1)(ik_{[12]})(\varphi^2(k_{[12]}) - \varphi^2(k_2)) + \varphi^2(k_1)\varphi^2(k_{[12]}) i(k_2 - k_1)}{k_{[12]}^2 (k_1^2 + k_2^2 + k_{[12]}^2)} \\
&= -2t \int_{\mathbb{R}^2} \mathrm{d}k_1 \mathrm{d}k_2 \left[\varphi^2(k_1)(\varphi^2(k_{[12]}) - \varphi^2(k_2)) + \varphi^2(k_1)\varphi^2(k_{[12]}) \frac{(k_2 - k_1)}{k_{[12]}} \right] \frac{1}{k_1^2 + k_2^2 + k_{[12]}^2},
\end{aligned}$$

which is indeed finite.

10 Stochastic data for the Sasamoto-Spohn model

10.1 Convergence of the RBE-enhancement

Here we study the convergence of the data

$$X_N, X_N^{\vee}, X_N^{\vee\vee}, X_N^{\vee\vee\circ}, X_N^{\vee\vee\vee}, Q_N, B_N(Q_N \circ X_N)$$

for the discrete Burgers equation (76). We will pick up some correction terms as we pass to the limit, which is due to the fact that in the continuous setting not all kernels were absolutely integrable and at some points we used certain symmetries that are violated now. We work in the following setting:

Assumption (f,g,h)**.** *Let* $f, g, h \in C_b^1(\mathbb{R}, \mathbb{C})$ *satisfy* $f(0) = g(0) = h(0,0) = 1$ *and assume that* f *is a real valued even function with* $f(x) \geqslant c_f > 0$ *for all* $|x| \leqslant \pi$, *that* $g(-x) = g(x)^*$, *and that* $h(x,y) = h(y,x)$ *and* $h(-x,-y) = h(x,y)^*$; *we write* $\bar{h}(x) = h(x,0)$. *Define for* $N \in \mathbb{N}$

$$\mathscr{F}\Delta_N \varphi(k) = -|k|^2 f_\varepsilon(k) \mathscr{F}\varphi(k), \qquad \mathscr{F}\mathrm{D}_N \varphi(k) = ikg_\varepsilon(k)\mathscr{F}\varphi(k),$$
$$\mathscr{F}B_N(\varphi, \psi)(k) = (2\pi)^{-1} \sum_\ell \mathscr{F}\varphi(\ell)\mathscr{F}\psi(k-\ell) h_\varepsilon(\ell, (k-\ell)),$$

where $\varepsilon = 2\pi/N$ *and* $f_\varepsilon(x) = f(\varepsilon x)$ *and similarly for* g *and* h*. Let* ξ *be a space-time white noise and set*

$$\mathbb{X}_N(\xi) = (X_N(\xi), X_N^{\vee}(\xi), X_N^{\curlyvee}(\xi), X_N^{\veebar}(\xi), X_N^{\vee\vee}(\xi), B_N(Q_N \circ X_N)(\xi)),$$

where

$$\begin{aligned}
\mathscr{L}X_N(\xi) &= \mathrm{D}_N\mathcal{P}_N\xi, \\
\mathscr{L}X_N^{\vee}(\xi) &= \mathrm{D}_N\Pi_N B_N(X_N(\xi), X_N(\xi)), \\
\mathscr{L}X_N^{\curlyvee}(\xi) &= \mathrm{D}_N\Pi_N B_N(X_N(\xi), X_N^{\vee}(\xi)), \\
\mathscr{L}X_N^{\veebar}(\xi) &= \mathrm{D}_N\Pi_N B_N(X_N^{\curlyvee}(\xi) \circ X_N(\xi)), \\
\mathscr{L}X_N^{\vee\vee}(\xi) &= \mathrm{D}_N\Pi_N B_N(X_N^{\vee}(\xi), X_N^{\vee}(\xi)), \\
\mathscr{L}Q_N(\xi) &= \mathrm{D}_N B_N(X_N(\xi), 1),
\end{aligned} \tag{88}$$

all with zero initial conditions except $X_N(\xi)$ *for which we choose the "stationary" initial condition*

$$X_N(\xi)(0) = \int_{-\infty}^0 e^{-sf_\varepsilon(\mathrm{D})|\mathrm{D}|^2}\mathrm{D}_N\mathcal{P}_N\xi(s)\mathrm{d}s.$$

We are ready to state the main result of this section:

Theorem 10.1. *Let* ξ *be a space-time white noise and make Assumption (f,g,h). Then for all* $0 < \delta < T$ *and* $p \geq 1$ *the sequence* $\mathbb{X}_N = \mathbb{X}_N(\xi)$ *converges to*

$$\tilde{\mathbb{X}} = (X, X^{\vee}, X^{\curlyvee} + 2cQ, X^{\vee\vee}, X^{\veebar} + cQ^{\vee} + 2cQ^{Q\circ X}, Q \circ X + c),$$

in $L^p[\Omega, C_T\mathscr{C}^{\alpha-1} \times C_T\mathscr{C}^{2\alpha-1} \times \mathscr{L}_T^{\alpha} \times \mathscr{L}_T^{2\alpha} \times \mathscr{L}_T^{2\alpha} \times C([\delta, T], \mathscr{C}^{2\alpha-1})]$ *and is uniformly bounded in* $L^p[\Omega, C_T\mathscr{C}^{\alpha-1} \times C_T\mathscr{C}^{2\alpha-1} \times \mathscr{L}_T^{\alpha} \times \mathscr{L}_T^{2\alpha} \times \mathscr{L}_T^{2\alpha} \times C_T\mathscr{C}^{2\alpha-1}]$*. Here we wrote*

$$c = -\frac{1}{4\pi}\int_0^\pi \frac{\mathrm{Im}(g(x)\bar{h}(x))}{x}\frac{h(x,-x)|g(x)|^2}{|f(x)|^2}\mathrm{d}x \in \mathbb{R}. \tag{89}$$

and

$$\mathscr{L}Q^{\vee} = \mathrm{D}X^{\vee}, \qquad \mathscr{L}Q^{Q\circ X} = \mathrm{D}(Q \circ X),$$

both with initial condition 0.

The proof of this theorem will occupy us for the remainder of the section. Define the kernel

$$H_t^N(k) = \mathbb{1}_{t\geqslant 0}\mathbb{1}_{|k|<N/2}ikg_\varepsilon(k)e^{-tk^2f_\varepsilon(k)}$$

which satisfies

$$\int_{\mathbb{R}} H_{s-\sigma}^N(k)H_{t-\sigma}^N(-k)\mathrm{d}\sigma = \mathbb{1}_{|k|<N/2}\frac{|g_\varepsilon(k)|^2}{f_\varepsilon(k)}\frac{e^{-k^2f_\varepsilon(k)|t-s|}}{2} \tag{90}$$

for all $k \in E$. Let us start then by analyzing the resonant product $B_N(Q_N \circ X_N)$.

Lemma 10.2. *For all* $0 < \delta < T$*, the resonant products* $(B_N(Q_N \circ X_N))$ *are bounded in* $L^p(\Omega, C_T\mathscr{C}^{2\alpha-1})$ *and converge to* $Q \circ X + c$ *in* $L^p(\Omega, C([\delta, T], \mathscr{C}^{2\alpha-1}))$.

Proof. We have the chaos decomposition

$$B_N(Q_N \circ X_N)(t,x) = \int_{(\mathbb{R}\times E)^2} K_N(t,x,\eta_{12}) W(\mathrm{d}\eta_{12}) + (2\pi)^{-1} \int_{\mathbb{R}\times E} K_N(t,x,\eta_{1(-1)}) \mathrm{d}\eta_1,$$

where

$$K_N(t,x,\eta_{12}) = \int_0^t \mathrm{d}\sigma e^{ik_{[12]}x} \psi_\circ(k_1,k_2) h_\varepsilon(k_1,k_2) H^N_{t-\sigma}(k_1) \bar{h}_\varepsilon(k_1) H^N_{\sigma-s_1}(k_1) H^N_{t-s_2}(k_2).$$

The kernel K_N satisfies the same bounds as the kernel in the definition of $Q \circ X$ and converges pointwise to it, so that the convergence of the second order Wiener-Itô integral over K_N to $Q \circ X$ in $L^p(\Omega, C_T \mathscr{C}^{2\alpha-1})$ follows from the dominated convergence theorem. For the term in the chaos of order 0 we use (90) to obtain

$$\begin{aligned}
&\int_{\mathbb{R}\times E} K^N(t,x,\eta_{1(-1)}) \mathrm{d}\eta_1 \\
&\quad = \int_E \mathrm{d}k_1 \int_0^t \mathrm{d}\sigma h_\varepsilon(k_1,-k_1) H^N_{t-\sigma}(k_1) \bar{h}_\varepsilon(k_1) \frac{|g_\varepsilon(k_1)|^2}{f_\varepsilon(k_1)} \frac{e^{-k_1^2 f_\varepsilon(k_1)(t-\sigma)}}{2} \\
&\quad = \int_E \mathrm{d}k_1 \int_0^t \mathrm{d}\sigma h_\varepsilon(k_1,-k_1) \mathbb{1}_{|k_1|<N/2} i k_1 g_\varepsilon(k_1) \bar{h}_\varepsilon(k_1) \frac{|g_\varepsilon(k_1)|^2}{f_\varepsilon(k_1)} \frac{e^{-2k_1^2 f_\varepsilon(k_1)(t-\sigma)}}{2} \\
&\quad = \varepsilon \sum_{|k_1|<N/2} \mathbb{1}_{k_1\neq 0} i \frac{g(\varepsilon k_1)\bar{h}(\varepsilon k_1)}{\varepsilon k_1} \frac{h(\varepsilon k_1, -\varepsilon k_1)|g(\varepsilon k_1)|^2}{4 f^2(\varepsilon k_1)} (1 - e^{-2k_1^2 f_\varepsilon(k_1) t}).
\end{aligned} \tag{91}$$

It is not hard to see that the sum involving the exponential correction term converges to zero uniformly in $t \in [\delta, T]$ whenever $\delta > 0$. For the remaining sum, note that

$$i\frac{g(-\varepsilon k_1)\bar{h}(-\varepsilon k_1)}{-\varepsilon k_1} = -i\Big(\frac{g(\varepsilon k_1)\bar{h}(\varepsilon k_1)}{\varepsilon k_1}\Big)^*,$$

while the second fraction on the right hand side of (91) is an even function of k_1, and thus

$$\begin{aligned}
\varepsilon & \sum_{|k_1|<N/2} \mathbb{1}_{k_1\neq 0} i \frac{g(\varepsilon k_1)\bar{h}(\varepsilon k_1)}{\varepsilon k_1} \frac{h(\varepsilon k_1, -\varepsilon k_1)|g(\varepsilon k_1)|^2}{4 f^2(\varepsilon k_1)} \\
&\qquad = -\varepsilon \sum_{0<k_1<N/2} \frac{\mathrm{Im}(g(\varepsilon k_1)\bar{h}(\varepsilon k_1))}{\varepsilon k_1} \frac{h(\varepsilon k_1, -\varepsilon k_1)|g(\varepsilon k_1)|^2}{2 f^2(\varepsilon k_1)}.
\end{aligned}$$

Now $\mathrm{Im}(g\bar{h}) \in C_b^1$ with $\mathrm{Im}(g\bar{h})(0) = 0$, and therefore $\mathrm{Im}(g(x)\bar{h}(x))/x$ is a bounded continuous function. Since furthermore $f(\varepsilon k_1) \geqslant c_f > 0$ for all $|k_1| < N/2$, the right hand side is a Riemann sum and converges to $2\pi c$ which is what we wanted to show. □

Next, define

$$E_N = E \cap (-N/2, N/2)$$

and the vertex function

$$V_N^{\vee}(\sigma, k_1) = \mathbb{1}_{|k_1| \leqslant N/2} \int_{E_N} h_\varepsilon(k_{[12]}^N, k_{-2}) h_\varepsilon(k_1, k_2) H_\sigma^N(k_{[12]}^N) \frac{|g_\varepsilon(k_2)|^2}{f_\varepsilon(k_2)} \frac{e^{-k_2^2 f_\varepsilon(k_2)|\sigma|}}{2} \mathrm{d}k_2. \tag{92}$$

We also set

$$V_N^{\between}(\sigma, k_1) = \mathbb{1}_{|k_1| \leqslant N/2} \int_{E_N} \mathrm{d}k_2 \psi_\circ(k_{[12]}^N, k_{-2}) h_\varepsilon(k_{[12]}^N, k_{-2}) \times h_\varepsilon(k_1, k_2) H_\sigma^N(k_{[12]}^N) \frac{|g_\varepsilon(k_2)|^2}{f_\varepsilon(k_2)} \frac{e^{-k_2^2 f_\varepsilon(k_2)|\sigma|}}{2}.$$

Lemma 10.3. *Let $c \in \mathbb{R}$ be the constant defined in* (89)*. Then $V_N^{\vee}(\cdot, k)$ converges weakly to $V^{\vee}(\sigma, k) + 2\pi c\delta(\sigma)$ for all $k \in E$, in the sense that*

$$\lim_{N \to \infty} \int_0^\infty \varphi(\sigma) V_N^{\vee}(\sigma, k) \mathrm{d}\sigma = \int_0^\infty \varphi(\sigma) V^{\vee}(\sigma, k) \mathrm{d}\sigma + 2\pi c \varphi(0)$$

for all measurable $\varphi\colon [0, \infty) \to \mathbb{R}$ with $|\varphi(\sigma) - \varphi(0)| \lesssim |\sigma|^\kappa$ for some $\kappa > 0$. Similarly, $V_N^{\between}(\cdot, k)$ converges weakly to $V^{\between}(\sigma, k) + 2\pi c\delta(\sigma)$ for all $k \in E$. Moreover, for all $\delta > 0$ we have

$$\sup_N \int_0^\infty |V_N^{\vee}(\sigma, k)| \mathrm{d}\sigma + \sup_N \int_0^\infty |V_N^{\between}(\sigma, k)| \mathrm{d}\sigma \lesssim |k|^\delta.$$

Proof. We write $V_N^{\vee}(\sigma, k) = (V_N^{\vee}(\sigma, k) - \mathbb{1}_{|k| \leqslant N/2} V_N^{\vee}(\sigma, 0)) + \mathbb{1}_{|k| \leqslant N/2} V_N^{\vee}(\sigma, 0)$. Let us first concentrate on the second term:

$$\begin{aligned}
&\int_0^\infty \mathrm{d}\sigma \varphi(\sigma) V_N^{\vee}(\sigma, 0) \\
&\quad = \int_{E_N} \mathrm{d}k_2 h_\varepsilon(k_2, -k_2) \bar{h}_\varepsilon(k_2)(ik_2) g_\varepsilon(k_2) \frac{|g_\varepsilon(k_2)|^2}{2 f_\varepsilon(k_2)} \int_0^\infty \mathrm{d}\sigma \varphi(\sigma) e^{-2k_2^2 f_\varepsilon(k_2)\sigma} \\
&\quad = - \sum_{0 < |k_2| \leqslant N/2} h_\varepsilon(k_2, -k_2) k_2 \mathrm{Im}(\bar{h}_\varepsilon(k_2) g_\varepsilon(k_2)) \frac{|g_\varepsilon(k_2)|^2}{f_\varepsilon(k_2)} \int_0^\infty \mathrm{d}\sigma \varphi(\sigma) e^{-2k_2^2 f_\varepsilon(k_2)\sigma}.
\end{aligned} \tag{93}$$

Now add and subtract $\varphi(0)$ in the integral over σ and observe that

$$\Big| \int_0^\infty \mathrm{d}\sigma (\varphi(\sigma) - \varphi(0)) e^{-2k_2^2 f_\varepsilon(k_2)\sigma} \Big| \lesssim \int_0^\infty \mathrm{d}\sigma \sigma^\kappa e^{-2k_2^2 c_f \sigma} \lesssim |k_2|^{-2\kappa} \int_0^\infty \mathrm{d}\sigma e^{-ck_2^2\sigma} \lesssim |k_2|^{-2-2\kappa},$$

and then

$$\begin{aligned}
&\Big| \sum_{0 < |k_2| \leqslant N/2} h_\varepsilon(k_2, -k_2) k_2 \mathrm{Im}(\bar{h}_\varepsilon(k_2) g_\varepsilon(k_2)) \frac{|g_\varepsilon(k_2)|^2}{f_\varepsilon(k_2)} |k_2|^{-2-2\kappa} \Big| \\
&\lesssim \varepsilon^{\kappa'} \Big(\varepsilon \sum_{0 < |k_2| \leqslant N/2} |h(\varepsilon k_2, -\varepsilon k_2) \mathrm{Im}(\bar{h}(\varepsilon k_2) g(\varepsilon k_2))| \frac{|g(\varepsilon k_2)|^2}{f(\varepsilon k_2)} |\varepsilon k_2|^{-1-\kappa'} \Big)
\end{aligned}$$

whenever $\kappa' \in [0, 1 \wedge 2\kappa)$. The term in the brackets is a Riemann sum and since $|Im(\bar{h}g)(x)| \lesssim |x|$ and $\kappa' < 1$, it converges to a finite limit. Thus we may replace $\varphi(\sigma)$ by $\varphi(0)$ and end up with

$$-\varphi(0) \sum_{0 < |k_2| \leqslant N/2} h_\varepsilon(k_2, -k_2) \mathrm{Im}(\bar{h}_\varepsilon(k_2) g_\varepsilon(k_2)) \frac{|g_\varepsilon(k_2)|^2}{2 f_\varepsilon^2(k_2) k_2},$$

which converges to $2\pi c \varphi(0)$.

It remains to treat the term

$$V_N^{\Upsilon}(\sigma, k) - \mathbb{1}_{|k| \leqslant N/2} V_N^{\Upsilon}(\sigma, 0) = \mathbb{1}_{|k| \leqslant N/2} \int_{E_N} \mathrm{d}k_2 (W_N^{\Upsilon}(\sigma, k, k_2) - W_N^{\Upsilon}(\sigma, 0, k_2)),$$

where the right hand side is to be understood as the (indirect) definition of W_N^{Υ}. Since for fixed (k, k_2, σ) the integrand converges to $(H_\sigma(k+k_2) - H_\sigma(k_2)) e^{-k_2^2 |\sigma|}/2$, it suffices to bound $|W_N^{\Upsilon}(\sigma, k, k_2) - W_N^{\Upsilon}(\sigma, 0, k_2)|$ uniformly in N by an expression which is integrable over $(\sigma, k_2) \in [0, \infty) \times E$. For the remainder of the proof let us write $\psi_\varepsilon(\ell, m) = h_\varepsilon((\ell + m)^N, -m) h_\varepsilon(\ell, m) g_\varepsilon((\ell + m)^N)$, which satisfies uniformly over $|\ell|, |m| < N/2$

$$|\psi_\varepsilon(\ell, m) - \psi_\varepsilon(0, m)| \lesssim \varepsilon |\ell| \lesssim (\varepsilon |\ell|)^\delta$$

for all $\delta \in [0, 1]$ (to bound $|(\ell + m)^N - m|$ note that $(\ell + m)^N = \ell + m + j(m, \ell) N$ for some $j(m, \ell) \in \mathbb{Z}$ and that if $|\ell| < N/2$, then $|\ell + jN| \geqslant |\ell|$ for all $j \in \mathbb{Z}$). We now have to estimate the term

$$\begin{aligned}
&|\psi_\varepsilon(k, k_2)(k + k_2)^N e^{-\sigma f_\varepsilon((k+k_2)^N)((k+k_2)^N)^2} - \psi_\varepsilon(0, k_2) k_2 e^{-\sigma f_\varepsilon(k_2) k_2^2} | \mathbb{1}_{|k|, |k_2| < N/2} e^{-k_2^2 f_\varepsilon(k_2) \sigma} \\
&\quad \leqslant |\psi_\varepsilon(k, k_2) - \psi_\varepsilon(0, k_2)| |k_2| e^{-\sigma f_\varepsilon(k_2) k_2^2} \mathbb{1}_{|k|, |k_2| < N/2} e^{-k_2^2 c_f \sigma} \\
&\qquad + |\psi_\varepsilon(k, k_2)| |(k + k_2)^N e^{-\sigma f_\varepsilon((k+k_2)^N)((k+k_2)^N)^2} - k_2 e^{-\sigma f_\varepsilon(k_2) k_2^2} | \mathbb{1}_{|k|, |k_2| < N/2} e^{-k_2^2 c_f \sigma} \\
&\quad \lesssim ((\varepsilon |k|)^\delta |k_2| + |(k + k_2)^N e^{-\sigma f_\varepsilon((k+k_2)^N)((k+k_2)^N)^2} - k_2 e^{-\sigma f_\varepsilon(k_2) k_2^2} |) \mathbb{1}_{|k|, |k_2| < N/2} e^{-k_2^2 c_f \sigma}.
\end{aligned}$$

The first addend on the right hand side is bounded by $|k|^\delta |k_2|^{1-\delta} e^{-\sigma c_f k_2^2}$, and for $\delta > 0$ this is integrable in (σ, k_2) and the integral is $\lesssim |k|^\delta$. To bound the second addend, let us define $\varphi_{\varepsilon, \sigma}(x) = x e^{-\sigma f_\varepsilon(x) x^2}$ and note that for $|x| < N/2$

$$|\varphi'_{\varepsilon, \sigma}(x)| = |1 - x\sigma f'(\varepsilon x) \varepsilon x^2 - x \sigma f_\varepsilon(x) 2x | e^{-\sigma f_\varepsilon(x) x^2} \lesssim (1 + \sigma |x|^2) e^{-c_f \sigma x^2} \lesssim e^{-\sigma c x^2}$$

for some $c > 0$. From here we can use the same arguments as in the proof of Lemma 9.5 to show that

$$|(k + k_2)^N e^{-\sigma f_\varepsilon((k+k_2)^N)((k+k_2)^N)^2} - k_2 e^{-\sigma f_\varepsilon(k_2) k_2^2} |) \mathbb{1}_{|k|, |k_2| < N/2} e^{-k_2^2 c_f \sigma} \leqslant F(k, k_2, \sigma)$$

for a function F with $\int_E \mathrm{d}k_2 \int_0^\infty \mathrm{d}\sigma F(k, k_2, \sigma) \lesssim |k|^\delta$, and thus we get the convergence and the bound for V_N^{Υ}. The term V_N^{8} can be treated using the same arguments. □

Proof of Theorem 10.1. We introduce analogous kernels as in the continuous setting: Define $G_N^{\bullet}(t,x,\eta) = e^{ikx} H_{t-s}^N(k)$ and and then inductively for $\tau = (\tau_1\tau_2)$

$$\begin{aligned} G_N^\tau(t,x,\eta_\tau) &= e^{ik_{[\tau]}^N x} H_N^\tau(t,\eta_\tau) \\ &= e^{ik_{[\tau]}^N x} \int_0^t \mathrm{d}\sigma H_{t-\sigma}^N(k_{[\tau]}^N) h_\varepsilon(k_{[\tau_1]}^N, k_{[\tau_2]}^N) H_N^{\tau_1}(\sigma,\eta_{\tau_1}) H_N^{\tau_2}(\sigma,\eta_{\tau_2}). \end{aligned}$$

A first consequence of this recursive description is that the contribution to the 0-th chaos always vanishes: just as for the kernels G^τ of Section 9 we have

$$G_N^\tau(t,x,\sigma(\eta_{1\dots n(-1)\dots(-n)})) = 0$$

whenever $\sigma \in \mathcal{S}_{2n}$, because $G_N^\tau(t,x,\eta_{1\dots 2n}) \propto [k_{1\dots 2n}^N]$.

We decompose every G_N^τ into two parts:

$$\begin{aligned} G_N^\tau(t,x,\eta_\tau) &= G_N^\tau(t,x,\eta_\tau)\mathbb{1}_{|k_1|,\dots,|k_{d(\tau)}|\leqslant N/(2d(\tau))} + G_N^\tau(t,x,\eta_\tau)(1-\mathbb{1}_{|k_1|,\dots,|k_{d(\tau)}|\leqslant N/(2d(\tau))}) \\ &= K_N^\tau(t,x,\eta_\tau) + F_N^\tau(t,x,\eta_\tau). \end{aligned} \tag{94}$$

Let us first indicate how to deal with F_N^τ, which gives a vanishing contribution in the limit. Define

$$q_N(\theta) = \mathbb{1}_{|k|\leqslant N/2} \frac{ikg_\varepsilon(k)}{i\omega + f_\varepsilon(k)k^2},$$

which satisfies $|q_N(\theta)| \lesssim |\theta|^{-1}$, uniformly in N (recall that we defined $|\theta| = |\omega|^{1/2} + |k|$). Similarly as in Section 9.5 we can bound every integrand that we need to control by a product of terms of the form $q_N(\theta_{[\tau']}^N)$, discrete vertex functions, and factors like $\psi_\circ(k_{[\tau']}^N, k_{[\tau'']}^N)$, where $\theta^N = (\omega, k^N)$. Moreover, every integrand contains a factor $|q_N(\theta_i)|$ for each its integration variables θ_i, and we can decompose $1 - \mathbb{1}_{|k_1|,\dots,|k_{d(\tau)}|\leqslant N/(2d(\tau))}$ into a finite sum of terms that each contain a factor $\mathbb{1}_{|\theta_i|>N/(2d(\tau))}$ for some i. We can therefore estimate $|q_N(\theta_i)|\mathbb{1}_{|\theta_i|>N/(2d(\tau))} \lesssim N^{-\delta}|\theta_i|^{1-\delta}$ for an arbitrarily small $\delta > 0$, which gives us a small factor. We now only have to show that the remaining integral stays bounded. To do so, we need to control the basic integral

$$\int_{\mathbb{R}\times E_N} |\theta|^{-\alpha} |(\theta'-\theta)^N|^{-\beta} \mathrm{d}\theta \mathbb{1}_{|k'|\leqslant N/2}.$$

But $(k'-k)^N = k' - k + jN$ for some $|j| \leqslant 1$ and if $|k'| \leqslant N/2$, then $|k' + jN| \geqslant |k'|$ for all $j \in \mathbb{Z}$. So by Lemma 9.8 we have

$$\begin{aligned} \int_{\mathbb{R}\times E_N} |\theta|^{-\alpha} |(\theta'-\theta)^N|^{-\beta} \mathrm{d}\theta \mathbb{1}_{|k'|\leqslant N/2} &\leqslant \sum_{|j|\leqslant 1} \int_{\mathbb{R}\times E_N} |\theta|^{-\alpha} |\theta' + (0,jN) - \theta|^{-\beta} \mathrm{d}\theta \mathbb{1}_{|k'|\leqslant N/2} \\ &\lesssim \sum_{|j|\leqslant 1} |\theta' + (0,jN)|^{-\rho} \mathbb{1}_{|k'|\leqslant N/2} \lesssim |\theta'|^{-\rho} \mathbb{1}_{|k'|\leqslant N/2}, \end{aligned}$$

where α, β, ρ are as in the announcement of the lemma. The second integral in Lemma 9.8 is estimated using the same argument, which also allows us to show that all discrete vertex functions apart from $V_N^{\triangledown}$ and $V_N^{\diamond}$ satisfy the same bounds as their continuous counterparts. Since we estimated $V_N^{\triangledown}$ and $V_N^{\diamond}$ in Lemma 10.3, we now simply need to repeat the calculations of Section 9.5 to show that the Wiener-Itô integral over F_N^τ converges to 0 in L^p in the appropriate Besov space.

We still have to treat the term K_N^τ in the decomposition (94) of G_N^τ. In the description of K_N^τ we can now replace every $k_{[\tau']}^N$ by the usual sum $k_{[\tau']}$ (where τ' is an arbitrary subtree of τ). Then K_N^τ satisfies the same bounds as G^τ, uniformly in N, and converges pointwise to it. So whenever G^τ is absolutely integrable we can apply the dominated convergence theorem to conclude. However, in the continuous case we used certain symmetries to derive the bounds for $Q \circ X$, $V^{\triangledown}$, and $V^{\diamond}$, and these symmetries turn out to be violated in the discrete case. This is why we separately studied the convergence of $B_N(Q_N \circ X_N)$ in Lemma 10.2, and in Lemma 10.3 we showed that $V_N^{\triangledown}$ and $V_N^{\diamond}$ satisfy the same bounds as $V^{\triangledown}$ and $V^{\diamond}$, uniformly in N, and converge to $V^{\triangledown} + 2\pi c\delta(\sigma)$ and $V^{\diamond} + 2\pi c\delta(\sigma)$ respectively. It thus remains to see which correction terms we pick up from the additional Dirac deltas.

Let us start with the contraction of G_N^{Ψ}. Here we have

$$X_N^{\Psi}(t,x) = \int_{(\mathbb{R}\times E)^3} G_N^{\Psi}(t,x,\eta_{123}) W(\mathrm{d}\eta_{123}) + \int_{\mathbb{R}\times E} G_{N,1}^{\Psi}(t,x,\eta_1) W(\mathrm{d}\eta_1)$$

with

$$G_{N,1}^{\Psi}(t,x,\eta_1) = \pi^{-1} G_N^{\triangledown}(t,x,\eta_1) = \pi^{-1} e^{ik_1 x} \int_0^t \mathrm{d}\sigma \int_0^\sigma \mathrm{d}\sigma' H_{t-\sigma}^N(k_1) H_{\sigma'-s_1}^N(k_1) V_N^{\triangledown}(\sigma-\sigma',k_1)$$

for $V_N^{\triangledown}$ as defined in (92). Now by Lemma 10.3, the right hand side converges to

$$\pi^{-1} G^{\triangledown}(t,x,\eta_1) + 2ce^{ik_1 x} \int_0^t \mathrm{d}\sigma H_{t-\sigma}(k_1) H_{\sigma-s_1}(k_1),$$

and we have

$$\int_{\mathbb{R}\times E} \Big(2ce^{ik_1 x} \int_0^t \mathrm{d}\sigma H_{t-\sigma}(k_1) H_{\sigma-s_1}(k_1) \Big) W(\mathrm{d}\eta_1) = 2cQ(t,x).$$

In conclusion, X_N^{Ψ} converges to $X^{\Psi} + 2cQ$ in $L^p(\Omega, \mathscr{L}_T^\alpha)$.

The only other place where $V_N^{\triangledown}$ appears is in the contractions of $X_N^{\Psi\!\circ}$. We have

$$X_N^{\Psi\!\circ}(t,x) = \int_{(\mathbb{R}\times E)^4} G_N^{\Psi\!\circ}(t,x,\eta_{1234}) W(\mathrm{d}\eta_{1234}) + \int_{(\mathbb{R}\times E)^2} G_{N,2}^{\Psi\!\circ}(t,x,\eta_{12}) W(\mathrm{d}\eta_{12})$$

with

$$G_{N,2}^{\Psi\!\circ}(t,x,\eta_{12}) = (2\pi)^{-1} (G_N^{\diamond}(t,x,\eta_{12}) + 2G_N^{\diamond'}(t,x,\eta_{12}) + 2G_N^{\triangledown'}(t,x,\eta_{12})).$$

Now $G_N^{\curlyvee}(t,x,\eta_{12})$ can be factorized as

$$G_N^{\curlyvee}(t,x,\eta_{12}) = e^{ik_{[12]}x}h_\varepsilon(k_1,k_2)\int_0^t \mathrm{d}\sigma \int_0^\sigma \mathrm{d}\sigma' \int_0^{\sigma'} \mathrm{d}\sigma'' H^N_{t-\sigma}(k_{[12]})H^N_{\sigma'-\sigma''}(k_{[12]}) \\ \times H^N_{\sigma''-s_1}(k_1)H^N_{\sigma''-s_2}(k_2)V_N^{\curlyvee}(\sigma-\sigma',k_{[12]}),$$

and Lemma 10.3 shows that the right hand side converges to

$$G^{\curlyvee}(t,x,\eta_{12}) + 2\pi c e^{ik_{[12]}x}\int_0^t \mathrm{d}\sigma \int_0^\sigma \mathrm{d}\sigma'' H_{t-\sigma}(k_{[12]})H_{\sigma-\sigma''}(k_{[12]})H_{\sigma''-s_1}(k_1)H_{\sigma''-s_2}(k_2),$$

with

$$\int_{(\mathbb{R}\times E)^2}\Big(e^{ik_{[12]}x}\int_0^t \mathrm{d}\sigma \int_0^\sigma \mathrm{d}\sigma'' H_{t-\sigma}(k_{[12]})H_{\sigma-\sigma''}(k_{[12]})H_{\sigma''-s_1}(k_1)H_{\sigma''-s_2}(k_2)\Big)W(\mathrm{d}\eta_{12}) \\ = Q^{\vee}(t,x).$$

Similarly, we factorize $G_N^{\curlyvee}(t,x,\eta_{12})$ as

$$G_N^{\curlyvee}(t,x,\eta_{12}) = e^{ik_{[12]}x}h_\varepsilon(k_1,k_2)\psi_\circ(k_1,k_2)\int_0^t \mathrm{d}\sigma \int_0^\sigma \mathrm{d}\sigma' \int_0^{\sigma'} \mathrm{d}\sigma'' H^N_{t-\sigma}(k_{[12]})H^N_{\sigma-s_2}(k_2) \\ \times H^N_{\sigma-\sigma'}(k_1)H^N_{\sigma''-s_1}(k_1)V_N^{\curlyvee}(\sigma'-\sigma'',k_1),$$

and Lemma 10.3 shows that the right hand side converges to

$$G^{\curlyvee}(t,x,\eta_{12})+2\pi c e^{ik_{[12]}x}\psi_\circ(k_1,k_2)\int_0^t \mathrm{d}\sigma \int_0^\sigma \mathrm{d}\sigma' H_{t-\sigma}(k_{[12]})H_{\sigma-s_2}(k_2)H_{\sigma-\sigma'}(k_1)H_{\sigma'-s_1}(k_1)$$

with

$$\int_{(\mathbb{R}\times E)^2}\Big(e^{ik_{[12]}x}\psi_\circ(k_1,k_2)\int_0^t \mathrm{d}\sigma \int_0^\sigma \mathrm{d}\sigma' H_{t-\sigma}(k_{[12]})H_{\sigma-s_2}(k_2)H_{\sigma-\sigma'}(k_1)H_{\sigma'-s_1}(k_1)\Big)W(\mathrm{d}\eta_{12}) \\ = Q^{Q\circ X}(t,x).$$

In conclusion, $X_N^{\curlyvee}$ converges to $X^{\curlyvee} + cQ^{\vee} + 2cQ^{Q\circ X}$ in $L^p(\Omega,\mathscr{L}_T^{2\alpha})$. $\square$

10.2 Convergence of the random operator

The purpose of this subsection is to prove that the random operator

$$A_N(f) := \Pi_N\Big(\int(\Pi_N(f \prec \tau_{-\varepsilon y}Q_N)\circ(\tau_{-\varepsilon z}X_N)\mu(\mathrm{d}y,\mathrm{d}z)\Big) \\ - \mathcal{P}_N\Big(\int((f\prec\tau_{-\varepsilon y}Q_N))\circ(\tau_{-\varepsilon z}X_N)\mu(\mathrm{d}y,\mathrm{d}z)\Big) \tag{95}$$

converges to zero in probability.

Theorem 10.4. *Let $\alpha \in (1/4, 1/2)$. Then we have for all $\delta, T > 0$ and $r \geqslant 1$*

$$\mathbb{E}[\|A_N\|^r_{C_T L(\mathscr{C}^\alpha, \mathscr{C}^{2\alpha-1})}]^{1/r} \lesssim N^{\alpha - 1/2 + \delta}. \tag{96}$$

We will prove the theorem by building on three auxiliary lemmas.

Lemma 10.5. *The operator A_N is given by*

$$A_N(f)(t,x) = \sum_{q,p} \Delta_q A_N(\Delta_p f)(t,x) = \sum_{q,p} \int_{\mathbb{T}} g^N_{p,q}(t,x,y) \Delta_p f(y) \mathrm{d}y$$

with

$$\mathscr{F} g^N_{p,q}(t,x,\cdot)(k) = \sum_{k_1,k_2} \Gamma^N_{p,q}(x;k,k_1,k_2) \mathscr{F} Q_N(t,k_1) \mathscr{F} X_N(t,k_2), \tag{97}$$

where

$$\begin{aligned}\Gamma^N_{p,q}(x;k,k_1,k_2) = (2\pi)^{-2} \tilde{\rho}_p(k) \psi_\prec(k,k_1) h_\varepsilon(k_1,k_2) \\ \times [e^{i(k_{[12]}-k)^N x} \rho_q((k_{[12]}-k)^N) \psi_\circ((k_1-k)^N, k_2) \\ - e^{i(k_{[12]}-k)x} \rho_q(k_{[12]}-k) \psi_\circ(k_1-k, k_2) \mathbb{1}_{|k_{[12]}-k| \leqslant N/2}]\end{aligned}$$

and where $\tilde{\rho}_p$ is a smooth function supported in an annulus $2^p \mathscr{A}$ such that $\tilde{\rho}_p \rho_p = \rho_p$.

Proof. Parseval's formula gives

$$\begin{aligned}\Delta_q A_N(\Delta_p f)(t,x) &= \int_{\mathbb{T}} g^N_{p,q}(t,x,y) \Delta_p f(y) \mathrm{d}y \\ &= (2\pi)^{-1} \sum_k \mathscr{F} g^N_{p,q}(t,x,\cdot)(-k) \rho_p(k) \mathscr{F} f(k).\end{aligned}$$

It suffices to verify that the same identity holds if we replace $\mathscr{F} g^N_{p,q}(t,x,\cdot)(k)$ with the right hand side of (97) and if we fix some k and consider $f(x) = (2\pi)^{-1} e^{ikx}$, i.e. $\mathscr{F} f(\ell) = \delta_{\ell,k}$. Let us look at the first term on the right hand side of (95)

$$\begin{aligned}&\int \Delta_q \Pi_N[(\Pi_N(\Delta_p((2\pi)^{-1} e^{ik\cdot}) \prec \tau_{-\varepsilon z_1}((2\pi)^{-1} e^{ik_1 \cdot}))) \circ (\tau_{-\varepsilon z_2}((2\pi)^{-1} e^{ik_2 \cdot}))] \mu(\mathrm{d}z_1, \mathrm{d}z_2) \\ &\quad = (2\pi)^{-3} \Delta_q \Pi_N[(\Pi_N(\rho_p(k) \psi_\prec(k,k_1) e^{i(k+k_1)\cdot})) \circ (e^{ik_2 \cdot})] \int e^{ik_1 \varepsilon z_1 + ik_2 \varepsilon z_2} \mu(\mathrm{d}z_1, \mathrm{d}z_2) \\ &\quad = (2\pi)^{-3} \Delta_q \Pi_N[\psi_\circ((k+k_1)^N, k_2) \rho_p(k) \psi_\prec(k,k_1) e^{i((k+k_1)^N + k_2)\cdot}] h_\varepsilon(k_1,k_2) \\ &\quad = (2\pi)^{-3} e^{i(k_{[12]}+k)^N \cdot} \rho_q((k_{[12]}+k)^N) \psi_\circ((k+k_1)^N, k_2) \tilde{\rho}_p(k) \rho_p(k) \psi_\prec(k,k_1) h_\varepsilon(k_1,k_2),\end{aligned}$$

where in the last step we used that $((k+k_1)^N + k_2)^N = (k+k_1+k_2)^N$. The second term in (95) is treated with the same arguments, and writing Q_N and X_N as inverse Fourier transforms equation (97) follows. □

Lemma 10.6. *For all $r \geqslant 1$ and all $\alpha, \beta \in \mathbb{R}$ we can estimate*

$$\mathbb{E}[\|A_N(t) - A_N(s)\|^r_{L(\mathscr{C}^\alpha, B^\beta_{r,r})}] \lesssim \sum_{q,p} 2^{qr\beta} 2^{-pr\alpha} \Big(\sup_{x \in \mathbb{T}} \sum_k \mathbb{E}[|(\mathscr{F} g^N_{p,q}(t,x,\cdot) - \mathscr{F} g^N_{p,q}(s,x,\cdot))(k)|^2] \Big)^{r/2}.$$

Proof. From the decomposition in Lemma 10.5 we get

$$\begin{aligned}\|A_N(f)(t) - A_N(f)(s)\|^r_{B^\beta_{r,r}} &= \sum_q 2^{qr\beta} \Big\| \sum_p \int_{\mathbb{T}} (g^N_{p,q}(t,x,y) - g^N_{p,q}(s,x,y)) \Delta_p f(y) \mathrm{d}y \Big\|^r_{L^r_x(\mathbb{T})} \\ &\leqslant \sum_{q,p} 2^{qr\beta} 2^{-pr\alpha} \Big\| \int_{\mathbb{T}} |g^N_{p,q}(t,x,y) - g^N_{p,q}(s,x,y)| \mathrm{d}y \Big\|^r_{L^r_x(\mathbb{T})} \|f\|^r_\alpha,\end{aligned}$$

so that

$$\begin{aligned}\mathbb{E}[\|A_N(t) - A_N(s)\|^r_{L(\mathscr{C}^\alpha, B^\beta_{r,r})}] &\lesssim \sum_{q,p} 2^{qr\beta} 2^{-pr\alpha} \mathbb{E}\Big[\Big\| \int_{\mathbb{T}} |g^N_{p,q}(t,x,y) - g^N_{p,q}(s,x,y)| \mathrm{d}y \Big\|^r_{L^r_x(\mathbb{T})}\Big] \\ &\lesssim \sum_{q,p} 2^{qr\beta} 2^{-pr\alpha} \mathbb{E}\Big[\Big\| \Big(\int_{\mathbb{T}} |g^N_{p,q}(t,x,y) - g^N_{p,q}(s,x,y)|^2 \mathrm{d}y \Big)^{1/2} \Big\|^r_{L^r_x(\mathbb{T})}\Big] \\ &\lesssim \sum_{q,p} 2^{qr\beta} 2^{-pr\alpha} \int_{\mathbb{T}} \mathbb{E}\Big[\Big(\sum_k |(\mathscr{F} g^N_{p,q}(t,x,\cdot) - \mathscr{F} g^N_{p,q}(s,x,\cdot))(k)|^2 \Big)^{r/2} \Big] \mathrm{d}x,\end{aligned}$$

where we applied Parseval's formula. Now $\sum_k |(\mathscr{F} g^N_{p,q}(t,x,\cdot) - \mathscr{F} g^N_{p,q}(s,x,\cdot))(k)|^2$ is a random variable in the second inhomogeneous chaos generated by ξ, and therefore our claim follows from Nelson's estimate. □

Lemma 10.7. *For all $p, q \geqslant -1$, all $0 \leqslant t_1 < t_2$, and all $\lambda \in [0,1]$ we have*

$$\sum_k \mathbb{E}[|\mathscr{F} g_{p,q}(t_2,x,\cdot)(k) - \mathscr{F} g_{p,q}(t_1,x,\cdot)(k)|^2] \lesssim \mathbb{1}_{2^p, 2^q \lesssim N} 2^{p(1-\lambda)+q} N^{-1+3\lambda} |t_2 - t_1|^\lambda. \quad (98)$$

Proof. We write

$$\begin{aligned}&\mathscr{F} Q_N(t_2,k_1) \mathscr{F} X_N(t_2,k_2) - \mathscr{F} Q_N(t_1,k_1) \mathscr{F} X_N(t_1,k_2) \\ &\quad = (\mathscr{F} Q_N(t_2,k_1) - \mathscr{F} Q_N(t_1,k_1)) \mathscr{F} X_N(t_2,k_2) \\ &\qquad + \mathscr{F} Q_N(t_1,k_1)(\mathscr{F} X_N(t_2,k_2) - \mathscr{F} X_N(t_1,k_2))\end{aligned}$$

and estimate both terms on the right hand side separately. We only concentrate on the second one, which is slightly more difficult to treat, the first one being accessible to essentially the same arguments. We have the following chaos decomposition:

$\mathscr{F} Q_N(t_1,k_1)(\mathscr{F} X_N(t_2,k_2) - \mathscr{F} X_N(t_1,k_2))$

$$= \int_{(\mathbb{R}\times E_N)^2} \Big((2\pi)^2 \delta_{\ell_1,k_1}\delta_{\ell_2,k_2} \int_0^{t_1} H^N_{t_1-r}(\ell_1) H^N_{r-s_1}(\ell_1) \mathrm{d}r (H^N_{t_2-s_2}(\ell_2) - H^N_{t_1-s_2}(\ell_2)) \Big) W(\mathrm{d}\eta_{12})$$
$$+ 2\pi \int_{\mathbb{R}\times E_N} \Big(\delta_{\ell_1,k_1}\delta_{-\ell_1,k_2} \int_0^{t_1} H^N_{t_1-r}(\ell_1) H^N_{r-s_1}(\ell_1) \mathrm{d}r (H^N_{t_2-s_1}(-\ell_1) - H^N_{t_1-s_1}(-\ell_1)) \Big) \mathrm{d}\eta_1.$$

Using the same arguments as in Section 9.5, we can estimate the second term on the right hand side by

$$\Big| \int_{\mathbb{R}\times E_N} \Big(\delta_{\ell_1,k_1}\delta_{-\ell_1,k_2} \int_0^{t_1} H^N_{t_1-r}(\ell_1) H^N_{r-s_1}(\ell_1) \mathrm{d}r (H^N_{t_2-s_1}(-\ell_1) - H^N_{t_1-s_1}(-\ell_1)) \Big) \mathrm{d}\eta_1 \Big|$$
$$\lesssim \frac{\delta_{k_{-1},k_2}}{|k_1|} |t_2 - t_1|^{\lambda/2} |k_1|^{\lambda}$$

for any $\lambda \in [0, 2]$. Based on the decomposition of Lemma 10.5, the contribution from the chaos of order 0 is thus bounded by

$$\sum_k \Big| \sum_{k_1,k_2\in E_N} \Gamma^N_{p,q}(x; k, k_1, k_2) \mathbb{E}[\mathscr{F}Q_N(t_1, k_1)(\mathscr{F}X_N(t_2, k_2) - \mathscr{F}X_N(t_1, k_2))] \Big|^2 \tag{99}$$
$$\lesssim \sum_k \Big| \sum_{k_1\in E_N} \Gamma^N_{p,q}(x; k, k_1, k_{-1}) \frac{|t_2 - t_1|^{\lambda/2}}{|k_1|^{1-\lambda}} \Big|^2$$
$$\lesssim \sum_k \Big| \rho_q(k)\tilde{\rho}_p(k) \sum_{k_1\in E_N} \psi_{\prec}(k, k_1)(\psi_\circ((k_1 - k)^N, k_1) - \psi_\circ(k_1 - k, k_1)) \frac{|t_2 - t_1|^{\lambda/2}}{|k_1|^{1-\lambda}} \Big|^2,$$

where we used that $|k| < N/2$ on the support of $\psi_{\prec}(\cdot, k_1)$. Now $\psi_\circ((k_1 - k)^N, k_1) = \psi_\circ(k_1 - k, k_1)$ unless $|k_1 - k| > N/2$ and $|k_1| \simeq N$, and there are at most $|k|$ values of k_1 with $|k| < |k_1| < N/2$ and $|k_1 - k| > N/2$. Therefore, the right hand side of (99) is bounded by

$$\lesssim |t_2 - t_1|^{\lambda} \sum_k \mathbb{1}_{|k|<N/2} |\rho_q(k)\tilde{\rho}_p(k)|^2 |k|^2 N^{-2+2\lambda} \lesssim \mathbb{1}_{p\sim q} \mathbb{1}_{2^q \lesssim N} |t_2 - t_1|^{\lambda} 2^{3q} N^{-2+2\lambda}$$
$$\lesssim \mathbb{1}_{2^p,2^q\lesssim N} 2^{p(1-\lambda)+q} N^{-1+3\lambda} |t_2 - t_1|^{\lambda}.$$

Next consider the contribution from the second chaos, which is given by

$$(2\pi)^2 \int_{(\mathbb{R}\times E)^2} \Big(\Gamma^N_{p,q}(x; k, k_1, k_2) \int_0^{t_1} H^N_{t_1-r}(k_1) H^N_{r-s_1}(k_1) \mathrm{d}r$$
$$\times (H^N_{t_2-s_2}(k_2) - H^N_{t_1-s_2}(k_2)) \Big) W(\mathrm{d}\eta_{12}).$$

Taking the expectation of the norm squared, we obtain (up to a multiple of 2π)

$$\int_{(\mathbb{R}\times E_N)^2} \Big| \Gamma^N_{p,q}(x; k, k_1, k_2) \int_0^{t_1} H^N_{t_1-r}(k_1) H^N_{r-s_1}(k_1) \mathrm{d}r (H^N_{t_2-s_2}(k_2) - H^N_{t_1-s_2}(k_2)) \Big|^2 \mathrm{d}\eta_{12}. \tag{100}$$

As in Section 9.5 we can show that

$$\int_{\mathbb{R}^2} \Big| \int_0^{t_1} H^N_{t_1-r}(k_1) H^N_{r-s_1}(k_1) \mathrm{d}r (H^N_{t_2-s_2}(k_2) - H^N_{t_1-s_2}(k_2)) \Big|^2 \mathrm{d}s_{12} \lesssim \frac{|k_2|^{2\lambda}|t_2-t_1|^{\lambda}}{|k_1|^2}$$

whenever $\lambda \in [0,2]$, and plugging this into (100) we get

$$\begin{aligned}
(100) &\lesssim \int_{E_N^2} |\Gamma^N_{p,q}(x;k,k_1,k_2)|^2 \frac{|k_2|^{2\lambda}|t_2-t_1|^{\lambda}}{|k_1|^2} \mathrm{d}k_{12} \\
&\lesssim N^{2\lambda}|t_2-t_1|^{\lambda} \int_{E_N^2} \mathrm{d}k_{12} \tilde{\rho}_p^2(k) \psi_{\prec}^2(k,k_1)|k_1|^{-2} \\
&\qquad \times \Big| e^{i(k_{[12]}-k)^N x} \rho_q((k_{[12]}-k)^N) \psi_{\circ}((k_1-k)^N, k_2) \\
&\qquad\qquad - e^{i(k_{[12]}-k)x} \rho_q(k_{[12]}-k) \psi_{\circ}(k_1-k,k_2) \mathbb{1}_{|k_{[12]}-k| \leqslant N/2} \Big|^2 .
\end{aligned}$$

First observe that the difference on the right hand side is zero unless $|k_1| \simeq N$, so that we can estimate $|k_1|^{-2} \lesssim N^{-1+\lambda}|k|^{-1-\lambda}$, and also we only have to sum over $|k_1| > |k|$. Moreover, the summation over k_2 gives $O(2^q)$ terms which leads to

$$(100) \lesssim \mathbb{1}_{2^p,2^q \lesssim N} 2^q N^{-1+3\lambda} |t_2-t_1|^{\lambda} \tilde{\rho}_p^2(k) |k|^{-\lambda},$$

and the sum over k of the right hand side is bounded by

$$\lesssim \mathbb{1}_{2^p,2^q \lesssim N} 2^{p(1-\lambda)+q} N^{-1+3\lambda} |t_2-t_1|^{\lambda},$$

which completes the proof of (98). □

Proof of Theorem 10.4. Let $\alpha \in (1/4,1/2)$, $r \geqslant 1$, and write $\gamma = 2\alpha - 1 + 1/r$. Combining Lemma 10.6 and Lemma 10.7, we get for all $\lambda \geqslant 0$ with $(1-\lambda)/2 > \alpha$

$$\begin{aligned}
\mathbb{E}[\|A_N(t) - A_N(s)\|^r_{L(\mathscr{C}^{\alpha},\mathscr{C}^{2\alpha-1})}]^{1/r} &\lesssim \mathbb{E}[\|A_N(f)(t) - A_N(f)(s)\|^r_{L(\mathscr{C}^{\beta},B^{\gamma}_{r,r})}]^{1/r} \\
&\lesssim \Big(\sum_{q,p \lesssim \log_2 N} 2^{qr\gamma} 2^{-pr\alpha} (2^{p(1-\lambda)+q} N^{-1+3\lambda} |t-s|^{\lambda})^{r/2} \Big)^{1/r} \\
&\simeq |t-s|^{\lambda/2} N^{\alpha-1/2+1/r+3\lambda/2}
\end{aligned}$$

Since $A_N(0) = 0$, Kolmogorov's continuity criterion gives

$$\mathbb{E}[\|A_N(f)\|^r_{C_T L(\mathscr{C}^{\alpha},\mathscr{C}^{2\alpha-1})}]^{1/r} \lesssim N^{\alpha-1/2+1/r+3\lambda/2}$$

whenever $\lambda/2 > 1/r$. Choosing $1/r + 3\lambda/2 \leqslant \delta$, the estimate (96) follows. □

A Proofs of some auxiliary results

Proof of Lemma 6.4. We prove the second claim, the first one follows from the same arguments. We also assume that $f \in \mathcal{M}_t^\gamma L^p$, the case $f \in \mathcal{M}_t^\gamma L^\infty$ is analogous. Let $t \geqslant 0$ and $j \geqslant 0$ and write $M = \|f\|_{\mathcal{M}_t^\gamma L^p} \|g(t)\|_\beta$. Then

$$\Big\| \int_0^t 2^{2j} \varphi(2^{2j}(t-s)) S_{j-1} f(s \vee 0) \mathrm{d}s \Delta_j g(t) \Big\|_{L^p} \leqslant 2^{-j\beta} M \int_0^t 2^{2j} |\varphi(2^{2j}(t-s))| s^{-\gamma} \mathrm{d}s.$$

We now split the integral at $t/2$. The Schwartz function φ satisfies $|\varphi(r)| \lesssim r^{-1}$, so using the fact that $\gamma \in [0,1)$ we get

$$\int_0^{t/2} 2^{2j} |\varphi(2^{2j}(t-s))| s^{-\gamma} \mathrm{d}s \lesssim \int_0^{t/2} (t-s)^{-1} s^{-\gamma} \mathrm{d}s \lesssim t^{-1} \int_0^{t/2} s^{-\gamma} \mathrm{d}s \lesssim t^{-\gamma}.$$

The remaining part of the integral can be simply estimated by

$$\int_{t/2}^t 2^{2j} |\varphi(2^{2j}(t-s))| s^{-\gamma} \mathrm{d}s \leqslant (t/2)^{-\gamma} \int_{\mathbb{R}} 2^{2j} |\varphi(2^{2j}(t-s))| \mathrm{d}s \lesssim t^{-\gamma},$$

which concludes the proof. □

Proof of Lemma 6.5. Let us begin with the first claim. By spectral support properties it suffices to control

$$\begin{aligned} &\Big\| \Big(\int_0^t 2^{2j} \varphi(2^{2j}(t-s)) S_{j-1} f(s) \mathrm{d}s - S_{j-1} f(t) \Big) \Delta_j g(t) \Big\|_{L^p} \\ &\qquad \leqslant \Big\| \int_{-\infty}^0 2^{2j} \varphi(2^{2j}(t-s)) S_{j-1} f(t) \mathrm{d}s \Delta_j g(t) \Big\|_{L^p} \\ &\qquad\quad + \Big\| \int_0^t 2^{2j} \varphi(2^{2j}(t-s)) S_{j-1} f_{t,s} \mathrm{d}s \Delta_j g(t) \Big\|_{L^p}, \end{aligned} \tag{101}$$

where we used that φ has total mass 1. Using that $|\varphi(r)| \lesssim |r|^{-1-\alpha/2}$, we can estimate the first term on the right hand side by

$$\lesssim 2^{-j\beta} t^{-\gamma+\alpha/2} \|f\|_{\mathscr{L}_p^{\gamma,\alpha}(t)} \|g(t)\|_\beta \int_{2^{2j}t}^\infty |r|^{-1-\alpha/2} \mathrm{d}r \lesssim 2^{-j(\alpha+\beta)} t^{-\gamma} \|f\|_{\mathscr{L}_p^{\gamma,\alpha}(t)} \|g(t)\|_\beta.$$

As for the second term in (101), we split the domain of integration into the intervals $[0, t/2]$ and $[t/2, t]$. On the first interval we use again that $|\varphi(r)| \lesssim |r|^{-1-\alpha/2}$ and then simply estimate $\|f_{t,s}\|_{L^p} \leqslant (t^{\alpha/2-\gamma} + s^{\alpha/2-\gamma}) \|f\|_{\mathscr{L}_p^{\gamma,\alpha}(t)}$ and $(t-s)^{-1-\alpha/2} \lesssim t^{-1-\alpha/2}$, and the required bound follows. On the second interval, an application of the triangle inequality shows that

$$\|f_{t,s}\|_{L^p} = \|s^{-\gamma}(s^\gamma f(s)) - t^{-\gamma}(t^\gamma f(t))\|_{L^p} \lesssim t^{-\gamma} |t-s|^{\alpha/2} \|f\|_{\mathscr{L}_p^{\gamma,\alpha}(t)},$$

from where we easily deduce the claimed estimate.

Let us now get to the bound $t^\gamma \left\|\left(\mathscr{L}(f \prec\!\!\prec g) - f \prec\!\!\prec (\mathscr{L}g)\right)(t)\right\|_{\mathscr{C}_p^{\alpha+\beta-2}} \lesssim \|f\|_{\mathscr{L}_p^{\gamma,\alpha}(t)} \|g(t)\|_\beta$. Given Lemma 6.4 and the product rule for the Laplacian, only the estimate for

$$\sum_j \partial_t \left(\int_0^t 2^{2j} \varphi(2^{2j}(t-s)) S_{j-1} u(s) \mathrm{d}s \right) \Delta_j v(t).$$

is non-trivial. For fixed j we recall that $\mathrm{supp}(\varphi) \subset \mathbb{R}_+$, and therefore

$$\int_0^t 2^{4j} \varphi'(2^{2j}(t-s)) S_{j-1} u(s) \mathrm{d}s = 2^{2j} \int_{\mathbb{R}} 2^{2j} \varphi'(2^{2j}(t-s)) S_{j-1} u(s) \mathbb{1}_{s \geqslant 0} \mathrm{d}s.$$

Since φ' integrates to zero we can subtract $2^{2j} \int_{\mathbb{R}} 2^{2j} \varphi'(2^{2j}(t-s)) S_{j-1} u(t) \mathrm{d}s$. The result then follows as in the first part of the proof. □

Proof of Lemma 6.6. We start by observing that exactly the same arguments as in the proof of Lemma 2.9 (replacing L^∞ by L^p at the appropriate places) yield

$$\|If\|_{\mathscr{L}_p^{0,\alpha}(T)} \lesssim \|f\|_{C_T \mathscr{C}_p^{\alpha-2}}, \qquad \|s \mapsto P_s u_0\|_{\mathscr{L}_p^{0,\alpha}(T)} \lesssim \|u_0\|_{\mathscr{C}_p^\alpha}.$$

It remains to include the possible blow up at 0. The estimate (50),

$$\|If\|_{\mathcal{M}_T^\gamma \mathscr{C}_p^\alpha} \lesssim \|f\|_{\mathcal{M}_T^\gamma \mathscr{C}_p^{\alpha-2}} \tag{102}$$

is shown in Lemma A.9 of [GIP15] (for $p = \infty$, but the extension to general p is again trivial). For the temporal Hölder regularity of If that we need in (49), we note that the estimate

$$\|P_t u\|_{\mathscr{C}_p^\alpha} \lesssim t^{-(\alpha+\beta)/2} \|u\|_{\mathscr{C}_p^{-\beta}}$$

for $\beta \geqslant -\alpha$ is again a simple extension from $p = \infty$ to general p, this time of Lemma A.7 in [GIP15]. An interpolation argument then yields

$$\|P_t u\|_{L^p} \lesssim t^{-\beta/2} \|u\|_{\mathscr{C}_p^{-\beta}} \tag{103}$$

whenever $\beta < 0$, and from here we obtain

$$\begin{aligned} \|(P_t - \mathrm{id})u\|_{L^p} &= \Big\| \int_0^t \partial_s P_s u \mathrm{d}s \Big\|_{L^p} = \Big\| \int_0^t P_s \Delta u \mathrm{d}s \Big\|_{L^p} \\ &\lesssim \int_0^t s^{-1+\alpha/2} \|u\|_{\mathscr{C}_p^\alpha} \mathrm{d}s \lesssim t^{\alpha/2} \|u\|_{\mathscr{C}_p^\alpha} \end{aligned} \tag{104}$$

for $\alpha \in (0,2)$. To estimate the temporal regularity, we now have to control

$$\begin{aligned} &\Big\| t^\gamma \int_0^t P_{t-r} f_r \mathrm{d}r - s^\gamma \int_0^s P_{s-r} f_r \mathrm{d}r \Big\|_{L^p} \\ &\qquad \leqslant (t^\gamma - s^\gamma) \int_0^t \|P_{t-r} f_r\|_{L^p} \mathrm{d}r + s^\gamma \int_s^t \|P_{t-r} f_r\|_{L^p} \mathrm{d}r + s^\gamma \Big\| (P_{t-s} - \mathrm{id}) If(s) \Big\|_{L^p} \end{aligned}$$

$$\lesssim \Big((t^\gamma - s^\gamma)\int_0^t (t-r)^{\alpha/2-1} r^{-\gamma}\mathrm{d}r + s^\gamma \int_s^t (t-r)^{\alpha/2-1} r^{-\gamma}\mathrm{d}r\Big)\|f\|_{\mathcal{M}_T^\gamma \mathscr{C}_p^{\alpha-2}}$$
$$+ s^\gamma |t-s|^{\alpha/2}\|If(s)\|_{\mathscr{C}_p^\alpha},$$

where we used (103) for the first two terms and (104) for the last term. Now (102) allows us to further simplify this to

$$\lesssim \Big((t^\gamma - s^\gamma)t^{\alpha/2-\gamma}\int_0^1 (1-r)^{\alpha/2-1} r^{-\gamma}\mathrm{d}r + \int_s^t (t-r)^{\alpha/2-1}\mathrm{d}r + |t-s|^{\alpha/2}\Big)\|f\|_{\mathcal{M}_T^\gamma \mathscr{C}_p^{\alpha-2}}.$$

Since $\alpha > 0$ and $\gamma < 1$, the first time integral is finite. Moreover it is not hard to see that $(t^\gamma - s^\gamma)t^{\alpha/2-\gamma} \lesssim |t-s|^{\alpha/2}$ (distinguish for example the cases $s \leqslant t/2$ and $s \in (t/2, t]$), and eventually we obtain

$$\|t^\gamma If(t) - s^\gamma If(s)\|_{L^p} \lesssim |t-s|^{\alpha/2}\|f\|_{\mathcal{M}_T^\gamma \mathscr{C}_p^{\alpha-2}}.$$

It remains to control the temporal regularity of $s \mapsto s^\gamma P_s u_0$, and this can be done using similar (but simpler) arguments as above, so that the proof is complete. □

Proof of Lemma 6.8. For the spatial regularity, observe that

$$t^\gamma \|\Delta_j u(t)\|_{L^p} \leqslant \min\{2^{-j\alpha}, t^{\alpha/2}\}\|u\|_{\mathscr{L}_p^{\gamma,\alpha}(T)}.$$

Interpolating, we obtain

$$t^\gamma \|\Delta_j u(t)\|_{L^p} \leqslant 2^{-j(\alpha-\varepsilon)} t^{\varepsilon/2}\|u\|_{\mathscr{L}_p^{\gamma,\alpha}(T)},$$

or in other words $t^{\gamma-\varepsilon/2}\|u(t)\|_{\alpha-\varepsilon} \leqslant \|u\|_{\mathscr{L}_p^{\gamma,\alpha}(T)}$. The statement about the temporal regularity is a special case of the following lemma. □

Lemma A.1. *Let $\alpha \in (0,1)$, $\varepsilon \in [0,\alpha)$, and let $f\colon [0,\infty) \to X$ be an α–Hölder continuous function with values in a normed vector space X, such that $f(0) = 0$. Then*

$$\|t \mapsto t^{-\varepsilon} f(t)\|_{C_T^{\alpha-\varepsilon}X} \lesssim \|f\|_{C_T^\alpha X}.$$

Proof. Let $0 \leqslant s < t$. If $s = 0$, the required estimate easily follows from the fact that $f(0) = 0$, so assume $s > 0$. If $t > 2s$, then we use again that $f(0) = 0$ to obtain

$$\frac{\|t^{-\varepsilon} f(t) - s^{-\varepsilon} f(s)\|_X}{|t-s|^{\alpha-\varepsilon}} \leqslant (t^{\alpha-\varepsilon} + s^{\alpha-\varepsilon})|t-s|^{\varepsilon-\alpha}\|f\|_{C_T^\alpha X}.$$

Now $t^{\alpha-\varepsilon} \leqslant (|t-s| + s)^{\alpha-\varepsilon}$, and $s/|t-s| < 1$ by assumption, thus the result follows. If $s < t \leqslant 2s$, we apply a first order Taylor expansion to $t^{-\varepsilon} - s^{-\varepsilon}$ and obtain

$$\frac{\|t^{-\varepsilon} f(t) - s^{-\varepsilon} f(s)\|_X}{|t-s|^{\alpha-\varepsilon}} \leqslant \frac{t^{-\varepsilon}}{|t-s|^{-\varepsilon}}\frac{\|f(t) - f(s)\|_X}{|t-s|^\alpha} + \frac{\varepsilon r^{-\varepsilon-1}(t-s)\|f(s)\|_X}{|t-s|^{\alpha-\varepsilon}}$$

for some $r \in (s,t)$. Now clearly the first term on the right hand side is bounded by $\|f\|_{C_T^\alpha X}$. For the second term we use $\|f(s)\|_X \leqslant s^\alpha \|f\|_{C_T^\alpha X}$ and get

$$\frac{\varepsilon r^{-\varepsilon-1}(t-s)\|f(s)\|_X}{|t-s|^{\alpha-\varepsilon}} \leqslant \varepsilon s^{-\varepsilon-1+\alpha}|t-s|^{\varepsilon+1-\alpha}\|f\|_{C_T^\alpha X} \leqslant \varepsilon \|f\|_{C_T^\alpha X},$$

using $t - s \leqslant s$ in the last step. □

References

[AC15] Romain Allez and Khalik Chouk, *The continuous Anderson hamiltonian in dimension two*, arXiv preprint arXiv:1511.02718 (2015). ↑37

[ACQ11] Gideon Amir, Ivan Corwin, and Jeremy Quastel, *Probability distribution of the free energy of the continuum directed random polymer in* $1+1$ *dimensions*, Comm. Pure Appl. Math. **64** (2011), no. 4, 466–537. MR2796514 (2012b:60304) ↑3

[AKQ14] Tom Alberts, Konstantin Khanin, and Jeremy Quastel, *The intermediate disorder regime for directed polymers in dimension* $1+1$, Ann. Probab. **42** (2014), no. 3, 1212–1256. MR3189070 ↑4

[Ass02] Sigurd Assing, *A pregenerator for Burgers equation forced by conservative noise*, Comm. Math. Phys. **225** (2002), no. 3, 611–632. MR1888875 (2003b:60094) ↑35

[Ass13] ———, *A rigorous equation for the Cole-Hopf solution of the conservative KPZ equation*, Stoch. Partial Differ. Equ. Anal. Comput. **1** (2013), no. 2, 365–388. MR3327511 ↑35

[BB16] Ismaël Bailleul and Frederic Bernicot, *Heat semigroup and singular PDEs*, J. Funct. Anal. **270** (2016), no. 9, 3344–3452. ↑5

[BBF15] Ismaël Bailleul, Frederic Bernicot, and Dorothee Frey, *Higher order paracontrolled calculus and 3d-PAM equation*, arXiv preprint arXiv:1506.08773 (2015). ↑5

[BCD11] Hajer Bahouri, Jean-Yves Chemin, and Raphael Danchin, *Fourier analysis and nonlinear partial differential equations*, Springer, 2011. ↑8, 64, 85

[BCK14] Yuri Bakhtin, Eric Cator, and Konstantin Khanin, *Space-time stationary solutions for the Burgers equation*, J. Amer. Math. Soc. **27** (2014), no. 1, 193–238. MR3110798 ↑6

[BD98] Michelle Boué and Paul Dupuis, *A variational representation for certain functionals of Brownian motion*, Ann. Probab. **26** (1998), no. 4, 1641–1659. MR1675051 (99m:60124) ↑44

[BG97] Lorenzo Bertini and Giambattista Giacomin, *Stochastic Burgers and KPZ equations from particle systems*, Comm. Math. Phys. **183** (1997), no. 3, 571–607. MR1462228 (99e:60212) ↑4, 5, 8, 59

[Bon81] Jean-Michel Bony, *Calcul symbolique et propagation des singularites pour les équations aux dérivées partielles non linéaires*, Ann. Sci. École Norm. Supér. (4) **14** (1981), 209–246. ↑9, 12

[CC13] Rémi Catellier and Khalil Chouk, *Paracontrolled distributions and the 3-dimensional stochastic quantization equation*, arXiv preprint arXiv:1310.6869 (2013). ↑5

[CC15] Giuseppe Cannizzaro and Khalil Chouk, *Multidimensional SDEs with singular drift and universal construction of the polymer measure with white noise potential*, arXiv preprint arXiv:1501.04751 (2015). ↑5, 6, 48, 50, 51, 54

[CF14] Khalil Chouk and Peter Friz, *Support theorem for a singular semilinear stochastic partial differential equation*, arXiv preprint arXiv:1409.4250 (2014). ↑5

[CFG15] Giuseppe Cannizzaro, Peter Friz, and Paul Gassiat, *Malliavin Calculus for regularity structures: the case of gPAM*, arXiv preprint arXiv:1511.08888 (2015). ↑48, 51

[CK14] Le Chen and Kunwoo Kim, *On comparison principle and strict positivity of solutions to the nonlinear stochastic fractional heat equations*, arXiv preprint arXiv:1410.0604 (2014). ↑7, 48

[CO14] Pedro Catuogno and Christian Olivera, *Strong solution of the stochastic Burgers equation*, Appl. Anal. **93** (2014), no. 3, 646–652. MR3176977 ↑35

[Cor12] Ivan Corwin, *The Kardar-Parisi-Zhang equation and universality class*, Random Matrices Theory Appl. **1** (2012), no. 1, 1130001, 76. MR2930377 ↑3

[CT15] Ivan Corwin and Li-Cheng Tsai, *KPZ equation limit of higher-spin exclusion processes*, arXiv preprint arXiv:1505.04158 (2015). ↑4

[DD16] François Delarue and Roland Diel, *Rough paths and 1d SDE with a time dependent distributional drift: application to polymers*, Probab. Theory Related Fields **165** (2016), no. 1, 1–63. ↑6, 48

[DPZ14] Giuseppe Da Prato and Jerzy Zabczyk, *Stochastic equations in infinite dimensions*, Vol. 152, Cambridge University Press, 2014. ↑29

[DT13] Amir Dembo and Li-Cheng Tsai, *Weakly asymmetric non-simple exclusion process and the Kardar-Parisi-Zhang equation*, Comm. Math. Phys. **341** (2013), no. 1, 219–261. ↑4

[Ech82] Pedro Echeverría, *A criterion for invariant measures of Markov processes*, Z. Wahrsch. Verw. Gebiete **61** (1982), no. 1, 1–16. MR671239 (84a:60088) ↑59

[EKMS00] Weinan E, K. Khanin, A. Mazel, and Ya. Sinai, *Invariant measures for Burgers equation with stochastic forcing*, Ann. of Math. (2) **151** (2000), no. 3, 877–960. MR1779561 (2002e:37134) ↑6

[FH14] Peter K. Friz and Martin Hairer, *A course on rough paths*, Universitext, Springer, Cham, 2014. With an introduction to regularity structures. MR3289027 ↑5, 24

[FNS77] Dieter Forster, David R. Nelson, and Michael J. Stephen, *Large-distance and long-time properties of a randomly stirred fluid*, Physical Review A **16** (August 1977), no. 2, 732–749. ↑78

[FQ14] Tadahisa Funaki and Jeremy Quastel, *KPZ equation, its renormalization and invariant measures*, Stochastic Partial Differential Equations: Analysis and Computations **3** (2014), no. 2, 159–220. ↑8, 59

[FS13] Peter Friz and Atul Shekhar, *Doob-Meyer for rough paths*, Bull. Inst. Math. Acad. Sin. (N.S.) **8** (2013), no. 1, 73–84. MR3097417 ↑23

[Fur14] Marco Furlan, *Stochastic Navier-Stokes equation in 3 dimensions*, Master's Thesis, 2014. supervised by Massimiliano Gubinelli. ↑13

[Gär88] Jürgen Gärtner, *Convergence towards Burgers' equation and propagation of chaos for weakly asymmetric exclusion processes*, Stochastic Process. Appl. **27** (1988), no. 2, 233–260. MR931030 (89e:60200) ↑4

[Gub04] Massimiliano Gubinelli, *Controlling rough paths*, J. Funct. Anal. **216** (2004), no. 1, 86–140, DOI 10.1016/j.jfa.2004.01.002. ↑24

[GIP16] Massimiliano Gubinelli, Peter Imkeller, and Nicolas Perkowski, *A Fourier analytic approach to pathwise stochastic integration*, Electron. J. Probab. **21** (2016), no. 2. ↑5

[GIP15] ———, *Paracontrolled distributions and singular PDEs*, Forum Math. Pi **3** (2015), no. 6. ↑5, 8, 11, 13, 37, 108

[GJ13] Massimiliano Gubinelli and Milton Jara, *Regularization by noise and stochastic Burgers equations*, Stoch. Partial Differ. Equ. Anal. Comput. **1** (2013), no. 2, 325–350. MR3327509 ↑4, 35

[GJ14] Patrícia Gonçalves and Milton Jara, *Nonlinear fluctuations of weakly asymmetric interacting particle systems*, Arch. Ration. Mech. Anal. **212** (2014), no. 2, 597–644. MR3176353 ↑4, 35

[GLT06] Massimiliano Gubinelli, Antoine Lejay, and Samy Tindel, *Young integrals and SPDEs*, Potential Anal. **25** (2006), no. 4, 307–326. MR2255351 (2007k:60182) ↑31

[GP15a] Massimiliano Gubinelli and Nicolas Perkowski, *Lectures on singular stochastic PDEs*, Ensaios Mat. **29** (2015), 1–89. ↑5

[GP15b] ———, *Energy solutions of KPZ are unique*, arXiv preprint arXiv:1508.07764 (2015). ↑4

[Hai11] Martin Hairer, *Rough stochastic PDEs*, Comm. Pure Appl. Math. **64** (2011), no. 11, 1547–1585. ↑7

[Hai13] ———, *Solving the KPZ equation*, Ann. Math. **178** (2013), no. 2, 559–664. ↑5, 7, 15, 24, 26

[Hai14] ———, *A theory of regularity structures*, Invent. Math. **198** (2014), no. 2, 269–504. MR3274562 ↑5, 24

[HL15] Martin Hairer and Cyril Labbé, *Multiplicative stochastic heat equations on the whole space*, arXiv preprint arXiv:1504.07162 (2015). ↑35

[HM15] Martin Hairer and Konstantin Matetski, *Discretisations of rough stochastic PDEs*, arXiv preprint arXiv:1511.06937 (2015). ↑8, 56, 59

[HM12] Martin Hairer and Jan Maas, *A spatial version of the Itô-Stratonovich correction*, Ann. Probab. **40** (2012), no. 4, 1675–1714. MR2978135 ↑7, 55

[HMW14] Martin Hairer, Jan Maas, and Hendrik Weber, *Approximating rough stochastic PDEs*, Comm. Pure Appl. Math. **67** (2014), no. 5, 776–870. MR3179667 ↑7, 8, 55

[HP13] Martin Hairer and Natesh S. Pillai, *Regularity of laws and ergodicity of hypoelliptic SDEs driven by rough paths*, Ann. Probab. **41** (2013), no. 4, 2544–2598. MR3112925 ↑23

[HQ15] Martin Hairer and Jeremy Quastel, *A class of growth models rescaling to KPZ*, arXiv preprint arXiv:1512.07845 (2015). ↑5

[HS15] Martin Hairer and Hao Shen, *A central limit theorem for the KPZ equation*, arXiv preprint arXiv:1507.01237 (2015). ↑5

[KPZ86] Mehran Kardar, Giorgio Parisi, and Yi-Cheng Zhang, *Dynamic scaling of growing interfaces*, Physical Review Letters **56** (1986), no. 9, 889–892. ↑3

[KS88] Ioannis Karatzas and Steven E. Shreve, *Brownian motion and stochastic calculus*, Springer, 1988. ↑43

[KS91] Joachim Krug and Herbert Spohn, *Kinetic roughening of growing surfaces*, C. Godreche, Cambridge University Press, Cambridge **1** (1991), no. 99, 1. ↑7, 58

[Kup14] Antti Kupiainen, *Renormalization group and stochastic PDEs*, Annales Henri Poincaré (2014), 1–39. ↑5

[LS98] Chi-Hang Lam and Franklin G Shin, *Improved discretization of the Kardar-Parisi-Zhang equation*, Phys. Rev. E **58** (1998), no. 5, 5592–5595. ↑58

[Mey81] Yves Meyer, *Remarques sur un théorème de J.-M. Bony*, Rendiconti del Circolo Matematico di Palermo. Serie II, 1981, pp. 1–20. ↑9

[MF14] Gregorio R. Moreno Flores, *On the (strict) positivity of solutions of the stochastic heat equation*, Ann. Probab. **42** (2014), no. 4, 1635–1643. MR3262487 ↑7, 48

[Mue91] Carl Mueller, *On the support of solutions to the heat equation with noise*, Stochastics Stochastics Rep. **37** (1991), no. 4, 225–245. MR1149348 (93e:60122) ↑7, 48

[MW14] Jean-Christophe Mourrat and Hendrik Weber, *Convergence of the two-dimensional dynamic Ising-Kac model to ϕ^4_2*, arXiv preprint arXiv:1410.1179 (2014). ↑8

[Per14] Nicolas Perkowski, *Studies of robustness in stochastic analysis and mathematical finance*, Ph.D. Thesis, 2014. ↑34

[PT16] David J Prömel and Mathias Trabs, *Rough differential equations driven by signals in Besov spaces*, J. Differential Equations **260** (2016), no. 6, 5202–5249. ↑5, 37

[QS15] Jeremy Quastel and Herbert Spohn, *The One-Dimensional KPZ Equation and its Universality Class*, J. Stat. Phys. **160** (2015), no. 4, 965–984. MR3373647 ↑3

[Qua14] Jeremy Quastel, *The Kardar-Parisi-Zhang equation and universality class*, XVIIth International Congress on Mathematical Physics, 2014, pp. 113–133. MR3204148 ↑3

[SS09] Tomohiro Sasamoto and Herbert Spohn, *Superdiffusivity of the 1D lattice Kardar-Parisi-Zhang equation*, J. Stat. Phys. **137** (2009), no. 5-6, 917–935. MR2570756 (2011a:82063) ↑7, 58, 59

[SS10] ———, *Exact height distributions for the KPZ equation with narrow wedge initial condition*, Nuclear Phys. B **834** (2010), no. 3, 523–542. MR2628936 (2011c:82067) ↑3

[Üst14] Ali Süleyman Üstünel, *Variational calculation of Laplace transforms via entropy on Wiener space and applications*, J. Funct. Anal. **267** (2014), no. 8, 3058–3083. MR3255483 ↑44

[ZK65] Norman J Zabusky and Martin D Kruskal, *Interaction of "solitons" in a collisionless plasma and the recurrence of initial states*, Physical Rev. Lett. **15** (1965), no. 6, 240. ↑58

[ZZ14] Rongchan Zhu and Xiangchan Zhu, *Approximating three-dimensional Navier-Stokes equations driven by space-time white noise*, arXiv preprint arXiv:1409.4864 (2014). ↑5, 8

[ZZ15a] ———, *Three-dimensional Navier-Stokes equations driven by space-time white noise*, J. Differential Equations **259** (2015), no. 9, 4443–4508. ↑5

[ZZ15b] ———, *Lattice approximation to the dynamical Φ^4_3 model*, arXiv preprint arXiv:1508.05613 (2015). ↑59