Transportation cost inequalities for singular SPDEs

I. BAILLEUL, M. HOSHINO and R. TAKANO

Abstract. We prove that the laws of the BPHZ random models satisfy some transportation cost inequalities in the full subcritical regime if there is no 'variance blowup' and the law of the noise is translation invariant and satisfies some transportation cost inequality. As a consequence, the laws of a number of solutions to some singular stochastic partial differential equations also satisfy some transportation cost inequalities. This is in particular the case of the $\Phi^4_{4-\delta}$ measures over the 4-dimensional torus, for all $0 < \delta < 4$.

1 – Introduction

Denote by $\mathsf{P}(E)$ the set of probability measures on a measurable space $(E, \mathcal{E})$. Fix a measurable function $c : E \times E \to [0, +\infty]$, usually called a cost function. The c-transportation cost between two probability measures μ_1, μ_2 on $(E, \mathcal{E})$ is defined as

$$\mathcal{T}_c(\mu_1, \mu_2) := \inf \mathbb{E}\big[c(X_1, X_2)\big]$$

for an infimum over the set of all random variables defined on some probability space such that X_1 has law μ_1 and X_2 has law μ_2. Given μ_1, μ_2 in $\mathsf{P}(E)$, the relative entropy $\mathcal{H}(\mu_2|\mu_1)$ of μ_2 with respect to μ_1 is defined as $\int_E \log\big(\frac{d\mu_2}{d\mu_1}\big) d\mu_2$, if μ_2 is absolutely continuous with respect to μ_1, otherwise we set it equal to $+\infty$. Let $\ell : [0, +\infty] \to [0, +\infty]$ be a continuous increasing unbounded function with $\ell(0) = 0$. A probability measure μ on $(E, \mathcal{E})$ is said to satisfy an (ℓ, c)-***transportation cost inequality*** if

$$\ell\big(\mathcal{T}_c(\nu, \mu)\big) \leq \mathcal{H}(\nu|\mu) \qquad (\forall\, \nu \in \mathsf{P}(E)). \tag{1.1}$$

This type of information theoretic property of a probability measure was first introduced by K. Marton in [28] in relation with the phenomenon of concentration of measure in metric spaces, for (E, d) a metric space, $c(x, y) = d(x, y)^2$ and $\ell(t)$ proportional to t. We talk in that case of a 2-transportation cost inequality. Gozlan & Léonard's review [17] provides a nice overview of transportation cost inequalities.

Talagrand [36] proved that Gaussian measures on a Euclidean space satisfy a dimension-free version of (1.1). This seminal work was followed by numerous other works in different directions. We single out the very influencial work [29] of Otto & Villani which clarified among other things the relation between the 2-transportation cost inequality and the log-Sobolev and Poincaré inequalities, in a Euclidean or Riemannian setting. We also mention Feyel & Üstünel's extension of Talagrand's result to an abstract Wiener space setting [14], which opened the door to proving a number of transportation cost inequalities for probability measures that are the laws of some solutions of some stochastic (possibly partial) differential equations, as in the works of Djellout, Guillin & Wu [12] and Saussereau [31], or the works of Wu & Zhang [37], Khoshnevisan & Sarantsev [24], Shang & Zhang [34], Shang & Wang [33], Dai & Li [11] or Li & Wang [25], to cite but a few references.

All the preceding results involve some classical Itô type dynamics. The pathwise solution theories of stochastic dynamics come with some powerful tools that provide some robust and deep informations on the laws of these solutions. The theory of rough paths [27] gives a pathwise picture of Itô or Stratonovich stochastic differential equations in which the solution map appears as the composition of a probabilistic lifting map, lifting the noise into a random rough path, and a deterministic solution map, defined on the space of all rough paths and with values in the space of paths. Building on this picture, Riedel proved in [30] that the law of the solution path to a rough differential equation driven by certain classes of Gaussian noises satisfies a transportation cost inequality if the law of the noise satisfies such an inequality.

On the SPDE side, the different pathwise theories developed within the last ten years offer a similar opportunity. It is the purpose of the present work to prove a very general result that provides automatically some transportation cost inequalities for the laws of the solutions of some singular stochastic partial differential equations.

We will work in the setting of regularity structures [18]. The architecture of that theory is similar to the architecture of rough paths theory. Given any equation in a well identified class of equations, there is a probabilistic lifting map that enriches the noise in the form of a more complex object, called a model. There is also a deterministic, locally Lipschitz, solution map that associates to any model a solution field. Gasteratos & Jacquier were able to prove in [15] that a transportation cost inequality for some Gaussian noises can be propagated to the solution field of two examples of mildly singular stochastic PDEs, the rough volatility model in finance and the two dimensional parabolic Anderson model equation on the torus. Compared to the rough paths setting, the singular SPDE setting involves an additional difficulty related to the fact that the probabilistic definition of a random model typically involves a non-trivial renormalization procedure. This explains why it is difficult to extend the hand-crafted computations of [15] to some more singular equations, as the renormalization process gets more complex when the equation gets more singular.

The map lifting a noise into a model is not unique; we will work here with the Bogoliubov–Parasiuk–Hepp–Zimmermann (BPHZ) prescription first introduced by Bruned, Hairer & Zambotti in [6]. The lifting map is defined by a limiting procedure whose convergence was first proved in Chandra & Hairer's unpublished work [7] under a set of assumptions on the noise giving some controls on its cumulants. The limit model is called the BPHZ model. The convergence of the procedure was proved under a different set of assumptions in Hairer & Steele's work [22] and Bailleul & Hoshino's work [1], where one asks that the law of the noise satisfies a ***spectral gap***. Here is what it means. The noise is defined on a Banach space Ω which contains a dense subspace H which is itself a Hilbert space for some norm $\|\cdot\|_H$ stronger than the norm inherited from Ω. One says that the law of the noise satisfies a spectral gap inequality with constant $a_1 > 0$ iff one has

$$\mathbb{E}\big[(F - \mathbb{E}[F])^2\big] \leq a_1 \mathbb{E}\big[\|dF\|_{H^*}^2\big] \tag{1.2}$$

for all cylindrical functions $F(\omega) = f\big(\varphi_1(\omega), \dots, \varphi_n(\omega)\big)$ with $\varphi_i \in \Omega^*$ and some smooth function $f : \mathbf{R}^n \to \mathbf{R}$ with at most polynomial growth. The H-differential dF of F at point $\omega \in \Omega$ is given for any $h \in H$ by $(d_\omega F)(h) = \sum_{i=1}^n \partial_i f\big(\varphi_1(\omega), \dots, \varphi_n(\omega)\big)\varphi_i(h)$. We denoted here by H^* and Ω^* the topological duals of H and Ω respectively. We define the cost

$$c_H(\omega_1, \omega_2) = \begin{cases} \|h\|_H^2 & \text{if } \omega_1 - \omega_2 = h \in H, \\ +\infty & \text{otherwise.} \end{cases}$$

Different tools were used in [22] and [1] to prove the convergence of the BPHZ models. We make the following assumption throughout this work.

Assumption A – *We are given a separable Banach space Ω of spacetime functions/distributions which contains a dense subspace H which is itself a Hilbert space for some norm $\|\cdot\|_H$ stronger than the norm inherited from Ω. We equip Ω with its Borel σ-algebra. The noise is defined as the identity map on Ω. Its law $\mathbb{P} \in \mathsf{P}(\Omega)$ is centered, invariant under spacetime shifts, and satisfies the transportation cost inequality*

$$a_\circ \mathcal{T}_{c_H}(\mathbb{Q}, \mathbb{P}) \leq \mathcal{H}(\mathbb{Q}|\mathbb{P}) \qquad (\forall\, \mathbb{Q} \in \mathsf{P}(\Omega)). \tag{1.3}$$

for some positive finite constant $a_\circ$.

In our situation, the space Ω will be a certain Banach space of distributions on $\mathbf{R}^d$. It is a folklore result that the spectral gap inequality (1.2) holds under Assumption A, with $a_1 = 1/(2a_\circ)$. A proof of this result is given in Appendix A as we could not find an easily accessible direct proof in the literature. The BPHZ model is thus well defined under Assumption A, using the constructions of [22] or [1].

We use in the present work a refinement of the construction of the BPHZ model given in Bailleul & Hoshino's work [1]. The introduction of this refinement is motivated by Gasteratos & Jacquier's extended contraction principle for transportation cost inequalities that takes in our setting the following form – see Lemma 2.11 in [15] for the original statement and Section 2.1 for a proof. Let M be a metric space.

1 – Proposition. *Let $M:\Omega\rightarrow\mathsf{M}$ be a random variable and $c_{\mathsf{M}}:\mathsf{M}\times\mathsf{M}\rightarrow[0,+\infty]$ be a measurable function such that*

$$c_{\mathsf{M}}\big(M(\omega_1),M(\omega_2)\big)\leq\mathbb{L}(\omega_2)\max\left(\sqrt{c_H(\omega_1,\omega_2)},\sqrt{c_H(\omega_1,\omega_2)}^{1/r}\right),\tag{1.4}$$

for a finite exponent $r\geq1$ and all ω_1,ω_2 in some measurable subset of Ω of $\mathbb{P}$-probability 1, for $\mathbb{L}\in\bigcap_{2\leq p<\infty}L^p(\Omega,\mathbb{P})$. Assume further that Assumption A holds. Then the law $P_M=\mathbb{P}\circ M^{-1}$ of M satisfies for any $1\leq\alpha<2$ the transportation cost inequality

$$a_\alpha\mathcal{T}_{c_{\mathsf{M}}^\alpha}(Q,P_M)\leq\mathcal{H}(Q|P_M)^{\alpha/(2r)}+\mathcal{H}(Q|P_M)^{\alpha/2}\qquad(\forall\,Q\in\mathsf{P}(\mathsf{M}))\tag{1.5}$$

where $a_\alpha=(2\|\mathbb{L}^\alpha\|_{L^{2/(2-\alpha)}(\mathbb{P})})^{-1}\min(a_\circ^{\alpha/(2r)},a_\circ^{\alpha/2})$.

Equation (1.5) corresponds to an $(\ell_\alpha,c_{\mathsf{M}}^\alpha)$-transportation cost inequality (1.1) with ℓ_α the convex and continuous inverse function of the non-negative increasing continuous concave function $(t\in[0,\infty))\mapsto a_\alpha^{-1}(t^{\alpha/(2r)}+t^{\alpha/2})$. Given the statement of Proposition 1, it is not surprising that the crux for proving a transportation cost inequality for the BPHZ random model $\widetilde{\mathsf{M}}^{\infty;\varepsilon}$ lies in proving a pathwise inequality of the form

$$\big\|\widetilde{\mathsf{M}}^{\infty;\varepsilon}(\omega+h):\widetilde{\mathsf{M}}^{\infty;\varepsilon}(\omega)\big\|_{\mathsf{M}}\leq\mathbb{L}(\omega)\max\big(\|h\|_H,\|h\|_H^{1/r}\big)\tag{1.6}$$

for all $\omega\in\Omega$, all $h\in H$, for some exponent $r\geq1$ and some random variable $\mathbb{L}$ in all the $L^p(\Omega,\mathbb{P})$ spaces $(1\leq p<\infty)$, and where M stands in the remainder of this introduction for some Polish space of models. This will be proved in Theorem 12 below, with all the notations defined in the meantime. Proposition 1 will then entail that the law of the BPHZ model satisfies a family of transportation cost inequalities, from which it will classically follow that its norm has a Gaussian tail. We note here that the norm $\|\cdot\|_{\mathsf{M}}$ involved in (1.6) will be homogeneous with respect to some natural dilation operation on the space of models.

The question of the longtime existence of some solutions to some singular SPDEs is subtler than the corresponding question for rough differential equations. It is only very recently that some first longtime existence results were obtained for some classes of equations, by Chandra, Feltes & Weber [9], Shen, Zhu & Zhu [35] and Chevyrev & Gubinelli [10]. Independently of these examples, under the assumption that a given subcritical singular stochastic PDE is defined on a fixed time interval, it is possible to infer from the transportation cost inequality satisfied by the BPHZ model that the law of the solution field satisfies as well a transportation cost inequality, assuming some a priori pathwise bounds on the solution field. Such a statement will be the object of Corollary 7 below. Incidentally, Theorem 12 and Corollary 7 allow to extend Riedel's transportation cost result [30] on the law of solutions to some rough differential equations driven by some Gaussian p-rough paths with $2<p<3$ to some arbitrary geometric, or branched, p-rough paths $(1\leq p<\infty)$ whose laws satisfy some transportation cost inequalities. This includes some non-Gaussian examples.

In a different direction, we obtain as a consequence of Corollary 7 some transportation cost inequalities for some Φ^4 measures in the full subcritical regime. Let us work over the 4-dimensional torus $\mathbf{T}^4$. For $\nu>\kappa>0$, with ν irrational and $u_0\in C^{-1-\kappa+\nu}(\mathbf{T}^4)$, consider the subcritical singular stochastic PDE

$$(\partial_t-\Delta)u=-u^3+(-\Delta)^{-\nu}(\zeta)\tag{1.7}$$

where ζ stands for a spacetime white noise over $\mathbf{T}^4$. Denote by $\widetilde{\mathsf{M}}^{\infty;\varepsilon}$ the BPHZ model over the regularity structure associated with this equation; it satisfies the Hölder bound (1.6) for some exponent $r(\nu)$ depending on ν. Equation (1.7) has a unique solution $u_t(u_0)$ in the setting of regularity structures. It defines a Markovian dynamics on $C^{-1-\kappa+\nu}(\mathbf{T}^4)$, defined globally in time and which satisfies the u_0-uniform polynomial bound

$$\|u_1(u_0)\|_{C^{-1-\kappa+\nu}}\lesssim\big(1+\|\widetilde{\mathsf{M}}^{\infty;\varepsilon}\|_{\mathsf{M}}\big)^{p(\nu)}$$

for some finite exponent $p(\nu) \geq 1$ depending on ν. This a priori estimates was proved by Chandra, Moinat & Weber in [8]. Denote by $\pi_t(u_0) \in \mathsf{P}(C^{-1-\kappa+\nu}(\mathbf{T}^3))$ the law of $u_t(u_0)$. The bound above entails that the family of probability measures

$$\Big(\frac{1}{s}\int_1^{s+1} \pi_t(u_0)dt\Big)_{s\geq 1}$$

on $C^{-1-\kappa+\nu}(\mathbf{T}^3)$ is tight. Any limit point of this family as s goes to $+\infty$ is an invariant probability measure for the Markovian dynamics (1.7). It can be inferred from the deep results of Hairer & Mattingly [20] on the strong Feller property, and the results of Hairer & Schönbauer [21] on the support theorem for the BPHZ model, that the Markovian dynamics (1.7) has a unique invariant probability measure. We call such a measure the $\Phi^4_{4-\nu}$ measure.

Set $\underline{\ell}(t) := t^2$ and define for any v_1, v_2 in $C^{-1-\kappa+\nu}(\mathbf{T}^4)$ the cost function

$$c_\nu(v_1, v_2) = \|v_2 - v_1\|_{C^{-1-\kappa+\nu}}^{1/p(\nu)}.$$

2 – Theorem. *There is a constant $a \in (0, +\infty)$ such that the $\Phi^4_{4-\nu}$ probability measure on $C^{-1-\kappa+\nu}(\mathbf{T}^4)$ satisfies an $(a\underline{\ell}, c_\nu)$-transportation cost inequality.*

Theorem 8 below is more general and shows that a whole class of invariant probability measures of some Markovian dynamics satisfies some explicit transportation cost inequalities. It is expected that the $\Phi^4_{4-\nu}$ measures satisfy some stronger log-Sobolev inequalities. The methods of the present work rely on some pathwise local Hölder estimates of the form (1.6); they do not allow a direct transfer of some log-Sobolev inequality unlike some global Lipschitz estimates. On the other hand, they are robust enough to apply in some situations where it is not clear that the methods used to prove some log-Sobolev inequalities have a chance to be effective, as for instance the methods developed in Bauerschmidt, Bodineau and Dagallier's works [3, 4].

Organisation of the work – We dedicate Section 2 to drawing some of the consequences that the fundamental Hölder regularity estimate (1.6) has for the laws of the solutions of some singular stochastic PDEs that are well defined globally in time. We prove in particular in Corollary 5 that the homogeneous norms of the BPHZ models have some Gaussian tails. The above mentioned Theorem 8 and Theorem 2 are proved in this section. We describe our functional setting in Section 3. Section 4 is dedicated to the description of a particular regularity-integrability structure and to the construction of an extended BPHZ model over this setting. This construction is a variation of the construction of the BPHZ model given in [1]. We prove that the extended BPHZ model satisfies the Hölder regularity estimate (1.6) in Section 5. We prove in Appendix A that the spectral gap inequality (1.2) holds under Assumption A, with $a_1 = 1/(2a_\circ)$.

We refer the reader to Bailleul & Hoshino's Tourist guide [2] for some background on regularity structures.

Notations – We collect here a few notations that are used throughout the text. For any multiindex $k = (k^j)_{j=1}^{1+d} \in \mathbf{N}^{1+d}$ we define

$$|k|_{\mathfrak{s}} := 2k_1 + \sum_{j=2}^{1+d} k^j, \qquad k! := \prod_{j=1}^{1+d} k^j!, \qquad \partial^k := \prod_{j=1}^{1+d} \partial_j^{k^j},$$

where ∂_j is the partial derivative with respect to the j-th variable x_j on $\mathbf{R}^{1+d}$. For $k, l \in \mathbf{N}^{1+d}$, we write $k \leq l$ if $k^j \leq l^j$ for any $j \in \{1, \dots, 1+d\}$, and we sets $\binom{l}{k} := \prod_{j=1}^{1+d}\binom{l^j}{k^j}$. For every $x = (x_j)_{j=1}^{1+d} \in \mathbf{R}^{1+d}$ and $k = (k^j)_{j=1}^{1+d} \in \mathbf{N}^{1+d}$ we define

$$x^k := \prod_{j=1}^{1+d} x_j^{k^j}, \qquad \|x\|_{\mathfrak{s}} := \sqrt{|x_1|} + \sum_{j=2}^{1+d} |x_j|.$$

2 – Transportation cost inequalities for solutions of singular SPDEs

We consider a subcritical singular parabolic stochastic PDEs

$$(\partial_t - \Delta)u = \mathcal{F}(u, \nabla u; \xi) \qquad (t > 0,\ x \in \mathbf{T}^d) \tag{2.1}$$

with initial condition $u_0 \in C^\gamma(\mathbf{T}^d)$ for some exponent $\gamma \in \mathbf{R}$ and ξ a random (spacetime) noise. The function $\mathcal{F}$ is affine in its ξ argument. We will show below that there is a certain extension of the Bruned–Hairer–Zambotti (BHZ) regularity structure [6] associated with the equation over which one can construct an extended version of the BPHZ random model. We will denote this extended version by $\widetilde{\mathsf{M}}^{\infty;\varepsilon}$ in the sequel. The extension of the BHZ regularity structure will be described in Section 4.1 and the random model $\widetilde{\mathsf{M}}^{\infty;\varepsilon}$ will be constructed in Section 4.2. The metric space

$$\big(\mathbf{M}(\mathscr{V}_\varepsilon)_{w_a}, \|\cdot:\cdot\|_{\mathbf{M}_{\mathrm{hom}}(\mathscr{V}_\varepsilon)_{w_a}}\big)$$

in which the BPHZ model $\widetilde{\mathsf{M}}^{\infty;\varepsilon}$ takes its values, and all the notations involved, will be explained in Section 4; they do not matter presently. In this section we will simply denote this space by $(\mathbf{M}, \|\cdot:\cdot\|_{\mathbf{M}})$. For $\mathsf{M} \in \mathbf{M}$ we set $\|\mathsf{M}\|_{\mathbf{M}} := \|\mathbf{0} : \mathsf{M}\|_{\mathbf{M}}$ where $\mathbf{0}$ stands for the zero model over the given regularity structure.

We will prove in Theorem 12 of Section 5 that if the law of ξ satisfies Assumption A and the condition (4.6) below holds – the phenomenon called 'variance blow-up' in [19] does not occur, then one has for all $\omega \in \Omega$ and $h \in H$ the inequality

$$\big\|\widetilde{\mathsf{M}}^{\infty;\varepsilon}(\omega + h) : \widetilde{\mathsf{M}}^{\infty;\varepsilon}(\omega)\big\|_{\mathbf{M}} \leq \mathbb{L}(\omega)\max\big(\|h\|_H, \|h\|_H^{1/m}\big) \tag{2.2}$$

for some random variable $\mathbb{L} \in \bigcap_{1\leq p<\infty} L^p(\Omega, \mathbb{P})$ and some integer m that depends only on $\mathcal{F}$ and the law of ξ. (We note here that (2.2) would take a different form if we were using the usual inhomogeneous norm – see (5.6) below.) This regularity estimate of Hölder type is all we need to draw some strong conclusions on the law of the solution to the equation. We obtain them from two elementary facts that we now recall.

2.1 – Two useful facts on transportation cost inequalities.

To make this work self-contained, we give a proof of Proposition 1 following Gasteratos & Jacquier's elementary proof of their extended contraction principle – Lemma 2.11 in [15].

Proof of Proposition 1 – Denote by $P_M = \mathbb{P} \circ M^{-1} \in \mathsf{P}(\mathbf{M})$ the law of M, and pick $Q \in \mathsf{P}(\mathbf{M})$ with $\mathcal{H}(Q|P_M) < \infty$. Since Ω is a Polish space, for any $Q \in \mathsf{P}(\mathbf{M})$ with $\mathcal{H}(Q|P_M) < \infty$ one has

$$\mathcal{H}(Q|P_M) = \inf\Big\{\mathcal{H}(\mathbb{Q}|\mathbb{P})\,;\, \mathbb{Q} \in \mathsf{P}(\Omega),\ Q = \mathbb{Q} \circ M^{-1}\Big\}. \tag{2.3}$$

Write $\Theta(Q, P_M) \subset \mathsf{P}(\mathbf{M} \times \mathbf{M})$ for the set of couplings of $Q \in \mathsf{P}(\mathbf{M})$ and P_M, and denote by E the expectation operator associated with any of these couplings. Similarly, write $\Theta(\mathbb{Q}, \mathbb{P}) \subset \mathsf{P}(\Omega \times \Omega)$ the set of couplings of $\mathbb{Q} \in \mathsf{P}(\Omega)$ and $\mathbb{P}$ and denote by $\mathbb{E}$ the expectation operator associated with any of these couplings. For all $1 \leq \alpha < 2$, one has $2r/(2r - \alpha) \leq 2/(2 - \alpha)$ and

$$\begin{aligned}
\inf_{\Theta(Q,P_M)} E\big[c^\alpha_{\mathbf{M}}(\cdot,\cdot)\big] &\leq \inf_{\Theta(\mathbb{Q},\mathbb{P})} \mathbb{E}\big[c^\alpha_{\mathbf{M}}(M\cdot, M\cdot)\big] \\
&\overset{(1.4)}{\leq} \inf_{\Theta(\mathbb{Q},\mathbb{P})} \Big(\mathbb{E}\big[\mathbb{L}^\alpha c_H^{\alpha/2}(\cdot,\cdot)\mathbf{1}_{c_H\geq 1}\big] + \mathbb{E}\big[\mathbb{L}^\alpha c_H^{\alpha/(2r)}(\cdot,\cdot)\mathbf{1}_{c_H<1}\big]\Big) \\
&\leq \inf_{\Theta(\mathbb{Q},\mathbb{P})} \Big(\|\mathbb{L}^\alpha\|_{L^{2/(2-\alpha)}(\mathbb{P})}\mathbb{E}[c_H]^{\alpha/2} + \|\mathbb{L}^\alpha\|_{L^{2r/(2r-\alpha)}(\mathbb{P})}\mathbb{E}[c_H]^{\alpha/(2r)}\Big) \\
&\overset{(1.3)}{\leq} \|\mathbb{L}^\alpha\|_{L^{2/(2-\alpha)}(\mathbb{P})}\left\{\Big(\frac{\mathcal{H}(\mathbb{Q}|\mathbb{P})}{a_\circ}\Big)^{\frac{\alpha}{2}} + \Big(\frac{\mathcal{H}(\mathbb{Q}|\mathbb{P})}{a_\circ}\Big)^{\frac{\alpha}{2r}}\right\} \\
&\leq b_\alpha \max\big(\mathcal{H}(\mathbb{Q}|\mathbb{P})^{\alpha/2}, \mathcal{H}(\mathbb{Q}|\mathbb{P})^{\alpha/(2r)}\big)
\end{aligned}$$

with $b_\alpha = 2\|\mathbb{L}^\alpha\|_{L^{2/(2-\alpha)}(\mathbb{P})}\max\big(a_\circ^{-\alpha/(2r)}, a_\circ^{-\alpha/2}\big)$, using Hölder inequality in the third inequality. One then gets the conclusion from (2.3). $\triangleright$

The proof makes clear what conclusion can be obtained if we only assume in Proposition 1 that $\mathbb{L} \in L^p(\Omega, \mathbb{P})$ for some finite $p \geq 2$. The following fact follows from Proposition 1 and the estimate (2.2).

3 – Corollary. *The law $\widetilde{P}^{\infty;\varepsilon} \in \mathsf{P}(\mathbf{M})$ of the* BPHZ *model $\widetilde{\mathsf{M}}^{\infty;\varepsilon}$ satisfies for each $1 \leq \alpha < 2$ an $(\ell_\alpha, \|\cdot:\cdot\|_{\mathbf{M}}^{\alpha})$-transportation cost inequality, where ℓ_α stands for the convex and continuous inverse function of the non-negative increasing continuous concave function $(t \in [0,\infty)) \mapsto a_\alpha^{-1}(t^{\alpha/(2r)} + t^{\alpha/2})$, and the random variable $\mathbb{L}$ involved in the definition of the constant a_α is the one that appears in* (2.2)*.*

In the particular case where the cost function is the distance function of a Polish space, one can relate some transportation cost inequalities to some exponential integrability properties of the reference probability measures. The following statement was proved by Gozlan in Theorem 1.13 of [16].

4 – Theorem. *Let $\ell : [0,\infty) \to [0,\infty)$ be a convex, continuous increasing function such that $\ell(0) = 0$. Its conjugate function $\ell^* : [0,\infty) \to [0,\infty]$ is defined for $0 \leq s < \infty$ as*

$$\ell^*(s) = \sup_{r \geq 0} \big\{ sr - \ell(r) \big\}.$$

Let (E, d) be a Polish space and μ be a probability measure on E. Assume that

- *one has $\{\ell^* < \infty\} = [0, s_0)$ for some $s_0 > 0$, and there are some positive constants ρ and $s_1 < s_0$ such that $\ell^*(s) \geq \rho s^2$ for any $s \in [0, s_1]$;*
- *there is a reference point $x_0 \in E$ such that $\int_E d(x_0, x)\mu(dx) < \infty$.*

In this setting, the following properties are equivalent.

(a) *The probability μ satisfies an $\big(\ell(\frac{\cdot}{\kappa_1}), d(\cdot,\cdot)\big)$-transportation cost inequality for some constant $0 < \kappa_1 < \infty$.*

(b) *The integral $\int_E e^{\ell(\kappa_2 d(x_0,x))}\mu(dx)$ is finite for some constant $0 < \kappa_2 < \infty$.*

We deduce from Corollary 3 with $\alpha = 1$ and the (a) $\Rightarrow$ (b) part of Theorem 4 that $\|\widetilde{\mathsf{M}}^{\infty;\varepsilon}\|_{\mathbf{M}}$ has a Gaussian tail.

5 – Corollary. *There is a finite positive constant κ_3 such that*

$$\mathbb{E}\big[\exp\big(\kappa_3 \|\widetilde{\mathsf{M}}^{\infty;\varepsilon}\|_{\mathbf{M}}^2\big)\big] < \infty.$$

Proof – We work here with $\alpha = 1$ in the conclusion of Corollary 3. It is straightforward to show that the concave function ℓ_1 satisfies the conditions of Theorem 4. By applying this theorem in the case where $(E, d) = \big(\mathbf{M}, \|\cdot:\cdot\|_{\mathbf{M}}\big)$ and the reference point in E is the zero model, we obtain the existence of a constant κ_2 such that $\mathbb{E}\big[\exp\big(\ell_1\big(\kappa_2 \|\widetilde{\mathsf{M}}^{\infty;\varepsilon}\|_{\mathbf{M}}\big)\big)\big]$ is finite. The result then follows from the fact that $\ell_1(t) \gtrsim t^2$ for $t \geq 1$.
One can also deduce this statement from (2.2) and Theorem 3.3 of Riedel's work [30]. ▷

We deduce the following fact from Corollary 5 and the (b) $\Rightarrow$ (a) part of Theorem 4. Recall that $\underline{\ell}(t) = t^2$.

6 – Corollary. *Let $(E, \|\cdot\|_E)$ be a Banach space and $X : \Omega \to E$ be a random variable that satisfies the upper bound*

$$g(\|X\|_E) \leq 1 + \big\|\widetilde{\mathsf{M}}^{\infty;\varepsilon}\big\|_{\mathbf{M}}$$

for some continuous concave strictly increasing function $g : [0,\infty) \to [0,\infty)$ with $g(0) = 0$. Then there is a constant $0 < a < \infty$ such that the law of X satisfies an $\big(a\underline{\ell}, g(\|\cdot - \cdot\|_E)\big)$-transportation cost inequality.

Note that $g(\|\cdot - \cdot\|_E)$ is a metric on E and $\big(E, g(\|\cdot - \cdot\|_E)\big)$ is a Polish space. In a typical situation $g(\cdot)$ would be the inverse function of a constant multiple of one of the functions z^p or $\exp\big(\kappa z^p\big) - 1$, for some constants $1 \leq p < \infty$ and $0 < \kappa < \infty$.

2.2 – Transportation cost inequalities for solutions of singular SPDEs. We consider the subcritical singular parabolic stochastic PDEs (2.1). We refer to Hairer's original work [18] or Bailleul & Hoshino's Tourist guide [2] for some background on the regularity structure formulation of Equation (2.1) in a space of modelled distributions depending on $\widetilde{\mathsf{M}}^{\infty;\varepsilon}(\omega)$, and the fact that this reformulation is locally well-posed on a random time interval. This conclusion holds under some regularity assumptions on $\mathcal{F}$ that are irrelevant here. We denote by bold $\boldsymbol{u}$ the solution modelled distribution and by $u = \mathsf{R}^{\widetilde{\mathsf{M}}^{\infty;\varepsilon}}(\boldsymbol{u})$ its reconstruction via the reconstruction operator associated with the BPHZ model $\widetilde{\mathsf{M}}^{\infty;\varepsilon}$.

One needs some assumptions on $\mathcal{F}$ to ensure the longtime existence of $\boldsymbol{u}$ and u. Some systematic investigations for finding such conditions were only done recently in the works of Chandra, Feltes & Weber [9], Shen, Zhu & Zhu [35] and Chevyrev & Gubinelli [10]. The first two works deal with some mildly singular equations for which $\mathcal{F}(u, \nabla u; \xi) = \sigma(u)\xi$ with $\sigma \in C_b^2(\mathbf{R})$ and ξ has almost surely some regularity comparable to the regularity of the two dimensional space white noise. The third work deals with some Φ^4-like dynamics where $\mathcal{F}(u, \nabla u; \xi) = P(u, \nabla u) + f(u, \nabla u)\xi$ and P is coercive in some sense, as in the Φ^4 example where $P(u, \nabla u) = -u^3$ and f is constant. They can also deal with some non-constant functions f, depending on the almost sure regularity of ξ – see Section 6.3 of [10]. We note that unlike [9, 35], the results of [10] hold in the full subcritical regime and for some dynamics set in the full Euclidean space.

We assume for the remainder of this section that the singular SPDE (2.1) *is globally well-posed; we denote by u the solution.*

In all the situations considered in [9, 35, 10] one gets the longtime existence of $\boldsymbol{u}$ and u from some a priori estimates that control the size of u in some spacetime domains in terms of the (BPHZ) model. Corollary 6 then takes the following form, again with $\underline{\ell}(t) = t^2$.

7 – Corollary. *Assume that the random solution u to Equation* (2.1) *takes its values in a Banach space $(E, \|\cdot\|_E)$ and one has*

$$g(\|u\|_E) \leq 1 + \|\widetilde{\mathsf{M}}^{\infty;\varepsilon}\|_{\mathsf{M}}$$

for some continuous concave strictly increasing function $g : [0,\infty) \to [0,\infty)$ with $g(0) = 0$. Then the law of u satisfies an $\big(a\underline{\ell}, g(\|\cdot - \cdot\|_E)\big)$-transportation cost inequality for some finite positive constant a.

The function space E is typically of the form $C\big((0,T]; C^\gamma(\mathbf{T}^d)\big)$ for some fixed positive finite time T, some $\gamma \in \mathbf{R}$ and some norm of the form $\sup_{0<t\leq T} t^q \|u(t)\|_{C^\gamma}$ with $0 < q < \infty$. One has for instance $g(z) = z^{1/p}$ for some constant $1 \leq p < \infty$ for the solution of the dynamical Φ^4_3 equation [8] and the two-dimensional parabolic Anderson model [9]. To emphasize the dependence of u on the initial condition u_0, we denote by $u(t; u_0)$ the value at time t of u. We denote below by $(\theta_s)_{s\in[0,\infty)}$ a family of shift operators acting on time-space functions/distributions via the formula $\theta_s(\cdot)(r,x) = (\cdot)(r+s,x)$. In our setting, the probability space Ω is a distribution space and the shifts act on Ω. Recall from Assumption A that the probability $\mathbb{P}$ is invariant by the action of the shifts.

8 – Theorem. *Assume that there exists an open interval $(\alpha,\beta) \subset \mathbf{R}$ such that the dynamics* (2.1) *defines a Feller process on $C^\gamma(\mathbf{T}^d)$ for any $\gamma \in (\alpha,\beta)$. Assume as well the existence of a continuous concave strictly increasing unbounded function $g : [0,\infty) \to [0,\infty)$, with $g(0) = 0$, such that $u(1; u_0)$ satisfies the u_0-uniform bound*

$$g(\|u(1;u_0)\|_{C^\gamma}) \leq 1 + \|\widetilde{\mathsf{M}}^{\infty;\varepsilon}\|_{\mathsf{M}}. \tag{2.4}$$

Then the dynamics (2.1) *on $C^\gamma(\mathbf{T}^d)$ has an invariant probability measure, which further satisfies an $\big(a\underline{\ell}, g(\|\cdot - \cdot\|_{C^\gamma})\big)$-transportation cost inequality for some constant $0 < a < \infty$.*

The condition (2.4) is of course reminiscent of the 'coming down from infinity' property satisfied by the solution of the Φ^4_3 dynamics.

Proof – By the Markov property, we have for any $t \geq 1$

$$g\big(\|u(t;u_0)\|_{C^\gamma}\big) = g\big(\|u(1;u(t-1))\|_{C^\gamma}\big) \leq 1 + \|\theta_{t-1}\widetilde{\mathsf{M}}^{\infty;\varepsilon}\|_{\mathsf{M}}.$$

It follows form the shift invariance of $\mathbb{P}$ that the law of $\widetilde{\mathsf{M}}^{\infty;\varepsilon}$ is also shift-invariant. We have as a consequence

$$\sup_{1\leq t<\infty} \mathbb{E}\big[\exp\big(\kappa_2 g(\|u(t;u_0)\|_{C^\gamma})^2\big)\big] < \infty \tag{2.5}$$

for some positive constant κ_2. Denote by $\pi_t(u_0) \in \mathsf{P}\big(C^\gamma(\mathbf{T}^d)\big)$ the law of $u(t;u_0)$ and set $\overline{\pi}_s(u_0) := \frac{1}{s}\int_1^{s+1}\pi_t(u_0)dt$ for $s \geq 1$. The uniform bound (2.5) classically implies the weak convergence of $\overline{\pi}_s(u_0)$ in $\mathsf{P}\big(C^\gamma(\mathbf{T}^d)\big)$ to a probability measure μ invariant for the dynamics (2.1), as s moves along a particular sequence $(s_n)_{n\geq 1}$ diverging to infinity. The reasoning goes as follows: Define for any constant $0 < m < \infty$ the set $K_m = \big\{f \in C^{\gamma+\delta}(\mathbf{T}^d)\,;\, \|f\|_{C^{\gamma+\delta}} \leq m\big\} \subset C^\gamma(\mathbf{T}^d)$, for $\delta > 0$ small. The set K_m is bounded in $C^{\gamma+\delta}(\mathbf{T}^d)$ hence compact in $C^\gamma(\mathbf{T}^d)$. We deduce from the t-uniform bound (2.5) with γ replaced by $\gamma+\delta$ implies by Chebychev inequality the upper bound $\overline{\pi}_s(u_0)(C^\gamma(\mathbf{T}^d)\setminus K_m) \lesssim e^{-\kappa_2 g(m)^2}$, from which the tightness of the $(\overline{\pi}_s(u_0))_{1\leq s<\infty}$ follows since g is increasing with $g(+\infty) = +\infty$.
The probability measure μ satisfies

$$\begin{aligned}
\int_{C^\gamma(\mathbf{T}^d)} e^{\kappa_2 g(\|\phi\|_{C^\gamma})^2}\mu(d\phi) &\leq \liminf_{n\to\infty}\int_{C^\gamma(\mathbf{T}^d)} e^{\kappa_2 g(\|\phi\|_{C^\gamma})^2}\overline{\pi}_{s_n}(u_0)(d\phi) \\
&\leq \liminf_{n\to\infty}\frac{1}{s_n}\int_1^{s_n+1}\left(\int_{C^\gamma(\mathbf{T}^d)} e^{\kappa_2 g(\|\phi\|_{C^\gamma})^2}\pi_t(u_0)(d\phi)\right)dt \\
&\leq \sup_{1\leq t<\infty}\mathbb{E}\big[\exp\big(\kappa_2 g(\|u(t;u_0)\|_{C^\gamma})^2\big)\big] < \infty.
\end{aligned}$$

The conclusion of the statement follows from the (b) $\Rightarrow$ (a) part of Theorem 4. ▷

As for the $\Phi^4_{4-\nu}$ measure, the uniqueness of an invariant probability measure from the dynamics is a consequence of the fact that the transition semigroup of the Markovian dynamics on $C^\gamma(\mathbf{T}^d)$ has the strong Feller property, a fact proved in great generality by Hairer & Mattingly in [20], and the support theorem for the BPHZ model proved by Hairer & Schönbauer in [21]. The latter implies that the law of u on the interval $[0,T]$ has support in $C([0,T],C^\gamma(\mathbf{T}^d))$ given by all functions with values u_0 at time 0. (This is Theorem 1.12 in [21].) The reasoning in the proof of Theorem 1.13 in [21] applies and gives the uniqueness of an invariant measure. Therefore we obtain Theorem 2 on the $\Phi^4_{4-\nu}$ measure as a direct corollary of Theorem 8 and the polynomial estimates on the solution u to (1.7) proved by Chandra, Moinat & Weber in [8]. Consult [10] for a number of other examples where the reasoning of the proof of Theorem 8 can be applied and gives some transportation cost inequalities for some invariant probability measures of some Feller dynamics.

3 – Functional setting

We describe in this short section the functional setting involved in the sequel – see Section 2 of [1] and references therein for the details. Set

$$\mathcal{L} := \partial_1^2 - \big(1 - \partial_2^2 - \cdots - \partial_{1+d}^2\big)^2.$$

Its associated semigroup $(\mathcal{Q}_s := e^{s\mathcal{L}})_{s>0}$ gives a representation of the inverse of the heat operator $\partial_t - \Delta + 1$ by

$$\big(\partial_1 - \Delta + 1\big)^{-1} = -\int_0^\infty (\partial_1 + \Delta - 1)\, e^{s\mathcal{L}}\, ds.$$

Denote by Q_s the smooth integral kernel of $\mathcal{Q}_s$

$$(\mathcal{Q}_s f)(x) = \mathcal{Q}_t(x,f) = \int_{\mathbf{R}^{1+d}} Q_s(x-y)f(y)dy.$$

We define the family of weight functions $(w_a)_{a\geq 0}$ on $\mathbf{R}^{1+d}$ by

$$w_a(x) := \big(1+\|x\|_{\mathfrak{s}}\big)^{-a}.$$

For $a \geq 0$ and $p \in [1,\infty]$ we define the weighted L^p norm by

$$\|f\|_{L^p(w_a)} := \|fw_a\|_{L^p(\mathbf{R}^{1+d})}.$$

For any non-negative integer m we set $\mathcal{Q}_t^{(m)} = (-t\mathcal{L})^m \mathcal{Q}_t$. For every $r < 4m$ and $p,q \in [1,\infty]$ we define the ***Besov space*** $B^{r,\mathcal{Q}}_{p,q}(w_a)$ as the completion of $C(\mathbf{R}^{1+d}) \cap L^p(w_a)$ under the norm

$$\|f\|^{(m)}_{B^{r,\mathcal{Q}}_{p,q}(w_a)} := \|\mathcal{Q}_1 f\|_{L^p(w_a)} + \big\|t^{-r/4}\|\mathcal{Q}_t^{(m)} f\|_{L^p(w_a)}\big\|_{L^q((0,1];\frac{dt}{t})}.$$

The topological space $B^{r,\mathcal{Q}}_{p,q}(w_a)$ is defined independently to the choice of m as long as $m > r/4$, since the norms $\|\cdot\|^{(m_1)}_{B^{r,\mathcal{Q}}_{p,q}(w_a)}$ and $\|\cdot\|^{(m_2)}_{B^{r,\mathcal{Q}}_{p,q}(w_a)}$ are equivalent for $m_1, m_2 > r/4$. We set

$$H^{r,\mathcal{Q}}(w_a) := B^{r,\mathcal{Q}}_{2,2}(w_a), \qquad C^{r,\mathcal{Q}}(w_a) := B^{r,\mathcal{Q}}_{\infty,\infty}(w_a).$$

We define a family $(\widetilde{K}_s)_{0<s\leq 1}$ of 2-regularizing operators setting

$$\widetilde{K}_s = -(\partial_1 + \Delta - 1)e^{s(\partial_1^2 - (\Delta-1)^2)}.$$

For technical reasons we fix a compactly supported smooth function χ which is equal to 1 on a neighborhood of 0 and we consider the modified operators

$$K_t = \big(1-\chi(\partial)\big)\widetilde{K}_t,$$

where

$$\chi(\partial)f := \mathscr{F}^{-1}\big(\chi\mathscr{F}f\big)$$

is a Fourier multiplier operator defined by Fourier transform $\mathscr{F}$ and its inverse $\mathscr{F}^{-1}$. This modification ensures that

$$\int_{\mathbf{R}^d} x^k \partial^l K_t(x)dx = 0$$

for any $k,l \in \mathbf{N}^{1+d}$, where we denote the operator K_t and its integral kernel K_t by the same symbol. Finally, for any $f \in C(\mathbf{R}^{1+d}) \cap L^p(w_a)$ and $k \in \mathbf{N}^{1+d}$, we define

$$\partial^k \mathcal{K}(x,f) := \int_0^1 \int_{\mathbf{R}^{1+d}} \partial_x^k K_s(x-y) f(y)\, dyds. \tag{3.1}$$

The operator $\mathcal{K}$ coincides with $(\partial_1 - \Delta + 1)^{-1}$ modulo some regularizing operators.

4 – BPHZ model over a particular regularity-integrability structure

We will prove in Theorem 12 in Section 5 the local Hölder regularity estimate (2.2) of the BPHZ map.

The BPHZ model was constructed in [22] and [1] using two different functional settings and different sets of tools. We use here the apparatus of [1], as a variation on the arguments used therein happen to lead directly to (2.2).

A notion of regularity-integrability structure was introduced in [23] and [1] to implement an iterative construction of the BPHZ model involving the Malliavin derivative operator and a spectral gap assumption on the law of the noise. The index set of a regularity structure is a subset of $\mathbf{R}$; the index set of a regularity-integrability structure is a subset of $\mathbf{R}\times[1,+\infty]$. This is related to the fact that we quantify the regularity of some analytical objects associated with any symbol τ of the structure in some τ-dependent Besov space $B^{r(\tau)}_{i(\tau),\infty}$ rather than in a fixed function space.

We define a strict partial order on $\mathbf{R}\times[1,\infty]$ by setting

$$(r,i) < (s,j) \quad \overset{\text{def}}{\Longleftrightarrow} \quad \Big\{ r < s \text{ and } \frac{1}{i} \leq \frac{1}{j} \Big\}.$$

We recall from [23] and [1] that a ***regularity-integrability structure*** (A,T,G) consists of the following elements.

- The index set A is a subset of $\mathbf{R}\times[1,\infty]$ such that for every $(s,j)\in\mathbf{R}\times[1,\infty]$ the set $\{(r,i)\in A\,;\,(r,i)<(s,j)\}$ is finite.
- The vector space $T=\bigoplus_{\mathbf{a}\in A}T_{\mathbf{a}}$ is an algebraic sum of Banach spaces $(T_{\mathbf{a}},\|\cdot\|_{\mathbf{a}})$.
- The structure group G is a group of continuous linear operators on T such that one has for all $\Gamma\in G$ and $\mathbf{a}\in A$
$$(\Gamma-\mathrm{Id})(T_{\mathbf{a}})\subset\bigoplus_{\mathbf{b}\in A,\,\mathbf{b}<\mathbf{a}}T_{\mathbf{b}}.$$

We introduce in Section 4.1 a particular regularity-integrability structure. This section is dedicated to proving that one can define the BPHZ model over this extended regularity-integrability structure under the assumption that the law of the noise satisfies Assumption A. We assume from now on that the Hilbert space H involved in that assumption is the Hilbert space $H^{-s_0,\mathcal{Q}}(w_b)$, for some $s_0\in\mathbf{R}$ and $b\geq 0$ which will be specified in Theorem 10 below.

4.1 – A particular regularity-integrability structure.

We introduce in this section a regularity-integrability structure that differs from the regularity-integrability structure of [1] by the fact that

- we include some trees corresponding to all order Malliavin derivatives, while only first order derivatives are considered in [1],
- we introduce two copies of the H space in our type set
$$\mathfrak{L}=\big\{\mathbb{K},\Omega,\mathbb{H},\underline{\mathbb{H}}\big\}.$$
The type $\mathbb{H}$ will represent some elements of H estimated in some $B_{\infty,\infty}$-type Besov norm, while the type $\underline{\mathbb{H}}$ will represent some elements of H estimated in some $B_{p,\infty}$-type Besov norm, with p varying freely in the interval $[2,\infty]$.

We invite the reader familiar with the constructions of the trees from [6] to skip the next paragraph.

§1. Sets of trees – We denote by $\widetilde{\mathbf{T}}$ the set of all 4-tuples $(\tau,\mathfrak{t},\mathfrak{n},\mathfrak{e})$, with τ a non-planar rooted tree τ with vertex/node set N_τ and edge set E_τ, with $\mathfrak{t}:E_\tau\to\mathfrak{L}$ a label map, and with two decoration maps $\mathfrak{n}:N_\tau\to\mathbf{N}^{1+d}$ and $\mathfrak{e}:E_\tau\to\mathbf{N}^{1+d}$. We sometimes abuse notations and write τ for a decorated tree $(\tau,\mathfrak{t},\mathfrak{n},\mathfrak{e})$ if the context makes clear what the decoration is. A *leaf* in a decorated tree is an edge $e=(u,v)$ such that there is no outgoing edge from v. (Every edge is represented as an ordered pair (u,v) of two nodes, where u is nearer to the root than v.)

Denote by $\widetilde{T}$ the linear space spanned by $\widetilde{\mathbf{T}}$. The tree product of $\tau,\sigma\in\widetilde{\mathbf{T}}$, denoted by $\tau\sigma$, is obtained by identifying the roots of τ and σ in the disjoint union $\tau\sqcup\sigma$, whose decorations are inherited from those of τ and σ except that the $\mathfrak{n}$-decoration at the root of $\tau\sigma$ is the sum of $\mathfrak{n}$-decorations at the roots of τ and σ. We extend that tree product linearly to $\widetilde{T}$, which turns it into an algebra.

For each $(\tau,\mathfrak{t},\mathfrak{n},\mathfrak{e})\in\widetilde{\mathbf{T}}$ we define $\widetilde{\Delta}(\tau,\mathfrak{t},\mathfrak{n},\mathfrak{e})$ as the infinite sum
$$\sum_{\sigma}\sum_{\mathfrak{n}_\sigma,\mathfrak{e}_{\partial\sigma}}\frac{1}{\mathfrak{e}_{\partial\sigma}!}\binom{\mathfrak{n}}{\mathfrak{n}_\sigma}\big(\sigma,\mathfrak{t}|_{E_\sigma},\mathfrak{n}_\sigma+\pi\mathfrak{e}_{\partial\sigma},\mathfrak{e}|_{E_\sigma}\big)\otimes\Big(\tau/\sigma,\mathfrak{t}|_{E_\tau\setminus E_\sigma},[\mathfrak{n}-\mathfrak{n}_\sigma]_\sigma,\mathfrak{e}|_{E_\tau\setminus E_\sigma}+\mathfrak{e}_{\partial\sigma}\Big),\qquad(4.1)$$
where

- σ runs over all subtrees of τ which contain the root of τ. Then the quotient tree τ/σ is obtained by identifying all nodes of σ in the tree τ. The edge set of τ/σ is $E_\tau\setminus E_\sigma$.
- $\mathfrak{n}_\sigma$ runes over all maps $N_\sigma\to\mathbf{N}^{1+d}$ such that $\mathfrak{n}_\sigma(v)\leq\mathfrak{n}(v)$ for any $v\in N_\sigma$. The map
$$[\mathfrak{n}-\mathfrak{n}_\sigma]_\sigma:N_{\tau/\sigma}\to\mathbf{N}^{1+d}$$

is defined by $[\mathfrak{n}-\mathfrak{n}_\sigma]_\sigma(v)=\mathfrak{n}(v)$ for non-root $v\in N_{\tau/\sigma}$ and by $[\mathfrak{n}-\mathfrak{n}_\sigma]_\sigma(\varrho)=\sum_{v\in N_\sigma}(\mathfrak{n}(v)-\mathfrak{n}_\sigma(v))$ for the root ϱ of τ/σ. Moreover one sets

$$\binom{\mathfrak{n}}{\mathfrak{n}_\sigma}:=\prod_{v\in N_\sigma}\binom{\mathfrak{n}(v)}{\mathfrak{n}_\sigma(v)}.$$

– $\partial\sigma$ is the boundary of σ, that is, the set of all edges $(u,v)\in E_\tau$ such that $u\in N_\sigma$ and $v\in N_\tau\setminus N_\sigma$. $\mathfrak{e}_{\partial\sigma}$ runs over all maps $\partial\sigma\to\mathbf{N}^{1+d}$. For any $u\in N_\sigma$, $\pi\mathfrak{e}_{\partial\sigma}(u):=\sum_{(u,v)\in\partial\sigma}\mathfrak{e}_{\partial\sigma}((u,v))$. Moreover one sets

$$\mathfrak{e}_{\partial\sigma}!:=\prod_{e\in\partial\sigma}\mathfrak{e}_{\partial\sigma}(e)!.$$

A proper definition of the infinite sum (4.1) involves the notions of 'bigraded space' and tensor product of bigraded spaces. See Section 2.3 of [6] for the details. Since we only consider below some truncations of $\widetilde{\Delta}$ into some finite sums, we do not touch the details here. It was shown in Proposition 3.23 of [6] that $\widetilde{T}$, equipped with the tree product and the coproduct $\widetilde{\Delta}$, is a Hopf algebra (as a bigraded space).

The map $\widetilde{\Delta}$ admits a useful recursive characterization. A 4-tuple $(\bullet,\varnothing,k,\varnothing)$ made up of a single node with decoration $\mathfrak{n}(\bullet)=k\in\mathbf{N}^{1+d}$ is denoted by X^k. We use the notations $\mathbf{1}:=X^0$ and $X_j:=X^{e_j}$, where $e_j=(0,\dots,1,\dots,0)$ is the j-th vector of the canonical basis of $\mathbf{N}^{1+d}$. We write $\mathcal{I}_k^{\mathfrak{l}}(\tau)$ for the planted tree obtained from $\tau\in\widetilde{\mathbf{T}}$ by grafting it to a new root with $\mathfrak{n}$-decoration 0, along an edge with label $\mathfrak{l}\in\mathfrak{L}$ and $\mathfrak{e}$-decoration $k\in\mathbf{N}^{1+d}$; we consider $\mathcal{I}_k^{\mathfrak{l}}:\widetilde{T}\to\widetilde{T}$ as a linear map. Then $\widetilde{\Delta}:\widetilde{T}\to\widetilde{T}\otimes\widetilde{T}$ is characterized by the recursive formulae

$$\widetilde{\Delta}(\tau\sigma)=(\widetilde{\Delta}\tau)(\widetilde{\Delta}\sigma),\qquad \widetilde{\Delta}(X^k)=\sum_{l\le k}\binom{k}{l}X^l\otimes X^{k-l},$$

$$\widetilde{\Delta}(\mathcal{I}_k^{\mathfrak{l}}(\tau))=(\mathcal{I}_k^{\mathfrak{l}}\otimes\mathrm{Id})\widetilde{\Delta}\tau+\sum_{l\in\mathbf{N}^{1+d}}\frac{X^l}{l!}\otimes\mathcal{I}_{k+l}^{\mathfrak{l}}(\tau)\qquad(\forall\,\tau,\sigma).$$

See Proposition 4.17 of [6] for a proof. Similarly to [1], we do not consider the trees that have some $\mathbb{K}$-type leaves. Because of this, we actually consider the quotient Hopf algebra

$$T=\widetilde{T}/I,$$

where I is the Hopf ideal spanned by trees that have a $\mathbb{K}$-type leaf. We identify T with the subspace spanned by the set $\mathbf{T}$ of all trees without $\mathbb{K}$-type leaves. Then the tree product $T\otimes T\to T$ and the coproduct

$$\Delta:=(\pi\otimes\pi)\circ\widetilde{\Delta}\circ inj:T\to T\otimes T$$

are well-defined, where $inj:T\to\widetilde{T}$ is the natural injection map and $\pi:\widetilde{T}\to T$ is a canonical surjection.

Fix $s_0\ge-\frac{2+d}{2}$, $\beta_0\in(0,2)$, and $r_0<-\frac{2+d}{2}-s_0$ such that

$$r_0\notin\Big\{2\beta_0+2q_1+\sum_{j=2}^{1+d}q_j\,;\,r,q_1,\dots,q_{1+d}\in\mathbf{Q}\Big\}.$$

(The lower bound of s_0 is not essential: it excludes the case $r_0>0$ where renormalization is no longer necessary. The parameter β_0 is chosen as an arbitrary number slightly smaller than 2 because of the same technical reason as [1].) Recall the embeddings of function spaces over the $(1+d)$-dimensional space $\mathbf{R}^{1+d}$

$$H^{-s_0,\mathcal{Q}}(w_b)\hookrightarrow B_{2,\infty}^{r_0+\frac{2+d}{2},\mathcal{Q}}(w_a)\hookrightarrow B_{p,\infty}^{r_0+\frac{2+d}{p},\mathcal{Q}}(w_a)\hookrightarrow C^{r_0,\mathcal{Q}}(w_a)=\Omega,$$

for any $a\ge b\ge 0$. For any $\varepsilon\ge0$ and $p\in[2,\infty]$, we define the degree map $r_{\varepsilon,p}:\mathbf{T}\to\mathbf{R}$ by

$$r_{\varepsilon,p}(\mathbb{\Omega}) = r_{\varepsilon,p}(\mathbb{H}) := r_0 - \varepsilon, \qquad r_{\varepsilon,p}(\underline{\mathbb{H}}) := r_0 - \varepsilon + \frac{2+d}{p}, \qquad r_{\varepsilon,p}(\mathbb{K}) := \beta_0,$$
$$r_{\varepsilon,p}(\tau_{\mathfrak{e}}^{\mathfrak{n}}) := \sum_{v\in N_\tau} |\mathfrak{n}(v)|_{\mathfrak{s}} + \sum_{e\in E_\tau} \big(r_{\varepsilon,p}(\mathfrak{t}(e)) - |\mathfrak{e}(e)|_{\mathfrak{s}}\big).$$

The parameter ε is introduced for some technical purpose which will not be fundamental here. We refer to the work [1] for more on this point. For each $\varepsilon \geq 0$ and $p \in [2,\infty]$ we introduce the projection map $P^+_{\varepsilon,p}$ from T to the subalgebra $T^+_{\varepsilon,p}$ spanned by the symbols $X^k \prod_{i=1}^n \mathcal{I}_{k_i}^{\mathfrak{l}_i}(\tau_i)$ with $n \in \mathbf{N}$, $k, k_i \in \mathbf{N}^{1+d}$, $\mathfrak{l}_i \in \mathfrak{l}$ and $\tau_i \in \mathbf{T}$ such that $r_{\varepsilon,p}(\tau_i) + r_{\varepsilon,p}(\mathfrak{l}_i) > |k_i|_{\mathfrak{s}}$ for each i. We define

$$\Delta_{\varepsilon,p} := (\mathrm{Id} \otimes P^+_{\varepsilon,p})\Delta : T \to T \otimes T^+_{\varepsilon,p},$$
$$\Delta^+_{\varepsilon,p} := \big(P^+_{\varepsilon,p} \otimes P^+_{\varepsilon,p}\big)\Delta : T^+_{\varepsilon,p} \to T^+_{\varepsilon,p} \otimes T^+_{\varepsilon,p}.$$

The ranges are algebraic tensor products since $P^+_{\varepsilon,p}$ restricts the choices of $\mathfrak{e}_{\partial\sigma}$ in (4.1). By a similar argument to [18, 6] we can see that $T^+_{\varepsilon,p}$ is a Hopf algebra with coproduct $\Delta^+_{\varepsilon,p}$ and T has a right comodule structure with coaction $\Delta_{\varepsilon,p}$. Denote by $S^+_{\varepsilon,p}$ the antipode of $(T^+_{\varepsilon,p}, \Delta^+_{\varepsilon,p})$. The p-dependence of $\Delta_{\varepsilon,p}$ was described in an example (3.6) in [1].

The linear space T is too large for our purpose and we fix until the end of this section the subset $\mathbf{B} \subset \mathbf{T}$ of all trees $\tau \in \mathbf{T}$ with only $\mathbb{K}$ or $\mathbb{\Omega}$ type edges that strongly conform to a given complete subcritical rule and such that $r_{0,\infty}(\tau) < C_0$ for a fixed number C_0. (Its choice is dictated by each equation when we apply the setting of regularity structures to the study of any given subcritical singular stochastic PDE.) Moreover, for each $(\tau, \mathfrak{t}, \mathfrak{n}, \mathfrak{e}) \in \mathbf{B}$, we assume that if $e = (u,v) \in \mathfrak{t}^{-1}(\mathbb{\Omega})$ then e is a leaf, $\mathfrak{n}(v) = \mathfrak{e}(e) = 0$, and there are no other edges $e' = (u, v') \in \mathfrak{t}^{-1}(\mathbb{\Omega})$ with $v' \neq v$.

§2. A regularity-integrability structure – For each $(\tau, \mathfrak{t}, \mathfrak{n}, \mathfrak{e}) \in \mathbf{B}$ and disjoint subsets $V, \underline{V}$ of $\subset \mathfrak{t}^{-1}(\mathbb{\Omega})$, we denote by $D_{V,\underline{V}}(\tau, \mathfrak{t}, \mathfrak{n}, \mathfrak{e})$ the decorated tree $(\tau, \mathfrak{t}', \mathfrak{n}, \mathfrak{e})$ with the modified node decoration $\mathfrak{t}'$ defined by

$$\mathfrak{t}'(e) = \begin{cases} \mathbb{H} & (e \in V), \\ \underline{\mathbb{H}} & (e \in \underline{V}), \\ \mathfrak{t}(e) & (\text{otherwise}). \end{cases}$$

We set as well

$$\dot{\mathbf{B}} := \Big\{ D_{V,\varnothing}\tau \ ; \ \tau \in \mathbf{B},\ \varnothing \neq V \subset \mathfrak{t}^{-1}(\mathbb{\Omega}) \Big\},$$
$$\underline{\dot{\mathbf{B}}} := \Big\{ D_{V,\{v\}}\tau \ ; \ \tau \in \mathbf{B},\ V \subset \mathfrak{t}^{-1}(\mathbb{\Omega}),\ v \in \mathfrak{t}^{-1}(\mathbb{\Omega}) \setminus V \Big\},$$
$$\widetilde{\mathbf{B}} := \mathbf{B} \cup \dot{\mathbf{B}} \cup \underline{\dot{\mathbf{B}}},$$

and define the linear spaces

$$V := \mathrm{span}(\mathbf{B}), \qquad U := \mathrm{span}(\mathbf{B} \cup \dot{\mathbf{B}}), \qquad W := \mathrm{span}(\widetilde{\mathbf{B}}).$$

We set

$$i_p(\tau) = \begin{cases} \infty & \text{if } \tau \in \mathbf{B} \cup \dot{\mathbf{B}} \\ p & \text{if } \tau \in \underline{\dot{\mathbf{B}}}. \end{cases}$$

For each $(\varepsilon, p) \in [0,\infty) \times [2,\infty]$ we define V^+_ε, U^+_ε, and $W^+_{\varepsilon,p}$ as the subalgebras of $T^+_{\varepsilon,p}$ generated by the families of symbols

$$\mathbf{V}^+_\varepsilon := \{X_j\}_{j=1}^{1+d} \cup \{\mathcal{I}_k^{\mathbb{K}}(\tau)\}_{k\in\mathbf{N}^{1+d},\, \tau\in\mathbf{B}\setminus\{X^l\}_l,\, r_{\varepsilon,p}(\tau)+\beta_0>|k|_{\mathfrak{s}}},$$
$$\mathbf{U}^+_\varepsilon := \{X_j\}_{j=1}^{1+d} \cup \{\mathcal{I}_k^{\mathbb{K}}(\tau)\}_{k\in\mathbf{N}^{1+d},\, \tau\in(\mathbf{B}\cup\dot{\mathbf{B}})\setminus\{X^l\}_l,\, r_{\varepsilon,p}(\tau)+\beta_0>|k|_{\mathfrak{s}}},$$
$$\mathbf{W}^+_{\varepsilon,p} := \{X_j\}_{j=1}^{1+d} \cup \big\{\mathcal{I}_k^{\underline{\mathbb{H}}}(\bullet)\big\}_{k\in\mathbf{N}^{1+d},\, r_{\varepsilon,p}(\underline{\mathbb{H}})>|k|_{\mathfrak{s}}} \cup \big\{\mathcal{I}_k^{\mathbb{K}}(\tau)\big\}_{k\in\mathbf{N}^{1+d},\, \tau\in\widetilde{\mathbf{B}}\setminus\{X^l\}_l,\, r_{\varepsilon,p}(\tau)+\beta_0>|k|_{\mathfrak{s}}},$$

respectively. We write $\mathbf{V}_\varepsilon^+$ and $\mathbf{U}_\varepsilon^+$ rather than $\mathbf{V}_{\varepsilon,p}^+$ and $\mathbf{U}_{\varepsilon,p}^+$ since p has no influences on the trees in $\mathbf{B}\cup\dot{\mathbf{B}}$. (From its definition, the quantity $r_{\varepsilon,p}(\tau)$ varies with p only on trees with at least one $\underline{\mathbb{H}}$-type edge.)

The proof of the following statement is omitted since it is almost identical to the proof of Lemma 6 in [1].

9 – Lemma. *The set V_ε^+, resp. U_ε^+ and $W_{\varepsilon,p}^+$, is a sub-Hopf algebra of $T_{\varepsilon,p}^+$, and V, resp. U and W, is a right comodule over V_ε^+, resp. U_ε^+ and $W_{\varepsilon,p}^+$, with coaction $\Delta_{\varepsilon,p}$.*

The group $\mathsf{G}_{\varepsilon,p}^+$ of characters on the Hopf algebra $(W_{\varepsilon,p}^+,\Delta_{\varepsilon,p}^+)$ has a representation in $GL(W)$ where $\mathsf{g}\in\mathsf{G}_{\varepsilon,p}^+$ is mapped to $(\mathrm{id}\otimes\mathsf{g})\Delta_{\varepsilon,p}$. Denote by $\mathsf{G}_{\varepsilon,p}$ the image group. Lemma 9 ensures that V and U are stable under the action of $\mathsf{G}_{\varepsilon,p}$.

We finally define a grading on $Z\in\{V,U,W\}$ setting for any $\mathbf{a}\in\mathbf{R}\times[1,\infty]$ and $\mathbf{Z}\in\{\mathbf{B},\mathbf{B}\cup\dot{\mathbf{B}},\widetilde{\mathbf{B}}\}$

$$Z_{\mathbf{a}} := \mathrm{span}\Big\{\tau\in\mathbf{Z}\,;\ \big(r_{\varepsilon,p}(\tau),i_p(\tau)\big)=\mathbf{a}\Big\}.$$

With this grading, we define some regularity-integrability structures

$$\begin{aligned}\mathscr{V}_\varepsilon &:= (A_\varepsilon, V, \mathsf{G}_{\varepsilon,p}|_V),\\ \mathscr{U}_\varepsilon &:= (A_\varepsilon, U, \mathsf{G}_{\varepsilon,p}|_U),\\ \mathscr{W}_{\varepsilon,p} &:= (A_{\varepsilon,p}, W, \mathsf{G}_{\varepsilon,p}),\end{aligned}$$

where

$$A_\varepsilon := \big\{\big(r_{\varepsilon,\infty}(\tau),\infty\big)\,;\ \tau\in\mathbf{B}\cup\dot{\mathbf{B}}\big\}$$

and

$$A_{\varepsilon,p} := \big\{\big(r_{\varepsilon,p}(\tau),i_p(\tau)\big)\,;\ \tau\in\widetilde{\mathbf{B}}\big\}.$$

These regularity-integrability structures are described below in their ***concrete form***

$$\begin{aligned}\mathscr{V}_\varepsilon &= \big((V,\Delta_{\varepsilon,p}),(V_\varepsilon^+,\Delta_{\varepsilon,p}^+)\big),\\ \mathscr{U}_\varepsilon &= \big((U,\Delta_{\varepsilon,p}),(U_\varepsilon^+,\Delta_{\varepsilon,p}^+)\big),\\ \mathscr{W}_{\varepsilon,p} &= \big((W,\Delta_{\varepsilon,p}),(W_{\varepsilon,p}^+,\Delta_{\varepsilon,p}^+)\big).\end{aligned}$$

Such a representation is useful when we want to make the generators of the structure group explicit.

§3. Models on regularity-integrability structures – We recall from [1] some basic facts about models on regularity-integrability structures that are involved in the statement and the proof of Theorem 10 below. A reader familiar with regularity structures can skip this paragraph after looking at (4.2) and (4.3). Denoting by $P_{\mathbf{a}}:T\to T_{\mathbf{a}}$ the canonical projection, we set with a slight abuse of notations

$$\|\tau\|_{\mathbf{a}} := \|P_{\mathbf{a}}\tau\|_{\mathbf{a}}$$

for any $\tau\in T$ and $\mathbf{a}\in A$.

Models – Assume we are given a pair of maps $\mathsf{M}=(\mathsf{\Pi},\mathsf{g})$ such that

$$\mathsf{\Pi}: W\to C^{r_0,\mathcal{Q}}(w_a),\qquad \mathsf{g}:\mathbf{R}^{1+d}\to\mathsf{G}_{\varepsilon,p}^+$$

and $\mathsf{\Pi}$ is continuous and linear. The map $\mathsf{\Pi}$ is called an ***interpretation map***. For any $\tau\in W$ and $\mu\in W_{\varepsilon,p}^+$ set

$$\begin{aligned}\mathsf{\Pi}_x^{\varepsilon,p}(\tau) &:= \big(\mathsf{\Pi}\otimes\mathsf{g}_x^{-1}\big)\Delta_{\varepsilon,p}(\tau),\\ \mathsf{g}_{yx}^{\varepsilon,p}(\mu) &:= \big(\mathsf{g}_y\otimes\mathsf{g}_x^{-1}\big)\Delta_{\varepsilon,p}^+(\mu),\end{aligned}$$

where g_x^{-1} is an inverse of g_x in the group $\mathsf{G}_{\varepsilon,p}^+$. A pair of maps $\mathsf{M}=(\mathsf{\Pi},\mathsf{g})$ as above is called a ***model on*** $\mathscr{W}_{\varepsilon,p}$ (with weight w_a) if

$$\|\mathsf{\Pi}:\tau\|_{\varepsilon,p;w_a} := \sup_{0<t\le 1} t^{-r_{\varepsilon,p}(\tau)/4}\big\|\mathcal{Q}_t(x,\mathsf{\Pi}_x^{\varepsilon,p}(\tau))\big\|_{L_x^{i_p(\tau)}(w_a)} < \infty \tag{4.2}$$

for any $\tau \in \widetilde{\mathbf{B}}$ and

$$\|\mathsf{g}:\mu\|_{\varepsilon,p;w_a} := \|\mathsf{g}_x(\mu)\|_{L_x^{ip(\mu)}(w_a)} + \sup_{y\in\mathbf{R}^{1+d}\setminus\{0\}} \left(w_a(y) \frac{\left\|\mathsf{g}^{\varepsilon,p}_{(x+y)x}(\mu)\right\|_{L_x^{ip(\mu)}(w_a)}}{\|y\|_{\mathfrak{s}}^{r_{\varepsilon,p}(\mu)}} \right) < \infty \tag{4.3}$$

for any $\mu \in \mathbf{W}^+_{\varepsilon,p}$. For any models $\mathsf{M} = (\mathsf{\Pi}, \mathsf{g})$, we set

$$\|\mathsf{M}\|_{\mathbf{M}(\mathscr{W}_{\varepsilon,p})_{w_a}} := \max_{\tau\in\widetilde{\mathbf{B}}} \|\mathsf{\Pi}:\tau\|_{\varepsilon,p;w_a} + \max_{\mu\in\mathbf{W}^+_{\varepsilon,p}} \|\mathsf{g}:\mu\|_{\varepsilon,p;w_a}.$$

Moreover, we define a metric $\|\mathsf{M}_1:\mathsf{M}_2\|_{\mathbf{M}(\mathscr{W}_{\varepsilon,p})_{w_a}}$ between two models M_1 and M_2 by

$$\|\mathsf{M}_1:\mathsf{M}_2\|_{\mathbf{M}(\mathscr{W}_{\varepsilon,p})_{w_a}} := \max_{\tau\in\widetilde{\mathbf{B}}} \|\mathsf{\Pi}_1,\mathsf{\Pi}_2:\tau\|_{\varepsilon,p;w_a} + \max_{\mu\in\mathbf{W}^+_{\varepsilon,p}} \|\mathsf{g}_1,\mathsf{g}_2:\mu\|_{\varepsilon,p;w_a}.$$

The quantities in the right hand side are defined in the same way as (4.2) and (4.3) but with $(\mathsf{\Pi}_1)^{\varepsilon,p}_x(\tau) - (\mathsf{\Pi}_2)^{\varepsilon,p}_x(\tau)$, $(\mathsf{g}_1)_x(\mu) - (\mathsf{g}_2)_x(\mu)$ and $(\mathsf{g}_1)^{\varepsilon,p}_{yx}(\mu) - (\mathsf{g}_2)^{\varepsilon,p}_{yx}(\mu)$ in place of $\mathsf{\Pi}^{\varepsilon,p}_x(\tau)$, $\mathsf{g}_x(\mu)$ and $\mathsf{g}^{\varepsilon,p}_{yx}(\mu)$, respectively. Finally, we define the space $\mathbf{M}(\mathscr{W}_{\varepsilon,p})_{w_a}$ as the completion of collections of all models $\mathsf{M} = (\mathsf{\Pi}, \mathsf{g})$ such that $\|\mathsf{M}\|_{\mathbf{M}(\mathscr{W}_{\varepsilon,p})_{w_a}}$ is finite and $\mathsf{\Pi}\tau$ and $\mathsf{g}(\mu)$ are smooth for any τ and μ, under the metric $\|\cdot:\cdot\|_{\mathbf{M}(\mathscr{W}_{\varepsilon,p})_{w_a}}$. This definition ensures that $\mathbf{M}(\mathscr{W}_{\varepsilon,p})_{w_a}$ is a Polish space.

• It will also be useful for later purposes to introduce a ***homogeneous norm*** on the space of models on $\mathscr{U}_\varepsilon$. Denote by $n_\circ(\tau)$ the number of Ω-type edges in an arbitrary symbol τ. Set

$$\|\mathsf{\Pi}:\tau\|_{\varepsilon;w_a,\mathrm{hom}} := \left(\sup_{0<t\le 1} t^{-r_\varepsilon(\tau)/4} \left\| \mathcal{Q}_t(x, \mathsf{\Pi}^\varepsilon_x(\tau)) \right\|_{L^\infty_x(w_a)} \right)^{1/n_\circ(\tau)},$$

for all $\tau \in \mathbf{B} \cup \dot{\mathbf{B}}$, and

$$\|\mathsf{g}:\mu\|_{\varepsilon;w_a,\mathrm{hom}} := \left(\|\mathsf{g}_x(\mu)\|_{L^\infty_x(w_a)} + \sup_{y\in\mathbf{R}^{1+d}\setminus\{0\}} \left(w_a(y) \frac{\left\|\mathsf{g}^{\varepsilon}_{(x+y)x}(\mu)\right\|_{L^\infty_x(w_a)}}{\|y\|_{\mathfrak{s}}^{r_\varepsilon(\mu)}} \right) \right)^{1/n_\circ(\mu)}$$

for all $\mu \in \mathbf{U}^+_\varepsilon$. Note that we can drop off the letter p from these notations since p has no influences for objects defined on $\mathscr{U}_\varepsilon$. We set

$$\|\mathsf{M}\|_{\mathbf{M}_{\mathrm{hom}}(\mathscr{U}_\varepsilon)_{w_a}} := \max_{\tau\in\mathbf{B}\cup\dot{\mathbf{B}}} \|\mathsf{\Pi}:\tau\|_{\varepsilon;w_a,\mathrm{hom}} + \max_{\mu\in\mathbf{U}^+_\varepsilon} \|\mathsf{g}:\mu\|_{\varepsilon;w_a,\mathrm{hom}}$$

for any $\mathsf{M} \in \mathbf{M}(\mathscr{U}_\varepsilon)_{w_a}$. We also define similarly a metric $\|\mathsf{M}_1:\mathsf{M}_2\|_{\mathbf{M}_{\mathrm{hom}}(\mathscr{U}_\varepsilon)_{w_a}}$ on $\mathbf{M}(\mathscr{U}_\varepsilon)_{w_a}$ and a metric $\|\mathsf{M}_1:\mathsf{M}_2\|_{\mathbf{M}_{\mathrm{hom}}(\mathscr{V}_\varepsilon)_{w_a}}$ on $\mathbf{M}(\mathscr{V}_\varepsilon)_{w_a}$. The metrics $\|\cdot:\cdot\|_{\mathbf{M}(\mathscr{V}_\varepsilon)_{w_a}}$ and $\|\cdot:\cdot\|_{\mathbf{M}_{\mathrm{hom}}(\mathscr{V}_\varepsilon)_{w_a}}$ generate the same topology on $\mathbf{M}(\mathscr{V}_\varepsilon)_{w_a}$.

• For the sub-concrete regularity-integrability structure of the form

$$\mathscr{W}'_{\varepsilon,p} = \big((Z, \Delta_{\varepsilon,p}), (Z^+, \Delta^+_{\varepsilon,p})\big)$$

where $Z = \mathrm{span}(\mathbf{C})$, for some subset $\mathbf{C} \subset \widetilde{\mathbf{B}}$, and Z^+ is a subalgebra of $W^+_{\varepsilon,p}$ generated by some subset $\mathbf{C}^+$ of $\mathbf{W}^+_{\varepsilon,p}$, it is useful to define the restricted quantity

$$\|\mathsf{M}\|_{\mathbf{M}(\mathscr{W}'_{\varepsilon,p})_{w_a}} := \|\mathsf{\Pi}^{\varepsilon,p}:\mathbf{C}\|_{\varepsilon,p;w_a} + \|\mathsf{g}^{\varepsilon,p}:\mathbf{C}^+\|_{\varepsilon,p;w_a} := \max_{\tau\in\mathbf{C}} \|\mathsf{\Pi}^{\varepsilon,p}:\tau\|_{\varepsilon,p;w_a} + \max_{\mu\in\mathbf{C}^+} \|\mathsf{g}^{\varepsilon,p}:\mu\|_{\varepsilon,p;w_a}.$$

In particular the restriction of any $\mathsf{M} \in \mathbf{M}(\mathscr{W}_{\varepsilon,p})_{w_a}$ to $\mathscr{V}_\varepsilon$ is a model in the usual sense of [18].

A family of renormalized models – We choose $\mathbf{B}$ so that all the edges of an arbitrary $\tau \in \widetilde{\mathbf{B}}$ with labels other than $\mathbb{K}$ can be contracted into some labeled nodes. We write

$$\bigcirc = \mathcal{I}^{\Omega}_0(\bullet), \qquad \odot = \mathcal{I}^{\mathbb{H}}_0(\bullet), \qquad \underline{\odot} = \mathcal{I}^{\mathbb{H}}_0(\bullet).$$

Recall from (3.1) the definition of the operator $\mathcal{K}$ acting on functions over $\mathbf{R}^{1+d}$. Denote by $C^\infty_\star$ be the set of all smooth functions over $\mathbf{R}^{1+d}$ whose derivatives of all orders are in the class $\bigcup_{a>0} L^\infty(w_a)$. An interpretation map $\mathsf{\Pi}$ is said to be $\mathcal{K}$***-admissible*** if it satisfies

$$(\mathsf{\Pi} X^k)(x) = x^k, \qquad \mathsf{\Pi}(\mathcal{I}^{\mathbb{K}}_k\tau) = \partial^k \mathcal{K}(\cdot, \mathsf{\Pi}\tau)$$

for all $k \in \mathbf{N}^{1+d}$ and τ. For any $\zeta, j \in C^\infty_\star$ there is a unique *multiplicative* $\mathcal{K}$-admissible map $\Pi^{\zeta,j} : W \to C^\infty_\star$ such that

$$\Pi^{\zeta,j}(\bigcirc) = \zeta, \qquad \Pi^{\zeta,j}(\odot) = \Pi^{\zeta,j}(\underline{\odot}) = j.$$

We call a model $\mathsf{M}^{\zeta,j;\varepsilon,p} = (\Pi^{\zeta,j}, \mathsf{g}^{\zeta,j;\varepsilon,p})$ on $\mathscr{W}_{\varepsilon,p}$ (with weight w_a for any large $a > 0$) defined from the following recursive definition of $(\mathsf{g}^{\zeta,j;\varepsilon,p}_x)^{-1}$ the ***naive model*** associated to (ζ, j)

$$\begin{aligned}
\big(\mathsf{g}^{\zeta,j;\varepsilon,p}_x\big)^{-1}(X^k) &= (-x)^k, \\
\big(\mathsf{g}^{\zeta,j;\varepsilon,p}_x\big)^{-1}(\tau\sigma) &= (\mathsf{g}^{\zeta,j;\varepsilon,p}_x)^{-1}(\tau)(\mathsf{g}^{\zeta,j;\varepsilon,p}_x)^{-1}(\sigma), \\
\big(\mathsf{g}^{\zeta,j;\varepsilon,p}_x\big)^{-1}(\mathcal{I}^{\mathfrak{l}}_k(\sigma)) &= -\sum_{l\in\mathbf{N}^{1+d}} \frac{(-x)^l}{l!}\mathsf{f}^{\zeta,j;\varepsilon,p}_x(\mathcal{I}^{\mathfrak{l}}_{k+l}(\sigma)), \qquad (4.4)\\
\mathsf{f}^{\zeta,j;\varepsilon,p}_x(\tau) = \mathbf{1}_{r_{\varepsilon,p}(\tau)>0} &\times \begin{cases} \partial^k j(x) & (\tau = \mathcal{I}^{\mathbb{H}}_k(\bullet)), \\ \partial^k \mathcal{K}(x, \Pi^{\zeta,j;\varepsilon,p}_x\sigma) & (\tau = \mathcal{I}^{\mathbb{K}}_k(\sigma)). \end{cases}
\end{aligned}$$

Denote by $\mathbf{B}_-$ the set of all $\tau \in \mathbf{B}$ such that $r_{0,\infty}(\tau) < 0$ and $\tau \notin \{\bigcirc\} \cup \{\mathcal{I}^{\mathbb{K}}_k(\sigma)\}_{k\in\mathbf{N}^{1+d},\,\sigma\in\mathbf{B}}$. We extend any map $\chi : \mathbf{B}_- \to \mathbf{R}$ into a linear map $\chi : W \to \mathbf{R}$ by setting $\chi(\tau) = 0$ for $\tau \in \widetilde{\mathbf{B}} \setminus \mathbf{B}_-$, and define the linear map $R_\chi : W \to W$ from the formula

$$R_\chi(\tau) := \tau + (\chi \otimes \mathrm{id})\Delta\tau \qquad (\tau \in \widetilde{\mathbf{B}}). \tag{4.5}$$

Given such a map χ, we define $\widehat{M}^\chi : W \to W$ as the unique linear map fixing X^k, $\bigcirc$, $\odot$, and $\underline{\odot}$, such that $\widehat{M}^\chi$ is *multiplicative* and satisfies

$$\widehat{M}^\chi(\mathcal{I}^{\mathbb{K}}_k\tau) = \mathcal{I}^{\mathbb{K}}_k\big(\widehat{M}^\chi(R_\chi\tau)\big)$$

for all k and τ. We also define the linear map

$$M^\chi := \widehat{M}^\chi R_\chi.$$

Then for each naive model $\mathsf{M}^{\zeta,j;\varepsilon,p}$, we define the model

$$\mathsf{M}^{\zeta,j,\chi;\varepsilon,p} := \big(\Pi^{\zeta,j,\chi}, \mathsf{g}^{\zeta,j,\chi;\varepsilon,p}\big)$$

with

$$\Pi^{\zeta,j,\chi} := \Pi^{\zeta,j} M^\chi$$

and $\mathsf{g}^{\zeta,j,\chi;\varepsilon,p}$ defined by the version of the recursive rules (4.4) with $(\Pi^{\zeta,j,\chi}, \mathsf{g}^{\zeta,j,\chi;\varepsilon,p})$ in place of $(\Pi^{\zeta,j}, \mathsf{g}^{\zeta,j;\varepsilon,p})$. See Section 3 of Bruned's work [5] for the well-defined character of $\mathsf{M}^{\zeta,j,\chi;\varepsilon,p}$ as a model.

4.2 – A BPHZ convergence result on the regularity-integrability structure $\mathscr{W}_{\varepsilon,p}$.

We can now state and prove the BPHZ convergence theorem that constructs the BPHZ model which we will study in the next section. The statement and its proof are similar to the main result of [1]; we will only describe below the structure of the argument, emphasizing the points where some modifications are needed.

For a family of compactly supported smooth functions $\varrho_n \in C^\infty(\mathbf{R}^{1+d})$ that converge weakly to a Dirac mass at 0 as $n \in \mathbf{N}$ goes to ∞, denote by $*$ the convolution operator and define a random variable ξ_n on Ω setting for $\omega \in \Omega$ and $h \in H$

$$\xi_n(\omega) := \varrho_n * \omega \in \Omega, \qquad h_n := \varrho_n * h.$$

We define an $\mathbf{M}(\mathscr{W}_{\varepsilon,p})_{w_a}$-valued random variable setting for any deterministic map $\chi_n : \mathbf{B}_- \to \mathbf{R}$

$$\mathsf{M}^{\xi_n,h_n,\chi_n;\varepsilon,p}(\omega) := \mathsf{M}^{\xi_n(\omega),h_n,\chi_n;\varepsilon,p}.$$

The main result of [6] yields that, for any $\mathbb{P}$ such that the law of ξ_n is centered, transition invariant, has moments of all orders, we can find a unique $c_n : \mathbf{B}_- \to \mathbf{R}$ such that

$$\mathbb{E}\big[(\Pi^{\xi_n,h_n,\chi_n;0,\infty}\tau)(x)\big] = 0$$

for all $\tau \in \mathbf{B}$ of non-positive $r_{0,\infty}$-degree. We call the renormalized model $\mathsf{M}^{\xi_n,h_n,\chi_n;\varepsilon,p}$ on the regularity-integrability structure $\mathscr{W}_{\varepsilon,p}$ associated with this special choice of map χ_n the BPHZ model, and use for it the shorthand notation

$$\mathsf{M}^{n,h_n;\varepsilon,p} = \mathsf{M}^{\xi_n,h_n,\chi_n;\varepsilon,p}.$$

The following notation is involved in the next statement: $\mathsf{J}_p := \big\{\varepsilon \geq 0\,;\, \mu \in \mathbf{W}^+_{0,2},\ r_{\varepsilon,p}(\mu) = 0\big\}$, for an arbitrary $p \in [2,\infty]$.

10 – Theorem. *Pick $s_0 \geq -\frac{2+d}{2}$ and $r_0 < -\frac{2+d}{2} - s_0$.*

- *Let $\mathbb{P}$ be a Borel probability measure on $\Omega = C^{r_0,\mathcal{Q}}(w_a)$, for some $a > 0$, that satisfies Assumption A with $H = H^{-s_0,\mathcal{Q}}(w_b)$ for some $b \geq 0$.*
- *Assume that*

$$r_{0,\infty}(\tau) > -\frac{2+d}{2} \tag{4.6}$$

for all $\tau \in \mathbf{B} \setminus \{\bigcirc\}$.

Then there exist constants $\varepsilon_0 > 0$ and $\overline{a} > 0$ depending only on b, d, and $\mathbf{B}$ such that for any $a > \overline{a}$, $p \in [2,\infty], \varepsilon \in (0,\varepsilon_0) \setminus \mathsf{J}_p$ and $q \in [1,\infty)$, one has

$$\sup_{n\in\mathbf{N}} \mathbb{E}\Bigg[\sup_{\|h\|_H\leq 1} \|\mathsf{M}^{n,h_n;\varepsilon,p}\|^q_{\mathbf{M}(\mathscr{W}_{\varepsilon,p})_{w_a}}\Bigg] < \infty.$$

and there exists a $\mathbf{M}(\mathscr{W}_{\varepsilon,p})_{w_a}$-valued random variable $\mathsf{M}^{\infty,h;\varepsilon,p}$ such that

$$\lim_{n\to\infty} \mathbb{E}\Bigg[\sup_{\|h\|_H\leq 1} \|\mathsf{M}^{n,h_n;\varepsilon,p} : \mathsf{M}^{\infty,h;\varepsilon,p}\|^q_{\mathbf{M}(\mathscr{W}_{\varepsilon,p})_{w_a}}\Bigg] = 0.$$

The limit model $\mathsf{M}^{\infty,h;\varepsilon,p}$ does not depend on the approximation of the identity $(\varrho_n)_{n\geq 1}$.

The use of the spectral gap assumption in the proof of convergence of the BPHZ random models was pioneered by Linares, Otto, Tempelmayr & Tsatsoulis in [26]. We note that Theorem 10 is valid for spatially periodic noises ξ on $\mathbf{R} \times \mathbf{T}^d$ satisfying Assumption A with $H = H^{-s_0}(\mathbf{R} \times \mathbf{T}^d)$, a usual flat L^2-Sobolev space with parabolic scaling. Indeed, since $H^{-s_0}(\mathbf{R} \times \mathbf{T}^d) \hookrightarrow H^{-s_0,\mathcal{Q}}(w_b)$ holds for any $b > 0$, ξ satisfies the spectral gap inequality (1.2) with $H = H^{-s_0,\mathcal{Q}}(w_b)$. Although H is of the form $H^{-s_0,\mathcal{Q}}(w_0)$ in [1], the same proof works for the weighted H with only a slight modification of Lemma 15 of [1]. When ξ and h are spatially periodic, then the model $\mathsf{M}^{n,h;\varepsilon,p}$ is also spatially periodic in the sense that

$$\mathsf{\Pi}^{n,h;\varepsilon,p}_{x+e_i}(\tau)(\cdot + e_i) = \mathsf{\Pi}^{n,h;\varepsilon,p}_{x}(\tau)(\cdot), \qquad \mathsf{g}^{n,h;\varepsilon,p}_{(y+e_i)(x+e_i)} = \mathsf{g}^{n,h;\varepsilon,p}_{yx}$$

for any $2 \leq i \leq 1+d$. Recall that $e_i = (0,\dots,1,\dots,0)$ is the i-th vector of the canonical basis of $\mathbf{N}^{1+d}$.

Proof of Theorem 10 – The proof proceeds by induction, as illustrated in the picture below. We first decompose $\mathbf{B} = \{X^k\}_{|k|_\mathfrak{s}<C_0} \cup \mathbf{B}_\circ$, where all $\tau \in \mathbf{B}_\circ$ has at least one Ω-type edge. (The constant C_0 is involved in the description of the set $\mathbf{B}$.) Recall that $n_\circ(\tau)$ denotes the number of noise symbols $\bigcirc$ in an arbitrary symbol τ. We align the elements of $\mathbf{B}_\circ$ as

$$\mathbf{B}_\circ = \big\{\tau_1 \preceq \tau_2 \preceq \cdots \preceq \tau_N\big\},$$

where $\sigma \preceq \tau$ means that

$$\big(n_\circ(\sigma), |E_\sigma|, r_{0,\infty}(\sigma)\big) \leq \big(n_\circ(\tau), |E_\tau|, r_{0,\infty}(\tau)\big)$$

in the lexicographical order. We have in particular $\tau_1 = \bigcirc$. For $0 \leq i \leq N$ set

$$\mathbf{B}_i := \big\{\tau_j\,;\, j \leq i\big\} \cup \{X^k\}_{|k|_\mathfrak{s}<C_0}.$$

The following statement is proved as Lemma 14 in [1].

11 – Lemma. *For $1 \leq i \leq N$, set*

$$V_i := \mathrm{span}(\mathbf{B}_i), \qquad U_i := \mathrm{span}\Big(\mathbf{B}_{i-1} \cup \dot{\mathbf{B}}_i\Big), \qquad W_i := \mathrm{span}\Big(\mathbf{B}_{i-1} \cup \dot{\mathbf{B}}_{i-1} \cup \underline{\dot{\mathbf{B}}_i}\Big).$$

For each $(\varepsilon, p) \in [0, \varepsilon_0) \times [2, \infty]$ *we define* $V_{i,\varepsilon}^{+}$, $U_{i,\varepsilon}^{+}$, *and* $W_{i,\varepsilon,p}^{+}$ *as the subalgebras of* $T_{\varepsilon,p}^{+}$ *generated by the families of symbols*

$$\begin{aligned}
\mathbf{V}_{i,\varepsilon}^{+} &:= \{X_j\}_{j=1}^{1+d} \cup \{\mathcal{I}_k^{\mathbb{K}}(\tau)\}_{k\in\mathbf{N}^{1+d},\, \tau\in\mathbf{B}_i\setminus\{X^l\}_l,\, r_{\varepsilon,p}(\tau)+\beta_0>|k|_{\mathfrak{s}}},\\
\mathbf{U}_{i,\varepsilon}^{+} &:= \{X_j\}_{j=1}^{1+d} \cup \{\mathcal{I}_k^{\mathbb{K}}(\tau)\}_{k\in\mathbf{N}^{1+d},\, \tau\in(\mathbf{B}_i\cup\dot{\mathbf{B}}_i)\setminus\{X^l\}_l,\, r_{\varepsilon,p}(\tau)+\beta_0>|k|_{\mathfrak{s}}},\\
\mathbf{W}_{i,\varepsilon,p}^{+} &:= \{X_j\}_{j=1}^{1+d} \cup \big\{\mathcal{I}_k^{\underline{\mathbb{H}}}(\bullet)\big\}_{k\in\mathbf{N}^{1+d},\, r_{\varepsilon,p}(\underline{\mathbb{H}})>|k|_{\mathfrak{s}}} \cup \big\{\mathcal{I}_k^{\mathbb{K}}(\tau)\big\}_{k\in\mathbf{N}^{1+d},\, \tau\in\widetilde{\mathbf{B}}_i\setminus\{X^l\}_l,\, r_{\varepsilon,p}(\tau)+\beta_0>|k|_{\mathfrak{s}}},
\end{aligned}$$

respectively. Then each of

$$\begin{aligned}
\mathscr{V}_{i,\varepsilon} &:= \Big((V_i, \Delta_{\varepsilon,\infty}), \big(V_{i-1}^{+}, \Delta_{\varepsilon,\infty}^{+}\big)\Big),\\
\mathscr{U}_{i,\varepsilon} &:= \Big((U_i, \Delta_{\varepsilon,\infty}), \big(U_{i-1}^{+}, \Delta_{\varepsilon,\infty}^{+}\big)\Big),\\
\mathscr{W}_{i,\varepsilon,p} &:= \Big((W_i, \Delta_{\varepsilon,p}), \big(W_{i-1,\varepsilon,p}^{+}, \Delta_{\varepsilon,p}^{+}\big)\Big)
\end{aligned}$$

is a sub-regularity-integrability structure of $\mathscr{W}_{\varepsilon,p}$.

For any fixed $p \in [2, \infty]$ we write below $\mathbf{bd}(\mathscr{W}, i, p)$ to mean that

$$\sup_{n\in\mathbf{N}} \mathbb{E}\bigg[\sup_{\|h\|_H\le 1} \|\mathsf{M}^{n,h_n;\varepsilon,p}\|_{\mathbf{M}(\mathscr{W}_{i,\varepsilon,p})_{w_a}}^{q}\bigg] < \infty$$

for any small $\varepsilon > 0$ and large $a > 2 + d$ and $q \in [1, \infty)$. Similarly, we write $\mathbf{cv}(\mathscr{W}, i, p)$ to mean that

$$\lim_{n,m\to\infty} \mathbb{E}\bigg[\sup_{\|h\|_H\le 1} \|\mathsf{M}^{n,h_n;\varepsilon,p} : \mathsf{M}^{m,h_m;\varepsilon,p}\|_{\mathbf{M}(\mathscr{W}_{\varepsilon,p})_{w_a}}^{q}\bigg] = 0$$

for any small $\varepsilon > 0$ and large $a > 2 + d$ and all $q \in [1, \infty)$. We write

$$\mathbf{bd}(\mathscr{W}, i), \quad \text{respectively} \quad \mathbf{cv}(\mathscr{W}, i),$$

to mean that $\mathbf{bd}(\mathscr{W}, i, p)$, resp. $\mathbf{cv}(\mathscr{W}, i, p)$ holds for any $p \in [2, \infty]$. Moreover, we also write $\mathbf{bd}(\mathscr{V}, i)$, $\mathbf{bd}(\mathscr{U}, i)$, $\mathbf{cv}(\mathscr{V}, i)$ and $\mathbf{cv}(\mathscr{U}, i)$ to mean the analogue statements with $\mathscr{V}_{i,\varepsilon}$ or $\mathscr{U}_{i,\varepsilon}$ in place of $\mathscr{W}_{i,\varepsilon,p}$. These statements do not depend on p so we do not record this parameter in the notation. The following diagram summarizes the order of the proof. The dashed arrows mean that the corresponding steps are probabilistic. The solid arrows are independent of the probability measure.

$$\begin{tikzcd}
\mathbf{cv}(\mathscr{V},1) \arrow[ddr] & \mathbf{cv}(\mathscr{V},2) \arrow[ddr] & & {} \arrow[ddr] & \mathbf{cv}(\mathscr{V},i) \\
\mathbf{cv}(\mathscr{U},1) \arrow[u, dashed] \arrow[dr] & \mathbf{cv}(\mathscr{U},2) \arrow[u, dashed] \arrow[dr] & \cdots & {} \arrow[dr] & \mathbf{cv}(\mathscr{U},i) \arrow[u, dashed] & \cdots \\
\mathbf{cv}(\mathscr{W},1) \arrow[u] \arrow[r] & \mathbf{cv}(\mathscr{W},2) \arrow[u] \arrow[r] & \mathbf{cv}(\mathscr{W},3) & {} \arrow[r] & \mathbf{cv}(\mathscr{W},i) \arrow[u]
\end{tikzcd}$$

The details of each step are as follows.

(0) The initial case: $\mathbf{cv}(\mathscr{W}, 1)$ is reduced to the fact that h is an elements of $H = H^{-s_0,\mathcal{Q}}(w_b)$. This step is identical to the corresponding step in Section 4.2 of [1].

(1) $\mathbf{cv}(\mathscr{W}, i) \to \mathbf{cv}(\mathscr{U}, i)$: This step is trivial because $\mathbf{cv}(\mathscr{U}, i)$ can be seen as a particular case of $\mathbf{cv}(\mathscr{W}, i, \infty)$. Indeed, by the construction of models, we have

$$\mathsf{\Pi}_x^{n,h_n;\varepsilon,\infty}(\tau) = \mathsf{\Pi}_x^{n,h_n;\varepsilon,\infty}(T(\tau))$$

for any $\tau \in \underline{\dot{\mathbf{B}}}$, where $T : \underline{\dot{\mathbf{B}}} \to \dot{\mathbf{B}}$ is the linear map that replaces $\underline{\odot}$ by $\odot$.

(2) $\mathbf{cv}(\mathscr{U}, i) \to \mathbf{cv}(\mathscr{V}, i)$: This step is identical to the corresponding step in Section 4.3 of [1].
   - If $r_{0,\infty}(\tau_i) > 0$, we obtain the result from $\mathbf{cv}(\mathscr{V}, i - 1)$, which is included in $\mathbf{cv}(\mathscr{U}, i)$. This is a consequence of the reconstruction theorem. See Lemma 17 of [1].

 - Otherwise, we obtain a stochastic estimate from $\mathsf{cv}(\mathscr{U}, i)$ by using the spectral gap inequality. See Lemma 19 of [1].

(3) $\big\{\mathsf{cv}(\mathscr{W}, i), \mathsf{cv}(\mathscr{U}, i), \mathsf{cv}(\mathscr{V}, i)\big\} \to \mathsf{cv}(\mathscr{W}, i+1)$: This step is split into three parts.

(3-1) First, we obtain some n-uniform bounds and the convergence of the g-parts
$$\mathsf{g}^{n,h_n;\varepsilon,p}\big(\mathcal{I}_k^{\mathbb{K}}(\tau)\big)$$
for any $\tau \in \widetilde{\mathbf{B}}_i$ from the assumptions $\big\{\mathsf{cv}(\mathscr{W}, i), \mathsf{cv}(\mathscr{U}, i), \mathsf{cv}(\mathscr{V}, i)\big\}$; a consequence of the multilevel Schauder estimate in regularity-integrability structures. See Lemma 21 of [1].

(3-2) Then one shows some n-uniform bounds and the convergence of the Π-parts
$$\Pi^{n,h_n;\varepsilon,p}(\tau)$$
for any $\tau \in \underline{\dot{\mathbf{B}}_{i+1}}$. It is important for that purpose that the second assumption of Theorem 10 implies that $r_{0,2}(\tau) > 0$ for any $\tau \in \widetilde{\mathbf{B}} \setminus \{\bigcirc, \odot, \underline{\odot}\}$. Hence, if $p = 2$, the result is obtained as an application of the Reconstruction Theorem in regularity-integrability structures. See Lemma 22 of [1].

(3-3) To extend the result $\mathsf{cv}(\mathscr{W}, i+1, 2)$ into $\mathsf{cv}(\mathscr{W}, i+1, p)$ for all $p \in [2, +\infty]$, we need a comparison formula between $\Pi^{n,h_n;\varepsilon,2}$ and $\Pi^{n,h_n;\varepsilon,p}$ – see Lemma 8 of [1]. The proof of this lemma is exactly the same as in [1], except that the role of $\mathbb{H}$ is replaced here by $\underline{\mathbb{H}}$. Once that formula is established, the proof of $\mathsf{cv}(\mathscr{W}, i+1, p)$ follows the same steps as the proof of Lemma 24 in [1].

This concludes the description of the structure of the proof of Theorem 10. ▷

Despite some similar notations, the spaces $\mathscr{W}$ and $\mathscr{V}$ above are not the same as the corresponding spaces of [1]. The above statement is strictly stronger than the corresponding statement in [1], and Theorem 10 cannot be obtained as a direct corollary of Theorem 10 in [1].

5 – Transportation cost inequalities for the BPHZ models

We discuss in this section the limit BPHZ models on $\mathscr{V}_\varepsilon$ and $\mathscr{U}_\varepsilon$, so we can drop off from the notations the letter p in the sequel. When discussing the restriction of BPHZ models to $\mathscr{V}_\varepsilon$, we also write $\mathsf{M}^{n;\varepsilon}$ instead of $\mathsf{M}^{n,h_n;\varepsilon}$ since this restriction is independent of h. By Theorem 10, there exists a subsequence $\big(\mathsf{M}^{n_k,h_{n_k};\varepsilon}(\omega)\big)_{k\ge1}$ such that
$$\lim_{k\to\infty} \sup_{\|h\|_H\le1} \big\|\mathsf{M}^{n_k,h_{n_k};\varepsilon}(\omega) : \mathsf{M}^{\infty,h;\varepsilon}(\omega)\big\|_{\mathsf{M}(\mathscr{U}_\varepsilon)_{w_a}} = 0$$
almost surely. We denote by Ω_0 the set of all $\omega \in \Omega$ for which this convergence holds. By the multilinearity for h, if $\omega \in \Omega_0$, then $\big(\mathsf{M}^{n_k,h_{n_k};\varepsilon}(\omega)\big)_{k\ge1}$ converges for any $h \in H$. We define
$$\overline{\mathsf{M}}^{\infty,h;\varepsilon}(\omega) = \begin{cases} \lim_{k\to\infty} \mathsf{M}^{n_k,h_{n_k};\varepsilon}(\omega) & (\omega \in \Omega_0), \\ \mathbf{0} & (\omega \notin \Omega_0). \end{cases}$$
Recall that $\mathbf{0}$ is the zero model over $\mathscr{U}_\varepsilon$. On the other hand, we denote by Ω_1 the set of all $\omega \in \Omega$ for which the restrictions $\mathsf{M}^{n_k;\varepsilon}(\omega)$ of $\mathsf{M}^{n_k,h_{n_k};\varepsilon}(\omega)$ into a model on $\mathscr{V}_\varepsilon$ converge as k goes to ∞. We define
$$\widetilde{\mathsf{M}}^{\infty;\varepsilon}(\omega) = \begin{cases} \lim_{k\to\infty} \mathsf{M}^{n_k;\varepsilon}(\omega) & (\omega \in \Omega_1), \\ \mathbf{0} & (\omega \notin \Omega_1). \end{cases}$$
Note that $\Omega_0 \subset \Omega_1$ and the restriction of $\overline{\mathsf{M}}^{\infty,h;\varepsilon}(\omega)$ onto $\mathscr{V}_\varepsilon$ coincides with $\widetilde{\mathsf{M}}^{\infty;\varepsilon}(\omega)$ for any $\omega \in \Omega_0$. Set
$$m(\mathbf{B}) := \max\big\{n_\circ(\tau)\,;\, \tau \in \mathbf{B}\big\}.$$

12 – Theorem. *One has $\Omega_0 + h \subset \Omega_1$ for all $h \in H$, and*
$$\big\|\widetilde{\mathsf{M}}^{\infty;\varepsilon}(\omega + h) : \widetilde{\mathsf{M}}^{\infty;\varepsilon}(\omega)\big\|_{\mathsf{M}_{\mathrm{hom}}(\mathscr{V}_\varepsilon)_{w_a}} \le \mathbb{L}(\omega) \max\big(\|h\|_H, \|h\|_H^{1/m(\mathbf{B})}\big) \tag{5.1}$$

for all $\omega \in \Omega_0$ and $h \in H$, for a random variable $\mathbb{L}(\omega)$ proportional to

$$\left(1 + \sup_{\|h\|_H \leq 1} \big\|\overline{\mathsf{M}}^{\infty,h;\varepsilon}(\omega)\big\|_{\mathbf{M}_{\mathrm{hom}}(\mathscr{U}_\varepsilon)_{w_a}}\right)^{1-1/m(\mathbf{B})}$$

on Ω_0 and infinite elsewhere.

We note that the random variable $\mathbb{L}$ is in all the $L^p(\mathbb{P})$ spaces for $1 \leq p < \infty$, from Theorem 10. Hence the random variable $\widetilde{\mathsf{M}}^{\infty;\varepsilon} : \Omega \to \mathbf{M}(\mathscr{V}_\varepsilon)_{w_a}$ satisfies the inequality (1.4) for all $\omega_1 \in \Omega$ and $\omega_2 \in \Omega_0$. It follows from Proposition 1 that the law of $\widetilde{\mathsf{M}}^{\infty;\varepsilon}$ on the space $\mathbf{M}(\mathscr{V}_\varepsilon)_{w_a}$ satisfies for any $1 \leq \alpha < 2$ an $(\ell_\alpha, c^\alpha_{\mathbf{M}})$-transportation cost inequality (1.1) with ℓ_α the convex and continuous inverse function of the non-negative increasing continuous concave function $(t \in [0,\infty)) \mapsto a_\alpha^{-1}(t^{\alpha/(2r)} + t^{\alpha/2})$.

The inequality (5.1) is obtained from the following algebraic relations. For the sake of generality, we consider a deterministic model $\mathsf{M}^{\zeta,j,\chi}$ defined by some inputs $\zeta, j \in C^\infty_\star$ and some renormalization map $\chi : \mathbf{B}_- \to \mathbf{R}$. We suppress from the notations the exponents ε, p to simplify the notations here and in the remainder of this section; these exponents have some fixed values.

13 – Proposition. *For all x, y, for any $\tau \in \mathbf{B}$ one has*

$$\mathsf{\Pi}_x^{\zeta+j,0,\chi}(\tau) = \sum_{V \subset \bigcirc_\tau} \mathsf{\Pi}_x^{\zeta,j,\chi}(D_V\tau), \tag{5.2}$$

and for any $\sigma \in \mathbf{V}^+$ one has

$$\mathsf{g}_{yx}^{\zeta+j,0,\chi}(\sigma) = \sum_{V \subset \bigcirc_\sigma} \mathsf{g}_{yx}^{\zeta,j,\chi}(D_V\sigma), \tag{5.3}$$

where $\bigcirc_\tau := \mathfrak{t}^{-1}(\mathfrak{n})$ and $D_V\tau := D_{V,\varnothing}\tau$.

Proof of Theorem 12 from Proposition 13 – First note that, if $\omega \in \Omega_0$, the subsequence $\{\mathsf{M}^{\xi_{n_k}+h_{n_k},0,\chi_{n_k}}\}$ converges via the decompositions (5.2) and (5.3), so $\omega + h \in \Omega_1$.
As an application of (5.2) we have

$$\begin{aligned}
&\big\|\mathsf{\Pi}^{n;\varepsilon}(\omega+h), \mathsf{\Pi}^{n;\varepsilon}(\omega) : \tau\big\|_{\varepsilon,p;w_a,\mathrm{hom}} \\
&\quad = \left(\sup_{0<t\leq 1} t^{-r_\varepsilon(\tau)/4} \big\|\mathcal{Q}_t\big(x, \mathsf{\Pi}_x^{\xi_n+h_n,0,\chi_n}(\tau) - \mathsf{\Pi}_x^{\xi_n,0,\chi_n}(\tau)\big)\big\|_{L^\infty_x(w_a)}\right)^{\frac{1}{n_\circ(\tau)}} \\
&\quad \leq \sum_{\varnothing \neq V \subset \bigcirc_\tau} \left(\sup_{0<t\leq 1} t^{-r_\varepsilon(\tau)/4} \big\|\mathcal{Q}_t\big(x, \mathsf{\Pi}_x^{\xi_n,h_n,\chi_n}(D_V\tau)\big)\big\|_{L^\infty_x(w_a)}\right)^{\frac{1}{n_\circ(\tau)}} \\
&\quad \leq \sum_{\varnothing \neq V \subset \bigcirc_\tau} \left(\sup_{\|h\|_H \leq 1} \|\mathsf{\Pi}^{n,h_n;\varepsilon} : D_V\tau\|_{\varepsilon,p;w_a,\mathrm{hom}}\right)^{\frac{n_\circ(D_V\tau)}{n_\circ(\tau)}} \|h\|_H^{\frac{|V|}{n_\circ(\tau)}}.
\end{aligned}$$

In the last inequality, we use the fact that $\mathsf{\Pi}^{n,h;\varepsilon}(D_V\tau)$ is homogeneous with respect to h. Note that $|V| + n_\circ(D_V\tau) = n_\circ(\tau)$. Since $\frac{1}{m(\mathbf{B})} \leq \frac{|V|}{n_\circ(\tau)} \leq 1$ and $\frac{n_\circ(D_V\tau)}{n_\circ(\tau)} = 1 - \frac{|V|}{n_\circ(\tau)} \leq 1 - \frac{1}{m(\mathbf{B})}$, we have

$$\begin{aligned}
&\big\|\mathsf{\Pi}^{n;\varepsilon}(\omega+h), \mathsf{\Pi}^{n;\varepsilon}(\omega) : \tau\big\|_{\varepsilon,p;w_a,\mathrm{hom}} \\
&\quad \lesssim \left(1 + \sup_{\|h\|_H \leq 1} \|\mathsf{\Pi}^{n,h_n;\varepsilon}(\omega)\|_{\mathbf{M}_{\mathrm{hom}}(\mathscr{U}_\varepsilon)_{w_a}}\right)^{1-1/m(\mathbf{B})} \max\big(\|h\|_H, \|h\|_H^{1/m(\mathbf{B})}\big).
\end{aligned}$$

We obtain a similar estimate for $\big\|\mathsf{g}^{n;\varepsilon}(\omega+h), \mathsf{g}^{n;\varepsilon}(\omega) : \mu\big\|_{\varepsilon,p;w_a,\mathrm{hom}}$ starting from (5.3). Taking the (subsequential) limit, we have the inequality (5.1). ▷

We need the following two lemmas for the proof of Proposition 13. Recall from (4.5) the definition of the linear map R_χ.

14 – Lemma. *One has $R_\chi(D_V\tau) = D_V(R_\chi(\tau))$ for all $\tau \in \mathbf{B}$ and $V \subset \bigcirc_\tau$. In the right hand side, D_V acts only on $\sigma \in \mathbf{B}$ such that $V \subset \bigcirc_\sigma$. Otherwise $D_V\sigma = 0$.*

Proof – Let us use the notation $\Delta\tau = \sum_{\tau_1,\tau_2} \delta(\tau_1,\tau_2)\tau_1 \otimes \tau_2$ with a sum over decorated trees τ_1, τ_2 and with some coefficients $\delta(\tau_1,\tau_2)$. Since Δ acts on Ω-type edges and $\mathbb{H}$-type edges in the same way we can write

$$\Delta(D_V\tau) = \sum_{\tau_1,\tau_2} \delta(\tau_1,\tau_2)\big(D_{V\cap\bigcirc_{\tau_1}}(\tau_1)\big) \otimes \big(D_{V\cap\bigcirc_{\tau_2}}(\tau_2)\big).$$

Since χ vanishes on all the trees with at least one $\mathbb{H}$-type edge we have the consequence by applying $\chi \otimes \mathrm{Id}$ to the above formula. Indeed one has

$$(\chi \otimes \mathrm{Id})\Delta(D_V\tau) = \sum_{\tau_1,\tau_2; V\subset\bigcirc_{\tau_2}} \delta(\tau_1,\tau_2)\chi(\tau_1) \otimes \big(D_V(\tau_2)\big) = D_V(R_\chi(\tau)).$$

$\rhd$

The next formula appears in Section 5.5 of [1]; we state it here with a proof for completeness. Recall from (4.4) the definition of the character $\mathsf{f}_x^{\zeta,j,\chi}$. We extend $\mathsf{g}_x^{\zeta,j,\chi}$ and $\mathsf{f}_x^{\zeta,j,\chi}$ linearly and multiplicatively by setting $\mathsf{g}_x^{\zeta,j,\chi}(\eta) = \mathsf{f}_x^{\zeta,j,\chi}(\eta) = 0$ for any planted trees η such that $r_{\varepsilon,p}(\eta) \leq 0$.

15 – Lemma. *For any $\sigma \in \mathbf{B}$ and $k \in \mathbf{N}^{1+d}$ we have*

$$\mathsf{g}_{yx}^{\zeta,j,\chi}(\mathcal{I}_k^{\mathbb{K}}\sigma) = (\mathsf{f}_y^{\zeta,j,\chi} \otimes \mathsf{g}_{yx}^{\zeta,j,\chi})(\mathcal{I}_k^{\mathbb{K}} \otimes \mathrm{Id})(\Delta\sigma) - \sum_{l\in\mathbf{N}^{1+d}} \frac{(y-x)^l}{l!}\, \mathsf{f}_x^{\zeta,j,\chi}(\mathcal{I}_{k+l}^{\mathbb{K}}\sigma). \tag{5.4}$$

Proof – Let us write in this proof $\mathsf{g}_x = \mathsf{g}_x^{\zeta,j,\chi}$ and $\mathsf{f}_x = \mathsf{f}_x^{\zeta,j,\chi}$. By applying the operator $\mathsf{g}_x \otimes \mathsf{g}_x^{-1} \otimes \mathsf{g}_x$ to the identity

$$\sum_{m\in\mathbf{N}^{1+d}} \frac{X^m}{m!} \otimes \Delta^+(\mathcal{I}_{k+m}^{\mathbb{K}}\sigma) = \sum_{m\in\mathbf{N}^{1+d}} \frac{X^m}{m!} \otimes (\mathcal{I}_{k+m}^{\mathbb{K}} \otimes \mathrm{Id})\Delta\sigma + \sum_{m,l\in\mathbf{N}^{1+d}} \frac{X^m}{m!} \otimes \frac{X^l}{l!} \otimes \mathcal{I}_{k+l+m}^{\mathbb{K}}\sigma,$$

and using the fact $(\mathsf{g}_x^{-1} \otimes \mathsf{g}_x)\Delta^+(\mathcal{I}_{k+m}^{\mathbb{K}}\sigma) = 0$ and the binomial theorem, we have the explicit representation of g_x

$$\mathsf{g}_x(\mathcal{I}_k^{\mathbb{K}}\sigma) = (\mathsf{f}_x \otimes \mathsf{g}_x)(\mathcal{I}_k^{\mathbb{K}} \otimes \mathrm{Id})\Delta\sigma.$$

Then we have as a consequence

$$\begin{aligned}
(\mathsf{f}_y \otimes \mathsf{g}_{yx})(\mathcal{I}_k^{\mathbb{K}} \otimes \mathrm{id})\Delta\sigma &= (\mathsf{f}_y \otimes \mathsf{g}_y \otimes \mathsf{g}_x^{-1})(\mathcal{I}_k^{\mathbb{K}} \otimes \mathrm{Id} \otimes \mathrm{Id})(\mathrm{id} \otimes \Delta^+)\Delta\sigma \\
&= (\mathsf{f}_y \otimes \mathsf{g}_y \otimes \mathsf{g}_x^{-1})(\mathcal{I}_k^{\mathbb{K}} \otimes \mathrm{Id} \otimes \mathrm{Id})(\Delta \otimes \mathrm{id})\Delta\sigma \\
&= (\mathsf{g}_y \otimes \mathsf{g}_x^{-1})(\mathcal{I}_k^{\mathbb{K}} \otimes \mathrm{Id})\Delta\sigma \\
&= (\mathsf{g}_y \otimes \mathsf{g}_x^{-1})\left(\Delta^+\mathcal{I}_k^{\mathbb{K}}\sigma - \sum_{l\in\mathbf{N}^{1+d}} \frac{X^l}{l!} \otimes \mathcal{I}_{k+l}^{\mathbb{K}}\sigma\right) \\
&= \mathsf{g}_{yx}(\mathcal{I}_k^{\mathbb{K}}\sigma) + \sum_{l,m\in\mathbf{N}^{1+d}} \frac{y^l}{l!}\frac{(-x)^m}{m!}\, \mathsf{f}_x(\mathcal{I}_{k+l+m}^{\mathbb{K}}\sigma) \\
&= \mathsf{g}_{yx}(\mathcal{I}_k^{\mathbb{K}}\sigma) + \sum_{b\in\mathbf{N}^{1+d}} \frac{(y-x)^b}{b!}\, \mathsf{f}_x(\mathcal{I}_{k+b}^{\mathbb{K}}\sigma),
\end{aligned}$$

which concludes the proof of the lemma. $\rhd$

Proof of Proposition 13 – Recall from Proposition 3.15 in Bruned's work [5] that we can factorize the renormalized interpretation map as

$$\mathsf{\Pi}_x^{\zeta,j,\chi} = \widehat{\mathsf{\Pi}}_x^{\zeta,j,\chi} R_\chi,$$

where $\widehat{\Pi}_x^{\zeta,j,\chi}$ is the linear and multiplicative map defined by $\widehat{\Pi}_x^{\zeta,j,\chi} X^k = (\cdot - x)^k$ and

$$\widehat{\Pi}_x^{\zeta,j,\chi}(\mathcal{I}_k^{\mathbb{K}}\tau) = \partial^k \mathcal{K}(\cdot, \Pi_x^{\zeta,j,\chi}\tau) - \sum_{l\in\mathbf{N}^{1+d},\, |l|_{\mathfrak{s}}<r(\mathcal{I}_k^{\mathbb{K}}\tau)} \frac{(\cdot - x)^l}{l!} \partial^{k+l}\mathcal{K}\big(x, \Pi_x^{\zeta,j,\chi}\tau\big).$$

First we prove (5.2) and the auxiliary equality

$$\widehat{\Pi}^{\zeta+j,0,\chi}(\tau) = \sum_{V\subset\bigcirc_\tau} \widehat{\Pi}^{\zeta,j,\chi}(D_V\tau) \tag{5.5}$$

simultaneously. The proof is an induction on the size of τ. Both (5.2) and (5.5) are obvious for the initial cases $\tau \in \{\bigcirc, X^k\}$. Next let τ be a planted tree of the form $\mathcal{I}_k^{\mathbb{K}}(\sigma)$ with $\sigma \in \mathbf{B}$. If σ satisfies (5.2) then, by the definition of the operator $\widehat{\Pi}_x^{\zeta,j,\chi}$, we have

$$\begin{aligned}
\widehat{\Pi}_x^{\zeta+j,0,\chi}(\tau) &= \partial^k\mathcal{K}\big(\cdot, \Pi_x^{\zeta+j,0,\chi}(\sigma)\big) - \sum_{|l|_{\mathfrak{s}}<r(\tau)} \frac{(\cdot - x)^l}{l!}\partial^{k+l}\mathcal{K}\big(x, \Pi_x^{\zeta+j,0,\chi}(\sigma)\big)\\
&= \sum_{V\subset\bigcirc_\sigma}\bigg(\partial^k\mathcal{K}\big(\cdot, \Pi_x^{\zeta,j,\chi}(D_V\sigma)\big) - \sum_{|l|_{\mathfrak{s}}<r(\tau)} \frac{(\cdot - x)^l}{l!}\partial^{k+l}\mathcal{K}\big(x, \Pi_x^{\zeta,j,\chi}(D_V\sigma)\big)\bigg)\\
&= \sum_{V\subset\bigcirc_\sigma} \widehat{\Pi}_x^{\zeta,j,\chi}(\mathcal{I}_k^{\mathbb{K}}(D_V\sigma)).
\end{aligned}$$

In the last equality we use the fact that $D_V\sigma$ is a linear combination of decorated trees with degree $r(\tau)$. Since $\bigcirc_\sigma = \bigcirc_\tau$ and $\mathcal{I}_k^{\mathbb{K}} D_V\sigma = D_V\mathcal{I}_k^{\mathbb{K}}\sigma$, we obtain that τ satisfies (5.5). For any non-planted trees $\tau = \prod_{i=0}^N \eta_i$, we also obtain (5.5) by the multiplicativity of $\widehat{\Pi}_x^{\zeta,j,\chi}$. Finally, by using the result of Lemma 14, we see that

$$\begin{aligned}
\Pi^{\zeta+j,0,\chi}(\tau) = \widehat{\Pi}^{\zeta+j,0,\chi}(R_\chi\tau) &= \sum_{V\subset\bigcirc_\tau}\widehat{\Pi}^{\zeta,j,\chi}(D_V(R_\chi\tau))\\
&= \sum_{V\subset\bigcirc_\tau}\widehat{\Pi}^{\zeta,j,\chi}(R_\chi(D_V\tau)) = \sum_{V\subset\bigcirc_\tau}\Pi^{\zeta,j,\chi}(D_V\tau).
\end{aligned}$$

Finally, we prove (5.3) by an induction on the size of trees. By multicativity, it is sufficient to consider the planted tree $\sigma = \mathcal{I}_k^{\mathbb{K}}(\eta)$. Note that $\mathsf{f}_x^{\zeta,j,\chi}$ satisfies from (5.2) the formula

$$\mathsf{f}_x^{\zeta+j,0,\chi}(\tau) = \sum_{V\subset\bigcirc_\tau}\mathsf{f}_x^{\zeta,j,\chi}(D_V\tau).$$

By applying (5.4) of Lemma 15 to $\Delta(\eta) =: \sum_{\eta_1,\eta_2}\delta(\eta_1,\eta_2)\eta_1\otimes\eta_2$, for some decorated trees η_1,η_2 and coefficients $\delta(\eta_1,\eta_2)$, we have

$$\begin{aligned}
\mathsf{g}_{yx}^{\zeta+j,0,\chi}(\mathcal{I}_k^{\mathbb{K}}\eta) &= \sum_{\eta_1,\eta_2}\delta(\eta_1,\eta_2)\mathsf{f}_y^{\zeta+j,0,\chi}(\mathcal{I}_k^{\mathbb{K}}\eta_1)\,\mathsf{g}_{yx}^{\zeta+j,0,\chi}(\eta_2) - \sum_{l\in\mathbf{N}^{1+d}}\frac{(y-x)^l}{l!}\,\mathsf{f}_x^{\zeta+j,0,\chi}(\mathcal{I}_{k+l}^{\mathbb{K}}\eta)\\
&= \sum_{\eta_1,\eta_2}\delta(\eta_1,\eta_2)\sum_{V_1\subset\bigcirc_{\eta_1},V_2\subset\bigcirc_{\eta_2}}\mathsf{f}_y^{\zeta,j,\chi}(\mathcal{I}_k^{\mathbb{K}}D_{V_1}\eta_1)\,\mathsf{g}_{yx}^{\zeta,j,\chi}(D_{V_2}\eta_2)\\
&\quad - \sum_{l\in\mathbf{N}^{1+d}}\frac{(y-x)^l}{l!}\sum_{V\subset\bigcirc_\eta}\mathsf{f}_x^{\zeta,j,\chi}\big(\mathcal{I}_{k+l}^{\mathbb{K}}(D_V\eta)\big)\\
&= \sum_{V\subset\bigcirc_\eta}\bigg(\sum_{\eta_1,\eta_2}\delta(\eta_1,\eta_2)\,\mathsf{f}_y^{\zeta,j,\chi}(\mathcal{I}_k^{\mathbb{K}}D_{V\cap\bigcirc_{\eta_1}}\eta_1)\,\mathsf{g}_{yx}^{\zeta,j,\chi}(D_{V\cap\bigcirc_{\eta_2}}\eta_2)\\
&\quad - \sum_{l\in\mathbf{N}^{1+d}}\frac{(y-x)^l}{l!}\,\mathsf{f}_x^{\zeta,j,\chi}\big(\mathcal{I}_{k+l}^{\mathbb{K}}(D_V\eta)\big)\bigg)\\
&= \sum_{V\subset\bigcirc_\eta}\bigg((\mathsf{f}_y^{\zeta,j,\chi}\otimes\mathsf{g}_{yx}^{\zeta,j,\chi})(\mathcal{I}_k^{\mathbb{K}}\otimes\mathrm{id})\Delta(D_V\eta) - \sum_{l\in\mathbf{N}^{1+d}}\frac{(y-x)^l}{l!}\mathsf{f}_x^{\zeta,j,\chi}\big(\mathcal{I}_{k+l}^{\mathbb{K}}(D_V\eta)\big)\bigg)
\end{aligned}$$

$$= \sum_{V \subset \bigcirc_\eta} \mathsf{g}_{yx}^{\zeta,j,\chi}(\mathcal{I}_k^{\mathbb{K}} \eta).$$

This concludes the proof of Proposition 13. ▷

We note that Schönbauder [32] also studied the extended BPHZ model including Malliavin derivatives of all orders, to show Malliavin differentiability of solutions to singular stochastic PDEs. In [32], building up the convergence of the BPHZ model over $\mathscr{V}_\varepsilon$ from [7], a lift map from $h \in H$ to the extended BPHZ model $\widetilde{\mathsf{M}}^{\infty,h;\varepsilon}$ over $\mathscr{U}_\varepsilon$ is constructed. In contrast, in our proof, BPHZ models over $\mathscr{V}_\varepsilon$ and $\mathscr{U}_\varepsilon$ are constructed simultaneously by the self-contained proof. Moreover, while only Gaussian noises are considered in [32], our proof is valid for all noises satisfying transportation cost inequalities. While our result (5.1) implies only Hölder continuity for $\|h\|_H \le 1$, we can also obtain a locally Lipschitz continuity

$$\big\|\widetilde{\mathsf{M}}^{\infty;\varepsilon}(\omega+h) : \widetilde{\mathsf{M}}^{\infty;\varepsilon}(\omega)\big\|_{\mathbf{M}(\mathscr{V}_\varepsilon)_{w_a}} \le \mathbb{L}(\omega) \max\big(\|h\|_H, \|h\|_H^{m(\mathbf{B})}\big) \tag{5.6}$$

for the inhomogeneous metric, by slightly modifying the above proof, which is consistent with Theorem 7 of [32].

A – A relation between transportation cost and spectral gap inequalities

Denote by $\iota : H \to \Omega$ the continuous injection of H into Ω.

16 – Lemma. *Under Assumption A one has for all small enough $\lambda > 0$*

$$\int_\Omega e^{\lambda\|\omega\|^2}\,\mathbb{P}(d\omega) < \infty.$$

Proof – Following the arguments of the proofs of Proposition 1.6 and Theorem 1.7 in Gozlan & Léonard's review [17], we have the existence of some positive finite constants a_1, a_2 such that one has

$$\mathbb{P}(F - m_F > r) \le a_1 e^{-a_2 r^2} \tag{A.1}$$

for all $r > 0$ and any 1-Lipschitz function $F : \Omega \to \mathbf{R}$ with respect to the extended distance d_H, with m_F is the median of F. One gets the conclusion by applying the inequality (A.1) to the function $F(\omega) = \|\omega\|_\Omega / \|\iota\|_{H\to\Omega}$. ▷

17 – Theorem. *A probability measure satisfying Assumption A also satisfies the spectral gap inequality* (1.2) *with $a_1 = 1/(2a_\circ)$.*

Proof – Let μ be a Borel probability measure on Ω satisfying Assumption A, and let F be a cylindrical function of the form $F(\omega) = f\big(\varphi_1(\omega), \dots, \varphi_n(\omega)\big)$ with $\varphi_1, \dots, \varphi_n \in \Omega^*$ and at most polynomially growing $f \in C^\infty(\mathbf{R}^n)$. By applying the Gram–Schmidt process, we may assume that $\{\iota^*(\varphi_k)\}_{k=1}^n \subset H^*$ is an orthonormal system in H^*. Then note that the projection map $p_n : \Omega \to \mathbf{R}^n$ defined by $p_n(\omega) = (\varphi_1(\omega), \dots, \varphi_n(\omega))$ is a contraction with respect to the extended distance $d_H = \sqrt{c_H}$ on Ω and the Euclidean distance d_n on $\mathbf{R}^n$ as

$$d_n\big(p_n(\omega+\iota(h)), p_n(\omega)\big)^2 = \sum_{k=1}^n \big|\varphi_k(\iota(h))\big|^2 = \sum_{k=1}^n \big|\iota^*(\varphi_k)(h)\big|^2 \le \|h\|_H^2 = d_H(\omega+\iota(h), \omega)^2.$$

Let $\mu_n = \mu \circ p_n^{-1}$ be the image measure of μ on $\mathbf{R}^n$ by the application p_n. For any probability measure $\nu_n(dx) = g_n(x)\mu_n(dx)$ on $\mathbf{R}^n$ with $g_n \ge 0$, define a probability measure $\rho_n(d\omega) := g_n(p_n(\omega))\mu(d\omega)$ on Ω. Then since $\frac{d\rho_n}{d\mu}(\omega) = \frac{d\nu_n}{d\mu_n}(p_n(\omega))$, we have $\mathcal{H}(\nu_n|\mu_n) = \mathcal{H}(\rho_n|\mu)$. Moreover, for any coupling Θ_n of ρ_n and μ, by setting $\theta_n := \Theta_n \circ (p_n \times p_n)^{-1}$ we have

$$\begin{aligned}
\mathcal{T}_{d_n^2}(\nu_n, \mu_n) \le \int_{\mathbf{R}^n \times \mathbf{R}^n} d_n(x,y)^2 \theta_n(dxdy) &= \int_{\Omega\times\Omega} d_n(p_n(\omega_1), p_n(\omega_2))^2 \Theta_n(d\omega_1 d\omega_2) \\
&\le \int_{\Omega\times\Omega} d_H(\omega_1,\omega_2)^2\, \Theta_n(d\omega_1 d\omega_2),
\end{aligned}$$

which implies $\mathcal{T}_{d_n^2}(\nu_n,\mu_n) \le \mathcal{T}_{d_H^2}(\rho_n,\mu)$. Hence μ_n satisfies a $(a_\circ\mathrm{Id}, d_n^2)$-transportation cost inequality. Then by Otto and Villani's theorem [29] – see also [17, Proposition 8.11]), the probability μ_n satisfies the spectral gap inequality

$$2a_\circ\mathbb{E}_{\mu_n}[|f-\mathbb{E}_{\mu_n}[f]|^2] \le \mathbb{E}_{\mu_n}[|\nabla f|^2],$$

where $\nabla f = (\partial_1 f, \cdots, \partial_n f)$ is the gradient vector. We note that we have $\mathbb{E}_{\mu_n}[|f-\mathbb{E}_{\mu_n}[f]|^2] = \mathbb{E}_\mu[|F-\mathbb{E}_\mu[F]|^2]$ by definition of μ_n. By taking a dual system $\{e_k\}_{k=1}^n \subset H$ of $\{\iota^*(\varphi_k)\} \subset H^*$ such that $\iota^*(\varphi_k)(e_\ell) = \mathbf{1}_{k=\ell}$, we have

$$|\nabla f|^2(p_n(\omega)) = \sum_{k=1}^n |\partial_k f(p_n(\omega))|^2 = \sum_{k=1}^n |d_\omega F(e_k)|^2 \le \|d_\omega F\|_{H^*}^2.$$

This shows the spectral gap inequality $2a_\circ\mathbb{E}_\mu[|F-\mathbb{E}_\mu[F]|^2] \le \mathbb{E}_\mu[\|d_\omega F\|_{H^*}^2]$.

We actually use the spectral gap inequality for functionals of the form $\mathcal{Q}_t(x, \Pi_x^{n;\varepsilon}\tau)$. These functionals are not cylindrical functions but they have the form

$$F(\omega) = \int_{\mathbf{R}^n} f(\varphi_x(\omega))\theta(dx) \tag{A.2}$$

for a continuous bounded function $x \mapsto \varphi_x$ taking values in the space of bounded linear operators $\Omega \to \mathbf{R}^n$, for an at most polynomially growing function $f \in C^\infty(\mathbf{R}^n)$ and a finite signed Borel measure θ on $\mathbf{R}^n$. Functions $\Pi_x^{n;\varepsilon}\tau$ are generated from regularized functions $\varrho*\omega$ with some smooth mollifier ϱ and by applying the three kinds of operations inductively; (1) linear combinations, (2) multiplications, and (3) integral operator $\mathcal{K}$. Therefore, functionals $\mathcal{Q}_t(x, \Pi_x^{n;\varepsilon}\tau)$ can be written in the form (A.2).

Let F be of the form (A.2). By the bounded continuity of the map $x \mapsto f(\varphi_x(\omega))$, there is a finite partition $\{A_{m,k}\}_{k=1}^{N_m}$ of $\mathbf{R}^n$ and a set of reference points $\{a_{m,k} \in A_{m,k}\}_{k=1}^{N_m}$ for each $m \in \mathbf{N}$ such that

$$F_m(\omega) := \sum_{k=1}^{N_m} f(\varphi_{a_{m,k}}(\omega))\theta(A_{m,k})$$

converges to $F(\omega)$ as m goes to ∞ for all $\omega \in \Omega$. Moreover, $d_\omega F_m(h)$ also converges to $d_\omega F(h)$ uniformly over $\|h\|_H \le 1$ for all $\omega \in \Omega$. Since F_m is a cylindrical function as discussed before, it satisfies the inequality

$$2a_\circ\mathbb{E}_\mu[|F_m-\mathbb{E}_\mu[F_m]|^2] \le \mathbb{E}_\mu[\|d_\omega F_m\|_{H^*}^2] = \mathbb{E}\bigg[\sup_{\|h\|_H\le 1} |d_\omega F_m(h)|^2\bigg].$$

By the pointwise convergence and the uniform integrabilities of F_m and $d_\omega F_m$ given by Lemma 16, we can take the limit $m \to \infty$ in the above inequality and obtain the conclusion. $\rhd$

References

[1] Bailleul I. and Hoshino M., *Random models on regularity-integrability structures.* arXiv:2310.10202.

[2] Bailleul I. and Hoshino M., *A tourist's guide to to regularity structures and singular stochastic PDEs.* doi:10.4171/EMSS/87. To appear in EMS Surv. Math. Sci. (2025+).

[3] Bauerschmidt R. and Bodineau T., *Log-Sobolev inequality for the continuum sine-Gordon model.* Comm. Pure Appl. Math. **74**(10):2064–2113, (2021).

[4] Bauerschmidt R. and Dagallier B., *Log-Sobolev inequality for the Φ_3^4 and Φ_3^4.* Comm. Pure Appl. Math. **77**(5):2579–2612, (2024).

[5] Bruned Y., *Recursive formulae in regularity structures.* Stoch PDE: Anal Comp **6**:525–564 (2018).

[6] Bruned Y., Hairer M. and Zambotti L., *Algebraic renormalisation of regularity structures.* Invent. Math. **215**(3):1039–1156 (2019).

[7] Chandra A. and Hairer M., *An analytic BPHZ theorem for Regularity Structures.* arXiv:1612.08138 (2016).

[8] Chandra A., Moinat A. and Weber H., *A Priori Bounds for the Equation in the Full Sub-critical Regime.* Arch Rational Mech Anal **247**, 48 (2023).

[9] Chandra A., de Lima Feltes G. and Weber H., *A priori bounds for 2-d generalised Parabolic Anderson Model.* arXiv:2402.05544 (2024).

[10] Chevyrev I. and Gubinelli M., *Large field problem in coercive singular PDEs.* arXiv:2510.20716 (2025).

[11] Dai Y. and Li R.N., *Transportation Inequalities for Stochastic Heat Equation with Rough Dependence in Space.* Acta. Math. Sin.-English Ser. **38**:2019–2038 (2022).

[12] Djellout H., Guillin A. and Wu L., *Transportation cost-information inequalities and applications to random dynamical systems and diffusions.* Ann. Probab. **32**:2702–2732 (2004).

[13] Esquivel S. and Weber H., *A priori bounds for the dynamic fractional Φ^4 model on $\mathbb{T}^3$ in the full subcritical regime.* arXiv:2411.16536 (2024).

[14] Feyel D. and Üstünel A. S., *Measure transport on Wiener space and the Girsanov theorem.* Comptes Rendus Mathematique **334**:1025–1028 (2002).

[15] Gasteratos I. and Jacquier A., *Transportation-cost inequalities for non-linear Gaussian functionals.* arXiv:2310.05750 (2023).

[16] Gozlan N., *Integral criteria for transportation cost inequalities.* Elect. Comm. in Probab. **11**:64–77 (2006).

[17] Gozlan N. and Léonard C., *Transport inequalities. A survey.* Markov Proc. Rel. Fields **16**:635–736 (2010).

[18] Hairer M., *A theory of regularity structures*, Invent. Math. **198**:269–504 (2014).

[19] Hairer M., *Renormalisation in the presence of variance blowup.* Ann. Probab. **53**(5):1958–1985 (2025).

[20] Hairer M. and Mattingly, J., *The strong Feller property for singular stochastic PDEs.* Ann. Inst. H. Poincaré **54**(3):1314–1340 (2018).

[21] Hairer M. and Schönbauer P., *The support of singular stochastic partial differential equations.* Forum Math. Pi **10**(e1):1–127 (2022).

[22] Hairer M. and Steele R., *The* BPHZ *Theorem for Regularity Structures via the Spectral Gap Inequality.* Arch. Rat. Mech. Anal. **248**, 9 (2024).

[23] Hoshino M., *A semigroup approach to reconstruction theorems and multilevel Schauder estimates.* Annales Henri Lebesgue, **8**:151–180, (2025).

[24] Khoshnevisan D. and Sarantsev A., *Talagrand concentration inequalities for stochastic partial differential equations.* Stoch. PDE: Anal. Comp. **7**:679–698 (2019).

[25] Li R. and Wang X., *Talagrand's transportation inequality for SPDEs with locally monotone drifts.* Statistics & Probability Letters **204**:109945 (2024).

[26] Linares P., Otto F., Tempelmayr M. and Tsatsoulis P., *A diagram-free approach to the stochastic estimates in regularity structures.* Invent. math. **237**:1469–1565, (2024).

[27] Lyons T.J., *Differential equations driven by rough signals.* Rev. Mat. Iberoamericana **14**(2):215–310 (1998).

[28] Marton K., *A simple proof of the blowing-up lemma.* IEEE Trans. Inform. Theory, **32**(3):445–446 (1986).

[29] Otto F. and Villani C., *Generalization of an inequality by Talagrand and links with the logarithmic Sobolev inequality.* J. Funct. Anal. **173**:361–400 (2000).

[30] Riedel S., *Transportation-cost inequalities for diffusions driven by Gaussian processes.* Electronic Journal of Probability, **22** (2017), pp. 1–26.

[31] Saussereau B., *Transportation inequalities for stochastic differential equations driven by a fractional Brownian motion.* Bernoulli, (2012), pp. 1–23.

[32] Schönbauer, P., *Malliavin calculus and densities for singular stochastic partial differential equations.* Probab. Theory Related Fields **186** (2023), 643–713.

[33] Shang S. and Wang R., *Transportation Inequalities Under Uniform Metric for a Stochastic Heat Equation Driven by Time-White and Space-Colored Noise.* Acta Appl Math **170**, 81-97 (2020).

[34] Shang S. and Zhang T., *Talagrand concentration inequalities for stochastic heat-type equationsunder uniform distance.* Electron. J. Probab., **24**:(129), 1-15, (2019).

[35] Shen H., Zhu, R. and Zhu Z., *Global well-posedness for* 2d *generalized parabolic Anderson model via paracontrolled calculus.* Stoch. PDEs: Anal. and Comput. (2025)

[36] Talagrand M., *Transportation cost for Gaussian and other product measures.* Geometric & Functional Analysis **6** (1996), pp. 587–600.

[37] Wu L. and Zhang Z., *Talagrand's T_2-transportation inequality and log-Sobolev inequality for dissipative SPDEs and applications to reaction-diffusion equations.* Chinese Ann. Math. Ser. B, **27**:(3), 243-262, 2006.

• Bailleul I. – Univ Brest, CNRS, LMBA - UMR 6205, F-29238 Brest, France.
E-mail: ismael.bailleul@univ-brest.fr

• Hoshino M. – Department of Mathematics, Institute of Science Tokyo, Japan.
E-mail: hoshino@math.sci.isct.ac.jp

• Takano R. – The University of Osaka, Japan.
E-mail: rtakano@sigmath.es.osaka-u.ac.jp