Diffusive limit for 3-dimensional KPZ equation. (1) PDE estimates.

Jacques Magnen\textsuperscript{a} and Jérémie Unterberger\textsuperscript{b}

\textsuperscript{a}Centre de Physique Théorique\textsuperscript{1} Ecole Polytechnique, 91128 Palaiseau Cedex, France jacques.magnen@cpht.polytechnique.fr

\textsuperscript{b}Institut Elie Cartan\textsuperscript{2} Université de Lorraine, B.P. 239, F – 54506 Vandœuvre-lès-Nancy Cedex, France jeremie.unterberger@univ-lorraine.fr

We study in the present series of articles the Kardar-Parisi-Zhang (KPZ) equation

$$\partial_th(t, x) = \nu \Delta h(t, x) + \lambda |\nabla h(t, x)|^2 + \sqrt{D} \eta(t, x), \quad x \in \mathbb{R}^d$$

in $d \geq 3$ dimensions in the perturbative regime, i.e. for $\lambda > 0$ small enough and an initial condition $h_0 = h(t = 0, \cdot)$ such that $\lambda |\nabla h_0| = o(1)$. The forcing term $\eta$ in the right-hand side is a regularized white noise. We prove a large-scale diffusive limit for the solution, in particular a heat-kernel behaviour for the covariance in a parabolic scaling. The proof is generally based on perturbative estimates obtained by a multi-scale cluster expansion, and a rigorous implementation of K. Wilson’s renormalization group scheme; it extends to equations in the KPZ universality class,

$$\partial_t h(t, x) = \nu \Delta h(t, x) + \lambda V(\nabla h(t, x)) + \sqrt{D} \eta(t, x), \quad x \in \mathbb{R}^d$$

for a large class of isotropic, convex deposition rates $V \geq 0$. Our expansion is meant to be a rigorous substitute for the response field formalism whenever the latter makes predictions. An important part of the proof relies however more specifically on a priori bounds for the solutions of the KPZ equation, coupled with large deviation estimates.

The present article is dedicated to PDE estimates. Our results extend those previously obtained for the noiseless equation. We prove in particular a comparison principle for sub- and supersolutions of the KPZ equation in an adapted functional space containing unbounded functions, which may be of independent interest for the study of viscous Hamilton-Jacobi equations in general. The comparison to the linear heat equation through a Cole-Hopf transform is an essential ingredient in the proofs, and our results are accordingly valid only for a function $V$ with at most quadratic growth at infinity.

**Keywords:** KPZ equation, viscous Hamilton-Jacobi equations, maximum principle, response formalism, constructive field theory, renormalization, cluster expansions.

**Mathematics Subject Classification (2010):** 35B50, 35B51, 35D40, 35K55, 35R60, 35Q82, 60H15, 81T08, 81T16, 81T18, 82C41.

\textsuperscript{1}Laboratoire associé au CNRS UMR 7644

\textsuperscript{2}Laboratoire associé au CNRS UMR 7502
0 General introduction

The KPZ equation \[^{[40]}\] is a stochastic partial differential equation describing the growth by normal deposition of an interface in \((d + 1)\) space dimensions, see e.g. \([6, 17]\). By definition the time evolution of the height \(h(t, x), x \in \mathbb{R}^d\), is given by

\[
\frac{\partial}{\partial t} h(t, x) = \nu \Delta h(t, x) + \lambda \left( \sqrt{1 + |\nabla h(t, x)|^2} - 1 \right) + \sqrt{D} \eta(t, x), \quad x \in \mathbb{R}^d
\]  

(0.1)

where \(\eta(t, x)\) is some (possibly regularized) white noise. The gradient \(|\nabla h|\) (the slope of the interface) is assumed to remain throughout small so that the evolution makes physically sense, precluding e.g. any overhang, so that the non-linear term \(\sqrt{1 + |\nabla h(t, x)|^2} - 1 \approx \frac{1}{2} |\nabla h(t, x)|^2\) is essentially quadratic;
using this approximation gives the most common form of this equation in the literature. Following these preliminary remarks, we shall call KPZ equation any equation of the type

$$\partial_t h(t, x) = \nu \Delta h(t, x) + \lambda V(\nabla h(t, x)) + \sqrt{D} \eta(t, x), \quad x \in \mathbb{R}^d$$ (0.2)

where the deposition rate $V$ is isotropic and convex (hence $V(\nabla h(t, x)) = a + b|\nabla h(t, x)|^2 + \ldots$ around 0, with $b \geq 0$). The interest is generally in the large-scale limit of this equation, for $t$ large. A well-known naive rescaling argument gives some ideas about the dependence on the dimension of this limit. Namely, the linearized equation, a stochastic heat equation which is a particular instance of Ornstein-Uhlenbeck process,

$$\partial_t \phi(t, x) = \nu \Delta \phi(t, x) + \sqrt{D} \eta(t, x), \quad (t, x) \in \mathbb{R}_+ \times \mathbb{R}^d$$ (0.3)

is invariant under the rescaling $\phi(t, x) \mapsto \phi^\varepsilon(t, x) := \varepsilon^{-d_\phi} \phi(\varepsilon^{-1} t, \varepsilon^{-\frac{d}{2}} x)$, where

$$d_\phi := \frac{1}{2} (d - 1)$$ (0.4)

is the scaling dimension of the field $\phi$ (or rather half the scaling dimension, in the physicists’ convention); we used here the equality in distribution, $\eta(\varepsilon^{-1} t, \varepsilon^{-\frac{d}{2}} x) \overset{(d)}{=} \varepsilon^{2(1 + \frac{d}{2})} \eta(t, x)$. Assuming that $\phi$ is a solution of the KPZ equation instead (say, with quadratic deposition rate $|\nabla \phi|^2$) yields after rescaling

$$\partial_t \phi^\varepsilon(t, x) = \nu \Delta \phi^\varepsilon(t, x) + \varepsilon^\varepsilon \lambda \frac{1}{2} |\nabla \phi^\varepsilon(t, x)|^2 + \sqrt{D} \eta^\varepsilon(t, x),$$ (0.5)

where $\eta^\varepsilon \overset{(d)}{=} \eta$. For $d > 2$, the scaling exponent $d_\phi$ is $> 0$, and the non-linear term scaling coefficient, $\varepsilon^\varepsilon$, vanishes in the limit $\varepsilon \to 0$; in other terms, the KPZ equation is sub-critical at large scales in $\geq 3$ dimensions and believed to behave like the corresponding linearized equation up to a redefinition (called renormalization) of the diffusion constant $\nu$ and of the noise strength $D$. Let $M > 1$ (hereafter called scale parameter) be some fixed arbitrary parameter (e.g. $M = 2$). More precisely, according to the general scheme due to K. Wilson [69, 70], the fluctuations of the solution field at time scale of order $\varepsilon^{-1} \approx M^j$ and space scale of order $\varepsilon^{-\frac{d}{2}} \approx M^{j/2}$ should be approximately governed for $j$ large by a linearized equation with scale-dependent coefficients $\nu^j, D^j (j \geq 0)$, themselves solutions of a certain complicated but explicit discrete dynamical system. We prove this diffusive limit in the present series of articles following Wilson’s scheme, thus establishing on firm mathematical ground old predictions of physicists. The space dimension $d$ does not really matter as long as $d \geq 3$, and we shall refer to this equation as KPZ$_3$.

Let us emphasize the differences with the one-dimensional KPZ$_1$ equation – traditionally considered with a quadratic term $|\nabla h(t, x)|^2$ in the right-hand side making it partially integrable [24, 56, 47] –, which is perhaps more familiar to the mathematical community due to the striking progress made by several groups in the study of that equation in the last years, see e.g. [24, 20, 61, 65]. Recent work [34], using arguments coming from PDE’s and rough paths, shows that KPZ$_1$ is a.s. well-defined for short time; showing existence beyond some short random time is possible only through the Cole-Hopf transformation $h \mapsto \exp(h)$ which turns it into a linear equation. Note that $h$ represents the free energy of directed polymers in random environment [20]. On the other hand (reversing the previous scaling argument) it seems impossible to solve KPZ$_3$ for short time (short scale) unless $\eta$ is regularized; the most natural regularization here (due to the origin of this problem in real-world statistical physics) would be to use a simple lattice cut-off. However (the choice of regularization scheme
being essentially irrelevant) we choose here for convenience a regularization by space convolution with a smooth ’bump’ function.

What we consider here is the opposite regime, namely the large-scale limit of KPZ. The large-scale limit of KPZ\(_1\) is an extremely interesting, partially understood object: it is a strongly-coupled theory, solved by means of sophisticated methods coming from mathematical physics and integrable systems (Bethe Ansatz, Tracy-Widom distribution of the largest eigenvalue of random matrices, free fermions and determinant al processes, matrix formulations... see e.g. [20] for a detailed review). Here we consider instead KPZ\(_3\), and prove a rather trivial limit, furthermore in a non-physical dimension since a three-dimensional interface requires a four-dimensional space (note that KPZ\(_2\) is believed by perturbative QFT arguments to be strongly coupled at large scales [7,17] and its large-scale limit is not at all understood). As in dimension 1, the exponential of the solution represents the free energy of directed polymers in random environment when \(V\) is quadratic; several authors discussed the associated polymer measure, see e.g. [33] [18] [19]; the paper [19] proves that the regime we consider (\(d \geq 3, \lambda \ll 1\)) is a weak disorder regime for the polymer measure. Their results are not unrelated to ours – we also use a random path representation –, but the analytic and probabilistic methods (martingale theory viz. multi-scale analysis), the object (study of the polymer measure viz. law of the KPZ field) and the range of applicability (limited to the quadratic case in the above cited papers) differ widely.

We believe that the interest of our result lies in the wide scope of applicability of our perturbation methods. Many dynamical problems in statistical physics, modeled as interacting particle systems, or as parabolic SPDEs heuristically derived by some mesoscopic limit, have been turned into a functional integral form using the so-called response (or Martin-Siggia-Rose) formalism and studied by using standard perturbative expansions originated from quantum field theory (QFT); for reviews see e.g. [17] or [3]. Despite the lack of mathematical rigour, this formalism yields a correct description of the qualitative behaviour of such dynamical problems in the large scale limit. Here we provide an efficient way of making this approach rigorous.

Let us here comment on the specific difficulties of dynamical problems from a QFT viewpoint, taking KPZ\(_3\) with quadratic non-linearity as an example but meaning to be general. The MSR formalism – which may be viewed as a careless Girsanov transform in infinite dimension – consists in rewriting formally the solution \(h\) of the KPZ equation as the first marginal of the measure \(\int \text{d}x (L_0(h,h)(t,x)+L_{\text{int}}(h,h)(t,x))\), where \(L_0(h,h)(t,x) = \int \text{d}h(t,x)(\partial_t h(t,x) - \nu \Delta h(t,x) + \frac{\lambda}{2} h^2(t,x))\) is the free (quadratic) part, and \(L_{\text{int}}(h,h)(t,x) = \int \text{d}h(t,x)\Delta h(t,x)\) is the interaction term. The ”infinite-dimensional Lebesgue measures” ”\(\text{d}h(t,x)\)” do not exist, but the product \(\text{d}h \text{d}h^{-1} L_0\) can be given a sense (using some regularization procedure) as a Gaussian measure. The problem is, it is a complex-valued Gaussian measure. The Feynman diagram perturbative approach consists in expanding \(e^{-\int L_{\text{int}}\text{d}h(t,x)}\) into a series and making a clever resummation of some truncation of it; it is non rigorous since it yields \(n\)-point functions in terms of an asymptotic expansion in the coupling parameter \(\lambda\) which is divergent in all interesting cases. Constructive approaches developed by mathematical physicists from the 70es have developed sophisticated, systematic truncation methods making it possible to control the error terms. The partial resummations are interpreted in the manner of K. Wilson as a scale-by-scale renormalization of the parameters \(\nu, \Delta, \lambda\) of the Lagrangian \(L_0 + L_{\text{int}}\). However, these approaches are for the moment essentially limited to equilibrium statistical physics models for which the Lagrangian is real and bounded below by some non-degenerate quadratic form in the fields. Here the absolute value of the measure involves a degenerate weight \(e^{-\frac{1}{2} \int \text{d}x \text{d}x h(t,x)^2}\) from which \(h\) is absent. Less systematic ap-
proaches have been carried out successfully in diverse cases where only the propagator (here, $e^{\nu \Delta}$) needs to be renormalized and the state space is finite-dimensional (and thus, despite the similarities of the issues, not connected to the response formalism): diffusive limit of random walks in random environments [15], proof of the KAM theorem [14]. One should also cite the article by Bricmont, Kupiainen and Lin [16] which gives long-time asymptotics of sub-critical non-linear deterministic PDE’s by using successive rescalings; the methods of the latter article have inspired analysts working on the blow-up of non-linear PDE’s, and also the authors of the present article, although the successive re-elaborations necessitated by the incorporation of the noise term have made the influences difficult to recognize.

The reader will find a description of the results and of the main steps of the proof in the three upcoming articles, (2) Generalized PDE estimates through Hamilton-Jacobi-Bellman formalism; (3) The multi-scale expansion. We shall mainly be concerned in this first part with PDE estimates for the noiseless or homogeneous KPZ equation, and some preliminary applications to estimates for the noisy equation.

Here is a brief outline of the article. Section 2 is concerned with bounded solutions of the homogeneous KPZ equation, relying in particular on the comparison principle for non-linear parabolic PDE’s. Some results are derived with little effort from those already existing in the literature; on the other hand, the bounds on the gradient and on the higher derivatives of the solution, see Theorems 2.1 and 2.2 may not be found elsewhere. The really original material starts in section 3 with the search for solutions in new spaces of unbounded functions with good averaging properties, called hereafter $\mathcal{H}_\alpha^1$ or $W_{\text{loc}}^{1,\infty}$ (see sections 3.2 and 3.3). Since we aim at proving that the solutions of the KPZ equation are essentially obtained from those of the linearized equation by applying perturbation theory, such spaces should contain the stationary solution $\phi$ of the Ornstein-Uhlenbeck process (0.3), and actually they do, as follows from a non-standard large-deviation theory for random variables with heavy-tailed distributions (see Appendix B). Note that $\phi$ is space-translation invariant in law, hence unbounded. We prove in particular a comparison principle (see Theorem 3.1) for sub- and supersolutions of the homogeneous KPZ equation in these spaces, implying existence and unicity for viscosity solutions, which are proved to be classical. The study of the full inhomogeneous equation (0.2) is much more difficult and will be carried out only starting from the third article in the series. However, it can be done without using perturbation theory provided one uses a scale cut-off, i.e. one selects the fluctuations of the solution on time ranges of order $M^j$ and space ranges of order $M^{j/2}$ for some integer $j \geq 0$. Such cut-offs must be implemented in some coherent way both on the propagator $G = (\partial_t - \nu \Delta)^{-1}$ and on the noise $\eta$. We show in section 4 how to solve uniquely and bound along essentially the same lines as in section 3 some $j$-scale infra-red cut-off inhomogeneous equation with a force term $g$ of scale $j$, satisfying Assumption 4.2. Finally sections 5 and 6 are appendices containing multi-scale decompositions of $G$ and $\eta$ and large-deviation estimates, implying the applicability of the general arguments developed in section 4 to the case of the noisy KPZ equation (0.2).

1 Model and notations

We consider throughout the present article either the homogeneous (or noiseless) equation

$$\partial_t h = \nu \Delta h + \lambda V(\nabla h)$$

(1.1)
where $\lambda > 0$ is a fixed, arbitrary constant, or the infra-red cut-off, inhomogeneous equation,
\[
\partial_t h = \nu \Delta h - \varepsilon h + V(\nabla h) + g,
\] (1.2)
where the constant $\varepsilon = M^{-j}$ ($M > 1$, $j \geq 0$) is an infra-red cut-off of scale $j$. The assumptions on $V$ are the following:

**Assumption 1.1** The deposition rate $V$ satisfies the following assumptions,

1. $V$ is $C^2$;
2. $V$ is isotropic, i.e. $V(\nabla h)$ is a function of $y = |\nabla h|$; by abuse of notation we shall consider $V$ either as a function of $\nabla h$ or of $y$;
3. $V$ is convex;
4. $V(0) = 0$ and $V(y) \geq 0$ for all $y \geq 0$;
5. (quadratic growth at infinity) $V(y) \leq y^2$ for all $y \geq 0$.

It follows immediately from Assumptions (1), (2) and (4) that $V(y) = O(y^2)$ near $y = 0$. Assumption (5) is thus equivalent (up to a redefinition of the constant $\lambda$) to requiring that $V$ has quadratic growth at infinity.

As for the force term $g$, it is assumed to be regular enough and have good averaging properties, depending on the cut-off scale $j \approx -\log \varepsilon$; the regularized white noise $\eta$ (as shown in Appendix A) satisfies these properties (see sections 4 and 5).

Assumption (3) is a key assumption to get a time decay of the gradient of the solution, and is also used in the proof of the comparison theorem for unbounded solutions; Proposition $2.3$ (ii), (iii) hold under a stronger assumption. Assumption (5) allows a comparison of the solutions to those of the usual KPZ equation corresponding to $V(y) = y^2$, which is linearizable. For the proof of the diffusive limit, we shall also need very weak estimates on the derivatives of any order of $V$, satisfied for all ‘reasonable’ functions.

**Notations.** The scale parameter is an arbitrary constant $M > 1$ fixed once and for all. The notation: $f(u) \leq g(u)$, resp. $f(u) \geq g(u)$ means: $|f(u)| \leq C|g(u)|$, resp. $|f(u)| \geq C|g(u)|$, where $C > 0$ is an unessential constant (depending only on $d, \nu, D$ and $M$). Similarly, $f(u) \approx g(u)$ means: $f(u) \leq g(u)$ and $g(u) \leq f(u)$. We denote by $L^p, p \in [1, \infty]$ the usual Lebesgue spaces with associated norm $\| \cdot \|_p$, by $H^{1,\infty}$ the Sobolev space of bounded functions with bounded generalized derivative, and by $C^{1,2}$ the space of functions which are $C^1$ in time and $C^2$ in space. The positive, resp. negative part of a function $f$ is denoted by $f_+$, resp. $f_-$; by definition, $f_+, f_- \geq 0$, $f = f_+ - f_-$ and $f_+ f_- = 0$. The oscillation $\text{osc}_\Omega f$ of a continuous function $f$ on a domain $\Omega$ is defined as $\sup_\Omega f - \inf_\Omega f$; the average $\frac{1}{\Omega} \int_\Omega f$ of $f$ on a bounded domain $\Omega$ is denoted by $\frac{1}{\Omega} f$. The space of lower, resp. upper semicontinuous functions on a domain $\Omega$ is denoted by $\text{LSC}(\Omega)$, resp. $\text{USC}(\Omega)$.

**2 Bounds for the homogeneous equation: the case of a bounded initial condition**

We consider in this section the homogeneous equation,
\[
\partial_t h = \nu \Delta h + \lambda V(|\Delta h|)
\] (2.1)
with initial condition \( h_0(x) = h(0, x) \) in \( W^{1,\infty} \). One finds in the literature a detailed study of the particular case \( V(y) = y^q, \ q > 1 \). Most basic results (including existence), based on the principle of maximum or on a short-time series expansion of the mild solution, depend very little on the precise form of \( V \), provided it is regular enough and, say, polynomially bounded. We quickly review them now and leave it to the reader to check that they extend to a rate \( V \) satisfying Assumptions 1.1 (1), (2), (4).

By [5] and [10], the Cauchy problem has a unique, global solution \( u \) which is classical for positive times, that is, \( u \in C([0, +\infty) \times \mathbb{R}^d) \cap C^{1,2}(0, \infty) \times \mathbb{R}^d) \). The comparison principle, in the form proved by Kaplan [38] for classical, bounded solutions of non-linear parabolic equations on unbounded spatial domains, implies that \( h_t \geq 0 \) for all \( t \geq 0 \) (resp. \( h_t \leq 0 \) for all \( t \geq 0 \) if \( h_0 \geq 0 \) (resp. \( h_0 \leq 0 \)) and yields the a priori estimates

\[
\|h_t\|_\infty \leq \|h_0\|_\infty, \quad \|\nabla h_t\|_\infty \leq \|\nabla h_0\|_\infty \quad (t \geq 0).
\]  

We now prove time-decay estimates of the solution for various norms, emphasizing those which are not straightforward extension of previously known results for \( V(y) = y^q \). Such estimates come roughly from three different sources. Generally speaking, constants appearing in the inequalities deteriorate when \( \nu \to 0 \) whenever parabolic estimates are involved (see below); Proposition 2.3 (ii), (iii) is an outstanding exception.

We recall here briefly for non-specialists the maximum principle and the comparison principle for parabolic PDE’s, in a weak form which is sufficient for section 2. Standard references on the subject are e.g. [26], [21], [7].

**Proposition 2.1 (maximum and comparison principle)** Let \( u(t, x), (t, x) \in [0, T] \times \mathbb{R}^d \) be a classical solution of the parabolic PDE \( \partial_t u(t, x) = \nu \Delta u(t, x) + W(t, x, \nabla u(t, x)) \), where \( W \) is a smooth function, bounded in any subset of the form \( \mathbb{R} \times \mathbb{R}^d \times K, K \subset \mathbb{R}^d \) compact. Assume that \( \sup_{[0, T] \times \mathbb{R}^d} |u| < \infty \) and \( \sup_{[0, T] \times \mathbb{R}^d} |\nabla u| < \infty \). Then:

(i) (weak maximum principle) \( \forall t \in [0, T], \ |u_t|_\infty \leq |u_0|_\infty \).

(ii) (weak comparison principle) Let \( \bar{U}, \bar{U} \) be a super-, resp. sub-solution of the above PDE, namely, \( \bar{U}, \bar{U} \in C^{1,2}([0, T] \times \mathbb{R}^d) \) and \( \partial_t \bar{U}(t, x) \geq \nu \Delta \bar{U}(t, x) + W(t, x, \nabla \bar{U}(t, x)), \partial_t \bar{U}(t, x) \leq \nu \Delta \bar{U}(t, x) + W(t, x, \nabla \bar{U}(t, x)) \). Assume \( \bar{U}_0 \geq \bar{U}_t \). Then \( \bar{U}_t \geq \bar{U}_t \) for all \( t \geq 0 \).

Note that the above proposition extends under appropriate monotonicity hypothesis to parabolic PDE’s of the form \( \partial_t u(t, x) = \nu \Delta u(t, x) + W(t, x, u(t, x), \nabla u(t, x)) \). However, it is precisely the absence of dependence of \( W \) on \( u(t, x) \) that makes two-sided a priori estimates like (2.2) so easy.

### 2.1 Comparison to the linear heat equation

Assumptions 1.1 (4)-(5), \( 0 \leq V(y) \leq y^2 \), allows (as we shall presently see) a direct comparison with the linear heat equation if either \( h_0 \geq 0 \) or \( h_0 \leq 0 \). Bounds for signed initial conditions follow then from the comparison principle: namely, letting \( \underline{h}, \overline{h} \) be the solution of (2.1) with initial condition \( h_0^\pm \), resp. \( -h_0^\pm \), one has

\[
\underline{h} \leq \overline{h} \leq 0; \quad h_0 \leq h \leq \overline{h}.
\]

Also, \( t \mapsto ||\overline{h}||_1 \) is increasing, while \( t \mapsto ||\underline{h}||_1 \) is decreasing.
Considering first $\tilde{h}$, the comparison principle allows one to bound the solution of (2.1) by the solution of the linear heat equation with same initial condition, namely,

$$|\tilde{h}(t)| \leq e^{\nu \Delta} h_0^+.$$  \hspace{1cm} (2.4)

We now turn to $\tilde{h}$ and bound similarly the solution of (2.1) with positive initial condition by the solution $u$ of the standard KPZ equation, $\partial_t u = \nu \Delta u + \lambda |\nabla u|^2$ with the same initial condition. The exponential transformation $w := e^{\frac{1}{\nu}u} - 1$ turns it into the linear equation $\partial_t w = \nu \Delta w$, with positive initial condition $w_0 = e^{\frac{1}{\nu}h_0^+} - 1$. The inequality $x \leq \frac{1}{\nu}(e^{\frac{1}{\nu}x} - 1)$, $x \geq 0$ yields

$$||\tilde{h}||_\infty \leq ||w||_\infty \leq e\frac{\nu}{\lambda}||w_0||_\infty.$$  \hspace{1cm} (2.5)

To go further, we assume $w_0 \in L^1$ and use the following standard parabolic estimates [62] for $q = 1$.

**Proposition 2.2 (parabolic estimates)** There exist constants $C_k$, $k = 0, 1, \ldots$ depending only on $d$ such that, for every regular enough function $f_0 : \mathbb{R}^d \to \mathbb{R}$ and $p \geq q \geq 1$,

$$\|\nabla^k e^{\nu \Delta} f_0\|_p \leq C_k(\nu t)^{-\frac{d}{2} + \frac{k}{p} - \frac{1}{p}} \|f_0\|_q, \quad k \geq 0.$$  \hspace{1cm} (2.6)

Let $\mu$ be the Lebesgue measure on $\mathbb{R}^d$. The well-known identity $\int f(u(x))dx = \int_0^{+\infty} \mu(u > a)f'(a)da$ valid for $u : \mathbb{R}^d \to \mathbb{R}_+$ measurable and $f : \mathbb{R}_+ \to \mathbb{R}$ smooth yields for $f(u) = \frac{1}{\lambda}(e^{\frac{1}{\nu}h_0^+} - 1)$

$$||\tilde{h}_1||_1 \leq \frac{\nu}{\lambda}||w_1||_1 \leq \frac{\nu}{\lambda}||w_0||_1 \leq \int h_0^+(x)1_{h_0^+(x) \leq \nu/\lambda}dx + \int_{\nu/\lambda}^{+\infty} \mu(h_0^+ > a)e^{\frac{1}{\nu}a}da$$

$$\leq ||h_0^+||_1(1 + \int_{\nu/\lambda}^{+\infty} e^{\frac{1}{\nu}a}da)$$

$$\leq ||h_0^+||_1 e^{\frac{1}{\nu}||h_0^+||_\infty}.$$  \hspace{1cm} (2.7)

so

$$\tilde{I}_\infty := ||h_0^+||_1 e^{\frac{1}{\nu}||h_0^+||_\infty}$$  \hspace{1cm} (2.8)

is an upper bound for $\sup_{t \geq 0} ||\tilde{h}_1||_1$ (see [42], Proposition 2 (iii)); at the same time, one gets

$$||\tilde{h}_1||_\infty \leq \frac{\nu}{\lambda}||w_1||_\infty \leq \frac{\nu}{\lambda}||w_0||_1 t^{-d/2} \leq \tilde{I}_\infty t^{-d/2}.$$  \hspace{1cm} (2.9)

On the other hand, (2.4) gives immediately if $h_0^- \in L^1$

$$||\tilde{h}_1||_\infty \leq ||h_0^-||_1 t^{-d/2}.$$  \hspace{1cm} (2.10)

Thus one has shown a global bound for the $L^1$-norm, and a time-decay in $O(t^{-d/2})$ for the sup-norm of solutions of (2.1) with arbitrary integrable initial condition $h_0 \in \mathcal{W}^{1,\infty} \cap L^1$. 

8
2.2 Time-decay of solutions of viscous Hamilton-Jacobi equations

A second series of results is a particular case of the more general time-decay of the gradients of solutions of viscous Hamilton-Jacobi equations, which can itself be seen as (1) an extension to non-linear equations of the standard parabolic estimates; (2) or a multi-dimensional extension of the decay of solutions of scalar conservation laws, see [31], [10], [8], or [9], section 3 for further results concerning in particular a single-sided bound on the Hessian. Generally speaking, such results rely on convexity assumptions on $V$. Here we shall only state the following estimate, which is an extension of [10], Theorem 1. In Proposition 2.3 by exception, constants, explicit and implicit (i.e. hidden by the sign $\lesssim$) are $\nu$-independent.

**Proposition 2.3 (time-decay of the gradient)**

(i) If $yV'(y) - V(y) \geq 0$, then

$$||\nabla h_t||_\infty \lesssim ||h_0||_\infty (\nu t)^{-\frac{1}{2}}.$$  (2.11)

(ii) Under the stronger assumption

$$yV'(y) - V(y) \geq C \min(y^2, y^q), \quad y \geq 0$$  (2.12)

for some constant $C > 0$ and some exponent $q \in (1, 2]$, one has

$$||\nabla h_t||_\infty \lesssim \left(\frac{||h_0||_\infty / \lambda}{t}\right)^{1/q} \quad , \quad t \leq \frac{||h_0||_\infty}{\lambda}$$  (2.13)

$$||\nabla h_t||_\infty \lesssim \left(\frac{||h_0||_\infty / \lambda}{t}\right)^{1/2} \quad , \quad t \geq \frac{||h_0||_\infty}{\lambda}$$  (2.14)

(iii) Under the even stronger assumption

$$yV'(y) - V'(y) \geq Cy^2,$$  (2.15)

one has for all $t \geq 0$

$$|\nabla h_t(x)| \lesssim \left(\frac{||h_t(x)||/\lambda}{t}\right)^{1/2}, \quad x \in \mathbb{R}^d.$$  (2.16)

Hence in particular

$$||\nabla h_t||_\infty \lesssim \left(\frac{||h_0||_\infty / \lambda}{t}\right)^{1/2}.$$  (2.17)

Note that condition (i), $yV'(y) - V(y) \geq 0$ is a consequence of the convexity of $V$ (see Assumption 1.1 (3)). On the other hand, the hypothesis (2.12) in (ii) holds true for functions $V(y)$ that behave like $y^2$ for $y$ small, and like $y^q$, $1 < q \leq 2$ for $y$ large; the stronger hypothesis (2.15) in (iii) for functions that behave like $y^2$ both for $y$ small and $y$ large. Note that the decay in (2.11) is produced by the diffusion term $\nu \Delta$ in the equation, so it might be called a generalized parabolic estimate; while (2.13,2.14) or (2.17) are diffusion-independent effects of the non-linear term in the equation, and would also hold true for viscosity solutions of the first-order Hamilton-Jacobi equation obtained by letting $\nu \to 0^+$. 

9
Proof. We first rescale $h$ and $x$ by letting $x \to x' = v^{-\frac{1}{2}}x$, $h \to u = \frac{1}{v}h$ so that $\nabla_{x'}u = (\frac{1}{v})^{\frac{1}{2}}\nabla_xu$, and $W(y) = \frac{1}{v^2}V((\frac{1}{v})^{\frac{1}{2}}y)$, so that the equation for $u$

$$\partial_t u = \Delta u + W(|\nabla u|)$$

is independent of the parameters $\nu, \lambda$.

(i) Referring to the proof of Lemma 3 in [31], from which [31], Theorem 2 follows immediately, and letting directly $\epsilon = 0$, we see that $\frac{\nabla u^2}{\partial t(u)}$ is a super-solution for the parabolic operator

$$\tilde{N}(w) := \Delta w + b \cdot \nabla w + cw^2 + ew - \partial_t w,$$

where $c := 2\theta(u)\theta''(u)$, $e = -2(W(\nabla u) - \nabla u \cdot W'(\nabla u))$ (note that for $V$ homogeneous, $V(|\nabla u|) = |\nabla u|^q$, $\tilde{N}(\frac{\nabla u^2}{\partial t(u)}) = N(\frac{\nabla u^2}{\partial t(u)})$ where $N$ [31] has instead of $ew$ a sum of two terms, $2(q-1)\theta'\theta^{-1}(u)\theta'(u)w^{1+q/2} - 2\theta''(u)H(\nabla u)w$, with $H(\xi) = W(\xi) - \xi \cdot \nabla W(\xi) + (q-1)|\xi|^q$, the first term $2(q-1)\theta'\theta^{-1}(u)\theta'(u)w^{1+q/2}$ compensating the last term $(q-1)|\xi|^q$ in $H$). Choose $\theta = \theta_1$ as in [31], eq. (24), so that $c = -1$. Now, for $V$ isotropic, $e = -2\theta''(u)(W(y) - yW'(y))$, $y = |\nabla u|$; this is $\leq 0$ under the assumptions (1.1). Hence $r^{-1}$ is a sub-solution of $\tilde{N}$, and the comparison principle yields $|\nabla u_t(x)|^2 \leq \frac{\theta''(u_0(x))}{t}$. Now $||u_1||_{\infty} \leq 2||u_0||_{\infty}$. Scaling back to the original variables $h, x$ yields the first bound, $||\nabla h||_{\infty} \leq ||h_0||_{\infty}(vn)^{-\frac{1}{2}}$.

(ii) The second bound is an extension of [10]. Exactly as in (i), one may assume that $u_0 \leq 0$. Up to an overall change of sign, $u \mapsto -u$, we are in the conditions of [10], Theorem 1, with $p = 2$, except that $V$ is not necessarily a power function. Letting again $\epsilon = 0$, the function $\Theta \in \Omega$ in eq. (20) p. 2005 is here $\Theta(r) = \Theta(v^2) = 2\gamma^2 \frac{\lambda}{\nu} W(y) - W(y) = yW'(y) - W(y)$; by assumption,

$$\Theta(r) \geq \min(r \left(\frac{\lambda^2}{\nu} \right)^{\frac{1}{2}} \nu^0/2), r \leq \frac{\nu}{\lambda^2} + \left(\frac{\lambda^2}{\nu} \right)^{\frac{1}{2}} r^2 1_{r > \frac{\nu}{\lambda^2}}.$$  

Eq. (14) p. 2003 implies

$$L_w + C\nu^{-2}\Theta(v^2 w)w \leq 0$$

with $L = \partial_t - \Delta$ up to some gradient term vanishing on functions $h$ which are independent of $x$ (see eq. (10) p. 2002 for a precise definition), $\nu := \sqrt{\omega}$, $w = |\nabla w|$. Now $\nu \leq ||u_0||_{\infty}$, hence

$$v^{-2}\Theta(v^2 w)w \geq w^2 \quad (w \leq \left(\frac{\nu ||u_0||_{\infty}}{\lambda^2} \right)^{-1}),$$

$$v^{-2}\Theta(v^2 w)w \geq \left(\frac{\nu ||u_0||_{\infty}}{\lambda^2} \right)^{\frac{1}{2}} w^{1+\frac{1}{2}} \quad (w \geq \left(\frac{\nu ||u_0||_{\infty}}{\lambda^2} \right)^{-1}),$$

so $N^w \leq 0$, where $N$ is the parabolic differential operator

$$N : h \mapsto \mathcal{L}h + CN(h), \quad N(h) = h^2 \mathbf{1}_{h \leq \left(\frac{\nu ||u_0||_{\infty}}{\lambda^2} \right)^{-1}} + \left(\frac{\nu ||u_0||_{\infty}}{\lambda^2} \right)^{\frac{1}{2}} h^{1+\frac{1}{2}} 1_{h > \left(\frac{\nu ||u_0||_{\infty}}{\lambda^2} \right)^{-1}}.$$  

(2.22)
Note that $N$ is an increasing function. The comparison principle thus implies that $w \leq h$ if $Nh \geq 0$. Such a function $h = h(t)$ is easily constructed by solving the ordinary differential equations $\partial_t h = -\left(\frac{\|u_0\|_\infty}{h}\right)^{1/q} h^{1+\frac{1}{q}}$ for $h \leq \left(\frac{\|u_0\|_\infty}{h}\right)^{1/q}$, $\partial_t h = -h^2$ for $h \geq \left(\frac{\|u_0\|_\infty}{h}\right)^{1/q}$, yielding up to unimportant constants $h(t) = \left(\frac{\|u_0\|_\infty}{h}\right)^{1/q} t^{1-\frac{1}{q}}$ ($t \leq \frac{\|u_0\|_\infty}{h}$), $h(t) = \frac{1}{t}$ ($t \geq \frac{\|u_0\|_\infty}{h}$). This gives bounds for $\|\nabla u\|_\infty$ by taking the square-root, and then bounds for $\|\nabla u\|_\infty$ by noting that $\nabla u = 2u^{1/2} \nabla v$ and $u^{1/2}(x) \leq \|u_0\|_\infty^{1/2}$ (see [9], proof of Proposition 3.1), namely,

$$\|\nabla u\|_\infty \leq \left(\frac{\|u_0\|_\infty}{t}\right)^{1/q} \|\nabla u_0\|_\infty^{1/q}, \quad t \leq \frac{\|u_0\|_\infty}{\lambda^2};$$

$$\|\nabla u\|_\infty \leq \left(\frac{\|u_0\|_\infty}{t}\right)^{1/q}, \quad t \geq \frac{\|u_0\|_\infty}{\lambda^2};$$

Hence (2.13, 2.14) by rescaling.

(iii) Under the stronger assumption (2.15), the previous computations yields $h(t) = \frac{1}{t}$ for all $t > 0$. Hence $|\nabla u(t, x)| \leq \left(\frac{\|u_0\|_\infty}{\lambda^2}\right)^{1/q}$. Eq. (2.16), and then (2.17), follow by rescaling and using the a priori bound $\|h\|_\infty \leq \|u_0\|_\infty$.

\[\square\]

### 2.3 Bounds through integral representation of mild solutions

The third source of results is the integral form of the equation,

$$h_t = e^{\nu \Delta} h_0 + \lambda \int_0^t e^{(t-s)\nu \Delta} V(\nabla h_s) ds,$$

(2.25)

the solutions of which, traditionally called *mild solutions*, are not necessary twice differentiable in space. (2.25) is used to prove local-in-time well-posedness of the equation, while the a priori estimates (2.2) imply global existence [5]. We shall not come back to this; instead, we give an application to the proof of various bounds for the gradient and for higher derivatives of the solution. Generally speaking, $\nu$-dependent constants (throughout denoted by $C$ and possibly varying from line to line) come out of the computations everywhere. From [5], the solution obtained by iterating (2.25),

$$h_t^{(0)} = h_0, \quad h_t^{(k+1)} = e^{\nu \Delta} h_t^{(k)} + \lambda \int_0^t e^{(t-s)\nu \Delta} V(\nabla h_s^{(k)}) ds \quad (k \geq 0)$$

(2.26)

in search for a fixed point is obtained as a converging series for $t < T_1^*$,

$$T_1^* = C(\lambda \|\nabla h_0\|_\infty)^{-2}$$

(2.27)

for some constant $C$, and shown to be uniformly smooth: namely, for every $k \geq 0$, $\|\nabla^k h_t\|_\infty \leq C_k \|\nabla^k h_0\|_\infty$. $0 \leq t \leq T_1^*$ provided the initial solution has bounded derivatives of order $\leq k$. By an appropriate choice of $C$ one may assume that $C_2 = C_3 = 2$. This, in turn, shows, using the a priori bound, $\|\nabla h_t\|_\infty \leq \|\nabla h_0\|_\infty$, that the solution at any later time also has bounded derivatives of arbitrary order. We are interested here in quantitative bounds that can be shown to be close to optimal in some case where explicit computations are possible (see next paragraph).
We shall give two different results. Recall \( I_{t_{\infty}} = ||h_{t_{\infty}}||_1 e^{\frac{t_{\infty}}{2}||h_{0}||_{\infty}} \) (see (2.8)). Our first result uses hypothesis (2.15), \( yV(y) - V(y) \geq C\gamma^2 \) (see Proposition (2.3)(iii)).

**Theorem 2.1 (decay in \( L^1 \)-norm of the gradient)** Assume \( h_0 \in W^{1,\infty} \cap L^1 \) and let \( h_t \) be the solution of the KPZ equation (2.7), with \( V \) satisfying the hypothesis (2.15). Then

\[
||\nabla h_t||_1 \leq \max(J_{t_{\infty}}, J_{t_{\infty} + \frac{\gamma}{2}}) (1 + t)^{-\frac{\gamma}{2}}, \quad J_{t_{\infty}} = \sup \left( (1, \frac{||\nabla h_0||_1}{||h_0||_1} + \lambda ||\nabla h_0||_{\infty} (1 + O(||h_0||_{\infty}))) \right) ||h_0||_1 e^{C||h_0||_{\infty}}.
\] (2.28)

The time-decay in \( O(t^{-\frac{\gamma}{2}}) \) of \( ||\nabla h_t||_1 \) is shown in [9] for \( V(y) = \gamma^2 \), but with a constant \( J_{t_{\infty}} \) which is roughly \( e^{t_{\infty}} \) and thus far from optimal (see p. 1290 and 1291). The emphasis there was on the asymptotic convergence for \( t \to \infty \) of the solution to a multiple of the heat-kernel (see Theorem 2.3 (a)), an interesting result in itself to which we do not come back here.

**Proof.**

Our proof, based on intuition derived from the explicit computations of the next paragraph, shows that there are different time regimes for \( ||\nabla h_t||_1 \). Initially (i) the \( L^1 \)-norm of the gradient may increase (as is the case for the \( L^1 \)-norm of the solution when the initial condition is positive); for later times (iii) it decreases like the square-root of time. There also appears a regime (ii) for intermediate times, during which the \( L^1 \)-norm of the gradient is shown to be essentially constant.

These three regimes come from the three essentially different bounds one has on \( ||\nabla h_t||_{\infty} \); namely, (i) \( ||\nabla h_t||_{\infty} \leq ||\nabla h_0||_{\infty} \) by the comparison principle; (ii) \( ||\nabla h_t||_{\infty} \leq \sqrt{\frac{||h_0||_{\infty}}{t}} t^{-\frac{\gamma}{2}} \) by Proposition (2.3)(iii); (iii)

\[
||\nabla h_t||_{\infty} \leq ||h_t||_2 t^{-\frac{\gamma}{2}} \leq I_{t_{\infty}} t^{-(d+1)/2},
\] (2.29)

as follows from a combination of Proposition (2.3)(i) and of the parabolic estimates developed in the lines following eq. (2.8), where

\[
I_{t_{\infty}} = ||h_0||_1 e^{C||h_0||_{\infty}}
\] (2.30)

is an upper bound for \( \sup_{t \geq 0} ||h_t||_1 \) (in order to get not too complicated formulas, we avoid the unpleasant task of optimizing the constants, and choose \( C \) large enough).

(i) For \( t \) small one uses the trivial bound (i), \( ||\nabla h_t||_1 \leq ||\nabla h_0||_{\infty} \), and applies the iterative scheme (2.26) in uniform time slices \([T_0, T_1] = [0, T_1], [T_1, T_2] = [T_1, 2T_1], \ldots, [T_{n_0-1}, T_{n_0}] = [(n_0 - 1)T_1, n_0 T_1]\) where \( n_0 \approx \lambda ||h_0||_{\infty} \), so that \( ||\nabla h_0||_{\infty} \approx \sqrt{\frac{||h_0||_{\infty}}{t}} (T_{n_0})^{-\frac{\gamma}{2}} \). At some time comparable with \( T_{n_0} \), the bound (ii) on \( ||\nabla h_t||_{\infty} \) becomes better. We let \( M_n^{(k)} := \sup_{[T_n, T^{*}_{n+1}]} ||\nabla u_t^{(k)}||_1 \) and \( M_n := \sup_{[T_n, T^{*}_{n+1}]} ||\nabla u_t||_1 = \lim_{k \to \infty} M_n^{(k)} \). By (2.26) and the parabolic estimates recalled in (2.6).

\[
M_n^{(k+1)} \leq ||\nabla h_{T_n}||_1 + \lambda \sup_{t \in [T_n, T^{*}_{n+1}]} \int_{T_n}^{t} (t - s)^{-\frac{\gamma}{2}} ||(\nabla h^{(k-1)}(s))^{2}||_1 ds,
\] (2.31)

together with the interpolation inequality,

\[
||\nabla h^{(k-1)}(s)||_2 \leq ||\nabla h^{(k-1)}(s)||_1 ||\nabla h^{(k-1)}(s)||_{\infty},
\] (2.32)

one obtains

\[
M_n^{(k+1)} \leq ||\nabla h_{T_n}||_1 + C^{-1} M_n^{(k)}
\] (2.33)
where $C^{-1}$ is proportional to the inverse of the constant $C$ in the definition of $T_1^*$. For $C$ large enough, this yields $\sup_{k} M_{n}^{(k)} \leq 2M_{n-1}$ and $M_n \leq 2M_{n-1}$. Thus

$$M_{n_0} \leq \|\nabla h_0\|_1 e^{C\|h_0\|_\infty} \leq \|\nabla h_0\|_1 I_\infty$$  \hspace{1cm} (2.34)

with an appropriate definition of the constant in (2.30).

(ii) For $n \geq n_0$ one defines inductively $T_n^*$ by $T_{n+1}^* - T_n^* = C(\lambda\|\nabla h_{T_n^*}\|_\infty)^{-2}$. Note that $T_{n_0}^* \approx \|h_0\|_\infty$, by the second estimate (ii) on $\|\nabla h_t\|_\infty$, $T_{n+1}^* - T_n^* \geq T_{n_0}^*$, so

$$T_n^* \geq \frac{\|h_0\|_\infty}{\|\nabla h_0\|_\infty^2} \left(1 + \frac{C}{\lambda\|h_0\|_\infty}\right)^{n-n_0}, \hspace{1cm} n \geq n_0.$$  \hspace{1cm} (2.35)

Instead of the bound $\|\nabla e^{(t-T_n^*)\Delta} h_{T_n^*}\|_1 \leq \|\nabla h_{T_n^*}\|_1$ used in (2.31), it is more clever for $n$ and $t - T_n^* \geq n_0$ large enough to use the parabolic estimate $\|\nabla e^{(t-t_n^*)\Delta} h_{T_n^*}\|_1 \leq (t - T_n^*)^{-\frac{1}{2}} I_\infty$ if the latter expression is $\leq \|\nabla h_{T_n^*}\|_1$. Thus one gets the improved estimate

$$M_{n}^{(k+1)} \leq \sup_{t \in [T_{n_0}^*, T_{n+1}^*]} \left(\inf(\|\nabla h_{T_n^*}\|_1, I_\infty(t - T_n^*)^{-\frac{1}{2}} + \lambda\|\nabla h_{T_n^*}\|_\infty M_{n}^{(k)}(t - T_n^*)^{\frac{1}{2}})\right).$$  \hspace{1cm} (2.36)

If $\|\nabla h_{T_n^*}\|_1 \geq I_\infty(T_{n+1}^* - T_n^*)^{-\frac{1}{2}} \approx \lambda I_\infty\|\nabla h_{T_n^*}\|_\infty$, the improved estimate (2.36) is better than (2.31) and yields

$$M_{n}^{(k+1)} \leq \sup\left(\|\nabla h_{T_n^*}\|_1 + \lambda\|\nabla h_{T_n^*}\|_\infty M_{n}^{(k)} \frac{\|h_{T_n^*}\|_1}{\|\nabla h_{T_n^*}\|_1}, \right.$$  

$$\left.\sup_{T_n^* \leq t \leq T_{n+1}^* \leq \|h_{T_n^*}\|_1/\|\nabla h_{T_n^*}\|_1} \left(I_\infty(t - T_n^*)^{-\frac{1}{2}} + \lambda\|\nabla h_{T_n^*}\|_\infty M_{n}^{(k)}(t - T_n^*)^{\frac{1}{2}}\right)\right).$$  \hspace{1cm} (2.37)

The function $x \mapsto \frac{a}{x} + bx$, here $x = \sqrt{t - T_n^*}$, is bounded on any interval of $\mathbb{R}_+$ by the max of its values at the two ends of the interval. Hence

$$M_{n}^{(k+1)} \leq \sup\left(\|\nabla h_{T_n^*}\|_1 + \lambda\|\nabla h_{T_n^*}\|_\infty M_{n}^{(k)} \frac{I_\infty}{\|\nabla h_{T_n^*}\|_1}, I_\infty(T_{n+1}^* - T_n^*)^{-\frac{1}{2}} + C^{-1} M_{n}^{(k)}\right)$$  \hspace{1cm} (2.38)

with $C^{-1} < 1$. Iterating these affine inequalities yields either

$$M_n \leq I_\infty(T_{n+1}^* - T_n^*)^{-\frac{1}{2}} \approx \lambda I_\infty\|\nabla h_{T_n^*}\|_\infty \leq I_\infty\|\nabla h_0\|_\infty;$$  \hspace{1cm} (2.39)

or, assuming on the contrary that $\|\nabla h_{T_n^*}\|_1 \geq I_\infty(T_{n+1}^* - T_n^*)^{-\frac{1}{2}}$,

$$M_n \leq \frac{\|\nabla u_{T_n^*}\|_1}{1 - \lambda I_\infty\|\nabla u_{T_n^*}\|_\infty}.$$  \hspace{1cm} (2.40)
Since the sequence \( n \mapsto \lambda I_n \| \nabla h_{T_n} \|_\infty \) is exponentially decreasing (as follows from the bound (ii) and the fact that the sequence \( (T'_n) \) is exponentially increasing, see (2.35)), the recursive sequence

\[
x_{n+1} = \frac{x_n}{1 - \lambda I_n \| \nabla h_{T_n} \|_\infty} \approx x_n + \lambda I_n \| \nabla h_{T_n} \|_\infty
\]

starting from \( x_1 \approx I_0 (T'_{n+1} - T'_{n})^{-\frac{1}{2}} \leq \lambda \| \nabla h_0 \|_\infty I_0 \), converges to

\[
x_\infty \leq \lambda I_0 \| \nabla h_0 \|_\infty (1 + O(\| h_0 \|_\infty)),
\]

This gives a global bound for \( M_n, n \geq 0 \),

\[
\sup_{n \geq 0} M_n \leq \bar{I}_\infty := \sup \left( \frac{\| \nabla h_0 \|_1}{\| h_0 \|_1}, \lambda \| \nabla h_0 \|_\infty (1 + O(\| h_0 \|_\infty)) \right) I_\infty,
\]

but no time decay yet in general.

(iii) For \( t \geq t_\infty^{2/d} \) we use the estimate (iii), \( \| \nabla h_t \|_\infty \leq I_\infty t^{-(d+1)/2} \) and prove the time decay in \( O(t^{-\frac{d}{2}}) \).

Let \( \bar{M}_n := \sup_{t\in[2^n,2^{n+1}]} \| \nabla h_t \|_1 \) for \( n \geq n_2 := 1 + \log_2 I_\infty^{2/d} \). For all \( t \in [2^n, 2^{n+1}) \),

\[
\| \nabla h_t \|_1 \leq 2^{-n/2} \| h_{t - 2^n} \|_1 + \lambda \int_{t - 2^n}^t ds (t - s)^{-\frac{1}{2}} \| \nabla h_s \|_1 \| \nabla h_t \|_\infty
\]

\[
\leq 2^{-n/2} I_\infty + \lambda 2^{n/2} (\bar{M}_n + \bar{M}_{n+1}) I_\infty (2^{-n/2} t^{-(d+1)/2}),
\]

hence

\[
\bar{M}_{n+1} \leq 2^{-n/2} I_\infty + \lambda I_\infty 2^{-nd/2} \bar{M}_n + \lambda \bar{M}_{n+1} 2^{-nd/2} \bar{M}_{n+1}.
\]

For \( n \geq n_2 \) one has by definition \( I_\infty 2^{-nd/2} \leq 1 \), so

\[
\bar{M}_{n+1} \leq (1 + O(\lambda))(2^{-n/2} I_\infty + \lambda \bar{M}_n),
\]

while \( \bar{M}_{n_2} \leq \bar{I}_\infty \) by (ii), implying by a straightforward induction

\[
\bar{M}_n \leq 2^{-n/2} I_\infty + \bar{I}_\infty \leq 2^{-n/2} (I_\infty + I_\infty^{1/d})
\]

and finally

\[
\sup_{t \leq t_\infty^{2/d}} \sqrt{t} \| \nabla h_t \|_1 \leq I_\infty + I_\infty^{1/d} \bar{I}_\infty.
\]

Finally,

\[
\sup_{t \leq t_\infty^{2/d}} \sqrt{t} \| \nabla h_t \|_1 \leq I_\infty^{1/d} \bar{I}_\infty.
\]

Hence the result.

\[\Box\]

**Remark.** If \( V \) does not satisfy (2.15), then the beginning of the proof is modified as follows: substituting to (ii) the bound \( \| \nabla h_t \|_\infty \leq \| h_t \|_\infty \) for \( t \geq n_0 \), see Proposition 2.3 (i), leads to \( n_0 \) defined such as to satisfy \( \| \nabla h_0 \|_\infty \approx \| h_0 \|_\infty (T'_n)^{-\frac{1}{2}} \), namely, \( n_0 \approx (\lambda \| h_0 \|_\infty)^2 \), and (compare with (2.34)) \( M_{n_0} \leq \| \nabla h_0 \|_\infty e^{C(\lambda \| h_0 \|_\infty)^2} \). A bound comparable to (2.28) probably holds with the quadratic exponential \( e^{C(\lambda \| h_0 \|_\infty)^2} \) substituting \( e^{C(\lambda \| h_0 \|_\infty)^2} \), which is clearly not optimal for the quadratic KPZ equation (see next paragraph).

Our second result is valid under our general assumptions on \( V \) stated in section 1.
Theorem 2.2 (bounds on higher derivatives) Let $h$ be the solution of eq. (2.1) with initial condition $h_0 \in W^{3,\infty}$. Then:

\[
\||\nabla^2 h_t||_\infty \leq P_1(\|h_0\|_\infty, \|\nabla h_0\|_\infty, \|\nabla^2 h_0\|_\infty) \frac{\ln(1 + t)}{t} \tag{2.50}
\]

\[
\||\nabla^3 h_t||_\infty \leq P_2(\|h_0\|_\infty, \|\nabla h_0\|_\infty, \|\nabla^2 h_0\|_\infty, \|\nabla^3 h_0\|_\infty) \left( \frac{\ln^2(1 + t)}{t^{3/2}} \right) \tag{2.51}
\]

where $P_1, P_2$ are polynomials.

Proof.

We already know that $\sup_{[0,T^*_1]} \||\nabla^k h_t||_\infty \leq 2\||\nabla^k h_0||_\infty$, $k = 2, 3$ for $T^*_1 \approx (\lambda\|\nabla h_0\|_\infty)^{-1}$. For $t \geq T^*_1$, $\nabla^2 h_t$ is the solution of an integral equation,

\[
\nabla^2 h_t = e^{\nu \lambda \Delta} \nabla^2 h_0 + \lambda \left( \int_{(1-\varepsilon)t}^{t} ds (e^{(1-s)\nu \Delta}) \nabla (V(\nabla h_s)) + \int_{0}^{(1-\varepsilon)t} ds (e^{(t-s)\nu \Delta}) \nabla (V(\nabla h_s)) \right). \tag{2.52}
\]

The idea is to commute the gradient with the heat operator $e^{(t-s)\nu \Delta}$ in order to make the most of parabolic estimates; the implied decay may be put to good use only for $t - s$ large enough, and we shall choose the parameter $\varepsilon$ accordingly. First, $\|e^{\nu \lambda \Delta} \nabla^2 h_0\|_\infty = \|\nabla^2 e^{\nu \lambda \Delta} h_0\|_\infty \leq ||h_0||_\infty t^{-1}$. Then, using $\nabla (V(\nabla h_s)) = V'(\nabla h_s) \cdot \nabla^2 h_s$ and Proposition 2.3(i), together with the inequality $V'(y) \leq y$, consequence of Assumption (1.1) (3), (4) (namely, $(2y)^2 \geq V(2y) \geq V(y) + yV'(y)$)

\[
\lambda \int_{(1-\varepsilon)t}^{t} ds (e^{(t-s)\nu \Delta}) \nabla (V(\nabla h_s)) \leq \lambda \int_{(1-\varepsilon)t}^{t} ds \frac{ds}{\sqrt{t-s}} \|\nabla^2 h_s\|_\infty \sup_{(1-\varepsilon)t, t} \|\nabla^2 h_s\|_\infty \leq \lambda \|h_0\|_\infty \sqrt{\varepsilon} \sup_{(1-\varepsilon)t, t} \|\nabla^2 h_s\|_\infty
\]

provided $(1 - \varepsilon)t \geq t$. To get a useful inequality we choose $\varepsilon$ so that $\lambda \|h_0\|_\infty \sqrt{\varepsilon} \leq \frac{1}{\lambda};$ namely, $\varepsilon \approx \min\left(\frac{1}{\lambda}, \frac{1}{\lambda \|h_0\|_\infty}\right)$. Finally (if $t \geq 1$) we split the second integral, $\int_{(1-\varepsilon)t}^{t}$, into several pieces:

\[
\lambda \int_{1/2}^{(1-\varepsilon)t} ds (e^{(t-s)\nu \Delta}) V(\nabla h_s) \leq \lambda \int_{1/2}^{(1-\varepsilon)t} ds \frac{ds}{t-s} \left( \frac{\|h_0\|_\infty}{\sqrt{t}} \right)^2 \leq \lambda \ln(\varepsilon^{-1}) \|h_0\|_\infty^2 t^{-1}; \tag{2.53}
\]

\[
\lambda \int_{1/2}^{t/2} ds (e^{(t-s)\nu \Delta}) V(\nabla h_s) \leq \lambda t^{-1} \int_{1/2}^{t/2} ds \|h_0\|_\infty^2 \leq \lambda \frac{\ln t}{t} \|h_0\|_\infty^2; \tag{2.54}
\]

\[
\lambda \int_{0}^{1} ds (e^{(t-s)\nu \Delta}) V(\nabla h_s) \leq \lambda t^{-1} \int_{0}^{1} ds \|\nabla h_0\|_\infty^2 = \lambda \|\nabla h_0\|_\infty^2 t^{-1}. \tag{2.55}
\]

If $t < 1$, we merge (2.54) into

\[
\lambda \int_{0}^{t/2} ds (e^{(t-s)\nu \Delta}) V(\nabla h_s) \leq \lambda t^{-1} \int_{0}^{t/2} ds \|\nabla h_0\|_\infty^2 = \frac{\lambda}{2} \|\nabla h_0\|_\infty^2. \tag{2.56}
\]

We finish as in the proof of Theorem 2.1 (iii), namely, letting $M_0 := \sup_{[0,T^*_1]} \||\nabla^2 h_t||_\infty$ and $M_n := \sup_{[2^{-n-1}T^*_1, 2^{-n}T^*_1]} \||\nabla^2 h_t||_\infty (n \geq 1)$, one has $M_0 \leq 2\||\nabla^2 h_0||_\infty$ and

\[
M_{n+1} \leq C^{-1} M_n + Q_1(\|h_0\|_\infty, \|\nabla h_0\|_\infty) 2^{-n} + \lambda \|\nabla h_0\|_\infty^2 \tag{2.57}
\]
for \( t < 1 \),
\[
M_{n+1} \leq C^{-1} M_n + Q_2(\|h_0\|_\infty, \|\nabla h_0\|_\infty) 2^{-n} + Q_3(\|h_0\|_\infty, \|\nabla h_0\|_\infty) n 2^{-n}
\]  
(2.58)
for \( t \geq 1 \). Hence (2.50).

This result is used as input to get similar a bound for \( \|\nabla^3 h\|_\infty \). This time we must move around three gradients in the best way; this gives three integrals rewritten as
\[
\int_{(1-\epsilon)t}^t ds e^{(t-s)v}\Delta \nabla^2 (V(\nabla h_s)), \quad \int_{t/2}^{(1-\epsilon)t} ds \nabla^2 e^{(t-s)v}\Delta \nabla (V(\nabla h_s)), \quad \int_0^{t/2} ds \nabla^3 e^{(t-s)v}\Delta V(\nabla h_s).
\]

One has \( \nabla^2(V(\nabla h_s)) = V'(\nabla h_s) \cdot \nabla^2 h_s + V''(\nabla h_s)(\nabla^2 h_s)^2 \), yielding the same constraints on \( \epsilon \), plus a supplementary quadratic term in \( \nabla^2 h_s \). The other terms are computed as before. Details are left to the reader.

\[ \square \]

### 2.4 An explicit example: decay of a 'bump' for the quadratic KPZ equation

We consider here the time-decay (pointwise and with respect to various norms) of the solution of the quadratic, homogeneous KPZ equation,
\[
\partial_t h = \Delta h + |\nabla h|^2
\]  
(2.59)
with initial "bump" condition \( h_0(x) = A 1_{|x| \leq L} \), where \( A, L > 0 \). The coefficient \( \lambda \) in front of the nonlinearity has been disposed of by a simple rescaling. Note that, if \( A = \|h_0\|_\infty \leq 1 \) (i.e. for a small initial condition), then the decay of the solution and of its derivatives in \( L^p \)-norms, \( 1 \leq p \leq \infty \) follow the parabolic estimates as for the solutions of the linear heat equation; thus we may assume that \( A \gg 1 \). We want to compare the decays obtained by explicit computation to those obtained in much greater generality in the previous paragraphs.

Through the exponential transformation, \( w = \exp h \), the equation becomes simply the heat equation,
\[
\partial_t w = \Delta w, \quad w_0 = e^{h_0}
\]  
(2.60)
so that
\[
h_t(x) = \ln \left( 1 + e^A - 1 + e^A - 1 + e^A - 1 + e^A \right)
\]  
(2.61)

Though the initial data is not in \( \mathcal{W}^{1,\infty} \), this defines a solution. We are interested in its behaviour for \( t \geq L^2 \), corresponding to the approximate amount of time necessary for the solution to smoothen up. Then
\[
h_t(x) \approx \ln \left( 1 + e^A \left( \frac{L^2}{t} \right)^{d/2} e^{-x^2/2t} \right).
\]  
(2.62)
There are two regimes:
(i) (initial regime) Assume $L^2 \leq t \leq L^2 e^{2A}$. Define $x_{\text{max}}(t) \in \mathbb{R}_+$ as the solution of the equation $e^{A\left(L^2 t\right)^{d/2}} e^{-x_{\text{max}}^2(t)/2t} = 1$; explicitly, $x_{\text{max}}(t) = \sqrt{2t(A - \frac{d}{2} \log(t/L^2))}$. If $|x| \geq x_{\text{max}}(t)$ then $h_t(x) \approx e^{A\left(L^2 t\right)^{d/2}} e^{-x^2/2t} \approx 1$. On the other hand, if $|x| \leq x_{\text{max}}(t)$, then $h_t(x)$ is still large, $h_t(x) \approx A - \frac{d}{2} \log(t/L^2) - \frac{x^2}{2t}$. In particular,

$$\|h_t\|_\infty \approx A - \frac{d}{2} \log(t/L^2) \geq 1 \quad (2.63)$$

and

$$\|h_t\|_1 \approx \|h_0\|_\infty \cdot \text{Vol}(B(0, x_{\text{max}}(t))) + e^{A\left(L^2 t\right)^{d/2}} \cdot t^d x_{\text{max}}^2(t). \quad (2.64)$$

The reader may easily check that dividing (2.63), resp. (2.64) by $t^{q/2}$ gives bounds for $\|\nabla^q h_t\|_\infty$, resp. $\|\nabla^q h_t\|_1$ if $q \in \mathbb{N}_*$.

(ii) (final regime) Assume $t \geq L^2 e^{2A}$. Then $h_t(x) \approx e^{A\left(L^2 t\right)^{d/2}} e^{-x^2/2t} \approx 1$; in other words, the bump has essentially disappeared. Furthermore,

$$\|h_t\|_\infty \approx e^{A L^d t^{-d/2}} \approx \|h_0\|_1 e^{\|h_0\|_\infty} t^{-d/2} \quad (2.65)$$

and

$$\|h_t\|_1 \approx e^{A L^d t} \approx \|h_0\|_1 e^{\|h_0\|_\infty}. \quad (2.66)$$

Again, bounds for $\|\nabla^q h_t\|_\infty$ or $\|\nabla^q h_t\|_1$ are obtained by dividing by $t^{q/2}$.

The above computations make it clear that the general bounds for $\|\nabla h_t\|_\infty$, see eq. (2.29), $\|\nabla h_t\|_1$ and $\|\nabla^q h_t\|_\infty$, $q = 2, 3$ obtained in Theorems 2.1 and 2.2 are essentially optimal.

3 Bound for the homogeneous equation: the case of unbounded initial conditions

We now want to prove existence of and bound solutions of the homogeneous KPZ equation,

$$\partial_t h = \nu \Delta h + V(\nabla h) \quad (3.1)$$

with unbounded initial condition $h_0$. We would typically like to consider a random initial condition which is a smoothened white noise (see Appendix A). This raises various problems. First (1), one would like to identify a functional space preserved by the linear heat equation, for which generalized parabolic estimates hold. Second (2), one would like to extend the comparison principle to such a functional space, in such a way as to prove existence of and bound the solution. Finally (3), one would like to identify the solution as the limit of solutions of (3.1) associated to a sequence of compactly supported (hence bounded) initial conditions converging to the original initial condition, so as to extend to the limit regularity results and estimates obtained in the previous section.

Let us now answer questions (1), (2), (3).
3.1 The functional spaces \( \mathcal{H}_0^0 \)

For \( f \in L^1_{\text{loc}}(\mathbb{R}^d) \) and \( \alpha \geq 0, x \in \mathbb{R}^d \), one may define

\[
f^\alpha(x) = \sup_{\tau > 0} (1 + \tau)^{\alpha} e^{\tau \Delta} |f|(x) \in [0, +\infty]
\]

and in particular

\[
f^\ast(x) := f^\alpha_0(x) = \sup_{\tau > 0} e^{\tau \Delta} |f|(x) \in [0, +\infty].
\]

Note that \( f^\ast \leq f^\alpha_0 \leq f^\beta_0 \) if \( \alpha \leq \beta \). If \( f \) is bounded, then \( f^\ast(x) \leq \|f\|_\infty \). On the other hand, the kind of random initial conditions we are interested in (see Appendix A) are a.s. unbounded, but satisfy a.s. \( f^\ast(x) < \infty \) for every \( x \) (see Lemma 6.5 and discussion thereafter); compare with the standard parabolic estimates, \( f_{d/2}^\ast(x) \leq \|f\|_1 \) for \( f \in L^1(\mathbb{R}^d) \). Note that, if \( \alpha = 0 \), the pointwise estimates

\[
|e^{\tau \Delta} f(x)| \leq f^\ast(x), \quad t \geq 0
\]

and, better still,

\[
(e^{\tau \Delta} f)^\ast(x) \leq f^\ast(x), \quad t \geq 0
\]

(generalizing to

\[
(e^{\tau \Delta} f)^\alpha(x) \leq f^\alpha_0(x), \quad t \geq 0
\]

are improvements on the global estimate \( \|e^{\tau \Delta} f\|_\infty \leq \|f\|_\infty \) which is useless for unbounded functions. If \( f^\alpha_0(x) < \infty \) for some \( \alpha > 0 \), then

\[
|e^{\tau \Delta} f(x)| \leq (1 + t)^{\alpha} f^\ast(x)
\]

decays in time.

**Definition 3.1** Let, for \( 0 \leq \alpha \leq d/2 \),

\[
\mathcal{H}_0^\alpha := \{ f \in L^1_{\text{loc}}(\mathbb{R}^d) \mid \forall x \in \mathbb{R}^d, f^\alpha_0(x) < \infty \}
\]

and

\[
\mathcal{H}^0 := \mathcal{H}_0^0 = \{ f \in L^1_{\text{loc}}(\mathbb{R}^d) \mid \forall x \in \mathbb{R}^d, f^\ast(x) < \infty \}.
\]

Since \( \alpha \leq d/2, \mathcal{H}_0^\alpha \supset L^1(\mathbb{R}^d) \cap L^\infty_{\text{loc}}(\mathbb{R}^d) \) (actually, it is easy to prove that \( \mathcal{H}_0^0 = \{0\} \) for \( \alpha > d/2 \)).

Another closely related definition is by averaging: if \( f \in L^1_{\text{loc}}(\mathbb{R}^d) \), and \( r \geq 0, x \in \mathbb{R}^d \), one may define

\[
f^d_\alpha(x) = \sup_{\rho > 0} (1 + \rho^2)^{\alpha} \frac{\int_{B(x, \rho)} |f(y)| dy}{\text{Vol}(B(x, \rho))}
\]

(called maximal function in real analysis for \( \alpha = 0 \), see the classical book by Stein [64]) where \( B(x, \rho) = \{ y \in \mathbb{R}^d; |y - x| < \rho \} \) is the Euclidean ball and \( \text{Vol}(B(x, \rho)) \) its volume. Here also, \( f^\#_0 \leq f^\#_\alpha \leq f^\#_\beta \) if \( 0 \leq \alpha \leq \beta \). It is convenient to denote averages by barred integrals, so that, by definition,

\[
\bar{f} = \frac{\int f}{\text{Vol}(\Omega)}.
\]

A simple result is the following:

**Lemma 3.2** There exists constants \( c, C > 0 \) such that, for every \( f \in C(\mathbb{R}^d) \), \( cf^\ast_\alpha(x) \leq f^d_\alpha(x) \leq Cf^\ast_\alpha(x) \).
Lemma 3.3

1. Let \( f \) given in terms of these two functions. We assume in the following lines that

\[
\int dy \frac{e^{-(x-y)^2/2t}}{(2\pi t)^{d/2}} |f(y)| = \int dr \frac{e^{-r^2/2t}}{(2\pi t)^{d/2}} \int_{\partial B(x,r)} dy |f(y)|
\]

Thus \( f_0^\alpha \leq f_0^\# \). If \( \alpha > 0 \) and \( t \geq 1 \) then

\[
(1 + t)^\alpha \int dy \frac{e^{-(x-y)^2/2t}}{(2\pi t)^{d/2}} |f(y)| \leq f_0^\# (1 + t)^{-1 - \frac{d}{2} + \alpha} \int dr (1 + r)^{1 + d - 2\alpha} e^{-r^2/2t} \leq f_0^\#(x),
\]  

so \( f_0^\alpha \leq f_0^\# \).

Conversely,

\[
e^{2\Delta}|f|(x) = \int \frac{e^{-\frac{1}{2}(|x-y|^2)}(2\pi)^{d/2} e^{r^2/2t} |f(y)|dy} \geq C \int_{B(x,r)} |f|.
\]  

In particular, an equivalent definition for \( \mathcal{H}_0^0 \) is:

\[
\mathcal{H}_0^0 = \{ f \in L^\infty_{loc}(\mathbb{R}^d) \mid \forall x \in \mathbb{R}^d, f_\alpha^\#(x) < \infty \}.
\]  

Note also that, if \( f \) is lower semicontinuous (in particular, if \( f \) is continuous), \( f_\alpha^\#(x) \geq \lim_{r \to 0} \int_{B(x,r)} |f| \geq |f(x)| \), and similarly \( f_\alpha^*(x) \geq |f(x)| \).

**Lemma 3.3**

1. Let \( f \in \mathcal{H}_0 \cap C(\mathbb{R}^d) \). Then \( f_\alpha^\# \) and \( f_\alpha^* \) are continuous.

2. Let furthermore

\[
\beta := \sup \{ \gamma \in [0, 1] \mid (x, y) \mapsto \frac{|f(x) - f(y)|}{|x - y|^\gamma} \text{ in } L^\infty_{loc} \} \in [0, 1]
\]  

be the maximum local Hölder exponent of \( f \), and assume \( \beta > 0 \). Then \( f_\alpha^\# \) and \( f_\alpha^* \) are Hölder continuous, with Hölder exponent \( \frac{\beta}{1 + \beta} \in [0, \frac{1}{2}] \).

In particular, \( f_\alpha^\# \) and \( f_\alpha^* \) are \( \left( \frac{1}{2} \right) \)-Hölder continuous if \( f \in C^1 \).

**Proof.** For the sake of the proof we choose a bounded function \( \phi : B(0, 1) \to \mathbb{R}_+ \) such that \( \phi(u) = \phi(|u|) \) is strictly increasing, \( \phi(0) = 0 \), \( \frac{\phi(u)}{u} > 2 \) and \( \frac{\phi(u)}{u} \to u \to \infty \); we assume furthermore that \( \phi(u) = o_{u \to 0}(u^{1/3}) \), so that the function \( \chi(u) = \chi(|u|) = \phi(u) \sqrt{\frac{\phi(u)}{u}} \) satisfies the same properties but \( \frac{\chi(u)}{\phi(u)} \to u \to \infty \). The core of the proof is a bound on the modulus of continuity of \( f^* := f^\alpha_0^*, f^\# := f^\alpha_0^\# \) given in terms of these two functions. We assume in the following lines that \( |y - x| \leq 1 \).
We now obtain a modulus of continuity for \( f^\sharp \), (from which it follows by Lemma 3.2 that \( f^* \) is also). Since \(|y - x| \leq 1\), then

\[
\int_{B(y,r)} |f| \leq \sup_{B(x,2)} |f|, \quad r \leq 1
\]

and

\[
\int_{B(y,r)} |f| \leq \left( \frac{r + 1}{r} \right)^d \int_{B(x,r+1)} |f| \leq 2^d f^\sharp(x), \quad r > 1.
\]

So

\[
\sup_{B(x,1)} f^\sharp \leq 2^\alpha \max(\sup_{B(x,2)} |f|, 2^d f^\sharp(x)).
\]

(i) Let us first prove that \( f^\sharp \) is locally bounded (from which it follows by Lemma 3.2 that \( f^* \) is also). Since \(|y - x| \leq 1\), then

\[
\int_{B(y,r)} |f| \leq \sup_{B(x,2)} |f|, \quad r \leq 1
\]

and

\[
\int_{B(y,r)} |f| \leq \left( \frac{r + 1}{r} \right)^d \int_{B(x,r+1)} |f| \leq 2^d f^\sharp(x), \quad r > 1.
\]

So

\[
\sup_{B(x,1)} f^\sharp \leq 2^\alpha \max(\sup_{B(x,2)} |f|, 2^d f^\sharp(x)).
\]

(ii) We now obtain a modulus of continuity for \( f^\sharp \). Fix \( x \in \mathbb{R}^d \) and let \( y \) vary in \( B(x,1) \). Consider first \( r \leq \phi(x - y) \). Then, letting \( \tau_{x-y}f(z) = f(z - (x - y)) \),

\[
\left| \int_{B(x,r)} |f| - \int_{B(y,r)} |f| \right| \leq \int_{B(x,r)} |f - \tau_{x-y}f| \leq \sup_{x' \in B(x,\phi(x - y)), y' \in B(y,\phi(x - y))} |f(x') - f(y')| \leq \text{osc}_{B(x,2\phi(x - y))} f(x).
\]

Since \( f \) is continuous in an neighbourhood of \( x \) this quantity goes to zero when \( y \to x \).

Consider now \( r > \phi(x - y) \). Letting \( r' = r + |x - y| \) so that \( B(x, r') \supset B(y, r) \),

\[
\int_{B(x,r')} |f| \geq \frac{\text{Vol}(B(y,r))}{\text{Vol}(B(x,r'))} \int_{B(y,r)} |f|
\]

hence

\[
(1 + r'^2)^\alpha \int_{B(x,r')} |f| - (1 + r^2)^\alpha \int_{B(y,r)} |f| \geq \left( \frac{r}{r + |x - y|} \right)^d - 1 \int_{B(y,r')} |f| \geq - \left( \frac{|x - y|}{\phi(x - y)} \right) f^\sharp(y).
\]

Similarly, with \( r'' = r - |x - y| \) (note that \( r'' > |x - y| \) by hypothesis),

\[
(1 + r'')^\alpha \int_{B(y,r')} |f| - (1 + r''^2)^\alpha \int_{B(x,r'')} |f| \geq - \left( \frac{|x - y|}{\phi(x - y)} \right) f^\sharp(x).
\]

Thus, with \( M = \sup_{B(x,1)} f^\sharp \), \( M < \infty \) by (i),

\[
\sup_{r}(1 + r^2)^\alpha \int_{B(y,r)} |f| - \sup_{r}(1 + r^2)^\alpha \int_{B(x,r)} |f| \leq \max \left\{ \sup_{r \leq \phi(x-y)} (1 + r^2)^\alpha \left( \int_{B(y,r)} |f| - \int_{B(x,r)} |f| \right), \right.
\]

\[
\left. \sup_{r > \phi(x-y)} \left( (1 + r^2)^\alpha \int_{B(y,r)} |f| - (1 + r^2)^\alpha \int_{B(x,r')} |f| \right) \right\} \leq \max \left( \text{osc}_{B(x,\phi(x-y)) \cup B(y,\phi(x-y))} f(x), M \frac{|x - y|}{\phi(x - y)} \right).
\]
Exchanging $x$ and $y$ gives the same inequality. Hence we have shown that $f^u$ is continuous, and obtained more precisely that, for every function $\phi$ satisfying the above hypotheses,

$$\text{osc}_{B(x,u)} f^u \leq \max(\text{osc}_{B(x,2\phi(u))}(f), M \frac{u}{\phi(u)}), \quad u \in (0, 1).$$  \hspace{1cm} (3.24)$$

In particular, choosing $\phi(u) = 2u^{1/(1+\beta)}$ if $\beta > 0$ yields $\text{osc}_{B(x,u)} f^u \leq u^{\beta/(1+\beta)}$, so $f^u$ is $\frac{\beta}{1+\beta}$ Hölder continuous.

(iii) Let us finally obtain a modulus of continuity for $f^\ast$. The proof is a slightly different from (ii) because the support of the heat kernel is the whole space; hence we must deal with the queue of the exponential $e^{-u^2 / t}$ for $u \gg \sqrt{t}$. Assume first $\sqrt{t} \leq \phi(x - y)$. Then

$$\left| e^{t\Delta} f(x) - e^{t\Delta} f(y) \right| \leq I_1 + I_2 + I_3,$$

where

$$I_1 = \int_{B(x,\alpha)} \frac{e^{-|x-y|^2/2t}}{(2\pi t)^{d/2}} |f(z) - \tau_{x-y} f(z)|dz \leq \text{osc}_{B(x,\alpha)} f = \text{osc}_{B(x,\alpha)} f$$

and

$$I_2 = 2^d \int_{\mathbb{R}^d \setminus B(x,\alpha)} \frac{e^{-|x-y|^2/2t}}{(2\pi t)^{d/2}} |f(z)|dz = \int_{B(x,\alpha \setminus 1)} \frac{e^{-|x|^2/2t}}{(2\pi t)^{d/2}} f(x + u)du$$

$$\leq 2^d \left( \frac{e^{-|x|^2/8t}}{(8\pi t)^{d/2}} |f(x + u)|du \right) \cdot e^{-3\alpha^2 (x-y)^2 / 8t}$$

$$\leq e^{-\frac{1}{1+\alpha(t)}} f^\ast(x),$$

while $I_3 = \int_{B(y,\alpha \setminus 1)} \frac{e^{-|x-y|^2/2t}}{(2\pi t)^{d/2}} |f(z)|dz$ is similar to $I_2$. The exponential factor in front of $f^\ast$ decreases to 0 when $y \to x$.

Assume now $\sqrt{t} > \phi(x - y)$. Then $|x - z|^2 \leq |x - y|^2 + |y - z|^2 + 2|x - y||y - z| \leq (1 + \varepsilon)|y - z|^2 + (1 + e^{-1})(x - y)^2$ for every $z \in \mathbb{R}^d$ and $\varepsilon > 0$. Choose $\varepsilon = \frac{|x-y|}{\phi(x-y)} < \frac{1}{2}$ so that $\frac{(1+e^{-1})|x-y|^2}{t} \leq \varepsilon$. Letting $t' = t(1 + \varepsilon)$, one obtains

$$(1 + t')^\alpha \int \frac{e^{-|x-y|^2/2t'}}{(2\pi t')^{d/2}} |f(z)|dz - (1 + t)^\alpha \int \frac{e^{-|x-y|^2/2t}}{(2\pi t)^{d/2}} |f(z)|dz$$

$$\geq -\varepsilon(1 + t)^\alpha \int \frac{e^{-|y-z|^2/2t}}{(2\pi t)^{d/2}} |f(z)|dz \geq -\varepsilon f^\ast(x).$$

Exchanging $x$ and $y$ gives a similar inequality, and one concludes to (3.24) as in (ii) by noting that $e^{-\frac{t}{1+\alpha(t)}} \leq M \frac{u}{\phi(u)}$ for $u < 1$.

\[\square\]

A result in the same direction is

**Lemma 3.4** Let $f \in \mathcal{H}_0^u \cap C(\mathbb{R}^d)$. Then, for every $t > 0$ and $\alpha > 0$,

$$|e^{t\Delta} f(x) - f(x)| \leq \text{osc}_{B(x,t^{\alpha/2})} f + e^{-\frac{1}{1+\alpha(t)}} f_0^\ast(x).$$  \hspace{1cm} (3.29)$$

Consequently, $e^{t\Delta} f \to_{t \to \infty} f$ uniformly on every compact.
Proof. (3.29) follows directly from the inequality
\[
|e^{\Delta} f(x) - f(x)| \leq \int_{y \in B(x, \varepsilon)} e^{-|x-y|^2/4t} |f(y) - f(x)| dy + \int_{y \in B(x, \varepsilon)^c} e^{-|x-y|^2/4t} (|f(y)| + |f(x)|) dy.
\]
(3.30)
Taking \(\varepsilon = t^\alpha\) with \(\alpha < \frac{4}{2}\), using the local boundedness of \(f_\alpha^n\) (proved in the previous lemma) and letting \(t \to 0\) yields the uniform convergence on a compact set. \(\square\)

Finally, we shall later on need to approximate functions in \(\mathcal{H}_\alpha^0\) by functions with compact support, and use the following lemma:

**Lemma 3.5** Let \(\chi : \mathbb{R}^d \to \mathbb{R}_+\) be a smooth 'bump' scale 1 function, i.e. \(\chi|_{B(0,1)} = 1, \chi|_{\mathbb{R}^d \setminus B(0,2)} = 0\). Denote by \(\chi_n(x) = \chi(\frac{x}{n})\) its dilations for \(n \in \mathbb{N}\). Then, if \(f \in \mathcal{H}_\alpha^0\), the functions \(f_n := f \cdot \chi_n, n \geq 1\) also belong to \(\mathcal{H}_\alpha^0\), and \((f_n)^* \to f^*\), \((f_n)^\sharp \to f^\sharp\) uniformly on every compact.

Uniform convergence implies in particular that \(f_n^*, f_n^\sharp\) are continuous, but does not give any bound on the modulus of continuity.

**Proof.** Let \(K \subset \mathbb{R}^d\) compact containing 0. We skip \(\alpha\) lower indices as in the previous lemma and prove that \((f_n)^* \to f^*\) uniformly on \(K\). Let \(B(0, r)\) a ball containing \(K\), and assume \(n \gg r\). Then \(|y-x|^2 \leq |t|^2\) for all \(t, x, y\) with \(t > 0, x \in K, |y| > n\). Hence
\[
0 \leq e^{\Delta}|f(x) - e^{\Delta}|f_n(x)| \leq 2^{d/2} \int_{|y| > n} e^{-|y|^2/4t} |f(y)| dy = 2^{d/2} e^{2\Delta}(|f| - |f_n|)(0),
\]
from which uniform convergence follows provided simple convergence holds at 0. But
\[
f^*(0) = \sup_{t > 0} (1 + t)^\alpha e^{\Delta}(\sup_n |f_n|)(0) = \sup_{t > 0} (1 + t)^\alpha (e^{\Delta}|f_n|)(0) = \lim_{n} (f_n)^*(0)
\]
by monotone convergence.

The proof for \((f_n)^\sharp\) is similar: let us just state that
\[
\int_{B(x, R)} (|f| - |f_n|) = 0
\]
(3.33)
if \(B(x, R) \subset B(0, n)\), and
\[
\int_{B(x, R)} (|f| - |f_n|) \leq \left(\frac{R + r}{R}\right)^d \int_{B(0, R+r)} (|f| - |f_n|)
\]
(3.34)
otherwise. Details are left to the reader. \(\square\)

We may now finally write down our pointwise parabolic estimates:

**Lemma 3.6 (pointwise parabolic estimates)** Let \(f \in \mathcal{H}_\alpha^0\). For every \(k \geq 0\),
\[
|\nabla^k e^{\Delta} f(x)| \leq t^{-\alpha-k/2} f_n^*(x)
\]
(3.35)
and
\[
(\nabla^k e^{\Delta} f)^\sharp_n(x) \leq t^{-\alpha} f_n^*(x)
\]
(3.36)
Proof.
By differentiating \( k \) times the computations leading to (3.11), one gets
\[
|\nabla^k e^{\tau \Delta} f(x)| \leq \int dr \left( \frac{r}{t} \right)^{k+1} \frac{e^{-r^2/2t}}{(2\pi t)^{d/2}} \int_{B(x,r)} dy |f(y)|
\]
\[
\leq f_0^\alpha(x) \int dr \left( \frac{r}{t} \right)^{k+1} \frac{e^{-r^2/2t}}{(2\pi t)^{d/2}} (1 + r^2)^{-\alpha} \text{Vol}(B(x,r))
\]
\[
\leq t^{-\alpha-k/2} f_0^\alpha(x) \int dr \frac{r}{t} \frac{e^{-r^2/2t}}{(2\pi t)^{d/2}} \text{Vol}(B(x,r)) = t^{-\alpha-k/2} f_0^\alpha(x). \quad (3.37)
\]
In particular, if \( \alpha = 0 \), one gets
\[
(\nabla^k e^{\tau \Delta} f)_0^\alpha(x) \leq t^{-k/2} (f_0^\alpha)_0^\alpha(x) = t^{-k/2} \sup_{\tau,\sigma > 0} (1 + \tau)^\alpha e^{(\tau+\sigma)\Delta} |f|(x) \leq t^{-k/2} f_0^\alpha(x). \quad (3.38)
\]

In the sequel we restrict for simplicity to the case \( \alpha = 0 \). All results below are easily adapted to the case \( \alpha > 0 \) or to similar functional spaces with pointwise bounds of the form \( |||f||(x) = \sup_{\tau > 0} F(\tau, e^{\tau \Delta} f)(x) \).

### 3.2 The comparison principle

We now want to use as initial condition of (3.1) functions \( h_0 \) such that \( h_0 \in \mathcal{H}^1 \cap C(\mathbb{R}^d) \), where
\[
\mathcal{H}^1 := \{ h_0 \in L^1_{\text{loc}}(\mathbb{R}^d) \mid e^{\lambda |h_0|} \in \mathcal{H}^0 \}. \quad (3.39)
\]
The comparison to the linear heat equation (see subsection 2.1) actually suggests to consider initial conditions in the unpleasant-looking space,
\[
\tilde{\mathcal{H}}^1 := \{ h_0 \in L^1_{\text{loc}}(\mathbb{R}^d) \mid e^{\lambda h_0}, h_0 \in \mathcal{H}^0 \}
\]
However, by Jensen’s inequality, \( e^{\lambda |h_0|^\alpha} \leq (e^{\lambda |h_0|})^\alpha \), so \( \mathcal{H}^1 \subset \tilde{\mathcal{H}}^1 \). Note that the definition is compatible with that of \( \mathcal{H}^0 \) in the previous paragraph, in the sense that \( \frac{1}{\lambda} (e^{\lambda |h_0|} - 1) \to |h_0(x)| \) when \( \lambda \to 0 \). Also, by Jensen’s inequality, \( \mathcal{H}^1 \) is for \( \lambda > 0 \) a convex subset (but not a vector subspace) of \( \mathcal{H}^0 \), and
\[
|||h_0|||_{\mathcal{H}^1}(x) := \frac{1}{\lambda} \sup_{\tau > 0} \ln \left( (e^{\lambda \Delta} e^{\lambda |h_0|})(x) \right) = \frac{1}{\lambda} \ln \left( (e^{\lambda |h_0|})^\alpha(x) \right)
\]
defines a family of pointwise ”quasi-norms”, in the sense that
\[
|||f|||_{\mathcal{H}^1}(x) \leq |||f|||_{\mathcal{H}^1}(x) \quad (\lambda \leq \lambda');
\]
\[
|||\mu f|||_{\mathcal{H}^1}(x) \leq |\mu| |||f|||_{\mathcal{H}^1}(x) \quad (\mu \in \mathbb{R})
\]
(the last inequality is actually an equality); \[
|||f_1 + f_2|||_{\mathcal{H}^1}(x) \leq \frac{1}{p_1} |||p_1 f_1|||_{\mathcal{H}^1}(x) + \frac{1}{p_2} |||p_2 f_2|||_{\mathcal{H}^1}(x) \quad (p_1, p_2 \geq 1, \frac{1}{p_1} + \frac{1}{p_2} = 1). \quad (3.44)
\]
We then expect the solution of \((3.1)\) to be "uniformly bounded in \(\mathcal{H}^1\)”, at least locally in time (thus allowing for further generalizations to equations with time-dependent coefficients), and thus to lie for all \(T > 0\) in the functional space
\[
\mathcal{H}^1([0, T]) := \{ h \in L^1_{\text{loc}}(\mathbb{R} \times \mathbb{R}^d) \mid \forall x \in \mathbb{R}^d, \sup_{t \in [0, T]} (e^{i|h|})^*(x) < \infty \}.
\]

As mentioned previously, the comparison principle in its different forms usually requires as a cornerstone assumption the boundedness of the solutions. However, various authors have proved ad hoc comparison principles for PDE’s with unbounded coefficients; the solution lies in functional spaces including functions growing at infinity. The KPZ equation is a very particular class of Hamilton-Jacobi-Bellman equations for which a comparison principle holds under quadratic growth conditions, see Ito [30], Da Lio-Ley [22], [23]. Lemma 2.1 and Theorem 2.1 in [22] state the following in our case:

**Proposition 3.7 [22]**

Let \( \bar{U} \in USC([0, T] \times \mathbb{R}^d) \) (resp. \( \bar{U} \in LSC([0, T] \times \mathbb{R}^d) \)) be a viscosity sub-solution (resp. super-solution) of \((3.7)\). Assume there exists \( C > 0 \) such that \( |\bar{U}(t, x)|, |\bar{\bar{U}}(t, x)| \leq C(1 + |x|^2) \) for all \( x \in \mathbb{R}^d, t \leq T \). Then \( \bar{U} \leq \bar{\bar{U}} \) in \([0, T] \times \mathbb{R}^d\).

A continuous function \( h_0 \) with quadratic growth at infinity, \( |h_0(x)| \leq 1 + |x|^2 \), is in general not in \( \mathcal{H}^1 \) for any \( \alpha \geq 0 \) or \( \lambda \geq 0 \). Conversely, a function in \( \mathcal{H}^1 \), \( \alpha, \lambda \geq 0 \) may grow arbitrarily fast in small domains \( \Omega_n, n \to \infty \) with \( d(0, \Omega_n) \to_{n \to \infty} \infty \) provided the Lebesgue measure of \( \Omega_n \) decreases to zero fast enough. On the other hand, since the supremum of \( n \) i.i.d. random variables grows like \( O(\sqrt{\log n}) \), one does expect random initial data \( h_0 \) to have a.s. quadratic growth at infinity. Actually, if \( \varepsilon > 0 \), then a.s. a random initial condition grows more slowly at infinity than \( |x|^\varepsilon \). Thus the above comparison principle holds for such data, and the existence of a sub-solution and a super-solution in this class of functions entails by Perron’s method the existence and unicity of a viscosity solution of \((3.1)\).

It seems however much more natural in our setting to prove a comparison principle for functions in \( \mathcal{H}^1 \) since the bounds one expects for the solution will depend on the pointwise maximal estimates \( (h_0)^* \) and \( (e^{ih_0})^* \) (on the contrary, solutions are expected to have a finite explosion time for initial conditions with quadratic growth, showing that this is in some sense too large a functional space). As it happens, we get such a comparison principle, but only for solutions in spaces \( \mathcal{H}^{\lambda'} \) with parameter \( \lambda' \geq 2\lambda \) (our proof does not hold for \( \lambda' = \lambda \)). In some sense \( \mathcal{H}^0 \) is the largest natural functional space for globally defined solutions of parabolic PDE’s. We conjecture that this extension of the viscosity solution theory to spaces modelled after \( \mathcal{H} \) (like \( \mathcal{H}^1 \) in the present case) is valid and of interest not only for the KPZ equation, but probably much beyond for many nonlinear parabolic PDE’s.

Let us state our main theorem, following closely the strategy of Da Lio and Ley:

**Theorem 3.1 (comparison principle)** Let \( \bar{U} \in USC([0, T] \times \mathbb{R}^d) \cap H^2_{\lambda}([0, T]) \) (resp. \( \bar{\bar{U}} \in LSC([0, T] \times \mathbb{R}^d) \cap H^2_{\lambda}([0, T]) \)) be a viscosity sub-solution (resp. super-solution) of \((3.7)\) for some \( \alpha \geq 0 \). Then \( \bar{U} \leq \bar{\bar{U}} \) in \([0, T] \times \mathbb{R}^d\).

The proof is very similar to [22], section 2. The essential element is the following lemma.
Lemma 3.8 Let \( U \in USC([0, T] \times \mathbb{R}^d) \cap H^{2,1}([0, T]) \) be a sub-solution, and \( \tilde{U} \in LSC([0, T] \times \mathbb{R}^d) \cap H^{2,1}([0, T]) \) be a super-solution of (3.7). Then \( \Psi_{\mu} := U - \mu \tilde{U}, \mu \in (0, 1) \) is a sub-solution of the quadratic KPZ equation,

\[
\partial_t \Psi = \Delta \Psi + \frac{\lambda}{1 - \mu} |\nabla \Psi|^2.
\]

Note that \( (U, \tilde{U} \in H^{2,1}([0, T])) \Rightarrow \Psi_{\mu} \in H^{1,1}([0, T]) \) (hence our choice of parameter, \( \lambda' \geq 2\lambda \), see discussion above).

**Proof.**

If \( U, V \in C^{1,2} \) then the proof is elementary. First,

\[
\partial_t \Psi_{\mu} \leq \Delta \Psi_{\mu} + \lambda (V(U) - \mu V(\tilde{U})).
\]

Then, since \( V \) is convex,

\[
V(a) \leq \mu V(b) + (1 - \mu) V\left(\frac{a - \mu b}{1 - \mu}\right), \quad a, b \in \mathbb{R}^d.
\]

Finally, applying this inequality to \( a = \nabla U, b = \nabla \tilde{U} \), and using Assumption [14] (4), \( V(y) \leq y^2 \), yields the result.

Otherwise the proof is essentially a very particular case of [22], Lemma 2.2. Let us reproduce the main arguments for the sake of the reader. Let \( \psi \in C^2([0, T] \times \mathbb{R}^d) \) and \( (\bar{i}, \bar{x}) \) a strict local maximum of \( \Psi - \psi \); we must prove that \( \partial_t \psi(\bar{i}, \bar{x}) \leq \nu \Delta \psi(\bar{i}, \bar{x}) + \frac{1}{1 - \mu} |\nabla \psi(\bar{i}, \bar{x})|^2 \). This is done by the standard doubling of variables argument, namely, we let \( \Theta(t, x, y) := \psi(t, x) + \frac{|x - y|^2}{2\nu} \), and \( M_{\varepsilon} = (\Psi - \Theta)(t_{\varepsilon}, x_{\varepsilon}) \) be the maximum of \( \Psi - \Theta \) in a small ball centered at \( (\bar{i}, \bar{x}) \); it is known that \( |x_{\varepsilon} - y_{\varepsilon}| = o(\varepsilon) \) and \( M_{\varepsilon} \rightarrow_{\varepsilon \to 0} \Psi(\bar{i}, \bar{x}) - \psi(\bar{i}, \bar{x}) \). By Theorem 8.3 in the User’s guide, see in [22] for the details of computations, one finds, exploiting the hypotheses on \( U, \tilde{U} \),

\[
\partial_t \psi(t_{\varepsilon}, x_{\varepsilon}) + H(t_{\varepsilon}, x_{\varepsilon}, \nabla \psi(t_{\varepsilon}, x_{\varepsilon}) + p_{\varepsilon}, X) - \mu H(t_{\varepsilon}, x_{\varepsilon}, \frac{p_{\varepsilon}}{\mu}, Y) \leq 0
\]

where \( p_{\varepsilon} = 2H_{x_{\varepsilon}} \), \( H(x, t, p, X) := -\lambda V(p) - \nu \text{Tr}(X) \), and \( X, Y \) are symmetric \( d \times d \) matrices, depending on \( \varepsilon \) and on a parameter \( \rho > 0 \), such that \( \text{Tr}(X - Y) \leq \Delta \psi(t_{\varepsilon}, x_{\varepsilon}) + O(\rho/\varepsilon^4) \). Hence

\[
\partial_t \psi(t_{\varepsilon}, x_{\varepsilon}) \leq \nu \Delta \psi(t_{\varepsilon}, x_{\varepsilon}) + \lambda \left[ V(\nabla \psi(t_{\varepsilon}, x_{\varepsilon}) + p_{\varepsilon}) - \mu V\left(\frac{p_{\varepsilon}}{\mu}\right) \right] + O(\rho/\varepsilon^4).
\]

Letting \( \rho \to 0 \) and using (3.48) as above yields

\[
\partial_t \psi(t_{\varepsilon}, x_{\varepsilon}) \leq \nu \Delta \psi(t_{\varepsilon}, x_{\varepsilon}) + \frac{\lambda}{1 - \mu} |\nabla \psi(t_{\varepsilon}, x_{\varepsilon})|^2.
\]

Finally, letting \( \varepsilon \to 0 \) gives the result. \( \square \)

We shall also need a non-standard comparison lemma for the linear heat equation:

Lemma 3.9 Let \( U \in USC([0, T] \times \mathbb{R}^d) \cap H^{0,1}([0, T]) \) (resp. \( \tilde{U} \in LSC([0, T] \times \mathbb{R}^d) \cap H^{0,1}([0, T]) \)) be a viscosity sub-solution (resp. super-solution) of the linear heat equation for some \( \alpha \geq 0 \). Then \( \underline{U} \leq \tilde{U} \) in \([0, T] \times \mathbb{R}^d \).
In other words, Theorem 3.1 holds for $\lambda = 0$.

**Proof.** Since the equation is linear, we may (by replacing $\bar{U}$ with $(U - \bar{U}) - e^{\lambda t}(U_0 - \bar{U}_0)$) assume that $\bar{U} = 0$ and $U_0 = 0$. Now, we have no bound at infinity available for $U_0$, and the classical maximum principle does not hold. Instead we choose a smooth function $\chi \geq 0$ with $\chi|_{(-\infty, 1]} \equiv 1$, supp$(\chi) \subset (-\infty, 1]$, define $X_n(x) := \chi(|x| - n)$ and obtain the following inequality for $U_n := \bar{U}X_n$,

$$(\partial_t - \Delta) U_n + \Delta X_n U + 2\nabla X_n \cdot \nabla U \leq 0. \quad (3.52)$$

Assume that $U \in C^{1,2}$ is a classical sub-solution to begin with. Then, since $U_n$, $\Delta X_n U$ and $\nabla X_n \cdot \nabla U$ are bounded, the classical comparison principle entails

$$U_n(t, x) \leq -\int_0^t ds e^{(s-t)\lambda} (\Delta X_n U(s)) - 2\int_0^t ds e^{(s-t)\lambda} (\nabla X_n \cdot \nabla U(s))$$

$$= \int_0^t ds e^{(s-t)\lambda} (\Delta X_n U(s)) - 2\int_0^t ds \nabla e^{(s-t)\lambda} \cdot (\nabla X_n U(s))$$

$$\leq \int_0^t ds e^{(s-t)\lambda} (|\tilde{X}_n U(s)|) + \int_0^t ds |\nabla e^{(s-t)\lambda} (\nabla X_n U(s))|, \quad (3.53)$$

where $\tilde{X}_n = \max(|\nabla X_n|, |\Delta X_n|)$. Now $\sum_n \tilde{X}_n \leq 1$, so (by the pointwise parabolic estimates)

$$\sum_n \int_0^t ds e^{(s-t)\lambda} |\tilde{X}_n U(s)| \leq \int_0^t ds e^{(s-t)\lambda} |\bar{U}(s)| \leq T \sup_{s \in [0, T]} |\bar{U}(s)|. \quad (3.54)$$

Hence (for $x$ fixed) $\int_0^t ds e^{(s-t)\lambda} |\tilde{X}_n U(s)| \rightarrow \bar{U} \rightarrow \infty 0$. Lemma 3.6 yields the same bound as (3.54), with $T$ replaced by $\sqrt{T}$, for the term with the gradient.

The above proof does not seem to extend to functions in $USC([0, T] \times \mathbb{R}^d) \cap H^0([0, T])$ by a density argument (in particular, if $\chi$ is a smooth, positive ‘bump’ function, then $\chi \ast U$ is a smooth sub-solution in the classical sense if $U$ is since $(\partial_t - \Delta)(\chi \ast U) = \chi \ast (\partial_t - \Delta)U \leq 0$, but not in the viscosity sense in general if $U$ is only upper-semicontinuous). Instead we use another truncation argument, which could also have been used in the classical case. We fix $x \in \mathbb{R}^d$ and let $n \rightarrow \infty$ as above. Since $U \in H^0$, it is locally bounded, so the function $U \ast \psi$ is a bounded sub-solution of the heat equation on $[0, T] \times B(0, n)$ with boundary value $U|_{\partial B(0, n)}$. Thus the classical maximum principle and Green’s formula imply that

$$U(t, x) \leq \int_0^T ds \int_{\partial B(0, n)} \nabla_n G_n(t, x; s, y) U(s, y), \quad t \leq T, \; x \in B(0, n) \quad (3.55)$$

where $\nabla_n$ is the normal derivative and $G_n$ is the Green function of the heat equation on $\mathbb{R}^n \times B(0, n)$. By standard estimates, $|\nabla G_n(t, x; s, y)| \leq (t - s)^{-\frac{d}{2}} G(t, x; s, y)$ if $y \in \partial B(0, n)$, where $G(t, x; s, y)$ is the usual heat kernel on $\mathbb{R} \times \mathbb{R}^d$. One has thus obtained an estimate very similar to (3.53), and the end of the proof is the same.

**Proof of Theorem 3.1.**

By Lemma 3.8 $\partial_t \Psi_{\mu} \leq \Delta \Psi_{\mu} + \frac{1}{1 - \mu} |\nabla \Psi_{\mu}|^2$. Equivalently, $(\partial_t - \Delta) \left( e^{\frac{|\Psi_{\mu}|^2}{4(1 - \mu)}} \right) \leq 0$. By Lemma $3.9$, $\Psi_{\mu}(t, x) \leq \frac{1 - \mu}{\lambda} \ln \left[ \int e^{-\lambda |x - y|^2 / 2(1 - \mu)} e^{\beta_0 |y|} dy \right]. \quad (3.56)$

Letting $\mu \rightarrow 1$, one finds $\Psi_{\mu} \leq 0$. \qed
3.3 Bounds for the solution

Let $h_0 \in \mathcal{H}^{2,1} \cap C(\mathbb{R}^d)$. Then $\mathring{h}_t := -e^{\nu\Delta} h_0$ is a sub-solution, and $\breve{h}_t := \frac{1}{t} \ln(e^{\nu\Delta} e^{\frac{t}{\nu} a})$ a supersolution of (3.1), and the pointwise parabolic estimates, together with Jensen’s inequality, imply that $\mathring{h}, \breve{h} \in C([0, T] \times \mathbb{R}^d) \cap \mathcal{H}^{2,1}([0, T])$. Perron’s method (see User’s guide [21], Theorem 4.1), in combination with the comparison principle of the previous paragraph, shows that

$$h(x) := \sup \{ \mathring{h}(x) \mid \mathring{h} \leq \breve{h} \}$$

is the unique viscosity solution in $C([0, T] \times \mathbb{R}^d) \cap \mathcal{H}^{2,1}([0, T])$ for every $T > 0$. We simply call $h$ the solution on $[0, T]$ of (3.1) with initial condition $h_0$.

The analogue of the space $\mathcal{W}^{1,\infty}$ in our setting is

$$\mathcal{W}_{loc}^{1,\infty,2,1} := \left\{ h_0 \in \mathcal{H}^{2,1} \mid \forall x \in \mathbb{R}^d, \sup_{\epsilon \in B(0, 1) \setminus \{0\}} \exp \left\{ 2\lambda \frac{|h_0(\epsilon + \cdot) - h_0(\cdot)|}{|\epsilon|} \right\}^* (x) < \infty \right\}. \quad (3.58)$$

Note that the choice of $B(0, 1)$ in the supremum in the above definition is largely irrelevant (any smaller neighbourhood of the origin would do, larger ones would just require a rescaling). Just like $\mathcal{H}^{2,1}$ before, $\mathcal{W}_{loc}^{1,\infty,2,1}$ is a convex subset of $C(\mathbb{R}^d)$, and

$$|||h_0|||_{\mathcal{W}_{loc}^{1,\infty,2,1}}(x) := |||h_0|||_{\mathcal{H}^{2,1}}(x) + \sup_{\epsilon \in B(0, 1)} \left|\left| |h_0(\epsilon + \cdot) - h_0(\cdot)| \right| \right|_{\mathcal{H}^{2,1}}(x) \quad (3.59)$$

defines a family of pointwise ”quasi-norms”.

A first easy result is:

**Lemma 3.10** Let $h$ be the viscosity solution on $[0, T]$ of (3.1) with initial condition $h_0 \in \mathcal{H}^{2,1} \cap C(\mathbb{R}^d)$. Then, for every $t \in [0, T]$, $$(e^{a|h_t|})^*(x) \leq (e^{a|h_0|})^*(x), \quad a \geq 1. \quad (3.60)$$

In particular,

$$|h_t(x)| \leq |||h_0|||_{\mathcal{H}^{2,1}}(x). \quad (3.61)$$

**Proof.** By the comparison principle, Theorem (3.1), $|h| \leq u$, where $u$ is the solution of the quadratic KPZ equation $\partial_t u = \Delta u + \lambda \nabla u^2$ with initial condition $|h_0|$. Then $e^{au}$ is a solution of the linear heat equation, hence (by Jensen’s theorem and pointwise parabolic estimates)

$$(e^{a|h_t|})^*(x) \leq (e^{a|h_0|})^*(x) = (e^{\Delta t} e^{a|h_0|})^*(x) \leq (e^{\Delta t} e^{a|h_0|})^*(x) \leq (e^{a|h_t|})^*(x) \quad (3.62)$$

for $a \geq 1$.

Thus $h$ extends to $t \in \mathbb{R}_+$ and satisfies (3.60) for arbitrary $t$.

3.4 Bounds for the gradient

We now assume $h_0 \in \mathcal{W}_{loc}^{1,\infty,2,1}$ or, better still, $h_0 \in \mathcal{W}_{loc}^{1,\infty,4,1}$.

**Lemma 3.11** (i) Assume $h_0 \in \mathcal{W}_{loc}^{1,\infty,2,1} \cap C(\mathbb{R}^d)$. Then the solution $h$ is classical for $t > 0$. Furthermore,

$$|\nabla h(t, x)| \leq |||h_0|||_{\mathcal{W}_{loc}^{1,\infty,2,1}}(x). \quad (3.63)$$

27
(ii) Assume \( h_0 \in \mathcal{W}^{1,\infty,1}_{loc} \cap C(\mathbb{R}^d) \). Then \( h_t \in \mathcal{W}^{1,\infty,1}_{loc} \) and
\[
e^{\varepsilon \| h_0 \|_{\mathcal{W}^{1,\infty,1}(x)}} \leq \left( \left( e^{4\varepsilon \| h_0 \|_{\mathcal{W}^{1,\infty,1}(x)}} \right)^{1/2} \right) \cdot \left( \left( e^{4\varepsilon \| h_0 \|_{\mathcal{W}^{1,\infty,1}(x)}} \right)^{1/4} \right) .
\]

**Proof.**

(i) Let \( \varepsilon \in B(0,1) \setminus \{ 0 \} \). By Lemma 3.8 \( \Psi_{\varepsilon}(t, \cdot) := h(t, \cdot + \varepsilon) - (1 - |\varepsilon|)h(t, \cdot) \) is a sub-solution of the KPZ equation \( \partial_t \psi = \Delta \psi + \frac{1}{2} |\nabla \psi|^2 \), hence \( t \mapsto (e^{4\lambda \Psi_{\varepsilon}(t, \cdot)})(x) \) is decreasing and
\[
\Psi_{\varepsilon}(t, x) \leq \frac{|\varepsilon|}{\lambda} \ln \left( e^{\Delta \left( e^{4\lambda |\nabla \Psi_{\varepsilon}(t, \cdot)|} \right)} \right) \leq \frac{|\varepsilon|}{2\lambda} \left\{ \ln e^{\Delta \left( e^{4\lambda |\nabla \Psi_{\varepsilon}(t, \cdot)|} \right)} \right\} + \ln e^{\Delta \left( e^{4\lambda |\nabla h_0|(x)} \right)} .
\]
Since \( h_0 \in \mathcal{W}^{1,\infty,2}_{loc} \), the functions \( \|h_0\|_{\mathcal{H}^{2,1}(x)} \) and \( \tilde{h}_0(x) := \exp \left( 2\lambda \left( \frac{|\nabla h_0(x)|}{\| h_0 \|_{\mathcal{H}^{2,1}(x)}} \right) \right) \) are locally bounded; exchanging the roles of \( x \) and \( x + \varepsilon \) gives a two-sided inequality for \( \Psi_{\varepsilon} \),
\[
|\Psi_{\varepsilon}(t, x)| \leq |\varepsilon| \sup_{y \in B(x,1)} \|h_0\|_{\mathcal{H}^{2,1}(y)} + \frac{|\varepsilon|}{2\lambda} \ln \sup_{y \in B(x,1)} \tilde{h}_0(y).
\]
Thus
\[
\frac{|h(t, x + \varepsilon) - h(t, x)|}{|\varepsilon|} \leq \frac{\ln \sup_{y \in B(x,1)} \tilde{h}_0(y)}{2\lambda} + \sup_{y \in B(x,1)} \|h_0\|_{\mathcal{H}^{2,1}(y)} + \sup_{y \in B(x,1)} |h(t, y)| .
\]
Letting \( \varepsilon \to 0 \), one shows that \( h \in \mathcal{W}^{1,\infty}_{loc} \). Now, fix a compact \( K \subset \mathbb{R}^d \) with smooth boundary. Solving the KPZ equation on \( \mathbb{R}^+ \times K \) with boundary condition \( h(t, \cdot) \big|_{\partial K} \) as in [5] and using the above a priori bound shows that the solution is classical in \( K \) for \( t > 0 \).

(ii) We introduce the following notations,
\[
\delta f(\cdot) := \frac{f(\cdot + \varepsilon) - f(\cdot)}{|\varepsilon|}, \quad \tilde{\delta} f(\cdot) := \frac{f(\cdot + \varepsilon) - (1 - |\varepsilon|)f(\cdot)}{|\varepsilon|} .
\]
Note that \( \delta h_t(\cdot) = \frac{1}{|\varepsilon|} \Psi_{\varepsilon}(t, \cdot) \) and \( \tilde{\delta} f = \delta f + f \). Then we get for every \( a \geq 1 \)
\[
\left( e^{4\lambda \tilde{\delta} h_t} \right)^*(x) = \left( e^{4\lambda \Psi_{\varepsilon}(t)} \right)^*(x) \leq \left( e^{4\lambda \Psi_{\varepsilon}(0)} \right)^*(x) = \left( e^{4\lambda \delta h_0} \right)^*(x) .
\]
Now we use twice Hölder’s inequality,
\[
\left( e^{2\lambda \tilde{\delta} h_t} \right)^*(x) \leq \left[ \left( e^{2\lambda \tilde{\delta} h_t} \right)^*(x) \right]^{1/2} \left[ \left( e^{2\lambda \delta h_0} \right)^*(x) \right]^{1/2} \leq \left[ \left( e^{2\lambda \tilde{\delta} h_t} \right)^*(x) \right]^{1/2} \left[ \left( e^{2\lambda \delta h_0} \right)^*(x) \right]^{1/2} \left[ \left( e^{4\lambda \delta h_0} \right)^*(x) \right]^{1/4} .
\]
Hence the result.
\[\Box\]

We now consider a rate \( V = V(y) \) satisfying assumption (2.15), i.e. behaving like \( y^2 \) for \( y \) small or large, and show how to generalize the conclusions of Proposition 2.3 (iii).
Let \( \chi \) be a 'bump' function as in Lemma 3.5 and \( \chi^{(n)}(x) := \chi \left( \frac{x}{n} \right) \).
Lemma 3.12 Let $h^{(n)}$ be the solution of the homogeneous KPZ equation with initial condition $h^{(n)}_0 = h_0(x)\chi^{(n)}(x)$. Then $h^{(n)}$ converges to $h$ uniformly on every compact.

Proof. It follows from Lemma 3.11 that $h^{(n)}$ is locally uniformly differentiable. Thus, by Ascoli’s theorem and the classical diagonal extraction procedure, one may construct a subsequence $h^{(n)}_{m}$ converging locally uniformly. By the stability principle for continuous viscosity solutions (see e.g. [7], Theorem 3.1), the limit is a solution of the KPZ equation with initial condition $h^0$. Since the solution is unique, we have shown that $h^{(n)}_{m} \rightarrow m \rightarrow \infty h$ in $C(\mathbb{R}_+ \times \mathbb{R}^d)$. Since the sequence $(h^{(n)})_n$ is pre-compact and all subsequences converge to $h$, the sequence $(h^{(n)})$ itself converges to $h$. \hfill \Box

Corollary 3.13 Let $V$ satisfy assumption (2.15), $yV'(y) - V(y) \geq Cy^2$. Then

$$|\nabla h_t(x)| \lesssim \left( \frac{||h_0||_{H^1}^d}{t} \right)^{1/2}.$$  \hfill (3.72)

Proof. By (2.16),

$$|\nabla h^{(n)}_t(x)| \lesssim \left( \frac{|h^{(n)}_t(x)|}{t} \right)^{1/2}, \quad x \in \mathbb{R}^d. \hfill (3.73)$$

By Lemma 3.10 $|h^{(n)}_t(x)| \leq ||h_0||_{H^1}$ of $V$. Hence, for every $\varepsilon \in B(0,1)$, $\frac{1}{|\varepsilon|} |h^{(n)}_t(x + \varepsilon) - h^{(n)}_t(x)| \leq \sup_{B(x,1)} \left( \frac{||h^{(n)}_0||_{H^1}^d}{t} \right)^{1/2}$. The corollary follows from letting first $n \rightarrow \infty$ and then $\varepsilon \rightarrow 0$. \hfill \Box

4 Bounds for the infra-red cut-off inhomogeneous equation

4.1 General philosophy of scale decompositions

In this section, we start our study of the inhomogeneous KPZ equation,

$$\partial_t \psi(t, x) = \nu \Delta \psi(t, x) + \lambda V(\nabla \psi(t, x)) + g(t, x) \hfill (4.1)$$

where $g(t, x)$ is a $C^1$ force term. For the time being, we only consider an infra-red cut-off version of this equation, see (1.2) or Definition 4.1 below. We only require here good scale-dependent averaging properties for $g$ (see precise assumptions below). For the complete study (to be developed in the further articles) we shall take for $g$ a regularized white noise, denoted by $\eta$.

The general motivation in the subsequent analysis is to get an effective scale separation mechanism. In other words, let $G$ be the Green kernel,

$$G : g \mapsto (Gg)(t) := \int_0^{+\infty} e^{\nu s} g(t - s)ds \hfill (4.2)$$

(called propagator in the physics literature). Eq. (4.1) is equivalent to the integral equation,

$$\psi = G(\lambda V(\nabla \psi) + g)). \hfill (4.3)$$

Now we want to write $G$ as a sum $G = \sum_{j \geq 0} G^j$ over scales, and similarly for the force term, $g = \sum_{j \geq 0} g^j$, in such a way that
(1) $\psi$ be well approximated by the sum $\sum_j \psi_j$, where $\psi_j$ is the solution of the integral equation

$$\psi_j = G^j(\lambda^j V(\nabla \psi_j) + g^j);$$  \hspace{1cm} (4.4)

and

(2) $\psi^j$ by the solution $\phi^j$ of the linearized equation, $\phi^j = G^j g^j$, at least for $\lambda$ small enough or $j$ large enough.

The approximations in (1) and (2) are responsible for the renormalization procedure in which $\lambda$ becomes the scale-dependent parameter $\lambda^j$, and $G^j$, $g^j$ also receive correction terms (see third article in the series, (3) *The multi-scale expansion in the small-field regime* for further explanations).

At this point we are not interested in the renormalization procedure and would like in principle to consider a single-scale equation such as (4.4),

$$\psi^j = G^j(\lambda V(\nabla \psi^j) + g^j)$$  \hspace{1cm} (4.5)

or simply (since a priori $G^j g^j \approx G^j g$)

$$\psi^j = G^j(\lambda V(\nabla \psi^j) + g).$$  \hspace{1cm} (4.6)

The easiest way to select fluctuations at time, resp. space distances of order $M^j$, resp. $M^{j/2}$ is to set

$$(G^j f)(t) = \int_{M^j}^{M^{j/2}} ds e^{\delta j^2} f(t - s)$$  \hspace{1cm} (4.7)

(see Definition 5.2 for a better definition, in which a smooth time cut-off is used instead). Coming back to $g = \eta$ to mimic the behaviour of the noisy KPZ equation, we are led to set $\phi^j = G^j \eta$, $\eta^j = (\partial_t - \Delta) \phi^j$. Recall

$$d_\phi := \frac{1}{2}(d - 1)$$  \hspace{1cm} (4.8)

is the *scaling dimension* of the solution of the Ornstein-Uhlenbeck or of the KPZ equation, see Introduction. It is proved in Appendix A that

$$E[\phi^j(t, x)\phi^{j'}(t', x')] \leq M^{-j-j'}(M^{-\max(j, j')}2^d \delta e^{-cM^{-\max(j, j')}|t-t'| - cM^{-\max(j, j')}|x-x'|}$$  \hspace{1cm} (4.9)

$$E[\eta^j(t, x)\eta^{j'}(t', x')] \leq M^{-2j-j'}(M^{-\max(j, j')}2^d \delta e^{-cM^{-\max(j, j')}|t-t'| - cM^{-\max(j, j')}|x-x'|}$$  \hspace{1cm} (4.10)

for some constant $c > 0$. Consider first the diagonal covariance ($j = j'$): since $\phi^j$ and $\eta^j$ are Gaussian, (4.9), (4.10) essentially mean that the following *scalings* hold,

$$\phi^j(t, x) = O(M^{-jd_\phi}), \quad \eta^j(t, x) = O((M^{-j})^{1+d_\phi})$$  \hspace{1cm} (4.11)

with random prefactors. The bounds in Appendix A also yield an order of magnitude of the gradients, with a supplementary $M^{-j/2}$ factor,

$$\nabla \phi^j(t, x) = O((M^{-j})^{\frac{1}{2}+d_\phi}), \quad \nabla \eta^j(t, x) = O((M^{-j})^{\frac{1}{2}+d_\phi}).$$  \hspace{1cm} (4.12)

For $j \neq j'$ one has an extra decaying exponential factor in $M^{-|j-j'|}$ which lies at the root of the scale separation mechanism.
Our actual mechanism of separation of scales (see third article in the series) is actually different. The reason is that the integral equation is a delay, non-local equation which does not satisfy at all the maximum principle, and we have no a priori bounds for its solutions. So we introduce instead in the sequel a very simple infra-red cut-off of scale $j$ for the propagator, namely, we replace $\nu \Delta$ by $\nu \Delta - M^{-j}$. Denoting by $G^{j \rightarrow}$ the Green kernel of the operator $\nu \Delta - M^{-j}$, one has the explicit formula

$$G^{j \rightarrow}(t, x; t', x') = 1_{t > t'} e^{-M^{-j} (t - t')} \frac{e^{-|x - x'|^2/2v(t-t')}}{(2\pi v(t-t'))^{d/2}},$$  \hspace{1cm} (4.13)$$

which makes apparent an exponential decay in time and space: since $\inf_{s>0} (\frac{|x|^2}{2s^2} + sM^{-j}) \approx M^{-j/2} |x - x'|$, 

$$G^{j \rightarrow}(t, x; t', x') \lesssim (t - t')^{-d/2} e^{-cM^{-j}(t-t') - cM^{-j/2}|x - x'|},$$  \hspace{1cm} (4.14)$$

for some constant $c > 0$. The idea is that $G^{j \rightarrow}$ is a good substitute for the sum $\sum_{k \leq j} G^k$. We also replace the force term $g$ by $g^j$ such that $g^j(t, x) = O((M)^{-j})$ as for $\eta^j$, see (4.11). Thus the new equation is the following.

**Definition 4.1 (inhomogeneous KPZ equation with scale $j$ infra-red cut-off)** The inhomogeneous KPZ equation with scale $j$ infra-red cut-off is

$$\partial_t \psi = (\nu \Delta - M^{-j}) \psi + \lambda V(\nabla \psi) + g^j. \hspace{1cm} (4.15)$$

The integral form of this equation is

$$\psi^j = G^{j \rightarrow}(V(\nabla \psi) + g^j). \hspace{1cm} (4.16)$$

Note that the kernel $G^{j \rightarrow}$ has no ultra-violet cut-off, in the sense that it behaves like the full Green kernel $G$ for time separations $|t - t'| \ll M^j$. It actually turns out that the solution $\psi^j$ of (4.15) has the correct scaling, $\psi^j(t, x) = O(M^{-jd})$, see (4.11), under appropriate assumptions on $g^j$ that we now proceed to write down.

### 4.2 Functional spaces of scale $j$

**Assumption 4.2** Let $g^j$ be a continuous function such that

(i) for all $t \in [0, T]$ and $x \in \mathbb{R}^d$,

$$\|g^j\|_{L^1([0, t], x)} := \frac{1}{\lambda} \sup_{0 < \delta \leq M^j} \int_0^t \int_0^t \left( e^{-M^{-j} \delta} \right)^p \ln \left( e^{\lambda M^{-j} \int_{(t - \delta)'}^t g^j(s) ds} \right)^* \left( x \right) \, ds \, dx < \infty; \hspace{1cm} (4.17)$$

(ii) for all $t \in [0, T]$ and $x \in \mathbb{R}^d$,

$$\frac{1}{\lambda} \sup_{0 < \delta \leq M^j, 0 < \delta \leq 1} M^{-j} \delta \int_0^t \int_0^t \left( e^{-M^{-j} \delta} \right)^p \ln \left( e^{\lambda M^{-j} \int_{(t - \delta)'}^t g^j(s) ds} \right)^* \left( x \right) \, ds \, dx \quad \text{is the average over the corresponding subinterval.} \hspace{1cm} (4.18)$$

where $\int_{(t - \delta)'}^t \int_0^t \left( e^{-p \delta t} \right)^* \, ds$ is the average over the corresponding subinterval.
The conditions \(0 \leq \delta t \leq M^j, 0 < |\delta x| \leq 1\) are arbitrary (any smaller neighbourhood of the origin would do). We denote by \(\mathcal{W}^{l,\infty,\lambda}_j((0, T]) \subset L^1_{\text{loc}}(\mathbb{R} \times \mathbb{R}^d)\) the convex subset of functions \(g\) satisfying (4.17) and (4.18) for all \(t \in [0, T]\) and \(x \in \mathbb{R}^d\). Note that the quantity in (4.18) is \(\sup_{0 < |\delta x| \leq 1} \| M^{j/2} g^{(\cdot + \delta x) - g(\cdot)} / |\delta x| \|_{L^1_j}([0, t], x)\). Summing (4.17) and (4.18) yields a pointwise "quasinequality",

\[
\|g\|_{\mathcal{W}^{l,\infty,\lambda}_j([0, t])} := \|g\|_{L^1_j}([0, t], x) + \sup_{0 < |\delta x| \leq 1} \| M^{j/2} g^{(\cdot + \delta x) - g(\cdot)} / |\delta x| \|_{L^1_j}([0, t], x). \tag{4.19}
\]

By Jensen’s inequality, \(\| M^{j/2} g^{(\cdot + \delta x) - g(\cdot)} / |\delta x| \|_{L^1_j}([0, t], x) \geq M^{j/2} \| g^{(\cdot + \delta x) - g(\cdot)} / |\delta x| \|_{L^1_j}([0, t], x)\).

Finally, we introduce a new, scale-dependent convex subspace \(\mathcal{W}^{l,\infty,\lambda}_j \subset \mathcal{W}^{l,\infty,\lambda}_j\).

**Definition 4.3 (the space \(\mathcal{W}^{l,\infty,\lambda}_j\))** Let

\[
\mathcal{W}^{l,\infty,\lambda}_j := \{h_0 \in \mathcal{H}^1, \forall x \in \mathbb{R}^d, \| h_0 \|_{\mathcal{W}^{l,\infty,\lambda}_j}(x) < \infty\}, \tag{4.20}
\]

where

\[
\| h_0 \|_{\mathcal{W}^{l,\infty,\lambda}_j} := \| h_0 \|_{\mathcal{H}^1} + \sup_{\varepsilon \in B(0, 1)} \| M^{j/2} \| h_0(\varepsilon + \cdot) - h_0(\cdot) \|_{\mathcal{H}^1}(x). \tag{4.21}
\]

Let us now see (as promised just before eqs. (4.17) and (4.18)) how to get a scaling of order \(O(M^{-\delta k})\) for \(\psi_j(t, x)\). Very roughly, letting \(C(s)\) be an order of magnitude of \(g^j(\cdot, s, \cdot)\), and \(C\) a global order of magnitude of \(g^j(\cdot, s)\), \(\| g^j \|_{L^1_j}([0, t], x) \approx \int_0^t ds e^{-M^{-j} \| C(t - s) \| M^j C}\). If \(\| g^j \|_{L^1_j}([0, T], x) = O(M^{-\delta k})\), then \(C = O((M^{-j})^{1 + \delta k})\) as in (4.11). To get also the correct scaling (4.12) of \(O((M^{-j})^{2 + \delta k})\) for \(\nabla \psi_j(t, x)\), we should therefore assume that

\[
\|g^j\|_{\mathcal{W}^{l,\infty,\lambda}_j([0, T])}(x) = O(M^{-\delta k}). \tag{4.22}
\]

This condition is made precise in Corollary 4.6. It is essentially correct up to the replacement of \(\lambda\) by \(4\lambda\).

### 4.3 Bounds

In order to take into account the right-hand side, we need the following Trotter-type formula. We use the following notations in this paragraph. The homogeneous nonlinear semi-group generated by the homogeneous KPZ equation (3.1) is denoted by \(\Phi^j(t)\), i.e. \(\Phi^j(t) h_0\) is the solution at time \(t\) of the homogeneous KPZ equation with initial condition \(h_0 \in \mathcal{W}^{l,\infty}\). In the next two lemmas we assume \(\psi_0 \in \mathcal{W}^{l,\infty}\), let \(g(t, x), (t, x) \in [0, T] \times \mathbb{R}^d\) be a function in \(C([0, T], C^2(\mathbb{R}^d))\) and define \(\bar{g}_k(s) : C(\mathbb{R}^d) \to C(\mathbb{R}^d), f \mapsto \bar{g}_k(s) f \) by \(\bar{g}_k(s) f(x) := \int_{k/n}^{(k+1)/n} g(u, x) du + f(x) (0 \leq k \leq n)\).

**Lemma 4.4 (Trotter formula)** Let, for \(k = 0, \ldots, n\,

\[
\psi_{k/n}(t) := \left(e^{-M^{-j}t/n} \Phi^j(t/n) \tau_k(t/n)\right) \left(e^{-M^{-j}t/n} \Phi^j(t/n) \tau_{k-1}(t/n)\right) \cdots \left(e^{-M^{-j}t/n} \Phi^j(t/n) \tau_0(t/n)\right) \psi_0. \tag{4.23}
\]

Then \(\psi_{k/n} \to_{n \to \infty} \psi_t\) in \(\| \cdot \|_{\infty}\)-norm.
Trotter formulas have been shown with some generality for non-linear monotonic operators acting on Hilbert spaces \[13\]. However here the natural spaces, \(L^\infty\) or \(W^{1,\infty}\), are not Hilbert spaces.

To show this lemma we therefore follow instead the proof of convergence of "viscous splitting" algorithms for the Navier-Stokes equation, as found in \[11\], §3.4, resting on their stability and consistency. The proof is however deferred after the following crucial bound in \(\mathcal{H}^\ell\) norm.

**Lemma 4.5** The following bounds hold,

\[
(e^{[p]f^n})^*(x) \leq \left[ (e^{[p]f^n})^*(x) \right]^{(p+1)/p} (e^{[p+1]f^n})^*(x) \left[ (e^{[p+1]f^n})^*(x) \right]^{1/p}.
\]  

(4.24)

and

\[
\left( e^{(2M)^{1/2} \|\psi^n\|_{\|\lambda^n\|_{H^\infty}}} \right)^* (x) \leq \left[ (e^{(2M)^{1/2} \|\psi^n\|_{\|\lambda^n\|_{H^\infty}}} \right)^*(x) \left( e^{(2M)^{1/2} \|\psi^n\|_{\|\lambda^n\|_{H^\infty}}} \right)^*(x) \left[ (e^{(2M)^{1/2} \|\psi^n\|_{\|\lambda^n\|_{H^\infty}}} \right)^*(x) \right]^{1/2}.
\]

\( (4.25)\)

where \(\|\cdot\|_{\|\lambda^n\|_{H^\infty}}(t, x)\) is the family of pointwise "quasi-norms" appearing in (4.17).

**Proof.** We shall rely on the following two elementary bounds,

\[
(e^{[p]f^n})^*(x) \leq (e^{[p]f^n})^*(x), \quad (f \in W^{1,\infty}, \ a \geq 1)
\]  

(4.26)

(see (3.62) and Hölder’s inequality

\[
(e^{[p]f^n})^*(x) \leq \left[ (e^{(p+1)\|\psi^n\|_{\|\lambda^n\|_{H^\infty}}} \right)^*(x) \left[ (e^{(p+1)\|\psi^n\|_{\|\lambda^n\|_{H^\infty}}} \right)^*(x) \right]^{1/p}.
\]  

(4.27)

Choose \(p\) in (4.27) defined by \(1/p = e^{-M^{-1/\ell(n)}}\). Note that \(1/p+1 \sim_{n \to \infty} M^{-1/\ell(n)}\). By (4.26),

\[
\left( e^{(2M)^{1/2} \|\psi^n\|_{\|\lambda^n\|_{H^\infty}}} \right)^*(x) = \left( e^{(2M)^{1/2} \|\psi^n\|_{\|\lambda^n\|_{H^\infty}}} \right)^*(x) \leq \left[ (e^{(2M)^{1/2} \|\psi^n\|_{\|\lambda^n\|_{H^\infty}}} \right)^*(x) \left[ (e^{(2M)^{1/2} \|\psi^n\|_{\|\lambda^n\|_{H^\infty}}} \right)^*(x) \right]^{1/2}.
\]

(4.28)

By induction on \(k\), this gives

\[
\left( e^{[p]f^n} \right)^*(x) \leq \left( e^{[p+1]f^n} \right)^*(x),
\]  

(4.29)

where

\[
A := \frac{1}{p+1} \sum_{k=1}^{n-1} \left( e^{-M^{-1/\ell(n)}} \right)^{k} \ln \left[ \left( e^{(p+1)\|\psi^n\|_{\|\lambda^n\|_{H^\infty}}} \right)^*(x) \right] \]

(4.30)

is asymptotically bounded by the quantity appearing in (4.17) with \(g = g^j\). Letting \(n \to \infty\) yields the first bound (4.24).

The proof of (4.25) is similar but requires a further elaboration on the arguments developed in the course of the proof of Lemma \[3.11\] to which we refer the reader. Let

\[ \delta f(\cdot) := \frac{f(\cdot + \delta x) - f(\cdot)}{|\delta x|}, \quad \tilde{\delta} f(\cdot) := \frac{f(\cdot + \delta x) - (1 - M^{-1/\ell} |\delta x|) f(\cdot)}{|\delta x|} \]

(4.31)
for $f \in C(\mathbb{R}^d)$. Note that $\tilde{\delta} f(\cdot) = \delta f(\cdot) + M^{-i/2} f(\cdot)$. Then

$$
(e^{\lambda M^{ij/2} \Phi_{t_0/n}^j(\delta(\cdot))})^*(x) = (e^{\lambda M^{ij/2} \Phi_{t_0/n}^j(\delta(\cdot) + M^{ij/2}(\cdot))})^*(x) \leq (e^{\lambda M^{ij/2} \tilde{\delta}(\delta(\cdot)))^*(x) \quad (4.32)
$$

for $\lambda \geq 1$. Hence

$$
\begin{align*}
\left(e^{2\lambda M^{ij/2} (\frac{p+1}{p})^2 \tilde{\delta}(\delta(\cdot))} \right)^*(x) & \leq \left(e^{2\lambda M^{ij/2} (\frac{p+1}{p})^2 \tilde{\delta}(\delta(\cdot))} \right)^*(x) \\
& = \left(e^{2\lambda M^{ij/2} (\frac{p+1}{p})^2 \tilde{\delta}(\delta(\cdot))} \right)^*(x) \\
& \leq \left[\left(e^{2\lambda M^{ij/2} (\frac{p+1}{p})^2 \tilde{\delta}(\delta(\cdot))} \right)^*(x) \right]^{1/2} \left[\left(e^{2\lambda M^{ij/2} (\frac{p+1}{p})^2 \tilde{\delta}(\delta(\cdot))} \right)^*(x) \right]^{1/2} \quad (4.33)
\end{align*}
$$

The end of the proof goes as before, up to the corrections necessitated by the difference between $\delta$ and $\tilde{\delta}$, which simply consist in using Holder’s inequality twice (hence the use of “quasi-norms” with index $4i1$),

$$
\begin{align*}
\left(e^{2\lambda M^{ij/2} (\frac{p+1}{p})^2 \tilde{\delta}(\delta(\cdot))} \right)^*(x) & \leq \left[\left(e^{4\lambda M^{ij/2} (\frac{p+1}{p})^2 \tilde{\delta}(\delta(\cdot))} \right)^*(x) \right]^{1/2} \left[\left(e^{4\lambda M^{ij/2} (\frac{p+1}{p})^2 \tilde{\delta}(\delta(\cdot))} \right)^*(x) \right]^{1/2} \quad (4.34)
\end{align*}
$$

and

$$
\left(e^{2\lambda M^{ij/2} \tilde{\delta}(\delta(\cdot))} \right)^*(x) \leq \left[\left(e^{4\lambda M^{ij/2} \tilde{\delta}(\delta(\cdot))} \right)^*(x) \right]^{1/2} \left[\left(e^{4\lambda M^{ij/2} \tilde{\delta}(\delta(\cdot))} \right)^*(x) \right]^{1/2} \quad (4.35)
$$

and

$$
\left(e^{2\lambda M^{ij/2} \tilde{\delta}(\delta(\cdot))} \right)^*(x) \leq \left[\left(e^{4\lambda M^{ij/2} \tilde{\delta}(\delta(\cdot))} \right)^*(x) \right]^{1/2} \left[\left(e^{4\lambda M^{ij/2} \tilde{\delta}(\delta(\cdot))} \right)^*(x) \right]^{1/2} \quad (4.36)
$$

\textbf{Proof of Lemma 4.4} Following [11], we split the proof into three points.

(i) (stability) As follows from the bounds proved in the above lemma, $\|\psi_{t/n}^{(n)}\|_{W^{1,\infty}} \leq C$, where $C$ is a constant depending on the initial condition and on $g$, but independent from $n$ and $k = 0, 1, \ldots, n - 1$.

(ii) (consistency) For every $n \geq 1$, $k = 0, 1, \ldots, n - 1$ and $s \in [0, t/n]$, we let $\psi_{s+k/l/n}^{(n)} := \tau_k(s)e^{-M^{ij/s}\Phi^4(s)\psi_{kl/n}^{(n)}} = \int_{s+k/l/n}^{s+k/l/n} g(u)du + e^{-M^{ij/s}\Phi^4(s)\psi_{kl/n}^{(n)}}$. Then

$$
\begin{align*}
\partial_t \psi_{s+k/l/n}^{(n)} & = g_{s+k/l/n} - M^{ij} e^{-M^{ij/s}\Phi^4(s)\psi_{kl/n}^{(n)}} + e^{-M^{ij/s}\Phi^4(s)\psi_{kl/n}^{(n)}} + \lambda V(\nabla(\Phi^4(s)\psi_{kl/n}^{(n)})) \\
& = g_{s+k/l/n} - \nabla \Delta M^{ij/s}(\psi_{s+k/l/n}^{(n)}) + \lambda V(\nabla(\Phi^4(s)\psi_{kl/n}^{(n)})) \\
& = \nabla \Delta M^{ij/s}(\psi_{s+k/l/n}^{(n)}) + \lambda V(\nabla(\Phi^4(s)\psi_{kl/n}^{(n)})) + O(\lambda \sup_{y \in C} |V(y)|) + O(\lambda C \sup_{y \in C} |V(y)| M^{ij/s}),
\end{align*}
$$

(4.37)
Therefore
\[ \partial_s \psi_{s+kt/n} = (\nu \Delta - M^{-j}) \psi_{s+kt/n} + \lambda V(\nabla \psi_{s+kt/n}) + g_{s+kt/n}, \] (4.38)

By the comparison principle,
\[ \|\psi_{s+kt/n}^{(n)} - \psi_{s+kt/n}\|_{\infty} \leq \|\psi_{kt/n}^{(n)} - \psi_{kt/n}\|_{\infty} + C' s \int_0^s \left( M^{-j} + M^{-j} \|g_{u+kt/n}\|_{\infty} + \nu \|\Delta g_{u+kt/n}\|_{\infty} \right) du \] (4.39)
(where \( C' \) depends like \( C \) on the initial condition and on \( g \)). Hence
\[ \|\psi_{s+kt/n}^{(n)} - \psi_{s+kt/n}\|_{\infty} \leq \|\psi_{kt/n}^{(n)} - \psi_{kt/n}\|_{\infty} + O(s^2). \] (4.40)

(iii) We apply (4.40) with \( s = t/n \). By induction on \( k \),
\[ \|\psi_{kt/n}^{(n)} - \psi_{t}\|_{\infty} \leq \frac{t^2}{n} \rightarrow 0, \] (4.41)
\[ \square \]

**Corollary 4.6** Assume \( \psi_0 \in \mathcal{W}^{1,\infty;4k}_j \) and \( g \in \mathcal{W}^{1,\infty;4k}_j([0,T]). \) Then the KPZ equation (4.15) has a unique viscosity solution \( \psi \) on \([0,T]\); the solution \( \psi \) is classical, is uniformly bounded in \( \mathcal{W}^{1,\infty;4k}_j \), and satisfies the following estimates on \([0,T] \times \mathbb{R}^d\),
\[ (e^{t} )^{(\psi)}(x) \leq \left( e^{\lambda t} \psi_0 \right)^{(x)} \mathbb{E}^{(\psi)}_{\lambda t}(x); \] (4.42)
\[ (e^{\lambda t} |\nabla \psi| )^{(x)} \leq \left[ e^{\lambda t} \psi_0 \right]^{(x)} \mathbb{E}^{(\psi)}_{\lambda t}(x); \] (4.43)
\[ \|\psi_t\|_{\mathcal{W}^{1,\infty;4k}_j(x)} \leq \|\psi_0\|_{\mathcal{W}^{1,\infty;4k}_j(x)} + \|g\|_{\mathcal{W}^{1,\infty;4k}_j([0,T])}(x). \] (4.44)
Therefore
\[ \psi_t(x) = O(M^{-j(\lambda \psi)}) \quad \nabla \psi_t(x) = O((M^{-j(\lambda \psi)})^{1/2}) \] (4.45)
if \( \|\psi_0\|_{\mathcal{W}^{1,\infty;4k}_j(x)} = O(M^{-j(\lambda \psi)}) \) and \( \|g\|_{\mathcal{W}^{1,\infty;4k}_j([0,T])}(x) = O(M^{-j(\lambda \psi)}) \).

**Proof.** The comparison principle Theorem 3.11 extends without any substantial modification to the inhomogeneous equation (4.15). By Lemma 4.4 the bounds proved in Lemma 4.5 for \( \psi^{(n)} \) hold for \( \psi \), as follows from taking the limit \( n \rightarrow \infty \). Details are left to the reader. \( \square \)

5 **Appendix A. Scale decompositions**

As a general motivation for this section, consider the Ornstein-Uhlenbeck process (0.3),
\[ \partial_t \phi = \nu \Delta \phi + \eta \] (5.1)
where \( \eta \) is a regularized white noise. Let \( G = (\partial_t - \nu \Delta)^{-1} \) be the Green kernel of the linear heat equation; formally, \( \phi = G \eta \). Thus scale \( j \) fluctuation fields \( \phi^j \) and \( \eta^j \) should be in direct link,
namely, \( \phi^j = G\eta^j \). A natural way to accomplish this is to cut \( G \) itself into scales, \( G = \sum_j G^j \), and set \( \phi^j = G^j \eta, \eta^j = (\partial_t - v\Delta)\phi^j \).

Before discussing various multi-scale decompositions of \( G \), let us first introduce \( \eta \). Let \( \bar{\eta}(t, x) \) be the standard white noise on \( \mathbb{R} \times \mathbb{R}^d \). We recall that \( \bar{\eta} \) is the distribution-valued centered Gaussian process with integrated covariance

\[
\mathbb{E} \left[ \left( \int dt dx f_1(t, x)\bar{\eta}(t, x) \right) \left( \int dt dx f_2(t, x)\bar{\eta}(t, x) \right) \right] = \int dt dx f_1(t, x)f_2(t, x) \tag{5.2}
\]

\((f_1, f_2 \in C^\infty_c(\mathbb{R} \times \mathbb{R}^d))\).

Let also \( \chi = \chi(x) \) be a 'bump' function as in Lemma 5.5.

**Definition 5.1 (regularized white noise)** Let \( \eta(t, x) = \int dy \chi(x - y)\bar{\eta}(t, y) \).

Formally, \( \mathbb{E}[\eta(t, x)\eta(t', x')] = \chi^2(x - x')\delta(t - t') \), where \( \chi^2 := \chi \ast \chi \) is another 'bump' function.

The stationary Ornstein-Uhlenbeck process,

\[
\phi(t, x) = \int dse^{(t-s)\nu\Delta} \eta_s(x), \tag{5.3}
\]

solution of (0.3), has covariance kernel (assuming e.g. \( t \geq t' \))

\[
\mathbb{E}[\phi(t, x)\phi(t', x')] = \int_0^\infty ds \int_0^{t'} ds' \int dy dy' e^{(t-s)\Delta}(x - y)e^{(t' - s')\Delta}(x' - y') \delta(s - s') \chi^2(y - y') = \int_0^\infty du \int dy dy' e^{(t-s+iu)\Delta}(x - y)e^{(t'-s')\Delta}(x' - y') \chi^2(y - y') \tag{5.4}
\]

The regularization has a measurable effect only around the diagonal \( t = t', x = x', u = 0 \). Away from the diagonal the last integral (5.4) behaves like \( \int_{-\infty}^{\infty} du \int dy dy' e^{(t-s+iu)\Delta} = \int_{-\infty}^{\infty} du \int dy \int_{-\infty}^{\infty} dy' \chi^2(x - x') e^{(t-s+iu)\Delta} = \pi \nu \int_{|x - x'|^2} d^2s = C_{|x - x'|^2} \) (the Green kernel of the Laplacian on \( \mathbb{R}^d \)) in the contrary case.

We now want to cut \( \phi \) into scales, i.e. understand how it behaves typically for time separations of order \( M^j (j \geq 0) \), or space separations of order \( M^{1/2} \), where \( M > 1 \) (the scale parameter) is an arbitrary but fixed parameter. The main task is to cut \( G \) into \( M \)-adic scales, \( G = \sum_{j \geq 0} G^j \); then (as discussed above) we define

\[
\phi^j = G^j \eta, \quad \eta^j = (\partial_t - v\Delta)\phi^j. \tag{5.6}
\]

With these definitions, \( \sum_{j \geq 0} \phi^j = G\eta = \phi \) is the Ornstein-Uhlenbeck field, and \( \sum_{j \geq 0} \eta^j = (\partial_t - v\Delta)G\eta = \eta \).

We proceed as follows. Let \( \tilde{\chi} : \mathbb{R}^+ \to \mathbb{R}^+ \) be a smooth 'bump' function of scale 1 supported away from the origin, say, \( \tilde{\chi}_{[M^{1/2}, M^{1/2}]} \equiv 1, \tilde{\chi}_{[M^{-1}, M^{-1}]} \equiv 0 \), chosen in such a way that

\[
\tilde{\chi}^0(\cdot) := \sum_{n \geq 0} \tilde{\chi}(M^n \cdot), \quad \tilde{\chi}^j(\cdot) := \tilde{\chi}(M^{-j}) (j \geq 1) \tag{5.7}
\]

form a partition of unity, i.e. \( \sum_{j \geq 0} \tilde{\chi}^j \equiv 1 \) on \( \mathbb{R}^+ \).
\textbf{Definition 5.2 (cut-off)} Let $G^j$ be the operator
\begin{equation}
(G^j g)(t) := \int \tilde{\chi}^j(s) e^{\nu \Delta} g(t - s) ds, \quad j \geq 0
\end{equation}
and
\begin{equation}
\phi^j = G^j \eta, \quad \eta^j = (\partial_t - \nu \Delta) \phi^j.
\end{equation}
Clearly, $\sum_{j \geq 0} G^j = G$ and $\sum_{j \geq 0} \phi^j = \phi$ is the solution of the Ornstein-Uhlenbeck equation (0.3).

Let $t \geq t'$. Assume $j \geq 1$. The diagonal covariance kernel $C^j_\phi(t, x, t', x') = \mathbb{E}[\phi^j_t(x) \phi^j_{t'}(x')]$ is non-zero only for $t - t' \leq M^j(M - M^{-1})$, in which case (recall $d_\phi := \frac{1}{2}(\frac{d}{\pi} - 1)$)
\begin{align*}
C^j_\phi(t, x, t', x') &= \int_{t - t'}^{t} du \left( e^{\nu(M^{-1} u)} e^{\nu(M^{-1} u - (t - t'))} \right) \chi^2(x - x') \\
&\leq M^j e^{-M^{-j}(t - t')^2/4\nu} M^{j/2} \\
&= (M^{-j})^{2d_\phi^2} e^{-M^{-j}(t - t')^2/4\nu} \leq (M^{-j})^{2d_\phi^2} e^{-cM^{-j/2}|x - x'|}
\end{align*}
for some constant $c > 0$. A similar formula holds for $j = 0$: if $t - t' \leq M$,
\begin{equation}
C^0_\phi(t, x, t', x') = \int_{t - t'}^{t} du \left( e^{\nu u} e^{\nu u - (t - t')} \right) \chi^2(x - x') \leq e^{-M^{-1}(t - t')^2/4\nu} \leq e^{-c|x - x'|}
\end{equation}
since $\chi^2$ is smooth and compactly supported.

Then the off-diagonal covariances
\begin{equation}
C^{j,j}_\phi(t, x, t', x') = \mathbb{E}[\phi^j_t(x) \phi^j_{t'}(x')]
\end{equation}
are similarly shown to satisfy for $j \geq j'$ the estimate
\begin{equation}
|C^{j,j}_\phi(t, x, t', x')| \leq \frac{M^j e^{-M^{-j}(t - t')^2/4\nu}}{M^{j/2}} \leq M^{-j} (M^{-j})^{2d_\phi^2} e^{-cM^{-j/2}|x - x'|}.
\end{equation}
Since $C^{j,j}_\phi(t, \cdot, \cdot) = 0$ for $|t - t'| \gg M^j$, one may clearly also write
\begin{equation}
|C^{j,j}_\phi(t, x, t', x')| \leq \frac{M^{-j} (M^{-j})^{2d_\phi^2} e^{-cM^{-j/2}|t - t'| - cM^{-j/2}|x - x'|}}{M^{j/2}}.
\end{equation}

Finally gradients, resp. time-derivatives applied to the heat kernel produce small factors of order $O(M^{-\max(j,j')/2})$, resp. $O(M^{-\max(j,j')})$. Let us recapitulate.

\textbf{Lemma 5.3 (covariance kernel estimates)} Let
\begin{equation}
C^{j,j}_\phi(t, x, t', x') = \mathbb{E}[\phi^j_t(x) \phi^j_{t'}(x')], C^{j,j}_\eta(t, x, t', x') = \mathbb{E}[\eta^j_t(x) \eta^j_{t'}(t', x')]
\end{equation}
and
\begin{equation}
C^j_\phi := C^{j,j}_\phi, \quad C^j_\eta := C^{j,j}_\eta.
\end{equation}
Then, for $j \geq j'$,
\begin{equation}
\left| \nabla_x^p \nabla_{x'}^q \frac{\partial^r}{\partial t^r} C^{j,j}_\phi(t, x, t', x') \right| \leq M^{-j} M^{4(j + p + q)} e^{-cM^{-j/2}|t - t'| - cM^{-j/2}|x - x'|}
\end{equation}
and
\[\left| \nabla_x \phi_j(t, x, t', x') \right| \leq M^{1-j|\lambda|} \frac{1}{2} (p+q+1) (M^{-j})^2 2d_0 \exp(-cM^{-j}|\lambda| - cM^{-j/2}|x-x'|). \quad (5.18)\]

Furthermore, if \( j \geq 0 \),
\[\mathbb{E}[(\eta_j(t) - \eta_j(y))^2] \leq (M^{-j})^{3+2d_0}|x| \quad (5.19)\]
and
\[\mathbb{E}[(\eta_j(t) - \eta_j(x))^2] \leq (M^{-j})^{4+2d_0}|t-x|^2. \quad (5.20)\]

The last two estimates follow immediately from Taylor’s formula: letting \( v := \frac{y-x}{|y-x|} \),
\[\mathbb{E}[(\eta_j(t) - \eta_j(y))^2] \leq \int_0^{|y-x|} \int_0^{|y-x|} d\zeta \left| \nabla_x \nabla_y \eta_j \right| \quad (5.21)\]
and similarly for \( \mathbb{E}[(\eta_j(t) - \eta_j(x))^2] \).

One has thus obtained a very elaborate version of the scalings, \( \phi_j(t, x) = O(M^{-j\lambda_0}) \), \( \eta_j(t, x) = O((M^{-j})^{1+d_0}) \), together with a first indication of the scale-separation mechanism: the prefactors in powers of \( M^{-j|\lambda|} \) show clearly that fields of widely separated scales are effectively independent.

6 Appendix B. Large deviations estimates for the single-scale noisy equation

6.1 Introduction

We consider here the noisy KPZ equation with scale \( j \) infra-red cut-off,
\[\partial_t \psi = (v\Delta - M^{-j})\psi + \lambda V(\nabla \psi) + \eta^j \quad (6.1)\]
with right-hand side \( \eta^j = G^j\eta \) defined as in Appendix A, subsection 2. Recall the outcome of the computations leading to Corollary 4.6 if \( ||\eta^j||_{W^{1,\infty,1}}([0, t], x) = O(M^{-j\lambda_0}) \), then \( ||\psi(t, x)|| = O((M^{-j})^{1+d_0}) \).

We show in this section that \( ||\eta^j||_{W^{1,\infty,1}}([0, \infty), x) \) is a.s. bounded, and prove large deviation estimates for this quantity when it is much larger than \( O(M^{-j\lambda_0}) \).

The random variables appearing in the definition of the pointwise “quasi-norms” associated with \( W^{1,\infty,1;k} \) are essentially space- and time-averages of a large number of independent log-normal variables, such as \( e^{A(t,x)} \eta_j(t, x) \). Log-normal variables have large tails in \( e^{-a(ln x)^2} \) and thus no exponential moment, hence standard large-deviation theory (notably Cramér’s theorem) does not give any valuable information on the probability that such averages become large. Some authors have been considering this problem, notably Russians, starting from the 60es; one may cite Linnik [45], Nagaev [52, 53], Rozovski [59], see also e.g. Klüppelberg and Mikosch [41] for a renewal of the theory with a view to applications in insurance. The theory is not easily accessible, partly because written originally in Russian journals in the 60es and 70es (in particular in Teoriya Veroyatnostei i ee Primeneniya, later translated to English as Theory of Probability and its Applications), partly for the lack of a theory as general and satisfactory as the standard large-deviation theory.

Let us just point out the difficulties (this very short abstract is taken from an inspiring review in [51]). Choose a random variable \( X \) with finite first and second moments; by translation and rescaling
we may assume that $\mathbb{E}[X] = 0$, $\mathbb{E}[X^2] = 1$. Let $S_n := X_1 + \ldots + X_n$, $M_n := \max(X_1, \ldots, X_n)$, where $X_1, \ldots, X_n$ are independent copies of $X$. Let finally $\bar{F}_n^X(x) := \mathbb{P}(X > x)$, $\bar{F}_n^X(x) := \mathbb{P}(M_n > x)$ and $\text{Errfc}(x) := \int_{x}^{\infty} \frac{e^{-t^2/2}}{\sqrt{2\pi}} \, dt$ be resp. the queues of $X$, of $S_n$ and of a standard Gaussian variable. By the central limit theorem, one expects

$$
\bar{F}_n^X(x) \approx \text{Errfc}(x/\sqrt{n}),
$$

(6.2)

at least if $x \approx \sqrt{n}$. On the other hand, one clearly has if $X \geq 0$

$$
\bar{F}_n^X(x) \geq \mathbb{P}(M_n > x) \sim_{x \to \infty} n\bar{F}_n^X(x).
$$

(6.3)

Subexponential distributions (including log-normal distributions) are precisely defined by the asymptotic relation $\bar{F}_n^X(x) \sim_{x \to \infty} n\bar{F}_n^X(x)$, implying a heavy queue. For distribution with lighter queues (such as e.g. Gaussian distributions), the inequality in (6.3) is very rough, in the sense that typically $n\bar{F}_n^X(x) \ll \bar{F}_n^X(x)$ for every $x \geq x_0$, with $x_0$ independent from $n$.

Thus, one expects, specifically for subexponential distributions, a central limit theorem behaviour as in (6.2) for $x \ll c_n$, with $c_n$ defined by $\text{Errfc}(c_n/\sqrt{n}) \approx \bar{F}_n^X(c_n)$, and an extreme-value regime,

$$
\bar{F}_n^X(x) \sim n\bar{F}_n^X(x), \quad x \gg d_n
$$

(6.4)

with $d_n \geq c_n$, in which $n\bar{F}_n^X(x) \gg \text{Errfc}(x/\sqrt{n})$. Optimal sequences $c_n, d_n$ have been identified for various types of subexponential distributions; for a standard log-normal variable $X = e^Z$, $Z \sim \mathcal{N}(0, 1)$, one finds $c_n, d_n \approx n^2 \ln n$. One major drawback of this picture is that it doesn’t say anything about the behaviour of $\bar{F}_n^X(x)$ in the window $c_n \leq x \leq d_n$ (in our case, for $x \approx c_n$ since $c_n = d_n$), which is expected to be a mixture of (6.2) and (6.3). The complicated asymptotics, valid on the whole real line, proved by Rozovski [59] — a veritable tour de force — give a more complete answer.

This being said, our problem does not fit exactly into this frame, since (1) we are only interested in upper bounds for $\bar{F}_n^X$, moreover in the extreme-value regime, with $x \geq n$; on the other hand (2) the variables $X_1, \ldots, X_i, \ldots, X_j, \ldots, X_n$ (chosen as local space or space-time averages of the noise) are not independent, but have correlations which decrease exponentially with the scaled distance $d^j$ (see below) or equivalently with $|j - i|$; (3) we need scale-dependent estimates for $\bar{F}_n^X$ since $X \approx e^{M^{-j}d^jZ}$, $Z \sim \mathcal{N}(0, 1)$ is strongly $j$-dependent. However all the previous results are strongly dependent on the particular form of the distribution, in particular on the first and second moments, and it is often difficult to retrace the $j$-dependence of the constants in the bounds.

Our main result is the following.

**Theorem 6.1** Let $j \in \mathbb{N}$ and $\lambda > 0$. Then the function $x \mapsto \|\eta^j\|_{W^1_{j,\text{loc},\mathbb{R}_+}}(x)$ is a.s. everywhere defined (i.e. finite). Furthermore, the following large deviation estimates holds for every $x \in \mathbb{R}^d$,

$$
\mathbb{P}\left[ \sup_{B(x,M^j/2)} \|\eta^j\|_{W^1_{j,\text{loc},\mathbb{R}_+}}(x) > AM^{-j}d^j \right] \leq A^{-c\ln(A)}, \quad A \geq 1
$$

(6.5)

where $c > 0$ is some constant.

As follows from Corollary [4.6] this implies (up to the replacement of $\lambda$ by $4\lambda$) that the solution $\psi$ of the full KPZ equation with scale $j$ infra-red cut-off (4.13) is defined a.s. for all positive times.
t ≥ 0 and sits in the space $\mathcal{W}^1_{\text{loc}}(\mathbb{R})$, with $\|\varphi\|_{\mathcal{W}^1_{\text{loc}}(\mathbb{R})} = O(M^{-jd_\varepsilon})$ for every $x \in \mathbb{R}^d$, with a random multiplicative prefactor $A(x)$ whose queue is bounded \textit{locally in} $x$ by that of a log-normal distribution. (Note that the prefactor $A(x)$ is \textit{not globally} bounded!)

The proof includes both Gaussian inequalities taken from the monograph [2], and an adaptation to weakly correlated variables of a result about large deviations for subexponential distributions [53]. We shall need quite a few preliminary results before the proof, given at the very end of the present section.

We finish this introductory paragraph with the tiny bit of stochastic domination and Gaussian inequalities used in the sequel, and a little bit of geometry.

**Definition 6.1** Let $X : \Omega \to \mathbb{R}, Y : \Omega' \to \mathbb{R}$ be two real-valued random variables, defined a priori on two different probability spaces. Then $X$ is stochastically dominated by $Y$ if

$$\forall x \in \mathbb{R}, \mathbb{P}[X > x] \leq \mathbb{P}[Y > x]. \quad (6.6)$$

We then write $X \preceq Y$.

By Strassen’s theorem [45], if $X \preceq Y$, there exists a coupling between $X$ and $Y$, i.e. random variables $X', Y' : \Omega' \to \mathbb{R}$ defined on the same probability space, with $X' \overset{(d)}{=} X$, $Y' \overset{(d)}{=} Y$, and such that $X' \preceq Y'$.

**Proposition 6.2** (see [2]) Let $(Z_1, \ldots, Z_n)$ be a centered Gaussian vector, and $\phi : \mathbb{R}^n \to \mathbb{R}$ be a convex function with polynomial growth at infinity. Then $\mathbb{E}[\phi(Z_1, \ldots, Z_n)]$ is an increasing function of the coefficients $c_{ij} = \mathbb{E}[Z_i Z_j], i, j = 1, \ldots, n$.

This technical lemma, due to Slepian (whose short proof relies on a Gaussian integration by parts) is one of the main tools for Gaussian inequalities. It extends to Gaussian fields $(Z_x)_{x \in \mathbb{R}^d}$ and convex functionals $\phi$ under adequate regularity assumptions.

**Proposition 6.3** (Borell-Tsirelson-Ibragimov-Sudakov or BTIS inequality) (see [2]) Let $(Y_x)_{x \in D}$ be a centered Gaussian process, such that $\sigma^2_D = \sup_{x \in D} \mathbb{E}[Y_x^2] < \infty$, and $\delta(x, y) := \sqrt{\mathbb{E}[(Y_x - Y_y)^2]} \leq |x - y|$. Let $\|Y\|_\infty := \max_{x \in D} |Y_x|$. Then a.s. $\|Y\|_\infty < \infty$, $\mathbb{E}[\|Y\|_\infty] \leq 1$ and

$$\mathbb{P}[\|Y\|_\infty - \mathbb{E}[\|Y\|_\infty] > u] \leq e^{-u^2/2\sigma^2_D}. \quad (6.7)$$

This is actually a particular case of the BTIS inequality. For a Gaussian process $Y$ indexed by an abstract set $\mathcal{D}$, $\mathbb{E}[\|Y\|_\infty]$ is bounded by the integral of the square-root of the entropy log $N(\varepsilon)$, $\mathbb{E}[\|Y\|_\infty] \leq \int_0^{\infty} \sqrt{\log N(\varepsilon)} d\varepsilon$, where $N(\varepsilon)$ is the minimum number of balls of diameter $\varepsilon$ (with respect to the metric induced by $\delta(\cdot, \cdot)$) covering $\mathcal{D}$. In our proposition, $\log N(\varepsilon) = 0$ for $\varepsilon \gg 1$ since $\sup_{x, y \in \mathcal{D}} \delta(x, y) \leq 1$, and $N(\varepsilon) = O(\varepsilon^{d_\varepsilon})$ otherwise by hypothesis, hence the result.

The above proposition applies for fixed $t_0, x_0$ to $Y_x := M^{1+d_\varepsilon} \eta^i(t_0, x_0 + M^{1/2} x)$, with $\mathcal{D} = B(0, 1)$. It follows from Lemma 5.3 in Appendix A that $\sigma^2_D \approx 1$ and $d(x, y) \leq |x - y|$. Thus

$$\mathbb{E} \left[ \sup_{B(t_0, M^{1/2})} |\eta^i| \right] \leq M^{-j(1+d_\varepsilon)} \quad (6.8)$$
and there exists a constant $C \leq 1$ such that
\[
\mathbb{P}[M^{(1+d_0)} B(x_0, M^{(1/2)}) \sup |\eta^j_{t_0}| > u + C] \leq e^{-u^2/2C}.
\] (6.9)

One easily deduces that
\[
M^{(1+d_0)} B(x_0, M^{(1/2)}) |\eta^j_{t_0}| \leq C(|Z| + 1)
\] (6.10)
if $Z \sim \mathcal{N}(0, 1)$.

Recall from section 3.1 that $f^* \leq f^2$ ($f \in C(\mathbb{R}^d, \mathbb{R})$) – note, and this is very important, that the inequality is exact, with a coefficient one $\sim$, where $f^2(x) = \sup_{\rho > 0} \int_{B(x, \rho)} |f|$. We cannot bound directly a supremum over a continuous parameter (here $\rho$), so it is natural to start by rewriting $(\eta^j_{t_0})^*$ in terms of its local averages or suprema on balls of radius $M^{(1/2)}$, over which we have a good control. However, we cannot obviously cover $\mathbb{R}^d$ (nor $B(x, \rho)$) by disjoint balls of fixed radius, and taking into account error terms due to overlaps or boundary effects would cost a multiplicative coefficient, which we cannot afford to do. Hence we first transform balls centered at $x$ into cubes by letting
\[
\Phi : \mathbb{R}^d \to \mathbb{R}^d, \quad y \mapsto \Phi(y) = x + \frac{|y - x|}{\|y - x\|_\infty} (y - x)
\] (6.11)
where $\|y - x\|_\infty := \sup(|y_1 - x_1|, \ldots, |y_d - x_d|)$ is the supremum norm. The Euclidean norm $| \cdot |$ and the supremum norm $\| \cdot \|_\infty$ are equivalent, hence (the easy proof is left to the reader) $\Phi$ and $\Phi^{-1}$ are uniformly Lipschitz. Thus $\sup_{B(x, \rho)} |\eta^j_{t_0}| = \sup_{B(x, \rho)} |\eta^j_{t_0} \circ \Phi^{-1}|$, where $B(x, \rho) = \{y \in \mathbb{R}^d | \|y - x\|_\infty \leq \rho\}$ is a cube. The field $\eta^j_{t_0} \circ \Phi^{-1}$ has the same general properties as $\eta^j_{t_0}$ (scaling, exponentially decreasing covariance) as stated in Lemma 5.3, so (by abuse of notation) we simply denote $\eta^j_{t_0} \circ \Phi^{-1}$ by $\eta^j_{t_0}$ in the sequel.

**Definition 6.4 (scale $j$ cubes)** Let $\mathbb{D}^j$ be the set of all scale $j$ cubes, i.e. of all primitive cells $[k_1 M^{(1/2)}, (k_1 + 1) M^{(1/2)}] \times \cdots \times [k_d M^{(1/2)}, (k_d + 1) M^{(1/2)}]$, $k_1, \ldots, k_d \in \mathbb{Z}$ of the square lattice $M^{(1/2)} \mathbb{Z}^d$.

We denote by $x_\Delta = (x_{\Delta,1}, \ldots, x_{\Delta,d})$ the center of a cube $\Delta \in \mathbb{D}^j$.

We now show how to bound an average $\int_{B(x, \rho)} |f|$, $f \in C(\mathbb{R}^d, \mathbb{R})$ over a cube of arbitrary radius in terms of the local suprema $f_\Delta := \sup_{\Delta} |f|$, $\Delta \in \mathbb{D}^j$. We give the proof in dimension 2 to simplify notations (in general, we would need the whole cellular decomposition of a cube). Let, for $\rho > 0$,
\[
\tilde{B}^j(x, \rho) := \{\Delta \in \mathbb{D}^j | \Delta \subset \tilde{B}(x, M^{(1/2)} \rho)\}, \quad \partial \tilde{B}^j(x, \rho) := \{\Delta \in \mathbb{D}^j | \Delta \cap \tilde{B}(x, M^{(1/2)} \rho) \neq \emptyset\} \setminus \tilde{B}^j(x, \rho),
\] (6.12)
and $n := \partial \tilde{B}^j(x, \rho)$. The boundary $\partial \tilde{B}^j(x, \rho)$ decomposes into 8 disjoint subsets,
\[
\partial \tilde{B}^j_{\text{right}}(x, \rho) := \{\Delta = [x_{\Delta, \min}, x_{\Delta, \max}] \times [y_{\Delta, \min}, y_{\Delta, \max}] | a < x_{\Delta, \min} < b < x_{\Delta, \max}, c \leq y_{\Delta, \min} < y_{\Delta, \max} \leq d\}
\] (6.13)
and similarly $\partial \tilde{B}^j_{\text{left}}(x, \rho), \partial \tilde{B}^j_{\text{up}}(x, \rho), \partial \tilde{B}^j_{\text{down}}(x, \rho)$ for the sides of the square;
\[
\partial \tilde{B}^j_{\text{up,right}}(x, \rho) = \{\Delta = [x_{\Delta, \min}, x_{\Delta, \max}] \times [y_{\Delta, \min}, y_{\Delta, \max}] | a < x_{\Delta, \min} < b < x_{\Delta, \max}, c < y_{\Delta, \min} < d < y_{\Delta, \max}\}
\] (6.14)
and similarly for the three other corners. We let $c_{\text{right}} := \frac{\text{Vol}(\partial B^j_{\text{right}}(x, \rho)) \cap \tilde{B}^j(x, \rho)}{\text{Vol}(\partial B^j_{\text{right}}(x, \rho))}$, and similarly $c_{\text{left}}, \ldots$ be the corresponding volume ratios. Let

$$F(c_{\text{right}}, c_{\text{left}}, \ldots) := \sum_{\Delta \in \partial B^j(x, \rho)} f_\Delta + c_{\text{right}} \sum_{\Delta' \in \partial B^j_{\text{right}}(x, \rho)} f_{\Delta'} + \ldots; \quad (6.15)$$

note that $\int_{B^j(x, \rho)} f \leq F(c_{\text{right}}, \ldots)$ since $c_{\text{right}}$ is the uniform volume ratio $\frac{\text{Vol}(\Delta) \cap \tilde{B}^j(x, \rho)}{\text{Vol}(\Delta)}$ of all scale $j$ cubes at the right border, as follows from the fact that the border is straight. Then trivially $F(c_{\text{right}}, c_{\text{left}}, \ldots) \leq \max \{ F(0, c_{\text{left}}, \ldots), F(0, c_{\text{right}}, \ldots) \}$; this same elementary remark may be repeated for the eight $c$ coefficients. Thus we have proved that

$$\int_{B^j(x, \rho)} f \leq \max_{B^j} \frac{\sum_{B^j \cap \Delta} f_\Delta}{\# B^j}, \quad (6.16)$$

where the $B^j$ range among $2^8$ subsets of squares, and by definition $\tilde{B}^j(x, \rho) \subset B^j \subset \tilde{B}^j(x, \rho) \cup \partial B^j(x, \rho)$.

6.2 A first preliminary result: large deviations for the noise

We prove in this paragraph the following result.

Lemma 6.5 Let $j \in \mathbb{N}$ and $t_0 \in \mathbb{R}_+$. Then the function $x \mapsto (\eta^j)^*(t_0, x)$ is a.s. everywhere defined (i.e. finite). Furthermore, the following large deviation estimates holds,

$$\mathbb{P}[ \sup_{B(x, M^j/2)} (\eta^j)^*(t_0) > AM^{-j(1+d_0)}] \leq e^{-c(A-C)^2} \quad (6.17)$$

for some constants $c, C > 0$, where $(A-C)^2 = (A-C)^2 1_{A>C}$.

It is actually reasonable to expect, on account of the central limit theorem, that $|\eta^j(t_0, x)| - \mathbb{E}[|\eta^j(t_0, x)|] \in \mathcal{H}_0^\alpha$ for every $\alpha < d/4$, and that the norm in $\mathcal{H}_0^\alpha$ satisfies large deviation estimates as in (6.17), but we do not prove this. The above result, however natural it may be, is not really needed anywhere in the article, but the proof of Theorem 6.2 is based on the arguments developed for the proof of the lemma.

Proof. In the sequel, $c, c', C > 0$ are constants possibly varying from line to line (contrary to $c_0, m_0$, see below, which are fixed once and for all). As already recalled, $(\eta^j)^*(t_0, x) \leq (\eta^j)^j(t_0, x) = \sup_{y > 0} \int_{B(x, M^j/2 \rho)} dy |\eta^j(t_0, y)|$. Also, from the results of Appendix A, the correlations of the field $(\eta^j_0(x))_{x \in \mathbb{R}^d}$ decay exponentially with the scaled distance $d^j(x, x') := M^j|\eta^j_0(x')|$, in the sense that, for a certain constant $c_0$,

$$\mathbb{E}[\eta^j_0(x)\eta^j_0(x')] \leq M^{-2j(1+d_0)} e^{-c_0 d^j(x,x')} \quad (6.18)$$

We split the proof into several points.
(i) In order to use the exponential decay, we first choose \( m_0 \geq 2 \) large enough (depending on further considerations), and partition \( \mathbb{D} \) into \( m_0 \) disjoint susets \( \mathbb{D}_\mu \), \( \mu \in \{1, \ldots, m_0\} \), with \( \Delta = [k_1 M^{1/2}, (k_1 + 1) M^{1/2}] \times \cdots \times [k_d M^{1/2}, (k_d + 1) M^{1/2}] \in \mathbb{D}_\mu \Leftrightarrow k_i \equiv \mu_i \mod m_0 \). Two points \( x, x' \) located in disjoint cubes \( \Delta \neq \Delta' \) in the same sublattice \( \mathbb{D}_\mu \) are thus at distance \( d^1(x, x') \gtrsim m_0 \), which amounts (up to rescaling) to replacing \( c_0 \) by \( m_0 c_0 \) in \((6.18)\); in the sequel, we may thus assume that \( c_0 \) is large enough. By abuse of notation, we also denote by \( \mathbb{D}_\mu \) the subset \( \cup \{\Delta; \Delta \in \mathbb{D}_\mu\} \subset \mathbb{R}^d \). Clearly,

\[
\int_{B(x, M^{1/2} \rho)} dy \eta^j(t_0, y) \leq \sup_{\mu} \int_{B(x, M^{1/2} \rho) \cap \mathbb{D}_\mu} dy \eta^j(t_0, y). \tag{6.19}
\]

If \( \psi : \mathbb{R} \to \mathbb{R} \) is increasing, then

\[
\mathbb{E} \left[ \psi \left( \int_{B(x, M^{1/2} \rho)} dy \eta^j(t_0, y) \right) \right] \leq \sum_{\mu} \mathbb{E} \left[ \psi \left( \int_{B(x, M^{1/2} \rho) \cap \mathbb{D}_\mu} dy \eta^j(t_0, y) \right) \right]. \tag{6.20}
\]

(ii) Next, we want to bound the average \( \int_{B(x, M^{1/2} \rho) \cap \mathbb{D}_\mu} dy \eta^j(t_0, y) \) over some fixed sublattice by the average of a finite number of variables representing the supremum of \( |\eta^j| \) on each cube. For that (note that the following construction is \( \mu \)-dependent, which we do not always specify) we introduce i.i.d. copies \( \eta^j_A \) of the field \( \eta^j \) restricted to some reference cube, and define a new random field \( \tilde{\eta}^j \) on \( \mathbb{D}_\mu \),

\[
\tilde{\eta}^j(x) := \sum_{\Delta' \in \mathcal{B}(x, \rho)} e^{-c d^1(\Delta, \Delta')} \eta^{j}_{\Delta}(x-x_D), \quad x \in \Delta \tag{6.21}
\]

separately on each cube \( \Delta \in \mathbb{D}_\mu \), where

\[
d^1(\Delta, \Delta') := M^{-1/2} \sup_{x \in \Delta} \inf_{y \in \Delta'} |x-y| \tag{6.22}
\]

is the set distance measured in scaled units, and

\[
\mathcal{B}(x, \rho) := \{\Delta \in \mathbb{D}_\mu \mid \Delta \subset \mathcal{B}(x, M^{1/2} \rho)\} \tag{6.23}
\]

(compare with the previous definition, \((6.12)\)). By a simple computation, one finds

\[
\mathbb{E}[\tilde{\eta}^j(x) \tilde{\eta}^j(x')] \approx (1+2d(\Delta, \Delta'))^d e^{-c d^1(\Delta, \Delta')} \mathbb{E}[(\eta^j_0(x-x_D)\eta^j_0(x'-x_D)) \gtrsim \mathbb{E}[\eta^j_0(x)\eta^j_0(x')] \tag{6.24}
\]

if \( x \in \Delta, x' \in \Delta' \) and

\[
\Delta, \Delta' \in \mathcal{B}(x, \rho) \cup \partial \mathcal{B}(x, \rho) := \{\Delta \in \mathbb{D}_\mu \mid \Delta \subset \mathcal{B}(x, M^{1/2} \rho) \neq \emptyset\}. \tag{6.25}
\]

Applying Proposition \ref{prop:4.2} with \( \phi(\eta^j_0) = \psi \left( M^{j(1+\rho)} \int_{B(x, M^{1/2} \rho) \cap \mathbb{D}_\mu} dy \eta^j_0(y) \right) \) where \( \psi \) is any convex, increasing function \( \mathbb{R} \to \mathbb{R} \),

\[
\mathbb{E} \left[ \psi \left( M^{j(1+\rho)} \int_{B(x, M^{1/2} \rho) \cap \mathbb{D}_\mu} dy \eta^j_0(y) \right) \right] \leq \mathbb{E} \left[ \psi \left( M^{j(1+\rho)} \int_{B(x, M^{1/2} \rho) \cap \mathbb{D}_\mu} dy \tilde{\eta}^j(y) \right) \right]. \tag{6.26}
\]

\(^{3}\text{Observe that } \psi_1 \circ \psi_2 \text{ is convex if } \psi_1 : \mathbb{R} \to \mathbb{R} \text{ is convex and increasing and } \psi_2 : \mathbb{R}^n \to \mathbb{R} \text{ is convex, since } \nabla^2(\psi_1 \circ \psi_2) = \psi_1' \circ \psi_2 \cdot \nabla \psi_2 \otimes \nabla \psi_2 + \psi_1' \circ \psi_2' \cdot \nabla^2 \psi_2.\)
As follows from the discussion in the previous paragraph,

\[
\int_{B(x,M^{1/2}p) \cap \mathbb{D}_\mu} dy |\tilde{\eta}^j(y)| \leq \max_{B^j} \frac{\sum_{\Delta \in B^j} \dot{Y}_\Delta}{\#B^j},
\]

(6.27)

where

\[
\dot{Y}_\Delta := \sup_\Lambda |\tilde{\eta}^j|
\]

(6.28)

and the \(B^j\) are a finite number (depending only on \(d\)) of subsets of cubes such that \(\tilde{B}^j(x,\rho) \subset B^j \subset \tilde{B}^j(x,\rho) \cup \partial \tilde{B}^j(x,\rho)\).

(iii) By construction, see (6.21),

\[
\dot{Y}_\Delta \leq \sum_{\Delta' \in B^j(x,\rho)} e^{-\epsilon_0 d^j(\Delta,\Delta')} \sup_\Lambda |\eta^j_{\Delta'}|.
\]

(6.29)

We have seen in (6.10) that \(M^{j+1+d_0} \sup_\Delta |\eta_j^\Delta| \leq C(|Z_\Delta| + 1)\) if \(Z_\Delta \sim \mathcal{N}(0,1)\). Since the fields \((\eta^j_\Delta)_\Delta\) are independent, we may by the above cited Strassen theorem define a coupling of the field \(\eta^j_{\Delta'}\) with i.i.d. standard Gaussian variables \((Z_\Delta)_{\Delta \in \mathcal{D}_\mu}\) in such a way that

\[
M^{j+1+d_0} \sup_\Delta |\dot{\eta}^j_\Delta| \leq C(|Z_\Delta| + 1).
\]

(6.30)

Hence

\[
M^{j+1+d_0} \sum_{\Delta \in B^j} \dot{Y}_\Delta \leq C \sum_{\Delta \in B^j(x,\rho)} (|Z_\Delta| + 1)
\]

(6.31)

– note that the bound in the right hand side does not depend on the choice of \(B^j\) – and

\[
\mathbb{E} \left[ \psi \left( M^{j+1+d_0} \int_{B(x,M^{1/2}p) \cap \mathbb{D}_\mu} d\tilde{\eta}^j(y) \right) \right] \leq \mathbb{E} \left[ \psi \left( \frac{C}{n} \sum_{\Delta \in B^j(x,\rho)} (|Z_\Delta| + 1) \right) \right].
\]

(6.32)

We rewrite the expectation as an integral by integration by parts,

\[
\mathbb{E} \left[ \psi \left( \frac{C}{n} \sum_{\Delta \in B^j} (|Z_\Delta| + 1) \right) \right] = \int_0^{+\infty} dA \psi'(A) \mathbb{P} \left[ \frac{C}{n} \sum_{\Delta \in B^j} (|Z_\Delta| + 1) > A \right] + \psi(0).
\]

(6.33)

Finally, \(\sum_{\Delta \in B^j} |Z_\Delta|\) is a sum of \(n\) independent copies of \(|Z|\), where \(Z \sim \mathcal{N}(0,1)\), to which we may apply standard large deviation arguments in a trivial setting,

\[
\mathbb{P} \left[ \sum_{\Delta \in B^j} |Z_\Delta| > nA \right] \leq \min \left( 1, \min_{t \geq 0} e^{-nA} \mathbb{E} \left[ e^{t \sum_{\Delta \in B^j} |Z_\Delta|} \right] \right) \leq \min \left( 1, 2^n \min_{t \geq 0} e^{-nA+nt^2/2} \right)
\]

\[
= \min \left( 1, 2^n e^{-nA^2/2} \right) \leq Ce^{-n(A-C)^2/2}.
\]

(6.34)

Thus we may choose \(\psi(A) = e^{-cn(A-C)^2} 1_{A>C} + 1_{A \leq C}\) so that \(\mathbb{E} \left[ \psi \left( \frac{C}{n} \sum_{\Delta \in B^j} (|Z_\Delta| + 1) \right) \right] \leq 1\). Collecting (6.20), (6.26) and (6.32), one obtains by Markov’s inequality

\[
\mathbb{P} \left[ M^{j+1+d_0} \int_{B(x,M^{1/2}p)} d\tilde{\eta}^j_0(y) > A \right] \leq \frac{1}{\psi(A)} \leq e^{-cn(A-C)^2}, \quad A \geq C.
\]

(6.35)
For each fixed $n \geq 1$, the set $\{ B^i(x, \rho), \rho \geq 0 \mid \# B^i(x, \rho) = n \} \cup \{ B^i(x, \rho) \cup \partial B^i(x, \rho), \rho \geq 0 \mid \# B^i(x, \rho) \cup \partial B^i(x, \rho) = n \}$ consists of 0, 1 or 2 elements. Thus, using (6.20),

$$\mathbb{P}[(\eta^j)^*(t_0, x) > AM^{-\varepsilon d_0}] \leq \min \left\{ 1, \sum_{n \geq 1} e^{-c n (A-C)^2} \right\} \leq e^{-c(A-C)^2}.$$  \hspace{1cm} (6.36)

Finally, we use a scaled version of (3.18),

$$\sup_{B(x,M^{1/2})} |\eta^j_{t_0}| \leq \sup_{B(x,2M^{1/2})} |\eta^j_{t_0}| + (\eta^j_{t_0})^d(x),$$  \hspace{1cm} (6.37)

from which we conclude that

$$\mathbb{P}[\sup_{B(x,M^{1/2})} (\eta^j_{t_0})^* > AM^{-\varepsilon d_0}] \leq e^{-c(A-C)^2}.$$  \hspace{1cm} (6.38)

In particular,

$$\mathbb{P}[\exists x \in \mathbb{R}^d \mid (\eta^j_{t_0})^*(x) = +\infty] \leq \sum_{\Delta \in \mathcal{D}} \mathbb{P}[\sup_{\Delta} (\eta^j_{t_0})^* = +\infty] = 0.$$  \hspace{1cm} (6.39)

\[\square\]

### 6.3 Large deviations for the exponential of the noise

We now turn to large deviation estimates for $(e^{AM|\eta^j_{t_0}|})^*(x)$ and prove the following result.

**Theorem 6.2** Let $j \in \mathbb{N}$, $\lambda > 0$ and $t_0 \in \mathbb{R}_+$. Then the function $x \mapsto (e^{AM|\eta^j_{t_0}|})^*(x)$ is a.s. everywhere defined (i.e. finite). Furthermore, the following large deviation estimates holds for every $x \in \mathbb{R}^d$,

$$\mathbb{P}[\sup_{B(x,M^{1/2})} \ln(e^{AM|\eta^j_{t_0}|})^*(x) > \varepsilon A] \leq A^{-\varepsilon \ln(A)}, \quad A \geq 1$$  \hspace{1cm} (6.40)

where $\varepsilon = AM^{-\varepsilon d_0}$ and $c > 0$ is some constant.

The proof is essentially similar to that of Lemma 6.5 except that it is based on large deviation estimates for log-normal variables. We cite a result by Nagaev, show how to apply it in our context, and prove a few technical lemmas before turning to the proof of Theorem 6.2.

#### 6.3.1 Log-normal large deviations

**Proposition 6.6** (see [53], Corollary 1.8) Let $X$ be a real-valued random variable such that $\mathbb{E}[X] = 0$ and $\mathbb{E}[|X|^t] < \infty$ for some $t \geq 2$, and $X_1, \ldots, X_n$ i.i.d. copies of $X$, $S_n := X_1 + \ldots + X_n$. Then

$$\mathbb{P}[S_n > A] \leq n \mathbb{E}[X^t 1_{X_0 > 0}] A^{-t} + e^{-2(t/2) - e^{-t/2}/2^n \mathbb{E}[X^2]}.$$  \hspace{1cm} (6.41)

Note that this general bound mixes the two regimes (6.2) and (6.3).

45
Corollary 6.7 Let \((Z_t)_{t=1,\ldots,n}, n \geq 1\) be i.i.d. standard Gaussian variables, and let

\[
S_n := \sum_{i=1}^{n} \left( e^{iZ_t} - \mathbb{E}[e^{iZ_t}] \right), \quad \bar{S}_n := \sum_{i=1}^{n} \left( e^{iZ_t} - 1 \right)
\]

where \(0 < \varepsilon \ll 1\). Let finally \(A \gg n\varepsilon\) and \(B \gg \ln(n)\). Then there exists a constant \(c > 0\) such that

(i) \[
\mathbb{P}[S_n > A] \leq (A/\varepsilon)^{-c\ln(A/\varepsilon)}
\]

or equivalently

(ii) \[
\mathbb{P}[\ln \left( \frac{S_n}{\varepsilon} \right) > B] \leq e^{-cB^2};
\]

(iii) \[
\mathbb{P}[\ln S_n > A] \leq (A/\varepsilon)^{-c\ln(A/\varepsilon)}.
\]

Furthermore, the same estimates (6.43), (6.44), (6.45) still hold if one replaces \(S_n\) by \(\bar{S}_n\).

Proof.

Note that (ii) is equivalent to (i), and (iii) follows directly from (ii) since \(e^A \gg n\varepsilon\) and \(\frac{e^A}{\varepsilon} \gg \frac{A}{\varepsilon}\) if \(A \gg n\varepsilon\). Also, since \(\mathbb{E}[e^{iZ_t}] = 1 + O(\varepsilon)\), \(S_n - S_n = O(n\varepsilon) \ll A\), so the same estimates hold indifferently for \(S_n\) or \(\bar{S}_n\) (up to the choice of \(c\)).

Therefore we need only prove (i) for \(S_n\). We apply the above Proposition with \(X = e^{iZ_t} - \mathbb{E}[e^{iZ_t}]\), where \(Z_t\) is any of the variables \(Z_t\). One finds \(\mathbb{E}X^2 = \mathbb{E}\{e^{2iZ_t} - (\mathbb{E}[e^{iZ_t}])^2\} \approx \varepsilon^2\) and \(\mathbb{E}[X'1_{X>0}] < \mathbb{E}[e^{iZ_t}] \leq 2e^{\frac{\varepsilon}{2}^2/2}\). The bound (6.41) is close to optimal if one chooses \(t = \frac{1}{2}\ln(A/\varepsilon) \gg 1\); we then find (using \(\ln^2(A/\varepsilon) \ll (A/\varepsilon)^k\) for all \(k > 0\))

\[
\mathbb{P}[S_n > A] \leq n\varepsilon^{1/2} \ln^2(A/\varepsilon) A^{-\frac{1}{2}\ln(A/\varepsilon)} + e^{-\frac{c}{2}(\ln(A/\varepsilon))^2}. \tag{6.46}
\]

(i) Assume first that \(A \geq e^{1/\varepsilon} \gg 1\). The second term in the right-hand side of (6.46) is then the smaller one since (for \(\kappa < \frac{1}{2}\))

\[
e^{-\frac{c}{2}(\ln(A/\varepsilon))} \leq e^{-\frac{c}{2}(\ln(A/\varepsilon))} \leq e^{-\frac{1}{2}\ln(A/\varepsilon)^2} = (A/\varepsilon)^{-\frac{1}{2}\ln(A/\varepsilon)} \leq A^{-\frac{1}{2}\ln(A/\varepsilon)}. \tag{6.47}
\]

As for the first term, it is bounded by \(A^{-c\ln(A/\varepsilon)}\) since (using \(A \gg 1\))

\[
e^{\frac{\varepsilon}{2} \ln^2(A/\varepsilon)} \leq e^{\frac{\varepsilon}{2} \ln(A/\varepsilon)} \leq e^{\frac{\varepsilon}{2} (\varepsilon \ln(A) + 1) \ln(A/\varepsilon)}
\]

\[
\leq e^{\frac{1}{2} \ln(A) \ln(A/\varepsilon)} = A^{\frac{1}{2} \ln(A/\varepsilon)}. \tag{6.48}
\]

and

\[
n \ll A/\varepsilon = \varepsilon^{1/2} \ln(A/\varepsilon) \leq A^{\frac{1}{2} \ln(A/\varepsilon)}. \tag{6.49}
\]

All together one has obtained

\[
\mathbb{P}[S_n > A] \leq A^{-c\ln(A/\varepsilon)} \ll (A/\varepsilon)^{-c\ln(A/\varepsilon)}, \quad A \geq e^{1/\varepsilon}. \tag{6.50}
\]
(ii) We now assume that \( A \leq e^{1/\varepsilon} \), implying that \( t = \frac{1}{\varepsilon} \ln(A/\varepsilon) \leq \frac{1}{\varepsilon} \). Then \( A' \) is not necessarily small, so we must first improve our bound on \( \mathbb{E}[X'_{1_{X>0}}] \):

\[
\mathbb{E}[X'_{1_{X>0}}] \leq 2\mathbb{E}[(e^{\varepsilon Z} - 1)'_{1_{Z>0}}] \\
\leq \int_0^{1/\varepsilon} dz (e^{\varepsilon z} - 1)' e^{-z^2/2} + \int_{1/\varepsilon}^{+\infty} dz (e^{\varepsilon z} - 1)' e^{-z^2/2} \\
\leq \int_0^{+\infty} dz (e^{\varepsilon z}) e^{-z^2/2} + \mathbb{E}[e^{\varepsilon Z}] e^{-1/2} \\
= 2^{-1/2} e^{\varepsilon \sqrt{2}/2}(t+1) + e^{\varepsilon (e^{\varepsilon^2 t} - 1/2)}. \quad (6.51)
\]

Since \( t \leq \frac{1}{\varepsilon} \), we find \( e^{\varepsilon (e^{\varepsilon^2 t} - 1/2)} \leq e^{-c/\varepsilon^2} \ll e^{-\frac{1}{2} \ln |\varepsilon|} \ll e^{-c \ln (A/\varepsilon)} = e'. \) Hence

\[
n\mathbb{E}[X'_{1_{X>0}}] A^{-j} \leq n t' (A/\varepsilon)^{-c \ln (A/\varepsilon)} \leq (A/\varepsilon)^{-c \ln (A/\varepsilon)}. \quad (6.52)
\]

As for the second term, clearly \( e^{-\frac{1}{2} \ln (A/\varepsilon)} \ll (A/\varepsilon)^{-c \ln (A/\varepsilon)} \) (see (6.47)). All together,

\[
\mathbb{P}[S_n > A] \leq (A/\varepsilon)^{-c \ln (A/\varepsilon)}, \quad A \leq e^{1/\varepsilon}. \quad (6.53)
\]

\( \square \)

**Remark.** The above results are actually valid as soon as \( A \gg n' \varepsilon \) with \( \kappa > \frac{1}{\varepsilon} \), as the reader may easily check (choose \( t = c \ln (A/\varepsilon) \) with \( c \) small enough and see how (6.46) and (6.47) are modified). The condition \( A \gg n' \varepsilon \) may certainly be further improved with some extra effort.

Corollary 6.7 has the following generalization.

**Corollary 6.8 (block large deviation estimates)** Let \( Z := \sum_{i=1}^{n'} |Z_i| \), where \( (Z_i)_{i=1,\ldots,n'} \) are i.i.d. standard Gaussian variables; \( n \in \mathbb{N}^+ \) a multiple of \( n' \), \( Z_i, i = 1,\ldots,n \) i.i.d. copies of \( Z \);

\[
S_n := \sum_{i=1}^{n/n'} (e^{Z_i} - \mathbb{E}[e^{Z_i}]), \quad \tilde{S}_n := \sum_{i=1}^{n/n'} (e^{Z_i} - 1) \quad (6.54)
\]

where \( 0 < \varepsilon \ll 1 \) and \( \varepsilon n' \ll 1 \). Let finally \( A \gg n \varepsilon \) and \( B \gg \ln(n) \). Then there exists a constant \( \kappa > 0 \) such that

(i) \( \mathbb{P}[S_n > A] \leq (A/n' \varepsilon)^{-c \ln (A/n' \varepsilon)} \) \quad (6.55)

or equivalently

(ii) \( \mathbb{P}[\ln \frac{S_n}{n'/\varepsilon} > B] \leq e^{-c B^2}; \) \quad (6.56)

(iii) \( \mathbb{P}[^{\ln S_n > A}] \leq (A/n' \varepsilon)^{-c \ln (A/n' \varepsilon)}. \) \quad (6.57)

Furthermore, the same estimates \( (6.55), (6.56), (6.57) \) still hold if one replaces \( S_n \) by \( \tilde{S}_n \).
The result is exactly the one stated in Corollary 6.7 if \( n' = 1 \). We want to prove the same kind of result for blocks of size \( n' \). Standard large deviation arguments apply to \( Z \), yielding (see (6.34)) \( P[Z > A] \leq ce^{-\frac{A^2}{2\varepsilon^2}(A-cn')^2} \), hence (letting as before \( X := e^Z - \mathbb{E}[e^Z] \)), \( \mathbb{E}[e^Z] = 1 + O(n'e) \), \( \mathbb{E}[X^2] \leq e^{2\mathbb{V}(Z)} = O(n'e^2) \), and

\[
\mathbb{E}[X'1_{X>0}] \leq \mathbb{E}[e^{tZ}] \leq t e \int_0^{+\infty} e^{tZ} e^{-\frac{t^2}{2\varepsilon^2}(Z-cn')^2} dz \\
\leq \int_0^{2cn'} e^{t(Z-cn')^2} + \int_{-\infty}^{+\infty} e^{t(Z-cn')^2} dz \\
\leq e^{2Cn'te} + e^{Cn'(t\varepsilon^2)/2}.
\]

We set \( t := \frac{1}{2} \ln(A/n' \sqrt{\varepsilon}) \) and distinguish two regimes according to whether \( A \geq e^{1/(n' \sqrt{\varepsilon})} \), corresponding to \( t \geq \frac{1}{\varepsilon \sqrt{n'}} \). Thus

\[
e^{2Cn'te} = e^{Cn't\ln(A/e \sqrt{n'})} \leq e^{\ln(A/e \sqrt{n'})} \ll A^{\frac{1}{2} \ln(A/e \sqrt{n'})}
\]

instead of (6.48), and

\[
\mathbb{E}[X'1_{X>0}] \leq \mathbb{E}[(e^Z - 1)^+] \\
\leq \int_0^{1/e} dz(e^Z - 1)^+ e^{-\frac{t^2}{2\varepsilon^2}(Z-cn')^2} + \int_{1/e}^{+\infty} dz(e^Z - 1)^+ e^{-\frac{t^2}{2\varepsilon^2}(Z-cn')^2} \\
\leq \int_0^{2cn'} dz(eZ)^+ + \int_0^{+\infty} dz(eZ)^+ e^{-\frac{t^2}{8n'}} + \int_{1/e}^{+\infty} dz e^{Z} e^{-\frac{t^2}{8n'} - \frac{1}{6t^2}} \\
\leq (2eCn')^{r+1} + (Cn')^{(r+1)} e^{(r+1)} e^{-\frac{t^2}{2} + e^{1/(n' \varepsilon^2 - \frac{1}{6t^2})}}
\]

instead of (6.51). Hence all estimates contained in the proof of Corollary 6.7 hold if one replaces \( A/e \) by \( A/n' \varepsilon \) or \( A/\sqrt{n'} \varepsilon \).

\[\Box\]

### 6.3.2 Mayer expansion

We also need in the course of the proof of Theorem 6.2 a technical result which we choose to state separately for the sake of clarity. In the sequel, \( B^j \) is one of the \( \mu \)-dependent subsets of cubes with \( \delta^j(x, \rho) \subset B^j \subset \delta^j(x, \rho) \cup \partial \delta^j(x, \rho) \) introduced in section 6.2.

**Lemma 6.9 ("Mayer expansion")** Let \((z_{\Delta})_{\Delta \in B^j(x, \rho)} \in \mathbb{R}_+\) and \( c > 0 \). Define

\[
y_{\Delta} := \sum_{\Delta \in B_j(x, \rho)} e^{-cyd(\Delta, \Delta')} z_{\Delta'},
\]

(see eq. (6.29)).

(i) ("Mayer expansion") Let

\[
S_\delta((z_{\Delta})) := \sum_{\Delta \in B_j(x, \rho)} \left(e^{-cyd} e_{\Delta'} - 1\right), \quad \delta \geq 0
\]

(6.62)
and
\[ S_0(y_\Delta) := \sum_{\Delta \in B_j} (e^{y_\Delta} - 1). \] (6.63)

Then
\[ 0 \leq S_0(y_\Delta) \leq \sum_{m \geq 1} \frac{1}{(m-1)!} \sum_{\delta_1 < \ldots < \delta_m} m! \prod_{p=1}^m S_{\delta_p}((z_\Delta)) \] (6.64)

where \( \delta_i, i = 1, 2, \ldots \) range among the set \( \{d^j(\Delta, \Delta'), \Delta, \Delta' \in B^j\} \).

More generally, if \( \delta_{\text{max}} \in \mathbb{R}_+ \), then
\[ S_0((y_\Delta)) \leq \sum_{m \geq 1} \frac{1}{(m-1)!} \sum_{\delta_{\text{max}} < \delta_1 < \ldots < \delta_m} m! \prod_{p=1}^m S_{\delta_p}((z_\Delta)) \] (6.65)

(ii) Let
\[ T((z_\Delta)) := \sum_{\Delta \in B_j(x, \rho)} e^{y_\Delta} \] (6.66)

and similarly
\[ T((y_\Delta)) := \sum_{\Delta \in B_j} e^{y_\Delta}. \] (6.67)

Then
\[ T((y_\Delta)) \leq (T((z_\Delta)))^{1+O(e^{-c})}. \] (6.68)

Note that (by invariance by translation) \( \delta_i \in \{d^j(\Delta, \Delta'), \Delta, \Delta' \in \mathbb{D}^j\} = \{d_1 < d_2 < \ldots\} \) where \( \Delta \) is some arbitrary fixed cube, and \( d_i \sim_{l \to \infty} i^{j/d} \). The indexation is easier if we choose a distorted distance instead of the Euclidean distance, i.e., if we define e.g. \( |x - y| = \sup_{l=1, \ldots, d} |(R(x - y))| \) where \( R \) is a generic rotation, so that the set \( \{\Delta' \in \mathbb{D}^j \mid d^j(\Delta, \Delta') = \delta\} \) contains at most one element for \( \Delta, \delta \) fixed, which we denote by \( \Delta(\delta) \). The nickname ”Mayer expansion” refers to a common expansion of the free energy in equilibrium statistical physics where \( e^{\beta H}, \beta H \) being the local energy density, is expanded into \((e^{\beta H} - 1) + 1\), which is exactly what we do in (i).

Proof.

(i) Let \( a_\Delta(\delta) := e^{-c_\delta e^{y_\Delta}} - 1 \) (\( \Delta \in B_j(x, \rho) \)) and \( a_\Delta := a_\Delta(0) = e^{y_\Delta} - 1 \). Rewriting \( S_0((y_\Delta)) \) in

49
terms of the z-variables and expanding the product of terms of the form \(a_{\Delta_0}(\delta) + 1\) yields

\[ S_0(y_\Delta) = \sum_{\Delta \in \B^1} \left[ \prod_{\delta} (a_{\Delta_0}(\delta) + 1) \right] - 1 = \sum_{\Delta \in \B^1} \sum_{m=1}^{n} \sum_{\delta_1 < \cdots < \delta_m} \prod_{p=1}^{m} a_{\Delta_0^p}(\delta_p) \]

\[ \leq \sum_{m=1}^{n} \frac{1}{(m-1)!} \sum_{\delta_1 < \cdots < \delta_m} \sum_{\Delta \in \B^1} a_{\Delta_1}(\delta_1) \left( \sum_{\Delta_2 \in \{\Delta \cup \{\delta_2\}, \ldots, \Delta \cup \{\delta_m\}\}} \prod_{\delta_2} a_{\Delta_2}(\delta_2) \right) \] \tag{6.69}

\[ \leq \sum_{m=1}^{n} \frac{1}{(m-1)!} \sum_{\delta_1 < \cdots < \delta_m} \sum_{\Delta \in \B^1} \prod_{\delta_2} a_{\Delta_2}(\delta_2) \] \tag{6.70}

For the proof of (6.65), we fix \(\delta_{\max} \geq 0\), start from (6.69) and pick its \(p\)-th factor, \(A_p = \sum_{\Delta \in \B^1} a_{\Delta}(\delta_p)\). If \(\delta_p > \delta_{\max}\) we bound \(A_p\) by \(S_{\delta_p}(\{\delta_\Delta\})\) as before. Otherwise we use the identity \((x - 1)^k \leq x^k - 1 (x \geq 1, k \geq 1)\) and Hölder’s inequality to get

\[ a_{\Delta_p}(\delta_p) \leq (e^{\varepsilon_{\Delta_p}} - 1)^{e^{-c_0 \rho}} = a_{\Delta_p}^{e^{-c_0 \rho}} \quad \text{and} \]

\[ A_p \leq (m - p + 1) \sum_{\Delta \in \B^1} a_{\Delta} = (m - p + 1)S_0(y_\Delta) e^{e^{-c_0 \rho}}. \tag{6.71} \]

Finally, \(\sum_p e^{-c_0 \rho} \leq 1 + e^{-c_0 \rho} + e^{-c_0 \rho^2} + \cdots = 1 + O(e^{-c_0}).\)

(ii) One finds (all sums or supremums in the next expressions range over subsets of \(\tilde{B}(x, \rho)\), unless otherwise stated)

\[ \sum_{\Delta \in \B^1} e^{y_{\Delta}} \leq \sum_{\Delta \in \B^1} e^{y_{\Delta}} \prod_{\Delta' \neq \Delta} e^{-c_0 \rho_{\Delta \Delta'}} \]

\[ \leq \frac{1}{(n-1)!} \sum_{\Delta_1} e^{c_\Delta_1} \left( \sum_{\Delta_2 \neq \Delta_1} e^{-c_0 \rho_{\Delta_1 \Delta_2}} \left( \sum_{\Delta_3 \neq \Delta_1, \Delta_2} e^{c_\Delta_3} e^{-c_0 \rho_{\Delta_1 \Delta_3}} (\ldots) \right) \right) \]

\[ \leq \frac{1}{(n-1)!} \left( \sum_{\Delta_1} e^{c_\Delta_1} \right) \sup_{\Delta_1} \left( \sum_{\Delta_2 \neq \Delta_1} e^{-c_0 \rho_{\Delta_1 \Delta_2}} \right) \ldots \]

\[ \leq \frac{1}{(n-1)!} \left( \sum_{\Delta_1} e^{c_\Delta_1} \right) \left[ (n-1) \left( \sum_{\Delta_2} e^{c_\Delta_1} \right) \right] \ldots \] \tag{6.72}

(Hölder’s inequality was used in the last line). The product of the prefactors in the last expression, \((n-1)(n-2)\cdots\) is exactly compensated by the factorial \(\frac{1}{(n-1)!}\), and there remains

\[ \sum_{\Delta \in \B^1} e^{y_{\Delta}} \leq \left( \sum_{\Delta \in \B^1(x, \rho)} e^{c_\Delta} \right)^{1+O(e^{-c_0})}. \tag{6.73} \]
Again, this lemma has a block generalization. Roughly speaking, we want to group together all cubes $\Delta'$ at distance $\delta \approx 3^k$ of a given cube $\Delta$ and sum over $k$, instead of summing over the $\delta_i$'s which (as a detailed computation proves) increase too slowly with $i$ to give a converging series. Actually, we bother to do so only for $\delta > \delta_{\text{max}}$, in a region where the exponential decay governs essentially the estimates; the value of $\delta_{\text{max}}$ is fixed later in the text. In order to avoid blocks with "holes" and overlaps between blocks, we introduce the following definitions. Let $\mathcal{D}_k \geq \log_3 \delta_{\text{max}}$ be the set of blocks $\mathcal{D}_k = [3^kM^{1/2}k_1, 3^kM^{1/2}(k_1+1)] \times \ldots \times [3^kM^{1/2}k_d, 3^kM^{1/2}(k_d+1)]$ of size $3^k$ included in $B(x, \rho)$. The $3^d - 1$ blocks of size $3^k$, $[x_{\Delta,1} + \varepsilon_1 3^kM^{1/2}(k_1 - \frac{1}{2}), x_{\Delta,1} + \varepsilon_1 3^kM^{1/2}(k_1 + \frac{1}{2})] \times \ldots \times [x_{\Delta,d} + \varepsilon_d 3^kM^{1/2}(k_d - \frac{1}{2})x_{\Delta,d} + \varepsilon_d 3^kM^{1/2}(k_d + \frac{1}{2})]$, where $\varepsilon = (\varepsilon_1, \ldots, \varepsilon_d) \in \{-1, 0, 1\}^d \setminus \{0, \ldots, 0\}$, are all situated at a scaled distance $\delta = 3^k$ of $\Delta$. We denote them by $\Delta(\delta), \delta = (\delta, \varepsilon)$ is a composite index including both the distance $\delta$ and a discrete index $\varepsilon$ ranging in a fixed finite set. Then, for smaller distances $\delta < 3^k$, we set $\Delta(\delta) = \Delta(\delta)$ as in the previous lemma, and $\delta = \delta$ simply. All together the blocks $\Delta(\delta)_k, \delta = (\delta, \varepsilon) (\delta \geq 3^k)$ or $\delta (\delta < 3^k)$ define for every fixed cube $\Delta$ a partition of $\mathbb{R}^d$. We choose in the sequel some arbitrary total ordering $< \delta$ of the indices $\delta$ such that $(\delta = (\delta, \varepsilon))$, or $\delta, \delta' = (\delta', \varepsilon') < \delta'$, or $\delta < \delta'$ $\Rightarrow \delta' < \delta'$.

**Lemma 6.10 (block "Mayer expansion")** Let $(z_{\Delta})_{\Delta \in B_j} \in \mathbb{R}_+$ and $c > 0$. Define as in the previous lemma

$$y_{\Delta} := \sum_{\Delta' \in B'(x, \rho)} e^{-co d/(\Delta, \Delta')} z_{\Delta'}, \quad S_0((y_{\Delta})) := \sum_{\Delta \in B'} (e^{y_{\Delta}} - 1), \quad S_0((z_{\Delta})) := \sum_{\Delta \in B'(x, \rho)} (e^{z_{\Delta}} - 1)$$

(6.74)

and let, for $k \in \mathbb{N}$,

$$S_{3^k}((z_{\Delta})) := \sum_{\Delta' \in \mathcal{D}_k} (e^{z_{\Delta}} - 1)$$

(6.75)

a block version of (6.62) distinguished by the boldface letter. Choose some value of $\delta_{\text{max}}$ and order the indices $\delta$ as indicated above. Then

$$S_0((y_{\Delta})) \leq \sum_{m \geq 1} \left( \frac{1}{(m-1)!} \right) \sum_{\delta_{\text{max}} < \delta_1 < \ldots < \delta_m} \prod_{p=1}^{m} S_{\delta_p}(z_{\Delta})) + \sum_{m \geq 1} \sum_{k=1}^{m-1} \sum_{\delta_1 < \ldots < \delta_m < \delta_{\text{max}}} S_0((z_{\Delta}))^{1+O(e^{-c})} \prod_{p=1}^{k} S_{\delta_p}(z_{\Delta})) + \sum_{m \geq 1} \sum_{\delta_1 < \ldots < \delta_m \leq \delta_{\text{max}}} S_0((z_{\Delta}))^{1+O(e^{-c})}$$

(6.76)

**Proof.** If $\Delta$ is a block of size $3^k$ for some $k \geq 0$, we let $a_{\Delta}(\delta) := e^{-c \delta} \sum_{\Delta A} z_{\Delta} - 1$. Thus

$$S_0((y_{\Delta})) \leq \sum_{\Delta \in B'} \left| \prod_{\delta} a_{\Delta(\delta)}(\delta) + 1 \right| - 1 = \sum_{\Delta \in B'} \sum_{m \geq 1} \sum_{\delta_1 > \ldots > \delta_m} \prod_{p=1}^{m} a_{\Delta(\delta_p)}(\delta_p).$$

(6.77)
We then expand as in (6.69) and (6.70), and forget the unnecessary factorials in the denominator (which would require a short discussion in any case since there is no symmetry factor for terms belonging to blocks with different sizes).

6.3.3 Proof of Theorem 6.2

We shall use several times the following elementary lemma.

**Lemma 6.11** Let $X_1, \ldots, X_n$, $n \geq 1$ be real-valued random variables, and $\lambda_1, \ldots, \lambda_n \in \mathbb{R}_+$ such that $\lambda_1 + \ldots + \lambda_n = 1$. Then

$$\mathbb{P}[\lambda_1 X_1 + \ldots + \lambda_n X_n > A] \leq \mathbb{P}[\sup_{p=1}^n X_p > A] \leq \sum_{p=1}^n \mathbb{P}[X_p > A]. \quad (6.78)$$

**Proof of Theorem 6.2** The general scheme of the proof is the same as that of Lemma 6.5. Applying Proposition 6.2 with $\phi(t) = \psi\left(\frac{1}{e} \int_{B(x,M^{1/2} \rho) \cap \Delta_p^j} dy\left(e^{\lambda M |\eta_0^j|} - 1 \right)\right)$ where $\psi : \mathbb{R}_+ \to \mathbb{R}$ is any convex, increasing function yields instead of (6.26)

$$\mathbb{E}\left[\psi\left(\frac{1}{e} \int_{B(x,M^{1/2} \rho) \cap \Delta_p^j} dy\left(e^{\lambda M |\eta_0^j|} - 1 \right)\right)\right] \leq \mathbb{E}\left[\psi\left(\frac{1}{e} \int_{B(x,M^{1/2} \rho) \cap \Delta_p^j} dy\left(e^{\lambda M |\eta_1^j|} - 1 \right)\right)\right].$$

Then we bound the last integral by sums of local suprema $\tilde{Y}_{\Lambda} = \sup_{\Lambda} |\eta^j(\gamma)|$,

$$\int_{B(x,M^{1/2} \rho) \cap \Delta_p^j} dy e^{\lambda M |\eta_1^j|} \leq \max_{\Lambda^j} \frac{\sum_{\Delta \in B^j} e^{\lambda M \tilde{Y}_{\Lambda}}}{\#B^j} \quad (6.79)$$

where $B^j$ is a union of cubes ranging over a finite set as in subsections 6.1 and 6.2, and $\#B^j \leq \#B(x, \rho) = n$. Eq. (6.79), (6.30) imply

$$\lambda M \tilde{Y}_{\Lambda} \leq C e \sum_{\Delta \in B(x, \rho)} e^{-c_0 d(\Delta, \Lambda')} |Z_{\Lambda'}| + 1 \leq C' e + C' e \sum_{\Delta \in B(x, \rho)} e^{-c_0 d(\Delta, \Lambda')} |Z_{\Lambda'}| \quad (6.80)$$

Finally, we use the formula

$$\mathbb{E}\left[\psi\left(\frac{1}{n^E} \sum_{\Delta \in B^j} \left(e^{\lambda M \tilde{Y}_{\Lambda}} - 1 \right)\right)\right] = \int_0^{+\infty} dA \psi'(A) \mathbb{P}\left[\sum_{\Delta \in B^j} e^{\lambda M \tilde{Y}_{\Lambda}} > n(1 + eA)\right] + \psi(0) \quad (6.81)$$

Clearly, $e^{-C' e}(1 + eA) \geq (1 - C' e)(1 + eA) \geq 1 + ceA$ if $A \gg 1$. Hence, assuming supp$\psi' \in [C, +\infty]$ with $C$ large enough, one finds

$$\mathbb{E}\left[\psi\left(\frac{1}{n^E} \sum_{\Delta \in B^j} \left(e^{\lambda M \tilde{Y}_{\Lambda}} - 1 \right)\right)\right] \leq \int_0^{+\infty} dA \psi'(A) \mathbb{P}\left[\sum_{\Delta \in B^j} e^{\lambda Y_{\Lambda}} > n(1 + eA)\right] + \psi(0) \quad (6.82)$$

with

$$Y_{\Lambda} := \sum_{\Delta \in B(x, \rho)} e^{-c_0 d(\Delta, \Lambda')} |Z_{\Lambda'}| \quad (6.83)$$
Hence \( P \left[ \sum \epsilon_\lambda Y > n(1 + \epsilon A) \right] \leq e^{-c \ln^2(nA)}, \quad A \gg 1. \) (6.84)

Hence \( P \left[ \sum \epsilon_\lambda Y > n(1 + \epsilon A) \right] \leq \psi^{-2}(A) \) with

\[
\psi(A) := e^{c \ln^2(nA-C)} \mathbf{1}_{A>C(1+\frac{1}{n})} + e^{c \ln^2(C)} \mathbf{1}_{A<C(1+\frac{1}{n})},
\]

Therefore,

\[
\mathbb{E} \left[ \psi \left( \frac{1}{n} \sum_{\Delta \in B^n} (e^{\lambda M \tilde{Y}_\lambda} - 1) \right) \right] \leq \left[ -\frac{1}{\psi(A)} \right]_0^{+\infty} + \psi(0) = O(1)
\]

and, by Markov’s inequality,

\[
P \left[ \frac{1}{n} \sum_{\Delta \in B^n} e^{\lambda M \tilde{Y}_\lambda} > 1 + \epsilon A \right] \leq \frac{1}{\psi(A)} \leq e^{-c \ln^2(nA)}, \quad A \gg 1
\]

(6.87)

so

\[
P \left[ \ln \left( \frac{1}{n} \sum_{\Delta \in B^n} e^{\lambda M \tilde{Y}_\lambda} \right) > \epsilon A \right] = P \left[ \frac{1}{n} \sum_{\Delta \in B^n} e^{\lambda M \tilde{Y}_\lambda} > e^{\epsilon A} \right] \leq e^{-c \ln^2(nA)}.
\]

(6.88)

(Note that, if \( \epsilon A \gg 1 \), one obtains in fact a Gaussian queue distribution,

\[
P \left[ \ln \left( \frac{1}{n} \sum_{\Delta \in B^n} e^{\lambda M \tilde{Y}_\lambda} \right) > \epsilon A \right] \leq e^{-c(\epsilon A + \ln \frac{1}{n})^2} \leq e^{-c(\epsilon A)^2},
\]

(6.89)

with a very bad coefficient \( \epsilon \) in front of \( A \) however.) Thus

\[
P \left[ \ln \left( e^{\lambda M \tilde{Y}_\lambda} \right) > \epsilon A \right] \leq \sum_{n \geq 1} e^{-c \ln^2(nA)} \leq A^{-c \ln(A)}.
\]

(6.90)

It remains to prove the key estimate (6.84). Recall \( Y_\lambda = \sum_{\Delta \in B(\lambda, \phi)} e^{\epsilon_\lambda y_\lambda |Z_\Delta|}, \) where \( (Z_\Delta)_\Lambda \) are i.i.d. standard Gaussian variables. It turns out that there are four different large deviation regimes, according to the value of \( S_0(\|Z_\Delta\|_0) = \sum_{\Delta \in B^n} \left( e^{\|Z_\Delta\|} - 1 \right) \), written as \( \hat{S}_n \) in Corollary (6.8) or \( S_0((Y_\lambda)_\Lambda) = \sum_{\Delta \in B^n} \left( e^{Y_\lambda} - 1 \right) \) (see Lemma (6.10). By assumption

\[
\sum_{\Delta \in B^n} e^{Y_\lambda} > n(1 + \epsilon A)
\]

(6.91)

with \( A \gg 1 \); in other terms,

\[
S_0((Y_\lambda)_\Lambda) > n \epsilon A.
\]

(6.92)
(i) (Gaussian regime) Assume \( S_0((\Delta Z)|\Delta) = \bar{S}_n \leq 2 \). Then \( \varepsilon[Z] = O(1) \) for all \( \Delta \), hence \( \varepsilon Y_{\Delta} = O(1) \) too, so \( S_0((\Delta Z)|\Delta) \leq S_0((\Delta Z)|\Delta) = \varepsilon \sum_{\Delta B(\Delta)} |Z_B| \). Therefore

\[
P[\sum_{\Delta B(\Delta)} e^{Y_{\Delta}} > n(1 + \varepsilon A)] \leq P[S_0((\Delta Z)|\Delta) > neA] \leq P[\sum_{\Delta B(\Delta)} |Z_B| > cnA] \leq e^{-c'n(A-C)^2}.
\] (6.93)

Clearly this last quantity is bounded by \( e^{-c''\ln^2(nA)} \) for \( A \) large enough.

(ii) (very large deviation regime) Assume \( S_0((\Delta Z)|\Delta) \gg n^{1+O(e^{-\varepsilon_0})} \), or equivalently \( A \varepsilon \gg n^{O(e^{-\varepsilon_0})} \). Then the "Mayer expansion" (see Lemma 6.9(i)) is not needed; we use Lemma 6.9(ii) and find successively \( S_0((\Delta Z)|\Delta) \approx T((\Delta Z)|\Delta) \leq (T((\Delta Z)|\Delta))^{1+O(e^{-\varepsilon_0})} \), so \( T((\Delta Z)|\Delta) \gg n \), hence again \( T((\Delta Z)|\Delta) \approx S_0((\Delta Z)|\Delta) \). All together we have found \( S_0((\Delta Z)|\Delta) \leq (S_0((\Delta Z)|\Delta))^{1+O(e^{-\varepsilon_0})} \).

So we may apply Corollary 6.7 to the result that

\[
P[\sum_{\Delta B(\Delta)} e^{Y_{\Delta}} > n(1 + \varepsilon A)] \leq P[\bar{S}_n > (n(1 + \varepsilon A))^{1/(1+O(e^{-\varepsilon_0}))}] \leq e^{-c''\ln^2(nA)}.
\] (6.94)

(iii) Assume \( ne \leq 1 \) and \( A \varepsilon \leq n^{O(e^{-\varepsilon_0})} \). Then we use the generalized block "Mayer expansion" (6.76) with \( \delta_{\text{max}} = 0 \):

\[
P[S_0((\Delta Z)|\Delta) > neA] \leq \prod_{m \geq 1} \prod_{\delta_1 > \ldots > \delta_m} \prod_{p=1}^{m} S_{\delta_p}((\Delta Z)|\Delta) > neA \]

Since \( \sum_{m \geq 1} \sum_{\delta_1 > \ldots > \delta_m} e^{-\sum_{p=1}^{m} (\delta_1 + \ldots + \delta_m)} \approx \sum_{m \geq 1} \frac{1}{m!} \sum_{\delta_1 > \ldots > \delta_m} e^{-\sum_{p=1}^{m} (\delta_1 + \ldots + \delta_m)} = 1 \), we get by Lemma 6.11

\[
P[S_0((\Delta Z)|\Delta) > neA] \leq \sum_{m \geq 1} \sum_{\delta_1 < \ldots < \delta_m} \prod_{p=1}^{m} S_{\delta_p}((\Delta Z)|\Delta) > neA \]

\[
\leq \sum_{m \geq 1} \sum_{\delta_1 < \ldots < \delta_m} \prod_{p=1}^{m} S_{\delta_p}((\Delta Z)|\Delta) > e^{\frac{n\varnothing}{(ne)^{\frac{1}{m}}}} 1^{1/m}
\]

(6.96)

where \( \bar{\delta} := \frac{1}{m}(\delta_1 + \ldots + \delta_m) \). Note that the expression \( S_{\delta_p} \) is a sum over blocks of size \( n' \approx \delta_p^{\delta_p} \ll (\varepsilon e^{-c'\varepsilon_0})^{-1} \), and \( e^{\frac{m\bar{\delta}}{(ne)^{\frac{1}{m}}} A 1/m} \gg 1 \), hence we are in the large deviation regime studied in Corollary 6.8.

Also,

\[
ne^{\frac{m\varnothing}{(ne)^{\frac{1}{m}}} A 1/m} \geq ne^{\frac{m\varnothing}{(ne)^{\frac{1}{m}}} A 1/m} \geq e^{\frac{m\bar{\delta}}{\varnothing} A 1/m - O(e^{-\varepsilon_0})}
\]

(6.97)

if \( m \geq 2 \). For \( m = 1 \) the estimates of Corollary 6.8 give directly a log-normal queue, so we sum over \( m \geq 2 \). By construction \( \delta_m \geq 3^{m/(3^{\delta_m})}, \) so \( \bar{\delta} \geq \frac{\delta_m}{m} \geq \sqrt{\delta_m} \) and \( n' \approx \delta_p^{\delta_p} \ll (\varepsilon e^{-c'\varepsilon_0})^{-1} \).

Hence

\[
P[S_{\delta_p}((\Delta Z)|\Delta) > \frac{n\varnothing}{(ne)^{\frac{1}{m}}} 1^{1/m}] \leq e^{-\ln^2\left(\frac{m\bar{\delta}}{(ne)^{\frac{1}{m}}} A\right)} \leq e^{-c\ln^2\left(\frac{n\varnothing}{A}\right)} \leq e^{-c\ln^2\left(nA - c\ln(nA)\right)}.
\]

(6.98)
Let $V_m(r) := \#(\delta_1, \ldots, \delta_m) \mid \delta < r = \#(\delta_1, \ldots, \delta_m) \mid \sum_{p=1}^m \delta_p < mr$: clearly $V_m(r) \leq \#(i_1, \ldots, i_m) \in \mathbb{N}^m \mid i_1 + \ldots + i_m = mr$, hence

$$V_m(r) \leq \int_0^{+\infty} dx_1 \ldots dx_m \sum_{\delta_p < mr} = \frac{(mr)^m}{m!} \leq (Cr)^m. \quad (6.99)$$

Thus (with an extra factor $\frac{1}{m}$ due to the ordering $\delta_1 > \ldots > \delta_m$)

$$\sum_{m \geq 2} \sum_{\delta_1 > \ldots > \delta_m} \mathbb{P} \left[ S_{\delta_p}((\lfloor Z \rfloor)_\Lambda) > e^{\frac{\delta}{c_0} \ln(nA)} \frac{1}{m} A^{1/m} \right] \leq \sum_{m \geq 1} \frac{1}{m!} (nA)^{-c \ln(nA)} \sum_{r=0}^{+\infty} (V_m(r) - V_m(r - 1))e^{-cr^2} \leq (nA)^{-c \ln(nA)} \sum_{r=0}^{+\infty} e^{-cr^2} \sum_{m \geq 0} V_m(r) \leq (nA)^{-c \ln(nA)} \leq \mathbb{P}((\lfloor 1, \ldots, \delta_{\max} \rfloor)) = 2^{\delta_{\max}} = (ne)^{2\ln(2)/c_0}, \quad (6.100)$$

(iv) Finally, assume $ne \geq 1$ and $Ae \leq n^{O(e^{-c_0})}$, and apply the generalized "Mayer expansion" with $\delta_{\max} = \frac{\delta}{c_0} \ln(nA)$ defined in such a way that $e^{\frac{\delta}{c_0}} \geq ne$ for $\delta \geq \delta_{\max}$. Since

$$\sum_{m \geq 1} \sum_{m' > 0} e^{-\frac{\delta}{c_0} \ln(nA)} \sum_{\delta_1 \ldots \delta_m} \mathbb{P} \left[ S_0((\lfloor Z \rfloor)_\Lambda)^{1+O(e^{-c_0})} \sum_{p=m+m' + 1}^m \frac{S_{\delta_p}((\lfloor Z \rfloor)_\Lambda)}{ne^{c_0} \delta_p} > (ne)^{1-2\ln(2)/c_0} \frac{A}{c_0} \prod_{p=m+m' + 1}^m \frac{e^{\frac{\delta}{c_0}}}{ne} \right]$$

$$\leq \sum_{m \geq 1} \left\{ \sum_{\delta_1 \ldots \delta_m} \mathbb{P} \left[ S_0((\lfloor Z \rfloor)_\Lambda) \frac{1+O(e^{-c_0})}{\delta_p} > e^{-2\ln(2)/c_0} \frac{A}{c_0} \prod_{p=m+m' + 1}^m \frac{e^{\frac{\delta}{c_0}}}{ne} \right] \right\} \quad (6.102)$$

$$\leq \sum_{m \geq 1} \left\{ \sum_{\delta_1 \ldots \delta_m} \mathbb{P} \left[ S_0((\lfloor Z \rfloor)_\Lambda) \frac{1+O(e^{-c_0})}{\delta_p} > e^{-2\ln(2)/c_0} \frac{A}{c_0} \prod_{p=m+m' + 1}^m \frac{e^{\frac{\delta}{c_0}}}{ne} \right] \right\} \quad (6.103)$$

where $\tilde{\delta} := \frac{1}{m-m'}(\delta_{m' + 1} + \ldots + \delta_m)$.

For $c_0$ large enough (recall $c_0$ has been multiplied by $m_0$, thus it suffices to choose $m_0$ large enough)

$$n^{1-2\ln(2)/c_0} e^{-2\ln(2)/c_0} e^{\frac{\delta}{c_0}} A \geq e^{\frac{\delta}{c_0}} A n^\kappa \geq n^\kappa \quad (6.104)$$

with $\kappa > \frac{1}{2}$, and

$$\frac{e^{\frac{\delta}{c_0}}}{n^\kappa} \geq \max \left( n, \frac{e^{\frac{\delta}{c_0}} (ne)^{2/3}}{\delta_p} \right) \geq \max \left( n, e^{\frac{\delta}{c_0}} A \frac{n^\kappa}{\kappa} O(e^{-c_0}) \right). \quad (6.105)$$
Thus we are in the large deviation regime (see remark after Corollary 6.7 and Corollary 6.8), and the lower bounds are as in (6.97), yielding a bound \(O((nA)^{-c\ln(nA)})\) for the sum (6.103) over \(m'\) and \(\delta_{m'+1}, \ldots, \delta_m\). As for the first sum over \(\delta_1 < \ldots < \delta_m \leq \delta_{\text{max}}\) in (6.102), or \(\delta_1 < \ldots < \delta_{m'}\) in (6.103), it produces as in (6.101) a supplementary multiplicative factor of order \((ne(2\ln(2)/c))^{nA} - c\ln(nA)\) of no incidence on the result since \((ne)^2(2\ln(2)/c)(nA)^{-c\ln(nA)} \leq n^2(2\ln(2)/c)(nA)^{-c\ln(nA)} \leq (nA)^{-c\ln(nA)}\).

\[ \square \]

### 6.4 Proof of Theorem 6.1

We may now finally prove Theorem 6.1. Let

\[
F(\delta t, [0, t]) := M^{-j} \delta t \sum_{p=0}^{\lfloor j/\delta t \rfloor} \left( e^{-M^{j/\delta t}} \right)^p \ln \left( \left( e^{\lambda M^{j/\delta t}} \delta t^{-\epsilon / \delta t} \eta(x) dx \right)^x \right),
\]

and

\[
c_{\delta t} := \left[ M^{-j} \delta t \sum_{p=0}^{\lfloor j/\delta t \rfloor} \left( e^{-M^{j/\delta t}} \right)^p \right]^{-1} \geq 1.
\]

Clearly,

\[
F(\delta t, [0, t]) \leq M^{-j} \delta t \sum_{p=0}^{\lfloor j/\delta t \rfloor} \left( e^{-M^{j/\delta t}} \right)^p \ln \left( \left( e^{\lambda M^{j/\delta t}} \delta t^{-\epsilon / \delta t} \eta(x) dx \right)^x \right)
\]

where \(\|\tilde{\eta}^j\|_{\infty, \{t-p\delta t, t-(p-1)\delta t\}}(x) := \sup_{s \in \{t-p\delta t, t-(p-1)\delta t\}} |\tilde{\eta}^j(s, x)|\). Thus

\[
c_{\delta t} F(\delta t) \leq \sup_p \left( e^{-M^{j/\delta t}} \right)^p \ln \left( \left( e^{\lambda M^{j/\delta t}} \delta t^{-\epsilon / \delta t} \eta(x) dx \right)^x \right)
\]

\[
\leq \sup_{q \in \mathbb{N}} e^{-q} \ln \left( \left( e^{\lambda M^{j/\delta t}} \delta t^{-\epsilon / \delta t} \eta(x) dx \right)^x \right).
\]

The estimates we developed for \(|\eta^j_j|\) in Theorem 6.2 extend to \(|\tilde{\eta}^j\|_{\infty, \{t-qM^{j}, t-(q-1)M^{j}\}}(x)\) by using the BTIS inequality once again. Hence

\[
\mathbb{P}(F(\delta t, [0, t]) > \epsilon A) \leq \sum_q \mathbb{P} \left( \ln \left( \left( e^{\lambda M^{j/\delta t}} \delta t^{-\epsilon / \delta t} \eta(x) dx \right)^x \right) > \epsilon \epsilon A \right)
\]

\[
\leq \sum_q \left( \epsilon A \right)^{-c\epsilon \ln(\epsilon A)} \leq A^{-c\epsilon \ln A}
\]

by Theorem 6.2.

Now, by Hölder’s inequality,

\[
F(\delta t, [0, t]) \leq \frac{1}{2} M^{-j} \delta t \sum_{p=0}^{\lfloor j/\delta t \rfloor} \left( e^{-M^{j/\delta t}} \right)^p \left\{ \ln \left( \left( e^{\lambda M^{(j-p+1)/\delta t}} \delta t^{-\epsilon / \delta t} \eta(x) dx \right)^x \right) + \ln \left( \left( e^{\lambda M^{j/\delta t}} \delta t^{-\epsilon / \delta t} \eta(x) dx \right)^x \right) \right\}
\]

\[
\leq e^{M^{j/\delta t}/2} F(\delta t/2, [0, t])
\]

(6.111)
so \( \|\eta\|_{L_1(R^+, x)} \leq \limsup_{\delta t \to 0, t \to +\infty} \frac{1}{t} F(\delta t, [0, t]) \) by (4.17), and by monotone convergence we get \( \mathbb{P}[\|\eta\|_{L_1(R^+, x)} > AM^{-j\delta t}] \leq \limsup_{\delta t \to 0, t \to +\infty} \mathbb{P}[F(\delta t, [0, t]) > \varepsilon A] \leq A^{-c \ln A} \).

The estimates for \( \|M^{1/2}(\cdots \delta x \cdots) - \eta(\cdots)\|_{L_1(R^+, x)} \) are proved in the same way since \( M^{1/2}(\cdots \delta x \cdots) - \eta(\cdots) = O((M^{-j})^{1 + \delta t}) \) scales like \( \eta_j(x) \) (see [5,20]).

\[ \Box \]

Acknowledgements. We wish to express our thanks to S. Benachour and G. Barles for useful discussions.

References

[1] A. Abdesselam, V. Rivasseau. Trees, forests and jungles: A botanical garden for cluster expansions, Lecture Notes in Physics 446, Springer Berlin/Heidelberg (1995).

[2] R. J. Adler. The geometry of random fields, SIAM (2010).

[3] A. Altland, B. Simons. Condensed matter field theory, Cambridge University Press (2010).

[4] G. Amir, I. Corwin, J. Quastel. Probability distribution of the free energy on the continuum directed random polymer in 1+1 dimensions, Comm. Pure Applied Math. 64, 466-537 (2011).

[5] L. Amour, M. Ben-Artzi. Global existence and decay for viscous Hamilton-Jacobi equations, Nonlinear Analysis, Theory, Methods and Applications 31, 621–628 (1998).

[6] A.-L. Barabasi, H. E. Stanley. Fractal concepts in surface growth. Cambridge University Press (1995).

[7] G. Barles. Solutions de viscosité et équations elliptiques du deuxième ordre, (graduate course given at the University of Tours, available at: www.lmpt.univ-tours.fr/~barles/Toulcours.pdf (1997).

[8] S. Benachour, M. Ben-Artzi, P. Laurençot. Sharp decay estimates and vanishing viscosity for diffusive Hamilton-Jacobi equations, Adv. Differential Equations 14, 1–25 (2009).

[9] S. Benachour, G. Karch, P. Laurençot. Asymptotic profiles of solutions to viscous Hamilton-Jacobi equations, J. Math. Pures Appl. 83, 1275–1308 (2004).

[10] S. Benachour, P. Laurençot. Global solutions to viscous Hamilton-Jacobi equations with irregular initial data, Comm. Partial Diff. Eq. 24, 1999-2021 (1999).

[11] A. J. Majda, A. L. Bertozzi. Vorticity and incompressible flow, Cambridge texts in applied mathematics, Cambridge University Press (2002).

[12] P. Billingsley. Convergence of probability measures, Wiley (1968).

[13] H. Brézis. Opérateurs maximaux monotones et semi-groupes de contractions dans les espaces de Hilbert, North-Holland (1973).

[14] J. Bricmont, K. Gawedzki, A. Kupiainen. KAM theorem and quantum field theory, Comm. Math. Phys. 201, 699-727 (1999).
[15] J. Bricmont, A. Kupiainen. *Random walks in asymmetric random environments*, Comm. Math. Phys. **142** (2), 345–420 (1991).

[16] J. Bricmont, A. Kupiainen, G. Lin. *Renormalization group and asymptotics of solutions of nonlinear parabolic equations*, Comm. Pure Appl. Math. **47**, 893–922 (1994).

[17] J. Cardy. *Field theory and nonequilibrium statistical mechanics*, cours de 3ème cycle de la Suisse Romande, semestre d’été de l’année académique 1998-1999 (in English).

[18] P. Carmona, Y. Hu. *On the partition function of a directed polymer in a random environment*, Probab. Theory Related Fields **124**, 431–457 (2002).

[19] F. Comets, N. Yoshida. *Directed polymers in random environment are diffusive at weak disorder*, Ann. Prob. **34**, 1746–1770 (2006).

[20] I. Corwin, *The Kardar-Parisi-Zhang equation and universality class*, Random Matrices Theory Appl. **1** (1), 113001 (2012).

[21] M. G. Crandall, H. Ishii, P.-L. Lions. *User’s guide to viscosity solutions of second order partial differential equations*, Bull. AMS **27** (1), 1–67 (1992).

[22] F. Da Lio, O. Ley. *Uniqueness results for second order Bellman-Isaacs equations under quadratic growth assumptions and applications*, SIAM J. Control Optim. **45** (1), 74–106 (2006).

[23] F. Da Lio, O. Ley. *Uniqueness results for convex Hamilton-Jacobi equations under p > 1 growth conditions on data*, arXiv:0810.1435.

[24] B. Derrida, M. R. Evans, V. Hakim, V. Pasquier. *Exact solution of a 1D asymmetric exclusion model using a matrix formulation*, J. Phys. A: Math. Gen. **26**, 1493–1517 (1993).

[25] *Constructive quantum field theory*, Proceedings of the 1973 Erice Summer School, ed. by G. Velo and A. Wightman, Lecture Notes in Physics **25**, Springer (1973).

[26] L. C. Evans. *Partial differential equations*, Graduate Studies in Mathematics **19**, AMS (1999).

[27] J. Feldman, J. Magnen, V. Rivasseau, R. Sénéor. *Construction and Borel summability of infrared $\Phi^4_4$ by a phase space expansion*, Comm. Math. Phys. **109**, 437–480 (1987).

[28] A. Friedman. *Partial differential equations of parabolic type*, Prentice-Hall(1964).

[29] J. Fröhlich, T. Spencer. *The Kosterlitz-Thouless transition in two-dimensional abelian spin systems and the Coulomb gas*, Comm. Math. Phys. **81**, 527–602 (1981).

[30] G. Gallavotti, K. Nicolò. *Renormalization theory in 4-dimensional scalar fields*, Comm. Math. Phys. **100**, 545–590 and **101**, 247–282 (1985).

[31] B. H. Gilding, M. Guedda, R. Kersner. *The Cauchy problem for $u_t = \Delta u + |\nabla u|^p$*, J. Math. Anal. Appl. **284**, 733-755 (2003).

[32] J. Glimm and A. Jaffe. *Quantum Physics, A Functional Point of View*, Springer (1987)
[33] N. Goldenfeld. *Lectures on phase transitions and the renormalization group*, Addison-Wesley (1992).

[34] M. Hairer. *Solving the KPZ equation*, arXiv:1109.6811.

[35] J. Z. Imbrie, T. Spencer. *Diffusion of directed polymer in a random environment*, J. Statist. Phys. 52, 609–626 (1988).

[36] K. Ito. *Existence of solutions to the Hamilton-Jacobi-Bellman equation under quadratic growth conditions*, J. Diff. Eq. 176, 11-28 (2001).

[37] K. Johansson. *Shape fluctuations and random matrices*, Comm. Math. Phys. 209, 437–476 (2000).

[38] S. Kaplan. *On the growth of solutions of quasi-linear parabolic equations*, Comm. Pure Appl. Math. 16, 305–330 (1963).

[39] I. Karatzas, S. Shreve. *Brownian motion and stochastic calculus*, Springer-Verlag (1991).

[40] M. Kardar, G. Parisi, Y.-C. Zhang. *Dynamic scaling of growing interfaces*, Phys. Rev. Lett. 56 (9), 889–892 (1986).

[41] C. Klüppelberg, T. Mikosch. *Large deviations of heavy-tailed random sums with applications to insurance and finance*, J. Appl. Probab. 34, 293–308 (1997).

[42] P. Laurençot, P. Souplet. *On the growth of mass for a viscous Hamilton-Jacobi equation*, Jour. Anal. Math. 89, 367–383 (2003).

[43] M. Laguës, A. Lesne. *Invariance d’´echelle. Des changements d’état à la turbulence*, Belin (2003).

[44] M. Le Bellac. *Quantum and statistical field theory*, Oxford Science Publications (1991).

[45] T. Lindvall. *Lectures on the coupling method*, John Wiley and sons (1992).

[46] Yu. V. Linnik. *Limit theorems allowing large deviations for the sums of independent variables I, II*, Theory Probab. Appl. 6, 145–161 and 377–391 (1961).

[47] J. Magnen, J. Unterberger. *From constructive theory to fractional stochastic calculus. (I) An introduction: rough path theory and perturbative heuristics*, Ann. Henri Poincaré 12, 1199–1226 (2011).

[48] J. Magnen, J. Unterberger. *From constructive theory to fractional stochastic calculus. (II) The rough path for \( \frac{1}{2} < \alpha < \frac{1}{4} \): constructive proof of convergence*, Ann. Henri Poincaré 13, 209–270 (2012).

[49] V. Mastropietro. *Non-perturbative renormalization*, World Scientific (2008).

[50] J. Magnen and D. Iagolnitzer. *Weakly self avoiding polymers in four dimensions*, Commun. Math. Phys., 162, 85–121 (1994)

[51] T. Mikosch, A. V. Nagaev. *Large deviations of heavy-tailed sums with applications in insurance*, Extremes 1, 81-110 (1998).
[52] A. V. Nagaev. *Integral limit theorems for large deviations when Cramér’s condition is not fulfilled I,II*, Theory Probab. Appl. 14, 51–64 and 193–208 (1969).

[53] A. V. Nagaev. *Large deviations of sums of independent random variables*, Ann. Prob. 7, 745–789 (1979).

[54] E. Nelson. *A quartic interaction in two dimensions*, in Mathematical Theory of Elementary Particles, R. Goodman and I. Segal eds, MIT Press (1966)

[55] B. Øksendal. *Stochastic differential equations. An introduction with applications*, Springer (2000).

[56] M. Prähofer, H. Spohn. *Scale invariance of the PNG droplet and the Airy process*, J. Stat. Phys. 108, 1071–1106 (2001).

[57] D. Revuz, M. Yor. *Continuous martingales and Brownian motion*, Springer (1999).

[58] V. Rivasseau. *From perturbative to constructive renormalization*, Princeton Series in Physics (1991).

[59] L. V. Rozovski. *Probabilities of large deviations on the whole axis*, Theory Probab. Appl. 38, 53–79 (1993).

[60] M. Salmhofer. *Renormalization: an introduction*, Springer Verlag (1999).

[61] T. Sasamoto, H. Spohn. *Exact height distributions for the KPZ equation with narrow wedge initial condition*, Nucl. Phys. B834, 523–542 (2010).

[62] M. E. Schonbek. *Uniform decay rates for parabolic conservation laws*, Nonlinear Anal., Theory, Methods and Appl. 10 (9), 943–956 (1986).

[63] I. Smears. *Hamilton-Jacobi-Bellman equations. Analysis and numerical analysis*, research report available on www.math.dur.ac.uk/Ug/projects/highlights/PR4/Smears_HJB_report.pdf

[64] E. M. Stein. *Harmonic analysis: real-variable methods, orthogonality and oscillatory integrals*, Princeton University Press (1993).

[65] C. A. Tracy, H. Widom. *Integral formulas for the asymmetric simple exclusion process*, Comm. Math. Phys. 279, 815–844 (2008).

[66] J. Unterberger. *Hölder-continuous paths by Fourier normal ordering*, Comm. Math. Phys. 298 (1), 1–36 (2010).

[67] J. Unterberger. *Mode d’emploi de la théorie constructive des champs bosoniques. Avec une application aux chemins rugueux*, to appear at: Confluentes Matematicae (in French with an English lexicon).

[68] F. Vignes-Tourneret. *Renormalisation des théories de champs non commutatives*, Thèse de doctorat de l’Université Paris 11, [arXiv:math-ph/0612014](http://arxiv.org/abs/math-ph/0612014)

[69] K. G. Wilson. *Renormalization Group and Critical Phenomena. I. Renormalization Group and the Kadanoff Scaling Picture*, Phys. Rev. B4, 3174–3184 (1971).
[70] K. G. Wilson and J. Kogut. *The renormalization group and the $\epsilon$ expansion*. Physics Reports 12, 75–200 (1974).