Mathematical
Physics


# Stochastic Burgers and KPZ Equations from Particle Systems*

**Lorenzo Bertini**[1], **Giambattista Giacomin**[2]

[1] Centre de Physique Theorique, Ecole Polytechnique, F-91128 Palaiseau Cedex, France.
E-mail: bertini@orphee.polytechnique.fr
[2] Institut für Angewandte Mathematik der Universität Zürich–Irchel, Winterthurerstr. 190, CH-8057 Zürich, Switzerland. E-mail: gbg@amath.unizh.ch

Received: 24 October 1995 / Accepted: 9 July 1996

**Abstract:** We consider two strictly related models: a solid on solid interface growth model and the weakly asymmetric exclusion process, both on the one dimensional lattice. It has been proven that, in the diffusive scaling limit, the density field of the weakly asymmetric exclusion process evolves according to the Burgers equation [8, 13, 18] and the fluctuation field converges to a generalized Ornstein-Uhlenbeck process [8, 10]. We analyze instead the density fluctuations beyond the hydrodynamical scale and prove that their limiting distribution solves the (non linear) Burgers equation with a random noise on the density current. For the solid on solid model, we prove that the fluctuation field of the interface profile, if suitably rescaled, converges to the Kardar–Parisi–Zhang equation. This provides a microscopic justification of the so called *kinetic roughening*, i.e. the non Gaussian fluctuations in some non-equilibrium processes. Our main tool is the Cole-Hopf transformation and its microscopic version. We also develop a mathematical theory for the macroscopic equations.

## 1. Introduction

The hydrodynamic behavior of physical systems is usually described by (non linear) PDE's. This description is in most of the cases approximate and, to model various neglected effects, a random forcing term can be added to the macroscopic equation. One is particularly interested in scale invariant forces, of which the space–time white noise is a typical example. This choice however introduces small scale singularities and poses the problem of the existence of the stochastic dynamics even when the deterministic equation is known to have good smoothing properties. Most rigorous results are restricted to one space dimension and several substantial problems appear in higher dimensions, see e.g. [1, 16]. On the other side, the question whether non linear stochastic PDE's, at

* This research has been partially supported by GNFM (L.B.) and by CNR and the Swiss National Science Foundation, Project 20-41'925.94 (G.G.).

least when they are well defined, are faithful descriptions of the evolution of suitable quantities defined on some particle dynamics is very natural. Moreover, by looking at some particle models, one may hope to understand how to construct infinite dimensional diffusions associated to some ill–posed stochastic PDE's. The aim of this paper is to derive a stochastic version of the viscous Burgers equation and the Kardar–Parisi–Zhang (KPZ) equation as scaling limits of microscopic particle models. We shall focus mainly on the latter topic referring to Appendix B for a precise statement of our results on the stochastic Burgers equation.

Consider a random *growth model*, as an example one may consider a Glauber dynamics for a ferromagnetic model at low enough temperature with an external positive magnetic field: if initially the system shows regions in which there is predominance of minus spin plus spins (*favoured phase*) and others in which there is predominance of minus spin (*unfavoured phase*), the favoured phase will expand, invading the regions occupied by the unfavoured phase. The separation layer between the two regions is then called interface; we refer to [20] for a more detailed introduction as well as an overview of different models. We stress that, unless the temperature is zero, the interface cannot be described microscopically as a separation line, because it has a thickness [20]. We are however going to study an *effective model* in which the interface is sharp and we will not face this problem.

The KPZ Eq.[17] has been proposed to describe the long scale behavior of the interface fluctuations. In this theory, through a *coarse groaning* procedure, the fluctuations of a $d$ dimensional interface (embedded in $\mathbf{R}^{d+1}$) are described, in a local coordinate system, by a single valued function $\mathbf{R}^d \ni r \mapsto h(r)$ which represents the interface height and evolves according to the (ill-posed) non linear stochastic PDE,

$$\partial_t h_t = \frac{1}{2}\Delta h_t - \frac{1}{2}(\nabla h_t)^2 + \dot{W}_t, \tag{1.1}$$

where $\nabla$ is the space gradient, $\Delta$ the Laplacian and $\dot{W}_t$ is the space-time white noise, i.e.

$$\mathbf{E}\left(\dot{W}_t(r)\,\dot{W}_t(r')\right) = \delta(t-t')\,\delta(r-r'). \tag{1.2}$$

A striking feature of growth processes is the roughness of the cluster surface, the so-called *kinetic roughening*. This is reflected by the presence of large and non Gaussian fluctuations, i.e. in the non-linearity of (1.1). The KPZ Eq.(1.1) is heuristically motivated within the *renormalization group* ideas: a general local dependence on the gradient $\nabla h$ is assumed, but the only *non irrelevant* term is the second order one. Accordingly, it is believed to be *universal*, i.e. independent, within a *certain* class, of the particular microscopic model.

We consider here only the one dimensional case, $d = 1$. We note that even in this case the mathematical interpretation of (1.1) is not obvious. This is best seen by trying to solve (1.1) as a perturbation of the associated Ornstein–Uhlenbeck process, defined by the same equation with the non linear term missing. The typical realizations of that process are in fact continuous in $r$, but not differentiable: the interpretation of the non linearity becomes then a non trivial point. Through a limiting procedure and a Wick renormalization of the non-linearity, see [1, 16] for the somehow analogous problem of the stochastic quantization of $P(\phi)_2$, we shall characterize uniquely a process associated to the Eq. (1.1). On the other hand we will not prove that it is the solution of a limiting equation. In fact we are not even able to write a meaningful version of (1.1).

We are going to focus on the derivation of the *stochastically perturbed* hydrodynamic Eq.(1.1), as a scaling limit of a microscopic interface model. Therefore proving, in a very



particular case, the universality hypothesis of the KPZ model. The conventional picture in the hydrodynamic limit [27] presents the deterministic equation as a law of large numbers for the density fields and the fluctuations around the hydrodynamic behavior as a central limit theorem. The fluctuations are thereby described by a Gaussian process (generalized Ornstein–Uhlenbeck process). We stress that such a process is the solution of a *linear* stochastic PDE, precisely the linearization of the hydrodynamic PDE, to which a random force term is added. To see the randomness and the non-linearity in the limiting process, as in (1.1), one is thus forced to analyze the fluctuations fields beyond the hydrodynamic scale. This possibility has been successfully pursued for one dimensional (non equilibrium) critical fluctuations in a local mean field theory [4, 12] and for other one dimensional long range models [25].

More precisely, in this paper we prove that Eq. (1.1) can be derived as the scaling limit of the fluctuation field for a microscopic growth model, the *weakly asymmetric single step solid on solid process* (SOS), which can be roughly described as follows. It is a model in which the microscopic interface is given as a single valued function $\zeta : \mathbf{Z} \mapsto \mathbf{Z}$. It furthermore obeys the *single step* constraint $|\zeta(x+1) - \zeta(x)| = 1$. The random dynamics is then specified by the following growth rules: local minima become maxima with rate $1/2 + \sqrt{\varepsilon}$ whereas local maxima become minima with rate $1/2$. This occurs independently at each site, no other transition is allowed. This process models a local evaporation/deposition and establishes a growth direction. We stress that, for future convenience, we have denoted by $\sqrt{\varepsilon}$ the strength of the asymmetry. We finally mention that this model can be obtained from the Metropolis dynamics for a two-dimensional Ising model in the limit in which the external magnetic field and the temperature converge to zero but their ratio is fixed and given by $\sqrt{\varepsilon}$, see [20].

To explain the scaling we shall consider, let us first recall the *hydrodynamic limit* and the Gaussian *hydrodynamic fluctuations* for this model. These results follow from the fact that SOS can be easily represented, as we shall see, in terms of the *weakly asymmetric exclusion process* (WASEP) for which the analogous results are proven in [8, 10, 13, 18].

Let us introduce a macroscopic coordinate $q = \sqrt{\varepsilon}\, x$ such that the strength of the asymmetry coincides with the scaling parameter. In this coordinate system the interface position is given by

$$m_s^\varepsilon(q) := \sqrt{\varepsilon}\, \zeta_{\varepsilon^{-1}s}\left(\varepsilon^{-1/2} q\right), \tag{1.3}$$

where $\zeta(x)$ is defined by linear interpolation for non integer $x$. Above we scaled the microscopic coordinate $x$, the interface height $\zeta$ and the microscopic time $t$ diffusively, namely $x \sim \zeta \sim \varepsilon^{-1/2}$, $t \sim \varepsilon^{-1}$. Assuming that $m_0^\varepsilon$ converges, as $\varepsilon \to 0$, to a differentiable function $m_0$, then (see [8, 13, 18] for a precise statement) $m_s^\varepsilon$ converges in probability to a function $m_s = m_s(q)$ which solves

$$\partial_s m_s = \frac{1}{2}\Delta m_s + \frac{1}{2}\left[1 - (\nabla m_s)^2\right] \tag{1.4}$$

in which the Laplacian is due to the symmetric part of the evolution and the other term to the asymmetric drift, which is of order one in this time scale since the force is $\sim \varepsilon^{1/2}$.

In order to understand the structure of the (random) forcing term which describes the corrections to (1.4) for non zero $\varepsilon$, one introduces the interface fluctuations as

$$Y_s^\varepsilon(q) = \varepsilon^{-1/4}\left[\varepsilon^{1/2}\zeta_{\varepsilon^{-1}s}(\varepsilon^{-1/2}q) - m_s(q)\right] \tag{1.5}$$

in which $\varepsilon^{-1/4}$ is the usual (CLT) normalization. In [8, 10] the central limit theorem associated to the law of large numbers (1.4) is proven. It states that $Y^\varepsilon$ converges in distribution to a process $Y_s = Y_s(q)$ which solves the *linear* stochastic PDE

$$\partial_s Y_s = \frac{1}{2}\,\Delta Y_s - \nabla m_s\,\nabla Y_s + \sqrt{1-(\nabla m_s)^2}\,\dot{W}_s, \tag{1.6}$$

in which the deterministic function $m_s$ is given and it is precisely the solution of (1.4).

As we mentioned above we shall consider the fluctuation field for SOS beyond the first non trivial hydrodynamic scaling, which is here the diffusive scaling $q = \varepsilon^{1/2}\,x$, $s = \varepsilon\,t$. To this end let us note that, according to (1.3), (1.4) the evolution of an homogenous flat interface $\bar\zeta_0 = $ const. is given by $\bar\zeta_t = \bar\zeta_0 + (1/2)\,\varepsilon^{1/2}t$. We shall consider a $\varepsilon^{1/2}$ perturbation which varies on the (longer) scale $r = \varepsilon\,x$. Set $Z_0^\varepsilon(r) := \varepsilon^{1/2}\left(\zeta_0(\varepsilon^{-1}r) - \bar\zeta_0\right)$; we therefore assume there exists a function (it may be random) $h_0 \in C(\mathbf{R})$ with at most linear growth for $r \to \pm\infty$ such that

$$Z_0^\varepsilon(r) = h_0(r) + o(1), \tag{1.7}$$

where $o(1)$ is infinitesimal as $\varepsilon \to 0$.

We then look at the interface height on the diffusive scaling $r = \varepsilon\,x$, $\tau = \varepsilon^2\,t$. Recall in fact that the asymmetry is $\varepsilon^{1/2}$, so that this new scaling is different from the previous one. Set $Z_\tau^\varepsilon(r) := \varepsilon^{1/2}\left(\zeta_{\varepsilon^{-2}\tau}(\varepsilon^{-1}r) - \bar\zeta_{\varepsilon^{-2}\tau}\right)$; we shall prove that

$$Z_\tau^\varepsilon(r) = h_\tau(r) + o(1), \tag{1.8}$$

where $h_\tau = h_\tau(r)$ is the solution of the KPZ equation (1.1) with initial condition $h_0$.

The KPZ equation is then to be interpreted as describing the long scale behavior of small fluctuations around a flat interface. We note in fact that in [17] is introduced a *small gradient* assumption. In our setting it is precisely formulated as the condition (1.7) on the initial profile. We also note that such a condition is quite natural in the analysis beyond the hydrodynamic scale. It is what has been called, in the context of the derivation of the Navier-Stokes equation as next order correction to the Euler Eq.[11], the *incompressible limit* condition. We finally mention that the (informal) statement (1.8) is not completely correct: in our analysis we shall find a contribution from the fluctuations to the deterministic growth of the homogeneous profile. In particular the function $\bar\zeta_t$ has to include a lower order correction, i.e. $\bar\zeta_t = \bar\zeta_0 + 1/2\,\varepsilon^{1/2}\,t - 1/24\,\varepsilon^{3/2}\,t$.

The real issue behind this result is whether a profile satisfying (1.7) is stable under the microscopic evolution. In other words we are asking what happens to a perturbation of order $\varepsilon^{1/2}$ (varying on the space scale $\varepsilon^{-1}$) after a time $\varepsilon^{-2}$ and, to derive the non linear Eq.(1.1), some *propagation of chaos* type result is needed. We note that for the weakly asymmetric exclusion process the *propagation of chaos* has been proven in a very strong form up to the hydrodynamic scale [8, 13], however those results do not hold in our regime; for instance we are exactly at the time scale in which the analysis in [9] breaks down.

Our results are obtained by using a non linear map, the Cole-Hopf transformation, which reduces (1.1) to a linear equation with multiplicative noise. For the microscopic process a similar transformation has been introduced in [13] and the transformed process solves a semimartingale equation with a linear drift. This technique is peculiar of the model introduced and the type of results we obtain do not seem to be, at the moment, in the domain of application of more general tools.



*Outline of the paper.* In the next section we explain the rigorous meaning of the KPZ equation, define the microscopic processes and state our main result on the convergence of the fluctuation fields to the solution of the KPZ equation. In Sect. 3 we introduce the Cole-Hopf transformation which reduces (1.1) to a linear equation with a multiplicative noise and the corresponding transformation for the microscopic process. At the end of that section it is also shown how the results on the *transformed process* imply the main result (convergence to the KPZ equation). The proof of the scaling limit for the transformed process is carried out in Sect. 4. Finally in Sect. 5 we prove some properties of the macroscopic Eqs. . The (equivalent) formulation of our results for the stochastic Burgers equation and the weakly asymmetric exclusion process can be found in Appendix B.

## 2. Notation and Main Results

Throughout all the paper the space of continuous functions on the real line $C(\mathbf{R})$ is equipped with the topology of uniform convergence over compact sets. As test functions we use the space $\mathcal{D}(\mathbf{R})$, i.e. the function in $C_0^\infty(\mathbf{R})$ with the inductive limit topology; its strong topological dual $\mathcal{D}'(\mathbf{R})$ is the space of distributions. For $\mathcal{H}$ a topological space and $T > 0$, $D([0,T];\mathcal{H})$ is the Skorohod space of $\mathcal{H}$-valued functions with *cadlag* trajectories [23]. Since the Skorohod topology relativized to $C([0,T];\mathcal{H})$ coincides with the uniform topology [5], we may (and sometimes do) consider random functions in $C([0,T];\mathcal{H})$ as elements of $D([0,T];\mathcal{H})$.

The processes we are dealing with are to be considered over an arbitrary but fixed macroscopic time interval $[0,T]$. We will be mainly concerned with the weak convergence of sequences of stochastic processes; following the usual notation we denote this convergence by the double arrow $\Rightarrow$ stressing this notion of convergence depends upon the topology used on the path space.

*2.1 The KPZ equation.* Before stating our main result on the convergence of the fluctuation field of the microscopic interface model, we explain the rigorous meaning of the KPZ Eq.(1.1).

Let $W_t$, $t \in [0,T]$, be the cylindrical Wiener process on $L^2(\mathbf{R}, dr)$. It is canonically realized as a distribution valued continuous process, i.e. the probability space is given by $\left(C([0,T];\mathcal{D}'(\mathbf{R})), \mathcal{A}, \mathcal{P}\right)$, here $\mathcal{A}$ is the $\sigma$-algebra generated by the cylindrical sets and $\mathcal{P}$ is the Gaussian measure with covariance

$$\int d\mathcal{P}\; W_t(\varphi_1)\,W_s(\varphi_2) = t \wedge s\,(\varphi_1, \varphi_2), \tag{2.1}$$

where $\varphi_1, \varphi_2 \in \mathcal{D}(\mathbf{R})$, are test functions, $a \wedge b := \min\{a, b\}$ and $(\cdot,\cdot)$ is the inner product in $L^2(\mathbf{R}, dr)$. We denote by $\mathcal{A}_t^0$ the natural filtration of $W_t$, i.e. the minimal $\sigma$-algebra such that $s \mapsto W_s$ is $\mathcal{A}_t^0$ measurable for all $s \in [0,t]$.

We shall characterize the solution of the KPZ equation through a limiting procedure. Accordingly we introduce a mollified version of the cylindrical Wiener process which will define a family of approximating problems. Let $J \in C_0^\infty(\mathbf{R})$ be an even positive function such that $\int dr\; J(r) = 1$. Introduce, for $\kappa > 0$, the mollifier $\delta_r^\kappa(r') := \kappa\,J(\kappa(r - r'))$ and define $W_t^\kappa(r) := W_t(\delta_r^\kappa)$; its covariance is

$$\mathbf{E}\left(W_t^\kappa(r) W_s^\kappa(r')\right) = t \wedge s\; C_\kappa(r - r') \quad , \quad C_\kappa(r) := \int dr'\; \delta_r^\kappa(r')\,\delta_0^\kappa(r'). \tag{2.2}$$

We then write a mollified KPZ equation against test functions as

$$h_t^\kappa(\varphi) = h_0(\varphi) + \int_0^t ds\, \frac{1}{2}\left\{h_s^\kappa(\varphi'') - \left[(\nabla h_s^\kappa)^2 - C_\kappa(0)\right](\varphi)\right\} + W_t^\kappa(\varphi), \qquad (2.3)$$

where $\varphi \in \mathcal{D}(\mathbf{R})$. It is formally obtained from (1.1) integrating by parts. We have also added the term $C_\kappa(0) \sim \kappa^{-1}$, which corresponds to take the *Wick product* of the non-linearity $(\nabla h^\kappa)^2$.

We assume the initial condition $h_0 = h_0(r)$ to be a random function in $C(\mathbf{R})$ which is independent of $\mathcal{P}$ and satisfies the following growth condition: for each $p > 0$ there exists $a = a_p$ such that

$$\sup_{r\in\mathbf{R}} e^{-a|r|}\, \mathbf{E}\left(e^{-p\, h_0(r)}\right) < \infty. \qquad (2.4)$$

We have thus assumed the initial datum to be a continuous trajectory $h_0 = h_0(r)$ which has for every $r \in \mathbf{R}$ an exponential moment that grows at most exponentially in $r$. One can obviously take $h_0$ to be a deterministic function with at most linear growth. We also introduce the filtration $\mathcal{A}_t := \sigma\{h_0\} \vee \mathcal{A}_t^0 = \sigma\{h_0, W_s\,;\, s \in [0,t]\}$.

The limit $\kappa \to \infty$ in (2.3) can be taken according to the following strategy. When the cutoff $\kappa$ is finite $W_t^\kappa(r)$ is smooth so that (2.3) makes sense in the space of differentiable functions; we thus obtain a processes $h^\kappa$ in $C([0,T]; C(\mathbf{R})) \cap C((0,T]; C^1(\mathbf{R}))$. Since a limiting process will not be differentiable in space we have no hope to get a convergent sequence in this space; however $\{h^\kappa\}_{\kappa>0}$ does form a weak convergent family as $\kappa \to \infty$ in the topology of $C([0,T]; C(\mathbf{R}))$. We have in fact the following result.

**Theorem 2.1.** *Let $h_0$ be a random function in $C(\mathbf{R})$ which satisfies the assumption (2.4), then:*

*(i) For all $\kappa > 0$, $T > 0$, there exists a process $h_t^\kappa = h_t^\kappa(r)$ in $C([0,T]; C(\mathbf{R})) \cap C((0,T]; C^1(\mathbf{R}))$, adapted to the filtration $\mathcal{A}_t$, which solves almost surely (2.3) for all $\varphi \in \mathcal{D}(\mathbf{R})$ and every $t \in [0,T]$. It is furthermore unique in the class of adapted processes $X_t = X_t(r)$ satisfying also the growth condition*

$$\sup_{t\in[0,T]} \sup_{r\in\mathbf{R}} e^{-a|r|} \int d\mathcal{P}\, e^{-2X_t(r)} < \infty \qquad (2.5)$$

*for some $a > 0$.*

*(ii) Consider $h_t^\kappa$, $t \in [0,T]$ as a random element in $C([0,T]; C(\mathbf{R}))$; then the family $\{h^\kappa\}$ is weakly convergent as $\kappa \to \infty$. The limiting process is denoted by $h$, i.e. $h^\kappa \Rightarrow h$ in $C([0,T]; C(\mathbf{R}))$.*

Although this result characterizes uniquely the solution of the KPZ equation through the approximating problems (2.3) it is not completely satisfactory since it avoids the issue of showing that $h$ satisfies a limiting equation. In particular we have not defined the Wick product in (2.3) for the limiting process; to our knowledge this problem has been solved only for infinite dimensional diffusions which are either Gaussian or constructed as perturbations of Gaussian measures [1, 16].

*2.2 Solid on solid model of growing surfaces.* We next define precisely the microscopic interface model we shall consider. The *weakly asymmetric single step solid on solid process* (SOS) is defined as follows.

On the state space

$$\widehat{\Omega} := \{\zeta \in \mathbf{Z}^{\mathbf{Z}} \,:\, \forall x \in \mathbf{Z}\ \ |\zeta(x+1) - \zeta(x)| = 1\} \qquad (2.6)$$

introduce the jump Markov process generated by

$$\widehat{L}_\varepsilon f(\zeta) := \sum_x \{c^+(x,\zeta)[f(\zeta + 2\delta_x) - f(\zeta)] + c^-(x,\zeta)[f(\zeta - 2\delta_x) - f(\zeta)]\}, \qquad (2.7)$$

where $\varepsilon > 0$, $f$ is a cylindrical function on $\widehat{\Omega}$, $\delta_x \in \mathbf{Z}^{\mathbf{Z}}$ is defined by $\delta_x(y) := \delta(x-y)$ (the Kronecker symbol) and

$$c^+(x,\zeta) := \begin{cases} 1/2 + \varepsilon^{1/2} & \text{if } \zeta(x+1) - 2\zeta(x) + \zeta(x-1) = 2, \\ 0 & \text{otherwise} \end{cases}$$

$$c^-(x,\zeta) := \begin{cases} 1/2 & \text{if } \zeta(x+1) - 2\zeta(x) + \zeta(x-1) = -2, \\ 0 & \text{otherwise.} \end{cases} \qquad (2.8)$$

The structure of the generator (2.7) preserves the single step constraint, $\zeta \in \widehat{\Omega}$ in the dynamic rules. The allowed transition are indeed either when local minima of $\zeta(x)$ are raised to local maxima, $\zeta(x) \mapsto \zeta(x)+2$, with rate $1/2+\sqrt{\varepsilon}$ or local maxima are lowered to local minima, $\zeta(x) \mapsto \zeta(x) - 2$, with rate $1/2$. In Sect. 2 we will introduce a mapping from a certain exclusion process which will, in particular, establish the existence of the dynamics we just introduced.

We shall consider $\zeta$ as a continuous function by linear interpolation on its value on the lattice $\mathbf{Z}$, i.e. for $r \in \mathbf{R}$ we define $\zeta(r)$ by

$$\zeta(r) := \zeta([r]) + (r - [r])(\zeta([r]+1) - \zeta([r])), \qquad (2.9)$$

where $[\cdot]$ is the integer part. When viewed this way, the process $\zeta_t$, $t \in \mathbf{R}^+$ is then a random element in $D(\mathbf{R}^+; C(\mathbf{R}))$.

**Definition 2.2.** *The initial profile. The initial distribution for SOS is given by a family $\{\hat{\mu}_\varepsilon\}_{\varepsilon>0}$ of probabilities on $\widehat{\Omega}$ which satisfies the following conditions:*

*(i) Set $\zeta^\varepsilon(r) := \sqrt{\varepsilon}\zeta(\varepsilon^{-1}r)$. There exists a random function with trajectories $h_0 = h_0(r)$ in $C(\mathbf{R})$ such that*

$$\zeta^\varepsilon \Rightarrow h_0 \qquad (2.10)$$

*as $\varepsilon \to 0$ in the topology of $C(\mathbf{R})$.*

*(ii) For each $n \in \mathbf{N}$ there are $a = a_n$, $c = c_n > 0$ such that*

$$\sup_{x\in\mathbf{Z}} e^{-a\varepsilon|x|} \int d\hat{\mu}_\varepsilon \exp\{-n\sqrt{\varepsilon}\,\zeta(x)\} \le c \qquad (2.11)$$

*for all $\varepsilon > 0$;*

*(iii) For each $n \in \mathbf{N}$ there are $a' = a'_n$, $c' = c'_n > 0$ such that*

$$\int d\hat{\mu}_\varepsilon \left(\sqrt{\varepsilon}[\zeta(x) - \zeta(y)]\right)^{2n} \le c'\, e^{a'\varepsilon(|x|+|y|)} \left(\varepsilon|x-y|\right)^n \qquad (2.12)$$

*for every $x, y \in \mathbf{Z}$ and all $\varepsilon > 0$.*

Condition (2.11) entails that the random function $h_0$ defined by (2.10) satisfies the assumption (2.4); the inequality (2.12) implies also $h_0$ is a.s. Hölder continuous with exponent less than 1/2. We remark that the moment conditions (2.11) and (2.12) can be somewhat weakened. A careful reading of Sect. 4 shows we really need them only for $n \leq 10$.

The conditions in Definition 2.2 are satisfied if the increments $\zeta(x+1) - \zeta(x) \in \{1, -1\}$ are independent with marginals

$$\hat{\mu}_\varepsilon\left(\zeta(x+1) - \zeta(x)\right) = \varepsilon^{-\frac{1}{2}}\left[m(\varepsilon x) - m(\varepsilon(x-1))\right] \tag{2.13}$$

in which $m = m(r)$ is a $\alpha$–Hölder continuous function ($\alpha \geq 1/2$) on $\mathbf{R}$ which satisfies the following condition: there is $a > 0$ such that for every $r \in \mathbf{R}$, $|m(r)| \leq a(1 + |r|)$. The initial datum for the KPZ equation is then the random function $h_0$, taking values in $C(\mathbf{R})$, given by (in law) $h_0 = m + b$, where $m$ is the deterministic function in (2.13) and $b = b(r)$ is a bilateral Brownian motion on $\mathbf{R}$, i.e. $b_1(r) := b(r), r \geq 0$ and $b_2(r) := b(-r)$, $r \geq 0$ are independent Brownians on $\mathbf{R}$.

Admissible initial data include however also deterministic profiles, i.e. the case in which $\hat{\mu}_\varepsilon$ is concentrated on a single configuration. For example take $\hat{\mu}_\varepsilon$ independent of $\varepsilon$ and concentrated on the configuration $\zeta_0$ such that $\zeta_0(x+1) - \zeta_0(x) = 1$ if $x$ is even, $\zeta_0(x+1) - \zeta_0(x) = -1$ if $x$ is odd. This gives $h_0 = 0$.

*2.3 The scaling limit.* We may now state our main result on the convergence of the fluctuation of SOS to the KPZ equation.

**Theorem 2.3.** *For $t \in [0, T]$, $r \in \mathbf{R}$, let*

$$Z_t^\varepsilon(r) := \sqrt{\varepsilon}\,\left(\zeta_{\varepsilon^{-2}t}(\varepsilon^{-1}r) - v_\varepsilon t\right), \tag{2.14}$$

*where $v_\varepsilon := \frac{1}{2}\varepsilon^{-3/2} - \frac{1}{24}\varepsilon^{-1/2}$ and regard $Z^\varepsilon$ as a random element in $D([0,T]; C(\mathbf{R}))$.*

*The family $\{Z^\varepsilon\}_{\varepsilon>0}$ is weakly convergent as $\varepsilon \to 0$; furthermore the weak limit is concentrated on $C([0,T]; C(\mathbf{R}))$ and coincides with the process $h$ defined in Theorem 2.1, i.e.*

$$Z^\varepsilon \Rightarrow h \tag{2.15}$$

*in the topology of $D([0,T]; C(\mathbf{R}))$.*

We remark that the Wick counterterm, which has been introduced *ad hoc* in (2.3) arises naturally in the scaling limit. This possibility was suggested, for the two dimensional Landau-Ginzburg equation with noise, in [15]; see also [6] for a related discussion.

We also note that Theorem 2.3 gives a rather strong convergence as a process in $D([0,T]; C(\mathbf{R}))$ and not in a distribution space. This is due to the fact that the microscopic process $\zeta_t(x)$ can be written as a function of the empirical average for a particle system which will be introduced below. More directly it follows from the fact that an elementary step of the dynamics changes $Z^\varepsilon$ by a factor of order $\varepsilon^{1/2}$ so that no space averaging is needed.

In the physical literature, see e.g. [20], it is well–known that the one–dimensional KPZ process has an invariant (but not reversible) state in which $\nabla h$ is distributed according to the white noise measure on $\mathcal{D}'(\mathbf{R})$, which is also the invariant (and reversible) state for the same equation without the non linear term. Once the KPZ process is rigorously constructed, the proof of this fact is a fairly simple computation; it is however remarkable that in our setting it is a straightforward consequence of the invariance of the Bernoulli measure on $\mathbf{Z}$ under the asymmetric exclusion process. The precise statement and proof are given, in Appendix B, as Proposition B.2.

*2.4 The weakly asymmetric simple exclusion process.* We conclude this section by explaining the relationship between SOS and the *weakly asymmetric simple exclusion process*, which will be the basic object in the proofs. Via the results in [22] on (infinite volume) exclusion processes, this mapping will also build a version of the SOS process and hence it will establish its existence.

For notation convenience, i.e. to have centered variables, we describe the particle model in terms of spin variables. The state space of the microscopic process is thus $\Omega := \{-1, 1\}^{\mathbf{Z}}$, its elements (*spin configurations*) are denoted by $\sigma = \{\sigma(x),\ x \in \mathbf{Z}\}$, where $\sigma(x) = +1$ (resp. $-1$) is interpreted as the site in $x$ being occupied (resp. empty).

The *weakly asymmetric simple exclusion process* (WASEP) is the process generated by

$$L_\varepsilon := \frac{1}{2}L^+ + \left(\frac{1}{2} + \sqrt{\varepsilon}\right) L^-, \tag{2.16}$$

where $L^\pm$ are the generators of the totally asymmetric exclusion processes, defined by

$$L^\pm f(\sigma) := \sum_x \frac{1+\sigma(x)}{2}\,\frac{1-\sigma(x\pm 1)}{2}\,\left[f(\sigma^{x,x\pm 1}) - f(\sigma)\right] \tag{2.17}$$

in which $f$ is a cylindrical function over $\Omega$ and given $x, y \in \mathbf{Z}$

$$\sigma^{x,y}(z) \;:=\; \begin{cases} \sigma(x) & \text{if} \quad z = y \\ \sigma(y) & \text{if} \quad z = x \\ \sigma(z) & \text{otherwise.} \end{cases} \tag{2.18}$$

We stress in (2.16) we adopted the (unusual) convention of an asymmetry $\sqrt{\varepsilon}$ toward the left. The details on the construction of the process can be found in Liggett [22]. We consider the WASEP $\sigma_t$, $t \geq 0$ canonically realized on the Skorohod space $D(\mathbf{R}^+; \Omega)$ and denote by $\mathbf{P}^\varepsilon_{\mu_\varepsilon}$ its law when the initial distribution is $\mu_\varepsilon$, probability on $\Omega$. The expectation with respect to $\mathbf{P}^\varepsilon_{\mu_\varepsilon}$ is denoted by $\mathbf{E}^\varepsilon_{\mu_\varepsilon}$.

Let $x_t^0$ be the position of the tagged particle for WASEP. We recall that the tagged particle is the particle that at time zero was closest to the origin on the positive half-axis, i.e. given $\sigma_0$, $x_0^0$ is defined by $x_0^0 := \min\{x \in \mathbf{Z} : x \geq 0,\ \sigma(x) = +1\}$ and $x_t^0$ is its position at time $t$ under the WASEP dynamics.

The generator of the joint Markov process $(\sigma_t, x_t^0)$ on the state space $\{(\sigma, x) \in \Omega \times \mathbf{Z} : \sigma(x) = 1\}$ is

$$H_\varepsilon := \frac{1}{2}\left(H^+ + (1 + 2\sqrt{\varepsilon})H^-\right), \tag{2.19}$$

where

$$\begin{aligned} H^\pm f(\sigma, x^0) \;:=\; & \frac{1}{4}\sum_{x \neq x^0}(1 + \sigma(x))(1 - \sigma(x \pm 1))\left[f(\sigma^{x,x\pm 1}, x^0) - f(\sigma, x^0)\right] \\ & + \frac{1}{2}(1 - \sigma(x^0 \pm 1))\left[f(\sigma^{x,x\pm 1}, x^0 \pm 1) - f(\sigma, x^0)\right] \end{aligned} \tag{2.20}$$

and $f$ is a cylindrical function.

It is then a simple check to verify that a version of SOS is given by

$$\zeta_t(x) = \begin{cases} \sum_{x_t^0 < y \le x} \sigma_t(y) - x_t^0 & \text{if} \quad x > x_t^0 \\ -\sum_{x < y \le x_t^0} \sigma_t(y) - x_t^0 & \text{if} \quad x < x_t^0 \\ -x_t^0 & \text{if} \quad x = x_t^0, \end{cases} \tag{2.21}$$

where $\sigma_t$ is WASEP. We shall work with this version and assume the initial distribution $\mu_\varepsilon$ be such that the hypotheses in Definition 2.2 hold.

In the reverse direction, given $\zeta_t$ SOS , to obtain a version on WASEP it is enough to look at the increments, i.e.

$$\sigma_t(x) = \zeta_t(x) - \zeta_t(x-1). \tag{2.22}$$

## 3. Non-Linear Transformation Method

In this section we introduce the main tool in proving the stated results. We shall use a non linear map, the so called Cole-Hopf transformation, which reduces the KPZ equation to a linear equation with multiplicative noise.

The scaling limit of SOS is analyzed through a microscopic analogue of this transformation, introduced by Gärtner in [13]. In this case we do not get a linear equation, but the process obtained is easier to handle and we shall prove that it converges to a linear equation in the limit $\varepsilon \to 0$; taking the inverse map, this will imply the convergence to the KPZ equation.

*3.1 Cole-Hopf transformation.* The heuristic observation is the following. Let $h_t$ be a solution of the KPZ Eq.(1.1) and introduce the process $\theta_t := \exp\{-h_t\}$, it then solves the *linear* equation

$$d\theta_t = \frac{1}{2}\Delta\theta_t dt - \theta_t dW_t \tag{3.1}$$

often called the *stochastic heat equation*.

The rigorous analysis starts by giving a meaning to the above equation when the stochastic differential is interpreted in the Ito sense. We assume the initial datum $\theta_0 = \theta_0(r)$ to be a random function in $C(\mathbf{R})$ which is independent on $\mathcal{P}$ and satisfies the following growth condition: for each $p > 0$ there exists $a = a_p$ such that

$$\sup_{r\in\mathbf{R}} e^{-a|r|}\, \mathbf{E}\left(|\theta_0(r)|^p\right) < \infty. \tag{3.2}$$

In the application to the KPZ equation we shall consider $\theta_0(r) = \exp\{-h_0(r)\}$, where $h_0$ satisfies the hypothesis (2.4).

Introduce the heat kernel

$$G_t(x) := \frac{1}{\sqrt{2\pi t}} \exp\left\{-\frac{x^2}{2t}\right\} \tag{3.3}$$

and formulate the stochastic heat equation in mild form as

$$\theta_t = G_t * \theta_0 - \int_0^t G_{t-s} * \theta_s dW_s, \tag{3.4}$$

where $*$ denotes convolution in space and



$$\int_0^t G_{t-s} * \theta_s dW_s\,(r) := \int_0^t \left(G_{t-s}(r-\cdot)\theta_s\,,\, dW_s\right) \tag{3.5}$$

is the Ito integral with respect to the cylindrical Wiener process.

It is convenient to formulate this problem also for the mollified Wiener process, i.e.

$$\theta_t^\kappa = G_t * \theta_0 - \int_0^t G_{t-s} * \theta_s^\kappa dW_s^\kappa. \tag{3.6}$$

We finally introduce

$$\mathcal{C}_+(\mathbf{R}) := \{f \in C(\mathbf{R}))\, :\, \forall r \in \mathbf{R}\;\; f(r) > 0\}. \tag{3.7}$$

**Theorem 3.1.** *Under the assumption (3.2) on $\theta_0$:*

*(i) For all $\kappa > 0$, $T > 0$, there exists a process $\theta_t^\kappa = \theta_t^\kappa(r)$ in $C([0,T]; C(\mathbf{R}))$, adapted to the filtration $\mathcal{A}_t$, which solves almost surely (3.6) for every $t \in [0,T]$. It is furthermore unique in the class of adapted processes $X_t = X_t(r)$ satisfying also the growth condition*

$$\sup_{t\in[0,T]} \sup_{r\in\mathbf{R}} e^{-a|r|} \int d\mathcal{P}\, X_t(r)^2 < \infty \tag{3.8}$$

*for some $a > 0$.*

*(ii) For all $p \ge 1$, $\theta_t^\kappa(r) \to \theta_t(r)$ in $L^p(\mathcal{P})$ and $\mathcal{P}$ a.s.; the convergence is uniform for $(t,r)$ in compact subsets of $[0,T] \times \mathbf{R}$. The limiting process is the unique solution of (3.4) in the class of $\mathcal{A}_t$ adapted processes satisfying (3.8).*

*(iii) If $\theta_0 \in \mathcal{C}_+(\mathbf{R})$ a.s. then $\theta \in C([0,T]; \mathcal{C}_+(\mathbf{R}))$ a.s.*

*Remark.* In Sect. 4.3 we shall introduce an equivalent formulation of the stochastic heat equation in the form of a martingale problem and we will use this new formulation to identify the limit.

The existence and uniqueness result (when $\mathbf{R}$ is replaced by a bounded interval) for the stochastic heat Eq. (3.1) goes back to Walsh [28]. Referring to the mollified process $W_t^\kappa$, the statement *(i)* is trivial. Point *(ii)* is proven in [2] where a Feynman-Kac formula for $\theta_t$ is also given. In that paper the initial condition is assumed to be bounded, but the extension to the exponential growth as in (3.2) is straightforward. The (for us fundamental) positivity property *(iii)* is due to Müller [24].

We are now ready to give an equivalent definition, at this point the third (and last), for the solution of the KPZ equation.

**Theorem 3.2.** *Let $\Psi\, :\, C([0,T]; \mathcal{C}_+(\mathbf{R})) \mapsto C([0,T]; C(\mathbf{R}))$ be defined by $f_t(r) \overset{\Psi}{\longmapsto} -\log f_t(r)$ and $\theta$ be the solution of the stochastic heat Eq.(3.4) with initial condition $\theta_0 = \exp\{-h_0\}$. Theorem 3.1 entails that $\Psi(\theta)$ is a.s. well defined and in $C([0,T]; C(\mathbf{R}))$. Then $\Psi(\theta)$ coincides (in law) with the process $h$ defined in Theorem 2.1.*

*3.2 The Gärtner transformation.* Here we introduce the microscopic analog of the Cole-Hopf transformation. A similar transformation has been used in [9, 10, 13]. In our case it maps the problem of the convergence of the fluctuation field of SOS to the KPZ equation into the problem of the convergence of the transformed process to the stochastic heat equation above discussed.

The transformed process $\xi_t = \xi_t(r)$, $r \in \mathbf{R}$ is defined as

$$\xi_t(r) := \exp\{-\gamma_\varepsilon \zeta_t(r) + \lambda_\varepsilon t\}, \tag{3.9}$$

where

$$\gamma_\varepsilon := \frac{1}{2}\log(1+2\varepsilon^{\frac{1}{2}}) \quad , \quad \lambda_\varepsilon := 1+\varepsilon^{\frac{1}{2}} - \sqrt{1+2\varepsilon^{\frac{1}{2}}} \tag{3.10}$$

and $\zeta_t(r)$, by linear interpolation on the values $\zeta_t(x)$, is SOS as constructed in (2.21).

As $\gamma_\varepsilon = \sqrt{\varepsilon} + O(\varepsilon)$ the first term at the exponent in (3.9) is the fluctuation field so that (3.9) is the microscopic analogue of the Cole-Hopf transformation. The other term $\lambda_\varepsilon t$ has been added in order to obtain the convergence of the process $\xi_t$ and takes into account the homogeneous deterministic drift of the interface. The inverse map to SOS is given by

$$\zeta_t(r) = -\frac{1}{\gamma_\varepsilon}\log \xi_t(r) + \frac{\lambda_\varepsilon}{\gamma_\varepsilon} t \tag{3.11}$$

and, for $x \in \mathbf{Z}$, to WASEP by

$$\sigma_t(x) = -\frac{1}{\gamma_\varepsilon}\Big[\log \xi_t(x) - \log \xi_t(x-1)\Big]. \tag{3.12}$$

For $f : \mathbf{Z} \to \mathbf{R}$, let us define the discrete gradients as

$$\nabla^{\pm} f(x) := \pm[f(x\pm 1) - f(x)]$$

and the discrete Laplacian as

$$\Delta f(x) := \nabla^+ \nabla^- f(x) \equiv f(x+1) + f(x-1) - 2f(x).$$

Let $\mathcal{F}_t$ be the natural filtration of $\sigma_t$, i.e. the $\sigma$-algebra generated by $\sigma_s$, $s \in [0,t]$. A direct computation [13] shows that $\xi_t(x)$, $x \in \mathbf{Z}$, satisfies the semimartingale equation

$$d\xi_t(x) = \frac{1}{2} e^{\gamma_\varepsilon} \Delta \xi_t(x)\, dt + dM_t(x), \tag{3.13}$$

where the $\mathcal{F}_t$ martingales $M_t(x)$ have brackets

$$\begin{aligned} & d\langle M(x), M(y)\rangle_t \\ &= \frac{\varepsilon}{2}\delta(x-y)\xi_t(x)^2 \qquad (3.14)\\ &\quad \cdot \left[\left(1+e^{-2\gamma_\varepsilon}\right)(1-\sigma_t(x)\sigma_t(x+1)) + \left(1-e^{-2\gamma_\varepsilon}\right)(\sigma_t(x)-\sigma_t(x+1))\right] dt \\ &= \frac{1}{2}\delta(x-y) \qquad (3.15)\\ &\quad \cdot \left[(1-e^{\gamma_\varepsilon})^2\left(2\xi_t(x)^2 + \xi_t(x)\Delta\xi_t(x)\right) - \left(1+e^{2\gamma_\varepsilon}\right)\nabla^+\xi_t(x)\nabla^-\xi_t(x)\right] dt. \end{aligned}$$

In the scaling limit Eq. (3.13) converges to the stochastic heat quation. The precise statement is the following.

**Theorem 3.3.** *For $t \in [0,T]$, $r \in \mathbf{R}$, let $\xi^\varepsilon_t(r) := \xi_{\varepsilon^{-2}t}(\varepsilon^{-1}r)$ and regard it as a random element in $D([0,T];C(\mathbf{R}))$. The family $\{\xi^\varepsilon\}_{\varepsilon>0}$ is weakly convergent as $\varepsilon \to 0$; furthermore the weak limit is concentrated on $C([0,T];C(\mathbf{R}))$ and coincides with the solution of the stochastic heat equation, i.e.*

$$\xi^\varepsilon \Rightarrow \theta \tag{3.16}$$

*in $D([0,T];C(\mathbf{R}))$. Here $\theta$ is the solution of (3.4) with initial datum $\theta_0 = \exp\{-h_0\}$, $h_0$ as in Definition 2.2.*

*3.3 The proof of Theorem 2.3.* Here we show how Theorem 3.3, together with the properties of the limit point $\theta$, implies our main result.

*Proof of Theorem 2.3.* Bearing in mind (3.11), the process $Z^\varepsilon$, as defined in (2.14), can be rewritten in terms of the scaled transformed process $\xi^\varepsilon$ as

$$Z^\varepsilon_t(r) = -\frac{\sqrt{\varepsilon}}{\gamma_\varepsilon}\log\left(\xi^\varepsilon_t(r)\right) + \sqrt{\varepsilon}\left[\frac{\lambda_\varepsilon}{\gamma_\varepsilon}\varepsilon^{-2} - v_\varepsilon\right] t. \tag{3.17}$$

Since $\lim_{\varepsilon\to 0}\sqrt{\varepsilon}/\gamma_\varepsilon = 1$ and there is a constant $c$ such that $|(\lambda_\varepsilon/\gamma_\varepsilon)\varepsilon^{-2} - v_\varepsilon| \le c$ for all $\varepsilon \in (0,1)$, we are left with establishing the weak convergence of $-\log\left(\xi^\varepsilon_t(r)\right)$ to $h$. In order to do this, we extend the map $\Psi$, as defined in Theorem 3.2, to the measurable map $\widehat{\Psi} : D([0,T];C(\mathbf{R})) \mapsto D([0,T];C(\mathbf{R}))$ defined by

$$\widehat{\Psi}(f_t(r)) = \begin{cases} -\log f_t(r) & \text{if } f \in D([0,T];\mathcal{C}_+(\mathbf{R})) \\ 0 & \text{otherwise} \end{cases} \tag{3.18}$$

and clearly $\widehat{\Psi}\left(\xi^\varepsilon_t(r)\right) = -\log\left(\xi^\varepsilon_t(r)\right)$.

By Theorem 3.3, $\xi^\varepsilon \Rightarrow \theta$ in $D([0,T];C(\mathbf{R}))$. We next note that $\widehat{\Psi}$ is continuous on the open subset $D([0,T];\mathcal{C}_+(\mathbf{R})) \subset D([0,T];C(\mathbf{R}))$. Using Theorem 3.1, (*iii*) we have $\mathcal{P}\left(\theta \in D([0,T];\mathcal{C}_+(\mathbf{R}))\right) = 1$, so that we can apply [5, Theorem 5.1] and conclude $\widehat{\Psi}(\xi^\varepsilon) \Rightarrow \widehat{\Psi}(\theta) = \Psi(\theta)$ a.s. By Theorem 3.2, $\Psi(\theta) = h$ (in law) and we are done. □

## 4. Convergence of the Transformed Process

In this section we prove the scaling limit for the transformed process $\xi^\varepsilon$. We first obtain, in Sect. 4.1, some moments estimates that imply the Hölder continuity of the process. In Sect. 4.2 we then establish a key estimate on the decay of the correlations. In Sect. 4.3 we finally complete the proof of Theorem 3.3 by showing compactness of $\xi^\varepsilon$ and that any weak limit solves the stochastic heat equation. We will use some properties of the transition probability for a random walk in $\mathbf{Z}$ that are proven in Appendix A.

We have to introduce some more notation. Let $p^\varepsilon_t(x)$ be the Green function associated with the drift term in Eq. (3.13), i.e. the function which solves

$$\begin{aligned} \partial_t p^\varepsilon_t(x) &= \frac{1}{2}e^{\gamma_\varepsilon}\Delta p^\varepsilon_t(x), \\ p^\varepsilon_0(x) &= \delta(x). \end{aligned} \tag{4.1}$$

We note that it can be interpreted as the transition probability for a symmetric random walk in continuum time over $\mathbf{Z}$.

The semimartingale Eq. (3.13) can then be rewritten as

$$\xi_t(x) = p_t^\varepsilon \circ \xi_0(x) + N_t^t(x), \tag{4.2}$$

where

$$p_t^\varepsilon \circ \xi_0(x) := \sum_y p_t^\varepsilon(x-y)\xi_0(y) \tag{4.3}$$

and for $s \le t$

$$N_s^t(x) := \int_0^s p_{t-\tau}^\varepsilon \circ dM_\tau(x) \equiv \sum_y \int_0^s p_{t-\tau}^\varepsilon(x-y)\, dM_\tau(y). \tag{4.4}$$

We note that for every fixed $t > 0$, $N_s^t(x)$, $s < t$ is a $\mathcal{F}_s$ martingale with bracket

$$\langle N^t(x), N^t(x)\rangle_s = \sum_y \int_0^s p_{t-\tau}^\varepsilon(x-y)^2\, d\langle M(y), M(y)\rangle_\tau. \tag{4.5}$$

*4.1 Moments estimates.* In this section we estimate the moments of $\xi_t(x)$, $\xi_t(x) - \xi_t(y)$ and $\xi_t(x) - \xi_s(x)$. By Eq. (3.14) we have

$$\frac{d}{dt}\langle M(x), M(x)\rangle_t \le c\,\varepsilon\,\xi_t(x)^2 \tag{4.6}$$

for some constant $c$. This bound allows *to close* the equations for the moments. In fact it is the only property of the martingale term in (3.13) we are going to use here.

Before starting we note that there is the following *a priori* bound:

$$\xi_t(x) \le \exp\left\{2\gamma_\varepsilon |x_t^0| + \gamma_\varepsilon |x| + \lambda_\varepsilon t\right\}, \tag{4.7}$$

and for all $n$, $\varepsilon$ and $t$ we can find $b = b_\varepsilon$ such that

$$\mathbf{E}_{\mu_\varepsilon}^\varepsilon\left(\exp\left\{2n\gamma_\varepsilon |x_t^0| + n\gamma_\varepsilon |x| + n\lambda_\varepsilon t\right\}\right) \le b. \tag{4.8}$$

This follows directly from the fact that $|x_t^0|$ is stochastically dominated by a Poisson random variable with mean $2t$ (since the jump rates for $x^0$ are bounded by 2). From the definitions (2.21) and (3.9), we obtain that for every $m$ there is $c$ such that for all $\varepsilon > 0$,

$$|\xi_t(x)^m - \xi_{t^-}(x)^m| \le c\,\varepsilon^{\frac{1}{2}}\,\xi_t(x)^m, \tag{4.9}$$

and by using again the fact that the jump rates are bounded (by 2) we have also that

$$\mathbf{P}_{\mu_\varepsilon}^\varepsilon\left(|\xi_t(x) - \xi_{t^-}(x)| > 0\right) = 0 \tag{4.10}$$

which, together with (4.7) and (4.8), implies that for all $n, m$,

$$\mathbf{E}_{\mu_\varepsilon}^\varepsilon\left(|\xi_t(x)^m - \xi_{t^-}(x)^m|^n\right) = 0. \tag{4.11}$$

We start by proving that the $L^p\left(\mathbf{P}_{\mu_\varepsilon}^\varepsilon\right)$ norm of $\xi_t(x)$ is bounded uniformly in $\varepsilon$.

**Lemma 4.1.** *For any $p \ge 1$, $T > 0$ there are $a, c > 0$ such that*

$$\sup_{t\in[0,\varepsilon^{-2}T]}\ \sup_{x\in\mathbb{Z}} e^{-a\varepsilon|x|}\ \|\xi_t(x)\|_{L^p\left(\mathbf{P}_{\mu_\varepsilon}^\varepsilon\right)} \le c \tag{4.12}$$

*for all $\varepsilon > 0$.*



*Proof.* For $s < t$ we have

$$\mathbf{E}_{\mu_\varepsilon}^\varepsilon\left(p_{t-s}^\varepsilon \circ \xi_s(x)\right)^n = \mathbf{E}_{\mu_\varepsilon}^\varepsilon\left(p_t^\varepsilon \circ \xi_0(x)\right)^n + \mathbf{E}_{\mu_\varepsilon}^\varepsilon \int_0^s d\tau\ (\partial_\tau + H_\varepsilon)\left(p_{t-\tau}^\varepsilon \circ \xi_\tau(x)\right)^n, \tag{4.13}$$

where we recall $H_\varepsilon$ is defined in (2.19), $\xi_t(x)$ in (3.9) and we shall use the representation (2.21) for $\zeta_t(x)$.

The first term on the right hand side of (4.13) can be bounded using the hypothesis on the initial condition. Indeed

$$\begin{aligned}\|p_t^\varepsilon \circ \xi_0(x)\|_{L^n\left(\mathbf{P}_{\mu_\varepsilon}^\varepsilon\right)} &\le \sum_y p_t^\varepsilon(x-y)\ \|\xi_0(y)\|_{L^n\left(\mathbf{P}_{\mu_\varepsilon}^\varepsilon\right)}\\ &\le c_1 \sum_y p_t^\varepsilon(x-y)\, e^{a_1\varepsilon|y|}\\ &\le c_1\, e^{a_1\varepsilon|x|} \sum_y p_t^\varepsilon(y)\, e^{a_1\varepsilon|y|} \le 2\,c_1\, e^{a_1\varepsilon|x|} \exp\left\{c_2 a_1^2 \varepsilon^2 t\right\}\end{aligned} \tag{4.14}$$

in which we used (2.11) and the fact that $\exp\{Ax\}$ is an eigenfunction of $\Delta$ with eigenvalue $2(\cosh A - 1)$.

Taking the limit $s \uparrow t$ in (4.13) and expanding the power on the right hand side we get

$$\begin{aligned}\mathbf{E}_{\mu_\varepsilon}^\varepsilon\left(\xi_t(y)\right)^n &\le c_3 e^{na_1\varepsilon|x|} \exp\left\{nc_2 a_1^2\varepsilon^2 t\right\}\\ &+\mathbf{E}_{\mu_\varepsilon}^\varepsilon \int_0^t d\tau \sum_{y_1,\dots,y_n} (\partial_\tau + H_\varepsilon)\prod_{i=1}^n \left(p_{t-\tau}^\varepsilon(x-y_i)\xi_\tau(y_i)\right),\end{aligned} \tag{4.15}$$

where we used (4.11), to exchange the limit with the expectation, and (4.14).

We next observe that

$$\begin{aligned}&H_\varepsilon \xi(x)^n - n\,\xi(x)^{n-1} H_\varepsilon \xi(x)\\ &= \frac{1}{2}\xi(x)^n \left\{\frac{1+\sigma(x)}{2}\left[e^{2n\gamma_\varepsilon} - 1 - n\left(e^{2\gamma_\varepsilon} - 1\right)\right]\right.\\ &\quad + \frac{1+\sigma(x+1)}{2} e^{2\gamma_\varepsilon}\left[e^{-2n\gamma_\varepsilon} - 1 - n\left(e^{-2\gamma_\varepsilon} - 1\right)\right]\\ &\quad \left. - \frac{1+\sigma(x)}{2}\,\frac{1+\sigma(x+1)}{2}\left[e^{2n\gamma_\varepsilon} - 1 - e^{2\gamma_\varepsilon}\left(e^{-2n\gamma_\varepsilon} - 1\right)\right]\right\}\\ &\le c_4\,\varepsilon\,\xi(x)^n.\end{aligned} \tag{4.16}$$

On the other hand, if we apply the generator to a product of $\xi$ computed at different points, $H_\varepsilon$ obeys the Leibniz rule. That is if $x_i \ne x_j$ for $i \ne j$,

$$H_\varepsilon\left[\xi(x_1)\cdots\xi(x_n)\right] = \sum_{i=1}^n \xi(x_1)\cdots\left[H_\varepsilon\xi(x_i)\right]\cdots\xi(x_n). \tag{4.17}$$

Recalling that $p_t^\varepsilon$ solves (4.1), formulae (4.15), (4.16) and (4.17) imply

$$\begin{aligned}\mathbf{E}_{\mu_\varepsilon}^\varepsilon\left(\xi_t(y)\right)^n &\le c_3 e^{na_1\varepsilon|x|}\exp\left\{nc_2 a_1^2\varepsilon^2 t\right\}\\ &+c_4\,\varepsilon\,\mathbf{E}_{\mu_\varepsilon}^\varepsilon \int_0^t d\tau \sum_{y_1,\dots,y_n\,:\,y_1=y_2}\ \prod_{i=1}^n p_{t-\tau}^\varepsilon(x-y_i)\xi_\tau(y_i)\end{aligned} \tag{4.18}$$

and by using Hölder inequality the second term in the right hand side of (4.18) can be bounded by

$$c_4\,\varepsilon \int_0^t d\tau \sum_{y_1,\cdots,y_n:y_1=y_2} \prod_{i=1}^n p^\varepsilon_{t-\tau}(x-y_i)\left[\mathbf{E}^\varepsilon_{\mu_\varepsilon}(\xi_\tau(y_i))^n\right]^{\frac{1}{n}}. \tag{4.19}$$

For $t \le T$, $a = na_1$, define

$$f(t) := \sup_{x\in\mathbf{Z}} e^{-a\varepsilon|x|}\,\mathbf{E}^\varepsilon_{\mu_\varepsilon}(\xi_{\varepsilon^{-2}t}(x)^n), \tag{4.20}$$

the bounds (4.18) and (4.19) yield

$$f(t) \le c_5 + c_6\,\varepsilon \int_0^{\varepsilon^{-2}t} d\tau \sum_{y_1,\cdots,y_n:y_1=y_2} \prod_{i=1}^n p^\varepsilon_{\varepsilon^{-2}t-\tau}(x-y_i)\,e^{\frac{a}{n}\varepsilon|x-y_i|}\cdot f(\varepsilon^2\tau). \tag{4.21}$$

It will be shown in Appendix A that for any $T > 0$ there is a constant $c = c(T)$ such that for all $t \in [0, \varepsilon^{-2}T]$, $\varepsilon > 0$,

$$\sup_{x\in\mathbf{Z}} p^\varepsilon_t(x) \le 1 \wedge c\,t^{-1/2}. \tag{4.22}$$

By repeating the estimate done in (4.14) we then obtain

$$\sum_{y_1,\cdots,y_n:y_1=y_2} \prod_{i=1}^n p^\varepsilon_{\varepsilon^{-2}t-\tau}(x-y_i)\,e^{\frac{a}{n}\varepsilon|x-y_i|} \le c_7\,(\varepsilon^{-2}t-\tau)^{-\frac{1}{2}}. \tag{4.23}$$

Finally, (4.21) and (4.23) imply that for every $n$ there exists $c_8 > 0$ such that

$$f(t) \le c_8\left(1 + \int_0^t ds\,\frac{f(s)}{\sqrt{t-s}}\right), \tag{4.24}$$

hence the estimate (4.12) follows from the *singular generalized Gronwall's Lemma* (see [14, Lemma 6, p. 33]) which states that (4.24) implies that for all $T$ there is a constant $C = C(T, c_8)$ such that $\sup_{t\le T} f(t) \le C$. □

We now state and prove an Hölder estimate in $x$ for the process $\xi_t(x)$.

**Lemma 4.2.** *For all $p \ge 1$, $T > 0$, $\alpha < 1/2$ there are $a, c > 0$ such that*

$$\sup_{t\in[0,\varepsilon^{-2}T]} \|\xi_t(x) - \xi_t(y)\|_{L^p(\mathbf{P}^\varepsilon_{\mu_\varepsilon})} \le c\,e^{a\varepsilon(|x|+|y|)}\,(\varepsilon|x-y|)^\alpha \tag{4.25}$$

*for any $x, y \in \mathbf{Z}$ and all $\varepsilon > 0$.*

*Proof.* Recall the decomposition (4.2) and set

$$R^t_s(x,y) := N^t_s(x) - N^t_s(y) = \sum_z \int_0^s q^\varepsilon_{t-\tau}(x,y;z)\,dM_\tau(z) \tag{4.26}$$

in which we have introduced

$$q^\varepsilon_{t-\tau}(x,y;z) := p^\varepsilon_{t-\tau}(x-z) - p^\varepsilon_{t-\tau}(y-z). \tag{4.27}$$

We shall apply the Burkholder-Davis-Gundy (BDG) inequality to the martingale $R^t_s(x,y)$:

$$\mathbf{E}^\varepsilon_{\mu_\varepsilon}\left(R^t_s(x,y)^{2n}\right) \le c_1\,\mathbf{E}^\varepsilon_{\mu_\varepsilon}\left([R^t(x,y), R^t(x,y)]^n_s\right), \tag{4.28}$$

where $c_1 = c_1(n)$ is a universal constant and

$$\begin{aligned}[R^t(x,y), R^t(x,y)]_s &= \sum_{\tau\le s}\left(R^t_\tau(x,y) - R^t_{\tau^-}(x,y)\right)^2 \\ &= \sum_{\tau\le s}\sum_z q^\varepsilon_{t-\tau}(x,y;z)^2\,[\xi_\tau(y) - \xi_{\tau^-}(y)]^2 \end{aligned} \tag{4.29}$$

is the quadratic variation of $R^t(x,y)$, since $R^t(x,y)$ is a bounded variation martingale. The sum over $\tau$ stands for the sum over the jump times $\tau$, i.e. the (a.s. countably many) times for which the term in the sum does not vanish. This form of the BDG inequality can be found in [23, 6 E.3].

We observe that if we could replace the quadratic variation on the right hand side of (4.28) with the bracket process

$$\langle R^t(x,y), R^t(x,y)\rangle_s = \sum_z \int_0^s q^\varepsilon_{t-\tau}(x,y;z)^2\,d\langle M(z), M(z)\rangle_\tau, \tag{4.30}$$

then the bound (4.25) would follow from Lemma 4.1 and the Hölder property for $p^\varepsilon_t(x)$.

We thus first show that the bracket process and the quadratic variation are close. To do this, we take advantage of the fact that the quadratic variation minus the bracket process is a martingale [23] and apply the BDG inequality also to this martingale. Let us introduce

$$D^t_s(x,y) := [R^t(x,y), R^t(x,y)]_s - \langle R^t(x,y), R^t(x,y)\rangle_s \tag{4.31}$$

which quadratic variation is

$$\begin{aligned}[D^t(x,y), D^t(x,y)]_s &= \sum_{\tau\le s}\left(R^t_\tau(x,y) - R^t_{\tau^-}(x,y)\right)^4 \\ &= \sum_{\tau\le s}\sum_z q^\varepsilon_{t-\tau}(x,y;z)^4\,[\xi_\tau(z) - \xi_{\tau^-}(z)]^4, \end{aligned} \tag{4.32}$$

since $\langle R^t(x,y), R^t(x,y)\rangle_s$ is continuous and of bounded variation.

Partition $[0,t]$ into subintervals $\mathcal{I}_i := [i, i+1]$ for $i = 0,\ldots,[t]-1$ and $\mathcal{I}_{[t]} := [[t], t]$. Using (4.9) we get

$$\begin{aligned}&\left\|[D^t(x,y), D^t(x,y)]_s\right\|_{L^n(\mathbf{P}^\varepsilon_{\mu_\varepsilon})} \\ &\le c_3\,\varepsilon^2 \sum_z \sum_{i=0}^{[t]} \sup_{s\in\mathcal{I}_i}\left[q^\varepsilon_{t-s}(x,y;z)\right]^4 \cdot \left\|\sum_{\tau\in\mathcal{I}_i}\xi_\tau(z)^4\right\|_{L^n(\mathbf{P}^\varepsilon_{\mu_\varepsilon})}. \end{aligned} \tag{4.33}$$

Calling $Q_i$ the number of exchanges between the sites $z$ and $z+1$ in the interval $\mathcal{I}_i$, we have

$$\sum_{\tau\in\mathcal{I}_i}\xi_\tau(z)^4 \le \xi_{t_i}(z)^4\,Q_i\,(1 + c\sqrt{\varepsilon})^{4Q_i}\,e^{4\lambda_\varepsilon} \tag{4.34}$$

as it follows from (2.21), (3.9) and (4.9). But the jump rates are bounded above by 2 and so by the fact that (4.34) is monotone in $Q_i$ and that the random variable $Q_i$ is stochastically dominated by a Poisson random variable with mean 2 we obtain

$$\begin{aligned} &\Big\| \sum_{\tau \in \mathcal{I}_i} \xi_\tau(z)^4 \Big\|_{L^n(\mathbf{P}^\varepsilon_{\mu_\varepsilon})} \\ &\le \left(\mathbf{E}^\varepsilon_{\mu_\varepsilon}\left\{\xi_{t_i}(z)^{4n}\, \mathbf{E}^\varepsilon_{\mu_\varepsilon}\left[Q_i^n\,(1+c\sqrt{\varepsilon})^{4nQ_i}\, e^{4n\lambda_\varepsilon} \middle| \mathcal{F}_{t_i}\right]\right\}\right)^{1/n} \\ &\le \left\|\xi_{t_i}(z)^4\right\|_{L^n(\mathbf{P}^\varepsilon_{\mu_\varepsilon})} (c_\varepsilon(n))^{1/n} \\ &\quad \exp\left\{2(1+c\sqrt{\varepsilon})^{4n}/n\right\} \exp(4\lambda_\varepsilon) \le c_4\, e^{a\varepsilon|z|} \end{aligned} \tag{4.35}$$

in which $c_\varepsilon(n)$ is the $n^{\text{th}}$ moment of a Poisson random variable with mean $2(1+c\sqrt{\varepsilon})^{4n}$. The last inequality is then a consequence of Lemma 4.1 and $c_4 = c_4(n)$ is independent on $\varepsilon \in (0,1)$.

From (4.33) and (4.35) we get

$$\begin{aligned} \left\| \left[D^t(x,y), D^t(x,y)\right]_s \right\|_{L^n(\mathbf{P}^\varepsilon_{\mu_\varepsilon})} &\le c_5\,\varepsilon^2 \sum_z \sum_{i=0}^{[t]} \sup_{s\in\mathcal{I}_i} \left(q^\varepsilon_{t-s}(x,y;z)^4\right) e^{a\varepsilon|z|} \\ \le\; 8\,c_5\,\varepsilon^2 \sum_z \sum_{i=0}^{[t]} &\left[\sup_{s\in\mathcal{I}_i} \left(p^\varepsilon_{t-s}(x,z)^4\right) + \sup_{s\in\mathcal{I}_i}\left(p^\varepsilon_{t-s}(y,z)^4\right)\right] e^{a\varepsilon|z|}. \end{aligned} \tag{4.36}$$

From (A.5) we have that for all $x$, $z$ and all $i = 0,\ldots,[t]$ and $\varepsilon$ sufficiently small $\sup_{s\in\mathcal{I}_i} p^\varepsilon_{t-s}(x,z) \le 2\,e\,p^\varepsilon_{t-i}(x,z)$. Using the inequality $\exp(a\varepsilon|y|) \le \exp(a\varepsilon y) + \exp(-a\varepsilon y)$, the fact that $\exp(a\varepsilon y)$ is an eigenfunction of $\Delta$ with eigenvalue $2(\cosh(a\varepsilon)-1)$ and (4.22) we finally obtain, since $t \le \varepsilon^{-2}T$,

$$\begin{aligned} &\left\| \left[D^t(x,y), D^t(x,y)\right]_s \right\|_{L^n(\mathbf{P}^\varepsilon_{\mu_\varepsilon})} \\ &\le c_6\,\varepsilon^2\, e^{c_7\varepsilon^2 t + a\varepsilon(|x|+|y|)} \int_0^t \frac{1}{(t-s)^{1-b}} ds \le c_8 \varepsilon^{2-2b} e^{a\varepsilon(|x|+|y|)}, \end{aligned} \tag{4.37}$$

where $b > 0$ can be arbitrarily chosen.

By using (4.6), (4.26), (4.28), (4.30), (4.32) and (4.37), we have

$$\begin{aligned} &\left\|N^t_s(x) - N^t_s(y)\right\|^2_{L^{2n}(\mathbf{P}^\varepsilon_{\mu_\varepsilon})} \\ \le\;& c_8\,\varepsilon^{2-2b}\, e^{a\varepsilon(|x|+|y|)} + c_9\,\varepsilon \sum_z \int_0^s d\tau\, q^\varepsilon_{t-\tau}(x,y;z)^2 \left\|\xi_\tau(z)^2\right\|^2_{L^n(\mathbf{P}^\varepsilon_{\mu_\varepsilon})} \\ \le\;& c_8\,\varepsilon^{2-2b}\, e^{a\varepsilon(|x|+|y|)} + c_{10}\,\varepsilon \sum_z \int_0^s d\tau\, q^\varepsilon_{t-\tau}(x,y;z)^2\, e^{a\varepsilon|z|}, \end{aligned} \tag{4.38}$$

where we used again Lemma 4.1.

As indicated in Appendix A, for any $T > 0$, $\beta \le 1/2$ there is a constant $c = c(\beta, T)$ such that for all $t \in [0, \varepsilon^{-2}T]$ and any $x, y \in \mathbf{Z}$,

$$\sup_{z\in\mathbf{Z}} \left|q^\varepsilon_{t-\tau}(x,y;z)\right| = \sup_{z\in\mathbf{Z}} \left|p^\varepsilon_t(x-z) - p^\varepsilon_t(y-z)\right| \le c\,\frac{1}{\sqrt{t}}\,\frac{|x-y|^{2\beta}}{t^\beta}. \tag{4.39}$$

Let us now take the limit $s \uparrow t$ in (4.38) keeping in mind that $\lim_{s\uparrow t} |N^t_s(x) - N^t_t(x)| = 0$ $\mathbf{P}^\varepsilon_{\mu_\varepsilon}$–a.s. (and that the same quantity is bounded by $c\sqrt{\varepsilon}\xi_t(x)$, for which all the moments exist). Choose $\beta = \alpha < 1/2$ in (4.39) and $b < 1/2$. We get

$$\left\|N^t_t(x) - N^t_t(y)\right\|_{L^{2n}(\mathbf{P}^\varepsilon_{\mu_\varepsilon})} \le c_{11}\, e^{a\varepsilon(|x|+|y|)} \left(\varepsilon + (\varepsilon|x-y|)^\alpha\right). \tag{4.40}$$

On the other hand, because of our hypothesis (2.12) on the initial measure $\mu_\varepsilon$, we have

$$\left\|p^\varepsilon_t \circ \xi_0(x) - p^\varepsilon_t \circ \xi_0(y)\right\|_{L^{2n}(\mathbf{P}^\varepsilon_{\mu_\varepsilon})} \le c_{12}\, e^{a\varepsilon(|x|+|y|)} \left(\varepsilon|x-y|\right)^{\frac{1}{2}}, \tag{4.41}$$

which used together with (4.40) in (4.2), yields (4.25). □

An analogous Hölder estimate holds in the time variable $t$.

**Lemma 4.3.** *For all $b > 0$, $p \ge 1$, $T > 0$, $\alpha < 1/4$ there are $a, c > 0$ such that*

$$\sup_{x\in\mathbf{Z}} e^{-a\varepsilon|x|} \left\|\xi_t(x) - \xi_s(x)\right\|_{L^p(\mathbf{P}^\varepsilon_{\mu_\varepsilon})} \le c\left[\left(\varepsilon^2|t-s|\right)^\alpha + \varepsilon^{1-b}\right] \tag{4.42}$$

*for any $t, s \in [0, \varepsilon^{-2}T]$ and all $\varepsilon > 0$.*

*Proof.* Let $h > 0$, in order to analyze the martingale term in (4.2) we write

$$N^{t+h}_{t+h}(x) - N^t_t(x) = \left[N^{t+h}_{t+h}(x) - N^{t+h}_t(x)\right] + \left[N^{t+h}_t(x) - N^t_t(x)\right]. \tag{4.43}$$

In Appendix A we verify that for all $\beta \le 1$ there is $c$ such that for all $t \in [0, \varepsilon^{-2}T]$, $h > 0$,

$$\sup_{x\in\mathbf{Z}} \left|p^\varepsilon_{t+h}(x) - p^\varepsilon_t(x)\right| \le c\,\frac{1}{\sqrt{t}}\left(\frac{h}{t}\right)^\beta. \tag{4.44}$$

By repeating the same steps of the previous lemma and choosing $\beta = 2\alpha < 1/2$ in the above inequality, we obtain

$$\begin{aligned} &\left\|\lim_{s\uparrow t}\left[N^{t+h}_s(x) - N^t_s(x)\right]\right\|^2_{L^{2n}(\mathbf{P}^\varepsilon_{\mu_\varepsilon})} \\ &\le c_1\, e^{a\varepsilon|x|}\, \varepsilon^{2-2b} + \\ &+ c_2\,\varepsilon \sum_y \int_0^t d\tau\, \left[p^\varepsilon_{t+h-\tau}(x-y) - p^\varepsilon_{t-\tau}(x-y)\right]^2 \left\|\xi_\tau(y)^2\right\|_{L^n(\mathbf{P}^\varepsilon_{\mu_\varepsilon})} \\ &\le c_1\, e^{a\varepsilon|x|}\, \varepsilon^{2-2b} + \\ &+ c_3\, e^{a\varepsilon|x|}\, h^{2\alpha}\, \varepsilon \int_0^t d\tau\, (t-\tau)^{-\frac{1}{2}-2\alpha} \le c_4\, e^{a\varepsilon|x|}\left[(\varepsilon^2 h)^{2\alpha} + \varepsilon^{2-2b}\right], \end{aligned} \tag{4.45}$$

where we used Lemma 4.1.

By the same argument

$$\left\| \lim_{s\uparrow t} \left[N_{s+h}^{t+h}(x) - N_s^{t+h}(x)\right] \right\|_{L^{2n}(\mathbf{P}^\varepsilon_{\mu_\varepsilon})}^2$$

$$\le c_1 e^{a\varepsilon|x|} \varepsilon^{2-2b} +$$

$$+ c_1 \varepsilon \sum_y \int_t^{t+h} d\tau\, p^\varepsilon_{t+h-\tau}(x-y)^2 \left\|\xi_\tau(y)^2\right\|_{L^n(\mathbf{P}^\varepsilon_{\mu_\varepsilon})}$$

$$\le c_1 e^{a\varepsilon|x|} \varepsilon^{2-2b} + c_2 e^{a\varepsilon|x|} \varepsilon \int_t^{t+h} d\tau\, (t+h-\tau)^{-1/2}$$

$$\le c_3 e^{a\varepsilon|x|} \left[(\varepsilon^2 h)^{\frac{1}{2}} + \varepsilon^{2-2b}\right] \tag{4.46}$$

in which we used Lemma 4.1 and the bound (4.22).

Hence the bound (4.42) is implied by (4.45), (4.46) and the following estimate on the initial datum $\xi_0$

$$\|p^\varepsilon_{t+h} \circ \xi_0(x) - p^\varepsilon_t \circ \xi_0(x)\|_{L^{2n}(\mathbf{P}^\varepsilon_{\mu_\varepsilon})} \le c_7 e^{a\varepsilon|x|} \left(\varepsilon^2 h\right)^\alpha,$$

which is a direct consequence of (2.12). □

The next step is to obtain uniform Hölder estimates from the lemmata we just proved. We will use Kolmogorov Theorem as stated in [26, Th 2.1]. In our notation it becomes

**Proposition 4.4.** *Let $K$ be a compact subset of $\mathbf{R}^d$ and let $\{X_z\}_{z\in K}$ be a real valued process for which there exist three strictly positive constants $\alpha_1, \alpha_2$ and $\alpha_3$ such that*

$$\mathbf{E}\left(|X_z - X_{z'}|^{\alpha_1}\right) \le \alpha_2 |z-z'|^{d+\alpha_3} \tag{4.47}$$

*for all $z, z' \in K$. Then there exists a modification $\widetilde{X}$ of $X$ such that for every $\alpha \in [0, \alpha_3/\alpha_1)$, there is a constant $c > 0$ for which*

$$\mathbf{E}\left[\left(\sup_{z\neq z'} \frac{\left|\widetilde{X}_z - \widetilde{X}_{z'}\right|}{|z-z'|^\alpha}\right)^{\alpha_1}\right] \le c. \tag{4.48}$$

Since we are interested only in the distribution of $\xi$ we shall not distinguish between $\xi$ and the modification in the above proposition. The additional problem comes from the fact that for $\varepsilon > 0$, $\xi_t(x)$ is not continuous in $t$, so that (4.47) cannot be satisfied. This problem will be overcome by defining a new process $\overline{\xi}_t(x)$, which is simply the interpolation of the values of the previous process at integer times, and by estimating the distance of $\xi$ and $\overline{\xi}$ uniformly in space and time.

We start by establishing the uniform Hölder estimate in the space variable for fixed time.

**Lemma 4.5.** *For any $p \ge 1$, $T > 0$, $b < 1/2$ and any compact set $K \subset \mathbf{R}$ there is $c > 0$ such that*

$$\sup_{t\in[0,T]} \left\| \sup_{r,r'\in K} \frac{|\xi_{\varepsilon^{-2}t}(\varepsilon^{-1}r) - \xi_{\varepsilon^{-2}t}(\varepsilon^{-1}r')|}{|r-r'|^b} \right\|_{L^p(\mathbf{P}^\varepsilon_{\mu_\varepsilon})} \le c \tag{4.49}$$

*for all $\varepsilon > 0$.*



*Proof.* From the definitions (2.21) and (3.9) we get

$$\xi_t(r) = \xi_t([r]) \left\{\cosh(\gamma_\varepsilon(r-[r])) - \sigma_t([r]+1)\sinh(\gamma_\varepsilon(r-[r]))\right\}. \tag{4.50}$$

It is then straightforward to verify for each $b' \in (0, 1/2)$ there is constant $c_1$ such for all $r_1, r_2 \in \mathbf{R}$ and $\varepsilon \in (0,1)$,

$$\frac{|\xi_s(r_1) - \xi_s(r_2)|}{|r_1 - r_2|^{b'}} \le 2^{b'} \frac{|\xi_s([r_1]) - \xi_s([r_2])|}{|[r_1] - [r_2]|^{b'}} + c_1 \sum_{i=1,2} (\xi_s([r_i]+1) + \xi_s([r_i])). \tag{4.51}$$

The $p$-moment of first term (resp. the second) on the right hand side can bounded by using Lemma 4.2 (resp. Lemma 4.1). We then obtain that for all $p > 0$ there is a constant $c_2$ such that for all $r, r' \in K$ we have

$$\mathbf{E}^\varepsilon_{\mu_\varepsilon}\left(\left|\xi_{\varepsilon^{-2}t}(\varepsilon^{-1}r') - \xi_{\varepsilon^{-2}t}(\varepsilon^{-1}r)\right|^p\right) \le c_2 |r'-r|^{b'p}. \tag{4.52}$$

Choose $p$ and $b'$ such that $b' = b + (1/p) < 1/2$ and apply Proposition 4.4 with $\alpha_1 = p$, $\alpha_2 = c_2$ and $\alpha_3 = b'p - d = bd$ to get the result. □

Let us consider now the time dependence. To this purpose we introduce a new process $\overline{\xi}_t(r)$, $t \in \mathbf{R}^+$, $r \in \mathbf{R}$, which is defined as

$$\overline{\xi}_t(r) := ([t] + 1 - t)\,\xi_{[t]}(r) + (t - [t])\,\xi_{[t]+1}(r). \tag{4.53}$$

**Lemma 4.6.** *For any $p \ge 1$, $T > 0$, $b < 1/4$ and any compact set $K \subset \mathbf{R}$ there is $c > 0$ such that*

$$\left\| \sup_{r,r'\in K} \sup_{t,t'\in[0,T]} \frac{|\overline{\xi}_{\varepsilon^{-2}t'}(\varepsilon^{-1}r) - \overline{\xi}_{\varepsilon^{-2}t}(\varepsilon^{-1}r')|}{(|t-t'| + |r-r'|)^b} \right\|_{L^p(\mathbf{P}^\varepsilon_{\mu_\varepsilon})} \le c \tag{4.54}$$

*for all $\varepsilon > 0$.*

*Proof.* It is very similar to the proof of Lemma 4.5. In fact

$$\frac{|\overline{\xi}_{\varepsilon^{-2}t'}(\varepsilon^{-1}r') - \overline{\xi}_{\varepsilon^{-2}t}(\varepsilon^{-1}r)|}{(|t'-t| + |r'-r|)^{b'}} \le \frac{|\overline{\xi}_{\varepsilon^{-2}t'}(\varepsilon^{-1}r') - \overline{\xi}_{\varepsilon^{-2}t'}(\varepsilon^{-1}r)|}{|r'-r|^{b'}} + \frac{|\overline{\xi}_{\varepsilon^{-2}t'}(\varepsilon^{-1}r) - \overline{\xi}_{\varepsilon^{-2}t}(\varepsilon^{-1}r)|}{|t'-t|^{b'}}. \tag{4.55}$$

Choose $b' \in (b, 1/4)$ and note that $\overline{\xi}_t(r) \le \xi_{[t]}(r) + \xi_{[t]+1}$ and $\overline{\xi}_t(r) \le \xi_t([r]) + \xi_t([r]+1)$. The result (4.54) follows then by using the bound (4.51) (and the analogous one for the time dependence) to estimate the right hand side of (4.55) and by applying Lemmata 4.1, 4.2, 4.3 and Proposition 4.4. □

We are left with proving that $\xi$ and $\overline{\xi}$ are uniformly close, i.e. if $K$ is a compact subset of $\mathbf{R}$, $T, \delta > 0$, we want the measure of the set

$$B(\delta) := \left\{ \sup_{t\in[0,T]} \sup_{r\in K} \left|\overline{\xi}_{\varepsilon^{-2}t}(\varepsilon^{-1}r) - \xi_{\varepsilon^{-2}t}(\varepsilon^{-1}r)\right| > \delta \right\} \tag{4.56}$$

to be small, i.e. vanishing with $\varepsilon$. We need only a statement in probability, but a stronger statement is available too, see formula (4.60) in the proof below.

**Lemma 4.7.** *For any $\delta > 0$, $T > 0$ and any compact set $K \subset \mathbf{R}$,*

$$\lim_{\varepsilon \to 0} \mathbf{P}^{\varepsilon}_{\mu_\varepsilon}(B(\delta)) = 0. \tag{4.57}$$

*Proof.* Define $I \subset \mathbf{Z}^+ \times \mathbf{Z}$ as follows

$$I := \{(n, x) \,:\, n = 0, \dots [\varepsilon^{-2}T] \;,\; \min_{r \in K} |x - \varepsilon^{-1} r| \leq 1\} \tag{4.58}$$

and observe that the cardinality of $I$ is bounded by $C_{K,T}\varepsilon^{-3}$, where $C_{K,T} < \infty$ for every $T \in (0, \infty)$ and any compact set $K$. We claim that for $b \in (0, 1/4)$ there is $c$ such that for all $(n, x) \in I$,

$$\left\| \sup_{r \in [x, x+1]} \sup_{t \in [n, n+1]} \left|\bar{\xi}_t(r) - \xi_t(r)\right| \right\|_{L^p\left(\mathbf{P}^{\varepsilon}_{\mu_\varepsilon}\right)} \leq c\,\varepsilon^b. \tag{4.59}$$

If (4.59) is granted, (4.57) follows because for all $p > 0$,

$$\begin{aligned}
&\mathbf{E}^{\varepsilon}_{\mu_\varepsilon}\left\{ \sup_{r \in K} \sup_{t \in [0,T]} \left|\bar{\xi}_{\varepsilon^{-2}t}(\varepsilon^{-1} r) - \xi_{\varepsilon^{-2}t}(\varepsilon^{-1} r)\right|^p \right\} \\
&\leq C_{K,T}\,\varepsilon^{-3} \sup_{(n,x) \in I} \mathbf{E}^{\varepsilon}_{\mu_\varepsilon}\left\{ \sup_{r \in [x,x+1]} \sup_{t \in [n,n+1]} \left|\bar{\xi}_t(r) - \xi_t(r)\right|^p \right\} \\
&\leq C_{K,T}\, c^p\, \varepsilon^{-3+pb} \, .
\end{aligned} \tag{4.60}$$

Choose then $p > 3/b$ and apply the Chebyshev inequality to obtain (4.57).

We are left with the proof of (4.59). Let us fix $(n, x)$ and start by observing that for $s \in [n, n+1]$ and $r \in [x, x+1]$,

$$|\bar{\xi}_s(r) - \xi_s(r)| \leq \sum_{i=0,1} \left(|\xi_n(x+i) - \xi_s(x+i)| + |\xi_{n+1}(x+i) - \xi_n(x+i)|\right) \tag{4.61}$$

by using Lemma 4.3 we thus get

$$\begin{aligned}
&\left\| \sup_{r \in [x,x+1]} \sup_{t \in [n,n+1]} \left|\bar{\xi}_t(r) - \xi_t(r)\right| \right\|_{L^p\left(\mathbf{P}^{\varepsilon}_{\mu_\varepsilon}\right)} \\
&\leq c\,\varepsilon^b + \sum_{i=0,1} \left\| \sup_{t \in [n,n+1]} |\xi_n(x+i) - \xi_t(x+i)| \right\|_{L^p\left(\mathbf{P}^{\varepsilon}_{\mu_\varepsilon}\right)} .
\end{aligned} \tag{4.62}$$

In order to bound the second term in the right hand side of formula (4.62) we recall that the jump rates are bounded by 2 so that if $m = [p] + 1$, we obtain

$$\begin{aligned}
&\mathbf{E}^{\varepsilon}_{\mu_\varepsilon}\left( \sup_{t \in [n,n+1]} |\xi_n(x+i) - \xi_t(x+i)|^m \right) \leq \mathbf{E}^{\varepsilon}_{\mu_\varepsilon}\left( \xi_n(x+i)^m \left[e^{\lambda_\varepsilon}(1 + c\sqrt{\varepsilon})^Q - 1\right]^m \right) \\
&\leq \mathbf{E}^{\varepsilon}_{\mu_\varepsilon}\left\{ \xi_n(x+i)^m \mathbf{E}^{\varepsilon}_{\mu_\varepsilon}\left\{ \left[e^{\lambda_\varepsilon}(1 + c\sqrt{\varepsilon})^Q - 1\right]^m \Big| \mathcal{F}_n \right\}\right\} \\
&\leq \mathbf{E}^{\varepsilon}_{\mu_\varepsilon}\{\xi_n(x+i)^m\} \sum_{q=0}^{\infty} \frac{2^q e^{-2}}{q!} \left[e^{\lambda_\varepsilon}(1 + c\sqrt{\varepsilon})^q - 1\right]^m \leq c_1\, \varepsilon^{m/2}
\end{aligned} \tag{4.63}$$



in which $Q$ is the number of jumps involving site $x + i$ in the time interval $[n, n+1]$. We have dominated $Q$ (in the stochastic sense) by the number of jumps which occur in a Poisson process with rate 2 in a unit time. The last inequality is finally a consequence of Lemma 4.1 and what follows. Given $N \in \mathbf{N}$ we have

$$\begin{aligned}
&\sum_{q=0}^{\infty} \frac{2^q e^{-2}}{q!} \left(e^{\lambda_\varepsilon}(1 + c\sqrt{\varepsilon})^q - 1\right)^m \\
&\leq \sum_{q=0}^{N-1} \frac{2^q e^{-2}}{q!} \left(e^{\lambda_\varepsilon}(1 + c\sqrt{\varepsilon})^q - 1\right)^m + \frac{2^{N(m+1)}}{N!} e^{(2^{m+1}-2)} .
\end{aligned} \tag{4.64}$$

Choose then $N = |\log \varepsilon|$ and $\varepsilon \leq \bar{\varepsilon}(< 1)$, so that $cN\sqrt{\varepsilon} \leq 1$. Elementary computations and the fact that $e^{\lambda_\varepsilon}(1 + c\sqrt{\varepsilon})^q - 1 \leq \exp(\lambda_\varepsilon + cq\sqrt{\varepsilon}) - 1 \leq 2ecq\sqrt{\varepsilon}$ for all $q < N$ yield that the right hand side of (4.64) is bounded by

$$\begin{aligned}
&(2ec)^m \varepsilon^{m/2} \left( \sum_{q=0}^{\infty} \frac{2^q e^{-2} q^m}{q!} \right) + \frac{2^{N(m+1)}}{N!} \exp\{2^{m+1} - 2\} \\
&\leq \left[ \left( (2ec)^m \sum_{q=0}^{\infty} \frac{2^q e^{-2} q^m}{q!} \right) \vee \left(\exp\{2^{m+1} - 2\}\right) \right] \left( \varepsilon^{m/2} + \frac{2^{N(m+1)}}{N!} \right) .
\end{aligned} \tag{4.65}$$

By the choice of $N$, the proof of (4.63) is completed. The claim (4.59) is then proven by using (4.63) in (4.62). □

*4.2 The key estimate.* Up to now we did not fully exploit the content of the martingale term in (3.13). In fact we only used the bound (4.6). In order to study the convergence of $\xi_t$ we have to show that, on the space-time scale we are considering, $\langle M(x), M(x)\rangle_t \to \varepsilon\xi_t(x)^2$, and this will require some control over the decay in time of the correlations for $\xi_t$. This is the key point in the proof of the scaling limit for $\xi_t$.

**Lemma 4.8.** *For all $T > 0$, $\delta > 0$, there are $a, c > 0$ such that*

$$\sup_{x \in \mathbf{Z}} e^{-a\varepsilon|x|} \mathbf{E}^{\varepsilon}_{\mu_\varepsilon} \left|\mathbf{E}^{\varepsilon}_{\mu_\varepsilon}\left(\nabla^+ \xi_{\varepsilon^{-2}t}(x)\, \nabla^- \xi_{\varepsilon^{-2}t}(x) \,\big|\, \mathcal{F}_{\varepsilon^{-2}s}\right)\right| \leq c\,\varepsilon^{\frac{3}{2}-\delta} (t-s)^{-\frac{1}{2}} \tag{4.66}$$

*for all $\sqrt{\varepsilon} \leq s < t \leq T$ and all $\varepsilon > 0$.*

The idea of the proof is to express the conditioned expectation in term of the process $\xi$ until one obtains a closed scheme that can be iterated. It is however important to note that the result does not reflect a scaling property: the statement of the lemma would be false if in the left hand side of (4.66) we replace $\nabla^+\xi\ \nabla^-\xi$ with $[\nabla^{\pm}\xi]^2$ and the bound we find ($\varepsilon^{3/2-\delta}$, not optimal) it is the result of a cancellation.

*Proof.* Let us introduce the notation

$$K_t(x) := \nabla^+ p^{\varepsilon}_t(x)\ \nabla^-\ p^{\varepsilon}_t(x). \tag{4.67}$$

For $0 \leq s_1 \leq s_2 < t$, recalling (4.5), we have

$$\mathbf{E}^{\varepsilon}_{\mu_\varepsilon}\left(\nabla^+ N^t_{s_2}(x)\ \nabla^-\ N^t_{s_2}(x) \,\middle|\, \mathcal{F}_{s_1}\right) = \nabla^+ N^t_{s_1}(x)\ \nabla^-\ N^t_{s_1}(x)$$
$$+\,\mathbf{E}^{\varepsilon}_{\mu_\varepsilon}\left(\sum_y \int_{s_1}^{s_2} K_{t-\tau}(x-y)\, d\langle M(y),\, M(y)\rangle_\tau \,\middle|\, \mathcal{F}_{s_1}\right). \qquad (4.68)$$

By using the decomposition

$$\nabla^\pm \xi_t(x) = \nabla^\pm p^\varepsilon_t \circ \xi_0(x) + \nabla^\pm N^t_t(x),$$

we thus get

$$\mathbf{E}^{\varepsilon}_{\mu_\varepsilon}\left(\nabla^+\xi_t(x)\ \nabla^-\ \xi_t(x) \,\middle|\, \mathcal{F}_s\right) = \nabla^+ p^\varepsilon_t \circ \xi_0(x)\ \nabla^-\ p^\varepsilon_t \circ \xi_0(x)$$
$$+\ \nabla^+ p^\varepsilon_t \circ \xi_0(x)\ \nabla^-\ N^t_s(x)\ +\ \nabla^+ N^t_s(x)\ \nabla^-\ p^\varepsilon_t \circ \xi_0(x)$$
$$+\ \nabla^+ N^t_s(x)\ \nabla^-\ N^t_s(x)\ +$$
$$\mathbf{E}^{\varepsilon}_{\mu_\varepsilon}\left(\sum_y \int_s^t K_{t-\tau}(x-y)\, d\langle M(y),\, M(y)\rangle_\tau \,\middle|\, \mathcal{F}_s\right), \qquad (4.69)$$

where we considered the limit $s_2 \uparrow t$ using the fact that, as in (4.11),

$$\mathbf{E}^{\varepsilon}_{\mu_\varepsilon}\left(\nabla^+\xi_{t-}(x)\ \nabla^-\ \xi_{t-}(x) \,\middle|\, \mathcal{F}_s\right) = \mathbf{E}^{\varepsilon}_{\mu_\varepsilon}\left(\nabla^+\xi_t(x)\ \nabla^-\ \xi_t(x) \,\middle|\, \mathcal{F}_s\right).$$

The lemma will be proven by estimating the $L^1\left(\mathbf{P}^{\varepsilon}_{\mu_\varepsilon}\right)$ norm of the various terms in the right hand side of (4.69).

We start by looking at the terms containing the intial datum $\xi_0$,

$$\mathbf{E}^{\varepsilon}_{\mu_\varepsilon}\left(\left[\nabla^\pm p^\varepsilon_t \circ \xi_0(x)\right]^2\right) \le \sum_{y,y'} \nabla^\pm p^\varepsilon_t(x-y)\ \nabla^\pm\ p^\varepsilon_t(x-y')\, \mathbf{E}^{\varepsilon}_{\mu_\varepsilon}\left(\xi_0(y)\,\xi_0(y')\right)$$
$$\le c_2\, e^{a\varepsilon|x|}\left[\sum_y \nabla^\pm p^\varepsilon_t(y)\, e^{a\varepsilon|y|}\right]^2$$
$$\le c_2\, e^{a\varepsilon|x|} \sum_y \left[\nabla^\pm p^\varepsilon_t(y)\right]^2 e^{(a+1)\varepsilon|y|} \cdot \sum_y e^{-\varepsilon|y|}$$
$$\le c_3\, e^{a\varepsilon|x|}\, \varepsilon^{-1}\, t^{-\frac{3}{2}}, \qquad (4.70)$$

where we used the hypothesis on $\mu_\varepsilon$, Cauchy-Schwarz inequality and the last inequality follows from

$$\sum_y \left[\nabla^\pm p^\varepsilon_t(y)\right]^2 e^{a\varepsilon|y|} \le c\, t^{-\frac{3}{2}} \qquad (4.71)$$

which is verified in Appendix A.

Now we control the martingale terms $\nabla^\pm N^t_s$. From (4.6) and (4.4) we get

$$\mathbf{E}^{\varepsilon}_{\mu_\varepsilon}\left(\left[\nabla^\pm N^t_s(x)\right]^2\right) = \mathbf{E}^{\varepsilon}_{\mu_\varepsilon} \sum_y \int_0^s \left[\nabla^\pm p^\varepsilon_{t-\tau}(x-y)\right]^2 d\langle M(y),\, M(y)\rangle_\tau$$
$$\le\ c_4\, \varepsilon \sum_y \int_0^s d\tau\, \left[\nabla^\pm p^\varepsilon_{t-\tau}(x-y)\right]^2 e^{a\varepsilon|y|} \le c_5\, e^{a\varepsilon|x|}\, \varepsilon\, (t-s)^{-\frac{1}{2}} \qquad (4.72)$$

in which we used Lemma 4.1 and (4.71).



In order to estimate the last term in (4.69), we express $\langle M(y), M(y)\rangle_\tau$ as a function of $\xi_t$ as in (3.15); keeping in mind (3.10) we obtain

$$\mathbf{E}^{\varepsilon}_{\mu_\varepsilon}\left(\sum_y \int_s^t K_{t-\tau}(x-y)\, d\langle M(y),\, M(y)\rangle_\tau \,\middle|\, \mathcal{F}_s\right)$$
$$= \left(1 - \sqrt{1+2\varepsilon^{\frac{1}{2}}}\right)^2 \left[I^1_{t,s}(x) + \frac{1}{2} I^2_{t,s}(x)\right] + \left(1+\varepsilon^{\frac{1}{2}}\right) I^3_{t,s}(x), \qquad (4.73)$$

where

$$I^1_{t,s}(x) \ :=\ \sum_y \int_s^t d\tau\, K_{t-\tau}(x-y)\, \mathbf{E}^{\varepsilon}_{\mu_\varepsilon}\left(\xi_\tau(y)^2 \,\middle|\, \mathcal{F}_s\right),$$
$$I^2_{t,s}(x) \ :=\ \sum_y \int_s^t d\tau\, K_{t-\tau}(x-y)\, \mathbf{E}^{\varepsilon}_{\mu_\varepsilon}\left(\xi_\tau(y)\, \Delta\, \xi_\tau(y) \,\middle|\, \mathcal{F}_s\right), \qquad (4.74)$$
$$I^3_{t,s}(x) \ :=\ \sum_y \int_s^t d\tau\, K_{t-\tau}(x-y)\, \mathbf{E}^{\varepsilon}_{\mu_\varepsilon}\left(\nabla^+\xi_\tau(y)\, \nabla^-\ \xi_\tau(y) \,\middle|\, \mathcal{F}_s\right).$$

We start by estimating the easiest, which is $I^2_{t,s}(x)$. By the definition of $\xi_t(x)$ we have

$$|\Delta\xi_t(x)| \le c_7\, \varepsilon^{\frac{1}{2}}\, \xi_t(x), \qquad (4.75)$$

and, by Lemma 4.1, we then conclude

$$\mathbf{E}^{\varepsilon}_{\mu_\varepsilon}\left|I^2_{t,s}(x)\right| \le c_8\, \varepsilon^{\frac{1}{2}} \sum_y \int_s^t d\tau\, \left|K_{t-\tau}(x-y)\right|\, \mathbf{E}^{\varepsilon}_{\mu_\varepsilon}\left(\xi_\tau(y)^2\right) \le c_9\, e^{a\varepsilon|x|}\, \varepsilon^{\frac{1}{2}}. \qquad (4.76)$$

To obtain a useful estimate of $I^1_{t,s}(x)$ is more difficult: we have to exploit a hidden cancellation which is revealed in Lemma A.1. As a consequence of that result $I^1_{t,s}(x)$ can be rewritten as

$$I^1_{t,s}(x) \ =\ \sum_y \int_s^t d\tau\, K_{t-\tau}(x-y)\, \mathbf{E}^{\varepsilon}_{\mu_\varepsilon}\left(\xi_\tau(y)^2 - \xi_t(x)^2 \,\middle|\, \mathcal{F}_s\right)$$
$$+\ \mathbf{E}^{\varepsilon}_{\mu_\varepsilon}\left(\xi_t(x)^2 \,\middle|\, \mathcal{F}_s\right) \cdot \sum_y \int_{t-s}^\infty d\tau\, K_\tau(x-y). \qquad (4.77)$$

By Lemma 4.1 the $L^1\left(\mathbf{P}^{\varepsilon}_{\mu_\varepsilon}\right)$ norm of the second term in the right hand side of (4.77) is bounded by

$$c_{10}\, e^{a\varepsilon|x|} \int_{t-s}^\infty d\tau \sum_x \left[\nabla^+ p^\varepsilon_\tau(x)\right]^2 \le c_{11}\, e^{a\varepsilon|x|} \int_{t-s}^\infty d\tau\, \tau^{-\frac{3}{2}} = 2\, c_{11}\, e^{a\varepsilon|x|}\, (t-s)^{-\frac{1}{2}} \qquad (4.78)$$

in which we used (4.71). But the $L^1\left(\mathbf{P}^{\varepsilon}_{\mu_\varepsilon}\right)$ norm of the first term on the right hand side of (4.77) is smaller than

$$\sum_y \int_s^t d\tau \; |K_{t-\tau}(x-y)| \\ \|\xi_\tau(y)+\xi_t(x)\|_{L^2(\mathbf{P}^\varepsilon_{\mu_\varepsilon})} \cdot \|\xi_\tau(y)-\xi_t(x)\|_{L^2(\mathbf{P}^\varepsilon_{\mu_\varepsilon})} \\ \leq \; c_{12}\, e^{a\varepsilon|x|} \sum_y \int_s^t d\tau \; \left|\nabla^+ p^\varepsilon_{t-\tau}(y)\; \nabla^-\; p^\varepsilon_{t-\tau}(y)\right| \\ e^{a\varepsilon|y|}\left[(\varepsilon|y|)^{2\alpha} + (\varepsilon^2|t-\tau|)^\alpha + \varepsilon^{\frac{1}{2}}\right]. \tag{4.79}$$

where $\alpha < 1/4$. The second inequality follows from Lemmata 4.2 and 4.3.

In order to estimate the right hand side of (4.79) we observe that, by (A.3),

$$\sum_y \int_0^{\varepsilon^{-2}T} d\tau \; \left|\nabla^+ p^\varepsilon_\tau(y)\; \nabla^-\; p^\varepsilon_\tau(y)\right| \; e^{a\varepsilon|y|} \; \left(|y|^{2\alpha}+\tau^\alpha\right) \\ \leq \; c_{14} \int_0^{\varepsilon^{-2}T} d\tau\; 1\wedge\tau^{-\frac{3}{2}} \sum_{y\in\mathbf{Z}} p^\varepsilon_\tau(y)\, e^{a\varepsilon|y|} \; \left(|y|^{2\alpha}+\tau^\alpha\right) \\ \leq \; c_{15} \int_0^{\varepsilon^{-2}T} d\tau\; \tau^\alpha\; 1\wedge\tau^{-\frac{3}{2}} \leq c_{16} \tag{4.80}$$

Keeping in mind that $\alpha < 1/4$, the right hand side of (4.79) is thus bounded by $c_{17}\, e^{a\varepsilon|x|}\, \varepsilon^{\frac{1}{2}-\delta}$. We have hence proven that, for some $a \in \mathbf{R}^+$ and every $\delta > 0$,

$$\mathbf{E}^\varepsilon_{\mu_\varepsilon} \left|I^1_{t,s}(x)\right| \leq c_{18}\, e^{a\varepsilon|x|} \left(\varepsilon^{\frac{1}{2}-\delta} + (t-s)^{-\frac{1}{2}}\right). \tag{4.81}$$

Instead of estimating $I^3_{t,s}$, we note that it is of the same form of the left hand side in (4.69). We have thus obtained a closed inequality. We next show it can be iterated giving the bound (4.66). For $\sqrt{\varepsilon} \leq s < t \leq T$ let us define

$$f_{t,s}(x) := \varepsilon^{-1}\, \mathbf{E}^\varepsilon_{\mu_\varepsilon} \left|\mathbf{E}^\varepsilon_{\mu_\varepsilon}\left(\nabla^+\xi_{\varepsilon^{-2}t}(x)\,\nabla^-\;\xi_{\varepsilon^{-2}t}(x)\;\middle|\;\mathcal{F}_{\varepsilon^{-2}s}\right)\right|.$$

From (4.69), (4.70), (4.72), (4.73), (4.76) and (4.81) we obtain

$$f_{t,s}(x) \leq c_{19}\, e^{a\varepsilon|x|}\, \varepsilon^{\frac{1}{2}-\delta}\left[1+\varepsilon^{\frac{1}{2}}(t-s)^{-\frac{1}{2}}\right] + \varepsilon^{-2}\left(1+\varepsilon^{\frac{1}{2}}\right) \\ \int_s^t d\tau \sum_{y\in\mathbf{Z}} \left|K_{\varepsilon^{-2}(t-\tau)}(x-y)\right|\, f_{\tau,s}(y), \tag{4.82}$$

and if we define

$$f_a(t,s) := \sup_{x\in\mathbf{Z}} e^{-a\varepsilon|x|}\, f_{t,s}(x),$$

the inequality (4.82) implies

$$f_a(t,s) \leq c_{20}\, \varepsilon^{\frac{1}{2}-\delta}\left[1+\varepsilon^{\frac{1}{2}}(t-s)^{-\frac{1}{2}}\right] + \varepsilon^{-2}\left(1+\varepsilon^{\frac{1}{2}}\right) \\ \int_s^t d\tau \sum_x \left|K_{\varepsilon^{-2}(t-\tau)}(x)\right|\, e^{a\varepsilon|x|}\, f_a(\tau,s). \tag{4.83}$$



Because of the results stated in Lemmata A.3 and A.4, the above inequality can be iterated, for $\varepsilon$ sufficiently small, and we obtain a convergent series. The estimate (4.66) is thus proven. □

*4.3 Proof of the scaling limit.* In this section we prove the scaling limit for the process $\xi$ as stated in Theorem 3.3. We study the family of random variables $\xi^\varepsilon_t(r) = \xi_{\varepsilon^{-2}t}(\varepsilon^{-1}r)$, $\varepsilon > 0$ over $D([0,T];C(\mathbf{R}))$ and we show that such a family is tight. We then prove that the limit is concentrated on $C([0,T];C(\mathbf{R}))$. Taking the limit of $\xi^\varepsilon$ along a convergent subsequence we show it satisfies a suitable martingale problem. By the existence of a unique solution to the martingale problem, we finally conclude the proof.

The space $C(\mathbf{R})$ is metrizable and a bounded metric which generates the topology of uniform convergence over compact sets is

$$\varrho(f,g) = \sum_{n=1}^\infty \frac{1}{2^n}\left\{1\wedge \max_{x\in[-n,n]}|f(x)-g(x)|\right\}. \tag{4.84}$$

We recall that for $a, b \in \mathbf{R}$, $a\wedge b = \min(a,b)$.

*Proof of Theorem 3.3.*

*– Tightness and support properties*

We use a tightness criterium due to Aldous and Kurtz [21, Th. 2.7], which in our notation becomes

**Proposition 4.9.** *Let $\xi^\varepsilon \in D([0,T];C(\mathbf{R}))$. Suppose that for each $t\in[0,T]$ the family of random functions $\{\xi^\varepsilon_t\}_{\varepsilon>0}$ is tight in $C(\mathbf{R})$ and that for any $\delta > 0$ there exists a process $\{A_\varepsilon(\delta)\}_{\varepsilon>0}$ such that*

$$\mathbf{E}^\varepsilon_{\mu_\varepsilon}\left\{\varrho\left(\xi^\varepsilon_{t+\delta},\xi^\varepsilon_t\right)\middle|\,\mathcal{F}_t\right\} \leq \mathbf{E}^\varepsilon_{\mu_\varepsilon}\left\{A_\varepsilon(\delta)\middle|\,\mathcal{F}_t\right\} \tag{4.85}$$

*and*

$$\lim_{\delta\to 0}\,\overline{\lim_{\varepsilon\to 0}}\,\mathbf{E}^\varepsilon_{\mu_\varepsilon}\{A_\varepsilon(\delta)\} = 0, \tag{4.86}$$

*then the family $\{\xi^\varepsilon\}_{\varepsilon>0}$ is tight in $D([0,T];C(\mathbf{R}))$ and hence, by Prohorov's Theorem, there is a sequence $\{\varepsilon_n\}_{n=1,2,\dots}$ such that $\xi^{\varepsilon_n}$ converges as $n\to\infty$.*

First of all observe that Lemma 4.1 and Lemma 4.5, together with [5, Theorem 8.2], imply that $\{\xi^\varepsilon_t\}_{\varepsilon>0}$ is tight in $C(\mathbf{R})$.

It is sufficient to choose

$$A_\varepsilon(\delta) = \sup_{t\in[0,T]} \varrho\left(\xi^\varepsilon_{t+\delta},\xi^\varepsilon_t\right). \tag{4.87}$$

The condition (4.85) is then obvious. We denote by $\mathbf{1}_B$ the characteristic function of the event $B(\delta)$ defined in (4.56) and we choose $K = [-N,N]$ in that definition. For any $\delta,\delta' > 0$ and every $N \in \mathbf{Z}^+$ we have

$$\mathbf{E}^\varepsilon_{\mu_\varepsilon}\{A_\varepsilon(\delta)\} \leq \sum_{n=N+1}^\infty \frac{1}{2^n} + \\ + \sum_{n=1}^N \frac{1}{2^n}\mathbf{E}^\varepsilon_{\mu_\varepsilon}\left\{\sup_{t\in[0,T]}\left[1\wedge\max_{x\in[-n,n]}\left|\xi_{\varepsilon^{-2}(t+\delta)}(x)-\xi_{\varepsilon^{-2}t}(x)\right|\right]\left(\mathbf{1}_{B(\delta')}+\mathbf{1}_{B^c(\delta')}\right)\right\}$$

$$\leq 2^{-N} + \sum_{n=1}^{N} \frac{1}{2^n}$$

$$\left\{ \mathbf{E}^{\varepsilon}_{\mu_\varepsilon}\left(\mathbf{1}_{B(\delta')}\right) + 2\delta' + \mathbf{E}^{\varepsilon}_{\mu_\varepsilon}\left[\sup_{t\in[0,T]}\max_{x\in[-n,n]} \left|\overline{\xi}_{\varepsilon^{-2}(t+\delta)}(x) - \overline{\xi}_{\varepsilon^{-2}t}(x)\right|\right]\right\},$$

hence, by Lemmata 4.6 and 4.7,

$$\overline{\lim_{\varepsilon\to 0}}\, \mathbf{E}^{\varepsilon}_{\mu_\varepsilon}\{A_\varepsilon(\delta)\} \leq 2^{-N} + 2\delta' + c\,\delta^{\alpha},$$

where $\alpha \in (0, 1/4)$. By the arbitrariety of $\delta'$ and $N$ we have verified also condition (4.86) and this completes the proof of the tightness.

We observe that by Lemma 4.5 any limit point $\xi_t(r)$ of $\xi^{\varepsilon}_t(r)$ is a.s. $\alpha$–hölderian, $\alpha \in (0, 1/2)$ in $r$ for every $t$. By using Lemmata 4.6 and 4.7 it is straightforward to see that $\xi \in C([0,T], C(\mathbf{R}))$.

*– Identification of the limit*

In order to identify the limit of $\xi^{\varepsilon}$ we formulate the stochastic heat Eq. (3.1) as a martingale problem which will be shown to be equivalent to the mild formulation (3.4). In the following $f_t(r)$ denotes the canonical coordinate in $C(\mathbf{R}^+, C(\mathbf{R}))$. Recall that $\theta_0$ is a random element in $C(\mathbf{R})$ satisfying the growth condition (3.2).

**Definition 4.10.** *The martingale problem. Let $Q$ be a probability measure on $C(\mathbf{R}^+, C(\mathbf{R}))$ such that for all $T > 0$,*

$$\sup_{t\in[0,T]}\sup_{r\in\mathbf{R}} e^{-a|r|}\int f_t(r)^2 dQ < \infty \tag{4.88}$$

*for some $a > 0$. The measure $Q$ solves the martingale problem if $Q(f_0 \in A) = \mathbf{P}(\theta_0 \in A)$ for all Borel sets $A \subset C(\mathbf{R})$ and for all $\varphi \in \mathcal{D}(\mathbf{R})$,*

$$M_t(\varphi) := (f_t, \varphi) - (f_0, \varphi) - \frac{1}{2}\int_0^t ds\,(f_s, \varphi''), \tag{4.89}$$

$$\Lambda_t(\varphi) := M_t(\varphi)^2 - \int_0^t ds \int dr\, f_s(r)^2 \varphi(r)^2 \tag{4.90}$$

*are $Q$–local martingales.*

The connection between the process $\theta$ as defined in (3.4) and the martingale problem in Definition 4.10 is made by the following proposition, which will be proven in Sect. 5.

**Proposition 4.11.** *For every random function $\theta_0 \in C(\mathbf{R})$ satisfying (3.2), the martingale problem has a unique solution $Q$. Moreover $Q$ coincides with the law of the process of the process $\theta$ which solves (3.4).*

By Lemma 4.1 and Fatou's Lemma, we obtain that any limit point $\xi$ of the family $\{\xi^{\varepsilon}\}_{\varepsilon>0}$ satisfies

$$\sup_{t\in[0,T]}\sup_{r\in\mathbf{R}} e^{-a|r|}\,\mathbf{E}^{\varepsilon}_{\mu_\varepsilon}\left(\xi_t(r)^2\right) < \infty$$

so that the condition (4.88) is met.

Since, by (2.10) and (3.9), $\xi^{\varepsilon}_0 \Rightarrow \theta_0$ in $C(\mathbf{R})$ the initial condition in the *martingale problem* is also satisfied.

*- Martingale relations.*

In order to analyze the martingale structure of $\xi^{\varepsilon}$, we introduce the *density field for the transformed process*. For all $t \in [0,T]$, $\varphi \in \mathcal{D}(\mathbf{R})$ let us define

$$Y^{\varepsilon}_t(\varphi) := \varepsilon \sum_x \varphi(\varepsilon x)\,\xi_{\varepsilon^{-2}t}(x). \tag{4.91}$$

We also introduce

$$M^{\varepsilon}_t(\varphi) := Y^{\varepsilon}_t(\varphi) - Y^{\varepsilon}_0(\varphi) - \frac{1}{2}e^{\gamma_\varepsilon}\int_0^t ds\,\varepsilon \sum_x \varepsilon^{-2}\,\Delta\,\xi_{\varepsilon^{-2}s}(x)\,\varphi(\varepsilon x),$$

$$\Lambda^{\varepsilon}_t(\varphi) := M^{\varepsilon}_t(\varphi)^2 - \int_0^t ds\,\varepsilon\sum_x \xi_{\varepsilon^{-2}s}(x)^2\varphi(\varepsilon x)^2 + R^{\varepsilon}_1(\varphi) + R^{\varepsilon}_2(\varphi), \tag{4.92}$$

where

$$R^{\varepsilon}_1(\varphi) := \int_0^t ds\,\varepsilon\sum_x \varphi(\varepsilon x)^2$$

$$\times\left[\left(\varepsilon - (1 - e^{\gamma_\varepsilon})^2\right)\xi_{\varepsilon^{-2}s}(x)^2 - \frac{1}{2}\varepsilon^{-1}(1 - e^{\gamma_\varepsilon})^2\,\xi_{\varepsilon^{-2}s}(x)\,\Delta\,\xi_{\varepsilon^{-2}s}(x)\right],$$

$$R^{\varepsilon}_2(\varphi) := \int_0^t ds\,\varepsilon\sum_x \varphi(\varepsilon x)^2\,\frac{1}{2}\varepsilon^{-1}$$

$$\left(1 + e^{2\gamma_\varepsilon}\right)\nabla^-\xi_{\varepsilon^{-2}s}(x)\,\nabla^+\xi_{\varepsilon^{-2}s}(x). \tag{4.93}$$

The semimartingale Eq. (3.13) and (3.15) directly imply that $M^{\varepsilon}_t(\varphi)$ e $\Lambda^{\varepsilon}_t(\varphi)$ are $\mathbf{P}^{\varepsilon}_{\mu_\varepsilon}$–martingales. We are then left with studying the limit (along subsequences) of (4.92). Let us first show that the error terms in (4.93) vanish in the limit $\varepsilon \to 0$. From Lemma 4.1 and (4.75) it follows that for all $\varphi \in \mathcal{D}(\mathbf{R})$ and all $t \in \mathbf{R}^+$ there is $c$ such that

$$\mathbf{E}^{\varepsilon}_{\mu_\varepsilon}\left(R^{\varepsilon}_1(\varphi)^2\right) \leq c\,\varepsilon^{1/2}, \tag{4.94}$$

and so the $L^2$–norm of $R^{\varepsilon}_1(\varphi)$ vanishes as $\varepsilon$ goes to zero. To prove the same for $R^{\varepsilon}_2(\varphi)$ we have to use the key Lemma 4.8. We have

$$\mathbf{E}^{\varepsilon}_{\mu_\varepsilon}\left(R^{\varepsilon}_2(\varphi)^2\right) = \left(1 + e^{2\gamma_\varepsilon}\right)^2\int_0^t ds\int_0^s ds'\,\varepsilon^2\sum_{x,y}\varphi(\varepsilon x)^2\varphi(\varepsilon y)^2\varepsilon^{-2}$$

$$\times\,\mathbf{E}^{\varepsilon}_{\mu_\varepsilon}\left\{\nabla^-\xi_{\varepsilon^{-2}s'}(x)\,\nabla^+\xi_{\varepsilon^{-2}s'}(x)\,\mathbf{E}^{\varepsilon}_{\mu_\varepsilon}\left(\nabla^-\xi_{\varepsilon^{-2}s}(y)\,\nabla^+\xi_{\varepsilon^{-2}s}(y)\middle|\,\mathcal{F}_{\varepsilon^{-2}s'}\right)\right\}$$

$$\leq c_2\int_0^t ds\int_0^s ds'\,\varepsilon^2\sum_{x,y}\varphi(\varepsilon x)^2\varphi(\varepsilon y)^2$$

$$\times\,\mathbf{E}^{\varepsilon}_{\mu_\varepsilon}\left\{\xi_{\varepsilon^{-2}s'}(x)^2\,\varepsilon^{-1}\left|\mathbf{E}^{\varepsilon}_{\mu_\varepsilon}\left(\nabla^-\xi_{\varepsilon^{-2}s}(y)\,\nabla^+\xi_{\varepsilon^{-2}s}(y)\middle|\,\mathcal{F}_{\varepsilon^{-2}s'}\right)\right|\right\}, \tag{4.95}$$

since

$$\left|\nabla^{\pm}\xi_t(x)\right| \leq c_1\,\varepsilon^{\frac{1}{2}}\,\xi_t(x). \tag{4.96}$$

By Lemma 4.1 and the above inequality, we may assume $s' \geq \sqrt{\varepsilon}$. By Lemma 4.8 for every $\delta > 0$ there are $c_2, a > 0$ such that for any $x, y \in \mathbf{Z}$, $\kappa \in \mathbf{R}^+$,

$$\mathbf{E}^{\varepsilon}_{\mu_\varepsilon} \left\{ \mathbf{1}_{\{\xi_{\varepsilon^{-2}s'}(x) \leq \kappa\}}\, \xi_{\varepsilon^{-2}s'}(x)^2\, \varepsilon^{-1} \left| \mathbf{E}^{\varepsilon}_{\mu_\varepsilon} \left( \nabla^- \xi_{\varepsilon^{-2}s}(y)\, \nabla^+ \xi_{\varepsilon^{-2}s}(y) \middle| \mathcal{F}_{\varepsilon^{-2}s'} \right) \right| \right\}$$
$$\leq c_2\, e^{a\varepsilon(|x|+|y|)}\, \kappa^2\, \varepsilon^{1/2-\delta}. \tag{4.97}$$

Moreover by (4.96), Lemma 4.1 and the Cauchy–Schwarz inequality, there exists a constant $c_3$ such that for all $x, y \in \mathbf{Z}$, $\kappa \in \mathbf{R}^+$,

$$\mathbf{E}^{\varepsilon}_{\mu_\varepsilon} \left\{ \mathbf{1}_{\{\xi_{\varepsilon^{-2}s'}(x) > \kappa\}}\, \xi_{\varepsilon^{-2}s'}(x)^2\, \varepsilon^{-1} \left| \mathbf{E}^{\varepsilon}_{\mu_\varepsilon} \left( \nabla^- \xi_{\varepsilon^{-2}s}(y)\, \nabla^+ \xi_{\varepsilon^{-2}s}(y) \middle| \mathcal{F}_{\varepsilon^{-2}s'} \right) \right| \right\}$$
$$\leq c_3\, e^{a\varepsilon(|x|+|y|)}\, \mathbf{P}^{\varepsilon}_{\mu_\varepsilon} \left( \xi_{\varepsilon^{-2}s'}(x) > \kappa \right) \leq c_4\, e^{a'\varepsilon(|x|+|y|)}\, \kappa^{-2}, \tag{4.98}$$

where we used, in the last step, Chebyshev inequality and again Lemma 4.1.

By using the bounds (4.97) and (4.98) in (4.95), letting $\varepsilon \to 0$ and then $\kappa \to \infty$ we have proven that

$$\lim_{\varepsilon \to 0} \mathbf{E}^{\varepsilon}_{\mu_\varepsilon} \left( R^{\varepsilon}_2(\varphi)^2 \right) = 0. \tag{4.99}$$

Fix $\varphi \in \mathcal{D}(\mathbf{R})$ and rewrite $M^{\varepsilon}_t(\varphi)$ as

$$M^{\varepsilon}_t(\varphi) = Y^{\varepsilon}_t(\varphi) - Y^{\varepsilon}_0(\varphi) - \frac{1}{2} \int_0^t ds\, Y^{\varepsilon}_s(\varphi'') + R^{\varepsilon}_3(\varphi), \tag{4.100}$$

where $R^{\varepsilon}_3(\varphi)$ contains the errors coming from the approximations $\Delta\varphi(\varepsilon x) \sim \varepsilon^2 \varphi''(\varepsilon x)$ and $e^{\gamma_\varepsilon} \sim 1$. They are easily controlled.

For all $N \in \mathbf{Z}^+$ let us define the stopping time $\tau_N$ on $D([0,\infty); C(\mathbf{R}))$ by

$$\tau_N := \inf\{t \,:\, |M_t(\varphi)| \wedge |\Lambda_t(\varphi)| > N\}$$

with the standard convention that $\tau_N = +\infty$ if $|M_t(\varphi)| \leq N$ and $|\Lambda_t(\varphi)| \leq N$ for all $t \in \mathbf{R}^+$. We denote by $M^{\tau_N}(\varphi)$ (respectively $\Lambda^{\tau_N}(\varphi)$) the stopped martingale $M_{\cdot \wedge \tau_N}(\varphi)$ (respectively $\Lambda_{\cdot \wedge \tau_N}(\varphi)$).

Let $\mathcal{Q}^\varepsilon$ the law of the canonical coordinate process over $D([0,T], C(\mathbf{R}))$ which has the same law as $\xi^\varepsilon$. We have already proven the tightness of $\{\mathcal{Q}^\varepsilon\}_{\varepsilon > 0}$. For all $s, t$, $0 \leq s < t \leq T$, and every function $F : D([0,T]; C(\mathbf{R})) \to \mathbf{R}$ which is measurable over $D([0,s]; C(\mathbf{R}))$, continuous and bounded if restricted to $C([0,T]; C(\mathbf{R}))$, we have that for all converging subsequences $\mathcal{Q}^{\varepsilon_n} \Rightarrow \mathcal{Q}$,

$$0 = \lim_n \int d\mathcal{Q}^{\varepsilon_n} \left[ M^{\tau_N}_t(\varphi) - M^{\tau_N}_s(\varphi) \right] F = \int d\mathcal{Q} \left[ M^{\tau_N}_t(\varphi) - M^{\tau_N}_s(\varphi) \right] F,$$
$$0 = \lim_n \int d\mathcal{Q}^{\varepsilon_n} \left[ \Lambda^{\tau_N}_t(\varphi) - \Lambda^{\tau_N}_s(\varphi) \right] F = \int d\mathcal{Q} \left[ \Lambda^{\tau_N}_t(\varphi) - \Lambda^{\tau_N}_s(\varphi) \right] F,$$

which follows from (4.92), (4.100), (4.94),(4.99) and [5, 5.2]. In fact $[M^{\tau_N}_t(\varphi) - M^{\tau_N}_s(\varphi)]F$ is continuous and bounded if restricted to $C([0,T]; C(\mathbf{R}))$. By the arbitrariety of $F, s, t$ we conclude that $M^{\tau_N}(\varphi)$ and $\Lambda^{\tau_N}(\varphi)$ are $\mathcal{Q}$–martingales. Moreover

$$\lim_{N \to \infty} \tau_N = +\infty \quad \mathcal{Q} - \text{a.s.},$$

which follows directly from the fact that any limit point $\mathcal{Q}$ is supported by $C([0,T], C(\mathbf{R}))$ for all $T > 0$. This implies that $M(\varphi)$ and $\Lambda(\varphi)$ are $\mathcal{Q}$–local martingales. □

## 5. Properties of the Macroscopic Equations

In this section we characterize the solution of the KPZ equation as the weak limit of the corresponding approximating problems. This is carried out by using the Cole-Hopf transformation and the existence and uniqueness theorem for the stochastic heat equation. We next prove the martingale formulation of the latter equation which has been used in the identification of the limit for the transformed process.

*Proof of Theorems 2.1 and 3.2.* Fix $\kappa > 0$ and let $\theta^\kappa_t(r)$ be the solution of (3.6). A straightforward application of Ito calculus, see [3] for a similar computation, shows that $h^\kappa_t(r) := -\log \theta^\kappa_t(r)$ is a.s. differentiable and solves

$$h^\kappa_t = h_0 + \frac{1}{2} \int_0^t ds\, \left\{ \Delta h^\kappa_s - \left[ (\nabla h^\kappa_s)^2 - C_\kappa(0) \right] \right\} + W^\kappa_t, \tag{5.1}$$

where we recall $C_\kappa$ is defined in (2.2). The renormalizing coefficient $C_\kappa(0)$ arises from the extra term in the Ito's calculus.

Since $h^\kappa_t(r)$ is a.s. differentiable, Eq. (5.1) is equivalent to (2.3). To complete the proof of (*i*) we need only to show (2.3) has a unique solution in the class of adapted processes satisfying (2.5). Using the Cole-Hopf transformation this is equivalent to the uniqueness of (3.6) under the growth condition (3.8), by Theorem 3.1, (*i*) we then conclude.

Recall that the map $\Psi$ is defined in Theorem 3.2 and note that the weak convergence of $\{h^\kappa\}_{\kappa>0}$ in $C([0,T]; C(\mathbf{R}))$ is now equivalent to the weak convergence of $\{\Psi(\theta^\kappa)\}_{\kappa>0}$. The latter follows from the a.s. convergence, uniform in compact subsets of $\mathbf{R}^+ \times \mathbf{R}$, of $\theta^\kappa_t(r)$ to $\theta_t(r)$, Theorem 3.1 (*ii*) and from the a.s. positivity of $\theta_t(r)$, Theorem 3.1 (*iii*).

This argument in fact proves Theorems 2.1, (*ii*) and 3.2 at the same time. □

*Proof of Proposition 4.11.* The existence result for the martingale problem may be easily deduced from Theorem 3.1. However we do not need it since it follows, as shown in the previous section, from the convergence (along subsequences) of $\xi^\varepsilon$.

We show next that we can extend the probability space to accommodate a suitable Wiener process $W_t$ together with the martingale solution $f_t$: on this extended space $f_t$ and $W_t$ solve (3.4) (this is usually referred to as a *Representation Theorem*). By the strong uniqueness of the stochastic heat equation, Theorem 3.1 (*ii*), we then conclude that $\mathcal{Q}$ coincides with the law of $\theta$. In particular this will prove that the martingale problem has a unique solution.

To prove the Representation Theorem we follow the line of [19, Lemma 2.4]. By Definition 4.10, there exists a sequence of stopping times $\{\tau_N\}$ such that $\lim_{N\to\infty} \tau_N = +\infty$ $\mathcal{Q}$–a.s., $M^{\tau_N}(\varphi)$ is a square integrable martingale and $\Lambda^{\tau_N}(\varphi)$ is a martingale. Let us then consider the *martingale measure*, see [28, Ch.2], $M(ds\,dr)$ associated to $M^{\tau_N}(\varphi)$ and notice that it is an *orthogonal martingale measure*, its *quadratic variation measure* $\langle M\rangle(ds\,dx)$ is $f_s(r)^2 ds\,dr$.

Possibly by extending the probability space (and consequently the filtration), we introduce a cylindrical Wiener process $\overline{W}_t$, independent of $f_t$. Let us denote by $\mathcal{Q}'$ the probability measure on the extended space. Set

$$W^N_t(\varphi) := -\int_0^{t\wedge\tau_N} \int \frac{1}{f_s(r)} \mathbf{1}_{\{f_s(r)\neq 0\}} \varphi(r) M(ds\,dr) - \int_0^{t\wedge\tau_N} \left( \mathbf{1}_{\{f_s(\cdot)=0\}} \varphi, d\overline{W}_s \right). \tag{5.2}$$

By direct computation one checks that $W_t^N = W_t^{\tau_N}$, where $W_t$ is a cylindrical Wiener process, and that

$$M_t^{\tau_N}(\varphi) = -\int_0^{t\wedge\tau_N} (f_s\varphi, dW_s) \tag{5.3}$$

so that

$$(f_{t\wedge\tau_N}, \varphi) - (f_0, \phi) = \frac{1}{2}\int_0^{t\wedge\tau_N} (f_s, \Delta\varphi) - \int_0^{t\wedge\tau_N} (f_s\varphi, dW_s), \tag{5.4}$$

$Q'$–a.s.. By letting $N$ go to infinity in (5.4) and using the hypothesis (4.88) one directly shows that

$$\int dQ' \left[(f_t, \varphi) - (G_t * f_0, \varphi) + \int_0^t (f_s(G_{t-s} * \varphi), dW_s)\right]^2 = 0 \tag{5.5}$$

so that the quantity between square brackets is zero $Q'$–a.s.. By using a countable family $\{\varphi\}$ which separates points in $C(\mathbf{R})$ we obtain that $f_t$ and $W_t$ solve the stochastic heat Eq. (3.4). □

## A. Some Properties of the Transition Probability

We here state and prove some elementary properties of $p_t^\varepsilon(x)$, which is the probability kernel for a symmetric random walk on $\mathbf{Z}$ defined in (4.1).

Let $q_t(x)$ be the transition probability kernel defined by

$$\partial_t q_t(x) = \frac{1}{2}\,\Delta\, q_t(x) \quad , \quad q_0(x) = \partial(x), \tag{A.1}$$

and let us express it in Fourier series

$$q_t(x) = \frac{1}{2\pi}\int_{-\pi}^{\pi} dk\, e^{ikx}\, e^{-t(1-\cos k)}. \tag{A.2}$$

Hence $p_t^\varepsilon(x) = q_{e^{\gamma_\varepsilon t}}(x)$. By using (A.2) it is straightforward to verify the inequalities (4.22), (4.39) and (4.44). In the same way one easily shows that there is $c > 0$ such that for all $t > 0$, $\varepsilon \in (0, 1)$,

$$\sup_{x\in\mathbf{Z}} \left|\nabla^\pm p_t^\varepsilon(x)\right| \le c\; 1 \wedge t^{-\frac{3}{2}} \tag{A.3}$$

and this implies (4.71).

Another useful representation for $q_t(x)$ is

$$q_t(x) = e^{-t}\sum_{n=0}^{\infty} \frac{t^n}{n!}\, p_n(x), \tag{A.4}$$

where $p_n(x)$ is the probability that $S_n = x$, where $\{S_n\}_{n=0,1,2,\ldots}$ is a simple symmetric random walk starting at 0. From this representation we immediately obtain that for any $0 \le T_1 \le T_2$ and any $x$

$$\sup_{t\in[T_1,T_2]} q_t(x) \le \exp(T_2 - T_1)\, q_{T_1}(x). \tag{A.5}$$



**Lemma A.1.** *Let $K_t(x)$ be defined as in (4.67), then*

$$\sum_x \int_0^\infty dt\; K_t(x) = 0. \tag{A.6}$$

*Proof.* It follows from a straightforward computation by using the representation of $p_t^\varepsilon(x)$ in Fourier series (A.2). □

**Lemma A.2.** *Let $q_t(x)$ be defined as in (A.1), then*

$$\sum_x \int_0^\infty dt\; \left|\nabla^+ q_t(x)\, \nabla^-\, q_t(x)\right| < 1. \tag{A.7}$$

*Proof.* By Cauchy-Schwarz inequality

$$\sum_x \left|\nabla^- q_t(x)\, \nabla^+\, q_t(x)\right| < \left\{\sum_x [\nabla^- q_t(x)]^2\right\}^{\frac{1}{2}} \left\{\sum_x [\nabla^+ q_t(x)]^2\right\}^{\frac{1}{2}},$$

and we stress that the inequality is strict because $\nabla^+ q_t(x) \neq \nabla^- q_t(x)$.

By summing by parts and using (A.1), we obtain

$$\sum_x \int_0^\infty dt\; [\nabla^\pm q_t(x)]^2 = -\int_0^\infty dt\, \sum_x \partial_t\, (q_t(x)^2) = 1,$$

and (A.7) is proven. □

**Lemma A.3.** *For each $T > 0$, $a \ge 0$, there exist $\varepsilon_0 > 0$ and $\beta < 1$ such that*

$$\sum_x \int_0^{\varepsilon^{-2}T} dt\; |K_t(x)|\, e^{a\varepsilon|x|} \le \beta \tag{A.8}$$

*for every $\varepsilon < \varepsilon_0$.*

*Proof.* Let $\kappa > 0$ and start by considering

$$\int_\kappa^{\varepsilon^{-2}T} dt\, \sum_x |K_t(x)|\; e^{a\varepsilon|x|}. \tag{A.9}$$

By using (A.3) it can be bounded by

$$\begin{aligned} &c_1 \int_\kappa^{\varepsilon^{-2}T} dt\, t^{-\frac{3}{2}} \sum_x p_t^\varepsilon(x)\, e^{a\varepsilon|x|} \\ &\le 2c_1 \int_\kappa^{\varepsilon^{-2}T} dt\, t^{-\frac{3}{2}}\, e^{a^2\varepsilon^2 t} \le 4\, c_1\, e^{a^2 T}\, \kappa^{-\frac{1}{2}} = c_2\, \kappa^{-\frac{1}{2}}, \end{aligned} \tag{A.10}$$

since

$$\sum_x p_t^\varepsilon(x) e^{a\varepsilon x} \le c_3\, \exp\left\{c_4\, a^2\, \varepsilon^2\, t\right\}.$$

On the other side

$$\lim_{\varepsilon\to 0}\int_0^\kappa dt \sum_x |K_t(x)|\, e^{a\varepsilon|x|} = \int_0^\kappa dt \sum_x \left|\nabla^+ q_t(x)\, \nabla^- q_t(x)\right| \tag{A.11}$$

by the dominated convergence Theorem.

Choose $\kappa$ sufficiently large and then $\varepsilon_0$ small. Inequality (A.8) is then a consequence of (A.10) and Lemma A.2. □

**Lemma A.4.** *For every $T > 0$, $a \geq 0$, there is $c > 0$ such that*

$$\sum_x \int_0^{\varepsilon^{-2}T} dt\ |K_t(x)|\, e^{a\varepsilon|x|}\, \left(T - \varepsilon^2 t\right)^{-\frac{1}{2}} \leq c \tag{A.12}$$

*for every $\varepsilon > 0$.*

*Proof.* Let us start by considering

$$\sum_x \int_{\varepsilon^{-2}T/2}^{\varepsilon^{-2}T} dt\ |K_t(x)|\, e^{a\varepsilon|x|}\, \left(T - \varepsilon^2 t\right)^{-\frac{1}{2}}$$

which can be bounded by proceeding as in (A.10). We obtain

$$\int_{\varepsilon^{-2}T/2}^{\varepsilon^{-2}T} dt\, t^{-\frac{3}{2}}\, e^{a^2 T}\, \left(T - \varepsilon^2 t\right)^{-\frac{1}{2}} \leq c_1\, \varepsilon.$$

On the other side

$$\sum_x \int_0^{\varepsilon^{-2}T/2} dt\ |K_t(x)|\, e^{a\varepsilon|x|}\, \left(T - \varepsilon^2 t\right)^{-\frac{1}{2}}$$

$$\leq (T/2)^{-\frac{1}{2}} \sum_x \int_0^{\varepsilon^{-2}T/2} dt\ |K_t(x)|\, e^{a\varepsilon|x|} \leq c_2. \tag{A.13}$$

By Lemma A.3. □

## B. Derivation of the Stochastic Burgers Equation as Scaling Limit of WASEP

The relation of the KPZ Eq. (1.1) with the (viscous) Burgers equation with conservative noise it is rather well-known. The latter can be in fact obtained by introducing $u_t := \nabla h_t$ and formally differentiating (1.1)

$$\partial_t u_t = \frac{1}{2}\Delta u_t - \frac{1}{2}\nabla\left(u_t^2\right) + \nabla \dot{W}_t. \tag{B.1}$$

The stochastic Burgers equation has been recently analyzed for the case of non conservative noise, i.e. when the term $\nabla \dot{W}_t$ in (B.1) is replaced by $\dot{W}_t$, and in that case it has been shown that the process has continuous (in time and space) trajectories [3, 7]. Due to the extra spatial gradient in front of the stochastic term of (B.1) the natural state space of the process is a distribution space. Thanks to the correspondence between SOS and WASEP, see Sect.2.4, we have also a microscopic derivation of (B.1).

In this case the macroscopic process $u = u_t$ is a distribution valued process on the path space $C\left([0,T];\mathcal{D}'(\mathbf{R})\right)$ and it defined by

$$u_t(\varphi) := -\int dr\ h_t(r)\,\varphi'(r) = \int dr\ \log\theta_t(r)\,\varphi'(r), \tag{B.2}$$

where $h_t = h_t(r)$ is the KPZ process and $\theta = \theta_t(r)$ the solution of the stochastic heat Eq. (3.4).

**Theorem B.1.** *For $\varphi \in \mathcal{D}(\mathbf{R})$, $t \in [0,T]$, introduce the fluctuation field for WASEP as*

$$X_t^\varepsilon(\varphi) := \sqrt{\varepsilon} \sum_x \varphi(\varepsilon x)\sigma_{\varepsilon^{-2}t}(x) \tag{B.3}$$

*and regard $X^\varepsilon = (X_t^\varepsilon)_{t\in[0,T]}$ as a random element in $D\left([0,T];\mathcal{D}'(\mathbf{R})\right)$.*

*Assume the initial distribution $\mu^\varepsilon$ is such that the law of $\hat{\mu}_\varepsilon$ of $\zeta_0(\cdot)$ (defined in 2.21) under $\mu^\varepsilon$ satisfies the hypotheses in Definition 2.2. Then*

$$X^\varepsilon \Rightarrow u \tag{B.4}$$

*in the topology of $D\left([0,T];\mathcal{D}'(\mathbf{R})\right)$.*

*Sketch of the proof.* It is analogous to Theorem 2.3. In fact it is easier to prove the weak convergence in a finer (and metrizable) topology than in the natural topology of $D([0,T];\mathcal{D}'(\mathbf{R}))$. This will be possible because $u_t(\cdot)$ is (a.s.) not worse than the derivative of an Hölder continuous function. Let

$$G := \Big\{X \in \mathcal{D}'(\mathbf{R}) \,:\, \exists f_X \in C(\mathbf{R})$$

$$\text{such that, } \forall\varphi \in \mathcal{D}(\mathbf{R}),\ X(\varphi) = -\int dr\ \varphi'(r)\, f_X(r)\Big\}, \tag{B.5}$$

note that $f_X$ is unique up to a constant. Eliminate this ambiguity by requiring $f_X(0) = 1$ and endow $G$ with the metric

$$\varrho_0(X,Y) := \varrho(f_X, f_Y), \tag{B.6}$$

where $\varrho$ is the metric in $C(\mathbf{R})$ defined in (4.84).

Using the Cole-Hopf transformation as in Theorem 2.3, one shows $X^\varepsilon \Rightarrow u$ in the topology of $D([0,T];G)$; the statement in Theorem B.1 follows since the topology of $D\left([0,T];\mathcal{D}'(\mathbf{R})\right)$ is coarser. □

As an application we now show that the white noise is an invariant measure for the stochastic Burgers equation.

Let $\nu$ the white noise measure on $\mathcal{D}'(\mathbf{R})$, i.e. the Gaussian measure with mean zero and covariance

$$\int d\nu\ X(\varphi_1)\, X(\varphi_2) = (\varphi_1\,,\,\varphi_2) \tag{B.7}$$

with $\varphi_i \in \mathcal{D}(\mathbf{R})$.

**Proposition B.2.** *The measure $\nu$ is invariant for the process $u_t$.*

*Proof.* We consider WASEP with initial distribution $\nu_{1/2}$, the Bernoulli measure on $\Omega$ with marginals $\nu_{1/2}(\sigma(x)=\pm 1)=1/2$. Let $F$ be a bounded and continuous function on $\mathcal{D}'(\mathbf{R})$, by [22, VIII 2.1], which characterizes the invariant measures for WASEP, we have

$$\mathbf{E}^{\varepsilon}_{\nu_{1/2}}\left(F\left(X^{\varepsilon}_t\right)\right)=\int d\nu_{1/2}\,F\left(X^{\varepsilon}\right). \tag{B.8}$$

Since

$$\lim_{\varepsilon\to 0}\int d\nu_{1/2}\,e^{i\,X^{\varepsilon}(\varphi)}=\exp\left\{-\frac{1}{2}(\varphi\,,\varphi)\right\} \tag{B.9}$$

for all $\varphi\in\mathcal{D}(\mathbf{R})$, the right hand side of (B.8) converges to $\int d\nu\,F$ and thus the invariance of $\nu$ follows from Theorem B.1. □

*Acknowledgement.* We are grateful to D.A. Dawson, G. Jona-Lasinio and E. Presutti for very fruitful discussions. L.B. thanks for the warm hospitality the Courant Institute, where he benefitted from a very stimulating atmosphere.

## References


1. Albeverio, S., Röckner, M.: *Stochastic differential equations in infinite dimension: Solutions via Dirichlet forms.* Probab. Theory Relat. Fields **89**, 347–386 (1992)
2. Bertini, L., Cancrini, N.: *The stochastic heat equation: Feynman-Kac formula and intermittence.* J. Stat. Phys. **78**, 1377–1401 (1995)
3. Bertini, L., Cancrini, N., Jona-Lasinio, G.: *The stochastic Burgers equation.* Commun. Math. Phys. **165**, 211–232 (1994)
4. Bertini, L., Presutti, E., Rüdiger, B., Saada E.: *Dynamical fluctuations at the critical point: Convergence to a non linear stochastic PDE.* Theory Probab. Appl. **38**, 689–741 (1993)
5. Billingsley, P.: *Convergence of probability measures.* New York: Wiley 1968
6. Cassandro, M., Marra, R., Presutti, E.: *Corrections to the critical temperature in 2d Ising systems with Kac potentials.* J. Stat. Phys. **78**, 1131–1138 (1995)
7. Da Prato, G., Debussche, A., Temam, R.: *Stochastic Burgers equation.* Nonlinear Differential Equations Appl. **1**, 389–629 (1994)
8. De Masi, A., Presutti, E., Scacciatelli E.: *The weakly asymmetric exclusion process.* Ann. Inst. Henri Poincaré, Probab. et Stat. **25** 1–38 (1989)
9. Dittrich, P.: *Travelling waves and longtime behaviour of the weakly asymmetric exclusion process.* Probab. Theory Relat. Fields **86**, 443–455 (1990)
10. Dittrich, P., Gärtner, J.: *A central limit theorem for the weakly asymmetric simple exclusion process.* Math. Nachr. **151**, 75–93 (1991)
11. Esposito, R., Marra, R., Yau, H.T.: *Diffusive limit of asymmetric simple exclusion;* Rev. Math. Phys. **6**, 1233–1267 (1994).
12. Fritz, J., Rüdiger, B. *Time dependent critical fluctuations of a one dimensional local mean field model.* Probab. Theory Relat. Fields **103**, 381–407 (1995)
13. Gärtner, J.: *Convergence towards Burgers equation and propagation of chaos for weakly asymmetric exclusion process.* Stoch. Processes Appl. **27**, 233–260 (1988)
14. Haraux, A.: *Nonlinear evolution equations - Global behavior of solutions.* Lecture Notes in Mathematics **841**, Berlin: Springer 1981
15. Jona-Lasinio, G.: *Stochastic reaction diffusion equations and interacting particle systems.* Ann. Inst. Henri Poincaré, Phys. Theor. **55**, 751–758 (1991)
16. Jona-Lasinio, G., Mitter, P.K.: *On the stochastic quantization of field theory.* Commun. Math. Phys. **101**, 409–436 (1985)
17. Kardar, M., Parisi, G., Zhang, Y.-C.: *Dynamical scaling of growing interfaces.* Phys. Rev. Lett. **56**, 889–892 (1986)
18. Kipnis, C., Olla, S., Varadhan, S.R.S.: *Hydrodynamics and large deviations for simple exclusion process.* Commun. Pure Appl. Math. **42** 115–137 (1989)
19. Konno, N., Shiga, T.: *Stochastic partial differential equations for some measure-valued diffusions.* Probab. Theory Relat. Fields **79**, 201–225 (1988)
20. Krug, J., Spohn, H.: *Kinetic roughening of growing surfaces.* In: "Solids far from equilibrium: Growth, morphology and defects". C. Godrèche ed. Cambridge: Cambridge University Press, 1991
21. Kurtz, T.G.: *Approximation of population processes.* CBMS-NSF Reg. Conf. Series in Appl. Math. 36 (1981)
22. Liggett, T.M.: *Interacting particles systems.* Berlin: Springer, 1985.
23. Métivier, M.: *Semimartingales: aA course on stochastic processes.* Berlin New York: de Gruyter, 1982
24. Müller, C.: *On the support of solutions to the heat equation with noise.* Stochastics and Stochastics Reports, **37**, 225–245 (1991)
25. Müller, C., Tribe, R.: *Stochastic p.d.e.'s arising from the long range contact an long range voter process.* Probab. Theory Relat. Fields **102**, 519–545 (1995)
26. Revuz, D., Yor, M.: *Continuous martingales and Brownian motion.* Berlin: Springer, 1991
27. Spohn, H.: *Large scale dynamics of interacting particles.* Berlin: Springer, 1991
28. Walsh: *An introduction to stochastic partial differential equations.* In: *École d'Été de Probabilités de Saint–Flour XIV*, Lecture Notes in Mathematics n. 1180. Berlin: Springer, 1986


Communicated by J.L. Lebowitz