FLUCTUATION EXponent
OF THE KPZ/STOCHASTIC BURGERS EQUATION

M. BALÁZS, J. QUASTEL, AND T. SEPPÄLÄINEN

1. Introduction

1.1. Background. The Kardar-Parisi-Zhang (KPZ) equation [14] is a formal stochastic partial differential equation for a random function $h(t, x), t > 0, x \in \mathbb{R}$,

$$\partial_t h = -\lambda (\partial_x h)^2 + \nu \partial_x^2 h + \sigma \dot{W},$$

where $\nu > 0$ and $\lambda \neq 0$ are fixed parameters and $\dot{W}(t, x)$ is Gaussian space-time white noise

$$E[\dot{W}(t, x)\dot{W}(s, y)] = \delta(t-s)\delta(y-x).$$

It is widely studied in physics as a model of randomly growing interfaces. The derivative $u = \partial_x h$ should satisfy the stochastic Burgers equation (SBE),

$$\partial_t u = -\lambda \partial_x u^2 + \nu \partial_x^2 u + \sigma \partial_x \dot{W},$$

which is a one-dimensional toy model for a randomly stirred fluid.

Using renormalization group methods physicists have computed the dynamic scaling exponent (1.3)

$$z = 3/2.$$ 

Roughly, this means that one expects nontrivial behaviour under the rescaling

$$h_\varepsilon(t, x) = \varepsilon^{1/2} h(\varepsilon^{-2} t, \varepsilon^{-1} x).$$

We will be considering these models in equilibrium, in which case $h(t, x) - h(t, 0)$ is a two-sided Brownian motion with variance $\nu^{-1} \sigma^2$ for each $t$. There are many physical arguments for (1.3), none of which are good starting points for rigorous analysis, and which are really only convincing in the sense that they are very well backed up by numerical work. Perhaps the simplest is to note that the rescaling $\varepsilon^{1/2} h(\cdot, \varepsilon^{-1} x)$
is essential in order to preserve the spatial Brownian motions. Furthermore, \( h_\varepsilon \) defined by (1.4) satisfies (1.1) with renormalized coefficients,

\[
\lambda_\varepsilon = \varepsilon^{2-z} \lambda, \quad \nu_\varepsilon = \varepsilon^{2-z} \nu, \quad \sigma_\varepsilon = \varepsilon^{\frac{2-z}{2}} \sigma,
\]

which suggests that (1.3) gives the only nontrivial scaling.

For the totally asymmetric exclusion process (TASEP) and the polynuclear growth model (PNG), which can be thought of as discretizations of (1.1), it is now known rigorously [10, 22] that in a weak sense,

\[
\text{Var}(h_\varepsilon(t,x)) \simeq t^{2/3} g_{sc}(t^{2/3}(x - vt))
\]

for an explicit \( v \) and scaling function \( g_{sc} \) related to the Tracy-Widom distribution. Note that these models are in some sense exactly solvable.

KPZ (1.1) is ill-posed because the quadratic nonlinear term cannot possibly make sense as any solution should be locally Brownian in \( x \) for fixed \( t \). Formally applying the Hopf-Cole transformation

\[
Z(t,x) = \exp\{-\nu^{-1} h(t,x)\}
\]

to (1.1) leads to the stochastic heat equation (SHE)

\[
\partial_t Z = \nu \partial_x^2 Z - \lambda \nu^{-1} \sigma Z \dot{W}.
\]

This is interpreted in the mild sense,

\[
Z(t,x) = \int p^\nu(t,x-y)Z(0,y)dy - \lambda \nu^{-1} \sigma \int_0^t \int p^\nu(t-s,x-y)Z(s,y)W(dyds),
\]

where \( p^\nu(t,x) = (4\pi \nu t)^{-1/2} \exp\{-x^2/4\nu t\} \) is the heat kernel and the white noise integral is the Itô integral [27]. The advantage is that (1.8) is well-posed [27] with a solution \( Z(t,x) \) strictly positive and continuous in both variables.

Starting from this \( Z(t,x) \), we define the Hopf-Cole solution of KPZ as

\[
h(t,x) \overset{\text{def}}{=} -\lambda^{-1} \nu \log Z(t,x)
\]

and \( u = \partial_x h \). These are the physically relevant solutions of KPZ (1.1) and SBE (1.2) as they are the ones that one obtains through appropriate limits of microscopic models. They are also what one obtains as the limit of solutions of the equations with noise smoothed in space:

Applying Itô’s formula to \( h(t,x) \) defined through (1.10) one obtains (1.1) modulo an infinite constant coming from the Itô term. More precisely, convolving \( h \) with a smoothing kernel in the \( x \) variable, one obtains (1.1) in the limit as the kernel is removed, after subtracting a diverging constant. The problem is inherent and cannot be finessed with the replacement of Itô integrals by Stratonovich integrals. What is going on is that the nonlinear term in (1.1) needs to be replaced by an appropriately renormalized object. We take a pragmatic view that (1.10) defines the solution and study it directly. We leave for future work the problem of showing that an appropriate interpretation of (1.1) is well defined, with (1.10) as the solutions.

We briefly recall the well-posedness argument for SHE (1.8) since it provides the rigorous basis for all our results. Let \( Z_1 \) and \( Z_2 \) be two solutions. Letting \( \tilde{Z} = Z_1 - Z_2 \) denote the difference,

\[
\tilde{Z}(t,x) = -\lambda \nu^{-1} \sigma \int_0^t \int p_\nu(t-s,x-y)\tilde{Z}(s,y)W(dyds).
\]
The expectation of the square of the Itô integral is

\[ \int_0^t \int E[Z^2(s, y)]p_\nu^2(t-s, x-y)dy. \]

Iterating the resulting inequality for \( \phi(t, x) = E[Z^2(t, x)] \) gives the uniqueness rigorously. It is worth noting that it only works in one dimension, and that it is not because of the linearity as might be supposed. It applies equally well to \( \partial_t Z = \nu \partial_x^2 Z - f(Z)W \) as long as \( f \) is a Lipschitz function.

The solution \( Z(t, x) \) can also be thought of as a continuum directed polymer model. There is a Feynman-Kac formula

\[ Z(t, x) = E_x[\exp\{-\beta \int_0^t \dot{W}(s, b(s))ds\}Z(0, b(t))], \]

where the expectation is over an independent Brownian motion \( b(s), s \geq 0 \), starting at \( x \), of variance \( \nu, \beta = \lambda \nu^{-1} \sigma \), and \( \exp\): is the Wick-ordered exponential. \( \log Z(t, x) = -\lambda \nu^{-1} h(t, x) \) is then interpreted as a free energy. The Wick ordering here means that after expanding the exponential and then taking the expectation over the Brownian paths, the resulting series of multiple Wiener-Itô integrals should be time ordered in the natural way. This series is obtained directly from (1.9) by iterating the equality, and is easily shown to be convergent in \( L^2 \) of the white noise.

The reason we write (1.13) is to draw attention to the analogy with discrete directed polymers, a typical model being

\[ z(n, x) = E_x[\exp\{-\beta \sum_{m=1}^n X(m, s_m)\}]; \]

where \( X(m, r), m \in \{1, 2, \ldots\}, r \in \{\ldots, -1, 0, 1, \ldots\} \) are independent and identically distributed random variables, and \( s_m \) is a simple random walk starting at \( x \). For details on the definition of (1.13) and how it can be obtained as an appropriate limit of discrete directed polymers, see [1].

Assuming reasonable decay on the tails of the \( X' \)s, it is expected [16] that for any \( \beta \),

(1.15) \[ \text{Var}(\log z(n, 0)) \sim cn^{2\chi} \quad \text{with} \quad \chi = 1/3. \]

This exact exponent has been verified for two polymer models in [25, 26], and for certain last passage percolation models, which are obtained in the \( \beta \to \infty \) (zero-temperature) limit [2, 13]. Suboptimal but nontrivial bounds in the range \( \chi \in [3/10, 1/2] \) are obtained for some other models in [7, 17, 18, 20, 21] and [27]. Note the contrast with dimensions \( d \geq 3 \) where the polymer is known to be diffusive for small \( \beta \). For a recent survey, see [15].

1.2. Hopf-Cole solution as the limit of smoothed out noise. We now survey what is known rigorously about (1.11). In terms of well-posedness, the technology at the present time [23] can only handle far smoother noise terms than the white noise. An unusual type of Wick product version of the problem has been introduced [12]. But besides requiring fairly smooth noises, this does not have the scaling expected [9], and is therefore believed not to be physically relevant.

The idea in [6], which leads to the physically relevant solution, is to smooth out the white noise in space a little, and then use the Hopf-Cole transformation and the tractability of (1.8) to remove the cutoff. As this is done, one finds one has to subtract a large constant from the equation. The resulting Hopf-Cole solution of
\([1.1]\) is given explicitly as the logarithm of the well-defined solution of \([1.8]\). We now recall the details.

Let \(W(t), t \geq 0\), be the cylindrical Wiener process, i.e. the continuous Gaussian process with
\[
\langle \varphi, W(t) \rangle \langle \psi, W(s) \rangle = \min(t, s) \langle \varphi, \psi \rangle
\]
for any \(\varphi, \psi \in C_c^\infty(\mathbb{R})\), the smooth functions with compact support in \(\mathbb{R}\). The distributional time derivative \(\dot{W}(t, x)\) is space-time white noise. Note the mild abuse of notation for the sake of clarity, as we write \(\dot{W}(t, x)\) even though it is a distribution on \((t, x) \in [0, \infty) \times \mathbb{R}\) as opposed to a classical function of \(t\) and \(x\).

Let \(\mathcal{F}(t)\), \(t \geq 0\), be the natural filtration, i.e. the smallest \(\sigma\)-field with respect to which \(W(s)\) are measurable for all \(0 \leq s \leq t\). Let \(G \in C_c^\infty(\mathbb{R})\) be an even, non-negative function with total integral 1 and for \(\kappa > 0\), \(G_\kappa(x) = \kappa^{-1} G(\kappa^{-1} x)\). The mollified Wiener process is given by \(W^\kappa(t, x) = \langle \tau_x G_\kappa, W(t) \rangle\), where \(\tau_x f(y) = f(y + x)\) is the shift operator. The distributional time derivative of \(W^\kappa(t, x)\) is the Gaussian field with covariance
\[
\begin{align*}
C_\kappa(x - y) \delta(t - s) &= E[\dot{W}^\kappa(t, x) \dot{W}^\kappa(s, y)],
\end{align*}
\]
where
\[
C_\kappa(x) = \kappa^{-1} \int G(\frac{\kappa^{-1} x - y}{2}) G(-\frac{\kappa^{-1} x - y}{2}) dy, \quad C_\kappa(0) = \kappa^{-1} \|G\|_2^2.
\]

Assume the initial data \(h(0, x)\) is a random continuous function on \(\mathbb{R}\), independent of the white noise \(\dot{W}\), such that for each \(p > 0\) there is an \(a = a(p) < \infty\) such that
\[
\sup_{x \in \mathbb{R}} e^{-a|x|} E[e^{-p h(x)}] < \infty.
\]
The mollified KPZ equation is
\[
\partial_t h^\kappa = -\lambda[(\partial_x h^\kappa)^2 - C_\kappa(0)] + \nu \partial_x^2 h^\kappa + \sigma \dot{W}^\kappa.
\]
The unique solution is a continuous Markov process. We denote by \(P_\kappa\) the corresponding probability distribution on \(C([0, T], C(\mathbb{R}))\). The process
\[
Z^\kappa(t, x) = \exp\{-\lambda \nu^{-1} h_\kappa(t, x)\}
\]
is the unique mild solution of the Itô equation,
\[
\partial_t Z^\kappa = \nu \partial_x^2 Z^\kappa + \lambda \nu^{-1} \sigma Z^\kappa \dot{W}^\kappa,
\]
with initial data \(Z^\kappa(0, x) = \exp\{-\lambda \nu^{-1} h(0, x)\}\).

The following summarizes previous results, mostly from \([6]\).

**Proposition 1.1.** 1. As \(\kappa \searrow 0\), \(Z^\kappa(t, x) \to Z(t, x)\) almost surely, uniformly on compact sets of \([0, \infty) \times \mathbb{R}\), where \(Z\) is the unique mild solution of \([1.8]\) adapted to \(\bar{\mathcal{F}}(t) = \sigma(h_0, W_s, s \leq t), t \geq 0\) with \(Z(0, x) = \exp\{-\lambda \nu^{-1} h(0, x)\}\) \([6]\). Furthermore, \(Z(t, x) > 0\) for all \(t \geq 0\), for Lebesgue-a.e. \(x \in \mathbb{R}\) \([19]\).

2. \([6]\) The \(P_\kappa\) are tight as measures on \(C([0, T], C(\mathbb{R}))\). The limit process \(h(t)\) coincides with
\[
h(t, x) = -\lambda^{-1} \nu \log Z(t, x).
\]

3. \([6]\) Define \(u = \partial_x h\) in the sense of distributions. If we start \(Z\) with initial data \(Z(0, x) = \exp\{B(x)\}\), where \(B(x)\) is a two-sided Brownian motion independent of \(W\) with variance \(\nu^{-1} \sigma^2\), then \(u\) is stationary in both space and time. In particular, for each \(t \geq 0\), \(h(t, x) - h(t, 0)\) is a two-sided Brownian motion. Furthermore, the
time reversed process \( u(T - t), t \in [0, T] \) is a solution of \( (1.2) \) with \( \lambda \) replaced by \(-\lambda\).

Note that the authors of \( [6] \) only consider the case \( \lambda = \nu = 1/2 \) and \( \sigma = 1 \), but their proofs work in general. Point 2 says that Gaussian white noise with variance \( \nu^{-1} \sigma^2 \) is invariant for \( (1.2) \). Stationarity here means stationarity of \( \langle \tau_x \varphi, u(t) \rangle \) for smooth functions of compact support \( \varphi(x) \), where \( \langle \tau_x \varphi \rangle(y) = \varphi(y - x) \). The corresponding \( h(t, x) \) and \( Z(t, x) \) are not stationary in time with these initial data, but the increments \( D\phi h(t, x) = h(t, x + \delta) - h(t, x) \) are space and time stationary. The time reversal property in point 3 was not stated before. It follows easily from the approximation by weakly asymmetric simple exclusion, which is our main tool and which we now describe.

1.3. Hopf-Cole solution as the limit of discrete models. We consider a weaker asymmetric nearest neighbour (i.e. simple) exclusion process on \( \mathbb{Z} \) (WASEP). This is a system of continuous time random walks jumping to the right at rate \( q = 1/2 + \varepsilon^{1/2} \) (with \( \varepsilon \in (0, 1/4) \)), with the rule that jumps to already occupied sites are not realized. Hence the occupation variable can be taken to be \( \eta(t, x) = 1 \) or 0 depending on whether or not there is a particle at \( x \in \mathbb{Z} \) at time \( t \). One of the most important properties of this system is that it preserves Bernoulli product measures with any density \( \rho \in [0, 1] \). Here we take \( \rho = 1/2 \); i.e. we take as the initial configuration the independent Bernoulli \( \{\eta(0, x)\}, x \in \mathbb{Z} \), with density 1/2.

Let \( \hat{\eta} = 2\eta - 1 \) and define the height function

\[
(1.23) \quad \zeta_\varepsilon(t, x) = \begin{cases} 
\sum_{0 < y \leq x} \hat{\eta}(t, y) - 2N(t, 0), & x > 0, \\
-2N(t, 0), & x = 0, \\
-\sum_{x < y \leq 0} \hat{\eta}(t, y) - 2N(t, 0), & x < 0,
\end{cases}
\]

where \( N(t, 0) \) is the current across the bond \((0,1)\) up to time \( t \), i.e. the number of particles that jump from 0 to 1 minus the number of particles that jump from 1 to 0 in the time interval \([0,t]\). For \( x \in \mathbb{R} \) and \( t \geq 0 \) let \( h_\varepsilon(t, x) \) denote the rescaled height function

\[
(1.24) \quad h_\varepsilon(t, x) = \varepsilon^{1/2} \left( \zeta_\varepsilon(\varepsilon^{-2}t, \lfloor \varepsilon^{-1}x \rfloor) - \nu_\varepsilon t \right),
\]

where \( \nu_\varepsilon = \frac{1}{2} \varepsilon^{-3/2} - \frac{1}{4} \varepsilon^{-1/2} \) and the closest integer \([x]\) is given by

\[
(1.25) \quad [x] = \lfloor x + \frac{1}{2} \rfloor.
\]

We think of \( h_\varepsilon \) as an element of the space \( D([0,\infty); D_u(\mathbb{R})) \), where \( D \) refers to right continuous paths with left limits. \( D_u(\mathbb{R}) \) indicates that in space these functions are equipped with the topology of uniform convergence on compact sets. Because the discontinuities of \( h_\varepsilon(t, \cdot) \) are restricted to \( \varepsilon(1/2 + \mathbb{Z}) \), it is measurable as a \( D_u(\mathbb{R}) \)-valued random function (see Sec. 18 of \( [8] \)). Since the jumps of \( h_\varepsilon(t, \cdot) \) are uniformly small, local uniform convergence works for us just as well as the standard Skorokhod topology. The probability distribution of the process \( h_\varepsilon \) on \( D([0,\infty); D_u(\mathbb{R})) \) will be denoted \( \mathcal{P}_\varepsilon \).

**Proposition 1.2** (\( [8] \)). As \( \varepsilon \searrow 0 \), the distributions \( \mathcal{P}_\varepsilon \) converge weakly to \( \mathcal{P} \), the distribution of the Hopf-Cole solution \( h \) \( (1.22) \) of \( (1.1) \) with \( \lambda = 1/2, \nu = 1/2 \), and \( \sigma = 1 \).
Proof. This was proved in [6] using the slightly different height function $\zeta^{BG}(t,x)$ related to ours by
\begin{equation}
\zeta^{BG}(t,x) = \zeta_e(t,x) - 2(1 - \eta(0,0)).
\end{equation}
The result follows for $\zeta_e(t,x)$ because the difference is bounded. Note also that [6] makes the height functions continuous in space by linear interpolation and uses the smaller path space $D([0,\infty);C(\mathbb{R}))$. This makes no difference because $h_x^{BG}(t,x)$ and its continuous version are uniformly within distance $\varepsilon^{1/2}$ of each other. \qed

1.4. Statement of results. We can now state our main results about $h$.

**Theorem 1.3.** Let $h(t,x)$ be the Hopf-Cole solution of \([1.1]\) as in \([1.22]\) with $Z(t,x)$ the solution of \([1.5]\), with initial data $Z(0,x) = \exp\{B(x)\}$, where $B(x)$ is a two-sided Brownian motion independent of $W$ with variance $\nu^{-1}\sigma^2$. Let $\text{Var}(h(t,x))$ denote the variance of $h(t,x)$. $\text{Var}(h(t,x))$ is a symmetric function of $x$, non-decreasing in $|x|$, and
\begin{equation}
\text{Var}(h(t,x)) - |x| \geq 0.
\end{equation}
Furthermore, there exist $c_0 = c_0(\sigma,\nu,\lambda) < \infty$, $C_1 = C_1(\sigma,\nu,\lambda) < \infty$, and $C_2 = C_2(\sigma,\nu,\lambda) < \infty$, such that for $t \geq c_0$ we have
\begin{equation}
C_1 t^{2/3} \leq \text{Var}(h(t,0)) \leq C_2 t^{2/3}
\end{equation}
and, for $1 < m < 3$,
\begin{equation}
C_4 t^{2m/3} \leq \int |x|^{m-2} [\text{Var}(h(t,x)) - |x|] \, dx \leq C_5 t^{2m/3}.
\end{equation}

Now we turn to results about the correlations of $u(t,x)$. Throughout we will assume that it is in equilibrium with initial data white noise with variance $\nu^{-1}\sigma^2$. Note that $u(t,x)$ is a distribution and not a function, and it is not at all obvious that the correlation functions even make sense. The first result says that they do, as a time parametrized family of probability measures $S(t,dx)$. Presumably, they even make sense as a nice function (i.e. $S(t,dx) = S(t,x)dx$), but we do not pursue this here, as it is not necessary for our purposes.

By definition $\langle \varphi, u(t) \rangle = -\int \varphi'(h(t,x)) \, dx$ for $\varphi \in C_c^\infty(\mathbb{R})$. For each fixed $t > 0$, define a bilinear functional on $C_c^\infty(\mathbb{R}) \times C_c^\infty(\mathbb{R})$,
\[ B_t(\varphi, \psi) = E[\langle \varphi, u(t) \rangle \langle \psi, u(0) \rangle]. \]
The following proposition provides us with our definition of the space-time correlation measure of the stochastic Burgers equation.

**Proposition 1.4.** For each $t > 0$ there is a unique probability measure $S(t,dx)$ on $\mathbb{R}$ such that for $\varphi, \psi \in C_c^\infty(\mathbb{R})$,
\begin{equation}
B_t(\varphi, \psi) = \int_\mathbb{R} \left[ \frac{1}{2} \int_\mathbb{R} \varphi\left(\frac{y+x}{2}\right) \psi\left(\frac{y-x}{2}\right) \, dy \right] S(t,dx).
\end{equation}
$S(t,\cdot)$ is symmetric: $S(t,A) = S(t,-A)$ for Borel sets $A$, where $-A = \{-x : x \in A\}$.

Here are the bounds on $S(t)$. 

Theorem 1.5. Let \( u(t) = \partial_t h(t) \) be the distributional derivative of the Hopf-Cole solution (1.22) of (1.1), in equilibrium with initial data white noise with variance \( \nu^{-1} \sigma^2 \) and let \( S(t, dx) \) be the space-time correlation measure defined through (1.30). With the same constants as in Theorem 1.3 for \( t \geq t_0 \) we have

\[
C_1 t^{2m/3} \leq \int |x|^m S(t, dx) \leq C_2 t^{2m/3}.
\]

The upper bound holds for \( 1 \leq m < 3 \). The lower bound holds for all \( m \geq 1 \). In particular, the bulk diffusivity defined by

\[
D_{\text{bulk}}(t) = \frac{1}{t} \int x^2 S(t, dx)
\]

satisfies

\[
C_1 t^{1/3} \leq D_{\text{bulk}}(t) \leq C_2 t^{1/3}.
\]

Remark 1.6. The dependence of the constants \( c_0 \) and \( C \) on \( m, \sigma, \nu, \lambda \) is as follows: We can take

\[
c_0 = \tilde{c}_0 \nu^{-4} \lambda^{-4}, \quad C_1 = \tilde{C}_1 \frac{2m+2}{3} \nu^{-m} \lambda^{-\frac{2m}{3}}, \quad C_2 = C(m) \sigma^{-2m+2} \nu^{-\frac{m}{3}} \lambda^{-\frac{2m}{3}}
\]

for some \( \tilde{c}_0, \tilde{C}_1 \in (0, \infty) \) and \( C(m) = \frac{C}{(m-1)(m-2)} \), \( 1 < m < 3 \) and \( C(1) = C/4 \). Here \( m = 1 \) refers to (1.28). The lower bound holds for all \( m \geq 1 \).

Remark 1.7. The bulk diffusivity is defined as the variance of the probability measure \( S(t, dx) \), and it measures how correlations spread in the stochastic Burgers equation. It is also the asymptotic variance of the rescaled second class particle from the approximating weakly asymmetric simple exclusion process. One can also define diffusivity of the underlying polymer paths. However, the relation between the two is murky. For \( m \in [1, 3], m \neq 2, \) (1.29) and (1.31) give further information on the spread of \( \text{Var}(h(t, x)) \) and \( S(t, dx) \), consistent with (1.6). The results for \( h \) and \( u \) are essentially equivalent once we have the identity

\[
\text{Var}(h(t, x)) = x + 2 \int x (z - x) S(t, dz)
\]

for \( x \in \mathbb{R} \) and \( t > 0 \).

Remark 1.8. Sometimes we want to indicate the dependence of the solution on the parameters by writing \( h(t, x; \lambda, \nu, \sqrt{\nu}) \). It is interesting to take \( \lambda = 1, \sigma^2 = \nu \) because \( h(t, x; 1, \nu, \sqrt{\nu}) = \nu h(\nu^{-3} t, \nu^{-2} x; 1, 1, 1) \) corresponds to the solution of

\[
\partial_t h = - (\partial_x h)^2 + \nu \partial^2_x h + \sqrt{\nu} \dot{W}
\]

and note that the result implies that there is a \( \nu_0 > 0 \) and fixed \( 0 < c_1 \leq c_2 < \infty \) such that

\[
c_1 t^{2/3} \leq \text{Var}(h(t, 0; 1, \nu, \sqrt{\nu})) \leq c_2 t^{2/3}
\]

for \( t \geq 1 \) for all \( \nu \leq \nu_0 \).

Remark 1.9. In this article we have concentrated solely on the equations forced by white noise. Similar behaviour is expected for other noises, as long as the decorrelation in space and time is sufficiently rapid; however, rigorous proofs are lacking at the time of publication. If one forces the equation with noises which do not rescale to white noise, one naturally expects different behaviour.

We turn to the proofs. The first issue is to further develop the connection with the exclusion process from Section 1.3.
2. WEAKLY ASYMMETRIC SIMPLE EXCLUSION

The route to constructing $S(t)$ and proving (1.34) is somewhat circuitous. The measure $S(t)$ is constructed as a weak limit from the rescaled correlations of the particle process from Section 1.3. Then we show that, in the sense of distributions,

$$\frac{1}{2} \partial_x^2 \text{Var}(h(t, x)) = S(t, x).$$

Finally, after studying solutions of the stochastic heat equation, we can deduce (1.34).

Recall the WASEP process defined in Section 1.3 and the definition of the rescaled height function (1.23). We can also define the rescaled velocity field as

$$u_\varepsilon(t, x) = \varepsilon^{-1/2} \tilde{\eta}(\varepsilon^{-2} t, [\varepsilon^{-1} x])$$

with $[\cdot]$ defined by (1.25). The rescaled space-time correlation functions are given by

$$S_\varepsilon(t, x) = E_\varepsilon[u_\varepsilon(t, x)u_\varepsilon(0, 0)].$$

For $x \in \mathbb{R}$ let us define a discrete Laplacian by

$$\Delta_\varepsilon f(x) = \frac{1}{2} \varepsilon^{-2} (f(x + \varepsilon) - 2f(x) + f(x - \varepsilon))$$

and a discrete absolute value in terms of the closest integer function by

$$|x|_\varepsilon = |\varepsilon^{-1} x|.$$

We begin by building on some well-known properties. The assumption that density is $1/2$ is used repeatedly.

**Proposition 2.1.** 1. For a fixed $t > 0$, $S_\varepsilon(t, x)$ is a probability density on $\mathbb{R}$ and symmetric in $x$ except at $x \in \varepsilon (1/2 + \mathbb{Z})$.

2. $S_\varepsilon(t, x) = \Delta_\varepsilon \text{Var}(h_\varepsilon(t, x))$. $\text{Var}(h_\varepsilon(t, x)) - |x|_\varepsilon$ is symmetric in $x \notin \varepsilon (1/2 + \mathbb{Z})$, nondecreasing in $|x|_\varepsilon$, nonnegative, and for each fixed $(\varepsilon, t)$ has exponentially decaying tails in $x$.

3. For $1 \leq m < \infty$,

$$\int_{\mathbb{R}} |x|^m S_\varepsilon(t, x) dx = \int_{\mathbb{R}} \Delta_\varepsilon (|x|^m) [\text{Var}(h_\varepsilon(t, x)) - |x|_\varepsilon] dx.$$

**Proof.** To see that $S_\varepsilon(t, x)$ is a probability density, use the well-known connection with the second class particle:

$$S_\varepsilon(t, x) = \varepsilon^{-1} \mathbf{P}_\varepsilon^{1/2} \{x(\varepsilon^{-2} t) = [\varepsilon^{-1} x]\}.$$

Here $\mathbf{P}_\varepsilon^{1/2}$ is the coupling measure of two ASEP’s that start with one discrepancy at the origin and Bernoulli(1/2) occupations elsewhere, and $x(\cdot)$ is the position of the second class particle. This setting is discussed in Section 4. A proof of (2.7) can be found for example in [4]. Symmetry of $S_\varepsilon(t, x)$ can be seen from the definition (2.2) and the fact that

$$\tilde{\eta}(t, k) = 1 - \eta(t, -k)$$

defines an ASEP $\tilde{\eta}$ equal in distribution to $\eta$. See [22] for the explicit computation that $S_\varepsilon(t, x) = \Delta_\varepsilon \text{Var}(h_\varepsilon(t, x))$. 
We now work towards the rest of 2. Let \( N(t, x) \) denote the current across the bond between site \( x \) and \( x + 1 \) up to time \( t \). We start by checking that

\[
\text{Cov}(N(t, 0), \sum_{y=-x+1}^{x} \eta(t, y)) = 0.
\]

With \( \tilde{\eta} \) as in (2.8), an \( \tilde{\eta} \)-particle jump from \( x \) to \( y \) is the same as an \( \eta \)-hole jump from \(-x\) to \(-y\). Hence \( \tilde{N}(t, 0) = N(t, -1) \). By the distributional equality \( \tilde{\eta} \equiv \eta \),

\[
\text{Cov}(N(t, 0), \sum_{y=-x+1}^{x} \eta(t, y)) = \text{Cov}(\tilde{N}(t, 0), \sum_{y=-x+1}^{x} \tilde{\eta}(t, y)) = \text{Cov}(N(t, -1), \sum_{y=-x+1}^{x} \eta(t, y)) = \text{Cov}(N(t, -1), \sum_{y=-x+1}^{x} \eta(t, y)) = \text{Cov}(N(t, 0), \sum_{y=-x+1}^{x} \eta(t, y)),
\]

and the first formula of (2.9) gives (2.10). Next, note that by the finite range of ASEP, for fixed \( \varepsilon > 0 \), \( v(x) \) and \( |x| \) have the same discrete Laplacian. By (2.10) and (2.11), both \( v(x) - |x| \) and \( \Delta_1 v(x) \) decay exponentially, and, by 1 and the first formula of 2, \( \Delta_1 v(x) \geq 0 \). Then for \( x \in \mathbb{Z}^+ \),

\[
v(x) - |x| = \sum_{k=x}^{\infty} \sum_{\ell=k}^{\infty} 2 \Delta_1 v(\ell + 1) > 0.
\]

This also shows that \( v(x) - |x| \) is strictly decreasing for \( x \in \mathbb{Z}^+ \), and thereby (2.10) is strictly decreasing in \( |x| \).

Putting the scaling into (2.10) gives

\[
\text{Var}(h_{\varepsilon}(t, x)) - |x|_{\varepsilon} = 4 \varepsilon \text{Cov}(N(\varepsilon^{-2}t, 0), N(\varepsilon^{-2}t, [\varepsilon^{-1}x])).
\]

This completes the proof of 2. To prove 3, start with the observation

\[
\int_{-N}^{N} |x|_{\varepsilon}^{m} \Delta_{\varepsilon}(|x|_{\varepsilon}) \, dx = 0.
\]

Then by 2 and by integration by parts (that is, by shifting the integration variable),

\[
\int_{-N}^{N} |x|_{\varepsilon}^{m} S_{\varepsilon}(t, x) \, dx = \int_{-N}^{N} |x|_{\varepsilon}^{m} \Delta_{\varepsilon} \text{Var}(h_{\varepsilon}(t, x)) - |x|_{\varepsilon} \, dx
\]

\[
= \int_{-N}^{N} \Delta_{\varepsilon}(|x|_{\varepsilon}^{m}) \text{Var}(h_{\varepsilon}(t, x)) - |x|_{\varepsilon} \, dx + B_{\varepsilon, t, m}(N),
\]

where \( B_{\varepsilon, t, m}(N) \) are sums of integrals of \( |x \pm \varepsilon|_{\varepsilon}^{m} \text{Var}(h_{\varepsilon}(t, x)) - |x|_{\varepsilon} \) over integrals of length \( \varepsilon \) around \( \pm N \). By 2 these are exponentially small in \( N \) as \( \varepsilon, t \) and \( m \) are fixed. Taking \( N \to \infty \) gives 3. \( \square \)
The key technical estimate, which will be proved in Section 4, is

**Theorem 2.2.** With the same constants as in Theorem 1.3 for all $0 < \varepsilon < 1/4$, $1 \leq m < 3$, and $t \geq c_0$,

$$C_1 t^{2m/3} \leq \int |x|^m S_\varepsilon(t, x) \, dx \leq C_2 t^{2m/3}. \tag{2.14}$$

**Corollary 2.3.** 1. For $0 < \varepsilon < 1/4$ and $t \geq c_0$,

$$C_1 t^{2/3} \leq \Var(h_\varepsilon(t, 0)) \leq C_2 t^{2/3}. \tag{2.15}$$

2. For each $t > 0$, the family of probability measures $\{S_\varepsilon(t, x) \, dx\}_{0 < \varepsilon < 1/4}$ is tight.

**Proof.** Part 1 follows from Theorem 2.2 and case $m = 1$ of part 3 of Proposition 2.1 because $\Delta_\varepsilon(|x|) = \varepsilon^{-1}$ for $x \in [-\varepsilon/2, \varepsilon/2)$ and vanishes elsewhere.

For $t \geq c_0$, tightness of $\{S_\varepsilon(t, x) \, dx\}_{0 < \varepsilon < 1/4}$ follows from the upper bound in (2.14). For $0 < t < c_0$ recall the second class particle connection (2.7). Proposition 4 in [24] proves that the second moment

$$\int |x|^2 S_\varepsilon(t, x) \, dx = \varepsilon^2 E(|x(\varepsilon^{-2} t)|^2)$$

is monotone nondecreasing in $t$. Thus the large-$t$ bound gives the tightness for all $t > 0$.

**Remark 2.4.** One may ask why we have up to this point restricted ourselves to WASEP with density $1/2$. In fact, in Section 4, we will consider other densities. After subtraction of an appropriate drift, the rescaled height functions in WASEP with density $p \in (0, 1)$ will converge to the Hopf-Cole solution of KPZ with appropriate parameters $\lambda$, $\nu$ and $\sigma$. If $h$ is the Hopf-Cole solution of (1.1), as in (1.22), then

$$h^\gamma(t, x) = \gamma^{-\alpha} h(\gamma^{-\beta} t, \gamma^{-1} x)$$

is the Hopf-Cole solution of (1.1) with new coefficients

$$\lambda_\gamma = \gamma^{\alpha-\beta+2} \lambda, \quad \nu_\gamma = \gamma^{-\beta+2} \nu, \quad \sigma_\gamma = \gamma^{-\beta+2-2\alpha} \sigma. \tag{2.17}$$

Hence the solutions obtained as the limit of WASEP with densities other than $1/2$ can simply be obtained through rescaling the solution obtained from density $1/2$. So in terms of the weakly asymmetric limit, there is no loss of generality in considering only density $1/2$.

**Remark 2.5.** Earlier works of the authors [3], [4], [24] also obtained bounds analogous to (2.14) for asymmetric simple exclusion processes. However, the bound required here is more subtle, as one needs the exact dependence on the asymmetry $\varepsilon$. This was not available in the earlier works. To be explicit, if we consider the second class particle $x(t)$ in the weakly asymmetric simple exclusion process on $\mathbb{Z}$, with $p = 1/2$ and $q = 1/2 + \varepsilon^1/2$, the statement (2.14) is that there are constants $0 < C_1 \leq C_2 < \infty$ and a $c_0 < \infty$ such that for $t \geq \varepsilon^{-2} c_0$, for $1 \leq m < 3$, for all $\varepsilon \in (0, 1/4)$,

$$C_1 \varepsilon^m 3^{2m/3} \leq E(|x(t)|^m) \leq C_2 \varepsilon^m 3^{2m/3}.$$

The factor $\varepsilon^{m/3}$ is the key technical novelty of the present work. It turns out that it can be achieved by careful extension of the coupling methods developed in [3], [4], but not by the resolvent method of [24].
3. Proofs of the main results

Once one has Corollary 2.3 and Proposition 2.1 the main results are obtained by careful passage to the limit \( \varepsilon \searrow 0 \). We now describe the details.

As a preliminary point we discuss the regularity of \( \operatorname{Var}(h(t,x)) \). The control comes from the weak limit \( h_\varepsilon \to h \). We have

\[
\operatorname{Var}(h_\varepsilon(t,0)) = \int |x| |S_\varepsilon(t, x)| \, dx \leq \left( \int |x|^2 \, S_\varepsilon(t, x) \, dx \right)^{1/2}.
\]

As mentioned in the proof of Corollary 2.3, this last quantity is nondecreasing in \( t \). Consequently by the upper bound in (2.11) and the i.i.d. mean zero spatial increments of \( h_\varepsilon(t,x) \) [see (1.22)] we conclude that \( \operatorname{Var}(h_\varepsilon(t,x)) \) is locally bounded as a function of \( (t,x) \), uniformly in \( \varepsilon > 0 \). By the weak limit \( \operatorname{Var}(h(t,x)) \)

\[
\leq \lim_{\varepsilon \to 0} \operatorname{Var}(h_\varepsilon(t,x)) \quad \text{and} \quad \operatorname{Var}(h(x,t)) \text{ is locally bounded.}
\]

The \( x \)-symmetry of \( \operatorname{Var}(h(t,x)) \) follows from part 3 of Proposition 2.1 or from the weak limit \( h_\varepsilon \to h \) and the distributional symmetry of \( \zeta_\varepsilon(t, \cdot) \). For any fixed \( x_0 \), \( h(t,x) - h(t,x_0) \) is a Brownian motion in \( x \) and hence the continuity of \( x \mapsto \operatorname{Var}(h(t,x)) \). By studying the stochastic heat equation we prove in the Appendix that

\[
\operatorname{Var}(h(t,x)) - |x| \to 0 \quad \text{as} \quad |x| \to \infty.
\]

We turn to proving the main results.

Proof of (1.30). From the definitions we have for test functions \( \varphi, \psi \in C_0^\infty(\mathbb{R}) \),

\[
E[\langle \varphi', h_\varepsilon(t) \rangle \langle \varphi', h_\varepsilon(0) \rangle] = E[\langle \varphi, u_\varepsilon(t) \rangle \langle \psi, u_\varepsilon(0) \rangle] + O(\varepsilon)
\]

(3.2)

\[
= \frac{1}{2} \int \left[ \int \varphi \left( y + x \right) y \left( y - x \right) dy \right] S_\varepsilon(t, x) \, dx + O(\varepsilon).
\]

Let \( S(t, dx) \) denote a weak limit point of \( S_\varepsilon(t, x) \) as \( \varepsilon \searrow 0 \). Taking the limit in (3.2), the last expression becomes the right-hand side of (1.30).

Convergence of the first expectation in (3.2) to the left-hand side of (1.30) follows from the weak convergence \( h_\varepsilon \to h \) with the following additional justification. Since the limit process \( h \) is continuous (in time), the pair \( (h_\varepsilon(0), h_\varepsilon(t)) \) converges weakly to \( (h(0), h(t)) \). Integration against a \( C_0^\infty(\mathbb{R}) \) function is a continuous function on \( D_1(\mathbb{R}) \). By an application of uniform integrability and Schwarz inequality, for the claimed convergence it is enough to show the \( L^1 \) boundedness of \( |\varphi', h_\varepsilon(t)\|^4 \). To see this we first transform the inner product (this is the beginning of the computation that one performs to check (3.2)):

\[
\langle \varphi', h_\varepsilon(t) \rangle = \varepsilon^{1/2} \sum_{k \in \mathbb{Z}} \zeta_\varepsilon(\varepsilon^{-2} t, k) \left[ \varphi((k + \frac{1}{2}) \varepsilon) - \varphi((k - \frac{1}{2}) \varepsilon) \right]
\]

\[
= -\varepsilon^{1/2} \sum_{k \in \mathbb{Z}} \varphi((k - \frac{1}{2}) \varepsilon) \psi(\varepsilon^{-2} t, k).
\]

(The constant term on the right-hand side of definition (1.24) vanishes since we are integrating the height function against a derivative.) This is a sum of independent mean zero random variables, and

\[
E[|\varphi', h_\varepsilon(t)|^4] \leq C \varepsilon^2 \sum_k \varphi((k - \frac{1}{2}) \varepsilon)^4 + C \varepsilon^2 \sum_{k, \ell} \varphi((k - \frac{1}{2}) \varepsilon)^2 \varphi((\ell - \frac{1}{2}) \varepsilon)^2,
\]

which is bounded uniformly in \( \varepsilon \). □
Proof of Theorem 1.3 and the upper bound of Theorem 1.3} Note first of all that by Remark 2.4, it suffices to prove all results with \( \lambda = 1/2, \nu = 1/2, \sigma = 1 \).

The upper bounds of (1.28) and (1.31) follow from the weak convergence and from the upper bounds in (2.14) and in \( I \) of Corollary 2.3.

Let \( 1 \leq m < 3 \). For the upper bound of (1.29) we collect these ingredients: inequality \( \text{Var}(h_\varepsilon(t,x)) - |x| \geq 0 \) from part 2 of Proposition 2.1, the fact that under the weak limit,

\[
\lim_{\varepsilon \searrow 0} \text{Var}(h_\varepsilon(t,x)) - |x| \geq \text{Var}(h(t,x)) - |x|,
\]

and for \( x \neq 0 \), \( \Delta_\varepsilon(|x|^m) \to m(m-1)|x|^{m-2}/2 \). Combine the upper bound in (2.14) with identity (2.6), let \( \varepsilon \searrow 0 \) in (2.6) and use Fatou’s Lemma.

To prove the lower bound of (1.31), let \( t > c_0 \) be fixed and choose a nonnegative smooth function \( f(x) \) with compact support such that \( f(x) \geq |x|^m \) for \( |x| \leq At^{2/3} \). We have

\[
\int f(x)S(t, dx) = \lim_{\varepsilon \searrow 0} \int f(x)S_\varepsilon(t, dx)
\]

and furthermore

\[
\int f(x)S_\varepsilon(t, dx) \geq \int_{|x| \leq At^{2/3}} |x|^m S_\varepsilon(t, dx).
\]

Choose \( \delta > 0 \) such that \( m + \delta < 3 \). By Chebyshev’s inequality and Theorem 2.2

\[
\int_{|x| \geq At^{2/3}} |x|^m S_\varepsilon(t, dx) \leq A^{-\delta} t^{-2\delta/3} \int |x|^{m+\delta} S_\varepsilon(t, dx) \leq A^{-\delta} Ct^{2m/3}.
\]

Hence, for appropriately chosen \( A \),

\[
\int f(x)S_\varepsilon(t, dx) \geq (C^{-1}/2)t^{2m/3}.
\]

Since this is true for all such \( f \), we conclude that the lower bound of (1.31) holds. \( \square \)

Proof of (2.1). We are unable to do this by direct approximation due to lack of control of moments of \( h_\varepsilon(t, x) \) higher than 2. By direct calculation \( E[h_\varepsilon(t,x)] = t/4! \) and we can take the \( \varepsilon \searrow 0 \) limit by uniform integrability that follows from the boundedness of \( \text{Var}(h_\varepsilon(t,x)) \) argued in the beginning of this section. Consequently

\[
E[h(t,x)] = t/4!.
\]

From

\[
\text{Var}(h(t,x)) = E[(h(t,x) - h(0,0) - t/4!)^2]
\]

we deduce

\[
\Delta_\delta \text{Var}(h(t,x)) = \delta^{-2} E[(\langle h(t,x), h(t,x) \rangle - h(0,0) - t/4!) (h(t,x + \delta) - h(t,x)) - (\langle h(t,x + \delta), h(t,x) \rangle - h(0,0) - t/4!) (h(t,x) - h(t,x + \delta))].
\]

Since increments are mean zero and stationary in space (part 3 of Proposition 1.1), the latter is equal to

\[
\delta^{-2} E[(h(0, \delta - x) - h(0, -x)) (h(t, \delta) - h(t, 0))].
\]

Define the “tent function” \( \varphi_\delta(x) = (\delta^{-1} - \delta^{-2} |x|)1_{|x| \leq \delta} \). We have shown that

\[
\Delta_\delta \text{Var}(h(t,x)) = \langle \tau_x \varphi_\delta, S(t) \rangle,
\]
where the angle brackets denote integration and $\tau_x \varphi_\delta(y) = \varphi_\delta(y - x)$ and we used the definition (1.30) of $S(t, dx)$. Integrating against a test function $\psi \in C_c^\infty(\mathbb{R})$ gives
\[
\int \Delta_\delta \psi(x) \text{Var}(h(t, x))dx = \langle \psi * \varphi_\delta, S(t) \rangle
\]
with $*$ denoting convolution. Let $\delta \downarrow 0$. Since $\text{Var}(h(t, x))$ is locally bounded we can take the limit on the left. In the limit we obtain
\[
\frac{1}{2} \int \partial_x^2 \psi(x) \text{Var}(h(t, x))dx = \langle \psi, S(t) \rangle.
\]
This completes the proof of Proposition 1.4.

We now complete the proof of Proposition 1.4 and Theorem 1.3 with the following

**Proposition 3.1.** For $x \in \mathbb{R}$ and $t > 0$,
\[
(3.5) \quad \text{Var}(h(t, x)) = x + 2 \int_x^\infty (z - x) S(t, dz).
\]

This proposition implies the remaining parts of Theorem 1.3 because symmetry implies that
\[
\text{Var}(h(t, x)) = |x| + 2 \int_{|x|}^\infty (z - |x|) S(t, dz).
\]
From this follow $\text{Var}(h(t, x)) - |x| \geq 0$ and the identities
\[
\text{Var}(h(t, 0)) = \int |x| S(t, dx),
\]
and for $1 < m < 3$,
\[
m(m - 1) \int |x|^{m - 2} [\text{Var}(h(t, x)) - |x|] dx = 2 \int |x|^m S(t, dx).
\]
Then we can apply the bounds from (1.31).

**Proof of Proposition 3.1.** We first prove an elementary lemma.

**Lemma 3.2.** Suppose $v$ is a continuous, symmetric function on $\mathbb{R}$ and in the sense of distributions $v'' = 2\mu$ for a symmetric probability measure $\mu$ on $\mathbb{R}$. Assume that $\int_{\mathbb{R}} |x| \mu(dx) < \infty$. Then
\[
(3.6) \quad v(x) = |x| + v(0) - \int_{\mathbb{R}} |z| \mu(dz) + 2 \int_{|x|}^\infty (z - |x|) \mu(dz).
\]

**Proof.** Suppose first that $v''/2$ is a continuous probability density. Then
\[
g'(x) = \int_x^\infty (z - x)v''(z)dz
\]
satisfies $g'' = v''$ and thereby
\[
v(x) = ax + b + \int_x^\infty (z - x)v''(z)dz
\]
for constants $a, b$. From symmetry deduce $a = 1$. Taking $x = 0$ identifies $b = v(0) - \frac{1}{2} \int_{\mathbb{R}} |z|v''(z)dz$. Now (3.6) holds for smooth $v$. Take a symmetric compactly supported smooth approximate identity $\{\phi_\delta\}_{\delta > 0}$, apply (3.6) to $\phi_\delta * v$ and let $\delta \downarrow 0$. 

\[\square\]
Continuing the proof of Proposition 3.1 apply \( v(x) = \text{Var}(h(t, x)) \) to get

\[
\text{Var}(h(t, x)) = |x| + \left( \text{Var}(h(t, 0)) - \int_{|x|}^{\infty} (z - |x|) S(t, dz) \right)
\]

(3.7)

From the Appendix we get \( \text{Var}(h(t, x)) \to 0 \). Combining this with the above gives first

(3.8) \( \text{Var}(h(t, 0)) = \int_{|x|}^{\infty} |z| S(t, dz) \)

and then

(3.9) \( \text{Var}(h(t, x)) = |x| + 2 \int_{|x|}^{\infty} (z - |x|) S(t, dz) = x + 2 \int_{|x|}^{\infty} (z - x) S(t, dz). \)

□

This completes the proof of the main results from Theorem 2.2. In a sense, this is the soft part of the argument, involving mainly passage to the limit of that estimate. We now turn to the proof of Theorem 2.2.

4. Second class particle estimate

In this section we prove the key estimate for the moment of a second class particle. The context is the weakly asymmetric simple exclusion process (WASEP) jumping to the right with rate \( p = 1/2 \) and to the left with rate \( q = 1/2 + \varepsilon^{1/2} \).

The proof is built on the method described in [4]. Throughout \( \varepsilon \in (0, 1/4) \), with the real interest being the limit \( \varepsilon \searrow 0 \). Probabilities associated to this process are denoted by \( P^{\rho}_{\varepsilon} \) when the process is stationary with Bernoulli \( \rho \) occupations. The macroscopic flux function is the expected current \( E^{\rho}_{\varepsilon} N(t, 0) \) of particles across the bond \((0,1)\) up to time \( t \) in the Bernoulli(\( \rho \)) stationary distribution and is given by

\[
H_{\varepsilon}(\rho) = -\varepsilon^{1/2}\rho(1 - \rho).
\]

The characteristic speed

(4.1) \( V_{\varepsilon}^{\rho} = H_{\varepsilon}'(\rho) = -\varepsilon^{1/2}(1 - 2\rho) \)

is the speed at which disturbances propagate in the macroscopic p.d.e. description of ASEP dynamics. Microscopically it is the average speed of the second class particle as we state below.

To define the notion of the second class particle, denote by \( P^{\rho}_{\varepsilon} \) the probability measure of the basic coupling of two processes \( \zeta^{-}(t) \leq \zeta(t) \) with this initial configuration: \( \zeta^{-}(0, 0) = 0 < 1 = \zeta(0, 0) \), and for \( x \neq 0 \), \( \zeta^{-}(0, x) = \zeta(0, x) \) have mean \( \rho \) and they are independent across the sites \( x \). Basic coupling means that the two processes obey the same Poisson clocks. The unique discrepancy between \( \zeta^{-}(t) \) and \( \zeta(t) \) is conserved during the coupled evolution and is called the second class particle. Its position is \( x(t) \), a process with nearest neighbor jumps on \( \mathbb{Z} \) and initial
position $x(0) = 0$. The mean speed of the second class particle is the characteristic speed (Corollary 2.5 in [3] or Theorem 2.1 in [4]):

$$E_x^\rho[x(t)] = V_x^\rho t.$$  

From (2.7), Theorem 2.2 is equivalent to the case $\rho = 1/2$ of the following theorem.

**Theorem 4.1.** With the same constants as in Theorem 1.3 for all $0 < \varepsilon < 1/4$, $1 \leq m < 3$, and $t \geq c_0 \varepsilon^{-2}$,

$$C_1 \varepsilon^{m/3} t^{2m/3} \leq E_x^\rho[|x(t) - V_x^\rho t|^m] \leq C_2 \varepsilon^{m/3} t^{2m/3}.$$  

The remainder of the section proves Theorem 4.1 with separate subsections for the upper and lower bound.

4.1. **Proof of the upper bound for Theorem 4.1.** We begin with a sketch of the upper bound proof. As described above, the second class particle $x(t)$ is the discrepancy between the pair of processes $(\zeta^-, \zeta)$ whose overall density is $\rho$. Additionally the pair $(\zeta^-, \zeta)$ is coupled with a third process $\eta$ at density $\lambda \in (0, \rho)$. Consequently there is a density $\rho - \lambda$ cloud of $\zeta - \eta$ second class particles. These $\zeta - \eta$ second class particles are also conserved by the evolution. Their current is the difference between the currents of the $\zeta$ and $\eta$ processes. Careful coupling makes it possible to compare $x(t)$ with the position of a tagged $\zeta - \eta$ second class particle denoted by $X_0(t)$ below.

The main estimate bounds a deviation of $x(t)$ in terms of its first moment, via several intermediate steps. A negative deviation of $x(t)$ implies the same for $X_0(t)$, which results in a deviation in the difference $J^\zeta - J^\eta$ of the currents in the $\zeta$ and $\eta$ processes. A Chebyshev bound brings in the variances of the currents $J^\zeta$ and $J^\eta$. As these currents correspond directly to the height functions $h_x$ in (2.6) and (2.7), their variances can be replaced by the first moment of $x(t)$ essentially via (2.6) and (2.7). Along the way we see that the optimal bound comes from

$$\lambda = \rho - c \varepsilon^{-1/6} t^{-1/3},$$  

de the vanishing of $H_x^\rho(\rho)$, the leading terms of the main estimate provide the correct order of magnitude in the final stages.

The main estimate alluded to above is part (i) of the first lemma, keeping in mind that $E_x^\rho[x(t)] = V_x^\rho t$.

**Lemma 4.2.** Let $B \in (0, \infty)$. There exist $C \in (0, \infty)$ and $c_1(B) \in (0, \infty)$ such that the following bounds hold for all $0 < \rho < 1$, $u \geq 1$, $0 < \varepsilon < 1/4$, and $t \geq c_1(B) \varepsilon^{-1/2}$.

(i) For $B \varepsilon^{1/3} t^{2/3} \leq u \leq 20 t/3$,

$$P_x^\rho(|x(t) - V_x^\rho t| \geq u) \leq C \varepsilon t^{2} u^{-4} E_x^\rho[|x(t) - V_x^\rho t|] + C \varepsilon t^{2} u^{-3} + \varepsilon^{-u^2/C t}.$$  

(ii) For $u \geq 20 t/3$,

$$P_x^\rho(|x(t) - V_x^\rho t| \geq u) \leq \varepsilon^{-u/C}.$$  

**Proof.** First we obtain the bounds for $P_x^\rho(x(t) \leq V_x^\rho t - u)$. By an adjustment of the constant $C$ we can assume that $u$ is a positive integer. Fix a density $0 < \rho < 1$
and let \( \lambda \in (0, \rho) \). Consider a basic coupling of three ASEP’s \( \zeta \geq \zeta^- \geq \eta \) with this initial configuration:

(a) Initially \( \{ \zeta(0, x) : x \neq 0 \} \) are i.i.d. Bernoulli(\( \rho \)) and \( \zeta(0, 0) = 1 \).

(b) Initially \( \zeta^-(0, x) = \zeta(0, x) - \delta_0(x) \).

(c) Initially \( \{ \eta(0, x) : x \neq 0 \} \) are i.i.d. Bernoulli(\( \lambda \)) and \( \eta(0, 0) = 0 \). The coupling of the initial occupations is such that \( \zeta(0, x) \geq \eta(0, x) \) for all \( x \neq 0 \).

Recall that basic coupling means that the processes share common Poisson clocks. It is important to understand that the initial distribution of the coupling is not in any way stationary for the joint process. There is the obvious point that we forced deterministic values at the origin. But even beyond that, a product measure for the pair \( (\eta, \zeta) \) would not be stationary for the coupled evolution. However, the \( \zeta \) and \( \eta \) marginals differ from stationary processes by at most one discrepancy, and this permits sharp enough computations.

Let \( x(t) \) be the position of the single second class particle between \( \zeta(t) \) and \( \zeta^-(t) \), initially at the origin. Let \( \{ X_i(t) : i \in \mathbb{Z} \} \) be the positions of the \( \zeta^- \eta \) second class particles, labeled so that initially

\[ \cdots < X_{-2}(0) < X_{-1}(0) < X_0(0) = 0 < X_1(0) < X_2(0) < \cdots \]

Let these second class particles preserve their labels in the dynamics and stay ordered. Thus the \( \zeta(t) \) configuration consists of first class particles (the \( \eta(t) \) process) and second class particles (the \( X_j(t) \)'s). Let \( \mathbf{P}_\varepsilon \) denote the joint probability distribution of these coupled processes. The marginal distribution of \( (\zeta, \zeta^-, x) \) under \( \mathbf{P}_\varepsilon \) is the same as under \( \mathbf{P}_\varepsilon^\rho \).

For \( x \in \mathbb{Z} \), \( J_x^\zeta(t) \) is the net left-to-right particle current in the \( \zeta \) process across the space-time line segment from point \((1/2, 0)\) to \((x+1, 2t)\). \( J_x^\eta(t) \) is defined similarly for the \( \eta \) process, and \( J_x^{\zeta^-\eta}(t) \) is the net current of second class particles. Current in the \( \zeta \) process is a sum of the first class particle current and the second class particle current: \( J_x^\zeta(t) = J_x^\eta(t) + J_x^{\zeta^-\eta}(t) \).

Since initially \( \zeta \geq \zeta^- \geq \eta \), it is not difficult to see that basic coupling preserves \( x(t) \in \{ X_j(t) \} \). Define the label \( m(t) \) by \( x(t) = X_{m(t)}(t) \) with initial value \( m(0) = 0 \). The label \( m(t) \) performs a walk on the labels of the \( \{ X_j \} \) with rates \( p \) to the left and \( q \) to the right, but jumps permitted only when \( X_j \) particles are adjacent. Details of how this follows from basic coupling can be found in Section 3 of [4]. Through a comparison with a reversible walk, Lemma 5.2 in [4] gives the bound

\[
\mathbf{P}_\varepsilon(m(t) \leq -k) \leq \exp\{-\varepsilon^{1/2}k\} \quad \text{for all} \ t \geq 0 \quad \text{and} \ k \geq 0.
\]

To get the first step of the estimation, note that if \( x(t) \leq V^\varepsilon t - u \) and \( m(t) > -k \), then \( X_{-k}(t) < [V^\varepsilon t] - u \). Then among the \( \zeta^- \eta \) particles only \( X_{-k+1}, \ldots, X_0 \) could have crossed from the left side of \( 1/2 \) to the right side of \( [V^\varepsilon t] - u + 1/2 \) during time \( (0, t) \). Thereby \( J^{\zeta^-\eta}_{[V^\varepsilon t]} - u(t) \leq k \), and by an appeal to (4.6) we have

\[
\mathbf{P}_\varepsilon(x(t) \leq V^\varepsilon t - u) \leq \mathbf{P}_\varepsilon(m(t) \leq -k) + \mathbf{P}_\varepsilon(J^{\zeta^-\eta}_{[V^\varepsilon t]} - u(t) \leq k) \leq \exp\{-\varepsilon^{1/2}k\} + \mathbf{P}_\varepsilon(J^{\zeta^-\eta}_{[V^\varepsilon t]} - u(t) \leq k).
\]

(4.7)

We work on the second probability on line (4.7). The first task is to center the current variables at their means. In a stationary process we know the means exactly: for example, in a stationary density \( \rho \) process \( E^\rho_t[J_x(t)] = H_t(\rho) - \rho x \) for \( x \in \mathbb{Z} \). Process \( \zeta \) can be coupled with a stationary density \( \rho \) process \( \zeta^{(\rho)} \) with at most one discrepancy. In this coupling \( |J^\zeta_x(t) - J^\rho_x^{(\rho)}(t)| \leq 1 \) and so we can use
expectations of stationary processes at the expense of small errors. Let $c_1$ below be a constant that absorbs the errors from using means of stationary processes and from ignoring integer parts. It satisfies $|c_1| \leq 3$. Then
\[
E_x J^\zeta_{[V^\rho t] - u}(t) - E_x J^n_{[V^\rho t] - u}(t) = t H_x(\rho) - (t H_x(\rho) - u) \rho - t H_x(\lambda) + (t H_x(\rho) - u) \lambda + c_1
\]
\[
= -\frac{1}{2} t H''_x(\rho)(\rho - \lambda)^2 + u(\rho - \lambda) + c_1
\]
\[
= -t \epsilon^{1/2}(\rho - \lambda)^2 + u(\rho - \lambda) + c_1.
\]
(4.8)

The computation above is quite crucial and in the end leads to the orders of magnitude we observe in the fluctuations. Its two key features are these: (i) The $H''_x(\rho)$ terms that come from the characteristic speed $V^\rho$ exactly kill the first order term from the expansion of $H_x(\rho) - H_x(\lambda)$. (ii) The nonvanishing of $H''_x(\rho)$ retains the $t \epsilon^{1/2}$ term. On the technical side, this allows us to make the mean of the current difference $J^\zeta_{[V^\rho t] - u}(t) - J^n_{[V^\rho t] - u}(t)$ positive and thereby create a deviation in (4.7).

Next the treatment splits into two cases.

Case 1. $B \epsilon^{1/3} \leq u \leq 5 \rho \epsilon^{1/2}$. Choose
\[
\lambda = \rho - \frac{1}{10} u \epsilon^{-1/2} t^{-1} \quad \text{and} \quad k = \left[\frac{1}{20} u^2 \epsilon^{-1/2} t^{-1}\right] - 3.
\]
By assuming $t \geq C(B) \epsilon^{-1/2}$ we guarantee that $u \geq 1$ and
\[
\frac{1}{20} u^2 \epsilon^{-1/2} t^{-1} \geq \frac{1}{20} B^2 \epsilon^{1/6} t^{1/3} \geq 4.
\]
Then
\[
k \geq \frac{1}{20} u^2 \epsilon^{-1/2} t^{-1} \geq 4.
\]
We continue with the second probability from line (4.7) by centering the currents. $X = X - E X$ denotes a centered random variable. In the next inequality the $-3$ in the definition (4.9) of $k$ absorbs $c_1$ from line (4.8). Choices (4.9) were made to get the most out of the inequality $J^\zeta_{[V^\rho t] - u}(t) - J^n_{[V^\rho t] - u}(t) \leq k$. Then
\[
P_x \{ J^\zeta_{[V^\rho t] - u}(t) - J^n_{[V^\rho t] - u}(t) \leq k \}
\]
\[
\leq P_x \left\{ J^\zeta_{[V^\rho t] - u}(t) - J^n_{[V^\rho t] - u}(t) \leq -\frac{1}{20} u^2 \epsilon^{-1/2} t^{-1} \right\}
\]
\[
\leq C \epsilon t u^{-4} \text{Var}_x \left[ J^\zeta_{[V^\rho t] - u}(t) - J^n_{[V^\rho t] - u}(t) \right]
\]
\[
\leq C \epsilon t u^{-4} \left( \text{Var}_x \left[ J^\zeta_{[V^\rho t] - u}(t) \right] + \text{Var}_x \left[ J^n_{[V^\rho t] - u}(t) \right] \right).
\]
(4.11)

$C$ is a constant that can change from line to line but is independent of all parameters.

We develop bounds on the variances above, first for $J^\zeta$. Utilize the coupling with a stationary density-$\rho$ process. Then apply the basic identity
\[
\text{Var}_x \left[ J_x(t) \right] = \rho(1 - \rho) E_x |x(t) - x|
\]
(4.12)
that links the variance of the current with the second class particle. (This is proved in Corollary 2.4 in [3] and in Theorem 2.1 in [4].) We find

\[
\Var_{\varepsilon}[J_{[\nu_{\varepsilon}^\rho t]} - u(t)] \leq 2 \Var_{\varepsilon}[J_{[\nu_{\varepsilon}^\rho t]} - u(t)] + 2 = 2\rho(1 - \rho)E_{\varepsilon}\left[\mathbf{x}(t) - [\nu_{\varepsilon}^\rho t] + u\right] + 2 \leq E_{\varepsilon}\left[\mathbf{x}(t) - V_{\varepsilon}^\rho t\right] + 4u.
\]

(4.13)

For the second variance on line (4.11) we begin in the same way:

\[
\Var_{\varepsilon}[J_{[\nu_{\varepsilon}^\lambda t]} - u(t)] \leq 2 \Var_{\varepsilon}[J_{[\nu_{\varepsilon}^\lambda t]} - u(t)] + 2 \leq 2\lambda(1 - \lambda)E_{\varepsilon}\left[\mathbf{x}^\lambda(t) - [\nu_{\varepsilon}^\lambda t] + u\right] + 2 \leq E_{\varepsilon}\left[\mathbf{x}^\lambda(t) - V_{\varepsilon}^\rho t\right] + 4u.
\]

(4.14)

Here we switched to a stationary density-\(\lambda\) process and introduced a second class particle \(\mathbf{x}^\lambda\) in this process. In order to get the same bound as on line (4.13) we wish to switch from \(\mathbf{x}^\lambda(t)\) to the second class particle \(\mathbf{x}(t)\) in the density-\(\rho\) process. To this end we utilize a coupling developed in Section 3 of [4]. Because the density-\(\rho\) process has higher particle density than the density-\(\lambda\) process, the second class particle in density \(\lambda\) moves on average faster in the direction of the drift. Theorem 3.1 of [4] allows us to couple \(\mathbf{x}^\lambda\) and \(\mathbf{x}\) so that \(\mathbf{x}(t) \geq \mathbf{x}^\lambda(t)\) with probability 1. Thus continuing from line (4.13),

\[
\Var_{\varepsilon}[J_{[\nu_{\varepsilon}^\lambda t]} - u(t)] \leq E_{\varepsilon}\left[\mathbf{x}(t) - \mathbf{x}^\lambda(t)\right] + E_{\varepsilon}\left[\mathbf{x}(t) - V_{\varepsilon}^\rho t\right] + 4u.
\]

(4.15)

Now \(E_{\varepsilon}\left[\mathbf{x}(t) - \mathbf{x}^\lambda(t)\right] = (V_{\varepsilon}^\rho - V_{\varepsilon}^\lambda)t = 2\varepsilon^{1/2}t(\rho - \lambda)\) and from the choice (4.9) of \(\lambda, 2\varepsilon^{1/2}t(\rho - \lambda) \leq u\); hence

\[
\Var_{\varepsilon}[J_{[\nu_{\varepsilon}^\lambda t]} - u(t)] \leq E_{\varepsilon}\left[\mathbf{x}(t) - V_{\varepsilon}^\rho t\right] + 5u.
\]

(4.16)

Insert bounds (4.13) and (4.15) into (4.11) to get

\[
P_{\varepsilon}[J_{[\nu_{\varepsilon}^\lambda t]} - u(t) \geq k] \leq C\varepsilon t^2 u^{-4}E_{\varepsilon}\left[\mathbf{x}(t) - V_{\varepsilon}^\rho t\right] + C\varepsilon t^2 u^{-3}.
\]

(4.17)

Insert (4.10) and (4.17) into line (4.7) to get

\[
P_{\varepsilon}(\mathbf{x}(t) \leq V_{\varepsilon}^\rho t - u) \leq C\varepsilon t^2 u^{-4}E_{\varepsilon}\left[\mathbf{x}(t) - V_{\varepsilon}^\rho t\right] + C\varepsilon t^2 u^{-3} + e^{-u^2/40t}
\]

and we have verified (4.24) for \(P_{\varepsilon}(\mathbf{x}(t) \leq V_{\varepsilon}^\rho t - u)\) for Case 1.

Case 2. \(u \geq 5\rho\varepsilon^{-1/2}t\).

This case is not important and we can do with crude error estimates. Let \(Z_t\) be a nearest-neighbor random walk with rates \(p = 1/2\) to the right and \(q = 1/2 + \varepsilon^{1/2}\) to the left. We have the stochastic domination \(Z_t \leq \mathbf{x}(t)\) because no matter what the environment next to \(\mathbf{x}(t)\), it has a weaker left drift than \(Z_t\). Then, since \(V_{\varepsilon}^p = -\varepsilon^{1/2}(1 - 2\rho), 2\rho\varepsilon^{1/2}t \leq 2u/5, \text{ and } \varepsilon < 1/4,\)

\[
P_{\varepsilon}\left[\mathbf{x}(t) \leq V_{\varepsilon}^\rho t - u\right] \leq P\{Z_t \leq -\varepsilon^{1/2}t - \frac{3}{5}u\}.
\]

(4.19)

For \(\alpha \in (0, 1],\) utilizing \((e^{\alpha} + e^{-\alpha})/2 \leq 1 + \alpha^2\) and \(e^{-\alpha} \geq 1 - \alpha,\)

\[
E_{\varepsilon}[e^{-\alpha Z_t}] = \exp\left(-(1 + \varepsilon^{1/2})t + t\frac{e^{\alpha} + e^{-\alpha}}{2}(1 + 2\varepsilon^{1/2}) - \varepsilon^{1/2}t e^{-\alpha}\right)
\]

\[
\leq \exp(\alpha^2 t + (\alpha + 2\alpha^2)\varepsilon^{1/2}t).
\]
We can estimate $P\{Z_t \leq -\varepsilon^{1/2}t - \frac{3}{5}u\} \leq \exp\left(-\frac{2}{5}au + 2\alpha t\right)$ and choose $\alpha = 1 \wedge \frac{3u}{20t}$ to obtain
\begin{equation}
P^\rho_\varepsilon \{x(t) \leq V^\rho t - u\} \leq \begin{cases}
\exp\left(-\frac{9}{200}u^2t^{-1}\right), & u \leq 20t/3, \\
\exp(-3u/10), & u > 20t/3.
\end{cases}
\end{equation}
Combining (4.18) and (4.20) gives Lemma 4.2 for $P^\rho_\varepsilon \{x(t) \leq V^\rho t - u\}$.

The corresponding upper tail bound $P^\rho_\varepsilon \{x(t) - V^\rho t \geq u\}$ is obtained from that for $P^\rho_\varepsilon \{x(t) - V^\rho t \leq -u\}$ by a particle-hole interchange followed by a reflection of the lattice. For details we refer to Lemma 5.3 in [4]. This completes the proof of Lemma 4.2.

**Proof of the upper bound of Theorem 4.1.** All that remains is to show that the terms that appeared in (4.14) integrate to get the upper bound in (4.3) for the moments of the second class particle. First for $m = 1$,
\[
E^\rho_\varepsilon |x(t) - V^\rho t| = \int_0^\infty P^\rho_\varepsilon \{|x(t) - V^\rho t| \geq u\} du
\leq \frac{1}{3}CtB^3E^\rho_\varepsilon |x(t) - V^\rho t| + \left(B + \frac{1}{2}CB^{-2}\right)e^{1/3}t^{2/3}
\quad + C_1(B)t^{1/3}\varepsilon^{-1/3}\exp\left(-\frac{1}{C_1(B)}e^{2/3}t^{1/3}\right) + \frac{20}{3}e^{-t/C}.
\]
$C_1(B)$ is a new constant that depends on $B$. Set $B = C^{1/3}$ to obtain
\[
E^\rho_\varepsilon |x(t) - V^\rho t| \leq \frac{2}{3}C^{1/3}t^{1/3}\varepsilon^{-1/3} + C_1t^{1/3}\varepsilon^{-1/3}\exp\left(-\frac{1}{C_1}t^{1/3}\varepsilon^{-2}\right) + \frac{20}{3}e^{-t/C}.
\]
We can fix a constant $c_0$ large enough so that, for a new constant $C$,
\begin{equation}
E^\rho_\varepsilon |x(t) - V^\rho t| \leq C\varepsilon^{1/3}t^{2/3} \quad \text{provided } t \geq c_0\varepsilon^{-2}.
\end{equation}
Restrict to those $t$ that satisfy this requirement and substitute (4.21) into Lemma 4.2. Then upon using $u \geq B\varepsilon^{1/3}t^{2/3}$ and redefining $C$ once more, we have for $B\varepsilon^{1/3}t^{2/3} \leq u \leq 20t/3$:
\begin{equation}
P^\rho_\varepsilon \{|x(t) - V^\rho t| \geq u\} \leq C\varepsilon^{2}u^{-3} + 2e^{-u^{2}/Ct}.
\end{equation}
Now take $1 < m < 3$ and use (4.22) together with Lemma 4.2 to obtain
\[
E^\rho_\varepsilon |x(t) - V^\rho t|^m = m \int_0^\infty P^\rho_\varepsilon \{|x(t) - V^\rho t| \geq u\} u^{m-1} du \leq B^m\varepsilon^{m/3}t^{2m/3}
\quad + Cm\varepsilon^{2}\int_{B\varepsilon^{1/3}t^{2/3}}^{\infty} (u^{m-4} + 2me^{-u^{2}/Ct}u^{m-1}) du + 2m\int_{20t/3}^{\infty} e^{-u/C}u^{m-1} du.
\]
This gives $E^\rho_\varepsilon |x(t) - V^\rho t|^m \leq C\varepsilon^{m/3}t^{2m/3}$ provided $t \geq c_0\varepsilon^{-2}$ for a large enough $c_0$.

### 4.2. Proof of the lower bound of Theorem 4.1

The proof of the lower bound uses similar ideas as the upper bound proof, with an extra twist. In the end we want to apply Chebyshev’s inequality again, but this time to yield a lower bound on the variance of the current. Thus the probability controlled by Chebyshev has to be in a sense a likely event.

The starting point is a coupled process pair $(\eta, \eta^\uparrow)$ of density $\lambda$, with one second class particle initially at position $n$ between them. Coupled to this pair is a process $\zeta \geq n$. Process $\zeta$ is essentially in density $\rho \in (\lambda, 1)$, except on the interval $\{0, 1, \ldots, n\}$, where $\zeta = n$ and consequently $\zeta$ also has initial density $\lambda$. 
We follow the single $\eta^+ - \eta$ second class particle $x^{(n)}(t)$ that starts at $n$. Position $n$ is chosen so that the expected position $n + V^\lambda t$ of $x^{(n)}(t)$ is sufficiently larger than $V^\rho t$ so that, by the upper bound already proved, $x^{(n)}(t) > V^\rho t$ is a likely event. Then, along the lines of the upper bound proof, from the event $x^{(n)}(t) > V^\rho t$ we derive an inequality for the current difference between the $\zeta$ and the $\eta$ processes.

Here things get more delicate than in the upper bound proof. We need to create a deviation in order to profitably apply a Chebyshev bound. The inequality derived for the current difference cannot be a deviation since, after all, it was a consequence of a likely event. However, this inequality would be a deviation if $\zeta$ started uniformly in density $\rho$. Thus we perform a change of measure where the initial distribution of $\zeta$ is replaced by the density-$\rho$ Bernoulli distribution, and we pay for this by estimating the $L^2$ norm of the Radon-Nikodym factor. Then we have a deviation which we can control, via Chebyshev, again with the variance of the current. But because the deviation originally came from a likely event, we obtain in the end a nonvanishing lower bound for the properly normalized variance of the current.

As in the upper bound proof, the key to getting the correct orders of magnitude in the end is to take $\rho - \lambda = bt^{-1/3} e^{-1/6}$.

We turn to the technical proof. By Jensen’s inequality it suffices to prove the lower bound for $m = 1$. Let $C_{UB}$ denote the constant in the upper bound statement that we just proved. We can also assume $c_0 \geq 1$. Fix a constant $b > 0$ and set
\[
a_1 = 2C_{UB} + 1 \quad \text{and} \quad a_2 = 8 + \sqrt{32b} + 8\sqrt{C_{UB}}.
\]
Increase $b$ if necessary so that
\begin{equation}
(4.23) \quad b^2 - 2a_2 \geq 1.
\end{equation}

Fix a density $\rho \in (0, 1)$ and define an auxiliary density $\lambda = \rho - bt^{-1/3} e^{-1/6}$. Define positive integers
\begin{equation}
(4.24) \quad u = \lfloor a_1 t^{2/3} e^{1/3} \rfloor \quad \text{and} \quad n = \lfloor V^\rho t \rfloor - \lfloor V^\lambda t \rfloor + u.
\end{equation}

By taking $c_0$ large enough in the statement of Theorem 4.1 we can ensure that $\lambda \in (\rho/2, \rho)$ and $u \in \mathbb{N}$.

Construct a basic coupling of three processes $\eta \leq \eta^+ \leq \zeta$ with the following initial state:

(a) Initially $\eta$ has i.i.d. Bernoulli($\lambda$) occupations $\{\eta(0, x) : x \neq n\}$ and $\eta(0, n) = 0$.

(b) Initially $\eta^+(0, x) = \eta(0, x) + \delta_n(x)$ for all $x \in \mathbb{Z}$. $x^{(n)}(t)$ is the location of the unique discrepancy between $\eta(t)$ and $\eta^+(t)$.

(c) Initially $\zeta$ has independent occupation variables, coupled with $\eta(0)$ as follows:

(c.1) $\zeta(0, x) = \eta(0, x)$ for $0 \leq x < n$ and $\zeta(0, n) = 1$.

(c.2) For $x > n$ and $x < 0$ variables $\zeta(0, x)$ are i.i.d. Bernoulli($\rho$) and $\zeta(0, x) \geq \eta(0, x)$.

Let $\mathbf{P}$ denote the probability measure of the coupled processes. Label the $\zeta - \eta$ second class particles as $\{X_m(t) : m \in \mathbb{Z}\}$ so that initially
\[
\cdots < X_{-1}(0) < 0 < X_0(0) = n = x^{(n)}(0) < X_1(0) < X_2(0) < \cdots.
\]

Again let the random label $m(t)$ satisfy $x^{(n)}(t) = X_{m(t)}(t)$, with initial value $m(0) = 0$. In basic coupling, $m(\cdot)$ jumps to the left with rate $q$ and to the right with rate $r$.
p, but only when there is an X particle adjacent to \(X_m(t)\). As in the proof of the upper bound, Lemma 5.2 in \cite{4} gives the bound

\begin{equation}
(4.25) \quad P\{m(t) \geq k\} \leq \exp\{-\varepsilon^{1/2}k\} \quad \text{for all } t \geq 0 \text{ and } k \geq 0.
\end{equation}

By the upper bound already proved and by the choice of \(a_1\),

\begin{equation}
(4.26) \quad P\{x^{(n)}(t) \leq |V^\rho t|\} = P\{x^{(n)}(t) \leq n + |V^\lambda t| - u\}
\end{equation}

\begin{equation}
\leq u^{-1}E|x^{(n)}(t)| - n - |V^\lambda t| \leq \frac{C_{UB}2^{3/3}1^{1/3}}{|a_1t^{2/3}1^{1/3}|} \leq \frac{1}{2}.
\end{equation}

This gives a lower bound for the complementary event,

\begin{equation}
(4.27) \quad \frac{1}{2} \leq P\{x^{(n)}(t) > |V^\rho t|\} \leq P\{m(t) \geq k\} + P\{J_{V^\rho t}^c(t) - J_{V^\rho t}^\rho(t) \geq -k\}.
\end{equation}

The reasoning behind the second inequality above is as follows: \(x^{(n)}(t) > |V^\rho t|\) and \(m(t) < k\) imply \(X_k(t) > |V^\rho t|\) and consequently \(-k \leq J_{V^\rho t}^\rho(t) - J_{V^\rho t}^\rho(t)\).

Put \(k = \lfloor a_2t^{1/3}1^{1/6} \rfloor - 2\). Observe from (4.25) that \(P\{m(t) \geq k\} \leq e^{-2} < 1/4\) follows from \(a_2t^{1/3}1^{1/6} \geq 2e^{-1/2} + 3\), which is guaranteed by \(t \geq c_0\varepsilon^{-2}\) and the definition of \(a_2\). Hence

\begin{equation}
(4.28) \quad \frac{1}{4} \leq P\{J_{V^\rho t}^c(t) - J_{V^\rho t}^\rho(t) \geq -a_2t^{1/3}1^{1/6} + 2\}

\leq P\{J_{V^\rho t}^\rho(t) \geq -a_2t^{1/3}1^{1/6} - t\varepsilon^{1/2}(2\rho\lambda - \lambda^2) + 1\}

+ P\{J_{V^\rho t}^\rho(t) \leq -a_2t^{1/3}1^{1/6} - t\varepsilon^{1/2}(2\rho\lambda - \lambda^2) - 1\}.
\end{equation}

Consider line (4.28). The \(\eta\) process can be coupled with a stationary \(P^\lambda\) process with at most one discrepancy. The mean current in the stationary process is

\(E\lambda[J_{V^\rho t}^\rho(t)] = tH(\lambda) - \lambda[V^\rho t] \geq tH(\lambda) - \lambda V^\rho t = -te^{1/2}(2\rho\lambda - \lambda^2)\).

Hence

\begin{equation}
\text{line (4.28)} \leq P\lambda[J_{V^\rho t}^\rho(t) \leq -a_2t^{1/3}1^{1/6} - t\varepsilon^{1/2}(2\rho\lambda - \lambda^2)]

\leq P\lambda[J_{V^\rho t}^\rho(t) \leq -a_2t^{1/3}1^{1/6}] \leq a_2^{-2}t^{-2/3}1^{-1/3} \text{Var}^\lambda[J_{V^\rho t}^\rho(t)]

\leq \frac{E\lambda|x(t)| - |V^\rho t|}{a_2^2t^{2/3}1^{1/3}} \leq \frac{E\lambda|x(t)| - V^\lambda t}{a_2^2t^{2/3}1^{1/3}} + \frac{2b}{a_2^2} + \frac{1}{a_2^3t^{2/3}1^{1/3}}

\leq C_{UB}a_2^{-2} + \frac{1}{16} + \frac{1}{16} \leq \frac{1}{8}.
\end{equation}

After Chebyshev above we applied the basic identity (4.12) for which we introduced a second class particle \(x(t)\) in a density-\(\lambda\) system under the measure \(P^\lambda\). Then we replaced \(|V^\rho t|\) with \(V^\lambda t\) and applied the upper bound and properties of \(a_2\).

Put this last bound back into line (4.28) to get

\begin{equation}
(4.30) \quad \frac{1}{8} \leq P(A), \quad \text{where } A = \{J_{V^\rho t}^\rho(t) \geq -2a_2t^{1/3}1^{1/6} - t\varepsilon^{1/2}(2\rho\lambda - \lambda^2) + 1\}.
\end{equation}

Let \(\gamma\) denote the distribution of the initial \(\zeta(0)\) configuration described by (a)–(c) in the beginning of this section. This is the initial distribution of process \(\zeta\) under the probability \(P\) in (4.30) above. Event \(A\) is not a deviation under \(P\). So in order to use a Chebyshev bound and bring in the variance of the current, we replace \(\gamma\)
with the density $\rho$ i.i.d. Bernoulli measure $\nu^\rho$. The Radon-Nikodym derivative of these two initial distributions is

$$
\frac{d\gamma}{d\nu^\rho}(\omega) = \frac{1}{\rho} \chi_{\omega_n = 1} \prod_{i=-n+1}^0 \left( \frac{\lambda}{\rho} \chi_{\omega_i = 1} + \frac{1-\lambda}{1-\rho} \chi_{\omega_i = 0} \right).
$$

Bound its second moment:

(4.31) \[ E^\rho[\left( \frac{d\gamma}{d\nu^\rho} \right)^2] = \frac{1}{\rho} \left( 1 + \frac{(\rho - \lambda)^2}{\rho(1-\rho)} \right)^n \leq \rho^{-1} e^{n(\rho-\lambda)^2/\rho(1-\rho)} \leq c_2(\rho). \]

Here condition $t \geq c_0 \varepsilon^{-2}$ implies a bound $c_2(\rho) < \infty$ independent of $t$ and $\varepsilon$. From (4.30) and Schwarz’s inequality,

$$
\frac{1}{8} \leq P(A) = \int P^\omega(A) \gamma(d\omega)
$$

(4.32) \[ = \int P^\omega(A) \frac{d\gamma}{d\nu^\rho}(\omega) \nu^\rho(d\omega) \leq c_2(\rho)^{1/2} (P^\rho(A))^{1/2}. \]

Note the stationary mean

$$
E^\rho[J_{\nu^\rho}(t)] = -t\varepsilon^{1/2}\rho^2 + \rho V^\rho t - \rho[V^\rho t] \leq -t\varepsilon^{1/2}\rho^2 + 1.
$$

Continue from line (4.32), recalling (4.23):

(64c_2(\rho))^{-1} \leq P^\rho(A) = P^\rho\{J_{\nu^\rho}(t) \geq -2a_2 t^{1/3}\varepsilon^{1/6} - t\varepsilon^{1/2}(2\rho\lambda - \lambda^2) + 1\}

\leq P^\rho\{J_{\nu^\rho}(t) \geq -2a_2 t^{1/3}\varepsilon^{1/6} + t\varepsilon^{1/2}(\rho - \lambda)^2\}

= P^\rho\{J_{\nu^\rho}(t) \geq (b^2 - 2a_2) t^{1/3}\varepsilon^{1/6}\} \leq P^\rho\{J_{\nu^\rho}(t) \geq t^{1/3}\varepsilon^{1/6}\}

\leq t^{-2/3}\varepsilon^{-1/3} \text{Var}^\rho[J_{\nu^\rho}(t)] \leq t^{-2/3}\varepsilon^{-1/3} E^\rho[|X(t) - V^\rho t|].

This completes the proof of the lower bound and thereby the proof of Theorem 4.1.

5. APPENDIX: PROPERTIES OF THE SOLUTION

Throughout this section $h(t, x) = -\log Z(t, x)$, where $Z(t, x)$ is the solution of (1.8) starting with a two-sided Brownian motion $\{B(x) : x \in \mathbb{R}\}$ with $B(0) = 0$. The goal is to prove

(5.1) \[ \text{Var}(h(t, x)) - |x| \to 0 \quad \text{as } |x| \to \infty. \]

Define the current of $u = \partial_x h$ across $x$ up to time $t$ by

$$
N(t, x) = h(t, x) - h(0, x)
$$

and the mass of $u$ in the interval $[0, x]$ at time $t$ by

$$
M(t, x) = h(t, x) - h(t, 0).
$$

**Proposition 5.1.** \[ \text{Var}(h(t, x)) - |x| = \text{Cov}(N(t, 0), N(t, x)). \]

**Proof.** Since $h(0, 0) = 0$ we have $h(t, x) = M(t, x) + N(t, 0)$. From the invariance of white noise (part 3 of Prop. 4.1), $\text{Var}(M(t, x)) = |x|$. Hence

(5.2) \[ \text{Var}(h(t, x)) - |x| = \text{Var}(N(t, 0)) + 2 \text{Cov}(M(t, x), N(t, 0)). \]

We claim that

$$
\text{Cov}(M(t, x), N(t, 0)) = -\text{Var}(N(t, 0)) + \text{Cov}(M(t, x), N(t, x)) + \text{Cov}(N(t, x), N(t, 0)).
$$

(5.3)
To see this, note that we always have the conservation law
\[ N(t,x) - N(t,0) = M(t,x) - M(0,x). \]

Hence
\[
\text{Cov}(M(t,x), N(t,0)) = \text{Cov}(M(t,x), N(t,x)) - \text{Cov}(M(t,x), (M(t,x) - M(0,x))).
\]

But Cov\((M(t,x), (M(t,x) - M(0,x))) = \frac{1}{2} \text{Var}(M(t,x) - M(0,x)). By the conservation law again, \( \frac{1}{2} \text{Var}(M(t,x) - M(0,x)) = \frac{1}{2} \text{Var}(N(t,x) - N(t,0)). \) Finally, by the translation invariance, \( \frac{1}{2} \text{Var}(N(t,x) - N(t,0)) = \text{Var}(N(t,x) - N(t,x), N(t,0)). \) This gives \((5.3)\).

From \((5.3)\) we can rewrite the right-hand side of \((5.2)\) as
\[
(5.4) \quad \text{Cov}(N(t,x), N(t,0)) + \text{Cov}(M(t,x), N(t,0) + N(t,x)).
\]

The proof is completed by noting that the second term vanishes by symmetry. To see it, note that \( \text{Var}(h(t,-x)) = \text{Var}(h(t,x)) \) and by translation invariance, \( \text{Cov}(N(t,-x), N(t,0)) = \text{Cov}(N(t,x), N(t,0)). \) Hence \( \text{Cov}(M(t,x), N(t,0) + N(t,x)) = \text{Cov}(M(t,-x), N(t,0) + N(t,-x)). \) But translating by \( x \) gives \( \text{Cov}(M(t,-x), N(t,0) + N(t,-x)) = \text{Cov}(-M(t,x), N(t,x) + N(t,0)). \)

**Proposition 5.2.** \([\lim_{|x| \to \infty} \text{Cov}(N(t,0), N(t,x)) = 0.\]

The two propositions combine to prove \((5.1)\).

The proof of Proposition \(5.2\) is based on the following lemma, for which we need to introduce some notation. Fix \( R > 0 \) and let \( W_1(t,x), W_2(t,x) \) be cylindrical Wiener processes and \( B_1(x), B_2(x) \) two-sided Brownian motions with \( B_1(0) = B_2(0) = 0 \), coupled as follows: For any \( \varphi \in C_c^\infty(\mathbb{R}) \) supported in \((-\infty,R)\), \( \langle \varphi, W_1(t) \rangle = \langle \varphi, W_2(t) \rangle \) are independent of \( \int \varphi dB_1 = \int \varphi dB_2 \), while for any \( \varphi \in C_c^\infty(\mathbb{R}) \) supported in \((R,\infty)\), \( \langle \varphi, W_1(t) \rangle, \langle \varphi, W_2(t) \rangle, \int \varphi dB_1 \) and \( \int \varphi dB_2 \) are independent. We will say that \((W_1,dB_1)\) and \((W_2,dB_2)\) are the same on \((-\infty,R)\) and independent on \((R,\infty)\).

**Lemma 5.3.** Let \( Z_i(t,x), i = 1,2, \) be the solutions of \((1.8)\) with \( W_i, i = 1,2 \) and initial data \( Z_i(0,x) = \exp\{B_i(x)\} \), where \((W_1,dB_1)\) and \((W_2,dB_2)\) are the same on \((-\infty,R)\) and independent on \((R,\infty)\). Then there is a finite \( C \) such that for \( R \geq |x| + 2t \),
\[
(5.5) \quad E[(Z_1(t,x) - Z_2(t,x))^2] \leq Ce^{-R+C(t+|x|)}.
\]

**Proof.** Let \( p(t,x) = \frac{1}{(2\pi t)^{1/2}} e^{-x^2/2t} \) be the heat kernel. We can write
\[
Z_i(t,x) = \int p(t,x-y)e^{B_i(y)} dy - \int_0^t \int p(t-s,x-y)Z_i(s,y)W_i(dsdy).
\]

First we obtain a preliminary bound on \( E[Z_i^2(t,x)] \). By Schwarz’s inequality it is bounded above by (dropping the \( i \) for clarity)
\[
2E[(\int p(t,x-y)e^{B(y)}dy)^2] + 2E[(\int \int p(t-s,x-y)Z(s,y)W(dsdy))^2].
\]

By Jensen’s inequality,
\[
E[(\int p(t,x-y)\exp\{B(y)\}dy)^2] \leq \int p(t,x-y)e^{2|y|}dy.
\]
Furthermore,
\[
E[(\int_0^t p(t-s, x-y)Z(s,y)W(dsdy))^2] = \int_0^t p^2(t-s, x-y)E[Z^2(s,y)]dsdy 
\leq \int_0^t \frac{1}{\sqrt{\pi(t-s)}} p(t-s, x-y)E[Z^2(s,y)]dsdy.
\]

Call \( g(t, x) = E[Z^2(t, x)] \) and let \( P_t \) denote the heat semigroup. \( g(0, x) = e^{2|x|} \)
and we have shown that
\[
g(t) \leq P_t g(0) + \int_0^t \frac{1}{\sqrt{\pi(t-s)}} P_{t-s} g(s)ds.
\]

Iterating once we see that this is bounded above by
\[
(1 + s\sqrt{t/\pi}) P_t g(0) + \int_0^t \frac{1}{\sqrt{\pi(t-s)}} P_{t-s} \int_0^s \frac{1}{\sqrt{\pi(s-u)}} P_{s-u} g(u)du ds.
\]
The last term can be simplified by noting that \( P_{t-s} P_{s-u} = P_{s-u} \), applying Fubini’s theorem, and using \( \int_0^s du = s \). The result is
\[
g(t) \leq (1 + s\sqrt{t/\pi}) P_t g(0) + \int_0^t P_{t-s} g(s)ds.
\]

If we let \( \tilde{g}(0) = g(0) \) and \( \tilde{g}(t) \) satisfy \( \tilde{g} - g \) with equality instead of inequality, then \( \tilde{g} - g \) satisfies \( (\tilde{g} - g)(t) \geq \int_0^t P_{t-s} (\tilde{g} - g)(s)ds \) with \( (\tilde{g} - g)(0) = 0 \). By the maximum principle for the heat equation, \( g \leq \tilde{g} \). \( \tilde{g}(t, x) \) is readily computed with the result that for some finite \( C \),
\[
E[Z^2_t(t, x)] = g(t, x) \leq Ce^{C(t+|x|)}.
\]

By \( \tilde{g}(0) \) again we have that \( f(t, x) := E[|Z_1(t, x) - Z_2(t, x)|^2] \) is bounded above by twice
\[
E\left[\left(\int_0^t p(t, x-y)\exp\{B_1(y)\} \exp\{B_2(y)\}dy\right)^2\right]
\]
\[
+ E\left[\left(\int_0^t p(t, x-y)(Z_1(s,y)W_1(dsdy) - Z_2(s,y)W_2(dsdy))\right)^2\right].
\]

Explicit computation gives that \( \text{(5.11)} \) is equal to
\[
\int_R^\infty \int_R^\infty p(t-x-y_1)p(t, x-y_2)E[(\exp{B_1(y_1)} - \exp{B_2(y_1)})(\exp{B_1(y_2)} - \exp{B_2(y_2)})]dy_1dy_2.
\]
By Schwarz’s inequality, \( \text{(5.11)} \) \( \leq \int p(t, x-y)\chi_{\{y \geq R\}}e^{2y}dy \). Another explicit computation gives that \( \text{(5.12)} \) is equal to
\[
\int_0^t \int_R^\infty p^2(t-s, x-y)f(s, y)dyds + \int_0^t \int_R^\infty p^2(t-s, x-y)E[Z^2_1(s,y)]dyds 
\leq \int_0^t \frac{1}{\sqrt{\pi(t-s)}} \int p(t-s, x-y)(f(s, y) + \chi_{\{y \geq R\}}Ce^{C(s+|y|)})dyds.
\]
Hence \( f(t) \) satisfies the same equation as \( g(t) \) in (5.13) except that this time \( f(0, x) \leq 1_{\{y \geq R\}} e^{2y} \). The same argument now shows that there is a finite \( C \) such that \( f(t, x) \leq Ce^{-R+C(t+|x|)} \) for \( R \geq |x| + 2t \).

\[ \square \]

**Proof of Proposition 5.2.** Let us use the notation \( N(t, x) \) for the normalized current \( N(t, x) - E[N(t, x)] \). First of all note that by (1.28) and 5 of Proposition 1.1 \( E[N^2(t, x)] \leq C(t) \) and does not depend on \( x \). Now let \( (W_1, dB_1), (W_2, dB_2), (W_3, dB_3) \) be coupled so that \( (W_1, dB_1) \) and \( (W_2, dB_2) \) are the same on \( (-\infty, x/2) \) and independent on \( (x/2, \infty) \), \( (W_2, dB_2) \) and \( (W_3, dB_3) \) are the same on \( (x/2, \infty) \) and independent on \( (-\infty, x/2) \), and \( (W_1, dB_1) \) and \( (W_3, dB_3) \) are independent. Let \( \bar{N}_1, \bar{N}_2, \bar{N}_3 \) be the currents corresponding to the three different pairs. Of course \( \text{Cov}(N(t, 0), N(t, x)) = E[\bar{N}_1(t, 0)\bar{N}_1(t, x)] \). By Schwarz’s inequality,

\[
|E[\bar{N}_1(t, 0)\bar{N}_1(t, x)] - E[\bar{N}_2(t, 0)\bar{N}_1(t, x)]| \leq C|E[(\bar{N}_1(t, 0) - \bar{N}_2(t, 0))^2]|^{1/2}
\]

and

\[
|E[\bar{N}_1(t, 0)\bar{N}_2(t, x)] - E[\bar{N}_1(t, 0)\bar{N}_3(t, x)]| \leq C|E[(\bar{N}_2(t, x) - \bar{N}_3(t, x))^2]|^{1/2}.
\]

By independence, \( E[\bar{N}_1(t, 0)\bar{N}_3(t, x)] = 0 \). By symmetry,

\[
E[\bar{N}_2(t, 0)\bar{N}_1(t, x)] = E[\bar{N}_1(t, 0)\bar{N}_2(t, x)].
\]

Hence

\[
\text{Cov}(N(t, 0), N(t, x)) \leq C(E[(h_1(t, 0) - h_2(t, 0))^2])^{1/2}.
\]

For each \( L > 0 \), let \( \log_L z = \log z \) for \( z \geq L^{-1} \) and \( \log_L z = -\log L \) for \( 0 < z < L^{-1} \). We have

\[
E[(h_1(t, 0) - h_2(t, 0))^2] \leq 2E[(\log_L Z_1(t, 0) - \log_L Z_2(t, 0))^2] + 4E[1_{0<Z_1< L^{-1}}(\log Z_1(t, 0))^2].
\]

(5.16)

Because \( \log_L \) is Lipschitz with constant \( L \) we have

\[
E[(\log_L Z_1(t, 0) - \log_L Z_2(t, 0))^2] \leq L^2 E[(Z_1(t, 0) - Z_2(t, 0))^2].
\]

(5.17)

By Lemma 5.3 for each fixed \( L \), this vanishes as \( |x| \to \infty \). On the other hand, since \( E[(\log Z_1(t, 0))^2] < \infty \), by the dominated convergence theorem,

\[
\lim_{L \to \infty} E[1_{0<Z_1< L^{-1}}(\log Z_1(t, 0))^2] = 0.
\]

This completes the proof of Proposition 5.2

\[ \square \]

**References**


License or copyright restrictions may apply to redistribution; see http://www.ams.org/journal-terms-of-use


