

## CONVERGENCE OF STOCHASTIC CELLULAR AUTOMATON TO BURGERS' EQUATION: FLUCTUATIONS AND STABILITY\*

Joel L. LEBOWITZ<sup>1</sup>, Enza ORLANDI<sup>2</sup> and Errico PRESUTTI<sup>2,3</sup>  
*Department of Mathematics, Rutgers University, New Brunswick, NJ 08903, USA*

Received 24 February 1988

*Dedicated to Joe Ford on the occasion of his 60th birthday*

We prove that for almost all realizations of the Boghosian-Levermore stochastic cellular automaton model the density profile converges, in the scaling limit, to the solution of Burgers' equation. The proof goes via the propagation of chaos and yields tight bounds on the fluctuations. These estimates also yield stability properties of the (smooth) shock front: at long times it remains well defined on a microscopic scale - but its location fluctuates.

### 1. Introduction

The use of cellular automata to model hydrodynamical flows of all varieties has grown explosively in the last two years from the spark started by Frisch, Hasslacher and Pomeau (FHP) in their seminal paper [1]: for earlier work see [2], for recent work [3, 4]. While the numerical results so far appear promising, at least to the devotees, the difficulties involved in proving convergence of the FHP model to the solutions of the Euler or Navier-Stokes equations appear almost as formidable as proving this for a system composed of atoms obeying Newtonian dynamics. In both cases the microscopic dynamics conserve mass, momentum and energy. The macroscopic fields corresponding to these quantities, obtained as sums of many microscopic variables, are then expected to evolve, on macroscopic space-time scales, according to suitable deterministic hydrodynamic equations [5]. More precisely, we expect the non-conserved rapidly varying variables to accommodate themselves to the "instantaneous" values of the slowly varying ones to produce states of "local equilibrium" with parameters specified by the values of the slow variables. The central mathematical problem is then to show that the correlations created by the dynamics do not take the system out from such a local equilibrium state. (This is of course true by definition for a global equilibrium state - but such a state is of no interest for time evolution.)

In order to get some insight into the questions posed by cellular automaton evolutions Boghosian and Levermore (BL) [6] invented a cellular automaton model for solving the one-dimensional Burgers'

\*Supported by NSF Grant No. DMR-86-12369. Part of this work was done while the authors were at the Institute for Theoretical Physics, University of California, Santa Barbara, California, where it was supported by NSF Grant No. Phy-82-17853 supplemented by funds from NASA.

<sup>1</sup>Also at Department of Physics.

<sup>2</sup>Permanent Address: Dipartimento di Matematica, Università di Roma "La Sapienza", P.le A.Moro 2, 00187 Roma, Italy.  
Partially supported by CNR PS-AITM grant.

<sup>3</sup>Present Address: Department of Mathematics, University of Colorado, Boulder, CO 80309, USA.

equation,

$$\frac{\partial}{\partial t} \rho(r, t) + c \frac{\partial}{\partial r} \left( \rho - \frac{\rho^2}{2} \right) = \nu \frac{\partial^2 \rho}{\partial r^2}. \quad (1.1)$$

Here  $c > 0$  is a velocity and  $\nu$  a viscosity. For  $\nu = 0$  (1.1) can lead to the formation of sharp shocks from smooth initial data. These are density discontinuities, low density on the left high density on the right, which travel at a fixed speed. For  $\nu \neq 0$  there are still such traveling fronts but the profile is now smooth having a thickness of  $\mathcal{O}(\nu/c)$ . We shall continue to refer to these fronts as shocks.

Boghosian and Levermore argued convincingly and demonstrated numerically that the density profile of their model does indeed follow the solution of Burgers' equation over a macroscopic time period of order unity. Their main difficulty in actually proving this involved controlling the correlations which may build up in the microscopic system – a problem which, as already mentioned, is central to all models of this kind.

In this paper we solve this problem for the BL model and prove very strong convergence to Burgers' equation. We do this by proving the propagation of chaos, i.e. we show that correlations do not in fact build up for fixed macroscopic times. We also discuss some results about the behavior of the system for “long” times when the main effect of the microscopic fluctuations is to shift in a random way the location of the shock profile – the shape remaining sharp on the microscopic level. This is consistent with stability properties of Burgers' equation.

Our analysis relies heavily on the results of DeMasi et al. [7] who obtained Burgers' equation for the density profile of particles on a lattice evolving via continuous time stochastic dynamics – the weakly asymmetric exclusion process.

We also mention here that the BL model generalizes in a natural way to higher dimensions and to cases where  $c$  depends on position leading, in  $d$ -dimensions, to the macroscopic equation

$$\frac{\partial \rho(r, t)}{\partial t} + \nabla \cdot [c(r) \rho (1 - \rho/2d)] = \nu \nabla^2 \rho. \quad (1.2)$$

## 2. Description of model and statement of results

Our model is the same as that investigated by BL [6]. We consider a system of particles on a one-dimensional lattice, each one with “velocity”  $\sigma = \pm 1$ . There is an exclusion rule that two particles at the same site cannot have the same velocity. Therefore each site  $q \in \mathbb{Z}$  may have at most two particles (with different velocities). Each unit time updating of the automaton consists of two steps, the first is stochastic the second one deterministic:

*Step 1 – Velocity flips.* This updating rule acts independently at each site  $q$ . If there are two or no particles at  $q$  then nothing happens. If, on the other hand, there is one particle at  $q$  with velocity  $\sigma(q)$  then its new velocity  $\sigma'(q)$  takes the value  $\pm 1[-1]$  with probability  $p = \frac{1}{2} + \epsilon[\frac{1}{2} - \epsilon]$ ,  $\epsilon > 0$ , i.e. the  $\sigma'(q)$  are independent random variables with mean  $2\epsilon$ .

*Step 2 – Advection.* Every particle at site  $q$  moves to  $q + \sigma'(q)$  keeping its velocity.

Notice that there is an effective interaction, "collisions", among the particles. The fact that a particle cannot flip its velocity if another one is present at the same site causes nontrivial space-time correlations in the process.

*Notation.* Let  $\mathcal{X}$  be the set of all particles configurations  $\eta$ ,  $\eta = \{\eta(q, \sigma); \eta(q, \sigma) \in \{0, 1\}, q \in \mathbb{Z}, \sigma(q) \in \{1, -1\}\}$ . Let  $P_\epsilon$  be the transition probability of the automaton described above from its state at time  $t$ ,  $t \in \mathbb{N}$ , an integer, to its state at  $t + 1$ ,

$$P_\epsilon(\eta(q, \sigma; t + 1) = 1 | \eta(t) = \eta) = \frac{1}{2} \{ \eta(q - \sigma, \sigma) + \eta(q - \sigma, -\sigma) \} + \epsilon \sigma \{ \eta(q - \sigma, \sigma) + \eta(q - \sigma, -\sigma) - 2\eta(q - \sigma, \sigma)\eta(q - \sigma, -\sigma) \}, \tag{2.1}$$

$$P_\epsilon(\eta(q, -1; t + 1) = 1, \eta(q + 2, 1; t + 1) = 1 | \eta(t) = \eta) = \eta(q + 1, 1) \cdot \eta(q + 1, -1). \tag{2.2}$$

We call the above process the weakly asymmetric automaton process (WAAP);  $E_\mu^\epsilon$  denotes the expectation with respect to the WAAP with initial measure  $\mu$ . If  $\epsilon = 0$  we call the process the symmetric automaton process (SAP) and write  $P$  and  $E_\mu$  for the above quantities.

We finally define  $\eta(q, t)$  to be the number of particles at  $(q, t)$ , i.e.

$$\eta(q, t) = \eta(q, 1; t) + \eta(q, -1; t) \tag{2.3}$$

and we set  $\eta(q, 0) = \eta(q)$ .

Our first result concerns the convergence to Burgers' equation and propagation of chaos for the WAAP.

*Theorem 2.1.* For  $\epsilon > 0$  let  $\mu_\epsilon$  be any probability measure on  $\mathcal{X}$  such that for all  $n \geq 1$ , for all distinct  $q_1, \dots, q_n$  and all  $\sigma_1, \dots, \sigma_n$

$$E_{\mu_\epsilon} \left[ \prod_{i=1, n} \eta(q_i, \sigma_i) \right] = \prod_{i=1, n} [\frac{1}{2} \rho(\epsilon q_i)], \tag{2.4}$$

where  $\rho(r)$ ,  $r \in \mathbb{R}$ , is a smooth function with values in  $[0, 2]$ . Then for any  $n \geq 1$

$$\lim_{\epsilon \rightarrow 0} \sup_{q_1, \sigma_1, \dots, q_n, \sigma_n} \left| E_{\mu_\epsilon}^\epsilon \left[ \prod_{i=1, n} \eta(q_i, \sigma_i; [\epsilon^{-2}t]) \right] - \prod_{i=1, n} [\frac{1}{2} \rho(\epsilon q_i, t)] \right| = 0, \tag{2.5}$$

where  $[\epsilon^{-2}t]$  is the largest integer  $\leq \epsilon^{-2}t$ .

The sup in (2.5) is over all distinct  $\{(q_1, \sigma_1), \dots, (q_n, \sigma_n)\}$  while  $\rho(r, t)$  solves Burgers' equation (1.1) with  $c = 2$ ,  $\nu = \frac{1}{2}$ , i.e.

$$\frac{\partial}{\partial t} \rho + 2 \frac{\partial}{\partial r} \left( \rho \left( 1 - \frac{\rho}{2} \right) \right) = \frac{1}{2} \frac{\partial^2}{\partial r^2} \rho, \tag{2.6a}$$

$$\rho(r, 0) = \rho(r). \tag{2.6b}$$

In Theorem 2.1 Burgers' equation is derived as the limiting equation for the ensemble average occupation numbers. In real systems, as well as in computer simulations, we are interested in what

happens in the system in a single run, see [6]. With this in mind we set for any  $\alpha > 0$ ,  $t \in \mathbb{N}$ ,  $q \in \mathbb{Z}$

$$M(q, t; \epsilon^{-\alpha}) = \epsilon^\alpha \sum_{|q' - q| \leq \frac{1}{2}\epsilon^{-\alpha}} \eta(q'; t) \quad (2.7)$$

the actual density of particles at time  $t$ , in the interval having length  $\epsilon^{-\alpha}$  centered around  $q$ . We shall choose  $\alpha < 1$  so that on a macroscopic scale where the unit of length is  $\epsilon^{-1}$ , the interval is infinitesimal.  $M(q, t; \epsilon^{-\alpha})$  is of course a random variable, different for each realization. The next theorem states our result about single realizations of the automaton system.

*Theorem 2.2.* Let  $\rho$  and  $\mu_\epsilon$  be as in theorem 2.1. Then for any  $\alpha$  and  $\xi$ , such that  $0 < \alpha < 1$  and  $0 < \xi < \alpha/2$ ,  $\xi < 1 - \alpha$ , and any positive  $T$  and  $R$

$$\lim_{\epsilon \rightarrow 0} P_{\mu_\epsilon}^\epsilon \left( \sup_{0 \leq t \leq \epsilon^{-2}T} \sup_{|q| \leq \epsilon^{-1}R} |M(q, t; \epsilon^{-\alpha}) - \rho(\epsilon q, \epsilon^2 t)| > \epsilon^\xi \right) = 0. \quad (2.8)$$

The convergence in (2.8) is faster than any power of  $\epsilon$ .

As mentioned in the introduction the study of fluctuations is important for establishing the stability properties of the evolution. We introduce the density fluctuation field  $X_t^\epsilon(\varphi)$  as

$$X_t^\epsilon(\varphi) = \sqrt{\epsilon} \sum_{q \in \mathbb{Z}} \varphi(\epsilon q) [\eta(q, \epsilon^{-2}t) - \rho(\epsilon q, t)]. \quad (2.9)$$

Here  $\varphi \in \mathcal{S}(\mathbb{R})$ , the Schwartz space of  $C^\infty$  real functions with "fast" decay at  $\infty$ , such that any derivative is uniformly bounded. One should think of  $\varphi$  as a smoothed characteristic function of an interval, say for instance the interval  $[0, 1]$ . Then  $X_t^\epsilon(\varphi)$  would be the total number of particles in  $[0, \epsilon^{-1}]$  minus its asymptotic value (as  $\epsilon \rightarrow 0$ ) divided by the square root of the volume. Hence  $X_t^\epsilon(\varphi)$  is just the usual density fluctuation random variable in that interval. The next theorem states that the density fluctuation field converges to a Gaussian distribution:

*Theorem 2.3.* Let  $\mu_\epsilon$  and  $\rho$  be as in theorem 2.1. Let  $P_\epsilon$  be the law of the process generated by all density fluctuation fields defined as in (2.9). Then  $P_\epsilon$  converges to  $P$  on finite distributions, where  $P$  is Gaussian with mean zero and covariance kernel  $C^*$ ,

$$\lim_{\epsilon \rightarrow 0} E_{\mu_\epsilon}^\epsilon (X_t^\epsilon(\varphi) X_{t+s}^\epsilon(\psi)) = \iint dr dr' \varphi(r) \psi(r') C_{t,s}^*(r, r'). \quad (2.10)$$

$C^*$  satisfies the equation

$$\frac{\partial}{\partial s} C_{t,t+s}^*(r, r') = \frac{1}{2} \frac{\partial^2}{\partial r^2} C_{t,t+s}^*(r, r') + 2 \frac{\partial}{\partial r} [1 - \rho(r, t+s)] C_{t,t+s}^*(r, r'),$$

with the initial condition at  $s = 0$  given by

$$C_{t,t}^*(r, r') = C_t(r, r') + \delta(r - r') \left[ \rho(r, t) \left( 1 - \frac{\rho(r, t)}{2} \right) \right],$$

where  $C_t(r, r')$  is the solution of the equation

$$\begin{aligned} \frac{\partial}{\partial t} C_t &= \frac{1}{2} \frac{\partial^2}{\partial r^2} C_t + 2 \frac{\partial}{\partial r} [1 - \rho(r, t)] C_t(r, r') + \frac{1}{2} \frac{\partial^2}{\partial r'^2} C_t + 2 \frac{\partial}{\partial r'} [1 - \rho(r', t)] C_t(r, r') \\ &+ \delta(r - r') \left\{ \rho(r, t) \left( 1 - \frac{\rho(r, t)}{2} \right) \frac{\partial}{\partial r} \rho(r, t) - \frac{1}{4} \left( \frac{\partial}{\partial r} \rho(r, t) \right)^2 \right\}, \end{aligned} \tag{2.11a}$$

with

$$C_0(r, r') = 0. \tag{2.11b}$$

*Remark.* Actually a statement sharper than the one given in theorem 2.3 holds. Let  $D(\mathbb{R}_+ \rightarrow S'(\mathbb{R}))$  be the space of functions from  $\mathbb{R}_+$  to the distribution space  $S'(\mathbb{R})$  which are right continuous with left limit, endowed with the Schorohod topology [8]. Let  $P^\epsilon$  be the law on  $D(\mathbb{R}_+ \rightarrow S'(\mathbb{R}))$  induced via (2.9) by the WAAP with initial measure  $\mu_\epsilon$ . Then  $P^\epsilon$  converges weakly to the Gaussian law described in theorem 2.3.

### 3. Outline of the proofs

We shall prove the theorems stated in section 2 using correlation functions techniques, as in [7]. Let  $x = (q, \sigma)$ ,  $q \in \mathbb{Z}$ ,  $\sigma = \pm 1$  and for each  $n \geq 1$  set

$$x = (x_1, \dots, x_n) \equiv (q, \sigma),$$

where  $x_i \neq x_j$  if  $i \neq j$ . The  $n$ -body correlation function is

$$u_n^\epsilon(x, t) = E_{\mu_\epsilon}^\epsilon \left( \prod_{i=1}^n \eta(x_i, t) \right), \tag{3.1}$$

$E_{\mu_\epsilon}^\epsilon$  being the expectation w.r.t.  $P_{\mu_\epsilon}^\epsilon$ , the law of the WAAP with initial measure  $\mu_\epsilon$ . Using the evolution rules for the automaton described in the preceding section we can easily write an equation for the time increments of the correlation functions:

$$u_n^\epsilon(x, t) = \sum_{k=n}^{2n} \sum_y c(k, y, \epsilon; x) u_k^\epsilon(y, t-1). \tag{3.2}$$

The coefficients in (3.2) satisfy the bound

$$\sum_y |c(k, y, \epsilon; x)| \leq c_k \epsilon \quad \text{if } k > n, \tag{3.3}$$

where the  $c_k$  are suitable constants.

This bound reflects the fact that when  $\epsilon = 0$ , i.e. in the symmetric case, the right-hand side of (3.2) contains only  $n$ -body correlations, a rather surprising fact at first sight, since there is a real interaction

among particles. More precisely (3.2) becomes

$$u_n^0(x, t) = \sum_y Q(x \rightarrow y) u_n^0(y, t-1), \quad (3.4)$$

$u_n^0$  being the correlation functions for the SAP and  $Q$  a probability transition matrix which describes the backward evolution of the  $n$ -particle symmetric automaton.  $Q$  is analysed in more detail in the appendix.

To go from the SAP to the WAAP one might try to view (3.2) as a perturbation of (3.4). The difficulty is that the bound (3.3) is only linear in  $\epsilon$ . As a consequence the perturbation series can only converge for times of the order of  $\epsilon^{-1}$ . Such times are infinitesimal on a macroscopic scale, where times are scaled by  $\epsilon^{-2}$ , hence the expansion does not allow to go past time 0, macroscopically. The problem comes from the fact that since we are looking for a derivative in (3.2), cf. (2.6), we cannot afford to neglect the signs of the coefficients, namely we need to do better than bounding their absolute values, as done in (3.3).

Let us assume, for the moment, that we can in fact extract a lattice derivative from the expression

$$\sum_{k=n}^{2n} \sum_y c(k, y, \epsilon; x) u_k^0(y, t-1) - \sum_y Q(x \rightarrow y) u_n^0(y, t-1) \quad (3.5)$$

and that the powers of  $Q$  are well approximated by random walk transition probabilities. (This is trivially so if  $n=1$ , otherwise some estimates are needed to control the effect of the exclusion interaction, we do this in the appendix.) This gives (after some work) an estimate on (3.5) of the form

$$\sum_k c^k \sum_{t > s_1 > \dots > s_k > 0} \epsilon^k [t - s_1]^{-1/2} \dots [s_{k-1} - s_k]^{-1/2} n(n+1) \dots (n+k). \quad (3.6)$$

where  $c$  is a suitable constant. The last factor takes into account the combinatorics of the expansion and  $[s_i - s_{i+1}]^{-1/2}$  estimates the difference between the probabilities that a symmetric random walk goes to nearest neighbor sites after a time  $s_i - s_{i+1}$ .

Unfortunately the estimate (3.6) is only sufficient for proving convergence at times of the order  $\epsilon^{-2+\beta}$  with  $\beta > 0$ . For any such time in fact the terms in (3.6) with  $k \geq 1$  are infinitesimal w.r.t.  $\epsilon$ , so that by stopping the series after a suitable number of terms, we get an expression for the correlation functions which is as accurate as desired. To reach later times we cannot simply iterate the above procedure, we would inevitably run into convergence problems once macroscopic times are reached. To break the growth of the correlation functions we have to fix, i.e. condition on, the value of the particle configuration at the times  $k\epsilon^{-2+\beta}$ , the end steps of our iterative procedure. By choosing  $\beta$  suitably small the iteration then works for times of order  $\epsilon^{-2}$ .

To realize this program we need to prove estimates like (3.6) for a process which starts from a single particle configuration and not from the nice smooth measure  $\mu_\epsilon$ , which we assumed describes the initial state of the system. Secondly we need to prove that the configurations on which we are conditioning are suitably close to the macroscopic density profile which solves the limiting Burgers' equation. In the next section we shall see that closeness may be expressed by the condition that for all relevant  $q'$

$$\epsilon^\alpha \sum_{|q-q'| < \frac{1}{2}\epsilon^{-\alpha}} (\eta(q, t) - \rho(\epsilon q, \epsilon^2 t)) \quad (3.7)$$

is suitably small. Here  $\eta(q, t)$  denotes the number of particles at  $q \in \mathbb{Z}$  at time  $t$ ,  $t$  is one of the times

$k\epsilon^{-2+\beta}$ , while  $\rho(r, t)$  is the solution of (2.6). It will be convenient in the actual estimates to replace  $\rho$  in (3.7) by the solution of a discretized version of (2.6).

The probability that (3.7) is larger than some quantity can be bounded using the Chebishev inequality. This states that if  $Y$  is a random variable with distribution  $P$ , then for any  $\delta > 0$  and  $n \geq 1$

$$P(|Y| > \delta) \leq \delta^{-2n} E(Y^{2n}),$$

$E$  denoting the expectation w.r.t.  $P$ . Hence

$$P_{\mu_\epsilon}^\epsilon \left[ \left| \epsilon^\alpha \sum_{|q-q'| < \frac{1}{2}\epsilon^{-\alpha}} (\eta(q, t) - \rho(\epsilon q, \epsilon^2 t)) \right| > \epsilon^\xi \right] \leq \epsilon^{-2n\xi} \sum_{q_1, \dots, q_{2n}; |q_i - q_j| \leq \frac{1}{2}\epsilon^{-\alpha}} \epsilon^{2n\alpha} E_{\mu_\epsilon}^\epsilon \left[ \prod_{i=1}^{2n} (\eta(q_i, t) - \rho(\epsilon q_i, \epsilon^2 t)) \right]. \tag{3.8}$$

This motivates the introduction of the following quantities which are slight, technically convenient, modifications of the expression appearing in the right-hand side of (3.8):

$$v_n^\epsilon(x, t|\eta) = E_\eta^\epsilon \left[ \prod_{i=1}^n (\eta(x_i, t) - \rho_\epsilon(x_i, t|\eta)) \right], \tag{3.9a}$$

$$v_n^\epsilon(x, t|\mu_\epsilon) = E_{\mu_\epsilon}^\epsilon \left[ \prod_{i=1}^n (\eta(x_i, t) - \rho_\epsilon(x_i, t|\mu_\epsilon)) \right], \tag{3.9b}$$

where  $\eta$  is an arbitrary configuration while  $\mu_\epsilon$  is the initial measure defined in theorem 2.1. Finally  $\rho_\epsilon(x, t|\eta)$  is the solution of the following equation:

$$\rho_\epsilon(q, \sigma, t+1|\eta) = \left[ \frac{1}{2} + \epsilon\sigma \right] \sum_{\tau=\pm 1} \rho_\epsilon(q-\sigma, \tau, t|\eta) - 2\epsilon\sigma \prod_{\tau=\pm 1} \rho_\epsilon(q-\sigma, \tau, t|\eta), \tag{3.10}$$

with initial condition  $\rho_\epsilon(x, 0) = \eta(x)$ ; while  $\rho_\epsilon(x, t|\mu_\epsilon)$  satisfies the same equation with initial condition  $\rho_\epsilon(x, 0) = E_{\mu_\epsilon}^\epsilon(\eta(x))$ . Notice that eq. (3.10) is obtained from (2.2) by assuming that the correlation functions factorize.

The main estimate, from which all other results follow more or less directly, is stated in the following proposition:

*Proposition 3.1.* For any  $n \geq 1, T > 0$ , there is a constant  $c$  such that for any  $\eta$  and any  $t \leq \epsilon^{-2}T$

$$\sup_x |v_n^\epsilon(x, t|\eta)| \leq ct^{-n/8}, \tag{3.11a}$$

while if  $\mu_\epsilon$  is as in theorem 2.1, then

$$\sup_x |v_n^\epsilon(x, t|\mu_\epsilon)| \leq c\epsilon^{n/2}. \tag{3.11b}$$

*Proof of theorem 2.1.* As proven in [6] the theorem is a consequence of the factorization of the correlation functions, which we have from proposition 3.1. One of the main results in [6] is that  $\rho_\epsilon(\cdot, \cdot|\mu_\epsilon)$  actually approximates the solution to (2.6) in the limit  $\epsilon \rightarrow 0$ .

*Proof of theorem 2.2.* We shall use the Chebishev inequality to estimate the probability in (2.8), but first we need to rewrite (2.8) in such a way that after having used the Chebishev inequality, we have an expression containing  $v$ -functions, so that we can use proposition 3.1. To do this we note that we can rewrite the difference in (2.8) as the expression in (3.7) with an error which goes like  $\epsilon^{1-\alpha}$ . We then use the following estimate, proven in [6]:

$$|2\rho_\epsilon(q, \sigma, t | \mu_\epsilon) - \rho(\epsilon q, \epsilon^2 t)| \leq c\epsilon$$

uniformly in  $\epsilon, q, t \leq \epsilon^{-2}T$ , where  $c$  is a suitable constant. Such a bound allows us to replace  $\rho$  by  $\rho_\epsilon$  in (3.7). The error goes like  $\epsilon$ . Since by assumption  $\xi < 1 - \alpha$ , the event in (2.8) is bounded by  $\epsilon^{-2}T\epsilon^{-1}R$  times an expression similar to that on the right-hand side of (3.8) with  $\rho_\epsilon$  in place of  $\rho$ . Such an expression in turn is bounded by the maximum of  $\{c\epsilon^{\alpha n}\epsilon^{-\xi 2n}, c\epsilon^n\epsilon^{-\xi 2n}\}$ , where as usual  $c$  denotes a suitable constant independent of  $\epsilon$ . The first term bounds the contribution coming from the sum over  $q_1, \dots, q_{2n}$  when they are pairwise equal, the second term deals with the case where all  $q_i$  are mutually distinct. We have used (3.11b) for estimating these last terms, while the former just counts how many terms are present, each giving a bounded contribution, because all occupation variables are bounded by 1. It is easy to see that the intermediate cases are also bounded by the same quantity. Since  $\xi < \alpha/2$ , we can, by choosing  $n$  large enough, make this term arbitrarily small when  $\epsilon \rightarrow 0$ , so that it still vanishes when multiplied by  $\epsilon^{-3}$ . In this way the theorem is proven; we shall omit the details.

The proof of theorem 2.3 is slightly more delicate, it requires a further analysis of the expression in (3.2). On the other hand it can be reduced, by using abstract arguments, to the proof that for any positive  $t$ ,  $|\epsilon^{-1}v_2^\xi(x, \epsilon^{-2}t | \mu_\epsilon)| \leq c$  where  $c$  does not depend on  $x$ . This follows from (3.11b) and so the theorem is proven. We refer to [7] for more details.

We conclude this section with some comments on the relation between the BL automaton and the simple exclusion process. The two systems share many common features and their differences disappear in the hydrodynamical limit, at least in the weakly asymmetric case; it does not seem, however, that they are truly isomorphic, at least we have not succeeded in establishing such a relation. To compare their collective behavior we should remember that the BL automaton may have twice as many particles per site as the exclusion process, hence when going from the automaton to the exclusion we should divide the density by 2. Furthermore in order to compare the WAAP to the weakly asymmetric simple exclusion considered in [7] we need to change the flip rates. In our case in fact the drift of a single particle is  $2\epsilon$  while in [7] it is only  $\epsilon$ . Therefore if we set the probability that the velocity flips to  $\pm 1$  equal to  $(1 \pm \epsilon)/2$ , we then get (1.1) instead of (2.6) with  $c = 1$  and  $v = \frac{1}{2}$ . Then  $\bar{\rho} = \rho/2$  solves the same Burgers' equation as the limiting density profile in [7].

#### 4. Proofs

We shall first write an equation for the time increments of the  $v$ -functions (unless otherwise specified  $v_n^\xi$  stands both for (3.9a) and (3.9b)). The equation will have a structure similar but more complex than for the usual  $u$ -correlation functions, cf. (3.2). It is convenient to interpret as number of particles the degree of a  $v$ -correlation function, its argument specifying locations and velocities of the particles. In such language the equation for the  $v$ -correlation functions describes a birth-death process with displacements. Besides



the basic *unperturbed* motion determined by  $Q$ , the same as that in eq. (3.4), we have more displacements, births (already appearing in eq. (3.2)) and deaths.

This last is a novelty with respect to (3.2) where the number of bodies could not decrease. This is not surprising since the  $v$ -functions are linear combinations of  $u$ -functions with equal or smaller degree. Naively one might think that this should not worry us at all, since the iteration becomes simpler when there are less particles. This is indeed so, but remember that we are trying to prove an upper bound for the  $v$ -functions which decreases when the number of particles increases, see (3.9). The real problem is in fact to recognize a suitably small factor for each particle which dies. In this way we shall deal with deaths. The problems coming from the births (convergence of the corresponding perturbative series) is dealt with by studying the process only for a time interval  $\epsilon^{-2+\beta}$  and then iterating, in the way explained in the preceding section. This second part is very similar to what is done in [7] so we shall only outline the main ideas. The first part is also pretty similar to [7], but with a somewhat more complex algebra. In particular we have to treat separately different type of configurations (including velocities) occurring in  $v_n^\epsilon(x, t)$ . We shall consider in detail the new problems while just outlining what can be reduced to [7]. We begin with estimates for  $t \leq \epsilon^{-2+\beta}$  for free configurations.

We have already mentioned that in the equation for the  $v$ -functions particles may die. This does not happen for *free* configurations: we give some definitions and then state a lemma to clarify this point.

*Definition.* We shall say that a configuration  $x = (x_1, \dots, x_n)$ ,  $x_i = (q_i, \sigma_i)$ ,  $i = 1, \dots, n$ , is *free* if  $q_i - q_j \neq 2$  for all  $i \neq j$ .

*Definition.*  $T$  is the transition probability in the configuration space of  $n$  particles which describes the process of randomization of the velocities defined in step 1 of section 2, for  $\epsilon = 0$ . Namely each particle changes or keeps its velocity with equal probability if no other particle is at the same site, if there is another particle then no change occurs.  $S_\pm$  is the advection step in the evolution, namely each particle moves by  $\pm\sigma$ ,  $\sigma$  being its velocity. With the above notation  $Q$ , as defined in section 3, equals  $TS_-$ ,  $P = S_+T$  is the forward evolution,  $R = TS_\pm T$  is the self-adjoint evolution operator. Notice that  $TS_+T = TS_-T$ ,  $T^2 = T$ ,  $T$  is self-adjoint while  $(S_+)^{\text{adj.}} = S_-$  and that  $Q^n T = R^n$ . Note also that by its definition

$$\sum_{x'} Q(x \rightarrow x') f(x') = \sum_{x', x''} S_-(x \rightarrow x'') T(x'' \rightarrow x') f(x')$$

so that in matrix notation  $\langle x | Q | x' \rangle = \langle x | S_- T | x' \rangle$  or  $Q = S_- T$ . Finally applying  $T(x \rightarrow x')$  to a function  $f(x') = f(q', \sigma')$  corresponds to replacing  $q'$  by  $q$ , summing over the velocities and dividing by a normalization constant.

*Lemma 4.1.* Let  $x$  be a free configuration of  $n$  particles and let  $t$  be an integer,  $t \geq 0$ . Then there exist coefficients  $a(k, y, \epsilon, x, t)$ ,  $k = n, \dots, 2n$ , such that

$$\sum_{x'} T(x \rightarrow x') v_n^\epsilon(x', t+1) = \sum_y R(x \rightarrow y) v_n^\epsilon(y, t) + \sum_{k=n}^{2n} \sum_y a(k, y, \epsilon, x, t) v_k^\epsilon(y, t). \tag{4.1}$$

The coefficients  $a$  are equal to 0 unless the position of each particle in  $y$  is nearest neighbor to some

particle in  $x$ . Furthermore for all  $k$

$$\sum_y |a(k, y, \epsilon, x, t)| \leq c\epsilon, \quad (4.2)$$

where  $c$  is a suitable constant which only depends on  $n$ . Finally for any  $k$  as above it is possible to partition the set of all  $y$  into pairs  $y', y''$  so that 1)  $y'$  differs from  $y''$  by displacing a particle by two sites and 2) there is a constant  $c$ , which only depends on  $n$ , such that

$$|a(k, y', \epsilon, x, t) + a(k, y'', \epsilon, x, t)| \leq c\epsilon t^{-1/2}. \quad (4.3)$$

*Proof.* Let us first consider the case  $n = 1$ . We have

$$u_1^\epsilon(q, \sigma, t+1) = \left(\frac{1}{2} + \epsilon\sigma\right) \sum_{\tau=\pm 1} u_1^\epsilon(q-\sigma, \tau, t) - 2\epsilon\sigma u_2^\epsilon(q-\sigma, 1, q-\sigma, -1, t)$$

from which we subtract  $\rho_\epsilon(q, \sigma, t+1)$ , as given in (3.10). On the left we then have  $v_1^\epsilon(q, \sigma, t+1)$  and on the right the linear terms reconstruct a  $v$ -function. On the other hand setting  $q' = q - \sigma$  we get

$$u_2^\epsilon(q', 1, q', -1, t) = v_2^\epsilon(q', 1, q', -1, t) + \sum_{\tau=\pm 1} v_1^\epsilon(q', \tau, t) \rho_\epsilon(q' - \tau, t) + \prod_{\tau=\pm 1} \rho_\epsilon(q, \tau, t),$$

so that

$$\begin{aligned} v_1^\epsilon(q, \sigma, t+1) &= \left(\frac{1}{2} + \epsilon\sigma\right) \sum_{\tau=\pm 1} v_1^\epsilon(q', \tau, t) - 2\epsilon\sigma v_2^\epsilon(q', 1, q', -1, t) \\ &\quad - 2\epsilon\sigma \sum_{\tau=\pm 1} v_1^\epsilon(q', \tau, t) \rho_\epsilon(q', -\tau, t). \end{aligned} \quad (4.4)$$

This proves (4.1) and (4.2) in the case  $n = 1$ . The same argument easily extends to  $n > 1$  when no two particles in  $x$  are at the same site (at  $t + 1$ ). We shall consider later the case when there are two particles at the same site.

Notice that in (4.4)  $\sigma$  multiplies  $\epsilon$ . Therefore in the equation for  $Tv_1^\epsilon$ , obtained by summing (4.4) over  $\sigma$  and then dividing by the normalization factor [2 in this case], the first order terms in  $\epsilon$  have opposite signs. The arguments of the corresponding  $v^\epsilon$  and  $\rho_\epsilon$  functions are different, the sites are displaced by 2. Then (4.3) is a consequence of the following estimate valid for all  $q$  and  $\sigma$ :

$$|\rho_\epsilon(q, \sigma, t) - \rho_\epsilon(q-2, \sigma, t)| \leq c\epsilon t^{-1/2}, \quad (4.5)$$

which can be derived from an analysis of (3.10). Actually for  $\rho_\epsilon(\cdot, \cdot, \cdot | \mu_\epsilon)$  the factor  $t^{-1/2}$  on the right-hand side of (4.5) can be set equal to  $\epsilon$ . The same argument holds for  $n > 1$  proving (4.3) and therefore the lemma for free configurations with at most one particle per site.

The new case is when there are two particles at the same site. We have

$$u_2^\epsilon(q, q, t+1) = E^\epsilon \left\{ \prod_{\sigma=\pm 1} \left[ \left(\frac{1}{2} + \epsilon\sigma\right) \eta(q-\sigma, t) - 2\epsilon\sigma \eta(q-\sigma, s, t) \eta(q-\sigma, -s, t) \right] \right\},$$

where we have omitted writing the velocities in the arguments of the  $u$ -functions: we use the convention

that each  $u$  equals the sum of  $u$ -functions over all possible velocities; above, on the left side of the equation, the sum is trivial, since the velocities must be of opposite signs. We have that

$$v_2^\epsilon(q, q, t + 1) = u_2^\epsilon(q, q, t + 1) - \sum_{\sigma = \pm 1} u_1^\epsilon(q, \sigma, t + 1) \rho_\epsilon(q, -\sigma, t + 1) + \prod_{\sigma = \pm 1} \rho_\epsilon(q, \sigma, t + 1). \quad (4.6)$$

Proceeding as before we find after some elementary manipulations that the zero order terms in  $\epsilon$  reconstruct the action of  $R$  on  $v_2^\epsilon$ , the linear order in  $\epsilon$  have the form of a lattice derivative (sites displaced by 2). The remaining terms are of the order  $\epsilon^2$ , notice that  $\epsilon < \text{const. } t^{-1/2}$ , since we do not go past times of order  $\epsilon^{-2}$ . Combining this with the previous analysis we then obtain the proof of the lemma.

*Remark.* Some comments on lemma 4.1 may be useful to relate what we are doing now with the outline given in the preceding section. The whole game was and is to extract lattice derivatives from the equation for the time increments of the  $v$ -functions since the factors  $\epsilon$  in (3.3) and (4.2) are not enough to make the perturbation series converge. Lemma 4.1 tells us that for each  $k$  the sum over  $y$  breaks up into a sum over suitable pairs  $y', y''$ . Each pair, by (4.3), is bounded by the sum of a lattice derivative of a  $v$ -function and a term of the order of  $\epsilon t^{-1/2}$ . Such result is responsible for the terms  $[s_i - s_{i+1}]^{-1/2}$  appearing in (3.6).

As explained above the validity of (4.3), crucial for our analysis, is directly related to the fact that we are computing  $Tv_n^\epsilon$ . It would not hold for  $v_n^\epsilon$  alone. Notice however that the following *telescopic* identity holds:

$$v_n^\epsilon(x, t) = \sum_{s=1}^t \sum_y Q^{t-s}(x \rightarrow y) \left[ v_n^\epsilon(y, s) - \sum_z Q(y \rightarrow z) v_n^\epsilon(z, s-1) \right], \quad (4.7a)$$

since by definition  $v_n^\epsilon(\cdot, 0) = 0$ . Hence, recalling that  $Q = TS_-$  and defining for  $s < t$ ,  $Q'_{t-s} = S_- Q^{t-s-1}$ , the right side of (4.7a) can be written in the form

$$\sum_y Q'_{t-s}(x \rightarrow y) \left[ \sum_z T(y \rightarrow z) v_n^\epsilon(z, s) - \sum_z R(y \rightarrow z) v_n^\epsilon(z, s-1) \right], \quad (4.7b)$$

so that if  $y$  is free we can again use lemma 4.1. On the other hand for the SAP the  $Q$ -probability that after time  $t - s$  the configuration in (4.7) is not free is small if  $t - s$  is large. This follows from the already mentioned fact [which we shall prove in the appendix] that the powers of  $Q$  behave asymptotically like those of the transition probability of a random walk.

We next examine a particular subset of non free configurations.

*k-type configurations*

*Definition.* A configuration  $x = (q_1, \sigma_1, \dots, q_n, \sigma_n)$  is of  $k$ -type if  $n = 2k$  and  $q_{2i} - \sigma_{2i} = q_{2i-1} - \sigma_{2i-1}$  for all  $i \leq k$ .

In the next lemma we examine the time increments of a  $v_2^\epsilon$  function evaluated at a type-1 configuration, the analysis of  $v_{2k}^\epsilon$ ,  $k > 1$ , is similar because different pairs of particles do not interfere in a unit time step.

*Lemma 4.2.* Let  $t \geq 1$  and  $q \in \mathbb{Z}$ . Then there are uniformly bounded functions  $b'$ ,  $b''$ ,  $b'''$  so that

$$\begin{aligned} v_2^\epsilon(q+1, 1, q-1, -1, t+1) &= v_2^\epsilon(q, 1, q, -1, t) + \frac{1}{2} \left[ \sum_{\tau=\pm 1} -\tau v_1^\epsilon(q, \tau, t) \right] \left[ \sum_{\tau=\pm 1} \tau \rho_\epsilon(q, \tau, t) \right] \\ &\quad - \frac{1}{4} [\rho_\epsilon(q, -1, t) - \rho_\epsilon(q, 1, t)]^2 + \epsilon^2 \left[ b'(q, t, \epsilon) v_2^\epsilon(q, 1, q, -1, t) \right. \\ &\quad \left. + \sum_{\tau=\pm 1} v_1^\epsilon(q, \tau, t) b''(q, \tau, t, \epsilon) + b'''(q, t, \epsilon) \right]. \end{aligned} \quad (4.8)$$

The proof of the lemma is essentially computational and is omitted.

#### Other configurations

There are, of course, configurations which are neither free nor of  $k$ -type. The simplest example is that of a configuration  $x$  with 2 particles at distance 2 having the same velocities. If we compute the time increments of the corresponding  $v$ -function we may argue as in lemma 4.1, and we get the same estimate except for the bound given in (4.3). This is so because in the sum on the left-hand side of (4.1) there is a configuration  $x'$  which is of type-1, for which the arguments used in the proof of lemma 4.1 do not apply. No problem about this particular configuration: we can use lemma 4.2 which gives us the *right* contribution. The troubles come from the remaining cases. We need to prove the estimate (4.3) for the time increments of  $\sum_x T'(x \rightarrow x') v_2^\epsilon(x', t)$  where  $T'$  is like  $T$  except that it equals 0 at type-1 configurations. The problem is then to find, even in such a restricted ensemble, the right cancellations which produce the desired lattice derivative, in the sense of the inequality (4.3), cf. the remarks at the end of the proof of lemma 4.1. In the present case it is easy to check that this is possible. In fact the time increment of  $v_2^\epsilon(q, -1, 1, q+1, -1, t)$  satisfies the estimates of lemma 4.1 (it is just the same computation as for the case when two particles are at the same site and such a case has been considered in lemma 4.1). Again it is a matter of simple computation to check that the same estimates hold for the combination  $\sum_{\tau=\pm 1} v_2^\epsilon(q-1, \tau, q+1, \tau)$ .

We can generalize the above argument to an arbitrary configuration where, however, the structure of the cancellations which produce the desired lattice derivative has a rather complex form. Their origin however is quite transparent: we shall group together configurations in such a way that in each group the average velocity is 0, like in the previous example. We start with a definition.

*Definition.* Given a configuration  $x = (q_1, \sigma_1, \dots, q_n, \sigma_n) \equiv (q, \sigma)$  [sometimes we shall write more explicitly  $q(x)$ ],  $q^{(f)} = \{q \in q : |q - q_i| \neq 2 \text{ for all } i = 1, \dots, n\}$  [or equivalently  $q^{(f)}(x)$ ],  $x^{(f)} = \{(q_i, \sigma_i) : q_i \in q^{(f)}\}$ ,  $x^{(1)} = \{(q_i, \sigma_i) : \text{there is } j \neq i \text{ such that } q_i - \sigma_i = q_j - \sigma_j\}$ ,  $x^{(2)} = \{(q_i, \sigma_i) : q_i \notin q^{(f)}, (q_i, \sigma_i) \notin x^{(1)}\}$ . We finally set  $q^{(\lambda)} = \{q_i : (q_i, \sigma_i) \in x^{(\lambda)}\}$  for  $\lambda = 1, 2$  [or equivalently  $q^{(\lambda)}(x)$ ].

In the above definition  $x^{(f)}$  is the free component of the configuration  $x$ ,  $x^{(1)}$  is the type- $k$  [for some  $k$ ] component of  $x$ . The sets  $q^{(1)}$  and  $q^{(2)}$  obtained from  $x$  are, as explained above, respectively unions of maximal connected components, two sites being *connected if they are at distance 2 from each other*. Using such notions we make the following definition.

Definition. Given  $x$  let  $q^{(2),i}$ ,  $i = 1, \dots, m$  be the set of all maximal connected components of  $q^{(2)}$ , and for  $i = 1, \dots, m$  let

$$S^i = \sum_{q_j \in q^{(2),i}} \sigma_j$$

be the total velocity of the particles in the  $i$ th connected component of  $q^{(2)}$ .

We still need one more definition which establishes the notation for the subsets of configurations where cancellations will occur.

Definition. Two configurations  $x$  and  $y$  are equivalent if (i)  $q(x) = q(y)$  [hence  $q^{(f)}(x) = q^{(f)}(y)$ ] (ii)  $x^{(1)} = y^{(1)}$  [hence  $q^{(2)}(x) = q^{(2)}(y)$  and they have the same connected components, say there are  $m$  such components] and finally (iii)  $|S^{(2),i}(x)| = |S^{(2),i}(y)|$ , for  $i = 1, \dots, m$ .

Given a configuration  $x$  we denote by  $\pi(x)$  the set of all configurations  $y$  which are equivalent to  $x$ . We then denote by  $\Pi_x$  the collection of all mutually distinct subsets of the form  $\pi(y)$  for any  $y$  such that  $T(x \rightarrow y) > 0$ . For  $a \in \Pi_x$  set  $a^{(\lambda)} = \{y^{(\lambda)} : y \in a\}$ , where  $\lambda = f, 1, 2$ .

Let  $x$  be a configuration with  $n$  particles and assume that there are  $k$  sites where there is one and just one particle of  $x$ . Then with the above notation

$$\sum_y T(x \rightarrow y) v_n^c(y, t) = 2^{-k} \sum_{a \in \Pi_x} \sum_{y \in a} v_n^c(y, t). \tag{4.9}$$

The cancellations will occur among configurations belonging to the same atom  $a$ . In the sequel therefore we fix an  $a$  and analyse the sum over  $y \in a$ . We shall first formalize a property already used several times, namely that the effects on the time increments of  $v^c$  coming from particles which do not interact in a unit time step are independent.

Let us define the increments of  $v^c$  as

$$v_{|z|}^c(z, t + 1) = \sum_{z'} v_{|z|}^c c(\epsilon, z, z', t), \tag{4.10}$$

where  $|z|$  denotes the number of particles in  $z$ . To simplify the writing we set  $v^c(x, t) \oplus v^c(y, t)$  for  $v_{|z|}^c(z, t)$  with  $z$  being the disjoint union of  $x$  and  $y$ . An analogous definition holds for the  $\oplus$  product of more than two  $v$ -functions. Thus given  $x$  and  $a \in \Pi_x$

$$\sum_{y \in a} v^c(y, t + 1) = \bigoplus_{\lambda=f,1,2} \sum_{y^{(\lambda)} \in a^{(\lambda)}} \sum_z v^c(z, t) c(\epsilon, y^{(\lambda)}, z, t). \tag{4.11}$$

We have already studied the structure of the increments for configurations which are either free or of type- $k$  in lemmas 4.1 and 4.2. Eq. (4.11) tells us that such estimates can be applied also for general configurations, because of the product structure of (4.11). So we are left with the problem of estimating the contribution of the sum over all  $y^{(2)} \in a^{(2)}$ .

Using again the independence of the effects coming from different connected components we can actually reduce ourselves to the case where  $a^{(2)}$  has a unique connected component. In this case the positions of the particles in any of the configurations  $y^{(2)} \in a^{(2)}$  are fixed and are  $q, q + 2, \dots, q + 2N$  for

suitable  $q$  and  $N$ . Call  $\sigma_0, \dots, \sigma_N$  the corresponding velocities, then, necessarily  $\sigma_i \geq \sigma_{i+1}$  for  $i = 0, \dots, N-1$ , otherwise there would be a type- $k$  component in the configuration. Therefore the allowed configurations are characterized by an index  $I$  which takes the values  $0, \dots, N+1$ :  $I = k$  means that  $\sigma_i = 1$  for  $i < k$  and  $\sigma_i = -1$  for  $i \geq k$ . If  $I = N+1$  this means that all  $\sigma_i = 1$ . The configuration with  $I = k$  has therefore total velocity opposite to that of the configuration with  $I = N+1-k$ . Since an atom  $a$  is determined by the requirement that all its configurations have the same absolute value of the total velocity there are just two configurations in  $a$  if  $I = k$  and  $k \neq N+1-k$  and only one otherwise.

Let  $\sigma_i = 1$  for  $i < k$  and  $= -1$  otherwise, i.e.  $I = k$ , call  $y^{(2)}$  such a configuration and let it be in the atom  $a$  we are considering. Then using (4.4)

$$\begin{aligned} \sum_z v^\epsilon(z, t) c(\epsilon, y^{(2)}, z, t) = & \bigoplus_{i=0, N} \left[ \left( \frac{1}{2} + \epsilon \sigma_i \right) \sum_{\tau=\pm 1} v_1^\epsilon(q + 2i - \sigma_i, \tau, t) \right. \\ & - 2\epsilon \sigma_i \left\{ v_2^\epsilon(q + 2i - \sigma_i, 1, q + 2i - \sigma_i, -1, t) \right. \\ & \left. \left. + \sum_{\tau=\pm 1} v_1^\epsilon(q + 2i - \sigma_i, \tau, t) \rho_\epsilon(q + 2i - \sigma_i, -\tau, t) \right\} \right]. \end{aligned} \quad (4.12)$$

Developing the product, the zero order in  $\epsilon$  gives  $Q$  acting on  $v^\epsilon(\cdot, t)$ , while the first order term is

$$\begin{aligned} & \sum_{i=0, N} \epsilon \sigma_i \left[ \bigoplus_{j \neq i} \frac{1}{2} v_1^\epsilon(q + 2j - \sigma_j, \tau, t) \right] \\ & \bigoplus \left[ \sum_{\tau=\pm 1} v_1^\epsilon(q + 2i - \sigma_i, \tau, t) - 2v_2^\epsilon(q + 2i - \sigma_i, 1, q + 2i - \sigma_i, -1, y) \right. \\ & \left. - 2 \sum_{\tau=\pm 1} v_1^\epsilon(q + 2i - \sigma_i, \tau, t) \rho_\epsilon(q + 2i - \sigma_i, -\tau, t) \right]. \end{aligned} \quad (4.13)$$

If we consider the same expression starting from the other configuration  $w^{(2)}$  in the same atom  $a^{(2)}$  as  $y^{(2)}$ , where, to be definite, we assume that  $a^{(2)}$  has two elements, we get the same expression as in (4.13) but with the new configuration  $\sigma'$ . Since  $I = N+1-k$  in this configuration it then follows that there are as many positive values of  $\sigma'_i$  as there are negative values of  $\sigma_i$ . By putting these together we reconstruct differences between  $v$ -functions and/or products of  $v$ -functions and  $\rho$  functions with displaced arguments. In this way we obtain lattice derivatives of  $v$ -functions and/or  $\rho$  functions. The latter are estimated by (4.5) and in this way we derive the same estimates as in lemma 4.1 also for this case, where the lattice derivatives involve differences of  $v$ -functions with sites displaced by at most  $2N$ .

### Summary of estimates

With the above considerations we have completed the characterization of the time increments of the  $v$ -functions as they appear in the perturbation expansion (4.7). We summarize below the main conclusions. (1) If no particle disappears when developing the square bracket term in (4.7b), then we have a factor  $\epsilon$  together with either a lattice derivative or a factor  $s^{-1/2}$ . (2) If however  $s = t$ , cf. (4.7a) and (4.7b), we only get a factor  $\epsilon$ . [This will be enough because we do not have to sum over  $s$  as we had to in case (1).] (3) When particles die there is a type- $k$  component in  $y$  at time  $s$ , which implies that  $k$  pairs of particles at time  $s$  have to be close [in each pair the distance between the two particles of the pair is 2]. Furthermore

for each pair there are three possibilities; (i) no particle dies, (ii) one particle dies contributing a factor  $s^{-1/2}$ , cf. (4.8) and (4.5), and there is a lattice derivative, (iii) both particles in the pair die, contributing a factor  $s^{-1}$ , cf. again (4.8) and (4.5). Instead of this in any of the above three subcases of (3) there might be a contribution of the order of  $\epsilon^2$ , cf. (4.8).

For the convergence of the perturbation series it is enough at each step to have one single lattice derivative and we can disregard the possible presence of other ones.

The same general features as above appear also in the weakly asymmetric simple exclusion process considered in [7], the number of terms at each step of the expansion, as well as their specific structure are different than in the present case, but they lead to the same conclusion (1), ..., (3) as above. There is however still an important difference, namely that the unperturbed motion described here by  $Q$  is not the same as the one appearing in [7], i.e. the transition probability of the symmetric simple exclusion process. In [7] the main point was to use the closeness of such process to that of independent symmetric random walks, here we obtain similar closeness between our  $n$ -particle symmetric automaton, with transition probability  $Q$ , and the process of  $n$  symmetric independent particles. This is done in the appendix, and since we can derive the same bounds as in [7] for the difference between the interacting and independent processes we are then reduced to the case considered in [7]. We do not consider it necessary to reproduce in the present context the arguments used in [7] to conclude the still long proof, since this requires essentially only a change of language. On the other hand the reader who has reached this point of the paper might like to have a somewhat more detailed outline than that in section 3, of how the proof proceeds, and this is what we are going to do below.

*Sketch of remaining steps*

As already mentioned in section 3 we cannot iterate indefinitely eq. (4.7a), even if we limit  $t$  to be less than  $\epsilon^{-2+\beta}$ ,  $\beta > 0$ . All our estimates in this section depend strongly on the number of particles [degree of the  $v$ -functions]. We are forced to keep this number fixed and take the limit as  $\epsilon \rightarrow 0$ , no diagonal limiting procedure [ $n \rightarrow \infty$  as  $\epsilon \rightarrow 0$ ] is allowed by our estimates. So let us fix the value of  $n$  for which we want to prove (3.11a) at  $t \leq \epsilon^{-2+\beta}$ . We then choose an integer  $N$  such that

$$\frac{n}{8} < \beta \frac{N}{4} \tag{4.14}$$

and we iterate (4.7)  $N$  times. We get a finite sum of terms. Part of these terms do not have any  $v$ -functions left, all particles have died; the remaining terms still have  $v$ -functions and their contribution is bounded by

$$c \sum_{t < s_1 < \dots < s_N} (t - s_1)^{-1/2} (s_1 - s_2)^{-1/2} \dots (s_{N-1} - s_N)^{-1/2} \epsilon^N \leq c' \epsilon^{-(2-\beta)N/2+N}. \tag{4.15}$$

The estimate in (4.15) is obtained by proving that the largest contribution comes from terms in which no particle dies. We can then use the estimates in lemma 4.1. We shall see that each lattice derivative produces a factor  $(s_i - s_{i+1})^{-1/2}$ ,  $i$  being the  $i$ th step in the iteration. The constant  $c$  in (4.15) is a combinatorial factor which takes into account the number of terms present in the remainder. Notice in fact that any  $v$ -function is bounded by 1. The right-hand side in (4.15) is smaller than  $c' \epsilon^{\beta N/2}$ . By (4.14) it is therefore smaller than  $t^{-n/8}$ , because  $t < \epsilon^{-2}$ .

We shall next analyse the other terms obtained when iterating (4.7)  $N$  times, namely those where all particles die. We group together terms by fixing the number of iterations [of (4.7)], say  $k$ , and the times

when the iteration takes place, say  $s_1, \dots, s_k$ . We further specify which particles die and which are the newly created particles [recall that when a new particle is created it starts from the same site where a previously existing particle was, we specify which was the *ancestor* but not its position]. Once all this is fixed, what we are left with is a process with initially  $n$  particles. They move according to the law specified by  $Q$  till time  $t - s_1$ . Then some birth and/or death occurs which determines a new particle configuration at time  $t - (s_1 - 1)$ . The specifications allow to reconstruct such a new configuration if that at time  $t - s_1$  is known. After this the  $Q$  process with a possibly different number of particles runs for a time  $s_1 - s_2$ . Then a new birth and death process takes place, and so forth. The game would be over [almost] if  $Q$  were the transition probability of the process of independent particles, we could then estimate the probabilities of events like *two particles are at distance 2* by using classical results for random walks. To reduce ourselves to this case we introduce a *coupling* between the exclusion and the independent processes, namely we introduce a larger space where both processes live and a joint law for the two whose marginals are the exclusion and independent process. The main feature of such coupling is that with large probability the interacting and the corresponding independent particles do not get too far from each other, which means that at time  $t$  they are not farther than  $t^\alpha$  as  $t \rightarrow \infty$ , for any  $\alpha > \frac{1}{4}$ . An application of this result is that with large probability the condition that two particles at time  $t$  are at distance 2 from each other, is a consequence of the condition that the corresponding independent particles are at distance  $2 + 2t^\alpha$ . This explains the reason for the bound in (3.11a). It turns out that the largest contribution in the iteration comes from terms where, at each step, there are only deaths, and just 2 particles die. Assume  $n = 2k$  and denote by  $P$  the law of the process starting with  $n$  particles from  $x$  and by  $P^0$  that for the independent particles, starting from the same initial configuration. When referring to independent particles we add a superscript 0. We have a bound which goes like

$$\begin{aligned} & \sum_{t > s_1 \dots > s_k > 0} P \left[ \bigcap_{i=1, k} \{ |q_{2i-1}(t - s_i) - q_{2i}(t - s_i)| = 2 \} \right] \prod_{i=1, k} s_i^{-1} \\ & \leq \sum_{t > s_1 \dots > s_k > 0} P^0 \left[ \bigcap_{i=1, k} \{ |q_{2i-1}^0(t - s_i) - q_{2i}^0(t - s_{2i})| \leq 2 + 2(t - s_i)^\alpha \} \right] \prod_{i=1, k} s_i^{-1} \\ & \leq \text{const.} \sum_{t > s_1 \dots > s_k > 0} (t - s_1)^{\alpha-1/2} \dots (t - s_k)^{\alpha-1/2} s_1^{-1} \dots s_k^{-1} \end{aligned}$$

times a large but fixed constant which takes into account the number of terms when such situations may occur. The last sum vanishes like  $t^{-\gamma k}$  for any  $\gamma < \frac{1}{2} - \alpha$ , i.e. like  $t^{-n(1/8-\zeta)}$  for any  $\zeta > \alpha/2 - \frac{1}{8} > 0$  and arbitrarily close to 0 since  $\alpha$  may be any number larger than  $\frac{1}{4}$ . To really get the exponent in (3.11a) equal to  $\frac{1}{8}$  we need to exploit one more feature of the coupling, namely that the first two [or any other chosen pair of] particles have the same distance in the exclusion and independent processes, with at most an error of 2. So that we can replace the condition that particles 1 and 2 are at distance 2 at time  $t - s_1$  by requiring that the independent particles 1 and 2 are, at the same time, at distance not larger than 4. In this way the first factor in the last sum namely  $(t - s_1)^{\alpha-1/2}$  is replaced by  $(t - s_1)^{-1/2}$  and so (3.11a) is obtained.

Of course the proof that the above is the *worst case* requires an analysis of all the others. This is done in a way similar to the above one, but we first need to specify how to properly take into account the contributions coming from the lattice derivatives. This is the last point we want to mention about the analysis of the case  $t \leq \epsilon^{-2+\beta}$ . Assume  $x$  is a configuration with  $n + 1$  particles [notice  $n + 1$  and not  $n$ ].



Denote by  $x^{(i)}$ ,  $i = 1, 2$  the same configuration without particle  $i$ . By using (4.7a)

$$v_n^\epsilon(x^{(2)}, t) - v_n^\epsilon(x^{(1)}, t) = \sum_{s=1}^t \sum_y [Q^{t-s}(x^{(2)} \rightarrow y) - Q^{t-s}(x^{(1)} \rightarrow y)] \psi(y, \epsilon, s),$$

where  $\psi$  is simply a shorthand for the square bracket term in (4.7a). We now use a further property of the evolution defined by  $Q$ , i.e. of the symmetric automaton, which states that it is possible to label the particles and follow them in time in such a way that the marginal over the motion of  $n$  particles of the process with  $n + 1$  particles, has the law of the automaton with those  $n$  particles. We discuss this point in the appendix, it is a nontrivial feature of the process, but perhaps not so surprising since we have seen that in some respects the symmetric automaton looks like the independent process, cf. section 3, and for the independent process the above property holds trivially. Assume therefore that we have introduced such a labeled process and consider the event  $\{ \text{particle 1 and particle 2 are at same site at some time} \leq t - s \}$ . Conditioned on this event the configuration at time  $t - s$  will be symmetric under the exchange of particles 1 and 2 and it will not contribute to (4.16). Therefore the expression (4.16) does not change if we add the condition that the above event does not occur. We do this and then consider separately the two terms in the above difference, the contribution coming from the lattice derivative is in fact sufficiently well taken into account by the characteristic function that the two particles involved do not meet in the time interval associated with the iterative step we are considering. As before this projects into a property of the independent particles: here we are *lucky*, since the event refers to the distance between two particles. By choosing, in the given time interval under consideration, the coupling with the independent particles in such a way that the two particles distances are the same [or differing at most by 2] in the exclusion and independent process we reduce the previous condition to one for the independent particles. We had already mentioned this feature of the coupling. Note that (1) we have to choose properly the coupling in the different steps of the iteration and (2) the probability that two random walks do not meet in a time interval  $s_i - s_{i+1}$  starting from a distance  $d$  is bounded by  $cd(s_i - s_{i+1})^{-1/2}$ ,  $c$  being a suitable constant. This is the origin of the corresponding factors in (3.6).

By combining this and the previous estimates one can bound all the terms coming from the  $N$  iterations of (4.7) and prove the validity of (3.11a) for  $t \leq \epsilon^{-2+\beta}$ ; we omit the details.

### *Macroscopic times*

We outline below the various steps necessary to reach the desired time  $\epsilon^{-2}\tau$  [its integer part to be more precise], where  $\tau$  is any given, fixed positive number.

First let us assume that the system lives in a finite interval of  $Z$  of length  $\epsilon^{-2}$  with periodic boundary conditions. From a physical point of view this is just the same as before, since we are interested in what happens in regions of size  $\epsilon^{-1}$ . Technically this simplifies the proofs a little bit, the infinite case, however, can be treated as well, cf. [7]. At time  $t \equiv \epsilon^{-2+\beta}$  we compare the actual [random] configuration  $\eta^{(1)}$  with the solution  $\rho_\epsilon(\cdot, t|\eta^{(0)})$  to (3.10) with initial condition  $\eta^{(0)}$ . We estimate their difference by means of the following expression:

$$\| \rho_\epsilon(\cdot, t|\eta^{(0)}) - \eta^{(1)} \| \doteq \sup_{q, \sigma} \left| \sum_{q', \sigma'} Q^{t(\epsilon^{-1/2})}((q, \sigma) \rightarrow (q', \sigma')) [ \rho_\epsilon(q', \sigma', t|\eta^{(0)}) - \eta^{(1)}(q', \sigma') ] \right|. \quad (4.16)$$

Using (3.11a) and proceeding as in the proof of theorem 2.2, cf. section 3, we can easily see that the probability that (4.16) is larger than  $\epsilon^\zeta$  vanishes faster than any power of  $\epsilon$  if  $\zeta < \frac{1}{8}$ . [Since  $Q$  is essentially a random walk transition probability,  $q'$  ranges typically in an interval of size  $\epsilon^{-1/4}$ . The normal fluctuations of the total density in such a region are of order  $\epsilon^{1/8}$ , hence the choice of  $\zeta$ ]. We choose  $\eta^{(1)}$  so that (4.16) is less than  $\epsilon^\zeta$ . We then take  $\eta^{(1)}$  as the new initial condition and let the system run for another time interval of length  $\epsilon^{-2+\beta}$ . We then compare the new random configuration  $\eta^{(2)}$  with  $\rho_\epsilon(\cdot, \epsilon^{-2+\beta} | \eta^{(1)})$  and, as before, we can assume that their  $\|\cdot\|$  difference is less than  $\epsilon^\zeta$ . We keep doing this until we reach the time  $\epsilon^{-2\tau}$ . We also choose the initial configuration  $\eta^{(0)}$  so that its  $\|\cdot\|$  difference from the initial profile  $\rho_\epsilon(q, \sigma) \equiv \frac{1}{2}\rho_\epsilon(\epsilon q)$  is also smaller than  $\epsilon^\zeta$ . In doing this we are neglecting configurations whose probability vanishes faster than any power of  $\epsilon$ , since  $\mu_\epsilon$  is a product measure.

The following statement can be proven by slightly modifying the proof that the solution to (3.10) converges to that of (2.6), cf. also [7], and it expresses the stability of (3.10) with respect to small perturbations.

Let  $\eta^{(k)}$ ,  $k \geq 0$ , be as above. Then if  $\beta$  is small enough, there are  $\delta > 0$  and  $c$  such that for all  $1 \leq k \leq \epsilon^{-2\tau}$  and all  $t$  such that  $k\epsilon^{-2+\beta} \leq t \leq (k+1)\epsilon^{-2+\beta}$

$$\sup_q |\rho_\epsilon(q, t | \eta^{(0)}) - \rho_\epsilon(q, t - (k-1)\epsilon^{-2+\beta} | \eta^{(k-1)})| \leq c\epsilon^\delta. \quad (4.17)$$

The next step in the proof of proposition 3.1 consists in proving (3.11a) with some positive exponent, not necessarily as large as  $\frac{1}{8}$ . This is a rather easy consequence of (4.17). Let  $t > \epsilon^{-2+\beta}$ , the proof in the other case has already been obtained. Assume also that  $t \leq \epsilon^{-2\tau}$ . Call  $k$  the largest integer such that  $(k+1)\epsilon^{-2+\beta} \leq t$ . We then have

$$v_n^\epsilon(x, t | \eta) = E_\eta^\epsilon \left[ \mathcal{E}_{k\epsilon^{-2+\beta}}^\epsilon \left( \prod_{i=1, n} \{ \eta(x_i, t) - \rho_\epsilon(x_i, t | \eta) \} \right) \right], \quad (4.18)$$

where  $\mathcal{E}_s^\epsilon$  denotes the expectation conditioned on the value of the process in the whole time interval  $[0, s]$ . Since this is a Markov process, such conditional expectation only depends on the configuration  $\eta_s$  at time  $s$ .

We assume that the initial configuration  $\eta$  is such that

$$\|\eta - \rho_\epsilon\| \leq \epsilon^\zeta,$$

where, as before,  $\rho_\epsilon(q, \sigma) \equiv \frac{1}{2}\rho_\epsilon(\epsilon q)$ . Denote by  $\chi$  the characteristic function of the event

$$\|\eta_s - \rho_\epsilon(\cdot, \cdot, s | \mu_\epsilon)\| \leq \epsilon^\zeta \quad \text{for all } s = i\epsilon^{-2+\beta}, i \leq k,$$

where  $\rho_\epsilon(\cdot, \cdot, s | \mu_\epsilon)$  is the solution to (3.10) with initial condition  $\rho_\epsilon(q, \sigma)$  as above. Setting  $\chi^c = 1 - \chi$  we have

$$|v_n^\epsilon(x, t)| \leq E_\eta^\epsilon(\chi^c) + E_\eta^\epsilon \left[ \chi E_{\eta(\cdot, \cdot, k\epsilon^{-2+\beta})}^\epsilon \left( \prod_{i=1, n} \{ \eta(x_i, t - k\epsilon^{-2+\beta}) - \rho_\epsilon(x_i, t | \mu_\epsilon) \} \right) \right].$$

As already noted the expectation of  $\chi^c$  vanishes faster than any power of  $\epsilon$ , cf. also the proof of theorem

2.2 in section 3. In the second term we add and subtract

$$\rho_\epsilon(x_i, t - k\epsilon^{-2+\beta} | \eta(\cdot, \cdot, k\epsilon^{-2+\beta})),$$

namely the solution to (3.10) with initial condition  $\eta(\cdot, \cdot, k\epsilon^{-2+\beta})$ . We expand the product. Since the difference of the two  $\rho$  functions depends only on the configuration at time  $k\epsilon^{-2+\beta}$  it drops out of the conditional expectation, which therefore defines a  $\nu$ -function at a time less than  $2\epsilon^{-2+\beta}$ . For this we use (3.11a) already proven in such a time interval. To estimate the difference of the two  $\rho$  functions we use (4.17), so that we get a bound which goes like  $\epsilon^\delta$  [i.e. like  $t^{-\delta/2}$  when  $t$  is of the order of  $\epsilon^{-2}$ ], because of the presence of the characteristic function  $\chi$ . In this way we have proven a bound like on the right of (3.11a) with the exponent being the smallest of  $\delta/2$  and  $\frac{1}{8}$ , it can be seen that actually the latter is the largest of the two.

It is now easy to prove (3.11a) with the right exponent  $\frac{1}{8}$ . Let  $n$  be fixed and  $t = \epsilon^{-2}\tau$ . We go back to (4.7a) and use it again to rewrite the terms containing  $\nu^\epsilon$  in its right hand side. We keep doing that for each term which contains a  $\nu$ -function with less than  $N$  particles, where  $N$  is such that  $N\delta/2 > n/8$ . For such terms we use the previous estimate with exponent  $\delta/2$ . For the others we use the same analysis as at short times, i.e. at  $t < \epsilon^{-2+\beta}$ . Since we have limited the growth of the number of particles we do not have convergence problems. What was accomplished earlier by assuming  $t < \epsilon^{-2+\beta}$  is now obtained using (3.11a) with exponent  $\delta/2$ .

The proof of (3.11b) is very similar to the previous one, one only needs to use (4.5) with  $\epsilon$  on the right side instead of  $t^{-1/2}$ , and this is correct because of the assumed smoothness of the initial datum. Actually the whole proof becomes much simpler, cf. for instance [7]. Notice however that the proof still requires an a priori bound for  $\nu_n^\epsilon$  which goes like  $\epsilon^{\delta n}$ , so that we cannot shortcut the previous analysis and examine directly the case of the initial state is described by  $\mu_\epsilon$ , at least using our type of approach.

### 5. Long time behavior: Micro-macro connection

As mentioned in the introduction the Burgers' equation (1.1) has traveling solutions, namely solutions of the form  $\rho(r, t) = \rho^*(r - vt)$ , where  $\rho^*$  is a smooth increasing function such that

$$\lim_{r \rightarrow \pm\infty} \rho^*(r) = \rho_\pm, \quad \rho_+ > \rho_- \tag{5.1}$$

and

$$v = c \left[ 1 - \frac{1}{2}(\rho_+ + \rho_-) \right]. \tag{5.2}$$

While the velocity  $v$  is independent of the viscosity  $\nu$  appearing in the right of (1.1), the shape of  $\rho^*$  does depend on  $\nu$  and in the limit as  $\nu \rightarrow 0$ ,  $\rho^*$  converges to a step function [shock wave] with values  $\rho_-$  and  $\rho_+$ , cf. for instance [6].

The stability of the traveling solutions to (1.1) with  $\nu = \frac{1}{2}$  can be analysed in terms of the automaton model since its density profile converges to the solution of (2.6), i.e. of (1.1) with  $c = 2$  and  $\nu = \frac{1}{2}$ . Therefore we choose the initial condition  $\rho(r) = \rho^*(r)$ , a traveling solution with velocity  $v > 0$ . If we then observe the system at a fixed macroscopic time  $t$  ( $\epsilon^{-2}t$  in microscopic units) from a frame which moves with velocity  $v$ , we will see, when  $\epsilon \rightarrow 0$ , the same profile  $\rho^*$  as at time 0. The question then is what

happens if we look at microscopic times  $\tau$  much larger than  $\epsilon^{-2}$ . More precisely we would like to know on which time scale, if any, we start observing macroscopic deviations from the traveling profile solution to (2.6). Remember that theorem 2.1 requires in an essential way that  $\tau\epsilon^2$  stays bounded when  $\epsilon \rightarrow 0$ . If there are going to be deviations for longer times, say microscopic times of the order of  $\epsilon^{-3}$ , then we would like to know in which way the macroscopic profile changes. We refer to [5] for a detailed analysis of such questions in a more general frame, here we make just a few remarks. First notice that the stability analysis on (1.1) does not give unique answers: a traveling profile  $\rho^*$  is (i) stable with respect to local perturbations which preserve the total mass; (ii) if the perturbation is still local but does not preserve the total mass then the shape of the traveling wave is stable but, asymptotically in time, the profile is spatially shifted with respect to the unperturbed one by an amount which depends on the mass variation. Finally (iii) the profile is unstable for macroscopic changes of the density at infinity. Which one of the above cases, if any applies to the description of the WAAP?

By looking at the automaton it seems natural to measure stability in terms of the size of the fluctuations. Theorem 2.3 provides an explicit expression for the covariance of such fluctuations. In particular one can see that  $C_t(r, r')$  diverges linearly in  $t$  as  $t \rightarrow \infty$ . The analysis is the same as in [7] to which we refer. Actually it is possible to deduce from the asymptotics of the covariance that the fluctuations' growth is due to rigid random spatial shifts of the profile of the traveling wave around its average position. Such effect should become macroscopic after macroscopic times of the order of  $\epsilon^{-1}$ , hence at microscopic times of the order of  $\epsilon^{-3}$  for the WAAP.

This phenomenon has been observed in the asymmetric simple exclusion for a shock with left density 0, i.e.  $\rho_- = 0$ , cf. [9], [10]. It is proven that the macroscopic shape of the density profile is stable and that the shock's location fluctuates like Brownian motion and the diffusion coefficient is explicitly computed. Most interestingly the shape of the shock is *well defined even at the particle level*. There is a stationary measure for the process as seen from the leftmost particle which converges exponentially to the Bernoulli measure with density  $\rho_+$ , far to the right of the leftmost particle.

It is an open question whether similar results hold for the asymmetric exclusion in general, i.e.  $\rho_- > 0$ , and for the automaton process of Boghosian and Levermore. The computer simulation in [6] seems to suggest that the shock broadens with time, contrary to what is indicated by the analysis of the fluctuations discussed above. On the other hand there are very recent computer simulations on the exclusion process [11] which indicate that the profile of the shock is stable, its location fluctuates like a Brownian motion and its diffusion coefficient agrees with that found by studying the fluctuations of the weakly asymmetric simple exclusion process. Actually the main purpose of the simulation was to see whether there is a stationary shock at the particle level, when  $\rho_- > 0$ . Their answer is positive, and we shall now describe their idea for finding microscopically the location of the shock in the context of the Boghosian-Levermore automaton [unlike the case  $\rho_- = 0$ , where the shock could be identified microscopically with the position of the leftmost particle, if  $\rho_- > 0$  it is not so clear what to do].

Let us consider the asymmetric automaton,  $p > \frac{1}{2}$ , cf. section 2 for notation. Start with a state corresponding to density  $\rho_{\pm}$  to the right, resp. left, of the origin. Call the particles in this state *first class particles*. Add other particles, *the second class particles*, to the left of the origin so that overall first and second class particles have density  $\rho_+$ . Let the particles move according to the rules given in the appendix, considering the first class particles as having a smaller label than the second class ones. In this way the first class particles move like they were alone, while first and second class particles together move also according to the Boghosian-Levermore automaton. Overall the system is in equilibrium and one might conjecture that the microscopic shock starts from the position of the rightmost second class particle. A similar procedure applies to the exclusion process and in the above referred simulation [11] one does

indeed see a stationary distribution when viewed from the rightmost second class particle. Far away from this particle the measure appears very close to the equilibrium measure with parameter  $\rho_{\pm}$ . It would be very interesting to have a rigorous proof of this behavior and to know whether the Boghosian–Levermore automaton has similar behavior.

We may add here that finding the level at which microscopic models with the same macroscopic behavior differ is a problem of general interest. In our case this means knowing how long we have to wait and how hard we have to look, to see a difference between the simple exclusion and the BL automaton. Put differently, what are the *universality class* for hydrodynamical behavior? This question has been investigated extensively for equilibrium phenomena, but only little is known for models of non-equilibrium behavior [5].

### Acknowledgements

We are indebted to Anna DeMasi, Pablo Ferrari, David Levermore and Herbert Spohn for many helpful discussions, comments and suggestions.

### Appendix

In this appendix we study the *backward* symmetric automaton with finitely many particles. Let us begin with the case  $n = 1$ . Assume the particle is initially at  $x = (q, \sigma)$ . Then at time 1 it will be at  $(q - \sigma, \sigma')$  where  $\sigma' = \pm 1$  with probability  $\frac{1}{2}$ , the corresponding transition probability is denoted by  $Q$ , cf. section 4, [we shall use the same symbol  $Q$  also when there will be more particles]. At time  $k$ , i.e. after  $k$  time steps, the random position  $q'$  of the particles will have the same distribution as a symmetric random walk initially at  $q - \sigma$  [not  $q$ ] and moving for a time  $k - 1$ , [not  $k$ ]. Its velocity at this time is  $\pm 1$  with equal probability.

The randomness is all in the velocities, given the velocities the motion is deterministic. It is therefore natural to introduce the probability space  $(\{-1, 1\}^N, \lambda)$  where  $\lambda$  is the Bernoulli measure on  $\{-1, 1\}^N$  which gives equal probability to  $\pm 1$ . We interpret  $\omega = (\omega(i), i \geq 1) \in \{-1, 1\}^N$  as a possible choice of the velocities at the times  $1, 2, \dots$ , each value  $\omega(i)$  occurring with probability  $\frac{1}{2}$  independently of the other values. It is convenient to denote the velocities by  $\omega$  and not by  $\sigma$  for reasons which will become clear in a while.

Let us now consider  $n$  particles,  $n > 1$ . We introduce  $n$  independent copies of the previous probability space  $(\{-1, 1\}^N, \lambda)$ , and denote the elements of the new space by  $\omega = (\omega_1, \dots, \omega_n)$ . We label the  $n$  particles and define  $x(k)$ ,  $k \geq 0$  starting from the initial configuration  $x$  as follows:  $x(k) \equiv (q(k), \sigma(k))$ ,  $x(0) = x$ , and for all  $i = 1, \dots, n$

$$q_i(k+1) = q_i(k) - \sigma(k), \tag{A.1}$$

$$\sigma_i(k+1) = \omega_i(k+1), \quad \text{if } q_j(k+1) \neq q_i(k+1) \forall j < i, \tag{A.2a}$$

$$\sigma_i(k+1) = -\omega_j(k+1), \quad \text{if } \exists j < i: q_j(k+1) = q_i(k+1), \tag{A.2b}$$

i.e. in a *collision* the lower labeled particle goes its own way while the higher labeled one accommodates to preserve the exclusion rule. We shall call this process the *symmetric labeled automaton*. It is trivial to see that if we neglect the labeling, i.e. we identify configurations which differ by a permutation of the labeling, then the process becomes the symmetric automaton process, with transition probability  $Q$ .

It is also easy to see that the process of any subset  $i_1, \dots, i_m$  of the  $n$  particles  $1, \dots, n$  has the same law as the process with the  $i_1, \dots, i_m$  particles alone. In particular any pair of particles move like independent particles until they collide with each other. Such properties are needed at some stage in the proof of proposition 3.1, cf. the end of the analysis of the estimates for  $t \leq \epsilon^{-2+\beta}$  in section 4.

Let us now discuss the coupling of this *exclusion* automaton with the independent one. Notice that the above realization of the labeled automaton process with exclusion *naturally* defines a coupling. Let us denote by  $\oplus_{i=1, n}(\{-1, 1\}^N, \lambda)$ ,  $\oplus$  standing here for the independent product of probability spaces, the space where we have realized the labeled automaton process. Define in such a space the process  $x^0(k)$ ,  $k \geq 0$ , as  $x^0(0) = x$  and

$$q_i^0(k+1) = q_i^0(k) - \sigma_i^0(k), \quad \sigma_i^0(k) = \omega_i(k), \quad (\text{A.3})$$

for  $i = 1, \dots, n$ . Of course this is the independent process, hence we have realized a coupling with the labeled automaton, both processes living in the same space. This is a very *effective* coupling indeed, as we are going to discuss. Notice first that particles with the same labels have the same velocities, except perhaps at the collision times, when two particles are at the same site. In fact from (A.2) and (A.3) we have

$$\sigma_i(t) - \sigma_i^0(t) = \sum_{j < i} \Delta_{i,j}(t), \quad (\text{A.4})$$

where

$$\Delta_{i,j}(t) = \delta_{q_i(t), q_j(t)} [-\omega_j(t) - \omega_i(t)]. \quad (\text{A.5})$$

Therefore

$$q_i(t) - q_i^0(t) = \sum_{j < i} \sum_{s=0}^{t-1} \Delta_{i,j}(s) \quad (\text{A.6})$$

thus the total displacement at a time  $t$  between the exclusion and the independent particles with label  $i$  is the sum over  $j < i$  of the displacements up to time  $t$  caused by the collisions with particle  $j$ . We are going to show that the probability that  $|q_i(t) - q_i^0(t)| > t^\alpha$  with  $\alpha > \frac{1}{2}$  goes to 0 faster than any inverse power of  $t$ .

Let us examine the effect of the collisions with particle  $j$ ,  $j < i$ , namely

$$\sum_{s=0}^{t-1} \Delta_{i,j}(s). \quad (\text{A.7})$$

To compute this quantity we need to know the trajectory of the exclusion particles  $i$  and  $j$  [up to time  $t-1$ ] and the values  $\omega_i(s)$  and  $\omega_j(s)$  at the times  $s$  when particles  $i$  and  $j$  collide with each other. Note