Derivation of the Leroux system as the hydrodynamic limit of a two-component lattice gas

József Fritz   Bálint Tóth

Institute of Mathematics
Budapest University of Technology and Economics

December 15, 2003

Abstract

The long time behavior of a couple of interacting asymmetric exclusion processes of opposite velocities is investigated in one space dimension. We do not allow two particles at the same site, and a collision effect (exchange) takes place when particles of opposite velocities meet at neighboring sites. There are two conserved quantities, and the model admits hyperbolic (Euler) scaling; the hydrodynamic limit results in the classical Leroux system of conservation laws, *even beyond the appearence of shocks.* Actually, we prove convergence to the set of entropy solutions, the question of uniqueness is left open. To control rapid oscillations of Lax entropies via logarithmic Sobolev inequality estimates, the symmetric part of the process is speeded up in a suitable way, thus a slowly vanishing viscosity is obtained at the macroscopic level. Following [4, 5], the stochastic version of Tartar–Murat theory of compensated compactness is extended to two-component stochastic models.

Key words: hydrodynamic limit, hyperbolic scaling, systems of conservation laws, compensated compactness

AMS 2001 subject classification: 60K35, 82C24

1 Introduction

The main purpose of this paper is to derive a couple of Euler equations (hyperbolic conservation laws) in a regime of shocks. While the case of smooth macroscopic solutions is quite well understood, see [24] and [14], serious difficulties emerge when the existence of classical solutions breaks down. A general method to handle attractive systems has been elaborated in [16], see also [4] and [9] for further references. Hyperbolic models with two conservation laws, however, can not be attractive in the usual sense because the phase space is not ordered in a natural way. We have to extend some advanced methods of PDE theory of hyperbolic conservation laws to stochastic (microscopic) systems. Lax entropy and compensated compactness are the main key words here, see [10], [11], [13], [19], [20], [2] for the first ideas, and the textbook [17] for a systematic treatment. The project has been initiated in [4], a full exposition of techniques in the case of a one-component asymmetric Ginzburg–Landau model is presented in [5]. Here we investigate the simplest possible, but nontrivial two-component lattice gas with collisions, further models are to be discussed in a forthcoming paper [6]. Since the underlying PDE theory is restricted to one space dimension, we also have to be satisfied with such models. The proof is based on a strict control of entropy pairs at the microscopic level as prescribed by P. Lax, L. Tartar and F. Murat for approximate solutions to hyperbolic conservation laws. A Lax entropy is macroscopically conserved along classical solutions, but the microscopic system can not have any extra conservation law, thus we are facing with rapidly oscillating quantities. These oscillations are to be controlled by means of logarithmic Sobolev inequality estimates, and effective bounds are obtainable only if the symmetric part of the microscopic evolution is strong enough. That is why the *microscopic viscosity* of the model goes to infinity, i.e. the model is changed when we rescale it. Of course, the *macroscopic viscosity* vanishes in the limit and thus the effect of speeding up the symmetric part of the microscopic infinitesimal generator is not seen in the hydrodynamic limit.

Unfortunately, compensated compactness yields only *existence* of weak solutions, the Lax entropy condition is not sufficient for weak *uniqueness* in the case of two component systems. That is why we can prove convergence of the conserved fields to the set of entropy solutions only, we do not know whether this set consists of a single trajectory specified by its initial data. Let us remark that [15] has the same difficulty concerning the derivation of the incompressible Navier–Stokes equation in 3 space dimensions. The Oleinik type conditions of weak uniqueness are out of reach of our methods because they require a one sided uniform Lipschitz continuity of the Riemann invariants of the macroscopic system, see [1] for most recent results of PDE theory in this direction. It is certainly not easy to get such bounds at the microscopic level.

The paper is organized as follows. The microscopic model and the macroscopic equations are introduced in the next two sections. The main result and its conditions are formulated in Section 4. Proofs are presented in Section 5, while some technical details are postponed to the Appendix.

2 Microscopic model

2.1 State space, conserved quantities, infinitesimal generator

We consider a pair of coupled asymmetric exclusion processes on the discrete torus, particles move with an average speed $+1$ and -1, respectively. Since we allow at most one particle per site, the individual state space consists of three elements. There is another effect in the interaction, something like a collision: if two particles of opposite velocities meet at neighboring sites, then they are also exchanged after some exponential holding times. We can associate velocities ± 1 to particles according to their categories, thus particle number and momentum are the natural conserved quantities; the numbers of $+1$ and -1 particles could have been another choice.

Throughout this paper we denote by $\mathbb{T}^n$ the discrete torus $\mathbb{Z}/n\mathbb{Z}$, $n \in \mathbb{N}$, and by $\mathbb{T}$ the continuous torus $\mathbb{R}/\mathbb{Z}$. The local spin space is $S = \{-1, 0, 1\}$. The state space of the interacting particle system of size n is

$$\Omega^n := S^{\mathbb{T}^n}.$$

Configurations will generally be denoted as

$$\underline{\omega} := (\omega_j)_{j \in \mathbb{T}^n} \in \Omega^n,$$

We need to separate the symmetric (reversible) part of the dynamics. This will be speeded up sufficiently in order to enhance convergence to local equilibrium also at a mesoscopic scale. The phenomenon of compensated compactness is materialized at this scale in the hydrodynamic limiting procedure. So (somewhat artificially) we consider separately the asymmetric and symmetric parts of the rate functions $r : S \times S \to \mathbb{R}_+$, respectively, $s : S \times S \to \mathbb{R}_+$. The dynamics of the system consists of elementary jumps exchanging nearest neighbor spins: $(\omega_j, \omega_{j+1}) \to (\omega'_j, \omega'_{j+1}) = (\omega_{j+1}, \omega_j)$, performed with rate $\lambda r(\omega_j, \omega_{j+1}) + \kappa s(\omega_j, \omega_{j+1})$, where $\lambda, \kappa > 0$ are speed-up factors, depending on the size of the system in the limiting procedure.

The rate functions are chosen as follows:

$$\begin{aligned} r(1,-1) &= 0, & r(-1,1) &= 2, \\ r(0,-1) &= 0, & r(-1,0) &= 1, \\ r(1,0) &= 0, & r(0,1) &= 1, \end{aligned}$$

that is the rate of collisions is twice as large as that of simple jumps, and

$$r(\omega_j, \omega_{j+1}) = \omega_j^-(1 - \omega_{j+1}^-) + \omega_{j+1}^+(1 - \omega_j^+)\,,$$

where $\omega_j^+ := 1\!\!1_{\{\omega_j = 1\}}$, $\omega_j^- := 1\!\!1_{\{\omega_j = -1\}}$ and $1\!\!1_A$ denotes the indicator of a set A. The rates of the symmetric component are simply

$$s(\omega_j, \omega_{j+1}) = 1\!\!1_{\{\omega_j \neq \omega_{j+1}\}}\,.$$

The rates r define a *totally asymmetric* dynamics, while the rates s define a *symmetric* one. The infinitesimal generators defined by these rates are:

$$\begin{aligned} L^n f(\underline{\omega}) &:= \sum_{j \in \mathbb{T}^n} r(\omega_j, \omega_{j+1})(f(\Theta_{j,j+1}\underline{\omega}) - f(\underline{\omega})) \\ K^n f(\underline{\omega}) &= \sum_{j \in \mathbb{T}^n} s(\omega_j, \omega_{j+1})(f(\Theta_{j,j+1}\underline{\omega}) - f(\underline{\omega}))\,, \end{aligned}$$

where $\Theta_{i,j}$ is the spin-exchange operator,

$$(\Theta_{i,j}\underline{\omega})_k = \begin{cases} \omega_j & \text{if} \quad k = i \\ \omega_i & \text{if} \quad k = j \\ \omega_k & \text{if} \quad k \neq i, j. \end{cases}$$

Recall that periodic boundary conditions are assumed in the definition of L^n and K^n.

To get exactly the familiar Leroux system (4) as the limit, the two conserved quantities, η and ξ should be chosen as

$$\eta_j = \eta(\omega_j) := 1 - |\omega_j| \quad \text{and} \quad \xi_j = \xi(\omega_j) := \omega_j.$$

The microscopic dynamics of the model has been defined so that $\sum_j \xi_j$ and $\sum_j \eta_j$ are conserved, we shall see that there is no room for other (independent) hidden conserved observables. In terms of the conservative quantities we have

$$r(\omega_j,\omega_{j+1}) = \frac{1}{4}(1-\eta_j-\xi_j)(1+\eta_{j+1}+\xi_{j+1}) + \frac{1}{4}(1+\eta_j-\xi_j)(1-\eta_{j+1}+\xi_{j+1})\,. \tag{1}$$

The rate functions are so chosen that the product measures

$$\pi^n_{\rho,u}(\underline{\omega}) = \prod_{j\in\mathbb{T}^n} \pi_{\rho,u}(\omega_j),$$

with one-dimensional marginals

$$\pi_{\rho,u}(0)=\rho, \quad \pi_{\rho,u}(\pm 1) = \frac{1-\rho\pm u}{2}.$$

are stationary in time. We shall call these Gibbs measures. The parameters take values from the set

$$\mathcal{D} := \{(\rho,u)\in[0,1]\times[-1,1] \,:\, \rho+|u|\le 1\},$$

and the uniform $\pi^n := \pi^n_{1/3,0}$ will serve as a reference measure. Due to conservations, the stationary measures $\pi^n_{\rho,u}$ are not ergodic. Expectation with respect to the measures $\pi^n_{\rho,u}$ will be denoted by $\mathbf{E}_{\rho,u}(\cdot)$. In particular, given a local observable $v_i := v(\omega_{i-m},\ldots,\omega_{i+m})$ with m fixed, its equilibrium expectation will be denoted as

$$\Upsilon(\rho,u) := \mathbf{E}_{\rho,u}(v_i).$$

The system of microscopic size n will be driven by the infinitesimal generator

$$G^n = nL^n + n^2\sigma K^n,$$

where $\sigma=\sigma(n)$ is the *macroscopic viscosity*, the factor $n\sigma(n)$ can be interpreted as the *microscopic viscosity*. A priori we require that $\sigma(n)\ll 1$ as $n\to\infty$. A very important restriction, $\sqrt{n}\sigma(n)\gg 1$ will be imposed on $\sigma(n)$, see condition (A) in subsection 4.2.

Let μ^n_0 be a probability distribution on Ω^n , which is the initial distribution of the microscopic system of size n, and denote

$$\mu^n_t := \mu^n_0 e^{tG^n}$$

the distribution of the system at (macroscopic) time t. The Markov process on the state space Ω^n driven by the infinitesimal generator G^n, started with initial distribution μ^n_0 will be denoted by $\mathcal{X}^n_t$.

2.2 Fluxes

Elementary computations show that the infinitesimal generators L^n and K^n act on the conserved quantities as follows, see (1).

$$\begin{aligned}
L^n\eta_i = & -\psi(\omega_i,\omega_{i+1}) + \psi(\omega_{i-1},\omega_i) & =: -\psi_i + \psi_{i-1},\\
L^n\xi_i = & -\phi(\omega_i,\omega_{i+1}) + \phi(\omega_{i-1},\omega_i) & =: -\phi_i + \phi_{i-1},\\
K^n\eta_i = & -\psi^s(\omega_i,\omega_{i+1}) + \psi^s(\omega_{i-1},\omega_i) & =: -\psi^s_i + \psi^s_{i-1},\\
K^n\xi_i = & -\phi^s(\omega_i,\omega_{i+1}) + \phi^s(\omega_{i-1},\omega_i) & =: -\phi^s_i + \phi^s_{i-1},
\end{aligned}$$

where

$$\begin{aligned}
\psi_i &= r(\omega_i,\omega_{i+1})\,(\eta_i-\eta_{i+1})\\
&= \frac{1}{2}\{\eta_i\xi_{i+1}+\eta_{i+1}\xi_i\} + \frac{1}{2}\{\eta_i-\eta_{i+1}\}\\
\phi_i &= r(\omega_i,\omega_{i+1})\,(\xi_i-\xi_{i+1})\\
&= \frac{1}{2}\{\eta_i+\eta_{i+1}-2+2\xi_i\xi_{i+1}\} + \frac{1}{2}\{\xi_{i+1}\eta_i-\xi_i\eta_{i+1}\} + \{\xi_i-\xi_{i+1}\},\\
\psi^s_i &= \eta_i-\eta_{i+1},\\
\phi^s_i &= \xi_i-\xi_{i+1}.
\end{aligned} \tag{2}$$

Note that the microscopic fluxes of the conserved observables induced by the symmetric rates $s(\omega_j,\omega_{j+1})$ are (discrete) gradients of the corresponding conserved variables.

It is easy to compute the macroscopic fluxes:

$$\begin{aligned}
\Psi(\rho,u) &:= \mathbf{E}_{\rho,u}(\psi_j) = \rho u\\
\Phi(\rho,u) &:= \mathbf{E}_{\rho,u}(\phi_j) = \rho + u^2 - 1
\end{aligned} \tag{3}$$

3 Leroux's equation – a short survey

Having the macroscopic fluxes (3) computed, the Euler equations of the system considered are expected to be

$$\begin{cases} \partial_t\rho + \partial_x(\rho u) = 0\\ \partial_t u + \partial_x(\rho + u^2) = 0. \end{cases} \tag{4}$$

with given initial data

$$u(0,x) = u_0(x), \quad \rho(0,x) = \rho_0(x). \tag{5}$$

This is exactly Leroux's equation well known in the PDE literature, see [17]. In the present section we shortly review the main facts about this PDE. The first striking fact is that such equations may have classical solutions only for some special initial data, in general shocks are developed in a finite time. Therefore solutions should be understood in a weak (distributional) sense, and there are many weak solutions for the same initial values.

The following vectorial notations sometimes make our formulas more compact:

$$\boldsymbol{u} := \begin{pmatrix} \rho \\ u \end{pmatrix}, \quad \boldsymbol{\Phi} := \begin{pmatrix} \Psi \\ \Phi \end{pmatrix},$$

$$\nabla := \begin{pmatrix} \dfrac{\partial}{\partial \rho} & \dfrac{\partial}{\partial u} \end{pmatrix}, \qquad \nabla^2 := \begin{pmatrix} \dfrac{\partial^2}{\partial \rho^2} & \dfrac{\partial^2}{\partial \rho \partial u} \\ \dfrac{\partial^2}{\partial \rho \partial u} & \dfrac{\partial^2}{\partial u^2} \end{pmatrix}$$

We shall use alternatively, at convenience, the compact vectorial and the explicit notation.

3.1 Lax entropy pairs

In the case of classical solutions (4) can be written as $\partial_t \boldsymbol{u} + D(\boldsymbol{u}) \partial_x \boldsymbol{u} = 0$, where

$$D(\rho, u) := \nabla \boldsymbol{\Phi}(\rho, u) = \begin{pmatrix} u & \rho \\ 1 & 2u \end{pmatrix}$$

is the matrix of the linearized system. The eigenvalues of D are just

$$\lambda = \lambda(\rho, u) := u + \frac{1}{2} \left\{ \sqrt{u^2 + 4\rho} + u \right\},$$
$$\mu = \mu(\rho, u) := u - \frac{1}{2} \left\{ \sqrt{u^2 + 4\rho} - u \right\}.$$

This means that (4) is *strictly hyperbolic* in the domain

$$\{(\rho, u) : \rho \geq 0, u \in \mathbb{R}, (\rho, u) \neq (0,0)\},$$

with marginal degeneracy (i.e. coincidence of the two characteristic speeds, $\lambda = \mu$) at the point $(\rho, u) = (0,0)$.

Lax entropy/flux pairs $\big(S(\boldsymbol{u}), F(\boldsymbol{u})\big)$ are solutions of the linear hyperbolic system $\nabla F(\boldsymbol{u}) = \nabla S(\boldsymbol{u}) \cdot \nabla \boldsymbol{\Phi}(\boldsymbol{u})$, that is $\partial_t S(\boldsymbol{u}) + \partial_x F(\boldsymbol{u}) = 0$ along classical solutions. This means that an entropy S is a conserved observable. In our particular case this reads

$$\begin{cases} F'_\rho = & uS'_\rho + S'_u, \\ F'_u = & \rho S'_\rho + 2uS'_u. \end{cases} \tag{6}$$

or, written as a second order linear equation for S:

$$\rho S''_{\rho\rho} + uS''_{\rho u} - S''_{uu} = 0. \tag{7}$$

This equation is known to have many convex solutions, see [10]. We call an entropy/flux pair *convex* if the map $(\rho, u) \mapsto S(\rho, u)$ is convex. In particular, a globally convex Lax entropy/flux pair defined on the whole half plane $\mathbb{R}_+ \times \mathbb{R}$ is

$$S(\rho, u) := \rho \log \rho + \frac{u^2}{2}, \quad F(\rho, u) := u\rho + u\rho \log \rho + \frac{2u^3}{3}.$$

Weak solutions of (6) are called *generalized entropy/flux pairs.* Riemann's method of solving second order linear hyperbolic PDEs in two variables (see Chapter 4 of [8]) and compactness of $\mathcal{D}$ imply that generalized entropy/flux pairs can be approximated pointwise by twice differentiable entropy/flux pairs.

An *entropy solution* of the Cauchy problem (4), (5) is a measurable function $[0, T] \times \mathbb{T} \ni (t, x) \mapsto \boldsymbol{u}(t, x) \in \mathbb{R}_+ \times \mathbb{R}$ which for any convex entropy/flux pair (S, F), and any nonnegative test function $\varphi : [0, T] \times \mathbb{T} \to \mathbb{R}$ with support in $[0, T) \times \mathbb{T}$ satisfies

$$\int_0^T \int_{\mathbb{T}} \left(\partial_t \varphi(t, x) S(\boldsymbol{u}(t, x)) + \partial_x \varphi(t, x) F(\boldsymbol{u}(t, x))\right) dx\, dt \\ + \int_{\mathbb{T}} \varphi(0, x) S(\boldsymbol{u}_0(x))\, dx \geq 0 \tag{8}$$

Note that $S(\rho, u) = \pm\rho, F(\rho, u) = \pm\rho u$, respectively, $S(\rho, u) = \pm u, F(\rho, u) = \pm(\rho + u^2)$ are entropy/flux pairs, thus entropy solutions are (a special class of) weak solutions. Entropy solutions of the Cauchy problem (4), (5) form a (strongly) closed subset of the Lebesgue space $L^p([0, T] \times \mathbb{T},\, dt\, dx) =: L^p_{t,x}$ for any $p \in [1, \infty)$.

3.2 Young measures, measure valued entropy solutions

A Young measure on $([0, T] \times \mathbb{T}) \times \mathcal{D}$ is $\nu = \nu(t, x; d\boldsymbol{v})$, where
(1) for any $(t, x) \in [0, T] \times \mathbb{T}$ fixed, $\nu(t, x; d\boldsymbol{v})$ is a probability measure on

$\mathcal{D}$, and,
(2) for any $A \subset \mathcal{D}$ fixed the map $(t,x) \mapsto \nu(t,x;A)$ is measurable.
Given a probability measure ν on $\mathbb{R}_+ \times \mathbb{R}$, we shall use the notation

$$\langle \nu\,,\, f\rangle := \int_{\mathcal{D}} f(\boldsymbol{v})\,\nu(d\boldsymbol{v}).$$

The set of Young measures will be denoted by $\mathcal{Y}$. A sequence $\nu^n \in \mathcal{Y}$ *converges vaguely* to $\nu \in \mathcal{Y}$, denoted $\nu^n \rightharpoonup \nu$, if for any $f \in C([0,T]\times\mathbb{T}\times\mathcal{D})$

$$\lim_{n\to\infty} \int_0^T \int_{\mathbb{T}} \langle \nu^n(t,x)\,,\, f(t,x,\cdot)\rangle\, dt\, dx = \int_0^T \int_{\mathbb{T}} \langle \nu(t,x)\,,\, f(t,x,\cdot)\rangle\, dt\, dx,$$

or, equivalently, if for any test function $\varphi \in C([0,T]\times\mathbb{T})$ and any $g \in C(\mathcal{D})$

$$\lim_{n\to\infty} \int_0^T \int_{\mathbb{T}} \varphi(t,x)\langle \nu^n(t,x)\,,\, g\rangle\, dt\, dx = \int_0^T \int_{\mathbb{T}} \varphi(t,x)\langle \nu(t,x)\,,\, g\rangle\, dt\, dx.$$

The set $\mathcal{Y}$ of Young measures will be endowed with the vague topology induced by this notion of convergence. $\mathcal{Y}$ endowed with the vague topology is metrizable, separable and compact. We also consider (without explicitly denoting this) the Borel structure on $\mathcal{Y}$, induced by the vague topology.

We say that the Young measure $\nu(t,x;d\boldsymbol{v})$ is *Dirac-type* if there exists a measurable function $\boldsymbol{u} : [0,T]\times\mathbb{T} \to \mathcal{D}$ such that for almost all $(t,x) \in [0,T]\times\mathbb{T}$, $\nu(t,x;d\boldsymbol{v}) = \delta_{\boldsymbol{u}(t,x)}(d\boldsymbol{v})$. We denote the subset of Dirac-type Young measures by $\mathcal{U} \subset \mathcal{Y}$. It is a fact (see Chapter 9 of [17]) that

$$\mathcal{Y} = \overline{\mathrm{co}(\mathcal{Y})} = \overline{\mathrm{co}(\mathcal{U})} = \overline{\mathcal{U}},$$

where 'co' stands for convex hull and closure is meant according to the vague topology.

We say that the Young measure $\nu(t,x;d\boldsymbol{v})$ is a *measure valued entropy solution* of the Cauchy problem (4), (5) iff for any convex entropy/flux pair (S,F) and any positive test function $\varphi : [0,T]\times\mathbb{T} \to \mathbb{R}_+$ with support in $[0,T)\times\mathbb{T}$,

$$\int_0^T \int_{\mathbb{T}} \big(\partial_t\varphi(t,x)\langle \nu(t,x)\,,\, S\rangle + \partial_x\varphi(t,x)\langle \nu(t,x)\,,\, F\rangle\big)\, dx\, dt + \int_{\mathbb{T}} \varphi(0,x)\, S(\boldsymbol{u}_0(x))\, dx \geq 0 \qquad (9)$$

holds true. Measure valued entropy solutions of the Cauchy problem (4), (5) form a (vaguely) *closed* subset of $\mathcal{Y}$.

Clearly, if $\boldsymbol{u} : [0,T] \times \mathbb{T} \to \mathcal{D}$ is an entropy solution of the Cauchy problem (4), (5) in the sense of (8), then the Dirac-type Young measure $\nu(t,x;d\boldsymbol{v}) := \delta_{\boldsymbol{u}(t,x)}(d\boldsymbol{v})$ is a measure valued entropy solution in the sense of (9). The convergence of subsequences of approximate solutions to measure solutions is almost immediate by vague compactness, the crucial issue is to show the Dirac property of measure valued entropy solutions. This is the aim of the theory of compensated compactness.

3.3 Tartar factorization

A probability measure $\nu(d\rho, du)$ on $\mathbb{R}^2$ satisfies the *Tartar factorization* property with respect to a couple (S_i, F_i), $i = 1,2$ of entropy/flux pairs if

$$\langle \nu, S_1F_2 - S_2F_1 \rangle = \langle \nu, S_1 \rangle \langle \nu, F_2 \rangle - \langle \nu, S_2 \rangle \langle \nu, F_1 \rangle \,. \tag{10}$$

Dirac measures certainly posses this property, and in some cases, there is a converse statement, too. The following one-parameter families of entropy/flux pairs play an essential role in the forthcoming argument:

$$\begin{aligned} S_a(\rho,u) &:= \rho + au - a^2\,, \quad & F_a(\rho,u) &:= (a+u)S_a(\rho,u)\,, \\ \bar{S}_a(\rho,u) &:= |\rho + au - a^2|\,, \quad & \bar{F}_a(\rho,u) &:= (a+u)\bar{S}_a(\rho,u)\,, \end{aligned} \tag{11}$$

where the parameter, $a \in \mathbb{R}$. The case of (S_a, F_a) is obvious because it is a linear function of the basic conserved observables and their fluxes. The pair $(\bar{S}_a, \bar{F}_a)$ satisfies (6) in the generalized (weak) sense. This is due to the facts that the line of non-differentiability, $\rho + au - a^2 = 0$, is just a characteristic line of the PDE (6), and $(\bar{S}_a, \bar{F}_a)$ coincides with $(\pm S_a, \pm F_a)$ on the domains $D_\pm := \{\pm(\rho + au - a^2) > 0\}$. See also Proposition 13.1.4 of [17].

Lemma 1. *Suppose that a compactly supported probability measure, ν on $\mathbb{R}_+ \times \mathbb{R}$ satisfies* (10) *for any two entropy/flux pairs of type* (11)*. Then ν is concentrated to a single point, i.e. it is a Dirac mass.*

Proof. This is Exercise 9.1 in [17], where detailed instructions are also added. For the Reader's convenience we reproduce the easy proof.

Define the function $\mathbb{R} \ni a \mapsto g(a)$ by

$$g(a) := \frac{\langle \nu, F_a \rangle}{\langle \nu, S_a \rangle} - a = \frac{\langle \nu, u(\rho + au - a^2) \rangle}{\langle \nu, (\rho + au - a^2) \rangle}. \tag{12}$$

Note that $\mathbb{R} \ni a \mapsto g(a)$ is a rational function

$$g(a) = \frac{\langle \nu, u \rangle a^2 - \langle \nu, u^2 \rangle a - \langle \nu, \rho u \rangle}{(a - a_1)(a - a_2)},$$

with *possible* poles at the real points

$$a_{1,2} = \frac{1}{2} \left\{ \langle \nu, u \rangle \pm \sqrt{\langle \nu, u \rangle^2 + 4 \langle \nu, \rho \rangle.} \right\}$$

Applying (10) to $(S_1, F_1) = (S_a, F_a)$ and $(S_2, F_2) = (\bar{S}_a, \bar{F}_a)$ we obtain

$$g(a) = \frac{\langle \nu, u | \rho + au - a^2 | \rangle}{\langle \nu, | \rho + au - a^2 | \rangle},$$

and hence

$$\sup_{a \in \mathbb{R}} |g(a)| \leq \sup\{|u| : (\rho, u) \in \operatorname{supp}(\nu)\} < \infty.$$

Since $\mathbb{R} \ni a \mapsto g(a)$ is rational function with real (possible) poles and also bounded, we conclude that it is actually constant. Taking $a \to \pm\infty$ in the definition (12), we obtain

$$g(a) \equiv \langle \nu, u \rangle.$$

From the definition (12) it follows immediately that

$$u = \langle \nu, u \rangle, \qquad \nu - \text{a.s.} \tag{13}$$

Next we apply (10) to $(S_1, F_1) = (S_a, F_a)$ and $(S_2, F_2) = (S_b, F_b)$ and get

$$(b - a) \left(\langle \nu, S_a S_b \rangle - \langle \nu, S_a \rangle \langle \nu, S_b \rangle \right) = \langle \nu, S_a \rangle \langle \nu, u S_b \rangle - \langle \nu, S_b \rangle \langle \nu, u S_a \rangle. \tag{14}$$

Using (13), from (14) it follows that for any $a, b \in \mathbb{R}$

$$\langle \nu, S_a S_b \rangle = \langle \nu, S_a \rangle \langle \nu, S_b \rangle.$$

Hence $\langle \nu, \rho^2 \rangle = \langle \nu, \rho \rangle^2$ and, consequently

$$\rho = \langle \nu, \rho \rangle, \qquad \nu - \text{a.s.} \tag{15}$$

also follows. Finally, (13) and (15) imply the statement of the Lemma. □

This lemma establishes that measure-valued solutions satisfying Tartar's factorization property (10) are, in fact, weak solutions.

4 The hydrodynamic limit under Eulerian scaling

4.1 Block averages

We choose a *mesoscopic* block size $l = l(n)$. A priori

$$1 \ll l(n) \ll n,$$

but more serious restrictions will be imposed, see condition (B) in subsection 4.2. and define the *block averages* of local observables in the following way: We fix once for ever a weight function $a : \mathbb{R} \to \mathbb{R}_+$. It is assumed that:
(1) $x \mapsto a(x)$ has support in the compact interval $[-1, 1]$,
(2) it has total weight $\int a(x)\,dx = 1$,
(3) it is even: $a(-x) = a(x)$, and
(4) it is twice continuously differentiable.

Given a local variable v_i its block average *at macroscopic space* x is defined as

$$\widehat{v}^n(x) = \widehat{v}^n(\underline{\omega}, x) := \frac{1}{l} \sum_j a\left(\frac{nx - j}{l}\right) v_j. \tag{16}$$

Note that, since $l = l(n)$, we do not denote explicitly dependence of the block average on the mesoscopic block size l.

We shall use the handy (but slightly abused) notation

$$\widehat{v}^n(t, x) := \widehat{v}^n(\mathcal{X}_t^n, x).$$

This is the empirical block average process of the local observable v_i.

In accordance with the compact vectorial notation introduced at the beginning of Section 3 we shall denote

$$\boldsymbol{\xi}_j := \begin{pmatrix} \eta_j \\ \xi_j \end{pmatrix}, \ \boldsymbol{\phi}_j := \begin{pmatrix} \psi_j \\ \phi_j \end{pmatrix}, \ \widehat{\boldsymbol{\xi}}^n(x) := \begin{pmatrix} \widehat{\eta}^n(x) \\ \widehat{\xi}^n(x) \end{pmatrix}, \ \widehat{\boldsymbol{\phi}}^n(x) := \begin{pmatrix} \widehat{\psi}^n(x) \\ \widehat{\phi}^n(x) \end{pmatrix},$$

and so on.

Let $\widehat{\boldsymbol{\xi}}^n(t, x)$ be the sequence of empirical block average processes of the conserved quantities, as defined above, regarded as elements of $L^1_{t,x} := L^1([0, T] \times \mathbb{T})$. We denote by $\mathbb{P}^n$ the distribution of these in $L^1_{t,x}$:

$$\mathbb{P}^n(A) := \mathbf{P}\left(\widehat{\boldsymbol{\xi}}^n \in A\right), \tag{17}$$

where $A \in L^1_{t,x}$ is (strongly) measurable. Tightness and weak convergence of the sequence of probability measures $\mathbb{P}^n$ will be meant according to the norm (strong) topology of $L^1_{t,x}$. Weak convergence of a subsequence $\mathbb{P}^{n'}$ will be denoted $\mathbb{P}^{n'} \Rightarrow \mathbb{P}$.

Further on, we denote by ν^n the sequence of Dirac-type random Young measures concentrated on the trajectories of the empirical averages $\widehat{\boldsymbol{\xi}}^n(t,x)$ and by $\mathbb{Q}^n$ their distributions on $\mathcal{Y}$:

$$\nu^n(t,x;d\boldsymbol{v}) := \delta_{\widehat{\boldsymbol{\xi}}^n(t,x)}(d\boldsymbol{v}), \qquad \mathbb{Q}^n(A) := \mathbf{P}\left(\nu^n \in A\right), \tag{18}$$

where $A \in \mathcal{Y}$ is (vaguely) measurable. Due to vague compactness of $\mathcal{Y}$, the sequence of probability measures $\mathbb{Q}^n$ is automatically tight. Weak convergence of a subsequence $\mathbb{Q}^{n'}$ will be meant according to the vague topology of $\mathcal{Y}$ and will be denoted $\mathbb{Q}^n \rightharpoonup \mathbb{Q}$. In this case we shall also say that the subsequence of random Young measures $\nu^{n'}$ (distributed according to $\mathbb{Q}^{n'}$) *converges vaguely in distribution* to the random Young measure ν (distributed according to $\mathbb{Q}$), also denoted $\nu^n \rightharpoonup \nu$.

4.2 Main result

All results are valid under the following conditions

(A) The macroscopic viscosity $\sigma = \sigma(n)$ satisfies

$$n^{-1/2} \ll \sigma \ll 1.$$

(B) The mesoscopic block size $l = l(n)$ is chosen so that

$$n^{2/3}\sigma^{1/3} \ll l \ll n\sigma$$

(C) The initial density profiles converge weakly in probability (or, equivalently in any L^p, $1 \leq p < \infty$). That is: for any test function $\boldsymbol{\varphi} : \mathbb{T} \to \mathbb{R} \times \mathbb{R}$

$$\lim_{n\to\infty} \mathbf{E}\Big(\Big| \int_{\mathbb{T}} \boldsymbol{\varphi}(x) \cdot \big(\widehat{\boldsymbol{\xi}}^n(0,x) - \boldsymbol{u}_0(x)\big)\, dx \Big| \Big) = 0.$$

Our main result is the following

Theorem 1. *Conditions (A), (B), and (C) are in force. The sequence of probability measures $\mathbb{P}^n$ on $L^1_{t,x}$, defined in (17) is tight (according to the norm topology of $L^1_{t,x}$). Moreover, if $\mathbb{P}^{n'}$ is a subsequence which converges weakly (according to the norm topology of $L^1_{t,x}$), $\mathbb{P}^{n'} \Rightarrow \mathbb{P}$, then the limit probability measure $\mathbb{P}$ is concentrated on the entropy solutions of the Cauchy problem (4), (5).*

Remark: Assuming *uniqueness* of the entropy solution $\boldsymbol{u}(t,x)$ of the Cauchy problem (4), (5), we could conclude that

$$\widehat{\boldsymbol{\xi}}^n \stackrel{L^1_{t,x}}{\longrightarrow} \boldsymbol{u}, \qquad \text{in probability.}$$

5 Proof

5.1 Outline of proof

We broke up the proof into several subsections according to what we think to be a logical and transparent structure.

In subsection 5.2 we state the precise *quantitative form* of the convergence to local equilibrium: the logarithmic Sobolev inequality valid for our model and Varadhan's large deviation bound on space-time averages of block variables. As main consequence of these we obtain our a priori estimates: the so-called *one-block estimate* and a version of the so-called *two-block estimate*, formulated for spatial derivatives of the empirical block averages. These estimates are of course the main probabilistic ingredients of the further arguments. The proof of these estimates is postponed to the Appendix of the paper.

In subsection 5.3 we write down an identity which turns out to be the stochastic approximation of the PDE (4). Various error terms are defined here which will be estimated in the forthcoming subsections.

In subsection 5.4 we introduce the relevant *Sobolev norms* and by using the previously proved a priori estimates we prove the necessary upper bounds on the apropriate Sobolev norms of the error terms.

In subsection 5.5 we show that choosing a subsequence of the random Young measures (18) which converges vaguely in distribution, the limiting (random) Young measure is almost surely a measure valued entropy solution of the Cauchy problem (4), (5).

Subsection 5.6 contains the stochastic version of the method of *compensated compactness.* It is further broken up into two sub-subsections as follows. In sub-subsection 5.6.1 we present the stochastic version of Murat's Lemma: we prove that for any smooth Lax entropy/flux pair the entropy production process is tight in the Sobolev space $H^{-1}_{t,x}$. In sub-subsection 5.6.2 we apply (an almost sure version of) Tartar's Div-Curl Lemma leading to the desired almost sure factorization property of the limiting random Young measures. Finally, as main consequence of Tartar's Lemma, we conclude that choosing any subsequence of the random Young measures (18) which converges vaguely in distribution, the limit (random) Young measure is almost surely of Dirac type.

The results of subsection 5.5 and sub-subsection 5.6.2 imply the Theorem. The concluding steps are presented in subsection 5.7.

5.2 Local equilibrium and a priori bounds

The hydrodynamic limit relies on macroscopically fast convergence to (local) equilibrium in blocks of mesoscopic size l. Fix the block size l and $(N, Z) \in \mathbb{N} \times \mathbb{Z}$ with the restriction $N + |Z| \leq l$ and denote

$$\begin{aligned} \Omega^l_{N,Z} &:= \big\{\underline{\omega} \in \Omega^l : \sum_{j=1}^{l} \eta_j = N, \sum_{j=1}^{l} \xi_j = Z\big\}, \\ \pi^l_{N,Z}(\underline{\omega}) &:= \pi^l_{\rho,u}(\underline{\omega} \,|\, \sum_{j=1}^{l} \eta_j = N, \sum_{j=1}^{l} \xi_j = Z), \end{aligned}$$

and, for $f : \Omega^l_{N,Z} \to \mathbb{R}$

$$\begin{aligned} K^l_{N,Z} f(\underline{\omega}) &:= \sum_{j=1}^{l-1} \big(f(\Theta_{j,j+1}\underline{\omega}) - f(\underline{\omega})\big), \\ D^l_{N,Z}(f) &:= \frac{1}{2} \sum_{j=1}^{l-1} \mathbf{E}^l_{N,Z} \left(\big(f(\Theta_{j,j+1}\underline{\omega}) - f(\underline{\omega})\big)^2 \right). \end{aligned}$$

In plain words: $\Omega^l_{N,Z}$ is the hyperplane of configurations $\underline{\omega} \in \Omega^l$ with fixed values of the conserved quantities, $\pi^l_{N,Z}$ is the *microcanonical distribution* on this hyperplane, $K^l_{N,Z}$ is the symmetric infinitesimal generator restricted to the hyperplane $\Omega^l_{N,Z}$, and finally $D^l_{N,Z}$ is the Dirichlet form associated to $K^l_{N,Z}$. Note, that $K^l_{N,Z}$ is defined with free boundary conditions. Expectations with respect to the measures $\pi^l_{N,Z}$ are denoted by $\mathbf{E}^l_{N,Z}\big(\,\cdot\,\big)$. The convergence to local equilibrium is *quantitatively controlled* by the following uniform logarithmic Sobolev estimate:

Lemma 2. *There exists a finite constant* $\aleph$ *such that for any* $l \in \mathbb{N}$, $(N,Z) \in \mathbb{N} \times \mathbb{Z}$ *with* $N + |Z| \leq l$ *and any* $h : \Omega^l_{N,Z} \to \mathbb{R}_+$ *with* $\mathbf{E}^l_{N,Z}(h) = 1$ *the following bound holds:*

$$\mathbf{E}^l_{N,Z}\big(h \log h\big) \leq \aleph\, l^2 D^l_{N,Z}\left(\sqrt{h}\right). \tag{19}$$

Remark: In [25] (see also [12]) the similar statement is proved (inter alia) for symmetric simple exclusion process. That proof can be easily adapted to our case. Instead of stirring configurations of two colors we have stirring of configurations of three colors. No really new ideas are involved. For sake of completeness however, we sketch the proof in subsection 6.1 of the Appendix.

The following large deviation bound goes back to Varadhan [23]. See also the monographs [9] and [4].

Lemma 3. *Let* $l \leq n$, $\mathcal{V} : S^l \to \mathbb{R}_+$ *and denote* $\mathcal{V}_j(\underline{\omega}) := \mathcal{V}(\omega_j, \ldots, \omega_{j+l-1})$. *Then for any* $\beta > 0$

$$\frac{1}{n} \sum_{j \in \mathbb{T}^n} \int_0^T \mathbf{E}_{\mu^n_s}\left(\mathcal{V}_j\right)\, ds \leq C \frac{l^3}{\beta n^2 \sigma} + \frac{T}{\beta} \max_{N,Z} \log \mathbf{E}^l_{N,Z}\big(\exp\{\beta \mathcal{V}\}\big) \tag{20}$$

Remarks: (1) Assuming only uniform bound of order l^{-2} on the spectral gap of $K^l_{N,Z}$ (rather than the stronger logarithmic Sobolev inequality (19)) and using Rayleigh-Schrödinger perturbation (see Appendix 3 of [9]) we would get

$$\frac{1}{n} \sum_{j \in \mathbb{T}^n} \int_0^T \mathbf{E}_{\mu^n_s}\left(\mathcal{V}_j\right)\, ds \leq$$

$$C \frac{l^3 \|\mathcal{V}\|_\infty}{n^2 \sigma} + T \|\mathcal{V}\|_\infty \left(\frac{\max_{N,Z} \mathbf{E}^l_{N,Z}\big(\mathcal{V}\big)}{\|\mathcal{V}\|_\infty} + \frac{\max_{N,Z} \mathbf{Var}^l_{N,Z}\big(\mathcal{V}\big)}{4\|\mathcal{V}\|^2_\infty} \right),$$

which wouldn't be sufficient for our needs.

(2) The proof of the bound (20) explicitly relies on the logarithmic Sobolev

inequality (19). It appears in [26] and it is reproduced in several places, see e.g. [4, 5]. We do not repeat it here.

The main probabilistic ingredients of our proof are the following two consequences of Lemma 3. These are variants of the celebrated *one block estimate*, respectively, *two blocks estimate* of Varadhan and co-authors.

Proposition 1. *Assume conditions (A) and (B). Given a local variable v_j there exists a constant C (depending only on v_j) such that the following bounds hold:*

$$\mathbf{E}\Big(\int_0^T \int_{\mathbb{T}} \Big|\widehat{v}^n(s,x) - \Upsilon(\widehat{\boldsymbol{\xi}}^n(s,x))\Big|^2 \, dx\, dt\Big) \quad \leq \quad C\frac{l^2}{n^2\sigma} \tag{21}$$

$$\mathbf{E}\Big(\int_0^T \int_{\mathbb{T}} |\partial_x \widehat{v}^n(s,x)|^2 \, dx\, dt\Big) \quad \leq \quad C\sigma^{-1} \tag{22}$$

The proof of Proposition 1 is postponed to subsection 6.3 in the Appendix. It relies on the large deviation bound (20) and an elementary probability lemma stated in subsection 6.2 of the Appendix.

We shall refer to (21) as the *block replacement bound* and to (22) as the *gradient bound.*

5.3 The basic identity

Given a smooth function $f : \mathcal{D} \to \mathbb{R}$ we write

$$\partial_t f(\widehat{\boldsymbol{\xi}}^n(t,x)) = G^n f(\widehat{\boldsymbol{\xi}}^n(t,x)) + \partial_t M_f^n(t,x),$$

where the process $t \mapsto M_f^n(t,x)$ is a martingale. Here and in the future $\partial_t f(\widehat{\boldsymbol{\xi}}^n(t,x))$ and $\partial_t M_f^n(t,x)$ are meant as *distributions* in their time variable.

In this order we compute the action of the infinitesimal generator $G^n = nL^n + n^2\sigma K^n$ on $f(\widehat{\boldsymbol{\xi}}^n(x))$. First we compute the asymmetric part:

$$nL^n f(\widehat{\boldsymbol{\xi}}^n(x)) \quad = \quad -\nabla f(\widehat{\boldsymbol{\xi}}^n(x)) \cdot \partial_x \widehat{\boldsymbol{\phi}}^n(x) + A_f^{1,n}(x) \tag{23}$$

where

$$A_f^{1,n}(x) = A_f^{1,n}(\underline{\omega}, x) := n \sum_{j\in\mathbb{T}} r(\omega_j, \omega_{j+1}) \times \tag{24}$$

$$\Big\{ f\big(\widehat{\boldsymbol{\xi}}^n(x) - \frac{1}{l}\big(a(\frac{nx-j}{l}) - a(\frac{nx-j-1}{l})\big)\big(\boldsymbol{\xi}_j - \boldsymbol{\xi}_{j+1}\big)\big) - f\big(\widehat{\boldsymbol{\xi}}^n(x)\big) + \frac{1}{l^2}\, a'(\frac{nx-j}{l}) \nabla f\big(\widehat{\boldsymbol{\xi}}^n(x)\big) \cdot \big(\boldsymbol{\xi}_j - \boldsymbol{\xi}_{j+1}\big) \Big\}.$$

See formula (2) for the definition of $\boldsymbol{\phi}$. $A^{1,n}_f$ is a numerical error term which will be easy to estimate.

Next, the symmetric part:

$$n^2\sigma K^n f(\widehat{\boldsymbol{\xi}}^n(x)) \quad = \quad \sigma\nabla f(\widehat{\boldsymbol{\xi}}^n(x))\cdot\partial_x^2\widehat{\boldsymbol{\xi}}^n(x) + A^{2,n}_f(x) \tag{25}$$

where

$$\begin{aligned} &A^{2,n}_f(x) = A^{2,n}_f(\underline{\omega},x) := \\ &n^2\sigma\sum_{j\in\mathbb{T}}\Big\{f\big(\widehat{\boldsymbol{\xi}}^n(x) - \frac{1}{l}\big(a(\frac{nx-j}{l}) - a(\frac{nx-j-1}{l})\big)\big(\boldsymbol{\xi}_j - \boldsymbol{\xi}_{j+1}\big)\big) \\ &\qquad\qquad - f\big(\widehat{\boldsymbol{\xi}}^n(x)\big) \quad + \frac{1}{l^3}a''(\frac{nx-j}{l})\nabla f\big(\widehat{\boldsymbol{\xi}}^n(x)\big)\cdot\boldsymbol{\xi}_j\Big\}. \end{aligned} \tag{26}$$

This is another numerical error term easy to estimate.

Hence our *basic identity*

$$\begin{aligned} &\partial_t f(\widehat{\boldsymbol{\xi}}^n(t,x)) + \nabla f(\widehat{\boldsymbol{\xi}}^n(t,x))\cdot\nabla\boldsymbol{\Phi}(\widehat{\boldsymbol{\xi}}^n(t,x))\cdot\partial_x\widehat{\boldsymbol{\xi}}^n(t,x) = \\ &\qquad\sum_{i=1}^{2}\Big(A^{i,n}_f(t,x) + B^{i,n}_f(t,x) + C^{i,n}_f(t,x)\Big) + \partial_t M^n_f(t,x). \end{aligned} \tag{27}$$

The various terms on the right hand side are

$$B^{1,n}_f(x) = B^{1,n}_f(\underline{\omega},x) := \partial_x\Big\{\nabla f(\widehat{\boldsymbol{\xi}}^n(x))\cdot\big(\boldsymbol{\Phi}(\widehat{\boldsymbol{\xi}}^n(x)) - \widehat{\boldsymbol{\phi}}^n(x)\big)\Big\} \tag{28}$$

$$B^{2,n}_f(x) = B^{2,n}_f(\underline{\omega},x) := \sigma\partial_x^2 f(\widehat{\boldsymbol{\xi}}^n(x)) = \partial_x\Big\{\sigma\nabla f(\widehat{\boldsymbol{\xi}}^n(x))\cdot\partial_x\widehat{\boldsymbol{\xi}}^n(x)\Big\} \tag{29}$$

$$C^{1,n}_f(x) = C^{1,n}_f(\underline{\omega},x) := -\big(\partial_x\widehat{\boldsymbol{\xi}}^n(x)\big)^\dagger\cdot\nabla^2 f(\widehat{\boldsymbol{\xi}}^n(x))\cdot\big(\boldsymbol{\Phi}(\widehat{\boldsymbol{\xi}}^n(x)) - \widehat{\boldsymbol{\phi}}^n(x)\big) \tag{30}$$

$$C^{2,n}_f(x) = C^{2,n}_f(\underline{\omega},x) := -\sigma\big(\partial_x\widehat{\boldsymbol{\xi}}^n(x)\big)^\dagger\cdot\nabla^2 f(\widehat{\boldsymbol{\xi}}^n(x))\cdot\big(\partial_x\widehat{\boldsymbol{\xi}}^n(x)\big) \tag{31}$$

and

$$\begin{aligned} A^{i,n}_f(t,x) &:= A^{i,n}_f(\mathcal{X}^n_t,x), \\ B^{i,n}_f(t,x) &:= B^{i,n}_f(\mathcal{X}^n_t,x), \\ C^{i,n}_f(t,x) &:= C^{i,n}_f(\mathcal{X}^n_t,x). \end{aligned}$$

In the present paper we shall apply the basic identity (27) only for Lax entropies $f(\boldsymbol{u}) = S(\boldsymbol{u})$. In this special case the left hand side gets the form of a conservation law:

$$\partial_t S(\widehat{\boldsymbol{\xi}}^n(t,x)) + \partial_x F(\widehat{\boldsymbol{\xi}}^n(t,x)) = \sum_{i=1}^{2}\Big(A_S^{i,n}(t,x) + B_S^{i,n}(t,x) + C_S^{i,n}(t,x)\Big) + \partial_t M_S^n(t,x), \tag{32}$$

5.4 Bounds

We fix $T < \infty$ and use the L^p norms

$$\|g\|_{L^p_{t,x}}^p := \int_0^T \int_{\mathbb{T}} |g(t,x)|^p \, dx\, dt$$

and the Sobolev norms

$$\|g\|_{W^{-1,p}_{t,x}} := \sup\Big\{ \int_0^T \int_{\mathbb{T}} \varphi(t,x) g(t,x)\, dx\, dt \;:\; \|\partial_t \varphi\|_{L^q_{t,x}}^q + \|\partial_x \varphi\|_{L^q_{t,x}}^q \le 1 \Big\}$$

where $p^{-1} + q^{-1} = 1$ and $\varphi : [0,T]\times\mathbb{T} \to \mathbb{R}$ is a test function. We use the standard notation $W^{-1,2}_{t,x} =: H^{-1}_{t,x}$.

Remark on notation: The numerical error terms $A_f^{i,n}(t,x)$, $i = 1,2$, will be estimated in $L^\infty_{t,x}$ norm. In these estimates only Taylor expansion bounds are used, no probabilistic argument is involved. The more sophisticated terms $B_f^{i,n}(t,x)$, $i = 1,2$, respectively, $C_f^{i,n}(t,x)$, $i = 1,2$, will be estimated in $H^{-1}_{t,x}$, respectively, $L^1_{t,x}$ norms. The martingale derivative $\partial_t M_f^n(t,x)$ will be estimated in $H^{-1}_{t,x}$ norm.

By straightforward numerical estimates (which do not rely on any probabilistic arguments) we obtain

Lemma 4. *Assume conditions (A) and (B). Let $f : \mathcal{D} \to \mathbb{R}$ be a twice continuously differentiable function with bounded derivatives. Then almost surely*

$$\Big\|A_f^{1,n}\Big\|_{L^\infty_{t,x}} = o(1) \quad \text{and} \quad \Big\|A_f^{2,n}\Big\|_{L^\infty_{t,x}} = o(1)$$

as $n \to \infty$.

Proof. Indeed, using nothing more than Taylor expansion and boundedness of the local variables we readily obtain

$$\sup_{x\in\mathbb{T}} \sup_{\underline{\omega}\in\Omega^n} \Big|A_f^{1,n}(\underline{\omega},x)\Big| \le C\frac{n}{l^2} = o(1) \tag{33}$$

$$\sup_{x\in\mathbb{T}} \sup_{\underline{\omega}\in\Omega^n} \Big|A_f^{2,n}(\underline{\omega},x)\Big| \le C\frac{n^2\sigma}{l^3} = o(1). \tag{34}$$

We omit the tedious but otherwise straightforward details. $\square$

Applying Proposition 1 we obtain the following more sophisticated bounds

Lemma 5. *Assume conditions (A) and (B). Let $f : \mathcal{D} \to \mathbb{R}$ be a twice continuously differentiable function with bounded derivatives. The following asymptotics hold, as $n \to \infty$:*

$$(i) \qquad \mathbf{E}\left(\left\|B_f^{1,n}\right\|_{H^{-1}_{t,x}}\right) = o(1)$$

$$(ii) \qquad \mathbf{E}\left(\left\|B_f^{2,n}\right\|_{H^{-1}_{t,x}}\right) = o(1)$$

$$(iii) \qquad \mathbf{E}\left(\left\|C_f^{1,n}\right\|_{L^{1}_{t,x}}\right) = o(1)$$

$$(iv) \qquad \mathbf{E}\left(\left\|C_f^{2,n}\right\|_{L^{1}_{t,x}}\right) = \mathcal{O}(1)$$

Proof.
(i) We use the block replacement bound (21):

$$\begin{aligned}
&\mathbf{E}\Big(\Big|\int_0^T\int_{\mathbb{T}} v(t,x)B_f^{1,n}(t,x)\,dx\,dt\Big|\Big) \\
&\quad= \mathbf{E}\Big(\Big|\int_0^T\int_{\mathbb{T}} \partial_x v(t,x)\nabla f(\widehat{\boldsymbol{\xi}}^n(t,x))\cdot\big(\boldsymbol{\Phi}(\widehat{\boldsymbol{\xi}}^n(t,x)) - \widehat{\boldsymbol{\phi}}^n(t,x)\big)\,dx\,dt\Big|\Big) \\
&\quad\le \sup_{\boldsymbol{u}\in\mathcal{D}}|\nabla f(\boldsymbol{u})|\,\|\partial_x v\|_{L^2_{t,x}}\,\mathbf{E}\Big(\int_0^T\int_{\mathbb{T}}\Big|\boldsymbol{\Phi}(\widehat{\boldsymbol{\xi}}^n(t,x)) - \widehat{\boldsymbol{\phi}}^n(t,x)\Big|^2\,dx\,dt\Big)^{1/2} \\
&\quad\le C\,\|\partial_x v\|_{L^2_{t,x}}\,\frac{l}{n\sqrt{\sigma}}.
\end{aligned}$$

(ii) We use the gradient bound (22):

$$\begin{aligned}
&\mathbf{E}\Big(\Big|\int_0^T\int_{\mathbb{T}} v(t,x)B_f^{2,n}(t,x)\,dx\,dt\Big|\Big) \\
&\quad= \mathbf{E}\Big(\Big|\int_0^T\int_{\mathbb{T}} \partial_x v(t,x)\nabla f(\widehat{\boldsymbol{\xi}}^n(t,x))\cdot\sigma\big(\partial_x\widehat{\boldsymbol{\xi}}^n(t,x)\big)\,dx\,dt\Big|\Big) \\
&\quad\le \sup_{\boldsymbol{u}\in\mathcal{D}}|\nabla f(\boldsymbol{u})|\,\|\partial_x v\|_{L^2_{t,x}}\,\sigma\mathbf{E}\Big(\int_0^T\int_{\mathbb{T}}\Big|\partial_x\widehat{\boldsymbol{\xi}}^n(t,x)\Big|^2\,dx\,dt\Big)^{1/2} \\
&\quad\le C\,\|\partial_x v\|_{L^2_{t,x}}\,\sigma^{1/2}.
\end{aligned}$$

(iii) We use both, the block replacement bound (21) and the gradient bound (22):

$$\begin{aligned}
&\mathbf{E}\Big(\int_0^T\int_{\mathbb{T}}\Big|C_f^{1,n}(t,x)\Big|\,dx\,dt\Big)\\
&\qquad\le \sup_{\boldsymbol{u}\in\mathcal{D}}\big|\nabla^2 f(\boldsymbol{u})\big|\,\mathbf{E}\Big(\int_0^T\int_{\mathbb{T}}\Big|\widehat{\boldsymbol{\phi}}^n(s,x)-\boldsymbol{\Phi}(\widehat{\boldsymbol{\xi}}^n(s,x))\Big|^2\,dx\,dt\Big)^{1/2}\times\\
&\qquad\qquad\qquad\mathbf{E}\Big(\int_0^T\int_{\mathbb{T}}\Big|\partial_x\widehat{\boldsymbol{\xi}}^n(s,x)\Big|^2\,dx\,dt\Big)^{1/2}\\
&\qquad\le C\frac{l}{n\sigma}.
\end{aligned}$$

(iv) We use again the gradient bound (22):

$$\begin{aligned}
&\mathbf{E}\Big(\int_0^T\int_{\mathbb{T}}\Big|C_f^{2,n}(t,x)\Big|\,dx\,dt\Big)\\
&\qquad\le \sup_{\boldsymbol{u}\in\mathcal{D}}\big|\nabla^2 f(\boldsymbol{u})\big|\,\sigma\mathbf{E}\int_0^T\int_{\mathbb{T}}\Big|\partial_x\widehat{\boldsymbol{\xi}}^n(s,x)\Big|^2\,dx\,dt\\
&\qquad\le C.
\end{aligned}$$

□

Lemma 6. *Assume conditions (A) and (B). Let $f:\mathcal{D}\to\mathbb{R}$ be a twice continuously differentiable function with bounded derivatives. There exists a constant C (depending only on f) such that the following asymptotics holds as $n\to\infty$:*

$$\mathbf{E}\left(\big\|\partial_t M_f^n\big\|_{H^{-1}_{t,x}}\right)=o(1)$$

Proof. Since

$$\big\|\partial_t M_f^n\big\|^2_{H^{-1}_{t,x}}\le\big\|M_f^n\big\|^2_{L^2_{t,x}},$$

we have to bound the expectation of the right hand side.

$$\mathbf{E}\Big(\int_0^T\int_{\mathbb{T}}\big(M_f^n(t,x)\big)^2\,dx\,dt\Big)=\mathbf{E}\Big(\int_0^T\int_{\mathbb{T}}\langle M_f^n(t,x)\rangle\,dx\,dt\Big),$$

where $t\mapsto\langle M_f^n(t,x)\rangle$ is the conditional variance process of the martingale $M_f^n(t,x)$:

$$\begin{aligned}
\langle M_f^n(t,x)\rangle \quad=\quad & n\left(L^n f^2(\widehat{\boldsymbol{\xi}}^n(t,x))-2f(\widehat{\boldsymbol{\xi}}^n(t,x))L^n f(\widehat{\boldsymbol{\xi}}^n(t,x))\right)\\
&+n^2\sigma\left(K^n f^2(\widehat{\boldsymbol{\xi}}^n(t,x))-2f(\widehat{\boldsymbol{\xi}}^n(t,x))K^n f(\widehat{\boldsymbol{\xi}}^n(t,x))\right).
\end{aligned}$$

Using the expressions (23) and (25) we obtain

$$\begin{aligned}\langle M^n_f(t,x)\rangle &= A^{1,n}_{f^2}(t,x) - 2f(\widehat{\boldsymbol{\xi}}^n(t,x))A^{1,n}_f(t,x)\\ &\quad + A^{2,n}_{f^2}(t,x) - 2f(\widehat{\boldsymbol{\xi}}^n(t,x))A^{2,n}_f(t,x).\end{aligned}$$

Hence, by the bounds (33) and (34) (which apply as well of course to the function f^2), we obtain

$$\sup_{t\in[0,T]}\sup_{x\in\mathbb{T}}\langle M^n_f(t,x)\rangle \le C\frac{n^2\sigma}{l^3} = o(1),$$

which proves the lemma. □

5.5 Convergence to measure valued entropy solutions

Proposition 2. *Conditions (A), (B), and (C) are in force. Let $\mathbb{Q}^{n'}$ be a subsequence of the probability distributions defined in (18), which converges weakly in the vague sense: $\mathbb{Q}^{n'} \rightharpoonup \mathbb{Q}$. Then the probability measure $\mathbb{Q}$ is concentrated on the measure valued entropy solutions of the Cauchy problem (4), (5).*

Proof. Due to separability of $C([0,T]\times\mathbb{T})$ it is sufficient to prove that for any convex Lax entropy/flux pair (S,F) and any nonnegative test function φ with support in $[0,T]\times\mathbb{T}$, (9) holds $\mathbb{Q}$-almost-surely. So we fix (S,F) and φ, and denote the real random variable

$$\begin{aligned}X^n &:= -\int_0^T\int_{\mathbb{T}}\varphi(t,x)\big(\partial_t S(\widehat{\boldsymbol{\xi}}^n(t,x)) + \partial_x F(\widehat{\boldsymbol{\xi}}^n(t,x))\big)\,dx\,dt\\ &= \int_0^T\int_{\mathbb{T}}\big(\partial_t\varphi(t,x)\langle\nu^n(t,x)\,,\,S\rangle + \partial_x\varphi(t,x)\langle\nu^n(t,x)\,,\,F\rangle\big)\,dx\,dt\\ &\qquad + \int_{\mathbb{T}}\varphi(0,x)S(\widehat{\boldsymbol{\xi}}^n(0,x))\,dx.\end{aligned}$$

In view of assumption (C), the last term on the right hand side converges to

$$\int_{\mathbb{T}}\varphi(0,x)\,S(\boldsymbol{u}_0(x))\,dx\,,$$

while the space-time integrals are continuous functionals of the Young measure, thus from assumption $\mathbb{Q}^n \rightharpoonup \mathbb{Q}$ it follows that

$$X^n \Rightarrow X, \tag{35}$$

where

$$X \quad := \quad \int_0^T \int_{\mathbb{T}} \big(\partial_t \varphi(t,x)\langle \nu(t,x)\,,\, S\rangle + \partial_x \varphi(t,x)\langle \nu(t,x)\,,\, F\rangle\big)\,dx\,dt$$
$$+ \int_{\mathbb{T}} \varphi(0,x) S(\boldsymbol{u}_0(x))\,dx.$$

and ν is distributed according to $\mathbb{Q}$.

We apply the basic identity (27) specified for $f(\boldsymbol{u}) = S(\boldsymbol{u})$, that is identity (32). It follows that

$$X^n = Y^n + Z^n \tag{36}$$

where

$$\begin{aligned} Y^n \quad &:= \quad \int_0^T \int_{\mathbb{T}} \varphi(t,x) C_S^{2,n}(t,x)\,dx\,dt \\ &= \quad \sigma \int_0^T \int_{\mathbb{T}} \varphi(t,x)\big(\partial_x \widehat{\boldsymbol{\xi}}^n(t,x)\big)^\dagger \cdot \nabla^2 S(\widehat{\boldsymbol{\xi}}^n(t,x)) \cdot \big(\partial_x \widehat{\boldsymbol{\xi}}^n(t,x)\big) \end{aligned}$$

and

$$Z^n := \int_0^T \int_{\mathbb{T}} \varphi(t,x) \left(\sum_{i=1}^{2} \Big(A_S^{i,n} + B_S^{i,n} \Big) + C_S^{1,n} + \partial_t M_S^n \right)(t,x)\,dx\,dt.$$

Due to convexity of S and positivity of φ we have

$$Y^n \geq 0, \qquad \text{almost surely.} \tag{37}$$

On the other hand, from Lemmas 4, 5, 6 we conclude that

$$\lim_{n\to\infty} \mathbf{E}\big(\,|Z^n|\,\big) = 0. \tag{38}$$

Finally, from (35), (36), (37) and (38) the statement of the Proposition follows. □

5.6 Compensated compactness

5.6.1 Murat's lemma

Lemma 7. *Assume conditions (A) and (B). Given a twice continuously differentiable Lax entropy/flux pair (S,F), the sequence*

$$X^n(t,x) := \partial_t S(\widehat{\boldsymbol{\xi}}^n(t,x)) + \partial_x F(\widehat{\boldsymbol{\xi}}^n(t,x))$$

is tight in $H^{-1}_{t,x}$.

Proof. Note that $X^n(t,x)$ is exactly the left hand side of the basic identity (32) and recall that this expression (in particular $\partial_t S(\widehat{\boldsymbol{\xi}}^n(t,x))$) is a random distribution in its t variable.

By definition and a priori boundedness of the domain $\mathcal{D}$, there exists a constant $C < \infty$ such that

$$\mathbf{P}\Big(\|X^n\|_{W^{-1,\infty}_{t,x}} \leq C \Big) = 1. \tag{39}$$

We decompose

$$X^n(t,x) = Y^n(t,x) + Z^n(t,x), \tag{40}$$

where

$$\begin{aligned} Y^n(t,x) &:= B^{1,n}_S(t,x) + B^{2,n}_S(t,x) + \partial_t M^n_S(t,x), \\ Z^n(t,x) &:= A^{1,n}_S(t,x) + A^{2,n}_S(t,x) + C^{1,n}_S(t,x) + C^{2,n}_S(t,x). \end{aligned}$$

For the definitions of the terms $A^{i,n}_S$, $B^{i,n}_S$, $C^{i,n}_S$, $i = 1, 2$, see (24), (26) and (28)–(31).

¿From Lemmas 4, 5 and 6 it follows that

$$\mathbf{E}\Big(\|Y^n\|_{H^{-1}_{t,x}} \Big) \to 0, \tag{41}$$

and

$$\mathbf{E}\Big(\|Z^n\|_{L^1_{t,x}} \Big) \leq C. \tag{42}$$

Further on, from (41), respectively, (42) it follows that for any $\varepsilon > 0$ one can find a *compact* subset K_ε of $H^{-1}_{t,x}$ and a *bounded* subset L_ε of $L^1_{t,x}$ such that

$$\mathbf{P}\Big(Y^n \notin K_\varepsilon\Big) < \varepsilon/2, \qquad \mathbf{P}\Big(Z^n \notin L_\varepsilon\Big) < \varepsilon/2. \tag{43}$$

On the other hand, Murat's lemma (see [13] or Chapter 9 of [17]) says that

$$M_\varepsilon := \big(K_\varepsilon + L_\varepsilon\big) \cap \{X \in H^{-1}_{t,x} : \|X\|_{W^{-1,\infty}_{t,x}} \leq C\}$$

is compact in $H^{-1}_{t,x}$. From (39), (40) and (43) it follows that

$$\mathbf{P}\Big(X^n \notin M_\varepsilon\Big) < \varepsilon,$$

uniformly in n, which proves the lemma. □

5.6.2 Tartar's lemma and its consequence

Lemma 8. *Assume conditions (A) and (B). Let $\mathbb{Q}^{n'}$ be a subsequence of the probability measures on $\mathcal{Y}$ defined in (18), which converges weakly in the vague sense: $\mathbb{Q}^{n'} \rightharpoonup\!\!\!\rightharpoonup \mathbb{Q}$. Then $\mathbb{Q}$ is concentrated on the (vaguely closed) subset of Young measures satisfying* (10). *That is, $\mathbb{Q}$-a.s. for any two generalized Lax entropy/flux pairs (S_1, F_1) and (S_2, F_2) and any test function $\varphi : [0,T] \times \mathbb{T} \to \mathbb{R}$,*

$$\int_0^T \int_{\mathbb{T}} \varphi(t,x) \langle \nu(t,x)\,,\, S_1F_2 - S_2F_1 \rangle\, dx\, dt = \tag{44}$$

$$\int_0^T \int_{\mathbb{T}} \varphi(t,x) \big(\langle \nu(t,x)\,,\, S_1 \rangle \langle \nu(t,x)\,,\, F_2 \rangle - \langle \nu(t,x)\,,\, S_2 \rangle \langle \nu(t,x)\,,\, F_1 \rangle \big)\, dx\, dt.$$

Proof. First we prove (44) for twice continuously differentiable entropy/flux pairs. Due to separability of $C([0,T] \times \mathbb{T})$ it is sufficient to prove that for any two twice continuously differentiable Lax entropy/flux pairs (S_1, F_1) and (S_2, F_2) and any test function $\varphi : [0,T] \times \mathbb{T} \to \mathbb{R}$, (44) holds $\mathbb{Q}$-almost-surely. So we fix (S_1, F_1), (S_2, F_2) and φ. Note that

$$\begin{aligned} X_j^n(t,x) &:= \partial_t S_j(\widehat{\boldsymbol{\xi}}^n(t,x)) + \partial_x F_j(\widehat{\boldsymbol{\xi}}^n(t,x)) \\ &= \partial_t \langle \nu^n(t,x)\,,\, S_j \rangle + \partial_x \langle \nu^n(t,x)\,,\, F_j \rangle \end{aligned}$$

$j = 1, 2$.

Due to Skorohod's representation theorem (see Theorem 1.8 of [3]) and Lemma 7 we can realize the random Young measures $\nu^n(t,x;d\boldsymbol{v})$ and $\nu(t,x;d\boldsymbol{v})$ *jointly* on an enlarged probability space $(\Xi, \mathcal{A}, \mathbf{P})$ so that $\mathbf{P}$-*almost-surely*

$$\nu^{n'} \rightharpoonup \nu, \qquad \text{and} \qquad \{ X_j^{n'} : n',\ j = 1,2 \} \text{ is relatively compact in } H_{t,x}^{-1}.$$

So, applying Tartar's Div-Curl Lemma (see [19], [20], or Chapter 9 of [17]) we conclude that (in this realization) almost surely the factorization (44) holds true.

Since $\mathcal{D}$ is compact, from Riemann's method of solving the linear hyperbolic PDE (7) (see Chapter 4 of [8]) it follows that generalized entropy/flux pairs are approximated pointwise by smooth ones. Thus the Tartar factorization (44) extends from smooth to generalized entropy/flux pairs. Hence the lemma. □

The main consequence of Lemma 8 is the following

Proposition 3. *Assume conditions (A) and (B). Let $\mathbb{Q}^{n'}$ be a subsequence of the probability measures on $\mathcal{Y}$ defined in (18), which converges weakly in the vague sense: $\mathbb{Q}^{n'} \rightharpoonup \mathbb{Q}$. Then the probability measure $\mathbb{Q}$ is concentrated on a set of Dirac-type Young measures, that is $\mathbb{Q}(\mathcal{U}) = 1$.*

Proof. In view of Lemma 8 this is a direct consequence of Lemma 1. □

Remark: This is the only point where we exploit the very special features of the PDE (4). Note that the proof of Lemma 1 relies on elementary explicit computations. In case of general 2×2 hyperbolic systems of conservation laws, instead of these explicit computations we should refer to DiPerna's arguments from [2], possibly further complicated by the existence of singular (non-hyperbolic) points isolated at the boundary of the domain $\mathcal{D}$. More general results will be presented in the forthcoming paper [6].

5.7 End of proof

From Propositions 2 and 3 it follows that from any subsequence n' one can extract a sub-subsequence n'' such that $\mathbb{Q}^{n''} \rightharpoonup \mathbb{Q}$ and $\mathbb{Q}$ is concentrated on the set of Dirac-type measure valued entropy solutions of the Cauchy problem. ¿From now on we denote simply by n this sub-subsequence. Referring again to Skorohod's Representation Theorem we realize the Dirac-type random Young measures $\nu^n_{t,x}(d\boldsymbol{v}) := \delta_{\widehat{\boldsymbol{\xi}}^n(t,x)}(d\boldsymbol{v})$ and $\nu_{t,x}(d\boldsymbol{v}) := \delta_{\boldsymbol{u}(t,x)}(d\boldsymbol{v})$ jointly on an enlarged probability space $(\Xi, \mathcal{A}, \mathbf{P})$, so that $\nu^n \rightharpoonup \nu$ almost surely and $(t,x) \mapsto \boldsymbol{u}(t,x)$ is almost surely entropy solution of the Cauchy problem. From basic functional analytic considerations (see e.g. Chapter 9 of [17]) it follows that, in case that the limit Youg measure is also Dirac-type, the vague convergence $\nu^n \rightharpoonup \nu$ implies strong (i.e. norm) convergence of the underlying functions,

$$\widehat{\boldsymbol{\xi}}^n \to \boldsymbol{u} \quad \text{in} \quad L^1_{t,x}. \tag{45}$$

So, we have realized jointly on the probability space $(\Xi, \mathcal{A}, \mathbf{P})$ the empirical block average processes $\widehat{\boldsymbol{\xi}}^n(t,x)$ and the random function $\boldsymbol{u}(t,x)$ so that the latter one is almost surely entropy solution of the Cauchy problem, and (45) almost surely holds true. This proves the theorem. □

6 Appendix

6.1 The logarithmic Sobolev inequality for random stirring of r colors on the linear graph $\{1, 2, \ldots, l\}$

Let $r \geq 2$ be a fixed integer. For $l \in \mathbb{N}$ we consider r-tuples of integers $N = (N_1, \ldots, N_r)$ such that

$$N_\alpha \geq 0, \quad \alpha = 1, \ldots, r \qquad \text{and} \qquad N_1 + \cdots + N_r = l, \tag{46}$$

$$\Omega_N^l := \Big\{ \underline{\omega} \in \{1, \ldots, r\}^l : \sum_{j=1}^{l} 1\!\!1_{\{\omega_j = \alpha\}} = N_\alpha,\ \alpha = 1, \ldots, r \Big\}.$$

Let π_N^l denote the uniform probability measure on Ω_N^l:

$$\pi_N^l(\underline{\omega}) = \frac{N_1! \cdots N_r!}{l!}, \qquad \underline{\omega} \in \Omega_N^l.$$

The one dimensional marginals of π_N^l are

$$\pi_N^{l,1}(\alpha) = \frac{N_\alpha}{l}.$$

The random element of Ω_N^l distributed according to π_N^l will be denoted $\underline{\zeta} = (\zeta_1, \zeta_2, \ldots, \zeta_l)$. Expectation with respect to π_N^l, respectively, $\pi_N^{l,1}$ will be denoted by $\mathbf{E}_N^l\big(\cdots\big)$, respectively, $\mathbf{E}_N^{l,1}\big(\cdots\big)$. Conditional expectation, given the first coordinate ζ_1 will be denoted $\mathbf{E}_N^l\big(\cdots \big| \zeta_1\big)$. Note that

$$\mathbf{E}_N^l\big(f(\underline{\zeta}) \big| \zeta_1 = \alpha\big) = \mathbf{E}_{N^\alpha}^{l-1}\big(f(\alpha, \zeta_2, \ldots, \zeta_l)\big)$$

where $\mathbf{E}_{N^\alpha}^{l-1}\big(\cdots\big)$ stands for expectation with respect to $(\zeta_2, \ldots, \zeta_l)$ distributed according to $\pi_{N^\alpha}^{l-1}$ and, given $N = (N_1, \ldots, N_\alpha, \ldots, N_r)$ with $N_\alpha \geq 1$, $N^\alpha := (N_1, \ldots, N_\alpha - 1, \ldots, N_r)$.

Given a probability density h over (Ω_N^l, π_N^l), its entropy is

$$H_N^l\big(h\big) := \mathbf{E}_N^l\big(h(\underline{\zeta}) \log h(\underline{\zeta})\big).$$

Further on, for $i, j \in \{1, \ldots, l\}$ let $\Theta_{i,j} : \Omega_N^l \to \Omega_N^l$ be the spin exchange operator

$$(\Theta_{i,j}\underline{\omega})_k = \begin{cases} \omega_j & \text{if } k = i, \\ \omega_i & \text{if } k = j, \\ \omega_k & \text{if } k \neq i, j, \end{cases} .$$

For $f:\Omega^l_N\to\mathbb{R}$ we define the Dirichlet form and the conditional Dirichlet form, given ζ_1

$$\begin{aligned}
D^l_N(f) &:= \frac{1}{2}\sum_{i=1}^{l-1}\mathbf{E}^l_N\left(\left(f(\Theta_{i,i+1}\underline{\zeta})-f(\underline{\zeta})\right)^2\right),\\
D^l_N(f|\zeta_1) &:= \frac{1}{2}\sum_{i=1}^{l-1}\mathbf{E}^l_N\left((f(\Theta_{i,i+1}\underline{\zeta})-f(\underline{\zeta}))^2|\zeta_1\right)\\
&= D^{l-1}_{N^{\zeta_1}}\left(f(\zeta_1,\cdot)\right).
\end{aligned}$$

The logarithmic Sobolev inequality is formulated in the following

Proposition 4. *There exist a finite constant $\aleph$ such that for any number of colors r, any block size $l\in\mathbb{N}$, any distribution of colors $N=(N_1,\dots,N_r)$ satisfying (46) and any probability density h over (Ω^l_N,π^l_N), the following inequality holds:*

$$H^l_N(h)\le\aleph\, l^2 D^l_N(\sqrt{h}). \tag{47}$$

Remark: The proof follows [25] (see also [12]). Due to exchangeability of the measures π^l_N some steps are considerably simpler than there.

Proof. We shall prove the Proposition by induction on l. Denote

$$W(l):=\sup_N\sup_h\frac{H^l_N(h)}{D^l_N(\sqrt{h})}.$$

The following identity is straightforward

$$\begin{aligned}
H^l_N(h) \;=\;& \mathbf{E}^{l,1}_N\left(\mathbf{E}^l_N(h(\underline{\zeta})|\zeta_1)\mathbf{E}^l_N(h_1(\underline{\zeta})\log h_1(\underline{\zeta})|\zeta_1)\right)\\
&+\mathbf{E}^{l,1}_N\left(\mathbf{E}^l_N(h(\underline{\zeta})|\zeta_1)\log\mathbf{E}^l_N(h(\underline{\zeta})|\zeta_1)\right),
\end{aligned}\tag{48}$$

where in the first term of the right hand side

$$h_1(\underline{\zeta}):=\frac{h(\underline{\zeta})}{\mathbf{E}^l_N(h(\underline{\zeta})|\zeta_1)}.$$

First we bound the first term on the right hand side of (48). By the induction hypothesis

$$\begin{aligned}
&\mathbf{E}_N^{l,1}\big(\mathbf{E}_N^l\big(h(\underline{\zeta})\big|\zeta_1\big)\mathbf{E}_N^l\big(h_1(\underline{\zeta})\log h_1(\underline{\zeta})\big|\zeta_1\big)\big) \\
&\quad= \mathbf{E}_N^{l,1}\big(\mathbf{E}_N^l\big(h(\underline{\zeta})\big|\zeta_1\big)\mathbf{E}_{N^{\zeta_1}}^{l-1}\big(h_1(\underline{\zeta})\log h_1(\underline{\zeta})\big)\big) \\
&\quad\le W(l-1)\mathbf{E}_N^{l,1}\big(\mathbf{E}_N^l\big(h(\underline{\zeta})\big|\zeta_1\big)D_{N^{\zeta_1}}^{l-1}\big(\sqrt{h_1}\big)\big) \\
&\quad= W(l-1)\mathbf{E}_N^{l,1}\big(D_{N^{\zeta_1}}^{l-1}\big(\sqrt{h(\zeta_1,\cdot)}\big)\big) \\
&\quad\le W(l-1)D_N^l\big(\sqrt{h}\big). \qquad (49)
\end{aligned}$$

Next we turn to the second term on the right hand side of (48). In order to simplify notation in the next argument we denote

$$\varrho_\alpha := \frac{N_\alpha}{l}, \qquad q_\alpha(j) := \mathbf{E}_N^l\big(h(\underline{\zeta})\mathbb{1}_{\{\zeta_j=\alpha\}}\big). \qquad (50)$$

It is straightforward that for any $0 < K < \infty$ there exists a finite constant $C = C(K)$ such that for any $v \in [0, K]$

$$v\log v \le (v-1) + C\big(\sqrt{v}-1\big)^2$$

and, furthermore, the constant C can be chosen so that for any $v > K$

$$v\log v \le Cv^{3/2}.$$

Hence, with the notation introduced in (50), we get the following upper bound for the second term on the right hand side of (48)

$$\begin{aligned}
&\mathbf{E}_N^{l,1}\Big(\mathbf{E}_N^l\big(h(\underline{\zeta})\big|\zeta_1\big)\log\mathbf{E}_N^l\big(h(\underline{\zeta})\big|\zeta_1\big)\Big) = \sum_{\alpha=1}^r \varrho_\alpha\frac{q_\alpha(1)}{\varrho_\alpha}\log\frac{q_\alpha(1)}{\varrho_\alpha} \qquad (51) \\
&\le C\sum_{\alpha=1}^r \varrho_\alpha\left\{\left(\sqrt{\frac{q_\alpha(1)}{\varrho_\alpha}}-1\right)^2\mathbb{1}_{\{\frac{q_\alpha(1)}{\varrho_\alpha}\le K\}} + \left(\frac{q_\alpha(1)}{\varrho_\alpha}\right)^{3/2}\mathbb{1}_{\{\frac{q_\alpha(1)}{\varrho_\alpha}> K\}}\right\}.
\end{aligned}$$

We use the straightforward inequality

$$\sum_{\alpha=1}^r \varrho_\alpha\left(\frac{q_\alpha(1)}{\varrho_\alpha}-1\right)\mathbb{1}_{\{\frac{q_\alpha(1)}{\varrho_\alpha}\le K\}} \le 0.$$

We choose K sufficiently large in order that Lemma 4.1 of [25] can be applied to $\{1,2,\ldots,l\} \ni j \mapsto \sqrt{q_\alpha(j)/\varrho_\alpha}$. Thus we obtain the upper bound

$$\left(\sqrt{\frac{q_\alpha(1)}{\varrho_\alpha}}-1\right)^2 1\!\!1_{\{\frac{q_\alpha(1)}{\varrho_\alpha}\le K\}} + \left(\frac{q_\alpha(1)}{\varrho_\alpha}\right)^{3/2} 1\!\!1_{\{\frac{q_\alpha(1)}{\varrho_\alpha}>K\}}$$

$$\le C'l\sum_{j=1}^{l-1}\left(\sqrt{\frac{q_\alpha(j+1)}{\varrho_\alpha}}-\sqrt{\frac{q_\alpha(j)}{\varrho_\alpha}}\right)^2, \qquad (52)$$

where C' is again a universal constant. Putting together (51) and (52) and returning to the explicit notation we obtain the following upper bound for the second term on the right hand side of (48):

$$\begin{aligned}
&\mathbf{E}_N^{l,1}\left(\mathbf{E}_N^l\left(h(\underline{\zeta})\big|\zeta_1\right)\log\mathbf{E}_N^l\left(h(\underline{\zeta})\big|\zeta_1\right)\right)\\
&\le C''l\sum_{j=1}^{l-1}\sum_{\alpha=1}^{r}\left(\sqrt{\mathbf{E}_N^l\left(h(\underline{\zeta})1\!\!1_{\{\zeta_{j+1}=\alpha\}}\right)}-\sqrt{\mathbf{E}_N^l\left(h(\underline{\zeta})1\!\!1_{\{\zeta_j=\alpha\}}\right)}\right)^2\\
&= C''l\sum_{j=1}^{l-1}\sum_{\alpha=1}^{r}\left(\sqrt{\mathbf{E}_N^l\left(h(\Theta_{j,j+1}\underline{\zeta})1\!\!1_{\{\zeta_j=\alpha\}}\right)}-\sqrt{\mathbf{E}_N^l\left(h(\underline{\zeta})1\!\!1_{\{\zeta_j=\alpha\}}\right)}\right)^2\\
&= C''l\sum_{j=1}^{l-1}\left(\sqrt{\mathbf{E}_N^l\left(h(\Theta_{j,j+1}\underline{\zeta})\right)}-\sqrt{\mathbf{E}_N^l\left(h(\underline{\zeta})\right)}\right)^2\\
&\le C''l\sum_{j=1}^{l-1}\mathbf{E}_N^l\left(\left(\sqrt{h(\Theta_{j,j+1}\underline{\zeta})}-\sqrt{h(\underline{\zeta})}\right)^2\right) = C''lD_N^l\left(\sqrt{h}\right). \qquad (53)
\end{aligned}$$

In the second step we used *exchangeability* of the canonical measures π_N^l. In the last inequality we note that the map

$$\mathbb{R}_+\times\mathbb{R}_+ \ni (x,y)\mapsto\left(\sqrt{x}-\sqrt{y}\right)^2$$

is *convex* and we use Jensen's inequality.

From (48), (49) and (53) eventually we obtain

$$W(l)\le W(l-1)+C''l,$$

which yields (47). □

6.2 An elementary probability lemma

The contents of the present subsection, in particular Lemma 9 and its Corollary 1 are borrowed from [22]. For their proofs see that paper.

Let (Ω, π) be a finite probability space and ω_i, $i \in \mathbb{Z}$ i.i.d. Ω-valued random variables with distribution π. Further on let

$$\boldsymbol{\xi} : \Omega \to \mathbb{R}^d, \qquad \boldsymbol{\xi}_i := \boldsymbol{\xi}(\omega_i),$$
$$v : \Omega^m \to \mathbb{R}, \qquad v_i := v(\omega_i \dots, \omega_{i+m-1})$$

and denote π^m the product measure on Ω^m with identical marginals π; $\mathbf{E}_{\pi^m}$ is expectation with respect to π^m . For $\boldsymbol{x} \in \mathrm{co}(\mathrm{Ran}(\boldsymbol{\xi}))$ let

$$\Upsilon(\boldsymbol{x}) := \frac{\mathbf{E}_{\pi^m}\big(v_1 \exp\{\sum_{i=1}^m \boldsymbol{\lambda} \cdot \boldsymbol{\xi}_i\}\}\big)}{\mathbf{E}_{\pi^m}\big(\exp\{\boldsymbol{\lambda} \cdot \boldsymbol{\xi}_1\}\}\big)^m},$$

where $\mathrm{co}(\mathrm{Ran}(\boldsymbol{\xi}))$ denotes the convex hull of the range of $\boldsymbol{\xi}$, and $\boldsymbol{\lambda} \in \mathbb{R}^d$ is chosen so that

$$\frac{\mathbf{E}_{\pi^m}\big(\boldsymbol{\xi}_1 \exp\{\boldsymbol{\lambda} \cdot \boldsymbol{\xi}_1\}\}\big)}{\mathbf{E}_{\pi^m}\big(\exp\{\boldsymbol{\lambda} \cdot \boldsymbol{\xi}_1\}\}\big)} = \boldsymbol{x}.$$

For $l \in \mathbb{N}$ we denote *plain* block averages by

$$\overline{\boldsymbol{\xi}}_l := \frac{1}{l} \sum_{j=1}^{l} \boldsymbol{\xi}_j.$$

Finally, let $b : [0, 1] \to \mathbb{R}$ be a fixed smooth function and denote

$$M(b) := \int_0^1 b(s)\, ds.$$

We also define the block averages *weighted by* b as

$$\langle b\,,\, \boldsymbol{\xi} \rangle_l := \frac{1}{l} \sum_{j=0}^{l} b(j/l) \boldsymbol{\xi}_j, \quad \langle b\,,\, v \rangle_l := \frac{1}{l} \sum_{j=0}^{l} b(j/l) v_j,$$

The following lemma relies on elementary probability arguments:

Lemma 9. *There exists a constant $C < \infty$, depending only on m, on the joint distribution of $(v_i, \boldsymbol{\xi}_i)$ and on the function b, such that the following bounds hold uniformly in $l \in \mathbb{N}$ and $\boldsymbol{x} \in (\mathrm{Ran}(\boldsymbol{\xi}) + \cdots + \mathrm{Ran}(\boldsymbol{\xi}))/l$:*
(i) If $M(b) = 0$, then

$$\mathbf{E}\Big(\exp\big\{\gamma \sqrt{l} \langle b\,,\, v \rangle_l\big\} \,\Big|\, \overline{\boldsymbol{\xi}}_l = \boldsymbol{x}\Big) \leq \exp\{C(\gamma^2 + \gamma/\sqrt{l})\}. \tag{54}$$

(ii) If $M(b) = 1$ then

$$\mathbf{E}\Big(\exp\big\{\gamma \sqrt{l}\big(\langle b\,,\, v \rangle_l - \Upsilon(\langle b\,,\, \boldsymbol{\xi} \rangle_l)\big)\big\} \,\Big|\, \overline{\boldsymbol{\xi}}_l = \boldsymbol{x}\Big) \leq \exp\{C(\gamma^2 + \gamma/\sqrt{l})\}. \tag{55}$$

The proof of this lemma appears in [22].

Corollary 1. *There exists a $\gamma_0 > 0$, depending only on m, on the joint distribution of $(v_i, \boldsymbol{\xi}_i)$ and on the function b, such that the following bounds hold uniformly in $l \in \mathbb{N}$ and $\boldsymbol{x} \in (\mathrm{Ran}(\boldsymbol{\xi}) + \cdots + \mathrm{Ran}(\boldsymbol{\xi}))/l$:*
(i) If $M(b) = 0$, then

$$\mathbf{E}\Big(\exp\left\{\gamma_0 l \langle b\,,\, v\rangle_l^2\right\} \Big| \overline{\boldsymbol{\xi}}_l = \boldsymbol{x}\Big) \leq \sqrt{2}. \tag{56}$$

(ii) If $M(b) = 1$ then

$$\mathbf{E}\Big(\exp\left\{\gamma_0 l \big(\langle b\,,\, v\rangle_l - \Upsilon(\langle b\,,\, \boldsymbol{\xi}\rangle_l)\big)^2\right\} \Big| \overline{\boldsymbol{\xi}}_l = \boldsymbol{x}\Big) \leq \sqrt{2}. \tag{57}$$

Proof. The bounds (56) and (57) follow from (54), respectively, (55) by exponential Gaussian averaging. □

6.3 Proof of the a priori bounds (Proposition 1)

6.3.1 Proof of the block replacement bound (21)

We note first that by simple numerical approximation (no probability bounds involved)

$$\left| \int_{\mathbb{T}} \left| \widehat{v}^n(x) - \Upsilon(\widehat{\boldsymbol{\xi}}^n(x)) \right|^2 dx - \frac{1}{n}\sum_{j=1}^{n} \left| \widehat{v}^n(j/n) - \Upsilon(\widehat{\boldsymbol{\xi}}^n(j/n)) \right|^2 \right| \leq C l^{-1} = o\left(\frac{l^2}{n^2\sigma}\right).$$

We apply Lemma 3 with

$$\mathcal{V}_j = \left| \widehat{v}^n(j/n) - \Upsilon(\widehat{\boldsymbol{\xi}}^n(j/n)) \right|^2.$$

We use the bound (57) of Corollary 1 with the function $b = a$ of (16). Note that $\beta = \gamma_0 l$ can be chosen in (20). This yields the bound (21).

6.3.2 Proof of the gradient bound (22)

Again, we start with numerical approximation:

$$\left| \int_{\mathbb{T}} |\partial_x \widehat{v}^n(x)|^2 \, dx - \frac{1}{n}\sum_{j=1}^{n} |\partial_x \widehat{v}^n(j/n)|^2 \right| \leq C\frac{n^2}{l^3} = o(\sigma^{-1}).$$

We apply Lemma 3 with

$$\mathcal{V}_j = |\partial_x \widehat{v}^n(j/n)|^2 .$$

We use now the bound (56) of Corollary 1 with the function $b = a'$, where a is the weighting function from (16). The same choice $\beta = \gamma_0 l$ applies. This will yield the bound (22).

Acknowledgement: We thank the kind hospitality of Institut Henri Poincaré where part of this work was done. We also acknowledge the financial support of the Hungarian Science Foundation (OTKA), grants T26176 and T037685. Thanks are also due to an anonymous referee for carefully reading the manuscript and pointing out some errors in the first submitted version.

References

[1] A. Bressan: *Hyperbolic Systems of Conservation Laws: The One Dimensional Cauchy Problem.* Oxford Lecture Series in Math. Appl. **20**. Oxford 2000.

[2] R.J. DiPerna: Convergence of approximate solutions to conservation laws. *Arch. Rat. Mech. Anal.* **82**: 27–70 (1983)

[3] S.N. Ethier, T.G. Kurtz: *Markov Processes: Characterization and Convergence.* J. Wiley, New York 1986.

[4] J. Fritz: *An Introduction to the Theory of Hydrodynamic Limits.* Lectures in Mathematical Sciences **18**. Graduate School of Mathematics, Univ. Tokyo, 2001.

[5] J. Fritz: Entropy pairs and compensated compactness for weakly asymmetric systems. *Advanced Studies in Pure Mathematics* (2003) (to appear), www.math.bme.hu/ jofri.

[6] J. Fritz, B. Tóth: in preparation (2003)

[7] M.Z. Guo, G.C. Papanicolaou, S.R.S. Varadhan: Nonlinear diffusion limit for a system with nearest neighbour interactions. *Commun. Math. Phys.* **118**: 31–59 (1988)

[8] F. John: *Partial Differential Equations.* Applied Mathematical Sciences, vol. 1, Springer, New York-Heidelberg-Berlin, 1971.

[9] C. Kipnis, C. Landim: *Scaling Limits of Interacting Particle Systems.* Springer, Berlin 1999.

[10] P. Lax: Shock waves and entropy. In: *Contributions to Nonlinear Functional Analysis*,606–634, ed. E.A. Zarantonello. Academic Press, 1971.

[11] P. Lax: *Systems of Conservation Laws and the Mathematical Theory of Shock Waves.* SIAM, CBMS-NSF 11, 1973.

[12] T.-Y. Lee, H.-T. Yau: Logarithmic Sobolev inequality for some models of random walks. *Ann. Probab.* **26**:1855-1873 (1998)

[13] F. Murat: Compacité par compensation. *Ann. Sci. Scuola Norm. Sup. Pisa* **5**: 489-507 (1978)

[14] S. Olla, S.R.S. Varadhan and H.-T. Yau: Hydrodynamic limit for a Hamiltonian system with weak noise. *Commun. Math. Phys.* **155**: 523–560 (1993)

[15] J. Quastel, H.-T. Yau: Lattice gases, large deviations, and the incompressible Navier–Stokes equation. *Ann. Math.* **148**: 51–108, (1998)

[16] F. Rezakhanlou: Hydrodynamic limit for attractive particle systems on $\mathbb{Z}^d$. *Commun. Math. Phys.* **140**: 417-448 (1991)

[17] D. Serre: *Systems of Conservation Laws.* Vol. 1-2. Cambridge University Press, 2000.

[18] J. Smoller: *Shock Waves and Reaction Diffusion Equations*, Second Edition, Springer, New York 1994.

[19] L. Tartar: Compensated compactness and applications to partial differential equations. In: *Nonlinear Analysis and Mechanics, Heriot-Watt Symposium* ed. R.J. Knops, *Pitman Research Notes in Mathematics* **39**: 136–212, 1979.

[20] L. Tartar: The compensated compactness method applied to systems of conservation laws. In: *Systems of Nonlinear PDEs*, ed. J.B. Ball,

NATO ASI Series C/Math. and Phys. Sci. vol **111**: 263–285, Reidel, Dordrecht 1983.

[21] B. Tóth, B. Valkó: Onsager relations and Eulerian hydrodynamic limit for systems with several conservation laws. *J. Statist. Phys.* **112**: 497-521 (2003)

[22] B. Tóth, B. Valkó: Perturbation of singular equilibrium for systems with two conservation laws — hydrodynamic limit. *in preparation* (2003)

[23] S.R.S. Varadhan: Nonlinear diffusion limit for a system with nearest neighbor interactions II. In: *Asymptotic Problems in Probability Theory, Sanda/Kyoto 1990* 75–128. Longman, Harlow 1993.

[24] H.T. Yau: Relative entropy and hydrodynamics of Ginzburg-Landau models. *Lett. Math. Phys.* **22**: 63-80 (1991)

[25] H.T. Yau: Logarithmic Sobolev inequality for generalized simple exclusion processes. *Probability Theory and Related Fields* **109**: 507-538 (1997)

[26] H.T. Yau: Scaling limit of particle systems, incompressible Navier-Stokes equations and Boltzmann equation. In: *Proceedings of the International Congress of Mathematics, Berlin 1998*, vol 3, pp 193-205, Birkhäuser (1999)

József Fritz
Institute of Mathematics
Technical University Budapest
Egry József u. 1.
H-1111 Budapest, Hungary
jofri@math.bme.hu

Bálint Tóth
Institute of Mathematics
Technical University Budapest
Egry József u. 1.
H-1111 Budapest, Hungary
balint@math.bme.hu