RIGOROUS DERIVATION OF THE LIGHTHILL-WHITHAM-RICHARDS MODEL FROM THE FOLLOW-THE-LEADER MODEL AS MANY PARTICLE LIMIT

M. DI FRANCESCO AND M.D. ROSINI

Abstract. We prove that the unique entropy solution to the macroscopic Lighthill-Witham-Richards model for traffic flow can be rigorously obtained as the large particle limit of the microscopic follow-the-leader model, which is interpreted as the discrete Lagrangian approximation of the former. More precisely, we prove that the empirical measure (respectively the discretised density) obtained from the follow-the-leader system converges in the 1–Wasserstein topology (respectively in $\mathbf{L^1_{loc}}$) to the unique entropy solution of the Lighthill-Witham-Richards equation in the Kruzkov sense. The initial data are taken in $\mathbf{L}^\infty$ with compact support, hence we are able to handle densities with vacuum. Our result holds for a reasonably general class of velocity maps (including all the relevant examples in the applications) with possible degenerate slope near the vacuum state. The proof of the result is based on discrete $\mathbf{BV}$ estimates and on a discrete version of the one-sided Oleinik-type condition. In particular, we prove that the regularizing effect $\mathbf{L}^\infty \mapsto \mathbf{BV}$ is intrinsic of the discrete model.

Keywords: micro-macro limit, Lighthill-Whitham-Richards models, follow-the-leader models, Oleinik condition, entropy solutions, particle method.

2010 AMS Subject classification: 35L65 , 35L45 , 90B20 , 65N75 , 82C22 .

1. Introduction

The modeling of vehicular traffic flow can be considered as one of the most important challenges of applied mathematics in the last seventy years. Among its several repercussions on real-world applications, we mention e.g. the development of smart traffic management systems for integrated applications of communications, control, and information processing technologies to the whole transport system. Other important resultant benefits are the implementation of a joint problem-solving in traffic management, and the addressing of practical problem such as reducing congestion and related costs. These goals can be achieved by optimizing the use of transport resources and infrastructures of the transport system as a whole, by bringing more efficiency in terms of traffic fluidity, and by providing procedures for system stabilization.

Several analytical models for vehicular traffics have been developed in the recent decades. In the first instance, they are classified into two main classes: microscopic models – taking into account each single vehicle – and macroscopic ones – dealing with averaged quantities. We refer to [5, 30, 39, 42] for a survey of the models currently available in the literature.

Recently, the availability of on-line data allows to implementing real-time strategies aiming at avoiding (or mitigating) congested traffic. To address this task, the development and the application of analytical models that are easy-to-use and with a high performance in terms of time and reliability are essential requirements. In this sense, opposed to direct numerical 'individual based' simulations of large number of interacting agents – as typical when dealing with microscopic models – many researchers recommend using macroscopic (e.g. fluid-dynamic) models for traffic flows. The main advantages of the macroscopic approach with respect to the microscopic one are

- the model is completely evolutive and is able to rapidly describe traffic situations at every time;
- the resulting description of queues evolution and of traveling times is accurate as the position of shock waves can be exactly computed and corresponds to queues tails;

- the theory helps developing efficient numerical schemes suitable to describe very large number of agents;
- the model can be easily calibrated, validated and implemented as the number of parameters is low;
- the theory allows to state and possibly solve optimal management problems.

The basic macroscopic approach to traffic flow is given by nonlinear hyperbolic conservation laws [11, 18, 46]. Such choice is a very natural consequence of the assumption that the total number of vehicles is constant as the traffic dynamics evolves. Among the macroscopic models, we can distinguish two main approaches: first order and second order models. The former – sometimes referred to as equilibrium models – are based on the assumption that the velocity can be expressed a priori as an explicit function of the density alone, see for instance [33, 41]. The latter correspond to the non-equilibrium models, in which the velocity and the density are coupled through a further evolution equation, which can be regarded as a continuum analogue of Newton's law, as for instance in [4, 50]. We underline that the only accurate physical law in vehicular traffic theory is the conservation of the total number of vehicles. All other assumptions result from coarse approximations of empirical observations. However, as the dynamics of any living system are influenced by decision-making and psychological effects, nobody would expect that traffic models could reach an accuracy comparable to that attained in other domains of science (e.g. Newtonian physics or thermodynamics). Nevertheless, they potentially have sufficient descriptive power for the specific application-driven purpose, and they help understanding non-trivial properties of traffic flows. In this paper we shall focus on equilibrium macroscopic traffic models by assuming a general constitutive equation for the velocity.

The use of macroscopic models relies on the *continuum assumption*, namely on the assumption that the medium is indefinitely divisible without changing its physical nature. Such assumption is not justifiable in the context of vehicular traffics, as the number of vehicles is typically far lower than the typical number of molecules e.g. in fluid dynamics. Indeed, in order to motivate the use of a continuum model, the number of vehicles should be large enough to give sense to concepts like macroscopic density and average flow. Usually, the continuum hypothesis is accepted as a technical approximation of the physical reality, regarding macroscopic quantities as measures of traffic features. In order to justify and make more clear the continuum hypothesis, the study of the discrete to continuum limit for second order models has been proposed in [3, 6]. First attempts at analyzing the same connection for first order macroscopic models have been recently proposed in [14, 15, 43]. In this paper we address this latter task in a rigorous and constructive form. More precisely we prove that, under reasonable assumptions on the velocity field, the continuum (macroscopic) model can be *solved* as a *many particle limit* of a discrete (microscopic) one.

Our approach can be sketched as follows. We fix $L > 0$ to be the total length of the vehicles on a highway, namely the total space occupied by the all vehicles (i.e. the total mass in a 'continuum PDEs' language), and we consider an initial continuum density $\bar{\rho}$ with total length L. For a given positive integer N, we split $\bar{\rho}$ in N platoons of 'possibly fractional' vehicles, each one of equal length $\ell := L/N$ (more precisely, ℓ is the space occupied by the vehicles belonging to each single platoon), with the endpoints of each platoon positioned at $\bar{x}_i \in \mathbb{R}$, $i = 0, \ldots, N$. The points $\bar{x}_i$ are interpreted as (ordered) particles, and they are taken as initial condition to an ODE system describing the evolution of vehicles in the discrete setting, namely to the *follow-the-leader system*

$$\dot{x}_i(t) = v\left(\frac{\ell}{x_{i+1}(t) - x_i(t)}\right), \qquad i = 0, \ldots, N-1, \tag{1.1a}$$

$$\dot{x}_N(t) = v_{\max}, \tag{1.1b}$$

where $v_{\max} > 0$ is the maximum speed possible. Here $v = v(\rho)$ is the empirical law for the velocity as a function of the density ρ. The points $x_i(t)$ are interpreted as moving particles on the real line. We remark here that no collisions occur between them, as the distance between two consecutive points is expected to be larger than or equal to the quantity ℓ for all times, see Lemma 1.1 below. We shall describe this model in detail below in Subsection 1.2. We then consider the empirical measure

$$\rho^N(t) = \ell \sum_{i=0}^{N-1} \delta_{x_i(t)}, \tag{1.2}$$

and prove in Theorem 1.3 that its limit (in a measure sense to be explained later on) as N goes to infinity is actually an $\mathbf{L}^1$ density ρ, which satisfies the Lighthill-Whitham-Richards model [33, 41]

$$\rho_t + f(\rho)_x = 0, \qquad \text{with} \quad f(\rho) = \rho\, v(\rho), \tag{1.3}$$

in the Oleinik-Kružkov entropy sense [31, 37], see Definition 1.1 below. Our convergence result has a natural interpretation as a *large particle limit* for the scalar conservation law (1.3). In this sense, it can be seen as an abstract particle method for (1.3) which can be applied in the context of numerics. On the other hand, the discrete model (1.1) can be also interpreted as a discrete *Lagrangian* formulation of (1.3), which makes our result meaningful from a physical point of view.

The main novelty with respect to previous results in the literature is that our result is purely *constructive*, in the sense that it can be considered as an alternative tool to *solve* a scalar conservation law. No property of the limiting solution is used, except the uniqueness of entropy solutions in [31, 37] which is used to prove that the scheme has a unique limit. Furthermore, differently from [3], we do not shrink the length of the vehicles to zero and we do not let the size of the highway or the number of vehicles under consideration tend to infinity. In fact, our approximation algorithm rather lets the number of platoons under consideration tend to infinity, but keeps both the length of the highway and the total length L of the vehicles constant. Finally, another important difference from [3] is that our approach allows to handle vacuum regions. This introduces further technical difficulties that are rigourously treated and solved in the present paper.

Although the literature on nonlinear conservation laws is extremely rich of effective numerical schemes (we mention here the pioneering work of Glimm [24] for systems, and the wave-front tracking algorithm proposed by Dafermos in [17] and improved later on by Di Perna [21] and Bressan [10], see [11] and the references therein for more details), to our knowledge the rigorous approximation of an entropy solution to a scalar conservation law by the empirical solution to an ODE system of lagrangian particles in the spirit of (1.1) has not been covered yet. The recent paper [15] provides preliminary results, but it does not contain the needed estimates to justify the limiting procedure.

Our approach differs from most of the numerical approaches to the solution to a scalar conservation law in that it interprets the microscopic limit as a *mean field limit of a system of interacting particles with nearest neighbour type interaction*, in the spirit of (locally and non-locally) interacting particles systems in statistical mechanics, probability, kinetic theory, mathematical biology, etc. In this sense, our result can be cast in the framework of large (deterministic) particle limits with application to several contexts in fluid mechanics, see e.g. the classical references [22, 35, 38]. In one space dimension, a key result in the context of deterministic approximations is the one by Russo [45], which applies to the linear diffusion equation, in which the diffusion operator is replaced by a nearest neighbour interaction term (see also later generalizations to nonlinear diffusion in [34]). We also mention here the paper by Brenier and Grenier [8], which provides a particle justification of the pressureless Euler system (and a particle approximation for a scalar conservation law, although with a completely different approach and interpretation). Our approach can be considered as more in the spirit of [45], applied to a scalar conservation law of traffic type.

The existing numerical method for scalar conservation laws which most resembles our particle method is probably the wave-front tracking algorithm, in which the solution is approximated by a piecewise constant profile which is discontinuous on a finite number of moving fronts. Such a structure naturally suggests the *total variation* as the natural quantity to look at in order to perform efficient uniform estimates, and the space $\mathbf{L}^1$ as the natural environment to set up the problem and to measure the error in the approximation procedure. In our case, the approximating sequence is a *linear combination of Dirac's deltas.* Therefore, a *measure topology* is needed to compare the approximating solution and its limit. Our choice (which will appear as the most natural one) for such a topology is (a scaled version of) the $1-$*Wasserstein distance*, see [1, 49].

The main advantage in using the Wasserstein distance relies on its identification with the $\mathbf{L}^1$–topology in the space of *pseudo-inverses of cumulative distributions*. Roughly speaking, let ρ be the solution to (1.3) and let

$$F(t,x) := \int_{-\infty}^{x} \rho(t,x)\,\mathrm{d}x \in [0,L],$$

be its primitive. The pseudo inverse variable

$$X(t,z) := \inf\{x \in \mathbb{R} :\ F(x) > z\}, \qquad z \in [0,L[,$$

formally satisfies the *Lagrangian PDE*

$$X_t(t,z) = v\left(\frac{1}{X_z(t,z)}\right).$$

Now, if we replace the z–derivative of X by a forward finite difference

$$X_z \approx \frac{X(t,z+\ell)-X(t,z)}{\ell},$$

and assume that X is piecewise constant on intervals of length ℓ, the ODE system (1.1) is immediately recovered, with the structure

$$X(t,z) = \sum_i x_i(t)\,\chi_{[i\ell,(i+1)\ell[}(z).$$

We shall explain the above formal computation more in detail in Section A in the Appendix.

The use of pseudo-inverse variables and Wasserstein distances in the framework of scalar conservation laws is not totally new. In [12], a contraction estimate in the so-called ∞–Wasserstein distance for genuinely nonlinear scalar conservation laws was derived. The case of non-decreasing solution was treated earlier in [7]. As far as the LWR model is concerned, we also remark here that in [36] a simplified version of the model (1.3) is derived by introducing as new variable the cumulative number of vehicles passing through a location x at time t starting from the passage of some reference vehicle, see [2, 19] for recent developments of this theory.

From the technical point of view, our convergence result relies first of all on proving that the empirical measure (1.2) has the same (weak) $N \to +\infty$ limit as the piecewise constant approximation

$$\hat{\rho}^N(t,x) = \sum_{i+1}^{N-1} y_i(t)\chi_{[x_i(t),x_{i+1}(t)[}, \qquad y_i(t) := \frac{\ell}{x_{i+1}(t)-x_i(t)},$$

in which $y_i(t)$ is the discrete lagrangian version of the density. The most important step, however, lies in providing strong $\mathbf{L}^1$ compactness of $\hat{\rho}^N$. This task is performed in two different ways. In the case of $\mathbf{BV}$ initial data, we are able to provide a direct estimate of the total variation of the discrete density (see Proposition 2.5). On the other hand, our main result concerns with the case of general $\mathbf{L}^\infty$ data: in this case, a key estimate on the particle model (see Lemma 2.4), which can be considered as a discrete version of the Oleinik condition for the scalar conservation law, allows to provide strong compactness even if the initial total variation is unbounded. In some sense, this proves that the one-sided Lipschitz regularizing effect of the scalar conservation law (1.3) is somehow an intrinsic property of the discrete Lagrangian formulation of the model. We defer to [25] and the references therein for general results on the regularizing effect for scalar conservation laws.

For numerical purposes, the use of discrete Oleinik conditions has been addressed before for the Lax-Friedrichs and Godunov schemes in [9, 26, 37, 47]. There is also a similar result for second order systems in [6]. The striking novelty in our approach is the fact that our discrete Oleinik condition is only posed in terms of the velocity field, whereas the classical Oleinik condition is stated in terms of the derivative of the flux, see [29]. This is due to the fact that the discrete model is a Lagrangian one, and is therefore characterised by the velocity law. The advantage of having the discrete one-sided Lipschitz condition in terms of the velocity is that we can also consider velocity laws with degenerate slopes at $\rho = 0$. An interesting numerical feature (which is however quite natural when considering particle approximations) is that the discrete approximation $\hat{\rho}^N$ for the density has *no vacuum regions* in the interior of its support, no matter whether or not the (continuum) initial condition is made up by more than one hump. Finally, let us mention that our discrete density $\hat{\rho}^N$ is always discontinuous on at most $N+1$ fronts, unlike in the wave front tracking approximation in which the number of jumps may increase in time.

Our paper is structured as follows. We introduce the (continuum) LWR model (1.3) and the (discrete) FTL model (1.1) in detail, in subsections 1.1 and 1.2 respectively. In Subsection 1.3 we recall the basics on the Wasserstein distance in one space dimension. We set up the approximating scheme and state our main result in Section 2. The precise statement of the main result is contained in Theorem 1.3 and its proof is split into the subsections 2.1, 2.2, 2.3, and 2.4. More precisely, Subsection 2.1 is devoted to the proof of the weak convergence of our approximating scheme, in Subsection 2.2 we prove the two basic compactness estimates mentioned above, in Subsection 2.3 we provide the needed time-continuity and prove strong compactness in $\mathbf{L}^1$, and finally in Subsection 2.4 we prove that the limit is the unique entropy solution in the Oleinik-Kružkov entropy sense [31, 37].

1.1. **The LWR model.** The LWR model is the first and most popular equilibrium model for traffic flows. It was independently introduced by Lighthill, Whitham [33] and Richards [41]. It is based on the assumption that the velocity of the vehicles depends only on their local density, and that the number of vehicles is conserved, namely that the total number of vehicles on a given segment of the road $x \in [a, b]$ only varies due to the incoming flux in $x = a$ and the outgoing flux in $x = b$. With these assumptions the model is expressed by the following Cauchy problem for a scalar conservation law

$$\rho_t + [\rho\, v(\rho)]_x = 0, \qquad t > 0,\ x \in \mathbb{R}, \tag{1.4a}$$

$$\rho(0, x) = \bar{\rho}(x), \qquad x \in \mathbb{R}, \tag{1.4b}$$

where $\rho = \rho(t, x) \in [0, 1]$ is the (dimensionless) normalized density of vehicles in $x \in \mathbb{R}$ at time $t \geq 0$, v is the (mean) velocity and $\bar{\rho}$ is the initial distribution of the vehicles with compact support.

The velocity law $\rho \mapsto v(\rho)$ is defined on $[0, 1]$ and with values in $[0, v_{\max}]$, with $v_{\max} > 0$ being the maximum speed corresponding to a free highway. Typically, lower velocities correspond to higher densities, namely $\rho \mapsto v(\rho)$ is a non-increasing function with $v(0) = v_{\max}$. Moreover, in order not to allow a vehicle to move in case of maximum density, one prescribes $v(1) = 0$. The traffic *flux* function

$$f(\rho) := \rho\, v(\rho)$$

is typically assumed to be concave with $f(0) = f(1) = 0$. Since $f(0) = 0$ and for the finite speed of propagation $\lim_{|x|\to+\infty} \rho(t, x) = 0$, we have that the total space $L := \|\rho(t)\|_{\mathbf{L}^1(\mathbb{R};[0,1])}$ occupied by all vehicles at time t is time independent, namely $L = \|\bar{\rho}\|_{\mathbf{L}^1(\mathbb{R};[0,1])}$ for all $t \geq 0$.

According to the theory of nonlinear conservation laws, see e.g. [11, 18, 46], solutions to (1.4) may develop discontinuities in a finite time, also for regular initial data. For this reason, one has to consider weak solutions ρ to (1.4), more precisely ρ in $\mathbf{L}^\infty\left([0, +\infty[\,; \mathbf{L}^1(\mathbb{R}; [0, 1])\right)$ that satisfy (1.4) in the sense of distributions, namely

$$\int_{\mathbb{R}} \int_0^{+\infty} \Big[\rho(t, x)\, \varphi_t(t, x) + f(\rho(t, x))\, \varphi_x(t, x)\Big]\, \mathrm{d}t\, \mathrm{d}x + \int_{\mathbb{R}} \bar{\rho}(x)\, \varphi(0, x)\, \mathrm{d}x = 0 \tag{1.5}$$

for all $\varphi \in \mathbf{C}^\infty_{\mathbf{c}}([0, +\infty[\, \times \mathbb{R}; \mathbb{R})$. The choice of $\mathbf{L}^1(\mathbb{R}; [0, 1])$ as the functional space to deal with the $x-$regularity appears as the most reasonable one in order to obtain existence of weak solutions when the approximating procedure is performed via a vanishing viscosity argument, see e.g. [18, Section 6.3]. However, the space $\mathbf{BV}(\mathbb{R}; [0, 1])$ is more reminiscent of the typical structure of solutions featuring shocks and rarefaction waves, and turns out to be a natural choice when the problem is e.g. solved by the polygonal approximation algorithm also known as the wave-front tracking algorithm [17], see also [11] and the references therein.

It is well known that the notion of weak solution introduced above is not enough to provide uniqueness of solutions to (1.4). The concept of entropy solution formulated in [31, 32, 37] (see also [18] and the references therein), provides the most natural and efficient way to single out a unique (physically relevant) solution to (1.4). Such concept can be be formulated in several ways, also depending on the regularity of ρ, the most general one being the one proposed by Kružkov [31], which holds for a reasonably wide class of fluxes (namely $\rho \mapsto f(\rho)$ being locally Lipschitz) and in arbitrary space dimension.

Definition 1.1 (Entropy solutions)**.** *Assume that the flux $\rho \mapsto f(\rho)$ is locally Lipschitz. A function ρ in $\mathbf{L}^\infty\left([0,+\infty[\,;\mathbf{L}^1(\mathbb{R};[0,1])\right)$ is an* entropy solution *to* (1.4) *if it satisfies the entropy inequality*

$$\begin{aligned}&\int_{\mathbb{R}}\int_0^{+\infty}\Big[|\rho(t,x)-k|\,\varphi_t(t,x)+\operatorname{sgn}\left(\rho(t,x)-k\right)\left[f\left(\rho(t,x)\right)-f(k)\right]\varphi_x(t,x)\Big]\,\mathrm{d}t\,\mathrm{d}x\\&\quad+\int_{\mathbb{R}}\varphi(0,x)|\bar\rho(x)-k|\,\mathrm{d}x\geq 0\end{aligned}\tag{1.6}$$

for all $\varphi \in \mathbf{C}^\infty_{\mathbf{c}}\left([0,+\infty[\times\mathbb{R};\mathbb{R}\right)$ with $\varphi\geq 0$, and for all constants $k\in\mathbb{R}$.

Clearly, any entropy solution is a weak solution to (1.4) in the sense of (1.5). Moreover uniqueness follows from (1.6).

Theorem 1.1 (Kružkov [31])**.** *Assume that the flux f is locally Lipschitz. Then, for any given initial condition $\bar\rho$ in $\mathbf{L}^\infty$ with $0\leq\bar\rho\leq 1$ a.e. and with compact support, there exists a unique entropy solution to* (1.4) *in the sense of Definition 1.1.*

It is easy to check that any function ρ satisfying the entropy inequality (1.6) satisfies also the following property of (weak) $\mathbf{L}^1$–continuity in time

$$\lim_{T\to 0+}\frac{1}{T}\int_0^T\int_{|x|\leq r}|\rho(t,x)-\bar\rho(x)|\,\mathrm{d}x\,\mathrm{d}t=0$$

for all $r>0$. However, depending on the way we attempt at constructing entropy solutions, an important issue is related with detecting the trace at $t=0$ in a strong enough topology. This is often the case when the approximating scheme lacks of compactness when t approaches zero. A theorem due to Chen and Rascle [13] states that the uniqueness of the entropy solution is preserved also for a notion of entropy solution relaxed at $t=0$, provided the flux f satisfies a.e. a genuine nonlinearity condition.

Theorem 1.2 (Chen and Rascle [13])**.** *Assume there exists no nontrivial interval on which f is affine. If $\bar\rho$ is in $\mathbf{L}^\infty$ with $0\leq\bar\rho\leq 1$ a.e. and with compact support, then there exists a unique ρ in $\mathbf{L}^\infty\left([0,+\infty[\,;\mathbf{L}^1(\mathbb{R};[0,1])\right)$ weak solution to* (1.4) *in the sense of* (1.5) *that satisfies also*

$$\int_{\mathbb{R}}\int_0^{+\infty}\Big[|\rho(t,x)-k|\,\varphi_t(t,x)+\operatorname{sgn}(\rho(t,x)-k)[f\left(\rho(t,x)\right)-f(k)]\,\varphi_x(t,x)\Big]\,\mathrm{d}t\,\mathrm{d}x\geq 0\tag{1.7}$$

for all $\varphi\in\mathbf{C}^\infty_{\mathbf{c}}\left(]0,+\infty[\times\mathbb{R};\mathbb{R}\right)$ with $\varphi\geq 0$, and for all constants $k\in\mathbb{R}$. Moreover, ρ is the unique entropy solution in the sense of Definition 1.1.

Let us finally recall that, for $\mathbf{C^1}$–fluxes f which are concave or convex, another classical tool to uniquely determine all weak solutions by their $\mathbf{L}^\infty$–initial values is the so called Oleinik-type condition [29]

$$\int_{\mathbb{R}}\int_0^{+\infty}f'(\rho(t,x))\,\varphi_x(t,x)\,\mathrm{d}t\,\mathrm{d}x\geq-\int_{\mathbb{R}}\int_0^{+\infty}\frac{1}{t}\,\varphi(t,x)\,\mathrm{d}t\,\mathrm{d}x\tag{1.8}$$

for all $\varphi\in\mathbf{C}^\infty_{\mathbf{c}}\left([0,+\infty[\times\mathbb{R};\mathbb{R}\right)$ with $\varphi\geq 0$, and for all $t>0$. Moreover, if f' has Lipschitz continuous inverse, then (1.8) implies that $\rho(t,\cdot)$ has locally bounded total variation for all $t>0$ even if the initial datum is not in $\mathbf{BV}$.

1.2. **The FTL model.** Microscopic models of vehicular traffic are typically based on the so called Follow-The-Leader (FTL) model, that is the subject of this section.

Consider a single lane road parameterized by $x \in \mathbb{R}$, with traffic moving in the direction of increasing x, with $N+1$ ordered Reference Vehicles (RVs). Denote by $t \mapsto x_i(t)$ the position of the i–th RV for $i = 0, \ldots, N$. Then, according to the FTL model, the evolution of the traffic along the road is described inductively by the following Cauchy problem for an ODE system

$$\dot{x}_N(t) = v_{\max}, \tag{1.9a}$$

$$\dot{x}_i(t) = v\left(\frac{\ell}{x_{i+1}(t) - x_i(t)}\right), \qquad i = 0, \ldots, N-1, \tag{1.9b}$$

$$x_i(0) = \bar{x}_i, \qquad i = 0, \ldots, N, \tag{1.9c}$$

where $v \in \mathbf{C^1}([0,1];[0,v_{\max}])$ is the velocity map, $\bar{x}_0 < \ldots < \bar{x}_N$ are the initial positions of the RVs, $\ell > 0$ is the length of each RV, and $v_{\max}$ is the maximum velocity, reached by vehicles with free road ahead, i.e. only by the leading vehicle x_N. Coherently with the definition of ℓ, we assume that

$$\bar{x}_{i+1} - \bar{x}_i \geq \ell, \qquad i = 0, \ldots, N-1. \tag{1.10}$$

System (1.9) can be solved inductively starting from $i = N$. Indeed, from (1.9a), we immediately deduce that

$$x_N(t) = \bar{x}_N + v_{\max}\, t.$$

Then, we can compute $t \mapsto x_i(t)$ once we know $t \mapsto x_{i+1}(t)$. In fact, according with the system (1.9) the velocity of the i–th RV depends on its distance from the $(i+1)$–th RV alone via the smooth velocity map v, that is assumed to be non-increasing and with $v(1) = 0$. The latter assumption can be interpreted as $\dot{x}_i(t) = 0$ when $x_{i+1}(t) - x_i(t) = \ell$, namely, if at time t the vehicles $x_i(t)$ and $x_{i+1}(t)$ are bumper-to-bumper, then the i–th RV is not moving. As we will see in the next lemma, this ensures that $x_{i+1}(t) - x_i(t) \geq \ell$, $i = 0, \ldots, N-1$, for all times $t \geq 0$ and, therefore, that (1.9) admits a global-in-time solution.

Lemma 1.1 (Discrete maximum principle, [43])**.** *For all $i = 0, \ldots, N-1$, we have*

$$\ell \leq x_{i+1}(t) - x_i(t) \leq \bar{x}_N - \bar{x}_0 + v_{\max}\, t \qquad \text{for all times } t \geq 0. \tag{1.11}$$

Proof. The upper bound is obvious. Hence, it is sufficient to prove the lower bound. Consider the Cauchy problem obtained from (1.9) by substituting v with its extension $V : [0,+\infty[\to [0, v_{\max}]$ defined by $V := v\,\chi_{[0,1]}$ (note that the extension on $]-\infty, 0[$ is not of interest as the argument of v is always detached from zero). Denote by $(X_0, \ldots, X_N)$ the corresponding solution. By (1.10) we have that $X_{i+1}(0) - X_i(0) \geq \ell$. Assume by contradiction that there exists $i \in \{0, \ldots, N-1\}$ and $t_2 > t_1 \geq 0$ such that $X_{i+1}(t_1) - X_i(t_1) = \ell$ and $X_{i+1}(t) - X_i(t) < \ell$ for all $t \in]t_1, t_2]$. Since $t \mapsto X_{i+1}(t) - X_i(t)$ is $\mathbf{C^1}$, we have that $\dot{X}_{i+1}(t_1) - \dot{X}_i(t_1) < 0$. On the other hand, by (1.9) we have that for all $t \in [t_1, t_2]$

$$\dot{X}_{i+1}(t) - \dot{X}_i(t) = V\left(\frac{\ell}{X_{i+2}(t) - X_{i+1}(t)}\right) \geq 0 \qquad \text{if } i = 0, \ldots, N-2,$$

$$\dot{X}_N(t) - \dot{X}_{N-1}(t) = v_{\max} > 0 \qquad \text{if } i = N-1,$$

but this is a contradiction. Hence $X_{i+1}(t) - X_i(t) \geq \ell$ for all $t \geq 0$. As a consequence of the uniqueness of the solution to (1.9), we have that $(X_1, \ldots, X_N)$ is in fact also the solution of (1.9) with the original v. ◇

1.3. **Notation and preliminaries on measure distances.** In this section we recall basic properties of pseudo-inverse operators that we shall use extensively in the rest of the paper. We defer to [49] for further details.

For a fixed $L > 0$, introduce the pseudo-inverse operators

$$\mathcal{X}\colon \mathbf{L}^\infty(\mathbb{R};[0,L]) \to \mathbf{L}^\infty([0,L[\,;\mathbb{R}),$$
$$\mathcal{F}\colon \mathbf{L}^\infty([0,L[\,;\mathbb{R}) \to \mathbf{L}^\infty(\mathbb{R};[0,L]),$$

defined by

$$\mathcal{X}[F](z) := \inf\{x \in \mathbb{R} :\ F(x) > z\} \qquad \text{for } z \in [0,L[,$$
$$\mathcal{F}[X](x) := \operatorname{meas}\{z \in [0,L] :\ X(z) \le x\} \qquad \text{for } x \in \mathbb{R},$$

and consider the space

$$\mathcal{M}_L := \{\rho \text{ Radon measure on } \mathbb{R} \text{ with compact support} :\ \rho \ge 0,\ \rho(\mathbb{R}) = L\}.$$

For a given $\rho \in \mathcal{M}_L$, we denote $x^\rho_{\min} := \min(\operatorname{spt}(\rho))$ and $x^\rho_{\max} := \max(\operatorname{spt}(\rho))$, and by $F_\rho : \mathbb{R} \to [0,L]$ its cumulative distribution, namely $F_\rho(x) := \rho(]-\infty,x])$. We observe that $F_\rho \in \mathbf{L}^\infty(\mathbb{R};[0,L])$ is non-decreasing, right-continuous with $F_\rho(x) = 0$ for all $x \le x^\rho_{\min}$ and $F_\rho(x) = L$ for all $x \ge x^\rho_{\max}$. Therefore we can define its pseudo-inverse $X_\rho := \mathcal{X}[F_\rho]$. Clearly, $X_\rho \in \mathbf{L}^\infty([0,L[\,;[x^\rho_{\min}, x^\rho_{\max}])$ is non-decreasing, right-continuous with $X_\rho(0) = x^\rho_{\min}$. By abuse of notation, we shall adopt the notation ρ to denote an absolutely continuous measure in $\mathcal{M}_L$ with $\mathbf{L}^1$–density ρ.

Lemma 1.2 (Change of variable)**.** *If $\rho \in \mathcal{M}_L$, then for all $\varphi \in \mathbf{C^0}(\mathbb{R};\mathbb{R})$ we have*

$$\int_{\mathbb{R}} \varphi(x)\,\mathrm{d}\rho(x) = \int_0^L \varphi(X_\rho(z))\,\mathrm{d}z.$$

We recall that, for $L = 1$, the one-dimensional 1–*Wasserstein distance* between $\rho_1,\rho_2 \in \mathcal{M}_1$ (defined in terms of optimal plans in the Monge-Kantorovich problem, see e.g. [49]) can be defined as

$$d_1(\rho_1,\rho_2) := \|F_{\rho_1} - F_{\rho_2}\|_{\mathbf{L}^1(\mathbb{R};\mathbb{R})} = \|X_{\rho_1} - X_{\rho_2}\|_{\mathbf{L}^1([0,1];\mathbb{R})}.$$

For a general strictly positive L, we introduce the *scaled* 1–*Wasserstein distance* between $\rho_1,\rho_2 \in \mathcal{M}_L$ as

$$d_{L,1}(\rho_1,\rho_2) := \|F_{\rho_1} - F_{\rho_2}\|_{\mathbf{L}^1(\mathbb{R};\mathbb{R})} = \|X_{\rho_1} - X_{\rho_2}\|_{\mathbf{L}^1([0,L];\mathbb{R})}. \tag{1.12}$$

Indeed, straightforward computation yields

$$d_{L,1}(\rho_1,\rho_2) = L\, d_1(\rho_1/L, \rho_2/L).$$

The distance $d_{L,1}$ inherits all the topological properties of the 1–Wasserstein distance for probability measures. In particular, a sequence $(\rho_n)_{n\in\mathbb{N}}$ in $\mathcal{M}_L$ converges to $\rho \in \mathcal{M}_L$ in $d_{L,1}$ if and only if

$$\lim_{n\to+\infty} \int_{\mathbb{R}} \varphi(x)\,\mathrm{d}\rho_n(x) = \int_{\mathbb{R}} \varphi(x)\,\mathrm{d}\rho(x),$$

for all $\varphi \in \mathbf{C^0}(\mathbb{R};\mathbb{R})$ growing at most linearly at infinity.

1.4. **Statement of the main result.** In this subsection we state our main result, which provides a rigorous description of the unique entropy solution ρ to the Cauchy problem (1.4) as the limit for N that goes to infinity of a density associated to the microscopic model (1.9) to be constructed as described below.

We shall work under the standing assumption on the initial datum

(In) The initial datum $\bar\rho$ is in $\mathcal{M}_L \cap \mathbf{L}^\infty(\mathbb{R};\mathbb{R})$ and $0 \le \bar\rho \le 1$ almost everywhere.

In some cases we shall require the stronger condition

(InBV) The initial datum $\bar\rho$ is in $\mathcal{M}_L \cap \mathbf{BV}(\mathbb{R};[0,1])$.

As for the velocity function v, we shall require throughout the paper

(V1) $v \in \mathbf{C^1}([0,1];[0,v_{\max}])$, v strictly decreasing on $[0,1]$, $v_{\max} > 0$.

(V2) $v(0) = v_{\max}$ and $v(1) = 0$.

The assumption (V1) is a minimal requirement for having a unique local solution to the system (1.9). Assumption (V2) is a sufficient condition to guarantee that such solution is globally defined, as seen in Lemma 1.1. From the modelling point of view, condition (V2) is a natural requirement, as it prescribes speed zero at maximal density. The monotonicity condition in (V1) is also a natural requirement for a traffic model (all vehicles drive faster in lower densities).

In some cases, we shall use the extra assumption

(V3) $\rho\, v''(\rho) + v'(\rho) \le 0$ for all $\rho \in [0,1]$.

Notice that the assumption (V3) implies in particular that the flux $\rho \mapsto f(\rho)$ is strictly concave. On the other hand, (V3) is a slightly stricter requirement than strict concavity, but is verified in many examples of velocities arising in traffic flow models.

Remark 1.1 (Examples of velocities)**.** Clearly, the prototype for the velocity $v(\rho) = v_{\max}\,(1-\rho)$ by Greenshields [28] satisfies the assumptions (V1), (V2), (V3). The same holds for the Pipes-Munjal velocity [40]

$$v(\rho) = v_{\max}\,(1-\rho^{\alpha}) \qquad \alpha > 0,$$

in which the concavity of the flux $\rho\, v(\rho)$ degenerates at $\rho = 0$. Further examples of speed-density relations that satisfy (V1), (V2), (V3) can be obtained by a slight modification of the Greenberg model [27]

$$v(\rho) = v_{\max} \left[\log\left(\frac{1+\alpha}{\alpha}\right)\right]^{-1} \log\left(\frac{1+\alpha}{\rho+\alpha}\right), \qquad \alpha > 0,$$

or of the Underwood model [48]

$$v(\rho) = v_{\max}\,\frac{e^{-\rho} - e^{-1}}{1-e^{-1}}.$$

We shall denote by $\bar{x}_{\min} < \bar{x}_{\max}$ the extremal points of the convex hull of the support of $\bar{\rho}$, namely $\bigcap_{[a,b]\supseteq \operatorname{spt}(\bar{\rho})} [a,b] = [\bar{x}_{\min}, \bar{x}_{\max}]$.

We now introduce our atomization scheme. Let n be a positive integer sufficiently large. We split the total length of vehicles $L := \|\bar{\rho}\|_{\mathbf{L}^1(\mathbb{R};\mathbb{R})}$ in $N_n := 2^n$ platoons of length $\ell_n := 2^{-n}L$ as follows. We set

$$\bar{x}_0^n := \bar{x}_{\min}, \tag{1.13a}$$

and recursively

$$\bar{x}_i^n := \sup\left\{x \in \mathbb{R} : \int_{\bar{x}_{i-1}^n}^{x} \bar{\rho}(y)\,\mathrm{d}y < \ell_n\right\}, \qquad i = 1, \ldots, N_n. \tag{1.13b}$$

It is easily seen that $\bar{x}_{N_n}^n = \bar{x}_{\max}$, $\bar{x}_{N_n-i}^n = \bar{x}_{N_{n+m}-2^m i}^{n+m}$, and since $0 \le \bar{\rho} \le 1$ a.e. we have

$$\ell_n = \int_{\bar{x}_i^n}^{\bar{x}_{i+1}^n} \bar{\rho}(y)\,\mathrm{d}y \le \bar{x}_{i+1}^n - \bar{x}_i^n, \qquad i = 0, \ldots, N_n - 1.$$

Thus the condition (1.10) is satisfied with $\ell = \ell_n$, and we can take the values $\bar{x}_0^n, \ldots, \bar{x}_{N_n}^n$ as the initial positions of the (N_n+1) reference vehicles in the n–depending version of the follow-the-leader (1.9)

$$\dot{x}_{N_n}^n(t) = v_{\max}, \tag{1.14a}$$

$$\dot{x}_i^n(t) = v\left(\frac{\ell_n}{x_{i+1}^n(t) - x_i^n(t)}\right), \qquad i = 0, \ldots, N_n - 1, \tag{1.14b}$$

$$x_i^n(0) = \bar{x}_i^n, \qquad i = 0, \ldots, N_n. \tag{1.14c}$$

The existence of a global-in-time solution to (1.14) follows from Lemma 1.1. Moreover, from (1.14a) we immediately deduce that

$$x_{N_n}^n(t) = \bar{x}_{\max} + v_{\max}\, t.$$

By introducing in (1.14) the new variable

$$y_i^n(t) := \frac{\ell_n}{x_{i+1}^n(t) - x_i^n(t)}, \qquad i = 0, \ldots, N_n - 1, \tag{1.15}$$

we obtain

$$\dot{y}_{N-1}^n = -\frac{(y_{N-1}^n)^2}{\ell_n} \left[v_{\max} - v(y_{N-1}^n)\right], \tag{1.16a}$$

$$\dot{y}_i^n = -\frac{(y_i^n)^2}{\ell_n} \left[v(y_{i+1}^n) - v(y_i^n)\right], \qquad i = 0, \ldots, N_n - 2, \tag{1.16b}$$

$$y_i^n(0) = \bar{y}_i^n := \frac{\ell_n}{\bar{x}_{i+1}^n - \bar{x}_i^n}, \qquad i = 0, \ldots, N_n - 1. \tag{1.16c}$$

Observe that $\ell_n / \left[\bar{x}_{\max} - \bar{x}_{\min} + v_{\max}\, t\right] \le y_i^n(t) \le 1$ for all $t \ge 0$ in view of Lemma 1.1. The quantity y_i^n can be seen as a discrete version of the density of cars in Lagrangian coordinates, and the ODEs (1.16a)–(1.16b) are a discrete Lagrangian version of the scalar conservation law (1.4a).

We are now ready to state the main result of this paper.

Theorem 1.3. *Let $\bar{\rho}$ satisfy the condition (In) and v the condition (V1) and (V2). Assume further that either*

- *$\bar{\rho}$ satisfies (InBV),*

or

- *v satisfies (V3).*

Define the piecewise constant (with respect to x) density

$$\hat{\rho}^n(t,x) := \sum_{i=0}^{N_n-1} y_i^n(t)\, \chi_{\left[x_i^n(t),\, x_{i+1}^n(t)\right[}(x), \tag{1.17}$$

and the empirical measure

$$\tilde{\rho}^n(t,x) := \ell_n \sum_{i=0}^{N_n-1} \delta_{x_i^n(t)}(x). \tag{1.18}$$

Then the sequence $(\hat{\rho}^n)_{n\in\mathbb{N}}$ converges to the unique entropy solution ρ of the Cauchy problem (1.4) almost everywhere and in $\mathbf{L^1_{loc}}([0,+\infty[\times\mathbb{R};[0,1])$. Moreover, the sequence $(\tilde{\rho}^n)_{n\in\mathbb{N}}$ converges to ρ in the topology of $\mathbf{L^1_{loc}}([0,+\infty[;\, d_{L,1})$.

2. Proof of the main result

Our strategy for the proof of Theorem 1.3 can be resumed as follows:

(i) Following the notation introduced in Subsection 1.3, we set $\hat{F}^n = F_{\hat{\rho}^n}$ and $\hat{X}^n = \mathcal{X}[\hat{F}^n]$, respectively $\tilde{F}^n = F_{\tilde{\rho}^n}$ and $\tilde{X}^n = \mathcal{X}[\tilde{F}^n]$, as the cumulative distribution of $\hat{\rho}^n$, respectively $\tilde{\rho}^n$, and its pseudo inverses. Introduce the *discrete Lagrangian density*

$$\check{\rho}^n = \hat{\rho}^n \circ \hat{X}^n.$$

(ii) We first prove that the sequence of piecewise constant pseudo-inverse distributions $(\tilde{X}^n)_{n\in\mathbb{N}}$ has a strong limit X in $\mathbf{L^1_{loc}}([0,+\infty[\times[0,L];\mathbb{R})$, which is equivalent to having $(\tilde{\rho}^n)_{n\in\mathbb{N}}$ converging to a measure ρ in the $\mathbf{L^1_{loc}}([0,+\infty[;\, d_{L,1})$ topology. At the same time, we shall also prove that $(\hat{X}^n)_{n\in\mathbb{N}}$ converges in $\mathbf{L^1_{loc}}([0,+\infty[\times[0,L];\mathbb{R})$ to the same limit X, i.e. $(\hat{\rho}^n)_{n\in\mathbb{N}}$ converges to ρ in $\mathbf{L^1_{loc}}([0,+\infty[;\, d_{L,1})$.

(iii) We then prove that the limit pseudo-inverse function X has difference quotients bounded below by 1. This fact allows to prove that the limit measure ρ in (ii) is actually in $\mathbf{L}^\infty$ and is a.e. bounded by 1. At the same time, we easily infer weak–$*$ convergence of $(\check{\rho}^n)_{n\in\mathbb{N}}$ to a limit $\check{\rho}$ in $\mathbf{L}^\infty$. It remains to prove that $\check{\rho}\circ F = \tilde{\rho}$, and that such limit is the unique entropy solution to (1.4). This requires stronger estimates on $\hat{\rho}^n$.

(iv) A direct proof of a uniform $\mathbf{BV}$ estimate for $\hat{\rho}^n$ can be performed in the case of $\mathcal{M}_L\cap\mathbf{BV}$ initial datum. In the case of general $\mathcal{M}_L\cap\mathbf{L}^\infty$ initial datum we shall prove that the discrete Lagrangian density $\check{\rho}^n$ satisfies a (uniform) *discrete version of the Oleinik condition*, which implies automatically a $\mathbf{BV}$ uniform estimate for $\check{\rho}^n$, and hence for $\hat{\rho}^n$. This step crucially requires condition (V3) on v.

(v) The definition of weak solution (1.5) for ρ follows from the $n\to+\infty$ limit of the formulation of (1.14) as a PDE
$$\tilde{X}^n_t = v(\check{\rho}^n). \tag{2.1}$$

(vi) We finally recover the entropy condition (1.6) in the discrete setting, and use the strong $\mathbf{L}^1$ compactness to pass it to the limit.

2.1. **Weak convergence of the approximating scheme.** Throughout this subsection we shall assume that v satisfies (V1) and (V2). Let $\hat{\rho}^n$ and $\tilde{\rho}^n$ be defined as in (1.17) and (1.18) respectively. We have that $\hat{\rho}^n(t),\tilde{\rho}^n(t)\in\mathcal{M}_L$ for all $t\geq 0$. Thus we can consider the cumulative distributions associated to $\hat{\rho}^n$ and $\tilde{\rho}^n$ (recall that $\tilde{\rho}^n$ is an empirical measure)
$$\hat{F}^n(t,x) := \int_{-\infty}^{x}\hat{\rho}^n(t,y)\,\mathrm{d}y, \qquad\qquad \tilde{F}^n(t,x) := \tilde{\rho}^n(]-\infty,x]),$$
and their pseudo-inverses
$$\hat{X}^n := \mathcal{X}\left[\hat{F}^n\right], \qquad\qquad \tilde{X}^n := \mathcal{X}\left[\tilde{F}^n\right],$$
extended to $z=L$ by taking $\hat{X}^n(t,L)=x^n_{N_n}(t)=\tilde{X}^n(t,L)$. By definition, see figures 1 and 2, for all $t\geq 0$,

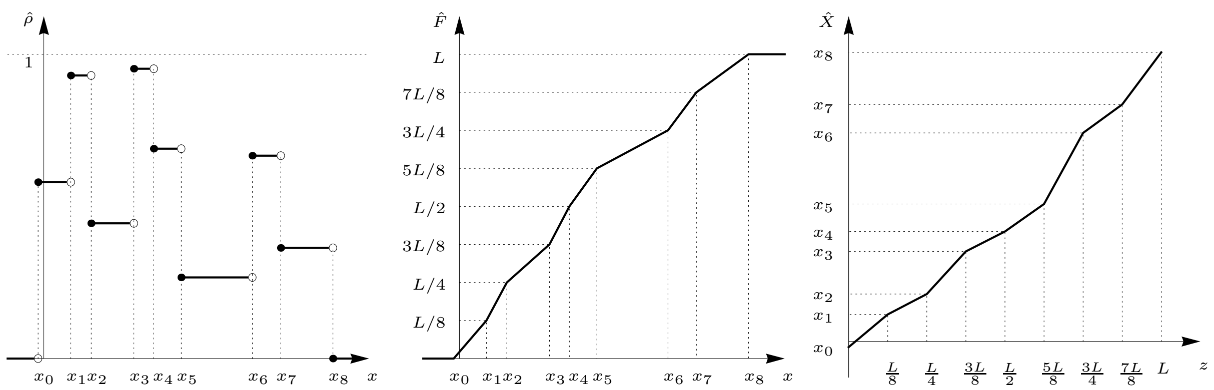

FIGURE 1. Maps of the form, respectively from the left, (1.17), (2.2) and (2.3) with $n(=3)$ and $t(\geq 0)$ omitted.

$z\in[0,L]$ and $x\in\mathbb{R}$ we have
$$\hat{F}^n(t,x) = \sum_{i=0}^{N_n-1}\left[i\,\ell_n + y^n_i(t)\,[x-x^n_i(t)]\right]\chi_{[x^n_i(t),\,x^n_{i+1}(t)[}(x) + L\,\chi_{[x^n_{N_n}(t),\,+\infty[}(x), \tag{2.2}$$
$$\hat{X}^n(t,z) = \sum_{i=0}^{N_n-2}\left[x^n_i(t)+\frac{z-i\,\ell_n}{y^n_i(t)}\right]\chi_{[i\,\ell_n,\,(i+1)\,\ell_n[}(z) + \left[x^n_{N_n-1}(t)+\frac{z-L+\ell_n}{y^n_{N_n-1}(t)}\right]\chi_{[L-\ell_n,\,L]}(z), \tag{2.3}$$

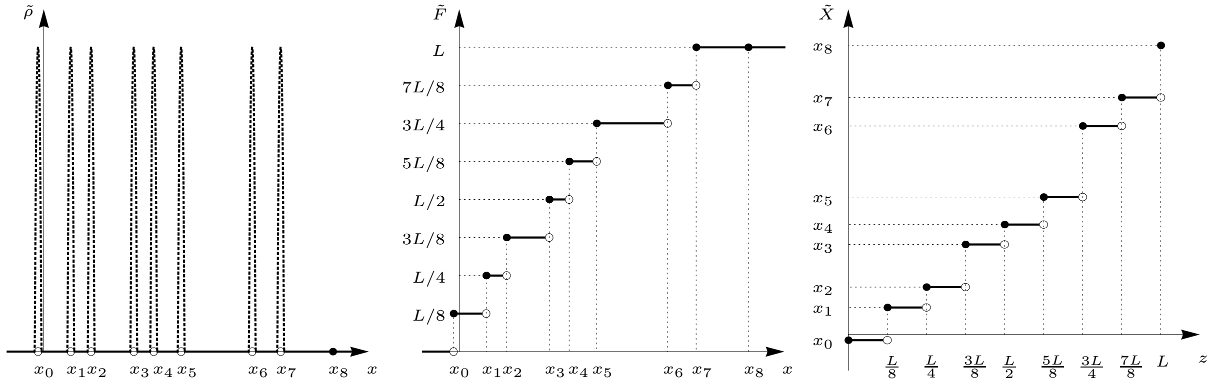

FIGURE 2. Maps of the form, respectively from the left, (1.18), (2.4) and (2.5) with $n(=3)$ and $t(\geq 0)$ omitted.

$$\tilde F^n(t,x) = \sum_{i=0}^{N_n-2} \ell_n \,(i+1)\,\chi_{\left[x_i^n(t),\, x_{i+1}^n(t)\right[}(x) + L\,\chi_{\left[x_{N_n-1}^n(t),\, +\infty\right[}(x), \tag{2.4}$$

$$\tilde X^n(t,z) = \sum_{i=0}^{N_n-1} x_i^n(t)\,\chi_{\left[i\,\ell_n,\, (i+1)\,\ell_n\right[}(z) + x_{N_n}^n(t)\,\chi_{\{L\}}(z). \tag{2.5}$$

Observe that for any fixed $t \geq 0$

- $z \mapsto \hat X^n(t,z)$ and $x \mapsto \hat F^n(t,x)$ are piecewise linear continuous and non-decreasing;
- $\hat F^n(t) : [x_0^n(t), x_{N_n}^n(t)] \to [0,L]$ and $\hat X^n(t) : [0,L] \to [x_0^n(t), x_{N_n}^n(t)]$ are strictly increasing and are inverse functions of each other in the classical sense;
- $z \mapsto \tilde X^n(t,z)$ and $x \mapsto \tilde F^n(t,x)$ are piecewise constant with N_n jumps of discontinuity, right continuous and non-decreasing;
- $\hat F^n(t,x) \leq \tilde F^n(t,x)$ for any $x \in \mathbb{R}$ and $\tilde X^n(t,z) \leq \hat X^n(t,z)$ for any $z \in [0,L]$;
- $\tilde F^{n+1}(t,x) \leq \tilde F^n(t,x)$ for any $x \in \mathbb{R}$ and $\tilde X^n(t,z) \leq \tilde X^{n+1}(t,z)$ for any $z \in [0,L]$;
- $\hat\rho^n(t,x) = \hat F_x^n(t,x)$ for all $x \neq x_i^n(t)$, $i = 1, \ldots, N_n$, while $\tilde\rho^n = \tilde F_x^n$ in the sense of distributions.

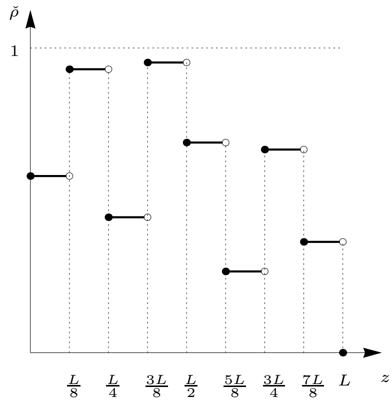

FIGURE 3. Map of the form (2.6) with $n(=3)$ and $t(\geq 0)$ omitted.

For later use, see Figure 3, we introduce also the *discrete Lagrangian density*

$$\check{\rho}^n(t,z) := \hat{\rho}^n\left(t, \hat{X}^n(t,z)\right) = \sum_{i=0}^{N_n-1} y_i^n(t)\,\chi_{[i\,\ell_n,\,(i+1)\,\ell_n[}(z) \tag{2.6}$$

and observe that

$$\tilde{X}_t^n(t,z) = v\left(\check{\rho}^n(t,z)\right), \qquad\qquad t>0,\ z\in[0,L]. \tag{2.7}$$

As a first step, we want to prove that $(\tilde{X}^n)_{n\in\mathbb{N}}$ and $(\hat{X}^n)_{n\in\mathbb{N}}$ have the same unique limit in $\mathbf{L^1_{loc}}([0,+\infty[\times[0,L];\mathbb{R})$.

Proposition 2.1 (Definition of X)**.** *There exists a unique $X \in \mathbf{L}^\infty\left([0,+\infty[\,\times\,[0,L]\,;\mathbb{R}\right)$, monotone non-decreasing and right continuous with respect to z, such that*

$$(\hat{X}^n)_{n\in\mathbb{N}} \textit{ and } (\tilde{X}^n)_{n\in\mathbb{N}} \textit{ converge to } X \textit{ in } \mathbf{L^1_{loc}}\left([0,+\infty[\,\times\,[0,L]\,;\mathbb{R}\right),$$

and for any $t,s>0$

$$\mathrm{TV}\left[X(t)\right] \le |\bar{x}_{\max} - \bar{x}_{\min} + v_{\max}\,t|, \tag{2.8a}$$

$$\|X(t)\|_{\mathbf{L}^\infty([0,L];\mathbb{R})} \le \max\left\{|\bar{x}_{\min}|, |\bar{x}_{\max} + v_{\max}\,t|\right\}, \tag{2.8b}$$

$$\int_0^L |X(t,z) - X(s,z)|\,\mathrm{d}z \le v_{\max}\,L\,|t-s|. \tag{2.8c}$$

Moreover, $(\tilde{X}^n)_{n\in\mathbb{N}}$ converges to X a.e. on $[0,+\infty[\,\times\,[0,L]$.

Proof. Fix $T>0$, and let $n>0$.
• **Step 1: $\tilde{\mathbf{X}}^{\mathbf{n}} \to \mathbf{X}$.** Since $z \mapsto \tilde{X}^n(t,z)$ is non-decreasing with $\tilde{X}^n(t,0) = x_0^n(t) \ge \bar{x}_0^n = \bar{x}_{\min}$ and $\tilde{X}^n(t,L) = \bar{x}_{\max} + v_{\max}\,t$, we have that

$$\mathrm{TV}\left[\tilde{X}^n(t)\right] \le |\bar{x}_{\max} - \bar{x}_{\min} + v_{\max}\,t|,$$
$$\left\|\tilde{X}^n(t)\right\|_{\mathbf{L}^\infty([0,L];\mathbb{R})} \le \max\{|\bar{x}_{\min}|, |\bar{x}_{\max} + v_{\max}\,t|\}.$$

Moreover, if $s<t$, then by (1.14b) and (2.5)

$$\int_0^L \left|\tilde{X}^n(t,z) - \tilde{X}^n(s,z)\right|\mathrm{d}z = \sum_{i=0}^{N_n-1} \ell_n\left[x_i^n(t) - x_i^n(s)\right] = \sum_{i=0}^{N_n-1} \ell_n \left[\int_s^t v\left(y_i^n(\tau)\right)\mathrm{d}\tau\right] \le v_{\max}\,L\,(t-s)\,.$$

Therefore, by applying Helly's theorem in the form [11, Theorem 2.4], up to a subsequence, $(\tilde{X}^n)_{n\in\mathbb{N}}$ converges in $\mathbf{L^1_{loc}}\left([0,+\infty[\,\times\,[0,L]\,;\mathbb{R}\right)$ to a function X right continuous w.r.t. z and satisfying (2.8). Finally, since $\tilde{X}^{n+1}(t,z) \le \tilde{X}^n(t,z)$ for all $t\ge0$ and $z\in[0,L]$, the whole sequence $(\tilde{X}^n)_{n\in\mathbb{N}}$ converges to X and a.e. on $[0,+\infty[\,\times\,[0,L]$.
• **Step 2: $\hat{\mathbf{X}}^{\mathbf{n}} \to \mathbf{X}$.** By definition, see (1.15), (2.3) and (2.5), we have for all $t\in[0,T]$

$$\int_0^L \left|\hat{X}^n(t,z) - \tilde{X}^n(t,z)\right|\mathrm{d}z = \sum_{i=0}^{N_n-1} y_i^n(t)^{-1} \int_{i\,\ell_n}^{(i+1)\,\ell_n} \left[z - i\,\ell_n\right]\mathrm{d}z$$
$$= \frac{\ell_n}{2}\sum_{i=0}^{N_n-1}\left[x_{i+1}^n(t) - x_i^n(t)\right] = \frac{\ell_n}{2}\left[x_{N_n}^n(t) - x_0^n(t)\right] \le \frac{\ell_n}{2}\left[\bar{x}_{\max} - \bar{x}_{\min} + v_{\max}\,T\right],$$

and the proof is complete as $(\tilde{X}^n)_{n\in\mathbb{N}}$ converges to X in view of Step 1. ◇

In the next lemma we prove that X inherits the maximum principle property satisfied by $\tilde{X}^n$ proven in Lemma 1.1.

Lemma 2.1. *For all $t \geq 0$ and for a.e. $z_1, z_2 \in [0, L]$ with $z_1 < z_2$ we have*
$$z_2 - z_1 \leq X(t, z_2) - X(t, z_1) \leq \bar{x}_{\max} - \bar{x}_{\min} + v_{\max}\, t. \tag{2.9}$$

Proof. The upper bound is obvious. Take $0 \leq z_1 < z_2 \leq L$. For $n > 0$ sufficiently large, we can take $i, j \in \{0, 1, \ldots, N_n\}$ such that $i < j$, $i\,\ell_n \leq z_1 < (i+1)\,\ell_n$ and $\ell_n\, j \leq z_2 < \ell_n\,(j+1)$. By (2.5) and Lemma 1.1 we have
$$\begin{aligned}\frac{\tilde{X}^n(t, z_2) - \tilde{X}^n(t, z_1)}{z_2 - z_1} &\geq \frac{x_j^n(t) - x_i^n(t)}{(j+1)\,\ell_n - i\,\ell_n} \geq \frac{(j-i)\,\ell_n}{(j+1)\,\ell_n - i\,\ell_n}\\ &= 1 - \frac{1}{j - i + 1} \geq 1 - \frac{1}{(z_2\,\ell_n^{-1} - 1) - z_1\,\ell_n^{-1} + 1} = 1 - \frac{\ell_n}{z_2 - z_1}.\end{aligned}$$
By letting n go to infinity we conclude the proof. Indeed, $\lim_{n\to+\infty} \ell_n/[z_2 - z_1] = 0$ and $(\tilde{X}^n)_{n\in\mathbb{N}}$ converges to X a.e. on $[0, +\infty[\, \times [0, L]$ in view of Proposition 2.1. ◇

Proposition 2.2 (Definition of F)**.** *$(\hat{F}^n)_{n\in\mathbb{N}}$ and $(\tilde{F}^n)_{n\in\mathbb{N}}$ converge to $F := \mathcal{F}[X]$ in $\mathbf{L^1_{loc}}([0, +\infty[\times\mathbb{R}; [0, L])$. Moreover, $(\tilde{F}^n)_{n\in\mathbb{N}}$ converges to F a.e. on $[0, +\infty[\, \times \mathbb{R}$.*

Proof. We first observe that by Lemma 2.1 for any fixed $t \geq 0$, the map $z \mapsto X(t, z)$ is strictly increasing and for all $z \in [0, L]$
$$z + X(t, 0) \leq X(t, z) \leq \bar{x}_{\max} + v_{\max}\, t - L + z.$$
Thus, F is well defined. The convergence of $(\hat{F}^n)_{n\in\mathbb{N}}$ and $(\tilde{F}^n)_{n\in\mathbb{N}}$ to F follows from the basic property (1.12) of the scaled Wasserstein distance and from Proposition 2.1. Indeed, for any $T > 0$ we have
$$\begin{aligned}\lim_{n\to+\infty} \int_0^T \int_{\mathbb{R}} \left|\hat{F}^n(t, x) - F(t, x)\right| \mathrm{d}x\, \mathrm{d}t &= \lim_{n\to+\infty} \int_0^T \int_0^L \left|\hat{X}^n(t, z) - X(t, z)\right| \mathrm{d}z\, \mathrm{d}t = 0,\\ \lim_{n\to+\infty} \int_0^T \int_{\mathbb{R}} \left|\tilde{F}^n(t, x) - F(t, x)\right| \mathrm{d}x\, \mathrm{d}t &= \lim_{n\to+\infty} \int_0^T \int_0^L \left|\tilde{X}^n(t, z) - X(t, z)\right| \mathrm{d}z\, \mathrm{d}t = 0.\end{aligned}$$
Finally, $(\tilde{F}^n)_{n\in\mathbb{N}}$ converges to F a.e. on $[0, +\infty[\, \times \mathbb{R}$ because $\tilde{F}^{n+1}(t, x) \leq \tilde{F}^n(t, x)$ for all $t \geq 0$ and $x \in \mathbb{R}$. ◇

Lemma 2.2. *For all $t \geq 0$ and for a.e. $x_1, x_2 \in \mathbb{R}$ with $x_1 < x_2$ we have*
$$0 \leq F(t, x_2) - F(t, x_1) \leq x_2 - x_1. \tag{2.10}$$

Proof. Fix $x_1 < x_2$ and denote $z_1 = F(t, x_1) \leq z_2 = F(t, x_2)$. Since the lower bound is obvious, it is sufficient to prove that
$$z_2 - z_1 \leq x_2 - x_1.$$
If $z_1 = z_2$, then there is nothing to prove. Assume therefore that $z_1 \neq z_2$ and fix $\eta \in\,]0, z_2 - z_1[$. By definition, $X(t, z) = \mathcal{X}[F](t, z) = \inf\{x \in \mathbb{R} : F(t, x) > z\}$. Since $F(t, x_2) = z_2 > z_2 - \eta$, we have that $X(t, z_2 - \eta) \leq x_2$. Moreover, $X(t, z_1) \geq x_1$ because $z \mapsto X(t, z)$ is strictly increasing and right continuous. Therefore, by Lemma 2.1 we have
$$x_2 - x_1 \geq X(t, z_2 - \eta) - X(t, z_1) \geq z_2 - \eta - z_1.$$
Since $\eta > 0$ is arbitrary, we have $z_2 - z_1 \leq x_2 - x_1$. ◇

Proposition 2.3 (Definition of ρ)**.** *For any $t \geq 0$, let $\rho(t)$ be the distributional derivative of $x \mapsto F(t, x)$, with F defined in Lemma 2.1. Then:*
- *$\rho(t, \cdot) \in \mathcal{M}_L$ for all $t \geq 0$,*
- *$0 \leq \rho(t, x) \leq 1$ for a.e. $t \geq 0$ and $x \in \mathbb{R}$,*
- *$(\tilde{\rho}^n)_{n\in\mathbb{N}}$ and $(\hat{\rho}^n)_{n\in\mathbb{N}}$ converge to ρ in the topology of $\mathbf{L^1_{loc}}([0, +\infty[;\ d_{L,1})$,*

Proof. For any fixed $t \geq 0$, by Lemma 2.2 we have that $x \mapsto F(t,x)$ is a Lipschitz function with $\mathrm{Lip}\,(F(t)) \leq 1$. Hence its weak derivative $\rho(t)$ is well defined in the space of distributions and is essentially bounded with $\|\rho(t)\|_{\mathbf{L}^\infty(\mathbb{R};\mathbb{R})} \leq 1$. Moreover, $x \mapsto F(t,x)$ is non-decreasing, and therefore $\rho(t) \geq 0$ a.e. in $\mathbb{R}$. By Proposition 2.1 and (1.12) we easily obtain that for any $T > 0$

$$\lim_{n\to+\infty} \int_0^T d_{L,1}\left(\hat\rho^n(t), \rho(t)\right) \mathrm{d}t = \lim_{n\to+\infty} \int_0^T \int_0^L \left|\hat X^n(t,z) - X(t,z)\right| \mathrm{d}z\, \mathrm{d}t = 0,$$
$$\lim_{n\to+\infty} \int_0^T d_{L,1}\left(\tilde\rho^n(t), \rho(t)\right) \mathrm{d}t = \lim_{n\to+\infty} \int_0^T \int_0^L \left|\tilde X^n(t,z) - X(t,z)\right| \mathrm{d}z\, \mathrm{d}t = 0.$$

Thus, ρ satisfies also the last condition and $\rho(t) \in \mathcal{M}_L$. ◇

Lemma 2.3 (Definition of $\check\rho$). *There exists $\check\rho$ in $\mathbf{L}^\infty([0,+\infty[\times[0,L];\mathbb{R})$ such that, up to a subsequence, $(\check\rho^n)_{n\in\mathbb{N}}$ converges weakly-* in $\mathbf{L}^\infty([0,+\infty[\times[0,L];\mathbb{R})$ to $\check\rho$.*

Proof. It is sufficient to observe that for any $n > 0$ we have $\|\check\rho^n\|_{\mathbf{L}^\infty([0,+\infty[\times[0,L];\mathbb{R})} \leq 1$ because, by Lemma 1.1, $\|y_i^n\|_{\mathbf{L}^\infty([0,+\infty[;\mathbb{R})} \leq 1$. ◇

We conclude this subsection by checking that the scheme is consistent with the prescribed initial condition in the limit.

Proposition 2.4. *The sequences $(\tilde\rho^n|_{t=0})_{n\in\mathbb{N}}$ and $(\hat\rho^n|_{t=0})_{n\in\mathbb{N}}$ both converge to $\bar\rho$ in the $d_{L,1}$–Wasserstein distance.*

Proof. By definitions (1.17) and (1.18) we have that

$$\hat\rho^n(0,x) = \sum_{i=0}^{N_n-1} \bar y_i^n\, \chi_{\left[\bar x_i^n,\, \bar x_{i+1}^n\right[}(x), \qquad\qquad \tilde\rho^n(0,x) = \ell_n \sum_{i=0}^{N_n-1} \delta_{\bar x_i^n}(x).$$

Therefore $F_{\hat\rho^n|_{t=0}} = \hat F^n|_{t=0}$, $F_{\tilde\rho^n|_{t=0}} = \tilde F^n|_{t=0}$ and by (1.12), (1.16c) we have

$$\begin{aligned} d_{L,1}(\tilde\rho^n|_{t=0}, \hat\rho^n|_{t=0}) &= \left\|\tilde F^n|_{t=0} - \hat F^n|_{t=0}\right\|_{\mathbf{L}^1(\mathbb{R};\mathbb{R})} = \sum_{i=0}^{N_n-2} \int_{\bar x_i^n}^{\bar x_{i+1}^n} \left[\ell_n - \bar y_i^n\left[x - \bar x_i^n\right]\right] \mathrm{d}x \\ &= \ell_n \sum_{i=0}^{N_n-2} \int_{\bar x_i^n}^{\bar x_{i+1}^n} \frac{\bar x_{i+1}^n - x}{\bar x_{i+1}^n - \bar x_i^n}\, \mathrm{d}x \leq \ell_n \left[\bar x_{\max} - \bar x_{\min}\right]. \end{aligned}$$

Hence, it is sufficient to prove that $(\tilde\rho^n|_{t=0})_{n\in\mathbb{N}}$ converges to $\bar\rho$ in the $d_{L,1}$–Wasserstein distance. By (1.13) we have that

$$\begin{aligned} d_{L,1}(\tilde\rho^n|_{t=0}, \bar\rho) &= \left\|\tilde F^n|_{t=0} - F_{\bar\rho}\right\|_{\mathbf{L}^1(\mathbb{R};\mathbb{R})} = \sum_{i=0}^{N_n-2} \int_{\bar x_i^n}^{\bar x_{i+1}^n} \left[\ell_n\,(i+1) - \int_{-\infty}^x \bar\rho(y)\,\mathrm{d}y\right] \\ &= \sum_{i=0}^{N_n-2} \int_{\bar x_i^n}^{\bar x_{i+1}^n} \left[\ell_n - \int_{\bar x_i^n}^x \bar\rho(y)\,\mathrm{d}y\right] \mathrm{d}x \leq \ell_n \left[\bar x_{\max} - \bar x_{\min}\right] \end{aligned}$$

and this concludes the proof. ◇

2.2. **BV estimates and discrete Oleinik condition.** Let us sum up what we have proven so far. The family of empirical measures $(\tilde{\rho}^n)_{n\in\mathbb{N}}$ converges in the scaled 1–Wasserstein sense to a limit ρ, with $\rho \in \mathbf{L}^\infty\left([0,+\infty[\,;\mathcal{M}_L\right)$ and $0 \le \rho \le 1$ almost everywhere. Moreover, the empirical measure $\tilde{\rho}^n$ has a pseudo-inverse distribution function $\tilde{X}^n$ satisfying the PDE

$$\tilde{X}^n_t(t,z) = v\left(\check{\rho}^n(t,z)\right), \qquad (t,z) \in [0,+\infty[\times[0,L],$$

with the family $(\check{\rho}^n)_{n\in\mathbb{N}}$ being weakly–$*$ compact in $\mathbf{L}^\infty\left([0,+\infty[\times[0,L];[0,1]\right)$. The Wasserstein topology is a proper tool to pass to the limit the time derivative term in the above PDE, as this term is linear. But on the other hand, the weak–$*$ topology is too weak to pass to the limit $v(\check{\rho}^n)$ for a general nonlinear v. Moreover, there is the additional difficulty of having to check that the two limits are related in some sense.

A typical way to overcome the difficulty stated above is to provide a **BV** estimate for the approximating sequence $(\check{\rho}^n)_{n\in\mathbb{N}}$. We tackle this task in two ways. First of all, we perform a direct estimate of the total variation of $\check{\rho}^n$, and prove that such a quantity decreases in time, and is therefore uniformly bounded provided the initial datum $\bar{\rho}$ is **BV**. However, this result is only partly satisfactory, as it is well known that the solution ρ to (1.4a) is **BV** even for an initial datum in $\mathbf{L}^1\cap\mathbf{L}^\infty$. We shall therefore prove that a uniform **BV** estimate of $\check{\rho}^n$ is available for an initial datum in $\mathbf{L}^1\cap\mathbf{L}^\infty$ provided the additional property (V4) of v is prescribed. The latter task is performed by a one-sided estimate of the difference quotients of $\check{\rho}^n$, in the spirit of a discrete version of the classical Oleinik-type condition (1.8), which can be considered as the main technical achievement of this paper.

We start with the following proposition.

Proposition 2.5 (**BV** contractivity for **BV** initial data). *Assume v satisfies (V1) and (V2). If $\bar{\rho}$ satisfies (InBV), then for any $n \in \mathbb{N}$*

$$\mathrm{TV}\left[\hat{\rho}^n(t)\right] = \mathrm{TV}\left[\check{\rho}^n(t)\right] \le \mathrm{TV}\left[\bar{\rho}\right] \qquad \textit{for all } t \ge 0.$$

Proof. For notational simplicity, we shall omit the dependence on t and n whenever not necessary. By construction, see (1.15) and (1.17), we have that

$$\begin{aligned}\mathrm{TV}\left[\hat{\rho}(0)\right] &= \bar{y}_0 + \bar{y}_{N-1} + \sum_{i=0}^{N-2} \left|\bar{y}_i - \bar{y}_{i+1}\right| \\ &= \fint_{\bar{x}_{\min}}^{\bar{x}_1} \bar{\rho}(y)\,\mathrm{d}y + \fint_{\bar{x}_{N-1}}^{\bar{x}_{\max}} \bar{\rho}(y)\,\mathrm{d}y + \sum_{i=0}^{N-2}\left|\fint_{\bar{x}_i}^{\bar{x}_{i+1}} \bar{\rho}(y)\,\mathrm{d}y - \fint_{\bar{x}_{i+2}}^{\bar{x}_{i+1}} \bar{\rho}(y)\,\mathrm{d}y\right| \le \mathrm{TV}\left[\bar{\rho}\right].\end{aligned}$$

Moreover

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\mathrm{TV}\left[\hat{\rho}(t)\right] &= \frac{\mathrm{d}}{\mathrm{d}t}\left[y_0 + y_{N-1} + \sum_{i=0}^{N-2}\left|y_i - y_{i+1}\right|\right] = \dot{y}_0 + \dot{y}_{N-1} + \sum_{i=0}^{N-2}\mathrm{sgn}\left[y_i - y_{i+1}\right]\left[\dot{y}_i - \dot{y}_{i+1}\right] \\ &= \Big[1 + \mathrm{sgn}\left[y_0 - y_1\right]\Big]\dot{y}_0 + \Big[1 - \mathrm{sgn}\left[y_{N-2} - y_{N-1}\right]\Big]\dot{y}_{N-1} + \sum_{i=1}^{N-2}\Big[\mathrm{sgn}\left[y_i - y_{i+1}\right] - \mathrm{sgn}\left[y_{i-1} - y_i\right]\Big]\dot{y}_i.\end{aligned}$$

We claim that the latter right hand side above is ≤ 0. Indeed, assumptions (V1) and (V2) together with (1.16) imply

$$\begin{aligned}\Big[1 + \mathrm{sgn}\left[y_0 - y_1\right]\Big]\dot{y}_0 &= -\Big[1 + \mathrm{sgn}\left[y_0 - y_1\right]\Big]\frac{y_0^2}{\ell}\left[v(y_1) - v(y_0)\right] \le 0, \\ \Big[1 - \mathrm{sgn}\left[y_{N-2} - y_{N-1}\right]\Big]\dot{y}_{N-1} &= -\Big[1 - \mathrm{sgn}\left[y_{N-2} - y_{N-1}\right]\Big]\frac{y_{N-1}^2}{\ell}\left[v_{\max} - v(y_{N-1})\right] \le 0,\end{aligned}$$

$$\Big[\operatorname{sgn}[y_i - y_{i+1}] - \operatorname{sgn}[y_{i-1} - y_i]\Big]\,\dot{y}_i = -\Big[\operatorname{sgn}[y_i - y_{i+1}] - \operatorname{sgn}[y_{i-1} - y_i]\Big]\,\frac{y_i^2}{\ell}\,[v(y_{i+1}) - v(y_i)] \le 0.$$

Therefore, $\mathrm{TV}\,[\hat{\rho}(t)] \le \mathrm{TV}\,[\bar{\rho}]$ for all $t \ge 0$. Finally, since $\check{\rho}^n = \hat{\rho} \circ \hat{X}$, $\check{\rho}^n$ is piecewise constant and on $i\,\ell$ has the same traces as $\hat{\rho}$ on x_i, the statement for $\check{\rho}^n$ follows easily. $\diamond$

We now perform our discrete Oleinik-type condition, which holds for general initial data in $\mathbf{L}^1 \cap \mathbf{L}^\infty$.

Lemma 2.4 (Discrete Oleinik-type condition)**.** *Assume v satisfies (V1), (V2), and (V3), and let $\bar{\rho}$ satisfy (In). Then, for any $i = 0, \ldots, N_n - 2$ we have*

$$t\,y_i^n(t)\,\left[v\left(y_{i+1}^n(t)\right) - v\left(y_i^n(t)\right)\right] \le \ell_n \qquad \text{for all } t \ge 0. \tag{2.11}$$

Proof. For notational simplicity, we shall omit the dependence on t and n whenever not necessary. Let

$$\begin{aligned} z_i &:= t\,y_i\,\left[v\left(y_{i+1}\right) - v\left(y_i\right)\right], && i = 0, \ldots, N-2,\\ z_{N-1} &:= t\,y_{N-1}\,[v_{\max} - v(y_{N-1})]. \end{aligned}$$

• **Step 0: $\mathbf{z_{N-1} \le \boldsymbol{\ell}}$.** By (1.16a) and (V1)

$$\begin{aligned} \dot{z}_{N-1} &= y_{N-1}\,[v_{\max} - v(y_{N-1})] + t\,\dot{y}_{N-1}\,[v_{\max} - v(y_{N-1})] - t\,y_{N-1}\,v'(y_{N-1})\,\dot{y}_{N-1}\\ &= y_{N-1}\,[v_{\max} - v(y_{N-1})] - \frac{t\,y_{N-1}^2}{\ell}\,[v_{\max} - v(y_{N-1})]^2 + \frac{t\,v'(y_{N-1})\,y_{N-1}^3}{\ell}\,[v_{\max} - v(y_{N-1})]\\ &\le y_{N-1}\,[v_{\max} - v(y_{N-1})]\left[1 - \frac{z_{N-1}}{\ell}\right]. \end{aligned}$$

Since $z_{N-1}(0) = 0$, from the above estimate we get $z_{N-1}(t) \le \ell$ for all $t \ge 0$.

• **Step 1: $\mathbf{z_{i+1} \le \boldsymbol{\ell} \Rightarrow z_i \le \boldsymbol{\ell}}$.** Let $i \in \{0, \ldots, N-3\}$ and assume $z_{i+1} \le \ell$. From (1.16b) and (V1) we get

$$\begin{aligned} \dot{z}_i =\,& y_i\,\left[v\left(y_{i+1}\right) - v\left(y_i\right)\right] + t\,\dot{y}_i\left[v\left(y_{i+1}\right) - v\left(y_i\right)\right] + t\,y_i\left[v'(y_{i+1})\,\dot{y}_{i+1} - v'(y_i)\,\dot{y}_i\right]\\ =\,& y_i\,\left[v\left(y_{i+1}\right) - v\left(y_i\right)\right] - \frac{t\,y_i^2}{\ell}\,[v(y_{i+1}) - v(y_i)]^2\\ &+ t\,y_i\left[-\frac{v'(y_{i+1})\,y_{i+1}^2}{\ell}\,[v(y_{i+2}) - v(y_{i+1})] + \frac{v'(y_i)\,y_i^2}{\ell}\,[v(y_{i+1}) - v(y_i)]\right]\\ =\,& y_i\,\left[v\left(y_{i+1}\right) - v\left(y_i\right)\right] - \frac{y_i}{\ell}\,[v(y_{i+1}) - v(y_i)]\,z_i - \frac{v'(y_{i+1})\,y_i\,y_{i+1}}{\ell}\,z_{i+1} + \frac{v'(y_i)\,y_i^2}{\ell}\,z_i. \end{aligned}$$

Since $\operatorname{sgn}_+[z_i] = \operatorname{sgn}_+\left[v\left(y_{i+1}\right) - v\left(y_i\right)\right] = \operatorname{sgn}_+[y_i - y_{i+1}]$ for all $t > 0$, from the assumption on z_{i+1} we easily obtain

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t}[z_i]_+ =\,& y_i\,\left[v\left(y_{i+1}\right) - v\left(y_i\right)\right]_+ - \frac{y_i}{\ell}\,[v(y_{i+1}) - v(y_i)]_+\,[z_i]_+\\ &- \frac{v'(y_{i+1})\,y_i\,y_{i+1}}{\ell}\,\operatorname{sgn}_+[z_i]\,z_{i+1} + \frac{v'(y_i)\,y_i^2}{\ell}\,[z_i]_+\\ \le\,& y_i\,\left[v\left(y_{i+1}\right) - v\left(y_i\right)\right]_+\left[1 - \frac{[z_i]_+}{\ell}\right] - v'(y_{i+1})\,y_i\,y_{i+1}\,\operatorname{sgn}_+[z_i] + \frac{v'(y_i)\,y_i^2}{\ell}\,[z_i]_+. \end{aligned}$$

Condition (V3) implies that the function $y \mapsto y\,v'(y)$ is non-increasing, which gives

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t}[z_i]_+ &\le y_i\,\left[v\left(y_{i+1}\right) - v\left(y_i\right)\right]_+\left[1 - \frac{[z_i]_+}{\ell}\right] - v'(y_i)\,y_i^2\,\operatorname{sgn}_+[z_i] + \frac{v'(y_i)\,y_i^2}{\ell}\,[z_i]_+\\ &= y_i\left[\left[v\left(y_{i+1}\right) - v\left(y_i\right)\right]_+ - v'(y_i)\,y_i\right]\left[1 - \frac{[z_i]_+}{\ell}\right]. \end{aligned}$$

Now, as $v' \le 0$, and since $z_i(0) = 0$, we get that $z_i(t)_+ \le \ell$ for all $t \ge 0$.

• **Step 2:** $\mathbf{z_{N-2} \leq \ell}$. From analogous computations as in previous step, we get

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}[z_{N-2}]_+ =& y_{N-2}\,\left[v\left(y_{N-1}\right)-v\left(y_{N-2}\right)\right]_+ - \frac{y_{N-2}}{\ell}\left[v(y_{N-1})-v(y_{N-2})\right]_+[z_{N-2}]_+\\ &- \frac{v'(y_{N-1})\,y_{N-2}\,y_{N-1}}{\ell}\,\mathrm{sgn}_+[z_{N-2}]\,z_{N-1} + \frac{v'(y_{N-2})\,y_{N-2}^2}{\ell}\,[z_{N-2}]_+,\end{aligned}$$

and we can use the monotonicity of $y \mapsto y\,v'(y)$ and Step 0 to get

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}[z_{N-2}]_+ \leq& y_{N-2}\,\left[v\left(y_{N-1}\right)-v\left(y_{N-2}\right)\right]_+\left[1-\frac{[z_{N-2}]_+}{\ell}\right]\\ &- v'(y_{N-2})\,y_{N-2}^2\,\mathrm{sgn}_+[z_{N-2}] + \frac{v'(y_{N-2})\,y_{N-2}^2}{\ell}\,[z_{N-2}]_+\\ =& y_{N-2}\left[\left[v\left(y_{N-1}\right)-v\left(y_{N-2}\right)\right]_+ - v'(y_{N-2})\,y_{N-2}\right]\left[1-\frac{[z_{N-2}]_+}{\ell}\right].\end{aligned}$$

Again, $v' \leq 0$ and $z_{N-2}(0)=0$ imply that $z_{N-2}(t)_+ \leq \ell$ for all $t \geq 0$.
• **Conclusion**. The estimate (2.11) is proven recursively: Step 2 provides the first step with $i = N-2$, whereas Step 1 proves that the estimate holds for all $i \in \{0,\ldots,N-3\}$. ◇

Corollary 2.1. *Assume v satisfies (V1), (V2), and (V3), and let $\bar\rho$ satisfy (In). Then, for any $i \in \{0,\ldots,N-2\}$ we have*

$$v\left(\hat\rho^n\left(t,x_i^n(t)\right)\right) - v\left(\hat\rho^n\left(t,x_{i+1}^n(t)\right)\right) \leq \frac{x_{i+1}^n(t)-x_i^n(t)}{t} \qquad \textit{for all } t>0. \tag{2.12}$$

Proof. The statement follows from Lemma 2.4 and from the definitions of $\hat\rho^N$ and y_i. ◇

In the following proposition we prove uniform bounds on the total variation of $v\left(\check\rho^n\right)$ and $v\left(\hat\rho^n\right)$. Let us emphasize that the regularising effect $\mathbf{L}^\infty \mapsto \mathbf{BV}$ implies that the $\mathbf{BV}$ estimate necessarily *blows up* as $t \searrow 0$.

Proposition 2.6 (Uniform $\mathbf{BV}$ estimates for $v\left(\check\rho^n\right)$ and $v\left(\hat\rho^n\right)$)**.** *Assume v satisfies the properties (V1), (V2), and (V3), and let $\bar\rho$ satisfy (In). Let $\delta > 0$. Then*

(i) the sequence $\left(v\left(\hat\rho^n\right)\right)_{n\in\mathbb{N}}$ is uniformly bounded in $\mathbf{L}^\infty\left([\delta,+\infty[\,;\mathbf{BV}(\mathbb{R};[0,v_{\max}])\right)$;
(ii) the sequence $\left(v\left(\check\rho^n\right)\right)_{n\in\mathbb{N}}$ is uniformly bounded in $\mathbf{L}^\infty\left([\delta,+\infty[\,;\mathbf{BV}([0,L];[0,v_{\max}])\right)$.

More precisely, for any $n \in \mathbb{N}$

$$\mathrm{TV}\left[v(\hat\rho^n(t))\right] = \mathrm{TV}\left[v(\check\rho^n(t))\right] \leq C_\delta \qquad \textit{for all } t \geq \delta,$$

where $C_\delta := \left[3\,v_{\max} + 2\,\frac{|\bar x_{\max}|+|\bar x_{\min}|}{\delta}\right]$.

Proof. For notational simplicity, we shall omit the dependence on n. We set

$$\hat\sigma(t,x) := v(\hat\rho(t,x)) + \frac{1}{t}\sum_{i=0}^{N-1} x_i(t)\,\chi_{[x_i(t),\,x_{i+1}(t)[}(x) \qquad \text{for all } x \in \mathbb{R}.$$

We claim that, for any fixed $t \geq 0$, the map $x \mapsto \hat\sigma(t,x)$ is a piecewise constant, non-decreasing function on $[x_0(t),x_N(t)[$. To see this, we first notice that the map $x \mapsto \hat\sigma(t,x)$ is constant on the interval $[x_i(t),x_{i+1}(t)[$, $i = 0,\ldots,N-1$. On the other hand, $\hat\sigma(t)$ is non-decreasing on the potential discontinuity points $x_i(t)$, $i = 1,\ldots,N-1$, in view of (2.12). Now, from (1.14) and the discrete maximum principle in Lemma 1.1 we know that for any $x \in [x_0(t),x_N(t)[$

$$\frac{\bar x_{\min}}{t} \leq \hat\sigma(t,x) \leq v_{\max} + \frac{1}{t}\left[\bar x_{\max} + v_{\max}\,t\right] = 2v_{\max} + \frac{\bar x_{\max}}{t}.$$

Hence $\hat{\sigma}$ is uniformly bounded in $\mathbf{L}^\infty\left([\delta,+\infty[\,;\mathbf{BV}(\mathbb{R};\mathbb{R})\right)$ with

$$\sup_{t\geq\delta}\mathrm{TV}\left[\hat{\sigma}(t)\right]\leq\left[2v_{\max}+\frac{|\bar{x}_{\max}|+|\bar{x}_{\min}|}{\delta}\right].$$

Therefore, also $v(\hat{\rho})$ is uniformly bounded in $\mathbf{L}^\infty\left([\delta,+\infty[\,;\mathbf{BV}(\mathbb{R};\mathbb{R})\right)$ because by triangular inequality

$$\begin{aligned}\mathrm{TV}\left[v(\hat{\rho}(t))\right]&\leq\mathrm{TV}\left[\hat{\sigma}(t)\right]+\mathrm{TV}\left[\frac{1}{t}\sum_{i=0}^{N-1}x_i(t)\,\chi_{[x_i(t),\,x_{i+1}(t)[}\right]\\&=\mathrm{TV}\left[\hat{\sigma}(t)\right]+\frac{1}{t}\left|\bar{x}_{\max}-\bar{x}_{\min}+v_{\max}\,t\right|\leq C_\delta,\end{aligned}$$

for $t\geq\delta$. Finally, since $\check{\rho}^n=\hat{\rho}\circ\hat{X}$, $\check{\rho}^n$ is piecewise constant and on $i\,\ell$ has the same traces as $\hat{\rho}$ on x_i, the statement for $\check{\rho}^n$ follows easily. ◇

2.3. **Time continuity and compactness.**

Proposition 2.7 (Uniform $\mathbf{L}^1$–continuity in time of $\check{\rho}^n$)**.** *For any $\delta>0$ we have*

$$\int_0^L\left|\check{\rho}^n(t,z)-\check{\rho}^n(s,z)\right|\mathrm{d}z\leq\left[C_\delta+v_{\max}\right]|t-s|\qquad\textit{for all } t,s\geq\delta,$$

with C_δ defined in Proposition 2.6.

Proof. By (1.16), we compute for $t>s>\delta$,

$$\begin{aligned}&\int_0^L\left|\check{\rho}^n(t,z)-\check{\rho}^n(s,z)\right|\mathrm{d}z=\sum_{i=0}^{N_n-1}\ell_n\left|y_i^n(t)-y_i^n(s)\right|=\sum_{i=0}^{N_n-1}\ell_n\left|\int_s^t\dot{y}_i^n(\tau)\,\mathrm{d}\tau\right|\\&=\sum_{i=0}^{N_n-2}\left|\int_s^t y_i^n(\tau)^2\left[v\left(y_{i+1}^n(\tau)\right)-v\left(y_i^n(\tau)\right)\right]\mathrm{d}\tau\right|+\int_s^t y_{N_n-1}^n(\tau)^2\left[v_{\max}-v\left(y_{N_n-1}^n(\tau)\right)\right]\mathrm{d}\tau.\end{aligned}$$

Therefore, by Lemma 1.1

$$\begin{aligned}&\int_0^L\left|\check{\rho}^n(t,z)-\check{\rho}^n(s,z)\right|\mathrm{d}z\\&\leq\int_s^t\left[\sum_{i=0}^{N_n-2}y_i^n(\tau)^2\left|v\left(y_{i+1}^n(\tau)\right)-v\left(y_i^n(\tau)\right)\right|+y_{N_n-1}^n(\tau)^2\left[v_{\max}-v\left(y_{N_n-1}^n(\tau)\right)\right]\right]\mathrm{d}\tau\\&\leq\int_s^t\left[\sum_{i=0}^{N_n-2}\left|v\left(y_{i+1}^n(\tau)\right)-v\left(y_i^n(\tau)\right)\right|+v_{\max}\right]\mathrm{d}\tau\leq\int_s^t\left[\mathrm{TV}\left[v\left(\check{\rho}^n(\tau)\right)\right]+v_{\max}\right]\mathrm{d}\tau.\end{aligned}$$

Then it is sufficient to apply the estimate in Proposition 2.6 to complete the proof. ◇

Proposition 2.8 (Uniform Wasserstein time continuity of $\hat{\rho}^n$)**.** *For any $n\in\mathbb{N}$ we have*

$$d_{L,1}\left(\hat{\rho}^n(t),\hat{\rho}^n(s)\right)\leq 2\,L\,v_{\max}\,|t-s|\qquad\textit{for all } s,t\geq0.$$

Proof. By (1.12), (1.15) and (1.14), we compute for any $t>s\geq0$

$$\begin{aligned}&d_{L,1}\left(\hat{\rho}^n(t),\hat{\rho}^n(s)\right)=\left\|\hat{X}^n(t)-\hat{X}^n(s)\right\|_{\mathbf{L}^1([0,L];\mathbb{R})}=\sum_{i=0}^{N_n-1}\int_{i\,\ell_n}^{(i+1)\,\ell_n}\left[\hat{X}^n(t,z)-\hat{X}^n(s,z)\right]\mathrm{d}z\\&=\sum_{i=0}^{N_n-1}\int_{i\,\ell_n}^{(i+1)\,\ell_n}\left[x_i^n(t)+\frac{z-i\,\ell_n}{y_i^n(t)}-x_i^n(s)-\frac{z-i\,\ell_n}{y_i^n(s)}\right]\mathrm{d}z\end{aligned}$$

$$
\begin{aligned}
&= \sum_{i=0}^{N_n-1} \ell_n \left[x_i^n(t) - x_i^n(s)\right] + \sum_{i=0}^{N_n-1} \left[y_i^n(t)^{-1} - y_i^n(s)^{-1}\right] \int_{i\,\ell_n}^{(i+1)\,\ell_n} (z - i\,\ell_n)\,\mathrm{d}z \\
&= \sum_{i=0}^{N_n-1} \ell_n \int_s^t v\left(y_i^n(\tau)\right) \mathrm{d}\tau + \sum_{i=0}^{N_n-1} \frac{\ell_n^2}{2} \int_s^t \frac{\mathrm{d}}{\mathrm{d}\tau} \left[y_i^n(\tau)^{-1}\right] \mathrm{d}\tau \\
&\le L\, v_{\max}\,(t-s) + \frac{\ell_n}{2} \int_s^t \left[\sum_{i=0}^{N_n-2} \left[v\left(y_{i+1}^n(\tau)\right) - v\left(y_i^n(\tau)\right)\right] + v_{\max} - v\left(y_{N_n-1}^n(\tau)\right)\right] \mathrm{d}\tau \\
&= L\, v_{\max}\,(t-s) + \frac{\ell_n}{2} \int_s^t \left[v_{\max} - v\left(y_0^n(\tau)\right)\right] \mathrm{d}\tau \\
&\le \left[1 + 2^{-n-1}\right] L\, v_{\max}\,(t-s) \le 2\,L\, v_{\max}\,(t-s)
\end{aligned}
$$

and this concludes the proof. ◇

We now recall a generalization of Aubin-Lions lemma, which uses the Wasserstein distance as a replacement of a negative Sobolev norm, proven in [44, Theorem 2], which we present here in a version adapted to our case. In order to have the paper self-contained, we first recall the precise statement of [44, Theorem 2] (see also the adapted version in [20]).

Theorem 2.1 (Theorem 2 from [44])**.** *On a separable Banach space $\mathbb{X}$, let be given*

- *(F)* a normal coercive integrand $\mathfrak{F} : \mathbb{X} \to [0, +\infty]$*, i.e., $\mathfrak{F}$ is lower semi-continuous and its sublevels are relatively compact in $\mathbb{X}$;*
- *(g)* a pseudo-distance $\mathfrak{g} : \mathbb{X} \times \mathbb{X} \to [0, +\infty]$*, i.e., $\mathfrak{g}$ is lower semi-continuous, and if $\nu, \mu \in \mathbb{X}$ are such that $\mathfrak{g}(\nu, \mu) = 0$, $\mathfrak{F}[\nu] < +\infty$ and $\mathfrak{F}[\mu] < +\infty$, then $\nu = \mu$.*

Let further U be a set of measurable functions $\nu : \ ]0, T[\to \mathbb{X}$, with a fixed $T > 0$. Under the hypotheses that

$$
\sup_{\nu \in U} \int_0^T \mathfrak{F}\left[\nu(t)\right] \mathrm{d}t < +\infty \qquad \text{and} \qquad \lim_{h \downarrow 0} \left[\sup_{\nu \in U} \int_0^{T-h} \mathfrak{g}\left(\nu(t+h), \nu(t)\right) \mathrm{d}t\right] = 0, \tag{2.13}
$$

Then U is strongly relatively compact in $\mathbf{L}^1(]0, T[\,; \mathbb{X})$.

Theorem 2.2 (Generalized Aubin-Lions lemma)**.** *Let T, $L > 0$ and $I \subset \mathbb{R}$ be a bounded open convex interval. Assume $w : \mathbb{R} \to \mathbb{R}$ is a Lipschitz continuous and strictly monotone function. Let $(\rho^n)_{n\in\mathbb{N}}$ be a sequence in $\mathbf{L}^\infty\left(]0, T[\times \mathbb{R}; \mathbb{R}\right) \cap \mathcal{M}_L$ such that*

- *(1)* $\rho^n : \ ]0, T[\to \mathbf{L}^1\left(\mathbb{R}; \mathbb{R}\right)$ *is measurable for all $n \in \mathbb{N}$;*
- *(2)* $\operatorname{spt}\left(\rho^n(t)\right) \subseteq I$ *for all $t \in]0, T[$ and $n \in \mathbb{N}$;*
- *(3)* $\sup_{n\in\mathbb{N}} \int_0^T \left[\left\|w\left(\rho^n(t)\right)\right\|_{\mathbf{L}^1(I;\mathbb{R})} + \mathrm{TV}\left[w\left(\rho^n(t)\right)\right]\right] \mathrm{d}t < +\infty$*;*
- *(4) there exists a constant C depending only on T such that $d_{L,1}\left(\rho^n(s), \rho^n(t)\right) \le C\,|t - s|$ for all $s, t \in]0, T[$ and $n \in \mathbb{N}$.*

Then, $(\rho^n)_{n\in\mathbb{N}}$ is strongly relatively compact in $\mathbf{L}^1(]0, T[\times \mathbb{R}; \mathbb{R})$.

Proof. We want to use Theorem 2.1 with

$$
\mathbb{X} := \mathbf{L}^1\left(I; \mathbb{R}\right), \quad \mathfrak{F}[\nu] := \|w(\nu)\|_{\mathbf{L}^1(I;\mathbb{R})} + \mathrm{TV}\left[w(\nu)\right], \quad U := (\rho^n)_{n\in\mathbb{N}}, \quad \mathfrak{g}(\nu, \mu) := \begin{cases} d_{L,1}(\nu, \mu) & \text{if } \ \nu, \mu \in \mathcal{M}_L, \\ +\infty & \text{otherwise.} \end{cases}
$$

We first have to prove that $\mathfrak{F}$, $\mathfrak{g}$ and U satisfy the corresponding hypotheses in Theorem 2.1.
(F) Assume that $(\nu^n)_{n\in\mathbb{N}}$ converges to ν in $\mathbf{L}^1\left(I; \mathbb{R}\right)$. Since w is Lipschitz continuous, $\left(w(\nu^n)\right)_{n\in\mathbb{N}}$ converges to $w(\nu)$ in $\mathbf{L}^1\left(I; \mathbb{R}\right)$. Hence, for the lower semi-continuity of the total variation w.r.t. the $\mathbf{L}^1$–norm, see [23, Theorem 1 on page 172], we have that $\mathrm{TV}\left[w(\nu)\right] \le \liminf_{n\to+\infty} \mathrm{TV}\left[w(\nu^n)\right]$. Thus $\mathfrak{F}\left[\nu\right] \le \liminf_{n\to+\infty} \mathfrak{F}\left[\nu^n\right]$

and $\mathfrak{F}$ is l.s.c. in $\mathbb{X}$. Finally, consider a sequence $(\nu^n)_{n\in\mathbb{N}}$ belonging to a sublevel of $\mathfrak{F}$, namely $\sup_{n\in\mathbb{N}} \mathfrak{F}[\nu^n] < +\infty$. For the compactness of $\mathbf{BV}$ in $\mathbf{L}^1$ on bounded open convex intervals and for basic properties of the $\mathbf{L}^1$–convergence, see [23, Theorem 4 on page 176 and Theorem 5 on page 21], up to a subsequence $(w(\nu^n))_{n\in\mathbb{N}}$ converges to $\bar{w}$ in $\mathbf{L}^1$ and a.e. on I. Since w is continuous and strictly monotone, $(\nu^n)_{n\in\mathbb{N}}$ is uniformly bounded in $\mathbf{L}^\infty$ (consequence of the uniform bound on the total variation) and converges to $\bar{\nu} := w^{-1}(\bar{w})$ a.e. on I and therefore, by the Lebesgue dominated convergence theorem, the convergence is also in $\mathbf{L}^1$.
(g) Proceeding as before and applying lower semi-continuity of the 1–Wasserstein distance w.r.t. the $\mathbf{L}^1$–norm give that $\mathfrak{g}$ is l.s.c. in $\mathbb{X}\times\mathbb{X}$. Finally, if $\mathfrak{F}[\nu] < +\infty$, $\mathfrak{F}[\mu] < +\infty$ and $\mathfrak{g}(\mu,\nu) = 0$, then $w(\mu), w(\nu)$ are in $\mathbf{BV}$, $\nu,\mu \in \mathcal{M}_L$, and $d_{L,1}(\mu,\nu) = 0$. Hence we have $\mu = \nu$.
(U) Conditions in (2.13) follow directly from the hypotheses (3) and (4).
Hence we can apply Theorem 2.1 and obtain the convergence in $\mathbf{L}^1(]0,T[\times I;\mathbb{R})$. Finally, recalling the hypothesis (2) concludes the proof. ◇

2.4. **Convergence to entropy solutions.** In the next proposition we collect the previous compactness results to get strong convergence.

Proposition 2.9. *Let $\check{\rho}$ be defined as in Lemma 2.3 and ρ as in Proposition 2.3. Under the assumptions in Theorem 1.3 we have that*

(i) the sequence $(\check{\rho}^n)_{n\in\mathbb{N}}$ converges to $\check{\rho}$ almost everywhere and strongly in $\mathbf{L^1_{loc}}$ on $]0,+\infty[\times[0,L]$;
(ii) the sequence $(\hat{\rho}^n)_{n\in\mathbb{N}}$ converges to ρ almost everywhere and strongly in $\mathbf{L^1_{loc}}$ on $]0,+\infty[\times\mathbb{R}$;
(iii) if $\bar{\rho}$ satisfies also (InBV), then the sequence $(\check{\rho}^n)_{n\in\mathbb{N}}$ converges to $\check{\rho}$ strongly in $\mathbf{L^1_{loc}}$ on $[0,+\infty[\times[0,L]$.

Proof. We already know from Proposition 2.3 that both $(\hat{\rho}^n)_{n\in\mathbb{N}}$ and $(\tilde{\rho}^n)_{n\in\mathbb{N}}$, defined respectively by (1.17) and (1.18), converge in the topology of $\mathbf{L^1_{loc}}([0,+\infty[;\, d_{L,1})$ to the density $\rho \in \mathbf{L}^\infty([0,+\infty[;\mathcal{M}_L)$ with $0 \le \rho \le 1$. From Proposition 2.1 we know that both $(\hat{X}^n)_{n\in\mathbb{N}}$ and $(\tilde{X}^n)_{n\in\mathbb{N}}$, defined respectively by (2.3) and (2.5), converge strongly in $\mathbf{L^1_{loc}}([0,+\infty[\times[0,L];\mathbb{R})$ to $X \in \mathbf{L}^\infty([0,+\infty[\times[0,L];\mathbb{R})$, the pseudo-inverse of F, the cumulative distribution of ρ. Finally, from Lemma 2.3 we know that $(\check{\rho}^n)_{n\in\mathbb{N}}$, defined by (2.6), converges to $\check{\rho}$ weakly-* in $\mathbf{L}^\infty([0,+\infty[\times[0,L];\mathbb{R})$.

• **Step 1.** *Strong convergence of $(\check{\rho}^n)_{n\in\mathbb{N}}$ for general initial datum in $\mathcal{M}_L\cap\mathbf{L}^\infty$.*
Let $\bar{\rho}$ satisfy (In). For any fixed $\delta > 0$, we know from Proposition 2.6 that $(v(\check{\rho}^n))_{n\in\mathbb{N}}$ is uniformly bounded in $\mathbf{L}^\infty([\delta,+\infty[;\,\mathbf{BV}([0,L];[0,v_{\max}]))$. Furthermore, from Proposition 2.7 we easily obtain that

$$\int_0^L |v(\check{\rho}^n(t,z)) - v(\check{\rho}^n(s,z))|\,\mathrm{d}z \le \mathrm{Lip}\,(v)\,[C_\delta + v_{\max}]\,|t-s| \qquad \text{for all } t,s \ge \delta.$$

Therefore, we can once again apply Helly's theorem in the form [11, Theorem 2.4] to get that $(v(\check{\rho}^n))_{n\in\mathbb{N}}$ is strongly compact in $\mathbf{L^1_{loc}}([\delta,+\infty[\times[0,L];[0,v_{\max}])$. Hence, by the monotonicity of v and the uniqueness of the limit in the $\mathbf{L^1_{loc}}([\delta,+\infty[;\, d_{L,1})$ topology, up to a subsequence $(\check{\rho}^n)_{n\in\mathbb{N}}$ converges strongly in $\mathbf{L^1_{loc}}$ and a.e. on $[\delta,+\infty[\times[0,L]$ to $\check{\rho}$. Finally, since $\delta > 0$ is arbitrary, the proof of (i) is complete.

• **Step 2.** *Strong convergence of $(\hat{\rho}^n)_{n\in\mathbb{N}}$ for general initial datum in $\mathcal{M}_L\cap\mathbf{L}^\infty$.*
Let $\bar{\rho}$ satisfy (In) and fix $T,\delta > 0$ with $\delta < T$. We want to prove that $(\rho^n)_{n\in\mathbb{N}}$ with $\rho^n(t,x) := \hat{\rho}^n(t+\delta,x)$ satisfies the hypotheses of Theorem 2.2 with $I =]\bar{x}_{\min} - 1, \bar{x}_{\max} + v_{\max}(T+\delta) + 1[$ and $w = v$. The hypotheses (1) and (2) are satisfied because by definition (1.17) we have that $\|\hat{\rho}^n(t)\|_{\mathbf{L}^1(\mathbb{R};\mathbb{R})} = L$ and $\mathrm{spt}\,(\hat{\rho}^n(t)) \subset I$ for all $t \in [0,T+\delta]$. By Proposition 2.6, the hypothesis (3) holds true because

$$\int_\delta^{T+\delta} \left[\|v(\hat{\rho}^n(t))\|_{\mathbf{L}^1(I;\mathbb{R})} + \mathrm{TV}\,[v(\hat{\rho}^n(t))] \right] \mathrm{d}t \le [v_{\max}\,|I| + C_\delta]\,T.$$

Finally, the hypothesis (4) follows directly from Proposition 2.8. Hence, we can apply Theorem 2.2 to obtain that $(\hat{\rho}^n)_{n\in\mathbb{N}}$ is strongly compact in $\mathbf{L}^1(]\delta,T[\times\mathbb{R};\mathbb{R})$. By the uniqueness of the limit in the $\mathbf{L}^1(]\delta,T[;\, d_{L,1})$ topology, up to a subsequence $(\hat{\rho}^n)_{n\in\mathbb{N}}$ converges strongly in $\mathbf{L}^1$ and a.e. on $]\delta,T[\times\mathbb{R}$ to ρ. Finally, since $T > \delta > 0$ are arbitrary, the proof of (ii) is complete.

• **Step 3.** *Strong convergence for initial datum in* **BV**.
Let $\bar\rho$ satisfy (InBV). The result in Proposition 2.5 ensures that both $(\hat\rho^n)_{n\in\mathbb{N}}$ and $(\check\rho^n)_{n\in\mathbb{N}}$ are uniformly bounded in $\mathbf{L}^\infty([0,+\infty[;\mathbf{BV}(\mathbb{R};[0,1]))$. Hence, we can repeat the proof of Proposition 2.7 (we omit the details) to obtain that

$$\int_0^L |\check\rho^n(t,z)-\check\rho^n(s,z)|\,\mathrm{d}z \le [\mathrm{Lip}\,(v)\,\mathrm{TV}(\bar\rho)+v_{\max}]\,|t-s| \qquad \text{for all } t,s\ge 0.$$

Therefore, Helly's theorem implies the desired compactness. Moreover, we can use Theorem 2.2 with w being the identity function on $[0,1]$, and obtain the desired compactness of $\hat\rho^n$. ⋄

We now prove that the two limits $\check\rho$ and ρ are related.

Proposition 2.10. *Let F be the cumulative distribution of ρ as defined in Proposition 2.2. Then*

$$\check\rho\,(t,F(t,x))=\rho(t,x) \qquad \textit{for a.e. } (t,x) \textit{ in } \mathrm{spt}(\rho).$$

Proof. By definition (2.6) and Lemma 1.2, for any $\varphi\in\mathbf{C}^\infty_{\mathbf{c}}([0,T]\times\mathbb{R};\mathbb{R})$ we have

$$\begin{aligned}\int_0^T\int_0^L \check\rho^n(t,z)\,\varphi\left(t,\hat X^n(t,z)\right)\mathrm{d}z\,\mathrm{d}t &= \int_0^T\int_0^L \hat\rho^n\left(t,\hat X^n(t,z)\right)\,\varphi\left(t,\hat X^n(t,z)\right)\mathrm{d}z\,\mathrm{d}t\\ &=\int_0^T\int_{\mathbb{R}}\hat\rho^n(t,x)^2\,\varphi(t,x)\,\mathrm{d}x\,\mathrm{d}t.\end{aligned}$$

By extracting the a.e. convergent subsequence provided in Proposition 2.9, we can send $n\to+\infty$ in the above identity and use the Lebesgue dominated convergence theorem (as the support of $\check\rho^n$ and $\hat\rho^n$ are uniformly bounded w.r.t. n) to get

$$\int_0^T\int_0^L\check\rho(t,z)\,\varphi\,(t,X(t,z))\,\mathrm{d}z\,\mathrm{d}t=\int_0^T\int_{\mathbb{R}}\rho(t,x)^2\,\varphi(t,x)\,\mathrm{d}x\,\mathrm{d}t.$$

By changing variable $z=F(t,x)$ in the first integral above, we get

$$\int_0^T\int_{\mathbb{R}}\check\rho\,(t,F(t,x))\,\rho(t,x)\,\varphi(t,x)\,\mathrm{d}x\,\mathrm{d}t=\int_0^T\int_{\mathbb{R}}\rho(t,x)^2\,\varphi(t,x)\,\mathrm{d}x\,\mathrm{d}t, \tag{2.14}$$

and this ends the proof. ⋄

In the next proposition we prove that ρ is a weak solution in the sense of (1.5).

Proposition 2.11. *The limit function ρ defined in Proposition 2.3 is a weak solution in the sense of* (1.5).

Proof. Let $\varphi\in\mathbf{C}^\infty_{\mathbf{c}}([0,+\infty[\times\mathbb{R};\mathbb{R})$. By (1.14), (2.5) and (2.6), for all n we have

$$\begin{aligned}&\int_0^{+\infty}\int_0^L\left[v\left(\check\rho^n(t,z)\right)\varphi_x\left(t,\tilde X^n(t,z)\right)\right]\mathrm{d}z\,\mathrm{d}t=\sum_{i=0}^{N_n-1}\int_0^{+\infty}\int_{i\,\ell_n}^{(i+1)\,\ell_n}\left[v\left(y_i^n(t)\right)\varphi_x\left(t,x_i^n(t)\right)\right]\mathrm{d}z\,\mathrm{d}t\\ &=\sum_{i=0}^{N_n-1}\int_0^{+\infty}\int_{i\,\ell_n}^{(i+1)\,\ell_n}\left[\dot x_i^n(t)\,\varphi_x\left(t,x_i^n(t)\right)\right]\mathrm{d}z\,\mathrm{d}t=\sum_{i=0}^{N_n-1}\int_0^{+\infty}\int_{i\,\ell_n}^{(i+1)\,\ell_n}\left[\frac{\mathrm{d}}{\mathrm{d}t}\varphi\left(t,x_i^n(t)\right)-\varphi_t\left(t,x_i^n(t)\right)\right]\mathrm{d}t\\ &=-\int_0^L\varphi\left(0,\tilde X^n(0,z)\right)\mathrm{d}z-\int_0^{+\infty}\int_0^L\varphi_t\left(t,\tilde X^n(t,z)\right)\mathrm{d}z\,\mathrm{d}t.\end{aligned}$$

Since $(\tilde X^n)_{n\in\mathbb{N}}$ and $(\check\rho^n)_{n\in\mathbb{N}}$ converge strongly in $\mathbf{L}^1([0,T]\times[0,L];\mathbb{R})$, and in view of Proposition 2.4, we get by sending $n\to+\infty$

$$\int_0^{+\infty}\int_0^L\left[\varphi_t\left(t,X(t,z)\right)+v\left(\check\rho(t,z)\right)\varphi_x\left(t,X(t,z)\right)\right]\mathrm{d}z\,\mathrm{d}t+\int_0^L\varphi\left(0,X_{\bar\rho}(z)\right)\,\mathrm{d}z=0.$$

We now apply the change of variable $x = X(t,z)$, see Lemma 1.2, and obtain

$$\int_0^{+\infty}\int_{\mathbb{R}}\Big[\rho(t,x)\,\varphi_t(t,x)+\rho(t,x)\,v\left(\check{\rho}\left(t,F(t,x)\right)\right)\varphi_x(t,x)\Big]\,\mathrm{d}x\,\mathrm{d}t+\int_{\mathbb{R}}\bar{\rho}(x)\,\varphi(0,x)\,\mathrm{d}x=0.$$

Finally, by Proposition 2.10 we have $\check{\rho}\left(t,F(t,x)\right)=\rho(t,x)$ a.e. on $\mathrm{spt}(\rho)$, and therefore ρ satisfies (1.5). $\diamond$

We are now ready to complete the proof of our main result.

Proof of Theorem 1.3. In view of Theorem 1.2, the entropy inequality (1.7) is sufficient in order to show that ρ is the unique entropy solution in the sense of Definition 1.1.

Let $\varphi\in C_c^\infty(]0,+\infty[\times\mathbb{R};\mathbb{R})$ with $\varphi\geq 0$ and $k\geq 0$ be a constant. We shall prove that the limit ρ satisfies the entropy inequality (1.7). We consider the quantity

$$\begin{aligned}&\int_0^{+\infty}\int_{\mathbb{R}}\Big[|\hat{\rho}^n(t,x)-k|\,\varphi_t(t,x)+\mathrm{sgn}(\hat{\rho}^n(t,x)-k)\,[f(\hat{\rho}^n(t,x))-f(k)]\,\varphi_x(t,x)\Big]\,\mathrm{d}x\,\mathrm{d}t\\&=B_0+B_N+\sum_{i=0}^{N_n-1}I_i,\end{aligned}$$

with

$$\begin{aligned}B_0&:=\int_0^{+\infty}\int_{-\infty}^{x_0^n(t)}\Big[k\,\varphi_t(t,x)+f(k)\,\varphi_x(t,x)\Big]\,\mathrm{d}x\,\mathrm{d}t,\\B_N&:=\int_0^{+\infty}\int_{x_{N_n}^n(t)}^{+\infty}\Big[k\,\varphi_t(t,x)+f(k)\,\varphi_x(t,x)\Big]\,\mathrm{d}x\,\mathrm{d}t,\\I_i&:=\int_0^{+\infty}\int_{x_i^n(t)}^{x_{i+1}^n(t)}\Big[|y_i^n(t)-k|\,\varphi_t(t,x)+\mathrm{sgn}(y_i^n(t,x)-k)\,[f(y_i^n(t,x))-f(k)]\,\varphi_x(t,x)\Big]\,\mathrm{d}x\,\mathrm{d}t.\end{aligned}$$

For simplicity in the notation, from now on we shall drop the n index and the (t,x) dependency, except in cases in which $t=0$. Moreover we define $y_N\equiv 0$. We next observe by (1.14) that

$$\frac{\mathrm{d}}{\mathrm{d}t}\left[\int_{x_i}^{x_{i+1}}\varphi\,\mathrm{d}x\right]=v(y_{i+1})\,\varphi(t,x_{i+1})-v(y_i)\,\varphi(t,x_i)+\int_{x_i}^{x_{i+1}}\varphi_t\,\mathrm{d}x, \tag{2.15}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}\left[\int_{-\infty}^{x_0}\varphi\,\mathrm{d}x\right]=v(y_0)\,\varphi(t,x_0)+\int_{-\infty}^{x_0}\varphi_t\,\mathrm{d}x, \tag{2.16}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}\left[\int_{x_N}^{+\infty}\varphi\,\mathrm{d}x\right]=-v_{\max}\varphi(t,x_N)+\int_{x_N}^{+\infty}\varphi_t\,\mathrm{d}x. \tag{2.17}$$

In view of (2.16) and (2.17), the terms B_0 and B_N can be rewritten as follows

$$B_0=\int_0^{+\infty}k\,[v(k)-v(y_0)]\,\varphi(x_0)\,\mathrm{d}t,\qquad\qquad B_N=\int_0^{+\infty}k\,[v_{\max}-v(k)]\,\varphi(x_N)\,\mathrm{d}t.$$

As for the term I_i, we have for $i=0,\ldots,N-1$

$$\begin{aligned}I_i=&\int_0^{+\infty}|y_i-k|\left\{\frac{\mathrm{d}}{\mathrm{d}t}\left[\int_{x_i}^{x_{i+1}}\varphi\,\mathrm{d}x\right]-v(y_{i+1})\,\varphi(x_{i+1})+v(y_i)\,\varphi(x_i)\right\}\mathrm{d}t\\&+\int_0^{+\infty}\mathrm{sgn}(y_i-k)\,[f(y_i)-f(k)]\,[\varphi(x_{i+1})-\varphi(x_i)]\,\mathrm{d}t.\end{aligned}$$

By (1.16), we compute the term

$$\int_0^{+\infty}|y_i-k|\frac{\mathrm{d}}{\mathrm{d}t}\left[\int_{x_i}^{x_{i+1}}\varphi\,\mathrm{d}x\right]\mathrm{d}t=-\int_0^{+\infty}\left[\int_{x_i}^{x_{i+1}}\varphi\,\mathrm{d}x\right]\frac{\mathrm{d}}{\mathrm{d}t}|y_i-k|\,\mathrm{d}t$$

$$
\begin{aligned}
&= -\int_0^{+\infty} \operatorname{sgn}(y_i - k) \left[-\frac{y_i^2}{\ell} [v(y_{i+1}) - v(y_i)] \right] \left[\int_{x_i}^{x_{i+1}} \varphi \, \mathrm{d}x \right] \mathrm{d}t \\
&= \int_0^{+\infty} \operatorname{sgn}(y_i - k) \, y_i \, [v(y_{i+1}) - v(y_i)] \left[\fint_{x_i}^{x_{i+1}} \varphi \, \mathrm{d}x \right] \mathrm{d}t.
\end{aligned}
$$

Hence, we have

$$
\sum_{i=0}^{N-1} I_{i-1} = \sum_{i=1}^{N} \int_0^{+\infty} J_i \, \mathrm{d}t + \sum_{i=1}^{N} \int_0^{+\infty} K_i \, \varphi(x_i) \, \mathrm{d}t + \int_0^{+\infty} L \, \varphi(t, x_0) \, \mathrm{d}t - \int_0^{+\infty} M \, \varphi(t, x_N) \, \mathrm{d}t,
$$

with

$$
\begin{aligned}
&J_i := \operatorname{sgn}(y_{i-1} - k) \, y_{i-1} \, [v(y_i) - v(y_{i-1})] \left[\fint_{x_{i-1}}^{x_i} \varphi \, \mathrm{d}x - \varphi(x_i) \right], \\
&K_i := \operatorname{sgn}(y_{i-1} - k) \, y_{i-1} \, [v(y_i) - v(y_{i-1})] + |y_i - k| \, v(y_i) - \operatorname{sgn}(y_i - k)[f(y_i) - f(k)], \\
&\qquad - |y_{i-1} - k| \, v(y_i) + \operatorname{sgn}(y_{i-1} - k) \, [f(y_{i-1}) - f(k)], \\
&L := |y_0 - k| \, v(y_0) - \operatorname{sgn}(y_0 - k) \, [f(y_0) - f(k)], \\
&M := k \, v_{\max} - f(k).
\end{aligned}
$$

We observe that

$$
B_N - \int_0^{+\infty} M \varphi(t, x_N) \, \mathrm{d}t = 0.
$$

We now compute L. If $k < y_0$, we have

$$
L = k \, [v(k) - v(y_0)] \geq 0,
$$

as v is non increasing. Therefore, for $k < y_0$

$$
B_0 + \int_0^{+\infty} L \, \varphi(t, x_0) \, \mathrm{d}x = 2 \int_0^{+\infty} k \, [v(k) - v(y_0)] \, \varphi(x_0) \, \mathrm{d}t \geq 0.
$$

Similarly, for $k \geq y_0$ we have

$$
L = k \, [v(y_0) - v(k)] \geq 0,
$$

which gives

$$
B_0 + \int_0^{+\infty} L \varphi(t, x_0) \, \mathrm{d}x = 0.
$$

We now compute the term K_i for $i = 1, \ldots, N$. After some easy manipulations, we get

$$
K_i = k \, [v(k) - v(y_i)] \, \{ \operatorname{sgn}(y_i - k) - \operatorname{sgn}(y_{i-1} - k) \}.
$$

We consider all the possible cases for k. If either $k < \min\{y_i, y_{i-1}\}$, or $k > \max\{y_i, y_{i-1}\}$, then we easily get $K_i = 0$. If $y_i \leq k \leq y_{i-1}$, then $K_i = 2k[v(y_i) - v(k)] \geq 0$ as v is non increasing. Finally, if $y_{i-1} \leq k \leq y_i$, then $K_i = 2k[v(k) - v(y_i)] \geq 0$. In all cases, we get $K_i \geq 0$ for all $i = 1, \ldots, N$. Putting all the terms together, we get

$$
\int_0^{+\infty} \int_{\mathbb{R}} \left[|\hat{\rho} - k| \, \varphi_t + \operatorname{sgn}(\hat{\rho} - k) \, [f(\hat{\rho}) - f(k)] \, \varphi_x \right] \mathrm{d}x \, \mathrm{d}t \geq \sum_{i=1}^{N} \int_0^{+\infty} J_i \, \mathrm{d}t. \tag{2.18}
$$

We now estimate the terms J_i. For some $\delta > 0$, assuming that the support of φ is contained in the strip $t \in [\delta, T]$, we have by Proposition 2.6

$$\begin{aligned}
\left|\sum_{i=1}^{N}\int_{0}^{+\infty} J_i \,\mathrm{d}t\right| &= \left|\sum_{i=1}^{N}\int_{0}^{+\infty} \mathrm{sgn}(y_{i-1}-k)\, y_{i-1}\, [v(y_i)-v(y_{i-1})] \left[\fint_{x_{i-1}}^{x_i} \varphi \,\mathrm{d}x - \varphi(x_i)\right] \mathrm{d}t\right| \\
&\leq \int_{0}^{+\infty}\sum_{i=1}^{N}\left[\frac{y_{i-1}^2}{\ell}\,|v(y_i)-v(y_{i-1})| \int_{x_{i-1}}^{x_i} |\varphi(x)-\varphi(x_i)|\,\mathrm{d}x\right]\mathrm{d}t \\
&\leq \mathrm{Lip}(\varphi)\int_{\delta}^{T} \sup_{i=1,\ldots,N}\left[\frac{y_{i-1}^2\, (x_i - x_{i-1})^2}{\ell}\right]\sum_{i=1}^{N} |v(y_i)-v(y_{i-1})|\,\mathrm{d}t \\
&\leq \ell\,\mathrm{Lip}(\varphi)\, T \sup_{t\geq\delta} \mathrm{TV}\left[v(\hat{\rho}^n(t))\right] \leq \ell\,\mathrm{Lip}(\varphi)\, T\, C_\delta.
\end{aligned}$$

As a consequence

$$\lim_{n\to+\infty}\sum_{i=1}^{N}\int_{0}^{+\infty} J_i \,\mathrm{d}t = 0$$

and letting n go to infinity in (2.18) we obtain the entropy inequality (1.7). $\diamond$

2.5. **Concluding remarks.** We conclude this paper with the some technical remarks which help motivating our choices in the strategy of the proof at several stages in the paper.

- In the case of v such that $v' \leq -c < 0$, then the Oleinik-type estimate (2.12) gives a one sided estimate for $\hat{\rho}^n_x$ in the sense of distributions. Such an estimate can be then passed to the limit very easily, and one obtains an analogous estimate for the limit. In this way, one can check that the limit ρ is an entropy solution in much easier way than the above proof. In the general case of v' possibly degenerating, such a strategy fails. Indeed, surprisingly enough the Oleinik estimate one gets in the limit from (2.12) is *not* equivalent (in general) to the estimate (1.8). For this reason, we preferred getting the entropy condition in the Kružkov sense rather than the one sided Lipschitz condition. This strategy allows in particular to get the entropy condition in the limit also in the case of v not satisfying (V3) and $\bar{\rho}$ satisfying (InBV).
- In the case of linear velocity v, e.g. $v(\rho) = v_{\max}(1-\rho)$, the convergence to a weak solution (1.5) can be obtained without the need of the **BV** estimates, as the velocity term in (2.7) is linear. This is somehow intrinsic in using a Lagrangian description.
- In order to get continuity in time for the sequence $\hat{\rho}^n$, the most natural try would be getting $\mathbf{L}^1$–continuity. Encouraged by the $\mathbf{L}^1$ time equi-continuity of $\check{\rho}^n$, we have attempted at proving such a property in many ways without success. This is the reason why use the generalized Aubin-Lions lemma, which allows to take advantage of the Wasserstein equi-continuity of $\hat{\rho}^n$, and still get the same $\mathbf{L}^1$–compactness in the end. The only drawback of this strategy is that we can't get any $\mathbf{L}^1$ time continuity for the limit.
- As pointed out in the introduction, the proposed Lagrangian approach has the advantage of providing a piecewise constant approximation with a non increasing number of jumps. The price to pay for such a simplification is that we lose the classical shock structure at a microscopic level. Indeed, as pointed out in [15, 43], the explicit solution to the FTL system even for simple Riemann–type initial conditions is not immediate. On the other hand, this aspect gives an added value to our result, as we show that shocks and rarefaction waves are still achieved in the macroscopic limit, despite not being easily detectable at the microscopic level.

Appendix A. Heuristic derivation of the FTL model from the LWR model

In this appendix we formally provide our derivation from the LWR model (1.4) of a discrete approximating model of the form (1.9).

Let ρ be an entropy solution of (1.4) in the sense of Definition 1.1 and $\kappa := R\,\rho$. We assume for simplicity that ρ is compactly supported. The total space occupied by the vehicles present in $]-\infty, x]$ at time $t \geq 0$ is

$$F(t,x) := \int_{-\infty}^{x} \rho(t,y)\,\mathrm{d}y. \tag{A.1}$$

Clearly, F takes values in $[0, L]$, where $L := \|\bar\rho\|_{\mathbf{L}^1(\mathbb{R};[0,1])}$, and for any fixed $t \geq 0$ the map $x \mapsto F(t,x)$ is continuous and non-decreasing, $F(t,-\infty) = 0$ and $F(t,+\infty) = L$. The result in the next proposition shows that (1.4a) is equivalent to requiring that the weak partial derivatives of F with respect to time and space commute in the sense of distributions.

Proposition A.1 ([16])**.** *The partial derivatives of F satisfy in the sense of distributions*

$$F_x = \rho, \qquad\qquad F_t = -f(\rho). \tag{A.2}$$

Proof. The first equality in (A.2) is obvious. For any test function $\psi \in \mathbf{C_c^\infty}(]0,+\infty[\times\mathbb{R};\mathbb{R})$ we have that by (1.5)

$$\begin{aligned}
\int_{\mathbb{R}}\int_{\mathbb{R}_+} F(t,x)\,\partial_t\psi_x(t,x)\,\mathrm{d}t\,\mathrm{d}x &= \int_{\mathbb{R}}\int_{\mathbb{R}_+} F(t,x)\,\partial_x\psi_t(t,x)\,\mathrm{d}t\,\mathrm{d}x\\
= -\int_{\mathbb{R}}\int_{\mathbb{R}_+} \rho(t,x)\,\psi_t(t,x)\,\mathrm{d}t\,\mathrm{d}x &= \int_{\mathbb{R}}\int_{\mathbb{R}_+} f(\rho(t,x))\,\psi_x(t,x)\,\mathrm{d}t\,\mathrm{d}x.
\end{aligned}$$

This shows that for any $t \geq 0$, the map $x \mapsto [F_t(t,x) + f(\rho(t,x))]$ is constant (as a distribution). Therefore there exists $c \in \mathbf{L^1_{loc}}([0,+\infty[;\mathbb{R})$ such that

$$\int_{\mathbb{R}}\int_0^{+\infty}\Big[F(t,x)\,\varphi_t(t,x) - f(\rho(t,x))\,\varphi(t,x) + c(t)\,\varphi(t,x)\Big]\,\mathrm{d}t\,\mathrm{d}x = 0.$$

Choose now, for any integer $k \in \mathbb{N}$,

$$\varphi(t,x) = \eta(t)\,\psi(x-k),$$

where $\eta \in \mathbf{C_c^\infty}(]0,+\infty[;\mathbb{R})$ and $\psi \in \mathbf{C_c^\infty}(\mathbb{R};[0,+\infty[)$ such that $\|\psi\|_{\mathbf{L}^1(\mathbb{R};\mathbb{R})} = 1$. We get

$$\begin{aligned}
0 &= \int_{\mathbb{R}}\int_0^{+\infty}\Big[F(t,x)\,\dot\eta(t) - f(\rho(t,x))\,\eta(t) + c(t)\,\eta(t)\Big]\,\psi(x-k)\,\mathrm{d}t\,\mathrm{d}x\\
&= \int_{\mathbb{R}}\int_0^{+\infty}\Big[F(t,x+k)\,\dot\eta(t) - f(\rho(t,x+k))\,\eta(t) + c(t)\,\eta(t)\Big]\,\psi(x)\,\mathrm{d}t\,\mathrm{d}x.
\end{aligned}$$

By Lebesgue dominated convergence theorem, we can send k to $+\infty$ and get

$$0 = \int_{\mathbb{R}}\int_0^{+\infty}\Big[L\,\dot\eta(t) + c(t)\,\eta(t)\Big]\,\psi(x)\,\mathrm{d}t\,\mathrm{d}x = \int_0^{+\infty} c(t)\,\eta(t)\,\mathrm{d}t,$$

and the above expression on the right hand side can be easily made nonzero by suitably choosing η, unless $c(t) = 0$ for a.e. $t \geq 0$, which proves the assertion. ◇

For any $t \geq 0$ the map $x \mapsto F(t,x)$ is strictly increasing on the intervals where the density $x \mapsto \rho(t,x)$ is not zero and otherwise it is constant. Therefore we can introduce $X := \mathcal{X}[F]$, the pseudo-inverse of F. Now, assume for simplicity that $\rho(t,x) > 0$ for all $(t,x) \in \mathrm{spt}(\rho) = \{(t,x) \in \mathbb{R}_+ \times \mathbb{R} : a(t) \leq x \leq b(t)\}$. Then, for any $t \geq 0$ by Proposition A.1 we have that $x \mapsto F(t,x)$ is strictly increasing on $\mathrm{spt}(\rho(t))$. This implies that

$\mathcal{X}[F]$ is the inverse of F on the support of ρ, namely $F(t, X(t,z)) = z$ on $(t,z) \in \mathbb{R}_+ \times [0,L]$, and, assuming that all the derivatives below are well defined, we have that

$$F_x(t, X(t,z)) = \rho(t, X(t,z)) > 0 \qquad\qquad \text{for a.e. } (t,z) \in \mathbb{R}_+ \times [0,L].$$

Therefore,

$$\begin{aligned} 1 &= \frac{\mathrm{d}}{\mathrm{d}z} F(t, X(t,z)) = F_x(t, X(t,z))\, X_z(t,z), \\ 0 &= \frac{\mathrm{d}}{\mathrm{d}t} F(t, X(t,z)) = F_t(t, X(t,z)) + F_x(t, X(t,z))\, X_t(t,z), \end{aligned}$$

which yields, once again by Proposition A.1, that $X(t,z)$ is indeed a solution of the PDE

$$X_t(t,z) = v\left(\frac{1}{X_z(t,z)}\right). \tag{A.3}$$

The initial condition $X(0,z)$ is determined by

$$\int_{-\infty}^{X(0,z)} \bar{\rho}(y)\,\mathrm{d}y = z.$$

The computation above is only rigorous on the sets in which $\rho(t,x) > 0$.

The last step needed in order to (formally) recognize the discrete model (1.9) in (A.3) is by replacing the z–derivative of X in (A.3) by the (forward) finite differences

$$X_z \approx \frac{X(t, z+\ell) - X(t,z)}{\ell}, \tag{A.4}$$

which gives

$$X_t(t,z) \approx v\left(\frac{\ell}{X(t,z+\ell) - X(t,z)}\right).$$

Then the desired model (1.9) is obtained by assuming that $X(t)$ is piecewise constant on intervals of measure ℓ, with $X(t, j\,\ell) = x_j(t)$, $j = 1, \ldots, N-1$. For any fixed $z \in \{i\,\ell :\ i = 0, 1, \ldots, N\}$, the map $t \mapsto X(t,z)$ can be ideally interpreted as the path described by the 'infinitesimal vehicle' labelled with $z \in [0,L]$. Therefore, (A.3) can be interpreted as the expression in the Lagrangian coordinates (t,z) of the Cauchy problem (1.4).

Acknowledgments

The authors acknowledge useful discussions with R. M. Colombo, P. Degond, and P. Marcati. MDF is supported by the Marie Curie CIG (Career Integration Grant) DifNonLoc - Diffusive Partial Differential Equations with Nonlocal Interaction in Biology and Social Sciences and by the Ministerio de Ciencia e Innovación, grant MTM2011-27739-C04-02. MDR was partially supported by ICM, University of Warsaw, and Narodowe Centrum Nauki, grant 4140.

References

[1] L. Ambrosio, N. Gigli, and G. Savaré. *Gradient flows in metric spaces and in the space of probability measures.* Lectures in Mathematics ETH Zürich. Birkhäuser Verlag, Basel, 2005. 1

[2] J.-P. Aubin. Macroscopic traffic models: Shifting from densities to 'celerities'. *Applied Mathematics and Computation*, 217(3):963 – 971, 2010. 1

[3] A. Aw, A. Klar, T. Materne, and M. Rascle. Derivation of continuum traffic flow models from microscopic follow-the-leader models. *SIAM Journal on Applied Mathematics*, 63(1):259–278, 2002. 1, 1

[4] A. Aw and M. Rascle. Resurrection of 'second order' models of traffic flow. *SIAM Journal on Applied Mathematics*, 60(3):916–938, 2000. 1

[5] N. Bellomo and C. Dogbe. On the modeling of traffic and crowds: A survey of models, speculations, and perspectives. *SIAM Review*, 53(3):409–463, 2011. 1

[6] F. Berthelin, P. Degond, M. Delitala, and M. Rascle. A model for the formation and evolution of traffic jams. *Arch. Ration. Mech. Anal.*, 187(2):185–220, 2008. 1, 1

[7] F. Bolley, Y. Brenier, and G. Loeper. Contractive metrics for scalar conservation laws. *J. Hyperbolic Differ. Equ.*, 2(1):91–107, 2005. 1

[8] Y. Brenier and E. Grenier. Sticky particles and scalar conservation laws. *SIAM J. Numer. Anal.*, 35(6):2317–2328 (electronic), 1998. 1

[9] Y. Brenier and S. Osher. The discrete one-sided Lipschitz condition for convex scalar conservation laws. *SIAM J. Numer. Anal.*, 25(1):8–23, 1988. 1

[10] A. Bressan. Global solutions of systems of conservation laws by wave-front tracking. *J. Math. Anal. Appl.*, 170(2):414–432, 1992. 1

[11] A. Bressan. *Hyperbolic systems of conservation laws*, volume 20 of *Oxford Lecture Series in Mathematics and its Applications*. Oxford University Press, Oxford, 2000. 1, 1, 1.1, 1.1, 2.1, 2.4

[12] J. A. Carrillo, M. Di Francesco, and C. Lattanzio. Contractivity of Wasserstein metrics and asymptotic profiles for scalar conservation laws. *J. Differential Equations*, 231(2):425–458, 2006. 1

[13] G.-Q. Chen and M. Rascle. Initial layers and uniqueness of weak entropy solutions to hyperbolic conservation laws. *Arch. Ration. Mech. Anal.*, 153(3):205–220, 2000. 1.1, 1.2

[14] R. M. Colombo and A. Marson. A Hölder continuous ODE related to traffic flow. *Proc. Roy. Soc. Edinburgh Sect. A*, 133(4):759–772, 2003. 1

[15] R. M. Colombo and E. Rossi. On the micro-macro limit in traffic flow. *To appear on Rendiconti dell'Università di Padova*, 2013. 1, 1, 2.5

[16] R. Courant and K. O. Friedrichs. *Supersonic Flow and Shock Waves*. Applied Mathematical Sciences. Springer-Verlag, 1976. A.1

[17] C. M. Dafermos. Polygonal approximations of solutions of the initial value problem for a conservation law. *J. Math. Anal. Appl.*, 38:33–41, 1972. 1, 1.1

[18] C. M. Dafermos. *Hyperbolic conservation laws in continuum physics*, volume 325 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, 2000. 1, 1.1, 1.1

[19] C. F. Daganzo. A variational formulation of kinematic waves: basic theory and complex boundary conditions. *Transportation Research Part B: Methodological*, 39(2):187–196, 2005. 1

[20] M. Di Francesco and D. Matthes. Curves of steepest descent are entropy solutions for a class of degenerate convection-diffusion equations. *Calc. Var. PDEs*, DOI 10.1007/s00526-013-0633-5, 2013. 2.3

[21] R. J. DiPerna. Global existence of solutions to nonlinear hyperbolic systems of conservation laws. *J. Differential Equations*, 20(1):187–212, 1976. 1

[22] R. L. Dobrušin. Vlasov equations. *Funktsional. Anal. i Prilozhen.*, 13(2):48–58, 96, 1979. 1

[23] Lawrence Craig Evans and Ronald F Gariepy. *Measure theory and fine properties of functions*, volume 5. CRC press, 1991. 2.3

[24] J. Glimm. Solutions in the large for nonlinear hyperbolic systems of equations. *Comm. Pure Appl. Math.*, 18:697–715, 1965. 1

[25] F. Golse and B. Perthame. Optimal regularizing effect for scalar conservation laws. *Rev. Mat. Iberoam.*, 29(4):1477–1504, 2013. 1

[26] J. B. Goodman and R. J. LeVeque. A geometric approach to high resolution TVD schemes. *SIAM J. Numer. Anal.*, 25(2):268–284, 1988. 1

[27] H. Greenberg. An analysis of traffic flow. *Operations Research*, 7(1):79–85, 1959. 1.1

[28] B. Greenshields. A study of traffic capacity. *Proceedings of the Highway Research Board*, 14:448–477, 1935. 1.1

[29] D. Hoff. The Sharp Form of Oleinik's Entropy Condition in Several Space Variables. *Transactions of the American Mathematical Society*, 276(2):707–714, 1983. 1, 1.1

[30] S. P. Hoogendoorn and P. H. L. Bovy. State-of-the-art of vehicular traffic flow modelling. In *Delft University of Technology, Delft, The*, pages 283–303, 2001. 1

[31] S. N. Kružkov. First order quasilinear equations with several independent variables. *Mat. Sb. (N.S.)*, 81 (123):228–255, 1970. 1, 1.1, 1.1

[32] P. D. Lax. *Hyperbolic systems of conservation laws and the mathematical theory of shock waves*. Society for Industrial and Applied Mathematics, Philadelphia, Pa., 1973. Conference Board of the Mathematical Sciences Regional Conference Series in Applied Mathematics, No. 11. 1.1

[33] M. J. Lighthill and G. B. Whitham. On kinematic waves. II. A theory of traffic flow on long crowded roads. *Proc. Roy. Soc. London. Ser. A.*, 229:317–345, 1955. 1, 1, 1.1

[34] R. C. MacCamy and E. Socolovsky. A numerical procedure for the porous media equation. *Comput. Math. Appl.*, 11(1-3):315–319, 1985. Hyperbolic partial differential equations, II. 1

[35] C. B. Morrey, Jr. On the derivation of the equations of hydrodynamics from statistical mechanics. *Comm. Pure Appl. Math.*, 8:279–326, 1955. 1

[36] G. F. Newell. A simplified theory of kinematic waves in highway traffic. *Transportation Research Part B: Methodological*, 27(4):281–313, 1993. 1

[37] O. A. Oleinik. Discontinuous solutions of non-linear differential equations. *Amer. Math. Soc. Transl. (2)*, 26:95–172, 1963. 1, 1.1

[38] L. Onsager. Crystal statistics. I. A two-dimensional model with an order-disorder transition. *Phys. Rev. (2)*, 65:117–149, 1944. 1

[39] B. Piccoli and A. Tosin. Vehicular traffic: A review of continuum mathematical models. In Robert A. Meyers, editor, *Encyclopedia of Complexity and Systems Science*. Springer New York, 2009. 1

[40] L. A. Pipes. Car following models and the fundamental diagram of road traffic. *Transp. Res.*, 1:21–29, 1967. 1.1

[41] P. I. Richards. Shock waves on the highway. *Operations Res.*, 4:42–51, 1956. 1, 1, 1.1

[42] M. D. Rosini. *Macroscopic models for vehicular flows and crowd dynamics: theory and applications*. Understanding Complex Systems. Springer, Heidelberg, 2013. 1

[43] E. Rossi. On the micro–macro limit in traffic flow. Master's thesis, Università Cattolica del Sacro Cuore, Brescia, 2013. 1, 1.1, 2.5

[44] R. Rossi and G. Savaré. Tightness, integral equicontinuity and compactness for evolution problems in Banach spaces. *Ann. Sc. Norm. Super. Pisa Cl. Sci. (5)*, 2(2):395–431, 2003. 2.3, 2.1

[45] G. Russo. Deterministic diffusion of particles. *Comm. Pure Appl. Math.*, 43(6):697–733, 1990. 1

[46] D. Serre. *Systems of conservation laws. 1 & 2*. Cambridge University Press, Cambridge, 1999. 1, 1.1

[47] E. Tadmor. The large-time behavior of the scalar, genuinely nonlinear Lax-Friedrichs scheme. *Math. Comp.*, 43(168):353–368, 1984. 1

[48] R. T. Underwood. Speed, volume, and density relationship. In *Quality and theory of traffic flow: a symposium*, pages 141–188. Greenshields, B.D. and Bureau of Highway Traffic, Yale University, 1961. 1.1

[49] C. Villani. *Topics in optimal transportation*, volume 58 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2003. 1, 1.3, 1.3

[50] H. M. Zhang. A non-equilibrium traffic model devoid of gas-like behavior. *Transportation Research Part B: Methodological*, 36(3):275–290, 2002. 1

Marco Di Francesco - Mathematical Sciences, University of Bath, Claverton Down, Bath (UK), BA2 7AY.
E-mail address: mdifrance@gmail.com

Massimiliano D. Rosini - ICM, University of Warsaw, ul. Prosta 69, 00-838 Warsaw, Poland
E-mail address: mrosini@icm.edu.pl