Convergence rates of the front tracking method for conservation laws in the Wasserstein distances

Susanne Solem

Abstract

We prove that front tracking approximations to scalar conservation laws with convex fluxes converge at a rate of Δx^2 in the 1-Wasserstein distance W_1. Assuming positive initial data, we also show that the approximations converge at a rate of Δx in the ∞-Wasserstein distance W_∞. Moreover, from a simple interpolation inequality between W_1 and W_∞ we obtain convergence rates in all the p-Wasserstein distances: $\Delta x^{1+1/p}$, $p \in [1, \infty]$.

1 Introduction

In this paper we will consider front tracking approximations to the scalar conservation law

$$\begin{aligned} &u_t + f(u)_x = 0, \qquad x \in \mathbb{R},\ t > 0, \\ &u(x,0) = u_0(x), \end{aligned} \tag{1.1}$$

where f is convex, u_0 is of compact support and Lip^+ bounded. A function v is said to be Lip^+ bounded if

$$\frac{v(x+z) - v(x)}{z} \leqslant C, \qquad \forall\ x, z \in \mathbb{R}, \quad z \neq 0, \tag{1.2}$$

for some constant C. Under these assumptions on f and u_0, it is well-established that (1.1) admits an unique entropy solution u and that $u(t)$ satisfies (1.2) for any $t > 0$, see for example [20, 21, 29, 15]. The aim of this paper is to prove the following theorem.

Main Theorem. *Let f be convex and let u_0 be of compact support. Then the front tracking approximations of* (1.1) *converge to the unique entropy solution of* (1.1) *at a rate of Δx^2 in the 1-Wasserstein distance. If in addition $u_0 \geqslant 0$, the approximations converge at a rate of Δx in the ∞-Wasserstein distance, and hence, they converge at a rate of $\Delta x^{1+1/p}$ in the p-Wasserstein distance W_p for any $p \in [1, \infty]$.*

See Section 2 for a complete statement of the theorem. This result demonstrates that front tracking approximations (for the class of initial data considered here) converge at a higher rate in every Wasserstein distance than the optimal rate of Δx in the usual metric L^1. Thus it supports the argument that the Wasserstein distances are well-suited to measure the approximation error of solutions to (1.1), which we started in [7] (for W_1) and continue below. Furthermore, these convergence results gives the front tracking method an advantage, in terms of guaranteed convergence rate, over formally higher-order finite volume approximations to (1.1) (for which no convergence rate estimate exists for general initial data in either L^1 or W_p).

In order to prove the main theorem, we will need (and establish) some properties of solutions to(1.1). Among other things, we will prove stability estimates in W_1 and W_∞, show that the support of the solution $u(t)$ is connected if the support of u_0 is, and revive a result by Oleĭnik [20].

1.1 The Wasserstein distances

The p-Wasserstein distance (or W_p-distance), also called the p-Monge–Kantorovich distance, is a metric on the set of probability measures with finite pth order moment, and for two probability measures μ and ν on $\mathbb{R}^d$ it takes the form

$$W_p(\mu, \nu) = \left(\inf_{\pi \in \Pi(\mu,\nu)} \int_{\mathbb{R}^{2d}} |x-y|^p d\pi(x,y) \right)^{1/p}, \quad p \in [1, \infty), \tag{1.3}$$

where the infimum is taken over all measures π on $\mathbb{R}^{2d}$ with marginals μ and ν. See [28] for further details. The W_∞-distance,

$$W_\infty(\mu,\nu) = \lim_{p\to\infty} W_p(\mu,\nu), \tag{1.4}$$

is a metric on the space of probability measures with bounded support. Although normally only defined for probability measures, the p-Wasserstein distance between two Borel measurable functions $u, v \geqslant 0$, each of the same finite mass and with finite pth order moment for $1 \leqslant p < \infty$,

$$\int_{\mathbb{R}^d} (u-v)(x)dx = 0, \quad \int_{\mathbb{R}^d} |x|^p u(x)dx < \infty, \quad \int_{\mathbb{R}^d} |x|^p v(x)dx < \infty, \tag{1.5}$$

and of compact support for $p = \infty$, $W_p(u,v) := W_p(u\mathcal{L}, v\mathcal{L})$ is well-defined. Here $\mathcal{L}$ denotes the Lebesgue measure.

All the W_p-distances are suited to measure the difference between (approximate) solutions to (1.1). If u_0, v_0 initially fulfil the conditions (1.5), then the two solutions $u(t), v(t)$ of (1.1) (possibly with different flux functions f, g for u, v respectively) will satisfy (1.5) at any later time t due to conservation of mass and finite speed of propagation. Hence, $W_p\big(u(t), v(t)\big)$ will be well-defined and finite as long as $W_p(u_0, v_0)$ is. To some extent, one can argue that the Wasserstein metrics are natural distances associated to (1.1). Indeed, heuristically the W_p metrics measure the minimum "cost" of transporting mass from one measure to another, and transporting quantities (of "mass") is exactly what (1.1) does.

The 1-Wasserstein distance seems to be particularly suitable in the context of (1.1). To see why, consider the shock and its approximation (stipled) in Figure 1(a). The L^1-distance, which is commonly used to measure approximation errors of (1.1), measures the area (in grey) between the two solutions. The height is $O(1)$ and the width $O(\Delta x)$. Hence, the L^1-error between the two solutions is $O(1)\cdot O(\Delta x) = O(\Delta x)$. The W_1-distance can be thought of as measuring the minimal amount of work (mass $\times$ distance) required to move mass from one measure to another. In Figure 1(a) this means that W_1 measures the work needed to move the surplus of mass to the right of the shock (light grey) to the shortage of mass to the left of the shock (dark grey). The mass (area) to be moved is $O(\Delta x)$, and it needs to be moved a distance $O(\Delta x)$. It follows that the W_1-error is $O(\Delta x)\cdot O(\Delta x) = O(\Delta x^2)$. The difference in the convergence rate between L^1 and W_1 for shock solutions has already been observed in the case of monotone finite volume scheme approximations. Teng and Zhang [27] obtained a convergence rate of $O(\Delta x)$ in L^1 for solutions consisting of a finite number of decreasing shocks, whereas the rate was improved to $O(\Delta x^2)$ in W_1 in [7].

We apply the same reasoning to the W_p-distance by replacing the distance function $|\cdot|$ with $|\cdot|^p$ and taking the pth root to find that the W_p-approximation error in Figure 1(a) is $(O(\Delta x)\cdot O(\Delta x^p))^{1/p} = O(\Delta x^{1+1/p})$.

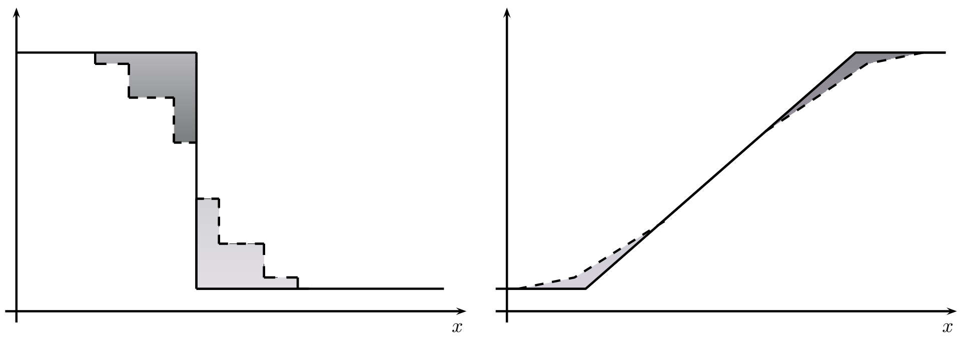
It is not given that there is always a gain in the convergence rate by utilizing one of the Wasserstein distances instead of the L^1-distance. Figure 1(b) depicts one such counterexample. Let the L^1-error between the solution and its approximation (stipled) be $O(\Delta x)$. If the distance between the surplus of mass (light grey) and the shortage of mass (dark grey) is $O(1)$, the error will be $(O(\Delta x)\cdot O(1))^{1/p} = O(\Delta x^{1/p})$ in W_p. Therefore, to obtain a higher rate in the W_p-distances, the approximation of the initial data can only redistribute small amounts of mass over small intervals. Furthermore, this redistribution of mass between the approximate and exact solution has to be (close to) preserved at any later time. In this paper we will see that this is the case for the front tracking approximation (which is a first order approximation in L^1) and, as a consequence, obtain the $O(\Delta x^{1+1/p})$-rate in W_p.

Lastly, Carrillo et al. [2] have shown that the W_∞-distance is contractive with respect to initial data for solutions of (1.1) — a property that will be exploited in this paper.

1.2 Front tracking, finite volume methods and convergence rates

The front tracking method was first proposed by Dafermos [4]. Later, Holden et al. [8] rediscovered it, extended it to non-convex fluxes and showed that it is a viable numerical method in one dimension. The main strength of the one-dimensional front tracking method is that the approximation is itself an entropy solution to a conservation law. We will make use of this strength in this paper by first proving general stability results of (1.1) in both W_1 and W_∞ and then applying them to the front tracking method in order to obtain the respective Δx^2 and Δx rates.

Up to this point the W_1-distance is the only one among the Wasserstein distances that has been applied in order to study convergence rates of approximations to (1.1). Tadmor et al. [24, 18, 19]



(a) Exact and approximate shock solution of (1.1). (b) Exact and bad (in a W_1 sense) approximate solution of (1.1).

Figure 1: The W_1-distance measures the amount of work required to move mass from one place (dark grey) to another (light grey).

extensively examined it in the context of conservation laws, but under the different name of the *Lip'-norm*. They showed (among other things) that a large class of monotone finite difference methods converge at a rate of Δx in the Lip'-norm for initial data u_0 of compact support satisfying (1.2). By applying their technique to the front tracking approximation, one obtains the rate Δx in W_1.

Due to the structure of the front tracking method, the (provable) convergence rate of this approximation is usually higher than the one for monotone methods. This can be observed in the above for the rates in W_1 (Δx^2 for front tracking and Δx for monotone schemes), and can also be noticed for the rates in L^1. By applying a well-known stability result in the L^1 norm, first proved by Lucier [17], one attains the (optimal) convergence rate Δx in L^1 of the front tracking approximation. However, the most generic result on convergence rates of monotone methods is the $O(\Delta x^{1/2})$ rate in L^1 due to Kuznetsov [16]. A counterexample due to Şabac shows that the $\Delta x^{1/2}$ rate for monotone methods is sharp and cannot be improved without further assumptions on the initial data [23]. But, even though it is not proved, it should be noted that numerical evidence indicates that the convergence rate is close to Δx for monotone schemes as well in the case of more "natural" initial data. The rate of $O(\Delta x)$ in L^1 for a finite number of travelling shocks in [27] endorses these observations.

The first proof of a second-order convergence rate of any numerical method to (1.1) was provided in L^1 by Lucier for a specific piecewise linear extension of the front tracking method [17]. In this paper we prove the same rate in W_1 without modifying the original method.

In [14] Karlsen and Risebro demonstrate the equivalence between entropy solutions of conservation laws and viscosity solutions of the Hamilton–Jacobi equations by utilizing the front tracking method. As a by-product they discover the rate Δx^2 in the L^∞ distance between the primitives of the front tracking approximation and the entropy solution. This result is closely related to the rate of Δx^2 in W_1 that we obtain in this paper, see Remark 3.7. Hong [11] proved a stability result in L^∞ for the Hamilton–Jacobi equations, from which one can also deduce a Δx^2 rate.

Apart from the second-order rate results for front tracking type methods in [17, 14], the only other proof of a second-order rate in any norm of any numerical method for (1.1) is, to the authors knowledge, the Δx^2 rate in the 1-Wasserstein distance in [7].

Except for a piecewise constant projection of the initial data, the one-dimensional front tracking method is grid independent. A simple way to extend the method to the multi-dimensional case, is by dimensional splitting, see [10]. The multi-dimensional extension is no longer the exact solution of a conservation law, and the accuracy of the method depends on the temporal grid size Δt. Thus, the method is no longer grid independent, and the convergence rate of the method might decrease. The two-dimensional method is proven to converge at a rate of $O(\Delta t^{1/2} + \Delta x^{1/2})$ in L^1 [13, 26], which is the same rate as the one Kuznetsov proved for multi-dimensional monotone schemes. Whether the rate improves in W_1, is not known. All studies of convergence rates of approximations to (1.1) in W_1, including the one in this paper, rely on a one dimensional interpretation of W_1 which does not extend to multiple dimensions.

The W_1 rate in this paper shows that the front tracking approximation to (1.1) can be considered a second-order method when applying a suitable metric (although this might be restricted to one dimension). Indeed, one can observe numerically that the front tracking approximation converges at the same rate Δx^2 as a second-order finite volume method in W_1. Furthermore, the Δx rate in W_∞ conveys that displaced mass in the front tracking approximation compared to the exact solution of (1.1) is moved at most a distance Δx.

Next follows an outline of this paper. In Section 2 we provide a short basis for the upcoming results and associated proofs before stating the main theorem. Section 3 contains stability estimates in the W_1-distance, which provide the convergence rate in W_1. Section 4 is devoted to the proof of the rate in W_∞. Lastly, Section 5 contains remarks on possible extensions of the main theorem.

2 Front tracking, Wasserstein metrics and main theorem

2.1 The front tracking method

These are the main ingredients in the front tracking method. Approximate the initial data u_0 in (1.1) by a piecewise constant function $u_0^{\Delta x}$ and the flux f by a piecewise linear function f^δ. Then solve the resulting conservation law

$$u_t + f^\delta(u)_x = 0, \qquad u(x,0) = u_0^{\Delta x}(x), \tag{2.1}$$

exactly. As $u_0^{\Delta x}$ is piecewise constant, the initial problem will be to solve a series of independent Riemann problems, each of them having a wave-front traveling with constant speed, due to f being piecewise linear, as a solution. Whenever two fronts meet, we restart the procedure by solving (2.1) with initial data $u(x,0) = u^{\delta,\Delta x}(x,t^c)$, where t^c is a interaction time. In this way we can find $u^{\delta,\Delta x}(x,t)$ for all times. The resulting solution $u^{\delta,\Delta x}$ is the unique entropy solution to (2.1).

As the Wasserstein distances require that the functions to be compared have equal mass, we approximate the initial data as

$$u_0^{\Delta x}(x) = u_i := \frac{1}{\Delta x}\int_{\mathcal{C}_i} u_0(y)dy, \qquad x \in \mathcal{C}_i := [x_{i-1/2}, x_{i+1/2}), \quad i \in \mathbb{Z}, \tag{2.2}$$

where $x_{i+1/2} = (i + 1/2)\Delta x$, $\Delta x > 0$, to ensure that $\int_{\mathbb{R}}(u_0 - u_0^{\Delta x})\,dx = 0$. (In general one can use piecewise constant approximations that are not necessarily tied to the grid on $\mathbb{R}$ or preserves the mass.) The front tracking flux f^δ is a piecewise linear approximation to f of the following form

$$f^\delta(u) = f(j\delta) + (u - j\delta)\frac{f\big((j+1)\delta\big) - f(j\delta)}{\delta}, \qquad u \in (j\delta, (j+1)\delta], \tag{2.3}$$

for $j \in \mathbb{Z} \cap [-(M+1)/\delta, M/\delta]$, where $M = \|u_0\|_{L^\infty(\mathbb{R})}$ and $\delta = O(\Delta x)$. See [4, 8, 9, 17] for more details on the method.

2.2 The Wasserstein distances in one dimension

Without loss of generality, let $\int_{\mathbb{R}} |u(x)|\,dx = 1$ in (1.1) from this point on. We define the two spaces

$$\mathcal{B}_p := \left\{ u \in L^\infty(\mathbb{R}) : u \geqslant 0, \int_{\mathbb{R}} u(x)\,dx = 1, \int_{\mathbb{R}} |x|^p u(x) dx < \infty \right\}, \quad p \in [1,\infty),$$
$$\mathcal{B} := \left\{ u \in L^\infty(\mathbb{R}) : u \geqslant 0,\ \mathrm{supp}(u) \text{ compact}, \int_{\mathbb{R}} u(x)\,dx = 1 \right\},$$

for ease of notation.

In one dimension, the p-Wasserstein distance (1.3) between u and v both in $\mathcal{B}_p$, has a simple interpretation as the L^p-distance between the pseudo-inverses of the distribution functions [3, 28],

$$U(x) = \int_{-\infty}^{x} u(y)\,dy, \quad V(x) = \int_{-\infty}^{x} v(y)\,dy. \tag{2.4}$$

The pseudo-inverses $U^{-1} : [0,1] \to \mathbb{R}$, $V^{-1} : [0,1] \to \mathbb{R}$ are defined as

$$U^{-1}(\xi) = \inf\{x :\ U(x) > \xi\}, \quad V^{-1}(x) = \inf\{x :\ V(x) > \xi\}.$$

Then the p-Wasserstein distance, $p \in [1,\infty)$ is

$$W_p(u,v) = \|U^{-1} - V^{-1}\|_{L_p([0,1])}. \tag{2.5}$$

When $u, v \in \mathcal{B}$, we can interpret the W_∞-distance in the same way using (1.4),

$$W_\infty(u,v) = \lim_{p\to\infty} \|U^{-1} - V^{-1}\|_{L_p([0,1])} = \|U^{-1} - V^{-1}\|_{L_\infty([0,1])}. \tag{2.6}$$

In particular, the W_1-distance takes the very simple form

$$W_1(u,v) = \int_{\mathbb{R}} |U - V|\, dx, \tag{2.7}$$

which can be found by using Fubini's theorem with (2.5). Notice that for the alternative form (2.7) of W_1 to be well-defined, u and v only need to satisfy

$$\int_{\mathbb{R}} (u-v)(x)dx = 0, \qquad \int_{\mathbb{R}} |x||u-v|(x)dx < \infty. \tag{2.8}$$

2.3 Connection to the Hamilton–Jacobi equation

There is a well-known equivalence between the viscosity solution of the Hamilton–Jacobi equation

$$U_t + f(U_x) = 0, \qquad \int^x u_0\; dx := U_0 \in BUC(\mathbb{R}) \tag{2.9}$$

and the entropy solution of (1.1) with $u_0 \in BV(\mathbb{R})$ through the relation

$$\partial_x U = u, \qquad U = \int^x u\; dx, \tag{2.10}$$

see [14] and references therein. If U_0 is Lipschitz continuous, bounded and f is convex and superlinear,

$$\lim_{|u|\to\infty} \frac{f(u)}{|u|} = \infty,$$

then the unique viscosity solutions of (2.9) can be found by the Hopf–Lax formula

$$U(x,t) = \min_{y\in\mathbb{R}} \left\{ tf^*\left(\frac{x-y}{t}\right) + U_0(y)\right\}, \tag{2.11}$$

where f^* is the Legendre transform of f,

$$f^*(p) = \sup_{u\in\mathbb{R}} \{pu - f(u)\} \tag{2.12}$$

(see for example Evans [6, Ch. 3.3, Ch. 10.3.4]). Then according to (2.10), we get the entropy solution of (1.1) by differentiating (2.11) with respect to x.

2.4 Main theorem

The main theorem relies on the following interpolation result.

Lemma 2.1. *If $u, v \in \mathcal{B}$, then*

$$W_p(u,v) \leqslant W_1(u,v)^{1/p} W_\infty(u,v)^{1-1/p}, \tag{2.13}$$

for $1 < p < \infty$.

Proof. By Hölder's inequality and the representation (2.5),

$$\begin{aligned} W_p(u,v)^p = \||F^{-1} - G^{-1}|^p\|_{L^1([0,1])} &\leqslant \|(F^{-1} - G^{-1})\|_{L^1([0,1])} \||F^{-1} - G^{-1}|^{p-1}\|_{L^\infty([0,1])} \\ &= \|F^{-1} - G^{-1}\|_{L^1([0,1])} \|F^{-1} - G^{-1}\|^{p-1}_{L^\infty([0,1])} \\ &= W_1(u,v) W_\infty(u,v)^{p-1}. \end{aligned}$$

By taking the pth root, we get the interpolation inequality (2.13). ☐

Theorem 2.2. *Let u be the entropy solution of* (1.1) *where f is twice continuously differentiable and convex, and $u_0 \in BV(\mathbb{R})$ be* Lip^+ *bounded and of compact support. Then the front tracking approximation $u^{\delta,\Delta x}$ of u satisfies*

$$W_1\big(u(t), u^{\delta,\Delta x}(t)\big) \leqslant C\Delta x^2$$

for all $t \in [0,T)$ for any $T > 0$. If in addition $u_0 \in \mathcal{B}$, then

$$W_p\big(u(t), u^{\delta,\Delta x}(t)\big) \leqslant C\Delta x^{1+1/p}, \quad p \in [1,\infty],$$

where $C := C_{T,f,\mathrm{supp}(u_0)}$.

Proof. This follows directly from the second order rate in W_1 and the first order rate in W_∞ to be proved in Theorem 3.6 and Theorem 4.6 respectively, and from the interpolation inequality (2.13). □

3 The convergence rate in W_1

We begin by providing two stability estimates in the 1-Wasserstein distance that will yield the second-order convergence rate.

Proposition 3.1. *Assume that f is continuously differentiable such that f' is locally Lipschitz. Let f and g both be convex, and let $u_0, v_0 \in BV(\mathbb{R})$ satisfy* (2.8) *and* (1.2)*. Then the entropy solutions u and v of*

$$\begin{aligned} u_t + f(u)_x &= 0, \qquad u(x,0) = u_0(x), \\ v_t + g(v)_x &= 0, \qquad v(x,0) = v_0(x), \end{aligned} \tag{3.1}$$

satisfy

$$W_1\big(u(t), v(t)\big) \leqslant C(t)\left[W_1\big(u_0, v_0\big) + \int_0^t \|(f-g)(v(s))\|_{L^1(\mathbb{R})} ds\right]. \tag{3.2}$$

Proof. As previously mentioned, since $u_0 - v_0$ satisfies (2.8), $(u-v)(t)$ will also fulfil the same conditions by conservation of mass and finite speed of propagation. Hence $W_1(u(t), v(t))$ is well-defined and finite.

We start by differentiating (2.7) with respect to t and using (2.4),

$$\begin{aligned} \frac{d}{dt}W_1(u,v) &= \int_{\mathbb{R}} \partial_t \left|(U-V)(x)\right| dx \\ &= \int_{\mathbb{R}} \mathrm{sgn}\,(U-V)\,(x)\partial_t(U-V)dx = -\int_{\mathbb{R}} \mathrm{sgn}\,(U-V)\,(x)\left(f(\partial_x U) - g(\partial_x V)\right)dx \\ &= -\int_{\mathbb{R}} a(u,v)\partial_x \left|U-V\right| dx + \int_{\mathbb{R}} \mathrm{sgn}\,(U-V)\,(x)\left((g-f)(v(x))\right)dx, \end{aligned} \tag{3.3}$$

where

$$a(u,v) = \int_0^1 f'(\alpha u + (1-\alpha)v)d\alpha.$$

Note that $U - V$ is differentiable in t due to the Lipschitz continuity in time of u and v with respect to the L^1 norm. From an integration by parts we find that the first term in (3.3) is

$$-\int_{\mathbb{R}} a(u,v)\partial_x \left|U-V\right| dx = \int_{\mathbb{R}} \left|U-V\right| D_x a(u,v)dx,$$

where we give meaning to $D_x a(u,v)$ as a distributional derivative ($|U-V|$ is Lipschitz, but $a(u,v)$ might contain decreasing jumps). This leads to the following upper bound on the time derivative of the W_1-distance,

$$\partial_t W_1(u(t),v(t)) \leqslant \sup_{x\in\mathbb{R}} \left(\partial_x a(u,v)(t)\right)^+ W_1(u(t),v(t)) + \|(g-f)(v(t))\|_{L^1(\mathbb{R})}.$$

Note that $(D_x a(u,v))^+ = (\partial_x a(u,v))^+$ as u and v are Lip^+ bounded. By Grönwall's inequality we deduce that (3.2) holds with

$$C(t) := \exp\left(\|f'\|_{\mathrm{Lip}} Ct\right) \geqslant \exp\left(\int_0^t \sup_{x\in\mathbb{R}} (\partial_x a(u,v)(s))^+ \, ds\right), \tag{3.4}$$

as $a(u,v)$ is increasing in both u and v and both $u(t)$ and $v(t)$ satisfy (1.2). The constant C is the constant in (1.2). □

A similar stability result was established by Nessyahu and Tadmor [18] by studying the dual equation of $u - v$, i.e. the backward in time equation for the dual φ in the Kantorovich–Rubinstein formulation of the W_1-distance, see [28, Thm. 1.14] for the definition.

Remark 3.2. *If u_0, v_0 are non-increasing and $f = g$ in Proposition 3.1, then* (3.2) *reduces to the contraction estimate*

$$W_1\big(u(t), v(t)\big) \leqslant W_1\big(u_0, v_0\big).$$

In (3.2) it is necessary that $u(t)$ and $v(t)$ satisfy (1.2). As front tracking approximations consist of piecewise constants, they will in general not fulfil this condition. In order to overcome this obstacle without risking to sacrifice the second-order convergence rate, we will utilize an old result by Oleĭnik [20, Theorem 2]:

Theorem 3.3 (Oleĭnik [20])**.** *Let f be twice continuously differentiable (and not necessarily convex). Assume that u and v are two piecewise smooth solutions of* (1.1) *which satisfy Oleĭniks condition E. Then if*

$$\left|\int_{y_1}^{y_2} u_0(y) - v_0(y)\, dy\right| \leqslant c \quad \text{for all } y_1, y_2 \in [a,b], \text{ then} \quad \left|\int_{x_1}^{x_2} u(y,t) - v(y,t)\, dy\right| \leqslant c$$

for all x_1, x_2 in the smaller interval $\big[a + Qt, b - Qt\big]$, $Q = \|f\|_{\mathrm{Lip}}$.

We will extend the above result to all $y_1, y_2, x_1, x_2 \in \mathbb{R}$ and to f only locally Lipschitz. To ensure that the piecewise smoothness assumption in Theorem 3.3 is satisfied, we will assume that f is convex.

Lemma 3.4. *Let $u_0, v_0 \in BV(\mathbb{R})$. Consider the respective entropy solutions u and v of* (1.1) *where f is assumed to be convex. If there exists $c > 0$ s.t.*

$$\left|\int_{y_1}^{y_2} u_0(y) - v_0(y)\, dy\right| \leqslant c \tag{3.5}$$

for all pairs $y_1, y_2 \in \mathbb{R}$, then

$$\left|\int_{x_1}^{x_2} u(x,t) - v(x,t)\, dx\right| \leqslant c$$

for all $x_1, x_2 \in \mathbb{R}$ for any finite time $t > 0$.

Proof. We start by approximating the initial data u_0 and v_0 by smooth functions of compact support, $\tilde{u}_0$ and $\tilde{v}_0$, such that

$$\|\tilde{u}_0 - u_0\|_{L^1(\Omega)} < \varepsilon \quad \text{and} \quad \|\tilde{v}_0 - v_0\|_{L^1(\Omega)} < \varepsilon \tag{3.6}$$

on a finite interval Ω (to be determined). Then if f is strictly convex and smooth, $\tilde{u}(t)$ and $\tilde{v}(t)$ will be piecewise smooth, see [5, 25] for example, and Oleĭniks condition E in [20] will be satisfied. As both the approximate initial data are of compact support, the lemma then follows directly for $\tilde{u}(t), \tilde{v}(t)$ from Theorem 3.3 for strictly convex and smooth f.

As $u_0 \in BV(\mathbb{R})$, $\|u_0\|_{L^\infty(\mathbb{R})} \leqslant M$ for some constant M. We extend the result to Lipschitz continuous f by approximating f by a sequence of twice continuously differentiable strictly convex flux functions f^ε such that $\|f - f^\varepsilon\|_{L^\infty(\mathbb{R})} < \varepsilon^3$. Such a function can be found by mollifying f and then adding $a_\varepsilon u^2$ for a suitably small a_ε such that $a_\varepsilon \to 0$ when $\varepsilon \to 0$. Notice that we can choose f^ε such that $\|f^\varepsilon\|_{\mathrm{Lip}} \leqslant \|f\|_{\mathrm{Lip}} + a_\varepsilon M$ on $[-M, M]$. Let $\tilde{u}^\varepsilon$ be the entropy solution of (1.1) with f^ε as a flux function and $\tilde{u}_0$ as initial data. From a L^1-stability estimate by Bouchut and Perthame [1, Thm. 3.1 *(iii)*], we find that

$$\begin{aligned}\int_{\Omega(t)} |\tilde{u}(x,t) - \tilde{u}^\varepsilon(x,t)|\, dx &\leqslant K\Big[(|\Omega(t)| + Qt)\, TV(\tilde{u}_0)t\|f - f^\varepsilon - (f - f^\varepsilon)(0)\|_{L^\infty}\Big]^{1/2} \\ &\leqslant K\Big[(|\mathrm{supp}(\tilde{u}_0)| + 3Qt)\, TV(u_0)t\|f - f^\varepsilon - (f - f^\varepsilon)(0)\|_{L^\infty}\Big]^{1/2} \\ &< K\Big[(|\mathrm{supp}(\tilde{u}_0)| + 3Qt)\, TV(u_0)t\Big]^{1/2} \varepsilon^{3/2}\end{aligned}$$

where K is an absolute constant and $\Omega(t)$ is the maximal support of $\tilde{u}(t)$ and $\tilde{u}^\varepsilon(t)$, and $Q = \|f\|_{\mathrm{Lip}} + a_\varepsilon M$. The same can be done for $\tilde{v}$. We extend the result to $u_0, v_0 \in BV(\mathbb{R})$ by choosing the support of $\tilde{u}_0$ (and of $\tilde{v}_0$) to be $\big[-\frac{1}{2\varepsilon}, \frac{1}{2\varepsilon}\big]$ in order to get

$$\int_{\Omega(t)} |\tilde{u}(x,t) - \tilde{u}^\varepsilon(x,t)|\, dx < C(t)\varepsilon, \tag{3.7}$$

where $C(t) := K\left[1 + (3Qt\varepsilon)^{1/2}\right](TV(u_0)t)^{1/2}$, from the estimate above. By choosing the (smooth) approximations to have the above support, the interval for which (3.6) holds has to be slightly smaller. We can choose it to be $\Omega = \big[-\frac{1}{2\varepsilon} + \varepsilon, \frac{1}{2\varepsilon} - \varepsilon\big]$.

Assume that (3.5) holds for $u_0, v_0 \in BV(\mathbb{R})$. Then, by the triangle inequality, (3.7) and the L^1 contraction property of (1.1),

$$\begin{aligned}\left|\int_{x_1}^{x_2} u(x,t) - v(x,t)\, dx\right| &\leqslant \left|\int_{x_1}^{x_2} \tilde{u}^\varepsilon(x,t) - \tilde{v}^\varepsilon(x,t)\, dx\right| \\ &\quad + \int_{x_1}^{x_2} |\tilde{u}(x,t) - \tilde{u}^\varepsilon(x,t)|\, dx + \int_{x_1}^{x_2} |\tilde{v}(x,t) - \tilde{v}^\varepsilon(x,t)|\, dx \\ &\quad + \int_{x_1}^{x_2} |\tilde{u}(x,t) - u(x,t)|\, dx + \int_{x_1}^{x_2} |\tilde{v}(x,t) - v(x,t)|\, dx \\ &< c + 2C(t)\varepsilon + \|\tilde{u}(t) - u(t)\|_{L^1([x_1,x_2])} + \|\tilde{v}(t) - v(t)\|_{L^1([x_1,x_2])} \\ &\leqslant c + 2C(t)\varepsilon + \|\tilde{u}_0 - u_0\|_{L^1([x_1-Qt,x_2+Qt])} + \|\tilde{v}_0 - v_0\|_{L^1([x_1-Qt,x_2+Qt])} \\ &\leqslant c + 2C(t)\varepsilon + 2\varepsilon,\end{aligned}$$

for any $x_1, x_2 \in \big[-\frac{1}{2\varepsilon} + \varepsilon + Qt, \frac{1}{2\varepsilon} - \varepsilon - Qt\big]$, as the theorem holds for f^ε twice continuously differentiable. Letting $\varepsilon \to 0$ now yields the result. ☐

Remark 3.5. *As the main result in this paper relies on f being convex, we simply assumed convexity in Lemma 3.4 in order to obtain the piecewise smoothness needed to apply Theorem 3.3. Jennings has shown that piecewise smooth solutions of* (1.1) *do exist for a certain class of u_0 and non-convex f, see [12]. By an appropriate approximation of u_0 and f by functions in this class, Lemma 3.4 should be extendible to non-convex f.*

We are now ready to prove that the convergence rate of front tracking approximations is $O(\Delta x^2)$ when measured in W_1.

Theorem 3.6. *Assume that $u_0 \in BV(\mathbb{R})$ is of compact support and satisfies* (1.2). *Let f be twice continuously differentiable and convex. Furthermore, let $u^{\delta,\Delta x}$ be the front tracking solution of* (2.1) *with initial data* (2.2) *and flux* (2.3) *such that $\delta = O(\Delta x)$. Then*

$$W_1\left(u(t), u^{\delta,\Delta x}(t)\right) \leqslant \tilde{C}(t)\Delta x^2, \tag{3.8}$$

where u is the entropy solution of (1.1). *The constant $\tilde{C}(t)$ is defined in* (3.11).

Proof. First observe that with $u_0^{\Delta x}$ as in (2.2), $u_0^{\Delta x} - u_0$ satisfies (2.8), and the W_1-distance is well-defined. Also, by a simple calculation using (2.7),

$$W_1\big(u_0, u_0^{\Delta x}\big) \leqslant \Delta x^2 TV(u_0). \tag{3.9}$$

In order to use Proposition 3.1, we fix an intermediate solution $u^{\delta,\sigma}$ by solving (2.1) replacing $u_0^{\Delta x}$ in (2.2) with the slightly regularized initial data

$$u_0^\sigma(x) = \begin{cases} u_i + \frac{u_{i+1}-u_i}{\Delta x}(x - i\Delta x), \quad x \in [i\Delta x, (i+1)\Delta x], & \text{when } u_i < u_{i+1}, \\ u_0^{\Delta x}(x), & \text{otherwise.} \end{cases}$$

As u_0 satisfies (1.2), it is not hard to see that u_0^σ also will. By the triangle inequality,

$$W_1\left(u(t), u^{\delta,\Delta x}(t)\right) \leqslant W_1\left(u(t), u^{\delta,\sigma}(t)\right) + W_1\left(u^{\delta,\sigma}(t), u^{\delta,\Delta x}(t)\right) = I + II.$$

Applying Proposition 3.1 to I, we get

$$I = W_1\left(u(t), u^{\delta,\sigma}(t)\right) \leqslant C(t)\left[W_1\big(u_0, u_0^\sigma\big) + \int_0^t \|(f - f^\delta)(u^{\delta,\sigma}(s))\|_{L^1(\mathbb{R})} ds\right],$$

where $W_1\big(u_0, u_0^\sigma\big) \leqslant \Delta x^2 TV(u_0)$ follows from (3.9). Furthermore,

$$\begin{aligned} \int_0^t \|(f - f^\delta)(u^{\delta,\sigma}(s))\|_{L^1(\mathbb{R})} ds &\leqslant t \max_{s\in[0,t]} \left|\operatorname{supp}\big(u^{\delta,\sigma}(s)\big)\right| \sup_{u\in[-M,M]} \left|(f - f^\delta)(u)\right| \\ &\leqslant \left[t \max_{s\in[0,t]} \left|\operatorname{supp}\big(u^{\delta,\sigma}(s)\big)\right| \sup_{u\in[-M-\delta,M+\delta]} f''(u)\right] \delta^2, \end{aligned}$$

where the first inequality follows from $(f - f^\delta)(0) = 0$ and the compact support of $u^{\delta,\sigma}$ and the second inequality by (2.3) and a Taylor expansion of f around $j\delta$ where $u \in [j\delta, (j+1)\delta]$. The number M is the constant such that $|u_0| \leqslant M$ (which is finite since $u_0 \in BV(\mathbb{R})$). It follows from (2.2) that also $|u_0^\sigma|, |u_0^{\Delta x}| \leqslant M$. Hence, $|u^{\delta,\Delta x}(t)|, |u^{\delta,\sigma}(t)| \leqslant M$. Thus

$$\begin{aligned} \max_{s\in[0,t]} \left|\operatorname{supp}\big(u^{\delta,\sigma}(s)\big)\right| &\leqslant |\operatorname{supp}(u_0)| + 2\Delta x + \|f^\delta\|_{\mathrm{Lip}([-M,M])} t \\ &\leqslant |\operatorname{supp}(u_0)| + 2\Delta x + \|f\|_{\mathrm{Lip}([-M-\delta,M+\delta])} t =: K(t), \end{aligned} \tag{3.10}$$

and similarly for $u^{\delta,\Delta x}$.

As $u^{\delta,\sigma}(t)$ and $u^{\delta,\Delta x}(t)$ are of compact support and $\int_{\mathbb{R}} u^{\delta,\sigma}(t) - u^{\delta,\Delta x}(t)\, dx = 0$, we estimate II as follows,

$$\begin{aligned} II = W_1\left(u^{\delta,\sigma}(t), u^{\delta,\Delta x}(t)\right) &= \int_{\mathbb{R}} \left| \int_{-\infty}^x \left(u^{\delta,\sigma} - u^{\delta,\Delta x}\right)(y,t) dy \right| dx \\ &\leqslant \max\left\{|\operatorname{supp}(u^{\delta,\sigma})|, |\operatorname{supp}(u^{\delta,\Delta x})|\right\} \sup_x \left| \int_\gamma^x \left(u^{\delta,\sigma} - u^{\delta,\Delta x}\right)(y,t) dy \right| \\ &\leqslant K(t) \sup_x \left| \int_\gamma^x \left(u^{\delta,\sigma} - u^{\delta,\Delta x}\right)(y,t) dy \right|, \end{aligned}$$

where γ is the smallest value in the support of $u^{\delta,\sigma}(t) - u^{\delta,\Delta x}(t)$. Also, for any $x_1, x_2 \in \mathbb{R}$,

$$\left| \int_{x_1}^{x_2} \left(u_0^\sigma - u_0^{\Delta x}\right)(y,t) dy \right| \leqslant \frac{1}{8} C \Delta x^2,$$

where C comes from (1.2). Thus we can apply Lemma 3.4 to conclude that

$$\sup_x \left| \int_\gamma^x \left(u^{\delta,\sigma} - u^{\delta,\Delta x}\right)(y,t) dy \right| \leqslant \frac{1}{8} C \Delta x^2.$$

Combining the two estimates I and II gives (3.8) with

$$\tilde{C}(t) = \left[C(t)\left(TV(u_0) + t\lambda^2 K(t) \sup_{u\in[-M,M]} f''(u)\right) + \frac{K(t)}{8} C\right], \tag{3.11}$$

where $\lambda = \delta/\Delta x$, $C(t)$ is defined in (3.4) and $K(t)$ in (3.10). □

Remark 3.7 (A different approach to the Δx^2 rate)**.** *As mentioned in the introduction, a rate of* Δx^2 *in the* W_1*-distance can be deduced from a result by Karlsen and Risebro [14, Remark 2.2] for* $f \in C^2$ *(not necessarily convex) and* $u_0 \in C_c^1$*. In their paper the focus is on proving the equivalence between entropy solutions of the conservation law* (1.1) *and viscosity solutions of the corresponding Hamilton–Jacobi equation* (2.9) *through the relation* (2.10)*. This is done by translating the front tracking method to a method for* (2.9)*. As a bonus, the authors find that front tracking approximations to* (2.9) *converge at a rate of* Δx^2 *in the* L^∞*-norm when* $u_0 \in C_c^1$*. From* (2.7)*, it is not hard to see that this translates into a rate of* $O(\Delta x^2)$ *in* W_1 *for* (1.1)*,*

$$W_1(u, u^{\delta,\Delta x}) \leqslant C|\Omega(t)|\|U - U^{\delta,\Delta x}\|_{L^\infty} = C|\Omega(t)|\Delta x^2, \tag{3.12}$$

where $|\Omega(t)| = |\mathrm{supp}(u(t)) \cup \mathrm{supp}(u^{\delta,\Delta x}(t))| \leqslant |\mathrm{supp}(u_0)| + 2\Delta x + 2\|f\|_{\mathrm{Lip}([-M-\delta, M+\delta])} t$*, and* C *depends on* $\|f''\|_{L^\infty}$*,* $\|u_0''\|_{L^\infty}$ *and the time* t*.*

Although the convergence rate (3.12) *is the same as the one we prove in this paper in* W_1*, the approach and the assumptions made differ. The results in this paper rely directly on inequalities involving the* W_1*-metric and does not go via front tracking approximations to solutions of* (2.9)*. With this approach we can prove a* Δx^2 *rate in* W_1 *also in the case of* u_0 *with decreasing jump discontinuities, whereas with the approach in [14] the rate will reduce to* Δx *for such* u_0*. The drawback is that we have to assume convexity of* f *which is not required in [14].*

4 The convergence rate in W_∞

In order to prove the Δx rate in W_∞, we require stability estimates of solutions to (1.1) with respect to both the initial data and the flux functions. To obtain these estimates, we will extend the W_∞-contractivity with respect to initial data proved by Carrillo et al. [2] to cover the case of the front tracking equation (2.1). Furthermore, inspired by the proof of the W_∞-contractivity, we will prove a stability estimate with respect to the fluxes.

We will from now on assume that $f(0) = 0$ is the minimum of f. As in [2], we restrict ourselves to initial data in $\mathcal{B}$, and start by assuming that the support of u_0 consists of one connected component. Then, under a certain condition on f, we can ensure that the support of the solution to (1.1) is connected at any later time $t > 0$:

Lemma 4.1. *Assume that* f *is convex and that*

$$\|a\|_{L^\infty([0,M])} \leqslant C, \qquad a(v) := \frac{f'(v)v - f(v)}{v^2}, \tag{4.1}$$

holds for some constant C*, where* $M = \|u_0\|_{L^\infty(\mathbb{R})}$*. Furthermore, let* $u_0 \in \mathcal{B}$ *satisfy* (1.2) *and assume that* $\mathrm{supp}(u_0)$ *consists of one connected component. Then the support of the solution to* (1.1)*,* $u(t)$*, is connected for any* $t > 0$*.*

Proof. Consider the transport equation

$$\partial_t v + \partial_x (bv) = 0, \quad u(0) = u_0, \qquad b(x,t) = \frac{f\,(u(x,t))}{u(x,t)}, \tag{4.2}$$

where $u(t)$ is the entropy solution to (1.1), with its associated characteristics equation

$$\frac{dX(t)}{dt} = b(x,t), \quad X(0) = x. \tag{4.3}$$

If $b(x,t)$ is Lip^+ bounded with respect to x for all $t > 0$, the generalized characteristics of are unique. It follows that the transport equation (4.2) has a unique measure solution $v(t) = X(t)\#u_0$, the pushforward of u_0 by the map $X(t)$, see Poupaud and Rascle [22]. Furthermore, the map $X(x,t)$ is continuous on $\mathbb{R}_+ \times \mathbb{R}$. Thus, if the support of u_0 is connected, the support of $v(t) = X(t)\#u_0$ has to be connected as well.

As u_0 is Lip^+ bounded and f is convex, $u(t)$ is also Lip^+ bounded. Then, under the condition (4.1), one can check that $b(t)$ is indeed Lip^+ bounded. (Note that $a \geqslant 0$ in (4.1) as $f(0) = 0$ is the minimum of f.) It follows that there is a unique measure solution $v(t) = X(t)\#u_0$ to (4.2) with connected support.

The entropy solution $u(t)$ of (1.1) also solves the transport equation (4.2). Hence, as the measure solution of (4.2) is unique, the entropy solution has to satisfy $u(t) = X(t)_{\#}u_0$. Thus, the support of $u(t)$ has to be connected. ☐

Note that the above condition (4.1) holds for f twice continuously differentiable.

The above lemma makes it possible to find an expression for the inverse of the primitive of $u(t)$ such that we can utilize the interpretation (2.6) of W_∞. In the case of an uniformly convex flux function, Carrillo et al. [2] make use of the Hopf–Lax formula (2.11) to explicitly express the primitive U of u and then find the inverse. We will do the same, but under the slightly different assumption that f satisfies (4.1) and is convex and superlinear to include fluxes of the form (2.3). The resulting explicit expression for the inverse is stated in the lemma below.

Lemma 4.2. *Let $u_0 \in \mathcal{B}$ be such that $\operatorname{supp}(u_0)$ consists of one connected component and let f satisfy (4.1). Then the inverse of $U(t) = \int^x u(t)$, where $u(t)$ solves (1.1) is*

$$U^{-1}(\gamma, t) = \max_{0 \leqslant \omega \leqslant \gamma} \left\{ t\tilde{f}\left(\frac{\gamma - \omega}{t}\right) + U_0^{-1}(\omega) \right\},$$

where $\tilde{f}$ is the inverse of f^*, see (2.12), restricted to $[0, \infty)$.

Proof. As $u_0 \in \mathcal{B}$, $u(t) \in \mathcal{B}$, and it follows that $U(t) = \int^x u(t)$ is Lipschitz continuous and bounded. Then, as f is convex and superlinear, $U(t)$ can be expressed with the Hopf–Lax formula (2.11).

As $\operatorname{supp}(u_0)$ is connected, U_0 is strictly increasing from 0 to 1 on a finite interval and we can find its inverse. Furthermore, we know from Lemma 4.1 that the support of $u(t)$ is connected. It follows that $U(t)$ is strictly increasing. Hence, its inverse exists and can be implicitly defined by the Hopf–Lax formula.

Note that as $f(0) = 0$ is the minimum, $f^*(0) = 0$. As f is convex, it is (strictly) increasing on $[0, \infty)$. Then $f^*(p)$ has to be increasing for $p \in [0, \infty)$. The rest of the proof is exactly like the proof of [2, Lemma 2.3]. □

Next follows a contraction result in the W_∞-distance with respect to the initial data. The result proposed here is [2, Thm. 2.4, Thm. 2.5] adjusted to include the front tracking flux (2.3). As we initially assume that f is only convex, we do not need the approximation procedure of the flux in C^1 which is needed in the proofs of [2, Thm. 2.4, Thm. 2.5] to make the contraction estimate valid for convex fluxes. We restate the main details of the proof here for completeness.

Proposition 4.3. *Let $u_0, v_0 \in \mathcal{B} \cap BV(\mathbb{R})$ and let f satisfy (4.1) and be convex and superlinear. Then the respective entropy solutions $u(t)$ and $v(t)$ of (1.1) satisfy*

$$W_\infty\big(u(t), v(t)\big) \leqslant W_\infty(u_0, v_0). \tag{4.4}$$

Proof. This proof is very similar to the ones of [2, Thm. 2.4, Thm. 2.5].

Due to the assumptions on u_0 and f, the primitive of u can be found by the Hopf–Lax formula (2.11). We start by assuming that $\operatorname{supp}(u_0)$ consists of one connected component. Then Lemma 4.2 holds, and we can look at the difference between explicit expressions of the inverses,

$$U^{-1}(\gamma, t) - V^{-1}(\gamma, t) = \max_{0 \leqslant \omega \leqslant \gamma} \left\{ t\tilde{f}\left(\frac{\gamma - \omega}{t}\right) + U_0^{-1}(\omega) \right\} - \max_{0 \leqslant \omega \leqslant \gamma} \left\{ t\tilde{f}\left(\frac{\gamma - \omega}{t}\right) + V_0^{-1}(\omega) \right\}. \tag{4.5}$$

Assume that ω_m realizes the maximum in the first expression. Then

$$\begin{aligned} U^{-1}(\gamma, t) - V^{-1}(\gamma, t) &= t\tilde{f}\left(\frac{\gamma - \omega_m}{t}\right) + U_0^{-1}(\omega_m) - \max_{0 \leqslant \omega \leqslant \gamma} \left\{ t\tilde{f}\left(\frac{\gamma - \omega}{t}\right) + V_0^{-1}(\omega) \right\} \\ &\leqslant t\tilde{f}\left(\frac{\gamma - \omega_m}{t}\right) + U_0^{-1}(\omega_m) - t\tilde{f}\left(\frac{\gamma - \omega_m}{t}\right) - V_0^{-1}(\omega_m) \\ &= U_0^{-1}(\omega_m) - V_0^{-1}(\omega_m) \leqslant \sup_{\omega \in [0,1]} \left| \left(U_0^{-1} - V_0^{-1}\right)(\omega) \right| \end{aligned}$$

By interchanging the roles of U^{-1} and V^{-1}, we find that

$$\left| U^{-1}(\gamma, t) - V^{-1}(\gamma, t) \right| \leqslant \sup_{\omega \in [0,1]} \left| \left(U_0^{-1} - V_0^{-1}\right)(\omega) \right|.$$

Taking the supremum on the left hand side yields (4.4) for initial data in $\mathcal{B}$ with support consisting of one connected component.

We extend the result to general initial data in $\mathcal{B} \cap BV(\mathbb{R})$. Consider two sequences $u_0^n, v_0^n \in \mathcal{B}$ with $\text{supp}(u_0^n)$ and $\text{supp}(v_0^n)$ connected, such that $u_0^n \to u_0$ and $v_0^n \to v_0$ in $L^1(\mathbb{R})$, and $\|u_0^n\|_{L^\infty}, \|v_0^n\|_{L^\infty} \leqslant \max\{\|u_0\|_{L^\infty}, \|v_0\|_{L^\infty}\}$. Then, as proven in [2, Th. 5.5], for any ε, we can choose sequences $u_0^n, v_0^n \in \mathcal{B}$ with connected support such that

$$W_\infty\big(u_0^n, v_0^n\big) \leqslant W_\infty(u_0, v_0) + \varepsilon,$$

and as u_0^n, v_0^n have connected supports, we know that

$$W_p\big(u^n(t), v^n(t)\big) \leqslant W_\infty\big(u^n(t), v^n(t)\big) \leqslant W_\infty(u_0, v_0) + \varepsilon. \tag{4.6}$$

It is well-known that scalar conservation laws satisfy an L^1-contraction property for any $t > 0$,

$$\|u^n(t) - u(t)\|_{L^1(\mathbb{R})} \leqslant \|u_0^n - u_0\|_{L^1(\mathbb{R})},$$
$$\|v^n(t) - v(t)\|_{L^1(\mathbb{R})} \leqslant \|v_0^n - v_0\|_{L^1(\mathbb{R})}.$$

Hence, for any $t \geqslant 0$, $u^n(t) \to u(t)$, $v^n(t) \to v(t)$ in $L^1(\mathbb{R})$ as $n \to \infty$. Furthermore, $\|u^n(t)\|_{L^\infty} \leqslant \|u_0^n\|_{L^\infty} \leqslant \|u_0\|_{L^\infty}$, and similarly for $v^n(t)$. It follows that $\text{supp}\,(u^n(t))$ and $\text{supp}\,(v^n(t))$ are uniformly bounded in n. Due to the bounded supports, the pth order moments of both $u^n(t)$ and $v^n(t)$ will also converge. As convergence in W_p is equivalent to weak convergence and convergence of the pth order moment [28, Thm. 7.12], we can now take the limit as $n \to \infty$ to the left in (4.6),

$$W_p\big(u(t), v(t)\big) \leqslant W_\infty(u_0, v_0) + \varepsilon.$$

We conclude the proof by letting $p \to \infty$ and, as the left hand side does not depend on ε, we send ε to zero. □

We now turn to the stability estimate in W_∞ with respect to the flux functions.

Proposition 4.4. *Let $u_0 \in \mathcal{B}$, and let f and g satisfy* (4.1) *and be convex and superlinear. Then the respective entropy solutions u and v satisfy*

$$W_\infty\big(u(t), v(t)\big) \leqslant t \sup_{\gamma\in[0,1]} \left|\left(\tilde{f} - \tilde{g}\right)\left(\frac{\gamma}{t}\right)\right|, \tag{4.7}$$

where $\tilde{f}, \tilde{g}$ are the inverses of the Legendre transforms f^* and g^*, restricted to $[0, \infty)$, of f and g.

Proof. As in Proposition 4.3, the primitive of u can be found by the Hopf–Lax formula (2.11), and we start by assuming that $\text{supp}(u_0)$ consists of one connected component such that we have an explicit expression for the difference between the inverses (4.5), due to Lemma 4.2. Again, assume that ω_m realizes the maximum in the first expression. Then

$$\begin{aligned} U^{-1}(\gamma, t) - V^{-1}(\gamma, t) &= t\tilde{f}\left(\frac{\gamma - \omega_m}{t}\right) + U_0^{-1}(\omega_m) - \max_{0\leqslant\omega\leqslant\gamma}\left\{t\tilde{g}\left(\frac{\gamma-\omega}{t}\right) + U_0^{-1}(\omega)\right\} \\ &\leqslant t\tilde{f}\left(\frac{\gamma - \omega_m}{t}\right) - t\tilde{g}\left(\frac{\gamma - \omega_m}{t}\right) \leqslant t \sup_{0\leqslant\omega\leqslant\gamma}\left|\tilde{f}\left(\frac{\gamma - \omega_m}{t}\right) - \tilde{g}\left(\frac{\gamma - \omega_m}{t}\right)\right|, \end{aligned}$$

and, after interchanging the roles of $U^{-1}(\gamma, t)$ and $V^{-1}(\gamma, t)$, we get

$$\left|U^{-1}(\gamma, t) - V^{-1}(\gamma, t)\right| \leqslant t \sup_{0\leqslant\omega\leqslant\gamma}\left|\tilde{f}\left(\frac{\gamma-\omega}{t}\right) - \tilde{g}\left(\frac{\gamma-\omega}{t}\right)\right|.$$

Taking the supremum over $\gamma \in [0, 1]$ results in (4.7) for u_0 with support consisting of one connected component.

We extend the result to general initial data in $\mathcal{B}$. Again we consider a sequence $u_0^n \in \mathcal{B}$ with $\text{supp}(u_0^n)$ connected, such that $u_0^n \to u_0$ in $L^1(\mathbb{R})$ and $\|u_0^n\|_{L^\infty} \leqslant \|u_0\|_{L^\infty}$. Then for $u^n(t), v^n(t)$,

$$W_p\big(u^n(t), v^n(t)\big) \leqslant W_\infty\big(u^n(t), v^n(t)\big) \leqslant t \sup_{\gamma\in[0,1]}\left|\left(\tilde{f} - \tilde{g}\right)\left(\frac{\gamma}{t}\right)\right|. \tag{4.8}$$

We use the L^1-contraction property,

$$\|u^n(t) - u(t)\|_{L^1(\mathbb{R})} \leqslant \|u_0^n - u_0\|_{L^1(\mathbb{R})},$$

$$\|v^n(t) - v(t)\|_{L^1(\mathbb{R})} \leqslant \|u_0^n - u_0\|_{L^1(\mathbb{R})},$$

to conclude that $u^n(t) \to u(t)$, $v^n(t) \to v(t)$ in $L^1(\mathbb{R})$, and that the pth order moments of both $u^n(t)$ and $v^n(t)$ converge. Then we can take the limit as $n \to \infty$ in (4.8),

$$W_p\big(u(t), v(t)\big) \leqslant t \sup_{\gamma \in [0,1]} \left| \left(\tilde{f} - \tilde{g}\right) \left(\frac{\gamma}{t}\right) \right|.$$

Letting $p \to \infty$ then yields (4.7). □

The inequality (4.7) does not provide much information in general. But, if we consider a convex function f and its piecewise linear interpolation, we show that the right hand side of (4.7) can be made small in the upcoming lemma. Recall the definition of the Legendre transform in (2.12). One can check that the Legendre transform of a piecewise linear, convex and continuous function,

$$g(u) = g_j + \sigma_j(u - u_j), \quad \sigma_j = \frac{g_{j+1} - g_j}{u_{j+1} - u_j}, \quad u_j \leqslant u \leqslant u_{j+1}, \quad j \in \mathbb{Z} \cap [-J, J], \tag{4.9}$$

is

$$g^*(p) = \begin{cases} +\infty, & p > \sigma_{-J}, \quad p < \sigma_J, \\ u_j p - g_j, & \sigma_{j-1} \leqslant p \leqslant \sigma_j. \end{cases} \tag{4.10}$$

Lemma 4.5. *Let f be convex and superlinear, and let g be the piecewise linear interpolation* (4.9) *of f with $g_j = f(u_j)$. Then*

$$f^*(p) - g^*(p) = 0, \qquad \text{for some} \quad p \in [\sigma_{j-1}, \sigma_j], \qquad \text{for all} \quad j \in \mathbb{Z}.$$

Proof. As f is convex and superlinear g will also be. The same is true for f^* and g^*. Also notice that $f \leqslant g$, so that $f^* \geqslant g^*$.

From (4.10) we see that $g^*(p) = u_j p - f(u_j)$ for $p \in [\sigma_{j-1}, \sigma_j]$ so that

$$f^*(p) - g^*(p) = \sup_{u \in \mathbb{R}} \{pu - f(u)\} - (u_j p - f(u_j)), \quad p \in [\sigma_{j-1}, \sigma_j]. \tag{4.11}$$

We claim that there exists $p \in [\sigma_{j-1}, \sigma_j]$ such that u_j realizes the supremum in the expression for f^*. For this p (4.11) will be zero. For $p \in \partial f(u_j)$ (the sub-differential at u_j), u_j realizes the supremum in (4.11). If $\partial f(u_j) \subset [\sigma_{j-1}, \sigma_j]$ we're done. But this has to be true as $f \leqslant g$, $f(u_j) - g(u_j) = 0$ and f is convex. □

Fully equipped with estimates in W_∞, we prove the Δx convergence rate.

Theorem 4.6. *Let $u_0 \in \mathcal{B} \cap BV(\mathbb{R})$ and let f be twice continuously differentiable and convex. Then the front tracking approximation converges towards the entropy solution of* (1.1) *at a rate of $O(\Delta x)$, i.e.*

$$W_\infty(u(t), u^{\delta,\Delta x}(t)) \leqslant L(t)\Delta x,$$

where $L(t) = 1 + t \max_{u \in [0, M+\delta]} |f''(u)|$.

Proof. By the triangle inequality we have

$$W_\infty(u(t), u^{\delta,\Delta x}(t)) \leqslant W_\infty(u(t), u^{\Delta x}(t)) + W_\infty(u^{\Delta x}(t), u^{\delta,\Delta x}(t)).$$

As f is C^2, one can check that the condition (4.1) holds for both f and the approximation f^δ in (2.3). Then, by Proposition 4.3,

$$W_\infty(u(t), u^{\Delta x}(t)) \leqslant W_\infty(u_0, u_0^{\Delta x}) \leqslant \Delta x,$$

where the last inequality follows from the fact that the primitives of u_0 and (2.2) are both increasing and satisfy $U_0((i + 1/2)\Delta x) - U_0^{\Delta x}((i + 1/2)\Delta x) = 0$ for all $i \in \mathbb{Z}$.

From Proposition 4.4, we have

$$\begin{aligned} W_\infty(u^{\Delta x}(t), u^{\delta,\Delta x}(t)) &\leqslant t \sup_{\gamma \in [0,1]} \left| \left(\tilde{f} - \tilde{g}\right) \left(\frac{\gamma}{t}\right) \right| \\ &= t \sup_{\gamma \in [0,1]} \left| \left((f^*)^{-1} - (g^*)^{-1}\right) \left(\frac{\gamma}{t}\right) \right| \\ &\leqslant t \max_i |\sigma_{j+1} - \sigma_{j-1}| = t \max_{u \in [0, M+\delta]} |f''(u)| 2\delta, \end{aligned}$$

where we in the last step have used Lemma 4.5 and the fact that f^* and g^* are increasing for $\gamma/t > 0$. □

Finally, having established Theorem 4.6, we can conclude that the main theorem of this paper, Theorem 2.2, holds.

5 Concluding remarks

In this paper we have shown that the front tracking approximations to scalar one-dimensional conservation laws with convex fluxes converge at a rate of $\Delta x^{1+1/p}$ in the p-Wasserstein distance. This gives the front tracking method an advantage, in terms of guaranteed convergence rate, over (formally) second-order finite volume schemes for which no second-order convergence rate has been proven for general initial data.

The convergence rate results in this paper are limited to Lip^+ bounded initial data u_0. In the case of Lip^+ unbounded u_0, it is well-known that the solution to (1.1) satisfies

$$\frac{u(x+z,t)-u(x,t)}{z} \leqslant \frac{C}{t}, \quad t>0, \tag{5.1}$$

whenever f is strongly convex, $f'' \geqslant \alpha > 0$. Let $u(t)$ and $v(t)$ be solutions to (3.1), where v_0 is the piecewise constant projection (2.2) to u_0, and $\|g-f\|_{L^\infty} \leqslant O(\Delta x^2)$, where g is strongly convex as well. Then (by an approach similar to the one in Proposition 3.1) preliminary calculations indicate that $W_1\big(u(t),v(t)\big) = O\left(\Delta x^2 \log|\Delta x|\right)$ for Lip^+ unbounded u_0. The front tracking flux $g = f^\delta$ is piecewise linear, but as it is an approximation to the strongly convex function f, the front tracking approximation should satisfy a discrete version of (5.1). This might be sufficient to prove a convergence rate of $\Delta x^2 \log|\Delta x|$ in the Lip^+ unbounded case, but it needs to be investigated further.

The main theorem in this paper strongly depend on the convexity of the flux f. As mentioned in Remark 3.7, the Δx^2 convergence rate in W_1 that one can deduce from [14] can be extended to non-convex fluxes as long as $u_0 \in C_c^1$. Remark 3.7 and the discussion on Lip^+ unbounded u_0 indicate that the rate might be lower for more general initial data. Whether the Δx rate in W_∞ can be extended to the non-convex case is unclear. The proofs in Section 4 depend on an explicit expression for the generalized inverse of the primitive. Due to the more complex nature of $u(t)$ in the non-convex case, a feasible expression for the generalized inverse is currently out of reach.

References

[1] F. Bouchut and B. Perthame. Kružkov's estimates for scalar conservation laws revisited. *Trans. Amer. Math. Soc.*, 350(7):2847–2870, 1998.

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

[3] J. A. Carrillo and G. Toscani. Wasserstein metric and large-time asymptotics for nonlinear diffusion equations. in *New Trends in Mathematical Physics(In Honour of the Salvatore Rionero 70th Birthday)*, 2005.

[4] C. Dafermos. Polygonal approximation of solutions of the initial value problem for a conservation law. *J. Math. Anal.*, 38:33–41, 1972.

[5] C. Dafermos. Generalized characteristics and the structure of solutions of hyperbolic conservation laws. *Indiana Uni. Math. J.*, 26(6):1097–1119, 1977.

[6] L. C. Evans. *Partial Differential Equations.* AMS, Providence, RI, 1998.

[7] U. S. Fjordholm and S. Solem. Second-order convergence of monotone schemes for conservation laws. *SIAM J. Numer. Anal.*, 54(3):1920–1945, 2016.

[8] H. Holden, L. Holden and R. Høegh-Krohn. A numerical method for first order nonlinear scalar hyperbolic conservation laws in one-dimension. *Comput. Math. Appl.*, 15:595–602, 1988.

[9] H. Holden and N. H. Risebro. *Front Tracking for Hyperbolic Conservation Laws, 2nd edn.* Springer, Berlin Heidelberg, 2015.

[10] H. Holden and N. H. Risebro. A method of fractional steps for scalar conservation laws without the CFL condition. *Math. Comp.*, 60:221–232, 1993.

[11] B. I. Hong. Regularity for Hamilton–Jacobi equations via approximation. *Bull. Austral. Math. Soc.*, 51:195–213, 1995.

[12] G. Jennings. Piecewise smooth solutions of a single conservation law exist. *Adv. in Math.*, 33:192–205, 1979.

[13] K. H. Karlsen. On the accuracy of a numerical method for two-dimensional scalar conservation laws based on dimensional splitting and front tracking. Preprint Series 30, Department of Mathematics, University of Oslo, 1994.

[14] K. H. Karlsen and N. H. Risebro. A note on front tracking and the equivalence between viscosity solutions of Hamilton-Jacobi equations and entropy solutions of scalar conservation laws. *Nonlin. Anal.*, 50:455–469, 2002.

[15] S. N. Kružkov First order quasi-linear equations in several independent variables. *Math. USSR Sbornik*, 10:217–243, 1970.

[16] N. N. Kuznetsov. Accuracy of some approximate methods for computing the weak solutions of a first-order quasi-linear equation. *USSR Computational Mathematics and Mathematical Physics*, 16(6):105–119, 1976.

[17] B. J. Lucier. A Moving Mesh Numerical Method for Hyperbolic Conservation Laws. *Math. Comp.*, 46(173):59–69, 1986.

[18] H. Nessyahu and E. Tadmor. The convergence rate of approximate solutions for nonlinear scalar conservation laws. *SIAM J. Numer. Anal.*, 29:1505–1519, 1992.

[19] H. Nessyahu, E. Tadmor, and T. Tassa. The convergence rate of Godunov type schemes. *SIAM J. Numer. Anal.*, 31:1–16, 1994.

[20] O. A. Oleĭnik. Uniqueness and stability of the generalized solution of the Cauchy problem for a quasi-linear equation. *Uspekhi Mat. Nauk*, 14:165–170, 1959.

[21] O. A. Oleĭnik. Discontinuous solutions of non-linear differential equations. *Amer. Math. Soc. Transl. Ser.*, 26:95–172, 1963.

[22] F. Poupaud and M. Rascle. Measure Solutions to the Linear Multi-dimensional Transport Equation with Non-smooth Coefficients. *Commun. Part. Diff. Eq.*, 22:337–358, 1997.

[23] F. Şabac. The Optimal Convergence Rate of Monotone Finite Difference Methods for Hyperbolic Conservation Laws. *SIAM J. Numer. Anal.*, 34(6):2306–2318, 1997.

[24] E. Tadmor. Local Error Estimates for Discontinuous Solutions of Nonlinear Hyperbolic Equations. *SIAM J. Numer. Anal.*, 28(4):891–906, 1991.

[25] E. Tadmor and T. Tassa. On the piecewise smoothness of entropy solutions to scalar conservation laws. *Comm. on PDE's*, 18:1631–1652, 1993.

[26] Z.-H. Teng. On the accuracy of fractional step methods for conservation laws in two dimensions. *SIAM J. Num. Anal.*, 31:43–63, 1994.

[27] Z.-H. Teng and P. Zhang. Optimal L1-Rate of Convergence for The Viscosity Method and Monotone Scheme to Piecewise Constant Solutions with Shocks. *SIAM J. Numer. Anal.*, 34:959–978, 1997.

[28] C. Villani. *Topics in Optimal Transportation.* AMS, Providence, RI, 2003.

[29] A. Vol'pert. The spaces BV and quasilinear equations. *USSR Math. Sb*, 2:225-267, 1967.