European Journal of Applied Mathematics (2025), **36**, pp. 613–637
doi:10.1017/S0956792524000317

PAPER

Optimal transport through a toll station

Arthur Stephanovitch[1], Anqi Dong[2] and Tryphon T. Georgiou[2]

[1]École normale supérieure Paris-Saclay, Gif-sur-Yvette, Île-de-France 91190, France.
[2]Mechanical and Aerospace Engineering, University of California, Irvine, CA 92697, USA.
Corresponding author: Anqi Dong; Email: anqid2@uci.edu

Received: 19 April 2023; **Revised:** 01 December 2023; **Accepted:** 26 May 2024; **First published online:** 18 September 2024

Keywords: optimal transport; flux constraint; convex optimization

2020 Mathematics Subject Classification: 49-XX (Primary); 26B25, 52A41 (Secondary)

Abstract
We address the problem of optimal transport with a quadratic cost functional and a constraint on the flux through a constriction along the path. The constriction, conceptually represented by a toll station, limits the flow rate across. We provide a precise formulation which, in addition, is amenable to generalization in higher dimensions. We work out in detail the case of transport in one dimension by proving existence and uniqueness of solution. Under suitable regularity assumptions, we give an explicit construction of the transport plan. Generalization of flux constraints to higher dimensions and possible extensions of the theory are discussed.

1. Introduction

In recent years, the Monge-Kantorovich theory of optimal mass transport has impacted a wide range of mathematical and scientific disciplines from probability theory to geophysics and from thermodynamics to machine learning [1, 7, 10, 20, 21]. Indeed, the Monge-Kantorovich paradigm of transporting one distribution to another, by seeking to minimize a suitable cost functional, has proved enabling in many ways. It gave rise to a class of control problems [5, 6], underlies variational principles in physics [13, 16], provided natural regularization penalties in inverse problems [2], led to new identification techniques in data science [12, 17], in graphical models [9], and linked to large deviations in probability theory [4, 15].

Historically, the Monge-Kantorovich theory proved especially relevant in economics when physical commodities were the object to be transported – a fact that contributed to L. Kantorovich receiving the Nobel prize. Extensions that pertain to physical constraints along the transport naturally were soon brought up. For instance, moment-type constraints have been considered in [18, Section 4.6.3] and, more recently, far generalized in [8]. Congestion being a significant impediment to transport has also drawn the attention of theorists and practitioners alike. For instance, besides optimizing for transportation, considerations of an added path-dependent cost to alleviate congestion have been considered in [3], see also [19, Section 4] for a comprehensive study of this research direction. Along a different direction, constraints have been introduced for probability densities as part of the optimization problem. Such bounds can capture the capacity of the transportation medium and as such have been studied in [14] or dynamical flow constraints as in [11].

In the present work, we formulate and address a natural variant of the standard optimal mass transport problem by imposing a hard constraint on the flux rate at a point along the path between distributions. Specifically, we pose and resolve the most basic problem where the restriction on throughput of the transport plan takes place at a single point. With this constraint in place, we seek to minimize a usual quadratic cost functional.

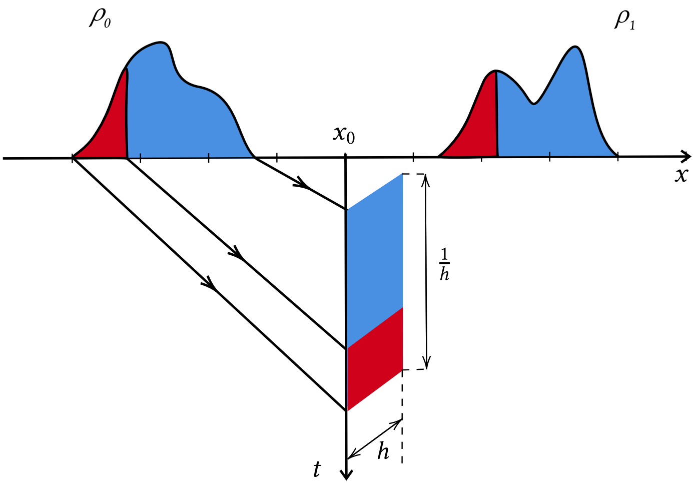

Figure 1. *Illustration of optimal transport through a toll with finite throughput.*

The analysis we provide focuses on one-dimensional distributions, with transport taking place on $\mathbb{R}$. We prove existence and uniqueness of an optimal transport plan and, under suitable regularity conditions, give an explicit construction. A slight generalization of our formulation, where the distributions have support on $\mathbb{R}^d$ but the transportation is to take place through a specified 'constriction' point, with a similar throughput constraint, can be worked out in the same manner, and it is sketched in the concluding remarks. The more general case where the transport takes place on higher dimensional manifolds with the throughput through possibly multiple points, curves, or surfaces similarly restricted is substantially more challenging and much remains open.

The problem formulation and ideas in the mathematical analysis that follows can be visualized by appealing to Figure 1. We begin with two probability densities ρ_0, ρ_1 having support on $\mathbb{R}$ and finite second-order moments and seek to transport one to the other, ρ_0 to ρ_1, within a window of time (herein, of duration normalized to 1) while minimizing a quadratic cost in the local velocity. That is, we seek to minimize the action integral of kinetic energy along the transport path. The minimal cost of the unrestricted transport is the so-called Wasserstein distance $\mathcal{W}_2(\rho_0, \rho_1)$ (a metric on the space of probability measures); we refer to standard references [20, 21] for the unconstrained optimal transport problem. The schematic in Figure 1 exemplifies a constraint at a pre-specified point, x_0, that can be seen as the location of constriction, or, of a toll along the transport, where throughput is bounded. That is, the flow rate across x_0 for mass times velocity is bounded by a value h. A vertical axis pointing downwards at x_0 marks the time when a specific mass-element crosses the toll, necessitating at least $1/h$ duration for the unit mass of the probability density ρ_0 to go through, in the most favourable case where the throughput rate is maintained for the duration (that is normalized to 1 time unit).

In the body of the paper, we prove existence and uniqueness of an optimal transport plan, and, assuming suitable regularity of the distributions, we provide an explicit construction for the solution. We further explore consequences of the toll being kept maximally 'busy' while mass is being transported through, in conjunction with minimizing the quadratic cost criterion on the kinetic energy, and we highlight ensuing properties of the optimal plan.

Specifically, in Section 2 we develop the formulation of the flux constraint and give a precise definition of the problem (Problem 1). In Section 3, we prove existence and uniqueness (Theorem 1) of solution, while conveniently recasting the problem in terms of a flux variable (Problem 2). Section 5 deals with the structural form of the transport and properties of solutions; we summarize the basic elements that allow an explicit construction of the solution in Algorithm 1. Section 6 provides a rudimentary example of transporting between uniform distributions, that highlights the essential

https://doi.org/10.1017/S0956792524000317 Published online by Cambridge University Press

property that speed needs to be suitably adjusted so as to fully utilize the throughput of the toll, while minimizing the quadratic cost. We close (Section 7) with a discussion on possible extensions of the problem to higher dimensions and multiple tolls. While the theory may be readily extended in certain cases, much remains to be understood. Such problems are of natural engineering and scientific interest.

2. Problem formulation

We consider two probability measures ρ_0 and ρ_1 on $\mathbb{R}$, having finite second-order moments and that each admits a density with respect to the Lebesgue measure. Throughout, we follow a standard (slight) abuse of notation and use the same symbols $\rho_i : \mathbb{R} \to \mathbb{R}_{\geq 0}$, for $i = 1, 2$, for the corresponding probability densities of the two measures. In this case, there exists a unique non-decreasing optimal transport map T from ρ_0 to ρ_1 (Theorem 2.5, [19]). Following a standard formulation of transport problems, we consider $Y : [0, 1] \times \mathbb{R} \to \mathbb{R}$ such that[1] $Y_{0\#\rho_0} = \rho_0$ and $Y_{1\#\rho_0} = \rho_1$, and we are interested in minimizing

$$J(\partial_t Y) := \int_0^1 \int_{\mathbb{R}} (\partial_t Y_t(x))^2 \rho_0(x) dx dt. \tag{2.1}$$

In the absence of any additional constraint on Y, the solution is $Y_t^\star(x) = x + t(T(x) - x)$ for T the optimal transport map between ρ_0 and ρ_1 and $J(\partial_t Y_t^\star) = \mathcal{W}_2^2(\rho_0, \rho_1)$, the squared Wasserstein-2 distance between the two [20]. Here, however, for a certain $x_0 \in \mathbb{R}$, we introduce a constraint on the flux passing through x_0, as explained below. Throughout the paper, T will always denote the optimal transportation map in the absence of any such constraint. The purpose of the present work of course is to develop a theory that addresses the case of transport with a bound on the flux through x_0.

When all functions are smooth and well defined, a flux constraint at x_0 can be expressed as

$$|\rho_t(x_0) v_t(x_0)| \leq h \quad \forall t \in (0, 1)$$

for ρ_t the density of $Y_{t\#\rho_0}$ and $v_t(x_0) = \partial_t Y_t(Y_t^{-1}(x_0))$. However in the general case, if ρ_t is not continuous (or does not even exist), this constraint is not well defined. One way to deal with such a situation is to recast the constraint as requiring that,[2] $\forall t \in (0, 1)$,

$$\limsup_{\substack{\alpha_1 \to 0 \\ \alpha_2 \to 0}} \frac{1}{|\alpha_2 - \alpha_1|} \int \mathbb{1}_{\{x_0 \in (Y_t(x)+\alpha_1, Y_t(x)+\alpha_2)\}} |\partial_t Y_t(x)| \rho_0(x) dx \leq h. \tag{2.2}$$

Then, if ρ_0 is continuous and Y_t is a C^1 diffeomorphism, the left-hand side (LHS) of (2.2) amounts to

$$\begin{aligned} \text{LHS (2.2)} &= \limsup_{\substack{\alpha_1 \to 0 \\ \alpha_2 \to 0}} \frac{1}{|\alpha_2 - \alpha_1|} \int \mathbb{1}_{\{x_0 \in (y+\alpha_1, y+\alpha_2)\}} |\partial_t Y_t(Y_t^{-1}(y))| \rho_t(y) dy \\ &= \rho_t(x_0) v_t(x_0). \end{aligned}$$

Interestingly, when Y_t fails to be a C^1 diffeomorphism, special care is needed. For instance, take $x_0 = 0$ and $Y_t(x) = \mathbb{1}_{\{x \in [-2,-1]\}}(1 - 2t)^3 x$. The constraint (2.2) is satisfied since $\partial_t Y_t(x) = 0$ at $t = 1/2$, and no mass sits near the toll for any $t \neq 1/2$. Thus, the formulation (2.2) fails to capture the situation where infinite mass passes through with zero velocity. We reformulate so as to avoid this technicality.

Consider the modified constraint that bounds the flux passing through x_0, expressed as requiring that $\forall t \in (0, 1)$

$$\limsup_{\substack{\alpha_1 \to 0 \\ \alpha_2 \to 0}} \frac{1}{|\alpha_2 - \alpha_1|} \int \mathbb{1}_{\{x_0 \in (Y_{t+\alpha_1}(x), Y_{t+\alpha_2}(x))\}} \rho_0(x) dx \leq h. \tag{2.3}$$

[1] As is common, $Y_{t\#\rho_0}$ denotes the push-forward of ρ_0 under Y_t, see [20].
[2] We use the standard notation $\mathbb{1}_A$ for the characteristic function of the set A.

https://doi.org/10.1017/S0956792524000317 Published online by Cambridge University Press

In the case where Y_t is C^1, using the Taylor expansion of Y in time, the left-hand side (LHS) of (2.3) amounts to

$$\text{LHS (2.3)} = \limsup_{\substack{\alpha_1\to 0\\ \alpha_2\to 0}} \frac{1}{|\alpha_2-\alpha_1|}\int \mathbb{1}_{\{x_0\in(Y_t(x)+\partial_t Y_t(x)\alpha_1+o(\alpha_1),Y_t(x)+\partial_t Y_t(x)\alpha_2+o(\alpha_2))\}}\rho_0(x)dx$$

$$= \limsup_{\substack{\alpha_1\to 0\\ \alpha_2\to 0}} \int \Big(\mathbb{1}_{\{Y_t(x)\in(x_0-\partial_t Y_t(x)\alpha_2+o(\alpha_2),x_0-\partial_t Y_t(x)\alpha_1+o(\alpha_1))\}}\frac{|\partial_t Y_t(x)|\mathbb{1}_{\{|\partial_t Y_t(x)|>0\}}}{|\alpha_2-\alpha_1||\partial_t Y_t(x)|}$$

$$+\mathbb{1}_{\{Y_t(x)\in(x_0+o(\alpha_1),x_0+o(\alpha_2))\}}\frac{\mathbb{1}_{\{\partial_t Y_t(x)=0\}}}{|\alpha_2-\alpha_1|}\Big)\rho_0(x)dx.$$

Using a change of variables, we readily see that (2.3) implies (2.2) and that if Y_t is a C^1 diffeomorphism, the two constraints are identical. Note also that, $\forall t\in(0,1)$, condition (2.3) is equivalent to

$$\forall\alpha_1,\alpha_2\in\mathbb{R},\qquad \int \mathbb{1}_{\{x_0\in(Y_{t+\alpha_1}(x),Y_{t+\alpha_2}(x))\}}\rho_0(x)dx\le h|\alpha_2-\alpha_1|. \tag{2.4}$$

Define

$$\Omega=\{x\in\text{Supp}\,(\rho_0)\mid x_0\in(x,T(x))\text{ or }x_0\in(T(x),x)\}, \tag{2.5}$$

where T is the optimal transport map of the unconstrained problem. Thus, Ω contains the support of mass that needs to cross the toll station, at some point in time, in either direction. From (2.4), it is evident that $h\ge\rho_0(\Omega)$ is necessary for the existence of a map satisfying the constraint (since the transport will take place over the time interval [0, 1]). Typically, $h>\rho_0(\Omega)$ is required, except in some special cases where $h=\rho_0(\Omega)$ may suffice, as for example when $\rho_0=\mathbb{1}_{\{[0,1]\}}$, $x_0=1$ and $\rho_1=\mathbb{1}_{\{[1,2]\}}$. From here on we assume that $h>\rho_0(\Omega)$.

We are now in a position to cast our optimization problem in terms of a velocity field $v_t(x)$ that will effect the transport; formally, $v_t(x)=\partial_t Y_t(x)$ relates to our earlier notation when functions are smooth. For any $v\in L^2([0,1]\times\mathbb{R},\mathbb{R})$, define the map $Y^v:[0,1]\times\mathbb{R}\to\mathbb{R}$ as the flow of v:

$$Y_t^v=\text{Id}+\int_0^t v_\tau d\tau, \tag{2.6}$$

with Id denoting the identity map in $\mathbb{R}$.

Definition 2.1. *Let $\mathcal{V}$ denote the set of functions $v:[0,1]\times\mathbb{R}\to\mathbb{R}$ for which $\int_0^1\int_{\mathbb{R}}v(x,t)^2\rho_0(x)dxdt<\infty$ and are such that Y^v (the flow of v) satisfies:*

(i) $Y^v_{1\#\rho_0}=\rho_1$

(ii) $\forall t\in(0,1)$, *Y_t^v satisfies the constraint (2.4).*

Our problem can now be stated as follows.

Problem 1. *Determine*

$$\inf_{v\in\mathcal{V}}J(v), \tag{2.7}$$

over the class $\mathcal{V}$ of Definition 2.1, and assert existence, uniqueness, and the functional form of minimizing solutions.

We first show that there exist velocity fields belonging to $\mathcal{V}$.

Proposition 2.2. *Supposing that ρ_0, ρ_1 have finite second-order moments and are absolutely continuous with respect to the Lebesgue measure, the set of functions $\mathcal{V}$ of Definition 2.1 is non-empty.*

Proof. Let us build an explicit map $v\in\mathcal{V}$. First, for $x\notin\Omega$ (with Ω defined in (2.5)), we can set $v_t(x)=T(x)-x$, as mass at the point x does not cross the toll. Let us now take care of the points that lie in Ω

https://doi.org/10.1017/S0956792524000317 Published online by Cambridge University Press

by splitting it into two sets:

$$\Omega_1 = \{x \in \Omega | x < x_0\} \quad \text{and} \quad \Omega_2 = \{x \in \Omega | x > x_0\}.$$

The velocity $v_t(x)$ for $x \in \Omega$ will be piecewise constant with 3 different pieces.

(i) Define T_1 the optimal transport map between the densities $\rho_{0|\Omega_1}$ (the restriction of ρ_0 to Ω_1) and $\mu_1(x) = h\mathbb{1}_{\{x\in[x_0-\rho_0(\Omega_1)/h, x_0]\}}$. Likewise, define T_2 the optimal transport map between the densities $\rho_{0|\Omega_2}$ and $\mu_2(x) = h\mathbb{1}_{\{x\in[x_0, x_0+\rho_0(\Omega_2)/h]\}}$. Let

$$t^\star = \rho_0(\Omega)$$

and $a : \Omega \to \mathbb{R}$ defined by

$$a(x) = \mathbb{1}_{\{x\in\Omega_1\}} \frac{2}{1 - t^\star/h}(T_1(x) - x) + \mathbb{1}_{\{x\in\Omega_2\}} \frac{2}{1 - t^\star/h}(T_2(x) - x).$$

Then, the flow of a transports the densities $\rho_{0|\Omega_1}$ onto μ_1 and $\rho_{0|\Omega_2}$ onto μ_2 in a time $(1 - t^\star/h)/2$.

(ii) Define $b : [x_0 - \rho_0(\Omega_1)/h, x_0 + \rho_0(\Omega_2)/h] \to \mathbb{R}$ by

$$b(x) = \mathbb{1}_{\{x\in[x_0-\rho_0(\Omega_1)/h, x_0]\}} \frac{\rho_0(\Omega_1)}{t^\star} - \mathbb{1}_{\{x\in[x_0, x_0+\rho_0(\Omega_2)/h]\}} \frac{\rho_0(\Omega_2)}{t^\star}.$$

Likewise, the flow of b transports the densities μ_1 onto $\mu_1^+(x) = \mu_1(x - \frac{\rho_0(\Omega_1)}{h})$ and μ_2 onto $\mu_2^-(x) = \mu_2(x + \frac{\rho_0(\Omega_2)}{h})$ in a time $\frac{t^\star}{h}$. Writing T for the optimal transport map between ρ_0 and ρ_1, let be T_1' the optimal transport map between the densities μ_1^+ and $T_{\#\rho_{0|\Omega_1}}$. Likewise, let be T_2' the optimal transport map between the densities μ_2^- and $T_{\#\rho_{0|\Omega_2}}$ respectively.

(iii) Finally, define $c : [x_0 - \rho_0(\Omega_2)/h, x_0 + \rho_0(\Omega_1)/h] \to$ by

$$c(x) = \mathbb{1}_{\{x\in[x_0-\rho_0(\Omega_2)/h, x_0]\}} \frac{2}{1 - t^\star/h}(T_2'(x) - x) + \mathbb{1}_{\{x\in[x_0, x_0+\rho_0(\Omega_1)/h]\}} \frac{2}{1 - t^\star/h}(T_1'(x) - x).$$

Then, the flow of c transports the densities μ_1^+ onto $T_{\#\rho_{0|\Omega_1}}$ and μ_2^- onto $T_{\#\rho_{0|\Omega_2}}$ in a time $(1 - t^\star/h)/2$.

Therefore, the flow of the three velocity fields a, b, c applied successively transports $\rho_{0|\Omega}$ onto $T_{\#\rho_{0|\Omega}}$ in a time $(1 - t^\star/h)/2 + t^\star/h + (1 - t^\star/h)/2 = 1$. Furthermore, mass crosses the toll only during the interval $[(1 - t^\star/h)/2, t^\star/h + (1 - t^\star/h)/2]$ and the flow rate at the toll values $h\frac{\rho_0(\Omega_1)}{t^\star} + h\frac{\rho_0(\Omega_2)}{t^\star} = h$ during this time. Finally, as ρ_0 and ρ_1 have finite second-order moment, we easily see that a, b and c verify the L^2 condition. □

3. Existence of a solution

Let us first define the set of flows of velocity fields in the class $\mathcal{V}$.

Definition 3.1. *The class of functions $\mathcal{Y}$ is defined as the set of maps $Y : [0, 1] \times \mathbb{R} \to \mathbb{R}$ such that there exist $v \in \mathcal{V}$ so that Y is the flow of v, i.e., $Y_t = Y_t^v = \mathrm{Id} + \int_0^t v_\tau d\tau$.*

From here on, the v in the notation Y_t^v is suppressed as we are truly interested in the transport map. We first derive certain useful properties of candidate minimizers of our problem. To this end, for any $Y \in \mathcal{Y}$ and $x \in \Omega$, we define

$$toll_Y(x) = \inf\{t \mid x_0 = Y_t(x)\}.$$

Thus, the function $toll_Y$ specifies the times of transit through the toll station of mass that is initially located at x and then transported via Y.

https://doi.org/10.1017/S0956792524000317 Published online by Cambridge University Press

It is clear that the function $toll_Y$ must be injective[3] for a minimizing solution and that mass flow takes place always in the same direction across the toll station. Then, $\forall t \in (0, 1)$, (2.3) is equivalent to

$$\limsup_{\substack{\alpha_1 \to 0 \\ \alpha_2 \to 0}} \frac{1}{|\alpha_2 - \alpha_1|} toll_{Y\#\rho_0}((t+\alpha_1, t+\alpha_2)) \le h,$$

and so, if $toll_{Y\#\rho_0}$ (the measure on $[0, 1]$ that weighs the mass that goes through x_0 at different times $t \in [0, 1]$) admits a continuous density ϱ_{toll}, the constraint amounts to $\varrho_{toll}(t) \le h$. Note also that this condition is different than simply stating $\rho_t(x_0) \le h$, as the latter does not take into account the speed of transport. Then we see that for $x \notin \Omega$, we can restrict ourselves to considering maps $Y \in \mathcal{Y}$ such that $Y_t(x) = x + t(T(x) - x)$ for T the optimal transport map between ρ_0 and ρ_1. Thus, in the sequel, without loss of generality we always suppose that $\Omega = \text{Supp}(\rho_0)$ and that $\sup \text{Supp}(\rho_0) \le x_0 \le \inf \text{Supp}(\rho_1)$. For clarity, we group together the set of assumptions used on probability measures ρ_0 and ρ_1.

Assumptions 1. *The probability measures ρ_0 and ρ_1 satisfy the following:*

- *They have finite second-order moments.*
- *They are absolutely continuous with respect to the Lebesgue measure.*
- $\sup \text{Supp}(\rho_0) \le x_0 \le \inf \text{Supp}(\rho_1)$.

Note that Ω in (2.5) is precisely $\text{Supp}(\rho_0)$, since all the mass in ρ_0 needs to be transported across the toll. Also note that if h is large enough so that the optimal transport map T verifies the constraint, then T automatically provides a solution to Problem 1. Thus, the problem is equivalent to that in the unconstrained case.

For $Y \in \mathcal{Y}$ and $toll_Y$, its corresponding transit-time function define the map $\overline{Y} : [0, 1] \times \mathbb{R} \to \mathbb{R}$ by

$$\overline{Y}_t(x) = \begin{cases} x + t\frac{x_0 - x}{toll_Y(x)} & \text{if } t \le toll_Y(x) \\ x_0 + (t - toll_Y(x))\frac{T(x) - x_0}{1 - toll_Y(x)} & \text{if } t \ge toll_Y(x) \end{cases} \tag{3.1}$$

and note $\overline{\mathcal{Y}} = \{\overline{Y} \mid Y \in \mathcal{Y}\}$ the set of functions of this type. The next statement states that we can restrict our minimization problem to functions of the form (3.1). Specifically, it states that for any candidate minimizer $Y \in \mathcal{Y}$, the speed of transport needs to remain constant at all times prior to transit, and again, constant at all times after transit. In addition, from the functional form, we see that $Y_1 = T(x)$ for all x. This last statement says that the final destination of mass originally located at x is the same, whether we apply T or the optimal plan that abides by the constraint; the only thing that changes in the two cases is the speed while the mass traverses the segment before x_0 and after (cf. example in Section 6).

Proposition 3.2. *For ρ_0 and ρ_1 satisfying Assumptions 1, we have*

$$\inf_{Y \in \mathcal{Y}} J(\partial_t Y) = \inf_{\overline{Y} \in \overline{\mathcal{Y}}} J(\partial_t \overline{Y})$$

Proof. For $Y \in \mathcal{Y}$ and $toll_Y$, define

$$Y^c_t(x) = \begin{cases} x + t\frac{x_0 - x}{toll_Y(x)} & \text{if } t \le toll_Y(x) \\ x_0 + (t - toll_Y(x))\frac{Y_1(x) - x_0}{1 - toll_Y(x)} & \text{if } t \ge toll_Y(x) \end{cases} \tag{3.2}$$

Thus, Y^c maintains the terminal destination $Y_1(x)$ and the crossing time $toll_Y(x)$, for the mass that was initially at x, while the speed of each particle remains constant before and after crossing. It follows that $Y^c \in \mathcal{Y}$ and that $J(\partial_t Y^c) \le J(\partial Y)$ as for any $x \in \text{Supp}(\rho_0)$ and path $\gamma_x \in C^1([0, toll_Y(x)], [x, x_0])$, the mean squared velocity of this path is always larger than the one of the path of constant speed, i.e.,

$$\int_0^{toll_Y(x)} \dot{\gamma}_x(t)^2 dt \ge \int_0^{toll_Y(x)} \left(\frac{x_0 - x}{toll_Y(x)}\right)^2 dt = \frac{(x_0 - x)^2}{toll_Y(x)}.$$

[3]This is followed by cyclic monotonicity since the cost is convex, see [21, Section 2.3].

https://doi.org/10.1017/S0956792524000317 Published online by Cambridge University Press

We can then restrict $\mathcal{Y}$ to the set of functions that are of the form (3.2) since candidate minimizers will always be of that form.

As the position $Y_1(x)$ in (3.2) doesn't impact the constraint (2.3), we consider how $Y_1(x)$ may depend on the time of crossing $toll_Y(x)$. Specifically, Y_1 must be a minimum over the set of functions

$$\{f : \text{Supp}\,(\rho_0) \to \text{Supp}\,(\rho_1) \mid f_{\#\rho_0} = \rho_1\},$$

for the cost

$$\int_{toll_Y(x)}^1 \int_{\mathbb{R}} \left(\frac{Y_1(x) - x_0}{1 - toll_Y(x)}\right)^2 \rho_0(x)dxdt = \int_{\mathbb{R}} \frac{(Y_1(x) - x_0)^2}{1 - toll_Y(x)} \rho_0(x)dx.$$

From this, we deduce that almost everywhere, $toll_Y(x) \le toll_Y(y)$ iff $Y_1(x) \ge Y_1(y)$. Indeed, for $x_1, x_2, t_1, t_2 \in \mathbb{R}_{>0}$, we have that if $x_1 < x_2$ and $t_1 \le t_2$ then $\frac{x_1}{t_1} + \frac{x_2}{t_2} < \frac{x_1}{t_2} + \frac{x_2}{t_1}$. Then if $toll_Y(x) \le toll_Y(y)$ and $Y_1(x) < Y_1(y)$, then

$$\frac{(Y_1(x) - x_0)^2}{1 - toll_Y(x)} + \frac{(Y_1(y) - x_0)^2}{1 - toll_Y(y)} > \frac{(Y_1(x) - x_0)^2}{1 - toll_Y(y)} + \frac{(Y_1(y) - x_0)^2}{1 - toll_Y(y)}$$

so Y_1 would not be optimal. Furthermore, as the problem is reversible (we can switch ρ_0 and ρ_1), we can deduce in the same way that $toll_Y(x) \ge toll_Y(y)$ iff $x \le y$. Therefore, $Y_1(x)$ is increasing, and we conclude that it is identical to T the optimal transport map between ρ_0 and ρ_1. □

From Proposition 3.2, we deduce that for Y, the flow of a (candidate) optimal solution, the map $x \mapsto toll_Y$ is strictly decreasing on the support of ρ_0 and that Y_t is one to one, for all t. We also deduce that for almost every $x \in \text{Supp}\,(\rho_0)$, the velocity $v_t(x)$ is constant (in t) before crossing the toll, changes at the toll, and then stays the same until $Y_t(x)$ reaches $T(x)$ at time $t = 1$. This can be formally written as

$$v_t(x) = v_0(x)\mathbb{1}_{\{t<toll_Y(x)\}} + v_1(x)\mathbb{1}_{\{t\ge toll_Y(x)\}}.$$

Let us write $\mathrm{v}(x) = \frac{x_0 - x}{toll_Y(x)}$ for the velocity of transport *prior* to crossing the toll, for the mass initially located at x at the start. Then, in light of Proposition 3.2, our problem is reduced to the following formulation.

Problem 1′. *Determine*

$$\begin{aligned} &\arg\min_{\mathrm{v}} \int_0^1 \int_{\mathbb{R}} \left(\mathrm{v}(x)^2 \mathbb{1}_{\{t \le \frac{x_0 - x}{\mathrm{v}(x)}\}} + \left(\frac{T(x) - x_0}{1 - \frac{x_0 - x}{\mathrm{v}(x)}}\right)^2 \mathbb{1}_{\{t \ge \frac{x_0 - x}{\mathrm{v}(x)}\}}\right) \rho_0(x)dxdt \\ &= \arg\min_{\mathrm{v}} \int_{\mathbb{R}} \left(\mathrm{v}(x)(x_0 - x) + \frac{(T(x) - x_0)^2}{1 - \frac{x_0 - x}{\mathrm{v}(x)}}\right) \rho_0(x)dx, \end{aligned} \tag{3.3}$$

subject to $x \mapsto \frac{x_0 - x}{\mathrm{v}(x)} = toll_{\mathrm{v}}(x)$ *being decreasing and bounded between 0 and 1, (since* $\frac{x_0 - x}{\mathrm{v}(x)} = toll(x)$*) and*

$$\limsup_{\substack{\alpha_1 \to 0 \\ \alpha_2 \to 0}} \frac{1}{|\alpha_2 - \alpha_1|} \int \mathbb{1}_{\{(t+\alpha_1)\mathrm{v}(x) < x_0 - x < (t+\alpha_2)\mathrm{v}(x)\}} \rho_0(x)dx \le h. \tag{3.4}$$

We now argue the existence of a minimizer $\mathrm{v}^\star$.

Proposition 3.3. *For ρ_0 and ρ_1 satisfying Assumptions 1, Problem 1 admits a solution.*

Proof. Let $(\mathrm{v}_n)_n$ be a minimizing sequence of Problem 2 (therefore also of Problem 1) and write $toll_n : \text{Supp}\,(\rho_0) \to (0, 1)$ the associated toll function: $toll_n(x) = \frac{x_0 - x}{\mathrm{v}_n(x)}$. Let $(\alpha_k)_k$ be a dense sequence in $\text{Supp}\,(\rho_0)$ (for example the rational numbers). By compactness, we have that $\forall k \in \mathbb{N}$, $toll_n(\alpha_k)$ admits a converging subsequence in n. Then using a diagonal argument, there exists a subsequence $(\mathrm{v}_{\varphi(n)})_n$ and $\beta_k \in [0, 1]$ such that, $\forall k \in \mathbb{N}$, $toll_{\varphi(n)}(a_k) \xrightarrow[n\to+\infty]{} \beta_k$, and $\alpha_k \le \alpha_l \iff \beta_l \le \beta_k$. For $x \in \text{Supp}\,(\rho_0)$, and $(\alpha_{\psi_x(k)})_k$ a decreasing subsequence converging to x, let be $toll(x) = \lim_k \beta_{\psi_x(k)} = \lim_k \lim_n toll_{\varphi(n)}(a_{\psi_x(k)})$, which is well defined as $\beta_{\psi_x(k)}$ is increasing. Then, $toll_{\varphi(n)}(x)$ converges to $toll(x)$ for any x being a point of continuity of $toll$. As $toll$ is a non-increasing map, it has at most a countable number of points of discontinuity; therefore, $toll_{\varphi(n)}$ converges to $toll$ a.e. Let be $f : \mathbb{R} \to \mathbb{R}$ a continuous bounded map, we then have by dominated convergence

$$\int_{\mathbb{R}} f(toll_{\varphi(n)}(x))\rho_0(x)dx \xrightarrow[n\to+\infty]{} \int_{\mathbb{R}} f(toll(x))\rho_0(x)dx,$$

so $toll_{\varphi(n)\#\rho_0}$ converges weakly to $toll_{\#\rho_0}$. For $x \in \mathrm{Supp}\,(\rho_0) \setminus \{x \mid toll(x) = 0\}$, define $\mathrm{v}(x) = \frac{x_0 - x}{toll(x)}$, it is well defined a.e. because $\{x \mid toll(x) = 0\}$ has measure 0 as $(\mathrm{v}_n)_n$ is a minimizing sequence. Then, $\mathrm{v}_{\varphi(n)}$ converges a.e. to v and as the constraint (3.4) is equivalent to

$$\forall \alpha_1, \alpha_2 \in \mathbb{R}, \qquad toll_{\mathrm{v}}((t+\alpha_1, t+\alpha_2)) \leq h|\alpha_2 - \alpha_1|,$$

v verifies the constraint. Finally, using Fatou's lemma we have $\lim_n J(\mathrm{v}_n) \geq J(\mathrm{v})$ so v is a minimizer of Problem 1. □

4. Uniqueness of the solution

Before we proceed with the proof of uniqueness of the minimizer, we recast our problem in terms of flux as the optimization variable. For $u \in L^1([0, 1] \times \mathbb{R}, \mathbb{R})$, a candidate flux (i.e., mass times velocity) defines a corresponding mass-measure ρ_t^u on $\mathbb{R}$ by duality via: $\forall \phi \in C_c^\infty(\mathbb{R}, \mathbb{R})$,

$$\int_{\mathbb{R}} \phi(x)d\rho_t^u(x) = \int_{\mathbb{R}} \phi(x)\rho_0(x)dx + \int_0^t \int_{\mathbb{R}} (\nabla\phi(x))u_r(x)dxdr.$$

Equivalently, we have that ρ^u solves in the weak sense the continuity equation

$$\begin{cases} \partial_t \rho_t^u = -\nabla \cdot u \\ \rho_0^u = \rho_0 \end{cases}.$$

For a flux u such that $\forall t \in (0, 1)$, ρ_t^u admits a positive density, let us express the cost of u as

$$\mathrm{J}(u) = \int_0^1 \int_{\mathbb{R}} \frac{u_t(y)^2}{\rho_t^u(y)} dydt \tag{4.1}$$

In the above, by a slight abuse of notation as it is often done, we used ρ^u to denote both the measure and the corresponding density, allowing these to be distinguished by the specific usage and context. Let us define the set of admissible fluxes.

Definition 4.1. *The class of functions $\mathcal{U}$ is defined as the set of flux $u : [0, 1] \times \mathbb{R} \to \mathbb{R}$ defined a.e. such that $\mathrm{J}(u) < \infty$ and*

- *(i′)* $\forall t \in (0, 1)$, *ρ_t^u admits a positive density function and $\rho_1^u = \rho_1$*
- *(ii′) for T_t^u the optimal transport map between ρ_0 and ρ_t^u, for $x \in \mathrm{Supp}\,(\rho_0)$ a.e. the map $t \mapsto T_t^u(x)$ is a bijection from $[0, 1]$ to $[x, T(x)]$ and is differentiable a.e.*
- *(iii′) satisfy*

$$\forall t \in (0, 1), \ \limsup_{x_1 \to x_0,\ x_2 \to x_0} \frac{1}{|x_2 - x_1|} \int_{x_1}^{x_2} |u_t(y)|dy \leq h. \tag{4.2}$$

In the proof of Proposition 4.2, we show that $\mathcal{U}$ is non-empty as any solution $v \in \mathcal{V}$ of Problem 2 generates a flux $u \in \mathcal{U}$. We can now recast our problem in terms of flux as the optimization variable.

Problem 2. *Consider*

$$\inf_{u \in \mathcal{U}} \mathrm{J}(u). \tag{4.3}$$

https://doi.org/10.1017/S0956792524000317 Published online by Cambridge University Press

over the class $\mathcal{U}$ of Definition 4.1. Determine existence, uniqueness, and functional form for a minimizing solution u.

We will first prove the equivalence of the above formulation in Problem 2 with that in Problem 1. The advantage of Problem 2 is that the constraint is now convex which will be convenient in proving uniqueness. Note that here we use roman J with argument the flux field, to echo the earlier usage in (2.1) where the action integral J first appeared with argument the velocity.

Proposition 4.2. *For ρ_0 and ρ_1 satisfying Assumptions 1, Problems 1 and 3 are equivalent.*

Proof. Let $Y \in \overline{\mathcal{Y}}$ be a solution of Problem 1, $v_t(\cdot) = \partial_t Y_t(Y_t^{-1}(\cdot))$ the associated velocity (defined everywhere except at the points $(toll(x), x)$, for all $x \in \mathrm{Supp}\,(\rho_0)$) and $\rho_t = Y_{t\#\rho_0}$ the associated mass flow. Then for $\phi \in C_c^\infty(\mathbb{R}, \mathbb{R})$, we have

$$\begin{aligned}\int_{\mathbb{R}} \phi(x) d\rho_t(x) &= \int_{\mathbb{R}} \phi(Y_t(x))\rho_0(x)dx \\ &= \int_{\mathbb{R}} \left(\phi(Y_0(x)) + \int_0^t \partial_t\phi(Y_r(x))dr\right)\rho_0(x)dx \\ &= \int_{\mathbb{R}} \left(\phi(Y_0(x)) + \int_0^t \nabla\phi(Y_r(x))v_r(Y_r(x))dr\right)\rho_0(x)dx \\ &= \int_{\mathbb{R}} \phi(x)\rho_0(x)dx + \int_0^t \int_{\mathbb{R}} \nabla\phi(x)v_r(x)d\rho_r(x)dr.\end{aligned}$$

Therefore, Y defines a unique flux $u \in L^1([0, 1] \times \mathbb{R}, \mathbb{R})$ (u is L^1 by Jensen inequality) by $u_t(x) = v_t(x)\rho_t(x)$ with $\mathrm{J}(u) = J(\partial_t Y)$. Furthermore, writing T_t for the optimal transport map between ρ_0 and ρ_t, we have for $x \in \mathrm{Supp}\,(\rho_0)$ almost everywhere, the map $t \mapsto T_t(x)$ is a bijection from $[0, 1]$ to $[x, T(x)]$ and its differential at x equals $v_t(T_t(x))$. On the other hand, for $\mathrm{v}(x) = \partial_t Y_0(x) = \frac{x_0 - x}{toll_{\mathrm{v}}(x)}$ we have that the left-hand side of (3.4) amounts to

$$\begin{aligned}\text{LHS (2.4)} &= \limsup_{\substack{\alpha_1\to 0\\ \alpha_2\to 0}} \int \mathbb{1}_{\{Y_t(x)+\alpha_1\mathrm{v}(x)<x_0<Y_t(x)+\alpha_2\mathrm{v}(x)\}} \frac{\mathrm{v}(x)}{|\alpha_2-\alpha_1|\mathrm{v}(x)}\rho_0(x)dx \\ &= \limsup_{\substack{\epsilon_1\to 0\\ \epsilon_2\to 0}} \frac{1}{|\epsilon_2-\epsilon_1|}\int \mathbb{1}_{\{y+\epsilon_1<x_0<y+\epsilon_2\}} v_t(y)\rho_t(y)dy.\end{aligned}$$

Therefore, $u \in \mathcal{U}$ and we conclude that $\inf_{u\in\mathcal{U}} \mathrm{J}(u) \le \min_{Y\in\mathcal{Y}} J(\partial_t Y)$.

For establishing the reverse direction, let be $u \in \mathcal{U}$ and define T_t^u the optimal transport map between ρ_0 and ρ_t^u. For $F_t(x) = \int_{-\infty}^x \rho_t^u(y)dy$, the cumulative distribution function of ρ_t^u, it is well known that $T_t^u(x) = F_t^{-1}(F_0(x))$, see [21, Chapter 1]. For $x \in \mathrm{Supp}\,(\rho_0)$, write $f_x : [x, T(x)] \to [0, 1]$ for the inverse in time of $t \mapsto T_t^u(x)$, i.e., for $y \in [x, T(x)]$, $T^u_{f_x(y)}(x) = y$. Since $\forall t \in [0, 1]$ we have $F_t(F_t^{-1}(F_0(x))) = F_0(x)$, we obtain for all $\phi \in C_c^\infty((0, 1) \times \mathbb{R}, \mathbb{R})$,

$$\int_0^1 \int_{\mathbb{R}} \partial_t\phi_t(x)F_t(F_t^{-1}(F_0(x)))dxdt = \int_0^1 \int_{\mathbb{R}} \partial_t\phi_t(x)F_0(x)dxdt$$

$$\int_0^1 \int_{\mathbb{R}} \partial_t\phi_t(x) \int_{\mathbb{R}} \mathbb{1}_{\{y\le T_t^u(x)\}}\rho_t^u(y)dydxdt = 0. \tag{4.4}$$

https://doi.org/10.1017/S0956792524000317 Published online by Cambridge University Press

Furthermore, for $x, y \in \mathbb{R}$ using integration by parts we have

$$\int_0^1 \partial_t \phi_t(x) \mathbb{1}_{\{y \le T_t^u(x)\}} \rho_t^u(y) dt = \int_0^1 \partial_t \phi_t(x) \mathbb{1}_{\{f_x(y) \le t\}} \rho_t^u(y) dt$$
$$= -\int_0^1 \phi_t(x) \mathbb{1}_{\{f_x(y) \le t\}} \partial_t \rho_t^u(y) dt - \phi_{f_x(y)}(x) \rho_{f_x(y)}^u(y),$$

so the left-hand side of (4.4) equals

$$-\int_{\mathbb{R}} \int_{\mathbb{R}} \left(\int_0^1 \phi_t(x) \mathbb{1}_{\{y \le T_t^u(x)\}} \partial_t \rho_t^u(y) dt + \phi_{f_x(y)}(x) \rho_{f_x(y)}^u(y) \right) dxdy. \tag{4.5}$$

Through a change of variable $s = f_x(y)$ and recalling that $T_{f_x(y)}^u(x) = y$, we obtain that

$$\int_{\mathbb{R}} \phi_{f_x(y)}(x) \rho_{f_x(y)}^u(y) dy = \int_0^1 \phi_s(x) \rho_s^u(T_s^u(x)) \partial_t T_s^u(x) ds.$$

On the other hand, as (ρ_t^u, u_t) solves the continuity equation we have

$$\int_{\mathbb{R}} \mathbb{1}_{\{y \le T_t^u(x)\}} \partial_t \rho_t^u(y) dy = -\int_{\mathbb{R}} \mathbb{1}_{\{y \le T_t^u(x)\}} \nabla u_t(y) dy = -u_t(T_t^u(x)).$$

Then, we have that (4.5) equals

$$-\int_{\mathbb{R}} \int_0^1 \phi_t(x) \big(-u_t(T_t^u(x)) + \rho_t^u(T_t^u(x)) \partial_t T_t^u(x) \big) dtdx.$$

Therefore, from (4.4) we deduce that (t, x) almost everywhere we have

$$\partial_t T_t^u(x) = \frac{u_t(T_t^u(x))}{\rho_t^u(T_t^u(x))}.$$

Therefore, T_t^u defines a map in $\mathcal{Y}$ such that

$$J(\partial_t T_t^u) = \int_0^1 \int_{\mathbb{R}} (\partial_t T_t^u(x))^2 dt \rho_0(x) dx = \int_0^1 \int_{\mathbb{R}} \left(\frac{u_t(T_t^u(x))}{\rho_t^u(T_t^u(x))} \right)^2 dt \rho_0(x) dx$$
$$= \int_0^1 \int_{\mathbb{R}} \frac{u_t(y)^2}{\rho_t^u(y)} dt dy = \mathrm{J}(u).$$

Then, we deduce that $\inf_{u \in \mathcal{U}} \mathrm{J}(u) \ge \min_{Y \in \mathcal{Y}} J(\partial_t Y)$. Furthermore, as every $y \in \arg\,\min_{Y \in \mathcal{Y}} J(\partial_t Y)$ defines a flux $u \in \mathcal{U}$ with $\mathrm{J}(u) \le J(\partial_t y)$, in particular, we have

$$\min_{u \in \mathcal{U}} \mathrm{J}(u) = \min_{Y \in \mathcal{Y}} J(\partial_t Y).$$

□

Using the equivalence of Problems 1 and 2, we can now prove the uniqueness of the minimizer.

Theorem 1. *For ρ_0 and ρ_1 satisfying Assumptions 1, Problem 2 (and so Problem 1) admits a unique solution.*

Proof. Suppose that we have u^1 and u^2, two solutions of $\min_{u \in \mathcal{U}} \mathrm{J}(u)$. For $\lambda \in (0, 1)$, $t \in [0, 1]$ and $x \in \mathbb{R}$, by convexity of the map $\mathbb{R} \times \mathbb{R}_{>0} \ni (a, b) \mapsto \frac{a^2}{b}$, we have that

$$\frac{(\lambda u_t^1(x) + (1-\lambda) u_t^2(x))^2}{\lambda \rho_t^{u^1}(x) + (1-\lambda) \rho_t^{u^2}(x)} \le \lambda \frac{u_t^1(x)^2}{\rho_t^{u^1}(x)} + (1-\lambda) \frac{u_t^2(x)^2}{\rho_t^{u^2}(x)}. \tag{4.6}$$

As u^1 and u^2 are both solutions, we have that

$$\lambda \mathrm{J}(u^1) + (1-\lambda) \mathrm{J}(u^2) = \min_{u \in \mathcal{U}} \mathrm{J}(u) \le \mathrm{J}(\lambda u^1 + (1-\lambda) u^2),$$

https://doi.org/10.1017/S0956792524000317 Published online by Cambridge University Press

so (4.6) is an equality almost everywhere. To lighten the notation, let us write $a_1 = u_t^1(x)$, $b_1 = \rho_t^1(x)$, $a_2 = u_t^2(x)$ and $b_2 = \rho_t^2(x)$. Then, when (4.6) is an equality, we have that the polynomial

$$P(\lambda) = (\lambda a_1 + (1-\lambda)a_2)^2 - (\lambda b_1 + (1-\lambda)b_2)(\lambda \frac{a_1^2}{b_1} + (1-\lambda)\frac{a_2^2}{b_2})$$

is identically zero. One can prove that the polynomial P is identically zero iff $\frac{a_1}{b_1} = \frac{a_2}{b_2}$ i.e., $\frac{u_t^1(x)}{\rho_t^{u^1}(x)} = \frac{u_t^2(x)}{\rho_t^{u^2}(x)}$. In this case, we then have $\partial_t T_t^{u^1}(T_t^{u^1}(x)) = \partial_t T_t^{u^2}(T_t^{u^2}(x))$ for $T_t^{u^i}$ the optimal transport map from ρ_0 to $\rho_t^{u^i}$ with $i = 1, 2$. Then, as $T_0^{u^1} = T_0^{u^2} = \mathrm{Id}$, we deduce that $\rho^{u^1} = \rho^{u^2}$ so $u^1 = u^2$ almost everywhere. □

5. Properties and structural form of the solution under smoothness assumption

We are now in a position to build explicitly the solution $\mathrm{v}^\star$ of Problem 1 in the case when ρ_0 and ρ_1 have additional smoothness assumptions. In the process of building the solution, we also the establish structural properties of the solution. Throughout this section, we will assume that ρ_0 and ρ_1 satisfy the following set of assumptions.

Assumptions 2. *The probability measures ρ_0 and ρ_1 satisfy the following:*

- *They are absolutely continuous with respect to the Lebesgue measure with probability density functions that are continuous, have bounded convex support, and are bounded from below on their support.*
- $\sup \mathrm{Supp}\,(\rho_0) < x_0 < \mathrm{Supp}\,(\rho_1)$.

Under the stated assumptions on ρ_0, ρ_1, by using the closed-form expression for the optimal transport map T in dimension one [21, Chapter 1], it is immediate to see that T is C^1. For $\mathrm{v} : \mathrm{Supp}\,(\rho_0) \to \mathbb{R}$ such that[4] $x \mapsto toll_{\mathrm{v}}(x) = \frac{x_0 - x}{\mathrm{v}(x)}$ is decreasing and bounded between 0 and 1 on $\mathrm{Supp}\,(\rho_0)$, the expression

$$C_y(\mathrm{v}) = \limsup_{\substack{\alpha_1 \to 0 \\ \alpha_2 \to 0}} \frac{1}{|\alpha_2 - \alpha_1|} \int \mathbb{1}_{\{(toll_{\mathrm{v}}(y)+\alpha_1)\mathrm{v}(x) < x_0 - x < (toll_{\mathrm{v}}(y)+\alpha_2)\mathrm{v}(x)\}} \rho_0(x) dx$$

gives the value of the flux passing through the toll station when the mass initially at y is crossing. Let us first prove that from the additional assumptions on ρ_0 and ρ_1, we have that the solution is continuous.

Proposition 5.1. *For ρ_0 and ρ_1 satisfying Assumptions 2, the solution $\mathrm{v}^\star \in L^2$ admits a continuous representative.*

Proof. From Section 1, we know that the solution $\mathrm{v}^\star \in L^2$ admits a representative such that the function $x \mapsto toll_{\mathrm{v}^\star}(x) = \frac{x_0 - x}{\mathrm{v}^\star(x)}$ is decreasing. Now by absurd, suppose that $\mathrm{v}^\star$ is not continuous on the interior of $\mathrm{Supp}\,(\rho_0)$. Then, there exists y in the interior of $\mathrm{Supp}\,(\rho_0)$ and $\epsilon > 0$ such that $\forall \delta > 0$, $\exists y_\delta \in \mathrm{Supp}\,(\rho_0)$ with $|y - y_\delta| < \delta$ and $|\mathrm{v}^\star(y) - \mathrm{v}^\star(y_\delta)| > \epsilon$. Suppose that $\forall \delta > 0$, $y_\delta - y > 0$ (the proof would be the same for $y_\delta - y < 0$). As $toll_{\mathrm{v}^\star}(x) = \frac{x_0 - x}{\mathrm{v}^\star(x)}$ is not continuous in y and it is decreasing, we have that

$$\lim_{\substack{x \to y \\ x > y}} toll_{\mathrm{v}^\star}(x) < toll_{\mathrm{v}^\star}(y).$$

Let us define $\alpha = T(y) - y - \mathrm{v}^\star(y)$. Suppose first that $\alpha > 0$, then as $toll_{\mathrm{v}^\star}$ is decreasing we have for $x \le y$,

$$\mathrm{v}^\star(x) = (toll_{\mathrm{v}^\star}(x))^{-1}(x_0 - x) \le (toll_{\mathrm{v}^\star}(y))^{-1}(x_0 - x) = \frac{x_0 - x}{x_0 - y}(T(y) - y - \alpha)$$

$$= T(x) - x - \alpha \frac{x_0 - x}{x_0 - y} + \frac{x_0 - x}{x_0 - y}(T(y) - y) - T(x) - x.$$

[4]The notation $toll_{\mathrm{v}}$ signifies $toll_Y$, for the corresponding Y obtained via (2.6).

https://doi.org/10.1017/S0956792524000317 Published online by Cambridge University Press

Then, as T is continuous, we have that $-\alpha\frac{x_0-x}{x_0-y}+\frac{x_0-x}{x_0-y}(T(y)-y)-T(x)-x$ tends to $-\alpha$ when x tends to y. We deduce that for $\gamma>0$ small enough,

$$\mathrm{v}^\star(x)+\gamma(T(x)-x)\leq T(x)-x$$

for all $x\in(y-\gamma,y]$. Then by strict convexity of J, the function

$$\mathrm{v}_2(x)=\mathrm{v}^\star(x)+\gamma\mathbb{1}_{\{x\in(y-\gamma,y]\}}(T(x)-x-\mathrm{v}^\star(x))$$

verifies that

$$\begin{aligned}
&J(\mathrm{v}^\star)-J(\mathrm{v}_2)\\
&=\gamma\int_{y-\delta}^{y}\left((\mathrm{v}^\star(x)-T(x)-x)(x_0-x)+\frac{(T(x)-x_0)^2}{1-\frac{x_0-x}{\mathrm{v}^\star(x)}}-\frac{(T(x)-x_0)^2}{1-\frac{x_0-x}{T(x)-x}}\right)\rho_0(x)dx\\
&=\gamma\int_{y-\delta}^{y}\left(\mathrm{v}^\star(x)(x_0-x)+\frac{(T(x)-x_0)^2}{1-\frac{x_0-x}{\mathrm{v}^\star(x)}}-(T(x)-x)^2\right)\rho_0(x)dx\\
&>0,
\end{aligned}$$

as the path of constant speed $t\mapsto x+t(x-T(x))$ has a smaller total squared velocity than the path $t\mapsto x+t\big(\mathrm{v}^\star(x)\mathbb{1}_{\{t\leq toll_Y(x)\}}+\frac{T(x)-x_0}{1-\frac{x_0-x}{\mathrm{v}^\star(x)}}\mathbb{1}_{\{t>toll_Y(x)\}}\big)$. Furthermore for $\mathcal{Y}$ small enough, we have that $C_x(\mathrm{v}_2)<h$ for all $x\in(y-\gamma,y)$, as ρ_0 is continuous and $toll_{\mathrm{v}^\star}$ is decreasing so $C_x(\mathrm{v}^\star)<C_{y_\delta}(\mathrm{v}^\star)$ for δ small enough. Therefore, we have that v_2 is a better solution to the problem.

If now $\alpha\leq 0$, then by continuity of T we have that for $\gamma>0$ small enough, $\mathrm{v}^\star(x)+2\gamma\geq T(x)-x$, for all $x\in(y,y+\gamma]$. As previously we can find a better solution $\mathrm{v}_2(x)=\mathrm{v}^\star(x)-\gamma\mathbb{1}_{\{x\in(y-\gamma,y]\}}(T(x)-x-\mathrm{v}^\star(x))$ to the problem which contradicts the fact that $\mathrm{v}^\star$ is the minimizer. □

The next proposition states that at the points where $\mathrm{v}^\star$ does not saturate the constraint, $\mathrm{v}^\star$ is equal to the unconstrained transport $T-\mathrm{Id}$.

Proposition 5.2. *For ρ_0 and ρ_1 satisfying Assumptions 2, if there exists $y\in\mathrm{Supp}\,(\rho_0)$ such that $C_y(\mathrm{v}^\star)<h$, then we have $\mathrm{v}^\star(y)=T(y)-y$.*

Proof. Suppose $\exists y\in\mathrm{Supp}\,(\rho_0)$ such that $C_y(\mathrm{v}^\star)<h$ and $\mathrm{v}^\star(y)\neq T(y)-y$. Define $g_\epsilon(x)=\mathbb{1}_{\{x\in(y-\epsilon,y+\epsilon)\}}\epsilon^3\exp\big(-\frac{1}{\epsilon^2-(x-y)^2}+\frac{1}{\epsilon^2}\big)$. Then, there exist $\epsilon\neq 0\in\mathbb{R}$ and $\delta>0$ such that $\forall x\in(y-\delta,y+\delta)$ we have $|\mathrm{v}^\star(x)+g_\epsilon(x)-(T(x)-x)|<|\mathrm{v}^\star(x)-(T(x)-x)|$ and $C_x(\mathrm{v}^\star+g_\epsilon)<h$, since g_ϵ introduces a vanishingly small bump at a suitable location. Through the flow of $\mathrm{v}^\star+g_\epsilon$, the points in $(y-\delta,y+\delta)$ travel with a velocity that is closer to the velocity of the path $t\mapsto x+t(T(x)-x)$. The velocity of the path $t\mapsto x+t(T(x)-x)$ being constant, it is the one that is minimal for the cost of the total squared velocity. By strict convexity of J, we have that $J(\mathrm{v}^\star)>J(\mathrm{v}^\star+g_\epsilon)$. This contradicts the optimality of $\mathrm{v}^\star$. □

We can now deduce some regularity of the function $\mathrm{v}^\star$.

Corollary 5.3. *For ρ_0 and ρ_1 satisfying Assumptions 2, the optimal solution $\mathrm{v}^\star$ of Problem 1 is C^1 almost everywhere.*

Proof. As T is C^1, then $\mathrm{v}^\star$ is also C^1 at points y that lie in the interior of the closed set $\{y\in\mathrm{Supp}\,(\rho_0)\mid \mathrm{v}^\star(y)=T(y)-y\}$. Otherwise if for some y it holds that $\mathrm{v}^\star(y)\neq T(y)-y$, then $\exists\delta>0$ such that $\forall x\in(y-\delta,y+\delta)$, $\mathrm{v}^\star(x)\neq T(x)-x$ which implies by Proposition 5.2 that $C_x(\mathrm{v}^\star)=h$. Solve the ordinary differential equation

$$\begin{cases}\partial_x\mathrm{v}(x)=\frac{\mathrm{v}(x)^2\rho_0(x)-h\mathrm{v}(x)}{h(x_0-x)} & \text{for } y-\delta\leq x\leq y+\delta\\ \mathrm{v}(y+\delta)=\mathrm{v}^\star(y+\delta)\end{cases}\tag{5.1}$$

for $\mathrm{v}(x)$. Note that this equation is solved backwards, starting from a terminal condition at $y+\delta$.

https://doi.org/10.1017/S0956792524000317 Published online by Cambridge University Press

It can be shown that the function v is well defined by establishing existence and uniqueness of the solution to (5.1) using the Cauchy-Lipschitz theorem and inherent boundedness. Indeed, if

$$\mathrm{v}(x) > \frac{h}{\inf\{\rho_0(y) \mid y \in \mathrm{Supp}\,(\rho_0)\}},$$

then $\partial_x \mathrm{v}(x) > 0$, and so v is decreasing with decreasing value of its argument on a small interval $[x-\epsilon, x]$, i.e., $\tau \mapsto \mathrm{v}(x-\tau)$ is decreasing for $\tau \in [0, \epsilon]$. Likewise, if

$$0 < \mathrm{v}(x) < \frac{h}{\sup\{\rho_0(y) \mid y \in \mathrm{Supp}\,(\rho_0)\}},$$

then $\partial_x \mathrm{v}(x) < 0$, and so v is increasing (again with decreasing value of its argument) on a small interval $[x-\epsilon, x]$. Therefore, we can conclude that if v exists, it is bounded on its interval of definition. Now, as $\mathrm{v}^\star(y+\delta) > 0$, and $\mathrm{v} \mapsto \frac{\mathrm{v}^2\rho_0(x) - h\mathrm{v}}{h(x_0 - x)}$ is Lipschitz on any compact set, we can apply the Cauchy-Lipschitz theorem to establish existence and uniqueness. We then have that v is well defined as the unique solution of (5.1) on $[y-\delta, y+\delta]$ and is C^1 on this interval. From the definition of v, it follows that $C_x(\mathrm{v}) = h$, and therefore v has the same flux as $\mathrm{v}^\star$. By uniqueness, v which is C^1 on $[y-\delta, y+\delta]$ is optimal, i.e., $\mathrm{v} = \mathrm{v}^*$. Finally as $\mathrm{v}^\star$ is C^1 on the interior of the set $\{y \in \mathrm{Supp}\,(\rho_0) \mid \mathrm{v}^\star(y) = T(y) - y\}$ and is also C^1 on the set $\{y \in \mathrm{Supp}\,(\rho_0) \mid \mathrm{v}^\star(y) \neq T(y) - y\}$, we deduce that $\mathrm{v}^\star$ is C^1 almost everywhere as the boundary of those two sets is at most countable. Note that the set $\{y \in \mathrm{Supp}\,(\rho_0) \mid \mathrm{v}^\star(y) = T(y) - y\}$ might have an empty interior. However, it does not change the fact that the set of points where $\mathrm{v}^\star$ is not C^1 lies in the boundary of the set $\{y \in \mathrm{Supp}\,(\rho_0) \mid \mathrm{v}^\star(y) \neq T(y) - y\}$. As the cardinality of the boundary of this set is at most countable, we can still conclude that $\mathrm{v}^\star$ is C^1 almost everywhere. □

Now that we have established that $\mathrm{v}^\star$ is C^1 a.e., we can write the constraint (3.4) for functions $\mathrm{v} \in C^1(\mathrm{Supp}\,(\rho_0), \mathbb{R})$ as: for $x \in \mathrm{Supp}(\rho_0)$, a.e.

$$C_x(\mathrm{v}) = \frac{\mathrm{v}(x)\rho_0(x)}{1 + \frac{x_0 - x}{\mathrm{v}(x)}\partial_x \mathrm{v}(x)} \leq h. \tag{5.2}$$

We now define the class of candidate functions that we focus on in the present section.

Definition 5.4. *The class of functions V is defined as the set of maps* $\mathrm{v} : \mathrm{Supp}\,(\rho_0) \to \mathbb{R}$, *that are* C^1 *a.e. and are such that*

(i) the map $x \mapsto \frac{x_0 - x}{\mathrm{v}(x)}$ *is decreasing and bounded between 0 and 1*

(ii) v *verifies condition* (5.2) *a.e.*

Note that from Corollary 5.3, we have that V is non-empty. For $\mathrm{v} : \mathrm{Supp}\,(\rho_0) \to \mathbb{R}$, define

$$J(\mathrm{v}) = \int_{\mathbb{R}} \left(\mathrm{v}(x)(x_0 - x) + \frac{(T(x) - x_0)^2}{1 - \frac{x_0 - x}{\mathrm{v}(x)}} \right) \rho_0(x) dx.$$

We can then rewrite Problem 1 in the present case where ρ_0 and ρ_1 are continuous, have bounded convex support, and are bounded from below on the interior of their support, as follows.

Problem 3. *Consider*

$$\min_{\mathrm{v} \in V} J(\mathrm{v}) \tag{5.3}$$

over the class of functions V of Definition 5.4. Determine existence, uniqueness, and a functional form for a minimizing solution v.

To solve Problem 3, we define velocity fields v on all of $\mathbb{R}$, even outside $\mathrm{Supp}\,(\rho_0)$, as this suitably defined *extension* of v will be conveniently expressed as a solution of a differential equation. To this end, we note that the constraint (5.2) can be alternatively expressed in the form

$$C_x^{\mathrm{alt}}(\mathrm{v}) := \frac{\mathrm{v}(x)^2\rho_0(x) - h(x_0 - x)\partial_x \mathrm{v}(x)}{\mathrm{v}(x)} \leq h. \tag{5.4}$$

https://doi.org/10.1017/S0956792524000317 Published online by Cambridge University Press

This alternative formulation applies even for points x where $\rho_0 = 0$ and will help define the sought extension for $\mathrm{v}^\star$.

Let us first extend on all of $\mathbb{R}$ the optimal transport map between ρ_0 and ρ_1. To this end, define $\alpha_0 = \inf \mathrm{Supp}\,(\rho_0)$, $\beta_0 = \sup \mathrm{Supp}\,(\rho_0)$, $\alpha_1 = \inf \mathrm{Supp}\,(\rho_1)$, $\beta_1 = \sup \mathrm{Supp}\,(\rho_1)$, and set

$$T^+(x) = \begin{cases} T(x) & \text{when } x \in \mathrm{Supp}\,(\rho_0) \\ \beta_1 + x - \beta_0 & \text{when } x \geq \beta_0 \\ \alpha_1 + x - \alpha_0 & \text{when } x \leq \alpha_0. \end{cases}$$

Let $\gamma_0, \gamma_1 \in \mathbb{R}$ be the uniquely defined points such that $\frac{x_0-\alpha_0}{\mathrm{v}^\star(\alpha_0)} = \frac{x_0-\gamma_0}{T^+(\gamma_0)-\gamma_0}$ and $\frac{x_0-\beta_0}{\mathrm{v}^\star(\beta_0)} = \frac{x_0-\gamma_1}{T^+(\gamma_1)-\gamma_1}$. The point γ_1 is the point that, when transported by $T - \mathrm{Id}$, crosses the toll at the same time and β_0 crosses the toll when being transported by $\mathrm{v}^\star$. Note that we have $\gamma_0 \leq \alpha_0$ and $\gamma_1 \geq \beta_0$. We also extend $\mathrm{v}^\star$ on the whole $\mathbb{R}$ as

$$\mathrm{v}^{\star+}(x) = \begin{cases} T(x) - x & \text{when } x \leq \gamma_0 \text{ or } x \geq \beta_0 \\ \mathrm{v}^\star(\alpha_0)\frac{x_0-x}{x_0-\alpha_0} & \text{when } \gamma_0 \leq x \leq \alpha_0 \\ \mathrm{v}^\star(x) & \text{when } x \in \mathrm{Supp}\,(\rho_0) \\ \mathrm{v}^\star(\beta_0)\frac{x_0-x}{x_0-\beta_0} & \text{when } \beta_0 \leq x \leq \gamma_1 \end{cases}$$

For notational simplicity, in the sequel, we suppress the labelling on T^+, $\mathrm{v}^{\star+}$ and use T, $\mathrm{v}^\star$ instead for the extended versions as well. To build $\mathrm{v}^\star$, we first establish that on the points where $T - \mathrm{Id}$ does not satisfy the constraint, $\mathrm{v}^\star$ actually saturates the constraint. As an immediate consequence of Proposition 5.2, we have the following lemma:

Lemma 5.5. *For ρ_0 and ρ_1 satisfying Assumptions 2, for all $x \in \mathrm{Supp}\,(\rho_0)$ such that $C_x(T - \mathrm{Id}) > h$ we have $C_x(\mathrm{v}^\star) = h$.*

We next characterize a leading segment of the distribution corresponding to points with velocity faster than that of the optimal unconstrained transport. It is essential that the leading edge ‘speeds up’ to allow the trailing portion to pass through and meet the time constraint. Specifically, we show that $\mathrm{v}^\star$ is greater than $T - \mathrm{Id}$ at the points to the right of points where $T - \mathrm{Id}$ does not satisfy the constraint.

Lemma 5.6. *For ρ_0 and ρ_1 satisfying Assumptions 2, for $x_1 = \sup\{x \in \mathrm{Supp}\,(\rho_0) \mid C_x(T - \mathrm{Id}) > h\}$ and $y_1 = \sup\{x \in \mathbb{R} \mid C_x^{\mathrm{alt}}(\mathrm{v}^\star) = h\}$ we have that $\forall x \in (x_1, y_1)$, $\mathrm{v}^\star(x) \geq T(x) - x$.*

Proof. First note that $y_1 \geq x_1$ by Lemma 5.5. Suppose that

$$\{\mathrm{v}^\star(x) < T(x) - x\} \cap (x_1, y_1) \neq \emptyset$$

and let $a = \sup\{x \in (x_1, y_1) \mid \mathrm{v}^\star(x) < T(x) - x\}$. We consider separately the two cases $\rho_0(a) = 0$ and $\rho_0(a) > 0$ below:

(i) If $\rho_0(a) = 0$ then $\forall x \geq a,\ \rho_0(x) = 0$, so

$$\mathrm{v}^\star(x) = \frac{T(y_1) - y_1}{y_1 - x_0}(x - x_0),$$

as $C_x^{\mathrm{alt}}(\mathrm{v}^\star) = h$ for all x$\in [a, y_1]$. Furthermore, $T(a) - a = \beta_1 + a - \beta_0 - a = T(y_1) - y_1$ and $T(a) - a = \mathrm{v}^\star(a)$ so necessarily $a = y_1$ and $\mathrm{Supp}\,(\rho_0) = [\alpha_0, y_1]$. Then, $\exists z \in (x_1, y_1)$ such that, $\rho_0(z) > 0$, $\mathrm{v}^\star(z) < T(z) - z$, and $\partial_x \mathrm{v}^\star(z) > T'(z) - 1$.

(ii) If $\rho_0(a) > 0$, then by convexity of $\mathrm{Supp}\,(\rho_0)$ we also have existence of that $z \in (x_1, y_1)$ with the same properties. In both cases, we have

$$\frac{\mathrm{v}^\star(z)\rho_0(z)}{1 + \frac{x_0-z}{\mathrm{v}^\star(z)}\partial_x \mathrm{v}^\star(z)} < \frac{(T(z) - z)\rho_0(z)}{1 + \frac{x_0-z}{T(z)-z}(T'(z) - 1)} \leq h$$

which contradicts the definition of y_1. □

https://doi.org/10.1017/S0956792524000317 Published online by Cambridge University Press

The following lemma states that if $v^\star$ saturates the constraint on a maximal interval (i.e., such that, the points just outside do not saturate the constraint), then either $v^\star = T - \mathrm{Id}$ throughout, or it is strictly greater than $T - \mathrm{Id}$ on a portion of the interval and strictly less than $T - \mathrm{Id}$ on another portion of the interval. This property is inherited by the convexity of the cost.

Lemma 5.7. *Suppose that ρ_0 and ρ_1 satisfy Assumptions 2 and let be $[a, b] \subset \{x \in \mathbb{R} \mid C_x^{\mathrm{alt}}(v^\star) = h\}$ with $[a, b]$ of maximal size. Then, $[a, b]$ has the following properties:*

- *There exists $x \in [a, b]$ such that $v^\star(x) > T(x) - x$ if and only if $\exists y \in [a, b]$ such that $v^\star(y) < T(y) - y$.*
- *If for all $\delta > 0$, there exists $y_1 \in (a, a + \delta)$ such that $v^\star(y_1) > T(y_1) - y_1$ then for all $\delta > 0$, there also exists $y_2 \in (a, a + \delta)$ such that $v^\star(y_2) < T(y_2) - y_2$.*
- *If for all $\delta > 0$, there exists $y_1 \in (b - \delta, b)$ such that $v^\star(y_1) < T(y_1) - y_1$ then for all $\delta > 0$, there also exists $y_2 \in (b - \delta, b)$ such that $v^\star(y_2) > T(y_2) - y_2$.*

Proof. Suppose that $\forall x \in [a, b]$, we have $v^\star(x) \geq T(x) - x$ and we do not have equality on the whole interval. Define

$$\Psi_a(\epsilon) = \int_a^b \Big((x_0 - x)(v^\star(x) + \epsilon) + \frac{(T(x) - x_0)^2}{1 - \frac{x_0 - x}{v^\star(x) + \epsilon}}\Big)\rho_0(x)dx.$$

Then, we have $\partial_x \Psi_a(0) = \int_a^b (x_0 - x)(1 - \frac{(T(x)-x_0)^2}{(v^\star(x)-(x_0-x))^2})\rho_0(x)dx > 0$. Let be $c < a$ such that $\partial_x \Psi_c(0) > 0$ and $\exists \delta > 0$ with $\frac{v^\star(c)^2\rho_0(c) - hv^\star(c)}{h(x_0 - c)} - \partial_x v^\star(c) = -\delta$. Then, there exist $d \in (c, a)$ such that $\partial_x \Psi_d(0) > 0$ and $\exists \delta > 0$ with $\frac{v^\star(d)^2\rho_0(d) - hv^\star(d)}{h(x_0 - d)} - \partial_x v^\star(d) = -\delta/2$. Let us define k_ϵ as the function solving the ODE

$$\begin{cases} \partial_x k_\epsilon(x) = -\partial_x v^\star(x) + \frac{(v^\star(x)+k_\epsilon)^2\rho_0(x) - h(v^\star(x)+k_\epsilon(x))}{h(x_0 - x)} & \text{for } x \leq d, \\ k_\epsilon(d) = -\epsilon. \end{cases}$$

Then for $\epsilon > 0$ small enough, we have $\partial_x k_\epsilon(x) < -\delta/4$, $\forall x \in (c, d)$. Therefore, for $\epsilon > 0$ small enough $\exists y \in (c, d)$ such that $k_\epsilon(y) = 0$. Define

$$v_\epsilon(x) = \begin{cases} v^\star(x) & \text{if } x \notin (y, b), \\ v^\star(x) - \epsilon & \text{if } x \in (d, b), \\ v^\star(x) + k_\epsilon(x) & \text{if } x \in (y, d]. \end{cases}$$

Then for $\epsilon > 0$ small enough, v_ϵ verifies the constraint and $J(v_\epsilon) < J(v^\star)$.

Furthermore, recalling that the speed of a particle at $x \in \mathrm{Supp}(\rho_0)$ after having crossed the toll is $\frac{T(x)-x_0}{1-\frac{x_0-x}{v^\star(x)}}$, then for $\overline{v}^\star : \mathrm{Supp}(\rho_1) \to \mathbb{R}_{<0}$ defined by

$$\overline{v}^\star(y) = \frac{x_0 - y}{1 - \frac{x_0 - T^{-1}(y)}{v^\star(T^{-1}(y))}},$$

we have that $\overline{v}^\star$ is a solution to the problem of transfering between ρ_1 and ρ_0. Therefore, applying the previous derivations to $\overline{v}^\star$, we deduce that for $[T(a), T(b)] \subset \{x \in \mathbb{R} \mid C_x^{\mathrm{alt}}(\overline{v}^\star) = h\}$ with $[T(a), T(b)]$ of maximal size, if there exists $y_1 \in [T(a), T(b)]$ with $\overline{v}^\star(y_1) < T^{-1}(y_1) - y_1$, then there exists $y_2 \in [T(a), T(b)]$ with $\overline{v}^\star(x) > T^{-1}(y_2) - y_2$. Now as we have the relation $v^\star(x) < T(x) - x$ iff $\overline{v}^\star(T(x)) < -(T(x) - x)$, we deduce that if there exists $x_1 \in [a, b]$ such that $v^\star(x_1) < T(x_1) - y_1$ then there exists $x_2 \in [a, b]$ such that $v^\star(x_2) < T(x_2) - y_2$.

Let us now prove the second and third point. Suppose that there exists $\delta > 0$ and $y_1 \in (a, a + \delta)$ such that $v^\star(y_1) > T(y_1) - y_1$ and $\forall x \in (a, a + \delta)$, we have $v^\star(x) \geq T(x) - x$. Then defining

$$\Psi_a(\epsilon) = \int_a^{a+\delta} \Big((x_0 - x)(v^\star(x) + \epsilon) + \frac{(T(x) - x_0)^2}{1 - \frac{x_0 - x}{v^\star(x) + \epsilon}}\Big)\rho_0(x)dx,$$

https://doi.org/10.1017/S0956792524000317 Published online by Cambridge University Press

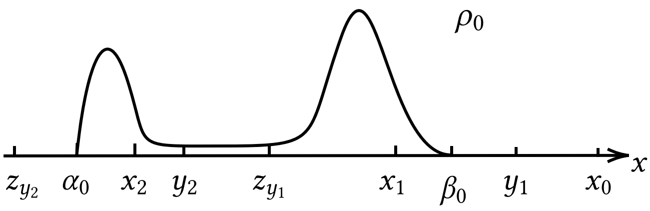

Figure 2. *Density $\rho_0(x)$ vs. x.*

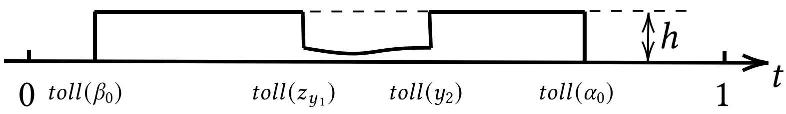

Figure 3. *Flux $\rho_t(x_0)v_t(x_0)$ at crossing.*

and carrying out the same derivation as previously we obtain that $v^\star$ is not optimal. Now applying the same reasoning to $\overline{v}^\star$, we obtain the third point. □

We are now in a position to build explicitly $v^\star$ using the lemmas. The process of building v^* consists of determining its value successively on intervals $[z_{y_i}, y_i]$ and $[y_{i+1}, z_{y_i}]$, with

$$\ldots > y_i > z_{y_i} > y_{i+1} > z_{y_{i+1}} > \ldots$$

such that $v^\star(x) \neq T(x) - x$ for $x \in [z_{y_i}, y_i]$ a.e., while $v^\star(x) = T(x) - x$ on the complement where $x \notin \bigcup_i [z_{y_i}, y_i]$. By Proposition 5.2, we know that $C_x^{\text{alt}}(v^\star) = h$ on intervals $[z_{y_i}, y_i]$, a fact that will help us determine $v^\star$ and the succession of points that define these intervals.

We explain the process in Figures 2-4 with an example. This example presents a situation where the behaviour of the corresponding optimal solution $v^\star$ is characterized by two distinct intervals $[z_{y_i}, y_i]$ $i = 1, 2$, where the constraint saturates. Thus, for this example, we identify three intervals of interest, $[z_{y_2}, y_2]$, $[y_2, z_{y_1}]$, and $[z_{y_1}, y_1]$. In the first and the last, the constraint saturates, whereas in the middle interval it does not. We proceed by working our way from right to left, always assuming that Supp (ρ_0) is to the left of the toll, as in the figures.

In general, the process begins by first computing the optimal transport map T, without involving the constraint. Then, we identify x_1 as the rightmost point where the throughput hits the limit set at x_0. Naturally, if the optimal transport map satisfies the throughput constraint, then it is the optimal map and specifies v^* throughout. Assuming that x_1 is finite, then a search to the right of x_1, that we explain later on, identifies y_1 as the rightmost point where v needs to be adjusted so as to abide by the throughput constraint while minimizing the transportation cost. In the example depicted in Figure 2, y_1 is shown located to the right of β_0 (= the supremum of the support of ρ_0), though this is not always the case, and depends on the terminal distribution ρ_1 via the optimization problem that specifies y_1. We choose to explain this case, where y_1 is to the right of β_0 so as to highlight that this is indeed possible.

Continuing on with our specific example, for the interval $[z_{y_1}, y_1]$, we have $v^\star = v_{y_1}$, with v_y defined in equation (5.6) explained below, which ensures that $C_x^{\text{alt}}(v) = h$. Then, on $[y_2, z_{y_1}]$ we have once again that the velocity is specified by the 'unconstrained' optimal map T, i.e, that $v^\star = T - \text{Id}$, and so $C_x^{\text{alt}}(v) = C_x^{\text{alt}}(T - \text{Id})$. Finally on $[z_{y_2}, y_2]$, we have $v^\star = v_{y_2}$ as $C_x^{\text{alt}}(v) = h$. Note that in this specific example where $y_1 \geq \beta_0$ and $z_{y_2} \leq \alpha_0$, we have $\forall x \in [z_{y_2}, \alpha_0]$, $toll(x) = toll(\alpha_0)$ and $\forall x \in [\beta_0, y_1]$, $toll(x) = toll(\beta_0)$.

We now detail how to build explicitly $v^\star$ in the general case. As noted, if $T - \text{Id}$ verifies the constraint throughout, which can now be explicitly stated as in (5.2), then $v^* = T - \text{Id}$ is the optimal solution. Otherwise define $x_1 = \sup\{x \in \text{Supp}\,(\rho_0) \mid C_x(T - \text{Id}) > h\}$, and thereby we determine $y_1 \in [x_1, x_0]$

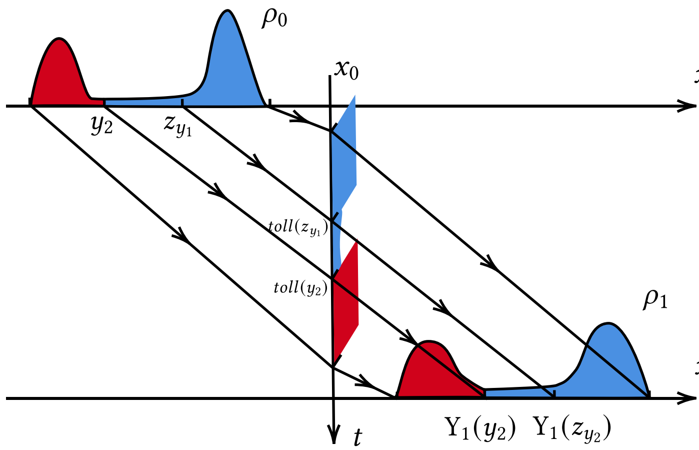

Figure 4. *Illustration of the flow through the toll. The middle segment* $[y_2, z_{y_1}]$ *transports through the toll unimpeded by the constraint towards the final destination, via the optimal transport map T, designed for unconstrained transport; each point in this interval maintains the same velocity before and after the toll. In contrast, the segments to the left and right,* $[z_{y_2}, y_2]$ *and* $[z_{y_1}, y_1]$*, respectively, are adjusted accordingly so as to saturate the constraint. The exact position of their respective end points (that may even be outside the support of* ρ_0*, as a matter of computational simplicity, in which case they correspond to zero density) is computed via the solution of an optimization problem and depend on the terminal distribution* ρ_1 *as well.*

(cf. Lemma 5.5) such that

$$y_1 = \sup\{x \in \mathbb{R} \mid C_x^{\text{alt}}(\mathrm{v}^\star) = h\}. \tag{5.5}$$

For any $y \in \mathbb{R}$ with $x_1 \leq y < x_0$, define the velocity $\mathrm{v}_y(x)$ as the solution of the differential equation

$$\begin{cases} \partial_x \mathrm{v}_y(x) = \frac{\mathrm{v}_y(x)^2 \rho_0(x) - h\mathrm{v}_y(x)}{h(x_0 - x)} & \text{for } x \leq y. \\ \mathrm{v}_y(y) = T(y) - y \end{cases} \tag{5.6}$$

Note that this equation is solved backwards, starting from a terminal condition at y. This value for the velocity ensures that the transport will saturate the constraint to left of y (i.e., $C_x^{\text{alt}}(\mathrm{v}_y) = h$ will hold for $x \leq y$). The functional form of $\mathrm{v}_y(x)$ will be used next to identify the first interval $[z_{y_1}, y_1]$, where the velocity will depart from that of the unconstrained transport T, via solving a suitable optimization problem to determine y_1. Since we know that the equality $C_x^{\text{alt}}(\mathrm{v}^\star) = h$ will be true on a certain interval $[z_{y_1}, y_1]$, on that interval we will have $\mathrm{v}^\star = v_{y_1}$.

Let $w_y^{x_1} = \inf\{x \leq x_1 \mid \forall s \in (x, x_1), \mathrm{v}_y(s) \geq T(s) - s\}$ (well defined by Lemma 5.6) and

$$z_y^{x_1} = \inf\{x \leq w_y^{x_1} \mid \forall s \in (x, w_y^{x_1}),\ \mathrm{v}_y(s) < T(s) - s\}. \tag{5.7}$$

Then, we have that $\mathrm{v}^\star(x) = \mathrm{v}_{y_1}(x)$, $\forall x \in (z_{y_1}^{x_1}, x_1)$ by Lemma 5.7 and Proposition 5.2.

We now determine y_1 by solving a suitable optimization problem. For $x \leq y < x_0$, define

$$J_x(y) = \int_{\mathbb{R}} \Big((T(s) - s)^2 \mathbb{1}\{s \notin (z_y^x, y)\} + ((x_0 - s)\mathrm{v}_y(s) + \frac{(T(s) - x_0)^2}{1 - \frac{x_0 - s}{\mathrm{v}_y(s)}}) \mathbb{1}\{s \in (z_y^x, y)\} \Big) \rho_0(s) ds.$$

We have $J_x(y) = J(\mathrm{v}_y^+)$ for the function v_y^+ such that $\mathrm{v}_y^+ = \mathrm{v}_y$ on $[z_y^x, y]$ and $\mathrm{v}_y^+ = T - \mathrm{Id}$ on $\mathbb{R} \setminus [z_y^x, y]$. Then, the first step of the building process of $\mathrm{v}^\star$ is to find y_1 solution of

$$y_1 = \arg\min_{y \geq x_1} J_{x_1}(y).$$

Such a y_1 is well defined as J_{x_1} is continuous on $[x_1, x_0]$. Once y_1 has been determined, we define $x_2 = \sup\{x < z_{y_1} \mid C_x(T - \mathrm{Id}) > h\}$. If x_2 is not defined then

$$\mathrm{v}^\star(x) = \begin{cases} \mathrm{v}_{y_1}(x) & \text{if } x \in (z_1, y_1), \\ T(x) - x & \text{if } x \notin (z_1, y_1), \end{cases}$$

otherwise we start again the same process to determine y_2 as

$$y_2 = \arg\min_{y > x_2} J_{x_2}(y).$$

If $y_2 < z_{y_1}$, it suggests that there is an interval $[y_2, z_{y_1}]$ where the transport follows the unconstrained map T, and we continue in the same way.

However, it is possible that the condition $y_i \leq z_{y_{i-1}}$ fails at some point, for some $i \geq 2$. In that case, intervals where the velocity departs from being $T(x) - x$, will merge. For instance, if we obtain $y_i > z_{y_{i-1}}$ then as $(y, y') \mapsto J_{x_{i-1}}(y) + J_{x_i}(y')$ is convex on $\{(y, y') \mid y' \leq z_y\}$, it means that $C_x^{\mathrm{alt}}(\mathrm{v}^\star) = h$, $\forall x \in (x_i, x_{i-1})$, and therefore, we have to start the optimization again and determine y_{i-1} as

$$y_{i-1} = \arg\min_{y > x_{i-1}} J_{x_i}(y).$$

If we obtain a value $y_{i-1} > z_{y_{i-2}}$, we reset x_{i-1} as being equal to x_i and, once again, we have to redetermine

$$y_{i-2} = \arg\min_{y > x_{i-2}} J_{x_{i-1}}(y).$$

Otherwise, i.e., if we obtain a value $y_{i-1} \leq z_{y_{i-2}}$, we reset x_i as $x_i = \sup\{x < z_{y_{i-1}} \mid C_x(T - \mathrm{Id}) > h\}$ for this updated value y_{i-1}. Once again, if x_i is well defined we continue the process by finding

$$y_i = \arg\min_{y > x_i} J_{x_i}(y).$$

We continue this iterative process until $\mathrm{v}^\star$ is defined on all of the support of ρ_0. The construction process is summarized in Algorithm 1.

To establish the correctness of this algorithm in constructing the solution $\mathrm{v}^\star$, we start with a key property of the function J_x

Proposition 5.8. *Let be* $y \in \mathrm{Supp}\,(\rho_0)$ *and* $x = \sup\{z \leq y | C_z(T - \mathrm{Id}) > h\}$. *There exists* $y^\star \in [x, x_0]$ *such that the function* $y \mapsto J_x(y)$ *is non-increasing on* $[x, y^\star]$ *and then non-decreasing on* $[y^\star, x_0]$. *Furthermore, if* $C_z(T - Id) < h$ *then* $\forall \epsilon > 0$, $y \mapsto J_x(y)$ *is not constant on* $[z - \epsilon, z + \epsilon]$.

Proof. Let v_y be the solution of (5.6) and its associated toll function $toll_y(x) = (x_0 - x)/\mathrm{v}_y(x)$. We have

$$\partial_x toll_y(x) = -\frac{1}{\mathrm{v}_y(x)} - \frac{x_0 - x}{\mathrm{v}_y(x)^2}\partial_x \mathrm{v}_y = -\frac{1}{\mathrm{v}_y(x)} - \frac{1}{h}(\rho_0(x) - \frac{h}{\mathrm{v}_y(x)}) = -\frac{\rho_0(x)}{h},$$

so in particular for $x \leq y$ we have

$$toll_y(x) = toll(y) + \rho_0([x, y])/h,$$

https://doi.org/10.1017/S0956792524000317 Published online by Cambridge University Press

Algorithm 1. Building Process

$x_1 \leftarrow sup\{x \in \text{Supp}\,(\rho_0) \mid C_x(T - \text{Id}) > h\}$
$z_0 \leftarrow x_0$
$i \leftarrow 1$
while $x_i > \min \text{Supp}\,(\rho_0)$ **do**
  $y_i \leftarrow \arg\ \min_{y \geq x_i} J_{x_i}(y)$
  while $y_i > z_{i-1}$ **do**
    $y_{i-1} \leftarrow \arg\ \min_{y \geq x_{i-1}} J_{x_i}(y)$
    $x_{i-1} \leftarrow x_i$
    $i \leftarrow i - 1$
  end while
  $w_{y_i}^{x_i} \leftarrow \inf\{x \leq x_i \mid \forall s \in (x, x_i), \mathrm{v}_y(s) \geq T(s) - s\}$
  $z_i \leftarrow \inf\{x \leq w_{y_i}^{x_i} \mid \forall s \in (x, w_{y_i}^{x_i}),\ \mathrm{v}_y(s) < T(s) - s\}$
  $x_{i+1} \leftarrow \sup\{x \leq z_i \mid C_x(T - \text{Id}) > h\}$
  $i \leftarrow i + 1$
end while
return $(z_k, y_k)_{k=1,\ldots,i-1}$

with $toll(y) = (x_0 - y)/(T(y) - y)$. Then as $\mathrm{v}_y(x) = (x_0 - x)/toll_y(x)$, we can rewrite J_x as

$$\begin{aligned} J_x(y) &= \int_{\mathbb{R}} \Big((T(s) - s)^2 \mathbb{1}\{s \notin (z_y^x, y)\} \\ &\qquad + (\frac{(x_0 - s)^2}{toll(y) + \rho_0([s, y])/h} + \frac{(T(s) - x_0)^2}{1 - toll(y) - \rho_0([s, y])/h}) \mathbb{1}\{s \in (z_y^x, y)\} \Big) \rho_0(s) ds. \\ &= \int_{\mathbb{R}} \Big((T(s) - s)^2 \mathbb{1}\{s \notin (z_y^x, y)\} + (f_1(s, y) + f_2(s, y)) \mathbb{1}\{s \in (z_y^x, y)\} \Big) \rho_0(s) ds, \end{aligned}$$

with

$$f_1(s, y) = \frac{(x_0 - s)^2}{toll(y) + \rho_0([s, y])/h} \text{ and } f_2(s, y) = \frac{(T(s) - x_0)^2}{1 - toll(y) - \rho_0([s, y])/h}.$$

Now as $f_1(z_y^x, y) + f_2(z_y^x, y) = (T(z_y^x) - z_y^x)^2$ and $f_1(y, y) + f_2(y, y) = (T(y) - y)^2$, we deduce that

$$\begin{aligned} \partial_y J_x(y) &= \int_{\mathbb{R}} \Big(\partial_y f_1(s, y) + \partial_y f_2(s, y) \Big) \mathbb{1}\{s \in (z_y^x, y)\} \rho_0(s) ds \\ &= - \int_{\mathbb{R}} \Bigg(\left(\frac{(x_0 - s)}{toll(y) + \rho_0([s, y])/h} \right)^2 (toll'(y) + \rho_0(y)/h) \\ &\qquad - \left(\frac{(T(s) - x_0)}{1 - toll(y) - \rho_0([s, y])/h} \right)^2 (toll'(y) + \rho_0(y)/h) \Bigg) \mathbb{1}\{s \in (z_y^x, y)\} \rho_0(s) ds \\ &= \int_{\mathbb{R}} \left(\left(\frac{(T(s) - x_0)}{1 - toll(y) - \rho_0([s, y])/h} \right)^2 - \left(\frac{(x_0 - s)}{toll(y) + \rho_0([s, y])/h} \right)^2 \right) \\ &\qquad \times (toll'(y) + \rho_0(y)/h) \mathbb{1}\{s \in (z_y^x, y)\} \rho_0(s) ds. \end{aligned}$$

Furthermore, for $s \leq y$, the function $y \mapsto toll(y) + \rho_0([s, y])/h = toll_y(s)$ is non-increasing as all the points $x \in (s, y)$ verify that $C_x(\mathrm{v}_y) = h$ i.e., the flow rate is maximized between s and y. We can then

https://doi.org/10.1017/S0956792524000317 Published online by Cambridge University Press

conclude that $toll'(y) + \rho_0(y)/h \leq 0$ and that the function

$$y \mapsto \left(\frac{(T(s) - x_0)}{1 - toll(y) - \rho_0([s, y])/h}\right)^2 - \left(\frac{(x_0 - s)}{toll(y) + \rho_0([s, y])/h}\right)^2$$

is non-negative and then non-positive, which gives the result.

Now if for $z \geq x$, $C_z(T - \mathrm{Id}) < h$, then $toll'(z) + \rho_0(z)/h < 0$ so $J_x(y)$ is not constant around z. □

Let now

$$n = \sup_{(x_i)_i} \sum_{i=1}^{\infty} \mathbb{1}_{\{C_{x_i}(T-Id)>h, C_{x_{i+1}}(T-id)<h\}}, \tag{5.8}$$

where the supremum is taken over all increasing sequence $(x_i)_i \in \mathrm{Supp}\,(\rho_0)^{\mathbb{N}}$. It corresponds to the number of times the function $x \mapsto C_x(T - \mathrm{Id})$ oscillates around the value h. We are going to prove that the problem is reduced to finding the intervals where $C_x(\mathrm{v}^\star) = h$. To this end, let us define a class of candidate minimizers that are characterized by the intervals where they saturate the constraint.

Definition 5.9. *For $N \in \mathbb{N}$, the class of functions V_N is defined as the set of maps $\mathrm{v}^{(z_i, y_i)_i} : \mathrm{Supp}\,(\rho_0) \to \mathbb{R}$ such that*

i) there exists a sequence of disjoints intervals $(z_i, y_i) \subset \mathrm{Supp}\,(\rho_0)$ $i = 1, \ldots, N$ with

$$\mathrm{v}^{(z_i, y_i)_i}(x) = \sum_i \mathbb{1}\{x \in (z_i, y_i)\} v_{y_i}(x) + \mathbb{1}\{x \notin \bigcup_i (z_i, y_i)\}(T(x) - x),$$

ii) for all $x \notin \bigcup_i (z_i, y_i)$, we have $C_x(T - Id) \leq h$,

iii) for all $i \in \{1, \ldots, N\}$ there exists $x \in (z_i, y_i)$ such that $C_x(T - \mathrm{Id}) > h$.

Let us prove that the solution $v^\star$ belongs to a certain class V_N for $N \leq n + 1$.

Lemma 5.10. *Suppose that ρ_0 and ρ_1 satisfy Assumptions 2 and that n defined in (5.8) is finite. There exists $N \in \{0, \ldots, n + 1\}$ such that Problem 3 is equivalent to $\inf_{v \in V_N} J(v)$ and the infimum is attained.*

Proof. First note that we have that $\inf_{v \in V_N} J(v) \geq \min_{\mathrm{v} \in V} J(\mathrm{v})$ as every candidate in V_N defines a candidate in V. Let be

$$E = \{x \in \mathrm{Supp}\,(\rho_0) |\ C_x(T - Id) > h\}.$$

Then as E is open, it can be written as a union of disjoint intervals : $E = \bigcup_{k \geq 1} (a_k, b_k)$. For all $k \geq 1$, let be

$$c_k = \sup\{x \leq a_k |\ C_x(T - Id) < h\} \text{ and } d_k = \inf\{x \geq b_k |\ C_x(T - Id) < h\}.$$

Then, the set

$$A = \bigcup_{k \geq 1} (c_k, d_k),$$

can actually be written as a finite union:

$$A = \bigcup_{k=1}^{m} (\alpha_k, \beta_k),$$

with $m \leq n + 1$ and α_k, β_k points belonging to $\partial\{x \in \mathrm{Supp}\,(\rho_0) |\ C_x(T - Id) < h\}$. Then for all $k \in \{1, \ldots, m\}$ and $x \in (\alpha_k, \beta_k)$, we have $C_x(\mathrm{v}^\star) = h$ so in particular, $\mathrm{v}^\star = \mathrm{v}_{\beta_k}$ on (α_k, β_k). Let be $Y_k = \sup\{x \geq \beta_k | \forall y \in (\alpha_k, x), \mathrm{v}^\star(y) = \mathrm{v}^\star_x(y)\}$ and $(y_k)_{k=1,\ldots,N}$ (with $N \leq m$) the decreasing sequence such that $\{y_k |\ k = 1, \ldots, N\} = \{Y_k |\ k = 1, \ldots, m\}$. Likewise, define $Z_k = \inf\{x \leq \alpha_k | \forall y \in (x, \beta_k), \mathrm{v}^\star(y) = \mathrm{v}^\star_{\beta_k}(y)\}$ and $(z_k)_{k=1,\ldots,N}$ the decreasing sequence such that $\{z_k |\ k = 1, \ldots, N\} = \{Z_k |\ k = 1, \ldots, m\}$. We have by construction that $\mathrm{v}^\star = \mathrm{v}^{(z_k, y_k)_{k=1,\ldots,N}}$ and that $\mathrm{v}^\star$ belongs to V_N. □

Using this result, we can now prove that when n is finite, Algorithm 1 builds the solution $\mathrm{v}^\star$.

Theorem 2. *Suppose that ρ_0 and ρ_1 satisfy Assumptions 2 and that n defined in (5.8) is finite. There exist $N \in \mathbb{N}$ and a sequence of disjoints intervals $\big((z_i, y_i)\big)_i \subset \operatorname{Supp}(\rho_0)^N$ such that for all $i \in \{1, \dots, N\}$ we have the following properties.*

- *(a) For $x \in [z_i, y_i]$, we have $\mathrm{v}^\star(x) = \mathrm{v}_{y_i}(x)$ (defined in (5.6)) and $C_x^{\mathrm{alt}}(\mathrm{v}^\star) = h$.*
- *(b) There exists $w_i \in (z_{y_i}, y_i)$ such that $\forall y \in [w_i, y_i)$, $\mathrm{v}^\star(y) \geq T(y) - y$ and $\forall y \in (z_{y_i}, w_i]$, $\mathrm{v}^\star(y) \leq T(y) - y$.*
- *(c) For all $x \notin \bigcup_{i=1}^N [z_i, y_i]$, $\mathrm{v}^\star(x) = T(x) - x$.*
- *(d) The sequence $\big((z_i, y_i)\big)_i$ is the output of Algorithm 1.*

Proof. We first build the intervals satisfying the points (a), (b) and (c). Let be $(Z_k, Y_k)_{k=1,\dots,N}$ the disjoint intervals given by Lemma 5.10 such that $\mathrm{v}^\star = \mathrm{v}^{(Z_k,Y_k)_{k=1,\dots,N}}$. From the definition of $\mathrm{v}^{(Z_k,Y_k)_{k=1,\dots,N}}$, we know that $(Z_k, Y_k)_{k=1,\dots,N}$ satisfy the points (a) and (c). Let us show that we can subdivide each interval (Z_k, Y_k) into a finite numbers of sub-intervals satisfying (b). Define for all $k \in \{1, \dots, N\}$,

$$\beta_k = \sup\{x \leq Y_k | C_x(T - \mathrm{Id}) > h\} \text{ and } \alpha_k = \inf\{x \geq Z_k | C_x(T - \mathrm{Id}) > h\}.$$

We first argue that we actually have $\beta_k < Y_K$. Suppose to the contrary that is not true, i.e., that $\beta_k = Y_K$. Then from the third point of Lemma 5.7, the function $\mathrm{v}^\star - (T - \mathrm{Id})$ is either strictly positive on an interval $(Y_k - \delta, Y_k)$ or it oscillates around 0. If it oscillates around 0, then there exists $x \in (Y_k - \delta, Y_k)$ such that $\mathrm{v}^\star(x) > T(x) - x$ and $\partial_x \mathrm{v}^\star(x) < T'(x) - 1$. Therefore, we have that

$$\frac{\mathrm{v}(x)\rho_0(x)}{1 + \frac{x_0 - x}{\mathrm{v}(x)}\partial_x \mathrm{v}(x)} > \frac{(T(x) - x)\rho_0(x)}{1 + \frac{x_0 - x}{T(x) - x}(T'(x) - 1)} \geq h, \tag{5.9}$$

which is impossible, so we deduce that $\mathrm{v}^\star - (T - \mathrm{Id})$ is strictly positive on an interval $(Y_k - \delta, Y_k)$. But again, as $\mathrm{v}^\star(Y_k) = T(Y_k) - Y_k$, we deduce that there exists $x \in (Y_k - \delta, Y_k)$ such that $\partial_x \mathrm{v}^\star(x) < T'(x) - 1$, which implies that $\mathrm{v}^\star$ does not satisfies the constraint at x. As this is not possible, we deduce that $\beta_k < Y_k$. Then, we also deduce from Lemma 5.6 that $\mathrm{v}^\star(x) \geq T(x) - x$ for all $x \in (\beta_k, Y_k)$. In the same way, we can prove that $\alpha_k > Z_k$ and $\mathrm{v}^\star \leq T - \mathrm{Id}$ on (Z_k, α_k).

Therefore, the interval (Z_k, Y_K) can be subdivided into a finite number of intervals $(z_{k,i}, y_{k,i})_{i=1,\dots,n_k}$ that verify (a) and (b) of the theorem. The number of sub-intervals is finite as the function $x \mapsto v^\star(x) - T(x) - x$ cannot oscillate around 0 at the point where $C_x(T - \mathrm{Id}) \geq h$ (same arguments as (5.9)). For points where $C_x(T - \mathrm{Id}) < h$, there exists $\delta > 0$ such that for all $y \in (x - \delta, x + \delta)$ we have $C_y(T - \mathrm{Id}) < h$, so $x \mapsto v^\star(x) - T - x$ cannot oscillate around 0, otherwise there would be points where $C_y(\mathrm{v}^\star) < h$ which is not possible as $\mathrm{v}^\star$ saturates the constraint on the interval. Now the sequence of intervals $(z_{k,i}, y_{k,i})_{i=1,\dots,n_k,k=1,\dots,N}$ is our desired sequence that verifies (a), (b) and (c) of the theorem.

Let us now prove (d) by induction. For $a, b \in \operatorname{Supp}(\rho)$, define $P_{a,b}$ the solution to Problem 3 but for the measure $\rho_{0|_{[a,b]}}$ and $\rho_{1|_{[T(a),T(b)]}}$. We prove that each time Algorithm 1 checks the condition of the 'while loop' at line 4, the current solution $(z_k, y_k)_{k=1,\dots,i-1}$ in the memory of the algorithm is optimal on the interval $[z_{i-1}, x_0]$ (i.e., $\mathrm{v}^{(z_k,y_k)_{k=1,\dots,i-1}} = P_{z_{i-1},x_0}$). This is true at the first iteration as $i = 1$ and $z_0 = x_0$.

Now fix $i \geq 2$ and suppose that after a number of iterations of the 'while loop' at line 4, for each $m \in \{1, \dots, i\}$, the solution $(z_k, y_k)_{k=1,\dots,m-1}$ is optimal on the interval $[z_{m-1}, x_0]$ (i.e., $\mathrm{v}^{(z_k,y_k)_{k=1,\dots,m-1}} = P_{z_{m-1},x_0}$). Compute line 5: $y_i = \arg\min_{y \geq x_i} J_{x_i}(y)$ and suppose that $y_i < z_{i-1}$. We have

$$J(P_{z_i,x_0}) = \min_{j=0,\dots,N} \min_{(\tau_k,\delta_k)_{k=1,\dots,j}} \left\{ J(\mathrm{v}^{(\tau_k,\delta_k)_{k=1,\dots,j}}) \mid C_x(\mathrm{v}^{(\tau_k,\delta_k)_{k=1,\dots,j}}) \leq h, \;\; x \in (z_i, x_0) \right\} \tag{5.10}$$

$$\geq \min_{\substack{j=0,\dots,N \\ l=0,\dots,j-1}} \min_{\substack{(\tau_k,\delta_k)_{k=1,\dots,j} \\ x_i \in (\tau_l,\delta_l)}} \left\{ J(\mathrm{v}^{(\tau_k,\delta_k)_{k=1,\dots,j}}) \;\middle|\; \begin{array}{l} C_x(\mathrm{v}^{(\tau_k,\delta_k)_{k=1,\dots,l}}) \leq h, \;\; x \in (z_{i-1}, x_0) \\ C_x(\mathrm{v}^{(\tau_k,\delta_k)_{k=l+1,\dots,j}}) \leq h, \;\; x \in (z_i, \delta_{l+1}) \end{array} \right\}. \tag{5.11}$$

The first equality is given by Lemma 5.10. Then, the inequality is due to the fact that every candidate minimizer $\mathrm{v}^{(\tau_k,\delta_k)_{k=1,\dots,j}}$ to (5.10) is a candidate minimizer to (5.11). As the solution to (5.11) is $\mathrm{v}^{(z_k,y_k)_{k=1,\dots,i}}$ which actually verifies $C_x(\mathrm{v}^{(z_k,y_k)_{k=1,\dots,i}}) \leq h, \forall x \in (z_i, x_0)$, we then deduce that $\mathrm{v}^{(z_k,y_k)_{k=1,\dots,i}}$ is equal to P_{z_i,x_0}.

https://doi.org/10.1017/S0956792524000317 Published online by Cambridge University Press

Let us now treat the case where $y_i \geq z_{i-1}$. We prove that the solution P_{z_i,x_0} verifies that $C_x(P_{z_i,x_0}) = h$ for all $x \in (x_i, z_{i-1})$. By contradiction, suppose that it is not the case and let be $\alpha = \sup\{x \geq x_i | C_x(P_{z_i,x_0}) = h\}$. Then, there exists $\beta \in (\alpha, z_{i-1})$ such that $C_\beta(P_{z_i,x_0}) < h$. From Proposition 5.8, we have that as y_i is the minimum of J_{x_i} and that $\beta < y_i$, then $y \mapsto J_{x_i}(y)$ is decreasing at β. Therefore, as $\alpha < \beta$ we deduce that there exists $\gamma \in (\beta, y_i)$ such that the function $y \mapsto P_{z_i,x_0}(y)\mathbb{1}_{\{y \notin (\alpha,\gamma)\}} + v_\gamma(y)\mathbb{1}_{\{y \in (\alpha,\gamma)\}}$ would be a better minimizer than P_{z_i,x_0}, which should not be possible. We therefore have

$$P_{z_i,x_0} \in \arg\ \min_{j=1,\ldots,i-1} \min_{\substack{(\tau_k,\delta_k)_{k=1,\ldots,j} \\ (x_i,z_{i-1}) \subset (\tau_j,\delta_j)}} \{ J(\mathrm{v}^{(\tau_k,\delta_k)_{k=1,\ldots,j}}) \mid C_x(\mathrm{v}^{(\tau_k,\delta_k)_{k=1,\ldots,j}}) \leq h\}.$$

We now compute the line 7 of the algorithm: $y_{i-1} = \arg\ \min_{y \geq x_{i-1}} J_{x_i}(y)$. We can apply the same reasoning whether $y_{i-1} > z_{i-2}$ or not and obtain by induction that once line 12 of the algorithm has been executed, we indeed have $P_{z_i,x_0} = \mathrm{v}^{(z_k,y_k)_{k=1,\ldots,i}}$. □

Next, we treat the case $n = \infty$. We first prove a stability result with respect to the constant h.

Lemma 5.11. *Suppose that ρ_0 and ρ_1 satisfy Assumptions 2. Then, there exists a constant $C > (h-1)^{-1}$ such that for all $\epsilon \in (0, C^{-1})$, writing $\mathrm{v}^\star_\epsilon$ the solution to Problem 3 for the flux constraint at $h - \epsilon$, we have*

$$J(\mathrm{v}^\star_\epsilon) \leq J(\mathrm{v}^\star) + C\epsilon.$$

Proof. Writing $\delta = \sup_{x \in \mathrm{Supp}\,(\rho_0)} toll_{\mathrm{v}^\star}(x)$, as $\mathrm{v}^\star$ is the solution to Problem 3, we have that $\delta < 1$. Suppose that

$$\epsilon \leq \min\{\frac{h(1-\delta)}{2}, h-1\}$$

and define $v_\epsilon = (1 - \frac{\epsilon}{h})\mathrm{v}^\star$. Then we have for all $x \in \mathrm{Supp}\,(\rho_0)$,

$$C_x(v_\epsilon) = (1 - \frac{\epsilon}{h})C_x(\mathrm{v}^\star) \leq h - \epsilon,$$

so v_ϵ is a candidate solution to the problem with constraint $h - \epsilon$. Furthermore,

$$\begin{aligned} J(v_\epsilon) &= \int_{\mathbb{R}} \left((1 - \frac{\epsilon}{h})\mathrm{v}^\star(x)(x_0 - x) + \frac{(T(x)-x_0)^2}{1 - \frac{x_0 - x}{(1-\frac{\epsilon}{h})\mathrm{v}^\star(x)}}\right)\rho_0(x)dx \\ &\leq J(\mathrm{v}^\star) + \frac{\epsilon}{h}\int_{\mathbb{R}} \frac{(T(x)-x_0)^2(1-\frac{\epsilon}{h})\mathrm{v}^\star(x)}{((1-\frac{\epsilon}{h})\mathrm{v}^\star(x) - (x_0 - x))(\mathrm{v}^\star(x) - (x_0 - x))}\rho_0(x)dx. \\ &\leq J(\mathrm{v}^\star) + \frac{\epsilon}{h}\left(\frac{1-\delta}{2\delta}(x_0 - \sup \mathrm{Supp}\,(\rho_0))\right)^{-1} \int_{\mathbb{R}} \frac{(T(x)-x_0)^2(1-\frac{\epsilon}{h})\mathrm{v}^\star(x)}{\mathrm{v}^\star(x) - (x_0 - x)}\rho_0(x)dx. \\ &\leq J(\mathrm{v}^\star) + C\epsilon. \end{aligned}$$

□

Now if $n = \infty$, it implies that the map $x \mapsto C_x(T - \mathrm{Id})$ oscillates indefinitely around h. For $\epsilon > 0$, take U a uniform random variable on $[0, \epsilon]$. As $x \mapsto C_x(T - \mathrm{Id})$ cannot oscillate around more than a countable number of values, we have that with probability 1, $x \mapsto C_x(T - \mathrm{Id})$ oscillates only a finite number of times around U. Therefore, we can solve Problem 3 with the flux constraint $h - U$ using Theorem 2. For v_U the solution given by Algorithm 1, we have that for all $x \in \mathrm{Supp}\,(\rho_0)$,

$$C_x(v_U) \leq h - U \leq h,$$

and by Lemma 5.11,

$$J(v_U) \leq J(\mathrm{v}^\star) + C\epsilon.$$

https://doi.org/10.1017/S0956792524000317 Published online by Cambridge University Press

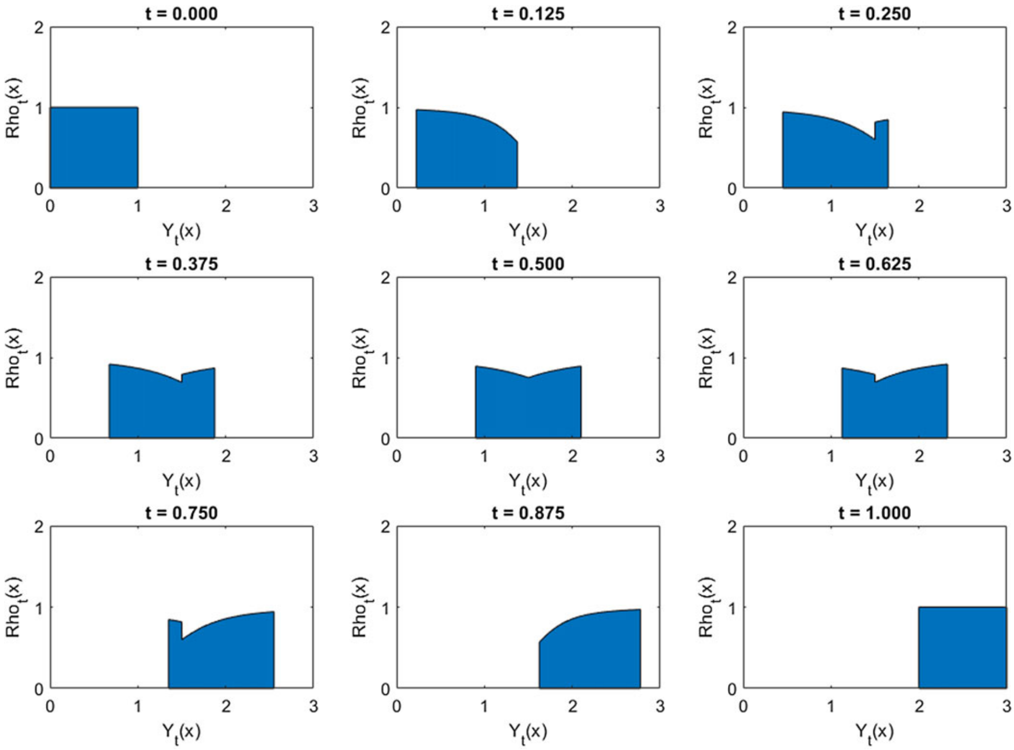

Figure 5. *Example of transporting a uniform distribution through a constriction (with $h = 1.5$) to a similar uniform terminal distribution. While the optimal unconstrained transport will preserve the shape of marginals at each time t, the flux constraint necessitates an optimal velocity that changes with x, stretching the leading edge of the distribution as it approaches the toll. Note that the snapshots of the transported distributions $Y_{t\sharp\rho_0}$ 'squeeze' while crossing the toll, and that the flow is symmetric with time.*

6. Numerical example

We provide an example to highlight the departure of the optimal transport plan through a toll with a bound on the flux, from the ideal unconstrained transport T. The example we have selected is basic, with uniform probability densities $\rho_0(x) = \mathbb{1}\{x \in [0, 1]\}$, $\rho_1(x) = \mathbb{1}\{x \in [2, 3]\}$, and a toll at $x_0 = 3/2$ with a bound h on the flux, with $1 < h \le 2$. The stringent constraint on the flux, which necessitates varying velocities so as to redistribute the mass flow as it traverses the toll, is clearly seen in the succession of distributions $Y_{t\sharp\rho_0}$ displayed in Figure 5. Evidently, these readily contrast with the unconstrained transport that pushes forward ρ_0 with constant speed giving $\rho_t(x) = \rho_0(x - 2t)$.

Specifically, with the flux constraint in place, we obtain that the optimal transport is effected by

$$Y_t(x) = \begin{cases} x + t\mathrm{v}(x) & \text{for } \ t \le toll(x) = \frac{3/2 - x}{\mathrm{v}(x)}, \\ 3/2 + (t - toll(x))g(x) & \text{for } \ t \ge toll(x). \end{cases}$$

Then, the constraint (5.2) gives that v solves the ODE

$$\frac{\mathrm{v}(x)}{1 + \frac{3/2 - x}{\mathrm{v}(x)}\partial_x \mathrm{v}(x)} = h.$$

It follows that $\mathrm{v}(x) = \frac{h(2x-3)}{2x-3+\alpha}$ for a certain value $\alpha \in \mathbb{R}$. Using the fact that the optimal solution must be symmetric in time ($\mathrm{v}(x) = g(1 - x)$) and that $g(x) = \frac{x+0.5}{1-toll(x)}$, we finally obtain that $\mathrm{v}(x) = \frac{h(2x-3)}{2x-1-h}$. Snapshots of the flow along the path from ρ_0 to ρ_1 are depicted in Figure 5.

7. Discussion and conclusion

We have presented theory for the most basic optimal transport problem in $\mathbb{R}$, through a constriction where a throughput constraint is imposed. We modelled the formulation after the standard Monge-Kantorovich optimal transport with a quadratic cost. We have shown that an optimal transport exists and is unique under general assumptions. Under some suitable assumptions on the densities to be transported to one another, we have shown explicitly how to construct the transport plan. Moreover, we have highlighted natural properties of the transport plan.

More generally, in the case where ρ_0 and ρ_1 are densities on $\mathbb{R}^d$ and that all the trajectories have to pass through a single point $x_0 \in \mathbb{R}^d$, we can readily extend the result presented as follows. For $\lambda_{\alpha S^{d-1}}$, the Lebesgue measure on the sphere of radius α and centre x_0 is defined as

$$\nu^0(\alpha) = \int_{\alpha S^{d-1}} \rho_0(x) d\lambda_{\alpha S^{d-1}}(x)$$

and ν^1 the same way. Then, the problem in $\mathbb{R}^d$ is equivalent to solving the problem in dimension 1 between the measure ν^0, ν^1 defined as $\nu^0(x) = \mathbb{1}\{x < 0\}\nu^0(-x)$ and $\nu^1(x) = \mathbb{1}\{x > 0\}\nu^1(x)$.

A significant departure from the current setting arises in the case of multiple tolls, or of a continuum of tolls, where the flux rate is bounded on a curve, surface, etc. The case where a sequence of tolls, possibly even zero-dimensional (points), where mass has to flow through all in succession, is of particular interest in engineering applications. Indeed, in the modern information age, knowledge of obstructions 'down the road' can undoubtedly be used to optimize transportation cost upstream. On the other hand, the paradigm of multiple alternative tolls that one can choose to cross is expected to have a more combinatorial flavour. Lastly, one could generalize the problem presented in this paper to transport of densities in dimension d, with a flux constraint on a measurable set with respect to the p-dimensional Hausdorff measure $\mathcal{H}^p$ (with $p \le d$). For instance, an analogous flux constraint on a measurable set $A \subset \mathbb{R}^d$ with $0 < \mathcal{H}^p(A) < \infty$ can be cast as $\forall B \subset A$ measurable with $\mathcal{H}^p(B) > 0$ and $t \in (0, 1)$

$$\forall \alpha_1, \alpha_2 \in \mathbb{R}, \quad \frac{\mathcal{H}^p(A)}{\mathcal{H}^p(B)} \int \mathbb{1}_{\{\exists \tau \in (t+\alpha_1, t+\alpha_2) \ | \ Y_\tau(x) \in B\}} \rho_0(x) dx \le h|\alpha_2 - \alpha_1|.$$

The proof of existence and uniqueness of a solution should follow using similar arguments. However, to completely characterize the behaviour of the solution as in the simpler case treated herein is expected to be considerably more challenging; one would need a finer description of how the mass distributes while traversing the toll.

Transport problems with a throughput restriction are quite natural in a variety of scientific disciplines. Of course, transportation through tolls on highways represents perhaps the most rudimentary paradigm in an engineering setting. Likewise, throughput through servers with a throughput bound is common in queuing systems. A continuum theory as envisioned herein, in higher dimensions and with multiple serial tolls, may produce useful practical insights. Finally, while fluid flow, passing through constrictions or porous media, though not directly abiding by the rigid setting of bounded throughput, could provide an idealized pertinent model in certain situations. Evidently, for an accurate model for fluid past constrictions, besides distinguishing between compressible and incompressible, throughput must be dictated by pressure, which in turn may be introduced in a suitable cost functional to be optimized for a further broadening of the general programme.

Financial support. The research has been supported in part by the AFOSR under grant FA9550-23-1-0096 and ARO under W911NF-22-1-0292.

Competing interests. The author(s) declare none.

References

[1] Ambrosio, L., Gigli, N. & Savaré, G. (2005) *Gradient Flows: In Metric Spaces and in the Space of Probability Measures*, Springer Science & Business Media, Birkhäuser Verlag, Basel – Boston – Berlin.

[2] Bredies, K., Carioni, M., Fanzon, S. & Romero, F. (2023) A generalized conditional gradient method for dynamic inverse problems with optimal transport regularization. *Found. Comput. Math.* **23**(3), 833–898.
[3] Carlier, G., Jimenez, C. & Santambrogio, F. (2008) Optimal transportation with traffic congestion and wardrop equilibria. *SIAM J. Control Optim.* **47**(3), 1330–1350.
[4] Chen, Y., Georgiou, T. T. & Pavon, M. (2016) On the relation between optimal transport and Schrödinger bridges: A stochastic control viewpoint. *J. Optim. Theory Appl.* **169**(2), 671–691.
[5] Chen, Y., Georgiou, T. T. & Pavon, M. (2021) Controlling uncertainty. *IEEE Control Syst. Mag.* **41**(4), 82–94.
[6] Chen, Y., Georgiou, T. T. & Pavon, M. (2021) Stochastic control liaisons: Richard Sinkhorn meets Gaspard Monge on a Schrödinger bridge. *SIAM Rev.* **63**(2), 249–313.
[7] Cullen, M. J. P. & Douglas, R. J. (1999) Applications of the Monge-Ampere equation and monge transport problem to meteorology and oceanography. *Contemp. Math.* **226**, 33.
[8] Ekren, I. & Soner, H. M. (2018) Constrained optimal transport. *Arch. Ration. Mech. Anal.* **227**(3), 929–965.
[9] Fan, J., Haasler, I., Karlsson, J. & Chen, Y. (2022) On the complexity of the optimal transport problem with graph-structured cost. In *International Conference on Artificial Intelligence and Statistics*, PMLR, pp. 9147–9165.
[10] Gangbo, W. & McCann, R. J. (1996) The geometry of optimal transportation. *Acta Math-DJURSHOLM* **177**(2), 113–161.
[11] Gladbach, P. & Kopfer, E. (2022) Limits of density-constrained optimal transport. *Calc. Var. Partial Differ. Equations* **61**(2), 1–31.
[12] Haker, S., Zhu, L., Tannenbaum, A. & Angenent, S. (2004) Optimal mass transport for registration and warping. *Int. J. Comput. Vision* **60**(3), 225–240.
[13] Jordan, R., Kinderlehrer, D. & Otto, F. (1998) The variational formulation of the Fokker–Planck equation. *SIAM J. Math. Anal.* **29**(1), 1–17.
[14] Korman, J. & McCann, R. J. (2013) Insights into capacity-constrained optimal transport. *Proc. Natl. Acad. Sci.* **110**(25), 10064–10067.
[15] Léonard, C. (2007) A large deviation approach to optimal transport. arXiv preprint arXiv:0710.1461.
[16] Otto, F. (2001) The geometry of dissipative evolution equations: The porous medium equation. *Commun. Partial Differ. Equations* **26**(1-2), 101–174.
[17] Peyré, G. & Cuturi, M. (2019) Computational optimal transport: With applications to data science. *Found. Trends® Mach. Learn.* **11**(5-6), 355–607, Publishers inc. Boston.
[18] Svetlozar, T. (1998) Rachev and Ludger Rüschendorf. In *Mass Transportation Problems: Volume I: Theory*, Vol. 1, Springer Science & Business Media, New York, USA.
[19] Santambrogio, F. (2015) Optimal transport for applied mathematicians. *Birkäuser, NY* **55**(58-63), 94.
[20] Villani, C. (2009) *Optimal Transport: Old and New*, vol. 338, Springer, Berlin.
[21] Villani, C. (2021) *Topics in Optimal Transportation*, vol. 58. American Mathematical Society, Providence, Rhode Island, USA.

Cite this article: Stephanovitch A., Dong A. and Georgiou T. T. (2025). Optimal transport through a toll station. *European Journal of Applied Mathematics*, **36**, 613–637. https://doi.org/10.1017/S0956792524000317