DYNAMIC OPTIMAL TRANSPORT ON NETWORKS

MARTIN BURGER[1], INA HUMPERT[2] AND JAN-FREDERIK PIETSCHMANN[3,*]

Abstract. We study a dynamic optimal transport problem on a network. Despite the cost for transport along the edges, an additional cost, scaled with a parameter κ, has to be paid for interchanging mass between edges and vertices. We show existence of minimisers using duality and discuss the relationship of the model to other metrics such as Fisher–Rao and the classical Wasserstein metric. Finally, we examine the limiting behaviour of the model in terms of the parameter κ.

1. INTRODUCTION

Transport on networks is an important problem that arises in many areas of science, *e.g.* traffic on road networks [6], distribution of gas [2, 19], or transport of vesicles within neurites, [12, 13, 25, 30]. On the other hand, there is the generic theory of optimal transport that describes how to move a given amount of mass at the lowest cost (see for example the books of Villani [28, 29], Santambrogio [23] or the survey of Brasco [5] for the necessary background) and its dynamic variant introduced by Benamou and Brenier [3].

In this paper, we aim to combine these notions by introducing a dynamic formulation of optimal transport on a network, where mass is transported along edges but can also be stored in vertices.

In the classical theory of optimal transport the *Wasserstein-distance* of order p between two probability measures μ_1 and μ_2 on $\Omega \subset \mathbb{R}^n$ is defined by as

$$d(\mu_1, \mu_2)^p := \inf_{p \in \mathbb{P}(\mu_1, \mu_2)} \int_{\Omega \times \Omega} |x - y|^p \, dp(x, y), \tag{1.1}$$

where the symbol $|\cdot|$ denotes the Euclidean norm on $\mathbb{R}^n$ and $\mathbb{P}$ denotes the set of probability measures with marginals μ_1 and μ_2.

Initially introduced as a numerical algorithm, Benamou and Brenier [3] introduced a dynamic version of the optimal transport problem. They showed that calculating the Wasserstein distance is equivalent to minimising an action functional representing the kinetic energy of curves connecting the two measures ρ_1 and ρ_2, subject

to a constraint given by a continuity equation, *i.e.*

$$W_2^2(\rho_1, \rho_2) := \inf_{(\mu_t, F_t)} \left\{ \int_0^1 \int_\Omega \frac{|F_t(x)|^2}{\mu_t(x)} \, \mathrm{d}x\mathrm{d}t \; \middle| \; \partial_t\mu_t + \mathrm{div}(F_t) = 0, \; \mu_t|_{t=0,1} = \rho_1, \rho_2 \right\},$$

where F_t denotes the flux and μ_t a curve in the space of probability measures. This formulation has the additional merit that can be easily generalised, *e.g.* to include non-linear mobilities [7, 17]. One case which is of particular interest here is when the initial and terminal measures ρ_1 and ρ_2 have different mass. Then, the classical Wasserstein distance can be replaced by the Wasserstein–Fisher–Rao metric whose dynamic formulation is given as

$$
\begin{aligned}
\mathcal{WFR}_\kappa^2(\rho_0, \rho_1) = \inf_{(\gamma_t, G_t, f_t)} & \left\{ \iint_{[0,1]\times\Omega} \frac{|G_t|^2}{2\gamma_t} \, \mathrm{d}x\mathrm{d}t + \kappa^2 \iint_{[0,1]\times\Omega} \frac{|f_t|^2}{2\gamma_t} \, \mathrm{d}x\mathrm{d}t \right. \\
& \left. \text{s.t.} \quad \partial_t\gamma_t + \mathrm{div}\, G_t = f, \; \gamma_t|_{t=0,1} = \rho_1, \rho_2 \right\},
\end{aligned}
\tag{1.2}
$$

for $\gamma_t \in \mathcal{M}^+(\Omega)$. This allows to compute distances between measures with different masses, see [8] for an existence result. Recently it was shown that a static version in the spirit of (1.1) exits, see [9] for details. The general theory that deals with probability measures of different mass is called *unbalanced optimal transport* and was simultaneously introduced and studied in [8, 9, 11, 14–16, 21].

More recently, [20] introduced a new transportation model on the closure of a domain $\overline{\Omega}$ that behaves differently in the interior and on the boundary while allowing for interaction between these two. This setting can be motivated if we think of Ω as a city with a ring road $\partial\Omega$ which can only be entered upon paying a fee denoted by κ. The overall density of cars is then $\rho = (\omega, \gamma) \in \mathcal{M}^+(\overline{\Omega}) \times \mathcal{M}^+(\partial\Omega)$, where the first entry corresponds to the cars in the inner city and the second entry to the those on the ring road. Informally, this model is given as

$$
\begin{aligned}
\mathcal{W}_{\mathrm{M}}^2(\rho_0, \rho_1) = \inf_{(\omega_t, \gamma_t, F_t, G_t, f_t)} & \left\{ \iint_{[0,1]\times\overline{\Omega}} \frac{|F_t|^2}{2\omega_t} \, \mathrm{d}x\mathrm{d}t + \iint_{[0,1]\times\partial\Omega} \frac{|G_t|^2}{2\gamma_t} \, \mathrm{d}x\mathrm{d}t + \kappa^2 \iint_{[0,1]\times\partial\Omega} \frac{|f_t|^2}{2\gamma_t} \, \mathrm{d}\sigma\mathrm{d}t \right. \\
& \left. \text{s.t} \quad \begin{matrix} \partial_t\omega_t + \mathrm{div}\, F_t = 0 & \text{in } \Omega, \\ F_t \cdot \nu = f_t & \text{in } \partial\Omega, \end{matrix} \quad \text{and} \quad \partial_t\gamma_t + \mathrm{div}\, G_t = f_t \text{ in } \partial\Omega \right\},
\end{aligned}
\tag{1.3}
$$

where the initial concentration is defined by $\rho_0 = (\omega_0, \gamma_0)$, the terminal concentration by $\rho_1 = (\omega_1, \rho_1)$, F_t denotes the momentum in Ω, G_t the momentum on $\partial\Omega$ and f_t is the normal outflux $f_t = F_t \cdot \nu$. Existence of minimisers was shown based on duality. Understanding the one-dimensional model as a trivial network with only one edges and two vertices (in spatial dimension one) serves as a starting point for our investigation.

In this work we introduce a dynamic formulation on a planar network. We identify edges with one-dimensional intervals on which a classical action functional is minimized while at vertices, mass may be transferred onto or off from a vertex by reaction terms as in the Fisher–Rao metric (1.2), rendering the transport problem on each edge similar to (1.3). After the formulation of the problem, we show that it is well-defined again using Fenchel–Rockafellar duality, and also analyse the asymptotic behaviour in terms of the cost parameter κ.

While our model is dynamic and allows for the storage of mass at the vertices, the static 1-Wasserstein distance on metric graphs has been studies in [18]. In their work, the authors focus on a connection between Kantorovich potentials and solutions of a p-Laplace problem. More recently in [10], the authors consider a similar setting as ours, yet again without explicit dynamics at the vertices. They introduce a p-Wasserstein distance in the spirit of Benamou and Brenier and show that, as in the classical setting, absolutely continuous curves admit vector fields that solve the continuity equation. Using this characterisation of geodesics, they observe

that the entropy functional is not displacement convex. However, they are still able to characterise solutions to a drift-diffusion-attraction equation as a gradient flow with respect to their distance.

This paper is organised as follows: In Section 2 we provide details on the network setting, in Section 3 we introduce our model and show existence of minimisers by means of Rockafellar duality. Section 4 we discuss the relationships of the distance-functional with other metrics as the Fisher–Rao and the classical Wasserstein-metric. In the limit case, where the costs for transporting mass over the vertices diverges to infinity, we recover that our distance converges either to infinity if masses are incompatible or to the classical Wasserstein-metric if masses are compatible. In the Appendix we present a formal calculation of the first order optimality conditions.

2. Network setting

We consider a planar network where edges can be identified with one-dimensional intervals. We denote the complete network by $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ with $\mathcal{V} = \{V^1, \ldots, V^n\}$ the set of vertices for $n \in \mathbb{N}$ and $\mathcal{E} = \{E^1, \ldots, E^m\}$ the set of edges for $m \in \mathbb{N}$. Every vertex is defined via its coordinates in the two-dimensional space $\mathbb{R}^2$, $i.e.$ $V^i \in \mathbb{R}^2$ for every $i \in \{1, \ldots, n\}$ and every edge is homeomorphic to a one-dimensional, open interval. To each edge we assign a starting and an end point and define two functions $\alpha, \omega \colon \mathcal{E} \to \mathcal{V}$ that assign to every edge its starting or its end point thus determining an orientation. The functions $\bar{\alpha}, \bar{\omega} \colon \{1, \ldots, m\} \to \{1, \ldots, n\}$ map a given edge the indices of the respective vertices. By $Z(V^i)$ we denote the indices of all edges originating or ending at V^i for $i \in \{1, \ldots, n\}$, $i.e.$

$$Z(V^i) = \big\{ j \in \{1, \ldots, m\} \colon \alpha(E^j) = V^i \vee \omega(E^j) = V^i \big\}.$$

Finally, for all $(i, j) \in \{(i, j) \; : \; i \in \{1, \ldots, n\} \text{ and } j \in Z(V^i)\}$ we denote by $\nu_{i,j}$ the outward normal vector of edge j at the point where it is connected to vertex i. With this notation, $\nu_{\bar{\alpha}(j),j}$ gives the normal at the starting point of E_j. Moreover, we denote by $\mathcal{M}_+(X)$ the set of non-negative bounded measures on a given space X and more precisely the set of non-negative measures on the set of edges (vertices) by

$$\mathcal{M}_+(\mathcal{E}) = \mathcal{M}_+(E^1) \times \ldots \times \mathcal{M}_+(E^m),$$
$$\mathcal{M}_+(\mathcal{V}) = \mathcal{M}_+(V^1) \times \ldots \times \mathcal{M}_+(V^n).$$

Since $V^i \in \mathbb{R}^d$, we have that each measure in $\mathcal{M}_+(\mathcal{V})$ is of the form

$$\sum_{i=1}^n c_i \delta_{V_i},$$

and therefore we identify $\mathcal{M}_+(\mathcal{V})$ with $\mathbb{R}_+^n$ from now on.

To formulate the dynamic optimal transport problem on the network let $\rho_0^j, \rho_1^j \in \mathcal{M}_+(E^j)$, $j = 1, \ldots, m$ and $\gamma_0^i, \gamma_1^i \in \mathbb{R}$, $i = 1, \ldots, n$ be given, and denote by $\boldsymbol{\rho_0} = (\rho_0^1, \ldots, \rho_0^m)$, $\boldsymbol{\rho_1} = (\rho_1^1, \ldots, \rho_1^m)$, the vector of all concentrations on edges at time $t = 0$ and $t = 1$ and by $\boldsymbol{\gamma_0} = (\gamma_0^1, \ldots, \gamma_0^n)$, $\boldsymbol{\gamma_1} = (\gamma_1^1, \ldots, \gamma_1^n)$ the vectors of the concentration on the vertices at time $t = 0$ and $t = 1$. Next, we define $\Omega_{\mathcal{E}} = \bigcup_{j=1}^m E^j$ as well as $\Omega_{\mathcal{V}} = \bigcup_{i=1}^n V^i$ and on the closed set $\Omega_{\mathcal{G}} = \Omega_{\mathcal{V}} \cup \Omega_{\mathcal{E}}$ we define the measure that translates to the total density on the network by

$$\varsigma_l = \sum_{j=1}^m \rho_l^j + \sum_{i=1}^n \gamma_l^i, \quad l \in \{0, 1\}, \tag{2.1}$$

and make the assumption that our initial and final data $(\boldsymbol{\rho_0}, \boldsymbol{\rho_1}, \boldsymbol{\gamma_0}, \boldsymbol{\gamma_1})$ are such that $\varsigma_0, \varsigma_1 \in \mathcal{P}(\Omega_{\mathcal{G}})$ holds.
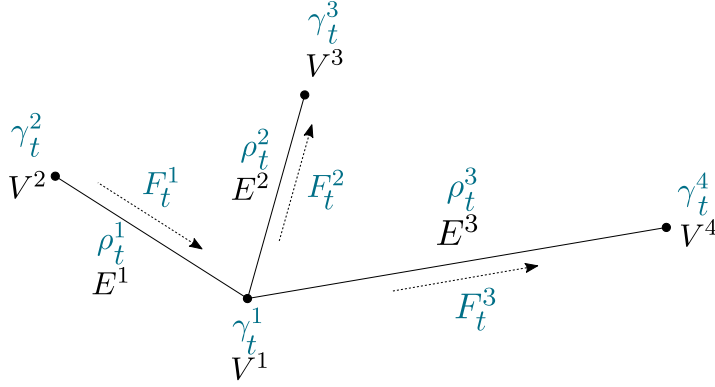
FIGURE 1. Example of a network configuration with three edges $\mathcal{E} = \{E^1, E^2, E^3\}$ and four vertices $\mathcal{V} = \{V^1, V^2, V^3, V^4\}$. The concentrations on E^j are given by ρ_t^j for $j \in \{1, 2, 3\}$ and the concentration on the vertices is given by γ_t^i on V^i with $i \in \{1, 2, 3, 4\}$. Labels of edges and vertices are shown in black and concentrations on those are shown in blue. The index set corresponding to the edges originating or ending at V^1 is given by $Z(V^1) = \{1, 2, 3\}$ as $\omega(E^1) = \alpha(E^2) = \alpha(E^3) = V^1$.

On every edge and every vertex, *i.e.* for every $j \in \{1, \ldots, m\}$ and $i \in \{1, \ldots, n\}$, we consider the following *continuity equation* on the network $\mathcal{G}$

$$\partial_t \rho_t^j + \partial_x F_t^j = 0 \text{ in } E^j, \quad \partial_t \gamma_t^i = f_t^i \text{ on } V^i \quad \text{with } f_t^i = \sum_{j \in Z(V^i)} F_t^j(V_i) \cdot \nu_{i,j}, \quad (2.2)$$

where $F_t^j \colon E^j \times (0, T] \to \mathbb{R}$, $f_t^i(0, T] \to \mathbb{R}$, the space derivative $\partial_x F_t^j$ is calculated with respect to the orientation of the edge. We will give a rigorous definition of weak solution in the next section. A sketch of this situation is shown in Figure 1. Note that as $\mathcal{V}$ is discrete, the time-dependent measures γ_t^i, f_t^i for $i \in \{1, \ldots, n\}$ are given by

$$\gamma_t^i = C_1(i, t)\delta_{V^i}, \qquad f_t^i = C_2(i, t)\delta_{V^i}$$

where δ_{V^i} denotes the Dirac-measure at the vertex V^i and $C_1(i, t), C_2(i, t) \in \mathbb{R}^+$ are time-dependent constants. Thus, we identify the measures with their respective time-dependent constants while, by abuse of notation, still writing γ_t^i and f_t^i in the following.

For a given network concentration $(\boldsymbol{\rho_0}, \boldsymbol{\rho_1}, \boldsymbol{\gamma_0}, \boldsymbol{\gamma_1}) \in \mathcal{M}_+(\mathcal{E}) \times \mathcal{M}_+(\mathcal{E}) \times \mathcal{M}_+(\mathcal{V}) \times \mathcal{M}_+(\mathcal{V})$ with $\varsigma_0, \varsigma_1 \in \mathcal{P}(\Omega_{\mathcal{G}})$, we consider the *minimisation-problem* of an action being the combination of Wasserstein and the Fisher–Rao terms

$$\min_{(\boldsymbol{\rho_t}, \boldsymbol{F_t}, \boldsymbol{\gamma_t}, \boldsymbol{f_t}) \in \mathcal{CE}(\boldsymbol{\rho_0}, \boldsymbol{\rho_1}, \boldsymbol{\gamma_0}, \boldsymbol{\gamma_1})} \left\{ \sum_{j=1}^m \iint_{\overline{E^j} \times [0,1]} \frac{|F_t^j|^2}{2\rho_t^j} \, \mathrm{d}x\mathrm{d}t + \kappa^2 \sum_{i=1}^n \int_{[0,1]} \frac{|f_t^i|^2}{2\gamma_t^i} \, \mathrm{d}t \right\}, \quad (2.3)$$

where $\kappa > 0$ is a given constant and with

$$\mathcal{CE}(\boldsymbol{\rho_0}, \boldsymbol{\rho_1}, \boldsymbol{\gamma_0}, \boldsymbol{\gamma_1}) = \left\{ (\boldsymbol{\rho_t}, \boldsymbol{F_t}, \boldsymbol{\gamma_t}, \boldsymbol{f_t}) \text{ that fulfil } (2.2) \text{ and } \boldsymbol{\rho_{t=0}} = \boldsymbol{\rho_0}, \boldsymbol{\rho_{t=1}} = \boldsymbol{\rho_1}, \boldsymbol{\gamma_{t=0}} = \boldsymbol{\gamma_0}, \boldsymbol{\gamma_{t=1}} = \boldsymbol{\gamma_1} \right\}.$$

In this setting - at least on a formal level - the total mass on the network is conserved as no mass may enter or leave the system, *i.e.*

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathrm{Mass}(\mathcal{G},t) = \frac{\mathrm{d}}{\mathrm{d}t}\left[\sum_{j=1}^{m}\int_{\overline{E^j}}\rho_t^j\,\mathrm{d}x + \sum_{i=1}^{n}\gamma_t^i\right] = 0$$

for every $t \in [0,T]$. Moreover, for brevity we write $\|\rho_t^j\| = \|\rho_t^j\|_{L^1(E^j)}$ in the whole paper.

Remark 2.1 (Generalisations).   a) An interesting generalisation of the current setting is the case where we allow mass in- and outflow at the vertices of grad 1, *i.e.* the outer vertices that are only connected with one edge.

   b) The generalisation to a non-connected network with finite many connected components is possible but we omit the proof for readability.

Remark 2.2 (Kantorovich formulation and limit problem). Another interesting question for further research is a static formulation of both (1.3) and (2.3). For the first one, based on the explicit calculation of geodesics in [20] between two point masses, one being located within the domain and one at the boundary, we conjecture that (1.3) allows for a Kantorovich formulation with cost

$$c(x,y) = \begin{cases} \frac{1}{2}|x-y|^2 & x,\,y \in \Omega \text{ or } x,\,y \in \partial\Omega, \\ \frac{1}{2}|x-y|^2 + \kappa\sqrt{|x-y|^2 + \kappa^2} + \kappa^2 & x \in \Omega,\,y \in \partial\Omega \text{ or } x \in \partial\Omega,\,y \in \Omega. \end{cases} \tag{2.4}$$

Furthermore, one might ask if one starts with a standard optimal transport problem on domain $\Omega = I_1 \cup I_2$ and let the size of I_2 go to zero, is there an appropriate rescaling of the cost so that we obtain (1.3) in the limit? This could be examined either on the level of optimality conditions or in the respective static solutions where the limit cost would then need to be (2.4).

3. Transport model and existence of minimisers

In this section, we show existence of minimisers of (2.3), based on a duality argument and by extending the strategy of [20] to the network setting.

3.1. Continuity equations and action functional

We introduce the notations

$$Q_{\overline{E^j}} := [0,1] \times \overline{E^j}, \quad \text{and} \quad Q_{V^i} := [0,1] \times V^i.$$

The precise notion of weak solution for the continuity equation (2.2) is given as follows.

Definition 3.1 (weak solution). Given $(\boldsymbol{\rho_0}, \boldsymbol{\rho_1}, \boldsymbol{\gamma_0}, \boldsymbol{\gamma_1}) \in \mathcal{M}_+(\mathcal{E}) \times \mathcal{M}_+(\mathcal{E}) \times \mathcal{M}_+(\mathcal{V}) \times \mathcal{M}_+(\mathcal{V})$, we denote by $\mathcal{CE}(\boldsymbol{\rho_0}, \boldsymbol{\rho_1}, \boldsymbol{\gamma_0}, \boldsymbol{\gamma_1})$ the set of measures $(\boldsymbol{\rho_t}, \boldsymbol{F_t}, \boldsymbol{\gamma_t}, \boldsymbol{f_t})$ which satisfy the continuity equation (2.2) in the following weak sense

$$\sum_{j=1}^{m}\Big[\iint_{Q_{\overline{E^j}}}(\partial_t\varphi_t\,\mathrm{d}\rho_t^j + \partial_x\varphi_t\,dF_t^j)\,\mathrm{d}t\Big] - \sum_{i=1}^{n}\varphi_t(V_i)f_t^i$$
$$= \sum_{j=1}^{m}\Big[\int_{\overline{E^j}}\varphi_1\,\mathrm{d}\rho_1^j - \int_{\overline{E^j}}\varphi_0\,\mathrm{d}\rho_0^j\Big] \tag{3.1}$$

for all test functions $\varphi_t \in C^1(\Omega_\mathcal{G} \times [0,1])$ and

$$\sum_{i=1}^{n} \Big[\int_{[0,1]} \partial_t \varphi_t^i \, \gamma_t^i \mathrm{d}t \Big] = \sum_{i=1}^{n} \Big[\int_{[0,1]} \varphi_t^i \, df_t^i \Big] + \sum_{i=1}^{n} \Big[\varphi_1^i \gamma_1^i - \varphi_0^i \gamma_0^i \Big] \tag{3.2}$$

for all $\varphi^i \in C^1([0,1])$, $i = 1, \ldots, n$ denoting the test function corresponding to the vertex V^i.

Remark 3.2. Note that as we are integrating over $\overline{E^j}$, formally each continuity equation produces a boundary term $\varphi_t(\omega(E_j))F_j \cdot \nu_{\bar\omega(j),j} - \varphi_t(\alpha(E_j))F_j \cdot \nu_{\bar\alpha(j),j}$. Summing over all equations this yields

$$\sum_{j=1}^{m} (\varphi_t(\omega(E_j))F_t^j(\omega(E_j)) \cdot \nu_{\bar\omega(j),j} + \varphi_t(\alpha(E_j))F_t^j(\alpha(E_j)) \cdot \nu_{\bar\alpha(j),j}$$

$$= \sum_{i=1}^{n} \varphi_t(V_i) \sum_{j \in Z(V_i)} F_t^j(V_i) \cdot \nu_{i,j} = \sum_{i=1}^{n} \varphi_t(V_i) f_t^i,$$

using the coupling of fluxes in (2.2) which is therefore incorporated in the weak formulation.

The continuity equation (2.2) can also be considered in the global sense, *i.e.* can be formulated as an equation on the whole network. This results in:

Proposition 3.3. *(global continuity equation) Let* $(\boldsymbol{\rho_t}, \boldsymbol{F_t}, \boldsymbol{\gamma_t}, \boldsymbol{f_t}) \in \mathcal{CE}(\boldsymbol{\rho_0}, \boldsymbol{\rho_1}, \boldsymbol{\gamma_0}, \boldsymbol{\gamma_1})$ *and define by* $\tilde{F}_t^j$ *the trivial extension of* F_t^j *on the whole network, i.e.* $\tilde{F}_t^j = 0$ *on all other edges* E^k *with* $k \neq j$, *(the measures* $\tilde{\gamma}_t^j, \tilde{f}_t^i$ *and* $\tilde{\rho}_t^j$ *are defined analogue). We define the following global variables*

$$\varsigma_t = \sum_{j=1}^{m} \tilde{\rho}_t^j + \sum_{i=1}^{n} \tilde{\gamma}_t^i, \quad H_t = \sum_{j=1}^{m} \tilde{F}_t^j \quad and \quad h_t = \sum_{i=1}^{n} \tilde{f}_t^i.$$

Then, the existence of the global continuity equation

$$\begin{cases} \partial_t \varsigma_t + \partial_x H_t = h_t & in \ \Omega_\mathcal{G} \times [0,1], \\ \sum_{j \in Z(V^i)} F_t^j \cdot \nu_{i,j} = f_t^i & in \ \Omega_\mathcal{V} \times [0,1]. \end{cases}$$

in the weak sense with initial $(\boldsymbol{\rho_0}, \boldsymbol{\gamma_0})$ *and terminal data* $(\boldsymbol{\rho_1}, \boldsymbol{\gamma_1})$ *such that* $\varsigma_0, \varsigma_1 \in \mathcal{P}(\Omega_\mathcal{G})$ *for* ς_0, ς_1 *defined in* (2.1). *This is a consequence of the existence of* 3.1.

Proof. Choosing $\varphi_t^i = \varphi|_{V^i}$ in (3.2) (where the pointwise restriction exists as $\varphi \in C^1(\Omega_\mathcal{G} \times [0,1])$) and summing up (3.1) and (3.2), we recover the claimed weak formulation

$$\iint_{\Omega_\mathcal{G} \times [0,1]} \partial_t \varphi_t \, d\Big(\sum_{j=1}^{m} \tilde{\rho}_t^j + \sum_{i=1}^{n} \tilde{\gamma}_t^i \Big) + \iint_{\Omega_\mathcal{G} \times [0,1]} \partial_x \varphi_t \, d\Big(\sum_{j=1}^{m} \tilde{F}_t^j \Big) - \iint_{\mathcal{V} \times [0,1]} \varphi_t \, dh_t$$

$$= \iint_{\Omega_\mathcal{G} \times [0,1]} \varphi_t \, d\Big(\sum_{i=1}^{n} \tilde{f}_t^i \Big) + \int_\mathcal{G} \varphi_1 \, d\Big(\sum_{j=1}^{m} \tilde{\rho}_1^j + \sum_{i=1}^{n} \tilde{\gamma}_1^i \Big) - \int_\mathcal{G} \varphi_0 \, d\Big(\sum_{j=1}^{m} \tilde{\rho}_0^j + \sum_{i=1}^{n} \tilde{\gamma}_0^i \Big)$$

for all $\varphi_t \in C^1(\Omega_\mathcal{G} \times [0,1])$. $\qquad\square$

To rigorously define the minimisation of the action functional, we need the following definition.

Definition 3.4 (generalised Lagrangian). For $(a, b) \in \mathbb{R}^+ \times \mathbb{R}$ we define the action density

$$A(a, b) = \begin{cases} \frac{|b|^2}{2a} & \text{if } a > 0, \\ 0 & \text{if } (a, b) = (0, 0), \\ +\infty & \text{otherwise.} \end{cases}$$

For $\boldsymbol{a} \in \mathbb{R}^n$, $\boldsymbol{b} \in \mathbb{R}^n$ and $\boldsymbol{\eta} = (\boldsymbol{\eta}^1, \ldots, \boldsymbol{\eta}^n) \in \mathbb{R}^{|Z(V^1)|} \times \cdots \times \mathbb{R}^{|Z(V^n)|}$ with $\boldsymbol{\eta}^i = (\eta^{i,j})_{j \in Z(V^i)}$, we also introduce the extended action density as

$$\bar{A}(\boldsymbol{a}, \boldsymbol{b}; \boldsymbol{\eta}) = \sum_{i=1}^n \left(\kappa^2 A(a^i, b^i) + \begin{cases} 0 & \text{if } a^i + \sum_{k \in Z(V^i)} \eta^{i,k} = 0, \\ +\infty & \text{otherwise.} \end{cases} \right) \tag{3.3}$$

Let us remark that $A(a, b)$ is convex in both variables, lower semi-continuous and 1-homogeneous. This allows us to give a rigorous definition of the action functional.

Definition 3.5 (action functional). For $\boldsymbol{\mu_t} = (\boldsymbol{\rho_t}, \boldsymbol{F_t}; \boldsymbol{\gamma_t}, \boldsymbol{f_t})$, we define the action functional as

$$\mathcal{A}(\boldsymbol{\mu_t}) := \sum_{j=1}^m \iint_{Q_{\overline{E^j}}} A\left(\frac{\mathrm{d}\rho_t^j}{\mathrm{d}\theta^j}, \frac{\mathrm{d}F_t^j}{\mathrm{d}\theta^j}\right) \mathrm{d}\theta^j + \kappa^2 \sum_{i=1}^n \int_0^1 A(\gamma_t^i, f_t^i)\, \mathrm{d}t, \tag{3.4}$$

where θ^j are non-negative Borel reference measures such that $|\rho_t^j| \ll \theta^j, |F_t^j| \ll \theta^j$. Since A is jointly 1-homogeneous, this definition does not depend on the choice of θ^j.

Note that this action functional is lower semi-continuous, see [4], Theorem 3.3. We now define the bounded Lipschitz-distance between two measures $\rho_0^j, \rho_1^j \in \mathcal{M}^+(E^j)$ on an edge and, respectively, for two measures $\gamma_0^i, \gamma_1^i \in \mathcal{M}^+(V^i)$ on a vertex, as

$$d_{\mathrm{BL}, \mathcal{E}}(\rho_0^j, \rho_1^j) = \sup \left\{ \left| \int_{\overline{E^j}} \Phi d(\rho_1^j - \rho_0^j) \right| \quad \text{s.t.} \quad \|\Phi\|_\infty + \mathrm{Lip}(\Phi) \le 1 \right\},$$
$$d_{\mathrm{BL}, \mathcal{V}}(\gamma_0^i, \gamma_1^i) = |\gamma_0^i - \gamma_1^i|,$$

where $\Phi \colon E^j \to \mathbb{R}$ is a Lipschitz-continuous function with $\mathrm{Lip}(\Phi)$ denoting its Lipschitz-constant (again, by abuse of notion we denote by γ_t^i a measure as well as the corresponding constant). It is well known that the bounded Lipschitz-distance metrises the narrow convergence of probability measures. Let us continue by summarising some properties of solutions to the continuity equation.

Proposition 3.6 (properties of solutions of continuity equations). *Any admissible quadruple* $(\boldsymbol{\rho_t}, \boldsymbol{F_t}; \boldsymbol{\gamma_t}, \boldsymbol{f_t}) \in \mathcal{CE}(\boldsymbol{\rho_0}, \boldsymbol{\rho_1}, \boldsymbol{\gamma_0}, \boldsymbol{\gamma_1})$ *can be disintegrated in time as*

$$d\rho_t^j(x, t) = d\rho_t^j(x)\mathrm{d}t$$

for all $j \in \{1, \ldots, m\}$. *If the action functional* $\mathcal{A}(\boldsymbol{\rho_t}, \boldsymbol{F_t}, \boldsymbol{\gamma_t}, \boldsymbol{f_t}) < \infty$ *is finite, then we obtain the following:*

(i) *For every* t, *the measures* ρ_t^j, γ_t^i *are non-negative for all* $i \in \{1, \ldots, n\}, j \in \{1, \ldots, m\}$. *Moreover, the total mass* ς_t *defined in* (2.1) *is conserved, i.e.*

$$\|\varsigma_t\| = \sum_{j=1}^m \|\rho_t^j\| + \sum_{i=1}^n \gamma_t^i = 1, \qquad \text{for almost every } t \in [0, 1].$$

(ii) *For every* $j \in \{1, \ldots, m\}$ *the Radon–Nikodym densities*

$$u_t^j(x) := \frac{\mathrm{d}F_t^j}{\mathrm{d}\rho_t^j},$$

are well-defined $\mathrm{d}\rho_t^j$ *almost everywhere and we obtain an alternative formulation of* $\mathcal{A}(\boldsymbol{\mu_t})$ *as*

$$\mathcal{A}(\boldsymbol{\rho_t}, \boldsymbol{F_t}; \boldsymbol{\gamma_t}, \boldsymbol{f_t}) = \frac{1}{2} \sum_{j=1}^{m} \iint_{\mathcal{Q}_{\overline{E^j}}} |u_t^j|^2 \mathrm{d}\rho_t^j \, \mathrm{d}t + \frac{1}{2} \sum_{i=1}^{n} \int_0^1 \kappa^2 \frac{|f_t^i|^2}{\gamma_t^i} \, \mathrm{d}t$$

$$= \frac{1}{2} \sum_{j=1}^{m} \iint_{\mathcal{Q}_{\overline{E^j}}} |u_t^j|^2 \mathrm{d}\rho_t^j \mathrm{d}t + \frac{1}{2} \sum_{i=1}^{n} \int_0^1 \kappa^2 \frac{|f_t^i|^2}{\gamma_t^i} \, \mathrm{d}t \, .$$

(iii) *The curves* $t \mapsto \rho_t^j \in \mathcal{M}(\overline{E^j})$ *and* $t \mapsto \gamma_t^i \in \mathcal{M}(V^i)$ *are narrowly continuous for all* $i \in \{1, \ldots, n\}, j \in \{1, \ldots, m\}$ *and satisfy the bounded-Lipschitz estimate*

$$\sum_{j=1}^{m} d_{BL,\overline{E^j}}(\rho_s^j, \rho_t^j) + \sum_{i=1}^{n} d_{BL,V^i}(\gamma_s^i, \gamma_t^i) \leq C_\kappa \sqrt{\mathcal{A}((\boldsymbol{\rho_t}, \boldsymbol{F_t}; \boldsymbol{\gamma_t}, \boldsymbol{f_t}))} |t - s|^{\frac{1}{2}} \tag{3.5}$$

for every time $s, t \in [0, 1]$ *and the constant* $C_\kappa = 2\sqrt{2(n+m)} \max\{1, \frac{1}{\kappa}\}$. *In particular the initial/terminal conditions are taken in the narrow sense, i.e.*

$$\lim_{t \to 0} \int_{\overline{E^j}} \phi(x) \, \mathrm{d}\rho_t^j(x) = \int_{\overline{E^j}} \phi(x) \, \mathrm{d}\rho_0^j(x), \quad \lim_{t \to 1} \int_{\overline{E^j}} \phi(x) \, \mathrm{d}\rho_t^j(x) = \int_{\overline{E^j}} \phi(x) \, \mathrm{d}\rho_1^j(x),$$

for all $\phi \in C^0(\overline{E^j})$.

Proof. For (i) and (ii) we refer to Proposition 3.5 in [20]. For (iii) note that the bounded Lipschitz-distance metrises the narrow convergence of measures, thus it suffices to establish equation (3.5). We only sketch the proof for $t \mapsto \rho_t^j$, as the argument is similar for the curve $t \mapsto \gamma_t^i$. For fixed $\Phi \in C^1(\mathcal{G})$ we will estimate from below the time derivatives of

$$l_t^j := \int_{\overline{E^j}} \Phi(x) \, \mathrm{d}\rho_t^j(x).$$

Using the weak form of the continuity equations and arguing as in Proposition 3.5 of [20] we obtain

$$\sum_{j=1}^{m} \left| \frac{\mathrm{d}l_t^j}{\mathrm{d}t} \right| \leq \|\partial_x \Phi\|_\infty \sum_{j=1}^{m} \int_{\overline{E^j}} |u_t^j| \, d\rho_t^j + \|\Phi\|_\infty \sum_{i=1}^{n} |f_t^i| \, .$$

Next, we use the facts that

- $\|u_t^j\|_{L^1(\overline{E}_j; \mathrm{d}\rho_t^j)} \leq \|u_t^j\|_{L^2(\overline{E}_j; \mathrm{d}\rho_t^j)}$ as well as $\|\rho_t^j\| \leq 1$ and $\gamma_t^i \leq 1$ (using (i)),
- $\sum_{i=1}^{n} \sqrt{a_i} \leq \sqrt{n} \left(\sum_{i=1}^{n} a_i \right)^{\frac{1}{2}}$ for all $a_i \geq 0$, by Hölder's inequality,

to calculate

$$\sum_{j=1}^{m} \Big| \frac{\mathrm{d}l_t^j}{\mathrm{d}t} \Big| \leq \big(\|\partial_x \Phi\|_\infty + \|\Phi\|_\infty \big) \Big(\sum_{j=1}^{m} \int_{\overline{E^j}} |u_t^j| \ d\rho_t^j + \sum_{i=1}^{n} |f_t^i| \Big)$$

$$\leq \sqrt{2(n+m)} \big(\|\partial_x \Phi\|_\infty + \|\Phi\|_\infty \big) \Big(\sum_{j=1}^{m} \int_{\overline{E^j}} |u_t^j|^2 \ d\rho_t^j + \sum_{i=1}^{n} |f_t^i|^2 \Big)^{\frac{1}{2}}. \tag{3.6}$$

As

$$\int_0^1 \Big(\sum_{j=1}^{m} \int_{\overline{E^j}} |u_t^j|^2 \ d\rho_t^j + \sum_{i=1}^{n} |f_t^i|^2 \Big) \ \mathrm{d}t \leq 2 \max \Big\{ 1, \frac{1}{\kappa^2} \Big\} \mathcal{A}(\boldsymbol{\mu_t}) < \infty,$$

the expression on the right hand side of (3.6) is an element of $L^2(0,1)$ and thus l is absolutely continuous. We obtain

$$\sum_{j=1}^{m} \Big| \int_{\overline{E^j}} \Phi \ d(\rho_t^j - \rho_s^j) \Big| = \sum_{j=1}^{m} |l_t^j - l_s^j| \leq \sum_{j=1}^{m} \int_s^t \Big| \frac{\mathrm{d}l_\tau^j}{\mathrm{d}\tau} \Big| \ \mathrm{d}\tau \leq \sum_{j=1}^{m} \Big\| \frac{\mathrm{d}l_t^j}{\mathrm{d}t} \Big\|_{L^2(0,1)} |t-s|^{\frac{1}{2}}$$

$$\leq \sqrt{2(n+m)} \big(\|\partial_x \Phi\|_\infty + \|\Phi\|_\infty \big) \Big(\sum_{j=1}^{m} \iint_{[0,1] \times \overline{E^j}} |u_t^j|^2 \ d\rho_t^j \mathrm{d}t + \sum_{i=1}^{n} \int_0^1 |f_t^i|^2 \ \mathrm{d}t \Big)^{\frac{1}{2}} |t-s|^{\frac{1}{2}}$$

$$\leq 2\sqrt{n+m} \big(\|\partial_x \Phi\|_\infty + \|\Phi\|_\infty \big) \Big(\sum_{j=1}^{m} \iint_{[0,1] \times \overline{E^j}} \frac{|u_t^j|^2}{2} \ d\rho_t^j \mathrm{d}t$$

$$+ \frac{1}{\kappa^2} \sum_{i=1}^{n} \int_0^1 \kappa^2 \frac{|f_t^i|^2}{2\gamma^i} \ \mathrm{d}t \Big)^{\frac{1}{2}} |t-s|^{\frac{1}{2}}$$

$$\leq 2\sqrt{n+m} \big(\|\partial_x \Phi\|_\infty + \|\Phi\|_\infty \big) \max \Big\{ 1, \frac{1}{\kappa} \Big\} \sqrt{\mathcal{A}(\boldsymbol{\mu_t})} |t-s|^{\frac{1}{2}} \ .$$

Taking the supremum over all Φ with $\|\Phi\|_\infty + \mathrm{Lip}(\Phi) \leq 1$, we recover the desired estimate for the bounded Lipschitz distance. Analogously we obtain the same inequality for γ_t^i, and thus the constant in the proposition is given by

$$C_\kappa = 4\sqrt{n+m} \max \Big\{ 1, \frac{1}{\kappa} \Big\} \ .$$

$\square$

3.2. Duality and Existence

Definition 3.7. For any admissible network concentration $(\boldsymbol{\rho_0}, \boldsymbol{\rho_1}, \boldsymbol{\gamma_0}, \boldsymbol{\gamma_1})$ and $\boldsymbol{\mu_t} \in \mathcal{CE}(\boldsymbol{\rho_0}, \boldsymbol{\rho_1}, \boldsymbol{\gamma_0}, \boldsymbol{\gamma_1})$, we define the quantity

$$\mathcal{W}_\kappa^2(\boldsymbol{\rho_0}, \boldsymbol{\rho_1}, \boldsymbol{\gamma_0}, \boldsymbol{\gamma_1}) \coloneqq \inf_{\boldsymbol{\mu_t} \in \mathcal{CE}(\boldsymbol{\rho_0}, \boldsymbol{\rho_1}, \boldsymbol{\gamma_0}, \boldsymbol{\gamma_1})} \mathcal{A}(\boldsymbol{\rho_t}, \boldsymbol{F_t}; \boldsymbol{\gamma_t}, \boldsymbol{f_t}). \tag{3.7}$$

It will turn out in Proposition 3.13 that this quantity is even a distance but first of all we have to show that it is always well-defined.

Lemma 3.8. *For every $(\boldsymbol{\rho_0}, \boldsymbol{\rho_1}, \boldsymbol{\gamma_0}, \boldsymbol{\gamma_1}) \in \mathcal{M}_+(\mathcal{E}) \times \mathcal{M}_+(\mathcal{E}) \times \mathcal{M}_+(\mathcal{V}) \times \mathcal{M}_+(\mathcal{V})$ such that $\varsigma_0, \varsigma_1 \in \mathcal{P}(\Omega_{\mathcal{G}})$, the quantity $\mathcal{W}_\kappa^2(\boldsymbol{\rho_0}, \boldsymbol{\rho_1}, \boldsymbol{\gamma_0}, \boldsymbol{\gamma_1})$ is finite.*

Proof. Let $V^k \in \mathcal{V}$ be an arbitrary vertex. We will show that any element $(\boldsymbol{\rho_0}, \boldsymbol{\gamma_0})$ can be connected to $(\mathbf{0}, \boldsymbol{\delta_{V^k}})$ with finite cost, where $\boldsymbol{\delta_{V^k}} = (0, \ldots, \delta_{V^k}, \ldots, 0) \in \mathbb{R}^n$ denotes the concentration on the vertices with δ being the Dirac-measure. By symmetry $(\mathbf{0}, \boldsymbol{\delta_{V^k}})$ can also be connected to any other $(\boldsymbol{\rho_1}, \boldsymbol{\gamma_1})$, thus connecting $\boldsymbol{\rho_0}$ with $\boldsymbol{\rho_1}$ with finite cost. As proven in Lemma 3.7 of [20], it is enough to show that this can be done in a finite number of elementary steps with finite cost combined with a time re-scaling argument. As we assumed that our network is finite and connected, the assertion follows from finitely many consecutive applications of Lemma 3.7 in [20]. □

Closely following to [9], we now proof existence of minimisers of (3.7) using a duality argument and the Fenchel–Rockafellar theorem, see [26] or Theorem 31.1 in [27].

Theorem 3.9 (Fenchel–Rockafellar theorem). *Let X_1, X_2 be normed vector spaces with topological duals X_1^*, X_2^* and $L\colon X_1 \to X_2$ be a bounded linear operator with adjoint $L^*\colon X_2^* \to X_1^*$. Furthermore, let $\mathcal{F}\colon X_1 \to \mathbb{R} \cup \{-\infty\}$ and $\mathcal{G}\colon X_2 \to \mathbb{R} \cup \{-\infty\}$ be two proper, concave functions. If there exists $x \in X_1$ such that $\mathcal{F}(x)$ is finite and $\mathcal{G}$ is continuous at $y = Lx$, then*

$$\sup_{x \in X_1} \{\mathcal{F}(x) + \mathcal{G}(Lx)\} = \min_{y^* \in X_2^*} \{-\mathcal{F}^*(L^*y^*) - \mathcal{G}^*(y^*)\},$$

where $\mathcal{F}^*$ denotes the Fenchel–Legendre conjugate of $\mathcal{F}$, respectively $\mathcal{G}$. Moreover, if there exists $y^* \in X_2^*, x \in X_1$ such that $L^*y^* \in \partial(-\mathcal{F})(x)$ and $Lx \in \partial(-\mathcal{G}^*)(y^*)$ then x achieves the sup *and y^* is a minimizer.*

Next, we define two subsolution sets, one corresponding to the edges and one to the vertices,

$$S_E := \left\{(\alpha, \beta) \in \mathbb{R} \times \mathbb{R}\colon \quad \alpha + \frac{|\beta|^2}{2} \le 0\right\},$$

$$S_{V^i}^\kappa := \left\{(a, b, \boldsymbol{c}) \in \mathbb{R} \times \mathbb{R} \times \mathbb{R}^{|Z(V^i)|}\colon \quad a + \frac{\left|b - \frac{1}{|Z(V^i)|}\sum_{j \in Z(V^i)} c_j\right|^2}{2\kappa^2} \le 0\right\} \tag{3.8}$$

and the convex indicator functions of these sets

$$\iota_{S_E}(\alpha, \beta) := \begin{cases} 0 & \text{if } (\alpha, \beta) \in S_E, \\ +\infty & \text{otherwise,} \end{cases} \quad \text{and} \quad \iota_{S_{V^i}^\kappa}(a, b, \boldsymbol{c}) := \begin{cases} 0 & \text{if } (a, b, \boldsymbol{c}) \in S_{V^i}^\kappa, \\ +\infty & \text{otherwise.} \end{cases}$$

For given i, j, the variables α, β will be dual multipliers to ρ_t^j, F_t^j and a will be dual to γ_t^i. The variable b will be dual to f_t^i and $\boldsymbol{c}$ will be dual to $\sum_{j \in Z(V^i)} F_t^j \cdot \nu_{i,j}$. Their sum $b - \sum c$ translates to the fact that no mass gets lost at a vertex. In the following we will take $(\alpha, \beta) = (\partial_t \phi_t^j, \partial_x \phi_t^i)$ and $(a, b, c) = (\partial_t \psi_t^i, \psi, \phi|_{\mathcal{V}})$ for suitable test functions $\phi_t^j \in C^1, \psi_t^i \in C^1$. Then, $(\partial_t \phi_t^j, \partial_x \phi_t^i) \in S_E$ and $\partial_t \psi_t^i, \psi, \phi|_{\mathcal{V}} \in S_{V^i}^\kappa$ mean that ϕ_t^j, ψ_t^i are (smooth) sub-solutions of the Hamilton–Jacobi system

$$\partial_t \phi_t^j + \frac{1}{2}|\partial_x \phi_t^j|^2 \le 0 \quad \text{and} \quad \partial_t \psi_t^i + \frac{1}{2\kappa^2}\left|\psi_t^i - \frac{1}{|Z(V^i)|}\sum_{j \in Z(V^i)} \phi_t^j\right|^2 \le 0, \tag{3.9}$$

(see Appendix A for the derivation of the Hamilton–Jacobi equations by formal Lagrangian calculus). As in Monsaingeon [20], this system of coupled Hamilton–Jacobi equations is invariant under the addition of a common constant to ψ_t^i and ϕ_t^j, $j \in Z(V^i)$. Consequently, the convex closed set $S_{V^i}^\kappa$ is thus invariant under diagonal shifts $b + k, \boldsymbol{c} + k$, for every constant $k \in \mathbb{R}$.

Lemma 3.10. *For $(\rho_t^j, F_t^j) \in \mathbb{R}^+ \times \mathbb{R}, (\gamma_t^i, f_t^i, \boldsymbol{\eta_t^i}) \in \mathbb{R}^+ \times \mathbb{R} \times \mathbb{R}^{|Z(V^i)|}$ the convex conjugates $\iota_{S_E}^*$ and $\iota_{S_{V^i}^\kappa}^*$ can be identified with the generalised Lagrangians, i.e.*

$$\iota_{S_E}^*(\rho_t^j, F_t^j) = A(\rho_t^j, F_t^j) = \begin{cases} \frac{|F_t^j|^2}{2\rho_t^j} & \text{if } \rho_t^j > 0, \\ 0 & \text{if } (\rho_t^j, F_t^j) = (0,0), \\ +\infty & \text{otherwise.} \end{cases}$$

$$\iota_{S_{V^i}^\kappa}^*(\gamma_t^i, f_t^i, \boldsymbol{\eta_t^i}) = \begin{cases} \kappa^2 \frac{|f_t^i|^2}{2\gamma_t^i} & \text{if } \gamma_t^i > 0 \text{ and } f_t^i + \sum_{k \in Z(V^i)} \eta_t^{i,k} = 0, \\ 0 & \text{if } (\gamma_t^i, f_t^i + \sum_{k \in Z(V^i)} \eta_t^{i,k}) = (0,0), \\ +\infty & \text{otherwise.} \end{cases} \qquad (3.10)$$

where $\boldsymbol{\eta_t^{i,1}} = \eta_t^{i,1}, \ldots, \eta_t^{i,Z(V^i)}$ denotes the vector of neighbouring edges originating or ending at the vertex V^i. The conditions $f_t^i + \sum_{k \in Z(V^i)} \eta_t^{i,k} = 0$ reflects by duality the invariance of $S_{V^i}^\kappa$ under diagonal shifts $b + kC, \boldsymbol{c} + kC$ discussed in the previous paragraph. An alternative version of (3.10) *in terms of the extended action is given by*

$$\sum_{i=1}^n \iota_{S_{V^i}^\kappa}^*(\gamma_t^i, f_t^i, \boldsymbol{\eta_t^i}) = \sum_{i=1}^n \kappa^2 A(\gamma_t^i, f_t^i) + \begin{cases} 0 & \text{if } f_t^i + \sum_{k \in Z(V^i)} \eta_t^{i,k} = 0, \\ +\infty & \text{otherwise} \end{cases} = \bar{A}(\boldsymbol{\gamma_t}, \boldsymbol{f_t}, \boldsymbol{\eta_t}). \qquad (3.11)$$

The proof of this lemma uses calculations similar to those in Lemma 5.17 of [23]. In particular note that the normalisation factor $\frac{1}{|Z(V^i)|}$ that appears in the definition of $S_{V^i}^\kappa$ yields the constraint $f_t^i + \sum_{k \in Z(V^i)} \eta_t^{i,k} = 0$.

As a next step we will define the concept of recession functions that will be needed in the following duality theorem:

Definition 3.11 (recession function [22], Chap. 4)**.** Let T be an arbitrary set. A function f^∞ on $[0,T] \times \mathbb{R}^n$ is called recession-function of a function $f^*: [0,T] \times \mathbb{R}^n \to \mathbb{R}^1 \cup \{+\infty\}$, if

$$f^\infty(t, w) = \lim_{\varepsilon \to +\infty} [f^*(t, x^* + \varepsilon w) - f^*(t, x^*)]/\varepsilon$$

whenever $x^* \in \mathbb{R}^n$ satisfies $f^*(t, x^*) < \infty$.

Note that we wrote f^* to emphasise that this function will be a convex conjugate of a function f in the following. The fact that convex-conjugates $f^*(t, \cdot)$ are lower semi-continuous, convex and not identically to $+\infty$, implies that $f^\infty(t, \cdot)$ is a well-defined, lower continuous, positively homogeneous, convex function from $\mathbb{R}^n$ to $\mathbb{R} \cup \{+\infty\}$, vanishing at 0, see Section 8 in [27].

To improve the readability of the following duality theorem, we define the space

$$\mathcal{C} := \left[C^1(\mathcal{Q}_{\overline{E}^1}) \times \ldots \times C^1(\mathcal{Q}_{\overline{E}^m}) \right] \times \left[C^1(\mathcal{Q}_{V^1}) \times \ldots \times C^1(\mathcal{Q}_{V^n}) \right] \qquad (3.12)$$

and for $\phi_t^j \in C^1(\mathcal{Q}_{\overline{E}^j}), \psi_t^i \in C^1(Q_{V^i}), \boldsymbol{\phi_t} = (\phi_t^1, \ldots, \phi_t^m), \boldsymbol{\psi_t} = (\psi_t^1, \ldots, \psi_t^n)$ we define the primal functional

$$\mathcal{J}^\kappa(\boldsymbol{\phi_t}, \boldsymbol{\psi_t}) := \sum_{j=1}^m \int_{E^j} \left(\phi_1^j d\rho_1^j - \phi_0^j \, d\rho_0^j \right) + \sum_{i=1}^n (\psi_1^i \gamma_1^i - \psi_0^i \gamma_0^i)$$

$$+ \sum_{j=1}^m \iint_{\mathcal{Q}_{\overline{E}^j}} \iota_{S_{\bar{E}^j}}(\partial_t \phi_t^j, \partial_x \phi_t^j) \, \mathrm{d}x\mathrm{d}t - \sum_{i=1}^n \int_0^1 \iota_{S_{V^i}^\kappa}(\partial_t \psi_t^i, \psi_t^i, \phi_t|_{V^i}) \, \mathrm{d}t.$$

These definitions enable us to prove the duality theorem:

Theorem 3.12 (Duality theorem). *Given admissible $(\boldsymbol{\rho_0}, \boldsymbol{\rho_1}, \boldsymbol{\gamma_0}, \boldsymbol{\gamma_1})$ we have the duality*

$$\mathcal{W}_\kappa^2(\boldsymbol{\rho_0}, \boldsymbol{\rho_1}, \boldsymbol{\gamma_0}, \boldsymbol{\gamma_1}) = \sup_{\mathcal{C}} \mathcal{J}^\kappa(\boldsymbol{\phi_t}, \boldsymbol{\psi_t}) \tag{3.13}$$

where inf = min *is attained in* (3.7).

Proof. With $\partial_t \boldsymbol{\psi_t} = (\partial_t \psi_t^1, \ldots, \partial_t \psi_t^n)$ etc., we start by defining the unfolding operator

$$\begin{aligned} L\colon \quad \mathcal{C} &\to \text{Range}(L) \\ (\boldsymbol{\phi_t}, \boldsymbol{\psi_t}) &\mapsto \left(\partial_t \boldsymbol{\phi_t}, \partial_x \boldsymbol{\phi_t}; \partial_t \boldsymbol{\psi_t}, \boldsymbol{\psi_t}, (\boldsymbol{\phi_t}^1, \ldots, \boldsymbol{\phi_t}^n)\right). \end{aligned}$$

with

$$\boldsymbol{\phi_t}^i = \left(\phi_t^j\big|_{V^i}\right)_{j \in Z(V^i)}. \tag{3.14}$$

As pointwise restriction and differentiation are continuous operations, the unfolding operator is continuous for the natural topology on $\mathcal{C}$ and

$$\begin{aligned} \text{Range}(L) = &\left[C(\mathcal{Q}_{\overline{E}^1}) \times \ldots \times C(\mathcal{Q}_{\overline{E}^m})\right]^2 \times \left[C(\mathcal{Q}_{V^1}) \times \ldots \times C(\mathcal{Q}_{V^n})\right] \times \left[C^1(\mathcal{Q}_{V^1}) \times \ldots \times C^1(\mathcal{Q}_{V^n})\right] \\ &\times \left[[C^1(\mathcal{Q}_{V^1})]^{|Z(V^1)|} \times \ldots \times [C^1(\mathcal{Q}_{V^n})]^{|Z(V^n)|}\right]. \end{aligned}$$

With these definitions, we can express the primal problem as

$$\sup_{(\boldsymbol{\phi_t}, \boldsymbol{\psi_t}) \in \mathcal{C}} \left\{\mathcal{F}(\boldsymbol{\phi_t}, \boldsymbol{\psi_t}) + \mathcal{G}(L(\boldsymbol{\phi_t}, \boldsymbol{\psi_t}))\right\}$$

with

$$\mathcal{F}(\boldsymbol{\phi_t}, \boldsymbol{\psi_t}) = \sum_{j=1}^m \int_{\overline{E^j}} \left(\phi_1^j \, d\rho_1^j - \phi_0^j \, d\rho_0^j\right) + \sum_{i=1}^n \left(\psi_1^i \gamma_1^i - \psi_0^i \gamma_0^i\right)$$

and

$$\mathcal{G}(L(\boldsymbol{\phi_t}, \boldsymbol{\psi_t})) = -\sum_{j=1}^m \iint_{\mathcal{Q}_{\overline{E^j}}} \iota_{S_{\overline{E}^j}} (\partial_t \phi_t^j, \partial_x \phi_t^j) \, \mathrm{d}x\mathrm{d}t - \sum_{i=1}^n \int_0^1 \iota_{S_{V^i}^\kappa} (\partial_t \psi_t^i, \psi_t^i, \phi_t|_{V^i}) \, \mathrm{d}t.$$

Note that $\mathcal{F}(\boldsymbol{\phi_t}, \boldsymbol{\psi_t})$, defined on the product space $\mathcal{C}$, is a sum of functions that are linear continuous with respect to their corresponding variables. Moreover, each $\iota_{S_{\overline{E}^j}}$ is convex, proper and l.s.c., so both $\mathcal{F}$ and $\mathcal{G}$ are concave, proper and upper semi-continuous functionals on $\mathcal{C}$. To apply the Fenchel–Rockafellar theorem, we need a pair $(\boldsymbol{\phi_t}, \boldsymbol{\psi_t})$ such that $\mathcal{G}(L(\boldsymbol{\phi_t}, \boldsymbol{\psi_t}))$ is continuous at $L(\boldsymbol{\phi_t}, \boldsymbol{\psi_t})$ and $\mathcal{F}(\boldsymbol{\phi_t}, \boldsymbol{\psi_t})$ is finite. An example for such a pair is given by $\boldsymbol{\phi_t} = -\mathbf{1}t, \boldsymbol{\psi_t} = -t$, as

$$\mathcal{F}(-\mathbf{1}t, -t) = \sum_{j=1}^m \int_{\overline{E^j}} -d\rho_1^j + \sum_{i=1}^n -\gamma_1^i < \infty \qquad \text{and}$$

$$\mathcal{G}(L(-\mathbf{1}t, -t)) = -\sum_{j=1}^m \iint_{\mathcal{Q}_{\overline{E^j}}} \iota_{S_{\overline{E}^j}}(-1, 0) \, \mathrm{d}x\mathrm{d}t - \sum_{i=1}^n \int_0^1 \iota_{S_{V^i}^\kappa}(-1, -t, -\mathbf{1}t) \, \mathrm{d}t = 0,$$

where $\mathcal{G}$ is continuous at $\mathcal{G}(L(-\mathbf{1}t, -t))$ as $\mathcal{G}(L(-\mathbf{1}t \pm \varepsilon, -t \pm \varepsilon)) = 0$ for $\varepsilon \ll 1$. This pair is clearly a solution of the Hamilton–Jacobi equations (3.9).

Thus, the Fenchel–Rockafellar theorem 3.9 guarantees that

$$\sup_{\mathcal{C}} \mathcal{J}^\kappa(\boldsymbol{\phi_t}, \boldsymbol{\psi_t}) = \inf_{\hat{\boldsymbol{\mu}}_t \in \mathcal{C}^*} \left\{ -\mathcal{F}^*(-L^*\hat{\boldsymbol{\mu}}_t) - \mathcal{G}^*(\hat{\boldsymbol{\mu}}_t) \right\}, \tag{3.15}$$

where $-\mathcal{F}^* = (-\mathcal{F})^*, -\mathcal{G}^* = (-\mathcal{G})^*$ are the Fenchel–Legendre (convex) conjugates of the convex functional $-\mathcal{F}, -\mathcal{G}$, respectively, and $L^* \colon (\mathrm{Range}(L))^* \to \mathcal{C}^*$ is the adjoint operator of the unfolding operator L and the target dual space identifies to

$$(\mathrm{Range}(L))^* = \left[\mathcal{M}(\mathcal{Q}_{\overline{E}^1}) \times \ldots \times \mathcal{M}(\mathcal{Q}_{\overline{E}^m}) \right]^2 \times \left[\mathcal{M}(\mathcal{Q}_{V^1}) \times \ldots \times \mathcal{M}(\mathcal{Q}_{V^n}) \right]^3$$

with elements denoted by

$$\hat{\boldsymbol{\mu}}_t = (\hat{\boldsymbol{\mu}}_{E_t}, \hat{\boldsymbol{\mu}}_{V_t}) = (\boldsymbol{\rho_t}, \boldsymbol{F_t}; \boldsymbol{\gamma_t}, \boldsymbol{f_t}, \boldsymbol{\eta}) \in (\mathrm{Range}(L))^*. \tag{3.16}$$

Similar to Theorem 2 in [20], one can show that

$$-\mathcal{F}^*(-L^*\hat{\boldsymbol{\mu}}_t) = \sup_{\boldsymbol{\phi_t}, \boldsymbol{\psi_t} \in \mathcal{C}} \left\{ \mathcal{F}(\boldsymbol{\phi_t}, \boldsymbol{\psi_t}) - \langle \hat{\boldsymbol{\mu}}_t, L(\boldsymbol{\phi_t}, \boldsymbol{\psi_t}) \rangle_{(\mathrm{Range}(L))^*, \mathrm{Range}(L)} \right\} \tag{3.17}$$

$$= \begin{cases} 0 & \text{if } \begin{cases} \sum_{j=1}^m \partial_t \rho_t^j + \partial_x F_t^j = 0 \text{ in } E^j & \text{with } F_t^j \cdot \nu_{i,j} = -\eta_t^{i,j} \text{ and } \rho^j|_{t=0,1} = \rho_{0,1}^j, \\ \sum_{i=1}^n \partial_t \gamma_t^i = f_t^i \text{ on } V^i & \text{with } \gamma^i|_{t=0,1} = \gamma_{0,1}^i, \end{cases} \\ +\infty & \text{otherwise,} \end{cases} \tag{3.18}$$

where $\alpha(E^j)$ denotes the initial vertex of E^j, $\omega(E^j)$ its terminal vertex and the equations and initial-terminal/boundary conditions should be understood in the integral sense as in Definition 3.1. Moreover, for a generic element $\boldsymbol{\zeta_t} = (\boldsymbol{\alpha_t}, \boldsymbol{\beta_t}; \boldsymbol{a_t}, \boldsymbol{b_t}, \boldsymbol{c_t}) \in \mathrm{Range}(L)$, we compute

$$-\mathcal{G}^*(\hat{\boldsymbol{\mu}}_t) = \sup_{\boldsymbol{\zeta_t} \in \mathrm{Range}(L)} \left\{ \langle \hat{\boldsymbol{\mu}}_t, \boldsymbol{\zeta_t} \rangle_{(\mathrm{Range}(L))^*, \mathrm{Range}(L)} + \mathcal{G}(\boldsymbol{\zeta_t}) \right\}$$

$$= \sup_{(\boldsymbol{\alpha_t}, \boldsymbol{\beta_t})} \left\{ \sum_{j=1}^m \iint_{\mathcal{Q}_{\overline{E}^j}} \alpha_t^j \, d\rho_t^j + \sum_{j=1}^m \iint_{\mathcal{Q}_{\overline{E}^j}} \beta_t^j \cdot dF_t^j - \sum_{j=1}^m \iint_{\mathcal{Q}_{\overline{E}^j}} \iota_{S_{\bar{E}^j}}(\alpha_t^j, \beta_t^j) \, \mathrm{d}x\mathrm{d}t \right\}$$

$$+ \sup_{(\boldsymbol{a_t}, \boldsymbol{b_t}, \boldsymbol{c_t})} \left\{ \sum_{i=1}^n \int_0^1 a_t^i \gamma_t^i \, \mathrm{d}t + \sum_{i=1}^n \int_0^1 b_t^i f_t^i \, \mathrm{d}t + \sum_{i=1}^n \int_0^1 \boldsymbol{c}_t^i \cdot \boldsymbol{\eta}_t^i \, \mathrm{d}t - \sum_{i=1}^n \int_0^1 \iota_{S_{V^i}^\kappa}(a_t^i, b_t^i, \boldsymbol{c}_t^i) \, \mathrm{d}t \right\},$$

where $\boldsymbol{c}_t^i$ are defined as $\boldsymbol{c}_t^i$ in (3.14), we used that $(\boldsymbol{\alpha_t}, \boldsymbol{\beta_t})$ and $(\boldsymbol{a_t}, \boldsymbol{b_t}, \boldsymbol{c_t})$ are uncoupled. Moreover, as each summand is uncoupled to the other others, we obtain

$$= \sum_{j=1}^m \sup_{(\alpha_t^j, \beta_t^j)} \left\{ \iint_{\mathcal{Q}_{\overline{E}^j}} \alpha_t^j \, d\rho_t^j + \iint_{\mathcal{Q}_{\overline{E}^j}} \beta_t^j \cdot dF_t^j - \iint_{\mathcal{Q}_{\overline{E}^j}} \iota_{S_{\bar{E}^j}}(\alpha_t^j, \beta_t^j) \, \mathrm{d}x\mathrm{d}t \right\}$$

$$+ \sum_{i=1}^n \sup_{(a_t^i, b_t^i, \boldsymbol{c}_t^i)} \left\{ \int_0^1 a_t^i \gamma_t^i \, \mathrm{d}t + \int_0^1 b_t^i f_t^i \, \mathrm{d}t + \int_0^1 \boldsymbol{c}_t^i \cdot \boldsymbol{\eta}_t^i \, \mathrm{d}t - \int_0^1 \iota_{S_{V^i}^\kappa}(a_t^i, b_t^i, \boldsymbol{c}_t^i) \, \mathrm{d}t \right\}.$$

Applying Theorem 5 in [22] allows us to interpret the previous equation as a sum of two convex conjugates and switch the operations of convex conjugation with the integration. Next, exploiting lemma 3.10 we obtain

$$-\mathcal{G}^*(\hat{\boldsymbol{\mu}}_{\boldsymbol{t}}) = \left(\sum_{j=1}^{m} \iint_{Q_{\overline{E^j}}} A\Big(\frac{\mathrm{d}\hat{\mu}_{E^j}}{\mathrm{d}\mathcal{L}_E^j}\Big) \, d\mathcal{L}_{E^j} + \sum_{j=1}^{m} \iint_{Q_{\overline{E^j}}} A^\infty\Big(\frac{\mathrm{d}\hat{\mu}_{E^j}}{\mathrm{d}\hat{\mu}_{E^j}^S}\Big) \, d\hat{\mu}_{E^j}^S \right) + \int_{[0,1]} \overline{A}(\hat{\boldsymbol{\mu}}_{\boldsymbol{V_t}}) \, \mathrm{d}t.$$

Here, $\mathcal{L}_{E^j}$ denotes the space-time Lebesgue measure on E^j for a given edge. The measures $\hat{\mu}_{E^j}^S, \hat{\mu}_{V^i}^S$ are any non-negative measures dominating the singular parts of $|\hat{\mu}_{E^j}|, |\hat{\mu}_{V^i}|$ and A^∞ denotes the recession function of A. Since A is 1-homogeneous, its recession function is $A^\infty = A$, see definition 3.11 and [27], Corollary 8.5.2. Then, we can write

$$-\mathcal{G}^*(\hat{\boldsymbol{\mu}}_{\boldsymbol{t}}) = \sum_{j=1}^{m} \iint_{Q_{\overline{E^j}}} A\Big(\frac{\mathrm{d}\hat{\mu}_{E^j}}{\mathrm{d}\lambda_{E^j}}\Big) d\lambda_{E^j} + \int_{[0,1]} \bar{A}(\hat{\boldsymbol{\mu}}_{\boldsymbol{V_t}}) \mathrm{d}t$$

for any dominating measure $\lambda_{E^j} \gg |\hat{\mu}_{E^j}|$. Now we introduce a different segmentation for $\hat{\boldsymbol{\mu}}_{\boldsymbol{t}}$ and replace the notation given in (3.16) by

$$\hat{\boldsymbol{\mu}}_{\boldsymbol{t}} = (\boldsymbol{\mu}_{\boldsymbol{E_t}}, \boldsymbol{\mu}_{\boldsymbol{V_t}}, \boldsymbol{\eta_t}) \coloneqq (\boldsymbol{\rho_t}, \boldsymbol{F_t}; \boldsymbol{\gamma_t}, \boldsymbol{f_t}; \boldsymbol{\eta_t})$$

in order to relate $-\mathcal{G}^*(\hat{\boldsymbol{\mu}}_{\boldsymbol{t}})$ to the action functional (3.4). With this choice (and by (3.11)) we obtain

$$
\begin{aligned}
-\mathcal{G}^*(\hat{\boldsymbol{\mu}}_{\boldsymbol{t}}) &= \sum_{j=1}^{m} \iint_{Q_{\overline{E^j}}} A\Big(\frac{\mathrm{d}\mu_{E^j}}{\mathrm{d}\lambda_{E^j}}\Big) \, d\lambda_{E^j} + \sum_{i=1}^{n} \int_{[0,1]} A(\mu_{V^i}) \, \mathrm{d}t + \sum_{i=1}^{n} \begin{cases} 0 & \text{if } f_t^i + \sum_{k \in Z(V^i)} \eta_t^{i,k} = 0 \\ +\infty & \text{otherwise} \end{cases} \\
&= \mathcal{A}(\boldsymbol{\rho_0}, \boldsymbol{\rho_1}, \boldsymbol{\gamma_0}, \boldsymbol{\gamma_1}) + \sum_{i=1}^{n} \begin{cases} 0 & \text{if } f_t^i + \sum_{k \in Z(V^i)} \eta_t^{i,k} = 0 \\ +\infty & \text{otherwise} . \end{cases}
\end{aligned}
\tag{3.19}
$$

Combining (3.15), (3.17) and (3.19) yields the correct flux conditions $\sum_{j \in Z(V^i)} F_t^j \cdot \nu_{i,j} = f_t^i$ (since $F_t^j \cdot \nu_{i,j} = -\eta_t^{i,j}$ from $-\mathcal{F}^*(-L^*\hat{\boldsymbol{\mu}}_{\boldsymbol{t}}) < +\infty$ and $f_t^i + \sum_{j \in Z(V^i)} \eta_t^{i,j} = 0$ from $-\mathcal{G}^*(\hat{\boldsymbol{\mu}}_{\boldsymbol{t}}) < +\infty$), we end up with the claimed duality

$$\sup_{\mathcal{C}} \mathcal{J}^\kappa(\boldsymbol{\phi_t}, \boldsymbol{\psi_t}) = \inf_{\boldsymbol{\mu_t} \in \mathcal{CE}(\boldsymbol{\rho_0}, \boldsymbol{\rho_1}, \boldsymbol{\gamma_0}, \boldsymbol{\gamma_1})} \mathcal{A}(\boldsymbol{\mu_t}) = \mathcal{W}_\kappa^2(\boldsymbol{\rho_0}, \boldsymbol{\rho_1}, \boldsymbol{\gamma_0}, \boldsymbol{\gamma_1}).$$

This completes the proof as $\mathcal{W}_\kappa^2$ is finite as shown in Lemma 3.8 and the Fenchel–Rockafellar theorem assures that the supremum in (3.13) is attained. $\qquad\square$

As expected the functional $\mathcal{W}_\kappa^2$ forms a metric. Using Theorem 3.12 and Prop. 3.6 (iii), the proof of the following proposition is analogue to Proposition 3.10 in [20].

Proposition 3.13 ([20], Prop. 3.10). *The quantity $\mathcal{W}_\kappa^2(\boldsymbol{\rho_0}, \boldsymbol{\rho_1}, \boldsymbol{\gamma_0}, \boldsymbol{\gamma_1})$, defined in 3.7, is a distance on $\mathcal{P}(\mathcal{G})$, i.e. the set of probability measures on the network $\mathcal{G}$.*

4. LIMITING CASES AND RELATIONSHIP TO OTHER METRICS

In this section we want to explore the influence of κ on the transport problem putting special emphasis on the limiting case $\kappa \to \infty$. This case translates to fact that the costs for transporting mass over the vertices goes to infinity.

The first step is to analyse the relationship between our metric $\mathcal{W}_\kappa^2$ and some other metrics on the graph. There are several special metrics we can consider, *e.g.* separate metrics on edges and nodes. Above we have already introduced the bounded Lipschitz distances on edges and nodes. The simplest case related to our metric is the Fisher–Rao metric $\mathcal{FR}_\kappa$ on nodes, *i.e.*,

$$\mathcal{FR}_\kappa^2(\boldsymbol{\gamma_0}, \boldsymbol{\gamma_1}) := \min \mathcal{A}^{\mathcal{FR}} = \min \left\{ \sum_{i=1}^n \int_0^1 \kappa^2 \frac{|f_t^i|^2}{2\gamma_t^i} \, \mathrm{d}t \quad \text{s.t.} \quad \partial_t \gamma_t^i = f_t^i \text{ in } V^i \right\}.$$

Existence of minimisers for the Fisher–Rao distance is given in Theorem 3.1 of [8] which allows us to write 'min' instead of 'inf'.

If the masses in the edges are compatible, *i.e.* $\|\rho_0^j\| = \|\rho_1^j\|$, for $j = 1, \ldots, m$, we can also consider the the classical Wasserstein distance on each edge $\mathcal{W}_{E^j}^2$ as the problem completely decouples to

$$\sum_{j=1}^m \mathcal{W}_{E^j}^2(\rho_0^j, \rho_1^j) = \sum_{j=1}^m \min \left\{ \iint_{\overline{E^j} \times [0,1]} \frac{|F_t^j|^2}{2\rho_t^j} \, \mathrm{d}x \mathrm{d}t \quad \text{s.t.} \quad \begin{array}{ll} \partial_t \rho_t^j + \partial_x F_t^j = 0 & \text{in } E^j, \\ F_t^j \cdot \nu_{i,j} = 0 & \text{in } \partial \overline{E^j} \end{array} \right\}. \tag{4.1}$$

In the case of incompatible masses on individual edges, but compatible overall edge mass, *i.e.*

$$\sum_j \|\rho_0^j\| = \sum_j \|\rho_1^j\|$$

we can introduce a Wasserstein metric on the edges only which, at the nodes, are connected via Kirchhoff's law, *i.e.*

$$\mathcal{W}_{\mathcal{E}}^2(\boldsymbol{\rho_0}, \boldsymbol{\rho_1}) = \min \left\{ \sum_{j=1}^m \iint_{\overline{E^j} \times [0,1]} \frac{|F_t^j|^2}{2\rho_t^j} \, \mathrm{d}x \mathrm{d}t \quad \text{s.t.} \quad \begin{array}{ll} \partial_t \rho_t^j + \partial_x F_t^j = 0 & \text{in } E^j, \\ \sum_{j \in Z(V^i)} F_t^j \cdot \nu_{i,j} = 0 & \text{in } \partial \overline{E^j}, \end{array} \right\}$$

The existence of a minimizer can be shown analogous to the original Wasserstein metric, since Kirchhoff's law directly allows a weak formulation of the constraint equation on the whole network of edges. Note that for comparison the original distance with $\kappa = 0$,

$$\mathcal{W}_0^2(\varsigma_0, \varsigma_1) := \min \left\{ \iint_{\mathcal{G} \times [0,1]} \frac{|H|^2}{2\varsigma} \, \mathrm{d}x \mathrm{d}t \quad \text{s.t.} \quad \begin{array}{ll} \partial_t \varsigma + \partial_x H = h & \text{in } \Omega_{\mathcal{G}} \times [0,1], \\ \sum_{j \in Z(V^i)} F_t^j \cdot \nu_{i,j} = f_t^i & \text{in } \Omega_{\mathcal{V}} \times [0,1]. \end{array} \right\},$$

with ς_0, ς_1 defined in (2.1) and H, h as in Proposition 3.3.

Proposition 4.1 (Sandwich theorem adapted to Proposition 5.1 in [20]). *For any admissible initial and terminal network concentrations $(\boldsymbol{\rho_0}, \boldsymbol{\rho_1}, \boldsymbol{\gamma_0}, \boldsymbol{\gamma_1})$, it holds that*

$$(i) \qquad\qquad\qquad \mathcal{FR}_\kappa^2(\boldsymbol{\gamma_0}, \boldsymbol{\gamma_1}) \leq \mathcal{W}_\kappa^2(\boldsymbol{\rho_0}, \boldsymbol{\rho_1}, \boldsymbol{\gamma_0}, \boldsymbol{\gamma_1}),$$

$$(ii) \qquad \sum_{j=1}^m d_{BL,\overline{E^j}}(\rho_s^j, \rho_t^j) + \sum_{i=1}^n d_{BL,V^i}(\gamma_s^i, \gamma_t^i) \leq C_\kappa \mathcal{W}_\kappa(\boldsymbol{\rho_0}, \boldsymbol{\rho_1}, \boldsymbol{\gamma_0}, \boldsymbol{\gamma_1}) \qquad and$$

if the node masses are compatible, i.e. $\boldsymbol{\gamma_0} = \boldsymbol{\gamma_1}$,

$$(iii) \qquad\qquad\qquad \mathcal{W}_\kappa^2(\boldsymbol{\rho_0}, \boldsymbol{\rho_1}, \boldsymbol{\gamma_0}, \boldsymbol{\gamma_1}) \leq \mathcal{W}_{\mathcal{E}}^2(\boldsymbol{\rho_0}, \boldsymbol{\rho_1}).$$

Proof. (i) As proven in Theorem 3.12, we know that $\mathcal{W}_\kappa^2$ geodesics exist and we pick any geodesic $\hat{\boldsymbol{\mu}}_{\boldsymbol{t}} = (\hat{\boldsymbol{\mu}}_{\boldsymbol{E}}, \hat{\boldsymbol{\mu}}_{\boldsymbol{V}}) = (\boldsymbol{\rho_t}, \boldsymbol{F_t}; \boldsymbol{\gamma_t}, \boldsymbol{f_t}, \boldsymbol{\eta})$ connecting $\boldsymbol{\rho_0}$ and $\boldsymbol{\rho_1}$ and by definition $\mathcal{A}(\hat{\boldsymbol{\mu}}_{\boldsymbol{t}}) = \mathcal{W}_\kappa^2$ holds. For $i \in \{1, \dots, n\}$ each component μ_{V^i} of the vector $\hat{\boldsymbol{\mu}}_{\boldsymbol{V}}$ solves the continuity equations $\partial_t \gamma_t^i = f_t^i$, thus connecting γ_0^i and γ_1^i on V^i, and is therefore admissible for the action $\mathcal{A}^{\mathcal{FR}}$ of the Fisher–Rao metric. Thus

$$\mathcal{FR}_\kappa^2(\boldsymbol{\gamma_0}, \boldsymbol{\gamma_1}) = \min_{\hat{\boldsymbol{\mu}}_{\boldsymbol{V}}{}'} \mathcal{A}^{\mathcal{FR}}(\hat{\boldsymbol{\mu}}_{\boldsymbol{V}}{}') \leq \mathcal{A}^{\mathcal{FR}}(\hat{\boldsymbol{\mu}}_{\boldsymbol{V}}) \leq \mathcal{A}(\hat{\boldsymbol{\mu}}_{\boldsymbol{t}}) = \mathcal{W}_\kappa^2(\boldsymbol{\rho_0}, \boldsymbol{\rho_1}, \boldsymbol{\gamma_0}, \boldsymbol{\gamma_1}) \,,$$

which shows the assertion.

(ii) Applying Theorem 3.12, we choose a geodesic $\boldsymbol{\mu_t} \in \mathcal{CE}(\boldsymbol{\rho_0}, \boldsymbol{\rho_1}, \boldsymbol{\gamma_0}, \boldsymbol{\gamma_1})$. By Proposition 3.6 (iii) we obtain

$$\sum_{j=1}^m d_{\mathrm{BL}, \overline{E^j}}(\rho_s^j, \rho_t^j) + \sum_{i=1}^n d_{\mathrm{BL}, V^i}(\gamma_s^i, \gamma_t^i) \leq C_\kappa \sqrt{\mathcal{A}(\boldsymbol{\mu_t})} |t - s|^{\frac{1}{2}} = C_\kappa \mathcal{W}_\kappa(\boldsymbol{\rho_0}, \boldsymbol{\rho_1}, \boldsymbol{\gamma_0}, \boldsymbol{\gamma_1}) |t - s|^{\frac{1}{2}} \,.$$

(iii) Pick an interior Wasserstein geodesic $(\boldsymbol{\rho_t}, \boldsymbol{F_t})$ and the trivial boundary geodesic $(\boldsymbol{\gamma_0}, \boldsymbol{0})$. Together these geodesics form a solution of the generalised continuity equation in the sense of Definition 3.1. Consequently, $\boldsymbol{\mu_t} \coloneqq (\boldsymbol{\rho_t}, \boldsymbol{F_t}, \boldsymbol{\gamma_t}, \boldsymbol{0})$ forms an admissible candidate in 3.7 and we obtain

$$\mathcal{W}_\kappa^2(\boldsymbol{\rho_0}, \boldsymbol{\rho_1}, \boldsymbol{\gamma_0}, \boldsymbol{\gamma_1}) \leq \mathcal{A}(\boldsymbol{\mu_t}) = \mathcal{W}_{\mathcal{E}}^2(\boldsymbol{\rho_0}, \boldsymbol{\rho_1}) \,.$$

$\square$

We can also show that, unless the initial and final mass on the vertices is zero, every geodesic is such that for every edge with unequal initial and final mass, the mass of at least one of the attached vertices must change.

Proposition 4.2. *For any admissible initial and terminal network concentrations* $(\boldsymbol{\rho_0}, \boldsymbol{\rho_1}, \boldsymbol{\gamma_0}, \boldsymbol{\gamma_1})$ *such that* $\boldsymbol{\gamma_0} = \boldsymbol{\gamma_1} \neq 0$. *Then for every edge* j *with* $\rho_0^j \neq \rho_1^j$ *either* $f^{\bar{\alpha}(j)}$ *or* $f^{\bar{\omega}(j)}$ *must be nonzero.*

Proof. We argue by contradiction and use that the dual potentials $(\boldsymbol{\phi_t}, \boldsymbol{\psi_t})$ are (smooth) subsolutions of the Hamilton–Jacobi equations (3.9) and satisfy the relations

$$\frac{F_t^j}{\rho_t^j} = \partial_x \phi_t^j \quad \text{and} \quad \kappa^2 \frac{f_t^i}{\gamma_t^i} = \psi_t^i - \frac{1}{|Z(V_i)|} \sum_{j \in Z(V_i)} \phi_t^j,$$

for $i = 1, \dots, n$. Fix and edge $j^* \in \{1, \dots, m\}$ with $\rho_0^j \neq \rho_1^j$ and assume $f_t^{\bar{\alpha}(j^*)} = 0$ as well as $f_t^{\bar{\omega}(j^*)} = 0$ for a.e. $t \in (0, 1)$. This implies, for $i^* \in \{\bar{\alpha}(j^*), \bar{\omega}(j^*)\}$, that $\gamma_t^{i^*} = C$ with $C > 0$ due to $\gamma_0^{i^*} = \gamma_1^{i^*} \neq 0$. Together with

$$0 = \kappa^2 f_t^{i^*} = \gamma_t^{i^*} \left(\psi_t^{i^*} - \frac{1}{|Z(V_{i^*})|} \sum_{j \in Z(V_{i^*})} \phi_t^j \right), \tag{4.2}$$

this yields $\psi_t^{i^*} - \frac{1}{|Z(V_{i^*})|} \sum_{j \in Z(V_{i^*})} \phi_t^j = 0$. Inserting this into (3.9) yields $\partial_t \psi_t^{i^*} = 0$ and, differentiating (4.2) in time, we obtain

$$0 = \partial_t \sum_{j \in Z(V_{i^*})} \phi_t^j = \frac{1}{2} \sum_{j \in Z(V_{i^*})} |\partial_x \phi_t^j|^2.$$

Therefore the fluxes F_t^j vanish for all edges j connected to either one of the two vertices $\bar{\alpha}(j^*)$ and $\bar{\omega}(j^*)$. In particular $F_t^{j^*} \cdot \nu_{i^*, j^*} = 0$ for $i^* \in \{\bar{\alpha}(j^*), \bar{\omega}(j^*)\}$ meaning that on j^*, the continuity equation has no-flux boundary conditions. This is a contradiction to $\rho_0^{j^*} \neq \rho_1^{j^*}$. $\square$

Proposition 4.3 (adapted to [20], Prop. 6.1)**.** *For fixed $(\boldsymbol{\rho_0}, \boldsymbol{\rho_1}, \boldsymbol{\gamma_0}, \boldsymbol{\gamma_1})$ the map $\kappa \mapsto \mathcal{W}_\kappa^2(\boldsymbol{\rho_0}, \boldsymbol{\rho_1}, \boldsymbol{\gamma_0}, \boldsymbol{\gamma_1})$ is non-decreasing.*

Proof. Note that the set of smooth subsolutions

$$
S^\kappa := \left\{ (\boldsymbol{\phi_t}, \boldsymbol{\psi_t}) \in \mathcal{C}: \quad \sum_{j=1}^m \partial_t \phi_t^j + \frac{1}{2} |\partial_x \phi_t^j|^2 \leq 0 \quad \text{and} \quad \sum_{i=1}^n \partial_t \psi_t^i + \frac{1}{2\kappa^2} \left| \psi_t^i - \frac{1}{|Z(V^i)|} \sum_{j \in Z(V^i)} \phi_t^j \right|^2 \leq 0 \right\}
$$

is non-decreasing in κ. The monotonicity immediately follows from the duality in Theorem 3.12, *i.e.*

$$
\mathcal{W}_\kappa^2(\boldsymbol{\rho_0}, \boldsymbol{\rho_1}, \boldsymbol{\gamma_0}, \boldsymbol{\gamma_1}) = \sup_{(\boldsymbol{\phi_t}, \boldsymbol{\psi_t}) \in S^\kappa} \left\{ \sum_{j=1}^m \int_{\overline{E^j}} \left(\phi_1^j \, d\rho_1^j - \phi_0^j \, d\rho_0^j \right) + \sum_{i=1}^n \psi_1^i \gamma_1^i - \psi_0^i \gamma_0^i \right\} .
$$

$\square$

Now, let us finally discuss the limit case $\kappa \to \infty$. In this case, transporting mass over the vertices becomes more and more costly and, intuitively, we expect that in the limit transporting mass over the vertices is prohibited.

In more detail, we recover two interesting scenarios depending on the initial concentration and the terminal concentration on the network. The first scenario is the case where the masses are incompatible, *i.e.* $\|\rho_0^j\| \neq \|\rho_1^j\|$ for one $1 \leq j \leq m$ or $\boldsymbol{\gamma_0^i} \neq \boldsymbol{\gamma_1^i}$ for a given $1 \leq i \leq n$. More precisely, this means that there exists at least one edge or one vertex where the corresponding initial mass is different than the terminal one. As mass cannot be transported over the vertices, we expect that the distance functional $\mathcal{W}_\kappa^2$ goes to infinity in this case. In the second scenario the masses are compatible. As we shall see, the distance functional will converge to a sum of classical Wasserstein metrics on each edge. In order to proof this assumption, we will need a lemma for each scenario. We start with the lemma for the incompatible scenario:

Lemma 4.4 (adapted to [20], Prop. 6.2)**.** *For any initial and terminal network concentrations $(\boldsymbol{\rho_0}, \boldsymbol{\rho_1}, \boldsymbol{\gamma_0}, \boldsymbol{\gamma_1})$ and any geodesic $(\boldsymbol{\rho_t}, \boldsymbol{F_t}; \boldsymbol{\gamma_t}, \boldsymbol{f_t})$, there holds*

$$
\mathcal{W}_\kappa^2(\boldsymbol{\rho_0}, \boldsymbol{\rho_1}, \boldsymbol{\gamma_0}, \boldsymbol{\gamma_1}) \geq \frac{\kappa^2}{2} \sum_{i=1}^n \|f_t^i\|^2 \geq \frac{\kappa^2}{2} \sum_{i=1}^n \left| \gamma_1^i - \gamma_0^i \right|^2 .
$$

Proof. Applying Theorem 3.12 allows us to choose a geodesic $\hat{\boldsymbol{\mu_t}} = (\boldsymbol{\rho_t}, \boldsymbol{F_t}; \boldsymbol{\gamma_t}, \boldsymbol{f_t}, \boldsymbol{\eta_t})$ from $\boldsymbol{\rho_0}$ to $\boldsymbol{\rho_1}$ and by definition we obtain

$$
\mathcal{W}_\kappa^2(\boldsymbol{\rho_0}, \boldsymbol{\rho_1}, \boldsymbol{\gamma_0}, \boldsymbol{\gamma_1}) = \mathcal{A}(\boldsymbol{\mu_t}) \geq \frac{\kappa^2}{2} \sum_{i=1}^n \int_0^1 \frac{|f_t^i|^2}{\gamma_t^i} \, dt \geq \frac{\kappa^2}{2} \sum_{i=1}^n \left(\int_0^1 f_t^i \, dt \right)^2 = \frac{\kappa^2}{2} \sum_{i=1}^n \|f_t^i\|^2 .
$$

where we used Proposition 3.6 (ii), *i.e.* $\gamma_t^i \leq \|\varsigma\| = 1$, as ς is a probability measure, and Jensen's inequality. The continuity equation $\partial_t \gamma_t^i = f_t^i$ finally controls the mass difference as (again using Jensen)

$$
\|f_t^i\| \geq \left| \int_0^1 1 \, df_t^i \right| \overset{\text{Def. 3.1}}{=} \left| \gamma_1^i - \gamma_0^i \right| .
$$

$\square$

The following result helps to solve weighted optimisation problems and will be applied in the proof of the next proposition in the case of compatible masses:

Lemma 4.5 ([20], Lem. 6.3). *Let K be a compact set, take $\mathfrak{f}, \mathfrak{g} \colon K \to \mathbb{R}^+ \cup \{+\infty\}$ two proper, lower semi-continuous functions, and consider*

$$\mathfrak{h}_\kappa(x) := \mathfrak{f}(x) + \kappa^2 \mathfrak{g}(x), \qquad \kappa > 0.$$

Assume that for all κ there is a minimiser $x_\kappa \in K$ of $\mathfrak{h}_\kappa$. Then as $\kappa \to +\infty$ any cluster point x_* of $\{x_\kappa\}$ minimises $\mathfrak{f}$ in $\operatorname{argmin} \mathfrak{g}$.

Proposition 4.6 (adapted to [20], Thm. 6). *For fixed $(\boldsymbol{\rho_0}, \boldsymbol{\rho_1}, \boldsymbol{\gamma_0}, \boldsymbol{\gamma_1})$ we obtain*

$$\mathcal{W}_\kappa^2(\boldsymbol{\rho_0}, \boldsymbol{\rho_1}, \boldsymbol{\gamma_0}, \boldsymbol{\gamma_1}) \underset{\kappa \,\to\, +\infty}{\to} \begin{cases} \mathcal{W}_\mathcal{E}^2(\boldsymbol{\rho_0}, \boldsymbol{\rho_1}) & \text{if } \boldsymbol{\gamma_0} = \boldsymbol{\gamma_1}, \\ +\infty & \text{otherwise,} \end{cases} \tag{4.3}$$

Let moreover $\boldsymbol{\mu_t}^\kappa = (\boldsymbol{\rho_t}^\kappa, \boldsymbol{F_t}^\kappa; \boldsymbol{\gamma_t}^\kappa, \boldsymbol{f_t}^\kappa)$ be any $\mathcal{W}_\kappa^2(\boldsymbol{\rho_0}, \boldsymbol{\rho_1}, \boldsymbol{\gamma_0}, \boldsymbol{\gamma_1})$-geodesic. If the node masses are compatible, i.e. $\boldsymbol{\gamma_0} = \boldsymbol{\gamma_1}$, $\|\rho_0^j\| = \|\rho_1^j\|$ for all $j \in \{1, \ldots, m\}$ then up to a subsequence

$$(\boldsymbol{\rho_t}^\kappa, \boldsymbol{F_t}^\kappa) \to (\boldsymbol{\rho_t}, \boldsymbol{F_t}), \quad \text{and} \quad \|f_t^\kappa\| \to 0$$

where $(\boldsymbol{\rho_t}, \boldsymbol{F_t})$ is a $\mathcal{W}_\mathcal{E}^2$-geodesic .

Proof. For incompatible node masses, *i.e.* $\boldsymbol{\gamma_0} \neq \boldsymbol{\gamma_1}$, we conclude from Lemma 4.4 that

$$\mathcal{W}_\kappa^2(\boldsymbol{\rho_0}, \boldsymbol{\rho_1}, \boldsymbol{\gamma_0}, \boldsymbol{\gamma_1}) \underset{\kappa \,\to\, +\infty}{\to} +\infty \ .$$

In the case of compatible node masses, we can apply Proposition 4.1 and Lemma 4.4 to control the boundary flux term via

$$\sum_{i=1}^n \|f^\kappa\|^2 \leq \frac{2}{\kappa^2} \mathcal{W}_\kappa^2(\boldsymbol{\rho_0}, \boldsymbol{\rho_1}, \boldsymbol{\gamma_0}, \boldsymbol{\gamma_1}) \leq \frac{2}{\kappa^2} \mathcal{W}_\mathcal{E}^2(\boldsymbol{\rho_0}, \boldsymbol{\rho_1}) \to 0.$$

To control the momentum $\boldsymbol{F}^\kappa$, we observe from Proposition 4.1 that any geodesic satisfies

$$\frac{1}{2} \sum_{j=1}^m \iint_{\overline{E^j} \times [0,1]} \frac{|F_t^{j,\kappa}|^2}{\rho_t^{j,\kappa}} \, \mathrm{d}x \mathrm{d}t \leq \mathcal{W}_\kappa^2(\boldsymbol{\rho_0}, \boldsymbol{\rho_1}, \boldsymbol{\gamma_0}, \boldsymbol{\gamma_1}) \leq \mathcal{W}_\mathcal{E}^2(\boldsymbol{\rho_0}, \boldsymbol{\rho_1}) \tag{4.4}$$

uniformly in $\kappa > 0$. With $\boldsymbol{F_t}^\kappa = \boldsymbol{u_t}^\kappa \boldsymbol{\rho_t}^\kappa$ and using Jensen's inequality

$$\|\boldsymbol{F_t}^\kappa\|^2 = \sum_{j=1}^m \Big(\iint_{\overline{E^j} \times [0,1]} |u_t^\kappa| \mathrm{d}\rho_t^\kappa \Big)^2 \leq \sum_{j=1}^m \iint_{\overline{E^j} \times [0,1]} |u_t^\kappa|^2 \, \mathrm{d}\rho_t^\kappa = \sum_{j=1}^m \iint_{\overline{E^j} \times [0,1]} \frac{|F_t^{j,\kappa}|^2}{\rho_t^{j,\kappa}} \, \mathrm{d}x \mathrm{d}t \ .$$

Therefore $\boldsymbol{F_t}^\kappa$ is bounded by $\|\boldsymbol{F_t}^\kappa\| \leq \mathcal{W}_\mathcal{E}^2(\boldsymbol{\rho_0}, \boldsymbol{\rho_1}) = C$ using (4.4), where C does not dependent on κ. Moreover $\boldsymbol{\rho_t}^\kappa, \boldsymbol{\omega_t}^\kappa \leq 1$, and the geodesic $\boldsymbol{\mu_t}^\kappa$ is bounded.

Our aim now is to apply Lemma 4.5, so we need compactness which is guaranteed by Prokhorov's Theorem: In a Polish space, bounded measures are tight ($\mathcal{G}$ is a Polish space as a closed subset of the Polish space $\mathbb{R}^2$). Thus we can apply Prokhorov's theorem and we obtain the narrow compactness $(\boldsymbol{\rho_{t}}_t^\kappa, \boldsymbol{F_{t}}_t^\kappa; \boldsymbol{\gamma_{t}}_t^\kappa, \boldsymbol{f_{t}}_t^\kappa) \overset{*}{\rightharpoonup} (\boldsymbol{\rho_t}, \boldsymbol{F_t}; \boldsymbol{\gamma_t}, \boldsymbol{0})$ up to subsequences, and we only have to proof that the limit $(\boldsymbol{\rho_t}, \boldsymbol{F_t})$ is a $\mathcal{W}_\mathcal{E}^2$-geodesic.

We now use Lemma 4.5 and set K to be the set of all geodesics for all values of $\kappa > 1$, which is narrowly compact by the previous discussion and because the linear continuity equations (3.1) and (3.2) are stable under

narrow limits. The functions

$$\mathfrak{f}(c, \boldsymbol{f_t}) := \frac{1}{2} \sum_{j=1}^{m} \iint_{\overline{E^j} \times [0,1]} \frac{|F_t^j|^2}{\rho_t^j} \, \mathrm{d}x \mathrm{d}t \quad \text{and} \quad \mathfrak{g}(\boldsymbol{\rho_t}, \boldsymbol{F_t}, \boldsymbol{\gamma_t}, \boldsymbol{f_t}) := \frac{1}{2} \sum_{i=1}^{n} \int_0^1 \frac{f_t^{i\,2}}{\gamma_t^i} \, \mathrm{d}t$$

are convex, proper, semi-continuous w.r.t. the narrow convergence of measures [4], Theorem 3.3, and geodesics are minimisers of $\mathcal{A} = \mathfrak{f} + \kappa^2 \mathfrak{g}$. The definition of $\frac{f_t^{i\,2}}{\gamma_t^i}$ in the extended sense implies that the minimisers of $\mathfrak{g}$ are solutions of the continuity equations (3.1) and (3.2) of the form $\boldsymbol{\mu_t} = (\boldsymbol{\rho_t}, \boldsymbol{F_t}; \boldsymbol{\gamma_t}, \boldsymbol{0})$ and assign the value $\iint \frac{f_t^{i\,2}}{\gamma_t^i} = 0$ to $\mathfrak{g}$. This set of solutions of $\mathcal{CE}(\boldsymbol{\rho_0}, \boldsymbol{\rho_1}, \boldsymbol{\gamma_0}, \boldsymbol{\gamma_1})$ with $f_t^i = 0$ obviously identifies with the whole set of pairs $(\boldsymbol{\rho_t}, \boldsymbol{F_t})$ of independent solutions of $\partial_t \rho_t^j + \partial_x F_t^j$ for all j with Kirchhoff's law

$$\partial_t \gamma_t^i = \sum_{j \in Z(V^i)} F_t^j \cdot \nu_{i,j} = 0$$

at the boundary. Consequently, the limit-geodesic $(\boldsymbol{\rho_t}, \boldsymbol{F_t})$ is a geodesic for $\mathcal{W}_{\mathcal{E}}^2$.

Finally, let us analyse the convergence in distance (4.3). Applying Proposition 4.1 (iii) and the lower semi-continuity of the actions with $(\boldsymbol{\rho_t}^\kappa, \boldsymbol{F_t}^\kappa) \overset{*}{\rightharpoonup} (\boldsymbol{\rho_t}, \boldsymbol{F_t})$ gives that

$$\limsup_{\kappa \to +\infty} \mathcal{W}_\kappa^2(\boldsymbol{\rho_0}, \boldsymbol{\rho_1}, \boldsymbol{\gamma_0}, \boldsymbol{\gamma_1}) \leq \mathcal{W}_{\mathcal{E}}^2(\boldsymbol{\rho_0}, \boldsymbol{\rho_1}) = \sum_{j=1}^{m} \iint_{\overline{E^j} \times [0,1]} \frac{|F_t^j|^2}{2\rho_t^j} \, \mathrm{d}x \mathrm{d}t \leq \liminf_{\kappa \to +\infty} \sum_{j=1}^{m} \iint_{\overline{E^j} \times [0,1]} \frac{|F_t^{j,\kappa}|^2}{2\rho_t^{j,\kappa}} \, \mathrm{d}x \mathrm{d}t$$

$$\leq \liminf_{\kappa \to +\infty} \Big(\iint_{\overline{E^j} \times [0,1]} \frac{|F_t^{j,\kappa}|^2}{2\rho_t^\kappa} \, \mathrm{d}x \mathrm{d}t + \int_0^1 \kappa^2 \frac{|f_t^{j,\kappa}|^2}{2\gamma_t^\kappa} \, \mathrm{d}t \Big) = \liminf_{\kappa \to +\infty} \mathcal{W}_\kappa^2(\boldsymbol{\rho_0}, \boldsymbol{\rho_1}, \boldsymbol{\gamma_0}, \boldsymbol{\gamma_1}),$$

where we used that $(\boldsymbol{\rho_t}^\kappa, \boldsymbol{F_t}^\kappa; \boldsymbol{\gamma_t}^\kappa, \boldsymbol{f_t}^\kappa)$ is a geodesic in the last equality. Thus, $\liminf = \limsup = \lim$ in the previous chain of inequalities, which finishes the proof. $\square$

5. GRADIENT FLOWS

Given the metric structure it appears natural to study corresponding gradient flows, which opens various questions for future research beyond the scope of this paper. Here we will only provide a formal derivation of the gradient flow equations for energy functionals of the form

$$\mathcal{E}(\boldsymbol{\rho}, \boldsymbol{\gamma}) = \sum_{j=1}^{m} \mathcal{G}_j(\rho^j) + \sum_{j=1}^{n} \mathcal{H}_i(\gamma^i). \tag{5.1}$$

The gradient flow structures can be derived in a standard way from a minimizing movement scheme, as in [1, 24], which constructs a time-discrete sequence $(\boldsymbol{\rho}^{\tau,k}, \boldsymbol{\gamma}^{\tau,k})$, $k \in \mathbb{N}$, whose iterates are obtained by minimizing the functional

$$\frac{1}{2\tau} \mathcal{W}_\kappa^2(\boldsymbol{\rho}^\tau, \boldsymbol{\rho}^{\tau,k-1}, \boldsymbol{\gamma}^\tau, \boldsymbol{\gamma}^{\tau,k-1}) + \mathcal{E}(\boldsymbol{\rho}^\tau, \boldsymbol{\gamma}^\tau)$$

with respect to $(\boldsymbol{\rho}^\tau, \boldsymbol{\gamma}^\tau)$ for a given pair $(\boldsymbol{\rho}^{\tau,k-1}, \boldsymbol{\gamma}^{\tau,k-1})$, By a formal limit procedure this yields that the gradient flow satisfies the transport problem with fluxes

$$F_t^j = \rho_t^j \partial_x \phi_t^j$$

$$f_t^i = \kappa^{-2}\gamma_t^i(\psi_t^i - \frac{1}{|Z(V_i)|}\sum_{j\in Z(V_i)}\phi_t^j)$$

and dual potentials

$$\phi_t^j = -\mathcal{G}_j'(\rho_t^j), \qquad \psi_t^i = -\mathcal{H}_i'(\gamma_t^i).$$

This implies that the gradient flows are of the form

$$\partial_t\rho_t^j = \partial_x(\rho_t^j\partial_x\mathcal{G}_j'(\rho_t^j))$$

$$\partial_t\gamma_t^i = -\kappa^{-2}\gamma_t^i(\mathcal{H}_i'(\gamma_t^i) - \frac{1}{|Z(V_i)|}\sum_{j\in Z(V_i)}\mathcal{G}_j'(\rho_t^j))$$

with the Kirchhoff condition

$$\sum_{j\in Z(V_i)}\rho_t^j\partial_x\mathcal{G}_j'(\rho_t^j)\cdot\nu_{i,j} = \kappa^{-2}\gamma_t^i(\mathcal{H}_i'(\gamma_T^i) - \frac{1}{|Z(V_i)|}\sum_{j\in Z(V_i)}\mathcal{G}_j'(\rho_t^j))$$

and the continuity

$$\mathcal{G}_j'(\rho_t^j) = \mathcal{G}_k'(\rho_t^k) \qquad \forall j,k\in Z(V_i).$$

A surprising effect of the coupling condition is the one-sided coupling of the variations of the edge energies, whose trace appears in the vertex equations. On the other hand we have the standard form as in Wasserstein gradient flows on the edges.

Example 5.1. Let us consider the standard case of a drift-diffusion equation on the edges, *i.e.*

$$\mathcal{G}_j(\rho^j) = \int_{E^j}(\rho^j\log\rho^j + \rho^jW_j)\,\mathrm{d}x$$

with some potential W_j, and some cost on the vertex concentrations $\mathcal{H}_i(\gamma^i) = h_i(\gamma^i)$.

$$\partial_t\rho_t^j = \partial_{xx}\rho_t^j + \partial_x(\rho_t^jW_j')$$

$$\partial_t\gamma_t^i = -\kappa^{-2}\gamma_t^i(h_i'(\gamma_t^i) - \frac{1}{|Z(V_i)|}\sum_{j\in Z(V_i)}(1+\log\rho_t^j+W_j))$$

with the Kirchhoff condition

$$\sum_{j\in Z(V_i)}(\partial_x\rho_t^j + \rho_t^jW_j)\cdot\nu_{i,j} = \kappa^{-2}\gamma_t^i(\langle_i'(\gamma_T^i) - \frac{1}{|Z(V_i)|}\sum_{j\in Z(V_i)}(1+\log\rho_t^j+W_j)).$$

The continuity condition in the vertices becomes $\log\rho_t^j + W_j = \log\rho_t^k + W_k$, which can be reformulated as the linear transmission condition

$$\rho_t^j\mathrm{e}^{W_j} = \rho_t^k\mathrm{e}^{W_k} \qquad \forall j,k\in Z(V_i).$$

APPENDIX A. FORMAL DERIVATION OF THE FIRST ORDER OPTIMALITY CONDITIONS

Based on the Lagrange calculus we present a formal derivation of the optimality conditions of (2.3) which include the Hamilton–Jacobi equations already introduced in (3.9).

Using the notation $\boldsymbol{\phi_t} = (\phi_t^1, \ldots, \phi_t^m)$ as well as $\boldsymbol{\psi_t} = (\psi_t^1, \ldots, \psi_t^n)$ and $\boldsymbol{\lambda_t} = (\lambda_t^1, \ldots, \lambda_t^n)$, we introduce the Lagrange functional

$$
\begin{aligned}
&\mathcal{L}(\boldsymbol{\rho_t}, \boldsymbol{\gamma_t}, \boldsymbol{F_t}, \boldsymbol{f_t}, \boldsymbol{\phi_t}, \boldsymbol{\psi_t}, \boldsymbol{\lambda_t}) \\
&= \sum_{j=1}^m \iint_{\overline{E^j} \times [0,1]} \frac{|F_t^j|^2}{2\rho_t^j} \, \mathrm{d}x\mathrm{d}t + \kappa^2 \sum_{i=1}^n \int_{[0,1]} \frac{|f_t^i|^2}{2\gamma_t^i} \, \mathrm{d}t + \sum_{j=1}^m \iint_{\overline{E^j} \times [0,1]} (\partial_t \rho_t^j + \partial_x F_t^j)\phi_t^j \, \mathrm{d}x\mathrm{d}t \\
&\quad + \sum_{i=1}^n \int_0^1 (\partial_t \gamma_t^i - f_t^i)\psi_t^i \, \mathrm{d}t + \sum_{i=1}^n \int_0^1 (f_t^i - \sum_{j \in Z(V^i)} F_t^j(V_i)\nu_{i,j})\lambda_t^i \, \mathrm{d}t \\
&= \sum_{j=1}^m \iint_{\overline{E^j} \times [0,1]} \frac{|F_t^j|^2}{2\rho_t^j} \, \mathrm{d}x\mathrm{d}t + \kappa^2 \sum_{i=1}^n \int_{[0,1]} \frac{|f_t^i|^2}{2\gamma_t^i} \, \mathrm{d}t + \sum_{j=1}^m \iint_{\overline{E^j} \times [0,1]} (-\rho_t^j \partial_t \phi_t^j - F_t^j \partial_x \phi_t^j) \, \mathrm{d}x\mathrm{d}t \\
&\quad + \sum_{j=1}^m \int_{\overline{E^j}} (\rho_1^j \phi_t^j(x,1) - \rho_0^j \phi_t^j(x,0)) \, \mathrm{d}x + \sum_{i=1}^n \sum_{j \in Z(V_i)} \int_0^1 F_t^j(V_i)\nu_{ij}\phi_t^j \, \mathrm{d}t \\
&\quad + \sum_{i=1}^n \int_0^1 (-\gamma_t^i \partial_t \psi_t^i - f_t^i \psi_t^i) \, \mathrm{d}t + \sum_{i=1}^n (\gamma_1^i \psi_t^i(1) - \gamma_0^i \psi_t^i(0)) \\
&\quad + \sum_{i=1}^n \int_0^1 (f_t^i - \sum_{j \in Z(V^i)} F_t^j(V_i)\nu_{i,j})\lambda_t^i \, \mathrm{d}t.
\end{aligned}
$$

Calculating the gradient of $\mathcal{L}$ into direction $\boldsymbol{\varphi} = [\boldsymbol{\varphi_{\rho_t}}, \boldsymbol{\varphi_{F_t}}, \boldsymbol{\varphi_{f_t}}, \boldsymbol{\varphi_{\gamma_t}}]$ yields

$$
\begin{aligned}
&\nabla_{\boldsymbol{\rho_t}, \boldsymbol{F_t}, \boldsymbol{f_t}, \boldsymbol{\gamma_t}} \mathcal{L}[\boldsymbol{\varphi}] \\
&= \sum_{j=1}^m \iint_{\overline{E^j} \times [0,1]} \left(-\frac{|F_t^j|^2}{2(\rho_t^j)^2} - \partial_t \phi_t^j \right)\varphi_{\rho_t^j} \, \mathrm{d}x\mathrm{d}t + \sum_{j=0}^m \int_{\overline{E^j}} (\varphi_{\rho_t^j}(x,1)\phi_t^j(x,1) - \varphi_{\rho_t^j}(x,0)\phi_t^j(x,0)) \, \mathrm{d}x \\
&\quad + \sum_{j=1}^m \iint_{\overline{E^j} \times [0,1]} \left(\frac{F_t^j}{\rho_t^j} - \partial_x \phi_t^j \right)\varphi_{F_t^j} \, \mathrm{d}x\mathrm{d}t + \sum_{i=1}^n \sum_{j \in Z(V_i)} \int_0^1 (\phi_t^j + \lambda_t^i)\varphi_{F_t^j}\nu_{ij} \, \mathrm{d}t \\
&\quad + \sum_{i=1}^n \int_0^1 \left(\kappa^2 \frac{f_t^i}{\gamma_t^i} - \psi_t^i + \lambda_t^i \right)\varphi_{f_t^i} \, \mathrm{d}t \\
&\quad + \sum_{i=1}^n \int_0^1 \left(-\kappa^2 \frac{|f_t^i|^2}{2(\gamma_t^i)^2} - \partial_t \psi_t^i \right)\varphi_{\gamma_t^i} \, \mathrm{d}t + \sum_{i=1}^n (\varphi_{\gamma_1^i}\psi_t^i(1) - \varphi_{\gamma_0^i}\psi_t^i(0)) \,.
\end{aligned}
$$

Choosing the direction $\boldsymbol{\varphi} = [\boldsymbol{0}, \boldsymbol{\varphi_{F_t}}, \boldsymbol{0}, \boldsymbol{0}]$ s.t. $F_t^i \nu_{ij} = 0$ for all $i \in \{0, \ldots, n\}$, $j \in Z(V_i)$ as well as $\boldsymbol{\varphi} = [\boldsymbol{0}, \boldsymbol{0}, \boldsymbol{\varphi_{f_t}}, \boldsymbol{0}]$ and setting the respective derivative to zero we obtain

$$
\frac{F_t^j}{\rho_t^j} - \partial_x \phi_t^j = 0 \quad \text{and} \quad \kappa^2 \frac{f_t^i}{\gamma_t^i} = \psi_t^i - \lambda_t^i,
$$

for $j = 1, \ldots m$ and $i = 1, \ldots n$. From $\boldsymbol{\varphi} = [\mathbf{0}, \boldsymbol{\varphi_{F_t}}, \mathbf{0}, \mathbf{0}]$ s.t. $F_t^i \nu_{ij} = 0$ s.t. $\varphi_{F_t^i} \nu_{ij} = \varphi_{F_t^i} \nu_{ij'}$ for all $j, j' \in Z(V_i)$ we have

$$\lambda_t^i = -\frac{1}{|Z(V_i)|} \sum_{j \in Z(V_i)} \phi_t^j, \quad i = 1, \ldots, n.$$

Finally, the choices $\boldsymbol{\varphi} = [\boldsymbol{\varphi_{\rho_t}}, \mathbf{0}, \mathbf{0}, \mathbf{0}]$, supported in the interior, as well as $\boldsymbol{\varphi} = [\mathbf{0}, \mathbf{0}, \mathbf{0}, \boldsymbol{\varphi_{\gamma_t}}]$ result in the equations

$$\partial_t \phi_t^j + \frac{1}{2} |\partial_x \phi_t^j|^2 = 0, \quad \text{and} \quad \partial_t \psi_t^i + \frac{1}{2\kappa^2} \left| \psi_t^i - \frac{1}{|Z(V_i)|} \sum_{j \in Z(V_i)} \phi_t^j \right|^2 = 0.$$

This corresponds exactly to (3.9).

References

[1] L. Ambrosio, N. Gigli and G. Savaré, Gradient Flows in Metric Spaces and in the Space of Probability Measures, 2nd ed. Lectures in Mathematics ETH Zürich. Birkhäuser Verlag, Basel (2008).

[2] M.K. Banda, M. Herty and A. Klar, Gas flow in pipeline networks. *Networks Heterogeneous Media* **1** (2006) 41.

[3] J.-D. Benamou and Y. Brenier, A computational fluid mechanics solution to the Monge–Kantorovich mass transfer problem. *Numer. Math.* **84** (2000) 375–393.

[4] G. Bouchitté and G. Buttazzo, New lower semicontinuity results for nonconvex functionals defined on measures. *Nonlinear Anal. Theory Methods Appl.* **15** (1990) 679–692.

[5] L. Brasco, A survey on dynamical transport distances. *J. Math. Sci.* **181** (2012) 755–781.

[6] A. Bressan, S. Čanić, M. Garavello, M. Herty and B. Piccoli, Flows on networks: recent results and perspectives. *EMS Surv. Math. Sci.* **1** (2014) 47–111.

[7] J.A. Carrillo, S. Lisini, G. Savaré and D. Slepčev, Nonlinear mobility continuity equations and generalized displacement convexity. *J. Funct. Anal.* **258** (2010) 1273–1309.

[8] L. Chizat, G. Peyré, B. Schmitzer and F.-X. Vialard, An Interpolating Distance between Optimal Transport and Fisher-Rao. Foundations of Computational Mathematics. Springer Verlag (2010).

[9] L. Chizat, G. Peyré, B. Schmitzer and F.-X. Vialard, Unbalanced optimal transport: dynamic and Kantorovich formulations. *J. Funct. Anal.* **274** (2018) 3090–3123.

[10] M. Erbar, D. Forkert, J. Maas and D. Mugnolo, Gradient flow formulation of diffusion equations in the wasserstein space over a metric graph, 2021.

[11] T.T. Georgiou, J. Karlsson and M.S. Takyar, Metrics for power spectra: An axiomatic approach. *IEEE Trans. Signal Process.* **57** (2009) 859–867.

[12] D.B. Hill, M.J. Plaza, K. Bonin and G. Holzwarth, Fast vesicle transport in pc12 neurites: velocities and forces. *Eur. Biophys. J.* **33** (2004) 623–632.

[13] I. Humpert, D. Di Meo, A.W. Püschel and J.-F. Pietschmann, On the role of vesicle transport in neurite growth: Modeling and experiments. *Math. Biosci.* **338** (2021) 108632.

[14] S. Kondratyev, L. Monsaingeon and D. Vorotnikov, A new optimal transport distance on the space of finite Radon measures. *Adv. Differ. Equ.* **21** (2016) 1117–1164.

[15] M. Liero, A. Mielke and G. Savaré, Optimal transport in competition with reaction: the Hellinger-Kantorovich distance and geodesic curves. *SIAM J. Math. Anal.* **48** (2016) 2869–2911.

[16] M. Liero, A. Mielke and G. Savaré, Optimal entropy-transport problems and a new Hellinger—Kantorovich distance between positive measures. *Invent. Math.* **211** (2018) 969–1117.

[17] S. Lisini and A. Marigonda, On a class of modified wasserstein distances induced by concave mobility functions defined on bounded intervals. *Manuscr. Math.* **133** (2020) 197–224.

[18] J.M. Mazón, J.D. Rossi and J. Toledo, Optimal mass transport on metric graphs. *SIAM J. Optim.* **25** (2015) 1609–1632.

[19] P. Mindt, J. Lang and P. Domschke, Entropy-preserving coupling of hierarchical gas models. *SIAM J. Math. Anal.* **51** (2019) 4754–4775.

[20] L. Monsaingeon, A new transportation distance with bulk/interface interactions and flux penalization. *Calc. Variations Partial Differ. Equ.* **60** (2021).

[21] B. Piccoli and F. Rossi, Generalized wasserstein distance and its application to transport equations with source. *Arch. Rational Mech. Anal.* **211** (2013) 335–358.

[22] R.T. Rockafellar, Integrals which are convex functionals. II. *Pac. J. Math.* **39** (1971) 439–469.

[23] F. Santambrogio, Optimal Transport for Applied Mathematicians: Calculus of Variations, PDEs, and Modeling. Progress in Nonlinear Differential Equations and Their Applications. Birkhäuser Basel (2015).

[24] F. Santambrogio, {Euclidean, metric, and Wasserstein} gradient flows: an overview. *Bull. Math. Sci.* **7** (2017) 87–154.

[25] K. Tsaneva-Atanasova, N. Azzopardi, T. Galli and D. Holcman, Modeling vesicle trafficking and neurite growth. *Biophys. J.* (2007) 115A–115A.

[26] R. Tyrrell Rockafellar, Duality and stability in extremum problems involving convex functions. *Pac. J. Math.* **21** (1967) 167–187.

[27] R. Tyrrell Rockafellar, Convex Analysis. Princeton University Press (1970).

[28] C. Villani, Optimal Transport: Old and New. Grundlehren der mathematischen Wissenschaften. Springer Berlin Heidelberg (2008).

[29] C. Villani, Topics in Optimal Transportation, Vol. 58 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI (2003).

[30] A. Yadaw, M. Siddiq, V. Rabinovich, R. Tolentino, J. Hansen and R. Iyengar, Dynamic balance between vesicle transport and microtubule growth enables neurite outgrowth. *PLOS Computat. Biol.* **15** (2019) e1006877.