Euro. Jnl of Applied Mathematics (2020), *vol.* 31, *pp.* 450–469 © Cambridge University Press 2019.
doi:10.1017/S0956792519000123

Nonlinear systems coupled through multi-marginal transport problems

M. LABORDE

Department of Mathematics and Statistics, McGill University, Montreal, Canada
email: maxime.laborde@mcgill.ca

(*Received 2 October 2018*; *revised 14 March 2019*; *accepted 19 March 2019*; *first published online 29 April 2019*)

In this paper, we introduce a dynamical urban planning model. This leads to the study of a system of nonlinear equations coupled through multi-marginal optimal transport problems. A first example consists in solving two equations coupled through the solution to the Monge–Ampère equation. We show that the Wasserstein gradient flow theory provides a very good framework to solve these highly nonlinear systems. At the end, a uniqueness result is presented in dimension one based on convexity arguments.

Key words: Wasserstein gradient flows, multi-marginal optimal transport, urban planning

2010 Mathematics Subject Classification: 35K40, 49J45 (Primary); 24K65, 35Q91 (Secondary)

1 Introduction

Recently, Kinderlehrer et al. proposed in [18] a gradient flow approach to solve the Poisson–Nernst–Planck system

$$\begin{cases} \partial_t \rho_1 - \alpha \Delta \rho_1^m - \operatorname{div}(\rho_1 \nabla (V_1 + \varphi)) = 0, \\ \partial_t \rho_2 - \beta \Delta \rho_2^m - \operatorname{div}(\rho_2 \nabla (V_2 - \varphi)) = 0, \\ -\Delta \varphi = \rho_1 - \rho_2. \end{cases}$$

This system is, for instance, used to model ionic transport of several interacting species. Inspired by this work, we are interested in a 'nonlinear' version where species ρ_1 and ρ_2 are coupled through the Monge–Ampère equation instead of the Poisson equation,

$$\begin{cases} \partial_t \rho_1 - \alpha \Delta \rho_1^m - \operatorname{div}\big(\rho_1 \nabla (V_1 + \varphi)\big) = 0, \\ \partial_t \rho_2 - \beta \Delta \rho_2^m - \operatorname{div}\big(\rho_2 \nabla (V_2 + \varphi^c)\big) = 0, \\ \rho_2 (Id - \nabla \varphi) \det \big[I - D^2 \varphi\big] = \rho_1, \end{cases} \tag{1.1}$$

where α, β are nonnegative constants, $m \geqslant 1$ is a fixed nonlinear diffusion exponent, V_1, V_2 are two given external potentials sufficiently smooth which may depend on ρ_1 and ρ_2 and φ^c is the c-transform of φ, $\varphi^c(x) = \sup_y |x-y|^2 - \varphi(y)$ and $|x|^2 - \varphi$ is convex. In the sequel, we will see that this kind of systems can arise naturally in evolution model for urban planning.

The modelling of optimal urban planning can be very challenging. Several mathematical models have been investigated to describe the equilibrium structure of residential and service areas

https://doi.org/10.1017/S0956792519000123 Published online by Cambridge University Press

in a city in the economic literature, see for instance the textbook of Fujita [16] and the paper by Lucas and Rossi-Hansberg [20]. Then, Carlier and Ekeland proposed a new equilibrium model using the theory of optimal transportation in order to overcome some difficulties appearing in [20] especially in the shape of cities [9, 10]. The recent advances in the optimal transportation theory have given new perspectives to study the structure of a city. In [6, 7, 24, 25], Buttazzo and Santambrogio introduced new variational models to obtain optimal distributions of important elements of the structure of urban regions, such as residential and service areas. Here the equilibrium condition is not considered which is the big difference with models proposed in [9]. The optimal behaviour comes from the fact that the distributions of residents and services optimise a total welfare functional. We want the densities that satisfy several criteria which are included in this functional. In particular, it can be composed by a transportation term from residential areas to service areas, a congestion term which penalises the concentration of inhabitants and a concentration term for services in order to increase efficiency and decrease management costs.

Specifically, these models consist in studying distributions of residents and services, given by two probability measures ρ_1 and ρ_2, on a given urban region $\Omega \subset \mathbb{R}^2$, which minimise the total cost functional, $\mathcal{E}(\rho_1, \rho_2)$, in order to reach an ideal organisation. Particularly, the cost functional $\mathcal{E}$ can be taken as

$$\mathcal{E}(\rho_1, \rho_2) = W_c(\rho_1, \rho_2) + \mathcal{F}(\rho_1) + \mathcal{G}(\rho_2), \tag{1.2}$$

where W_c is the optimal transport problem between ρ_1 and ρ_2 with cost c and $c(x, y)$ is the commuting cost to move from residences at x to services at y. Several interpretations may be given to this cost. For example, it might represent the gas cost paid by workers to reach service areas and then workers want to live close to services in order to decrease car travel. The congestion term is represented by $\mathcal{F}$, an internal energy given by a convex superlinear function F,

$$\begin{cases} \int_\Omega F(\rho(x))\,dx & \text{if } F(\rho) \in L^1(\Omega), \\ +\infty & \text{otherwise.} \end{cases}$$

Since F is superlinear and convex, $\mathcal{F}$ can be rewritten as

$$\mathcal{F}(\rho) = \int_\Omega \frac{F(\rho)}{\rho}\rho,$$

where $\rho \mapsto \frac{F(\rho)}{\rho}$ is an increasing function which can be seen as the unhappiness of a citizen when he lives in a place where the population density is ρ.

Finally, the concentration term, $\mathcal{G}$, is of the form

$$\mathcal{G}(\rho) = \frac{1}{2}\iint_{\Omega\times\Omega} h(|x-y|)\,d\rho(x)d\rho(y),$$

where h is an increasing function modelling interactions between different services.

Notice that other models with different transport problems have been studied. For example, in [12], the optimal transport problem is replaced by a transportation cost taking into account congestion effects. This term is given by the squared distance norm of $\rho_1 - \rho_2$ in the dual space of a subspace of $H^1(\Omega)$.

However, since nowadays we have to deal with existing cities that are not configurated optimally outright, it seems natural to study how ρ_1 and ρ_2 evolve in time starting from the existing

https://doi.org/10.1017/S0956792519000123 Published online by Cambridge University Press

distributions in order to decrease $\mathcal{E}$. This leads to study the gradient flow of $\mathcal{E}$ in a Wasserstein product space. This theory was introduced by Jordan et al. in [17] and use the notion of minimising movements by De Giorgi [13] in the space of probability measures endowed with the Wasserstein metric. For an energy $\mathcal{S} : \mathcal{P}(\Omega) \to \mathbb{R} \cup \{+\infty\}$, it consists in constructing a time step approximation using an implicit Euler scheme: given a small time step $h > 0$ and an initial data $\rho_0 \in \mathcal{P}(\Omega)$, we define by induction a sequence $(\rho_h^k)_{k\in\mathbb{N}}$ as $\rho_h^0 := \rho_0$ and for all $k \geqslant 0$,

$$\rho_h^{k+1} = \text{argmin}_{\rho\in\mathcal{P}(\Omega)} \frac{1}{2h} W_2^2(\rho_h^k, \rho) + \mathcal{S}(\rho),$$

where W_2 represents the 2-Wasserstein distance ($c(x, y) = |x - y|^2$). This scheme is called JKO scheme. Then, at least formally, the constant piecewise interpolation in time of $(\rho_h^k)_{k\in\mathbb{N}}$, ρ_h, converges when $h \searrow 0$ to a solution to

$$\partial_t \rho - \text{div}\left(\rho \nabla \frac{\delta \mathcal{S}}{\delta \rho}(\rho)\right) = 0,$$

where $\frac{\delta \mathcal{S}}{\delta \rho}$ is the first variation of $\mathcal{S}$ [3, 17, 26]. In particular, if $\mathcal{S}$ is of the form

$$\mathcal{S}(\rho) := \int_\Omega F(\rho(x))\, dx + \int_\Omega V(x)\rho(x)\, dx + \frac{1}{2} \iint_{\Omega\times\Omega} K(x - y)\rho(x)\rho(y)\, dxdy,$$

then the first variation of $\mathcal{S}$ is $\frac{\delta \mathcal{S}}{\delta \rho} = F'(\rho) + V + K * \rho$.

In the case where c is the quadratic cost, for example according to [26], the first variation of the optimal transport problem is

$$\frac{\delta W_2^2(\cdot, \mu)}{\delta \rho}(\rho) = \varphi \text{ and } \frac{\delta W_2^2(\rho, \cdot)}{\delta \mu}(\mu) = \psi,$$

where (φ, ψ) is a pair of Kantorovich potentials associated with $W_2(\rho, \mu)$ and then formally,

$$\frac{\delta \mathcal{E}(\cdot, \rho_2)}{\delta \rho_1}(\rho_1) = \varphi + F'(\rho_1) \text{ and } \frac{\delta \mathcal{E}(\rho_1, \cdot)}{\delta \rho_2}(\rho_2) = \psi + K * \rho_2,$$

where $K(x) = h(|x|)$. And since $\psi = \varphi^c$ and the optimal transport map from ρ to μ, $\mathrm{T} = \mathrm{Id} - \nabla\varphi$, satisfies the Monge–Ampère equation

$$\rho_2(\mathrm{T})\, \det[\mathrm{DT}] = \rho_1,$$

see Section 2, then at least formally, the gradient flow of $\mathcal{E}$ is a solution to a system of the form (1.1) with $F(z) = \frac{\alpha z^m}{m-1}$, $\beta = 0$, $V_1 = 0$ and $V_2 = K * \rho_2$.

In this paper, we propose to investigate a generalisation of problem (1.1). We extend it to more than two populations, and thus the transport problem becomes a multi-marginal transport problem. In addition, the cost that workers want to minimise can be different from the one of services or firms. Indeed, they have to take into account the gas cost to reach their work, whereas this cost is not relevant for services. Thus it is natural to assume that each population wants to minimise a transport problem with its own cost. Therefore, the system is not a gradient flow anymore, and we will use a semi-implicit JKO scheme introduced in [14] to deal with these different costs.

https://doi.org/10.1017/S0956792519000123 Published online by Cambridge University Press

The organisation of the paper is as follows. Section 2 recalls the results from Optimal Transport and Multi-Marginal Transport theories. In Section 3, we specify our problem and state our main result. Section 4 is devoted to the demonstration of the existence of solutions to system (3.3)–(3.5) which is a generalisation of the evolution problem (1.1) to multi-marginal transport problems with different costs. The proof is based on a semi-implicit JKO scheme and on an extension of the Aubin–Lions lemma in order to obtain strong regularity. At the end of Section 5, by convexity arguments, we give a uniqueness result in dimension one for some class of functionals.

2 Preliminaries

In the sequel, Ω represents a smooth open bounded subset of $\mathbb{R}^n$.

2.1 Wasserstein space

For a detailed exposition, we refer to reference textbooks [3, 26–28]. We denote $\mathcal{M}^+(\Omega)$ the set of nonnegative finite Radon measures on Ω, $\mathcal{P}(\Omega)$ the space of probability measures on Ω and $\mathcal{P}^{ac}(\Omega)$ the subset of $\mathcal{P}(\Omega)$ of probability measures on Ω absolutely continuous with respect to the Lebesgue measure. For all $\rho, \mu \in \mathcal{P}(\Omega)$, we denote $\Pi(\rho, \mu)$, the set of probability measures on $\Omega \times \Omega$ having ρ and μ as first and second marginals, respectively. If $\gamma \in \Pi(\rho, \mu)$, then γ is called a transport plan between ρ and μ.
For all $\rho, \mu \in \mathcal{P}(\Omega)$, we denote by $W_2(\rho, \mu)$ the Wasserstein distance between ρ and μ,

$$W_2^2(\rho, \mu) = \inf \left\{ \iint_{\Omega\times\Omega} |x-y|^2 \, d\gamma(x,y) \, : \, \gamma \in \Pi(\rho, \mu) \right\}.$$

Since this optimal transportation problem is a linear problem under linear constraints, it admits a dual formulation given by

$$W_2^2(\rho, \mu) = \sup \left\{ \int_\Omega \varphi(x) \, d\rho(x) + \int_\Omega \psi(y) \, d\mu(y) \, : \, \varphi(x) + \psi(y) \leqslant |x-y|^2 \right\}.$$

Optimal solutions of the dual problem are called Kantorovich potentials between ρ and μ. A pair of Kantorovich potentials (φ, ψ) is such that ψ is the c-conjugate of φ where c is the quadratic cost, i.e., $\psi = \varphi^c := \inf_y |\cdot - y|^2 - \varphi(y)$.

If $\rho \in \mathcal{P}^{ac}(\Omega)$, Brenier proves in [5] that the optimal transport plan γ is unique and induced by an optimal transport map T, i.e., γ is of the form $(Id \times \mathrm{T})_{\#}\rho$, where $\mathrm{T}_{\#}\rho = \mu$ and T is the gradient of a convex function. Moreover, the optimal transport map is given by $\mathrm{T} = \mathrm{Id} - \nabla\varphi$ where φ is a Kantorovich potential between ρ and μ and T satisfies the Monge–Ampère equation

$$\mu(\mathrm{T}) \det[\mathrm{DT}] = \rho.$$

It is well known that $\mathcal{P}(\Omega)$ endowed with the Wasserstein distance defines a metric space and W_2 metrizes the narrow convergence of probability measures.

Finally, if $\boldsymbol{\rho} = (\rho_1, \ldots, \rho_l)$ and $\boldsymbol{\mu} = (\mu_1, \ldots, \mu_l)$ are in $\mathcal{P}(\Omega)^l$, we define the product distance by

$$\boldsymbol{W_2}(\boldsymbol{\rho}, \boldsymbol{\mu}) = \left(\sum_{i=1}^{l} W_2^2(\rho_i, \mu_i) \right)^{1/2}.$$

https://doi.org/10.1017/S0956792519000123 Published online by Cambridge University Press

2.2 Multi-marginal transportation problem

In this section, we recall some results from the multi-marginal transport theory that we will use in the sequel. We refer, for instance, to [15, 22] for a complete survey on this topic. The usual optimal transport can be extended to several marginals $\rho_1, \dots, \rho_l \in \mathcal{P}(\Omega)$. Let c be a cost function from Ω^l to $\mathbb{R}$, the multi-marginal transport problem, $\mathcal{W}_c$, is defined by

$$\mathcal{W}_c(\rho_1, \dots, \rho_l) := \inf \left\{ \int_{\Omega^l} c(x_1, \dots, x_l)\, d\lambda(x_1, \dots, x_l) \,:\, \lambda \in \Pi(\rho_1, \dots, \rho_l) \right\},$$

where $\Pi(\rho_1, \dots, \rho_l) := \left\{ \lambda \in \mathcal{P}(\Omega^l) \,:\, {\pi^i}_{\#}\lambda = \rho_i \right\}$ and π^i denotes the canonical projection from Ω^l to Ω. By standard arguments, the existence of an optimal transport plan is guaranteed as in the two marginals case. Then, if we assume that c is continuous on $\overline{\Omega}^l$, the following dual formulation holds:

$$\mathcal{W}_c(\rho_1, \dots, \rho_l) = \sup \left\{ \sum_{i=1}^{l} \int_{\Omega} u_i(x_i)\, d\rho_i(x_i) \,:\, \sum_{i=1}^{l} u_i(x_i) \leqslant c(x_1, \dots, x_l) \right\}.$$

Any optimal $u_1, \dots, u_l$ for the dual formulation are called Kantorovich potentials and are c-conjugate functions, i.e.,

$$u_i(x_i) = \inf \left\{ c(x_1, \dots, x_l) - \sum_{j=1, j\neq i}^{l} u_j(x_j),\ x_j \in \Omega \right\}, \text{ for all } i = 1, \dots, l.$$

For any λ optimal transport plan and $u_1, \dots, u_l$ Kantorovich potentials, we get

$$\sum_{i=1}^{l} u_i(x_i) = c(x_1, \dots, x_l), \qquad \lambda - \text{a.e.}$$

In addition, assuming that ρ_i is absolutely continuous with respect to the Lebesgue measure and c is differentiable in the i-th variable, then u_i is a Lipschitz function and

$$\nabla u_i(x_i) = \nabla_{x_i} c(x_1, \dots, x_l), \qquad \lambda - \text{a.e.} \tag{2.1}$$

3 Assumptions and main result

In the following, we assume that we have $l > 1$ different populations. The congestion functional associated with the population ρ_i is given by

$$\mathcal{F}_i(\rho) := \begin{cases} \int_\Omega F_i(\rho(x))\, dx & \text{if } F_i(\rho) \in L^1(\Omega), \\ +\infty & \text{otherwise,} \end{cases} \tag{3.1}$$

where $F_i : \mathbb{R}_+ \to \mathbb{R}$ is a strictly convex superlinear function of class $\mathcal{C}^2$. Defining $P_i(x) := xF_i'(x) - F_i(x)$ the pressure associated with F_i, we assume

$$F_i(0) = 0 \text{ and } P_i(x) \leqslant C(1 + F_i(x)). \tag{3.2}$$

The typical examples of energies are $F(\rho) := \rho \log(\rho)$, which gives a linear diffusion driven by the Laplacian, and $F(\rho) := \rho^m$ $(m > 1)$, which corresponds to the porous medium diffusion.

https://doi.org/10.1017/S0956792519000123 Published online by Cambridge University Press

The multi-marginal interaction energy $\mathcal{W}_i : \mathcal{P}(\Omega)^l \to \mathbb{R}$ is defined by

$$\mathcal{W}_i(\rho_1, \dots, \rho_l) := \inf \left\{ \int_{\Omega^l} c_i(x_1, \dots, x_l)\, d\lambda(x_1, \dots, x_l) \,:\, \lambda \in \Pi(\rho_1, \dots, \rho_l) \right\},$$

where the cost function $c_i : \Omega^l \to \mathbb{R}$ is assumed to be continuous on $\overline{\Omega}^l$ and differentiable with respect to x_i such that $\nabla_{x_i} c_i$ is continuous on $\overline{\Omega}^l$ and bounded on $\overline{\Omega}^l$.

Example 3.1 (Barycentre) Assume $l = 3$ and ρ_1 evolves minimising at each step the functional

$$(\rho_1, \rho_2, \rho_3) \mapsto \alpha W_2^2(\rho_1, \rho_2) + \beta W_2^2(\rho_1, \rho_3).$$

That means that ρ_1 wants to reach the barycentre in the Wasserstein space of ρ_2, ρ_3 with weight $\alpha, \beta > 0$, see [2].

This functional can be rewritten as the multi-marginal problem

$$\inf_{\gamma \in \Pi(\rho_1, \rho_2, \rho_3)} \int_{\Omega} c(x, y, z)\, d\gamma(x, y, z),$$

where $c(x, y, z) = \alpha |x - y|^2 + \beta |x - z|^2$ satisfies the assumptions above.

The goal of this paper is to study the existence and uniqueness of solution to the following nonlinear diffusion system with nonlocal interactions

$$\begin{cases} \partial_t \rho_i = \Delta P_i(\rho_i) + \operatorname{div}(\rho_i \nabla u_i) \ \text{ in } \mathbb{R}^+ \times \Omega, \\ \rho_{i|t=0} = \rho_{i,0}, \end{cases} \qquad i \in \{1, \dots, l\}, \tag{3.3}$$

where u_i is an optimal Kantorovich potential of

$$\mathcal{W}_i(\rho_1, \dots, \rho_l) := \inf \left\{ \int_{\Omega^l} c_i(x_1, \dots, x_l)\, d\lambda(x_1, \dots, x_l) \,:\, \lambda \in \Pi(\rho_1, \dots, \rho_l) \right\}. \tag{3.4}$$

Since Ω is a bounded subset of $\mathbb{R}^n$, (3.3) is supplemented with Neumann boundary conditions on $\partial\Omega$,

$$\big(\nabla P_i(\rho_i) + \rho_i \nabla u_i\big) \cdot \nu = 0 \qquad \text{on } \mathbb{R}^+ \times \partial\Omega, \tag{3.5}$$

where ν is the outward normal to $\partial\Omega$.

To simplify the exposition, we do not treat potentials or nonlocal interactions in (3.4) even if this can be added easily.

The main difficulty is to handle the nonlinear cross term $\operatorname{div}(\rho_i \nabla u_i)$. However, we remark that if λ_i is an optimal transport plan in (3.4) and ρ_i is absolutely continuous with respect to the Lebesgue measure, then by (2.1),

$$\nabla u_i(t, x_i) = \nabla_{x_i} c_i(x_1, \dots, x_l), \qquad \lambda_i(t) - \text{a.e.} \tag{3.6}$$

Consequently, for all $\Phi \in \mathcal{C}_c^\infty([0, +\infty) \times \mathbb{R}^n)$,

$$\int_0^{+\infty} \int_\Omega \rho_i(t, x) \nabla u_i(t, x) \cdot \nabla \Phi(t, x)\, dx dt = \int_0^{+\infty} \int_{\Omega^l} \nabla_{x_i} c_i(x_1, \dots, x_l) \cdot \nabla \Phi(t, x_i)\, d\lambda_i(t, x_1, \dots, x_l) dt,$$

https://doi.org/10.1017/S0956792519000123 Published online by Cambridge University Press

since $\lambda_i(t)$ solves $\mathcal{W}_i(\rho_1(t), \dots, \rho_l(t))$, t-a.e., and (3.6). Since the right-hand side is a linear term with respect to λ_i, it is easier to work with this one, and then we define a weak solution to (3.3)–(3.5) in the following way.

Definition 3.2 A weak solution to (3.3)–(3.5) is a curve $t \in (0, +\infty) \mapsto (\rho_1(t), \dots, \rho_l(t)) \in \mathcal{P}^{ac}(\Omega)^l$ such that $\nabla P_i(\rho_i) \in \mathcal{M}^n((0, T) \times \Omega)$, for all $T < +\infty$, and

$$\int_0^{+\infty} \left(\int_\Omega \partial_t \Phi \rho_i \, dx - \int_\Omega \nabla \Phi \cdot d \nabla P_i(\rho_i) \right) dt$$
$$- \int_0^{+\infty} \int_{\Omega^l} \nabla_{x_i} c_i(x_1, \dots, x_l) \cdot \nabla \Phi(t, x_i) \, d\lambda_i(t, x_1, \dots, x_l) \, dt = - \int_\Omega \Phi(0, x) \rho_{i,0}(x) \, dx,$$

for every $\Phi \in \mathcal{C}_c^\infty([0, +\infty) \times \mathbb{R}^n)$, where $\lambda_i(t)$ is an optimal transport plan of $\mathcal{W}_i(\rho_1(t), \dots, \rho_l(t))$, t-a.e.

Our main result is the following:

Theorem 3.3 *If $\rho_{1,0}, \dots, \rho_{l,0} \in \mathcal{P}^{ac}(\Omega)$ satisfy*

$$\mathcal{F}_i(\rho_{i,0}) < +\infty, \text{ for all } i \in \{1, \dots l\}, \tag{3.7}$$

where $\mathcal{F}_i$ is given by (3.1), then (3.3)–(3.5) admit at least one weak solution.

Remark 3.4 To simplify the analysis, we assume that each population has an individual diffusion. This implies that solutions are absolutely continuous with respect to the Lebesgue measure and then Kantorovich potentials are Lipschitz and (3.6) holds. However, Theorem 3.3 can be generalised replacing ∇u_i by

$$U_i(x_i) = \int_{\Omega^{l-1}} \nabla_{x_i} c_i(x_1, \dots, x_l) \, d\lambda_i^{x_i}(\boldsymbol{x_{-i}}),$$

where $\boldsymbol{x_{-i}} = (x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_l)$ and $\lambda_i^{x_i}$ is obtained by disintegrating the optimal transport plan λ_i with respect to ρ_i, $\lambda_i = \lambda_i^{x_i} \otimes \rho_i$.

4 Existence result

The proof of Theorem 3.3 is based on a variant of the well-known JKO scheme introduced by Jordan et al. [17]. We construct by induction with a semi-implicit Euler scheme l sequences $(\rho_{i,h}^k)_{k \in \mathbb{N}} \subset \mathcal{P}^{ac}(\Omega)$, where $h > 0$ is a given time step. Since the cost of the multi-marginal functional $\mathcal{W}_i$ depends on the density i, system (3.3)–(3.5) is not a gradient flow in the Wasserstein product space.

To deal with this issue, we introduce the functional $\overline{\mathcal{W}}_i(\cdot|\boldsymbol{\mu})$, where $\boldsymbol{\mu} = (\mu_1, \dots, \mu_l)$, defined by

$$\overline{\mathcal{W}}_i(\rho|\boldsymbol{\mu}) := \mathcal{W}_i(\mu_1, , \mu_{i-1}, \rho, \mu_{i+1}, \dots, \mu_l).$$

In other words, $\overline{\mathcal{W}}_i(\rho|\boldsymbol{\mu})$ is the multi-marginal problem with marginals $\mu_1, \dots, \mu_{i-1}, \rho$, $\mu_{i+1}, \dots, \mu_l$.

https://doi.org/10.1017/S0956792519000123 Published online by Cambridge University Press

Sequences $(\rho_{i,h}^k)_{k\in\mathbb{N}}$ are then constructed using the following semi-implicit JKO scheme: for all $i \in [\![1, I]\!]$, $\rho_{i,h}^0 = \rho_{i,0}$ and for all $k \geqslant 0$, $\rho_{i,h}^{k+1}$ minimises

$$\mathcal{E}_{i,h}(\rho|\boldsymbol{\rho}_h^k) := W_2^2(\rho, \rho_{i,h}^k) + 2h\left(\mathcal{F}_i(\rho) + \overline{\mathcal{W}}_i(\rho|\boldsymbol{\rho}_h^k)\right), \tag{4.1}$$

on $\rho \in \mathcal{P}^{ac}(\Omega)$, where $\boldsymbol{\rho}_h^k := (\rho_{1,h}^k, \ldots, \rho_{I,h}^k)$. At each step, all the marginals are frozen except the i-th marginal in the functional (3.4). These sequences are well defined by standard compactness and lower semi-continuity arguments.

Now define the piecewise constant interpolations by $\rho_{i,h}(0) := \rho_{i,0}$, and for all $t > 0$,

$$\rho_{i,h}(t) := \rho_{i,h}^{k+1} \text{ if } t \in (hk, h(k+1)]. \tag{4.2}$$

Let $\lambda_{i,h}^{k+1}$ be an optimal transport plan for $\mathcal{W}_i\left(\rho_{1,h}^k, \ldots, \rho_{i-1,h}^k, \rho_{i,h}^{k+1}, \rho_{i+1,h}^k, \ldots, \rho_{I,h}^k\right)$ and $\lambda_{i,h}$ be the piecewise constant interpolation defined by

$$\lambda_{i,h}(t) := \lambda_{i,h}^{k+1} \text{ if } t \in (hk, h(k+1)]. \tag{4.3}$$

4.1 Basic *a priori* estimates

In this section, we retrieve the usual estimates in the Wasserstein gradient flow theory. First, we show that $\overline{\mathcal{W}}_i$ is Lipschitz in the Wasserstein space.

Lemma 4.1 *There exists a constant $C > 0$ such that, for all $\boldsymbol{\mu} := (\mu_1, \ldots, \mu_I) \in \mathcal{P}(\Omega)^I$, and for all $\rho_1, \rho_2 \in \mathcal{P}^{ac}(\Omega)$,*

$$\overline{\mathcal{W}}_i(\rho_1|\boldsymbol{\mu}) - \overline{\mathcal{W}}_i(\rho_2|\boldsymbol{\mu}) \leqslant C W_2(\rho_1, \rho_2).$$

Proof Let γ be the W_2-optimal transport plan between ρ_1 and ρ_2 and T the W_2-optimal transport map associated with γ, i.e., $\gamma = (\mathrm{T} \times \mathrm{I})_{\#}\rho_2$. Let $\lambda \in \Pi(\mu_1, \ldots, \mu_{i-1}, \rho_2, \mu_{i+1}, \ldots, \mu_I)$ optimal for $\overline{\mathcal{W}}_i(\rho_2|\boldsymbol{\mu})$. Define λ_T by

$$\int_{\Omega^I} \varphi(x_1, \ldots, x_I)\, d\lambda_T(x_1, \ldots, x_I) := \int_{\Omega^I} \varphi(x_1, \ldots, \mathrm{T}(x_i), \ldots, x_I)\, d\lambda(x_1, \ldots, x_I),$$

for all $\varphi \in \mathcal{C}(\Omega^I)$. By definition, $\lambda_T \in \Pi(\mu_1, \ldots, \mu_{i-1}, \rho_1, \mu_{i+1}, \ldots, \mu_I)$. Then,

$$\begin{aligned}
\overline{\mathcal{W}}_i(\rho_1|\boldsymbol{\mu}) - \overline{\mathcal{W}}_i(\rho_2|\boldsymbol{\mu}) &\leqslant \int_{\Omega^I} \left[c_i(x_1, \ldots, \mathrm{T}(x_i), \ldots, x_I) - c_i(x_1, \ldots, x_I)\right] d\lambda(x_1, \ldots, x_I) \\
&\leqslant \|\nabla_{x_i} c_i\|_{L^\infty} \int_{\Omega^I} |\mathrm{T}(x_i) - x_i|\, d\lambda(x_1, \ldots, x_I) \\
&\leqslant C W_2(\rho_1, \rho_2),
\end{aligned}$$

where we used the fact that c_i is uniformly Lipschitz with respect to the i-th variable and Cauchy–Schwarz inequality. □

In the next proposition, we state usual estimates from JKO scheme.

Proposition 4.2 *For all $T > 0$, there exists $C_T > 0$ such that, for all h, k, with $hk < T$, $N = \lfloor \frac{T}{h} \rfloor$, for $i \in [\![1, I]\!]$, we have*

https://doi.org/10.1017/S0956792519000123 Published online by Cambridge University Press

$$\mathcal{F}_i\left(\rho_{i,h}^k\right) \leqslant C_T, \tag{4.4}$$

$$\sum_{k=0}^{N-1} W_2^2\left(\rho_{i,h}^k, \rho_{i,h}^{k+1}\right) \leqslant C_T h. \tag{4.5}$$

Proof We first prove (4.5). Since $\rho_{i,h}^{k+1}$ is optimal in the minimisation of (4.1) and $\rho_{i,h}^k$ is a competitor, we have

$$W_2^2\left(\rho_{i,h}^k, \rho_{i,h}^{k+1}\right) \leqslant 2h\left(\mathcal{F}_i\left(\rho_{i,h}^k\right) - \mathcal{F}_i\left(\rho_{i,h}^{k+1}\right) + \overline{W}_i\left(\rho_{i,h}^k|\boldsymbol{\rho}_h^k\right) - \overline{W}_i\left(\rho_{i,h}^{k+1}|\boldsymbol{\rho}_h^k\right)\right). \tag{4.6}$$

Then using Lemma 4.1 in (4.6) and Young's inequality, we obtain

$$\begin{aligned} W_2^2\left(\rho_{i,h}^k, \rho_{i,h}^{k+1}\right) &\leqslant 2h\left(\mathcal{F}_i\left(\rho_{i,h}^k\right) - \mathcal{F}_i\left(\rho_{i,h}^{k+1}\right) + CW_2\left(\rho_{i,h}^k, \rho_{i,h}^{k+1}\right)\right) \\ &\leqslant 2h\left(\mathcal{F}_i\left(\rho_{i,h}^k\right) - \mathcal{F}_i\left(\rho_{i,h}^{k+1}\right) + \frac{1}{4h} W_2^2\left(\rho_{i,h}^k, \rho_{i,h}^{k+1}\right) + 4C^2 h\right). \end{aligned}$$

We can thus absorb the W_2^2 term in the left-hand side,

$$\frac{1}{2} W_2^2\left(\rho_{i,h}^k, \rho_{i,h}^{k+1}\right) \leqslant 2h\left(\mathcal{F}_i\left(\rho_{i,h}^k\right) - \mathcal{F}_i\left(\rho_{i,h}^{k+1}\right) + Ch\right).$$

Summing over k, we find

$$\begin{aligned} \sum_{k=0}^{N-1} W_2^2\left(\rho_{i,h}^k, \rho_{i,h}^{k+1}\right) &\leqslant 4h\left(\sum_{k=0}^{N-1}\left(\mathcal{F}_i\left(\rho_{i,h}^k\right) - \mathcal{F}_i\left(\rho_{i,h}^{k+1}\right)\right) + C^2 T\right) \\ &\leqslant 4h\left(\mathcal{F}_i(\rho_{i,0}) - \mathcal{F}_i\left(\rho_{i,h}^N\right) + C^2 T\right). \end{aligned} \tag{4.7}$$

Since Ω is bounded, $\mathcal{F}_i$ is bounded from below and using the assumption (3.7), we conclude (4.5). The proof is completed noticing that the estimate (4.4) comes from (4.7) and (3.7). □

4.2 Refined *a priori* estimates

The goal of this section is to obtain stronger estimates on $P_i(\rho_{i,h})$ in order to deal with the nonlinear diffusion term.

Proposition 4.3 *For all $i \in [\![1, l]\!]$ and for all $k \geqslant 0$, we have $P_i\left(\rho_{i,h}^{k+1}\right) \in W^{1,1}(\Omega)$ and*

$$h\left(\nabla u_{i,h}^{k+1} \rho_{i,h}^{k+1} + \nabla P_i\left(\rho_{i,h}^{k+1}\right)\right) = -\nabla \varphi_{i,h}^{k+1} \rho_{i,h}^{k+1} \qquad a.e, \tag{4.8}$$

where $\varphi_{i,h}^{k+1}$ is a Kantorovich potential (so that its gradient is unique $\rho_{i,h}^{k+1}$-a.e.) from $\rho_{i,h}^{k+1}$ to $\rho_{i,h}^k$ for W_2 and $u_{i,h}^{k+1}$ is a Kantorovich potential for $\overline{\mathcal{W}}_i(\rho_{i,h}^{k+1}|\boldsymbol{\rho}_h^k)$.

Proof The proof is the same as in [1, 19] for example. We start by taking the first variation in the semi-implicit JKO scheme. Let $\xi \in \mathcal{C}_c^\infty(\Omega; \mathbb{R}^n)$ be given and Φ_τ the corresponding flow defined by

$$\partial_\tau \Phi_\tau = \xi \circ \Phi_\tau, \ \Phi_0 = Id.$$

https://doi.org/10.1017/S0956792519000123 Published online by Cambridge University Press

Define the perturbation ρ_τ of $\rho_{i,h}^{k+1}$ by $\rho_\tau := \Phi_{\tau\#}\rho_{i,h}^{k+1}$. Then we get

$$\frac{1}{\tau}\left(\mathcal{E}_{i,h}(\rho_\tau|\boldsymbol{\rho}_h^k) - \mathcal{E}_{i,h}(\rho_{i,h}^{k+1}|\boldsymbol{\rho}_h^k)\right) \geqslant 0. \tag{4.9}$$

By standard computations, we have

$$\limsup_{\tau\searrow 0}\frac{1}{\tau}\left(W_2^2(\rho_\tau, \rho_{i,h}^k) - W_2^2(\rho_{i,h}^{k+1}, \rho_{i,h}^k)\right) \leqslant \int_{\Omega\times\Omega}(x-y)\cdot\xi(x)\,d\gamma_{i,h}^{k+1}(x,y), \tag{4.10}$$

where $\gamma_{i,h}^{k+1}$ is a W_2-optimal transport plan in $\Pi(\rho_{i,h}^{k+1}, \rho_{i,h}^k)$ and $\gamma_{i,h}^{k+1} = (Id \times \mathrm{T}_{i,h}^{k+1})_{\#}\rho_{i,h}^{k+1}$ with $\mathrm{T}_{i,h}^{k+1} = \mathrm{Id} - \nabla\varphi_{i,h}^{k+1}$. Moreover, by (3.2), (4.4) and Lebesgue's dominated convergence theorem, we obtain

$$\limsup_{\tau\searrow 0}\frac{1}{\tau}\left(\mathcal{F}_i(\rho_\tau) - \mathcal{F}_i(\rho_{i,h}^{k+1})\right) \leqslant -\int_\Omega P_i(\rho_{i,h}^{k+1}(x))\mathrm{div}(\xi(x))\,dx. \tag{4.11}$$

Finally, by definition of $\lambda_{i,h}^{k+1}$, we have

$$\limsup_{\tau\searrow 0}\frac{1}{\tau}\left(\overline{\mathcal{W}}_i(\rho_\tau|\boldsymbol{\rho}_h^k) - \overline{\mathcal{W}}_i(\rho_{i,h}^{k+1}|\boldsymbol{\rho}_h^k)\right) \leqslant \int_{\Omega^l}\nabla_{x_i}c_i(x_1,\ldots,x_l)\cdot\xi(x_i)\,d\lambda_{i,h}^{k+1}(x_1,\ldots,x_l). \tag{4.12}$$

Combining (3.6), (4.9), (4.10), (4.11) and (4.12) and replacing ξ by $-\xi$, we find, for all $\xi \in \mathcal{C}_c^\infty(\Omega;\mathbb{R}^n)$,

$$\int_\Omega \nabla\varphi_{i,h}^{k+1}\cdot\xi\rho_{i,h}^{k+1} - h\int_\Omega P_i(\rho_{i,h}^{k+1})\mathrm{div}(\xi) + h\int_\Omega \nabla u_{i,h}^{k+1}\cdot\xi\rho_{i,h}^{k+1} = 0. \tag{4.13}$$

Now, we claim that $P_i(\rho_{i,h}^{k+1}) \in W^{1,1}(\Omega)$. Indeed, since P_i is controlled by F_i, (4.4) gives $P_i(\rho_{i,h}^{k+1}) \in L^1(\Omega)$ and, by (4.13), we obtain

$$\begin{aligned}\left|\int_\Omega P_i(\rho_{i,h}^{k+1})\mathrm{div}(\xi)\right| &\leqslant \left[\int_\Omega \frac{|\nabla\varphi_{i,h}^k(y)|}{h}\rho_{i,h}^{k+1} + \|\nabla_{x_i}c_i\|_{L^\infty}\right]\|\xi\|_{L^\infty(\Omega)}\\ &\leqslant \left[\frac{W_2(\rho_{i,h}^k, \rho_{i,h}^{k+1})}{h} + C\right]\|\xi\|_{L^\infty(\Omega)}.\end{aligned}$$

By duality, this implies $P_i(\rho_{i,h}^{k+1}) \in BV(\Omega)$ and $\nabla P_i(\rho_{i,h}^{k+1}) = \left(-\nabla u_{i,h}^{k+1}\rho_{i,h}^{k+1} - \frac{\nabla\varphi_{i,h}^{k+1}}{h}\rho_{i,h}^{k+1}\right)$ in $\mathcal{M}^n(\Omega)$. In fact, $P_i(\rho_{i,h}^{k+1})$ is in $W^{1,1}(\Omega)$ because $\nabla u_{i,h}^{k+1}\rho_{i,h}^{k+1} + \frac{\nabla\varphi_{i,h}^k}{h}\rho_{i,h}^{k+1} \in L^1(\Omega)$ and then (4.8) is proved. □

We deduce from (4.8) an $L^1((0,T), BV(\Omega))$ estimate for $P_i(\rho_{i,h})$.

Corollary 4.4 *For all $T > 0$, we have*

$$\|P_i(\rho_{i,h})\|_{L^1((0,T);W^{1,1}(\Omega))} \leqslant CT.$$

Proof Integrating (4.8), we obtain

$$h\int_\Omega |\nabla P_i(\rho_{i,h}^{k+1})| \leqslant W_2(\rho_{i,h}^k, \rho_{i,h}^{k+1}) + Ch.$$

https://doi.org/10.1017/S0956792519000123 Published online by Cambridge University Press

Then summing from $k = 0$ to $N - 1$ and thanks to (4.5), we have

$$\int_0^T \int_\Omega |\nabla P_i(\rho_{i,h})| \leqslant CT.$$

The conclusion follows with (3.2) and (4.4). □

4.3 Convergences and proof of Theorem 3.3

4.3.1 *Weak and strong convergences of $\rho_{i,h}$*

From the total square distance estimate (4.5), we deduce the classical W_2-convergence.

Proposition 4.5 *For all $T > 0$ and $i \in \{1, \dots, l\}$, there exists $\rho_i \in \mathcal{C}^{1/2}([0, T], \mathcal{P}^{ac}(\Omega))$ such that, up to extraction of a discrete subsequence,*

$$\sup_{t\in[0,T]} W_2\big(\rho_{i,h}(t), \rho_i(t)\big) \to 0.$$

Proof The proof is classical and is a consequence of (4.5) and a refined version of Arzelà–Ascoli's theorem [3, Proposition 3.3.1]. □

In order to handle the nonlinear diffusion term, the next proposition proves strong convergence in time and space.

Proposition 4.6 *Up to a subsequence, for all $i \in \{1, \dots, l\}$, $\rho_{i,h}$ converges strongly in $L^1((0, T) \times \Omega)$ to ρ_i and $\nabla P_i(\rho_{i,h})$ converges narrowly to $\nabla P_i(\rho_i)$.*

Proof The proof is now well known. We apply an extension of Aubin–Lions lemma proved by Rossi and Savaré [23, Theorem 2], see for example [11, 19]. Then we obtain that $\rho_{i,h}$ converges to ρ_i strongly in $L^1((0, T) \times \Omega)$.

It remains to prove that $P_i(\rho_{i,h})$ converges strongly to $P_i(\rho_i)$ in $L^1((0, T) \times \Omega)$. First, we know that $P_i(\rho_{i,h})$ is uniformly bounded in $L^\infty((0, T), L^1(\Omega))$, using (3.2), and thanks to Corollary 4.4, we have that $P_i(\rho_{i,h})$ is uniformly bounded in $L^1((0, T), W^{1,1}(\Omega))$. Then the Sobolev embedding gives that $P_i(\rho_{i,h})$ is uniformly bounded in $L^\infty((0, T), L^1(\Omega)) \cap L^1((0, T), L^{n/(n-1)}(\Omega))$. We deduce that $P_i(\rho_{i,h})$ is uniformly bounded in $L^{(n+1)/n}((0, T) \times \Omega)$, [11, Lemma 5.3]. This implies that $P_i(\rho_{i,h})$ is uniformly integrable and Vitali's convergence theorem gives that $P_i(\rho_{i,h})$ converges strongly to $P_i(\rho_i)$ in $L^1((0, T) \times \Omega)$. Then we conclude the narrow convergence of $\nabla P_i(\rho_{i,h})$ to $\nabla P_i(\rho_i)$ in $\mathcal{M}^n((0, T) \times \Omega)$, thanks to Corollary 4.4. □

4.3.2 *Convergence of $\overline{\mathcal{W}}_i$-optimal transport plans*

First, let us recall some notations. Let $\lambda_{i,h}^{k+1}$ be an optimal transport plan for $\mathcal{W}_i\left(\rho_{1,h}^k, \dots, \rho_{i-1,h}^k, \rho_{i,h}^{k+1}, \rho_{i+1,h}^k, \dots, \rho_{l,h}^k\right)$ and $\lambda_{i,h}$ the piecewise constant interpolation of $(\lambda_{i,h}^k)_k$, defined in (4.3).

In this section, the goal is to prove that $\lambda_{i,h}$ converges to λ_i, where $\lambda_i(t)$ is an optimal transport plan for $\mathcal{W}_i\left(\rho_1(t), \dots, \rho_l(t)\right)$, t-a.e. To simplify the exposition, we focus on the case $i = 1$ and the analysis is similar for $i > 1$. We introduce the shifted piecewise constant interpolations for all $i \in \{2, \dots, l\}$,

https://doi.org/10.1017/S0956792519000123 Published online by Cambridge University Press

$$\tilde{\rho}_{i,h}(t) := \rho_{i,h}^k \text{ if } t \in (hk, h(k+1)] \text{ and } \tilde{\rho}_{i,h}(0) := \rho_{i,0} \text{ if } t = 0,$$

and we denote, $\tilde{\boldsymbol{\rho}}_{-\mathbf{1},\boldsymbol{h}}$, the $(l-1)$-tuple $(\tilde{\rho}_{2,h}, \dots, \tilde{\rho}_{l,h})$ so that $\lambda_{1,h}(t) \in \Pi\left(\rho_{1,h}(t), \tilde{\boldsymbol{\rho}}_{-\mathbf{1},\boldsymbol{h}}(t)\right)$, for all $t > 0$.

Proposition 4.7 *For all $T > 0$, $\lambda_{1,h}$ narrowly converges to λ_1 in $\mathcal{P}([0,T] \times \Omega^l)$ and $\lambda_1(t) \in \Pi(\rho_1(t), \dots, \rho_l(t))$, t-a.e.*

Proof Proposition 4.5 implies that $\tilde{\boldsymbol{\rho}}_{-\mathbf{1},\boldsymbol{h}}$ narrowly converges to $\boldsymbol{\rho}_{-\mathbf{1}} := (\rho_2, \dots, \rho_l)$ in $L^\infty([0,T], \mathcal{P}^{ac}(\Omega)^{l-1})$. Define $\lambda_{1,h}^T := T^{-1}\lambda_{1,h}(t) \otimes dt \in \mathcal{P}([0,T] \times \Omega^l)$. Since $[0,T] \times \Omega^l$ is bounded, the sequence $\lambda_{1,h}^T$ is tight, then by Prokhorov's theorem, $\lambda_{1,h}^T$ narrowly converges to λ_1^T in $\mathcal{P}([0,T] \times \Omega^l)$. It remains to show that λ_1^T can be written as $T^{-1}\lambda_1(t) \otimes dt$, where $\lambda_1(t) \in \Pi(\rho_1(t), \dots, \rho_l(t))$ t-a.e. Denote $\pi^1, \pi^{1,i}$ the projections from $[0,T] \times \Omega^l$ to $[0,T] \times \Omega$ and $[0,T]$ with $\pi^{1,i}(t, x_1, \dots, x_l) = (t, x_i)$, and $\pi^1(t, x_1, \dots, x_l) = t$. Then we have $\pi_\#^{1,1}\lambda_{1,h} = \rho_{1,h}(t)dt$, $\pi_\#^{1,i}\lambda_{1,h} = \tilde{\rho}_{i,h}(t)dt$, for $i \neq 1$ and $\pi_\#^1\lambda_{1,h} = T^{-1}\mathcal{L}_{|[0,T]}$. When h goes to 0, since $\rho_{1,h}(t)dt$ and $\tilde{\rho}_{i,h}(t)dt$ narrowly converge to $\rho_1(t)dt$ and $\rho_i(t)dt$, we obtain $\pi_\#^{1,i}\lambda_1 = \rho_i(t)dt$ and $\pi_\#^1\lambda_i = T^{-1}\mathcal{L}_{|[0,T]}$, which concludes the proof. □

It remains to prove that the transport plan obtained in the last Proposition 4.7, $\lambda_1(t)$, is optimal for $\mathcal{W}_1(\rho_1(t), \dots, \rho_l(t))$. We start establishing an approximation result for an optimal transport plan between $\rho_1(t), \rho_2(t), \dots, \rho_l(t)$.

Lemma 4.8 *Let $\overline{\lambda}_1(t)$ be an optimal transport plan for $\mathcal{W}_1(\rho_1(t), \dots, \rho_l(t))$. There exists a sequence of transport plans $\overline{\lambda}_{1,h}(t) \in \Pi(\rho_{1,h}(t), \tilde{\boldsymbol{\rho}}_{-\mathbf{1},\boldsymbol{h}}(t))$ such that*

$$\sup_{t \in [0,T]} W_1(\overline{\lambda}_1(t), \overline{\lambda}_{1,h}(t)) \to 0.$$

Proof The proof is an adaptation of the one from [4, Lemma 6.2]. Let $\gamma_1(t) \in \Pi(\rho_1(t), \rho_{1,h}(t))$ be the optimal transport plan for W_2 and, for $i > 1$, let $\tilde{\gamma}_i(t) \in \Pi(\rho_i(t), \tilde{\rho}_{i,h}(t))$ be the optimal transport plan for W_2. Let us disintegrate $\gamma_1(t)$ and $\tilde{\gamma}_i(t)$ as $\gamma_1(t) = \rho_1(t) \otimes \gamma_1^x(t)$ and $\tilde{\gamma}_i(t) = \rho_i(t) \otimes \tilde{\gamma}_i^x(t)$. Now define $\overline{\lambda}_{1,h}(t)$ by

$$\overline{\lambda}_{1,h}(t) = \int_{\Omega^l} \gamma_1^{x_1}(t) \otimes \tilde{\gamma}_2^{x_2}(t) \otimes \cdots \otimes \tilde{\gamma}_l^{x_l}(t)\, d\overline{\lambda}_1(t, x_1, \dots, x_l).$$

By construction, $\overline{\lambda}_{1,h}(t) \in \Pi(\rho_{1,h}(t), \tilde{\boldsymbol{\rho}}_{-\mathbf{1},\boldsymbol{h}}(t))$. Then we introduce π a transport plan between $\overline{\lambda}_1(t)$ and $\overline{\lambda}_{1,h}(t)$ defined, for all $\varphi \in \mathcal{C}(\Omega^{2l})$, by

$$\int_{\Omega^{2l}} \varphi(\boldsymbol{x}, \boldsymbol{y})\, d\pi(\boldsymbol{x}, \boldsymbol{y}) = \int_{\Omega^l} \left(\int_{\Omega^l} \varphi(\boldsymbol{x}, \boldsymbol{y})\, \gamma_1^{x_1}(t, dy_1) \tilde{\gamma}_2^{x_2}(t, dy_2) \otimes \cdots \otimes \tilde{\gamma}_l^{x_l}(t, dy_l) \right) \overline{\lambda}_1(t, d\boldsymbol{x}),$$

where $\boldsymbol{x} = (x_1, \dots, x_l), \boldsymbol{y} = (y_1, \dots, y_l)$ are in Ω^l. Since $\pi \in \Pi(\overline{\lambda}_1(t), \overline{\lambda}_{1,h}(t))$, we have

$$\begin{aligned} W_1(\overline{\lambda}_1(t), \overline{\lambda}_{1,h}(t)) &\leqslant \int_{\Omega^{2l}} \left(|x_1 - y_1| + \sum_{i=2}^{l} |x_i - y_i| \right) d\pi(\boldsymbol{x}, \boldsymbol{y}) \\ &\leqslant \int_{\Omega^2} |x_1 - y_1| \gamma_1^{x_1}(t, dy_1) \rho_1(t, dx_1) + \sum_{i=2}^{l} \int_{\Omega^2} |x_i - y_i| \tilde{\gamma}_i^{x_i}(t, dy_i) \rho_i(t, dx_i) \end{aligned}$$

https://doi.org/10.1017/S0956792519000123 Published online by Cambridge University Press

$$\leqslant \int_{\Omega^2} |x-y|\gamma_1(t,dx,dy) + \sum_{i=2}^{l} \int_{\Omega^2} |x-y|\tilde{\gamma}_i(t,dx,dy)$$

$$\leqslant W_2(\rho_{1,h}(t),\rho_1(t)) + \sum_{i=2}^{l} W_2(\tilde{\rho}_{i,h}(t),\rho_i(t)).$$

Then Proposition 4.5 concludes the proof. □

From the previous lemma, we show that $\lambda_1(t)$ is optimal for $\mathcal{W}_1\ (\rho_1(t),\dots,\rho_l(t))$ t-a.e in $[0,T]$.

Proposition 4.9 *For almost every $t\in[0,T]$, $\lambda_1(t)$ is optimal for $\mathcal{W}_1\ (\rho_1(t),\dots,\rho_l(t))$.*

Proof Let $\overline{\lambda}_1(t)$ be an optimal transport plan for $\mathcal{W}_1\ (\rho_1(t),\dots,\rho_l(t))$. First, define $\overline{\lambda}_{1,h}(t)$ as in Lemma 4.8. Since, by definition, for all $t\in[0,T]$, $\lambda_{1,h}(t)$ is optimal for $\mathcal{W}_1(\rho_{1,h}(t),\tilde{\boldsymbol{\rho}}_{\boldsymbol{-1,h}}(t))$ and $\overline{\lambda}_{1,h}(t)\in\Pi(\rho_{1,h}(t),\tilde{\boldsymbol{\rho}}_{\boldsymbol{-1,h}}(t))$, we have

$$\int_{\Omega^l} c_1(\boldsymbol{x})\overline{\lambda}_{1,h}(t,d\boldsymbol{x}) \geqslant \int_{\Omega^l} c_1(\boldsymbol{x})\lambda_{1,h}(t,d\boldsymbol{x}),$$

where $\boldsymbol{x}=(x_1,\dots,x_l)$. So, for all nonnegative function $\varphi\in\mathcal{C}^\infty([0,T])$, we get

$$\int_0^T\int_{\Omega^l} c_1(\boldsymbol{x})\overline{\lambda}_{1,h}(t,d\boldsymbol{x})\varphi(t)\,dt \geqslant \int_0^T\int_{\Omega^l} c_1(\boldsymbol{x})\lambda_{1,h}(t,d\boldsymbol{x})\varphi(t)\,dt.$$

Since Ω is bounded and according to Lemma 4.8,

$$\int_0^T\int_{\Omega^l} c_1(\boldsymbol{x})\overline{\lambda}_{1,h}(t,d\boldsymbol{x})\varphi(t)\,dt \to \int_0^T\int_{\Omega^l} c_1(\boldsymbol{x})\overline{\lambda}_1(t,d\boldsymbol{x})\varphi(t)\,dt,$$

as $h\to 0$. In addition, since $\lambda_{1,h}$ narrowly converges to λ_1 in $\mathcal{P}([0,T]\times\Omega^l)$, we have

$$\int_0^T\int_{\Omega^l} c_1(\boldsymbol{x})\lambda_{1,h}(t,d\boldsymbol{x})\varphi(t)\,dt \to \int_0^T\int_{\Omega^l} c_1(\boldsymbol{x})\lambda_1(t,d\boldsymbol{x})\varphi(t)\,dt.$$

And then

$$\int_0^T\int_{\Omega^l} c_1(\boldsymbol{x})\overline{\lambda}_1(t,d\boldsymbol{x})\varphi(t)\,dt \geqslant \int_0^T\int_{\Omega^l} c_1(\boldsymbol{x})\lambda_1(t,d\boldsymbol{x})\varphi(t)\,dt.$$

The inequality holds for all nonnegative function $\varphi\in\mathcal{C}^\infty([0,T])$, we thus obtain, for almost every $t\in[0,T]$,

$$\mathcal{W}_1\ (\rho_1(t),\dots,\rho_l(t)) \geqslant \int_{\Omega^l} c_1(\boldsymbol{x})\lambda_1(t,d\boldsymbol{x}),$$

and the proof is concluded. □

4.3.3 *Proof of Theorem 3.3*

First, we show that $(\rho_{1,h},\dots,\rho_{l,h})$ is solution to a discrete approximation of system (3.3).

https://doi.org/10.1017/S0956792519000123 Published online by Cambridge University Press

Proposition 4.10 *Let $h > 0$, for all $T > 0$, let N such that $Nh = T$ and for all $\phi \in \mathcal{C}_c^\infty([0,T) \times \Omega)$, then*

$$\int_0^T \int_\Omega \rho_{i,h}(t,x)\partial_t\phi(t,x)\,dxdt = h\sum_{k=0}^{N-1}\int_\Omega \nabla P_i(\rho_{i,h}^{k+1}(x)) \cdot \nabla\phi(t_k,x)\,dx + h\sum_{k=0}^{N-1}\int_{\Omega^l} \nabla_{x_i} c_i(\boldsymbol{x}) \cdot \nabla\phi(t_k,x_i)\,d\lambda_{i,h}^{k+1}(\boldsymbol{x}) + \sum_{k=0}^{N-1}\int_{\Omega\times\Omega} \mathcal{R}[\phi(t_k,\cdot)](x,y)d\gamma_{i,h}^{k+1}(x,y) - \int_\Omega \rho_{i,0}(x)\phi(0,x)\,dx,$$

with, for all $\phi \in \mathcal{C}_c^\infty([0,T) \times \mathbb{R}^n)$,

$$|\mathcal{R}[\phi](x,y)| \leqslant \frac{1}{2}\|D^2\phi\|_{L^\infty([0,T)\times\Omega)}|x-y|^2,$$

and $\gamma_{i,h}^{k+1}$ is an optimal transport plan in $\Gamma(\rho_{i,h}^k, \rho_{i,h}^{k+1})$.

Proof This is a consequence of (4.8) (see [1, 19]). □

Now, we have to take the limit in the system of Proposition 4.10. Linear terms (with time derivative) and diffusion terms converge to the desired result, thanks to Proposition 4.6. The remainder terms goes to 0 as h goes to 0 because of (4.5). So it remains to check the convergence of multi-marginal interaction terms. By Proposition 4.7, $\lambda_{i,h}$ converges to λ_i in $\mathcal{P}([0,T] \times \Omega^l)$ and then,

$$h\sum_{k=0}^{N-1}\int_{\Omega^l} \nabla_{x_i} c_i(\boldsymbol{x}) \cdot \nabla\phi(t_k,x_i)\,d\lambda_{i,h}^{k+1}(\boldsymbol{x}) \to \int_0^T\int_{\Omega^l} \nabla_{x_i} c_i(\boldsymbol{x}) \cdot \nabla\phi(t,x_i)\,d\lambda_i(t,\boldsymbol{x})dt.$$

and, by Proposition 4.9, $\lambda_i(t)$ is an optimal transport plan for $\mathcal{W}_i(\rho_1(t), \ldots, \rho_l(t))$, t-a.e.

5 Uniqueness in dimension one

In this section, Ω is a compact convex subset in $\mathbb{R}$. We give a uniqueness result based on a displacement convexity argument and some examples of functionals satisfying this condition. Although Theorem 5.7 holds in dimension higher than one, we restrict ourselves to the dimension one because, as far as we know, there is no example of multi-marginal functional geodesically convex in higher dimension.

5.1 Displacement convexity in product Wasserstein space

For the purpose of this paper, it is enough to restrict ourselves to absolutely continuous probability measures. Given ρ_0 and ρ_1 in $\mathcal{P}^{ac}(\Omega)$, there exists a unique optimal transport map T between ρ_0 and ρ_1, i.e., $\mathrm{T}_\#\rho_0 = \rho_1$ and

$$W_2^2(\rho_0,\rho_1) = \int_\Omega |x - \mathrm{T}(x)|^2\rho_0(x)\,dx.$$

https://doi.org/10.1017/S0956792519000123 Published online by Cambridge University Press

In addition, $\mathrm{T} : \Omega \to \Omega$ is a nondecreasing map.

The Wasserstein geodesic between ρ_0 and ρ_1 is the curve $t \in [0,1] \mapsto \rho_t$ given by the McCann's interpolation

$$\rho_t := \mathrm{T}^t{}_{\#}\rho_0,$$

where $\mathrm{T}^t = (1-t)\mathrm{Id} + t\mathrm{T}$ is the optimal transport map between ρ_0 and ρ_t, and ρ_t is a constant speed geodesic:

$$W_2(\rho_t, \rho_s) = |t-s| W_2(\rho_0, \rho_1).$$

Now we recall the definition of geodesically convex functional in Wasserstein product space, see [3, 26–28].

Definition 5.1 Let $\lambda \in \mathbb{R}$. A functional $\mathcal{W} : \mathcal{P}(\Omega)^l \to (-\infty, +\infty]$ is said λ-geodesically convex in $\mathcal{P}(\Omega)^l$ if for every $i \in [\![1, l]\!]$ and for every couple $(\mu_i^0, \mu_i^1) \in \mathcal{P}(\Omega)^2$,

$$\begin{aligned}\mathcal{W}(\mu_1^t, \dots, \mu_l^t) \leqslant (1-t)\mathcal{W}(\mu_1^0, \dots, \mu_l^0) + t\mathcal{W}(\mu_1^1, \dots, \mu_l^1) \\ - \frac{\lambda}{2} t(1-t)\boldsymbol{W_2^2}((\mu_1^0, \dots, \mu_l^0), (\mu_1^1, \dots, \mu_l^1)),\end{aligned}$$

where μ_i^t is a constant speed geodesic between μ_i^0 and μ_i^1 and $\boldsymbol{W_2}$ is the product distance on $\mathcal{P}(\Omega)^l$.

Note that if $F : [0, +\infty) \to \mathbb{R}$ satisfies McCann's condition, i.e., the map

$$r \in (0, +\infty) \mapsto r^n F(r^{-n}) \text{ is convex nonincreasing,} \tag{5.1}$$

then it is well known that

$$\mathcal{F}(\rho) := \begin{cases} \int_\Omega F(\rho) & \text{if } F(\rho) \in L^1(\Omega), \\ +\infty & \text{otherwise,} \end{cases}$$

is geodesically convex ($\lambda = 0$), see [21].

In the following, we give a class of multi-marginal functionals geodesically convex. First, we provide a characterisation of the co-monotone transport plan as in [26, Lemma 2.8].

Lemma 5.2 *For $l \geqslant 2$, let γ be a transport plan having $\rho_1, \dots, \rho_l$ as marginals. If γ satisfies the property*

$$(x_1, \dots, x_l), (y_1, \dots, y_l) \in \operatorname{supp} \gamma \Rightarrow [x_1 < y_1 \Rightarrow \forall i,\ x_i \leqslant y_i], \tag{5.2}$$

then $\gamma = \gamma_{mon} := (G_1^{[-1]}, \dots, G_l^{[-1]})_{\#}\mathcal{L}^1_{|[0,1]}$, where $G_i^{[-1]}$ is the pseudo-inverse of the cumulative distribution function of ρ_i, $G_i(a) = \rho_i((-\infty, a])$.

Proof This lemma is an extension of [26, Lemma 2.8] (where the case $l = 2$ is studied), and the proof is similar. First, for all $a_1, \dots, a_l \in \mathbb{R}$, we know that $\gamma_{mon}((-\infty, a_1] \times \cdots \times (-\infty, a_l]) =$

https://doi.org/10.1017/S0956792519000123 Published online by Cambridge University Press

$\min_i G_i(a_i)$. Indeed, by the definition of $\gamma_{mon} \in \Pi(\rho_1, \dots, \rho_l)$,

$$\begin{aligned}\gamma_{mon}\big((-\infty, a_1] \times \cdots \times (-\infty, a_l]\big) &= \Big|\big\{x \in [0,1] : \forall i,\ G_i^{[-1]}(x) \leqslant a_i\big\}\Big| \\ &= \Big|\big\{x \in [0,1] : \forall i,\ x \leqslant G_i(a_i)\big\}\Big| \\ &= \min_i G_i(a_i).\end{aligned}$$

Since the knowledge of $\gamma((-\infty, a_1] \times \cdots \times (-\infty, a_l])$, for all $a_1, \dots, a_l \in \mathbb{R}$ is enough to characterise γ, we just need to show that

$$\gamma\big((-\infty, a_1] \times \cdots \times (-\infty, a_l]\big) = \min_i G_i(a_i),$$

to conclude the proof. Define for all i, the set $A_i = \Pi_{j=1}^{i-1}[a_j, +\infty) \times (-\infty, a_i] \times \Pi_{j=i+1}^{l}[a_j, +\infty)$. Since γ satisfies (5.2), for all $i \neq j$, we cannot have both $\gamma(A_i) > 0$ and $\gamma(A_j) > 0$. Then,

$$\begin{aligned}\gamma\big((-\infty, a_1] \times \cdots \times (-\infty, a_l]\big) &= \min_i \gamma\big((-\infty, a_1] \times \cdots \times (-\infty, a_l] \cup A_i\big) \\ &= \min_i \gamma\big(\mathbb{R} \times \cdots \times \mathbb{R} \times (-\infty, a_i] \times \mathbb{R} \cdots \times \mathbb{R}\big) \\ &= \min_i \rho_i((-\infty, a_i]) \\ &= \min_i G_i(a_i).\end{aligned}$$

□

This lemma allows us to study the geodesic convexity of multi-marginal functionals for a large class of costs.

Proposition 5.3 *Let $c : \Omega^l \to \mathbb{R}$ be a $\mathcal{C}^2$ convex function satisfying $\partial_{i,j} c \leqslant 0$ for all $i \neq j$. The functional $\mathcal{W}_c : \mathcal{P}^{ac}(\Omega)^l \to \mathbb{R}$ defined by*

$$\mathcal{W}_c(\rho_1, \dots, \rho_l) := \inf\left\{\int_{\Omega^l} c(x_1, \dots, x_l)\, d\gamma(x_1, \dots, x_l) : \gamma \in \Pi(\rho_1, \dots, \rho_l)\right\}$$

is geodesically convex in $\mathcal{P}^{ac}(\Omega)^l$.

Proof Given $(\rho_1^0, \dots, \rho_l^0)$ and $(\rho_1^1, \dots, \rho_l^1)$ in $\mathcal{P}^{ac}(\Omega)^l$, define the constant speed geodesic between ρ_i^0 and ρ_i^1, $\rho_i^t = \mathrm{T}_{i\#}^t \rho_i^0$. Let γ_0 be an optimal transport plan for the multi-marginal problem $\mathcal{W}_c(\rho_1^0, \dots, \rho_l^0)$. By [8, Theorem 4.1], there exist $l-1$ nondecreasing maps $\mathrm{S}_2, \dots, \mathrm{S}_l$ such that $\gamma_0 = (\mathrm{Id}, \mathrm{S}_2, \dots, \mathrm{S}_l)_\# \rho_1^0$.

Define the interpolation plan γ_t by

$$\gamma_t = \big(\mathrm{T}_1^t, \dots, \mathrm{T}_l^t\big)_\# \gamma_0 = \big(\mathrm{T}_1^t, \mathrm{T}_2^t \circ \mathrm{S}_2, \dots, \mathrm{T}_l^t \circ \mathrm{S}_l\big)_\# \rho_1^0.$$

Observe that $\gamma_1 = (\mathrm{T}_1, \mathrm{T}_2 \circ \mathrm{S}_2, \dots, \mathrm{T}_l \circ \mathrm{S}_l)_\# \rho_1^0$ and since T_1 and for all $i \geqslant 2$, $\mathrm{T}_i \circ \mathrm{S}_i$ are nondecreasing maps, γ_1 satisfies (5.2).

We want to show that γ_1 is an optimal transport plan for $\mathcal{W}_c(\rho_1^1, \dots, \rho_l^1)$. Theorem 4.1 from [8] says that the optimal transport plan γ_{opt} of $\mathcal{W}_c(\rho_1^1, \dots, \rho_l^1)$ is of the form $\gamma_{opt} = (\mathrm{Id}, \tilde{\mathrm{S}}_2, \dots, \tilde{\mathrm{S}}_l)_\# \rho_1^1$, where $\tilde{\mathrm{S}}_i$ is nondecreasing. Then γ_{opt} also satisfies (5.2) and then by Lemma 5.2, we conclude that γ_1 is an optimal transport plan for $\mathcal{W}_c(\rho_1^1, \dots, \rho_l^1)$.

https://doi.org/10.1017/S0956792519000123 Published online by Cambridge University Press

By convexity of c, we have

$$\begin{aligned}\mathcal{W}_c(\rho_1^t,\dots,\rho_l^t) &\leqslant \int_{\Omega^l} c(x_1,\dots,x_l)\,d\gamma_t\\ &\leqslant \int_{\Omega^l} c\big(\mathrm{T}_1^t(x_1),\dots,\mathrm{T}_l^t(x_l)\big)\,d\gamma_0\\ &\leqslant (1-t)\int_{\Omega^l} c(x_1,\dots,x_l)\,d\gamma_0 + t\int_{\Omega^l} c\big(\mathrm{T}_1(x_1),\dots,\mathrm{T}_l(x_l)\big)\,d\gamma_0\\ &\leqslant (1-t)\mathcal{W}_c(\rho_1^0,\dots,\rho_l^0) + t\mathcal{W}_c(\rho_1^1,\dots,\rho_l^1),\end{aligned}$$

which concludes the proof. □

Remark 5.4 This result cannot be generalised in higher dimensions. Indeed, in dimension $n>1$, it is well known that $W_2(\cdot,\sigma)$ is not λ-convex along geodesic on $\mathcal{P}(\Omega)$ (see Example 9.1.5 from [3]).

5.2 Wasserstein contraction

First, let us define the Fréchet subdifferential for $\mathcal{W}:\mathcal{P}^{ac}(\Omega)^l\to(-\infty,+\infty]$ by extending the definition given in [3].

Definition 5.5 Let $\mathcal{W}:\mathcal{P}^{ac}(\Omega)^l\to(-\infty,+\infty]$ be a functional and let $\boldsymbol{\xi}=(\xi_1,\dots,\xi_l)\in L^2((\mu_1,\dots,\mu_l),\Omega)$, i.e.,

$$\sum_{i=1}^{l}\int_\Omega |\xi_i|^2\mu_i<+\infty.$$

We say that $\boldsymbol{\xi}$ is in the Fréchet subdifferential $\partial\mathcal{W}(\mu_1,\dots,\mu_l)$ if

$$\liminf_{\boldsymbol{\nu}\to\boldsymbol{\mu}}\frac{\mathcal{W}(\boldsymbol{\nu})-\mathcal{W}(\boldsymbol{\mu})-\sum_{i=1}^l\int_\Omega\langle\xi_i(x),\mathrm{T}_{\mu_i}^{\nu_i}(x)-x\rangle\mu_i(x)\,dx}{\boldsymbol{W_2}(\boldsymbol{\nu},\boldsymbol{\mu})}\geqslant 0, \tag{5.3}$$

where $\boldsymbol{\mu}:=(\mu_1,\dots,\mu_l)$, $\boldsymbol{\nu}:=(\nu_1,\dots,\nu_l)$ and $\mathrm{T}_{\mu_i}^{\nu_i}$ is the optimal transport map between μ_i and ν_i.

The next proposition characterises the subdifferential of λ-geodesically convex functionals.

Proposition 5.6 *Let $\mathcal{W}:\mathcal{P}^{ac}(\Omega)^l\to(-\infty,+\infty]$ be a λ-geodesically convex functional. Then a vector $\boldsymbol{\xi}\in L^2(\boldsymbol{\mu},\Omega)$ belongs to the Fréchet subdifferential of $\mathcal{W}$ at $\boldsymbol{\mu}$ if and only if*

$$\mathcal{W}(\boldsymbol{\nu})-\mathcal{W}(\boldsymbol{\mu})\geqslant\sum_{i=1}^{l}\int_\Omega\langle\xi_i(x),\mathrm{T}_{\mu_i}^{\nu_i}(x)-x\rangle\mu_i(x)\,dx+\frac{\lambda}{2}\boldsymbol{W_2^2}(\boldsymbol{\nu},\boldsymbol{\mu}), \tag{5.4}$$

for all $\boldsymbol{\nu}\in\mathcal{P}^{ac}(\Omega)^l$. Moreover, if $\boldsymbol{\xi}\in\partial\mathcal{W}(\boldsymbol{\mu})$ and $\boldsymbol{\kappa}\in\partial\mathcal{W}(\boldsymbol{\nu})$, then

$$\sum_{i=1}^{l}\int_\Omega\langle\xi_i(x)-\kappa_i(\mathrm{T}_{\mu_i}^{\nu_i}(x)),\mathrm{T}_{\mu_i}^{\nu_i}(x)-x\rangle\mu_i(x)\,dx\leqslant-\lambda\boldsymbol{W_2^2}(\boldsymbol{\mu},\boldsymbol{\nu}). \tag{5.5}$$

https://doi.org/10.1017/S0956792519000123 Published online by Cambridge University Press

Proof The proof is the same as in the characterisation by Variational inequalities and monotonicity done in [3, p. 231]. □

We can now prove the following uniqueness result.

Theorem 5.7 *Assume F_i satisfies (5.1) and $\mathcal{W}_i$ is a λ_i-geodesically convex functional. Let $\boldsymbol{\rho}^1 := (\rho_1^1, \dots, \rho_l^1)$ and $\boldsymbol{\rho}^2 := (\rho_1^2, \dots, \rho_l^2)$ be two weak solutions of (3.3)–(3.5) with initial conditions $\rho_i^1(0,\cdot) = \rho_{i,0}^1$ and $\rho_i^2(0,\cdot) = \rho_{i,0}^2$. If for all $T < +\infty$,*

$$\int_0^T \sum_{i=1}^{l} \|v_{i,t}^1\|_{L^2(\rho_{i,t}^1)}\, dt + \int_0^T \sum_{i=1}^{l} \|v_{i,t}^2\|_{L^2(\rho_{i,t}^2)}\, dt < +\infty, \tag{5.6}$$

with, for $j \in \{1,2\}$,

$$v_{i,t}^j := -\nabla F_i'(\rho_{i,t}^j) - \nabla u_i^j,$$

then for every $t \in [0,T]$,

$$\boldsymbol{W_2^2}(\boldsymbol{\rho}_t^1, \boldsymbol{\rho}_t^2) \leqslant e^{-\left(2\sum_{i=1}^{l}\lambda_i\right)t} \boldsymbol{W_2^2}(\boldsymbol{\rho}_0^1, \boldsymbol{\rho}_0^2).$$

Proof Using Theorem 5.24 and Corollary 5.25 from [26] and assumption (5.6), we obtain

$$\frac{d}{dt}\left(\frac{1}{2} W_2^2(\rho_{i,t}^1, \rho_{i,t}^2)\right) = \int_\Omega \langle x - \mathrm{T}_i^t(x),\, v_{i,t}^1(x) - v_{i,t}^2(\mathrm{T}_i^t(x))\rangle \rho_i^1(x)\, dx,$$

where T_i^t is the optimal transport map between $\rho_{i,t}^1$ and $\rho_{i,t}^2$. Since F_i satisfies McCann's condition (5.1), we have

$$\int_\Omega \langle x - \mathrm{T}_i^t(x),\, \nabla F_i'(\rho_{i,t}^2(\mathrm{T}_i^t(x))) - \nabla F_i'(\rho_{i,t}^1(x))\rangle \rho_{i,t}^1(x)\, dx \leqslant 0.$$

In addition, $\mathcal{W}_i$ is λ_i-geodesically convex, and then (5.5) gives

$$\sum_{i=1}^{l} \int_\Omega \langle \nabla u_{i,t}^1(x) - \nabla u_{i,t}^2(\mathrm{T}_i^t(x)),\, \mathrm{T}_i^t(x) - x\rangle \rho_{i,t}^1(x)\, dx \leqslant -\sum_{i=1}^{l} \lambda_i \boldsymbol{W_2^2}(\boldsymbol{\rho}_t^1, \boldsymbol{\rho}_t^2).$$

Summing over i and combining these inequalities, we obtain

$$\frac{d}{dt}\left(\boldsymbol{W_2^2}(\boldsymbol{\rho}_t^1, \boldsymbol{\rho}_t^2)\right) \leqslant -\left(2\sum_{i=1}^{l}\lambda_i\right) \boldsymbol{W_2^2}(\boldsymbol{\rho}_t^1, \boldsymbol{\rho}_t^2).$$

Gronwall's lemma concludes the proof. □

Remark 5.8 Assumption (5.6) in Theorem 5.7 is made to ensure the absolute continuity of $W_2(\rho_{i,t}^1, \rho_{i,t}^2)$ and can be checked using (4.8) (see for example [19, Proposition 7.3]).

6 Conclusion

In this work, we have studied the existence and uniqueness of nonlinear systems coupled through multi-marginal transport problems. These systems can arise in dynamical urban planning models.

https://doi.org/10.1017/S0956792519000123 Published online by Cambridge University Press

We have introduced a constructive method using a semi-implicit Euler scheme in Wasserstein space to solve it. This method has permitted to deal with the high nonlinearity of the interactions.

At the end, we have given a uniqueness result in dimension one for multi-marginal problems with convex costs. We have used the geodesic convexity of these problems to prove it.

A perspective of this work is to replace the multi-marginal transport by its entropic regularisation. That means that noise is introduced in the transport problems between populations. Then using the same argument, we want to prove the existence and uniqueness and in addition obtain a numerical scheme to solve it.

Conflicts of interest

None.

References

[1] AGUEH, M. (2005) Existence of solutions to degenerate parabolic equations via the Monge-Kantorovich theory. *Adv. Differ. Equations* **10**(3), 309–360.

[2] AGUEH, M. & CARLIER, G. (2011) Barycenters in the Wasserstein space. *SIAM J. Math. Anal.* **43**(2), 904–924.

[3] AMBROSIO, L., GIGLI, N. & SAVARÉ, G. (2005) *Gradient Flows in Metric Spaces and in the Space of Probability Measures*. Lectures in Mathematics, ETH Zürich, Birkhäuser Verlag, Basel.

[4] BAILLON, J.-B. & CARLIER, G. (2012) From discrete to continuous Wardrop equilibria. *Netw. Heterog. Media* **7**(2), 219–241.

[5] BRENIER, Y. (1991) Polar factorization and monotone rearrangement of vector-valued functions. *Comm. Pure Appl. Math.* **44**(4), 375–417.

[6] BUTTAZZO, G. & SANTAMBROGIO, F. (2005) A model for the optimal planning of an urban area. *SIAM J. Math. Anal.* **37**(2), 514–530.

[7] BUTTAZZO, G. & SANTAMBROGIO, F. (2009) A mass transportation model for the optimal planning of an urban region. *SIAM Rev.* **51**(3), 593–610.

[8] CARLIER, G. (2003) On a class of multidimensional optimal transportation problems. *J. Convex Anal.* **10**(2), 517–529.

[9] CARLIER, G. & EKELAND, I. (2004) The structure of cities. *J. Global Optim.* **29**(4), 371–376.

[10] CARLIER, G. & EKELAND, I. (2007) Equilibrium structure of a bidimensional asymmetric city. *Nonlinear Anal. Real World Appl.* **8**(3), 725–748.

[11] CARLIER, G. & LABORDE, M. (2017) A splitting method for nonlinear diffusions with nonlocal, nonpotential drifts. *Nonlinear Anal. Theory Methods Appl.* **150**, 1 – 18.

[12] CARLIER, G. & SANTAMBROGIO, F. (2005) A variational model for urban planning with traffic congestion. *ESAIM Control Optim. Calc. Var.* **11**(4), 595–613.

[13] DE GIORGI, E. (1993) New problems on minimizing movements. In: *Boundary Value Problems for Partial Differential Equations and Applications*, RMA Res. Notes Appl. Math., Vol. 29, Masson, Paris, pp. 81–98.

[14] DI FRANCESCO, M. & FAGIOLI, S. (2013) Measure solutions for non-local interaction PDEs with two species. *Nonlinearity* **26**(10), 2777–2808.

[15] DI MARINO, S., GEROLIN, A. & NENNA, L. (2017) Optimal transportation theory with repulsive costs. In: *Topological Optimization and Optimal Transport*, Radon Ser. Comput. Appl. Math., Vol. 17, De Gruyter, Berlin, pp. 204–256.

[16] FUJITA, M. (1989) *Urban Economic Theory: Land Use and City Size*, Cambridge University Press, Cambridge.

[17] JORDAN, R., KINDERLEHRER, D. & OTTO, F. (1998) The variational formulation of the Fokker-Planck equation. *SIAM J. Math. Anal.* **29**(1), 1–17.

https://doi.org/10.1017/S0956792519000123 Published online by Cambridge University Press

[18] KINDERLEHRER, D., MONSAINGEON, L. & XU, X. (2017) A Wasserstein gradient flow approach to Poisson-Nernst-Planck equations. *ESAIM Control Optim. Calc. Var.* **23**(1), 137–164.
[19] LABORDE, M. (2017) On some nonlinear evolution systems which are perturbations of Wasserstein gradient flows. In: *Topological Optimization and Optimal Transport*, Radon Ser. Comput. Appl. Math., Vol. 17, De Gruyter, Berlin, pp. 304–332.
[20] LUCAS, R. E. & ROSSI-HANSBERG, E. (2002) On the internal structure of cities. *Econometrica* **70**(4), 1445–1476.
[21] MCCANN, R. J. (1997) A convexity principle for interacting gases. *Adv. Math.* **128**(1), 153–179.
[22] PASS, B. (2015) Multi-marginal optimal transport: theory and applications. *ESAIM Math. Model. Numer. Anal.* **49**(6), 1771–1790.
[23] ROSSI, R. & SAVARÉ, G. (2003) Tightness, integral equicontinuity and compactness for evolution problems in Banach spaces. *Ann. Sc. Norm. Super. Pisa Cl. Sci.* (5) 2(2), 395–431.
[24] SANTAMBROGIO, F. (2007) Transport and concentration problems with interaction effects. *J. Global Optim.* **38**(1), 129–141.
[25] SANTAMBROGIO, F. *Variational problems in transport theory with mass concentration,* Tesi. Scuola Normale Superiore di Pisa (Nuova Series)], Vol. 4, Edizioni della Normale, Pisa, 2007.
[26] SANTAMBROGIO, F. (2015) *Optimal Transport for Applied Mathematicians*. Progress in Nonlinear Differential Equations and Their Applications, Vol. 87, Birkhäuser Verlag, Basel.
[27] VILLANI, C. (2003) *Topics in Optimal Transportation*, Graduate Studies in Mathematics, Vol. 58. American Mathematical Society, Providence, RI.
[28] VILLANI, C. (2009) *Optimal Transport*, Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], Vol. 338, Springer-Verlag, Berlin. Old and new.

https://doi.org/10.1017/S0956792519000123 Published online by Cambridge University Press