

# A note on Cournot-Nash equilibria and Optimal Transport between unequal dimensions\*

Luca Nenna<sup>†</sup>      Brendan Pass<sup>‡</sup>

September 29, 2022

## Abstract

This note is devoted to study a class of games with a continuum of players for which Cournot-Nash equilibria can be obtained by minimisation of some cost related to Optimal Transport. In particular we focus on the case of an Optimal Transport term between unequal dimension. We also present several numerical simulations.

## Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
<b>2</b>	<b>Optimal Transport between unequal dimensions</b>	<b>2</b>
2.1	Multi-to one-dimensional optimal transport . . . . .	3
<b>3</b>	<b>A variational approach to Cournot-Nash equilibria</b>	<b>4</b>
3.1	Nestedness of Cournot-Nash equilibrium . . . . .	5
<b>4</b>	<b>Numerics</b>	<b>7</b>
4.1	Congestion case . . . . .	7
4.2	Best reply scheme . . . . .	8
4.3	The Entropic Regularization . . . . .	9

## 1 Introduction

Since Aumann’s seminal works [1, 2], equilibria in games with a continuum of players have received a lot of attention from the Economics and Game Theory communities. Following Mas-Colell’s approach [13], we consider a type space  $X \subset \mathbb{R}^m$  endowed with a probability measure  $\mu$ . Each player of type  $x$  has to choose a strategy  $y \in Y$  (where  $Y \subset \mathbb{R}^n$ ) in order to minimise a cost  $\Phi : X \times Y \times \mathcal{P}(Y) \rightarrow \mathbb{R}$  which depends both on his type  $x$  and strategy  $y$  as well as the distribution of strategies  $\nu \in \mathcal{P}(Y)$  resulting from the other players’ behaviour. Since the cost depends on other players’ strategies only through the distribution  $\nu$ , it is not important who plays a specific strategy, but how many players chose it (i.e. the game is anonymous). Thus, a Cournot-Nash equilibrium is defined as a probability measure  $\gamma \in \mathcal{P}(X \times Y)$ , whose first marginal is  $\mu$ , such that

$$\gamma(\{(x, y) \in X \times Y : \Phi(x, y, \nu) = \min_{z \in Y} \Phi(x, z, \nu)\}) = 1$$

where the strategy distribution  $\nu$  is the second marginal of  $\gamma$ . Notice that if we consider an homogenous population of players (which mean that we do not need a

---

\*B.P. is pleased to acknowledge the support of National Sciences and Engineering Research Council of Canada Discovery Grant number 04658-2018. L.N. was supported by a public grant as part of the Investissement d’avenir project, reference ANR-11-LABX-0056-LMH, LabEx LMH and from the CNRS PEPS JCJC (2022).

<sup>†</sup>Université Paris-Saclay, CNRS, Laboratoire de mathématiques d’Orsay, 91405, Orsay, France. [luca.nenna@universite-paris-saclay.fr](mailto:luca.nenna@universite-paris-saclay.fr)

<sup>‡</sup>Department of Mathematical and Statistical Sciences, 632 CAB, University of Alberta, Edmonton, Alberta, Canada, T6G 2G1 [pass@ualberta.ca](mailto:pass@ualberta.ca)

distribution  $\mu$  anymore) then we retrieve exactly the definition of Nash equilibrium. Let us now focus on the additively separate case where the total cost  $\Phi$  can be written as  $\Phi(x, y, \nu) = c(x, y) + F(\nu, y)$ . It has been recently showed in [6] (see also [4]) that Cournot-Nash equilibria can be obtained by the minimization of a certain functional on the set of measures on the space of strategies, which actually means that the games considered in [6] belongs to a class of models which have the structure of a potential game (i.e. variational problems). This functional typically involves two terms: an optimal transport cost and a more standard integral functional which may capture both congestion and attractive effects. This variational approach is somehow more constructive and informative (but less general since it requires a separable cost) than the one relying on fixed-point arguments as in [13] but the optimal transport cost term is delicate to handle. It is indeed costly in general to solve numerically an optimal transport problem and compute an element of the subdifferential of the optimal cost as a function of its marginals. However it has been recently introduced (see for instance [3, 11, 12]) a very efficient numerical method relying on an entropic regularization of optimal transport which turned to be useful also to approximate Cournot-Nash equilibria (see [7]), in the case of potential games.

In [6] the authors treat the case where the measures  $\mu$  and  $\nu$  are probabilities on  $X \subset \mathbb{R}^m$  and  $Y \subset \mathbb{R}^n$ , respectively, where  $m = n$ . In this article we want to address the case in which  $m > n$  so that the optimal transport term becomes an *unequal dimensional Optimal Transport term* [8, 9]. In a recent paper [15] we have showed that for a variety of different forms of  $\mathcal{F}[\nu](y)$  a nestedness condition, which is known to yield much improved tractability of the optimal transport problem, holds for any minimizer. Depending on the exact form of the functional, we exploit this to find local differential equations characterizing solutions, prove convergence of an iterative scheme to characterise the solution, and prove regularity results. The aim of this paper is to illustrate some numerical methods to compute the solution in the case of Cournot-Nash equilibria by using the characterization results established in [15].

**Motivation** We want to briefly remark through an example that unequal dimensional Optimal Transport is actually quite natural in the Economics/Game theory framework as the one we are dealing with.

Let think that we have a population of physicians who are characterized by a type  $x$  belonging to  $X \subset \mathbb{R}^m$  where the dimension  $m$  is the number of characteristics: age, gender, university where they get their diploma, their hometown, etc. The probability measure  $\mu$  on  $X$  gives then the distribution of types in the physician population. The main idea of Optimal Transport is to match the physician with a city, for example, where they would like to open their private practice. The cities are characterized by a type  $y$  which belongs to  $Y \subset \mathbb{R}^n$  where  $n$  is now the number of characteristics and  $Y$  is endowed with a probability measure  $\nu$  giving the distribution of types of cities. Notice that the matching between physicians and cities must minimize a given transportation cost  $c(x, y)$  which can indeed model the cost for physician  $x$  to commute to the city  $y$  where he/she has his/her practice. One can also think, in more economical terms, that the cost is of the form  $c(x, y) = -s(x, y)$  where  $s$  is a surplus and in this case the optimal matching will maximize the given surplus. It is actually quite important to highlight now that the number of characteristics  $n$  of the cities can be much smaller than the one for the physician; indeed it is natural to take  $n = 1$  saying that just a characteristic (i.e. the population) of the city is taken into account by the physicians.

**Remark 1** (Notation). *With a slight abuse of notation we will identify the measure with its density throughout all the paper.*

## 2 Optimal Transport between unequal dimensions

Given two probability measures  $\mu$  on  $X \subset \mathbb{R}^m$  and  $\nu$  on  $Y \subset \mathbb{R}^n$ , the Monge-Kantorovich problem consists in transporting  $\mu$  onto  $\nu$  so as to optimise a given cost function  $c(x, y)$ . Indeed we look for a probability measure  $\gamma$  on the product space  $X \times Y$  whose marginals are  $\mu$  and  $\nu$  such that it solves the following maximisation problem

$$\mathcal{T}_c(\mu, \nu) := \inf_{\gamma \in \Pi(\mu, \nu)} \int_{X \times Y} c(x, y) d\gamma(x, y), \quad (1)$$

where  $\Pi(\mu, \nu) := \{\eta \in \mathcal{P}(X \times Y) : \pi_x(\eta) = \mu, \pi_y(\eta) = \nu\}$  and  $\pi_x : X \times Y \rightarrow X$ .

We refer the reader to [18, 19] for results about the characterisation of optimal maps for the case in which  $m = n$ . Before discussing the case  $m > n$ , it is important to highlight that the Monge-Kantorovich problem admits a dual formulation which is useful in order to understand the solution to (1)

$$\mathcal{T}_c^{dual}(\mu, \nu) := \sup_{(u, v) \in \mathcal{U}} \int_X u(x) d\mu(x) + \int_Y v(y) d\nu(y), \quad (2)$$

where  $\mathcal{U} := \{(u, v) \in L^1(\mu) \oplus L^1(\nu) : u(x) + v(y) \geq c(x, y) \text{ on } X \times Y\}$ . The optimal  $u, v$  are called Kantorovich potentials. The interesting fact is that  $\mathcal{T}_c = \mathcal{T}_c^{dual}$ .

Under mild conditions (for instance, if  $c$  is differentiable, as we are assuming here,  $X$  connected and  $\bar{\mu}(x) > 0$  for all  $x \in X$  [18]), there is a unique solution  $(u, v)$  to the dual problem, up to the addition  $(u, v) \mapsto (u + C, v - C)$  of a constant, known as the Kantorovich potentials and these potentials are  $c$ -concave; that is, they satisfy

$$u(x) = v^c(x) := \min_{y \in Y} [c(x, y) - v(y)], \quad v(y) = u^c(y) := \min_{x \in X} [c(x, y) - u(x)],$$

In particular if both  $\nu$  and  $\mu$  are absolutely continuous, then the potential  $v$  satisfies the Monge-Ampere type equation almost everywhere [14]<sup>1</sup>:

$$v(y) = \int_{\partial^c v(y)} \frac{\det(D_{yy}^2 c(x, y) - D^2 v(y))}{\sqrt{|\det(D_{yx}^2 c D_{xy}^2 c)(x, y)|}} \mu(x) d\mathcal{H}^{m-n}(x), \quad (3)$$

where  $\partial^c v(y) := \{x : u(x) + v(y) = c(x, y)\}$ . In general, this is a *non-local* differential equation for  $v(y)$ , since the domain of integration  $\partial^c v(y)$  is defined using the values of  $v$  and  $u = v^c$  throughout  $Y$ ; however, when the model satisfies the generalized nestedness condition (namely  $\partial^c v(y) = X_=(y, Dv(y))$ , where  $X_=(y, p) = \{x \in X : D_y c(x, y) = p\}$ , see [15][Definition 1] for more details), it reduces to the local equation [14]:

$$v(y) = \int_{X_=(y, Dv(y))} \frac{\det(D_{yy}^2 c(x, y) - D^2 v(y))}{\sqrt{|\det(D_{yx}^2 c D_{xy}^2 c)(x, y)|}} \mu(x) d\mathcal{H}^{m-n}(x). \quad (4)$$

## 2.1 Multi-to one-dimensional optimal transport

We consider now the optimal transport problem in the case in which  $m > n = 1$  (for more details we refer the reader to [9]). Let us define the super-level set  $X_{\geq}(y, k)$  as follows

$$X_{\geq}(y, k) := \{x \in X : \partial_y c(x, y) \geq k\},$$

and its strict variant  $X_{>}(y, k) := X_{\geq}(y, k) \setminus X_=(y, k)$ . In order to build an optimal transport map  $T$  we take the unique level set splitting the mass proportionately with  $y$ ; that is  $k(y)$  such that

$$\mu(X_{\geq}(y, k(y))) = \nu((-\infty, y]), \quad (5)$$

then we set  $y = T(x)$  for all  $x$  which belongs to  $X_=(y, k(y))$ . Notice that if  $x \in X_=(y_0, k(y_0)) \cap X_=(y_1, k(y_1))$  then the map  $T$  is not well-defined. In order to avoid such a degenerate case we ask that the model  $(c, \mu, \nu)$  satisfies the following property

**Definition 2** (Nestedness  $n = 1$ ). *The model  $(c, \mu, \nu)$  is nested if*

$$\forall y_0, y_1 \text{ with } y_1 > y_0, \nu([y_0, y_1]) > 0 \implies X_{\geq}(y_0, k(y_0)) \subset X_{>}(y_1, k(y_1)).$$

<sup>1</sup>Note that, here and below, our notation differs somewhat from [9] and [14], since we have adopted the convention of minimizing, rather than maximizing, in (1).

Thus, if the model  $(c, \mu, \nu)$  is nested then [9][Theorem 4] assures that  $\gamma_T = (\text{id}, T)_\# \mu$ , where the map  $T$  is built as above, is the unique maximiser of (1) on  $\Pi(\mu, \nu)$ . Moreover, the optimal potential  $v(y)$  is given by  $v(y) = \int^y k(t) dt$ .

We remark that nestedness depends on all the data  $(c, \mu, \nu)$  of the problem, in the following we give a condition which can ensure nestedness when only  $c$  and  $\mu$  are fixed.

Consider now the equality  $\nu([-\infty, y]) = \mu(X_{\leq}(y, k(y)))$ , then differentiating it gives the following equation

$$\nu(y) = \int_{X_{=}(y, k(y))} \frac{D_{yy}^2 c(x, y) - k'(y)}{|D_{xy}^2 c(x, y)|} \mu(x) d\mathcal{H}^{m-1}(x), \quad (6)$$

which can be viewed as the multi-to-one counterpart of the Monge-Ampère equation in classical Optimal Transport theory (note also that it is exactly the  $n = 1$  case of (4)).

We now briefly recall some results of [15] which guarantee the nestedness of the model by checking that a lower bound on  $\nu$  satisfies a certain condition.

Fix  $y_0 < y_1$  (where  $y_0, y_1 \in Y$ ),  $k_0 \in D_y c(X, y_0)$  and set  $k_{max}(y_0, y_1, k_0) = \sup\{k : X_{\geq}(y_0, k_0) \subseteq X_{\geq}(y_1, k)\}$ . We then define the *minimal mass difference*,  $D_{\mu}^{min}$ , as follows:

$$D_{\mu}^{min}(y_0, y_1, k_0) = \mu(X_{\geq}(y_1, k_{max}(y_0, y_1, k_0)) \setminus X_{\geq}(y_0, k_0)).$$

The minimal mass difference represents the smallest amount of mass that can lie between  $y_0$  and  $y_1$ , and still have the corresponding level curves  $X_{=}(y_0, k_0)$  and  $X_{=}(y_1, k_1)$  not intersect.

**Theorem 3** (Thm. 3 [15]). *Assume that  $\mu$  and  $\nu$  are absolutely continuous with respect to Lebesgue measure. If  $D_{\mu}^{min}(y_0, y_1, k(y_0)) < \nu([y_0, y_1])$  for all  $y_0 < y_1$  where  $k(y)$  is defined by (5), then  $(c, \mu, \nu)$  is nested. Conversely, if  $(c, \mu, \nu)$  is nested, we must have  $D_{\mu}^{min}(y_0, y_1, k(y_0)) \leq \nu([y_0, y_1])$  for all  $y_1 > y_0$ .*

then the following condition on the density of  $\nu$  follows

**Corollary 4** (Cor. 4 [15]). *Assume that  $\mu$  and  $\nu$  are absolutely continuous with respect to Lebesgue measure. If for each  $y_0 \in Y$ , we have*

$$\sup_{y_1 \in Y, y_0 \leq y \leq y_1} \left[ \frac{D_{\mu}^{min}(y_0, y_1, k(y_0))}{y_1 - y_0} - \nu(y) \right] < 0,$$

then  $(c, \mu, \nu)$  is nested.

### 3 A variational approach to Cournot-Nash equilibria

We briefly recall here the variational approach to Cournot-Nash via Optimal Transport proposed in [4–6]. Agents are characterised by a type  $x$  belonging to a compact metric space  $X$  which is endowed with a given probability measure  $\mu \in \mathcal{P}(X)$  which gives the distribution of types in the agent population. For sake of simplicity we will consider probability measures absolutely continuous with respect to the Lebesgue measure and, with a slight abuse of notation we will identify the measure with its density. Each agent must choose a strategy  $y$  from a space strategy space  $Y$  (a compact metric space again) by minimizing a given cost  $\Phi : X \times Y \times \mathcal{P}(Y) \rightarrow \mathbb{R}$ . Notice that the cost  $\Phi(x, y, \nu)$  of one agent does not depend only on the type of the agent and the chosen strategy but also on the other agents' choice through the probability distribution  $\nu$  resulting from the whole population strategy choice. Then, an equilibrium can be described by a joint probability distribution  $\gamma$  on  $X \times Y$  which gives the joint distribution of types and strategies and is consistent with the cost-minimising behaviour of agents.

**Definition 5** (Cournot-Nash equilibrium). *Given a cost  $\Phi : X \times Y \times \mathcal{P}(Y) \rightarrow \mathbb{R}$ , a probability measure  $\mu \in \mathcal{P}(X)$ , a Cournot-Nash equilibrium is a probability measure on  $\mathcal{P}(X \times Y)$  such that*

$$\gamma(\{(x, y) \in X \times Y : \Phi(x, y, \nu) = \max_{z \in Y} \Phi(x, z, \nu)\}) = 1, \quad (7)$$

and  $\pi_x(\gamma) = \mu$ ,  $\pi_y(\gamma) = \nu$ .

Here we consider Cournot-Nash equilibria in the separable case that is

$$\Phi(x, y, \nu) = c(x, y) + F(\nu, y),$$

where  $c$  is a transport cost and the term  $F(\nu, y)$  captures all the strategic interactions. In particular  $F$  can be of the form

$$F(\nu, y) = f(\nu) + \int_Y \phi(y, z) d\nu(z),$$

where

- $f(\nu)$  captures the congestion effects: frequently played strategies are costly;
- $\int_Y \phi(y, z) d\nu(z)$  captures the positive interactions.

**Remark 6.** *In order to understand better the role played by the term  $F(\nu, y)$ , let us consider again the example of the physicians and the cities that we explained in the introduction. In this case the distribution  $\nu$  of the types of the cities, namely the strategies the physician have to choose, is unknown. The main idea of the Cournot-Nash equilibria is to find the matching  $\gamma$  such that the agents minimize a transport cost (again this can represent the cost of commuting) and an extra term which takes into account*

- *a repulsive effect (the congestion): it is self-defeating opening a practice in a city chosen already by many other physicians;*
- *an attractive effect (the positive interactions term): the physicians want to share their own experiences with their peers.*

In [6] the authors observed that  $F$  is the first variation of the energy

$$\mathcal{F}(\nu) = \int_Y f(\nu(y)) dy + \int_{Y \times Y} \phi(y, z) d\nu(y) d\nu(z),$$

and consequently proved that the condition defining the equilibria is in fact the Euler-Lagrange equation for the following variational problem

$$\inf\{\mathcal{J}(\mu, \nu) \mid \nu \in \mathcal{P}(Y)\}, \quad (8)$$

where

$$\mathcal{J}(\mu, \nu) = \mathcal{T}_c(\mu, \nu) + \mathcal{F}(\nu).$$

Indeed if  $\nu$  solves (8) and  $\gamma$  is the optimal transport plan between  $\mu$  and  $\nu$  for the cost  $c(x, y)$  then  $\gamma$  is a Cournot-Nash equilibrium for the cost  $\Phi(x, y, \nu)$  defined as above. Notice that once the cost function satisfies the twisted condition (also known in economics as the Spence-Mirrlees condition) then  $\gamma$  is determinist, that is  $\gamma$  is a pure equilibrium (see, for example, [18]).

### 3.1 Nestedness of Cournot-Nash equilibrium

We now briefly recall the main results in [15] assuring that nestedness condition holds. Moreover, in the next section, we will see how nestedness can lead to the development of numerical methods to solve (8).

**The congestion case** We firstly consider the case  $n = 1$ , meaning that we are in the multi-to-one Optimal Transport problem, and the functional  $\mathcal{F}(\nu)$  takes into account only the congestion effects. Indeed we consider the following functional

$$\mathcal{F}(\nu) = \int_Y f(\nu) dy$$

with  $f : [0, \infty) \rightarrow \mathbb{R}$  continuously differentiable on  $[0, \infty)$ , strictly convex with super-linear growth at infinity and satisfying

$$\lim_{x \rightarrow 0^+} f'(x) = -\infty.$$

An example is the entropy  $f(\nu) = \nu \log(\nu)$ . In particular we have this result establishing lower and upper bounds on the density  $\nu$  and the nestedness of the model.

**Theorem 7** (Thm. 11 [15]). *Assume that  $\nu \in \mathcal{P}(Y)$ , with  $Y = (0, \bar{y})$  is a minimizer of (8) then  $\nu$  is absolutely continuous w.r.t. the Lebesgue measure and there exist two constants  $M_1$  and  $M_2$  depending on the cost function and  $Y$  such that*

$$(f')^{-1}(M_1) \leq \nu(y) \leq (f')^{-1}(M_2)$$

. Moreover,  $(c, \mu, \nu)$  is nested provided

$$\sup_{y_1 \in Y, y_0 \leq y \leq y_1} \frac{D_\mu^{\min}(y_0, y_1, k(y_0))}{y_1 - y_0} - (f')^{-1}(M_1) < 0$$

for all  $y_0 \in Y$ .

Notice that once we have established the model is nested then by setting  $k(y) = v'(y)$  we get

$$\nu(y) = \int_{X=(y, k(y))} \frac{D_{yy}^2 c(x, y) - k'(y)}{|D_{xy}^2 c(x, y)|} \mu(x) d\mathcal{H}^{m-1}(x) := G(y, k(y), k'(y)) \quad (9)$$

and differentiating the first order condition of (8) we get a second order differential equation for  $k$

$$k(y) + f''(G(y, k(y), k'(y))) \frac{d}{dy} G(y, k(y), k'(y)) = 0.$$

In section 4.1 we detail how helpful nestedness is in order to design a suitable numerical method.

**The interaction case** We consider now the case  $m > n \geq 1$  and a functional  $\mathcal{F}$  which captures only the interaction effects, that is

$$F(y, \nu) = V(y) + \int_Y \phi(y, z) d\nu(z),$$

where  $\phi$  is symmetric that is  $\phi(y, z) = \phi(z, y)$ . In this case it is actually more difficult to find lower and upper bounds as in the congestion, but still the following nestedness result holds

**Theorem 8** (Thm. 15 [15]). *Assume that  $y \mapsto c(x, y) + V(y) + \phi(y, z)$  is uniform convex throughout  $X \times Y \times Y$  and*

$$\langle (D_y c(x, y) + DV(y) + D_y \phi(y, z)) \mid n_Y(y) \rangle \geq 0 \quad \forall x \in X, z \in Y, y \in \partial Y,$$

where  $n_Y(y)$  is the outward normal to  $Y$ . Then the model  $(c, \mu, \nu)$  is nested.

Moreover in this special case one can characterize the minimizer by using the best-reply scheme that we will detail in section 4.2.

## 4 Numerics

We now describe three different numerical approaches to solve (8). In particular we are now considering discretized measures that is  $\mu = \sum_{i=1}^M \mu_i \delta_{x_i}$  ( $x_i \in \mathbb{R}^n$ ) and  $\nu = \sum_{i=1}^N \nu_i \delta_{y_i}$  ( $y_i \in \mathbb{R}^n$ ), where  $M$  and  $N$  is the number of gridpoints chosen to discretize  $X$  and  $Y$  respectively. To reduce the amount of notation here, we use the same notations for the continuous problem as for the discretized one where integrals are replaced by finite sums and  $c$  and  $\gamma$  are now  $M \times N$  matrices.

### 4.1 Congestion case

In this section we deal with a numerical method exploiting the nestedness result we provided for the congestion case. Consider the

$$\mathcal{F}(\nu) = \int_Y f(\nu) dy$$

with  $f$  satisfying the hypothesis we gave in the previous section and  $Y = (0, \bar{y})$  such that the model is nested. Then, the main idea of the iterative algorithm is actually to combine the optimality condition with the fact that in the nestedness case we can explicitly build the optimal transport map and the Kantorovich potential. We firstly recall that in this case the optimality condition of (8) reads

$$v(y) + f'(\nu) = C. \quad (10)$$

Since the model is nested we have that  $v(y) = \int_0^y k(s) ds$  which leads to

$$\nu(y) = (f')^{-1} \left( C - \int_0^y k(s) ds \right). \quad (11)$$

The idea is now to use (11) to obtain the following iterative algorithm: fix an initial density  $\nu^{(0)}$

$$\begin{cases} k^{(n)}(y) \text{ s.t } \mu(X_{\geq}(y, k^{(n)}(y))) = \nu^{(n-1)}((0, y]), \\ \nu^{(n)}(y) = (f')^{-1} \left( C^{(n)} - \int_0^y k^{(n)}(s) ds \right), \end{cases} \quad (12)$$

where  $C^{(n)}$  is computed such that  $\nu^{(n)}$  is a probability density.

**Remark 9.** *This algorithm is actually an adaptation to the unequal dimensional case of the one proposed in [5]. For both the equal and unequal case we proved in [16] a convergence result by choosing a suitable metric.*

**Remark 10.** *Take  $f(\nu) = \nu \log(\nu)$  then the update of  $\nu^{(n)}$  reads*

$$\nu^{(n)}(y) = \frac{\exp(-\int_0^y k^{(n)}(s) ds)}{\int_Y \exp(-\int_0^y k^{(n)}(s) ds) dy}$$

**Numerical results** In all the simulations we present in this paragraph we have taken the uniform density on the arc  $(0, \frac{\pi}{6})$  as initial density  $\nu^{(0)}$ ,  $\mu$  is always the uniform on quarter disk  $X = \{x_1, x_2 > 0 : x_1^2 + x_2^2 < 1\}$  and  $f(\nu)$  is the entropy as in the remark above. By Corollary 12 in [15], when the cost is given by  $c(x, y) = |x_1 - \cos(y)|^2 + |x_2 - \sin(y)|^2$  we know that the model is nested provided  $\bar{y} \leq 0.61$  (which is exactly the case we consider for the numerical tests). Moreover, we have also considered an additional potential term, which will favour a concentration in a certain area of  $(0, \frac{\pi}{6})$ , so that the functional  $\mathcal{F}$  has the form

$$\mathcal{F}(\nu) = \int_Y f(\nu) dy + \int V(y) d\nu(y),$$

where  $V(y) = 10|y - 0.1|^2$ . The second equation in (12) then becomes

$$\nu^{(n)}(y) = (f')^{-1} \left( C^{(n)} - V(y) - \int_0^y k^{(n)}(s) ds \right).$$

In Figure 1 we show the final density we have obtained (on the left) and the intersection between the level-set  $X_{\geq}(y, k^*(y))$  for the final solution and the support of the

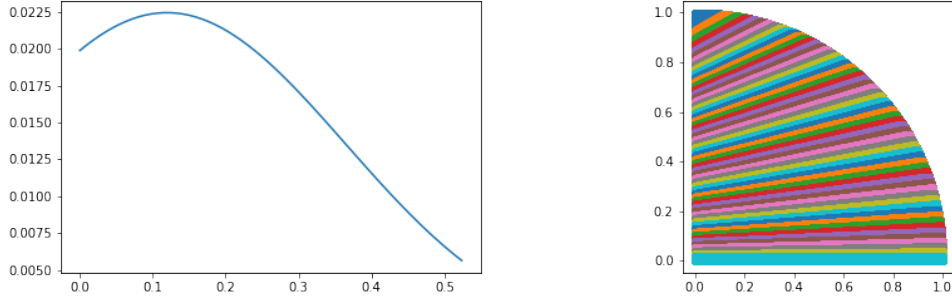


Figure 1: (Left) Final density  $\nu^*$ . (Right) Intersection between the level-set  $X_{\geq}(y, k^*(y))$  for the final solution and the support of the fix measure  $\mu$

fix measure  $\mu$ . Notice that, as expected, the level set  $X_{\geq}(y, k^*(y))$  do not cross each other meaning that the model is nested.

In Figure 2, we present some simulations by taking the same data set as above, but the cost function is now given by

$$c(x, y) = \left(1 + |x_1 - \cos(y)|^2 + |x_2 - \sin(y)|^2\right)^{\frac{p}{2}}$$

with  $p > 2$ . Notice that now the level set are not straight lines anymore.

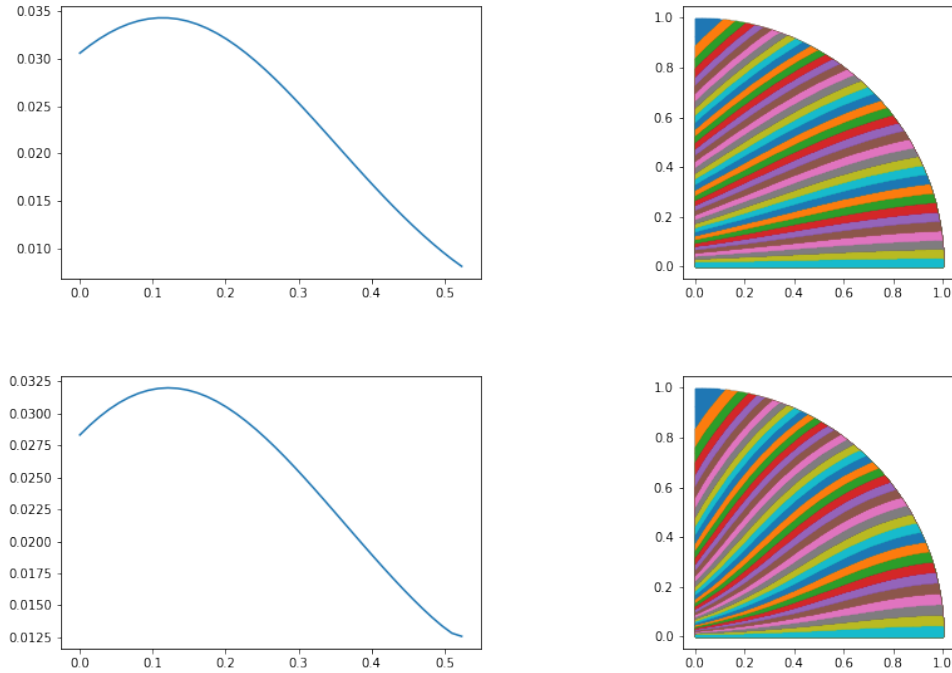


Figure 2: First row:  $p = 4$ . (Left) Final density  $\nu^*$ . (Right) Intersection between the level-set  $X_{\geq}(y, k^*(y))$  for the final solution and the support of the fix measure  $\mu$ . Second row:  $p = 8$ . (Left) Final density  $\nu^*$ . (Right) Intersection between the level-set  $X_{\geq}(y, k^*(y))$  for the final solution and the support of the fixed measure  $\mu$ .

## 4.2 Best reply scheme

Assume now that

$$F(y, \nu) = V(y) + \int_Y \phi(y, z) d\nu(z)$$



and consider the first optimality condition of (8) which reads

$$c(x, y) + F(y, \nu) = C, \quad (13)$$

where  $C$  is a constant assuring that the minimizer  $\nu$  is a probability distribution. Notice that by differentiating (13) with respect to  $y$  we get

$$D_y c(x, y) + DF(y, \nu) = 0, \quad (14)$$

we denote by  $B_\nu : X \rightarrow Y$  the map such that

$$D_y c(x, B_\nu(x)) + DF(\nu, B_\nu(x)) = 0, \quad (15)$$

which is well defined under the hypothesis of theorem 8. Then, the best-reply scheme, firstly introduced in [5], consists in iterating the application defined as

$$\mathcal{B}(\nu) := (B_\nu)_\# \mu. \quad (16)$$

We refer the reader to [15][Theorem 19] for a convergence result. Notice that [15][Theorem 19] assures also that  $\mathcal{B}(\nu)$  is absolutely continuous with respect to the Lebesgue measure, meaning that the density, and so (16), can be computed by using the co-area formula. We obtain then the following algorithm: fix  $\nu^{(0)}$  then

$$\nu^{(n+1)} = \mathcal{B}(\nu^{(n)}) := (B_{\nu^{(n)}})_\# \mu.$$

**Numerical results** As in the previous section the initial density  $\nu^{(0)}$  is the uniform on the arc  $(0, \frac{\pi}{6})$  and  $\mu$  is the uniform on  $X = \{x_1, x_2 > 0 : x_1^2 + x_2^2 < 1\}$ . Moreover, we take here a potential  $V(y) = |y - \frac{\pi}{12}|^2$  and a quadratic interaction  $\phi(y, z) = |y - z|^2$ . In Figure 3 we plot the final density obtained in the case in which the cost function is  $c(x, y) = |x_1 - \cos(y)|^2 + |x_2 - \sin(y)|^2$ . Notice that since there is not a congestion term like the entropy, the final density has a support much smaller than the interval  $(0, \frac{\pi}{6})$ .

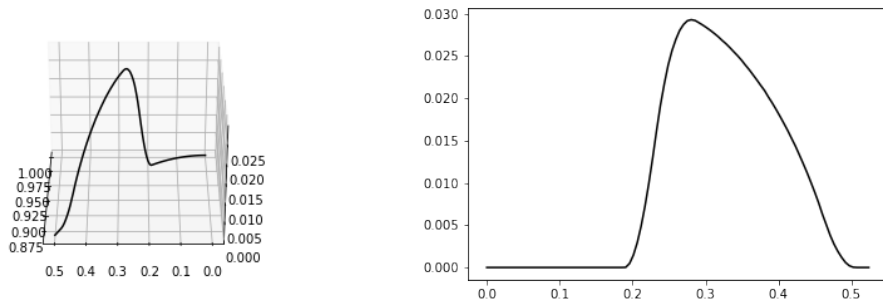


Figure 3: (Left) Final density  $\nu^*$  concentrated on the arc  $(0, \frac{\pi}{6})$ . (Right) Final density  $\nu^*$ .

In Figure 4 we repeat the simulation with the same data set but taking the cost function  $c(x, y) = \left(1 + |x_1 - \cos(y)|^2 + |x_2 - \sin(y)|^2\right)^{\frac{p}{2}}$  with  $p = 4, 8$ . Notice that the cost function plays the role of an attractive term with respect to the fix measure  $\mu$  which makes the final density have a more spread support.

### 4.3 The Entropic Regularization

The last numerical approach we consider is the entropic regularization of Optimal Transport (we refer the reader to [3, 6, 10, 11, 17] for more details) which can be easily used to solve (8) no matter the choice of the transportation cost  $c(x, y)$  and the functional  $\mathcal{F}(\nu)$ . This actually implies that the entropic regularization is clearly more flexible than the methods we have previously introduced but in some cases (i.e.

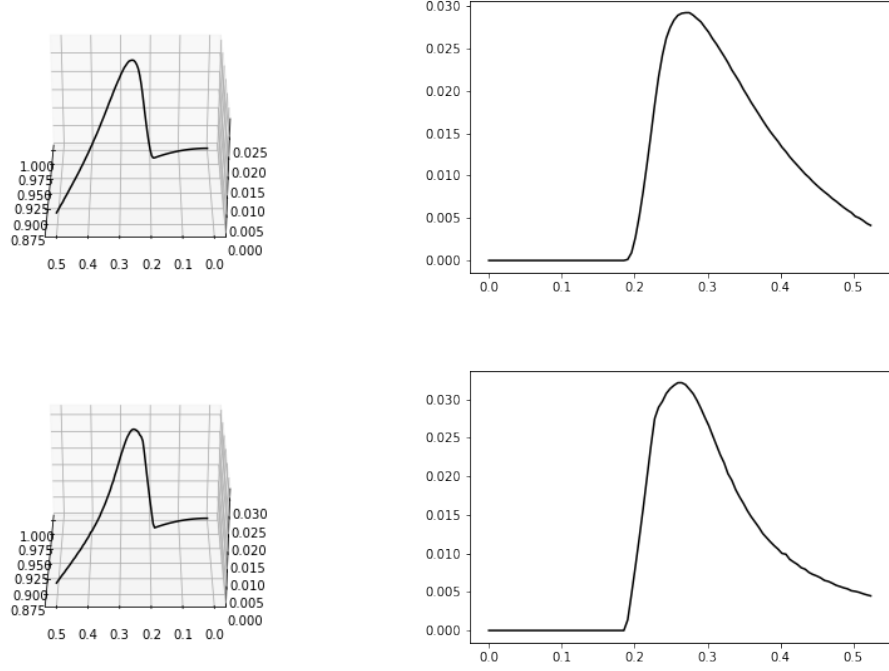


Figure 4: (Left) Final density  $\nu^*$  concentrated on the arc  $(0, \frac{\pi}{6})$ . (Right) Final density  $\nu^*$ .

the congestion term is the entropy) the major drawback of this approach is an extra diffusion. The entropic regularization of Optimal Transport can be stated as follows

$$\mathcal{T}_{c,\epsilon}(\mu, \nu) := \inf \left\{ \int_{X \times Y} cd\gamma + \epsilon \int_{X \times Y} (\log(\gamma) - 1)d\gamma \mid \gamma \in \Pi(\mu, \nu) \right\}, \quad (17)$$

where we assume that  $0(\log(0) - 1) = 0$ . The problem now is strictly convex and it can be re-written as

$$\mathcal{T}_{c,\epsilon}(\mu, \nu) := \inf \{ \epsilon \mathcal{H}(\gamma | \bar{\gamma}) \mid \gamma \in \Pi(\mu, \nu) \},$$

where  $\mathcal{H}(\gamma | \bar{\gamma}) := \int_{X \times Y} (\log(\frac{\gamma}{\bar{\gamma}}) - 1)d\gamma$  is the relative entropy and  $\bar{\gamma} = \exp(\frac{-c}{\epsilon})$ . It can be proved that (17) is strictly convex problem and it admits a solution of the form

$$\gamma^* = \exp\left(\frac{u^*}{\epsilon}\right) \bar{\gamma} \exp\left(\frac{v^*}{\epsilon}\right),$$

where  $u^*$  and  $v^*$  are the Lagrange multipliers associated to the marginal constraints. We also highlight that by adding the entropic term we have penalized the non-negative constraint meaning that the solution  $\gamma^*$  is always positive.

The regularized version of (8) can be stated in the following way

$$\inf \{ \epsilon \mathcal{H}(\gamma | \bar{\gamma}) + \mathcal{F}(\pi_y(\gamma)) + \mathcal{G}(\pi_x(\gamma)) \} \quad (18)$$

where  $\mathcal{F}$  is the functional capturing the congestion and interactions effects and  $\mathcal{G}(\rho) = i_\mu(\rho)$  is the indicator function in the convex analysis sense and it is used to enforce the prescribed first marginal. Before introducing the generalization of Sinkhorn algorithm proposed in [10] we briefly recall without proof a classical duality result

**Proposition 11.** *The dual problem of (18)*

$$\sup_{u,v} -\mathcal{G}^*(-u) - \mathcal{F}^*(-v) - \epsilon \int_{X \times Y} \exp\left(\frac{u^*}{\epsilon}\right) \bar{\gamma} \exp\left(\frac{v^*}{\epsilon}\right) dx dy. \quad (19)$$

Moreover strong duality holds.

The generalized Sinkhorn algorithm is then obtained by relaxation of the maximizations on the dual problem (19). We get the iterative method computing a sequence of potentials: given 2 vectors  $u^{(0)}$  and  $v^{(0)}$ , then the update at step  $n$  is defined as

$$\begin{cases} u^{(n)} := \operatorname{argmax}_u -\mathcal{G}^*(-u) - \epsilon \int_{X \times Y} \exp\left(\frac{u}{\epsilon}\right) \bar{\gamma} \exp\left(\frac{v^{(n-1)}}{\epsilon}\right), \\ v^{(n)} := \operatorname{argmax}_v -\mathcal{F}^*(-v) - \epsilon \int_{X \times Y} \exp\left(\frac{u^{(n)}}{\epsilon}\right) \bar{\gamma} \exp\left(\frac{v}{\epsilon}\right). \end{cases} \quad (20)$$

**Remark 12.** For many interesting functionals  $\mathcal{F}$  the relaxed maximizations can be computed point-wise in space and analytically. Notice that a problem can arise in treating the interaction term, but one can easily overcome this difficulty by adopting a semi-implicit approach as the one proposed in [7].

**Numerical results**  $\mu$  is the uniform measure on  $X = \{x_1, x_2 > 0 : x_1^2 + x_2^2 < 1\}$ . Moreover, we take here a functional  $\mathcal{F}$ , as in the best reply scheme section, given by the sum of a potential  $V(y) = |y - \frac{\pi}{12}|^2$  and a quadratic interaction  $\phi(y, z) = |y - z|^2$ . We can then compare the numerical results obtained by the best reply scheme and generalized Sinkhorn, with a regularization parameter  $\epsilon = 10^{-3}$ , in the case of the quadratic cost function.

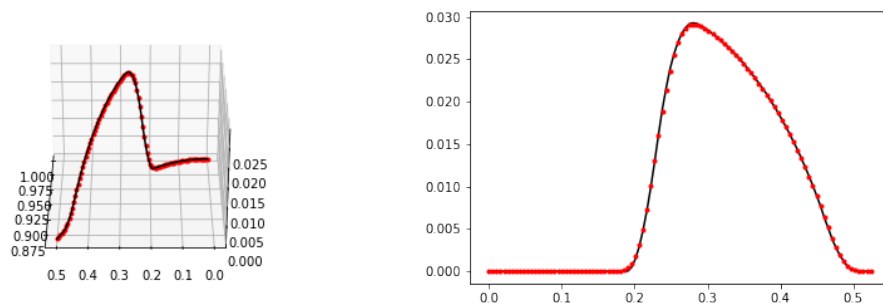


Figure 5: (Left) Final density  $\nu^*$  obtained by the best reply scheme (dark solid line) and by entropic regularization (red dots) concentrated on the arc  $(0, \frac{\pi}{6})$ . (Right) Final density  $\nu^*$  obtained by the best reply scheme (dark solid line) and by entropic regularization (red dots).

**Remark 13.** Notice that we have not compared the iterative method for the congestion case with the entropic regularization. This is almost due to the fact that adding an other entropy term to the problem will cause an extra diffusion which could be difficult to treat even by choosing a small regularization parameter.

## References

- [1] R. Aumann. Existence of competitive equilibria in markets with a continuum of traders. *Econometrica*, 32:39–50, 1964.
- [2] R. Aumann. Markets with a continuum of traders. *Econometrica*, 34:1–17, 1966.
- [3] Jean-David Benamou, Guillaume Carlier, Marco Cuturi, Luca Nenna, and Gabriel Peyré. Iterative Bregman projections for regularized transportation problems. *SIAM J. Sci. Comput.*, 37(2):A1111–A1138, 2015.
- [4] Adrien Blanchet and Guillaume Carlier. From Nash to Cournot-Nash equilibria via the Monge-Kantorovich problem. *Philos. Trans. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci.*, 372(2028):20130398, 11, 2014.
- [5] Adrien Blanchet and Guillaume Carlier. Remarks on existence and uniqueness of Cournot–Nash equilibria in the non-potential case. *Mathematics and Financial Economics*, 8(4):417–433, 2014.

- [6] Adrien Blanchet and Guillaume Carlier. Optimal transport and Cournot-Nash equilibria. *Math. Oper. Res.*, 41(1):125–145, 2016.
- [7] Adrien Blanchet, Guillaume Carlier, and Luca Nenna. Computation of cournot-nash equilibria by entropic regularization. *Vietnam Journal of Mathematics*, Sep 2017.
- [8] P.-A. Chiappori, R. McCann, and B. Pass. Multidimensional matching. *ArXiv e-prints*, April 2016.
- [9] Pierre-André Chiappori, Robert J. McCann, and Brendan Pass. Multi-to one-dimensional optimal transport. *Communications on Pure and Applied Mathematics*, pages n/a–n/a.
- [10] L. Chizat, G. Peyré, B. Schmitzer, and F.-X. Vialard. Scaling algorithms for unbalanced transport problems. Technical report, <http://arxiv.org/abs/1607.05816>, 2016.
- [11] M. Cuturi. Sinkhorn distances: Lightspeed computation of optimal transport. In *Advances in Neural Information Processing Systems (NIPS) 26*, pages 2292–2300, 2013.
- [12] A. Galichon and B. Salanié. Matching with trade-offs: Revealed preferences over competing characteristics. Technical report, Preprint SSRN-1487307, 2009.
- [13] A. Mas-Colell. On a theorem of Schmeidler. *J. Math. Econ.*, 3:201–206, 1984.
- [14] Robert J McCann and Brendan Pass. Optimal transportation between unequal dimensions. *arXiv preprint arXiv:1805.11187*, 2018.
- [15] Luca Nenna and Brendan Pass. Variational problems involving unequal dimensional optimal transport. *Journal de Mathématiques Pures et Appliquées*, 139:83–108, 2020.
- [16] Luca Nenna and Brendan Pass. Transport type metrics on the space of probability measures involving singular base measures. *arXiv preprint arXiv:2201.00875*, 2022.
- [17] Gabriel Peyré. Entropic approximation of Wasserstein gradient flows. *SIAM J. Imaging Sci.*, 8(4):2323–2351, 2015.
- [18] Filippo Santambrogio. *Optimal transport for applied mathematicians*. Progress in Nonlinear Differential Equations and their Applications, 87. Birkhäuser/Springer, Cham, 2015. Calculus of variations, PDEs, and modeling.
- [19] C. Villani. *Optimal Transport: Old and New*, volume 338 of *Grundlehren der mathematischen Wissenschaften*. Springer, 2009.