Computation of Cournot-Nash equilibria by entropic regularization

Tom Rossa

January 2025

Computational Optimal Transport - MVA

Keywords: Cournot-Nash equilibria - Optimal Transport - Entropic Regularization - Proximal Splitting

Abstract

This project investigates a class of continuum games, frequently encountered in economics and game theory. In this context, each player seeks to maximize their utility, which depends on their individual strategy choice and the aggregated strategy distribution, on which they have negligible influence. A common objective in such settings is to find an equilibrium distribution, known as the Cournot-Nash equilibrium. This project is built upon the work of Blanchet and al. 2018, which formulates this problem using a variational approach linking it to optimal transport theory. Then, the introduction of an entropic regularization term simplify its numerical resolution and transforms the problem into an instance of Bregman proximal splitting. By leveraging a method from the literature (Peyré 2015), they propose an efficient algorithm for solving the regularized problem. This report provides a review of the historical connection between Cournot-Nash equilibria and optimal transport theory, followed by a detailed exposition of the proximal splitting method used to find these equilibria. Finally, numerical experiments are presented, demonstrating the performance of the proposed method under various settings, depending on the strength of the regularization.

1 Introduction

The concept of Nash equilibrium, first introduced by John Nash in his seminal work in 1950 (Nash 1951), has become a cornerstone of game theory and economic modeling. This equilibrium represents a state in which no player can improve their payoff by unilaterally changing their strategy, given the strategies of others. By formalizing decision-making in competitive and strategic environments, Nash's framework has provided a powerful lens for analyzing non-cooperative interactions among rational agents.

Definition: Nash Equilibrium
Consider a setting with N players, each having a strategy space Y_i and a cost function $J_i : Y_i \times Y_{-i} \to \mathbb{R}$. A Nash equilibrium is formally defined as a collection of strategies $\bar{y} = (\bar{y}_1, \bar{y}_2, \dots, \bar{y}_N)$ such that:

$$J_i(\bar{y}_i, \bar{y}_{-i}) \le J_i(y_i, \bar{y}_{-i}) \quad \forall y_i \in Y_i, \quad \forall i \in \{1, 2, \dots, N\},$$

where $y_{-i} = \{y_1, \dots, y_{i-1}, y_{i+1}, \dots, y_N\}$ represents the strategies of all players except player i.

In this project, we focus on a well-known class of games in economics involving a continuum of players, where each individual player has a negligible marginal impact on the cost functions of others. In this framework, an equilibrium can be determined by minimizing a specific functional, which reformulates the problem as an instance of optimal transport.

Games with a Continuum of Players. Introduced in the seminal work of Aumann 1966, these games extend the classical framework of game theory by considering an infinite set of players modeled as a probability space:

$$(\mathcal{I}, \Sigma, \mu) = \begin{cases} \mathcal{I} & : \text{ the set of players,} \\ \Sigma & : \text{ a } \sigma\text{-algebra defining measurable subsets,} \\ \mu & : \text{ a positive, nonatomic measure on } (\mathcal{I}, \Sigma). \end{cases}$$

The strategy space is similarly modeled as a probability space $(\mathcal{Y}, \mathcal{A}, \nu)$. In such games, payoffs depend on an aggregate function of individual strategies, and the equilibrium is defined as a distribution over the strategy space. A defining feature of these games is that individual decisions are negligible due to the nonatomicity of μ, meaning that the actions of a single player do not influence the aggregate behavior of the population.

In this framework, a Cournot-Nash equilibrium is a distribution $\nu^* \in \mathcal{P}(\mathcal{Y})$ such that no player has an incentive to deviate from their strategy, given the aggregate distribution ν^*. In Schmeidler 1973, the author establishes the existence of such an equilibrium under the following conditions:

Theorem: (Existence of a Nash Equilibrium)
Let $\mathcal{X}$ be compact, the cost function J_i continuous and convex, and the distribution of player types μ nonatomic. Then, there exists ν^* such that

$$\forall i \in \mathcal{I}, \; y_i^* = \arg\min_{y \in \mathcal{Y}_i} J_i(y, \nu^*) \neq \emptyset,$$

where the distribution of strategies across players $\pi(y^*) = \nu^*$ is consistent under reconstruction.

This theorem is proven using arguments from convex optimization and the fixed-point theorem. It is shown that under the stated conditions, the optimal strategy distribution is a fixed point of the best response operator $\mathcal{F}$, which maps a given strategy distribution ν to the best responses of all players:

$$\forall \nu \in \mathcal{P}(\mathcal{Y}), \ \mathcal{F}(\nu) = \rho,$$

where $d\rho(y) = \arg\min_{y \in \mathcal{Y}} J(y, \nu)$. Building on these foundational results, Mas-Colell 1984 reformulated this problem, revealing profound connections between this field of study and Optimal Transport Theory.

Cournot-Nash Equilibrium. In this reformulation, we now consider γ as a joint distribution over $\Theta \times \mathcal{Y}$. For each player, a type x is associated, along with a strategy y chosen by that player. The cost function becomes

$$\Phi : \Theta \times \mathcal{Y} \times \mathcal{P}(\mathcal{Y}) \mapsto \mathbb{R},$$

where it takes a player's type, an individual strategy, and an aggregate set of strategies, providing the individual utility of the decision for that player. This approach is more suited to games with incomplete information, as it operates at the distribution level rather than the individual choice level. Additionally, it allows for the exploitation of geometric and topological properties of measure spaces. Intuitively, a Cournot-Nash equilibrium is characterized similarly to the traditional definition: each player has no incentive to deviate from their strategy, and the aggregation of strategies must align with the individual choices.

Problem Formulation. The problem that we will solve numerically in the following stages of this project is formulated in the paper by A. Blanchet and Carlier 2016. Staying within the same framework as before, and with μ as the distribution of player types, a Cournot-Nash equilibrium is defined as follows:

Cournot-Nash Equilibrium

A Cournot-Nash equilibrium is a joint distribution over $\mathcal{P}(\Theta \times \mathcal{Y})$ such that

$$\begin{cases} P_{1\#}\gamma = \mu, \ \text{Feasibility Condition} \\ P_{2\#}\gamma = \nu, \ \text{Self-consistency Condition} \\ \gamma(\{(x,y) \in \Theta \times \mathcal{Y}; \quad \Phi(x,y,\nu) = \arg\min_{z \in \mathcal{Y}} \Phi(x,z,\nu)\}) = 1 \end{cases}$$

where $P_1(x,y) = x$ and $P_2(x,y) = y$. This equilibrium is called "pure" if γ is supported by a graph, i.e., each player type is uniquely associated with a strategy y.

The self-consistency condition of the strategy distribution is ensured by the constraint on the second marginal of γ. Thus, when the cost function Φ is linearly separable, in the form

$$\forall (x,y,\nu) \in \Theta \times \mathcal{Y} \times \mathcal{P}(\mathcal{Y}), \quad \Phi(x,y,\nu) = c(x,y) + V(\nu)(y),$$

where $c(x,y)$ represents an individual cost associated with the choice y for the type x, and $V(\nu)$ represents a cost dependent on the aggregated distribution of strategies. The term $V(\nu)$ captures two distinct types of effects that may be:

- Repulsive or congestion effects, meaning that if a large number of players choose specific strategies, the cost increases.
- Attractive or mimetic effects, where a strategy that is very different from others becomes more costly.

Thus, we can describe $V(\nu)$ as:

$$\forall y \in \mathcal{Y},\ V(\nu)(y) = f(y,\nu) + \int \phi(y,u)d\nu(u),$$

where f is a continuous and increasing congestion function, and ϕ is a continuous interaction kernel.

Variationnal Approach. Let $\Pi(\mu,\nu)$ denote the set of transport plans on $\Theta \times \mathcal{Y}$ whose marginals are μ and ν. We define a certain energy functional E as follows:

$$\forall \nu \in \mathcal{P}(\mathcal{Y}), \quad E(\nu) := \int_{\mathcal{Y}} F(y)\, d\nu(y) + \frac{1}{2}\int_{\mathcal{Y}^2} \phi(y_1,y_2)\, d\nu(y_1)d\nu(y_2)$$

where $F(y) = \int_0^t f(s)\, ds$.

With this specification, we can reduce the problem to an optimal transport problem by defining the transport cost function as:

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

Thus, the associated problem becomes:

$$\inf_{\nu\in\mathcal{P}(\mathcal{Y})} \mathcal{W}_c(\mu,\nu) + E(\nu), \tag{1}$$

This formulation allows us to approach the problem using a variational method for determining the equilibrium, which has several advantages compared to fixed-point methods. Notably, the following result from Nenna and Pass 2024:

Theorem: (Characterization of the Cournot-Nash Equilibrium)
If

1. ν is a solution to problem (1),
2. γ is the optimal transport plan between μ and ν for the cost $c(x,y)$, such that

$$\int_{\Theta\times\mathcal{Y}} c(x,y)\, d\gamma(x,y) = \mathcal{W}_c(\mu,\nu),$$

then γ is a Cournot-Nash equilibrium for the cost $\Phi(x,y,\nu)$ defined as above.

Notice that once the cost function satisfies the Spence-Mirrlees condition, then γ is deterministic, i.e., γ is a pure equilibrium. The cost function $c(x,y)$ satisfies the Spence-Mirrlees condition if

$$\frac{\partial^2 c(x,y)}{\partial x \partial y} > 0, \quad \forall (x,y) \in \mathcal{X}\times\mathcal{Y}.$$

This condition ensures that $c(x,y)$ is strictly convex in y for each x and strictly increasing in x for each y. In this case, $\gamma = (Id,T)_{\#}\mu$ for some Borel map $T : \Theta \mapsto \mathcal{Y}$. The equilibrium is thus induced by an optimal transport map that solves the Monge problem (...), and this optimal map acts in the same way as the optimal response operator, associating a unique response in the strategy space to each player type:

$$\forall y \in \mathcal{Y}, \quad c(x,T(x)) + V(T(x),\nu) \le c(x,y) + V(y,\nu).$$

In this context, we can apply various numerical methods from optimal transport theory to compute the Cournot-Nash equilibrium numerically. It is, in fact, established in Kahn 1989 and recalled in A. Blanchet and Carlier 2016 that conditions for the existence of such an equilibrium are as follows:

Conditions for the existence of a Cournot-Nash equilibrium
If

1. m_0 is a non-negative Borel reference measure,
2. The cost function is additively separable,
3. $\nu \in \mathcal{P}(\mathcal{Y}) \cap \mathcal{L}^1(m_0)$,
4. $\nu \mapsto V[\nu]$ is continuous from $(\mathcal{P}(\mathcal{Y}), w\_^*)$ to $(C(\mathcal{Y}), \|\cdot\|_\infty)$,

then there exists a Cournot-Nash equilibrium γ.

It is important to note that condition 4, where $\mathcal{P}(\mathcal{Y})$ is endowed with the $weak\_^*$ topology, is satisfied when $V[\nu](y) = \int_{\mathcal{Y}} \phi(y,z) d\nu(z)$ with ϕ continuous. In practice, although we have existence results for such an equilibrium (Kahn 1989), we cannot be certain that this equilibrium will be unique. Additionally, several challenges arise when seeking explicit solutions to the problem. Therefore, there is a need to establish efficient numerical methods for computing approximations to the solution.

2 Numerical Methods

We now work within the framework of the article Blanchet and al. 2018, where the distributions of player types and strategies (denoted by μ and ν, respectively) are considered discrete. This allows us to rewrite problem (1) as:

$$\inf_{\nu \in \mathcal{P}(\mathcal{Y})} MK(\nu) + E(\nu) \tag{2}$$

where

- $(\phi_{i,j})_{i,j}$ is the interaction cost matrix,
- $MK(\nu) = \inf_{\gamma \in \Pi(\mu,\nu)} \{c \cdot \gamma = \sum_{i,j} c_{i,j}\gamma_{i,j}\}$,
- $E(\nu) := \sum_j F_j(\nu_j) + \frac{1}{2}\sum_{k,j \in J \times J} \varphi_{kj}\nu_k\nu_j$.

2.1 Entropic Regularization

In practice, solving problem (2) can be challenging, even when the energy function E is convex. Indeed, the term $MK(\nu)$ itself presents a complex problem to solve. Therefore, the authors propose introducing an entropic regularization term, which simplifies the problem considerably in practice. This results in the introduction of a new term:

$$MK_\epsilon(\nu) := \inf_{\gamma \in \Pi(\mu,\nu)} \{c \cdot \gamma + \epsilon \sum_{i,j \in I \times J} \gamma_{i,j}(\ln(\gamma_{i,j}) - 1)\} \tag{3}$$

$$= \inf_{\gamma \in \Pi(\mu,\nu)} \{c \cdot \gamma + \epsilon KL(\gamma | \mathbf{1})\} \tag{4}$$

where $KL(\gamma|\mathbf{1})$ is simply the negative of the Shannon-Boltzmann entropy. The Kullback-Leibler divergence is strongly convex, and therefore, the introduction of this regularization term makes the objective function ϵ-strongly convex. As a result, the minimization problem (3) admits a unique solution. Thus, an approximation of the solution to (2) is given by the smooth minimization problem:

$$\inf_{\nu \in \mathcal{P}(\mathcal{Y})} MK_\epsilon(\nu) + E(\nu) \tag{5}$$

or equivalently:

$$\inf_{\nu \in \mathcal{P}(\mathcal{Y}), \gamma \in \pi(\mu,\nu)} \{c \cdot \gamma + \epsilon \sum_{i,j \in I \times J} \gamma_{i,j}(\ln(\gamma_{i,j}) - 1) + E(\nu)\} \tag{6}$$

The first-order optimality condition yields the Gibbs distribution for the solution γ^*:

$$\forall i, j, \quad \gamma_{ij} = a_i \exp\left(-\frac{1}{\epsilon}\left(c_{ij} + f_j(\nu_j) + \sum_{k \in J} \phi_{kj}\nu_k\right)\right)$$

where a_i is the normalization constant inherent to the feasibility condition.

In addition to allowing the reformulation of the initial problem in a smoother form, facilitating the development of efficient numerical methods, this reformulation retains an intuitive interpretation within game theory. The regularization term can be viewed as including a random component ($\epsilon X_{i,j}$) in the total cost, where ($X_{i,j}$) are i.i.d. random variables. These variables can model, for instance, the presence of information asymmetry or a factor of irrationality among the players.

2.2 Proximal Splitting Algorithm

To efficiently solve the minimization program (6), the authors propose using a method known as Proximal Splitting, introduced by Peyré 2015. This algorithm can be seen as an extension of iterative Bregman projections to broader cases than indicator functions of convex sets.

Bregman Proximal Problem. We will see that problem (6) is a particular instance of Bregman proximal problems. A Bregman divergence is a measure of dissimilarity between two points in a convex space. For a convex functional Γ, the Bregman divergence $D_\Gamma(\pi, \xi)$ between two points π and ξ is defined as:

$$D_\Gamma(\pi, \xi) = \Gamma(\pi) - \Gamma(\xi) - \langle \nabla\Gamma(\xi), \pi - \xi \rangle$$

A Bregman proximal problem seeks to minimize an objective function f with a Bregman-type regularization. It can be formulated as:

$$\min_{\tilde{\pi} \in \mathcal{D}} (f(\tilde{\pi}) + \lambda D_\Gamma(\tilde{\pi}, \pi))$$

where

- $f(\pi)$ is the objective function,
- λ is a regularization parameter,
- $D_\Gamma(\tilde{\pi}, \pi)$ is the Bregman divergence relative to a reference point π,
- Γ is a smooth convex functional.

Dykstra's Algorithm. Keeping the previous notation, consider a problem of the form:

$$\min_{\pi\in\mathcal{D}}\{D_\Gamma(\pi\|\xi)+\varphi_1(\pi)+\varphi_2(\pi)\} \tag{7}$$

Consider two functions, $\varphi_1(\cdot)$ and $\varphi_2(\cdot)$, defined on $\mathcal{D}=\mathrm{dom}(\Gamma)$. Under the following assumptions:

- φ_1 and φ_2 are proper convex lower semi-continuous functions,
- φ_1 and φ_2 are coercive, ensuring the existence and uniqueness of the solution to the proximal problem for the divergence D_Γ,
- A domain qualification condition on φ_1, φ_2, and Γ, ensuring that their intersection domains are non-empty.

The Dykstra algorithm, defined later, converges to the unique solution of problem (7). This result generalizes the convergence of Bregman projection iterations, which is limited to the case where φ_1 and φ_2 are indicator functions of convex sets. In this case, the proximal problems reduce to projections according to the Bregman divergence.
A notable example is the regularized Kantorovich problem, where the Gibbs kernel is projected according to the Kullback-Leibler (KL) divergence onto the set $\mathcal{U}(\mu,\nu)$ of linear constraints on the marginals. This set is defined as:

$$\mathcal{U}(\mu,\nu)=\left\{\gamma\in\mathcal{M}_+(I\times J)\mid \gamma\mathbf{1}_J=\mu,\ \gamma^\top\mathbf{1}_I=\nu\right\},$$

where $\mathbf{1}_I$ and $\mathbf{1}_J$ are appropriately sized vectors filled with 1's. This framework provides a concrete case where Dykstra's algorithm is successfully applied.

Algorithm 1 Dykstra's Algorithm

0: **Initialization:** Set $\pi^{(0)}=\xi$, $v^{(0)}=0$, and $v^{(-1)}=0$
0: **for** $l=1,2,\dots$ **do**
0: Define $[l]_2=\mathrm{mod}(l,2)+1$ (alternating between 1 and 2)
0: Compute:

$$\pi^{(l)}\leftarrow \mathrm{Prox}_{\varphi_{[l]_2}}\left(\nabla\Gamma^*(\nabla\Gamma(\pi^{(l-1)})+v^{(l-2)})\right)$$

0: Update:

$$v^{(l)}\leftarrow v^{(l-2)}+\nabla\Gamma(\pi^{(l-1)})-\nabla\Gamma(\pi^{(l)})$$

0: **end for**=0

Thus, the paper by Peyré 2015 formalizes the proof of convergence of the successive iterations $\pi^{(l)}$ towards the solution of problem (7) under the previous conditions, using the fact that the iterations correspond to an alternate block minimization of the dual problem.

This algorithm is therefore highly useful in the context of solving the regularized problem (6), as it turns out that we can rewrite it as follows:

$$\inf_{\nu\in\mathcal{P}(\mathcal{Y}),\gamma\in\pi(\mu,\nu)}\left\{c\cdot\gamma+\epsilon\sum_{i,j\in I\times J}\gamma_{i,j}\left(\ln(\gamma_{i,j})-1\right)+E(\nu)\right\}$$

$$\iff \inf_{\nu\in\mathcal{P}(\mathcal{Y}),\gamma\in\pi(\mu,\nu)}\left\{\epsilon\sum_{i,j\in I\times J}\gamma_{i,j}\left(\ln\left(\frac{\gamma_{i,j}}{e^{-\frac{c_{i,j}}{\epsilon}}}\right)-1\right)+E(\nu)\right\}$$

Noting that $\bar{\gamma} = e^{-\frac{c_{i,j}}{\epsilon}}$, we reduce the problem to:

$$\inf_{\nu \in \mathcal{P}(\mathcal{Y}), \gamma \in \pi(\mu,\nu)} \left\{ \epsilon \big(KL(\gamma \| \bar{\gamma}) + E(\nu) \big) \right\} \iff \inf_{\gamma \in \mathbb{R}_+^{I \times J}} \left\{ KL(\gamma \| \bar{\gamma}) + G(\nu) \right\}$$

where

$$G(\gamma) := \chi_{\{\Lambda_1(\gamma)=\mu\}} + \frac{1}{\epsilon} E(\Lambda_2(\gamma))$$

Here, $\chi_{\{\Lambda_1(\gamma)=\mu\}}$ is the indicator function for the set of joint distributions having the first marginal μ, and $E(\Lambda_2(\gamma))$ is the previously presented convex and smooth energy function, applied to the second marginal of γ. We can now draw the connection with the problem of the form (7), noting that the Kullback-Leibler divergence KL is a Bregman divergence, associated with the convex functional of Shannon-Boltzmann entropy. To adapt Dykstra's algorithm to the Kullback-Leibler divergence, we make the following substitutions:

- $\Gamma = \mathcal{H}$, which is Shannon entropy, implying that $D_\Gamma = KL$.
- The variable z is defined by $z^{(l)} = \nabla \Gamma(v^{(l)})$.

With these substitutions, the iterations of the algorithm become:

Algorithm 2 Dykstra's Algorithm for KL Divergence

0: Initialize:

$$\pi^{(0)} = \bar{\gamma}, \quad z^{(0)} = \mathbf{1}, \quad z^{(-1)} = \mathbf{1}$$

0: **for** $l > 0$ **do**

0: Compute:

$$\pi^{(l)} = \mathrm{Prox}_{G_l} \left(\pi^{(l-1)} \odot z^{(l-2)} \right),$$

0: Update:

$$z^{(l)} = z^{(l-2)} \odot \pi^{(l)},$$

0: **end for**=0

The divisions $\odot$ are coordinate-wise, and $(G_l)_{i=1}^3$ is a sequence of elementary functionals for which the proximal problem is simple to compute, and such that $\sum_{l=1}^3 G_l = G$. The sequence $(G_l)_l$ is then extended by periodicity.

Practical case of Cournot-Nash equilibria. It is recalled that in our study, the energy function E is the sum of a repulsive congestion term and an attractive interaction term, as defined after (2). We decompose it as follows:

$$E(\nu) := E_2(\nu) + E_3(\nu),$$

where:

- $E_2(\nu) := \frac{1}{2} \sum_{j \in J} \nu_j^2 + \sum_{j,k \in J \times J} \varphi_{kj} \nu_k \nu_j$
- $E_3(\nu) := \sum_{j \in J} H_j(\nu_j)$ is a sum of convex congestion terms.

In this context, we can express the optimization problem as follows:

$$\inf_{\gamma \in \mathbb{R}^{I \times J}} \left\{ \mathrm{KL}(\gamma \mid \gamma) + G_1(\gamma) + G_2(\gamma) + G_3(\gamma) \right\},$$

where $G_1(\gamma)$, $G_2(\gamma)$, and $G_3(\gamma)$ are functional terms defined as follows:

- $G_1(\gamma) = \chi_{\Lambda_1(\gamma)=\mu}$
- $G_2(\gamma) = \frac{1}{\epsilon} E_2(\Lambda_2(\gamma))$
- $G_3(\gamma) = \frac{1}{\epsilon} E_3(\Lambda_2(\gamma))$

We can then apply the Dykstra algorithm, which has been explicitly defined earlier and converges well toward the optimal solution if the functions G_l satisfy the conditions given after (7). At each iteration, it is necessary to solve the proximal problem of the form $\text{prox}_{G_l}^{KL}(\bar{\gamma})$.

- **Proximal of** G_1: The proximal of G_1 corresponds to the fixed marginal constraint $\Lambda_1(\gamma) = \mu$. It is well-known and can be written in closed form as:

  $$\text{prox}_{G_1}^{KL}(\bar{\gamma}) = \mathbb{P}_\mu(\bar{\gamma}),$$

  where $\mathbb{P}_\mu(\theta)$ is the projection of θ onto the set of matrices whose margins are μ.

- **Proximal of** G_2: The proximal of G_2 is given by:

  $$\gamma_{ij} = \theta_{ij} \exp\left(-\frac{\nu_j + \sum_{k \in J} \varphi_{kj} \nu_k}{\epsilon} \right)$$

  where ν is the second margin of γ, and to obtain ν, one must solve a system using Newton's method to ensure the self-consistency of the marginals.

- **Proximal of** G_3: The proximal of G_3 is computed as follows:

  $$\gamma_{ij} = \theta_{ij} \exp\left(-\frac{H_j'(\nu_j)}{\epsilon} H_j'(\nu_j) \right)$$

  where $H_j'(\cdot)$ denotes the derivative of the convex function H_j. The ν_j's are obtained by solving the system on the marginals using Newton's method.

Thus, this scheme provides a relatively simple algorithm for implementation, which allows the regularized problem to be solved in just a few iterations.

2.3 Semi-Implicit Scheme

Then, the authors address a more general scenario where E is not convex due to the interaction term involving the symmetric matrix φ_{kj}. While there is no formal convergence guarantee for this approach, we note that if the scheme converges, it will converge to an equilibrium. Despite the lack of a theoretical convergence proof, the semi-implicit method we present has proven to be effective in practice. The core idea is to approximate the non-convex interaction term by its linearization at each step.

Instead of solving (5) directly, we transform it into a sequence of convex optimization problems. We begin with an initial guess $\nu_0 \in \mathcal{P}(\mathcal{Y})$, and for each iteration $n \geq 1$, we solve:

$$\nu^{(n+1)} = \arg \min_{\nu \in P(Y)} \left(MK_\varepsilon(\nu) + E^{(n)}(\nu) \right)$$

where in $E^{(n)}$, the interaction term is replaced by its linear approximation:

$$E^{(n)}(\nu) = \sum_{j \in J} F_j(\nu_j) + \sum_{j \in J} V^{(n)} \nu_j, \quad V^{(n)} := \sum_{k \in J} \varphi_{kj} \nu_k^{(n)} \tag{8}$$

This linearized problem can be efficiently solved using the Dykstra proximal-splitting method discussed earlier. The linear term is absorbed into the Kullback-Leibler (KL) term, simplifying the optimization to two proximal steps: one for projecting onto the fixed marginal constraint and another for handling the congestion cost.

2.4 Extension to Several Populations

The proposed approach extends naturally to the case of multiple interacting populations. For simplicity, consider two populations defined on type spaces $\Theta_1 = \{x_1^i\}_{i \in I_1}$ and $\Theta_2 = \{x_2^i\}_{i \in I_2}$, sharing a common strategy space $\mathcal{Y} = \{y_j\}_{j \in J}$. Given type distributions $\mu_1 \in P(\Theta_1)$ and $\mu_2 \in P(\Theta_2)$, and transport cost matrices $c_1 \in \mathbb{R}^{I_1 \times J}$ and $c_2 \in \mathbb{R}^{I_2 \times J}$, the optimization problem becomes

$$\inf_{(\nu_1,\nu_2) \in P(Y) \times P(Y)} \left\{ \mathrm{MK}_1^{\epsilon_1}(\nu_1) + \mathrm{MK}_2^{\epsilon_2}(\nu_2) + E_1(\nu_1) + E_2(\nu_2) + F(\nu_1 + \nu_2) \right\},$$

where $\mathrm{MK}_l^{\epsilon_l}$ is the regularized transport cost for population $l = 1, 2$, $E_l(\nu_l)$ models individual interaction costs, and $F(\nu_1 + \nu_2)$ represents a total congestion cost. Notably, the proximal step associated with F can be computed as in the single-population case by summing the contributions of ν_1 and ν_2. This formulation generalizes the framework while maintaining efficient resolution using proximal splitting methods.

3 Numerical Experiments

In this section, we present several numerical experiments to demonstrate the practical efficiency of the proximal splitting method proposed in this paper. We first consider a one-dimensional problem where the distribution μ of player types is modeled as a mixture of two Gaussians, discretized over the interval $[0, 5]$ using 150 points. The fixed cost matrix is defined as:

$$c_{i,j} = |x_i - y_j|^2,$$

where $x_i \in \Theta$ represents a specific player type, and $y_j \in \mathcal{Y}$ denotes a strategy. Additionally, the energy function is defined as:

$$E(\nu) = \sum_{j \in J} \frac{1}{2} \nu_j^2 + \frac{1}{2} \sum_{k,j \in J \times J} \phi_{k,j} \nu_k \nu_j + \sum_{j \in J} |y_j - 2|^2,$$

where $\phi_{k,j} = 10^{-2} |x_k - y_j|^2$, and the third term acts as a confinement potential that encourages players to select strategies close to the specified potential. Notably, there is no need to explicitly implement a proximal step for this term during the algorithm's execution, as it is absorbed into the Kullback-Leibler divergence. Below, we visualize the distribution of player types and the fixed cost matrix combined with the confinement potential:

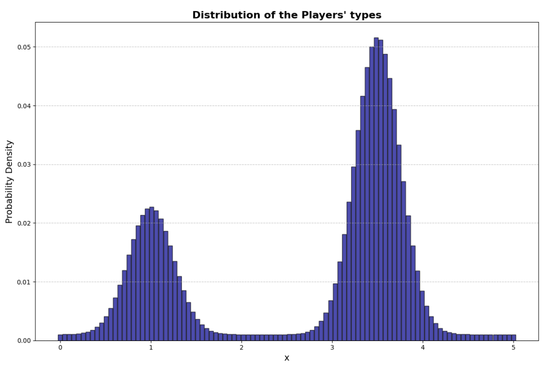

Figure 1: Distribution of players' types

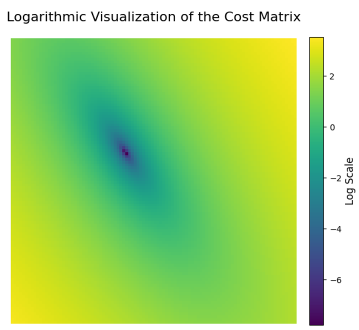

Figure 2: Visualization of Cost Matrix and Confinement Potential

In this relatively simple example, the optimization problem is strongly convex and low-dimensional. Therefore, only for this experiment, we solved the original minimization problem:

$$\inf_{\nu \in \mathcal{P}(\mathcal{Y})} MK(\nu) + E(\nu),$$

using the Python library CVXPY. This approach allows us to compute the true solution of the non-regularized problem, which we can then compare with the solutions obtained using the proximal splitting method. The solution obtained is visualized below:

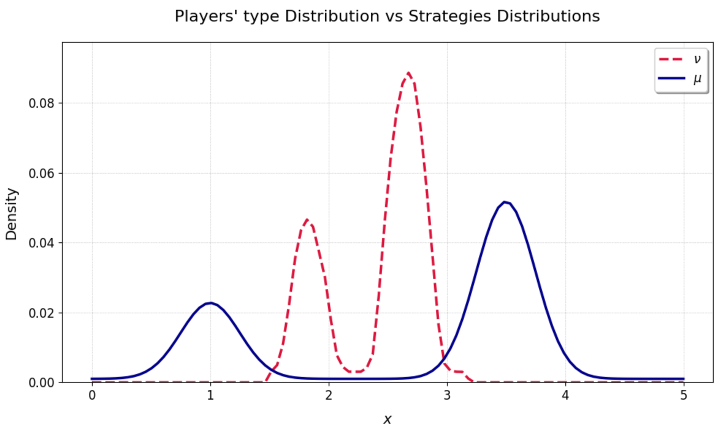

Figure 3: Solution of Non-Regularized Problem

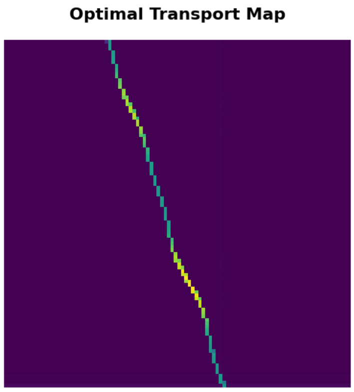

Figure 4: Optimal Transport Map

The solution to the non-regularized problem reveals that the resulting strategy distribution is significantly concentrated around the confinement potential. However, not all players choose strategies that solely minimize the fixed costs. This is due to the repulsive effect of congestion costs, which is clearly observable. Moreover, the optimal transport plan γ^* corresponding to the Nash equilibrium can be supported on a graph. The plot on the right illustrates that the equilibrium associates each player with a specific strategy, indicating that the Cournot-Nash equilibrium is pure:

$$\gamma^* = (Id, T)_{\#}\mu.$$

Subsequently, we implemented the proximal splitting algorithm using a semi-implicit scheme, which demonstrates easier practical convergence. For various values of the regularization parameter ϵ, we compare the resulting solutions after approximately 100 iterations.

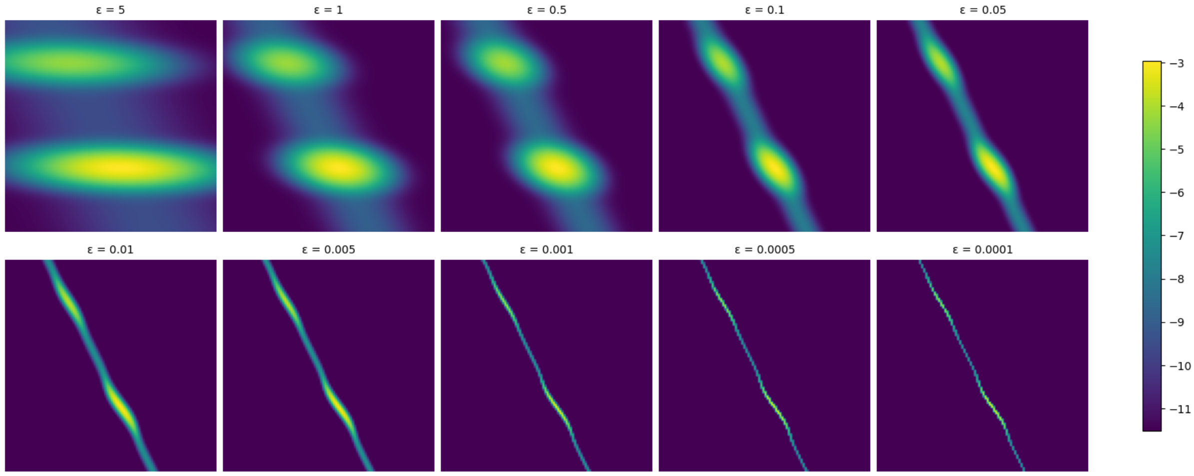

Figure 5: Comparison of Transport Plans for Different Values of ϵ

As shown, when the regularization parameter ϵ decreases, the transport plan obtained from the regularized problem approaches the theoretical Cournot-Nash equilibrium derived earlier. Unsurprisingly, higher values of ϵ produce more diffuse solutions, which also converge faster. The resulting strategy distributions for different values of ϵ are visualized below, further highlighting this behavior.

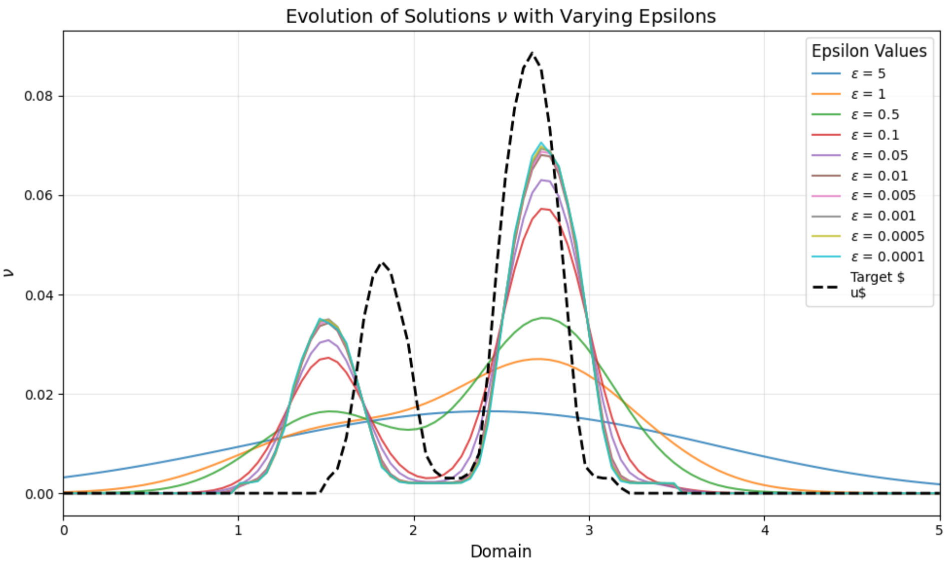

Figure 6: Comparison Distribution of Strategies with different ϵ

It is observed, however, that as ϵ decreases, the solution approaches the optimal strategy distribution. Nevertheless, a certain bias remains, as evidenced by the comparison with the theoretical distribution represented by the dashed line. This bias may stem from the approach adopted within the semi-implicit framework. This double approximation provides, in practice, a method that converges reasonably well to the optimal solution, albeit with a seemingly irreducible bias in this example. This observation is further verified by examining the evolution of the entropic bias as a function of $\log(\epsilon)$:

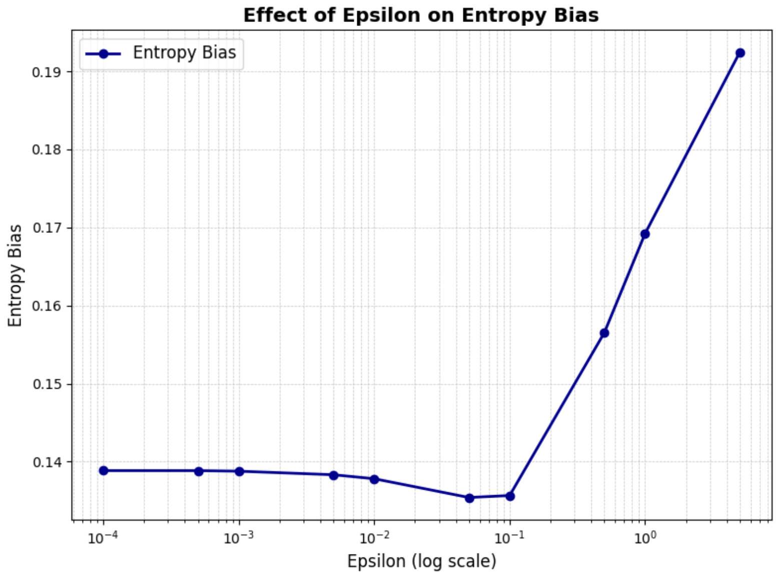

Figure 7: Evolution du biais entropique avec ϵ

The entropic bias quantifies the difference between the second marginal of γ^* and the second marginals of the various γ_ϵ, which are the respective solutions of the regularized problems. While this bias initially decreases rapidly, it stabilizes around a strictly positive value and does not appear to converge to zero as $\epsilon \to 0$.

Case in Two Dimensions. For the two-dimensional case, we consider a Gaussian mixture for the distribution of player types. The cost function is defined as $c(x, y) = \|x - y\|_2^2$, with a congestion term $F_j(\nu_j) = \frac{1}{2}\nu_j^2$, quadratic interactions $\varphi_{kj} = 10^{-4}\|y_k - y_j\|^2$, and a confinement potential $v_j = \|y_j - (2, 2)\|^4$. The simulations were conducted with $N = 25$, a discretization of the domain $[0, 5]^2$, and by treating the interaction term using a semi-implicit approach. Below, we illustrate the initial distribution as well as the average fixed cost associated with each position:

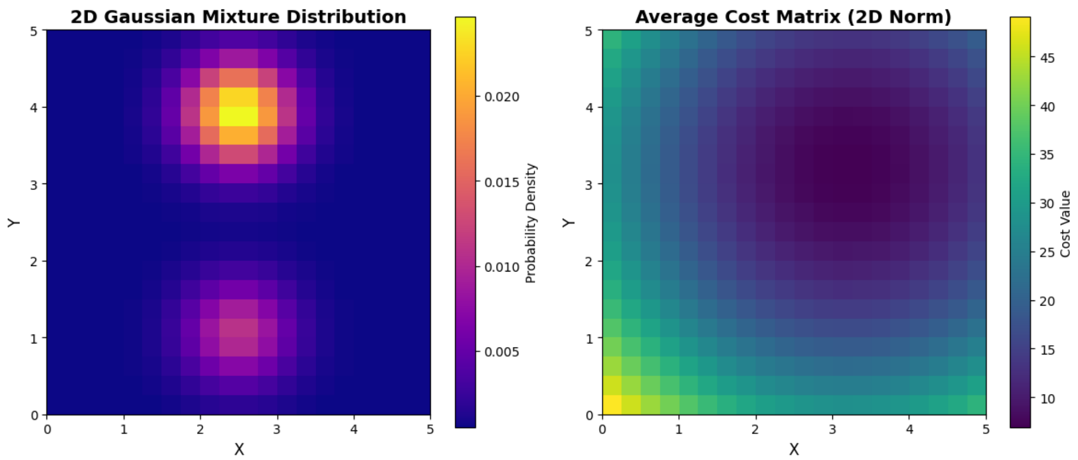

Figure 8: Initial distribution and average fixed cost associated with each position.

Similarly, using the proximal splitting algorithm, we solve the problem for various values of the regularization term ϵ. The resulting equilibrium strategy distributions are shown below:

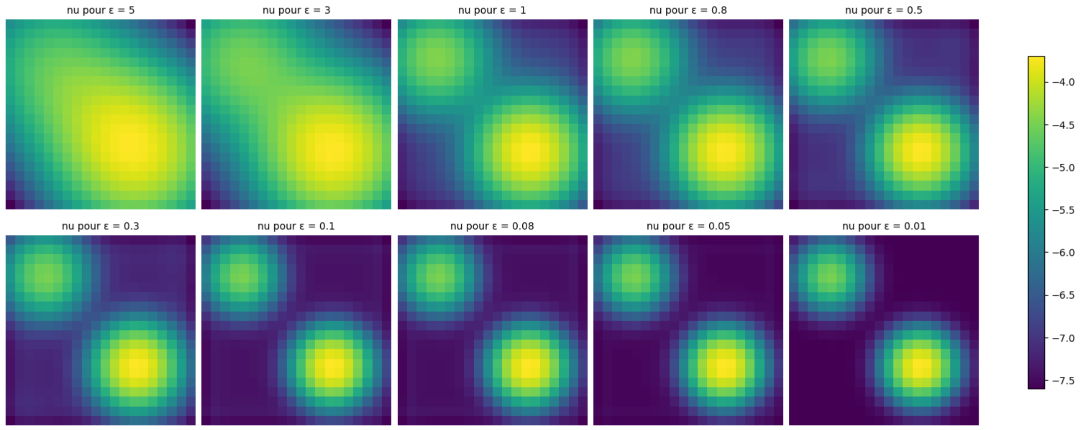

Figure 9: Equilibrium strategy distributions for various values of ϵ (2D view).

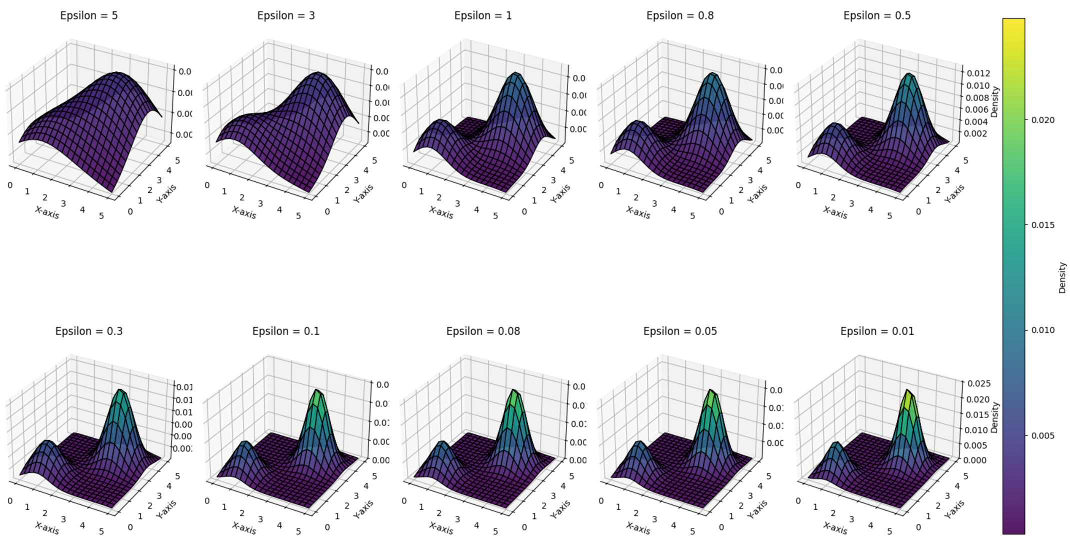

Figure 10: Equilibrium strategy distributions for various values of ϵ (3D view).

Case of Two Populations. For the case involving two populations, the distributions of player types are defined as two Gaussian mixtures, represented as follows:

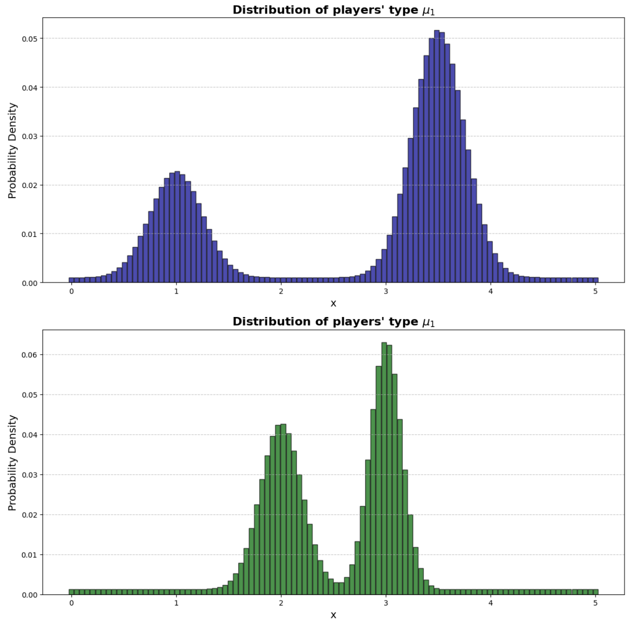

Figure 11: Distributions of player types for the two-population setup.

The energy functions $(E_l)_{l=1}^2$ are defined as:

$$E_l(\nu_l) = \sum_{j\in J}(\nu_j^l)^3 + \sum_{k,j\in J\times J}\varphi_{kj}^l \nu_j^l \nu_k^l + \sum_{j\in J}|y_j - 1|^4,$$

where the interaction terms are given by $\varphi_{kj}^l = 10^{-2}|y_k - y_j|^2$. The total congestion F_j is expressed as:

$$F_j(\nu_j^1 + \nu_j^2) = (\nu_j^1 + \nu_j^2)^3.$$

The cost functions are defined as $c_{ij} = |x_i - y_j|^2$, and the objective is to study the evolution of the solution for varying values of the regularization parameter ϵ.

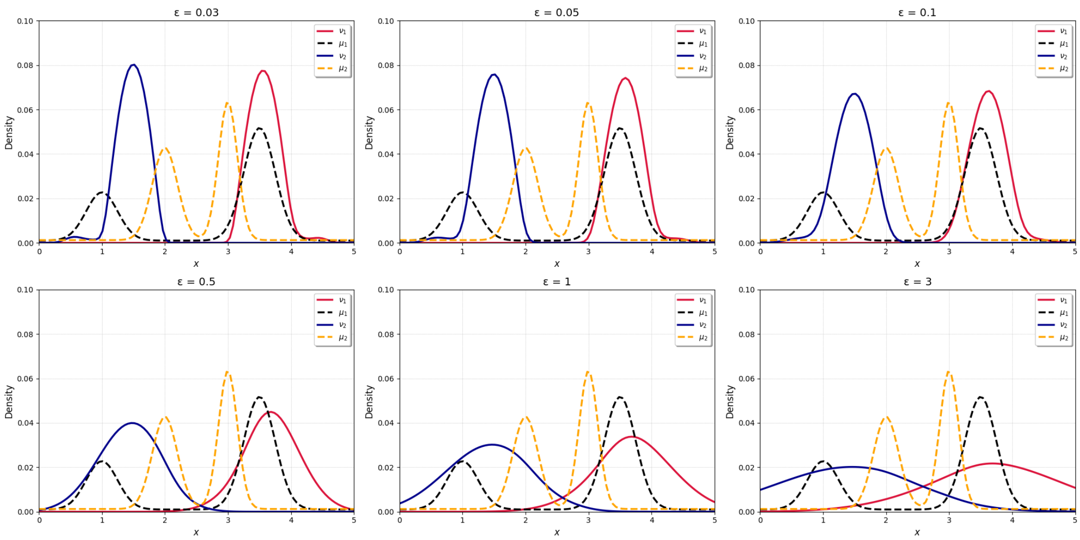

Figure 12: Comparison of equilibrium solutions for two populations under different values of ϵ.

As observed in this scenario, a decrease in the regularization parameter results in less diffuse solutions. The populations exhibit significant sensitivity to the cumulative congestion term, with individuals tending to adopt strategies that distance themselves from the other population. Additionally, as ϵ approaches zero, the convergence of the algorithm becomes noticeably slower.

4 Conclusion and Perspectives

In conclusion, this project demonstrated how the problem of finding a Cournot-Nash equilibrium in a game with a continuum of players can be reformulated as an optimal transport problem. By introducing an entropic regularization term, proximal splitting methods were leveraged to efficiently approximate the Cournot-Nash equilibrium. We empirically tested the algorithm proposed in the referenced article, particularly its semi-implicit version, and analyzed its performance across three different settings as the regularization term ϵ approached zero. In all cases, we observed that as ϵ decreased, the solution became less diffuse, and the algorithm's convergence slowed significantly. However, in scenarios where the optimal solution of the non-regularized problem was available, we noted that the solution did not appear to converge to the theoretical optimum as ϵ approached zero. This instability is likely due to the fact that the semi-implicit method introduces a double approximation of the original problem, which may limit its accuracy in recovering the exact equilibrium.

5 Connexion with the Course

The proposed project closely aligns with topics covered in the course, particularly the section on the Sinkhorn algorithm and the regularized Kantorovich problem. These concepts serve as a foundation for understanding optimal transport in the context of entropy-regularized formulations. Specifically, the iterative Bregman projection procedure, which forms the core of the Sinkhorn algorithm, solves the regularized Kantorovich dual problem through successive updates of dual variables. This method leverages the entropic penalty to ensure both computational tractability and smooth convergence. Our algorithm builds upon this framework, extending it to a more general setting that incorporates additional constraints, such as congestion effects and interaction potentials. By generalizing the iterative Bregman projection procedure, our approach enables the resolution of a broader class of optimization problems, offering enhanced flexibility and applicability in contexts beyond classical regularized optimal transport. This generalization highlights the relevance of the theoretical concepts discussed in the course while providing a practical implementation to address more complex scenarios.

6 References

References

Aumann, Robert J. (1966). "Markets with a Continuum of Traders". In: *Econometrica* 34, pp. 1–17.

Blanchet, Adrien and Guillaume Carlier (2016). "Optimal transport and Cournot-Nash equilibria". In: *Mathematics of Operations Research* 41.1, pp. 125–145.

Blanchet and al. (2018). "Computation of Cournot–Nash Equilibria by Entropic Regularization". In: *Vietnam Journal of Mathematics* 46, pp. 15–31.

Kahn, M. A. (1989). "On Cournot-Nash equilibrium distributions for games with a nonmetrizable action space and upper semi-continuous payoffs". In: *Transactions of the American Mathematical Society* 1, pp. 127–146.

Mas-Colell, A. (1984). "On a theorem of Schmeidler". In: *Journal of Mathematical Economics* 3, pp. 201–206.

Nash, John F. (1951). "Non-Cooperative Games". In: *Annals of Mathematics* 54.2, pp. 286–295.

Nenna, L. and B. Pass (2024). "A Note on Cournot-Nash Equilibria and Optimal Transport Between Unequal Dimensions". In: *Optimal Transport Statistics for Economics and Related Topics*. Vol. 483. Studies in Systems, Decision and Control. Springer, Cham.

Peyré, Gabriel (2015). "Entropic Approximation of Wasserstein Gradient Flows". In: *SIAM Journal on Imaging Sciences* 8.4, pp. 2323–2351. DOI: 10.1137/15M1010087.

Schmeidler, D. (1973). "Equilibrium points of nonatomic games". In: *Journal of Statistical Physics* 7, pp. 295–300.