

Optimal Transport in Cosmology

Vytis Krupovnickas

April 2024

1 Introduction

We begin by discussing the convex Kantorovich potential and the Monge-Ampère equation, which are central to understanding Monge's optimal transport problem. The convexity of the Kantorovich potential is shown to have important implications for the dynamics of mass transport, particularly in ensuring the absence of multi-streaming. The focus then shifts to the practical problem of finding an assignment map between points at different times, given their respective densities. This involves minimizing a specific function subject to mass conservation constraints. The Kantorovich potential, as a solution to the Monge-Ampère equation, plays a pivotal role in this process.

To solve the assignment problem numerically, we review existing methods that rely on discretizing the densities at the initial and current times. We also introduce a new semi-discrete approach that combines a continuous representation of the initial density with a discrete representation of the current density. In the discrete-discrete MAK reconstruction, we consider a scenario where both densities are represented by a set of particles. This leads to a combinatorial assignment problem, which, despite its complexity, can be tackled efficiently using polynomial-time algorithms.

We further delve into MAK duality, examining the structure of the discrete Monge problem and its connection to the gravitational potential. The Kantorovich optimization problem is presented as an alternative formulation, along with its dual problem involving Lagrange multipliers. We discuss entropic regularized methods and their advantages and drawbacks, particularly in the context of cosmological analyses. Lastly, we explain baryonic acoustic oscillations (BAO) in the universe. BAO serve as a standard ruler for measuring cosmological distances and play a crucial role in our understanding of the large-scale structure of the universe.

1.1 Optimal Transport and Cosmology

The optimal transport problem is closely related to cosmological transport in the context of studying the large-scale structure of the universe and the evolution of cosmic mass distribution. In cosmology, one of the key goals is to

understand how matter has moved and clustered over time, from the nearly uniform distribution in the early universe to the structure we observe today.

We can apply the optimal transport problem to map the distribution of matter in the early universe (the cosmic microwave background era) to its distribution at a later time (the current distribution of galaxies). Cosmological transport problems involve reconstructing the trajectories of mass elements (or galaxies) from their observed positions back to their initial positions. The optimal transport framework provides a mathematical foundation for this reconstruction, allowing astronomers to infer the dynamics of the universe's expansion and the growth of cosmic structures.

The optimal transport problem can be also be used to study Baryonic Acoustic Oscillations, or BAO, which are fluctuations in the density of visible matter in the universe. By analyzing the transport of mass in the context of BAO, cosmologists can gain insights into the properties of dark matter, dark energy, and the overall geometry of the universe. In cosmological transport problems, both discrete and continuous approaches are used. Discrete methods involve representing the mass distribution as a set of particles (galaxies or simulation particles), while continuous methods model the mass distribution as a continuous field.

The optimal transport problem can be formulated in both discrete and continuous settings, making it a versatile tool for cosmological studies. Numerical simulations of cosmic evolution often rely on solving transport problems to model the movement of mass under the influence of gravity. The optimal transport framework provides a rigorous way to define and solve these transport problems, ensuring that the simulations accurately capture the dynamics of the universe.

The static version of the optimal transport problem has been applied to astrophysical data, where the initial density fluctuation field, dating back to the early Universe, has been successfully extracted from observational data at the present epoch. This application to astrophysics was carried out in a multidisciplinary collaboration that included a few of the organizers (Brenier, Matarrese, Mohayaee) starting in 2002. The research gained substantial rapidity through the development of a new semi-discrete code by Bruno Lévy and Quentin Mérigot.

1.2 Cosmological Model

In the study of cosmology, the universe is understood to be composed of various massive objects on small scales, including but not limited to asteroids, moons, rocky planets, gas giants, stars, black holes, as well as hot and cold gases, and dark matter. These components are commonly found within galaxies, which are themselves gravitationally bound to form larger structures such as local groups and superclusters. Superclusters represent significant aggregations of galaxies that maintain coherence despite the overarching expansion of the universe, known as the Hubble flow.

The Hubble flow characterizes the recession of all objects in the universe from

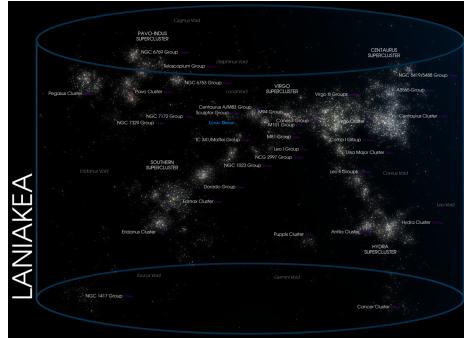


Figure 1: Laniakea Supercluster, an example of a gravitationally bound set of galaxies. Taken from Wikimedia Commons.

each other, a phenomenon that is accelerating and is predicted to lead to an infinite separation of all objects in the framework of the standard cosmological model. However, the effects of the Hubble flow are not dominant on small scales where gravitational binding prevails. The Hubble flow is given by the equation:

$$H(t) = \left(\frac{\dot{a}(t)}{a(t)} \right)^2 = \frac{8\pi G}{3} \rho(t) - \frac{\kappa c^2}{a^2(t)} + \frac{\Lambda c^2}{3} \quad (1)$$

For current time, we describe the Hubble flow as a Hubble constant, although for the purposes of describing the cosmology of the universe this constant is only useful for ascertaining the evolution of the Hubble flow.

In cosmological analysis we consider the universe on large scales, typically on the order of 100 megaparsecs (Mpc) cubes. At these scales, the universe exhibits isotropy and homogeneity, which are critical for describing its large-scale structure (Isotropy means that the universe is the same no matter which direction you are looking in, homogeneity means that the universe has the same structure and composition at every point on large scales). In this context, the universe transitions from being perceived as gravitationally bound on small scales to being dominated by the effects of the Hubble flow. This shift allows for the application of fluid equations to describe the universe's form and evolution, akin to treating it as a fluid in a colloidal solid analogy.

The following paper discussed in this lecture presents an accelerated method for solving the Monge–Ampère–Kantorovich (MAK) problem, which is essential for advanced cosmological modeling. The MAK problem has historically been challenging to compute due to its cubic time complexity. The Wasserstein distance, which corresponds to the action integral in an incompressible Euler fluid, plays a pivotal role in this context. It provides a natural metric for comparing entities of different natures, facilitating the application of the optimal transport problem to cosmological models. The optimal transport problem is closely related to the least action principle, a fundamental principle in physics that describes the path of least resistance for a system. This relationship under-

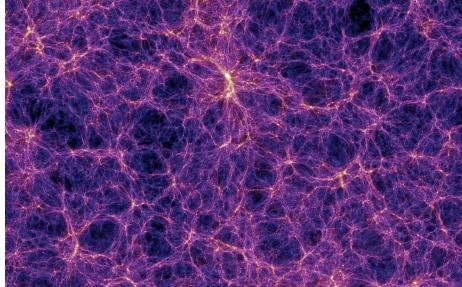


Figure 2: The filamentary structure of the universe. All visible matter is generally linked together along lines as shown in the image, leaving large voids containing negligible quantities of visible matter. Taken from Wikimedia Commons.

scores the significance of the MAK problem in understanding the dynamics of the universe and highlights the potential for more sophisticated and accurate cosmological models.

To begin, we want to look at a few important properties of describing the universe in order to clarify what problems we are approaching and how these problems are described. Thus, it is important to expand on the definition of Baryonic acoustic oscillations and the cosmological least action principle.

1.3 Cosmological Least Action Principle

Jim Peebles' cosmological least action principle is a variational principle applied to the field of cosmology. It is an extension of the classical principle of least action to a cosmological setting, where the action is minimized for the entire history of the universe. In the context of cosmology, the action is typically expressed in terms of the Einstein-Hilbert action for general relativity, coupled with additional terms for matter and radiation.

The principle states that the evolution of the universe follows a path that minimizes the action, which is a functional of the metric tensor and matter fields. This leads to the Einstein field equations when the variation of the action with respect to the metric tensor is set to zero.

The cosmological least action principle can be written as follows:

$$S = \int d^4x \sqrt{-g} \left(\frac{R}{16\pi G} + \mathcal{L}_{\text{matter}} \right)$$

where: S is the action, d^4x represents the four-dimensional volume element, g is the determinant of the metric tensor, R is the Ricci scalar, G is the gravitational constant, $\mathcal{L}_{\text{matter}}$ is the Lagrangian density for matter and radiation.

The principle of least action in cosmology provides a powerful framework for understanding the dynamics of the universe, from the large-scale structure to the behavior of individual galaxies.

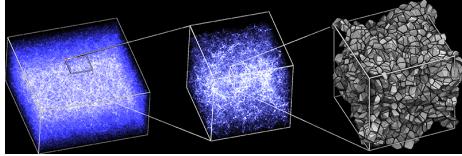


Figure 3: Model of the universe using optimal transport methods. Taken from iap.fr

2 Baryonic Acoustic Oscillations

Baryonic Acoustic Oscillations (BAO) are a key feature of the large-scale structure of the universe. They are the result of sound waves that propagated through the hot, dense plasma of the early universe. These sound waves were generated by the gravitational interaction between dark matter and baryonic matter, leading to oscillations in the density of the baryonic matter.

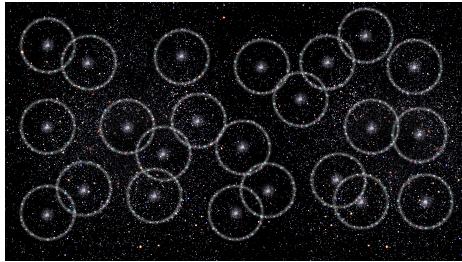


Figure 4: Baryonic acoustic oscillations, taken from bigthink.com

In the early universe - before recombination - photons, electrons, and baryons were tightly coupled in a hot plasma. Density fluctuations in this plasma created pressure gradients, leading to the propagation of sound waves. These sound waves caused periodic compressions and rarefactions in the baryon-photon fluid.

At recombination, approximately 380,000 years after the Big Bang, the universe cooled enough for electrons to combine with protons to form neutral hydrogen atoms. This decoupling of photons and baryons caused the sound waves to freeze in place, leaving an imprint on the distribution of matter in the form of a characteristic peak in the correlation function at a specific scale, known as the sound horizon scale.

The sound horizon scale, which is the distance the sound waves traveled in the early universe, serves as a standard ruler for cosmological distance measurements. By observing the large-scale structure of the universe and identifying the BAO feature, astronomers can measure the expansion history of the universe and constrain cosmological parameters, such as the dark energy equation of state and the curvature of the universe.

BAO measurements are typically made using large galaxy surveys, which

map the three-dimensional distribution of galaxies in the universe. The characteristic scale of the BAO can be detected in the two-point correlation function of galaxies or in the power spectrum of the galaxy distribution.

The precise measurement of BAO has become a crucial tool in modern cosmology. It provides independent confirmation of the cosmological model derived from observations of the cosmic microwave background (CMB) and supernovae. BAO measurements have played a significant role in establishing the Λ CDM (Lambda Cold Dark Matter) model as the standard model of cosmology, which describes a universe dominated by dark energy and dark matter.

3 The Problem

We're looking at a problem where we want to figure out how the initial conditions of self-gravitating matter (like galaxies or galaxy clusters) have changed over time. This matter is governed by the momentum equation (or Euler equation), continuity equation, and the Poisson equation.

The momentum, or Euler, equation describes how the velocity of the matter changes over time due to gravity and the expansion of the universe. The gravity part tries to pull matter together, while the expansion part (called the Hubble-Lemaître drag) tries to slow down this collapsing effect. The continuity equation ensures that mass is conserved, meaning that the total amount of matter doesn't change over time. Finally, the Poisson equation governs the gravitational potential, which is related to how gravity acts on the matter.

The equations are described below:

The Poisson equation:

$$\nabla^2 \Phi = 4\pi G\rho$$

This equation relates the gravitational potential Φ to the mass density ρ , where G is the gravitational constant.

The Euler equation:

$$\frac{\partial \mathbf{v}}{\partial t} + (\mathbf{v} \cdot \nabla) \mathbf{v} = -\frac{\nabla P}{\rho} - \nabla \Phi + \frac{2}{3} H \mathbf{v}$$

This equation describes the change in velocity \mathbf{v} of the matter over time, accounting for pressure gradients (∇P), gravitational potential ($\nabla \Phi$), and the Hubble-Lemaître drag ($\frac{2}{3} H \mathbf{v}$), where H is the Hubble parameter.

The Continuity equation:

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{v}) = 0$$

This equation ensures mass conservation, stating that the change in mass density ρ over time is balanced by the divergence of the mass flux ($\rho \mathbf{v}$).

We use a special time variable, τ , which goes from 0 (the initial condition) to 1 (the present time). The goal is to reconstruct the initial fluctuations in the density of matter from its present distribution, even though this problem doesn't

have a unique solution. However, by making some reasonable assumptions, we can turn it into a well-defined optimization problem.

We consider the problem of reconstructing the fluctuations in the initial condition of self-gravitating matter governed by the following equations, in Eulerian form:

$$\frac{\partial v}{\partial \tau} + (v \cdot \nabla_x) v = -\frac{3}{2\tau} (\nabla_x \phi + v) \quad (1)$$

$$\frac{\partial \rho}{\partial \tau} + \nabla_x \cdot (\rho v) = 0 \quad (2)$$

$$\nabla_x^2 \phi = \frac{\rho - 1}{\tau} \quad (3)$$

where ρ is the density field defined over $V \subset \mathbb{R}^3$, ϕ is the gravitational potential, τ is the growth rate of structures used as a time variable, and x denotes the co-moving coordinates. The peculiar velocity field v is also expressed as a function of the co-moving coordinates x .

Equation (1) is the momentum (Euler) equation, Equation (2) enforces mass conservation (continuity equation), and Equation (3) is the Poisson equation that governs the gravitational potential.

Given the density field ρ_F at time $\tau_F = 1$ that corresponds to the present distribution of mass, our goal is to reconstruct the initial fluctuations of $\rho(\cdot, \tau_I + \epsilon)$ for a small ϵ . The problem is under-determined, but under some reasonable simplifying assumptions, it can be replaced by a well-posed convex optimization problem.

4 Lagrangian Perturbation Theory

At the initial condition $\tau_I = 0$, for the right-hand side of the Poisson equation (3) to be defined, the density needs to be uniform, $\rho_I(\cdot) = \rho(\cdot, \tau_I) = 1$. The same consideration for the right-hand side of the momentum equation (1) implies that at the initial condition, the velocity coincides with the (negated) gradient of the potential $v_I(\cdot) = v(\cdot, \tau_I) = -\nabla_x \phi(\cdot, \tau_I)$. This condition, which also equivalently arises as a solution to the linearized set of equations (1)-(3), is sometimes referred to as slaving.

Consider now the Lagrangian point of view, and focus on the mass element that is at position q at time τ_I . Denote its trajectory by $x(q, \tau)$. Its initial speed at time τ_I is given by $v_I(q) = -\nabla_q \phi(q, \tau_I) = -\nabla_q \phi_I(q)$. In 1D, one can prove that the speed of the mass element remains constant at any time τ . In other words, integrating (1), (2), (3) in 1D results in a uniform rectilinear motion for all mass elements:

$$D_\tau x(q, \tau) = \frac{1}{\tau_F} (x_F(q) - q) = -\nabla_q \phi_I(q) \quad \forall \tau \quad (5)$$

$$x(q, \tau) = (1 - \frac{\tau}{\tau_F})q + \frac{\tau}{\tau_F} x_F(q) \quad (6)$$

where $x_F(q) = x(q, \tau_F)$ and where D_τ denotes the Lagrangian derivative with respect to time τ . It also means that in 1D, to determine the entire motion, one only needs to know the initial potential $\phi_I(\cdot) = \phi(\cdot, \tau_I)$ at time $\tau_I = 0$.

In 3D, for a small time τ , the speed of a mass element is still given by (4), but it is no longer strictly the case at any time. However, it is considered to be a reasonable approximation. This approximation means that the right-hand side of the momentum equation (1) vanishes. Physically, it means that the Hubble-Lemaître drag exactly counterbalances the effect of gravity, implying that each mass element has a uniform rectilinear motion (5). In this setting, to reconstruct the full dynamics, one just needs to determine the $\mathbf{V} \rightarrow \mathbf{V}$ map $q \mapsto x_F(q)$. This map is in turn completely determined by the potential $\phi_I(\cdot) = \phi(\cdot, 0)$, using the relation

$$x_F(q) = q + \tau_F v_I(q) = q - \tau_F \nabla_q \phi_I(q). \quad (7)$$

At this point, given the potential ϕ_I at the initial condition (we will see how to compute it in Section 4), we can reconstruct the Zel'dovich approximation. This gives us for the mass particle that was located at q at time τ_I its position $x_F(q) = q - \tau_F \nabla_q \phi_I(q)$ at the present time $\tau_F = 1$. In other words, this gives us the assignment between the initial condition and the present distribution of mass from which we can obtain the particle positions at arbitrary times (τ), up to the first-order Lagrangian perturbation theory, as

$$x(\tau, q) = q + \frac{\tau}{\tau_F} (x_F(q) - q). \quad (8)$$

Although the assignment between q and x_F is valid for as long as the convexity holds, we limit ourselves to the first order only for obtaining the particle positions at intermediate times. The main reason is that here we test our method with the goodness of reconstruction of BAO. It happens that often one adds additional, broad-band, fitting terms to the power spectrum which takes care of the mode-coupling as well as other effects such as the shot noise. The implementation of the second and higher-order Lagrangian perturbation theory into our algorithm shall be reported in forthcoming works.

5 Least Action Principle and Optimal Transport

One can also obtain the momentum equation (1) by extremizing the following action integral (Brenier et al. (2003) Appendix D):

$$I = \frac{1}{2} \int_{\tau_I}^{\tau_F} \int_V \left(\rho |v|^2 + \frac{3}{2} |\nabla_x \phi|^2 \right) \tau^{3/2} d^3 x d\tau \quad (9)$$

subject to mass conservation (2), to the Poisson equation for the potential (3), and to the boundary conditions:

$$\rho(\cdot, \tau_I) = \rho_I(\cdot) = 1; \quad \rho(\cdot, \tau_F) = \rho_F(\cdot) \quad (10)$$

where ρ_I denotes the (uniform) density map at the initial condition and ρ_F denotes the density map at present time $\tau = \tau_F$. Using the method of Lagrange multipliers and varying ρ , one obtains the momentum equation (1).

Consider now an approximation, where the second term of the integrand and the $\tau^{3/2}$ factor are removed (Giavalisco et al. 1993), which may be thought of as replacing the $\frac{3}{2}$ coefficient by $\frac{3\alpha}{2}$ and making α tend to zero. The action integral (7) then becomes:

$$I = \frac{1}{2} \int_{\tau_I}^{\tau_F} \int_V \rho |v|^2 d^3x d\tau. \quad (11)$$

Note that the integrand has only the kinetic energy, and no longer any potential energy. This again corresponds to the Zel'dovich approximation. Given the boundary conditions (8), now we want to find the motion that minimizes the kinetic energy. If we have a single mass particle, it is easy to see that minimizing the action results in a uniform rectilinear motion (Landau & Lifshitz 1975). It can be proved (Benamou & Brenier 2000) that this is still the case for any number of particles, or even for a continuous density field ρ : extremizing the action (9) means that all mass particles (or all elementary mass elements for a continuous ρ) move in a uniform rectilinear manner. In other words, finding the motion $x(q, \tau)$ that minimizes the action I on $V \times [\tau_I, \tau_F]$ is equivalent to finding the map $x_F : V \rightarrow V$ that gives the position at present time $x_F(q) = x(q, \tau_F)$ of the mass element that was at position q at the initial condition $\tau_I = 0$. The map x_F minimizes the following functional:

$$\inf_{x_F} \int_V \rho(q) |q - x_F(q)|^2 d^3q \quad (12)$$

subject to mass conservation (2) and to the boundary conditions (8). Now, it may be more natural to write mass conservation in Lagrangian coordinates. Using $\rho(x(q, \tau), \tau) = \rho(q) / (\det \nabla_q x)$, the mass conservation constraint writes:

$$\rho_F(x_F(q)) \det(\nabla_q x_F(q)) = \rho_I(q) \quad \forall q \quad (13)$$

The minimization of expression (10) subject to (11) is referred to as Monge's optimal transport problem (Monge).

6 The Convex Kantorovich Potential and the Monge-Ampère Equation

Consider Monge's optimal transport problem:

$$\sup_{x_F} \inf_{\Psi} \left\{ J(x_F) = \int_V \rho_I(q) x_F(q) \cdot q d^3q - \int_V \Psi(x_F(q)) \rho(q) d^3q + \int_V \Psi(x) \rho_F(x) d^3x \right\}. \quad (14)$$

The first-order optimality condition with respect to x_F is:

$$\frac{\partial J}{\partial x_F} = 0 \Rightarrow \rho_I(q)q = \rho_I(q)\nabla\Psi(x_F(q)) \Rightarrow q = \nabla_x\Psi(x_F(q)). \quad (15)$$

The second-order optimality condition is:

$$\frac{\partial^2 J}{\partial x_F^2} \geq 0 \Rightarrow D^2\Psi \geq 0 \Rightarrow \Psi \text{ is a convex function.} \quad (16)$$

The function $x \mapsto q$ that maps a mass element x at present time τ_F back to its initial position q at time τ_I is the gradient of a convex potential Ψ (called the Kantorovich potential).

The forward map $q \mapsto x_F$ can be expressed as:

$$x_F(q) = \nabla_q\Phi(q); \Phi \text{ is a convex function.} \quad (17)$$

Φ and Ψ are related by the Legendre-Fenchel transform:

$$\forall q, \Phi(q) = \Psi^*(q) \text{ where } \Psi^*(q) = \sup_x[x \cdot q - \Psi(x)]. \quad (18)$$

The relation between the gravitational potential ϕ_I at time τ_I and the convex Kantorovich potential Φ associated with the map $q \mapsto x_F(q)$ is:

$$x_F(q) = q - \nabla_q\phi_I(q) = \nabla_q\Phi(q) \Rightarrow \Phi(q) = \frac{1}{2}q^2 - \phi_I(q). \quad (19)$$

The insertion of $x_F(q) = \nabla_q\Phi(q)$ into the mass conservation constraint yields the Monge-Ampère equation:

$$\rho_F(\nabla_x\Phi(q))D^2\Phi(q) = \rho_I(q). \quad (20)$$

The convexity of the Kantorovich potential implies that there is no multistreaming in the reconstructed dynamics.

To summarize, given the density at present time ρ_F and the density at the initial condition $\rho_I = 1$, our goal is to find the assignment map $q \mapsto x_F(q)$ that determines the assignment between the points q at the initial condition and the points x at the current time. It has the following properties:

- $x_F(\cdot)$ is the minimizer of $\inf_{x_F} \left\{ \int_V |q - x_F(q)|^2 \rho_I(q) d^2q \right\}$ subject to mass conservation;
- $x_F(\cdot)$ is also the gradient of the (convex) Kantorovich potential Φ ;
- Φ is the solution of the Monge-Ampère equation;
- The convexity of Φ implies that there is no multistreaming in the reconstructed dynamics;
- Φ is related to the gravitational potential at the initial condition ϕ_I by: $\Phi(q) = \frac{1}{2}q^2 - \phi_I(q)$.

From the assignment map $x_F(\cdot)$, it is (optionally) possible to reconstruct higher-order dynamics using Lagrangian perturbation theory and at first order using the expression we have given earlier.

7 Solving the Assignment Problem

In this section, we describe numerical solution mechanisms to compute the assignment map $x_F(\cdot)$. We first review the existing methods, which are based on a discretization of the density ρ_I at the initial condition and a discretization of the density ρ_F at the current time (Sections 3.1, 3.2). Then we present our method (based on semi-discrete optimal transport), that uses a continuous representation of the initial density ρ_I and a discrete representation of the density ρ_F , hence a semi-discrete method (Section 3.3).

7.1 Discrete-discrete MAK Reconstruction

We consider that the density at the initial condition ρ_I and the density at present time ρ_F are both represented in discrete form, by a set of N particles, with each particle i having a mass $\mu_i = 1/N$:

- At the initial condition τ_I , the mass distribution ρ_I is represented by a set of N points q_i , $i = 1 \dots N$. Since the initial distribution of mass is uniform at τ_I , the points q_i are organized on a regular grid;
- At present time τ_F , the distribution of mass ρ_F is represented by a set of (the same number N) of points x_j , $j = 1 \dots N$.

In this setting, the initial problem of finding the (continuous) map $x_F(\cdot) : V \rightarrow V$ is replaced with finding which point x_j corresponds to each point q_i . The discrete version of Monge's problem writes:

$$\inf_{\pi} \sum_j |q_j - x_{\pi(j)}|^2 \quad (21)$$

where $\pi : [1 \dots N] \rightarrow [1 \dots N]$ is a permutation of the indices.

7.2 MAK Duality

We now exhibit more structure in the discrete Monge problem and its relation with the gravitational potential ϕ_I , that we will use to design a more efficient algorithm. Instead of searching for the (combinatorial) assignment $i \mapsto j = \pi(i)$, we consider now the following optimization problem, introduced by Kantorovich (1942):

$$\inf_{\gamma} \sum_i \sum_j \gamma_{ij} |x_i - q_j|^2 \quad \text{subject to} \quad (22)$$

$$\sum_i \gamma_{ij} = \mu_i \quad \forall j, \quad (23)$$

$$\sum_j \gamma_{ij} = \mu_j \quad \forall i, \quad (24)$$

$$\gamma_{ij} \geq 0 \quad \forall i, j. \quad (25)$$

The objective function depends on an array of coefficients γ_{ij} . Intuitively, each coefficient γ_{ij} indicates how much matter goes from q_j to x_i . In this setting, matter can split and merge between different particles, for instance, a particle q_j can send half of its matter to particle x_k and the other half to particle x_l (using $\gamma_{jk} = \gamma_{jl} = 0.5$). Clearly, the mass of all the matter that gathers at a particle x_i should sum as the mass μ_i of the particle, and the mass of all the matter originated from a particle q_j should sum as the mass μ_j of the particle. Since no matter disappears, all coefficients γ_{ij} should be positive. An array of coefficients γ_{ij} that satisfies the three constraints is called a transport plan (and an optimal transport plan if it minimizes the objective function).

At first sight, it may seem to be a rather convoluted reformulation of Monge's problem, in particular, we now need to find N^2 unknowns, to be compared with the $N \rightarrow N$ permutation we had to find initially. However, it can be observed that the problem is a linear optimization problem with linear constraints. Introduce $\psi \in \mathbb{R}^N$ and $\phi \in \mathbb{R}^N$ the Lagrange multipliers associated with constraints respectively. The dual of the optimization problem writes:

$$\sup_{\psi, \phi} \left\{ \sum_i \psi_i \mu_i + \sum_j \phi_j \mu_j \right\} \quad \text{subject to} \quad \psi_i + \phi_j \leq \frac{1}{2} |x_i - q_j|^2 \quad \forall i, j. \quad (26)$$

In addition, given a pair ψ, ϕ that satisfies the constraint, it is easy to check that replacing ϕ with ψ^c still satisfies the constraints while always increasing the objective function, where ψ^c is defined by:

$$\psi_j^c = \inf_i \left\{ \frac{1}{2} |x_i - q_j|^2 - \psi_i \right\}. \quad (27)$$

There exist several methods that exploit the structure of the problem and its dual, we refer the reader to the literature for a survey. Among these methods, we mention the entropic regularized method, that solves:

$$\inf_{\gamma} \left\{ \sum_i \sum_j \gamma_{ij} |x_i - q_j|^2 + \varepsilon \sum_i \sum_j \gamma_{ij} \log(\gamma_{ij}) \right\} \quad (28)$$

where ε is a (small) regularization parameter. If both the q_j 's and x_i 's are supported by regular grids, it is possible to exploit the structure of the problem to design a fast and efficient algorithm. The advantage of this algorithm is its speed and simplicity. The drawbacks are the need for re-sampling everything on regular grids and the difficulty of tuning the parameter ε (a too large value of ε results in a blurry, imprecise transport plan, and a too small value of ε makes the algorithm slow to converge). Various ways of leveraging its speed while overcoming its drawbacks are currently being studied.

In the next subsection we describe a different method and although the algorithm that we eventually obtain is more complicated than those based on entropic regularized schemes, it does not depend on a regularization parameter ε , and does not require ρ_F to be re-sampled on a regular grid.

7.3 Semi-discrete MAK Reconstruction

We consider the discrete assignment problem expressed by equation (21). The initial density ρ_I is represented by a set of N particles $\{q_j\}_{j=1}^N$ located on a regular grid, depicted in blue in Figure 1-A. The current mass distribution ρ_F consists of N particles $\{x_i\}_{i=1}^N$, shown in red in the figure. To enhance precision, we can utilize a finer grid for the q_j particles. Although this increases computational time, as the number of q_j particles approaches infinity and their mass tends to zero, the initial density ρ_I becomes uniform, while ρ_F remains concentrated at the points x_i .

This scenario mirrors our early universe reconstruction problem, where the initial density ρ_I is uniform, and the present-time density ρ_F corresponds to galaxies represented by single points x_i . In this context, it can be demonstrated that each point x_i is coupled to a polyhedral region, computable through an efficient algorithm, resulting in a more precise outcome compared to the combinatorial approach.

Delving into the details, we explore the dual problem, leading to the optimization problem dependent on a vector of N values $\{\psi_i\}$ and a function φ . This problem can be simplified to depend solely on the vector ψ . Partitioning the space V into N regions $\{V_{\psi_i}\}$ according to the index i that minimizes the integrand, we reformulate the integral and impose a constraint ensuring the existence of at least one pair φ, ψ satisfying the conditions.

The resulting partition, known as a Laguerre diagram, is a generalization of the Voronoi diagram, parameterized by the vector ψ . The solution to the optimization problem yields a vector $\{\psi_i\}$, from which we can deduce the gravitational potential φ_I . This potential maps each point q in a Laguerre cell V_{ψ_i} to a corresponding point x_i at the current time.

In summary, solving the optimization problem provides a set of coefficients $\{\psi_i\}$ defining a partition of V into Laguerre cells. Each cell corresponds to the set of initial points that collapses into a specific point x_i at the present time.

8 Numerical Solution

Let us denote by $K(\psi) : \mathbb{R}^N \rightarrow \mathbb{R}$ the objective function of the optimization problem (32):

$$K(\psi) = \sum_i \left(\int_V \frac{1}{2} |\mathbf{x}_i - \mathbf{q}|^2 \psi_i - \psi_i \rho I(\mathbf{q}) d^3 q \right) + \sum_i \psi_i \mu_i \quad (36)$$

It can be shown that $K(\cdot)$ is a concave C^2 function, which suggests that it can be efficiently maximized by a Newton algorithm. The extensive algorithmic details of this Newton algorithm are given in Appendix A. Briefly, the algorithm iteratively maximizes second-order approximations of $K(\cdot)$. A 2-dimensional example of the Laguerre diagrams corresponding to each iteration is shown in Figure 3. The algorithm starts with $\psi = 0$, then updates ψ by solving a series of linear systems. Ultimately, the algorithm finds the unique solution, and

all the Laguerre cells have the prescribed volumes. This numerical algorithm outperforms the previous combinatorial ones by fully exploiting the variational nature of the problem.

The computation time scales as $O(N \log N)$, massively outperforming previous approaches. To provide a realistic setting within which we aim to use this algorithm, we employ snapshots from the cosmological N -body simulation suite AbacusCosmos. The convergence time of the reconstruction slightly depends on the degree to which non-linear clustering has occurred in the samples. Therefore, we run the complexity analysis on snapshots of different redshifts, keeping the particle number N consistent across redshifts, and find that the computation time increases with decreasing redshift.

9 References

Levy, B., Mohayaee, R., & von Hausegger, S. A fast semi-discrete optimal transport algorithm for a unique reconstruction of the early universe.

<https://www.iap.fr/actualites/laune/2022/TransportOptimal/OptimalTransport-en.html>

<https://lgarrison.github.io/AbacusCosmos/>