A Numerical Scheme for Approximating the Support and the Value of an Optimal Solution to the Mass Transfer Problem via Wavelets and Multiresolution Analysis

Miguel Angel Larruz-Medina , José Rigoberto Gabriel-Argüelles * , Eloisa Benitez-Mariño , Alberto Santamaria-Pang

Posted Date: 18 March 2026

doi: 10.20944/preprints202603.1455.v1

Keywords: optimal transport; wavelet theory; multiresolution analysis; computational efficiency; numerical optimization

 doi:10.20944/preprints202603.1455.v1

Article

A Numerical Scheme for Approximating the Support and the Value of an Optimal Solution to the Mass Transfer Problem via Wavelets and Multiresolution Analysis

Miguel Angel Larruz-Medina [1], José Rigoberto Gabriel-Argüelles [1,*], Eloisa Benitez-Mariño [1] and Alberto Santamaria-Pang [4,5]

[1] Universidad Veracruzana, México
[2] Microsoft Healthcare AI, Redmond, WA 98052, USA
[3] Johns Hopkins School of Medicine, Baltimore, MD 21205, USA
* Correspondence: jgabriel@uv.mx

Abstract

In this paper, we present a scheme for approximating the support and optimum value of an optimal measure for the Monge–Kantorovich (MK) mass transfer problem. Obtaining exact solutions to the MK problem is difficult; such solutions are only found in a few specific cases. Using an algorithm to approximate the optimum value is computationally expensive, particularly in high-dimensional or large-scale scenarios. To address this challenge, we developed an innovative method that integrates wavelet theory and multiresolution analysis. This method uses wavelet-based techniques to approximate the support of an optimal measure, further reducing the number of variables in the linear programs and thus decreasing the dimensionality and computational complexity of each step of the scheme. This method generates a sequence of transport problems with optimal measures of finite support. We then demonstrate that the optimum values of the transport problems converge to the optimum value of the MK problem and that the supports of the finite optimal measures converge to the support of an optimal measure for the MK problem. We present some numerical experiments to demonstrate the efficiency of the scheme. We observe that the method has potential applications in various fields, such as image processing, economics, resource allocation, and machine learning, where finding efficient solutions to large-scale optimal transportation problems is essential.

Keywords: optimal transport; wavelet theory; multiresolution analysis; computational efficiency; numerical optimization

1. Introduction

The Monge–Kantorovich (MK) mass transfer problem is an optimization problem on probability measure spaces with applications in several fields of mathematics, for example: functional analysis, differential geometry, statistics, economics, and dynamical systems Rachev [1], Villani [2]. Obtaining computational solutions to the MK problem is fundamental. However, approximating a continuous problem using discrete processes requires a large number of computational operations, especially in high-dimensional or large-scale scenarios, which is very costly. Therefore, computationally efficient approximation schemes are crucial for practical implementations [3].

In this paper, we propose a numerical approximation scheme inspired by the work of Gabriel [4], where the MK problem is studied in compact metric spaces and the optimal value is approximated by a sequence of the MK problems (transportation problems), where the MK problem is studied in compact metric spaces and the optimal value is approximated by a sequence of MK problems (transportation problems), where their marginal measures are probability measures with finite support.

 doi:10.20944/preprints202603.1455.v1

The transportation problems are solved to identify an optimal measure with finite support. This approach is refined by wavelet theory and multiresolution analysis, which significantly reduces the number of variables compared to traditional linear programming methods.

Recent research on algorithms for solving the MK problem includes Lee [5], which studies an algorithm to solve the Monge–Ampère equation using gradient flow, an elliptic formulation and a Hamilton–Jacobi equation. In Chi [6], stochastic optimization with a stable optimization process is used to approximate the MK problem with an algorithm termed Cop-OT. In Tanguy [7], numerical methods are studied for L-Lipschitz gradients of l-strongly convex potentials and the convergence of stochastic gradient descent methods for neural networks, where optimal transport is applied. In Acosta [8], are used wavelet methods as in this work, but its approach is based on apply the wavelet transform in the cost function.

Our wavelet-based multiresolution approximation method not only improves computational efficiency but also ensures that the supports of a sequence of the measures converge to the support of the optimal measure of the MK problem. Furthermore, we provide a rigorous mathematical proof and a theoretical framework underpinning this optimization approach. Numerical experiments clearly highlight the significant improvements in computational complexity and accuracy, making our approach broadly applicable across fields demanding efficient solutions to optimal transport problems, such as large-scale training of artificial neural networks via backpropagation.

The proposed algorithm can be applied to problems that require both an optimal measure and the corresponding value of the Monge–Kantorovich (MK) problem because it efficiently approximates, the optimal transport plan while preserving high accuracy in estimating the transport cost. By leveraging wavelet-based multiresolution analysis, the method adaptively refines the computational domain to focus on the regions that contribute most to the optimal coupling, thus reducing complexity without losing essential structural information. This capability makes it suitable for applications where the MK solution provides meaningful physical, geometric, or statistical correspondences between distributions. For example, in Bonne [9], the optimal transport is used for tasks such as image processing, geometric processing, rendering, fluid simulation, and computational optics, all of which rely on precise transport measures. In Kamsu [10], recent advances in optimal transport for data science and machine learning are reviewed, highlighting the importance of accurate and scalable solvers. Likewise, Chen [11] applies optimal transport to model fluid flows in the brain, demonstrating its relevance in biomedical imaging, while Kreinovich [12] discusses applications in economics, where the computation of optimal transport values guides efficient resource allocation and market equilibrium analysis.

2. Wavelet Theory

A wavelet function ψ is a complex-valued, square-integrable function, $\psi \in L^2(\mathbf{R})$. Wavelets provide a versatile mathematical tool for analyzing signals at multiple resolutions and scales. Formally, a wavelet function $\psi \in L^2(\mathbf{R})$ satisfies admissibility conditions (see Mallat [13]).

Given a function $f \in L^2(\mathbf{R})$, its continuous wavelet transform (CWT) at position u and scale $s > 0$ is given by:

$$W_f(u,s) = \left\langle f(t),\, s^{-1/2}\,\psi\left(\frac{t-u}{s}\right)\right\rangle.$$

The original function f can then be reconstructed from its wavelet coefficients:

$$f(t) = \frac{1}{C_\psi}\int_0^\infty \int_{-\infty}^\infty W_f(u,s)\, s^{-1/2}\psi\left(\frac{t-u}{s}\right)\frac{du\,ds}{s^2}.$$

Multiresolution analysis (MRA) is an efficient computational framework and it uses nested subspaces $\{V_j\}_{j\in\mathbf{Z}} \subset L^2(\mathbf{R})$ defined by scaling functions $\phi(t)$, where $\mathbf{Z}$ is the set of integers. Moreover the sequence $\{V_j\}_{j\in\mathbf{Z}}$ form a nested sequence of closed subspaces and satisfy:

- *Nested structure:* $\cdots \subset V_j \subset V_{j+1} \subset \cdots$
- *Completeness:* $\overline{\bigcup_{j\in\mathbf{Z}} V_j} = L^2(\mathbf{R})$

 doi:10.20944/preprints202603.1455.v1

- *Trivial intersection:* $\bigcap_{j\in\mathbf{Z}} V_j = \{0\}$
- *Scaling property:* $f(t) \in V_j \Leftrightarrow f(2^{-j}t) \in V_0$
- *Translation invariance:* $\phi(t-n) \in V_0,\ \forall n \in \mathbf{Z}$

Each subspace V_{j+1} decomposes orthogonally as $V_{j+1} = V_j \oplus W_j$, where W_j could be spanned by wavelet functions $\psi_{j,n}(t) = 2^{j/2}\psi(2^j t - n)$ (see Hardle [14]). Hence, any function $f \in L^2(\mathbf{R})$ can be represented as:

$$f(t) = \sum_{k\in\mathbf{Z}} \alpha_{j_0,k}\phi_{j_0,k}(t) + \sum_{j\geq j_0}\sum_{k\in\mathbf{Z}} \beta_{j,k}\psi_{j,k}(t),$$

with coefficients $\alpha_{j,k} = \langle f, \phi_{j,k}\rangle$ and $\beta_{j,k} = \langle f, \psi_{j,k}\rangle$.

This approach naturally extends to higher-dimensional spaces $L^2(\mathbf{R}^p)$, with multidimensional wavelets constructed via tensor products of 1D wavelets. In 2D, wavelets are defined in Mallat [13] as:

$$\psi^{(\epsilon)}(x,y) = \psi^{\epsilon_1}(x)\psi^{\epsilon_2}(y), \quad \epsilon = (\epsilon_1,\epsilon_2),\ \epsilon_i \in \{0,1\},$$

with $\psi^0 = \phi$ (scaling function) and $\psi^1 = \psi$ (wavelet function). Our experiments address 2D cases, though the method generalizes efficiently to higher dimensions.

3. Mass Transfer Problem

To define the mass transfer problem, we need the following:

a) Two compact metric spaces X and Y, endowed with the Borel σ-algebras $\mathbf{B}(X)$ and $\mathbf{B}(Y)$, respectively.
b) A continuous cost function $c : X \times Y \to \mathbf{R}$.
c) A probability measure ν_1 on $\mathbf{B}(X)$ and a probability measure ν_2 on $\mathbf{B}(Y)$.

We denote $\mathbf{M}(X\times Y)$ as the linear space of all finite signed measures on $\mathbf{B}(X\times Y)$ and assume these spaces are endowed with the weak convergence topology. If $\mu \in \mathbf{M}(X\times Y)$, then the marginals (or projections) of μ are denoted by $\Pi_1\mu$ and $\Pi_2\mu$, which are signed finite measures on $\mathbf{B}(X)$ and $\mathbf{B}(Y)$, respectively, defined for $A \in \mathbf{B}(X)$ and $B \in \mathbf{B}(Y)$ by:

$$\Pi_1\mu(A) = \mu(A\times Y) \quad \text{and} \quad \Pi_2\mu(B) = \mu(X\times B).$$

The Monge–Kantorovich mass transfer problem is:

$$\text{MK}: \quad \min_{\mu\geq 0} \int_{X\times Y} c\, d\mu \tag{1}$$

$$\text{s.t.} \quad \Pi_1\mu = \nu_1, \quad \Pi_2\mu = \nu_2, \quad \mu \in \mathbf{M}(X\times Y). \tag{2}$$

A measure $\mu \in \mathbf{M}_+(X\times Y)$ is called feasible if it satisfies the constraints in (2) and if $\left|\int_{X\times Y} c\, d\mu\right| < \infty$.

In Gabriel [4], a method is developed to approximate the value of the Monge–Kantorovich problem in compact metric spaces. indeed, since X and Y are compact metric spaces, there exist countable dense sets $X_\infty = \{x_1, x_2, \ldots\} \subset X$ and $Y_\infty = \{y_1, y_2, \ldots\} \subset Y$. Now, take a sequence of real numbers $r_n \downarrow 0$. It is possible to construct disjoint sets around points x_k and y_j, each with diameter less than or equal to $2r_n$, such that $X = \bigcup_{k=1}^{s_n} A_k^n$ and $Y = \bigcup_{j=1}^{t_n} D_j^n$, with $x_k \in A_k^n$ and $y_j \in D_j^n$.

Define:

$$M_n = \{x_1, x_2, \ldots, x_{s_n}\} \times \{y_1, y_2, \ldots, y_{t_n}\} = X_n \times Y_n,$$

and call a refinement of M_n the set

$$M_{n+1} = \{x_1, x_2, \ldots, x_{s_{n+1}}\} \times \{y_1, y_2, \ldots, y_{t_{n+1}}\} = X_{n+1} \times Y_{n+1}.$$

 doi:10.20944/preprints202603.1455.v1

Then, we build the following sequence of linear programming problems (transportation problems):

$$\begin{aligned} \mathrm{MK}_n: \quad \min \quad & \sum_{k=1}^{s_n}\sum_{j=1}^{t_n} c(x_k, y_j)\lambda_{k,j}^n \\ \text{s.t.} \quad & \sum_{j=1}^{t_n} \lambda_{k,j}^n = a_k^n, \quad 1 \le k \le s_n \\ & \sum_{k=1}^{s_n} \lambda_{k,j}^n = b_j^n, \quad 1 \le j \le t_n \\ & \lambda_{k,j}^n \ge 0 \quad \forall 1 \le k \le s_n, 1 \le j \le t_n \end{aligned}$$

with $a_k^n = \nu_1(A_k^n)$ and $b_j^n = \nu_2(D_j^n)$. It has been proved that with its solutions, we can define a sequence of measures:

$$\mu_n^*(E) = \sum_{k=1}^{s_n}\sum_{j=1}^{t_n} \lambda_{k,j}^n \delta_{(x_k,y_j)}(E),$$

where $\delta_{x_k,y_j}(E) = 1$ if $(x_k, y_j) \in E$ and $= 0$ otherwise, that converge weakly to the optimal solution μ^* of the MK problem.

4. Wavelet-Based Mass Transfer Algorithm

We propose a five-step iterative algorithm to solve the Monge–Kantorovich (MK) optimal transport problem using a wavelet-based multiresolution strategy that achieves computational efficiency without sacrificing accuracy. The method is designed for 2D domains but generalizes to higher-dimensional settings. Also, the method is defined for $X = Y = [0, 1]$, but it can be used in any compact sets of $\mathbb{R}$ finding an isomorphism.

To ensure the convergence of the algorithm, it is important to set some assumptions. The first one is that we must ensure that the optimal μ has no atoms, which means, $\mu(\{z\}) = 0\ \forall z \in X \times Y$. The second one is that the optimal plan μ must be concentrated along the graph of a Lipschitz (or piecewise smooth) function.

Step 1: Discretization. Define ordered dyadic grids $X_n = Y_n = \left\{ \frac{i}{2^n} \,\middle|\, 0 \le i \le 2^n \right\}$ and solve the relaxed MK problem:

$$\mathrm{MK}_n: \min_{\lambda_{i,j}^n \ge 0} \sum_{(x_i^n, y_j^n)} c(x_i^n, y_j^n)\lambda_{i,j}^n,$$

subject to marginal constraints that ensure consistency with prescribed measures. In this step, we consider $A_k^n = (x_{k-1}^n, x_k^n]$ and $D_j^n = (y_{j-1}^n, y_j^n]$.It is important to use the dyadic grids, because of the uniform sampling properties of the wavelets, (see Mallat [13]).

Step 2: Support Extraction. Collect the support of the solution:

$$S_n = \left\{ (x_i^n, y_j^n) \mid \lambda_{i,j}^n > 0 \right\},$$

which defines a sparse approximation of the optimal plan.

Step 3: Wavelet Analysis. Represent the support set S_n as a piecewise-constant function $T_n(x, y)$ over the sets $A_k^n \times D_j^n$, then apply a 2D DWT (see Daubechies [15]) to compute detail coefficients:

$$\beta_{n-1,(i,j)}^3 = \langle T_n, \psi_{n-1,(i,j)}^{(3)} \rangle.$$

Wavelet coefficients above a threshold ϵ define a region of interest W_n, expanded to a neighborhood B_n,

 doi:10.20944/preprints202603.1455.v1

$$B_n = \bigcup_{\{(x_k^{n-1}, y_j^{n-1}) : \beta_{n-1,(k,j)} > \epsilon\}} A_k^{n-1} \times D_j^{n-1}$$

The detail coefficients that are above a threshold gives information of the regularity of the functions that we are approximating, because if the value of the coefficient is a local maximun of the wavelet transform function, then in this localization we can find an irregularity of the function, see Mallat [13]. Due to the uniformity assumption of the transport plan, those coefficients give us some areas of interest.

Step 4: Reduced Problem (WMK). The next resolution level defines a refined grid $X_{n+m} \times Y_{n+m}$, $m \in \mathbb{N}$, but the transport plan is solved only over $O_n = B_n \cap (X_{n+m} \times Y_{n+m})$, subject to marginal conditions:

$$\text{WMK}_{n+m} : \min_{\lambda_{i,j}^{n+m} \geq 0} \sum_{(x_i, y_j) \in O_n} c(x_i, y_j) \lambda_{i,j}^{n+m}.$$

Step 5: Iteration. Repeat Steps 2–4. Under mild assumptions, $\mu_n \to \mu^*$, and the supports $\text{supp}(\mu_n)$ converge in Hausdorff distance.

The algorithm is summarized in the Figure 1.

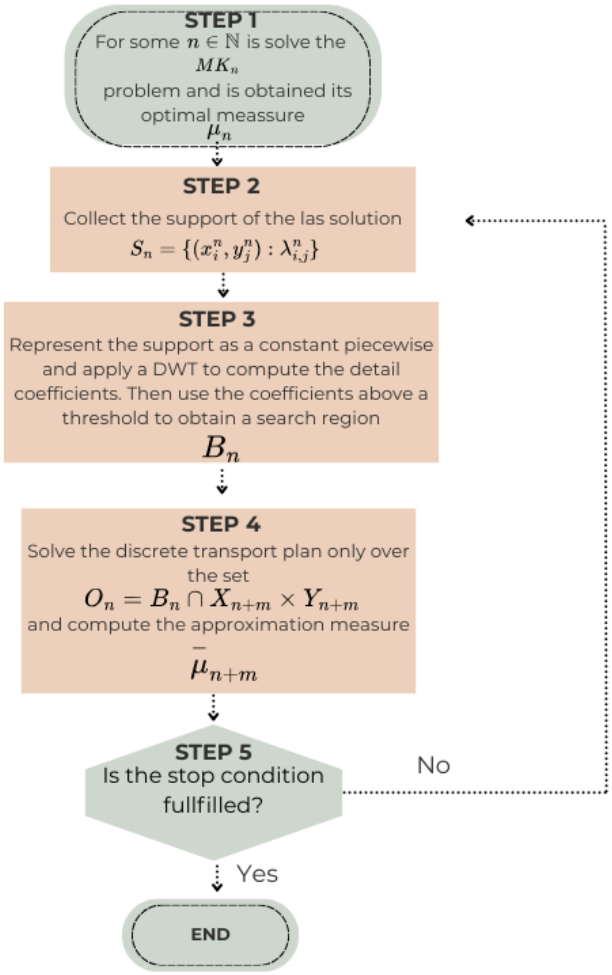

Figure 1. Diagram of the mass transport algorithm with wavelets

Convergence. To ensure the convergence we have proved a pair of technical, but intuitive, propositions. The first one tells us that when we have a sequence of measures $\{\mu_n\}$ such that converge weakly to μ,

 doi:10.20944/preprints202603.1455.v1

then the support of the μ_n's surround the support of μ, and the main atoms of the support converge in a Hausdorff distance. To prove it, we use de Theorem 6.1 in Parthasarathy [16], which establishes that for a measurable set B whose boundary has μ-zero measure, it is fulfilled $\mu_n(B) \to \mu(B)$.

Proposition 1. *Let $\{\mu_n\}$ be a sequence of measures that converge weakly to μ and B a measurable set whose boundary has μ-zero measure. Let $\{B_n\}$, with $B_1 = B$, a sequence of sets such that $B_n \downarrow supp(\mu) \cap B$. Then there exist a sequence $\{\mu_{m_k}(B_{m_k})\}$, with $m_k \uparrow \infty$, such that $\mu_{m_k}(B_{m_k}) \to \mu(B)$.*

Proof. It is built the sequence.

Let us observe that because $\mu_n(B \cap supp(\mu)) \to \mu(B)$, there exist $m_k \in \mathbb{N}$ such that,

$$|\mu(B) - \mu_{m_k}(B \cap supp(\mu)| < \frac{1}{2k}$$

If m_k is set, then $\mu_{m_k}(B_n) \to \mu_{m_k}(B \cap supp(\mu))$ when $n \to \infty$, therefore, there exist $n_k \in \mathbb{N}$ such that,

$$|\mu_{m_k}(B_{n_k}) - \mu_{m_k}(B \cap supp(\mu)| < \frac{1}{2k}$$

furthermore, we can set $n_k \geq m_k$. This procedure can be done for each $k \in \mathbb{N}$ and we can choose the index such that $m_k < m_{k+1}$ and $n_k < n_{k+1}$ for each $k \in \mathbb{N}$. Then,

$$\begin{aligned} |\mu_{m_k}(B_{n_k}) - \mu(B)| &\leq |\mu_{m_k}(B_{n_k}) - \mu_{m_k}(B \cap supp(\mu))| + |\mu_{m_k}(B \cap supp(\mu)) - \mu(B)| \\ &< \frac{1}{2k} + \frac{1}{2k} = \frac{1}{k} \end{aligned}$$

With the above, it can be ensured that $\lim_{k\to\infty} \mu_{m_k}(B_{n_k}) = \mu(B)$.

Because $n_k \geq m_k$ then $B_{n_k} \subset B_{m_k}$ for each k. Then, it can be written:

$$\mu_{m_k}(B_{m_k}) = \mu_{m_k}(B_{n_k}) + \mu_{m_k}(B_{m_k} \setminus B_{n_k})$$

So it is enough to prove that $\mu_{m_k}(B_{m_k} \setminus B_{n_k})$ tends to zero.

Let $A_k = B_1 \setminus B_{n_k}$, where it can be observed that $B_{m_k} \setminus B_{n_k} \subset A_k$. Then,

$$\begin{aligned} \lim_{k\to\infty} \mu_{m_k}(A_k) &= \lim_{k\to\infty} [\mu_{m_k}(B_1) - \mu_{m_k}(B_{n_k})] \\ &= \mu(B) - \mu(B) = 0 \end{aligned}$$

Because $0 \leq \mu_{m_k}(B_{m_k} \setminus B_{n_k}) \leq \mu_{m_k}(A_k)$, it can be concluded that $\lim_{k\to\infty} \mu_{m_k}(B_{m_k}) = \mu(B)$, as desired. □

We must observe that the optimal measures $\overline{\mu}_n^*$ of each subprogram WMK$_n$ are not necessary the optimal measure μ_n^* of the transport problem MK$_n$, but the following proposition prove that the minimal value of a subsequence WMK$_{n_k}$ converge to the optimal value of the MK problem with a suitable setup.

Proposition 2. *Let μ^* be the optimal measure of the MK problem and let $\{\mu_k^*\}$ the optimal measures of the MK$_{n_k}$ problems such that, for a sequence of sets with $B_k' \downarrow supp(\mu^*)$, it is fulfilled that $\mu_k^*(B_k') \to 1$. Let $\{\overline{\mu}_k\}$ be feasible measures for the MK$_{n_k}$ problems such that $supp(\overline{\mu}_k) \subset B_k' \cap X_{n_k} \times Y_{n_k}$. Then $\langle \overline{\mu}_k^*, c \rangle \to \langle \mu^*, c \rangle$.*

Proof. Lets prove:

$$|\langle \mu_k^*, c \rangle - \langle \overline{\mu}_k, c \rangle| \to 0$$

since $\langle \mu_k^*, c \rangle \to \langle \mu^*, c \rangle$, converge to the minimum value of the MK problem.

 doi:10.20944/preprints202603.1455.v1

Let $B_k = B'_k \cap X_{n_k} \times Y_{n_k}$, then we can write the discrete probability measures as the following sum,

$$\mu_k^* = \sum_{z \in X_{n_k} \times Y_{n_k}} \lambda_z^k \delta_z$$
$$\overline{\mu}_k = \sum_{z \in B_k} \overline{\lambda}_k^k \delta_z$$

where δ_z is the Dirac measure in z, i.e. $\delta_z(E) = 1$ if $z \in E$ and $\delta_z(E) = 0$ otherwise.

Let us observe that since μ_k^* is the optimal measure of the MK_{n_k} problem, then

$$\begin{aligned} 0 &\leq \langle \overline{\mu}_k, c \rangle - \langle \mu_k^*, c \rangle \\ &= \sum_{z \in B_k} \overline{\lambda}_z^k c(z) - \sum_{z \in X_{n_k} \times Y_{n_k}} \lambda_z^k c(z) \end{aligned}$$

which implies:

$$\langle \overline{\mu}_k, c \rangle - \langle \mu_k^*, c \rangle = \sum_{z \in B_k} \left(\overline{\lambda}_z^k - \lambda_z^k \right) c(z) - \sum_{z \in B_k^c} \lambda_z^k c(z) \tag{3}$$

where $B_z^c = X_{n_k} \times Y_{n_k} \setminus B_k$.

Since c is a continuous function in the compact $X \times Y$, then is bounded, $s = \sup_{z \in X \times Y} |c(z)| < \infty$. Hence, we have:

$$\left| \sum_{z \in B_k^c} \lambda_z^k c(z) \right| \leq s \sum_{z \in B_k^c} \lambda_z^k = s \mu_k^*(B_k^c)$$

It can be observed that $\mu_k^*(B_k^c) \to 0$ when $k \to \infty$, which implies:

$$\sum_{z \in B_k^c} \lambda_z^k c(z) \to 0 \tag{4}$$

Analogous, we observe that:

$$\begin{aligned} \left| \sum_{z \in B_k} \left(\overline{\lambda}_z^k - \lambda_z^k \right) c(z) \right| &\leq s \left| \sum_{z \in B_k} \left(\overline{\lambda}_z^k - \lambda_z^k \right) \right| \\ &= s \left| \sum_{z \in B_k} \overline{\lambda}_z^k - \sum_{z \in B_k} \lambda_z^k \right| \\ &= s \left| 1 - \sum_{z \in B_k} \lambda_z^k \right| = s \mu_k^*(B_k^c) \end{aligned}$$

and then

$$\sum_{z \in B_k} \left(\overline{\lambda}_z^k - \lambda_z^k \right) c(z) \to 0 \tag{5}$$

With the right side of the equation (3) and the equations (4) and (5), the proposition is proved. □

Let $B = X \times Y$ and we consider some sequence of sets such that $B_0 = B$ and $B_n \downarrow supp(\mu)$, so for the Proposition 1, there exist a sequence of optimal measures $\{\mu_{n_k}^*\}_{k \in \mathbb{N}}$ for the MK_{n_k} problems such that each measure's support surround the support of the optimal measure μ^* for the MK problem. Furthermore, the relevant support of each $\mu_{n_k}^*$ measure converge in Hausdorff distance to $supp(\mu^*)$.

For notation, lets call B'_n the sets defined in the Step 3 of the algorithm. If we could stablish the sequence $\{n_k\}_{k \in \mathbf{N}}$, then the algorithm will perform as follows; the first step will compute $\mu_{n_1}^*$ solving the MK_{n_1} problem, which will give us information about the localization of $supp(\mu^*)$; the next step is to define a set B'_1 with $supp(\mu_{n_1}^*)$ where $supp(\mu^*) \subset B'_{n_1}$ and which will be an approximation of the

 doi:10.20944/preprints202603.1455.v1

set B_1; the following is to solve the WMK_{n_2} problem in the intersection of B'_{n_1} and the grid $X_{n_2} \times Y_{n_2}$; the solution of the WMK_{n_2} problem will give us a measure $\overline{\mu}^*_{n_2}$; this is repeated until some favorable conditions are fulfilled. Let us observe that for each $k \in \mathbb{N}$, $\overline{\mu}^*_{n_k}$ it is not necessarily equal to $\mu^*_{n_k}$, but because of the Proposition 2, the measures defined with this method will converge to μ^*.

The wavelets play a fundamental role in this procedure. Because we are searching an optimal plan (the support of the optimal measure) that lies in a continuous function, the wavelets irregularities localization properties help us to determine where a continuous function could lie based on the information given by each approximation $supp(\overline{\mu}_{n_k})$. It is important to highlight that the efficiency of the algorithm will depend on the wavelet length, given that as it defines the thickness of the sets B'_{n_k}.

Complexity analysis. Let $X_n = Y_n = \left\{ \frac{i}{2^n} \mid 0 \leq i \leq 2^n \right\}$. Solving the optimal transport problem on the full grid $X_n \times Y_n$ requires $|X_n \times Y_n| = 2^n \cdot 2^n = 4^n$ transport variables. Hence, the complexity is $O(4^n)$.

Wavelet-based sparsity. Suppose the optimal plan μ^* is concentrated along the graph of a Lipschitz (or piecewise smooth) function. Then, the transport support lies on a low-dimensional manifold. A wavelet transform of the indicator function $T_n(x, y)$ for this support yields sparse coefficients due to vanishing moments. Using results from nonlinear approximation theory Mallat [13], the number of significant 2D wavelet coefficients needed to approximate such a transport plan with accuracy ϵ is $k_n = O(n2^n)$. This estimate accounts for $O(2^n)$ coefficients per level summed across n levels.

The WMK formulation restricts the transport plan to a sparse domain O_n defined by the wavelet support. As a result, the number of variables is $O(k_n) = O(n2^n) \ll O(4^n)$, yielding an exponential reduction in problem size. Although the number of variables in the new algorithm also grows exponentially, it remains much smaller than in the classical case. For example, the classical algorithm for $n = 12$ requires 16,777,216 variables, whereas the proposed method requires 94,208 variables for the db10 case.

This complexity reduction enables the WMK scheme to scale to higher resolutions where traditional MK solvers become intractable. The localization properties of wavelets further guarantee that the recovered support converges to the true optimal plan with high fidelity, even at coarse scales.

5. Experimental Setup

We consider a well-known problem used in Gabriel [4] to analyze a wavelet-based algorithm. Let $X = Y = [0, 1]$, and let the marginal distributions ν_1, ν_2 be uniform (Lebesgue measure). The cost function is $c(x, y) = x^2y - xy^2$, a nonlinear, non-symmetric function that introduces a moderate level of complexity for testing.

The exact solution to this problem is known and corresponds to the optimal transport map:

$$f(t) = \begin{cases} \frac{1}{4} + t, & \text{for } t \in \left[0, \frac{3}{4}\right), \\ 1 - t, & \text{for } t \in \left[\frac{3}{4}, 1\right]. \end{cases}$$

The graph of this piecewise-defined optimal transport function is shown in Figure 3 (fourth subfigure), which visualizes the structure of the optimal coupling under the cost function $c(x, y) = x^2y - xy^2$. The exact minimum value of the MK problem is $\min(\text{MK}) = -\frac{9}{256} = -0.03515625$.

We construct the discretized sets $X_n = Y_n = \left\{ \frac{k}{2^n} \,\middle|\, k = 0, 1, \ldots, 2^n \right\}$. For multiresolution analysis, we use the scaling functions **Coiflet1** (coif1), **Haar**, and **Daubechies10** (db10), which have compact support and vanishing moments. The filter lengths are 2, 6, 10, for Haar, Coiflet1, Daubechies10, scaling functions (the properties of each scaling function can be seen in Hardle [14])).

For the numerical applications, we used Python (v3.12.11) on a standard computer (13th Gen Intel(R) Core(TM) i7-13650HX; 16 GB RAM). The libraries used were `PyWavelets` for the wavelet transform, `PuLP` to solve the linear programming problems, and `NumPy` for matrix calculations.

We compare the classical discretized linear programming approach Gabriel [4] (denoted MK_n) with our wavelet-guided refinement scheme (denoted WMK_n). At each stage:

1. MK_n is solved to obtain a discrete optimal transport plan μ^*_n;

 doi:10.20944/preprints202603.1455.v1

2. A wavelet transform identifies singularities in the support of μ_n^*;
3. A reduced domain $O_{n+1} \subset X_{n+1} \times Y_{n+1}$ is definedd;
4. The transport problem is solved again, but only within O_{n+1} to obtain $\overline{\mu}_{n+1}^*$.

The adaptive search regions B_k of the Step 3 for $k \in \{7, 9, 12\}$ and the scaling function Coiflet1 are illustrated in Figure 3. Each image represents the search region B_k (yellow) as a subset of $X \times Y$. For each B_k image, it is illustrated as a red dotted line the localization of the optimal transport function. It should be noted that each search region B_k surrounds the optimal transport function. The region in black denotes the subset of $X \times Y$ where the algorithm does not search for the support of each approximate measure $\overline{\mu}_k^*$.

The regions for db10 are similar to those shown in Figure 3. In the Figure 2 we observe that as well as with the scaling function Coiflet1, the yellow region surrounds the optimal transport function, but because the filter length of the Daubechies10 function is wider than the filter length of the Coiflet1 function, the yellow regions in this case are bigger, which could be helpful to ensure that inside the region is the $supp(\mu^*)$, but it implies more variables to deal with in the sub-process of the linear program.

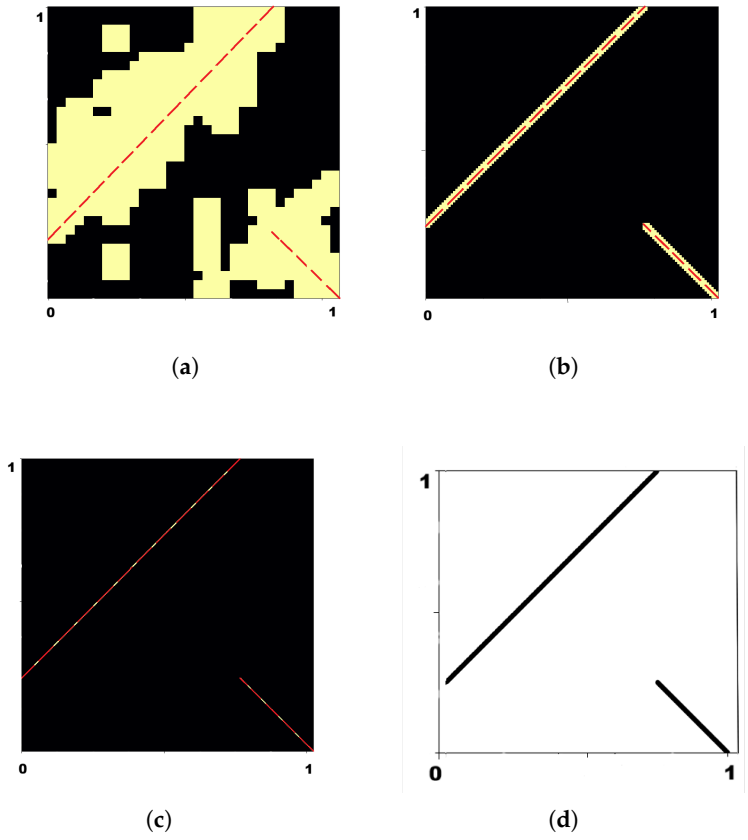

Figure 2. Visualization of B_7 (**a**), B_9 (**b**), and B_{12} (**c**), obtained using the Coiflet1 wavelet, together with the optimal coupling function F_1 (**d**). The red dotted line in subfigures (**a**)–(**c**) indicates the localized region of the optimal coupling.

 doi:10.20944/preprints202603.1455.v1

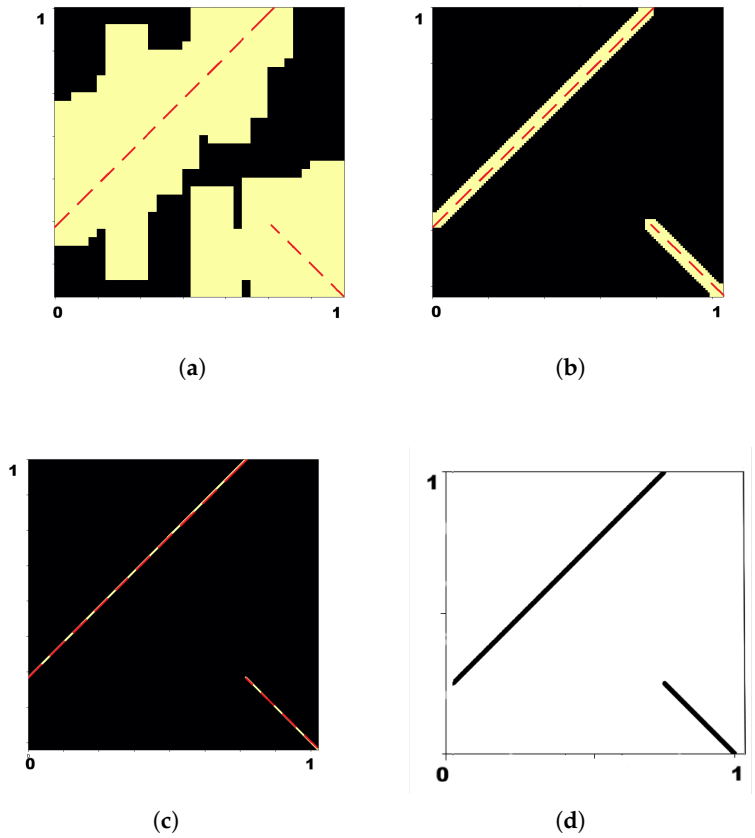

Figure 3. Visualization of B_7 (**a**), B_9 (**b**), and B_{12} (**c**), obtained using the Daubechies10 wavelet, together with the optimal coupling function F_1 (**d**). The red dotted line in subfigures (**a**)–(**c**) indicates the localized region of the optimal coupling.

In the case of the Haar wavelet, the regions do not converge to the optimal coupling, see Figure 4. That is because the filter of the length of the Haar scaling function is smaller than the other two scaling function, and then the searching regions (yellow) in this case would not capture the support of μ^* inside it. If the Haar wavelet is needed (for example, because it is symmetrical), then the problem of the filter length could be improved if its considered a deeper level in the wavelet transform (Step 3), because in a deeper level V_k the support of the generating function $\phi_{k,0}$ is wider.

 doi:10.20944/preprints202603.1455.v1

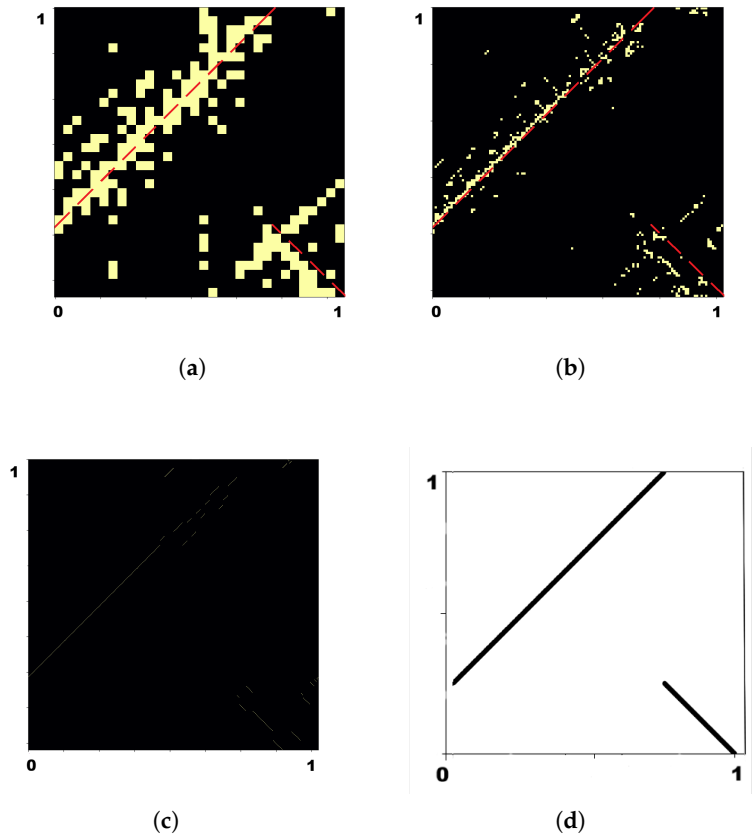

Figure 4. Visualization of B_7 (**a**), B_9 (**b**), and B_{12} (**c**), obtained using the Haar wavelet, together with the optimal coupling function F_1 (**d**). The red dotted line in subfigures (**a**)–(**b**) indicates the localized region of the optimal coupling.

6. Results

Table 1 summarize the number of variables used in the linear program, the approximation error, and the computed minimum values for both schemes and each scaling function. The classical method becomes infeasible beyond level $n = 8$ with the current system configuration, while the wavelet-enhanced method maintains computational efficiency and accuracy up to $n = 12$.

The wavelet-guided method offers a scalable and accurate approach to solving the Monge–Kantorovich mass transfer problem by leveraging multiresolution wavelet analysis to adaptively refine regions of interest. By concentrating computational effort on the most relevant areas of the transport plan's support, it achieves high precision while significantly reducing the number of optimization variables compared to classical uniform discretization schemes. Although not all wavelet bases perform equally well—such as the Haar wavelet, which suffers from information loss, the method demonstrates strong convergence properties when suitable scaling functions are chosen.

Numerical experiments on synthetic test cases with known analytical solutions confirm the method's efficiency and robustness. The algorithm successfully reconstructs piecewise-defined transport maps, reaching high resolutions (up to $n = 12$) without exhausting memory resources, illustrating its capacity to handle structurally complex cost functions and domains. These results highlight the method's computational advantage and its potential for large-scale problems that are infeasible using traditional Monge–Kantorovich solvers.

Beyond theoretical validation, the framework has practical relevance across various domains, including image registration, edge detection, shape morphing, logistics, and optimal transport applications in machine learning and biomedical imaging.

 doi:10.20944/preprints202603.1455.v1

Table 1. Comparison of the classical method (MK_n) and our proposed method (WMK_n) for multiple resolutions levels and wavelet types.

Setting	No. Var.	Error	Approx. Value
MK_7 (no wavelet)	16,384	3.8147×10^{-6}	−0.0351524353
WMK_7 (coif1)	8,000	3.8147×10^{-6}	−0.0351524353
WMK_7 (db10)	11,072	3.8147×10^{-6}	−0.0351524353
WMK_7 (Haar)	2,656	1.1805×10^{-3}	−0.0333511084
MK_8 (no wavelet)	65,536	9.537×10^{-7}	−0.0351552963
WMK_8 (coif1)	5,072	9.537×10^{-7}	−0.0351552963
WMK_8 (db10)	9,072	9.537×10^{-7}	−0.0351552963
WMK_8 (Haar)	3,888	3.4115×10^{-3}	−0.0317446963
MK_9 (no wavelet)	262,144	N/A	Insufficient memory
WMK_9 (coif1)	10,192	2.385×10^{-7}	−0.0351560115
WMK_9 (db10)	18,288	2.385×10^{-7}	−0.0351560115
WMK_9 (Haar)	7,584	4.2620×10^{-3}	−0.0351560115
MK_{10} (no wavelet)	1,048,576	N/A	Insufficient memory
WMK_{10} (coif1)	20,432	5.97×10^{-8}	−0.0351561903
WMK_{10} (db10)	36,720	5.97×10^{-8}	−0.0351561903
WMK_{10} (Haar)	4,096	4.650×10^{-3}	−0.0305061907
MK_{11} (no wavelet)	4,194,304	N/A	Insufficient memory
WMK_{11} (coif1)	40,912	1.50×10^{-8}	−0.0351562350
WMK_{11} (db10)	73,584	1.49×10^{-8}	−0.0351562351
WMK_{11} (Haar)	8,192	4.8349×10^{-3}	−0.0303212723
MK_{12} (no wavelet)	16,777,216	N/A	Insufficient memory
WMK_{12} (coif1)	94,208	1.47×10^{-8}	−0.0351562353
WMK_{12} (db10)	159,680	1.62×10^{-8}	−0.0351562338
WMK_{12} (Haar)	16,384	4.9251×10^{-3}	−0.0302310522

Future work will focus on extending the approach to handle multimodal and singular transport supports, integrating adaptive wavelet bases tuned to problem-specific features, and combining it with entropic regularization methods for enhanced scalability and generalization in higher-dimensional and multi-marginal transport problems.

While the data are artificial, the structure and algorithm used are applicable to real-world problems such as image registration and edge detection in computer vision, shape matching and morphing in computer graphics, resource allocation and logistics in operations research, optimal transport in machine learning (e.g., domain adaptation, generative modeling), and biomedical imaging tasks such as anatomical alignment or intensity normalization Bonne [9], Kamsu [10], Chen [11], Kreinovich [12].

References

1. Rachev, S., Rüschendorf, L.: *Mass Transportation Problems*. Vols. I and II. Springer, New York (1998)
2. Villani, C.: *Optimal Transport: Old and New*. Vol. 338. Springer, Berlin (2008).
3. Hasan, M.M.A., Zaman, M., Jawad, A., Santamaria-Pang, A., Lee, H.H., Tarapov, I., See, K.B., Imran, M.S., Roy, A., Pourmohammadi Fallah, Y., Asadizanjani, N., Forghani, R.: WaveFormer: A 3D Transformer with Wavelet-Driven Feature Representation for Efficient Medical Image Segmentation. In: de Bruijne, M., Veni, G., Naji, K. (eds) *Medical Image Computing and Computer Assisted Intervention – MICCAI 2025*. Lecture Notes in Computer Science, vol. 15963, pp. 684–694. Springer, Cham (2025).
4. Gabriel, J.R., González-Hernández, J., López-Martínez, R.R.: Numerical approximations to the mass transfer problem on compact spaces. *IMA Journal of Numerical Analysis* **30**(4), 1121–1136 (2010)
5. Lee, W., Lai, R., Li, W., Osher, S.: Generalized unnormalized optimal transport and its fast algorithms. *Journal of Computational Physics* **436**, 110041 (2021)
6. Chi, J., Wang, B., Chen, H., Zhang, L., Li, X., Ouyang, J.: Approximate continuous optimal transport. *International Journal of Intelligent Systems* **37**(8), 5354–5380 (2022)

 doi:10.20944/preprints202603.1455.v1

7. Tanguy, E., Desolneux, A., Delon, J.: Constrained approximation optimal transport maps. *ESAIM: Control, Optimisation and Calculus of Variations* **31**, 70 (2025)
8. Acosta-Portilla, Juan Rafael, Gonzáles-Flores, Carlos, López-Martinez, Raquiel Rufino, Sánchez-Nungaray Armando: Efficient Methond to Solve the Monge–Kantorovich Problem Using Wavelet Analysis *Axioms* **12**(2023)
9. Bonneel, N., Digne, J.: A survey of optimal transport for computer graphics and computer vision. *Computer Graphics Forum* **42**(2), 439–460 (2023)
10. Kamsu-Foguem, B., Msouobu Gueuwou, S.L., Kounta, C.A.K.A.: Generative adversarial networks based on optimal transport: a survey. *Artificial Intelligence Review* **56**(7), 6723–6773 (2023)
11. Chen, X., Benveniste, H., Tannenbaum, A.R.: Unbalanced regularized optimal mass transport with applications to fluid flows in the brain. *Scientific Reports* **14**(1), 1111 (2024)
12. Kreinovich, V., Yamaka, W., Leurcharusmee, S.: *Applications of Optimal Transport to Economics and Related Topics*. Springer, Cham (2008)
13. Mallat, S.: *A Wavelet Tour of Signal Processing*. 3rd edn. Academic Press (2009)
14. Härdle, W., Kerkyacharian, G., Picard, D., Tsybakov, A.: *Wavelets, Approximation and Statistical Applications*. 1st edn. Springer, New York, NY (2012)
15. Daubechies, I.: *Ten Lectures on Wavelets*. SIAM (1992)
16. Parthasarathy, K. R.: *Probability Measures on Metric Spaces*. Academic Press (1967)