Article

Efficient Method to Solve the Monge–Kantarovich Problem Using Wavelet Analysis

Juan Rafael Acosta-Portilla [1], Carlos González-Flores [2], Raquiel Rufino López-Martínez [3] and Armando Sánchez-Nungaray [3,*]

1   Instituto de Investigaciones y Estudios Superiores Económicos y Sociales, Universidad Veracruzana, Veracruz 94294, Mexico; juaacosta@uv.mx
2   Escuela Superior de Ingeniería Mecánica y Eléctrica Zacatenco, Instituto Politécnico Nacional, Mexico City 07738, Mexico; cfgonzalez@esimez.mx
3   Facultad de Matemáticas, Universidad Veracruzana, Veracruz 94294, Mexico; ralopez@uv.mx
*   Correspondence: armsanchez@uv.mx

Abstract: In this paper, we present and justify a methodology to solve the Monge–Kantorovich mass transfer problem through Haar multiresolution analysis and wavelet transform with the advantage of requiring a reduced number of operations to carry out. The methodology has the following steps. We apply wavelet analysis on a discretization of the cost function level j and obtain four components comprising one corresponding to a low-pass filter plus three from a high-pass filter. We obtain the solution corresponding to the low-pass component in level $j-1$ denoted by μ_{j-1}^*, and using the information of the high-pass filter components, we get a solution in level j denoted by $\hat{\mu}_j$. Finally, we make a local refinement of $\hat{\mu}_j$ and obtain the final solution μ_j^σ.

Keywords: mass transfer problem; wavelets; multiresolution analysis

MSC: 42C40; 49Q20; 65T60

1. Introduction

In recent years, schemes to approximate infinite linear programs have become very important in theory. The authors of [1] showed that under suitable assumptions, the program's optimum value can be approximated by the values of finite-dimensional linear programs and that every accumulation point of a sequence of optimal solutions for the approximating programs is an optimal solution for the original problem. In particular, in [2] the authors studied the Monge–Kantorovich mass transfer (MT) problem on metric spaces. They considered conditions under the MT problem as solvable and, furthermore, that an optimal solution can be obtained as the weak limit of a sequence of optimal solutions to suitably approximate MT problems.

Moreover, in [3], the authors presented a numerical approximation for the value of the mass transfer (MT) problem on compact metric spaces. A sequence of transportation problems was built, and it proved that the value of the MT problem is a limit of the optimal values of these problems. Moreover, they gave an error bound for the numerical approximation. A generalization of this scheme of approximation was presented in [4,5]. They proposed an approximation scheme for the Monge–Kantorovich (MK) mass transfer problem on compact spaces that consisted of reducing to solve a sequence of finite transport problems. The method presented in that work uses a metaheuristic algorithm inspired by a scatter search in order to reduce the dimensionality of each transport problem. Finally, they provided some examples of that method.

On the other hand, the authors of [6] provided orthonormal bases for $L^2(\mathbb{R}^n)$ that have properties that are similar to those enjoyed by the classical Haar basis for $L^2(\mathbb{R})$. For example, each basis consists of appropriate dilates and translates of a finite collection of

"piecewise constant" functions. The construction is based on the notion of multiresolution analysis and reveals an interesting connection between the theory of compactly supported wavelet bases and the theory of self-similar tilings. Recent applications of the wavelet filter methodology have been used in various problems arising in communication systems and detection of thermal defects (see, for example, [7,8], respectively).

In [9], the authors gave a scheme to approximate the MK problem based on the symmetries of the underlying spaces. They took a Haar-type MRA constructed according to the geometry of the spaces. Thus, they applied the Haar-type MRA based on symmetries to the MK problem and obtained a sequence-of-transport problem that approximates the original MK problem for each MRA space. Note that in the case of Haar's classical wavelet, this methodology coincides with the methods presented in [2,3].

It is important to note that various scientific problems are modeled through the Monge–Kantorovich approach; therefore, providing new efficient methodologies to find approximations of such problems turns out to be very useful. Within the applications of problems whose solutions are the Monge–Kantorovich problem are found: the use of the transport problem for the analysis of elastic image registration (see, for example, [10–12]). Other optimization problems related to this topic and differential equation tools can be found in recent works such as [13,14].

The main goal of this paper is to present a scheme of approximation of the MK problem based on wavelet analysis in which we use wavelet filters to split the original problem. That is, we apply the filter to the discrete cost function in level j, which results in a cost function of level $j - 1$ and three components of wavelet analysis. Using the information of the cost function given by the low-pass filter, which belongs to level $j - 1$, we construct μ^*_{j-1} a solution of the MK problem for that level $j - 1$, and using the additional information, the other three components of wavelet analysis are extended to $\hat{\mu}_j$, which is a solution to level j, where the projection of $\hat{\mu}_j$ to level $j - 1$ is μ^*_{j-1}. Finally, we make a local analysis of the solution $\hat{\mu}_j$ to obtain an improved solution based on the type of points of that solution (we have two type of points that are defined in the base in the connectedness of the solution).

This work has three non-introductory sections. In the first of them we present the Haar multiresolution analysis (MRA) in one and two dimensions. Next, we relate this to the absolutely continuous measures over a compact in $\mathbb{R}^2$. We finish with the definition of the Monge–Kantorovich mass transfer problem and its relation to the MRA.

In the second section, we define a proximity criterion for the components of the support of the simple solutions of the MK problem and study in detail the problem of, given a solution μ^*_{j-1} at level $j - 1$ of resolution for the MK problem, construct a feasible solution $\hat{\mu}_j$ for the MK problem at level of resolution j such that it is a refinement of the solution with lees resolution.

On the other hand, in the third section we present a methodological proposal to solve the MK^j problem such that it can be summarized in a simple algorithm of six steps:

Step 1. We consider a discretization of the cost function for the level j, denoted by c_j.

Step 2. We apply the wavelet transform to c_j; we obtain the low-pass component c_{j-1} and three high-pass components, denoted by Ψ_1, Ψ_2 and Ψ_3, respectively.

Step 3. Using c_{j-1} and the methodology of [3,4,9], we obtain a solution μ^*_{j-1} for MK^{j-1} associated with this cost function.

Step 4. We classify the points of the support of the solution μ^*_{j-1} by proximity criteria as points of Type I or Type II.

Step 5. Using the solution μ^*_{j-1}, the information of the high-pass components and Lemma 1, we obtain a feasible solution for the level j, which is denoted by $\hat{\mu}_j$. This feasible solution has the property that its projection to the level $j - 1$ is equal to μ^*_{j-1}; moreover, the support of $\hat{\mu}_j$ is contained in the support of μ^*_{j-1}.

Step 6. The classification of the points of μ^*_{j-1} induce classification of the points in $\hat{\mu}_j$ by contention in the support. Over the points of Type I of the solution $\hat{\mu}_j$, we do not move those points. For the points of Type II, we apply a permutation to the

solution over the two points that better improves the solution, and we repeat the process with the rest of the points.

Finally, we present a series of examples that use the proposed methodology based on wavelet analysis and compare their results with those obtained applying the methodology of [3,4,9].

2. Preliminaries

2.1. One-Dimensional MRA

The results of this and the following subsection are well known, and for a detailed exposition, we recommend consulting [15–17]. We begin by defining a general multiresolution analysis and developing the particular case of the Haar multiresolution analysis on $\mathbb{R}$. Given $a > 0$ and $b \in \mathbb{R}$, the dilatation operator D_a and the translation operator T_b are defined by

$$(D_a f)(x) = a^{1/2} f(ax) \quad \text{and} \quad (T_b f)(x) = f(x - b) \tag{1}$$

for every $f \in L^2(\mathbb{R})$, where the latter denotes the usual Hilbert space of square integrable real functions defined on $\mathbb{R}$. A multiresolution analysis (MRA) on $\mathbb{R}$ is a sequence of subspaces $(V_j)_{j \in \mathbb{Z}}$ of $L^2(\mathbb{R})$ such that it satisfies the following properties:

(1) $V_j \subset V_{j+1}$ for every $j \in \mathbb{Z}$.
(2) $L^2(\mathbb{R}) = \overline{\text{span}} \bigcup_{j \in \mathbb{Z}} V_j$.
(3) $\bigcap_{j \in \mathbb{Z}} V_j = \{0\}$.
(4) $V_j = D_{2^j} V_0$.
(5)   There exists a function $\varphi \in L^2$, called the scaling function, such that the collection $\{T_j \varphi\}_{j \in \mathbb{Z}}$ is an orthonormal system of translates and

$$V_0 = \overline{\text{span}}\{T_j \varphi\}_{i \in \mathbb{Z}}. \tag{2}$$

We denote as χ_A the characteristic function of the set A. Then, the Haar scaling function is defined by

$$\varphi(x) = \chi_{[0,1)}(x). \tag{3}$$

For each pair $j, k \in \mathbb{Z}$, we call

$$I_{j,k} = [2^{-j}k, 2^{-j}(k+1)). \tag{4}$$

Hence, we define the function

$$\begin{aligned}
\varphi_{j,k}(x) &= (T_k D_{2^j} \varphi)(x) \\
&= 2^{j/2} \varphi(2^j x - k) \\
&= 2^{j/2} \chi_{I_{j,k}}(x).
\end{aligned} \tag{5}$$

The collection $\{\varphi_{j,k}\}_{j,k \in \mathbb{Z}}$ is called the system of Haar scaling functions. For $j_0 \in \mathbb{Z}$, the collection $\{\varphi_{j_0,k}\}_{k \in \mathbb{Z}}$ is referred to as the system of scale j_0 Haar scaling functions. The Haar function is defined by

$$\psi(x) = \chi_{[0,1/2)}(x) - \chi_{[1/2,1)}(x). \tag{6}$$

For each pair $j, k \in \mathbb{Z}$, we define the function

$$\begin{aligned}
\psi_{j,k}(x) &= (T_k D_{2^j} \psi)(x) \\
&= 2^{j/2} \psi(2^j x - k) \\
&= 2^{j/2} (\chi_{I_{j+1,2k}} - \chi_{I_{j+1,2k+1}})(x).
\end{aligned} \tag{7}$$

The collection $\{\psi_{j,k}\}_{j,k\in\mathbb{Z}}$ is referred to as the Haar system on $\mathbb{R}$. For $j_0 \in \mathbb{Z}$, the collection $\{\psi_{j_0,k}\}_{k\in\mathbb{Z}}$ is referred to as the system of scale j_0 Haar functions. It is well known that with respect to the usual inner product $\langle\cdot,\cdot\rangle$ in $L^2(\mathbb{R})$, the Haar system on $\mathbb{R}$ is an orthonormal system. Moreover, for each $j_0 \in \mathbb{Z}$, the collection of scale j_0 Haar scaling functions is an orthonormal system. Thus, for each $j \in \mathbb{Z}$, the approximation operator P_j on $L^2(\mathbb{R})$ is defined by

$$(P_j f)(x) = \sum_k \langle f, \varphi_{j,k}\rangle \varphi_{j,k}(x), \text{ for all } f \in L^2(\mathbb{R}), \tag{8}$$

and the approximation space V_j by

$$V_j = \overline{\text{span}}\{\varphi_{j,k}\}_{k\in\mathbb{Z}}. \tag{9}$$

The collection $\{V_j\}_{j\in\mathbb{Z}}$ is called Haar multiresolution analysis. Similarly, we have that for each $j_0 \in \mathbb{Z}$, the detail operator Q_j on $L^2(\mathbb{R})$ is defined by

$$(Q_j f)(x) = (P_{j+1}f)(x) - (P_j f)(x), \text{ for all } f \in L^2(\mathbb{R}), \tag{10}$$

and the wavelet space W_j by

$$W_j = \overline{\text{span}}\{\psi_{j,k}\}_{k\in\mathbb{Z}}. \tag{11}$$

Note that $V_{j+1} = V_j \oplus W_j$ for all $j \in \mathbb{Z}$. Hence, the collection

$$\{\varphi_{j_0,k}, \psi_{j,k} : j \geq j_0, k \in \mathbb{Z}\} \tag{12}$$

is a complete orthonormal system on $\mathbb{R}$; this system is called the scale j_0 Haar system on $\mathbb{R}$. As a consequence, the Haar system $\{\psi_{j,k}\}_{j,k\in\mathbb{Z}}$ is a complete orthonormal system on $\mathbb{R}$.

2.2. Two-Dimensional MRA

To obtain the Haar MRA on $\mathbb{R}^2$, we consider the Haar MRA on $\mathbb{R}$ defined in the previous subsection with scaling and Haar functions φ and ψ, and from them, through a tensor product approach we can construct a two-dimensional scaling and Haar function. First, we define the four possible products:

$$\begin{aligned}\Phi(x,y) &= \varphi(x)\varphi(y), \quad \Psi^{(1)}(x,y) = \varphi(x)\psi(y) \\ \Psi^{(2)}(x,y) &= \psi(x)\varphi(y), \quad \Psi^{(3)}(x,y) = \psi(x)\psi(y),\end{aligned} \tag{13}$$

which are the scaling function associated with the unitary square and three Haar functions, respectively. Hence, for each $j, k_1, k_2 \in \mathbb{Z}$, we define naturally the scaling and Haar function systems:

$$\begin{aligned}\Phi_{j,k_1,k_2}(x,y) &= \varphi_{j,k_1}(x)\varphi(y)_{j,k_2}, \quad \Psi^{(1)}_{j,k_1,k_2}(x,y) = \varphi(x)_{j,k_1}\psi(y)_{j,k_2} \\ \Psi^{(2)}_{j,k_1,k_2}(x,y) &= \psi(x)_{j,k_1}\varphi(y)_{j,k_2}, \quad \Psi^{(3)}_{j,k_1,k_2}(x,y) = \psi_{j,k_1}(x)\psi(y)_{j,k_2}.\end{aligned} \tag{14}$$

Then for $j_0 \in \mathbb{Z}$, as in the one-dimensional case, we have that the collection

$$\{\Phi_{j_0,k_1,k_2} : k_1, k_2 \in \mathbb{Z}\} \cup \{\Psi^{(i)}_{j,k_1,k_2} : 1 \leq i \leq 3, j \geq j_0\} \tag{15}$$

is an orthonormal basis on $L^2(\mathbb{R}^2)$. Thus, the collection

$$\{\Psi^{(i)}_{j,k_1,k_2} : 1 \leq i \leq 3, j, k_1, k_2 \in \mathbb{Z}\} \tag{16}$$

is an orthonormal basis on $L^2(\mathbb{R}^2)$. Then for $j \in \mathbb{Z}$ and $f \in L^2(\mathbb{R}^2)$, the approximation operator is defined by

$$(\mathbf{P}_j f)(x,y) = \sum_{k_1}\sum_{k_2}\langle f, \Phi_{j,k_1,k_2}\rangle \Phi_{j,k_1,k_2}(x,y), \tag{17}$$

and for $i = 1, 2, 3$, the detail operators are

$$(\mathbf{Q}_j^{(i)} f)(x,y) = \sum_{k_1}\sum_{k_2}\langle f, \Psi_{j,k_1,k_2}^{(i)}\rangle \Psi_{j,k_1,k_2}^{(i)}(x,y). \tag{18}$$

Hence, the projection can be written as

$$(\mathbf{P}_{j+1} f)(x,y) = (\mathbf{P}_j f)(x,y) + (\mathbf{Q}_j^{(1)} f)(x,y) + (\mathbf{Q}_j^{(2)} f)(x,y) + (\mathbf{Q}_j^{(3)} f)(x,y). \tag{19}$$

We will describe the approximation $\mathbf{P}_j$ and detail operators $\mathbf{Q}_j^{(i)}$ from the geometric point of view. First of all, we fix some $j, k_1, k_2 \in \mathbb{Z}$ and define the square

$$\begin{aligned}
S(j,k_1,k_2) &= I_{j,k_1} \times I_{j,k_2} \\
&= [2^{-j}k_1, 2^{-j}(k_1+1)) \times [2^{-j}k_2, 2^{-j}(k_2+1)).
\end{aligned} \tag{20}$$

Then, we have that Φ_{j,k_1,k_2} is the characteristic function of $S(j,k_1,k_2)$, in symbols

$$\Phi_{j,k_1,k_2}(x,y) = \chi_S(x,y) \tag{21}$$

where $S = S(j,k_1,k_2)$. Therefore for $f \in L^2(\mathbb{R}^2)$, the operator $\mathbf{P}_j f$ acts as a discretization of f that is constant over the disjointed $S(j,k_1,k_2)$ squares. On the other hand, we can split $S(j,k_1,k_2)$ as follows (see Figure 1):

$$\Phi_{j,k_1,k_2}(x,y) \qquad \Psi_{j,k_1,k_2}^{(1)}(x,y) \qquad \Psi_{j,k_1,k_2}^{(2)}(x,y) \qquad \Psi_{j,k_1,k_2}^{(3)}(x,y)$$
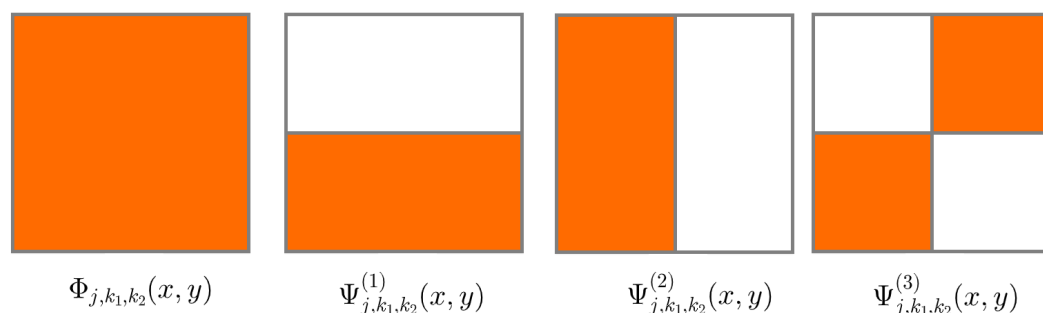
Figure 1. Functions Φ_{j,k_1,k_2}, $\Psi_{j,k_1,k_2}^{(1)}$, $\Psi_{j,k_1,k_2}^{(2)}$ and $\Psi_{j,k_1,k_2}^{(3)}$.

(i)   Into two rectangles with half the height and the same width as $S(j,k_1,k_2)$, namely, the sets

$$\begin{aligned}
A_1(j,k_1,k_2) &= [2^{-j}k_1, 2^{-j}(k_1+1)) \times [2^{-j-1}2k_2, 2^{-j-1}(2k_2+1)) \text{ and} \\
A_2(j,k_1,k_2) &= [2^{-j}k_1, 2^{-j}(k_1+1)) \times [2^{-j-1}2k_2+1, 2^{-j-1}(2k_2+2)).
\end{aligned} \tag{22}$$

(ii) Into two rectangles with the same height and half the width as $S(j,k_1,k_2)$, namely,

$$\begin{aligned}
B_1(j,k_1,k_2) &= [2^{-j-1}2k_1, 2^{-j-1}(2k_1+1)) \times [2^{-j}k_2, 2^{-j}(k_2+1)) \text{ and} \\
B_2(j,k_1,k_2) &= [2^{-j-1}2k_1+1, 2^{-j-1}(2k_1+2)) \times [2^{-j}k_2, 2^{-j}(k_2+1)).
\end{aligned} \tag{23}$$

(iii) Into four squares with half the side lengths of $S(j, k_1, k_2)$, namely,

$$
\begin{aligned}
C_1(j, k_1, k_2) &= [2^{-j-1}2k_1, 2^{-j-1}(2k_1 + 1)) \times [2^{-j-1}2k_2, 2^{-j-1}(2k_2 + 1)), \\
C_2(j, k_1, k_2) &= [2^{-j-1}2k_1 + 1, 2^{-j-1}(2k_1 + 2)) \times [2^{-j-1}2k_2, 2^{-j-1}(2k_2 + 1)), \\
C_3(j, k_1, k_2) &= [2^{-j-1}2k_1, 2^{-j-1}(2k_1 + 1)) \times [2^{-j-1}2k_2 + 1, 2^{-j-1}(2k_2 + 2)) \text{ and} \\
C_4(j, k_1, k_2) &= [2^{-j-1}2k_1 + 1, 2^{-j-1}(2k_1 + 2)) \times [2^{-j-1}2k_2 + 1, 2^{-j-1}(2k_2 + 2)).
\end{aligned}
\tag{24}
$$

Hence, for $i = 1, 2, 3$, the function $\Psi^{(i)}_{j, k_1, k_2}$ is defined by

$$
\begin{aligned}
\Psi^{(1)}_{j, k_1, k_2}(x, y) &= \chi_{A_1}(x, y) - \chi_{A_2}(x, y) \text{ with } A_l = A_l(j, k_1, k_2) \text{ for } l = 1, 2. \\
\Psi^{(2)}_{j, k_1, k_2}(x, y) &= \chi_{B_1}(x, y) - \chi_{B_2}(x, y) \text{ with } B_l = B_l(j, k_1, k_2) \text{ for } l = 1, 2. \\
\Psi^{(3)}_{j, k_1, k_2}(x, y) &= \chi_{C_1 \cup C_4}(x, y) - \chi_{C_2 \cup C_3}(x, y) \text{ with } C_l = C_l(j, k_1, k_2) \text{ for } l = 1, 2, 3, 4.
\end{aligned}
\tag{25}
$$

(see Figure 1). Thus, the image of $f \in L^2(\mathbb{R}^2)$ under the detail operator $\mathbf{Q}^{(i)}_j$ is a function $\mathbf{Q}^{(i)}_j f$ formed by pieces that oscillate symmetrically on each square $S(j, k_1, k_2)$.

Now we have the elements to define an MRA on $\mathbb{R}^2$; to do this, we will use the notation introduced for the one-dimensional MRA on $\mathbb{R}$ defined in the previous subsection. For each $j \in \mathbb{Z}$, we define

$$
\mathbf{V}_j = V_j \otimes V_j.
\tag{26}
$$

Then the collection $\{\mathbf{V}_j\}_{j \in \mathbb{Z}}$ is the Haar multiresolution analysis on $\mathbb{R}^2$, where the dilatation and translation operators are defined by

$$
\mathbf{D}_{2^j} = D_{2^j} \otimes D_{2^j} \text{ and } \mathbf{T}_{m, n} = T_m \otimes T_n,
\tag{27}
$$

respectively. Note that we have the following relation:

$$
\mathbf{V}_{j+1} = \mathbf{V}_j \oplus \mathbf{W}^{(1)}_j \oplus \mathbf{W}^{(2)}_j \oplus \mathbf{W}^{(3)}_j,
\tag{28}
$$

where $\mathbf{W}^{(1)}_j = \mathbf{V}_j \otimes \mathbf{W}_j$, $\mathbf{W}^{(2)}_j = \mathbf{W}_j \otimes \mathbf{V}_j$ and $\mathbf{W}^{(3)}_j = \mathbf{W}_j \otimes \mathbf{W}_j$. For more detail with respect to the Haar MRA, see [15].

2.3. Measures and MRA

In this subsection, we will use the two previous ones and the [9] approach to relate measures over $\mathbb{R}^2$ with the Haar MRA on $\mathbb{R}^2$. The results and definitions presented in this and the following subsection can be found in the classical references [18,19]. We consider a compact subset X of $\mathbb{R}^2$ and a measure μ such that it is absolutely continuous with respect to the Lebesgue measure λ. We call

$$
f_\mu = \frac{d\mu}{d\lambda}
\tag{29}
$$

the Radon–Nikodym derivative of μ with respect to λ. By construction, it necessary holds that $f_\mu \in L^1(X)$. We additionally suppose that $f_\mu \in L^2(X)$. Then, as a consequence of the Haar MRA on $\mathbb{R}^2$, we have that

$$
\|\mathbf{P}_j f_\mu - f_\mu\|_2 \to 0 \text{ as } j \to \infty.
\tag{30}
$$

Moreover, the compactness of X ensures that

$$
\|\mathbf{P}_j f_\mu - f_\mu\|_1 \leq \lambda(X)^{1/2} \|\mathbf{P}_j f_\mu - f_\mu\|_2.
\tag{31}
$$

The above allow us to define μ_j, the approximation of the measure μ to the level $j \in \mathbb{Z}$ of the Haar MRA on $\mathbb{R}^2$, as the measure induced by the projection of the Radon–Nikodym to the level j. That is, μ_j is defined by the relation

$$d\mu_j = \mathbf{P}_j f_\mu \, d\lambda. \tag{32}$$

We denote the expectation of a function $g \in L(\mathbb{R}^2)$ with respect to λ as

$$\mathbb{E}[g] = \int_{\mathbb{R}^2} g d\lambda. \tag{33}$$

Then Theorem 4 and Corollary 5 in [9] ensure that for each $g \in L^2(\mathbb{R}^2)$ and $j \in \mathbb{Z}$, it is fulfilled that

$$\mathbb{E}[\mathbf{P}_j g] = \mathbb{E}[g]. \tag{34}$$

Thus, μ_j is absolutely continuous with respect to the Lebesgue measure. If, additionally, we suppose that μ is a measure with support on X, then (μ_j) converges to μ in the L^1 and L^2 sense. That is, by the Riesz Theorem, we can associate each of these measures μ_j to an integrable function $\hat{\mu}_j \in \mathbf{P}_j L^2(\mathbb{R}^2)$ such that for each Lebesgue measurable set $A \subset \mathbb{R}^2$,

$$
\begin{aligned}
\mu_j(A) &= \int_A \hat{\mu}_j d\lambda \\
&= \int \hat{\mu}_j \chi_A d\lambda \\
&= \mathbb{E}[\hat{\mu}_j \chi_A],
\end{aligned}
\tag{35}
$$

and we apply the respective convergence mode. Further, the compact support of measures ensures that the sequence (μ_j) converges weakly to the measure μ, proof of which can be found in [9] Theorem 7.

2.4. M-K Problem and MRA

In this subsection, we study the Monge–Kantorovich problem from the point of view of the Haar MRA on $\mathbb{R}^2$ (for a detailed exposition, we recommend consulting [9]). Let X and Y be two compact subsets of $\mathbb{R}^2$. We denote by $M^+(X \times Y)$ the family of finite measures on $X \times Y$. Given $\mu \in M^+(X \times Y)$, we denote its marginal measures on X and Y as

$$\Pi_1 \mu(E_1) = \mu(E_1 \times Y) \tag{36}$$

and

$$\Pi_2 \mu(E_2) = \mu(X \times E_2) \tag{37}$$

for each μ-measurable set $E_1 \subset X$ and $E_2 \subset Y$. Let c be a real function defined on $X \times Y$, and η^1, η^2 are two measures defined on X and Y, respectively. The Monge–Kantorovich mass transfer problem is given as follows:

$$
\begin{aligned}
\text{MK: minimize} \quad & \langle \mu, c \rangle := \int c d\mu \\
\text{subject to:} \quad & \Pi_1 \mu = \eta^1, \quad \Pi_2 \mu = \eta^2, \quad \mu \in M^+(X \times Y).
\end{aligned}
\tag{38}
$$

A measure $\mu \in M^+(X \times Y)$ is said to be a feasible solution for the MK problem if it satisfies (38) and $\langle \mu, c \rangle$ is finite. We said that the MK problem is solvable if there is a feasible solution μ^* that attains the optimal value for it. So μ^* is called an optimal solution for (38). If, additionally, we assume that μ, η^1 and η^2 are absolutely continuous with respect to the Lebesgue measures on $\mathbb{R}^2$ and $\mathbb{R}$, then in a natural way, we can discretize the MK problem

through the Haar MRA on $\mathbb{R}^2$ as follows. For $j \in \mathbb{Z}$, we define the MK problem of level j by:

$$\text{MK}_j: \text{minimize} \quad \langle \mu_j, c \rangle := \int c \, d\mu_j$$
$$\text{subject to:} \quad \Pi_1 \mu_j = \eta_j^1, \quad \Pi_2 \mu_j = \eta_j^2, \quad \mu \in M^+(X \times Y). \tag{39}$$

where μ_j, η_j^1 and η_j^2 are the projections to level j of the measures μ, η^1 and η^2, respectively, to the Haar MRA.

3. Technical Results

In this section, we present a series of results that ensure the good behavior of the methodological proposal of the next section. In order to do this, we start by assuming an MK problem with cost function $c = c(x, y)$, base sets $X = Y = [0, 1]$ and measure restrictions $\eta^1 = \eta^2 = \lambda|_{[0,1]}$. In other words, we consider the problem of moving a uniform distribution to a uniform one with the minimum movement cost. Since in applications we work with discretized problems, then as a result of applying the MRA on $\mathbb{R}^2$, we have that our objective is to solve:

$$\text{MK}_j: \text{minimize} \quad \sum_{i,k} \mu_{i,k}^j c_{i,k}^j$$
$$\text{subject to:} \quad \sum_i \mu_{i,k}^j = \frac{1}{2^j}, \quad \text{for each} \quad k = 1, \dots, 2^j.$$
$$\sum_k \mu_{i,k}^j = \frac{1}{2^j}, \quad \text{for each} \quad i = 1, \dots, 2^j. \tag{40}$$
$$\sum_{i,k} \mu_{i,k}^j = 1.$$
$$\mu_{i,k}^j \geq 0, \quad \text{for each } i, k = 1, \dots, 2^j.$$

where $\mu_{i,k}^j$ is the portion of the initial value $\frac{1}{2^j}$ in the position $I_{j,i}$ of the x-axis allocated to the position $I_{j,k}$ of the y-axis. We call j-discrete unit square the grid formed by the squares $S(j, k_1, k_2)$ (see (20)), dividing the set $[0, 1] \times [0, 1]$ in $2^j \times 2^j$ blocks, in a such way that each one is identified with the point (k_1, k_2). We suppose that there is a simple solution μ for (40). That is, μ is a feasible solution such that, given $i_0, k_0 \in \{1, \dots, 2^j\}$ with $\mu_{i_0, k_0} \neq 0$, it necessarily holds that $\mu_{i_0, k} = 0$ for each $k \neq k_0$ and $\mu_{i, k_0} = 0$ for each $i \neq i_0$. Geometrically, if the measure μ is plotted as a discrete heat map in the j-discrete unit square, then no color element in the plot has another color element in its same row and column, as can be seen in Figure 2.

Definition 1. *We define a proximity criteria in the j-discrete unit square as follows: (i, k) is a neighbor of (l, m) if*

$$|i - l| = 1 \text{ or } |k - m| = 1. \tag{41}$$

In Figure 2, we plot the support of the hypothetical simple solution μ. Hence, the neighbors of the position in the middle of the cross are those that touch the yellow stripes. Then in this example, the middle point has four neighbors.
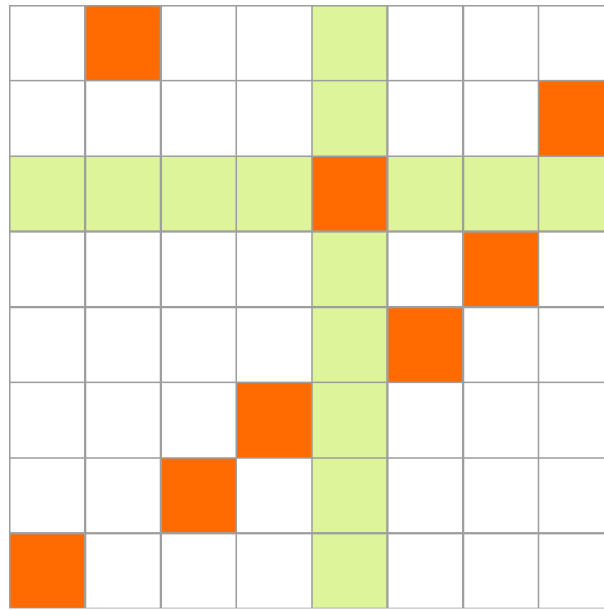
Figure 2. Support of μ and the proximity criteria.

With this in mind, we can classify the points in support(μ) as follows.

Definition 2. *We say that $(k_1, k_2) \in$ support(μ) is a border point if k_1 or k_2 equals 0 or 1; otherwise, we call it an interior point. It is clear that a border point has at least one neighbor and at most three, whereas an interior point has at least two neighbors and at most four. Hence, we can partition* support(μ) *into two sets as follows.*
The set of the points of Type I is given by

$$T_1 = \{(i, k) \in \text{support}(\mu) \mid (i, k) \text{ has minimun neighbors}\}, \tag{42}$$

and the set of the points of Type II is given by

$$T_2 = \text{support}(\mu) \setminus T_1. \tag{43}$$

Intuitively, the set T_1 is composed of well-controlled points, whereas the set T_2 has the points that admit permutations between them, since, as we will see in the next section, in the proposed algorithm they will be permuted. See Figures 3 and 4. Naturally, since μ is a feasible solution for (40), then given elements $(k_1, k_2), (m_1, m_2) \in$ support(μ) and for $i = 1, 2$ permutations σ_i over $\{k_i, m_i\}$, the measure μ^σ defined by

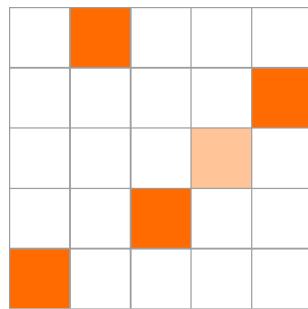$$\mu^\sigma(a, b) = \mu(\sigma_1(a), \sigma_2(b)) \tag{44}$$

is a feasible solution.
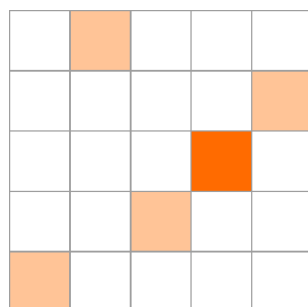
Figure 3. Classification of the points in support(μ): Type I points.

Figure 4. Classification of the points in support(μ): Type II points.

Refining Projections

In this subsection, we study the problem of improving an optimal solution μ_{j-1}^* for (40) on level $j-1 \in \mathbb{Z}$ to a feasible solution $\hat{\mu}_j$ for the next level j. Let μ_{j-1}^* be an optimal solution for level $j-1$. Then we are looking for $\hat{\mu}_j$ such that:

$$
\begin{aligned}
&(1) \quad \hat{\mu}_j \text{ is a feasible solution.} \\
&(2) \quad \mathbf{P}_{j-1}\hat{\mu}_j = \mu_{j-1}^*.
\end{aligned} \tag{45}
$$

As described in the previous section, the measure $\hat{\mu}_j \in L^2(\mathbb{R}^2)$ can be decomposed in

$$
\hat{\mu}_j = \mu_{j-1}^* + \nu_1 + \nu_2 + \nu_3 \tag{46}
$$

where

$$
\begin{aligned}
\nu_i(x,y) &= (\mathbf{Q}_j^{(i)}\hat{\mu}_j)(x,y) \\
&= \sum_{k_1}\sum_{k_2}\langle \hat{\mu}_j, \Psi_{j,k_1,k_2}^{(i)}\rangle \Psi_{j,k_1,k_2}^{(i)}(x,y).
\end{aligned} \tag{47}
$$

From the geometric point of view, the projections ν_i are formed from differences of characteristic functions, as we mentioned in Section 2.2. So we have the following result:

Lemma 1. *Let $j \in \mathbb{Z}$ and μ_{j-1}^* be an optimal solution for (40) at level $j-1$. Then for each positive measure $\hat{\mu}_j \in \mathbf{P}_j L^2(\mathbb{R}^2)$ such that $\mathbf{P}_{j-1}\hat{\mu}_j = \mu_{j-1}^*$ and $\hat{\mu}_j = \mu_{j-1}^* + \nu_1 + \nu_2 + \nu_3$, it necessarily holds that*

$$
\mu_{j-1}^*(A) = 0 \text{ implies } \nu_i(A) = 0 \tag{48}
$$

for each Lebesgue measurable set $A \subset \mathbb{R}^2$ and each $i = 1, 2, 3$. Therefore, the support of $\hat{\mu}_j$ is contained in the support of μ_{j-1}^*.

Proof. We only make the proof for the case $i = 1$, since the other two are very similar. To simplify the notation, we use the symbols μ_- and ν_- as measures or functions in the respective subspace of $L^2(\mathbb{R}^2)$. Since when setting a level $j \in \mathbb{Z}$ all the measures in question are constant in pairs of rectangles dividing $S(j, k_1, k_2)$, as we prove in Section 2.2, then it is enough to prove that (65) is valid on this rectangles. Let

$$
\begin{aligned}
A_1 &= A_1(j, k_1', k_2') \\
&= [2^{-j}k_1', 2^{-j}(k_1'+1)) \times [2^{-j-1}2k_2', 2^{-j-1}(2k_2'+1)).
\end{aligned} \tag{49}
$$

and

$$
\begin{aligned}
A_2 &= A_2(j, k_1', k_2') \\
&= [2^{-j}k_1', 2^{-j}(k_1'+1)) \times [2^{-j-1}2k_2'+1, 2^{-j-1}(2k_2'+2)),
\end{aligned} \tag{50}
$$

as in (22). Then for $l = 1, 2$, we have that

$$
\begin{aligned}
\hat{\mu}_j(A_l) &= \mu_{j-1}^*(A_l) + \sum_{i=1}^{3} \nu_i(A_l) \\
&= \mathbb{E}[\mu_{j-1}^* \chi_{A_l}] + \sum_{i=1}^{3} \mathbb{E}[\nu_i \chi_{A_l}].
\end{aligned}
\tag{51}
$$

Now we will calculate each one of the expectations separately. By (17), (18) and (25), we have that

$$
\mu_{j-1}^*(x, y) = \sum_{k_1} \sum_{k_2} \langle \mu_{j-1}^*, \Phi_{j,k_1,k_2} \rangle \Phi_{j,k_1,k_2}(x, y)
\tag{52}
$$

and

$$
\begin{aligned}
\nu_1(x, y) &= \sum_{k_1} \sum_{k_2} \langle \nu_1, \Psi_{j,k_1,k_2}^{(1)} \rangle \Psi_{j,k_1,k_2}^{(1)}(x, y) \\
&= \sum_{k_1} \sum_{k_2} \langle \nu_1, \Psi_{j,k_1,k_2}^{(1)} \rangle [\chi_{A_1(j,k_1,k_2)}(x, y) - \chi_{A_2(j,k_1,k_2)}(x, y)].
\end{aligned}
\tag{53}
$$

Then for $l = 1$ and by (15), (16) and (22), we have that

$$
\begin{aligned}
\mathbb{E}[\mu_{j-1}^* \chi_{A_1}] &= \mathbb{E}[\chi_{A_1} \sum_{k_1} \sum_{k_2} \langle \mu_{j-1}^*, \Phi_{j,k_1,k_2} \rangle \Phi_{j,k_1,k_2}] \\
&= \langle \mu_{j-1}^*, \Phi_{j,k_1',k_2'} \rangle \mathbb{E}[\chi_{A_1} \Phi_{j,k_1',k_2'}]] \\
&= d \mathbb{E}[\chi_{A_1}] \\
&= d \frac{1}{2^{2j+1}}
\end{aligned}
\tag{54}
$$

where $d = \langle \mu_{j-1}^*, \Phi_{j,k_1',k_2'} \rangle$ and

$$
\begin{aligned}
\mathbb{E}[\nu_1 \chi_{A_1}] &= \mathbb{E}[\chi_{A_1} \sum_{k_1} \sum_{k_2} \langle \nu_1, \Psi_{j,k_1,k_2}^{(1)} \rangle (\chi_{A_1(j,k_1,k_2)} - \chi_{A_2(j,k_1,k_2)})] \\
&= \langle \nu_1, \Psi_{j,k_1',k_2'}^{(1)} \rangle \mathbb{E} \chi_{A_1} \\
&= c \mathbb{E}[\chi_{A_1}] \\
&= c \frac{1}{2^{2j+1}}
\end{aligned}
\tag{55}
$$

where $c = \langle \nu_1, \Psi_{j,k_1',k_2'}^{(1)} \rangle$. Similarly, but using (23) and (24), we can prove that

$$
\mathbb{E}[\nu_2 \chi_{A_1}] = 0
\tag{56}
$$

and

$$
\mathbb{E}[\nu_3 \chi_{A_1}] = 0.
\tag{57}
$$

Then

$$
\begin{aligned}
\hat{\mu}_j(A_1) &= d \frac{1}{2^{2j+1}} + c \frac{1}{2^{2j+1}} \\
&= \mu_{j-1}^*(A_1) + \nu(A_1)
\end{aligned}
\tag{58}
$$

Now, we will make an analogous argument for the case $l = 2$. Hence,

$$
\begin{aligned}
\mathbb{E}[\mu_{j-1}^* \chi_{A_2}] &= \mathbb{E}[\chi_{A_2} \sum_{k_1} \sum_{k_2} \langle \mu_{j-1}^*, \Phi_{j,k_1,k_2} \rangle \Phi_{j,k_1,k_2}] \\
&= \langle \mu_{j-1}^*, \Phi_{j,k_1,k_2} \rangle \mathbb{E}[\chi_{A_2} \Phi_{j,k_1',k_2'}]] \\
&= d \mathbb{E}[\chi_{A_2}] \\
&= d \frac{1}{2^{2j+1}}
\end{aligned}
\tag{59}
$$

and

$$
\begin{aligned}
\mathbb{E}[\nu_1 \chi_{A_2}] &= \mathbb{E}[\chi_{A_2} \sum_{k_1} \sum_{k_2} \langle \nu_1, \Psi_{j,k_1,k_2}^{(1)} \rangle (\chi_{A_1(j,k_1,k_2)} - \chi_{A_2(j,k_1,k_2)})] \\
&= -\langle \nu_1, \Psi_{j,k_1',k_2'}^{(1)} \rangle \mathbb{E}\chi_{A_2} \\
&= -c \mathbb{E}[\chi_{A_2}] \\
&= -c \frac{1}{2^{2j+1}}.
\end{aligned}
\tag{60}
$$

Thus,

$$
\begin{aligned}
\hat{\mu}_j(A_2) &= d \frac{1}{2^{2j+1}} - c \frac{1}{2^{2j+1}} \\
&= \mu_{j-1}^*(A_2) + \nu(A_2)
\end{aligned}
\tag{61}
$$

and by (54), (55), (59) and (60), we have that

$$
\mu_{j-1}^*(A_1) = \mu_{j-1}^*(A_2)
\tag{62}
$$

and

$$
\nu_1(A_1) = -\nu_1(A_2).
\tag{63}
$$

Therefore, it follows from (58), (61), (62), (63) and the fact that $\hat{\mu}_j$ is a positive measure, that

$$
\begin{aligned}
|\nu_1(A_1)| &= |\nu_1(A_2)| \\
&\leq |\mu_{j-1}^*(A_1)| \\
&= |\mu_{j-1}^*(A_2)|
\end{aligned}
\tag{64}
$$

From which we conclude that $\mu_{j-1}^*(A) = 0$ implies $\nu_1(A) = 0$. Similarly, it can be shown that $\mu_{j-1}^*(A) = 0$ implies $\nu_i(A) = 0$ for $i = 2, 3$; for this, analogous proofs are carried out, with the difference being that for $i = 2$, the sets to be considered are B_1 and B_2 as in (23), whereas $C_1 \cup C_4$ and $C_2 \cup C_3$ as in (24) are the respective sets when $i = 3$. □

We have proved that if it is intended to go back to the preimage of the projection of the approximation operator **P** from a level $j - 1$ to a level j, the support of the level $j - 1$ delimits that of the j level. Now, we will prove that for every measure $\hat{\mu}_j$ that satisfies (45) and (46), it necessarily holds that $\nu_1 = \nu_2 = 0$.

Lemma 2. *Let $j \in \mathbb{Z}$ and μ_{j-1}^* be an optimal simple solution for (40) at level $j - 1$. Then for each feasible solution $\hat{\mu}_j \in \mathbf{P}_j L^2(\mathbb{R}^2)$ to (40) at level j such that $\mathbf{P}_{j-1}\hat{\mu}_j = \mu_{j-1}^*$ and $\hat{\mu}_j = \mu_{j-1}^* + \nu_1 + \nu_2 + \nu_3$. It necessarily holds that*

$$
\nu_1 = \nu_2 = 0.
\tag{65}
$$

Proof. In order to perform this proof, we use the restrictions of the MK problem (40), which in turn, are related to the marginal measures. Therefore, we will only complete the proof for one of the projections, since the other is analogous. From the linearity of the Radon–Nikodym derivative, it follows that

$$
\begin{aligned}
\eta_j^1 &= \Pi_1 \hat{\mu}_j \\
&= \Pi(\mu_{j-1}^* + \nu_1 + \nu_2 + \nu_3) \\
&= \Pi \mu_{j-1}^* + \Pi \nu_1 + \Pi \nu_2 + \Pi \nu_3.
\end{aligned}
\tag{66}
$$

Let $k' \in \mathbb{Z}$ and $I_{j+1,2k'} = [2^{-j-1}2k', 2^{-j-1}(2k'+1))$. Thus, by (39) and (66), we have that

$$
\begin{aligned}
\Pi_1 \hat{\mu}_j(I_{j+1,2k'}) &= \eta_j^1(I_{j+1,2k'}) \\
&= \Pi_1 \mu_{j-1}^*(I_{j+1,2k'}) \\
&= \alpha.
\end{aligned}
\tag{67}
$$

That is, we are evaluating feasible solutions on rectangles whose height is half the size of the squares with which they are discretized at the j level of the Haar MRA. Now, we will develop in detail (66) evaluated on $I_{j,k'}$. Since μ_{j-1}^* is a simple solution, we call $l' \in \mathbb{Z}$ the only number such that $\mu_{j-1}^+(S(j,k',l')) > 0$, where $S(j,k_1,k_2)$ is defined as in (20). With the aim of simplifying the notation, we define $\mathcal{I} = I_{j+1,2k'} \times \mathbb{R}$. By (16) and (25), we have that

$$
\begin{aligned}
\Pi_1 \nu_1(I_{j+1,2k'}) &= \nu_i(\mathcal{I}) \\
&= \mathbb{E}[\chi_{\mathcal{I}} \nu_1] \\
&= \mathbb{E}\Big[\chi_{\mathcal{I}} \sum_{k_1} \sum_{k_2} \langle \nu_1, \Psi_{j,k_1,k_2}^{(1)} \rangle (\chi_{A_1(j,k_1,k_2)} - \chi_{A_2(j,k_1,k_2)})\Big] \\
&= \mathbb{E}\Big[\chi_{\mathcal{I}} \sum_{k_2} \langle \nu_1, \Psi_{j,k',k_2}^{(1)} \rangle (\chi_{A_1(j,k',k_2)} - \chi_{A_2(j,k',k_2)})\Big] \\
&= \mathbb{E}\Big[\chi_{\mathcal{I}} \langle \nu_1, \Psi_{j,k',l'}^{(1)} \rangle (\chi_{A_1(j,k',l')} - \chi_{A_2(j,k',l')})\Big]
\end{aligned}
\tag{68}
$$

By the way $\mathcal{I}$ was defined, necessarily in the last equality it must be fulfilled that one of the terms in the expectation is equal to 0. Hence, it is fulfilled that

$$
\begin{aligned}
\Pi_1 \nu_1(I_{j+1,2k'}) &= \langle \nu_1, \Psi_{j,k',l'}^{(1)} \rangle \mathbb{E}[\chi_{\mathcal{I}}] \\
&= \frac{1}{2^{2j+1}} \langle \nu_1, \Psi_{j,k',l'}^{(1)} \rangle.
\end{aligned}
\tag{69}
$$

By a similar argument, it can be proved that

$$
\Pi \nu_2(I_{j+1,2k'}) = 0
\tag{70}
$$

and

$$
\Pi \nu_2(I_{j+1,2k'}) = 0.
\tag{71}
$$

Then from (66) to (71), it follows that

$$
\alpha = \alpha + \frac{1}{2^{2j+1}} \langle \nu_1, \Psi_{j,k',l'}^{(1)} \rangle + 0 + 0.
\tag{72}
$$

Hence,

$$
\langle \nu_1, \Psi_{j,k',l'}^{(1)} \rangle = 0.
\tag{73}
$$

Therefore, $\nu_1 = 0$. In a similar way, we can prove that $\nu_2 = 0$. □

Suppose we have a simple optimal solution μ^*_{j-1} for the MK problem discretized through the Haar MRA at level $j-1 \in \mathbb{Z}$ and that we are interested in refining that solution to the next level j. By Lemmas 1 and 2, any $\hat{\mu}_j$ that satisfies (45) has its support contained in the support of μ^*_{j-1} and has components $\nu_1 = \nu_2 = 0$. Then the problem of constructing a feasible solution $\hat{\mu}_j$ such that it refines μ^*_{j-1} is reduced to construct

$$
\begin{aligned}
\hat{\mu}_j(x,y) &= \mu^*_{j-1}(x,y) + \nu_3(x,y) \\
&= \mu^*_{j-1}(x,y) + \sum_{k_1}\sum_{k_2} \langle \hat{\mu}_j, \Psi^{(i)}_{j,k_1,k_2} \rangle \Psi^{(i)}_{j,k_1,k_2}(x,y),
\end{aligned}
\tag{74}
$$

which is equivalent to chosing for each $k_1, k_2 \in \mathbb{Z}$ a value $\nu^3_{k_1,k_2}$ such that

$$
\hat{\mu}^j_{k_1,k_2} = \mu^{*,j-1}_{k_1,k_2} + \nu^3_{k_1,k_2}.
\tag{75}
$$

By Lemma 1 for each $k_1, k_2 \in \mathbb{Z}$, it is fulfilled that

$$
|\nu^3_{k_1,k_2}| \le \mu^{*,j-1}_{k_1,k_2}.
\tag{76}
$$

Therefore, the choice of $\nu^3_{k_1,k_2}$ is restricted to a compact collection, and since $\hat{\mu}_j$ is a solution of the linear program (40), then

$$
|\nu^3_{k_1,k_2}| = \mu^{*,j-1}_{k_1,k_2}.
\tag{77}
$$

Thus, the sign of $\nu^3_{k_1,k_2}$ must be such that it minimizes $\nu^3_{k_1,k_2} c^j_{k_1,k_2}$. That is,

$$
\nu^3_{k_1,k_2} c^j_{k_1,k_2} \le 0
\tag{78}
$$

4. Methodological Proposal

In this section, we show through examples a process that builds solutions to the MK problem with a reduced number of operations. First, we consider the MK problem with cost function $c : [0,1] \times [0,1] \to \mathbb{R}$ defined by

$$
c(x,y) = 4x^2y - xy^2
\tag{79}
$$

and homogeneous restrictions ν_1, ν_2 over $[0,1]$. So that the algorithm can be graphically appreciated, we take a small level of discretization, namely $j = 6$. Thus, in the Haar MRA over $\mathbb{R}^2$ at level $j = 6$, the cost function has the form shown in Figure 5, which can be stored in a vector of size $2^{2j} = 2^{2(6)} = 2^{12}$.

Now, we apply the filtering process to the cost function at level $j = 6$, which results in four functions

$$
c_6(x,y) = c_5(x,y) + \Psi^{(1)}(x,y) + \Psi^{(2)}(x,y) + \Psi^{(3)}(x,y);
\tag{80}
$$

see Figures 6–9.
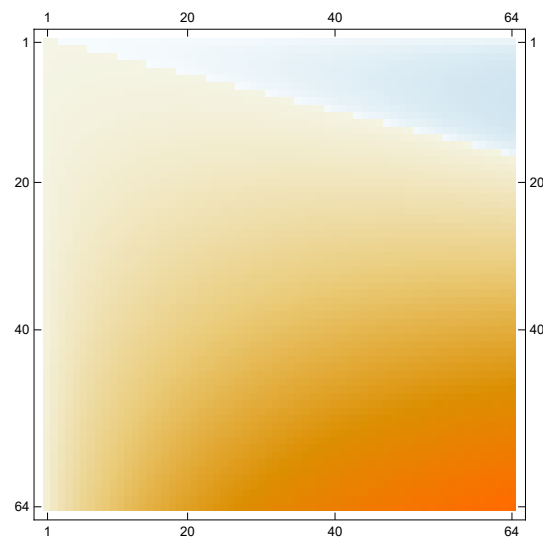
Figure 5. Step 1. Discretization of the cost function to the level j, which is denoted by c_j. In particular, the cost function is $c(x, y) = 4x^2y - xy^2$ for lever $j = 6$.

Step 2. Filtering the original discrete function using the high-pass filter, which yield three discrete functions denoted by Ψ_1, Ψ_2 and Ψ_3, that functions correspond to Figures 7, 8 and 9, respectively, each describing local changes in the original discrete function. It is then low-pass filtered to produce the approximate discrete function c_5, which is given by Figure 6.
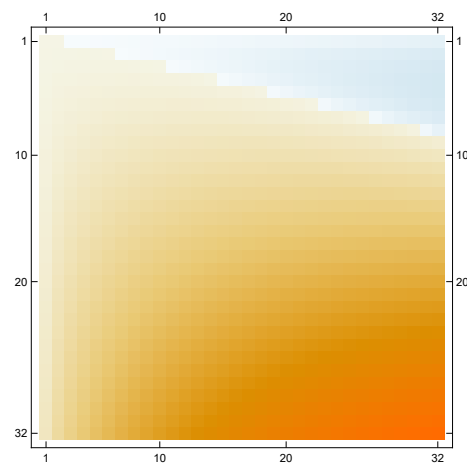
Figure 6. $c_5(x, y)$.
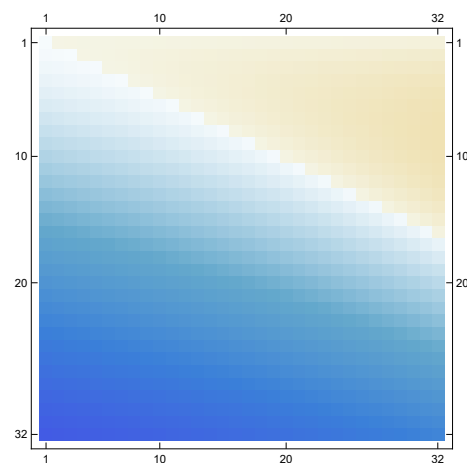
Figure 7. $\Psi^{(1)}(x, y)$.
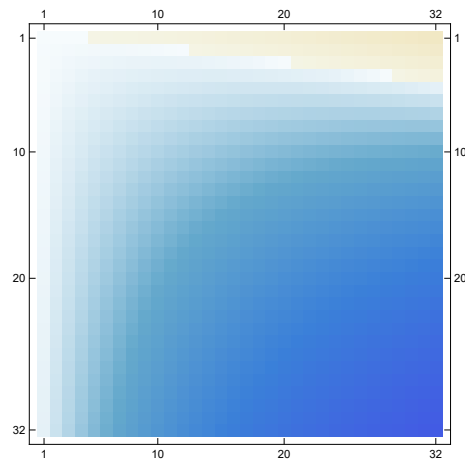
Figure 8. $\Psi^{(2)}(x,y)$.
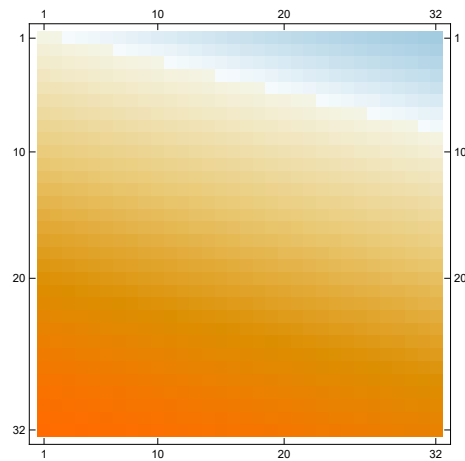
Figure 9. $\Psi^{(3)}(x,y)$.

We then solve the MK problem for the level $j - 1 = 5$. That is, we find a measure $\mu_{j-1}^* = \mu_5^*$ that is an optimal solution for the MK problem with cost function c_5. Such data can be stored in a vector of size $2^{2j-2} = 2^{10}$; see Figure 10. For each entry k, the formal application that plots this vector in a square is defined by $k \rightarrow (k_1 + 1, k_2 + 1)$, where $0 \leq k_1, k_2 < 2^{n-1}$ and $k = k_1 \cdot 2^{j-1} + k_2$.
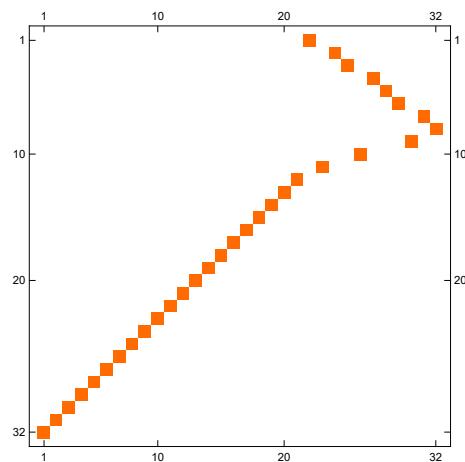
Figure 10. Step 3. We obtain a solution μ_5^* for MK^5 associated with the cost function c_5 given in Figure 6.

Since the measure $\mu_{j-1}^* = \mu_5^*$ is an optimal simple solution for the MK problem, then we can represent its support in a simple way, as we show below:

$$\text{support}(\mu_{j-1}^*) = \bigcup_{k_1,k_2} S_{k_1,k_2} \tag{81}$$

where $S_{k_1,k_2} = S(j, k_1, k_2)$ is the square in (20). Next, we split each block $S_{k_1,k_2} \subset \text{support}(\mu_{j-1}^*)$ into four parts as in (24); see Figure 11.
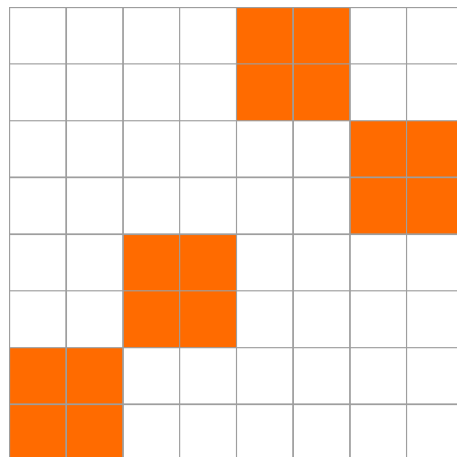
Figure 11. Division of the components of support(μ_{j-1}^*) into four parts.

From the technical point of view, in the discretization at level $j - 1 = 5$, we have a grid of $2^{j-1} \times 2^{j-1} = 32 \times 32$ squares that we call S_{k_1,k_2} and identify with the points (k_1, k_2). Thus, we refine to a grid of 64×64, splitting each square into four, which in the new grid are determined by points

$$(2k_1 - 1, 2k_2), (2k_1 - 1, 2k_2 - 1), (2k_1, 2k_2 - 1) \text{ and } (2k_1, 2k_2); \tag{82}$$

see Figure 12.

Figure 12. Refinement of grid from level $j - 1$ to j of discretization.

As we prove in Lemma 1, any feasible solution $\hat{\mu}_j = \hat{\mu}_6$ that refines μ_5^* has its support contained in support(μ_5^*). Therefore, we must only deal with the region delimited by the support of μ_5^*. By Lemma 2, in order to construct the solution $\hat{\mu}_6$, we only need to determine the values n_{k_1,k_2}^3 corresponding to the coefficients of the wavelet part ν_3; however, by (77) and (78), those values are well determined and satisfy that when added with the scaling part μ_{k_1,k_2}^*, the result is a scalar multiple of a characteristic function. For example if the

square S_{k_1,k_2} has scaling part coefficient $\mu^*_{k_1,k_2} = c$, then we choose $\nu^3_{k_1,k_2} = -c$. Hence, by (21) and (25), we have that

$$
\begin{aligned}
\hat{\mu}^6_{k_1,k_2}(x,y) &= \mu^*_{k_1,k_2}(x,y) + \nu^3_{k_1,k_2}(x,y) \\
&= c\chi_S(x,y) - c[\chi_{C_1 \cup C_4}(x,y) - \chi_{C_2 \cup C_3}(x,y)] \\
&= 2c\chi_{C_2 \cup C_3}(x,y).
\end{aligned}
\tag{83}
$$

Thus, from an operational point of view, we only need to chose between two options of supporting each division of square S_{k_1,k_2}, as we illustrate in Figures 13 and 14.
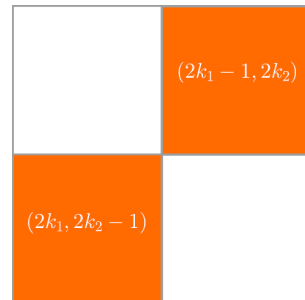
Figure 13. Supports for refinement of the element S_{k_1,k_2} corresponding to the level j : Option I.
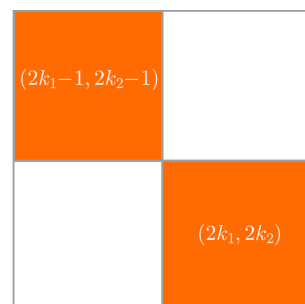
Figure 14. Supports for refinement of the element S_{k_1,k_2} corresponding to the level j: Option II.

This coincides with our geometric intuition. Hence, the resulting feasible solution $\hat{\mu}_6$ has its support contained in support(μ^*_5), and its weight within each square S_{k_1,k_2} is presented in a diagonal within that block; see Figure 15. Finally, we can improve $\hat{\mu}_6$ by observing the way the filtering process acts. To do this, we apply the proximity criteria (41) and split the support of $\hat{\mu}_6$ into points of Type I and II. In Figure 16, we identify the points of Type I and II of solution μ^*_5, whereas in Figure 15, we do the same but for $\hat{\mu}_6$.

Intuitively, the division of the support into points of Type I and II allows us to classify the points so that they have an identity function form and, consequently, that come from the discretization of a continuous function—points of Type I—and in points that come from the discretization of a discontinuous function—points of Type II. Thus, the points of Type I are located in such a way that they generate a desired solution, and therefore it is not convenient to move them, whereas Type II points are free to be changed as this does not lead to the destruction of a continuous structure in the solution. As we mentioned in the previous section, each permutation of rows or columns of one weighted element S_{k_1,k_2} with another $S_{k'_1,k'_2}$ constructs a feasible solution; see (44). Thus, as a heuristic technique to improve the solution, we check the values $\langle c_6, \mu^\sigma_6 \rangle$ associated with each solution μ^σ_6 obtained by permuting rows or columns of points of Type II of the solution $\hat{\mu}_6$. We call $\mu^{\sigma^*}_6$ the solution for which its permutation gives it the best performance. See Figure 17.
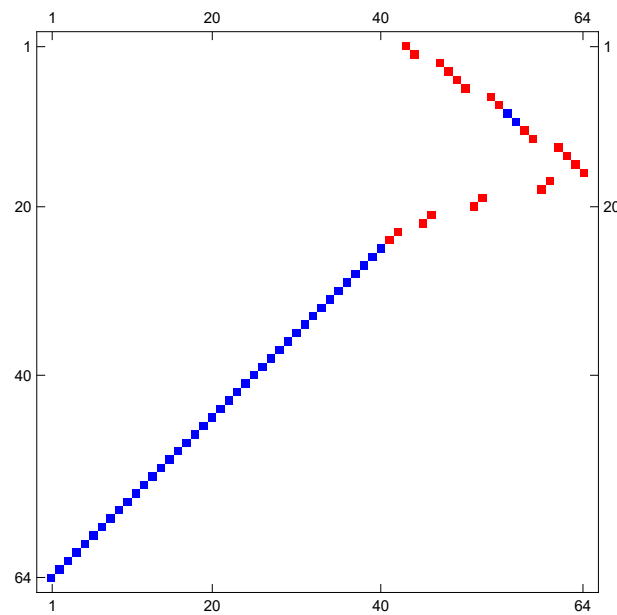
Figure 15. Step 5. Using the solution μ_5^* which is given by Figure 10, the information of the high-pass components (Figures 7, 8 and 9) and Lemma 1, we obtain a feasible solution for Level 6, which is denoted by $\hat{\mu}_6$.
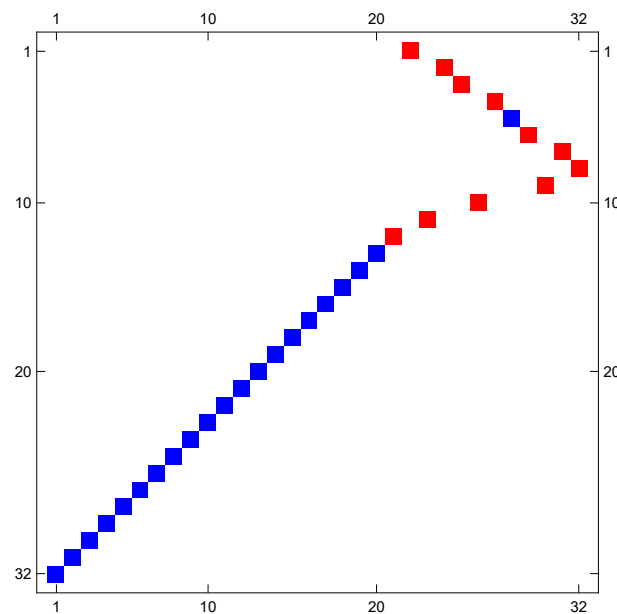
Figure 16. Step 4. We classify the points of the support of the solution μ_5^* by proximity criteria as points of Type I ■ or Type II ■ (the measure μ_5^* corresponds to Figure 10).

Finally, we present Table 1 that compare the solutions of the MK problem, in which MK^5 is the value associated with the optimal solution μ_5^* at level of discretization $j - 1 = 5$, MK$_6$ is the value associated with an optimal solution μ_6^* at level of discretization $j = 6$, and MK$_6^{\sigma^*}$ is the value associated with the solution $\mu_6^{\sigma^*}$ obtained by the heuristic method described in the previous paragraph.
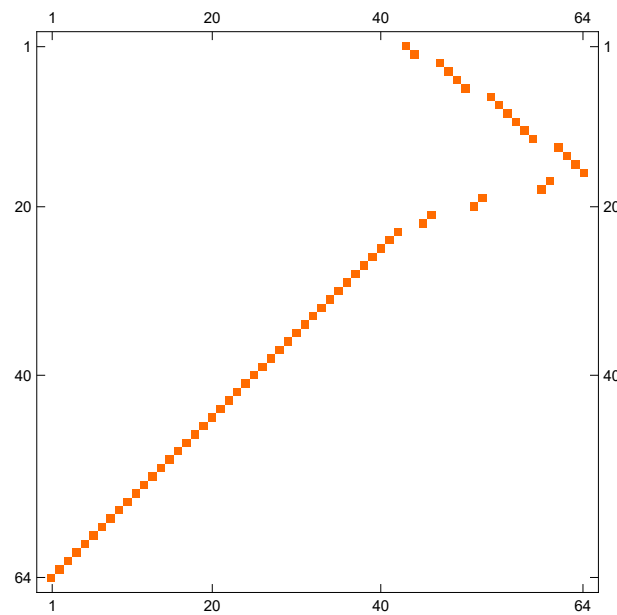
Figure 17. Step 6. Classification of the points of μ_5^* induces classification of the points in $\hat{\mu}_6$ by contention in the support. Over the points of Type I of the solution $\hat{\mu}_6$, we do not move those points. For the points of Type II, we apply a permutation to the solution over the two points that improve the solution and repeat the process with the rest of the points.

Table 1. Comparison of the values corresponding to MK_5, $MK_6^{\sigma^*}$ and MK_6.

MK_5	$MK_6^{\sigma^*}$	MK_6
0.24785	0.247738	0.247726

5. Other Examples of This Methodology

We conclude this work with a series of examples in which we apply the proposed methodology. Each of them is divided into the six-step algorithm introduced in the previous section and corresponds to a classical example existing in the literature.

5.1. Example with Cost Function $c(x,y) = x^2y - xy^2$

Let MK be the MK problem with cost function $c(x,y) = x^2y - xy^2$ and uniform restrictions η_1 and η_2 over the unitary interval $[0,1]$. In order to be more didactic, we consider the level of discretization $j = 6$. Next, we present the reduced algorithm.

Step 1. Discretize the cost function at level j—that is, over a grid of $2^j \times 2^j$. Figure 18.
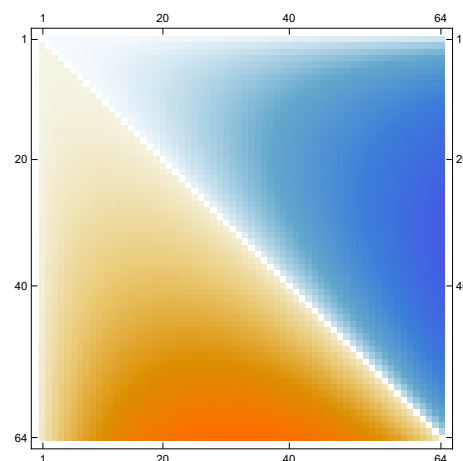
Figure 18. Discretization of the cost function $c(x,y) = x^2y - xy^2$ at level $j = 6$.

Step 2. Apply the filtering process to the cost function at level $j-1$, obtaining the filtered cost function c_5, which is plotted on a grid of $2^{j-1} \times 2^{j-1}$. Figure 19.
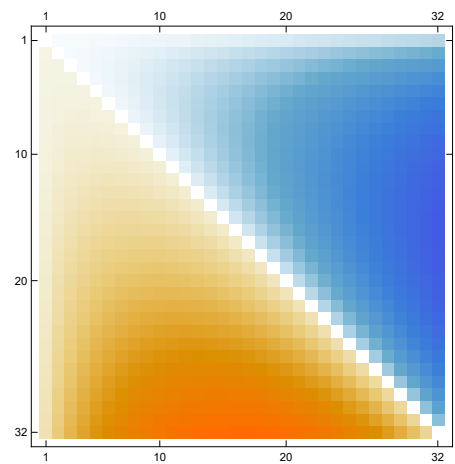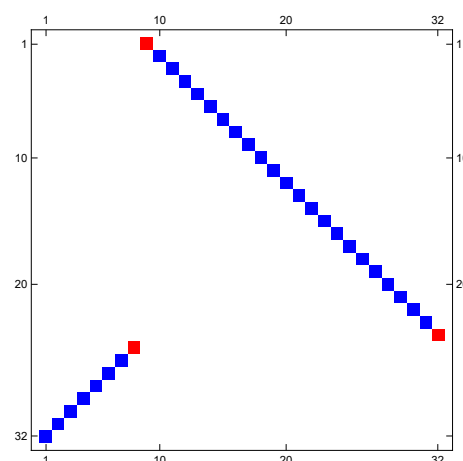

Figure 19. Filtering of the cost function c at level $j-1 = 5$.

Step 3. Find an optimal simple solution μ_{j-1}^* for the discretized MK_j problem. That is, solve for the cost function c_5 and obtain an optimal simple solution μ_5^*. Figure 20.
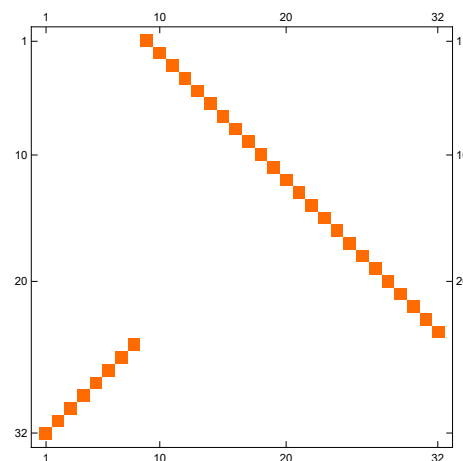
Figure 20. MK solution for the filtered function c at level $j-1 = 5$.

Step 4. Apply the proximity criteria to the support of μ_{j-1}^*. Figure 21.

Figure 21. Classification of points of support(μ_5^*) into Type I ■ and Type II ■.

Step 5. Refine the optimal simple solution μ_{j-1}^* to a feasible solution $\hat{\mu}_j$, dividing each weighted square S_{k_1,k_2} at level $j-1$ into four squares at level j (see (82)) and place mass according to the criteria (83). Figure 22.
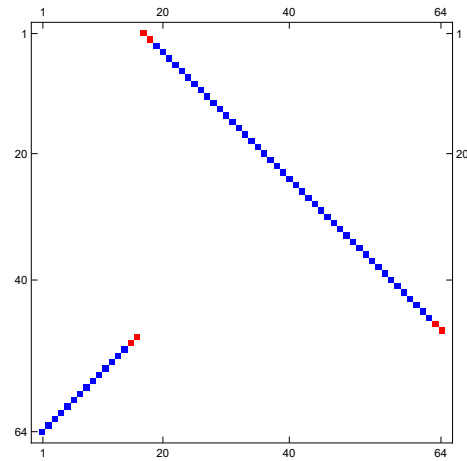
Figure 22. Solution $\hat{\mu}_6$ from refinement of μ_5^*.

Step 6. Permute the rows and columns of the points of Type II in support$(\hat{\mu}_j)$ using (44) to construct feasible solutions μ_j^σ and chose the one that has better performance. Figure 23.
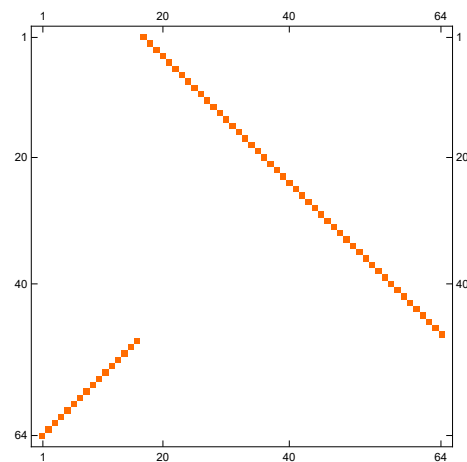
Figure 23. Final result $\mu_6^{\sigma*}$.

The following Table 2 contains the comparison of the proposed methodology with the two immediate levels of resolution.

Table 2. Comparison of the values corresponding to MK$_5$, MK$_6^{\sigma*}$ and MK$_6$.

MK$_5$	MK$_6^{\sigma*}$	MK$_6$
-0.0350952	-0.035141	-0.035141

5.2. Example with Cost Function $c(x,y) = xy$

Let MK be the MK problem with cost function $c(x,y) = xy$ and uniform restrictions η_1 and η_2 over the unitary interval $[0,1]$. In order to be more didactic, we consider the level of discretization $j = 6$. Next, we present the reduced algorithm.

**Step 1**. Discretize the cost function at level j—that is, over a grid of $2^j \times 2^j$. Figure 24.
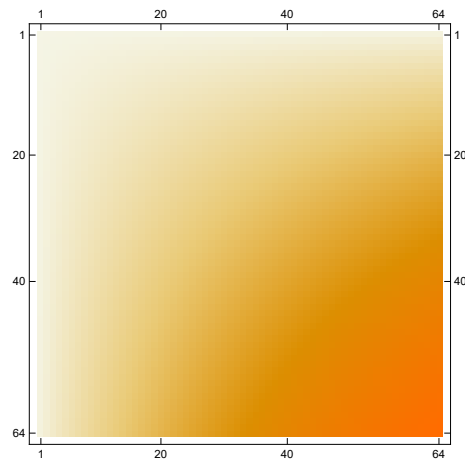
Figure 24. Discretization of the cost function $c(x, y) = xy$ at level $j = 6$.

Step 2. Apply the filtering process to the cost function at level $j - 1$, obtaining the filtered cost function c_5, which is plotted on a grid of $2^{j-1} \times 2^{j-1}$. Figure 25.
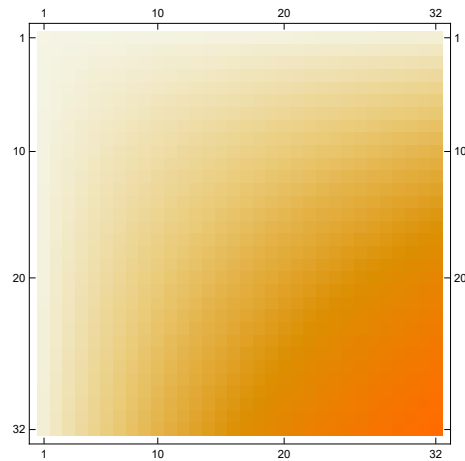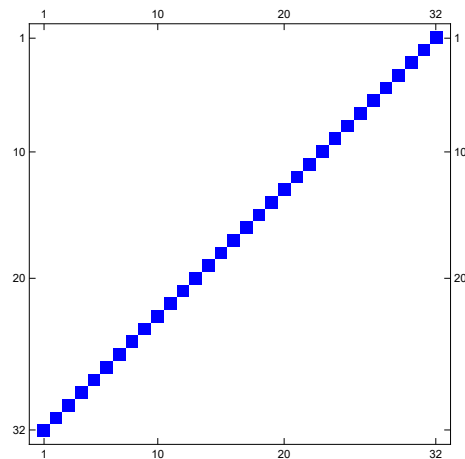

Figure 25. Apply the filter to the cost function c at level $j - 1 = 5$.

Step 3. Find an optimal simple solution μ_{j-1}^* for the discretized MK_j problem. That is, solve for the cost function c_5 and obtain an optimal simple solution μ_5^*. Figure 26.
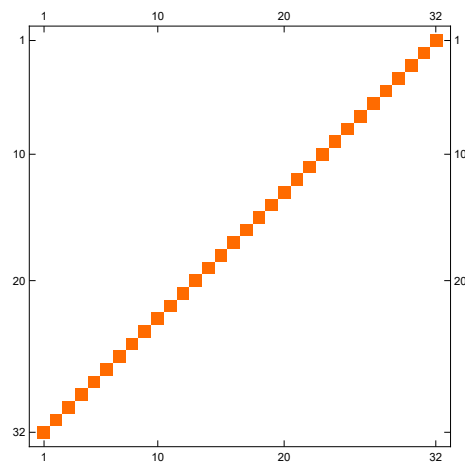

Figure 26. Solution of the MK for the function c at level $j - 1 = 5$.

Step 4. Apply the proximity criteria to the support of μ_{j-1}^*. Figure 27.

Figure 27. In this example there are only Type I points ■.

Step 5. Refine the optimal simple solution μ_{j-1}^* to a feasible solution $\hat{\mu}_j$, dividing each weighted square S_{k_1,k_2} at level $j-1$ into four squares at level j (see (82)) and placing mass according to the criteria (83). Figure 28.
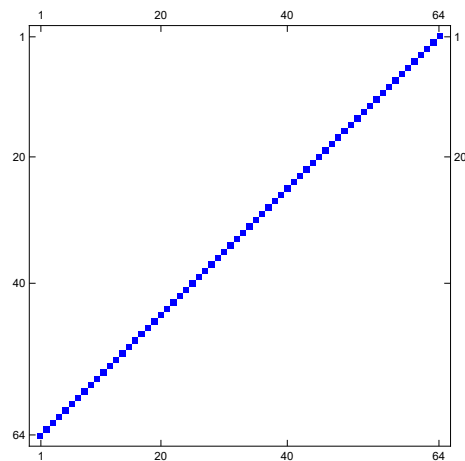
Figure 28. Solution $\hat{\mu}_6$ from refinement of μ_5^*.

Step 6. Permute the rows and columns of the points of Type II in support$(\hat{\mu}_j)$ using (44) to construct feasible solutions μ_j^σ and chose the one that has better performance. Figure 29.
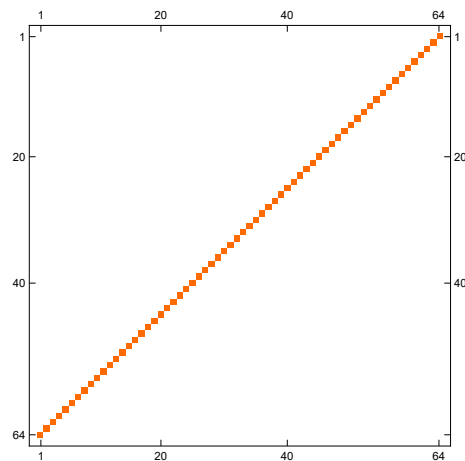
Figure 29. Final result $\mu_6^{\sigma^*}$.

The following Table 3 contains the comparison of the proposed methodology with the two immediate levels of resolution.

Table 3. Comparison of the values corresponding to MK_5, $MK_6^{\sigma^*}$ and MK_6.

MK_5	$MK_6^{\sigma^*}$	MK_6
0.166748	0.166687	0.166687

5.3. Example with Cost Function $c(x, y) = (2y - x - 1)^2(2y - x)^2$

We take the MK problem with cost function $c(x, y) = (2y - x - 1)^2(2y - x)^2$ and homogeneous restrictions over the unitary interval. Again, we consider the level of discretization $j = 6$.

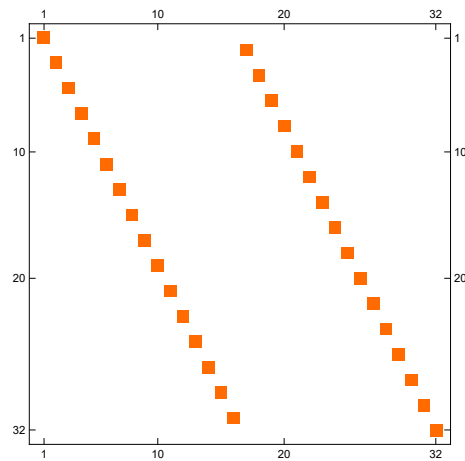**Step 1**. Discretize the cost function at level j. Figure 30.
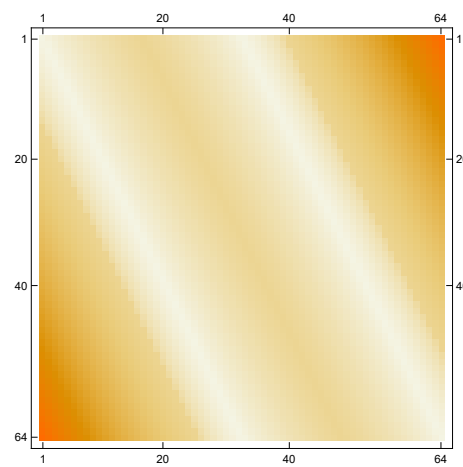
Figure 30. Discretization of cost function $c(x, y) = (2y - x - 1)^2(2y - x)^2$ for level $j = 6$.

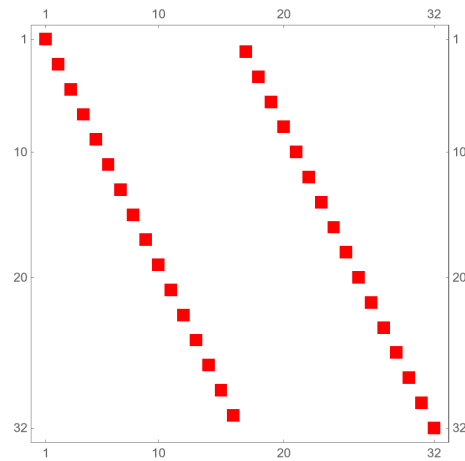**Step 2**. Apply the filtering process at level $j - 1$ to the cost function. Figure 31.
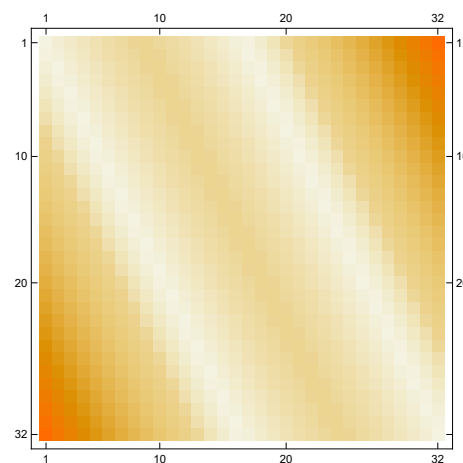
Figure 31. Filtered cost function for $j = 5$.

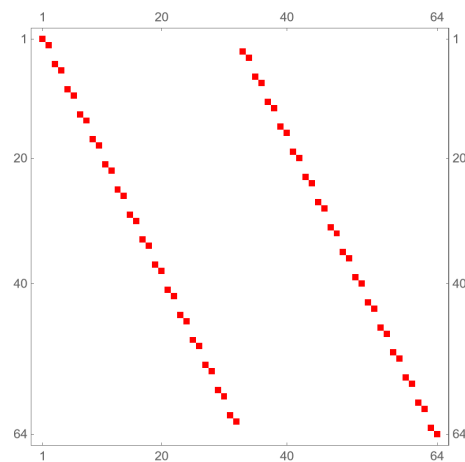**Step 3**. Find an optimal simple solution μ_{j-1}^* for the MK_{j-1} problem. Figure 32.

Figure 32. Solution of the MK_5 problem.

Step 4. Apply the proximity criteria to the support of μ^*_{j-1}. Figure 33.

Figure 33. In this example, there are only Type II points ■.

Step 5. Refine the solution μ^*_{j-1} to a feasible solution $\hat{\mu}_6$ applying the criteria (82) and (83). Figure 34.

Figure 34. Feasible solution $\hat{\mu}_6$ from refinement of μ^*_5.

Step 6. Permute points of Type II of $\hat{\mu}_j$ using (44) to construct feasible solutions μ^σ_j and chose which has better performance. Figure 35.
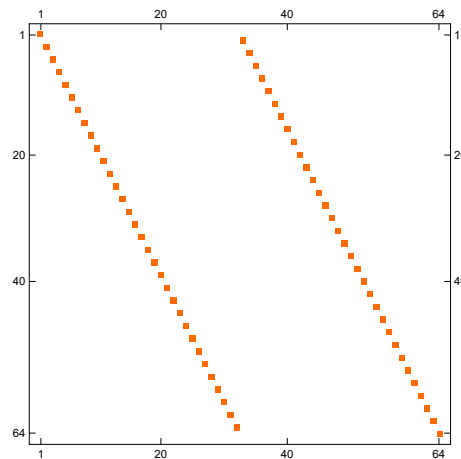
Figure 35. Final result $\mu_6^{\sigma^*}$.

The Table 4 summarizes the results obtained.

Table 4. Comparison of values MK_5, $MK_6^{\sigma^*}$ and MK_6.

$\mathbf{MK_5}$	$\mathbf{MK_6^{\sigma^*}}$	$\mathbf{MK_6}$
0.000236571	0.0000600852	0.0000597889

6. Conclusions and Future Work

Note that with the methodology of [3,4,9], the authors obtain a solution of MK^j. For this, they need to resolve a transport problem with 4^j variables. We call this methodology an exhaustive method. For our methodology, in Step 3, we need to resolve a transport problem with 4^{j-1} variables, and the other steps of the methodology are methods of classification, ordering and filtering; with 2^j data for classification and ordering and 4^j data for filtering, it is clear that this method requires fewer operations to resolve transport problems.

In summary, we have the following table comparing the results of solving the examples more often used in the literature with our methodology versus the exhaustive method (using all variables).

Cost Function	MK^6	MK_6^{σ}	**Error**
xy	0.166687	0.166687	
$x^2y - xy^2$	-0.035141	-0.035141	1.52588×10^{-5}
$4x^2y - xy^2$	0.247726	0.247738	5.31762×10^{-5}
$(2y - x - 1)^2(2y - x)^2$	0.0000600947	0.0000600947	
Cost Function	MK^7	MK_7^{σ}	**Error**
xy	0.166672	0.166672	
$x^2y - xy^2$	-0.0351524	-0.0351524	3.8147×10^{-6}
$4x^2y - xy^2$	0.247695	0.247698	1.24066×10^{-5}
$(2y - x - 1)^2(2y - x)^2$	0.0000600852	0.0000600852	
Cost Function	MK^8	MK_8^{σ}	**Error**
xy	0.166668	0.166668	
$x^2y - xy^2$	-0.0351553	-0.0351553	9.53674×10^{-7}
$4x^2y - xy^2$	0.247688	0.247688	3.21631×10^{-6}
$(2y - x - 1)^2(2y - x)^2$			

Note that our method always improves the solution of the level $j - 1$ and for some examples give an exact solution; we use Mathematica$^{©}$ and basic computer equipment

for programming this methodology, and maybe we can improve the results with software focused on numerical calculus and better computer equipment. It is also important to mention that the methodology presented in this work has some weaknesses. In our computational experiments, we noticed that if we did not start with a sufficient amount of information, then the methodology tended to give very distant results. In other words, if the initial level of discretization was not fine enough, then because the algorithm lowers the resolution level when executed, such loss of information generates poor performance. However, when starting with an adequate level of discretization, experimentally it can be observed that the distribution of the solutions for the discretized problems, as well as the respective optimal values, have stable behavior with a clear trend. The question that arises naturally is: "In practice, what are the parameters that determine good or bad behavior of the algorithm?" Clearly, if the cost function is fixed and we rule out the possible technical problems associated with programming and computing power, the only remaining parameter is the initial refinement level at which the algorithm is going to work—that is, the level j. However, if we reflect more deeply on the reasons why there is a practical threshold beyond which at a certain level of discretization the algorithm has stable behavior, we only have as a possible causes the level of information of the cost function that captures the MR analysis. In other words, if the oscillation of the cost function at a certain level of resolution is well determined by MR analysis, then the algorithm will have good performance.

The approach presented in this paper is far from exhaustive and, on the contrary, opens the possibility for a number of new proposals for approximating solutions to the MK problem. The above is due to the fact that in the work [9], it was proven that discretization of the MK problem can be performed from any MR analysis over $\mathbb{R}^2$. Therefore, the possibility of implementing other types of discretions remains open. In principle, as we mentioned in the previous paragraph, the most natural thing is to expect better performance if the nature of the cost function and the types of symmetric geometric structures that it induces in space are studied in order to use an MR analysis that fits this information and therefore has more efficient performance.

References

1. Hernández-Lerma, O.; Lasserre, J.B. Approximation schemes for infinite linear programs. *SIAM J. Optim.* **1998**, *8*, 973–988. [CrossRef]
2. González-Hernández, J.; Gabriel-Argüelles, J.R.; Hernández-Lerma, O. On solutions to the mass transfer problem. *SIAM J. Optim.* **2006**, *17*, 485–499. [CrossRef]
3. Gabriel-Argüelles, J.R.; González-Hernández, J.; López-Martínez, R.R. Numerical approximations to the mass transfer problem on compact spaces. *IMA J. Numer. Anal.* **2010**, *30*, 1121–1136. [CrossRef]
4. Avendaño-Garrido, M.L.; Gabriel-Argüelles, J.R.; Quintana-Torres, L.; Mezura-Montes, E. A metaheuristic for a numerical approximation to the mass transfer problem. *Int. J. Appl. Math. Comput. Sci.* **2016**, *26*, 757–766. [CrossRef]
5. Avendaño-Garrido, M.L.; Gabriel-Argüelles, J.R.; Quintana-Torres, L.; Mezura-Montes, E. An Efficient Numerical Approximation for the Monge–Kantorovich Mass Transfer Problem. In *Machine Learning, Optimization, and Big Data. MOD 2015*; Lecture Notes in Computer Science; Pardalos, P., Pavone, M., Farinella, G., Cutello, V., Eds.; Springer: Cham, Switzerland, 2016; Volume 9432, pp. 233–239. [CrossRef]

6.   Gröchenig, K.; Madych, W.R. Multiresolution analysis, Haar bases, and self-similar tilings of $\mathbb{R}^n$. *Inst. Electr. Electron. Eng. Trans. Inf. Theory* **1992**, *38*, 556–568. [CrossRef]

7.   Liu, Y.; Xu, K.D.; Li, J.; Guo, Y.J.; Zhang, A.; Chen, Q. Millimeter-wave E-plane waveguide bandpass filters based on spoof surface plasmon polaritons. *IEEE Trans. Microw. Theory Tech.* **2022**, *70*, 4399–4409. [CrossRef]

8.   Liu, K.; Yang, Z.; Wei, W.; Gao, B.; Xin, D.; Sun, C.; Gao, G.; Wu, G. Novel detection approach for thermal defects: Study on its feasibility and application to vehicle cables. *High Volt.* **2023**, *8*, 358–367. [CrossRef]

9.   Sánchez-Nungaray, A.; González-Flores, C.; López-Martínez, R.R. Multiresolution Analysis Applied to the Monge–Kantorovich Problem. *Abstr. Appl. Anal.* **2018**, *2018*, 1764175. [CrossRef]

10.   Haker, S.; Tannenbaum, A.; Kikinis, R. *Mass Preserving Mappings and Surface Registration*; MICCAI 2001, Lecture Notes in Computer Science; Springer: Berlin/Heidelberg, Germany, 2001; Volume 2208, pp. 120–127.

11.   Zhu, L.; Yang, Y.; Haker, S.; Tannenbaum, A. Optimal mass transport for registration and warping. *Int. J. Comput. Vis.* **2004**, *60*, 225–240. .:VISI.0000036836.66311.97. [CrossRef]

12.   Haber, E.; Rehman, T.; Tannenbaum, A. An efficient numerical method for the solution of the L_2 optimal mass transfer problem. *SIAM J. Sci. Comput.* **2010**, *32*, 7–211. [CrossRef] [PubMed]

13.   Xie, X.; Wang, T.; Zhang, W. Existence of solutions for the (p, q)-Laplacian equation with nonlocal Choquard reaction. *Appl. Math. Lett.* **2023**, *135*, 108418. [CrossRef]

14.   Peng, Z.; Hu, J.; Shi, K.; Luo, R.; Huang, R.; Ghosh, B.K.; Huang, J. A novel optimal bipartite consensus control scheme for unknown multi-agent systems via model-free reinforcement learning. *Appl. Math. Comput.* **2020**, *369*, 124821. [CrossRef]

15.   Walnut, D.F. *An Introduction to Wavelet Analysis*; Birkhäuser: Boston, MA, USA, 2004; pp. 115–138.

16.   Guo, K.; Labate, D.; Lim, W.Q.; Weiss, G.; Wilson, E. Wavelets with composite dilations and their MRA properties. *Appl. Comput. Harmon. Anal.* **2006**, *20*, 202–236. [CrossRef]

17.   Krishtal, I.A.; Robinson, B.D.; Weiss, G.L.; Wilson, E.N. Some simple Haar-type wavelets in higher dimensions. *J. Geom. Anal.* **2007**, *17*, 87–96. [CrossRef]

18.   Bazaraa-Mokhtar, S.; John, J.; Hanif, D. *Linear Programming and Network Flows*; John Wiley & Sons: Hoboken, NJ, USA, 2010; pp. 513–528.

19.   Billingsley, P. *Convergence of Probability Measures*; John Wiley & Sons: New York, NY, USA, 1999; pp. 27–29.