ESAIM: M2AN 60 (2026) 1055–1079
https://doi.org/10.1051/m2an/2026030

ESAIM: Mathematical Modelling and Numerical Analysis
www.esaim-m2an.org

SEMI-DISCRETE MULTI-TO-ONE DIMENSIONAL VARIATIONAL PROBLEMS

Omar Abdul Halim*, Daniyar Omarov and Brendan Pass

Abstract. We study a class of semi-discrete variational problems that arise in economic matching and game theory, where agents with continuous attributes are matched to a finite set of outcomes with a one dimensional structure. Such problems appear in applications including Cournot–Nash equilibria, and hedonic pricing, and can be formulated as problems involving optimal transport between spaces of unequal dimensions. In our discrete strategy space setting, we establish analogues of results developed for a continuum of strategies in Nenna and Pass [*J. Math. Pures Appl.* **139** (2020) 83–108], ensuring solutions have a particularly simple structure under certain conditions. This has important numerical consequences, as it is natural to discretize when numerically computing solutions. We adapt standard semi-discrete optimal transport techniques to the variational setting in which the target measure is unknown. By leveraging discrete nestedness when it holds, our sequential algorithms improve robustness and achieve computational gains, together with rigorous convergence guarantees, as demonstrated through numerical experiments.

Mathematics Subject Classification. 49Q22, 65K10, 91B52.

Received August 13, 2025. Accepted March 25, 2026.

1. Introduction

A broad class of problems in economics, game theory, and urban planning can be formulated as variational problems on the space of probability measures involving optimal transport. More explicitly, they amount minimizing a functional of the form:

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

where μ is a fixed distribution of agents, ν a free distribution of strategies, $\mathcal{W}_c$ is the cost of optimal transport between them with cost function c and $\mathcal{F}$ a functional on the space $\mathcal{P}(Y)$ of probability measures, with different forms depending on the precise problem under consideration. Examples include *Cournot–Nash equilibria* models [4–7], where agents interact strategically in markets with congestion or competition effects; *urban planning* problems [2, 9–12], which involve the optimal spatial allocation of infrastructure or public goods; and *hedonic pricing* models [13, 18], where agents evaluate goods or services based on individual preferences.

These problems are computationally challenging due to the complexity of optimal transport. However, in at least one important regime they can be simplified considerably. It is often reasonable to parametrize agents by

Keywords and phrases. Optimal transport, variational problems, numerical algorithms, Cournot–Nash equilibria, hedonic pricing.

University of Alberta, Edmonton, Alberta, Canada.
*Corresponding author: oabdulh1@ualberta.ca

a distribution μ on a high dimensional space X (reflecting a high degree of heterogeneity), whereas available strategies have more limited variability and are therefore modeled by a one-dimensional space Y. The corresponding *multi-to one-dimensional* optimal transport problem admits a greatly simplified characterization of solutions, provided that a certain condition on c, μ and ν, recently developed in [14] and known as *nestedness*, is satisfied. Furthermore, hypotheses on $\mathcal{F}$, c and μ have recently been developed ensuring that for the optimal ν in (1), the optimal transport problem arising there is nested [25]. These advances greatly enhance the theoretical tractability of (1), but a significant issue arises when trying to apply them numerically. In practice, when computing solutions, one typically discretizes the target space, so that the one-dimensional space Y is replaced by a finite set Y_N and the unknown measure $\nu \in \mathcal{P}(Y_N)$ becomes discrete. However, the original notion of nestedness relies on the continuous structure of Y in a crucial way, and so the simple characterization of solutions in [25] cannot be directly exploited numerically.[1]

On the other hand, an analogue of nestedness for optimal transport problems with discrete target spaces has been recently introduced in [1], together with an essentially closed form characterization of solutions when this condition is satisfied. The purpose of this paper is twofold. First, to translate the analysis in [25] to the discrete target setting, and establish conditions on μ, c, F, and Y_N under which solutions to (1) satisfy discrete nestedness, and, secondly, to exploit these results to develop efficient numerical algorithms.

Our focus here is therefore on discrete versions of the two main types of functionals $\mathcal{F}$ studied in [25]: *internal energies*, or *congestion terms*, modeling situations where agents prefer strategies which differ from those chosen by other agents, and those modeling *hedonic pricing problems*, in which both buyers and sellers seek to maximize their utilities while exchanging goods chosen from a set Y_N of feasible goods or contracts; in this case, the functional $\mathcal{F}$ is itself the optimal transport distance to another fixed distribution μ_1.

On the numerical side, several computational methods for semi-discrete OT problems with fixed target points have already been developed; see, for example, [3, 8, 15, 21, 22] for 2-norm squared costs, [20] for 2-norm costs, and [16, 17] for more general p-norms. In addition, there are several extensive available packages for solving standard problems with fixed target measures such as *pysdot* and *geogram*. On the other hand, there are not built-in packages for variational problem with varying target measures. However, one can easily adapt standard techniques, such as the *lifting* algorithm of [22] and analytical Jacobian expressions from [15, 21] to the problems considered in this work, and we do so here. Our main numerical goal is to leverage the discrete nested structure (when it holds) to develop alternative algorithms which, as we show below, have certain numerical advantages. In particular, we develop new damped Newton algorithms in which the damping exploits nestedness, which our numerical experiments indicate are faster than standard Newton algorithms. We also develop what we call sequential methods, which directly use the nested structure of solutions. Though these are slower than the Newton algorithms, they are much more robust with respect to initialization. In fact, for the internal energy, we prove that our sequential method always converges (for any initialization) when the solution is nested; this stands in contrast to Newton methods which are only locally convergent.

The paper is organized as follows. In Section 2, we review relevant background on optimal transport and discrete nestedness. Section 3 considers the semi-discrete framework of the variational problem with a congestion term and introduces a numerical method for computing the solution. In Section 4, we present the hedonic pricing problem in the semi-discrete setting and propose an algorithm that utilizes the nested structure of the solution to improve computational efficiency in the case of nested solution.

2. Nestedness of optimal transport

2.1. Optimal transport

Here we briefly introduce the optimal transport problem; a fuller introduction can be found in, for example, [28, 29]. Given probability measures μ and ν on bounded sets $X \subset \mathbb{R}^{d_X}$ and $Y \subseteq \mathbb{R}^{d_Y}$, respectively, and a cost

[1]Some related computations with discrete measures are carried out in [26]. Calculations there are done using variables in a continuous ambient space Y, and the effect on the solution of passing back and forth between discrete and continuous problems is neglected.

function $c \in C(X \times Y)$, the optimal transport problem is to minimize

$$\mathcal{W}_c(\mu, \nu) = \min_{\gamma \in \Gamma(\mu,\nu)} \int_{X \times Y} c(x, y)\, \mathrm{d}\gamma(x, y), \tag{2}$$

where $\Gamma(\mu, \nu)$ consists of all probability measures γ on $X \times Y$ with marginals μ and ν.

As a linear program, (2) admits a dual problem and it is well known that strong duality holds; that is,

$$\mathcal{W}_c(\mu, \nu) = \sup_{(u,v)\in\mathcal{V}} \left\{ \int_X u(x)\mathrm{d}\mu(x) + \int_Y v(y)\, \mathrm{d}\nu(y) \right\}, \tag{3}$$

where $\mathcal{V} = \{(u, v) \in L^1(\mu) \times L^1(\nu) \mid u(x) + v(y) \le c(x, y),\ \forall (x, y) \in X \times Y\}$ is the set of feasible potentials. Moreover, it is well known that a minimizer γ in (2) and a maximizer (u, v) in (3) both exist, and that $u(x) + v(y) = c(x, y)$, γ almost everywhere. Furthermore, the optimal dual potentials can be chosen to be *c-concave*, meaning that $u(x) = \inf_{y\in Y}\{c(x, y) - v(y)\}$ and $v(y) = \inf_{x\in X}\{c(x, y) - u(x)\}$.

We are especially interested here in the semi-discrete setting, where the source measure μ is continuous, while the target measure $\nu = \sum_{i=1}^N \nu_i \delta_{y_i}$ is supported on finitely many points $\{y_i\}_{i=1}^N$ (see [19, 24, 27] for a detailed review). In this case, the optimal dual potential $v = (v_1, \dots, v_N)$ becomes a vector in $\mathbb{R}^N$ where $v_i = v(y_i)$. It induces the measurable sets

$$X_i := \{x \in X \mid c(x, y_i) - v_i \le c(x, y_j) - v_j \text{ for all } j\}.$$

Any optimal measure γ can only transport x to y_i if $x \in X_i$, and the dual potential u is piecewise defined on these sets by $u(x) = c(x, y_i) - v_i \quad$ for $x \in X_i$. On the boundary between regions X_i and X_j, the potential satisfies $c(x, y_i) - v_i = c(x, y_j) - v_j$, which determines the geometry of the transport cells and the interfaces between them.

2.2. A sufficient condition for discrete nestedness

Throughout the rest of the paper, we specialize to the case where $d_Y = 1$ and simply denote $d_X = d$. Let $X \subset \mathbb{R}^d$ and $Y = \{y_1, \dots, y_N\} \subset \mathbb{R}$ be bounded sets such that X is open. We will make the following assumption, which is a slight weakening of the well known twist condition, on $c \in C^2(X \times Y)$:

(H1) $D_x c(x, y_{i+1}) - D_x c(x, y_i) \neq 0$ for all $x \in X$ and $1 \le i \le N - 1$.

Let $\mu \in \mathcal{P}(X)$ be an absolutely continuous measure with positive density, and let $\nu \in \mathcal{P}(Y)$ be given by $\nu = \sum_{i=1}^N \nu_i \delta_{y_i}$, where δ_{y_i} is the Dirac mass at $y_i \in Y$ and $\nu_i \in (0, 1]$. We define the following level and sub-level sets:

$$X^N_=(y_i, k) := \{x \in X : c(x, y_{i+1}) - c(x, y_i) = k\},$$
$$X^N_\ge(y_i, k) := \{x \in X : c(x, y_{i+1}) - c(x, y_i) \ge k\},$$

and $X^N_>(y, k) := X^N_\ge(y, k) \setminus X^N_=(y, k)$.

Definition 2.1. We say that the optimal transport problem (c, μ, ν) is discretely nested if for all $1 \le i \le N-2$, we have

$$X^N_\ge(y_i, k^N(y_i)) \subset X^N_>(y_{i+1}, k^N(y_{i+1})),$$

where $k^N(y_r)$ is chosen such that $\mu(X^N_\ge(y_r, k^N(y_r))) = \sum_{i=1}^r \nu_i$.

Remark 2.2 (Choice of the thresholds $k^N(y_r)$)**.** Fix $r \in \{1, \dots, N - 1\}$ and define $h_r(k) := \mu(X^N_\ge(y_r, k)) - \sum_{i=1}^r \nu_i$. By (H1) and the Implicit Function Theorem, for each k the level set $X^N_=(y_r, k) = \{c(\cdot, y_{r+1}) - c(\cdot, y_r) = k\}$ is an $(m-1)$-dimensional hypersurface in X. Since μ is absolutely continuous, it follows that $\mu(X^N_=(y_r, k)) = 0$ for all k. Therefore h_r is continuous.

Since $X^N_{\geq}(y_r,k_1) \subseteq X^N_{\geq}(y_r,k_2)$ for $k_1 > k_2$, the function h_r is nonincreasing in k. Moreover, h_r is constant and equals to $1-\sum_{i=1}^r \nu_i$ for $k \leq \underline{k_r} = \max\{k : X^N_{\geq}(y_r,k) = X\}$ and constant equals to $-\sum_{i=1}^r \nu_i$ for $k \geq \overline{k_r} = \min\{k : X^N_{\geq}(y_r,k) = \emptyset\}$. On the interval $[\underline{k_r}, \overline{k_r}]$, the positivity of the density of μ together with $\mu(X^N_{=}(y_r,k)) = 0$ implies that h_r is strictly decreasing.

As $h_r(\underline{k_r}) = 1-\sum_{i=1}^r \nu_i > 0$ and $h_r(\overline{k_r}) = -\sum_{i=1}^r \nu_i < 0$, by the Intermediate Value Theorem, there exists a unique $k^N(y_r) \in [\underline{k_r}, \overline{k_r}]$ such that $\mu(X^N_{\geq}(y_r,k^N(y_r))) = \sum_{i=1}^r \nu_i$.

Note that this definition is equivalent to the definition of discrete nestedness introduced in [1], where, whenever $\sum_{r=i+1}^j \nu_r > 0$, we require $X^N_{\geq}(y_i,k^N(y_i)) \subset X^N_{>}(y_j,k^N(y_j))$.

The discrete nestedness condition imposes a monotonicity structure on the family of superlevel sets $X^N_{\geq}(y_i,k^N(y_i))$. This structure ensures that the regions associated with successive outcomes are ordered in a way that enables a piecewise construction of the optimal transport map. The following result provides a precise characterization of the solution under this condition. It is proven in [1], and can be seen as a semi-discrete version of Theorem 4 in [14].

Theorem 2.3 ([1])**.** *Assume the optimal transport problem (c,μ,ν) is discretely nested. Then, setting $X_1 = X^N_{>}(y_1,k^N(y_1))$, $X_N = X \setminus X^N_{\geq}(y_{N-1},k^N(y_{N-1}))$, and $X_i = X^N_{>}(y_i,k^N(y_i)) \setminus X^N_{\geq}(y_{i-1},k^N(y_{i-1}))$ for all $1 < i < N$, the optimal transport plan is unique and pairs each $x \in X_i$ with y_i for all $1 \leq i \leq N$. Furthermore, the optimal potentials (u,v) are defined by $u(x) = c(x,y_i) - v_i$ for all $x \in X_i$, where*

$$v_i = \sum_{j=1}^{i-1} k^N(y_j),$$

for $1 < i \leq N$ where $v_1 = 0$.

We now turn to the task of establishing sufficient conditions for discrete nestedness. Our work in the remainder of this section mirrors the theory developed for continuous one-dimensional targets in Section 2.2 of [25]. For $1 \leq i \leq N-2$, we define the minimal mass difference as follows,

$$D^{\min}_{\mu}(y_i,k_i) = \mu\big(X^N_{\geq}(y_{i+1},k_{\max}(y_i,k_i)) \setminus X^N_{\geq}(y_i,k_i)\big)$$

where $k_{\max}(y_i,k_i) = \sup\big\{k \in \mathbb{R} : X^N_{\geq}(y_i,k_i) \subseteq X^N_{\geq}(y_{i+1},k)\big\}$ for some k_i. This minimal mass difference tells us the least additional μ-mass required so that a superlevel set $X^N_{\geq}(y_i,k_i)$ becomes contained in a corresponding set for y_{i+1}. This leads to a necessary condition for discrete nestedness, as well as a sufficient condition under strict inequality.

Theorem 2.4. *If (c,μ,ν) is discretely nested, then $D^{\min}_{\mu}(y_i,k^N(y_i)) \leq \nu_{i+1}$ for all $1 \leq i \leq N-2$. Conversely, if $D^{\min}_{\mu}(y_i,k^N(y_i)) < \nu_{i+1}$ for all $1 \leq i \leq N-2$, then (c,μ,ν) is discretely nested.*

Proof. Assume that (c,μ,ν) is discretely nested. Then, for all $1 \leq i \leq N-2$ we have $X^N_{\geq}(y_i,k^N(y_i)) \subset X^N_{>}(y_{i+1},k^N(y_{i+1})) \subset X^N_{\geq}(y_{i+1},k^N(y_{i+1}))$ which implies $k^N(y_{i+1}) \leq k_{\max}(y_i,k^N(y_i))$. Hence,

$$\begin{aligned} D^{\min}_{\mu}(y_i,k^N(y_i)) &= \mu(X^N_{\geq}(y_{i+1},k_{\max}(y_i,k^N(y_i))) \setminus X^N_{\geq}(y_i,k^N(y_i))) \\ &\leq \mu(X^N_{\geq}(y_{i+1},k^N(y_{i+1})) \setminus X^N_{\geq}(y_i,k^N(y_i)) \\ &= \mu(X^N_{\geq}(y_{i+1},k^N(y_{i+1}))) - \mu(X^N_{\geq}(y_i,k^N(y_i))) = \nu_{i+1} \end{aligned}$$

for all $1 \leq i \leq N-2$.

Assume that $D^{\min}_{\mu}(y_i,k^N(y_i)) < \nu_{i+1}$ for all $1 \leq i \leq N-2$. If (c,μ,ν) is not discretely nested, then there exists $1 \leq j \leq N-2$ such that $X^N_{\geq}(y_j,k^N(y_j)) \not\subset X^N_{>}(y_{j+1},k^N(y_{j+1}))$. Thus, $k^N(y_{j+1}) \geq k_{\max}(y_j,k^N(y_j))$ and

$$\mu(X^N_{\geq}(y_{j+1},k^N(y_{j+1})) \setminus X^N_{\geq}(y_j,k^N(y_j))) \leq D^{\min}_{\mu}(y_j,k^N(y_j)).$$

But,

$$\begin{aligned}\nu_{j+1} &= \mu(X^N_{\geq}(y_{j+1}, k^N(y_{j+1}))) - \mu(X^N_{\geq}(y_j, k^N(y_j))) \\ &\leq \mu(X^N_{\geq}(y_{j+1}, k^N(y_{j+1})) \setminus X^N_{\geq}(y_j, k^N(y_j)) \leq D^{\min}_{\mu}(y_j, k^N(y_j))\end{aligned}$$

which is a contradiction. □

The minimal mass condition in Theorem 2.4 yields a verifiable sufficient condition for discrete nestedness that does not require prior knowledge of the splitting levels $k^N(y_i)$. In particular, by uniformly bounding the minimal mass difference across all splitting levels, we obtain the following corollary.

Corollary 2.5. *If for all* $1 \leq i \leq N-2$ *we have* $\sup_{k\in\mathbb{R}} D^{\min}_{\mu}(y_i, k) - \nu_{i+1} < 0$, *then* (c, μ, ν) *is discretely nested.*

Proof. The condition $\sup_{k\in\mathbb{R}} D^{\min}_{\mu}(y_i, k) - \nu_{i+1} < 0$ implies that $D^{\min}_{\mu}(y_i, k) - \nu_{i+1} < 0$ for all k in particular $D^{\min}_{\mu}(y_i, k^N(y_i)) - \nu_{i+1} < 0$ for all $1 \leq i \leq N-2$ and by Theorem 2.4 we get discrete nestedness. □

Example 2.6. Let μ be the uniform measure on $X = (0,1)^2$ with $c(x, y_i) = -x_1 y_i - x_2 F(y_i)$ for all $x = (x_1, x_2) \in X$ and $y_i \in Y$, where F is an increasing convex function. In this case, the set $X^N_{=}(y_i, k)$ is a line orthogonal to the vector $\Big(y_{i+1} - y_i, F(y_{i+1}) - F(y_i)\Big)$.

For a fixed k_0, the set $X^N_{\geq}(y_{i+1}, k_{\max}(y_i, k_0)) \setminus X^N_{\geq}(y_i, k_0)$ is contained within a triangle whose vertices are given by the intersections of $X^N_{=}(y_i, k_0)$, $X^N_{=}(y_{i+1}, k_{\max}(y_i, k_0))$, and the x_1-axis. It follows that

$$D^{\min}_{\mu}(y_i, k_0) \leq \sup_{k\in\mathbb{R}} D^{\min}_{\mu}(y_i, k) \leq \frac{1}{2}\left(\frac{F(y_{i+2}) - F(y_{i+1})}{y_{i+2} - y_{i+1}} - \frac{F(y_{i+1}) - F(y_i)}{y_{i+1} - y_i}\right).$$

This supremum is bounded above by the area of the triangle formed by the intersection of $X^N_{=}(y_i, k)$, $X^N_{=}(y_{i+1}, k_{\max}(y_i, k))$, and the x_1-axis for some k such that the intersection of $X^N_{=}(y_i, k)$ and $X^N_{=}(y_{i+1}, k_{\max}(y_i, k))$ lies on the line $x_2 = 1$.

By Corollary 2.5, the problem (c, μ, ν) is discretely nested whenever

$$\sup_{k\in\mathbb{R}} D^{\min}_{\mu}(y_i, k) \leq \frac{1}{2}\left(\frac{F(y_{i+2}) - F(y_{i+1})}{y_{i+2} - y_{i+1}} - \frac{F(y_{i+1}) - F(y_i)}{y_{i+1} - y_i}\right) < \nu_{i+1}$$

for all $1 \leq i \leq N-2$.

Remark 2.7. Setting $z_i = (y_i, F(y_i)) \in \mathbb{R}^2$, Example 2.6 corresponds to the classical planar semi-discrete optimal transport problem with linear cost $\tilde{c}(x, z) = -\langle x, z\rangle$. We write the cost in the form $c(x, y) = -\langle x, (y, F(y))\rangle$ to remain within the multi-to-one dimensional framework ($d_Y = 1$), with F encoding the planar geometry of the embedded target points.

3. Internal energy

As above, we take $Y = \{y_1, \dots, y_N\}$ to be discrete throughout this section. We consider variational problems of the form

$$\min_{\nu\in\mathcal{P}(Y)} \{\mathcal{W}_c(\mu, \nu) + \mathcal{F}(\nu)\}, \tag{4}$$

for a fixed $\mu \in \mathcal{P}(X)$, where $\mathcal{F}(\nu) = \sum_{i=1}^N f(\nu_i)\nu_i$ and $f : [0, \infty) \to \mathbb{R}$ is continuously differentiable on $(0, \infty)$, strictly convex, has superlinear growth, and satisfies $\lim_{s\to 0^+} f'(s) = -\infty$. These conditions on f guarantee the existence and uniqueness of a minimizer. Moreover, the strict convexity of f implies that for each fixed v the function $s \mapsto \sum_{i=1}^N (f')^{-1}(s - v)$ is strictly increasing (where $(f')^{-1}$ denotes the inverse function of f'), and therefore admits an inverse, which we denote by J_v. By [6], under the above assumptions, there exists a unique solution of (4). Our work in this section closely follows the corresponding analysis in the continuous case in Section 3.1 of [25].

3.1. Discrete nestedness of the solutions

To proceed with our analysis, we first obtain bounds on the optimal weights ν_i. In particular, the following proposition provides lower and upper estimates on each ν_i in terms of f, the geometry of Y and the cost function c.

Proposition 3.1. *The minimizer ν of (4) satisfies*

$$(f')^{-1}(J_{-M_c|y_i-y_1|}(1) - M_c|y_i - y_1|) \leq \nu_i \leq (f')^{-1}(J_{M_c|y_i-y_1|}(1) + M_c|y_i - y_1|)$$

where $M_c = \sup_{(x,y_j,y_k)} \frac{|c(x,y_j)-c(x,y_k)|}{|(x,y_j)-(x,y_k)|}$ *and* $|\cdot|$ *is the Euclidean norm.*

Proof. Let ν be the minimizer of (4) and let $v = (v_i)_{i=1}^N$ such that (u, v) is the solution to the Kantorovich dual problem (3) between the source measure μ on X and the target measure $\nu = \sum_{i=1}^N \nu_i \delta_{y_i}$.

Without loss of generality, we normalize by taking $v_1 = 0$. From [6], we get the first order optimality condition of (4) which is

$$v_i + f'(\nu_i) = C, \tag{5}$$

for some constant C. From (5), we get $\nu_i = (f')^{-1}(C - v_i)$ and using the fact that ν is a probability measure we deduce $1 = \sum_{i=1}^N \nu_i = \sum_{i=1}^N (f')^{-1}(C - v_i)$ and so $J_v(1) = C$.

From [23] we know that v is a Lipschitz function with the constant $M_c = \sup_{(x,y_j,y_k)} \frac{|c(x,y_j)-c(x,y_k)|}{|(x,y_j)-(x,y_k)|}$ where $|\cdot|$ is the Euclidean norm, thus $-M_c|y_i - y_1| \leq v_i \leq M_c|y_i - y_1|$. By the monotonicity of $(f')^{-1}$, we get

$$\sum_{i=1}^N (f')^{-1}(C - M_c|y_i - y_1|) \leq \sum_{i=1}^N (f')^{-1}(C - v_i) = 1 \leq \sum_{i=1}^N (f')^{-1}(C + M_c|y_i - y_1|).$$

Apply $J_{-M_c|y_i-y_1|}$ to $1 \leq \sum_{i=1}^N (f')^{-1}(C + M_c|y_i - y_1|)$ to get $J_{-M_c|y_i-y_1|}(1) \leq C$. Similarly, we get $C \leq J_{M_c|y_i-y_1|}(1)$. Therefore,

$$(f')^{-1}(J_{-M_c|y_i-y_1|}(1) - M_c|y_i - y_1|) \leq (f')^{-1}(C - v_i) = \nu_i \leq (f')^{-1}(J_{M_c|y_i-y_1|}(1) + M_c|y_i - y_1|).$$

□

Using the lower bound on the components of ν from Proposition 3.1, we now derive a concrete sufficient condition ensuring that (c, μ, ν) is discretely nested. This is done by inserting the lower estimate into the mass comparison criterion of Corollary 2.5.

Theorem 3.2. *Assume that*

$$\sup_{k \in \mathbb{R}} D_\mu^{\min}(y_i, k) - (f')^{-1}(J_{-M_c|y_{i+1}-y_1|}(1) - M_c|y_{i+1} - y_1|) < 0$$

for all $1 \leq i \leq N - 2$. Then, letting ν be the minimizer of (4), (c, μ, ν) is discretely nested.

Proof. The proof follows from Corollary 2.5 and Proposition 3.1. □

The following example illustrates how the preceding result can be used to ensure nestedness of the solution on a concrete example.

Example 3.3. Going back to Example 2.6, when μ is the uniform measure on $X = (0,1)^2$ with $c(x, y_i) = -x_1 y_i - x_2 F(y_i)$ for all $y_i \in Y$ with $0 = y_1 \leq y_i < y_{i+1}$, and F is an increasing convex function, let $f(s) = s\ln(s)$. In this case, apart from an irrelevant additive constant we get $f'(s) = \ln(s)$ and J_v is the inverse of

$s \mapsto \sum_{i=1}^N e^{s-v_i} = e^s \sum_{i=1}^N e^{-v_i}$ where v_i is the Kantorovich potential. Hence, $J_v(1) = \ln\Big(\frac{1}{\sum_{i=1}^N e^{-v_i}}\Big)$. Also, we have $M_c = 1$, and by Proposition 3.1, we get

$$\nu_{i+1} \geq e^{\ln\Big(\frac{1}{\sum_{p=1}^N e^{y_{i+1}}}\Big) - y_{i+1}} = \frac{e^{-y_{i+1}}}{Ne^{y_{i+1}}}.$$

By Theorem 3.2, the problem (c, μ, ν) is nested where ν is the minimizer whenever

$$\sup_{k\in\mathbb{R}} D_\mu^{\min}(y_i, k) \leq \tfrac{1}{2}\Big(\tfrac{F(y_{i+2})-F(y_{i+1})}{y_{i+2}-y_{i+1}} - \tfrac{F(y_{i+1})-F(y_i)}{y_{i+1}-y_i}\Big) < \tfrac{1}{Ne^{2y_{i+1}}}$$

for all $1 \leq i \leq N-2$. When $F(y) = \frac{y^2}{A}$ for some constant A and $y_i = \frac{i}{N}$, we have $\sup_{k\in\mathbb{R}} D_\mu^{\min}(y_i, k) \leq \frac{1}{AN}$ which implies that the problem (c, μ, ν) is discretely nested when $A > e^2$.

The structure described above admits a meaningful interpretation in practical decision-making contexts. Inspired by the holiday choice in Section 2.1 of [4], consider, for instance, a population of individuals (represented by the measure μ) choosing among a finite set of national parks, indexed by $y_i \in Y$. Each park is characterized by a scalar attribute y_i, which quantifies its scenic quality – reflecting features such as landscape diversity, iconic viewpoints, and biodiversity. We can interpret $F(y_i)$ as the (normalized) quality of visitor-facing utilities and services at the park, such as accessibility of trailheads, availability and reliability of shuttle services, cleanliness and density of restrooms, potable water points, and safety infrastructure. Empirically, these utilities tend to improve with scenic quality, and often do so in an accelerating fashion – top-tier scenic parks attract disproportionately larger investments in visitor infrastructure.

Each individual is described by a vector $x = (x_1, x_2)$, where x_1 represents their preference for scenic quality, and x_2 encodes their preference for high-quality visitor utilities. The cost function $c(x, y_i) = -x_1 y_i - x_2 F(y_i)$ then encodes preferences that weigh the park's natural scenery (through y_i) against the quality of its visitor amenities (through $F(y_i)$). The allocation of individuals across parks arises as a Cournot–Nash equilibrium [6]: each visitor chooses the park that maximizes their own utility while considering the overall distribution of other visitors. Congestion effects, modeled by the term $f(s) = s \ln s$, capture the population's aversion to overcrowding – visitors prefer not to choose parks that are already popular and heavily frequented by others.

3.2. Numerical algorithm

In this section, we present a numerical algorithm for computing the solution of (4) under the assumption of discrete nestedness. We first present the numerical formulation of the problem introduced earlier. Let $c \in C^2(X \times Y)$ satisfy (H1).

Definition 3.4 (Laguerre cells and tessellation)**.** Given a finite set of points $\{y_i\}_{i=1}^N \subset Y$, a weight vector $\mathbf{v} = (v_i)_{i=1}^N \in \mathbb{R}^N$, and a cost function $c : X \times Y \to \mathbb{R}_{\geq 0}$, the *Laguerre cell* associated with y_i is defined as

$$\mathrm{Lag}_i(\mathbf{v}) := \{x \in X \mid c(x, y_i) - v_i \leq c(x, y_j) - v_j, \ \forall j \neq i\}.$$

The collection of cells $\{\mathrm{Lag}_i(\mathbf{v})\}_{i=1}^N$ constitutes a *Laguerre tessellation* of the domain X. Furthermore, the tessellation is said to be *nested* if each Laguerre cell $\mathrm{Lag}_i(\mathbf{v})$ shares a boundary with at most two adjacent cells, namely $\mathrm{Lag}_{i-1}(\mathbf{v})$ and $\mathrm{Lag}_{i+1}(\mathbf{v})$, for all $1 < i < N$.

For the semi-discrete OT problem $\mathcal{W}_c(\mu, \nu)$, an optimal dual vector $\mathbf{v}$ induces Laguerre cells $\mathrm{Lag}_i(\mathbf{v})$ satisfying the mass balance $\mu(\mathrm{Lag}_i(\mathbf{v})) = \nu_i$ (see, *e.g.* [24]).

By [6] and as recalled in the proof of Proposition 3.1, the minimizer ν of the variational problem (4) satisfies the sufficient first-order optimality condition $\nu_i = (f')^{-1}(C - v_i)$ for some constant C, where $\mathbf{v} = (v_i)_{i=1}^N$ is the optimal dual vector associated with $\mathcal{W}_c(\mu, \nu)$. Together with the mass balance condition $\mu(\mathrm{Lag}_i(\mathbf{v})) = \nu_i$, this yields a coupled nonlinear system $\mu(\mathrm{Lag}_i(\mathbf{v})) = (f')^{-1}(C - v_i)$, for the weights $\mathbf{v}$ and the scalar C, which we now seek to solve numerically.

Problem 3.5. Given a set of points $\{y_i\}_{i=1}^N$, a probability measure $\mu(x)$ on the domain X, and a cost function c, we seek to solve for the unknown weights $\mathbf{v} = (v_i)_{i=1}^N \in \mathbb{R}^N$ and the scalar constant $C \in \mathbb{R}$ the following system of equations:

$$\mu(\mathrm{Lag}_i(\mathbf{v})) - (f')^{-1}(C - v_i) = 0, \quad i = 1, 2, \ldots, N. \tag{6}$$

In particular, if $f(z) = z \log z$, then $(f')^{-1}(z) = e^z$. Substituting this expression into the system (6) yields:

$$\mu(\mathrm{Lag}_i(\mathbf{v})) - e^{C - v_i} = 0, \quad i = 1, 2, \ldots, N. \tag{7}$$

Remark 3.6. To satisfy the compatibility condition in (7), the following normalization must hold:

$$\sum_{i=1}^N \mu(\mathrm{Lag}_i(\mathbf{v})) = 1 \quad \Longrightarrow \quad \sum_{i=1}^N e^{C - v_i} = 1.$$

One approach to enforce this condition is to determine the value of C that satisfies the constraint, and then solve the resulting system of equations. Specifically, C is given by:

$$C = \log\left(\frac{1}{\sum_{i=1}^N e^{-v_i}}\right).$$

Substituting this expression into the system (7) yields the following:

$$\mu(\mathrm{Lag}_i(\mathbf{v})) - \frac{e^{-v_i}}{\sum_{k=1}^N e^{-v_k}} = 0, \quad i = 1, 2, \ldots, N. \tag{8}$$

With the problem reformulated in a form amenable to numerical computation, we adopt Newton's method as a baseline for solving the system in (8). Known for its fast local convergence, Newton's method is a natural first choice for root-finding problems. However, it suffers from stability and global convergence issues, which motivate the exploration of alternative methods. We include standard Newton's method here both to clarify its implementation and to provide a benchmark for comparison.

3.2.1. Standard Newton's method

In all examples, the Jacobian matrix used in Newton's method was computed using the explicit analytic formula given in Theorem 1.3 of [21], adapted to the internal-energy problem. It is important to note that, as observed in (8), the vector $\mathbf{v}$ is unique only up to scalar addition. Consequently, the inverse of the Jacobian matrix is computed in the orthogonal complement of its kernel, which is spanned by $\mathbf{1}$.

Problem 3.7 (Standard Newton's method)**.** Consider the problem of determining $\mathbf{v}^*$, a vector in $\mathbb{R}^N$, such that $G(\mathbf{v}^*) = \mathbf{0}$ for the function $G(\mathbf{v}) : \mathbb{R}^N \to \mathbb{R}^N$. The components of this function are defined by the equation

$$\{G(\mathbf{v})\}_i = \mu(\mathrm{Lag}_i(\mathbf{v})) - \frac{e^{-v_i}}{\sum_{k=1}^N e^{-v_k}}, \quad i = 1, 2, \ldots, N\ . \tag{9}$$

In order to find the root of equation (9), Newton's method is used, which iteratively updates the estimate according to the formula:

$$\mathbf{v}^{(k+1)} = \mathbf{v}^{(k)} - [\nabla G(\mathbf{v}^{(k)})]^\dagger G(\mathbf{v}^{(k)})\ ,$$

where $[\nabla G]^\dagger$ denotes the inverse of the derivative ∇G over the orthogonal complement of the kernel of ∇G, specifically $\ker(\nabla G) = \mathbf{1}$.

3.2.2. Damped Newton's method

When the solution corresponds to a nested tessellation, we apply a damped Newton's method to ensure iterates remain within the nested domain.

Problem 3.8 (Damped Newton's method)**.** Starting from the current iterate $\mathbf{v}^{(k)}$, the update is given by

$$\mathbf{v}^{(k+1)} = \mathbf{v}^{(k)} - \frac{1}{2^s}[\nabla G(\mathbf{v}^{(k)})]^{\dagger} G(\mathbf{v}^{(k)}),$$

where $s \in \mathbb{Z}_{\geq 0}$ is the smallest integer such that the new iterate $\mathbf{v}^{(k+1)}$ lies inside the nested domain.

A practical criterion for verifying the **nestedness** condition is that the triple intersections of successive Laguerre cells are empty:

$$\mathrm{Lag}_{i-1}(\mathbf{v}) \cap \mathrm{Lag}_i(\mathbf{v}) \cap \mathrm{Lag}_{i+1}(\mathbf{v}) = \emptyset, \quad i = 2, 3, \ldots, N-1.$$

3.2.3. Nested methods

Unlike Newton's method, which seeks to solve the full nonlinear system simultaneously, the nested structure of the problem admits a sequential approach that greatly simplifies computation. This method exploits the monotonicity and ordered behavior of Laguerre cells under appropriate conditions, allowing the potentials to be computed one at a time in a forward pass. By fixing $v_1 = 0$ and constraining the range of C (as discussed in Rem. B.3), the vector $\mathbf{v}$ can be constructed component-by-component, preserving the nestedness of the Laguerre tessellation at each step.

This approach is particularly effective when the solution is known or expected to lie within the nested regime, as it avoids costly Jacobian inversions and improves numerical stability. The full formulation is provided below, with pseudocode in Algorithms 2, 3, and 4. Convergence guarantees are established in Appendix B.

Problem 3.9. Let $H(\mathbf{v}, C) : \mathbb{R}^{N+1} \to \mathbb{R}^N$ be defined component-wise by

$$\{H(\mathbf{v}, C)\}_i = \mu(\mathrm{Lag}_i(\mathbf{v})) - e^{C - v_i}, \quad i = 1, 2, \ldots, N. \tag{10}$$

The goal is to find $(\mathbf{v}^*, C^*)$ such that $H(\mathbf{v}^*, C^*) = \mathbf{0}$. Due to the nested structure of the solution, this system can be solved sequentially. Fixing $v_1 = 0$ and choosing $C \in (-\infty, 0]$, we compute each successive v_i by solving the scalar equations:

$$\begin{aligned}
\{H(\mathbf{v}, C)\}_1 &= \mu(\mathrm{Lag}_1(v_1, v_2)) - e^{C - v_1} = 0 \quad \Rightarrow \quad \text{Solve for } v_2, \\
\{H(\mathbf{v}, C)\}_2 &= \mu(\mathrm{Lag}_2(v_1, v_2, v_3)) - e^{C - v_2} = 0 \quad \Rightarrow \quad \text{Solve for } v_3, \\
&\vdots \\
\{H(\mathbf{v}, C)\}_{N-1} &= \mu\big(\mathrm{Lag}_{N-1}(v_{N-2}, v_{N-1}, v_N)\big) - e^{C - v_{N-1}} = 0 \quad \Rightarrow \quad \text{Solve for } v_N.
\end{aligned}$$

The final term of $H(\mathbf{v}^*, C^*)$ defines an error function:

$$\mathrm{Error}(C) := \{H(\mathbf{v}, C)\}_N = \mu(\mathrm{Lag}_N(v_{N-1}, v_N)) - e^{C - v_N}.$$

The root C^* satisfying $\mathrm{Error}(C^*) = 0$ can be found using bisection or Newton's method. The derivative $\frac{d}{dC}\mathrm{Error}(C)$ may be approximated using a Centered Finite Difference scheme.

Before addressing key practical considerations, we first examine potential challenges in implementing the sequential approach – particularly those related to the nestedness of the Laguerre tessellation and its sensitivity to the parameter C. The following remarks outline these issues and suggest strategies for resolving them effectively.

Remark 3.10 (Insufficiency of mass)**.** Recall the internal problem in the nested method: given a value of C, find $\mathbf{v}^*$ such that $v_1^* = 0$ and the following hold:

$$\begin{aligned}
\mu(\mathrm{Lag}_1(v_1^*, v_2)) &= e^{C-v_1^*} \Rightarrow v_2^* \quad \text{with} \quad \mu(\mathrm{Lag}_2(v_1^*, v_2^*)) > e^{C-v_2^*},\\
\mu(\mathrm{Lag}_2(v_1^*, v_2^*, v_3)) &= e^{C-v_2^*} \Rightarrow v_3^* \quad \text{with} \quad \mu(\mathrm{Lag}_3(v_1^*, v_2^*, v_3^*)) > e^{C-v_3^*},\\
\mu(\mathrm{Lag}_3(v_2^*, v_3^*, v_4)) &= e^{C-v_3^*} \Rightarrow v_4^* \quad \text{with} \quad \mu(\mathrm{Lag}_4(v_2^*, v_3^*, v_4^*)) > e^{C-v_4^*},\\
&\vdots\\
\mu(\mathrm{Lag}_{N-2}(v_{N-3}^*, v_{N-2}^*, v_{N-1})) &= e^{C-v_{N-2}^*} \Rightarrow v_{N-1}^*\\
&\quad \text{with} \quad \mu(\mathrm{Lag}_{N-1}(v_{N-3}^*, v_{N-2}^*, v_{N-1}^*)) > e^{C-v_{N-1}^*},\\
\mu(\mathrm{Lag}_{N-1}(v_{N-2}^*, v_{N-1}^*, v_N)) &= e^{C-v_{N-1}^*} \Rightarrow v_N^*,
\end{aligned}$$

defining the residual

$$\mathrm{Error}(C) := \mu(\mathrm{Lag}_N(v_{N-2}^*, v_{N-1}^*, v_N^*)) - e^{C-v_N^*}.$$

The strict inequalities ensure sufficient mass remains at each step to satisfy the next constraint. If any such condition fails, the exponential terms dominate and no feasible solution exists for the given C. In this case, C must be decreased. (This corresponds to the situation where $h(C) = -\infty$, as discussed in Appendix B.)

Beyond ensuring structural properties such as nestedness, it is essential to understand the behavior of the error function as the parameter C varies. This understanding is not only of theoretical interest but also plays a critical role in initializing and bounding root-finding algorithms, particularly bisection and Newton's method. The following remark analyzes the limiting behavior of $\mathrm{Error}(C)$ and its implications for numerical stability and algorithmic initialization.

Remark 3.11 (Limiting behavior of C)**.** Let y_i denote six target points distributed along a quarter circle. Figure 1 presents three plots illustrating the limiting behavior of the error function $\mathrm{Error}(C)$. Several key observations follow:

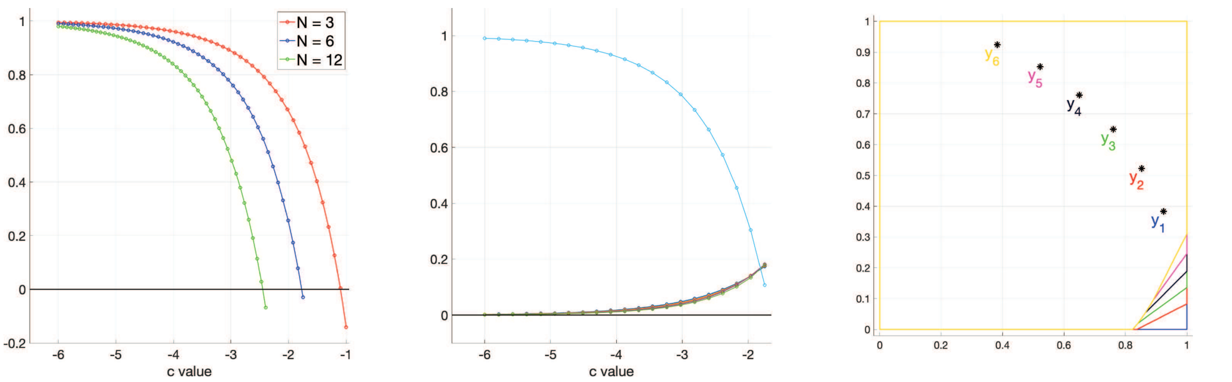
- As shown in Figure 1a, the slope of the error curve becomes steeper as the number of target points increases. This supports the hypothesis that the initial interval for bisection in solving Problem 3.9 becomes increasingly sensitive with finer discretization.
- There exists an upper bound $\overline{C} < 0$ such that $\mathrm{Error}(C)$ is undefined for all $C > \overline{C}$. This phenomenon, discussed in Remark 3.10, reflects the inability to construct a valid tessellation due to insufficient mass. Notably, $\overline{C}$ decreases as the number of target points increases.
- Regarding the lower limit, the error function appears to converge to 1 as $C \to -\infty$. This is due to the exponential terms vanishing in that limit, leading to $\mu(\mathrm{Lag}_i) \to 0$ for $i = 1, \dots, N-1$, and thus:

$$\mathrm{Error}(C) = \mu(\mathrm{Lag}_N) - e^{C-v_N} \to \mu(\mathrm{Lag}_N) \to 1 - \sum_{i=1}^{N-1} \mu(\mathrm{Lag}_i) \to 1.$$

  This asymptotic behavior is clearly visible in Figure 1b.
- Similar to the upper limit, there exists a lower bound $\underline{C} < 0$ below which the solution ceases to be nested; see Figure 1c. However, observing this behavior requires selecting very small values of C, which rarely occurs in practice. As a result, the lower bound $\underline{C}$ is typically less restrictive than the upper bound $\overline{C}$.

Remark 3.12 (Algorithmic simplification)**.** Recall the intermediate problems from Problem 3.9. For $i = 1, 2, \dots, N-1$ we have

$$\{H(\mathbf{v}, C)\}_i \;=\; \mu\big(\mathrm{Lag}_i(v_{i-1}, v_i, v_{i+1})\big) - e^{C-v_i} \;=\; 0 \quad \Longrightarrow \quad \text{Solve for } v_{i+1}\;.$$



(A) Error as a function of C (B) Measures as a function of C (C) Laguerre cells for small C value

FIGURE 1. Limiting behavior of the error function.

This is the semi-discrete optimal-transport (OT) problem with three points. When the equations are solved sequentially, the same condition can be written as

$$\{H(\mathbf{v}, C)\}_i \;=\; \mu\big(\mathrm{Lag}_i(v_i, v_{i+1})\big) - \sum_{j=1}^{i} e^{C-v_j} \;=\; 0 \quad \Longrightarrow \quad \text{Solve for } v_{i+1}, \quad i = 1, \dots, N-1 \,.$$

Thus, for each $i \in \{1, \dots, N-1\}$ the step reduces to a two-point semi-discrete OT problem with target locations $\{y_i, y_{i+1}\}$ and target masses $\big\{\sum_{j=1}^{i} e^{C-v_j},\, 1-\sum_{j=1}^{i} e^{C-v_j}\big\}$. Accordingly, each intermediate step is solved using the standard Newton's method for two-point semi-discrete OT, with the Jacobian matrix computed analytically.

3.3. Numerical examples

In this section, we present numerical results that illustrate the performance of the methods described above. We compare the convergence behavior and computational efficiency of standard and damped Newton's method and the bisection and Newton's methods applied to solve the nested problem. For the initial guesses, we used $\mathbf{v}^0 = \mathbf{0}$ for both the Newton and damped Newton's methods. For the nested method solved using the Newton algorithm, the initial value was set to $C^0 = -5$ (or $C^0 = -7.5$ when $N = 192$). When employing the bisection method for the nested formulation, the initial interval for C^0 was taken as $[-5,\ 0]$ (or $[-7.5,\ 0]$ when $N = 192$).

We consider the following examples:

$$\text{Straight line:} \quad y(t) = \left(\begin{smallmatrix} t \\ t \end{smallmatrix}\right), \qquad t \in \left[\frac{1}{10}, \frac{9}{10}\right], \tag{E1}$$

$$\text{Scaled parabola:} \quad y(t) = \begin{pmatrix} t \\ (\frac{t}{e})^2 \end{pmatrix}, \qquad t \in [0, 1], \tag{E2}$$

$$\text{Quarter circle:} \quad y(t) = \begin{pmatrix} \cos(t) \\ \sin(t) \end{pmatrix}, \qquad t \in \left[\frac{\pi}{8}, \frac{3\pi}{8}\right], \tag{E3}$$

$$\text{Parabola:} \quad y(t) = \left(\begin{smallmatrix} t \\ t^2 \end{smallmatrix}\right), \qquad t \in \left[\frac{1}{N+1}, \frac{N}{N+1}\right], \tag{E4}$$

where we fix the cost function to be $c(x, y) = \|x-y\|_2^2$. In this case, the Laguerre cells can be efficiently evaluated using a *lifting* algorithm (see Observation 7 in [22]).

For a given value of N, the target points are placed equidistantly over the respective parameter intervals.

TABLE 1. Computation time (and iterations) with error tolerance 10^{-5} and uniform measure $d\mu = dx_1 dx_2$.

Number of Target Points	$N = 3$	$N = 6$	$N = 12$	$N = 24$	$N = 48$	$N = 96$	$N = 192$
Example (E1)							
Solution	$C = -1.1532$	$C = -1.8491$	$C = -2.531$	$C = -3.2152$	$C = -3.9028$	$C = -4.5928$	$C = -5.2842$
Standard Newton	0.3627 (3)	0.4924 (3)	1.0829 (3)	2.5331 (3)	5.1 (3)	8.9818 (3)	17.271 (3)
Damped Newton	0.6109 (3, 0)	0.5101 (3, 0)	0.892 (3, 0)	2.029 (3, 0)	4.7076 (3, 0)	7.9782 (3, 0)	15.861 (3, 0)
Nested Bisection	10.554 (17)	25.705 (17)	102.21 (17)	312.90 (17)	626.22 (15)	970.42 (11)	3397.6 (18)
Nested Newton	12.827 (6)	23.728 (5)	87.669 (5)	217.80 (4)	806.45 (6)	1543.1 (5)	2820.6 (5)
Example (E2)							
Solution	$C = -1.1609$	$C = -1.8478$	$C = -2.5315$	$C = -3.2188$	$C = --3.9087$	$C = -4.6002$	$C = -5.2925$
Standard Newton	0.2193 (2)	0.3611 (2)	0.9704 (3)	1.9962 (3)	7.7496 (4)	20.228 (4)	96.869 (5)
Damped Newton	0.1972 (2, 0)	0.3451 (2, 0)	0.9618 (3, 0)	1.9753 (3, 0)	7.7644 (4, 0)	19.693 (4, 0)	96.482 (5, 0)
Nested Bisection	4.8444 (17)	10.615 (15)	40.960 (17)	93.164 (17)	198.24 (16)	447.34 (17)	860.33 (15)
Nested Newton	5.3689 (5)	15.739 (6)	41.827 (5)	63.049 (4)	179.00 (5)	415.40 (5)	716.60 (4)
Example (E3)							
Solution	$C = -1.0985$	$C = -1.7721$	$C = -2.4504$	$C = -3.1345$	$C = -3.8225$	$C = -4.5128$	$C = -5.2045$
Standard Newton	0.1178 (1)	0.3941 (3)	NAN	NAN	NAN	NAN	NAN
Damped Newton	0.1281 (1, 0)	0.4084 (3, 0)	NAN	NAN	NAN	NAN	NAN
Nested Bisection	10.009 (17)	22.413 (17)	40.900 (17)	90.623 (17)	182.91 (17)	368.51 (17)	796.60 (18)
Nested Newton	7.9133 (5)	20.906 (5)	44.879 (5)	73.941 (4)	197.68 (5)	393.94 (5)	843.77 (5)

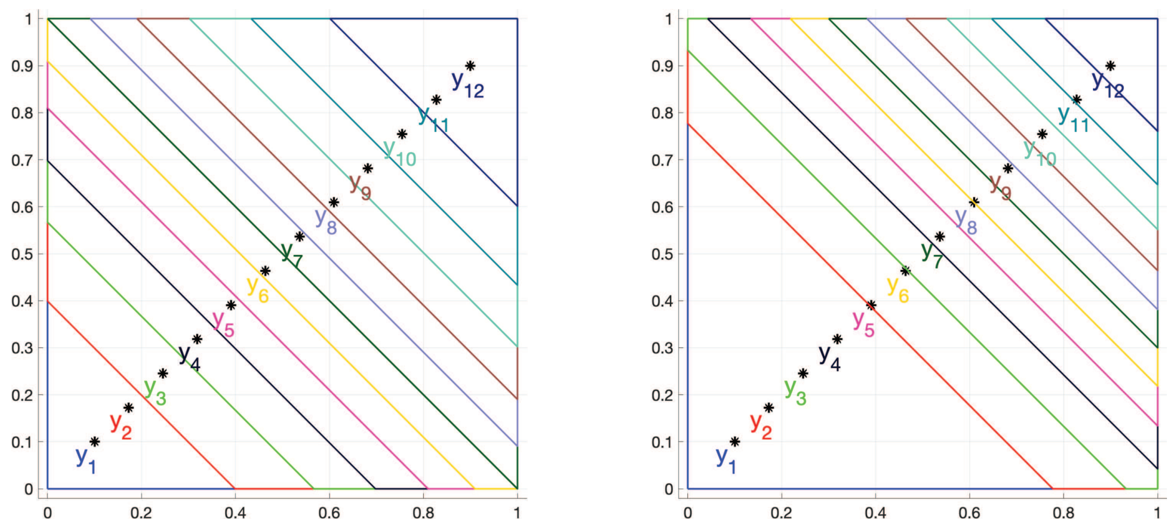
Notes. Green highlighted values indicate the best (lowest) computation time for the corresponding example and value of N. Red highlighted values indicate failure/non-convergence or cases where the relevant nestedness condition does not hold, so the corresponding method is not applicable.

TABLE 2. Computation time (and iterations) with error tolerance 10^{-5} and non-uniform measure $d\mu = 4x_1x_2 dx_1 dx_2$.

Number of Target Points	$N = 3$	$N = 6$	$N = 12$	$N = 24$	$N = 48$	$N = 96$	$N = 192$
Example (E1)							
Solution	$C = -1.403$	$C = -2.1136$	$C = -2.8009$	$C = -3.4869$	$C = -4.175$	$C = -4.8649$	$C = -5.5562$
Standard Newton	NAN	NAN	NAN	NAN	NAN	NAN	NAN
Damped Newton	0.3368 (3, 1)	1.1826 (5, 2)	1.8107 (5, 2)	4.2793 (5, 2)	7.4674 (4, 2)	19.189 (5, 2)	NAN
Nested Bisection	13.706 (17)	31.338 (16)	88.938 (17)	212.89 (14)	561.90 (16)	1146.9 (17)	2656.4 (18)
Nested Newton	9.8735 (4)	42.402 (6)	105.97 (5)	309.17 (5)	530.24 (4)	990.89 (4)	2535.3 (4)
Example (E2)							
Solution	$C = -1.3624$	$C = -2.0517$	$C = -2.7334$	$C = -3.4182$	$C = -4.1063$	$C = -4.7967$	$C = -5.4884$
Standard Newton	0.3614 (4)	NAN	NAN	NAN	NAN	NAN	NAN
Damped Newton	0.3791 (4, 0)	0.7854 (4, 1)	1.841 (4, 2)	4.3981 (5,3)	12.006 (6, 5)	25.653 (6, 6)	95.626 (6, 7)
Nested Bisection	10.332 (15)	32.209 (17)	76.697 (17)	203.54 (16)	469.83 (17)	927.91 (15)	2518.9 (18)
Nested Newton	7.4183 (3)	34.355 (6)	76.088 (5)	197.57 (5)	321.19 (3)	941.35 (4)	2081.0 (4)
Example (E4)							
Solution	$C = -1.3593$	$C = -2.1325$	$C = -2.8632$	$C = -3.573$	$C = -4.2735$	$C = -4.97$	
Standard Newton	NAN	NAN	NAN	NAN	NAN	NAN	NAN
Damped Newton	0.4449 (4, 2)	NAN	NAN	NAN	NAN	NAN	Solution is not nested
Nested Bisection	12.756 (15)	42.239 (17)	116.31 (16)	291.23 (16)	659.34 (17)	972.56 (13)	Solution is not nested
Nested Newton	11.921 (5)	27.639 (4)	126.32 (5)	310.75 (5)	614.98 (4)	603.78 (2)	Solution is not nested

Notes. Green highlighted values indicate the best (lowest) computation time for the corresponding example and value of N. Red highlighted values indicate failure/non-convergence or cases where the relevant nestedness condition does not hold, so the corresponding method is not applicable.

It is well known that the solution to Example (E1) is always nested. Moreover, as demonstrated in Example 3.3, the solution to Example (E2) with a uniform measure is also nested. Tables 1 and 2 report the compu-



(A) Uniform measure: $d\mu = dx_1 dx_2$ (B) Non-uniform measure: $d\mu = 4x_1x_2dx_1dx_2$

FIGURE 2. Solution for Example (E1) with 12 target points.

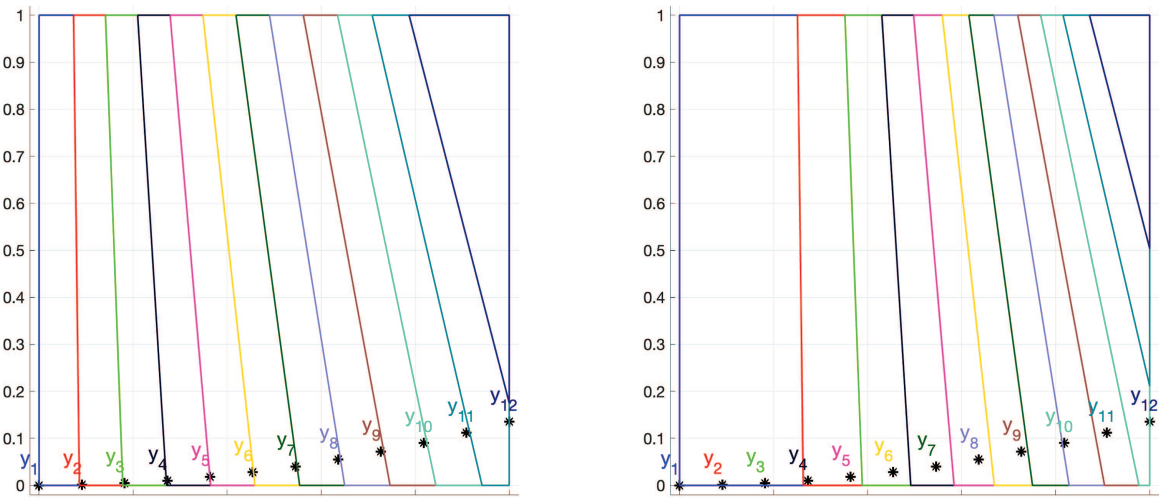
tation time and number of iterations for each method, together with the number of damping steps applied, for the uniform ($\mathrm{d}\mu = \mathrm{d}x_1\mathrm{d}x_2$) and non-uniform ($\mathrm{d}\mu = 4x_1x_2\mathrm{d}x_1\mathrm{d}x_2$) measures, respectively.

Across all examples, and for both uniform and non-uniform measures, the nested methods (damped Newton, bisection, and Newton) exhibit distinct advantages over the standard full-system Newton's method. In particular, the damped Newton's method is, on average, slightly faster than the standard Newton's method and exhibits more stable runtime behavior as N increases. Nested bisection and nested Newton incur higher computational costs due to the repeated solution of subproblems.

In terms of robustness, exploiting the nested structure is clearly advantageous. The standard Newton's method fails to converge in several examples – due to its sensitivity to the initial guess – particularly for larger N and non-uniform measures, whereas the damped Newton's method converges more reliably. When full-system Newton-type methods fail altogether, the sequential nested approaches remain robust and continue to produce solutions, as demonstrated in Table 2. Figures 2–4 illustrate the solutions for each example in the case of $N = 12$ target points. Notably, while the solution to Example (E3) is nested under a uniform measure (see Tab. 1 and Fig. 4a), it ceases to be nested under a non-uniform measure even with three target points, as shown in Figure 4b. A similar pattern is observed in Example (E4): the solution remains nested under a uniform measure but becomes non-nested when a non-uniform measure is applied with $N = 192$ target points, as illustrated in Table 2.

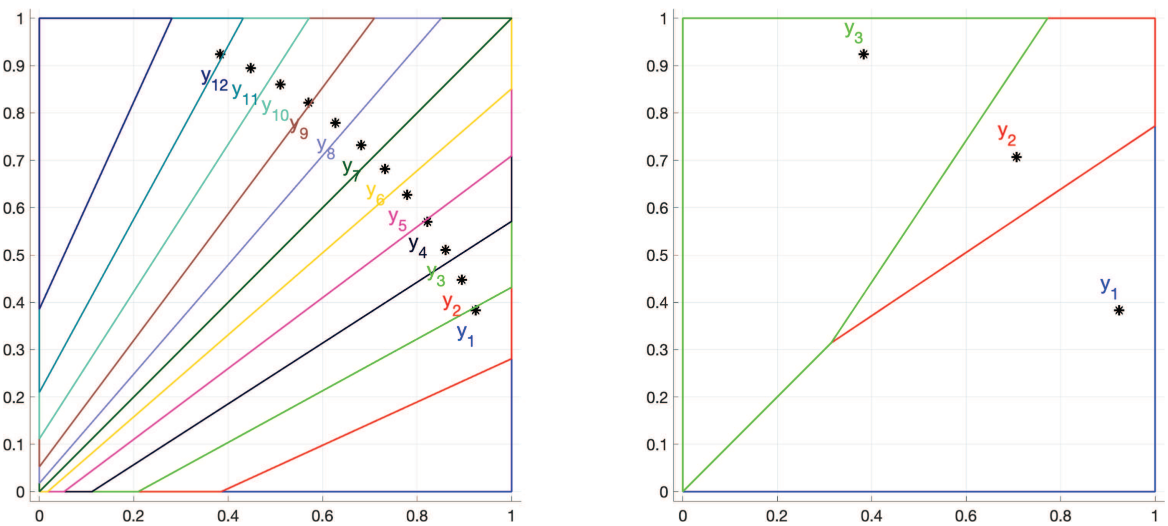
4. HEDONIC PRICING PROBLEM

We now turn to a class of problems that arise naturally in two-sided matching and market design, often referred to as hedonic pricing problems [18]. In these settings, agents from two distinct populations are matched indirectly through a shared distribution of outcomes. This leads to a variational formulation where the objective is to find a probability measure ν on a finite set $Y = \{y_1, \ldots, y_N\}$ that balances transport costs from both populations.



(A) Uniform measure: $d\mu = dx_1 dx_2$ (B) Non-uniform measure: $d\mu = 4x_1x_2 dx_1 dx_2$

FIGURE 3. Solution for Example (E2) with 12 target points.



(A) Uniform measure: $d\mu = dx_1 dx_2$, $N = 12$ (B) Non-uniform measure: $d\mu = 4x_1x_2 dx_1 dx_2$, $N = 3$

FIGURE 4. Solution for Example (E3) with 12 target points.

This corresponds to minimizing the sum of two optimal transport costs:

$$\min_{\nu \in \mathcal{P}(Y)} \{\mathcal{W}_{c_1}(\mu_1, \nu) + \mathcal{W}_{c_2}(\mu_2, \nu)\}, \tag{10}$$

where μ_1 and μ_2 are continuous distributions representing the two agent populations, and c_1, c_2 are their respective cost functions.

In [25], the notion of nestedness was extended to the two-population setting. In this work, we further extend it to the semi-discrete case. Specifically, we define a solution to be *hedonically nested* when both optimal transport plans from the respective populations to the common outcome distribution exhibit discrete nestedness.

Definition 4.1. The solution ν of problem (10) is *hedonically nested* if both (c_1, μ_1, ν) and (c_2, μ_2, ν) are discretely nested.

4.1. Problem statement

Throughout this section, we restrict ourselves to the setting where $X = X^1 = X^2 \subset \mathbb{R}^d$ and $c_1 = c_2 \in C^2(X \times Y)$ where c satisfies the generalized Spence-Mirlees, or twist, condition (found in, for example, [28]). In addition, we are given two absolutely continuous probability measures μ_1 and μ_2 supported on the bounded, open domain X, and a collection of target points $y_i \in Y \subset \mathbb{R}^d$, $i = 1, 2, \ldots, N$. From [13], it is known that the solution to the corresponding hedonic problem is unique, and the optimality condition

$$v^1 + v^2 = C$$

holds for some constant C, where v^1 and v^2 are the dual Kantorovich potentials corresponding to the optimal transport problems (c, μ_1, ν) and (c, μ_2, ν), respectively.

This motivates the following reformulation:

Problem 4.2. Given target points $y_i \in Y \subset \mathbb{R}^d$, $i = 1, \ldots, N$, a domain $X \subset \mathbb{R}^d$, and a cost function c, find $(\mathbf{v}, C) \in \mathbb{R}^N \times \mathbb{R}$ such that $v_1 = 0$ and

$$\mu_1(\mathrm{Lag}_i(\mathbf{v})) = \mu_2(\mathrm{Lag}_i(C - \mathbf{v})), \quad \text{for } i = 1, \ldots, N,$$

where μ_1 and μ_2 are probability measures on X, and

$$\mathrm{Lag}_i(\mathbf{v}) = \{x \in X \mid c(x, y_i) - v_i \le c(x, y_j) - v_j, \ \forall j \ne i\}.$$

This reformulation is justified by observing that for the equilibrium measure ν, optimal transport yields $\mu_1(\mathrm{Lag}_i(v^1)) = \mu_2(\mathrm{Lag}_i(v^2)) = \nu_i$ for each i, where v^1 and v^2 are the corresponding dual potentials. Using the optimality condition $v^1 + v^2 = C$, we can set $v^1 = \mathbf{v} = (v_i)_{i=1}^N$ and $v^2 = (C - v_i)_{i=1}^N$, thereby reducing the problem to a single potential vector $\mathbf{v}$ and a constant C.

The formulation in Problem 4.2 provides a natural framework for analyzing the structure of the hedonic pricing problem. Moreover, in the special case where the solution is hedonically nested, further simplifications are possible. Specifically, the nested structure of the optimal partitions allows for a sequential solution method, which underpins the nested formulation described in the next section.

Nested Algorithm for the Hedonic Problem. Assume that the solution to Problem 4.2, denoted by $(\mathbf{v}^*, C^*)$, is *hedonically nested*; that is, both Laguerre tessellations $\{\mathrm{Lag}_i(\mathbf{v}^*)\}_{i=1}^N$ and $\{\mathrm{Lag}_i(C^* - \mathbf{v}^*)\}_{i=1}^N$ are nested. Under this assumption, the hedonic pricing problem admits a sequential formulation:

$$\begin{aligned}
v_1 = 0 \ \Rightarrow \ \mu_1(\mathrm{Lag}_1(v_1, v_2)) &= \mu_2(\mathrm{Lag}_1(C - v_1, C - v_2)) \ \Rightarrow \ \text{Solve for } v_2, \\
\mu_1(\mathrm{Lag}_2(v_1, v_2, v_3)) &= \mu_2(\mathrm{Lag}_2(C - v_1, C - v_2, C - v_3)) \ \Rightarrow \ \text{Solve for } v_3, \\
&\vdots
\end{aligned}$$

$$\mu_1(\mathrm{Lag}_{N-1}(v_{N-2},v_{N-1},v_N)) = \mu_2(\mathrm{Lag}_{N-1}(C-v_{N-2},C-v_{N-1},C-v_N)) \;\Rightarrow\; \text{Solve for } v_N,$$
$$\mathrm{Error}(C,\mathbf{v}) := \mu_1(\mathrm{Lag}_N(v_{N-1},v_N)) - \mu_2(\mathrm{Lag}_N(C-v_{N-1},C-v_N)).$$

To justify this algorithm, we will show that if the first $N-1$ equations are satisfied, then the final error term $\mathrm{Error}(C,\mathbf{v})$ necessarily vanishes.

Proposition 4.3. *Assume that the solution to Problem 4.2, denoted by* $(\mathbf{v}^*,C^*)$, *is* hedonically nested. *Further, suppose that* $(\mathbf{v}^*,C^*)$ *satisfies the first* $N-1$ *equations:*

$$\begin{aligned}\mu_1(Lag_1(v_1^*,v_2^*)) &= \mu_2(Lag_1(C^*-v_1^*,C^*-v_2^*)),\\ \mu_1(Lag_2(v_1^*,v_2^*,v_3^*)) &= \mu_2(Lag_2(C^*-v_1^*,C^*-v_2^*,C^*-v_3^*)),\\ &\vdots\\ \mu_1(Lag_{N-1}(v_{N-2}^*,v_{N-1}^*,v_N^*)) &= \mu_2(Lag_{N-1}(C^*-v_{N-2}^*,C^*-v_{N-1}^*,C^*-v_N^*)).\end{aligned}$$

Then the error function is zero,

$$Error(C^*,\mathbf{v}^*) := \mu_1(Lag_N(v_{N-1}^*,v_N^*)) - \mu_2(Lag_N(C^*-v_{N-1}^*,C^*-v_N^*)) = 0.$$

Proof. Since the Laguerre tessellations $\{\mathrm{Lag}_i(\mathbf{v}^*)\}_{i=1}^N$ and $\{\mathrm{Lag}_i(C^*-\mathbf{v}^*)\}_{i=1}^N$ form partitions of X, we have:

$$\sum_{i=1}^N \mu_1(\mathrm{Lag}_i(\mathbf{v}^*)) = \sum_{i=1}^N \mu_2(\mathrm{Lag}_i(C^*-\mathbf{v}^*)) = 1.$$

As a result, we obtain:

$$\begin{aligned}\mathrm{Error}(C^*,\mathbf{v}^*) &= \mu_1(\mathrm{Lag}_N(v_{N-1}^*,v_N^*)) - \mu_2(\mathrm{Lag}_N(C^*-v_{N-1}^*,C^*-v_N^*))\\ &= \left(1-\sum_{i=1}^{N-1}\mu_1(\mathrm{Lag}_i(\mathbf{v}^*))\right) - \left(1-\sum_{i=1}^{N-1}\mu_2(\mathrm{Lag}_i(C^*-\mathbf{v}^*))\right)\\ &= \sum_{i=1}^{N-1}[\mu_2(\mathrm{Lag}_i(C^*-\mathbf{v}^*)) - \mu_1(\mathrm{Lag}_i(\mathbf{v}^*))].\end{aligned}$$

By assumption, the terms in the sum cancel individually, yielding $\mathrm{Error}(C^*,\mathbf{v}^*)=0$. □

Remark 4.4. By Proposition 4.3, the hedonic pricing problem has rank $N-1$ with $N+1$ variables. As a result, when the solution is hedonically nested, one can freely specify two degrees of freedom - typically C and v_1 - and then sequentially solve the remaining $N-1$ scalar equations to recover the full solution. In our numerical experiments, we fix $C=0$ and $v_1=0$. Once the remaining components of $\mathbf{v}$ are determined through this sequential procedure, the uniqueness of the solution ensures that the induced measure ν is the unique minimizer of the hedonic pricing problem.

Remark 4.5 (Algorithmic simplification)**.** Similar to Remark 3.12, one can derive a two-point formulation for the intermediate problems arising in the hedonic problem. Recall that for $i=1,2,\dots,N-1$ we have

$$\mu_1\big(\mathrm{Lag}_i(v_{i-1},v_i,v_{i+1})\big) \;=\; \mu_2\big(\mathrm{Lag}_i(C-v_{i-1},C-v_i,C-v_{i+1})\big) \quad\Longrightarrow\quad \text{Solve for } v_{i+1}.$$

When the equations are solved sequentially, the same condition can be written as

$$\mu_1\big(\mathrm{Lag}_i(v_i,v_{i+1})\big) \;=\; \mu_2\big(\mathrm{Lag}_i(C-v_i,C-v_{i+1})\big) \quad\Longrightarrow\quad \text{Solve for } v_{i+1}.$$

Hence, for each $i\in\{1,\dots,N-1\}$ the step reduces to a two-point semi-discrete variational OT problem with target locations $\{y_i,y_{i+1}\}$.

TABLE 3. Computation time of different methods (error tolerance is 10^{-7}, $\mathrm{d}\mu_1(x) = \mathrm{d}x_1\mathrm{d}x_2$, $\mathrm{d}\mu_2(x) = 4x_1x_2\mathrm{d}x_1\mathrm{d}x_2$).

$v^0 = 0$; $v_1 = 0$, $C = 0$	$N = 3$	$N = 6$	$N = 12$	$N = 24$	$N = 48$	$N = 96$	$N = 192$
Straight line: $y(t) = \binom{t}{t}$, $t \in [\frac{1}{10}, \frac{9}{10}]$							
Standard Newton	0.7275 (3)	1.2052 (3)	3.1526 (4)	6.4427 (4)	17.526 (4)	25.977 (4)	NAN
Damped Newton	0.6569 (3, 0)	1.1642 (3, 0)	3.2731 (4, 0)	6.4408 (4, 0)	17.595 (4, 0)	25.743 (4, 0)	NAN
Sequential Method	1.4969	4.0464	8.4965	21.731	46.262	100.21	211.59
Curve: $y(t) = \begin{pmatrix} t \\ t^{1.5} \end{pmatrix}$, $t \in [\frac{1}{N+1}, \frac{N}{N+1}]$							
Standard Newton	0.7728 (3)	2.7488 (4)	3.3387 (4)	10.565 (5)	17.626 (5)	37.749 (5)	112.81 (6)
Damped Newton	0.6549 (3, 0)	2.8531 (4, 0)	3.1942 (4, 0)	10.806 (5, 0)	17.530 (5, 0)	38.288 (5, 0)	113.94 (6, 0)
Sequential Method	0.9958	3.6050	7.7683	19.526	45.524	103.48	222.94
Scaled Parabola: $y(t) = \begin{pmatrix} t \\ (\frac{t}{e})^2 \end{pmatrix}$, $t \in [0, 1]$							
Standard Newton	0.7188 (3)	1.2179 (3)	2.5451 (3)	4.9134 (3)	14.817 (4)	NAN	135.49 (4)
Damped Newton	0.6423 (3, 0)	1.306 (3, 0)	2.5061 (3, 0)	4.818 (3, 0)	14.089 (4, 0)	Solution is not *hedonically* nested	134.91 (4, 0)
Sequential Method	0.9839	2.5997	6.0309	14.897	35.801	Solution is not *hedonically* nested	168.27

Notes. Green highlighted values indicate the best (lowest) computation time for the corresponding example and value of N. Red highlighted values indicate failure/non-convergence or cases where the relevant nestedness condition does not hold, so the corresponding method is not applicable.

4.2. Numerical results

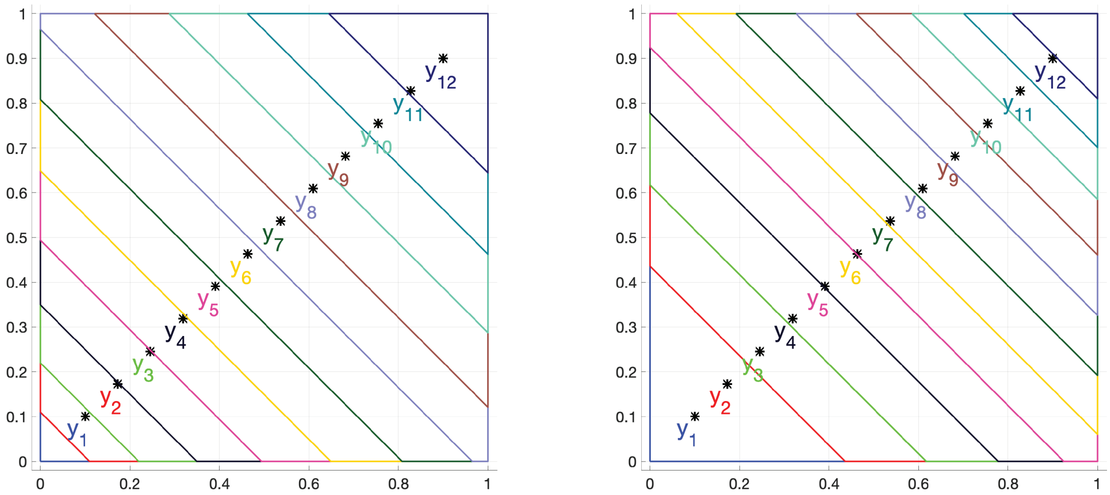
We present numerical experiments comparing Newton's method and the nested method for solving the hedonic pricing problem. In the following examples, we focus on the case $c(x, y) = \|x - y\|_2^2$, where $x \in X \subseteq \mathbb{R}^2$ and the discrete variable y lies along a curve $y(t) \in \mathbb{R}^2$. We compare the convergence behavior and computational efficiency of three methods: full-step Newton, damped Newton, and the sequential method.

Remark 4.6. Proposition 4.3 implies that the nested algorithm reduces the hedonic-pricing problem to solving $N - 1$ scalar equations, each depending only on previously computed values. Moreover, Remark 4.5 yields a further reduction: at each step one needs only to solve a two-point variational semi-discrete OT problem. Consequently, the sequential method solves these subproblems using the standard Newton's method, with the Jacobian computed analytically and tailored to the hedonic formulation.

Table 3 compares the performance of the full-system Newton's method with the nested strategies – *damped Newton* and *Sequential method* – across a range of matching scenarios.

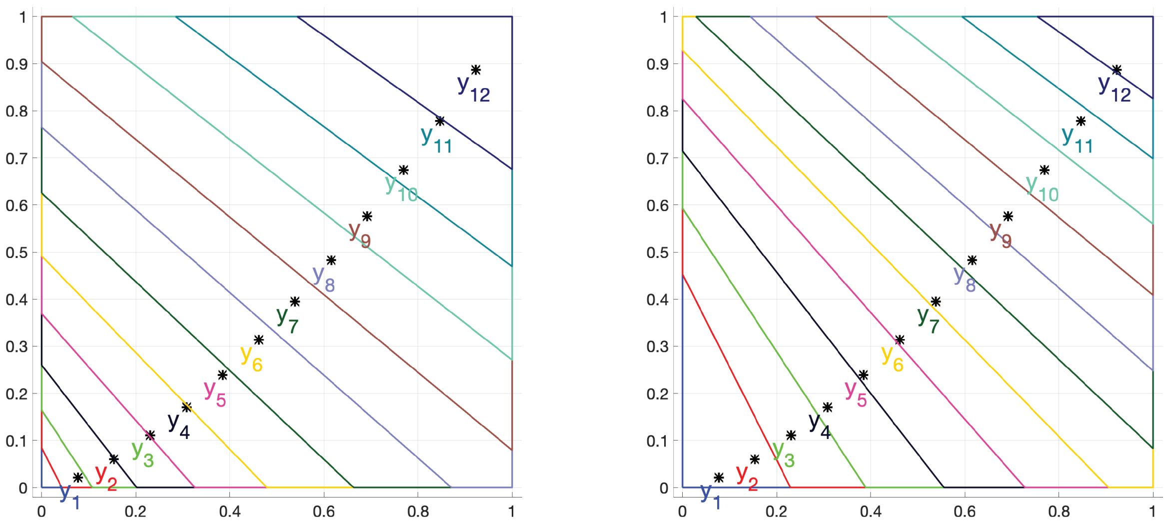
Figures 5–7 illustrate cases where the solution is hedonically nested, as the region boundaries under both cases (c, μ_1, ν) and (c, μ_2, ν) remain non-intersecting within the domain X. In contrast, for target points distributed along the scaled parabola $x_2 = \left(\frac{x_1}{e}\right)^2$, the solution is nested for some values of N, fails at $N = 96$, and is recovered at $N = 192$, illustrating the somewhat counterintuitive fact that nestedness is a nontrivial, problem-dependent property that can vary non-monotonically with problem size.

Overall, the nested damped Newton method is, on average, a bit faster than the full-system Newton's method. In addition, the sequential method exhibited the greatest robustness: it produced a solution in every example for which the solution is nested, including cases in which the full-system Newton (standard or damped) failed.



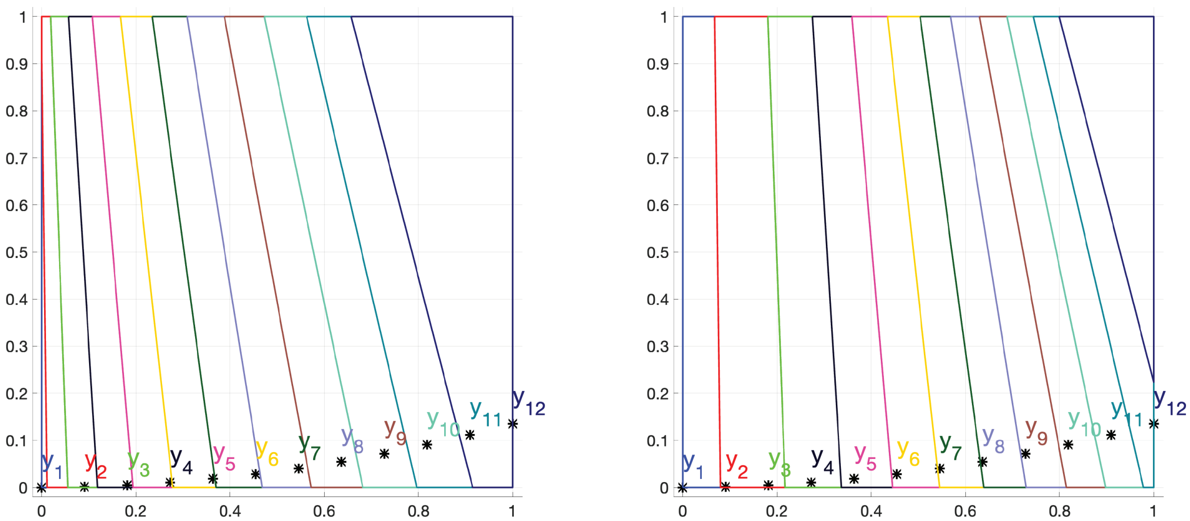
(A) $d\mu_1 = dx_1dx_1$ (B) $d\mu_2 = 4x_1x_2dx_1dx_1$

FIGURE 5. Target points on a straight line $x_2 = x_1$, $N = 12$.



(A) $d\mu_1 = dx_1dx_1$ (B) $d\mu_2 = 4x_1x_2dx_1dx_1$

FIGURE 6. Target points on a curve $x_2 = x_1^{1.5}$, $N = 12$.



(A) $d\mu_1 = dx_1 dx_1$ (B) $d\mu_2 = 4x_1 x_2 dx_1 dx_1$

FIGURE 7. Target points on a scaled parabola $x_2 = \left(\frac{x_1}{e}\right)^2$, $N = 12$.

ACKNOWLEDGMENTS

The work of OAH was completed in partial fulfillment of the requirements for a doctoral degree in mathematics at the University of Alberta. BP is pleased to acknowledge the support of Natural Sciences and Engineering Research Council of Canada Discovery Grant number 04864-2024. DO gratefully acknowledge that this research was supported in part by the Pacific Institute for the Mathematical Sciences.

REFERENCES

[1] O. Abdul Halim and B. Pass, Multi-to one-dimensional screening and semi-discrete optimal transport. Preprint arXiv:2506.21740 (2025).

[2] M. Agueh and G. Carlier, Barycenters in the Wasserstein space. *SIAM J. Math. Anal.* **43** (2011) 904–924.

[3] F. Aurenhammer, Power diagrams: properties, algorithms and applications. *SIAM J. Comput.* **16** (1987) 78–96.

[4] A. Blanchet and G. Carlier, From Nash to Cournot–Nash equilibria via the Monge–Kantorovich problem. *Philos. Trans. R. Soc. A* **372** (2014) 20130398.

[5] A. Blanchet and G. Carlier, Remarks on existence and uniqueness of Cournot–Nash equilibria in the non-potential case. *Math. Financ. Econ.* **8** (2014) 417–433.

[6] A. Blanchet and G. Carlier, Optimal transport and Cournot–Nash equilibria. *Math. Oper. Res.* **41** (2016) 125–145.

[7] A. Blanchet, G. Carlier and L. Nenna, Computation of Cournot–Nash equilibria by entropic regularization. *Vietnam J. Math.* (2017).

[8] D.P. Bourne and S.M. Roper, Centroidal power diagrams, Lloyd's algorithm, and applications to optimal location problems. *SIAM J. Numer. Anal.* **53** (2015) 2545–2569.

[9] G. Buttazzo and F. Santambrogio, A model for the optimal planning of an urban area. *SIAM J. Math. Anal.* **37** (2005) 514–530.

[10] G. Buttazzo and F. Santambrogio, A mass transportation model for the optimal planning of an urban region. *SIAM Rev.* **51** (2009) 593–610.

[11] G. Carlier and I. Ekeland, The structure of cities. *J. Glob. Optim.* **29** (2004) 371–376.

[12] G. Carlier and F. Santambrogio, A variational model for urban planning with traffic congestion. *ESAIM: Control Optim. Calc. Var.* **11** (2005) 595–613.

[13] P.-A. Chiappori, R. McCann and L. Nesheim, Hedonic price equilibria, stable matching and optimal transport; equivalence, topology and uniqueness. *Econ. Theory* **42** (2010) 317–354.

[14] P.-A. Chiappori, R.J. McCann and B. Pass, Multi- to one-dimensional optimal transport. *Commun. Pure Appl. Math.* **70** (2017) 2405–2444.

[15] F. De Gournay, J. Kahn and L. Lebrat, Differentiation and regularity of semi-discrete optimal transport with respect to the parameters of the discrete measure. *Numer. Math.* **141** (2019) 429–453.

[16] L. Dieci and D. Omarov, Solving semi-discrete optimal transport problems: star shapedeness and Newton's method. *Numer. Algorithms* **99** (2025) 949–1004.

[17] L. Dieci and J.D. Walsh III, The boundary method for semi-discrete optimal transport partitions and Wasserstein distance computation. *J. Comput. Appl. Math.* **353** (2019) 318–344.

[18] I. Ekeland, An optimal matching problem. *ESAIM: Control Optim. Calc. Var.* **11** (2005) 57–71.

[19] A. Galichon, Optimal Transport Methods in Economics. Princeton University Press, Princeton (2016).

[20] V.N. Hartmann and D. Schuhmacher, Semi-discrete optimal transport: a solution procedure for the unsquared Euclidean distance case. *Math. Method. Oper. Res.* (2020) 1–31.

[21] J. Kitagawa, Q. Mérigot and B. Thibert, Convergence of a Newton algorithm for semi-discrete optimal transport. *J. Eur. Math. Soc.* **21** (2019) 2603–2651.

[22] B. Lévy, A numerical algorithm for L_2 semi-discrete optimal transport in 3D. *ESAIM: Math. Model. Numer. Anal.* **49** (2015) 1693–1715.

[23] R.J. McCann, Polar factorization of maps on Riemannian manifolds. *Geom. Funct. Anal.* **11** (2001) 589–608.

[24] Q. Merigot and B. Thibert, Optimal transport: discretization and algorithms, in *Handbook of Numerical Analysis*, Vol. 22 (2021) 133–212.

[25] L. Nenna and B. Pass, Variational problems involving unequal dimensional optimal transport. *J. Math. Pures Appl.* **139** (2020) 83–108.

[26] L. Nenna and B. Pass, A note on Cournot–Nash equilibria and optimal transport between unequal dimensions, in Optimal Transport Statistics for Economics and Related Topics (2023) 117–130.

[27] G. Peyré and M. Cuturi, Computational optimal transport. *Found. Trends Mach. Learn.* **11** (2019) 355–607.

[28] F. Santambrogio, Optimal Transport for Applied Mathematicians: Calculus of Variations, PDEs, and Modeling. Birkhäuser (2015).

[29] C. Villani, Optimal Transport: Old and New. Springer, Berlin (2009).

Appendix A. Algorithm pseudocodes

Algorithm 1. *DampedNewton* – Damped Newton's method.

Input: Ω, $\mu(x)$, $\mathbf{v^0}$, y_i, $i = 1, 2 \ldots, N$, `MAXIT`, `TOL` ▷ Default: $\mathbf{v^0} = \mathbf{0}$, `MAXIT` = 20, `TOL` = 10^{-5}

  $\{err^0\}_i \leftarrow \mu(Lag_i(y, \mathbf{v^0})) - \frac{e^{-v_i^0}}{\sum_{k=1}^N e^{-v_k^0}}$, $\forall i$
  for $j \leftarrow 0$ to `MAXIT` **do**
    $H^j \leftarrow \nabla G(\mathbf{v}^{(j)})$
    $\mathbf{s} \leftarrow -(H^j)^\dagger \mathbf{err}^j$, $\mathbf{v}^{j+1} \leftarrow \mathbf{v}^j + \mathbf{s}$
    while $\mathrm{Lag}(y_{i-1}) \cap \mathrm{Lag}(y_i) \cap \mathrm{Lag}(y_{i+1}) \neq \emptyset, \quad i = 2, 3, \ldots, N-1$ **do** ▷ Damping if not nested
      $\mathbf{s} \leftarrow \frac{1}{2}\mathbf{s}$, $\mathbf{v}^{j+1} \leftarrow \mathbf{v}^j + s$
    end while
    $\{err^{j+1}\}_i \leftarrow \mu(Lag_i(y, \mathbf{v^{j+1}})) - \frac{e^{-v_i^{j+1}}}{\sum_{k=1}^N e^{-v_i^{j+1}}}$, $\forall i$
    if $||\mathbf{err}^{j+1}||_\infty \leq$ `TOL` **then**
      return $\mathbf{v^{j+1}}$
    end if
  end for

Algorithm 2. *ErrorFunc* – Error computation.

Input: $\{\Omega,\ \mu(x)\}$, y_i, $i = 1, 2, \ldots, N$, C
  $v_1 \leftarrow 0$
  for $j \leftarrow 2$ to N **do**
    $\mu(\mathrm{Lag}_{j-1}(v_1, \ldots, v_j) - e^{C^* - v_{j-1}} = 0 \quad \Rightarrow \quad$ Solve for v_j
    if $e^{C^* - v_j} - \mu(\mathrm{Lag}_j(v_1, \ldots, v_j) > 0$ **then return** $Error \leftarrow NAN$ ▷ Remark 3.10
    end if
  end for
  $Error \leftarrow \mu(\mathrm{Lag}_N(v_1, \ldots, v_N) - e^{C^* - v_N}$

Algorithm 3. *NestBisection* – Nested bisection on C value.

Input: $\{\Omega,\ \mu(x)\}$, y_i, $i = 1, 2, \ldots, N$, C^0, C^1, `TOL` ▷ Default: `TOL` = 10^{-5}
  while $|Error| >$ `TOL` **do**
    $C^* \leftarrow \frac{1}{2}(C^0 + C^1)$
    $Error \leftarrow ErrorFunc(C^*)$ ▷ Algorithm 2
    if $Error > 0$ **then**
      $C^0 \leftarrow C^*$
    else if $Error > 0 \vee Error = NAN$ **then**
      $C^1 \leftarrow C^*$
    end if
  end while

Algorithm 4. *NestNewton* – Nested Newton on C value.

```
Input: {Ω, μ(x)}, y_i, i = 1,2,...,N, C^0, TOL          ▷ Default: TOL = 10^{-5}
  h ← eps^{1/3}, Error ← ErrorFunc(C^0)                  ▷ Algorithm 2
  while |Error| > TOL do
      H ← (ErrorFunc(C^0+h) − ErrorFunc(C^0−h)) / (2h)   ▷ Centered Finite Difference Scheme
      s ← −Error/H, C^1 ← C^0 + s
      while C^1 ≥ 0 do                                   ▷ Solution is always negative
          s ← (1/2)s, C^1 ← C^0 + s
      end while
      Error ← ErrorFunc(C^1)
      while Error = NAN do                               ▷ Remark 3.10
          s ← (1/2)s, C^1 ← C^0 + s, Error ← ErrorFunc(C^1)
      end while
      C^0 ← C^1,
  end while
```

Appendix B. Convergence of the nested algorithm of the internal energy problem

While Subsection 3.2 focused on the numerical implementation, we now shift to a more theoretical perspective. Accordingly, we return to the original notation used in Section 2, which is better suited for analytical arguments. In this section, we will prove the convergence of the numerical algorithm introduced in Subsection 3.2 for computing the solution of (4) under the assumption that the solution exhibits discrete nestedness.

We begin by describing the algorithm in full detail. It aims to find a measure $\nu \in \mathcal{P}(Y)$ that satisfies the first-order optimality condition (5) associated with (4). To solve this system, we exploit the correspondence between ν and v. Specifically, we express ν_i as $\nu_i = (f')^{-1}(C - v_i)$, and associate each ν_i with a subset $X_i(v) \subset X$ such that $\mu(X_i(v)) = \nu_i$. For each chosen C, we determine $v^C = (v_i^C)_{i=1}^N$ using the relation between $\nu_i^C = \mu(X_i(v^C))$ and $(f')^{-1}(C - v_i^C)$. We then define an error function $h(C)$ which we will prove is monotone. Consequently, the solution can be computed using the bisection method to solve $h(C) = 0$.

We define $h(C)$ as follows:

For each C, we set $v_1^C = 0$ and define $X_1(v^C) = X_{\geq}^N(y_1, k_1)$ such that

$$\mu(X_{\geq}^N(y_1, k_1)) = \min\{1, (f')^{-1}(C - v_1^C)\}.$$

Then, for each $i \geq 2$, we set $v_i^C = v_{i-1}^C + k_{i-1}$ and define $\overline{k}_i$ such that

$$\mu(X_{\geq}^N(y_i, \overline{k}_i)) = (f')^{-1}(C - v_i^C) + \sum_{j=1}^{i-1} \mu(X_j(v^C)).$$

If $\overline{k}_i < k_{\max}(y_{i-1}, k_{i-1})$, we set $k_i = \overline{k}_i$; otherwise, we set $k_i = k_{\max}(y_{i-1}, k_{i-1})$. We then define

$$X_i(v^C) = X_{\geq}^N(y_i, k_i) \setminus X_{\geq}^N(y_{i-1}, k_{i-1}).$$

This process continues until we reach some $m^C = \max\{i : 1 - \mu(X_{\geq}^N(y_{i-2}, k_{i-2})) \geq (f')^{-1}(C - v_{i-1}^C)\} \leq N$, at which point we set $X_{m^C}(v^C) = X \setminus X_{\geq}^N(y_{m^C-1}, k_{m^C-1})$, and $X_i(v^C) = \emptyset$ for all $i > m^C$. If $m^C < N$, we set $h(C) = -\infty$; otherwise, we set $h(C) = 1 - \mu(X_{\geq}^N(y_{N-1}, k_{N-1})) - (f')^{-1}(C - v_N^C)$. Also, we set $\nu_i^C = \mu(X_i(v^C))$ for all $1 \leq i \leq m^C$ and $\nu^C = \sum_{i=1}^{m^C} \nu_i^C \delta_{y_i}$.

Remark B.1. If $k_i < k_{\max}(y_{i-1}, k_{i-1})$, then $\nu_i^C = \mu(X_i(v^C)) = (f')^{-1}(C - v_i^C)$. Otherwise, $\nu_i^C = \mu(X_i(v^C)) \geq (f')^{-1}(C - v_i^C)$.

Lemma B.2. *The function $h(C)$ is decreasing in C.*

Proof. Let $C < C'$ and let $v^C = (v_i^C)_{i=1}^N$ and $v^{C'} = (v_i^{C'})_{i=1}^N$ be associated with $(X_i(v^C))_{i=1}^N$ and $(X_i(v^{C'}))_{i=1}^N$, respectively. We adapt $v_i^{C'}$ and k_i' such that they are defined similar to v_i^C and k_i respectively, when C is replaced by C'.

By the strict monotonicity of $(f')^{-1}$, we obtain

$$\mu(X_1(v^C)) = (f')^{-1}(C - v_1^C) = \min\{1, (f')^{-1}(C - v_1^C)\} < \min\{1, (f')^{-1}(C' - v_1^{C'})\} = \mu(X_1(v^{C'}))$$

in the case where $(f')^{-1}(C - v_1^C) < 1$ since $v_1^C = v_1^{C'} = 0$. This yields

$$\mu(X_{\geq}^N(y_1, k_1)) = \mu(X_1(v^C)) < \mu(X_1(v^{C'})) = \mu(X_{\geq}^N(y_1, k_1')),$$

and since $\mu(X_{\geq}^N(y_1, k))$ is a decreasing function of k, this implies $k_1 > k_1'$.

On the other hand, if $(f')^{-1}(C - v_1^C) \geq 1$, then $\mu(X_1(v^C)) = \mu(X_1(v^{C'})) = 1$, and hence $h(C) = h(C') = -\infty$, so the comparison of k_1 and k_1' becomes irrelevant in this case.

Proceeding to the case $(f')^{-1}(C - v_1^C) < 1$, we have $v_2^C = k_1 + v_1^C > k_1' + v_1^{C'} = v_2^{C'}$.

We now prove that $k_{i+1} > k_{i+1}'$ and $v_{i+2}^C > v_{i+2}^{C'}$ whenever $k_i > k_i'$ and $v_{i+1}^C > v_{i+1}^{C'}$. Suppose that $k_{i+1} \leq k_{i+1}'$, then $X_{\geq}^N(y_{i+1}, k_{i+1}') \subseteq X_{\geq}^N(y_{i+1}, k_{i+1})$ and so

$$\begin{aligned}\mu(X_{i+1}(v^C)) &= \mu(X_{\geq}^N(y_{i+1}, k_{i+1}) \setminus X_{\geq}^N(y_i, k_i)) \\ &\geq \mu(X_{\geq}^N(y_{i+1}, k_{i+1}') \setminus X_{\geq}^N(y_i, k_i')) \\ &= \mu(X_{i+1}(v^{C'})) \geq (f')^{-1}(C' - v_{i+1}^{C'}) > (f')^{-1}(C - v_{i+1}^C),\end{aligned}$$

since $C' - v_{i+1}^{C'} > C - v_{i+1}^C$. Consequently, $k_{i+1} = k_{\max}(y_i, k_i)$. Since $k_i > k_i'$, we have

$$X_{\geq}^N(y_i, k_i) \subset X_{\geq}^N(y_i, k_i') \subseteq X_{\geq}^N(y_{i+1}, k_{\max}(y_i, k_i')) \subseteq X_{\geq}^N(y_{i+1}, k_{i+1}'),$$

which implies that $k_{i+1} = k_{\max}(y_i, k_i) \geq k_{i+1}'$. Thus, $k_{i+1}' = k_{\max}(y_i, k_i) = k_{i+1}$.

We claim that there exists $x \in \overline{X_{=}^N(y_i, k_i)} \cap \overline{X_{=}^N(y_{i+1}, k_{\max}(y_i, k_i))}$. Consider a sequence k^n that decreases to k_{i+1}. By the definition of $k_{\max}(y_i, k_i)$, we have $X_{\geq}^N(y_i, k_i) \not\subset X_{>}^N(y_{i+1}, k^n)$, so we take $x_n \in X_{\geq}^N(y_i, k_i) \setminus X_{>}^N(y_{i+1}, k^n)$. Then,

$$c(x_n, y_{i+1}) - c(x_n, y_i) \geq k_i \quad \text{and} \quad c(x_n, y_{i+2}) - c(x_n, y_{i+1}) \leq k^n.$$

Taking a subsequential limit $\overline{x} \in \overline{X}$ of (x_n) gives

$$c(\overline{x}, y_{i+1}) - c(\overline{x}, y_i) \geq k_i \quad \text{and} \quad c(\overline{x}, y_{i+2}) - c(\overline{x}, y_{i+1}) \leq k_{\max}(y_i, k_i).$$

Since $X_{\geq}^N(y_i, k_i) \subset X_{\geq}^N(y_{i+1}, k_{\max}(y_i, k_i))$, we conclude that $\overline{x} \in \overline{X_{=}^N(y_i, k_i)} \cap \overline{X_{=}^N(y_{i+1}, k_{\max}(y_i, k_i))}$, proving our claim.

Since $\overline{x} \in \overline{X_{=}^N(y_i, k_i)} \subset \overline{X_{>}^N(y_i, k_i')} \setminus \overline{X_{=}^N(y_i, k_i')}$, we have $c(\overline{x}, y_{i+1}) - c(\overline{x}, y_i) > k_i'$ and by the continuity of $c(\cdot, y_{i+1}) - c(\cdot, y_i)$ at $\overline{x}$, there exists a neighborhood U of $\overline{x}$ such that $c(x, y_{i+1}) - c(x, y_i) > k_i'$ for all $x \in U \cap X$. Also, by (H1) and the fact $\overline{x} \in \overline{X_{=}^N(y_{i+1}, k_{\max}(y_i, k_i))}$, there exists $x \in U \cap X \subseteq X_{\geq}^N(y_i, k_i')$ such that $c(x, y_{i+2}) - c(x, y_{i+1}) < k_{\max}(y_i, k_i)$, which contradicts $X_{\geq}^N(y_i, k_i') \subseteq X_{\geq}^N(y_{i+1}, k_{\max}(y_i, k_i))$. Hence, $k_{i+1}' \neq k_{\max}(y_i, k_i)$, so $k_{i+1}' < k_{i+1}$ and $v_{i+2}^C = k_{i+1} + v_{i+1}^C > k_{i+1}' + v_{i+1}^{C'} = v_{i+2}^{C'}$.

Let $C_0 \in \Big\{C \mid \sup\{i \mid \mu(X_i(v^C)) > 0\} = N\Big\}$, then for all $1 \leq i \leq N-2$ we have

$$1 - \mu(X_{\geq}^N(y_i, k_i)) - (f')^{-1}(C_0 - v_{i+1}^C) > 0.$$

Since v_i^C and k_i decrease as C increases, $1 - \mu(X_{\geq}^N(y_i, k_i))$ decreases and $(f')^{-1}(C - v_{i+1}^C)$ increases, which implies $1 - \mu(X_{\geq}^N(y_i, k_i)) - (f')^{-1}(C - v_{i+1}^C) > 0$ decreases for all $1 \leq i \leq N-2$. In particular, we have $1-\mu(X_{\geq}^N(y_{N-2}, k_{N-2}))-(f')^{-1}(C-v_{N-1}^C) > 0$ for all $C \leq C_0$ and so $\mu(X_N(v^C)) = 1-\mu(X_{\geq}^N(y_{N-1}, k_{N-1})) > 0$ for all $C \leq C_0$. We conclude that

$$(-\infty, C^*) \subseteq \left\{C \mid \sup\{i \mid \mu(X_i(v^C)) > 0\} = N\right\} \subseteq (-\infty, C^*]$$

for $C^* = \sup\left\{C \mid \sup\{i \mid \mu(X_i(v^C)) > 0\} = N\right\}$.

When $C > C^*$, we have $h(C) = -\infty$. When $C = C^*$, then

$$h(C) = \begin{cases} -\infty & \text{if } m^C < N \\ 1 - \mu(X_{\geq}^N(y_{N-1}, k_{N-1})) - (f')^{-1}(C - v_N^C) & \text{if } m^C = N \end{cases}$$

where m^C is defined as above. When $C < C^*$, since v_i^C and k_i decrease as C increases, the quantity $1 - \mu(X_{\geq}^N(y_{N-1}, k_{N-1}))$ decreases, while $(f')^{-1}(C-v_N^C)$ increases. This implies that $h(C)$ is decreasing, completing the proof. □

Remark B.3. By Theorem 3.1 we obtain

$$(f')^{-1}(J_{-M_c|y_i-y_1|}(1) - M_c|y_i - y_1|) \leq \nu_i^C = (f')^{-1}(C - v_i^C) \leq (f')^{-1}(J_{M_c|y_i-y_1|}(1) + M_c|y_i - y_1|),$$

which implies

$$J_{-M_c|y_i-y_1|}(1) - M_c|y_i - y_1| \leq C - v_i^C \leq J_{M_c|y_i-y_1|}(1) + M_c|y_i - y_1|.$$

When $i = 2$, we get

$$J_{-M_c|y_2-y_1|}(1) - M_c|y_2 - y_1| \leq C \leq J_{M_c|y_2-y_1|}(1) + M_c|y_2 - y_1|,$$

as $v_1^C = 0$, which provides bounds on C that can be used to find C using the bisection method on $h(C) = 0$.

Remark B.4. In this theoretical algorithm, the issue discussed in the last bullet point of Remark 3.11 – namely, that certain values of C may produce a Laguerre tessellation that is not nested – is resolved within this framework. In such cases, nestedness can be enforced using the correction described above: whenever

$$\overline{k}_i = v_{i+1}^* - v_i^* > k_{\max}(y_{i-1}, v_i^* - v_{i-1}^*),$$

we replace v_{i+1}^* by

$$v_{i+1}^* := v_i^* + k_{\max}(y_{i-1}, v_i^* - v_{i-1}^*),$$

where the notation v^* refers to the updated numerical values, as also used in Remark 3.10. However, in all the examples shown in Subsection 3.3, this situation does not arise, and hence this correction step is omitted in the implementation.

Lemma B.2 implies that if there is some C such that $h(C) = 0$, the bisection algorithm will converge to this C. The proposition below, in turn, verifies that if the solution is discretely nested, the C corresponding to the solution satisfies $h(C) = 0$. Together, these two results ensure that if the solution is discretely nested, the bisection algorithm applied to the error function h constructed above will converge to it.

Proposition B.5. *Let ν be the solution of (4). If (c, μ, ν) is discretely nested, then $\nu_i = \nu_i^C = \mu(X_i(v^C))$ for all $1 \leq i \leq N$ where C satisfies $h(C) = 0$.*

Proof. Since ν is the solution of (4), ν satisfies $\nu_i = (f')^{-1}(C - v_i)$ for all $1 \le i \le N$, where v is the Kantorovich potential of (c, μ, ν) with $v_1 = 0$ for some C. Let $\nu_i^C = \mu(X_i(v^C))$ and we claim that $\nu_i^C = \nu_i$. From the discrete nestedness of (c, μ, ν), we have $\mu(X_{\geq}^N(y_1, v_2 - v_1)) = \nu_1$. Then, $\mu(X_1(v^C)) = \mu(X_{\geq}^N(y_1, v_2 - v_1)) = (f')^{-1}(C) = \nu_1$. which implies $v_2 - v_1 = k_1$, and so $v_2^C = k_1 + v_1 = v_2$. Also, from the fact $\nu_2 = (f')^{-1}(C - v_2)$, and the discrete nestedness, we get $\mu(X_{\geq}^N(y_2, v_3 - v_2)) = \nu_1 + \nu_2$, and $v_3 - v_2 < k_{\max}(y_1, v_2 - v_1)$. As $v_2^C = v_2$, we conclude that $\nu_2^C = (f')^{-1}(C - v_2^C)$ and $\overline{k_2} = v_3 - v_2 < k_{\max}(y_1, k_1)$ and we set $k_2 = \overline{k}_2$ where $\mu(X_{\geq}^N(y_2, \overline{k}_2)) = \nu_1^C + \nu_2^C$. So, $v_3 - v_2 = k_2$ and $v_3 = k_2 + v_2^C = v_3^C$.

Proceeding inductively, we get $v_i^C = v_i$ and $\nu_i^C = \nu_i$ for all $1 \le i \le N$ and $h(C) = 1 - \mu(X_{\geq}^N(y_{N-1}, k_{N-1})) - (f')^{-1}(C - v_N^C) = \nu_N - (f')^{-1}(C - v_N^C) = 0$, which completes the proof. □

As a consequence of Lemma B.2 and Proposition B.5, if the solution is discretely nested, then there exists $C^\star$ such that $h(C^\star) = 0$, and we conclude that the numerical algorithm based on the bisection method applied to h converges to $C^\star$.

Theorem B.6 (Convergence of the nested algorithm)**.** *Under the assumptions of Sections 2 and 3 and assuming that (c, μ, ν) is discretely nested, the error function h is decreasing with a unique root $C^\star$, and the nested bisection algorithm converges to $C^\star$.*

Proof. By Proposition B.5, if (c, μ, ν) is discretely nested, then the constant $C^\star$ associated with the solution satisfies $h(C^\star) = 0$. Lemma B.2 shows that h is decreasing. Remark B.3 provides bounds $C_L \le C^\star \le C_U$, so the root is initially bracketed.

The bisection method preserves this bracketing due to monotonicity of h, and the interval length halves at each iteration. Hence the iterates converge to a value $\bar{C}$ with $h(\bar{C}) = 0$. Monotonicity of h implies uniqueness of the root and so $\bar{C} = C^\star$. □