Performance Evaluation 154 (2022) 102286

Contents lists available at ScienceDirect

Performance Evaluation

journal homepage: www.elsevier.com/locate/peva

Analysis of an optimal policy in dynamic bipartite matching models

Arnaud Cadas [a,b], Josu Doncel [c,*], Ana Bušić [a,b]

[a] Inria, Paris, France
[b] Département d'informatique de l'ENS, ENS, CNRS, PSL University, Paris, France
[c] University of the Basque Country, UPV/EHU, Leioa, Spain

ARTICLE INFO

Article history:
Received 16 June 2021
Received in revised form 1 December 2021
Accepted 15 February 2022
Available online 22 February 2022

Keywords:
Dynamic matching models
Optimal control
Markov decision process

ABSTRACT

A dynamic bipartite matching model is given by a bipartite matching graph which determines the possible matchings between the various types of supply and demand items. Both supply and demand items arrive to the system according to a stochastic process. Matched pairs leave the system and the others wait in the queues, which induces a holding cost. We model this problem as a Markov Decision Process and study the discounted cost and the average cost problem. We assume that the cost function is linear on the queue sizes. We show that for the N-shaped matching graph, an optimal matching control prioritizes the matchings in the pendant edges and is of threshold type for the diagonal edge. In addition, for the average cost problem, we compute the optimal threshold value. We then show how the obtained results can be used to characterize the structure of an optimal matching control for a quasi-complete graph with an arbitrary number of nodes. For arbitrary bipartite graphs, we show that, when the cost of the pendant edges is larger than in the neighbors, an optimal matching policy prioritizes the items in the pendant edges. We also study the W-shaped matching graph and, when the cost of the pendant edges is larger than the cost of the middle edge, we conjecture that an optimal matching policy is also of threshold type with priority to the pendant edges; however, when the cost of the middle edge is larger, we present simulations that show that it is not optimal to prioritize items in the pendant edges.

1. Introduction

We study matching models with a bipartite compatibility graph. In this model, one supply and one demand item arrive to the system in each time slot according to a given stochastic process. Compatible supply and demand items can be matched, in which case they leave the system, and items that are not matched stay in the system. We assume that supply and demand items are divided in classes. Thus, each class of supply items is compatible with a different subset of demand item classes and, likewise, each class of demand items is compatible with a different subset of supply item classes. In Fig. 1 we represent an example of a compatibility graph with three demand nodes and three supply nodes. In this case, when the system is empty and there is an arrival of the demand class 2 and the supply class 2, the arriving items can be matched and leave the system. However, in case of an arrival of the demand class 1 and the supply class 3 when the system is empty, both items stay in the system since they cannot be matched.

* Corresponding author.
E-mail address: josu.doncel@ehu.eus (J. Doncel).

https://doi.org/10.1016/j.peva.2022.102286

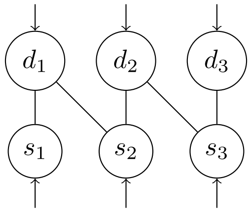

Fig. 1. A matching graph with three supply classes and three demand classes.

Once the compatibility graph and the probability distribution of arrivals of supply and demand classes are fixed, the stochastic process of the number of items present in the system depends clearly on how compatible items are matched, i.e., on the matching policy. For instance, the First Come First Matched policy is a popular matching policy that consists of matching the incoming items with the oldest compatible items, if any. Another example is the Matching the Longest policy, which matches each incoming item with its compatible item with the largest number of items (or the largest queue). Unmatched supply and demand items incur a cost (for instance, when the available kidneys are not compatible with the patient that requires the donation). An optimal matching policy can be hence defined as how items are matched so as to minimize the cost of the system. In this work, we consider bipartite matching graphs and we aim to characterize the optimal matching policy for this case.

When the compatibility graph is complete, since supply and demand items arrive in pairs, any newly arriving pair can be matched, so there is never any queue. This shows that an optimal matching control consists of matching all the items for this matching graph. For the rest of the compatibility graphs, the characterization of an optimal matching policy is not that easy. Indeed, we model this problem as a Markov Decision Process and use arguments of structured policies in this article. We consider the discounted cost problem as well as the average cost problem. We assume that the instantaneous cost is a linear function on the queue sizes (i.e., on the number of items of each class).

The main contributions of this article are summarized as follows:

- We first consider the N-shaped matching graph, which is formed by two supply and two demand classes. For this system, we show that an optimal matching policy matches all the items of the pendant edges and is of threshold type in the diagonal edge. Furthermore, for the average cost problem, we provide an analytical expression of the optimal threshold.
- We then focus on quasi-complete matching graphs with an arbitrary number of supply and demand nodes. We provide conditions on the holding costs under which a threshold type policy has the same cost as a threshold type policy of the N-shaped graph. As a result, using that an optimal matching policy for the N-shaped graph is of threshold type, we show that an optimal policy of a quasi-complete graph with an arbitrary number of supply and demand nodes is also of threshold type.
- We study optimal matching policies of an arbitrary bipartite graph in which the holding cost of each pendant edge is larger than the holding cost of its neighbors. For this case, we show that an optimal matching policy consists of matching all the items of the pendant edges.
- Finally, we consider the W-shaped graph, which is formed by two supply and three demand classes. We differentiate two cases. First, we consider that the cost of the pendant demand nodes is larger than of the middle demand node. We present the properties that the value function must satisfy to prove that an optimal matching policy is of threshold type with priority to the pendant edges. Unfortunately, given the difficulty of these properties, we did not succeed in showing that these properties are preserved by the Dynamic Programming operator. However, we show that, if there is a set of properties (containing those required to show the optimality of the threshold type policy) that are preserved under the Dynamic Programming operator, an optimal matching policy is of threshold type with priority to the pendant edges for this case. Furthermore, we consider the case when the cost of the pendant nodes is smaller than of the middile demand node. For this case, we present our numerical work that shows that the matching policy that prioritizes the pendant edges is not optimal.

A conference version of this article appeared in [1].

The remainder of the article is organized as follows. In Section 2, we put our work in the context of the existing literature. In Section 3, we describe the optimal control problem we investigate in this paper. We characterize an optimal matching policy for the N-shaped graph in Section 4. Then, in Section 5 we study the optimal policy of a quasicomplete matching graph and in Section 6 of arbitrary bipartite matching graphs. We also consider in Section 7 the W-shaped matching graph. Finally, we provide the main conclusions of our work in Section 8.

2. Related work

The study of how to optimally match compatible items has been widely studied. This problem was introduced by with Petersen and König and it was analyzed first considering that the population is fixed. For this case, its known that the

Hopcroft–Karp algorithm [2] solves this problem in a bipartite graph with a time complexity of $O(m\sqrt{n})$ where m is the number of edges and n is the number of nodes. We refer to [3, Table I] for a historical review of algorithms that compute maximum matchings. Then, this problem was extended to the dynamic setting in which one population is static and the other arrives according to a stochastic process [4–7]. Our work differs from this large literature since we explore fully dynamic matching models, i.e., all the items arrive to the system according to a random process.

In 1984, Kaplan studied the tenant assignment process of public housing in Boston [8] in which, when a public housing unit becomes available, it is assigned to the longest waiting family that listed the corresponding housing project. Kaplan was interested in the matching rate, i.e., the fraction of families having the same preferences that are assigned to a specific housing project. The authors in [9] modeled this problem as the First Come First Served (FCFS) infinite bipartite matching model. This problem is defined by a connected bipartite graph, where nodes represent the class of items and the edges their compatibilities. Several articles followed the aforementioned work. The authors in [10] provided necessary and sufficient conditions for the ergodicity of the Markov chain associated to this model. They also proved the product form of its stationary distribution. Furthermore, the authors in [11] considered a more detailed Markov chain and proved its reversibility. The authors in [12] showed that the stationary distribution of the FCFS infinite bipartite matching model coincides with that of the following queueing systems: the redundant model of [13] and the skill-based parallel server system of [14].

In [15] the authors considered the bipartite matching model with other matching policies such as Last Come First Served, Random, Match the Longest and Priority and established necessary and sufficient conditions for the stability of these models. They also showed that the Match the Longest policy has the maximum stability region and introduced the Extended Bipartite Matching Model, which extends the FCFS infinite matching policy by considering any joint distribution for the classes of arrival pairs.

We would like to remark that some authors also investigated matching models where the compatibility graph is not bipartite. This alternative model was introduced by [16] and the main particularity is that items arrive to the system one by one. The authors in [17] showed that the stationary distribution of items for this model with the FCFS matching policy satisfies a product-form expression.

In this work, we aim to find study an optimal matching policy with holding costs on the size of the queues. The authors in [18] also considered holding costs on the size of the queues in a non-bipartite matching model and for the finite horizon case. In our work, we consider a bipartite matching model and the discounted and average problems. Indeed, to the best of our knowledge, optimality results for bipartite matching models have been obtained only in asymptotic regimes. The authors in [19] analyze the heavy-traffic regime in which the difference between the arrivals of one item class and its compatible items tends to zero for the average cost problem and derive a new matching policy that is asymptotically optimal with bounded regret. A related optimization problem of maximizing rewards on edges has been considered in [20]. In that paper, the authors consider compatibilities given by a hypergraph and develop a matching policy based on an extension of the Greedy Primal–Dual algorithm and, using fluid limits, they show asymptotic optimality of their algorithm.

One of the main contributions of this work is to show the optimality of threshold-type matching policies in dynamic bipartite matching graphs. Similar policies have been also studied in a different context in which the goal is to optimally assign jobs to servers. For instance, the authors in [21,22] consider that N-shaped model and show that there is a threshold policy that is asymptotically optimal. The authors in [23] extended this result to a parallel server system with an arbitrary topology.

3. Model description

We consider a bipartite matching graph $(\mathcal{D} \cup \mathcal{S}, \mathcal{E})$ where $\mathcal{D} = \{d_1, d_2, \dots, d_{n_D}\}$ and $\mathcal{S} = \{s_1, s_2, \dots, s_{n_S}\}$ are, respectively, the set of demand nodes (or queues) and the set of supply nodes. $\mathcal{E} \subset \mathcal{D} \times \mathcal{S}$ is the set of allowed matching pairs. In each time slot n, a demand item and a supply item arrive to the system according to the i.i.d. arrival process $A(n)$. We assume independence between demand and supply arrivals. The demand item arrives to the queue d_i with probability α_i and the supply item arrives to the queue s_j with probability β_j, i.e:

$$\forall (i,j) \in \mathcal{A} \quad \mathbb{P}(A(n) = e_{(i,j)}) = \alpha_i \beta_j > 0$$

with $\sum_{i=1}^{n_\mathcal{D}} \alpha_i = 1$, $\sum_{j=1}^{n_\mathcal{S}} \beta_j = 1$ and where $\mathcal{A} = \mathcal{D} \times \mathcal{S}$ is the set of allowed arrival pairs, $e_{(i,j)} = e_{d_i} + e_{s_j}$ and $e_k \in \mathbb{N}^{n_\mathcal{D}+n_\mathcal{S}}$ is the vector of all zeros except in the kth coordinate where it is equal to one, $k \in \mathcal{D} \cup \mathcal{S}$. We assume that the α_i and β_j are chosen such that the arrival distribution satisfies the necessary and sufficient conditions for stabilizability of the MDP model: Ncond given in [24], i.e. $\forall D \subsetneq \mathcal{D}$, $\forall S \subsetneq \mathcal{S}$:

$$\sum_{d_i \in D} \alpha_i < \sum_{s_j \in \mathcal{S}(D)} \beta_j \text{ and } \sum_{s_j \in S} \beta_j < \sum_{d_i \in \mathcal{D}(S)} \alpha_i \tag{1}$$

where $\mathcal{D}(j) = \{i \in \mathcal{D} : (i,j) \in \mathcal{E}\}$ is the set of demand classes that can be matched with a class j supply and $\mathcal{S}(i) = \{j \in \mathcal{S} : (i,j) \in \mathcal{E}\}$ is the set of supply classes that can be matched with a class i demand. The extension to subsets $S \subset \mathcal{S}$ and $D \subset \mathcal{D}$ is $\mathcal{D}(S) = \bigcup_{j \in S} \mathcal{D}(j)$ and $\mathcal{S}(D) = \bigcup_{i \in D} \mathcal{S}(i)$. The main notation of this article is presented in Table 1.

Table 1
Summary of the main notation of this article.

$\mathcal{D} = \{d_1, d_2, \dots, d_{n_D}\}$	Set of demand nodes
$\mathcal{S} = \{s_1, s_2, \dots, s_{n_S}\}$	Set of supply nodes
α_i	Probability that a demand item arrives to node d_i
β_i	Probability that a supply item arrives to node s_i
(i, j)	Edge that connects d_i and s_j
e_{d_i}	The d_ith vector of the canonical basis of $\mathbb{R}^{n_D+n_S}$
e_{s_j}	The $n_D + s_j$th vector of the canonical basis of $\mathbb{R}^{n_D+n_S}$
$e_{(i,j)}$	The sum of e_{d_i} and e_{s_j} (i.e., $e_{(i,j)} = e_{d_i} + e_{s_j}$)
$\mathcal{D}(j)$	The set of demand classes that can be matched with s_j
$\mathcal{S}(i)$	The set of supply classes that can be matched with d_i
$Q(n) = (q_k(n))_{k\in\mathcal{D}\cup\mathcal{S}}$	Vector of queue lengths at time slot n
$X(n) = (x_k(n))_{k\in\mathcal{D}\cup\mathcal{S}}$	Vector of queue lengths at time slot n after arrivals
U_x	Set of admissible matchings when the vector of lengths after arrivals is x
c_k	Holding cost of node $k \in \mathcal{D} \cup \mathcal{S}$

We denote by $q_k(n)$ the queue length of node k at time slot n, where $k \in \mathcal{D}\cup\mathcal{S}$. Let $Q(n) = (q_k(n))_{k\in\mathcal{D}\cup\mathcal{S}}$ be the vector of the queue length of all the nodes and $\mathcal{Q} = \{q \in \mathbb{N}^{n_\mathcal{D}+n_\mathcal{S}} : \sum_{k\in\mathcal{D}} q_k = \sum_{k\in\mathcal{S}} q_k\}$ be the set of all the possible queues length. We must have $q(n) \in \mathcal{Q}$ for all n. Matchings at time n are carried out after the arrivals at time n. Hence, $Q(n)$ evolves over time according to the following expression:

$$Q(n+1) = Q(n) + A(n) - u(Q(n), A(n)), \tag{2}$$

where u is a deterministic Markovian decision rule which maps the current state $Y(n) = (Q(n), A(n))$ to the vector of the items that are matched at time n. Thus, Y is a Markov Decision Process where the control is denoted by u. It is sufficient to consider only deterministic Markovian decision rules and not all history-dependent randomized decision rules as proved in [25, Theorem 5.5.3] and [25, Proposition 6.2.1]. Let us define $X(n) = Q(n) + A(n)$ as the vector of the queue length of all the nodes just after the arrivals. In order to ease the notations and because the matchings only depend on the queues length after the arrivals, we use the following notation in the remainder of the paper: $u(Q(n), A(n)) = u(X(n))$. When the state of the system is $Y(n) = (q, a)$, $x = q + a$, $u(x)$ must belong to the set of admissible matchings which is defined as:

$$U_x = \left\{ u = \sum_{(i,j)\in\mathcal{E}} u_{(i,j)} e_{(i,j)} \in \mathbb{N}^{n_\mathcal{D}+n_\mathcal{S}} : (a)\ \forall i \in \mathcal{D}, \sum_{k\in S(i)} u_{(i,k)} \le x_{d_i},\ (b)\ \forall j \in \mathcal{S}, \sum_{k\in D(j)} u_{(k,j)} \le x_{s_j} \right\} \tag{3}$$

where $u_{(i,j)}$ is the number of matchings in the edge (i, j). U_x is defined for all $x \in \mathcal{Q}$. We consider a linear cost function on the buffer size of the nodes: $c(Q(n), A(n)) = c(X(n)) = \sum_{k\in\mathcal{D}\cup\mathcal{S}} c_k x_k(n)$.

A matching policy π is a sequence of deterministic Markovian decision rules, i.e. $\pi = (u(X(n)))_{n\ge1}$. The goal is to obtain an optimal matching policy for two optimization problems:

- The average cost problem:

$$g^* = \inf_\pi g^\pi \quad \text{with } g^\pi(y) = \lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} \mathbb{E}_y^\pi [c(Y(n))]$$

- The discounted cost problem:

$$v_\theta^* = \inf_\pi v_\theta^\pi \quad \text{with } v_\theta^\pi(y) = \lim_{N\to\infty} \sum_{n=0}^{N-1} \theta^t \mathbb{E}_y^\pi [c(Y(n))]$$

where $\theta \in [0, 1)$ is the discount factor and $y \in \mathcal{Y} = \mathcal{Q} \times \mathcal{A}$ is the starting state. Both problems admit an optimal stationary policy, i.e. the decision rule depends only on the state of the system and not on the time [25]. The notation $\mathbb{E}_y^\pi$ indicates that the expectation is over the arrival process, given that $Y(0) = y$ and using the matching policy π to determine the matched items $u(X(n))$ for all n.

As $A(n)$ are i.i.d., to ease the notation from now on, we denote by A a random variable with the same distribution as $A(1)$. For a given function v, $Y(n) = (q, a)$, $x = q + a$, $u \in U_x$, we define for all $0 \le \theta \le 1$:

$$L_u^\theta v(q, a) = c(q, a) + \theta\mathbb{E}[v(q + a - u, A)] = c(x) + \theta\mathbb{E}[v(x - u, A)]$$
$$L^\theta v(q, a) = c(q, a) + \min_{u\in U_x} \theta\mathbb{E}[v(q + a - u, A)] = c(x) + \min_{u\in U_x} \theta\mathbb{E}[v(x - u, A)]$$

and in particular, we define $T_u = L_u^1$ and $T = L^1$. A solution of the discounted cost problem can be obtained as a solution of the Bellman fixed point equation $v = L^\theta v$. In the average cost problem, the Bellman equation is given by $g^* + v = Tv$.

We say that a value function v or a decision rule u is structured if it satisfies a special property, such as being increasing, decreasing or convex. Throughout the article, by increasing we mean nondecreasing and we will use strictly increasing

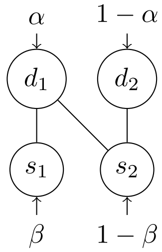

Fig. 2. The N-shaped matching graph.

for increasing. For a collection of properties σ, we denote by V^σ the set of structured value functions that satisfy those properties. Similarly, we define D^σ as the set of structured decision rules induced by those structured value functions. A policy is called structured when it only uses structured decision rules and the set of structured stationary matching policies is denoted by $\Pi^\sigma = \{\pi = (u(X(n)))_{n\geq 1} : u \in D^\sigma\}$. The framework of this work is that of property preservation when we apply the Dynamic Programming operator. For the average cost problem, we use an adapted version of [23,Theorem 6.11.3] and we proceed to characterize the structure of an optimal matching policy as follows: first, we identify a set of structured value functions V^σ and a set of structured deterministic Markovian decision rules D^σ such that if the value function belongs to V^σ an optimal decision rule belongs to D^σ. Then, we show that the properties of V^σ are preserved by the Dynamic Programming operator, i,e. that $Lv \in V^\sigma$ if $v \in V^\sigma$. Finally, we show that these properties hold in the limit as well.

In the case of the average cost problem, our proofs are based on [23,Theorem 8.11.1], which uses the results of the discounted cost problem. In fact, the average cost problem is considered as a limit when θ tends to one and, therefore, it is enough to show that the properties still hold for this limit. The statement of the theorems we use for the discounted cost problem and the average cost problem are provided in Appendix A. These theorems present some technical requiments due to the unboundness of the costs. In Appendix A we show that the technical requirements of the theorems we use are satisfied and, therefore, when we study an optimal matching policy in the discounted cost problem, we only need to check that conditions (a), (b) and (c) of Theorem 6 are verified, i.e,.

(a) $v \in V^\sigma$ implies that $L^\theta v \in V^\sigma$;
(b) $v \in V^\sigma$ implies that there exists a decision $u' \in D^\sigma$ such that $u' \in \arg\min_u L_u^\theta v$;
(c) V^σ is a closed subset of the set of value functions under pointwise convergence.

whereas when we study an optimal matching policy in the average cost problem, we only need to check that conditions (a) and (b) of Theorem 7 are verified, i.e.,

(a) for any sequence $(\theta_n)_{n\geq 0}$, $0 \leq \theta_n < 1$, for which $\lim_{n\to+\infty} \theta_n = 1$,

$$\lim_{n\to+\infty} [v^*_{\theta_n} - v^*_{\theta_n}(0)e] \in V^\sigma_H \quad \text{with } e(y) = 1 \text{ for all } y \in \mathcal{Y}$$

(b) $v \in V^\sigma_H$ implies that there exists a decision $u' \in D^\sigma$ such that $u' \in \arg\min_u T_u v$;

4. *N*-Shaped graph

We now focus on the N-shaped matching graph, which is formed by two supply nodes and two demand nodes as well as a N-shaped set of edges (see Fig. 2). Specifically, we have $\mathcal{D} = \{d_1, d_2\}$, $\mathcal{S} = \{s_1, s_2\}$ and $\mathcal{E} = \{(1, 1), (1, 2), (2, 2)\}$. We also define $(2, 1)$ as the imaginary edge between d_2 and s_1 (imaginary because $(2, 1) \notin \mathcal{E}$) that we introduce to ease the notations. To ensure stability, we assume that $\alpha > \beta$.

In this section, we show that an optimal policy for the N-shaped matching graph has a specific structure. For this purpose, we first present the properties of the value function. Then, we show how these properties characterize the optimal decision rule and how they are preserved by the Dynamic Programming operator. Finally, we prove the desired results in Theorems 1 and 2.

4.1. Value function properties

We now present the properties of the value function. We first define the increasing property as follows:

Definition 1 (*Increasing Property*)**.** Let $(i, j) \in \mathcal{E}$. We say that a function v is increasing in (i, j) or $v \in \mathcal{I}_{(i,j)}$ if

$$\forall a \in \mathcal{A}, \forall q \in \mathcal{Q}, \quad v(q + e_{(i,j)}, a) \geq v(q, a).$$

Remark 1. The increasing property in (2, 1) can be interpreted as the preference to match items of (1, 1) and of (2, 2) than of (1, 2). Indeed, $v(q+e_{(1,1)}+e_{(2,2)}-e_{(1,2)},a)=v(q+e_{(2,1)},a)\geq v(q,a)$.

We now define the convexity property as follows:

Definition 2 (*Convexity Property*)**.** A function v is convex in (1, 2) or $v\in\mathcal{C}_{(1,2)}$ if $\forall a\in\mathcal{A}$, $\forall q\in\mathcal{Q}$ such that $q_{d_1}\geq q_{s_1}$, we have

$$v(q+2e_{(1,2)},a)-v(q+e_{(1,2)},a)\geq v(q+e_{(1,2)},a)-v(q,a).$$

Likewise, v is convex in (2, 1) or $v\in\mathcal{C}_{(2,1)}$ if $\forall a\in\mathcal{A}$, $\forall q\in\mathcal{Q}$ such that $q_{s_1}\geq q_{d_1}$, we have

$$v(q+2e_{(2,1)},a)-v(q+e_{(2,1)},a)\geq v(q+e_{(2,1)},a)-v(q,a).$$

We also define the boundary property next:

Definition 3 (*Boundary Property*)**.** A function $v\in\mathcal{B}$ if

$$\forall a\in\mathcal{A},\quad v(0,a)-v(e_{(2,1)},a)\leq v(e_{(1,2)},a)-v(0,a).$$

We remark that, the properties $\mathcal{I}_{(1,1)}$, $\mathcal{I}_{(2,2)}$, $\mathcal{I}_{(2,1)}$ and $\mathcal{C}_{(1,2)}$ are used to characterize the optimal decision rule, whereas $\mathcal{C}_{(2,1)}$ and $\mathcal{B}$ are required to show that $\mathcal{C}_{(1,2)}$ is preserved by the operator L^{θ}. In the remainder of the section, we consider the following set of structured value functions

$$V^{\sigma}=\mathcal{I}_{(1,1)}\cap\mathcal{I}_{(2,2)}\cap\mathcal{I}_{(2,1)}\cap\mathcal{C}_{(1,2)}\cap\mathcal{C}_{(2,1)}\cap\mathcal{B},\tag{4}$$

which means that we consider that the value function satisfies the following properties: it is increasing with (1, 1), (2, 2) and (2, 1), convex in (1, 2) and (2, 1) and satisfies the boundary property.

4.2. Optimal decision rule

A decision rule is of threshold type in (1, 2) with priority to (1, 1) and (2, 2) if it matches all the items of (1, 1) and (2, 2) and it matches the items of (1, 2) only if the remaining items (in d_1 and s_2) exceed a specific threshold $t\in\mathbb{N}\cup\infty$. The decision rule of threshold type in (1, 2) with priority to (1, 1) and (2, 2) is defined formally now.

Definition 4 (*Threshold-type Decision Rule*)**.** A decision rule u_x is of threshold type in (1, 2) with priority to (1, 1) and (2, 2) when $u_x=\min(x_{d_1},x_{s_1})e_{(1,1)}+\min(x_{d_2},x_{s_2})e_{(2,2)}+k_t(x)e_{(1,2)}$ where

$$k_t(x)=\begin{cases}0, & \textit{if } x_{d_1}-x_{s_1}\leq t,\\ x_{d_1}-x_{s_1}-t, & \textit{otherwise}.\end{cases}$$

Let us note that if $t=\infty$, the above decision rule consists of never matching (1, 2). However, if $t<\infty$, the decision rule matches the items of (1, 2) until the remaining items in d_1 and s_2 after matching all the possible items in (1, 1) and (2, 2) do not exceed the threshold t. In fact, the state of the system after a decision rule of threshold type in (1, 2) with priority to (1, 1) and (2, 2) is of the form $(0,l,l,0)$ if $x_{d_1}\leq x_{s_1}$, of the form $(l,0,0,l)$ if $x_{d_1}>x_{s_1}$ and $l<t$ and the form $(t,0,0,t)$ otherwise.

In the rest of this section, we consider that D^{σ} is the set of decision rules that are of threshold type in (1, 2) with priority to (1, 1) and (2, 2) for any $t\in\mathbb{N}\cup\infty$.

We now show that there exists an optimal decision rule that matches all the items in (1, 1) and (2, 2).

Proposition 1. *Let $v\in\mathcal{I}_{(1,1)}\cap\mathcal{I}_{(2,2)}\cap\mathcal{I}_{(2,1)}$ and $0\leq\theta\leq1$. For any $q\in\mathcal{Q}$ and $a\in\mathcal{A}$, let $x=q+a$. Thus, there exists $u^*\in U_x$ such that $u^*\in\arg\min_{u\in U_x}L_u^{\theta}v(q,a)$, $u^*_{(1,1)}=\min(x_{d_1},x_{s_1})$ and $u^*_{(2,2)}=\min(x_{d_2},x_{s_2})$. In particular, this result holds for the average operator: T_u.*

Proof. See Appendix B. □

From this result, it follows that there exists an optimal decision rule that matches all possible items of (1, 1) and (2, 2). We denote by K_x the set of possible matching in (1, 2) after matching all the items of (1, 1) and (2, 2).

Definition 5. Let $0\leq\theta\leq1$, $x\in\mathcal{Q}$. Then,

$$K_x=\begin{cases}\{0\}, & \textit{if } x_{d_1}\leq x_{s_1},\\ \{0,\dots,\min(x_{d_1}-x_{s_1},x_{s_2}-x_{d_2})\}, & \textit{otherwise}.\end{cases}$$

We now prove that a decision rule of threshold type in (1, 2) with priority to (1, 1) and (2, 2) is optimal.

Proposition 2. *Let $v\in\mathcal{I}_{(1,1)}\cap\mathcal{I}_{(2,2)}\cap\mathcal{I}_{(2,1)}\cap\mathcal{C}_{(1,2)}$. Let $0\leq\theta\leq1$. For any $q\in\mathcal{Q}$ and for any $a\in\mathcal{A}$, $x=q+a$, there exists $u^*\in D^{\sigma}$ such that $u^*\in\arg\min_{u\in U_x}L_u^{\theta}v(q,a)$. In particular, this result holds for the average operator: T_u.*

Proof. The idea of the proof is to show that for any admissible matching u, the matching u^* that prioritizes the pendant edges satisfies that $L^{\theta}_{u^*}v(q,a) \leq L^{\theta}_{u}v(q,a)$ for all q and a. Detailed proof is provided in Appendix C. □

4.3. Value function property preservation

In this section, we show that the properties of the value function defined in Section 4.1 are preserved by the Dynamic Programming operator. In other words, we show that if v is increasing with $(1,1)$, $(2,2)$ and $(2,1)$, convex in $(1,2)$ and $(2,1)$ and satisfies the boundary property, so does $L^{\theta}v$.

We first focus on the increasing property in $(1,1)$, $(2,2)$ and $(2,1)$ and we show that they are preserved by the Dynamic Programming operator.

Lemma 1. *If a function $v \in \mathcal{I}_{(1,1)} \cap \mathcal{I}_{(2,2)} \cap \mathcal{I}_{(2,1)}$, then $L^{\theta}v \in \mathcal{I}_{(1,1)} \cap \mathcal{I}_{(2,2)} \cap \mathcal{I}_{(2,1)}$.*

Proof. The idea of the proof is to consider any q and a and define $\bar{q}$ as q with two additional items for each edge we consider. Then, since $v(q,a) \leq v(\bar{q},a)$ by assumption, it is enough to show that $L^{\theta}v(q,a) \leq L^{\theta}v(\bar{q},a)$. Detailed proof is provided in Appendix D.1. □

We now aim to show that the convexity in $(1,2)$ is preserved by the Dynamic Programming operator. It is important to note that $L^{\theta}v \in C_{(1,2)}$ when the value function is not only increasing in $(1,1)$, $(2,2)$ and $(2,1)$ and convex in $(1,2)$, but also satisfies the boundary condition.

Lemma 2. *If $v \in \mathcal{I}_{(1,1)} \cap \mathcal{I}_{(2,2)} \cap \mathcal{I}_{(2,1)} \cap \mathcal{C}_{(1,2)} \cap \mathcal{B}$, then $L^{\theta}v \in \mathcal{C}_{(1,2)}$.*

Proof. The idea of the proof is to consider any q and a and define $\bar{q}$ as q with two additional items for each that we consider as well as $\bar{\bar{q}}$ as $\bar{q}$ with two more additional items in the same edge. Then, since $v \in \mathcal{C}_{(1,2)}$, we know that $v(\bar{q},a) - v(q,a) \leq v(\bar{\bar{q}},a) - v(\bar{q},a)$ by assumption, it is enough to show that $L^{\theta}v(\bar{q},a) - L^{\theta}v(q,a) \leq L^{\theta}v(\bar{\bar{q}},a) - L^{\theta}v(\bar{q},a)$. Detailed proof is provided in Appendix D.2. □

Finally, to show that the Dynamic Programming operator preserves the boundary property, we need to use the convexity property in $(2,1)$. The preservation of the boundary and the convexity properties are proven in the following lemma.

Lemma 3. *If $v \in \mathcal{I}_{(1,1)} \cap \mathcal{I}_{(2,2)} \cap \mathcal{I}_{(2,1)} \cap \mathcal{C}_{(1,2)} \cap \mathcal{C}_{(2,1)} \cap \mathcal{B}$, then $L^{\theta}v \in \mathcal{C}_{(2,1)} \cap \mathcal{B}$.*

Proof. For the Boundary condition, the idea is to show that $L^{\theta}v(0,a) - L^{\theta}v(e_{(2,1)},a) \leq L^{\theta}v(e_{(1,2)},a) - L^{\theta}v(0,a)$ for any $a \in \mathcal{A}$, whereas for the convexity property we proceed similarly than in the proof of Lemma 2. Detailed proof is provided in Appendix D.3. □

4.4. Structure of an optimal policy

Now, using the result of Theorem 6, we show that there exists an optimal matching policy which is formed of a sequence of decision rules that belongs to D^{σ} (with a fixed threshold).

Theorem 1. *The optimal control for the discounted cost problem is of threshold type in* $(1,2)$ *with priority to* $(1,1)$ *and* $(2,2)$*.*

Proof. We apply Theorem 6 where V^{σ} is the set of functions defined in (4) and D^{σ} the set defined in Definition 4. We focus on the structural conditions of the theorem. From Lemmas 1–3 of Section 4.3, it follows (a) since they show that if $v \in V^{\sigma}$, then $L^{\theta}v \in V^{\sigma}$. The result of Proposition 2 shows (b) because the policy that belongs to D^{σ} minimizes $L^{\theta}_{u}v$ if $v \in V^{\sigma}$. Finally, since limits preserve inequalities, the point-wise convergence of functions of V^{σ} belong to this set, which shows (c). □

The following theorem shows that the previous result also holds for the average cost problem.

Theorem 2. *The optimal control for the average cost problem is of threshold type in* $(1,2)$ *with priority to* $(1,1)$ *and* $(2,2)$*.*

Proof. We want to apply Theorem 7 using the same value function set V^{σ} and the same decision rule set D^{σ} as in the proof of the previous theorem. Assumptions $(A1)$ to $(A4)$ hold because of Lemma 9. Let $(\theta_n)_{n\in\mathbb{N}}$ be a sequence such that $0 \leq \theta_n < 1$ for all $n \in \mathbb{N}$ and $\lim\limits_{n\to+\infty} \theta_n = 1$. Let $n \in \mathbb{N}$. We know that $v^*_{\theta_n} \in V^{\sigma}$ (see the proof of Theorem 1). The inequalities in the definitions of the properties used in V^{σ} still hold if we add a constant to v, thus $v^*_{\theta_n} - v^*_{\theta_n}(0)e \in V^{\sigma}$. Using Assumption $(A3)$ and Assumption $(A4)$, we have $H \leq v^*_{\theta_n} - v^*_{\theta_n}(0)e \leq M$, so $v^*_{\theta_n} - v^*_{\theta_n}(0)e \in V^{\sigma}_{H}$. This last result holds for each $n \in \mathbb{N}$ and since limits preserve inequalities V^{σ}_{H} is a closed set, $\lim\limits_{n\to+\infty} [v^*_{\theta_n} - v^*_{\theta_n}(0)e] \in V^{\sigma}_{H}$ which shows (a). The result of Proposition 2 shows (b) because the policy that belongs to D^{σ} minimizes $L^{1}_{u}v = T_{u}v$ if $v \in V^{\sigma}_{H} \subset V^{\sigma}$. □

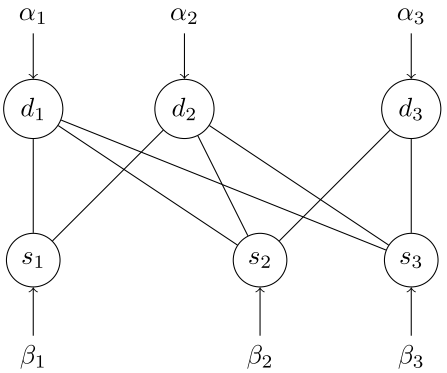

Fig. 3. A complete matching graph minus the edge (3, 1).

4.5. Computing the optimal threshold

We consider the matching policy of threshold type in $(1, 2)$ with priority to $(1, 1)$ and $(2, 2)$ in the average cost case. In this result, we provide an analytical expression of the threshold at which items of $(2, 1)$ are matched.

Proposition 3. *Let* $\rho = \frac{\beta(1-\alpha)}{\alpha(1-\beta)} \in (0, 1)$, $R = \frac{c_{s_1}+c_{d_2}}{c_{d_1}+c_{s_2}}$ *and* $\Pi^{T_{(1,2)}}$ *be the set of matching policy of threshold type in* $(1, 2)$ *with priority to* $(1, 1)$ *and* $(2, 2)$*. The optimal threshold* t^**, which minimizes the average cost on* $\Pi^{T_{(1,2)}}$*, is*

$$t^* = \begin{cases} \lceil k \rceil & \text{if } f(\lceil k \rceil) \le f(\lfloor k \rfloor) \\ \lfloor k \rfloor & \text{otherwise} \end{cases}$$

where $k = \frac{\log \frac{\rho-1}{(R+1)\log\rho}}{\log \rho} - 1$ *and* $f(x) = (c_{d_1} + c_{s_2})x + (c_{d_1} + c_{d_2} + c_{s_1} + c_{s_2})\frac{\rho^{x+1}}{1-\rho} - (c_{d_1} + c_{s_2})\frac{\rho}{1-\rho} + ((c_{d_1} + c_{s_1})\alpha\beta + (c_{d_2} + c_{s_2})(1-\alpha)(1-\beta) + (c_{d_2} + c_{s_1})(1-\alpha)\beta + (c_{d_1} + c_{s_2})\alpha(1-\beta))$.

The threshold t^* *is positive.*

Proof. The idea of the proof is to look at the Markov chain derived from the policy $u_t^\infty \in \Pi^{T_{(1,2)}}$. We show that the Markov chain is positive recurrent and we compute the stationary distribution. Using the strong law of large numbers for Markov chains, we show that the average cost $g^{u_t^\infty}$ is equal to the expected cost of the system under the stationary distribution. Then, we find an analytical form for the expected cost which depends on the threshold on $(1, 2)$, i.e, t. Finally, we minimize the function over t. Detailed proof is provided in Appendix E. □

From this result, it is important to note that the value of k (and therefore, the value of t^*) is increasing with ρ and it tends to 0 when $\rho \to 0$ whereas it tends to infinity when $\rho \to 1$. This means that, when $\beta \ll \alpha$, the threshold is zero, i.e., an optimal matching policy always matches the items of the diagonal edge. On the other hand, when $\beta \to \alpha$ (recall that $\alpha > \beta$ to ensure stability), we have that the threshold tends to infinity. This means that an optimal matching policy never matches items of the diagonal edge.

5. Optimal policy of quasi-complete graphs

We aim to generalize the results of the N-shaped graph to a quasi-complete matching graph with an arbitrary number of supply and demand nodes. A quasi-complete matching graph is a matching graph where all the supply and demand nodes are connected, except for one supply node which is connected to all but one demand node. Let us denote a quasi-complete matching graph as $\mathcal{G} = (\mathcal{D} \cup \mathcal{S}, \mathcal{E})$ with $\mathcal{E} = (\mathcal{D} \times \mathcal{S}) \setminus \{(i^*, j^*)\}$, where (i^*, j^*) ($i^* \in \mathcal{D}$ and $j^* \in \mathcal{S}$) is the missing edge (see Fig. 3 for an example). We show that, under certain assumptions on the holding costs, this matching graph is related to the N-shaped matching graph. Hence, throughout this section, when we refer to the N-shaped matching graph we use the superscript N.

First, we define the transformations to move from Y (the Markov Decision Process defined on $\mathcal{G}$) to Y^N (the Markov Decision Process defined on $\mathcal{G}^N$).

Definition 6. Let $\mathcal{Q}$ and $\mathcal{A}$ be the sets of possible states and arrivals of $\mathcal{G}$. Let $\mathcal{Q}^N$ and $\mathcal{A}^N$ be the sets of possible states and arrivals of $\mathcal{G}^N$. We define the projection from $\mathcal{Q}$ to $\mathcal{Q}^N$ and the projection from $\mathcal{A}$ to $\mathcal{A}^N$ as

$$p_{\mathcal{Q}}^N(q) = \left(\sum_{i \in \mathcal{D}(j^*)} q_{d_i}, q_{d_{i^*}}, q_{s_{j^*}}, \sum_{j \in \mathcal{S}(i^*)} q_{s_j} \right) \text{ and}$$

$$p_{\mathcal{A}}^N(a) = \begin{cases} e_{(1,1)} & \text{if } a = e_{(i,j^*)}, i \in \mathcal{D}(j^*) \\ e_{(2,2)} & \text{if } a = e_{(i^*,j)}, j \in \mathcal{S}(i^*) \\ e_{(2,1)} & \text{if } a = e_{(i^*,j^*)} \\ e_{(1,2)} & \text{otherwise} \end{cases}.$$

We can easily show that $p_{\mathcal{Q}}^N(x) = p_{\mathcal{Q}}^N(q) + p_{\mathcal{A}}^N(a)$ where $x = q + a$. Let $a^N \in \mathcal{A}^N$, we also define $(p_{\mathcal{A}}^N)^{-1}(a^N) = \{a \in \mathcal{A} : p_{\mathcal{A}}^N(a) = a^N\}$.

Let A be the arrival process associated to $\mathcal{G}$, we construct an arrival process A^N for $\mathcal{G}^N$ in the following way:

$$\forall a^N \in \mathcal{A}^N, \quad \mathbb{P}(A^N = a^N) = \sum_{a \in (p_{\mathcal{A}}^N)^{-1}(a^N)} \mathbb{P}(A = a)$$

We assume that the quasi-complete graphs we consider in this section satisfy the following property: the holding cost is the same for all the demand nodes that are compatible with all the supply nodes and the holding cost is the same for all the supply nodes that are compatible with all the demand nodes. In other words, there exist c_1 and c_2 such that $c_{d_i} = c_1$ for all $i \in \mathcal{D}(j^*)$ and $c_{s_j} = c_2$ for all $j \in \mathcal{S}(i^*)$. In the example of Fig. 3, this means that $c_{d_1} = c_{d_2}$ as well as $c_{s_2} = c_{s_3}$. From this assumption on the holding costs, we can construct the N-shaped matching graph such that $c_{d_1}^N = c_{d_i}$, for all $i \in \mathcal{D}(j^*)$, $c_{d_2}^N = c_{d_{i^*}}$, $c_{s_1}^N = c_{s_{j^*}}$ and $c_{s_2}^N = c_{s_j}$ for all $j \in \mathcal{S}(i^*)$. Therefore, using Definition 6, it follows that

$$c(x) = c^N(p_{\mathcal{Q}}^N(x)) \text{ for all } x \in \mathcal{Q}. \tag{5}$$

We now define the decision rule of threshold-type that we study in this section.

Definition 7 (*Threshold-type decision rule*)**.** A decision rule u_x is said to be of threshold type with priority to i^* and j^* if:

1. it matches all the items of (i, j^*) for all $i \in \mathcal{D}(j^*)$ and all the items of (i^*, j) for all $j \in \mathcal{S}(i^*)$.
2. it matches the items of (i, j) ($i \in \mathcal{D}(j^*)$ and $j \in \mathcal{S}(i^*)$) only if the remaining items (in the sum of d_i for $i \in \mathcal{D}(j^*)$) are above a specific threshold, denoted by t (with $t \in \mathbb{N} \cup \infty$).

i.e, u_x is such that $\sum_{i \in \mathcal{D}(j^*)} (u_x)_{(i,j^*)} = \min(\sum_{i \in \mathcal{D}(j^*)} x_{d_i}, x_{s_{j^*}})$, $\sum_{j \in \mathcal{S}(i^*)} (u_x)_{(i^*,j)} = \min(x_{d_{i^*}}, \sum_{j \in \mathcal{S}(i^*)} x_{s_j})$ and $\sum_{i \in \mathcal{D}(j^*)} \sum_{j \in \mathcal{S}(i^*)} (u_x)_{(i,j)} = k_t(x)$ where

$$k_t(x) = \begin{cases} 0 & \text{if } \sum_{i \in \mathcal{D}(j^*)} x_{d_i} - x_{s_{j^*}} \le t \\ \sum_{i \in \mathcal{D}(j^*)} x_{d_i} - x_{s_{j^*}} - t & \text{otherwise} \end{cases}.$$

We define D^* as the set of decision rules that are of threshold type with priority to i^* and j^* for any $t \in \mathbb{N} \cup \infty$.

We aim to show that the stationary policy based on the above decision rules is optimal for a quasi-complete matching graph with an arbitrary number of supply and demand nodes. For this purpose, we first need to show the following properties.

Lemma 4. *Let $0 \le \theta < 1$, let $\pi^* = (u(X(n)))_{n \ge 0}$ (with $u \in D^*$ as defined in Definition 7) be a threshold-type policy with priority in i^* and j^* and let $(\pi^N)^* = (u^N(X^N(n)))_{n \ge 0}$ (with $u^N \in D^\sigma$ as defined in Definition 4) be a threshold-type policy with priority in $(1, 1)$ and $(2, 2)$. Thus, we have $v_\theta^{\pi^*}(y) = v_\theta^{(\pi^N)^*}(y^N)$ for all $y = (q, a) \in \mathcal{Q} \times \mathcal{A}$, with $y^N = (p_{\mathcal{Q}}^N(q), p_{\mathcal{A}}^N(a))$. The result remains true for the average cost problem with g^{π^*} and $g^{(\pi^N)^*}$.*

Proof. The idea of the proof is to first show that for π^*, we have that $p_{\mathcal{Q}}^N(x_n) = x_n^N$ and then that the expected cost in both matching models coincides. Detailed proof is provided in Appendix F. □

Lemma 5. *Let $0 \le \theta < 1$, let π be a stationary policy on Y, $y = (q, a) \in \mathcal{Q} \times \mathcal{A}$ and $y^N = (p_{\mathcal{Q}}^N(q), p_{\mathcal{A}}^N(a))$. Thus, there exists a policy π^N on Y^N such that $v_\theta^\pi(y) = v_\theta^{\pi^N}(y^N)$. The result remains true for the average cost problem with g^π and g^{π^N}.*

Proof. The idea of the proof is that, for any π, we define π^N such that we have that $p_{\mathcal{Q}}^N(x_n) = x_n^N$ and also that the expected cost in both cases coincides. Detailed proof is provided in Appendix G □

We now prove that an optimal matching policy for a quasi-complete graph with an arbitrary number of supply and demand nodes is formed by decision rules which are as defined in Definition 7.

Theorem 3. *The optimal control for the discounted cost problem and the average cost problem is of threshold type with priority to i^* and j^*.*

Proof. Let $0 \le \theta < 1$, let π be a stationary policy on Y and $\pi^* = (u(X(n)))_{n \ge 0}$ (with $u \in D^*$ as defined in Definition 7) be a threshold-type policy with priority in i^* and j^*. Let $y_0 = (q_0, a_0) \in \mathcal{Q} \times \mathcal{A}$ and $y_0^N = (p_{\mathcal{Q}}^N(q_0), p_{\mathcal{A}}^N(a_0))$. Using Lemma 5,

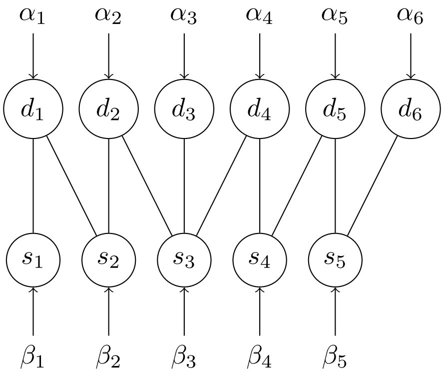

Fig. 4. An example of the bipartite matching graphs under study.

there exists a policy π^N such that:

$$v^{\pi}_{\theta}(y_0) = v^{\pi^N}_{\theta}(y^N_0) \geq v^{(\pi^N)^*}_{\theta}(y^N_0) = v^{\pi^*}_{\theta}(y_0)$$

where $(\pi^N)^* = (u^N(X^N(n)))_{n\geq 0}$ (with $u^N \in D^\sigma$ as defined in Definition 4) is a threshold-type policy with priority in (1, 1) and (2, 2). The inequality comes from our optimality result on $\mathcal{G}^N$: Theorem 1. The last equality comes from Lemma 4. This was done for any $y_0 = (q_0, a_0) \in \mathcal{Q} \times \mathcal{A}$ and the average cost problem follows easily (with Theorem 2), giving us the final result. □

6. Optimal policy of the pendant edges of arbitrary bipartite graphs

In this section, we focus on an arbitrary bipartite matching graph.

We say that an edge (i_1, j_1) belongs to the neighborhood of an edge (i_2, j_2) if $i_1 = i_2$ or $j_1 = j_2$. We denote by $N((i, j))$ the neighborhood of an edge (i, j). We assume that in the neighborhood of an pendant edge there are not pendant edges (we discuss how our results extend to several connected pendant edges in Remark 2). We denote by $\mathcal{E}^*$ the set of pendant nodes. An example of the matching graphs under study is provided in Fig. 4, where the set of pendant edges is $\mathcal{E}^* = \{(1, 1), (3, 3), (6, 5)\}$ and we have, for instance, that $N((1, 1)) = \{(1, 2)\}$ and $N((3, 3)) = \{(2, 3), (4, 3)\}$.

We show that, under certain assumptions on the costs of the nodes, an optimal matching policy consists of matching all the items of the pendant edges. To this end, we follow the same structure as in Section 4, i.e., we first present the properties of the value function; then, we characterize the optimal decision rule and show how they are preserved by the Dynamic Programming operator and, finally, we characterize the optimal matching policy in the pendant edges using Theorems 6 and 7.

6.1. Value function properties

We now define the undesirability property for a given edge.

Definition 8 (*Undesirability Property*). Let $(i_1, j_1) \in \mathcal{E}$. We say that a function v is undesirable in (i_1, j_1) or $v \in \mathcal{U}_{(i_1,j_1)}$ if for all $(i_2, j_2) \in N((i_1, j_1))$

$$v(q + e_{(i_1,j_1)} - e_{(i_2,j_2)}, a) \geq v(q, a),$$

for all $a \in \mathcal{A}$ and $q \in \mathcal{Q}$ such that $q_{s_{j_2}} \geq 1$ if $i_1 = i_2$ and $q_{d_{i_2}} \geq 1$ if $j_1 = j_2$.

Let us note that the above property is the same as saying that, $v(q+e_{(i_1,j_1)}, a) \geq v(q+e_{(i_2,j_2)}, a)$, for all $(i_2, j_2) \in N((i_1, j_1))$, that is, it is preferable to match the items in (i_1, j_1) than in (i_2, j_2). In the remainder of this section, we consider the following structured set:

$$V^\sigma = \bigcap_{(i,j)\in\mathcal{E}^*} \left(\mathcal{I}_{(i,j)} \cap \mathcal{U}_{(i,j)}\right), \tag{6}$$

where $I_{(i,j)}$ denotes the increasing property in (i, j) as defined in Definition 1. We also assume that the following property on the function c: for all $(i, j) \in \mathcal{E}^*$, $c_{d_i} \geq c_{d_{i'}}$ for all $(i', j) \in N((i, j))$ (or $c_{s_j} \geq c_{s_{j'}}$ for all $(i, j') \in N((i, j))$). This means that the cost on an pendant edges is larger than in its neighbors.

6.2. Optimal decision rule

In this section, we show that, for any $v \in V^\sigma$, there is a decision rule u that prioritizes the pendant edges and minimizes $L_u^\theta v$. We say that a matching policy prioritizes the pendant edges if it matches all the items in the pendant edges. This means that, for all $(i,j) \in \mathcal{E}^*$, $u_{(i,j)} = \min(x_{d_i}, x_{s_j})$. We consider that D^σ is the set of decision rules that prioritizes the pendant edges. We now show that, if $v \in V^\sigma$, there exists a decision rule that belongs to D^σ that minimizes $L_u^\theta v$.

Proposition 4. *Let $a \in \mathcal{A}$, $q \in \mathcal{Q}$, $x = q + a$, $v \in V^\sigma$, $0 \le \theta \le 1$. There exists $u^* \in U_x$ such that $u^* \in \arg\min_{u \in U_x} L_u^\theta v(q,a)$ and $u^*_{(i,j)} = \min(x_{d_i}, x_{s_j})$ for all $(i,j) \in \mathcal{E}^*$. In particular, this result holds for the average cost operator T_u.*

Proof. The idea of the proof is to show that for any matching u, there exists another matching policy u^* that prioritizes the items of the pendant edges and such that $L_{u^*}^\theta v(q,a) \le L_u^\theta v(q,a)$ for any q and a. Detailed proof is provided in Appendix H. □

6.3. Value function property preservation

We now show that the properties of the value function that belong to V^σ are preserved by the Dynamic Programming operator, i.e., considering that $v \in V^\sigma$, we aim to show that $Lv \in V^\sigma$.

We first focus on the preservation of the increasing property in the pendant edges by the Dynamic Programming operator. This can be shown using the same arguments as Lemma 1 and therefore we omit its proof.

Lemma 6. *Let $(i,j) \in \mathcal{E}^*$. If $v \in \mathcal{I}_{(i,j)} \cap \mathcal{U}_{(i,j)}$, then $Lv \in \mathcal{I}_{(i,j)}$.*

We also show that the undesirability property is preserved by the Dynamic Programming operator.

Lemma 7. *Let $(i,j) \in \mathcal{E}^*$. If $v \in \mathcal{I}_{(i,j)} \cap \mathcal{U}_{(i,j)}$, then $Lv \in \mathcal{U}_{(i,j)}$.*

Proof. The idea of the proof is to consider any q with at least one item in the neighborhood of a pendant edge and we define $\bar{q}$ as q plus two items in the pendant edge under consideration minus two items of the neighborhood. Since $v(q,a) \le v(\bar{q},a)$ by assumption for any a, we show that $L^\theta v(q,a) \le L^\theta v(\bar{q},a)$ for any a. Detailed proof is provided in Appendix I. □

6.4. Structure of an optimal policy

In this section, we show that there exists an optimal policy that consists of a sequence of decision rules that belongs to D^σ, i.e., that prioritizes the pendant edges. We first focus on the discounted cost problem.

Theorem 4. *The optimal control for the discounted cost problem prioritizes the pendant edges.*

Proof. We apply Theorem 6 where V^σ is the set of functions defined in (6) and D^σ the set defined in Proposition 4. From Lemmas 6 and 7, it follows (a). The result of Proposition 4 shows (b). And, since limits preserve inequalities, the point-wise convergence of functions of V^σ belong to this set, which shows (c). □

The following theorem shows that the previous result is also verified for the average cost problem.

Theorem 5. *The optimal control for the average cost problem prioritizes the pendant edges.*

Proof. We want to apply Theorem 7 using the same value function set V^σ and the same decision rule set D^σ as in the proof of the previous proposition. Let $(\theta_n)_{n \in \mathbb{N}}$ be a sequence such that $0 \le \theta_n < 1$ for all $n \in \mathbb{N}$ and $\lim_{n \to +\infty} \theta_n = 1$. Let $n \in \mathbb{N}$. We know that $v^*_{\theta_n} \in V^\sigma$ (see the proof of Theorem 4). The inequalities in the definitions of the properties used in V^σ still hold if we add a constant to v, thus $v^*_{\theta_n} - v^*_{\theta_n}(0)e \in V^\sigma$. Using Assumption $(A3)$ and Assumption $(A4)$, we have $H \le v^*_{\theta_n} - v^*_{\theta_n}(0)e \le M$, so $v^*_{\theta_n} - v^*_{\theta_n}(0)e \in V_H^\sigma$. This last result holds for each $n \in \mathbb{N}$ and since limits preserve inequalities V_H^σ is a closed set, $\lim_{n \to +\infty} [v^*_{\theta_n} - v^*_{\theta_n}(0)e] \in V_H^\sigma$ which shows (a). The result of Proposition 4 shows (b) because the policy that belongs to D^σ minimizes $L_u^1 v = T_u v$ if $v \in V_H^\sigma \subset V^\sigma$. □

We now present the following remark.

Remark 2. For the above results, we assume that the pendant edges do not have other pendant edges in their neighborhoods. We now explain how the results of this section also hold when in the neighborhood of an pendant edge there are other pendant edges. An example of such a matching model consists of the matching graph of Fig. 4 with an additional demand node d_7 and an edge $(7,5)$.

We first note that, if the cost of the pendant edges that are neighbors is the same, these edges can be merged and seen as a single edge whose arrival probability is equal to the sum of the arrival probabilities of the merged nodes. Otherwise, we require that the undesirability and increasing properties to be satisfied by the pendant edge with the highest cost and the above arguments can be used to show that an optimal policy prioritizes the most expensive pendant edges.

7. Optimal policy of the W-shaped graph

We now focus on a W-shaped matching graph. This matching graph is formed by three demand nodes, two supply nodes and the following set of edges: $\mathcal{E} = \{(1,1),(2,1),(2,2),(3,2)\}$. Let us also define $(1,2)$ and $(3,1)$ as the imaginary edge between d_1 and s_2 and between d_3 and s_1, respectively, that we introduce to ease the notations. Note that $(1,2) \notin \mathcal{E}$ and $(3,1) \notin \mathcal{E}$.

We assume that $\alpha_1 < \beta_1$ and $\alpha_3 < \beta_2$ to ensure the stability of the system.

We remark that the W-shaped matching graph is the simplest bipartite graph that one can consider (except for the N-shaped matching graph, which has been studied in Section 4). Therefore, the goal of this section is twofold: (*i*) under the condition of the holding costs of Section 6, we aim to go beyond the result of Theorem 4 and provide the structure of an optimal policy of this matching graph; and (*ii*) we study an optimal policy when the conditions of Section 6 do not hold. We first focus on the former case and then in the later one.

7.1. Higher cost on the pendant nodes

We now consider that the holding cost of the pendant nodes is larger than the holding cost of the middle node, i.e, $c_{d_1} \geq c_{d_2}$ and $c_{d_3} \geq c_{d_2}$.

7.1.1. Value functions properties

We present the properties that are needed to characterize an optimal matching policy for the W-shaped graph. First, we consider the increasing and undesirability properties in the pendant edges $(1,1)$ and $(3,2)$ as defined in Definitions 1 and 8. We also consider the convexity property in $(2,1)$ and $(2,2)$ (see Definition 2). Finally, we present two additional properties: the exchangeability and the modularity properties. These properties are required to prove that the optimal decision rule has a threshold in $(2,1)$ and in $(2,2)$.

Definition 9 (*Exchangeable Property*)**.** A function v is exchangeable in $(2,1)$ and $(3,1)$ or $v \in \mathcal{H}_{(2,1),(3,1)}$ if $\forall a \in \mathcal{A}$, $\forall q \in \mathcal{Q}$,

$$v(q + e_{(2,1)}, a) - v(q, a) = v(q + e_{(3,1)}, a) - v(q - e_{(2,1)} + e_{(3,1)}, a).$$

Likewise, v is exchangeable in $(2,2)$ and $(1,2)$ or $v \in \mathcal{H}_{(2,2),(1,2)}$ if $\forall a \in \mathcal{A}$, $\forall q \in \mathcal{Q}$,

$$v(q + e_{(2,2)}, a) - v(q, a) = v(q + e_{(1,2)}, a) - v(q - e_{(2,2)} + e_{(1,2)}, a).$$

Definition 10 (*Modularity Property*)**.** A function v is modular in $(2,1)$ and $(2,2)$ or $v \in \mathcal{M}_{(2,1),(2,2)}$ if $\forall a \in \mathcal{A}$, $\forall q \in \mathcal{Q}$,

$$v(q + e_{(2,1)} + e_{(2,2)}, a) - v(q + e_{(2,1)}, a) = v(q + e_{(2,2)}, a) - v(q, a).$$

7.1.2. Optimal decision rule

In this section, we show that for any v satisfying the aforementioned properties, there is decision rule which is of threshold-type in $(2,1)$ and $(2,2)$ and prioritizes the items in the pendant edges. Let us first present a formal definition of the decision rule under consideration.

Definition 11 (*Threshold-type Decision Rule*)**.** A decision rule u_x is said to be of threshold type in $(2,1)$ and $(2,2)$ with priority to $(1,1)$ and $(3,2)$ if:

1. it matches all of $(1,1)$ and $(3,2)$.
2. it matches items of $(2,1)$ only if the remaining jobs in s_1 are above a specific threshold, denoted by $t_{(2,1)}$ with $t_{(2,1)} \in \mathbb{N} \cup \infty$ and matches items in $(2,2)$ only if the remaining jobs in s_2 are above a specific threshold, denoted by $t_{(2,2)}$ with $t_{(2,2)} \in \mathbb{N} \cup \infty$.

In other words, $u_x = \min(x_{d_1}, x_{s_1})e_{(1,1)} + \min(x_{d_3}, x_{s_2})e_{(3,2)} + \min(k_{t_{(2,1)}}(x), x_{d_2})e_{(2,1)} + \min(j_{t_{(2,2)}}(x), x_{d_2})e_{(2,2)}$ where

$$k_{t_{(2,1)}}(x) = \begin{cases} 0 & \text{if } x_{s_1} - x_{d_1} \leq t_{(2,1)} \\ x_{s_1} - x_{d_1} - t_{(2,1)} & \text{otherwise} \end{cases}$$

and $$j_{t_{(2,2)}}(x) = \begin{cases} 0 & \text{if } x_{s_2} - x_{d_3} \leq t_{(2,2)} \\ x_{s_2} - x_{d_3} - t_{(2,2)} & \text{otherwise} \end{cases}$$

Let D^σ be the set of decision rules as defined above. We now show that the above decision rule minimizes $L_u^\theta v$.

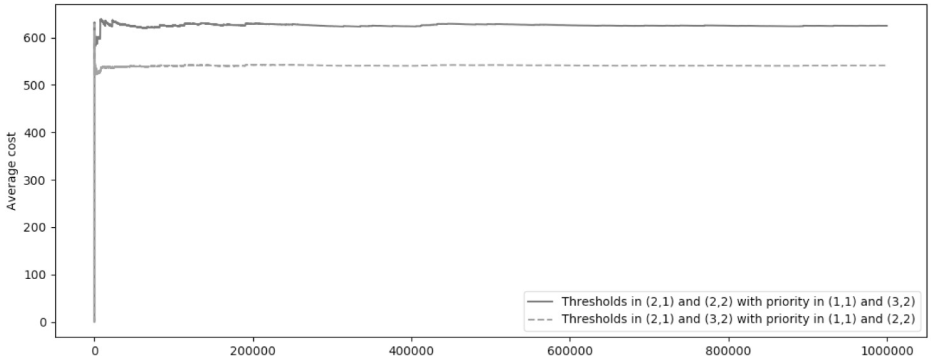

Fig. 5. W-shaped model. Average cost of our conjecture and a similar policy but with a priority in $(2, 2)$ instead of $(3, 2)$.

Proposition 5. *Let $v \in \mathcal{I}_{(1,1)} \cap \mathcal{I}_{(3,2)} \cap \mathcal{U}_{(1,1)} \cap \mathcal{U}_{(3,2)} \cap \mathcal{C}_{(2,1)} \cap \mathcal{C}_{(2,2)} \cap \mathcal{H}_{(2,1),(3,1)} \cap \mathcal{H}_{(2,2),(1,2)} \cap \mathcal{M}_{(2,1),(2,2)}$. Let $0 \le \theta \le 1$. There exists $u^* \in U_x$ such that u^* is a matching policy of threshold type in $(2, 1)$ and $(2, 2)$ with priority to $(1, 1)$ and $(3, 2)$ (as defined in Definition 11) and $u^* \in \arg\min_{u \in U_x} L_u^{\theta} v(x)$. In particular, this result holds for the average operator: $L_u^1 = T_u$.*

Proof. The idea of the proof is similar to the one of Proposition 2. However, one has to use the property $\mathcal{H}_{(2,1),(3,1)}$ to handle $(3, 1)$ left over items when defining the threshold in $(2, 1)$ and also the property $\mathcal{H}_{(2,2),(1,2)}$ to handle $(1, 2)$ left over items when defining the threshold in $(2, 2)$. It is also needed the property $\mathcal{M}_{(2,1),(2,2)}$ to prove the independence between the two thresholds. Detailed proof is provided in Appendix J. □

7.1.3. Structure of an optimal policy

We want to characterize an optimal matching policy in the W-shaped graph with higher costs in the pendant edges. It follows directly from the reasoning of Section 6 that matching all the items of the pendant edges $(1, 1)$ and $(3, 2)$ is optimal. Therefore, the question now is what to do with the items of $(2, 1)$ and $(2, 2)$. In the N-shaped graph, the non-pendant edge is matched only if the number of items in this edge is above a threshold. Therefore, in the W-shaped graph with higher costs in the pendant edges, an optimal matching policy seems to be of threshold type as defined in Definition 11. However, if we want to apply Theorem 6 and Theorem 7 to characterize an optimal matching policy, we need to show the preservation by the Dynamic Programming of the following properties of the value function: $\mathcal{I}_{(1,1)} \cap \mathcal{I}_{(3,2)} \cap \mathcal{U}_{(1,1)} \cap \mathcal{U}_{(3,2)} \cap \mathcal{C}_{(2,1)} \cap \mathcal{C}_{(2,2)} \cap \mathcal{H}_{(2,1),(3,1)} \cap \mathcal{H}_{(2,2),(1,2)} \cap \mathcal{M}_{(2,1),(2,2)}$. Given the difficulty of the aforementioned set of properties, we did not succeed in showing that all the properties are preserved by the Dynamic Programming operator (even if we consider additional properties such as, for instance, the boundary property of Definition 3). Hence, we conjecture the existence of a set V^{σ} that contains the properties $\mathcal{I}_{(1,1)} \cap \mathcal{I}_{(3,2)} \cap \mathcal{U}_{(1,1)} \cap \mathcal{U}_{(3,2)} \cap \mathcal{C}_{(2,1)} \cap \mathcal{C}_{(2,2)} \cap \mathcal{H}_{(2,1),(3,1)} \cap \mathcal{H}_{(2,2),(1,2)} \cap \mathcal{M}_{(2,1),(2,2)}$ and such that if $v \in V^{\sigma}$, then $L^{\theta} v \in V^{\sigma}$. Under this conjecture, we can use the result of Theorems 6 and 7 to prove that an optimal matching policy for the W-shaped graph with higher costs on the pendant edges consists of decision rules that are of threshold-type as defined in Definition 11.

7.2. Higher cost on the middle node

We consider the W-shaped graph when the cost on the pendant nodes is not larger. We now present our numerical work that shows that prioritizing the items of the pendant edges is not optimal for this case. We set $c_{d_1} = c_{d_2} = 10$, $c_{d_3} = c_{s_1} = 1$, $c_{s_2} = 1000$, $\alpha_1 = 0.4$, $\alpha_2 = 0.35$, $\alpha_3 = 0.25$, $\beta_1 = 0.5$ and $\beta_2 = 0.5$. For these parameters, we illustrate in Fig. 5 the evolution of the average cost of the system for two policies: (i) a policy of threshold type in $(2, 1)$ and $(3, 2)$ with priority in $(1, 1)$ and $(2, 2)$ and (ii) policy of threshold type in $(2, 1)$ and $(2, 2)$ with priority in $(1, 1)$ and $(3, 2)$. For both policies, we selected the best thresholds by numerical experiments which are $t_{(2,1)} = 14$ and $t_{(3,2)} = 0$ for the former policy and $t_{(2,1)} = 11$ and $t_{(2,2)} = 0$ for the latter policy. We observe that the former policy outperforms the later policy. This experiment shows that the matching policy that seems to be optimal in the W-shaped matching graph with higher cost in the pendant edges is not optimal when the largest cost is in the middle edge. As a result, we can claim that, in an arbitrary matching graph, an optimal policy is still unknown, but it seems to depend not only on the topology of the matching graph, but also on the costs of the nodes. This makes the characterization of an optimal matching policy of an arbitrary matching graph an pendantly difficult and challenging task.

8. Conclusion

We consider bipartite matching graphs with linear costs in the buffer size. We model this problem as a Markov Decision Process where the goal is to find an optimal matching control, that is, how to match items so as to minimize the cost of the system. We study the derived optimal control problem for the discounted cost and the average cost problems. In both cases, we characterize the structure of an optimal policy in a wide range of matching models. First, for the N-shaped matching graph, we show that there exists an optimal policy that is of threshold type for the diagonal edge with priority to the end edges of the matching graph. Furthermore, we characterize the optimal threshold for the average cost problem. Then, we show how the results obtained for the N-shaped graph can be used to provide the structure of a quasicomplete graph with an arbitrary number of nodes. Additionally, we consider an arbitrary bipartite graph and we show that, when the cost of the pendant nodes is larger or equal than the cost of its neighbor nodes, an optimal policy always matches all the items in the pendant edges. Finally, we investigate the W-shaped matching model. For this case, we conjecture that an optimal matching policy is also of threshold type with priority to the pendant edges when the cost of the pendant nodes is large and we provide numerical evidence that shows that, when the cost of the middle node is large, an optimal matching control does not prioritize the pendant edges. The conclusion of the analysis carried out for the W-shaped graph is that an optimal matching control seems to depend on the cost of the nodes.

The characterization of an optimal matching policy in an arbitrary bipartite matching graph remains as an open question. For future work, it would be interesting to prove our conjecture of an optimal matching policy on the W-shaped matching graph when the cost of the pendant nodes is large and characterize an optimal policy when the cost of the intermediate node is large. Another possible future research is the study of an optimal matching policy for general (non-bipartite) graphs. In this case, items arrive one by one to the system and there are not two different sets of nodes. In spite of these differences with respect to bipartite matching graphs, we think that the techniques applied in this article can be also used to characterize an optimal matching policy for non-bipartite matching graphs.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

The work of Josu Doncel has been supported by the Department of Education of the Basque Government, Spain through the Consolidated Research Group MATHMODE (IT1294-19), by the Marie Sklodowska-Curie, Spain grant agreement No 777778 and by the Spanish Ministry of Science and Innovation, Spain with reference PID2019-108111RB-I00 (FEDER/AEI).

This work was also funded by the French National Research Agency, France grant ANR-16-CE05-0008.

Appendix A. Statement and discussion of Theorem 6 and of Theorem 7

Theorem 6 (*Adapted from [25, Theorem 6.11.3]*). *Assume that the following properties hold: there exists positive function w on $\mathcal{Y}$ such that*

$$\sup_{y} \frac{c(y)}{w(y)} < +\infty, \tag{7}$$

$$\sup_{(y,u)} \frac{1}{w(y)} \sum_{y'} \mathbb{P}(y'|y,u)w(y') < +\infty, \tag{8}$$

and for every μ, $0 \le \mu < 1$, there exist η, $0 \le \eta < 1$ and some integer J such that for every J-tuple of Markov deterministic decision rules $\pi = (u_1, \dots, u_J)$ and every y

$$\mu^J \sum_{y'} P_\pi(y'|y)w(y') < \eta w(y), \tag{9}$$

where P_π denotes the J-step transition matrix under policy π. Let $0 \le \theta < 1$. Let V_w the set of functions in the state space which have a finite w-weighted supremum norm, i.e., $\sup_y |v(y)/w(y)| < +\infty$. Assume that

(*) for each $v \in V_w$, there exists a deterministic Markov decision rule u such that $L^\theta v = L^\theta_u v$.

Let V^σ and D^σ be such that

(a) $v \in V^\sigma$ implies that $L^\theta v \in V^\sigma$;
(b) $v \in V^\sigma$ implies that there exists a decision $u' \in D^\sigma$ such that $u' \in \arg\min_u L^\theta_u v$;
(c) V^σ is a closed subset of the set of value functions under pointwise convergence.

Then, there exists an optimal stationary policy $\pi^* = (u^*(X(n)))_{n\ge1}$ that belongs to Π^σ with $u^* \in \arg\min_u L^\theta_u v$.

The above statement has been previously presented in [26, Theorem 1] and it has the following nice properties: (i) it removes the need to verify that $V^\sigma \subset V_w$ and (ii) its statement separates the structural requirements (*(a)*, *(b)* and *(c)*) from the technical requirements related to the unboundedness of the cost function ((7), (8), (9) and $(*)$).

In order to prove the optimality in the average cost case, we will use the following result:

Theorem 7. *[25, Theorem 8.11.1] Suppose that the following properties hold:*

(A1) $\exists\, C \in \mathbb{R}, \forall y = (q, a) \in \mathcal{Y}, x = q + a,\ -\infty < C \le c(x) < +\infty,$
(A2) $\forall y \in \mathcal{Y}, \forall\, 0 \le \theta < 1,\ v_\theta^*(y) < +\infty$
(A3) $\exists\, H \in \mathbb{R}, \forall y \in \mathcal{Y}, \forall\, 0 \le \theta < 1,\ -\infty < H \le v_\theta^*(y) - v_\theta^*(0)$
(A4) There exists a nonnegative function $M(y)$ *such that*

- *(a)* $\forall y \in \mathcal{Y},\ M(y) < +\infty$
- *(b)* $\forall y \in \mathcal{Y}, \forall\, 0 \le \theta < 1,\ v_\theta^*(y) - v_\theta^*(0) \le M(y)$
- *(c) There exists* $u \in U_0$ *for which* $\sum_y \mathbb{P}(y|0, u)M(y) < +\infty$

Let H *and* M *be defined in Assumptions (A3) and (A4). We define a subset* V_H^σ *of* V^σ *which contains all the value functions* $v \in V^\sigma$ *such that* $H \le v(y) - v(0) \le M(y)$ *for all* $y \in \mathcal{Y}$*. Then, if*

(a) for any sequence $(\theta_n)_{n\ge 0}$, $0 \le \theta_n < 1$, *for which* $\lim_{n\to+\infty} \theta_n = 1$,

$$\lim_{n\to+\infty} [v_{\theta_n}^* - v_{\theta_n}^*(0)e] \in V_H^\sigma \quad \text{with } e(y) = 1 \text{ for all } y \in \mathcal{Y}$$

and

(b) $v \in V_H^\sigma$ *implies that there exists a decision* $u' \in D^\sigma$ *such that* $u' \in \arg\min_u T_u v$*;*

Then $D^\sigma \cap \arg\min_u T_u v \neq \emptyset$ *and* $u^* \in D^\sigma \cap \arg\min_u T_u v$ *implies that the stationary matching policy which uses* u^* *is lim sup average optimal.*

Remark 3. When $v_\theta(0)$ is used in this theorem, we mean $v_\theta(y)$ with $y = (0, a)$ for any $a \in \mathcal{A}$.

We now show that the technical requirements of Theorem 6 due to the unboundness of the costs are verified in our model since we are considering that the cost function is linear.

Lemma 8. *The conditions ((7), (8), (9) and* $(*)$*) are verified.*

Proof. We show that the technical details given in (7)–(9) are verified when c is linear. Let $y = (q, a) \in \mathcal{Q} \times \mathcal{A}$, $x = q + a$, we choose $w(y) = \sum_i x_i$. In our case, the cost is a linear function of x therefore, $c(y)/w(y) \le \max_{i\in\mathcal{D}\cup\mathcal{S}} c_i$. This shows (7). In addition,

$$\frac{1}{w(y)} \sum_{y'} \mathbb{P}(y'|y, u)w(y') = \mathbb{E}\left[\frac{w(x - u + a')}{w(y)}\Big|A = a'\right] \le E\left[\frac{w(x + a')}{w(y)}\Big|A = a'\right] = \frac{\sum_i x_i + 2}{\sum_i x_i} \le 2$$

since $w(y)$ is increasing and two items arrive to the system in each step following a process which is independent of the state of the system. This shows (8). Finally, we can repeat the previous argument to show that for all J-step matching policy π

$$\sum_{y'} P_\pi(y'|y)w(y') \le \sum_{y'} P_{\pi_0}(y'|y)w(y') = w(y) + 2J.$$

where $\pi_0 = (0, \ldots, 0)$ is the policy which does not match any items. Therefore, (9) is satisfied if there exist J integer and $\eta < 1$ such that

$$\mu^J(w(y) + 2J) \le \eta w(y) \iff \eta > \frac{\mu^J(w(y) + 2J)}{w(y)}.$$

Since it is decreasing with J and when $J \to \infty$ it tends to zero, there exists a J integer such that η is less than one and, therefore, (9) is also verified.

Since for each state of the system, the set of admissible matching policies is finite, it follows that (*) holds. □

We now show that Assumptions (*A1*) to (*A4*) of Theorem 7 hold in our model since the cost function under consideration is a linear function.

Lemma 9. *Assumptions (A1) to (A4) hold.*

Proof. First, we note that the cost function is nonnegative and therefore Assumption (*A1*) holds using $C = 0$. Let π_{ML} be the policy Match the Longest as defined in [24, Definition 2.6]. This policy is stable for any bipartite matching graph as proved in [24, Theorem 7.1], which means that the derived Markov chain is positive recurrent and $g^{\pi_{ML}} < \infty$. Moreover, the set $\{y \in \mathcal{Y} : c(y) < g^{\pi_{ML}}\}$ is nonempty because $g^{\pi_{ML}} > \min_{(i,j)\in\mathcal{A}} c_{d_i} + c_{s_j}$ almost surely and $c(0, a) = c_{d_i} + c_{s_j}$ (for any $a = e_{(i,j)} \in \mathcal{A}$). It is also finite because $g^{\pi_{ML}} < \infty$ and c is increasing in y (because c is linear). Therefore, we can use [25, Theorem 8.10.9] and Assumptions (*A2*) to (*A4*) hold. □

Appendix B. Proof of Proposition 1

Let $v \in \mathcal{I}_{(1,1)} \cap \mathcal{I}_{(2,2)} \cap \mathcal{I}_{(2,1)}$, $0 \le \theta \le 1$, $q \in \mathcal{Q}$, $a \in \mathcal{A}$, $x = q + a$ and $u \in U_x$. The maximum number of items that can be matched in $(1, 1)$ is denoted by $m_{(1,1)} = \min(x_{d_1}, x_{s_1})$ and in $(2, 2)$ by $m_{(2,2)} = \min(x_{d_2}, x_{s_2})$.

Let $p_{(1,2)} = \min(u_{(1,2)}, x_{s_1} - u_{(1,1)}, x_{d_2} - u_{(2,2)})$ be the number of items that are matched in $(1, 2)$ under the control u that can be transformed from matching items in $(1, 1)$ and $(2, 2)$ at the same time. We define a policy $u^{p_{(1,2)}}$ that removes the $p_{(1,2)}$ items in $(1, 2)$ and matches $p_{(1,2)}$ items in $(1, 1)$ and $(2, 2)$, that is, $u^{p_{(1,2)}} = u + p_{(1,2)}(e_{(1,1)} + e_{(2,2)} - e_{(1,2)})$. Using (3), we verify that this policy is admissible, i.e. $u^{p_{(1,2)}} \in U_x$: $u^{p_{(1,2)}} \in \mathbb{N}^4$ is true because $u \in U_x$. *(a)* is true because $u^{p_{(1,2)}}_{(2,2)} = u_{(2,2)} + p_{(1,2)} \le x_{d_2}$. *(b)* is true because $u^{p_{(1,2)}}_{(1,1)} = u_{(1,1)} + p_{(1,2)} \le x_{s_1}$. Then, since $v \in \mathcal{I}_{(2,1)}$, it follows that $L^{\theta}_{u^{p_{(1,2)}}} v(q, a) \le L^{\theta}_u v(q, a)$.

We now define u^* as the decision rule that matches all the items $(1, 1)$ and $(2, 2)$ of $x - u^{p_{(1,2)}}$, that is, of the remaining items when we apply $u^{p_{(1,2)}}$. Hence, we have that $u^* = u^{p_{(1,2)}} + e_{(1,1)}(m_{(1,1)} - u^{p_{(1,2)}}_{(1,1)}) + e_{(2,2)}(m_{(2,2)} - u^{p_{(1,2)}}_{(2,2)})$. Using (3), we also verify that $u^* \in U_x$: $u^* \in \mathbb{N}^4$ is true because $u^{p_{(1,2)}} \in U_x$, $m_{(1,1)} \ge 0$ and $m_{(2,2)} \ge 0$. We immediately get that $u^*_{(2,2)} = m_{(2,2)} \le x_{d_2}$ and $u^*_{(1,1)} = m_{(1,1)} \le x_{s_1}$. For x_{d_1} and x_{s_2}, we distinguish the following cases:

1. If $p_{(1,2)} = u_{(1,2)}$. Then we have $u^*_{(1,2)} = u^{p_{(1,2)}}_{(1,2)} = 0$. Thus,
$$u^*_{(1,1)} + u^*_{(1,2)} = m_{(1,1)} \le x_{d_1}$$
$$u^*_{(2,2)} + u^*_{(1,2)} = m_{(2,2)} \le x_{s_2}$$
2. If $p_{(1,2)} = x_{s_1} - u_{(1,1)}$. Then,
$$u^*_{(1,1)} + u^*_{(1,2)} = m_{(1,1)} + u_{(1,2)} - x_{s_1} + u_{(1,1)} \le u_{(1,2)} + u_{(1,1)} \le x_{d_1}$$
$$u^*_{(2,2)} + u^*_{(1,2)} = m_{(2,2)} + u_{(1,2)} - x_{s_1} + u_{(1,1)} \le x_{d_2} + u_{(1,2)} - x_{s_1} + u_{(1,1)} = x_{s_2} + u_{(1,2)} - x_{d_1} + u_{(1,1)} \le x_{s_2}$$
3. If $p_{(1,2)} = x_{d_2} - u_{(2,2)}$. Then,
$$u^*_{(1,1)} + u^*_{(1,2)} = m_{(1,1)} + u_{(1,2)} - x_{d_2} + u_{(2,2)} \le x_{s_1} + u_{(1,2)} - x_{d_2} + u_{(2,2)} = x_{d_1} + u_{(1,2)} - x_{s_2} + u_{(2,2)} \le x_{d_1}$$
$$u^*_{(2,2)} + u^*_{(1,2)} = m_{(2,2)} + u_{(1,2)} - x_{d_2} + u_{(2,2)} \le u_{(1,2)} + u_{(2,2)} \le x_{s_2}$$

In all the cases, *(a)* and *(b)* are true. Hence, since $v \in \mathcal{I}_{(1,1)} \cap \mathcal{I}_{(2,2)}$, it results that $L^{\theta}_{u^*} v(q, a) \le L^{\theta}_{u^{p_{(1,2)}}} v(q, a)$ and, as a consequence, $L^{\theta}_{u^*} v(q, a) \le L^{\theta}_u v(q, a)$ for any $u \in U_x$.

Appendix C. Proof of Proposition 2

Let $q \in \mathcal{Q}$, $a \in \mathcal{A}$, $x = q + a$ and $u \in U_x$. We define $m_{(1,1)} = \min(x_{s_1}, x_{d_1})$ and $m_{(2,2)} = \min(x_{s_2}, x_{d_2})$. Since $v \in \mathcal{I}_{(1,1)} \cap \mathcal{I}_{(2,2)} \cap \mathcal{I}_{(2,1)}$, it follows from Proposition 1 that there exists $u' \in U_x$ that matches all the items of the pendant edges and $L^{\theta}_{u'} v(q, a) \le L^{\theta}_u v(q, a)$. Therefore, u' is of the following form: $u' = m_{(1,1)}e_{(1,1)} + m_{(2,2)}e_{(2,2)} + ke_{(1,2)}$ with $k \in K_x$. We now prove that there exists $t \in \mathbb{N} \cup \infty$ such that

$$L^{\theta}_{u^*} v(q, a) \le L^{\theta}_{u'} v(q, a), \quad \forall k \in K_x \tag{10}$$

where u^* is a decision rule of threshold type in the diagonal edge with priority to the pendant edges (see Definition 4). If $x_{s_1} \ge x_{d_1}$, then $K_x = 0$ and, as a result, we have that $u^* = u'$. We now consider that $x_{s_1} < x_{d_1}$, in which case $K_x \ne \{0\}$. For this case, the state of the system after applying a threshold type decision rule (u^* or u') is of the form $(l, 0, 0, l)$. Therefore, to compare $L^{\theta}_{u^*} v(q, a)$ with $L^{\theta}_{u'} v(q, a)$, we only need to compare $\mathbb{E}[v(j^* e_{(1,2)}, A)]$ with $\mathbb{E}[v(j' e_{(1,2)}, A)]$, where $j^*, j' \in K_x$. Since v is convex in $(1, 2)$, we distinguish the following cases:

- First, we consider that $\forall j \in \mathbb{N}$, $\mathbb{E}[v((j + 1)e_{(1,2)}, A)] - \mathbb{E}[v(je_{(1,2)}, A)] \le 0$. For this case, we define u^* as follows: $u^* = m_{(1,1)}e_{(1,1)} + m_{(2,2)}e_{(2,2)}$ (note that this is the same as choosing $t = \infty$ and therefore $k_t(x) = 0$). Since $\mathbb{E}[v((j + 1)e_{(1,2)}, A)] - \mathbb{E}[v(je_{(1,2)}, A)] \le 0$, it follows that
$$L^{\theta}_{u^*} v(q, a) \le L^{\theta}_{u^*+e_{(1,2)}} v(q, a) \le \cdots \le L^{\theta}_{u^*+ke_{(1,2)}} v(q, a)$$
for all $k \in K_x$. Therefore, it follows (10).

- We now consider that $\mathbb{E}[v(e_{(1,2)}, A)] - \mathbb{E}[v(0, A)] \geq 0$. For this case, we define u^* as follows: $u^* = m_{(1,1)}e_{(1,1)} + m_{(2,2)}e_{(2,2)} + (x_{d_1} - x_{s_1})e_{(1,2)}$ (note that this is the same as choosing $t = 0$ and $k_t(x) = x_{d_1} - x_{s_1}$). Using that $\mathbb{E}[v(e_{(1,2)}, A)] - \mathbb{E}[v(0, A)] \geq 0$ and also that v is convex in $(1, 2)$, it results that
$$L^{\theta}_{u^*} v(q, a) \leq L^{\theta}_{u^*-e_{(1,2)}} v(q, a) \leq \cdots \leq L^{\theta}_{u^*-ke_{(1,2)}} v(q, a)$$
for all $k \in K_x$ and since $u^* - (x_{d_1} - x_{s_1} - k)e_{(1,2)} = u'$ (with $x_{d_1} - x_{s_1} - k \in K_x$, $\forall\, k \in K_x$), it follows (10).
- Finally, we consider that $\exists j \in \mathbb{N}^*,\ \mathbb{E}[v((j+1)e_{(1,2)}, A)] - \mathbb{E}[v(je_{(1,2)}, A)] \geq 0$. Let $\underline{j} = \min\{j \in \mathbb{N}^* :\ \mathbb{E}[v((j+1)e_{(1,2)}, A)] - \mathbb{E}[v(je_{(1,2)}, A)] \geq 0\}$. By definition of $\underline{j}$ and by convexity of v in $(1, 2)$, we have
$$\mathbb{E}[v((\underline{j} - l)e_{(1,2)}, A)] - \mathbb{E}[v((\underline{j} - l - 1)e_{(1,2)}, A)] \leq 0 \quad \forall l \in [\![0\ ;\ \underline{j} - 1]\!] \tag{11}$$
and
$$\mathbb{E}[v((\underline{j} + 1 + l)e_{(1,2)}, A)] - \mathbb{E}[v((\underline{j} + l)e_{(1,2)}, A)] \geq 0 \quad \forall l \in \mathbb{N} \tag{12}$$
We define u^* as the decision rule of Definition 4 with $t = \underline{j}$. We first consider that $x_{d_1} - x_{s_1} \leq \underline{j}$, then we have $k_t(x) = 0$ and $L^{\theta}_{u^*} v(q, a) \leq L^{\theta}_{u'} v(q, a)$ for all $k \in K_x$ by (11) (note that $0 \leq k \leq x_{d_1} - x_{s_1} \leq \underline{j}$), which shows (10). We now consider that $x_{d_1} - x_{s_1} > \underline{j}$, then $k_t(x) = x_{d_1} - x_{s_1} - \underline{j}$ and $L^{\theta}_{u^*} v(q, a) = c(x) + \theta\mathbb{E}[v(\underline{j}e_{(1,2)}, A)]$. Therefore, for all $k \in K_x$, $L^{\theta}_{u^*} v(q, a) \leq L^{\theta}_{u'} v(q, a)$ by (11) if $k \geq \underline{j}$ or by (12) if $k \leq \underline{j}$, which proves (10).

Appendix D. Proofs of Section 4.3

D.1. *Proof of Lemma 1*

We first show that if $v \in I_{(1,1)}$, then $L^{\theta} v \in I_{(1,1)}$. Let $\underline{q} \in \mathcal{Q}$ and $a \in \mathcal{A}$, $\underline{x} = \underline{q} + a$. We define $\overline{q} = \underline{q} + e_{(1,1)}$, $\overline{x} = \overline{q} + a$. Since v is increasing with $(1, 1)$, we have that $v(\overline{q}, a) \geq v(\underline{q}, a)$. We aim to show that $L^{\theta} v(\overline{q}, a) \geq L^{\theta} v(\underline{q}, a)$.

Let $u_{\overline{x}} \in \arg\min_{u \in U_{\overline{x}}} L^{\theta}_{u} v(\overline{q}, a)$. Since $(\overline{x})_{s_1} \geq 1$ and $(\overline{x})_{d_1} \geq 1$, from Proposition 1, it follows that $(u_{\overline{x}})_{(1,1)} = \min(\overline{x}_{d_1}, \overline{x}_{s_1}) \geq 1$. Therefore, we define $u_{\underline{x}} = u_{\overline{x}} - e_{(1,1)}$. Thus, it is easy to see that $u_{\underline{x}} \in U_{\underline{x}}$ because $u_{\overline{x}} \in U_{\overline{x}}$. Moreover, we observe that $\overline{x} - u_{\overline{x}} = \underline{x} - u_{\underline{x}}$ and, since c is a linear function, $c(\overline{x}) > c(\underline{x})$. Hence,
$$\begin{aligned}
L^{\theta}_{u_{\underline{x}}} v(\underline{q}, a) &= c(\underline{x}) + \theta\mathbb{E}[v(\underline{x} - u_{\underline{x}}, A)] \\
&= c(\underline{x}) - c(\overline{x}) + c(\overline{x}) + \theta\mathbb{E}[v(\overline{x} - u_{\overline{x}}, A)] \\
&= c(\underline{x}) - c(\overline{x}) + L^{\theta} v(\overline{q}, a) \\
&< L^{\theta} v(\overline{q}, a).
\end{aligned}$$
And, since $u_{\underline{x}} \in U_{\underline{x}}$, then by definition $L^{\theta} v(\underline{q}, a) \leq L^{\theta}_{u_{\underline{x}}} v(\underline{q}, a)$ and, as a result, the desired result follows.

To prove that if $v \in I_{(2,2)}$, then $L^{\theta} v \in I_{(2,2)}$, one can use the same arguments as above with $\overline{q} = \underline{q} + e_{(2,2)}$. We omit it for clarity of the presentation.

The proof of the preservation of the increasing property in $\mathcal{I}_{(2,1)}$ is also similar but requires to handle the case when no items can be matched in $(1, 2)$. Let $\underline{q} \in \mathcal{Q}$ and $a \in \mathcal{A}$, $\underline{x} = \underline{q} + a$. Let $\overline{q} = \underline{q} + e_{(1,1)} + e_{(2,2)} - e_{(1,2)}$, $\overline{x} = \overline{q} + a$. Since $v \in \mathcal{I}_{(2,1)}$, we know that $v(\underline{q}, a) \leq v(\overline{q}, a)$. Besides, $c(\underline{x}) < c(\overline{x})$ holds because c is a linear function of x. We aim to show that $L^{\theta} v(\underline{q}, a) \leq L^{\theta} v(\overline{q}, a)$.

Let $u_{\overline{x}} \in \arg\min_{u \in U_{\overline{x}}} L^{\theta}_{v} v(\overline{q}, a)$. From Proposition 1, we know that $(u_{\overline{x}})_{(2,2)} = \min(\overline{x}_{d_2}, \overline{x}_{s_2})$ and $(u_{\overline{x}})_{(1,1)} = \min(\overline{x}_{d_1}, \overline{x}_{s_1})$. We first consider that $\underline{x}_{d_1} \geq 1$ and $\underline{x}_{s_2} \geq 1$. We define $u_{\underline{x}} = u_{\overline{x}} - e_{(1,1)} - e_{(2,2)} + e_{(1,2)}$. Thus, we have that $\underline{x} - u_{\underline{x}} = \overline{x} - u_{\overline{x}}$ and $u_{\underline{x}} \in U_{\underline{x}}$ because $u_{\overline{x}} \in U_{\overline{x}}$ as well as $\underline{x}_{s_1} = \overline{x}_{s_1} - 1 \geq (u_{\overline{x}})_{(1,1)} - 1 = \min(\overline{x}_{d_1}, \overline{x}_{s_1}) - 1 = \min(\underline{x}_{d_1}, \underline{x}_{s_1} + 1) - 1 \geq 0$, $\underline{x}_{d_2} = \overline{x}_{d_2} - 1 \geq (u_{\overline{x}})_{(2,2)} - 1 = \min(\overline{x}_{d_2}, \overline{x}_{s_2}) - 1 = \min(\underline{x}_{d_2} + 1, \underline{x}_{s_2}) - 1 \geq 0$. Thus,
$$\begin{aligned}
L^{\theta}_{u_{\underline{x}}} v(\underline{q}, a) &= c(\underline{x}) + \theta\mathbb{E}[v(\underline{x} - u_{\underline{x}}, A)] \\
&= c(\underline{x}) - c(\overline{x}) + c(\overline{x}) + \theta\mathbb{E}[v(\overline{x} - u_{\overline{x}}, A)] \\
&= c(\underline{x}) - c(\overline{x}) + L^{\theta} v(\overline{q}, a) \\
&< L^{\theta} v(\overline{q}, a).
\end{aligned}$$
And since $u_{\underline{x}} \in U_{\underline{x}}$, then $L^{\theta} v(\underline{q}, a) \leq L^{\theta}_{u_{\underline{x}}} v(\underline{q}, a)$ and from the above inequality, the desired result follows.

We now consider that $\underline{x}_{d_1} = 0$ or $\underline{x}_{s_2} = 0$. In that case, we cannot match more items in state $\bar{x}$ than in state $\underline{x}$, which implies that $u_{\bar{x}} \in U_{\underline{x}}$. As a result,

$$\begin{aligned} L^{\theta}_{u_{\bar{x}}} v(\underline{q}, a) &= c(\underline{x}) + \theta\mathbb{E}[v(\underline{x} - u_{\bar{x}}, A)] \\ &\le c(\underline{x}) + \theta\mathbb{E}[v(\bar{x} - u_{\bar{x}}, A)] \text{ since } v \in \mathcal{I}_{(2,1)} \\ &= c(\underline{x}) - c(\bar{x}) + c(\bar{x}) + \theta\mathbb{E}[v(\bar{x} - u_{\bar{x}}, A)] \\ &= c(\underline{x}) - c(\bar{x}) + L^{\theta} v(\bar{q}, a) \\ &< L^{\theta} v(\bar{q}, a). \end{aligned}$$

And since $u_{\underline{x}} \in U_{\underline{x}}$, then $L^{\theta} v(\underline{q}, a) \le L^{\theta}_{u_{\underline{x}}} v(\underline{q}, a)$ and from the above inequality, the desired result follows.

D.2. Proof of Lemma 2

Let $\underline{q} \in \mathcal{Q}$, $\underline{q}_{d_1} \ge \underline{q}_{s_1}$, $a \in \mathcal{A}$, $\underline{x} = \underline{q} + a$, $\bar{q} = \underline{q} + e_{(1,2)}$, $\bar{x} = \bar{q} + a$, $\bar{\bar{q}} = \bar{q} + e_{(1,2)}$ and $\bar{\bar{x}} = \bar{\bar{q}} + a$. Since v is convex in $(1, 2)$, we have $v(\bar{q}, a) - v(\underline{q}, a) \le v(\bar{\bar{q}}, a) - v(\bar{q}, a)$. We aim to show that $L^{\theta} v(\bar{q}, a) - L^{\theta} v(\underline{q}, a) \le L^{\theta} v(\bar{\bar{q}}, a) - L^{\theta} v(\bar{q}, a)$. For $y \in \left\{\underline{x}, \bar{x}, \bar{\bar{x}}\right\}$, let $u_y \in \arg\min_u L^{\theta}_u v(y)$. From Proposition 2, we can choose u_y such that $u_y = \min(y_{d_1}, y_{s_1}) e_{(1,1)} + \min(y_{d_2}, y_{s_2}) e_{(2,2)} + k_t(y) e_{(1,2)}$.

Let us also define $m = \underline{x} - u_{\underline{x}}$. Suppose that $\underline{q}_{d_1} \ge \underline{q}_{s_1} + 1$ or $a \in \mathcal{A} \setminus \{e_{(2,1)}\}$, we can distinguish 3 cases: (a) $k_t(\bar{x}) > 0$, (b) $k_t(\bar{x}) = 0$ and $k_t(\bar{\bar{x}}) > 0$ and (c) $k_t(\bar{x}) = 0$ and $k_t(\bar{\bar{x}}) = 0$:

(a) If $k_t(\bar{x}) > 0$. Then,

$$\begin{aligned} L^{\theta} v(\bar{q}, a) - L^{\theta} v(\underline{q}, a) &= c(\bar{x}) - c(\underline{x}) + \theta\mathbb{E}[v(m, A) - v(m, A)] \\ &= c(\bar{\bar{x}}) - c(\bar{x}) + \theta\mathbb{E}[v(m, A) - v(m, A)] \\ &= L^{\theta} v(\bar{\bar{q}}, a) - L^{\theta} v(\bar{q}, a). \end{aligned}$$

(b) If $k_t(\bar{x}) = 0$ and $k_t(\bar{\bar{x}}) > 0$. Then,

$$\begin{aligned} L^{\theta} v(\bar{q}, a) - L^{\theta} v(\underline{q}, a) &= c(\bar{x}) - c(\underline{x}) + \theta\mathbb{E}[v(m + e_{(1,2)}, A) - v(m, A)] \\ &= c(\bar{x}) - c(\underline{x}) + L^{\theta} v(\bar{q}, a) - L^{\theta}_{u_{\bar{x}} + e_{(1,2)}} v(\bar{q}, a) \\ &= c(\bar{\bar{x}}) - c(\bar{x}) + L^{\theta} v(\bar{q}, a) - L^{\theta}_{u_{\bar{x}} + e_{(1,2)}} v(\bar{q}, a) \\ &\le c(\bar{\bar{x}}) - c(\bar{x}) \\ &= c(\bar{\bar{x}}) - c(\bar{x}) + \theta\mathbb{E}[v(m + e_{(1,2)}, A) - v(m + e_{(1,2)}, A)] \\ &= L^{\theta} v(\bar{\bar{q}}, a) - L^{\theta} v(\bar{q}, a) \end{aligned}$$

where the inequality is given because $k_t(\underline{x}) = k_t(\bar{x}) = 0$ and $1 \in K_{\bar{x}}$.

(c) If $k_t(\bar{x}) = 0$ and $k_t(\bar{\bar{x}}) = 0$. Then,

$$\begin{aligned} L^{\theta} v(\bar{q}, a) - L^{\theta} v(\underline{q}, a) &= c(\bar{x}) - c(\underline{x}) + \theta\mathbb{E}[v(m + e_{(1,2)}, A) - v(m, A)] \\ &= c(\bar{\bar{x}}) - c(\bar{x}) + \theta\mathbb{E}[v(m + e_{(1,2)}, A) - v(m, A)] \\ &\le c(\bar{\bar{x}}) - c(\bar{x}) + \theta\mathbb{E}[v(m + 2e_{(1,2)}, A) - v(m + e_{(1,2)}, A)] \\ &= L^{\theta} v(\bar{\bar{q}}, a) - L^{\theta} v(\bar{q}, a) \end{aligned}$$

where the inequality is true since $v \in \mathcal{C}_{(1,2)}$.

Suppose now that $\underline{q}_{d_1} = \underline{q}_{s_1}$ and $a = e_{(2,1)}$, we can distinguish 2 cases: $k_t(\bar{\bar{x}}) > 0$ and $k_t(\bar{\bar{x}}) = 0$:

- If $k_t(\bar{\bar{x}}) > 0$. Then,

$$\begin{aligned} L^{\theta} v(\bar{q}, a) - L^{\theta} v(\underline{q}, a) &= c(\bar{x}) - c(\underline{x}) + \theta\mathbb{E}[v(m - e_{(2,1)}, A) - v(m, A)] \\ &\le c(\bar{x}) - c(\underline{x}) + \theta\mathbb{E}[v(m, A) - v(m, A)] \\ &= c(\bar{\bar{x}}) - c(\bar{x}) + \theta\mathbb{E}[v(m, A) - v(m, A)] \\ &= c(\bar{\bar{x}}) - c(\bar{x}) + \theta\mathbb{E}[v(m - e_{(2,1)}, A) - v(m - e_{(2,1)}, A)] \\ &= L^{\theta} v(\bar{\bar{q}}, a) - L^{\theta} v(\bar{q}, a) \end{aligned}$$

  where the inequality is given because $v \in \mathcal{I}_{(2,1)}$.

- If $k_t(\bar{\bar{x}}) = 0$. Then,

$$\begin{aligned} L^\theta v(\bar{q}, a) - L^\theta v(\underline{q}, a) &= c(\bar{x}) - c(\underline{x}) + \theta \mathbb{E}[v(m - e_{(2,1)}, A) - v(m, A)] \\ &= c(\bar{\bar{x}}) - c(\bar{x}) + \theta \mathbb{E}[v(m - e_{(2,1)}, A) - v(m, A)] \\ &\le c(\bar{\bar{x}}) - c(\bar{x}) + \theta \mathbb{E}[v(m - e_{(2,1)} + e_{(1,2)}, A) - v(m - e_{(2,1)}, A)] \\ &= L^\theta v(\bar{\bar{q}}, a) - L^\theta v(\bar{q}, a) \end{aligned}$$

where the inequality is given because $v \in \mathcal{B}$.

D.3. Proof of Lemma 3

D.3.1. Preservation of $\mathcal{B}$

Proof. Let $a \in \mathcal{A}$. Since $v \in \mathcal{B}$, we have $v(0, a) - v(e_{(2,1)}, a) \le v(e_{(1,2)}, a) - v(0, a)$. We aim to show that $L^\theta v(0, a) - L^\theta v(e_{(2,1)}, a) \le L^\theta v(e_{(1,2)}, a) - L^\theta v(0, a)$. For any $x \in \mathcal{Q}$, we know from Proposition 2 that there exists $u_x \in \arg\min_u L_u^\theta v(x)$ such that $u_x = \min(x_{d_1}, x_{s_1}) e_{(1,1)} + \min(x_{d_2}, x_{s_2}) e_{(2,2)} + k_t(x) e_{(1,2)}$. We show the desired result for each possible value of a:

- If $a = e_{(1,1)}$ or $a = e_{(2,2)}$. Suppose that $k_t(e_{(1,2)}) = 0$. Then,

$$\begin{aligned} L^\theta v(0, a) - L^\theta v(e_{(2,1)}, a) &= c(a) - c(a + e_{(2,1)}) + \theta \mathbb{E}[v(0, A) - v(e_{(2,1)}, A)] \\ &< c(a + e_{(1,2)}) - c(a) + \theta \mathbb{E}[v(0, A) - v(e_{(2,1)}, A)] \\ &\le c(a + e_{(1,2)}) - c(a) + \theta \mathbb{E}[v(e_{(1,2)}, A) - v(0, A)] \\ &= L^\theta v(e_{(1,2)}, a) - L^\theta v(0, a) \end{aligned}$$

where the second inequality is given since $v \in \mathcal{B}$. Suppose now that $k_t(e_{(1,2)}) > 0$. Then,

$$\begin{aligned} L^\theta v(0, a) - L^\theta v(e_{(2,1)}, a) &= c(a) - c(a + e_{(2,1)}) + \theta \mathbb{E}[v(0, A) - v(e_{(2,1)}, A)] \\ &\le c(a) - c(a + e_{(2,1)}) \\ &< c(a + e_{(1,2)}) - c(a) \\ &= c(a + e_{(1,2)}) - c(a) + \theta \mathbb{E}[v(0, A) - v(0, A)] \\ &= L^\theta v(e_{(1,2)}, a) - L^\theta v(0, a) \end{aligned}$$

where the first inequality is given because $v \in \mathcal{I}_{(2,1)}$.
- If $a = e_{(1,2)}$. Suppose that $k_t(2e_{(1,2)}) = 0$. Then,

$$\begin{aligned} L^\theta v(0, a) - L^\theta v(e_{(2,1)}, a) &= c(a) - c(a + e_{(2,1)}) + \theta \mathbb{E}[v(e_{(1,2)}, A) - v(0, A)] \\ &< c(a + e_{(1,2)}) - c(a) + \theta \mathbb{E}[v(e_{(1,2)}, A) - v(0, A)] \\ &\le c(a + e_{(1,2)}) - c(a) + \theta \mathbb{E}[v(2e_{(1,2)}, A) - v(e_{(1,2)}, A)] \\ &= L^\theta v(e_{(1,2)}, a) - L^\theta v(0, a) \end{aligned}$$

where the second inequality follows since $v \in \mathcal{C}_{(1,2)}$. Suppose now that $k_t(2e_{(1,2)}) > 0$ and $k_t(e_{(1,2)}) = 0$. Then,

$$\begin{aligned} L^\theta v(0, a) - L^\theta v(e_{(2,1)}, a) &= c(a) - c(a + e_{(2,1)}) + \theta \mathbb{E}[v(e_{(1,2)}, A) - v(0, A)] \\ &= c(a) - c(a + e_{(2,1)}) + L^\theta v(0, a) - L^\theta_{u_a + e_{(1,2)}} v(0, a) \\ &\le c(a) - c(a + e_{(2,1)}) \\ &< c(a + e_{(1,2)}) - c(a) \\ &= c(a + e_{(1,2)}) - c(a) + \theta \mathbb{E}[v(e_{(1,2)}, A) - v(e_{(1,2)}, A)] \\ &= L^\theta v(e_{(1,2)}, a) - L^\theta v(0, a), \end{aligned}$$

where the first inequality follows since $1 \in U_{\bar{x}}$ Finally, suppose that $k_t(2e_{(1,2)}) > 0$ and $k_t(e_{(1,2)}) > 0$. Then,

$$\begin{aligned} L^\theta v(0, a) - L^\theta v(e_{(2,1)}, a) &= c(a) - c(a + e_{(2,1)}) + \theta \mathbb{E}[v(0, A) - v(0, A)] \\ &< c(a + e_{(1,2)}) - c(a) + \theta \mathbb{E}[v(0, A) - v(0, A)] \\ &= L^\theta v(e_{(1,2)}, a) - L^\theta v(0, a). \end{aligned}$$

- If $a = e_{(2,1)}$. Then,

$$\begin{aligned} L^{\theta}v(0,a) - L^{\theta}v(e_{(2,1)},a) &= c(a) - c(a+e_{(2,1)}) + \theta\mathbb{E}[v(e_{(2,1)},A) - v(2e_{(2,1)},A)] \\ &\le c(a) - c(a+e_{(2,1)}) + \theta\mathbb{E}[v(0,A) - v(e_{(2,1)},A)] \\ &< c(a+e_{(1,2)}) - c(a) + \theta\mathbb{E}[v(0,A) - v(e_{(2,1)},A)] \\ &= L^{\theta}v(e_{(1,2)},a) - L^{\theta}v(0,a) \end{aligned}$$

where the first inequality follows since $v \in \mathcal{C}_{(2,1)}$. □

D.3.2. Preservation of $\mathcal{C}_{(2,1)}$

Proof. Let $a \in \mathcal{A}$ and $\underline{q} \in \mathcal{Q}$ such that $\underline{q}_{s_1} \ge \underline{q}_{d_1}$, $\underline{x} = \underline{q} + a$. We define $\overline{q} = \underline{q} + e_{(2,1)}$, $\overline{x} = \overline{q} + a$, $\overline{\overline{q}} = \overline{q} + e_{(2,1)}$ and $\overline{\overline{x}} = \overline{\overline{q}} + a$. Since v is convex in (2, 1), we have $v(\overline{q},a) - v(\underline{q},a) \le v(\overline{\overline{q}},a) - v(\overline{q},a)$. We aim to show that $L^{\theta}v(\overline{q},a) - L^{\theta}v(\underline{q},a) \le L^{\theta}v(\overline{\overline{q}},a) - L^{\theta}v(\overline{q},a)$. For $y \in \left\{\underline{x}, \overline{x}, \overline{\overline{x}}\right\}$, let $u_y \in \arg\min_u L_u^{\theta}v(y)$. From Proposition 2, we can choose u_y such that $u_y = \min(y_{d_1}, y_{s_1})e_{(1,1)} + \min(y_{d_2}, y_{s_2})e_{(2,2)} + k_t(y)e_{(1,2)}$. Let us also define $m = \underline{x} - u_{\underline{x}}$. We can distinguish 2 cases: (a) $\underline{q}_{s_1} \ge \underline{q}_{d_1} + 1$ or $a \in \mathcal{A} \setminus \{e_{(1,2)}\}$ and (b) $\underline{q}_{s_1} = \underline{q}_{d_1}$ and $a = e_{(1,2)}$:

(a) If $\underline{q}_{s_1} \ge \underline{q}_{d_1} + 1$ or $a \in \mathcal{A} \setminus \{e_{(1,2)}\}$. Then,

$$\begin{aligned} L^{\theta}v(\overline{q},a) - L^{\theta}v(\underline{q},a) &= c(\overline{x}) - c(\underline{x}) + \theta\mathbb{E}[v(m+e_{(2,1)},A) - v(m,A)] \\ &\le c(\overline{x}) - c(\underline{x}) + \theta\mathbb{E}[v(m+2e_{(2,1)},A) - v(m+e_{(2,1)},A)] \\ &< c(\overline{\overline{x}}) - c(\overline{x}) + \theta\mathbb{E}[v(m+2e_{(2,1)},A) - v(m+e_{(2,1)},A)] \\ &= L^{\theta}v(\overline{\overline{q}},a) - L^{\theta}v(\overline{q},a) \end{aligned}$$

where the first inequality is given because $v \in \mathcal{C}_{(2,1)}$.

(b) If $\underline{q}_{s_1} = \underline{q}_{d_1}$ and $a = e_{(1,2)}$. Suppose that $k_t(\underline{x}) = 0$. Then,

$$\begin{aligned} L^{\theta}v(\overline{q},a) - L^{\theta}v(\underline{q},a) &= c(\overline{x}) - c(\underline{x}) + \theta\mathbb{E}[v(m-e_{(1,2)},A) - v(m,A)] \\ &\le c(\overline{\overline{x}}) - c(\overline{x}) + \theta\mathbb{E}[v(m-e_{(1,2)}+e_{(2,1)},A) - v(m-e_{(1,2)},A)] \\ &= L^{\theta}v(\overline{\overline{q}},a) - L^{\theta}v(\overline{q},a) \end{aligned}$$

because c is a linear function and $v \in \mathcal{B}$. Suppose now that $k_t(\underline{x}) > 0$. Then,

$$\begin{aligned} L^{\theta}v(\overline{q},a) - L^{\theta}v(\underline{q},a) &= c(\overline{x}) - c(\underline{x}) + \theta\mathbb{E}[v(m+e_{(2,1)},A) - v(m,A)] \\ &\le c(\overline{x}) - c(\underline{x}) + \theta\mathbb{E}[v(m+2e_{(2,1)},A) - v(m+e_{(2,1)},A)] \\ &< c(\overline{\overline{x}}) - c(\overline{x}) + \theta\mathbb{E}[v(m+2e_{(2,1)},A) - v(m+e_{(2,1)},A)] \\ &= L^{\theta}v(\overline{\overline{q}},a) - L^{\theta}v(\overline{q},a) \end{aligned}$$

where the first equality follows since $v \in \mathcal{C}_{(2,1)}$. □

Appendix E. Proof of Proposition 3

Proof. Let $u_t^{\infty} \in \Pi^{T_{(1,2)}}$. Let us look at the Markov chain derived from this policy. The set of possible states (except for Y_0) is $\mathcal{S}^{\mathcal{A}} = \{(s_i, a) : i \in \mathbb{N}, a \in \mathcal{A}\}$ with $s_i = (t-i, 0, 0, t-i)$ if $i \le t$ and $s_i = (0, i-t, i-t, 0)$ otherwise. The s_i are all possible states after matching the items using u_t^{∞}. In order to see more clearly the behavior of the Markov chain, we group together some states. Let us define $\mathcal{S} = \{S_i : i \in \mathbb{N}\}$ with $S_0 = \{(s_0, e_{(1,2)}), (s_0, e_{(1,1)}), (s_0, e_{(2,2)}), (s_1, e_{(1,2)})\}$ and $S_i = \{(s_{i-1}, e_{(2,1)}), (s_i, e_{(1,1)}), (s_i, e_{(2,2)}), (s_{i+1}, e_{(1,2)})\}$ for all $i \in \mathbb{N}^*$. Fig. 6 shows that this Markov chain defined on $\mathcal{S}$ is clearly irreducible.

The balance equations are the following:

$$\beta(1-\alpha)\pi_{S_i} = \alpha(1-\beta)\pi_{S_{i+1}} \qquad i = 0, 1, \ldots$$

Solving these equations under the constraint that $\sum_{i=0}^{\infty} \pi_{S_i} = 1$ give:

$$\pi_{S_i} = \rho^i(1-\rho) \qquad i = 0, 1, \ldots \tag{13}$$

with $\rho = \frac{\beta(1-\alpha)}{\alpha(1-\beta)} \in (0, 1)$. So (13) is the stationary distribution of the Markov chain on $\mathcal{S}$, which thus is positive recurrent. Using these results, we can easily conclude that the Markov chain on $\mathcal{S}^{\mathcal{A}}$ is also irreducible. Then, since the arrival process

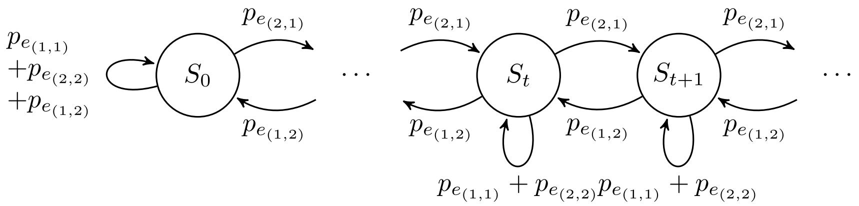

Fig. 6. The graph associated to the Markov chain derived from u_t^∞ and defined on the state space $\mathcal{S}$. $p_{e_{(1,1)}} = \alpha\beta$, $p_{e_{(2,2)}} = (1-\alpha)(1-\beta)$, $p_{e_{(1,2)}} = \alpha(1-\beta)$ and $p_{e_{(2,1)}} = \beta(1-\alpha)$.

does not depend on the state and because the following condition must be satisfied $\pi_{S_i} = \pi_{(s_{i-1},e_{(2,1)})} + \pi_{(s_i,e_{(1,1)})} + \pi_{(s_i,e_{(2,2)})} + \pi_{(s_{i+1},e_{(1,2)})}$, we can think of the following as the stationary distribution:

$$\pi_{(s_i,a)} = \rho^i(1-\rho)p_a \tag{14}$$

with p_a as defined in Fig. 6 (i.e. $p_{e_{(1,1)}} = \alpha\beta$, $p_{e_{(2,2)}} = (1-\alpha)(1-\beta)$, $p_{e_{(1,2)}} = \alpha(1-\beta)$ and $p_{e_{(2,1)}} = \beta(1-\alpha)$), for all $i \in \mathbb{N}$ and $a \in \mathcal{A}$. Let us verify that (14) is indeed a stationary distribution:

$$\begin{aligned}\sum_{k\in\mathbb{N}}\sum_{a\in\mathcal{A}} \pi_{(s_k,a)} p((s_k,a),(s_i,a')) &= p_{a'}(\pi_{(s_{i-1},e_{(2,1)})} + \pi_{(s_i,e_{(1,1)})} + \pi_{(s_i,e_{(2,2)})} + \pi_{(s_{i+1},e_{(1,2)})}) \\ &= p_{a'}(\rho^{i-1}(1-\rho)p_{e_{(2,1)}} + \rho^i(1-\rho)p_{e_{(1,1)}} + \rho^i(1-\rho)p_{e_{(2,2)}} + \rho^{i+1}(1-\rho)p_{e_{(1,2)}}) \\ &= p_{a'}\rho^i(1-\rho)(\frac{\beta(1-\alpha)}{\rho} + \alpha\beta + (1-\alpha)(1-\beta) + \rho\alpha(1-\beta)) \\ &= p_{a'}\rho^i(1-\rho) \\ &= \pi_{(s_i,a')}\end{aligned}$$

Hence, $\sum_{i\in\mathbb{N}}\sum_{a\in\mathcal{A}} \pi_{(s_i,a)} = \sum_{i\in\mathbb{N}} \rho^i(1-\rho)\sum_{a\in\mathcal{A}} p_a = \sum_{i\in\mathbb{N}} \rho^i(1-\rho) = 1$ and, therefore, (14) is the stationary distribution of the Markov chain derived from the policy u_t^∞ and this Markov chain is positive recurrent. Using the monotone convergence theorem and the strong law of large number for Markov chains, we can compute the average cost $g^{u_t^\infty}$:

$$g^{u_t^\infty}(y) = \lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} \mathbb{E}_y^{u_t^\infty}[c(Y(n))] = \mathbb{E}_\pi[c(Y)]$$

where $\mathbb{E}_\pi$ means the expectation over the stationary distribution π defined as (14). Finally, we compute this value:

$$\begin{aligned}\mathbb{E}_\pi[c(Y)] &= \sum_{i=1}^{t}\sum_{a\in\mathcal{A}} c(s_{t-i}+a)\pi_{(s_{t-i},a)} + \sum_{i\in\mathbb{N}}\sum_{a\in\mathcal{A}} c(s_{t+i}+a)\pi_{(s_{t+i},a)} \\ &= \sum_{i=1}^{t}((c_{d_1}+c_{s_2})i + c_{d_1} + c_{s_1})\rho^{t-i}(1-\rho)\alpha\beta + ((c_{d_1}+c_{s_2})i + c_{d_2} + c_{s_1})\rho^{t-i}(1-\rho)(1-\alpha)\beta \\ &\quad + ((c_{d_1}+c_{s_2})i + c_{d_1} + c_{s_2})\rho^{t-i}(1-\rho)\alpha(1-\beta) + ((c_{d_1}+c_{s_2})i + c_{d_2} + c_{s_2})\rho^{t-i}(1-\rho)(1-\alpha)(1-\beta) \\ &\quad + \sum_{i\in\mathbb{N}}((c_{d_2}+c_{s_1})i + c_{d_1} + c_{s_1})\rho^{t+i}(1-\rho)\alpha\beta + ((c_{d_2}+c_{s_1})i + c_{d_2} + c_{s_1})\rho^{t+i}(1-\rho)(1-\alpha)\beta \\ &\quad + ((c_{d_2}+c_{s_1})i + c_{d_1} + c_{s_2})\rho^{t+i}(1-\rho)\alpha(1-\beta) + ((c_{d_2}+c_{s_1})i + c_{d_2} + c_{s_2})\rho^{t+i}(1-\rho)(1-\alpha)(1-\beta) \\ &= \sum_{i=1}^{t}(c_{d_1}+c_{s_2})i\rho^{t-i}(1-\rho) + (c_{d_1}+c_{s_1})\rho^{t-i}(1-\rho)\alpha\beta + (c_{d_2}+c_{s_1})\rho^{t-i}(1-\rho)(1-\alpha)\beta \\ &\quad + (c_{d_1}+c_{s_2})\rho^{t-i}(1-\rho)\alpha(1-\beta) + (c_{d_2}+c_{s_2})\rho^{t-i}(1-\rho)(1-\alpha)(1-\beta) \\ &\quad + \sum_{i\in\mathbb{N}}(c_{d_2}+c_{s_1})i\rho^{t+i}(1-\rho) + (c_{d_2}+c_{s_1})\rho^{t+i}(1-\rho)\alpha\beta + (c_{d_2}+c_{s_1})\rho^{t+i}(1-\rho)(1-\alpha)\beta \\ &\quad + (c_{d_2}+c_{s_1})\rho^{t+i}(1-\rho)\alpha(1-\beta) + (c_{d_2}+c_{s_2})\rho^{t+i}(1-\rho)(1-\alpha)\end{aligned} \tag{15}$$

Using basic algebra, one can show that, for any c, the following properties hold:

$$\sum_{i=1}^{t} c\cdot i\cdot\rho^{t-i} = c\left(t - \rho\frac{1-\rho^t}{1-\rho}\right) \text{ and } \sum_{i\in\mathbb{N}} c\cdot i\cdot\rho^{t-i} = c\frac{\rho^{t+1}}{1-\rho}.$$

Moreover, for any c and q, we have that

$$q\cdot c\cdot(1-\rho)\sum_{i=1}^{t}\rho^{t-i}=q\cdot c\cdot(1-\rho^t)\text{ and }q\cdot c\cdot(1-\rho)\sum_{i\in\mathbb{N}}\rho^{t+i}=q\cdot c\cdot\rho^t.$$

Using these properties in (15), we obtain that

$$\begin{aligned}\mathbb{E}_\pi[c(Y)]&=(c_{d_1}+c_{s_2})\left(t-\rho\frac{1-\rho^t}{1-\rho}\right)+(1-\rho^t)((c_{d_1}+c_{s_1})\alpha\beta+(c_{d_1}+c_{s_2})\alpha(1-\beta)\\&\quad+(c_{d_2}+c_{s_1})\alpha(1-\beta)+(c_{d_2}+c_{s_2})(1-\alpha)(1-\beta))+(c_{d_2}+c_{s_1})\frac{\rho^{t+1}}{1-\rho}+\rho^t((c_{d_1}+c_{s_1})\alpha\beta\\&\quad+(c_{d_1}+c_{s_2})\alpha(1-\beta)+(c_{d_2}+c_{s_1})\alpha(1-\beta)+(c_{d_2}+c_{s_2})(1-\alpha)(1-\beta))\\&=(c_{d_1}+c_{s_2})t+(c_{d_1}+c_{d_2}+c_{s_1}+c_{s_2})\frac{\rho^{t+1}}{1-\rho}-(c_{d_1}+c_{d_2})\frac{\rho}{1-\rho}+((c_{d_1}+c_{s_1})\alpha\beta\\&\quad+(c_{d_1}+c_{s_2})\alpha(1-\beta)+(c_{d_2}+c_{s_1})\alpha(1-\beta)+(c_{d_2}+c_{s_2})(1-\alpha)(1-\beta)).\end{aligned}\tag{16}$$

We aim to obtain the value of t that minimizes (16). To this end, we compute the second derivative of (16) with respect to t and it results in

$$(c_{d_1}+c_{d_2}+c_{s_1}+c_{s_2})\frac{\rho^{t+1}}{1-\rho}(\log\rho)^2,$$

which is positive for $\rho\in(0,1)$, i.e., (16) is convex. As a consequence, the minimum of (16) is achieved when its derivative with respect to t is equal to zero:

$$c_{d_1}+c_{s_2}+(c_{d_1}+c_{d_2}+c_{s_1}+c_{s_2})\frac{\rho^{t+1}}{1-\rho}(\log\rho)=0\iff 1+(1+R)\frac{\rho^{t+1}}{1-\rho}(\log\rho)=0,$$

where $R=\frac{c_{s_1}+c_{d_2}}{c_{d_1}+c_{s_2}}$. The root of the previous equation is

$$t_0=\frac{1}{\log\rho}\log\left(\frac{\rho-1}{(R+1)\log\rho}\right)-1.$$

Since t_0 is not necessarily integer, the optimal threshold t^* is obtained by computing the value of (16) for the ceil of t_0 and the floor of t_0 and choosing the minimum of these values. □

We now show that t^* is always positive.

Lemma 10. *t^* is always positive.*

Proof. From Taylor's Theorem, it follows that for all $\rho\in(0,1)$

$$\frac{1-\rho}{(1+R)log(\rho)}\in\left(0,\frac{1}{R+1}\right).$$

Since $R>0$, we have that for all $\rho\in(0,1)$

$$\log\left(\frac{1-\rho}{(1+R)log(\rho)}\right)<0.$$

Hence,

$$\frac{\log\left(\frac{1-\rho}{(1+R)}\right)}{\log(\rho)}>0.$$

As a result, t^* is positive since t_0 is positive.

Appendix F. Proof of Lemma 4

Let $0\le\theta<1$, let $\pi^*=(u(X(n)))_{n\ge0}$ (with $u\in D^*$ as defined in Definition 7) be a threshold-type policy with priority in i^* and j^* and let $(\pi^N)^*=(u^N(X^N(n)))_{n\ge0}$ (with $u^N\in D^\sigma$ as defined in Definition 4) be a threshold-type policy with priority in (1, 1) and (2, 2). Let $y_0=(q_0,a_0)\in\mathcal{Q}\times\mathcal{A}$, with $y_0^N=(p_{\mathcal{Q}}^N(q_0),p_{\mathcal{A}}^N(a_0))$.

First, let us note that $(\pi^N)^*$ is the "projection" of π^* on $\mathcal{G}^N$ in the following sense. Let $x\in\mathcal{Q}$, $u(x)\in D^*$ the matching of state x ($u(x)\in U_x$) under the policy π^* and $u^N(p_{\mathcal{Q}}^N(x))\in D^\sigma$ the matching of state $p_{\mathcal{Q}}^N(x)$ ($u^N(p_{\mathcal{Q}}^N(x))\in U^N_{p_{\mathcal{Q}}^N(x)}$) under the policy $(\pi^N)^*$. We can easily see that $u^N_{(1,1)}(p_{\mathcal{Q}}^N(x))=\sum_{i\in\mathcal{D}(j^*)}u_{(i,j^*)}(x)$, $u^N_{(2,2)}(p_{\mathcal{Q}}^N(x))=\sum_{j\in\mathcal{S}(i^*)}u_{(i^*,j)}(x)$ and $u^N_{(1,2)}(p_{\mathcal{Q}}^N(x))=\sum_{i\in\mathcal{D}(j^*)}\sum_{j\in\mathcal{S}(i^*)}u_{(i,j)}(x)$.

Then, we are going to prove by induction that for any $a_1^N \in \mathcal{A}^N, \dots, a_n^N \in \mathcal{A}^N$ and any $a_1 \in (p_{\mathcal{A}}^N)^{-1}(a_1^N), \dots, a_n \in (p_{\mathcal{A}}^N)^{-1}(a_n^N)$, we have $p_{\mathcal{Q}}^N(x_n) = x_n^N$.

We already specifically chose x_0^N to verify this property: $p_{\mathcal{Q}}^N(x_0) = p_{\mathcal{Q}}^N(q_0) + p_{\mathcal{A}}^N(a_0) = q_0^N + a_0^N = x_0^N$. Now, assume that $p_{\mathcal{Q}}^N(x_{n-1}) = x_{n-1}^N$. Then,

$$\begin{aligned}
p_{\mathcal{Q}}^N(q_n) &= \left(\sum_{i\in\mathcal{D}(j^*)} (q_n)_{d_i}, (q_n)_{d_{i^*}}, (q_n)_{s_{j^*}}, \sum_{j\in\mathcal{S}(i^*)} (q_n)_{s_j} \right) \\
&= \left(\sum_{i\in\mathcal{D}(j^*)} (x_{n-1} - u(x_{n-1}))_{d_i}, (x_{n-1} - u(x_{n-1}))_{d_{i^*}}, \right. \\
&\quad \left. (x_{n-1} - u(x_{n-1}))_{s_{j^*}}, \sum_{j\in\mathcal{S}(i^*)} (x_{n-1} - u(x_{n-1}))_{s_j} \right) \\
&= \left(\sum_{i\in\mathcal{D}(j^*)} (x_{n-1})_{d_i} - \sum_{j\in\mathcal{S}(i)} u(x_{n-1})_{(i,j)}, \right. \\
&\quad (x_{n-1})_{d_{i^*}} - \sum_{j\in\mathcal{S}(i^*)} u(x_{n-1})_{(i^*,j)}, (x_{n-1})_{s_{j^*}} - \sum_{i\in\mathcal{D}(j^*)} u(x_{n-1})_{(i,j^*)}, \\
&\quad \left. \sum_{j\in\mathcal{S}(i^*)} (x_{n-1})_{s_j} - \sum_{i\in\mathcal{D}(j)} u(x_{n-1})_{(i,j)} \right) \\
&= p_{\mathcal{Q}}^N(x_{n-1}) - \sum_{i\in\mathcal{D}(j^*)} u(x_{n-1})_{(i,j^*)} e_{(1,1)} - \sum_{j\in\mathcal{S}(i^*)} u(x_{n-1})_{(i^*,j)} e_{(2,2)} \\
&\quad - \sum_{i\in\mathcal{D}(j^*)} \sum_{j\in\mathcal{S}(i^*)} u(x_{n-1})_{(i,j)} e_{(1,2)} \\
&= p_{\mathcal{Q}}^N(x_{n-1}) - u^N(p_{\mathcal{Q}}^N(x_{n-1}))_{(1,1)} e_{(1,1)} - u^N(p_{\mathcal{Q}}^N(x_{n-1}))_{(2,2)} e_{(2,2)} \\
&\quad - u^N(p_{\mathcal{Q}}^N(x_{n-1}))_{(1,2)} e_{(1,2)} \\
&= p_{\mathcal{Q}}^N(x_{n-1}) - u^N(p_{\mathcal{Q}}^N(x_{n-1})) \\
&= x_{n-1}^N - u^N(x_{n-1}^N) \\
&= q_n^N
\end{aligned}$$

and $p_{\mathcal{A}}^N(a_n) = a_n^N$ (because $p_{\mathcal{A}}^N((p_{\mathcal{A}}^N)^{-1}(a^N)) = a^N$ for all $a^N \in \mathcal{A}^N$). Thus, $p_{\mathcal{Q}}^N(x_n) = p_{\mathcal{Q}}^N(q_n) + p_{\mathcal{A}}^N(a_n) = q_n^N + a_n^N = x_n^N$. Finally, using this property and (5), we have

$$\begin{aligned}
\mathbb{E}_{y_0}^{\pi^*}[c(Y(n))] &= \sum_{a_1\in\mathcal{A},\dots,a_n\in\mathcal{A}} c(x_n) \prod_{k=1}^{n} \mathbb{P}(A = a_k) \\
&= \sum_{a_1^N\in\mathcal{A}^N,\dots,a_n^N\in\mathcal{A}^N} \sum_{a_1\in(p_{\mathcal{A}}^N)^{-1}(a_1^N)} \cdots \sum_{a_n\in(p_{\mathcal{A}}^N)^{-1}(a_n^N)} c(x_n) \prod_{k=1}^{n} \mathbb{P}(A = a_k) \\
&= \sum_{a_1^N\in\mathcal{A}^N,\dots,a_n^N\in\mathcal{A}^N} \sum_{a_1\in(p_{\mathcal{A}}^N)^{-1}(a_1^N)} \cdots \sum_{a_n\in(p_{\mathcal{A}}^N)^{-1}(a_n^N)} c^N(p_{\mathcal{Q}}^N(x_n)) \prod_{k=1}^{n} \mathbb{P}(A = a_k) \\
&= \sum_{a_1^N\in\mathcal{A}^N,\dots,a_n^N\in\mathcal{A}^N} \sum_{a_1\in(p_{\mathcal{A}}^N)^{-1}(a_1^N)} \cdots \sum_{a_n\in(p_{\mathcal{A}}^N)^{-1}(a_n^N)} c^N(x_n^N) \prod_{k=1}^{n} \mathbb{P}(A = a_k) \\
&= \sum_{a_1^N\in\mathcal{A}^N,\dots,a_n^N\in\mathcal{A}^N} c^N(x_n^N) \prod_{k=1}^{n} \sum_{a_k\in(p_{\mathcal{A}}^N)^{-1}(a_k^N)} \mathbb{P}(A = a_k)
\end{aligned}$$

$$= \sum_{a_1^N \in \mathcal{A}^N, \ldots, a_n^N \in \mathcal{A}^N} c^N(x_n^N) \prod_{k=1}^{n} \mathbb{P}(A^N = a_k^N)$$
$$= \mathbb{E}_{y_0^N}^{(\pi^N)^*}[c^N(Y^N(n))].$$

This equality is true for any $n \in \mathbb{N}$. Therefore, $v_\theta^{\pi^*}(y_0) = v_\theta^{(\pi^N)^*}(y_0^N)$ and $g^{\pi^*}(y_0) = g^{(\pi^N)^*}(y_0^N)$. This was done for any $y_0 = (q_0, a_0) \in \mathcal{Q} \times \mathcal{A}$ with $y_0^N = (p_{\mathcal{Q}}^N(q_0), p_{\mathcal{A}}^N(a_0))$, giving the final result.

Appendix G. Proof of Lemma 5

Proof. Let $0 \le \theta < 1$, let π be a stationary policy on Y, $y_0 = (q_0, a_0) \in \mathcal{Q} \times \mathcal{A}$ and $y_0^N = (p_{\mathcal{Q}}^N(q_0), p_{\mathcal{A}}^N(a_0))$. We start by constructing a history dependent policy $\pi^N = (u_n^N)_{n\ge 0}$ on Y^N that will make Y^N "follow" (in some sense) the projection of Y on $\mathcal{G}^N$. First, let us introduce new independent random variables $\hat{A}(n)$ that we sample just after the arrivals on $\mathcal{G}^N$. $\hat{A}(n)$ is defined on $\mathcal{A}$ with the following distribution: $\forall n \in \mathbb{N}^*$, $\forall a \in \mathcal{A}$, $\forall a^N \in \mathcal{A}^N$,

$$\mathbb{P}(\hat{A}(n) = a | A^N(n) = a^N) = \begin{cases} \frac{\mathbb{P}(A(n)=a)}{\mathbb{P}(A^N(n)=a^N)} & \text{if } a \in (p_{\mathcal{A}}^N)^{-1}(a^N) \\ 0 & \text{Otherwise} \end{cases}.$$

Then, we define u_n^N the decision rule of π^N at time n based on the history of the trajectory, i.e.

$$(A^N(1), \hat{A}(1), \ldots, A^N(n), \hat{A}(n)) = (a_1^N, \hat{a}_1^N, \ldots, a_n^N, \hat{a}_n^N),$$

the initial state $y_0 = (q_0, a_0)$ and the stationary policy π. Let $\hat{x}_n \in \mathcal{Q}$ be the state we end up by starting in $x_0 = q_0 + a_0$ and following the dynamics (2) with the sequence of arrivals $\hat{a}_1^N, \ldots, \hat{a}_n^N$ and under the policy π. Let $u \in U_{\hat{x}_n}$ be the decision rule applied for the state $\hat{x}_n$ under the policy π. We construct $u_n^N \in U_{p_{\mathcal{Q}}^N(\hat{x}_n)}$ such that $(u_n^N)_{(1,1)} = \sum_{i \in \mathcal{D}(j^*)} u_{(i,j^*)}$, $(u_n^N)_{(2,2)} = \sum_{j\in\mathcal{S}(i^*)} u_{(i^*,j)}$ and $(u_n^N)_{(1,2)} = \sum_{i\in\mathcal{D}(j^*)} \sum_{j\in\mathcal{S}(i^*)} u_{(i,j)}$.

Now, let us show by induction that, under π^N, we have $p_{\mathcal{Q}}^N(x_n) = x_n^N$ for any $a_1^N \in \mathcal{A}^N, \ldots, a_n^N \in \mathcal{A}^N$ and any $\hat{a}_1 \in (p_{\mathcal{A}}^N)^{-1}(a_1^N), \ldots, \hat{a}_n \in (p_{\mathcal{A}}^N)^{-1}(a_n^N)$ such that $\hat{a}_1 = a_1, \ldots, \hat{a}_n = a_n$.

We already specifically chose y_0^N to verify this property: $p_{\mathcal{Q}}^N(x_0) = p_{\mathcal{Q}}^N(q_0) + p_{\mathcal{A}}^N(a_0) = q_0^N + a_0^N = x_0^N$. Now, assume that $p_{\mathcal{Q}}^N(x_{n-1}) = x_{n-1}^N$. First, let us note that $x_{n-1} = \hat{x}_{n-1}$ because $\hat{a}_1 = a_1, \ldots, \hat{a}_{n-1} = a_{n-1}$ and they both follow the same dynamics under the same policy π. Then,

$$p_{\mathcal{Q}}^N(q_n) = \left(\sum_{i\in\mathcal{D}(j^*)} (q_n)_{d_i}, (q_n)_{d_{i^*}}, (q_n)_{s_{j^*}}, \sum_{j\in\mathcal{S}(i^*)} (q_n)_{s_j} \right)$$
$$= \left(\sum_{i\in\mathcal{D}(j^*)} (x_{n-1} - u(x_{n-1}))_{d_i}, (x_{n-1} - u(x_{n-1}))_{d_{i^*}} \right.$$
$$\left., (x_{n-1} - u(x_{n-1}))_{s_{j^*}}, \sum_{j\in\mathcal{S}(i^*)} (x_{n-1} - u(x_{n-1}))_{s_j} \right)$$
$$= \left(\sum_{i\in\mathcal{D}(j^*)} (x_{n-1})_{d_i} - \sum_{j\in\mathcal{S}(i)} u(x_{n-1})_{(i,j)}, \right.$$
$$(x_{n-1})_{d_{i^*}} - \sum_{j\in\mathcal{S}(i^*)} u(x_{n-1})_{(i^*,j)}, (x_{n-1})_{s_{j^*}} - \sum_{i\in\mathcal{D}(j^*)} u(x_{n-1})_{(i,j^*)},$$
$$\left. \sum_{j\in\mathcal{S}(i^*)} (x_{n-1})_{s_j} - \sum_{i\in\mathcal{D}(j)} u(x_{n-1})_{(i,j)} \right)$$
$$= p_{\mathcal{Q}}^N(x_{n-1}) - \sum_{i\in\mathcal{D}(j^*)} u(x_{n-1})_{(i,j^*)} e_{(1,1)} - \sum_{j\in\mathcal{S}(i^*)} u(x_{n-1})_{(i^*,j)} e_{(2,2)}$$
$$- \sum_{i\in\mathcal{D}(j^*)} \sum_{j\in\mathcal{S}(i^*)} u(x_{n-1})_{(i,j)} e_{(1,2)}$$

$$
\begin{aligned}
&= p_{\mathcal{Q}}^N(x_{n-1}) - \sum_{i\in\mathcal{D}(j^*)} u(\hat{x}_{n-1})_{(i,j^*)} e_{(1,1)} - \sum_{j\in\mathcal{S}(i^*)} u(\hat{x}_{n-1})_{(i^*,j)} e_{(2,2)} \\
&\quad - \sum_{i\in\mathcal{D}(j^*)} \sum_{j\in\mathcal{S}(i^*)} u(\hat{x}_{n-1})_{(i,j)} e_{(1,2)} \\
&= p_{\mathcal{Q}}^N(x_{n-1}) - (u_n^N)_{(1,1)} e_{(1,1)} - (u_n^N)_{(2,2)} e_{(2,2)} - (u_n^N)_{(1,2)} e_{(1,2)} \\
&= x_{n-1}^N - u_n^N \\
&= q_n^N
\end{aligned}
$$

and $p_{\mathcal{A}}^N(a_n) = a_n^N$ (because $p_{\mathcal{A}}^N((p_{\mathcal{A}}^N)^{-1}(a^N)) = a^N$ for all $a^N \in \mathcal{A}^N$). Thus, $p_{\mathcal{Q}}^N(x_n) = p_{\mathcal{Q}}^N(q_n) + p_{\mathcal{A}}^N(a_n) = q_n^N + a_n^N = x_n^N$.
Then, using this property and (5), we have

$$
\begin{aligned}
\mathbb{E}_{y_0}^{\pi}[c(Y(n))] &= \sum_{a_1\in\mathcal{A},\dots,a_n\in\mathcal{A}} c(x_n) \prod_{k=1}^{n} \mathbb{P}(A(k) = a_k) \\
&= \sum_{a_1^N\in\mathcal{A}^N,\dots,a_n^N\in\mathcal{A}^N} \sum_{a_1\in(p_{\mathcal{A}}^N)^{-1}(a_1^N)} \cdots \sum_{a_n\in(p_{\mathcal{A}}^N)^{-1}(a_n^N)} c(x_n) \prod_{k=1}^{n} \mathbb{P}(A(k) = a_k) \\
&= \sum_{a_1^N\in\mathcal{A}^N,\dots,a_n^N\in\mathcal{A}^N} \sum_{a_1\in(p_{\mathcal{A}}^N)^{-1}(a_1^N)} \cdots \sum_{a_n\in(p_{\mathcal{A}}^N)^{-1}(a_n^N)} c^N(p_{\mathcal{Q}}^N(x_n)) \prod_{k=1}^{n} \mathbb{P}(A(k) = a_k) \\
&= \sum_{a_1^N\in\mathcal{A}^N,\dots,a_n^N\in\mathcal{A}^N} \sum_{\hat{a}_1\in(p_{\mathcal{A}}^N)^{-1}(a_1^N)} \cdots \sum_{\hat{a}_n\in(p_{\mathcal{A}}^N)^{-1}(a_n^N)} c^N(x_n^N) \prod_{k=1}^{n} \mathbb{P}(A(k) = \hat{a}_k) \\
&= \sum_{a_1^N\in\mathcal{A}^N,\dots,a_n^N\in\mathcal{A}^N} \sum_{\hat{a}_1\in(p_{\mathcal{A}}^N)^{-1}(a_1^N)} \cdots \sum_{\hat{a}_n\in(p_{\mathcal{A}}^N)^{-1}(a_n^N)} c^N(x_n^N) \prod_{k=1}^{n} \mathbb{P}(\hat{A}(k) = \hat{a}_k | A^N(k) = a_k^N) \mathbb{P}(A^N(k) = a_k^N) \\
&= \sum_{a_1^N\in\mathcal{A}^N,\dots,a_n^N\in\mathcal{A}^N} \sum_{\hat{a}_1\in(p_{\mathcal{A}}^N)^{-1}(a_1^N)} \cdots \sum_{\hat{a}_n\in(p_{\mathcal{A}}^N)^{-1}(a_n^N)} c^N(x_n^N) \prod_{k=1}^{n} \mathbb{P}(A^N(k) = a_k^N, \hat{A}(k) = \hat{a}_k) \\
&= \mathbb{E}_{y_0^N}^{\pi^N}[c^N(Y^N(n))].
\end{aligned}
$$

This equality is true for any $n \in \mathbb{N}$. Therefore, $v_\theta^{\pi}(y_0) = v_\theta^{\pi^N}(y_0^N)$ and $g^{\pi}(y_0) = g^{\pi^N}(y_0^N)$. □

Appendix H. Proof of Proposition 4

Let $q \in \mathcal{Q}$, $a \in \mathcal{A}$, $x = q + a$ and $u \in U_x$. We define $x_{(i,j)} = \min(x_{d_i}, x_{s_j})$ for all $(i,j) \in \mathcal{E}$ to ease the notations. We assume that the matching graph has m pendant edges, i.e., $\mathcal{E}^* = \{(i_1,j_1),(i_2,j_2),\dots,(i_m,j_m)\}$. For $k = 1,\dots,m$, we define

$$
p_k = \min(x_{(i_k,j_k)} - u_{(i_k,j_k)}, \sum_{(i,j)\in N((i_k,j_k))} u_{(i,j)}).
$$

We now observe that, for all $(i,j) \in N((i_k,j_k))$, we define $0 \le p_{i,j} \le u_{(i,j)}$ such that $p_k = \sum_{(i,j)\in N((i_k,j_k))} p_{i,j}$. Hence, we define

$$
u' = u + \sum_{k=1}^{m} \left(p_k e_{(i_k,j_k)} - \sum_{(i,j)\in N((i_k,j_k))} p_{i,j} e_{(i,j)} \right).
$$

We assume that $u' \in U_x$. Since $v \in V^\sigma$, it follows that

$$
L_{u'}^{\theta} v(q,a) \le L_u^{\theta} v(q,a). \tag{17}
$$

We now define $m_k = x_{(i_k,j_k)} - u'_{(i_k,j_k)}$ for all $k = 1,\dots,m$ and $u^* = u' + \sum_{k=1}^{m} m_k e_{(i_k,j_k)}$. We assume that $u^* \in U_x$. Using that $v \in V^\sigma$, we have that $L_{u^*}^{\theta} v(q,a) \le L_{u'}^{\theta} v(q,a)$, and, taking into account (17), it follows that $L_{u^*}^{\theta} v(q,a) \le L_u^{\theta} v(q,a)$, which holds for any $u \in U_x$ and therefore also for $u \in \arg\min L^\theta v(q,a)$. As a result, $u^* \in \arg\min L^\theta v(q,a)$. Besides, for any pendant edge (i_k,j_k), we have that

$$
u^*_{(i_k,j_k)} = u'_{(i_k,j_k)} + x_{(i_k,j_k)} - u'_{(i_k,j_k)} = x_{(i_k,j_k)},
$$

and the desired result follows if we show that $u' \in U_x$ and $u^* \in U_x$.

- We aim to show that $u' \in U_x$. To prove this, we only need to show that, for any $u \in U_x$ and any $k \in \{1, \dots, m\}$, $u_0 \in U_x$ where

$$u_0 = u + p_k e_{(i_k,j_k)} - \sum_{(i,j)\in N((i_k,j_k))} p_{i,j} e_{(i,j)}.$$

We start by showing that $u_0 \in \mathbb{N}^{n_\mathcal{D}+n_\mathcal{S}}$. First, $(u_0)_{(i_k,j_k)} \geq 0$ because $u \in U_x$ and $p_k \geq 0$. Then, for all $(i,j) \in N((i_k,j_k))$, we have $u_{(i,j)} \geq p_{i,j}$ by definition. Thus, it follows that $(u_0)_{(i,j)} \geq 0$.
Now, we show *(a)* and *(b)* assuming that d_{i_k} is of degree one (the proof for the case that s_{j_k} is of degree one is symmetric and therefore we omitted it). We first show *(a)* for d_{i_k} as follows:

$$(u_0)_{(i_k,j_k)} = u_{(i_k,j_k)} + p_k \leq x_{(i_k,j_k)} \leq x_{d_{i_k}}$$

where the first inequality holds by definition of p_k and the second by definition of $x_{(i_k,j_k)}$. We also show *(a)* for all $i \in \mathcal{D}(j_k)$ as follows:

$$\begin{aligned}\sum_{r\in\mathcal{S}(i)} (u_0)_{(i,r)} &= (u_0)_{(i,j_k)} + \sum_{r\in\mathcal{S}(i)\setminus\{j_k\}} (u_0)_{(i,r)} \\ &= u_{(i,j_k)} - p_{i,j_k} + \sum_{r\in\mathcal{S}(i)\setminus\{j_k\}} u_{(i,r)} \\ &\leq \sum_{r\in\mathcal{S}(i)} u_{(i,r)} \\ &\leq x_{d_i},\end{aligned}$$

where the first inequality holds since $p_{i,j_k} \geq 0$ and the second since $u \in U_x$.
We now show *(b)* for j_k as follows:

$$\sum_{r\in\mathcal{D}(j_k)} (u_0)_{(r,j_k)} = \sum_{r\in\mathcal{D}(j_k)} u_{(r,j_k)} \leq x_{s_{j_k}},$$

where the inequality holds since $u \in U_x$.

- We now aim to show that $u^* \in U_x$ and we observe that it is enough to show that, for any $k \in \{1, \dots, m\}$, $u'_0 = u' + m_k e_{(i_k,j_k)} \in U_x$.
We first observe that, if $m_k = 0$, then $u_0 = u'$ and therefore $u'_0 \in U_x$. Therefore, we now consider than $m_k > 0$. For this case,

$$m_k > 0 \iff x_{(i_k,j_k)} > u'_{(i_k,j_k)} = u_{(i_k,j_k)} + p_k.$$

We observe that $p_k = x_{(i_k,j_k)} - u_{(i_k,j_k)}$ cannot be given since the above expression gives a contradiction. Therefore, we have that
$p_k = \sum_{(i,j)\in N((i_k,j_k))} u_{(i,j)}$. For this case, $p_{i,j} = u_{(i,j)}$ for all $(i,j) \in N((i_k,j_k))$ and therefore

$$u'_{(i,j)} = 0. \tag{18}$$

First, we have $u'_0 \in \mathbb{N}^{n_\mathcal{D}+n_\mathcal{S}}$ because $m_k \geq 0$ and $u' \in U_x$. Then, we show *(a)* and *(b)* assuming that d_{i_k} is of degree one (the proof for the case that s_{j_k} is of degree one is symmetric and therefore we omitted it). We first show *(a)* for i_k as follows:

$$(u'_0)_{(i_k,j_k)} = u'_{(i_k,j_k)} + m_k = x_{(i_k,j_k)} \leq x_{d_{i_k}}.$$

Finally, we show *(b)* for j_k as follows:

$$\begin{aligned}\sum_{r\in\mathcal{D}(j_k)} (u'_0)_{(r,j_k)} &= (u'_0)_{(i_k,j_k)} + \sum_{r\in\mathcal{D}(j_k)\setminus\{i_k\}} (u'_0)_{(r,j_k)} \\ &= (u'_0)_{(i_k,j_k)} + \sum_{r\in\mathcal{D}(j_k)\setminus\{i_k\}} (u')_{(r,j_k)} \\ &= (u'_0)_{(i_k,j_k)} \\ &= x_{(i_k,j_k)} \\ &\leq x_{s_{j_k}},\end{aligned}$$

where the third equality holds by (18).

Appendix I. Lemma 7

Let $(i_1, j_1) \in \mathcal{E}^*$ and $(i_2, j_2) \in N((i_1, j_1))$. We consider that $i_1 = i_2$ (the case $j_1 = j_2$ is symmetric and therefore it can be proven analogously). Let $\underline{a} \in \mathcal{A}$, $\underline{q} \in \mathcal{Q}$ such that $\underline{q}_{j_2} \geq 1$ and let $\underline{x} = \underline{q} + a$. We define $\overline{q} = \underline{q} + e_{(i_1,j_1)} - e_{(i_2,j_2)}$ and $\overline{x} = \overline{q} + a$. Since $v \in \mathcal{U}_{(i,j)}$, we know that $v(\overline{q}, a) \geq v(\underline{q}, a)$. We aim to show that $L^\theta v(\overline{q}, a) \geq L^\theta v(\underline{q}, a)$.

Let $u_{\overline{x}} \in \arg\min_{u \in U_x} L_u^\theta v(\overline{q}, a)$, using Proposition 4, we can choose $u_{\overline{x}}$ such that $(u_{\overline{x}})_{(i_1,j_1)} = \min(\overline{x}_{d_{i_1}}, \overline{x}_{s_{j_1}})$. Consider that $\overline{x}_{d_{i_1}} \geq 1$, then $(u_{\overline{x}})_{(i_1,j_1)} \geq 1$ because $\overline{x}_{j_1} = \underline{x}_{j_1} + 1 \geq 1$. We define $u_{\underline{x}} = u_{\overline{x}} + e_{(i_2,j_2)} - e_{(i_1,j_1)}$. We know that $u_{\underline{x}} \in U_{\underline{x}}$ since $0 \leq (u_{\overline{x}})_{(i_1,j_1)} - 1 \leq \overline{x}_{s_{j_1}} - 1 = \underline{x}_{s_{j_1}}$, $\sum_{r \in \mathcal{D}(j_2)} (u_{\underline{x}})_{(r,j_2)} = \left(\sum_{r \in \mathcal{D}(j_2)} (u_{\overline{x}})_{(r,j_2)}\right) + 1 \leq \overline{x}_{s_{j_2}} + 1 = \underline{x}_{s_{j_2}}$ and $u_{\overline{x}} \in U_{\overline{x}}$. Besides, $\overline{x} - u_{\overline{x}} = \underline{x} - u_{\underline{x}}$ and the desired result follows since

$$\begin{aligned}
L^\theta v(\underline{q}, a) &\leq L^\theta_{u_{\underline{x}}} v(\underline{q}, a) \\
&= c(\underline{x}) + \theta \mathbb{E}[v(\underline{x} - u_{\underline{x}}, A)] \\
&= c(\underline{x}) + \theta \mathbb{E}[v(\overline{x} - u_{\overline{x}}, A)] \\
&= c(\underline{x}) - c(\overline{x}) + L^\theta v(\overline{q}, A) \\
&< L^\theta v(\overline{q}, A),
\end{aligned}$$

where the first inequality follows since $u_{\underline{x}} \in U_{\underline{x}}$ and the last one since $c(\overline{x}) > c(\underline{x})$. We now consider that $\overline{x}_{d_{i_1}} = 0$, then we cannot match more items in $\overline{x}$ than we could do in $\underline{x}$, i.e. $u_{\overline{x}} \in U_{\underline{x}}$. Indeed, we have $(u_{\overline{x}})_{(i_1,j_1)} \leq \overline{x}_{d_{i_1}} = 0 \leq \underline{x}_{s_{j_1}}$ and $\sum_{r \in \mathcal{D}(j_2)} (u_{\overline{x}})_{(r,j_2)} \leq \overline{x}_{s_{j_2}} \leq \underline{x}_{s_{j_2}}$. Therefore,

$$\begin{aligned}
L^\theta v(\underline{q}, a) &\leq L^\theta_{u_{\overline{x}}} v(\underline{q}, a) \\
&= c(\underline{x}) + \theta \mathbb{E}[v(\underline{x} - u_{\overline{x}}, A)] \\
&\leq c(\underline{x}) + \theta \mathbb{E}[v(\overline{x} - u_{\overline{x}}, A)] \\
&< c(\overline{x}) + \theta \mathbb{E}[v(\overline{x} - u_{\overline{x}}, A)] \\
&= L^\theta v(\overline{q}, A),
\end{aligned}$$

where the first inequality follows since $u_{\overline{x}} \in U_{\underline{x}}$, the second one since $v \in V^\sigma$ and the last one since $c(\overline{x}) > c(\underline{x})$.

Appendix J. Proof of Proposition 5

Proof. Let $a \in \mathcal{A}$, $q \in \mathcal{Q}$, $x = q + a$ and $u \in U_x$. First of all, we need to introduce some notations. Let $m_{(1,1)} = \min(x_{s_1}, x_{d_1})$ (resp. $m_{(3,2)} = \min(x_{s_2}, x_{d_3})$) be the maximal number of matchings that can be done in $(1, 1)$ (resp. in $(3, 2)$). Let

$$K_x = \begin{cases} \{0\} & \text{if } x_{s_1} \leq x_{d_1} \\ \{0, \ldots, \min(x_{s_1} - x_{d_1}, x_{d_2})\} & \text{else} \end{cases}$$

be the set of possible matchings in $(2, 1)$ after matching all the items in $(1, 1)$ and in $(3, 2)$. Let

$$J_x = \begin{cases} \{0\} & \text{if } x_{s_2} \leq x_{d_3} \\ \{0, \ldots, \min(x_{s_2} - x_{d_3}, x_{d_2})\} & \text{else} \end{cases}$$

be the set of possible matchings in $(2, 2)$ after matching all the items in $(1, 1)$ and in $(3, 2)$.

We now remark that K_x does not depend on the number of matchings in $(2, 2)$. Indeed, if $x_{s_1} \geq x_{d_1}$ and $x_{s_2} \geq x_{d_3}$, then $x_{d_2} = x_{s_1} - x_{d_1} + x_{s_2} - x_{d_3}$ because $x \in \mathcal{Q}$, thus $\min(x_{s_1} - x_{d_1}, x_{d_2}) = x_{s_1} - x_{d_1}$ which cannot be modified by any matching in $(2, 2)$. If $x_{s_1} \geq x_{d_1}$ and $x_{s_2} \leq x_{d_3}$, then $J_x = \{0\}$, i.e., we cannot match items in $(2, 2)$. If $x_{s_1} \leq x_{d_1}$, then $K_x = \{0\}$, i.e., we cannot match items in $(2, 1)$ and this cannot be changed with matchings in $(2, 2)$. A symmetric argument can be use to show that J_x does not depend on the number of matchings in $(2, 1)$.

Since $v \in \mathcal{I}_{(1,1)} \cap \mathcal{I}_{(3,2)} \cap \mathcal{U}_{(1,1)} \cap \mathcal{U}_{(3,2)}$, we can use Proposition 4: $\exists u' \in U_x$ such that $L^\theta_{u'} v(q, a) \leq L^\theta_u v(q, a)$ and $u' = m_{(1,1)} e_{(1,1)} + m_{(3,2)} e_{(3,2)} + k e_{(2,1)} + j e_{(2,2)}$ with $k \in K_x$ and $j \in J_x$. Therefore, the desired result follows if we prove that there exist $t_{(2,1)} \in \mathbb{N} \cup \infty$ and $t_{(2,2)} \in \mathbb{N} \cup \infty$ such that

$$L^\theta_{u^*} v(x) \leq L^\theta_{u'} v(x), \quad \forall k \in K_x, \forall j \in J_x \tag{19}$$

where u^* is defined as in Definition 11. Let us first define the threshold in $(2, 1)$ as $t_{(2,1)} = \min\{k \in \mathbb{N} : \mathbb{E}[v((k + 1)e_{(2,1)}, A) - v(k e_{(2,1)}, A)] \geq 0\}$ and the threshold in $(2, 2)$ as $t_{(2,2)} = \min\{j \in \mathbb{N} : \mathbb{E}[v((j + 1)e_{(2,2)}, A) - v(j e_{(2,2)}, A)] \geq 0\}$ (with the convention that $\min\{\emptyset\} = \infty$). Then, given x, there are four cases:

1. $x_{s_1} = x_{d_1}$ and $x_{s_2} = x_{d_3}$. In that case $K_x = \{0\}$ and $J_x = \{0\}$, $k_{t_{(2,1)}} = 0$ and $j_{t_{(2,2)}} = 0$. Thus we have $u^* = u'$.

2. $x_{s_1} > x_{d_1}$ and $x_{s_2} > x_{d_3}$. In that case, we have $x_{d_2} = x_{s_1} - x_{d_1} + x_{s_2} - x_{d_3}$. So the number of matchings in (2, 1) ((2, 2)) for u^* is exactly $k_{t_{(2,1)}}$ ($j_{t_{(2,2)}}$). We define $u^2 = u' - (k - k_{t_{(2,1)}}(x))e_{(2,1)}$. Suppose that $k < k_{t_{(2,1)}}(x)$ (this is only possible if $k_{t_{(2,1)}}(x) = x_{s_1} - x_{d_1} - t_{(2,1)} > 0$), then by definition of $t_{(2,1)}$ and convexity in (2, 1) ($v \in \mathcal{C}_{(2,1)}$) we have: $\forall p \in \{0, \dots, x_{s_1} - x_{d_1} - k - t_{(2,1)} - 1\}$,

$$\mathbb{E}[v((t_{(2,1)} + p + 1)e_{(2,1)}, A) - v((t_{(2,1)} + p)e_{(2,1)}, A)] \geq 0$$
$$\iff L^\theta_{u^2-(p+1)e_{(2,1)}-(j-x_{s_2}+x_{d_3})e_{(2,2)}} v(q, a) - L^\theta_{u^2-pe_{(2,1)}-(j-x_{s_2}+x_{d_3})e_{(2,2)}} v(q, a) \geq 0$$
$$\iff L^\theta_{u^2-(p+1)e_{(2,1)}} v(q, a) - L^\theta_{u^2-pe_{(2,1)}} v(q, a) \geq 0$$

because $v \in \mathcal{M}_{(2,1),(2,2)}$. This means that

$$L^\theta_{u'} v(q, a) \geq L^\theta_{u'+e_{(2,1)}} v(q, a) \geq \cdots \geq L^\theta_{u^2} v(q, a)$$

Suppose now that $k > k_{t_{(2,1)}}(x)$ (this is only possible if $t_{(2,1)} > 0$), then by definition of $t_{(2,1)}$ and convexity in (2, 1) ($v \in \mathcal{C}_{(2,1)}$) we have: $\forall p \in \{0, \dots, \min\{x_{s_1} - x_{d_1}, t_{(2,1)}\} - x_{s_1} + x_{d_1} + k - 1\}$,

$$\mathbb{E}[v((\min\{x_{s_1} - x_{d_1}, t_{(2,1)}\} - p)e_{(2,1)}, A) - v((\min\{x_{s_1} - x_{d_1}, t_{(2,1)}\} - p - 1)e_{(2,1)}, A)] \leq 0$$
$$\iff L^\theta_{u^2+pe_{(2,1)}-(j-x_{s_2}+x_{d_3})e_{(2,2)}} v(q, a) - L^\theta_{u^2+(p+1)e_{(2,1)}-(j-x_{s_2}+x_{d_3})e_{(2,2)}} v(q, a) \leq 0$$
$$\iff L^\theta_{u^2+pe_{(2,1)}} v(q, a) - L^\theta_{u^2+(p+1)e_{(2,1)}} v(q, a) \leq 0$$

because $v \in \mathcal{M}_{(2,1),(2,2)}$ which means that

$$L^\theta_{u^2} v(q, a) \leq \cdots \leq L^\theta_{u'-e_{(2,1)}} v(q, a) \leq L^\theta_{u'} v(q, a)$$

Thus, we showed that $L^\theta_{u^2} v(x) \leq L^\theta_{u'} v(x)$ for any $k \in K_x$ and any $j \in J_x$. Now, let us compare u^2 and u^*. We can do a similar proof as we just did but with (2, 2) instead of (2, 1). Suppose that $j < j_{t_{(2,2)}}(x)$ (this is only possible if $j_{t_{(2,2)}}(x) = x_{s_2} - x_{d_3} - t_{(2,2)} > 0$), then by definition of $t_{(2,2)}$ and convexity in (2, 2) ($v \in \mathcal{C}_{(2,2)}$) we have: $\forall p \in \{0, \dots, x_{s_2} - x_{d_3} - j - t_{(2,2)} - 1\}$,

$$\mathbb{E}[v((t_{(2,2)} + p + 1)e_{(2,2)}, A) - v((t_{(2,2)} + p)e_{(2,2)}, A)] \geq 0$$
$$\iff L^\theta_{u^*-(p+1)e_{(2,2)}-k_{t_{(2,1)}}e_{(2,1)}} v(q, a) - L^\theta_{u^*-pe_{(2,2)}-k_{t_{(2,1)}}e_{(2,1)}} v(q, a) \geq 0$$
$$\iff L^\theta_{u^*-(p+1)e_{(2,2)}} v(q, a) - L^\theta_{u^*-pe_{(2,2)}} v(q, a) \geq 0$$

because $v \in \mathcal{M}_{(2,1),(2,2)}$. This means that

$$L^\theta_{u^2} v(q, a) \geq L^\theta_{u^2+e_{(2,2)}} v(q, a) \geq \cdots \geq L^\theta_{u^*} v(q, a)$$

Suppose now that $j > j_{t_{(2,2)}}(x)$ (this is only possible if $t_{(2,2)} > 0$), then by definition of $t_{(2,2)}$ and convexity in (2, 2) ($v \in \mathcal{C}_{(2,2)}$) we have: $\forall p \in \{0, \dots, \min\{x_{s_2} - x_{d_3}, t_{(2,2)}\} - x_{s_2} + x_{d_3} + j - 1\}$,

$$\mathbb{E}[v((\min\{x_{s_2} - x_{d_3}, t_{(2,2)}\} - p)e_{(2,2)}, A) - v((\min\{x_{s_2} - x_{d_3}, t_{(2,2)}\} - p - 1)e_{(2,2)}, A)] \leq 0$$
$$\iff L^\theta_{u^*+pe_{(2,2)}-k_{t_{(2,1)}}e_{(2,1)}} v(q, a) - L^\theta_{u^*+(p+1)e_{(2,2)}-k_{t_{(2,1)}}e_{(2,1)}} v(q, a) \leq 0$$
$$\iff L^\theta_{u^*+pe_{(2,2)}} v(q, a) - L^\theta_{u^*+(p+1)e_{(2,2)}} v(q, a) \leq 0$$

because $v \in \mathcal{M}_{(2,1),(2,2)}$ which means that

$$L^\theta_{u^*} v(q, a) \leq \cdots \leq L^\theta_{u^2-e_{(2,2)}} v(q, a) \leq L^\theta_{u^2} v(q, a)$$

Thus, we showed that $L^\theta_{u^*} v(q, a) \leq L^\theta_{u^2} v(q, a) \leq L^\theta_{u'} v(q, a)$ for any $k \in K_x$ and any $j \in J_x$.

3. $x_{s_1} \geq x_{d_1}$ and $x_{s_2} \leq x_{d_3}$. In that case, we have $J_x = \{0\}$ and $j_{t_{(2,2)}} = 0$. If $x_{s_1} = x_{d_1}$, then $K_x = \{0\}$ and $k_{t_{(2,1)}} = 0$. Thus, we have $u^* = u'$. Otherwise, suppose that $k < \min\{k_{t_{(2,1)}}(x), x_{d_2}\}$ (this is only possible if $k_{t_{(2,1)}}(x) = x_{s_1} - x_{d_1} - t_{(2,1)} > 0$ and $x_{d_2} > 0$), then by definition of $t_{(2,1)}$ and because $v \in \mathcal{C}_{(2,1)} \cap \mathcal{H}_{(2,1),(3,1)}$, we have: $\forall p \in \{0, \dots, x_{d_2} - k - \max\{t_{(2,1)} - x_{d_3} + x_{s_2}, 0\} - 1\}$,

$$\mathbb{E}[v((\max\{t_{(2,1)}, x_{d_3} - x_{s_2}\} + p + 1)e_{(2,1)}, A) - v((\max\{t_{(2,1)}, x_{d_3} - x_{s_2}\} + p)e_{(2,1)}, A)] \geq 0$$
$$\iff \mathbb{E}[v((\max\{t_{(2,1)} - x_{d_3} + x_{s_2}, 0\} + p + 1)e_{(2,1)} + (x_{d_3} - x_{s_2})e_{(3,1)}, A)$$
$$- v((\max\{t_{(2,1)} - x_{d_3} + x_{s_2}, 0\} + p)e_{(2,1)} + (x_{d_3} - x_{s_2})e_{(3,1)}, A)] \geq 0$$
$$\iff L^\theta_{u^*-(p+1)e_{(2,1)}} v(q, a) - L^\theta_{u^*-pe_{(2,1)}} v(q, a) \geq 0$$

which means that

$$L^{\theta}_{u'}v(q,a) \geq L^{\theta}_{u'+e_{(2,1)}}v(q,a) \geq \cdots \geq L^{\theta}_{u^*}v(q,a)$$

Suppose now that $k > \min\{k_{t_{(2,1)}}(x), x_{d_2}\}$ (this is only possible if $k_{t_{(2,1)}}(x) < x_{d_2}$), then by definition of $t_{(2,1)}$ and because $v \in \mathcal{C}_{(2,1)} \cap \mathcal{H}_{(2,1),(3,1)}$, we have: $\forall p \in \{0,\ldots,\min\{x_{d_2}, t_{(2,1)} - x_{d_3} + x_{s_2}\} - x_{d_2} + k - 1\}$,

$$\begin{aligned}
&\mathbb{E}[v((\min\{x_{s_1} - x_{d_1}, t_{(2,1)}\} - p)e_{(2,1)}, A) - v((\min\{x_{s_1} - x_{d_1}, t_{(2,1)}\} - p - 1)e_{(2,1)}, A)] \leq 0\\
\iff &\mathbb{E}[v((\min\{x_{d_2}, t_{(2,1)} - x_{d_3} + x_{s_2}\} - p)e_{(2,1)} + (x_{d_3} - x_{s_2})e_{(3,1)}, A)\\
&\qquad - v((\min\{x_{d_2}, t_{(2,1)} - x_{d_3} + x_{s_2}\} - p - 1)e_{(2,1)} + (x_{d_3} - x_{s_2})e_{(3,1)}, A)] \leq 0\\
\iff &L^{\theta}_{u^*+pe_{(2,1)}}v(q,a) - L^{\theta}_{u^*+(p+1)e_{(2,1)}}v(q,a) \leq 0
\end{aligned}$$

which means that

$$L^{\theta}_{u^*}v(q,a) \leq \cdots \leq L^{\theta}_{u'-e_{(2,1)}}v(q,a) \leq L^{\theta}_{u'}v(q,a)$$

Thus, we showed that $L^{\theta}_{u^*}v(q,a) \leq L^{\theta}_{u'}v(q,a)$ for any $k \in K_x$ and any $j \in J_x$.

4. $x_{s_1} \leq x_{d_1}$ and $x_{s_2} \geq x_{d_3}$. In that case, we have $K_x = \{0\}$ and $k_{t_{(2,1)}} = 0$. If $x_{s_2} = x_{d_3}$, then $J_x = \{0\}$ and $j_{t_{(2,2)}} = 0$. Thus, we have $u^* = u'$. Otherwise, suppose that $j < \min\{j_{t_{(2,2)}}(x), x_{d_2}\}$ (this is only possible if $j_{t_{(2,2)}}(x) = x_{s_2} - x_{d_3} - t_{(2,2)} > 0$ and $x_{d_2} > 0$), then by definition of $t_{(2,2)}$ and because $v \in \mathcal{C}_{(2,2)} \cap \mathcal{H}_{(2,2),(1,2)}$, we have: $\forall p \in \{0,\ldots,x_{d_2} - j - \max\{t_{(2,2)} - x_{d_1} + x_{s_1}, 0\} - 1\}$,

$$\begin{aligned}
&\mathbb{E}[v((\max\{t_{(2,2)}, x_{d_1} - x_{s_1}\} + p + 1)e_{(2,2)}, A) - v((\max\{t_{(2,2)}, x_{d_1} - x_{s_1}\} + p)e_{(2,2)}, A)] \geq 0\\
\iff &\mathbb{E}[v((\max\{t_{(2,2)} - x_{d_1} + x_{s_1}, 0\} + p + 1)e_{(2,2)} + (x_{d_1} - x_{s_1})e_{(1,2)}, A)\\
&\qquad - v((\max\{t_{(2,2)} - x_{d_1} + x_{s_1}, 0\} + p)e_{(2,2)} + (x_{d_1} - x_{s_1})e_{(1,2)}, A)] \geq 0\\
\iff &L^{\theta}_{u^*-(p+1)e_{(2,2)}}v(q,a) - L^{\theta}_{u^*-pe_{(2,2)}}v(q,a) \geq 0
\end{aligned}$$

which means that

$$L^{\theta}_{u'}v(q,a) \geq L^{\theta}_{u'+e_{(2,2)}}v(q,a) \geq \cdots \geq L^{\theta}_{u^*}v(q,a)$$

Suppose now that $j > \min\{j_{t_{(2,2)}}(x), x_{d_2}\}$ (this is only possible if $j_{t_{(2,2)}}(x) < x_{d_2}$), then by definition of $t_{(2,2)}$ and because $v \in \mathcal{C}_{(2,2)} \cap \mathcal{H}_{(2,2),(1,2)}$, we have: $\forall p \in \{0,\ldots,\min\{x_{d_2}, t_{(2,2)} - x_{d_1} + x_{s_1}\} - x_{d_2} + j - 1\}$,

$$\begin{aligned}
&\mathbb{E}[v((\min\{x_{s_2} - x_{d_3}, t_{(2,2)}\} - p)e_{(2,2)}, A) - v((\min\{x_{s_2} - x_{d_3}, t_{(2,2)}\} - p - 1)e_{(2,2)}, A)] \leq 0\\
\iff &\mathbb{E}[v((\min\{x_{d_2}, t_{(2,2)} - x_{d_1} + x_{s_1}\} - p)e_{(2,2)} + (x_{d_1} - x_{s_1})e_{(1,2)}, A)\\
&\qquad - v((\min\{x_{d_2}, t_{(2,2)} - x_{d_1} + x_{s_1}\} - p - 1)e_{(2,2)} + (x_{d_1} - x_{s_1})e_{(1,2)}, A)] \leq 0\\
\iff &L^{\theta}_{u^*+pe_{(2,2)}}v(q,a) - L^{\theta}_{u^*+(p+1)e_{(2,2)}}v(q,a) \leq 0
\end{aligned}$$

which means that $L^{\theta}_{u^*}v(q,a) \leq \cdots \leq L^{\theta}_{u'-e_{(2,2)}}v(q,a) \leq L^{\theta}_{u'}v(q,a)$. Thus, we showed that $L^{\theta}_{u^*}v(q,a) \leq L^{\theta}_{u'}v(q,a)$ for any $k \in K_x$ and any $j \in J_x$. □

References

[1] A. Cadas, A. Bušić, J. Doncel, Optimal control of dynamic bipartite matching models, in: Proceedings of the 12th EAI International Conference on Performance Evaluation Methodologies and Tools, in: Valuetools 2019, ACM, New York, NY, USA, 2019, pp. 39–46.
[2] J.E. Hopcroft, R.M. Karp, An n^5/2 algorithm for maximum matchings in bipartite graphs, SIAM J. Comput. 2 (4) (1973) 225–231.
[3] R. Duan, S. Pettie, Linear-time approximation for maximum weight matching, J. ACM 61 (1) (2014).
[4] R.M. Karp, U.V. Vazirani, V.V. Vazirani, An optimal algorithm for on-line bipartite matching, in: Proceedings of the Twenty-Second Annual ACM Symposium on Theory of Computing, 1990, pp. 352–358.
[5] V.H. Manshadi, S.O. Gharan, A. Saberi, Online stochastic matching: Online actions based on offline statistics, Math. Oper. Res. 37 (4) (2012) 559–573.
[6] J. Feldman, A. Mehta, V.S. Mirrokni, S. Muthukrishnan, Online stochastic matching: Beating 1-1/e, in: IEEE Computer Society (Ed.), Proceedings of the 2009 50th Annual IEEE Symposium on Foundations of Computer Science, ACM Press, 2009, pp. 117–126.
[7] P. Jaillet, X. Lu, Online stochastic matching: New algorithms with better bounds, Math. Oper. Res. 39 (3) (2014) 624–646.
[8] E.H. Kaplan, Managing the Demand for Public Housing (Ph.D. thesis), Massachusetts Institute of Technology, 1984.
[9] R. Caldentey, E.H. Kaplan, G. Weiss, Fcfs infinite bipartite matching of servers and customers, Adv. Appl. Probab. 41 (3) (2009) 695–730.
[10] I. Adan, G. Weiss, Exact FCFS matching rates for two infinite multitype sequences, Oper. Res. 60 (2) (2012) 475–489.
[11] I. Adan, A. Bušić, J. Mairesse, G. Weiss, Reversibility and further properties of FCFS infinite bipartite matching, Math. Oper. Res. 43 (2) (2018) 598–621.
[12] I. Adan, I. Kleiner, R. Righter, G. Weiss, FCFS parallel service systems and matching models, Perform. Eval. 127–128 (2018) 253–272.

[13] K. Gardner, S. Zbarsky, S. Doroudi, M. Harchol-Balter, E. Hyytiä, A. Scheller-Wolf, Queueing with redundant requests: exact analysis, Queueing Syst. 83 (3–4) (2016) 227–259.
[14] I. Adan, G. Weiss, A skill based parallel service system under FCFS-ALIS — Steady state, overloads, and abandonments, Stoch. Syst. 4 (1) (2014) 250–299.
[15] A. Bušić, V. Gupta, J. Mairesse, Stability of the bipartite matching model, Adv. Appl. Probab. 45 (2) (2013) 351–378.
[16] J. Mairesse, P. Moyal, et al., Stability of the stochastic matching model, J. Appl. Probab. 53 (4) (2016) 1064–1077.
[17] P. Moyal, A. Busic, J. Mairesse, A product form for the general stochastic matching model, 2017, arXiv preprint arXiv:1711.02620.
[18] I. Gurvich, A. Ward, On the dynamic control of matching queues, Stoch. Syst. 4 (2) (2014) 479–523.
[19] A. Busic, S.P. Meyn, Approximate optimality with bounded regret in dynamic matching models, 2014, arXiv e-prints.
[20] M. Nazari, A.L. Stolyar, Reward maximization in general dynamic matching systems, Queueing Syst. (2018).
[21] S.L. Bell, R.J. Williams, Dynamic scheduling of a system with two parallel servers in heavy traffic with resource pooling: Asymptotic optimality of a threshold policy, Ann. Appl. Probab. 11 (3) (2001) 608–649.
[22] J.M. Harrison, Heavy traffic analysis of a system with parallel servers: asymptotic optimality of discrete-review policies, Ann. Appl. Probab. (1998) 822–848.
[23] S. Bell, R. Williams, Dynamic scheduling of a parallel server system in heavy traffic with complete resource pooling: Asymptotic optimality of a threshold policy, Electron. J. Probab. 10 (2005) 1044–1115.
[24] A. Busic, V. Gupta, J. Mairesse, Stability of the bipartite matching model, Adv. Appl. Probab. 45 (2) (2013) 351–378.
[25] M.L. Puterman, Markov Decision Processes: Discrete Stochastic Dynamic Programming, in: Wiley series in probability and statistics, Wiley-Interscience, 2005.
[26] E. Hyon, A. Jean-Marie, Scheduling services in a queuing system with impatience and setup costs, Comput. J. 55 (5) (2012) 553–563.