J. Appl. Prob. **53,** 1064–1077 (2016)
*doi:*10.1017/jpr.2016.65
© *Applied Probability Trust* 2016

STABILITY OF THE STOCHASTIC MATCHING MODEL

JEAN MAIRESSE,* *CNRS and UPMC*
PASCAL MOYAL,** *Université de Technologie de Compiègne and Northwestern University*

Abstract

We introduce and study a new model that we call the *matching model*. Items arrive one by one in a buffer and depart from it as soon as possible but by pairs. The items of a departing pair are said to be *matched*. There is a finite set of classes $\mathcal{V}$ for the items, and the allowed matchings depend on the classes, according to a *matching graph* on $\mathcal{V}$. Upon arrival, an item may find several possible matches in the buffer. This indeterminacy is resolved by a *matching policy*. When the sequence of classes of the arriving items is independent and identically distributed, the sequence of buffer-content is a Markov chain, whose stability is investigated. In particular, we prove that the model may be stable if and only if the matching graph is nonbipartite.

Keywords: Markovian queueing theory; stability; matching; graph

2010 Mathematics Subject Classification: Primary 60J10
Secondary 68M20; 05C75

1. Introduction

A *matching model*, as described in the abstract, is formally specified by a triple (G, Φ, μ) formed by:

- a *matching graph* $G = (\mathcal{V}, \mathcal{E})$, that is, an undirected graph whose vertices $\mathcal{V}$ are the classes of items and whose edges $\mathcal{E}$ are the allowed matchings between classes,
- a *matching policy* Φ which defines the new buffer-content given the pair formed by the old buffer-content and the arriving item,
- a probability μ on $\mathcal{V}$, the common law of the independent and identically distributed (i.i.d.) classes of the arriving items.

The sequence of buffer-content forms a Markov chain. The stability problem consists of determining the conditions on (G, Φ, μ) for the Markov chain to be positive recurrent.

As such, despite being simple and natural, the matching model seems to be original. It has a queueing model flavor, with the crucial specificity that items play the roles of both customers and servers. In spirit, it is related to the general models of 'constrained queueing networks' [12], 'input-queued crossbar switches' [11], or 'call centers with skills-based routing' [8, Section 5].

The present model can be seen as a particular case in discrete time of the *matching queues* introduced in [9], where items may be matched by groups of more than two, and where a control is performed to minimize the holding cost, allowing to keep 'matchable' jobs in line, in order

Received 27 May 2015; revision received 13 January 2016.
* Postal address: Sorbonne Universités, UPMC Univ Paris 06, CNRS, LIP6, 4 place Jussieu, 75252 Paris Cedex 05, France.
** Postal address: LMAC, Université de Technologie de Compiègne, 60203 Compiègne Cedex, France.
Email address: pascal.moyal@utc.fr

https://doi.org/10.1017/jpr.2016.65 Published online by Cambridge University Press

to wait for a more profitable match in the future. However, our approach is widely different in that we consider a fixed matching policy, which prohibits the type of control studied in [9].

The closest connection with existing models in the literature has to be made with the recent 'bipartite matching (BM) model'. This connection plays a central role in several proofs. The BM model was introduced in [6]; see also [1] and [2]. In this context, items arrive by pairs in a buffer and depart from it, as soon as possible, also by pairs. There is a finite number of classes partitioned into 'customer' classes and 'server' classes. Each pair, arriving or departing, is formed by exactly one customer and one server. For departing pairs, an additional requirement is that the customer and the server should be matched, with the allowed matchings depending on the classes only. The sequence of classes of arriving items is i.i.d. and, in each arriving pair, the customer is independent of the server. In [5], the same model was studied without the restriction that the arriving customer and server should be independent. For convenience, let us denote this last model by EBM (extended BM).

Clearly, the (E)BM model and the matching model are close. In fact, the matching model may be viewed as a particular case of the EBM model. Indeed, consider a matching model with graph $(\mathcal{V}, \mathcal{E})$ and sequence of arriving items $(v_n)_n$. Let $\widetilde{\mathcal{V}}$ be a disjoint copy of $\mathcal{V}$. Define a bipartite matching model with customer classes $\mathcal{V}$, server classes $\widetilde{\mathcal{V}}$, possible matches $\{(u, \tilde{v}) \mid (u, v) \in \mathcal{E}\}$, and arriving sequence $(v_n, \widetilde{v}_n)_n$. If the matching policies are the same, then, at any time, the buffer-content of the bipartite matching model is $(U, \widetilde{U})$ if the buffer-content of the original matching model is U. In this bipartite matching model, there is a perfect correlation between the arriving customer and server, so this is indeed an EBM model and not a BM model.

Due to the above connection, we can transfer several results proved for the EBM model in [5] to the matching model. But, on the other hand, we are able to get more precise results in the present context.

1.1. Content

Isolating the matching model as an interesting object of study is the first contribution of the present paper. The second contribution is to show that the matching model may be stable if and only if the matching graph is nonbipartite (see (14), below).

In a nutshell, the situation is as follows. A connected graph G is either bipartite or not. In the first case, we may construct a stable *bipartite* matching model on G (see [5]) but not a stable matching model. In the second case, we may construct a stable matching model on G (and the bipartite matching model is not even defined). Additional results are provided for matching models on a nonbipartite matching graph:

(i) the model is always stable under the natural conditions for the 'match the longest' policy (see (15), below),

(ii) this is not true for all matching policies (see Proposition 3, below). This last result is reminiscent of queueing systems which do not achieve their full capacity region; see, for example, the model with re-entrant lines in [10]. The result (i) on the optimality of 'match the longest', has connections with a result of [12], which stated that in their 'constrained queueing network', the 'max-weight' policy has a maximal stability region. Also, the proofs have the same flavor, as they both use a quadratic Lyapunov function.

1.2. Convention

By default, a *graph* is finite, simple, and undirected, that is, of the form $G = (\mathcal{V}, \mathcal{E})$, with $0 < \#\mathcal{V} < \infty$ and $\mathcal{E} \subset (\mathcal{V} \times \mathcal{V}) \setminus \{(v, v),\ v \in \mathcal{V}\}$, and $(u, v) \in \mathcal{E}$ implies that $(v, u) \in \mathcal{E}$.

https://doi.org/10.1017/jpr.2016.65 Published online by Cambridge University Press

Write $u - v$ for $(u, v) \in \mathcal{E}$ and $u \not\sim v$ for $(u, v) \notin \mathcal{E}$. For $U \subset \mathcal{V}$, define

$$U^c = \mathcal{V} \setminus U, \qquad \mathcal{E}(U) = \{v \in \mathcal{V} \mid \text{there exists } u \in U,\ u - v\}.$$

For $u \in V$, write $\mathcal{E}(u) = \mathcal{E}(\{u\})$. For $U \subset \mathcal{V}$, the *subgraph induced by* U is the graph $(U, \mathcal{E} \cap (U \times U))$. An *independent set* of a graph G is a nonempty subset $\mathcal{I} \subset \mathcal{V}$ which does not include any pair of neighbors, that is, for all $i \neq j \in \mathcal{I}$, $i \not\sim j$. Let $\mathbb{I}$ be the set of independent sets of G.

Given a finite set S, denote by $\mathcal{M}^+(S)$ the set of probability measures μ on S such that, for all i in S, $\mu(i) > 0$.

Denote by $\mathbb{N}$ the set of nonnegative integers. Let A^* be the set of finite words over the alphabet A. For any word $w \in A^*$ and any letter $a \in A$, let $|w|_a$ be the number of occurrences of a in w. Let $|w| = \sum_{a \in A} |w|_a$ be the *length* of w. Let $[w] := (|w|_a)_{a \in A} \in \mathbb{N}^A$ be the *commutative image* of w.

2. The matching model

The *matching model* associated with a graph G, called the *matching graph*, is defined as follows. Start with an empty 'buffer' and, for any n in $\mathbb{N}$, draw an element v_n of $\mathcal{V}$ and apply the following rule:

(i) if there is no element j of $\mathcal{V}$ in the buffer such that $v_n - j$, then add v_n to the buffer,

(ii) otherwise, do not add v_n and remove from the buffer an element j such that $v_n - j$ (we say that v_n and j are *matched* together).

If several elements j of the buffer are such that $v_n - j$, the one to be removed depends on a *matching policy* to be specified.

The sequence $(v_n)_{n \in \mathbb{N}}$ is assumed to be i.i.d. Throughout the paper, we denote by μ the common law over $\mathcal{V}$ of the elements v_n, $n \in \mathbb{N}$. We always assume that $\mu \in \mathcal{M}^+(\mathcal{V})$.

The stability problem of the matching model can be described in the following rough terms: what are the conditions on G, the matching policy, and the distribution μ such that the system is stable, in the sense that the buffer reaches an equilibrium behavior?

Example 1. Consider the matching graph $G = (\mathcal{V}, \mathcal{E})$ with $\mathcal{V} = \{1, 2, 3, 4\}$ and $\mathcal{E} = \{(1, 2), (2, 3), (2, 4), (3, 4)\}$; see Figure 1.

Consider the sequence $(v_n)_n = 3, 1, 4, 2, 4, 4, 1, 2, 3, 2, 1, 4, \ldots$. Denote by U_n the ordered sequence of elements in the buffer after the arrivals of $v_0, \ldots, v_{n-1}$. We have $U_1 = (3)$, $U_2 = (3, 1)$, $U_3 = (1)$, $U_4 = \varnothing$, $U_5 = (4)$, $U_6 = (4, 4)$, $U_7 = (4, 4, 1)$, and U_8 depends on the matching policy: indeed $v_7 = 2$ can be matched either with $v_4 = 4$, $v_5 = 4$, or $v_6 = 1$. A convenient way of visualizing the dynamic is given in Figure 2 (assuming that v_7 is matched with v_6).

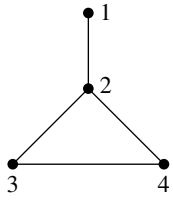


FIGURE 1: The matching graph of Example 1.

https://doi.org/10.1017/jpr.2016.65 Published online by Cambridge University Press

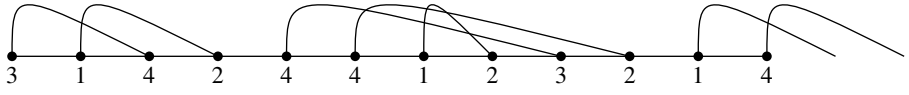


FIGURE 2: The matching model in action, on the matching graph of Figure 1.

3. Structural properties of the matching graph

The conditions Ncond, defined below, will turn out to be necessary for the stability of the matching model (Proposition 2, below). This justifies a thorough study of these conditions, which is the purpose of this section.

Let $G = (\mathcal{V}, \mathcal{E})$ and let $\mu \in \mathcal{M}^+(\mathcal{V})$. Define the following condition on μ.

Ncond(G) : $\quad \mu(\mathcal{I}) < \mu(\mathcal{E}(\mathcal{I}))$ for all $\mathcal{I} \in \mathbb{I}$.

First we obtain the following result.

Lemma 1. *For any connected graph G and $\mu \in \mathcal{M}^+(\mathcal{V})$,* Ncond$(G)$ *is equivalent to*

$$\mu(U) < \mu(\mathcal{E}(U)) \quad \textit{for all } U \subset \mathcal{V},\ U \neq \varnothing,\ U \neq \mathcal{V}. \tag{1}$$

Proof. Fix $\mu \in \mathcal{M}^+(\mathcal{V})$. It is clear that (1) entails Ncond(G), now let us focus on the converse. Consider $U \subset \mathcal{V}, U \neq \varnothing, U \neq \mathcal{V}$, such that U is not an independent set. Note that this implies, in particular, that $|U| > 1$.

(i) Assume first that the subgraph induced by U is connected. Then, for all $u \in U$, there exists $v \in U, u - v$. This implies that $U \subset \mathcal{E}(U)$. Also, since G is connected and $U \neq \mathcal{V}$, we have that $U \subsetneq \mathcal{E}(U)$. Therefore, $\mu(U) < \mu(\mathcal{E}(U))$.

(ii) Assume now that the subgraph induced by U has several connected components. Let $U = U_1 \cup U_2$, where U_1 is the union of the connected components of cardinality 1, and U_2 is the union of the other connected components. The set U_2 is nonempty, otherwise U would be an independent set. Moreover, for all $u \in U_2$, there exists $v \in U_2, u-v$; hence, exactly as in case (i), we have

$$U_2 \subsetneq \mathcal{E}(U_2). \tag{2}$$

Now, if U_1 is empty, we have $U_2 = U$ and (2) allows us to conclude. If not, U_1 is an independent set and, from Ncond(G), we obtain $\mu(U_1) < \mu(\mathcal{E}(U_1))$. Also, (2) entails that $(\mathcal{E}(U_1) \cup U_2) \subset \mathcal{E}(U)$ and since, by definition, $\mathcal{E}(U_1) \cap U_2 = \varnothing$, we obtain

$$\mu(U) = \mu(U_1) + \mu(U_2) < \mu(\mathcal{E}(U_1)) + \mu(U_2) = \mu(\mathcal{E}(U_1) \cup U_2) \leq \mu(\mathcal{E}(U)).$$

This concludes the proof. □

With some abuse of notation, let us denote by Ncond(G), the subset of probability measures $\mu \in \mathcal{M}^+(\mathcal{V})$ satisfying the condition Ncond(G).

Example 2. For the matching graph of Figure 1, the set of independent sets is $\mathbb{I} = \{\{1\}, \{2\}, \{3\}, \{4\}, \{1, 3\}, \{1, 4\}\}$. Therefore, as the total mass of μ is 1, we have

$$\text{Ncond} = \left\{\mu(1) < \mu(2) < \tfrac{1}{2}, \mu(1) + \mu(3) < \tfrac{1}{2}, \mu(1) + \mu(4) < \tfrac{1}{2}\right\}. \tag{3}$$

https://doi.org/10.1017/jpr.2016.65 Published online by Cambridge University Press

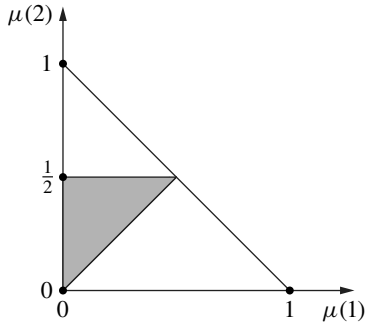


FIGURE 3: The shaded area is the projection of the region $\mathrm{Ncond}(G) \cap \{\mu(3) = \mu(4)\}$.

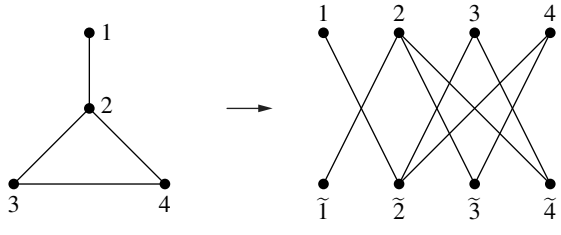


FIGURE 4: The matching graph of Example 1 and its bipartite double cover.

Making the simplifying assumption $\mu(3) = \mu(4)$, we obtain

$$\mathrm{Ncond} \cap \{\mu(3) = \mu(4)\} = \left\{\mu(1) < \mu(2) < \tfrac{1}{2}\right\};$$

see Figure 3.

3.1. Specific conditions for bipartite graphs

Assume that $G = (\mathcal{V}, \mathcal{E})$ is bipartite and let $\mathcal{V} = \mathcal{V}_1 \cup \mathcal{V}_2$ be a bipartition of the vertices in two independent sets. The following condition on the probability measures $\mu \in \mathcal{M}^+(\mathcal{V})$ will be a useful tool in several proofs.

$\mathrm{Ncond}_{1/2}(G)$: $\quad \mu(\mathcal{V}_1) = \mu(\mathcal{V}_2) = \frac{1}{2}$ for all $\mathcal{I} \in \mathbb{I} \setminus \{\mathcal{V}_1, \mathcal{V}_2\}$, $\mu(\mathcal{I}) < \mu(\mathcal{E}(\mathcal{I}))$.

The set of measures $\mathrm{Ncond}_{1/2}(G)$ is defined as for $\mathrm{Ncond}(G)$.

3.2. Bipartite double cover

Given a graph $G = (\mathcal{E}, \mathcal{V})$, its *bipartite double cover* (see, for example, [4]) is the bipartite graph $2 \circ G = (2 \circ \mathcal{V}, 2 \circ \mathcal{E})$ defined by

$$2 \circ \mathcal{V} = \mathcal{V} \cup \{\tilde{u} \mid u \in \mathcal{V}\}, \qquad 2 \circ \mathcal{E} = \{(u, \tilde{v}), (v, \tilde{u}) \mid (u, v) \in \mathcal{E}\},$$

where the set $\tilde{\mathcal{V}} = \{\tilde{u} \mid u \in \mathcal{V}\}$ is a disjoint copy of $\mathcal{V}$. Also denote by $2 \circ \mathbb{I}$ the set of independent sets of $2 \circ G$ and, for all $U \subset 2 \circ \mathcal{V}$, let $2 \circ \mathcal{E}(U)$ be the set of neighbors of the elements of U in $2 \circ G$. The bipartite double cover of the graph of Example 1 is given in Figure 4.

Consider a probability μ on $\mathcal{V}$, and define the probability $2 \circ \mu \in \mathcal{M}^+(2 \circ \mathcal{V})$ by

$$\text{for all } u \in \mathcal{V}, \qquad 2 \circ \mu(u) = 2 \circ \mu(\tilde{u}) = \tfrac{1}{2}\mu(u).$$

Observe the following connection between the conditions $\mathrm{Ncond}(\cdot)$ and $\mathrm{Ncond}_{1/2}(\cdot)$.

https://doi.org/10.1017/jpr.2016.65 Published online by Cambridge University Press

Lemma 2. *For any graph G, we have*

$$\mu \in \mathrm{Ncond}(G) \quad \Longleftrightarrow \quad 2 \circ \mu \in \mathrm{Ncond}_{1/2}(2 \circ G). \tag{4}$$

Proof. ('$\Longrightarrow$') Let $\mathcal{I} \in 2 \circ \mathbb{I}$. We can then write $\mathcal{I} = A \cup \tilde{B}$, $A \subset \mathcal{V}$, $\tilde{B} \subset \tilde{\mathcal{V}}$. Observe that the corresponding subset $A \cup B$ of $\mathcal{V}$ is not an independent set of G in general, because neither A nor B are so. But in view of Lemma 1, we may write

$$\begin{aligned} 2 \circ \mu(\mathcal{I}) &= \tfrac{1}{2}\mu(A) + \tfrac{1}{2}\mu(B) \\ &< \tfrac{1}{2}\mu(\mathcal{E}(A)) + \tfrac{1}{2}\mu(\mathcal{E}(B)) \\ &= 2 \circ \mu(2 \circ \mathcal{E}(A)) + 2 \circ \mu(2 \circ \mathcal{E}(\tilde{B})) \\ &= 2 \circ \mu(2 \circ \mathcal{E}(\mathcal{I})), \end{aligned}$$

where the last equality follows from the fact that $2 \circ \mathcal{E}(A)$ and $2 \circ \mathcal{E}(\tilde{B})$ form a partition of $2 \circ \mathcal{E}(\mathcal{I})$.

('$\Longleftarrow$') Let $\mathcal{I} \in \mathbb{I}$ and let $\tilde{\mathcal{I}}$ be its copy in $\tilde{\mathcal{V}}$. Clearly, $\mathcal{I} \cup \tilde{\mathcal{I}} \in 2 \circ \mathbb{I}$; therefore,

$$\begin{aligned} \mu(\mathcal{I}) &= 2 \circ \mu(\mathcal{I} \cup \tilde{\mathcal{I}}) \\ &< 2 \circ \mu(2 \circ \mathcal{E}(\mathcal{I} \cup \tilde{\mathcal{I}})) \\ &= 2 \circ \mu(2 \circ \mathcal{E}(\mathcal{I})) + 2 \circ \mu(2 \circ \mathcal{E}(\tilde{\mathcal{I}})) \\ &= \tfrac{1}{2}\mu(\mathcal{E}(\mathcal{I})) + \tfrac{1}{2}\mu(\mathcal{E}(\mathcal{I})) \\ &= \mu(\mathcal{E}(\mathcal{I})). \end{aligned}$$

□

3.3. Checking the conditions Ncond

Given G and μ, how can we check efficiently whether the conditions $\mathrm{Ncond}(G)$ hold?

The cardinality of $\mathcal{I}$ is exponential in $|\mathcal{V}|$, so checking directly all the inequalities yields an algorithm of exponential time complexity. But it is possible to do better.

Proposition 1. *Given a graph $G = (\mathcal{V}, \mathcal{E})$ and a probability μ on $\mathcal{V}$, there exists an algorithm of time complexity $O(|\mathcal{V}|^3)$ to decide if μ satisfies* $\mathrm{Ncond}(G)$.

Proof. The result [5, Proposition 3.5] implies in particular that the checking of $\mathrm{Ncond}_{1/2}(2 \circ G)$ can be done with an algorithm of time complexity $O(|\mathcal{V}|^3)$. Using Lemma 2, we obtain the result for $\mathrm{Ncond}(G)$ as a direct corollary. □

3.4. Main result

Theorem 1. *Let G be a connected graph. We have*

$$G \text{ nonbipartite} \quad \Longleftrightarrow \quad \mathrm{Ncond}(G) \neq \varnothing. \tag{5}$$

Proof. Let G be a connected graph. We first prove that

$$G \text{ bipartite} \quad \Longrightarrow \quad \mathrm{Ncond}(G) = \varnothing. \tag{6}$$

Suppose that G is bipartite, and let $\mathcal{V} = \mathcal{V}_1 \cup \mathcal{V}_2$ be a bipartition of the vertices of G. Since G is connected, we have $\mathcal{E}(\mathcal{V}_1) = \mathcal{V}_2$ and $\mathcal{E}(\mathcal{V}_2) = \mathcal{V}_1$. The corresponding conditions in $\mathrm{Ncond}(G)$ are

$$\mu(\mathcal{V}_1) < \mu(\mathcal{V}_2), \qquad \mu(\mathcal{V}_2) < \mu(\mathcal{V}_1); \tag{7}$$

https://doi.org/10.1017/jpr.2016.65 Published online by Cambridge University Press

hence, (6) holds. The following implication was proved in [5, Theorem 4.2]:

$$G \text{ bipartite} \quad \Longrightarrow \quad \mathrm{Ncond}_{1/2}(G) \neq \varnothing. \tag{8}$$

Consequently, comparing (6) and (8) and using Lemma 2, we see that (7) is the only contradiction preventing $\mathrm{Ncond}(G)$ from holding whenever G is connected and bipartite.

It remains to prove that

$$G \text{ nonbipartite} \quad \Longrightarrow \quad \mathrm{Ncond}(G) \neq \varnothing. \tag{9}$$

For this we first need to recall an auxiliary result. Consider a *directed* bipartite graph $D = (\mathcal{V}_1 \cup \mathcal{V}_2, \mathcal{E}_1 \cup \mathcal{E}_2)$, $\mathcal{E}_1 \subset \mathcal{V}_1 \times \mathcal{V}_2$, $\mathcal{E}_2 \subset \mathcal{V}_2 \times \mathcal{V}_1$. Given $\nu \in \mathcal{M}^+(\mathcal{V}_2)$, define $\bar{\nu} \in \mathcal{M}^+(\mathcal{V}_1 \cup \mathcal{V}_2)$ by

$$\bar{\nu}(u) = \tfrac{1}{2}\nu(\mathcal{V}_2 \times \{u\}) \quad \text{for all } u \in \mathcal{V}_1, \qquad \bar{\nu}(u) = \tfrac{1}{2}\nu(\{u\} \times \mathcal{V}_1) \quad \text{for all } u \in \mathcal{V}_2.$$

The next statement is a direct consequence of [5, Theorem 4.2]: if D is strongly connected, then, since G is connected and nonbipartite, the graph

$$UD = (\mathcal{V}_1 \cup \mathcal{V}_2, \mathcal{E}_1 \cup \{(v, u) \mid (u, v) \in \mathcal{E}_1\})$$

is itself connected. Thus, there exist

$$\nu \in \mathcal{M}^+(\mathcal{V}_2), \qquad \bar{\nu} \in \mathrm{Ncond}_{1/2}(UD). \tag{10}$$

Let us return to the proof of (9). The next result is standard and was proved in [4, Theorem 3.4]. If G is connected then

$$G \text{ nonbipartite} \quad \Longleftrightarrow \quad 2 \circ G \text{ connected}.$$

Now, assume that G is connected and nonbipartite, then its bipartite double cover $2 \circ G$ is connected. Consider the *directed* graph D defined by

$$\text{nodes:} \quad 2 \circ \mathcal{V} = \mathcal{V} \cup \tilde{\mathcal{V}}, \qquad \text{arcs:} \quad \{u \to \tilde{v} \mid (u, v) \in \mathcal{E}\} \cup \{\tilde{u} \to u \mid u \in \mathcal{V}\}.$$

It is easy to prove that D is strongly connected. Let us apply (10) to D with $\mathcal{E}_2 = \{\tilde{u} \to u \mid u \in \mathcal{V}\}$. We obtain the existence of $\bar{\nu} \in \mathrm{Ncond}_{1/2}(2 \circ G)$ and, by construction, $\bar{\nu}(u) = \bar{\nu}(\tilde{u})$ for all $u \in \mathcal{V}$. Therefore, according to (4), the probability measure $\mu \in \mathcal{M}^+(\mathcal{V})$ defined by

$$\mu(u) = \bar{\nu}(u) + \bar{\nu}(\tilde{u}), \qquad u \in \mathcal{V},$$

belongs to $\mathrm{Ncond}(G)$. This completes the proof. □

4. Stability of the matching model

To formalize the definition given in Section 2, the matching model is specified by a triple (G, Φ, μ), where the following conditions hold.

- $G = (\mathcal{V}, \mathcal{E})$ is the matching graph defined as in Section 2, and assumed to be connected.

https://doi.org/10.1017/jpr.2016.65 Published online by Cambridge University Press

- Φ is the matching policy defined as follows. We view the state of the buffer as a word over the alphabet $\mathcal{V}$. More precisely, the state space is
$$\mathcal{U} = \{u \in \mathcal{V}^* \mid \text{for all } (i, j) \in \mathcal{E},\ |u|_i \times |u|_j = 0\},$$
and we denote by $U_n \in \mathcal{U}$, the state of the system just before the arrival of item v_n, for any $n \in \mathbb{N}$. The *matching policy* is a mapping $\Phi\colon \mathcal{U} \times \mathcal{V} \to \mathcal{U}$. In words, $\Phi(U, v)$ is the new buffer-content after the arrival of an element v in a buffer of content U. Observe that only the current state of the buffer is taken into account, which is a restriction, but a reasonable one.
- $\mu \in \mathcal{M}^+(\mathcal{V})$ is the probability distribution of the arrivals. Precisely, the sequence of arriving items $(v_n)_{n\in\mathbb{N}}$ is i.i.d. of common law μ.

Let $\mathbf{0}$ be the empty word of $\mathcal{V}^*$. Given a matching model (G, Φ, μ) and a sequence of arrivals $(v_n)_{n\in\mathbb{N}}$, the sequence of buffer-content $(U_n)_{n\in\mathbb{N}}$ is a Markov chain over the state space $\mathcal{U}$ satisfying
$$U_0 = \mathbf{0}, \qquad U_{n+1} = \Phi(U_n, v_n), \quad n \in \mathbb{N}.$$
This Markov chain is clearly irreducible and periodic of period 2. We say that the matching model is *stable* if $(U_n)_{n\in\mathbb{N}}$ is positive recurrent.

Consider the pair (G, Φ) formed by the matching graph and the matching policy. The *stability region* of (G, Φ) is the subset of $\mathcal{M}^+(\mathcal{V})$ formed by the probability measures μ such that (G, Φ, μ) is stable.

4.1. More on matching policies

The matching policy may depend on the order of the items (that is, on their arrival dates). An example is first-come–first-served, where an arriving item of class j is matched with the oldest (if any) item of class i in the buffer such that $j-i$.

Other matching policies are independent of the arrival dates. In such cases, the matching decision at time n depends only on the commutative image $[U]$ of the state $U \in \mathcal{U}$. In other words, the sequence $([U_n])_{n\in\mathbb{N}}$ is a Markov chain on the state space
$$[\mathcal{U}] = \{u \in \mathbb{N}^{\mathcal{V}} \mid \text{for all } (i, j) \in \mathcal{E},\ u_i \times u_j = 0\}.$$
Two such policies are considered below: 'match the longest' and 'priority'. For $i \in \mathcal{V}$, let $e_i \in \mathbb{N}^{\mathcal{V}}$ be defined by $(e_i)_i = 1$ and $(e_i)_j = 0,\ j \neq i$.

Match the longest is the matching policy $\mathrm{ML}\colon [\mathcal{U}] \times \mathcal{V} \to [\mathcal{U}]$ defined by
$$(U, i) \longmapsto \begin{cases} U + e_i & \text{if } j \in \mathcal{E}(i) \implies U_j = 0, \\ U - e_j,\ j = \max\{\arg\max U_{|\mathcal{E}(i)}\} & \text{otherwise,} \end{cases} \tag{11}$$
where $\arg\max U_{|\mathcal{E}(i)}$ is the set of indices k of $\mathcal{E}(i)$ for which U_k is positive and maximal. This set is nonempty and j is the maximum with respect to some given total order on $\mathcal{V}$. In words, the ML gives priority to the more represented compatible class in the buffer.

Let us now define the priority policies. For each $i \in \mathcal{V}$, define the *preferences* of i as a total order on the set $\mathcal{E}(i)$. *Priority* is the matching policy $\Phi\colon [\mathcal{U}] \times \mathcal{V} \to [\mathcal{U}]$ defined by
$$(U, i) \longmapsto \begin{cases} U + e_i & \text{if } j \in \mathcal{E}(i) \implies U_j = 0, \\ U - e_j,\ j = \max\{\mathcal{E}(i) \cap \operatorname{supp} U\} & \text{otherwise,} \end{cases} \tag{12}$$
where $\operatorname{supp} U = \{j \in \mathcal{V} \mid U_j > 0\}$. In the above second case, the set $\mathcal{E}(i) \cap \operatorname{supp} U$ is nonempty and j is the maximum with respect to the preferences of i on $\mathcal{E}(i)$.

https://doi.org/10.1017/jpr.2016.65 Published online by Cambridge University Press

4.2. The results

Our first result states that Ncond are necessary stability conditions. An analog result holds for the bipartite matching model (see [5, Lemma 3.2]).

Proposition 2. *Consider a connected matching graph G and a matching policy Φ. We have, for all $\mu \in \mathcal{M}^+(\mathcal{V})$,*

$$(G, \Phi, \mu) \text{ stable} \quad \Longrightarrow \quad \mu \in \mathrm{Ncond}(G).$$

Proof. First assume that we have $\mu(\mathcal{I}) > \mu(\mathcal{E}(\mathcal{I}))$ for some independent set $\mathcal{I} \subset \mathbb{I}$. For any $n \in \mathbb{N}$, let $\mathcal{I}_n$ be the number of elements of $\mathcal{I}$ in the system at time n, F_n be the number of arrivals of type-$\mathcal{I}$ up to time n, and E_n be the number of arrivals of type-$\mathcal{E}(\mathcal{I})$ up to time n. Finally, denote $H_n = F_n - E_n$. Observe that

$$|U_n| \geq \mathcal{I}_n \geq H_n. \tag{13}$$

By the strong law of large numbers, we obtain

$$\frac{|U_n|}{n} \geq \frac{H_n}{n} \to \mu(\mathcal{I}) - \mu(\mathcal{E}(\mathcal{I})) > 0 \quad \text{as } n \to \infty \text{ a.s.}$$

This implies that $(U_n)_n$ is transient.

Suppose now that for some $\mathcal{I} \in \mathbb{I}$, we have $\mu(\mathcal{I}) = \mu(\mathcal{E}(\mathcal{I}))$. In that case, the Markov chain $(H_n)_n$ is null recurrent. Again, in view of (13), the Markov chain $(U_n)_n$ cannot be positive recurrent. □

The graph G is said to be *separable of order* p, $p \geq 2$, if there exists a partition of $\mathcal{V}$ into independent sets $\mathcal{I}_1, \ldots, \mathcal{I}_p$, such that

$$u - v \quad \text{for all } i \neq j,\ u \in \mathcal{I}_i,\ v \in \mathcal{I}_j.$$

In other words, G is separable of order p if its complementary graph can be partitioned into p cliques. Note that separable graphs of order 2 are bipartite (see Figure 5), whereas separable graphs of order 3 or more are nonbipartite.

Theorem 2. *Consider a connected matching graph G. Let* pol *be the set of matching policies. Let* ML *be the 'match the longest' policy; see (11). We have*

$$G \text{ nonbipartite} \iff \text{there exist } \Phi \in \mathrm{pol},\ \mu \in \mathrm{Ncond}(G) \text{ such that } (G, \Phi, \mu) \text{ is stable,} \tag{14}$$

$$G \text{ nonbipartite} \implies (G, \mathrm{ML}, \mu) \text{ is stable for all } \mu \in \mathrm{Ncond}(G), \tag{15}$$

$$G \text{ separable}, p \geq 3 \implies (G, \Phi, \mu) \text{ is stable for all } \Phi \in \mathrm{pol},\ \mu \in \mathrm{Ncond}(G). \tag{16}$$

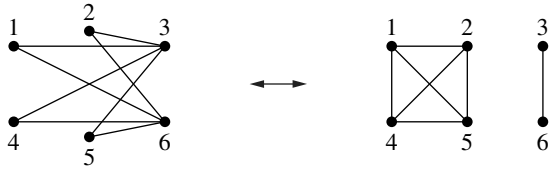


FIGURE 5: Separable graph of order 2 (*left*) and the complementary graph (*right*).

https://doi.org/10.1017/jpr.2016.65 Published online by Cambridge University Press

By merging Theorem 2 with the results from [5], we obtain the following. A connected matching graph G is either bipartite or not. In the first case, we may construct a stable *bipartite* matching model on G (as in [5]) but not a stable matching model. In the second case, we may construct a stable matching model on G (and the bipartite matching model is not even defined).

Proof of Theorem 2. Fix the connected graph G. According to (5), the set $\mathrm{Ncond}(G)$ is nonempty if and only if G is nonbipartite. Therefore, we have

$$\begin{aligned}(G,\Phi,\mu)\text{ stable}\quad\text{there exist }\Phi\in\mathrm{pol},\ \mu\in\mathrm{Ncond}(G)\quad&\Longrightarrow\quad\mathrm{Ncond}(G)\neq\varnothing\\&\Longrightarrow\quad G\text{ nonbipartite.}\end{aligned}\tag{17}$$

Let us now prove that (15) holds. Together with (17), it will also prove (14). Let $\mu\in\mathrm{Ncond}(G)$. Consider the bipartite double cover $2\circ G=(2\circ\mathcal{V},2\circ\mathcal{E})$ of G. According to Lemma 2, we have $2\circ\mu\in\mathrm{Ncond}_{1/2}(2\circ G)$.

Consider the bipartite matching model on the graph $2\circ G$ with matching policy ML and i.i.d. arriving sequence $(v_n,\widetilde{v}_n)_n$ of common law $2\circ\mu$. In the latter, let $([W_n],[\widetilde{W}_n])_n$ be the corresponding buffer-content Markov chain, as defined in [5, Equation (3)], that is, for any $i\in\mathcal{V}$ and $\tilde{i}\in\widetilde{\mathcal{V}}$, let $W_n(i)$ and $\widetilde{W}_n(\tilde{i})$ count the number of buffered items of respective classes i and $\tilde{i}$. Then, if $U_0=W_0$, one can easily check by induction that, for all n, we have $U_n=W_n$ almost surely. Since $2\circ\mu\in\mathrm{Ncond}_{1/2}(2\circ G)$, according to [5, Theorem 7.1], the Markov chain $([W_n],[\widetilde{W}_n])_n$ is positive recurrent. We deduce that $([U_n])_n$ is also positive recurrent, which completes the proof of (15).

The only point that remains to be proved is (16). Assume that G is separable of order $p\geq 3$ and let $\mathcal{I}_1,\dots,\mathcal{I}_p$ be the independent sets partitioning $\mathcal{V}$. For any $\mathcal{I}\in\mathbb{I}$, there exists i such that $\mathcal{I}\subset\mathcal{I}_i$ and $\mathcal{E}(\mathcal{I})^c=\mathcal{I}_i$. Therefore,

$$\mathcal{E}(\mathcal{I})^c\in\mathbb{I}\quad\text{for all }\mathcal{I}\in\mathbb{I}.$$

For $\mathcal{I}\in\mathbb{I}$, define $\mathcal{I}'=\mathcal{E}(\mathcal{I})^c$. Observe that $\mathcal{E}(\mathcal{I})=\mathcal{E}(\mathcal{I}')$. In particular, we have

$$\mu(\mathcal{E}(\mathcal{I})^c)<\mu(\mathcal{E}(\mathcal{I}))\quad\Longleftrightarrow\quad\mu(\mathcal{I}')<\mu(\mathcal{E}(\mathcal{I}')).\tag{18}$$

Since G is nonbipartite, $\mathrm{Ncond}(G)$ is nonempty. Assume that $\mu\in\mathrm{Ncond}(G)$ so that the right-hand side of (18) holds. Therefore, the left-hand side of (18) holds as well for all $\mathcal{I}\in\mathbb{I}$. Consider the Lyapunov function L, defined for all $u\in\mathcal{U}$ by $L(u)=|u|$. Fix $U_n=u\in\mathcal{U}\setminus\{\mathbf{0}\}$, and consider the independent set $\mathcal{I}^u=\{i\in\mathcal{V};\ |u|_i>0\}$. For any matching policy, the size of the buffer decreases (respectively, increases) at time $n+1$ if and only if $v_{n+1}\in\mathcal{E}(\mathcal{I}^u)$ (respectively, $v_{n+1}\notin\mathcal{E}(\mathcal{I}^u)$). Hence,

$$\mathbb{E}[L(U_{n+1})-L(u)\mid U_n=u]=\mu(\mathcal{E}(\mathcal{I}^u)^c)-\mu(\mathcal{E}(\mathcal{I}^u))<0.$$

We conclude that the model is stable by applying the Lyapunov–Foster theorem (see, for example, [3, Section 5.1]). □

5. Detailed study of the model of Example 1

In this section, consider again the matching graph G of Figure 1. For simplicity, fix $\mu\in\mathcal{M}^+(\mathcal{V})$ such that $\mu(3)=\mu(4)$. Let us fix a matching policy and denote by stab the stability region of the model. According to (3), we have $\mathrm{Ncond}(G)=\{\mu(1)<\mu(2)<\frac{1}{2}\}$. By Proposition 2, we have $\mathrm{stab}\subset\{\mu(1)<\mu(2)<\frac{1}{2}\}$. Let us refine this statement with a nontrivial sufficient stability condition.

https://doi.org/10.1017/jpr.2016.65 Published online by Cambridge University Press

Lemma 3. *The stability region satisfies*

$$\mathrm{Ncond}(G) \cap \{\mu(1)(1-\mu(1)) < \mu(2)^2\} \subset \mathrm{stab} \subset \mathrm{Ncond}(G).$$

Proof. We have to prove only the left-hand inclusion. Assume that $\mathrm{Ncond}(G)$ is satisfied. For u in the state space $\mathcal{U}$, set $|u|_{34} = |u|_3 + |u|_4$. Fix η such that

$$\mu(1)\mu(2)^{-1} < 1 - \eta < 1 \tag{19}$$

and consider the Lyapunov function

$$L_\eta \colon \mathcal{U} \to \mathbb{R}_+, \qquad u \longmapsto (1-\eta)|u|_1 + |u|_2 + \mu(1)\mu(2)^{-1}|u|_{34}.$$

Let us compute, for all $n \in \mathbb{N}$,

$$\Delta_\eta = \mathbb{E}[L_\eta(U_{n+1}) - L_\eta(U_n) \mid U_n = u]$$

in the different regions of the state space. If $|u|_2 > 0$, we have

$$\Delta_\eta = \mu(2) - (1-\mu(2)) = 2\mu(2) - 1,$$

so $\Delta_\eta < 0$ according to $\mathrm{Ncond}(G)$. If $|u|_{34} > 0$, we have

$$\Delta_\eta = (1-\eta)\mu(1) + \mu(3)\mu(1)\mu(2)^{-1} - \mu(3)\mu(1)\mu(2)^{-1} - \mu(2)\alpha = (1-\eta)\mu(1) - \mu(2)\alpha,$$

where $\alpha = 1-\eta$ if the arriving item of type-2 is matched with a buffered item of type-1, and $\alpha = \mu(1)\mu(2)^{-1}$ otherwise. From (19), we obtain

$$\Delta_\eta \le (1-\eta)\mu(1) - \mu(2)\mu(1)\mu(2)^{-1} = -\eta\mu(1) < 0.$$

If $|u|_1 > 0$ and $|u|_{34} = 0$, we have

$$\Delta_\eta = (1-\eta)\mu(1) + 2\mu(3)\mu(1)\mu(2)^{-1} - (1-\eta)\mu(2).$$

Replacing $2\mu(3)$ by $1 - \mu(1) - \mu(2)$ and symplifying, we obtain

$$\Delta_\eta < 0 \quad \Longleftrightarrow \quad \eta\mu(2)(\mu(2) - \mu(1)) + \mu(1)(1-\mu(1)) < \mu(2)^2.$$

Applying again the Lyapunov–Foster theorem to the subset $A = \{\mathbf{0}\}$, the model is stable on any region $\mathrm{Ncond}(G) \cap \{\eta\mu(2)(\mu(2) - \mu(1)) + \mu(1)(1-\mu(1)) < \mu(2)^2\}$, for η satisfying (19). By letting η go to 0, we obtain the left-hand inclusion of Lemma 3. □

According to Theorem 2, the ML policy has a maximal stability region and reaches the right-hand bound in Lemma 3. It is then natural to wonder whether there exists a matching policy with the smallest possible stability region, that is, reaching the left-hand bound in Lemma 3. To investigate this question, let us introduce two matching policies of the *priority* type, see (12).

Policy A. 2 gives priority to '3 or 4' over 1.

Policy B. 2 gives priority to 1 over '3 or 4'.

Denote by $\mathrm{stab}(A)$ and $\mathrm{stab}(B)$ the stability regions of policies A and B, respectively.

https://doi.org/10.1017/jpr.2016.65 Published online by Cambridge University Press

We use a simplified state space description $\breve{\mathcal{U}}$, by considering the commutative image of the states and by merging items 3 and 4 as follows:

$$\breve{\mathcal{U}} = \{(0, \ell, 0),\ \ell \in \mathbb{N}\} \cup \{(k, 0, m),\ k, m \in \mathbb{N}\} = \breve{\mathcal{U}}_2 \cup \breve{\mathcal{U}}_{134}.$$

The buffer-content is described by the $\breve{\mathcal{U}}$-valued Markov chain $(\breve{U}_n)_n$, where

$$\breve{U}_n(1) = |U_n|_1, \qquad \breve{U}_n(2) = |U_n|_2, \qquad \breve{U}_n(3) = |U_n|_{34}.$$

Observe that $(\breve{U}_n)_n$ has to go through state $(0, 0, 0)$ to go from $\breve{\mathcal{U}}_2$ to $\breve{\mathcal{U}}_{134}$ (or the other way round). Due to this property, $(\breve{U}_n)_n$ is positive recurrent if and only if the induced Markov chains on $\breve{\mathcal{U}}_2$ and $\breve{\mathcal{U}}_{134}$ are both positive recurrent.

Let us first consider the induced Markov chain on $\breve{\mathcal{U}}_2$. It is the same for the two priority policies and its transition matrix $\boldsymbol{P}$ satisfies

$$\boldsymbol{P}_{i,i-1} = 1 - \mu(2), \qquad \boldsymbol{P}_{i,i+1} = \mu(2), \qquad \text{for all } i \in \mathbb{N} \setminus \{0\}.$$

So the stability condition of the induced chain is $\mu(2) < 1 - \mu(2)$ or in other words $\mu(2) < \frac{1}{2}$.

Now consider the induced Markov chains on $\breve{\mathcal{U}}_{134}$, which depend on the priority policy. The two induced chains are random walks on $\mathbb{Z}_+^2$, meaning that the transition probabilities are homogeneous in the interior of the state space, and along each of the axes. Denote by $\boldsymbol{Q}_A$ and $\boldsymbol{Q}_B$ the transition matrices of the induced chains under the policies A and B respectively. The graphs of $\boldsymbol{Q}_A$ and $\boldsymbol{Q}_B$ are represented in Figure 6, where (i, j) corresponds to the state $(i, 0, j)$.

Let us justify, for example, the coefficients $(\boldsymbol{Q}_A)_{ij,i(j-1)} = \mu(2) + \mu(3)$, $i \geq 0$, $j > 0$. In state (i, j), there are either j items of type-3 or j items of type-4. In the first case (respectively, second case), one of the j items is removed if an item of type-4 (respectively, of type-3) arrives. In both cases, such an event occurs with the same probability $\mu(3) = \mu(4)$. Further, due to the priority policy, one of the j items is also removed whenever an item of type-2 arrives (probability $\mu(2)$).

The detailed study of random walks in $\mathbb{Z}_+^2$ is carried out in the monograph [7]. The salient result [7, Theorem 3.3.1], is the necessary and sufficient condition for positive recurrence in terms of the one-step drifts of the random walk on the interior of the quadrant, and on each of the axes. It applies directly to our context.

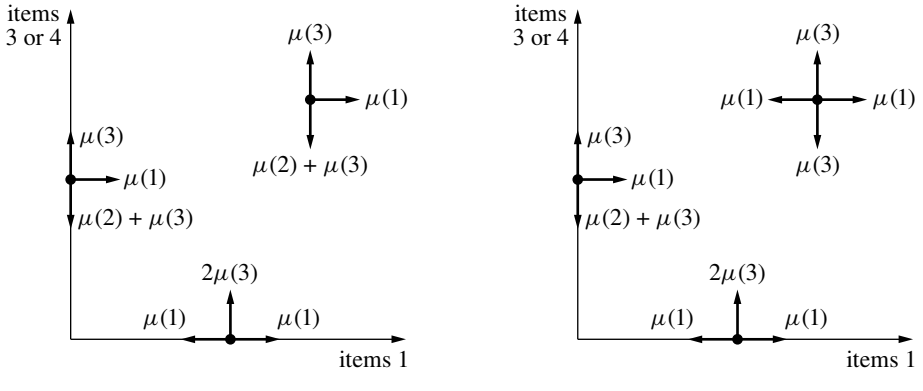


FIGURE 6: The graph of $\boldsymbol{Q}_A$ (*left*) and that of $\boldsymbol{Q}_B$ (*right*).

https://doi.org/10.1017/jpr.2016.65 Published online by Cambridge University Press

Let us first consider policy A. The drifts of the Markov chain are as follows.

Interior $D_x = \mu(1), D_y = -\mu(2)$.

First axis. $D'_x = \mu(1) - \mu(2), D'_y = 2\mu(3)$.

Second axis. $D''_x = \mu(1), D''_y = -\mu(2)$.

Since $D_x > 0$ and $D_y < 0$, the Markov chain is stable if and only if $D_x D'_y - D_y D'_x < 0$; see [7, Theorem 3.3.1]. We have

$$\begin{aligned} D_x D'_y - D_y D'_x < 0 \quad &\iff \quad 2\mu(1)\mu(3) + \mu(2)(\mu(1) - \mu(2)) < 0 \\ &\iff \quad \mu(1)(1 - \mu(1)) < \mu(2)^2. \end{aligned}$$

We now turn to the priority policy B. The drifts of the Markov chain areas follows.

Interior. $D_x = \mu(1) - \mu(2), D_y = 0$.

First axis. $D'_x = \mu(1) - \mu(2), D'_y = 2\mu(3)$.

Second axis. $D''_x = \mu(1), D''_y = -\mu(2)$.

Since $D_x < 0$ and $D_y = 0$, the Markov chain is stable if and only if $D_y D''_x - D_x D''_y < 0$; see [7, Theorem 3.3.1]. We have

$$D_y D''_x - D_x D''_y < 0 \quad \iff \quad \mu(2)(\mu(1) - \mu(2)) < 0 \quad \iff \quad \mu(1) < \mu(2).$$

Summarizing all of the above, we obtain the next proposition.

Proposition 3. *The stability regions under policies A and B are, respectively,*

$$\text{stab}(A) = \text{Ncond}(G) \cap \{\mu(1)(1 - \mu(1)) < \mu(2)^2\} \quad \textit{and} \quad \text{stab}(B) = \text{Ncond}(G).$$

Using Lemma 3, we can rephrase Proposition 3 by saying that policy A has the smallest possible stability region, while policy B has the largest possible stability region. The two stability regions are represented in Figure 7.

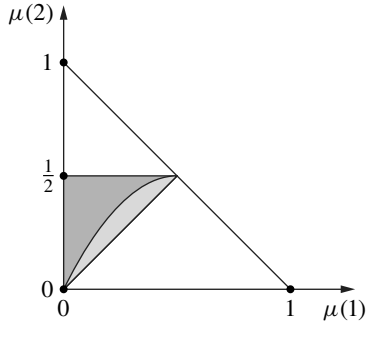


FIGURE 7: The area of stab(A) is shown in darker shading, and stab(B) is the union of the areas with darker and lighter shading.

https://doi.org/10.1017/jpr.2016.65 Published online by Cambridge University Press

References

[1] Adan, I. and Weiss, G. (2012). Exact FCFS matching rates for two infinite multitype sequences. *Operat. Res.* **60,** 475–489.

[2] Adan, I., Bušić, A., Mairesse, J. and Weiss, G. (2016). Reversibility and further properties of FCFS infinite bipartite matching. Available at http://arxiv.org/abs/1507.05939v1.

[3] Brémaud, P. (1999). *Markov Chains: Gibbs Fields, Monte Carlo Simulation, and Queues* (Texts Appl. Math. **31**). Springer, New York.

[4] Brualdi, R. A., Harary, F. and Miller, Z. (1980). Bigraphs versus digraphs via matrices. *J. Graph Theory* **4,** 51–73.

[5] Bušić, A., Gupta, V. and Mairesse, J. (2013). Stability of the bipartite matching model. *Adv. Appl. Prob.* **45,** 351–378.

[6] Caldentey, R., Kaplan, E. H. and Weiss, G. (2009). FCFS infinite bipartite matching of servers and customers. *Adv. Appl. Prob.* **41,** 695–730.

[7] Fayolle, G., Malyshev, V. A. and Menshikov, M. (1995). *Topics in the Constructive Theory of Countable Markov Chains*. Cambridge University Press.

[8] Gans, N., Koole, G. and Mandelbaum, A. (2003). Telephone call centers: tutorial, review, and research prospects. *Manufacturing Serv. Operat. Manag.* **5,** 79–141.

[9] Gurvich, I. and Ward, A. (2014). On the dynamic control of matching queues. *Stoch. Systems* **4,** 479–523.

[10] Kumar, P. R. (1993). Re-entrant lines. *Queueing Systems Theory Appl.* **13,** 87–110.

[11] McKeown, N., Mekkittikul, A., Anantharam, V. and Walrand, J. (1999). Achieving 100% throughput in an input-queued switch. *IEEE Trans. Commun.* **47,** 1260–1267.

[12] Tassiulas, L. and Ephremides, A. (1992). Stability properties of constrained queueing systems and scheduling policies for maximum throughput in multihop radio networks. *IEEE Trans. Automatic Control* **37,** 1936–1948.

https://doi.org/10.1017/jpr.2016.65 Published online by Cambridge University Press