Check for updates

J. Appl. Probab. **62,** 1360–1381 (2025)
*doi:*10.1017/jpr.2025.22

ONLINE MATCHING FOR THE MULTICLASS STOCHASTIC BLOCK MODEL

NAHUEL SOPRANO-LOTO,* ** AND
MATTHIEU JONCKHEERE,* *** *LAAS-CNRS*
PASCAL MOYAL,**** *Université de Lorraine*

Abstract

We consider the problem of sequential matching in a stochastic block model with several classes of nodes and generic compatibility constraints. When the probabilities of connections do not scale with the size of the graph, we show that under the NCOND condition, a simple max-weight type policy allows us to attain an asymptotically perfect matching while no sequential algorithm attains perfect matching otherwise. The proof relies on a specific Markovian representation of the dynamics associated with Lyapunov techniques.

Keywords: Online matching; random graph; stochastic block model; Markov chain

2010 Mathematics Subject Classification: Primary 05C80; 60J10
Secondary 91B68; 93D05; 68Q87

1. Introduction

A matching in a graph is a subset of edges where no two edges share a common vertex. Finding algorithms to determine matchings of maximum cardinality is an important problem in computer science and combinatorial optimization (see [17] and references therein).

Recently, online matching algorithms have gained significant attention because of their wide range of applications in fields such as job markets, advertisements, dates, and transportation businesses. Unlike traditional matching algorithms, which operate on a fixed input graph, online matching algorithms must handle sequentially arriving nodes and return a matching on the fly at each step. This added complexity presents a significant challenge for algorithm design, since the algorithm must make decisions with incomplete information and a limited view of the overall problem.

In their seminal paper, Karp, Vazirani, and Vazirani [10] investigate randomized online matching algorithms for certain classes of non-random graphs. Specifically, they focus on graphs with one-sided bipartite vertex arrivals. The authors use the competitive ratio between the worst online and best offline scenarios as the performance metric for evaluating a policy, where an offline algorithm is one that can wait until all requests are received before determining

Received 14 May 2024; accepted 7 February 2025.
* Postal address: LAAS-CNRS, 7 Av. du Colonel Roche, 31400 Toulouse, France.
** Email address: nahuel.soprano-loto@laas.fr
*** Email address: matthieu.jonckheere@laas.fr
**** Postal address: Université de Lorraine, CNRS, Inria, IECL, F-54000 Nancy, France. Email: pascal.moyal@univ-lorraine.fr

© The Author(s), 2025. Published by Cambridge University Press on behalf of Applied Probability Trust.

https://doi.org/10.1017/jpr.2025.22 Published online by Cambridge University Press

its responses. They establish that a specific randomization strategy can achieve an asymptotically optimal value of order $n(1 - e^{-1})$ for the number of online matchings in a bipartite graph of size n.

More recently, an important line of research focused on online algorithms with, or without, knowledge of the arrival parameters, and showed that structured policies are asymptotically optimal by taking the system to the fluid scale; see, e.g., [6, 11, 15], and references therein. In these references, (asymptotic) optimality is intended as the ability to maximize a given reward or to minimize a cost where, along the various cases, the costs are put on the time elapsed before finding a match (*holding costs*, in the settings of matching queues), and/or the rewards can be put on the edges themselves, allowing, in some cases, heterogeneity in the rewards and costs between the various classes of nodes and/or the various types of edges. In the model considered here, we measure the number of unmatched nodes over time, regardless of their classes, so the costs can be considered as homogeneous in these settings. It is worth mentioning that our techniques can be easily adapted to max-weight policies taking into account heterogeneous rewards in the edges, of the type considered in [8].

On the other end of the spectrum, there has been an important research effort into characterizing the so-called matching number (the cardinality of a maximum matching) of random graphs. Matching numbers on sparse configuration models are studied in detail in [4], generalizing the well-known results of [2, 9]. Randomized online matching algorithms and their properties in terms of competitive ratio have been studied in [7] for regular graphs, and more recently for more general sparse graphs: see [1, 16] for configuration model graphs, and [18] for geometric random graphs.

To the best of our knowledge, there are no results on multiclass matching on random graphs, either from the perspective of matching number characterization or from characterizing the performance of online algorithms, though it is of course a crucial model for applications.

In this study, we address an online matching problem within a stochastic block model consisting of n nodes subject to per-class compatibility constraints. More precisely, a sequence of random graphs is constructed on the fly in the sense that nodes arrive one by one, and their adjacencies are determined upon arrival. This can be seen as an exploration of the large graph of size n. A random class is assigned to every node according to a fixed distribution over a (finite) set of classes, and each node of class i connects with a node of class j with probability $0 \leq \rho(i, j) \leq 1$ (that do not depend on n). The matchings are online in the sense that at most one edge can be added to the matching per time unit.

To tackle this problem, we introduce a family of matching policies based on simple monotonic rules, which extend classical max-weight policies. We then use a joint construction of the stochastic block model and the matching process to reach a Markovian representation of the number of unmatched nodes. This representation, combined with the specification of matching policies, allows us to use Lyapunov techniques by establishing a series of drift inequalities. To achieve this, we build on the techniques introduced in [8]. This in turn allows us to prove stability properties for the underlying Markov process.

This yields relevant results on the asymptotic optimality of online matching algorithms under specific matching policies. We show that, when a certain condition (known as NCOND in the literature [12]) on the probabilities $\rho(i, j)$ is satisfied, the policies that we study are able to achieve asymptotically optimal $O(n)$ matchings. Conversely, we prove that no algorithm can stabilize the process formed by unmatched nodes if the NCOND condition is not met, providing in this way a characterization of the maximum stability region. These findings reveal a phase transition in the difficulty of the matching problem in a multiclass setting, and offer valuable

https://doi.org/10.1017/jpr.2025.22 Published online by Cambridge University Press

insights into the fundamental limits of online matching competitive ratios in such contexts. While our study does not address the case of sparse random graphs, this remains an open question for future research.

Section 2 is dedicated to presenting some necessary background information. In Subsection 2.1 we present the construction of the online stochastic block model, together with the online matching algorithm on that graph. In Subsection 2.2, we introduce the NCOND condition, and state that we cannot expect to have an asymptotically optimal online matching algorithm unless this condition holds. Finally, in Subsection 2.3 we present the max-weight type policies that we consider in this work, and give the Markovian representation of the process of unmatched nodes under the latter policies.

Section 3 is dedicated to presenting the main results of this paper. In Subsection 3.1, Theorem 1 provides a necessary and a sufficient condition for the existence of a perfect matching infinitely often almost surely, under max-weight type matching policies. In Subsection 3.2 we present a consequence of our main result regarding the ergodicity and asymptotic optimality of the corresponding matching algorithm, while Subsection 3.3 discusses the long-run behavior of the process. The proof of Theorem 1 is given in Section 4, while Section 5 is devoted to the proof of our result of asymptotic optimality, Corollary 2.

2. Preliminaries

2.1. Online matching on the stochastic block model

Let $\mathbb{N}_0$ be the set of non-negative integers, and $\mathbb{N}$ the subset of positive integers. In the online version of the stochastic block model, nodes arrive one by one, and the graph is constructed on the fly, generating an increasing sequence of random graphs $\{\mathbf{G}_t\}_{t\in\mathbb{N}_0} = \{(\mathbf{V}_t, \mathbf{E}_t)\}_{t\in\mathbb{N}_0}$. Its construction requires the following ingredients: a finite set of classes $\mathcal{C}$, a probability ν defined on $\mathcal{C}$, and a symmetric matrix $\rho = (\rho(i,j))_{i,j\in\mathcal{C}} \in [0,1]^{\mathcal{C}\times\mathcal{C}}$. Without loss of generality, we assume that ν has full support $\mathcal{C}$, i.e. $\nu(i) > 0$ for every $i \in \mathcal{C}$. If this were not the case, we could simply restrict our study to the support of ν.

The sequence of graphs is defined inductively, together with a sequence of associated matchings on these graphs. Specifically, let $\mathbf{G}_0 = (\mathbf{V}_0, \mathbf{E}_0) = (\emptyset, \emptyset)$ and $\mathbf{M}_0 = \emptyset$. Also, for any $i \in \mathcal{C}$, we introduce the auxiliary elements $\mathbf{V}_0^i = \emptyset$ and $\mathbf{U}_0^i = \emptyset$. We interpret, at time t, the set $\mathbf{M}_t$ as the matching in the graph $\mathbf{G}_t$ ($\mathbf{M}_t$ is a subset of $\mathbf{E}_t$ containing only non-adjacent edges), the set $\mathbf{V}_t^i$ as the set of all nodes of $\mathbf{V}_t$ having class i, and the set $\mathbf{U}_t^i$ as the subset of unmatched nodes of class i. By *unmatched node*, we mean that the node is not the endpoint of any edge in the matching. The class of a node $v \in \mathbf{V}_t$ will be denoted by $c(v) \in \mathcal{C}$, so $v \in \mathbf{V}_t^{c(v)}$. We will say in this case that v is a *c(v) -node*. Also, for a class subset $\mathcal{A} \subset \mathcal{C}$, we say that v is an *$\mathcal{A}$-node* when $c(v) \in \mathcal{A}$.

Suppose we have defined $\mathbf{G}_{t-1} = (\mathbf{V}_{t-1}, \mathbf{E}_{t-1})$, $\mathbf{M}_{t-1} \subset \mathbf{E}_{t-1}$, $\{\mathbf{V}_{t-1}^i\}_{i\in\mathcal{C}}$, and $\{\mathbf{U}_{t-1}^i\}_{i\in\mathcal{C}}$. Then the corresponding objects at time t are defined as follows:

(i) A node v_t is added to the graph, and we set $\mathbf{V}_t = \mathbf{V}_{t-1} \cup \{v_t\}$. We draw the class $c(v_t) \in \mathcal{C}$ of node v_t from the distribution ν on $\mathcal{C}$, independently of everything else, and we set $\mathbf{V}_t^{c(v_t)} = \mathbf{V}_{t-1}^{c(v_t)} \cup \{v_t\}$ and $\mathbf{V}_t^i = \mathbf{V}_{t-1}^i$ for $i \in \mathcal{C} \setminus \{c(v_t)\}$. Next, we associate with every node $v \in \mathbf{V}_{t-1}$ a Bernoulli random variable $\psi_{t,v}$ of parameter $\rho(c(v_t), c(v))$, i.e. $\mathbb{P}(\psi_{t,v} = 1) = 1 - \mathbb{P}(\psi_{t,v} = 0) = \rho(c(v_t), c(v))$, independent of everything else, and we set $\mathbf{E}_t = \mathbf{E}_{t-1} \cup \{\{v_t, v\} : v \in \mathbf{V}_{t-1} \text{ such that } \psi_{t,v} = 1\}$.

https://doi.org/10.1017/jpr.2025.22 Published online by Cambridge University Press

(ii) The matching $\mathbf{M}_t$ is obtained from $\mathbf{M}_{t-1}$ either by keeping it as it was, or by adding one edge in such a way that the resulting family of edges is a matching of $\mathbf{G}_t$. Beyond these restrictions, the choice of $\mathbf{M}_t$ is absolutely arbitrary. For concreteness, the reader can think that $\mathbf{M}_t$ depends on the current and past states of the system, and on an extra independent source of randomness. More precisely, we may view $\mathbf{M}_t$ as a function depending on $\{(\mathbf{G}_s, \{\mathbf{V}_s^i\}_{i\in\mathcal{C}}, \mathbf{M}_s)\}_{s\in\{0,\dots,t-1\}}$, on the class of the incoming node $c(v_t)$, on the Bernoulli random variables $\{\psi_{t,v}\}_{v\in\mathbf{V}_{t-1}}$, and possibly on an external – independent of everything else – source of randomness. The latter function, and thereby the criterion for choosing the edge to be added to the matching at t, if any, is called the *online matching policy*. We have the following alternatives:

- If $\mathbf{M}_t = \mathbf{M}_{t-1}$, set $\mathbf{U}_t^{c(v_t)} = \mathbf{U}_{t-1}^{c(v_t)} \cup \{v_t\}$ and $\mathbf{U}_t^i = \mathbf{U}_{t-1}^i$ for $i \neq c(v_t)$.
- If $\mathbf{M}_t \setminus \mathbf{M}_{t-1}$ has one element, say $\{v, w\}$, set $\mathbf{U}_t^{c(v_t)} = (\mathbf{U}_{t-1}^{c(v_t)} \cup \{v_t\}) \setminus \{v, w\}$ and $\mathbf{U}_t^i = \mathbf{U}_{t-1}^i \setminus \{v, w\}$ for $i \neq c(v_t)$.

In the second case, in which the edge $\{v, w\}$ is added to the matching, we say that the nodes v and w are *matched*. In addition to describing the disjoint set of edges, with some abuse of terminology the word 'matching' will also be used to refer to a pair of nodes that have been matched.

2.1.1. *A note on the classical stochastic block model.* The stochastic block model is a widely used random graph model for studying clustering and community detection. It can be easily defined as follows: a fixed, finite set of nodes V divided into r communities $V^1, \dots, V^r$. For each pair of distinct nodes, a link is created between them in a random fashion, independently of everything else, with a probability that depends only on the communities to which the nodes belong. In the construction of the online stochastic block model, conditioned to having communities $\{\mathbf{V}_t^i\}_{i\in\mathcal{C}}$ at time t, the distribution of the set of nodes $\mathbf{E}_t$ corresponds to that of the stochastic block model with these communities, and the connection probabilities are given by the matrix ρ.

2.2. The NCOND condition

The connection probabilities $(\rho(i,j))_{i,j\in\mathcal{C}}$ define an adjacency matrix $(\mathcal{E}_\rho(i,j))_{i,j\in\mathcal{C}}$ in the following natural way: $\mathcal{E}_\rho(i,j) = \mathbf{1}\{\rho(i,j) > 0\}$ for every $i, j \in \mathcal{C}$. If two classes $i, j \in \mathcal{C}$ are such that $\mathcal{E}_\rho(i,j) = 1$, we say that they are *compatible*. Call $\mathcal{G}_\rho = (\mathcal{C}, \mathcal{E}_\rho)$ the graph defined by this adjacency matrix, a graph that we will refer to as the *root graph*. Observe that this graph is undirected, in the sense that i is compatible with j if and only if j is compatible with i, and that it admits self-loops, which refer to classes that are compatible with themselves. A class-subset $\mathcal{I} \subset \mathcal{C}$ is said to be a $\mathcal{G}_\rho$-independent set if $i, j \in \mathcal{I}$ implies $\mathcal{E}_\rho(i,j) = 0$. According to this definition, the empty set is a $\mathcal{G}_\rho$-independent set. Observe also that if $\mathcal{I}$ is a $\mathcal{G}_\rho$-independent set, then the classes lying in $\mathcal{I}$ cannot have self-loops. It is convenient to define the set

$$\mathbb{I}_\rho = \{\mathcal{I} \subset \mathcal{C} : \mathcal{I} \text{ is a non-empty } \mathcal{G}_\rho\text{-independent set}\}.$$

Also, for a class-subset $\mathcal{A} \subset \mathcal{C}$ we let $\mathcal{N}_\rho(\mathcal{A}) = \{j \in \mathcal{C} : \mathcal{E}_\rho(i,j) = 1 \text{ for some } i \in \mathcal{A}\}$ denote the set of classes that are compatible with some class belonging to $\mathcal{A}$. For $i \in \mathcal{C}$, with some abuse of notation, we write $\mathcal{N}_\rho(i)$ instead of $\mathcal{N}_\rho(\{i\})$. Let

$$\eta = \eta(\mathcal{G}_\rho, \nu) = \min\{\nu(\mathcal{N}_\rho(\mathcal{I})) - \nu(\mathcal{I}) : \mathcal{I} \in \mathbb{I}_\rho\}, \tag{1}$$

https://doi.org/10.1017/jpr.2025.22 Published online by Cambridge University Press

where we write $\nu(\mathcal{A}) = \sum_{i\in\mathcal{A}} \nu(i)$ for $\mathcal{A} \subset \mathcal{C}$. The parameter η will be a key parameter to characterize the stability of the system at hand. The following definition is reminiscent of those given, for instance, in [3, 8].

Definition 1. We say that the pair $(\mathcal{G}_\rho, \nu)$ satisfies NCOND if $\eta > 0$ or, in other words, if $\nu(\mathcal{I}) < \nu(\mathcal{N}_\rho(\mathcal{I}))$ for all $\mathcal{I} \in \mathbb{I}_\rho$.

We emphasize that the definition of NCOND depends on ρ only through $\mathcal{G}_\rho$. In other words, the only thing that matters in deciding whether NCOND is satisfied is not the precise values of the entries of the matrix ρ, but only whether they are zero or not.

For $t \in \mathbb{N}_0$, the vector $X_t = (X_t(i))_{i\in\mathcal{C}}$ defined as $X_t(i) = |\mathbf{U}_t^i|$, $i \in \mathcal{C}$, encodes the number of unmatched nodes in each class. The process $X = \{X_t\}_{t\in\mathbb{N}_0}$, which will be referred to as the *unmatched process*, will be our main object of study. Note that X clearly depends on the online matching policy. Without further assumption, it is a priori not a Markov chain. The following proposition investigates the sequence of random variables $\{\|X_t\|\}_{t\in\mathbb{N}_0}$, where $\|x\| = \max_{i\in\mathcal{C}} |x(i)|$ stands for the supremum norm of a vector $x \in \mathbb{R}^{\mathcal{C}}$. More precisely, if NCOND is not satisfied, the proposition establishes a criterion that implies that the sequence is not *tight*, where by tight we mean that $\lim_{\kappa\to\infty} \limsup_{t\in\mathbb{N}_0} \mathbb{P}[\|X_t\| > \kappa] = 0$. This notion of tightness coincides with the tightness of the sequence of distributions $\{\mu_t\}_{t\in\mathbb{N}_0}$ defined on $\mathbb{N}_0^{\mathcal{C}}$ as $\mu_t(\,\cdot\,) = \mathbb{P}(X_t \in \cdot)$, according to which the sequence $\{\mu_t\}_{t\in\mathbb{N}_0}$ is said to be *tight* if, for every $\varepsilon > 0$ there exists a finite set $\mathscr{F} = \mathscr{F}(\varepsilon) \subset \mathbb{N}_0^{\mathcal{C}}$ such that the inequality $\mu_t(\mathscr{F}) \geq 1 - \varepsilon$ holds for every $t \in \mathbb{N}_0$.

Proposition 1. *For any online matching policy, if* NCOND *is not satisfied, then there exists* $c = c(\nu, \rho) > 0$ *such that*

$$\liminf_{t\to\infty} \mathbb{P}[\|X_t\| \geq \sqrt{t}] \geq c. \tag{2}$$

In particular, the sequence of random variables $(\|X_t\|)_{t\in\mathbb{N}_0}$ *is not tight. Moreover, if* $\eta < 0$ *then*

$$\liminf_{t\to\infty} \frac{\|X_t\|}{t} > 0 \quad \textit{almost surely (a.s.).} \tag{3}$$

Observe that the requirement $\eta < 0$ is stronger than the lack of NCOND. Equation (3) is to be understood as follows: under this requirement, it is not possible to get an online matching algorithm producing a matching of a size of the order of the size of the graph. We will see later that under NCOND this is no longer the case. If $\eta = 0$, Proposition 1 does not make any statement in this regard.

Proof of Proposition 1. We start by establishing a lower bound for the unmatched process. For a fixed $\mathcal{G}_\rho$-independent set $\mathcal{I}$, and for $t \in \mathbb{N}_0$, let $\xi_t^{\mathcal{I}} = \mathbf{1}\{c(i_t) \in \mathcal{I}\} - \mathbf{1}\{c(i_t) \in \mathcal{N}_\rho(\mathcal{I})\}$. Observe that $\{\xi_t^{\mathcal{I}}\}_{t\in\mathbb{N}}$ is an independent and identically distributed sequence of random variables with common distribution

$$\begin{cases} \mathbb{P}[\xi_t^{\mathcal{I}} = 1] & = \nu(\mathcal{I}), \\ \mathbb{P}[\xi_t^{\mathcal{I}} = -1] & = \nu(\mathcal{N}_\rho(\mathcal{I})), \\ \mathbb{P}[\xi_t^{\mathcal{I}} = 0] & = 1 - \nu(\mathcal{I}) - \nu(\mathcal{N}_\rho(\mathcal{I})). \end{cases}$$

https://doi.org/10.1017/jpr.2025.22 Published online by Cambridge University Press

If we define

$$S_t^{\mathcal{I}} = \sum_{s=1}^{t} \xi_s^{\mathcal{I}}, \quad t \in \mathbb{N}_0, \tag{4}$$

then $\{S_t^{\mathcal{I}}\}_{t\in\mathbb{N}_0}$ is a lazy random walk on $\mathbb{Z}$. The key observation is that

$$|\mathcal{C}|\|X_t\| \geq \sum_{i\in\mathcal{I}} X_t(i) \geq S_t^{\mathcal{I}} \quad \text{for all } t \in \mathbb{N}_0, \tag{5}$$

since the number of unmatched $\mathcal{I}$-nodes at time t is equal to the number of $\mathcal{I}$-node arrivals minus the number of matchings involving $\mathcal{I}$-nodes, and all these matchings necessarily involve $\mathcal{N}_\rho(\mathcal{I})$-nodes. Assume now that NCOND does not hold, i.e. $\eta \leq 0$, and let $\mathcal{I} \in \mathbb{I}_\rho$ be such that $\nu(\mathcal{N}_\rho(\mathcal{I})) - \nu(\mathcal{I}) = \eta$. Since $\mu = -\eta$ and $\sigma^2 = \nu(\mathcal{I}) + \nu(\mathcal{N}_\rho(\mathcal{I})) > 0$ are respectively the mean and the variance of $\xi_t^{\mathcal{I}}$, we get

$$\begin{aligned}\mathbb{P}(\|X_t\| \geq \sqrt{t}) \geq \mathbb{P}\Big(S_t^{\mathcal{I}} \geq |\mathcal{C}|\sqrt{t}\Big) &= \mathbb{P}\left(\frac{S_t^{\mathcal{I}} - \mu t}{\sqrt{t}\sigma} \geq \frac{|\mathcal{C}|\sqrt{t} - \mu t}{\sqrt{t}\sigma}\right) \\ &\geq \mathbb{P}\left(\frac{S_t^{\mathcal{I}} - \mu t}{\sqrt{t}\sigma} \geq \frac{|\mathcal{C}|}{\sigma}\right) \xrightarrow[t\to\infty]{} 1 - \Phi\left(\frac{|\mathcal{C}|}{\sigma}\right)\end{aligned}$$

due to the central limit theorem, where Φ represents the cumulative distribution function of the standard Gaussian. Finally, take $c = 1 - \Phi(|\mathcal{C}|/\sigma)$ to get (2).

Assume now that $\eta < 0$, and take $\mathcal{I} \in \mathbb{I}_\rho$ achieving the minimum, that is, such that $\nu(\mathcal{N}_\rho(\mathcal{I})) - \nu(\mathcal{I}) = \eta$. The strong law of large numbers gives $\lim_{t\to\infty} S_t^{\mathcal{I}}/t = -\eta > 0$ a.s., and (3) follows by stochastic comparison. □

Remark 1. The previous proof can be reformulated in a slightly different, but insightful, way. Suppose that $\eta < 0$, let $\mathcal{I} \in \mathbb{I}_\rho$ be a minimizer set, and let $\{S_t^{\mathcal{I}}\}_{t\in\mathbb{N}_0}$ be its associated lazy random walk. Again by the strong law of large numbers,

$$\frac{1}{t}\left[\sum_{i\in\mathcal{I}} |\mathbf{V}_t^i| - \sum_{i\in\mathcal{N}_\rho(\mathcal{I})} |\mathbf{V}_t^i|\right] = \frac{S_t^{\mathcal{I}}}{t} \xrightarrow[t\to\infty]{} -\eta > 0 \quad \text{a.s.}$$

In particular, there exists a (random) point in time after which the set $\bigcup_{i\in\mathcal{I}} \mathbf{V}_t^i$ always has strictly more nodes than $\mathbf{W}_t = \bigcup_{i\in\mathcal{N}_\rho(\mathcal{I})} \mathbf{V}_t^i$. For all such t, the set $\bigcup_{i\in\mathcal{I}} \mathbf{V}_t^i$ is itself a $\mathbf{G}_t$-independent set, so the subgraph of $\mathbf{G}_t$ induced by the node-subset $\mathbf{V}_t \setminus \mathbf{W}_t$ has more than $\sum_{i\in\mathcal{I}} |\mathbf{V}_t^i|$ odd connected components (every singleton that is a subset of an independent set is an odd connected component). Summarizing, $\mathbf{W}_t$ is a node-subset such that the subgraph in $\mathbf{G}_t$ induced by its complement has a number of odd connected components that exceeds its cardinality. Hence, $\mathbf{W}_t$ does not satisfy Tutte's condition [19], implying that $\mathbf{G}_t$ cannot have a perfect matching.

2.3. Matching policies and Markov representation

2.3.1. *ϕ-policies.* We now define a particular family of online matching policies, with the important feature that they will have an associated Markov representation. Since the policies in this family are defined in terms of a function ϕ, we will refer to them as *ϕ-policies*. Fix $t \in \mathbb{N}$,

https://doi.org/10.1017/jpr.2025.22 Published online by Cambridge University Press

suppose that $\mathbf{V}_{t-1}$, $\{\mathbf{V}_{t-1}^i\}_{i\in\mathcal{C}}$, $\mathbf{E}_{t-1}$, $\mathbf{M}_{t-1}$, and $\{\mathbf{U}_{t-1}^i\}_{i\in\mathcal{C}}$ have been defined, and that a node v_t arrives (with class $c(v_t)$ drawn using ν). A class $i_t^* = \phi(X_{t-1}, c(v_t))$ is chosen as a function of the vector X_{t-1} and the class $c(v_t)$, where $\phi\colon \mathbb{N}_0^{\mathcal{C}} \times \mathcal{C} \to \mathcal{C}$ is an arbitrary function. The next step is to construct $\mathbf{E}_t$ by adding edges to $\mathbf{E}_{t-1}$ of the form $\{v_t, v\}$ with $v \in \mathbf{V}_{t-1}$ (through the random variables $\{\psi_{t,v}\}_{v\in\mathbf{V}_{t-1}}$). We have the following alternatives:

- If any of these new edges has an endpoint in $\mathbf{U}_{t-1}^{i_t^*}$, i.e. if the edge-set

$$\tilde{\mathbf{E}}_t = \{e \in \mathbf{E}_t\colon e \cap \mathbf{U}_{t-1}^{i_t^*} \neq \emptyset \text{ and } e \cap \{v_t\} \neq \emptyset\}$$

  is non-empty, then we define $\mathbf{M}_t$ from $\mathbf{M}_{t-1}$ by adding any one of the edges of $\tilde{\mathbf{E}}_t$, no matter which one. Observe that, in this case, we have $X_t(i_t^*) = X_{t-1}(i_t^*) - 1$.

- If, otherwise, $\tilde{\mathbf{E}}_t = \emptyset$, then we set $\mathbf{M}_t = \mathbf{M}_{t-1}$.

In other words, i_t^* is the class of the match of v_t if the latter node is at all matched at time t. We note that, in the second case, an edge is not added to the matching even in the presence of other nodes available to be matched with v_t, i.e. even if there existed a class $i \neq i_t^*$ and a node $v \in \mathbf{U}_{t-1}^i$ with $\{v_t, v\} \in \mathbf{E}_t$.

It is convenient to give an alternative, equivalent way of viewing ϕ-policies. As before, suppose that the elements at time $t-1$ have been defined, that a node v_t has arrived, and that the class i_t^* has been chosen as a function of X_{t-1} and $c(v_t)$. If $\mathbf{U}_{t-1}^{i_t^*} = \emptyset$, then we set $\mathbf{M}_t = \mathbf{M}_{t-1}$. Otherwise, consider an arbitrary enumeration $\mathbf{U}_{t-1}^{i_t^*} = \{w_1, \ldots, w_L\}$. If $\{v_t, w_1\} \in \mathbf{E}_t$, i.e. if $\psi_{t,w_1} = 1$, we set $\mathbf{M}_t = \mathbf{M}_{t-1} \cup \{\{v_t, w_1\}\}$. If, otherwise, $\{v_t, w_1\} \notin \mathbf{E}_t$, we 'try' with w_2; that is, if $\psi_{t,w_2} = 1$ we set $\mathbf{M}_t = \mathbf{M}_{t-1} \cup \{\{v_t, w_2\}\}$. In case we also have that $\{v_t, w_2\} \notin \mathbf{E}_t$, we try with w_3, and so on. If we do not succeed with any of the vertices in $\mathbf{U}_{t-1}^{i_t^*}$, i.e. if $\psi_{t,w_1} = \cdots = \psi_{t,w_L} = 0$, then we set $\mathbf{M}_t = \mathbf{M}_{t-1}$, no matter if there were more edges in $\mathbf{E}_t$ available to be added to $\mathbf{M}_{t-1}$. Note that, to define this algorithm, not all the information about $\mathbf{E}_t$ was used but only that regarding the Bernoulli's random variables until the first success, i.e. until the first index $l \in \{1, \ldots, L\}$ for which $\psi_{t,l} = 1$. This observation could result in an inexpensive implementation of the algorithm.

Under this policy, the stochastic process $\{X_t\}_{t\in\mathbb{N}_0}$ is a discrete-time Markov chain (DTMC) on the state-space $\mathbb{N}_0^{\mathcal{C}}$; we will set out its transition probabilities below. For every $i \in \mathcal{C}$, we set $1_i \in \mathbb{N}_0^{\mathcal{C}}$ to be the canonical vector defined as $1_i(j) = \mathbf{1}\{i = j\}$. Suppose that, at step $t-1$, the state of the DTMC is $x \in \mathbb{N}_0^{\mathcal{C}}$, namely $X_{t-1} = x$. A new node v_t is added to the graph, whose class $i = c(v_t) \in \mathcal{C}$ is drawn from the distribution ν independently of everything else, and then another class $j = i_t^* = \phi(x, i) \in \mathcal{C}$ is chosen. We stress that nothing forbids i and j being the same class a priori. Next, a node $u_t \in \mathbf{U}_{t-1}^j$ is possibly matched with v_t. The non-zero transition probabilities of the DTMC determined by this dynamic are given by

$$P(x, x + 1_i) = \nu(i)(1 - \rho(i, \phi(x, i)))^{x(\phi(x,i))}, \qquad x \in \mathbb{N}_0^{\mathcal{C}}, \tag{6}$$

$$P(x, x - 1_j) = \sum_{i:\phi(x,i)=j} \nu(i)[1 - (1 - \rho(i, j))^{x(j)}], \qquad x \in \mathbb{N}_0^{\mathcal{C}}\colon x - 1_j \in \mathbb{N}_0^{\mathcal{C}}. \tag{7}$$

The transition $P(x, x + 1_i)$ corresponds to the addition of one unmatched i-node. For this to happen, there has to be an arrival of an i -node that is not matched with anyone. This lack of

https://doi.org/10.1017/jpr.2025.22 Published online by Cambridge University Press

matching occurs when no $\phi(x, i)$-node that was unmatched at time $t-1$ became adjacent to the arriving node in the construction of $\mathbf{E}_t$, an event with probability $(1-\rho(i, \phi(x, i)))^{x(\phi(x,i))}$ (when $\mathbf{U}^j_{t-1} \neq \emptyset$, this corresponds to the event $\psi_{t,w_1} = \cdots = \psi_{t,w_L} = 0$ in the aforementioned alternative construction). This explains the right-hand-side of (6). We note that this probability is 1 even in the case in which $\rho(i, \phi(x, i)) = 1$ and $x(\phi(x, i)) = 0$, in concordance with the convention $0^0 = 1$. The transition $P(x, x - 1_j)$ corresponds to the departure of a j -node. For this to occur, there are two necessary conditions. The first one is the arrival of a node whose class i is such that the class chosen by the policy for its match is j (hence the sum on the right-hand-side of (24)). The second is that the matching actually takes place, an event whose complement has probability $(1-\rho(i, j))^{x(j)}$.

2.3.2. *Max-weight type policies.* The kind of policies we will mainly focus on, which we term max-weight policies, are defined as ϕ-policies with a particular choice of ϕ. Let $h\colon \mathcal{C} \to \{1, \ldots, |\mathcal{C}|\}$ be an arbitrary bijective function that establishes a hierarchy or priority between classes, and let $w\colon \mathbb{N}_0 \times [0, 1] \to [0, \infty)$ denote an abstract weight-function. A *max-weight policy* is defined as a ϕ-policy with the following particular choice of ϕ associated to w:

$$\phi_w(x, i) = \arg\max_{j \in \mathcal{C}} (w(x(j), \rho(i, j)), h(j)).$$

The argmax is defined in terms of the lexicographical order between ordered pairs. In other words, $\phi_w(x, i)$ is the class j that maximizes the weight $w(x(j), \rho(i, j))$, where the hierarchy h is used as a tie-breaker.

We now identify some properties over the weight-function w that will be used later as hypotheses in our results.

Assumption 1. (Hypotheses over w.) *The weight-function $w\colon \mathbb{N}_0 \times [0, 1] \to [0, \infty)$ is assumed to satisfy the following conditions:*

(i) $w(n, r) > 0$ *if and only if* $n, r > 0$.

(ii) *w is coordinate-wise non-decreasing: $w(n, r) \leq w(n', r)$ and $w(n, r) \leq w(n, r')$ for every $n \leq n'$ and every $r \leq r'$.*

(iii) *For every $r \in (0, 1]$, the following minimum exists:*

$$n^*(r) = \min\{m \in \mathbb{N}\colon\ w(n-1, 1) < w(n, r) \text{ for all } n \geq m\}. \tag{8}$$

The first condition prevents the selection of a class with which there is no opportunity for a matching to take place (unless all classes meet the same conditions). We will expand on this in Remark 3. The second one is somehow natural: on the one hand, the more individuals there are of a certain class, the more favorable it is to perform a match in order to prevent nodes from accumulating. On the other hand, the larger r is, the higher the probability that the involved matching attempts will actually take place.

Example 1. Two relevant examples are $w_1(n, r) = n\mathbf{1}\{r > 0\}$ and $w_2(n, r) = n(1-(1-r)^n)$. The weight-function w_1 defines a 'match the longest' type of policy, since it only considers the number of stored nodes and not the connection probabilities. For the second one, the quantity $w_2(x(j), \rho(i, j)) = x(j)(1-(1-\rho(i, j))^{x(j)})$ is nothing but the number of j-nodes times the probability of performing a match with one of them, if the class j is chosen.

Remark 2. The – less transparent – third condition in Assumption 1 states that, when the number of nodes of a certain class is large enough, all the relevance of the weight relies on

https://doi.org/10.1017/jpr.2025.22 Published online by Cambridge University Press

the first coordinate and not on the second one (Lemma 3 sheds light on this direction). This makes all such weights equivalent to w_1 in saturated regimes. Nevertheless, we think that it is valuable to consider general weight-functions depending also on the connection probabilities, such as w_2, since they may have better performance in practice.

Remark 3. Let us present an important feature of the family of max-weight policies associated to a w satisfying Assumption 1. Suppose that we stand at $X_{t-1} = x$, and let i be the class of the incoming node. Let

$$\mathcal{S}_x = \{i \in \mathcal{C} : x(i) > 0\} \tag{9}$$

be the support of x. Assume that $\mathcal{N}_\rho(i) \cap \mathcal{S}_x \neq \emptyset$, or, in other words, that there are nodes present in the systems that are likely to be connected to the incoming node. Then the class $j = \phi(x, i)$ chosen by the policy lies in $\mathcal{N}_\rho(i) \cap \mathcal{S}_x$. This (desired) property follows from the first condition over w, from which a class j is such that $w(x(j), \rho(i, j)) > 0$ precisely if and only if $j \in \mathcal{N}_\rho(i) \cap \mathcal{S}_x$.

2.3.3 *State-space.* For simplicity, and without loss of generality in what follows, we assume hereafter that the initial condition of our DTMC X is the empty state $\mathbf{0} \in \mathbb{N}_0^{\mathcal{C}}$. We then define the following set, gathering all states that are reachable from the initial condition:

$$\mathscr{X} = \{x \in \mathbb{N}_0^{\mathcal{C}} : \mathbb{P}(X_t = x) > 0 \text{ for some } t \in \mathbb{N}_0\}.$$

From now on we restrict the DTMC X to the state-space $\mathscr{X}$. The following result provides a sufficient condition for the irreducibility of the DTMC on the latter set.

Proposition 2. *Let P be the transition matrix defined by (6) and (7) associated to a max-weight policy whose weight-function w satisfies Assumption 1. Suppose that, for every $i \in \mathcal{C}$ there exists $j \in \mathcal{C}$ such that $\rho(i, j) > 0$ or, in other words, that $\mathcal{N}_\rho(i) \neq \emptyset$ for every $i \in \mathcal{C}$. Then the transition matrix $(P(x, y))_{x,y \in \mathscr{X}}$ is irreducible.*

Proof. By the very definition of $\mathscr{X}$, for every $x \in \mathscr{X}$ there exists $t \in \mathbb{N}_0$ such that $P^t(\mathbf{0}, x) > 0$ (here P^t denotes the tth power of the matrix P). It remains to show that we can go from any $x \in \mathscr{X}$ to $\mathbf{0}$, namely that there exists $t \in \mathbb{N}_0$ such that $P^t(x, \mathbf{0}) > 0$. Of course, we can assume $x \neq \mathbf{0}$, and hence $\mathcal{S}_x \neq \emptyset$. The hypothesis of our proposition guarantees that $\mathcal{N}_\rho(\mathcal{S}_x) \neq \emptyset$, so the class of the arriving node, say i, belongs to $\mathcal{N}_\rho(\mathcal{S}_x)$ with positive probability. Due to Remark 3, the chosen class $j = \phi(x, i)$ lies in $\mathcal{N}_\rho(i) \cap \mathcal{S}_x$, and two nodes are matched with positive probability. Summarizing, for every $x \in \mathscr{X} \setminus \{\mathbf{0}\}$, there exists $y \in \mathscr{X}$ such that $P(x, y) > 0$ and $\sum_{i \in \mathcal{C}} x(i) = 1 + \sum_{i \in \mathcal{C}} y(i)$. Proceeding in this way finitely many times, we can find $t \in \mathbb{N}$ such that $P^t(x, \mathbf{0}) > 0$. □

Corollary 1. *If* NCOND *holds, then the transition matrix P defined by (6) and (7), and associated to a max-weight policy whose weight-function satisfies Assumption 1, is irreducible.*

Proof. For any node $i \in \mathcal{S}$, the non-empty independent set $\{i\}$ satisfies $\nu(i) < \nu(\mathcal{N}_\rho(i))$. Then $\nu(\mathcal{N}_\rho(i)) > 0$, and hence $\mathcal{N}_\rho(i) \neq \emptyset$. □

3. Main result, and consequences

3.1. Perfect matching infinitely often

From Corollary 1, under NCOND the DTMC $X = \{X_t\}_{t \in \mathbb{N}_0}$ associated to a max-weight type policy is irreducible, and we then say that it is *stable* if it admits a unique invariant distribution.

https://doi.org/10.1017/jpr.2025.22 Published online by Cambridge University Press

Let $\tau_{\mathbf{0}} = \inf\{t \in \mathbb{N} \colon X_t = \mathbf{0}\}$ be the first positive visit to the empty state. Then it is standard that stability is equivalent to the finiteness of the expectation of $\tau_{\mathbf{0}}$, and to the tightness of the sequence of distributions $\{\mu_t\}_{t\in\mathbb{N}_0}$ (defined in Subsection 2.2). This implies in particular that $\mathbb{P}[\tau_{\mathbf{0}} < \infty] = 1$, and in turn, that

$$\mathbb{P}[X_t = \mathbf{0} \text{ infinitely often}] = \mathbb{P}\Big[\limsup_{t\to\infty}\{X_t = \mathbf{0}\}\Big] = 1. \tag{10}$$

Since the empty state $\mathbf{0}$ corresponds to a *perfect matching*, i.e. a matching containing all the nodes, we can say equivalently that, almost surely, a perfect matching occurs infinitely often (i.o.). Our main result, Theorem 1, proves that these good properties hold whenever the condition NCOND is satisfied (i.e. $\eta > 0$), and establishes the opposite conclusion if $\eta < 0$.

Theorem 1. *Let $\rho \in [0,1]^{\mathcal{C}\times\mathcal{C}}$ be a symmetric matrix, and ν a fully supported probability on $\mathcal{C}$. Assume that the pair $(\mathcal{G}_\rho, \nu)$ satisfies* NCOND. *Let $w\colon \mathbb{N}_0 \times [0,1] \to [0,\infty)$ be a weight-function satisfying Assumption 1, and consider the max-weight policy associated to w. Recall the definition of η given in (1). Then the associated DTMC X is stable if and only if* NCOND *holds, that is, $\eta > 0$. In particular, if $\eta > 0$, then almost surely the online matching is perfect infinitely often. If $\eta < 0$, then the probability of having a perfect matching infinitely often is zero.*

Theorem 1 is proved in Section 4. In the statement of the theorem, the necessity of NCOND for stability is simply an application of Proposition 1. In the latter result, the more general concept of tightness can be understood as a measurement of stability even under the lack of Markovianity. In this respect, Proposition 1 and Theorem 1 can be interpreted as meaning that NCOND is the maximum stability region for general online matching policies: the absence of NCOND precludes any stable (tight) policy, whereas its presence allows us to find a stable one.

On another hand, as previously stated, the ability of the online matching algorithm to perform a perfect matching i.o. is reminiscent of the stability of an underlying matching queue, gathering the unmatched nodes. As stated in Theorem 1, this property is guaranteed if and only if NCOND holds (up to the critical case $\eta = 0$), provided that the matching algorithm is induced by a policy ϕ_w whose weight function w satisfies Assumption 1. This provides a simple example of a class of matching policies that have a maximal stability region, or equivalently, of a class of matching algorithms that achieve a perfect matching i.o. almost surely under the minimal condition NCOND. It also immediately follows from [8, Proposition 2] that there exists a measure ν generating the stochastic block model and satisfying NCOND if and only if the underlying graph $\mathcal{G}_\rho$ is not a bipartite simple graph, which is equivalent to saying that the resulting SBM is itself not a bipartite graph i.o.

Observe that the present model is a variant (without reneging) of the matching model with impatience introduced in [8], in which the existence of edges between nodes of compatible classes is random, instead of binary as in [8]. In the binary case, it is easy to construct unstable models, that is, models for which the conclusions of Theorem 1 fail, even though NCOND holds. Specifically, in [12, Section 5], for a simple graph with four nodes, a priority-type policy is introduced for which it is shown that NCOND is not sufficient for the stability of the corresponding stochastic matching system, thereby implying in the present context that, a.s., there does *not* exist a perfect matching i.o. The latter result is widely extended in [14, Theorem 3], where it is shown that NCOND is in general not sufficient for the stability of the considered system, and thereby, for the existence of a perfect matching almost surely i.o., for a wide class of non-bipartite graphs, and for matching policies of the priority type and of the 'random' type.

https://doi.org/10.1017/jpr.2025.22 Published online by Cambridge University Press

Then, extending these results to the present non-binary case is immediate, because performing perfect matchings i.o. is even harder if the existence of edges between elements of respective classes i and j such that $\mathcal{E}_\rho(i,j)=1$ is random (as is the case in the present stochastic block model), instead of certain (as in the case for the binary matching models addressed, e.g., in [8, 12, 14]). It then follows that there exist (non-trivial) online algorithms (e.g. of the 'priority' type) that fail to produce perfect matchings i.o., despite the fact that NCOND holds. In other words, satisfying NCOND is *necessary* for the existence of an efficient online algorithm in that sense (whence the terminology NCOND, for 'necessary condition'), but not sufficient in itself, because online algorithms fail in general to produce perfect matchings i.o., despite the fact that NCOND holds.

3.2. Ergodicity and asymptotic optimality

By using classical results in conjunction with the fundamental drift inequality stated in Proposition 2, we can derive the following result.

Corollary 2. *Under the assumptions of Theorem 1, assume that* NCOND *holds, and let* π *be the unique invariant distribution associated to the DTMC X. Then*

$$\lim_{t\to\infty}\frac{1}{t}\sum_{s=1}^{t}\|X_s\| = \int \|x\|\,\pi(\mathrm{d}x) < \infty \quad \text{a.s.} \tag{11}$$

and, as a consequence,

$$\lim_{t\to\infty}\frac{\|X_t\|}{t} = 0 \quad \text{a.s.} \tag{12}$$

The proof of Corollary 2 is provided in Subsection 5. The first assertion (11) is twofold: on the one hand, it asserts the finiteness of the invariant mean $\int \|x\|\,\pi(\mathrm{d}x)$; on the other, it establishes the convergence of the ergodic averages. A detailed bound for this invariant mean is given in the proof of Corollary 2. Since $\|X_t\|$ is a measure of the number of unmatched nodes at time t (recall that $\|\cdot\|$ is the supremum norm), (12) can be interpreted as meaning that, almost surely, the size of the matching is of the order of the size of the graph. In other words, the matching algorithm is asymptotically optimal.

3.3. Long-run behavior

Assume that NCOND holds (and hence that X is irreducible, due to Corollary 1). Let π be the unique invariant distribution on $\mathscr{X}$, whose existence is guaranteed by Theorem 1. Since the DTMC X is 2-periodic, we cannot expect to have a long-run convergence to π. Nevertheless, a periodic version can be formulated. For $l\in\{0,1\}$, let $\mathscr{X}_l=\left\{x\in\mathscr{X}\colon\ \sum_{i\in\mathcal{C}}x(i)=l\ (\mathrm{mod}\ 2)\right\}$ be the subspace containing the states whose total number of nodes is congruent with l modulo 2, and let π_l be defined as $\pi_l(x)=2\pi(x)\mathbf{1}\{x\in\mathscr{X}_l\}$, $x\in\mathscr{X}$. For $l\in\{0,1\}$, π_l is in fact a probability measure, due to the ergodic theorem (see, for instance, [5, Subsection 4.2.1]). The following result is then simply a restatement of [5, Theorem 4.2.3].

Corollary 3. *Under the assumptions of Corollary 2, the DTMC X satisfies*

$$\lim_{t\to\infty}\sum_{x\in\mathscr{X}_l}\left|\mathbb{P}(X_{2t+l}=x)-\pi_l(x)\right| = 0, \quad l\in\{0,1\}.$$

https://doi.org/10.1017/jpr.2025.22 Published online by Cambridge University Press

4. Proof of Theorem 1

In this section, we prove our main result. The strategy of the proof is that we first prove our fundamental drift inequality, Proposition 2. Its consequences in terms of the stability and existence of a perfect matching i.o. are given in Subsections 4.2 and 4.3.

4.1. Fundamental drift inequality

For a function $f\colon \mathscr{X} \to [0, \infty)$, we define its drift at $x \in \mathscr{X}$ as

$$\mathrm{d}f(x) = Pf(x) - f(x) = \sum_{y \in \mathscr{X}} P(x, y)f(y) - f(x), \tag{13}$$

where P is the transition matrix defined by (6) and (7). After some necessary preliminaries, we will establish an inequality that controls the drift of the quadratic function defined as $q(x) = \sum_{i \in \mathcal{C}} x(i)^2$.

Definition 2. Let r^* be the minimum positive connection probability defined as

$$r^* = \min\{\rho(i, j)\colon \rho(i, j) > 0\}.$$

For every $r \in (0, 1]$, let $n^*(r) \in \mathbb{N}$ be defined as

$$n^*(r) = \min\{m \in \mathbb{N}\colon w(n-1, 1) < w(n, r) \text{ and } w(n, 1) < w(n+1, r) \text{ for all } n \geq m\}.$$

If w satisfies Assumption 1, then the set on which this minimum is taken is non-empty (due to the third condition), and hence the definition of $n^*(r)$ is proper. Also let $n^* = n^*(r^*)$ and $K = \max_{n \in \mathbb{N}_0} n(1 - r^*)^n$.

For $x \in \mathscr{X}$ and $\mathcal{A} \subset \mathcal{C}$, let $\|x\|_{\mathcal{A}} = \max_{i \in \mathcal{A}} x(i)$, and observe that $\|x\|_{\mathcal{C}}$ coincides with the supremum norm $\|x\|$. Finally, define the set of classes with and without self-loops respectively as $\mathcal{C}_\rho^+ = \{i \in \mathcal{C}\colon \rho(i, i) > 0\}$ and $\mathcal{C}_\rho^0 = \{i \in \mathcal{C}\colon \rho(i, i) = 0\}$. Then we have the following result.

Proposition 3. *Let $\rho \in [0, 1]^{\mathcal{C} \times \mathcal{C}}$ be a symmetric matrix, and ν be a fully supported probability on $\mathcal{C}$. Assume that the pair $(\mathcal{G}_\rho, \nu)$ satisfies* NCOND. *Let $w\colon \mathbb{N}_0 \times [0, 1] \to [0, \infty)$ be a weight-function satisfying Assumption 1, and consider the max-weight policy associated to w. Then the quadratic function q satisfies*

$$\mathrm{d}q(x) \leq -2\eta \|x\|_{\mathcal{C}_\rho^0} - \sum_{i \in \mathcal{C}_\rho^+ \colon x(i) \geq n^*} 2x(i)\nu(i) + 1 + 2n^* + 4K\big(1 + \big|\mathcal{C}_\rho^+\big|\big) \tag{14}$$

for any $x \in \mathscr{X}$, and for η defined by (1).

The proof of Proposition 2 is given at the end of this subsection. To get to this, we need to prepare the field through a series of previous results, the proofs of which are given in the Appendix. One of these results, Proposition 8, is a result that was obtained in [8] for a more general framework.

The first step is to compare the drift $\mathrm{d}q$ with the drift with respect to another DTMC with homogeneous values of the connection probabilities. By doing so, we will also deal with classes without too many nodes. Recall the definition of r^*, and let $\hat{\rho}$ be the connection probability matrix defined by $\hat{\rho}(i, j) = r^* \mathbf{1}\{\mathcal{E}_\rho(i, j) = 1\}$, $i, j \in \mathcal{C}$. Consider the transition probabilities

https://doi.org/10.1017/jpr.2025.22 Published online by Cambridge University Press

associated to these connection probabilities. For $i \in \mathcal{C}$,

$$\hat{P}(x, x+1_i) = \nu(i)\big(1-\hat{\rho}_{i,\hat{\phi}(x,i)}\big)^{x(\hat{\phi}(x,i))} \qquad \text{if } x \in \mathscr{X},$$
$$\hat{P}(x, x-1_i) = \sum_{j:\hat{\phi}(x,j)=i} \nu(j)(1-(1-\hat{\rho}(j,i))^{x(i)}) \quad \text{if } x-1_i \in \mathscr{X},$$

with $\hat{\phi}(x,i) = \arg\max_{j\in\mathcal{C}}(w(x(j), \hat{\rho}(i,j)), h(j))$. Let $\hat{\mathrm{d}}q$ be the generator of q associated to these transition probabilities, defined as in (13), by replacing $P(\cdot,\cdot)$ with $\hat{P}(\cdot,\cdot)$.

Lemma 1. (Control of $\tilde{\mathrm{d}}q$ in terms of $\hat{\mathrm{d}}q$.) *Let $\chi : \mathscr{X} \to \mathscr{X}$ be the function that conserves only the classes with more than n^* nodes, that is, for all $x \in \mathscr{X}$,*

$$(\chi(x))(i) = x(i)\mathbf{1}\{x(i) \geq n^*\}, \quad i \in \mathcal{C}. \tag{15}$$

Then $\mathrm{d}q(x) \leq \hat{\mathrm{d}}q(\chi(x)) + 2n^*$, $x \in \mathscr{X}$.

Proof of Lemma 1. We first need to give an expression for the generator of q. First observe that, for all $i \in \mathcal{C}$,

$$q(x+1_i) - q(x) = 1 + 2x(i) \quad \text{for every } x \in \mathscr{X},$$
$$q(x-1_i) - q(x) = 1 - 2x(i) \quad \text{for every } x \in \mathscr{X} \text{ such that } x - 1_i \in \mathscr{X}.$$

Hence, for all $x \in \mathscr{X}$ such that $x - 1_i \in \mathscr{X}$,

$$\begin{aligned}\mathrm{d}q(x) &= \sum_{y\in\mathscr{X}} P(x,y)(q(y)-q(x)) \\ &= \sum_{i\in\mathcal{C}} (1+2x(i))P(x, x+1_i) + \sum_{i\in\mathcal{S}_x} (1-2x(i))P(x, x-1_i) \\ &= 1 + \sum_{i\in\mathcal{S}_x} 2x(i)P(x, x+1_i) - \sum_{i\in\mathcal{S}_x} 2x(i)P(x, x-1_i),\end{aligned} \tag{16}$$

where the last identity follows because, after distributing, the factor accompanying the value 1 represents the total probability of all possible outcomes. Replacing by the values of the transition probabilities (6) and (7), we get

$$\begin{aligned}\mathrm{d}q(x) = 1 &+ \sum_{i\in\mathcal{S}_x} 2x(i)\nu(i)(1-\rho(i,\phi(x,i)))^{x(\phi(x,i))} \\ &- \sum_{i\in\mathcal{S}_x} 2x(i) \sum_{i:\phi(x,i)=j} \nu(i)(1-(1-\rho(i,j))^{x(j)}).\end{aligned} \tag{17}$$

We will make use of the following lemma as follows: if the class of the incoming node has at least one neighbor class with more than n^* nodes, then, for both policies ϕ and $\hat{\phi}$, it will necessarily choose a class whose number of nodes is a maximum. This is the main raison d'être for the third condition of Assumption 1. We have the following preliminary result.

Lemma 2. *Let $x \in \mathscr{X}$ and, recalling (8), let $i \in \mathcal{C}$ be such that $x(j) \geq n^*$ for some $j \in \mathcal{N}_\rho(i)$. Then $x(\phi(x,i)) = x(\hat{\phi}(x,i)) = \max\{x(j) : j \in \mathcal{N}_\rho(i)\}$.*

https://doi.org/10.1017/jpr.2025.22 Published online by Cambridge University Press

Proof of Lemma 2. If $k \in \mathcal{C}$ is such that $x(k) < n^*$, then $k \neq \phi(x, i)$. To see why this is true, take $j \in \mathcal{N}_\rho(i)$ such that $x(j) \geq n^*$. Then, applying the various conditions of Assumption 1,

$$\begin{aligned} w(x(k), \rho(i, k)) \leq w(n^* - 1, \rho(i, k)) \leq w(n^* - 1, 1) &< w(n^*, r^*) \\ &\leq w(x(j), r^*) \\ &\leq w(x(j), \rho(i, j)), \end{aligned}$$

and so the maximum defining $\phi(x, i)$ cannot be attained at k. Hence $x(\phi(x, i)) \geq n^*$. Similarly, $k \notin \mathcal{N}_\rho(i)$ implies that $k \neq \phi(x, i)$, and hence $\phi(x, i) \in \mathcal{N}_\rho(i)$. If there existed $j \in \mathcal{N}_\rho(i)$ such that $x(j) > x(\phi(x, i))$, we would have

$$\begin{aligned} w(x(j), \rho(i, j)) \geq w(x(j), r^*) &> w(x(j) - 1, 1) \\ &\geq w(x(\phi(x, i)), 1) \\ &\geq w(x(\phi(x, i)), \rho(i, \phi(x, i))), \end{aligned}$$

contradicting the definition of $\phi(x, i)$. Hence, such a j does not exist, and so

$$x(\phi(x, i)) = \max\{x(j) \colon j \in \mathcal{N}_\rho(i)\},$$

as desired. The proof for $x(\hat{\phi}(x, i))$ is analogous. □

We can now come back to the proof of Lemma 1. Fix $x \in \mathscr{X}$ and, recalling (15), let $y = \chi(x)$. Bear in mind the following analogous expression of (16) for $\hat{\mathrm{d}}$ instead of d and y instead of x:

$$\hat{\mathrm{d}}q(y) = 1 + \sum_{i \in \mathcal{S}_y} 2y(i)\hat{P}(y, y + 1_i) - \sum_{i \in \mathcal{S}_y} 2y(i)\hat{P}(y, y - 1_i). \tag{18}$$

First of all, we compare the first sums in (16) and (18). For this, define the sets $\mathcal{A} = \{i \in \mathcal{C} \colon x(i) \geq n^*\}$ and

$$\mathcal{A}_1 = \mathcal{A} \cap \mathcal{N}_\rho(\mathcal{A}) = \{i \in \mathcal{A} \colon \mathcal{E}_\rho(i, j) = 1 \text{ for some } j \in \mathcal{A}\}, \qquad \mathcal{A}_2 = \mathcal{A} \setminus \mathcal{A}_1.$$

Then

$$\begin{aligned} &\sum_{i \in \mathcal{S}_x} 2x(i)P(x, x + 1_i) - \sum_{i \in \mathcal{S}_y} 2y(i)\hat{P}(y, y + 1_i) \\ &\quad = \sum_{i \in \mathcal{S}_x \setminus \mathcal{A}} 2x(i)P(x, x + 1_i) + \sum_{i \in \mathcal{A}} 2x(i)P(x, x + 1_i) - \sum_{i \in \mathcal{A}} 2x(i)\hat{P}(y, y + 1_i) \\ &\quad \leq 2n^* + \sum_{i \in \mathcal{A}} 2x(i)[P(x, x + 1_i) - \hat{P}(y, y + 1_i)] \\ &\quad = 2n^* + \sum_{i \in \mathcal{A}_1} 2x(i)\nu(i)\{[1 - \rho(i, \phi(x, i))]^{x(\phi(x,i))} - (1 - r^*)^{y(\phi(y,i))}\} \\ &\qquad + \sum_{i \in \mathcal{A}_2} 2x(i)\nu(i)\big\{[1 - \rho(i, \phi(x, i))]^{x(\phi(x,i))} - 1\big\} \leq 2n^*, \end{aligned}$$

where in the first identity we used that $\mathcal{S}_y = \mathcal{A}$, and that x and y coincide there. In the last inequality, we used that $x(\phi(x, i)) = y(\phi(y, i))$ if $i \in \mathcal{A}_1$, due to Lemma 2.

https://doi.org/10.1017/jpr.2025.22 Published online by Cambridge University Press

We now control the difference between the second sums in (16) and (18). Write the set $\{x(i)\colon i \in \mathcal{A}\}$ (which readily coincides with $\{y(i)\colon i \in \mathcal{A}\}$) as $\{n_1, \dots, n_L\}$, where L is a positive integer and $n_l < n_{l+1}$ for every $l \in \{1, \dots, L-1\}$. Also let $\mathcal{H}_l = \{i \in \mathcal{A}\colon x(i) = n_l\}$ for every $l \in \{1, \dots, L\}$. The $\mathcal{H}_l$ are nothing but the level sets of x (or y) at levels greater than or equal to n^*. Then we easily get that

$$
\begin{aligned}
&-\sum_{i\in\mathcal{S}_x} 2x(i)P(x, x-1_i) + \sum_{i\in\mathcal{S}_y} 2y(i)\hat{P}(y, y-1_i)\\
&\quad\le \sum_{i\in\mathcal{A}} 2x(i)[-P(x, x-1_i) + \hat{P}(y, y-1_i)]\\
&\quad\le \sum_{l=1}^{L} 2n_l \sum_{i\in\mathcal{H}_l} [-P(x, x-1_i) + \hat{P}(y, y-1_i)]\\
&\quad= \sum_{l=1}^{L} 2n_l \left[-\sum_{i\in\mathcal{H}_l} \sum_{j:\phi(x,j)=i} \nu(j)(1-(1-\rho(i,j))^{n_l}) + \sum_{i\in\mathcal{H}_l} \sum_{j:\hat{\phi}(x,j)=i} \nu(j)(1-(1-\hat{\rho}(i,j))^{n_l}) \right]\\
&\quad\le \sum_{l=1}^{L} 2n_l \left[-\sum_{i\in\mathcal{H}_l} \sum_{j:\hat{\phi}(x,j)=i} \nu(j)(1-(1-\hat{\rho}(j,i))^{n_l}) + \sum_{i\in\mathcal{H}_l} \sum_{j:\hat{\phi}(x,j)=i} \nu(j)(1-(1-\hat{\rho}(j,i))^{n_l}) \right]\\
&\quad= 0,
\end{aligned}
$$

where the last identity holds because each of the two sums in the square brackets, due again to Lemma 2, are equal to $(1-(1-r^*)^{n_l})\nu(\{j \in \mathcal{N}_\rho(\mathcal{H}_l)\colon \max_{k\in\mathcal{N}_\rho(j)} x(k) = n_l\})$. Indeed, for the first sum, for instance, after switching the sums we get

$$
\begin{aligned}
\sum_{i\in\mathcal{H}_l} \sum_{j:\phi(x,j)=i} \nu(j)(1-(1-\hat{\rho}(j,i))^{n_l}) &= \sum_{j:\hat{\phi}(x,j)\in\mathcal{H}_l} \nu(j)(1-(1-\hat{\rho}(j,\hat{\phi}(x,j)))^{n_l})\\
&= \sum_{j:\hat{\phi}(x,j)\in\mathcal{H}_l} \nu(j)(1-(1-r^*)^{n_l})\mathbf{1}\{j \in \mathcal{N}_\rho(\mathcal{H}_l)\}\\
&= (1-(1-r^*)^{n_l}) \sum_{j\in\mathcal{N}_\rho(\mathcal{H}_l):\hat{\phi}(x,j)\in\mathcal{H}_l} \nu(j).
\end{aligned}
$$

This concludes the proof. □

For a class-subset $\mathcal{A} \subset \mathcal{C}$, let $1_\mathcal{A}$ be the configuration defined as $(1_\mathcal{A})(i) = \mathbf{1}\{i \in \mathcal{A}\}$. For $x \in \mathcal{X}$, the configuration $x1_\mathcal{A} \in \mathcal{X}$ is defined as the coordinate-wise product, namely $(x1_\mathcal{A})(i) = x(i)\mathbf{1}\{i \in \mathcal{A}\}$ for any $i \in \mathcal{C}$. Recall the definition in (9). The following result allows us to deal with the nodes whose classes have self-loops.

Lemma 3. (Reduction to self-loop-free supports.) *If $x \in \mathcal{X}$ is such that $x(i) \ge n^*$ for every $i \in \mathcal{S}_x$, then $\hat{\mathrm{d}}q(x) \le \hat{\mathrm{d}}q\big(x1_{\mathcal{C}_\rho^0}\big) + 4K|\mathcal{C}_\rho^+| - \sum_{i\in\mathcal{C}_\rho^+} 2x(i)\nu(i)$.*

Proof of Lemma 3. For $k \in \mathcal{C}_\rho^+$, we will prove that $\hat{\mathrm{d}}q(x) \le \hat{\mathrm{d}}q(x1_{\mathcal{C}\setminus\{k\}}) + 4K - 2x(k)\nu(k)$. A simple iteration of this result then proves the required result. Similarly to (17), we have

https://doi.org/10.1017/jpr.2025.22 Published online by Cambridge University Press

$$\hat{\mathrm{d}}q(x) = 1 + \sum_{i\in\mathcal{S}_x} 2x(i)\nu(i)[1-\hat{\rho}(i,\hat{\phi}(x,i))]^{x(\hat{\phi}(x,i))} - \sum_{i\in\mathcal{S}_x} 2x(i) \sum_{j:\hat{\phi}(x,j)=i} \nu(j)(1-(1-\hat{\rho}(j,i))^{x(i)}). \tag{19}$$

We start by controlling the first sum. If $i=k$, we have

$$2x(k)\nu(k)(1-\hat{\rho}(k,\hat{\phi}(x,k)))^{x(\hat{\phi}(x,k))} \le 2x(k)\nu(k)(1-r^*)^{x(k)} \le 2K,$$

where we have used that $x(\hat{\phi}(x,k)) \ge x(k)$ due to Lemma 3. For the same reason, if $i \ne k$,

$$2x(i)\nu(i)(1-\hat{\rho}(i,\hat{\phi}(x,i)))^{x(\hat{\phi}(x,i))} \le 2y(i)\nu(i)(1-\hat{\rho}(i,\hat{\phi}(y,i)))^{y(\hat{\phi}(y,i))},$$

where we have defined $y = x1_{C\setminus\{k\}}$. This inequality is actually an identity if i is not an element of $\mathcal{N}_\rho(k)$. To control the second sum, we first observe that, after switching the sums, we have

$$\sum_{i\in\mathcal{S}_x} 2x(i) \sum_{j:\hat{\phi}(x,j)=i} \nu(j)(1-[1-\hat{\rho}(j,i)]^{x(i)}) = \sum_j 2x(\hat{\phi}(x,j))\nu(j)(1-[1-\hat{\rho}(j,\hat{\phi}(x,j))]^{x(\hat{\phi}(x,j))}).$$

For $j=k$, we have

$$\begin{aligned} 2x(\hat{\phi}(x,k))\nu(k)(1-[1-\hat{\rho}(k,\hat{\phi}(x,k))]^{x(\hat{\phi}(x,k))}) &\ge 2x(k)\nu(k)(1-(1-r^*)^{x(k)}) \\ &= 2x(k)\nu(k) - 2x(k)\nu(k)(1-r^*)^{x(k)} \\ &\ge 2x(k)\nu(k) - 2K. \end{aligned}$$

If $j \ne k$, we again deduce from Lemma 3 that

$$2x(\hat{\phi}(x,j))\nu(j)(1-(1-\hat{\rho}(j,\hat{\phi}(x,j)))^{x(\hat{\phi}(x,j))}) \ge 2y(\hat{\phi}(y,j))\nu(j)(1-(1-\hat{\rho}(j,\hat{\phi}(y,j)))^{y(\hat{\phi}(y,j))}),$$

with equality if $j \notin \mathcal{N}_\rho(k)$. Putting things together, we get

$$\begin{aligned} \hat{\mathrm{d}}q(x) &= 1 + \sum_{i\in\mathcal{S}_x} 2x(i)\nu(i)(1-\hat{\rho}(i,\hat{\phi}(x,i)))^{x(\hat{\phi}(x,i))} \\ &\quad - \sum_j 2x(\hat{\phi}(x,j))\nu(j)(1-(1-\hat{\rho}(j,\hat{\phi}(x,j)))^{x(\hat{\phi}(x,j))}) \\ &\le 1 + \sum_{i\ne k} 2y(i)\nu(i)(1-\hat{\rho}(i,\hat{\phi}(y,i)))^{y(\hat{\phi}(y,i))} \\ &\quad - \sum_{j\ne k} 2y(\hat{\phi}(y,j))\nu(j)(1-(1-\hat{\rho}(j,\hat{\phi}(y,j)))^{x(\hat{\phi}(y,j))}) + 4K - 2x(k)\nu(k) \\ &= \hat{\mathrm{d}}q(y) + 4K - 2x(k)\nu(k), \end{aligned}$$

as desired. □

https://doi.org/10.1017/jpr.2025.22 Published online by Cambridge University Press

The next step reduces the case to configurations whose supports are $\mathcal{G}_\rho$-independent sets.

Lemma 4. (Reduction to supports that are $\mathcal{G}_\rho$-independent sets.) *Let $x \in \mathscr{X}$ be such that $\mathcal{S}_x \subset \mathcal{C}_\rho^0$, and such that $x(i) \geq n^*$ for every $i \in \mathcal{S}_x$. Then there exists $y \in \mathscr{X}$ such that $\|x\| = \|y\|$, $\mathcal{S}_y \subset \mathcal{C}_\rho^0$, $\mathcal{S}_y$ is a $\mathcal{G}_\rho$-independent set, and $\hat{\mathrm{d}}q(x) \leq \hat{\mathrm{d}}q(y) + 2K$.*

We first state the following lemma.

Lemma 5. *Let $x \in \mathscr{X}$ be such that $\mathcal{S}_x \subset \mathcal{C}_\rho^0$ and $x(i) \geq n^*$ for every $i \in \mathcal{S}_x$, and assume that $\mathcal{S}_x$ is not a $\mathcal{G}_\rho$-independent set. Let $\mathcal{A}$ be a connected component of $\mathcal{S}_x$ with $|\mathcal{A}| > 1$, and let $k = \arg\min_{i \in \mathcal{A}}(x(i), h(i))$. Then $\mathrm{d}q(x) \leq \mathrm{d}q(x1_{\mathcal{C}\setminus\{k\}}) + 2\nu(k)K$.*

Proof. Let $y = x1_{\mathcal{C}\setminus\{k\}}$. Recall the expression of the generator in (19). We first dominate the first sum. If $i = k$, we have

$$2x(k)\nu(k)(1 - \hat{\rho}(k, \hat{\phi}(x, k)))^{x(\hat{\phi}(x,k))} \leq 2\nu(k)K.$$

If $i \in \mathcal{S}_x \setminus \{k\}$ is such that $\hat{\phi}(x, i) \neq k$, then $\hat{P}(x, x + 1_i) = \hat{P}(y, y + 1_i)$. Finally, if $i \in \mathcal{S}_x \setminus \{k\}$ is such that $\hat{\phi}(x, i) = k$, then $\mathcal{N}_\rho(i) \cap \mathcal{S}_x = \{k\}$ and hence

$$2x(i)\nu(i)(1 - \hat{\rho}(i, \hat{\phi}(x, i)))^{x(\hat{\phi}(x,i))} \leq 2y(i)\nu(i) = 2y(i)\nu(i)(1 - \hat{\rho}(i, \hat{\phi}(y, i)))^{y(\hat{\phi}(y,i))}.$$

(The identity follows because $y(\hat{\phi}(y, i)) = 0$ due to the fact that $\mathcal{N}_{\hat{\rho}}(i) \cap \mathcal{S}_y = \mathcal{N}_\rho(i) \cap \mathcal{S}_y = \emptyset$.) Putting these together, we get

$$\sum_{i \in \mathcal{S}_x} 2x(i)\nu(i)(1 - \hat{\rho}(i, \hat{\phi}(x, i)))^{x(\hat{\phi}(x,i))} \leq \sum_{i \in \mathcal{S}_y} 2y(i)\nu(i)(1 - \hat{\rho}(i, \hat{\phi}(y, i)))^{y(\hat{\phi}(y,i))} + 2\nu(k)K.$$

The control of the second sum is more delicate. Consider the sets

$$\mathcal{B}_1 = \{i \in \mathcal{N}_\rho(k) \colon \hat{\phi}(x, i) = k\}, \qquad \mathcal{B}_2 = (\mathcal{N}_\rho(\mathcal{B}_1) \cap \mathcal{S}_x) \setminus \{k\}.$$

On the one hand, observe that if $i \in \mathcal{S}_x \setminus (\mathcal{B}_2 \cup \{k\})$ we have $\hat{P}(x, x - 1_i) = \hat{P}(y, y - 1_i)$. On the other hand, we get

$$\begin{aligned}
\sum_{i \in \mathcal{B}_2} 2y(i)\hat{P}(y, y - 1_i) &= \sum_{i \in \mathcal{B}_2} 2y(i) \sum_{j : \hat{\phi}(y,j) = i} \nu(j)(1 - (1 - \hat{\rho}(j, i))^{y(i)}) \\
&= \sum_{j : \hat{\phi}(y,j) \in \mathcal{B}_2} 2y(\hat{\phi}(y, j))\nu(j)(1 - (1 - \hat{\rho}(j, \hat{\phi}(y, j)))^{y(\hat{\phi}(y,j))}) \\
&= \sum_{j \in \mathcal{B}_1^c : \hat{\phi}(y,j) \in \mathcal{B}_2} 2y(\hat{\phi}(y, j))\nu(j)(1 - (1 - \hat{\rho}(j, \hat{\phi}(y, j)))^{y(\hat{\phi}(y,j))}) \\
&\quad + \sum_{j \in \mathcal{B}_1} 2y(\hat{\phi}(y, j))\nu(j)(1 - (1 - \hat{\rho}(j, \hat{\phi}(y, j)))^{y(\hat{\phi}(y,j))}) \\
&\leq \sum_{j : \hat{\phi}(x,j) \in \mathcal{B}_2} 2x(\hat{\phi}(x, j))\nu(j)(1 - (1 - \hat{\rho}(j, \hat{\phi}(x, j)))^{x(\hat{\phi}(x,j))}) \\
&\quad + \sum_{j \in \mathcal{B}_1} 2x(k)\nu(j)(1 - (1 - \hat{\rho}(j, k))^{x(k)}) \\
&= \sum_{i \in \mathcal{B}_2} 2x(i)\hat{P}(x, x - 1_i) + 2x(k)\hat{P}(x, x - 1_k).
\end{aligned}$$

https://doi.org/10.1017/jpr.2025.22 Published online by Cambridge University Press

The inequality holds because if $j \in \mathcal{B}_1$ then $\hat{\phi}(x, j) = k$ and hence $y(\hat{\phi}(y, j)) \le x(k)$. All in all, we get

$$\begin{aligned}
&-\sum_{i\in\mathcal{S}_x} 2x(i)\hat{P}(x, x-1_i) \\
&\quad = -\sum_{i\in\mathcal{S}_x\setminus(\mathcal{B}_2\cup\{k\})} 2x(i)\hat{P}(x, x-1_i) - \sum_{i\in\mathcal{B}_2} 2x(i)\hat{P}(x, x-1_i) - 2x(k)\hat{P}(x, x-1_k) \\
&\quad \le -\sum_{i\in\mathcal{S}_y\setminus\mathcal{B}_2} 2y(i)\hat{P}(y, y-1_i) - \sum_{i\in\mathcal{B}_2} 2y(i)\hat{P}(y, y-1_i) \\
&\quad = -\sum_{i\in\mathcal{S}_y} 2y(i)\hat{P}(y, y-1_i),
\end{aligned}$$

which concludes the proof. □

Proof of Lemma 4. Iterating Lemma 6 readily completes the proof of Lemma 5. □

For $i, j \in \mathcal{C}$, let $\tilde{\rho}(i, j) = \mathcal{E}_\rho(i, j)$, and let $\tilde{\mathrm{d}}$ be the associated generator. In this dynamic, if the chosen class has nodes to match, the match is performed with probability 1. The next step is to compare the generators $\hat{\mathrm{d}}$ and $\tilde{\mathrm{d}}$.

Lemma 6. (Comparison between $\hat{\mathrm{d}}$ and $\tilde{\mathrm{d}}$.) *If $x \in \mathscr{X}$ is such that $\mathcal{S}_x \subset \mathcal{C}_\rho^0$, $x(i) \ge n^*$ for every $i \in \mathcal{S}_x$, and $\mathcal{S}_x$ is a $\mathcal{G}_\rho$-independent set, then $\hat{\mathrm{d}}q(x) \le \tilde{\mathrm{d}}q(x) + 2K$.*

Proof. Let $x \in \mathscr{X}$ be such that $\mathcal{S}_x \subset \mathcal{C}_\rho^0$, $x(i) \ge n^*$ for $i \in \mathcal{S}_x$, and such that $\mathcal{S}_x$ is a $\mathcal{G}_\rho$-independent set. Then

$$\hat{\mathrm{d}}q(x) = 1 + \sum_{i\in\mathcal{S}_x} 2x(i)\hat{P}(x, x+1_i) - \sum_{i\in\mathcal{S}_x} 2x(i)\hat{P}(x, x-1_i),$$

and the analogous expression for $\tilde{\mathrm{d}}q(x)$. On the one hand, since $\mathcal{S}_x$ is an independent set, we have

$$\sum_{i\in\mathcal{S}_x} 2x(i)\hat{P}(x, x+\eta_i) = \sum_{i\in\mathcal{S}_x} 2x(i)\tilde{P}(x, x+\eta_i). \tag{20}$$

On the other, for $n \ge n^*$, Lemma 2 gives

$$\begin{aligned}
&\sum_{i\in\mathcal{S}_x:x(i)=n} 2x(i)\tilde{P}(x, x-\eta_i) - \sum_{i\in\mathcal{S}_x:x(i)=n} 2x(i)\hat{P}(x, x-\eta_i) \\
&\quad = \sum_{i\in\mathcal{S}_x:x(i)=n} 2n \sum_{j:\tilde{\phi}(x,j)=i} \nu(j) - \sum_{i\in\mathcal{S}_x:x(i)=n} 2n \sum_{j:\hat{\phi}(x,j)=i} \nu(j)(1-(1-\hat{\rho}(j,i))^n) \\
&\quad = 2n \sum_{j:\|x\|_{\mathcal{N}_\rho(j)}=n} \nu(j) - 2n \sum_{j:\|x\|_{\mathcal{N}_\rho(j)}=n} \nu(j)(1-(1-r^*)^n) \\
&\quad = 2n(1-r^*)^n \nu\{j\colon \|x\|_{\mathcal{N}_\rho(j)} = n\} \le 2K\nu\{j\colon \|x\|_{\mathcal{N}_\rho(j)} = n\}.
\end{aligned}$$

https://doi.org/10.1017/jpr.2025.22 Published online by Cambridge University Press

Summing over $n \geq n^*$, we get

$$\sum_{i \in \mathcal{S}_x} 2x(i)\tilde{P}(x, x - \eta_i) - \sum_{i \in \mathcal{S}_x} 2x(i)\hat{P}(x, x - \eta_i) \leq 2K. \tag{21}$$

We conclude by gathering (20) and (21). □

The following result is the last one we need. Its proof was given in [8] in a more general context. For the benefit of the reader, we present here a proof adapted to this case.

Lemma 7. (Control of $\tilde{d}$.) *Assume that* $(\mathcal{G}_\rho, \nu)$ *satisfies* NCOND. *If* $x \in \mathscr{X}$ *is such that* $\mathcal{S}_x \subset \mathcal{C}_\rho^0$ *and it is a* $\mathcal{G}_\rho$*-independent set, then* $\tilde{\mathrm{d}}q(x) \leq 1 - 2\eta\|x\|$.

Proof. Let $x \in \mathscr{X}$ be such that $\mathcal{S}_x$ is a $\mathcal{G}_\rho$-independent set that is included in $\mathcal{C}_\rho^0$. In this case, $\tilde{\mathrm{d}}q(x)$ can we written as

$$\tilde{\mathrm{d}}q(x) = 1 + \sum_{i \in \mathcal{S}_x} 2x(i)[\nu(i) - \tilde{P}(x, x - 1_i)].$$

Since the case where $x = 0$ leads to a trivial result, we proceed by assuming that $x \neq 0$. Let $\{a_1, \ldots, a_L\}$ be the image of the function

$$\begin{cases} \mathcal{S}_x & \to & \mathbb{N}_0, \\ i & \mapsto & x(i), \end{cases}$$

assuming $a_l > a_{l+1}$ for every $l \in \{1, \ldots, L-1\}$. For any $l \in \{1, \ldots, L\}$ we let $\mathcal{H}_l = \{i \in \mathcal{S}_x : x(i) = a_l\}$. We can then write

$$\sum_{i \in \mathcal{S}_x} x(i)[\nu(i) - \tilde{P}(x, x - 1_i)] = \sum_{i \in \mathcal{S}_x} x(i) r_i = \sum_{l=1}^{L} a_l \sum_{i \in \mathcal{H}_l} r_i = \sum_{l=1}^{L} a_l R_l,$$

where we have defined $r_i = \nu(i) - \tilde{P}(x, x - 1_i)$ and $R_l = \sum_{i \in \mathcal{H}_l} r_i$.

If $R_l \leq 0$ for every $l \in \{1, \ldots, L\}$, we are done since

$$\sum_{l=1}^{L} a_l R_l \leq a_1 R_1 = \|x\| \left[\nu(\mathcal{H}_1) - \sum_{i \in \mathcal{H}_1} \tilde{P}(x, x - 1_i) \right] = \|x\| [\nu(\mathcal{H}_1) - \nu(\mathcal{N}_\rho(\mathcal{H}_1))] \leq -\|x\|\eta.$$

If not, we define $L_1 = \max\{l \in \{1, \ldots, L\} : R_l > 0\}$. Then

$$\sum_{l=1}^{L} a_l R_l \leq \sum_{l=1}^{L_1} a_l R_l \leq \sum_{l=1}^{L_1 - 2} a_l R_l + a_{L_1 - 1}(R_{L_1 - 1} + R_{L_1}). \tag{22}$$

If $R_{L_1-1} + R_{L_1} \leq 0$, we control the last quantity by $\sum_{l=1}^{L_1-2} a_l R_l$ and restart the procedure. If, otherwise, $R_{L_1-1} + R_{L_1} > 0$, we control the right-hand side of (22) by

$$\sum_{l=1}^{L_1 - 3} a_l R_l + a_{L_1 - 2}(R_{L_1 - 2} + R_{L_1 - 1} + R_{L_1}).$$

https://doi.org/10.1017/jpr.2025.22 Published online by Cambridge University Press

We again separate the cases $R_{L_1-2}+R_{L_1-1}+R_{L_1}\le 0$ and $R_{L_1-2}+R_{L_1-1}+R_{L_1}>0$, and proceed as before. Finally, if we reach $\mathcal{H}_1$ with this procedure, we have

$$\begin{aligned} a_1\sum_{l=1}^{L_1} R_l &= \|x\|\left[\nu(\mathcal{H}_1\cup\cdots\cup\mathcal{H}_{L_1}) - \sum_{i\in\mathcal{H}_1\cup\cdots\cup\mathcal{H}_{L_1}} \tilde{P}(x, x-1_i)\right] \\ &= \|x\|[\nu(\mathcal{H}_1\cup\cdots\cup\mathcal{H}_{L_1}) - \nu(\mathcal{N}_\rho(\mathcal{H}_1\cup\cdots\cup\mathcal{H}_{L_1}))] \le -\eta\|x\|, \end{aligned}$$

which concludes the proof. □

We are now in a position to prove the fundamental drift inequality in a few lines.

Proof of Proposition 3. Fix $x\in\mathscr{X}$. Applying Lemmas 1, 3, 4, 6, and 7, we get

$$\begin{aligned} \mathrm{d}q(x) \le \hat{\mathrm{d}}q(\chi(x)) + 2n^* &\le \hat{\mathrm{d}}q\big(\chi(x)1_{\mathcal{C}_\rho^0}\big) + 4K|\mathcal{C}_\rho^+| - \sum_{i\in\mathcal{C}_\rho^+} 2\chi(x)(i)\nu(i) + 2n^* \\ &\le \hat{\mathrm{d}}q(y) + 2K + 4K|\mathcal{C}_\rho^+| - \sum_{i\in\mathcal{C}_\rho^+} 2\chi(x)(i)\nu(i) + 2n^* \\ &\le \tilde{\mathrm{d}}q(y) + 4K + 4K|\mathcal{C}_\rho^+| - \sum_{i\in\mathcal{C}_\rho^+} 2\chi(x)(i)\nu(i) + 2n^* \\ &le 1 - 2\eta\|y\| + 4K + 4K|\mathcal{C}_\rho^+| - \sum_{i\in\mathcal{C}_\rho^+} 2\chi(x)(i)\nu(i) + 2n^*, \end{aligned}$$

where $y\in\mathscr{X}$ is such that $\|y\| = \|\chi(x)1_{\mathcal{C}_\rho^0}\| = \|x\|_{\mathcal{C}_\rho^0}$ and $\mathcal{S}_y\in\mathbb{I}_\rho$. The proof is completed by observing that the last expression coincides with the right-hand side of (14). □

4.2. Stability

As already mentioned, that NCOND is necessary for the stability of the DTMC follows immediately from Proposition 1. We now prove that NCOND implies stability. We say that a function $f\colon \mathscr{X}\to[0,\infty)$ is a *Lyapunov function* if there exist $\varepsilon>0$ and a finite set $\mathscr{F}\subset\mathscr{X}$ such that

$$Pf(x) = \sum_{y\in\mathscr{X}} P(x,y)f(y) < \infty \quad \text{for every } x\in\mathscr{F}, \tag{23}$$

$$\mathrm{d}f(x) \le -\varepsilon \qquad \text{for every } x\in\mathscr{X}\setminus\mathscr{F}. \tag{24}$$

The Foster–Lyapunov theorem (see, e.g., [5, Theorem 7.1.1]) guarantees stability if a Lyapunov function is found. In our case, stability will be obtained by proving, as a direct consequence of Proposition 2, that the quadratic function q is Lyapunov. Since the set $\{y\in\mathscr{X}\colon P(x,y)\neq 0\}$ is finite for every $x\in\mathscr{X}$, condition (23) is fulfilled automatically. The real challenge is condition (24), which follows, for instance, by choosing $\varepsilon=1$ and

$$\mathscr{F} = \left\{x\in\mathscr{X}\colon \|x\|_{\mathcal{C}_\rho^0} \le \frac{1+n^*+2K(1+|\mathcal{C}_\rho^+|)}{\eta},\ \|x\|_{\mathcal{C}_\rho^+} \le \frac{2+2n^*+4K(1+|\mathcal{C}_\rho^+|)}{\min_{i\in\mathcal{C}_\rho^+}\nu(i)}\right\}$$

(the minimum over an empty set being defined as $+\infty$), and by using inequality (14).

https://doi.org/10.1017/jpr.2025.22 Published online by Cambridge University Press

4.3. Perfect matching infinitely often

The fact that the online matching is perfect i.o. almost surely for $\eta > 0$ follows readily from (10). Now assume that $\eta < 0$, and let $\mathcal{I} \in \mathbb{I}_\rho$ attain the minimum, which is equivalent to saying that $\nu(\mathcal{I}) - \nu(\mathcal{N}_\rho(\mathcal{I})) = -\eta > 0$. In this case, the associated lazy random walk $\{S_t^{\mathcal{I}}\}_{t\in\mathbb{N}_0}$ defined by (4) is transient to infinity, since its drift is precisely $\nu(\mathcal{I}) - \nu(\mathcal{N}_\rho(\mathcal{I}))$. On the other hand, similarly to (5), it is clear that for all t, the number of unmatched $\mathcal{I}$-nodes at time t is a.s. lower-bounded by $S_t^{\mathcal{I}}$. So the number of unmatched $\mathcal{I}$-nodes also grows to infinity with t almost surely, proving in turn that, almost surely, there is no perfect matching i.o. This concludes the proof of Theorem 1.

5. Proof of Corollary 2

For any $x \in \mathscr{X}$ we have

$$\eta\|x\|_{\mathcal{C}_\rho^0} + \sum_{i\in\mathcal{C}_\rho^+ \,:\, x(i)\geq n^*} x(i)\nu(i) \geq (\|x\| - n^*)\left(\eta \wedge \min_{i\in\mathcal{C}_\rho^+} \nu(i)\right). \tag{25}$$

To see why this it true, we separate the cases $\|x\| = \|x\|_{\mathcal{C}_\rho^0}$, $\|x\| = \|x\|_{\mathcal{C}_\rho^+} \geq n^*$, and $\|x\| = \|x\|_{\mathcal{C}_\rho^+} < n^*$. Now we use inequality (25) in (14) to get that

$$\mathrm{d}q(x) \leq -2\|x\|\left(\eta \wedge \min_{i\in\mathcal{C}_\rho^+} \nu(i)\right) + 1 + 2n^*\left[1 + \left(\eta \wedge \min_{i\in\mathcal{C}_\rho^+} \nu(i)\right)\right] + 4K(1 + |\mathcal{C}_\rho^+|).$$

We can use [13, Theorem 14.3.7] to obtain

$$\int \|x\|\,\pi(\mathrm{d}x) \leq \frac{1 + 2n^*[1 + (\eta \wedge \min_{i\in\mathcal{C}_\rho^+} \nu(i))] + 4K(1 + |\mathcal{C}_\rho^+|)}{2(\eta \wedge \min_{i\in\mathcal{C}_\rho^+} \nu(i))}.$$

Once the finiteness of the invariant expectation has been obtained, we can apply [5, Theorem 3.3.2] to obtain the convergence of the ergodic averages. Finally, (12) follows because

$$0 \leq \frac{\|X_t\|}{t} = \frac{1}{t}\sum_{s=1}^{t} \|X_s\| - \frac{t-1}{t}\frac{1}{t-1}\sum_{s=1}^{t-1} \|X_s\| \xrightarrow[t\to\infty]{} 0 \quad \text{a.s.}$$

Funding information

There are no funding bodies to thank relating to the creation of this article.

Competing interests

There were no competing interests to declare which arose during the preparation or publication process of this article.

References

[1] Aoudi, M. H. A. D., Moyal, P. and Robin, V. (2022). Markovian online matching algorithms on large bipartite random graphs. *Methodology Comput. Appl. Prob.* **24**, 3195–3225.

[2] Aronson, J., Frieze, A. M. and Pittel, B. G. (1998). Maximum matchings in sparse random graphs: Karp–Sipser revisited. *Random Structures Algorithms* **12**, 111–177.

https://doi.org/10.1017/jpr.2025.22 Published online by Cambridge University Press

[3] BEGEOT, J., MARCOVICI, I., MOYAL, P. AND RAHME, Y. (2021). A general stochastic matching model on multigraphs. *ALEA Lat. Am. J. Prob. Math. Stat.* **18**, 1325–1351.
[4] BORDENAVE, C., LELARGE, M. AND SALEZ, J. (2013). Matchings on infinite graphs. *Prob. Theory Relat. Fields* **157**, 183–208.
[5] BRÉMAUD, P. (2020). *Markov Chains – Gibbs Fields, Monte Carlo Simulation and Queues* (Texts Appl. Math. **31**), 2nd edn. Springer, Cham.
[6] CADAS, A., BUŠIĆ, A. AND DONCEL, J. (2019). Optimal control of dynamic bipartite matching models. In *Proc. 12th EAI Int, Conf. Performance Evaluation Methodologies and Tools*, pp. 39–46.
[7] COHEN, I. R. AND WAJC, D. (2018). Randomized online matching in regular graphs. In *Proc. 29th Ann. ACM-SIAM Symp. Discrete Algorithms*. Society for Industrial and Applied Mathematics, Philadelphia, PA, pp. 960–979.
[8] JONCKHEERE, M., MOYAL, P., RAMREZ, C. AND SOPRANO-LOTO, N. (2023). Generalized max-weight policies in stochastic matching. *Stoch. Systems* **13**, 40–58.
[9] KARP, R. M. AND SIPSER, M. (1981). Maximum matching in sparse random graphs. In *Proc. 22nd Ann. Symp. Foundations of Computer Science*, pp. 364–375.
[10] KARP, R. M., VAZIRANI, U. V. AND VAZIRANI, V. V. (1990). An optimal algorithm for on-line bipartite matching. In *Proc. 22nd Ann. ACM Symp. Theory of Computing*. Association for Computing Machinery, New York, pp. 352–358.
[11] KERIMOV, S., ASHLAGI, I. AND GURVICH, I. (2021). Dynamic matching: Characterizing and achieving constant regret. *SSRN Electron. J.* 3824407.
[12] MAIRESSE, J. AND MOYAL, P. (2016). Stability of the stochastic matching model. *J. Appl. Prob.* **53**, 1064–1077.
[13] MEYN, S. AND TWEEDIE, R. L. (2009). *Markov Chains and Stochastic Stability*, 2nd edn. Cambridge University Press.
[14] MOYAL, P. AND PERRY, O. (2017). On the instability of matching queues. *Ann. Appl. Prob.* 27, 3385–3434.
[15] NAZARI, M. AND STOLYAR, A. L. (2019). Reward maximization in general dynamic matching systems. *Queueing Systems* **91**, 143–170.
[16] NOIRY, N., PERCHET, V. AND SENTENAC, F. (2021). Online matching in sparse random graphs: Non-asymptotic performances of greedy algorithm. In *Proc. 35th Int. Conf. Neural Information Processing Systems*, ed. A. Beygelzimer, Y. Dauphin, P. Liang and J. Wortman Vaughan. Association for Computing Machinery, New York, pp. 21400–21412.
[17] Schrijver, A. (2003). *Combinatorial Optimization. Polyhedra and Efficiency*, vol. A (Algorithms and Combinatorics **24**). Springer, Berlin.
[18] SENTENAC, F., NOIRY, N., LERASLE, M., MÉNARD, L. AND PERCHET, V. (2023). Online matching in geometric random graphs. Preprint, arXiv:2306.07891.
[19] TUTTE, W. T. (1947). The factorization of linear graphs. *J. London Math. Soc.* **22**, 107–111.

https://doi.org/10.1017/jpr.2025.22 Published online by Cambridge University Press