

Availability is all you need: achieving optimal regret with minimal information for dynamic matching

Süleyman Kerimov, Pengyu Qian, Mingwei Yang, and Sophie H. Yu*

February 18, 2026

Abstract

We study a centralized discrete-time dynamic two-way matching model with finitely many agent types. Agents arrive stochastically over time and join their type-dedicated queues waiting to be matched. We focus on *availability-based* policies that make matching decisions based solely on agent availability across types (i.e., whether queues are empty or not), rather than relying on complete queue-length information (e.g., the longest-queue policy). We aim to achieve constant regret at all times with optimal scaling in terms of the general position gap, ϵ , which measures the distance of the fluid relaxation from degeneracy.

We classify availability-based policies into *global* and *local* policies based on the scope of information they utilize. First, for general networks (possibly cyclic), we propose a global availability-based policy, *probabilistic matching* (**PM**), and prove that it achieves the optimal all-time regret scaling of $O(\epsilon^{-1})$, matching the known lower bound established by [KAG24]. Second, for acyclic networks, we focus on the class of local availability-based policies, specifically static priority policies that prioritize matches based on a fixed order. Within this class, we derive the first explicit regret bound for the previously proposed *tree priority* policy (**TP**), showing all-time regret scaling of $O(\epsilon^{-(d+1)/2})$, where d is the network depth. Next, we introduce a new *truncated tree priority* (**TTP**) policy and prove that it is the first static priority policy to achieve the optimal all-time regret scaling of $O(\epsilon^{-1})$. These policies are appealing for matching systems such as queueing and load balancing; they reduce operational costs by using minimal information while effectively balancing the trade-off between immediate and future rewards.

Contents

1	Introduction	3
1.1	Our contributions	4
1.2	Related literature	5
1.3	Notation	6
2	Model	6
2.1	Optimality criterion	7
2.2	Static-planning and general position gap	7
2.3	Policy information classes	9

*S. Kerimov is with the Jones Graduate School of Business, Rice University, Houston TX, USA, kerimov@rice.edu. P. Qian is with the Questrom School of Business, Boston University, Boston MA, USA, pqian@bu.edu. M. Yang is with the Department of Management Science and Engineering, Stanford University, Stanford CA, USA, mwyang@stanford.edu. S. H. Yu is with The Wharton School of Business, University of Pennsylvania, Philadelphia PA, USA, hysophie@wharton.upenn.edu.

3	Main results	9
3.1	Global availability-based policy	10
3.2	Local availability-based policy	11
4	Analysis	12
4.1	Analysis of probabilistic matching policy	13
4.2	Analysis of tree priority policy on acyclic graphs	15
4.3	Analysis of truncated tree priority policy on acyclic graphs	16
5	Numerical experiments	19
6	Conclusion	19
A	Preliminaries	25
A.1	Martingale and concentration inequalities	25
A.2	Lyapunov function analysis	27
A.3	Markov chain	28
A.4	Consistent policy	28
A.4.1	Proof of Lemma 2	30
B	Discussion on tree priority policy on acyclic graphs	31
B.1	Tightness of regret analysis on paths	32
B.1.1	Proof of Claim 1	35
B.1.2	Proof of Claim 2	36
B.2	Towards optimal regret scaling	36
C	Deferred proofs	39
C.1	Proof of Lemma 1	39
C.2	Deferred proofs for Section 4.2	39
C.2.1	Proof of Lemma 4	39
C.2.2	Proof of Proposition 2	40
C.3	Deferred proofs for Section 4.3	45
C.3.1	Proof of Proposition 3	45
C.3.2	Proof of Proposition 4	46
C.3.3	Proof of Proposition 5	49
C.3.4	Proof of Proposition 6	52
C.3.5	Proof of Theorem 3	54

1 Introduction

Dynamic matching markets, where agents arrive over time and must be matched based on compatibility, arise in numerous applications including kidney exchange markets (e.g., [ABB⁺18]), on-line carpooling and ride-hailing platforms (e.g., [ÖW20]), and logistics and delivery systems (e.g., [ET25]). In such settings, agents of different types arrive stochastically to type-specific queues, waiting to be matched. The allowable matches are captured by an underlying compatibility graph, where edge weights encode edge-specific matching rewards. A central design problem is to construct online greedy policies that perform well at any time, i.e., balancing immediate and future rewards, while remaining simple enough to implement in large, distributed systems.

A substantial body of recent literature analyzes dynamic matching policies relative to an offline omniscient benchmark, with a particular focus on the *general position gap* (GPG) condition. This condition is generically satisfied when the optimal basic matches (those corresponding to a basic optimal solution of the fluid relaxation) remain stable under small perturbations of the model parameters [MC89, JK12, CLY24]. The GPG parameter ϵ quantifies the magnitude of such allowable perturbations, effectively measuring the distance of the fluid relaxation from degeneracy [WXY23, KAG24, Gup24, KAG25]. Intuitively, ϵ captures both the inherent “thickness” of the market and its operational stability: a larger ϵ implies a wider separation between optimal and suboptimal actions, making it easier for the central planner to consistently perform “correct” matches. Understanding how a policy’s all-time regret scales with ϵ is therefore crucial, as this relationship reveals the robustness of the policy across diverse operating environments.

Under the GPG condition, the optimal all-time regret, defined as the maximum expected difference between the performances of the offline optimum and the policy at any moment, is known to scale as $\Theta(\epsilon^{-1})$ [KAG24]. To achieve this optimal scaling, state-of-the-art policies such as longest-queue [KAG25], sum-of-squares [Gup24], and primal-dual [WXY23] typically rely on sophisticated, information-intensive rules that require real-time access to exact queue-length information (see Figure 1). While the longest-queue policy achieves this optimal scaling using local queue lengths, and others do so using global queue lengths, it remains an open question whether policies using substantially less information can match this fundamental $O(\epsilon^{-1})$ limit.

This question is not merely of theoretical interest; it is driven by the operational costs and vulnerabilities inherent in maintaining fine-grained state information. In many real-world systems, maintaining and communicating such detailed state information is costly, fragile, or even infeasible. In large-scale load balancing, for instance, policies requiring instantaneous queue lengths incur prohibitive communication overheads compared to those using partial information [Mit01, L XK⁺11, dBBVLM22]. Furthermore, policies sensitive to exact queue lengths are vulnerable to manipulation by strategic agents [ES07, ERIZ25] and are less robust to measurement noise or delayed updates [Ibr18]. These practical limitations motivate the search for *availability-based* policies that make decisions based solely on whether queues are empty or not, rather than how long they are. Indeed, availability information is the minimal information requirement for executing a match, forming the baseline for state granularity.

To systematically analyze the information requirements for policies achieving the optimal all-time regret scaling, we classify policies along two dimensions: granularity and scope. The first dimension distinguishes coarse availability information (binary empty/non-empty status) from fine-grained queue-length information (exact counts). The second dimension distinguishes *local* information (restricted to the neighbors of the arriving agent) from *global* information (encompassing the state of the entire network). Motivated by this classification, our work investigates the sufficiency of minimal-information policies, asking whether coarse, availability-based rules can achieve the optimal all-time regret scaling as their information-intensive counterparts.

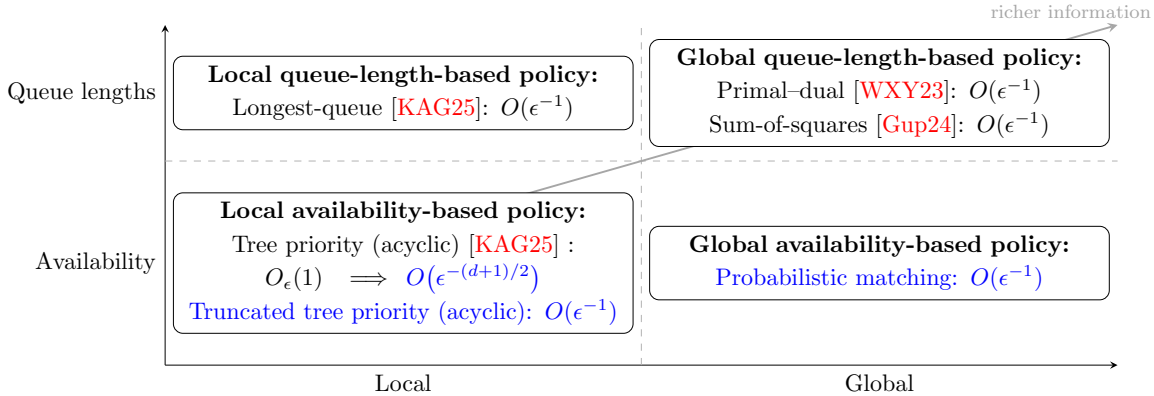


Figure 1: Information landscape and regret scaling of dynamic matching policies, with the results in this paper highlighted in blue. The policies are positioned according to the granularity (availability vs. queue lengths) and scope (local vs. global) of the state information they use, along with their all-time regret scaling in terms of the GPG parameter ϵ , where $O_\epsilon(1)$ refers to a constant bound with an unknown dependence on ϵ , and d denote the depth of the acyclic graph.

1.1 Our contributions

In this paper, we provide a comprehensive analysis of availability-based greedy policies in two-way matching networks under the GPG condition. Throughout our analysis, we focus on optimal basic matches. This restriction is justified by the GPG condition, which guarantees that optimal basic matches remain stable under small perturbations of the model parameters. Our results, summarized in Figure 1, demonstrate that restricting attention to such availability-based policies need not sacrifice performance.

For general networks (including those with cycles), we introduce the *probabilistic matching (PM)* policy, which is a *global availability-based policy*, by carefully adapting the probability of performing each match based on the state of the system. We prove that this policy achieves the optimal regret scaling of $O(\epsilon^{-1})$ (Theorem 1). This result establishes that fine-grained queue-length information is not strictly necessary for optimality; coarse availability information suffices, provided one has a global view of the network.

Next, we turn to acyclic networks, where the reduced network formed by the optimal basic matches contains no cycles. Central to this setting are under-demanded agent types, defined as those for which the static planning problem (and consequently any “reasonable” policy) matches only a fraction of their arrivals. Crucially, every connected component of such a reduced network possesses a unique under-demanded node [KAG25]. We utilize this unique node as the root to orient the connected component as a rooted tree, capturing the intuition that nodes further from the root are increasingly demanded. This structural perspective is pivotal as it allows us to focus on a family of *local availability-based policies*, specifically *static priority policies*. These policies match agents according to a fixed ordering and have been widely studied in the literature [MP17, ADSW25, KAG25]. In our setting, we prioritize children over the parent node, meaning that we prioritize the more demanded nodes over the less demanded ones.

First, we revisit the *tree priority (TP)* policy proposed by [KAG25]. Leveraging the rooted tree structure defined above, **TP** prioritizes matches that are farther away from the under-demanded root (i.e., prioritizing “children” over “parents”). While **TP** is known to admit bounded all-time regret, the scaling of the regret in terms of ϵ was previously unknown. By heavily exploiting the hierarchical structure of the tree, we derive the first explicit regret bound for **TP**, showing it scales

as $O(\epsilon^{-(d+1)/2})$, where d is the network depth (Theorem 2).

Next, we propose the *truncated tree priority* (**TTP**) policy, which is the first static priority policy proven to achieve the optimal $O(\epsilon^{-1})$ regret scaling (Theorem 3). While **TTP** shares the same prioritization strategy as **TP** to efficiently clear over-demanded queues, the critical distinction lies in the fallback mechanism: upon the arrival of an agent, **TTP** strictly forbids matching with its “parent” node, whereas **TP** allows it as a last resort. This constraint effectively decouples the system, ensuring one-way propagation of stochasticity. This property enables us to employ a careful multi-step drift analysis, a novel and necessary technical departure from the standard one-step drift techniques used in prior work [WXY23, Gup24, KAG25], to establish the optimal queue-length bound. Remarkably, this result implies that for acyclic graphs, optimality is achievable using the minimal possible information set: the binary availability of immediate neighbors without requiring either global visibility or exact queue counts.

1.2 Related literature

Dynamic matching has been extensively studied in various settings. We review several streams of literature most relevant to our work.

Multi-way dynamic matching. Starting from [KAG24], the multi-way dynamic matching problem has recently received extensive attention, and the optimal regret scaling of $O(\epsilon^{-1})$ has been achieved via several policies. [KAG24] propose a batching policy, which performs a maximum weighted matching periodically. [Gup24] develops a sum-of-square policy that achieves an optimal regret scaling even under a more general model. Later on, [WXY23] design a policy based on the primal-dual framework, which can achieve the optimal regret scaling even for unknown arrival rates. [KAG25] consider two-way matching networks and focus on policies that perform a match whenever possible. They show that the longest-queue policy achieves the optimal regret scaling, and they propose the **TP** policy that achieves constant regret at all times only for acyclic matching networks without explicitly characterizing the regret.

Dynamic matching on fixed graphs. Similar to our work, numerous papers study the dynamic matching problem with a fixed matching configuration by imposing different modeling assumptions. Most literature assumes that each match consists of exactly two agents, where the underlying matching network can be either bipartite [ÖW20, KSSW22, ADSW25, CHS24] or non-bipartite [CILB⁺20, AS22]. A common feature of these literature is that they all assume that agents depart stochastically, and different objectives are considered, such as minimizing the holding costs [BM15], maximizing the match-specific rewards [CILB⁺20], or maximizing the generated rewards minus the holding costs [ADSW25].

Dynamic matching on random graphs. There is a vast literature on dynamic matching models based on random graphs, where agents arrive over time and can possibly form edges with the existing agents with fixed probabilities [AAGK17, ABJM19, ALG20]. A common assumption in this literature is that match values are homogeneous so that the focus is on minimizing the number of unmatched agents, and the general finding of this literature is that acting greedily is asymptotically optimal. Recently, [BRS⁺22] consider a setting where match values are heterogeneous, and as the market grows large, they show that greedy threshold policies are asymptotically optimal. Another line of research assumes that the agents lie in a metric space, and the weight of an edge between two agents is determined by the distance between them [Kan21, BFP23, YY26, LUY26].

Network revenue management. In the network revenue management (NRM) problem, there exist offline resources, and online requests that consume certain amounts of offline resources arrive dynamically. [TVR98] achieve $O(\sqrt{T})$ regret in the (quantity-based) NRM model via the bid-price policy. Later on, [JK12] improve the regret to be $O(1)$ via a re-solving policy under the GPG condition. Since then, constant regret is achieved when the arrival rates are unknown [Jas15], when the GPG condition does not hold [AG19, VB19, BW20, VBG21, JMZ22], and for a variety of other related problems [BBP24]. In contrast, for the multi-way dynamic matching problem considered in this paper, without the GPG condition, a regret lower bound of $\Omega(\sqrt{T})$ exists [KAG24].

Stability of stochastic matching systems. The multi-way dynamic matching model is closely related to the study of stability of stochastic matching systems [MM16, RM21, JMRS23]. The connection is established by, e.g., [KAG24, Lemma 4.1] and [Gup24, Lemma 1], which assert that bounded all-time regret is implied by the stability of queues that are fully utilized by the fluid relaxation, provided that only the matches active in the fluid relaxation are performed. For the special case of two-way matching systems, [MM16] identify sufficient and necessary conditions for stability, and [JMRS23] show that several policies, including the longest-queue policy and a generalized max-weight policy, achieve the maximal stability region.

1.3 Notation

For $x, y \in \mathbb{R}$, we use $x \wedge y$ to denote $\min\{x, y\}$ and x^+ to denote $\max\{x, 0\}$. For $n \in \mathbb{Z}_{>0}$, we use $[n]$ to denote the set $\{1, \dots, n\}$. For a vector $v \in \mathbb{R}^n$ and an index set $J \subset [n]$, we use v^+ to denote (v_1^+, \dots, v_n^+) and use $v_J \in \mathbb{R}^{|J|}$ to denote the vector obtained by restricting v on J . We use \mathbf{e}_i to denote the i -th standard basis, i.e., the i -th entry of \mathbf{e}_i is 1 with all other entries being 0. We use standard asymptotic notation: for two positive sequences $\{x_n\}$ and $\{y_n\}$, we write $x_n = O(y_n)$ if $x_n \leq Cy_n$ for an absolute constant C and for all n ; $x_n = \Omega(y_n)$ if $y_n = O(x_n)$.

For a rooted tree, let $\mathcal{C}(i)$ be the set of children of a node i , and let $P(i)$ be the parent of a non-root node i . For each node i , denote $\mathcal{T}(i)$ as the set of nodes in the subtree rooted at i (including i), and denote $\mathcal{T}^-(i) \triangleq \mathcal{T}(i) \setminus \{i\}$.

2 Model

We study a centralized dynamic matching market. There are n types of agents $\mathcal{A} = [n]$ and k types of matches $\mathcal{M} = [k]$. For each $m \in \mathcal{M}$, the value of performing a match m is denoted by $r_m > 0$. Following [KAG25], we focus on two-way matching structures, i.e., each match $m \in \mathcal{M}$ consists of exactly two agent types. We encode the matching structure into a *matching matrix* $M \in \{0, 1\}^{\mathcal{A} \times \mathcal{M}}$, where $M_{im} = 1$ if and only if match m involves agent type i for all $i \in \mathcal{A}$ and $m \in \mathcal{M}$. Assume without loss of generality that each agent type participates in at least one match. We denote $(\mathcal{A}, \mathcal{M})$ as the (undirected) graph where vertices are formed by agent types and edges are formed by matches. For each agent type $i \in \mathcal{A}$, denote $\mathcal{N}(i)$ as the set of neighbors of i in the graph $(\mathcal{A}, \mathcal{M})$. If there is a match in \mathcal{M} that contains agent types i and j , we will use $m(i, j)$ to denote this match.

We consider discrete-time arrivals where exactly one agent arrives and joins the type-dedicated queue at each time $t \in [T]$. For simplicity, we assume that there are no preexisting agents in the system.¹ Let $\lambda \in \mathbb{R}_{>0}^n$ denote a probability distribution over agent types, where λ_i represents the probability of an arriving agent being type i , and $\sum_{i=1}^n \lambda_i = 1$. For each type i , let $A_i(t)$ denote the

¹This is a common assumption made by the literature [WXY23, KAG25].

cumulative number of type i arrivals up to and including time t , with $\Delta A_i(t) \triangleq A_i(t) - A_i(t-1) \in \{0, 1\}$ indicating whether an agent of type i arrives at time $t > 0$. At each period t , the central decision maker can perform match $m \in \mathcal{M}$ only if there are waiting agents of each type contained in m . This match then generates a reward r_m , and the agents participating in the match depart. For the sequence of events within each period, an agent first arrives, and then the decision maker decides which matches to perform. We refer to the tuple $\mathcal{G} = (M, \lambda, r)$ as the *matching network*, which captures both the matching structure and the arrival process.

A (randomized) *matching policy* decides how to (randomly) perform matches at each period. We will restrict our attention to non-anticipative policies, whose decisions at each moment only depend on the events happened so far. Given a policy, for all match $m \in \mathcal{M}$ and time $t \geq 0$, we denote $D_m(t)$ as the number of match m performed by the policy up to and including time t , and denote $\Delta D_m(t) \triangleq D_m(t) - D_m(t-1)$ as the number of match m performed at time $t > 0$. We maintain a queue for each type $i \in \mathcal{A}$, which contains all agents of type i in the system that have not been matched. Define $Q_i(t) \triangleq (A(t) - MD(t))_i$ as the length of each queue $i \in \mathcal{A}$ after time $t \geq 0$, and we refer to $Q(t)$ as the *state* of the system after time t . We say that a state is *valid* under a policy if it is reachable by the policy from the all-zero state.

2.1 Optimality criterion

The expected total value generated by a policy Π during the first t periods is denoted by $\mathcal{R}^\Pi(t) \triangleq \mathbb{E}[r^T D(t)]$. Let $\mathcal{R}^*(t)$ be the expected value attained by the policy that takes no action before time t and performs matches that maximize the overall rewards up to time t . Formally,

$$\mathcal{R}^*(t) \triangleq \mathbb{E} \left[\max_{\substack{\text{s.t.} \\ My \leq A(t) \\ y \in \mathbb{Z}_{\geq 0}^k}} r^T y \right], \quad (1)$$

and $\mathcal{R}^*(t)$ is straightforwardly an upper bound for $\mathcal{R}^\Pi(t)$ for every policy Π . Define the *regret* of a policy Π after time t as $\mathcal{R}^*(t) - \mathcal{R}^\Pi(t)$, and define the *all-time regret* of Π after time T as

$$\text{Regret}(\Pi, T) \triangleq \sup_{0 \leq t \leq T} (\mathcal{R}^*(t) - \mathcal{R}^\Pi(t)).$$

We say that Π achieves constant regret at all times (or all-time constant regret), if $\text{Regret}(\Pi, T)$ is upper bounded by a constant that does not depend on T . Note that this performance metric differs from simply achieving exact hindsight optimality at the end of time horizon T , which can be attained by a trivial policy that delays all matches until the end of the time horizon T and then solves an optimization problem to maximize overall rewards. In other words, to achieve all-time constant regret, a policy must effectively balance short-term and long-term rewards, ensuring that decisions made at each step do not compromise overall performance.

2.2 Static-planning and general position gap

We consider the fractional relaxation of the integer program in (1). By Jensen's inequality,

$$\mathcal{R}^*(t) = \mathbb{E} \left[\max_{\substack{\text{s.t.} \\ My \leq A(t) \\ y \in \mathbb{Z}_{\geq 0}^k}} r^T y \right] \leq \max_{\substack{\text{s.t.} \\ Mx \leq t\lambda \\ x \in \mathbb{R}_{\geq 0}^k}} r^T x. \quad (2)$$

Replacing $z = x/t$ and adding slack variables $(s_i)_{i \in \mathcal{A}}$, the linear program in the RHS can be rewritten as

$$\text{SPP}(\lambda) \triangleq \begin{array}{ll} \max & r^T z \\ \text{s.t.} & Mz + s = \lambda \\ & z \in \mathbb{R}_{\geq 0}^k, s \in \mathbb{R}_{\geq 0}^n \end{array},$$

which we refer to as the *static-planning problem*.

Next, we introduce the notion of general position gap that captures the stability level of $\text{SPP}(\lambda)$. This condition is proved necessary for any policy to achieve constant regret at all times (see, e.g., [KAG24, Example 3.1]). We note that the definition of general position gap we adopt coincides with that in [KAG24, KAG25].

Definition 1 (General position gap). A matching network \mathcal{G} satisfies the *general position gap* (GPG) condition if $\text{SPP}(\lambda)$ has a non-degenerate optimal solution, i.e., all n basic variables in this solution are strictly positive. When \mathcal{G} satisfies the GPG condition with a non-degenerate optimal solution (z^*, s^*) , define the GPG parameter ϵ as the minimum value of all basic variables, i.e.,

$$\epsilon \triangleq \min_{m \in \mathcal{M}, z_m^* > 0} z_m^* \wedge \min_{i \in \mathcal{A}, s_i^* > 0} s_i^*. \quad (3)$$

The GPG condition is a standard assumption in online revenue management and dynamic matching literature (see, e.g., [JK12, CLY24, KAG24]), and any linear program can satisfy this condition with an arbitrarily small perturbation [MC89]. The power of the GPG condition comes from the following important property (see, e.g., [WXY23, Gup24, KAG25]).

Proposition 1 (Corollary 4.1 in [KAG25]). Suppose that \mathcal{G} satisfies the GPG condition, and let (z^*, s^*) be a non-degenerate optimal solution of $\text{SPP}(\lambda)$ with ϵ defined as (3). Then, for every $\lambda' \in \mathbb{R}_{\geq 0}^n$ with $\|\lambda - \lambda'\|_1 \leq \epsilon$, $\text{SPP}(\lambda')$ has an optimal solution with the same basic activities, i.e., non-zero components, as (z^*, s^*) .

Given a matching network \mathcal{G} that satisfies the GPG condition, let (z^*, s^*) be a non-degenerate optimal solution of $\text{SPP}(\lambda)$ with ϵ defined as (3). Define $\mathcal{M}_+ \triangleq \{m \in \mathcal{M} \mid z_m^* > 0\}$ and $\mathcal{M}_0 \triangleq \mathcal{M} \setminus \mathcal{M}_+$ as the set of active matches and the set of redundant matches, respectively. Also, define $\mathcal{A}_+ \triangleq \{j \in \mathcal{A} \mid s_j^* > 0\}$ and $\mathcal{A}_0 \triangleq \mathcal{A} \setminus \mathcal{A}_+$ as the set of under-demanded queues and the set of over-demanded queues, respectively. Note that

$$\epsilon = \min_{m \in \mathcal{M}_+} z_m^* \wedge \min_{j \in \mathcal{A}_+} s_j^* > 0.$$

When there exist multiple connected components in the graph $(\mathcal{A}, \mathcal{M})$, we can analyze the policy independently on each connected component, since the actions on one connected component do not affect the performance on another one. As a result, we assume without loss of generality that $(\mathcal{A}, \mathcal{M})$ only consists of one connected component.

As an important consequence of the GPG condition, which is also commonly used by prior work (see, e.g., [KAG25, Lemma 5.1] and [Gup24, Lemma 1]), bounding the all-time regret of a policy can be boiled down to analyzing the total length of the over-demanded queues, provided that the policy is restricted to active matches.

Lemma 1. Suppose that \mathcal{G} satisfies the GPG condition, and let (z^*, s^*) be a non-degenerate optimal solution of $\text{SPP}(\lambda)$. Suppose that the following conditions hold under a policy Π :

1. Only matches in \mathcal{M}_+ are performed, and

2. $\sum_{i \in \mathcal{A}_0} \mathbb{E}[Q_i(t)] \leq B$ for every $t > 0$, where $B > 0$ does not depend on t .

Then, $\text{Regret}(\Pi, T) \leq r_{\max} n B$, where $r_{\max} \triangleq \max_{m \in \mathcal{M}_+} r_m$.

The proof of Lemma 1 is deferred to Section C.1. In view of Lemma 1, to achieve constant regret at all times, it suffices to restrict the policies to only use active matches and control the lengths of over-demanded queues. We note that it is a common strategy to ignore redundant matches [KAG24, Gup24, KAG25] with few exceptions [WXY23]. From now on, we assume $\mathcal{M} = \mathcal{M}_+$.

We say that a queue $i \in \mathcal{A}$ is *truncated* if, whenever an agent in queue i arrives and is not matched within the same period, this agent is immediately discarded, where discarding agents can be equivalently viewed as setting these agents aside and never matching them. We allow policies to truncate the under-demanded queues, i.e., queues in \mathcal{A}_+ .²

2.3 Policy information classes

Throughout, we focus on greedy policies that perform at most one match at each period, where the performed match must include the newly arriving agent. Consequently, all policies studied by prior work [WXY23, Gup24, KAG25] are greedy policies, with the exception of the batching policy [KAG24]. We distinguish greedy policies in terms of the information they use when performing matches:

The first dimension of classification concerns the granularity of the state information. A policy is availability-based if its matching decision utilizes only availability information across agent types (i.e., whether a queue is empty or not), representing a coarse level of state information. Conversely, a policy is queue-length-based if its matching decision utilizes exact queue-length information across agent types.

The second dimension distinguishes greedy policies based on the scope of the utilized information. We classify a policy as local if the matching decision upon the arrival of an agent of type $i \in \mathcal{A}$ depends solely on the state information of i 's neighboring queues. If a policy utilizes state information beyond the local neighbors of the arriving agent, we classify it as global.

Taken together, these two dimensions allow us to classify any greedy policy into one of four distinct information classes, as illustrated in Figure 1. For example, the longest-queue policy is local queue-length-based, the **TP** policy of [KAG25] is local availability-based, and the primal-dual policy of [WXY23] and the sum-of-squares policy of [Gup24] are global queue-length-based.

3 Main results

In this section, we present our main results regarding the information requirements for optimal matching policies. Contrary to the intuition that strictly less information might degrade performance, our analysis demonstrates that coarse availability information is sufficient to achieve the optimal all-time regret scaling. We establish this first for general networks via a global availability-based policy in Section 3.1, and then for acyclic networks via local availability-based policies in Section 3.2.

Algorithm 1 Probabilistic matching (**PM**)

1: $Q(0) \leftarrow \mathbf{0}$
2: **for** $t = 1, \dots, T$ **do**
3: $\mathcal{U}_+(t) \leftarrow \{i \in \mathcal{A} \mid Q_i(t-1) > 0\}; \mathcal{U}_0(t) \leftarrow \mathcal{A} \setminus \mathcal{U}_+(t)$
4: Define $\tilde{\lambda}(t) \in \mathbb{R}^n$ such that

$$\tilde{\lambda}_i(t) = \begin{cases} \lambda_i, & i \in \mathcal{U}_0(t), \\ \lambda_i + \epsilon/n, & i \in \mathcal{U}_+(t), \end{cases}$$

5: Let $(\tilde{z}(t), \tilde{s}(t))$ be an optimal solution of $\text{SPP}(\tilde{\lambda}(t))$ with the same basic activities as (z^*, s^*)
 ▷ Guaranteed by Proposition 1.
6: An agent of type $j \in \mathcal{A}$ arrives
7: Match the arriving agent to an agent in each queue $i \in \mathcal{N}(j) \cap \mathcal{U}_+(t)$ with probability

$$\frac{\tilde{z}_{m(i,j)}(t)}{\sum_{k \in \mathcal{N}(j) \cap \mathcal{U}_+(t)} \tilde{z}_{m(k,j)}(t)}.$$

8: **if** the arriving agent is matched to an agent in queue i **then**
9: $Q_i(t) \leftarrow Q_i(t) - 1$
10: **else**
11: $Q_j(t) \leftarrow Q_j(t) + \mathbf{1}_{\{j \notin \mathcal{A}_+\}}$
12: **end if**
13: **end for**

3.1 Global availability-based policy

In this subsection, we propose a global availability-based policy *probabilistic matching* (**PM**), which is formally presented in Algorithm 1, that achieves optimal all-time regret. Fix $t > 0$, and we describe the dynamics of the policy at time t . Define $\mathcal{U}_+(t)$ and $\mathcal{U}_0(t)$ respectively as the set of non-empty queues and the set of empty queues at the beginning of time t . Define a new arrival vector $\tilde{\lambda}(t)$ as in Line 4, which does not necessarily satisfy $\sum_{i \in \mathcal{A}} \tilde{\lambda}_i(t) = 1$. Recall that (z^*, s^*) is a non-degenerate optimal solution of $\text{SPP}(\lambda)$. Since $\|\lambda - \tilde{\lambda}(t)\|_1 \leq \epsilon$, by Proposition 1, $\text{SPP}(\tilde{\lambda}(t))$ has an optimal solution $(\tilde{z}(t), \tilde{s}(t))$ with the same basic activities as (z^*, s^*) .

Suppose an agent of type $j \in \mathcal{A}$ arrives at time t , and j has at least one non-empty neighboring queue; otherwise, there are no matches we can perform. For each non-empty neighboring queue i of j , we match the arriving agent to an agent of type i with probability proportional to $\tilde{z}_{m(i,j)}(t)$. Finally, we update $Q(t)$ accordingly.

The following theorem bounds the all-time regret of **PM**, whose proof is presented in Section 4.1.

Theorem 1 (Probabilistic matching). *Suppose that \mathcal{G} satisfies the GPG condition with ϵ defined as (3). Then, **PM** is global and availability-based, and satisfies*

$$\text{Regret}(\mathbf{PM}, T) \leq O\left(r_{\max} \cdot \frac{n^{2.5}}{\epsilon}\right),$$

where $r_{\max} = \max_{m \in \mathcal{M}_+} r_m$.

²Under the GPG condition, without discarding, the expected length of every under-demanded queue will grow unbounded under any matching policy with bounded all-time regret [KAG24]. Hence, discarding is necessary to ensure the ergodicity of the Markov chain $(Q(t))_{t \geq 0}$.

The proof of Theorem 1 proceeds by showing that the standard quadratic Lyapunov function exhibits a negative one-step drift under the matching probabilities chosen by **PM**.

3.2 Local availability-based policy

In this subsection, we turn to local availability-based policies, assuming that the matching network $(\mathcal{A}, \mathcal{M})$ is acyclic (i.e., it forms a tree). In this setting, [KAG25, Lemma 3.1] establishes the existence of a unique under-demanded queue, denoted by r , allowing us to view $(\mathcal{A}, \mathcal{M})$ as a tree rooted at r . We focus on *static priority policies*, a simple yet powerful family of local availability-based policies widely studied in prior work (see, e.g., [MP17, ADSW25, KAG25]).

Definition 2 (Static priority policy). For each $i \in \mathcal{A}$, let \succ_i be a strict ordering over a subset of matches $\mathcal{M}(i) \subseteq \mathcal{M}$ that contain i . The static priority policy with priority orders $(\succ_i)_{i \in \mathcal{A}}$ is a policy that, upon the arrival of an agent of type $i \in \mathcal{A}$, performs at most one available match in $\mathcal{M}(i)$. The policy chooses the highest priority match according to \succ_i whenever one or more matches in $\mathcal{M}(i)$ are available.

First, we revisit the tree priority (**TP**) policy proposed by [KAG25]. **TP** is a static priority policy where $\mathcal{M}(i) = \{m(i, j)\}_{j \in \mathcal{N}(i)}$ includes all matchings that involve i . The priority orders are defined such that $m(i, j) \succ_i m(i, P(i))$ for all $i \in \mathcal{A}_0$ and $j \in \mathcal{C}(i)$. In other words, upon the arrival of an agent of type i , **TP** first attempts to match this agent to an agent in the children of i ; if no child node is available, **TP** then attempts to match this agent to an agent in its parent.

While [KAG25] establishes that the regret of **TP** is independent of T , their analysis does not characterize how this regret depends on the network structure or the GPG parameter ϵ . To address this gap, our contribution here is to provide the first explicit regret bound for **TP** in the following theorem, whose proof is presented in Section 4.2.

Theorem 2 (Tree priority). *Suppose that \mathcal{G} satisfies the GPG condition with ϵ defined as (3), and the graph $(\mathcal{A}, \mathcal{M})$ is acyclic. Then,*

$$\text{Regret}(\mathbf{TP}, T) \leq O \left(r_{\max} \cdot \frac{n^2 2^{d_r}}{\epsilon} \left(1 + \frac{1}{\epsilon} \right)^{\lfloor (d_r - 1)/2 \rfloor} \right),$$

where $r_{\max} = \max_{m \in \mathcal{M}_+} r_m$ and d_r is the depth of the tree $(\mathcal{A}, \mathcal{M})$ when rooted at r .

To prove Theorem 2, we construct a generalized quadratic Lyapunov function with a one-step negative drift by heavily exploiting the hierarchical structure of the tree, where the coefficients of the Lyapunov function are carefully chosen to capture the propagation of queue imbalances along different layers of the tree. While our analysis follows a similar proof strategy of [KAG25], their analysis fails to track the coefficients of the Lyapunov function. We overcome this barrier by employing much more refined and delicate arguments. We further illustrate in Section B.1 that the current bound can not be further improved by applying the one-step generalized quadratic Lyapunov function.

In addition to the regret bound provided in Theorem 2 for all acyclic networks, in Section B.2, we consider $(\mathcal{A}, \mathcal{M})$ to be a path of 4 nodes and show the regret bound of $O(\epsilon^{-1})$ for **TP**, matching the $\Omega(\epsilon^{-1})$ lower bound by [KAG24]. This new bound improves upon the $O(\epsilon^{-2})$ scaling implied by Theorem 2 for this specific case, confirming that our Lyapunov analysis for proving Theorem 2 is not tight. However, this tighter bound is achieved via tailored arguments which, as discussed at the end of Section B.2, do not generalize to arbitrary acyclic networks.

Next, we propose the truncated tree priority (**TTP**) policy. **TTP** is a static priority policy where, for every $i \in \mathcal{A}$, the set of allowable matches is restricted to the children of i , i.e., $\mathcal{M}(i) = \{m(i, j)\}_{j \in \mathcal{C}(i)}$, with \succ_i being an arbitrary order over $\mathcal{M}(i)$. In other words, when an agent of type i arrives, **TTP** attempts to match it with an available agent in any child node of i , breaking ties according to \succ_i . If no such agent exists, the arriving agent joins queue i . Crucially, **TTP** never attempts to match the arriving agent with agents in the parent node of i .

As our main result, we show that **TTP** is the first static priority policy to achieve the optimal regret scaling of $O(\epsilon^{-1})$ at all times. This indicates that local availability information is sufficient to achieve optimal performance in acyclic graphs. The proof of Theorem 3 is presented in Section C.3.5.

Theorem 3 (Truncated tree priority). *Suppose \mathcal{G} satisfies the GPG condition with parameter ϵ defined as (3), and the graph $(\mathcal{A}, \mathcal{M})$ is acyclic. Then,*

$$\text{Regret}(\mathbf{TTP}, T) \leq O\left(r_{\max} \cdot \frac{n^2 d_r^2}{\epsilon}\right),$$

where $r_{\max} = \max_{m \in \mathcal{M}_+} r_m$ and d_r is the depth of the tree $(\mathcal{A}, \mathcal{M})$ when rooted at r .

The matching logic of both **TP** and **TTP** aligns with the intuition that nodes further from the under-demanded root are more demanded, thus warranting higher priority for their matches. The distinction arises when no child nodes are available: **TTP** strictly forbids the arriving agent from matching with its parent, forcing it to wait in the queue, whereas **TP** allows the agent to match with its parent as a final resort. We illustrate the distinction between **TP** and **TTP** by the following example.

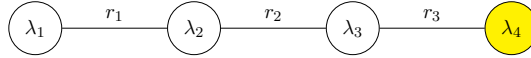


Figure 2: A path network where $\mathcal{A}_+ = \{4\}$ (the root is indicated with a yellow node).

Example 1. When the matching network forms a path as in Figure 2, both **TP** and **TTP** prioritize the match between any node i and its child $i - 1$ over the match between i and its parent $i + 1$. For arrival sequence $(3, 2, 1)$ (the first arrival is to queue 3, then to queue 2, and then to queue 1), **TP** performs match $m(2, 3)$ upon the arrival to queue 2, whereas **TTP** does not perform this match but instead performs the match $m(1, 2)$ upon the arrival to queue 1.

Intuitively, by preventing the arriving agent from matching with the parent node, **TTP** ensures that existing agent arrivals at a node do not influence the future dynamics of its subtree. This structural decoupling enables us to leverage a *multi-step drift analysis* for **TTP**, overcoming the limitations of the standard one-step drift analysis that yields suboptimal bounds for **TP**. Specifically, we generalize **TTP** to accommodate *fractional arrivals* and construct a geometric Lyapunov function by utilizing two technical ingredients: the Lipschitz continuity of the queue-length process, and the one-step negative drift under *fluid arrivals*. By further applying standard concentration inequalities for the true arrival process, we establish the desired multi-step drift, which yields the optimal $O(\epsilon^{-1})$ bound for **TTP**—a result that remains elusive for **TP**.

4 Analysis

In this section, we formally analyze the policies considered in this paper. We present the proofs of Theorems 1, 2, and 3 in Sections 4.1, 4.2, and 4.3, respectively.

To lay the groundwork for these analyses, we first introduce the formal definition of (geometric) Lyapunov functions.

Definition 3 (Lyapunov function). Let $(X(t))_{t \geq 0}$ be a discrete-time Markov chain defined on a complete metrizable state space \mathcal{X} . A function $\Phi : \mathcal{X} \rightarrow \mathbb{R}_{\geq 0}$ is a *Lyapunov function* with drift-size parameter $\gamma > 0$, drift-time parameter $t_0 > 0$, and exception parameter K if

$$\sup_{x \in \mathcal{X} : \Phi(x) > K} \{\mathbb{E}_x[\Phi(X(t_0))] - \Phi(x)\} \leq -\gamma.$$

A function $\Phi : \mathcal{X} \rightarrow \mathbb{R}_{\geq 0}$ is a *geometric Lyapunov function* with a geometric drift size $0 < \gamma < 1$, drift time t_0 , and exception parameter K if

$$\sup_{x \in \mathcal{X} : \Phi(x) > K} \frac{\mathbb{E}_x[\Phi(X(t_0))]}{\Phi(x)} \leq \gamma.$$

Next, we introduce tools for translating queue-length bounds under the steady state to all-time queue-length bounds. Per Lemma 1, since we are concerned with the all-time regret performance of matching policies, it is crucial to bound the expected queue lengths at all times. However, Lyapunov techniques only allow us to bound the expected queue lengths in the steady state. We show that, for the matching policies satisfying the following consistency property, queue-length bounds that hold at any time can be directly deduced from the bounds under the steady state.

The consistency property informally posits that, when starting from different initial states with the same arrival process, the difference between the induced states does not grow over time. Similar concepts also appear in prior work under the names Lipschitz continuity [MP17] or non-expansiveness [BB13].

Definition 4 (Consistency). We say that a policy is *consistent* if for all valid initial states $Q(0)$ and $Q'(0)$, for every possible arrival at time 1, there exists a coupling μ between $Q(1)$ and $Q'(1)$ such that

$$\mathbb{E}_{(Q(1), Q'(1)) \sim \mu} [\|Q(1) - Q'(1)\|_1] \leq \|Q(0) - Q'(0)\|_1,$$

where the randomness of $Q(1)$ and $Q'(1)$ comes from the randomness used by the policy.

As a consequence of a policy being consistent, every bound for the expected total queue-length under the stationary distribution can be directly translated into an all-time bound.

Lemma 2. *Let Π be a consistent policy. Suppose that the Markov chain $(Q(t))_{t \geq 0}$ is ergodic, and let π be its stationary distribution. If $\mathbb{E}_\pi[\|Q(0)\|_1] \leq B$, then $\mathbb{E}[\|Q(t)\|_1] \leq 2B$ for every $t \geq 0$.*

The proof of Lemma 2 is deferred to Section A.4.1. Moreover, we show in Section A.4 that all static priority policies and the longest-queue policy are consistent.

4.1 Analysis of probabilistic matching policy

In this subsection, we prove Theorem 1 by largely following a standard analysis for the quadratic Lyapunov function (see, e.g., the proof of [KAG25, Lemma 5.4]). The global and availability-based properties of **PM** come from its description. Moreover, since **PM** always matches the arriving agent whenever at least one of its neighboring queues is non-empty, as stated in the following fact, any two adjacent queues cannot be non-empty at the same time.

Fact 1. Let $(Q(t))_{t \geq 0}$ be the states induced by **PM** under an arbitrary sample path. Then, for all $t \geq 0$ and $m(i, j) \in \mathcal{M}$, $Q_i(t) \cdot Q_j(t) = 0$.

Then, we analyze the Markov chain $(Q(t))_{t \geq 0}$ by using the following quadratic Lyapunov function: For $t \geq 0$, define

$$\mathcal{L}(t) \triangleq \sum_{i \in \mathcal{A}_0} (Q_i(t))^2.$$

Fix $t \geq 0$, and we upper bound the drift $\mathbb{E}[\mathcal{L}(t+1) - \mathcal{L}(t) \mid Q(t)]$, where the expectation is taken over the randomness of the arrival process and **PM**. For each match $m \in \mathcal{M}$, let x_m denote the probability that m is performed by **PM** at time $t+1$. By standard calculation (see, e.g., [KAG25, Proposition 5.1]),

$$\mathbb{E}[\mathcal{L}(t+1) - \mathcal{L}(t) \mid Q(t)] \leq 2\langle Q(t), \lambda - Mx \rangle + 1. \quad (4)$$

Recall by 1 that any two adjacent queues cannot be non-empty at the same time. For each match $m \in \mathcal{M}$ that contains types i and j such that $i \in \mathcal{U}_+(t)$ and $j \in \mathcal{U}_0(t)$, m is performed at time $t+1$ if and only if an agent of type j arrives and **PM** decides to match this agent with an agent in queue i ; it follows that

$$\begin{aligned} x_m &= \lambda_j \cdot \frac{\tilde{z}_m(t)}{\sum_{k \in \mathcal{N}(j) \cap \mathcal{U}_+(t)} \tilde{z}_{m(k,j)}(t)} \geq \lambda_j \cdot \frac{\tilde{z}_m(t)}{\sum_{k \in \mathcal{N}(j)} \tilde{z}_{m(k,j)}(t)} \\ &= \lambda_j \cdot \frac{\tilde{z}_m(t)}{(M\tilde{z}(t))_j} \geq \lambda_j \cdot \frac{\tilde{z}_m(t)}{\tilde{\lambda}_j(t)} = \lambda_j \cdot \frac{\tilde{z}_m(t)}{\lambda_j} = \tilde{z}_m(t), \end{aligned}$$

where the second inequality holds since $(\tilde{z}(t), \tilde{s}(t))$ is a feasible solution of $\text{SPP}(\tilde{\lambda}(t))$. Hence, for each queue $i \in \mathcal{U}_+(t)$,

$$(Mx)_i = \sum_{j \in \mathcal{N}(i)} x_{m(i,j)} \geq \sum_{j \in \mathcal{N}(i)} \tilde{z}_{m(i,j)}(t) = (M\tilde{z}(t))_i = \tilde{\lambda}_i(t), \quad (5)$$

where the last equality holds since $(\tilde{z}(t), \tilde{s}(t))$ has the same basic activities as (z^*, s^*) , and hence $\tilde{s}_i(t) = s_i^* = 0$ (recall that $\tilde{s}_i(t)$ is the slack variable in the constraint $(M\tilde{z}(t))_i + \tilde{s}_i(t) = \tilde{\lambda}_i(t)$ of $\text{SPP}(\tilde{\lambda}(t))$). Combining (4) and (5), and since $Q_i(t) = 0$ for every $i \in \mathcal{U}_0(t)$, we get

$$\mathbb{E}[\mathcal{L}(t+1) - \mathcal{L}(t) \mid Q(t)] \leq 2 \sum_{i \in \mathcal{U}_+(t)} Q_i(t)(\lambda_i - \tilde{\lambda}_i(t)) + 1 = -2\frac{\epsilon}{n} \|Q(t)\|_1 + 1. \quad (6)$$

Next, we apply the negative drift (6) and Lemma 10 to upper bound the expected total queue length. Fix $t \geq 0$. The variation of $\|Q(t)\|_2$ is bounded by

$$\|Q(t+1)\|_2 - \|Q(t)\|_2 \leq \|Q(t+1) - Q(t)\|_2 = \|\Delta A(t+1) - M\Delta D(t+1)\|_2 \leq 1,$$

where the last inequality holds since at most one match is performed at time $t+1$, and the performed match must involve the arriving agent. Then, we bound the expected decrease of $\|Q(t)\|_2$, i.e., $\mathbb{E}[\|Q(t+1)\|_2 - \|Q(t)\|_2 \mid Q(t)]$. When $\|Q(t)\|_2 \geq n/\epsilon$, we get

$$\begin{aligned} \mathbb{E}[\|Q(t+1)\|_2 - \|Q(t)\|_2 \mid Q(t)] &\leq \mathbb{E}\left[\frac{\|Q(t+1)\|_2^2 - \|Q(t)\|_2^2}{2\|Q(t)\|_2} \mid Q(t)\right] \\ &\leq \mathbb{E}\left[\frac{-2\frac{\epsilon}{n}\|Q(t)\|_1 + 1}{2\|Q(t)\|_2} \mid Q(t)\right] \leq -\frac{\epsilon}{2n}, \end{aligned}$$

where the first inequality holds by $x - y \leq (x^2 - y^2)/(2y)$ for $x \in \mathbb{R}$ and $y > 0$ [CJK⁺06, Lemma 3.6], the second inequality holds by (6), and the last inequality holds since $\|Q(t)\|_1 \geq \|Q(t)\|_2$.

Now, we apply Lemma 10 on $\Psi(t) = \|Q(t)\|_2$ with $K = 1$, $\eta = \epsilon/(2n)$, and $B = n/\epsilon$ to get

$$\mathbb{E}[\|Q(t)\|_2] = \mathbb{E}[\|\Psi(t)\|] \leq 2 + \frac{n}{\epsilon} + \frac{1 - \frac{\epsilon}{2n}}{\frac{\epsilon}{n}} = \frac{3}{2} + \frac{2n}{\epsilon} \leq \frac{3n}{\epsilon},$$

where the last inequality holds since $\epsilon \leq 1$ and $n \geq 2$. Finally, the all-time regret bound for **PM** follows from $\|Q(t)\|_1 \leq \sqrt{n} \|Q(t)\|_2$ and Lemma 1.

4.2 Analysis of tree priority policy on acyclic graphs

In this subsection, we prove Theorem 2, and all the omitted proofs can be found in Section C.2.

For all $i, j \in \mathcal{A}$, let $d(i, j)$ be the (unweighted) distance between i and j . For each $i \in \mathcal{A}$, define $d_i \triangleq \max_{j \in \mathcal{T}(i)} d(i, j)$ as the depth of the subtree rooted at i . For each $i \in \mathcal{A}$, denote $\mathcal{P}(i) \triangleq \{j \in \mathcal{A}_0 \mid i \in \mathcal{T}^-(j) \text{ with } d(j, i) \equiv 0 \pmod{2}\}$ as the set of over-demanded ancestors of i whose depth has the same parity with i .

For each $i \in \mathcal{A}$, define

$$\epsilon_i \triangleq \begin{cases} z_{m(i, \mathcal{P}(i))}^*, & i \in \mathcal{A}_0, \\ s_i^*, & i \in \mathcal{A}_+. \end{cases} \quad (7)$$

By (3), we have $\epsilon = \min_{i \in \mathcal{A}} \epsilon_i$, and $\epsilon \leq \epsilon_i \leq 1$ for every $i \in \mathcal{A}$. The following lemma characterizes ϵ_i in terms of λ .

Lemma 3 (Theorem 4.1 in [KAG25]). *Suppose that \mathcal{G} satisfies the GPG condition, and the graph $(\mathcal{A}, \mathcal{M})$ is acyclic. Then, for every $i \in \mathcal{A}_0$,*

$$\epsilon_i = \sum_{j \in \mathcal{T}(i)} (-1)^{d(i, j)} \lambda_j.$$

We then recursively define the coefficients of our Lyapunov function: For every $i \in \mathcal{A}_0$, let

$$\alpha_i \triangleq 1 + \frac{1}{\epsilon_i} \sum_{j \in \mathcal{P}(i)} \alpha_j (\lambda_j - \epsilon_j). \quad (8)$$

Note that for every $i \in \{r\} \cup \mathcal{C}(r)$, we have $\mathcal{P}(i) = \emptyset$ and hence $\alpha_i = 1$. We consider the following generalized quadratic Lyapunov function: For every $t \geq 0$, define

$$\mathcal{L}(t) \triangleq \sum_{i \in \mathcal{A}_0} \alpha_i (f_i(Q(t))^+)^2, \quad (9)$$

where for every $i \in \mathcal{A}_0$ and $v \in \mathbb{R}^n$,

$$f_i(v) \triangleq \sum_{j \in \mathcal{T}^-(i)} (-1)^{d(i, j)+1} v_j. \quad (10)$$

Intuitively, to ensure a one-step negative Lyapunov drift, we need to recursively define the coefficients such that the positive drift appearing at lower levels of the tree (i.e., the levels further from the root) is dominated by the negative drift at higher levels.

Notably, our Lyapunov function recovers the one used in [KAG25] if we do not take the positive parts. As noted in [KAG25], the coefficients α_i 's are intractable when using their Lyapunov function. As we will show, this modification to the Lyapunov function is crucial for deriving an explicit upper bound for the queue length. We reiterate that our proof significantly differs from [KAG25] and relies on a more delicate drift analysis. It is also unclear how to adapt their arguments to give an explicit queue-length bound even using our Lyapunov function.

The following lemma upper bounds the coefficients α_i 's.

Lemma 4. *For every $i \in \mathcal{A}_0$, $\alpha_i \leq (1 + \epsilon^{-1})^{\lfloor (d(r,i)-1)/2 \rfloor}$.*

In the following proposition, we upper bound the drift $\mathbb{E}[\mathcal{L}(t+1) - \mathcal{L}(t) \mid Q(t)]$, where the expectation is taken over the randomness of the arrival at time $t+1$.

Proposition 2. *It holds that*

$$\mathbb{E}[\mathcal{L}(t+1) - \mathcal{L}(t) \mid Q(t)] \leq -\frac{\epsilon}{2^{d_r-1}} \|Q(t)\|_1 + n \left(1 + \frac{1}{\epsilon}\right)^{\lfloor (d_r-1)/2 \rfloor}.$$

Moreover, the Markov chain $(Q(t))_{t \geq 0}$ is ergodic.

By Proposition 2, the Markov chain $(Q(t))_{t \geq 0}$ is ergodic, and we denote its stationary distribution as π . By applying Lemma 9 with $f(Q(t)) = \frac{\epsilon}{2^{d_r-1}} \|Q(t)\|_1$, $g(Q(t)) = \mathcal{L}(t)$, and $c = n(1 + \epsilon^{-1})^{\lfloor (d_r-1)/2 \rfloor}$, the drift bound in Proposition 2 yields

$$\mathbb{E}_\pi [\|Q(0)\|_1] \leq \frac{n2^{d_r-1}}{\epsilon} \left(1 + \frac{1}{\epsilon}\right)^{\lfloor (d_r-1)/2 \rfloor}.$$

By Proposition 7 and Lemma 2, it follows that, for every $t \geq 0$,

$$\mathbb{E} [\|Q(t)\|_1] \leq \frac{n2^{d_r}}{\epsilon} \left(1 + \frac{1}{\epsilon}\right)^{\lfloor (d_r-1)/2 \rfloor}.$$

Finally, the regret bound for **TP** follows from Lemma 1.

4.3 Analysis of truncated tree priority policy on acyclic graphs

In this subsection, we present the proof of Theorem 3, which follows a multi-step drift analysis. All the omitted proofs in this subsection can be found in Section C.3.

Our primary goal is to construct a geometric Lyapunov function and then apply the following lemma to derive the desired queue-length bound under the stationary distribution.³

Lemma 5 (Theorem 5 in [GZ06]). *Let $(X(t))_{t \geq 0}$ be a discrete-time Markov chain defined on a complete metrizable state space \mathcal{X} that possesses a stationary distribution π . Suppose that $\Phi : \mathcal{X} \rightarrow \mathbb{R}_{\geq 0}$ is a geometric Lyapunov function with parameters γ , t_0 , and K . Then,*

$$\mathbb{E}_\pi [\Phi(X)] \leq \frac{K}{1 - \gamma} \cdot \sup_{x \in \mathcal{X}} \frac{\mathbb{E}_x [\Phi(X(t_0))]}{\Phi(x)}.$$

We will construct the geometric Lyapunov function by the following proposition, which is a modified version of [GZ06, Theorem 6].

³Proposition 5 is originally established for continuous-time Markov chains, but its proof also holds for discrete-time Markov chains.

Proposition 3. Let $(X(t))_{t \geq 0}$ be a discrete-time Markov chain defined on a complete metrizable state space \mathcal{X} . Suppose Φ is a Lyapunov function with δ -truncated drift for $\delta > 0$:

$$\sup_{x \in \mathcal{X}: \Phi(x) > K} \mathbb{E}_x [\max \{-\delta, \Phi(X(t_0)) - \Phi(x)\}] \leq -\gamma. \quad (11)$$

Moreover, suppose that there exists $\theta > 0$ such that

$$\theta L_3(\delta, \theta, t_0) \leq \gamma, \quad (12)$$

where

$$L_3(\delta, \theta, t) \triangleq \sup_{x \in \mathcal{X}} \mathbb{E}_x [(\max \{-\delta, \Phi(X(t)) - \Phi(x)\})^2 \exp(\theta(\Phi(X(t)) - \Phi(x))^+)]. \quad (13)$$

Then, $\exp(\theta \cdot \Phi(\cdot))$ is a geometric Lyapunov function with geometric drift size parameter $1 - \gamma\theta/2$, drift time parameter t_0 and exception parameter $e^{\theta K}$.

To apply Proposition 3, we need a Lyapunov function with two multi-step drift properties: a negative truncated drift (11) and a bounded second-order exponential moment (12). For our choice of the Lyapunov function, we use

$$\Phi(Q) \triangleq \sum_{i \in \mathcal{A}_0} Q_i \quad (14)$$

defined for $Q \in \mathbb{R}_{\geq 0}^n$.

TTP under fractional arrivals. To establish the multi-step drift properties for Φ , we need to generalize **TTP** to accommodate *fractional arrivals*, enabling us to utilize **TTP**'s desirable drift properties under the *fluid arrival*. Specifically, $\Delta A_i(t)$ can take any value in $\mathbb{R}_{\geq 0}$, and we allow arrivals at multiple nodes at each period. We also allow performing fractional matches, i.e., $\Delta D_m(t)$ can take any value in $\mathbb{R}_{\geq 0}$. We generalize **TTP** to match fractional arrivals as follows. At each period, we visit the nodes in \mathcal{A} in an order such that every node $i \in \mathcal{A}_0$ is visited before the parent of i . For the current node i , we match the new arrivals at i to as many agents in $\mathcal{C}(i)$ as possible. Notice that, under the true arrival, this generalized version of **TTP** behaves identically as the original **TTP**. We emphasize that this generalization only serves the purpose of analysis, and **TTP** does not perform fractional matches under the true arrival.

The following two propositions serve as the main technical ingredients of our analysis. The first proposition shows that, for two fractional arrival trajectories, when starting from the same initial state, the difference between the respective queue-length trajectories induced by **TTP** is controlled by the difference between the arrival trajectories.

Proposition 4 (Lipschitz continuity). For two fractional arrival trajectories $(A(t))_{t \geq 0}$ and $(A'(t))_{t \geq 0}$, let $(q(t))_{t \geq 0}$ and $(q'(t))_{t \geq 0}$ be the respective queue-length trajectories induced by **TTP** with $q(0) = q'(0)$. Then, for every $t \geq 0$,

$$\left| \sum_{i \in \mathcal{A}_0} q_i(t) - \sum_{i \in \mathcal{A}_0} q'_i(t) \right| \leq 2(d_r + 1) \sum_{i \in \mathcal{A}} \max_{0 \leq s \leq t} |A_i(s) - A'_i(s)|, \quad (15)$$

where d_r is the depth of the tree $(\mathcal{A}, \mathcal{M})$ when rooted at r .

The next proposition asserts that, under the fluid arrival, i.e., $A_i(t) = t\lambda_i$ for all $i \in \mathcal{A}$ and $t \geq 0$, Φ exhibits a negative one-step drift for the queue-length process induced by **TTP**. We will use $(q(t))_{t \geq 0}$ to denote the queue-length process induced by **TTP** under the fluid arrival.

Proposition 5 (Negative drift under fluid arrival). *Suppose that $A_i(t) = t\lambda_i$ for all $i \in \mathcal{A}$ and $t \geq 0$. Then, under **TTP**, for every $t \geq 0$,*

$$\Phi(q(t+1)) \leq (\Phi(q(t)) - \epsilon)^+. \quad (16)$$

With the above two propositions, we can now show that, under the true arrival, Φ satisfies the multi-step drift properties required by Proposition 3, yielding the desired geometric Lyapunov function. Specifically, by the one-step negative drift of Φ under the fluid arrival (Proposition 5), the total queue length decreases linearly under the fluid arrival (recall that Φ encodes the total queue length). Since the true arrival process concentrates around the fluid arrival process by standard concentration inequalities (Lemma 8), the Lipschitz-continuity property (Proposition 4) implies that the queue-length process under the true arrival also concentrates around the queue-length process under the fluid arrival. Choosing an appropriate step size then concludes the multi-step drift properties of Φ under the true arrival. We formalize the above arguments in the following proposition.

Proposition 6. *Let $(Q(t))_{t \geq 0}$ be the queue-length process induced by **TTP** under the true arrival. Then, there exist universal constants $n_0, \kappa_0 > 0$ such that for any $n \geq n_0$, any $0 < \epsilon < 1$, and any*

$$\theta \leq \frac{\kappa_0}{(d_r + 1)^2 n}, \quad (17)$$

the function $V(Q) \triangleq \exp(\epsilon\theta \cdot \Phi(Q))$ for $Q \in \mathbb{R}_{\geq 0}^n$ is a geometric Lyapunov function for $(Q(t))_{t \geq 0}$ with:

- *Geometric drift size parameter:* $1 - \frac{5}{2}(d_r + 1)^2 n\theta$,
- *Drift time parameter:* $K \triangleq t_0 \epsilon^{-2}$, where $t_0 \triangleq 25(d_r + 1)^2 n$,
- *Exception parameter:* $\exp(\theta t_0)$.

Now, we are ready to finish the proof of Theorem 3. We sketch the proof argument here and defer the complete proof to Section C.3.5.

Proof sketch of Theorem 3. By Proposition 6, under the true arrival, $V(Q) = \exp(\epsilon\theta \cdot \Phi(Q))$ is a geometric Lyapunov function with geometric drift size parameter $1 - \frac{5}{2}(d_r + 1)^2 n\theta$, drift time parameter $K = t_0 \epsilon^{-2}$, and exception parameter $\exp(\theta t_0)$. By Proposition 5, we get

$$\mathbb{E}_\pi \left[e^{\epsilon\theta \cdot \Phi(Q)} \right] \leq \frac{2e^{\theta t_0} \phi(K)}{5(d_r + 1)^2 n\theta} \leq O(1), \quad (18)$$

where $\phi(K) \triangleq \sup_{x \in \mathbb{R}_{\geq 0}^n} \mathbb{E}_x \left[e^{\epsilon\theta(\Phi(Q(K)) - \Phi(x))} \right]$ is the maximum expected overshoot, and the last inequality is derived by applying Propositions 4 and 5. By Jensen's inequality,

$$\mathbb{E}_\pi[\Phi(Q)] \leq \frac{1}{\epsilon\theta} \log \mathbb{E}_\pi[e^{\epsilon\theta \cdot \Phi(Q)}] \leq O\left(\frac{nd_r^2}{\epsilon}\right), \quad (19)$$

where the second inequality holds by (18) and the definition of θ .

Finally, **TTP** is consistent by Proposition 7, and hence applying Lemma 2 yields

$$\mathbb{E} \left[\sum_{i \in \mathcal{A}_0} Q(t) \right] \leq 2\mathbb{E}_\pi \left[\sum_{i \in \mathcal{A}_0} Q(0) \right] = 2\mathbb{E}_\pi[\Phi(Q)] \leq O\left(\frac{nd_r^2}{\epsilon}\right)$$

for every $t \geq 0$, where the last inequality holds by (19). Combining the above displayed equation with Lemma 1 concludes the proof. \square

Remark 1. Our analysis crucially relies on the Lipschitz-continuity property of **TTP** (Proposition 4). While it is difficult to directly establish drift properties under the true arrival, drift properties under the fluid arrival appear much more tractable. In addition, the proximity of the fluid arrival process and the true arrival process can be derived via standard concentration bounds. Applying the Lipschitz-continuity property then yields multi-step drift properties under the true arrival. Since our analysis does not require properties of the true arrival process other than it being well-concentrated, our result for **TTP** can be potentially extended to arrival processes beyond the multinomial one, as long as suitable concentration estimates are available.

While the Lipschitz-continuity property might seem natural, establishing it turns out to be notoriously challenging for most policies, with a notable example being Jackson networks (see, e.g., [GZ06, Theorem 1]). A salient feature of **TTP** is its “feedforward-ness”: an arriving agent is never matched with its parent, so the current state of any node does not affect the future dynamics in its subtree. In contrast, by allowing the match between the arriving agent and its parent, **TP** would create “feed-back” interactions, which appear to be the main obstacle to establishing its Lipschitz-continuity property. In fact, for systems with such feed-back interactions, it remains challenging to establish Lipschitz-continuity properties even for small networks; to the best of our knowledge, similar properties are known only for very simple networks such as a two-node system [DR00].

5 Numerical experiments

In this section, we conduct numerical experiments to evaluate the performance of the proposed greedy policies and compare them with benchmark policies from the literature. All simulations are based on 1000 replications.

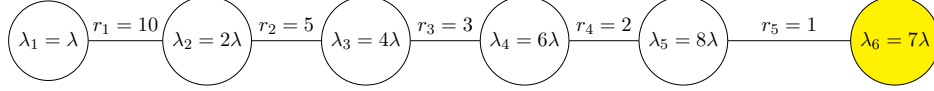
We first adopt the examples from [KAG25, Figures 5 and 10] as our first set of instances (see Figure 3). In Figure 4, we simulate **LG**, **PM**, **TP**, and **TTP** on these two networks, and the experiments suggest that there is no clear hierarchy between policies, i.e., there are instances, where **PM** and **LQ** are dominated both by **TP** and **TTP** (as shown in Figure 4a), and vice versa (as shown in Figure 4b). Nevertheless, the performance gap between all policies are remarkably small.

In Figure 5, we consider a cyclic network, and simulate the performance of **PM** and **LG** on this cyclic network, given that both **TP** and **TTP** requires the network to be acyclic. We observe that both policies perform very close to each other numerically, which is aligned with the fact that both policies achieve the optimal regret scaling of $O(\epsilon^{-1})$.⁴

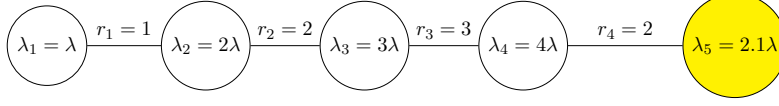
6 Conclusion

In this paper, we considered two-way dynamic matching and comprehensively studied availability-based policies, whose matching decisions depend only on agent availability across types, instead of

⁴We observe similar numerical performance between **PM** and **LQ** when the network structure exhibits more complexity, i.e., paths originating from cycle nodes.

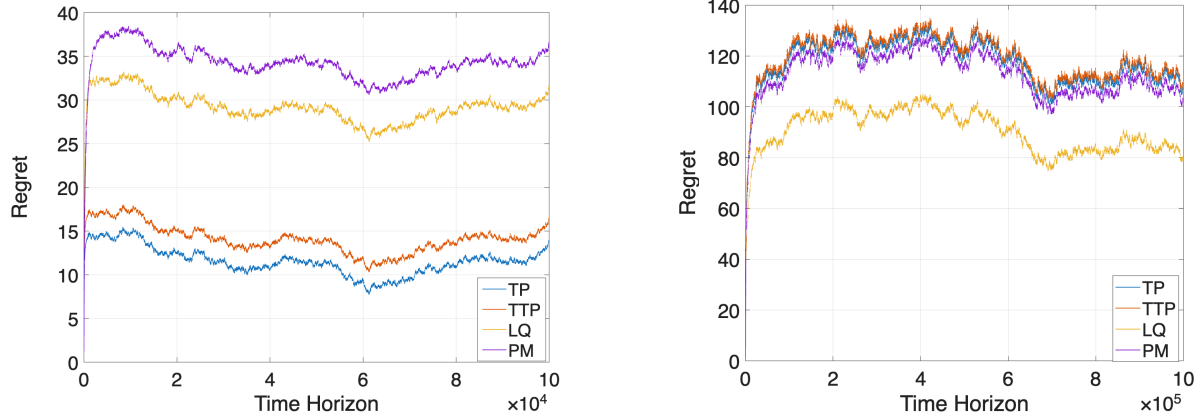


(a) $\epsilon = \lambda = 1/28$ and $\mathcal{A}_+ = \{6\}$ ([KAG25, Figure 5]).



(b) $\lambda \approx 0.08$, $\epsilon = 0.1\lambda$, and $\mathcal{A}_+ = \{5\}$ (Figure 10 from [KAG25, Figure 10])

Figure 3: Two example matching networks satisfying the GPG condition.



(a) Policy comparisons on the network in Figure 3a (b) Policy comparisons on the network in Figure 3b

Figure 4: Regret comparison of the tree priority (**TP**), the truncated tree priority (**TTP**), the longest-queue (**LQ**), and the randomized greedy (**PM**) policies.

utilizing complete queue-length information. We first proposed a global availability-based policy **PM** for general networks and a local availability-based policy **TTP** for acyclic networks, with both policies achieving the optimal regret scaling. It remains an exciting open problem to find a local availability-based policy that achieves the optimal regret scaling for general networks.

We also provided the first explicit regret bound for the local availability-based policy **TP** proposed by [KAG25]. Numerical outcomes suggest **TP** exhibits the optimal regret scaling, and we additionally show the optimal regret scaling of **TP** for a path of 4 nodes via tailored arguments, indicating that our Lyapunov analysis for **TP** is not tight. With the above evidences, we conjecture that **TP** exhibits the optimal regret scaling, and we believe the resolution of this conjecture would lead to novel insights and techniques.

Our numerical experiments show that there are instances where **PM** and **LQ** are dominated both by **TP** and **TTP**, and vice versa. Nevertheless, the performance gaps between all policies are remarkably small. Notably, we have not been able to construct an instance in which **TTP** achieves strictly better regret than **TP**, and therefore we reiterate the conjecture that **TP** also achieves the optimal regret scaling of $O(\epsilon^{-1})$ (see [KAG25, Example 6.1]).

For multi-way matching networks where each match can include more than two agent types, both the primal-dual policy of [WXY23] and the sum-of-square policy of [Gup24] achieve the optimal regret scaling. Nevertheless, both policies are global and queue-length-based, and it would be intriguing to achieve the optimal regret scaling for multi-way dynamic matching via (local)

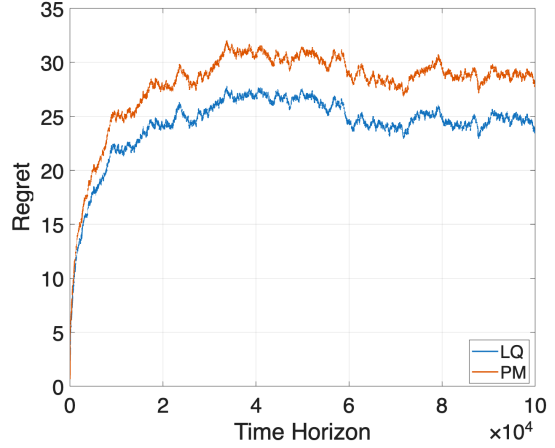
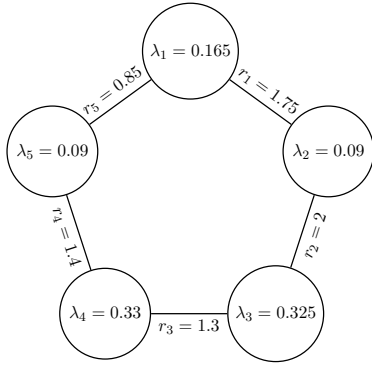


Figure 5: (LEFT) A cyclic two-way matching network that satisfies the GPG condition, where $z^* = \{0.085, 0.05, 0.32, 0.01, 0.08\}$ and $\epsilon = 0.01$. (RIGHT) Regret comparison of the longest-queue (**LQ**) and the probabilistic matching (**PM**) policies

availability-based policies.

While our work establishes strong theoretical guarantees for availability-based policies, an important future direction is to formalize their practical benefits of transparency and resistance against manipulation. Combining our results with the literature on strategic behaviors in queueing systems (e.g., [ERIZ25]) could help demonstrate how availability-based policies provide robustness in strategic environments under precise game-theoretic assumptions.

References

- [AAGK17] Ross Anderson, Itai Ashlagi, David Gamarnik, and Yash Kanoria. Efficient dynamic barter exchange. *Operations Research*, 65(6):1446–1459, 2017.
- [ABB⁺18] Itai Ashlagi, Adam Bingaman, Maximilien Burq, Vahideh Manshadi, David Gamarnik, Cathi Murphey, Alvin E Roth, Marc L Melcher, and Michael A Rees. Effect of match-run frequencies on the number of transplants and waiting times in kidney exchange. *American Journal of Transplantation*, 18(5):1177–1186, 2018.
- [ABJM19] Itai Ashlagi, Maximilien Burq, Patrick Jaillet, and Vahideh Manshadi. On matching and thickness in heterogeneous dynamic markets. *Operations Research*, 67(4):927–949, 2019.
- [ADSW25] Angelos Aveklouris, Levi DeValve, Maximiliano Stock, and Amy Ward. Matching impatient and heterogeneous demand and supply. *Operations Research*, 73(3):1637–1658, 2025.
- [AG19] Alessandro Arlotto and Itai Gurvich. Uniformly bounded regret in the multisecretary problem. *Stochastic Systems*, 9(3):231–260, 2019.
- [ALG20] Mohammad Akbarpour, Shengwu Li, and Shayan Oveis Gharan. Thickness and information in dynamic matching markets. *Journal of Political Economy*, 128(3):783–815, 2020.

- [AS22] Ali Aouad and Ömer Sarıtaç. Dynamic stochastic matching under limited time. *Operations Research*, 70(4):2349–2383, 2022.
- [BB13] Francois Baccelli and Pierre Brémaud. *Elements of queueing theory: Palm Martingale calculus and stochastic recurrences*, volume 26. Springer Science & Business Media, 2013.
- [BBP24] Santiago R Balseiro, Omar Besbes, and Dana Pizarro. Survey of dynamic resource-constrained reward collection problems: Unified model and analysis. *Operations Research*, 72(5):2168–2189, 2024.
- [BFP23] Eric Balkanski, Yuri Faenza, and Noémie Périvier. The power of greedy for online minimum cost matching on the line. In *Proceedings of the 24th ACM Conference on Economics and Computation*, pages 185–205, 2023.
- [BM15] Ana Bujić and Sean Meyn. Approximate optimality with bounded regret in dynamic matching models. *ACM SIGMETRICS Performance Evaluation Review*, 43(2):75–77, 2015.
- [BRS⁺22] Jose H Blanchet, Martin I Reiman, Virag Shah, Lawrence M Wein, and Linjia Wu. Asymptotically optimal control of a centralized dynamic matching market with general utilities. *Operations Research*, 70(6):3355–3370, 2022.
- [BW20] Pornpawee Bumpensanti and He Wang. A re-solving heuristic with uniformly bounded loss for network revenue management. *Management Science*, 66(7):2993–3009, 2020.
- [CHS24] Ziyun Chen, Zhiyi Huang, and Enze Sun. Stochastic online correlated selection. In *2024 IEEE 65th Annual Symposium on Foundations of Computer Science (FOCS)*, pages 2275–2294. IEEE, 2024.
- [CILB⁺20] Natalie Collina, Nicole Immorlica, Kevin Leyton-Brown, Brendan Lucier, and Neil Newman. Dynamic weighted matching with heterogeneous arrival and departure rates. In *Web and Internet Economics: 16th International Conference, WINE 2020, Beijing, China, December 7–11, 2020, Proceedings 16*, pages 17–30. Springer, 2020.
- [CJK⁺06] János Csirik, David S. Johnson, Claire Kenyon, James B. Orlin, Peter W. Shor, and Richard R. Weber. On the sum-of-squares algorithm for bin packing. *J. ACM*, 53(1):1–65, 2006.
- [CLY24] Guanting Chen, Xiaocheng Li, and Yinyu Ye. Technical note - an improved analysis of lp-based control for revenue management. *Operations Research*, 72(3):1124–1138, 2024.
- [CY01] Hong Chen and David D Yao. *Fundamentals of queueing networks: Performance, asymptotics, and optimization*, volume 46. Springer Science & Business Media, 2001.
- [dBBVLM22] Mark Van der Boor, Sem C Borst, Johan SH Van Leeuwen, and Debankur Mukherjee. Scalable load balancing in networked systems: A survey of recent advances. *SIAM Review*, 64(3):554–622, 2022.
- [DR00] Paul Dupuis and Kavita Ramanan. A multiclass feedback queueing network with a regular skorokhod problem. *Queueing Systems*, 36(4):327–349, 2000.

- [ERIZ25] Arturo Estrada Rodriguez, Rouba Ibrahim, and Dongyuan Zhan. On customer (dis-) honesty in unobservable queues: The role of lying aversion. *Management Science*, 71(1):844–860, 2025.
- [ES07] Atila Eryilmaz and Rayadurgam Srikant. Fair resource allocation in wireless networks using queue-length-based scheduling and congestion control. *IEEE/ACM transactions on networking*, 15(6):1333–1344, 2007.
- [ET25] Myungeun Eom and Alejandro Toriello. Batching and greedy policies: How good are they in dynamic matching? *Manufacturing & Service Operations Management*, 2025.
- [Gup24] Varun Gupta. Greedy algorithm for multiway matching with bounded regret. *Operations Research*, 72(3):1139–1155, 2024.
- [GZ06] David Gamarnik and Assaf Zeevi. Validity of heavy traffic steady-state approximations in generalized jackson networks. 2006.
- [GZ08] Peter W Glynn and Assaf Zeevi. Bounding stationary expectations of markov processes. In *Markov processes and related topics: a Festschrift for Thomas G. Kurtz*, volume 4, pages 195–215. Institute of Mathematical Statistics, 2008.
- [Ibr18] Rami Ibrahim. Sharing delay information in service systems: A literature survey. *Queueing Systems*, 89(1–2):49–79, 2018.
- [Jas15] Stefanus Jasin. Performance of an lp-based control for revenue management with unknown demand parameters. *Operations Research*, 63(4):909–915, 2015.
- [JK12] Stefanus Jasin and Sunil Kumar. A re-solving heuristic with bounded revenue loss for network revenue management with customer choice. *Mathematics of Operations Research*, 37(2):313–345, 2012.
- [JMRS23] Matthieu Jonckheere, Pascal Moyal, Claudia Ramírez, and Nahuel Soprano-Loto. Generalized max-weight policies in stochastic matching. *Stochastic Systems*, 13(1):40–58, 2023.
- [JM22] Jiashuo Jiang, Will Ma, and Jiawei Zhang. Degeneracy is ok: Logarithmic regret for network revenue management with indiscrete distributions. *arXiv preprint arXiv:2210.07996*, 2022.
- [KAG24] Süleyman Kerimov, Itai Ashlagi, and Itai Gurvich. Dynamic matching: Characterizing and achieving constant regret. *Management Science*, 70(5):2799–2822, 2024.
- [KAG25] Süleyman Kerimov, Itai Ashlagi, and Itai Gurvich. On the optimality of greedy policies in dynamic matching. *Operations Research*, 73(1):560–582, 2025.
- [Kan21] Yash Kanoria. Dynamic spatial matching. *arXiv preprint arXiv:2105.07329*, 2021.
- [KSSW22] Kristen Kessel, Ali Shameli, Amin Saberi, and David Wajc. The stationary prophet inequality problem. In *Proceedings of the 23rd ACM Conference on Economics and Computation*, pages 243–244, 2022.

- [LVY26] Yingxi Li, Ellen Vitercik, and Mingwei Yang. Smoothed analysis of online metric matching with a single sample: Beyond metric distortion. In *ITCS*, volume 362 of *LIPIcs*, pages 94:1–94:23. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2026.
- [LXK⁺11] Yi Lu, Qi Xie, Gabriel Kliot, Alan Geller, James R. Larus, and Albert Greenberg. Join-idle-queue: A novel load balancing algorithm for dynamically scalable web services. *Performance Evaluation*, 68(11):1056–1071, 2011.
- [MC89] Nimrod Megiddo and Ramaswamy Chandrasekaran. On the ϵ -perturbation method for avoiding degeneracy. *Operations Research Letters*, 8(6):305–308, 1989.
- [Mit01] Michael Mitzenmacher. The power of two choices in randomized load balancing. *IEEE Transactions on Parallel and Distributed Systems*, 12(10):1094–1104, 2001.
- [MM16] Jean Mairesse and Pascal Moyal. Stability of the stochastic matching model. *Journal of Applied Probability*, 53(4):1064–1077, 2016.
- [MP17] Pascal Moyal and Ohad Perry. On the instability of matching queues. 2017.
- [ÖW20] Erhun Özkan and Amy R Ward. Dynamic matching for real-time ride sharing. *Stochastic Systems*, 10(1):29–70, 2020.
- [RM21] Youssef Rahme and Pascal Moyal. A stochastic matching model on hypergraphs. *Advances in Applied Probability*, 53(4):951–980, 2021.
- [Rob03] Philippe Robert. Stochastic networks and queues. (*No Title*), 2003.
- [TVR98] Kalyan Talluri and Garrett Van Ryzin. An analysis of bid-price controls for network revenue management. *Management Science*, 44(11-part-1):1577–1593, 1998.
- [VB19] Alberto Vera and Siddhartha Banerjee. The bayesian prophet: A low-regret framework for online decision making. *ACM SIGMETRICS Performance Evaluation Review*, 47(1):81–82, 2019.
- [VBG21] Alberto Vera, Siddhartha Banerjee, and Itai Gurvich. Online allocation and pricing: Constant regret via bellman inequalities. *Operations Research*, 69(3):821–840, 2021.
- [Ver18] Roman Vershynin. *High-dimensional probability: An introduction with applications in data science*, volume 47. Cambridge university press, 2018.
- [WXY23] Yehua Wei, Jiaming Xu, and Sophie H Yu. Constant regret primal-dual policy for multi-way dynamic matching. *Available at SSRN 4357216*, 2023.
- [YY26] Mingwei Yang and Sophie H. Yu. Online metric matching: Beyond the worst case. *Operations Research*, 74(1):130–140, 2026.

A Preliminaries

A.1 Martingale and concentration inequalities

Lemma 6 (Doob's L^2 maximal inequality). *Let X_1, \dots, X_T be a discrete-time submartingale. Then,*

$$\mathbb{E} \left[\sup_{0 \leq s \leq T} |X_s|^2 \right] \leq 4\mathbb{E} [|X_T|^2] .$$

Lemma 7 (McDiarmid's Inequality). *Let $X_1, \dots, X_n \in \mathcal{X}$ be independent random variables. Suppose that function $f : \mathcal{X}^n \rightarrow \mathbb{R}$ satisfies the bounded differences property with bounds c_1, \dots, c_n : For all $i \in [n]$ and $x \in \mathcal{X}^n$,*

$$\sup_{x'_i \in \mathcal{X}} |f(x_1, \dots, x_{i-1}, x_i, x_{i+1}, \dots, x_n) - f(x_1, \dots, x_{i-1}, x'_i, x_{i+1}, \dots, x_n)| \leq c_i .$$

Then, for any $\delta > 0$,

$$\mathbb{P} (|f(X_1, \dots, X_n) - \mathbb{E}[f(X_1, \dots, X_n)]| \geq \delta) \leq 2 \exp \left(-\frac{2\delta^2}{\sum_{i=1}^n c_i^2} \right) .$$

Lemma 8. *Let $(X(k))_{k \geq 1}$ be i.i.d. n -dimensional random vectors with a categorical distribution with parameters λ , i.e., $\mathbb{P}(X(1) = e_i) = \lambda_i$, where $\sum_{i=1}^n \lambda_i = 1$ and $0 < \lambda_i < 1$ for $i \in [n]$. For any $T \geq 1$, define*

$$Z(T) \triangleq \frac{1}{\sqrt{T}} \sum_{i=1}^n \max_{1 \leq t \leq T} \left| \sum_{k=1}^t X_i(k) - \lambda_i t \right| .$$

Then, for all $\xi > 0$, there exist constants $C_1, C_2(\xi), C_3(\xi) < \infty$ such that

$$\sup_{T \geq 1} \mathbb{E} [Z(T)] \leq C_1 \triangleq 2\sqrt{n} , \tag{20}$$

$$\sup_{T \geq 1} \mathbb{E} [\exp (\xi Z(T))] \leq C_2(\xi) \triangleq \exp (2\sqrt{n}\xi + \sigma^2 \xi^2) , \tag{21}$$

$$\sup_{T \geq 1} \mathbb{E} \left[(Z(T))^2 \cdot \exp (\xi Z(T)) \right] \leq C_3(\xi) \triangleq 4e^{-2n} \cdot C_2 \left(\xi + \frac{1}{\sqrt{n}} \right) , \tag{22}$$

where $\sigma > 0$ is an absolute constant.

Proof. We prove (20), (21) and (22) sequentially.

Proof of (20) For any $t \geq 1$ and $i \in [n]$, let

$$Y_i(t) \triangleq \sum_{k=1}^t X_i(k) - \lambda_i t ,$$

then the process $\{Y_i(t)\}_{t \geq 0}$ is a martingale with increments $\Delta Y_i(k) = X_i(k) - \lambda_i$. For each $T \geq 1$, define the coordinate-wise maximum

$$M_i(T) \triangleq \max_{0 \leq t \leq T} |Y_i(t)|$$

and then $Z(T) = T^{-1/2} \sum_{i=1}^n M_i(T)$.

By Lemma 6, we have

$$\mathbb{E}[M_i(T)^2] \leq 4\mathbb{E}[Y_i(T)^2] = 4\text{Var}(Y_i(T)) = 4T\lambda_i(1 - \lambda_i).$$

By Jensen's inequality, we obtain

$$\mathbb{E}[M_i(T)] \leq \sqrt{4T\lambda_i(1 - \lambda_i)} \leq 2\sqrt{T}\sqrt{\lambda_i}.$$

Then, we have

$$\mathbb{E}[Z(T)] = T^{-1/2} \mathbb{E} \left[\sum_{i=1}^n M_i(T) \right] \leq 2 \sum_{i=1}^n \sqrt{\lambda_i} \leq 2\sqrt{n},$$

where the second inequality holds because $\sum \lambda_i = 1, \lambda_i \geq 0$, and by Cauchy-Schwarz, $\sum_{i=1}^n \sqrt{\lambda_i}$ is maximized when $\lambda_i = 1/n$ for all $i \in [n]$.

Proof of (21) We analyze $Z(T)$ as a function of the T independent random vectors. Define the function $f : (\{e_1, \dots, e_n\})^T \rightarrow \mathbb{R}$ by

$$f(x(1), \dots, x(T)) = \frac{1}{\sqrt{T}} \sum_{i=1}^n \max_{1 \leq t \leq T} \left| \sum_{k=1}^t x_i(k) - \lambda_i t \right|,$$

where $x_i(k)$ denotes the i -th component of the vector $x(k)$. Note that $Z(T) = f(X(1), \dots, X(T))$.

We check the bounded difference condition required by Lemma 7. Consider two sequences of vectors $\mathbf{x} = (x(1), \dots, x(T))$ and $\mathbf{x}' = (x'(1), \dots, x'(T))$ that differ only at a single index $\ell \in [T]$. Since the vectors are canonical basis vectors, let $x(\ell) = e_u$ and $x'(\ell) = e_v$ for some distinct $u, v \in [n]$. Let $r_i(t) \triangleq \sum_{k=1}^t x_i(k)$ and $r'_i(t) \triangleq \sum_{k=1}^t x'_i(k)$ denote the partial sums for the two sequences. The change at index ℓ affects the partial sums only for $t \geq \ell$. Specifically:

- The u -th coordinate decreases by 1: $r'_u(t) = r_u(t) - 1$ for all $t \geq \ell$.
- The v -th coordinate increases by 1: $r'_v(t) = r_v(t) + 1$ for all $t \geq \ell$.
- All other coordinates $i \notin \{u, v\}$ remain unchanged: $r'_i(t) = r_i(t)$ for all t .

Applying this to the u -th component term in f :

$$\left| \max_{1 \leq t \leq T} |r_u(t) - \lambda_u t| - \max_{1 \leq t \leq T} |r'_u(t) - \lambda_u t| \right| \leq \max_{\ell \leq t \leq T} |r_u(t) - r'_u(t)| = 1.$$

Similarly,

$$\left| \max_{1 \leq t \leq T} |r_v(t) - \lambda_v t| - \max_{1 \leq t \leq T} |r'_v(t) - \lambda_v t| \right| \leq 1.$$

The terms for $i \notin \{u, v\}$ do not change. Therefore, the total change in the function value is bounded by:

$$|f(\mathbf{x}) - f(\mathbf{x}')| \leq \frac{1}{\sqrt{T}} \sum_{i=1}^n \left| \max_t |r_i(t) - \lambda_i t| - \max_t |r'_i(t) - \lambda_i t| \right| \leq \frac{1}{\sqrt{T}} (1 + 1) = \frac{2}{\sqrt{T}}.$$

Thus, f satisfies the bounded difference condition with constants $c_k = 2/\sqrt{T}$ for each $k \in [T]$. The sum of squared constants is:

$$\sum_{k=1}^T c_k^2 = \sum_{k=1}^T \frac{4}{T} = 4.$$

By Lemma 7, for any $\delta > 0$:

$$\mathbb{P}(|Z(T) - \mathbb{E}[Z(T)]| \geq \delta) \leq 2 \exp\left(-\frac{2\delta^2}{\sum_{k=1}^T c_k^2}\right) = 2 \exp\left(-\frac{\delta^2}{2}\right).$$

Applying [Ver18, Proposition 2.6.1], which relates several equivalent sub-Gaussian properties, to the random variable $Z(T) - \mathbb{E}[Z(T)]$, we get that there exists an absolute constant $\sigma > 0$ such that

$$\mathbb{E}\left[e^{\delta(Z(T) - \mathbb{E}[Z(T)])}\right] \leq e^{\sigma^2 \delta^2}, \quad \forall \delta \in \mathbb{R}. \quad (23)$$

Then, it follows that

$$\sup_{T \geq 1} \mathbb{E}\left[e^{\xi Z(T)}\right] \leq \sup_{T \geq 1} e^{\xi \mathbb{E}[Z(T)]} \cdot \mathbb{E}\left[e^{\xi(Z(T) - \mathbb{E}[Z(T)])}\right] \leq \exp(2\sqrt{n}\xi + \sigma^2 \xi^2),$$

where the second inequality follows from (20) and (23), concluding $C_2(\xi) = \exp(2\sqrt{n}\xi + \sigma^2 \xi^2)$.

Proof of (22) We use the inequality $x^2 \leq \frac{4e^{-2}}{\alpha^2} e^{\alpha x}$, valid for all $x \geq 0, \alpha > 0$. Applying this with $x = Z(T)$ and choosing $\alpha = 1/\sqrt{n}$:

$$(Z(T))^2 e^{\xi Z(T)} \leq \frac{4e^{-2}}{(1/\sqrt{n})^2} e^{Z(T)/\sqrt{n}} e^{\xi Z(T)} = 4e^{-2} n e^{(\xi + 1/\sqrt{n})Z(T)}.$$

Taking expectations and applying (21) with parameter $\xi + 1/\sqrt{n}$:

$$\mathbb{E}\left[(Z(T))^2 e^{\xi Z(T)}\right] \leq 4e^{-2} n \mathbb{E}\left[e^{(\xi + 1/\sqrt{n})Z(T)}\right] \leq 4e^{-2} n C_2(\xi + 1/\sqrt{n}).$$

Hence, it follows that $C_3(\xi) = 4e^{-2} n C_2(\xi + 1/\sqrt{n})$. \square

A.2 Lyapunov function analysis

In this subsection, we review standard tools on Lyapunov function analysis, which is one of the prevalent approaches to bound the expected value of some function with respect to the steady state of a Markov chain.

The following lemma bounds the expected value of a function $f(\cdot)$ when the drift of the Lyapunov function depends also on $f(\cdot)$.

Lemma 9 (Corollary 4 in [GZ08]). *Let $X = (X(t))_{t \geq 0}$ be a discrete-time Markov chain on a discrete state space \mathcal{X} with transition kernel P , and suppose $f : \mathcal{X} \rightarrow \mathbb{R}_{\geq 0}$. If there exists a function $\Phi : \mathcal{X} \rightarrow \mathbb{R}_{\geq 0}$ and a constant c for which*

$$\int_{\mathcal{X}} P(x, dy) \Phi(y) - \Phi(x) \leq -f(x) + c, \quad \forall x \in \mathcal{X},$$

then

$$\int_{\mathcal{X}} \pi(dx) f(x) \leq c$$

for every stationary distribution π of X .

We will also apply the following lemma, which directly bounds the expectation of the Lyapunov function at all times instead of only under the stationary distribution.

Lemma 10 (Lemma 5 in [WXY23]). *Let $\Psi(t)$ be an $\{\mathcal{F}_t\}$ -adapted discrete-time stochastic process satisfying:*

- *Bounded variation:* $|\Psi(t+1) - \Psi(t)| \leq K$;
- *Expected decrease:* $\mathbb{E}[\Psi(t+1) - \Psi(t) \mid \mathcal{F}_t] \leq -\eta$, *when* $\Psi(t) \geq B$;
- $\Psi(0) \leq K + B$.

Then, for every $t \geq 0$, we have

$$\mathbb{E}[\Psi(t)] \leq K \left(1 + \left\lceil \frac{B}{K} \right\rceil\right) + K \left(\frac{K - \eta}{2\eta}\right).$$

A.3 Markov chain

The following lemma introduces a generic way of establishing the ergodicity of Markov chains.

Lemma 11 (Corollary 8.7 in [Rob03]). *Let $(X(t))_{t \geq 0}$ be a discrete-time, homogeneous, irreducible and aperiodic Markov chain on a countable state space \mathcal{X} . If there exist a function $f : \mathcal{X} \rightarrow \mathbb{R}_+$ and constants $K, \eta > 0$ such that*

- (i) $\mathbb{E}_x[f(X(1)) - f(x)] \leq -\eta$ *when* $f(x) > K$,
- (ii) $\mathbb{E}_x[f(X(1))] < \infty$ *when* $f(x) \leq K$, *and*
- (iii) *the set $\{x \in \mathcal{X} \mid f(x) \leq K\}$ is finite,*

then the Markov chain $(X(t))_{t \geq 0}$ is ergodic.

A.4 Consistent policy

In this subsection, we investigate the consistency property (Definition 4) of existing policies. Recall that [MP17, Lemma 4] establish consistency for certain static priority policies⁵ when no agents are discarded. Also, [KAG25, Lemma 5.3] claim that all (deterministic) policies that perform some match whenever possible are consistent, which is not accurate as we will demonstrate. Then, we show that all static priority policies and the longest-queue policy are consistent.

We start with an example to refute [KAG25, Lemma 5.3].

Example 2. Assume that n is sufficiently large and the graph $(\mathcal{A}, \mathcal{M})$ forms a path such that for every $i \in [n - 1]$, there is a match in \mathcal{M} containing types i and $i + 1$. To describe a policy that performs some match whenever possible, it suffices to specify, under a state q where any two adjacent queues are not both non-empty, which match to perform when an agent of type $i > 1$ arrives with queues $i - 1$ and $i + 1$ being non-empty at the same time. In this case, our policy Π decides the match as follows: If $q_1 = 0$, then Π performs the match $m(i - 1, i)$; otherwise, Π performs the match $m(i, i + 1)$. In other words, when queue 1 is empty, Π prioritizes matching to the queue with a smaller index; otherwise, Π prioritizes matching to the queue with a larger index.

⁵They consider the static priority policies that perform a match whenever possible and select matches according to a fixed priority order over all matches.

Next, we give the initial states and specify the arrival sequence. Let $Q(0) = \mathbf{0}$ and $Q'(0) = (1, 0, \dots, 0)$. The first four arriving agents are of types 3, 5, 4, 6, respectively. During the first four periods, Π will perform $m(3, 4)$ and $m(5, 6)$ under Q , and perform $m(4, 5)$ under Q' . Hence, $\|Q(4)\|_1 = 0$ and $\|Q'(4)\|_1 = 3$, which implies

$$\|Q(4) - Q'(4)\|_1 = 3 > 1 = \|Q(0) - Q'(0)\|_1.$$

Therefore, Π is not consistent.

We remark that, in the above example, one can repeat the arrival pattern to make the distance between $Q(t)$ and $Q'(t)$ arbitrarily large as t increases. That is, the next four arriving agents are of types respectively 8, 10, 9, 11, and so on.

Recall that the longest-queue policy adopts the following matching rule: When the arriving agent has multiple non-empty neighboring queues, select the one with the largest length to match to, with ties broken according to a fixed order.⁶ [KAG25] show that the longest-queue policy achieves the optimal regret scaling for general networks. Next, we show that all static priority policies and the longest-queue policy are consistent.

Proposition 7. *All static priority policies and the longest-queue policy are consistent.*

Proof. We first show that any static priority policy is consistent. Let $Q(0)$ and $Q'(0)$ be two valid initial states, and let $Q(1)$ and $Q'(1)$ be the corresponding states induced by a static priority policy Π after one-step transition. Suppose that an agent in queue i arrives at time 1.

- If the arriving agent is not matched under both $Q(0)$ and $Q'(0)$, or is matched to an agent in the same queue under both $Q(0)$ and $Q'(0)$, then $Q(1) - Q'(1) = Q(0) - Q'(0)$, implying

$$\|Q(1) - Q'(1)\|_1 = \|Q(0) - Q'(0)\|_1.$$

- If the arriving agent is matched to an agent in queue j under $Q(0)$ and is not matched under $Q'(0)$, then $Q(1) = Q(0) - \mathbf{e}_j$ and $Q'(1) = Q'(0) + \mathbf{1}_{\{i \in \mathcal{A}_0\}} \cdot \mathbf{e}_i$, and we must have $Q_j(0) > 0$ and $Q'_j(0) = 0$. As a result,

$$\|Q(1) - Q'(1)\|_1 = \|Q(0) - \mathbf{e}_j - Q'(0) - \mathbf{1}_{\{i \in \mathcal{A}_0\}} \cdot \mathbf{e}_i\|_1 \leq \|Q(0) - Q'(0)\|_1,$$

where the inequality holds since $|Q_i(0) - Q'_i(0) - \mathbf{1}_{\{i \in \mathcal{A}_0\}}| \leq |Q_i(0) - Q'_i(0)| + 1$ and $|Q_j(0) - Q'_j(0) - 1| = |Q_j(0) - Q'_j(0)| - 1$ given that $Q_j(0) > 0$ and $Q'_j(0) = 0$. The case where the arriving agent is not matched under $Q(0)$ and is matched under $Q'(0)$ can be handle analogously.

- Suppose that the arriving agent is match to an agent in queue j under $Q(0)$ and is matched to an agent in queue k under $Q'(0)$ with $j \neq k$, then $Q(1) = Q(0) - \mathbf{e}_j$ and $Q'(1) = Q'(0) - \mathbf{e}_k$. Assume by symmetry that $m(i, j) \succ_i m(i, k)$, then we must have $Q_j(0) > 0$ and $Q'_j(0) = 0$. As a result,

$$\|Q(1) - Q'(1)\|_1 = \|Q(0) - \mathbf{e}_j - Q'(0) + \mathbf{e}_k\|_1 \leq \|Q(0) - Q'(0)\|_1,$$

where the inequality holds since $|Q_k(0) - Q'_k(0) + 1| \leq |Q_k(0) - Q'_k(0)| + 1$ and $|Q_j(0) - Q'_j(0) - 1| = |Q_j(0) - Q'_j(0)| - 1$ given that $Q_j(0) > 0$ and $Q'_j(0) = 0$.

⁶In [KAG25], when there are multiple longest neighboring queues, the longest-queue policy breaks ties arbitrarily. However, breaking ties arbitrarily may render the longest-queue policy inconsistent. To illustrate, under two different states with identical sets of longest neighboring queues, if the policy breaks ties differently, then consistency would be violated.

Combining all the above concludes that Π is consistent.

We now turn to the longest-queue policy, whose proof follows a similar argument. In particular, the analysis on the first two cases proceeds identically, and we conclude the consistency of the longest-queue policy by analyzing the last case.

- Suppose that the arriving agent is match to an agent in queue j under $Q(0)$ and is matched to an agent in queue k under $Q'(0)$ with $j \neq k$, then $Q(1) = Q(0) - \mathbf{e}_j$ and $Q'(1) = Q'(0) - \mathbf{e}_k$. By the matching rule of the longest-queue policy,

$$Q_j(0) \geq Q_k(0) \quad \text{and} \quad Q'_j(0) \leq Q'_k(0); \quad (24)$$

moreover, one of the inequalities must be strict since the longest-queue policy breaks ties according to a fixed order. Notice that either $Q_j(0) > Q'_j(0)$ or $Q'_k(0) > Q_k(0)$, as otherwise,

$$Q'_j(0) \leq Q'_k(0) \leq Q_k(0) \leq Q_j(0) \leq Q'_j(0),$$

where the first and the third inequalities hold by (24), contradicting the fact that at least one inequality in (24) is strict. Assume by symmetry that $Q_j(0) > Q'_j(0)$, and it follows that

$$\|Q(1) - Q'(1)\|_1 = \|Q(0) - \mathbf{e}_j - Q'(0) + \mathbf{e}_k\|_1 \leq \|Q(0) - Q'(0)\|_1,$$

where the inequality holds since $|Q_k(0) - Q'_k(0) + 1| \leq |Q_k(0) - Q'_k(0)| + 1$ and $|Q_j(0) - Q'_j(0) - 1| = |Q_j(0) - Q'_j(0)| - 1$ given that $Q_j(0) > Q'_j(0)$.

□

A.4.1 Proof of Lemma 2

Recall that $(Q(t))_{t \geq 0}$ are the states induced by Π starting from $Q(0) = \mathbf{0}$. Let $(Q'(t))_{t \geq 0}$ be the states induced by Π when we start from $Q'(0) \sim \pi$. Fix $t \geq 0$. Since Π is consistent, by induction, there exists a coupling μ between $Q(t)$ and $Q'(t)$ such that

$$\mathbb{E}_{(Q(t), Q'(t)) \sim \mu} [\|Q(t) - Q'(t)\|_1] \leq \mathbb{E} [\|Q(0) - Q'(0)\|_1].$$

Hence,

$$\begin{aligned} \mathbb{E} [\|Q(t)\|_1] - \mathbb{E} [\|Q'(t)\|_1] &= \mathbb{E}_{(Q(t), Q'(t)) \sim \mu} [\|Q(t)\|_1 - \|Q'(t)\|_1] \\ &\leq \mathbb{E}_{(Q(t), Q'(t)) \sim \mu} [\|Q(t) - Q'(t)\|_1] \\ &\leq \mathbb{E} [\|Q(0) - Q'(0)\|_1] = \mathbb{E} [\|Q'(0)\|_1]. \end{aligned}$$

Therefore,

$$\mathbb{E} [\|Q(t)\|_1] \leq \mathbb{E} [\|Q'(t)\|_1] + \mathbb{E} [\|Q'(0)\|_1] = 2\mathbb{E} [\|Q'(0)\|_1] \leq 2B,$$

where the equality holds since both $Q'(0)$ and $Q'(t)$ follow the stationary distribution π .

B Discussion on tree priority policy on acyclic graphs

In this section, we discuss the possibility of improving the regret bound in Theorem 2 for **TP** on acyclic graphs. In Section B.1, we analyze the all-time regret of **TP** on a path of at most 4 nodes to demonstrate the barrier of directly improving the proof of Theorem 2. In Section B.2, we leverage an alternative approach to derive optimal all-time regret for **TP** on a path of 4 nodes, indicating that our Lyapunov analysis for Theorem 2 is not tight.

Since **TP** always matches the arriving agent whenever it has a non-empty neighboring queue, as stated in the following fact, any two adjacent queues cannot be non-empty at the same time.

Fact 2. Let $(Q(t))_{t \geq 0}$ be the states induced by **TP** under an arbitrary sample path. Then, for all $t \geq 0$ and $m(i, j) \in \mathcal{M}$, $Q_i(t) \cdot Q_j(t) = 0$.

Let $\mathcal{A}_{\text{odd}} \subseteq \mathcal{A}$ be the set of nodes with an odd depth, and let $\mathcal{A}_{\text{even}} \triangleq \mathcal{A} \setminus \mathcal{A}_{\text{odd}}$ be the set of nodes with an even depth. We first show in the following lemma that truncating a subset of queues with an even (resp. odd) depth would not decrease the length of each queue with an odd (resp. even) depth and would not increase the length of each other queue with an even (resp. odd) depth.

Lemma 12. Suppose that $(\mathcal{A}, \mathcal{M})$ is acyclic, and consider a static priority policy. For $\mathcal{A}' \subseteq \mathcal{A}_{\text{even}}$ (resp. $\mathcal{A}' \subseteq \mathcal{A}_{\text{odd}}$), let $(Q(t))_{t \geq 0}$ be the states of the original system \mathcal{S} , and let $(Q'(t))_{t \geq 0}$ be the states of a new system \mathcal{S}' where each queue $i \in \mathcal{A}'$ is truncated. If both systems have the same arrival process, then for all sample path and $t \geq 0$, $Q_i(t) \leq Q'_i(t)$ for every $i \in \mathcal{A}_{\text{odd}}$ (resp. $i \in \mathcal{A}_{\text{even}}$), and $Q_i(t) \geq Q'_i(t)$ for every $i \in \mathcal{A}_{\text{even}}$ (resp. $i \in \mathcal{A}_{\text{odd}}$).

Proof. We only prove the statement for truncating even-depth queues, and the statement for truncating odd-depth queues can be established analogously. The statement is true for $t = 0$ since $Q_i(0) = Q'_i(0)$ for every $i \in \mathcal{A}$. Assume for induction that the statement is true for time $t \geq 0$, and we show that it is also true for time $t + 1$ by inspecting the queues that are updated in either systems. Let i be the queue that the arriving agent at time $t + 1$ belongs to, and let $\mathcal{N}' \subseteq \mathcal{N}(i)$ be the set of neighbors j of i such that $m(i, j) \in \mathcal{M}(i)$. Note that the depth of each node in \mathcal{N}' and the depth of i have different parities.

- If $i \in \mathcal{A}_{\text{odd}}$, and no match is performed in \mathcal{S} , then $Q_j(t) = 0$ for every $j \in \mathcal{N}'$. By the inductive hypothesis, $Q'_j(t) \leq Q_j(t) = 0$ for every $j \in \mathcal{N}'$, and hence no match is performed in \mathcal{S}' . Consequently, $Q_i(t + 1) = Q_i(t) + 1 \leq Q'_i(t) + 1 = Q'_i(t + 1)$.
- Suppose that $i \in \mathcal{A}_{\text{odd}}$, and a match $m(i, j)$ is performed in \mathcal{S} for some $j \in \mathcal{N}$. Firstly, if no match is performed in \mathcal{S}' , which implies $Q'_j(t + 1) = Q'_j(t) = 0$, then $Q_i(t + 1) = Q_i(t) \leq Q'_i(t) = Q'_i(t + 1) - 1$, and $Q_j(t + 1) \geq 0 = Q'_j(t + 1)$. Moreover, if a match $m(i, j)$ is performed in \mathcal{S}' , then $Q_i(t + 1) = Q_i(t) \leq Q'_i(t) = Q'_i(t + 1)$, and $Q_j(t + 1) = Q_j(t) - 1 \geq Q'_j(t) - 1 = Q'_j(t + 1)$. Finally, if a match $m(i, k)$ is performed in \mathcal{S}' for some $k \in \mathcal{N}' \setminus \{j\}$, we must have $m(i, j) \succ_i m(i, k)$ and $Q'_j(t + 1) = Q'_j(t) = 0$ since $Q_k(t) \geq Q'_k(t)$; it follows that $Q_i(t + 1) = Q_i(t) \leq Q'_i(t) = Q'_i(t + 1)$, $Q_j(t + 1) \geq 0 = Q'_j(t + 1)$, and $Q_k(t + 1) = Q_k(t) \geq Q'_k(t) \geq Q'_k(t + 1)$.
- If $i \in \mathcal{A}_{\text{even}}$, and no match is performed in \mathcal{S}' , then we must have $Q_j(t) \leq Q'_j(t) = 0$ for every $j \in \mathcal{N}'$ by the inductive hypothesis, which implies that no match is performed in \mathcal{S} . Hence, $Q_i(t + 1) = Q_i(t) + 1 \geq Q'_i(t) + 1 \geq Q'_i(t + 1)$.
- Suppose that $i \in \mathcal{A}_{\text{even}}$, and a match $m(i, j)$ is performed in \mathcal{S}' for some $j \in \mathcal{N}'$. Firstly, if no match is performed in \mathcal{S} , which implies $Q_j(t + 1) = Q_j(t) = 0$, then $Q'_i(t + 1) = Q'_i(t) \leq Q_i(t) \leq$

$Q_i(t+1)$, and $Q'_j(t+1) \geq 0 = Q_j(t+1)$. Moreover, if a match $m(i, j)$ is performed in \mathcal{S} , then $Q'_i(t+1) = Q'_i(t) \leq Q_i(t) = Q_i(t+1)$, and $Q'_j(t+1) = Q'_j(t) - 1 \geq Q_j(t) - 1 = Q_j(t+1)$. Finally, if a match $m(i, k)$ is performed in \mathcal{S} for some $k \in \mathcal{N}' \setminus \{j\}$, we must have $m(i, j) \succ_i m(i, k)$ and $Q_j(t+1) = Q_j(t) = 0$ since $Q'_k(t) \geq Q_k(t)$; it follows that $Q'_i(t+1) = Q'_i(t) \leq Q_i(t) = Q_i(t+1)$, $Q_j(t+1) \geq 0 = Q'_j(t+1)$, and $Q'_k(t+1) = Q'_k(t) \geq Q_k(t) = Q'_k(t+1) + 1$.

Therefore, the statement holds for time $t + 1$, which completes the proof. \square

B.1 Tightness of regret analysis on paths

In this subsection, we assume $(\mathcal{A}, \mathcal{M})$ forms a path with at least four nodes, and we show how to bound the expected queue lengths for the three nodes farthest away from the root. Combining with Lemma 1, this implies a regret upper bound consistent with Theorem 2 when the path consists of at most four nodes. In particular, we will iteratively bound the queue lengths in a bottom-up manner starting from the leaf node, and our analysis clearly illustrates the necessity of the dependence on network depth in the regret bound achieved by our current approach.

Let $\mathcal{A} = \{1, 2, \dots, n\}$ and $\mathcal{M} = \{1, 2, \dots, n-1\}$, where $j \in \mathcal{M}$ denote the match of between j and $j+1$. Here, we consider the case when n is the under-demanded node with $s_n^* > 0$ and $\mathcal{A}_+ = \{n\}$. Then, the root of $(\mathcal{A}, \mathcal{M})$ is node n by construction. We illustrate the constructed matching network in Figure 6.

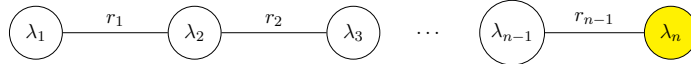


Figure 6: A path network where $\mathcal{A}_+ = \{n\}$.

Note that under **TP**, the priority orders $(\succ_i)_{i \in [n]}$ is unique: for all $i \in \{2, \dots, n-1\}$, $m(i-1, i) \succ_i m(i, i+1)$. We introduce the following family of (artificial) systems $\{\mathcal{S}_i\}_{1 \leq i < n}$, where we truncate queues $i+1, \dots, n$ in \mathcal{S}_i , and construct a coupling with our original system such that each system is equipped with the same arrival process as in the original system. Under \mathcal{S}_i , denote the number of agents of type j in the queue at the end of time t by $\bar{Q}_j^i(t)$ for $1 \leq j \leq n$. Note that for all $t \geq 0$ and $1 \leq j \leq i$, we have

$$\bar{Q}_j^i(t) = A_j(t) - \bar{D}_{j-1}^i(t) \mathbf{1}_{\{j>1\}} - \bar{D}_j^i(t) \mathbf{1}_{\{j<n\}}, \quad (25)$$

where for any $1 \leq j \leq n-1$, $\bar{D}_j^i(t)$ denotes the number of matches of $m(j, j+1)$ performed up to and including time t in \mathcal{S}_i , and we set $\bar{D}_0^i(t) = \bar{D}_n^i(t) = 0$. In particular, by the property of \mathcal{S}_i , $\bar{Q}_j^i(t) = 0$ for any $j \in \{i+1, \dots, n\}$.

The following corollary, as a direct consequence of Lemma 12, characterizes an alternating pattern in queue-length comparisons between each \mathcal{S}_i and the original system under **TP**.

Corollary 1. *Under **TP**, for any $t \geq 0$ and $1 \leq i < n$, if i is odd, then*

$$\bar{Q}_{2m}^i(t) \leq Q_{2m}(t), \quad \bar{Q}_{2m+1}^i(t) \geq Q_{2m+1}(t), \quad \forall 0 \leq 2m \leq i-1; \quad (26)$$

if i is even, then

$$\bar{Q}_{2m}^i(t) \geq Q_{2m}(t), \quad \bar{Q}_{2m+1}^i(t) \leq Q_{2m+1}(t), \quad \forall 0 \leq 2m \leq i. \quad (27)$$

Proof. For $i = n-1$, the statement follows directly from Lemma 12. For $i \in \{1, \dots, n-2\}$, since $m(i, i+1) \succ_i m(i+1, i+2)$, provided that queue $i+1$ is truncated, whether truncating queues $i+2, \dots, n$ or not does not affect the lengths of queues $1, \dots, i$. Hence, we can apply Lemma 12 to the system where only queue $i+1$ is truncated to conclude the statement. \square

Our goal is to use the coupling between our original system and \mathcal{S}_i 's to characterize the queue-lengths under **TP**. We also note that this coupling ensures the processes we are going to analyze to be Markovian; e.g., the process $(Q_1(t))_{t \geq 0}$ itself is not a Markov chain, since the transition probabilities depend on the state of queue 2, whereas $(\bar{Q}_1^1(t))_{t \geq 0}$ is a Markov chain. In general, by the construction of the artificial systems, $(\bar{Q}^i(t))_{t \geq 0}$ is a Markov chain for all $1 \leq i < n$.

We then introduce the following lemma that characterizes the optimal solution of the static planning problem $\text{SPP}(\lambda)$.

Lemma 13 (Theorem 4.1 in [KAG25]). *For any $1 \leq i < n$, $z_i^* + z_{i-1}^* \mathbf{1}_{\{i \geq 2\}} = \lambda_i$.*

Following Lemma 13, we get $z_1^* = \lambda_1$, $z_2^* = \lambda_2 - \lambda_1$, and $z_3^* = \lambda_3 - z_2^* = \lambda_3 - \lambda_2 + \lambda_1$, and note that $z_1^*, z_2^*, z_3^* \geq \epsilon$, which will be useful to prove the following results. Next, we discuss the intuition behind the construction of our Lyapunov functions. In \mathcal{S}_1 , we only need to focus on the length of queue 1, and we naturally adopt the quadratic Lyapunov function $\mathcal{L}(t) \triangleq (Q_1(t))^2$. When it comes to \mathcal{S}_2 , to achieve all-time constant regret via Lemma 1, we should have $D_1(t) \approx A_1(t)$ and $D_2(t) \approx A_2(t) - A_1(t)$. This implies that ideally we want both

$$A_1(t) - D_1(t) = Q_1(t) \quad \text{and} \quad A_2(t) - A_1(t) - D_2(t) = Q_2(t) - Q_1(t)$$

to be small. Hence, a natural choice of the Lyapunov function would be $\mathcal{L}(t) \triangleq \beta(Q_1(t))^2 + (Q_2(t) - Q_1(t))^2$ for appropriately chosen coefficient $\beta \geq 0$. However, we always have $Q_1(t) \cdot Q_2(t) = 0$ by 2, implying that we can safely drop the first term in $\mathcal{L}(t)$, giving rise to our final choice of the Lyapunov function for \mathcal{S}_2 . Similar derivations also lead to our construction of the Lyapunov function for \mathcal{S}_3 , which will become clear momentarily.

In the following two lemmas, we show that $\mathbb{E}[Q_1(t)]$ and $\mathbb{E}[Q_2(t)]$ can be bounded by $O(\epsilon^{-1})$ at all times respectively.

Lemma 14. $\mathbb{E}[Q_1(t)] \leq \epsilon^{-1}$ for all $t \geq 0$ under **TP**.

Proof. Consider the Lyapunov function $\mathcal{L}(t) \triangleq (\bar{Q}_1^1(t))^2$. Conditioned on $\mathcal{L}(t) > 0$, when an agent of type 1 arrives, we have $\bar{Q}_1^1(t+1) = \bar{Q}_1^1(t) + 1$; when an agent of type 2 arrives, match 1 is performed under **TP** and $\bar{Q}_1^1(t+1) = \bar{Q}_1^1(t) - 1$. Thus, for all $t \geq 0$, we have

$$\begin{aligned} \mathbb{E}[\mathcal{L}(t+1) - \mathcal{L}(t) \mid \bar{Q}_1^1(t), \mathcal{L}(t) > 0] &= \mathbb{E}[(\bar{Q}_1^1(t+1) + \bar{Q}_1^1(t))(\bar{Q}_1^1(t+1) - \bar{Q}_1^1(t)) \mid \bar{Q}_1^1(t), \mathcal{L}(t) > 0] \\ &\leq -2\bar{Q}_1^1(t)(\lambda_2 - \lambda_1) + 1, \end{aligned}$$

where $\lambda_2 - \lambda_1 = z_2^* \geq \epsilon$. Per Lemma 11, the Markov chain $(\bar{Q}_1^1(t))_{t \geq 0}$ is ergodic, and we denote its stationary distribution by π . Then by Lemma 9, we have

$$\mathbb{E}_\pi[\bar{Q}_1^1(0)] \leq \frac{1}{2(\lambda_2 - \lambda_1)} \leq \frac{1}{2\epsilon}.$$

Per Lemma 2 and Proposition 7, we have $\mathbb{E}[\bar{Q}_1^1(t)] \leq \epsilon^{-1}$ for all $t \geq 0$. Finally, it follows from 1 that $\mathbb{E}[Q_1(t)] \leq \epsilon^{-1}$ for all $t \geq 0$, since $Q_1(t) \leq \bar{Q}_1^1(t)$ for all $t \geq 0$. \square

Lemma 15. $\mathbb{E}[Q_2(t)] \leq 2\epsilon^{-1}$ for all $t > 0$ under **TP**.

Proof. Consider the Lyapunov function $\mathcal{L}(t) \triangleq (\bar{Q}_2^2(t) - \bar{Q}_1^2(t))^2$. Conditioned on $\mathcal{L}(t) > 0$, we can either have $\bar{Q}_1^2(t) > 0$ or $\bar{Q}_2^2(t) > 0$ by 2.

Claim 1. For all $t \geq 0$ and $i = 1, 2$, we have

$$\mathbb{E}[(\mathcal{L}(t+1) - \mathcal{L}(t) \mid \bar{Q}^2(t), \mathcal{L}(t) > 0, \bar{Q}_i^2(t) > 0] \leq -2\epsilon\bar{Q}_i^2(t) + 1. \quad (28)$$

The proof of Claim 1 is deferred to Appendix B.1.1. Using Claim 1, we have

$$\begin{aligned}\mathbb{E}[(\mathcal{L}(t+1) - \mathcal{L}(t) \mid \bar{Q}^2(t), \mathcal{L}(t) > 0)] &\leq -(2\epsilon\bar{Q}_1^2(t) + 1) \cdot \mathbf{1}_{\{\bar{Q}_1^2(t) > 0\}} - (2\epsilon\bar{Q}_2^2(t) + 1) \cdot \mathbf{1}_{\{\bar{Q}_2^2(t) > 0\}} \\ &\leq -2\epsilon|\bar{Q}_2^2(t) - \bar{Q}_1^2(t)| + 1,\end{aligned}$$

where the first inequality follows from Claim 1 and the second inequality follows from the fact that $\bar{Q}_1^2(t) \cdot \bar{Q}_2^2(t) = 0$ for all $t \geq 0$ by 2. Denote the stationary distribution of the Markov Chain $(\bar{Q}^2(t))_{t \geq 0}$ by π , which is granted by Lemma 11. Per Lemma 9, we have

$$\mathbb{E}_\pi[|\bar{Q}_2^2(0) - \bar{Q}_1^2(0)|] \leq \frac{1}{2\epsilon},$$

Per Lemma 2 and Proposition 7, we have $\mathbb{E}[|\bar{Q}_2^2(t) - \bar{Q}_1^2(t)|] \leq \epsilon^{-1}$ for all $t \geq 0$. Since $\bar{Q}_2^2(t) \geq Q_2(t)$ and $\bar{Q}_1^2(t) \leq Q_1(t)$ for all $t \geq 0$ by 1, together with Lemma 14, we get $\mathbb{E}[\bar{Q}_1^2(t)] \leq \epsilon^{-1}$ and $\mathbb{E}[Q_2(t)] \leq \mathbb{E}[\bar{Q}_2^2(t)] \leq 2\epsilon^{-1}$ for all $t \geq 0$. \square

Next, we upper bound $\mathbb{E}[Q_3(t)]$, for which we can no longer achieve the optimal scaling of $O(\epsilon^{-1})$.

Lemma 16. $\mathbb{E}[Q_3(t)] \leq O(\epsilon^{-2})$ for all $t \geq 0$ under **TP**.

Proof. Consider the following Lyapunov function

$$\mathcal{L}(t) \triangleq \beta_1 (\bar{Q}_1^3(t))^2 + \beta_2 (\bar{Q}_2^3(t) - \bar{Q}_1^3(t))^2 + (\bar{Q}_3^3(t) - \bar{Q}_2^3(t) + \bar{Q}_1^3(t))^2, \quad (29)$$

where we will determine $\beta_1, \beta_2 \in \mathbb{R}_{\geq 0}$ momentarily. Define $\mathcal{B}(t) \triangleq \{i \in \mathcal{A} \mid \bar{Q}_i^3(t) > 0\}$ as the set of non-empty queues at the end of time t . Define the following events $\mathcal{E}_1(t) \triangleq \{\mathcal{B}(t) = \{1\}\}$, $\mathcal{E}_2(t) \triangleq \{\mathcal{B}(t) = \{2\}\}$, $\mathcal{E}_3(t) \triangleq \{\mathcal{B}(t) = \{3\}\}$, and $\mathcal{E}_4(t) \triangleq \{\mathcal{B}(t) = \{1, 3\}\}$. Note that by 2, the union of these events forms a partition when $\mathcal{L}(t) > 0$. Next, we introduce the following claim on bounding $\mathbb{E}[\mathcal{L}(t+1) - \mathcal{L}(t) \mid \bar{Q}^3(t), \mathcal{L}(t) > 0, \mathcal{E}_i(t)]$ for $1 \leq i \leq 4$.

Claim 2. For all $t \geq 0$, we have

$$\begin{aligned}\mathbb{E}[\mathcal{L}(t+1) - \mathcal{L}(t) \mid \bar{Q}^3(t), \mathcal{L}(t) > 0, \mathcal{E}_1(t)] \\ \leq -2[(\beta_1 + \beta_2)(\lambda_2 - \lambda_1) - (\lambda_3 - \lambda_2 + \lambda_1)] \cdot |\bar{Q}_3^3(t) - \bar{Q}_2^3(t) + \bar{Q}_1^3(t)| + \beta_1 + \beta_2 + 1.\end{aligned}$$

Moreover, for all $k = 2, 3, 4$, and $t \geq 0$, we have

$$\mathbb{E}[\mathcal{L}(t+1) - \mathcal{L}(t) \mid \bar{Q}^3(t), \mathcal{L}(t) > 0, \mathcal{E}_k(t)] \leq -2\epsilon |\bar{Q}_3^3(t) - \bar{Q}_2^3(t) + \bar{Q}_1^3(t)| + 2(\beta_1 + \beta_2 + 1).$$

The proof of Claim 2 is deferred to Appendix B.1.2. Note that per Claim 2, $\mathcal{L}(t)$ has a negative drift under $\mathcal{E}_2(t), \mathcal{E}_3(t), \mathcal{E}_4(t)$ regardless of the choices of β_1 and β_2 , but the sign of the drift is unclear under $\mathcal{E}_1(t)$. Let $\delta \triangleq (\beta_1 + \beta_2)(\lambda_2 - \lambda_1) - (\lambda_3 - \lambda_2 + \lambda_1)$ be the coefficient in the drift of $\mathcal{L}(t)$ under $\mathcal{E}_1(t)$. In order to ensure that $\mathcal{L}(t)$ has a negative drift under $\mathcal{E}_1(t)$ as well, we will pick β_1 and β_2 to ensure that $\delta > 0$. Therefore, the overall drift is given by

$$\begin{aligned}\mathbb{E}[\mathcal{L}(t+1) - \mathcal{L}(t) \mid \bar{Q}^3(t), \mathcal{L}(t) > 0] &\leq -2 \min\{\delta, \epsilon\} |\bar{Q}_3^3(t) - \bar{Q}_2^3(t) + \bar{Q}_1^3(t)| + 2(\beta_1 + \beta_2 + 1) \\ &= -2 \min\{\delta, \epsilon\} |\bar{Q}_3^3(t) - \bar{Q}_2^3(t) + \bar{Q}_1^3(t)| + \frac{2(\delta + \lambda_3)}{\lambda_2 - \lambda_1}.\end{aligned}$$

It follows by Lemma 11 that the stationary distribution π of the Markov chain $(\bar{Q}^3(t))_{t \geq 0}$ exists. Moreover, by Lemma 9, and if we choose appropriate β_1 and β_2 such that $\delta = \epsilon$,

$$\mathbb{E}_\pi [|\bar{Q}_3^3(0) - \bar{Q}_2^3(0) + \bar{Q}_1^3(0)|] \leq \frac{2(\delta + \lambda_3)}{2(\lambda_2 - \lambda_1) \min\{\delta, \epsilon\}} = \frac{2(\epsilon + \lambda_3)}{2\epsilon(\lambda_2 - \lambda_1)} \leq O(\epsilon^{-2}), \quad (30)$$

where the last inequality holds since $\lambda_3 \leq 1$ and $\lambda_2 - \lambda_1 \geq \epsilon$. To translate the expected queue-length under the steady-state to that at all times, by Lemma 2 and Proposition 7, we have

$$\mathbb{E} [|\bar{Q}_3^3(t) - \bar{Q}_2^3(t) + \bar{Q}_1^3(t)|] \leq O(\epsilon^{-2})$$

for all $t \geq 0$. Moreover, by Lemma 15, we have $\mathbb{E}[Q_2(t)] \leq 2\epsilon^{-1}$ for all $t \geq 0$. Therefore, by 1, we conclude that $\mathbb{E}[Q_3(t)] \leq \mathbb{E}[\bar{Q}_3^3(t)] \leq O(\epsilon^{-2})$ for all $t \geq 0$. \square

Remark 2. Note that in the analysis of Lemma 16, we are unable to achieve the optimal scaling of $O(\epsilon^{-1})$ for the expected length of queue 3 in certain cases by using the generalized quadratic Lyapunov function defined in (29). Specifically, when $\lambda_3 = \Omega(1)$ and $\lambda_2 - \lambda_1 = O(\epsilon)$, the last inequality in (30) will be tight. Furthermore, when generalizing our analysis to matching networks with an arbitrary depth, similar situations would repetitively occur as the depth grows, indicating that it is inevitable for the resulting regret to depend on the depth.

B.1.1 Proof of Claim 1

. First, assume that $\bar{Q}_1^2(t) > 0$. If an agent of type 1 arrives, then we have

$$\begin{aligned} (\bar{Q}_2^2(t+1) - \bar{Q}_1^2(t+1)) - (\bar{Q}_2^2(t) - \bar{Q}_1^2(t)) &= -1, \\ (\bar{Q}_2^2(t+1) - \bar{Q}_1^2(t+1)) + (\bar{Q}_2^2(t) - \bar{Q}_1^2(t)) &= -2\bar{Q}_1^2(t) - 1. \end{aligned}$$

If an agent of type 2 arrives, then we have

$$\begin{aligned} (\bar{Q}_2^2(t+1) - \bar{Q}_1^2(t+1)) - (\bar{Q}_2^2(t) - \bar{Q}_1^2(t)) &= 1, \\ (\bar{Q}_2^2(t+1) - \bar{Q}_1^2(t+1)) + (\bar{Q}_2^2(t) - \bar{Q}_1^2(t)) &= -2\bar{Q}_1^2(t) + 1. \end{aligned}$$

And, if an agent of type 3 arrives, $\mathcal{L}(t+1) = \mathcal{L}(t)$. Thus, we have for all $t \geq 0$,

$$\begin{aligned} \mathbb{E} [(\mathcal{L}(t+1) - \mathcal{L}(t) \mid \bar{Q}^2(t), \mathcal{L}(t) > 0, \bar{Q}_1^2(t) > 0)] &\leq -2\bar{Q}_1^2(t)(\lambda_2 - \lambda_1) + \lambda_1 + \lambda_2 \\ &\leq -2\epsilon\bar{Q}_1^2(t) + 1, \end{aligned}$$

where we used the fact that $\lambda_2 - \lambda_1 \geq \epsilon$ and $\lambda_1 + \lambda_2 \leq 1$.

Now assume that $\bar{Q}_2^2(t) > 0$. If an agent of type 1 or 3 arrives, then we have

$$\begin{aligned} (\bar{Q}_2^2(t+1) - \bar{Q}_1^2(t+1)) - (\bar{Q}_2^2(t) - \bar{Q}_1^2(t)) &= -1, \\ (\bar{Q}_2^2(t+1) - \bar{Q}_1^2(t+1)) + (\bar{Q}_2^2(t) - \bar{Q}_1^2(t)) &= 2\bar{Q}_2^2(t) - 1, \end{aligned}$$

and if an agent of type 2 arrives, then we have

$$\begin{aligned} (\bar{Q}_2^2(t+1) - \bar{Q}_1^2(t+1)) - (\bar{Q}_2^2(t) - \bar{Q}_1^2(t)) &= 1, \\ (\bar{Q}_2^2(t+1) - \bar{Q}_1^2(t+1)) + (\bar{Q}_2^2(t) - \bar{Q}_1^2(t)) &= 2\bar{Q}_2^2(t) + 1. \end{aligned}$$

Thus, we have for all $t \geq 0$,

$$\begin{aligned} \mathbb{E} [(\mathcal{L}(t+1) - \mathcal{L}(t) \mid \bar{Q}^2(t), \mathcal{L}(t) > 0, \bar{Q}_2^2(t) > 0)] &\leq -2\bar{Q}_2^2(t)(\lambda_3 - \lambda_2 + \lambda_1) + \lambda_1 + \lambda_2 + \lambda_3 \\ &\leq -2\epsilon\bar{Q}_2^2(t) + 1, \end{aligned}$$

where we used the fact that $\lambda_3 - \lambda_2 + \lambda_1 \geq \epsilon$ and $\lambda_1 + \lambda_2 + \lambda_3 \leq 1$.

B.1.2 Proof of Claim 2

Under $\mathcal{E}_1(t)$, any arriving agent with types 1 or 3 increases $\bar{Q}_1^3(t)$ or $\bar{Q}_3^3(t)$ by 1, respectively. If the arriving agent is of type 2, $\bar{Q}_1^3(t)$ decreases by 1, and an arriving agent of type 4 does not affect the queue-lengths since $\bar{Q}_3^3(t) = 0$. Thus, we have

$$\begin{aligned} \mathbb{E}[\mathcal{L}(t+1) - \mathcal{L}(t) \mid \bar{Q}^3(t), \mathcal{L}(t) > 0, \mathcal{E}_1(t)] \\ \leq (-2\beta_1(\lambda_2 - \lambda_1)\bar{Q}_1^3(t) + \beta_1) + (-2\beta_2(\lambda_2 - \lambda_1)\bar{Q}_1^3(t) + \beta_2) + (2(\lambda_3 - \lambda_2 + \lambda_1)\bar{Q}_1^3(t) + 1) \\ = (-2(\beta_1 + \beta_2)(\lambda_2 - \lambda_1)\bar{Q}_1^3(t) + \beta_1 + \beta_2) + (2(\lambda_3 - \lambda_2 + \lambda_1)\bar{Q}_1^3(t) + 1) \\ = -2[(\beta_1 + \beta_2)(\lambda_2 - \lambda_1) - (\lambda_3 - \lambda_2 + \lambda_1)]|\bar{Q}_3^3(t) - \bar{Q}_2^3(t) + \bar{Q}_1^3(t)| + \beta_1 + \beta_2 + 1, \end{aligned}$$

where the last equality holds because under $\mathcal{E}_1(t)$, we have $\bar{Q}_3^3(t) = \bar{Q}_2^3(t) = 0$. Under $\mathcal{E}_2(t)$, any arriving agent of type 2 increases $\bar{Q}_2^3(t)$ by 1, and any arriving agent of types 1 or 3 decreases $\bar{Q}_2^3(t)$ by 1, while $\bar{Q}_1^3(t+1) = \bar{Q}_3^3(t+1) = 0$. An arriving agent of type 4 does not affect the queue-lengths. Thus, we have

$$\begin{aligned} \mathbb{E}[\mathcal{L}(t+1) - \mathcal{L}(t) \mid \bar{Q}^3(t), \mathcal{L}(t) > 0, \mathcal{E}_2(t)] &\leq -2(\beta_2 + 1)(\lambda_3 - \lambda_2 + \lambda_1)\bar{Q}_2^3(t) + \beta_2 + 1 \\ &\leq -2(\beta_2 + 1)\epsilon|\bar{Q}_3^3(t) - \bar{Q}_2^3(t) + \bar{Q}_1^3(t)| + \beta_2 + 1 \\ &\leq -2\epsilon|\bar{Q}_3^3(t) - \bar{Q}_2^3(t) + \bar{Q}_1^3(t)| + \beta_2 + 1. \end{aligned}$$

Under $\mathcal{E}_3(t)$, any arriving agent of type 1 increases $\bar{Q}_1^3(t)$ by 1, any arriving agent of type 2 decreases $\bar{Q}_3^3(t)$ by 1, any arriving agent of type 3 increases $\bar{Q}_3^3(t)$ by 1, and any arriving agent of type 4 decreases $\bar{Q}_3^3(t)$ by 1. Thus, we have

$$\begin{aligned} \mathbb{E}[\mathcal{L}(t+1) - \mathcal{L}(t) \mid \bar{Q}^3(t), \mathcal{L}(t) > 0, \mathcal{E}_3(t)] &\leq 2(\beta_1 + \beta_2)\lambda_1 - 2(\lambda_4 - \lambda_3 + \lambda_2 - \lambda_1)\bar{Q}_3^3(t) + 1 \\ &\leq -2\epsilon|\bar{Q}_3^3(t) - \bar{Q}_2^3(t) + \bar{Q}_1^3(t)| + 2(\beta_1 + \beta_2) + 1. \end{aligned}$$

Finally, under $\mathcal{E}_4(t)$, any arriving agent of types 1 or 3 increases $\bar{Q}_1^3(t)$ or $\bar{Q}_3^3(t)$ by 1, respectively. Any arriving agent of type 2 decreases $\bar{Q}_1^3(t)$ by 1, and any arriving agent of type 4 decreases $\bar{Q}_3^3(t)$ by 1. Thus, we have

$$\begin{aligned} \mathbb{E}[\mathcal{L}(t+1) - \mathcal{L}(t) \mid \bar{Q}^3(t), \mathcal{L}(t) > 0, \mathcal{E}_4(t)] \\ \leq -2(\lambda_2 - \lambda_1)(\beta_1 + \beta_2)\bar{Q}_1^3(t) + (\lambda_1 + \lambda_2)(\beta_1 + \beta_2) - 2(\lambda_4 - \lambda_3 + \lambda_2 - \lambda_1)(\bar{Q}_1^3(t) + \bar{Q}_3^3(t)) + 1 \\ \leq -2(\lambda_4 - \lambda_3 + \lambda_2 - \lambda_1)(\bar{Q}_1^3(t) + \bar{Q}_3^3(t)) + 1 + (\lambda_1 + \lambda_2)(\beta_1 + \beta_2) \\ \leq -2\epsilon|\bar{Q}_3^3(t) - \bar{Q}_2^3(t) + \bar{Q}_1^3(t)| + (\beta_1 + \beta_2) + 1, \end{aligned}$$

concluding the proof.

B.2 Towards optimal regret scaling

In this subsection, we assume that $(\mathcal{A}, \mathcal{M})$ is a path with 4 nodes, where 1 is the leaf node and 4 is the root. We aim to upper bound the expected length of queue 3 by $O(\epsilon^{-1})$ under the stationary distribution. It then follows by Proposition 7 and Lemma 2 that $\mathbb{E}[Q_3(t)] \leq O(\epsilon^{-1})$ for every $t \geq 0$.

Proposition 8. *Suppose that $(\mathcal{A}, \mathcal{M})$ is a path with 4 nodes, where 1 is the leaf node. Then, under **TP**, $\mathbb{E}_\pi[Q_3(t)] \leq O(\epsilon^{-1})$.*

Proof. Recall that queue 4 is truncated as it is under-demanded, and let \mathcal{S} denote the system obtained by further truncating queue 2. By Lemma 12, it suffices to show that $\mathbb{E}_\pi[Q_3(t)] \leq O(\epsilon^{-1})$ under \mathcal{S} . From now on, we use $(Q(t))_{t \geq 0}$ to denote the queue-length process under \mathcal{S} .

Next, we define another system \mathcal{S}' by modifying \mathcal{S} . In particular, we will “split” queue 3 into two sub-queues $3'$ and $3''$ with arrival probability $\lambda_{3'} \triangleq \lambda_2 - \lambda_1 - \epsilon/2$ and $\lambda_{3''} \triangleq \lambda_3 - (\lambda_2 - \lambda_1 - \epsilon/2)$, respectively, where agents in queue $3'$ can only match with agents in queue 2, and agents in queue $3''$ can only match with agents in queue 4. Notice that both $\lambda_{3'}$ and $\lambda_{3''}$ are strictly positive since, by Lemma 3,

$$\min\{\lambda_2 - \lambda_1, \lambda_3 - \lambda_2 + \lambda_1\} \geq \epsilon.$$

We implement the splitting by randomly assigning each agent in queue 3 to queue $3'$ and queue $3''$ with probability proportional to respectively $\lambda_{3'}$ and $\lambda_{3''}$. Under \mathcal{S}' , when an agent in queue 2 arrives, **TP** first tries to match it with an agent in queue 1, then tries to match it with an agent in queue $3'$ if queue 1 is empty, and finally discards this agent otherwise. Moreover, when an agent in queue 4 arrives, **TP** first tries to match it with an agent in queue $3''$, and then discards this agent if queue $3''$ is empty. Notice that every arriving agent to queues 1, 3, or $3''$ will immediately join the queue since its neighboring queues are all truncated and hence empty.

We couple the arrival processes of \mathcal{S} and \mathcal{S}' in the most natural way, where arrivals in queue $3'$ and queue $3''$ under \mathcal{S}' are coupled with arrivals in queue 3 under \mathcal{S} . Denote $(Q'(t))_{t \geq 0}$ as the queue-length process under \mathcal{S}' . We first show that, for every sample path, $Q_1(t) = Q'_1(t)$ for every $t \geq 0$.

Lemma 17. *Under the above coupling between the arrival processes of \mathcal{S} and \mathcal{S}' , for every sample path, $Q_1(t) = Q'_1(t)$ for every $t \geq 0$.*

Proof. The statement holds for $t = 0$ since $Q_1(0) = Q'_1(0) = 0$. Assume for induction that $Q_1(t) = Q'_1(t)$ for time $t \geq 0$, and we show that it also holds for time $t+1$. If the arrival under \mathcal{S} at time $t+1$ belongs to queue 3 or queue 4, then $Q_1(t+1) = Q_1(t) = Q'_1(t) = Q'_1(t+1)$. If the arrival under \mathcal{S} at time $t+1$ belongs to queue 1, then $Q_1(t+1) = Q_1(t) + 1$, and $Q'_1(t+1) = Q'_1(t) + 1$, implying $Q_1(t+1) = Q'_1(t+1)$. If the arrival under \mathcal{S} at time $t+1$ belongs to queue 2, since by the matching rule of **TP**, either both the arrivals under \mathcal{S} and \mathcal{S}' are matched to an agent in queue 1, or none of them does, implying $Q_1(t+1) = Q'_1(t+1)$. \square

Next, we show that under every sample path, $Q_3(t) \leq Q'_{3'}(t) + Q'_{3''}(t)$ for every $t \geq 0$.

Lemma 18. *Under the above coupling between the arrival processes of \mathcal{S} and \mathcal{S}' , for every sample path, $Q_3(t) \leq Q'_{3'}(t) + Q'_{3''}(t)$ for every $t \geq 0$.*

Proof. The statement holds for $t = 0$ since $Q_3(0) = Q'_{3'}(0) = Q'_{3''}(0) = 0$. Assume for induction that $Q_3(t) \leq Q'_{3'}(t) + Q'_{3''}(t)$ for time $t \geq 0$, and we show that it also holds for time $t+1$. Recall that we couple the arrival processes of \mathcal{S} and \mathcal{S}' in the most natural way, and we prove the statement for different possibilities of the arrival at time $t+1$. Suppose that the arrival under \mathcal{S} at time $t+1$ belongs to queue i .

- If $i = 1$, then $Q_3(t+1) = Q_3(t)$, $Q'_{3'}(t+1) = Q'_{3'}(t)$, and $Q'_{3''}(t+1) = Q'_{3''}(t)$, implying $Q_3(t+1) \leq Q'_{3'}(t+1) + Q'_{3''}(t+1)$.
- If $i = 3$, then $Q_3(t+1) = Q_3(t) + 1$ and $Q'_{3'}(t+1) + Q'_{3''}(t+1) = Q'_{3'}(t) + Q'_{3''}(t) + 1$, implying $Q_3(t+1) \leq Q'_{3'}(t+1) + Q'_{3''}(t+1)$.

- If $i = 2$, by Lemma 17, either both arrivals under \mathcal{S} and \mathcal{S}' are matched with an agent in queue 1, or none of them does. In the former case, we have $Q_3(t+1) = Q_3(t)$, $Q'_{3'}(t+1) = Q'_{3'}(t)$, and $Q'_{3''}(t+1) = Q'_{3''}(t)$, implying $Q_3(t+1) \leq Q'_{3'}(t+1) + Q'_{3''}(t+1)$.

Now, we focus on the latter case, where none of the arrivals under \mathcal{S} and \mathcal{S}' are matched with an agent in queue 1. If $Q_3(t) = 0$, then $Q_3(t+1) = Q_3(t) = 0 \leq Q'_{3'}(t+1) + Q'_{3''}(t+1)$. Otherwise, the arrival under \mathcal{S} is matched to an agent in queue 3, which gives $Q_3(t+1) = Q_3(t) - 1$. Since $Q'_{3'}(t+1) \geq Q'_{3'}(t) - 1$ and $Q'_{3''}(t+1) = Q'_{3''}(t)$, it follows that

$$Q_3(t+1) = Q_3(t) - 1 \leq Q'_{3'}(t) + Q'_{3''}(t) - 1 \leq Q'_{3'}(t+1) + Q'_{3''}(t+1),$$

as desired.

- Suppose that $i = 4$. If $Q_3(t) = 0$, then $Q_3(t+1) = Q_3(t) = 0 \leq Q'_{3'}(t+1) + Q'_{3''}(t+1)$ holds. Otherwise, the arrival under \mathcal{S} is matched to an agent in queue 3, which gives $Q_3(t+1) = Q_3(t) - 1$. Since $Q'_{3'}(t+1) = Q'_{3'}(t)$ and $Q'_{3''}(t+1) \geq Q'_{3''}(t) - 1$, it follows that

$$Q_3(t+1) = Q_3(t) - 1 \leq Q'_{3'}(t) + Q'_{3''}(t) - 1 \leq Q'_{3'}(t+1) + Q'_{3''}(t+1),$$

as desired.

Therefore, the statement holds for time $t + 1$, completing the proof. \square

Note that we can view the system consisting of queues 1, 2, and $3'$ as an $M/M/1$ queueing system with request arrival rate $\lambda_1 + \lambda_{3'} = \lambda_2 - \epsilon/2$ and service rate λ_2 , and view the system consisting of queues $3''$ and 4 as an $M/M/1$ queueing system with request arrival rate $\lambda_{3''} = \lambda_1 + \lambda_3 - \lambda_2 + \epsilon/2$ and service rate λ_4 . These two queueing systems are independent by the construction of \mathcal{S}' , and the service rate in every system is greater than its request arrival rate by $\epsilon/2$ since $\lambda_4 \geq \lambda_1 + \lambda_3 - \lambda_2 + \epsilon$ by Lemma 3. Hence, by Lemmas 17 and 18,

$$\mathbb{E}_\pi[Q_1(t) + Q_3(t)] \leq \mathbb{E}_\pi[Q'_1(t) + Q'_{3'}(t) + Q'_{3''}(t)] \leq \frac{4}{\epsilon},$$

where the second inequality holds by the standard queue-length bound for $M/M/1$ queueing systems, implying $\mathbb{E}_\pi[Q_3(t)] \leq O(\epsilon^{-1})$. \square

Discussion on generalizing Proposition 8. We show in Proposition 8 that the expected length of Q_3 is upper bounded by $O(\epsilon^{-1})$ in the stationary distribution via a careful splitting argument. It is then natural to ask whether one can generalize this approach to bound the expected queue lengths in more general networks. Unfortunately, as we argue now, this approach does not even apply to other queues in a longer path. Recall that our approach involves two key steps:

1. We first split certain queues to create multiple independent sub-systems, where each sub-system contains only one “server” queue or only one “request” queue so that the expected queue length can be bounded using standard results for $M/M/1$ queueing systems.
2. Then, we apply Lemmas 17 and 18 to argue that, under each sample path, the total queue length in these sub-systems upper bounds the total queue length in the original system.

Nevertheless, the second step no longer holds true in more general networks. Intuitively, this is because assigning agents into different sub-systems may accidentally result in better matching decisions, rendering a smaller total queue length.

We illustrate this barrier in the case where $(\mathcal{A}, \mathcal{M})$ is a path with 5 nodes. To bound $\mathbb{E}_\pi[Q_4(t)]$, one attempt would be to truncate the queues with an even depth, i.e., the queues with an odd index. Suppose that we split queue 2 into sub-queues $2'$ and $2''$, and we couple the original system \mathcal{S} and the new system \mathcal{S}' in the most natural way. We specify a sample path under which $Q_2(t) + Q_4(t) > Q'_{2'}(t) + Q'_{2''}(t) + Q'_4(t)$, where $(Q(t))_{t \geq 0}$ and $(Q'(t))_{t \geq 0}$ are the queue-length processes of \mathcal{S} and \mathcal{S}' , respectively. Consider the arrival sequence $(2, 4, 3, 1)$ under \mathcal{S} , and suppose the corresponding arrival sequence under \mathcal{S}' is $(2', 4, 3, 1)$. Only one match $(2, 3)$ is performed by **TP** under \mathcal{S} , whereas two matches $(3, 4)$ and $(1, 2')$ are performed by **TP** under \mathcal{S}' , implying

$$Q_2(4) + Q_4(4) = 1 > 0 = Q'_{2'}(4) + Q'_{2''}(4) + Q'_4(4).$$

Similar examples can be exploited to show that splitting other queues such as 3 or 4 does not always lead to a larger total queue length under all sample paths.

C Deferred proofs

C.1 Proof of Lemma 1

Recall that the constraints of $\text{SPP}(\lambda)$ can be written as

$$\begin{bmatrix} M & I \end{bmatrix} \begin{bmatrix} z \\ s \end{bmatrix} = \lambda,$$

where $I \in \mathbb{R}^{n \times n}$ is the identity matrix. Let $M_B \in \mathbb{R}^{n \times n}$ be the matrix obtained by selecting the columns of $\begin{bmatrix} M & I \end{bmatrix}$ corresponding to basic variables of $\text{SPP}(\lambda)$, which implies that M_B is full-rank. Hence, $(z_{\mathcal{M}_+}^*, s_{\mathcal{A}_+}^*)^T = M_B^{-1} \lambda$. For each vector $v \in \mathbb{R}^n$, we use $(M_B^{-1}v)_{\mathcal{M}_+}$ to denote the first $|\mathcal{M}_+|$ components of $M_B^{-1}v$. Fix $t > 0$. Note that

$$\mathcal{R}^*(t) \leq t \cdot r_{\mathcal{M}_+}^T z^* = t \cdot r_{\mathcal{M}_+}^T (M_B^{-1} \lambda)_{\mathcal{M}_+}$$

and

$$\mathcal{R}^\Pi(t) = \mathbb{E} [r^T D(t)] = \mathbb{E} \left[r_{\mathcal{M}_+}^T (M_B^{-1} (A(t) - Q(t)))_{\mathcal{M}_+} \right] = \mathbb{E} \left[r_{\mathcal{M}_+}^T (M_B^{-1} (t\lambda - Q(t)))_{\mathcal{M}_+} \right].$$

Therefore, the regret of Π after time t is

$$\mathcal{R}^*(t) - \mathcal{R}^\Pi(t) \leq \mathbb{E} \left[r_{\mathcal{M}_+}^T (M_B^{-1} Q(t))_{\mathcal{M}_+} \right] \leq r_{\max} \cdot \|M_B^{-1}\|_\infty \cdot \mathbb{E} [\|Q(t)\|_1] \leq r_{\max} B \cdot \|M_B^{-1}\|_\infty.$$

By [KAG25, Theorem 4.1], each entry of M_B^{-1} lies between $[-1, 1]$, implying that $\|M_B^{-1}\|_\infty \leq n$. This concludes the proof.

C.2 Deferred proofs for Section 4.2

C.2.1 Proof of Lemma 4

The upper bound holds straightforwardly for $i \in \mathcal{A}_0$ such that $d(r, i) \leq 2$ since $\mathcal{P}(i) = \emptyset$. For $i \in \mathcal{A}_0$ such that $d(r, i) > 2$, assume by induction that the upper bound holds for all $j \in \mathcal{P}(i)$, and we show that it also holds for i . By (8),

$$\begin{aligned} \alpha_i &= 1 + \frac{1}{\epsilon} \sum_{j \in \mathcal{P}(i)} \alpha_j (\lambda_j - \epsilon_j) \leq 1 + \frac{1}{\epsilon} \sum_{j \in \mathcal{P}(i)} \alpha_j \leq 1 + \frac{1}{\epsilon} \sum_{j \in \mathcal{P}(i)} \left(1 + \frac{1}{\epsilon}\right)^{\lfloor (d(r, j) - 1)/2 \rfloor} \\ &= 1 + \frac{1}{\epsilon} \sum_{t=0}^{\lfloor (d(r, i) - 3)/2 \rfloor} \left(1 + \frac{1}{\epsilon}\right)^t = \left(1 + \frac{1}{\epsilon}\right)^{\lfloor (d(r, i) - 1)/2 \rfloor}, \end{aligned}$$

where the first inequality holds since $\epsilon_i \geq \epsilon$ and $0 \leq \epsilon_j \leq \lambda_j \leq 1$ for every $j \in \mathcal{P}(i)$, and the second inequality holds by the inductive hypothesis. Hence, the upper bound also holds for i , concluding the proof.

C.2.2 Proof of Proposition 2

Fix $t \geq 0$, and we aim to upper bound the Lyapunov drift $\mathbb{E}[\mathcal{L}(t+1) - \mathcal{L}(t) \mid Q(t)]$, where the expectation is taken over the randomness of the arrival at time $t+1$. For each match $m \in \mathcal{M}$, let x_m denote the probability that m is performed by **TP** at time $t+1$. Let $\mathcal{E}_1 \triangleq \{i \in \mathcal{A}_0 \mid f_i(Q(t)) > 0\}$ be the set of over-demanded nodes i with a strictly positive $f_i(Q(t))$. The next lemma simplifies the Lyapunov drift.

Lemma 19. *It holds that*

$$\mathbb{E}[\mathcal{L}(t+1) - \mathcal{L}(t) \mid Q(t)] \leq 2 \sum_{i \in \mathcal{E}_1} \alpha_i \cdot f_i(Q(t)) \cdot f_i(\lambda - Mx) + n \left(1 + \frac{1}{\epsilon}\right)^{\lfloor (d_r-1)/2 \rfloor}. \quad (31)$$

Proof. Recall that $Q(t+1) = Q(t) + \Delta A(t+1) - M\Delta D(t+1)$. Since $f_i(\cdot)$ is linear, for every $i \in \mathcal{A}_0$,

$$f_i(Q(t+1)) - f_i(Q(t)) = f_i(Q(t+1) - Q(t)) = f_i(\Delta A(t+1) - M\Delta D(t+1)) \quad (32)$$

$$= \sum_{j \in \mathcal{T}^-(i)} (-1)^{d(i,j)+1} (\Delta A(t+1) - M\Delta D(t+1))_j. \quad (33)$$

Since **TP** always matches the arriving agent whenever it has a non-empty neighboring queue, there is precisely one non-zero entry in $\Delta A(t+1) - M\Delta D(t+1)$, which must be either 1 or -1 . Hence, (33) implies that, for every $i \in \mathcal{A}_0$,

$$|f_i(Q(t+1)) - f_i(Q(t))| \leq 1. \quad (34)$$

By (9),

$$\mathbb{E}[\mathcal{L}(t+1) - \mathcal{L}(t) \mid Q(t)] = \sum_{i \in \mathcal{A}_0} \alpha_i \cdot \mathbb{E}[(f_i(Q(t+1)))^+)^2 - (f_i(Q(t)))^+)^2 \mid Q(t)].$$

For each $i \in \mathcal{A}_0$ with $f_i(Q(t)) < 0$, (34) implies $f_i(Q(t+1)) \leq f_i(Q(t)) + 1 \leq 0$, and hence $f_i(Q(t))^{+} = f_i(Q(t+1))^{+} = 0$. Also, for each $i \in \mathcal{A}_0$ with $f_i(Q(t)) = 0$, (34) implies $f_i(Q(t+1)) \leq f_i(Q(t)) + 1 \leq 1$, and hence $(f_i(Q(t+1)))^+)^2 - (f_i(Q(t)))^+)^2 \leq 1$. Recall that $\mathcal{E}_1 \triangleq \{i \in \mathcal{A}_0 \mid f_i(Q(t)) > 0\}$ denotes the set of nodes i with a strictly positive $f_i(Q(t))$. As a result,

$$\mathbb{E}[\mathcal{L}(t+1) - \mathcal{L}(t) \mid Q(t)] \leq \sum_{i \in \mathcal{E}_1} \alpha_i \cdot \mathbb{E}[(f_i(Q(t+1)))^+)^2 - (f_i(Q(t)))^+)^2 \mid Q(t)] + \sum_{i \in \mathcal{A}_0 \setminus \mathcal{E}_1} \alpha_i. \quad (35)$$

Fix $i \in \mathcal{E}_1$. By (34), we have $f_i(Q(t+1)) \geq f_i(Q(t)) - 1 \geq 0$. It holds that

$$\begin{aligned} (f_i(Q(t+1)))^+)^2 - (f_i(Q(t)))^+)^2 &= f_i(Q(t+1))^2 - f_i(Q(t))^2 \\ &= (f_i(Q(t+1)) - f_i(Q(t)))^2 - 2f_i(Q(t))(f_i(Q(t)) - f_i(Q(t+1))) \\ &\leq 1 + 2f_i(Q(t))(f_i(Q(t+1)) - f_i(Q(t))) \\ &= 2f_i(Q(t)) \cdot f_i(\Delta A(t+1) - M\Delta D(t+1)) + 1, \end{aligned}$$

where the inequality holds by (34), and the last equality holds by (32). It follows that

$$\begin{aligned}\mathbb{E}[(f_i(Q(t+1))^+)^2 - (f_i(Q(t))^+)^2 \mid Q(t)] &\leq \mathbb{E}[2f_i(Q(t)) \cdot f_i(\Delta A(t+1) - M\Delta D(t+1)) + 1 \mid Q(t)] \\ &= 2f_i(Q(t)) \cdot f_i(\mathbb{E}[\Delta A(t+1) - M\Delta D(t+1) \mid Q(t)]) + 1 \\ &= 2f_i(Q(t)) \cdot f_i(\lambda - Mx) + 1,\end{aligned}$$

where the first equality holds since $f_i(\cdot)$ is linear. Combining the above displayed equation and (35), we obtain

$$\begin{aligned}\mathbb{E}[\mathcal{L}(t+1) - \mathcal{L}(t) \mid Q(t)] &\leq \sum_{i \in \mathcal{E}_1} \alpha_i (2f_i(Q(t)) \cdot f_i(\lambda - Mx) + 1) + \sum_{i \in \mathcal{A}_0 \setminus \mathcal{E}_1} \alpha_i \\ &= 2 \sum_{i \in \mathcal{E}_1} \alpha_i \cdot f_i(Q(t)) \cdot f_i(\lambda - Mx) + \sum_{i \in \mathcal{A}_0} \alpha_i \\ &\leq 2 \sum_{i \in \mathcal{E}_1} \alpha_i \cdot f_i(Q(t)) \cdot f_i(\lambda - Mx) + n \left(1 + \frac{1}{\epsilon}\right)^{\lfloor (d_r-1)/2 \rfloor},\end{aligned}$$

where the last inequality holds by Lemma 4 and the fact that $d(r, i) \leq d_r$ for every $i \in \mathcal{A}_0$. \square

It remains to upper bound the RHS of (31). Define $\mathcal{E}_2 \triangleq \{i \in \mathcal{A}_0 \mid \|Q_{\mathcal{C}(i)}(t)\|_1 > 0\}$ as the set of over-demanded nodes i such that at least one of its child nodes has a non-empty queue after time t . The following lemma establishes that $f_i(\lambda - Mx) = -\epsilon_i$ for every $i \in \mathcal{E}_2$, and $f_i(\lambda - Mx) \leq \lambda_i - \epsilon_i$ for every $i \in \mathcal{E}_1 \setminus \mathcal{E}_2$.

Lemma 20. $f_i(\lambda - Mx) = -\epsilon_i$ for every $i \in \mathcal{E}_2$, and $f_i(\lambda - Mx) \leq \lambda_i - \epsilon_i$ for every $i \in \mathcal{E}_1 \setminus \mathcal{E}_2$.

Proof. By (10),

$$\begin{aligned}f_i(\lambda - Mx) &= \sum_{j \in \mathcal{T}^-(i)} (-1)^{d(i,j)+1} (\lambda - Mx)_j \\ &= \lambda_i - \epsilon_i - \sum_{j \in \mathcal{T}^-(i)} (-1)^{d(i,j)+1} (Mx)_j,\end{aligned}\tag{36}$$

where the second equality holds by Lemma 3.

Hence, to upper bound $f_i(\lambda - Mx)$ given by (36), it suffices to lower bound

$$(\lambda_i - \epsilon_i) - f_i(\lambda - Mx) = \sum_{j \in \mathcal{T}^-(i)} (-1)^{d(i,j)+1} (Mx)_j = \mathbb{E} \left[\sum_{j \in \mathcal{T}^-(i)} (-1)^{d(i,j)+1} (M\Delta D(t+1))_j \mid Q(t) \right].$$

Define random variable $W_i \triangleq \sum_{j \in \mathcal{T}^-(i)} (-1)^{d(i,j)+1} (M\Delta D(t+1))_j$. Depending on the matches performed at time $t+1$, $W_i = 0$ happens in the following three cases:

1. No match is performed, implying $\Delta D(t+1) = \mathbf{0}$.
2. The performed match $m(\ell_1, \ell_2)$ satisfies $\ell_1, \ell_2 \in \mathcal{A} \setminus \mathcal{T}^-(i)$, implying $(M\Delta D(t+1))_j = 0$ for every $j \in \mathcal{T}^-(i)$.
3. The performed match $m(\ell_1, \ell_2)$ satisfies $\ell_1, \ell_2 \in \mathcal{T}^-(i)$. In this case, $(M\Delta D(t+1))_j = 0$ for every $j \in \mathcal{T}^-(i) \setminus \{\ell_1, \ell_2\}$, and

$$(-1)^{d(i,\ell_1)+1} (M\Delta D(t+1))_{\ell_1} + (-1)^{d(i,\ell_2)+1} (M\Delta D(t+1))_{\ell_2} = (-1)^{d(i,\ell_1)+1} + (-1)^{d(i,\ell_2)+1} = 0,$$

where the last equality holds since ℓ_1 and ℓ_2 are adjacent.

As the only remaining case, suppose that the performed match $m(\ell_1, \ell_2)$ satisfies $\ell_1 = i$ and $\ell_2 \in \mathcal{C}(i)$. In this case, $(M\Delta D(t+1))_j = 0$ for every $j \in \mathcal{T}^-(i) \setminus \{\ell_2\}$ and $(M\Delta D(t+1))_{\ell_2} = 1$, implying that $W_i = 1$. In all cases, we always have $W_i \geq 0$, and hence $f_i(\lambda - Mx) = \lambda_i - \epsilon_i - \mathbb{E}[W_i \mid Q(t)] \leq \lambda_i - \epsilon_i$. This concludes the second part of Lemma 20. Furthermore, if $i \in \mathcal{E}_2$, by the matching rule of **SP**, $W_i = 1$ holds if and only if an agent of type i arrives at time $t+1$, which happens with probability λ_i . Therefore, $\mathbb{E}[W_i \mid Q(t)] = \lambda_i$ for $i \in \mathcal{E}_2$, concluding the first part of Lemma 20. \square

We write

$$\sum_{i \in \mathcal{E}_1} \alpha_i \cdot f_i(Q(t)) \cdot f_i(\lambda - Mx) = \underbrace{\sum_{i \in \mathcal{E}_1 \cap \mathcal{E}_2} \alpha_i \cdot f_i(Q(t)) \cdot f_i(\lambda - Mx)}_{U_1} + \underbrace{\sum_{i \in \mathcal{E}_1 \setminus \mathcal{E}_2} \alpha_i \cdot f_i(Q(t)) \cdot f_i(\lambda - Mx)}_{U_2}.$$

Lemma 20 ensures that $U_1 < 0$ since $f_i(Q(t)) > 0$ for every $i \in \mathcal{E}_1$, and provides an upper bound for $f_i(\lambda - Mx)$ for $i \in \mathcal{E}_1 \setminus \mathcal{E}_2$. To show that U_2 is small, it remains to upper bound $f_i(Q(t))$ for $i \in \mathcal{E}_1 \setminus \mathcal{E}_2$. For every $i \in \mathcal{A}_0 \setminus \mathcal{E}_2$, let $\mathcal{H}(i)$ be the set of nodes $j \in \mathcal{T}^-(i) \cap \mathcal{E}_2$ that satisfy the following conditions:

1. $d(i, j) \equiv 0 \pmod{2}$, and
2. for every $k \in \mathcal{P}(j) \cap \mathcal{T}^-(i)$ (nodes k on the path between i and j (excluding i and j) with $d(i, k) \equiv 0 \pmod{2}$), $k \notin \mathcal{E}_2$.

Equivalently, we can constructively define $\mathcal{H}(i)$ as follows: Imagine walking from i down to the leaf nodes in the subtree rooted at i with step size 2. If we encounter a node in \mathcal{E}_2 , then we add the current node into $\mathcal{H}(i)$ and stop; otherwise, we continue walking. One can also see by this constructive definition that $\mathcal{T}^-(j) \cap \mathcal{T}^-(k) = \emptyset$ for all $j, k \in \mathcal{H}(i)$ with $j \neq k$, i.e., there is no overlapping between any two rooted subtrees with different roots $j, k \in \mathcal{H}(i)$.

The following lemma states that for every $i \in \mathcal{A}_0 \setminus \mathcal{E}_2$, we can upper bound $f_i(Q(t))$ by the sum of $f_j(Q(t))$ over all $j \in \mathcal{H}(i)$.

Lemma 21. *For every $i \in \mathcal{A}_0 \setminus \mathcal{E}_2$,*

$$f_i(Q(t)) \leq \sum_{j \in \mathcal{H}(i)} f_j(Q(t)).$$

Proof. Recall that $\mathcal{E}_2 \triangleq \{i \in \mathcal{A}_0 \mid \|Q_{\mathcal{C}(i)}(t)\|_1 > 0\}$. Fix $i \in \mathcal{A}_0 \setminus \mathcal{E}_2$. Recall that $\mathcal{T}^-(j) \cap \mathcal{T}^-(k) = \emptyset$ for all $j, k \in \mathcal{H}(i)$ with $j \neq k$. Define $\mathcal{U} \triangleq \bigcup_{j \in \mathcal{H}(i)} \mathcal{T}^-(j)$. Since $d(i, j) \equiv 0 \pmod{2}$ for every $j \in \mathcal{H}(i)$, we have

$$\sum_{k \in \mathcal{U}} (-1)^{d(i, k)+1} Q_k(t) = \sum_{j \in \mathcal{H}(i)} \sum_{k \in \mathcal{T}^-(j)} (-1)^{d(j, k)+1} Q_k(t) = \sum_{j \in \mathcal{H}(i)} f_j(Q(t)).$$

Moreover, by the definition of $\mathcal{H}(i)$, $j \notin \mathcal{E}_2$ for every $j \in \mathcal{T}^-(i) \setminus (\mathcal{U} \cup \mathcal{H}(i))$ such that $d(i, j) \equiv 0 \pmod{2}$, implying $Q_j(t) = 0$ for every $j \in \mathcal{T}^-(i) \setminus \mathcal{U}$ such that $d(i, j) \equiv 1 \pmod{2}$. Hence,

$$\sum_{j \in \mathcal{T}^-(i) \setminus \mathcal{U}} (-1)^{d(i, j)+1} Q_j(t) = \sum_{j \in \mathcal{T}^-(i) \setminus \mathcal{U}} \mathbf{1}_{\{d(i, j) \equiv 0 \pmod{2}\}} \cdot (-1)^{d(i, j)+1} Q_j(t) \leq 0.$$

Combining the above two displayed equations,

$$\begin{aligned} f_i(Q(t)) &= \sum_{j \in \mathcal{T}^-(i)} (-1)^{d(i,j)+1} Q_j(t) \\ &= \sum_{j \in \mathcal{U}} (-1)^{d(i,j)+1} Q_j(t) + \sum_{j \in \mathcal{T}^-(i) \setminus \mathcal{U}} (-1)^{d(i,j)+1} Q_j(t) \leq \sum_{j \in \mathcal{H}(i)} f_j(Q(t)), \end{aligned}$$

concluding the proof. \square

Next, we apply Lemmas 20 and 21 to formally upper bound the first term in the RHS of (31), showing that $\mathcal{L}(t)$ has a one-step negative drift.

Lemma 22. *It holds that*

$$\sum_{i \in \mathcal{E}_1} \alpha_i \cdot f_i(Q(t)) \cdot f_i(\lambda - Mx) \leq - \sum_{i \in \mathcal{E}_1 \cap \mathcal{E}_2} f_i(Q(t)) \cdot \epsilon_i.$$

Proof. We have

$$\begin{aligned} &\sum_{i \in \mathcal{E}_1} \alpha_i \cdot f_i(Q(t)) \cdot f_i(\lambda - Mx) \\ &= \sum_{i \in \mathcal{E}_1 \cap \mathcal{E}_2} \alpha_i \cdot f_i(Q(t)) \cdot f_i(\lambda - Mx) + \sum_{i \in \mathcal{E}_1 \setminus \mathcal{E}_2} \alpha_i \cdot f_i(Q(t)) \cdot f_i(\lambda - Mx) \\ &\leq \sum_{i \in \mathcal{E}_1 \cap \mathcal{E}_2} \alpha_i \cdot f_i(Q(t)) \cdot (-\epsilon_i) + \sum_{i \in \mathcal{E}_1 \setminus \mathcal{E}_2} \alpha_i \cdot f_i(Q(t)) \cdot (\lambda_i - \epsilon_i) \\ &\leq \sum_{i \in \mathcal{E}_1 \cap \mathcal{E}_2} \alpha_i \cdot f_i(Q(t)) \cdot (-\epsilon_i) + \sum_{i \in \mathcal{E}_1 \setminus \mathcal{E}_2} \alpha_i (\lambda_i - \epsilon_i) \sum_{j \in \mathcal{H}(i)} f_j(Q(t)) \\ &= \sum_{i \in \mathcal{E}_2} f_i(Q(t)) \left(-\mathbf{1}_{\{i \in \mathcal{E}_1\}} \cdot \alpha_i \epsilon_i + \sum_{j \in \mathcal{E}_1 \setminus \mathcal{E}_2} \mathbf{1}_{\{i \in \mathcal{H}(j)\}} \cdot \alpha_j (\lambda_j - \epsilon_j) \right), \end{aligned} \tag{37}$$

where the first inequality holds by Lemma 20 and the fact that $f_i(Q(t)) > 0$ for every $i \in \mathcal{E}_1$, the second inequality holds by Lemma 21, and the last equality holds since $\mathcal{H}(i) \subseteq \mathcal{E}_2$ for every $i \in \mathcal{E}_1 \setminus \mathcal{E}_2$. To further upper bound (37), we consider each $i \in \mathcal{E}_2$ individually. On one hand, for every $i \in \mathcal{E}_2 \setminus \mathcal{E}_1$, since $f_i(Q(t)) \leq 0$ by the definition of \mathcal{E}_1 , we have

$$\begin{aligned} &f_i(Q(t)) \left(-\mathbf{1}_{\{i \in \mathcal{E}_1\}} \cdot \alpha_i \epsilon_i + \sum_{j \in \mathcal{E}_1 \setminus \mathcal{E}_2} \mathbf{1}_{\{i \in \mathcal{H}(j)\}} \cdot \alpha_j (\lambda_j - \epsilon_j) \right) \\ &= f_i(Q(t)) \left(\sum_{j \in \mathcal{E}_1 \setminus \mathcal{E}_2} \mathbf{1}_{\{i \in \mathcal{H}(j)\}} \cdot \alpha_j (\lambda_j - \epsilon_j) \right) \leq 0. \end{aligned}$$

On the other hand, for every $i \in \mathcal{E}_1 \cap \mathcal{E}_2$, which satisfies $f_i(Q(t)) > 0$, it holds that

$$\begin{aligned} -\mathbf{1}_{\{i \in \mathcal{E}_1\}} \cdot \alpha_i \epsilon_i + \sum_{j \in \mathcal{E}_1 \setminus \mathcal{E}_2} \mathbf{1}_{\{i \in \mathcal{H}(j)\}} \cdot \alpha_j (\lambda_j - \epsilon_j) &\leq -\alpha_i \epsilon_i + \sum_{j \in \mathcal{P}(i) \cap (\mathcal{E}_1 \setminus \mathcal{E}_2)} \alpha_j (\lambda_j - \epsilon_j) \\ &\leq -\alpha_i \epsilon_i + \sum_{j \in \mathcal{P}(i)} \alpha_j (\lambda_j - \epsilon_j) = -\epsilon_i, \end{aligned}$$

where the first inequality holds since $i \in \mathcal{H}(j)$ implies $j \in \mathcal{P}(i)$, and the last equality holds by (8). Combining the above three displayed equations concludes the proof. \square

By Lemmas 19 and 22, and the fact that $\epsilon \leq \epsilon_i \leq 1$ for every $i \in \mathcal{A}$, we get

$$\mathbb{E}[\mathcal{L}(t+1) - \mathcal{L}(t) \mid Q(t)] \leq -2\epsilon \sum_{i \in \mathcal{E}_1 \cap \mathcal{E}_2} f_i(Q(t)) + n \left(1 + \frac{1}{\epsilon}\right)^{\lfloor (d_r-1)/2 \rfloor}. \quad (38)$$

Then, we apply the following lemma to translate the above Lyapunov drift in terms of f to a drift in terms of the total queue length.

Lemma 23. *For every $q \in \mathbb{Z}_{\geq 0}^n$, let $\mathcal{E}_1 \triangleq \{i \in \mathcal{A}_0 \mid f_i(q) > 0\}$ and $\mathcal{E}_2 \triangleq \{i \in \mathcal{A}_0 \mid \|q_{\mathcal{C}(i)}\|_1 > 0\}$. Then,*

$$\frac{1}{2^{d_r}} \sum_{i \in \mathcal{A}_0} q_i \leq \sum_{i \in \mathcal{E}_1 \cap \mathcal{E}_2} f_i(q).$$

Proof. Fix $q \in \mathbb{Z}_{\geq 0}^n$. We prove a stronger statement that

$$\frac{1}{2^{d_i}} \sum_{j \in \mathcal{T}^-(i)} q_j \leq \sum_{j \in \mathcal{T}(i) \cap \mathcal{E}_1 \cap \mathcal{E}_2} f_j(q) \quad (39)$$

for every $i \in \mathcal{A}$, and Lemma 23 follows by setting $i = r$.

For those $i \in \mathcal{A}$ with $d_i = 0$, i.e., i is a leaf node, (39) holds straightforwardly since both sides equal 0. Assume by induction that (39) holds for all $i \in \mathcal{A}$ with $d_i < k$ such that $k \in [d_r]$, and we show that (39) holds for all $i \in \mathcal{A}$ with $d_i = k$. Fix $i \in \mathcal{A}$ with $d_i = k$. Observe that

$$\begin{aligned} \sum_{j \in \mathcal{T}^-(i) \cap \mathcal{E}_1 \cap \mathcal{E}_2} f_j(q) &= \sum_{j \in \mathcal{C}(i)} \sum_{k \in \mathcal{T}(j) \cap \mathcal{E}_1 \cap \mathcal{E}_2} f_k(q) \\ &\geq \sum_{j \in \mathcal{C}(i)} \frac{1}{2^{d_j}} \sum_{k \in \mathcal{T}^-(j)} q_k \geq \frac{1}{2^{d_i-1}} \sum_{j \in \mathcal{C}(i)} \sum_{k \in \mathcal{T}^-(j)} q_k, \end{aligned}$$

where the first inequality holds by the inductive hypothesis. Hence,

$$\sum_{j \in \mathcal{T}(i) \cap \mathcal{E}_1 \cap \mathcal{E}_2} f_j(q) \geq \mathbf{1}_{\{i \in \mathcal{E}_1 \cap \mathcal{E}_2\}} \cdot f_i(q) + \frac{1}{2^{d_i-1}} \sum_{j \in \mathcal{C}(i)} \sum_{k \in \mathcal{T}^-(j)} q_k,$$

and it suffices to show that

$$\mathbf{1}_{\{i \in \mathcal{E}_1 \cap \mathcal{E}_2\}} \cdot f_i(q) + \frac{1}{2^{d_i-1}} \sum_{j \in \mathcal{C}(i)} \sum_{k \in \mathcal{T}^-(j)} q_k \geq \frac{1}{2^{d_i}} \sum_{j \in \mathcal{T}^-(i)} q_j,$$

which is equivalent to

$$\mathbf{1}_{\{i \in \mathcal{E}_1 \cap \mathcal{E}_2\}} \cdot f_i(q) \geq \frac{1}{2^{d_i}} \sum_{j \in \mathcal{C}(i)} \left(q_j - \sum_{k \in \mathcal{T}^-(j)} q_k \right). \quad (40)$$

We show that (40) holds under three different cases. Firstly, if $i \notin \mathcal{E}_1$, i.e., $f_i(q) \leq 0$, then $\mathbf{1}_{\{i \in \mathcal{E}_1 \cap \mathcal{E}_2\}} \cdot f_i(q) = 0$, and

$$\sum_{j \in \mathcal{C}(i)} \left(q_j - \sum_{k \in \mathcal{T}^-(j)} q_k \right) \leq \sum_{j \in \mathcal{T}^-(i)} (-1)^{d(i,j)+1} q_j = f_i(q) \leq 0,$$

implying (40). Next, if $i \notin \mathcal{E}_2$, i.e., $\|q_{\mathcal{C}(i)}\|_1 = 0$, then $\mathbf{1}_{\{i \in \mathcal{E}_1 \cap \mathcal{E}_2\}} \cdot f_i(q) = 0$, and

$$\sum_{j \in \mathcal{C}(i)} \left(q_j - \sum_{k \in \mathcal{T}^-(j)} q_k \right) = - \sum_{j \in \mathcal{C}(i)} \sum_{k \in \mathcal{T}^-(j)} q_k \leq 0,$$

implying (40). Finally, if $i \in \mathcal{E}_1 \cap \mathcal{E}_2$, then $f_i(q) > 0$ and $\mathbf{1}_{\{i \in \mathcal{E}_1 \cap \mathcal{E}_2\}} = 1$; by (10),

$$\begin{aligned} f_i(q) - \frac{1}{2^{d_i}} \sum_{j \in \mathcal{C}(i)} \left(q_j - \sum_{k \in \mathcal{T}^-(j)} q_k \right) &= \sum_{j \in \mathcal{T}^-(i)} (-1)^{d(i,j)+1} q_j - \frac{1}{2^{d_i}} \sum_{j \in \mathcal{C}(i)} \left(q_j - \sum_{k \in \mathcal{T}^-(j)} q_k \right) \\ &= \sum_{j \in \mathcal{C}(i)} \left(\left(1 - \frac{1}{2^{d_i}} \right) q_j + \sum_{k \in \mathcal{T}^-(j)} \left((-1)^{d(i,k)+1} + \frac{1}{2^{d_i}} \right) q_k \right) \\ &\geq \sum_{j \in \mathcal{T}^-(i)} \left(1 - \frac{1}{2^{d_i}} \right) (-1)^{d(i,j)+1} q_j \\ &= \left(1 - \frac{1}{2^{d_i}} \right) f_i(q) > 0, \end{aligned}$$

where the inequality holds since $q_j \geq 0$ for every $j \in \mathcal{A}$. This concludes the proof. \square

Combining (38) and Lemma 23, we obtain

$$\mathbb{E}[\mathcal{L}(t+1) - \mathcal{L}(t) \mid Q(t)] \leq -\frac{\epsilon}{2^{d_r-1}} \|Q(t)\|_1 + n \left(1 + \frac{1}{\epsilon} \right)^{\lfloor (d_r-1)/2 \rfloor},$$

as desired. By Lemma 11, the above drift bound also implies that the Markov chain $(Q(t))_{t \geq 0}$ is ergodic, concluding the proof of Proposition 2.

C.3 Deferred proofs for Section 4.3

C.3.1 Proof of Proposition 3

By the second-order Taylor's expansion at 0, for some $\theta' \in (0, \theta]$ and any $x_1, x_2 \in \mathcal{X}$, we have

$$\begin{aligned} \exp(\theta(\Phi(x_2) - \Phi(x_1))) &\leq \exp(\theta(\max\{-\delta, \Phi(x_2) - \Phi(x_1)\})) \\ &= 1 + \theta(\max\{-\delta, \Phi(x_2) - \Phi(x_1)\}) \\ &\quad + \frac{\theta^2}{2} (\max\{-\delta, \Phi(x_2) - \Phi(x_1)\})^2 \exp(\theta'(\max\{-\delta, \Phi(x_2) - \Phi(x_1)\})) \\ &\leq 1 + \theta(\max\{-\delta, \Phi(x_2) - \Phi(x_1)\}) \\ &\quad + \frac{\theta^2}{2} (\max\{-\delta, \Phi(x_2) - \Phi(x_1)\})^2 \exp(\theta(\Phi(x_2) - \Phi(x_1))^+) . \end{aligned}$$

We now fix an arbitrary $x \in \mathcal{X}$ such that $\Phi(x) > K$, set $x_1 = x$, $x_2 = X(t_0)$ and take the expectation of both sides above to obtain

$$\begin{aligned} \mathbb{E}_x[\exp(\theta(\Phi(X(t_0)) - \Phi(x)))] &\leq 1 + \theta \mathbb{E}_x[(\max\{-\delta, \Phi(X(t_0)) - \Phi(x)\})] \\ &\quad + \frac{\theta^2}{2} \mathbb{E}_x[(\max\{-\delta, \Phi(X(t_0)) - \Phi(x)\})^2 \exp(\theta(\Phi(X(t_0)) - \Phi(x))^+)] \\ &\stackrel{(a)}{\leq} 1 - \gamma\theta + \frac{\theta^2}{2} L_3(\delta, \theta, t_0) \\ &\stackrel{(b)}{\leq} 1 - \gamma\theta/2, \end{aligned}$$

where (a) follows from (11) and the definition of L_3 ; (b) holds by (12). Note that the condition $\Phi(x) > K$ is equivalent to $\exp(\theta \cdot \Phi(x)) > \exp(\theta K)$. Thus, $\exp(\theta \cdot \Phi(x))$ is a geometric Lyapunov function with geometric drift size parameter $1 - \gamma\theta/2$, drift time parameter t_0 , and exception parameter $\exp(\theta K)$.

C.3.2 Proof of Proposition 4

For all $i \in \mathcal{A}_0$ and $t \geq 0$, define

$$R_i(t) \triangleq q_i(0) + A_i(t) - \sum_{j \in \mathcal{C}(i)} D_{m(j,i)}(t)$$

as the number of agents arriving at i before and including time t that are not matched with agents in the children of i . By the matching rule of **TTP**, $R_i(t)$ also equals the number of agents arriving at i before and including time t that are not matched immediately upon arrival, i.e.,

$$R_i(t) = q_i(t) + D_{m(i,P(i))}(t) \quad (41)$$

for all $i \in \mathcal{A}_0$ and $t \geq 0$.

Analogous to the Lipschitz-continuity results for several classical queueing-network models [CY01], we establish Proposition 4 via Skorokhod reflection mappings. However, off-the-shelf reflection-map representations are generally not available for dynamic matching systems, and we derive an explicit characterization tailored to dynamic matching in the following lemma.

Lemma 24. *Under **TTP**, for all $i \in \mathcal{A}$ and $t \geq 0$,*

$$\sum_{j \in \mathcal{C}(i)} q_j(t) = \sum_{j \in \mathcal{C}(i)} R_j(t) - A_i(t) + \max_{0 \leq s \leq t} \left[A_i(s) - \sum_{j \in \mathcal{C}(i)} R_j(s) \right]^+. \quad (42)$$

Proof. We prove (42) by induction. Firstly, (42) holds straightforwardly for the base case of $t = 0$ since all terms in (42) equal zero. We assume that $t \geq 1$, and (42) holds for $t - 1$ and all $i \in \mathcal{A}$. We now show that (42) holds for t and all $i \in \mathcal{A}$. Fix $i \in \mathcal{A}$. It holds that

$$\begin{aligned} \sum_{j \in \mathcal{C}(i)} q_j(t) &= \sum_{j \in \mathcal{C}(i)} \left(q_j(t-1) + \Delta A_j(t) - \Delta D_{m(j,i)}(t) - \sum_{k \in \mathcal{C}(j)} \Delta D_{m(k,j)}(t) \right) \\ &= \sum_{j \in \mathcal{C}(i)} (R_j(t) - \Delta D_{m(j,i)}(t)) - A_i(t-1) + \max_{0 \leq s \leq t-1} \left[A_i(s) - \sum_{j \in \mathcal{C}(i)} R_j(s) \right]^+, \end{aligned} \quad (43)$$

where the second inequality holds by the inductive hypothesis and the identity $R_j(t) = R_j(t-1) + \Delta A_j(t) - \sum_{k \in \mathcal{C}(j)} \Delta D_{m(k,j)}(t)$. Since **TTP** only matches the new arrivals in queue- i with agents in the children of i , we have $\sum_{j \in \mathcal{C}(i)} \Delta D_{m(j,i)}(t) \leq \Delta A_i(t)$.

On one hand, if $\sum_{j \in \mathcal{C}(i)} \Delta D_{m(j,i)}(t) = \Delta A_i(t)$, then by (43),

$$\sum_{j \in \mathcal{C}(i)} q_j(t) = \sum_{j \in \mathcal{C}(i)} R_j(t) - A_i(t) + \max_{0 \leq s \leq t-1} \left[A_i(s) - \sum_{j \in \mathcal{C}(i)} R_j(s) \right]^+.$$

Since $\sum_{j \in \mathcal{C}(i)} q_j(t) \geq 0$, by rearranging the terms in the RHS, it follows that

$$\max_{0 \leq s \leq t-1} \left[A_i(s) - \sum_{j \in \mathcal{C}(i)} R_j(s) \right]^+ \geq \max \left\{ 0, A_i(t) - \sum_{j \in \mathcal{C}(i)} R_j(t) \right\} = \left[A_i(t) - \sum_{j \in \mathcal{C}(i)} R_j(t) \right]^+.$$

Combining the above two displayed equations yields

$$\sum_{j \in \mathcal{C}(i)} q_j(t) = \sum_{j \in \mathcal{C}(i)} R_j(t) - A_i(t) + \max_{0 \leq s \leq t} \left[A_i(s) - \sum_{j \in \mathcal{C}(i)} R_j(s) \right]^+,$$

concluding the inductive step.

On the other hand, if $\sum_{j \in \mathcal{C}(i)} \Delta D_{m(j,i)}(t) < \Delta A_i(t)$, which implies there are redundant new agents in queue- i after matching with all existing agents in the children of i , then we must have $\sum_{j \in \mathcal{C}(i)} q_j(t) = 0$. By rewriting (43), we get

$$\sum_{j \in \mathcal{C}(i)} q_j(t) = \sum_{j \in \mathcal{C}(i)} R_j(t) - A_i(t) + \max_{0 \leq s \leq t-1} \left[A_i(s) - \sum_{j \in \mathcal{C}(i)} R_j(s) \right]^+ + \Delta A_i(t) - \sum_{j \in \mathcal{C}(i)} \Delta D_{m(j,i)}(t).$$

Since $\sum_{j \in \mathcal{C}(i)} q_j(t) = 0$ and $\Delta A_i(t) - \sum_{j \in \mathcal{C}(i)} \Delta D_{m(j,i)}(t) > 0$, by rearranging the RHS of the above equation, it follows that

$$\max_{0 \leq s \leq t-1} \left[A_i(s) - \sum_{j \in \mathcal{C}(i)} R_j(s) \right]^+ < A_i(t) - \sum_{j \in \mathcal{C}(i)} R_j(t) \leq \left[A_i(t) - \sum_{j \in \mathcal{C}(i)} R_j(t) \right]^+, \quad (44)$$

which implies $A_i(t) - \sum_{j \in \mathcal{C}(i)} R_j(t) > 0$. Therefore,

$$\begin{aligned} \sum_{j \in \mathcal{C}(i)} q_j(t) &= 0 = \sum_{j \in \mathcal{C}(i)} R_j(t) - A_i(t) + \left[A_i(t) - \sum_{j \in \mathcal{C}(i)} R_j(t) \right]^+ \\ &= \sum_{j \in \mathcal{C}(i)} R_j(t) - A_i(t) + \max_{0 \leq s \leq t} \left[A_i(s) - \sum_{j \in \mathcal{C}(i)} R_j(s) \right]^+, \end{aligned}$$

where the second equality follows from $A_i(t) - \sum_{j \in \mathcal{C}(i)} R_j(t) > 0$, and the last equality holds by (44). This concludes the inductive step. \square

For all $i \in \mathcal{A}$ and $t \geq 0$, analogously define $R'_i(t)$, and $D'_{m(j,i)}(t)$ for $j \in \mathcal{C}(i)$ under the arrival trajectory $(A'(t))_{t \geq 0}$. For all $i \in \mathcal{A}$ and $t \geq 0$, we have

$$\begin{aligned} \left| \sum_{j \in \mathcal{C}(i)} q_j(t) - \sum_{j \in \mathcal{C}(i)} q'_j(t) \right| &\leq \left| \sum_{j \in \mathcal{C}(i)} R_j(t) - \sum_{j \in \mathcal{C}(i)} R'_j(t) \right| + |A_i(t) - A'_i(t)| \\ &\quad + \left| \max_{0 \leq s \leq t} \left[A_i(s) - \sum_{j \in \mathcal{C}(i)} R_j(s) \right]^+ - \max_{0 \leq s \leq t} \left[A'_i(s) - \sum_{j \in \mathcal{C}(i)} R'_j(s) \right]^+ \right| \\ &\leq 2 \left(\max_{0 \leq s \leq t} \left| \sum_{j \in \mathcal{C}(i)} R_j(s) - \sum_{j \in \mathcal{C}(i)} R'_j(s) \right| + \max_{0 \leq s \leq t} |A_i(s) - A'_i(s)| \right), \quad (45) \end{aligned}$$

where the first inequality holds by Proposition 24, and the second inequality holds since $|\max_i x_i - \max_i y_i| \leq \max_i |x_i - y_i|$ and $|x^+ - y^+| \leq |x - y|$.

The following lemma bounds the first max term in (45).

Lemma 25. *for all $i \in \mathcal{A}$ and $t \geq 0$,*

$$\max_{0 \leq s \leq t} \left| \sum_{j \in \mathcal{C}(i)} R_j(s) - \sum_{j \in \mathcal{C}(i)} R'_j(s) \right| \leq \sum_{j \in \mathcal{T}^-(i)} \max_{0 \leq s \leq t} |A_j(s) - A'_j(s)|. \quad (46)$$

To conclude (15), by (45),

$$\begin{aligned} \left| \sum_{i \in \mathcal{A}_0} q_i(t) - \sum_{i \in \mathcal{A}_0} q'_i(t) \right| &\leq \sum_{i \in \mathcal{A}} \left| \sum_{j \in \mathcal{C}(i)} q_j(t) - \sum_{j \in \mathcal{C}(i)} q'_j(t) \right| \\ &\leq 2 \sum_{i \in \mathcal{A}} \left(\max_{0 \leq s \leq t} \left| \sum_{j \in \mathcal{C}(i)} R_j(s) - \sum_{j \in \mathcal{C}(i)} R'_j(s) \right| + \max_{0 \leq s \leq t} |A_i(s) - A'_i(s)| \right) \\ &\leq 2 \sum_{i \in \mathcal{A}} \left(\sum_{j \in \mathcal{T}^-(i)} \max_{0 \leq s \leq t} |A_j(s) - A'_j(s)| + \max_{0 \leq s \leq t} |A_i(s) - A'_i(s)| \right) \\ &\leq 2(d_r + 1) \sum_{i \in \mathcal{A}} \max_{0 \leq s \leq t} |A_i(s) - A'_i(s)|, \end{aligned}$$

where the third inequality holds by Lemma 25.

It remains to prove Lemma 25.

Proof of Lemma 25. For all $i \in \mathcal{A}$ and $t \geq 0$,

$$\sum_{j \in \mathcal{C}(i)} D_{m(j,i)}(t) = \sum_{j \in \mathcal{C}(i)} (R_j(t) - q_j(t)) = A_i(t) - \max_{0 \leq s \leq t} \left[A_i(s) - \sum_{j \in \mathcal{C}(i)} R_j(s) \right]^+, \quad (47)$$

where the first equality holds by (41), and the second equality holds by Proposition 24. As a result,

$$\begin{aligned} \sum_{j \in \mathcal{C}(i)} R_j(t) &= \sum_{j \in \mathcal{C}(i)} \left(q_j(0) + A_j(t) - \sum_{k \in \mathcal{C}(j)} D_{m(k,j)}(t) \right) \\ &= \sum_{j \in \mathcal{C}(i)} \left(q_j(0) + \max_{0 \leq s \leq t} \left[A_j(s) - \sum_{k \in \mathcal{C}(j)} R_k(s) \right]^+ \right), \end{aligned} \quad (48)$$

where the first equality holds by the definition of $R_j(t)$, and the second inequality holds by (47); the same formula also holds for $\sum_{j \in \mathcal{C}(i)} R'_j(t)$. Applying (48) and the assumption that $q(0) = q'(0)$,

we get

$$\begin{aligned}
& \max_{0 \leq s \leq t} \left| \sum_{j \in \mathcal{C}(i)} R_j(s) - \sum_{j \in \mathcal{C}(i)} R'_j(s) \right| \\
&= \max_{0 \leq s \leq t} \left| \sum_{j \in \mathcal{C}(i)} \max_{0 \leq \ell \leq s} \left[A_j(\ell) - \sum_{k \in \mathcal{C}(j)} R_k(\ell) \right]^+ - \sum_{j \in \mathcal{C}(i)} \max_{0 \leq \ell \leq s} \left[A'_j(\ell) - \sum_{k \in \mathcal{C}(j)} R'_k(\ell) \right]^+ \right| \\
&\leq \max_{0 \leq s \leq t} \sum_{j \in \mathcal{C}(i)} \left(\max_{0 \leq \ell \leq s} \left| \sum_{k \in \mathcal{C}(j)} R_k(\ell) - \sum_{k \in \mathcal{C}(j)} R'_k(\ell) \right| + \max_{0 \leq \ell \leq s} |A_j(\ell) - A'_j(\ell)| \right) \\
&\leq \sum_{j \in \mathcal{C}(i)} \left(\max_{0 \leq s \leq t} \left| \sum_{k \in \mathcal{C}(j)} R_k(s) - \sum_{k \in \mathcal{C}(j)} R'_k(s) \right| + \max_{0 \leq s \leq t} |A_j(s) - A'_j(s)| \right), \tag{49}
\end{aligned}$$

where the first inequality holds since $|\max_i x_i - \max_i y_i| \leq \max_i |x_i - y_i|$ and $|x^+ - y^+| \leq |x - y|$.

Now, we are ready to inductively establish (46). Fix $t \geq 0$. First, (46) holds for every leaf node i since $\mathcal{C}(i) = \mathcal{T}^-(i) = \emptyset$. Assume for induction that $i \in \mathcal{A}$ is a non-leaf node, and (46) holds for all nodes in $\mathcal{T}^-(i)$. Then,

$$\begin{aligned}
& \max_{0 \leq s \leq t} \left| \sum_{j \in \mathcal{C}(i)} R_j(s) - \sum_{j \in \mathcal{C}(i)} R'_j(s) \right| \\
&\leq \sum_{j \in \mathcal{C}(i)} \left(\max_{0 \leq s \leq t} \left| \sum_{k \in \mathcal{C}(j)} R_k(s) - \sum_{k \in \mathcal{C}(j)} R'_k(s) \right| + \max_{0 \leq s \leq t} |A_j(s) - A'_j(s)| \right) \\
&\leq \sum_{j \in \mathcal{T}^-(i)} \max_{0 \leq s \leq t} |A_j(s) - A'_j(s)|,
\end{aligned}$$

where the first inequality holds by (49), and the second inequality holds by the inductive hypothesis. This concludes the inductive step. \square

C.3.3 Proof of Proposition 5

Fix $t \geq 0$. For any node $i \in \mathcal{A}$, define $\beta_i : \mathbb{R}_{\geq 0}^n \rightarrow \mathbb{R}_{\geq 0}$ such that for every $q \in \mathbb{R}_{\geq 0}^n$, $\beta_i(q)$ denotes the matching rate from i to its children under **TTP** at time $t+1$ assuming $q(t) = q$, i.e.,

$$\beta_i(q) \triangleq \sum_{j \in \mathcal{C}(i)} \Delta D_{m(j,i)}(t+1).$$

Notice that $\beta_i(q) = 0$ for every leaf node i . By the matching rule of **TTP**, for any node $i \in \mathcal{A}$, one has the recursion

$$\beta_i(q) = \min \left\{ \lambda_i, \sum_{j \in \mathcal{C}(i)} (\lambda_j + q_j - \beta_j(q)) \right\}, \tag{50}$$

reflecting the fact that i can receive at most λ_i units of flow from its children, and each child j has at most $\lambda_j + q_j - \beta_j(q)$ units of flow available to send upward.

Recall the definition of ϵ_i 's given in (7). By Lemma 3, we also have

$$\epsilon_i = \lambda_i - \sum_{j \in \mathcal{C}(i)} \epsilon_j, \quad (51)$$

for every $i \in \mathcal{A}$. For every $q \in \mathbb{R}_{\geq 0}^n$, define

$$F(q) \triangleq \beta_r(q) + 2 \sum_{i \in \mathcal{A}_0} \beta_i(q). \quad (52)$$

By construction, $F(q)$ is exactly the total one-step decrease of $\Phi(q)$ under **TTP**: a match between r and its child reduces $\Phi(q)$ by 1, while a match between two over-demanded nodes reduces $\Phi(q)$ by 2. Hence, to prove (16), since the total arrival to over-demanded queues at time $t+1$ is $1 - \lambda_r$, it suffices to show

$$F(q) \geq 1 - \lambda_r + \min \{\epsilon, \Phi(q)\} \quad (53)$$

for every $q \in \mathbb{R}_{\geq 0}^n$.

The proof of (53) relies on two properties of $F(\cdot)$ provided in the following lemma.

Lemma 26. *The following properties hold for $F(\cdot)$:*

1. *If $\Phi(q) \leq \epsilon$, then $F(q) = 1 - \lambda_r + \Phi(q)$.*
2. *$F(\cdot)$ is non-decreasing, i.e., $F(q') \leq F(q)$ if $q' \leq q$, where the comparison is made entry-wise.*

To see that Lemma 26 implies (53), on one hand, if $\Phi(q) \leq \epsilon$, then $F(q) = 1 - \lambda_r + \Phi(q)$ by Property 1. On the other hand, if $\Phi(q) > \epsilon$, consider an arbitrary $q' \in \mathbb{R}_{\geq 0}^n$ satisfying $\Phi(q') = \epsilon$ and $q' \leq q$, and it holds that

$$F(q) \geq F(q') = 1 - \lambda_r + \epsilon,$$

where the inequality holds by Property 2, and the equality holds by Property 1.

It remains to prove Lemma 26.

Proof of Lemma 26. We prove the desired properties separately.

Proof of Property 1. Fix $q \in \mathbb{R}_{\geq 0}^n$ with $\Phi(q) \leq \epsilon$. We first show that for every $i \in \mathcal{A}$,

$$\beta_i(q) = \sum_{j \in \mathcal{C}(i)} \epsilon_j - \sum_{j \in \mathcal{T}^-(i)} (-1)^{d(i,j)} q_j, \quad (54)$$

where $d(i, j)$ is the (unweighted) distance between i and j . By definition, if i is a leaf node, then $\beta_i(q) = 0$ and $\mathcal{C}(i) = \mathcal{T}^-(i) = \emptyset$, so (54) holds. Assume for induction that (54) holds for all children

$j \in \mathcal{C}(i)$ of a non-leaf node i . Then,

$$\begin{aligned}
\beta_i(q) &= \min \left\{ \lambda_i, \sum_{j \in \mathcal{C}(i)} (\lambda_j + q_j - \beta_j(q)) \right\} \\
&= \min \left\{ \lambda_i, \sum_{j \in \mathcal{C}(i)} \left(\lambda_j + q_j - \sum_{k \in \mathcal{C}(j)} \epsilon_k + \sum_{k \in \mathcal{T}^-(j)} (-1)^{d(j,k)} q_k \right) \right\} \\
&= \min \left\{ \lambda_i, \sum_{j \in \mathcal{C}(i)} \epsilon_j - \sum_{j \in \mathcal{T}^-(i)} (-1)^{d(i,j)} q_j \right\} \\
&= \sum_{j \in \mathcal{C}(i)} \epsilon_j - \sum_{j \in \mathcal{T}^-(i)} (-1)^{d(i,j)} q_j,
\end{aligned}$$

where the first equality holds by (50), the second equality holds by the induction hypothesis, the third equality holds by (51), and the last equality holds since $\sum_{j \in \mathcal{T}^-(i)} (-1)^{d(i,j)} q_j \leq \sum_{j \in \mathcal{T}^-(i)} q_j \leq \Phi(q) \leq \epsilon$ and $\lambda_i - \sum_{j \in \mathcal{C}(i)} \epsilon_j = \epsilon_i \geq \epsilon$. This completes the inductive step.

Next, we apply (54) to show Property 1. By the definition of $F(\cdot)$,

$$\begin{aligned}
F(q) &= \beta_r(q) + 2 \sum_{i \in \mathcal{A}_0} \beta_i(q) \\
&= \sum_{j \in \mathcal{C}(r)} \epsilon_j - \sum_{j \in \mathcal{T}^-(r)} (-1)^{d(r,j)} q_j + 2 \sum_{i \in \mathcal{A}_0} \left(\sum_{j \in \mathcal{C}(i)} \epsilon_j - \sum_{j \in \mathcal{T}^-(i)} (-1)^{d(i,j)} q_j \right) \\
&= \left(\sum_{j \in \mathcal{C}(r)} \epsilon_j + 2 \sum_{i \in \mathcal{A}_0} \sum_{j \in \mathcal{C}(i)} \epsilon_j \right) - \left(\sum_{j \in \mathcal{T}^-(r)} (-1)^{d(r,j)} q_j + 2 \sum_{i \in \mathcal{A}_0} \sum_{j \in \mathcal{T}^-(i)} (-1)^{d(i,j)} q_j \right),
\end{aligned}$$

where the second equality holds by (54). We bound the above two terms separately. Firstly,

$$\sum_{j \in \mathcal{C}(r)} \epsilon_j + 2 \sum_{i \in \mathcal{A}_0} \sum_{j \in \mathcal{C}(i)} \epsilon_j \stackrel{(a)}{=} 2 \sum_{j \in \mathcal{A}_0} \epsilon_j - \sum_{j \in \mathcal{C}(r)} \epsilon_j \stackrel{(b)}{=} 1 - \lambda_r,$$

where (a) holds because the sum over children of all $i \in \mathcal{A}_0$ counts every node in $\mathcal{A}_0 \setminus \mathcal{C}(r)$ exactly once, so $\sum_{i \in \mathcal{A}_0} \sum_{j \in \mathcal{C}(i)} \epsilon_j = \sum_{j \in \mathcal{A}_0} \epsilon_j - \sum_{j \in \mathcal{C}(r)} \epsilon_j$; (b) holds because summing (51) over \mathcal{A}_0 yields $\sum_{i \in \mathcal{A}_0} \lambda_i = \sum_{i \in \mathcal{A}_0} \epsilon_i + (\sum_{i \in \mathcal{A}_0} \epsilon_i - \sum_{j \in \mathcal{C}(r)} \epsilon_j)$, and we know $\sum_{i \in \mathcal{A}_0} \lambda_i = 1 - \lambda_r$. Moreover,

$$\begin{aligned}
&\sum_{j \in \mathcal{T}^-(r)} (-1)^{d(r,j)} q_j + 2 \sum_{i \in \mathcal{A}_0} \sum_{j \in \mathcal{T}^-(i)} (-1)^{d(i,j)} q_j \\
&= \sum_{i \in \mathcal{A}_0} q_i \left(2 \sum_{j: i \in \mathcal{T}^-(j)} (-1)^{d(j,i)} - (-1)^{d(r,i)} \right) = - \sum_{i \in \mathcal{A}_0} q_i = -\Phi(q),
\end{aligned}$$

where the second equality holds since $2 \sum_{i=1}^k (-1)^i - (-1)^k = -1$ for any $k \in \mathbb{Z}_{>0}$. Combining the above three displayed equations concludes Property 1.

Proof of Property 2. For all nodes $i \in \mathcal{A}$ and $q \in \mathbb{R}_{\geq 0}^n$, define $T_i(q)$ as the total matching rate in the subtree rooted at i :

$$T_i(q) \triangleq \sum_{j \in \mathcal{T}(i)} \beta_j(q).$$

We show by induction that $T_i(\cdot)$ is non-decreasing for every node $i \in \mathcal{A}$, which implies Property 2 since $F(q) = T_r(q) + \sum_{i \in \mathcal{C}(r)} T_i(q)$.

For all leaf nodes i and $q \in \mathbb{R}_{\geq 0}^n$, $T_i(q) = \beta_i(q) = 0$, implying that $T_i(\cdot)$ is non-decreasing. Consider a non-leaf node i , and assume for induction that $T_j(\cdot)$ is non-decreasing for every $j \in \mathcal{T}^-(i)$. We now show that $T_i(\cdot)$ is also non-decreasing. Notice that

$$T_i(q) = \sum_{j \in \mathcal{T}(i)} \beta_j(q) = \beta_i(q) + \sum_{j \in \mathcal{C}(i)} T_j(q) = \min \left\{ \lambda_i, \sum_{j \in \mathcal{C}(i)} (\lambda_j + q_j - \beta_j(q)) \right\} + \sum_{j \in \mathcal{C}(i)} T_j(q),$$

where the last equality holds by (50). Using the identity $\min\{A, B\} + C = \min\{A + C, B + C\}$:

$$\begin{aligned} T_i(q) &= \min \left\{ \lambda_i + \sum_{j \in \mathcal{C}(i)} T_j(q), \sum_{j \in \mathcal{C}(i)} (\lambda_j + q_j + (T_j(q) - \beta_j(q))) \right\} \\ &= \min \left\{ \lambda_i + \sum_{j \in \mathcal{C}(i)} T_j(q), \sum_{j \in \mathcal{C}(i)} \left(\lambda_j + q_j + \sum_{k \in \mathcal{C}(j)} T_k(q) \right) \right\}, \end{aligned}$$

and the monotonicity of $T_i(\cdot)$ follows immediately from the monotonicity of $T_j(\cdot)$ for $j \in \mathcal{C}(i)$ and $T_k(\cdot)$ for $k \in \mathcal{C}(j)$. \square

C.3.4 Proof of Proposition 6

To establish that $V(Q)$ is a geometric Lyapunov function for $(Q(t))_{t \geq 0}$, we utilize Proposition 3, which requires verifying two conditions: (1) Φ has a negative truncated drift, and (2) the second-order exponential moment L_3 is bounded.

Step 1: negative truncated drift. We first show that for any initial state $x \in \mathbb{R}_{\geq 0}^n$ with $\Phi(x) > \epsilon K$:

$$\mathbb{E}_x [\max \{-\epsilon K, \Phi(Q(K)) - \Phi(x)\}] \leq -5(d_r + 1)^2 n \epsilon^{-1}. \quad (55)$$

Let $(q(t))_{t \geq 0}$ be the deterministic queue-length trajectory induced by **TTP** under the fluid arrival starting from $q(0) = x$. For any $x \in \mathbb{R}_{\geq 0}^n$, we bound the positive part of the drift:

$$\begin{aligned} [\Phi(Q(K)) - \Phi(x) + \epsilon K]^+ &\leq [\Phi(Q(K)) - \Phi(q(K))]^+ + [\Phi(q(K)) - \Phi(x) + \epsilon K]^+ \\ &\stackrel{(a)}{\leq} |\Phi(Q(K)) - \Phi(q(K))| + 1_{\{\Phi(x) \leq \epsilon K\}} \epsilon K \\ &\stackrel{(b)}{\leq} 2(d_r + 1) \sum_{i \in \mathcal{A}} \sup_{0 \leq t \leq K} |A_i(t) - \lambda_i t| + 1_{\{\Phi(x) \leq \epsilon K\}} \epsilon K, \end{aligned} \quad (56)$$

where (a) holds because

$$\Phi(q(K)) - \Phi(x) + \epsilon K \leq [\Phi(x) - \epsilon K]^+ - \Phi(x) + \epsilon K \leq 1_{\{\Phi(x) \leq \epsilon K\}} \epsilon K, \quad (57)$$

in view of Proposition 5; and (b) holds by Proposition 4.

Hence, for any initial state x with $\Phi(x) > \epsilon K$, the indicator term is zero, and we get

$$\begin{aligned}
\mathbb{E}_x [\max \{-\epsilon K, \Phi(Q(K)) - \Phi(x)\}] &= -\epsilon K + \mathbb{E}_x [\Phi(Q(K)) - \Phi(x) + \epsilon K]^+ \\
&\stackrel{(a)}{\leq} -\epsilon K + 2(d_r + 1) \sum_{i \in \mathcal{A}} \sup_{0 \leq t \leq K} |A_i(t) - \lambda_i t| \\
&\stackrel{(b)}{\leq} -\epsilon K + 2(d_r + 1) \cdot (2\sqrt{n}\sqrt{K}) \\
&\stackrel{(c)}{=} \sqrt{t_0}\epsilon^{-1} (-\sqrt{t_0} + 4(d_r + 1)\sqrt{n}) \\
&\stackrel{(d)}{=} -5(d_r + 1)^2 n \epsilon^{-1},
\end{aligned}$$

where (a) holds by (56); (b) holds by (20) in Lemma 8; (c) holds by $K = t_0 \epsilon^{-2}$; (d) holds by substituting $t_0 = 25(d_r + 1)^2 n$.

Step 2: bounding the second-order exponential moment. Next, we verify the condition from Proposition 3 involving the second-order exponential moment L_3 . Specifically, we show that for small enough θ :

$$(\epsilon\theta) \cdot L_3(\epsilon K, \epsilon\theta, K) \leq 5(d_r + 1)^2 n \epsilon^{-1}. \quad (58)$$

Recall the definition of L_3 in (13) and $t_0 = 25(d_r + 1)^2 n$. It holds that

$$\begin{aligned}
L_3(\epsilon K, \epsilon\theta, K) &= \sup_{x \in \mathbb{R}_{\geq 0}^n} \mathbb{E}_x \left[(\max\{-\epsilon K, \Phi(Q(K)) - \Phi(x)\})^2 \cdot \exp(\epsilon\theta(\Phi(Q(K)) - \Phi(x))^+) \right] \\
&\stackrel{(a)}{\leq} \mathbb{E} \left[\left(2\epsilon K + 2(d_r + 1) \sum_{i \in \mathcal{A}} \sup_{0 \leq t \leq K} |A_i(t) - \lambda_i t| \right)^2 \exp \left(\epsilon\theta \cdot 2(d_r + 1) \sum_{i \in \mathcal{A}} \sup_{0 \leq t \leq K} |A_i(t) - \lambda_i t| \right) \right] \\
&\stackrel{(b)}{\leq} 8\epsilon^2 K^2 \mathbb{E} \left[\exp \left(\epsilon\theta \cdot 2(d_r + 1) \sum_{i \in \mathcal{A}} \sup_{0 \leq t \leq K} |A_i(t) - \lambda_i t| \right) \right] \\
&\quad + 8(d_r + 1)^2 \mathbb{E} \left[\left(\sum_{i \in \mathcal{A}} \sup_{0 \leq t \leq K} |A_i(t) - \lambda_i t| \right)^2 \exp \left(\epsilon\theta \cdot 2(d_r + 1) \sum_{i \in \mathcal{A}} \sup_{0 \leq t \leq K} |A_i(t) - \lambda_i t| \right) \right] \\
&\stackrel{(c)}{\leq} 8\epsilon^2 K^2 \cdot C_2 \left(2\epsilon\theta\sqrt{K}(d_r + 1) \right) + 8(d_r + 1)^2 K \cdot C_3 \left(2\epsilon\theta\sqrt{K}(d_r + 1) \right), \quad (59)
\end{aligned}$$

where (a) holds because for any state x ,

$$\begin{aligned}
|\max\{-\epsilon K, \Phi(Q(K)) - \Phi(x)\}| &= |-\epsilon K + [\Phi(Q(K)) - \Phi(x) + \epsilon K]^+| \\
&\leq \epsilon K + [\Phi(Q(K)) - \Phi(x) + \epsilon K]^+ \\
&\leq \epsilon K + \left(\epsilon K + 2(d_r + 1) \sum_{i \in \mathcal{A}} \sup_{0 \leq t \leq K} |A_i(t) - \lambda_i t| \right)
\end{aligned}$$

by (56); (b) holds because $(a + b)^2 \leq 2a^2 + 2b^2$; (c) holds by Lemma 8 with the scaling \sqrt{K} .

In the following, we aim to find the range of θ that satisfies

$$L_3(t_0 \epsilon^{-1}, \epsilon\theta, t_0 \epsilon^{-2}) \leq \frac{5(d_r + 1)^2 n \epsilon^{-1}}{\epsilon\theta}.$$

Rearranging, it suffices to show

$$\theta \leq \frac{5(d_r + 1)^2 n \epsilon^{-2}}{L_3(t_0 \epsilon^{-1}, \epsilon \theta, t_0 \epsilon^{-2})} \triangleq \Delta. \quad (60)$$

Recall the upper bound for $L_3(\epsilon K, \epsilon \theta, t_0 \epsilon^{-2})$ derived in (59) with coefficients substituted ($K = t_0 \epsilon^{-2}$, so $\epsilon K = t_0 \epsilon^{-1}$ and $\epsilon \sqrt{K} = \sqrt{t_0}$):

$$L_3(\epsilon K, \epsilon \theta, t_0 \epsilon^{-2}) \leq 8t_0^2 \epsilon^{-2} C_2(\tilde{\theta}) + 8(d_r + 1)^2 t_0 \epsilon^{-2} C_3(\tilde{\theta}),$$

where $\tilde{\theta} \triangleq 2\theta \sqrt{t_0}(d_r + 1)$. Substituting this into Δ :

$$\Delta = \frac{5(d_r + 1)^2 n \epsilon^{-2}}{\epsilon^{-2} [8t_0^2 C_2(\tilde{\theta}) + 8(d_r + 1)^2 t_0 C_3(\tilde{\theta})]} = \frac{5(d_r + 1)^2 n}{8t_0^2 C_2(\tilde{\theta}) + 8(d_r + 1)^2 t_0 C_3(\tilde{\theta})}. \quad (61)$$

Substituting $t_0 = 25(d_r + 1)^2 n$:

- $8t_0^2 = 8(625(d_r + 1)^4 n^2) = 5000(d_r + 1)^4 n^2$,
- $8(d_r + 1)^2 t_0 = 8(d_r + 1)^2 (25(d_r + 1)^2 n) = 200(d_r + 1)^4 n$.

We assume $\theta = \frac{\kappa}{(d_r + 1)^2 n}$ for some $\kappa > 0$ determined later. Then $\tilde{\theta} = \frac{10\kappa}{\sqrt{n}}$. Substituting $C_2(\tilde{\theta}) = e^{20\kappa + O(1/n)}$ and $C_3(\tilde{\theta}) = 4ne^{20\kappa + O(1/n)}$, the inequality $\theta \leq \Delta$ becomes:

$$\frac{\kappa}{(d_r + 1)^2 n} \leq \frac{5(d_r + 1)^2 n}{(d_r + 1)^4 n [5000ne^{20\kappa + O(1/n)} + 200(4ne^{20\kappa + O(1/n)})]}.$$

Simplifying:

$$\kappa \leq \frac{5}{5000e^{20\kappa + O(1/n)} + 800e^{20\kappa + O(1/n)}} = \frac{1}{1160e^{20\kappa + O(1/n)}}.$$

Since $f(\kappa) = 1160\kappa e^{20\kappa}$ is increasing, there exist constants $n_0 > 0$ and $\kappa_0 > 0$ such that for all $n \geq n_0$ and $\theta \leq \frac{\kappa_0}{(d_r + 1)^2 n}$, the condition holds.

Conclusion. Conditions (55) and (58) satisfy the requirements of Proposition 3 with drift magnitude $\gamma = 5(d_r + 1)^2 n \epsilon^{-1}$. Thus, $V(Q)$ is a geometric Lyapunov function for $(Q(t))_{t \geq 0}$ with drift size parameter $1 - \frac{\epsilon}{2}\gamma = 1 - \frac{5}{2}(d_r + 1)^2 n \theta$.

C.3.5 Proof of Theorem 3

By Proposition 6, under the true arrival, $V(Q) = \exp(\epsilon \theta \cdot \Phi(Q))$ is a geometric Lyapunov function with geometric drift size parameter: $1 - \frac{5}{2}(d_r + 1)^2 n \theta$, drift time parameter: $K = t_0 \epsilon^{-2}$, and exception parameter $\exp(\theta t_0)$. By Proposition 5, we get

$$\mathbb{E}_\pi [e^{\epsilon \theta \cdot \Phi(Q)}] \leq \frac{2e^{\theta t_0} \phi(K)}{5(d_r + 1)^2 n \theta}, \quad (62)$$

where $\phi(K) \triangleq \sup_{x \in \mathbb{R}_{\geq 0}^n} \mathbb{E}_x [e^{\epsilon \theta (\Phi(Q(K)) - \Phi(x))}]$ is the maximum expected overshoot. We upper bound $\phi(K)$ by

$$\begin{aligned} \phi(K) &= \sup_{x \in \mathbb{R}_{\geq 0}^n} \mathbb{E}_x [e^{\epsilon \theta (\Phi(Q(K)) - \Phi(x) + \epsilon K)}] \cdot e^{-\epsilon^2 \theta K} \leq \sup_{x \in \mathbb{R}_{\geq 0}^n} \mathbb{E}_x [e^{\epsilon \theta [\Phi(Q(K)) - \Phi(x) + \epsilon K]^+}] \cdot e^{-\epsilon^2 \theta K} \\ &\leq \mathbb{E} \left[\exp \left(2\epsilon \theta (d_r + 1) \sum_{i \in \mathcal{A}} \sup_{0 \leq t \leq K} |A_i(t) - \lambda_i t| \right) \right] \leq C_2(\tilde{\theta}), \end{aligned}$$

where the second inequality holds by (56), and $\tilde{\theta} \triangleq 2(d_r + 1)\sqrt{t_0}\theta$. By Jensen's inequality,

$$\begin{aligned} \mathbb{E}_\pi[\Phi(Q)] &\leq \frac{1}{\epsilon\theta} \log \mathbb{E}_\pi[e^{\epsilon\theta \cdot \Phi(Q)}] \\ &\leq \frac{1}{\epsilon\theta} \left[\log C_2(\tilde{\theta}) + \log \left(\frac{2e^{\theta t_0}}{5(d_r + 1)^2 n \theta} \right) \right] \\ &\leq \left(\frac{n(d_r + 1)^2}{\epsilon \kappa_0} \right) \underbrace{\left[\left(20\kappa_0 + \frac{50\kappa_0^2}{n} \right) + \left(25\kappa_0 + \log \frac{2}{5\kappa_0} \right) \right]}_{O(1)} = O\left(\frac{nd_r^2}{\epsilon} \right), \end{aligned} \quad (63)$$

where the second inequality holds by (62), and the last inequality holds because by $t_0 = 25(d_r + 1)^2 n$ and $\theta = \kappa_0(d_r + 1)^{-2}n^{-1}$ from Proposition 6, we get

$$\frac{1}{\epsilon\theta} = \frac{1}{\epsilon} \cdot \frac{(d_r + 1)^2 n}{\kappa_0} = \frac{1}{\kappa_0} \cdot \frac{n(d_r + 1)^2}{\epsilon},$$

$\theta t_0 = \left(\frac{\kappa_0}{(d_r + 1)^2 n} \right) (25(d_r + 1)^2 n) = 25\kappa_0$, and $\tilde{\theta} = 2(d_r + 1)(5(d_r + 1)\sqrt{n}) \left(\frac{\kappa_0}{(d_r + 1)^2 n} \right) = \frac{10\kappa_0}{\sqrt{n}}$, then

$$\log C_2(\tilde{\theta}) = 2\sqrt{n} \left(\frac{10\kappa_0}{\sqrt{n}} \right) + \frac{1}{2} \left(\frac{10\kappa_0}{\sqrt{n}} \right)^2 = 20\kappa_0 + \frac{50\kappa_0^2}{n},$$

and

$$\log \left(\frac{2e^{\theta t_0}}{5(d_r + 1)^2 n \theta} \right) = \log \left(\frac{2e^{25\kappa_0}}{5\kappa_0} \right) = 25\kappa_0 + \log \left(\frac{2}{5\kappa_0} \right).$$

Finally, **TTP** is consistent by Proposition 7, and hence applying Lemma 2 yields

$$\mathbb{E} \left[\sum_{i \in \mathcal{A}_0} Q(t) \right] \leq 2\mathbb{E}_\pi \left[\sum_{i \in \mathcal{A}_0} Q(0) \right] = 2\mathbb{E}_\pi[\Phi(Q)] \leq O\left(\frac{nd_r^2}{\epsilon} \right)$$

for every $t \geq 0$, where the last inequality holds by (63). Combining the above displayed equation with Lemma 1 concludes the proof.