Online Decision Making in Networked Marketplaces

Pengyu Qian

Submitted in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy
under the Executive Committee
of the Graduate School of Arts and Sciences

COLUMBIA UNIVERSITY

2021

ABSTRACT

Online Decision Making in Networked Marketplaces

Pengyu Qian

Modern, technologically-enabled markets are disrupting many industry sectors, including transportation, labor, lodging, dating services and others. While the system operator is able to collect data and deploy various control levers, these systems are highly complex, marked by a large number of interacting self-interested agents, uncertainty about the future and imperfect demand predictions. There remain major challenges in optimizing these marketplaces. In this dissertation, I describe work designing novel algorithms and performing theoretical analysis of networked systems, including those that arise in marketplaces. I demonstrate how to use tools from applied probability, modern optimization, and economics to develop methodologies for online decision making in contexts such as queueing control, revenue management, and running a matching platform.

The first part of the dissertation designs novel algorithms for dynamic assignment and revenue management. The work considers networked systems where agents or tasks arrive over time, which is broadly relevant to service platforms with heterogeneous services, for instance shared transportation systems. Firstly, we propose a near optimal "mirror backpressure" control methodology for joint entry/assignment/pricing control in a network where there are a fixed number of supply units (vehicles), and demands with different origin and destination nodes arrive over time. The MBP policy does not need demand arrival rate predictions at all, and we prove guarantees of near optimal performance over a finite horizon. Secondly, we study a special case of the network control problem where

the geographical imbalances in demand are small enough such that, ignoring stochasticity, they can be corrected using assignment control alone. The objective is to minimize the fraction of customers who are "lost" (not served) because there is no vehicle at a nearby location when the customer arrives. We show that for this setting we can achieve a refined notion of optimality, i.e., the large deviations optimality.

The second part of the dissertation analyzes equilibria in matching markets under different mechanisms. Firstly, we study the Gale-Shalpley "deferred acceptance" algorithm, which has been successfully adopted in contexts such as school choice and resident matching programs. Our research question is, "Which Gale-Shapley matching markets exhibit a short-side advantage?" I.e., in which markets does being on the short side of the market allow agents to obtain better match partners relative to a similar "balanced" market with equal numbers of agents on the two sides? We address this problem by looking at the "random matching market" model where each agent considers only a subset of potential partners on the other side, and sharply characterize the resulting (nearly unique) stable matching, overcoming significant technical challenges. Secondly, we study the waiting-list mechanism, which is commonly used in kidney assignment, public housing allocation, and beyond. We show that the waiting-list mechanism is near-optimal in terms of allocative efficiency for general systems with an arbitrary number of agent types and item types, and obtain tight bound on the efficiency loss. Comparing to existing works which could only analyze very simple systems, we tackle the general case by taking a completely different approach and establishing a novel connection with stochastic gradient descent.

Contents

List of Figures

List of Tables

Acknowledgments

To my advisor, Yash Kanoria. You have been the best advisor I could have hoped for – brilliant, patient, and trusting me with freedom to explore as I slowly learned to navigate. You gave me the support when I needed it the most, and pushed me outside of my comfort zone. You've taught me to enjoy doing research in its purest form, and the importance of practicing writing, communication, and presentation. You've been a great friend and took my success personally, for that I am profoundly grateful. It makes me proud to be one of your first (two) PhD students.

I am very fortunate to have worked with so many wonderful collaborators – Sid Banerjee, Itai Ashlagi, Amin Saberi, and Jacob Leshno. To Sid, for infecting me with your enthusiasm in tackling challenging problems and having a breadth of interest. To Itai and Amin, for inviting me to visit Stanford, which leads to the Chapter 4 of this dissertation, and for being proactive supporters. To Jacob, for possessing the economic intuition that I desire to have. Collaborating with them has been a privilege. I am deeply grateful to their generous support during my job market.

I am greatly indebted to the other members of my thesis committee, Assaf Zeevi, Costis Maglaras, Vineet Goyal, and Ilan Lobel, for their feedback, and to other professors in the DRO division of Columbia Business School and Columbia IEOR who have provided invaluable advice and mentorship throughout my years as a PhD student. To Assaf, for being a role model that I aspire to become someday, for showing me that one can possess such deep insights and powerful technical skills. To Omar, for being an encouraging mentor and supporter who is always ready to give advice and offer help when I need it. To Nick, Awi, Paul, Costis, Ciamac, Dan, Fanyin, Henry, Jay, Ward for their constructive feedback on my research in the last few years. To Santiago, Carri, Jing, Hongyao, Vineet, Adam for their encouragement and patient guidance during my job market year. To all the wonderful CBS staff, especially Liz, Dan, Clara, Winnie, Cristina and Maria, for their tireless support and warm friendship.

I will never forget the joy and laughter brought by my fellow PhD students at DRO and IEOR, and beyond. To Seungki, for the countless hours we spent walking around campus talking about research and life, for being an incredibly efficient collaborator, for being my wedding witness and friend for life. To Jiaqi, for being my first academic sister. To Barış, for the wonderful Turkish breakfast. To Kleanthis, for the neat lecture notes. To all my friends, too many to list here, who made my PhD experience fun, and from whom I've learned a lot.

I was introduced to the world of operations research by my undergraduate thesis supervisors, Zaiwen Wen and Zizhuo Wang. I am thankful to them for leading me into this wonderful field, and for advocating for me during my PhD application.

I am profoundly grateful to my parents, for always being patient with my naive questions since I was a little kid, for supporting me in my move halfway across the world, for always believing in me, unconditionally loving me. I hope my grandpa would be proud of me.

Lastly, I would like to thank Qiuyuan, who has been my pillar of strength since I knew her in college. She put her trust in me and moved with me across the ocean, overcame countless resulting obstacles in work and life, and excelled in her profession. She has been willing to listen to my research struggles, tolerate my bad moods, and enlightens me with new perspectives. Her constant caring, love and support have not only been essential during these years but have also made me realize what's important in life.

Introduction

Many of the marketplaces have been reshaped by technology. The deployment of various algorithms enabled billions of people to get a ride at the click of a button via Uber, Lyft or other ride-hailing platforms. Systems based on matching market algorithms and mechanisms are used to match people to their potential love interests, students to schools, and donor's organs to patients. While the technological development have greatly benefited the society, they lead to highly complex systems, and there remain major challenges in optimizing these marketplaces. For instance, most ride-hailing companies continue to make most tactical decisions in fairly "naïve" ways: e.g., without using demand predictions, and accounting for supply availability only near the origin location. The important issues for these systems lie at the intersection of engineering, mathematics and economics, and, specifically, the fields of operations research and operations management.

In this dissertation, we focus on algorithms and mechanisms that not only have good theoretical guarantees, but also are simple, robust, and hence practical for real-world systems. The dissertation has two parts: Part I designs novel near-optimal dynamic assignment and pricing algorithms that could be useful in a realistic environment. Part II addresses foundational questions regarding the nature of equilibria in matching markets under different mechanisms and market compositions.

In Part I, we study the design of dynamic assignment and pricing policies in networked systems where agents or tasks arrive over time. We focus on the queueing network model, which is a canonical model of these systems. This work is broadly relevant to service platforms with heterogeneous services, for instance shared transportation systems. The dynamic control problem in these systems are notoriously challenging for the following reasons: (i) Control decisions not only generate payoff, but also modulate the distribution of resources in the network over time, thus creating a tension. (ii) Complex network externality. For example, vehicles relocate in ride-hailing systems, hence any control decision affects the future availability of resources throughout the system within short timescales.

We aim at developing approaches to tackle these challenges systematically. Prior literature on the control of queueing networks typically solves the optimal control problem in the fluid limit or diffusion limit, and relies on the exact predictions of future arrival rates. As a result, these approaches are generally sensitive to the errors in predictions, and might have poor performance in a realistic environment. Part I designs simple policies that do not use demand predictions and still achieve near optimality, and therefore are more useful in practice.

Part I consists of two chapters. In Chapter 1, we propose a near optimal "mirror backpressure" control methodology for joint entry/assignment/pricing control in a network where there are a fixed number of supply units (vehicles), and demands with different origin and destination nodes arrive over time. Mirror backpressure (MBP) autocorrects geographical supply imbalances by aggressively protecting and replenishing supply where it is scarce, while deploying supply from regions where it is plentiful. The paper makes several notable contributions:

1. The MBP policy is "blind", i.e., it does not need demand arrival rate predictions at all, in sharp contrast to previous work which relies on perfect estimates of future demand arrival rates.
2. We prove guarantees of near optimal performance over a finite horizon, and moreover allow demand arrival rates to be (slowly) time-varying. This is a major improvement upon the steady state guarantees with stationary demand obtained in previous work.
3. Our methodology provides a systematic way to construct simple control policies across a variety of levers for queueing networks with provable guarantees. Our policy design uses two ideas: the celebrated backpressure policy for network control, and the mirror descent algorithm for optimization (a generalization of gradient descent). Backpres-

sure turns out to be inadequate for our problem, because it suffers from "underflow" when the controller wants to serve a customer but no resource is available at the customer's origin. Our MBP policy systematically resolves this issue. We make the crucial observation that under MBP, the queue length vector executes dual stochastic mirror descent on the fluid optimization problem. The policy takes a very simple form, making it easy to communicate in practice for its implementation.

In Chapter 2, we study a special case of the network control problem where the geographical imbalances in demand are small enough such that, ignoring stochasticity, they can be corrected using dispatch (i.e., the assignment of vehicles to customers) control alone. (The condition is known as the Complete Resource Pooling (CRP) condition in the queueing literature.) The objective is to minimize the fraction of customers who are "lost" (not served) because there is no vehicle at a nearby location when the customer arrives. We show that for this setting we can achieve a refined notion of optimality. We make the following contributions:

1. We show that a remarkably simple "MaxWeight" control policy serves almost all customers. The policy simply dispatches from that location near the customer which currently has the most vehicles. Note that the MaxWeight policy is also blind, i.e., it requires no knowledge of demand arrival rates.
2. We obtain a large deviations optimal dispatch control policy (in terms of demand arrival rates). Our policy, which we call "Scaled MaxWeight" is a straightforward generalization of MaxWeight: it employs a supply "scaling factor" for each location, and dispatches from the nearby location with the largest scaled number of vehicles. The optimal scaling factors can be computed using the demand arrival rates if the latter are known. Even if suboptimal scaling factors are used, very few customers are lost. We obtain these results by performing the first large deviations analysis of a queueing network under the CRP condition. This work may inspire similar large deviations analyses of other queueing network settings, and lead to new fine-grained control insights.

In Part II we try to demystify the nature of equilibria in various marketplaces. An important step towards designing better marketplaces is understanding the performance of currently widely used mechanisms/algorithms. Examples of popular mechanisms/algorithms that I study include the Gale-Shalpley "deferred acceptance" algorithm, which is the bedrock of Shapley's Nobel Prize in Economics, and has been successfully adopted in contexts such as school choice and resident matching programs. Another example is the waiting-list mechanism, which is commonly used in kidney assignment, public housing allocation, and beyond. However, due to the complex interaction of heterogeneous agents in these networked marketplaces, many foundational questions regarding the equilibria remains unanswered.

Part II also consists of two chapters. In Chapter 3, we study the Gale-Shapley two-sided matching market model, where agents on both sides have ordinal preferences over potential partners on the other side. This model has been instrumental in the design of numerous real-world marketplaces. We raise the following research question: "Which Gale-Shapley matching markets exhibit a short-side advantage?" I.e., in which markets does being on the short side of the market allow agents to obtain better match partners relative to a similar "balanced" market with equal numbers of agents on the two sides? We address this problem by looking at the "random matching market" model (with uniform and independent agent preference rankings on both sides) where each agent considers only a subset of potential partners on the other side, with $n+k$ men and n women which are "partially connected" (each agent considers only d potential partners on the other side), and sharply characterize the resulting (nearly unique) stable matching, overcoming significant technical challenges. The economic interpretation of our finding is striking and represents a significant advance in our understanding of matching markets without money: a market exhibits a short-side advantage if and only if the number of short side agents who remain unmatched is smaller than the market imbalance k. One nice consequence of this finding for researchers in the field is that they can now estimate whether a market exhibits a short-side advantage from publicly available summary statistics alone. At the heart of the paper is a very simple observation: #(unmatched on short side)+ #imbalance = #(unmatched on long side). If the first term on the left-hand side dominates, the numbers unmatched on the two sides are comparable and there is no significant short side advantage. In contrast, if the imbalance term dominates the left-hand side dominates, there are many more unmatched on the long side, and the short side is matched to more preferred partners (correspondingly, fewer short side agents entirely fail to find a partner).

In Chapter 4, we focus on the waiting list mechanisms. Waiting-lists are common assignment mechanisms for allocating scarce goods that arrive stochastically over time. The mechanism can be illustrated by the classic example of public housing allocation in Boston: each family (i.e., agents) eligible for public housing can join the waiting-list of a type of housing project, each type of housing projects (i.e., items) become available over time, and they are offered to the families on the waiting-list in a first-come-first-served manner. In these systems, a key metric is the quality of matches, i.e., the allocative efficiency. However, despite its widespread use, relatively little is known about the efficiency of the waiting-list mechanism. We show that the waiting-list mechanism is near-optimal in terms of allocative efficiency for general systems with an arbitrary number of agent types and item types, and obtain tight bound on the efficiency loss. The first fundamental theorem of welfare economics tells us that in markets with money, competitive equilibrium leads to efficient allocations. But the waiting-list mechanism is non-monetary, which makes it pleasantly surprising that it is allocatively efficient. Our approach is completely different from previous ones, and establishes a novel connection with stochastic gradient descent. Interestingly, the waiting costs of each waiting-list serve as shadow prices for the items.

Bibliography notes

Chapter 1 Joint with Yash Kanoria. Appeared as an extended abstract at the ACM Conference on Economics and Computation (EC), in 2020. A journal version is in submission.

Chapter 2 Joint with Siddhartha Banerjee, and Yash Kanoria. Appeared as an extended abstract at the ACM SIGMETRICS, in 2018. A journal version is in submission.

Chapter 3 Joint with Yash Kanoria, and Seungki Min. Appeared as an extended abstract at the ACM-SIAM Symposium on Discrete Algorithms (SODA), in 2021. A journal version is in submission.

Chapter 4 Joint with Itai Ashlagi, Jacob Leshno, and Amin Saberi. Appeared as an extended abstract at the ACM Conference on Economics and Computation (EC), in 2020. A journal version is in preparation.

For each work, I contributed substantially to all aspects of the research and writing process.

Part I:

Novel Algorithms for Dynamic Resource Allocation

CHAPTER 1

Blind Dynamic Resource Allocation in Closed Networks via Mirror Backpressure

1.1 Introduction

The control of complex systems with circulating resources such as shared transportation platforms and scrip systems has been heavily studied in recent years. The hallmark of such systems is that serving a demand unit causes a (reusable) supply unit to be relocated. Closed queueing networks (i.e., networks where a fixed number of supply units circulate in the system) provide a powerful abstraction for these applications ([1], [2]). The key challenge is *managing the distribution of supply* in the network. A widely adopted approach for this problem is to solve the deterministic optimization problem that arises in the continuum limit (often called *the static planning problem*), and show that the resulting control policy is near-optimal in a certain asymptotic regime. However, this approach only works under the restrictive assumption that *(1) the system parameters (demand arrival rates) are precisely known.* Furthermore, previous papers ([1], [2]) assume that *(2) the system is in steady state.* As is pointed out by Banerjee, Freund, and Lykouris [1], relaxing either of these assumptions has been of interest.

In this paper, we relax *both* assumptions.[1] We propose a family of simple, practical

[1]The paper Banerjee, Kanoria, and Qian [3] is similarly motivated, but restricts attention to a narrow special case: assignment control in networks satisfying a strong *complete resource pooling* (CRP) assumption, and conducts a sharp large deviations analysis. In particular, non-idling/greedy policies

control policies that are *blind* in that they use *no* prior knowledge of demand arrival rates, and prove strong transient and steady state performance guarantees for these policies, for demand arrival rates that are stationary or vary slowly in time. In simulations, our policies achieve excellent performance that beats the state-of-the-art policies even in an unequal contest where the latter policies are provided exact demand arrival rates whereas our proposed policies are given *no* prior information about demand arrival rates.

Informal description of the model. For ease of exposition, our baseline setting is one where entry control is the only available control lever, and demand is stationary. Later we allow other controls including dynamic pricing, and flexible assignment of resources, and moreover allow for time-varying demand arrival rates, and show that our machinery and guarantees extends seamlessly. In our baseline entry control model, we consider a closed queueing network consisting of m nodes (locations), and a fixed number K of supply units that circulate in the system. Demand units with different origin-destination node pairs arrive stochastically over slotted time with some stationary arrival rates which are unknown to the controller. The controller dynamically decides whether to admit each incoming demand unit. Each admission decision has two effects: it generates a certain *payoff* depending on the origin and destination of the demand unit, and it causes a supply unit to *relocate* from the origin to the destination instantaneously, if the origin node is non-empty. The goal of the system is to maximize the collected payoff over a period of time.

Notably, the greedy policy, which admits a demand unit if a supply unit is available, is generically far from optimal: even as $K \to \infty$, the optimality gap per demand unit of this policy is $\Omega(1)$ even in steady state; see Remark 1.1 in Section 1.2. The intuition is that some nodes have no available supply an $\Omega(1)$ fraction of the time in steady state under the greedy policy, and so the policy is forced to drop a significant proportion of the demand

suffice to achieve asymptotic optimality under CRP. In contrast, the present work is general: e.g., the JEA setting we solve in Section 1.6 generalizes the model of that paper by dropping the CRP assumption, necessitating a completely different non-greedy approach to control; already under our illustrative model (Section 1.2) the CRP assumption of Banerjee, Kanoria, and Qian [3] is automatically violated and the greedy policy fails to achieve asymptotic optimality.

which would have been served under the optimal policy. Furthermore, if demand arrival rates are imperfectly known, any state independent policy [such as that of 1] generically suffers a steady state optimality gap per demand unit of $\Omega(1)$; see Banerjee, Kanoria, and Qian [3, Proposition 4].

Preview of our main result. We propose a large class of simple and practical control policies that are blind (i.e., require *no* estimates of the demand arrival rates), and show that, under a mild connectivity assumption on the network, the policies are near optimal. Specifically, we show that our policies lose payoff (per demand unit) at most $O\left(\frac{K}{T} + \frac{1}{K}\right)$ relative to the optimal policy that knows the demand arrival rates, where K is the number of supply units, T is the number of demand units that arrive during the period of interest. Our result is *non-asymptotic*, i.e., our performance guarantee holds for finite K and T, and thus covers both transient and steady state performance. In particular, taking $T \to \infty$, we obtain a steady state optimality gap of $O(\frac{1}{K})$, matching that of the state-of-the-art policy of Banerjee, Freund, and Lykouris [1], though that policy requires perfect estimates of demand arrival rates, in sharp contrast to our policy which is completely blind. Our bound further provides a guarantee on transient performance: the horizon-dependent term K/T in our bound on optimality gap is small if the total number of arrivals T over the horizon is large compared to the number of supply units K. Notably, our bound does not deteriorate as the system size increases in the "large market regime" where the number of supply units K increases proportionally to the demand arrival rates (see the discussion after Theorem 1.1): here the number of arrivals $T = \Theta(K \cdot T^{\text{real}})$, where T^{real} is the time horizon measured in physical time, and we can rewrite our bound on the optimality gap as $O\left(\frac{1}{T^{\text{real}}} + \frac{1}{K}\right) \xrightarrow{K \to \infty} O\left(\frac{1}{T^{\text{real}}}\right)$.

Our policies retain their good performance if demand arrival rates vary slowly over time: We show (see Section 1.6.2) that the loss in payoff per customer under MBP is bounded by $O\left(\frac{K}{T} + \frac{1}{K} + \sqrt{\eta K}\right)$, where is η is the maximum change in demand arrival rates per customer arrival. In the aforementioned large market regime, the optimality gap per customer can be expressed as $O\left(\frac{1}{T^{\text{real}}} + \frac{1}{K} + \sqrt{\zeta}\right)$ where $\zeta \triangleq \eta K$ is rate of change of ϕ with respect to physical time.

We now motivate and introduce our control policies. First, we describe how our problem is one of controlling a closed queueing network. Next, we describe the celebrated backpressure methodology for blind control of queueing networks. We then outline the central challenge in using backpressure in settings like ours. Finally, we introduce our proposed policies which significantly generalize backpressure, and may be broadly useful.

Analogy with control of a closed queueing network. Our problem can be viewed as one of optimal control of a closed queueing network. In the terminology of classic queueing theory, the K supply units are "jobs", and each node in our model has both a queue of jobs (supply units) as well as a "server" which receives a "service token" each time a demand unit arrives with that location as the origin. (We emphasize the reversal of the usual mapping: in our setup supply units are "jobs" and demand units act as service tokens.) Our model also specifies the "routing" of jobs: service tokens are labeled with a destination queue to which the served job (supply unit) moves. Since jobs circulate in the system (they do not arrive or leave), our setup is a *closed* queueing network.[2] (Networks where jobs arrive, go through one or more services, and then leave, are called *open* networks.)

Backpressure. Our control approach is inspired by the celebrated backpressure methodology of Tassiulas and Ephremides [4] for the control of queueing networks. Backpressure simply uses queue lengths as congestion costs (the shadow prices to the flow constraints; the flow constraint for each queue is that the inflow must be equal to the outflow in the long run), and chooses a control decision at each time which maximizes the myopic payoff inclusive of congestion costs. Concretely, in our baseline entry control setting, backpressure admits a demand if and only if the payoff of serving the demand plus the origin queue length exceeds the destination queue length. This simple approach has been used very effectively in a range of settings arising in cloud computing, network-

[2]There are subtle differences between our model and "classical" closed queueing networks in the timing of when a job joins the destination queue, and when the "service" of a job is initiated. These differences are non-essential, see, e.g., Banerjee, Kanoria, and Qian [3, Section 8].

ing, etc.; see, e.g. [5]. Backpressure is provably near-optimal (in the large market limit) in many settings where payoffs accrue from serving jobs, because it has the property of executing dual stochastic gradient descent (SGD) on the controller's deterministic (continuum limit) optimization problem. As we discuss next, this property breaks down when the so-called "no-underflow constraint" binds, making it very challenging to use backpressure in our setting (indeed, this difficulty appears to be the reason that backpressure has not yet been proposed as a control approach in such settings with circulating resources).

Main challenge: no-underflow constraint. The control policy must satisfy the *no-underflow constraint*, namely, that each decision to admit a demand unit needs to be backed by an available supply unit at the origin node of the demand. This constraint couples together the present and future decisions, and presents a challenge in deploying the backpressure methodology in numerous settings, including ours.

In certain settings this constraint does *not* pose a problem: For example, in the well known "crossbar switch" problem in [6], there are no "payoffs" apart from the shadow prices (the goal is only to prevent queues from building up), so backpressure only recommends to serve a queue with positive length (after all, backpressure only serves a queue if it is longer than the destination queue) and so the no-underflow constraint does not bind. In several works that do include payoffs, the authors *make strong assumptions* to similarly ensure the constraint does not bind.[3] In our setting, payoffs are essential (there is value generated by serving a customer), and so the constraint *does* bind.

A machinery that introduces *virtual queues* has been developed to extend backpressure to settings where the constraint binds; see, e.g., [11]. The main idea is to introduce a "fake" supply unit into the network each time the constraint binds, to preserve the SGD property of backpressure. In open queueing networks, these fake supply units eventually

[3]For example, [7] assume that the network satisfies a so-called Extreme Allocation Available (EAA) condition, which ensures that the no-underflow constraint does not bind; [8] assumes that payoffs are generated only by the source nodes, which have infinite queue lengths. [9] consider networks where the payoffs are generated only by the output nodes, and show that a variant of backpressure avoids underflow entirely under this assumption. [10] assume that the network satisfies a so-called Dedicated Item (DI) condition.

leave the system, and so have a small effect (under appropriate assumptions). In our closed network setting, these fake supply units, once created, never leave and so would build up in the system, leading to very poor performance. In Section 1.4.3, we provide a detailed discussion of the challenge posed by the no-underflow constraint, and how it prevents us from using backpressure as is.

Our solution: Mirror Backpressure. In solving this problem, we introduce a novel class of policies which we call *Mirror Backpressure.* MBP generalizes the celebrated backpressure (BP) policy and is as simple and practical as BP. Whereas BP uses the queue lengths as congestion costs, MBP employs a flexibly chosen *congestion function* to translate from queue lengths to congestion costs. MBP features a simple and intuitive structure: for example, in the entry control setting, the platform admits a demand only if the payoff of serving it outweighs the difference between congestion costs at the destination and origin of the demand. Crucially, the congestion function is designed so that MBP has the property that it executes dual *stochastic mirror descent* [12, 13] on the platform's continuum limit optimization problem, with the chosen mirror map.[4] The mirror map can be flexibly chosen to fit the problem geometry arising from the no-underflow constraints. Roughly, we find better performance with congestion functions which are steep for small queue lengths, the intuition being that this makes MBP more aggressive in protecting the shortest queues (and hence preventing underflow). In case of finite buffers, we find it beneficial to use congestion functions which moreover increase steeply as the queue length approaches buffer capacity, to prevent buffer overflow (Section 1.6.1).

We develop a general machinery to prove performance guarantees for MBP, which draws inspiration from two distinct toolkits: the machinery for proving convergence of mirror descent from the optimization literature, and the Lyapunov drift method from the network control literature. We provide a ready Lyapunov function for any MBP policy. Furthermore, we improve upon the Lyapunov drift method to obtain a sharp bound on the

[4]The special case of the congestion function being the identity function corresponds to standard BP, which has the property of executing stochastic gradient descent, a special case of mirror descent [14].

suboptimality caused by the no-underflow constraint. Our analysis exploits the structure of the platform's continuum limit optimization problem in a novel way (see Section 1.5).

Our work fits into the broad literature on the control of stochastic processing networks [15]. Our MBP methodology for designing blind control policies with provable guarantees applies to open queueing networks as well. We are optimistic that MBP will prove broadly useful in the control of queueing networks.

Main contributions. To summarize, we make two main contributions in this paper:
(i) **Mirror Backpressure: a class of near-optimal control policies for queueing networks that are completely blind.** In general settings that consider entry control, pricing, and flexible assignment, we propose a family of dynamic control policies for queueing networks, the Mirror Backpressure policies, that have strong transient and steady state performance guarantees. The MBP policies are simple and practical, and do not require any prior knowledge of demand arrival rates (which are permitted to vary in time), making them promising for applications. Policy design boils down to choosing suitable congestion functions.
(ii) **A framework for systematic design and analysis of MBP control policies.** Our control framework has a tight connection with mirror descent, which makes the process of policy design and analysis both systematic and flexible, and allows us to handle the challenging no-underflow constraint. The general machinery we develop can be seamlessly leveraged to design policies with provable guarantees for a variety of settings. This is in contrast with various intricate approaches in the queueing literature that do not easily generalize.

In Section 1.6 we generalize the baseline model (which allows entry control only) and include pricing and flexible assignment as control levers. We study joint entry-assignment control (JEA) in Section 1.6.2 and joint pricing-assignment control (JPA) in Section 1.6.3. Our control policies and performance guarantees extend seamlessly.

Applications. Our general model (Section 1.6) includes a number of key ingredients common to many applications. We illustrate its versatility by discussing the application to shared transportation systems (Section 1.7) and the application to scrip systems (Section 1.8). These applications and the relevant settings in the paper are summarized in Table 1.1.

Application	*Control lever*	*Corresponding setting in this paper*
Ride-hailing in USA, Europe	Pricing & Dispatch	Joint pricing-assignment
Ride-hailing in China	Admission & Dispatch	Joint entry-assignment
Bike sharing	Reward points	Pricing (finite buffer queues)
Scrip systems	Admission & Provider selection	Joint entry-assignment

Table 1.1: Summary of applications of our model, the control levers therein and the corresponding settings in this paper. See Section 1.6 for the joint entry-assignment and joint pricing-assignment settings (which allow for finite buffers). For each setting, we design MBP policies that are near optimal.

Shared transportation systems include ride-hailing and bike sharing systems. Here the nodes in our model correspond to geographical locations, while supply units and demand units correspond to vehicles and customers, respectively. Bike sharing systems dynamically incentivize certain trips using point systems to minimize out-of-bike and out-of-dock events caused by demand imbalance. Our pricing setting is relevant for the design of a dynamic incentive program for bike sharing; in particular, it allows for a limited number of docks. Ride-hailing platforms make dynamic decisions to optimize their objectives (e.g., revenue, welfare, etc.). For ride-hailing, our pricing-assignment model is relevant in regions such as North America, and our entry-assignment control model is relevant in in regions where dynamic pricing is undesirable like in China. We perform realistic simulations of ride-hailing and find that our MBP policy, suitably adapted to account for positive travel times, performs well (Section 1.7.1 and Appendix A.4).

A scrip system is a nonmonetary trade economy where agents use scrips (tokens, coupons, artificial currency) to exchange services (because monetary transfer is undesirable or impractical), e.g., for babysitting or kidney exchange. A key challenge in these markets is the design of the admission-and-provider-selection rule: If an agent is running low on scrip balance, should they be allowed to request services? If yes, and if there are several possible providers for a trade, who should be selected as the service provider? In Section 1.8, we show that a natural model of a scrip system is a special case of our entry-assignment control setting, yielding a near optimal admission-and-provider-selection control rule.

1.1.1 Literature Review

MaxWeight/backpressure policy. Backpressure [also known as MaxWeight, see 4, 5] are well-studied dynamic control policies in constrained queueing networks for workload minimization [16, 17], queue length minimization [18] and utility maximization [14], etc. Attractive features of MaxWeight/backpressure policies include their simplicity and provably good performance, and that arrival/service rate information is not required beforehand. The main challenge in using backpressure is the no-underflow constraints, as described earlier. Most of this literature considers the open queueing networks setting, where packets/jobs enter and leave, and there is much less work on closed networks. An exception is a recent paper on assignment control of closed networks by Banerjee, Kanoria, and Qian [3], which shows the large deviations optimality of "scaled" MaxWeight policies. Importantly, in that paper the demand arrival rates are assumed to satisfy a strong near balance assumption ("complete resource pooling"), as a result of which it suffices to consider non-idling policies (i.e., a "greedy" policy with assignment control only). In the present paper, in contrast, we allow very general demand arrival rates, which makes it necessary to deploy idling policies (e.g., entry control, pricing) to achieve good performance. Indeed, already under our illustrative model (Section 1.2) the CRP assumption of Banerjee, Kanoria, and Qian [3] is automatically violated and the greedy policy fails to achieve asymptotic optimality; see Remark 1.1 in Section 1.2.

While previous works use queue lengths or their power as congestion costs [16], our MBP policies significantly generalize backpressure by allowing a general increasing function (e.g., the logarithm) of queue lengths as congestion costs. As with backpressure, MBP policies carry provable guarantees.

Mirror Descent. Mirror descent (MD) is a generalization of the gradient descent algorithm for optimization, which was proposed by [12], see also [13]. MD is much more flexible than gradient descent as one can freely choose a "mirror map" that captures the geometry of the problem (including its objective and its constraints). Recently, there have been several works that use MD to solve online decision-making problems [e.g., 19]. Notably, [20] uses MD to obtain an improved approximation factor for a worst-case version of the so-called "k-server problem"; the k-server problem bears a certain resemblance to our setting in that the controller needs to manage the spatial distribution of supply. A key difference between our work and the existing works is that our proposed simple control policies remarkably have the *property* that they induce the queue lengths to follow MD dynamics, whereas the existing works actively run MD to solve their algorithmic problems.

Applications: shared transportation, scrip systems. Most of the ride-hailing literature studied controls that require the exact knowledge of system parameters: Özkan and Ward [21] studied payoff maximizing assignment control in an open queueing network model, Braverman, Dai, Liu, and Ying [2] derived the optimal state independent routing policy that sends empty vehicles to under-supplied locations, Banerjee, Freund, and Lykouris [1] adopted the Gordon-Newell closed queueing network model and considered various controls that maximize throughput, welfare or revenue. Balseiro, Brown, and Chen [22] considered a dynamic programming based approach for dynamic pricing for a specific network of star structure. (Ma, Fang, and Parkes [23] studied the somewhat different issue of ensuring that drivers have the incentive to accept dispatches by setting prices which are sufficiently smooth in space and time, in a model with no demand stochasticity.) Banerjee, Kanoria, and Qian [3] which assumes a near balance condition on demands and equal pickup costs may be the only paper in this space that does not require knowledge of system parameters. Comparing with Banerjee, Freund, and Lykouris [1] which obtains a steady state optimality gap of $O(\frac{1}{K})$ (in the absence of travel times) assuming *perfect* knowledge of demand arrival rates which are assumed to be *stationary*, our control policy achieves the same steady state optimality gap with *no* knowledge of demand arrival rates, and further achieves a *transient* optimality gap under *time-varying* demand arrival rates of $O(\frac{K}{T} + \frac{1}{K} + \sqrt{\eta K})$ for a finite number of arrivals T and changes of up to η per period (i.e., per arrival) in demand arrival rates. Some of these papers are able to formally handle travel delays: Braverman, Dai, Liu, and Ying [2], Banerjee, Freund, and Lykouris [1], and Banerjee, Kanoria, and Qian [3] prove theoretical results for the setting with i.i.d. geometric/exponential travel delays; Ma, Fang, and Parkes [23] consider deterministic travel delays. On the other hand, Balseiro, Brown, and Chen [22] ignores travel delays in their theory and later heuristically adapt their policy to accommodate travel delay (the present paper follows a similar approach). On the other hand, [21] is the only paper among these which (like the present paper) allows time-varying demand.

Our model can be applied to the design of dynamic incentive programs for bike sharing systems [24] and service provider rules for scrip systems [25, 26]. For example, the "minimum scrip selection rule" proposed in [25] is a special case of our policy, and our methodology leads to control rules in much more general settings as described in Section 1.8.

Other related work. A related stream of research studies online stochastic bipartite matching, see, e.g., [27, 28, 29, 30]; the main difference between their setting and ours is that we study a *closed* system where supply units never enter or leave the system. Network revenue management is a classical set of (open network) dynamic resource allocation problems, e.g., see [31, 32], and recent works, e.g., [33]. [34, 35, 36] and others study how process flexibility can facilitate improved performance, analogous to our use of assignment control to maximize payoff (when all pickup costs are equal), but the focus there is more on network design than on control policies. Again, this is an open network setting in that each supply unit can be used only once.

1.1.2 Organization of the Paper

The remainder of our paper is organized as follows. From Section 1.2 to Section 1.5 we focus on the entry control setting as an illustrative example of our approach: Section 1.2 presents our model and the platform objective. Section 1.3 introduces the Mirror Backpressure policy and presents our main theoretical result, i.e., a performance guarantee for the MBP policies. Section 1.4 introduces the static planning problem and describes the connection between the MBP policies and mirror descent. Section 1.5 outlines the proof of our main result. In Section 1.6, we provide MBP policies for joint entry-assignment and joint pricing-assignment control settings and allow for time-varying demand arrival rates, demonstrating the versatility of our approach. In Sections 1.7 and 1.8 we discuss the applications to shared transportation systems and scrip systems, respectively.

Notation. All vectors are column vectors if not specified otherwise. The transpose of vector or matrix $\mathbf{x}$ is denoted as $\mathbf{x}^{\mathrm{T}}$. We use $\mathbf{e}_i$ to denote the i-th unit column vector with the i-th coordinate being 1 and all other coordinates being 0, and $\mathbf{1}$ ($\mathbf{0}$) to denote the all 1 (0) column vector, where the dimension of the vector will be indicated in the superscript when it is not clear from the context, e.g., $\mathbf{e}_i^n$.

1.2 Illustrative Model: Dynamic Entry Control

In this section, we formally define our model of dynamic entry control in closed queueing networks. We will use this model as an illustrative example of our methodology.

We consider a finite-state Markov chain model with slotted time $t = 0, 1, 2, \ldots$, where a fixed number (denoted by K) of identical *supply units* circulate among a set of *nodes* V (locations), with $m \triangleq |V| > 1$. In our model, t will capture the number of demand units (customers) who have arrived so far (minus 1).

Queues (system state). At each node $j \in V$, there is an infinite-buffer queue of supply units. (Section 1.6.1 shows how to seamlessly incorporate finite-buffer queues.) The *system state* is the vector of queue lengths at time t, which we denote by $\mathbf{q}[t] = [q_1[t], \cdots, q_m[t]]^{\mathrm{T}}$. Denote the state space of queue lengths by $\Omega_K \triangleq \{\mathbf{q} : \mathbf{q} \in \mathbb{Z}_+^m, \mathbf{1}^{\mathrm{T}}\mathbf{q} = K\}$, and the normalized state space by $\Omega \triangleq \{\mathbf{q} : \mathbf{q} \in \mathbb{R}_+^m, \mathbf{1}^{\mathrm{T}}\mathbf{q} = 1\}$.

Demand Types and Arrival Process. We assume exactly one demand unit (customer) arrives at each period t, and denote her type by $(o[t], d[t]) \in V \times V$, where $o[t]$ is her origin and $d[t]$ is her destination. With probability ϕ_{jk}, we have $(o[t], d[t]) = (j, k)$, independent of demands in earlier periods.[5] Let $\boldsymbol{\phi} \triangleq (\phi_{jk})_{j \in V, k \in V}$. Importantly, the system can observe the type of the arriving demand at the beginning of each time slot, but *the probabilities (arrival rates)* $\boldsymbol{\phi}$ *are not known.* Thus we substantially relax the assumption in previous works that the system has exact knowledge of demand arrival rates ([21], [1], [22]).

Entry Control and Payoff. At time t, after observing the demand type $(o[t], d[t]) = (j, k)$, the system makes a binary decision $x_{jk}[t] \in \{0, 1\}$ where $x_{jk}[t] = 1$ stands for serving the demand, $x_{jk}[t] = 0$ means rejecting the demand. A supply unit moves and payoff is collected (or not) accordingly as follows:

- If $x_{jk}[t] = 1$, then a supply unit relocates from j to k, immediately. Meanwhile, the platform collects payoff $v[t] = w_{jk}$ in this period. Without loss of generality, let $\max_{j,k \in V} |w_{jk}| = 1$.
- If $x_{jk}[t] = 0$, then supply units remain where they are and $v[t] = 0$.

Because the queue lengths are non-negative by definition, we require the following *no-underflow constraint* to be met at any t:

$$x_{jk}[t] = 0 \quad \text{if} \quad q_j[t] = 0 \,. \tag{1.1}$$

As a convention, we let $x_{j'k'}[t] = 0$ if $(o[t], d[t]) \neq (j', k')$. A *feasible policy* specifies, for each time $t \in \{0, 1, 2, \dots\}$, a mapping from the history so far of demand types

[5]This is equivalent to considering a continuous time model where the arrivals of different types of demands follow independent Poisson processes with rates proportional to the (ϕ_{jk})s. The discrete time model considered is the embedded chain of the continuous time model.

$\big(o[t'], d[t']\big)_{t' \le t}$ and states $(\mathbf{q}[t'])_{t' \le t}$ to a decision $x_{jk}[t] \in \{0,1\}$ satisfying (1.1), where $(j,k) = (o[t], d[t])$ as above. We allow $x_{jk}[t]$ to be randomized, although our proposed policies will be deterministic. The set of feasible policies is denoted by $\mathcal{U}$.

System Dynamics and Objective. The dynamics of system state $\mathbf{q}[t] \in \Omega_K$ is as follows:

$$\mathbf{q}[t+1] = \mathbf{q}[t] + x_{jk}[t](-\mathbf{e}_j + \mathbf{e}_k)\,. \tag{1.2}$$

We use $v^\pi[t]$ to denote the payoff collected at time t under control policy π. Let W_T^π denote the average payoff per period (i.e., per customer) collected by policy π in the first T periods, and let W_T^* denote the optimal payoff per period in the first T periods over all admissible policies. Mathematically, they are defined respectively as:

$$W_T^\pi \triangleq \min_{\mathbf{q}\in\Omega_K} \frac{1}{T}\sum_{t=0}^{T-1}\mathbb{E}[v^\pi[t]|\mathbf{q}[0]=\mathbf{q}]\,, \qquad W_T^* \triangleq \sup_{\pi\in\mathcal{U}}\max_{\mathbf{q}\in\Omega_K} \frac{1}{T}\sum_{t=0}^{T-1}\mathbb{E}[v^\pi[t]|\mathbf{q}[0]=\mathbf{q}]\,. \tag{1.3}$$

Define the infinite-horizon per period payoff W^π collected by policy π and the optimal per period payoff over all admissible policies W^* respectively as:

$$W^\pi \triangleq \liminf_{T\to\infty} W_T^\pi\,, \qquad W^* \triangleq \limsup_{T\to\infty} W_T^*\,. \tag{1.4}$$

We measure the performance of a control policy π by its per-customer *optimality gap* ("loss"):

$$L_T^\pi = W_T^* - W_T^\pi \qquad \text{and} \qquad L^\pi = W^* - W^\pi\,. \tag{1.5}$$

Note that we consider the worst-case initial system state when evaluating a given policy, and the best initial state for the optimal benchmark; see (1.3). Such a definition of optimality gap provides a conservative bound on policy performance and avoids the (unilluminating) discussion of the dependence of performance on initial state.

We make the following mild connectivity assumption on the demand arrival rates $\boldsymbol{\phi}$.

Condition 1.1 (Strong Connectivity of $\boldsymbol{\phi}$). *Define the* connectedness *of* $\boldsymbol{\phi}$ *as*

$$\alpha(\boldsymbol{\phi}) \triangleq \min_{S \subsetneq V, S \neq \emptyset} \sum_{j \in S, k \in V \setminus S} \phi_{jk} \,. \tag{1.6}$$

We assume that $\boldsymbol{\phi}$ *is* strongly connected, *namely, that* $\alpha(\boldsymbol{\phi}) > 0$.

Note that Condition 1.1 is equivalent to requiring that for every ordered pair of nodes (j,k), there is a sequence of demand types with positive arrival rate that would take a supply unit from j eventually to k.

We conclude this section with an example which shows that the greedy policy typically has steady state optimality gap $\Omega(1)$ per period, followed by the observation that the main assumption of Banerjee, Kanoria, and Qian [3] is automatically violated in our setting.

Example 1.1 (Greedy policy typically incurs $\Omega(1)$ loss). *Consider a network with three nodes* $V = \{1,2,3\}$, *demand arrival probabilities* $\phi_{12} = \epsilon$, $\phi_{23} = \frac{1}{3} + \epsilon$, $\phi_{21} = \phi_{32} = \frac{1}{3} - \epsilon$ *(where* $0 < \epsilon < \frac{1}{6}$*), and payoffs* $w_{23} = w > 0$, $w_{12} = w_{21} = w_{32} = \frac{w}{2}$. *Let* $\mathbf{x}^*$ *be the optimal solution to the SPP* (1.10)-(1.12). *By inspection,* $\mathbf{x}^*$ *should induce the maximum circulation in each of the two cycles* 1—2—1 *and* 2—3—2, *hence* $x^*_{12} = x^*_{32} = 1$, $x^*_{21} = \frac{\epsilon}{\frac{1}{3}-\epsilon}$, $x^*_{23} = \frac{\frac{1}{3}-\epsilon}{\frac{1}{3}+\epsilon}$. *We know that there exists a policy whose performance approaches the value of the SPP as* $K \to \infty$ *[1]. We will prove by contradiction that the greedy policy incurs an* $\Omega(1)$ *loss for this example, by showing that its payoff per period is* $\Omega(1)$ *below the value of the SPP. Consider the steady state under the greedy policy. Suppose the loss is vanishing, i.e., all but an* $o(1)$ *fraction of type* $(1,2)$ *and type* $(3,2)$ *demand are served. Suppose a* γ *fraction of the time there is a supply unit present at node* 2. *As a result, since the greedy policy is being used, a* γ *fraction of demands of type* $(2,1)$ *are served, and a* γ *fraction of demands of type* $(2,3)$ *are served. Flow-balance at nodes* 1 *and* 3, *respectively, implies that we have* $(\frac{1}{3} - \epsilon)\gamma = \epsilon - o(1)$, $(\frac{1}{3} + \epsilon)\gamma = \frac{1}{3} - \epsilon - o(1)$. *However, these two equations cannot both be satisfied as* $K \to \infty$ *unless* $\epsilon = \frac{1}{9}$. *We infer that the greedy policy incurs an* $\Omega(1)$ *loss in this network for any* $\epsilon \in (0, \frac{1}{6})$, $\epsilon \neq \frac{1}{9}$.

Remark 1.1. *The complete resource pooling (CRP) condition imposed in Banerjee,* *Kanoria, and Qian [3, Assumption 3] is automatically violated in the model we have defined in this section. Consider our setup including Condition 1.1. The CRP condition can be stated as follows: for each subset of nodes $S \subsetneq V, S \neq \emptyset$, the "net demand" $\mu_S \triangleq \sum_{i\in S}\sum_{j\in V\setminus S}\phi_{ij}$ is less than the "net supply" $\lambda_S \triangleq \sum_{j\in V\setminus S}\sum_{i\in S}\phi_{ji}$, i.e., $\mu_S < \lambda_S$. Clearly, any demand arrival rates $\boldsymbol{\phi}$ violate CRP, since if $\mu_S < \lambda_S$ for some $S \subsetneq V, S \neq \emptyset$ then this means that $\mu_{V\setminus S} > \lambda_{V\setminus S}$ (given that $\mu_{V\setminus S} = \lambda_S$ and $\lambda_{V\setminus S} = \mu_S$ by definition), i.e., CRP is violated. In Example 1.1, the subset $\{2\}$ (and the subset $\{1,2\}$) violates this constraint.*

1.3 The MBP Policies and Main Result

In this section, we propose a family of blind online control policies, and state our main result for these policies, which provides a strong transient and steady state performance guarantee for finite systems.

1.3.1 The Mirror Backpressure Policies

We propose a family of online control policies which we call *Mirror Backpressure* (MBP) policies. Each member of the MBP family is specified by a mapping of normalized queue lengths $\mathbf{f}(\bar{\mathbf{q}}) : \Omega \to \mathbb{R}^m$, where $\mathbf{f}(\bar{\mathbf{q}}) \triangleq [f(\bar{q}_1), \cdots, f(\bar{q}_m)]^{\mathrm{T}}$ and f is a monotone increasing function.[6] We will refer to $f(\cdot)$ as the *congestion function*, which maps each (normalized) queue length to a congestion cost at that node, based on which MBP will make its decisions. (We will defined normalized queue lengths $\bar{\mathbf{q}}$ below.)

We will later clarify the precise role of the congestion function: we will show that MBP executes dual stochastic mirror descent [13] on the fluid limit problem with mirror map equal to the inverse of the congestion function. Similar to the design of effective mirror descent algorithms, the choice of congestion function should depend on the constraints

[6]The methodology we will propose will seamlessly accommodate general mappings $\mathbf{f}(\cdot)$ such that $\mathbf{f} = \nabla F$ where $F(\cdot) : \Omega \to \mathbb{R}$ is a strongly convex function, a special case of which is $\mathbf{f}(\bar{\mathbf{q}}) \triangleq [f_1(\bar{q}_1), \cdots, f_m(\bar{q}_m)]^{\mathrm{T}}$ for some monotone increasing (f_j)s. Here it suffices to consider a single congestion function $f(\cdot)$, whereas in Section 1.6 we will employ queue-specific congestion functions $f_j(\cdot)$.

of the setting, leading to an interesting interplay between problem geometry and policy design.

For conciseness, in this section we will state our main result for the congestion function

$$f(\bar{q}_j) \triangleq -\sqrt{m} \cdot \bar{q}_j^{-\frac{1}{2}}, \tag{1.7}$$

and postpone the results for other choices of congestion functions to Appendix A.3 (see also Remark 1.2). For technical reasons, we need to keep $\bar{\mathbf{q}}$ in the *interior* of the normalized state space Ω, i.e., we need to ensure that all normalized queue lengths remain positive. This is achieved by defining the normalized queue lengths $\bar{\mathbf{q}}$ as

$$\bar{q}_i \triangleq \frac{q_i + \delta_K}{\tilde{K}} \quad \text{for} \quad \delta_K \triangleq \sqrt{K} \quad \text{and} \quad \tilde{K} \triangleq K + m\delta_K\,. \tag{1.8}$$

Note that this definition leads to $\mathbf{1}^{\mathrm{T}}\bar{\mathbf{q}} = 1$ and therefore $\bar{\mathbf{q}} \in \Omega$.

Our proposed MBP policy for the entry control problem is given in Algorithm 1. MBP admits a demand of type (j, k) if and only if the *score*

$$w_{jk} + f(\bar{q}_j) - f(\bar{q}_k) \tag{1.9}$$

is nonnegative *and* the origin node j has at least one supply unit (see Figure 1.1 for illustration of the score). The score (1.9) is nonnegative if and only if the payoff w_{jk} of serving the demand outweighs the difference of congestion costs (given by $f(\bar{q}_k)$ and $f(\bar{q}_j)$) between the demand's destination k and origin j. Roughly speaking, MBP is more willing to take a supply unit from a long queue and add it to a short queue, than vice versa; see Figures 1.1 and 1.2. The policy is not only completely blind, but also semi-local, i.e., it only uses the queue lengths at the origin and destination. Note that the congestion cost (1.7) increases with queue length (as required), and furthermore decreases sharply as queue length approaches zero. Observe that such a choice of congestion function makes MBP very reluctant to take supply units from short queues and helps to enforce the no-underflow constraint (1.1). See Section 1.4.3 for detailed discussion on the no-underflow constraint.

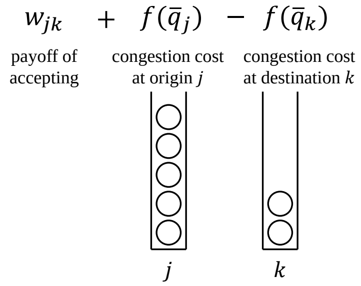

Figure 1.1: The score (1.9); MBP admits a demand unit only if the score is non-nonnegative,

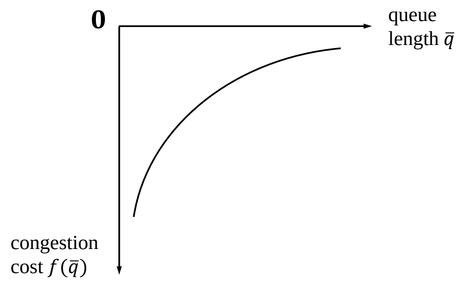

Figure 1.2: An example of a congestion function (a mapping from queue lengths to congestion costs) which aggressively protects supply units in near-empty queues.

ALGORITHM 1: Mirror Backpressure (MBP) Policy for Entry Control

---

At the start of period t, the platform observes $(o[t], d[t]) = (j, k)$.

if $w_{jk} + f(\bar{q}_j[t]) - f(\bar{q}_k[t]) \geq 0$ **and** $q_j[t] > 0$ **then**

  $x_{jk}[t] \leftarrow 1$, i.e., serve the incoming demand;

else

  $x_{jk}[t] \leftarrow 0$, i.e., drop the incoming demand;

end

The queue lengths update as $\bar{\mathbf{q}}[t+1] = \bar{\mathbf{q}}[t] - \frac{1}{K} x_{jk}[t](\mathbf{e}_j - \mathbf{e}_k)$.

1.3.2 Performance Guarantee for MBP Policies

We now formally state the main performance guarantee of our paper for the dynamic entry control model introduced in Section 1.2. We will outline the proof in Section 1.5, and extend the result to more general settings in Section 1.6.

Theorem 1.1. *Consider a set of m nodes and any demand arrival rates $\boldsymbol{\phi}$ that satisfy Condition 1.1. Then there exists $K_1 = \text{poly}(m, \frac{1}{\alpha(\boldsymbol{\phi})})$, and a universal constant $C < \infty$, such that the following holds.*[7] *For the congestion function $f(\cdot)$ defined in* (1.7)*, for any*

[7]Here "poly" indicates a polynomial. The constant C is universal in the sense that it does not depend on K, m or $\alpha(\boldsymbol{\phi})$.

$K \geq K_1$, the following guarantees hold for Algorithm 1

$$L_T^{\mathrm{MBP}} \leq M_1 \cdot \frac{K}{T} + M_2 \cdot \frac{1}{K}\,, \qquad \text{and} \qquad L^{\mathrm{MBP}} \leq M_2 \cdot \frac{1}{K}\,, \qquad \text{for } M_1 \triangleq Cm \text{ and } M_2 \triangleq Cm^2\,.$$

Remark 1.2. *In Section 1.6 we obtain results similar to Theorem 1.1 in broader settings that allow pricing and flexible assignment (Theorem 1.2, 1.3), and moreover allow for time-varying demand arrival rates in Section 1.6.2. In Appendix A.3 (Theorem A.1), we generalize Theorem 1.1 by showing similar performance guarantees for a whole class of congestion functions that satisfy certain growth conditions. Informally, the congestion function needs to be steep enough near zero to protect the nodes from being drained of supply units. For example, for both the logarithmic congestion function, i.e.* $f(\bar{q}) = c \cdot \log(\bar{q})$*, and the linear congestion function, i.e.* $f(\bar{q}) = c \cdot \bar{q}$ *with* $c > c_0$ *for some* $c_0 = \mathrm{poly}(m, \frac{1}{\alpha(\boldsymbol{\phi})})$*, the same guarantee as in Theorem 1.1 holds with* $K_1 = \mathrm{poly}(c, m, \frac{1}{\alpha(\boldsymbol{\phi})})$*,* $M_1 = \mathrm{poly}(c, m)$*,* $M_2 = \mathrm{poly}(c, m)$*. However, the specific polynomials depend on the choice of congestion function.*

There are several attractive features of the performance guarantee provided by Theorem 1.1 for the simple and practically attractive Mirror Backpressure policy:

(1) **The policy is completely blind.** In practice, the platform operator at best has access to an imperfect estimate of the demand arrival rates $\boldsymbol{\phi}$, so it is a very attractive feature of the policy that it does not need any estimate of $\boldsymbol{\phi}$ whatsoever. It is worth noting that the consequent bound of $O\left(\frac{1}{K}\right)$ on the steady state optimality gap remarkably matches that provided by Banerjee, Freund, and Lykouris [1] even though MBP requires *no* knowledge of $\boldsymbol{\phi}$, whereas the policy of Banerjee, Freund, and Lykouris [1] requires *exact* knowledge of $\boldsymbol{\phi}$: As shown in Banerjee, Kanoria, and Qian [3, Proposition 4], if the estimate of demand arrival rates is imperfect, any state independent policy [such as that of 1] generically suffers a long run (steady state) per customer optimality gap of $\Omega(1)$ (as $K \to \infty$). Note that the greedy policy (which admits a demand whenever a supply unit is available) also typically suffers a steady state per period optimality gap of $\Omega(1)$; see Example 1.1 in Section 1.2.

(2) **Guarantee on transient performance.** In contrast with Banerjee, Freund, and Lykouris [1] which provides only a steady state bound for finite K, we are able to provide a performance guarantee for finite horizon and finite (large enough) K. The horizon-dependent term K/T in our bound on optimality gap is small if the total number of arrivals T is large compared to the number of supply units K.

It is worth noting that our bound *does not* deteriorate as the system size increases in the "large market regime", where the number of supply units K increases proportionally to the demand arrival rates [this regime is natural in ride-hailing settings, taking the trip duration to be of order 1 in physical time, and where a non-trivial fraction of cars are busy at any time, see, e.g., 2]. Let T^{real} denote the horizon in physical time. As K increases in the large market regime, the primitive $\boldsymbol{\phi}$ remains unchanged, while $T = \Theta(K \cdot T^{\text{real}})$ since there are $\Theta(K)$ arrivals per unit of physical time. Hence, we can rewrite our performance guarantee as

$$W_T^* - W_T^{\text{MBP}} \leq M \left(\frac{1}{T^{\text{real}}} + \frac{1}{K} \right) \xrightarrow{K \to \infty} \frac{M}{T^{\text{real}}} .$$

Our bound on the optimality gap per customer in steady state is M_2/K, matching that of Banerjee, Freund, and Lykouris [1] in its scaling with K. (However, our constant M_2 is quadratic in the number of nodes m, whereas the constant in the other paper is linear in m.)

(3) **Flexibility in the choice of congestion function.** Because of the richness of the class of congestion functions covered in Appendix A.3 which generalizes Theorem 1.1, the system controller now has the additional flexibility to choose a suitable congestion function $f(\cdot)$. For example, in our setting the performance guarantee for the congestion function given in (1.7) (Theorem 1.1) is more attractive than that for the linear congestion function $f(\bar{q}) = c \cdot \bar{q}$ (Remark 1.2) in the following way: in the latter case the coefficient c needs to be larger than a threshold that depends on connectedness $\alpha(\boldsymbol{\phi})$ for a non-trivial performance guarantee to hold. (Thus, in order to choose c the platform needs to know $\alpha(\boldsymbol{\phi})$, whereas no knowledge of $\alpha(\boldsymbol{\phi})$ is needed when using the congestion function (1.7).)

From a practical perspective, this flexibility can allow significant performance gains to be unlocked by making an appropriate choice of $f(\cdot)$, as evidenced by our numerical experiments in Section 1.7.1 and Appendix A.4.

1.4 The MBP Policies and Mirror Descent

In this section, we describe the main intuition behind the success of MBP policies, namely, that they execute (dual) mirror descent on a certain deterministic optimization problem. In Section 1.4.1, we define the deterministic optimization problem which arises in the continuum limit: the *static planning problem* (SPP), whose value we use to upper bound the optimal finite (and infinite) horizon per period W_T^* (and W^*) defined in (1.3) and (1.4). In Section 1.4.2, we first review the interpretation of the celebrated Backpressure (BP) policy as a stochastic gradient descent algorithm on the dual of the SPP, and then proceed to generalize the argument to informally show that MBP executes mirror descent on the dual of SPP. In Section 1.4.3 we discuss the main challenge in turning the intuition into a proof, namely, the no-underflow constraint.

1.4.1 The Static Planning Problem

We first introduce a linear program (LP) that will be used to upper bound W_T^* and W^*. The LP, called the static planning problem (SPP) [see, e.g., 15, 7], is:

$$\text{maximize}_{\mathbf{x}} \sum_{j,k\in V} w_{jk} \cdot \phi_{jk} \cdot x_{jk} \tag{1.10}$$

$$\text{s.t.} \sum_{j,k\in V} \phi_{jk} \cdot x_{jk}(\mathbf{e}_j - \mathbf{e}_k) = \mathbf{0} \qquad \text{(flow balance)} \tag{1.11}$$

$$x_{jk} \in [0,1] \qquad \forall j,k \in V\,. \qquad \text{(demand constraint)} \tag{1.12}$$

One interprets x_{jk} as the fraction of type (j,k) demand which is accepted, and the objective (1.10) as the rate at which payoff is generated under the fractions $\mathbf{x}$. In the SPP (1.10)-(1.12), one maximizes the rate of payoff generation subject to the requirement that the average inflow of supply units to each node in V must equal the outflow (constraint (1.11)), and that $\mathbf{x}$ are indeed fractions (constraint (1.12)). Let W^{SPP} be the optimal value of SPP. The following proposition formalizes that, as is typical in such settings, W^{SPP} is an upper bound on the optimal steady state (per customer) payoff W^*. It further establishes that the optimal finite horizon per customer payoff W_T^* cannot be much larger than W^{SPP}.

Proposition 1.1. *For any horizon $T < \infty$ and any K, the finite and infinite horizon average payoff W_T^* and W^* are upper bounded as*

$$W_T^* \le W^{\text{SPP}} + m \cdot \frac{K}{T}, \qquad W^* \le W^{\text{SPP}}. \tag{1.13}$$

We obtain the finite horizon upper bound to W_T^* in (1.13) by slightly relaxing the flow constraint (1.11) to accommodate the fact that flow balance need not be exactly satisfied over a finite horizon.

1.4.2 MBP Executes Dual Stochastic Mirror Descent on SPP

The BP policy and our proposed MBP policies are closely related to the (partial) dual of the SPP:

$$\begin{aligned} &\text{minimize}_{\mathbf{y}}\ g(\mathbf{y}), \\ &\text{where } g(\mathbf{y}) \triangleq \sum_{j,k\in V} \phi_{jk} \cdot \max_{x_{jk}\in[0,1]} x_{jk}(w_{jk} + y_j - y_k) = \sum_{j,k\in V} \phi_{jk} \cdot \left(w_{jk} + y_j - y_k\right)^+, \end{aligned} \tag{1.14}$$

where $(x)^+ \triangleq \max\{0, x\}$. Here $\mathbf{y}$ are the dual variables corresponding to the flow balance constraints (1.11), and have the interpretation of "congestion costs" [37], i.e., y_j can be thought of as the "cost" of having one extra supply unit at node j.

In the rest of this subsection, we informally describe the interpretation of BP as stochastic gradient descent, and the interpretation of MBP as stochastic mirror descent, on problem (1.14).

Review of the interpretation of BP as dual stochastic subgradient descent. Rich dividends have been obtained by treating the (properly scaled) current queue lengths $\mathbf{q}$ as the dual variables $\mathbf{y}$, resulting in the celebrated backpressure (BP, also known as MaxWeight) control policy, introduced by Tassiulas and Ephremides [4], see also, e.g., [8, 14]. Formally, BP sets the current value of $\mathbf{y}$ to be proportional to the current normalized queue lengths, i.e., $\mathbf{y}[t] = c \cdot \bar{\mathbf{q}}[t]$ for some $\bar{\mathbf{q}} \in \Omega$ defined, e.g., as in (1.8), and some $c > 0$ and greedily maximizes the inner problem in (1.14) for every origin j and destination k, i.e.,

$$x_{jk}^{\mathrm{BP}}[t] = \begin{cases} 1 & \text{if } w_{jk} + c \cdot \bar{q}_j[t] - c \cdot \bar{q}_k[t] \geq 0 \text{ and } q_j[t] > 0\,, \\ 0 & \text{otherwise}\,. \end{cases} \tag{1.15}$$

The main attractive feature of this policy is that it is extremely simple and does not need to know demand arrival rates $\boldsymbol{\phi}$. The BP policy can be viewed as a *stochastic subgradient descent (SGD)* algorithm on the dual problem (1.14), when the current state is in the *interior* of the state space, i.e., when $q_j > 0$ for all $j \in V$ [38]. To see this, denote the subdifferential (set of subgradients) of function $g(\cdot)$ at $\mathbf{y}$ as $\partial g(\mathbf{y})$. Observe that *the expected change of queue lengths under BP is proportional to the negative of a subgradient of $g(\cdot)$ at $\mathbf{y} = c \cdot \bar{\mathbf{q}}[t]$*, in particular

$$-\frac{\tilde{K}}{c} \cdot \mathbb{E}[\mathbf{y}[t+1] - \mathbf{y}[t]] = -\mathbb{E}[\mathbf{q}[t+1] - \mathbf{q}[t]] = \sum_{j,k \in V} \phi_{jk} \cdot x_{jk}^{\mathrm{BP}}[t](\mathbf{e}_j - \mathbf{e}_k) \ \in \ \partial g(\mathbf{y}[t])\,, \tag{1.16}$$

where the first equality follows from the definition $\mathbf{y}[t] = c \cdot \bar{\mathbf{q}}[t]$ (and the definition of normalized queue length (1.8)) and second equality is just the expectation of the system dynamics (1.2). Here $\sum_{j,k \in V} \phi_{jk} \cdot x_{jk}^{\mathrm{BP}}[t](\mathbf{e}_j - \mathbf{e}_k) \in \partial g(\mathbf{y}[t])$ since g is a maximum of linear functions of $\mathbf{y}$ parameterized by $\mathbf{x}$, hence g is convex and the gradient of a linear function among these which is an argmax at $\mathbf{y}[t]$ (in particular, the linear function parameterized by $\mathbf{x}^{\mathrm{BP}}[t]$) is a subgradient of g at $\mathbf{y}[t]$.

Eq. (1.16) shows that the evolution of $\mathbf{y}[t]$ when $\mathbf{q}[t] > 0$ is exactly an iteration of SGD with step size $\frac{c}{\tilde{K}}$. This interpretation of BP as stochastic subgradient descent leads to desirable properties including stability, approximate minimization of delay/workload, and approximate revenue maximization in certain networks [see, e.g., 5, 14, etc.]. However, as we will see in Section 1.4.3, in our setting the SGD property of backpressure breaks on the *boundary* of state space, i.e., when there exists $j' \in V$ such that $q_{j'} = 0$, due to the *no-underflow* constraints $\mathbf{q} \geq \mathbf{0}$.

MBP executes dual stochastic mirror descent on the SPP. The key innovation of our approach is to design a family of policies generalizing BP (MBP given in Algorithm 1) that *executes stochastic mirror descent on the partial dual problem* (1.14) *(with flow constraints dualized), with* $\bar{\mathbf{q}}[t]$ *given by* (1.8) *being the mirror point and the inverse mirror map being the (vector) congestion function* $\mathbf{f}(\bar{\mathbf{q}}) \triangleq [f(\bar{q}_1), \cdots, f(\bar{q}_m)]^{\mathrm{T}}$*.* Mathematically, if $\mathbf{q} > 0$, we have

$$-\tilde{K} \cdot \mathbb{E}[\bar{\mathbf{q}}[t+1] - \bar{\mathbf{q}}[t]] = -\mathbb{E}[\mathbf{q}[t+1] - \mathbf{q}[t]] = \sum_{j,k \in V} \phi_{jk} \cdot x_{jk}^{\mathrm{MBP}}[t](\mathbf{e}_j - \mathbf{e}_k) \in \partial g(\mathbf{y})\Big|_{\mathbf{y}=\mathbf{f}(\bar{\mathbf{q}}[t])}, \tag{1.17}$$

where $\mathbf{x}^{\mathrm{MBP}}[t]$ is the control defined in Algorithm 1; notice that the entry rule $\mathbf{x}^{\mathrm{MBP}}[t]$ has the same form as that for BP (1.15) except that it uses a general congestion function $f(\bar{q}_j)$, leading to (1.17) for MBP via the same reasoning that led to (1.16) for BP. Thus, MBP performs stochastic mirror descent on the partial dual problem (1.14), which generalizes the previously known fact that BP performs stochastic gradient descent.

A main advantage of mirror descent over gradient descent is that it can better capture the geometry of the state space via an appropriate choice of mirror map [see, e.g., 12, 13]. In our setting, the congestion function $\mathbf{f}(\bar{\mathbf{q}})$ is the inverse mirror map and can be flexibly chosen.

Our approach blending backpressure and mirror descent with a flexibly chosen mirror map is novel. We believe it can serve as a general framework for systematic design of provably near optimal backpressure-like control policies for queueing networks in settings with hairy practical constraints.

1.4.3 Challenge: No-underflow Constraints

As we have discussed earlier, the no-underflow constraints pose a challenge when applying backpressure to various settings. The following simple example illustrates how BP fails when the proportionality constant c is not chosen to be sufficiently large.

Example 1.2 (BP is far from optimal if c is not large enough)**.** *Consider the network introduced in Example 1.1. Suppose the platform employs backpressure where the shadow prices are taken to be proportional to (normalized) queue lengths* $\mathbf{y}[t] = c \cdot \bar{\mathbf{q}}[t]$ *with* $c < \frac{3}{2}w$.

Let $\mathbf{y}^*$ *be the optimal dual variables in* (1.14)*. By complementary slackness we have that the set of dual optima are* $\mathbf{y}^*$ *which satisfy*

$$\frac{w}{2} + y_1^* - y_2^* \geq 0\,, \quad \frac{w}{2} + y_2^* - y_1^* = 0\,, \quad w + y_2^* - y_3^* = 0\,, \quad \frac{w}{2} + y_3^* - y_2^* \geq 0\,.$$

Hence $\mathbf{y}^*$ *takes the form* $\mathbf{y}^* = (y_1^*, y_1^* - \frac{w}{2}, y_1^* + \frac{w}{2})$ *for arbitrary* $y_1^* \in \mathbb{R}$*. Let* $\bar{\mathbf{q}}^* \triangleq \mathbf{y}^*/c$ *be the queue lengths corresponding to the optimal dual variables in* (1.14) *with the additional constraint that the normalized queue lengths sum to* 1*. Simple algebra yields* $\bar{\mathbf{q}}^* = (\frac{1}{3}, \frac{2c-3w}{6c}, \frac{2c+3w}{6c})$*. Because* $c < \frac{3}{2}w$ *we have* $\bar{q}_2^* < 0$*, and so* $\bar{\mathbf{q}}^*$ *lies outside the normalized state space* $\bar{\mathbf{q}}^* \notin \Omega$*. Hence, the* $\bar{\mathbf{q}}[t]$ *will never converge to* $\bar{\mathbf{q}}^*$ *and BP is far from optimal.*

Even if the platform uses BP with sufficiently large c to ensure that $\bar{\mathbf{q}}^* \in \Omega$, the existing analysis of BP still fails, as is demonstrated below.

Example 1.3 (BP has positive Lyapunov drift at a certain state)**.** *Again consider Example 1.1 and let* $c \geq \frac{3}{2}w$*. A typical analysis of BP is based on establishing that the "drift" defined by*

$$\mathbb{E}\Big[\, \|\bar{\mathbf{q}}[t+1] - \bar{\mathbf{q}}^*\|_2^2 \,\Big|\, \bar{\mathbf{q}}[t] \,\Big] - \|\bar{\mathbf{q}}[t] - \bar{\mathbf{q}}^*\|_2^2$$

is strictly negative when $\|\bar{\mathbf{q}}[t] - \bar{\mathbf{q}}^*\|_2 = \Omega(1)$*. Suppose at time* t *we have*[8] $\bar{\mathbf{q}}[t] = (\frac{2}{3}, 0, \frac{1}{3})$*;*

[8] The integrality of the components of $\mathbf{q}[t]$ is non-essential, hence we assume all components of $\mathbf{q}[t]$ are integers. Also, here we take the normalized queue lengths to be defined as $\bar{\mathbf{q}}[t] \triangleq \mathbf{q}[t]/K$ to simplify the expressions.

in particular, queue 2 is empty. Note that at $\bar{\mathbf{q}}[t]$*, BP can only fulfill the demand going from* 1 *to* 2 *and from* 3 *to* 2 *because of the no-underflow constraint. Straightforward calculation shows that the "drift" is positive for large enough* K *if* $\epsilon < \frac{w}{2c+3w}$.

In the following analysis, we show that the underflow problem is provably alleviated by MBP policies with an appropriately chosen congestion function. For example, the MBP policy with congestion function given in (1.7) is more aggressive in preserving supply units in near-empty queues compared to BP, making the system less likely to violate the no-underflow constraints. Besides carrying formal guarantees, the MBP policy also achieves better performance than BP in simulations (Section 1.7.1 and Appendix A.4).

1.5 Proof of Theorem 1.1

In this section we provide the key lemmas that lead to a proof of Theorem 1.1. Our analysis generalizes and refines the so-called Lyapunov drift method in the network control literature [see, e.g., 37]. It consists of three steps:

(1) In Section 1.5.1, we use Lyapunov analysis to upper bound the suboptimality that MBP incurs in one period by the sum of several auxiliary terms (Lemma 1.1). The auxiliary terms are easier to control and have clear interpretations.

(2) In Section 1.5.2, we utilize the structure of the dual problem (1.14) to bound the auxiliary terms introduced in the first step (Lemmas 1.2 and 1.3).

(3) In Section 1.5.3, we average the one-step optimality gap obtained in previous steps over a finite/infinite horizon, and conclude the proof of Theorem 1.1.

We use the antiderivative of $\mathbf{f}(\cdot)$ as our Lyapunov function; for the congestion function f in (1.7), this is

$$F(\bar{\mathbf{q}}) \triangleq -2\sqrt{m}\sum_{j\in V}\sqrt{\bar{q}_j}\,. \tag{1.18}$$

Motivation for our choice of Lyapunov function. We utilize our key observation that MBP executes mirror descent on the dual of SPP (see Section 1.4.2) to find a suitable (uncentered) Lyapunov function. The standard proof of convergence of mirror descent uses the Bregman divergence $B_F(\bar{\mathbf{q}}, \bar{\mathbf{q}}^*)$, generated by the antiderivative $F(\cdot)$ of the inverse mirror map, as the Lyapunov function (note that $B_F(\bar{\mathbf{q}}, \bar{\mathbf{q}}^*)$ is a "centered" function in that it achieves its minimum at $\bar{\mathbf{q}}^*$; this function generalizes the centered quadratic function used to analyze stochastic gradient descent). We use the "uncentered" version of the Bregman divergence, which is nothing but F itself, as our Lyapunov function; this choice turns out to be natural for studying the time-averaged performance (rather than convergence of the last iterate). Since the congestion function corresponds to the inverse mirror map, our F is simply the antiderivative of the congestion function.[9]

1.5.1 Single Period Analysis of MBP via Lyapunov Function

This part of the proof relies on the key observation we made in Section 1.4, i.e., that MBP policy executes stochastic mirror descent on the dual objective function $g(\mathbf{y})$ (the dual problem was defined in (1.14)) except when underflow happens. As a result, our analysis combines (a modification of) the standard approach for stochastic mirror descent algorithms [see, e.g., 12, 13] with a novel argument that bounds the suboptimality contributed by underflow.

Recall that W^{SPP} is the optimal value of SPP (1.10)-(1.12), $v^{\text{MBP}}[t]$ denotes the payoff collected under the MBP policy in the t-th period, and $g(\cdot)$ is the dual problem (1.14). We have the following result (proved in Appendix A.2):

Lemma 1.1 (Suboptimality of MBP in one period). *Consider congestion functions $f(\cdot)$s that are strictly increasing and continuously differentiable. We have the following decomposition:*

$$W^{\text{SPP}} - \mathbb{E}[v^{\text{MBP}}[t] | \bar{\mathbf{q}}[t]] \leq \underbrace{\tilde{K}\left(F(\bar{\mathbf{q}}[t]) - \mathbb{E}[F(\bar{\mathbf{q}}[t+1]) | \bar{\mathbf{q}}[t]]\right)}_{\mathcal{V}_1} + \underbrace{\frac{1}{2\tilde{K}} \cdot \max_{j \in V} \left| f'(\bar{q}_j[t]) \right|}_{\mathcal{V}_2}$$

[9] An alternate viewpoint is that our setting and policy fit into the "drift-plus-penalty" framework in the network control literature [37], with the Lyapunov function which is the antiderivative of the congestion functions. Previous work focuses on the quadratic Lyapunov function.

$$+\underbrace{\left(W^{\mathrm{SPP}} - g(\mathbf{f}(\bar{\mathbf{q}}[t]))\right)}_{\mathcal{V}_3} + \underbrace{\mathbb{1}\left\{q_j[t] = 0, \exists j \in V\right\}}_{\mathcal{V}_4}. \tag{1.19}$$

In Lemma 1.1, the LHS of (1.19) is the suboptimality incurred by MBP (benchmark against the value of SPP) in a single period. On the RHS of (1.19), $\mathcal{V}_1$ and $\mathcal{V}_2$ come from the standard analysis of mirror descent; $\mathcal{V}_3$ is the negative of the dual suboptimality at $\mathbf{y} = (\bar{\mathbf{q}}[\mathbf{t}])$, hence it is always non-positive; $\mathcal{V}_4$ is the payoff loss because of underflow.

In the next subsection, we outline our novel analysis showing that the sum of the last three terms $\mathcal{V}_2 + \mathcal{V}_3 + \mathcal{V}_4$ is small. As a result, $\mathcal{V}_1$ is the main term on the right-hand side. Observe that it is proportional to the *Lyapunov drift*: the negative of the expected change in the Lyapunov function in one time step. The main intuition leading to the finite horizon performance guarantee in Theorem 1.1 is then that if the suboptimality of MBP in some period is large, then (1.19) implies that there is also a large negative Lyapunov drift, and this cannot be the case on average since the Lyapunov function value must remain bounded.

1.5.2 Bounding Single Period Payoff Loss

In this section we proceed to upper bound $\mathcal{V}_2 + \mathcal{V}_3 + \mathcal{V}_4$ on the RHS of (1.19). Observe that the terms $\mathcal{V}_2$ and $\mathcal{V}_4$ are non-negative, while $\mathcal{V}_3$ is non-positive, thus the goal is to show that $\mathcal{V}_3$ compensates for $\mathcal{V}_2 + \mathcal{V}_4$. First notice that $\mathcal{V}_2$ is large when there exist very short queues (because the congestion function (1.7) changes rapidly only for short queue lengths), and $\mathcal{V}_4$ is non-zero only when some queues are empty. Helpfully, it turns out that $\mathcal{V}_3$ is more negative in these same cases; we show this by exploiting the structure of the dual problem (1.14).

In Lemma 1.2 we provide an upper bound for $\mathcal{V}_3$ that becomes more negative as the shortest queue length decreases.

Lemma 1.2. *Consider congestion functions $f(\cdot)$s that are strictly increasing and contin-*

uously differentiable, and any $\boldsymbol{\phi}$ with connectedness $\alpha(\boldsymbol{\phi}) > 0$. We have

$$\mathcal{V}_3 \leq \; -\alpha(\boldsymbol{\phi}) \cdot \left[\max_{j \in V} f(\bar{q}_j) - \min_{j \in V} f(\bar{q}_j) - 2m\right]^+ .$$

We prove Lemma 1.2 in Appendix A.2 by utilizing complementary slackness for the SPP (1.10)-(1.12).

The following lemma bounds $\mathcal{V}_2 + \mathcal{V}_3 + \mathcal{V}_4$. The proof is in Appendix A.3. (In fact we prove a general version of the lemma which applies to all congestion functions that satisfy certain growth conditions formalized in Condition A.1 in Appendix A.3. The growth conditions serve to ensure that $\mathcal{V}_3$ compensates for $\mathcal{V}_2 + \mathcal{V}_4$.)

Lemma 1.3. *Consider the congestion function* (1.7)*, and any $\boldsymbol{\phi}$ with connectedness $\alpha(\boldsymbol{\phi}) > 0$. Then there exists $K_1 = \mathrm{poly}\left(m, \frac{1}{\alpha(\boldsymbol{\phi})}\right)$ such that for $K \geq K_1$,*

$$\mathcal{V}_2 + \mathcal{V}_3 + \mathcal{V}_4 \leq M_2 \cdot \frac{1}{\tilde{K}}$$

for $M_2 = Cm^2$, where $C > 0$ is a universal constant (which does not depend on K, m or $\alpha(\boldsymbol{\phi})$). Here $\tilde{K}$ was defined in (1.8)*.*

1.5.3 Proof of Theorem 1.1: Optimality Gap of MBP

Putting Lemma 1.1 and Lemma 1.3 together leads to the following proof of Theorem 1.1. The main idea is to use the so-called *Lyapunov drift argument* of [37], namely, to sum the expectation of (1.19) (the bound in Lemma 1.1) over the first T time steps. The terms $\mathcal{V}_1$ form a telescoping sum.

Proof of Theorem 1.1. Plugging in Lemma 1.3 into (1.19) in Lemma 1.1 and taking expectation, we obtain

$$W^{\mathrm{SPP}} - \mathbb{E}[v^{\mathrm{MBP}}[t]] \leq \tilde{K}\left(\mathbb{E}[F(\bar{\mathbf{q}}[t])] - \mathbb{E}[F(\bar{\mathbf{q}}[t+1])]\right) + M_2 \frac{1}{\tilde{K}} \qquad \text{for } K \geq K_1 \,. \quad (1.20)$$

Take the sum of both sides of the inequality (1.20) from $t = 0$ to $t = T - 1$, and divide the sum by T. This yields

$$W^{\text{SPP}} - W_T^{\text{MBP}} \leq \frac{\tilde{K}}{T}\big(\mathbb{E}[F(\bar{\mathbf{q}}[0])] - \mathbb{E}[F(\bar{\mathbf{q}}[T])]\big) + M_2\frac{1}{\tilde{K}} \qquad \text{for } K \geq K_1\,.$$

Using Proposition 1.1 and the inequality above, we have

$$\begin{aligned} L_T^{\text{MBP}} = W_T^* - W_T^{\text{MBP}} &\leq W^{\text{SPP}} + m\frac{K}{T} - W_T^{\text{MBP}} \\ &\leq \frac{\tilde{K}}{T}\big(m + \mathbb{E}[F(\bar{\mathbf{q}}[0])] - \mathbb{E}[F(\bar{\mathbf{q}}[T])]\big) + M_2\frac{1}{\tilde{K}} \\ &\leq \frac{\tilde{K}}{T}\left(m + \sup_{\bar{\mathbf{q}}_1,\bar{\mathbf{q}}_2\in\Omega}\big(F(\bar{\mathbf{q}}_1) - F(\bar{\mathbf{q}}_2)\big)\right) + M_2\frac{1}{\tilde{K}}\,, \end{aligned}$$

Let $M_1 \triangleq m + \sup_{\bar{\mathbf{q}}_1,\bar{\mathbf{q}}_2\in\Omega}\big(F(\bar{\mathbf{q}}_1) - F(\bar{\mathbf{q}}_2)\big)$. Observe that the function $F(\bar{\mathbf{q}})$ given in (1.18) is negative $F(\bar{\mathbf{q}}) \leq 0$ for all $\bar{\mathbf{q}} \in \Omega$, and is a convex function which achieves its minimum at $\bar{\mathbf{q}} = \frac{1}{m}\mathbf{1}$. Therefore we have

$$M_1 \leq m - \inf_{\bar{\mathbf{q}}\in\Omega} F(\bar{\mathbf{q}}) \leq m - F\left(\frac{1}{m}\mathbf{1}\right) = 3m\,.$$

Hence the finite-horizon optimality gap of MBP is upper bounded by $M_1\frac{\tilde{K}}{T} + M_2\frac{1}{\tilde{K}}$ where $M_1 = Cm$, $M_2 = Cm^2$ and C does not depend on m, K, or $\alpha(\boldsymbol{\phi})$. Moreover, $\tilde{K} = K + m\sqrt{K} \in [K, 2K]$ taking $K_1 \geq m^2$. This concludes the proof. □

1.6 Generalizations and Extensions

In this section, we allow the platform to have additional control levers beyond entry control and consider two general settings, namely, joint entry-assignment control (JEA) and joint pricing-assignment control (JPA). We also allow the queues to have finite buffers. We show that the extended models enjoy similar performance guarantees to that in Theorem 1.1 under mild conditions on the model primitives.

1.6.1 Congestion Functions for Finite Buffer Queue

Suppose the queues at a subset of nodes $V_{\text{b}} \subset V$ have a finite buffer constraint. For $j \in V_{\text{b}}$, denote the buffer size by $d_j = \bar{d}_j K$ for some scaled buffer size $\bar{d}_j \in (0, 1)$. (If

$\bar{d}_j \geq 1$, the buffer size exceeds the number of supply units $d_j \geq K$ and there is no constraint as a result, i.e., $j \notin V_{\rm b}$.) We will find it convenient to define $\bar{d}_j = 1$ for each $j \in V \backslash V_{\rm b}$. To avoid the infeasible case where the buffers are too small to accommodate all supply units, we assume that $\sum_{j\in V} \bar{d}_j > 1$. Throughout Section 1.6, the normalized state space will be

$$\Omega \triangleq \left\{ \bar{\mathbf{q}} : \mathbf{1}^{\rm T}\bar{\mathbf{q}} = 1,\ \mathbf{0} \leq \bar{\mathbf{q}} \leq \bar{\mathbf{d}} \right\}, \quad \text{where } \bar{d}_j \triangleq d_j/K\,.$$

Similar to the case of entry control, we need to keep $\bar{\mathbf{q}}$ in the interior of Ω, which is achieved by defining the normalized queue lengths $\bar{\mathbf{q}}$ as

$$\bar{q}_j \triangleq \frac{q_j + \bar{d}_j \delta_K}{\tilde{K}} \quad \text{for} \quad \delta_K = \sqrt{K} \quad \text{and} \quad \tilde{K} \triangleq K + \left(\sum_{j\in V} \bar{d}_j \right) \delta_K\,. \tag{1.21}$$

One can verify that $\bar{\mathbf{q}} \in \Omega$ for any feasible state $\mathbf{q}$. When $\bar{d}_j = 1$ for all $j \in V$, the definition of $\bar{q}_j$ in (1.21) reduces to the one in (1.8). The congestion functions $(f_j(\cdot))_{j\in V}$ are monotone increasing functions that map (normalized) queue lengths to congestion costs. Here we will state our main results for the congestion function vector

$$f_j(\bar{q}_j) \triangleq \begin{cases} \sqrt{m} \cdot C_b \cdot \left(\left(1 - \frac{\bar{q}_j}{\bar{d}_j}\right)^{-\frac{1}{2}} - \left(\frac{\bar{q}_j}{\bar{d}_j}\right)^{-\frac{1}{2}} - D_b \right), & \forall j \in V_{\rm b}\,, \\ -\sqrt{m} \cdot \bar{q}_j^{-\frac{1}{2}} & \forall j \in V \backslash V_{\rm b}\,. \end{cases} \tag{1.22}$$

Here C_b and D_b are normalizing constants[10] chosen to ensure that (i) for all $j, k \in V$, we have that $f_j(\bar{q}_j) = f_k(\bar{q}_k)$ when both queues are empty $q_j = q_k = 0$; (ii) for all $j, k \in V_{\rm b}$, we have that $f_j(\bar{q}_j) = f_k(\bar{q}_k)$ when both queues are full $q_j = d_j$, $q_k = d_k$. (We state the results for other choices of congestion functions in Appendix A.3.)

Note that $f_j(\cdot)$ in (1.22) is identical to $f(\cdot)$ in (1.7) for $j \notin V_b$, i.e., (1.22) is a generalization of (1.7) to the case where some queues have buffer constraints. The intuitive reason (1.22) is a suitable congestion function is that it enables MBP to focus on queues

[10]Define $\epsilon \triangleq \frac{\delta_K}{\tilde{K}}$. Let $h_{\rm b}(\bar{q}) \triangleq (1-\bar{q})^{-\frac{1}{2}} - \bar{q}^{-\frac{1}{2}}$ and $h(\bar{q}) \triangleq -\bar{q}^{-\frac{1}{2}}$. Define $C_b \triangleq \frac{h(\epsilon) - h(1/\sum_{j\in V} \bar{d}_j)}{h_{\rm b}(\epsilon) - h_{\rm b}(1/\sum_{j\in V} \bar{d}_j)}$ and $D_b \triangleq h_{\rm b}(1/\sum_{j\in V} \bar{d}_j) - C_b^{-1} h(1/\sum_{j\in V} \bar{d}_j)$. In addition to the properties listed in the main text, we also have that $f_j(\bar{d}_j/\sum_{j\in V} \bar{d}_j)$ has the same value for all $j \in V$. These properties are useful in the following analysis.

which are currently either almost empty or almost full (the congestion function values for those queues take on their smallest and largest values, respectively), and use the control levers available to make the queue lengths for those queues trend strongly away from the boundary they are close to.

1.6.2 Joint Entry-Assignment Setting

We first generalize the entry control setting introduced in Section 1.2 by allowing the system to choose a flexible pickup and dropoff node for each demand, and furthermore allowing demand arrival rates to vary in time. Formally, instead of an origin node and a destination node, in this setting each demand unit has an abstract *type* $\tau \in \mathcal{T}$, and the type for the demand unit in period t is drawn from distribution $\boldsymbol{\phi}^t = (\phi^t_\tau)_{\tau \in \mathcal{T}}$, independently across t. The demand type at period t is denoted by $\tau[t]$. Each demand type $\tau \in \mathcal{T}$ has a pick-up neighborhood $\mathcal{P}(\tau) \subset V, \mathcal{P}(\tau) \neq \emptyset$ and drop-off neighborhood $\mathcal{D}(\tau) \subset V, \mathcal{D}(\tau) \neq \emptyset$. The sets $(\mathcal{P}(\tau))_{\tau \in V}$ and $(\mathcal{D}(\tau))_{\tau \in V}$ are model primitives. (In shared transportation systems, each demand type τ may correspond to an (origin, destination) pair in V^2, with $\mathcal{P}(\tau)$ being nodes close to the origin and $\mathcal{D}(\tau)$ being nodes close to the destination. In the special case that $\mathcal{P}(\tau)$ and $\mathcal{D}(\tau)$ are singletons for each $\tau \in \mathcal{T}$ we recover the illustrative model in Section 1.2.)

The platform control and payoff in this setting are as follows. At time t, after observing the demand type $\tau[t] = \tau$, the system makes a decision

$$(x_{j\tau k}[t])_{j \in \mathcal{P}(\tau), k \in \mathcal{D}(\tau)} \in \{0,1\}^{|\mathcal{P}(\tau)| \cdot |\mathcal{D}(\tau)|} \quad \text{such that} \quad \sum_{j \in \mathcal{P}(\tau), k \in \mathcal{D}(\tau)} x_{j\tau k}[t] \leq 1 \,. \tag{1.23}$$

Here $x_{j\tau k}[t] = 1$ stands for the platform choosing pick-up node $j \in \mathcal{P}(\tau)$ and drop-off node $k \in \mathcal{D}(\tau)$, causing a supply unit to be relocated from j to k. The constraint in (1.23) captures that each demand unit is either served by one supply unit, or not served. With $x_{j\tau k}[t] = 1$, the system collects payoff $v[t] = w_{j\tau k}$. Without loss of generality, we assume the scaling

$$\max_{\tau\in\mathcal{T},j\in\mathcal{P}(\tau),k\in\mathcal{D}(\tau)} |w_{j\tau k}| = 1\,.$$

Because the queue lengths are non-negative and upper bounded by buffer sizes, we require the following constraint to be met at any t:

$$x_{j\tau k}[t] = 0 \quad \text{if} \quad q_j[t] = 0 \text{ or } q_k[t] = d_k\,.$$

As a convention, let $x_{j\tau' k} = 0$ if $\tau' \neq \tau$. The dynamics of system state $\mathbf{q}[t]$ is as follows:

$$\mathbf{q}[t+1] = \mathbf{q}[t] + \sum_{\tau\in\mathcal{T},j\in\mathcal{P}(\tau),k\in\mathcal{D}(\tau)} (-\mathbf{e}_j + \mathbf{e}_k)x_{j\tau k}[t]\,. \tag{1.24}$$

The definition of a feasible policy is similar to the case of entry control, hence we skip the details. We once again define the transient and steady state optimality gaps L_T^π and L^π as in Section 1.2 via (1.3)-(1.5).

The dual problem to the SPP in period t in the JEA setting (see Appendix A.1.1 for the SPP, which we denote by SPP^t) is

$$\text{minimize}_{\mathbf{y}}\; g^t_{\text{JEA}}(\mathbf{y}), \quad \text{for } g^t_{\text{JEA}}(\mathbf{y}) \triangleq \sum_{\tau\in\mathcal{T}} \phi^t_\tau \max_{j\in\mathcal{P}(\tau),k\in\mathcal{D}(\tau)} \left(w_{j\tau k} + y_j - y_k\right)^+\,. \tag{1.25}$$

As before, MBP is defined to achieve the argmax in the definition of the dual objective g_{JEA}, with the ys replaced by congestion costs: (i) Again, decisions are made based on payoffs adjusted by congestion costs, and demand units which generate (weakly) positive adjusted payoff are admitted. (ii) The pickup and dropoff locations are chosen to maximize the adjusted payoff.

ALGORITHM 2: Mirror Backpressure (MBP) Policy for Joint Entry-Assignment

At the start of period t, the system observes demand type $\tau[t] = \tau$.

$(j^*, k^*) \leftarrow \text{argmax}_{j \in \mathcal{P}(\tau), k \in \mathcal{D}(\tau)} w_{j\tau k} + f_j(\bar{q}_j[t]) - f_k(\bar{q}_k[t])$

if $w_{j^*\tau k^*} + f_{j^*}(\bar{q}_{j^*}[t]) - f_{k^*}(\bar{q}_{k^*}[t]) \geq 0$ **and** $q_{j^*}[t] > 0,\ q_{k^*}[t] < d_{k^*}$ **then**

  $x_{j^*\tau k^*}[t] \leftarrow 1$, i.e., serve the incoming demand using a supply unit from j^* and relocate it to k^* ;

else

  $x_{j^*\tau k^*}[t] \leftarrow 0$, i.e., drop the incoming demand;

end

The queue lengths update as $\bar{\mathbf{q}}[t+1] = \bar{\mathbf{q}}[t] - \frac{1}{K} x_{j^*\tau k^*}[t](\mathbf{e}_{j^*} - \mathbf{e}_{k^*})$.

We make the following connectivity assumption on the primitives $(\boldsymbol{\phi}^t, \mathcal{P}, \mathcal{D})$ for all t in the horizon.

Condition 1.2 (Strong Connectivity of $(\boldsymbol{\phi}^t, \mathcal{P}, \mathcal{D})$)**.** *For any demand arrival rates $\boldsymbol{\phi}$, define the connectedness of triple $(\boldsymbol{\phi}, \mathcal{P}, \mathcal{D})$ as*

$$\alpha(\boldsymbol{\phi}, \mathcal{P}, \mathcal{D}) \triangleq \min_{S \subsetneq V, S \neq \emptyset} \sum_{\tau \in \mathcal{P}^{-1}(S) \cap \mathcal{D}^{-1}(V \backslash S)} \phi_\tau \, . \tag{1.26}$$

Here $\mathcal{P}^{-1}(S) \triangleq \{\tau \in \mathcal{T} : \mathcal{P}(\tau) \cap S \neq \emptyset\}$ is the set of demand types for which nodes S can serve as a pickup node; and $\mathcal{D}^{-1}(\cdot)$ is defined similarly. We assume that for some $\alpha_{\min} > 0$, for all t in the horizon it holds that $(\boldsymbol{\phi}^t, \mathcal{P}, \mathcal{D})$ is $\alpha_{\min}$-strongly connected, namely, $\alpha(\boldsymbol{\phi}^t, \mathcal{P}, \mathcal{D}) \geq \alpha_{\min}$.

If each type $\tau \in \mathcal{T}$ corresponds to an origin-destination pair $\tau = (j, k) \in V^2$ and $\mathcal{P}(\tau) = \{j\}$, $\mathcal{D}(\tau) = \{k\}$ and demand arrival rates are stationary $\boldsymbol{\phi}^t = \boldsymbol{\phi}$, then the JEA setting reduces to entry control model in Section 1.2 and $\alpha(\boldsymbol{\phi}, \mathcal{P}, \mathcal{D}) = \alpha(\boldsymbol{\phi})$ for $\alpha(\boldsymbol{\phi})$ defined in (1.6).

Definition 1.1. *We say that demand arrival rates vary η-slowly for some $\eta \geq 0$ if $\|\boldsymbol{\phi}^{t+1} - \boldsymbol{\phi}^t\|_1 \leq \eta$ for all $t \geq 0$ in the horizon of interest.*

Note that any sequence of demand arrival rates varies 2-slowly, so $\eta \in [0, 2]$, with $\eta = 0$ being the case of stationary demand arrival rates.

We show the following performance guarantee, analogous to Theorem 1.1.

Theorem 1.2. *Fix a set V of $m \triangleq |V| > 1$ nodes, a subset $V_\mathrm{b} \subseteq V$ of buffer-constrained nodes with scaled buffer sizes $\bar{d}_j \in (0,1)\ \forall j \in V_\mathrm{b}$ satisfying*[11] *$\sum_{j\in V} \bar{d}_j > 1$, and a minimum connectivity $\alpha_{\min} > 0$. Then there exists $K_1 = \mathrm{poly}\left(m, \bar{\mathbf{d}}, \frac{1}{\alpha_{\min}}\right)$, $M_1 = Cm$, and $M_2 = C\frac{\sqrt{m}}{\min_{j\in V} \bar{d}_j}\left(\frac{\sum_{j\in V}\bar{d}_j}{\min\{\sum_{j\in V}\bar{d}_j - 1, 1\}}\right)^{3/2}$ where C is a universal constant that does not depend on m, $\bar{\mathbf{d}}$, η or $\alpha_{\min}$, such that for the congestion functions $(f_j(\cdot))_{j\in V}$ defined in (1.22), the following guarantee holds for Algorithm 2. For any horizon T, any $K \geq K_1$, and any sequence of demand arrival rates $(\boldsymbol{\phi}^t)_{t=0}^{T-1}$ which varies η-slowly (for some $\eta \in [0,2]$) and pickup and dropoff neighborhoods $\mathcal{P}$ and $\mathcal{D}$ such that $(\boldsymbol{\phi}^t, \mathcal{P}, \mathcal{D})$ is $\alpha_{\min}$-strongly connected for all $t \leq T-1$, we have*

$$L_T^{\mathrm{MBP}} \leq M_1 \cdot \left(\frac{K}{T} + \sqrt{\eta K}\right) + M_2 \cdot \frac{1}{K}\,.$$

In Appendix A.3 we prove a general version of Theorem 1.2 which provides a performance guarantee for a large class of congestion functions.

1.6.3 Joint Pricing-Assignment Setting

In this section, we consider the joint pricing-assignment (JPA) setting and design the corresponding MBP policy. The platform's control problem is to set a price for each demand origin-destination pair, and decide an assignment at each period to maximize payoff. Our model here will be similar to that of Banerjee, Freund, and Lykouris [1], except that the platform does *not* know demand arrival rates, and we allow a finite horizon. The proposed algorithm will be a generalization of backpressure based joint-rate-scheduling control policies [see, e.g., 39, 14]. The demand types τ, pick-up neighborhood $\mathcal{P}(\tau)$ and drop-off neighborhood $\mathcal{D}(\tau)$ are defined in the same way as in section 1.6.2. For simplicity, we assume that the demand type distribution $\boldsymbol{\phi} = (\phi_\tau)_{\tau\in\mathcal{T}}$ is time invariant in this subsection.

[11] Recall that we define $\bar{d}_j \triangleq 1$ for all $j \in V \backslash V_\mathrm{b}$.

The platform control and payoff in this setting are as follows. At time t, after observing the demand type $\tau[t] = \tau$, the system chooses a *price* $p_\tau[t] \in [p_\tau^{\min}, p_\tau^{\max}]$ and a decision

$$(x_{j\tau k}[t])_{j\in\mathcal{P}(\tau),k\in\mathcal{D}(\tau)} \in \{0,1\}^{|\mathcal{P}(\tau)|\cdot|\mathcal{D}(\tau)|} \quad \text{such that} \quad \sum_{j\in\mathcal{P}(\tau),k\in\mathcal{D}(\tau)} x_{j\tau k}[t] \le 1\,. \tag{1.27}$$

As before we require

$$x_{j\tau k}[t] = 0 \quad \text{if} \quad q_j[t] = 0 \text{ or } q_k[t] = d_k\,.$$

The result of the platform control is as follows:

(1) Upon seeing the price, the arriving demand unit will decline (to buy) with probability $F_\tau(p_\tau[t])$, where $F_\tau(\cdot)$ is the cumulative distribution function of type τ demand's willingness-to-pay.

(2) If the demand accepts (i.e., buys), the system state updates as per

$$\mathbf{q}[t+1] = \mathbf{q}[t] + \sum_{j\in\mathcal{P}(\tau),k\in\mathcal{D}(\tau)} (-\mathbf{e}_j + \mathbf{e}_k)x_{j\tau k}[t]\,. \tag{1.28}$$

Meanwhile, the platform collects payoff $v[t] = p_\tau[t] - c_{j\tau k}$ where $c_{j\tau k}$ is the "cost" of serving a demand unit of type τ using pick-up node j and drop-off node k.

(3) If the demand unit declines, the supply units do not move and $v[t] = 0$.

We assume the following regularity conditions to hold for demand functions $\big(F_\tau(p_\tau)\big)_\tau$. These assumptions are quite standard in the revenue management literature, [see, e.g., 31].

Condition 1.3. *(1) Assume*[12] $F_\tau(p_\tau^{\min}) = 0$ *and that* $F_\tau(p_\tau^{\max}) = 1$.

(2) Each demand type's willingness-to-pay is non-atomic with support $[p_\tau^{\min}, p_\tau^{\max}]$ *and positive density everywhere on the support; hence* $F_\tau(p_\tau)$ *is differentiable and strictly increasing on* $(p_\tau^{\min}, p_\tau^{\max})$. *(If the support is a subinterval of* $[p_\tau^{\min}, p_\tau^{\max}]$*, we redefine* $p_\tau^{\min}$ *and* $p_\tau^{\max}$ *to be the boundaries of this subinterval.)*

(3) The revenue functions $r_\tau(\mu_\tau) \triangleq \mu_\tau \cdot p_\tau(\mu_\tau)$ *are concave and twice continuously differentiable, where* μ_τ *denotes the fraction of demand of type* τ *which is realized (i.e.,*

[12]The assumption $F_\tau(p_\tau^{\min}) = 0$ is without loss of generality, since if a fraction of demand is unwilling to pay $p_\tau^{\min}$, that demand can be excluded from $\boldsymbol{\phi}$ itself.

willing to pay the price offered).

As a consequence of Condition 1.3 parts 1 and 2, the willingness to pay distribution $F_\tau(\cdot)$ has an inverse denoted as $p_\tau(\mu_\tau) : [0,1] \to [p_\tau^{\min}, p_\tau^{\max}]$ which gives the price which will cause any desired fraction $\mu_\tau \in [0,1]$ of demand to be realized. (The concavity assumption in part 3 of the condition is stated in terms of this function $p_\tau(\cdot)$.) Without loss of generality, let $\max_{\tau\in\mathcal{T}} p_\tau^{\max} + \max_{j,k\in V,\tau\in\mathcal{T}} |c_{j\tau k}| = 1$.

In the JPA setting, the net demand $\phi_\tau\mu_\tau$ plays a role in myopic revenues but also affects the distribution of supply, and the chosen prices need to balance myopic revenues with maintaining a good spatial distribution of supply. Intuitively, when sufficiently flexible pricing is available as a control lever, the system should modulate the quantity of demand through changing the prices (and serving all the demand which is then realized) rather than apply entry control (i.e., dropping some demand proactively). Our MBP policy for this setting will have this feature.

The dual problem to the SPP in the JPA setting (the SPP is stated in Appendix A.1.2) is[13]

$$\text{minimize}_{\mathbf{y}}\, g_{\text{JPA}}(\mathbf{y}) \quad \text{for } g_{\text{JPA}}(\mathbf{y}) \triangleq \sum_{\tau\in\mathcal{T}} \phi_\tau \max_{\{0\le\mu_\tau\le 1\}} \left(r_\tau\left(\mu_\tau\right) + \mu_\tau \max_{j\in\mathcal{P}(\tau),k\in\mathcal{D}(\tau)} \left(-c_{j\tau k} + y_j - y_k\right)\right). \tag{1.29}$$

Once again, the MBP policy (Algorithm 3 below) is defined to achieve the argmaxes in the definition of the dual objective $g_{\text{JPA}}(\cdot)$ with the ys replaced by congestion costs: MBP dynamically sets prices p_τ such that mean fraction of demand realized under the policy is the outer argmax in the definition (1.29) of $g_{\text{JPA}}(\cdot)$, and the assignment decision of MBP achieves the inner argmax in the definition (1.29) of $g_{\text{JPA}}(\cdot)$. The policy again has the property that it executes stochastic mirror descent on the dual objective $g_{\text{JPA}}(\cdot)$.

The MBP policy retains the advantage that it does not require any prior knowledge of gross demand $\boldsymbol{\phi}$. We assume that the willingness-to-pay distributions $F_\tau(\cdot)$s are exactly known to the platform; it may be possible to relax this assumption via a modified policy

[13]The derivation of the dual objective is in Appendix A.2.

which "learns" the $F_\tau(\cdot)$s, however, pursuing this direction is beyond the scope of the present paper.

ALGORITHM 3: Mirror Backpressure (MBP) Policy for Joint Pricing-Assignment

At the start of period t, the system observes $\tau[t] = \tau$.

$(j^*, k^*) \leftarrow \arg\max_{j\in\mathcal{P}(\tau),k\in\mathcal{D}(\tau)} \left\{-c_{j\tau k} + f_j(\bar{q}_j[t]) - f_k(\bar{q}_k[t])\right\}$;

if $q_{j^*}[t] > 0,\ q_{k^*}[t] < d_{k^*}$ **then**

$\mu_\tau[t] \leftarrow \operatorname{argmax}_{\mu_\tau\in[0,1]} \left\{r_\tau(\mu_\tau) + \mu_\tau \cdot (-c_{j^*\tau k^*} + f_{j^*}(\bar{q}_{j^*}[t]) - f_{k^*}(\bar{q}_{k^*}[t]))\right\}$;

$p_\tau[t] \leftarrow F_\tau^{-1}(\mu_\tau[t])$;

$x_{j^*\tau k^*}[t] \leftarrow 1$, i.e., serve the incoming demand (if it stays) by pick up from j^* and drop off at k^*;

else

$x_{j^*\tau k^*}[t] \leftarrow 0$, i.e., drop the incoming demand;

end

The queue lengths update as $\bar{\mathbf{q}}[t+1] = \bar{\mathbf{q}}[t] - \frac{1}{K} x_{j^*\tau k^*}[t](\mathbf{e}_{j^*} - \mathbf{e}_{k^*})$.

Condition 1.3 ensures that Algorithm 3 has two key desirable properties:

(1) The computed prices satisfy $p_\tau[t] \in [p_\tau^{\min}, p_\tau^{\max}]$ (by the observation following Condition 1.3).

(2) The optimization problem for computing $\mu_\tau[t]$ is a one-dimensional concave maximization problem (Condition 1.3 part 3), hence $\mu_\tau[t]$ can be efficiently computed.

We have the following performance guarantee for Algorithm 3, analogous to Theorem 1.1.

Theorem 1.3. *Fix a set V of $m = |V| > 1$ nodes, scaled buffer sizes $\bar{\mathbf{d}} = (\bar{d}_j)_{j\in V}$ with*[14] *$\bar{d}_j \in (0,1]$ and $\sum_{j\in V} \bar{d}_j > 1$, minimum and maximum allowed prices $(p_\tau^{\min}, p_\tau^{\max})_{\tau\in\mathcal{T}}$, any $(\boldsymbol{\phi}, \mathcal{P}, \mathcal{D})$ that satisfy Condition 1.2 (strong connectivity), and willingness-to-pay distributions $(F_\tau)_{\tau\in\mathcal{T}}$ that satisfy Condition 1.3. Then there exist $K_1 < \infty$, $M_1 = \text{poly}\left(m, \bar{\mathbf{d}}\right)$, and $M_2 = \text{poly}\left(m, \bar{\mathbf{d}}\right)$ such that for the congestion functions $(f_j(\cdot))$s defined in (1.22), the following guarantee holds for Algorithm 3. For any horizon T and for any $K \geq K_1$,*

[14] Recall that we use $\bar{d}_j \triangleq 1$ for nodes $j \in V\backslash V_b$ which do not have a buffer size constraint.

we have

$$L_T^{\mathrm{MBP}} \leq M_1 \cdot \frac{K}{T} + M_2 \cdot \frac{1}{K}, \qquad \text{and} \qquad L^{\mathrm{MBP}} \leq M_2 \cdot \frac{1}{K}.$$

1.7 Application to Shared Transportation Systems

Our setting can be mapped to shared transportation systems such as bike sharing and ride-hailing systems. In this context, the nodes in our model correspond to geographical locations, while supply units and demand units correspond to vehicles and customers, respectively.

Dynamic incentive program for bike sharing systems. A major challenge faced by bike sharing systems such as Citi Bike in New York City is the frequent out-of-bike and out-of-dock events caused by demand imbalance. One popular solution is to dynamically incentivize certain trips by awarding points (with cash value) depending on a trip's pickup and dropoff locations [24]. Thus the problem of designing a dynamic incentive program is addressed (in a stylized way) by the pricing setting we study (the joint pricing-assignment setting studied Section 1.6.3, but with no assignment flexibility). MBP tells the system operator, quantitatively, how to reward rides that relocate bikes to locations which have a scarcity of bikes. In docked bike sharing systems, there is a constraint on the number of docks available at each location. Such constraints are seamlessly handled in our framework as detailed earlier in Section 1.6.1. One concern may be that our model ignores travel delays. However, in most bike sharing systems, the fraction of bikes in transit at any time is typically quite small (under 10-20%).[15] As a result, we expect our control insights to retain their power despite the presence of delays. (Indeed, we will numerically demonstrate in Section 1.7.1 that this is the case in a realistic ridehailing setting; see the

[15]The report https://nacto.org/bike-share-statistics-2017/ tells us that U.S. dock-based systems produced an average of 1.7 rides/bike/day, while dockless bike share systems nationally had an average of about 0.3 rides/bike/day. Average trip duration was 12 minutes for pass holders (subscribers) and 28 mins for casual users. In other words, for most systems, each bike was used less than 1 hour per day, which implies that less than 10% of bikes are in use at any given time during day hours (in fact the utilization is below 20% even during rush hours).

excess supply case where MBP performs well even when the vast majority of supply is in transit at any time.) We leave a detailed study of bike sharing platforms to future work.

Online control of ride-hailing platforms. Ride-hailing platforms make dynamic decisions to optimize their objectives (e.g., revenue, welfare, etc.). For most ride-hailing platforms in North America, pricing is used to modulate demand. In certain countries such as China, however, pricing is a less acceptable lever, hence admission control of customers is used as a control lever instead. In both cases, the platform further decides where (near the demand's origin) to dispatch a car from, and where (near the demand's destination) to drop off a customer. These scenarios are captured, respectively, by the joint entry-assignment (JEA)[16] and joint pricing-assignment (JPA) models studied in Section 1.6. A concern may be that travel delays play a significant role in ride-hailing, whereas delays are ignored in our theory. In the following subsection, we summarize a numerical investigation of ride-hailing focusing on entry and assignment controls only (a full description is provided in Appendix A.4). We find that MBP performs well despite the presence of travel delays. In order to address the case where the available supply is scarce, we heuristically adapt MBP to incorporate the Little's law constraint (Section 1.7.1).

1.7.1 Numerical simulations in a realistic ride-hailing setting

We simulate the MBP policy in a realistic ride-hailing environment using yellow cab data from NYC Taxi & Limousine Commission and travel times from Google Maps. In the interest of space, we provide only a summary of our simulations here and defer a full description to Appendix A.4.

We allow the platform two control levers: entry control and assignment/dispatch

[16]The JEA setting can be mapped to ride-hailing as follows: there is a demand type τ corresponding to each (origin, destination) pair $(j, k) = V^2$, with $\mathcal{P}(\tau)$ being nodes close to the origin j and $\mathcal{D}(\tau)$ being nodes close to the destination k.

control, similar to the JEA setting[17,18] in Section 1.6.2. Our theoretical model made the simplifying assumption that pickup and service of demand are *instantaneous*. We relax this assumption in our numerical experiments by adding realistic travel times. We consider the following two cases:

(1) *Excess supply.* The number of cars in the system is slightly (5%) above the "fluid requirement" (see Appendix A.4.1 for details on the "fluid requirement") to achieve the value of the static planning problem.

(2) *Scarce supply.* The number of cars fall short (by 25%) of the "fluid requirement", i.e., there are not enough cars to realize the optimal solution of static planning problem (ignoring stochasticity).

Summary of findings. We make a natural modification of the MBP policy (with congestion function (1.7)) to account for finite travel times; specifically, we employ a *supply-aware MBP* policy which estimates and uses a shadow price of keeping a vehicle (supply unit) occupied for one unit of time. This policy is described below in Section 1.7.1. We find that in both the excess supply and the scarce supply cases, the MBP policy, which is given no information about the demand arrival rates, significantly outperforms the static (fluid-based) policy, even when the latter is provided with prior knowledge of exact demand arrival rates. The MBP policy also vastly outperforms the greedy non-idling policy, which demonstrates the practical importance and value of proactively dropping demand.

The Supply-Aware MBP Policy

In order to heuristically modify MBP to account for travel times, we begin by observing that the SPP must now include a Little's law constraint. (The same observation

[17] The correspondence between our (ride-hailing) simulation setting and the JEA setting is as follows: In the ride-hailing setting, the type of a demand is its origin-destination pair, i.e. $\mathcal{T} = V \times V$. For type (j, k) demand, its supply neighborhood is the neighboring locations of j, which we denote by (with a slight abuse of notation) $\mathcal{P}(j)$. We do not consider flexible drop-off, therefore $\mathcal{D}(j, k) = \{k\}$.

[18] In our simulations, we focus on the special case where demand is stationary instead of time-varying, even though MBP policies are expected to work well if demand varies slowly over time. We make this choice because it allows us to compare performance against that of the policy proposed in [1] for the stationary demand setting.

was previously leveraged by [2] and [1] to formally handle travel times, albeit under the assumption that travel times are i.i.d. exponentially distributed.) Our heuristic modification of MBP will maintain an estimate of the shadow price corresponding to the Little's law constraint, and penalize rides appropriately.

Applying Little's Law, if the optimal solution $\mathbf{z}^*$ of the SPP (A.1)-(A.3) (see Appendix A.1.1; here we work with the special case where $\boldsymbol{\phi}$ does not depend on t) is realized as the average long run assignment, the mean number of cars which are occupied in picking up or transporting customers is $\sum_{j,k\in V}\sum_{i\in\mathcal{P}(j)} D_{ijk}\cdot z^*_{ijk}$, for $D_{ijk} \triangleq \tilde{D}_{ij} + \hat{D}_{jk}$, where $\tilde{D}_{ij}$ is the pickup time from i to j and $\hat{D}_{jk}$ is the travel time from j to k. We augment the SPP with the additional supply constraint

$$\sum_{j,k\in V}\sum_{i\in\mathcal{P}(j)} D_{ijk}\cdot z_{ijk} \leq K\,. \tag{1.30}$$

We propose and test in the simulation the following heuristic policy inspired by MBP, that additionally incorporates the supply constraint. We call it *supply-aware MBP*. Given a demand arrival with origin j and destination k, the policy makes its decision as per:

$$i^* \leftarrow \arg\max_{i\in\mathcal{P}(j)} \left\{w_{ijk} + f(\bar{q}_i[t]) - f(\bar{q}_k[t]) - v[t]D_{ijk}\right\}$$

If $w_{i^*jk} + f(\bar{q}_{i^*}[t]) - f(\bar{q}_k[t]) - v[t]D_{i^*jk} \geq 0$ **and** $q_{i^*}[t] > 0$, dispatch from i^*, **else** Drop,

We define the tightened supply constraint as

$$\sum_{j,k\in V}\sum_{i\in\mathcal{P}(j)} D_{ijk}\cdot z_{ijk} \leq 0.95K\,, \tag{1.31}$$

where the coefficient of K is the flexible "utilization" parameter, that we have set 0.95, meaning that we are aiming to keep 5% vehicles free on average, systemwide.[19] Here $v[t]$ is the current estimate of the shadow price for a "tightened" version of supply constraint (1.30). We use the congestion function $f_j(\bar{q}_j) = \sqrt{m}\cdot\bar{q}_j^{-1/2}$, i.e. the one given in (1.7), in our numerical simulations. An adaptation here is that the queue lengths are normalized by the estimated number of *free* cars instead of K, which we set as $0.05K$ to be consistent

[19] Keeping a small fraction of vehicles free is helpful in managing the stochasticity in the system. Note that the present paper does not study how to systematically choose the utilization parameter.

with the "utilization" parameter we choose. We update $v[t]$ as

$$v[t+1] = \left[v[t] + \frac{1}{K} \left(\sum_{j,k \in V} \sum_{i \in \mathcal{P}(j)} D_{ijk} \cdot \mathbb{1}\{(o[t], d[t]) = (j,k), \text{MBP dispatches from } i\} - 0.95K \right) \right]^{+}.$$

An iteration of supply-aware MBP is equivalent to executing a (dual) stochastic mirror descent step on the supply-aware SPP with objective (A.1) and constraints (A.2), (A.3) and (1.31).

1.8 Application to Scrip Systems

In this section, we illustrate the application of our model to scrip systems. A scrip system is a nonmonetary trade economy where agents use scrips (tokens, coupons, artificial currency) to exchange services. These systems are typically implemented when monetary transfer is undesirable or impractical. For example, [26] suggest that in kidney exchange, to align the incentives of hospitals, the exchange should deploy a scrip system that awards points to hospitals that submit donor-patient pairs to the central exchange, and deducts points from hospitals that conduct transplantations. Another well-known example is Capitol Hill Babysitting Co-op [25, 40], where married couples pay for babysitting services by another couples with scrips. A key challenge in these markets is the design of the admission-and-provider-selection rule: If an agent is running low on scrip balance, should they be allowed to request services? If yes, and if there are several possible providers for a trade, who should be selected for service?

We introduce a natural model of a scrip system with multiple agents and heterogeneous services, where agents exchange scrips (i.e., artificial currency) for services. There is a central planner who tries to maximize social welfare by making decisions over whether a trade should occur when a service request arises, and if so, who the service provider should be. The setting is seen to be a special case of the joint entry-assignment (JEA) setting studied in Section 1.6; yielding a simple MBP control rule that comes with the guarantee that it asymptotically maximizes social welfare.

1.8.1 Model of Scrip Systems

We now describe a model of a service exchange (i.e., a scrip system). Consider an economy with a finite number of agents indexed by $j \in V$. There are finitely many types of service types Σ indexed by $\sigma \in \Sigma$. A demand type $\tau = (j, \sigma)$ is specified by the requestor $j \in V$ along with the requested service type $\sigma \in \Sigma$, i.e., the set of demand types $\mathcal{T} \subseteq V \times \Sigma$. If the demand is served, the requestor pays a scrip to the service provider. Accordingly, for each demand type $\tau = (j, \sigma)$, we define the *compatible* set of agents who can serve it as $\mathcal{D}(\tau) \subseteq V \backslash \{j\}$. We again consider a slotted time model, where in each period exactly one service request arises, with demand type drawn i.i.d. from the distribution[20] $\boldsymbol{\phi} = (\phi_\tau)_{\tau \in \mathcal{T}}$. There are a fixed number K of scrips in circulation, distributed among the agents. For each $\tau = (j, \sigma) \in \mathcal{T}$, serving a demand type $\tau = (j, \sigma)$ generates payoff $w_{j\sigma}$.

Observe that our model here is a special case of the JEA setting.[21]

Comparison with the model in Johnson et al. [25]. The work [25] consider the case where there is only *one* type of service which *all* agents can provide, and requests arrive at the same rate from all agents. One one hand, we significantly generalize their model by considering heterogeneous service types, general compatibility structures, and asymmetric service request arrivals. They obtain an optimal rule for the symmetric fully connected setting, whereas we develop an asymptotically optimal control rule for the general setting. On the other hand, we only focus on the central planner setting, and leave the incentives of agents for future work (see the remarks in Section 1.8.2).

[20] Time-varying demand arrival rates can be seamlessly handled since they are permitted in the JEA setting; we work with stationary arrival rates only for the sake of brevity.

[21] This can be seen as follows: For each demand type $\tau \in \mathcal{T}$, the compatible set of service providers $\mathcal{D}(\tau)$ is identified with the "dropoff neighborhood" for τ. The "pickup neighborhood" is a singleton set consisting of the requestor $\mathcal{P}(\tau) = \{j\}$. Finally, for each $k \in \mathcal{D}(\tau)$ we define the payoff $w_{j\tau k} \triangleq w_{j\sigma}$. The primitives $V, \mathcal{P}, \mathcal{D}, \boldsymbol{\phi}$ and $(w_{j\tau k})_{\tau=(j,\sigma)\in\mathcal{T}, k\in\mathcal{D}(\tau)}$ fully specify the JEA setting.

1.8.2 MBP Control Rule and Asymptotic Optimality

Since the model above is a special case of the JEA setting, we immediately obtain an MBP control rule for scrip systems that achieves asymptotic optimality as a special case of Algorithm 2 and Theorem 1.2. This control rule is specified in Algorithm 4 below. The congestion function $f(\cdot)$ can again be chosen flexibly; we state our formal guarantee for the congestion function in (1.7). Denote the normalized number of scrips (defined in (1.8)) in the possession of agent i by $\bar{q}_i$.

ALGORITHM 4: MBP Admission-and-provider-selection rule for scrip systems

---

At the start of period t, the central planner receives a request from agent j for service type σ, i.e., demand type $\tau = (j, \sigma)$ arises.

if $w_{j\sigma} + f(\bar{q}_j[t]) - \min_{k \in \mathcal{D}(\tau)} f(\bar{q}_k[t]) \geq 0$ **and** $\bar{q}_i[t] > 0$ **then**

  $k^* \leftarrow \mathrm{argmin}_{k \in \mathcal{D}(\tau)} f(\bar{q}_k[t])$,

  Let agent k^* provide the service to j, and agent j gives one scrip to agent k^* ;

else

  Reject the service request from agent j;

end

Theorem 1.2 immediately implies the following performance guarantee for Algorithm 4.

Corollary 1.1. *Consider a set of m agents and any demand type distribution and compatibilities $(\boldsymbol{\phi}, \mathcal{P}, \mathcal{D})$ (where $\mathcal{P}$ is identity) that satisfy Condition 1.2. Then there exists $K_1 = \mathrm{poly}\left(m, \frac{1}{\alpha(\boldsymbol{\phi}, \mathcal{P}, \mathcal{D})}\right)$ and a universal $C > 0$ that does not depend on m, K or $\alpha(\boldsymbol{\phi}, \mathcal{P}, \mathcal{D})$, such that for the congestion function $f(\cdot)$ defined in (1.7), for any $K \geq K_1$, the following guarantee holds for Algorithm 4*

$$L_T^{\mathrm{MBP}} \leq M_1 \cdot \frac{K}{T} + M_2 \cdot \frac{1}{K}, \qquad \text{and} \qquad L^{\mathrm{MBP}} \leq M_2 \cdot \frac{1}{K}, \qquad \text{for } M_1 \triangleq Cm \text{ and } M_2 \triangleq Cm^2 .$$

A few remarks on the model and results are in order:

1. *Necessity of declining trades.* By considering a more general setting than in [25], we obtain qualitatively different insights on the optimal control rule by central planner. In

[25], it is optimal for the central planner to always approve trades, and let the agent with fewest scrips be the service provider. In our general setting, however, in many cases the central planner has to decline a non-trivial fraction of the trades to sustain flow balance of scrips in the system (constraint (1.11)).[22] When a trade is approved, our policy also chooses the compatible trade partner with the fewest scrips as service provider.

2. *Incentives.* Our analysis of scrip systems is meant to illustrate the versatility of MBP control policies, hence we only focused on the central planner setting. It would be interesting to study the MBP control rule in the decentralized setting where the agents recommended to be potential trading partners can decide whether to trade, but that is beyond the scope of the current paper. (At a high level, we expect that agents will have an incentive to provide service whenever requested by the MBP policy as long as (i) agents are sufficiently patient, and (ii) agents benefit from trading, i.e., agents derive more value from receiving service than the cost they incur from providing service.)

1.9 Discussion

In this paper we considered the payoff maximizing dynamic control of a closed network of resources. We proposed a novel family of policies called Mirror Backpressure (MBP), which generalize the celebrated backpressure policy such that it executes mirror descent with the desired mirror map, while retaining the simplicity of backpressure. The MBP policy overcomes the challenge stemming from the no-underflow constraint and it does not require any knowledge of demand arrival rates. We proved that MBP achieves good transient performance for demand arrival rates which are stationary or vary slowly over time, losing at most $O\left(\frac{K}{T}+\frac{1}{K}+\sqrt{\eta K}\right)$ payoff per customer, where K is the number of supply units, T is the number of customers over the horizon of interest, and η is the maximum change in demand arrival rates per customer arrival. We considered a variety of

[22]For example, consider a setting with two agents j_1 and j_2. Denote the demand type requested by j_1 as τ_1 (this demand type can be served by j_2) and similarly define τ_2. Under the mild condition $\phi_{\tau_1} \neq \phi_{\tau_2}$, the planner will be forced to decline a positive fraction of requests.

control levers: entry control, assignment control and pricing, and allowed for finite buffer sizes. We discussed the application of our results to the control of shared transporation systems and scrip systems.

One natural question is whether our bounds capture the right scaling of the per customer optimality gap of MBP with K, T and η, relative to the best policy which is given exact demand arrival rates and horizon length T in advance. Consider the joint entry-assignment setting (Section 1.6.2). It is not hard to construct examples showing that each of the terms in our bound is unavoidable: a $1/K$ optimality gap arises in steady state (under stationary demand arrival rates) for instance in a two-node entry-control-only example where the two demand arrival rates are exactly equal to each other, the K/T term arises because over a finite horizon the flow balance constraints need not be satisfied exactly and MBP does not exploit this flexibility fully, and the $\sqrt{\eta K}$ term arises in examples where demand arrival rates oscillate (with a period of order $\sqrt{K/\eta}$) but MBP does not take full advantage of the flexibility to allow queue lengths to oscillate alongside. We omit these examples in the interest of space.

We point out some interesting directions that emerge from our work:

1. *Improved performance via "centering" MBP based on demand arrival rates.* If the optimal shadow prices $\mathbf{y}^*$ are known (or learned by learning $\boldsymbol{\phi}$ via observing demand), we can modify the congestion function to $\tilde{f}_j(\bar{q}_j) = y_j^* + f(\bar{q}_j)$. For the resulting "centered" MBP policy, based on the result of [38] and the convergence of mirror descent, we are optimistic that the steady state regret will decay exponentially in K.

Another promising direction is to pursue the viewpoint that there is an MBP policy which (very nearly) maximizes the steady state rate of payoff generation, specifically for the choice of congestion functions $f_j(\cdot)$ that are the discrete derivatives of the relative value function $F(\bar{\mathbf{q}})$ (for the average payoff maximization dynamic programming problem) with respect to $\bar{q}_j$; see Chapter 7.4 of [41] for background on dynamic programming. Thus, estimates of the relative value function $F(\bar{\mathbf{q}})$ can guide the choice of congestion function.

2. *Other applications of MBP.* MBP appears to be a powerful and general approach to

obtain near optimal performance despite no-underflow constraints in the control of queueing networks. It does not necessitate a heavy traffic assumption, and provides guarantees on both transient and steady state performance, as well as performance under demand arrival rates which vary slowly in time. We conclude with a concrete problem which one may try to address using MBP: The matching queues problem studied by [10] is hard due to no-underflow constraints and to handle them that paper makes stringent assumptions on the network structure. MBP may be able to achieve near optimal performance for more general matching queue systems.

CHAPTER 2

Dynamic Assignment Control of Closed Networks under Complete Resource Pooling

2.1 Introduction

Several real-world systems such as shared transportation platforms and scrip systems involve resource (supply) units circulating in a network. The hallmark of such systems is that serving a demand unit causes a (reusable) supply unit to be relocated. Closed queueing networks provide a powerful abstraction for these applications [see, e.g., 42, 1, 2, 25, 43]. The platform operator makes tactical control decisions with the aim of maximizing longer-term system performance, which necessitates that the operator *manage the distribution of the supply* to ensure continued availability of supply throughout the network. In this paper, we focus on *dynamic assignment control* of a closed queueing network given *limited flexibility*, i.e., when a demand unit arrives at a node, from which compatible (e.g., nearby) node should a supply unit be assigned to serve it?

A central challenge in such systems is that of distributional mismatch between supply and demand: to fulfill a demand which arrives at a node, there has to be an available supply unit at a compatible node when the demand arrives. There are two sources of distributional supply-demand asymmetry: *structural imbalance* (some nodes may have a tendency to have a systematic net inflow, or outflow, of supply units) and *stochasticity*. Previous works have studied assignment (or control) decisions made in a *state-independent* manner which handles structural imbalance by solving the fluid limit problem which arises as the number of supply units K is taken to ∞. However, this approach fails to react to stochasticity leading to optimality gap (fraction of demand lost) which shrinks to zero only (slowly) as $1/K$ [1] as K grows if demand arrival rates are *exactly* known, and non-vanishing optimality gap as $K \to \infty$ if demand arrival rates are not perfectly known (see Proposition 2.3 in Section 2.4.2). In this paper we propose simple and practical *state-dependent* assignment control policies which automatically handle both structural imbalance and stochasticity. Our policies come with a strong performance guarantee and do not require demand arrival rates to be known (if these rates are known, even better performance can be obtained).

We focus on demand arrival rates satisfying an approximate balance condition (very similar to Hall's condition in matching and Complete Resource Pooling in queueing), which ensures that in the absence of stochasticity (i.e., in the fluid limit), all demand can be satisfied. The control problem remains non-trivial: all state-independent policies provide unsatisfactory performance as summarized above (Proposition 2.3), and a naive state-dependent policy similarly suffers $\Omega(1)$ optimality gap as $K \to \infty$ (Example 2.4). We provide a very simple "maximum weight" (MaxWeight) control policy which does not use demand arrival rate information and achieves optimality gap (loss) which decays *exponentially* in K. This result motivates the large deviations question: *Which policy maximizes the loss exponent?* We propose a natural family of Scaled MaxWeight (SMW) policies generalizing MaxWeight, and show that all SMW policies achieve exponentially small loss. We then prove the surprising result that there is always an SMW policy which is *exponent-optimal* among all assignment control policies, and characterize how the parameters of the optimal SMW policy are determined by the demand arrival rates.

Our Model. We adopt a stylized model which isolates the challenge of managing the distribution of (reusable) supply in the network given limited flexibility. (Later, we suitably augment this baseline model to incorporate salient features of specific applications.)

In our model, the system consists of a network with two sets of nodes, namely, the supply nodes and the demand nodes. A fixed number of supply units circulate among the supply nodes. Demand units arrive stochastically at demand nodes with supply node destinations, at some time-invariant rates. For each demand node, a subset of the supply nodes are *compatible* with it, and the platform dynamically decides from which compatible supply node to assign a supply unit to serve the incoming demand unit. Thus, compatibilities capture the *limited flexibility* available to the platform. After a supply unit is assigned to a demand unit, it becomes available again at the destination of the demand unit. (Supply units relocate only while serving demand.) Supply units do not enter or leave the system. The platform's goal is to meet as much demand as possible in steady state. (Our results will extend to transient performance as well.)

Our model assumes that the supply units relocate instantaneously in the process of serving a demand unit. This assumption facilitates a sharp theoretical analysis of general network structures, and moreover ensures transparency about the role of supply units: all K supply units are free when a demand unit arrives, and thus K quantifies the total available "buffer" of free supply units. The controller's challenge is that of managing the distribution of the K supply units to ensure the continued availability of supply throughout the network.

To obtain tight characterizations, we consider the asymptotic regime where the number of supply units in the system K goes to infinity, and perform a large deviations analysis.

Complete Resource Pooling condition. A main assumption in our model is an approximate balance condition on the demand arrival rates. This condition is very similar to the complete resource pooling (CRP) condition in the queueing literature, therefore we will refer to it as CRP hereafter. CRP is a standard assumption in the heavy traffic analysis of queueing systems [see, e.g., 44, 17, 36]. It can be interpreted as requiring enough overlap in the processing ability of servers (demand nodes in our model) so that they form a "pooled server". The CRP condition under our model is closely related to the condition in Hall's marriage theorem in bipartite matching theory. If any CRP inequality is strictly violated, this forces a positive fraction of demand to be lost even as $K \to \infty$.

Analogy with a classic closed queueing network scheduling problem. Using the terminology of classic queueing theory, the K supply units are "jobs", each demand location is a "server", each supply location is a "buffer", inter-arrival times of demand units with origin i are "service times" at server i. The distribution of demand destinations given an origin node captures "routing probabilities". "Servers" are flexible (i.e., they can serve multiple queues), and assignment is equivalent to "scheduling". We emphasize the reversal of the usual mapping: in our setup supply units are "jobs" and demand units act as service tokens. As a consequence, intuition based on traditional queueing systems does not easily extend to our setup.

2.1.1 Main Contributions

We show that a simple and practical MaxWeight assignment policy effectively manages the distribution of supply in the network, leading to a fraction of demand lost that decays exponentially fast in K. Each time a demand arrives, MaxWeight simply assigns a supply unit from the compatible node which currently has the largest number of supply units. In particular, MaxWeight requires no knowledge of demand arrival rates.

This finding motivates a thorough *large deviations analysis* which yields surprisingly elegant results. As a function of system primitives, we derive a large deviations rate-optimal assignment policy that minimizes lost demand. Our optimal policy is a close cousin of MaxWeight and its parameters depend in a natural way on demand arrival rates. Our contribution is threefold:

1. **A family of simple policies.** We propose a family of state-dependent assignment policies called Scaled MaxWeight (SMW) policies, and prove that all of them guarantee exponential decay of demand-loss probability under the CRP condition. An SMW policy is parameterized by a vector of scaling factors, one for each (supply) node; each demand is served by assigning a supply from the compatible node with the

largest scaled number of supply units. SMW policies are simple, explicit and promising for practical applications (Section 2.6.2 and Appendix B.10 demonstrate stellar performance in a realistic simulation environment).

2. **The value of (intelligent) state-dependent control.** We show (Proposition 2.3) that no state-independent assignment policy can achieve loss which decays exponentially in K, and that if demand arrival rates are not perfectly known, then the loss of a state-independent policy (generically) does not vanish as $K \to \infty$. Also, a naive state-dependent control policy suffers $\Omega(1)$ loss as $K \to \infty$ (Example 2.4). Our SMW policies provide vastly superior performance: even the naive unscaled ("vanilla") MaxWeight assignment policy requiring no knowledge of demand arrival rates achieves loss which decays exponentially in K.

3. **Exponent-optimal policy and qualitative insights.** For general network structures, we obtain an explicit specification for the optimal scaling factors for SMW based on compatibilities and demand arrival rates. Further, we obtain the surprising finding that the optimal SMW policy is, in fact, *exponent-optimal* among all state-dependent policies (Theorem 2.1). A key ingredient of this result is that SMW policies satisfy the *critical subset* property: for each SMW policy, there is a corresponding (fluid) equilibrium state, and for this state there are "critical" subsets of demand nodes that are most vulnerable to the depletion of supply in compatible supply nodes. Each SMW policy simultaneously "protects" all critical subsets maximally by maintaining high supply levels near structurally under-supplied nodes.

We consider the natural "large market" scaling where the demand arrival rate is proportional to K, and show that each supply unit is frequently in use.

Technical contributions. To the best of our knowledge, we are the first to perform a large deviations analysis under CRP, leading to the challenging problem of deriving an exponent optimal control. One key difficulty in the mathematical analysis is the necessity to deal with a multi-dimensional system even in the limit. Usually CRP renders the control problem "easy" because it leads to the "collapse" of the system state to a lower dimensional space in the heavy traffic limit, as in many existing works that establish the asymptotic optimality of a certain policy in minimizing the workload/holding costs of a queueing system. In contrast, in our setting, the limit system remains m-dimensional, where m is the number of supply nodes. A second key challenge we face is that the ideal state for the system is a priori unknown, making it unclear how to define a Lyapunov function. We overcome these difficulties via a novel approach. We construct a *policy-specific* Lyapunov function to facilitate a sharp large deviations analysis of a given SMW policy leveraging the machinery of [45]. The analysis applies to general network structures, and reveals that the SMW policy maximally protects all the "critical subsets" of demand nodes. We deduce the existence of an exponent optimal SMW policy, and characterize its scaling factors in terms of demand arrival rates. Happily, the fluid equilibrium for this optimal policy is revealed as the ideal state.

Though our setting considers a closed network, we think that it could inspire similar analyses in open networks, e.g., when there is a shared finite buffer (e.g., a common waiting room) for multiple queues. Our technical machinery may also be broadly useful in deriving large-deviation optimal controls in settings where the ideal state is a priori unclear.

2.1.2 Applications

Our main model and analysis can serve as a building block towards studying various applications. We discuss two broad applications later in the paper.

Shared transportation systems. Shared transportation platforms such as those for ride-hailing and bikesharing make assignment control decisions under limited flexibility to manage the distribution of supply. In these applications, the nodes in our model correspond to geographical locations,[1] while supply units and demand units correspond to vehicles and customers, respectively. The assignment control in ride-hailing takes the form of dispatch, i.e., the platform can decide where (near the demand's origin) to

[1]The set of supply nodes and demand nodes are replicas of each other in these applications.

dispatch a car from. Bikesharing platforms can execute assignment control by suggesting to the customer where (near the customer's origin or destination) to pick up (or drop off) a bike.[2]

We discuss the application to shared transportation systems in Section 2.6. Transportation involves positive travel times. We incorporate travel times into our theory and show that SMW policies retain their good performance, and also demonstrate excellent performance in realistic simulations of ridehailing:

(i) We extend our theory by letting demand have independent exponential travel times with mean that can depend on the origin-destination pair, and assume zero pickup times. We consider the large market scaling and assume that the total service requirement (the average number of demands in service at any time assuming no lost demand) is a fraction of supply which is strictly below 1, consistent with the reality in shared transportation. We prove that for any SMW policy, the loss is again exponentially small in K.

(ii) We demonstrate excellent performance of SMW policies in simulations of ridehailing based on the NYC taxi dataset. We propose data-driven approaches for "learning" SMW scaling factors via simulations, and observe close alignment of the resulting SMW scaling factors with those suggested by our theoretical analysis.

We also describe how state-independent "empty" relocation of vehicles can be seamlessly incorporated in our setup.

Scrip systems. A scrip system is a nonmonetary trade economy where agents use scrips (tokens, coupons, artificial currency) to exchange services. These systems are typically implemented when monetary transfer is undesirable or impractical. For example, [26] suggest that in kidney exchange, to align the incentives of hospitals, the exchange should deploy a scrip system that awards points to hospitals that submit donor-patient pairs to the central exchange, and deducts points from hospitals that conduct transplan-

[2]For example, the Bike Angels program of CitiBike implicitly makes these suggestions to members by awarding "points for taking bikes from crowded stations and bringing them to empty ones or stations expected to soon become empty". Notice the resemblance to a MaxWeight approach. A live map of point awards is shown to customers.

tations. Another well-known example is Capitol Hill Babysitting Co-op ([40], see also [25]), where married couples pay for babysitting services by another couples with scrips. A key challenge in these markets is the design of the service provider selection rule: among the possible providers for a requested service/trade, who should be selected for service? The platform operator tries to minimize discarded requests (which happen when the service requester runs out of scrips) by choosing this rule appropriately. We will show in Section 2.7 that with only cosmetic modifications to the setup, our results translate fully to a model of scrip systems; in particular we derive exponent-optimal control policies for these systems.

2.1.3 Literature Review

MaxWeight scheduling. MaxWeight is a simple scheduling policy in constrained queueing networks which (roughly speaking) chooses the feasible control decision that serves the queues with largest total weight (e.g. queue length, head-of-line waiting time, etc.), at each time. MaxWeight scheduling has been shown to exhibit good performance in various settings (see, e.g., [4, 7, 16, 17, 18, 6]), including by [36] who study an open one-hop network version of our setting. In contrast, we find that MaxWeight achieves a suboptimal exponent in our closed network setting.

Large deviations in queueing systems. There is a large literature on characterizing the probability of building up long queues in *open* queueing networks, including controlled [see, e.g., 46, 47] and uncontrolled [see, e.g., 48, 49] networks. The work closest to ours is that of [45], who established the relationship between Lyapunov functions and buffer overflow probability for open queueing networks. The key difficulty in extending the Lyapunov approach to closed queueing networks is the lack of a natural reference state where the Lyapunov function equals to 0 (in an open queueing network the reference state is simply $\mathbf{0}$). It turns out that as we optimize the MaxWeight parameters we are also solving for the best reference state.

Applications: shared transportation systems, scrip systems. [21] studied revenue-maximizing state-independent assignment control by solving a minimum cost flow problem in the fluid limit. [2] modeled the system by a closed queueing network and derived the optimal static routing policy that sends empty vehicles to under-supplied locations. [1] adopted the Gordon-Newell closed queueing network model and considered static pricing/repositioning/matching policies that maximizes throughput, welfare or revenue. In contrast to our work, which studies state-dependent control, these works consider static control that completely relies on system parameters. In terms of convergence rate to the fluid-based solution, [21] did not study the convergence rate of their policy, [2] observed from simulation an $O(1/\sqrt{K})$ convergence rate as the number of supply units in the closed system K goes to infinity,[3] while [1] showed finite system bounds with an $O(1/K)$ convergence rate as $K \to \infty$ in the absence of service times and an $O(1/\sqrt{K})$ convergence rate with service times. All these works propose static policies, and we show that no static policy can achieve exponentially small loss. In contrast, under the CRP condition, we obtain exponentially small loss in K, and further obtain the optimal exponent.

Our approach of studying control while initially ignoring travel delays is mirrored in several papers in this literature, starting with [42]. The main model in [1] ignores travel delays, and the paper subsequently shows that all its findings are robust to that assumption. Similarly, subsequent to the present paper, [22] study the control of (large) networks of circulating resources by ignoring travel delays and then show robustness of their results to delays.

There have been a few papers that model and analyze scrip systems, e.g., [50, 51, 25, 43] etc. The closest paper to ours is [25], which considers the case where the compatibility graph is fully connected and the demand arrival rates are identical for each demand type. They propose a service selection rule which is the same as the vanilla version of

[3]In the setting of [2], the loss probability can remain positive even as K grows, in contrast with our setting where the loss probability can always be sent to 0 because of our CRP condition under which the flows in the network can potentially be balanced. The comparison of convergence rates is most meaningful if we restrict attention to instances in their setting where the loss probability goes to zero as K grows.

our proposed policy and show that it is optimal in their symmetric setting. We significantly generalize their model by considering asymmetric demand arrivals and general skill compatibility graphs. For other examples of scrip systems, see, e.g., [40, 26], etc.

Online stochastic bipartite matching. There is a related stream of research on online stochastic bipartite matching, see, e.g., [27, 28, 29, 30]. Different types of supplies and demands arrive over time, and the system manager matches supplies with demands of compatible types using a specific matching policy, and then discharges the matched pairs from the system. Our work is different in that we study a *closed* system where supply units never enter or leave the system. Moreover, this literature focuses on the stability and other properties under a given policy instead of looking for the optimal control [except 29].

Other related work. [34, 35, 36] and others study how process flexibility can facilitate improved performance, analogous to our use of dispatch control to improve demand fulfillment. Along similar lines, network revenue management is a classical dynamic resource allocation problem, see, e.g., Gallego and Van Ryzin [31] and Talluri and Van Ryzin [32], and recent works, e.g., Jasin and Kumar [52] and Bumpensanti and Wang [33]. Different types of demands arrive over time, and a centralized decision is made at each arrival. Again, each of these settings is "open" in that each service token or supply unit can be used only once, in contrast to our closed setting.

2.1.4 Organization of the paper

The remainder of our paper is organized as follows. In Section 2.2 we introduce the basic notation and formally describe our baseline model together with the performance metric. In Section 2.3 we introduce the family of Scaled MaxWeight policies. In Section 2.4 we present our main theoretical result, i.e., that there is an exponent optimal SMW policy for any set of primitives satisfying our main assumption. In Section 2.5 we prove the exponent optimality of SMW policies. In Section 2.6 we discuss the application to shared transportation systems. In Section 2.7 we discuss the application to scrip systems.

We conclude in Section 2.8.

Notation. We use $\mathbf{e}_i$ to denote the i-th unit vector, and $\mathbf{1}$ the all-1 vector. The dimensions of the vectors will be clear from the context. For a finite index set A, define $\mathbf{1}_A \triangleq \sum_{i\in A}\mathbf{e}_i$. For a set Ω in Euclidean space $\mathbb{R}^n$, denote its relative interior by $\mathrm{relint}(\Omega)$. For event C, we define the indicator random variable $\mathbb{I}\{C\}$ to equal 1 when C is true, else 0. All vectors are column vectors if not specified otherwise.

2.2 The Model and Preliminaries

2.2.1 Basic Setting

We study the dynamic assignment problem in networks with circulating resources. We consider an infinite-horizon continuous-time model, with a fixed number K of identical *supply units* that circulate in the network. Formally, we consider a sequence of systems indexed by $K \in \mathbb{Z}_+$.

The (Assignment) Compatibility Graph. The assignment compatibility structure is described by a bipartite *compatibility graph* $G = (V_S \cup V_D, E)$, where the K supply units are distributed over the *supply nodes* V_S, and demand units arrive at the *demand nodes*[4] V_D. We add a prime symbol to the indices of nodes in V_D to distinguish between the two. Let $m \triangleq |V_S|$ and $n \triangleq |V_D| \in \mathbb{Z}_+$ be the number of supply and demand nodes, respectively. Each edge $(i, j') \in E$ represents a compatible pair of supply and demand nodes, i.e., a supply unit currently stationed at $i \in V_S$ can serve demand arriving at $j' \in V_D$. See Figure 2.1 for an illustration. We denote the neighborhood of a supply node $i \in V_S$ (resp. demand node $j' \in V_D$) in G as $\partial(i) \subseteq V_D$ (resp. $\partial(j') \subseteq V_S$). Moreover, for any set of supply nodes $A \subseteq V_S$, we also use $\partial(A)$ to denote its demand neighborhood (and vice versa).

Demand Types and Arrival Process. We denote the type of a demand as $(j', k) \in$

[4]The physical meaning of the nodes depends on the application. For example, in ride-hailing the supply nodes and demand nodes are replicas of each other and both stand for physical locations. However, our result does not require the symmetry between these two sets of nodes.

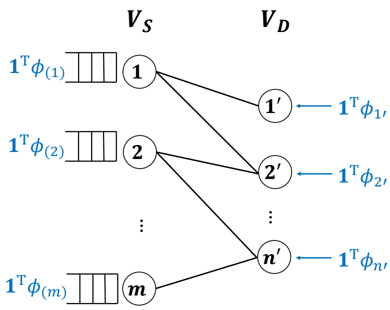

Figure 2.1: The bipartite (assignment) compatibility graph: On the left are supply nodes $i \in V_S$, and on the right are demand nodes $j' \in V_D$. The edges entering a demand node j' encode compatible (e.g., nearby) supply nodes that can serve node j'. The (normalized) rate of arrival of demand with origin j' is $\mathbf{1}^{\mathrm{T}}\phi_{j'}$. Assuming no demand is lost, the (normalized) rate of arrival of supply units to i is $\mathbf{1}^{\mathrm{T}}\phi_{(i)}$ (this is the normalized arrival rate of demand with destination i).

$V_D \times V_S$, where j' is its *origin* node and k is its *destination* node. Demand units of each type (j', k) arrive sequentially following independent Poisson processes with rates $\hat{\phi}^K_{j'k}$. We use $\hat{\boldsymbol{\phi}}^K$ to denote the $n \times m$ matrix of demand arrival rates.

We will consider the asymptotic regime where both the number of supply units K and demand arrival rates $\hat{\boldsymbol{\phi}}^K \triangleq K\hat{\boldsymbol{\phi}}$ (for some $\hat{\boldsymbol{\phi}}$ which does not depend on K) go to infinity together. We call this scaling the *large market regime*. We will later show that the large market scaling ensures that each supply unit waits an $O(1)$ amount of time in expectation between two consecutive assignments under the family of policies we prescribe (see Section 2.4).

The *demand type distribution* is $\boldsymbol{\phi} \triangleq \frac{\hat{\boldsymbol{\phi}}}{\mathbf{1}^{\mathrm{T}}\hat{\boldsymbol{\phi}}\mathbf{1}}$, which is the normalized version of $\hat{\boldsymbol{\phi}}$. We will find it convenient to carry out our technical development and analysis in terms of $\boldsymbol{\phi} \in \mathbb{R}^{n\times m}$ instead of $\hat{\boldsymbol{\phi}}$ wherever the total arrival rate $\mathbf{1}^{\mathrm{T}}\hat{\boldsymbol{\phi}}\mathbf{1}$ does not play a role. We denote the k-th column of $\boldsymbol{\phi}$ (i.e., the normalized arrival rates at different origins of demands with destination k) as $\phi_{(k)}$, and the transpose of the j'-th row of $\boldsymbol{\phi}$ (i.e., the normalized arrival rates of demands with origin j' and different destination nodes) as $\phi_{j'}$. Thus, the (normalized) rate of a demand units arriving at node j' is $\mathbf{1}^{\mathrm{T}}\phi_{j'}$, and, assuming all demands are matched, the (normalized) rate of supply units arriving at node k is $\mathbf{1}^{\mathrm{T}}\phi_{(k)}$. Without loss of generality, we exclude demand nodes with zero demand arrival rate from V_D.

We use the term *network* to refer to a given set of primitives: an assignment compatibility graph G and demand type distribution matrix $\boldsymbol{\phi}$. We make two mild assumptions on the network.

Assumption 2.1 (Connectedness). *A network* $(G, \boldsymbol{\phi})$ *is* connected *if for every ordered pair of distinct supply nodes* $(k_0, i) \in V_S \times V_S$, $k_0 \neq i$, *there is a finite sequence of demand types* $(j'_1, k_1), \cdots, (j'_\ell, k_\ell = i)$ *such that* $\phi_{j'_r k_r} > 0$ *for all* $r = 1, \cdots, \ell$, *and* $k_{r-1} \in \partial(j'_r)$ *for all* $r = 1, \cdots, \ell$.

Assumption 2.1 requires that for every pair of supply nodes, there is a sequence of demand types with positive arrival rates and corresponding compatible supply nodes that would take a supply unit from one node eventually to the other node.

We now observe that if the compatibility graph affords ample flexibility, specifically, if the destination for every demand type belongs to the compatible neighborhood of the origin, then the control problem is trivial. The reason is simple: we can "reserve" a supply unit for each demand origin node $j' \in V_D$, and each reserved supply unit will never leave the corresponding neighborhood $\partial(j')$, ensuring that no demand is ever lost. (We formalize this observation in Appendix B.6.) This motivates the following assumption to ensure that the flexibility available is sufficiently limited that the assignment control problem at hand is non-trivial.

Assumption 2.2 (Limited flexibility). *A network* $(G, \boldsymbol{\phi})$ *has* limited flexibility *if there exists an origin-destination pair* $j' \in V_D$ *and* $k \in V_S$ *such that* $k \notin \partial(j')$ *and* $\phi_{j'k} > 0$, *i.e., the destination* k *for these demand units is not a supply node compatible with their origin* j'.

Simplifying assumptions regarding relocation of supply. We make the simplifying assumptions that the relocation of a supply unit upon serving a demand is instantaneous, and that a supply unit does not move unless assigned. These assumptions parallel that in an emerging line of works studying control of systems with circulating resources, e.g. [1, 22]. The assumptions keep the state space manageable while retaining the complex supply externalities between nodes (namely, serving a demand redistributes the supply by causing a supply unit to relocate to a specific destination), which is the key challenge that we focus on. We relax the instantaneous relocation assumption in Section 2.6.1 and in Section 2.6.2 (simulations) and show that our insights are robust to this assumption. In Section 2.8 we observe that "empty" relocation (as may occur in ride-hailing) which is state independent can be seamlessly integrated into our framework.

System State. For the K-th system, its state at any time is given by $\mathbf{X}^K$, an m-dimensional vector that tracks the number of supply units at each supply node. The state space of the K-th system is thus given by $\Omega_K \triangleq \left\{\mathbf{x} \in \{0, 1, 2, \dots\}^m \,\middle|\, \mathbf{1}^{\mathrm{T}}\mathbf{x} = K\right\}$. Note that the normalized state $\frac{1}{K}\mathbf{X}^K$ lies in the m-probability simplex $\Omega = \{\mathbf{x} \in \mathbb{R}^m | \mathbf{x} \geq 0, \mathbf{1}^{\mathrm{T}}\mathbf{x} = 1\}$. We use $\mathbf{X}^K(0)$ to denote the initial state.

2.2.2 Optimal Assignment Control

Given the above setting, the problem we want to study is how to design assignment policies which minimize the probability of losing demand. For fixed K, this problem can be formulated as an average cost Markov decision process on a finite (albeit, very large) state space, and is thus known to admit a stationary optimal policy (i.e., where the assignment rule at any time only depends on the current system state $\mathbf{X}^K$; see Proposition 5.1.3 in [41]).

Assignment policies. Upon the arrival of an incoming demand of type (j', k), the platform must immediately assign a supply unit from a compatible node of j'; subsequently, after serving the demand, the supply unit becomes available at the destination node k. If no supply unit is available at any compatible node of j', then we experience a *demand loss*, wherein the demand unit leaves the system without being served. Let $\mathcal{U}^K$ be the set of stationary policies for the K-th system. An assignment policy $U \in \mathcal{U}$ consists of, for each $j' \in V_D$, $k \in V_S$, a sequence of mappings $\left(U^K \in \mathcal{U}^K\right)_{K=1}^{\infty}$, which map the current queue-length vector $\mathbf{X}^K$ and demand type (j', k) to $U^K[\mathbf{X}^K](j', k) \in \partial(j') \cup \{\emptyset\}$.

Here $U^K[\mathbf{X}^K](j',k) = i$ means given the current state $\mathbf{X}^K$, we assign a supply unit from $i \in \partial(j')$ to fulfill demand with origin j' and destination k, and $U^K[\mathbf{X}^K](j',k) = \emptyset$ means that the platform does not assign supplies to type (j',k) demands and hence any such demand is lost. When $\mathbf{X}_i^K = 0$ for all $i \in \partial(j')$, this forces $U^K[\mathbf{X}^K](j',k) = \emptyset$ since there is no supply at nodes compatible to j'. For simplicity of notation, we refer to the policies by U instead of U^K.

System Evolution. Let t_r be the r-th demand arrival epoch after time 0. Denote the state of the system just before t_r by $\mathbf{X}^K(t_r^-)$ (the initial state is $\mathbf{X}^K(0)$); note that this incorporates the state change due to serving the $(r-1)$-th demand arrival for $r > 1$. Now suppose the platform uses an assignment policy U, and the r-th demand arrival has origin node $o[r]$ with destination $d[r]$ (sampled from demand type distribution $\boldsymbol{\phi}$). Let $S[r] \triangleq U^K[\mathbf{X}^K(t_r^-)](o[r], d[r])$ be the chosen supply node (potentially $\emptyset$). Then, formally, the system state updates as per

$$\mathbf{X}^K(t_r) \triangleq \begin{cases} \mathbf{X}^K(t_r^-) - \mathbf{e}_{S[r]} + \mathbf{e}_{d[r]} & \text{if } S[r] \in V_S\,, \\ \mathbf{X}^K(t_r^-) & \text{if } S[r] = \emptyset\,. \end{cases}$$

Performance Measure. The platform's goal is to find an assignment policy that loses as few demands as possible in steady state. A natural performance measure is the *long-run average demand-loss probability*. Formally, for $U \in \mathcal{U}$ we define

$$\mathbb{P}_{\mathrm{o}}^{K,U} \triangleq \min_{\mathbf{X}^{K,U}(0)\in\Omega_K} \mathbb{E}\left(\lim_{T\to\infty} \frac{1}{T}\sum_{r=1}^{T} \mathbb{I}\left\{U^K[\mathbf{X}^{K,U}(t_r^-)](o[r],d[r]) = \emptyset\right\}\right)\,, \tag{2.1}$$

$$\mathbb{P}_{\mathrm{p}}^{K,U} \triangleq \max_{\mathbf{X}^{K,U}(0)\in\Omega_K} \mathbb{E}\left(\lim_{T\to\infty} \frac{1}{T}\sum_{r=1}^{T} \mathbb{I}\left\{U^K[\mathbf{X}^{K,U}(t_r^-)](o[r],d[r]) = \emptyset\right\}\right)\,. \tag{2.2}$$

Here (2.1) is an *optimistic* (subscript "o" for optimistic) performance measure (which underestimates demand-loss probability), whereas (2.2) is a *pessimistic* (subscript "p" for pessimistic) performance measure (which overestimates demand-loss probability). Since $U \in \mathcal{U}$ is a stationary policy, the limits in (2.1) and (2.2) exist. Note that $\mathbb{P}_{\mathrm{o}}^{K,U} \leq \mathbb{P}_{\mathrm{p}}^{K,U}$. We will establish the exponent optimality of our policy by showing that its pessimistic measure decays as fast with K as any policy's optimistic measure can possibly decay.

The exact values of (2.1) and (2.2) for fixed K are challenging to study. To this end, the main performance measures of interest in this work are the decay rates of $\mathbb{P}_{\mathrm{o}}^{K,U}$ and $\mathbb{P}_{\mathrm{p}}^{K,U}$ as $K \to \infty$:

$$\gamma_{\mathrm{o}}(U) \triangleq -\liminf_{K\to\infty} \frac{1}{K} \log \mathbb{P}_{\mathrm{o}}^{K,U} , \tag{2.3}$$

$$\gamma_{\mathrm{p}}(U) \triangleq -\limsup_{K\to\infty} \frac{1}{K} \log \mathbb{P}_{\mathrm{p}}^{K,U} . \tag{2.4}$$

For brevity, we henceforth refer to these as the *demand-loss exponents.* Note that $\gamma_{\mathrm{o}}(U) \geq \gamma_{\mathrm{p}}(U)$. The definition (2.3) uses lim inf so that we can state a strong converse result by upper bounding $\sup_{U\in\mathcal{U}} \gamma_{\mathrm{o}}(U)$, since no policy can achieve a larger demand-loss exponent. Similarly, the definition (2.4) uses lim sup so that we can state a strong achievability result (for our proposed policies the limit will exist; when the limit exists we write $\gamma(U) \triangleq \gamma_{\mathrm{o}}(U) = \gamma_{\mathrm{p}}(U)$).

2.2.3 The Complete Resource Pooling (CRP) Condition

We now make a few additional definitions to allow us to state our main assumption.

We say that a subset of demand nodes $J \subsetneq V_D$ has *limited flexibility* if there is some demand node $j' \in J$ and supply node $k \notin \partial(J)$ such that $\phi_{j'k} > 0$. (Informally, there is a demand type which requires supply units to leave the neighborhood of J.) We denote the set of limited-flexibility subsets by $\mathcal{J}$. Assumption 2.2 guarantees that there is at least one non-trivial singleton J and hence that $\mathcal{J} \neq \emptyset$.

Observe that J has limited flexibility if and only if

$$\mu_J \triangleq \sum_{j'\in J} \sum_{k\notin\partial(J)} \phi_{j'k} > 0 . \tag{2.5}$$

We call μ_J the *net demand* of J, since it captures the probability that a demand arrival has origin in J and destination outside $\partial(J)$ (and hence requires a supply unit to leave

$\partial(J)$). Similarly, we define the (optimistic) *net supply* to J as

$$\lambda_J \triangleq \sum_{j' \notin J} \sum_{k \in \partial(J)} \phi_{j'k} \,. \tag{2.6}$$

Informally, λ_J the probability that a demand arrival is such that it can (depending on the assignment decision) cause a supply unit to enter $\partial(J)$.

The following is the main assumption of this paper.

Assumption 2.3 (Complete Resource Pooling)**.** *We assume that for all subsets of demand nodes J with limited flexibility (i.e., $J \subsetneq V_D$ with positive net demand $\mu_J > 0$) we have that $\lambda_J > \mu_J$, where the net supply λ_J was defined in* (2.6)*, and the net demand μ_J was defined in* (2.5)*.*

The intuition behind this assumption is simple: it assumes the system is "balanceable" in that for each subset $J \subsetneq V_D$ of demand nodes, supply arrives sufficiently fast at neighboring nodes to meet the demand arriving to J, on average. Assumption 2.3 is equivalent to a strict version of the condition in Hall's marriage theorem. It is also closely related to the Complete Resource Pooling (CRP) condition in queueing: we show (formalized in Proposition B.2 that in Appendix B.9) if the "open queueing network counterpart" of network $(G, \hat{\boldsymbol{\phi}})$ satisfies the CRP condition defined in [17], then the network $(G, \hat{\boldsymbol{\phi}})$ satisfies Assumption 2.3. The control problem under CRP is non-trivial: In Section 2.4.2 we will show that all state-independent policies and a naive state-dependent policy perform inadequately.

We remark that the condition $\lambda_J > \mu_J$ is equivalent to $\sum_{i \in \partial(J)} \mathbf{1}^{\mathrm{T}} \phi_{(i)} > \sum_{j' \in J} \mathbf{1}^{\mathrm{T}} \phi_{j'}$ (informally, that the total supply to J exceeds total demand of J), but the representation $\lambda_J > \mu_J$ will turn out to be more closely related to our analysis and our main theorem. We will find that the limited-flexibility subsets J with ratio λ_J/μ_J close to 1, i.e., only a small excess of supply over demand, will be pivotal in determining the performance of our policies and optimal policy design. We illustrate the quantities involved ($\mathcal{J}$, λ_J and μ_J) and their impact on policy performance and design via an example at the end of the next section (Example 2.1).

We show that Assumption 2.3 is necessary in order to obtain exponentially small loss in Proposition 2.1.

Proposition 2.1. *For any G and $\boldsymbol{\phi}$'s such that Assumption 2.3 is violated, it holds that for any policy U, the demand loss probability does not decay exponentially,*[5] *i.e.,* $\gamma_{\mathrm{o}}(U) = \gamma_{\mathrm{p}}(U) = 0$ *where* $\gamma_{\mathrm{o}}(U)$ *and* $\gamma_{\mathrm{p}}(U)$ *are defined in* (2.3) *and* (2.4).

In other words, if Assumption 2.3 is violated, this means the system has significant distributional imbalance of demand and demand loss is unavoidable. The intuition is similar to that of Hall's marriage theorem [53]: if there is a limited-flexibility subset J with net supply (weakly) less than the net demand, then it is impossible for any policy to ensure that all but an exponentially small fraction of demand originating in J is served. The proof of Proposition 2.1 is in Appendix B.6.

2.2.4 Sample Path Large Deviation Principle

Our main theoretical result is the culmination of a sharp *large deviations* analysis, characterizing the best possible demand loss exponent. We provide a brief introduction to classical large deviations theory in this subsection.

For each fixed $K \in \mathbb{Z}_+$ and $T \in (0, \infty)$, define a *scaled sample path* of accumulated demand arrivals $\bar{\mathbf{A}}^K(\cdot) \in (L^\infty[0,T])^{n\times m}$ as follows.[6] Let $\{A^K_{j'k}(\cdot)\}_{j'\in V_D, k\in V_S}$ be independent Poisson processes where $A^K_{j'k}(\cdot)$ has rate $K\hat{\phi}_{j'k}$. Let

$$\bar{\mathbf{A}}^K_{j'k}(t) \triangleq \frac{1}{K}\mathbf{A}^K_{j'k}(t) \qquad \forall t \in [0,T]\,. \tag{2.7}$$

Let μ_K be the law of $\bar{\mathbf{A}}^K(\cdot)$ in $(L^\infty[0,T])^{n\times m}$. For all $\mathbf{f} \in \mathbb{R}^{n\times m}_+$, let

$$\Lambda^*(\mathbf{f}) \triangleq \begin{cases} \sum_{j'\in V_D}\sum_{k\in V_S}\left(f_{j'k}\log\frac{f_{j'k}}{\hat{\phi}_{j'k}} - f_{j'k} + \hat{\phi}_{j'k}\right) & \text{if } \mathbf{f} > \mathbf{0}\,, \\ \infty & \text{otherwise}\,. \end{cases} \tag{2.8}$$

[5] If the inequality in Assumption 2.3 is strictly *reversed* for some $J \subsetneq V_D$, i.e., $\lambda_J < \mu_J$ then we have a demand loss probability which is at least $\epsilon > 0$ for all K, where $\epsilon = \sum_{j'\in J}\mathbf{1}^{\mathrm{T}}\phi_{j'} - \sum_{i\in\partial(J)}\mathbf{1}^{\mathrm{T}}\phi_{(i)}$.

[6] Here $L^\infty[0,T]$ denotes the space of bounded functions on $[0,T]$ equipped with the supremum norm.

For any set Γ, let $\bar{\Gamma}$ denote its closure, and Γ^o denote its interior. Below is the sample path large deviation principle (also known as Mogulskii's Theorem, see [54]):

Fact 2.1. *For measures $\{\mu_K\}$ defined above, and any arbitrary measurable set $\Gamma \subseteq (L^\infty[0,T])^{n\times m}$, we have*

$$-\inf_{\bar{\mathbf{A}}\in\Gamma^o} I_T(\bar{\mathbf{A}}) \le \liminf_{K\to\infty} \frac{1}{K}\log\mu_K(\Gamma) \le \limsup_{K\to\infty}\frac{1}{K}\log\mu_K(\Gamma) \le -\inf_{\bar{\mathbf{A}}\in\bar{\Gamma}} I_T(\bar{\mathbf{A}}), \qquad (2.9)$$

where the rate function[7] *is:*

$$I_T(\bar{\mathbf{A}}) \triangleq \begin{cases} \int_0^T \Lambda^*\left(\dot{\bar{\mathbf{A}}}(t)\right) dt & \text{if } \bar{\mathbf{A}}(\cdot)\in \mathrm{AC}[0,T],\ \bar{\mathbf{A}}(0)=\mathbf{0}, \\ \infty & \text{otherwise}. \end{cases} \qquad (2.10)$$

Here $\mathrm{AC}[0,T]$ is the space of absolutely continuous functions on $[0,T]$, and $\dot{\bar{\mathbf{A}}}(t)$ is the derivative of $\bar{\mathbf{A}}$ at time t when the derivative exists.

Informally, this fact says the following. (Suppose the leftmost term and rightmost term in (2.9) are equal.) The probability exponent (with respect to K) for the event Γ is equal to the exponent for the most likely fluid sample path (a limit of scaled sample paths, see Section 2.5.1) of demand $\bar{\mathbf{A}}$ such that the event occurs. The exponent for $\bar{\mathbf{A}}$ is the *time integral of the exponent for its time derivative*, and the latter is given by the function (2.8) where the summand is the large deviations exponent of a (sequence of) Poisson random variable(s) with mean $\hat{\phi}_{j'k}$.

In the present work, the relevant Γ will be the demand-loss event. The reason the sample paths of accumulated demand arrivals fully determine whether this event occurs is because given any deterministic policy (as the policies we propose will be), the arrival process $\mathbf{A}(\cdot)$ and the initial configuration $\mathbf{X}(0)$ uniquely determine the evolution of the system state $\mathbf{X}(\cdot)$, and hence determine demand loss. The key will be to understand the most likely sample paths of the arrival process which lead to demand loss. Our converse (impossibility) bound on the exponent will be established by constructing a fluid sample path of demand arrivals that *always* leads to demand loss regardless of the policy.

[7] Since absolutely continuous functions are differentiable almost everywhere, the rate function is well-defined.

2.3 Scaled MaxWeight Policies

The traditional MaxWeight policy is a celebrated approach to scheduling which has been effectively deployed in many applications such as cloud computing, communication networks, traffic management, etc., [see, e.g., 4, 55]. MaxWeight (hereafter referred to as vanilla MaxWeight) allocates the service capacity to the queue(s) with largest "weight" (where weight can be any relevant parameter such as queue length, head-of-the-line waiting time, etc.). In our setting, supply units form queues and demand is like service tokens, and vanilla MaxWeight would correspond to assigning from the compatible supply node with most supply units (with appropriate tie-breaking rules).

Besides its simplicity, one reason for the popularity of MaxWeight is that it is known to be asymptotically optimal in many problem settings (e.g., see [47, 16, 36, 6]). In our setting too, we will find that vanilla MaxWeight is asymptotically optimal. In fact, we will show that it achieves an exponentially small loss. However, we will find that, in general, vanilla MaxWeight does not achieve the largest possible loss exponent. (We will provide a concrete example at the end of this section.) Suboptimality of the exponent prompts us to consider alternate control policies.

We generalize vanilla MaxWeight by attaching a positive scaling parameter α_i to each queue $i \in V_S$, and assign from the compatible queue with largest *scaled* queue length $\mathbf{X}_i/\alpha_i$. Without loss of generality, we normalize $\boldsymbol{\alpha}$ s.t. $\mathbf{1}^{\mathrm{T}}\boldsymbol{\alpha} = 1$, or equivalently, $\boldsymbol{\alpha} \in \mathrm{relint}(\Omega)$. We call this family of policies *Scaled MaxWeight (*SMW*) policies*, and use SMW($\boldsymbol{\alpha}$) to denote SMW with parameter $\boldsymbol{\alpha}$.

The formal definition of SMW is as follows.

Definition 2.1 (Scaled MaxWeight SMW($\boldsymbol{\alpha}$)). *Fix $\boldsymbol{\alpha} \in \mathrm{relint}(\Omega)$, i.e., $\boldsymbol{\alpha} \in \mathbb{R}^m$ such that $\alpha_i > 0 \ \forall i \in V_S$ and $\sum_{i \in V_S} \alpha_i = 1$. Given system state $\mathbf{X}(t_r^-)$ just before the r-th demand arrival and for demand arriving at demand node j',* SMW*($\boldsymbol{\alpha}$) assigns from*

$$\mathrm{argmax}_{i \in \partial(j')} \frac{\mathbf{X}_i(t_r^-)}{\alpha_i}$$

if $\max_{i\in\partial(j')} \frac{\mathbf{X}_i(t_r^-)}{\alpha_i} > 0$*; otherwise the demand is lost. (If there are ties when determining the argmax, it assigns from the location with highest index.*[8]*)*

As may be expected, SMW policies tend to equalize the scaled queue lengths if CRP holds. The following fact is formalized later in Proposition 2.5 in Section 2.5.

Remark 2.1 (Resting state under SMW($\boldsymbol{\alpha}$))**.** *If Assumptions 2.1, 2.2 and 2.3 hold then for any* $\boldsymbol{\alpha} \in \text{relint}(\Omega)$*, the SMW(*$\boldsymbol{\alpha}$*) policy has a "resting state"* $\boldsymbol{\alpha}$*: Specifically, consider using SMW(*$\boldsymbol{\alpha}$*) on a sequence of systems indexed by the number of supply units* K*. Then there exists* $T_0 = T_0(\alpha) > 0$ *which does not depend on* K*, such that for any* $T > T_0$*,*

$$\limsup_{K\to\infty} \left(\max_{\mathbf{X}^K(0)\in\Omega_K} \left\| \tfrac{1}{K}\mathbf{X}^{K,\boldsymbol{\alpha}}(T) - \boldsymbol{\alpha} \right\|_2 \right) = 0 \quad \text{almost surely}\,,$$

where $\mathbf{X}^{K,\boldsymbol{\alpha}}(T)$ *is the state of the* K*-th system at time* T*.*

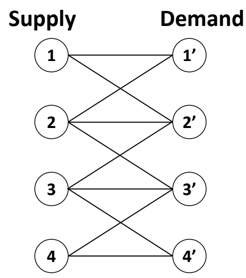

Figure 2.2: An example compatibility graph.

We conclude this section with an example which illustrates our model and SMW policies, and provides a brief preview of our main result.

Example 2.1. *Consider a network with "line-of-four-nodes" compatibility graph given as*

$G = (V_S \cup V_D, E) = (\{1,2,3,4\} \cup \{1',2',3',4'\}, \{11',12',21',22',23',32',33',34',43',44'\})\,;$

see Figure 2.2. Let the demand type distribution $\boldsymbol{\phi}$*, supported on types* $\{1'3, 2'4, 3'1, 4'2\}$*, be*

$$\phi_{1'3} = \phi_{2'4} = 0.25\,,\ \phi_{3'1} = 0.1\,,\ \phi_{4'2} = 0.4\,.$$

[8]Our analysis and results are unchanged if any other deterministic tie-breaking rule is employed instead.

It is easy to verify that the network $(G, \boldsymbol{\phi})$ satisfies Assumptions 2.1 and 2.2. It also satisfies the CRP condition (Assumption 2.3): Table 2.3 lists the limited-flexibility subsets $\mathcal{J}$, i.e., the demand node subsets J whose net demand $\mu_J > 0$, and their neighborhoods, net supply λ_J and net demand. For example, $\lambda_{\{1'\}} = \phi_{3'1} + \phi_{4'2} = 0.5$ and $\mu_{\{1'\}} = \phi_{1'3} = 0.25$. We see that the net supply exceeds net demand $\lambda_J > \mu_J$ for each limited-flexibility subset, as required. We also observe that the log ratio $\xi_J \triangleq \log\left(\frac{\lambda_J}{\mu_J}\right)$ is smallest for $J = \{4'\}$.

Our main result (in the next section) will tell us that because this network satisfies our assumptions, for any $\boldsymbol{\alpha} \in \mathrm{relint}(\Omega)$, the SMW($\boldsymbol{\alpha}$) policy achieves a loss which decays exponentially in K. The result will moreover say that the loss exponent achieved by SMW($\boldsymbol{\alpha}$) is explicitly given by $\gamma(\boldsymbol{\alpha}) = \min_{J\in\mathcal{J}} \mathbf{1}^{\mathrm{T}}_{\partial(J)}\boldsymbol{\alpha} \cdot \xi_J > 0$, and establish that there is an SMW policy which is globally exponent optimal. In particular, in this example:

- *(Optimal SMW policy) The SMW policy with*

$$\boldsymbol{\alpha} = \bar{\boldsymbol{\alpha}} = \begin{bmatrix} \frac{b}{2} & \frac{b}{2} & \frac{1-b}{2} & \frac{1-b}{2} \end{bmatrix}^{\mathrm{T}} \qquad \text{for } b = \frac{\log 1.25}{\log 2 + \log 1.25} \approx 0.244 \tag{2.11}$$

  has (normalized) resting state $\bar{\boldsymbol{\alpha}}$ and achieves loss exponent $\gamma(\bar{\boldsymbol{\alpha}}) = \frac{\log 1.25 \cdot \log 2}{\log 2 + \log 1.25} \approx 0.169$. SMW($\bar{\boldsymbol{\alpha}}$) maximizes $\gamma(\boldsymbol{\alpha})$ and is, in fact, exponent optimal among all *possible policies.*

- *(Vanilla MaxWeight achieves a suboptimal exponent) The vanilla MaxWeight policy has (normalized) resting state $\begin{bmatrix} \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \end{bmatrix}^{\mathrm{T}}$ and achieves a loss exponent $0.5 \log 1.25 \approx 0.112$.*

Note that the resting state $\bar{\boldsymbol{\alpha}}$ of the exponent optimal policy "protects" the subset $\{4'\}$ which has the smallest λ_J/μ_J by putting $\alpha_3 + \alpha_4 = 1 - b \approx 75.6\%$ fraction of supply in its neighborhood.[9]

[9] In this example, it turns out that the achieved exponent $\gamma(\boldsymbol{\alpha}) = \max\left((\alpha_1+\alpha_2)\xi_{\{1'\}}, (\alpha_3+\alpha_4)\xi_{\{4'\}}\right)$ hinges entirely on the tradeoff between protecting $\{1'\}$ and $\{4'\}$. Specifically, SMW with any $\boldsymbol{\alpha} \in \mathrm{relint}(\Omega)$ satisfying $\alpha_3 + \alpha_4 = 1 - b \approx 75.6\%$ is exponent optimal, and $\boldsymbol{\alpha}$ defined in (2.11) represents one such choice.

Table 2.1: Limited-flexibility subsets $J \in \mathcal{J}$ in Example 2.1, their neighborhood $\partial(J)$, net demand μ_J and net supply λ_J.

J	$\partial(J)$	μ_J	λ_J	$\xi_J \triangleq \log\left(\frac{\lambda_J}{\mu_J}\right)$
$\{1'\}$	$\{1,2\}$	0.25	0.5	0.69
$\{1',2'\}$	$\{1,2,3\}$	0.25	0.5	0.69
$\{3',4'\}$	$\{2,3,4\}$	0.1	0.5	1.61
$\{4'\}$	$\{3,4\}$	0.4	0.5	0.22

2.4 Main Result

In this section we present our main result, which says that for any network such that CRP holds: (i) All Scaled Maxweight (SMW) policies yield exponential decay of demand loss in the number of supply units K, with an exponent which we explicitly specify. (ii) For scaling parameter vector $\boldsymbol{\alpha}$ which maximizes the exponent among SMW policies, the SMW($\boldsymbol{\alpha}$) policy is exponent optimal among all possible policies. In sharp contrast, we show in Section 2.4.2 that that no state-independent assignment policy can achieve loss which decays exponentially in K, and moreover that if demand arrival rates are not perfectly known, then the loss of a state-independent policy (generically) does not vanish as $K \to \infty$. Also, a naive state-dependent control policy suffers $\Omega(1)$ loss as $K \to \infty$.

Recall from Section 2.2.3 the set of subsets of demand nodes with limited flexibility

$$\mathcal{J} = \left\{ J \subsetneq V_D : \sum_{j' \in J} \sum_{k \notin \partial(J)} \phi_{j'k} > 0 \right\}. \tag{2.12}$$

The following is our main result.

Theorem 2.1 (**Main Result**)**.** *For any network* $(G, \boldsymbol{\phi})$ *satisfying Assumptions 2.1, 2.2 and 2.3, we have:*

1. ***Exponentially small loss under any SMW policy****: For any* $\boldsymbol{\alpha} \in \mathrm{relint}(\Omega)$, SMW($\boldsymbol{\alpha}$) *achieves exponential decay of the demand loss probability with exponent*[10,11]

$$\gamma(\boldsymbol{\alpha}) = \min_{J \in \mathcal{J}} B_J \log\left(\frac{\lambda_J}{\mu_J}\right) > 0\,, \tag{2.13}$$

$$\text{where} \quad B_J \triangleq \mathbf{1}_{\partial(J)}^{\mathrm{T}} \boldsymbol{\alpha}\,, \qquad \lambda_J \triangleq \sum_{j' \notin J} \sum_{k \in \partial(J)} \phi_{j'k}\,, \quad \text{and} \quad \mu_J \triangleq \sum_{j' \in J} \sum_{k \notin \partial(J)} \phi_{j'k}\,.$$

2. ***There is an exponent optimal SMW policy:*** *Under* any *policy U, it must be that*

$$\gamma_{\mathrm{p}}(U) \le \gamma_{\mathrm{o}}(U) \le \bar{\gamma}\,, \qquad \text{where} \quad \bar{\gamma} = \sup_{\boldsymbol{\alpha} \in \mathrm{relint}(\Omega)} \gamma(\boldsymbol{\alpha})\,. \tag{2.14}$$

Thus, there is an SMW *policy that achieves an exponent arbitrarily close to the optimal one.*

The first part of the theorem states that for any SMW policy with $\boldsymbol{\alpha}$ in the relative interior of Ω, the policy achieves an explicitly specified positive demand loss exponent $\gamma(\boldsymbol{\alpha})$, i.e., the demand loss probability decays as $e^{-(\gamma(\boldsymbol{\alpha})-o(1))K}$ as $K \to \infty$. The second part of the theorem provides a universal upper bound $\bar{\gamma}$ on the exponent that *any* policy can achieve, i.e., for any assignment policy U, the demand loss probability is at least $e^{-(\bar{\gamma}+o(1))K}$. Crucially, $\bar{\gamma}$ is identical to the supremum over α of $\gamma(\boldsymbol{\alpha})$. In other words, *there is an (almost) exponent optimal* SMW *policy*, and moreover, the scaling parameters for this policy can be obtained as the solution to the explicit problem: $\text{maximize}_{\boldsymbol{\alpha} \in \mathrm{relint}(\Omega)} \gamma(\boldsymbol{\alpha})$.

We note that Theorem 2.1 is qualitatively different from the numerous results showing near optimality of (vanilla) maximum weight matching in various open queueing network settings [e.g., 16, 17, show that vanilla MaxWeight asymptotically minimizes workload in heavy-traffic in certain open queueing networks under the CRP condition]. Despite our objective (minimize demand loss) being symmetric in all the m queues, our result says that there is an optimal *scaled* maximum weight policy, that is *not* symmetric in

[10]We show that for SMW policies, the lim inf in (2.3) and lim sup in (2.4) are equal, i.e., $\gamma_{\mathrm{o}}(\boldsymbol{\alpha}) = \gamma_{\mathrm{p}}(\boldsymbol{\alpha})$. (We use $\boldsymbol{\alpha}$ to represent the policy SMW($\boldsymbol{\alpha}$) in the argument of the γs.)

[11]Note that the argument of the logarithm has a strictly larger numerator than denominator for every $J \subsetneq V_D$ since Assumption 2.3 holds, implying that $\gamma(\boldsymbol{\alpha})$ is the minimum of finitely many positive numbers, and hence is positive.

the m queues; rather, it is uses asymmetric scaling factors that optimally account for the network primitives.

Intuition for $\gamma(\boldsymbol{\alpha})$. Consider the expression for $\gamma(\boldsymbol{\alpha})$ in (2.13). It is a minimum over subsets $J \in \mathcal{J}$ of demand nodes of a certain "robustness" of the subset to demand loss. For subset J, the robustness of SMW($\boldsymbol{\alpha}$)'s ability to serve demand arising in J is the product of two terms $B_J \times \log\left(\frac{\lambda_J}{\mu_J}\right)$ (see Figure 2.3 for an illustration of the quantities involved):

- "Protection" due to $\boldsymbol{\alpha}$: At the resting point $\boldsymbol{\alpha}$ (see Remark 2.1) of SMW($\boldsymbol{\alpha}$), the supply at neighboring nodes is $B_J = \mathbf{1}^{\mathrm{T}}_{\partial(J)}\boldsymbol{\alpha}$, and the larger that is, the more unlikely it is that the subset will be deprived of supply.
- "Inherent robustness" arising from excess of supply over demand: The logarithmic term $\xi_J \triangleq \log(\lambda_J/\mu_J)$ captures the *inherent robustness* of that subset is to being drained of supply. Recall that λ_J is the (optimistic) net supply coming in to $\partial(J)$, and that μ_J is the net demand taking supply out of $\partial(J)$. The larger the ratio λ_J/μ_J, the more oversupplied and hence robust J is.

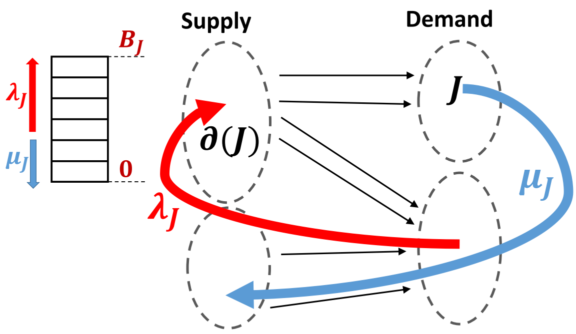

Figure 2.3: An illustration of the terms B_J, λ_J, and μ_J in Theorem 2.1.

Remarkably, the expression for robustness of subset J under SMW($\boldsymbol{\alpha}$) is as large (i.e., as good) as the demand loss exponent for subset J alone would be, with starting state $\boldsymbol{\alpha}$, under a "protect-J" policy which *exclusively protects J at the expense of all other nodes.* (Similar to standard buffer overflow probability calculations, the likelihood of the supply at $\partial(J)$ being depleted by KB_J units under a protect-J policy is $\Theta((\lambda_J/\mu_J)^{-KB_J}) = \Theta(\exp(-KB_J \log(\lambda_J/\mu_J)))$. We then set B_J to the starting scaled supply at $\partial(J)$, i.e.,

$B_J = \mathbf{1}_{\partial(J)}^{\mathrm{T}} \boldsymbol{\alpha}$, to establish the claim.) Thus, Theorem 2.1 part 1 says that given the resting state $\boldsymbol{\alpha}$, SMW($\boldsymbol{\alpha}$) achieves an exponent such that *it suffers no loss from the need to protecting multiple subsets J simultaneously.* Given this remarkable property, it is intuitive that the globally optimal exponent can be achieved via an SMW policy by choosing $\boldsymbol{\alpha}$ suitably (part 2 of the theorem).

Structural insights. The choice of scaling factors (resting state) $\boldsymbol{\alpha}$ for SMW which maximizes the exponent $\gamma(\boldsymbol{\alpha})$ as a function of network primitives $(G, \boldsymbol{\phi})$ is discussed in Section 2.4.1.

Proof approach. We establish Theorem 2.1 via a *novel Lyapunov analysis for a closed queueing network.* A key technical challenge we face in our closed queueing network setting is that it is a priori unclear what the ideal state for the system is. This is in contrast to open queueing network settings in which the ideal state is typically the one in which all queues are empty, and the Lyapunov functions considered typically achieve their minimum at this state. We overcome the challenge of unknown ideal state via an innovative approach as follows: We define a *policy-specific* Lyapunov function that achieves its minimum at the resting point of the SMW policy we are analyzing, and use this Lyapunov function to characterize its exponent $\gamma(\boldsymbol{\alpha})$. Moreover, given the optimal choice of $\boldsymbol{\alpha}$, our tailored Lyapunov function corresponding to this choice of $\boldsymbol{\alpha}$ helps us establish our converse result. In particular, the ideal state is finally revealed as a byproduct of our analysis to be equal to the optimal choice of $\boldsymbol{\alpha}$. Our technical machinery may be broadly useful in deriving large-deviation optimal controls in settings where the appropriate target state is apriori unclear. Our analysis is described in Section 2.5.

Transient performance. Our analysis extends readily to finite horizon performance: Considering transient behavior over a finite horizon (which is not too short), under a starting scaled state $\frac{\mathbf{X}^K(0)}{K} = \boldsymbol{\alpha} \in \mathrm{relint}(\Omega)$, we find that the optimal demand loss exponent is $\gamma(\boldsymbol{\alpha})$ given by (2.13) and SMW($\boldsymbol{\alpha}$) achieves it. The formal statement is provided in

Appendix B.4.4.

Utilization Rate of Supply Units. Recall that we consider the large market regime where the number of supply units K and the demand arrival rates $\hat{\boldsymbol{\phi}}^K \triangleq K\hat{\boldsymbol{\phi}}$ scale up at the same rate. The next proposition shows that in this regime under any SMW policy, supply units are "frequently" in use, in the sense that is formalized below.

Definition 2.2 (Resource utilization rate)**.** *Given a policy $U \in \mathcal{U}$, the resource utilization rate $\xi^{K,U}$ is the average number of demands served per supply unit per unit time in steady state in the K-th system.*

Proposition 2.2. *Consider any network $(G, \boldsymbol{\phi})$ satisfying Assumptions 2.1, 2.2 and 2.3 and any $\boldsymbol{\alpha} \in \mathrm{relint}(\Omega)$. Consider the SMW($\boldsymbol{\alpha}$) policy and denote its resource utilization rate by $\xi^{K,\boldsymbol{\alpha}}$.*

1. *(Utilization rate) There exists $c > 0$ such that for any $K > 0$ we have $\xi^{K,\boldsymbol{\alpha}} > c$.*
2. *(Waiting time) Suppose the head-of-line unit from the queue at the supply location is chosen in a first-in-first-out (FIFO) manner when implementing SMW($\boldsymbol{\alpha}$), then there exists $w < \infty$ such that for every $K > 0$, for every current state $\mathbf{X}(t)$, and every supply unit (distinguished by its location in V_S and its queue position), the expected waiting time before the supply unit is assigned is at most[12] w.*

Proposition 2.2 tells us that for any SMW policy, the resource utilization rate is bounded below by a positive constant which does not depend on K. See Appendix B.5 for the proof.

2.4.1 Optimal choice of scaling factors

In this subsection, we discuss the optimal choice of the scaling factors (resting state) $\boldsymbol{\alpha}$ based on Theorem 2.1. We illustrate the structure of the optimal $\boldsymbol{\alpha}$ via two examples (formal corollaries generalizing each example to arbitrary compatibility graphs are provided in Appendix B.5).

[12] The same result also holds when the supply unit is chosen uniformly at random from the queue.

We start by defining a *vulnerable subset* as one with small inherent robustness.

Definition 2.3 (Vulnerable subset). *Given a compatibility graph G and a sequence of demand type distributions $(\boldsymbol{\phi}^n)_{n\in\mathbb{Z}_+}$, we say that a limited-flexibility subset of demand nodes $J \subset \mathcal{J}$ is* vulnerable *if its inherent robustness vanishes as n grows:*

$$\xi_J \triangleq \log\left(\frac{\lambda_J^n}{\mu_J^n}\right) \xrightarrow{n\to\infty} 0\,. \tag{2.15}$$

Our first example considers the case of exactly one vulnerable subset.

Example 2.2 (If one subset of nodes is vulnerable, the optimal $\boldsymbol{\alpha}$ protects it). *Consider the "line-of-four-nodes" compatibility graph introduced in Example 2.1 and Figure 2.2, and the sequence of demand type distribution matrices.*

$$\boldsymbol{\phi}^n = \begin{array}{c} \\ 1' \\ 2' \\ 3' \\ 4' \end{array}\begin{array}{c} \begin{array}{cccc} 1 & 2 & 3 & 4 \end{array} \\ \begin{bmatrix} 0 & 0 & 1/4 & 1/4-\eta_n \\ 0 & 0 & 0 & \eta_n \\ \delta_n & 0 & 0 & 0 \\ 1/4-\delta_n & 1/4 & 0 & 0 \end{bmatrix} \end{array} \quad \text{for } n \in \mathbb{Z}_+\,. \tag{2.16}$$

We set $\delta_n = 1/n$ and $\eta_n = 1/8$ in this example (and consider $n > 4$). Note that $(G, \boldsymbol{\phi}^n)$ satisfies Assumptions 2.1, 2.2 and 2.3 for all $n > 4$.

The subsets of demand locations with limited flexibility are the same for all ϕ^n in the sequence $\mathcal{J} = \{\{1'\}, \{1', 2'\}, \{3', 4'\}, \{4'\}\}$. Consider these subsets one by one. We have $\lambda_{\{4'\}} = \frac{1}{2}$ and $\mu_{\{4'\}} = \frac{1}{2} - \frac{1}{n}$, which tells us that $\{4'\}$ is a "vulnerable" subset since

$$\xi_{\{4'\}} \triangleq \log\left(\frac{\lambda_{\{4'\}}}{\mu_{\{4'\}}}\right) = \frac{2}{n} + O\left(\frac{1}{n^2}\right) \xrightarrow{n\to\infty} 0^+\,.$$

Meanwhile, the other subsets are not vulnerable in the sense that $\xi_J \triangleq \log(\lambda_J/\mu_J)$ remains bounded away from zero: $\xi_{\{1'\}} = \log\left(\frac{1/2}{3/8}\right) \xrightarrow{n\to\infty} \log(4/3) > 0$, and $\xi_{\{1',2'\}} = \xi_{\{3',4'\}} = \log\left(\frac{1/2}{1/4}\right) = \log 2 > 0$. We deduce from Theorem 2.3 (as formalized in Corollary B.1 in Appendix B.5), that for any $\epsilon > 0$, there exists $n_0 < \infty$ such that, for all $n > n_0$, for network $(G, \boldsymbol{\phi}^n)$ we have

(i) *(Optimal exponent) The best achievable exponent* $\bar{\gamma}$ *is close to* $\xi_{\{4'\}}$*. Formally,* $\bar{\gamma} \in [(1-\epsilon)\xi_{\{4'\}}, \xi_{\{4'\}}]$ *and, as always, SMW policies suffice to achieve it* $\bar{\gamma} = \sup_{\boldsymbol{\alpha}\in \mathrm{relint}(\Omega)} \gamma(\boldsymbol{\alpha})$.

(ii) *(Near optimal* $\boldsymbol{\alpha}$ *protects vulnerable subset* $\{4'\}$*.) If SMW with scaling factors* $\boldsymbol{\alpha} \in \mathrm{relint}(\Omega)$ *achieves a demand-loss exponent* $\gamma(\boldsymbol{\alpha}) \geq (1-\epsilon)\xi_{\{4'\}}$*, then it must be that* $\alpha_3 + \alpha_4 \geq 1 - \epsilon$*. (Note that* $\partial(4') = \{3, 4\}$*.)*

(iii) *(Example of near optimal* $\boldsymbol{\alpha}$*.) The SMW policy with* $\boldsymbol{\alpha} = \begin{bmatrix} \frac{\epsilon}{2} & \frac{\epsilon}{2} & \frac{1-\epsilon}{2} & \frac{1-\epsilon}{2} \end{bmatrix}^{\mathrm{T}}$ *achieves* $\gamma(\boldsymbol{\alpha}) = (1-\epsilon)\xi_{\{4'\}}$.

Example 2.2 illustrates Corollary B.1 in Appendix B.5, which demonstrates that if there is just one vulnerable subset of demand nodes J_1, then the exponent optimal SMW policy has a resting state which puts almost all the supply in the neighborhood of J_1. The intuition is that the total supply located in $\partial(J_1)$ follows a random walk which has only slightly positive drift even if the assignment rule protects it (recall that the definition of the net supply λ_{J_1} assumes that the policy protects J_1), and hence it is optimal to keep the total supply in $\partial(J_1)$ at a high resting point, to minimize the likelihood of depletion.

Our next example illustrates the case of two non-overlapping vulnerable subsets.

Example 2.3 (If there are two non-overlapping vulnerable subsets, the optimal $\boldsymbol{\alpha}$ protects them in inverse proportion to their inherent robustness)**.** *Once again consider the same compatibility graph as in Example 2.2. We further take the sequence* $\boldsymbol{\phi}^n$ *given by* (2.16) *again with* $\delta_n = 1/n$ *but change the definition of* η_n *to* $\eta_n = \eta/n$ *for some fixed* $\eta > 0$ *(we consider* $n > 4/\min(1,\eta)$*). Note that* $\lim_{n\to\infty} \boldsymbol{\phi}^n = \boldsymbol{\phi}^*$ *where* $\boldsymbol{\phi}^*$ *is given by* (2.16) *with* δ_n *and* η_n *both replaced by* 0.

The limited-flexibility subsets of demand locations are the same for all ϕ^n *in the sequence* $\mathcal{J} = \{\{1'\}, \{1', 2'\}, \{3', 4'\}, \{4'\}\}$*. The two singleton subsets are vulnerable:*

$$\xi_{\{4'\}} \triangleq \log\left(\frac{\lambda_{\{4'\}}}{\mu_{\{4'\}}}\right) = \log\left(\frac{1/2}{1/2 - 1/n}\right) = \frac{2}{n} + O\big(\tfrac{1}{n^2}\big) \xrightarrow{n\to\infty} 0^+, \quad \xi_{\{1'\}} = \frac{2\eta}{n} + O\big(\tfrac{1}{n^2}\big) \xrightarrow{n\to\infty} 0^+,$$

and $\frac{\xi_{\{1'\}}}{\xi_{\{4'\}}} = \eta + O(\frac{1}{n})$*. The other subsets are not vulnerable since* $\xi_{\{1',2'\}} = \xi_{\{3',4'\}} = \log\big(\frac{1/2}{1/4}\big) = \log 2 > 0$*. We deduce from Theorem 2.3 (formalized in Corollary B.2 in*

Appendix B.5), that for any $\epsilon > 0$*, there exists* $n_0 < \infty$ *such that, for all* $n > n_0$*, for network* $(G, \boldsymbol{\phi}^n)$ *we have*

(i) (Optimal exponent) The best achievable exponent $\bar{\gamma}$ *is close to* $H \triangleq \frac{\xi_{\{4'\}}\xi_{\{1'\}}}{\xi_{\{4'\}}+\xi_{\{1'\}}} = \frac{1}{n} \cdot \frac{\eta}{1+\eta} + O\big(\frac{1}{n^2}\big)$ *. Formally,* $\bar{\gamma} \in [(1-\epsilon)H, H]$*, and, as always, SMW policies suffice to achieve it* $\bar{\gamma} = \sup_{\boldsymbol{\alpha} \in \mathrm{relint}(\Omega)} \gamma(\boldsymbol{\alpha})$*.*

(ii) (Near optimal $\boldsymbol{\alpha}$ *protects vulnerable subsets in inverse proportion to their inherent robustness.) If SMW with scaling factors* $\boldsymbol{\alpha} \in \mathrm{relint}(\Omega)$ *achieves a demand-loss exponent* $\gamma(\boldsymbol{\alpha}) \geq (1-\epsilon)H$*, then it must be that*

$$\alpha_1 + \alpha_2 \overset{\epsilon}{=} \frac{\xi_{\{4'\}}}{\xi_{\{4'\}}+\xi_{\{1'\}}} = \frac{1}{1+\eta} + O\big(\tfrac{1}{n}\big) \quad \text{and} \quad \alpha_3 + \alpha_4 \overset{\epsilon}{=} \frac{\xi_{\{1'\}}}{\xi_{\{4'\}}+\xi_{\{1'\}}} = \frac{\eta}{1+\eta} + O\big(\tfrac{1}{n}\big)\,,$$

where $a \overset{\epsilon}{=} b$ *represents* $|a-b| \leq \epsilon$*. (Recall that* $\partial(1') = \{1,2\}$ *and* $\partial(4') = \{3,4\}$*.)*

(iii) (Example of near optimal $\boldsymbol{\alpha}$*.) The SMW policy with*

$$\boldsymbol{\alpha} = \begin{bmatrix} \frac{\eta'}{2(1+\eta')} & \frac{\eta'}{2(1+\eta')} & \frac{1}{2(1+\eta')} & \frac{1}{2(1+\eta')} \end{bmatrix}^T \quad \text{for} \quad \eta' = \frac{\xi_{\{1'\}}}{\xi_{\{4'\}}} = \eta + O\big(\tfrac{1}{n}\big) \tag{2.17}$$

achieves $\gamma(\boldsymbol{\alpha}) \geq (1-\epsilon)H$*.*

Example B.2 illustrates Corollary B.2 in Appendix B.5 which tells us that if there are two non-overlapping vulnerable subsets of demand nodes J_1 and J_2, then the exponent optimal SMW policy has a resting state which divides the supply between the two neighborhoods in inverse proportion to the inherent robustness of the vulnerable subsets

$$\frac{\mathbf{1}^{\mathrm{T}}_{\partial(J_2)}\boldsymbol{\alpha}}{\mathbf{1}^{\mathrm{T}}_{\partial(J_1)}\boldsymbol{\alpha}} \approx \frac{\xi_{J_1}}{\xi_{J_2}} \approx \eta\,.$$

In this simple example, $\partial(J_1) \cup \partial(J_2) = V_S$. More generally, if $\partial(J_1) \cup \partial(J_2) \subsetneq V_S$, then the optimal $\boldsymbol{\alpha}$ places very little supply at nodes outside the union of neighborhoods $\partial(J_1) \cup \partial(J_2)$; see Corollary B.2.

While the examples above (and the corollaries they illustrate) focusing on the cases of one or two vulnerable subsets are interesting in themselves; we highlight that the optimal policy characterized in Theorem 2.1 goes much beyond to solve the general m-dimensional problem considering *all* subsets of V_S simultaneously. SMW with the optimal $\boldsymbol{\alpha}$ balances between the demands of protecting different subsets and is (provably) globally exponent optimal.

Knowledge requirements. We remark that choosing the exponent optimal $\boldsymbol{\alpha}$ requires exact knowledge of $\boldsymbol{\phi}$. However, if a noisy estimate of the demand type distribution is employed to choose $\boldsymbol{\alpha}$ (by maximizing the exponent for the estimated distribution), the resulting SMW policy will nevertheless perform well: (i) it will achieve exponentially small loss (as long as the true $\boldsymbol{\phi}$ satisfies our assumptions), (ii) if the estimate of $\boldsymbol{\phi}$ is close to the true distribution, then the exponent achieved by the chosen $\boldsymbol{\alpha}$ will be close to the estimated exponent based on the estimated distribution, since $\gamma(\boldsymbol{\alpha})$ given by (2.13) varies continuously in $\boldsymbol{\phi}$ for each $\boldsymbol{\alpha} \in \mathrm{relint}(\Omega)$.

2.4.2 State-independent policies and naive state-dependent policies are inferior

State-independent policies. Previous works studying control of circulating resources in networks, e.g., [21] and [1], have proposed state-independent control policies. We show that in our setting, such policies are not competitive with the SMW policies we have proposed.

We first formally define state-independent policies.

Definition 2.4 (State independent policy)**.** *We call an assignment policy U* state independent *if, for each*[13] *$K \geq 1$, it maps each $j' \in V_D$, $k \in V_S$, $r \in \mathbb{Z}_+$ to a distribution $u_{j'k}(t_r^-)$ over $\partial(j') \cup \{\emptyset\}$; for the r-th demand arrival with origin j' and destination k, the platform dispatches from i drawn independently from distribution $u_{j'k}(t_r^-)$, ignoring the current state $\mathbf{X}(t_r^-)$ and the history. If $i = \emptyset$ or there is no supply at the dispatch node, the demand is lost.*

The next proposition formalizes that for any state independent policy: (i) Exponentially small loss is impossible (even if demand arrival rates are exactly known), (ii) Given

[13] We suppress the dependence on K in our notation.

a compatibility graph G and a state independent policy, for "almost all" demand type distributions $\boldsymbol{\phi}$ the loss incurred under the policy does not vanish as $K \to \infty$; informally, asymptotic optimality fails if $\boldsymbol{\phi}$ is not exactly known. The proof is in Appendix B.6.

Proposition 2.3 (All state independent policies have inferior performance)**.** *Fix a compatibility graph G and any state-independent dispatch policy U. We have:*

1. *(Exponentially small loss is impossible.) For any demand type distribution $\boldsymbol{\phi}$, $\mathbb{P}_{\mathrm{o}}^{K,U} = \Omega\left(\frac{1}{K^2}\right)$. In particular, $\gamma_{\mathrm{o}}(U) = 0$, where $\gamma_{\mathrm{o}}(\cdot)$ is the optimistic exponent defined in* (2.3)*.*

2. *(For almost all $\boldsymbol{\phi}$, asymptotic optimality fails.) Let $\mathrm{Supp}(\boldsymbol{\phi}) \triangleq \{(j',k) \in V_D \times V_S : \phi_{j'k} > 0\}$. Fix any subset of demand types $S \subseteq V_D \times V_S$ such that each demand node $j' \in V_D$ has at least one demand type in S. Let $D(S) \triangleq \{\boldsymbol{\phi} : \mathrm{Supp}(\boldsymbol{\phi}) = S\}$ be the set of demand type distributions with support S. Then, then there is a subset of $D(S)$ which is open and dense in $D(S)$ such that for all $\boldsymbol{\phi}$ in this subset it holds that $\liminf_{K\to\infty} \mathbb{P}_{\mathrm{o}}^{K,U} > 0$.*

Proposition 2.3 makes it clear that as K grows, any state independent policy suffers from inferior performance. There are two possibilities regarding what is known about the demand type distribution $\boldsymbol{\phi}$:

1. **$\boldsymbol{\phi}$ exactly known.** In this case, part 1 of Proposition 2.3 tells us that any state independent policy has loss $\Omega(\frac{1}{K^2})$ whereas any SMW policy produces exponentially small loss (Theorem 2.1 part 1) and moroever SMW($\boldsymbol{\alpha}$) is exponent optimal for $\boldsymbol{\alpha}$ chosen to maximize $\gamma(\boldsymbol{\alpha})$ in (2.13).
2. **$\boldsymbol{\phi}$ is not exactly known.** In this case, any state independent policy typically fails to achieve asymptotic optimality (part 2 of Proposition 2.3) whereas vanilla MaxWeight (or any fixed SMW policy) achieves exponentially small loss.

A naive state-dependent policy. Would a naive state dependent policy do well in our setting? For a natural state dependent policy, we show via a simple example that the loss is $\Omega(1)$ as $K \to \infty$, even though the example network satisfies all our assumptions.

Define the *naive* policy as follows: each time a demand arrives, consider the supply nodes compatible with the origin in a uniformly random order (independently of the past), and assign a supply unit from the first compatible supply node which has at least one supply unit.

Example 2.4 (Naive state-dependent policy loses $\Omega(1)$). *Consider again the "line-of-four-nodes" compatibility graph introduced in Example 2.1 and Figure 2.2, and the demand type distribution matrix*

$$\phi = \begin{array}{c} \\ 1' \\ 2' \\ 3' \\ 4' \end{array} \begin{array}{c} \begin{array}{cccc} 1 & 2 & 3 & 4 \end{array} \\ \begin{bmatrix} 0 & 0 & 0.21 & 0.21 \\ 0.08 & 0 & 0 & 0 \\ 0 & 0.1 & 0 & 0 \\ 0.4 & 0 & 0 & 0 \end{bmatrix} \end{array} . \tag{2.18}$$

It is easy to verify that this network satisfies Assumptions 2.1, 2.2 and 2.3. Even so, the naive policy incurs $\Omega(1)$ *loss in this network (in fact, this is true for any demand type distribution in a ball of positive radius centered at the right-hand side of* (2.18)*). The proof is in Appendix B.6.*

Variants of the naive policy which sample a compatible supply using a non-uniform distribution can similarly be shown to fail in simple examples.

2.5 Analysis of Scaled MaxWeight Policies: Proof of Theorem 2.1

In this section, we analyze the large deviations behavior of the system and prepare all the elements needed to prove Theorem 2.1. In Section 2.5.1, we follow the standard approach for large deviations analyses and characterize the system behavior in the fluid scale through fluid sample paths and fluid limits. In Section 2.5.2 we take a novel approach to define a family of Lyapunov functions parameterized by the desired state, since we do not know the ideal state for the system. In Section 2.5.3 we follow [45] and show that if the Lyapunov function (centered at the starting state) is scale-invariant and sub-additive, a policy that performs steepest descent on this Lyapunov function is exponent optimal. In Section 2.5.4 we prove that each SMW policy performs steepest descent on the Lyapunov function centered at its resting state and is hence exponent optimal given its resting state. We also explicitly characterize the optimal exponent, the most likely sample paths leading to demand loss, and the critical subsets (i.e., the subsets that are most likely to be depleted of supply). Finally, we deduce Theorem 2.1.

2.5.1 Fluid Sample Paths and Fluid Limits

For any stationary assignment policy $U \in \mathcal{U}$ defined in Section 2.2, we define the scaled demand and queue-length sample paths by (the former was defined in (2.7))

$$\bar{\mathbf{A}}^K_{j'k}(t) \triangleq \frac{1}{K}\mathbf{A}^K_{j'k}(t)\,, \quad \bar{\mathbf{X}}^{K,U}_i(t) \triangleq \frac{1}{K}\mathbf{X}^{K,U}_i(t)\,, \tag{2.19}$$

Note that for a fixed policy (with specified tie-breaking rules), each given demand sample path and initial state uniquely determines the state sample path.

To obtain a large deviation result, we need to study the demand process and the queue-length process in the fluid scaling, as captured in (2.19). We take the standard approach of *fluid sample paths* (FSP) (see [47, 45]).

Definition 2.5 (Fluid sample paths)**.** *We call a pair* $(\bar{\mathbf{A}}(\cdot), \bar{\mathbf{X}}^U(\cdot))_T \triangleq (\bar{\mathbf{A}}(\cdot), \bar{\mathbf{X}}^U(\cdot))_{t\in[0,T]}$ *a fluid sample path on* $[0,T]$ *(under policy* U*) if there exists a sequence* $(\,(\bar{\mathbf{A}}^K(\cdot))_{t\in[0,T]}$, $(\bar{\mathbf{X}}^{K,U}(\cdot))_{t\in[0,T]}\,)$ *where* $\bar{\mathbf{A}}^K(\cdot)$ *are scaled demand sample paths and* $\bar{\mathbf{X}}^{K,U}(\cdot)$ *are state sample paths determined by the* $\bar{\mathbf{A}}^K(\cdot)$*'s, such that it has a subsequence which converges to* $((\bar{\mathbf{A}}(\cdot))_{t\in[0,T]}, (\bar{\mathbf{X}}^U(\cdot))_{t\in[0,T]})$ *uniformly on* $[0,T]$.

In short, FSPs include both typical and atypical sample paths. Recall Fact 1, which gives the likelihood for an unlikely event to occur based on the most likely fluid sample path that causes the event. Accordingly, the large deviations analysis in Section 2.5.4 will identify the most likely FSP that leads to demand loss.

Fluid limits are fluid sample paths that characterize *typical* system behavior, as they are the formal limits in the Functional Law of Large Numbers [56].

Definition 2.6 (Fluid limits). *We call a pair* $(\bar{\mathbf{A}}(\cdot), \bar{\mathbf{X}}^U(\cdot))_T$ *a fluid limit on* $[0, T]$ *(under policy* U*) if (i) the pair* $(\bar{\mathbf{A}}(\cdot), \bar{\mathbf{X}}^U(\cdot))_T$ *is a fluid sample path; (ii) we have* $\bar{\mathbf{A}}_{j'k}(t) = \hat{\phi}_{j'k}t$*, for all* $j' \in V_D, k \in V_S$ *and all* $t \in [0, T]$.

2.5.2 A Family of Lyapunov Functions

Lyapunov functions are a useful tool for analyzing complex stochastic systems. In open queuing networks the ideal state is one in which all queues are empty, and correspondingly the Lyapunov function is chosen to achieve its minimum value in the ideal state, e.g., the sum of squared queue lengths Lyapunov function is a popular choice [4, etc.], while others have also used piecewise linear Lyapunov functions ([57], etc.). Since our setting is a closed queueing network and ideal state is unknown, we instead construct a novel approach. We define a family of piecewise linear Lyapunov functions, parameterized by the desired state $\boldsymbol{\alpha}$, such that the function achieves its minimum at $\boldsymbol{\alpha}$.

Definition 2.7. *For each* $\boldsymbol{\alpha} \in \text{relint}(\Omega)$*, define Lyapunov function* $L_{\boldsymbol{\alpha}}(\mathbf{x}) : \Omega \to [0, 1]$ *as* $L_{\boldsymbol{\alpha}}(\mathbf{x}) \triangleq 1 - \min_i \frac{x_i}{\alpha_i}$.

The intuition behind our definition is as follows. The Lyapunov function value is jointly determined by the desired state $\boldsymbol{\alpha}$ of the system (under some policy) and our objective of avoiding demand loss, and can be interpreted as the energy of the system at each state. The desired state should have minimum energy, and the most undesirable states should have maximum energy. In our case the boundary $\partial\Omega$ of Ω is most undesirable since demand loss only happens there, and correspondingly, $L_{\boldsymbol{\alpha}}(\mathbf{x}) = 1$ for $\mathbf{x} \in \partial\Omega$, whereas $L_{\boldsymbol{\alpha}}(\boldsymbol{\alpha}) = 0$ as we want. See Figure 2.4 for an illustration. These functions moreover have the properties of being scale-invariant and sub-additive, which play a key role in our analysis. We state and prove these properties in Appendix B.1.

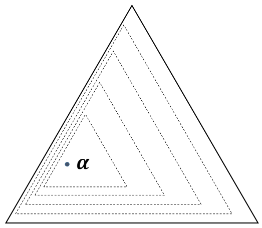

Figure 2.4: Sub-level sets of $L_{\boldsymbol{\alpha}}$ when $|V_S| = |V_D| = 3$. State space Ω is the probability simplex in $\mathbb{R}^3$, and its boundary coincides with $\{\mathbf{x} : L_{\boldsymbol{\alpha}}(\mathbf{x}) = 1, \mathbf{1}^{\mathrm{T}}\mathbf{x} = 1\}$. The minimum value is achieved at $\boldsymbol{\alpha}$; $L_{\boldsymbol{\alpha}}(\boldsymbol{\alpha}) = 0$.

2.5.3 Sufficient Conditions for Exponent Optimality

In this section, given a starting state, we provide a converse bound on the exponent for any stationary policy $U \in \mathcal{U}$, and derive sufficient conditions for a policy to achieve this bound.

We use the intuition from differential games (see, e.g., [58]) to informally illustrate the interplay between the control and the most likely sample path leading to demand loss. Consider a zero-sum game between the adversary (nature) who chooses the fluid-scale demand arrival process $\bar{\mathbf{A}}(\cdot)$, and the controller who decides the assignment rule U, where the adversary minimizes the large-deviation "cost" of a demand sample path that leads to demand loss. Specifically, the adversary's cost for a demand sample path $\bar{\mathbf{A}}(\cdot)$ is the rate function defined in (2.10), i.e., the exponent. The converse bound we will obtain next will correspond to the adversary playing first and choosing the minimum cost *time-invariant* demand sample path that ensures demand loss. The following pleasant surprises will emerge subsequently: (i) we will find an equilibrium in pure strategies to the aforementioned zero-sum game, (ii) the converse will turn out to be tight, i.e., the adversary's equilibrium demand sample path will be time invariant, (iii) the controller's equilibrium assignment strategy will be an SMW policy with specific $\boldsymbol{\alpha}$ (this simple policy will satisfy the sufficient conditions for achievability we will state immediately after our converse, in Proposition 2.4).

We provide a policy-independent upper bound on the exponent that only depends on the starting state. First, for any $\mathbf{f} \in \mathbb{R}_+^{n \times m}$, define

$$\mathcal{X}_{\mathbf{f}} \triangleq \left\{ \Delta\mathbf{x} \;\middle|\; \begin{array}{ll} \Delta x_i = \sum_{j' \in V_D} f_{j'i} - \sum_{j' \in \partial(i)} d_{ij'} \left(\sum_{k \in V_S} f_{j'k} \right), & \forall i \in V_S \\ \sum_{i \in \partial(j')} d_{ij'} = 1, \; d_{ij'} \geq 0, & \forall i \in V_S, j' \in V_D \end{array} \right\}, \tag{2.20}$$

which is the attainable change of (normalized) state in unit time, given that the average demand arrival rates during this period are $\mathbf{f}$ and assuming no demand is lost. (Here $(d_{ij'})_{i \in \partial(j')}$ is the chosen *assignment distribution* over supply nodes neighboring j' for assigning supply units to serve demand originating at j'.) Then given starting state $\boldsymbol{\alpha}$, the attainable states at time T belong to $\boldsymbol{\alpha} + T\mathcal{X}_{\mathbf{f}} \triangleq \{\mathbf{y} \in \mathbb{R}^m : \mathbf{y} = \boldsymbol{\alpha} + T\mathbf{x}, \mathbf{x} \in \mathcal{X}_{\mathbf{f}}\}$, if no demand is lost during $[0, T]$ and the average demand arrival rate is $\mathbf{f}$. We obtain an upper bound on the demand-loss exponent by considering the most likely $\mathbf{f}$ and T such that $\boldsymbol{\alpha} + T\mathcal{X}_{\mathbf{f}}$ lies entirely outside the state space Ω. Because the true state must lie in Ω, there must be demand loss during $[0, T]$, no matter the assignment rule $\mathbf{d}$ used by the controller.

Lemma 2.1 (Converse bound on the exponent)**.** *For any stationary policy $U \in \mathcal{U}$, it holds that*

$$-\liminf_{K \to \infty} \frac{1}{K} \log \mathbb{P}_{\mathrm{o}}^{K,U} \leq \sup\nolimits_{\boldsymbol{\alpha} \in \mathrm{relint}(\Omega)} \gamma_{\mathrm{CB}}(\boldsymbol{\alpha}), \tag{2.21}$$

where, for $\Lambda^*(\cdot)$ given by (2.8), $\gamma_{\mathrm{CB}}(\boldsymbol{\alpha}) \triangleq \inf\limits_{\mathbf{f} \in \mathbb{R}_+^{nm} : v_\alpha(\mathbf{f}) > 0} \dfrac{\Lambda^*(\mathbf{f})}{v_\alpha(\mathbf{f})}$, and $v_\alpha(\mathbf{f}) \triangleq \min\limits_{\Delta\mathbf{x} \in \mathcal{X}_{\mathbf{f}}} L_{\boldsymbol{\alpha}}(\boldsymbol{\alpha} + \Delta\mathbf{x})$.

We now provide an informal explanation for the form of this key lemma. The $\boldsymbol{\alpha}$ in (2.21) captures the most frequently visited (normalized) state (the "resting" state) in steady state under U, and $\gamma_{\mathrm{CB}}(\boldsymbol{\alpha})$ is an upper bound on the exponent given the most frequent state $\boldsymbol{\alpha}$. Let us informally describe the expression for $\gamma_{\mathrm{CB}}(\boldsymbol{\alpha})$. Suppose the system starts in state $\boldsymbol{\alpha}$. Then $v_{\boldsymbol{\alpha}}(\mathbf{f})$ is the minimum rate of increase of $L_{\boldsymbol{\alpha}}(\cdot)$ under demand arrival rates $\mathbf{f}$, no matter the assignment distributions $\mathbf{d}$. So, starting at $\boldsymbol{\alpha}$ and under time-invariant demand arrival rates $\mathbf{f}$, the state hits Ω and demand is lost in time at most $1/v_{\boldsymbol{\alpha}}(\mathbf{f})$, implying a demand-loss exponent of at most $\frac{\Lambda^*(\mathbf{f})}{v_{\boldsymbol{\alpha}}(\mathbf{f})}$. The upper bound $\gamma_{\text{CB}}(\boldsymbol{\alpha})$ follows from minimizing over $\mathbf{f}$ since nature can choose any $\mathbf{f}$. Finally, the bound in (2.21) takes the supremum over $\boldsymbol{\alpha}$ since the policy can choose its resting state. The proof of Lemma 2.1 is in Appendix B.2.

The following proposition provides sufficient conditions for a policy to achieve the converse bound exponent $\gamma_{\text{CB}}(\boldsymbol{\alpha})$. The conditions are requirements on the time derivative of $L_{\boldsymbol{\alpha}}(\bar{\mathbf{X}}^U(t))$. Recall that a time $t \in (0, T)$ is said to be a *regular point* of an FSP $(\bar{\mathbf{A}}(\cdot), \bar{\mathbf{X}}^U(\cdot))_T$ if $\bar{\mathbf{A}}(\cdot), \bar{\mathbf{X}}^U(\cdot), L_{\boldsymbol{\alpha}}(\bar{\mathbf{X}}^U(\cdot))$ are all differentiable at time t.

Proposition 2.4 (Sufficient conditions)**.** *Fix $\boldsymbol{\alpha} \in \text{relint}(\Omega)$. Let $U \in \mathcal{U}$ be a stationary, non-idling policy. Suppose that for each regular point t, the following hold:*

1. *(Steepest descent). For any demand fluid sample path $\bar{\mathbf{A}}(\cdot)$, we have*
$$\dot{L}_{\boldsymbol{\alpha}}(\bar{\mathbf{X}}^U(t)) = \inf_{U' \in \mathcal{U}_{\text{ni}}} \left\{ \dot{L}_{\boldsymbol{\alpha}}(\bar{\mathbf{X}}^{U'}(t)) \,\middle|\, \bar{\mathbf{X}}^{U'}(t) = \bar{\mathbf{X}}^U(t) \right\},$$
for corresponding queue-length sample paths satisfying $\bar{\mathbf{X}}^U(t) \neq \boldsymbol{\alpha}$ and $L_{\boldsymbol{\alpha}}(\bar{\mathbf{X}}^U(t)) < 1$, where $\mathcal{U}_{\text{ni}}$ is the set of non-idling policies;
2. *(Negative drift). There exists $\eta > 0$ and $\epsilon > 0$ such that for all FSPs $(\bar{\mathbf{A}}(\cdot), \bar{\mathbf{X}}^U(\cdot))$ satisfying $\dot{\bar{\mathbf{A}}}(t) \in B(\boldsymbol{\phi}, \epsilon)$ and $\bar{\mathbf{X}}(t) \neq \boldsymbol{\alpha}$, we have $\dot{L}_{\boldsymbol{\alpha}}(\bar{\mathbf{X}}^U(t)) \leq -\eta$. Here $B(\boldsymbol{\phi}, \epsilon)$ is a ball with radius ϵ centered on the typical demand type distribution $\boldsymbol{\phi}$.*

Then we have $\gamma_{\text{o}}(U) = \gamma_{\text{p}}(U) = \gamma_{\text{CB}}(\boldsymbol{\alpha})$, i.e., $\gamma(U) = \gamma_{\text{CB}}(\boldsymbol{\alpha})$.

Informally, the negative drift property requires the policy to have negative Lyapunov drift for near typical demand arrival rates, as long as the current state is not $\boldsymbol{\alpha}$. This property forces the state to return to $\boldsymbol{\alpha}$. Faced with a policy satisfying the above sufficient conditions, the adversary wants to force equality in (B.7) by forcing the queue-length sample path $\bar{\mathbf{X}}^U$ to go *radially* outward starting at $\boldsymbol{\alpha}$. This is why our converse in Lemma 2.1 based on a time invariant demand arrival process will turn out to be tight. We will formalize this intuition in Section 2.5.4 and explicitly characterize the most likely demand FSP forcing demand loss.

2.5.4 Optimality of SMW Policies, Explicit Exponent, and Critical Subsets

In this section, we verify that SMW policies satisfy the sufficient conditions in Proposition 2.4. In doing so, we reveal the critical subset structure of the most-likely sample paths for demand loss and derive the explicit exponent for SMW($\boldsymbol{\alpha}$). Proofs for this section are in Appendix B.4.

The following lemma shows that the Lyapunov drift only depends on the nodes with shortest scaled queue lengths, and that SMW($\boldsymbol{\alpha}$) minimizes its use of supplies from these queues.

Lemma 2.2 (SMW($\boldsymbol{\alpha}$) causes steepest descent)**.** *Let* $(\bar{\mathbf{A}}, \bar{\mathbf{X}}^U)$ *be any FSP under any non-idling policy* U *on* $[0,T]$*, and consider any* $\boldsymbol{\alpha} \in \mathrm{relint}(\Omega)$*. For a regular* $t \in [0,T]$*, define:*

$$S_1(\bar{\mathbf{X}}^U(t)) \triangleq \left\{ k \in V_S : k \in \mathrm{argmin} \frac{\bar{\mathbf{X}}^U_k(t)}{\alpha_k} \right\},$$

$$S_2\left(\bar{\mathbf{X}}^U(t), \dot{\bar{\mathbf{X}}}^U(t)\right) \triangleq \left\{ k \in S_1(\bar{\mathbf{X}}^U(t)) : k \in \mathrm{argmin} \frac{\dot{\bar{\mathbf{X}}}^U_k(t)}{\alpha_k} \right\}.$$

All the derivatives are well defined since t *is regular. We have*

$$\dot{L}_{\boldsymbol{\alpha}}(\bar{\mathbf{X}}^U(t)) = -\frac{\dot{\bar{\mathbf{X}}}^U_k(t)}{\alpha_k} \quad \text{for any } k \in S_2(\bar{\mathbf{X}}^U(t), \dot{\bar{\mathbf{X}}}^U(t)) \tag{2.22}$$

$$\geq -\frac{1}{\mathbf{1}_{S_2}^{\mathrm{T}}\boldsymbol{\alpha}} \left(\sum_{j' \in V_D, k \in S_2} \dot{A}_{j'k}(t) - \sum_{j' \in V_D : \partial(j') \subseteq S_2, k \in V_S} \dot{A}_{j'k}(t) \right) \tag{2.23}$$

for $\bar{\mathbf{X}}^U(t) \neq \boldsymbol{\alpha}$ *and* $L_{\boldsymbol{\alpha}}(\bar{\mathbf{X}}^U(t)) < 1$*. Inequality (2.23) holds with equality under* SMW($\boldsymbol{\alpha}$)*, i.e.,* SMW($\boldsymbol{\alpha}$) *satisfies the steepest descent property in Proposition 2.4.*

In Lemma 2.3, we prove that SMW($\boldsymbol{\alpha}$) satisfies the negative drift property. In particular, the drift η is related to the slack in the CRP condition.

Lemma 2.3 (SMW($\boldsymbol{\alpha}$) satisfies negative drift)**.** *For any* $\boldsymbol{\alpha} \in \mathrm{relint}(\Omega)$*, under Assumptions 2.1, 2.2 and 2.3, the policy* SMW*(*$\boldsymbol{\alpha}$*) satisfies the negative drift condition in Proposition 2.4.*

Before proceeding with our analysis, we point out that Lemma 2.3 implies that $\boldsymbol{\alpha}$ is the unique resting state of SMW($\boldsymbol{\alpha}$) policy.

Proposition 2.5 (Resting state of SMW($\boldsymbol{\alpha}$))**.** *Suppose Assumptions 2.1, 2.2 and 2.3 hold. For any* $\boldsymbol{\alpha} \in \text{relint}(\Omega)$*, there exists* $T_0 > 0$ *such that any fluid limit* $(\bar{\mathbf{A}}, \bar{\mathbf{X}})$ *on* $[0, T]$ *(where* $T > T_0$*) under SMW(*$\boldsymbol{\alpha}$*) satisfies* $\bar{\mathbf{X}}(t) = \boldsymbol{\alpha}$ *for all* $t \in [T_0, T]$*.*

Combining Proposition 2.4 with Lemmas 2.2 and 2.3, we immediately deduce that SMW($\boldsymbol{\alpha}$) achieves the best possible exponent given resting state $\boldsymbol{\alpha}$.

Corollary 2.1. *For any* $\boldsymbol{\alpha} \in \text{relint}(\Omega)$*, we have* $\gamma(\boldsymbol{\alpha}) = \gamma_{\text{CB}}(\boldsymbol{\alpha})$*.*

We argued in Section 2.5.3 that the most likely queue-length sample path leading to demand loss with initial state $\boldsymbol{\alpha}$ should be radial. From Lemma 2.2 we see that the rate at which the Lyapunov function increases depends on the (scaled) inflow and outflow rate of supply in each subset. This implies that the most likely sample path should drain the supply of one subset (the critical subset), and that subset will determine the demand loss exponent. We next lemma obtains an explicit expression for $\gamma_{\text{CB}}(\boldsymbol{\alpha})$ and the most likely demand FSP forcing demand loss.

Lemma 2.4. *Recall the definitions of* $\mathcal{J}$ *in* (2.12) *and* B_J, λ_J *and* μ_J *in* (2.13)*. For any* $\boldsymbol{\alpha} \in \text{relint}(\Omega)$*, we have* $\gamma_{\text{CB}}(\boldsymbol{\alpha}) = \min_{J \in \mathcal{J}} B_J \log(\lambda_J/\mu_J)$*. Moreover, the infimum in the definition of* $\gamma_{\text{CB}}(\boldsymbol{\alpha})$ *in Lemma 2.1 is achieved by the following* $\mathbf{f}^*$*: for any* $J^* \in \text{argmin}_{J \in \mathcal{J}} B_J \log(\lambda_J/\mu_J)$*,*

$$f^*_{j'k} \triangleq \begin{cases} \hat{\phi}_{j'k} \lambda_{J^*}/\mu_{J^*} & \text{for } j' \in J^*, k \notin \partial(J^*)\,, \\ \hat{\phi}_{j'k} \mu_{J^*}/\lambda_{J^*} & \text{for } j' \notin J^*, k \in \partial(J^*)\,, \\ \hat{\phi}_{j'k} & \textit{otherwise}\,. \end{cases} \tag{2.24}$$

We can now prove the main theorem.

Proof of Theorem 2.1. Lemma 2.1 along with the explicit expression for $\gamma_{\text{CB}}(\boldsymbol{\alpha})$ provided by Lemma 2.4 yields the converse result (part 2 of the theorem).

Achievability (part 1 of the theorem) follows from Corollary 2.1 along with the explicit expression for $\gamma_{\text{CB}}(\boldsymbol{\alpha})$ provided by Lemma 2.4.

□

2.6 Application to Shared Transportation Systems

In this section we discuss the application of our findings to shared transportation systems including ride-hailing and bike sharing systems, focusing on assignment control. In these systems, for each customer (demand unit), the platform must assign a vehicle (supply unit) which is sufficiently close to their origin location, and this limited flexibility leads to the compatibility graph G in our model. (In bikesharing, customers are willing to walk only a certain amount for pickup; within these constraints, they do respond to suggestions to prefer a given pickup location as in the Bike Angels program of CitiBike; see Section 2.1.2.) The number of bikes in a bikesharing system is typically held constant as in our model, and in ride-hailing drivers typically do a substantial number of trips in a session,[14] and so it is common for theoretical investigations of tactical control levers to make the approximation that cars do not enter or leave the system, e.g., [2, 22]. Shared transportation platforms typically aim to meet as much demand as possible.[15]

Notably, in shared transportation systems, a supply unit must spend positive time serving a demand before becoming available again at the destination. In Section 2.6.1, we incorporate travel times into our theory and show that SMW policies retain their superior performance and ensure loss which decays exponentially in K. In Section 2.6.2, we provide a summary of simulation experiments for ridehailing based on New York City yellow cab data. The simulation results validate our theoretical results and demonstrate excellent performance of our policies (a full description is provided in Appendix B.10).

[14] For example, the average number of trips per session is over 12 in New York City `https://toddwschneider.com/dashboards/nyc-taxi-ridehailing-uber-lyft-data/`.

[15] Though the formal objective in Section 2.2 was to maximize the fraction of demand served, note that all our results are unchanged if the platform is payoff-maximizing where the payoff of serving a demand depends on the demand's origin and destination. This is because we perform a large deviations analysis, and the payoff values have no impact on the large-deviation asymptotics.

Finally in Section 2.6.3 we briefly discuss additional aspects of ride-hailing and bike sharing systems.

2.6.1 Incorporating Travel Delays

In this subsection, we relax the assumption that supply units move instantaneously between nodes by adding travel delays. Even in the presence of travel delays, we will show that any SMW policy with scaling parameters $\boldsymbol{\alpha} \in \text{relint}(\Omega)$ achieves exponential decay of the demand loss probability in the large market regime (the practically relevant regime).

We first describe the model with travel delays. The following model inherits all the components of the model defined in Section 2.2 where K is the number of supply units, except that it has an enlarged state space to keep track of in-transit supply units, and additional parameters to characterize travel times.

Model with travel delays. Following a standard way to model travel delays which preserves tractability [see, e.g., 59, 2, 1], we assume that the travel delays of serving demand units are independent random variables drawn from exponential distributions with means which depend on the source and destination of the demand. Let the mean travel time from node $j' \in V_D$ to node $k \in V_S$ be denoted by $\tau_{j'k} \in \mathbb{R}_+$. We assume the τs do not depend on K. We make the simplifying assumption that pickup remains instantaneous, because travel times between neighboring locations are short relative to travel times to all other locations. The primitives of the extended model are $(G, \hat{\boldsymbol{\phi}}, \boldsymbol{\tau})$ and the demand type distribution is again $\boldsymbol{\phi} = \frac{\hat{\boldsymbol{\phi}}}{\mathbf{1}^{\mathrm{T}}\hat{\boldsymbol{\phi}}\mathbf{1}}$.

The augmented state space. The state of the K-th system is now $(\mathbf{X}^K(t), \mathbf{Y}^K(t))$, where $X_i^K(t)$ is the number of *available* supply units at (supply) node i at time t, and $Y_{j'k}^K(t)$ is the number of supply units *in transit* from node j' to node k at time t. Note that the travel delays follow exponential distributions, which have the memoryless property, and therefore $(\mathbf{X}^K(t), \mathbf{Y}^K(t))$ fully characterizes the system state.

Large market regime. As before, we consider the large market regime where the number of supply units K and the demand arrival rates $\hat{\boldsymbol{\phi}}^K \triangleq K\hat{\boldsymbol{\phi}}$ scale up proportionally. Since the mean travel times $(\tau_{j'k})_{j' \in V_D, k \in V_S}$ do not depend on K, if a $\Theta(1)$ fraction of demand is served on average, a $\Theta(K)$ number of supply units is in transit at any time, on average, meaning that an $\Theta(1)$ fraction of supply units is in service, consistent with the reality in shared transportation.

In order to order to serve (almost) all the demand, we need sufficiently many supply units. By Little's law, if all demand units are served, the expected number of in-transit supply units is $K \sum_{j' \in V_D} \sum_{k \in V_S} \hat{\phi}_{j'k} \tau_{j'k}$. This number must be smaller than K to satisfy all demand even if stochasticity is ignored. In order to obtain an exponentially small loss despite stochasticity, we will need a slightly stronger assumption:

Assumption 2.4. *The model primitives* $(G, \hat{\boldsymbol{\phi}}, \boldsymbol{\tau})$ *satisfy* $\sum_{j' \in V_D} \sum_{k \in V_S} \hat{\phi}_{j'k} \tau_{j'k} < 1$.

Let $\beta \triangleq 1 - \sum_{j' \in V_D} \sum_{k \in V_S} \hat{\phi}_{j'k} \tau_{j'k}$. Here β is the proportion of free supply units if all demands are served, and $1 - \beta$ is the ideal utilization rate (the utilization rate if all demands are served). Here utilization rate is the average proportion of time during which a supply unit is engaged in serving demand. Assumption 2.4 requires that $\beta \in (0, 1)$, which is consistent with the reality in shared transportation, e.g., the ride-hailing industry in New York City has an average driver utilization rate of 58% (Parrott and Reich 2018, NYC TLC and DoT 2019), i.e., on average 42% of drivers are free at any given time (moreover, most of these free drivers are not travelling to pick up a passenger[16]). In most bikesharing systems, the fraction of bikes in transit at any time is typically quite small (under 10%).[17]

The following is our main result for the setting with travel delay. For any assignment

[16]NYC TLC and DoT (2019) reports that the average trip duration is 20 minutes, and for each trip that occurs a driver spends nearly 14 minutes "cruising" (free), and less than half of that time, about 5.5 minutes, is the driver traveling to pick up a passenger. Thus a driver spends roughly 8 minutes waiting for their next trip.

[17]The report `https://nacto.org/bike-share-statistics-2017/` tells us that U.S. dock-based systems produced an average of 1.7 rides/bike/day, while dockless bike share systems nationally had an average of about 0.3 rides/bike/day. Average trip duration was 12 minutes for pass holders (subscribers) and 28 mins for casual users. In other words, for most systems, each bike was used less than 1 hour per day, which implies that less than 10% of bikes are in use at any given time during day hours (in fact the utilization is below 10% even during rush hours).

policy U, define the pessimistic performance measure $\gamma_{\mathrm{p}}(U)$ by (2.4).

Theorem 2.2 (Result with Travel Delays)**.** *Consider any network with travel delays* $(G, \hat{\boldsymbol{\phi}}, \boldsymbol{\tau})$*. If the network satisfies Assumptions 2.1, 2.2, 2.3 and 2.4, then for any* $\boldsymbol{\alpha} \in \mathrm{relint}(\Omega)$*,* SMW*(*$\boldsymbol{\alpha}$*) achieves exponential decay of the demand loss probability with strictly positive demand loss exponent, i.e.,* $\gamma_{\mathrm{p}}(\mathrm{SMW}(\boldsymbol{\alpha})) > 0$*.*

Theorem 2.2 shows that a key finding obtained from the analysis in previous sections (where there is no travel delay), i.e., that SMW policies achieve exponentially decaying demand loss probability as the number of supply units increases, is preserved when delay is incorporated. The scaling regime is the natural large market regime, along with the natural assumption that the system has a fleet size (of supply units) that is strictly larger than what is necessary to satisfy all demand (Assumption 2.4). Thus, SMW policies are able to deploy excess supply to effectively manage the stochasticity caused by travel time and demand uncertainty in the system.

Meanwhile, the negative results in Section 2.4.2 on state-independent policies and naive state-dependent policies are also preserved with travel delay, i.e., any state-independent policy can only achieve polynomially decaying demand loss and moreover (typically) fails asymptotic optimality if exact demand arrival rates are not known, and similarly a naive state-dependent policy can incur $\Omega(1)$ demand loss.

Remark 2.2 (State-independent/naive state-dependent policies remain inferior and utilization rate remains high)**.** *Augment the system in Propositions 2.1, 2.2 and 2.3 (and Example 2.4) to incorporate travel delays* $\boldsymbol{\tau}$ *as above. Then Propositions 2.1, 2.2 and 2.3 (and the claim in Example 2.4) continue to hold, and the proofs are unchanged.*

Thus, SMW policies remain substantially superior to alternative policies under travel delays.

We prove Theorem 2.2 in Appendix B.8. Similar to the previous analysis, the proof of Theorem 2.2 is based on a novel Lyapunov analysis. The analysis is more involved than the one in Section 2.5 because of the enlarged state space. For each $\boldsymbol{\alpha} \in \mathrm{relint}(\Omega)$, we construct a Lyapunov function that augments the prior Lyapunov function (see Definition 2.7) with additional terms that capture how much the number of in-transit supply units deviate from their typical values. We show that in the fluid limit, the Lyapunov function exhibits a strictly negative drift if the current state is not at its unique minimum. Using similar methodology as in Section 2.5, we show that the demand loss exponent can be lower bounded by a variational problem (more complicated than the one in Section 2.5) that has strictly positive value, leading to Theorem 2.2.

2.6.2 Simulation experiments

We use NYC yellow cab data (to estimate demand) and Google Maps (to estimate travel times) to simulate SMW-based dispatch policies in an environment that resembles the real-world ride-hailing system in Manhattan, New York City. In the interest of space, we provide only a brief summary of these experiments here and refer the interested reader to Appendix B.10 for a full description.

Our theoretical model in Section 2.2 made several simplifying assumptions:

1. Service is *instantaneous* (i.e., vehicles travel to their destination with no delay).
2. Pickup is *instantaneous* (i.e., vehicles travel to matched customers with no delay).
3. The objective is to minimize lost demand *in steady state* (though our characterization extends to transient performance as shown in Appendix B.4.4).

We relax these assumptions one by one in our numerical experiments. We study three settings: (i) steady state performance with *Service times* (Section B.10.2); (ii) steady state performance with *Service+Pickup times* (Section B.10.3); and (iii) *Transient* performance with *Service+Pickup times* (Section B.10.4). For the second and third settings, we modify SMW policies heuristically to incorporate pickup times. In each case, we let the number of cars in the system be only slightly ($\sim 3\%$) above the "fluid requirement" (see Appendix B.10.5 for a formal definition of the fluid requirement) to meet demand, and find that we are able to meet almost all demand nevertheless (the number of free cars in real systems is typically much larger and hence the real problem is easier along this dimension, see the paragraph following Assumption 2.4 in Section 2.6.1).

A highlight of SMW policies is that they are a simple family of policies with a manageable number of parameters (one per location). We propose a simulation-based optimization approach to choose the scaling parameters α in a practical setting.

Summary of findings (Appendix B.10). Consistently across all three settings, we find that the vanilla MaxWeight policy, which requires no knowledge of the demand arrival rates, outperforms static (fluid-based) control proposed in prior work by up to an order of magnitude, and loses very little demand even with small K (just ~ 10 free cars per location, whereas the static policy has a lot more free cars to work with since it loses so much more demand). Furthermore, in each of the settings, the SMW policy obtained using simulation-based optimization further significantly outperforms vanilla MaxWeight. Overall, we deduce that non-zero service times, non-exponential pickup times, and finite K do not diminish the effectiveness of the SMW family policies at managing the spatial distribution of supply. In addition, we observe that the simulation-based optimal scaling factors α in the Service time setting are similar to the theory-based optimal α, indicating robustness of our structural results (Section 2.4.1) to travel time.

2.6.3 Additional discussion

Role of supply as a buffer. As mentioned, less than 10% of bikes in a typical bike sharing system are in use at any time. The vast majority of bikes serve as a "buffer" against distributional mismatch between supply and demand, and *not* merely to fulfil the "service requirement". This aligns well with our focus in this paper on the role of supply as a buffer. In ride-hailing systems a larger fraction $\sim 60\%$ of cars are typically carrying passengers at any time, but this still leaves a substantial fraction $\sim 40\%$ free, and these free cars again serve as a buffer.

Empty relocation. It is quite costly for bike share system operators to relocate bikes, and they generally prefer to avoid (or minimize) this. In ride-hailing, empty relocation incurs gas costs (it also costs driver effort and causes road congestion), and may be beneficial to drivers in some settings and not in others.[18]

Incorporating empty relocation in our theory. Drivers may independently choose to relocate without a passenger, or the platform may make relevant suggestions to drivers (or incentivize drivers to relocate). For example, if CRP is violated in the absence of empty relocation, the ride-hailing platform may employ empty relocation to ensure that CRP holds.

We point out that state-independent relocation of free supply units can be seamlessly incorporated into our framework following the approach in Banerjee, Freund, and Lykouris [1, Section 5.1]: For every trip ending at node $k \in V_S$, the car is redirected to node $i \in V_S$ with probability r_{ki} for all $i \in V$, independently. Call $(r_{ki})_{k \in V_S, i \in V_S}$ the empty-relocation rule and i the "effective destination". This generalization of our model is straightforward to incorporate. Throughout the paper, the demand type distribution $\boldsymbol{\phi}$ is replaced with the "effective demand type distribution" $\boldsymbol{\phi}^{\text{eff}}$ whose definition is immediate from the empty-relocation rule: $\phi^{\text{eff}}_{j'i} \triangleq \sum_{k \in V_S} \phi_{j'k} r_{ki}$, and our entire formulation, analysis and results in Sections 2.2-2.5 remain unchanged. Section 2.6.1 incorporating travel delays also extends unchanged with the modified definition $\beta \triangleq 1 - \sum_{j' \in V_D} \sum_{k \in V_S} \sum_{i \in V_S} \hat{\phi}_{j'k} r_{ki} (\tau_{j'k} + \tau_{ki})$ and the assumption that this $\beta > 0$ in place of Assumption 2.4.

Future directions related to bike sharing. Our model in Section 2.2 captures pickup flexibility in dockless bike-sharing systems (e.g., Mobike in China, the world's largest shared bicycle operator by number of bicycles). Beyond our model, bike sharing may afford the platform the additional control lever of suggesting to customers where to drop off their bike, in which case we expect that SMW policies retain their guarantees with the recommended dropoff location being the location near the destination with the fewest (scaled) number of bikes. In docked bike-sharing systems (e.g., CitiBike in New

[18]For instance, this online article by Uber data scientists `https://www.uber.com/newsroom/semi-automated-science-using-an-ai-simulation-framework` finds that "...when dispatch distances are relatively longer, drivers maximize their earnings by using less gas by remaining stationary between trips" instead of gravitating to high demand areas, and that this behavior causes only a few additional trips to be lost.

York City), there is an additional wrinkle, namely, stations have a limited number of docks, and a bike cannot be dropped off at a location if no dock is available. We are optimistic that our analysis can be extended to such a setting, leading to generalized SMW policies which seek to ensure that both bikes and free docks remain available throughout the network.

2.7 Application to Scrip Systems

Scrip systems allow agents to exchange services like babysitting, and have been proposed as a way to improve the functioning of kidney exchanges (here hospitals play the role of agents). In a scrip system, a fixed amount of artificial currency (scrips) circulates among a set of agents, and when agent i services a request by agent k, then agent k "pays" agent i in scrip. Given a service request, the platform has limited flexibility in assigning the provider since, typically, only a subset of agents are able to provide the requested service. A loss occurs when an agent runs out of scrips and is hence unable to request service. We show that with only cosmetic modifications, our model and results translate fully to a model of a scrip system with heterogeneous services, thus providing novel prescriptive insights into dynamic assignment control of such systems. We show that for any scrip system such that CRP (formally reintroduced for this application later) holds, we can construct a family of simple service provider selection rules, which we name *Scaled Minimum Scrips* (SMS) policies, and prove a very strong performance guarantee analogous to Theorem 2.1 for these policies. In particular, SMS policies achieve exponentially small loss under complete resource pooling, and moreover, there is an SMS policy (which we characterize) which is exponent optimal among all policies.

We note that many features of our model align with real-world scrip systems. Transactions in scrip systems are typically quick, which justifies our instantaneous relocation assumption. Scrips only relocate as a result of transactions (no "empty" relocation). The number of scrips is typically held nearly constant over significant periods of time. Finally, the CRP assumption appears reasonable for many scrip systems: In the proposed scrip system between hospitals for kidney exchanges [26], approximate similarity of patient pools across hospitals and partial flexibility in matching donor-patient pairs with each other should ensure CRP. One would also expect CRP to hold for scrip systems in contexts like babysitting, as long as participants make themselves available as providers sufficiently often.

2.7.1 Model of Scrip Systems

We now provide a detailed description of our model of a scrip system.

Service exchange. The set of primitives is the same as in the previous model, i.e., it consists of a compatibility graph $G(V_S \cup V_D, E)$ and Poisson arrivals with a demand arrival rate matrix $\hat{\boldsymbol{\phi}}$ and consequent demand type distribution (normalized demand arrival rate) matrix $\boldsymbol{\phi} = \hat{\boldsymbol{\phi}}/(\mathbf{1}^{\mathrm{T}}\hat{\boldsymbol{\phi}}\mathbf{1})$ (let $m = |V_S|, n = |V_D|$). Here V_S is the finite set of agents, and V_D is the finite set of heterogeneous *types* of service. Each agent has a skill set, i.e., the service types he[19] can provide. The skill set structure is modeled by the *skill compatibility graph* G (see Figure 2.5 for an illustration). The neighborhood of $i \in V_S$ in G is his skill set, which is denoted by $\partial(i) \subseteq V_D$. The neighborhood of $j' \in V_D$ in G consists of the providers of type j' service, which is denoted by $\partial(j') \subseteq V_S$.

The main difference between the current model and the previous model is in the types of requests (i.e., demand). In the previous model, each demand originates from a demand node and has a supply node destination. The situation is reversed here: each service request originates from an agent (i.e., "supply node") and requires a certain service type (i.e., "demand node"). Therefore, the arrival rate matrix $\boldsymbol{\phi}$ is of dimension $m \times n$, and $\phi_{ij'}$ is the probability of a request to be of type (i, j') requests, i.e., it comes from agent i and requests type j' service. We assume that agent i does not request service types in $\partial(i)$ (i.e., service types belonging to i's own skill set); formally, $\phi_{ij'} = 0$ for all $i \in V_S, j' \in \partial(i)$. (This assumption does not impose any restriction, since, if $i \in \partial(j')$ but

[19]For expository simplicity, we refer to an agent as "he" and the central planner as "she".

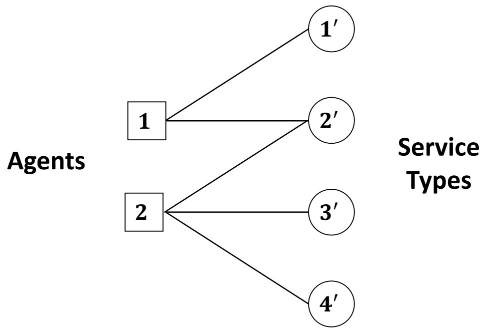

Figure 2.5: An example of skill compatibility graph in a service exchange with two service types and four agents.

i wants to request service type j', one can formally define an additional service type k' such that $\partial(k') = \partial(j')\backslash\{i\}$ and classify the request as type (i, k').) We also assume that each agent has a positive arrival rate of requesting *some* service type.

Scrips. There are a fixed number (denoted by K) of scrips in the K-th system, which are distributed among the agents. Denote the number of scrips each agent has at time t as $\mathbf{X}^K(t) = [X_1^K(t), \cdots, X_m^K(t)]$, hence $\mathbf{X}^K(t) \in \Omega_K$ where Ω_K is defined in Section 2.2.

We informally point out that there is a natural constraint on the total number of scrips a system operator can introduce: Whereas it is tempting to think that the efficiency of a scrip system can be increased simply by increasing the total number of scrips in circulation, this is the case only up to the point where the system experiences a "monetary crash", where money is sufficiently devalued that no agent is willing to perform a service; see, e.g., [51].

Service provider selection rule. The central planner's control lever is the *provider selection rule*: when a request of type (i, j') arrives, the planner chooses the provider of type j' service. Subsequently, after providing the service, agent i pays a scrip to the service provider. As is typical in scrip systems, if an agent i has no scrip, then his request is lost. As in the previous model, it suffices to consider stationary policies U, which is formally defined as a sequence of mappings, indexed by the total number of scrips K, that map the current distribution of scrips $\mathbf{X}^K$ and request type (i, j') to $\partial(j') \cup \{\emptyset\}$.

Let t_r be the r-th service request arrival epoch after time 0. Denote the state of the system just before t_r by $\mathbf{X}^K(t_r^-)$ (the initial state is $\mathbf{X}^K(0)$). Now suppose the platform uses an assignment policy U, and the r-th request comes from agent $o[r]$ and the requested service type is $d[r]$. Let $S[r] \triangleq U^K[\mathbf{X}^K(t_r^-)](o[r], d[r])$ be the chosen service provider (potentially $\emptyset$). Formally,

$$\mathbf{X}^K(t_r) \triangleq \begin{cases} \mathbf{X}^K(t_r^-) - \mathbf{e}_{o[r]} + \mathbf{e}_{S[r]} & \text{if } S[r] \in V_S \,, \\ \mathbf{X}^K(t_r^-) & \text{if } S[r] = \emptyset \,. \end{cases}$$

Performance measure. We consider a central planner who tries to maximize the fraction of requests served. We define the *optimistic* and *pessimistic* performance measures in exactly the same way as in (2.1) and (2.2). Similarly, for policy U, we define *demand-loss exponents* $\gamma_{\mathrm{o}}(U)$ and $\gamma_{\mathrm{p}}(U)$ in the same way as in (2.3) and (2.4).

Complete Resource Pooling condition (for scrip systems). We require the following CRP condition on the network primitives G and $\boldsymbol{\phi}$ for our main result in this section.

Assumption 2.5. *We assume that for all subsets $I \subsetneq V_S$ where $I \neq \emptyset$, it holds that $\lambda_I > \mu_I$ for $\lambda_I \triangleq \sum_{i \notin I} \sum_{j' \in \partial(I)} \phi_{ij'}$ and $\mu_I \triangleq \sum_{i \in I} \sum_{j' \notin \partial(I)} \phi_{ij'}$.*

Intuitively, Assumption 2.5 assumes that for each subset $I \subsetneq V_S$ of agents, requests (from outside I) which belong to the union of their skill sets arrive fast enough that they can earn enough scrips to finance their own service requests.[20]

Discussion of the model. The skill compatibility graph can capture intricate compatibility structures. For example, in scrip systems for kidney exchange, for each service (i.e., exchange) request, the ability of each other agent (hospital) to service the request may be thought of as stochastic or else arbitrary. Happily, arbitrary compatibilities can

[20]Let us clarify the relationship between Assumption 2.5 and the assumptions we made in the main model in Section 2.2: Assumption 2.5 is slightly stronger than Assumption 2.3 in that it requires strict inequality for *all* strict subsets of V_S and not just for subsets with $\mu_I > 0$. Though we do not need this stronger assumption for our analysis, we make it to simplify the exposition in this section by eliminating the need for other assumptions. In particular, Assumption 2.5 automatically implies connectivity (the analog of Assumption 2.1). Also, the analog of Assumption 2.2 (limited flexibility) holds automatically in the present setup since each individual agent forms a "limited flexibility" subset, i.e., for all $i \in V_S$ we have $\mu_{\{i\}} > 0$, which holds since $\forall i \in V_S \; \exists j' \in V_D$ such that $\phi_{ij'} > 0$, and moreover $\phi_{ij'} > 0 \Rightarrow j' \notin \partial(i) \Rightarrow \mu_{\{i\}} > 0$.

be captured in our framework by including a node in V_D for each element in 2^{V_S}, i.e., the power set of V_S.

2.7.2 Scaled Minimum Scrips (SMS) selection rules and main result

Leveraging the similarity between the current model and the previous model introduced in Section 2.2, we are easily able to define the following Scaled Minimum Scrip selection rule which is similar to SMW in spirit and achieves exponentially decaying demand loss. The formal definition of SMS is as follows.

Definition 2.8 (Scaled Minimum Scrip selection rule SMS($\boldsymbol{\alpha}$))**.** *Fix $\boldsymbol{\alpha} \in \mathrm{relint}(\Omega)$, i.e., $\boldsymbol{\alpha} \in \mathbb{R}^m$ such that $\alpha_i > 0\ \forall i \in V_S$ and $\sum_{i\in V_S} \alpha_i = 1$. Given system state $\mathbf{X}(t_r^-)$ just before the r-th demand arrival and for demand with type (i, j'), SMS($\boldsymbol{\alpha}$) chooses service provider*

$$\mathrm{argmin}_{k\in\partial(j')} \frac{X_k(t_r^-)}{\alpha_k}$$

if $X_i(t_r^-) > 0$; otherwise the request is lost. (If there are ties when determining the argmin, it assigns from the location with highest index.)

The following performance guarantee similar to Theorem 2.1 holds for Scaled Minimum Scrip($\boldsymbol{\alpha}$) under the CRP condition (Assumption 2.5).

Theorem 2.3 (Result for Scrip Systems)**.** *For any scrip system $(G, \boldsymbol{\phi})$ satisfying Assumption 2.5, we have:*

1. ***Exponentially small loss under any SMS policy**: For any $\boldsymbol{\alpha} \in \mathrm{relint}(\Omega)$, SMS($\boldsymbol{\alpha}$) achieves exponential decay of the demand loss with exponent,*

$$\gamma(\boldsymbol{\alpha}) = \min_{I\subsetneq V_S, I\neq\emptyset} B_I \log\left(\frac{\lambda_I}{\mu_I}\right) > 0\,, \tag{2.25}$$

$$\text{where}\quad B_I \triangleq \mathbf{1}_I^{\mathrm{T}}\boldsymbol{\alpha}\,, \qquad \lambda_I \triangleq \sum_{i\notin I}\sum_{j'\in\partial(I)} \phi_{ij'}\,, \quad \text{and} \quad \mu_I \triangleq \sum_{i\in I}\sum_{j'\notin\partial(I)} \phi_{ij'}\,. \tag{2.26}$$

2. ***There is an exponent optimal SMS policy****: Under* any *policy U, it must be that*

$$\gamma_{\mathrm{p}}(U) \leq \gamma_{\mathrm{o}}(U) \leq \bar{\gamma}\,, \qquad \text{where} \quad \bar{\gamma} = \sup_{\boldsymbol{\alpha} \in \mathrm{relint}(\Omega)} \gamma(\boldsymbol{\alpha})\,. \tag{2.27}$$

Thus, there is an SMS *rule that achieves an exponent arbitrarily close to the optimal one.*

The proof of Theorem 2.3 is very similar to that of Theorem 2.1; see Appendix B.7.

Remark 2.3 (Comparison with the model in [25])**.** *[25] consider the case where there is only one type of service which all agents can provide (i.e., G is a star graph), and $\phi_{ij'}$ is equal for all agents i. On one hand, we significantly generalize their model by considering heterogeneous services, asymmetric service request arrivals, and general skill compatibility graphs. They show that the* minimum scrip *selection rule, a special case of our SMS rule, is optimal for their symmetric setting, whereas we show that the family of SMS selection rules achieve exponentially small demand loss and that there exists an SMS rule that is globally exponent-optimal. On the other hand, our analysis of scrip systems is meant to illustrate the versatility of SMW type policies, hence we only focused on the central planner setting and leave a study of the incentives of agents for future work.*

2.8 Discussion

In this paper we study state-dependent assignment control of a shared transportation system modeled as a closed queueing network. We introduce a family of state-dependent assignment policies called Scaled MaxWeight (SMW) and prove that they have superior performance in terms of maximizing throughput, comparing with state-independent policies including previously proposed policies. In particular, we construct an SMW policy that (almost) achieves the optimal large deviation rate of decay of demand loss. Our analysis also uncovers the structure of the problem: given system state, demand loss is most likely to happen within state-dependent critical subsets of locations. The optimal SMW policy protects all critical subsets simultaneously.

SMW policies are simple and explicit, and hence have the potential to influence practice. We discuss two applications: Towards shared transportation applications, we show the SMW policies continue to have exponentially small loss if there are positive travel times, and obtain promising simulation results in a realistic ridehailing environment. We also also provide a model of a scrip system, and show that our entire formulation and results translate to that model with only cosmetic changes, leading us to propose Scaled Minimum Scrip (SMS) policies for service provider assignment in such systems. Our work may inspire similar analyses in open networks, e.g., obtaining exponent optimal controls when there is a shared finite buffer (e.g., a common waiting room) for multiple queues.

Part II:

Equilibria Analysis in Matching Markets

CHAPTER 3

In Which Random Matching Markets Does the Short Side Enjoy an Advantage?

3.1 Introduction

Ashlagi, Kanoria and Leshno [62] found that in a matching market with n men and $n+1$ women, and uniformly randomly complete preference lists, independent across agents, there is a nearly unique stable matching, where the average rank of men for their wives is just $\log n(1+o(1))$ (the same as it would have been under random serial dictatorship, where each man in turn selects their favorite remaining woman), whereas the average rank of women for their husbands is $\frac{n}{\log n}(1+o(1))$. For example, with $n = 1,000$, men get matched to their seventh most desired woman, whereas women are matched to only their 145th most preferred man. Of course the situation is completely reversed if, instead, there are 999 women, while the number of men is still $1,000$. This led [62] to conclude "... *we find that matching markets are extremely competitive, with even the slightest imbalance greatly benefiting the short side.*"

Meanwhile, over the past two decades, a large number of real world matching market datasets from deferred acceptance (DA) based clearinghouses have become available to different researchers in the field, e.g., from centralized labor markets like the National Residency Matching Program (NRMP) [63] and the Israel Psychology Masters Match [64], college admissions [e.g., 65], and school choice [e.g., 66]. Since these (and other) clearinghouses run the incentive-compatible deferred acceptance (DA) algorithm, the preference rankings collected may be assumed to reflect the true underlying preferences.[1] Notably, *none* of these real world data sets exhibit a "stark effect of competition" phenomenon, i.e., we do not see an abrupt change in the stable matching for a small change in market composition. As a representative example, we provide numerical counterfactuals for high school admissions data collected in one of the major cities in the U.S.: The data includes the preference lists provided by nearly 75,000 applicants, and 700 programs with a total capacity of 73,000. To study the effect of competition, we vary the market "imbalance" across a wide range by dropping up to 20,000 students from the data (uniformly at random) at one extreme, and duplicating up to 20,000 students (uniformly at random) at the other extreme, while holding the set of programs and their capacities fixed. We numerically evaluate the effect of thus varying the number of students on the resulting allocation of programs to students under the student-proposing DA algorithm. As per the usual practice, we summarize the allocation in terms of the fraction of students who are allotted to one of their top-k most preferred programs (for $k = 1, 3$) and the fraction who are unassigned; see the solid lines in Figure 3.1. Observe that the summary statistics vary extremely smoothly and slowly in the number of students over a wide range. In other words, *we observe no stark effect of competition in real world data*, which is at odds with the aforementioned conclusion of [62].

The stochastic model of matching markets considered in [62] is often called a "random matching market"; one where agents have independent, complete and uniformly random preference lists over the other side. The model was introduced by Knuth [71], heavily studied by Pittel and others [72, 73, 74] and this model (and variants) remains a workhorse for research in the area [e.g., 75, 76] and even for deriving operational insights, e.g., which tie-breaking rule to use [77], making it imperative that we understand how its predictions

[1]Incentive compatibility of DA for the proposing side was established by [67]. For the receiving side, approximate IC is strongly suggested by the findings of [68] and [62], among others, and the mechanism further seems very hard to manipulate in most practical settings. However, it is worth noting [69] has found empirical evidence of incorrect preference reporting in certain situations, while [70] suggests that participants may not report options they like if those options are infeasible. Since the extent and nature of misreporting in our data (if any) is unclear, we simply assume the preference reports to be truthful.

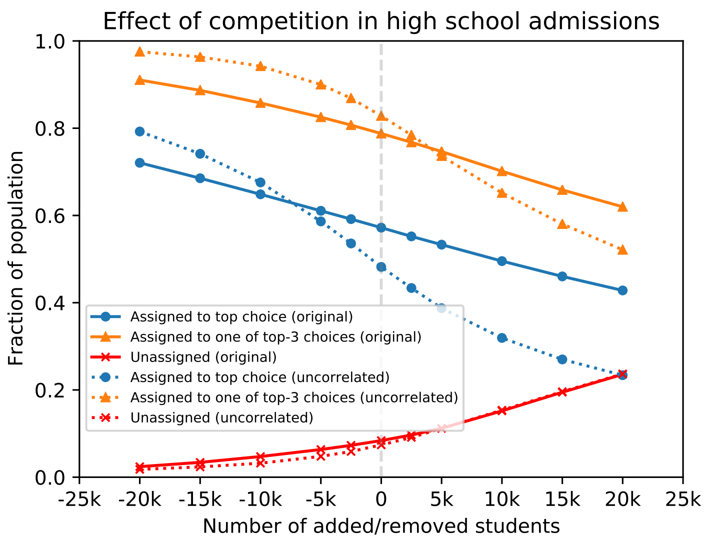

Figure 3.1: Fraction of students who are assigned to one of their top-k most preferred programs (for $k = 1, 3$) and the fraction of students who are unassigned, as a function of the number of students removed or duplicated uniformly at random. Simulations are based on the actual high school admissions data containing 75k applicants and 73k seats across 700 programs (averaged across 100 realizations). The solid lines use the student preference rankings and program priorities in the original data, and implement a single tie-breaking rule[3]. The dotted lines are based on randomizing preferences and priorities: Each student's preference list has unchanged length but its entries are drawn without replacement with the sampling probability of each program being proportional to the number of students who have applied to it in the original dataset, and each program uses a uniformly random and independent priority ordering over students.

might depart from reality, and the role played by each of the stylized assumptions in the model.

It is natural to ask whether correlation in preferences in real markets is the reason that they do not exhibit a strong effect of competition. Indeed, if preferences on the "men" side of the market are fully correlated (i.e., all agents have the same preference ordering) while the other "women" side has arbitrary preferences, then there is a unique stable matching which can be computed by running serial dictatorship by women (women serially pick their favorite available man), in the order of the womens' universal ranking by men. One would expect this unique stable matching to transform smoothly as the

[3]Under single tie-breaking, each student receives a random lottery number at the beginning of matching process, which is used by all programs for breaking ties between applicants with the same priority.

number of agents on one side of the market is varied. To test whether correlation is indeed the reason we see only a weak effect of competition in the actual high school admissions data introduced above, we ran an additional experiment with "randomized" preferences: We took only the student and program "degrees" (i.e., preference list lengths and number of times the program is listed) from the data, generated both student preference lists and program priorities independently at random, and studied the resulting allocation as a function of the number of students; see the dotted lines in Figure 3.1 (the distribution over preferences is precisely specified in the caption of the figure). While the effect of competition under randomized preferences is somewhat stronger than in the original data,[4] it bears no resemblance to the abrupt phase transition found by [62]. Thus it appears that even without correlation in preferences, and despite being very well connected,[5] realistic markets seem to lack a strong effect of competition. This prompts us to investigate the effect of the level of connectivity in the random matching market model on the effect of competition.

Model. Our model generalizes the random matching market model to allow "partially connected" markets with each agent having an average degree d in a random (undirected) connectivity graph. Each agent has a preference ranking over only their neighbors in the connectivity graph. We assume there are $n+k$ men and n women, where the "imbalance" k may be positive or negative but we restrict to "small" imbalances $|k| = o(n)$. For technical convenience, the random graph model we work with is one where each man is connected to a uniformly random subset of exactly d women, independent of other men.[6]

[4] Define the elasticity of the fraction of students who get their top choice as the percent change in this fraction for every 1% change in the number of students. Near the status quo number of students, we find that the elasticity of the high school market is nearly -1.0, whereas the elasticity of the randomized high school market is nearly -2.8.

[5] In the high school admissions setting, both without and with randomization of preferences, over 97% of student pairs are within two hops of each other and nearly 100% are within three hops of each other (a pair of students is within one hop of each other if they list a program in common).

[6] As a result, each woman has Binomial$(n+k, d/n) \xrightarrow[n\to\infty]{\text{d}}$ Poisson(d) neighbors where $\xrightarrow{\text{d}}$ denotes convergence in distribution. Throughout the paper we will restrict attention to $d = \omega(1)$, as a result of which Poisson$(d) \xrightarrow{\text{p}} d$, i.e., the degree of each woman is also very close to d, and so the asymmetry between the two sides in the model is mainly technical.

Main findings. We characterize stable matchings as a function of d and the number of women n and find that the short side enjoys a significant advantage only for d exceeding $\log^2 n$: For moderately connected markets, specifically any d such that $d = o(\log^2 n)$ and $d = \omega(1)$ and large n, we find that *there is no stark effect of competition*, namely, the short and long sides of the market are almost equally well off (for $|k| = O(n^{1-\epsilon})$ market imbalance), with agents on both sides getting a $\sqrt{d}(1 + o(1))$-ranked partner on average. Notably, this regime extends far beyond the connectivity threshold (above which the connectivity graph is connected with high probability) of $d = \Theta(\log n)$. On the other hand, for densely connected markets, specifically for any $d = \omega(\log^2 n)$ and large n, we find that there is a stark effect of competition: assuming a small imbalance $|k| = o(n)$, the short side agents get a partner of rank $\log n$ on average, while the long side agents get a partner of (much larger) rank $d/\log n$ on average. Numerical simulations of our model confirm the theoretical predictions, and in fact further enhance our understanding: they suggest a *sharp threshold between the two regimes close to* $d \approx 1.0 \times \log^2 n$ and that *this holds even for small* n *down to* $n \gtrsim 10$. Figure 3.2 provides a schematic depicting our main findings (including the $d \approx 1.0 \times \log^2 n$ threshold between regimes suggested by numerics).

Since preference list lengths in most real markets are much below $\log^2 n$ (the latter is nearly 48 for $n = 1000$ and nearly 117 for $n = 50000$), and correlation in preferences only appears to reduce the effect of competition (see Figure 3.1 for indicative evidence), our findings may explain why real world matching datasets do not exhibit a strong effect of competition.

We highlight that the "no stark effect" regime includes well connected markets for connectivity in the range $d \in (\Theta(\log n), o(\log^2 n))$. This is in sharp contrast to buyer-seller market, where, roughly, connectivity implies a stark effect of competition where the short side of the market captures all the surplus (see Remark 3.1 in Section 3.3 for a detailed discussion). In particular, our results imply that the informal claim in [62] of strong *similarity* between the two kinds of markets is incorrect for moderately connected markets (parallel to the fact that most real world matching markets are well connected but do not exhibit a stark effect of competition).

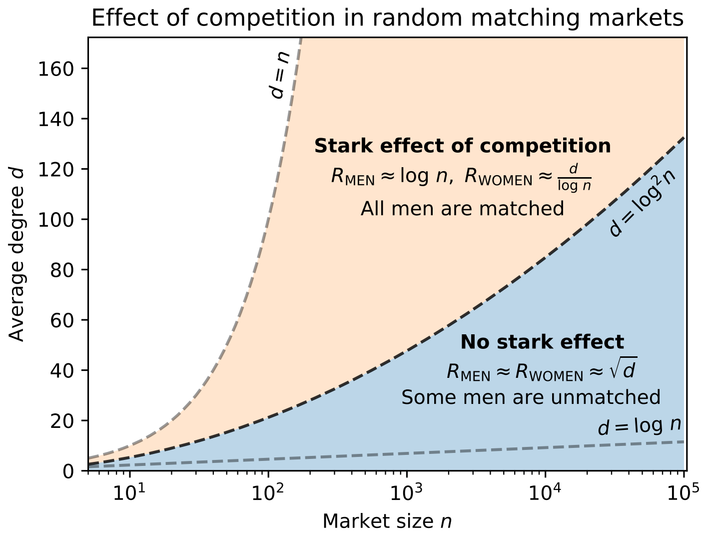

Figure 3.2: Schematic showing the two "competitiveness" regimes for partially connected random matching markets with small imbalance, n women, men being on the short side, and connectivity (average preference list length) d. R_{MEN} denotes the average rank of men for their wives, and R_{WOMEN} the average rank of women for their husbands.

Intuition for our findings. We now provide the high level intuition behind our main results. [62] showed a stark effect of competition for fully connected markets $d = n$. We find that as d is decreased, this phenomenon remains intact if all short side agents are matched: if women are on the long side ($k < 0$), though the average rank of women for their husbands $R_{\text{WOMEN}} \approx d/\log n$ decreases as d decreases, it remains true that men are significantly better off than women $R_{\text{WOMEN}}/R_{\text{MEN}} \geq 1 + \Omega(1)$. As d falls below a certain threshold, a positive number of men remains unmatched with high probability. The threshold turns out to be $\log^2 n$, corresponding to the fact that the maximum number of proposals made by any man in the fully connected random market is $\Theta(\log^2 n)$; see [78].

But does a few agents remaining unmatched have any bearing on the stark effect of competition phenomenon? A priori it is unclear that this should be the case. After all, it is easy to construct matching markets such that some short side agents remain unmatched, but where the short side is nevertheless significantly better off than the long side.[7] Remarkably, we find that in *random* matching markets there is *no stark effect of competition if some short side agents remain unmatched.* We now provide some informal intuition why random matching markets have this property: Clearly, due to the matching constraint the number of unmatched men must be exactly k plus the number of unmatched women. Hence, assuming a small imbalance k (the balanced market with $k = 0$ being a special case), the number of unmatched agents on the two sides must be nearly the same. But the number of unmatched men should grow with R_{MEN} (the more proposals men need to make in men-proposing DA, the larger the number of men that will reach the end of their preference list), whereas the number of unmatched women should similarly grow with R_{WOMEN} (e.g., one can consider women-proposing DA, and assume — or prove separately — that, as usual, the WOSM is close to the MOSM). We deduce that we should have $R_{\text{MEN}} \approx R_{\text{WOMEN}}$ in the $d \ll \log^2 n$ regime, i.e., there is almost no advantage from being on the short side of the market.

Next, we provide more detailed quantitative intuition leading informally to the sharp estimates of R_{MEN} and R_{WOMEN} in our characterization of moderately connected markets. This intuition is based on a detailed heuristic picture of the stable outcome in a random matching market (we do not formalize the full detailed picture in this paper, and instead prove our main theorem via a "shortcut" described below). Intuitively, both R_{MEN} and the number of unmatched men δ^m should be governed by the (endogenous) probability p_{MEN} that a neighboring woman j (independently of other women) is "interested" in given man i (the woman j is said to be interested if she receives no proposal which she prefers to i): in particular, the rank of man i for his wife (his most preferred woman who accepts his proposal) should be distributed as Geometric(p_{MEN}) truncated at d, leading to $R_{\text{MEN}} \approx 1/p_{\text{MEN}}$ (assuming $1/p_{\text{MEN}} \ll d$) and $\delta^m \approx n\mathbb{P}(\text{Geometric}(p_{\text{MEN}}) > d) = n(1 - p_{\text{MEN}})^{-d} \approx n \exp(-dp_{\text{MEN}})$. Analogously for women, letting p_{WOMEN} denote the (en-

[7] For instance, consider a densely connected random matching market with the modification that a few agents on the short side have empty (or short) preference lists. The latter agents will remain unmatched, but the short side agents will nevertheless have a much smaller average rank for the their partners than the long side agents.

dogeneous) probability that woman j receives a proposal from each neighboring man i, we expect $R_{\text{WOMEN}} \approx 1/p_{\text{WOMEN}}$ and $\delta^w \approx n\exp(-dp_{\text{WOMEN}})$. For k small, we have that both sides must have nearly the same number of unmatched agents $\delta^w \approx \delta^m$ and hence $p_{\text{WOMEN}} \approx p_{\text{MEN}}$ and $R_{\text{MEN}} \approx R_{\text{WOMEN}}$. But we can further get quantitative estimates: the average number of proposals received by women is nearly the same as the average number of proposals made by men $(n+k)R_{\text{MEN}}/n \approx R_{\text{MEN}} \approx 1/p_{\text{MEN}}$, and since $p_{\text{WOMEN}} \approx$ average number of proposals received/(typical length of preference list) $\approx 1/(dp_{\text{MEN}})$. We deduce that $p_{\text{MEN}} \approx p_{\text{WOMEN}} \approx \frac{1}{\sqrt{d}}$ and so $R_{\text{MEN}} \approx R_{\text{WOMEN}} \approx \sqrt{d}$ and $\delta^m \approx \delta^w \approx ne^{-\sqrt{d}}$.

Technical contributions. Our characterization of the stable matching in partially connected random matching markets (as a function of connectivity d) is novel: stable matchings have not been previously characterized either in balanced or in unbalanced random markets under partial connectivity $d < n$ and $d = \omega(1)$. (A few papers have studied the extreme case of sparsely connected markets $d = \Theta(1)$ under various preference models; see Section 3.1.1.)

Our characterization showing a stark effect of competition in densely connected markets $d = \omega(\log^2 n)$ is proved via an analysis similar to [62]. In contrast, our characterization showing no stark effect of competition in moderately connected markets $d = o(\log^2 n)$ and $d = \omega(1)$ overcomes significant technical difficulties via a novel approach as we now describe.

The main challenge we face relative to previous works studying fully connected markets [e.g., 62, 71] is the complexity in the way that DA terminates when there is a positive (but vanishing) fraction of unmatched agents on both sides of the market. Note that DA terminates when the number of women who have received at least one proposal equals the number of men who have not exhausted their lists. In the large d regime (including fully connected markets as studied in the prior literature), it is likely, assuming that men are on the short side ($k < 0$), that the event of $n + k$ distinct women each receiving at least one proposal happens before any man reaches the bottom of his list. As a result, the total number of proposals in DA can be well approximated by the solution of the coupon collector's problem where $n + k$ distinct coupons must be drawn, and key properties of the MOSM can be deduced from there. In the small to medium d regime, however, it is likely that a positive number of men have reached the end of their lists during the run of DA. As a result, to estimate the number of proposals of DA, it is necessary to get a handle on the number of men who have been rejected d times, which is considerably more complicated to analyze, especially for $d = \omega(1)$ as we consider, where the fraction of unmatched men is positive but vanishing (previous works, especially [79], have developed a machinery to handle the case of $d = \Theta(1)$ which leads to a $\Theta(1)$ fraction of unmatched agents). One of our technical contributions is resolving this difficulty. Instead of proving the detailed heuristic picture given in the previous paragraph, we control two main quantities: (i) the total number of proposals before DA terminates (this quantity is the one tracked in the related literature), and (ii) the number of unmatched men and women when DA terminates. The matching constraint tells us that the number of unmatched men is exactly k plus the number of unmatched women. Thus, to control (ii) it suffices to control the number of unmatched men. We estimate (bound) this quantity by constructing a "fake" process where a man who is accepted and then later rejected is allowed to make d additional proposals. It turns out this process is much easier to analyze and it yields a sufficiently good estimate of the number of men who end up unmatched under the assumption $d = o(\log^2 n)$.

3.1.1 Related work

The closest papers to our work are the ones studying random matching markets [71, 72, 73, 74, 62, 78]. All of these papers assume *complete* preference lists. Whereas the early papers focused on balanced random markets and found that the proposing side (in DA) has a substantial advantage, [62] and follow up papers found that in unbalanced markets, the short side has a substantial advantage. The main technical difficulty we face relative to these papers is that a positive number of agents remain unmatched on *both* sides of the market in moderately connected markets $d = o(\log^2 n)$, preventing us from directly leveraging the analogy with the coupon collector problem as in the previous works.

Notable papers by Immorlica and Mahdian and others [79, 68] show a small core (i.e., a small set of stable matchings) while working with short (constant-sized) preference lists, leading to a linear fraction of unmatched agents. Arnosti [80] and Menzel [76] characterize the (nearly unique) stable outcome in settings with constant-sized preference lists, and in particular, we expect their characterizations can be used to show that the outcome changes "smoothly" as a function of the market imbalance under short lists. In contrast to the aforementioned papers, our work restricts attention to the case $d = \omega(1)$ and indeed identifies the existence of a threshold at $d \sim \log^2 n$, as a result of which the fraction of unmatched agents in our setting is vanishing. Technically, the consequence of this phenomenon is that "rejection chains" in the progress of DA are $\omega(1)$ in length in our work, making them harder to analyze, and the (approximate) system "state" no longer has bounded dimension as in [80].

Our work belongs to a vast theoretical literature on matching markets, which began with the work of Gale and Shapley [81] introducing stable matching and the deferred acceptance algorithm, and has developed over the last six decades with major contributions by Roth, Sotomayor, and a large number of other prominent researchers [see, e.g., 82, 83]. Key combinatorial properties of stable matchings are extremely well understood for multiple decades now, and more recently, it has been generally accepted that in typical matching markets, the man optimal stable matching is nearly the same as the woman optimal stable matching [79, 68, 62], allowing one to talk about *the* stable matching in typical settings.

What still remains troublingly mysterious is the *nature* of the stable matching as a function of market primitives, especially in settings where there is a significant idiosyncratic/horizontal component to preferences and preference lists are not short (when there is a strong vertical component to preferences, the outcome is known to be approximately assortative, e.g., see [84]). [62] suggested that the outcome depends heavily on which side is the short side of the market, but innumerable datasets and the present theoretical work indicate that this is *not* the case in typical markets. The present paper aims to explain the relative lack of competitiveness of typical matching markets, and overall to take a small step towards a better understanding of how the stable matching depends on market primitives. Reasoning based on [62] has the potential to lead theorists (and perhaps practitioners) astray, given that we often want to derive operational insights, e.g., which tie breaking rule to use [77], based on the analysis of models resembling the random matching market model.

There is a robust and growing body of practical work on designing real world matching markets, especially in the contexts of school and college admissions [e.g., 65, 66, 85], and various labor markets [e.g., 63, 64]. Stability, namely, that no pair of agents should prefer to match with each other, has been found to be crucial in the design of centralized clearinghouses [86] and predictive of outcomes in decentralized matching markets [87, 88]. We are not aware of any real world matching dataset in which the short side of the market is vastly better off even if the imbalance is small. It further appears that most practitioners are aware that a platform operator cannot make one side of the market vastly better off by slightly tilting the market imbalance in favor of that side.

Organization of the paper. In Section 3.2, we introduce our model of partially connected random matching markets. In Section 3.3, we state our main theorems (Theorem 3.1 and 3.2) and discuss them. An overview of our proof of our characterization of moderately connected markets (Theorem 3.1) is provided in Section 3.4. In Section 3.5, we provide the simulation results that confirm and sharpen our theoretical predictions. Formal proofs are relegated to the appendix.

3.2 Model

We consider a two-sided market that consists of a set of men $\mathcal{M} = \{1, \ldots, n+k\}$ and a set of women $\mathcal{W} = \{1, \ldots, n\}$. Here k is a positive or negative integer, which we call the *imbalance*.

We fix a positive integer $d \leq n$ which we call the *connectivity* (or *average degree*) of the market. Each man i has a strict preference list $\succ_i$ over a uniformly random subset $\mathcal{W}_i \subset \mathcal{W}$ of $|\mathcal{W}_i| = d$ women (from among the $\binom{n}{d}$ possibilities), where the subsets $\mathcal{W}_i$ are drawn independently across men. Each woman j has strict preferences $\succ_j$ over only the men who include her in their preference list[8]

$$\mathcal{M}_j = \{i \in \mathcal{M} : j \in \mathcal{W}_i\}.$$

A *matching* is a mapping μ from $\mathcal{M} \cup \mathcal{W}$ to itself such that for every $i \in \mathcal{M}$, $\mu(i) \in \mathcal{W} \cup \{i\}$, and for every $j \in \mathcal{W}$, $\mu(j) \in \mathcal{M} \cup \{j\}$, and for every $i, j \in \mathcal{M} \cup \mathcal{W}$, $\mu(i) = j$ implies $\mu(j) = i$. We use $\mu(j) = j$ to denote that agent j is unmatched under μ.

A matching μ is *unstable* if there are a man i and a woman j such that $j \succ_i \mu(i)$ and $i \succ_j \mu(j)$. A matching is *stable* if it is not unstable.

A *random matching market* is generated by drawing, for each man i, a uniformly random preference list over $\mathcal{W}_i$ (from among the $|\mathcal{W}_i|!$ possibilities), and for each woman j, a uniformly random preference list over $\mathcal{M}_j$, independently across agents.

A stable matching always exists, and can be found using the Deferred Acceptance (DA) algorithm by Gale and Shapley [81]. They show that the men-proposing DA finds the *men-optimal stable matching* (MOSM), in which every man is matched with his most preferred stable woman. The MOSM matches every woman with her least preferred stable man. Likewise, the women-proposing DA produces the women-optimal stable matching (WOSM) with symmetric properties. All of our results will characterize the MOSM. Given the strong evidence from [79, 68, 62] and other works that the MOSM and WOSM are nearly the same in typical matching markets (with the exception of balanced and densely connected random markets, which we avoid by assuming $k < 0$ in Theorem 3.2), we omit to formally show this fact for our setting in the current version of the paper

[8]Equivalently, we sample an undirected bipartite random graph G connecting men $\mathcal{M}$ to women $\mathcal{W}$, where each man has degree exactly d and the d neighboring women of each man are selected uniformly at random and independently across men. Given G, for each agent has a strict preference ranking over all his/her neighbors in G and does not rank any other agents.

though we believe it can be done, e.g., using the method developed in [89] (the property MOSM $\approx$ WOSM is found to hold consistently in our numerical simulations of our model).

We are interested in how matched agents rank their assigned partners under stable matching, and in the number of agents who remain unmatched. Denote the rank of woman j in the preference list $\succ_i$ of man i by $\text{Rank}_i(j) \equiv |\{j' : j' \succeq_i j\}|$. A smaller rank is better, and i's most preferred woman has a rank of 1. Symmetrically, denote the rank of i in the preference list of j by $\text{Rank}_j(i)$.

Definition 3.1. *Given a matching μ, the* men's average rank of wives *is given by*

$$R_{\text{MEN}}(\mu) = \frac{1}{n+k}\left(|\bar{\mathcal{M}}(\mu)|(d+1) \; + \sum_{i \in \mathcal{M} \setminus \bar{\mathcal{M}}(\mu)} \text{Rank}_i(\mu(i)) \right) ,$$

where $\bar{\mathcal{M}}(\mu)$ is the set of men who are unmatched under μ, and the number of unmatched men *is denoted by $\delta^m(\mu)$, i.e., $\delta^m(\mu) = |\bar{\mathcal{M}}(\mu)|$.*

Similarly, the women's average rank of husbands *is given by*

$$R_{\text{WOMEN}}(\mu) = \frac{1}{n}\left(\sum_{j \in \bar{\mathcal{W}}(\mu)} (|\mathcal{M}_j| + 1) \; + \sum_{j \in \mathcal{W} \setminus \bar{\mathcal{W}}(\mu)} \text{Rank}_j(\mu(j)) \right)$$

where $\bar{\mathcal{W}}(\mu)$ is the set of women who are unmatched under μ, and the number of unmatched women *is denoted by $\delta^w(\mu)$, i.e., $\delta^w(\mu) = |\bar{\mathcal{W}}(\mu)|$.*

(Note here that if an agent is unmatched, we take the rank for the agent to be one more than the length of the agent's preference list.) By the rural hospital theorem [90], the set of unmatched agents ($\bar{\mathcal{M}}(\mu)$ and $\bar{\mathcal{W}}(\mu)$) is the same in every stable matching μ, and therefore we simply represent the number of unmatched men and women under stable matching by δ^m and δ^w respectively throughout the remainder of paper.

We remark that the only asymmetry in our model is that the lengths of men's preference lists are deterministically d, whereas each woman has $\text{Binomial}(n+k, d/n) \xrightarrow[n\to\infty]{\text{d}}$ $\text{Poisson}(d)$ neighbors where $\xrightarrow{\text{d}}$ denotes convergence in distribution. Since our theoretical analysis will assume $d = \omega(1)$, we have $\text{Poisson}(d) \xrightarrow{\text{p}} d$, i.e., the degree of each woman is also very close to d, and so the asymmetry between the two sides in the model is mainly technical.

3.3 Results

In this section we state and discuss our main results.

Before stating our results, we restate a main finding of [62] (Theorem 2 in that paper) on the structure of stable matchings in fully connected random markets. (The statement has been modified —and slightly weakened in the process— with the aim of allowing easy comparison with our main theorems.)

Theorem (Ashlagi, Kanoria, and Leshno [62], Fully connected markets)**.** *Consider a sequence of random matching markets indexed by* n*, with* $n+k$ *men and* n *women, for* $k = k(n) \in [-n/2, -1]$*, and complete preference lists on both sides of the market (i.e., connectivity* $d = n$*). For fixed* $\epsilon > 0$*, with high probability the following hold for every stable matching*[9] μ*:*

$$R_{\text{MEN}}(\mu) \le (1+\varepsilon)\left(\tfrac{n}{n+k}\right) \log\left(\tfrac{n}{|k|}\right),$$

$$R_{\text{WOMEN}}(\mu) \ge \frac{n+k}{1 + (1+\varepsilon)\left(\tfrac{n}{n+k}\right) \log\left(\tfrac{n}{|k|}\right)}\,,$$

and all men are matched.

The theorem shows that even a slight imbalance in the number of agents on the two sides of the market results in a stark effect on stable outcomes that strongly favors the agents on the short side of the market: agents on the short side are essentially able to freely *choose* their partners (as [62] explain, R_{MEN} is nearly the same as it would be under random serial dictatorship by the men), whereas agents on the long side do only a little better than being matched with a random partner. In particular, even with $k = -1$, it holds that $R_{\text{MEN}}(\mu) \le 1.01 \log n$ and $R_{\text{WOMEN}}(\mu) \ge \frac{0.99n}{\log n}$ in every stable matching,

[9] Though our definition of average rank is slightly different from that of [62], the bounds stated are nevertheless valid for our definition.

w.h.p. In the present paper, we investigate stable matchings in random *partially connected* matching markets, and compare with the above finding of [62].

Moderately and sparsely connected markets. In our first main result, we show that the short-side advantage disappears in partially connected markets (with small or zero imbalance) whose connectivity parameter d is below $\log^2 n$.

Theorem 3.1 (Moderately Connected Markets). *Consider a sequence of random matching markets indexed by n, with $n+k$ men and n women ($k = k(n)$ can be positive or negative or zero), and connectivity (average degree) $d = d(n)$, with $d = \omega(1)$ and $d = o(\log^2 n)$, and*[10] $|k| = O(ne^{-\sqrt{d}})$*. Then with high probability,*[11] *we have*

$$\left|R_{\text{MEN}}(\text{MOSM}) - \sqrt{d}\right| \leq d^{0.3},$$
$$\left|R_{\text{WOMEN}}(\text{MOSM}) - \sqrt{d}\right| \leq d^{0.3},$$
$$\left|\log \delta^m - \log\left(ne^{-\sqrt{d}}\right)\right| \leq d^{0.3},$$
$$\left|\log \delta^w - \log\left(ne^{-\sqrt{d}}\right)\right| \leq d^{0.3}.$$

Informally, in large random matching markets with average degree $d = o(\log^2 n)$ and a small imbalance $k = O(n^{1-\epsilon})$, under stable matching we have $R_{\text{MEN}} \approx R_{\text{WOMEN}} \approx \sqrt{d}$ irrespective of which side is the short side, and there are approximately $ne^{-\sqrt{d}} = \omega(1)$ unmatched agents on both sides of the market. Thus there is no short-side advantage and agents on both sides are matched to their $\sqrt{d}$-th ranked partner on average. A significant number of agents are left unmatched even on the short side, in contrast to a fully connected unbalanced matching market where all agents on the short side are matched. Though we only characterize the MOSM in the present version of the paper, we believe the same characterization extends to the WOSM as well. We give an overview of the proof of Theorem 3.1 in Section 3.4 and the formal proof in Appendix C.2.

The main intuition for Theorem 3.1 is that for $d = o(\log^2 n)$, a positive number of men remain unmatched with high probability, because they reach the end of their

[10] In particular, for arbitrary fixed $\epsilon > 0$, the result holds for any $k = k(n)$ that satisfies $|k(n)| = O(n^{1-\epsilon})$.

[11] Specifically, our characterization holds with probability at least $1 - O(\exp(-d^{1/4})) = 1 - o(1)$.

preference lists in men-proposing DA ([78] showed that some men need to go $\log^2 n$ deep in their preference lists in the fully connected market). Clearly, the number of unmatched men must be exactly k plus the number of unmatched women. Then, assuming a small imbalance k, the number of unmatched agents on the two sides must be nearly the same. But the number of unmatched men should grow with R_{MEN} (the more men need to propose, the larger the number that will reach the end of their preference lists), whereas the number of unmatched women should similarly grow with R_{WOMEN} (e.g., one can consider women proposing DA, and assume that, as usual, the WOSM is close to the MOSM). We deduce that we should have $R_{\text{MEN}} \approx R_{\text{WOMEN}}$ in the $d \ll \log^2 n$ regime. (Informal quantitative intuition leading to the precise estimates of R_{MEN} and δ^m is provided in the introduction; we avoid reproducing it here.)

We highlight that Theorem 3.1 encompasses a wide range of connectivity parameters $d = o(\log^2 n)$, which extends far beyond the connectivity threshold $d \approx \log n$ (this is the connectivity threshold in our model, the same as for Erdős-Rényi random graphs). Thus our "no stark effect of competition" result does not require a disconnected or fragmented market. Rather, the result applies even to very well connected markets.[12] This is in sharp contrast to buyer-seller markets, where, roughly, connectivity implies a stark effect of competition, as captured in the following remark.

Remark 3.1 (Connected buyer-seller markets exhibit a stark effect of competition)**.** *Consider a buyer-seller market where each of $n+k$ sellers is selling one unit of the same commodity, and each of n buyers wants to buy one unit and has value* 1 *for a unit. A bipartite graph G with sellers on one side and buyers on the other captures which trades are feasible. (This is a special case of the Shapley-Shubik assignment model [91].) We say that an unbalanced market with $k > 0$ (or $k < 0$) exhibits a stark effect of competition if, in any equilibrium, all trades occur at price* 0 *(or* 1*), i.e., the agents on the short side, namely buyers (sellers), capture all the surplus. Then we know [91] that for $k \neq 0$ the*

[12] For example, with $n = 1,000$, $\log^2 n \approx 48$. Taking $d = 10$ (much less than 48), numerics tell us that 9.6% of pairs of men are within 1 hop of each other (i.e., there is woman who is ranked by both men), and 99.98% of pairs of men are within 2 hops of each other.

market exhibits a stark effect of competition if the following requirement is satisfied:

$$\mathcal{E} \equiv \{\text{For each agent } j \text{ on the long side, there exists a matching in } G$$
$$\text{where all short side agents are matched but agent } j \text{ is unmatched}\}\,.$$

Requirement $\mathcal{E}$ is only slightly stronger than connectivity of G: Suppose, as in our model in Section 3.2, that each seller is connected to a uniformly random subset of d buyers. Under this stochastic model for G, for any sequence of k such that $1 \leq |k| = O(1)$, event $\mathcal{E}$ occurs (i.e., there is a stark effect of competition) for all d exceeding the connectivity threshold at $d = \log n$:

(i) For any $\epsilon > 0$ and $d \geq (1+\epsilon)\log n$, with high probability, G is connected and moreover, event $\mathcal{E}$ occurs, i.e., there is a stark effect of competition.

(ii) For any $\epsilon > 0$ and $d \leq (1-\epsilon)\log n$, with high probability, the connectivity graph G is disconnected (in fact a positive number of buyers have degree zero).

Numerical simulations in the Section 3.5 show that the finding in Theorem 3.1 holds up extremely well for all $d \lesssim 1.0 \log^2 n$ for realistic values of n (not just asymptotically in n for $d = o(\log^2 n)$). Now $\log^2 n$ is quite large for realistic market sizes (see Figure 3.2 in the introduction), far in excess of preference list lengths in many real markets: we have $\log^2 n \approx 48$ for $n = 1000$, 85 for $n = 10000$ and 132 for $n = 100000$. In contrast, we have $n \approx 80,000$ for the high school admissions data introduced in the Section 3.1 and preference lists have length no more than 12 (the average length is only around 6.9), $n \approx 30,000$ for the National Residency Matching Program and preference lists have length only about 11 on average. Thus, real preference list lengths are typically much smaller than $\log^2 n$. Moreover, correlation in preferences should only reduce the effect of competition (e.g., see the evidence in Figure 3.1), leading us to contend that *the vast majority of real matching markets live in the "no stark effect of competition" regime covered by Theorem 3.1.* This may explain why, in simulation experiments on real data like the one shown in Figure 3.1, only a relatively weak effect of competition is observed.

Densely connected markets. Our next main result shows that $d \sim \log^2 n$ is the threshold level of connectivity above which the finding of [62] holds true, i.e., the short side is markedly better off even in (large) markets with a small imbalance. Moreover, this benefit of being on the short side arises in conjunction with the key property that *all agents on the short side of the market are matched* (an implausible occurrence in real world markets).

Theorem 3.2 (Densely Connected Markets)**.** *Consider a sequence of random matching markets indexed by* n*, with* $n+k$ *men and* n *women, and connectivity (average degree)* $d = d(n)$*, with* $k = k(n) < 0$ *and* $|k| = o(n)$*,* $d = \omega(\log^2 n)$ *and* $d = o(n)$*. Then, with high probability, all men are matched under stable matching, and we have*

$$R_{\text{MEN}}(\text{MOSM}) \le (1 + o(1)) \log n \,,$$
$$R_{\text{WOMEN}}(\text{MOSM}) \ge (1 + o(1)) \frac{d}{\log n} \,.$$

This result shows that the short-side advantage emerges in densely connected markets even when the imbalance is small (including for an imbalance of one, i.e., $k = -1$). More specifically, when $d = \omega(\log^2 n)$, it predicts that the agents on the short side are matched to their $\log n$-th ranked partner on average whereas the agents on the long side are matched to their $\left(\frac{d}{\log n}\right)$-th ranked partner on average. Theorem 3.2 smoothly interpolates between the result in AKL [62] and our Theorem 3.1 (though the extremes $d = \Omega(n)$ and $d = \Theta(\log^2 n)$ are not covered by the formal statement in present form): as connectedness d increases, a phase transition happens at $d = \Theta(\log^2 n)$, and the short side advantage starts to emerge for $d = \omega(\log^2 n)$. The magnitude of the advantage increases as the market becomes denser. Combining Theorems 3.1 and 3.2, we conclude that, assuming a small imbalance, a short-side advantage exists if and only if a matching market is connected densely enough, and the threshold level of connectivity $d \sim \log^2 n$.

The analysis leading to Theorem 3.2 is similar to that leading to [62, Theorem 2]. The number of proposals in men-proposing DA remains unaffected; the only change is that women now have rank lists of approximate length d (instead of length $n+k$), and so, receiving about $\log n$ proposals leads to an average rank of husband of about $d/\log n$. The proof is in Appendix C.3.

3.4 Overview of the proof of Theorem 3.1

This section provides an overview of the proof of Theorem 3.1, which is our characterization of moderately connected random matching markets. Our proof uses the well-known analogy between DA and the coupon collector problem to bound women's average rank of their husbands, but also encounters and tackles the challenge of tracking the (strictly positive) number of men who have reached the bottom of their preference lists by constructing a novel bound using a tractable stochastic process. The latter challenge did not arise in the setting of [62] where *all* short side agents are matched under stable matching, and similarly doesn't arise in our "densely connected markets" setting (Theorem 3.2). Following [62] and the majority of other theoretical papers on matching markets, we prove our characterizations for large n (and then use numerics to demonstrate that they extend to small n; see Section 3.5). Alongside an overview of the proof this section provides parenthetical pointers to the relevant formal lemmas; their statements and proofs can be found in Appendix C.2.

Our analysis tracks the progress of the following McVitie-Wilson [92] (sequential proposals) version of the men-proposing Deferred Acceptance algorithm that outputs MOSM (the final outcome is known to be the MOSM, independent of the sequence in which proposals are made). Under this algorithm, only one man proposes at a time, and "rejection chains" are run to completion before the next man is allowed to make his first proposal. The algorithm takes the preference rankings of the agents as its input.

Algorithm 3.1 (Man-proposing Deferred Acceptance)**.** *Initialize "men who have entered" $\hat{\mathcal{M}} \leftarrow \phi$, unmatched women $\bar{\mathcal{W}} \leftarrow \mathcal{W}$, the number of proposals $t \leftarrow 0$, the number of unmatched men $\delta^m \leftarrow 0$.*

1. *If $\mathcal{M}\backslash\hat{\mathcal{M}}$ is empty then terminate. Else, let i be the man with the smallest index in*

$\mathcal{M}\backslash\hat{\mathcal{M}}$. Add i to $\hat{\mathcal{M}}$.

2. *If man i has not reached the end of his preference list, do $t \leftarrow t+1$ and man i proposes to his most preferred woman j whom he has not yet proposed. If he is at the end of his list, do $\delta^m \leftarrow \delta^m + 1$ go to Step 1.*

3. *Decision of j:*

   (a) *If $j \in \bar{\mathcal{W}}$, i.e., j is currently unmatched, then she accepts i. Remove j from $\bar{\mathcal{W}}$. Go to Step 1.*

   (b) *If j is currently matched, she accepts the better of her current partner and i, and rejects the other. Set i to be the rejected man and continue at Step 2.*

Principle of deferred decisions. As we are interested in the behavior of Algorithm 3.1 on a random matching market, we think of the deterministic algorithm on a random input as a randomized algorithm, which is easier to analyze. The randomized, or coin flipping, version of the algorithm does not receive preferences as input, but draws them through the process of the algorithm. This is often called the *principle of deferred decisions*. The algorithm reads the next woman in the preference of a man in step 2 and whether a woman prefers a man over her current proposal in step 3b. No man applies twice to the same woman during the algorithm, and therefore the algorithm never reads previously revealed preferences. In step 2 the randomized algorithm selects the woman j uniformly at random from those to whom man i has not yet proposed. In step 3b, the probability that j prefers i over her current match is $1/(\nu(j)+1)$ where $\nu(j)$ is the number of proposals previously received by woman j.

Stopping time. Algorithm 3.1 defines that "time" t ticks whenever a man makes a proposal. First observe that the current number of unmatched men $\delta^m[t] = \delta^m$ at time t, i.e., men who have reached the bottom of their lists and are still unmatched, is *non-decreasing* over time, whereas the current number of unmatched women $\delta^w[t] = |\bar{\mathcal{W}}|$ at time t, i.e., women who have yet to receive their first proposal, is *non-increasing* over time. The MOSM is found when the number of unmatched men exactly equals the number of unmatched women plus k. We view this total number of proposals τ when DA terminates as a stopping time:

$$\tau = \min\{t \geq 1 : \delta^m[t] = \delta^w[t] + k\}\,. \tag{3.1}$$

This total number of proposals τ serves as a key quantity enabling our formal characterization of the MOSM (see Figure 3.3 for an illustration). On the men's side, the sum

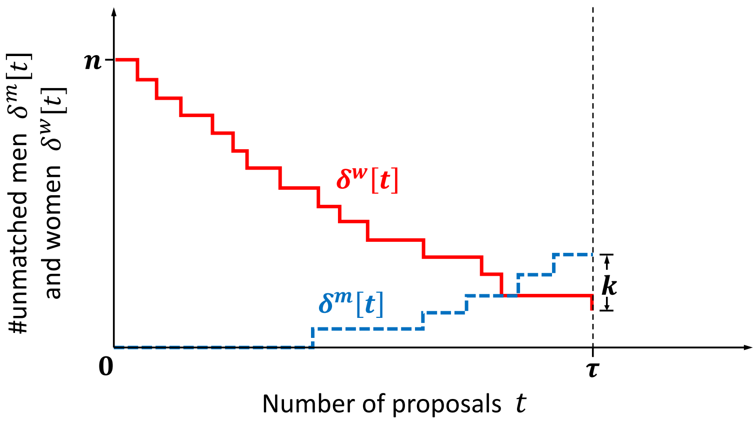

Figure 3.3: Illustration of a sample path of the current number of unmatched men $\delta^m[t]$ and unmatched women $\delta^w[t]$ under Man-proposing Deferred Acceptance (Algorithm 3.1). The algorithm terminates at $t = \tau$, the first time $\delta^m[t] = \delta^w[t] + k$. (In this illustration $k > 0$).

of men's rank of wives is approximately the total number of proposals τ (more precisely, this sum is $\tau + \delta^m[\tau]$ given that the rank for an unmatched agent is defined as one more than the length of the agent's preference list, but $\tau \gg \delta^m[\tau]$ is the dominant term). On women's side, since each proposal goes approximately to a uniformly random woman, as a function of the total number of proposals we can tightly control the distribution of the number of proposals received by individual women (this distribution is close to Poisson and tightly concentrates around its average) and therefore their average rank of husbands (Propositions C.5 and C.6), as well as the number of unmatched women (Propositions C.2

and[13] C.4).

Therefore, the bulk of the proof of Theorem 3.1 is dedicated to bounding the total number of proposals τ. Because of the aforementioned technical challenge that a positive number of agents remain unmatched on both sides, a direct application of the coupon collector analogy is not enough. Instead, we control the two stochastic processes that track the current number of unmatched men $\delta^m[t]$ and unmatched women $\delta^w[t]$ at each time t and make use of the identity (3.1) that $\delta^m[\tau] = \delta^w[\tau] + k$. (Upon termination, the number of unmatched men must be k plus the number of unmatched women.) For technical purposes, we extend the definition of $\delta^m[t]$ and $\delta^w[t]$ to $t > \tau$ as follows: if there are no men waiting to propose (i.e., a stable matching has been found), we introduce a fake man who is connected to d women (uniformly and independently drawn) with a uniformly random preference ranking over them, and keep running Algorithm 3.1.

Upper bound on the total number of proposals. We show (in Proposition C.1) that the total number of proposals cannot be too large, i.e., $\tau \leq (1+\epsilon)n\sqrt{d}$ with high probability for $\epsilon = d^{-1/4} = o(1)$. We establish this bound by showing that after a large enough number of proposals have been made, i.e., at time $t = (1+\epsilon)n\sqrt{d}$, the current number of unmatched women $\delta^w[t]$ has (with high probability) dropped below $ne^{-\sqrt{d}}$ whereas the current number of unmatched men $\delta^m[t]$ has (with high probability) increased above some level which is $\omega(ne^{-\sqrt{d}})$ and hence, since $k = O(ne^{-\sqrt{d}})$, the stopping event ($\delta^m[\tau] = \delta^w[\tau] + k$) must have happened earlier, i.e., $\tau \leq (1+\epsilon)n\sqrt{d}$. The upper bound on $\delta^w[(1+\epsilon)n\sqrt{d}]$ (see Lemma C.7) is derived using a standard approach that utilizes the analogy to the coupon collector problem. The lower bound on $\delta^m[(1+\epsilon)n\sqrt{d}]$ (see Lemma C.10) is obtained by counting the number of occurrences of d-rejections-in-a-row during the men-proposing DA procedure (whenever rejections take place d times in a row, at least one man becomes unmatched). Thus, our lower bound on $\delta^m[(1+\epsilon)n\sqrt{d}]$ ignores that some men are first accepted, and then later rejected causing them to reach the end

[13]In Proposition C.4, we first upper bound the number of unmatched women, and then use the aforementioned observation to lower bound the number of proposals.

of their preference lists via less than d consecutive rejections. Our conservative approach provides tractability and saves us from needing to track how far down their preference lists the currently matched men are. Nevertheless, the slack in this step necessitates our stronger assumption $d = o(\log^2 n)$, despite our conjecture that the characterization extends for all $d < 0.99 \log^2 n$.

Lower bound on the total number of proposals. We prove (in Proposition C.4) that the total number of proposals cannot be too small, i.e., $\tau \geq (1-\epsilon)n\sqrt{d}$ with high probability for some $\epsilon = o(1)$. We start with upper bounding (in Lemma C.13) the expected number of unmatched men in the stable matching, $\mathbb{E}[\delta^m]$, by showing that the probability of the last proposing man being rejected cannot be too large given that each woman has received at most $(1+\epsilon)\sqrt{d}$ proposals on average (recall that $\tau \leq (1+\epsilon)n\sqrt{d}$ w.h.p.). We then use Markov's inequality to derive an upper bound on δ^m which holds with high probability, and deduce (in Proposition C.3) an upper bound on δ^w using the identity $\delta^m = \delta^w + k$. Then we again use the coupon collector analogy to bound τ from below: the process cannot stop too early since the current number of unmatched women $\delta^w[t]$ does not decay fast enough to satisfy the upper bound on $\delta^w[\tau]$ $(= \delta^w)$ if τ is too small.

3.5 Numerical Simulations

This section provides simulation results that confirm and sharpen the theoretical predictions made in Section 3.3. Our simulations reveal (i) a sharp threshold at connectivity $d \approx 1.0 \log^2 n$ with no stark effect of competition observed for d below this threshold, and (ii) that our findings hold even for small values of n. We also investigate the role of imbalance k. Finally, we observe that the connectivity in the actual high school admissions data resembles that in a market with $n = 500$ and $d = 7 \ll \log^2 500 \approx 40$, providing some explanation for why that dataset does not exhibit a stark effect of competition.

We first examine the effect of connectivity on stable matchings in a random matching market of a fixed size. Specifically, we consider a market with 1,000 men and 1,001 women ($n = 1001$, $k = -1$) where the length of each man's preference list d varies from 5 to 150. For each degree d we generate 500 random realizations of matching markets according to the generative model described in Section 3.2, and for each realization we compute the MOSM via the men-proposing DA algorithm. Figure 3.4 reports the men's average rank of wives and the women's average rank of husbands (left) and the number of unmatched men and women (right) at each d. While not reported here to avoid cluttering the figures, we observe almost identical results for the WOSM. Observe that when $d < \log^2 n$ both men's average rank and women's average rank are highly concentrated at $\sqrt{d}$ and both the number of unmatched men and the number of unmatched women are close to $ne^{-\sqrt{d}}$, which confirms the estimates in Theorem 3.1. As d grows beyond $\log^2 n$, the average rank of men and women start to deviate from each other, and specifically, the average rank of short side (men) stops increasing whereas the average rank of long side (women) increases linearly: i.e., $R_{\text{MEN}} \approx \log n$ and $R_{\text{WOMEN}} \approx \frac{d}{\log n}$ when $d > \log^2 n$, confirming Theorem 3.2. We also remark that the number of unmatched men quickly vanishes as d increases beyond $\log^2 n$ (note that the y-axis of the plot has a log-scale).

The above observation extends to a wide range of market size n (even for small $n \leq 50$). To better illustrate, we investigate three kinds of threshold degree levels $d^*_{\text{rank}}(n)$, $d^*_{\delta}(n)$, and $d^*_{\text{conn}}(n)$ that sharply characterize the phase transitions that occur when degree d varies in random matching markets of size n. We define these thresholds as follows: given that $k = -1$ as above,

$$d^*_{\text{rank}}(n) = \min_d \left\{\mathbb{E}_{n,d}[R_{\text{WOMEN}}(\text{MOSM})]/\mathbb{E}_{n,d}[R_{\text{MEN}}(\text{MOSM})] \geq 1.15\right\}, \tag{3.2}$$

$$d^*_{\delta}(n) = \min_d \left\{\mathbb{E}_{n,d}[\delta^m] \leq 0.5\right\}, \tag{3.3}$$

$$d^*_{\text{conn}}(n) = \min_d \left\{\mathbb{E}_{n,d}[\text{the number of connected components}] \leq 2\right\}, \tag{3.4}$$

where $\mathbb{E}_{n,d}[\cdot]$ represents the expected value of some random variable in a random matching market with $n-1$ men each of whose degree is d and n women. The rank-gap threshold $d^*_{\text{rank}}(n)$ indicates the degree value beyond which men's average rank and women's

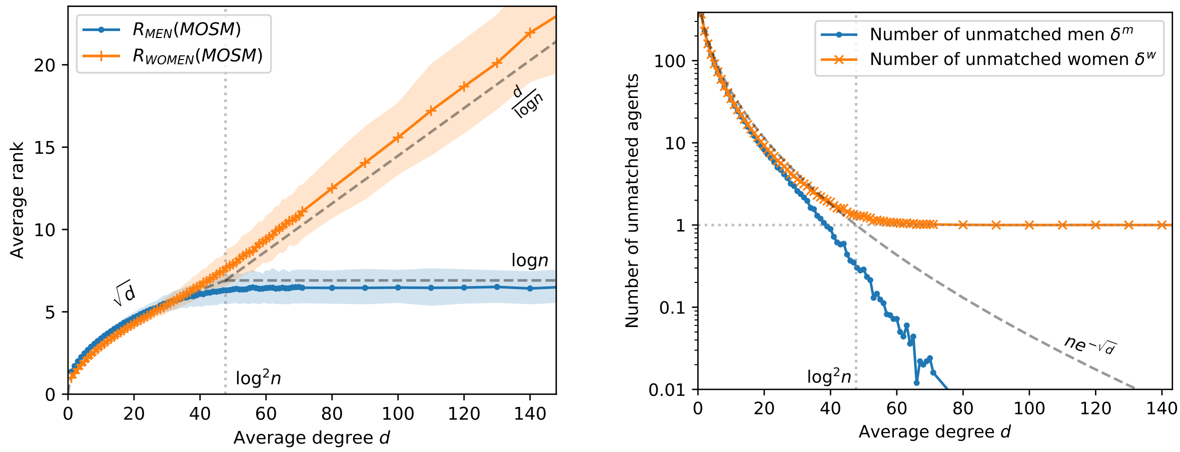

Figure 3.4: Men's average rank of wives R_{MEN} and women's average rank of husbands R_{WOMEN} (left) and the number of unmatched men δ^m and the number of unmatched women δ^w (right) under MOSM in random matching markets with 1,000 men and 1,001 women ($n = 1001$, $k = -1$), and a varying length of men's preference list d. In both figures, solid lines indicate the average value across 500 random realizations, and gray dashed lines indicate our theoretical predictions (Theorem 3.1 and 3.2) annotated with their expressions. In the left figure, the shaded areas surrounding solid lines represent the range between the top and bottom 10th percentiles of 500 realizations of men's and women's average rank.

average rank start to deviate from each other (in particular, we require a 15% or larger difference in the average ranks on the two sides of the market); the unmatched-man threshold $d^*_\delta(n)$ is the degree value beyond which all men are (typically) matched; and the connectivity threshold $d^*_{\text{conn}}(n)$ is the degree value beyond which the entire market is typically connected. We quantify these threshold values based on numerical simulations. More specifically, we vary the number of men n from 10 to 2,500, and for each n we use bisection method with a varying d to find the threshold degrees, where the expected values are approximated with sample averages across 500 random realizations. We find that bisection method is adequate since each of the measures on which the above thresholds are defined is observed to monotonically increase or decrease in d, and further to change rapidly near the threshold value d^* that we want to estimate.

Figure 3.5 plots the measured threshold degrees. Remarkably, the thresholds $d^*_{\text{rank}}(n)$ and $d^*_\delta(n)$ are very close to $\log^2 n$ for all tested values of n. This suggests that our predicted threshold is fairly sharp: the short-side advantage emerges if and only if $d \gtrsim$

$1.0 \times \log^2 n$. Also note that this threshold is much larger than the connectivity threshold $d^*_{\text{conn}}(n) \approx \log n$.

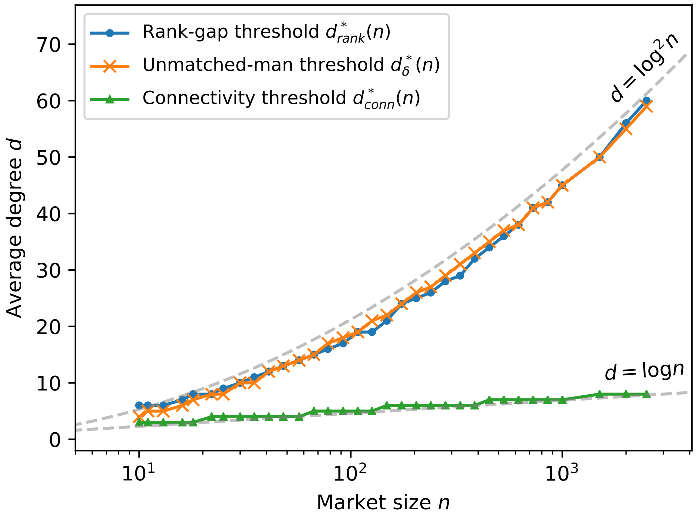

Figure 3.5: Threshold degrees $d^*_{\text{rank}}(n)$, $d^*_{\delta}(n)$, and $d^*_{\text{conn}}(n)$, defined in (3.2)–(3.4), in random matching markets with $n-1$ men and n women where n ranges from 10 to 2,500. For each n, the threshold values are found using bisection method in which we simulate 500 realizations at each attempted d. The gray dashed lines indicate the theoretical predictions annotated with their expressions.

We next investigate the effect of imbalance k on the stable outcomes and characterize it at the different levels of connectivity d. Analogous to the numerical experiment for the high school admissions discussed in Section 3.1, we fix the number of women $n = 500$ (so $\log^2 n \approx 40$), and measure men's average rank under MOSM (averaged across 500 realizations) where the number of men varies from 450 to 550. To facilitate easier comparison, we compute the normalized average rank R_{MEN}/d: e.g., $R_{\text{MEN}}/d \approx 0.2$ implies that in average a man is matched to his top-20% most preferred woman out of his preference list. Figure 3.6 shows how the (normalized) men's average rank changes as we add or remove men in the market, tested with different values of d. Observe that for large $d > \log^2 n$ (e.g., $d = 100, 450$) there is a stark effect when we inject a slight imbalance into the balanced market; compare 500 men vs. 501 men. In contrast, for small $d < \log^2 n$ (e.g., $d = 10, 20$), the stable outcome changes very "smoothly" across a wide range of imbalance, which is consistent with simulation results based on high school admissions data (see Figure 3.1).

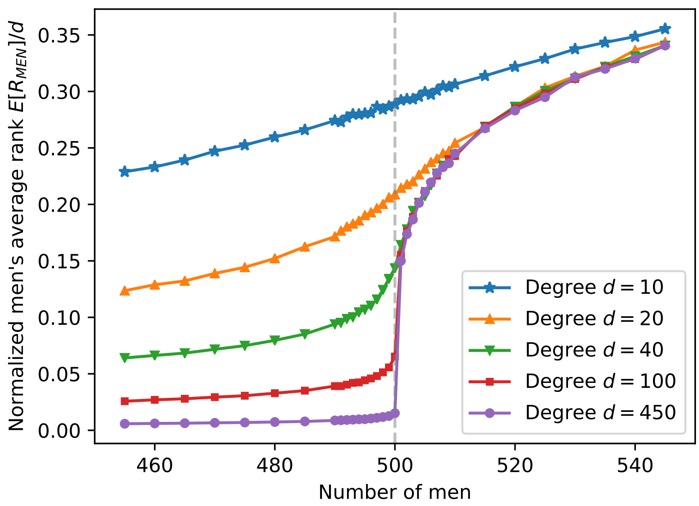

Figure 3.6: The effect of imbalance k on men's average rank in random matching markets with a fixed number of women $n = 500$ ($\log^2 n \approx 40$). For each $d \in \{10, 20, 40, 100, 450\}$, the corresponding curve reports men's average rank under MOSM normalized by d, i.e., $\mathbb{E}[R_{\text{MEN}}(\text{MOSM})]/d$, where the number of men varies from 450 to 550 (i.e., $k = -50, \dots, 50$). Each data point reports the average value across 500 realizations.

We conclude this section by providing some statistics that illustrate the level of connectivity in the high school admissions example and showing that random matching markets with $n = 500$ and $d = 7$ exhibit a similar level of connectivity. We focus on the pairwise distance among students as a measure of the connectivity of a matching market: e.g., the distance between two students is one hop if they applied to the same program. On the actual high school admissions data, we sample 1,000 students out of total 75,202 students, and measure the distance from each of selected students to all the other 75,201 students. We observe that 10.1% of student pairs are within 1 hop, 97.8% of pairs are within 2 hops, and 100.0% of pairs are within 3 hops. (Recall that the average preference list length in this high school admissions data was 6.9.) We apply the same analysis on our random matching market model and find that the model with $n = 500$ and $d = 7$ yields a comparable outcome: 9.4% of man pairs are within 1 hop, 98.1% of pairs are within 2 hops, and 100.0% of pairs are within 3 hops. Given that $\log^2 n \approx 40$

for $n = 500$, far in excess of $d = 7$, and since correlation in preferences seems only to reduce the effect of competition (see evidence in Figure 3.1), we deduce that the high school admissions market seems to lie well within the "no stark effect" regime covered by Theorem 3.1, which provides an explanation as to why we do not see a stark effect of competition (Figure 3.1).

3.6 Discussion

We investigated stable matchings in random matching markets which are partially connected, and asked which random matching markets exhibit a stark effect of competition. In particular, unlike many previous papers which study whether there is a nearly unique stable matching, we focus on the issue of how well (or poorly) agents do under stable matching, as a function of market primitives. The parameter d captured the connectivity (average degree), n captured the market size and k captured the imbalance, whereas preferences were assumed to be uniformly random and independent. We found that, in densely connected markets $d = \omega(\log^2 n)$, the short side of the market enjoys a significant advantage, generalizing the finding of [62] in fully connected markets. In contrast, in moderately connected markets $d = o(\log^2 n)$, we found that for any $k = o(n)$, the two sides of the market do almost equally well, challenging the claim of [62] that "*matching markets are extremely competitive*". Notably, this "no stark effect of competition" regime extends far beyond the connectivity threshold of $d = \log n$ and thus includes well connected markets. Numerical simulation results not only support our theory but further indicate that our findings extend to small n and that there is a sharp threshold between the two regimes at $d \approx 1.0 \log^2 n$. We argued informally that most real world matching markets lie in the no stark effect of competition regime, providing some explanation why matching market datasets do not exhibit a stark effect of competition.

Following the theoretical matching literature, we have analyzed a highly stylized model in the interest of tractability and obtaining sharp results. (Even so, we encounter and overcome significant new technical challenges.) We leave as interesting and challenging directions for future work to characterize stable matchings while incorporating various features of real world market such as many-to-one matching, correlation in preferences, and small market sizes.

CHAPTER 4

Price Discovery and Efficiency in Waiting Lists: A Connection to Stochastic Gradient Descent

4.1 Introduction

Public scarce resources are often allocated through waiting lists, in which agents can select which resource to wait for For example, the New York Public Housing Authority asks applicants to choose a project specific queue. If utilities are quasi-linear in waiting costs and total waiting costs are constant across assignments, waiting times act as prices that clear the market and create an efficient assignment. In this sense, waiting lists are similar to standard competitive equilibrium (CE) models, with the exception that prices are quoted in waiting times instead of monetary transfers. Waiting lists mechanisms have a natural price formation process. Indeed, waiting times, or prices, are naturally adjusted with arrival of agents and resources. This paper is concerned with how the price dynamics that are inherent to waiting lists impact allocative efficiency.

The economy considered has items of different types as well as agents arriving over time. A waiting list mechanism maintains one observable queue for each item type. Agents have unit demand, heterogeneous private values for items. The utility of an agent is quasi-linear in waiting costs. Upon arrival, the agent, who maximizes expected utility, chooses either a single queue and waits until she is assigned an item, or, leaves immediately and receives an outside option. When an item arrives, it is assigned to an agent in the respective queue or discarded if that queue is empty.

Prices in this economy depend on the current state of the queues; the price of an item type decreases when such an item is assigned to an agent, and increases when an agent joins the corresponding queue. This means that different agents may face different prices upon arrival. This stochastic evolution creates a challenge to analyze this simple economy.

To understand how price adjustments affect welfare, it is helpful to first relate the economy to a canonical CE model. Suppose that instead of dynamically assigning items, a center could wait for a large time frame and then simultaneously assign all accumulated items to the agents By the first welfare theorem, prices that clear the market and form a CE will generate an assignment of items to agents that maximizes the total value agents ascribe to their assigned items. Assuming that total waiting costs are constant across assignments, the value generated by the CE assignment gives an upper bound on the welfare in the dynamic model. Moreover, this welfare can be achieved if prices (expected waiting costs at each queue) remain constant at CE prices.

Since prices in the economy fluctuate, they may fail to generate the same welfare as CE prices. To illustrate this consider a simple example with a single type of item. Agents' valuations are distributed uniformly between 0 and 1 and they arrive twice as fast as items. Since only half the agents can be assigned, it is optimal to set a price equal to 1/2 and assign the item to the agents with valuation above that price. But in the waiting list the price fluctuates below and above 1/2. This results in some agents with lower values being assigned and some agents with high values selecting the outside option.

With multiple types of items, prices follow a multidimensional Markov process. While this process is easy to describe, the stationary distribution cannot be derived tractably. For this reason, several related papers restricted attention to economies with at most two item types (see, e.g., Baccara, Lee, and Yariv [93] and Leshno [94]). However, we can bound the welfare loss in the economy by drawing a connection between the price dynamics under the waiting list mechanism, and the stochastic gradient descent (SGD) optimization algorithm. The main insight is that while price adjustments are stochastic, prices tend to adjust towards their CE levels. Formally, the expected price adjustment is a subgradient of a dual of the welfare maximization problem.

A key quantity is the granularity of price adjustments Δ, defined as the maximal increase in price (waiting costs) due to the addition of one agent to a queue. Equivalently, Δ determines the size of the adjustment step taken by the SGD algorithm after seeing one sample. This granularity is a key distinction between the price adjustment process in our economy and standard usage of SGD in optimization: to make the SGD converge the step size is reduced to zero over time, while in waiting lists the step size is exogenously given by agent sensitivity to waiting costs and remains constant. Despite this distinction, the connection to SGD allows us to leverage tools from Lyapunov theory to obtain tractable bounds for welfare.

The main finding is that the welfare loss in waiting lists is bounded by the granularity of price adjustments Δ times a factor that depends only on the relative arrival rates of agents and items. The dynamic price adjustments keep prices "close" to the CE prices, despite never converging. The welfare loss due to constant fluctuation of prices depends on the magnitude of these fluctuations, which become small as Δ becomes small.

For further intuition, consider again the simple example and assume waiting costs are linear. If waiting behind 10 agents in the queue gives a waiting cost of $1/2$, then each arrival or departure of an agent changes the price by 10% of the item's maximal value. If agents are more patient and waiting behind 100 agents in the queue gives a waiting cost of $1/2$, then each arrival or departure of an agent changes the price by only 1% of the item's maximal value. So Δ is smaller when agents are more patient. Intuitively, prices will fluctuate less when agents are patient resulting in less welfare loss.

Two additional results complement the main finding. We show that the bound is essentially tight (up to constant factors) by explicitly constructing an economy in which welfare loss is high. A distinct and important feature of this economy is the multiplicity of CE prices.

Finally, it is shown that the loss is generically much smaller for economies with finitely many agent types. When the number of agent types is finite, the CE prices are unique for all but a zero measure of arrival rates. And when CE prices are unique the welfare loss becomes exponentially small as Δ tends to zero.

The analysis can offer insight for general price adaptation algorithms. Consider a planner who is able to choose the granularity of price adjustments Δ. Increasing Δ will make the prices adjust more quickly if the distribution of agent valuations changes. But a higher value of Δ will also cause losses when agent valuations remains the same, but prices change due to the random arrivals. This tension is inherent to any price adaptation mechanism that can only observe noisy signals, and needs to trade-off overreactions due to imprecise information and slow reaction time due to the need to accumulate sufficiently precise information. A particular feature may be of interest: firms that adjust prices according to an SGD heuristic may find it optimal to react slowly to changes in monetary policy if these changes have unclear implications for their demand new prices new to be learned through the dynamics.

The first area this paper is related is about dynamic matching motivated by applications in public housing Kaplan [95, 96] and organ allocation Zenios [97]. Several papers compare the efficiency across queuing and lottery mechanisms for restricted set of preferences [98, 99, 100]. Some papers consider optimal design of dynamic allocation without accounting for the stochastic process [101, 102, 103, 104]. Our paper focuses on understanding the relation between the dynamics in the queueing mechanism the economic efficiency.

Several papers are concerned with the stochasticty in dynamic two-sided trading markets in order to optimize clearing timing decisions Mendelson [105], Kelly and Yudovina [106], and Loertscher, Muir, and Taylor [107]. These papers also restrict attention to either a single asset or a binary type space.

This paper also relates to papers that are concerned with convergence of tâtonnement processes using gradient descent. Numerous paper analyze these processes in markets with multiple goods [108, 109, 110, 111] and in congestion or transportation settings [112, 113]. These papers consider static market and study identify an adjustment process that converges to equilibrium. In contrast, in the dynamic market considered here "prices" never converge.[1]

Similar ideas have been considered in the network control literature [see, e.g., 5], but the model consider a much more general settings (continuum of types, nonlinear waiting costs) and establish novel results (exponential loss).

4.2 A Simple Illustrative Example

We consider a simple market to illustrate the loss from price-fluctuations. There is a single kind of item, which arrives according to a Poisson process with rate 1. Agents arrive according to a Poisson process with rate 2. Arriving agents choose whether to join the queue and wait to receive an item, or be immediately assigned an outside option and receive 0 utility. The utility of an agent who joins the queue is

$$v - c \cdot w$$

where w is the amount of time the agent waits, $c = 0.02$ is equal to the expected waiting costs of waiting for one item's arrival, and v is the agent's value of receiving the item. Each agent's value v is independently drawn from the uniform distribution on $[0, 1]$. Agents observe the current number of agents in the queue when they arrive. Agents in the queue are assigned in the order in which they joined the queue (First-Come First-Served). When an agent with value v arrive, she will join the queue if $v - c \cdot \mathbb{E}[w|q] > 0$ where $\mathbb{E}[w|q] = q + 1$ is the expected wait given there are currently q agents in the queue.

As a benchmark, consider a planner who collects all agents and items that arrive up to time T and assigns all items at time T. If T is large, the planner collects approximately

[1]Some papers identify price adjustments processes that converge to market clearing prices in a static version of our assignment problem though with finite many types of agents (e.g., Bertsekas [114] and Demange, Gale, and Sotomayor [115]).

twice as many agents as items, and the distribution of agent values is approximately $U[0, 1]$. Because only half the agents can be assigned an item, allocative efficiency is maximized by allocating items only to agents with a value $v \in [1/2, 1]$.

Figure 4.2 shows the distribution of queue length an arriving agent faces. On average, the queue length implies an expected waiting cost equal to 1/2. If all agents faced a price (in waiting cost) equal to 1/2, only agents with a value $v \in [1/2, 1]$ would be assigned items and the maximal allocative efficiency would be generated.

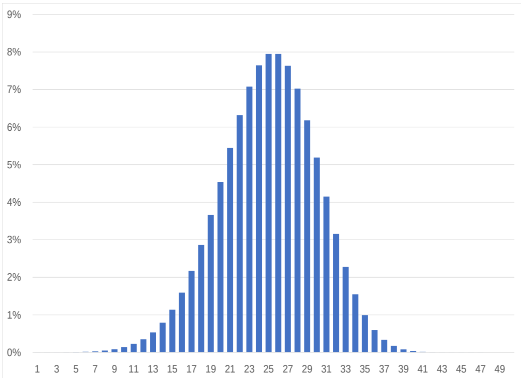

Figure 4.1: Queue length distribution for the example

However, Figure 4.2 shows there is considerable variation in the prices agents face. As a result, the allocation may be inefficient. An agent with a value of $v = 0.41$ may be assigned if she is lucky to arrive when the queue happens to be short. An agent with a value of $v = 0.59$ may not be assigned if she is arrives when the queue happens to be long. Figure 4.2 shows the implies probability that an agent is assigned as a function of the agent's value. It shows that agents with a value of $0.41, 0.59$ are assigned the item with probabilities 18% and 81%, respectively.

4.3 Model

We study an infinite horizon economy, in which agents and items arrive randomly over time. We describe the economy, set benchmarks for allocative efficiency, and describe the

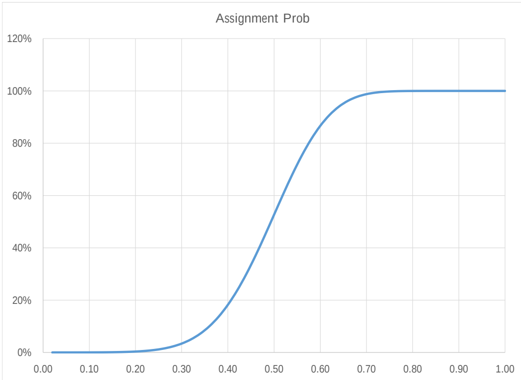

Figure 4.2: Implied probability of assignment given the agent's value

waiting list mechanism.

Economy We consider a market in which items and unit-demand agents arrive over time. Agents arrive according to a Poisson process with rate λ. Each agent has a type θ drawn independently according to distribution F over the set of types Θ. We assume that Θ is a compact subset of an Euclidean space, and allow for both finitely many agent types as well as a continuum of agents. We say that there are finitely many agent types if F corresponds to finitely many atoms.

Items arrive according to a Poisson process with total rate normalized to 1. The agent and item arrival processes are independent. Each arriving item is of a type $j \in \mathcal{J} = \{1, 2, \ldots, J\}$. An item is of type j with probability $\mu_j > 0$, where $\sum_{j \in \mathcal{J}} \mu_j = 1$. Denote by $\mu_{\min} \triangleq \min_{j \in \mathcal{J}} \mu_j > 0$, $\mu_{\max} \triangleq \max_{j \in \mathcal{J}} \mu_j > 0$. We define an auxiliary item type $\emptyset$, which denotes being unassigned and use $\mathcal{J}_\emptyset \triangleq \mathcal{J} \cup \{\emptyset\}$.

The value an agent of type $\theta \in \Theta$ obtains from getting assigned to an item of type $j \in \mathcal{J}_\emptyset$ is given by $v(\theta, j)$, where we normalize $v(\theta, \emptyset) = 0$. Agents' utilities are quasi-linear in waiting costs; An agent of type θ that is assigned to type j after waiting w units of time receives a utility of

$$u_\theta(j, w) \triangleq v(\theta, j) - c(w)\,,$$

where $c(w)$ is the cost of waiting w units of time.

We make the following technical assumptions. We assume that for each $j \in \mathcal{J}$, $v(\theta, j)$ is continuous in θ and bounded from above by $v_{\max} \in \mathbb{R}$. We assume that the waiting cost function is smooth, strictly increasing, weakly concave,[2] and $c(0) = 0$, $\lim_{w\to\infty} c(w) = \infty$.

To simplify notation, we consider an equivalent discrete time process[3] indexed by t which records the sequence of arrivals. For each arrival epoch t, the indicator ξ_t equals one if the t-th arrival is an agent, and equals zero if the t-th arrival is an item arrives. If $\xi_t = 1$, let θ_t denote the type of the agent arriving at t. If $\xi_t = 0$, let $j_t \in \mathcal{J}$ denote the type of item arriving at t.

Assignments and Allocative Efficiency An *allocation* η assigns each agent with one item, and each non-auxiliary item is assigned to at most one agent. The allocative efficiency of a matching is defined as the average item's value to its assigned agent. Formally, given allocation η, for each epoch t such that $\xi_t = 1$ let $\eta_t \in \mathcal{J}_\emptyset$ be the kind of item assigned under η to the agent of type θ_t that arrived in epoch t. Let $A_T = \sum_{t\le T} \xi_t$ be the total number of agents that arrived up to epoch T. Allocative efficiency under η is defined as

$$W(\eta) = \liminf_{T\to\infty} \frac{1}{A_T} \sum_{t=1}^{T} \xi_t v(\theta_t, \eta_t)\,. \tag{4.1}$$

We restrict attention to allocations that satisfy a *no-Ponzi* condition. Loosely speaking, this condition ensures the assignment is approximately valid if the market terminates at some large finite time.[4] Formally, let $\mathcal{R}_T(\eta)$ denote the number of agents and items that arrived by time T and are waiting to be assigned at time T.[5] The assignment η satisfies the no-Ponzi condition if there exists a finite $M \in \mathbb{R}$ such that $\mathcal{R}_T(\eta) < M$ for all T.

[2]Our results also extend to convex $c(w)$ such that both $c'(w)$ and $c''(w)$ are subexponential, i.e., there exists α such that $c'(w), c''(w) \le e^{\alpha w}$ for all $w \ge 0$.

[3]The equivalence is due to the Arrival Theorem of Poisson-driven processes [see, e.g., 116].

[4]For example, an allocation that assigns all agents to items in the market described in Section 4.2 is not valid if the market terminates at any finite time, as only (approximately) half the agents can be assigned items.

[5]In other words, $\mathcal{R}_T(\eta)$ counts the number of agents who arrived before time T and are assigned under η to an item that arrives after time T, plus the number of items that arrive before time T and are assigned to agents that arrive after time T.

Define optimal allocative efficiency to be

$$W^{\text{OPT}} = \mathbb{E}\left[\sup_{\eta \in H} W(\eta)\right] ,$$

where H is the set of no-Ponzi allocations and the expectation is taken over all possible realizations.

Allocation by a Waiting List Our main interest is to analyze the allocative efficiency of the allocation generated by a standard waiting list mechanism. The mechanism holds a separate First-Come-First-Served queue for each item. An arriving agent observes the length of the queue for each item and chooses to join the end of one of the queues, or take the auxiliary item immediately (i.e., balk). An agent who joins a queue will wait in that queue until receiving an item. When an item arrives, it is assigned to the agent at the head of its queue, if there is any; if the item's queue is empty, the item is discarded.

To formally describe the mechanism, let $\mathbf{q} = (q_1, \ldots, q_J) \in \mathbb{Z}_+^J$ denote the state where there are q_j agents in the queue for item j. An arriving agent of type θ who observes $\mathbf{q}$ and chooses to join the queue for item j will wait a random amount of time w_j before receiving item $j \in \mathcal{J}_\emptyset$, and will receive an expected utility of $v(\theta, j) - \mathbb{E}[c(w_j)|q_j]$. Thus, the agent will choose to join the queue for item $a(\theta, \mathbf{q}) \in$ defined by[6]

$$a(\theta, \mathbf{q}) = \arg\max_{j \in \mathcal{J}_\emptyset} \left\{v(\theta, j) - p_j(q_j)\right\} , \tag{4.2}$$

where we define $p_j(q_j) \triangleq \mathbb{E}[c(w_j)|q_j]$. We allow agents to leave without joining any queue, and simplify notation by setting $p_\emptyset(\cdot) \equiv 0$. For notation simplicity, denote $\mathbf{p}(\mathbf{q}) \triangleq [p_1(q_1), \cdots, p_J(q_J)]$. Denote the queue lengths just before the t-th arrival by $\mathbf{q}_t$. That is, an agent that arrives at epoch t will face prices $\mathbf{p}_t = \mathbf{p}(\mathbf{q}_t)$, which depend on the current state of the queues $\mathbf{q}_t$.

Given a realization, let η^{WL} denote the allocation induced by the waiting list. Under

[6] To simplify notation, we assume that the arg max is unique and implicitly rely on a tie-breaking rule to ensure a unique selection if the agent is indifferent between multiple items. Our results do not depend on the choice of tie-breaking rule. For example, agents may randomly choose an item in the arg max or choose the lowest index item within the arg max.

our assumptions, η^{WL} satisfies the no-Ponzi condition.[7] We denote the expected allocative efficiency of the waiting list mechanism by

$$W^{\mathrm{WL}} \triangleq \mathbb{E}\left[W(\eta^{\mathrm{WL}})\right] .$$

and refer to $W^{\mathrm{OPT}} - W^{\mathrm{WL}}$ as the *allocative efficiency loss*, or loss for short.

4.4 Bounding the Allocative Efficiency Loss

In the waiting list mechanism, each agent is presented a menu of items and associated expected waiting costs. We consider the waiting cost of an item as the item's price, and henceforth refer to the expected waiting cost of item j as the price of item j.

Standard intuition from competitive equilibria tells us that appropriately set prices can guide agent's choices and lead to an allocation that maximizes allocative efficiency. But in the waiting list there is no planner that sets prices. Instead, prices are determined by the current state of the queues. Prices adapt over time in a process that is similar to a tâtonnement process: the price of item j increases when an agent chooses to join queue j, and the price of item j decreases when item j arrives and one agent is removed from queue j.

Because prices are state-dependent, prices fluctuate over time. For example, if an agent θ arrives immediately after several copies of the item j happened to arrive, agent θ will face a lower price p_j (i.e., expected waiting cost for item j). The example in Section 4.2 shows that such stochastic price fluctuations can lead to lower allocative efficiency.

We seek to understand the allocative efficiency of the fluctuating state-dependent prices in the waiting list. A natural approach would be to calculate the stationary distribution of prices. Unfortunately, this stationary distribution is not tractable when $|\mathcal{J}| > 2$, that is, there are strictly more than 2 kinds of items.[8] We therefore take a

[7]Because no queue length can ever exceed $q_{\max} \triangleq \max_{j\in\mathcal{J}}\ p_j^{-1}(v_{\max})$ we have that $\mathcal{R}_t(\eta^{\mathrm{WL}}) \leq |\mathcal{J}| \cdot q_{\max}$.

[8]Because of this limitation, previous papers that relied on calculation of the exact stationary distri-

different approach that allows us to analyze general markets with any number of items, a general (possibly continuous) distribution of agent types, and nonlinear waiting costs.

Our analysis shows that the following attribute plays an central role in determining allocative efficiency:

Definition 4.1. *The step size* Δ *is the maximal change in price due to a single arrive and is given by*[9]

$$\Delta \triangleq \max_{j \in \mathcal{J}} \max_{0 \leq q \leq q_{\max}} \big(p_j(q) - p_j(q-1)\big).$$

In other words, each arrival of an item j reduces the price of item j by at most Δ. Each arrival of an agent who joins the queue of item j increases the price of item j by at most Δ. If waiting costs are linear, i.e., $c(t) = c \cdot t$ for some $c > 0$, we have $\Delta = c/\mu_{\min}$. That is, Δ is the expected cost of waiting for a single arrival of the least frequent item.

We can now state our main result.

Theorem 4.1. *The allocative efficiency under the waiting list is*

$$W^{\mathrm{WL}} \geq W^{\mathrm{OPT}} - \frac{\lambda + 2}{2\lambda}\Delta \,. \tag{4.3}$$

I redid the calculation for the case $p = c(q+1)/\mu$ (previously it was $p = cq/\mu$), and the numerator of the coefficient of Δ goes from $\lambda + 1$ to $\lambda + 2$.

Theorem 4.1 shows that the queueing mechanism achieves allocative efficiency that is close to optimal in general dynamic markets. To illustrate the result, suppose waiting costs are linear and that the agent arrival rate is equal to the item arrival rate (that is, $\lambda = \mu = 1$). In this case, the allocative efficiency loss is bounded by the cost of waiting for a single arrival of the least frequent item times $\frac{3}{2}$. In Section 4.4.2 we provide an example giving a lower bound on the allocative efficiency loss.

To gain intuition for Theorem 4.1, we draw connections to two related problems.

bution of the underlying Markov chain were limited to a model with 2 items (e.g., [94, 99, 93]).

[9]Let $p_j(-1) = 0$ so that Δ is well-defined. Note that $p_j(0) > 0$ because even if the agent arrives to an empty queue, she still needs to wait for the next to item to arrive.

Problem I: Duality for Static Allocation Consider the static allocation problem in which a planner chooses a static assignment of the expected amount of agents that arrive per unit time to the expected amount of items that arrive per unit time in order to maximize allocative efficiency. We refer to the maximal value per agent the assignment can generate as the optimal static allocative efficiency and denote it by W^*. The value of W^* is the optimal value of assignment problem (4.4).

$$
\begin{aligned}
W^* = \max_{\{x_{\theta j}\}_{\theta\in\Theta, j\in\mathcal{J}}} \quad & \sum_{j\in\mathcal{J}} \int_{\Theta} x_{\theta j}\, v(\theta, j)\, dF(\theta) & & \\
\text{subject to} \quad & \sum_{j\in\mathcal{J}} x_{\theta j} \le 1,\ x_{\theta j} \in [0,1] & \forall \theta \in \Theta & \quad (4.4) \\
& \int_{\Theta} \lambda x_{\theta j}\, dF(\theta) \le \mu_j & \forall j \in \mathcal{J} &
\end{aligned}
$$

In problem (4.4), $x_{\theta j}$ is the share of agents of type θ that are assigned item j. The first constraint requires that the shares $x_{\theta j}$ are well defined. The second constraint is the resource constraint, it requires that the expected amount of item j arriving per unit time should at least as large as the expected amount of agents that arrive per unit time and are assigned to item j.

Proposition 4.1. *The optimal allocative efficiency is $W^{\text{OPT}} = W^*$, where W^* is the optimal static allocative efficiency.*

The proof of Proposition 4.1 is in Appendix D.1.

It will be useful to consider the dual problem of the assignment problem (4.4), which optimizes over possible prices. The following strong duality result is well-known and we therefore omit the proof.[10]

Lemma 4.1 (Monge-Kantorovich duality)**.** *The optimal value W^* of assignment problem* (4.4)*, coincides with the optimal value of the following dual optimization problem:*

$$\underset{\mathbf{p}\ge\mathbf{0}}{\text{minimize}}\ \ h(\mathbf{p})$$

[10]Problem (4.4) is known as the (unbalanced) optimal transport problem, which has the strong duality property stated in Lemma 4.1. For further details, see, e.g., [117].

where (here we let $p_\phi \triangleq 0$)

$$h(\mathbf{p}) \triangleq \int_\Theta \max_{j \in \mathcal{J}_\emptyset} \left[v(\theta, j) - p_j \right] dF(\theta) + \frac{1}{\lambda} \sum_{j \in \mathcal{J}} \mu_j p_j \,. \tag{4.5}$$

We use $\mathbf{p}^*$ to denote some optimal value for the dual problem, and refer to $\mathbf{p}^*$ as optimal prices.

Problem II: Stochastic Gradient Descent We relate the price adaptation in the waiting list to the run of the stochastic gradient descent optimization algorithm. The SGD algorithm can be regarded as a stochastic version of the gradient descent optimization algorithm. Each step of the SGD is random, but the expected step of SGD correspond to a step of gradient descent. SGD optimization is commonly used in machine learning, e.g., for training a neural networks [118]. By understanding the connection between our problem and the SGD algorithm we are able to leverage the substantial theory on SGD algorithms.

The following lemma establishes the connection between the waiting list and the SGD algorithm:

Lemma 4.2. *If the system is in state $\mathbf{q}_t$ the expected change to the queue length from a single arrival $\mathbb{E}[\mathbf{q}_t - \mathbf{q}_{t+1}]$ equals $\frac{\lambda}{1+\lambda}$ times a subgradient of the dual objective $h(\mathbf{p}_t)$ at $\mathbf{p}_t = \mathbf{p}(\mathbf{q}_t)$.*

Proof. The expected adjustment to the length of queue j from a single arrival is

$$\begin{aligned} \mathbb{E}[q_{j,t+1} - q_{j,t}] &= \mathbb{E}\left[\mathbf{1}_{\{\xi_t = 1,\, a(\theta_t, \mathbf{q}_t) = j\}} - \mathbf{1}_{\{\xi_t = 0,\, j_t = j\}} \right] \\ &= \frac{\lambda}{1+\lambda} \int_\Theta \mathbf{1}_{\{j = \arg\max_{j \in \mathcal{J}_\emptyset} \{v(\theta, j) - p_j(q_{j,t})\}\}} dF(\theta) - \frac{1}{1+\lambda} \mu_j \,. \end{aligned} \tag{4.6}$$

It is straightforward to verify that $\frac{1+\lambda}{\lambda}\mathbb{E}[q_{j,t+1} - q_{j,t}]$ is a subgradient of $h(\mathbf{p}_t)$ at $\mathbf{p}_t = \mathbf{p}(\mathbf{q}_t)$. □

In other words, Lemma 4.2 says that the stochastic price adjustment from one arrival corresponds to a step of an SGD algorithm for the dual objective $h(\mathbf{p})$ (Lemma 4.1). Loosely speaking, the waiting lists adjusts prices in the right direction on average, but the adjustment is random because it depends on the realization of a single arrival. However, in order for gradient descent and SGD algorithms to converge, the step size must decrease to zero as the algorithm approaches the optimal value. In the waiting list, the size of the adjustment is fixed and bounded by the step size Δ. Therefore, the price adjustment in the waiting list corresponds to the run of an SGD with a fixed step size that will never converge.

Intuition for Theorem 4.1 The connection to SGD allows us to apply the techniques developed to better understand SGD algorithms. The main part of the proof uses a Lyapunov potential function to decompose the expected value from the next arrival. Each agent's arrival generates a value of assignment, and the value generated depends on the current state of the queues. This value is related to the dual objective (4.5) evaluated at the current prices $\mathbf{p}_t$, which by Lemma 4.1 is at least as high as W^*. Each arrival also changes the current state. We use a Lyapunov function to capture the "potential" given the current price $L(\mathbf{p}_t)$. We decompose the expected value from the next arrival into a combination of the objective, change in potential, and a per-period loss.

$$\mathbb{E}[v(\theta_t, a(\theta_t, \mathbf{q}_t))\xi_t|\mathbf{q}_t] \geq \frac{\lambda}{\lambda+1}W^* - \underbrace{\frac{1}{\mu_{\min} \cdot \Delta}\big(L(\mathbf{p}_t) - \mathbb{E}[L(\mathbf{p}_{t+1})|\mathbf{q}_t]\big)}_{\text{(I) change in potential}} - \underbrace{\frac{2+\lambda}{2(1+\lambda)}\Delta}_{\text{(II) loss}} \quad (4.7)$$

To interpret equation (4.7), observe that $\frac{\lambda}{\lambda+1}W^* = \frac{\lambda}{\lambda+1}W^{\text{OPT}}$ is the average per-arrival (including both agents and items) value under the optimal assignment. Equation (4.7) shows that the waiting list achieves this value minus change in potential and a per-period loss. Summing over many periods, the change in potential (I) forms a telescoping series, and will therefore remain bounded. Therefore, as we average over many periods, we have that (I) tends to zero. The loss term (II) is uniformly bounded for any $\mathbf{p}_t$, allowing us to obtain the bound in Theorem 4.1 without calculating the stationary distribution.

Proof of Theorem 4.1. The proof adopts the Lyapunov analysis approach. We give the main arguments of the proof for the special case of linear waiting cost here and relegate technical lemmas and the proof of general cases to the appendix.

Let $W_T(\eta^{\mathrm{WL}})$ be the total value of items assigned to agents that arrive before epoch T, that is

$$W_T(\eta^{\mathrm{WL}}) = \sum_{t=1}^{T} \xi_t \cdot v(\theta_t, a(\theta_t, \mathbf{q}_t)) \,.$$

We have that

$$W^{\mathrm{WL}} = \frac{1+\lambda}{\lambda} \mathbb{E}\left[\liminf_{T\to\infty} \frac{W_T(\eta^{\mathrm{WL}})}{T}\right]$$

We introduce several definitions for the analysis. As is standard in Lyapunov analysis, let the Lyapunov function be the following quadratic function:

$$L(\mathbf{p}) = \frac{1}{2}\sum_{j\in\mathcal{J}} \mu_j p_j^2 \,.$$

Let $\mathbf{a}_t$ and $\mathbf{d}_t$ be the vectors representing the arriving agent and item at time t, respectively:

$$\mathbf{a}_t \triangleq \mathbf{e}_{a(\theta_t,\mathbf{q}_t)}\xi_t \,, \qquad \mathbf{d}_t \triangleq \mathbf{e}_{j_t}(1-\xi_t) \,.$$

Let $u_{j,t} \triangleq \max\left\{0, d_{j,t} - q_{j,t} - a_{j,t}\right\}$ denote the number of discarded items of type j at time t.[11] The evolution of the length of queue j is governed by

$$q_{j,t+1} = \left[q_{j,t} + a_{j,t} - d_{j,t}\right]^{+} = q_{j,t} + a_{j,t} - d_{j,t} + u_{j,t} \,, \quad \text{for each } j \in \mathcal{J}.$$

By Lemma 4.3 we have that

$$\begin{aligned}\mathbb{E}[v(\theta_t, a(\theta_t,\mathbf{q}_t))\xi_t|\mathbf{q}_t] = & \frac{\lambda}{1+\lambda} h(\mathbf{p}_t) - \frac{1}{c}\left(L(\mathbf{p}_t) - \mathbb{E}[L(\mathbf{p}_{t+1}) \mid \mathbf{q}_t]\right) \\ & - \frac{1}{2}\sum_{j\in\mathcal{J}} \frac{c}{\mu_j}\mathbb{E}[(a_{j,t}-d_{j,t})^2 + u_{j,t}^2 \mid \mathbf{q}_t]\end{aligned} \tag{4.8}$$

By Lemma 4.4 we have that $\frac{1}{2}\sum_{j\in\mathcal{J}} \frac{c}{\mu_j}\mathbb{E}[(a_{j,t}-d_{j,t})^2 + u_{j,t}^2 \mid \mathbf{q}_t] \le \frac{2+\lambda}{2(1+\lambda)}\Delta$. By Lemma 4.1 we have that $h(\mathbf{p}_t) \ge W^*$. Together, we have that

$$\mathbb{E}[v(\theta_t, a(\theta_t,\mathbf{q}_t))\xi_t|\mathbf{q}_t] \ge \frac{\lambda}{1+\lambda} W^* - \frac{1}{c}\left(L(\mathbf{p}_t) - \mathbb{E}[L(\mathbf{p}_{t+1}) \mid \mathbf{q}_t]\right) - \frac{2+\lambda}{2(1+\lambda)}\Delta \,.$$

[11]Recall that under our definition of the waiting list mechanism, an item is discarded it the item finds its corresponding queue to be empty when it arrives.

Therefore, we have that

$$
\begin{aligned}
\mathbb{E}\left[W_T(\eta^{\mathrm{WL}})\right] &= \mathbb{E}\left[\sum_{t=1}^{T} \xi_t \cdot v(\theta_t, a(\theta_t, \mathbf{q}_t))\right] \\
&= \mathbb{E}\left[\mathbb{E}\left[\sum_{t=1}^{T} \xi_t \cdot v(\theta_t, a(\theta_t, \mathbf{q}_t)) \mid \mathbf{q}_t\right]\right] \\
&= \mathbb{E}\left[\sum_{t=1}^{T} \mathbb{E}\left[\xi_t \cdot v(\theta_t, a(\theta_t, \mathbf{q}_t)) \mid \mathbf{q}_t\right]\right] \\
&\geq \mathbb{E}\left[\sum_{t=1}^{T} \frac{\lambda}{1+\lambda} W^* - \frac{1}{c}\left(L(\mathbf{p}_t) - \mathbb{E}[L(\mathbf{p}_{t+1} \mid \mathbf{q}_t]\right) - \frac{2+\lambda}{2(1+\lambda)}\Delta\right] \\
&= T\frac{\lambda}{1+\lambda} W^* - \frac{1}{c}\left(L(\mathbf{p}_1) - \mathbb{E}[L(\mathbf{p}_{T+1})]\right) - T\frac{2+\lambda}{2(1+\lambda)}\Delta\,. \qquad (4.9)
\end{aligned}
$$

By Lemma 4.5 we can translate the bound for $\mathbb{E}[W_T(\eta^{\mathrm{WL}})]$ to a bound for W^{WL}, i.e.,

$$
W^{\mathrm{WL}} = \frac{1+\lambda}{\lambda} \lim_{T\to\infty} \frac{1}{T} \mathbb{E}\left[W_T(\eta^{\mathrm{WL}})\right]\,.
$$

Plugging in (4.9) to the above equality, we have

$$
W^{\mathrm{WL}} \geq W^* - \frac{2+\lambda}{2\lambda}\Delta\,.
$$

This concludes the proof. □

Lemma 4.3. *If $c(w) = c \cdot w$, we have that*

$$
\begin{aligned}
\mathbb{E}[v(\theta_t, a(\theta_t, \mathbf{q}_t))\xi_t | \mathbf{q}_t] = &\frac{\lambda}{1+\lambda} h(\mathbf{p}_t) - \frac{1}{c}\left(L(\mathbf{p}_t) - \mathbb{E}[L(\mathbf{p}_{t+1}) \mid \mathbf{q}_t]\right) \\
&- \frac{1}{2}\sum_{j\in\mathcal{J}} \frac{c}{\mu_j} \mathbb{E}[(a_{j,t} - d_{j,t})^2 + u_{j,t}^2 \mid \mathbf{q}_t]\,.
\end{aligned}
$$

Proof. We have that the drift of Lyapunov function $L(\mathbf{p})$ in one period is

$$
\begin{aligned}
&L(\mathbf{p}_t) - L(\mathbf{p}_{t+1}) \\
&= \frac{1}{2}\sum_{j\in\mathcal{J}} \mu_j \left[p_{j,t}^2 - \left(p_{j,t} + \frac{c}{\mu_j}(a_{j,t} - d_{j,t} + u_{j,t})\right)^2\right] \\
&= \frac{1}{2}\sum_{j\in\mathcal{J}} \mu_j \left[p_{j,t}^2 - \left(p_{j,t} + \frac{c}{\mu_j}(a_{j,t} - d_{j,t})\right)^2 - \frac{c^2}{\mu_j^2} u_{j,t}^2 - \frac{2c}{\mu_j}\left(p_{j,t} + \frac{c}{\mu_j}(a_{j,t} - d_{j,t})\right) u_{j,t}\right]
\end{aligned}
$$

$$\stackrel{(a)}{=} \frac{1}{2}\sum_{j\in\mathcal{J}} \mu_j \left[p_{j,t}^2 - \left(p_{j,t} + \frac{c}{\mu_j}(a_{j,t} - d_{j,t}) \right)^2 - \frac{c^2}{\mu_j^2} u_{j,t}^2 \right] ,$$

where the equality (a) follows from the fact that $\left(p_{j,t} + \frac{c}{\mu_j}(a_{j,t} - d_{j,t})\right) u_{j,t} \equiv 0$ for all $j \in \mathcal{J}$.

We further simplify the Lyapunov drift as follows.

$$\begin{aligned} L(\mathbf{p}_t) - L(\mathbf{p}_{t+1}) &= \frac{1}{2}\sum_{j\in\mathcal{J}} \mu_j \left[p_{j,t}^2 - \left(p_{j,t} + \frac{c}{\mu_j}(a_{j,t} - d_{j,t}) \right)^2 \right] - \frac{1}{2}\sum_{j\in\mathcal{J}} \frac{c^2}{\mu_j} u_{j,t}^2 \\ &= -\frac{1}{2}\sum_{j\in\mathcal{J}} \mu_j \left[\frac{2c}{\mu_j}(a_{j,t} - d_{j,t}) p_{j,t} + \frac{c^2}{\mu_j^2}(a_{j,t} - d_{j,t})^2 \right] - \frac{1}{2}\sum_{j\in\mathcal{J}} \frac{c^2}{\mu_j} u_{j,t}^2 \\ &= -c\sum_{j\in\mathcal{J}} (a_{j,t} - d_{j,t}) p_{j,t} - \frac{1}{2}\sum_{j\in\mathcal{J}} \frac{c^2}{\mu_j} \left((a_{j,t} - d_{j,t})^2 + u_{j,t}^2 \right) . \end{aligned} \tag{4.10}$$

We expanded the expected value of the next arrival $\mathbb{E}\left[v(\theta_t, a(\theta_t, \mathbf{q}_t))\xi_t \mid \mathbf{q}_t\right]$ plus the expected value of the term $\langle \mathbf{p}_t, \mathbf{a}_t - \mathbf{d}_t \rangle$ to show these are related to the dual objective given in (4.5).

$$\begin{aligned} &\mathbb{E}\left[v(\theta_t, a(\theta_t, \mathbf{q}_t))\xi_t - \langle \mathbf{p}_t, \mathbf{a}_t - \mathbf{d}_t \rangle \mid \mathbf{q}_t\right] \\ &= \mathbb{E}\left[v(\theta_t, a(\theta_t, \mathbf{q}_t))\xi_t - \sum_{j\in\mathcal{J}} p_{j,t}(a_{j,t} - d_{j,t}) \mid \mathbf{q}_t\right] \\ &= \mathbb{E}\left[\max_{j\in\mathcal{J}_\emptyset} \left[v(\theta_t, j) - p_{j,t}\right] \xi_t + \sum_{j\in\mathcal{J}} p_{j,t} d_{j,t} \mid \mathbf{q}_t\right] \\ &= \mathbb{E}[\xi_t]\mathbb{E}\left[\max_{j\in\mathcal{J}_\emptyset} \left[v(\theta_t, j) - p_{j,t}\right] \mid \mathbf{q}_t\right] + \mathbb{E}\left[\sum_{j\in\mathcal{J}} p_{j,t} d_{j,t} \mid \mathbf{q}_t\right] \\ &= \frac{\lambda}{1+\lambda} \int_\Theta \max_{j\in\mathcal{J}_\emptyset} \left[v(\theta_t, j) - p_{j,t}\right] dF(\theta) + \frac{1}{1+\lambda}\sum_{j\in\mathcal{J}} \mu_j p_{j,t} \\ &= \frac{\lambda}{1+\lambda} h(\mathbf{p}_t) . \end{aligned} \tag{4.11}$$

By adding $c \cdot v(\theta_t, a(\theta_t, \mathbf{q}_t))\xi_t$ to both sides of equation (4.10), taking expectation conditional on $\mathbf{q}_t$ and applying equation (4.11) we obtain the required identity. □

Lemma 4.4. *If $c(w) = c \cdot w$, we have that for any $\mathbf{q}_t$,*

$$\frac{1}{2}\sum_{j\in\mathcal{J}} \frac{c}{\mu_j} \mathbb{E}[(a_{j,t} - d_{j,t})^2 + u_{j,t}^2 \mid \mathbf{q}_t] \leq \frac{2+\lambda}{2(1+\lambda)} \Delta .$$

Proof. We have

$$\frac{c}{2\mu_j}(a_{j,t} - d_{j,t})^2 + \frac{c}{2\mu_j}u_{j,t}^2 = \begin{cases} \frac{c}{\mu_j} & \text{if } d_{j,t} = 1 \text{ and } q_j = 0\,, \\ \frac{c}{2\mu_j} & \text{otherwise }\,. \end{cases}$$

That is, the term above is equal to $c/\mu_j \leq \Delta$ for arrivals that correspond to an item that is discarded because its queue is empty, and equal to $c/2\mu_j \leq \Delta/2$ for all other arrivals. Note that the probability that an item is discarded must be at most $1/(1+\lambda)$. We thus have

$$\begin{aligned} &\frac{1}{2}\sum_{j\in\mathcal{J}} \frac{c}{\mu_j}\mathbb{E}[(a_{j,t} - d_{j,t})^2 + u_{j,t}^2 \mid \mathbf{q}_t] \\ &\leq \frac{1}{1+\lambda}\Delta + \frac{\lambda}{1+\lambda}\frac{\Delta}{2} \\ &= \frac{2+\lambda}{2(1+\lambda)}\Delta\,. \end{aligned}$$

□

Lemma 4.5. *For the model defined in Section 4.3, we have that*

$$W^{\mathrm{WL}} = \frac{1+\lambda}{\lambda}\lim_{T\to\infty}\frac{1}{T}\mathbb{E}\left[W_T(\eta^{\mathrm{WL}})\right]\,.$$

Proof. We need to argue that the limiting operator and the expectation can be interchanged. We proceed in two steps.

First, we show that the stochastic process $\{(\xi_t, v(\theta_t, a(\theta_t, \mathbf{q}_t)))\}_{t\geq 0}$ is ergodic. Note that $v(\theta_t, a(\theta_t, \mathbf{q}_t))$ only depends on $\mathbf{q}_t$ and independent variables θ_t. The finite state Markov chain $\{\mathbf{q}_t\}_{t\geq 0}$ is irreducible and aperiodic,[12] therefore it has a unique steady state distribution and $\{(\xi_t, v(\theta_t, a(\theta_t, \mathbf{q}_t)))\}_{t\geq 0}$ is ergodic.

Second, we exchange the order of limit and expectation. It follows from the Birkhoff's ergodic theorem that $\frac{W_T(\eta^{\mathrm{WL}})}{A_T}$ converges almost surely to $\mathbb{E}[v_\infty|\xi_\infty = 1]$, where (ξ_∞, v_∞) is the steady state distribution of $(\xi_t, v(\theta_t, a(\theta_t, \mathbf{q}_t)))$. Since $\frac{W_T(\eta^{\mathrm{WL}})}{A_T}$ is non-negative and

[12]Irreducibility follows from the fact that all states can go to $\mathbf{0}$ with positive probability. Aperiodicity comes from the fact that the state can stay at $\mathbf{0}$ for an arbitrary number of periods.

uniformly bounded from above by $v_{\max}$ for all $T > 0$, we have

$$W^{\text{WL}} = \mathbb{E}\left[\lim_{T\to\infty} \frac{W_T(\eta^{\text{WL}})}{A_T}\right] = \lim_{T\to\infty} \mathbb{E}\left[\frac{W_T(\eta^{\text{WL}})}{A_T}\right] = \mathbb{E}[v_\infty|\xi_\infty = 1],$$

where we apply the bounded convergence theorem in the second equality to exchange the limits; the last equality holds because the boundedness of $\frac{W_T(\eta^{\text{WL}})}{A_T}$ and its almost sure convergence implies L_1 convergence. Finally, observe that

$$\begin{aligned}
\mathbb{E}[v_\infty] &= \mathbb{E}[v_\infty|\xi_\infty = 1] \cdot \mathbb{P}(\xi_\infty = 1) + \mathbb{E}[v_\infty|\xi_\infty = 0] \cdot \mathbb{P}(\xi_\infty = 0) \\
&= \mathbb{E}[v_\infty|\xi_\infty = 1] \cdot \frac{\lambda}{1+\lambda} + 0 \\
&= W^{\text{WL}} \cdot \frac{\lambda}{1+\lambda},
\end{aligned}$$

where the second equality follows from the fact that all the rewards are collected when agents arrive, i.e., $\xi_t = 1$. Note that

$$\lim_{T\to\infty} \frac{1}{T}\mathbb{E}\left[W_T(\eta^{\text{WL}})\right] = \mathbb{E}[v_\infty],$$

we have

$$W^{\text{WL}} = \frac{1+\lambda}{\lambda}\mathbb{E}[v_\infty] = \frac{1+\lambda}{\lambda} \lim_{T\to\infty} \frac{1}{T}\mathbb{E}\left[W_T(\eta^{\text{WL}})\right].$$

□

4.4.1 Asymptotic Optimality

Theorem 4.1 bounds the allocate efficiency loss in the waiting list. The following corollaries imply that the loss tends to zero when agents' patience increases or when the market size grows large.

As agents are increasingly patient, the marginal expected waiting cost becomes smaller. For example if $c(w) = c \cdot w$, then $\Delta = c/\mu_{\min} \to 0$ as $c \to 0$. This implies:

Corollary 4.1. *Fix $F, \{\mu\}_{j\in\mathcal{J}}, \lambda$ and consider a sequence of markets indexed by ℓ in which waiting cost is $c(w) = c_\ell \cdot w$. Let W_ℓ^{WL} denote the allocative efficiency of the waiting list* *for market ℓ. If $c_\ell \to 0$ as $\ell \to \infty$, then*

$$W_\ell^{\mathrm{WL}} \xrightarrow[\ell\to\infty]{} W^{\mathrm{OPT}} .$$

Note that W^{OPT} and the optimal prices $\mathbf{p}^*$ are independent of c_ℓ. As c_ℓ tends to zero, the length of queue j that generates the optimal price, $q_j^* = p_j^*/c_\ell$, is increasing. Moreover, there are smaller fluctuations in queue lengths and there fore in prices.This implies lower variation in prices and higher allocative efficiency.

If the market thickens in the sense that arrivals of agents and items become more frequent, the expected cost of waiting for a single arrival becomes lower. For example, if $c(w) = c \cdot w$ and $\mu_{\min} \to \infty$, then $\Delta = c/\mu_{\min} \to 0$. Therefore:

Corollary 4.2. *Fix $F, \{\mu\}_{j\in\mathcal{J}}, \lambda$ and $c(\cdot)$. Consider a sequence of markets indexed by ℓ in the agent arrival rate is $\ell \cdot \lambda$ and the arrival rate of item j is $\ell \cdot \mu_j$. Let W_ℓ^{WL} denote the allocative efficiency of the waiting list for market ℓ. If $\ell \to \infty$ then*

$$W_\ell^{\mathrm{WL}} \xrightarrow[\ell\to\infty]{} W^{\mathrm{OPT}} .$$

4.4.2 Lower Bound for the Allocative Efficiency Loss

A natural question is whether the bound given in Theorem 4.1 is tight. We give a lower bound for the allocative efficiency loss by constructing an economy in which the loss is approximately Δ.

Example 4.1. *Consider a dynamic market in which $\Theta = \mathcal{J}$, that is, the set of items is $\mathcal{J} = \{1, 2, \ldots, J\}$ and there is a corresponding agent type for each agent type. The distribution of agent types and item types are uniform, i.e., $\mathbb{P}(\theta = j) = \mu_j = 1/J$, $\forall j \in \mathcal{J}$. The total agent arrival rate is $\lambda = 1$, and the waiting cost is linear, i.e. $c(w) = c \cdot w$. The value of agent θ for item j is*

$$v(\theta, j) = \begin{cases} \gamma & \text{if } \theta = j\,, \\ 0 & \text{if } \theta \neq j\,. \end{cases}$$

Proposition 4.2. *For any J and any $\Delta > 0$ and $\varepsilon > 0$, there exists a market with J items as in Example 4.1 in which the allocative efficiency loss under the waiting list is*

$$W^{\text{OPT}} - W^{\text{WL}} \geq \Delta - \varepsilon$$

Proof of Proposition 4.2. We prove the result by calculating the allocative efficiency loss in the market of Example 4.1. By Proposition 4.1, we have that $W^{\text{OPT}} = \gamma$.

Under the waiting list, an agent of type θ will only join the queue for item $j = \theta$. An agent arriving in epoch t of type $\theta_t = j$ will choose to join queue j to receive a value of 1 only if

$$\gamma \geq p_j(q_j) = \frac{c}{\mu_j}(1 + q_{t,j}) = \Delta(1 + q_{t,j}) \ ,$$

or

$$q_{t,j} \leq \frac{\gamma}{\Delta} - 1 \ .$$

Therefore, the possible states of each queue $j \in \mathcal{J}$ are $0, 1, \ldots, K$ with $K = \lfloor \gamma/\Delta \rfloor$. Let $\pi_j(k)_{0 \leq k \leq K}$ denote the steady distributions over the length of queue j. Because the length of the queue follows a reflected unbiased random walk, all states are equally likely and $\pi_j(k) = \frac{1}{K+1}$.[13]

The allocative efficiency under the waiting list is given by

$$\begin{aligned} W^{\text{OPT}} - W^{\text{WL}} &= \gamma - \sum_{j \in \Theta} F(j) \left(\pi_j(K) \cdot 0 + \sum_{k<K} \pi_j(k) \cdot \gamma \right) \\ &= \gamma - J \frac{1}{J} \left(\frac{1}{K+1} \cdot 0 + \frac{K}{K+1} \cdot \gamma \right) \\ &= \frac{1}{K+1} \cdot \gamma \\ &= \frac{1}{\lfloor \gamma/\Delta \rfloor + 1} \cdot \gamma \end{aligned}$$

By choosing γ such that $\lfloor \gamma/\Delta \rfloor \approx \gamma/\Delta - 1$ we get that

$$W^{\text{OPT}} - W^{\text{WL}} \geq \frac{1}{\gamma/\Delta - 1 + 1} \cdot \gamma - \varepsilon$$

[13] To see this directly, observe that equating probability flows across a cut gives for any $0 < k < K$ that $\pi_j(k)\lambda/J = \pi_j(k+1)\mu_j$, which implies that $\pi_j(0) = \pi_j(1) = \cdots = \pi_j(K)$.

$$= \Delta - \varepsilon$$

□

4.4.3 Welfare of the Queueing Mechanism

As the market grows large, the welfare converges to the optimal allocation minus the predicted waiting times. This follows from the following result that restricts attention to the convergence of prices.

Proposition 4.3. *Fix $F, \{\mu\}_{j\in\mathcal{J}}, \lambda$ and consider a sequence of markets indexed by ℓ in which waiting cost is $c(w) = c_\ell \cdot w$. Let $\mathcal{P}^*$ be the set of minimizers of the dual function $h(\mathbf{p})$, let $\mathbf{p}_{\ell,\infty}$ be the random variable of steady-state price in the market indexed by ℓ. If $c_\ell \to 0$ as $\ell \to \infty$, then*

$$\limsup_{\ell\to\infty} \ \mathbb{P}\left(\mathbf{p}_{\ell,\infty} \notin \mathcal{P}^*\right) = \ 0\,.$$

Proof sketch. Let $\mathbf{p}^*$ be any minimizer of the dual function $h(\mathbf{p})$. Consider the following centered Lyapunov function:

$$\bar{L}(\mathbf{p}) = \ \frac{1}{2}\sum_{j\in\mathcal{J}} \mu_j (p_j - p_j^*)^2\,.$$

We first fix ℓ and omit the subscript ℓ. Recall that $p_{j,t+1} = p_{j,t} + \frac{c}{\mu_j}(a_{j,t} - d_{j,t} + u_{j,t})$. Similar to the proof of Lemma 4.3, we have

$$\begin{aligned}
&\bar{L}(\mathbf{p}_{t+1}) - \bar{L}(\mathbf{p}_t)\\
&= \frac{1}{2}\sum_{j\in\mathcal{J}} \mu_j \left(\frac{c^2}{\mu_j^2}(a_{j,t} - d_{j,t} + u_{j,t})^2 + \frac{2c}{\mu_j}(p_{j,t} - p_j^*)(a_{j,t} - d_{j,t} + u_{j,t})\right)\\
&= c\langle \mathbf{p}_t - \mathbf{p}^*, \mathbf{a}_t - \mathbf{d}_t\rangle + \frac{1}{2}\sum_{j\in\mathcal{J}} \mu_j \left(\frac{c^2}{\mu_j^2}(a_{j,t} - d_{j,t} + u_{j,t})^2 + \frac{2c}{\mu_j}(p_{j,t} - p_j^*)u_{j,t}\right)\\
&\overset{(a)}{\le} c\langle \mathbf{p}_t - \mathbf{p}^*, \mathbf{a}_t - \mathbf{d}_t\rangle + \frac{1}{2}\sum_{j\in\mathcal{J}} \mu_j \left(\frac{c^2}{\mu_j^2}(a_{j,t} - d_{j,t} + u_{j,t})^2 + \frac{2c}{\mu_j}p_{j,t} \cdot u_{j,t}\right)\\
&= c\langle \mathbf{p}_t - \mathbf{p}^*, \mathbf{a}_t - \mathbf{d}_t\rangle + \frac{1}{2}\sum_{j\in\mathcal{J}} \mu_j \left(\frac{c^2}{\mu_j^2}\left((a_{j,t} - d_{j,t})^2 + u_{j,t}^2\right) + \frac{2c}{\mu_j}\left(p_{j,t} + \frac{c}{\mu_j}(a_{j,t} - d_{j,t})\right)u_{j,t}\right)\,.
\end{aligned}$$

Here inequality (a) is because $p_j^* \geq 0$ and $u_{j,t} \geq 0$. Note that $(p_{j,t} + \frac{c}{\mu_j}(a_{j,t} - d_{j,t}))u_{j,t} \equiv 0$. Using arguments similar to those in the proof of Lemma 4.4, we have

$$\bar{L}(\mathbf{p}_{t+1}) - \bar{L}(\mathbf{p}_t) \leq \ c\langle \mathbf{p}_t - \mathbf{p}^*, \mathbf{a}_t - \mathbf{d}_t \rangle + c \cdot \Delta \,.$$

Take expectations on both sides conditioned on $\mathbf{q}_t$, and use the fact that $\mathbb{E}[\mathbf{a}_t - \mathbf{d}_t | \mathbf{q}_t] \in -\partial h(\mathbf{p}_t)$, we have

$$\mathbb{E}[\bar{L}(\mathbf{p}_{t+1})|\mathbf{q}_t] - \bar{L}(\mathbf{p}_t) \leq \ -c\left(h(\mathbf{p}_t) - h(\mathbf{p}^*)\right) + c \cdot \Delta \,. \tag{4.12}$$

For $\epsilon > 0$, let $B(\epsilon) \triangleq \max\{\bar{L}(\mathbf{p}) : h(\mathbf{p}) \leq h(\mathbf{p}^*) + \epsilon\}$. It is straightforward to verify that for the Markov chain $\{\mathbf{p}_t\}_{t=1}^{\infty}$ and $\Delta \leq \frac{\epsilon}{2}$, the Lyapunov function $\bar{L}(\cdot)$ has negative drift $-\frac{c\epsilon}{2}$ when $h(\mathbf{p}) \geq h(\mathbf{p}^*) + \epsilon$. Notice that the maximum increase of $\bar{L}(\mathbf{p})$ in one period is $cv_{\max}$. Applying Theorem 1 in [57], we have

$$\mathbb{P}(\bar{L}(\mathbf{p}_\infty) > B(\epsilon) + 2cv_{\max}m) \leq \left(\frac{2v_{\max}}{2v_{\max} + \epsilon}\right)^{m+1}$$

Let $m = \frac{B(\epsilon)}{cv_{\max}}$. Let $\ell \to \infty$ therefore $c_\ell \to 0$, and this shows that $\mathbb{P}(\bar{L}(\mathbf{p}_{\ell,\infty}) > 3B(\epsilon)) \to 0$. Since ϵ can be chosen arbitrarily, this concludes the proof. □

4.5 Exponentially Small Loss for Generic Problems

The magnitude of the loss generated by miss-allocations depends naturally on the parameters of the economy and this section further isolates sources of the inefficiency. Indeed it is shown here that in economies with finitely many agents and item types, the loss is generically much smaller than the stated in Theorem 4.1. The following natural assumption will drive the result.

Assumption 4.1 (Unique shadow price)**.** *The dual problem* (4.5) *with finite types of items and agents has a unique minimizer* $\mathbf{p}^*$.

Assumption 4.1 holds generically when there are finitely many agent types.

Proposition 4.4. *Let $\boldsymbol{\lambda} \in \mathbb{R}_{++}^{|\Theta|}$ be the vector of arrival rates of each type of agent. The vectors $(\mathbf{v}, \boldsymbol{\lambda}, \boldsymbol{\mu})$ satisfying Assumption 4.1 are open and dense in $\mathbb{R}_{++}^{|\Theta|\times|\mathcal{J}|+|\Theta|+|\mathcal{J}|}$.*

Proof of Proposition 4.4. We say that the problem instance $(\boldsymbol{\lambda}, \boldsymbol{\mu})$ satisfies *generalized imbalance (GI)* if there are no pair of nonempty subsets of agent types $\mathcal{I}' \subset \Theta$ and item types $\mathcal{J}' \subset \mathcal{J}$ such that the total arrival rate of agents with type in $\mathcal{I}'$ exactly matches the total arrival rate of items with types in $\mathcal{J}'$. The proposition holds because the problem instances satisfying GI are open and dense in $\mathbb{R}_{++}^{|\mathcal{I}|+|\mathcal{J}|}$, and that GI implies Assumption 4.1 (see Proposition C.2 of [119]). □

Theorem 4.2. *Suppose there are a finite number of agent types, the primitives $\mathbf{v} = (v(\theta, j))_{\theta\in\Theta, j\in\mathcal{J}}$, $\boldsymbol{\lambda} = (\lambda_\theta)_{\theta\in\Theta}$, $\boldsymbol{\mu} = (\mu_j)_{j\in\mathcal{J}}$ satisfy Assumption 4.1, and consider a sequence of markets indexed by ℓ in which waiting cost is $c(w) = c_\ell \cdot w$. If $c_\ell \to 0$ as $\ell \to \infty$, then there exists $\alpha = \alpha(\mathbf{v}, \boldsymbol{\lambda}, \boldsymbol{\mu}) > 0$, such that*

$$W^{\mathrm{OPT}} - W_\ell^{\mathrm{WL}} = O\left(\exp(-\alpha/\Delta_\ell)\right) \quad \ell \to \infty\,.$$

The intuition behind Theorem 4.2 is the following. Assumption 4.1 led to the robustness of the optimal dual prices; indeed as long as prices are within some $\delta > 0$ distance from the unique dual prices $\mathbf{p}^*$, items are allocated efficiently due to complementary slackness. And when prices are further away from $\mathbf{p}^*$, they adjust towards the optimal dual prices at some minimal rate $\gamma > 0$. So prices follow a biased random walk towards $\mathbf{p}^*$, implying that that they are far from the optimal dual prices with probability that is exponentially small in step size Δ_ℓ's inverse. The exponent α in the theorem is roughly proportional to the product of δ and γ.

The remainder of this section formalizes this intuition. For this purpose consider an economy that satisfies Assumption 4.1 with the unique dual price $\mathbf{p}^*$ and assume that agents have linear waiting costs as stated in the theorem.

Robustness of dual prices. The following sets will serve to define the robustness of dual prices. Define the set of *active agent types* as:

$$\Theta^* \triangleq \left\{\theta \in \Theta : \max_{j\in\mathcal{J}} \left(v(\theta, j) - p_j^*\right) > 0\right\},$$

the set of *active item types* as:

$$\mathcal{J}^* \triangleq \left\{j \in \mathcal{J} : p_j^* > 0\right\},$$

and for each type of agent $\theta \in \Theta$, its set of *active matches* as:

$$\mathcal{J}_\theta^* \triangleq \text{argmax}_{j\in\mathcal{J}}[v(\theta, j) - p_j^*]^+ .$$

From complementary slackness conditions, an allocation $\mathbf{x}$ is an optimal solution of the static allocation problem (4.4) *if and only if* (i) all active agent types are assigned items, (ii) all active item types are assigned to agents, (iii) for each (θ, j) such that $x_{\theta,j} > 0$, (θ, j) is an active match, and, (iv) $\mathbf{x}$ satisfies the primal constraints.

For uniquely defined $\mathbf{p}^*$, observe that there exists a constant $\delta > 0$ such that when the price is within δ of $\mathbf{p}^*$, conditions (i)-(iii) are satisfied.

Definition 4.2. *Consider a problem instance* $(\mathbf{v}, \boldsymbol{\lambda}, \boldsymbol{\mu})$ *that satisfies Assumption 4.1. Let*

$$\delta_{\boldsymbol{\lambda}} \triangleq \min_{\theta\in\Theta^*} \max_{j\in\mathcal{J}} \left(v(\theta, j) - p_j^*\right),$$

$$\delta_{\boldsymbol{\mu}} \triangleq \min_{j\in\mathcal{J}^*} p_j^*,$$

$$\delta_{\mathbf{v}} \triangleq \min_{\{\theta\in\Theta:\mathcal{J}_\theta^*\neq\emptyset\}} \min_{j\in\mathcal{J}_\theta^*, j'\notin\mathcal{J}_\theta^*} \left([v(\theta, j) - p_j^*]^+ - v(\theta, j') + p_{j'}^*\right).$$

We refer to $\delta \triangleq \min\{\delta_{\boldsymbol{\lambda}}, \delta_{\boldsymbol{\mu}}, \delta_{\mathbf{v}}\} > 0$ *as the* robustness of dual price *of the problem instance.*

Define the following set of near-optimal dual prices.

$$\mathcal{P} \triangleq \left\{\mathbf{p} : ||\mathbf{p} - \mathbf{p}^*||_\infty < \frac{\delta}{2}\right\}.$$

The following lemma and its proof states that the loss arises only when $\mathbf{p} \notin \mathcal{P}$.

Lemma 4.6. *If prices in steady state satisfy that $\mathbb{P}(\mathbf{p}_\infty \notin \mathcal{P}) \leq \kappa$, then*

$$W^{\mathrm{OPT}} - W^{\mathrm{WL}} \leq \frac{1+\lambda}{\lambda} v_{\max}\kappa \,.$$

Proof. Let $x_{\theta,j,t}$ be the match in the t-th period, which equals to 1 if a type θ agent is assigned a type j item in the t-th period, and equals to zero otherwise. For prices that satisfy $\mathbf{p}_t \in \mathcal{P}$, it holds by definition (robustness of the dual price) that

$$\begin{aligned}
\sum_{j\in\mathcal{J}} x_{\theta,j,t} =&\ \mathbb{1}\{\text{a type } \theta \text{ agent arrives at } t\}\,, && \forall \theta \in \Theta^* \\
\sum_{\theta\in\Theta} x_{\theta,j,t} =&\ \mathbb{1}\{\text{a type } j \text{ item arrives at } t\}\,, && \forall j \in \mathcal{J}^* \\
x_{\theta,j,t} = 0\,, & && \forall \theta \in \Theta\,, \theta \notin \mathcal{J}^*_\theta\,.
\end{aligned}$$

Denote by w_∞ the match value obtained in a period in steady state (by the Arrival Theorem, the steady state of the discrete time process we are considering is the same as the steady state distribution of the original continuous time process). Note that

$$W^{\mathrm{WL}} = \frac{1+\lambda}{\lambda}\mathbb{E}[w_\infty] \geq \frac{1+\lambda}{\lambda}\mathbb{E}[w_\infty | \mathbf{p}_\infty \in \mathcal{P}] \cdot \mathbb{P}(\mathbf{p}_\infty \in \mathcal{P})\,. \tag{4.13}$$

The last term on the RHS is bounded from below $1-\kappa$ by assumption. It remains to show $\frac{1+\lambda}{\lambda}\mathbb{E}[w_\infty|\mathbf{p}_\infty \in \mathcal{P}] = W^{\mathrm{OPT}}$ because

$$W^{\mathrm{WL}} \geq (1-\kappa)W^{\mathrm{OPT}} \geq W^{\mathrm{OPT}} - \frac{1+\lambda}{\lambda} v_{\max}\kappa\,.$$

By linearity of expectation, we have

$$\mathbb{E}[w_\infty|\mathbf{p}_\infty \in \mathcal{P}] = \sum_{\theta\in\Theta, j\in\mathcal{J}} v(\theta,j)\mathbb{E}[x_{\theta,j,\infty}|\mathbf{p}_\infty \in \mathcal{P}]\,.$$

Note that

$$\begin{aligned}
\sum_{j\in\mathcal{J}} \mathbb{E}[x_{\theta,j,\infty}|\mathbf{p}_\infty \in \mathcal{P}] =&\ \lambda_\theta\,, && \forall \theta \in \Theta^* && (4.14)\\
\sum_{\theta\in\Theta} \mathbb{E}[x_{\theta,j,\infty}|\mathbf{p}_\infty \in \mathcal{P}] =&\ \mu_j\,, && \forall j \in \mathcal{J}^* && (4.15)\\
\mathbb{E}[x_{\theta,j,\infty}|\mathbf{p}_\infty \in \mathcal{P}] =&\ 0\,, && \forall \theta \in \Theta\,, j \notin \mathcal{J}^*_\theta\,. && (4.16)
\end{aligned}$$

This completes the proof since equalities (4.14)-(4.16) correspond to the complementary slackness condition of the static allocation problem. □

Rate of adjustment. Next we show that prices are very likely to be in $\mathcal{P}$, by showing that prices that deviates from $\mathbf{p}^*$ quickly adjusts back. The rate of price adjustment is related to the "sharpness" of the dual objective (4.5).

Lemma 4.7 (Geometry of dual function)**.** *Suppose Assumption 4.1 holds. Then there exists* $\gamma(\mathbf{v}, \boldsymbol{\lambda}, \boldsymbol{\mu}) > 0$ *such that for any* $\mathbf{p} \in \mathbb{R}^{|\mathcal{J}|}$*, we have*

$$h(\mathbf{p}) - h(\mathbf{p}^*) \geq \gamma(\mathbf{v}, \boldsymbol{\lambda}, \boldsymbol{\mu})||\mathbf{p} - \mathbf{p}^*||_2 \,. \tag{4.17}$$

Moreover,

$$\gamma(\mathbf{v}, \boldsymbol{\lambda}, \boldsymbol{\mu}) \geq \frac{1}{\lambda}\left\{\min_{\mathcal{I} \subset \mathcal{J}} \frac{\lambda - \sum_{j \in \mathcal{I}} \mu_j}{\sqrt{|\mathcal{I}|}}\,, \min_{\{\mathcal{I}: \mathcal{I} \subset \mathcal{J}, \mathcal{I} \supset \cup_{\theta \in \Theta^*} \mathcal{J}^*_\theta\}} \frac{\sum_{j \in \mathcal{I}} \mu_j - \sum_{\theta \in \Theta^*} \lambda_\theta}{\sqrt{|\mathcal{I}|}}\,, \min_{j \in \mathcal{J}} \mu_j\right\}.$$

The proof of this technical lemma appear in Appendix D.4.1. The proof shows that the rate of adjustment is positive for $\mathbf{p} \in \mathcal{P}$ and by convexity of $h(\mathbf{p}) - h(\mathbf{p}^*)$ this holds also also for prices not in $\mathcal{P}$. The proof Theorem 4.2 can now be completed.

current progress

Proof of Theorem 4.2. Using Lemma 4.6, we have that

$$W^{\mathrm{OPT}} - W^{\mathrm{WL}} \leq \frac{1+\lambda}{\lambda} v_{\max} \mathbb{P}(\mathbf{p}_\infty \notin \mathcal{P}) \,.$$

Lemma D.5 in the appendix provides the following concentration bound (using further Lyapunov analysis and using the established bound on the rate of adjustment):

$$\mathbb{P}(\mathbf{p}_\infty \notin \mathcal{P}) \leq \exp\left(-\log\left(1 + \frac{\gamma \mu_{\min}}{4}\right)\left(\frac{\delta}{12\Delta}\right)\right).$$

Therefore the allocative efficiency loss is bounded by:

$$W^{\mathrm{OPT}} - W^{\mathrm{WL}} \leq \frac{1+\lambda}{\lambda} v_{\max} \, \exp\left(-\log\left(1 + \frac{\gamma \mu_{\min}}{4}\right)\left(\frac{\delta}{12\Delta}\right)\right).$$

where δ is the robustness of price, and γ is the rate of price adjustment. □

4.6 Conclusion

This paper considers a dynamic economy, in which the waiting times play the role of prices in guiding the allocation and rationing items. It studies the impact of the fluctuations in waiting times resulting from the stochasticity in the arrival of demand and supply, and quantifies the allocative efficiency loss resulting from this fluctuation. We observe that the efficiency loss compared to an optimum offline assignment is fairly small and is captured by the marginal increase in expected waiting cost from having one more agent in the queue. Furthermore, when equilibrium waiting times are essentially unique, there is almost no efficiency loss. Our results show that despite the decentralized nature of these markets and the underlying stochasticity, simple waiting mechanisms obtain near optimum allocative efficiency. They also justify deterministic or "fluid" approximations for modeling the behavior of such queueing systems.

References

[1] S. Banerjee, D. Freund, and T. Lykouris, "Pricing and optimization in shared vehicle systems: An approximation framework", *CoRR abs/1608.06819*, 2016.

[2] A. Braverman, J. G. Dai, X. Liu, and L. Ying, "Empty-car routing in ridesharing systems", *Operations Research*, vol. 67, no. 5, pp. 1437–1452, 2019.

[3] S. Banerjee, Y. Kanoria, and P. Qian, "Dynamic assignment control of a closed queueing network under complete resource pooling", *arXiv preprint arXiv:1803.04959*, 2018, Also in ACM SIGMETRICS 2018.

[4] L. Tassiulas and A. Ephremides, "Stability properties of constrained queueing systems and scheduling policies for maximum throughput in multihop radio networks", *IEEE transactions on automatic control*, vol. 37, no. 12, pp. 1936–1948, 1992.

[5] L. Georgiadis, M. J. Neely, L. Tassiulas, *et al.*, "Resource allocation and cross-layer control in wireless networks", *Foundations and Trends® in Networking*, vol. 1, no. 1, pp. 1–144, 2006.

[6] S. T. Maguluri and R Srikant, "Heavy traffic queue length behavior in a switch under the maxweight algorithm", *Stochastic Systems*, vol. 6, no. 1, pp. 211–250, 2016.

[7] J. G. Dai and W. Lin, "Maximum pressure policies in stochastic processing networks", *Operations Research*, vol. 53, no. 2, pp. 197–218, 2005.

[8] A. L. Stolyar, "Maximizing queueing network utility subject to stability: Greedy primal-dual algorithm", *Queueing Systems*, vol. 50, no. 4, pp. 401–457, 2005.

[9] L. Huang and M. J. Neely, "Utility optimal scheduling in processing networks", *Performance Evaluation*, vol. 68, no. 11, pp. 1002–1021, 2011.

[10] I. Gurvich and A. Ward, "On the dynamic control of matching queues", *Stochastic Systems*, vol. 4, no. 2, pp. 479–523, 2014.

[11] L. Jiang and J. Walrand, "Stable and utility-maximizing scheduling for stochastic processing networks", in *2009 47th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, IEEE, 2009, pp. 1111–1119.

[12] A. S. Nemirovsky and D. B. Yudin, "Problem complexity and method efficiency in optimization.", 1983.

[13] A. Beck and M. Teboulle, "Mirror descent and nonlinear projected subgradient methods for convex optimization", *Operations Research Letters*, vol. 31, no. 3, pp. 167–175, 2003.

[14] A. Eryilmaz and R Srikant, "Fair resource allocation in wireless networks using queue-length-based scheduling and congestion control", *IEEE/ACM Transactions on Networking (TON)*, vol. 15, no. 6, pp. 1333–1344, 2007.

[15] J. M. Harrison, "Brownian models of open processing networks: Canonical representation of workload", *Annals of Applied Probability*, pp. 75–103, 2000.

[16] A. L. Stolyar, "Maxweight scheduling in a generalized switch: State space collapse and workload minimization in heavy traffic", *Annals of Applied Probability*, vol. 14, no. 1, pp. 1–53, 2004.

[17] J. G. Dai and W. Lin, "Asymptotic optimality of maximum pressure policies in stochastic processing networks", *Annals of Applied Probability*, vol. 18, no. 6, pp. 2239–2299, 2008.

[18] A. Eryilmaz and R Srikant, "Asymptotically tight steady-state queue length bounds implied by drift conditions", *Queueing Systems*, vol. 72, no. 3-4, pp. 311–359, 2012.

[19] V. Gupta and A. Radovanović, "Interior-point-based online stochastic bin packing", *Operations Research*, 2020.

[20] S. Bubeck, M. B. Cohen, Y. T. Lee, J. R. Lee, and A. Mądry, "K-server via multiscale entropic regularization", in *Proceedings of the 50th annual ACM SIGACT symposium on theory of computing*, 2018, pp. 3–16.

[21] E. Özkan and A. R. Ward, "Dynamic matching for real-time ride sharing", *Stochastic Systems*, vol. 10, no. 1, pp. 29–70, 2020.

[22] S. R. Balseiro, D. B. Brown, and C. Chen, "Dynamic pricing of relocating resources in large networks", *Management Science*, 2020.

[23] H. Ma, F. Fang, and D. C. Parkes, "Spatio-temporal pricing for ridesharing platforms", in *Proceedings of the 2019 ACM Conference on Economics and Computation*, 2019, pp. 583–583.

[24] H. Chung, D. Freund, and D. B. Shmoys, "Bike angels: An analysis of citi bike's incentive program", in *Proceedings of the 1st ACM SIGCAS Conference on Computing and Sustainable Societies*, ACM, 2018, p. 5.

[25] K. Johnson, D. Simchi-Levi, and P. Sun, "Analyzing scrip systems", *Operations Research*, vol. 62, no. 3, pp. 524–534, 2014.

[26] N. Agarwal, I. Ashlagi, E. Azevedo, C. R. Featherstone, and Ö. Karaduman, "Market failure in kidney exchange", *American Economic Review*, vol. 109, no. 11, pp. 4026–70, 2019.

[27] R. Caldentey, E. H. Kaplan, and G. Weiss, "Fcfs infinite bipartite matching of servers and customers", *Advances in Applied Probability*, vol. 41, no. 3, pp. 695–730, 2009.

[28] I. Adan and G. Weiss, "A loss system with skill-based servers under assign to longest idle server policy", *Probability in the Engineering and Informational Sciences*, vol. 26, no. 3, pp. 307–321, 2012.

[29] A. Bušić and S. Meyn, "Approximate optimality with bounded regret in dynamic matching models", *ACM SIGMETRICS Performance Evaluation Review*, vol. 43, no. 2, pp. 75–77, 2015.

[30] J. Mairesse and P. Moyal, "Stability of the stochastic matching model", *Journal of Applied Probability*, vol. 53, no. 4, pp. 1064–1077, 2016.

[31] G. Gallego and G. Van Ryzin, "Optimal dynamic pricing of inventories with stochastic demand over finite horizons", *Management science*, vol. 40, no. 8, pp. 999–1020, 1994.

[32] K. T. Talluri and G. J. Van Ryzin, *The theory and practice of revenue management*. Springer Science & Business Media, 2006, vol. 68.

[33] P. Bumpensanti and H. Wang, "A re-solving heuristic with uniformly bounded loss for network revenue management", *Management Science*, vol. 66, no. 7, pp. 2993–3009, 2020.

[34] W. C. Jordan and S. C. Graves, "Principles on the benefits of manufacturing process flexibility", *Management Science*, vol. 41, no. 4, pp. 577–594, 1995.

[35] A. Désir, V. Goyal, Y. Wei, and J. Zhang, "Sparse process flexibility designs: Is the long chain really optimal?", *Operations Research*, vol. 64, no. 2, pp. 416–431, 2016.

[36] C. Shi, Y. Wei, and Y. Zhong, "Process flexibility for multiperiod production systems", *Operations Research*, vol. 67, no. 5, pp. 1300–1320, 2019.

[37] M. J. Neely, "Stochastic network optimization with application to communication and queueing systems", *Synthesis Lectures on Communication Networks*, vol. 3, no. 1, pp. 1–211, 2010.

[38] L. Huang and M. J. Neely, "Delay reduction via lagrange multipliers in stochastic network optimization", in *Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks, 2009. WiOPT 2009. 7th International Symposium on*, IEEE, 2009, pp. 1–10.

[39] X. Lin and N. B. Shroff, "Joint rate control and scheduling in multihop wireless networks", in *2004 43rd IEEE Conference on Decision and Control (CDC)(IEEE Cat. No. 04CH37601)*, IEEE, vol. 2, 2004, pp. 1484–1489.

[40] J. Sweeney and R. J. Sweeney, "Monetary theory and the great capitol hill baby sitting co-op crisis: Comment", *Journal of Money, Credit and Banking*, vol. 9, no. 1, pp. 86–89, 1977.

[41] D. P. Bertsekas, *Dynamic programming and optimal control*, 2. Athena scientific Belmont, MA, 1995, vol. 1.

[42] A. Waserhole and V. Jost, "Pricing in vehicle sharing systems: Optimization in queuing networks with product forms", *EURO Journal on Transportation and Logistics*, vol. 5, no. 3, pp. 293–320, 2016.

[43] I. A. Kash, E. J. Friedman, and J. Y. Halpern, "An equilibrium analysis of scrip systems", *ACM Transactions on Economics and Computation (TEAC)*, vol. 3, no. 3, pp. 1–32, 2015.

[44] J. M. Harrison and M. J. López, "Heavy traffic resource pooling in parallel-server systems", *Queueing systems*, vol. 33, no. 4, pp. 339–368, 1999.

[45] V. Venkataramanan and X. Lin, "On the queue-overflow probability of wireless systems: A new approach combining large deviations with lyapunov functions", *IEEE Transactions on Information Theory*, vol. 59, no. 10, pp. 6367–6392, 2013.

[46] A. L. Stolyar and K. Ramanan, "Largest weighted delay first scheduling: Large deviations and optimality", *Annals of Applied Probability*, pp. 1–48, 2001.

[47] A. L. Stolyar, "Control of end-to-end delay tails in a multiclass network: Lwdf discipline optimality", *Annals of Applied Probability*, pp. 1151–1206, 2003.

[48] K. Majewski and K. Ramanan, "How large queue lengths build up in a jackson network", 2008.

[49] J. Blanchet, "Optimal sampling of overflow paths in jackson networks", *Mathematics of Operations Research*, vol. 38, no. 4, pp. 698–719, 2013.

[50] E. J. Friedman, J. Y. Halpern, and I. Kash, "Efficiency and nash equilibria in a scrip system for p2p networks", in *Proceedings of the 7th ACM conference on Electronic commerce*, 2006, pp. 140–149.

[51] I. A. Kash, E. J. Friedman, and J. Y. Halpern, "Optimizing scrip systems: Crashes, altruists, hoarders, sybils and collusion", *Distributed Computing*, vol. 25, no. 5, pp. 335–357, 2012.

[52] S. Jasin and S. Kumar, "A re-solving heuristic with bounded revenue loss for network revenue management with customer choice", *Mathematics of Operations Research*, vol. 37, no. 2, pp. 313–345, 2012.

[53] J. Marshall Hall, "Combinatorial theory", *NY: JohnWiley& Sons*, pp. 238–243, 1986.

[54] A. Dembo and O. Zeitouni, *Large deviations techniques and applications*, 2nd, ser. Application of Mathematics. Springer, 1998, vol. 38.

[55] S. T. Maguluri, R. Srikant, and L. Ying, "Stochastic models of load balancing and scheduling in cloud computing clusters", in *2012 Proceedings IEEE Infocom*, IEEE, 2012, pp. 702–710.

[56] J. G. Dai, "On positive harris recurrence of multiclass queueing networks: A unified approach via fluid limit models", *Annals of Applied Probability*, pp. 49–77, 1995.

[57] D. Bertsimas, D. Gamarnik, and J. N. Tsitsiklis, "Performance of multiclass markovian queueing networks via piecewise linear lyapunov functions", *Annals of Applied Probability*, pp. 1384–1428, 2001.

[58] R. Atar, P. Dupuis, and A. Shwartz, "An escape-time criterion for queueing networks: Asymptotic risk-sensitive control via differential games", *Mathematics of Operations Research*, vol. 28, no. 4, pp. 801–835, 2003.

[59] D. K. George, "Stochastic modeling and decentralized control policies for large-scale vehicle sharing systems via closed queueing networks", Ph.D. dissertation, The Ohio State University, 2012.

[60] J. A. Parrott and M. Reich, "An earnings standard for new york city's app-based drivers: Economic analysis and policy assessment", 2018.

[61] New York City Taxi and Limousine Commission and Department of Transportation, "Improving efficiency and managing growth in new york's for-hire vehicle sector", 2019.

[62] I. Ashlagi, Y. Kanoria, and J. D. Leshno, "Unbalanced random matching markets: The stark effect of competition", *Journal of Political Economy*, vol. 125, no. 1, pp. 69–98, 2017.

[63] A. E. Roth and E. Peranson, "The redesign of the matching market for american physicians: Some engineering aspects of economic design", *American Economic Review*, vol. 89, no. 4, pp. 748–780, 1999.

[64] A. Hassidim, A. Romm, and R. I. Shorrer, "Redesigning the israeli psychology master's match", *American Economic Review*, vol. 107, no. 5, pp. 205–09, 2017.

[65] I. Rios, T. Larroucau, G. Parra, and R. Cominetti, "Improving the chilean college admissions system", *Oper. Res*, 2019.

[66] A. Abdulkadiroglu, P. A. Pathak, and A. E. Roth, "The New York City high school match", *American Economic Review*, vol. 95, no. 2, pp. 364–367, 2005.

[67] L. E. Dubins and D. A. Freedman, "Machiavelli and the Gale-Shapley algorithm", *The American Mathematical Monthly*, vol. 88, no. 7, pp. 485–494, 1981.

[68] F. Kojima and P. A. Pathak, "Incentives and stability in large two-sided matching markets", *American Economic Review*, vol. 99, no. 3, pp. 608–27, 2009.

[69] A. Hassidim, D. Marciano, A. Romm, and R. I. Shorrer, "The mechanism is truthful, why aren't you?", *American Economic Review*, vol. 107, no. 5, pp. 220–24, 2017.

[70] F. Echenique, R. Gonzalez, A. Wilson, and L. Yariv, "Top of the batch: Interviews and the match", *arXiv preprint arXiv:2002.05323*, 2020.

[71] D. E. Knuth, *Mariages stables et leurs relations avec d'autres problemes combinatoires: introduction a l'analysis mathematique des algorithmes-*. Les Presses de l'Université de Montréal, 1976.

[72] B. Pittel, "The average number of stable matchings", *SIAM J. Disc. Math*, vol. 2, no. 4, pp. 530–549, 1989.

[73] D. E. Knuth, R. Motwani, and B. Pittel, "Stable husbands", *Random Structures & Algorithms*, vol. 1, no. 1, pp. 1–14, 1990.

[74] B. Pittel, "On likely solutions of a stable marriage problem", *The Annals of Applied Probability*, pp. 358–401, 1992.

[75] S. Lee, "Incentive compatibility of large centralized matching markets", *The Review of Economic Studies*, vol. 84, no. 1, pp. 444–463, 2016.

[76] K. Menzel, "Large matching markets as two-sided demand systems", *Econometrica*, vol. 83, no. 3, pp. 897–941, 2015.

[77] I. Ashlagi and A. Nikzad, "What matters in school choice tie-breakings? how competition guides design", in *Proceedings of the 2016 ACM Conference on Economics and Computation*, 2016, pp. 767–768.

[78] B. Pittel, "On likely solutions of the stable matching problem with unequal numbers of men and women", *Mathematics of Operations Research*, vol. 44, no. 1, pp. 122–146, 2019.

[79] N. Immorlica and M. Mahdian, "Marriage, honesty, and stability", in *Proceedings of the sixteenth annual ACM-SIAM symposium on Discrete algorithms*, Society for Industrial and Applied Mathematics, 2005, pp. 53–62.

[80] N. Arnosti, "Short lists in centralized clearinghouses", *EC '15: Proceedings of the Sixteenth ACM Conference on Economics and Computation*, p. 751, 2015.

[81] D. Gale and L. S. Shapley, "College admissions and the stability of marriage", *American Mathematical Monthly*, vol. 69, no. 1, pp. 9–15, 1962.

[82] A. E. Roth and M. A. O. Sotomayor, *Two-sided Matching: A Study in Game-Theoretic Modeling and Analysis*, 18. Cambridge University Press, Cambridge, UK, 1990.

[83] M. David, *Algorithmics of matching under preferences*. World Scientific, 2013, vol. 2.

[84] P. Legros and A. F. Newman, "Beauty is a beast, frog is a prince: Assortative matching with nontransferabilities", *Econometrica*, vol. 75, no. 4, pp. 1073–1102, 2007.

[85] U. Dur, S. D. Kominers, P. A. Pathak, and T. Sönmez, "Reserve design: Unintended consequences and the demise of boston's walk zones", *Journal of Political Economy*, vol. 126, no. 6, pp. 2457–2479, 2018.

[86] A. E. Roth, "A natural experiment in the organization of entry-level labor markets: Regional markets for new physicians and surgeons in the united kingdom", *The American economic review*, pp. 415–440, 1991.

[87] J. H. Kagel and A. E. Roth, "The dynamics of reorganization in matching markets: A laboratory experiment motivated by a natural experiment", *The Quarterly Journal of Economics*, vol. 115, no. 1, pp. 201–235, 2000.

[88] G. J. Hitsch, A. Hortaçsu, and D. Ariely, "Matching and sorting in online dating", *American Economic Review*, vol. 100, no. 1, pp. 130–63, 2010.

[89] L. Cai and C. Thomas, “The short-side advantage in random matching markets”, *arXiv preprint arXiv:1910.04406*, 2019.

[90] A. E. Roth, “On the allocations of residents to rural hospitals: A general property of two-sided matching markets”, *Econometrica*, vol. 54, no. 2, pp. 425–427, 1986.

[91] L. S. Shapley and M. Shubik, “The assignment game i: The core”, *International Journal of game theory*, vol. 1, no. 1, pp. 111–130, 1971.

[92] D. G. McVitie and L. B. Wilson, “The stable marriage problem”, *Communications of the ACM*, vol. 14, no. 7, pp. 486–490, 1971.

[93] M. Baccara, S. Lee, and L. Yariv, “Optimal dynamic matching”, CEPR Discussion Papers, Tech. Rep., 2018.

[94] J. Leshno, “Dynamic matching in overloaded waiting lists”, *Available at SSRN 2967011*, 2017.

[95] E. H. Kaplan, “Managing the demand for public housing”, Ph.D. dissertation, Massachusetts Institute of Technology, 1984.

[96] ——, “A public housing queue with reneging”, *Decision Sciences*, vol. 19, no. 2, pp. 383–391, 1988.

[97] S. A. Zenios, “Modeling the transplant waiting list: A queueing model with reneging”, *Queueing systems*, vol. 31, no. 3-4, pp. 239–251, 1999.

[98] X. Su and S. Zenios, “Patient choice in kidney allocation: The role of the queueing discipline”, *Manufacturing & Service Operations Management*, vol. 6, no. 4, pp. 280–301, 2004.

[99] F. Bloch and D. Cantala, “Dynamic assignment of objects to queuing agents”, *American Economic Journal: Microeconomics*, vol. 9, no. 1, pp. 88–122, 2017.

[100] N. Arnosti and P. Shi, “Design of lotteries and wait-lists for affordable housing allocation”, *Management Science*, 2020.

[101] V. Polterovich, “Rationing, queues, and black markets”, *Econometrica: Journal of the Econometric Society*, pp. 1–28, 1993.

[102] M. Braverman, J. Chen, and S. Kannan, “Optimal provision-after-wait in healthcare”, *Mathematics of Operations Research*, vol. 41, no. 1, pp. 352–376, 2016.

[103] P. Afeche, R. Caldentey, and V. Gupta, “On the optimal design of a bipartite matching queueing system”, *Available at SSRN 3345302*, 2019.

[104] B. Ata, Y. Ding, and S. Zenios, "An achievable-region-based approach for kidney allocation policy design with endogenous patient choice", *Manufacturing & Service Operations Management*, vol. 23, no. 1, pp. 36–54, 2021.

[105] H. Mendelson, "Market behavior in a clearing house", *Econometrica: Journal of the Econometric Society*, pp. 1505–1524, 1982.

[106] F. Kelly and E. Yudovina, "A markov model of a limit order book: Thresholds, recurrence, and trading strategies", *Mathematics of Operations Research*, vol. 43, no. 1, pp. 181–203, 2018.

[107] S. Loertscher, E. V. Muir, and P. G. Taylor, "Optimal market thickness and clearing", *Unpublished paper, Department of Economics, University of Melbourne.[1224]*, 2018.

[108] H. Uzawa, "Walras' tatonnement in the theory of exchange", *The Review of Economic Studies*, vol. 27, no. 3, pp. 182–194, 1960.

[109] R. Cole and L. Fleischer, "Fast-converging tatonnement algorithms for one-time and ongoing market problems", in *Proceedings of the fortieth annual ACM symposium on Theory of computing*, 2008, pp. 315–324.

[110] Y. K. Cheung, R. Cole, and Y. Tao, "Dynamics of distributed updating in fisher markets", in *Proceedings of the 2018 ACM Conference on Economics and Computation*, 2018, pp. 351–368.

[111] Y. K. Cheung, R. Cole, and N. R. Devanur, "Tatonnement beyond gross substitutes? gradient descent to the rescue", *Games and Economic Behavior*, 2019.

[112] W. B. Powell and Y. Sheffi, "The convergence of equilibrium algorithms with predetermined step sizes", *Transportation Science*, vol. 16, no. 1, pp. 45–55, 1982.

[113] J. R. Correa and N. E. Stier-Moses, "Wardrop equilibria", *Wiley encyclopedia of operations research and management science*, 2010.

[114] D. P. Bertsekas, "A distributed algorithm for the assignment problem", *Lab. for Information and Decision Systems Working Paper, MIT*, 1979.

[115] G. Demange, D. Gale, and M. Sotomayor, "Multi-item auctions", *Journal of Political Economy*, vol. 94, no. 4, pp. 863–872, 1986.

[116] R. W. Wolff, "Poisson arrivals see time averages", *Operations Research*, vol. 30, no. 2, pp. 223–231, 1982.

[117] A. Galichon, *Optimal transport methods in economics.* Princeton University Press, 2018.

[118] E. BackProp, Y. LeCun, L. Bottou, G. B. Orr, and K. Muller, *Neural networks: Tricks of the trade*, 1998.

[119] R. Johari, V. Kamble, and Y. Kanoria, "Matching while learning", *arXiv preprint arXiv:1603.04549*, 2016.

[120] D. P. Williamson, *Network Flow Algorithms.* Cambridge University Press, 2019.

[121] N. Buchholz, "Spatial equilibrium, search frictions and efficient regulation in the taxi industry", Technical report, University of Texas at Austin, Tech. Rep., 2015.

[122] A. L. Stolyar, "Large deviations of queues sharing a randomly time-varying server", *Queueing Systems*, vol. 59, no. 1, p. 1, 2008.

[123] H. L. Royden and P. Fitzpatrick, *Real analysis.* Macmillan New York, 1988, vol. 32.

[124] V. G. Subramanian, "Large deviations of max-weight scheduling policies on convex rate regions", *Mathematics of Operations Research*, vol. 35, no. 4, pp. 881–910, 2010.

[125] S. M. Ross, *Stochastic processes. 1996*, 1996.

[126] R. Durrett, *Probability: theory and examples.* Cambridge university press, 2010.

[127] K. L. Chung, *A course in probability theory.* Academic press, 2001.

[128] A. Shwartz and A. Weiss, *Large deviations for performance analysis: queues, communication and computing.* CRC Press, 1995, vol. 5.

[129] D. Dubhashi and D. Ranjan, "Balls and bins: A study in negative dependence", *Random Structures & Algorithms*, vol. 13, no. 2, pp. 99–124, 1998.

[130] K. Joan-Dev and F. Proschan, "Negative association of random variables with applications", *The Annals of Statistics*, vol. 11, no. 1, pp. 286–295, 1983.

[131] D. E. Knuth, "Stable marriage and its relation to other combinatorial problems", in *An introduction to the mathematical analysis of algorithms*, vol. 10, American Mathematical Society, 1996.

[132] I. Ashlagi, Y. Kanoria, and J. D. Leshno, "Unbalanced random matching markets: The stark effect of competition", *Journal of Political Economy*, vol. 25, no. 1, pp. 65–98, 2015.

[133] B. Pittel, "On likely solutions of the stable matching problem with unequal numbers of men and women", *Mathematics of Operations Research*, vol. 44, no. 1, pp. 122–146, 2019.

[134] W Kozakiewicz *et al.*, "On the convergence of sequences of moment generating functions", *The Annals of Mathematical Statistics*, vol. 18, no. 1, pp. 61–69, 1947.

[135] A. Postnikov, "Permutohedra, associahedra, and beyond", *International Mathematics Research Notices*, vol. 2009, no. 6, pp. 1026–1106, 2009.

[136] J. Gallier, "Notes on convex sets, polytopes, polyhedra, combinatorial topology, voronoi diagrams and delaunay triangulations", *arXiv preprint arXiv:0805.0292*, 2008.

APPENDIX A

Proofs in "Blind Dynamic Resource Allocation in Closed Networks via Mirror Backpressure"

Organization of the appendix. In this paper, we proved performance guarantees for three settings: entry control (Theorem 1.1), joint entry-assignment control (Theorem 1.2) and joint pricing-assignment control (Theorem 1.3). In the appendix, we will only prove the results for JEA and JPA since entry control is a special case of JEA. For most parts of the proof, the proof of JEA can be easily extended to JPA. For particular lemmas/propositions, the proofs of the JPA setting are more involved. For easier reading, we put analogous results together.

The appendix is organized as follows.

1. In Appendix A.1 we prove Proposition 1.1, i.e., that the value of SPP is an upper bound of the best achievable per customer payoff. We will prove the counterpart of Proposition 1.1 for JEA (Proposition A.1) and JPA (Proposition A.2) settings.

2. In Appendix A.2, we perform the Lyapunov analysis and analyze the geometry of the dual problem (1.14), and prove Lemma 1.1 and Lemma 1.2. We will prove the counterpart of these lemmas for JEA and JPA settings.

3. In Appendix A.3 we prove Lemma 1.3. We also prove a general result (Theorem A.1), and show that it implies Theorems 1.1, 1.2, and 1.3.

4. In Appendix A.4, we provide further details of the simulation setting.

A.1 Finite Horizon Payoff Upper Bound: Proof of Proposition 1.1

In this section, we prove the finite horizon payoff upper bounds for JEA (Proposition A.1) and JPA setting (Proposition A.2). Proposition 1.1 is implied by Proposition A.1.

A.1.1 Joint Entry-Assignment Setting

Consider the JEA setting defined in Section 1.6.2, which allows for flexible assignment and time-varying demand.

We will state and prove a generalization of Proposition 1.1 (fluid-based upper bound on the payoff) to the JEA setting. Before that we introduce some linear programs which generalize the static planning problem (1.10)-(1.12), and establish a lemma relating their values to each other.

Relevant linear programs

Fix a horizon T. We will consider the following linear program at time t, based on the current demand arrival rates $\boldsymbol{\phi}^t$:

$$\text{SPP}^t: \quad \text{maximize}_{\mathbf{z}} \sum_{\tau\in\mathcal{T},j\in\mathcal{P}(\tau),k\in\mathcal{D}(\tau)} w_{j\tau k}\, z_{j\tau k} \tag{A.1}$$

$$\text{s.t.} \sum_{\tau\in\mathcal{T},j\in\mathcal{P}(\tau),k\in\mathcal{D}(\tau)} z_{j\tau k}(\mathbf{e}_j - \mathbf{e}_k) = \mathbf{0} \quad \text{(flow balance)}, \tag{A.2}$$

$$\sum_{j\in\mathcal{P}(\tau),k\in\mathcal{D}(\tau)} z_{j\tau k} \le \phi^t_\tau\,, \ z_{j\tau k} \ge 0 \quad \forall j,k\in V,\ \tau\in\mathcal{T}\,. \ \text{(demand constraint)}\,. \tag{A.3}$$

The variable $z_{j\tau k}$ can be interpreted as the flow of demand type τ being served by pickup location j and dropoff location k. (Note that our LP formulation here has a cosmetic difference from that in (1.10)-(1.12): here we find that it simplifies our analysis to use the *flows* $z_{j\tau k}$ as the LP variables instead of using the *fractions* $x_{j\tau k}$ of demand of type τ served by pickup location j and dropoff location k as the variables. The correspondence is simply $z_{j\tau k} \leftrightarrow \phi^t_\tau x_{j\tau k}$.) We denote the value of SPP^t by W^{SPP^t}.

Define the average demand arrival rates

$$\overline{\boldsymbol{\phi}} \triangleq \frac{1}{T}\sum_{t=0}^{T-1} \boldsymbol{\phi}^t . \tag{A.4}$$

We define an "average" linear program $\overline{\mathrm{SPP}}$ as the linear program given by (A.1), (A.2), and the averaged demand constraint

$$\sum_{j\in\mathcal{P}(\tau),k\in\mathcal{D}(\tau)} z_{j\tau k} \le \bar{\phi}_\tau\,,\ z_{j\tau k} \ge 0 \quad \forall j,k \in V,\ \tau\in\mathcal{T}\,.\ \text{(demand constraint)}\,. \tag{A.5}$$

We denote the value of $\overline{\mathrm{SPP}}$ by $W^{\overline{\mathrm{SPP}}}$.

Although we will not use this property, note that $W^{\overline{\mathrm{SPP}}} \ge \frac{1}{T}\sum_{t=0}^{T-1} W^{\mathrm{SPP}^t}$ since if $\mathbf{z}^t$ is feasible for SPP^t for each t then $\bar{\mathbf{z}} = \frac{1}{T}\sum_{t=0}^{T-1}\mathbf{z}^t$ is feasible for $\overline{\mathrm{SPP}}$. Rather, we will prove and then leverage the property that $W^{\overline{\mathrm{SPP}}}$ is not much larger than W^{SPP^t} for any $t \le T-1$ if the demand arrival rates vary slowly with t.

Lemma A.1. *Suppose the demand arrival rates vary η-slowly (Definition 1.1) for some $\eta > 0$. Fix a horizon T. For any $0 \le t \le T-1$ we have*

$$W^{\mathrm{SPP}^t} \ge W^{\overline{\mathrm{SPP}}} - \eta T m/2\,. \tag{A.6}$$

Proof. Since $\|\boldsymbol{\phi}^{t'+1} - \boldsymbol{\phi}^{t'}\|_1 \le \eta$ for all t', we know that

$$\|\boldsymbol{\phi}^t - \overline{\boldsymbol{\phi}}\|_1 \le \eta T/2\,. \tag{A.7}$$

Let $\bar{\mathbf{z}}$ be an optimal solution to $\overline{\mathrm{SPP}}$. If $\bar{\mathbf{z}}$ is feasible for SPP^t we are done. Suppose not. Using the standard flow decomposition approach [see, e.g., 120, the interested reader can also find the flow decomposition argument in the proof of Lemma A.3 below], the flow $\bar{\mathbf{z}}$ can be decomposed into flows along directed cycles, since it satisfies the flow balance constraints (A.2): directed cycles $\mathcal{C}$ carrying flow $f_{\mathcal{C}} > 0$ in the decomposition take the form $\mathcal{C} = \big((j_1,\tau_1,j_2),(j_2,\tau_2,j_3),\cdots,(j_s,\tau_s,j_{s+1}=j_1)\big)$ where the nodes $j_1, j_2, \ldots, j_s$ are distinct from each other, and for each $r = 1, 2, \ldots, s$, there is a flow from j_r to j_{r+1} due to demand type τ_r. We have

$$\bar{z}_{j\tau k} = \sum_{\mathcal{C} \ni (j,\tau,k)} f_{\mathcal{C}} \qquad \text{for all } \tau \in \mathcal{T}, j \in \mathcal{P}(\tau), k \in \mathcal{D}(\tau)\,. \tag{A.8}$$

(The number of cycles in the decomposition is bounded above by $\sum_{\tau \in \mathcal{T}} |\mathcal{P}(\tau)||\mathcal{D}(\tau)|$, but our argument will not be affected by the number of cycles. In fact our argument can handle an infinity of demand types by replacing sums with integrals.)

Starting from the flow $\bar{\mathbf{z}}$ and the associated cycle decomposition (A.8), we reduce the flows $(f_{\mathcal{C}})$ along the cycles via the following iterative process, in order to obtain $\mathbf{z}^t$ which is feasible for the problem SPP^t:

Consider each demand type $\tau \in \mathcal{T}$ in turn and do the following. Define the (current) arrival rate violation as

$$\delta_\tau \triangleq \left(\sum_{\mathcal{C}} f_{\mathcal{C}} \cdot \text{count}(\mathcal{C}, \tau) \; - \; \phi_\tau^t \right)_+ \,.$$

where $\text{count}(\mathcal{C}, \tau)$ is the number of times demand type τ appears in cycle $\mathcal{C}$. If $\delta_\tau = 0$ do nothing. If $\delta_\tau > 0$, reduce the flows in cycles containing τ sufficiently that after the reduction $\sum_{\mathcal{C}} f_{\mathcal{C}} \cdot \text{count}(\mathcal{C}, \tau) = \phi_\tau^t$ holds (the reduction can be divided arbitrarily between the different cycles containing τ; subject to the constraints that no cycle-flow should increase and no cycle-flow should go below zero). Note that the payoff loss resulting from this reduction is bounded above by $\delta_\tau m$ since each cycle length is at most m (since no node is repeated in a cycle), the ws are assumed to be bounded by 1, and the total reduction in cycle flows is at most δ_τ.

This simple process maintains the following properties:

- The flow balance constraint (A.2) is satisfied throughout.
- Cycle-flows are non-increasing during the process. Cycle-flows never drop below zero.
- For all demand types which have already been processed so far, the arrival rate

constraint is satisfied. Formally: During the process, denote the current value of the right-hand side of (A.8) by $z_{j\tau k}$. Then $\sum_{j\in\mathcal{P}(\tau),k\in\mathcal{D}(\tau)} z_{j\tau k} = \sum_{\mathcal{C}} f_{\mathcal{C}}\cdot\text{count}(\mathcal{C},\tau) \leq \phi_\tau^t$ for all demand types τ which have already been processed.

In particular, at the end of the process, we arrive at flows $\mathbf{z}^t$ which are feasible for SPP^t. It remains to show that the payoff lost due to the reduction in flows is bounded by $\eta T m/2$.

Since flows are non-increasing and the initial flows are feasible for $\overline{\text{SPP}}$, we have that $\delta_\tau \leq \left(\bar{\phi}_\tau - \phi_\tau^t\right)_+$ for all $\tau \in \mathcal{T}$. Since the payoff lost while processing demand type τ is bounded above by $\delta_\tau m$ (as we argued above), the total loss in payoff lost is then bounded above by

$$m\sum_{\tau\in\mathcal{T}}\delta_\tau \leq m\sum_{\tau\in\mathcal{T}}\left(\bar{\phi}_\tau - \phi_\tau^t\right)_+ \leq m\|\boldsymbol{\phi}^t - \overline{\boldsymbol{\phi}}\|_1 \leq m\eta T/2\,,$$

where we used (A.7) in the last inequality. Thus, we have constructed a feasible solution $\mathbf{z}^t$ to SPP^t which achieves payoff at least $W^{\overline{\text{SPP}}} - \eta T m/2$. The lemma follows. □

Upper bound on the payoff

We state below the generalization of Proposition 1.1 to the JEA setting with time-varying demand arrival rates.

Proposition A.1. *For any horizon $T < \infty$, any K and any starting state $\mathbf{q}[0]$, the best achievable finite horizon average payoff W_T^* in the JEA setting is upper bounded as*

$$W_T^* \leq W^{\overline{\text{SPP}}} + m\cdot\frac{K}{T}\,.$$

Here $W^{\overline{\text{SPP}}}$ is the optimal value of $\overline{\text{SPP}}$ given by (A.1), (A.2) *and* (A.5).

The idea behind Proposition A.1 is as follows. As is typical in such settings, $W^{\overline{\text{SPP}}}$ is an upper bound on the payoff if the flow constraints are satisfied in expectation. However, since the flow constraints can be slightly violated in the finite horizon setting under consideration, we obtain an upper bound by slightly relaxing the flow constraint (A.2) in the SPP to

$$\left|\mathbf{1}_S^{\mathrm{T}}\left(\sum_{\tau\in\mathcal{T},j\in\mathcal{P}(\tau),k\in\mathcal{D}(\tau)} z_{j\tau k}(\mathbf{e}_j-\mathbf{e}_k)\right)\right|\le\frac{K}{T}\qquad\forall\, S\subseteq V\,,\tag{A.9}$$

where $\mathbf{1}_S$ is the vector with 1s at nodes in S and 0s at all other nodes.

We establish two key lemmas to facilitate the proof of Proposition A.1. The first lemma (Lemma A.2) shows that the expected payoff cannot exceed the value of the finite horizon demand-averaged SPP, i.e., the linear program defined by (A.1), (A.9) and (A.5).

Lemma A.2. *For any horizon $T<\infty$, any K and any starting state $\mathbf{q}[0]$, the expected payoff generated by any feasible joint entry-assignment control policy π is upper bounded by the value of the linear program defined by* (A.1)*, the approximate flow balance constraints* (A.9) *and time-averaged demand constraints* (A.5)*.*

Proof. Let π be any feasible policy. For each $\tau\in\mathcal{T}$ and $j\in\mathcal{P}(\tau),k\in\mathcal{D}(\tau)$, define

$$\bar{z}_{j\tau k}\triangleq\frac{1}{T}\sum_{t=0}^{T-1}\mathbb{E}[x_{j\tau k}[t]\mathbb{I}\{\tau[t]=\tau\}]\,.$$

In words, $\bar{z}_{j\tau k}$ is the average flow over $1\le t\le T$ of the demand type τ being served by pickup location j and dropoff location k. Since for each t, $z_{j\tau k}[t]\triangleq\mathbb{E}[\mathbb{I}\{\tau[t]=\tau\}x_{j\tau k}]$ satisfies the period-specific demand constraint (A.3) for all $\tau\in\mathcal{T},j\in\mathcal{P}[\tau],k\in\mathcal{D}(\tau)$, the averaged constraints (A.5) must hold for $\bar{\mathbf{z}}$.

We can write the expected per-period payoff collected in the first T periods as:

$$\begin{aligned}W_T^{\pi}=&\ \frac{1}{T}\sum_{t=0}^{T-1}\mathbb{E}\left[\sum_{\tau\in\mathcal{T},j\in\mathcal{P}(\tau),k\in\mathcal{D}(\tau)} w_{j\tau k}\cdot x_{j\tau k}[t]\mathbb{I}\{\tau[t]=\tau\}\right]\\ =&\sum_{\tau\in\mathcal{T},j\in\mathcal{P}(\tau),k\in\mathcal{D}(\tau)} w_{j\tau k}\cdot\bar{z}_{j\tau k}\,,\end{aligned}$$

where we only used linearity of expectation. In words, the expected per-period payoff is the objective (A.1) evaluated at $\bar{\mathbf{z}}$. Similarly, for the time-average of the change of queue length we have:

$$\frac{1}{T}\cdot\mathbb{E}[\mathbf{q}[T]-\mathbf{q}[0]]=\sum_{\tau\in\mathcal{T},j\in\mathcal{P}(\tau),k\in\mathcal{D}(\tau)}\bar{z}_{j\tau k}\cdot(\mathbf{e}_j-\mathbf{e}_k)\,,$$

which implies that $\bar{\mathbf{z}}$ satisfies the approximate flow constraints (A.9) since $|\sum_{j\in S} q_j[T] - q_j[0]| \leq K$ for all $S \subset V$. (Because there are only K resources circulating in the system, the net outflow from any subset of nodes $S \subseteq V$ should not exceed K in magnitude.)

We have shown that $\bar{\mathbf{z}}$ is feasible for the given linear program with constraints (A.9) and (A.5), and the expected payoff earned W_T^π is identical to objective (A.1) evaluated at $\bar{\mathbf{z}}$. It follows that W_T^π is upper bounded by the value of the optimization problem defined by (A.1), (A.9) and (A.5) regardless of the initial configuration $\mathbf{q}[0]$. This concludes the proof. □

In order to facilitate the second key lemma, we first prove a supporting lemma (Lemma A.3). We call $\mathbf{z}$ a (directed) *acyclic flow* if there is no (directed) cycle

$$\mathcal{C} = \big((j_1, \tau_1, j_2), (j_2, \tau_2, j_3), \cdots, (j_s, \tau_s, j_{s+1} = j_1) \big), \qquad \text{where } j_r \in V \text{ and } \tau_r \in \mathcal{T} \text{ for } r = 1, 2, \cdots, s,$$

such that

$$z_{j_r, \tau_r, j_{r+1}} > 0 \qquad \text{for all } r = 1, \cdots, s.$$

In words, there is no cycle $\mathcal{C}$ such that there is a positive flow along $\mathcal{C}$.

Lemma A.3. *Any feasible solution* $\mathbf{z}^F$ *of the finite horizon averaged SPP satisfying approximate flow balance* (A.9) *and the average demand constraint* (A.5) *can be decomposed as*

$$\mathbf{z}^{\mathrm{F}} = \mathbf{z}^{\mathrm{S}} + \mathbf{z}^{\mathrm{DAG}}, \tag{A.10}$$

where $\mathbf{z}^{\mathrm{S}}$ *is a feasible solution for the SPP satisfying exact flow balance* (A.2) *and* (A.5), *and* $\mathbf{z}^{\mathrm{DAG}}$ *is an acyclic flow satisfying* (A.9) *and* (A.5).

Proof. The existence of such a decomposition can be established using a standard flow decomposition argument [see, e.g., 120]: Start with $\mathbf{z}^{\mathrm{S}} = \mathbf{0}$ and $\mathbf{z}^{\mathrm{DAG}} = \mathbf{z}^{\mathrm{F}}$. Then, iteratively, if $\mathbf{z}^{\mathrm{DAG}}$ includes a cycle $\mathcal{C}$ with a positive flow along $\mathcal{C}$ as above, move a flow of $u(\mathcal{C}) \triangleq \min_{1\le r\le s} z_{j_r,\tau_r,j_{r+1}}$ along $\mathcal{C}$ from $\mathbf{z}^{\mathrm{DAG}}$ to $\mathbf{z}^{\mathrm{S}}$, via the updates

$$z^{\mathrm{S}}_{j_r,\tau_r,j_{r+1}} \leftarrow z^{\mathrm{S}}_{j_r,\tau_r,j_{r+1}} + u(\mathcal{C})\,, \qquad\qquad z^{\mathrm{DAG}}_{j_r,\tau_r,j_{r+1}} \leftarrow z^{\mathrm{DAG}}_{j_r,\tau_r,j_{r+1}} - u(\mathcal{C})\,,$$

for all $r = 1, 2, \ldots, s$. This iterative process maintains the following invariants which hold at the end of each iteration:

- $\mathbf{z}^{\mathrm{S}}$ remains feasible for the SPP, in particular, it satisfies flow balance (A.2).
- $\mathbf{z}^{\mathrm{F}} = \mathbf{z}^{\mathrm{S}} + \mathbf{z}^{\mathrm{DAG}}$ remains true.
- It remains true that

$$\sum_{\tau\in\mathcal{T}, j\in\mathcal{P}(\tau), k\in\mathcal{D}(\tau)} z^{\mathrm{DAG}}_{j\tau k}(\mathbf{e}_j - \mathbf{e}_k) = \sum_{\tau\in\mathcal{T}, j\in\mathcal{P}(\tau), k\in\mathcal{D}(\tau)} z^{\mathrm{F}}_{j\tau k}(\mathbf{e}_j - \mathbf{e}_k)\,.$$

  i.e., $\mathbf{z}^{\mathrm{DAG}}$ has the same net inflow/outflow from each supply node as $\mathbf{z}^{\mathrm{F}}$. In particular, $\mathbf{z}^{\mathrm{DAG}}$ satisfies approximate flow balance (A.9).

Moreover, the iterative process progresses monotonically: Observe that $\mathbf{z}^{\mathrm{S}}$ coordinate-wise (weakly) increases monotonically, whereas $\mathbf{z}^{\mathrm{DAG}}$ coordinate-wise (weakly) decreases monotonically (but preserves $\mathbf{z}^{\mathrm{DAG}} \ge \mathbf{0}$). Since we also know that $\mathbf{z}^{\mathrm{S}}$ is bounded, it follows that this iterative process converges. Moreover, in the limit it must be that there is no remaining cycle with positive flow in $\mathbf{z}^{\mathrm{DAG}}$ (else we observe a contradiction with the fact that the process has converged). Hence, $\mathbf{z}^{\mathrm{S}}$ and $\mathbf{z}^{\mathrm{DAG}}$ at the end of the process provide the claimed decomposition. □

Using this supporting lemma, we now establish the second key lemma which shows that the value of the averaged SPP with approximate flow balance constraints (A.9) cannot be much larger than the value of the averaged program $\overline{\mathrm{SPP}}$ which imposes exact flow balance constraints (A.2).

Lemma A.4. *The value of the linear program defined by* (A.1)*, the approximate flow balance constraints* (A.9) *and time-averaged demand constraints* (A.5) *is bounded above by*

$$W^{\overline{\mathrm{SPP}}} + m \cdot \frac{K}{T}\,.$$

where $W^{\overline{\mathrm{SPP}}}$ *is the value of the linear program* $\overline{\mathrm{SPP}}$ *which imposes exact flow balance constraints* (A.2)*.*

Proof. We appeal to the decomposition from Lemma A.3 to decompose any feasible solution $\mathbf{z}^{\mathrm{F}}$ to the finite horizon fluid problem as

$$\mathbf{z}^{\mathrm{F}} = \mathbf{z}^{\mathrm{S}} + \mathbf{z}^{\mathrm{DAG}}\,,$$

where $\mathbf{z}^{\mathrm{S}}$ is feasible for $\overline{\mathrm{SPP}}$ and $\mathbf{z}^{\mathrm{DAG}}$ is a directed acyclic flow that satisfies approximate flow balance (A.9) and the averaged demand constraints (A.3). Hence, the objective (A.1) can be written as the sum of two terms

$$\sum_{\tau\in\mathcal{T},j\in\mathcal{P}(\tau),k\in\mathcal{D}(\tau)} w_{j\tau k}\cdot z^{\mathrm{F}}_{j\tau k} = \sum_{\tau\in\mathcal{T},j\in\mathcal{P}(\tau),k\in\mathcal{D}(\tau)} w_{j\tau k}\cdot (z^{\mathrm{S}}_{j\tau k} + z^{\mathrm{DAG}}_{j\tau k})\,, \tag{A.11}$$

and each of the terms can be bounded from above. By definition of $W^{\overline{\mathrm{SPP}}}$ we know that

$$\sum_{\tau\in\mathcal{T},j\in\mathcal{P}(\tau),k\in\mathcal{D}(\tau)} w_{j\tau k}\cdot z^{\mathrm{S}}_{j\tau k} \le W^{\overline{\mathrm{SPP}}}\,.$$

We will now show that

$$\sum_{\tau\in\mathcal{T},j\in\mathcal{P}(\tau),k\in\mathcal{D}(\tau)} w_{j\tau k}\cdot z^{\mathrm{DAG}}_{j\tau k} \le (m-1)\cdot\frac{K}{T} < m\cdot\frac{K}{T}\,.$$

The lemma will follow, since this will imply an upper bound of $W^{\overline{\mathrm{SPP}}} + m\cdot\frac{K}{T}$ on the objective for any $\mathbf{z}^{\mathrm{F}}$ satisfying (A.9) and (A.5).

Consider $\mathbf{z}^{\mathrm{DAG}}$. Since it is an acyclic flow, there is an ordering $(j_1, j_2, \ldots, j_m)$ of the nodes in V such that all positive flows move supply from an earlier node to a later node in this ordering. More precisely, it holds that for any $\tau\in\mathcal{T}$,

$$z^{\mathrm{DAG}}_{j_l,\tau,j_r} = 0 \qquad \forall\, l > r \;\text{ s.t. }\; j_l\in\mathcal{P}(\tau), j_r\in\mathcal{D}(\tau)\,. \tag{A.12}$$

Now consider the subsets $A_\ell \triangleq \{j_1, j_2, \ldots, j_\ell\}\subset V$ for $\ell = 1, 2, \ldots, m-1$. Note that from (A.12), $\mathbf{z}^{\mathrm{DAG}}$ does not move any supply from $V\backslash A_\ell$ to A_ℓ. Hence we have

$$\mathbf{1}^{\mathrm{T}}_{A_\ell}\left(\sum_{\tau\in\mathcal{T},j\in\mathcal{P}(\tau),k\in\mathcal{D}(\tau)} z^{\mathrm{DAG}}_{j\tau k}(\mathbf{e}_j - \mathbf{e}_k)\right) = \sum_{\tau\in\mathcal{T},j\in\mathcal{P}(\tau)\cap A_\ell,k\in\mathcal{D}(\tau)\cap(V\backslash A_\ell)} z^{\mathrm{DAG}}_{j\tau k}$$

$$\leq \frac{K}{T} \qquad \forall\, l = 1, 2, \ldots, m-1\,, \tag{A.13}$$

We made use of (A.9) to obtain the upper bound. Further, note that for each $z^{\mathrm{DAG}}_{j_l,\tau,j_r}$ with $l < r$, the term $z^{\mathrm{DAG}}_{j_l,\tau,j_r}$ is part of the above sum for $\ell = l$. Motivated by this observation, we bound the expected payoff of $\mathbf{z}^{\mathrm{DAG}}$ by first using our assumption $\max_{j,k\in V,\tau\in\mathcal{T}} |w_{j\tau k}| \leq 1$ to bound the payoff by the sum of z^{DAG}s (the first inequality below), and then bounding the sum of z^{DAG}s by "allocating" $z^{\mathrm{DAG}}_{j_l,\tau,j_r}$ to the left-hand side of (A.13) with $\ell = l$ and summing over ℓ (the second inequality below):

$$\begin{aligned}
&\sum_{\tau\in\mathcal{T},j\in\mathcal{P}(\tau),k\in\mathcal{D}(\tau)} w_{j\tau k}\cdot z^{\mathrm{DAG}}_{j\tau k}\\
\leq &\sum_{\tau\in\mathcal{T},j\in\mathcal{P}(\tau),k\in\mathcal{D}(\tau)} z^{\mathrm{DAG}}_{j\tau k}\\
\leq &\sum_{1\leq\ell<m}\ \sum_{\tau\in\mathcal{T},j\in\mathcal{P}(\tau)\cap A_\ell,k\in\mathcal{D}(\tau)\cap(V\backslash A_\ell)} z^{\mathrm{DAG}}_{j\tau k}\\
\leq &\,(m-1)\cdot\frac{K}{T}\,.
\end{aligned}$$

The last inequality uses (A.13) summed over ℓ. This completes the proof. □

Proof of Proposition A.1. The proposition follows immediately from Lemmas A.2 and A.4. □

A.1.2 Joint Pricing-Assignment Setting

Consider the JPA setting defined in Section 1.6.3. Recall that we assumed stationary demand arrival rates (in contrast to the JEA setting). The static planning problem (SPP) in the JPA setting is

$$\text{maximize}_{\mathbf{x}} \sum_{\tau\in\mathcal{T}} \phi_\tau \left(r_\tau \left(\sum_{j\in\mathcal{P}(\tau),k\in\mathcal{D}(\tau)} x_{j\tau k} \right) - \sum_{j\in\mathcal{P}(\tau),k\in\mathcal{D}(\tau)} c_{j\tau k}\cdot x_{j\tau k} \right) \tag{A.14}$$

$$\text{s.t.} \sum_{\tau\in\mathcal{T}} \phi_\tau \sum_{j\in\mathcal{P}(\tau),k\in\mathcal{D}(\tau)} x_{j\tau k}(\mathbf{e}_j - \mathbf{e}_k) = \mathbf{0} \qquad \text{(flow balance)}$$

(A.15)

$$\sum_{j\in\mathcal{P}(\tau),k\in\mathcal{D}(\tau)} x_{j\tau k} \le 1\,,\ x_{j\tau k} \ge 0 \quad \forall j,k \in V\,,\ \tau\in\mathcal{T} \qquad \text{(demand constraint)}\,. \tag{A.16}$$

Proposition A.2. *For any horizon $T < \infty$, any K and any starting state $\mathbf{q}[0]$, the finite and infinite horizon average payoff W_T^* and W^* in the JPA setting are upper bounded as*

$$W_T^* \le W^{\text{SPP}} + m\cdot\frac{K}{T}\,, \qquad W^* \le W^{\text{SPP}}\,.$$

Here W^{SPP} is the optimal value of SPP (A.14)-(A.16)*.*

The main twist of the proof comparing to Proposition A.1 is that the objective function in (A.14) is no longer linear. We first prove a JPA version of Lemma A.2.

Lemma A.5. *For any horizon $T < \infty$, any K and any starting state $\mathbf{q}[0]$, the expected payoff generated by any JPA policy π is upper bounded by the value of the finite horizon SPP:*

$$\begin{aligned}
\text{maximize}_{\mathbf{x}} \ & \sum_{\tau\in\mathcal{T}} \phi_\tau \cdot \left(r_\tau\left(\sum_{j\in\mathcal{P}(\tau),k\in\mathcal{D}(\tau)} x_{j\tau k}\right) - \sum_{j\in\mathcal{P}(\tau),k\in\mathcal{D}(\tau)} c_{j\tau k}\cdot x_{j\tau k}\right)\\
\text{s.t. } & \mathbb{1}_S^{\mathrm{T}}\left(\sum_{\tau\in\mathcal{T}}\phi_\tau \sum_{j\in\mathcal{P}(\tau),k\in\mathcal{D}(\tau)} x_{j\tau k}(\mathbf{e}_j - \mathbf{e}_k)\right) \le \frac{K}{T} \qquad \forall S\subseteq V\\
& \sum_{j\in\mathcal{P}(\tau),k\in\mathcal{D}(\tau)} x_{j\tau k} \le 1\,,\ x_{j\tau k}\ge 0 \qquad \forall j,k\in V\,,\ \tau\in\mathcal{T}\,.
\end{aligned}$$

Proof. Let π be any feasible JPA policy. For each demand type $\tau\in\mathcal{T}$ and $j\in\mathcal{P}(\tau)$, $k\in\mathcal{D}(\tau)$, define

$$\bar{x}_{j\tau k} \triangleq \frac{1}{T}\sum_{t=0}^{T-1}\mathbb{E}[\bar{F}_\tau(p_\tau[t])\cdot x_{j\tau k}[t]|\tau[t]=\tau]\,.$$

In words, $\bar{x}_{j\tau k}$ is the average rate over the first T periods of picking up type τ demands from node j and dropping them off at node k.

Let $U_\tau[t]$ be the willingness-to-pay of a type τ demand arriving at time t. We decompose the time-average of payoff collected in the first T periods as:

$$W_T^\pi = \frac{1}{T}\sum_{t=0}^{T-1}\mathbb{E}\left[\sum_{\tau\in\mathcal{T}}\mathbb{1}\{\tau[t]=\tau, U_\tau[t]\ge p_\tau[t]\}\sum_{j\in\mathcal{P}(\tau),k\in\mathcal{D}(\tau)}(p_\tau[t]-c_{j\tau k})\cdot x_{j\tau k}[t]\right]$$

$$= \frac{1}{T}\sum_{t=0}^{T-1}\sum_{\tau\in\mathcal{T}}\phi_\tau\cdot\mathbb{E}\left[\mathbb{1}\{U_\tau[t]\geq p_\tau[t]\}\sum_{j\in\mathcal{P}(\tau),k\in\mathcal{D}(\tau)}(p_\tau[t]-c_{j\tau k})\cdot x_{j\tau k}[t]\,\middle|\,\tau[t]=\tau\right].$$

Because $U_\tau[t]$ is independent of $p_\tau[t]$ and $x_{j\tau k}[t]$, we have

$$W_T^\pi = \frac{1}{T}\sum_{t=0}^{T-1}\sum_{\tau\in\mathcal{T}}\phi_\tau\cdot\mathbb{E}\left[\bar{F}_\tau(p_\tau[t])\sum_{j\in\mathcal{P}(\tau),k\in\mathcal{D}(\tau)}(p_\tau[t]-c_{j\tau k})\cdot x_{j\tau k}[t]\,\middle|\,\tau[t]=\tau\right].$$

Let $\mu_\tau[t]\triangleq\bar{F}_\tau(p_\tau[t])$, and let $\hat{x}_{j\tau k}[t]\triangleq\mu_\tau[t]\cdot x_{j\tau k}[t]$, we have

$$\begin{aligned}
W_T^\pi &= \frac{1}{T}\sum_{t=0}^{T-1}\sum_{\tau\in\mathcal{T}}\phi_\tau\cdot\mathbb{E}\left[\left(\sum_{j\in\mathcal{P}(\tau),k\in\mathcal{D}(\tau)}\hat{x}_{j\tau k}[t]\right)\cdot\bar{F}_\tau^{-1}(\mu_\tau[t])-\sum_{j\in\mathcal{P}(\tau),k\in\mathcal{D}(\tau)}c_{j\tau k}\cdot\hat{x}_{j\tau k}[t]\,\middle|\,\tau[t]=\tau\right]\\
&\leq \frac{1}{T}\sum_{t=0}^{T-1}\sum_{\tau\in\mathcal{T}}\phi_\tau\cdot\mathbb{E}\Bigg[\left(\sum_{j\in\mathcal{P}(\tau),k\in\mathcal{D}(\tau)}\hat{x}_{j\tau k}[t]\right)\cdot\bar{F}_\tau^{-1}\left(\sum_{j\in\mathcal{P}(\tau),k\in\mathcal{D}(\tau)}\hat{x}_{j\tau k}[t]\right)\\
&\quad-\sum_{j\in\mathcal{P}(\tau),k\in\mathcal{D}(\tau)}c_{j\tau k}\cdot\hat{x}_{j\tau k}[t]\,\Big|\,\tau[t]=\tau\Bigg]\\
&= \frac{1}{T}\sum_{t=0}^{T-1}\sum_{\tau\in\mathcal{T}}\phi_\tau\cdot\mathbb{E}\left[r_\tau\left(\sum_{j\in\mathcal{P}(\tau),k\in\mathcal{D}(\tau)}\hat{x}_{j\tau k}[t]\right)-\sum_{j\in\mathcal{P}(\tau),k\in\mathcal{D}(\tau)}c_{j\tau k}\cdot\hat{x}_{j\tau k}[t]\,\middle|\,\tau[t]=\tau\right].
\end{aligned}$$

Here the first inequality follows from the fact that $\bar{F}_\tau^{-1}(\cdot)$ is non-increasing, the last equality uses the definition of revenue function $r_\tau(\cdot)$. Linearity of conditional expectation and conditional Jensen's inequality yields:

$$W_T^\pi\leq\sum_{\tau\in\mathcal{T}}\phi_\tau\cdot\left(-\sum_{j\in\mathcal{P}(\tau),k\in\mathcal{D}(\tau)}c_{j\tau k}\cdot\bar{x}_{j\tau k}+\frac{1}{T}\sum_{t=0}^{T-1}r_\tau\left(\mathbb{E}\left[\sum_{j\in\mathcal{P}(\tau),k\in\mathcal{D}(\tau)}\hat{x}_{j\tau k}[t]\,\middle|\,\tau[t]=\tau\right]\right)\right).$$

Use Jensen's inequality again, we have

$$W_T^\pi\leq\sum_{\tau\in\mathcal{T}}\phi_\tau\cdot\left(-\sum_{j\in\mathcal{P}(\tau),k\in\mathcal{D}(\tau)}c_{j\tau k}\cdot\bar{x}_{j\tau k}+r_\tau\left(\sum_{j\in\mathcal{P}(\tau),k\in\mathcal{D}(\tau)}\bar{x}_{j\tau k}\right)\right).$$

For the time-average of the change of queue length we have:

$$\frac{1}{T}\cdot\mathbb{E}[\mathbf{q}[T]-\mathbf{q}[0]]=\sum_{\tau\in\mathcal{T}}\phi_\tau\sum_{j\in\mathcal{P}(\tau),k\in\mathcal{D}(\tau)}\bar{x}_{j\tau k}\cdot(\mathbf{e}_j-\mathbf{e}_k).$$

Because there are only K resources in the system, the net outflow from any subset of nodes should not exceed K. Note that $\bar{\mathbf{x}}$ must satisfy constraint (A.16). Optimizing over $\bar{\mathbf{x}}$ yields the desired result. □

Proof Sketch of Proposition A.2. The rest of the proof proceeds almost exactly the same as in Proposition A.1. The only caveat is that the equation (A.11) should be replaced by inequality

$$\begin{aligned}
&\sum_{\tau\in\mathcal{T}}\phi_\tau\left(-\sum_{j\in\mathcal{P}(\tau),k\in\mathcal{D}(\tau)}c_{j\tau k}\cdot x^{\mathrm{F}}_{j\tau k}+r_\tau\left(\sum_{j\in\mathcal{P}(\tau),k\in\mathcal{D}(\tau)}x^{\mathrm{F}}_{j\tau k}\right)\right)\\
\leq&\sum_{\tau\in\mathcal{T}}\phi_\tau\left(-\sum_{j\in\mathcal{P}(\tau),k\in\mathcal{D}(\tau)}c_{j\tau k}\cdot x^{\mathrm{S}}_{j\tau k}+r_\tau\left(\sum_{j\in\mathcal{P}(\tau),k\in\mathcal{D}(\tau)}x^{\mathrm{S}}_{j\tau k}\right)\right)\\
&+\sum_{\tau\in\mathcal{T}}\phi_\tau\left(-\sum_{j\in\mathcal{P}(\tau),k\in\mathcal{D}(\tau)}c_{j\tau k}\cdot x^{\mathrm{DAG}}_{j\tau k}+r_\tau\left(\sum_{j\in\mathcal{P}(\tau),k\in\mathcal{D}(\tau)}x^{\mathrm{DAG}}_{j\tau k}\right)\right).
\end{aligned}$$

Here the inequality follows from $\mathbf{x}^{\mathrm{F}}=\mathbf{x}^{\mathrm{S}}+\mathbf{x}^{\mathrm{DAG}}$ and the fact that $r_\tau(\cdot)$ is subadditive by virtue of being a non-negative concave function. □

A.2 Lyapunov Analysis: Proof of Lemma 1.1 and Lemma 1.2

In this section, we prove the counterparts of Lemma 1.1 and Lemma 1.2 for JEA (Lemma A.6 and Lemma A.8, resp.) and JPA setting (Lemma A.7 and Lemma A.9, resp.). Lemma 1.1 is implied by Lemma A.6, and Lemma 1.2 is implied by Lemma A.8.

A.2.1 Decomposition of Optimality Gap

Generalization of Lemma 1.1 for the JEA Setting

The following lemma generalizes Lemma 1.1 for the JEA setting.

Lemma A.6. *Consider congestion functions $f_j(\cdot)$s that are strictly increasing and continuously differentiable, and that $f_j(\bar{q}_j)\leq f_k(\bar{q}_k)$ (i) for any $k\in V$ if $q_j=0$, and (ii) for* *any $j \in V$ if $q_k = d_k$, $k \in V_{\mathrm{b}}$. We have the following decomposition (W^{SPP^t} is defined in Appendix A.1.1 and g^t_{JEA} is defined in* (1.25)*):*

$$W^{\mathrm{SPP}^t} - \mathbb{E}[v^{\mathrm{MBP}}[t]|\bar{\mathbf{q}}[t]] \leq \underbrace{\tilde{K}\left(F(\bar{\mathbf{q}}[t]) - \mathbb{E}[F(\bar{\mathbf{q}}[t+1])|\bar{\mathbf{q}}[t]]\right)}_{\mathcal{V}_1} + \underbrace{\frac{1}{2\tilde{K}} \cdot \max_{j\in V} f'_j(\bar{q}_j[t])}_{\mathcal{V}_2}$$
$$+ \underbrace{\left(W^{\mathrm{SPP}^t} - g^t_{\mathrm{JEA}}(\mathbf{f}(\bar{\mathbf{q}}[t]))\right)}_{\mathcal{V}_3} + \underbrace{\mathbb{1}\left\{q_j[t] = 0 \text{ or } d_j, \exists j \in V\right\}}_{\mathcal{V}_4}.$$

Proof. For congestion functions $f_j(\bar{q}_j)$ that are strictly increasing and continuous for each j, we consider the Lyapunov function $F(\bar{\mathbf{q}})$ which is the antiderivative of $\mathbf{f}(\bar{\mathbf{q}})$. The Bregman divergence associated with $\mathbf{f}(\bar{\mathbf{q}})$ is defined as:

$$D_F(\bar{\mathbf{q}}_1, \bar{\mathbf{q}}_2) = F(\bar{\mathbf{q}}_1) - F(\bar{\mathbf{q}}_2) - \langle \mathbf{f}(\bar{\mathbf{q}}_1), \bar{\mathbf{q}}_1 - \bar{\mathbf{q}}_2 \rangle . \tag{A.17}$$

Plugging $\bar{\mathbf{q}}_1 = \bar{\mathbf{q}}[t+1]$, $\bar{\mathbf{q}}_2 = \bar{\mathbf{q}}[t]$ into (A.17) and rearranging the terms, we have:

$$F(\bar{\mathbf{q}}[t+1]) - F(\bar{\mathbf{q}}[t]) = \langle \mathbf{f}(\bar{\mathbf{q}}[t]), \bar{\mathbf{q}}[t+1] - \bar{\mathbf{q}}[t] \rangle + D_F(\bar{\mathbf{q}}[t+1], \bar{\mathbf{q}}[t]) .$$

Subtracting $\frac{1}{\tilde{K}} \sum_{\tau\in\mathcal{T}} \phi^t_\tau \sum_{j\in\mathcal{P}(\tau),k\in\mathcal{D}(\tau)} w_{j\tau k} \cdot x_{j\tau k}[t]$ on both sides and taking conditional expectation given $\bar{\mathbf{q}}[t]$, we have:

$$\mathbb{E}[F(\bar{\mathbf{q}}[t+1])|\bar{\mathbf{q}}[t]] - F(\bar{\mathbf{q}}[t]) - \frac{1}{\tilde{K}} \sum_{\tau\in\mathcal{T}} \phi^t_\tau \sum_{j\in\mathcal{P}(\tau),k\in\mathcal{D}(\tau)} w_{j\tau k} \mathbb{E}[x_{j\tau k}[t]|\bar{\mathbf{q}}[t]]$$
$$= \underbrace{-\frac{1}{\tilde{K}} \sum_{\tau\in\mathcal{T}} \phi^t_\tau \sum_{j\in\mathcal{P}(\tau),k\in\mathcal{D}(\tau)} w_{j\tau k} \mathbb{E}[x_{j\tau k}[t]|\bar{\mathbf{q}}[t]] + \langle \mathbf{f}(\bar{\mathbf{q}}[t]), \mathbb{E}[\bar{\mathbf{q}}[t+1]|\bar{\mathbf{q}}[t]] - \bar{\mathbf{q}}[t] \rangle}_{\text{(I)}} + \underbrace{\mathbb{E}\left[D_F(\bar{\mathbf{q}}[t+1], \bar{\mathbf{q}}[t])|\bar{\mathbf{q}}[t]\right]}_{\text{(II)}} \tag{A.18}$$

Let $x^{\mathrm{NOM}}_{j\tau k}[t]$ be the "nominal" control that ignores the no-underflow constraint, i.e.

$$(x^{\mathrm{NOM}}_{j\tau k})[t] = \begin{cases} 1 & \text{if } w_{j\tau k} + f_j(\bar{q}_j[t]) - f_k(\bar{q}_k[t]) \geq 0 \\ 0 & \text{otherwise.} \end{cases} \tag{A.19}$$

It immediately follows that

$$(x^{\mathrm{MBP}}_{j\tau k})[t] = (x^{\mathrm{NOM}}_{j\tau k})[t] \cdot \mathbb{1}\{q_j[t] > 0,\ q_k[t] < d_k\} . \tag{A.20}$$

With a slight abuse of notation, denote $\mathbf{x}^{\mathrm{NOM}}$ as $\tilde{\mathbf{x}}$, $\mathbf{x}^{\mathrm{MBP}}$ as $\mathbf{x}$. Rearranging the terms in (I) and plugging in (A.20), we have

$$\begin{aligned}
\text{(I)} = & -\frac{1}{\tilde{K}}\sum_{\tau\in\mathcal{T}}\phi_\tau^t \sum_{j\in\mathcal{P}(\tau),k\in\mathcal{D}(\tau)} \big(w_{j\tau k} + f_j(\bar{q}_j[t]) - f_k(\bar{q}_k[t])\big)\cdot\mathbb{E}[x_{j\tau k}[t]|\bar{\mathbf{q}}[t]] \\
= & -\frac{1}{\tilde{K}}\sum_{\tau\in\mathcal{T}}\phi_\tau^t \sum_{j\in\mathcal{P}(\tau),k\in\mathcal{D}(\tau)} \big(w_{j\tau k} + f_j(\bar{q}_j[t]) - f_k(\bar{q}_k[t])\big)\cdot\mathbb{E}[\tilde{x}_{j\tau k}[t]|\bar{\mathbf{q}}[t]] \\
& + \frac{1}{\tilde{K}}\sum_{\tau\in\mathcal{T}}\phi_\tau^t \sum_{j\in\mathcal{P}(\tau),k\in\mathcal{D}(\tau)} \big(w_{j\tau k} + f_j(\bar{q}_j[t]) - f_k(\bar{q}_k[t])\big)\cdot\mathbb{E}[\tilde{x}_{j\tau k}[t]|\bar{\mathbf{q}}[t]]\cdot\mathbb{1}\left\{q_j[t]=0 \text{ or } q_k[t]=d_k\right\}.
\end{aligned}$$

By definition of the nominal control $\tilde{\mathbf{x}}$, we have:

$$\begin{aligned}
& -\frac{1}{\tilde{K}}\sum_{\tau\in\mathcal{T}}\phi_\tau^t \sum_{j\in\mathcal{P}(\tau),k\in\mathcal{D}(\tau)} \big(w_{j\tau k} + f_j(\bar{q}_j[t]) - f_k(\bar{q}_k[t])\big)\cdot\mathbb{E}[\tilde{x}_{j\tau k}[t]|\bar{\mathbf{q}}[t]] \\
= & -\frac{1}{\tilde{K}}\sum_{\tau\in\mathcal{T}}\phi_\tau^t \sum_{j\in\mathcal{P}(\tau),k\in\mathcal{D}(\tau)} \big(w_{j\tau k} + f_j(\bar{q}_j[t]) - f_k(\bar{q}_k[t])\big)^+ \\
= & -\frac{1}{\tilde{K}}\cdot g^t_{\text{JEA}}(\mathbf{f}(\bar{\mathbf{q}}[t]))\,.
\end{aligned}$$

Using the fact that $\max_{j,k\in V,\tau\in\mathcal{T}}|w_{j\tau k}|=1$, we have

$$\begin{aligned}
& \frac{1}{\tilde{K}}\sum_{\tau\in\mathcal{T}}\phi_\tau^t \sum_{j\in\mathcal{P}(\tau),k\in\mathcal{D}(\tau)} \big(w_{j\tau k} + f_j(\bar{q}_j[t]) - f_k(\bar{q}_k[t])\big)\cdot\mathbb{E}[\tilde{x}_{j\tau k}[t]|\bar{\mathbf{q}}[t]]\cdot\mathbb{1}\left\{q_j[t]=0 \text{ or } q_k[t]=d_k\right\} \\
\le & \frac{1}{\tilde{K}}\sum_{\tau\in\mathcal{T}}\phi_\tau^t\cdot\mathbb{1}\left\{q_j[t]=0 \text{ or } d_j,\ \exists j\right\} \\
& + \frac{1}{\tilde{K}}\sum_{\tau\in\mathcal{T}}\phi_\tau^t \sum_{j\in\mathcal{P}(\tau),k\in\mathcal{D}(\tau)} \big(f_j(\bar{q}_j[t]) - f_k(\bar{q}_k[t])\big)^+\cdot\mathbb{1}\left\{q_j[t]=0 \text{ or } q_k[t]=d_k\right\} \\
\le & \frac{1}{\tilde{K}}\cdot\mathbb{1}\left\{q_j[t]=0 \text{ or } d_j,\ \exists j\right\}\,.
\end{aligned}$$

Here the last inequality follows from the assumption that $f_j(\bar{q}_j[t]) \le f_k(\bar{q}_k[t])$ for any $j,k\in V$ when $q_j[t]=0$ or $q_k[t]=d_k$. (Condition (ii) in the lemma as stated only covers $k\in V_{\text{b}}$. However, in case where $k\notin V_{\text{b}}$, i.e., $d_k=K$, and $q_k[t]=d_k$ holds, then we automatically have $q_k[t]=K \;\Rightarrow\; q_j[t]=0$ and condition (i) kicks in, i.e., condition (ii) in fact holds for all $k\in V$.) Note that when no queue has finite buffer constraints as in the illustrative model in Section 1.2, such assumption is satisfied by any congestion function such that $f_j(\bar{q}_j)=f(\bar{q}_j)$ for all $j\in V$ where $f(\cdot)$ is a monotonically increasing function.

Combining the above inequality and equality yields

$$(\mathrm{I}) \leq -\frac{1}{\tilde{K}} \cdot g^t_{\mathrm{JEA}}(\mathbf{f}(\bar{\mathbf{q}}[t])) + \frac{1}{\tilde{K}} \cdot \mathbb{1}\left\{q_j[t] = 0 \text{ or } d_j,\ \exists j\right\} .$$

Now we proceed to bound (II). By definition of Bregman divergence, (II) is the second order remainder of the Taylor series of $F(\cdot)$. Using the fact that $f(\cdot)$ is increasing, we have[1]

$$(\mathrm{II}) \leq \frac{1}{2}\sum_{j\in V} \mathbb{E}\left[f'_j(\bar{q}_j[t])(\bar{q}_j[t] - \bar{q}_j[t+1])^2 | \bar{\mathbf{q}}[t]\right] \leq \frac{1}{2\tilde{K}^2} \cdot \max_{j\in V} f'_j(\bar{q}_j[t]) .$$

Plugging the above bounds on (I) and (II) into (A.18), we have

$$\begin{aligned} &\mathbb{E}[F(\bar{\mathbf{q}}[t+1])|\bar{\mathbf{q}}[t]] - F(\bar{\mathbf{q}}[t]) - \frac{1}{\tilde{K}}\mathbb{E}[v^{\mathrm{MBP}}[t]|\bar{\mathbf{q}}[t]] \\ \leq& -\frac{1}{\tilde{K}} \cdot g^t_{\mathrm{JEA}}(\mathbf{f}(\bar{\mathbf{q}}[t])) + \frac{1}{2\tilde{K}^2} \cdot \max_{j\in V} f'_j(\bar{q}_j[t]) + \frac{1}{\tilde{K}} \cdot \mathbb{1}\left\{q_j[t] = 0 \text{ or } d_j, \exists j\right\} . \end{aligned}$$

Rearranging the terms yields:

$$\begin{aligned} -\mathbb{E}[v^{\mathrm{MBP}}[t]|\bar{\mathbf{q}}[t]] \leq \tilde{K}\left(F(\bar{\mathbf{q}}[t]) - \mathbb{E}[F(\bar{\mathbf{q}}[t+1])|\bar{\mathbf{q}}[t]]\right) + \frac{1}{2\tilde{K}} \cdot \max_{j\in V} f'_j(\bar{q}_j[t]) \\ - g^t_{\mathrm{JEA}}(\mathbf{f}(\bar{\mathbf{q}}[t])) + \mathbb{1}\left\{q_j[t] = 0 \text{ or } d_j, \exists j\right\} . \end{aligned}$$

Adding W^{SPP^t} to both sides concludes the proof. □

Joint Pricing-Assignment Setting

For JPA setting, we have the following lemma which is analogous to Lemma 1.1.

Lemma A.7. *Consider congestion functions $f_j(\cdot)$s that are strictly increasing and continuously differentiable, and that $f_j(\bar{q}_j) \leq f_k(\bar{q}_k)$ (i) for any $k \in V$ if $q_j = 0$, and (ii) for any $j \in V$ if $q_k = d_k$, $k \in V_{\mathrm{b}}$. We have the following decomposition:*

$$\begin{aligned} W^* - \mathbb{E}[v^{\mathrm{MBP}}[t]|\bar{\mathbf{q}}[t]] \leq \tilde{K}\left(F(\bar{\mathbf{q}}[t]) - \mathbb{E}[F(\bar{\mathbf{q}}[t+1])|\bar{\mathbf{q}}[t]]\right) + \frac{1}{2\tilde{K}} \cdot \max_{j\in V} f'_j(\bar{q}_j[t]) \\ + \left(W^{\mathrm{SPP}} - g_{\mathrm{JPA}}(\mathbf{f}(\bar{\mathbf{q}}[t]))\right) + \mathbb{1}\left\{q_j[t] = 0 \text{ or } d_j, \exists j\right\} , \end{aligned} \tag{A.21}$$

where $g_{\mathrm{JPA}}(\mathbf{y})$ is defined in (1.29).

[1] For exposition simplicity, we ignore the difference between $f'(\bar{q}_j[t])$ and $f'(\bar{q}_j[t+1])$ in the Taylor expansion.

Proof Sketch. The proof is analogous to Lemma A.6. To use the strong duality argument, we prove below that $g_{\text{JPA}}(\cdot)$ defined in (1.29) is indeed the partial dual function of the SPP (A.14)-(A.16). Then because the primal problem is a concave optimization problem with linear constraint, strong duality must hold.

Let $\mathbf{y}$ be the Lagrange multipliers corresponding to constraints (A.15). We have

$$\begin{aligned} g_{\text{JPA}}(\mathbf{y}) = &\max_{\sum_{j\in\mathcal{P}(\tau),k\in\mathcal{D}(\tau)} x_{j\tau k}\le 1, x_{j\tau k}\ge 0} \sum_{\tau\in\mathcal{T}} \phi_\tau \left(r_\tau \left(\sum_{j\in\mathcal{P}(\tau),k\in\mathcal{D}(\tau)} x_{j\tau k} \right) \right. \\ &\left. + \sum_{j\in\mathcal{P}(\tau),k\in\mathcal{D}(\tau)} \left(-c_{j\tau k} + y_j - y_k\right) x_{j\tau k} \right) \\ = &\sum_{\tau\in\mathcal{T}} \phi_\tau \max_{\sum_{j\in\mathcal{P}(\tau),k\in\mathcal{D}(\tau)} x_{j\tau k}\le 1, x_{j\tau k}\ge 0} \left(r_\tau \left(\sum_{j\in\mathcal{P}(\tau),k\in\mathcal{D}(\tau)} x_{j\tau k} \right) \right. \\ &\left. + \sum_{j\in\mathcal{P}(\tau),k\in\mathcal{D}(\tau)} \left(-c_{j\tau k} + y_j - y_k\right) x_{j\tau k} \right) . \end{aligned}$$

Let $\mu_\tau = \sum_{j\in\mathcal{P}(\tau),k\in\mathcal{D}(\tau)} x_{j\tau k}$, we have

$$\begin{aligned} g_{\text{JPA}}(\mathbf{y}) &= \sum_{\tau\in\mathcal{T}} \phi_\tau \max_{0\le\mu_\tau\le 1} \max_{\sum_{j\in\mathcal{P}(\tau),k\in\mathcal{D}(\tau)} x_{j\tau k}=\mu_\tau, x_{j\tau k}\ge 0} \left(r_\tau(\mu_\tau) + \sum_{j\in\mathcal{P}(\tau),k\in\mathcal{D}(\tau)} \left(-c_{j\tau k} + y_j - y_k\right) x_{j\tau k} \right) \\ &= \sum_{\tau\in\mathcal{T}} \max_{0\le\mu_\tau\le 1} \left(r_\tau(\mu_\tau) + \mu_\tau \max_{j\in\mathcal{P}(\tau),k\in\mathcal{D}(\tau)} \left(-c_{j\tau k} + y_j - y_k\right) \right) . \end{aligned}$$

□

A.2.2 Geometry of the Dual Function

Generalization of Lemma 1.2 for the JEA Setting

The following lemma generalizes Lemma 1.2 for the JEA setting.

Lemma A.8. *Consider congestion functions $(f_j(\cdot))_{j\in V}$ that are strictly increasing and continuously differentiable, and any triple $(\boldsymbol{\phi}^t, \mathcal{P}, \mathcal{D})$ with connectedness at least $\alpha_{\min} > 0$.*

Then the term $\mathcal{V}_3 = W^{\mathrm{SPP}^t} - g^t_{\mathrm{JEA}}(\bar{\mathbf{q}}[t])$ (see Lemma A.6) is bounded above as

$$\mathcal{V}_3 \leq \ -\alpha_{\min} \cdot \left[\max_{j\in V} f_j(\bar{q}_j[t]) - \min_{j\in V} f_j(\bar{q}_j[t]) - 2m\right]^+ .$$

Proof. Consider $\mathbf{y} \triangleq (f_j(\bar{q}_j[t])_{j\in V}$ and order the nodes in V in decreasing order of y_j as $y_{i_1} \geq y_{i_2} \geq \cdots y_{i_m}$. For $r = 1$ to $r = m-1$, we repeat the following procedure: if $y_{i_r} - y_{i_{r+1}} \leq 2$, then do nothing and move on to $r+1$; if otherwise, perform the following update:

$$y_{i_k} \leftarrow \ y_{i_k} - \big(y_{i_r} - y_{i_{r+1}} - 2\big) \qquad \forall 1 \leq k \leq r\,.$$

Recall that $g(\mathbf{y}) = \sum_{\tau\in\mathcal{T}} \phi_\tau \sum_{j\in\mathcal{P}(\tau),k\in\mathcal{D}(\tau)} [w_{j\tau k} + y_j - y_k]^+$. For the terms where $j,k \in \{i_1,\cdots,i_r\}$ or $j,k \in \{i_{r+1},\cdots,i_m\}$, their value are not affected by the update. Consider the terms where $j \in \{i_1,\cdots,i_r\}$, $k \in \{i_{r+1},\cdots,i_m\}$: If $y_{i_r} - y_{i_{r+1}} > 2$, then after the update, for $\tau \in \mathcal{P}^{-1}(j) \cap \mathcal{D}^{-1}(k)$,

$$w_{j\tau k} + y_j - y_k \geq \ w_{j\tau k} + y_{i_r} - \big(y_{i_r} - y_{i_{r+1}} - 2\big) - y_{i_{r+1}} \geq w_{j\tau k} + 2 > 0\,,$$

hence the update decrease these terms each by $y_{i_r} - y_{i_{r+1}} - 2$. Finally, for the terms where $j \in \{i_{r+1},\cdots,i_m\}, k \in \{i_1,\cdots,i_r\}$, it is easy to verify that their value stay at zero after the update. To sum up, such an update decreases $g(\mathbf{y})$ by at least

$$\left(\sum_{\tau\in\mathcal{P}^{-1}(\{i_1,\cdots,i_r\})\cap\mathcal{D}^{-1}(\{i_{r+1},\cdots,i_m\})} \phi_\tau\right) \cdot \big[y_{i_r} - y_{i_{r+1}} - 2\big]^+ .$$

Note that the first term is lower bounded by $\alpha_{\min}$ defined in (1.26). As a result, after the finishing the procedure, $g(\mathbf{y})$ decreased by at least:

$$\alpha_{\min} \cdot \sum_{r=1}^{m-1} \big[y_{i_r} - y_{i_{r+1}} - 2\big]^+ \geq \ \alpha_{\min} \cdot [y_{i_1} - y_{i_m} - 2m]^+ .$$

By strong duality we have $\min_{\mathbf{y}} g^t_{\mathrm{JEA}}(\mathbf{y}) = W^{\mathrm{SPP}^t}$, hence

$$g^t_{\mathrm{JEA}}(\mathbf{y}) - W^{\mathrm{SPP}^t} \geq \ \alpha_{\min} \cdot \left[\max_{j\in V} y_j - \min_{k\in V} y_k - 2m\right]^+ .$$

This concludes the proof. □

Joint Pricing-Assignment Setting

The following lemma is the counterpart of Lemma 1.2 for the JPA setting.

Lemma A.9. *Consider congestion functions* $(f_j(\cdot))_{j\in V}$ *that are strictly increasing and continuously differentiable, and any* $\boldsymbol{\phi}$ *with connectedness* $\alpha(\boldsymbol{\phi}, \mathcal{P}, \mathcal{D}) > 0$. *We have*

$$g_{\text{JPA}}(\mathbf{y}) - W^{\text{SPP}} \geq \ \alpha(\boldsymbol{\phi}, \mathcal{P}, \mathcal{D}) \cdot \left[\max_{j\in V} y_j - \min_{k\in V} y_k - 2m\right]^+ ,$$

where W^{SPP} *is the value of SPP* (A.14)-(A.16)*, and* $\alpha(\boldsymbol{\phi}, \mathcal{P}, \mathcal{D})$ *is defined in* (1.26).

Proof Sketch. The proof is a direct extension of the proof of Lemma 1.2. The key observation is that: if $y_j - y_k \geq 2 \geq 2\max_{j,k\in V,\tau\in\mathcal{T}} |c_{j\tau k}| + \bar{p}$, then for any $\tau \in \mathcal{P}^{-1}(j) \cap \mathcal{D}^{-1}(k)$ we have

$$\text{argmax}_{\{0\leq\mu_\tau\leq 1\}} \left(r_\tau(\mu_\tau) + \mu_\tau \cdot \max_{j\in\mathcal{P}(\tau),k\in\mathcal{D}(\tau)} \left(-c_{j\tau k} + y_j - y_k\right)\right) = 1\,,$$

for any $\tau \in \mathcal{P}^{-1}(k) \cap \mathcal{D}^{-1}(j)$ we have:

$$\text{argmax}_{\{0\leq\mu_\tau\leq 1\}} \left(r_\tau(\mu_\tau) + \mu_\tau \cdot \max_{k\in\mathcal{P}(\tau),j\in\mathcal{D}(\tau)} \left(-c_{k\tau j} + y_k - y_j\right)\right) = 0\,.$$

□

A.3 Proofs of Lemma 1.3 and Theorems 1.2 and 1.3

In this section, we first show that Lemma 1.3 and its counterparts for JEA and JPA settings hold if the congestion function satisfy certain growth conditions. Together with the lemmas derived in Appendix A.2, we conclude that Theorems 1.2 and 1.3 hold for any congestion function that satisfies the growth conditions. Finally, we verify the growth condition for several congestion functions including (1.22).

A.3.1 A Sufficient Condition of Lemma 1.3 and its Counterparts for JEA and JPA

Since the statements and proofs for the JEA and JPA settings are almost identical, we only provide them for the JEA setting to avoid redundancy. The generalization of Lemma 1.3 to the JEA setting is as follows:

Lemma A.10. *Consider the congestion function* (1.22). *Consider a set V of $m = |V| > 1$ nodes, a subset $V_{\mathrm{b}} \subset V$ of buffer-constrained nodes with scaled buffer sizes $\bar{d}_j \in (0,1)$ (recall that we define $\bar{d}_j \triangleq 1$ for all $j \in V \backslash V_{\mathrm{b}}$) satisfying $\sum_{j\in V} \bar{d}_j > 1$, and any $(\boldsymbol{\phi}^t, \mathcal{P}, \mathcal{D})$ that satisfies Condition 1.2 with $\alpha_{\min} > 0$. Then there exists $K_1 = \mathrm{poly}\left(m, \bar{\mathbf{d}}, \frac{1}{\alpha_{\min}}\right)$ such that for $K \geq K_1$,*

$$\mathcal{V}_2 + \mathcal{V}_3 + \mathcal{V}_4 \leq M_2 \cdot \frac{1}{\tilde{K}}, \qquad \text{for } M_2 = C\frac{\sqrt{m}}{\min_{j\in V} \bar{d}_j}\left(\frac{\sum_{j\in V} \bar{d}_j}{\min\{\sum_{j\in V} \bar{d}_j - 1, 1\}}\right)^{3/2},$$

where $\mathcal{V}_2, \mathcal{V}_3, \mathcal{V}_4$ were defined in Lemma A.6, $\tilde{K}$ was defined in (1.21)*, and $C > 0$ is a universal constant that is independent of m, $\bar{\mathbf{d}}$, K, or $\alpha_{\min}$.*

Since the congestion function (1.7) is a special case of (1.22), and the illustrative model introduced in Section 1.2 is a special case of JEA model, Lemma A.10 will imply Lemma 1.3.

We define below a growth condition for congestion functions. Lemma A.11 will imply that if the congestion function satisfies this growth condition (with certain parameters), then Lemma A.10 holds.

Condition A.1 (Growth condition for congestion functions)**.** *We say the congestion functions $(f_j(\cdot))_{j\in V}$ satisfy the growth condition with parameters $(\alpha, K_1, M_1, M_2) \in \mathbb{R}^4_{++}$ if the following holds:*

1. *For each $j \in V$, $f_j(\cdot)$ is strictly increasing and continuously differentiable. Moreover,*

(a) For any $K > K_1$, $f_j(\bar{q}_j) \le f_k(\bar{q}_k)$ (i) for any $k \in V$ if $q_j = 0$, and (ii) for any $j \in V$ if $q_k = d_k$, $k \in V_{\mathrm{b}}$.

(b) For any $j, k \in V$, we have $f_j\left(\frac{\bar{d}_j}{\sum_{j\in V}\bar{d}_j}\right) = f_k\left(\frac{\bar{d}_k}{\sum_{j\in V}\bar{d}_j}\right)$.

2. Define

$$\mathcal{B}(\mathbf{f}) \triangleq \left\{ \bar{\mathbf{q}} \in \Omega :\ \max_{j\in V}\left| f_j\left(\frac{\bar{d}_j}{\sum_{j\in V}\bar{d}_j}\right) - f_j(\bar{q}_j)\right| \le 4m \right\}.$$

Denote $\bar{\mathcal{B}}(\mathbf{f}) \triangleq \Omega\backslash\mathcal{B}(\mathbf{f})$.

(a) For any $K > K_1$, $\forall \bar{\mathbf{q}} \in \bar{\mathcal{B}}(\mathbf{f})$,

$$\alpha\left(\max_{j\in V}\left| f_j\left(\frac{\bar{d}_j}{\sum_{j\in V}\bar{d}_j}\right) - f_j(\bar{q}_j)\right| - 2m\right)^+ \ge \frac{1}{2\tilde{K}}\cdot \max_{j\in V} f_j'(\bar{q}_j) + \mathbb{1}\{q_j = 0 \text{ or } d_j, \exists j\}. \tag{A.22}$$

(b) Let $F(\bar{\mathbf{q}})$ be the antiderivative of $\mathbf{f}(\bar{\mathbf{q}}) \triangleq (f_j(\bar{q}_j))_{j\in V}$, we have $\sup_{\mathbf{q},\mathbf{q}'\in\Omega}(F(\bar{\mathbf{q}}) - F(\bar{\mathbf{q}}')) \le M_1$.

(c) We have $\sup_{\bar{\mathbf{q}}\in\mathcal{B}(\mathbf{f})}\max_{j\in V} f_j'(\bar{q}_j) \le M_2$.

(d) If $\exists j \in V$ such that $q_j = 0$ or $q_j = d_j$, then $\bar{\mathbf{q}} \in \bar{\mathcal{B}}(\mathbf{f})$.

Lemma A.11. *In the JEA setting, if the congestion functions $(f_j(\cdot))_{j\in V}$ satisfy the growth conditions (Condition A.1) with parameters $(\alpha_{\min}, K_1, M_1, M_2)$, then for $K \ge K_1$,*

$$\mathcal{V}_2 + \mathcal{V}_3 + \mathcal{V}_4 \le M_2 \cdot \frac{1}{\tilde{K}}, \tag{A.23}$$

where $\mathcal{V}_2, \mathcal{V}_3, \mathcal{V}_4$ were defined in Lemma A.6 and $\tilde{K}$ was defined in (1.21).

Proof of Lemma A.11. Recall that

$$\mathcal{V}_2 + \mathcal{V}_3 + \mathcal{V}_4 = \frac{1}{2\tilde{K}}\cdot\max_{j\in V} f_j'(\bar{q}_j[t]) + \left(W^{\mathrm{SPP}} - g^t_{\mathrm{JEA}}(\mathbf{f}(\bar{\mathbf{q}}[t]))\right) + \mathbb{1}\left\{q_j[t] = 0 \text{ or } d_j, \exists j \in V\right\}.$$

For $\bar{\mathbf{q}} \in \mathcal{B}(\mathbf{f})$, since the congestion functions satisfy Condition A.1, we have $\mathcal{V}_4 = 0$.

By definition of W^{SPP} we have $\mathcal{V}_3 \leq 0$. As a result, it follows from Condition A.1 that

$$\mathcal{V}_2 + \mathcal{V}_3 + \mathcal{V}_4 \leq \frac{1}{2\tilde{K}} \cdot \sup_{\bar{\mathbf{q}} \in \mathcal{B}(\mathbf{f})} \max_{j \in V} f'_j(\bar{q}_j[t]) = M_2 \cdot \frac{1}{\tilde{K}}\,.$$

For $\bar{\mathbf{q}} \in \bar{\mathcal{B}}(\mathbf{f})$, it follows from Lemma A.8 that

$$\mathcal{V}_3 \leq\ -\alpha_{\min} \cdot \left[\max_{j \in V} f_j(\bar{q}_j) - \min_{j \in V} f_j(\bar{q}_j) - 2m\right]^+ \,.$$

Note that

$$\begin{aligned}
&\max_{j \in V} \left| f_j\left(\frac{\bar{d}_j}{\sum_{j \in V} \bar{d}_j}\right) - f_j(\bar{q}_j)\right| \\
&\leq\ \max\left\{\max_{j \in V} f_j(\bar{q}_j) - \min_{j \in V} f_j\left(\frac{\bar{d}_j}{\sum_{j \in V} \bar{d}_j}\right), \max_{j \in V} f_j\left(\frac{\bar{d}_j}{\sum_{j \in V} \bar{d}_j}\right) - \min_{j \in V} f_j(\bar{q}_j)\right\}\,.
\end{aligned}$$

Note that there must exists $j^* \in V$ such that $\bar{q}_{j^*} \leq \frac{\bar{d}_{j^*}}{\sum_{j \in V} \bar{d}_j}$, hence $f_{j^*}(\bar{q}_{j^*}) \leq f_j\left(\frac{\bar{d}_{j^*}}{\sum_{j \in V} \bar{d}_j}\right)$. Because the congestion functions satisfy Condition A.1 point 1(d), we have $f_j\left(\frac{\bar{d}_j}{\sum_{j \in V} \bar{d}_j}\right)$ has the same value for all $j \in V$, therefore

$$\begin{aligned}
\max_{j \in V} f_j(\bar{q}_j) - \min_{j \in V} f_j\left(\frac{\bar{d}_j}{\sum_{j \in V} \bar{d}_j}\right) &=\ \max_{j \in V} f_j(\bar{q}_j) - f_{j^*}\left(\frac{\bar{d}_{j^*}}{\sum_{j \in V} \bar{d}_j}\right) \\
&\leq\ \max_{j \in V} f_j(\bar{q}_j) - f_{j^*}(\bar{q}_{j^*}) \\
&\leq\ \max_{j \in V} f_j(\bar{q}_j) - \min_{j \in V} f_j(\bar{q}_j)\,.
\end{aligned}$$

Similarly, we can show that

$$\max_{j \in V} f_j\left(\frac{\bar{d}_j}{\sum_{j \in V} \bar{d}_j}\right) - \min_{j \in V} f_j(\bar{q}_j) \leq\ \max_{j \in V} f_j(\bar{q}_j) - \min_{j \in V} f_j(\bar{q}_j)\,.$$

Combined, we have

$$\mathcal{V}_3 \leq\ -\alpha_{\min} \cdot \left(\max_{j \in V} \left| f_j\left(\frac{\bar{d}_j}{\sum_{j \in V} \bar{d}_j}\right) - f_j(\bar{q}_j)\right| - 2m\right)^+ \,.$$

Plugging in Condition A.1 point 2(a), we have for $\bar{\mathbf{q}} \in \bar{\mathcal{B}}(\mathbf{f})$,

$$\mathcal{V}_2 + \mathcal{V}_3 + \mathcal{V}_4 \leq\ 0\,.$$

Combine the above two cases, we conclude the proof. □

It remains to be shown that the congestion function (1.22) satisfies Condition A.1.

Lemma A.12. *The congestion function* (1.22) *satisfies the growth conditions (Condition A.1) with parameters* $(\alpha_{\min}, K_1, M_1, M_2)$ *where*

$$K_1 = \text{poly}\left(m, \bar{\mathbf{d}}, \frac{1}{\alpha_{\min}}\right), \quad M_1 = Cm\,, \quad M_2 = C\frac{1}{\min_{j\in V}\bar{d}_j}\left(\frac{\sum_{j\in V}\bar{d}_j}{\min\{\sum_{j\in V}\bar{d}_j - 1, 1\}}\right)^{3/2}\sqrt{m}\,.$$

Here C *is a universal constant that is independent of* m*,* $\bar{\mathbf{d}}$*,* K *and* $\alpha_{\min}$*.*

We delay the proof of Lemma A.12 to Appendix A.3.3. We are now ready to prove Lemma A.10.

Proof of Lemma A.10. Lemma A.10 immediately follows from Lemma A.11 and Lemma A.12. □

A.3.2 Proof of Main Theorems

Recall that we proved Theorem 1.1 in Section 1.5 using Lemmas 1.1, 1.2, and 1.3. Similarly, we can prove Theorem 1.2 and 1.3.

Proof of Theorem 1.2. We draw inspiration from the proof of Theorem 1.1, along with some additional work to handle time-varying demand arrival rates, for which we draw upon Lemma A.1 and Proposition A.1.

Note that for the congestion functions defined in (1.22), we have $f_j(\bar{q}_j[t]) \le f_k(\bar{q}_k[t])$ when $q_j[t] = 0$ or $q_k[t] = d_k$. Also, the functions are strictly increasing and continuously differentiable. Hence, Lemmas A.6, A.8, and A.10 (the JEA versions of Lemmas 1.1, 1.2 and A.10) apply to the congestion functions (1.22).

As in the proof of Theorem 1.1, we argue as follows: Plugging in Lemma A.10 into the bound in Lemma A.6 and taking expectation, we obtain

$$W^{\text{SPP}^t} - \mathbb{E}[v^{\text{MBP}}[t]] \le \tilde{K}\left(\mathbb{E}[F(\bar{\mathbf{q}}[t])] - \mathbb{E}[F(\bar{\mathbf{q}}[t+1])]\right) + M_2\frac{1}{\tilde{K}}\,, \tag{A.24}$$

for all $K \geq K_2 = \text{poly}\left(m, \bar{\mathbf{d}}, \frac{1}{\alpha_{\min}}\right)$, where $M_2 = C_2 \frac{\sqrt{m}}{\min_{j \in V} \bar{d}_j} \left(\frac{\sum_{j \in V} \bar{d}_j}{\min\{\sum_{j \in V} \bar{d}_j - 1, 1\}}\right)^{3/2}$ for a universal constant C_2. Consider the first T_0 periods. Take the sum of both sides of the inequality (A.24) from $t = 0$ to $t = T_0 - 1$, and divide the sum by T_0. This yields

$$
\begin{aligned}
\tfrac{1}{T_0} \sum_{t=0}^{T_0-1} W^{\mathrm{SPP}^t} \; - \; W_{T_0}^{\mathrm{MBP}} &\leq \frac{\tilde{K}}{T_0} \left(\mathbb{E}[F(\bar{\mathbf{q}}[0])] - \mathbb{E}[F(\bar{\mathbf{q}}[T_0])]\right) + M_2 \frac{1}{\tilde{K}} \\
&\leq \frac{\tilde{K}}{T_0} \left(\sup_{\bar{\mathbf{q}}, \bar{\mathbf{q}}' \in \Omega} \left(F(\bar{\mathbf{q}}) - F(\bar{\mathbf{q}}')\right) \right) + M_2 \frac{1}{\tilde{K}} \\
&\leq \frac{\tilde{K}}{T_0} C_1 m + M_2 \frac{1}{\tilde{K}} \,, \qquad \text{(A.25)}
\end{aligned}
$$

for all $K \geq K_2$, where C_1 is a universal constant. Here we used the bound $\sup_{\mathbf{q}, \mathbf{q}' \in \Omega}(F(\bar{\mathbf{q}}) - F(\bar{\mathbf{q}}')) \leq C_1 m$ from Lemma A.12 (specifically the part of the lemma about Condition A.1 part 2(b)).

Using Proposition A.1 and then Lemma A.1, we have

$$
\begin{aligned}
L_{T_0}^{\mathrm{MBP}} = W_{T_0}^* - W_{T_0}^{\mathrm{MBP}} &\leq W^{\overline{\mathrm{SPP}}} - W_{T_0}^{\mathrm{MBP}} + m \frac{K}{T_0} \\
&\leq \eta T_0 m/2 + \tfrac{1}{T_0} \sum_{t=0}^{T_0-1} W^{\mathrm{SPP}^t} \; - W_{T_0}^{\mathrm{MBP}} + m \frac{K}{T_0} \\
&\leq \frac{\tilde{K}}{T_0} m(C_1 + 1) + M_2 \frac{1}{\tilde{K}} + \eta T_0 \cdot m/2 \\
&\leq \frac{K}{T_0} 2m(C_1 + 1) + M_2 \frac{1}{K} + T_0 \eta \cdot m/2 \,. \qquad \text{(A.26)}
\end{aligned}
$$

where we used (A.25) in the third inequality, and $K \leq \tilde{K} \leq 2K$ for all $K \geq K_3 = m^2$ in the last inequality. It remains to choose T_0 appropriately, i.e., to divide the horizon T into intervals of appropriate length. Note that the bound on per period loss (A.26) is minimized for $T_0 = T_* = 2\sqrt{(C_1 + 1)K/\eta}$, which makes the first and third terms equal. This observation will guide our choice of T_0.

If $T \leq T_*$, we set $T_0 = T$ and we immediately have

$$
L_T^{\mathrm{MBP}} \leq \frac{K}{T} 4m(C_1 + 1) + M_2 \frac{1}{K} \qquad \forall T < T_* \,, \qquad \text{(A.27)}
$$

since the first term is larger than the third term in (A.26). If $T > T_*$ then we divide T into $\lceil T/T_* \rceil$ intervals of equal length (up to rounding error). In particular, each interval has length $T_0 \in [T_*/2, T_*]$, the first term is again larger than the third term in (A.26) and so the per period loss in each interval is bounded above by

$$\frac{K}{T_0} 4m(C_1+1) + M_2 \frac{1}{K} \le \frac{K}{T_*/2} 4m(C_1+1) + M_2 \frac{1}{K} = \sqrt{\eta K} 4m \sqrt{C_1+1} + M_2 \frac{1}{K}\,.$$

Since this bound holds for each interval, it holds for the full horizon of length T, i.e.,

$$L_T^{\mathrm{MBP}} \le \sqrt{\eta K} 4m \sqrt{C_1+1} + M_2 \frac{1}{K} \qquad \forall T \ge T_*\,. \tag{A.28}$$

Combining (A.27) and (A.28), we obtain that for any $K \ge K_1 \triangleq \max(K_2, K_3)$ and any horizon T, we have

$$L_T^{\mathrm{MBP}} \le \sqrt{\eta K} 4m \sqrt{C_1+1} + M_2 \frac{1}{K} + \frac{K}{T} 4m(C_1+1) \le M_1 \left(\frac{K}{T} + \sqrt{\eta K} \right) + M_2 \frac{1}{K}\,,$$

for $M_1 \triangleq 4m(C_1+1)$. Defining $C \triangleq \max(C_2, 4(C_1+1))$ we obtain the bound claimed in the theorem. □

Proof sketch for Theorem 1.3. The proof is a direct extension of the proof of Theorem 1.1, and follows from Lemmas A.7, A.9, and the JPA counterpart of Lemma A.10 (which is almost identical to Lemma A.10, and was hence omitted). We bound M_1 using Lemma A.12. □

Since Condition A.1 implies Lemma A.10 (using Lemma A.11), we have the following general version of Theorem 1.2 using the exact same proof as that of Theorem 1.2.

Theorem A.1 (General result for the JEA setting)**.** *Consider a set V of $m \triangleq |V| > 1$ nodes, a subset $V_{\mathrm{b}} \subseteq V$ of buffer-constrained nodes with scaled buffer sizes $\bar{d}_j \in (0,1)$ $\forall j \in V_{\mathrm{b}}$ satisfying*[2] *$\sum_{j \in V} \bar{d}_j > 1$, and a minimum connectivity $\alpha_{\min} > 0$. Consider any congestion functions $(f_j(\cdot))_{j \in V}$ that satisfy Condition A.1 with parameters $(\alpha = \alpha_{\min}, K_1, M_1, M_2) \in \mathbb{R}^4_{++}$. Then for any horizon T, any $K \ge K_1$, and any sequence of demand arrival rates $(\boldsymbol{\phi}^t)_{t=0}^{T-1}$ which varies η-slowly (for some $\eta \in [0,2]$) and pickup and dropoff neighborhoods $\mathcal{P}$ and $\mathcal{D}$ such that $(\boldsymbol{\phi}^t, \mathcal{P}, \mathcal{D})$ is $\alpha_{\min}$-strongly connected*

[2]Recall that we define $\bar{d}_j \triangleq 1$ for all $j \in V \backslash V_{\mathrm{b}}$.

(Condition 1.2) for all $t \leq T-1$, we have

$$L_T^{\mathrm{MBP}} \leq 4(M_1+m)\cdot\left(\frac{K}{T}+\sqrt{\eta K}\right)+M_2\cdot\frac{1}{K}\,.$$

In the following subsection, we will show examples of alternate congestion functions that satisfy Condition A.1 and obtain the corresponding parameters K_1, M_1 and M_2.

A.3.3 Validating Condition A.1 for Congestion Functions

In this section, we prove Lemma A.12. We will go a step further and show that Lemma A.12 holds several other congestion functions.

Recall the congestion function defined in (1.22): let $V_\mathrm{b} \subset V$ be the subset of buffer-constrained nodes with scaled buffer sizes $\bar{d}_j \in (0,1)$, and

$$f_j(\bar{q}_j) = \sqrt{m}\cdot C_b\cdot\left(\left(1-\frac{\bar{q}_j}{\bar{d}_j}\right)^{-\frac{1}{2}} - \left(\frac{\bar{q}_j}{\bar{d}_j}\right)^{-\frac{1}{2}} - D_b\right), \qquad \forall j \in V_\mathrm{b}$$

$$f_j(\bar{q}_j) = -\sqrt{m}\cdot\bar{q}_j^{-\frac{1}{2}}, \qquad \forall j \in V\backslash V_\mathrm{b}$$

Here C_b and D_b are normalizing constants chosen as follows. Define $\epsilon \triangleq \frac{\delta_K}{\tilde{K}}$ (where δ_K and $\tilde{K}$ were defined in (1.21)). Let $h_\mathrm{b}(\bar{q}) \triangleq (1-\bar{q})^{-\frac{1}{2}} - \bar{q}^{-\frac{1}{2}}$ and $h(\bar{q}) \triangleq -\bar{q}^{-\frac{1}{2}}$. Define $C_b \triangleq \frac{h(\epsilon)-h(1/\sum_{j\in V}\bar{d}_j)}{h_\mathrm{b}(\epsilon)-h_\mathrm{b}(1/\sum_{j\in V}\bar{d}_j)}$ and $D_b \triangleq h_\mathrm{b}(1/\sum_{j\in V}\bar{d}_j) - C_b^{-1}h(1/\sum_{j\in V}\bar{d}_j)$. These definitions ensure that Condition A.1 point 1(b) holds, and are useful in establishing Condition A.1 point 1(a).

Proof of Lemma A.12. (The proof of this lemma involves a lot of notations and computation. For readability, we use the following simplifying notation (with a slight abuse of notation): for $x_a, y_a \in \mathbb{R}_+$ where $a \in \mathcal{A} \subset \mathbb{Z}_+$, $\{x_a\} = O(\{y_a\})$ ($\{x_a\} = \Omega(\{y_a\})$, resp.) means that there exists a universal constant $C > 0$ that does not depend on $m, K, \bar{\mathbf{d}}$, or $\alpha_{\min}$ such that $x_a \leq Cy_a$ ($x \geq Cy_a$, resp.) for each $a \in \mathcal{A}$. We say $\{x_a\} = \Theta(\{y_a\})$ if $\{x_a\} = O(\{y_a\})$ and $\{x_a\} = \Omega(\{y_a\})$.) Denote $\bar{d}_\Sigma \triangleq \sum_{j\in V}\bar{d}_j$, $\bar{d}_g \triangleq \min\{1, \sum_{j\in V}\bar{d}_j - 1\}$, $\bar{d}_{\min} \triangleq \min_{j\in V}\bar{d}_j$. Recall that $\bar{d}_j \in (0,1)$ for any $j \in V_\mathrm{b}$, and that $\bar{d}_\Sigma > 1$.

- Point 1. It is not hard to see that the congestion functions $(f_j(\bar{q}_j))_{j\in V}$ are strictly increasing and continuously differentiable. For any $K>0$, we have $f_j(\bar{q}_j)=f_k(\bar{q}_k)$ for any $j,k\in V$ if $q_j=q_k=0$. As a result, if $q_j=0$, we have $f_j(\bar{q}_j)\le f_k(\bar{q}_k)$ for any $k\in V$. It can be easily verified that Point 1(b) is also satisfied by any $K>0$.

  It remains to be shown that there exists $K_1<\infty$ such that for $K\ge K_1$, we have $f_j(\bar{q}_j)\le f_k(\bar{q}_k)$ for any $j\in V$ if $q_k=d_k$ and $k\in V_{\mathrm{b}}$. To this end, if suffices to check the inequality $f_j(\bar{q}_j)\le f_k(\bar{q}_k)$ for $q_j=d_j$, $q_k=d_k$ where $j\in V\backslash V_{\mathrm{b}}$ and $k\in V_{\mathrm{b}}$: In this case, we have $f_j(\bar{q}_j)\le 0$; for $K=\Omega(\max\{\bar{d}_\Sigma^2,\frac{\bar{d}_\Sigma^2}{\bar{d}_g^2}\})$, we have $C_b=\Theta(1)$, $D_b=O(\sqrt{\frac{\bar{d}_\Sigma}{\bar{d}_g}})$ hence $f_k(\bar{q}_k)=\Omega(\sqrt{m}\frac{K^{1/4}}{\bar{d}_g^{1/2}})\ge 0$. Therefore point 1 is satisfied for $K_1=O(\max\{\bar{d}_\Sigma^2,\frac{\bar{d}_\Sigma^2}{\bar{d}_g^2}\})=O(\frac{\bar{d}_\Sigma^2}{\bar{d}_g^2})$.

- Point 2(a). For $\mathbf{q}$ such that $\bar{\mathbf{q}}\in\bar{\mathcal{B}}(\mathbf{f})$ and $0<q_j<d_j$ for any $j\in V$, we have, by definition of $\bar{\mathbf{q}}\in\bar{\mathcal{B}}(\mathbf{f})$,

$$\text{LHS of (A.22)}\ge 2m\alpha\,.$$

  On the other hand, we have for $K=\Omega(\frac{\bar{d}_\Sigma^2}{\bar{d}_g^2})$, we have $C_b=\Theta(1)$ hence

$$\text{RHS of (A.22)}=O\left(\frac{1}{K}\cdot\sqrt{m}\cdot K^{3/4}\cdot \bar{d}_{\min}^{-1}\bar{d}_g^{-3/2}\right)$$

  Here the RHS of (A.22) is maximized when $q_j=0$ or $q_j=d_j$. Therefore (A.22) holds for $K\ge K_1=\Omega\left(\max\left\{\frac{\bar{d}_\Sigma^2}{\bar{d}_g^2},\frac{1}{m^2\alpha^4\bar{d}_{\min}^4\bar{d}_g^6}\right\}\right)$. For $\mathbf{q}$ such that $\bar{\mathbf{q}}\in\bar{\mathcal{B}}(\mathbf{f})$ and $q_j=0$ or d_j for some $j'\in V$, we have

$$\text{LHS of (A.22)}\ge \alpha\sqrt{m}\cdot\Omega\left(K^{1/4}-\sqrt{\bar{d}_\Sigma}\right)\,,$$

  which is obtained by plugging in $q_{j'}$. For $K=\Omega(\frac{\bar{d}_\Sigma^2}{\bar{d}_g^2})$, we also have

$$\text{RHS of (A.22)}=O\left(\frac{1}{K}\cdot\sqrt{m}\cdot K^{3/4}\cdot \bar{d}_{\min}^{-1}\bar{d}_g^{-3/2}+1\right)\,,$$

  Using the analysis above, for $K\ge\Omega\left(\frac{m^2}{\bar{d}_{\min}^4\bar{d}_g^6}\right)$, the first term in the parentheses is

$O(1)$. In this case we have RHS of (A.22) $= O(1)$. Therefore (A.22) holds for

$$K = \Omega\left(\max\left\{\frac{1}{m^2\alpha^4\bar{d}_{\min}^4\bar{d}_g^6}, \frac{m^2}{\bar{d}_{\min}^4\bar{d}_g^6}, \frac{1}{\alpha^2\bar{d}_g^3}, \frac{\bar{d}_\Sigma^2}{\bar{d}_g^2}\right\}\right).$$

Combined, (A.22) holds for $K_1 = O\left(\max\left\{\frac{1}{m^2\alpha^4\bar{d}_{\min}^4\bar{d}_g^6}, \frac{m^2}{\bar{d}_{\min}^4\bar{d}_g^6}, \frac{1}{\alpha^2\bar{d}_g^3}, \frac{\bar{d}_\Sigma^2}{\bar{d}_g^2}\right\}\right)$.

- Point 2(b). Note that for $K = \Omega(\frac{\bar{d}_\Sigma^2}{\bar{d}_g^2})$,

$$\begin{aligned}
&\sup_{\mathbf{q},\mathbf{q}'\in\Omega^K}\left(F(\bar{\mathbf{q}}) - F(\bar{\mathbf{q}}')\right) \\
&\leq \max\{C_b, 1\}\cdot O\left(\sqrt{m}\sup_{\mathbf{q},\mathbf{q}'\in\Omega^K}\left(\sum_{j\in V}\sqrt{\bar{d}_j}\left(-\sqrt{\bar{q}_j} - \sqrt{\bar{d}_j - \bar{q}_j} + \sqrt{\bar{q}_j'} + \sqrt{\bar{d}_j - \bar{q}_j'}\right)\right)\right) \\
&\leq O\left(\sqrt{m}\max_{\mathbf{q}'\in\Omega^K}\sum_{j\in V}\sqrt{\bar{d}_j}\left(\sqrt{\bar{q}_j'} + \sqrt{\bar{d}_j - \bar{q}_j'}\right)\right) \\
&= O(m).
\end{aligned}$$

  Hence

$$M_1 = \text{poly}(m) = O(m).$$

- Point 2(c). Note that for $\bar{\mathbf{q}} \in \mathcal{B}_{\mathbf{f}}$, we have $\bar{q}_j = \Theta\left(\frac{\bar{d}_j}{\bar{d}_\Sigma}\right)$, hence

$$M_2 = \max_{\bar{\mathbf{q}}\in\mathcal{B}_{\mathbf{f}}}\max_{j\in V}|f'(\bar{q}_j)| \leq \frac{1}{\bar{d}_{\min}}\left(\frac{\bar{d}_\Sigma}{\bar{d}_g}\right)^{3/2} O(\sqrt{m}).$$

  For the special case where $V_{\mathrm{b}} = \emptyset$ hence $\bar{d}_j = 1$ for all $j \in V$, we have $\bar{d}_\Sigma = m$, $\bar{d}_{\min} = 1$, $\bar{d}_g = 1$ and $M_2 = O(m^2)$.

- Point 2(d). Note that for $\bar{\mathbf{q}} \in \mathcal{B}_{\mathbf{f}}$, we have $\bar{q}_j = \Theta\left(\frac{\bar{d}_j}{\bar{d}_\Sigma}\right)$, hence point 2(d) holds.

□

In the following (Lemma A.13), we verify Condition A.1 for two congestion functions other than the one given in (1.22).

Let $V_{\mathrm{b}} \subset V$ be the subset of buffer-constrained nodes with scaled buffer sizes $\bar{d}_j \in (0, 1)$, and define

- *Logarithmic congestion function.*

$$
\begin{aligned}
f_j(\bar{q}_j) &= c \cdot C_b \cdot \left(\log\left(\frac{\bar{q}_j}{\bar{d}_j}\right) - \log\left(1 - \frac{\bar{q}_j}{\bar{d}_j}\right) - D_b \right), && \forall j \in V_{\mathrm{b}} \\
f_j(\bar{q}_j) &= c \cdot \log \bar{q}_j \,, && \forall j \in V \backslash V_{\mathrm{b}} \qquad \text{(A.29)}
\end{aligned}
$$

Here C_b and D_b are normalizing constants chosen as follows. Define $\epsilon \triangleq \frac{\delta_K}{\tilde{K}}$ (where δ_K and $\tilde{K}$ were defined in (1.21)). Let $h_{\mathrm{b}}(\bar{q}) \triangleq \log \bar{q} - \log(1-\bar{q})$ and $h(\bar{q}) \triangleq \log \bar{q}$. Define $C_b \triangleq \frac{h(\epsilon) - h(1/\sum_{j\in V} \bar{d}_j)}{h_{\mathrm{b}}(\epsilon) - h_{\mathrm{b}}(1/\sum_{j\in V}\bar{d}_j)}$ and $D_b \triangleq h_{\mathrm{b}}(1/\sum_{j\in V}\bar{d}_j) - C_b^{-1} h(1/\sum_{j\in V}\bar{d}_j)$. Here $c = \Omega\left(\max\{\frac{1}{\alpha}, m\}\right)$.

- *Linear congestion function.*

$$
f_j(\bar{q}_j) = c \cdot \frac{\bar{q}_j}{\bar{d}_j}\,, \qquad \forall j \in V\,, \qquad \text{(A.30)}
$$

where $c = \Omega\left(\frac{\bar{d}_\Sigma}{\alpha}\right)$.

Lemma A.13. *Let $\bar{d}_\Sigma \triangleq \sum_{j\in V} \bar{d}_j$, $\bar{d}_g \triangleq \min\{1, \sum_{j\in V}\bar{d}_j - 1\}$ and $\bar{d}_{\min} \triangleq \min_{j\in V} \bar{d}_j$. The congestion functions* (A.29) *and* (A.30) *satisfy the growth conditions (Condition A.1) with parameters (α, K_1, M_1, M_2) where*

- *Logarithmic congestion function:*

$$
\begin{aligned}
K_1 =& C \cdot \max\left\{ \frac{c^2}{\bar{d}_{\min}^2 \bar{d}_g^2 m^2 \alpha^2}, \frac{c^2}{\bar{d}_{\min}^2 \bar{d}_g^2}, \frac{\bar{d}_\Sigma^2}{\bar{d}_{\min}^2 \bar{d}_g^2} \right\}, \\
M_1 =& C \cdot c \cdot \log m\,, \quad M_2 = C \cdot \frac{\bar{d}_\Sigma}{\bar{d}_{\min}} \cdot c
\end{aligned}
$$

for a universal constant $C > 0$.

- *Linear congestion function:*

$$
K_1 = C \cdot \max\left\{ \frac{c}{m\alpha}, \bar{d}_\Sigma^2 \right\}, \quad M_1 = C \cdot c\,, \quad M_2 = C \cdot m^2
$$

for a universal constant $C > 0$.

Proof of Lemma A.13. **Logarithmic function.** Point 1 in Condition A.1 is obvious. Now we verify the other points one by one:

- Point 1. It is not hard to see that the congestion functions $(f_j(\bar{q}_j))_{j\in V}$ are strictly increasing and continuously differentiable. For any $K > 0$, we have $f_j(\bar{q}_j) = f_k(\bar{q}_k)$ for any $j, k \in V$ if $q_j = q_k = 0$. As a result, if $q_j = 0$, we have $f_j(\bar{q}_j) \leq f_k(\bar{q}_k)$ for any $k \in V$. It can be easily verified that Point 1(b) is also satisfied by any $K > 0$.

It remains to be shown that there exists $K_1 < \infty$ such that for $K \geq K_1$, we have $f_j(\bar{q}_j) \leq f_k(\bar{q}_k)$ for any $j \in V$ if $q_k = d_k$ and $k \in V_{\mathrm{b}}$. To this end, if suffices to check the inequality $f_j(\bar{q}_j) \leq f_k(\bar{q}_k)$ for $q_j = d_j$, $q_k = d_k$ where $j \in V \backslash V_{\mathrm{b}}$ and $k \in V_{\mathrm{b}}$: In this case, we have $f_j(\bar{q}_j) \leq 0$; for $K = \Omega(\max\{\bar{d}_\Sigma^2, \frac{\bar{d}_\Sigma^2}{\bar{d}_g^2}\})$, we have $C_b = \Theta(1)$, $D_b = O(\log \frac{\bar{d}_\Sigma}{\bar{d}_g})$ hence $f_k(\bar{q}_k) = \Omega(c \cdot \log \frac{\sqrt{K}}{\bar{d}_g}) \geq 0$. Therefore point 1 is satisfied for $K_1 = O(\max\{\bar{d}_\Sigma^2, \frac{\bar{d}_\Sigma^2}{\bar{d}_g^2}\}) = O(\frac{\bar{d}_\Sigma^2}{\bar{d}_g^2})$.

- Point 2(a). For $\mathbf{q}$ such that $\bar{\mathbf{q}} \in \bar{\mathcal{B}}(\mathbf{f})$ and $0 < q_j < d_j$ for any $j \in V$, we have, by definition of $\bar{\mathbf{q}} \in \bar{\mathcal{B}}(\mathbf{f})$,

$$\text{LHS of (A.22)} \geq 2m\alpha\,.$$

On the other hand, we have for $K = \Omega(\frac{\bar{d}_\Sigma^2}{\bar{d}_g^2})$, we have $C_b = \Theta(1)$, hence

$$\text{RHS of (A.22)} = O\left(\frac{1}{K} \cdot c \cdot C_b \cdot \frac{\sqrt{K}}{\bar{d}_{\min}\bar{d}_g}\right).$$

Here the RHS of (A.22) is maximized when $q_j = 0$ or $q_j = d_j$. Therefore (A.22) holds for $K_1 = O\left(\max\left\{\frac{\bar{d}_\Sigma^2}{\bar{d}_g^2}, \frac{c^2}{\bar{d}_{\min}^2\bar{d}_g^2 m^2\alpha^2}\right\}\right)$.

For $\mathbf{q}$ such that $\bar{\mathbf{q}} \in \bar{\mathcal{B}}(\mathbf{f})$ and $q_j = 0$ or d_j for some $j' \in V$, we have

$$\text{LHS of (A.22)} \geq \alpha c \cdot \Omega\left(\frac{1}{2}\log K - \log(\bar{d}_{\min}\bar{d}_g) - \log \bar{d}_\Sigma\right),$$

which is obtained by plugging in $q_{j'}$. We also have

$$\text{RHS of (A.22)} \leq O\left(\frac{1}{K} \cdot c \cdot C_b \cdot \frac{\sqrt{K}}{\bar{d}_{\min}\bar{d}_g} + 1\right),$$

Using the analysis above, for $K = \Omega\left(\max\left\{\frac{\bar{d}_\Sigma^2}{\bar{d}_g^2}, \frac{c^2}{\bar{d}_{\min}^2\bar{d}_g^2}\right\}\right)$, the first term in the parentheses is $O(1)$. In this case we have RHS of (A.22) $\leq O(1)$. Therefore when $c = \Omega\left(\frac{1}{\alpha}\right)$, (A.22) holds for $K = \Omega\left(\left(\frac{\bar{d}_\Sigma}{\bar{d}_{\min}\bar{d}_g}\right)^2\right)$.

Combined, for $c = \Omega\left(\frac{1}{\alpha}\right)$, Point 2(a) holds for

$$K_1 = \Omega\left(\max\left\{\frac{c^2}{\bar{d}_{\min}^2 \bar{d}_g^2 m^2 \alpha^2}, \frac{c^2}{\bar{d}_{\min}^2 \bar{d}_g^2}, \frac{\bar{d}_\Sigma^2}{\bar{d}_{\min}^2 \bar{d}_g^2}\right\}\right).$$

- Point 2(b). Note that

$$\begin{aligned}
&\sup_{\mathbf{q},\mathbf{q}' \in \Omega^K} \left(F(\bar{\mathbf{q}}) - F(\bar{\mathbf{q}}')\right) \\
&= O\left(c \cdot \sup_{\mathbf{q},\mathbf{q}' \in \Omega^K} \left(\sum_{j \in V} \left(\bar{q}_j \log \bar{q}_j + (\bar{d}_j - \bar{q}_j)\log(\bar{d}_j - \bar{q}_j) - \bar{q}_j \log \bar{q}_j\right.\right.\right. \\
&\qquad \left.\left.\left. -(\bar{d}_j - \bar{q}'_j)\log(\bar{d}_j - \bar{q}'_j)\right)\right)\right) \\
&\le O\left(-c \cdot \min_{\mathbf{q}' \in \Omega^K} \sum_{j \in V} \left(\bar{q}'_j \log \bar{q}'_j + (\bar{d}_j - \bar{q}'_j)\log(\bar{d}_j - \bar{q}'_j)\right)\right) \\
&= O(c \cdot \log m),
\end{aligned}$$

where the inequality follows from the fact that $\bar{q}_j, \bar{d}_j - \bar{q}_j \in (0,1)$ hence $\bar{q}_j \log \bar{q}_j < 0$ and $(\bar{d}_j - \bar{q}_j)\log(\bar{d}_j - \bar{q}_j) < 0$. Hence

$$M_1 = \mathrm{poly}(c, m) = O(c \cdot \log m).$$

- Point 2(c). For $\bar{\mathbf{q}} \in \mathcal{B}_{\mathbf{f}}$, we have $\frac{\bar{d}_j}{\bar{d}_\Sigma} e^{-\frac{4m}{c \cdot C_b}} \le \bar{q}_j \le \frac{\bar{d}_j}{\bar{d}_\Sigma} e^{\frac{4m}{c \cdot C_b}}$. Choose $c \ge \frac{8m}{C_b} = \Omega(m)$, we have

$$M_2 = \max_{\bar{\mathbf{q}} \in \mathcal{B}_{\mathbf{f}}} \max_j |f'(\bar{q}_j)| \le \max_{\bar{\mathbf{q}} \in \mathcal{B}_{\mathbf{f}}} \max_{j \in V} |\bar{q}_j^{-1}| \le \mathrm{poly}(c, \bar{\mathbf{d}}) = \frac{\bar{d}_\Sigma}{\bar{d}_{\min}} \cdot O(c).$$

- Point 2(d). Note that $\bar{\mathbf{q}} \in \mathcal{B}_{\mathbf{f}}$, we have $\frac{\bar{d}_j}{\bar{d}_\Sigma} e^{-\frac{4m}{c \cdot C_b}} \le \bar{q}_j \le \frac{\bar{d}_j}{\bar{d}_\Sigma} e^{\frac{4m}{c \cdot C_b}}$. Given the choice of c derived in the last bullet point, we know point 2(d) holds.

Linear Function. Now we consider the linear congestion function.

- Point 1. It is easy to verify Point 1, therefore we omit the proof.

- Point 2(a). For $\mathbf{q}$ such that $\bar{\mathbf{q}} \in \bar{\mathcal{B}}(\mathbf{f})$ and $0 < q_j < d_j$ for any $j \in V$, we have, by definition of $\bar{\mathbf{q}} \in \bar{\mathcal{B}}(\mathbf{f})$,

$$\text{LHS of (A.22)} \geq 2m\alpha\,.$$

We also have

$$\text{RHS of (A.22)} \leq O\left(\frac{1}{K} \cdot c\right)$$

Therefore (A.22) holds for $K = \Omega\left(\frac{c}{m\alpha}\right)$. For $\mathbf{q}$ such that $\bar{\mathbf{q}} \in \bar{\mathcal{B}}(\mathbf{f})$ and $q_j = 0$ or d_j for some $j' \in V$, we have

$$\text{LHS of (A.22)} \geq \alpha c \cdot \Omega\left(\frac{1}{\bar{d}_\Sigma} - \frac{1}{\sqrt{K}}\right)\,,$$

which is obtained by plugging in $q_{j'}$. We also have

$$\text{RHS of (A.22)} \leq O\left(\frac{1}{K} \cdot c + 1\right)\,,$$

for $K = \Omega(c)$, the first term in the parenthese is $O(1)$, therefore we have RHS of (A.22) $= O\,(1)$. Choosing $c = \Omega\left(\frac{\bar{d}_\Sigma}{\alpha}\right)$, then (A.22) holds for $K = \Omega(\bar{d}_\Sigma^2)$.

Combined, when choosing $c = \Omega\left(\frac{\bar{d}_\Sigma}{\alpha}\right)$, Point 2(a) holds for $K_1 = O\left(\max\{\frac{c}{m\alpha}, \bar{d}_\Sigma^2\}\right)$.

- Point 2(b). Note that

$$\begin{aligned}
&\sup_{\mathbf{q},\mathbf{q}' \in \Omega^K} \left(F(\bar{\mathbf{q}}) - F(\bar{\mathbf{q}}')\right) \\
&= O\left(c \sup_{\mathbf{q},\mathbf{q}' \in \Omega^K} \left(\sum_{j \in V} \bar{d}_j^{-1}\left(\bar{q}_j^2 - (\bar{q}_j')^2\right)\right)\right) \\
&\leq O\left(c \max_{\mathbf{q}' \in \Omega^K} \sum_{j \in V} \bar{d}_j^{-1}\left(\bar{q}_j^2\right)\right) \\
&= O(c)\,.
\end{aligned}$$

Hence

$$M_1 = \text{poly}(c) = O(c)\,.$$

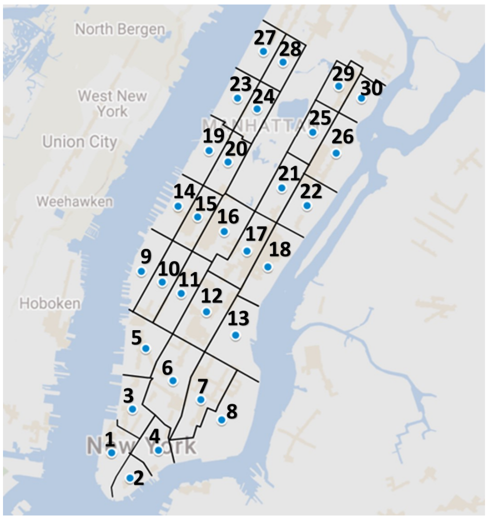

Figure A.1: A 30 location model of Manhattan below 110-th street, excluding the Central Park. (tessellation is based on [121])

- Point 2(c). For $\bar{\mathbf{q}} \in \mathcal{B}_{\mathbf{f}}$, we have $\bar{q}_j = \Theta\left(\frac{\bar{d}_j}{d_\Sigma}\right)$, hence

$$M_2 = \max_{\bar{\mathbf{q}} \in \mathcal{B}_{\mathbf{f}}} \max_{j \in V} |f'(\bar{q}_j)| = O(c)\,.$$

- Point 2(d). For $\bar{\mathbf{q}} \in \mathcal{B}_{\mathbf{f}}$, we have $\bar{q}_j = \Theta\left(\frac{\bar{d}_j}{d_\Sigma}\right)$. Hence point 2(d) holds.

□

A.4 Appendix to Section 1.7.1

In this section we provide a full description of our simulation environment and the benchmark we employ.

A.4.1 Simulation Setup and Benchmark Policies

Throughout the numerical experiments, we use the following model primitives.

- *Payoff structure.* In many scenarios, ride-hailing platforms take a commission proportional to the trip fare, which increases with trip distance/duration. Motivated by this, we present results for w_{ijk} set to be the travel time from j to[3] k.

[3] We tested a variety of payoff structures, and found that our results are robust to the choice of $\mathbf{w}$. One

• *Graph topology.* We consider a 30-location model of Manhattan below 110-th street excluding Central Park (see Figure A.1), as defined in Buchholz [121]. We let pairs of regions which share a non-trivial boundary be pickup compatible with each other, e.g., regions 23 and 24 are compatible but regions 23 and 20 are not.

• *Demand arrival process, and pickup/service times.* We consider a stationary demand arrival process, whose rate is the average decensored demand from 8 a.m. to 12 p.m. estimated in [121]. This period includes the morning rush hour and has significant imbalance of demand flow across geographical locations (for many customers the destination is in Midtown Manhattan).[4] We estimate travel times between location pairs using Google Maps.[5]

• *Number of cars, and steady state upper bound.*

— *Excess supply.* We use as a baseline the fluid requirement K_{fl} on number of cars needed to achieve optimal payoff. A simple workload conservation argument (using Little's Law) gives the fluid requirement as follows. Applying Little's Law, if the optimal solution $\mathbf{z}^*$ of SPP (A.1)-(A.3) is realized as the average long run assignment, the mean number of cars who are currently occupied, i.e. serving or picking up customers is $\sum_{j,k\in V}\sum_{i\in\mathcal{P}(j)} D_{ijk}\cdot z^*_{ijk}$, for $D_{ijk}\triangleq \tilde{D}_{ij}+\hat{D}_{jk}$, where $\tilde{D}_{ij}$ is the pickup time from i to j and $\hat{D}_{jk}$ is the travel time from j to k. In our case, it turns out that $K_{\text{fl}} = 7,307$. We use $1.05\times K_{\text{fl}}$ as the total number of cars in the system to study the excess supply case, i.e., there are 5% extra (idle) cars in the system beyond the number needed to achieve the W^{SPP} benchmark.

— *Scarce supply.* When the number of cars in the system is fewer than the fluid requirement, i.e., $K=\kappa K_{\text{fl}}$ for $\kappa<1$, no policy can achieve a steady state performance of

set of tests was to generate 100 random payoff vectors $\mathbf{w}$, with each w_{ijk} drawn i.i.d. from Uniform(0,1); we found that the results obtained are similar.

[4]We also simulated the MBP and greedy policy with time-varying demand arrival rates, where the demand arrival rate is estimated (from the real data) for every 5 min interval. Our MBP policy still significantly outperforms the greedy policy.

[5]We extract the pairwise travel time between region centroids (marked by the dots in Figure A.1) using Google Maps, denoted by $\hat{D}_{ij}$'s $(i,j=1,\cdots,30)$. We use $\hat{D}_{jk}$ as service time for customers traveling from j to k. For each customer at j who is picked up by a supply from i we add a pickup time [6] of $\tilde{D}_{ij}=\max\{\hat{D}_{ij}, 2\text{ minutes}\}$. The average travel time across all demand is 13.1 minutes, and the average pickup time is about 4 minutes (it is policy dependent).

W^{SPP}. A tighter upper bound on the steady state performance is then the value of the SPP (A.1)-(A.3) with the additional supply constraint

$$\sum_{j,k \in V} \sum_{i \in \mathcal{P}(j)} D_{ijk} \cdot z_{ijk} \le K\,.$$

We denote the value of this problem for $K = \kappa K_{\text{fl}}$ by $W^{\text{SPP}}(\kappa)$. We study the case $\kappa = 0.75$ as an example of scarce supply. For our simulation environment, it turns out that $W^{\text{SPP}}(0.75) \approx 0.86 W^{\text{SPP}}$, i.e., $0.86 W^{\text{SPP}}$ is an upper bound on the per period payoff achievable in steady state.

We compare the performance of our MBP-based policy against the following two policies:

1. *Static (fluid-based) policy.* The fluid-based policy is a static randomization based on the solution to the SPP, given exactly correct demand arrival rates [see, e.g., 1, 21]: Let $\mathbf{z}^*$ be a solution of SPP. When a type (j,k) demand arrives at location j, the randomized fluid-based policy dispatches from location $i \in \mathcal{P}(j)$ with probability z^*_{ijk}/ϕ_{jk}.
2. *Greedy non-idling policy.* For each demand type (j,k), the greedy policy dispatches from supply location i that has the highest payoff w_{ijk} among all compatible neighbors of j' which have at least one supply unit available. If there are ties (as is the case if the payoff w_{ijk} does not depend on i), the policy prefers a supply location with shorter pickup time.

A.4.2 The Excess Supply Case

We simulate the (stationary) system from 8 a.m. to 12 p.m. with 100 randomly generated initial states[7]. The simulation results on performance are shown in Figure A.2. The result confirms that the MBP policy significantly outperforms both the static policy

[7] We first uniformly sample 100 points from the simplex $\{\mathbf{q} : \sum_{i \in V} q_i = K\}$, which are used as the system's initial states at 6 a.m. (note that all the cars are free). Then we "warm-up" the system by employing the static policy from 6 a.m. to 8 a.m., assuming the demand arrival process during this period to be stationary (with the average demand arrival rate during this period as mean). Finally, we use the system's states at 8 a.m. as the initial states.

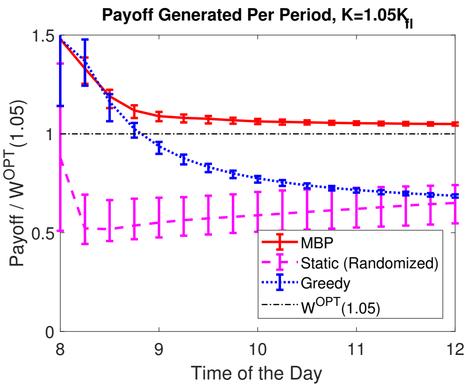

Figure A.2: Per period payoff under the MBP policy, static fluid-based policy and greedy policy (with 90% confidence intervals), relative to W^{SPP}.

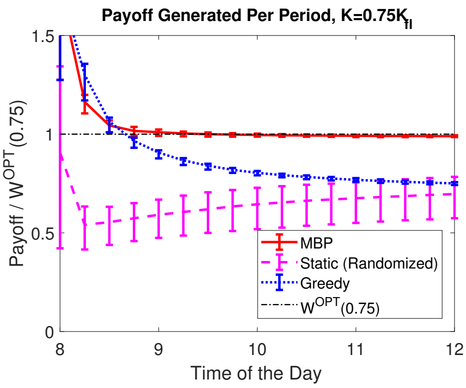

Figure A.3: Per period payoff under the modified MBP policy, static fluid-based policy and greedy policy (with 90% confidence intervals), relative to $W^{\mathrm{SPP}}(0.75)$, the value of SPP along with constraint (1.30) for $K = 0.75K_{\mathrm{fl}}$.

and the greedy policy: the average payoff under MBP over 4 hours is about 105% of W^{SPP} (here W^{SPP} is again an upper bound on the steady state performance[8]), while the static policy and greedy policy only achieve 65% and 68% of W^{SPP}, respectively. The performance of the static policy converges very slowly to W^{SPP}, leading to poor transient performance.[9] The performance of the greedy policy quickly deteriorates over time because it ignores the flow balance constraints and creates huge geographical imbalances in supply availability.

A.4.3 The Scarce Supply Case

In the scarce supply case, e.g., $K = 0.75K_{\mathrm{fl}}$, no policy can achieve a stationary performance of W^{SPP}; rather we have an steady state upper bound of $W^{\mathrm{SPP}}(0.75) \approx 0.86W^{\mathrm{SPP}}$. We use this as our benchmark.

Figure A.3 shows that the MBP policy also vastly outperforms the static policy and

[8] W^{SPP} is still an upper bound on stationary performance when pickup and service times are included in our model. However, in this case a transient upper bound is difficult to derive. As a result, we use the ratio of average per period payoff to W^{SPP} as a performance measure, with the understanding that it may exceed 1 at early times.

[9] For example, after running for 20 hours, and the average payoff generated by static policy in the 20-th hour is $0.96W^{\mathrm{SPP}}$.

greedy policy in the scarce supply case. MBP generates average per period payoff that is 99% of the benchmark $W^{\text{SPP}}(0.75)$ over 4 hours, while the static policy and greedy policy only achieves 69% and 74% resp. of the benchmark over the same period. Reassuringly, the mean value of $v(t)$ in our simulations of supply-aware MBP is within 10% of the optimal dual variable to the tightened supply constraint (1.31) in the SPP along with (1.31) (both values are close to 0.50). Again, we observe that the average performance of static policy improves (slowly) as the time horizon gets longer, while the performance of greedy deteriorates.

APPENDIX B

Proofs in "Dynamic Assignment Control of Closed Networks under Complete Resource Pooling"

This technical appendix is organized as follows.

- We prove our main result, Theorem 2.1, in Appendices B.1-B.4. In particular:
    - Appendix B.1 discusses fluid sample paths in detail and establishes key properties of our Lyapunov functions, including the proof of Lemma B.1.
    - Appendix B.2 includes the proof of Lemma 2.1, a converse bound on the demand loss exponent.
    - Appendix B.3 includes the proof of Proposition 2.4, containing sufficient conditions for a policy to achieve the optimal exponent.
    - Appendix B.4 shows that the SMW policy satisfies the sufficient conditions for exponent optimality, and derives explicitly the optimal exponent and most-likely sample paths, including the proofs of Lemma 2.2, Lemma 2.3, and Lemma 2.4. It also formally establishes exponent optimality of SMW policies for transient performance.
- Appendix B.5 includes the proof of Proposition 2.2 showing frequent utilization of supply units under SMW, and provides the structural corollaries (of Theorem 2.1) illustrated in Section 2.4.1.
- Appendix B.6 shows the necessity of the assumptions and state-dependent control,

including the proofs of Propositions B.1, 2.1 and 2.3, and the claim in Example 2.4.

- Appendix B.7 proves Theorem 2.3, the extension of our main result to scrip systems.
- Appendix B.8 proves Theorem 2.2, the extension of our main result to the shared transportation setting with travel delays.
- Appendix B.9 proves that the Assumption 3 in our paper is implied by the CRP condition defined in [17].
- Appendix B.10 provides the full description of our simulation experiments.

B.1 Lyapunov Functions and Fluid Sample Paths

B.1.1 Properties of the Lyapunov Functions

Scale-invariance and sub-additivity (about $\boldsymbol{\alpha}$)

Lemma B.1 (Key properties of $L_{\boldsymbol{\alpha}}(\cdot)$)**.** *For $L_{\boldsymbol{\alpha}}(\cdot)$ with $\alpha \in \mathrm{relint}(\Omega)$, we have:*

1. *Scale-invariance (about $\boldsymbol{\alpha}$). $L_{\boldsymbol{\alpha}}(\boldsymbol{\alpha}+c\Delta\mathbf{x}) = cL_{\boldsymbol{\alpha}}(\boldsymbol{\alpha}+\Delta\mathbf{x})$ for any $c>0$ and $\Delta\mathbf{x} \in \mathbb{R}^m$ such that $\mathbf{1}^{\mathrm{T}}\Delta\mathbf{x} = 0$ and $\boldsymbol{\alpha}+\Delta\mathbf{x} \in \Omega, \boldsymbol{\alpha}+c\Delta\mathbf{x} \in \Omega$.*
2. *Sub-additivity (about $\boldsymbol{\alpha}$). $L_{\boldsymbol{\alpha}}(\boldsymbol{\alpha}+\Delta\mathbf{x}+\Delta\mathbf{x}') \le L_{\boldsymbol{\alpha}}(\boldsymbol{\alpha}+\Delta\mathbf{x}) + L_{\boldsymbol{\alpha}}(\boldsymbol{\alpha}+\Delta\mathbf{x}')$ for any $\Delta\mathbf{x}, \Delta\mathbf{x}' \in \mathbb{R}^m$ such that $\mathbf{1}^{\mathrm{T}}\Delta\mathbf{x} = \mathbf{1}^{\mathrm{T}}\Delta\mathbf{x}' = 0$ and $\boldsymbol{\alpha}+\Delta\mathbf{x}+\Delta\mathbf{x}', \boldsymbol{\alpha}+\Delta\mathbf{x}, \boldsymbol{\alpha}+\Delta\mathbf{x}' \in \Omega$.*

Proof of Lemma B.1. (i) For $c>0$, $\boldsymbol{\alpha} \in \mathrm{relint}(\Omega)$, we have

$$L_{\boldsymbol{\alpha}}(\boldsymbol{\alpha}+c\Delta\mathbf{x}) = 1 - \min_i \frac{\alpha_i + c\Delta x_i}{\alpha_i} = -\min_i \frac{c\Delta x_i}{\alpha_i} = -c\min_i \frac{\Delta x_i}{\alpha_i} = cL_{\boldsymbol{\alpha}}(\boldsymbol{\alpha}+\Delta\mathbf{x})\,.$$

(ii) For $\boldsymbol{\alpha} \in \mathrm{relint}(\Omega)$, we have

$$\begin{aligned} L_{\boldsymbol{\alpha}}(\boldsymbol{\alpha}+\Delta\mathbf{x}+\Delta\mathbf{x}') &= 1 - \min_i \frac{\alpha_i + \Delta x_i + \Delta x_i'}{\alpha_i} = -\min_i \frac{\Delta x_i + \Delta x_i'}{\alpha_i} \\ \le\; -\min_i \frac{\Delta x_i}{\alpha_i} - \min_i \frac{\Delta x_i'}{\alpha_i} &= L_{\boldsymbol{\alpha}}(\boldsymbol{\alpha}+\Delta\mathbf{x}) + L_{\boldsymbol{\alpha}}(\boldsymbol{\alpha}+\Delta\mathbf{x}')\,. \end{aligned}$$

□

Regularity properties

The following lemma is a collection of regularity properties of $L_{\boldsymbol{\alpha}}(\mathbf{x})$ that are useful in the following proofs.

Lemma B.2. *For* $\boldsymbol{\alpha} \in \mathrm{relint}(\Omega)$ *and* $L_{\boldsymbol{\alpha}}(\mathbf{x})$ *specified in Definition 2.7, we have*

1. $L_{\boldsymbol{\alpha}}(\mathbf{x}) \geq 0$ *for all* $\mathbf{x} \in \Omega$*, and* $L_{\boldsymbol{\alpha}}(\mathbf{x}) = 0$ *if and only if* $\mathbf{x} = \boldsymbol{\alpha}$*.*

2. $L_{\boldsymbol{\alpha}}(\mathbf{x})$ *is globally Lipschitz on* Ω*, i.e. for any* $\mathbf{x}_1, \mathbf{x}_2 \in \Omega$*, we have*

$$|L_{\boldsymbol{\alpha}}(\mathbf{x}_1) - L_{\boldsymbol{\alpha}}(\mathbf{x}_2)| \leq \frac{1}{\min_i \alpha_i} ||\mathbf{x}_1 - \mathbf{x}_2||_\infty .$$

Proof of Lemma B.2. Property 1 is easy to verify hence we omit the proof.

For property 2, note that

$$|L_{\boldsymbol{\alpha}}(\mathbf{x}_1) - L_{\boldsymbol{\alpha}}(\mathbf{x}_2)| = \left| \min_i \frac{x_{1,i}}{\alpha_i} - \min_i \frac{x_{2,i}}{\alpha_i} \right| \leq \min_i \frac{|x_{1,i} - x_{2,i}|}{\alpha_i} \leq \frac{1}{\min_i \alpha_i} ||\mathbf{x}_1 - \mathbf{x}_2||_\infty .$$

□

B.1.2 Formal Definition of FSPs

We denote the correspondence from the given demand sample path and initial state to the uniquely determined state sample path by $\Psi^{K,U} : (\bar{\mathbf{A}}^K(\cdot), \bar{\mathbf{X}}^{K,U}(0)) \mapsto \bar{\mathbf{X}}^{K,U}(\cdot)$.

In this section, we discuss the existence of *fluid sample paths (FSPs)* and techniques related to FSP in large deviations analysis. FSP is a technique used to establish large deviation bounds of the *queue lengths* using the sample path large deviation principle of *demand arrival processes* (Fact 2.1), see, e.g., [122, 45].

We briefly comment on the existence of FSP. Consider a sequence of demand sample paths $\{\bar{\mathbf{A}}^K(\cdot)\}_{K=1}^\infty$ where in the K-th system the interarrival times of type (j', k) demand are deterministic with value $\frac{1}{K\hat{\phi}_{j'k}}$. It is trivial to show that $\{\bar{\mathbf{A}}^K(\cdot)\}_{K=1}^\infty$ converges uniformly on compact intervals (u.o.c.) to the fluid limit $\bar{\mathbf{A}}(t) = t\hat{\boldsymbol{\phi}}$. For any policy $U \in \mathcal{U}$, because at most one relocation happens at each demand arrival, each (normalized) queue length process $\bar{\mathbf{X}}^K(\cdot) = \Psi^{K,U}(\bar{\mathbf{A}}^K(\cdot), \bar{\mathbf{X}}^K(0))$ is Lipschitz continuous with Lipschitz constant $\mathbf{1}^{\mathrm{T}}\hat{\boldsymbol{\phi}}\mathbf{1}$, hence equicontinuous; see, for example, [123]. Thus, there must exist a subsequence of $\{\bar{\mathbf{X}}^K(\cdot)\}_{K=1}^{\infty}$ that converges u.o.c. to a continuous function $\bar{\mathbf{X}}(\cdot)$. Therefore $(\bar{\mathbf{A}}(\cdot), \bar{\mathbf{X}}(\cdot))$ is an FSP. This establishes the existence of FSP.

In the large-deviations literature, a technique named the "contraction principle" is often used to translate large deviations principles (LDP) for the arrival process to LDP for the state process, see [54]. The translation step is important in most of the large deviations analysis in the literature, including the one in this paper. However, to apply the contraction principle one needs to prove that the mapping from demand sample path $\bar{\mathbf{A}}(\cdot)$ to queue length sample path $\bar{\mathbf{X}}(\cdot)$ is continuous with respect to suitable topologies for the corresponding functional spaces. The continuity is usually technically challenging to establish (see [124] for an application of the contraction principle to MaxWeight policies under a different setting). The FSP technique partly circumvents this issue.

B.2 Converse Bound on the Exponent: Proof of Lemma 2.1

In this section, we prove Lemma 2.1, the converse bound on the exponent for any policy $U \in \mathcal{U}$. The proof consists of three steps:

- Step 1: For each stationary policy $U \in \mathcal{U}$ we define a state $\tilde{\boldsymbol{\alpha}} \in \mathrm{relint}(\Omega)$ such that the state visits the neighborhood of $\tilde{\boldsymbol{\alpha}}$ frequently enough. In the following steps we will bound the demand loss exponent of U by $\gamma_{\mathrm{CB}}(\tilde{\boldsymbol{\alpha}})$.

- Step 2: Given that the system's initial state is close to $\tilde{\boldsymbol{\alpha}}$, we construct a set of demand sample paths that are guaranteed to lead to a demand loss regardless of the policy used. To this end, we compute $v_{\tilde{\boldsymbol{\alpha}}}(\mathbf{f})$, which the minimum rate of increase of $L_{\tilde{\boldsymbol{\alpha}}}(\cdot)$ under demand arrival rates $\mathbf{f}$ no matter the assignment distributions. This step is used to lower bound the "one-shot" probability of demand-loss.

- Step 3: We use renewal-reward theorem to translate the one-shot demand loss probability to steady-state demand loss probability. The final bound in (2.21) takes the supremum over $\boldsymbol{\alpha}$ since the policy can choose its resting state.

The technique used in step 2 follows from Proposition 9 in [45]. The approach in steps 1 and 3 is novel (to the best of our knowledge) and tackles the key challenge of our closed network model, i.e., the policy has the flexibility to choose a resting state, as opposed to open network settings where the resting state is always $\mathbf{0}$.

Proof of Lemma 2.1. Step 1: Find the "frequently visited" state $\boldsymbol{\alpha}$. Fix a stationary policy $U \in \mathcal{U}$. For each K, the K-th system under policy U is a finite-state Markov chain, whose state space has cardinality smaller than K^m. Since we are considering the optimistic exponent, let the K-th system start within a communication class that minimizes steady state demand loss among all communication classes. Denote the stationary distribution (henceforth it refers to the stationary distribution of the communication class where the initial state belongs to) of (normalized) states as $\pi^K(\bar{\mathbf{X}}^K)$. Then there must exist a (normalized) state $\tilde{\mathbf{X}}^K$ such that $\pi^K(\tilde{\mathbf{X}}^K) \geq K^{-m}$. Take a subsequence $\{K_r\}$ of $\{K\}$ such that

$$\lim_{r\to\infty} \frac{1}{K_r} \log \mathbb{P}_{\text{o}}^{K_r,U} = \liminf_{K\to\infty} \frac{1}{K} \log \mathbb{P}_{\text{o}}^{K,U} .$$

By compactness of Ω, there must exist a further subsequence of $\{K_r\}$, which we denote by $\{K_{r'}\}$, and $\boldsymbol{\alpha} \in \Omega$ such that $\lim_{r'\to\infty} \tilde{\mathbf{X}}^{K_{r'}} = \boldsymbol{\alpha}$.

For any $0 < \epsilon_1 < \frac{1}{2}\left(\min_{j:\alpha_j>0} \alpha_j\right)$, define $\tilde{\boldsymbol{\alpha}} \in \operatorname{relint}(\Omega)$ such that

$$0 < \tilde{\alpha}_j < \epsilon_1/2 \quad \text{for } j \text{ such that } \alpha_j = 0\,,$$
$$|\tilde{\alpha}_j - \alpha_j| < \epsilon_1/2 \quad \text{for } j \text{ such that } \alpha_j > 0\,.$$

Since $\boldsymbol{\alpha}$ is the limit point of $\tilde{\mathbf{X}}^{K_{r'}}$, there exists $r_0'(\epsilon) > 0$ such that $\forall r' \geq r_0'(\epsilon)$,

$$0 \leq \tilde{X}_j^{K_{r'}} < \tilde{\alpha}_j \quad \text{for } j \text{ such that } \alpha_j = 0\,, \tag{B.1}$$
$$|\tilde{X}_j^{K_{r'}} - \alpha_j| < \epsilon_1/2 \quad \text{for } j \text{ such that } \alpha_j > 0\,. \tag{B.2}$$

Inequalities (B.1) and (B.2) imply that for $r' \geq r_0'(\epsilon)$

$$|\tilde{X}_j^{K_{r'}} - \tilde{\alpha}_j| \leq \ \tilde{\alpha}_j < \epsilon_1, \quad \text{for } j \text{ such that } \alpha_j = 0$$

$$|\tilde{X}_j^{K_{r'}} - \tilde{\alpha}_j| \leq \; |\tilde{X}_j^{K_{r'}} - \alpha_j| + |\tilde{\alpha}_j - \alpha_j| < \epsilon_1, \quad \text{for } j \text{ such that } \alpha_j > 0\,.$$

Hence $||\tilde{\mathbf{X}}^{K_{r'}} - \tilde{\boldsymbol{\alpha}}||_\infty < \epsilon_1$ for $r' \geq r_0'(\epsilon)$.

We quantify the fact that $\tilde{\boldsymbol{\alpha}}$ is a "frequently visited" state in the following claim.

Claim: Fix $K = K_{r'}$ that comes from the subsequence defined above. In the K-th system, define

$$\tau^K \triangleq \inf\left\{t > 0 : \bar{\mathbf{X}}^K(t) = \tilde{\mathbf{X}}^K | \bar{\mathbf{X}}^K(0) = \tilde{\mathbf{X}}^K\right\}\,, \tag{B.3}$$

then we have

$$\mathbb{E}[\tau^K] \leq \; \frac{K^m}{\mathbf{1}^{\mathrm{T}}\hat{\boldsymbol{\phi}}\mathbf{1}}\,.$$

Proof of claim: Consider the discrete-time embedded chain of $\{\bar{\mathbf{X}}^K(\cdot)\}$. Since the initial state $\tilde{\mathbf{X}}^K$ is positive recurrent within its communication class, the expected number of jumps between two consecutive visits to $\tilde{\mathbf{X}}^K$ is inversely proportional to its steady state measure $\pi^K(\tilde{\mathbf{X}}^K)$. By definition of $\tilde{\mathbf{X}}^K$, the expected number of jumps must be no larger than K^m. Since the time between two jumps are i.i.d. exponential variables with mean $(\mathbf{1}^{\mathrm{T}}\boldsymbol{\phi}\mathbf{1})^{-1}$, this concludes the proof.

Step 2: Lower bound on the "one-shot" demand-loss probability. Fix $K_{r'}$ and a demand sample path $\bar{\mathbf{A}}^{K_{r'}}(\cdot)$. For $t > 0$, define $f_{j'k}(t) \triangleq \frac{1}{t}\bar{\mathbf{A}}^{K_{r'}}(t)$, i.e. the average arrival rate of type (j', k) demand during $[0, t]$. For stationary policy U, denote the average fraction of demand arriving at j' that is served by supply at i during this period as $d_{ij'}^U(t)$ (we omit the superscript U in the following for notational simplicity). For $t \geq 0$, if $\bar{\mathbf{X}}^{K_{r'}}(0) = \tilde{\mathbf{X}}^{K_{r'}}$ and no demand is lost prior to t, we have for any $i \in V_S$

$$\bar{X}_i^{K_{r'}}(t) - \tilde{X}_i^{K_{r'}} = t\left(\sum_{j' \in V_D} f_{j'i}(t) - \sum_{j' \in \partial(i)} d_{ij'}(t)\left(\sum_{k \in V_S} f_{j'k}(t)\right)\right)\,.$$

Since $\tilde{\alpha}_j > 0$ for any $j \in V_S$, the Lyapunov function $L_{\tilde{\boldsymbol{\alpha}}}(\cdot)$ is well-defined. Evaluate the

Lyapunov function at $\left(\tilde{\boldsymbol{\alpha}} + \bar{\mathbf{X}}^{K_{r'}}(t) - \tilde{\mathbf{X}}^{K_{r'}}\right)$, we have:

$$\begin{aligned}
& L_{\tilde{\boldsymbol{\alpha}}}\left(\tilde{\boldsymbol{\alpha}} + \bar{\mathbf{X}}^{K_{r'}}(t) - \tilde{\mathbf{X}}^{K_{r'}}\right) \qquad \text{(B.4)} \\
&= L_{\tilde{\boldsymbol{\alpha}}}\left(\tilde{\boldsymbol{\alpha}} + t\left(\sum_{j' \in V_D} f_{j'i}(t) - \sum_{j' \in \partial(i)} d_{ij'}(t)\left(\sum_{k \in V_S} f_{j'k}(t)\right)\right)_{i \in V_S}\right) \\
&\overset{(a)}{=} tL_{\tilde{\boldsymbol{\alpha}}}\left(\tilde{\boldsymbol{\alpha}} + \left(\sum_{j' \in V_D} f_{j'i}(t) - \sum_{j' \in \partial(i)} d_{ij'}(t)\left(\sum_{k \in V_S} f_{j'k}(t)\right)\right)_{i \in V_S}\right) \\
&\geq t \min_{\Delta\mathbf{x} \in \mathcal{X}_{\mathbf{f}}} L_{\tilde{\boldsymbol{\alpha}}}(\tilde{\boldsymbol{\alpha}} + \Delta\mathbf{x}). \qquad \text{(B.5)}
\end{aligned}$$

Equality (a) holds because the Lyapunov function is scale-invariant with respect to $\tilde{\boldsymbol{\alpha}}$. Here $\Delta\mathbf{x}$ is the change of (normalized) state in unit time given average demand arrival rate during this period $\mathbf{f}$, and $\mathcal{X}_{\mathbf{f}}$ is defined in (2.20).

Define $v_{\tilde{\boldsymbol{\alpha}}}(\mathbf{f}) \triangleq \min_{\Delta\mathbf{x} \in \mathcal{X}_{\mathbf{f}}} L_{\tilde{\boldsymbol{\alpha}}}(\tilde{\boldsymbol{\alpha}} + \Delta\mathbf{x})$, which is the minimum rate the Lyapunov function increases under any policy, given demand arrival rate $\mathbf{f}$. Now we construct a set of demand sample paths that must lead to demand loss before the system returns to the starting state. First note that $\{\mathbf{f} : v_{\tilde{\boldsymbol{\alpha}}}(\mathbf{f}) > 0\}$ is non-empty. To see this, let $f'_{j'k}$ equal to 1 for some j' and $k \notin \partial(j')$, and 0 otherwise (such a pair (j', k) exists by Assumption 2.2). This $\mathbf{f}'$ results in a strictly positive[1] $v_{\tilde{\boldsymbol{\alpha}}}(\mathbf{f}')$. Therefore for any $\epsilon_2 > 0$ there exists demand arrival rate $\tilde{\mathbf{f}}$ such that

$$v_{\tilde{\boldsymbol{\alpha}}}(\tilde{\mathbf{f}}) > 0 \qquad \text{and} \qquad \frac{\Lambda^*(\tilde{\mathbf{f}})}{v_{\tilde{\boldsymbol{\alpha}}}(\tilde{\mathbf{f}})} \leq \inf_{\mathbf{f}: v_{\tilde{\boldsymbol{\alpha}}}(\mathbf{f}) > 0} \frac{\Lambda^*(\mathbf{f})}{v_{\tilde{\boldsymbol{\alpha}}}(\mathbf{f})} + \epsilon_2.$$

It is not hard to show that $v_{\tilde{\boldsymbol{\alpha}}}(\mathbf{f})$ is continuous in $\mathbf{f}$, hence there exists $\epsilon_3 > 0$ such that for any $\hat{\mathbf{f}} : ||\hat{\mathbf{f}} - \tilde{\mathbf{f}}||_\infty < \epsilon_3$, we have

$$v_{\tilde{\boldsymbol{\alpha}}}(\hat{\mathbf{f}}) > (1 - \epsilon_2) v_{\tilde{\boldsymbol{\alpha}}}(\tilde{\mathbf{f}}) > 0\,.$$

[1] To see this, notice that $L_{\tilde{\boldsymbol{\alpha}}}(\mathbf{x}) > 0$ for any $\mathbf{x} \in \Omega \backslash \{\tilde{\boldsymbol{\alpha}}\}$, hence it suffices to show that $\mathbf{0} \notin \mathcal{X}_{\mathbf{f}'}$. Because for any $\Delta\mathbf{x} \in \mathcal{X}_{\mathbf{f}'}$, we have $\Delta x_k = f'_{j'k} > 0$, hence $\mathbf{0} \notin \mathcal{X}_{\mathbf{f}'}$. This concludes the proof.

Denote $T \triangleq \frac{1+\frac{\epsilon_1}{\min_{j:\alpha_j>0}\alpha_j}}{(1-\epsilon_2)v_{\tilde{\boldsymbol{\alpha}}}(\tilde{\mathbf{f}})}$, define

$$\mathcal{B}_{\tilde{\boldsymbol{\alpha}}} \triangleq \left\{ \bar{\mathbf{A}}(\cdot) \in C\left[0,T\right] \middle| \sup_{t\in[0,T]} ||\bar{\mathbf{A}}(t) - t\tilde{\mathbf{f}}||_\infty \leq \epsilon_3 \right\}.$$

For any demand arrival sample path $\bar{\mathbf{A}}(\cdot) \in \mathcal{B}_{\tilde{\boldsymbol{\alpha}}}$, we will show that for $t \in [0,T]$ the followings are true: (i) normalized state $\bar{\mathbf{X}}^{K_{r'}}(t)$ does not hit $\tilde{\mathbf{X}}^{K_{r'}}$ before any demand is lost; (ii) at least one demand is lost.

To prove (i), define function $\tilde{L}_{\tilde{\boldsymbol{\alpha}}}(\bar{\mathbf{X}}) \triangleq L_{\tilde{\boldsymbol{\alpha}}}\left(\tilde{\boldsymbol{\alpha}} + \bar{\mathbf{X}} - \tilde{\mathbf{X}}^{K_{r'}}\right)$. By definition, we have $L_{\tilde{\boldsymbol{\alpha}}}(\mathbf{x}) > 0$ for any $\mathbf{x} \in \{\mathbf{x} \in \mathbb{R}^m : \mathbf{1}^{\mathrm{T}}\mathbf{x} = 1\}\backslash\{\tilde{\boldsymbol{\alpha}}\}$, hence we have that $\tilde{L}_{\tilde{\boldsymbol{\alpha}}}(\bar{\mathbf{X}}) > 0$ for any $\bar{\mathbf{X}} \in \Omega\backslash\{\tilde{\mathbf{X}}^{K_{r'}}\}$. By inequality (B.5), if no demand is lost during $[0,T]$ we have:

$$\tilde{L}_{\tilde{\boldsymbol{\alpha}}}\left(\bar{\mathbf{X}}^{K_{r'}}(t)\right) \geq tv\left(\frac{1}{t}\bar{\mathbf{A}}(t)\right) \geq t \min_{\bar{\mathbf{A}}(\cdot)\in\mathcal{B}} v\left(\frac{1}{t}\bar{\mathbf{A}}(t)\right) > t(1-\epsilon_2)v_{\tilde{\boldsymbol{\alpha}}}(\tilde{\mathbf{f}}) > 0.$$

We prove (ii) by contradiction. Suppose no demand is lost given (fluid scale) demand arrival sample path $\bar{\mathbf{A}}(\cdot) \in \mathcal{B}$, then

$$\tilde{L}_{\tilde{\boldsymbol{\alpha}}}\left(\bar{\mathbf{X}}^{K_{r'}}(T)\right) \geq T \min_{\bar{\mathbf{A}}(\cdot)\in\mathcal{B}} v\left(\frac{1}{T}\bar{\mathbf{A}}(T)\right) > \frac{1+\frac{\epsilon_1}{\min_{j:\alpha_j>0}\alpha_j}}{(1-\epsilon_2)v_{\tilde{\boldsymbol{\alpha}}}(\tilde{\mathbf{f}})}(1-\epsilon_2)v_{\tilde{\boldsymbol{\alpha}}}(\tilde{\mathbf{f}}) = 1 + \frac{\epsilon_1}{\min_{j:\alpha_j>0}\alpha_j}.$$

Expand the expression of $\tilde{L}_{\tilde{\boldsymbol{\alpha}}}\left(\bar{\mathbf{X}}^{K_{r'}}(T)\right)$ on the LHS, we have

$$1 - \min_j \frac{\bar{X}_j^{K_{r'}}(T) + \left(\tilde{\alpha}_j - \tilde{x}_j^{K_{r'}}\right)}{\tilde{\alpha}_j} > 1 + \frac{\epsilon_1}{\min_{j:\alpha_j>0}\alpha_j}.$$

Therefore

$$\begin{aligned}\min\left\{\min_{j:\alpha_j=0}\frac{\bar{X}_j^{K_{r'}}(T)}{\tilde{\alpha}_j}, \min_{j:\alpha_j>0}\frac{\bar{X}_j^{K_{r'}}(T)-\epsilon_1/2}{\tilde{\alpha}_j}\right\} &\leq \min_j \frac{\bar{X}_j^{K_{r'}}(T) + \left(\tilde{\alpha}_j - \tilde{x}_j^{K_{r'}}\right)}{\tilde{\alpha}_j} \\ &< -\frac{\epsilon_1}{\min_{j:\alpha_j>0}\alpha_j}. \end{aligned} \tag{B.6}$$

Note that the first inequality in (B.6) holds because of (B.1) and (B.2). Inequality (B.6) implies that $\min_j \bar{X}_j^{K_{r'}}(T) < 0$, which is impossible as queue lengths must be non-negative.

Step 3: Asymptotic steady-state lower bound on demand loss probability. We use renewal-reward theorem [see, e.g., 125] to lower bound the demand-loss probability. Consider the regenerative process that restarts each time $\bar{\mathbf{X}}^{K_{r'}}(t) = \tilde{\mathbf{X}}^{K_{r'}}$. Without loss of generality, let $\bar{\mathbf{X}}^{K_{r'}}(0) = \tilde{\mathbf{X}}^{K_{r'}}$. Recall the definition of τ^K in (B.3). Using the claim in step 1 and the result in step 2, we have:

$$
\begin{aligned}
\mathbb{P}_{\mathrm{o}}^{K_{r'},U} &= \frac{\mathbb{E}\left[\#\{\text{demand lost during } [0,\tau]\}\right]}{\mathbb{E}[\tau]} \\
&\geq {K_{r'}}^{-m}(\mathbf{1}^{\mathrm{T}}\hat{\boldsymbol{\phi}}\mathbf{1})\mathbb{E}\left[\#\{\text{demand lost during } [0,\tau]\}\right] \\
&\geq {K_{r'}}^{-m}(\mathbf{1}^{\mathrm{T}}\hat{\boldsymbol{\phi}}\mathbf{1})\mathbb{P}\left(\#\{\text{demand lost during } [0,\tau]\} \geq 1\right) \\
&\geq {K_{r'}}^{-m}(\mathbf{1}^{\mathrm{T}}\hat{\boldsymbol{\phi}}\mathbf{1})\mathbb{P}\left(\bar{\mathbf{A}}^{K_{r'}}(\cdot) \in \mathcal{B}_{\tilde{\boldsymbol{\alpha}}}\right).
\end{aligned}
$$

Take asymptotic limit on both sides, we have:

$$
\begin{aligned}
\liminf_{r'\to\infty} \frac{1}{K_{r'}} \log \mathbb{P}_{\mathrm{o}}^{K_{r'},U} &\geq \liminf_{r'\to\infty} \frac{1}{K_{r'}} \log \mathbb{P}\left(\bar{\mathbf{A}}^{K_{r'}}(\cdot) \in \mathcal{B}_{\tilde{\boldsymbol{\alpha}}}\right) \\
&\overset{(a)}{\geq} - \inf_{\bar{\mathbf{A}}(\cdot)\in\mathcal{B}^{\mathrm{o}}_{\tilde{\boldsymbol{\alpha}}}\cap \mathrm{AC}[0,T]} \int_0^T \Lambda^*\left(\dot{\bar{\mathbf{A}}}(t)\right) dt \\
&\overset{(b)}{\geq} -T\Lambda^*(\tilde{\mathbf{f}}) \\
&= -\frac{1+\frac{\epsilon_1}{\min_{j:\alpha_j>0}\alpha_j}}{(1-\epsilon_2)v_{\tilde{\boldsymbol{\alpha}}}(\tilde{\mathbf{f}})}\Lambda^*(\tilde{\mathbf{f}}) \\
&\geq -\frac{1+\frac{\epsilon_1}{\min_{j:\alpha_j>0}\alpha_j}}{1-\epsilon_2}\left(\inf_{\mathbf{f}:v_{\tilde{\boldsymbol{\alpha}}}(\mathbf{f})>0}\frac{\Lambda^*(\mathbf{f})}{v_{\tilde{\boldsymbol{\alpha}}}(\mathbf{f})}+\epsilon_2\right).
\end{aligned}
$$

Here (a) holds because of Mogulskii's Theorem (Fact 2.1), (b) holds because demand sample path $\bar{\mathbf{A}}(t) = t\tilde{\mathbf{f}} \in \mathrm{AC}[0,T]$ is a member of $\mathcal{B}_{\tilde{\boldsymbol{\alpha}}}$. For any $\delta > 0$, by choosing small enough $\epsilon_1(\delta), \epsilon_2(\delta) > 0$, we have

$$
-\liminf_{r'\to\infty} \frac{1}{K_{r'}} \log \mathbb{P}_{\mathrm{o}}^{K_{r'},U} \leq (1+\delta)(\gamma_{\mathrm{CB}}(\tilde{\boldsymbol{\alpha}}(\delta)) + \delta).
$$

Here the choice of $\tilde{\boldsymbol{\alpha}}$ depends on δ. To get rid of the multiplicative term $(1+\delta)$, it suffices to show that $\sup_{\boldsymbol{\alpha}\in\mathrm{relint}(\Omega)} \gamma_{\mathrm{CB}}(\boldsymbol{\alpha}) < \infty$. This can be proved by the following construction: let $\bar{\mathbf{A}}(t) = t\mathbf{f}'$ for $t \in [0,1]$ where $f_{j'k} = 1$ for some $j' \in V_D$ and $k \notin \partial(j')$. Because $\gamma_{\mathrm{CB}}(\boldsymbol{\alpha})$ is defined by an infimum $\gamma_{\mathrm{CB}}(\boldsymbol{\alpha}) \triangleq \inf_{\mathbf{f}\in\mathbb{R}_+^{nm}:v_\alpha(\mathbf{f})>0} \frac{\Lambda^*(\mathbf{f})}{v_\alpha(\mathbf{f})}$, we have $\gamma_{\mathrm{CB}}(\boldsymbol{\alpha}) \leq \frac{\Lambda^*(\mathbf{f}')}{v_{\tilde{\boldsymbol{\alpha}}}(\mathbf{f}')}$. By definition, $v_{\tilde{\boldsymbol{\alpha}}}(\mathbf{f}') = 1 - \max_{\Delta\mathbf{x}\in\mathcal{X}_{\mathbf{f}'}} \min_i \frac{\tilde{\alpha}_i+\Delta x_i}{\tilde{\alpha}_i} = -\max_{\Delta\mathbf{x}\in\mathcal{X}_{\mathbf{f}'}} \min_i \frac{\Delta x_i}{\tilde{\alpha}_i}$. Note that

$$\mathcal{X}_{\mathbf{f}'} = \{\Delta\mathbf{x} \in \mathbb{R}^{|V_S|} : \sum_{i\in\partial(j')} \Delta x_i = -1\,, \Delta x_i \le 0 \text{ for } i \in \partial(j')\,, \Delta x_k = 1$$
$$\Delta x_i = 0 \text{ for } i \notin \partial(j') \cup \{k\}\}\,.$$

Therefore

$$\max_{\Delta\mathbf{x}\in\mathcal{X}_{\mathbf{f}'}} \min_{i\in V_S} \frac{\Delta x_i}{\tilde{\alpha}_i} = \max_{\Delta\mathbf{x}\in\mathcal{X}_{\mathbf{f}'}} \min_{i\in\partial(j')} \frac{\Delta x_i}{\tilde{\alpha}_i} \le \max_{\Delta\mathbf{x}\in\mathcal{X}_{\mathbf{f}'}} \min_{i\in\partial(j')} \Delta x_i \le -\frac{1}{|\partial(j')|} \le -\frac{1}{m}\,.$$

Hence $v_{\tilde{\boldsymbol{\alpha}}}(\mathbf{f}') \ge \frac{1}{m}$. Hence $\gamma_{\text{CB}}(\alpha) \le \frac{\Lambda^*(\mathbf{f}')}{v_{\tilde{\boldsymbol{\alpha}}}(\mathbf{f}')} \le m\Lambda^*(\mathbf{f}') < \infty$. Therefore by choosing a small enough δ, we have

$$-\liminf_{r'\to\infty} \frac{1}{K_{r'}} \log \mathbb{P}_{\text{o}}^{K_{r'},U} \le \gamma_{\text{CB}}(\tilde{\boldsymbol{\alpha}}(\epsilon)) + \epsilon.$$

By the definition of subsequence $\{K_{r'}\}$, we have

$$-\liminf_{K\to\infty} \frac{1}{K} \log \mathbb{P}_{\text{o}}^{K,U} \le \gamma_{\text{CB}}(\tilde{\boldsymbol{\alpha}}(\epsilon)) + \epsilon.$$

As a result, for any $\epsilon > 0$ there exists $\alpha \in \Omega$ such that $-\liminf_{K\to\infty} \frac{1}{K} \log \mathbb{P}_{\text{o}}^{K,U} \le \sup_{\boldsymbol{\alpha}\in\text{relint}(\Omega)} \gamma_{\text{CB}}(\boldsymbol{\alpha}) + \epsilon$, therefore $-\liminf_{K\to\infty} \frac{1}{K} \log \mathbb{P}_{\text{o}}^{K,U} \le \sup_{\boldsymbol{\alpha}\in\text{relint}(\Omega)} \gamma_{\text{CB}}(\boldsymbol{\alpha})$. □

B.3 Sufficient Conditions for Exponent Optimality: Proof of Proposition 2.4

The proof of Proposition 2.4 consists of two parts. We first derive an achievability bound for policies that, for a given $\boldsymbol{\alpha} \in \text{relint}(\Omega)$, satisfy the negative drift property in Proposition 2.4; we then show it matches the converse bound in Lemma 2.1 for that specific $\boldsymbol{\alpha}$ (i.e., $\gamma_{\text{CB}}(\boldsymbol{\alpha})$) if the steepest descent property in Proposition 2.4 is also satisfied.

The full proof of Proposition 2.4 is quite technical, but the key idea is straightforward. Given starting state $\boldsymbol{\alpha}$, the (i) steepest descent property of U and (ii) the scale-invariance and sub-additivity of $L_{\boldsymbol{\alpha}}(\cdot)$, together ensure that the speed at which $L_{\boldsymbol{\alpha}}(\cdot)$ increases under U cannot exceed the minimum speed $v_{\boldsymbol{\alpha}}(\mathbf{f})$ in the converse construction (Lemma 2.1) for

$\mathbf{f} \triangleq \dot{\mathbf{A}}(t)$. Mathematically,

$$\begin{aligned}
&\dot{L}_{\boldsymbol{\alpha}}(\bar{\mathbf{X}}^U(t))\Big|_{\dot{\mathbf{A}}(t)=\mathbf{f}} \\
&= \inf_{U' \in \mathcal{U}_{\mathrm{ni}}} \left\{ \dot{L}_{\boldsymbol{\alpha}}(\bar{\mathbf{X}}^{U'}(t)) \,\Big|\, \dot{\mathbf{A}}(t) = \mathbf{f} \right\} && \text{(steepest descent)} \\
&= \min_{\Delta\mathbf{x} \in \mathcal{X}_{\mathbf{f}}} \lim_{\Delta t \to 0} \frac{L_{\boldsymbol{\alpha}}(\bar{\mathbf{X}}^U(t) + \Delta\mathbf{x}\Delta t) - L_{\boldsymbol{\alpha}}(\bar{\mathbf{X}}^U(t))}{\Delta t} && \text{(definition of } \mathcal{X}_{\mathbf{f}}) \\
&\leq \min_{\Delta\mathbf{x} \in \mathcal{X}_{\mathbf{f}}} \lim_{\Delta t \to 0} \frac{L_{\boldsymbol{\alpha}}(\boldsymbol{\alpha} + \Delta\mathbf{x}\Delta t)}{\Delta t} && \text{(sub-additivity of } L_{\boldsymbol{\alpha}}\text{, Lemma B.1)} \\
& && \text{(B.7)} \\
&= \min_{\Delta\mathbf{x} \in \mathcal{X}_{\mathbf{f}}} L_{\boldsymbol{\alpha}}(\boldsymbol{\alpha} + \Delta\mathbf{x}) = v_{\boldsymbol{\alpha}}(\mathbf{f})\,. && \text{(scale-invariance of } L_{\boldsymbol{\alpha}}\text{, Lemma B.1)}
\end{aligned}$$

As a result, the demand loss exponent under U is no worse than $\gamma_{\mathrm{CB}}(\boldsymbol{\alpha})$.

B.3.1 An achievability bound

The following lemma is an adaptation of Theorem 5 and Proposition 7 in [45] to our setting. It gives the achievability bound for the exponent of the steady state demand-loss probability, for any policy such that the negative drift condition in Proposition 2.4 is met for $L_{\boldsymbol{\alpha}}(\cdot)$ where $\boldsymbol{\alpha} \in \mathrm{relint}(\Omega)$. The main technical difficulty comes from the fact that it characterizes the *steady state* of the system. The analysis uses Freidlin-Wentzell theory and follows from [47, 45]. While the main proof idea follows that in [45], we refine the results there by dropping the assumption that all FSPs are Lipschitz continuous with a universal Lipschitz constant. This allows us to deal with Poisson-driven demand arrival processes which does not satisfy this assumption.

Lemma B.3 (Achievability bound)**.** *For the system being considered, if policy U satisfies the negative drift condition in Proposition 2.4 for $L_{\boldsymbol{\alpha}}(\cdot)$ where $\boldsymbol{\alpha} \in \mathrm{relint}(\Omega)$, we have (the subscript "AB" stands for achievability bound)*

$$-\limsup_{K \to \infty} \frac{1}{K} \log \mathbb{P}_{\mathrm{p}}^{K,U} \geq \gamma_{\mathrm{AB}}(\boldsymbol{\alpha})\,. \tag{B.8}$$

Here for fixed[2] $T > 0$,

$$\gamma_{\text{AB}}(\alpha) \triangleq \inf_{v>0,\mathbf{f},\bar{\mathbf{A}},\bar{\mathbf{X}}} \frac{\Lambda^*(\mathbf{f})}{v},$$

where $(\bar{\mathbf{A}}, \bar{\mathbf{X}})$ *is a FSP on* $[0,T]$ *under* U *such that for some regular* $t \in [0,T]$

$$\dot{\bar{\mathbf{A}}}(t) = \mathbf{f}, \quad L_{\boldsymbol{\alpha}}(\bar{\mathbf{X}}(t)) < 1, \quad \dot{L}_{\boldsymbol{\alpha}}(\bar{\mathbf{X}}(t)) = v.$$

Proof of Lemma B.3. *Step 1. Define stopping times and consider the sampling chain.* In this step, we mostly follow the approach in [45] (Freidlin-Wentzell theory) and decompose the expression for the likelihood of the Lyapunov function taking on a large value. There are minor differences between our proof and proof of Theorem 4 in [45] because of our closed queueing network setting, so we will write down each step for completeness.

Let $\bar{\mathbf{X}}_{\mathbf{z}}^{K,U}(\infty)$ be a random vector distributed as the stationary distribution of recurrent class associated with initial (normalized) state $\mathbf{z} \in \Omega$. For notation simplicity, we suppress the dependence on $\mathbf{z}$ and U and keep them fixed. We want to upper bound:

$$\limsup_{K\to\infty} \frac{1}{K} \log \mathbb{P}\left(L_{\boldsymbol{\alpha}}(\bar{\mathbf{X}}^K(\infty)) \geq 1\right).$$

Choose positive constants δ, ϵ such that $0 < \delta < \epsilon < 1$. Consider the following stopping times defined on a sample path $\bar{\mathbf{X}}^K(\cdot)$:

$$\begin{aligned}
\beta_1^K &\triangleq \inf\{t \geq 0 : L_{\boldsymbol{\alpha}}(\bar{\mathbf{X}}^K(t)) \leq \delta\}, \\
\eta_i^K &\triangleq \inf\{t \geq \beta_i^K : L_{\boldsymbol{\alpha}}(\bar{\mathbf{X}}^K(t)) \geq \epsilon\}, \quad i = 1, 2, \cdots \\
\beta_i^K &\triangleq \inf\{t \geq \eta_{i-1}^{K,U} : L_{\boldsymbol{\alpha}}(\bar{\mathbf{X}}^K(t)) \leq \delta\}, \quad i = 2, 3, \cdots
\end{aligned}$$

Let the discrete-time Markov chain $\hat{\mathbf{X}}^K[i]$ be obtained by sampling $\bar{\mathbf{X}}^K(t)$ at the stopping times η_i^K. Since $\bar{\mathbf{X}}^K(\cdot)$ is stationary, there must also exist a stationary distribution for Markov chain $\hat{\mathbf{X}}^K[\cdot]$. Let Θ^K denote the state space of the sampled chain $\hat{\mathbf{X}}^K[\cdot]$, $\hat{\pi}^K$ is the sampled chain's stationary distribution.

The above construction was based on the following idea: first divide time into cycles, where the i-th cycle is the interval of time between consecutive η_i's, i.e., a cycle is completed each time the value of $L_\alpha(\bar{\mathbf{X}}^K)$ goes down below δ and then rises above ϵ. Then

[2]The definition of quantity $\gamma_{\text{AB}}(\boldsymbol{\alpha})$ is based on the local behavior of $\bar{\mathbf{A}}$ and $\bar{\mathbf{X}}$ for times close to t. In particular, the value of T plays no role.

the fraction of time the Lyapunov function spent above 1 is equal to the ratio

$$\mathbb{E}[\text{time for which } L_{\boldsymbol{\alpha}}(\bar{\mathbf{X}}^K) \geq 1 \text{ during a cycle}]/(\mathbb{E}[\text{length of cycle}])$$

in steady state. We sample the initial state as $\bar{\mathbf{X}}^K(0) = \mathbf{x} \sim \hat{\pi}^K$, hence the first cycle itself characterizes the steady state ratio. Therefore, the stationary likelihood of event $\{L_{\boldsymbol{\alpha}}(\bar{\mathbf{X}}^K) \geq 1\}$ can be expressed as (see Lemma 10.1 in [47]):

$$\mathbb{P}\left(L_{\boldsymbol{\alpha}}(\bar{\mathbf{X}}^K) \geq 1\right) = \frac{\int_{\Theta^K} \hat{\pi}^K(d\mathbf{x}) \cdot \mathbb{E}\left(\int_0^{\eta_1^K} \mathbb{I}\left\{L_{\boldsymbol{\alpha}}(\bar{\mathbf{X}}^K(t)) \geq 1\right\} dt \,\middle|\bar{\mathbf{X}}^K(0) = \mathbf{x}\right)}{\int_{\Theta^K} \hat{\pi}^K(d\mathbf{x}) \cdot \mathbb{E}(\eta_1^K|\bar{\mathbf{X}}^K(0) = \mathbf{x})}. \tag{B.9}$$

Step 2. Bounding the RHS of (B.9). To upper bound $\mathbb{P}\left(L_{\boldsymbol{\alpha}}(\bar{\mathbf{X}}^K) \geq 1\right)$, we lower bound the denominator in the RHS of (B.9) and upper bound the numerator.

- *Step 2a. Bounding the Denominator.* To lower bound the denominator, we focus on the discrete-time embedded chain of $\{\bar{\mathbf{X}}^K(\cdot)\}$. Note each exactly one demand arrives at each jump of the chain, therefore $||\bar{\mathbf{X}}^K(\cdot)||_\infty$ change by at most $\frac{1}{K}$ at each jump. Using property 2 of $L_{\boldsymbol{\alpha}}(\cdot)$ in Lemma B.2, we further have that $L_{\boldsymbol{\alpha}}(\bar{\mathbf{X}}^K(\cdot))$ change by at most $\frac{1}{K\cdot\min_i \alpha_i}$ at each jump. Since the Lyapunov function $L_{\boldsymbol{\alpha}}(\bar{\mathbf{X}}^K(\cdot))$ has to increase from δ to ϵ during $[0, \eta_1^K]$, there exists $K_1 = K_1(\epsilon, \delta) > 0$ such that for any $K > K_1$, at least $\frac{K\cdot\min_i \alpha_i}{2}(\epsilon - \delta)$ jumps occur during $[0, \eta_1^K]$. Because the times between two consecutive jumps follow i.i.d. exponential distribution with rate $K\mathbf{1}^{\mathrm{T}}\hat{\boldsymbol{\phi}}\mathbf{1}$, therefore for any $K > K_1$,

$$\mathbb{E}(\eta_1^K|\bar{\mathbf{X}}^K(0) = \mathbf{x}) \geq \frac{K\cdot\min_i \alpha_i}{2}(\epsilon-\delta)\frac{1}{K\mathbf{1}^{\mathrm{T}}\hat{\boldsymbol{\phi}}\mathbf{1}} = \frac{\min_i \alpha_i}{2\cdot\mathbf{1}^{\mathrm{T}}\hat{\boldsymbol{\phi}}\mathbf{1}}(\epsilon-\delta)\,. \tag{B.10}$$

- *Step 2b. Bounding the Numerator.* This part is more complex, and we first decompose the numerator into several terms. Let $\rho \in (\epsilon, 1)$. Because each (normalized) queue length change by at most $\frac{1}{K}$ at each jump almost surely, and that $L_{\boldsymbol{\alpha}}(\cdot)$ is Lipschitz continuous, there exists $K_2 = K_2(\epsilon, \rho) > 0$, such that for all $K \geq K_2$, we have $L(\bar{\mathbf{X}}^K(\eta_i^K)) \leq \rho$.

We define another stopping time:

$$\eta^{K,\uparrow} \triangleq \inf\{t \geq 0 : L_{\boldsymbol{\alpha}}(\bar{\mathbf{X}}^K(t)) \geq 1\}\,.$$

Then for any $\mathbf{x} \in \Theta^K$, we must have:

$$\mathbb{E}\left(\int_0^{\eta_1^K} \mathbb{I}\{L_{\boldsymbol{\alpha}}(\bar{\mathbf{X}}^K(t)) \geq 1\}dt \,\middle|\bar{\mathbf{X}}^K(0) = \mathbf{x}\right) \leq\; \mathbb{E}\left(\mathbb{I}\{\eta^{K,\uparrow} \leq \beta_1^K\}(\beta_1^K - \eta^{K,\uparrow}) \,\middle|\bar{\mathbf{X}}^K(0) = \mathbf{x}\right)\,.$$

The above inequality holds because:

- if $\beta_1^K \leq \eta^{K,\uparrow}$, then both sides are zero (because the Lyapunov function will hit ϵ before 1);
- if $\beta_1^K > \eta^{K,\uparrow}$, then $L_{\boldsymbol{\alpha}}(\bar{\mathbf{X}}^K(t)) \geq 1$ can occur only for a subset of $t \in [\eta^{K,\uparrow}, \beta_1^K]$, and this time interval has length $\beta_1^K - \eta^{K,\uparrow}$.

Hence

$$\begin{aligned}
&\mathbb{E}\left(\int_0^{\eta_1^K} \mathbb{I}\left\{L_{\boldsymbol{\alpha}}(\bar{\mathbf{X}}^K(t)) \geq 1\right\} dt \,\middle|\bar{\mathbf{X}}^K(0) = \mathbf{x}\right) \\
\leq\; &\mathbb{E}\left(\beta_1^K - \eta^{K,\uparrow} \,\middle|\eta^{K,\uparrow} \leq \beta_1^K, \bar{\mathbf{X}}^K(0) = \mathbf{x}\right)\mathbb{P}\left(\eta^{K,\uparrow} \leq \beta_1^K \,\middle|\bar{\mathbf{X}}^K(0) = \mathbf{x}\right)\,.
\end{aligned}$$

Define

$$\beta^K(\mathbf{x}) \triangleq \inf\left\{t \geq 0 : L_{\boldsymbol{\alpha}}(\bar{\mathbf{X}}^K(t)) \leq \delta \,\middle|\bar{\mathbf{X}}^K(0) = \mathbf{x}\right\}\,.$$

Using the properties of Markov chains and conditional expectation, we have:

$$\begin{aligned}
&\mathbb{E}\left(\beta_1^K - \eta^{K,\uparrow} \,\middle|\eta^{K,\uparrow} \leq \beta_1^K, \bar{\mathbf{X}}^K(0) = \mathbf{x}\right) \\
=\; &\mathbb{E}\left(\mathbb{E}\left(\beta^K\left(\bar{\mathbf{X}}^K(\eta^{K,\uparrow})\right)\right) \,\middle|\eta^{K,\uparrow} \leq \beta_1^K, \bar{\mathbf{X}}^K(0) = \mathbf{x}\right) \\
\leq\; &\sup_{\mathbf{x} \in \Omega}\mathbb{E}\left(\beta_1^K \,\middle|\bar{\mathbf{X}}^K(0) = \mathbf{x}\right)\,.
\end{aligned}$$

Let T be a positive number which will be chosen later. Recall that $L_{\boldsymbol{\alpha}}(\mathbf{x}) \leq \rho$ for all $\mathbf{x} \in \Theta^K$ almost surely when $K \geq K_2$. Hence, for any such $\mathbf{x} \in \Theta^K$, we have,

$$\mathbb{E}\left(\int_0^{\eta_1^K} \mathbb{I}\left\{L_{\boldsymbol{\alpha}}(\bar{\mathbf{X}}^K(t)) \geq 1\right\} dt \,\middle|\bar{\mathbf{X}}^K(0) = \mathbf{x}\right)$$

$$
\begin{aligned}
&\le \mathbb{E}\left(\beta_1^K - \eta^{K,\uparrow} \,\middle|\, \eta^{K,\uparrow} \le \beta_1^K, \bar{\mathbf{X}}^K(0) = \mathbf{x}\right) \mathbb{P}\left(\eta^{K,\uparrow} \le \beta_1^K \,\middle|\, \bar{\mathbf{X}}^K(0) = \mathbf{x}\right) \\
&\le \left(\sup_{\mathbf{x}\in\Omega} \mathbb{E}\left(\beta_1^K \,\middle|\, \bar{\mathbf{X}}^K(0) = \mathbf{x}\right)\right) \Bigg[\mathbb{P}\left(\eta^{K,\uparrow} \le T \,\middle|\, \bar{\mathbf{X}}^K(0) = \mathbf{x}\right) \\
&\qquad + \mathbb{P}\left(\beta_1^K \ge T \,\middle|\, \bar{\mathbf{X}}^K(0) = \mathbf{x}\right)\Bigg] \qquad \left(\text{using } \eta^{K,\uparrow} \le \beta_1^K \Rightarrow \eta^{K,\uparrow} \le T \text{ or } T \le \beta_1^K\right) \\
&\le \underbrace{\left(\sup_{\mathbf{x}\in\Omega} \mathbb{E}\left(\beta_1^K \,\middle|\, \bar{\mathbf{X}}^K(0) = \mathbf{x}\right)\right)}_{\text{(a)}} \Bigg[\underbrace{\sup_{\mathbf{x}:L_{\boldsymbol{\alpha}}(\mathbf{x})\le\rho} \mathbb{P}\left(\eta^{K,\uparrow} \le T \,\middle|\, \bar{\mathbf{X}}^K(0) = \mathbf{x}\right)}_{\text{(b)}} \\
&\qquad + \underbrace{\sup_{\mathbf{x}:L_{\boldsymbol{\alpha}}(\mathbf{x})\le\rho} \mathbb{P}\left(\beta_1^K \ge T \,\middle|\, \bar{\mathbf{X}}^K(0) = \mathbf{x}\right)}_{\text{(c)}}\Bigg]. \qquad \text{(B.11)}
\end{aligned}
$$

— *Step 2b(i). Bounding term (a).* Term (a) is the upper bound of the expected time for the Lyapunov function to hit a lower level δ starting from a higher level ϵ. Because the policy U satisfies the negative drift condition, it follows from standard argument (see Part B(1) of the proof of Theorem 4 in [45], which applies the classical results in [56]) that there exists $K_3 = K_3(\delta, \epsilon)$ and constant $C > 0$ such that for $K \ge K_3$, we have $(a) \le C$.

— *Step 2b(ii). Asymptotics for (b).* Let $K \to \infty$ and apply Proposition 2 in [45] to $\bar{\mathbf{X}}^K(\cdot)$. We have:

$$
\begin{aligned}
&\limsup_{K\to\infty} \frac{1}{K} \log\left(\sup_{\mathbf{x}:L_{\boldsymbol{\alpha}}(\mathbf{x})\le\rho} \mathbb{P}\left(\eta^{K,\uparrow} \le T \,\middle|\, \bar{\mathbf{X}}^K(0) = \mathbf{x}\right)\right) \\
\le\ & -\inf_{\bar{\mathbf{A}},\bar{\mathbf{X}}} \int_0^T \Lambda^*\left(\dot{\bar{\mathbf{A}}}(t)\right) dt,\ \textit{where } (\bar{\mathbf{A}}, \bar{\mathbf{X}}) \textit{ is an FSP} \\
&\textit{such that } L_{\boldsymbol{\alpha}}(\bar{\mathbf{X}}(0)) \le \rho,\ L_{\boldsymbol{\alpha}}(\bar{\mathbf{X}}(t)) \ge 1 \textit{ for some } t \in [0, T].
\end{aligned}
$$

— *Step 2b(iii). Asymptotics for (c).* Intuitively, term (c) is the tail probability of the duration of a cycle that terminates when the Lyapunov function hit δ. It remains to be shown that this term is negligible comparing to (b) as $T \to \infty$. Let $K \to \infty$

and apply Proposition 2 in [45] to $\bar{\mathbf{X}}^K(\cdot)$. We obtain:

$$\begin{aligned}
&\limsup_{K\to\infty} \frac{1}{K} \log \left(\sup_{\mathbf{x}: L_{\boldsymbol{\alpha}}(\mathbf{x}) \le \rho} \mathbb{P}\left(\beta_1^K \ge T \,\middle|\, \bar{\mathbf{X}}^K(0) = \mathbf{x} \right) \right) \\
\le & - \inf_{\bar{\mathbf{A}}, \bar{\mathbf{X}}} \int_0^T \Lambda^*\left(\dot{\bar{\mathbf{A}}}(t) \right) dt, \ \textit{where } (\bar{\mathbf{A}}, \bar{\mathbf{X}}) \textit{ is an FSP} \\
& \textit{such that } L_{\boldsymbol{\alpha}}(\bar{\mathbf{X}}(0)) \le \rho,\ L_{\boldsymbol{\alpha}}(\bar{\mathbf{X}}(t)) \ge \delta \textit{ for all } t \in [0,T]\,.
\end{aligned}$$

We focus on the variational problem on the RHS. Note that any FSP that is feasible to the variational problem must satisfy:

$$\delta \le \ L_{\boldsymbol{\alpha}}(\bar{\mathbf{X}}(0)) + \int_{t=1}^T \dot{L}(\bar{\mathbf{X}}(t))dt \le \ \rho + \int_{t=1}^T \dot{L}(\bar{\mathbf{X}}(t))dt\,.$$

For any fixed FSP, define $\mathcal{T}_0 \triangleq \{t \in [0,T] : \dot{L}(\bar{\mathbf{X}}(t)) > -\eta\}$, where η is the negative drift parameter in the statement of Proposition 2.4. Denote the measure of $\mathcal{T}_0$ by t_0. Therefore it must hold that:

$$\begin{aligned}
\rho + \int_{t=1}^T \dot{L}(\bar{\mathbf{X}}(t))dt &= \rho + \int_{t\notin\mathcal{T}_0} \dot{L}(\bar{\mathbf{X}}(t))dt + \int_{t\in\mathcal{T}_0} \dot{L}(\bar{\mathbf{X}}(t))dt \\
&\le \rho - \eta(T - t_0) + \int_{t\in\mathcal{T}_0} \dot{L}(\bar{\mathbf{X}}(t))dt\,.
\end{aligned}$$

Hence

$$\int_{t\in\mathcal{T}_0} \dot{L}(\bar{\mathbf{X}}(t))dt \ge \ \eta(T-t_0) + \delta - \rho \ge \ \eta(T-t_0) - 1\,.$$

There are two cases:

Case 1: When $t_0 > \frac{T}{2}$. Define

$$\begin{aligned}
J_{\min} \triangleq \min \quad & \Lambda^*(\dot{\bar{\mathbf{A}}}(t)) \\
\text{subject to} \quad & \dot{L}(\bar{\mathbf{A}}(t)) \ge -\eta\,,\ t \in [0,T]\,,\ (\bar{\mathbf{A}}(t), \bar{\mathbf{X}}(t)) \text{ is an FSP.}
\end{aligned} \tag{B.12}$$

Note that $J_{\min} \ge \min_{\mathbf{f}\notin B(\phi,\epsilon')} \Lambda^*(\mathbf{f}) > 0$ and ϵ' is the ϵ specified in condition (2) of Proposition 2.4. Therefore a lower bound of the exponent of these sample paths is

$$\int_0^T \Lambda^*\left(\dot{\bar{\mathbf{A}}}(t) \right) dt \ge \ \frac{T}{2} J_{\min}\,.$$

Case 2: When $t_0 \le \frac{T}{2}$. We have

$$\int_{t\in\mathcal{T}_0} \dot{L}(\bar{\mathbf{X}}(t))dt \ge \ \eta(T-t_0) - 1 \ge \frac{\eta T}{2} - 1\,.$$

We choose $T > \frac{4}{\eta}$, therefore $\frac{\eta T}{2} - 1 \ge \frac{\eta T}{4}$. A lower bound of the exponent of these sample paths is the value of the following variational problem:

$$J(T) \triangleq -\inf_{\bar{\mathbf{A}},\bar{\mathbf{X}}} \int_0^T \Lambda^*\left(\dot{\bar{\mathbf{A}}}(t)\right) dt, \ \textit{where } (\bar{\mathbf{A}},\bar{\mathbf{X}}) \textit{ is an FSP}$$
$$\textit{such that} \int_0^T \max\{\dot{L}_{\boldsymbol{\alpha}}(\bar{\mathbf{X}}(t)), 0\}dt \ge \frac{\eta T}{4}\,.$$

We claim that $J(T) \to \infty$ as $T \to \infty$ and prove the claim in step 3.

Combine the two cases, we have:

$$\limsup_{K\to\infty} \frac{1}{K} \log\left(\sup_{\mathbf{x}:L_{\boldsymbol{\alpha}}(\mathbf{x})\le\rho} \mathbb{P}\left(\beta_1^K \ge T \,\middle|\, \bar{\mathbf{X}}^K(0) = \mathbf{x}\right)\right) \le -\min\left\{\frac{T}{2}J_{\min}\,, J(T)\right\}\,.$$

It is not hard to see that as $T \to \infty$, the exponent of term (c) tends to $-\infty$ hence is negligible.

Now combine all the terms. For fixed ϵ, δ, ρ, note that the denominator of (B.9) and (a) in (B.11) are bounded by a constant term, so they have no contribution to the exponent of (B.9). Since as $T \to \infty$, (c) in (B.11) have an exponent that is at most $-\liminf_{T\to\infty} J(T)$, we have

$$\begin{aligned}
&\limsup_{K\to\infty} \frac{1}{K} \log \mathbb{P}_{\mathrm{p}}^{K,U} \\
\le\ & -\liminf_{T\to\infty} J(T)\,,\ \limsup_{K\to\infty} \frac{1}{K} \log\left(\max_{\bar{\mathbf{X}}^K(0)\in\Omega} \mathbb{P}\left(L_{\boldsymbol{\alpha}}(\bar{\mathbf{X}}^K(\infty)) \ge 1\right)\right) && \text{(B.13)}\\
\le\ & -\inf_{T>0} \inf_{\bar{\mathbf{A}},\bar{\mathbf{X}}} \int_0^T \Lambda^*\left(\dot{\bar{\mathbf{A}}}(t)\right) dt \\
& \quad \textit{where } (\bar{\mathbf{A}},\bar{\mathbf{X}}) \textit{ is an FSP such that } L_{\boldsymbol{\alpha}}(\bar{\mathbf{X}}(0)) = \rho,\ L_{\boldsymbol{\alpha}}(\bar{\mathbf{X}}(T)) \ge 1\,. && \text{(B.14)}
\end{aligned}$$

Finally, let $\delta, \epsilon, \rho \to 0$, we have

$$\begin{aligned}
&\limsup_{K\to\infty} \frac{1}{K} \log \mathbb{P}_{\mathrm{p}}^{K,U} \\
\le\ & -\inf_{T>0} \inf_{\bar{\mathbf{A}},\bar{\mathbf{X}}} \int_0^T \Lambda^*\left(\dot{\bar{\mathbf{A}}}(t)\right) dt
\end{aligned}$$

where $(\bar{\mathbf{A}}, \bar{\mathbf{X}})$ *is an FSP such that* $L_{\boldsymbol{\alpha}}(\bar{\mathbf{X}}(0)) = 0$, $L_{\boldsymbol{\alpha}}(\bar{\mathbf{X}}(T)) \geq 1$.

We briefly summarize Step 2 and provide some intuition. The goal is to upper bound the stationary likelihood that the Lyapunov function equals 1. To study the stationary behavior, we first divide time into cycles, where a cycle is completed each time the Lyapunov function goes down below δ then rises above ϵ, where $\delta < \epsilon \ll 1$. Then using a variant of renewal-reward theorem (equation (B.9)), we only need to lower bound the expected cycle duration, and upper bound the expected time the Lyapunov function stays at 1 during a cycle. The Lipschitz property of the Lyapunov function ensures that the cycle duration is bounded away from 0 hence has no contribution to the *exponent* of the desired likelihood (Lemma B.2). Meanwhile, the negative drift condition ensures the expected time until the Lyapunov function returns to δ after hitting 1. This leaves the exponent of the desired likelihood to be solely dependent on the probability that the Lyapunov function ever hit 1 during a cycle. Finally we apply the sample path large deviation principle (Fact 2.1) to bound this quantity.

Step 3. Reduce (B.13) to an one-dimensional variational problem. This rest of the proof is exactly the same as the proof of Theorem 5 and Proposition 7 in [45]; we provide the intuition and omit the details.

The proof up until this point dealt with the *steady state* of the system. Recall the link between the exponent and value of a differential game described in Section 2.5.3. We now lower bound the exponent of the steady state demand loss probability by a variational problem (differential game), namely, (B.14). Since we are trying to lower bound the adversary's cost, we consider an "ideal adversary" who can increase $L_{\boldsymbol{\alpha}}(\mathbf{x})$ at the minimum cost at *each* level set. Mathematically,

$$\text{The quantity in } (B.14) \leq -\inf_{T>0} \theta_T \,, \tag{B.15}$$

where

$$\theta_T \triangleq \inf_{L_{\boldsymbol{\alpha}}(\cdot)} \int_0^T l_{\boldsymbol{\alpha},T}\left(L(t), \dot{L}(t)\right) dt$$

$$\text{s.t. } L(\cdot) \text{ is absolutely continuous and } L(0) = 0, \quad L(T) \geq 1\,.$$

$$\begin{aligned} l_{\boldsymbol{\alpha},T}(y,v) \triangleq \inf_{\bar{\mathbf{A}},\bar{\mathbf{X}}} \ & \Lambda^*(\mathbf{f}) \\ \text{s.t. } & (\bar{\mathbf{A}},\bar{\mathbf{X}}) \text{ is an FSP on } [0,T] \text{ such that for some regular } t \in [0,T] \\ & \dot{\bar{\mathbf{A}}}(t) = \mathbf{f}, \quad L_{\boldsymbol{\alpha}}(\bar{\mathbf{X}}(t)) = y, \quad \dot{L}_{\boldsymbol{\alpha}}(\bar{\mathbf{X}}(t)) = v\,. \end{aligned}$$

Using the scale-invariance property of $L_{\boldsymbol{\alpha}}(\mathbf{x})$ (Lemma B.1), we can show that $l_{\boldsymbol{\alpha},T}(y,v)$ is independent of y (Proposition 7 in [45]). As a result, the above variational problem reduces to an one-dimensional problem where the "ideal adversary" chooses a single rate (i.e., v in the statement of Lemma B.3) at which $L_{\boldsymbol{\alpha}}(\mathbf{x})$ increases. This problem is exactly the one in the statement of Lemma B.3.

(We prove the claim in step 2 that $\liminf_{T\to\infty} J(T) = \infty$ here. Using exactly the same argument as in step 3, we can show that $J(T) \geq \frac{\eta T}{4}\gamma_{\text{AB}}(\boldsymbol{\alpha})$ where the RHS is defined in (B.8). This concludes the proof.)

□

B.3.2 Converse Bound Matches Achievability Bound

In Lemma 2.1 we obtain a converse bound which holds for any state-dependent policy. However, for a given policy U can we obtain a tighter policy-specific converse bound? In the following Lemma, we show that for policies that satisfy the negative drift property in Proposition 2.4 for Lyapunov function $L_{\boldsymbol{\alpha}}(\cdot)$ where $\boldsymbol{\alpha} \in \text{relint}(\Omega)$, there is a tighter converse bound given by $\gamma_{\text{CB}}(\boldsymbol{\alpha})$.

Lemma B.4. *For policies $U \in \mathcal{U}$ that satisfy the negative drift condition in the statement of Proposition 2.4 for $\boldsymbol{\alpha} \in \text{relint}(\Omega)$, we have*

$$-\liminf_{K\to\infty} \frac{1}{K} \log P_{\text{o}}^{K,U} \leq \gamma_{\text{CB}}(\boldsymbol{\alpha})\,.$$

Proof. The following proof is very similar to the proof of Lemma 2.1. We will emphasize the parts that are different and skip the repetitive arguments. In the proof of Lemma

2.1, we divide the process into cycles and apply the renewal-reward theorem. We follow the same approach here except that we define the cycles differently.

Step 1: Show that $\boldsymbol{\alpha}$ is the "resting point" of U. Fix $\epsilon_1 > 0$ and define

$$\tau^K \triangleq \inf\left\{t \geq 0 : L_{\boldsymbol{\alpha}}(\bar{\mathbf{X}}^K(t)) \leq \epsilon_1\right\}.$$

Using the argument in Step 2b(i) of the proof of Lemma B.3, we can show that there exists $K_0 = K_0(\epsilon_1) > 0$ and constant $C > 0$ such that for $K \geq K_0$,

$$\sup_{\mathbf{x}\in\Omega} \mathbb{E}\left(\tau^K | \bar{\mathbf{X}}^K(0) = \mathbf{x}\right) \leq C.$$

In other words, starting from any state, the expected time for the system state to reach the $O(\epsilon_1)$-neighborhood of $\boldsymbol{\alpha}$ is bounded from above by a constant.

Step 2: Lower bound the demand-loss probability. Proceed exactly as Step 2 and Step 3 in the proof of Lemma 2.1, we explicitly construct a demand sample path that guarantees a demand loss within $\Theta(1)$ units of time given the starting state satisfies $L_\alpha\left(\bar{\mathbf{X}}^K(T+\tau^K)\right) < \epsilon_1$. Then we obtain the desired result.

□

Now we combine Lemma B.3 and Lemma B.4 to prove Proposition 2.4 by showing that $\gamma_{\text{AB}}(\boldsymbol{\alpha}) = \gamma_{\text{CB}}(\boldsymbol{\alpha})$. Lemma B.1 and the steepest descent property in Proposition 2.4 are crucial in showing $\gamma_{\text{AB}}(\boldsymbol{\alpha}) \geq \gamma_{\text{CB}}(\boldsymbol{\alpha})$ (the other direction is obvious).

Proof of Proposition 2.4. Let $U \in \mathcal{U}$ satisfy the conditions in Proposition 2.4. Then for regular t we have

$$\begin{aligned}
\dot{L}_{\boldsymbol{\alpha}}(\bar{\mathbf{X}}(t)) &\leq \inf_{U'\in\mathcal{U}} \left\{\dot{L}_{\boldsymbol{\alpha}}(\bar{\mathbf{X}}^{U'}(t)) \,\middle|\, \dot{\bar{\mathbf{A}}}'(t) = \mathbf{f}\right\} && \text{(steepest descent)}\\
&= \min_{\Delta\mathbf{x}\in\mathcal{X}_{\mathbf{f}}} \lim_{\Delta t\to 0} \frac{L_{\boldsymbol{\alpha}}(\bar{\mathbf{X}}^{U'}(t) + \Delta\mathbf{x}\Delta t) - L_{\boldsymbol{\alpha}}(\bar{\mathbf{X}}^{U'}(t))}{\Delta t}\\
&\leq \min_{\Delta\mathbf{x}\in\mathcal{X}_{\mathbf{f}}} \lim_{\Delta t\to 0} \frac{L_{\boldsymbol{\alpha}}(\boldsymbol{\alpha} + \Delta\mathbf{x}\Delta t)}{\Delta t} && \text{(sub-additivity, Lemma B.1)}\\
&= \min_{\Delta\mathbf{x}\in\mathcal{X}_{\mathbf{f}}} L_{\boldsymbol{\alpha}}(\boldsymbol{\alpha} + \Delta\mathbf{x}) = v_{\boldsymbol{\alpha}}(\mathbf{f}). && \text{(scale-invariance, Lemma B.1)}
\end{aligned}$$

Let $v = \dot{L}_{\boldsymbol{\alpha}}(\bar{\mathbf{X}}(t))$, from $v \leq v_{\boldsymbol{\alpha}}(\mathbf{f})$ we have $\{v > 0\} \subset \{v_{\boldsymbol{\alpha}}(\mathbf{f}) > 0\}$, hence using Lemma

B.3 we have

$$\gamma_{\text{AB}}(\boldsymbol{\alpha}) = \inf_{v>0,\mathbf{f},\bar{\mathbf{A}},\bar{\mathbf{X}}} \frac{\Lambda^*(\mathbf{f})}{v} \geq \inf_{\mathbf{f}:v_{\boldsymbol{\alpha}}(\mathbf{f})>0} \frac{\Lambda^*(\mathbf{f})}{v_{\boldsymbol{\alpha}}(\mathbf{f})} = \gamma_{\text{CB}}(\boldsymbol{\alpha})\,.$$

But since by Lemma B.4 we know $\gamma_{\text{CB}}(\boldsymbol{\alpha})$ is a converse bound for policy U, hence $\gamma_{\text{AB}}(\boldsymbol{\alpha}) \leq \gamma_{\text{CB}}(\boldsymbol{\alpha})$. Therefore $\gamma_{\text{AB}}(\boldsymbol{\alpha}) = \gamma_{\text{CB}}(\boldsymbol{\alpha})$. □

B.4 SMW Policies and Explicit Exponent

Appendix B.4 shows that the SMW policy satisfies the sufficient conditions for exponent optimality, and derives explicitly the optimal exponent and most-likely sample paths, including the proofs of Lemma 2.2, Lemma 2.3, and Lemma 2.4. The last subsection formally establishes exponent optimality of SMW policies for transient performance.

B.4.1 Lyapunov Drift of FSPs under SMW: Proof of Lemma 2.2

In this subsection we prove Lemma 2.2 which establishes that SMW($\boldsymbol{\alpha}$) policies perform steepest descent on $L_{\boldsymbol{\alpha}}(\cdot)$.

Proof of Lemma 2.2. For notation simplicity, we will write $S_1(\bar{\mathbf{X}}(t))$ as S_1, $S_2\left(\bar{\mathbf{X}}(t), \dot{\bar{\mathbf{X}}}(t)\right)$ as S_2, and $\min_{k\in S_1} \frac{\dot{\bar{X}}_k(t)}{\alpha_k}$ as c in the following. Let $(\bar{\mathbf{A}}, \bar{\mathbf{X}})$ be an FSP under policy $U \in \mathcal{U}$.

- *Proof of (2.22).* Note that t is a *regular* time, hence $L_{\boldsymbol{\alpha}}(\bar{\mathbf{X}}(\cdot))$ and $\bar{\mathbf{X}}(\cdot)$ are differentiable at t. It follows from the definition of derivatives that $\dot{L}_{\boldsymbol{\alpha}}(\bar{\mathbf{X}}(t))$ is determined by the queues in S_2 alone, hence we have $\dot{L}_{\boldsymbol{\alpha}}(\bar{\mathbf{X}}(t)) = -\min_{k\in S_1} \frac{\dot{\bar{X}}_k(t)}{\alpha_k} = -c$.
- *Proof of (2.23).* For the K-th system, define auxiliary processes:

$$\bar{E}^{K,U}_{ij'k}(t) \triangleq \#\{\text{Type } (j',k) \text{ demand units that arrive during } [0,t] \text{ and are served by supply units at } i \text{ under policy } U \in \mathcal{U}\} \quad i,k \in V_S\,,\ j' \in V_D\,.$$

Using standard argument [see, e.g., 7], we can extend the definition of FSP (Definition 2.5) to $(\bar{\mathbf{A}}(\cdot), \bar{\mathbf{X}}(\cdot), \bar{\mathbf{E}}(\cdot))$, where a subsequence of $\bar{\mathbf{E}}^{K,U}(\cdot)$ converges u.o.c. to $\bar{\mathbf{E}}(\cdot)$. We focus on the regular times t where $\dot{\bar{\mathbf{E}}}(t)$ exists, which includes almost all regular times because $\dot{\bar{\mathbf{E}}}(t)$ is differentiable almost everywhere.

Consider any non-idling policy $U' \in \mathcal{U}$, and $\bar{\mathbf{X}}^{U'}(t)$ such that $\bar{\mathbf{X}}^{U'}(t) \neq \boldsymbol{\alpha}$, $L_{\boldsymbol{\alpha}}(\bar{\mathbf{X}}^{U'}(t)) < 1$. The flow of supply units entering S_2 is $\sum_{j' \in V_D, k \in S_2} \dot{\bar{A}}_{j'k}(t)$ because U' is non-idling. The flow of supply units leaving S_2 is at least $\sum_{j' \in V_D : \partial(j') \subset S_2, k \in V_S} \dot{\bar{A}}_{j'k}(t)$ because U' is non-idling and that the supply units in $V_S \backslash S_2$ cannot be used to serve demand originating from $\{j' \in V_D : \partial(j') \subset S_2\}$. Therefore,

$$\sum_{k \in S_2} \dot{\bar{X}}_k^{U'}(t) \leq \sum_{j' \in V_D, k \in S_2} \dot{\bar{A}}_{j'k}(t) - \sum_{j' \in V_D : \partial(j') \subset S_2, k \in V_S} \dot{\bar{A}}_{j'k}(t)\,. \tag{B.16}$$

Now we consider SMW($\boldsymbol{\alpha}$) policies and $\bar{\mathbf{X}}^{\mathrm{SMW}(\boldsymbol{\alpha})}(t)$ such that $\bar{\mathbf{X}}^{\mathrm{SMW}(\boldsymbol{\alpha})}(t) \neq \boldsymbol{\alpha}$. For the process $\bar{\mathbf{E}}(t)$ (resp. $\bar{\mathbf{X}}(t)$), we use notation $\Delta\bar{\mathbf{E}}(t)$ (resp. $\Delta\bar{\mathbf{X}}(t)$) to denote $\bar{\mathbf{E}}(t+\Delta t) - \bar{\mathbf{E}}(t)$ (resp. $\bar{\mathbf{X}}(t+\Delta t) - \bar{\mathbf{X}}(t)$). It holds that

$$\sum_{k \in S_2} \Delta \bar{X}_k^{K,U}(t) = \sum_{j' \in V_D, k \in S_2} \sum_{i \in \partial(j')} \Delta \bar{E}_{ij'k}^{K,U}(t) - \sum_{i \in S_2, k \in V_S} \sum_{j' \in \partial(i)} \Delta \bar{E}_{ij'k}^{K,U}(t)\,.$$

For regular t, it follows from the definition of derivative that

$$\sum_{k \in S_2} \dot{\bar{X}}_k^{U}(t) = \sum_{j' \in V_D, k \in S_2} \sum_{i \in \partial(j')} \dot{\bar{E}}_{ij'k}^{U}(t) - \sum_{i \in S_2, k \in V_S} \sum_{j' \in \partial(i)} \dot{\bar{E}}_{ij'k}^{U}(t)\,.$$

For SMW($\boldsymbol{\alpha}$) policy, using exactly the same argument as in Lemma 4 of [7], we have

$$\dot{\bar{E}}_{ij'k}^{\mathrm{SMW}(\boldsymbol{\alpha})}(t) = 0 \qquad \text{if } \frac{\bar{X}_i^{\mathrm{SMW}(\boldsymbol{\alpha})}(t)}{\alpha_i} < \max_{\ell \in \partial(j')} \frac{\bar{X}_\ell^{\mathrm{SMW}(\boldsymbol{\alpha})}(t)}{\alpha_\ell}\,. \tag{B.17}$$

By definition of S_2, there exists $\epsilon > 0$ such that any (scaled) queue length in S_2 is strictly smaller than all (scaled) queue lengths in $V_S \backslash S_2$ in $(t, t+\epsilon)$, which also implies that the queue lengths in $V_S \backslash S_2$ remain strictly positive during $(t, t+\epsilon)$. Apply (B.17), we know that the system will use the supplies within $V_S \backslash S_2$ to serve all demands arriving at $\partial(V_S \backslash S_2)$ during $(t, t+\epsilon)$. Hence we have

$$\begin{aligned}
\sum_{k \in V_S \backslash S_2} \dot{\bar{X}}_k^{\mathrm{SMW}(\alpha)}(t) &= \sum_{j' \in V_D, k \in V_S \backslash S_2} \sum_{i \in \partial(j')} \dot{\bar{E}}_{ij'k}^{\mathrm{SMW}(\alpha)}(t) - \sum_{j' \in \partial(V_S \backslash S_2), k \in V_S} \dot{\bar{A}}_{j'k}(t) \\
&\leq \sum_{j' \in V_D, k \in V_S \backslash S_2} \dot{\bar{A}}_{j'k}(t) - \sum_{j' \in \partial(V_S \backslash S_2), k \in V_S} \dot{\bar{A}}_{j'k}(t)\,.
\end{aligned}$$

Since it is a closed system, we have:

$$
\begin{aligned}
\sum_{k\in S_2} \dot{\bar{X}}_k^{\mathrm{SMW}(\alpha)}(t) &= -\sum_{k\in V_S\setminus S_2} \dot{\bar{X}}_k^{\mathrm{SMW}(\alpha)}(t) \\
&\geq \sum_{j'\in\partial(V_S\setminus S_2),k\in V_S} \dot{\bar{A}}_{j'k}(t) - \sum_{j'\in V_D,k\in V_S\setminus S_2} \dot{\bar{A}}_{j'k}(t)\,. \qquad \text{(B.18)}
\end{aligned}
$$

Note that

$$
\begin{aligned}
\text{RHS of (B.18)} &= \sum_{j'\in\partial(V_S\setminus S_2),k\in V_S} \dot{\bar{A}}_{j'k}(t) - \sum_{j'\in V_D,k\in V_S\setminus S_2} \dot{\bar{A}}_{j'k}(t) \\
&= \left(\sum_{j'\in V_D,k\in V_S} \dot{\bar{A}}_{j'k}(t) - \sum_{j':\partial(j')\subset S_2,k\in V_S} \dot{\bar{A}}_{j'k}(t) \right) \\
&\quad - \left(\sum_{j'\in V_D,k\in V_S} \dot{\bar{A}}_{j'k}(t) - \sum_{j'\in V_D,k\in S_2} \dot{\bar{A}}_{j'k}(t) \right) \\
&= \sum_{j'\in V_D,k\in S_2} \dot{\bar{A}}_{j'k}(t) - \sum_{j':\partial(j')\subset S_2,k\in V_S} \dot{\bar{A}}_{j'k}(t) \\
&= \text{RHS of (B.16)}\,.
\end{aligned}
$$

Finally, observe that for any $k \in S_2$,

$$
\dot{L}_{\boldsymbol{\alpha}}(\bar{\mathbf{X}}(t)) = -\frac{\dot{\bar{X}}_k^{U'}(t)}{\alpha_k} = -\frac{1}{\mathbf{1}_{S_2}^{\mathrm{T}}\boldsymbol{\alpha}} \sum_{k\in S_2} \alpha_k \frac{\dot{\bar{X}}_k^{U'}(t)}{\alpha_k} = -\frac{1}{\mathbf{1}_{S_2}^{\mathrm{T}}\boldsymbol{\alpha}} \sum_{k\in S_2} \dot{\bar{X}}_k^{U'}(t)\,. \qquad \text{(B.19)}
$$

Plug (B.18) and (B.16) into (B.19), we know that inequality (2.22) holds, and it becomes equality for SMW($\boldsymbol{\alpha}$) policy.

□

B.4.2 Lyapunov Drift of FLs under SMW: Proof of Lemma 2.3

Proof of Lemma 2.3. **Negative drift.** Let $(\bar{\mathbf{A}}, \bar{\mathbf{X}})$ be a fluid limit of the system under SMW($\boldsymbol{\alpha}$), and t be its regular point. Simply plug in Lemma 2.2, and replace $\dot{\bar{A}}_{j'k}(t)$ with

$\hat{\phi}_{j'k}$, we have (S_2 is defined in Lemma 2.2, $S_2 \neq \emptyset$)

$$\begin{aligned}
\dot{L}_{\boldsymbol{\alpha}}(\bar{\mathbf{X}}(t)) =& -\frac{1}{\mathbf{1}_{S_2}^{\mathrm{T}}\boldsymbol{\alpha}}\left(\sum_{j'\in V_D, k\in S_2}\dot{\bar{A}}_{j'k}(t) - \sum_{j'\in V_D:\partial(j')\subseteq S_2, k\in V_S}\dot{\bar{A}}_{j'k}(t)\right)\\
\leq & -\min_{S_2\subsetneq V_S, S_2\neq\emptyset}\frac{1}{\mathbf{1}_{S_2}^{\mathrm{T}}\boldsymbol{\alpha}}\left(\sum_{j'\in V_D, k\in S_2}\hat{\phi}_{j'k} - \sum_{j'\in V_D:\partial(j')\subseteq S_2, k\in V_S}\hat{\phi}_{j'k}\right)\\
\leq & -\min_{S_2\subsetneq V_S, S_2\neq\emptyset}\left(\sum_{j'\in V_D, k\in S_2}\hat{\phi}_{j'k} - \sum_{j'\in V_D:\partial(j')\subseteq S_2, k\in V_S}\hat{\phi}_{j'k}\right)\\
\overset{(a)}{\leq} & -\min\{\xi, \hat{\phi}_{\min}\}\,.
\end{aligned}$$

Here (a) holds for the following reason. First note that when $\bar{\mathbf{X}}(t) \neq \boldsymbol{\alpha}$, we have $S_2 \neq V_S$. Let $J \triangleq \{j' \in V_D : \partial(j') \subset S_2; \exists k \in V_S\backslash S_2 \text{ s.t. } \phi_{j'k} > 0\}$. If $J = \emptyset$, we have

$$\begin{aligned}
& \sum_{j'\in V_D, k\in S_2}\hat{\phi}_{j'k} - \sum_{j'\in V_D:\partial(j')\subseteq S_2, k\in V_S}\hat{\phi}_{j'k}\\
= & \sum_{j'\in V_D:\partial(j')\cap(V_S\backslash S_2)\neq\emptyset, k\in S_2}\hat{\phi}_{j'k} - \sum_{j'\in V_D:\partial(j')\subseteq S_2, k\in V_S\backslash S_2}\hat{\phi}_{j'k}\\
\geq & \sum_{j'\in V_D:\partial(j')\cap(V_S\backslash S_2)\neq\emptyset, k\in S_2}\hat{\phi}_{j'k} \geq \hat{\phi}_{\min}\,,
\end{aligned}$$

where $\hat{\phi}_{\min} \triangleq \min_{j'\in V_S, k\in V_S, \hat{\phi}_{j'k}>0}\hat{\phi}_{j'k}$ is the minimum positive arrival rate for any demand type (j', k) (the last inequality holds because of Assumption 1). If $J \neq \emptyset$, we must have $J \in \mathcal{J}$, hence

$$\sum_{j'\in V_D, k\in S_2}\hat{\phi}_{j'k} - \sum_{j'\in V_D:\partial(j')\subseteq S_2, k\in V_S}\hat{\phi}_{j'k} \geq \sum_{j'\in V_D, k\in\partial(J)}\hat{\phi}_{j'k} - \sum_{j'\in J, k\in V_S}\hat{\phi}_{j'k} \geq \xi\,,$$

where $\xi \triangleq \min_{J\in\mathcal{J}}\left(\sum_{i\in\partial(J)}\mathbf{1}^{\mathrm{T}}\hat{\boldsymbol{\phi}}_{(i)} - \sum_{j'\in J}\mathbf{1}^{\mathrm{T}}\hat{\boldsymbol{\phi}}_{j'}\right) > 0$ is the Hall's gap of the system.

Robustness of drift. Define

$$G(\mathbf{f}) \triangleq \min_{S\subsetneq V_S, S\neq\emptyset}\left(\sum_{j'\in V_D, k\in S}f_{j'k} - \sum_{j':\partial(j')\subseteq S, k\in V_S}f_{j'k}\right)\,.$$

Note that $G(\mathbf{f})$ is continuous in $\mathbf{f}$. Since $G(\hat{\boldsymbol{\phi}}) \leq -\min\{\xi, \hat{\phi}_{\min}\} < 0$, by continuity there exists ϵ such that for any $\dot{\bar{\mathbf{A}}}(t) \in B(\hat{\boldsymbol{\phi}}, \epsilon)$,

$$\dot{L}_{\boldsymbol{\alpha}}(\bar{\mathbf{X}}(t)) = G\left(\dot{\bar{\mathbf{A}}}(t)\right) \leq -\frac{1}{2}\min\{\xi, \hat{\phi}_{\min}\}\,.$$

$\square$

B.4.3 Explicit Exponent and Most Likely Sample Path: Proof of Lemma 2.4

Proof of Lemma 2.4. **Explicit exponent.** Let $(\bar{\mathbf{A}}(\cdot), \bar{\mathbf{X}}(\cdot))$ be a fluid sample path under SMW($\boldsymbol{\alpha}$). For a regular point t of this FSP, denote $\mathbf{f} \triangleq \dot{\bar{\mathbf{A}}}(t)$.

For notation simplicity, for $S \subset V_S$ denote

$$\text{gap}_S(\mathbf{f}) \triangleq \sum_{j':\partial(j')\subseteq S, k\in V_S} f_{j'k} - \sum_{j'\in V_D, k\in S} f_{j'k}\,.$$

In words, $\text{gap}_S(\mathbf{f})$ is the minimum net rate at which supply in S is drained given current demand arrival rate $\mathbf{f}$, assuming no demand is dropped. Using the result of Lemma 2.2, we have:

$$\dot{L}_{\boldsymbol{\alpha}}(\bar{\mathbf{X}}(t)) = \frac{\text{gap}_{S_2}(\mathbf{f})}{\mathbf{1}_{S_2}^{\mathrm{T}}\boldsymbol{\alpha}}\,, \tag{B.20}$$

where $S_2 \triangleq S_2(\bar{\mathbf{X}}(t), \dot{\bar{\mathbf{X}}}(t))$ and the latter is defined in Lemma 2.2. Given $\dot{\bar{\mathbf{A}}}(t) = \mathbf{f}$, we define

$$\bar{v}(\mathbf{f}) \triangleq \sup_{\bar{\mathbf{X}}(t)\in\Omega\backslash\{\boldsymbol{\alpha}\}} \dot{L}_{\boldsymbol{\alpha}}(\bar{\mathbf{X}}(t)) = \max_{S\neq\emptyset, S\subsetneq V_S} \frac{\text{gap}_S(\mathbf{f})}{\mathbf{1}_S^{\mathrm{T}}\boldsymbol{\alpha}}\,.$$

Recall the definition of $\gamma_{\text{AB}}(\boldsymbol{\alpha})$ in Lemma B.3, we have

$$\begin{aligned}
\gamma_{\text{AB}}(\boldsymbol{\alpha}) &= \inf_{\mathbf{f}\geq\mathbf{0}:\bar{v}(\mathbf{f})>0} \frac{\Lambda^*(\mathbf{f})}{\bar{v}(\mathbf{f})} \\
&= \inf_{\mathbf{f}\geq\mathbf{0}:\max_{S\subseteq V_S}\text{gap}_S(\mathbf{f})>0} \frac{\Lambda^*(\mathbf{f})}{\max_{S\subseteq V_S}\frac{\text{gap}_S(\mathbf{f})}{\mathbf{1}_S^{\mathrm{T}}\boldsymbol{\alpha}}} \\
&= \inf_{\mathbf{f}\geq\mathbf{0}:\max_{S\subseteq V_S}\text{gap}_S(\mathbf{f})>0} \left\{ \min_{S\subseteq V_S:\text{gap}_S(\mathbf{f})>0} \left(\mathbf{1}_S^{\mathrm{T}}\boldsymbol{\alpha}\right) \frac{\Lambda^*(\mathbf{f})}{\text{gap}_S(\mathbf{f})} \right\} &\text{(B.21)} \\
&\overset{(a)}{=} \min_{S\subseteq V_S} \left\{ \inf_{\mathbf{f}\geq\mathbf{0}:\text{gap}_S(\mathbf{f})>0} \left(\mathbf{1}_S^{T}\boldsymbol{\alpha}\right) \frac{\Lambda^*(\mathbf{f})}{\text{gap}_S(\mathbf{f})} \right\}. &\text{(B.22)}
\end{aligned}$$

For completeness, define the minimum over the empty set as $+\infty$. Here (a) holds because: For a minimizer $\mathbf{f}^* \geq \mathbf{0}$ of the outer problem of (B.21) and a minimizer $S^* \subseteq V_S$ of the inner problem of (B.21), $S^* \subseteq V_S$ is feasible for the inner problem of (B.22) while $\mathbf{f}^* \geq \mathbf{0}$

is feasible for the outer problem of (B.22), hence $(B.21) \geq (B.22)$. Similarly we can show $(B.21) \leq (B.22)$.

We claim that

$$(B.22) = \min_{J \in \mathcal{J}} \left\{ \inf_{\mathbf{f} \geq \mathbf{0} : \mathrm{gap}_{\partial(J)}(\mathbf{f}) > 0} \left(\mathbf{1}_{\partial(J)}^{\mathrm{T}} \boldsymbol{\alpha} \right) \frac{\Lambda^*(\mathbf{f})}{\mathrm{gap}_{\partial(J)}(\mathbf{f})} \right\} . \tag{B.23}$$

Recall that the definition of $\mathcal{J}$:

$$\mathcal{J} = \left\{ J \subsetneq V_D : \sum_{j' \in J} \sum_{k \notin \partial(J)} \phi_{j'k} > 0 \right\} .$$

To see (B.23), first note that for $S \subseteq V_S$ where $\{j' \in V_D : \partial(j') \subset S\}$ is empty, $\mathrm{gap}_S(\mathbf{f})$ is non-positive regardless of $\mathbf{f} \geq \mathbf{0}$, hence such S can never be the minimizer. For other S, let $J \triangleq \{j' \in V_D : \partial(j') \subset S\}$, then $\partial(J) \subset S$. Note that

$$\begin{aligned}
\mathrm{gap}_{\partial(J)}(\mathbf{f}) &= \sum_{j' \in J, k \in V_S} f_{j'k} - \sum_{j' \in V_D, k \in \partial(J)} f_{j'k} \\
&= \sum_{j' : \partial(j') \subset S, k \in V_S} f_{j'k} - \sum_{j' \in V_D, k \in S} f_{j'k} + \sum_{j' \in V_D, k \in S \backslash \partial(J)} f_{j'k} \\
&= \mathrm{gap}_S(\mathbf{f}) + \sum_{j' \in V_D, k \in S \backslash \partial(J)} f_{j'k} \\
&\geq \mathrm{gap}_S(\mathbf{f}) .
\end{aligned}$$

As a result, for $\mathbf{f}$ such that $\mathrm{gap}_S(\mathbf{f}) > 0$, we have

$$\left(\mathbf{1}_S^{\mathrm{T}} \boldsymbol{\alpha} \right) \frac{\Lambda^*(\mathbf{f})}{\mathrm{gap}_S(\mathbf{f})} \geq \left(\mathbf{1}_{\partial(J)}^{\mathrm{T}} \boldsymbol{\alpha} \right) \frac{\Lambda^*(\mathbf{f})}{\mathrm{gap}_{\partial(J)}(\mathbf{f})} .$$

Hence only those $S \subseteq V_S$ where $S = \partial(J)$ for $J \subseteq V_D$ can be the minimizer. If $J \notin \mathcal{J}$, then $\mathrm{gap}_{\partial(J)}(\mathbf{f}) \leq 0$ regardless of $\mathbf{f} \geq \mathbf{0}$, so these sets are also ruled out. Therefore (B.23) holds.

Suppose the outer minimum of (B.23) is achieved by $J^* \in \mathcal{J}$. Denote the optimal value of the inner infimum of (B.23) as $(\mathbf{1}_{\partial(J^*)}^{\mathrm{T}} \boldsymbol{\alpha}) g(\hat{\boldsymbol{\phi}}, J) > 0$, then we have:

$$\inf_{\mathbf{f} \geq \mathbf{0} : \mathrm{gap}_{\partial(J)}(\mathbf{f}) > 0} \Lambda^*(\mathbf{f}) - g(\hat{\boldsymbol{\phi}}, J) \left(\sum_{j' \in J, k \in V_S} f_{j'k} - \sum_{j' \in V_D, k \in \partial(J)} f_{j'k} \right) = 0 . \tag{B.24}$$

We can get rid of the constraint on $\mathbf{f}$ because for $\mathbf{f}$ where $\mathrm{gap}_{\partial(J)}(\mathbf{f}) \leq 0$, the argument of minimization in (B.24) is negative; and for $\mathbf{f}$ that has negative components, its rate function is ∞ by definition. Using Legendre transform, we have:

$$
\begin{aligned}
&\inf_{\mathbf{f}} \Lambda^*(\mathbf{f}) - g(\hat{\boldsymbol{\phi}}, J)\left(\sum_{j'\in J, k\in V_S} f_{j'k} - \sum_{j'\in V_D, k\in\partial(J)} f_{j'k}\right)\\
=\ & \inf_{\mathbf{f}} \Lambda^*(\mathbf{f}) - \mathbf{f}^{\mathrm{T}}\left(g(\hat{\boldsymbol{\phi}}, J)\sum_{j'\in J, k\in V_S} \mathbf{e}_{j'k} - g(\hat{\boldsymbol{\phi}}, J)\sum_{j'\in V_D, k\in\partial(J)} \mathbf{e}_{j'k}\right)\\
=\ & -\Lambda\left(g(\hat{\boldsymbol{\phi}}, J)\sum_{j'\in J, k\in V_S} \mathbf{e}_{j'k} - g(\hat{\boldsymbol{\phi}}, J)\sum_{j'\in V_D, k\in\partial(J)} \mathbf{e}_{j'k}\right)\\
\overset{(b)}{=}\ & -\sum_{j'\in V_D, k\in V_S} \hat{\phi}_{j'k}\left(e^{g(\hat{\boldsymbol{\phi}}, J)\mathbb{I}\{j'\in J\} - g(\hat{\boldsymbol{\phi}}, J)\mathbb{I}\{k\in\partial(J)\}} - 1\right).
\end{aligned}
$$

In (b) we use the fact that the dual function of $\Lambda^*(\mathbf{f})$ is $\Lambda(\mathbf{x}) = \sum_{j'\in V_D, k\in V_S} \hat{\phi}_{j'k}(e^{x_{j'k}} - 1)$ where $\mathbf{x} \in \mathbb{R}^{n\times m}$. Hence Eq. (B.24) reduces to the nonlinear equation

$$
\left(\sum_{j'\notin J, k\in\partial(J)} \hat{\phi}_{j'k}\right) e^{-g(\hat{\boldsymbol{\phi}}, J)} + \left(\sum_{j'\in J, k\notin\partial(J)} \hat{\phi}_{j'k}\right) e^{g(\hat{\boldsymbol{\phi}}, J)} = \sum_{j'\notin J, k\in\partial(J)} \hat{\phi}_{j'k} + \sum_{j'\in J, k\notin\partial(J)} \hat{\phi}_{j'k}\,.
$$

Let $y \triangleq e^{g(\hat{\boldsymbol{\phi}}, J)}$, this becomes a quadratic equation:

$$
\left(\sum_{j'\in J, k\notin\partial(J)} \hat{\phi}_{j'k}\right) y^2 - \left(\sum_{j'\notin J, k\in\partial(J)} \hat{\phi}_{j'k} + \sum_{j'\in J, k\notin\partial(J)} \hat{\phi}_{j'k}\right) y + \left(\sum_{j'\notin J, k\in\partial(J)} \hat{\phi}_{j'k}\right) = 0\,.
$$

Hence

$$
y = \frac{\sum_{j'\notin J, k\in\partial(J)} \hat{\phi}_{j'k}}{\sum_{j'\in J, k\notin\partial(J)} \hat{\phi}_{j'k}} \text{ or } 1\,.
$$

Since $g(\hat{\boldsymbol{\phi}}, J) > 0$, we have

$$
g(\hat{\boldsymbol{\phi}}, J) = \log\left(\frac{\sum_{j'\notin J, k\in\partial(J)} \hat{\phi}_{j'k}}{\sum_{j'\in J, k\notin\partial(J)} \hat{\phi}_{j'k}}\right) = \log\left(\frac{\sum_{j'\notin J, k\in\partial(J)} \phi_{j'k}}{\sum_{j'\in J, k\notin\partial(J)} \phi_{j'k}}\right).
$$

Plugging into (B.23), we have:

$$
\gamma_{\mathrm{AB}}(\boldsymbol{\alpha}) = \min_{J\in\mathcal{J}} \left(\mathbf{1}^{\mathrm{T}}_{\partial(J)}\boldsymbol{\alpha}\right) \log\left(\frac{\sum_{j'\notin J, k\in\partial(J)} \phi_{j'k}}{\sum_{j'\in J, k\notin\partial(J)} \phi_{j'k}}\right).
$$

Remark: For $J \in \mathcal{J}$, if there exists $j' \in V_D$ such that $j' \notin J$ but $\partial(j') \subseteq \partial(J)$, then such subsets J are "spurious" in the sense that they cannot achieve the minimum in the expression of $\gamma_{\text{AB}}(\boldsymbol{\alpha})$ (the term corresponding to $J \cup \{j'\}$ is no larger than the term correpsonding to J). Therefore only the "maximal" J's matter to the value of exponent.

Most likely demand sample path leading to demand loss. Denote

$$\mathbf{c} \triangleq g(\hat{\boldsymbol{\phi}}, J) \left(\sum_{j' \in J, k \in V_S} \mathbf{e}_{j'k} - \sum_{j' \in V_D, k \in \partial(J)} \mathbf{e}_{j'k} \right) ,$$

denote $\mathbf{f}_J$ as the minimizer of the inner minimization problem on the RHS of (B.23). We have

$$\begin{aligned}
\mathbf{f}_J &= \operatorname{argmin}_{\mathbf{f} \geq \mathbf{0}} \sum_{j' \in V_D} \sum_{k \in V_S} \left(\Lambda^*_{j'k}(f_{j'k}) - c_{j'k} f_{j'k} \right) \\
&= \operatorname{argmin}_{\mathbf{f} \geq \mathbf{0}} \sum_{j' \in V_D} \sum_{k \in V_S} \left(f_{j'k} \log \frac{f_{j'k}}{\hat{\phi}_{j'k}} + \hat{\phi}_{j'k} - f_{j'k} - c_{j'k} f_{j'k} \right) .
\end{aligned}$$

First order condition implies: $(\mathbf{f}_J)_{j'k} = \hat{\phi}_{j'k} \frac{e^{c_{j'k}+1}}{\sum_{j',k} \hat{\phi}_{j'k} e^{c_{j'k}+1}} = \hat{\phi}_{j'k} \frac{e^{c_{j'k}}}{\sum_{j',k} \hat{\phi}_{j'k} e^{c_{j'k}}}$. Recall the definition of λ_J, μ_J in (2.13), we have

$$\begin{aligned}
\sum_{j',k} \hat{\phi}_{j'k} e^{c_{j'k}} &= \sum_{j' \in J, k \notin \partial(J)} \hat{\phi}_{j'k} \frac{\lambda_J}{\mu_J} + \sum_{j' \notin J, k \in \partial(J)} \hat{\phi}_{j'k} \frac{\mu_J}{\lambda_J} + \left(1 - \sum_{j' \in J, k \notin \partial(J)} \hat{\phi}_{j'k} - \sum_{j' \notin J, k \in \partial(J)} \hat{\phi}_{j'k} \right) \\
&= \mu_J \frac{\lambda_J}{\mu_J} + \lambda_J \frac{\mu_J}{\lambda_J} + (1 - \lambda_J - \mu_J) \\
&= 1 .
\end{aligned}$$

Hence

$$(\mathbf{f}_J)_{j'k} = \begin{cases} \hat{\phi}_{j'k}(\lambda_J / \mu_J), & \text{for } j' \in J, k \notin \partial(J) \\ \hat{\phi}_{j'k}(\mu_J / \lambda_J), & \text{for } j' \notin J, k \in \partial(J) \\ \hat{\phi}_{j'k}, & \text{otherwise} \end{cases} .$$

Let $J^* = \operatorname{argmin}_{J \in \mathcal{J}} B_J \log(\lambda_J / \mu_J)$, then demand sample path with constant derivative $\mathbf{f}_{J^*}$ is the most likely sample path leading to demand drop. □

B.4.4 Transient behavior

Consider transient behavior over $[0, T]$ of our model with starting state $\mathbf{X}^K(0) \in \Omega_K$. We modify our objective appropriately: For any policy U which may be time dependent, we define

$$\mathbb{P}^{K,U}(\mathbf{X}^K(0), T) \triangleq \mathbb{E}\left(\frac{1}{A_\Sigma(T)} \sum_{r: t_r \in [0,T]} \mathbb{I}\left\{U^K_{t_r}[\mathbf{X}^{K,U}(t_r^-)](o[r], d[r]) = \emptyset\right\}\right), \tag{B.25}$$

where $A_\Sigma \triangleq \sum_{j' \in V_D, k \in V_S} A_{j'k}(T)$ is the total number of demand arrivals during $[0, T]$, t_r is the r-th demand arrival epoch. We then define

$$\gamma_{\mathrm{o}}(U) \triangleq -\liminf_{K\to\infty} \frac{1}{K} \log \mathbb{P}^{K,U}(\mathbf{X}^K(0), T), \tag{B.26}$$

$$\gamma_{\mathrm{p}}(U) \triangleq -\limsup_{K\to\infty} \frac{1}{K} \log \mathbb{P}^{K,U}(\mathbf{X}^K(0), T). \tag{B.27}$$

If $\gamma_{\mathrm{o}}(U) = \gamma_{\mathrm{p}}(U)$, we denote this value by $\gamma(U)$ and call it the exponent achieved by policy U.

Theorem B.1. *Fix any $\boldsymbol{\alpha} \in \mathrm{relint}(\Omega)$ and any $T \geq T_0$ for $T_0 = \frac{1}{v_{\boldsymbol{\alpha}}(\mathbf{f}^*)}$, where $v_{\boldsymbol{\alpha}}(\cdot)$ was defined in Lemma 2.1 and $\mathbf{f}^*$ is given by Lemma 2.4. Consider a sequence of initial states $\mathbf{X}^K(0) \in \Omega_K$ such that $\frac{\mathbf{X}^K(0)}{K} \xrightarrow{K\to\infty} \boldsymbol{\alpha}$ and transient behavior over $[0, T]$. Then, the SMW($\boldsymbol{\alpha}$) policy achieves exponent $\gamma(\boldsymbol{\alpha})$ as given by (2.13). No other policy can do better: for any policy U, we have $\gamma_{\mathrm{p}}(U) \leq \gamma_{\mathrm{o}}(U) \leq \gamma(\boldsymbol{\alpha})$.*

Sketch of proof of Theorem B.1. The converse bound $\gamma_{\mathrm{o}}(U) \leq \gamma(\boldsymbol{\alpha})$ follows from the proof of Lemma 2.1. The adversary (nature) can ensure at least this much demand loss by using the demand arrival rates $\mathbf{f}^*$ given in Lemma 2.4.

Achievability is straightforward to show. The sufficient conditions for exponent optimality in Proposition 2.4 (steepest descent and negative drift) apply to transient behavior starting at scaled state $\boldsymbol{\alpha}$ and for any finite horizon $T \geq 1/v_{\boldsymbol{\alpha}}(\mathbf{f}^*)$: The proof of the proposition goes through verbatim since it is fundamentally an argument about what happens over a finite horizon. It then remains to check that SMW($\boldsymbol{\alpha}$) satisfies these conditions, but we know this is true from Lemmas 2.2 and 2.3. □

B.5 Proof of Proposition 2.2 and appendix to Section 2.4.1

In this appendix, the first subsection provides the proof of Proposition 2.2 showing frequent utilization of supply units under SMW. The second subsection provides the structural corollaries (of Theorem 2.1) illustrated in Section 2.4.1.

B.5.1 Utilization rate of supply units: Proof of Proposition 2.2

Proof of Proposition 2.2. 1. Because supply units relocate only when assigned to an incoming demand, we have

$$\begin{aligned}\xi^{K,\boldsymbol{\alpha}} &= \frac{\mathbb{E}(\text{number of demand fulfilled in unit time in steady state})}{(\text{number of supply units})} \\ &= \frac{K \cdot \mathbf{1}^{\mathrm{T}}\hat{\boldsymbol{\phi}}\mathbf{1} - \mathbb{E}(\text{number of lost demand in unit time in steady state})}{K} \\ &\geq \mathbf{1}^{\mathrm{T}}\hat{\boldsymbol{\phi}}\mathbf{1} - \mathbb{P}_p^{K,\boldsymbol{\alpha}}\,,\end{aligned}$$

where $\mathbb{P}_p^{K,\boldsymbol{\alpha}}$ is the pessimistic demand loss probability defined in (2.2). Apply Theorem 2.1, we have for any $\boldsymbol{\alpha} \in \text{relint}(\Omega)$, $\lim_{K\to\infty} \xi^{K,\boldsymbol{\alpha}} = \mathbf{1}^{\mathrm{T}}\hat{\boldsymbol{\phi}}\mathbf{1} > 0$. Note that the above argument only uses the fact that the probability of losing demand is diminishing as $K \to \infty$, hence it holds with travel delays as well (apply Theorem 2.2).

2. *Sketch of proof.* The key observation is that under FIFO, if a supply unit is not assigned, neither do all the supply units that join the same queue later. Fix a supply unit which is the end-of-line unit in the i-th queue at time 0. Let ϵ, ζ be positive constants to be speficied later. Let $T' \triangleq \frac{4}{\epsilon\eta} + \max\{T_0, \frac{2}{\lambda_{\min}}\} > 0$ where η is the Lyapunov drift under SMW($\boldsymbol{\alpha}$) defined in Lemma 2.3, T_0 is defined in Theorem B.1, and $\lambda_{\min} \triangleq \min_{i\in V_S} \sum_{j'\in V_D} \hat{\phi}_{j'i}$. We have $\eta > 0$, $\lambda_{\min} > 0$, where the former is ensured by Lemma 2.3, and the latter holds because of Assumption 2.1. We

consider the time intervals $[0, T'), [T', 2T'), \cdots$.

In the following, we upper bound the probability that the fixed unit is not assigned during $[kT', (k+1)T')$ given it is not assigned during $[0, kT')$ (here $k \geq 0$). Let $kT' + \tau^K$ be the first time $L_{\boldsymbol{\alpha}}(\bar{\mathbf{X}}^K(t))$ hit level $\frac{\zeta}{K}$ or below during $[kT', (k+1)T')$. Define the following three events:

$$
\begin{aligned}
\mathcal{E}_1^K &\triangleq \left\{ \tau^K \leq \frac{4}{\epsilon\eta} \middle| \bar{\mathbf{X}}^K(kT') \right\}, \\
\mathcal{E}_2^K &\triangleq \left\{ L_{\boldsymbol{\alpha}}(\bar{\mathbf{X}}^K(t)) < 1 \text{ for all } t \in [kT' + \tau^K, (k+1)T' + \tau^K] \right\}, \\
\mathcal{E}_3^K &\triangleq \left\{ \sum_{j' \in V_D} (\bar{A}_{j'i}((k+1)T') - \bar{A}_{j'i}(kT' + \tau)) \leq \frac{3}{2} \right\},
\end{aligned}
$$

Note that if event $\mathcal{E}_1^K \cap \mathcal{E}_2^K \cap \mathcal{E}_3^K$ happens, then the fixed supply unit must be assigned during $[kT', (k+1)T')$: otherwise, the length of the i-th queue will exceed $\frac{3}{2}K$, which is impossible. Now we use union bound to lower bound $\mathcal{E}_1^K \cap \mathcal{E}_2^K \cap \mathcal{E}_3^K$. Using the argument in the proof of Theorem 4 in [45], there exists $\epsilon > 0, \zeta > 0$ independent of $\bar{\mathbf{X}}^K(kT')$ such that for large enough K, $\mathbb{E}[\tau^K] \leq \frac{1}{\epsilon\eta}$. Let the undetermined constants ϵ, ζ to be such ϵ, ζ. Using Markov's inequality we have $\mathbb{P}(\mathcal{E}_1^K) \geq 1 - \frac{1}{4} = \frac{3}{4}$. Using Theorem B.1 we have for large enough K, the probability of $\mathcal{E}_2^K$ converges to 1, hence $\mathbb{P}(\mathcal{E}_2^K) \geq \frac{3}{4}$ for large enough K. Using Chernoff bound of Poisson arrivals, we have for large enough K, $\mathbb{P}(\mathcal{E}_3^K) \geq \frac{3}{4}$. As a result,

$$
\mathbb{P}(\mathcal{E}_1^K \cap \mathcal{E}_2^K \cap \mathcal{E}_3^K) \geq \frac{1}{4}.
$$

Let $\omega_i^K(\mathbf{x})$ be the waiting time of the fixed supply unit given the (normalized) initial state converges to $\mathbf{x}$ as $K \to \infty$. Then for large enough K, we have

$$
\mathbb{E}[\omega_i^K(\mathbf{x})] \leq \sum_{k=0}^{\infty} \left(1 - \frac{1}{4}\right)^k \frac{1}{4}(k+1)T' = 4T' < \infty.
$$

This concludes the proof. Note that the above argument only uses the fact that the probability of losing demand is diminishing as $K \to \infty$, and that the Lypuanov drift in fluid limit is negative, hence it should hold with travel delays as well (apply Theorem 2.2).

□

B.5.2 Appendix to Section 2.4.1: optimal choice of scaling factors

The following corollary of Theorem 2.1 considers the case where there is exactly one vulnerable subset of demand nodes (Definition 2.3).

Corollary B.1 (If one subset of nodes is vulnerable, the optimal $\boldsymbol{\alpha}$ protects it)**.** *Fix a compatibility graph G. Consider a sequence of demand type distributions $(\boldsymbol{\phi}^n)_{n=1}^{\infty}$ satisfying the following properties:*

- *(Limiting distribution) There is a demand type distribution $\boldsymbol{\phi}^*$ such that $\lim_{n\to\infty}\boldsymbol{\phi}^n = \boldsymbol{\phi}^*$ and such that $(G,\boldsymbol{\phi}^*)$ satisfies Assumptions 2.1 and 2.2.*
- *(Vulnerable subset) There is a subset $J_1 \in \mathcal{J}^*$ such that $\lambda^*_{J_1} = \mu^*_{J_1}$, whereas for all other subsets $J \in \mathcal{J}^*\backslash J_1$, we have $\lambda^*_J > \mu^*_J$, cf. Assumption 2.3 (here λ^*_J, μ^*_J and $\mathcal{J}^*$ are the quantities under distribution $\boldsymbol{\phi}^*$). The distributions $\boldsymbol{\phi}^n$ satisfy Assumption 2.3; in particular, $\lambda^n_{J_1}/\mu^n_{J_1} \to 1^+$.*

Fix any $\epsilon \in (0,1/2)$. There exists $n_0 = n_0(\epsilon) < \infty$ such that, for all $n > n_0$, the following holds on network (G,ϕ^n):

(i) (Optimal exponent) The best achievable exponent $\bar{\gamma}$ satisfies

$$\bar{\gamma} \in [(1-\epsilon)\xi_{J_1}, \xi_{J_1}] \qquad \text{for } \xi_{J_1} \triangleq \log(\lambda^n_{J_1}/\mu^n_{J_1})\,.$$

As always, SMW policies suffice to achieve it, i.e., $\bar{\gamma} = \sup_{\boldsymbol{\alpha}\in \mathrm{relint}(\Omega)} \gamma(\boldsymbol{\alpha})$.

(ii) (Near optimal $\boldsymbol{\alpha}$ protects supply near J_1.) If SMW with scaling factors $\boldsymbol{\alpha} \in \mathrm{relint}(\Omega)$ achieves a demand-loss exponent $\gamma(\boldsymbol{\alpha}) \geq (1-\epsilon)\xi_{J_1}$, then it must be that

$$\mathbf{1}^T_{\partial(J_1)}\boldsymbol{\alpha} \geq 1-\epsilon\,.$$

(iii) (Example of near optimal $\boldsymbol{\alpha}$.) The SMW($\boldsymbol{\alpha}$) policy with

$$\alpha_i \triangleq \begin{cases} \frac{1-\epsilon}{|\partial(J_1)|} & \text{for all } i \in \partial(J_1)\,, \\ \frac{\epsilon}{m-|\partial(J_1)|} & \text{for all } i \in V_S \backslash \partial(J_1)\,. \end{cases} \tag{B.28}$$

achieves $\gamma(\boldsymbol{\alpha}) = (1-\epsilon)\xi_{J_1}$.

Informally speaking, Corollary B.1 says that if there is just one vulnerable subset of demand nodes J_1, then the exponent optimal SMW policy has a resting state which puts almost all the supply in the neighborhood of J_1. The intuition is that the supply at $\partial(J_1)$ follows a random walk which has only slightly positive drift even if the assignment rule protects it (recall that the definition of the net supply λ_{J_1} is optimistic), and hence it is optimal to keep the total supply at these nodes at a high resting point, to minimize the likelihood of depletion.

It is easy to verify that Example 2.2 satisfies the conditions in Corollary B.1: Note that in the example $\lim_{n\to\infty} \boldsymbol{\phi}^n = \boldsymbol{\phi}^*$ where $\boldsymbol{\phi}^*$ is given by (2.16) with δ_n replaced by 0 and η_n replaced by 1/8. Clearly, the limit demand type distribution $\boldsymbol{\phi}^*$ satisfies Assumptions 2.1 and 2.2, and $\boldsymbol{\phi}^n$ satisfies Assumption 2.3 for all $n > 4$. Furthermore, the limited-flexibility subset $\{4'\}$ is vulnerable, whereas all the other limited-flexibility subsets (namely, $\{1'\}$, $\{1', 2'\}$ and $\{3', 4'\}$) are not vulnerable.

We now prove the corollary.

Proof of Corollary B.1. We are given that $(G, \boldsymbol{\phi}^n)$ satisfies Assumption 2.3 for all $n \in \mathbb{Z}_+$. We start by showing that for all large enough n, we have that $(G, \boldsymbol{\phi}^n)$ also satisfies Assumptions 2.1 and 2.2: We are given that $(G, \boldsymbol{\phi}^*)$ satisfies Assumptions 2.1 and 2.2. For any demand type distribution $\boldsymbol{\phi}$, let the *support* of $\boldsymbol{\phi}$ be the set of demand types which occur with positive probability

$$\text{support}(\boldsymbol{\phi}) \triangleq \{(j', i) \in V_D \times V_S : \phi_{j'i} > 0\}\,.$$

Since $\lim_{n\to\infty} \boldsymbol{\phi}^n = \boldsymbol{\phi}^*$, it is clear that there exists n_0 such that for all $n > n_0$, the support of $\boldsymbol{\phi}^n$ is a superset of the support of $\boldsymbol{\phi}^*$, i.e., $\text{support}(\boldsymbol{\phi}^n) \supseteq \text{support}(\boldsymbol{\phi}^*)$. It is then clear from the form of Assumptions 2.1 and 2.2 that $(G, \boldsymbol{\phi}^n)$ satisfies them, given that $(G, \boldsymbol{\phi}^*)$ satisfies them (the assumptions are requirements on the *support* of the demand type distribution, and if a given distribution satisfies them, then it is easy to see that any distribution supported on a superset of demand types also satisfies them).

For all $n > n_0$, since $(G, \boldsymbol{\phi}^n)$ satisfies all three assumptions, Theorem 2.1 is applicable. From Theorem 2.1 part 1, we know $\gamma(\boldsymbol{\alpha}) \le \mathbf{1}^T_{\partial(J_1)}\boldsymbol{\alpha} \log(\lambda^n_{J_1}/\mu^n_{J_1}) = \mathbf{1}^T_{\partial(J_1)}\xi_{J_1}$. We deduce both part (ii) of the corollary, as well as $\bar{\gamma} \le \xi_{J_1}$ towards part (i) (to reach the latter conclusion we further use $\mathbf{1}^T_{\partial(J_1)}\boldsymbol{\alpha} \le 1$ and Theorem 2.3 part 2).

We now prove part (iii), namely, that for $\boldsymbol{\alpha}$ defined in (B.28), SMW($\boldsymbol{\alpha}$) achieves an exponent

$$\gamma(\boldsymbol{\alpha}) = (1-\epsilon)\log(\lambda^n_{J_1}/\mu^n_{J_1})\,. \tag{B.29}$$

(It will follow immediately that $\bar{\gamma} \ge (1-\epsilon)\log(\lambda^n_{J_1}/\mu^n_{J_1})$, completing the proof of part (i) as well.) We will again use Theorem 2.1 part 1 to establish (B.29). It is clear from the definition (B.28) that $\mathbf{1}^T_{\partial(J_1)}\boldsymbol{\alpha} = 1-\epsilon$ and hence $\mathbf{1}^T_{\partial(J_1)}\boldsymbol{\alpha}\log(\lambda^n_{J_1}/\mu^n_{J_1}) = (1-\epsilon)\log(\lambda^n_{J_1}/\mu^n_{J_1})$. Hence, to show that (B.29) holds, it suffices to show that we have

$$\mathbf{1}^T_{\partial(J)}\boldsymbol{\alpha} \cdot \log(\lambda^n_J/\mu^n_J) \ge (1-\epsilon)\log(\lambda^n_{J_1}/\mu^n_{J_1}) \tag{B.30}$$

for all $J \in \mathcal{J}^n \backslash \{J_1\}$. We will show that this holds for all large enough n.

Consider any $J \ne J_1$ such that $J \in \mathcal{J}^n$ for infinitely many n (if $J \in \mathcal{J}^n$ for finitely many n, we can eliminate it from consideration simply by taking n large enough). We will show that (B.30) holds for J for all n large enough. Note that for the chosen $\boldsymbol{\alpha}$ we have $\mathbf{1}^T_{\partial(J)}\boldsymbol{\alpha} \ge \epsilon/m > 0$ (since $|\partial(J)| \ge 1$, using Assumption 2.3), and so it suffices to show that

$$\liminf_{n\to\infty} \log(\lambda^n_J/\mu^n_J) > 0\,, \tag{B.31}$$

since the right-hand side of (B.30) tends to 0 as $n \to \infty$. (Here we define any positive number divided by 0 as ∞.) If $J \in \mathcal{J}^*$, it is easy to see that (B.31) holds: we know that $\lambda_J^n \to \lambda_J^*$ and $\mu_J^n \to \mu_J^* > 0$, and so $\log(\lambda_J^n/\mu_J^n) \to \log(\lambda_J^*/\mu_J^*) > 0$. To complete the proof consider the complementary case $J \notin \mathcal{J}^*$, i.e., $\mu_J^* = 0$. We will establish (B.31) by showing that $\lambda_J^* > 0$. Since $J \in \mathcal{J}^n$ for some $n > n_0$, by definition of $\mathcal{J}^n$ we know that $\partial(J)$ is a strict subset of V_S (else there cannot be a demand type with origin in J and destination in $V_S \backslash \partial(J)$). Consider any $i_1 \in V_S \backslash \partial(J)$ and any $i_2 \in \partial(J)$. Since we know that $\boldsymbol{\phi}^*$ satisfies Assumption 2.1, there is a path to move supply from i_1 to i_2, and so there must exist a demand type (j', k) with $j' \in V_D \backslash J$ and $k \in \partial(J)$ with $\phi_{j'k}^* > 0$, which immediately implies $\lambda_J^* > 0$. We deduce from $\lambda_J^n \to \lambda_J^* > 0$ and $\mu_J^n \to \mu_J^* = 0$ that $\log(\lambda_J^n/\mu_J^n) \to \infty$, and hence that (B.31) holds.

Since there are only finitely many subsets J to consider, we deduce from (B.31) that there exists n_0 such that, for all $n > n_0$, (B.30) holds for all $J \in \mathcal{J}^n \backslash \{J_1\}$. □

The second corollary considers the case of two non-overlapping vulnerable subsets of nodes.

Corollary B.2 (If there are two non-overlapping vulnerable subsets, the optimal $\boldsymbol{\alpha}$ protects them in inverse proportion to their inherent robustness)**.** *Fix a compatibility graph G. Consider a sequence of demand type distributions $(\boldsymbol{\phi}^n)_{n=1}^{\infty}$ satisfying the following properties:*

- *(Limiting distribution) There is a demand type distribution $\boldsymbol{\phi}^*$ such that* $\lim_{n\to\infty} \boldsymbol{\phi}^n = \boldsymbol{\phi}^*$ *and such that $(G, \boldsymbol{\phi}^*)$ satisfies Assumptions 2.1 and 2.2.*

- *(Vulnerable subsets) There are two non-overlapping subsets $J_1, J_2 \in \mathcal{J}^*, J_1 \cap J_2 = \emptyset, \partial(J_1) \cap \partial(J_2) = \emptyset$ such that $\lambda_{J_1}^* = \mu_{J_1}^*$ and $\lambda_{J_2}^* = \mu_{J_2}^*$, whereas for all other subsets $J \in \mathcal{J}^* \backslash \{J_1, J_2\}$, we have $\lambda_J^* > \mu_J^*$, cf. Assumption 2.3 (here λ_J^*, μ_J^* and $\mathcal{J}^*$ are the quantities under distribution $\boldsymbol{\phi}^*$). The distributions $\boldsymbol{\phi}^n$ satisfy Assumption 2.3; in particular, $\lambda_{J_1}^n/\mu_{J_1}^n \to 1^+$ and $\lambda_{J_2}^n/\mu_{J_2}^n \to 1^+$.*

Fix any $\epsilon \in (0, 1/2)$. There exists $n_0 = n_0(\epsilon) < \infty$ such that, for all $n > n_0$, the following holds on network (G, ϕ^n):

(i) *(Optimal exponent) The best achievable exponent* $\bar{\gamma}$ *satisfies*

$$\bar{\gamma} \in [(1-\epsilon)H, H] \qquad \text{for } H \triangleq \frac{\xi_{J_1}\xi_{J_2}}{\xi_{J_1}+\xi_{J_2}}\,,\ \xi_J \triangleq \log(\lambda_J^n/\mu_J^n)\,.$$

As always, SMW policies suffice to achieve it, i.e., $\bar{\gamma} = \sup_{\boldsymbol{\alpha}\in \mathrm{relint}(\Omega)} \gamma(\boldsymbol{\alpha})$.

(ii) *(Near optimal* $\boldsymbol{\alpha}$ *protects supply near* J_1*.) If SMW with scaling factors* $\boldsymbol{\alpha} \in \mathrm{relint}(\Omega)$ *achieves a demand-loss exponent* $\gamma(\boldsymbol{\alpha}) \geq (1-\epsilon)H$*, then it must be that*

$$\mathbf{1}_{\partial(J_1)}^T\boldsymbol{\alpha} \overset{\epsilon}{=} \frac{\xi_{J_2}}{\xi_{J_1}+\xi_{J_2}} \qquad \text{and} \qquad \mathbf{1}_{\partial(J_2)}^T\boldsymbol{\alpha} \overset{\epsilon}{=} \frac{\xi_{J_1}}{\xi_{J_1}+\xi_{J_2}}\,,$$

where $a \overset{\epsilon}{=} b$ *represents* $|a-b| \leq \epsilon$.

(iii) *(Example of near optimal* $\boldsymbol{\alpha}$*.) The SMW(*$\boldsymbol{\alpha}$*) policy with*

$$\alpha_i \triangleq \begin{cases} \frac{1-\epsilon_1}{|\partial(J_1)|}\cdot\frac{\xi_{J_2}}{\xi_{J_1}+\xi_{J_2}} & \text{for all } i \in \partial(J_1)\,, \\ \frac{1-\epsilon_1}{|\partial(J_2)|}\cdot\frac{\xi_{J_1}}{\xi_{J_1}+\xi_{J_2}} & \text{for all } i \in \partial(J_2)\,, \\ \frac{\epsilon}{m-|\partial(J_1)|-|\partial(J_2)|} & \text{for all } i \in V_S\backslash(\partial(J_1)\cup\partial(J_2)) \end{cases} \tag{B.32}$$

for $\epsilon_1 \triangleq \epsilon\cdot\mathbb{I}\big(V_S\backslash(\partial(J_1)\cup\partial(J_2)) \neq \emptyset\big)$*, achieves* $\gamma(\boldsymbol{\alpha}) \geq (1-\epsilon)H$.

Corollary B.2 says that if there are two non-overlapping vulnerable subsets of demand nodes J_1 and J_2, then the exponent optimal SMW policy has a resting state (i) which puts almost all the supply in the union of their neighborhoods $\partial(J_1)\cup\partial(J_2)$, (ii) divides the supply between the two neighborhoods in inverse proportion to the inherent robustness of the vulnerable subsets

$$\frac{\mathbf{1}_{\partial(J_2)}^T\boldsymbol{\alpha}}{\mathbf{1}_{\partial(J_1)}^T\boldsymbol{\alpha}} \approx \frac{\xi_{J_1}}{\xi_{J_2}}\,.$$

Example 2.3 follows from Corollary B.2: Clearly, the limit demand type distribution $\boldsymbol{\phi}^*$ in the example satisfies Assumptions 2.1 and 2.2, and $\boldsymbol{\phi}^n$ satisfies Assumption 2.3 for all $n > 4/\min(1,\eta)$. Furthermore, the limited-flexibility subsets $\{1'\}$ and $\{4'\}$ are non-overlapping and vulnerable, whereas all the other limited-flexibility subsets (namely, $\{1',2'\}$ and $\{3',4'\}$) are not vulnerable. Note that $V_S\backslash(\partial(J_1)\cup\partial(J_2)) = \emptyset$ and hence $\epsilon_1 = 0$ in the example.

We now prove the corollary.

Proof of Corollary B.2. The proof is analogous to that of Corollary B.1.

From Theorem 2.1 part we know that for any α, it holds that

$$\gamma(\alpha) \le \mathbf{1}_{\partial(J_1)}^T \boldsymbol{\alpha} \cdot \xi_{J_1} \quad \text{and} \quad \gamma(\alpha) \le \mathbf{1}_{\partial(J_2)}^T \boldsymbol{\alpha} \cdot \xi_{J_2} \,. \tag{B.33}$$

Since $\partial(J_1) \cap \partial(J_2) = \emptyset$, we know that

$$\mathbf{1}_{\partial(J_1)}^T \boldsymbol{\alpha} + \mathbf{1}_{\partial(J_2)}^T \boldsymbol{\alpha} \le \mathbf{1}^T \boldsymbol{\alpha} = 1 \,.$$

We then deduce from (B.33) that

$$\gamma(\alpha) \le H = \frac{\xi_{J_1}\xi_{J_2}}{\xi_{J_1} + \xi_{J_2}} \,.$$

holds for all $\boldsymbol{\alpha} \in \text{relint}(\Omega)$, and hence, using Theorem 2.1 part 2, we obtain $\bar{\gamma} \le H$. This is the upper bound in part (i) of the corollary.

We now prove part (ii). If $\gamma(\boldsymbol{\alpha}) \ge (1-\epsilon)H$ then using (B.33) we have

$$\begin{aligned} & \mathbf{1}_{\partial(J_1)}^T \boldsymbol{\alpha} \cdot \xi_{J_1} \ge (1-\epsilon) \cdot \frac{\xi_{J_1}\xi_{J_2}}{\xi_{J_1} + \xi_{J_2}} \\ \Rightarrow \qquad & \mathbf{1}_{\partial(J_1)}^T \boldsymbol{\alpha} \ge (1-\epsilon) \cdot \frac{\xi_{J_2}}{\xi_{J_1} + \xi_{J_2}} \ge \frac{\xi_{J_2}}{\xi_{J_1} + \xi_{J_2}} - \epsilon \end{aligned} \tag{B.34}$$

and similarly

$$\mathbf{1}_{\partial(J_2)}^T \boldsymbol{\alpha} \ge \frac{\xi_{J_1}}{\xi_{J_1} + \xi_{J_2}} - \epsilon \,. \tag{B.35}$$

But (B.35) further implies

$$\mathbf{1}_{\partial(J_1)}^T \boldsymbol{\alpha} \le 1 - \mathbf{1}_{\partial(J_2)}^T \boldsymbol{\alpha} \le \frac{\xi_{J_2}}{\xi_{J_1} + \xi_{J_2}} + \epsilon \,.$$

Combining with (B.34) we have shown $\mathbf{1}_{\partial(J_1)}^T \boldsymbol{\alpha} \overset{\epsilon}{=} \frac{\xi_{J_2}}{\xi_{J_1}+\xi_{J_2}}$, and analogously obtain $\mathbf{1}_{\partial(J_2)}^T \boldsymbol{\alpha} \overset{\epsilon}{=} \frac{\xi_{J_1}}{\xi_{J_1}+\xi_{J_2}}$. This completes the proof of part (ii).

It remains to show part (iii) which will further imply the lower bound $\bar{\gamma} \ge H(1-\epsilon)$ in part (i). Part (iii) states that $\boldsymbol{\alpha}$ defined in (B.32), we have $\gamma(\boldsymbol{\alpha}) \ge (1-\epsilon)H$ for large enough n. Using Theorem 2.1 part 1, it suffices to show that for large enough n, we have

$$\mathbf{1}_{\partial(J)}^T \boldsymbol{\alpha} \xi_J \geq (1-\epsilon)H \tag{B.36}$$

for all $J \in \mathcal{J}^n$. For $J = J_1$, it clear that the left-hand side of (B.36) is $(1-\epsilon_1)H \geq (1-\epsilon)H$, and similarly for J_2. It remains to consider the other subsets. Note that $H \xrightarrow{n\to\infty} 0$. Now to prove that for large enough n, (B.36) holds for all $J \in \mathcal{J}^n \backslash \{J_1, J_2\}$, we can use the proof of (B.30) (in the proof of Corollary B.1) verbatim.

□

B.6 Necessity of the Assumptions and the Inferiority of State-Independent Control

This section shows the necessity of our assumptions, and of state-dependent control, including the proofs of Propositions B.1, 2.1 and 2.3. It also demonstrates poor performance of the naive state-dependent policy by establishing the claim in Example 2.4.

B.6.1 Necessity of Assumption 2.2: Proof of Proposition B.1

Proposition B.1 (Ample flexibility renders the control problem trivial)**.** *Consider any network* $(G, \boldsymbol{\phi})$ *which satisfies Assumption 2.1 and such that for all* $j' \in V_D$ *and* $k \in V_S$ *such* $\phi_{j'k} > 0$ *it holds that* $k \in \partial(j')$. *Then for any* $K \geq n \triangleq |V_D|$, *there is a control policy which loses an identically zero fraction of demand in the long run. Formally, there is a policy* U *such that* $\mathbb{P}_{\mathrm{p}}^{K,U} = 0$, *for* $\mathbb{P}_{\mathrm{p}}^{K,U}$ *defined in* (2.2) *below.*

Proof of Proposition B.1. We define the following policy U which ensures no demand loss in the long run, i.e., $\mathbb{P}_{\mathrm{p}}^{K,U} = 0$. Arbitrarily choose n of the K supply units and dedicate one of the chosen supply units to each of the demand nodes. Suppose the supply unit dedicated to demand node j' is initially at supply node i. Since Assumption 2.1 is satisfied, there is a way to move the supply unit from i to a supply node compatible with

j' in a finite (random) time. Move the supply unit to some node in $\partial(j')$. Similarly, move each of the n dedicated demand units into the neighborhood of the corresponding demand node. All this is completed in an initial transient of finite (random) duration (the expected duration is also finite). Thereafter, for each demand arrival, use the supply unit dedicated to the origin of the demand to serve it. We are guaranteed that the destination $k \in \partial(j')$, i.e., the supply unit remains within the neighborhood of j' after completing service (we are told that demand types with $k \notin \partial(j')$ have zero arrival rate $\phi_{j'k} = 0$). □

B.6.2 Necessity of CRP Condition: Proof of Proposition 2.1

Proof of Proposition 2.1. There are two cases:

Case 1: *There exists* $J \subsetneq V_D$ *s.t.* $\lambda_J < \mu_J \iff \sum_{i\in\partial(J)} \mathbf{1}^{\mathrm{T}}\phi_{(i)} < \sum_{j'\in J} \mathbf{1}^{\mathrm{T}}\phi_{j'}$.

The main proof idea in this case is simply that since the net supply to $\partial(J)$ is less than the net demand originating in J, a positive fraction of demand must be lost.

Consider the following balance equation:

$$\begin{aligned}
&\#\{\text{demands originating in } J \text{ during } [0,T] \text{ which are lost}\}\\
&= \#\{\text{demands originating in } J \text{ during } [0,T]\}\\
&\qquad - \#\{\text{demands originating in } J \text{ during } [0,T] \text{ which are fulfilled}\}\\
&\geq \#\{\text{demands originating in } J \text{ during } [0,T]\} - \#\{\text{supplies assigned from } \partial(J) \text{ during } [0,T]\}\\
&\geq \sum_{r:t_r\in[0,T]} \mathbb{I}\{o[r] \in J\} - \sum_{r:t_r\in[0,T]} \mathbb{I}\{d[r] \in \partial(J)\} - \#\{\text{initial supply in } \partial(J)\}\,.
\end{aligned}$$

The first inequality holds because the demands originating in J can only be fulfilled by supply units from $\partial(J)$. The second inequality holds because the total number of supply units assigned from $\partial(J)$ during $[0,T]$ cannot exceed the initial supply there plus the number of demand arrivals with destination in $\partial(J)$. Divide both sides by $A_\Sigma(T)$ which is the total number of demand arrivals during $[0,T]$, and let $T \to \infty$. By the strong law of large numbers, we have:

$$\liminf_{T\to\infty}\{\text{fraction of lost in } [0,T]\} \geq \sum_{j'\in J} \mathbf{1}^{\mathrm{T}}\phi_{j'} - \sum_{i\in\partial(J)} \mathbf{1}^{\mathrm{T}}\phi_{(i)} > 0\,.$$

Hence a positive fraction of demand will be lost in the long run, and the loss exponent is 0.

Case 2: *We have* $\lambda_{J'} \geq \mu_{J'}$ *for all* $J' \in \mathcal{J}$ *but there exists* $J \in \mathcal{J}$ *such that* $\lambda_J = \mu_J \Leftrightarrow \sum_{i\in\partial(J)} \mathbf{1}^{\mathrm{T}}\phi_{(i)} = \sum_{j'\in J} \mathbf{1}^{\mathrm{T}}\phi_{j'}$.

The high-level idea in this case is that if all the demand originating in J is served (if possible), then, at best, the total quantity of supply in $\partial(J)$ follows an unbiased random walk on $0, 1, \ldots, K$. Such a random walk spends a positive fraction of time at 0, and all demand originating in J when there is zero supply in $\partial(J)$ is lost. The proof is somewhat more intricate than this argument may suggest; in particular because we need to allow for idling policies (those which sometimes lose demand even though supply is available at a neighboring node).

Divide the demand arrivals into cycles with MK^2 arrivals each, where

$$M \triangleq \frac{1}{\mu_J},$$

for $\mu_J = \sum_{j'\in J, k\notin\partial(J)} \phi_{j'k} > 0$ as before. Without loss of generality, consider the first cycle $t_1, \cdots, t_{MK^2}$. Define random walk S_r with the following dynamics:

- $S_0 = \mathbf{1}^{\mathrm{T}}_{\partial(J)}\mathbf{X}(0)$.
- $S_{r+1} = S_r + 1$ if $o[r] \notin J, d[r] \in \partial(J)$.
- $S_{r+1} = S_r - 1$ if $o[r] \in J, d[r] \notin \partial(J)$.
- $S_{r+1} = S_r$ otherwise.

It is not hard to see that if no demand is lost during $r \leq MK^2$ under some policy U, then S_r is a pathwise upper bound on the number of supply units in $\partial(J)$, namely, $\mathbf{1}^{\mathrm{T}}_{\partial(J)}\mathbf{X}(t_r)$, for any $r \leq MK^2$. With this observation, we have:

$$\begin{aligned} &\mathbb{P}\left(\text{some demand is lost during } r \leq MK^2\right) \\ &\geq \mathbb{P}\left(S_{r'} = 0 \text{ for some } r' < MK^2\right) \cdot \left(\mathbf{1}^T\phi_{j'}\right). \end{aligned} \tag{B.37}$$

The above is true because when the event on RHS happens, either (1) some demand is lost before t'_r, or (2) no demand is lost before $t_{r'}$, then since $0 = S_{r'} \geq \mathbf{1}^{\mathrm{T}}_{\partial(J)}\mathbf{X}(t_{r'})$ we have $\mathbf{1}^{\mathrm{T}}_{\partial(J)}\mathbf{X}(t_{r'}) = 0$ and so any demand with origin in J is lost at $t_{r'+1}$. Importantly, (B.37) holds for *any* policy.

For the given J we have $\lambda_J = \mu_J > 0$ and so S_r is a "lazy" simple random walk, which takes a step with probability $2\mu_J$ independently at each r. Define the stopping time τ as

$$\tau \triangleq \inf\left\{r \in \mathbb{Z}_+ : S_r \in \left\{\mathbf{1}^{\mathrm{T}}_{\partial(J)}\mathbf{X}(0) - K, \mathbf{1}^{\mathrm{T}}_{\partial(J)}\mathbf{X}(0) + K\right\}\right\} .$$

Using [Example 4.1.6, 126] on the lazy simple random walk $S_r - \mathbf{1}^{\mathrm{T}}_{\partial(J)}\mathbf{X}(0)$, we obtain[3]

$$\mathbb{E}[\tau] = \frac{K^2}{2\mu_J} .$$

Using Markov's inequality, we have

$$\mathbb{P}\left(\tau \geq MK^2\right) \leq \frac{\mathbb{E}[\tau]}{MK^2} = \frac{1}{2}$$

By symmetry

$$\mathbb{P}\left(S_\tau - \mathbf{1}^{\mathrm{T}}_{\partial(J)}\mathbf{X}(0) = -K \text{ and } \tau < mK^2\right) = \frac{1}{2}\mathbb{P}\left(\tau < mK^2\right) \geq \frac{1}{2}\cdot\frac{1}{2} = \frac{1}{4} . \qquad \text{(B.38)}$$

Now, $S_\tau - \mathbf{1}^{\mathrm{T}}_{\partial(J)}\mathbf{X}(0) = -K$ and $\tau < MK^2$, i.e., S_r hits $\mathbf{1}^{\mathrm{T}}_{\partial(J)}\mathbf{X}(0) - K$ during $r < MK^2$, implies that S_r hits 0 during $t < MK^2$, since S_r must hit 0 (weakly) before it hits $\mathbf{1}^{\mathrm{T}}_{\partial(J)}\mathbf{X}(0) - K$. Hence, plugging (B.38) into (B.37) we obtain that

$$\mathbb{P}\left(\text{some demand is lost during } r \leq MK^2\right) \geq \frac{\mathbf{1}^{\mathrm{T}}\phi_{j'}}{4} ,$$

and this uniform and strictly positive lower bound holds for *any* policy, during *any* cycle consisting of MK^2 consecutive arrivals.

It follows that

$$\mathbb{P}^{K,U}_{\mathrm{p}} \geq \mathbb{P}^{K,U}_{\mathrm{o}} = \Omega\left(\frac{1}{K^2}\right) ,$$

[3] Since $S_r - \mathbf{1}^{\mathrm{T}}_{\partial(J)}\mathbf{X}(0)$ is a lazy version of a simple random walk, which takes a step with probability $2\mu_J$ independently at each time, the expectation of the time τ to hit $\pm K$ is inflated by a factor of $1/(2\mu_J)$ relative to that of a simple random walk (this follows from using the natural coupling between the steps in the two walks, and noting that the lazy walk takes expected time $1/(2\mu_J)$ between consecutive steps).

and hence $\gamma_{\mathrm{p}}(U) = \gamma_{\mathrm{o}}(U) = 0$ for any U. □

B.6.3 Necessity of State-Dependent Control: Proof of Proposition 2.3

Proof of Proposition 2.3.

- Proof of first part. For notation simplicity, denote $X(t_r)$ by $X[r]$, similar for another notations. Denote the probability mass function of distribution $u_{j'k}[t]$ by $u_{j'k}[t](\cdot)$. We first define an "augmented" policy $\tilde{U}$ for any state-independent policy U. Policy $\tilde{U}$ is also state independent with distribution $\tilde{u}_{j'k}[t]$, where:

  $$\tilde{u}_{j'k}[t](i) = u_{j'k}[t](i) + \frac{1}{|\partial(j')|} u_{j'k}[t](\emptyset) \quad \text{for } i \in \partial(j')\,,$$
  $$\tilde{u}_{j'k}[t](\emptyset) = 0\,.$$

  In the following analysis, we couple U and $\tilde{U}$ in such a way that if U dispatches from i to serve the t-th demand, then $\tilde{U}$ will do the same.

  Divide the demand arrivals into cycles with K^2 arrivals each. We will lower bound the probability of demand loss in any cycle. Without loss of generality, consider the first cycle $[1, K^2]$. Suppose $\mathbf{X}^{K,U}[0] = \mathbf{X}_0$. By Assumption 2.2, $\exists j' \in V_D,\ k \notin \partial(j') \subset V_S$ such that $\phi_{j'k} > 0$. Consider the random walk S_t with the following dynamics, which is the "virtual" net change of supply in $\partial(j')$:

  - $S_0 = 0$.
  - $S_{t+1} = S_t + 1$ if $d[t] \in \partial(j')$ and policy $\tilde{U}$ assigns a supply unit from outside of $\partial(j')$ to serve it (regardless of whether there is available supply to assign).
  - $S_{t+1} = S_t - 1$ if $d[t] \notin \partial(j')$ and policy $\tilde{U}$ assigns a supply unit from $\partial(j')$ to serve it (regardless of whether there is available supply to assign).
  - $S_{t+1} = S_t$ if otherwise.

Using similar argument as in eq. (B.37), we have

$$\begin{aligned}&\mathbb{P}\left(\text{some demand is lost in epoch } [1, K^2]\right)\\ \geq\ &\mathbb{P}\left(S_{K^2} + \mathbf{1}_{\partial(j')}^{\mathrm{T}} \mathbf{X}_0 > K \text{ or } < 0\right) \geq \mathbb{P}\left(|S_{K^2}| > K\right) .\end{aligned} \tag{B.39}$$

Note that S_{K^2} is the sum of K^2 independent random variables Z_t, where $Z_t = S_t - S_{t-1}$. Here independence holds because we ignore demand losses in the definition of the process. Here Z_t has support $\{-1, 0, 1\}$ and satisfies:

$$\mathbb{P}(Z_t = -1) \geq \delta \triangleq \phi_{j'k} > 0\,, \tag{B.40}$$

where $k \notin \partial(j)$. There are two cases:

1. If $\mathbb{E}[S_{K^2}] \leq -\frac{K^2}{2}$, then for $K \geq 8$, we have

$$\begin{aligned}&1 - \mathbb{P}\left(S_{K^2} \in [-K, K]\right)\\ \geq\ &1 - \mathbb{P}\left(S_{K^2} - \mathbb{E}[S_{K^2}] \geq -K + \frac{K^2}{2}\right)\\ \geq\ &1 - 2\exp\left(-\frac{K^2}{32}\right) \qquad \text{(Hoeffding's inequality, } -K + K^2/2 \geq K^2/4)\\ \geq\ &\frac{1}{2}\,.\end{aligned}$$

   Plugging into (B.39) establishes that demand is lost with likelihood at least 1/2.

2. If $\mathbb{E}[S_{K^2}] > -\frac{K^2}{2}$, then using linearity of expectation and simple algebra we obtain that the number of t's such that $\mathbb{E}[Z_t] \geq -\frac{3}{4}$ is at least $\frac{K^2}{7}$.

   Denote the set of these t's as $\mathcal{T}$. Hence

$$K^2 \geq \mathrm{Var}(S_{K^2}) = \sum_{t=1}^{K^2} \mathrm{Var}(Z_t) \geq \sum_{t\in\mathcal{T}} \mathrm{Var}(Z_t) \geq \frac{K^2}{7} \cdot \delta \left(1 - \frac{3}{4}\right)^2 = \frac{\delta}{102} K^2\,, \tag{B.41}$$

   using (B.40).

   Note from (B.39) that to show a constant lower bound of demand-loss probability on $[1, K^2]$, it suffices to derive a uniform upper bound on $\mathbb{P}\left(S_{K^2} \in [-K, K]\right)$ that is strictly smaller than 1. To this end, apply Theorem 7.4.1 in [127] (Berry-Esseen

Theorem) to obtain:

$$\sup_{x\in\mathbb{R}}\left|\mathbb{P}\left(S_{K^2}-\mathbb{E}[S_{K^2}]\le x\sqrt{\mathrm{Var}[S_{K^2}]}\right)-\Phi(x)\right|$$
$$\le \frac{\sum_{t=1}^{K^2}\mathbb{E}|Z_t-\mathbb{E}Z_t|^3}{\left(\mathrm{Var}[S_{K^2}]\right)^{3/2}}\le \frac{5000}{K\delta^{3/2}}\,, \tag{B.42}$$

where $\Phi(\cdot)$ is the cumulative distribution function of the standard normal distribution.

Denote $B(x,a)\triangleq[x-a,x+a]$. Note that there are two subcases (indexed 2(i) and 2(ii)):

$$[-K,K]\subset B\left(\mathbb{E}[S_{K^2}],4K\right)\,,\qquad [-K,K]\cap B\left(\mathbb{E}[S_{K^2}],2K\right)=\emptyset\,.$$

In subcase 2(i),

$$\mathbb{P}\left(S_{K^2}\in[-K,K]\right)\le\mathbb{P}\left(S_{K^2}\in B\left(\mathbb{E}[S_{K^2}],4K\right)\right)\,,$$

whereas in subcase 2(ii),

$$\mathbb{P}\left(S_{K^2}\in[-K,K]\right)\le 1-\mathbb{P}\left(S_{K^2}\in B\left(\mathbb{E}[S_{K^2}],2K\right)\right)\,.$$

Hence

$$\mathbb{P}\left(S_{K^2}\in[-K,K]\right)$$
$$\le\max\left\{\mathbb{P}\left(S_{K^2}\in B\left(\mathbb{E}[S_{K^2}],4K\right)\right),1-\mathbb{P}\left(S_{K^2}\in B\left(\mathbb{E}[S_{K^2}],2K\right)\right)\right\}\,. \tag{B.43}$$

Use (B.42) and $\mathrm{Var}(S_{K^2})\le K^2$ to obtain

$$\begin{aligned}\mathbb{P}\left(S_{K^2}\in B\left(\mathbb{E}[S_{K^2}],4K\right)\right)&\le\mathbb{P}\left(S_{K^2}-\mathbb{E}[S_{K^2}]\le\sqrt{\mathrm{Var}[S_{K^2}]}\frac{4K}{\sqrt{\mathrm{Var}[S_{K^2}]}}\right)\\&\le 5000\delta^{-3/2}K^{-1}+\Phi\left(\frac{4K}{\sqrt{\mathrm{Var}[S_{K^2}]}}\right)\\&\le 5000\delta^{-3/2}K^{-1}+\Phi\left(50\delta^{-1/2}\right)\,,\end{aligned}$$

$$1-\mathbb{P}\left(S_{K^2}\in B\left(\mathbb{E}[S_{K^2}],2K\right)\right)$$

$$
\begin{aligned}
&= \mathbb{P}\left(S_{K^2} - \mathbb{E}[S_{K^2}] \le \sqrt{\mathrm{Var}[S_{K^2}]}\frac{-2K}{\sqrt{\mathrm{Var}[S_{K^2}]}}\right) \\
&\quad + \mathbb{P}\left(S_{K^2} - \mathbb{E}[S_{K^2}] \ge \sqrt{\mathrm{Var}[S_{K^2}]}\frac{2K}{\sqrt{\mathrm{Var}[S_{K^2}]}}\right) \\
&\le 10000\delta^{-3/2}K^{-1} + 2\Phi\left(\frac{-2K}{\sqrt{\mathrm{Var}[S_{K^2}]}}\right) \\
&\le 10000\delta^{-3/2}K^{-1} + 2\Phi\left(-2\right) .
\end{aligned}
$$

Hence for $K > \max\left\{\frac{10000\delta^{-3/2}}{\bar{\Phi}\left(50\delta^{-1/2}\right)}, \frac{10000\delta^{-3/2}}{\frac{1}{2}-\Phi(-2)}\right\}$, plugging into (B.43) and then into (B.39), we obtain

$$
\mathbb{P}(\text{some demand is lost in } [1, K^2]) \ge \min\left\{\frac{1}{2}\bar{\Phi}\left(50\delta^{-1/2}\right), \frac{1}{2} - \Phi(-2)\right\} > 0 .
$$

Since we obtained a uniform lower bound on the likelihood of dropping demand in both cases, we conclude that the steady state demand-loss probability is $\Omega(1/K^2)$ as $K \to \infty$.

- Proof of second part. Consider any $k \in V_S$ such that $\exists j' \in V_D$ such that $(j', k) \in S$. Given a demand type distribution $\boldsymbol{\phi} \in D(S)$, suppose U achieves asymptotic optimality $\mathbb{P}_{\mathrm{o}}^{K,U} = o(1)$, i.e., $1-o(1)$ fraction of demand is served. This implies that a fraction $\sum_{j' \in V_D:(j',k)\in S} \phi_{j'k} - o(1)$ of demand has destination k *and* is served under U. And that a fraction $\sum_{(j',i)\in S} \phi_{j'i}u_{j'k} - o(1)$ of demand is assigned a supply unit from k *and* is served under U. (Our proof will focus on the case where $u_{j'k}$ is time invariant and independent of K. The proof for the general case of time varying $u_{j'k}(t)$ which can depend on K is very similar, though the latter fraction can now vary over time, increasing the notational burden. We omit the details.) But in steady state, the inflow of supply units to node k must be equal to the outflow of supply units, i.e., it must be that

$$
\sum_{j' \in V_D:(j',k)\in S} \phi_{j'k} = \sum_{(j',i)\in S} \phi_{j'i}u_{j'k} .
$$

This is a knife edge requirement. In particular, the set of $\boldsymbol{\phi} \in D(S)$ which *do not* satisfy this condition is clearly an open and dense subset of $D(S)$. For all such $\boldsymbol{\phi}$, the above argument implies that $\liminf_{K\to\infty} \mathbb{P}_{\mathrm{o}}^{K,U} > 0$, completing the proof.

□

B.6.4 Proof of Example 2.4

We will prove by contradiction that the naive policy incurs an $\Omega(1)$ loss. Suppose the loss is vanishing $\mathbb{P}_{\mathrm{o}}^{K} = o(1)$, i.e., all but a $o(1)$ fraction of demands are served. Consider the subset of supply nodes $\{3,4\}$ (demand type $(4'1)$ is entirely dependent on this subset). We will show that supply units arrive at these nodes slower than they are assigned from these nodes, which cannot possibly be the case in steady state: The fraction of demands which lead to a supply unit arriving to $\{3,4\}$ is at most $\sum_{j'\in V_D}\sum_{k\in\{3,4\}}\phi_{j'k} = \phi_{1'3}+\phi_{1'4} = 0.42$. All demands of type $(4'1)$ which are served are assigned a supply unit from $\{3,4\}$. Since all but $o(1)$ fraction of demands of type $(4'1)$ are served:

(i) There is a supply unit present in at least one of $\{3,4\}$ a $1-o(1)$ fraction of the time.

(ii) A fraction of demands $0.4-o(1)$ are of type $(4'1)$ *and* are assigned a supply unit from $\{3,4\}$.

Now consider demands of type $(3'2)$: When such a demand arrives, using point (i) above, with probability $1-o(1)$ there is a supply unit present in at least one of $\{3,4\}$. The other compatible supply (with the origin $3'$) is 2. In all cases where there is a supply unit present in at least one of $\{3,4\}$, the naive policy assigns a supply unit from one of $\{3,4\}$ with probability at least $1/2$, by definition of the policy. It follows that a fraction $1/2-o(1)$ of demands of type $(3'2)$ are assigned a supply unit from one of $\{3,4\}$, and hence a fraction $0.1\times 1/2-o(1) = 0.05-o(1)$ of demands are of type $(3'2)$ *and* are assigned a supply unit from one of $\{3,4\}$. In total (adding across the demand types $(4'1)$ and $(3'2)$), a supply unit from one of $\{3,4\}$ is assigned to serve at least a fraction $0.45-o(1)$ of all demand.

But this (minimum possible) "outflow rate" exceeds the maximum possible "inflow rate" of 0.42 established above, which is impossible in steady state. Thus we have obtained a contradiction. We infer that the naive policy incurs an $\Omega(1)$ loss in this network. We further observe that both the (minimum possible) outflow rate and the maximum possible inflow rate are continuous in $\boldsymbol{\phi}$, hence the above argument goes through for any demand type distribution which is sufficiently close to $\boldsymbol{\phi}$ given by (2.18).

B.7 Extension to Scrip Systems: Proof of Theorem 2.3

The proof of Theorem 2.3 is almost identical to the proof of Theorem 2.1. To avoid redundancy, we skip the parts of the proof which are mere repetitions of their counterparts in the proof of Theorem 2.1.

Proof of Theorem 2.3. Recall that the converse result in Theorem 2.1 follows from Lemmas 2.1 and 2.4, the achievablity result follows from Lemmas 2.2, 2.3, 2.4 and Proposition 2.4.

Here we can prove a result identical to Lemma 2.1 except that $v_{\boldsymbol{\alpha}}(\mathbf{f})$ is now defined as

$$v_\alpha(\mathbf{f}) \triangleq \min_{\Delta\mathbf{x}\in\mathcal{X}'_{\mathbf{f}}} L_{\boldsymbol{\alpha}}(\boldsymbol{\alpha}+\Delta\mathbf{x})\,,$$

where

$$\mathcal{X}'_{\mathbf{f}} \triangleq \left\{ \Delta\mathbf{x} \;\middle|\; \begin{array}{ll} \Delta x_i = \sum_{j'\in\partial(i)} d_{ij'}\left(\sum_{k\in V_S} f_{kj'}\right) - \sum_{j'\in V_D} f_{ij'}, & \forall i\in V_S \\ \sum_{i\in\partial(j')} d_{ij'} = 1,\ d_{ij'}\geq 0, & \forall i\in V_S, j'\in V_D \end{array} \right\}.$$

Here $(d_{ij'})_{i\in\partial(j')}$ is the chosen *service provider distribution* over agents neighboring j' for assigning agents to serve demand of service j'. Lemmas 2.2, 2.3, 2.4 are replaced by Lemmas B.5, B.6, B.7 below, respectively. Proposition 2.4 continues to hold. This concludes the proof. □

Lemma B.5 (SMS($\boldsymbol{\alpha}$) causes steepest descent)**.** *Let* $(\bar{\mathbf{A}}, \bar{\mathbf{X}}^U)$ *be any FSP under any* *non-idling policy U on $[0,T]$, and consider any $\boldsymbol{\alpha} \in \mathrm{relint}(\Omega)$. For a regular $t \in [0,T]$, define:*

$$S_1(\bar{\mathbf{X}}^U(t)) \triangleq \left\{ k \in V_S : k \in \operatorname{argmin} \frac{\bar{X}_k^U(t)}{\alpha_k} \right\},$$
$$S_2\left(\bar{\mathbf{X}}^U(t), \dot{\bar{\mathbf{X}}}^U(t)\right) \triangleq \left\{ k \in S_1(\bar{\mathbf{X}}^U(t)) : k \in \operatorname{argmin} \frac{\dot{\bar{X}}_k^U(t)}{\alpha_k} \right\}.$$

All the derivatives are well defined since t is regular. We have

$$\dot{L}_{\boldsymbol{\alpha}}(\bar{\mathbf{X}}^U(t)) = -\frac{\dot{\bar{\mathbf{X}}}_k^U(t)}{\alpha_k} \quad \text{for any } k \in S_2(\bar{\mathbf{X}}^U(t)) \tag{B.44}$$
$$\geq -\frac{1}{\mathbf{1}_{S_2}^{\mathrm{T}}\boldsymbol{\alpha}} \left(\sum_{i \in V_S, j' \in \partial(S_2)} \dot{\bar{A}}_{ij'}(t) - \sum_{i \in S_2, j' \in V_D} \dot{\bar{A}}_{ij'}(t) \right) \tag{B.45}$$

for $\bar{\mathbf{X}}^U(t) \neq \boldsymbol{\alpha}$ and $L_{\boldsymbol{\alpha}}(\bar{\mathbf{X}}^U(t)) < 1$. Inequality (B.45) holds with equality under SMS($\boldsymbol{\alpha}$)*, i.e.,* SMS($\boldsymbol{\alpha}$) *satisfies the steepest descent property in Proposition 2.4.*

Proof. We will write $S_1(\bar{\mathbf{X}}(t))$ as S_1, $S_2\left(\bar{\mathbf{X}}(t), \dot{\bar{\mathbf{X}}}(t)\right)$ as S_2, and $\min_{k \in S_1} \frac{\dot{\bar{X}}_k(t)}{\alpha_k}$ as c in the following. Let $(\bar{\mathbf{A}}, \bar{\mathbf{X}}^U)$ be an FSP under policy $U \in \mathcal{U}$.

- *Proof of* (B.44). The proof is exactly the same as the proof of (2.22).
- *Proof of* (B.45). For the K-th system, define auxiliary processes:

  $$\bar{E}_{ij'k}^{K,U}(t) \triangleq \#\{\text{Type } (i,j') \text{ demand units that arrive during } [0,t]$$
  $$\text{and are served by agents at } k \text{ under policy } U \in \mathcal{U}\} \quad i,k \in V_S\,,\ j' \in V_D\,.$$

  Similar to the proof of (2.23), extend the definition of FSP to $(\bar{\mathbf{A}}(\cdot), \bar{\mathbf{X}}(\cdot), \bar{\mathbf{E}}(\cdot))$. For regular times t, we have

  $$\sum_{i \in S_2} \dot{\bar{X}}_i^U(t) = \sum_{k \in V_S, i \in S_2} \sum_{j' \in \partial(i)} \dot{\bar{E}}_{kj'i}^U(t) - \sum_{i \in S_2, j' \in V_D} \sum_{k \in \partial(j')} \dot{\bar{E}}_{ij'k}^U(t)\,.$$

  Consider any non-idling policy $U' \in \mathcal{U}$, it cannot use the agents in S_2 to serve the demand of service types out of $\partial(S_2)$. Therefore for any policy U' we have

  $$\sum_{k \in S_2} \dot{\bar{X}}_k^{U'}(t) \leq \sum_{i \in V_S, j' \in \partial(S_2)} \dot{\bar{A}}_{ij'}(t) - \sum_{i \in S_2, j' \in V_D} \dot{\bar{A}}_{ij'}(t)\,. \tag{B.46}$$

  For SMS($\boldsymbol{\alpha}$) policy, using similar argument as in the proof of (2.23), we know that all the demands for service type $j' \in \partial(S_2)$ will be served by agents $i \in S_2$ during $(t, t+\epsilon)$ for some $\epsilon > 0$. Hence we have

$$\begin{aligned}\sum_{k\in S_2} \dot{\bar{X}}_k^{\mathrm{SMS}(\alpha)}(t) &= \sum_{i\in V_S, j'\in\partial(S_2)} \dot{\bar{A}}_{ij'}(t) - \sum_{i\in S_2, j'\in V_D} \sum_{k\in\partial(j')} \dot{\bar{E}}_{ij'k}^{\mathrm{SMS}(\alpha)}(t) \\ &\geq \sum_{i\in V_S, j'\in\partial(S_2)} \dot{\bar{A}}_{ij'}(t) - \sum_{i\in S_2, j'\in V_D} \dot{\bar{A}}_{ij'}(t)\,.\end{aligned}$$

Finally, observe that for any $k \in S_2$,

$$\dot{L}_{\boldsymbol{\alpha}}(\bar{\mathbf{X}}(t)) = -\frac{\dot{\bar{X}}_k^{U'}(t)}{\alpha_k} = -\frac{1}{\mathbf{1}_{S_2}^{\mathrm{T}}\boldsymbol{\alpha}} \sum_{k\in S_2} \alpha_k \frac{\dot{\bar{X}}_k^{U'}(t)}{\alpha_k} = -\frac{1}{\mathbf{1}_{S_2}^{\mathrm{T}}\boldsymbol{\alpha}} \sum_{k\in S_2} \dot{\bar{X}}_k^{U'}(t)\,. \tag{B.47}$$

Plug (B.46) into (B.47), we know that inequality (B.44) holds, and it becomes equality for SMS(α) policy.

□

Lemma B.6 (SMS($\boldsymbol{\alpha}$) satisfies negative drift)**.** *For any $\boldsymbol{\alpha} \in \mathrm{relint}(\Omega)$, under Assumption 2.5, the policy* SMS*($\boldsymbol{\alpha}$) satisfies the negative drift condition in Proposition 2.4.*

Proof. It follows from Lemma B.5 that for any fluid limit under SMS($\boldsymbol{\alpha}$) $(\bar{\mathbf{A}}(\cdot), \bar{\mathbf{X}}(t))$ and regular t, we have

$$\dot{L}_{\boldsymbol{\alpha}}(t) \leq -\min_{S_2 \subsetneq V_S, S_2 \neq \emptyset} \left(\sum_{i\in V_S, j'\in\partial(S_2)} \phi_{ij'} - \sum_{i\in S_2, j'\in V_D} \phi_{ij'} \right)\,.$$

Because of Assumption 2.5, we have $\dot{L}_{\boldsymbol{\alpha}}(t) < 0$, and the rest of the proof proceeds exactly the same as the proof of Lemma 2.3. □

Lemma B.7. *Recall the definitions of B_J, λ_J and μ_J in* (2.26)*. For any $\boldsymbol{\alpha} \in \mathrm{relint}(\Omega)$, we have $\gamma(\boldsymbol{\alpha}) = \min_{I\subsetneq V_S, I\neq\emptyset} B_I \log\left(\frac{\lambda_I}{\mu_I}\right)$.*

Proof. We omit the proof because it is almost identical to the proof of Lemma 2.4. □

B.8 SMW with Travel Delays: Proof of Theorem 2.2

This section provides a proof of Theorem 2.2, our guarantee of exponentially small loss under SMW in the presence of travel delays (Section 2.6.1).

B.8.1 Fluid Sample Paths, Fluid Limits, and Large Deviations Principle

Similar to the development in Section 2.5.1, we first define the fluid sample paths and fluid limits of the system with delay. Consider the K-th system under SMW($\boldsymbol{\alpha}$) policy. We make the following definitions:

- For $j' \in V_D$, $k \in V_S$, let $\mathbf{A}^K_{j'k}(\cdot)$ be an independent Poisson process with rate $\hat{\boldsymbol{\phi}}^K_{j'k} = K\hat{\boldsymbol{\phi}}_{j'k}$.
- For $\boldsymbol{\alpha} \in \text{relint}(\Omega)$ and $i \in V_S$, we denote by $X^{K,\boldsymbol{\alpha}}_i(t)$ the number of available supply units at node i at time t.
- For $j' \in V_D$, $k \in V_S$, we denote by $Y^{K,\boldsymbol{\alpha}}_{j'k}(t)$ the number of supply units transporting type (j', k) demands at time t.
- For $j' \in V_D$, $k \in V_S$, we denote by $R^{K,\boldsymbol{\alpha}}_{j'k}(t)$ be the cumulative number of supply units that arrive at node k carrying type (j', k) demand during time $[0, t]$.

Define the scaled version of the above sample paths as follows:

$$\bar{A}^K_{j'k}(t) \triangleq \frac{1}{K} A^K_{j'k}(t)\,, \qquad \bar{X}^{K,\boldsymbol{\alpha}}_i(t) \triangleq \frac{1}{K}\bar{X}^{K,\boldsymbol{\alpha}}_i(t)\,, \tag{B.48}$$

$$\bar{Y}^{K,\boldsymbol{\alpha}}_{j'k}(t) \triangleq \frac{1}{K} Y^{K,\boldsymbol{\alpha}}_{j'k}(t)\,, \qquad \bar{R}^{K,\boldsymbol{\alpha}}_{j'k}(t) \triangleq \frac{1}{K}\bar{R}^{K,\boldsymbol{\alpha}}_{j'k}(t)\,. \tag{B.49}$$

We define fluid sample paths and fluid limits as follows.

Definition B.1 (Fluid sample paths)**.** *We call* $(\bar{\mathbf{A}}(\cdot), \bar{\mathbf{X}}^{\boldsymbol{\alpha}}(\cdot), \bar{\mathbf{Y}}^{\boldsymbol{\alpha}}(\cdot), \bar{\mathbf{R}}^{\boldsymbol{\alpha}}(\cdot))_T$ *a fluid sample path (under SMW($\boldsymbol{\alpha}$)) on* $[0, T]$ *if there exists a sequence of sample paths* $(\bar{\mathbf{A}}^K(\cdot), \bar{\mathbf{X}}^{K,\boldsymbol{\alpha}}(\cdot), \bar{\mathbf{Y}}^{K,\boldsymbol{\alpha}}(\cdot), \bar{\mathbf{R}}^{K,\boldsymbol{\alpha}}(\cdot))_{K=1}^{\infty}$ *(which are defined in* (B.48) *and* (B.49)*), such that it has a subsequence which converges to* $(\bar{\mathbf{A}}(\cdot), \bar{\mathbf{X}}^{\boldsymbol{\alpha}}(\cdot), \bar{\mathbf{Y}}^{\boldsymbol{\alpha}}(\cdot), \bar{\mathbf{R}}^{\boldsymbol{\alpha}}(\cdot))$ *uniformly on* $[0, T]$*.*

Definition B.2 (Fluid limits)**.** *We call* $(\bar{\mathbf{A}}(\cdot), \bar{\mathbf{X}}^{\boldsymbol{\alpha}}(\cdot), \bar{\mathbf{Y}}^{\boldsymbol{\alpha}}(\cdot), \bar{\mathbf{R}}^{\boldsymbol{\alpha}}(\cdot))_T$ *a fluid limit (under SMW($\boldsymbol{\alpha}$)) on* $[0, T]$ *if (i) it is a fluid sample path; (ii) we have* $\bar{A}_{j'k}(t) = \hat{\phi}_{j'k}t$ *and* $\bar{R}_{j'k}(t) = \frac{1}{\tau_{j'k}}\int_{s=0}^{t} \bar{Y}^{\boldsymbol{\alpha}}_{j'k}(s)ds$*, for all* $j' \in V_D, k \in V_S$ *and all* $t \in [0, T]$*.*

Large deviations principle for $M/M/\infty$ queue. Because the system with travel delay consists of $M/M/\infty$ queues, the following result [Theorem 12.18, 128] is useful.

Let $Y^K(\cdot)$ be the sample path of the content of an $M/M/\infty$ queue with job arrival rate $K\hat{\phi}$ and service rate τ^{-1}; $A^K(t)$ be the number of job arrivals to the queue during $[0,t]$; $R^K(t)$ be the number of served jobs during $[0,t]$. Let

$$\bar{Y}^K(t) \triangleq \frac{1}{K} Y^K(t)\,, \quad \bar{A}^K(t) \triangleq \frac{1}{K} A^K(t)\,, \quad \bar{R}^K(t) \triangleq \frac{1}{K} R^K(t)\,.$$

Let μ_K be the law of $(\bar{Y}^K(\cdot), \bar{A}^K(\cdot), \bar{R}^K(\cdot))$ in $(L^\infty[0,T])^3$. Let $\Lambda^*(\ell,\cdot)$ be the large deviation rate function of Poisson random variable with mean ℓ:

$$\Lambda^*(\ell, f) \triangleq \begin{cases} f \log \frac{f}{\ell} - f + \ell & \text{if } f > 0\,, \\ \infty & \text{otherwise}\,. \end{cases} \tag{B.50}$$

For any set Γ, let $\bar{\Gamma}$ be its closure, and Γ^o be its interior. We have the following sample path large deviations principle.[4]

Fact B.1. *For measures $\{\mu_K\}$ defined above, and any arbitrary measurable set $\Gamma \subseteq (L^\infty[0,T])^3$, we have*

$$-\inf_{(\bar{Y},\bar{A},\bar{R})\in\Gamma^o} I_T(\bar{Y}) \le \liminf_{K\to\infty} \frac{1}{K} \log \mu_K(\Gamma) \le \limsup_{K\to\infty} \frac{1}{K} \log \mu_K(\Gamma) \le -\inf_{(\bar{Y},\bar{A},\bar{R})\in\bar{\Gamma}} I_T(\bar{Y})\,, \tag{B.51}$$

where the rate function is:

$$I_T(\bar{Y},\bar{A},\bar{R}) \triangleq \begin{cases} \int_0^T \left(\Lambda^*\left(\hat{\phi}, \dot{\bar{A}}(t)\right) + \Lambda^*\left(\frac{\bar{Y}(t)}{\tau}, \dot{\bar{R}}(t)\right)\right) dt & \textit{if } \bar{Y}(\cdot), \bar{A}(\cdot), \bar{R}(\cdot) \in \mathrm{AC}[0,T]\,, \\ & \bar{Y}(0) = 0\,, \\ \infty & \textit{otherwise}\,. \end{cases} \tag{B.52}$$

Here $\mathrm{AC}[0,T]$ *is the space of absolutely continuous functions on* $[0,T]$.

[4] The original formulation in [128] is more compact than the following one, but the following formulation turns out to be more useful in our analysis.

B.8.2 Lyapunov Functions and Drift

Our analysis relies on a novel family of piecewise linear Lyapunov functions, which we construct below. Let Ω^ℓ be the $(\ell-1)$-dimensional simplex.

Definition B.3. *For each $\boldsymbol{\alpha} \in \mathrm{relint}(\Omega)$, define Lyapunov function $L_{\boldsymbol{\alpha}}(\mathbf{x}, \mathbf{y}) : \Omega^{m+n\times m} \to \mathbb{R}$ as*

$$L_{\boldsymbol{\alpha}}(\mathbf{x}, \mathbf{y}) = \ L_{1,\boldsymbol{\alpha}}(\mathbf{x}) + \frac{2}{\min_{i\in V_S} \alpha_i} L_2(\mathbf{y})$$

where $L_{1,\boldsymbol{\alpha}}(\mathbf{x}) = \ \beta - \min_{i\in V_S} \ \frac{x_i}{\alpha_i}$, $L_2(\mathbf{y}) = \ \sum_{j'\in V_D, k\in V_S} |y_{j'k} - \tau_{jk}\hat{\phi}_{j'k}|$.

The intuition of such choices of Lyapunov functions is as follows. The first part of the Lyapunov function, $L_{1,\boldsymbol{\alpha}}(\mathbf{x})$, is almost identical to the Lyapunov function for the no-delay case (see Definition 2.7) except for the constant term since only β portion of the cars are available at the system equilibrium. It captures how much the current distribution of available supply units deviates from the distribution at equilibrium. The second part of the Lyapunov function characterizes the deviation of the number of in-transit cars from their typical values. The Lyapunov function attains minimum value 0 at $\Omega_{m+n\times m}$ at $(\,(\beta\alpha_i)_{i\in V_S}, (\tau_{j'k}\hat{\phi}_{j'k})_{j'\in V_D, k\in V_S}\,)$, and is strictly positive elsewhere on $\Omega_{m+n\times m}$.

Same as before, the demand-loss probability can be upper bounded by the probability that the Lyapunov function exceeds a certain value. Note that demand loss only happens when $x_i = 0$ for some $i \in V_S$, which implies $L_{1,\boldsymbol{\alpha}} = \beta$. In the following, we bound the probability of the event where $L_{\boldsymbol{\alpha}}(\bar{\mathbf{X}}, \bar{\mathbf{Y}}) \geq \beta$.

Because we only need an achievability bound, it suffices to prove a result analogous to Lemma B.3. As a first step, we establish in the following lemma that the Lyapunov function has negative drift under SMW($\boldsymbol{\alpha}$) policies in the fluid limit.

A time $t \in (0, T)$ is said to be a *regular point* of an FSP $(\bar{\mathbf{A}}(\cdot), \bar{\mathbf{X}}^{\boldsymbol{\alpha}}(\cdot), \bar{\mathbf{Y}}^{\boldsymbol{\alpha}}(\cdot), \bar{\mathbf{R}}^{\boldsymbol{\alpha}}(\cdot))_T$ if $\bar{\mathbf{A}}(\cdot), \bar{\mathbf{X}}^{\boldsymbol{\alpha}}(\cdot), \bar{\mathbf{Y}}^{\boldsymbol{\alpha}}(\cdot), \bar{\mathbf{R}}^{\boldsymbol{\alpha}}(\cdot), L_{\boldsymbol{\alpha}}(\bar{\mathbf{X}}^{\boldsymbol{\alpha}}(\cdot), \bar{\mathbf{Y}}^{\boldsymbol{\alpha}}(\cdot))$ are all differentiable at time t.

Because of the Large Deviations Principle (Facts 2.1 and B.1), it will suffice in our analysis to consider only the FSPs that have absolutely continuous demand sample paths

$\bar{\mathbf{A}}(\cdot)$. Now, if $\bar{\mathbf{A}}(\cdot)$ is absolutely continuous, then so are $\bar{\mathbf{X}}^{\boldsymbol{\alpha}}(\cdot)$ and $L_{\boldsymbol{\alpha}}(\bar{\mathbf{X}}^{\boldsymbol{\alpha}}(\cdot))$, and as a result almost all t are regular.

As a first step to bound the drift of $L_{\boldsymbol{\alpha}}$ we first bound the drift of $L_{1,\boldsymbol{\alpha}}$ in Lemma B.8. For notation simplicity, we drop the FSP's superscript $\boldsymbol{\alpha}$.

Lemma B.8. *Let* $(\bar{\mathbf{A}}(\cdot), \bar{\mathbf{X}}(\cdot), \bar{\mathbf{Y}}(\cdot), \bar{\mathbf{R}}(\cdot))_T$ *be any FSP under SMW($\boldsymbol{\alpha}$) on* $[0,T]$*, where* $\boldsymbol{\alpha} \in \operatorname{relint}(\Omega)$*. Define:*

$$S_1(\bar{\mathbf{X}}(t)) \triangleq \left\{ k \in V_S : k \in \operatorname{argmin} \frac{\bar{\mathbf{X}}_k(t)}{\alpha_k} \right\},$$
$$S_2\left(\bar{\mathbf{X}}(t), \dot{\bar{\mathbf{X}}}(t)\right) \triangleq \left\{ k \in S_1(\bar{\mathbf{X}}(t)) : k \in \operatorname{argmin} \frac{\dot{\bar{\mathbf{X}}}_k(t)}{\alpha_k} \right\}. \tag{B.53}$$

For a regular $t \in [0,T]$*, we have*

$$\dot{L}_{1,\boldsymbol{\alpha}}(\bar{\mathbf{X}}(t)) \leq -\frac{1}{\mathbf{1}_{S_2}^{\mathrm{T}}\boldsymbol{\alpha}} \left(\sum_{j' \in V_D, k \in S_2} \dot{\bar{R}}_{j'k}(t) - \sum_{j' \in V_D : \partial(j) \subset S_2, k \in V_S} \dot{\bar{A}}_{j'k}(t) \right).$$

Proof. From (B.19) we have

$$\dot{L}_{1,\boldsymbol{\alpha}}(\bar{\mathbf{X}}(t)) = -\frac{1}{\mathbf{1}_{S_2}^{\mathrm{T}}\boldsymbol{\alpha}} \sum_{k \in S_2} \dot{\bar{X}}_k(t). \tag{B.54}$$

Because we are considering a closed system, it holds that:

$$\sum_{j' \in V_D, k \in V_S} \dot{\bar{Y}}_{j'k}(t) + \sum_{k \in V_S} \dot{\bar{X}}_k(t) = 0. \tag{B.55}$$

Therefore

$$\sum_{k \in S_2} \dot{\bar{X}}_k(t) = -\sum_{j' \in V_D, k \in V_S} \dot{\bar{Y}}_{j'k}(t) - \sum_{k \in V_S \backslash S_2} \dot{\bar{X}}_k(t). \tag{B.56}$$

Note that

$$\dot{\bar{Y}}_{j'k}(t) \leq \dot{\bar{A}}_{j'k}(t) - \dot{\bar{R}}_{j'k}(t), \tag{B.57}$$

where the equality is achieved when no type (j', k) demand is lost at time t. Using the same argument as in the proof of Lemma 2.2, we know that under SMW($\boldsymbol{\alpha}$) policy all demand in $\partial(V_S \backslash S_2)$ are served by supplies in $V_S \backslash S_2$, and that no demand whose origin is in $\partial(V_S\backslash S_2)$ is lost. We have

$$\sum_{k\in V_S\backslash S_2} \dot{\bar{X}}_k(t) = \sum_{j'\in V_D, k\in V_S\backslash S_2} \dot{\bar{R}}_{j'k}(t) - \sum_{j'\in\partial(V_S\backslash S_2), k\in V_S} \dot{\bar{A}}_{j'k}(t)\,. \tag{B.58}$$

Plug in (B.57) and (B.58) to (B.56), we have

$$\begin{aligned}\sum_{k\in S_2} \dot{\bar{X}}_k(t) \geq & \sum_{j'\in V_D, k\in V_S} \left(\dot{\bar{R}}_{j'k}(t) - \dot{\bar{A}}_{j'k}(t)\right) - \sum_{j'\in V_D, k\in V_S\backslash S_2} \dot{\bar{R}}_{j'k}(t) + \sum_{j'\in\partial(V_S\backslash S_2), k\in V_S} \dot{\bar{A}}_{j'k}(t) \\ = & \sum_{j'\in V_D, k\in S_2} \dot{\bar{R}}_{j'k}(t) - \sum_{j'\in V_D:\partial(j)\subset S_2, k\in V_S} \dot{\bar{A}}_{j'k}(t)\,.\end{aligned}$$

Plugging the above to (B.54) and we conclude the proof. □

Now we are ready to bound the drift of $L_{\boldsymbol{\alpha}}$.

Lemma B.9. *Let* $(\bar{\mathbf{A}}(\cdot), \bar{\mathbf{X}}(\cdot), \bar{\mathbf{Y}}(\cdot), \bar{\mathbf{R}}(\cdot))_T$ *be any FSP under SMW($\boldsymbol{\alpha}$) on* $[0, T]$*, where* $\boldsymbol{\alpha} \in \mathrm{relint}(\Omega)$*. Recall the definition of* S_2 *in* (B.53)*.*

- *If for any* $i \in S_2$*,* $\bar{X}_i(t) > 0$ *or* $\bar{X}_i(t) = 0$*,* $\dot{\bar{X}}_i(t) > 0$*, we have*

$$\begin{aligned}& \dot{L}(\bar{\mathbf{X}}(t), \bar{\mathbf{Y}}(t)) \\ \triangleq\; & F_1(\bar{\mathbf{X}}(t), \bar{\mathbf{Y}}(t), \dot{\bar{\mathbf{A}}}(t), \dot{\bar{\mathbf{R}}}(t)) \\ \leq\; & -\frac{1}{\alpha_{\min}} \sum_{j'\in V_D, k\in V_S} \left|\hat{\phi}_{j'k} - \frac{\bar{Y}_{j'k}(t)}{\tau_{j'k}}\right| \\ & - \frac{1}{\mathbf{1}_{S_2}^{\mathrm{T}}\boldsymbol{\alpha}} \left(\sum_{j'\in V_D, k\in S_2} \dot{\bar{A}}_{j'k}(t) - \sum_{j':\partial(j')\subset S_2, k\in V_S} \dot{\bar{A}}_{j'k}(t)\right) \\ & + \frac{3}{\alpha_{\min}} \sum_{j'\in V_D, k\in V_S} \left(\left|\dot{\bar{A}}_{j'k}(t) - \hat{\phi}_{j'k}\right| + \left|\dot{\bar{R}}_{j'k}(t) - \frac{\bar{Y}_{j'k}(t)}{\tau_{j'k}}\right|\right)\,.\end{aligned}$$

- *If for* $i \in S_2$*,* $\bar{X}_i(t) = 0$ *and* $\dot{\bar{X}}_i(t) = 0$*, we have*

$$\begin{aligned}& \dot{L}(\bar{\mathbf{X}}(t), \bar{\mathbf{Y}}(t)) \\ \triangleq\; & F_2(\bar{\mathbf{X}}(t), \bar{\mathbf{Y}}(t), \dot{\bar{\mathbf{A}}}(t), \dot{\bar{\mathbf{R}}}(t)) \\ \leq\; & -\frac{2}{\alpha_{\min}} \left(\sum_{j'\in V_D, k\in S_2} \dot{\bar{A}}_{j'k}(t) - \sum_{j':\partial(j')\subset S_2, k\in V_S} \dot{\bar{A}}_{j'k}(t)\right) \\ & - \frac{2}{\alpha_{\min}} \sum_{j'\in V_D, k\in S_2} \left[\hat{\phi}_{j'k} - \frac{\bar{Y}_{j'k}(t)}{\tau_{j'k}}\right]^-\end{aligned}$$

$$- \frac{2}{\alpha_{\min}} \sum_{j' \in V_D, k \notin S_2} \left| \hat{\phi}_{j'k} - \frac{\bar{Y}_{j'k}(t)}{\tau_{j'k}} \right|$$

$$+ \frac{4}{\alpha_{\min}} \sum_{j' \in V_D, k \in V_S} \left(\left| \dot{\bar{A}}_{j'k}(t) - \hat{\phi}_{j'k} \right| + \left| \dot{\bar{R}}_{j'k}(t) - \frac{\bar{Y}_{j'k}(t)}{\tau_{j'k}} \right| \right) .$$

Proof. Recall the definition of S_2 in (B.53). To analyze the Lyapunov drift of $L_{\boldsymbol{\alpha}}$, we consider two cases depending on, roughly speaking, whether the queues in S_2 are empty at t and shortly after t.

- **Case 1**: for any $i \in S_2$, $\bar{X}_i(t) > 0$ or $\bar{X}_i(t) = 0$, $\dot{\bar{X}}_i(t) > 0$. Let $\alpha_{\min} \triangleq \min_{i \in V_S} \alpha_i$. We have

$$\begin{aligned}
&\dot{L}_{\boldsymbol{\alpha}}(\bar{\mathbf{X}}(t), \bar{\mathbf{Y}}(t)) \\
\leq\ & -\frac{1}{\mathbf{1}_{S_2}^{\mathrm{T}} \boldsymbol{\alpha}} \left(\sum_{j' \in V_D, k \in S_2} \dot{\bar{R}}_{j'k}(t) - \sum_{j' : \partial(j') \subset S_2, k \in V_S} \dot{\bar{A}}_{j'k} \right) && \text{(B.59)} \\
& - \frac{2}{\alpha_{\min}} \sum_{j' \in V_D, k \in V_S} \left(\dot{\bar{A}}_{j'k}(t) - \dot{\bar{R}}_{j'k}(t) \right) \\
& \left(\mathbb{I}\left\{ \bar{Y}_{j'k}(t) \leq \hat{\phi}_{j'k} \tau_{j'k} \right\} - \mathbb{I}\left\{ \bar{Y}_{j'k}(t) > \hat{\phi}_{j'k} \tau_{j'k} \right\} \right) && \text{(B.60)} \\
\triangleq\ & F_1(\bar{\mathbf{X}}(t), \bar{\mathbf{Y}}(t), \dot{\bar{\mathbf{A}}}(t), \dot{\bar{\mathbf{R}}}(t)) .
\end{aligned}$$

Here the term (B.59) comes from Lemma B.8. Note that

$$\begin{aligned}
&\sum_{j' \in V_D, k \in S_2} \dot{\bar{R}}_{j'k}(t) - \sum_{j' : \partial(j') \subset S_2, k \in V_S} \dot{\bar{A}}_{j'k} \\
= & \left(\sum_{j' \in V_D, k \in S_2} \dot{\bar{A}}_{j'k}(t) - \sum_{j' : \partial(j') \subset S_2, k \in V_S} \dot{\bar{A}}_{j'k}(t) \right) - \sum_{j' \in V_D, k \in S_2} \left(\dot{\bar{A}}_{j'k}(t) - \dot{\bar{R}}_{j'k}(t) \right)
\end{aligned}$$

Simple algebra yields that: for $j' \in V_D$, $k \in V_S$.

$$\dot{\bar{A}}_{j'k}(t) - \dot{\bar{R}}_{j'k}(t) \leq \ \hat{\phi}_{j'k} - \frac{\bar{Y}_{j'k}(t)}{\tau_{j'k}} + \left| \dot{\bar{A}}_{j'k}(t) - \hat{\phi}_{j'k} \right| + \left| \dot{\bar{R}}_{j'k}(t) - \frac{\bar{Y}_{j'k}(t)}{\tau_{j'k}} \right| .$$

Combined, we have

$$\sum_{j' \in V_D, k \in S_2} \dot{\bar{R}}_{j'k}(t) - \sum_{j' : \partial(j') \subset S_2, k \in V_S} \dot{\bar{A}}_{j'k}$$

$$\geq \left(\sum_{j' \in V_D, k \in S_2} \dot{\bar{A}}_{j'k}(t) - \sum_{j':\partial(j') \subset S_2, k \in V_S} \dot{\bar{A}}_{j'k}(t) \right) - \sum_{j' \in V_D, k \in S_2} \left(\hat{\phi}_{j'k} - \frac{\bar{Y}_{j'k}(t)}{\tau_{j'k}} \right)$$
$$- \sum_{j' \in V_D, k \in S_2} \left(\left| \dot{\bar{A}}_{j'k}(t) - \hat{\phi}_{j'k} \right| + \left| \dot{\bar{R}}_{j'k}(t) - \frac{\bar{Y}_{j'k}(t)}{\tau_{j'k}} \right| \right) .$$

Now we focus on the term (B.60). For $j' \in V_D$, $k \in V_S$, we have

$$\left(\dot{\bar{A}}_{j'k}(t) - \dot{\bar{R}}_{j'k}(t) \right) \left(\mathbb{I}\left\{ \bar{Y}_{j'k}(t) \leq \hat{\phi}_{j'k}\tau_{j'k} \right\} - \mathbb{I}\left\{ \bar{Y}_{j'k}(t) > \hat{\phi}_{j'k}\tau_{j'k} \right\} \right)$$
$$= \left(\dot{\bar{A}}_{j'k}(t) - \dot{\bar{R}}_{j'k}(t) \right) \left(\mathbb{I}\left\{ \hat{\phi}_{j'k} \geq \frac{\bar{Y}_{j'k}(t)}{\tau_{j'k}} \right\} - \mathbb{I}\left\{ \hat{\phi}_{j'k} < \frac{\bar{Y}_{j'k}(t)}{\tau_{j'k}} \right\} \right)$$
$$= \left(\left(\dot{\bar{A}}_{j'k}(t) - \hat{\phi}_{j'k} \right) - \left(\dot{\bar{R}}_{j'k}(t) - \frac{\bar{Y}_{j'k}(t)}{\tau_{j'k}} \right) \right)$$
$$\left(\mathbb{I}\left\{ \hat{\phi}_{j'k} \geq \frac{\bar{Y}_{j'k}(t)}{\tau_{j'k}} \right\} - \mathbb{I}\left\{ \hat{\phi}_{j'k} < \frac{\bar{Y}_{j'k}(t)}{\tau_{j'k}} \right\} \right)$$
$$+ \left(\hat{\phi}_{j'k} - \frac{\bar{Y}_{j'k}(t)}{\tau_{j'k}} \right) \left(\mathbb{I}\left\{ \hat{\phi}_{j'k} \geq \frac{\bar{Y}_{j'k}(t)}{\tau_{j'k}} \right\} - \mathbb{I}\left\{ \hat{\phi}_{j'k} < \frac{\bar{Y}_{j'k}(t)}{\tau_{j'k}} \right\} \right) .$$

Note that

$$\left(\hat{\phi}_{j'k} - \frac{\bar{Y}_{j'k}(t)}{\tau_{j'k}} \right) \left(\mathbb{I}\left\{ \hat{\phi}_{j'k} \geq \frac{\bar{Y}_{j'k}(t)}{\tau_{j'k}} \right\} - \mathbb{I}\left\{ \hat{\phi}_{j'k} < \frac{\bar{Y}_{j'k}(t)}{\tau_{j'k}} \right\} \right) = \left| \hat{\phi}_{j'k} - \frac{\bar{Y}_{j'k}(t)}{\tau_{j'k}} \right| ,$$

and that

$$\left(\left(\dot{\bar{A}}_{j'k}(t) - \hat{\phi}_{j'k} \right) - \left(\dot{\bar{R}}_{j'k}(t) - \frac{\bar{Y}_{j'k}(t)}{\tau_{j'k}} \right) \right)$$
$$\left(\mathbb{I}\left\{ \hat{\phi}_{j'k} \geq \frac{\bar{Y}_{j'k}(t)}{\tau_{j'k}} \right\} - \mathbb{I}\left\{ \hat{\phi}_{j'k} < \frac{\bar{Y}_{j'k}(t)}{\tau_{j'k}} \right\} \right)$$
$$\geq - \left| \dot{\bar{A}}_{j'k}(t) - \hat{\phi}_{j'k} \right| - \left| \dot{\bar{R}}_{j'k}(t) - \frac{\bar{Y}_{j'k}(t)}{\tau_{j'k}} \right| .$$

Therefore we have

$$\left(\dot{\bar{A}}_{j'k}(t) - \dot{\bar{R}}_{j'k}(t) \right) \left(\mathbb{I}\left\{ \bar{Y}_{j'k}(t) \leq \hat{\phi}_{j'k}\tau_{j'k} \right\} - \mathbb{I}\left\{ \bar{Y}_{j'k}(t) > \hat{\phi}_{j'k}\tau_{j'k} \right\} \right)$$
$$\geq \left| \hat{\phi}_{j'k} - \frac{\bar{Y}_{j'k}(t)}{\tau_{j'k}} \right| - \left(\left| \dot{\bar{A}}_{j'k}(t) - \hat{\phi}_{j'k} \right| + \left| \dot{\bar{R}}_{j'k}(t) - \frac{\bar{Y}_{j'k}(t)}{\tau_{j'k}} \right| \right) .$$

Plugging into (B.59) and (B.60), we have

$$\begin{aligned}
&F_1(\bar{\mathbf{X}}(t), \bar{\mathbf{Y}}(t), \dot{\bar{\mathbf{A}}}(t), \dot{\bar{\mathbf{R}}}(t)) \\
\leq\ & -\frac{1}{\alpha_{\min}} \sum_{j' \in V_D, k \in V_S} \left| \hat{\phi}_{j'k} - \frac{\bar{Y}_{j'k}(t)}{\tau_{j'k}} \right| - \frac{1}{\mathbf{1}_{S_2}^{\mathrm{T}} \boldsymbol{\alpha}} \left(\sum_{j' \in V_D, k \in S_2} \dot{\bar{A}}_{j'k}(t) - \sum_{j' : \partial(j') \subset S_2, k \in V_S} \dot{\bar{A}}_{j'k}(t) \right) \\
&+ \frac{3}{\alpha_{\min}} \sum_{j' \in V_D, k \in V_S} \left(\left| \dot{\bar{A}}_{j'k}(t) - \hat{\phi}_{j'k} \right| + \left| \dot{\bar{R}}_{j'k}(t) - \frac{\bar{Y}_{j'k}(t)}{\tau_{j'k}} \right| \right) .
\end{aligned}$$

- **Case 2:** for $i \in S_2$, $\bar{X}_i(t) = 0$ and $\dot{\bar{X}}_i(t) = 0$. In this case, $\dot{L}_{1,\boldsymbol{\alpha}}(\bar{\mathbf{X}}(t)) = 0$. Similar to the proof of Lemma 2.2, for $i, k \in V_S$, $j' \in V_D$, let $\bar{E}_{ij'k}(t)$ be the FSP of the number of type (j', k) demand served by supply units at i during $[0, t]$. Define

$$U_{j'k}(t) \triangleq A_{j'k}(t) - \sum_{i \in \partial(j')} \bar{E}_{ij'k}(t)$$

as the number of type (j', k) demand lost during $[0, t]$.

We have

$$\begin{aligned}
&\dot{L}_2(\bar{\mathbf{Y}}(t)) \\
=\ & -\sum_{j' \in V_D, k \in V_S} \left(\dot{\bar{A}}_{j'k}(t) - \dot{\bar{U}}_{j'k}(t) - \dot{\bar{R}}_{j'k}(t) \right) \mathbb{I}\left\{ \bar{Y}_{j'k}(t) \leq \hat{\phi}_{j'k} \tau_{j'k} \right\} \\
&+ \sum_{j' \in V_D, k \in V_S} \left(\dot{\bar{A}}_{j'k}(t) - \dot{\bar{U}}_{j'k}(t) - \dot{\bar{R}}_{j'k}(t) \right) \mathbb{I}\left\{ \bar{Y}_{j'k}(t) > \hat{\phi}_{j'k} \tau_{j'k} \right\} \\
\leq\ & -\sum_{j' \in V_D, k \in V_S} \left(\dot{\bar{A}}_{j'k}(t) - \dot{\bar{R}}_{j'k}(t) \right) \left(\mathbb{I}\left\{ \hat{\phi}_{j'k} \geq \frac{\bar{Y}_{j'k}(t)}{\tau_{j'k}} \right\} - \mathbb{I}\left\{ \hat{\phi}_{j'k} < \frac{\bar{Y}_{j'k}(t)}{\tau_{j'k}} \right\} \right) \\
&+ \sum_{j' \in V_D, k \in V_S} \dot{\bar{U}}_{j'k}(t) .
\end{aligned}$$

Note that by definition of the set S_2, no queue in $\partial(V_S \backslash S_2)$ loses demand at time t. We have

$$\sum_{j' \in V_D, k \in V_S} \dot{\bar{U}}_{j'k}(t) = \sum_{j' : \partial(j') \subset S_2, k \in V_S} \dot{\bar{A}}_{j'k}(t) - \sum_{j' \in V_D, k \in S_2} \dot{\bar{R}}_{j'k}(t) .$$

Combining, and using the same algebra as in Case 1, we have:

$$\dot{L}_2(\bar{\mathbf{Y}}(t))$$

$$
\begin{aligned}
&= -\sum_{j'\in V_D, k\in V_S}\left(\dot{\bar{A}}_{j'k}(t)-\dot{\bar{R}}_{j'k}(t)\right)\left(\mathbb{I}\left\{\hat{\phi}_{j'k}\geq\frac{\bar{Y}_{j'k}(t)}{\tau_{j'k}}\right\}-\mathbb{I}\left\{\hat{\phi}_{j'k}<\frac{\bar{Y}_{j'k}(t)}{\tau_{j'k}}\right\}\right)\\
&\quad+\sum_{j':\partial(j')\subset S_2, k\in V_S}\dot{\bar{A}}_{j'k}(t)-\sum_{j'\in V_D, k\in S_2}\dot{\bar{R}}_{j'k}(t)\\
&\leq -\sum_{j'\in V_D, k\in V_S}\left|\hat{\phi}_{j'k}-\frac{\bar{Y}_{j'k}(t)}{\tau_{j'k}}\right|+\sum_{j'\in V_D, k\in V_S}\left(\left|\dot{\bar{A}}_{j'k}(t)-\hat{\phi}_{j'k}\right|+\left|\dot{\bar{R}}_{j'k}(t)-\frac{\bar{Y}_{j'k}(t)}{\tau_{j'k}}\right|\right)\\
&\quad-\left(\sum_{j'\in V_D, k\in S_2}\dot{\bar{A}}_{j'k}(t)-\sum_{j':\partial(j')\subset S_2, k\in V_S}\dot{\bar{A}}_{j'k}(t)\right)+\sum_{j'\in V_D, k\in S_2}\left(\dot{\bar{A}}_{j'k}(t)-\dot{\bar{R}}_{j'k}(t)\right)\\
&\leq -\left(\sum_{j'\in V_D, k\in S_2}\dot{\bar{A}}_{j'k}(t)-\sum_{j':\partial(j')\subset S_2, k\in V_S}\dot{\bar{A}}_{j'k}(t)\right)-\sum_{j'\in V_D, k\in S_2}\left[\hat{\phi}_{j'k}-\frac{\bar{Y}_{j'k}(t)}{\tau_{j'k}}\right]^-\\
&\quad-\sum_{j'\in V_D, k\notin S_2}\left|\hat{\phi}_{j'k}-\frac{\bar{Y}_{j'k}(t)}{\tau_{j'k}}\right|+2\sum_{j'\in V_D, k\in V_S}\left(\left|\dot{\bar{A}}_{j'k}(t)-\hat{\phi}_{j'k}\right|+\left|\dot{\bar{R}}_{j'k}(t)-\frac{\bar{Y}_{j'k}(t)}{\tau_{j'k}}\right|\right).
\end{aligned}
$$

Here $[x]^- \triangleq -\min\{x,0\}$.

Therefore we have

$$
\begin{aligned}
&\dot{L}(\bar{\mathbf{X}}(t),\bar{\mathbf{Y}}(t))\\
&\triangleq F_2(\bar{\mathbf{X}}(t),\bar{\mathbf{Y}}(t),\dot{\bar{\mathbf{A}}}(t),\dot{\bar{\mathbf{R}}}(t))\\
&\leq -\frac{2}{\alpha_{\min}}\left(\sum_{j'\in V_D, k\in S_2}\dot{\bar{A}}_{j'k}(t)-\sum_{j':\partial(j')\subset S_2, k\in V_S}\dot{\bar{A}}_{j'k}(t)\right)\\
&\quad-\frac{2}{\alpha_{\min}}\sum_{j'\in V_D, k\in S_2}\left[\hat{\phi}_{j'k}-\frac{\bar{Y}_{j'k}(t)}{\tau_{j'k}}\right]^-\\
&\quad-\frac{2}{\alpha_{\min}}\sum_{j'\in V_D, k\notin S_2}\left|\hat{\phi}_{j'k}-\frac{\bar{Y}_{j'k}(t)}{\tau_{j'k}}\right|\\
&\quad+\frac{4}{\alpha_{\min}}\sum_{j'\in V_D, k\in V_S}\left(\left|\dot{\bar{A}}_{j'k}(t)-\hat{\phi}_{j'k}\right|+\left|\dot{\bar{R}}_{j'k}(t)-\frac{\bar{Y}_{j'k}(t)}{\tau_{j'k}}\right|\right).
\end{aligned}
$$

□

Using the result in Lemma B.9, we can show that the system has strictly negative Lyapunov drift in the fluid limit, and that the drift remains negative for perturbed demand arrival rates and travel times given the perturbation is small enough.

Lemma B.10. *Fix $\boldsymbol{\alpha} \in \mathrm{relint}(\Omega)$. Then there exists $\eta > 0$ and $\epsilon > 0$ such that for all FSPs $(\bar{\mathbf{A}}(\cdot), \bar{\mathbf{X}}(\cdot), \bar{\mathbf{Y}}(\cdot), \bar{\mathbf{R}}(\cdot))_T$ (under the SMW($\boldsymbol{\alpha}$) policy), if $t \in (0, T)$ is regular, $L_{\boldsymbol{\alpha}}(\bar{\mathbf{X}}(t), \bar{\mathbf{Y}}(t)) > \frac{\beta}{2}$, and that $\dot{\bar{\mathbf{A}}}(t) \in B(\hat{\boldsymbol{\phi}}, \epsilon)$, $\max_{j' \in V_D, k \in V_S} |\dot{\bar{R}}_{j'k} - \bar{Y}_{j'k}(t)/\tau_{j'k}| \le \epsilon$, we have $\dot{L}_{\boldsymbol{\alpha}}(\bar{\mathbf{X}}(t), \bar{\mathbf{Y}}(t)) \le -\eta$.*

Proof. Same as in the proof of Lemma B.9, we consider two cases. Recall the definition of S_2 in (B.53).

- If for any $i \in S_2$, $\bar{X}_i(t) > 0$ or $\bar{X}_i(t) = 0$, $\dot{\bar{X}}_i(t) > 0$, we have

$$
\begin{aligned}
&\dot{L}(\bar{\mathbf{X}}(t), \bar{\mathbf{Y}}(t)) \\
\triangleq\ & F_1(\bar{\mathbf{X}}(t), \bar{\mathbf{Y}}(t), \dot{\bar{\mathbf{A}}}(t), \dot{\bar{\mathbf{R}}}(t)) \\
\le\ & \underbrace{-\frac{1}{\alpha_{\min}} \sum_{j' \in V_D, k \in V_S} \left| \hat{\phi}_{j'k} - \frac{\bar{Y}_{j'k}(t)}{\tau_{j'k}} \right|}_{\text{(I)}} \\
& \underbrace{- \frac{1}{\mathbf{1}_{S_2}^{\mathrm{T}} \boldsymbol{\alpha}} \left(\sum_{j' \in V_D, k \in S_2} \hat{\phi}_{j'k}(t) - \sum_{j' : \partial(j') \subset S_2, k \in V_S} \hat{\phi}_{j'k}(t) \right)}_{\text{(II)}} .
\end{aligned}
$$

  Depending on whether $S_2 = V_S$, there are two sub-cases:

  - When $S_2 \neq V_S$, if follows from Assumption 2.3 that (II)> 0. Since (I)≥ 0, we have

    $$\dot{L}(\bar{\mathbf{X}}(t), \bar{\mathbf{Y}}(t)) \le -\text{(II)} \le -\min\{\lambda_{\min}, \xi\} .$$

    Here $\lambda_{\min} \triangleq \min_{i \in V_S} \mathbf{1}^{\mathrm{T}} \hat{\boldsymbol{\phi}}_{(i)} > 0$ is the minimum supply arrival rate at any node (that has positive arrival rate), and

    $$\xi \triangleq \min_{J \subsetneq V_D, J \neq \emptyset} \left(\sum_{i \in \partial(J)} \mathbf{1}^{\mathrm{T}} \hat{\boldsymbol{\phi}}_{(i)} - \sum_{j' \in J} \mathbf{1}^{\mathrm{T}} \hat{\boldsymbol{\phi}}_{j'} \right) > 0$$

    is the Hall's gap of the system.

  - When $S_2 = V_S$, observe that (II)$= 0$, hence we only analyze (I). Recall that we focus on the case where $L_{\boldsymbol{\alpha}}(\bar{\mathbf{X}}(t), \bar{\mathbf{Y}}(t)) > \frac{\beta}{2}$. Denote $\kappa \triangleq \sum_{i \in V_S} \bar{X}_i(t)$. We have

$$\sum_{j'\in V_D,k\in V_S} \left(\bar{Y}_{j'k}(t) - \hat{\phi}_{j'k}\tau_{j'k}\right) = \beta - \kappa \,,$$

hence

$$\sum_{j'\in V_D,k\in V_S} \left|\hat{\phi}_{j'k} - \frac{\bar{Y}_{j'k}(t)}{\tau_{j'k}}\right| \geq \frac{|\beta-\kappa|}{\tau_{\max}} \,. \tag{B.61}$$

Here $\tau_{\max} \triangleq \max_{j'\in V_D,k\in V_S} \tau_{j'k}$. Plug in to the expression of F_1, we have

$$\dot{L}(\bar{\mathbf{X}}(t), \bar{\mathbf{Y}}(t)) \leq -\frac{1}{\alpha_{\min}}\frac{|\beta-\kappa|}{\tau_{\max}} \,.$$

On the other hand, since $S_2 = V_S$, it must be that $\bar{X}_i(t) = \alpha_i \kappa$ for all $i \in V_S$, hence

$$L_{1,\boldsymbol{\alpha}}(\bar{\mathbf{X}}(t)) = \beta - \kappa \,.$$

Since $L_{\boldsymbol{\alpha}}(\bar{\mathbf{X}}(t), \bar{\mathbf{Y}}(t)) > \frac{\beta}{2}$, we have

$$L_2(\bar{\mathbf{Y}}(t)) > \left(\kappa - \frac{\beta}{2}\right)\frac{\alpha_{\min}}{2} \,.$$

When $\kappa < \frac{3}{4}\beta$, plugging into (B.61), we have

$$\dot{L}_{\boldsymbol{\alpha}}(\bar{\mathbf{X}}(t), \bar{\mathbf{Y}}(t)) \leq -\frac{\beta}{4\tau_{\max}} \,.$$

When $\kappa \geq \frac{3}{4}\beta$, we have

$$L_2(\bar{\mathbf{Y}}(t)) = \sum_{j'\in V_D,k\in V_S} |\bar{Y}_{j'k}(t) - \tau_{j'k}\hat{\phi}_{j'k}| \geq \frac{\alpha_{\min}\beta}{8} \,.$$

hence

$$\sum_{j'\in V_D,k\in V_S} \left|\hat{\phi}_{j'k} - \frac{\bar{Y}_{j'k}(t)}{\tau_{j'k}}\right| \geq \frac{\alpha_{\min}\beta}{8\tau_{\max}} \,,$$

therefore

$$\dot{L}_{\boldsymbol{\alpha}}(\bar{\mathbf{X}}(t), \bar{\mathbf{Y}}(t)) \leq -\frac{\alpha_{\min}\beta}{8\tau_{\max}} \,.$$

Combine all the above analysis, we have

$$\dot{L}_{\boldsymbol{\alpha}}(\bar{\mathbf{X}}(t), \bar{\mathbf{Y}}(t)) \leq \; -\min\left\{\lambda_{\min}, \xi, \frac{\alpha_{\min}\beta}{8\tau_{\max}}\right\}.$$

- If for $i \in S_2$, $\bar{X}_i(t) = 0$ and $\dot{\bar{X}}_i(t) = 0$, we have

$$\begin{aligned}
&\dot{L}(\bar{\mathbf{X}}(t), \bar{\mathbf{Y}}(t)) \\
&\triangleq F_2(\bar{\mathbf{X}}(t), \bar{\mathbf{Y}}(t), \dot{\bar{\mathbf{A}}}(t), \dot{\bar{\mathbf{R}}}(t)) \\
&\leq \; -\frac{2}{\alpha_{\min}}\left(\sum_{j'\in V_D, k\in S_2}\hat{\phi}_{j'k}(t) - \sum_{j':\partial(j')\subset S_2, k\in V_S}\hat{\phi}_{j'k}(t)\right) - \frac{2}{\alpha_{\min}}\sum_{j'\in V_D, k\in S_2}\left[\hat{\phi}_{j'k} - \frac{\bar{Y}_{j'k}(t)}{\tau_{j'k}}\right]^- \\
&\quad -\frac{2}{\alpha_{\min}}\sum_{j'\in V_D, k\notin S_2}\left|\hat{\phi}_{j'k} - \frac{\bar{Y}_{j'k}(t)}{\tau_{j'k}}\right|.
\end{aligned}$$

  - When $S_2 \neq V_S$, we have

$$\dot{L}(\bar{\mathbf{X}}(t), \bar{\mathbf{Y}}(t)) \leq -\min\{\lambda_{\min}, \xi\}.$$

  - When $S_2 = V_S$. Since all the cars are in-transit, we have

$$\sum_{j'\in V_D, k\in V_S}\hat{\phi}_{j'k}\tau_{j'k} - \sum_{j'\in V_D, k\in V_S}\bar{Y}_{j'k}(t) = \; -\beta$$

    Hence

$$-\sum_{j'\in V_D, k\in V_S}\left[\hat{\phi}_{j'k}\tau_{j'k} - \bar{Y}_{j'k}(t)\right]^- \leq \; -\beta,$$

    and

$$-\sum_{j'\in V_D, k\in V_S}\left[\hat{\phi}_{j'k} - \frac{\bar{Y}_{j'k}(t)}{\tau_{j'k}}\right]^- \leq \; -\frac{\beta}{\tau_{\max}},$$

    Therefore

$$\dot{L}(\bar{\mathbf{X}}(t), \bar{\mathbf{Y}}(t)) \leq -\frac{\beta}{\tau_{\max}}.$$

Combine all the cases above, we have for any fluid limit, when $L_{\boldsymbol{\alpha}}(\bar{\mathbf{X}}(t), \bar{\mathbf{Y}}(t)) > \frac{\beta}{2}$, we have

$$\dot{L}_{\boldsymbol{\alpha}}(\bar{\mathbf{X}}(t), \bar{\mathbf{Y}}(t)) \leq \; -\min\left\{\lambda_{\min}, \xi, \frac{\alpha_{\min}\beta}{8\tau_{\max}}\right\}.$$

Repeat the analysis above for FSP, we have

$$\dot{L}_{\boldsymbol{\alpha}}(\bar{\mathbf{X}}(t),\bar{\mathbf{Y}}(t)) \leq -\min\left\{\min_{J\in\mathcal{J}}\left(\sum_{j'\in V_D,k\in\partial(J)}\dot{\bar{A}}_{j'k}-\sum_{j'\in J,k\in V_S}\dot{\bar{A}}_{j'k}\right),\min_{i\in V_S}\sum_{j'\in V_S}\dot{\bar{A}}_{j'i},\frac{\alpha_{\min}\beta}{8\tau_{\max}}\right\} + \frac{4}{\alpha_{\min}}\sum_{j'\in V_D,k\in V_S}\left(\left|\dot{\bar{A}}_{j'k}(t)-\hat{\phi}_{j'k}\right|+\left|\dot{\bar{R}}_{j'k}(t)-\frac{\bar{Y}_{j'k}(t)}{\tau_{j'k}}\right|\right). \tag{B.62}$$

Using the same argument as at end of proof of Lemma 2.3, we conclude that the drift is strictly negative for small enough perturbation of demand arrival rates and travel times. □

B.8.3 Proof of Theorem 2.2

Now we are ready to prove Theorem 2.2.

Proof of Theorem 2.2. Since we only need an achievability result, it suffices to repeat Steps 1 and 2 in the proof of Lemma B.3. Since the technical analysis is almost identical, we make the following claim and omit its proof.

Claim: Consider the system under SMW($\boldsymbol{\alpha}$) policy for some $\boldsymbol{\alpha}\in\text{relint}(\Omega)$. Let $\mathbb{P}_{\text{p}}^{K,U}$ be the pessimistic demand-loss probability defined in (2.1), then we have

$$-\limsup_{K\to\infty}\frac{1}{K}\log\mathbb{P}_{\text{p}}^{K,U}\geq\frac{\beta}{2}\gamma_{\text{AB}}(\boldsymbol{\alpha}). \tag{B.63}$$

Here for fixed $T>0$,

$$\gamma_{\text{AB}}(\boldsymbol{\alpha})\triangleq\inf_{v>0,\mathbf{f},(\bar{\mathbf{A}},\bar{\mathbf{X}},\bar{\mathbf{Y}},\bar{\mathbf{R}})}\frac{\sum_{j'\in V_D,k\in V_S}\Lambda^*(\hat{\phi}_{j'k},f_{j'k})+\sum_{j'\in V_D,k\in V_S}\Lambda^*(\frac{\bar{Y}_{j'k}(t)}{\tau_{j'k}},r_{j'k})}{v},$$

where $(\bar{\mathbf{A}},\bar{\mathbf{X}},\bar{\mathbf{Y}},\bar{\mathbf{R}})$ is a FSP on $[0,T]$ under SMW($\boldsymbol{\alpha}$) such that for some regular $t\in(0,T)$

$$\dot{\bar{\mathbf{A}}}(t)=\mathbf{f},\quad\dot{\bar{\mathbf{R}}}(t)=\mathbf{r},\quad L_{\boldsymbol{\alpha}}(\bar{\mathbf{X}}(t),\bar{\mathbf{Y}}(t))\in\left(\frac{\beta}{2},\beta\right),\quad\dot{L}_{\boldsymbol{\alpha}}(\bar{\mathbf{X}}(t),\bar{\mathbf{Y}}(t))=v.$$

It remains to show that $\gamma_{\text{AB}}(\boldsymbol{\alpha})>0$. Recall eq. (B.62):

$$\dot{L}_{\boldsymbol{\alpha}}(\bar{\mathbf{X}}(t),\bar{\mathbf{Y}}(t)) \leq -\min\left\{\min_{J\in\mathcal{J}}\left(\sum_{j'\in V_D,k\in\partial(J)}\dot{\bar{A}}_{j'k}-\sum_{j'\in J,k\in V_S}\dot{\bar{A}}_{j'k}\right),\min_{i\in V_S}\sum_{j'\in V_S}\dot{\bar{A}}_{j'i},\frac{\alpha_{\min}\beta}{8\tau_{\max}}\right\}$$

$$+\frac{4}{\alpha_{\min}}\sum_{j'\in V_D,k\in V_S}\left(\left|\dot{\bar{A}}_{j'k}(t)-\hat{\phi}_{j'k}\right|+\left|\dot{\bar{R}}_{j'k}(t)-\frac{\bar{Y}_{j'k}(t)}{\tau_{j'k}}\right|\right).$$

For $v>0$, define

$$\begin{aligned}\gamma(v)\triangleq&\min_{\mathbf{f}>\mathbf{0},\mathbf{r}>\mathbf{0},\mathbf{y}\in\Omega_{n\times m}}\sum_{j'\in V_D,k\in V_S}\Lambda^*(\hat{\phi}_{j'k},f_{j'k})+\sum_{j'\in V_D,k\in V_S}\Lambda^*\left(\frac{y_{j'k}}{\tau_{j'k}},r_{j'k}\right)\\ &\text{s.t. }-\min\left\{\min_{J\in\mathcal{J}}\left(\sum_{j'\in V_D,k\in\partial(J)}f_{j'k}-\sum_{j'\in J,k\in V_S}f_{j'k}\right),\min_{i\in V_S}\sum_{j'\in V_S}f_{j'i},\frac{\alpha_{\min}\beta}{8\tau_{\max}}\right\}\\ &\quad+\frac{4}{\alpha_{\min}}\sum_{j'\in V_D,k\in V_S}\left(\left|f_{j'k}-\hat{\phi}_{j'k}\right|+\left|r_{j'k}-\frac{y_{j'k}}{\tau_{j'k}}\right|\right)\geq v\,.\end{aligned}$$

Then it holds that $\gamma_{\text{AB}}(\boldsymbol{\alpha})\geq\inf_{v>0}\frac{\gamma(v)}{v}$. We define the following three quantities:

$$\begin{aligned}\gamma_1(v)\triangleq&\min_{\mathbf{f}>\mathbf{0}}\sum_{j'\in V_D,k\in V_S}\Lambda^*(\hat{\phi}_{j'k},f_{j'k})\\ &\text{s.t. }-\min\left\{\min_{J\in\mathcal{J}}\left(\sum_{j'\in V_D,k\in\partial(J)}f_{j'k}-\sum_{j'\in J,k\in V_S}f_{j'k}\right),\min_{i\in V_S}\sum_{j'\in V_S}f_{j'i},\frac{\alpha_{\min}\beta}{8\tau_{\max}}\right\}\\ &\quad\geq-\frac{1}{2}\min\left\{\xi,\lambda_{\min},\frac{\alpha_{\min}\beta}{8\tau_{\max}}\right\}.\\ \gamma_2(v)\triangleq&\min_{\mathbf{f}>\mathbf{0}}\sum_{j'\in V_D,k\in V_S}\Lambda^*(\hat{\phi}_{j'k},f_{j'k})\\ &\text{s.t. }\frac{4}{\alpha_{\min}}\sum_{j'\in V_D,k\in V_S}\left|f_{j'k}-\hat{\phi}_{j'k}\right|\geq\frac{v}{2}+\frac{1}{4}\min\left\{\xi,\lambda_{\min},\frac{\alpha_{\min}\beta}{8\tau_{\max}}\right\}.\\ \gamma_3(v)\triangleq&\min_{\mathbf{r}>\mathbf{0},\mathbf{y}\in\Omega_{n\times m}}\sum_{j'\in V_D,k\in V_S}\Lambda^*\left(\frac{y_{j'k}}{\tau_{j'k}},r_{j'k}\right)\\ &\text{s.t. }\frac{4}{\alpha_{\min}}\sum_{j'\in V_D,k\in V_S}\left|r_{j'k}-\frac{y_{j'k}}{\tau_{j'k}}\right|\geq\frac{v}{2}+\frac{1}{4}\min\left\{\xi,\lambda_{\min},\frac{\alpha_{\min}\beta}{8\tau_{\max}}\right\}.\end{aligned}$$

Note that if $(\mathbf{f},\mathbf{r},\mathbf{y})$ satisfies the constraint in the definition of $\gamma(v)$, then it must satisfy at least one of the constraints in the definition of $\gamma_1(v)$, $\gamma_2(v)$, and $\gamma_3(v)$. Hence

$$\gamma(v)\geq\ \min\{\gamma_1(v),\gamma_2(v),\gamma_3(v)\}\,.$$

Therefore

$$\inf_{v>0}\frac{\gamma(v)}{v}\geq\ \min\left\{\inf_{v>0}\frac{\gamma_1(v)}{v},\inf_{v>0}\frac{\gamma_2(v)}{v},\inf_{v>0}\frac{\gamma_3(v)}{v}\right\}.$$

Now we bound the three quantities on the RHS one by one. Using the same argument as in the no-delay case, we can show that there exists $\delta_1 > 0$ such that $\inf_{v>0} \frac{\gamma_1(v)}{v} > \delta_1$.

For the other two quantities, we first prove the following bound. For $\ell > 0, f > 0$, since $\frac{d^2}{df^2}\Lambda^*(\ell, f) = \frac{1}{f}$, using Taylor expansion we have

$$\Lambda^*(\ell, f) = f \log \frac{f}{\ell} - f + \ell \geq \frac{1}{2f}(f - \ell)^2 .$$

If $f \leq 2\phi$ we have

$$\frac{1}{2f}(f - \phi)^2 \geq \frac{1}{4\phi}(f - \phi)^2$$

Otherwise $\frac{f-\phi}{f} \geq \frac{1}{2}$, hence

$$\frac{1}{2f}(f - \phi)^2 \geq \frac{1}{2}(f - \phi) .$$

Combined, we have

$$\Lambda^*(\ell, f) \geq \frac{1}{\max\{2, 4\phi\}} \min\left\{(f - \phi)^2, |f - \phi|\right\} .$$

Looking at the constraint in the definition of $\gamma_2(v)$, it can be deduced that there must exist $\tilde{j}' \in V_D, \tilde{k} \in V_S$ such that

$$|f_{\tilde{j}'\tilde{k}} - \hat{\phi}_{\tilde{j}'\tilde{k}}| \geq \frac{\alpha_{\min}}{4nm}\left(\frac{v}{2} + \frac{1}{4}\min\left\{\xi, \lambda_{\min}, \frac{\alpha_{\min}\beta}{8\tau_{\max}}\right\}\right) .$$

Denote $g \triangleq \frac{1}{4}\min\left\{\xi, \lambda_{\min}, \frac{\alpha_{\min}\beta}{8\tau_{\max}}\right\} > 0$. Hence

$$\begin{aligned}
\frac{\gamma_2(v)}{v} &\geq \frac{\Lambda^*_{\tilde{j}'\tilde{k}}(\hat{\phi}_{\tilde{j}'\tilde{k}}, f_{\tilde{j}'\tilde{k}})}{v} \\
&\geq \frac{1}{\max\{2, 4\hat{\phi}_{\max}\}}\frac{1}{v}\min\left\{\frac{\alpha_{\min}^2}{16n^2m^2}\left(\frac{v}{2} + g\right)^2, \frac{\alpha_{\min}}{8nm}v\right\} \\
&\geq \frac{1}{\max\{2, 4\hat{\phi}_{\max}\}}\frac{1}{v}\min\left\{\frac{\alpha_{\min}^2}{16n^2m^2}\left(g^2 + gv\right), \frac{\alpha_{\min}}{8nm}v\right\} \\
&\geq \frac{1}{\max\{2, 4\hat{\phi}_{\max}\}}\min\left\{\frac{\alpha_{\min}^2 g}{16n^2m^2}, \frac{\alpha_{\min}}{8nm}\right\} .
\end{aligned}$$

Note that the last term is independent of v and is strictly positive. Therefore there exists $\delta_2 > 0$ such that $\inf_{v>0} \frac{\gamma_2(v)}{v} > \delta_2$. Similarly, we can show that there exists $\delta_3 > 0$ such that $\inf_{v>0} \frac{\gamma_3(v)}{v} > \delta_3$. This establishes that $\gamma_{\text{AB}}(\boldsymbol{\alpha}) > 0$ and concludes the proof. □

B.9 Classical CRP Condition Implies Assumption 3

In this section, we show that the Assumption 3 in our paper is implied by the CRP condition defined in [17]. This justifies our naming of Assumption 3 as the CRP condition.

Note that the CRP condition is defined for open networks in the literature. To facilitate the comparison between the CRP condition and Assumption 3, we first define an open network counterpart of our model: Consider an one-hop queueing system with m queues (indexed by $i \in V_S$) and n servers (indexed by $j' \in V_D$). Jobs arrive to the i-th queue at rate

$$\lambda_i \triangleq \sum_{j' \in V_D} \hat{\phi}_{j'i} \,, \tag{B.64}$$

and the j'-th server processes jobs at rate

$$\mu_{j'} \triangleq \sum_{k \in V_S} \hat{\phi}_{j'k} \,. \tag{B.65}$$

Let $G = (V_S \cup V_D, E)$ be the compatibility graph defined in our paper, and denote the neighborhood of $i \in V_S$ (or $j' \in V_D$) in G by $\partial(i)$ (or $\partial(j')$). To defined the classical CRP condition, we first need to make the following definitions. Define the (primal) static planning problem as the following linear program:

$$\begin{aligned}
\text{minimize}_{\mathbf{x},\rho} \quad & \rho \\
\text{subject to} \quad & \sum_{j' \in \partial(i)} \mu_{j'} x_{ij'} = \lambda_i && \forall i \in V_S \,, \\
& \sum_{i \in \partial(j')} x_{ij'} \le \rho && \forall j' \in V_D \,, \\
& x_{ij'} \ge 0 && \forall i \in V_S \,,\ j' \in V_D \,.
\end{aligned}$$

The dual problem of the static planning problem is the following:

$$\text{maximize}_{\mathbf{y},\mathbf{z}} \quad \sum_{i \in V_S} \lambda_i y_i$$

$$
\begin{aligned}
\text{subject to} \quad & z_{j'} \geq y_i \mu_{j'} && \forall (i,j') \in E, \\
& \sum_{j' \in V_D} z_{j'} = 1 \\
& z_{j'} \geq 0 && \forall j' \in V_D \,.
\end{aligned}
$$

Assumption B.1 (Heavy-traffic CRP condition (Assumptions 1,2 in [17]))**.** *A triple* $(\boldsymbol{\lambda}, \boldsymbol{\mu}, G)$ *is said to be in heavy traffic if the primal static planning problem has a unique solution* $(\mathbf{x}^*, \rho^*)$*, where* $\sum_{i \in \partial(j')} x^*_{ij'} = 1$ *for all* $j' \in V_D$ *and* $\rho^* = 1$*. The triple is said to satisfy the CRP condition if the dual static planning problem has a nonnegative, unique optimal solution* $(\mathbf{y}^*, \mathbf{z}^*)$*.*

Proposition B.2. *For primitives* $(\hat{\boldsymbol{\phi}}, G)$*, define* $\boldsymbol{\lambda}, \boldsymbol{\mu}$ *according to* (B.66) *and* (B.67)*. If* $(\boldsymbol{\lambda}, \boldsymbol{\mu}, G)$ *satisfy Assumption B.1, then* $(\hat{\boldsymbol{\phi}}, G)$ *satisfy Assumption 3 in our paper.*

Proof of Proposition B.2. Consider $(\hat{\boldsymbol{\phi}}, G)$ such that $(\boldsymbol{\lambda}, \boldsymbol{\mu}, G)$ satisfy Assumption B.1. Let $(\mathbf{y}^*, \mathbf{z}^*)$ be the unique optimal solution to the dual static planning problem. Applying Corollary A.1 in [17], we have that $\mathbf{y}^*$ is the unique vector which satisfies

$$
\max_{\mathbf{v} \in V} \; \mathbf{y}^* \cdot \mathbf{v} = 0 \,, \tag{B.66}
$$

$$
\sum_{i \in V_S} \lambda_i y^*_i = 1 \,. \tag{B.67}
$$

Here V is defined as

$$
V \triangleq \left\{ \mathbf{v} \in \mathbb{R}^m \;\middle|\; \begin{array}{l} v_i = \sum_{j' \in \partial(i)} d_{ij'} \mu_{j'} - \lambda_i, \quad \forall i \in V_S \\ \sum_{i \in \partial(j')} d_{ij'} \leq 1, \;\; d_{ij'} \geq 0, \quad \forall i \in V_S, j' \in V_D \end{array} \right\},
$$

which is the set of all possible flow rates out of the queues.

Let $\tilde{\mathbf{y}} = \frac{1}{\sum_{i \in V_S} \lambda_i} \mathbf{1}$, we will show that $\tilde{\mathbf{y}}$ satisfies (B.66) and (B.67), and hence $\mathbf{y}^* = \tilde{\mathbf{y}}$. Eq. (B.67) is easy to verify. For (B.66), because $(\boldsymbol{\lambda}, \boldsymbol{\mu}, G)$ satisfy Assumption B.1, we have

$$
\tilde{\mathbf{y}} \cdot \mathbf{v} = \frac{1}{\sum_{i \in V_S} \lambda_i} \left(\sum_{i \in V_S} \sum_{j' \in \partial(i)} d_{ij'} \mu_{j'} - \sum_{i \in V_S} \lambda_i \right) = \frac{1}{\sum_{i \in V_S} \lambda_i} \left(\sum_{j' \in V_D} \mu_{j'} \sum_{i \in \partial(j')} d_{ij'} - \sum_{i \in V_S} \lambda_i \right) .
$$

According to the definition of V, we have $\sum_{i\in\partial(i)} d_{ij'} \leq 1$, hence

$$\tilde{\mathbf{y}} \cdot \mathbf{v} \leq \frac{1}{\sum_{i\in V_S} \lambda_i} \left(\sum_{j'\in V_D} \mu_{j'} - \sum_{i\in V_S} \lambda_i \right) ,$$

where the equality can be achieved. Applying Assumption B.1, we have

$$\sum_{i\in V_S} \lambda_i = \sum_{i\in V_S} \sum_{j'\in\partial(i)} \mu_{j'} x^*_{ij'} = \sum_{j'\in V_D} \mu_{j'} \sum_{i\in\partial(j')} x^*_{ij'} = \sum_{j'\in V_D} \mu_{j'} .$$

Hence $\tilde{\mathbf{y}}$ satisfies (B.66).

We now prove that $\sum_{i\in I} \lambda_i < \sum_{j'\in\partial(I)} \mu_{j'}$ for all $I \subsetneq V_S, I \neq \emptyset$. For any $I \subsetneq V_S, I \neq \emptyset$, consider the vector $\mathbf{v} \in \mathbb{R}^m$ where

$$\begin{aligned} v_i &= 1 && \text{for } i \in I\,, \\ v_i &= -\frac{|I|}{m-|I|} && \text{for } i \in V_S\backslash I\,. \end{aligned}$$

Because $(\boldsymbol{\lambda}, \boldsymbol{\mu}, G)$ satisfy Assumption B.1, by applying Lemma 5 in [17] we have: for $V^o \triangleq \{\mathbf{v} \in \mathbb{R}^m : \mathbf{1}^{\mathrm{T}}\mathbf{v} = 0\}$, there exists $\delta > 0$ such that $\{\mathbf{v} \in V^o : ||\mathbf{v}||_2 \leq \delta\} \subset V$. It can be easily verified that $\mathbf{v} \in V^o$. As a result, there exists $\delta > 0$ such that $\delta\mathbf{v} \in V$. We have:

$$0 < \delta|I| = \sum_{i\in I} \delta v_i \leq \sum_{j'\in\partial(I)} \mu_{j'} - \sum_{i\in I} \lambda_i\,,$$

We have so far proved that $\sum_{i\in I} \lambda_i < \sum_{j'\in\partial(I)} \mu_{j'}$ for all $I \subsetneq V_S, I \neq \emptyset$, we now show that this implies Assumption 3. First, we establish

$$\sum_{i\in\partial(J)} \lambda_i > \sum_{j'\in J} \mu_j \qquad \forall\ J \subsetneq V_D, J \neq \emptyset\,, \tag{B.68}$$

because, for $I \triangleq V_S\backslash\partial(J)$, we know

$$\begin{aligned} &\sum_{i\in I} \lambda_i < \sum_{j'\in\partial(I)} \mu_{j'} \\ \Rightarrow \sum_{i\in V_S} \lambda_i - \sum_{i\in\partial(J)} \lambda_i < \sum_{j'\in\partial(I)} \mu_{j'} \leq &\sum_{j'\in V_D} \mu_{j'} - \sum_{j'\in J} \mu_{j'} \\ \Rightarrow &\sum_{i\in\partial(J)} \lambda_i > \sum_{j'\in J} \mu_{j'}\,. \end{aligned}$$

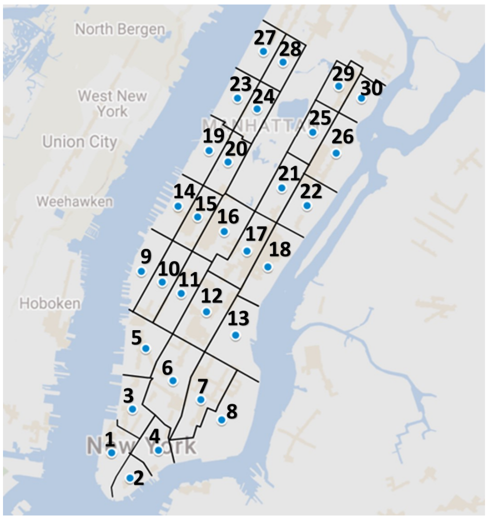

Figure B.1: A 30 location model of Manhattan below 110-th street, excluding the Central Park. (Source: tessellation is based on [121], the figure is generated using Google Maps.)

where we used $\partial(I) \cap J = \emptyset$ by definition of I in the second line, and we used $\sum_{i \in V_S} \lambda_i = \sum_{j' \in V_D} \mu_{j'}$ to get the third line. Our Assumption 3 follows by restricting attention to limited-flexibility subsets J and cancelling the terms which are common on the two sides of the inequality. This concludes the proof. □

B.10 Simulation experiments (full description)

In this appendix, we provide a full description of our simulations in an environment that resembles ride-hailing in Manhattan, New York City. We use demand estimates from [121] (the estimates are based on NYC yellow cab data) and Google Maps to estimate travel times, and simulate SMW-based dispatch policies.

B.10.1 The Data, Simulation Environment and Benchmark

Throughout this section, we use the following set of model primitives.

- *Graph topology.* We consider a 30-location model of Manhattan below 110-th street excluding Central Park (see Figure B.1), based on [121]. We let pairs of regions which share a non-trivial boundary be compatible with each other.

- *Demand arrival process, Pickup/service times, and number of cars.* Throughout this section, we consider a stationary demand arrival rate[5] that satisfies the CRP condition, which is obtained by "symmetrizing"[6] the decensored demand estimated in [121] (see subsection B.10.5 for a full description). We estimate travel times between location pairs using Google Maps, and use as a baseline the fluid requirement K_{fl} on number of cars needed to meet demand. We use K_{tot} (not K) to denote the total number of cars, and $K_{\text{slack}} = K_{\text{tot}} - K_{\text{fl}}$ to denote the excess over the fluid requirement. Here K_{slack} is similar to the K in our theory since it is the average number of free cars assuming all demand is met.

Simulation Design. We consider the following simulation settings:

1. *Stationary performance with Service time.* We investigate steady state performance; steady state is reached in ∼1-2 hours under SMW policies.
2. *Stationary performance with Service+Pickup time.* Same as above.
3. *Transient performance with Service + Pickup time.* We investigate performance over a short horizon (below 2 hours) for different initial configurations.

Benchmark policy: fluid-based policy. The benchmark policy we consider is a static randomization based on the solution to the fluid problem [1, 21]. See subsection B.10.5 for details.

Learning the optimal parameters. We use MATLAB's built-in `particleswarm` solver to learn the optimal SMW scaling parameters via simulation-based optimization in each setting.

[5]We leave the cases where demand is time-varying for future research. Our numerical study in Section B.10.4 regarding transient performance may be seen as a first step towards the time-varying case.

[6]Instead of symmetrizing, an alternative would be to consider an "empty" relocation rule (see Section 2.8) such that CRP holds. We obtained similar results under this alternative (we omit those results in the interest of space).

B.10.2 Steady state with Service times

A preliminary simulation of the setting in our paper (i.e., pickup and service are both instantaneous) showed that under vanilla MaxWeight policy we only need $K_{\text{slack}} = 120$ to obtain a demand-loss rate below 1%, under SMW(α) with α defined in Theorem 2.1 the number further reduces to 80. However, the demand-loss rate stays above 5% under the fluid-based policy even when $K_{\text{slack}} = 200$.[7] We then proceeded to simulate the Service time setting, and obtained similarly encouraging results. In this setting, the average trip time is 13.2 minutes, and the fluid requirement is $K_{\text{fl}} = 7,061$ cars.

Results. The simulation results on performance[8] are shown in Figure B.2, and the theoretical and learned α are shown in Figure B.3. Figure B.2 confirms that SMW policies including vanilla MaxWeight outperform the fluid-based policy; in fact only $K_{\text{slack}} = 100$ extra cars ($<1.5\%$ of K_{tot}, or <4 free cars per location on average if all demand is met) in the system lead to a negligible fraction of demand lost. The demand loss probability decays rapidly with K_{slack} under SMW policies, while it decays much slower under the fluid-based policy. SMW with parameters chosen based on Theorem 2.1 performs nearly as well as the learned SMW policy, despite small $K_{\text{slack}} = 100$. Figure B.3 shows that the learned α is very similar to the theoretically optimal α structurally. Both policies allocate larger parameters (i.e., give more protection to the supply) in the Upper West Side area which has a small Hall's gap (i.e., small slack in the CRP condition).

B.10.3 Steady state with Service and Pickup times

In the following experiment we further incorporate pickup times. The average pickup time is 5.5 minutes, and the fluid requirement increases to $K_{\text{fl}} = 10,002$ cars. Our objective here is to show that SMW policies can be heuristically adapted to more general settings, and retain their good performance. We propose the following SMW-based

[7] The results remain similar when service time is included, hence we only include the graph of the latter case.

[8] We also tested stochastic service times and found no significant difference in performance.

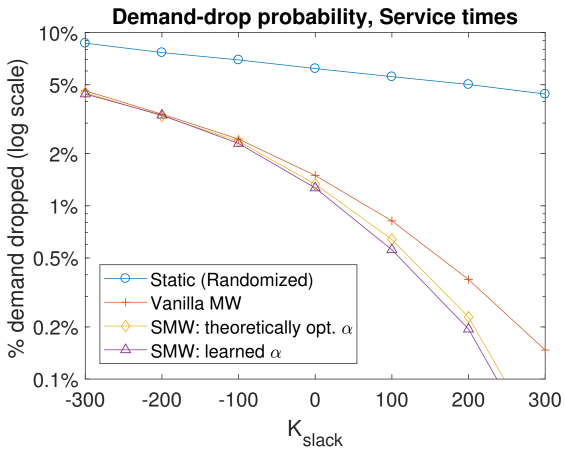

Figure B.2: *Service times* setting: Stationary demand-loss probability under the static fluid-based policy, vanilla MaxWeight policy, SMW policy with theoretically optimal α, and SMW policy with learned α. Note that the y-axis is in log-scale. Here $K_{\text{fl}} = 7,061$. The plots indicate significant separation between fluid and SMW policies at all values of K, and separation between vanilla MaxWeight and optimized SMW. For each data point we run 200 trials and take the average.

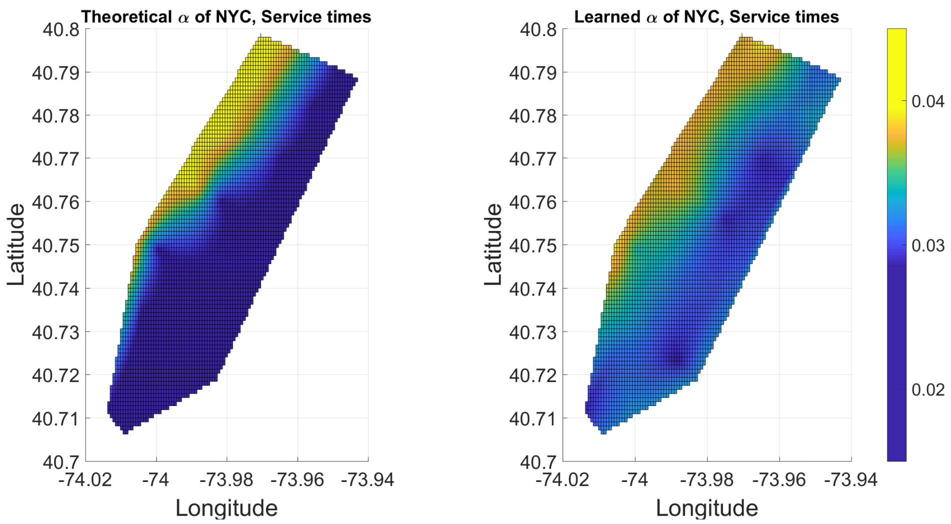

Figure B.3: *Service times* setting: Theoretically optimal α derived from Theorem 2.1 (left) and the α learned via simulation-based optimization (right), both for the NYC dataset with $K_{\text{slack}} = 200$. Darker shades indicate smaller values of α_i, while lighter shades correspond to larger values.

heuristic policy. Intuitively, pickup times need to be taken into consideration when making dispatch decisions, because every minute spent on picking up a customer leads to an opportunity cost. We consider policies of the following form. When demand arrives at

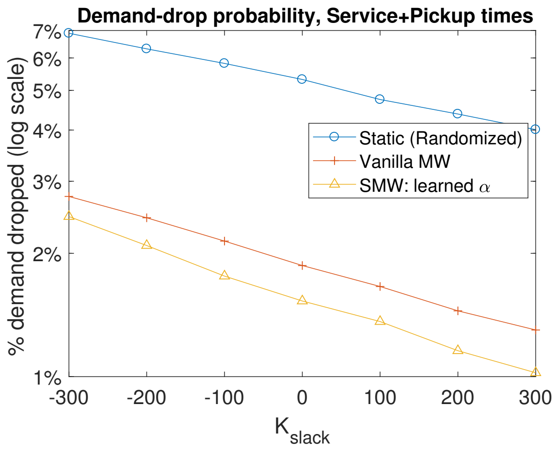

Figure B.4: *Service+Pickup times* setting: Stationary demand-loss probability under the fluid-based policy, the vanilla MaxWeight policy, and the SMW policy with α learned via simulation optimization. Here $K_{\text{fl}} = 10,002$ cars. For each data point we average over 200 trials.

location j, dispatch from

$$\text{argmax}_{i\in\partial(j)} \frac{x_i}{\alpha_i} - zD_{ij},$$

where x_i is the number of free cars at i, and D_{ij} is the pickup time between i and j. In addition to scaling parameters α, we have an additional parameter z which captures the importance given to pickup delay in making dispatch decisions.

Results. Simulation results are shown in Figure B.4. We observe that the SMW-based policies including vanilla MaxWeight significantly outperform the fluid-based policy. A few hundred extra cars ($< 3\%$ of K_{tot}) in the system suffice to ensure that only $\sim 1\%$ of demand is lost.

B.10.4 Transient Behavior with Service and Pickup times

In the last experiment, we consider transient behavior instead of steady state performance. We fix K_{slack} to be 200. For initial configurations, we sample 4 initial queue-length vectors uniformly from the simplex $\{\mathbf{x} : x_1 + \cdots + x_{30} = 200\}$, and the cars initially in transit are based on picking up all demand that arose in the last hour. For each initial state we consider 4 time horizons: 0.5, 1, 1.5 and 2 hours. We learn the optimal

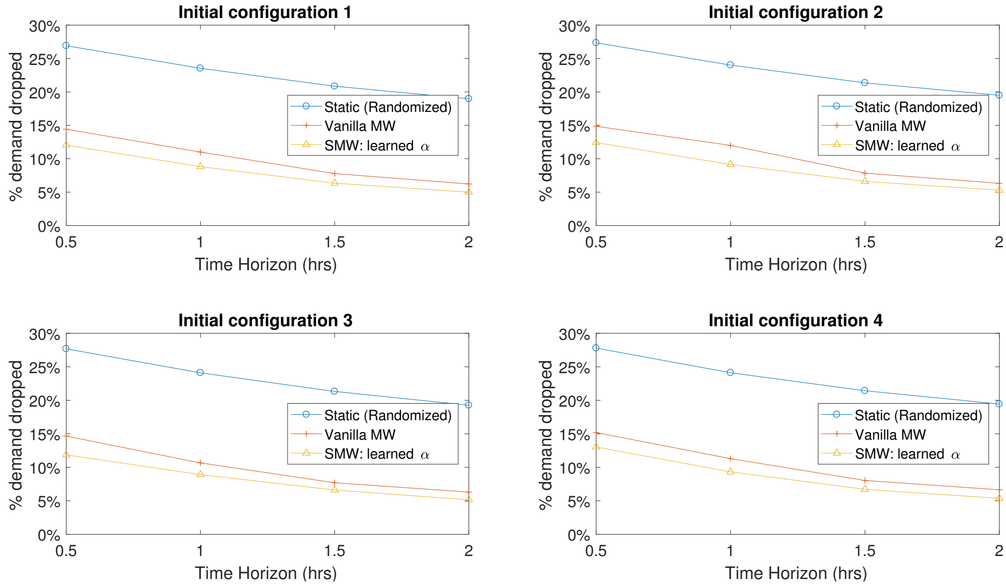

Figure B.5: *Transient Performance with Service+Pickup times*: The plots show the demand-loss probability under the fluid-based policy, the vanilla MaxWeight policy, and the SMW policy with learned α, with 4 different initial configurations, chosen randomly on the simplex. We fix $K_{\text{slack}} = 200$, and consider time horizons ranging from 0.5 to 2 hours. For each data point we run 200 trials and take the average.

SMW parameters for each initial state and time horizon pair to minimize the fraction of demand lost and then compare the performance of SMW policies, vanilla MaxWeight and the fluid-based policy. The results are shown in Figure B.5. It turns out that SMW policies outperform the fluid-based policy by an even larger margin in this case since they are able to quickly (in under an hour) spread the supply out across locations.

B.10.5 Simulation Settings

In this subsection, we fill in the missing details in the previous subsections.

Model Primitives.

- *Demand arrival process (ϕ).* Using the estimation in [121], which is based on Manhattan's taxi trip data during August and September in 2012, we obtain the (average) demand arrival rates for each origin-destination pair during the day (7 a.m. to 4 p.m.) denoted by $\tilde{\phi}_{ij}$ $(i, j = 1, \cdots, 30)$. However, we find that $\tilde{\phi}_{ij}$ violates CRP (there are a lot more rides *to* Midtown than *from* Midtown). We consider the following "sym-

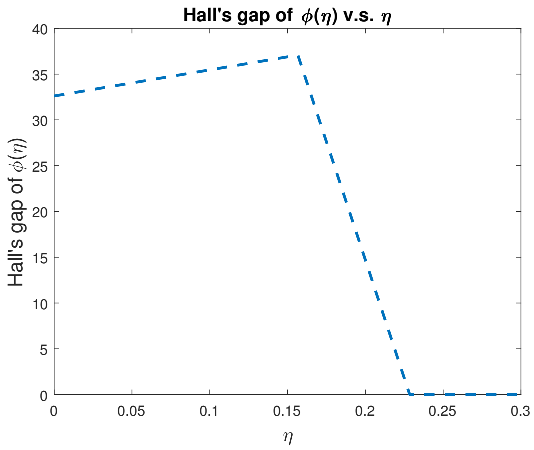

Figure B.6: Hall's gap of symmetrized matrix $\phi(\eta)$ (see Eq. (B.69)) versus parameter η, based on the demand arrival rates $\tilde{\phi}$ computed from the Manhattan taxi data. Our simulations use $\eta = 0.21$, which corresponds to a small but non-zero Hall's gap (< 10).

metrization" of $\tilde{\phi} \triangleq (\tilde{\phi}_{ij})_{30\times 30}$ to ensure that CRP holds (ride-hailing platforms may use spatially varying prices and repositioning to obtain CRP, see Section 2.1):

$$\phi(\eta) \triangleq \eta\tilde{\phi} + (1-\eta)\frac{1}{2}(\tilde{\phi} + \tilde{\phi}^{\mathrm{T}}), \qquad \eta \in (0,1). \tag{B.69}$$

Figure B.6 shows how the Hall's gap of $\phi(\eta)$ varies with η. We pick $\eta = 0.21$ such that CRP is "almost violated"[9]. The subset of locations with smallest Hall's gap is then the Upper West Side (locations 19, 23, 24, 27, 28 in Figure B.1).

- *Pickup/service times* ($D/\tilde{D}$). We extract the pairwise travel time between region centroids (marked by the dots in Figure B.1) using Google Maps, denoted by D_{ij}'s ($i, j = 1, \cdots, 30$). We use D_{ij} as service time for customers traveling from i to j. For each customer at i who is picked up by a supply from k we add a pickup time [10] of $\tilde{D}_{ki} = \max\{\frac{3}{2}D_{ki}, 3 \text{ minutes}\}$.

Benchmark policy: fluid-based policy. We consider the fluid-based randomized policy [1, 21] as a benchmark. Let $\mathcal{X}$ be the solution set of the feasibility problem

$$\sum_{j\in\partial(i)} x_{ij} = \lambda_i \quad \forall i, \qquad \sum_{i\in\partial(j)} x_{ij} = \mu_j \quad \forall j. \tag{B.70}$$

[9] We also ran simulations for $\eta = 0.15$ such that Hall's gap is large. There is no significant difference in the policies' relative performances, so we didn't include it here.

[10] We use the inflated D_{ij}'s as pickup times to account for delays in finding or waiting for the customer.

Since CRP holds, $\mathcal{X} \neq \emptyset$. Let $\mathbf{x}^* \triangleq \text{argmin}_{\mathbf{x} \in \mathcal{X}} \sum_{(i,j)\in E} \tilde{D}_{ij} x_{ij}$. When demand arrives at location j, the randomized fluid-based policy dispatches from location $i \in \partial(j)$ with probability x^*_{ij}/μ_j. Then x^*_{ij} is the rate of dispatching cars from i to serve demand at j. From [1], we know that $\mathbf{x}^*$ leads to a zero demand-loss as $K \to \infty$ with and without pickup times (assuming demand remains constant). Moreover, with pickup times, Little's Law gives that the fluid-based policy minimizes the expected number of cars on-route to pick up customers.

Benchmark fleet-size. In the *Service time* setting, a fraction of cars are in transit under the stationary distribution; in the *Service+Pickup time* setting, there is an additional fraction of cars on-route to pick up customers. A simple workload conservation argument (using Little's Law) gives the benchmark fleet-sizes as follows.

- *Service time.* Assuming no demand is lost, the mean number of cars in transit is: $K_{\text{fl}} = \sum_{i,j} \phi_{ij} D_{ij}$. In our setting, we have $K_{\text{in-transit}} \approx 7,061$. Since CRP holds and demand-loss probability goes to 0 under both fluid-based policy and SMW policies, $K_{\text{in-transit}}$ is a reasonable benchmark fleet-size K_{fl}. We will vary the number of cars in the system denoted by $K_{\text{tot}} = K_{\text{fl}} + K_{\text{slack}}$ and compare the performance of different policies. Here K_{slack} is the number of free cars in the system when no demand is lost.

- *Service+Pickup time.* Applying Little's Law, if no demand is lost, the mean number of cars picking up customers is at least $K_{\text{pickup}} = \min_{\mathbf{x} \in \mathcal{X}} \tilde{D}_{ij} x_{ij}$. In our case, we have $K_{\text{pickup}} \approx 2,941$. Hence, the benchmark fleet size is $K_{\text{fl}} = K_{\text{in-transit}} + K_{\text{pickup}} = 10,002$. Note that this number is close to the real-world fleet size: there were approximately 11,500 active medallions when [121] was written.

APPENDIX C

Proofs in "In Which Random Matching Markets Does the Short Side Enjoy an Advantage"

Organization of the appendix. The technical appendix is organized as follows.

- Appendix C.1 describes several concentration inequalities and auxiliary stochastic processes that will be heavily used in the following theoretical analysis.
- Appendix C.2 establishes Theorem 3.1, the main result for moderately connected markets. The proof is lengthy and will be further divided into several steps, with an overview provided at the beginning of each step.
- Appendix C.3 establishes Theorem 3.2, the main result for densely connected markets.

C.1 Preliminaries

C.1.1 Basic Inequalities

Lemma C.1. *The following inequalities hold:*

- *For any* $|x| \leq \frac{1}{2}$*, we have* $e^{-x-x^2} \leq 1 - x \leq e^{-x}$*.*
- *For any* $k > 0$ *and* $\epsilon \in \left(0, \frac{1}{k}\right)$*, we have* $1 + k\epsilon \leq \frac{1}{1-k\epsilon}$*.*

C.1.2 Negative Association of Random Variables

The concept of negative association provides a stronger notion of negative correlation, which is useful to analyze the concentration of the sum of dependent random variables.

Definition C.1 (Negatively Associated Random Variables [129])**.** *A set of random variables $X_1, X_2, \ldots, X_n$ are* negatively associated *(NA) if for any two disjoint index sets $I, J \subseteq \{1, \ldots, n\}$,*

$$\mathbb{E}\left[f(X_i : i \in I) \cdot g(X_j : j \in J)\right] \leq \mathbb{E}\left[f(X_i : i \in I)\right] \cdot \mathbb{E}\left[g(X_j : j \in J)\right]$$

for any two functions $f : \mathbb{R}^{|I|} \mapsto \mathbb{R}$ and $g : \mathbb{R}^{|J|} \mapsto \mathbb{R}$ that are both non-decreasing or both non-increasing (in each argument).

The following lemma formalizes that the sum of negatively associated (NA) random variables is as concentrated as the sum of independent random variables:

Lemma C.2 (Chernoff-Hoeffding Bound for Negatively Associated Random Variables [129])**.** *Let $X_1, X_2, \ldots, X_n$ be NA random variables with $X_i \in [a_i, b_i]$ always. Then, $S \triangleq \sum_{i=1}^n X_i$ satisfies the following tail bound:*

$$\mathbb{P}\left(|S - \mathbb{E}[S]| \geq t\right) \leq 2 \exp\left(-\frac{2t^2}{\sum_{i=1}^n (b_i - a_i)^2}\right). \tag{C.1}$$

We refer to [129] for the proof.

The following lemma provides sufficient conditions for a set of random variables to be NA. For each sufficient condition, we provide a pointer to a paper where it has been established.

Lemma C.3 (Sufficient Conditions for Negative Association)**.** *The followings hold:*

(i) (Permutation distribution [130, Theorem 2.11]) *Let $x_1, x_2, \ldots, x_n$ be n real numbers and let $X_1, X_2, \ldots, X_n$ be random variables such that $(X_1, X_2, \ldots, X_n)$ is a uniformly random permutation of $(x_1, x_2, \ldots, x_n)$. Then $X_1, X_2, \ldots, X_n$ are NA.*

(ii) (Union of independent sets of NA random variables [130, Property 7]) *If $X_1, X_2, \ldots, X_n$ are NA, $Y_1, Y_2, \ldots, Y_m$ are NA, and $\{X_i\}_i$ are independent of $\{Y_j\}_j$, then $X_1, \ldots, X_n, Y_1, \ldots, Y_m$ are NA.*

(iii) (Concordant monotone functions [130, Property 6]) *Increasing functions defined on disjoint subsets of a set of NA random variables are NA. More precisely, suppose $f_1, f_2, \ldots, f_k$ are all non-decreasing in each coordinate, or all non-increasing in each coordinate, with each $f_j : \mathbb{R}^{|I_j|} \mapsto \mathbb{R}$ defined on $(X_i)_{i \in I_j}$ for some disjoint index subsets $I_1, \ldots, I_k \subseteq \{1, \ldots, n\}$. If $X_1, X_2, \ldots, X_n$ are NA, then the set of random variables $Y_1 \triangleq f_1(X_i : i \in I_1), Y_2 \triangleq f_2(X_i : i \in I_2), \ldots, Y_k \triangleq f_k(X_i : i \in I_k)$ are NA.*

C.1.3 Balls-into-bins

A balls-into-bins process with T balls and n bins is defined as follows: at each time $t = 1, \ldots, T$, a ball is placed into one of n bins uniformly at random, independently of the past. Index the bins by $j \in \{1, \ldots, n\}$, and let $I_{j,t}$ be an indicator variable that equals one if the t^{th} ball is placed in the j^{th} bin and equals zero otherwise. Further define $W_j \triangleq \sum_{t=1}^{T} I_{j,t}$ representing the total number of balls placed into the j^{th} bin.

A particular random variable of interest is the number of empty bins at the end of the process, for which we have the following concentration inequality.

Lemma C.4 (Number of empty bins)**.** *Let X be the number of empty bins at the end of a balls-into-bins process with T balls and n bins. For any $\epsilon > 0$, we have*

$$\mathbb{P}\left(\frac{1}{n}X - \left(1 - \frac{1}{n}\right)^T \geq \epsilon\right) \leq \exp\left(-2n\epsilon^2\right) ,$$
$$\mathbb{P}\left(\frac{1}{n}X - \left(1 - \frac{1}{n}\right)^T \leq \epsilon\right) \leq \exp\left(-2n\epsilon^2\right) .$$

Proof. Observe that $\{I_{j,t}\}_{j \in \{1,\ldots,n\}, t \in \{1,\ldots,T\}}$ are negatively associated (NA) since $\{I_{j,t}\}_{j \in \{1,\ldots,n\}}$ are NA for each t (by Lemma C.3–((i)), since $\{I_{j,t}\}_{j \in \{1,\ldots,n\}}$ is a uniformly random permutation of $n-1$ zeros and a single one) and they are independent across t (Lemma C.3–((ii))). Consequently, $W_1, \ldots, W_n$ are NA due to Lemma C.3–((iii)), since $f_j(I_{j,1}, \ldots, I_{j,T}) \triangleq \sum_{t=1}^{T} I_{j,t}$

is non-decreasing in each coordinate.

Define $Y_j \triangleq \mathbb{I}(W_j = 0)$ indicating whether the j^{th} bin is empty at the end. Although Y_j's are not independent, they are NA (again, by Lemma C.3–((iii))). Because $Y_j \sim \text{Bernoulli}\left(\left(1-\frac{1}{n}\right)^T\right)$ and $X = \sum_{j=1}^n Y_j$, by applying Hoeffding's bound (Lemma C.2), we obtain the desired result. □

Lemma C.5. *Let W_j denotes the number of balls in the j^{th} bin at the end of a balls-into-bins process with T balls and n bins. For any $\Delta > 0$, we have*

$$\mathbb{P}\left(\frac{1}{n}\sum_{j=1}^{n}\frac{1}{W_j+1} \geq \frac{n}{T} + \Delta\right) \leq \exp\left(-2n\Delta^2\right).$$

Proof. Since $W_j \sim \text{Binomial}\left(T, \frac{1}{n}\right)$, we have

$$\begin{aligned}
\mathbb{E}\left[\frac{1}{W_j+1}\right] &= \sum_{k=0}^{T}\frac{1}{k+1}\times\binom{T}{k}\left(\frac{1}{n}\right)^k\left(1-\frac{1}{n}\right)^{T-k} \\
&= \frac{n}{T+1}\sum_{k=0}^{T}\binom{T+1}{k+1}\left(\frac{1}{n}\right)^{k+1}\left(1-\frac{1}{n}\right)^{(T+1)-(k+1)} \\
&= \frac{n}{T+1}\times\left(1-\left(1-\frac{1}{n}\right)^{T+1}\right) \leq \frac{n}{T}.
\end{aligned}$$

In the proof of Lemma C.4, we have shown that $W_1, \ldots, W_n$ are NA. By Lemma C.3–((iii)), $\frac{1}{W_1+1}, \ldots, \frac{1}{W_n+1}$ are also NA. Therefore, by applying Hoeffding's bound (Lemma C.2), we obtain the desired result. □

C.1.4 Chernoff's Bound on Random Sum

Lemma C.6. *Fix any $p \in (0,1)$ and any $p' \in (0,1)$. Define the random sum*

$$S \triangleq \sum_{i=1}^{N} X_i\,,$$

where X_i's are i.i.d. random variables and have distribution[1] $\text{Geometric}(p)$*, and $N \sim$* $\text{Geometric}(p')$ *and is independent of X_i's. Let S_i's be i.i.d. random variables and have*

[1]Here, by Geometric(p) we mean the distribution $\mathbb{P}(X_i = k) = p(1-p)^{k-1}$ for $k \geq 1$, i.e., the support of the distribution is $\{1, 2, \ldots\}$, and its expectation is $1/p > 1$.

the same distribution as S, for $\lambda > \mathbb{E}[S] = 1/(pp')$ we have

$$\mathbb{P}\left(\frac{1}{n}\sum_{i=1}^{n} S_i \geq \lambda\right) \leq \exp\left(-\frac{n}{2\lambda^2}\left(\lambda - \mathbb{E}[S]\right)^2\right).$$

Proof. Denote $q \triangleq 1-p$, $q' \triangleq 1-p'$. In the first step, we derive the moment generating function of S, which we denote by $M(t)$. Note that

$$M(t) = \mathbb{E}[e^{tS}] = \mathbb{E}\left[e^{t\sum_{i=1}^{N} X_i}\right] = \mathbb{E}\left[\left[\mathbb{E}e^{tX}\right]^N\right] = \mathbb{E}\left[\gamma^N\right].$$

where

$$\gamma \triangleq \begin{cases} \frac{pe^t}{1-qe^t} & \text{if } qe^t < 1\,, \\ \infty & \text{otherwise}\,. \end{cases} \tag{C.2}$$

we have

$$\mathbb{E}[\gamma^N] = \sum_{k=1}^{\infty} \gamma^k (1-p')^{k-1} p' = p'\gamma \sum_{k=1}^{\infty} \gamma^{k-1} (q')^{k-1} = \begin{cases} \frac{p'\gamma}{1-\gamma q'} & \text{if } \gamma q' < 1\,, \\ \infty & \text{otherwise}\,. \end{cases}$$

By plugging in γ, we obtain

$$M(t) = \begin{cases} \frac{p'\frac{pe^t}{1-qe^t}}{1-q'\frac{pe^t}{1-qe^t}} = \frac{p'pe^t}{1-e^t(q+q'p)} & \text{if } t < \bar{t} \triangleq \log(1/(q+q'p))\,, \\ \infty & \text{otherwise}\,. \end{cases}$$

Here we used that $q+q'p > q$ to simplify the condition for $M(t)$ to be finite to $e^t(q+q'p) < 1 \Leftrightarrow t < \bar{t}$.

Now we derive the convex conjugate of $\log M(t)$, a.k.a. the large deviation rate function. Note that $\mathbb{E}[S] = 1/(pp')$. Define $\Lambda^* : [1/(pp'), \infty) \to \mathbb{R}$ as

$$\Lambda^*(\lambda) \triangleq \sup_{t \geq 0} \left(\lambda t - \log M(t)\right) = \sup_{t \in [0,\bar{t})} \left(\lambda t - \log M(t)\right)$$

Fix $\lambda \geq 1/(pp')$ and let t^* be the maximizer of the supremum above. The derivative of $\lambda t - \log M(t)$ with respect to t for $t \in [0, \bar{t})$ is

$$\lambda - 1 - \frac{e^t(q+q'p)}{1-e^t(q+q'p)} = \lambda - \frac{1}{1-e^t(q+q'p)}\,,$$

and in particular it is decreasing in t, corresponding to the fact that $\lambda t - \log M(t)$ is concave in t (we already knew concavity holds because the log moment generating function is always convex). Note further that the derivative at $t = 0$ is non-negative since

$$\lambda - \frac{1}{1-(q+q'p)} = \lambda - 1/(pp') \geq 0\,,$$

and that the derivative eventually becomes negative since it tends to $-\infty$ as $t \to \bar{t}^-$. Hence the first order condition will give us the maximizer $t^* \in [0, \bar{t})$ of $\lambda t - \log M(t)$ as follows:

$$\lambda = \frac{1}{1 - e^{t^*}(q+q'p)} \qquad \Rightarrow \qquad e^{t^*} = \frac{1-\frac{1}{\lambda}}{q+q'p}\,.$$

Therefore, we have

$$\begin{aligned}\Lambda^*(\lambda) &= \lambda \log\left(1 - \frac{1}{\lambda}\right) - \lambda \log\left(q+q'p\right) - \log\left(\frac{p'p\frac{1-\frac{1}{\lambda}}{q+q'p}}{1/\lambda}\right)\\ &= \lambda \log\left(1 - \frac{1}{\lambda}\right) - \lambda \log\left(q+q'p\right) - \log\left(\lambda - 1\right) + C,\end{aligned}$$

where C is a constant. A short calculation tells us that

$$\frac{d\Lambda^*}{d\lambda}(\lambda) = \log\left(1 - \frac{1}{\lambda}\right) - \log\left(q+q'p\right)\,, \qquad \frac{d^2\Lambda^*}{d\lambda^2}(\lambda) = \frac{1}{\lambda(\lambda-1)}\,. \tag{C.3}$$

Let $S_1, \cdots, S_n$ be i.i.d. random variables with the same distribution as S. Using Chernoff's bound, for $\lambda \geq \mathbb{E}[S]$ we have

$$\mathbb{P}\left(\frac{1}{n}\sum_{i=1}^{n} S_i \geq \lambda\right) \leq \exp\left(-n\Lambda^*(\lambda)\right). \tag{C.4}$$

Since $\Lambda^*(\cdot)$ is a large deviation rate function, we have that $\Lambda^*(\mathbb{E}[S]) = 0$ and $\frac{d\Lambda^*}{d\lambda}(\mathbb{E}[S]) = 0$. We will now use Taylor's theorem taking terms up to second order for $\Lambda^*(\lambda)$ around $\mathbb{E}[S]$ to obtain the desired bound. Note that at any $\lambda' \in (\mathbb{E}[S], \lambda)$, using the explicit form of $\frac{d^2\Lambda^*}{d\lambda^2}$ in (C.3) we have

$$\frac{d^2\Lambda^*}{d\lambda^2}(\lambda') \geq \frac{1}{(\lambda')^2} \geq \frac{1}{\lambda^2}\,,$$

where we used $\mathbb{E}[S] > 1$. Now, using Taylor's theorem, we know that for some $\lambda' \in (0, \lambda)$

we have

$$\Lambda^*(\lambda) = \frac{1}{2}\frac{d^2\Lambda^*}{d\lambda^2}(\lambda')\left(\lambda - \mathbb{E}[S]\right)^2 \geq \frac{1}{2\lambda^2}\left(\lambda - \mathbb{E}[S]\right)^2 .$$

Plugging into (C.4), we obtain

$$\mathbb{P}\left(\frac{1}{n}\sum_{i=1}^{n} S_i \geq \lambda\right) \leq \exp\left(-\frac{n}{2\lambda^2}\left(\lambda - \mathbb{E}[S]\right)^2\right)$$

as required. □

C.1.5 Notations and Preliminary Observations

We here introduce the variables that formally describe the state of a random matching market over the course of the men-proposing deferred-acceptance (MPDA) procedure (Algorithm 3.1).

The *time* t ticks whenever a man makes a proposal. Let $I_t \in \mathcal{M}$ be the man who proposes at time t, and $J_t \in \mathcal{W}$ be the woman who receives that proposal. We define $M_{i,t} \triangleq \sum_{s=1}^{t} \mathbb{I}(I_s = i)$ that counts the number of proposals that a man i has made up to time t, and define $W_{j,t} \triangleq \sum_{s=1}^{t} \mathbb{I}(J_s = j)$ that counts the number of proposals that a woman j has received up to time t. We will often use $\vec{M}_t \triangleq (M_{i,t})_{i\in\mathcal{M}}$ and $\vec{W}_t \triangleq (W_{j,t})_{j\in\mathcal{W}}$ as vectorized notations. By definition, we have

$$\sum_{i\in\mathcal{M}} M_{i,t} = \sum_{j\in\mathcal{W}} W_{j,t} = t\,,$$

for any $0 \leq t \leq \tau$ where τ is the total number of proposals under MPDA.

Let $\mathcal{H}_t \subseteq \mathcal{W}$ be the set of women that the man I_t had proposed to before time t: i.e., $\mathcal{H}_t \triangleq \{J_s : I_s = i \text{ for some } s \leq t-1\}$ and we have $|\mathcal{H}_t| < d$.According to the principle of deferred decisions, the t^{th} proposal goes to one of women that the man I_t had not proposed to yet: i.e., J_t is sampled from $\mathcal{W} \setminus \mathcal{H}_t$ uniformly at random. And then, the proposal gets accepted by the woman J_t with probability $1/(W_{J_t,t-1}+1)$.

We denote *the current number of unmatched men* and *women* at time t by $\delta^m[t]$ and $\delta^w[t]$, respectively. More precisely, $\delta^m[t]$ represents the number of men who have exhausted all his preference list but left unmatched[2] at time t: i.e., $\delta^m[t] \triangleq \sum_{i\in\mathcal{M}} \mathbb{I}(M_{i,t} = d, \mu_t(i) = i)$ where μ_t is the current matching at time t. Also note that once a woman receives a proposal, she remains matched until the end of MPDA procedure: i.e., $\delta^w[t] \triangleq \sum_{j\in\mathcal{W}} \mathbb{I}(\mu_t(j) = j) = \sum_{j\in\mathcal{W}} \mathbb{I}(W_{j,t} = 0)$. We observe that $\delta^m[t]$ starts from zero (at $t = 0$) and is *non-decreasing* over time, and $\delta^w[t]$ starts from n and is *non-increasing* over time.

Recall that τ is the *the total number of proposals* that is made until the end of MPDA, i.e., the time at which the men-optimal stable matching (MOSM) is found. MPDA ends when there is no more man to make a proposal, i.e., when every unmatched man had already exhausted his preference list. In (3.1), we expressed τ as a stopping time, namely,

$$\tau = \min\{t \geq 1 : \delta^m[t] = \delta^w[t] + k\}\,.$$

In particular, we have

$$\delta^m[\tau] = \delta^w[\tau] + k\,,$$

since the number of matched men equals to the number of matched women under any feasible matching. Furthermore, we have

$$R_{\text{MEN}}(\text{MOSM}) = \frac{\tau + \delta^m[\tau]}{n + k}\,,$$

by the definition of men's rank.

An extended process. We introduce an *extended process* as a natural continuation of the MPDA procedure that continues to evolve even after the MOSM is found (i.e., the extended process continues for $t > \tau$). Recall that the MPDA procedure under the principle of deferred decisions works as follows: As described in Algorithm 3.1, $n + k$ men in $\mathcal{M}$ sequentially enter the market one by one, and whenever a new man enters, he makes a proposal and the acceptance/rejection process continues until all men who

[2]It is important that the definition of $\delta^m[t]$ does not count the men who have not entered the market until time t. In other words, it counts the number of men who are "confirmed" to be unmatched under MOSM, and correspond to the variable δ^m described in Algorithm 3.1. This quantity is different from the number of unmatched men under the current matching μ_t, which may decrease when a man proposes to a woman who has never received any proposal.

have entered are either matched or have reached the bottom of their preference lists (i.e., until it finds a new MOSM among the men who have entered including the newly entered man).

To define the extended process, we start by defining an extended market, which has the same n women but an infinite supply of men: $n+k$ "real" men $\mathcal{M}$ who are present in the original market, and an infinity of "fake" men $\mathcal{M}_{\text{fake}}$ in addition. The distribution of preferences in the extended market is again as described in Section 3.2 (in particular, the preference distribution does not distinguish real and fake men). We then define the *extended process* as tracking the progress of Algorithm 3.1 on the extended market: the $n+k$ real men enter first in Algorithm 3.1, as before, and we then continue Algorithm 3.1 after time τ for all $t > \tau$ by continuing to introduce additional (fake) men sequentially after time τ. In particular, the extended process is identical to the original MPDA process until the MOSM is found (i.e., for $t \leq \tau$).

Observe that in this extended process, the MOSM among $\mathcal{M} \cup \mathcal{W}$ can be understood as a stable outcome found after $n+k$ men have entered the market. Therefore, all the aforementioned notations (I_t, J_t, $M_{i,t}$, $W_{j,t}$, $\mathcal{H}_t$, μ_t, $\delta^m[t]$, $\delta^w[t]$) are well-defined for any time $t \geq 0$ while preserving all their properties characterized above, and we similarly denote by $\hat{\mathcal{M}}[t] \subset \mathcal{M} \cup \mathcal{M}_{\text{fake}}$ the set of men who have entered so far (consistent with the notation in Algorithm 3.1). In the later proofs, we utilize these notations and their properties (e.g., $\delta^m[\tau] \leq \delta^m[t]$ implies that $\tau \leq t$ since $\delta^m[t]$ is non-decreasing over time for $t = 0, 1, \ldots$).

Balls-into-bins process analogy. When we analyze the women side, we heavily utilize the balls-into-bins process as done in [71]. We make an analogy between the number of proposals that each of n women has received (denoted by $W_{j,t}$) and the number of balls that had been placed into each of n bins. For example, the number of unmatched women at time t corresponds to the number of empty bins after t balls had been placed.

Recall that, according to the principle of deferred decisions, the t^{th} proposal goes to one of women uniformly at random among whom he had not yet proposed to (i.e., $\mathcal{W} \setminus \mathcal{H}_t$), and thus the recipients of proposals, $J_1, J_2, \ldots$, are not independent. In the balls-into-bins process, in contrast, the t^{th} ball is placed into one of n bins uniformly at random, independently of the other balls' placement. Despite this difference (sampling without replacement v.s. sampling with replacement), the balls-into-bins process provides a good enough approximation as the number of proposals made by an individual man (i.e., $|\mathcal{H}_t|$) is much smaller than the total number of men and women. We will show that (e.g., in Lemma C.8 in the next section) that the corresponding error term can be effectively bounded.

C.2 Proof for Small to Medium-Sized d: the case of $d = o(\log^2 n)$, $d = \omega(1)$

In this section, we consider the case such that $d = o(\log^2 n)$ and $d = \omega(1)$. We will prove the following quantitative version of Theorem 3.1.

Theorem C.1 (Quantitative version of Theorem 3.1)**.** *Consider a sequence of random matching markets indexed by n, with $n+k$ men and n women ($k = k(n)$ can be positive or negative), and the men's degrees are $d = d(n)$. If $|k| = O(ne^{-\sqrt{d}})$, $d = \omega(1)$ and $d = o(\log^2 n)$, then with probability $1 - O(\exp(-d^{\frac{1}{4}}))$ we have*

1. (Men's average rank of wives)

$$\left| R_{\text{MEN}}(\text{MOSM}) - \sqrt{d} \right| \leq \; 6d^{\frac{1}{4}} \,.$$

2. (Women's average rank of husbands)

$$\left| R_{\text{WOMEN}}(\text{MOSM}) - \sqrt{d} \right| \leq \; 8d^{\frac{1}{4}} \,.$$

3. (The number of unmatched men)

$$\left| \log \delta^m - \log ne^{-\sqrt{d}} \right| \leq \; 3d^{\frac{1}{4}} \,.$$

4. (The number of unmatched women)

$$\left|\log \delta^w - \log n e^{-\sqrt{d}}\right| \leq \ 2.5 d^{\frac{1}{4}} .$$

The proofs are organized as follows:

- (Section C.2.1) We first show that with high probability, the stopping time of MPDA (Algorithm 3.1), namely, τ, is bounded above as $\tau \leq n\left(\sqrt{d} + d^{\frac{1}{4}}\right)$, by utilizing the coupled extended process defined in Section C.1.5. This yields a high probability upper bound on $R_{\text{MEN}}(\text{MOSM})$ and a lower bound on the number of unmatched men δ^m and unmatched women δ^w.

- (Section C.2.2) We prove the complementary bounds on $R_{\text{MEN}}(\text{MOSM})$, δ^m, and δ^w: a lower bound on $R_{\text{MEN}}(\text{MOSM})$ and an upper bound on the number of unmatched men δ^m and unmatched women δ^w. To this end, we start by analyzing the rejection chains triggered by the last man to enter in MPDA, and deduce upper bounds on $\mathbb{E}[\delta^m]$ and $\mathbb{E}[\delta^w]$, using the fact that the order in which men enter does not matter. Using Markov's inequality, we then obtain high probability upper bounds on δ^m and δ^w, which lead to lower bounds on τ and $R_{\text{MEN}}(\text{MOSM})$.

- (Section C.2.3) We construct the concentration bounds on $R_{\text{WOMEN}}(\text{MOSM})$ based upon the concentration results on τ. In this step, we utilizes the balls-into-bins process to analyze the women's side while carefully controlling the difference between the MPDA procedure and the balls-into-bins process. This completes the proof of Theorem C.1.

C.2.1 Step 1: Upper Bound on the Total Number of Proposals τ

We prove the following two propositions.

Proposition C.1 (Upper bound on men's average rank)**.** *Consider the setting of Theorem 3.1. With probability* $1 - O(\exp(-\sqrt{n}))$*, we have the following upper bounds on the total* *number of proposals and men's average rank:*

$$\tau \leq n\left(\sqrt{d} + d^{\frac{1}{4}}\right), \qquad R_{\text{MEN}}(\text{MOSM}) \leq \sqrt{d} + 2d^{\frac{1}{4}}.$$

Proposition C.2 (Lower bound on the number of unmatched women)**.** *Consider the setting of Theorem 3.1. With probability* $1 - O(\exp(-\sqrt{n}))$*, we have the following lower bounds on the number of unmatched men* δ^m *and unmatched women* δ^w*:*

$$\delta^m \geq n \exp\left(-\sqrt{d} - 3d^{\frac{1}{4}}\right), \qquad \delta^w \geq n \exp\left(-\sqrt{d} - 2d^{\frac{1}{4}}\right).$$

Throughout the proofs we utilize the extended process defined in Section C.1.5, which enables us to analyze the state dynamics even after the termination of original DA procedure. Most of the work is in proving Proposition C.1, which is done in Sections C.2.1–C.2.1. We then deduce Proposition C.2 from Proposition C.1 in Section C.2.1. The overall proof structure is as follows:

- (Sections C.2.1 and C.2.1) We first analyze the women side using balls-into-bins process analogy: Given that a sufficient number of proposals have been made (in particular, for $t = (1+\epsilon)n\sqrt{d}$), we construct a high probability upper bound on the current number of unmatched women $\delta^w[t]$ and the probability p_t of a proposal being accepted.

- (Sections C.2.1 and C.2.1) We then analyze the men side and obtain a lower bound on the current number of unmatched men $\delta^m[t]$ at $t = (1+\epsilon)n\sqrt{d}$ by utilizing the upper bound on acceptance probability p_t. Since this lower bound exceeds the upper bound on $\delta^w[t]$ (plus k) which holds at the same t, we deduce that, whp, the algorithm has already terminated, $\tau \leq t = (1+\epsilon)n\sqrt{d}$, since we know that $\delta^m[\tau] = \delta^w[\tau] + k$. See Figure 3.3 in Section 3.4 for illustration. Consequently, an upper bound on R_{MEN} follows from the identity $R_{\text{MEN}} = \frac{\tau + \delta^m}{n+k}$, thus completing the proof of Proposition C.1.

- (Section C.2.1) Given the upper bound on τ, we obtain a lower bound on δ^w using the balls-into-bins analogy again. This leads to a lower bound on δ^m due to the identity $\delta^m = \delta^w + k$, which completes the proof of Proposition C.2.

Upper bound on number of unmatched women after a sufficient number of proposals

The following result formalizes the fact that there cannot be too many unmatched women after a sufficient number of proposals have been made.

Lemma C.7. *Consider the setting of Theorem 3.1 and the extended process defined in Section C.1.5. For any $\epsilon \in (0, \frac{1}{2})$ and $n \in \mathbb{Z}_+$, we have*

$$\mathbb{P}\left(\delta^w[(1+\epsilon)n\sqrt{d}] > ne^{-(1+\frac{\epsilon}{2})\sqrt{d}}\right) \leq \exp\left(-\frac{1}{2}nd\epsilon^2 e^{-3\sqrt{d}}\right). \tag{C.5}$$

In words, after $t = (1+\epsilon)n\sqrt{d}$ proposals have been made, at most $ne^{-(1+\frac{\epsilon}{2})\sqrt{d}}$ women remain unmatched with high probability.

Proof. It is well known that for any $t > 0$, $\delta^w[t]$ is stochastically dominated by the number of empty bins at the end of a balls-into-bins process (defined in Section C.1.3) with t balls and n bins, which we denote by $X_{t,n}$. (See, e.g., [131]; the idea is that since men's preference lists sample women without replacement, the actual process has a weakly larger probability of proposing to an unmatched woman at each step relative to picking a uniformly random woman, and hence a stochastically smaller number of unmatched women at any given t.) Therefore we have

$$\begin{aligned}
&\mathbb{P}\left(\delta^w[(1+\epsilon)n\sqrt{d}] > ne^{-(1+\frac{\epsilon}{2})\sqrt{d}}\right) \\
&\leq \mathbb{P}\left(X_{(1+\epsilon)n\sqrt{d},n} > ne^{-(1+\frac{\epsilon}{2})\sqrt{d}}\right) \\
&= \mathbb{P}\left(\frac{1}{n}X_{(1+\epsilon)n\sqrt{d},n} - \left(1-\frac{1}{n}\right)^{(1+\epsilon)n\sqrt{d}} > e^{-(1+\frac{\epsilon}{2})\sqrt{d}} - \left(1-\frac{1}{n}\right)^{(1+\epsilon)n\sqrt{d}}\right).
\end{aligned}$$

By Lemma C.1 and Lemma C.4, we further have

$$\begin{aligned}
&\mathbb{P}\left(\delta^w[(1+\epsilon)n\sqrt{d}] > ne^{-(1+\frac{\epsilon}{2})\sqrt{d}}\right) \\
&\leq \mathbb{P}\left(\frac{1}{n}X_{(1+\epsilon)n\sqrt{d},n} - \left(1-\frac{1}{n}\right)^{(1+\epsilon)n\sqrt{d}} > e^{-(1+\frac{\epsilon}{2})\sqrt{d}} - e^{-(1+\epsilon)\sqrt{d}}\right)
\end{aligned}$$

$$\leq \exp\left(-2n\left(e^{-(1+\frac{\epsilon}{2})\sqrt{d}} - e^{-(1+\epsilon)\sqrt{d}}\right)^2\right).$$

For $0 < a < b$, using the convexity of function $f(x) = e^{-x}$ we have $e^{-a} - e^{-b} \geq e^{-b}(b-a)$, and therefore for $\epsilon \in (0, \frac{1}{2})$ and any $n \in \mathbb{Z}_+$,

$$\mathbb{P}\left(\delta^w[(1+\epsilon)n\sqrt{d}] > ne^{-(1+\frac{\epsilon}{2})\sqrt{d}}\right) \leq \exp\left(-2nd\left(\epsilon - \frac{\epsilon}{2}\right)^2 e^{-2(1+\epsilon)\sqrt{d}}\right) \leq \exp\left(-\frac{1}{2}nd\epsilon^2 e^{-3\sqrt{d}}\right).$$

This concludes the proof. □

Upper bound on ex-ante acceptance probability

We define the *ex-ante acceptance probability* as

$$p_t \triangleq \frac{1}{|\mathcal{W} \setminus \mathcal{H}_t|} \sum_{j \in \mathcal{W} \setminus \mathcal{H}_t} \frac{1}{W_{j,t-1} + 1}. \tag{C.6}$$

This is the probability that the t^{th} proposal is accepted after the proposer I_t is declared but the recipient J_t is not yet revealed (recall that I_t is the identity of the man who makes the t^{th} proposal, J_t is the identity of the woman who receives it, and $\mathcal{H}_t$ is the set of women whom I_t has previously proposed to). In the following lemma, we construct a high probability upper bound on the summation in (C.6), and the subsequent lemma will use it to obtain an upper bound on $p_{(1+\frac{\epsilon}{2})n\sqrt{d}}$ for small ϵ.

Lemma C.8. *For any $\Delta > 0$ and t such that $1 \leq t \leq nd$, we have*

$$\mathbb{P}\left(\frac{1}{n}\sum_{j \in \mathcal{W}} \frac{1}{W_{j,t} + 1} \geq \frac{n}{t} + \frac{d^2}{n} + \Delta\right) \leq 2\exp\left(-\frac{n\Delta^2}{8d}\right).$$

This is also valid for the extended process (i.e., when $t \geq \tau$).

Proof. Consider a balls-into-bins process with t balls and n bins, and let $\tilde{J}_s \in \{1, \ldots, n\}$ be the index of bin into which the s^{th} ball is placed, and let $\tilde{W}_{j,t} \triangleq \sum_{s=1}^{t} \mathbb{I}(\tilde{J}_s = j)$ be the total number of balls placed in the j^{th} bin. Recall that $\tilde{J}_s$ is being sampled from $\{1, \ldots, n\}$ $(= \mathcal{W})$ uniformly at random.

We make a coupling between the MPDA procedure and the balls-into-bins process as follows: when determining the s^{th} recipient J_s, we take $J_s \leftarrow \tilde{J}_s$ if $\tilde{J}_s \notin \mathcal{H}_s$, or otherwise, sample J_s among $\mathcal{W} \setminus \mathcal{H}_s$ uniformly at random. In other words, the man I_s first picks a woman $\tilde{J}_s$ among the entire $\mathcal{W}$ uniformly at random, and then proposes to her only if he had not proposed to her yet; if he already had proposed before, he proposes to another woman randomly sampled among $\mathcal{W} \setminus \mathcal{H}_s$. It is straightforward that the evolution of the recipient process J_s under this coupling is identical to that under the usual MPDA procedure.

Define $D_t \triangleq \sum_{s=1}^{t} \mathbb{I}(J_s \neq \tilde{J}_s)$ representing the total discrepancy between the MPDA procedure and its coupled balls-into-bins process. Observe that $\mathbb{I}(\tilde{J}_s \neq J_s) = \mathbb{I}\left(\tilde{J}_s \in \mathcal{H}_s\right)$ and thus $\mathbb{P}\left(\tilde{J}_s \neq J_s \middle| \mathcal{F}_{s-1}\right) = \mathbb{P}\left(\tilde{J}_s \in \mathcal{H}_s \middle| \mathcal{F}_{s-1}\right) \leq \frac{d}{n}$ where $\mathcal{F}_{s-1}$ represents all information revealed up to time $s-1$. Let $Z_s \triangleq D_s - \frac{d}{n}s$ and observe that $(M_s)_{s\geq 0}$ is a supermartingale with $Z_0 = 0$ and $|Z_{s+1} - Z_s| \leq 1$. By Azuma's inequality, we have for any $\Delta_0 > 0$,

$$\mathbb{P}\left(D_t \geq \frac{dt}{n} + \Delta_0\right) \leq \mathbb{P}(Z_t - Z_0 \geq \Delta_0) \leq \exp\left(-\frac{\Delta_0^2}{2t}\right).$$

On the other hand, since $0 \leq \frac{1}{w+1} \leq 1$ for any $w \geq 0$, we deduce that

$$\begin{aligned}\sum_{j\in\mathcal{W}} \frac{1}{W_{j,t}+1} - \sum_{j\in\mathcal{W}} \frac{1}{\tilde{W}_{j,t}+1} &\leq \sum_{j\in\mathcal{W}:W_{j,t}<\tilde{W}_{j,t}} \left(\frac{1}{W_{j,t}+1} - \frac{1}{\tilde{W}_{j,t}+1}\right) \\ &\leq \left|\{j \in \mathcal{W} : W_{j,t} < \tilde{W}_{j,t}\}\right| \leq D_t\,,\end{aligned}$$

where the last inequality follows from the fact that in order to observe $W_{j,t} < \tilde{W}_{j,t}$ for some j, at least one mismatch $\{\tilde{J}_s \neq J_s\}$ should take place. Based on the high probability upper bound on D_t obtained above, we have for any $\Delta_1 > 0$,

$$\mathbb{P}\left(\frac{1}{n}\sum_{j\in\mathcal{W}} \frac{1}{W_{j,t}+1} - \frac{1}{n}\sum_{j\in\mathcal{W}} \frac{1}{\tilde{W}_{j,t}+1} \geq \frac{dt}{n^2} + \Delta_1\right) \leq \mathbb{P}\left(D_t \geq \frac{dt}{n} + n\Delta_1\right) \leq \exp\left(-\frac{n^2\Delta_1^2}{2t}\right). \tag{C.7}$$

We now utilize the result derived for the balls-into-bins process. From Lemma C.5, we have for any $\Delta_2 > 0$,

$$\mathbb{P}\left(\frac{1}{n}\sum_{j\in\mathcal{W}} \frac{1}{\tilde{W}_{j,t}+1} \geq \frac{n}{t} + \Delta_2\right) \leq \exp\left(-2n\Delta_2^2\right).$$

Combined with (C.7),

$$
\begin{aligned}
&\mathbb{P}\left(\frac{1}{n}\sum_{j\in\mathcal{W}}\frac{1}{W_{j,t}+1}\geq\frac{n}{t}+\Delta_2+\frac{dt}{n^2}+\Delta_1\right)\\
\leq\ &\mathbb{P}\left(\frac{1}{n}\sum_{j\in\mathcal{W}}\frac{1}{W_{j,t}+1}\geq\frac{n}{t}+\Delta_2+\frac{dt}{n^2}+\Delta_1,\frac{1}{n}\sum_{j\in\mathcal{W}}\frac{1}{\tilde{W}_{j,t}+1}<\frac{n}{t}+\Delta_2\right)\\
&+\mathbb{P}\left(\frac{1}{n}\sum_{j\in\mathcal{W}}\frac{1}{\tilde{W}_{j,t}+1}\geq\frac{n}{t}+\Delta_2\right)\\
\leq\ &\mathbb{P}\left(\frac{1}{n}\sum_{j\in\mathcal{W}}\frac{1}{W_{j,t}+1}-\frac{1}{n}\sum_{j\in\mathcal{W}}\frac{1}{\tilde{W}_{j,t}+1}\geq\frac{dt}{n^2}+\Delta_1\right)+\exp\left(-2n\Delta_2^2\right)\\
\leq\ &\exp\left(-\frac{n^2\Delta_1^2}{2t}\right)+\exp\left(-2n\Delta_2^2\right),
\end{aligned}
$$

for any $\Delta_1>0$ and $\Delta_2>0$.

We are ready to prove the claim. Given any $\Delta>0$, take $\Delta_1=\Delta_2=\Delta/2$. Then,

$$
\begin{aligned}
&\mathbb{P}\left(\frac{1}{n}\sum_{j\in\mathcal{W}}\frac{1}{W_{j,t}+1}\geq\frac{n}{t}+\frac{d^2}{n}+\Delta\right)\\
\leq\ &\mathbb{P}\left(\frac{1}{n}\sum_{j\in\mathcal{W}}\frac{1}{W_{j,t}+1}\geq\frac{n}{t}+\frac{dt}{n^2}+\Delta_1+\Delta_2\right)\\
\leq\ &\exp\left(-\frac{n^2\Delta_1^2}{2t}\right)+\exp\left(-2n\Delta_2^2\right)\\
=\ &\exp\left(-\frac{n^2\Delta^2}{8t}\right)+\exp\left(-\frac{1}{2}n\Delta^2\right)\\
\leq\ &\exp\left(-\frac{n\Delta^2}{8d}\right)+\exp\left(-\frac{1}{2}n\Delta^2\right)\\
\leq\ &2\exp\left(-\frac{n\Delta^2}{8d}\right),
\end{aligned}
$$

where we utilized the fact that $\frac{dt}{n^2}\leq\frac{d^2}{n}$ and $\frac{n}{d}\leq\frac{n^2}{t}$ under the given condition $t\leq nd$. □

Lemma C.9. *Fix any $\alpha\in(0,1)$, $\epsilon<0.2$ and sequences $(d(n))_{n\in\mathbb{N}}$, and $(\gamma(n))_{n\in\mathbb{N}}$ such that $d=d(n)=\omega(1)$ and $d=o(\log^2 n)$, and $\gamma=\gamma(n)=\Theta\left(n^{-\alpha}\right)$. Define the maximal* *ex-ante acceptance probability (for any $t \le nd$) as*

$$\overline{p}_t \triangleq \max_{\mathcal{H}\subset\mathcal{W}:|\mathcal{H}|\le d} \left\{ \frac{1}{|\mathcal{W}\setminus\mathcal{H}|} \sum_{j\in\mathcal{W}\setminus\mathcal{H}} \frac{1}{W_{j,t-1}+1} \right\}. \tag{C.8}$$

Then there exists $n_0 < \infty$ such that for all $n > n_0$, we have

$$\mathbb{P}\left(\overline{p}_{(1+\frac{\epsilon}{2})n\sqrt{d}} \ge \frac{1+\gamma}{(1+\frac{\epsilon}{2})\sqrt{d}}\right) \le 2\exp\left(-\frac{\gamma^2}{32}\frac{n}{d^2}\right).$$

This is also valid for the extended process (i.e., when $(1+\frac{\epsilon}{2})n\sqrt{d} \ge \tau$).

Proof. Let $t \triangleq (1+\frac{\epsilon}{2})n\sqrt{d}$ and $\mathcal{H}^*$ be the maximizer of (C.8). Observe that $|\mathcal{W}\setminus\mathcal{H}^*| \ge n-d$ and $\sum_{j\in\mathcal{W}\setminus\mathcal{H}^*} \frac{1}{W_{j,t-1}+1} \le \sum_{j\in\mathcal{W}} \frac{1}{W_{j,t-1}+1}$, and hence

$$\overline{p}_t = \frac{1}{|\mathcal{W}\setminus\mathcal{H}^*|} \sum_{j\in\mathcal{W}\setminus\mathcal{H}^*} \frac{1}{W_{j,t-1}+1} \le \frac{1}{n-d}\sum_{j\in\mathcal{W}} \frac{1}{W_{j,t-1}+1} \le \frac{1}{n-d}\left(1+\sum_{j\in\mathcal{W}}\frac{1}{W_{j,t}+1}\right).$$

The last inequality uses that at most one of the terms in the summation decreases from $t-1$ to t, and the decrease in that term is less than 1.

Let $r \triangleq \frac{t}{n} = (1+\frac{\epsilon}{2})\sqrt{d}$, and $\Delta \triangleq \frac{\gamma}{2\sqrt{d}}$. Under the specified asymptotic conditions, for n large enough we have

$$r\Delta = (1+\tfrac{\epsilon}{2})\sqrt{d}\cdot\frac{\gamma}{2\sqrt{d}} \le 0.6\gamma\,,\quad \frac{r}{n} = \frac{(1+\frac{\epsilon}{2})\sqrt{d}}{n} \le 0.1\gamma\,,\quad \frac{rd^2}{n} \le \frac{(1+\frac{\epsilon}{2})d^{5/2}}{n} \le 0.1\gamma\,,\quad \frac{d}{n} \le 0.1\gamma\,.$$

Consequently, since $\gamma = o(1)$, for large enough n we have

$$\begin{aligned}\frac{n}{n-d}\left(\frac{1}{r}+\frac{d^2}{n}+\Delta+\frac{1}{n}\right) &= \frac{1}{1-d/n}\cdot\frac{1}{r}\cdot\left(1+\frac{rd^2}{n}+r\Delta+\frac{r}{n}\right)\\ &\le \frac{1}{r}\cdot\frac{1}{1-0.1\gamma}\cdot(1+0.1\gamma+0.6\gamma+0.1\gamma)\\ &\le \frac{1}{r}\cdot(1+\gamma) = \frac{1+\gamma}{(1+\frac{\epsilon}{2})\sqrt{d}}\,.\end{aligned}$$

As a result,

$$\begin{aligned}\mathbb{P}\left(\overline{p}_{(1+\frac{\epsilon}{2})\tau^*} \ge \frac{1+\gamma}{(1+\frac{\epsilon}{2})\sqrt{d}}\right) &\le \mathbb{P}\left(\frac{1}{n-d}\left(1+\sum_{j\in\mathcal{W}}\frac{1}{W_{j,t}+1}\right) \ge \frac{1+\gamma}{(1+\frac{\epsilon}{2})\sqrt{d}}\right)\\ &\le \mathbb{P}\left(\frac{1}{n-d}\left(1+\sum_{j\in\mathcal{W}}\frac{1}{W_{j,t}+1}\right) \ge \frac{n}{n-d}\times\left(\frac{1}{r}+\frac{d^2}{n}+\Delta+\frac{1}{n}\right)\right)\end{aligned}$$

$$= \mathbb{P}\left(\frac{1}{n}\sum_{j\in\mathcal{W}}\frac{1}{W_{j,t}+1} \geq \frac{n}{t} + \frac{d^2}{n} + \Delta\right)$$
$$\leq 2\exp\left(-\frac{n\Delta^2}{8d}\right) = 2\exp\left(-\frac{\gamma^2}{32}\frac{n}{d^2}\right),$$

where the last inequality follows from Lemma C.8. □

Lower bound on the number of unmatched men after a sufficient number of proposals

The following result formalizes the fact that there cannot be too few unmatched men after an enough number of proposals have been made.

Lemma C.10. *Consider the setting of Theorem 3.1 and the extended process defined in Section C.1.5. For any sequence $(\epsilon(n))_{n\in\mathbb{N}}$ such that $\epsilon = \epsilon(n) < 0.2$ and $\epsilon(n) = \omega\left(\frac{1}{n^{0.49}}\right)$, there exists $n_0 < \infty$ such that for all $n > n_0$, we have*

$$\mathbb{P}\left(\delta^m[(1+\epsilon)n\sqrt{d}] \leq \frac{\epsilon}{16}ne^{-(1-\frac{\epsilon}{3})\sqrt{d}}\right) \leq \exp\left(-\sqrt{n}\right). \tag{C.9}$$

In words, after $(1+\epsilon)n\sqrt{d}$ proposals have been made, at least $\frac{\epsilon}{16}ne^{-(1-\frac{\epsilon}{3})\sqrt{d}}$ men become unmatched with high probability.

Proof. Let $\tau^* \triangleq n\sqrt{d}$. To obtain a lower bound on the number of unmatched men at time $(1+\epsilon)\tau^*$, we count the number of *d-rejection-in-a-row* events that occur during $[(1+\frac{\epsilon}{2})\tau^*, (1+\epsilon)\tau^*]$. This will provide a lower bound since whenever the rejection happens d times in a row the number of unmatched men increases at least by one.

For this purpose, we first utilize the upper bound on the ex-ante acceptance probability. By Lemma C.9 we have: given that $\gamma = \gamma(n) = \Theta\left(\frac{1}{n^\alpha}\right)$ for some $\alpha \in (0,1)$, $\epsilon = \epsilon(n) < 0.2$, and that $d = d(n) = \omega(1)$ and $d = o(\log^2 n)$, there exists $n_0 > 0$ such that for all $n > n_0$,

$$\mathbb{P}\left(\overline{p}_{(1+\frac{\epsilon}{2})\tau^*} \geq \frac{1+\gamma}{(1+\frac{\epsilon}{2})\sqrt{d}}\right) \leq 2\exp\left(-\frac{\gamma^2}{32}\frac{n}{d^2}\right). \tag{C.10}$$

Let $\hat{p} \triangleq \frac{1+\gamma}{(1+\frac{\epsilon}{2})\sqrt{d}}$ and consider the events where $\overline{p}_{(1+\frac{\epsilon}{2})\tau^*} \leq \hat{p}$ is satisifed. Since $\overline{p}_t$ is non-increasing over time on each sample path, we have $p_t \le \hat{p}$ for all $t \ge (1+\frac{\epsilon}{2})\tau^*$ on this sample path: i.e., a proposal after time $(1+\frac{\epsilon}{2})\tau^*$ is accepted with probability at most $\hat{p}$.

As an analogy, we imagine a coin tossing process with head probability $\hat{p}$ (which is an exaggeration of the actual acceptance probability, making it underestimate the occurrence of rejections and provides a valid lower bound on the actual number of d-rejection-in-a-row events), and count how many times d-tail-in-a-row takes place during $\frac{\epsilon}{2}\tau^*$ coin tosses. With $X_i \overset{\text{i.i.d.}}{\sim} \text{Geometric}(\hat{p})$ representing the number of coin tosses required to observe one head (acceptance), the total number of coin tosses required to observe one d-tail-in-a-row is given by $\sum_{i=1}^{N} \min\{X_i, d\}$ where N is the smallest i such that $X_i > d$. Note that $N \sim \text{Geometric}\left((1-\hat{p})^d\right)$. However, N is correlated with X_i's. To upper bound the random sum, observe that conditioned on N, $\{X_1, \cdots, X_{N-1}\}$ are independent *truncated* Geomtric($\hat{p}$) variables that only take value on $\{1, \cdots, d\}$, which are stochastically dominated by Geomtric($\hat{p}$) random variables. Since $\min\{X_N, d\} \le d$, the random sum of interest is stochastically dominated by $d+S$, where $S = \sum_{i=1}^{N'} X_i$, and $N' \sim \text{Geometric}\left((1-\hat{p})^d\right)$ independent of X_i's. (Note that by Wald's identity we have $\mathbb{E}[S] = \hat{p}^{-1}(1-\hat{p})^{-d}$.) Consequently, the total number of coin tosses required to observe $\frac{\epsilon}{8} n e^{-d\hat{p}}$ d-tail-in-a-row's is stochastically dominated by

$$\sum_{j=1}^{\frac{\epsilon}{8} n e^{-d\hat{p}}} (d + S_j),$$

where $S_1, S_2, \ldots$ are i.i.d. random variables with the same distribution as S defined above.

Let R denote the total number of d-tail-in-a-row events that occur during $[(1+\frac{\epsilon}{2})\tau^*, (1+\epsilon)\tau^*]$. From the above argument, we deduce that

$$\mathbb{P}\left(R \le \frac{\epsilon}{8} n e^{-d\hat{p}}\right) \le \mathbb{P}\left(\sum_{j=1}^{\frac{\epsilon}{8} n e^{-d\hat{p}}} (d+S_j) \ge \frac{\epsilon}{2}\tau^*\right) = \mathbb{P}\left(\sum_{j=1}^{\frac{\epsilon}{8} n e^{-d\hat{p}}} S_j \ge \frac{\epsilon}{2}\tau^* - \frac{\epsilon}{8} n d e^{-d\hat{p}}\right). \tag{C.11}$$

We now proceed to bound the RHS of (C.11). Note that

$$\frac{\frac{\epsilon}{2}\tau^* - \frac{\epsilon}{8}nde^{-d\hat{p}}}{\frac{\epsilon}{8}ne^{-d\hat{p}}} = \frac{4n\sqrt{d}}{ne^{-\frac{1+\gamma}{1+\frac{\epsilon}{2}}\sqrt{d}}} - d = 4\sqrt{d}e^{\frac{1+\gamma}{1+\frac{\epsilon}{2}}\sqrt{d}} - d\,.$$

Recall that $\gamma = \Theta\left(\frac{1}{n^\alpha}\right)$, $\epsilon < 0.2$, and $d = \omega(1)$, we have for large enough n, $4\sqrt{d}e^{\frac{1+\gamma}{1+\frac{\epsilon}{2}}\sqrt{d}} - d > 3.9\sqrt{d}e^{\frac{1+\gamma}{1+\frac{\epsilon}{2}}\sqrt{d}}$. Plugging $\lambda \triangleq 3.9\sqrt{d}e^{\frac{1+\gamma}{1+\frac{\epsilon}{2}}\sqrt{d}}$ into Lemma C.6, we obtain

$$\begin{aligned}
\mathbb{P}\left(\sum_{j=1}^{\frac{\epsilon}{8}ne^{-d\hat{p}}} S_j \geq \frac{\epsilon}{2}\tau^* - \frac{\epsilon}{8}nde^{-d\hat{p}}\right) &\leq \exp\left(-\frac{\frac{\epsilon}{8}ne^{-d\hat{p}}}{2\lambda^2}\left(\lambda - \mathbb{E}[S]\right)^2\right) \\
&\leq \exp\left(-\frac{\epsilon n}{16}e^{-\frac{1+\gamma}{1+\frac{\epsilon}{2}}\sqrt{d}}\left(1 - \frac{\mathbb{E}[S]}{\lambda}\right)^2\right). \qquad \text{(C.12)}
\end{aligned}$$

We also have $\hat{p} = \frac{1+\gamma}{(1+\frac{\epsilon}{2})\sqrt{d}} = o(1)$ and thus for large enough n,

$$(1-\hat{p})^{-d} \leq \left(e^{-\hat{p}-\hat{p}^2}\right)^{-d} = e^{\frac{1+\gamma}{1+\frac{\epsilon}{2}}\sqrt{d} + \left(\frac{1+\gamma}{1+\frac{\epsilon}{2}}\right)^2},$$

where we use the fact that $1 - x \geq e^{-x-x^2}$ for any $|x| \leq 0.5$. Further observe that for large enough n,

$$\frac{1+\frac{\epsilon}{2}}{1+\gamma}e^{\left(\frac{1+\gamma}{1+\frac{\epsilon}{2}}\right)^2} \leq 1.2e < 3.3\,,$$

and therefore,

$$\mathbb{E}[S] = \hat{p}^{-1}(1-\hat{p})^{-d} \leq \frac{1+\frac{\epsilon}{2}}{1+\gamma}\sqrt{d}e^{\frac{1+\gamma}{1+\frac{\epsilon}{2}}\sqrt{d} + \left(\frac{1+\gamma}{1+\frac{\epsilon}{2}}\right)^2} \leq 3.3\sqrt{d}e^{\frac{1+\gamma}{1+\frac{\epsilon}{2}}\sqrt{d}}.$$

For RHS of (C.12), we deduce that for large enough n,

$$\begin{aligned}
\exp\left(-\frac{\epsilon n}{16}e^{-\frac{1+\gamma}{1+\frac{\epsilon}{2}}\sqrt{d}}\left(1 - \frac{\mathbb{E}[S]}{\lambda}\right)^2\right) &\leq \exp\left(-\frac{\epsilon n}{16}e^{-\frac{1+\gamma}{1+\frac{\epsilon}{2}}\sqrt{d}}\left(1 - \frac{3.3}{3.9}\right)^2\right) \\
&\leq \exp\left(-\frac{\epsilon n}{800}e^{-\frac{1+\gamma}{1+\frac{\epsilon}{2}}\sqrt{d}}\right).
\end{aligned}$$

Combining all these results, for large enough n, we obtain

$$\mathbb{P}\left(R \leq \frac{\epsilon}{8}ne^{-d\hat{p}}\right) \leq \exp\left(-\frac{\epsilon n}{800}e^{-\frac{1+\gamma}{1+\frac{\epsilon}{2}}\sqrt{d}}\right).$$

As a result, we obtain a high probability lower bound on the number of unmatched men for the sample paths satisfying $\overline{p}_{(1+\frac{\epsilon}{2})\tau^*} \leq \hat{p}$:

$$\begin{aligned}
&\mathbb{P}\left(\delta^m[(1+\epsilon)\tau^*] \leq \frac{\epsilon}{8} n e^{-\frac{1+\gamma}{1+\frac{\epsilon}{2}}\sqrt{d}} \middle| \overline{p}_{(1+\frac{\epsilon}{2})\tau^*} \leq \frac{1+\gamma}{(1+\frac{\epsilon}{2})\sqrt{d}}\right) \\
&\leq \mathbb{P}\left(R \leq \frac{\epsilon}{8} n e^{-\frac{1+\gamma}{1+\frac{\epsilon}{2}}\sqrt{d}}\right) \\
&\leq \exp\left(-\frac{\epsilon n}{800} e^{-\frac{1+\gamma}{1+\frac{\epsilon}{2}}\sqrt{d}}\right) .
\end{aligned}$$

Combining with (C.10), we obtain

$$\begin{aligned}
&\mathbb{P}\left(\delta^m[(1+\epsilon)\tau^*] \leq \frac{\epsilon}{8} n e^{-\frac{1+\gamma}{1+\frac{\epsilon}{2}}\sqrt{d}}\right) \\
&\leq \mathbb{P}\left(\delta^m[(1+\epsilon)\tau^*] \leq \frac{\epsilon}{8} n e^{-\frac{1+\gamma}{1+\frac{\epsilon}{2}}\sqrt{d}} \middle| \overline{p}_{(1+\frac{\epsilon}{2})\tau^*} \leq \frac{1+\gamma}{(1+\frac{\epsilon}{2})\sqrt{d}}\right) + \mathbb{P}\left(\overline{p}_{(1+\frac{\epsilon}{2})\tau^*} \geq \frac{1+\gamma}{(1+\frac{\epsilon}{2})\sqrt{d}}\right) \\
&\leq \exp\left(-\frac{\epsilon n}{800} e^{-\frac{1+\gamma}{1+\frac{\epsilon}{2}}\sqrt{d}}\right) + 2\exp\left(-\frac{\gamma^2}{32}\frac{n}{d^2}\right) .
\end{aligned}$$

Now we take $\gamma = n^{-1/5}$. First observe that, for large enough n, since $d = o(\log^2 n)$, we have

$$2\exp\left(-\frac{\gamma^2}{32}\frac{n}{d^2}\right) = 2\exp\left(-\frac{1}{32}\frac{n^{3/5}}{d^2}\right) \leq \frac{1}{2}\exp\left(-\sqrt{n}\right) ,$$

and furthermore, since $\epsilon < 0.2$,

$$e^{-\frac{1+\gamma}{1+\frac{\epsilon}{2}}\sqrt{d}} \geq e^{-(1+\gamma)(1-\frac{\epsilon}{3})\sqrt{d}} = e^{-(1-\frac{\epsilon}{3})\gamma\sqrt{d}} \cdot e^{-(1-\frac{\epsilon}{3})\sqrt{d}} \geq \frac{1}{2} e^{-(1-\frac{\epsilon}{3})\sqrt{d}} .$$

Therefore, we obtain

$$\begin{aligned}
\mathbb{P}\left(\delta^m[(1+\epsilon)\tau^*] \leq \frac{\epsilon}{16} n e^{-(1-\frac{\epsilon}{3})\sqrt{d}}\right) &\leq \mathbb{P}\left(\delta^m[(1+\epsilon)\tau^*] \leq \frac{\epsilon}{8} n e^{-\frac{1+\gamma}{1+\frac{\epsilon}{2}}\sqrt{d}}\right) \\
&\leq \exp\left(-\frac{\epsilon n}{800} e^{-\frac{1+\gamma}{1+\frac{\epsilon}{2}}\sqrt{d}}\right) + 2\exp\left(-\frac{\gamma^2}{32}\frac{n}{d^2}\right) \\
&\leq \exp\left(-\frac{\epsilon n}{1600} e^{-(1-\frac{\epsilon}{3})\sqrt{d}}\right) + \frac{1}{2}\exp\left(-\sqrt{n}\right) \\
&\leq \exp\left(-\frac{\epsilon n}{1600} e^{-\sqrt{d}}\right) + \frac{1}{2}\exp\left(-\sqrt{n}\right) .
\end{aligned}$$

Given that $\epsilon = \omega(\frac{1}{n^{0.49}})$, we further have for large enough n,

$$\frac{\epsilon n}{1600} e^{-\sqrt{d}} \geq \frac{1}{1600} n^{0.51} e^{-\sqrt{d}} \geq \sqrt{n} + \log 2 ,$$

thus,

$$\exp\left(-\frac{\epsilon}{1600}ne^{-2\sqrt{d}}\right) \le \frac{1}{2}\exp\left(-\sqrt{n}\right) ,$$

which concludes the proof. □

Upper bound on the total number of proposals τ and men's average rank $R_{\mathbf{MEN}}$ (Proposition C.1)

With the help of the coupling between the extended process and the men-proposing DA, we are now able to prove Proposition C.1.

Proof of Proposition C.1. We make use of Lemma C.7. Denote $n\sqrt{d}$ by τ^*. Plug $\epsilon = d^{-\frac{1}{4}}$ in (C.5). For the RHS of (C.5) we have

$$\exp\left(-\frac{1}{2}nd\epsilon^2 e^{-3\sqrt{d}}\right) = \exp\left(-\frac{1}{2}n\sqrt{d}e^{-3\sqrt{d}}\right) \le \exp\left(-\frac{1}{2}ne^{-3\sqrt{d}}\right) \le \exp\left(-\sqrt{n}\right) .$$

Here the last inequality holds because $d = o(\log^2 n)$, and it follows that for any $\alpha > 0$, $e^{-3\sqrt{d}} = \omega\left(\frac{1}{n^\alpha}\right)$. Therefore,

$$\mathbb{P}\left(\delta^w[(1+d^{-\frac{1}{4}})\tau^*] > ne^{-\sqrt{d}}\right) \le \exp\left(-\sqrt{n}\right) . \tag{C.13}$$

We further utilize Lemma C.10. Plug $\epsilon = d^{-\frac{1}{4}}$ in (C.9). For the LHS of (C.9), because $\frac{1}{16}\frac{1}{x}e^{\frac{1}{3}x} \ge e^{\frac{1}{4}x}$ for large enough x, we have for large enough n,

$$\frac{\epsilon}{16}ne^{-(1-\frac{\epsilon}{3})\sqrt{d}} = \frac{1}{16}ne^{-\sqrt{d}}d^{-\frac{1}{4}}e^{\frac{1}{3}d^{\frac{1}{4}}} \ge ne^{-\sqrt{d}}e^{\frac{1}{4}d^{\frac{1}{4}}} ,$$

and hence

$$\mathbb{P}\left(\delta^m[(1+\epsilon)\tau^*] \le ne^{-\sqrt{d}}e^{\frac{1}{4}d^{\frac{1}{4}}}\right) \le \mathbb{P}\left(\delta^m[(1+\epsilon)\tau^*] \le \frac{\epsilon}{16}ne^{-(1-\frac{\epsilon}{3})\sqrt{d}}\right) \le \exp\left(-\sqrt{n}\right) . \tag{C.14}$$

Note that by assumption on the imbalance k, i.e., $|k| = O(ne^{-\sqrt{d}})$, there exists some constant C such that $|k| \le Cne^{-\sqrt{d}}$ for large enough n. Consequently, since $C+1 \le e^{\frac{1}{4}d^{\frac{1}{4}}}$ for large enough d (and hence for large enough n as $d = \omega(1)$), we have for large enough

n,

$$|k| \le Cne^{-\sqrt{d}} \le ne^{-\sqrt{d}}\left(e^{\frac{1}{4}d^{\frac{1}{4}}} - 1\right) .$$

Recall that τ is the smallest t such that

$$\delta^m[t] - \delta^w[t] = k ,$$

where the process $\delta^m[t] - \delta^w[t]$ is non-decreasing over time. Therefore, we have

$$\begin{aligned}
\mathbb{P}\left(\tau \ge (1+d^{-\frac{1}{4}})\tau^*\right) &\le \mathbb{P}\left(\delta^m[(1+d^{-\frac{1}{4}})\tau^*] - \delta^w[(1+d^{-\frac{1}{4}})\tau^*] \le k\right) \\
&= \mathbb{P}\left(\delta^m[(1+d^{-\frac{1}{4}})\tau^*] - \delta^w[(1+d^{-\frac{1}{4}})\tau^*] \le k,\ \delta^w[(1+d^{-\frac{1}{4}})\tau^*] \le ne^{-\sqrt{d}}\right) \\
&\quad + \mathbb{P}\left(\delta^m[(1+d^{-\frac{1}{4}})\tau^*] - \delta^w[(1+d^{-\frac{1}{4}})\tau^*] \le k,\ \delta^w[(1+d^{-\frac{1}{4}})\tau^*] > ne^{-\sqrt{d}}\right) \\
&\le \mathbb{P}\left(\delta^m[(1+d^{-\frac{1}{4}})\tau^*] \le ne^{-\sqrt{d}} + k\right) + \mathbb{P}\left(\delta^w[(1+d^{-\frac{1}{4}})\tau^*] > ne^{-\sqrt{d}}\right) \\
&\le \mathbb{P}\left(\delta^m[(1+d^{-\frac{1}{4}})\tau^*] \le ne^{-\sqrt{d}}e^{\frac{1}{4}d^{\frac{1}{4}}}\right) + \mathbb{P}\left(\delta^w[(1+d^{-\frac{1}{4}})\tau^*] > ne^{-\sqrt{d}}\right) \\
&\le 2\exp\left(-\sqrt{n}\right) ,
\end{aligned}$$

where we made use of (C.13) and (C.14) in the last step.

As a result, when the imbalance satisfies $|k| = O(ne^{-\sqrt{d}})$, with probability $1 - O(\exp(-\sqrt{n}))$, we have

$$\tau \le\ n\left(\sqrt{d} + d^{\frac{1}{4}}\right) .$$

By definition of $R_{\text{MEN}}(\text{MOSM})$, we have

$$R_{\text{MEN}}(\text{MOSM}) = \frac{\tau + \delta^m}{n+k} \le \frac{\tau + n}{n+k} .$$

Hence for $\tau \le\ n\left(\sqrt{d} + d^{\frac{1}{4}}\right)$, we have for large enough n,

$$R_{\text{MEN}}(\text{MOSM}) \le \frac{n}{n+k}\left(\sqrt{d} + d^{\frac{1}{4}} + 1\right) \le \left(1 + 0.5d^{-\frac{1}{4}}\right)\left(\sqrt{d} + d^{\frac{1}{4}} + 1\right) \le \sqrt{d} + 2d^{\frac{1}{4}},$$

where we utilized the fact that $\frac{n}{n+k} \le \frac{n}{n-|k|} \le \frac{1}{1-Ce^{-\sqrt{d}}} \le 1 + 2Ce^{-\sqrt{d}} \le 1 + 0.5d^{-\frac{1}{4}}$ for large enough d.

□

Lower bounds on the number of unmatched women δ^w and unmatched men δ^m (Proposition C.2)

We now derive a lower on the number of unmatched women δ^w. Similar to the proof of Lemma C.7, we again make an analogy between balls-into-bins process and DA procedure, but we now consider a variation of balls-into-bins process that exaggerates the effect of "sampling without replacement" as opposed to the original balls-into-bins process that assumes sampling with replacement. The lower bound on the number of empty bins in this process provides a lower bound on the number of unmatched women δ^w, which immediately leads to a lower bound on the number of unmatched men δ^m by the identity $\delta^m = \delta^w + k$.

Lemma C.11. *For any $t \geq d$ and $\Delta > 0$, we have*

$$\mathbb{P}\left(\frac{\delta^w[t]}{n-d} - \left(1 - \frac{1}{n-d}\right)^{t-d} \leq -\Delta\right) \leq \exp\left(-2(n-d)\Delta^2\right).$$

This is also valid for the extended process defined in Section C.1.5.

Proof. Note that the t^{th} proposal goes to a woman chosen uniformly at random after excluding the set of women $\mathcal{H}_t$ that the man has previously proposed to. Therefore,

$$\mathbb{P}\left(t^{\text{th}} \text{ proposal goes to one of unmatched women} \middle| \, \delta^w[t-1], \mathcal{H}_t\right) = \frac{\delta^w[t-1]}{n - |\mathcal{H}_t|} \leq \frac{\delta^w[t-1]}{n-d},$$

since $|\mathcal{H}_t| \leq d$. Consider a process $\underline{\delta}^w[t]$ defined as

$$\underline{\delta}^w[t] = \underline{\delta}^w[t-1] - X_t \quad \text{where} \quad X_t \sim \text{Bernoulli}\left(\min\left\{\frac{\delta^w[t-1]}{n-d}, 1\right\}\right).$$

Since the process $\underline{\delta}^w[t]$ exaggerates the likelihood of an unmatched woman receiving a proposal and hence exaggerates the likelihood of decrementing by 1 at each level, $\delta^w[t]$ stochastically dominates $\underline{\delta}^w[t]$: i.e., $\mathbb{P}\left(\delta^w[t] \leq x\right) \leq \mathbb{P}\left(\underline{\delta}^w[t] \leq x\right)$ for all $x \in \mathbb{N}$. We also observe that $\underline{\delta}^w[t]$ counts the number of empty bins in a process (we refer to it below as the original process) similar to balls-into-bins process where d bins are occupied during the first d periods, and then the regular balls-into-bins process begins with $n-d$ empty bins. Consider Lemma C.4 applied to the "modified" balls-into-bins process of putting t'

balls into $n-d$ bins, where the bins correspond to those which are not occupied by the first d balls in the original process, and t' is the total number of balls which go into these bins in the original process up to t. Clearly, $t' \le t-d$, since the first d balls do not go into these bins. We hence deduce from Lemma C.4 that

$$\begin{aligned}\mathbb{P}\left(\frac{\delta^w[t]}{n-d} - \left(1-\frac{1}{n-d}\right)^{t-d} \le -\Delta\right) &\le \mathbb{P}\left(\frac{\delta^w[t]}{n-d} - \left(1-\frac{1}{n-d}\right)^{t'} \le -\Delta\right) \\ &\le \exp\left(-2(n-d)\Delta^2\right).\end{aligned}$$

□

Lemma C.12. *Consider the setting of Theorem 3.1 and the extended process defined in Section C.1.5. Then there exists $n_0 < \infty$ such that for all $n > n_0$, we have the following lower bounds on the number of unmatched women:*

$$\mathbb{P}\left(\delta^w[(1+d^{-\frac{1}{4}})n\sqrt{d}] \le ne^{-(1+2d^{-\frac{1}{4}})\sqrt{d}}\right) \le \exp\left(-\sqrt{n}\right), \tag{C.15}$$

$$\mathbb{P}\left(\delta^w[(1-5d^{-\frac{1}{4}})n\sqrt{d}] \le ne^{-(1-2.5d^{-\frac{1}{4}})\sqrt{d}}\right) \le \exp\left(-\sqrt{n}\right). \tag{C.16}$$

Proof. Let $\tau^* \triangleq n\sqrt{d}$.

Proof of (C.15)**.** Fix $t = (1+d^{-\frac{1}{4}})\tau^* = (1+d^{-\frac{1}{4}})n\sqrt{d}$. For large enough n, we have

$$\frac{d}{n} \le \frac{d}{e^{\sqrt{d}}} \le 0.1d^{-\frac{1}{4}}, \quad \frac{t-d}{n-d} \le \frac{t}{n}\cdot\frac{1}{1-d/n} \le \sqrt{d}\cdot\frac{1+d^{-\frac{1}{4}}}{1-0.1d^{-\frac{1}{4}}} \le \sqrt{d}(1+1.2d^{-\frac{1}{4}}).$$

Consequently, with $\Delta \triangleq e^{-(1+2d^{-\frac{1}{4}})\sqrt{d}}$, for large enough n we have

$$\begin{aligned}\frac{n-d}{n}\left[\left(1-\frac{1}{n-d}\right)^{t-d} - \Delta\right] &\ge \left(1-\frac{d}{n}\right)\cdot\left[\exp\left(-\frac{t-d}{n-d} - \frac{t-d}{(n-d)^2}\right) - e^{-(1+2d^{-\frac{1}{4}})\sqrt{d}}\right] \\ &\ge \frac{1}{2}\cdot\left[\exp\left(-\sqrt{d}(1+1.2d^{-\frac{1}{4}})\cdot(1+1/(n-d))\right) - e^{-(1+2d^{-\frac{1}{4}})\sqrt{d}}\right] \\ &\ge \frac{1}{2}\cdot\left[\exp\left(-\sqrt{d}(1+1.4d^{-\frac{1}{4}})\right) - \exp\left(-\sqrt{d}(1+2d^{-\frac{1}{4}})\right)\right] \\ &= e^{-\sqrt{d}} \times \frac{1}{2}\cdot\left(\exp\left(-1.4d^{\frac{1}{4}}\right) - \exp\left(-2d^{\frac{1}{4}}\right)\right) \\ &\ge e^{-\sqrt{d}} \times \frac{1}{2}\cdot e^{-2d^{\frac{1}{4}}}\left(2.0d^{\frac{1}{4}} - 1.4d^{\frac{1}{4}}\right) \\ &= e^{-\sqrt{d}} \times e^{-2d^{\frac{1}{4}}} \times 0.3d^{\frac{1}{4}} \ge e^{-(1+2d^{-\frac{1}{4}})\sqrt{d}}.\end{aligned}$$

In the second last inequality, we utilize the fact that $e^{-a} - e^{-b} \ge e^{-b}(b-a)$ for any

$0 < a < b$. Therefore, by Lemma C.11,

$$\begin{aligned}
\mathbb{P}\left(\delta^w[t] \le n e^{-(1+2d^{-\frac{1}{4}})\sqrt{d}}\right) &= \mathbb{P}\left(\frac{\delta^w[t]}{n} \le e^{-(1+2d^{-\frac{1}{4}})\sqrt{d}}\right)\\
&\le \mathbb{P}\left(\frac{\delta^w[t]}{n} \le \frac{n-d}{n}\left(\left(1-\frac{1}{n-d}\right)^{t-d} - \Delta\right)\right)\\
&\le \mathbb{P}\left(\frac{\delta^w[t]}{n-d} - \left(1-\frac{1}{n-d}\right)^{t-d} \le -\Delta\right)\\
&\le \exp\left(-2(n-d)\Delta^2\right) = \exp\left(-2(n-d)e^{-2(1+d^{-\frac{1}{4}})\sqrt{d}}\right).
\end{aligned}$$

The claim follows from the fact that $2(n-d)e^{-2(1+d^{-\frac{1}{4}})\sqrt{d}} \ge \sqrt{n}$ for large enough n.

Proof of (C.16)**.** Fix $t = (1-5d^{-\frac{1}{4}})\tau^* = (1-5d^{-\frac{1}{4}})n\sqrt{d}$. For large enough n, we have

$$\frac{d}{n} \le \frac{d}{e^{\sqrt{d}}} \le 0.1 d^{-\frac{1}{4}}, \quad \frac{t-d}{n-d} \le \frac{t}{n}\cdot\frac{1}{1-d/n} \le \sqrt{d}\cdot\frac{1-5d^{-\frac{1}{4}}}{1-0.1d^{-\frac{1}{4}}} \le \sqrt{d}(1-4.8d^{-\frac{1}{4}}),$$

Consequently, with $\Delta \triangleq e^{-(1-2.5d^{-\frac{1}{4}})\sqrt{d}}$,

$$\begin{aligned}
\frac{n-d}{n}\left[\left(1-\frac{1}{n-d}\right)^{t-d} - \Delta\right] &\ge \left(1-\frac{d}{n}\right)\cdot\left[\exp\left(-\frac{t-d}{n-d} - \frac{t-d}{(n-d)^2}\right) - e^{-(1-2.5d^{-\frac{1}{4}})\sqrt{d}}\right]\\
&\ge \frac{1}{2}\cdot\left[\exp\left(-\sqrt{d}(1-4.8d^{-\frac{1}{4}})\cdot(1+1/(n-d))\right) - e^{-(1-2.5d^{-\frac{1}{4}})\sqrt{d}}\right]\\
&\ge \frac{1}{2}\cdot\left[\exp\left(-\sqrt{d}(1-4.6d^{-\frac{1}{4}})\right) - \exp\left(-\sqrt{d}(1-2.5d^{-\frac{1}{4}})\right)\right]\\
&= e^{-\sqrt{d}} \times \frac{1}{2}\cdot\left(\exp\left(4.6d^{\frac{1}{4}}\right) - \exp\left(2.5d^{\frac{1}{4}}\right)\right)\\
&\overset{(a)}{\ge} e^{-\sqrt{d}} \times \frac{1}{2}\cdot e^{2.5d^{\frac{1}{4}}}\left(4.6d^{\frac{1}{4}} - 2.5d^{\frac{1}{4}}\right)\\
&\ge e^{-\sqrt{d}} \times e^{2.5d^{\frac{1}{4}}} \times d^{\frac{1}{4}} \ge e^{-(1-2.5d^{-\frac{1}{4}})\sqrt{d}},
\end{aligned}$$

for large enough n. Here (a) follows from the fact that $f(x) = e^x$ is convex hence $f(x_2) - f(x_1) \ge f'(x_1)(x_2 - x_1)$ for $x_2 > x_1$. Therefore, by Lemma C.11,

$$\begin{aligned}
\mathbb{P}\left(\delta^w[t] \le n e^{-(1-2.5d^{-\frac{1}{4}})\sqrt{d}}\right) &= \mathbb{P}\left(\frac{\delta^w[t]}{n} \le e^{-(1-2.5d^{-\frac{1}{4}})\sqrt{d}}\right)\\
&\le \mathbb{P}\left(\frac{\delta^w[t]}{n} \le \frac{n-d}{n}\left(\left(1-\frac{1}{n-d}\right)^{t-d} - \Delta\right)\right)
\end{aligned}$$

$$\leq \mathbb{P}\left(\frac{\delta^w[t]}{n-d} - \left(1 - \frac{1}{n-d}\right)^{t-d} \leq -\Delta\right)$$
$$\leq \exp\left(-2(n-d)\Delta^2\right) = \exp\left(-2(n-d)e^{-2(1-2.5d^{-\frac{1}{4}})\sqrt{d}}\right).$$

The claim follows from the fact that $2(n-d)e^{-2(1-2.5d^{-\frac{1}{4}})\sqrt{d}} \geq \sqrt{n}$ for large enough n. □

We are now able to prove Proposition C.2.

Proof of Proposition C.2. By Proposition C.1 and the monotonicity of $\delta^w[t]$, we have for large enough n,

$$\mathbb{P}\left(\delta^w[\tau] \leq e^{-2d^{\frac{1}{4}}} n e^{-\sqrt{d}}\right) \leq \mathbb{P}\left(\delta^w[\tau] \leq e^{-2d^{\frac{1}{4}}} n e^{-\sqrt{d}}, \tau \leq (1+d^{-\frac{1}{4}})\tau^*\right) + \mathbb{P}\left(\tau \geq (1+d^{-\frac{1}{4}})\tau^*\right)$$
$$\leq \mathbb{P}\left(\delta^w[(1+d^{-\frac{1}{4}})\tau^*] \leq e^{-2d^{\frac{1}{4}}} n e^{-\sqrt{d}}\right) + \exp\left(-\sqrt{n}\right). \quad \text{(C.17)}$$

Moreover, by Lemma C.12, we have for large enough n,

$$\mathbb{P}\left(\delta^w[(1+d^{-\frac{1}{4}})\tau^*] \leq e^{-2d^{\frac{1}{4}}} n e^{-\sqrt{d}}\right) \leq \exp\left(-\sqrt{n}\right).$$

From (C.17), we conclude that with probability $1 - 2\exp(-\sqrt{n})$,

$$\delta^w \geq n e^{-\sqrt{d}-2d^{\frac{1}{4}}}.$$

Since $|\delta^m - \delta^w| = |k| \leq O(ne^{-\sqrt{d}})$, it follows that with probability $1 - 2\exp(-\sqrt{n})$,

$$\delta^m \geq n e^{-\sqrt{d}-3d^{\frac{1}{4}}}.$$

□

C.2.2 Step 2: Lower Bound on the Total Number of Proposals τ

In this section, we prove the following two propositions.

Proposition C.3. *Consider the setting in Theorem 3.1. With probability* $1-O\left(\exp\left(-d^{\frac{1}{4}}\right)\right)$ *,we have the following upper bounds on the number of unmatched men* δ^m *and unmatched women* δ^w*:*

$$\delta^m \leq n\exp\left(-\sqrt{d} + 2.5d^{\frac{1}{4}}\right), \qquad \delta^w \leq n\exp\left(-\sqrt{d} + 2.5d^{\frac{1}{4}}\right).$$

Proposition C.4. *Consider the setting of Theorem 3.1. With probability* $1-O\left(\exp\left(-d^{\frac{1}{4}}\right)\right)$, *we have the following lower bound on the total number of proposals and men's average rank under the men-optimal stable matching:*

$$\tau \geq \ n\left(\sqrt{d}-5d^{\frac{1}{4}}\right)\,, \qquad R_{\text{MEN}}(\text{MOSM}) \geq \ \sqrt{d}-6d^{\frac{1}{4}}\,.$$

The proofs of Proposition C.3 and C.4 have the following structure:

- (Sections C.2.2 and C.2.2) Proof of Proposition C.3: We first derive an upper bound on the expected number of unmatched men $\mathbb{E}[\delta^m]$ in Lemma C.13, utilizing the fact that the probability of the last proposing man being rejected cannot be too large given that the total number of proposals τ is limited by its upper bound (Proposition C.1). We immediately deduce an upper bound $\mathbb{E}[\delta^w]$ by using the identity $\delta^m = \delta^w + k$. The high probability upper bounds on δ^m and δ^w follow by applying Markov's inequality.

- (Section C.2.2) Proof of Proposition C.4: We obtain a lower bound on the total number of proposals τ by showing that the current number of unmatched women $\delta^w[t]$ does not decay fast enough (again argued with a balls-into-bins analogy) and hence it will violate the upper bound on $\delta^w[\tau]$ ($= \delta^w$) derived in Proposition C.3 if τ is too small. The lower bound on τ immediately translates into the lower bound on $R_{\text{MEN}}(\text{MOSM})$ due to the identity $R_{\text{MEN}}(\text{MOSM}) = \frac{\tau+\delta^m}{n+k}$.

Upper bound on the expected number of unmatched women $\mathbb{E}[\delta^w]$

Using a careful analysis of the rejection chains triggered by the last proposing man's proposal, we are able to derive an upper bound on the expected number of unmatched women.

Lemma C.13. *Consider the setting of Theorem 3.1. There exists* $n_0 < \infty$ *such that for all* $n > n_0$, *we have the following upper bounds on the expected number of unmatched men* *and women under stable matching*

$$\mathbb{E}[\delta^m] \le n \exp(-\sqrt{d} + 1.4d^{1/4}), \quad \mathbb{E}[\delta^w] \le n \exp(-\sqrt{d} + 1.5d^{1/4}). \tag{C.18}$$

Proof. We will track the progress of the man proposing DA algorithm making use of the principle of deferred decisions, and further make use of a particular sequence of proposals: we will specify beforehand an arbitrary man i (before any information whatsoever is revealed), and then run DA to convergence on the other men, before man i makes a single proposal. We will show that the probability that the man i remains unmatched is bounded as

$$\mathbb{P}(\mu(i) = i) \le \exp(-\sqrt{d} + 1.4d^{1/4}) \tag{C.19}$$

for large enough n. This will imply that, by symmetry across men, the expected number of unmatched men under stable matching will be bounded above as

$$\mathbb{E}[\delta^m] \le (n + k) \exp(-\sqrt{d} + 1.4d^{1/4}).$$

Finally the number of unmatched women at the end is exactly $\delta^w = \delta^m - k$, and so

$$\mathbb{E}[\delta^w] = \mathbb{E}[\delta^m] - k \le (n + k) \exp(-\sqrt{d} + 1.4d^{1/4}) - k \le n \exp(-\sqrt{d} + 1.5d^{1/4})$$

for large enough n as required, using $k = O(ne^{-\sqrt{d}})$. The rest of proof is devoted to establishing (C.19).

Using Proposition C.1, we have that with probability $1 - O(\exp(-\sqrt{n}))$, at the end of DA, τ is bounded above as

$$\tau \le n\big(\sqrt{d} + d^{\frac{1}{4}}\big), \tag{C.20}$$

and using Proposition C.2, we have that with probability $1 - O(\exp(-\sqrt{n}))$,

$$\delta^w \ge ne^{-\sqrt{d} - 2d^{\frac{1}{4}}}, \tag{C.21}$$

at the end of DA. Note that if (C.20) holds at the end of DA, then the RHS of (C.20) is an upper bound on t throughout the run of DA. Similarly, since the number of unmatched woman $\delta^w[t]$ is monotone non-increasing in t, if (C.21) holds at the end of DA, then the RHS of (C.21) is a lower bound on $\delta^w[t]$ throughout the run of DA. If, at any stage during the run of DA either (C.20) (with t instead of τ) or (C.21) (with $\delta^w[t]$ instead of δ^w) is violated, declare a "failure" event $\mathcal{E} \equiv \mathcal{E}_\tau$. By union bound, we know that $\mathbb{P}(\mathcal{E}) = O(\exp(-\sqrt{n}))$. For $t \leq \tau$, let $\mathcal{E}_t$ denote the event that no failure has occurred during the first t proposals of DA. We will prove (C.19) by showing an upper bound on the likelihood that man i remains unmatched for sample paths where no failure occurs, and assuming the worst (i.e., that i certainly remains unmatched) in the rare cases where there is a failure.

Run DA to convergence on men besides i. Now consider proposals by i. At each such proposal, the recipient woman is drawn uniformly at random from among at least $n - d + 1$ "candidate" women (the ones to whom i has not yet proposed). Assuming $\mathcal{E}_t^c$, we know that

$$t \leq n\big(\sqrt{d} + d^{\frac{1}{4}}\big)\,, \tag{C.22}$$

and hence the total number of proposals received by candidate women is at most $n\big(\sqrt{d}+d^{\frac{1}{4}}\big)$, and hence the average number of proposals received by candidate women is at most $n(\sqrt{d} + d^{\frac{1}{4}})/(n - d + 1) \leq \sqrt{d}(1 + d^{-1/4} + \log^2 n/n) \leq \sqrt{d}(1 + 1.1d^{-\frac{1}{4}}) \leq \sqrt{d} + 1.1d^{\frac{1}{4}}$ for large enough n, using $d = o(\log^2 n)$. If the proposal goes to woman j, the probability of it being accepted is $\frac{1}{w_{j,t}+1}$. Averaging over the candidate women and using Jensen's inequality for the function $f(x) = \frac{1}{x+1}$, the probability of the proposal being accepted is at least $\frac{1}{\sqrt{d}+1.1d^{1/4}+1} \geq \frac{1}{\sqrt{d}+1.2d^{1/4}}$. If the proposal is accepted, say by woman j, this triggers a rejection chain. We show that it is very unlikely that this rejection chain will cause an additional proposal to woman j (which will imply that it is very unlikely that the rejection chain will cause i himself to be rejected): For every additional proposal in the rejection chain, the likelihood that it goes to an unmatched woman far exceeds the likelihood that it goes to woman j: if the current time is t' and $\mathcal{E}_{t'}^c$ holds, then, since all $\delta^w[t']$ unmatched women are certainly candidate recipients of the next proposal, the likelihood of the proposal being to an unmatched woman is at least $\delta^w[t'] \geq ne^{-\sqrt{d}-2d^{\frac{1}{4}}} \geq \sqrt{n}$ times the likelihood of it being to woman j for n large enough, using $d = o(\log^2 n)$. Now if the proposal is to an unmatched woman, this causes the rejection chain to terminate, hence the expected number of proposals to an unmatched woman in the rejection chain is at most 1. We immediately deduce that if a failure does not occur prior to termination of the chain, the expected number of proposals to woman j in the rejection chain is at most $\frac{1}{\sqrt{n}}$. It follows that

$$
\begin{aligned}
&\mathbb{P}(i \text{ is displaced from } j \text{ by the rejection chain triggered when } j \text{ accepts his proposal}) \\
&\leq \mathbb{P}(j \text{ receives a proposal in the rejection chain triggered}) \\
&\leq \mathbb{E}[\text{Number of proposals received by } j \text{ in the rejection chain triggered}] \\
&\leq \frac{1}{\sqrt{n}},
\end{aligned}
$$

for n large enough. Overall, the probability of the proposal by i being "successful" in that it is both (a) accepted, and then (b) man i is not pushed out by the rejection chain, is at least

$$\frac{1}{\sqrt{d} + 1.2d^{1/4}}\left(1 - \frac{1}{\sqrt{n}}\right) \leq \frac{1}{\sqrt{d} + 1.3d^{1/4}},$$

for large enough n. Hence the probability of an unsuccessful proposal (if there is no failure) is at most

$$1 - \frac{1}{\sqrt{d} + 1.3d^{1/4}} \leq \exp\left\{-\frac{1}{\sqrt{d} + 1.3d^{1/4}}\right\},$$

and so the probability of all d proposals being unsuccessful (if there is no failure) is at most

$$\exp\left\{-\frac{d}{\sqrt{d} + 1.3d^{1/4}}\right\} \leq \exp\left\{-\sqrt{d} + 1.3d^{1/4}\right\}.$$

Formally, what we have obtained is an upper bound on the quantity $\mathbb{E}\big[\mathbb{I}(\mu(i) = i)\mathbb{I}(\mathcal{E}^c)\big]$, namely,

$$\mathbb{E}\big[\mathbb{I}(\mu(i) = i)\mathbb{I}(\mathcal{E}^c)\big] \leq \exp\left\{-\sqrt{d} + 1.3d^{1/4}\right\}.$$

Since the probability of failure is bounded as $\mathbb{P}(\mathcal{E}) \le O(\exp(-\sqrt{n}))$, the overall probability that of man i remaining unmatched is bounded above as

$$\begin{aligned}\mathbb{P}(\mu(i) = i) &\le \mathbb{E}\big[\mathbb{I}(\mu(i) = i)\mathbb{I}(\mathcal{E}^c)\big] + \mathbb{P}(\mathcal{E}) \\ &\le \exp\left\{-\sqrt{d} + 1.3d^{1/4}\right\} + O(\exp(-\sqrt{n})) \le \exp\left\{-\sqrt{d} + 1.4d^{1/4}\right\}\end{aligned}$$

for large enough n, i.e., the bound (C.19) which we set out to show. □

Upper bound on the number of unmatched men δ^m and unmatched women δ^w (Proposition C.3)

Proof. Proof of Proposition C.3. Recall the results in Lemma C.13:

$$\mathbb{E}[\delta^m] \le n\exp(-\sqrt{d} + 1.4d^{1/4}), \quad \mathbb{E}[\delta^w] \le n\exp(-\sqrt{d} + 1.5d^{1/4}). \tag{C.23}$$

We use Markov's inequality for each δ^m and δ^w:

$$\begin{aligned}\mathbb{P}\left(\delta^m > n\exp(-\sqrt{d} + 2.4d^{1/4})\right) &\le \frac{\mathbb{E}[\delta^m]}{n\exp(-\sqrt{d} + 2.4d^{1/4})} \le \exp(-d^{1/4}), \\ \mathbb{P}\left(\delta^w > n\exp(-\sqrt{d} + 2.5d^{1/4})\right) &\le \frac{\mathbb{E}[\delta^w]}{n\exp(-\sqrt{d} + 2.5d^{1/4})} \le \exp(-d^{1/4}).\end{aligned}$$

□

Lower bound on the number of total proposals τ (Proposition C.4)

Proof. Proof of Proposition C.4. Consider the extended process defined in Appendix C.1.5, and let $\delta^w[t]$ be the number of unmatched woman at time t of the extended process. Let τ be the time when the men-optimal stable matching is found, i.e., $\delta^w = \delta^w[\tau]$. Let $\epsilon \triangleq d^{-1/4}$. We have

$$\begin{aligned}\mathbb{P}\left(\tau < (1-5\epsilon)n\sqrt{d}\right) &\le \mathbb{P}\left(\tau < (1-5\epsilon)n\sqrt{d},\ \delta^w[\tau] < ne^{-(1-2.5\epsilon)\sqrt{d}}\right) + \mathbb{P}\left(\delta^w[\tau] \ge ne^{-(1-2.5\epsilon)\sqrt{d}}\right) \\ &\le \mathbb{P}\left(\delta^w[(1-5\epsilon)n\sqrt{d}] < ne^{-(1-2.5\epsilon)\sqrt{d}}\right) + \mathbb{P}\left(\delta^w[\tau] \ge ne^{-(1-2.5\epsilon)\sqrt{d}}\right).\end{aligned} \tag{C.24}$$

Here the last inequality holds because $\delta^w[t]$ is non-increasing over t on each sample path. It follows from Proposition C.3 that the second term on the RHS of (C.24) is $O(e^{-d^{1/4}})$.

It remains to bound the first term on the RHS of (C.24). By Lemma C.12, we have

$$\mathbb{P}\left(\delta^w[(1-5\epsilon)n\sqrt{d}] < ne^{-(1-2.5\epsilon)\sqrt{d}}\right) \leq \exp(-\sqrt{n}),$$

for large enough n. By plugging this in the RHS of (C.24), we obtain

$$\mathbb{P}\left(\tau < (1-5\epsilon)n\sqrt{d}\right) = O\left(\exp(-d^{\frac{1}{4}})\right). \tag{C.25}$$

Note that by the definition of $R_{\text{MEN}}(\text{MOSM})$, we have

$$R_{\text{MEN}}(\text{MOSM}) \geq \frac{\tau}{n+k}.$$

Since $|k| = O(ne^{-\sqrt{d}})$, using an argument similar to the one at the end of the proof of Proposition C.1, we can deduce from (C.25) that

$$\mathbb{P}\left(R_{\text{MEN}}(\text{MOSM}) < (1-6\epsilon)\sqrt{d}\right) = O\left(\exp\left(-d^{\frac{1}{4}}\right)\right).$$

This concludes the proof. □

C.2.3 Step 3: Upper and Lower Bounds on Women's Average Rank R_{WOMEN}

In this section, we prove the following two propositions.

Proposition C.5 (Lower bound on women's average rank)**.** *Consider the setting of Theorem 3.1. With probability* $1-\frac{3}{n}$*, we have the following lower bound on women's average rank:*

$$R_{\text{WOMEN}}(\text{MOSM}) \geq \sqrt{d} - 3d^{\frac{1}{4}},$$

Proposition C.6 (Upper bound on women's average)**.** *Consider the setting of Theorem 3.1. With probability* $1-O(\exp(-d^{\frac{1}{4}}))$*, we have the following upper bound on the women's average rank:*

$$R_{\text{WOMEN}}(\text{MOSM}) \leq \sqrt{d} + 8d^{\frac{1}{4}}.$$

In order to characterize the women side, we introduce a different extended process which we call the *continue-proposing process* that is slightly different from one introduced in Section C.1.5. Until the MOSM is found (i.e., $t \leq \tau$), the continue-proposing process is identical to the original DA procedure. After the MOSM is found (i.e., $t > \tau$), the proposing man I_t is chosen arbitrarily among the men who have not yet exhausted their preference list (i.e., $\{i \in \mathcal{M} : M_{i,t-1} < d\}$), and we let him propose to his next candidate. We do not care about the matching nor the acceptance/rejection after τ, since we only keep track of the number of proposals that each man has made, $M_{i,t}$, and each woman has received, $W_{j,t}$. The continue-proposing process terminates at time $t = (n+k)d$, when all men exhaust their preference lists.

To analyze the concentration of R_{WOMEN}, we first construct upper and lower bounds on its conditional expectation. More formally, we define

$$\bar{R}[t] \triangleq \frac{1}{n} \sum_{j \in \mathcal{W}} \frac{W_{j,(n+k)d} - W_{j,t}}{W_{j,t} + 1}, \tag{C.26}$$

where $W_{j,(n+k)d}$ represents the degree of woman j in a random matching market so that $W_{j,(n+k)d} - W_{j,t}$ represents the number of remaining proposals that woman j will receive after time t. In Lemma C.15, we prove that $\bar{R}[\tau]$ is concentrated around $\sqrt{d}$ given $\tau \approx n\sqrt{d}$. In Lemma C.16, we show that $\bar{R}[\tau]$ (plus 1) is indeed the conditional expectation of R_{WOMEN} given $W_{j,\tau}$'s and $W_{j,(n+k)d}$'s, and further characterize the conditional distribution of R_{WOMEN} given $\bar{R}[\tau]$, which leads to the concentration bounds on R_{WOMEN}. Within the proofs, we also utilize the fact that $\bar{R}[t]$ is decreasing over time on each sample path.

Concentration of expected women's average rank $\bar{R}_t$

We first state a preliminary lemma that will be used to show the concentration of $\bar{R}_t$.

Lemma C.14. *Fix any t and T such that $t < T$ and positive numbers $c_1, \ldots, c_n$ such that $c_j \in [0, 1]$ for all j, and define*

$$Y_{t,T} \triangleq \sum_{j \in \mathcal{W}} c_j (W_{j,T} - W_{j,t}).$$

With $S \triangleq \sum_{j=1}^{n} c_j$*, we have*

$$\mathbb{P}\left(Y_{t,T} \geq (1+\epsilon)\frac{(T-t)S}{n-d} \middle| \vec{W}_t \right) \leq \exp\left(-\frac{1}{4}\epsilon^2 \times \frac{(T-t)S}{n} \right) \tag{C.27}$$

$$\mathbb{P}\left(Y_{t,T} \leq (1-\epsilon)\frac{(T-t)(S-d)}{n-d} \middle| \vec{W}_t \right) \leq \exp\left(-\frac{1}{4}\epsilon^2 \times \frac{(T-t)(S-d)}{n-d} \right) \tag{C.28}$$

for any $\epsilon \in [0,1]$.

Proof. Throughout this proof, we assume that $W_{1,t}, \dots, W_{n,t}$ are revealed, i.e. we consider the conditional probabilities/expectations given $W_{1,t}, \dots, W_{n,t}$. In addition, we assume that $c_1 \leq c_2 \leq \dots \leq c_n$ without loss of generality.

Proof of (C.27)**:** We first establish an upper bound using a coupling argument. Recall that $W_{j,s}$ counts the number of proposals that a woman j had received up to time s, which is governed by the recipient process J_s. We construct a coupled process $\left(\overline{W}_{j,s}\right)_{j\in\mathcal{W}, s\geq t}$ that counts based on $\overline{J}_s$ as follows:

(i) Initialize $\overline{W}_{j,t} \leftarrow W_{j,t}$ for all j.

(ii) At each time $s = t+1, t+2, \dots, T$, after the recipient J_s is revealed (which is uniformly sampled among $\mathcal{W} \setminus \mathcal{H}_s$), determine $\overline{J}_s \in \{d+1, \dots, n\}$:

- If $J_s \in \{d+1, \dots, n\}$, set $\overline{J}_s \leftarrow J_s$.
- If $J_s \in \{1, \dots, d\}$, sample $\overline{J}_s$ according to the probability distribution $p_s(\cdot)$ defined as (the motivation for this definition is provided below)

$$p_s(j) = \begin{cases} 0 & \text{if } j \in \{1,\dots,d\}, \\ \frac{1}{n-d} / \frac{|\{1,\dots,d\}\setminus\mathcal{H}_s|}{|\mathcal{W}\setminus\mathcal{H}_s|} & \text{if } j \in \{d+1,\dots,n\} \cap \mathcal{H}_s, \\ \left(\frac{1}{n-d} - \frac{1}{|\mathcal{W}\setminus\mathcal{H}_s|}\right) / \frac{|\{1,\dots,d\}\setminus\mathcal{H}_s|}{|\mathcal{W}\setminus\mathcal{H}_s|} & \text{if } j \in \{d+1,\dots,n\} \setminus \mathcal{H}_s. \end{cases}$$

(iii) Increase the counter of $\overline{J}_s$ instead of J_s: i.e., $\overline{W}_{j,s} \leftarrow \overline{W}_{j,s-1} + \mathbb{I}\{\overline{J}_s = j\}$ for all j.

In words, whenever a proposal goes to one of d women who have smallest c_j values (i.e., when $J_s \in \{1, \dots, d\}$), we randomly pick one among the other $n-d$ women (i.e.,

$\overline{J}_s \in \{d+1,\ldots,n\}$) and increase that woman's counter $\overline{W}_{\overline{J}_s}$. Otherwise (i.e., when $J_s \in \{d+1,\ldots,n\}$), we count the proposal as in the original process. In any case, we have $c_{\overline{J}_s} \geq c_{J_s}$.

Note that we do not alter the proposal mechanism in this coupled process, but just count the proposals in a different way. Therefore, we have

$$\sum_{j\in\mathcal{W}} c_j(W_{j,T} - W_{j,t}) \leq \sum_{j\in\mathcal{W}} c_j(\overline{W}_{j,T} - \overline{W}_{j,t}), \tag{C.29}$$

Also note that the (re-)sampling distribution $p_s(\cdot)$ was constructed in a way that $\overline{J}_s$ is chosen uniformly at random among $\{d+1,\ldots,n\}$, unconditioned on J_s, independently of $\mathcal{H}_s$. More formally, we have for any $j \in \{d+1,\ldots,n\} \setminus \mathcal{H}_s$,

$$\begin{aligned}\mathbb{P}(\overline{J}_s = j|\mathcal{H}_s) &= \mathbb{P}(J_s = j|\mathcal{H}_s) + \mathbb{P}(\overline{J}_s = j|\mathcal{H}_s, J_s \in \{1,\ldots,d\}) \cdot \mathbb{P}(J_s \in \{1,\ldots,d\}|\mathcal{H}_s) \\ &= \frac{1}{|\mathcal{W}\setminus\mathcal{H}_s|} + \left(\frac{1}{n-d} - \frac{1}{|\mathcal{W}\setminus\mathcal{H}_s|}\right) = \frac{1}{n-d}.\end{aligned}$$

Similarly it can be verified that $\mathbb{P}(\overline{J}_s = j|\mathcal{H}_s) = \frac{1}{n-d}$ also for any $j \in \{d+1,\ldots,n\} \cap \mathcal{H}_s$. The fact that $|\mathcal{H}_s| < d$ guarantees that $p_s(\cdot)$ is a well-defined probability mass function. Therefore,

$$\sum_{j\in\mathcal{W}} c_j(\overline{W}_{j,T} - \overline{W}_{j,t}) \stackrel{\mathrm{d}}{=} \sum_{j=d+1}^{n} c_j X_j,$$

where $X_j \sim \text{Binomial}\left(T-t, \frac{1}{n-d}\right)$ for $j \in \{d+1,\ldots,n\}$. Although X_j's are not independent, they are negatively associated as in the balls-into-bins process (see Section C.1.3). For any $\lambda \in \mathbb{R}$, $\exp(\lambda c_j X_j)$'s are also NA due to Lemma C.3–((iii)), and therefore,

$$\begin{aligned}\mathbb{E}\left[\exp\left(\lambda \sum_{j=d+1}^{n} c_j X_j\right)\right] &\leq \prod_{j=d+1}^{n} \mathbb{E}\left[e^{\lambda c_j X_j}\right] \\ &= \prod_{j=d+1}^{n} \left(1 - \frac{1}{n-d} + \frac{1}{n-d} e^{\lambda c_j}\right)^{T-t} \\ &\leq \prod_{j=d+1}^{n} \exp\left(-\frac{1}{n-d} + \frac{1}{n-d} e^{\lambda c_j}\right)^{T-t}\end{aligned}$$

$$= \exp\left\{(T-t)\left(-1+\frac{1}{n-d}\sum_{j=d+1}^{n} e^{\lambda c_j}\right)\right\}.$$

Since $c_j \in [0,1]$ and $e^x \le 1 + x + x^2$ for any $x \in (-\infty, 1]$, we have for any $\lambda \in [0,1]$,

$$-1+\frac{1}{n-d}\sum_{j=d+1}^{n} e^{\lambda c_j} \le -1+\frac{1}{n-d}\sum_{j=d+1}^{n}(1+\lambda c_j+\lambda^2 c_j^2) \le \frac{\lambda+\lambda^2}{n-d}\sum_{j=d+1}^{n} c_j.$$

By Markov's inequality, for any $\lambda \in [0,1]$,

$$\begin{aligned}
&\mathbb{P}\left(\sum_{j=d+1}^{n} c_j X_j \ge (1+\epsilon)\frac{T-t}{n-d}\sum_{j=d+1}^{n} c_j\right)\\
\le&\frac{\mathbb{E}\left[\exp\left(\lambda\sum_{j=d+1}^{n} c_j X_j\right)\right]}{\exp\left(\lambda(1+\epsilon)\frac{T-t}{n-d}\sum_{j=d+1}^{n} c_j\right)}\\
\le&\exp\left\{(T-t)\cdot\frac{\lambda+\lambda^2}{n-d}\sum_{j=d+1}^{n} c_j-\lambda(1+\epsilon)\frac{T-t}{n-d}\sum_{j=d+1}^{n} c_j\right\}\\
\le&\exp\left\{(\lambda^2-\lambda\epsilon)\cdot\frac{T-t}{n-d}\sum_{j=d+1}^{n} c_j\right\}.
\end{aligned}$$

By taking $\lambda \triangleq \frac{\epsilon}{2}$, we obtain

$$\mathbb{P}\left(\sum_{j=d+1}^{n} c_j X_j \ge (1+\epsilon)\frac{T-t}{n-d}\sum_{j=d+1}^{n} c_j\right) \le \exp\left(-\frac{1}{4}\epsilon^2 \times \frac{T-t}{n-d}\sum_{j=d+1}^{n} c_j\right).$$

Also note that

$$\frac{S}{n} = \frac{1}{n}\sum_{j=1}^{n} c_j \le \frac{1}{n-d}\sum_{j=d+1}^{n} c_j.$$

Therefore, together with (C.29),

$$\begin{aligned}
&\mathbb{P}\left(Y_{t,T} \ge (1+\epsilon)\frac{(T-t)S}{n-d}\,\middle|\, W_{1,t},\ldots,W_{n,t}\right)\\
\le&\mathbb{P}\left(Y_{t,T} \ge (1+\epsilon)\frac{T-t}{n-d}\sum_{j=d+1}^{n} c_j\,\middle|\, W_{1,t},\ldots,W_{n,t}\right)\\
\le&\exp\left(-\frac{1}{4}\epsilon^2 \times \frac{T-t}{n-d}\sum_{j=d+1}^{n} c_j\right)
\end{aligned}$$

$$\le \exp\left(-\frac{1}{4}\epsilon^2 \times \frac{(T-t)S}{n}\right).$$

Proof of (C.28)**:** Similarly to above, we can construct a coupled process $\left(\underline{W}_{j,s}\right)_{s\ge t}$ under which $\underline{J}_s$ is resampled among $\{1,\dots,n-d\}$ whenever a proposal goes to one of d women who have largest c_j values (i.e., when $J_s \in \{n-d+1,\dots,n\}$) while $\mathbb{P}\left(\underline{J}_s = j \middle| \mathcal{H}_s\right) = \frac{1}{n-d}$ for any $j \in \{1,\cdots,n-d\}$ and any $\mathcal{H}_s$. With this process, we have

$$\sum_{j\in\mathcal{W}} c_j(W_{j,T}-W_{j,t}) \ge \sum_{j\in\mathcal{W}} c_j(\underline{W}_{j,T}-\underline{W}_{j,t}) \overset{\mathrm{d}}{=} \sum_{j=1}^{n-d} c_j X_j,$$

where $X_j \sim \text{Binomial}\left(T-t, \frac{1}{n-d}\right)$ for $j \in \{1,\dots,n-d\}$ and X_j's are NA.

For any $\lambda \in [-1,0]$,

$$\begin{aligned}
\mathbb{E}\left[\exp\left(\lambda\sum_{j=1}^{n-d} c_j X_j\right)\right] &\le \prod_{j=1}^{n-d}\mathbb{E}\left[e^{\lambda c_j X_j}\right] \\
&= \prod_{j=1}^{n-d}\left(1-\frac{1}{n-d}+\frac{1}{n-d}e^{\lambda c_j}\right)^{T-t} \\
&\le \prod_{j=1}^{n-d}\exp\left(-\frac{1}{n-d}+\frac{1}{n-d}e^{\lambda c_j}\right)^{T-t} \\
&= \exp\left\{(T-t)\left(-1+\frac{1}{n-d}\sum_{j=1}^{n-d}e^{\lambda c_j}\right)\right\}.
\end{aligned}$$

Since $c_j \in [0,1]$ and $e^x \le 1+x+x^2$ for any $x \in (-\infty,1]$, we have for any $\lambda \in [-1,0]$,

$$-1+\frac{1}{n-d}\sum_{j=1}^{n-d}e^{\lambda c_j} \le -1+\frac{1}{n-d}\sum_{j=1}^{n-d}(1+\lambda c_j+\lambda^2 c_j^2) \le \frac{\lambda+\lambda^2}{n-d}\sum_{j=1}^{n-d}c_j\,.$$

Using Markov's inequality, we have

$$\begin{aligned}
&\mathbb{P}\left(\sum_{j=1}^{n-d} c_j X_j \le (1-\epsilon)\frac{T-t}{n-d}\sum_{j=1}^{n-d}c_j\right) \\
=\ &\mathbb{P}\left(\exp\left(\lambda\sum_{j=1}^{n-d}c_jX_j\right) \ge \exp\left(\lambda(1-\epsilon)\frac{T-t}{n-d}\sum_{j=1}^{n-d}c_j\right)\right) \\
\le\ &\frac{\mathbb{E}\left[\exp\left(\lambda\sum_{j=1}^{n-d}c_jX_j\right)\right]}{\exp\left(\lambda(1-\epsilon)\frac{T-t}{n-d}\sum_{j=1}^{n-d}c_j\right)}
\end{aligned}$$

$$\leq \exp\left\{(T-t)\cdot\frac{\lambda+\lambda^2}{n-d}\sum_{j=1}^{n-d}c_j-\lambda(1-\epsilon)\frac{T-t}{n-d}\sum_{j=1}^{n-d}c_j\right\}$$
$$\leq \exp\left\{(\lambda^2+\lambda\epsilon)\cdot\frac{T-t}{n-d}\sum_{j=1}^{n-d}c_j\right\}.$$

With $\lambda \triangleq -\frac{\epsilon}{2}$, we obtain

$$\mathbb{P}\left(\sum_{j=1}^{n-d}c_jX_j \leq (1-\epsilon)\frac{T-t}{n-d}\sum_{j=1}^{n-d}c_j\right) \leq \exp\left(-\frac{1}{4}\epsilon^2 \times \frac{T-t}{n-d}\sum_{j=1}^{n-d}c_j\right).$$

Consequently, since $S-d=\sum_{j=1}^{n}c_j - d \leq \sum_{j=1}^{n-d}c_j$,

$$\mathbb{P}\left(Y_{t,T} \leq (1-\epsilon)\frac{(T-t)(S-d)}{n-d}\middle| W_{1,t},\ldots,W_{n,t}\right)$$
$$\leq \mathbb{P}\left(Y_{t,T} \leq (1-\epsilon)\frac{T-t}{n-d}\sum_{j=1}^{n-d}c_j\middle| W_{1,t},\ldots,W_{n,t}\right)$$
$$\leq \exp\left(-\frac{1}{4}\epsilon^2\times\frac{T-t}{n-d}\sum_{j=1}^{n-d}c_j\right)$$
$$\leq \exp\left(-\frac{1}{4}\epsilon^2\times\frac{(T-t)(S-d)}{n-d}\right).$$

□

Lemma C.15. *Consider the setting of Theorem 3.1 and $\bar{R}[t]$ defined in* (C.26)*. There exists $n_0<\infty$ such that for all $n>n_0$, we have*

$$\mathbb{P}\left(\bar{R}\left[n(\sqrt{d}+d^{\frac{1}{4}})\right] \leq \sqrt{d}-2.3d^{\frac{1}{4}}\right) \leq \exp\left(-\frac{n}{8}\right). \tag{C.30}$$
$$\mathbb{P}\left(\bar{R}\left[n(\sqrt{d}-5d^{\frac{1}{4}})\right] \geq \sqrt{d}+7.5d^{\frac{1}{4}}\right) \leq \exp\left(-\frac{n}{d^4}\right). \tag{C.31}$$

Proof. **Proof of** (C.30)**:** Fix $t=n\left(\sqrt{d}+d^{\frac{1}{4}}\right)$ and let $S\triangleq\sum_{j\in\mathcal{W}}\frac{1}{W_{j,t}+1}$. Due to the convexity of $f(x)\triangleq\frac{1}{x+1}$, we have for large enough d (i.e. large enough n since $d=\omega(1)$),

$$\frac{S}{n}=\frac{1}{n}\sum_{j\in\mathcal{W}}f(W_{j,t}) \geq f\left(\frac{1}{n}\sum_{j\in\mathcal{W}}W_{j,t}\right)=f\left(\frac{t}{n}\right)=\frac{1}{t/n+1}\geq\frac{1}{\sqrt{d}+1.05d^{\frac{1}{4}}}. \tag{C.32}$$

Given the asymptotic condition, for large enough n,

$$\frac{|k|}{n} = O(e^{-\sqrt{d}}) \le 0.1 d^{-\frac{1}{4}}, \qquad \frac{t}{nd} \le \frac{\sqrt{d} + d^{\frac{1}{4}}}{d} = d^{-\frac{1}{2}} + d^{-\frac{3}{4}} \le 0.1 d^{-\frac{1}{4}},$$
$$\frac{S-d}{n/\sqrt{d}} \ge \frac{1}{1 + 1.05 d^{-\frac{1}{4}}} - \frac{d^{\frac{3}{2}}}{n} \ge 1 - 1.1 d^{-\frac{1}{4}}.$$

Therefore,

$$\begin{aligned}\frac{((n+k)d-t)(S-d)}{n-d} &\ge \frac{((n+k)d-t)(S-d)}{n} \\ &\ge d\left(1 - \frac{|k|}{n} - \frac{t}{nd}\right) \times \frac{n}{\sqrt{d}} \cdot \frac{S-d}{n/\sqrt{d}} \\ &\ge n\sqrt{d} \times \left(1 - 0.1 d^{-\frac{1}{4}} - 0.1 d^{-\frac{1}{4}}\right) \times \left(1 - 1.1 d^{-\frac{1}{4}}\right) \\ &\ge n\sqrt{d}\left(1 - 1.3 d^{-\frac{1}{4}}\right). \end{aligned} \tag{C.33}$$

Utilizing Lemma C.14, with[3] $c_j \triangleq \frac{1}{W_{j,t}+1}$, $T \triangleq (n+k)d$ and $\epsilon \triangleq d^{-\frac{1}{4}}$, we obtain

$$\begin{aligned}\mathbb{P}\left(\bar{R}[t] \le \sqrt{d} - 2.3 d^{\frac{1}{4}} \,\middle|\, \vec{W}_t\right) &= \mathbb{P}\left(n\bar{R}[t] \le n\sqrt{d}\left(1 - 2.3 d^{-\frac{1}{4}}\right) \,\middle|\, \vec{W}_t\right) \\ &\le \mathbb{P}\left(n\bar{R}[t] \le (1-\epsilon) \times n\sqrt{d}\left(1 - 1.3 d^{-\frac{1}{4}}\right) \,\middle|\, \vec{W}_t\right) \\ &\le \mathbb{P}\left(n\bar{R}[t] \le (1-\epsilon) \times \frac{((n+k)d-t)(S-d)}{n-d} \,\middle|\, \vec{W}_t\right) \\ &\le \exp\left(-\frac{1}{4}\epsilon^2 \times \frac{((n+k)d-t)(S-d)}{n-d}\right) \\ &\le \exp\left(-\frac{1}{4} d^{-\frac{1}{2}} \times n\sqrt{d}\left(1 - 1.3 d^{-\frac{1}{4}}\right)\right) \\ &\le \exp(-n/8), \end{aligned}$$

where the last inequality follows from the fact that $1 - 1.3 d^{-\frac{1}{4}} \ge \frac{1}{2}$ for large enough d. Since the above result holds for any realization of $\vec{W}_t$, the claim follows.

Proof of (C.31)**:** Fix $t = n(\sqrt{d} - 5 d^{\frac{1}{4}})$. Define

$$\mathcal{A} \triangleq \left\{ \vec{W}_t \,\middle|\, \frac{1}{n} \sum_{j \in \mathcal{W}} \frac{1}{W_{j,t}+1} \le \frac{n}{t} + \frac{d^2}{n} + d^{-\frac{3}{4}} \right\} \subset \mathbb{N}^{|\mathcal{W}|}.$$

[3]In Lemma C.14, we assume that c_j's are some deterministic constants whereas we set $c_j \triangleq \frac{1}{W_{j,t}+1}$ here. This is fine because the results of Lemma C.14 are stated in terms of conditional probability given $\vec{W}_t$.

Applying Lemma C.8 with $\Delta \triangleq d^{-\frac{3}{4}}$, we obtain

$$\mathbb{P}\left(\vec{W}_t \notin \mathcal{A}\right) = \mathbb{P}\left(\frac{S}{n} > \frac{n}{t} + \frac{d^2}{n} + d^{-\frac{3}{4}}\right) \leq 2\exp\left(-\frac{n\Delta^2}{8d}\right) = 2\exp\left(-\frac{1}{8}nd^{-\frac{5}{2}}\right) .$$

Regarding the last term, observe that for large enough d, we have

$$2\exp\left(-\frac{1}{8}nd^{-\frac{5}{2}}\right) \leq \frac{1}{2}\exp\left(-nd^{-4}\right) .$$

For any $\vec{W}_t \in \mathcal{A}$, we have for large enough n,

$$\begin{aligned}\frac{S}{n} \leq& \frac{n}{t} + \frac{d^2}{n} + d^{-\frac{3}{4}} = \frac{1}{\sqrt{d} - 5d^{\frac{1}{4}}} + \frac{d^2}{n} + d^{-\frac{3}{4}} = \frac{1}{\sqrt{d}}\left(\frac{1}{1-5d^{-\frac{1}{4}}} + \frac{d^{\frac{5}{2}}}{n} + d^{-\frac{1}{4}}\right)\\ \leq& \frac{1}{\sqrt{d}}\left(1+6.1d^{-\frac{1}{4}}\right),\end{aligned}$$

Furthermore, given the asymptotic conditions,

$$\begin{aligned}\frac{((n+k)d-t)S}{n-d} &\leq \frac{(n+k)dS}{n-d}\\ &\leq \frac{n+|k|}{n}\cdot\frac{n}{n-d}\cdot dS\\ &\leq \frac{n+|k|}{n}\cdot\frac{n}{n-d}\cdot n\sqrt{d}\left(1+6.1d^{-\frac{1}{4}}\right)\\ &= \left(1+\frac{|k|}{n}\right)\cdot\frac{1}{1-d/n}\cdot n\sqrt{d}\left(1+6.1d^{-\frac{1}{4}}\right)\\ &\leq \left(1+0.1d^{-\frac{1}{4}}\right)\cdot\left(1+0.1d^{-\frac{1}{4}}\right)\cdot n\sqrt{d}\left(1+6.1d^{-\frac{1}{4}}\right)\\ &\leq n\sqrt{d}\left(1+6.4d^{-\frac{1}{4}}\right) ,\end{aligned}$$

where we used the fact that $\frac{|k|}{n} = O(e^{-\sqrt{d}}) \leq 0.1d^{-\frac{1}{4}}$ and $\frac{d}{n} \leq 0.1d^{-\frac{1}{4}}$ for large enough n. We further utilize Lemma C.14: By taking $c_j \triangleq \frac{1}{W_{j,t+1}}$, $T \triangleq (n+k)d$ and $\epsilon \triangleq d^{-\frac{1}{4}}$, we obtain

$$\begin{aligned}\mathbb{P}\left(\bar{R}[t] \geq \sqrt{d} + 7.5d^{\frac{1}{4}} \middle| \vec{W}_t\right) &= \mathbb{P}\left(n\bar{R}[t] \geq n\sqrt{d}\left(1+7.5d^{-\frac{1}{4}}\right) \middle| \vec{W}_t\right)\\ &\leq \mathbb{P}\left(n\bar{R}[t] \geq (1+\epsilon)\times n\sqrt{d}\left(1+6.4d^{-\frac{1}{4}}\right) \middle| \vec{W}_t\right)\\ &\leq \mathbb{P}\left(n\bar{R}[t] \geq (1+\epsilon)\times\frac{((n+k)d-t)S}{n-d} \middle| \vec{W}_t\right)\\ &\leq \exp\left(-\frac{1}{4}\epsilon^2\times\frac{((n+k)d-t)S}{n}\right).\end{aligned}$$

for any $\vec{W}_t \in \mathcal{A}$. Also note that, from (C.33), for large enough n,

$$\frac{((n+k)d-t)S}{n} \geq \left(\left(1-\frac{|k|}{n}\right)d - \frac{t}{n}\right)S \geq \left(\left(1-0.1d^{-\frac{1}{4}}\right)d - \sqrt{d} + 5d^{\frac{1}{4}}\right)S\,,$$

where we used the fact that $\frac{|k|}{n} = O(e^{-\sqrt{d}}) \leq 0.1d^{-\frac{1}{4}}$. Because $\left(1-0.1d^{-\frac{1}{4}}\right)d - \sqrt{d} + 5d^{\frac{1}{4}} \geq d - 1.3d^{\frac{3}{4}}$ for large enough n, and that $S \geq \frac{n}{t/n+1}$ as derived in (C.32), we have

$$\frac{((n+k)d-t)S}{n} \geq \left(d-1.3d^{\frac{3}{4}}\right)\frac{n}{t/n+1} \geq \left(d-1.3d^{\frac{3}{4}}\right)\frac{n}{\sqrt{d}} \geq n\sqrt{d}\left(1-1.3d^{-\frac{1}{4}}\right)\,,$$

and therefore,

$$\exp\left(-\frac{1}{4}\epsilon^2 \times \frac{((n+k)d-t)S}{n}\right) \leq \exp\left(-\frac{1}{4}d^{-\frac{1}{2}} \times n\sqrt{d}\left(1-1.3d^{-\frac{1}{4}}\right)\right) \leq \exp\left(-n/8\right).$$

Combining all results, we obtain the desired result: for large enough n,

$$\begin{aligned}
\mathbb{P}\left(\bar{R}[t] \geq \sqrt{d} + 7.5d^{\frac{1}{4}}\right) &\leq \mathbb{P}\left(\bar{R}[t] \geq \sqrt{d} + 7.5d^{\frac{1}{4}} \middle| \vec{W}_t \in \mathcal{A}\right) \cdot \mathbb{P}\left(\vec{W}_t \in \mathcal{A}\right) + \mathbb{P}\left(\vec{W}_t \notin \mathcal{A}\right) \\
&\leq \exp\left(-n/8\right) + \frac{1}{2}\exp\left(-nd^{-4}\right) \\
&\leq \exp\left(-nd^{-4}\right)\,,
\end{aligned}$$

where the last inequality follows from that $\frac{n}{8} \geq \frac{n}{d^4} + \log 2$ for large enough n and d. □

Concentration of women's average rank R_{WOMEN}

The following lemma states that conditioned on $(\vec{W}_\tau, \vec{W}_{(n+k)d})$, $R_{\text{WOMEN}}(\text{MOSM})$ is concentrated around $\bar{R}[\tau]$.

Lemma C.16. *For any given n, k and d and $(\vec{W}_\tau, \vec{W}_{(n+k)d})$ which arises with positive probability, we have $\mathbb{E}[R_{\text{WOMEN}}(\text{MOSM})|\vec{W}_\tau, \vec{W}_{(n+k)d}] = 1 + \bar{R}[\tau]$. Furthermore, for any $\epsilon > 0$ we have*

$$\mathbb{P}\left(R_{\text{WOMEN}}(\text{MOSM}) \geq 1 + (1+\epsilon)\bar{R}[\tau] \,\middle|\, \vec{W}_\tau, \vec{W}_{(n+k)d}\right) \leq \exp\left(-\frac{2\epsilon^2 n^2 \bar{R}[\tau]^2}{\sum_{j\in\mathcal{W}} W_{j,(n+k)d}^2}\right)\,, \tag{C.34}$$

$$\mathbb{P}\left(R_{\text{WOMEN}}(\text{MOSM}) \le 1 + (1-\epsilon)\bar{R}[\tau] \,\Big|\, \vec{W}_\tau, \vec{W}_{(n+k)d} \right) \le \exp\left(-\frac{2\epsilon^2 n^2 \bar{R}[\tau]^2}{\sum_{j\in\mathcal{W}} W_{j,(n+k)d}^2} \right). \tag{C.35}$$

Proof. Within this proof, we assume that τ, $\vec{W}_\tau = \left(W_{j,\tau}\right)_{j\in\mathcal{W}}$, and $\vec{W}_{(n+k)d} = \left(W_{j,(n+k)d}\right)_{j\in\mathcal{W}}$ are revealed (and hence so is $\bar{R}[\tau]$). In what follows, $\mathbb{P}(\cdot)$ and $\mathbb{E}[\cdot]$ denote the associated conditional probability and the conditional expectation, respectively.

For brevity, let $w_j \triangleq W_{j,\tau}$, $w'_j \triangleq W_{j,(n+k)d} - W_{j,\tau}$, and $R_j \triangleq \text{Rank}_j(\text{MOSM})|\vec{W}_\tau, \vec{W}_{(n+k)d}$. Note that a woman j receives w_j proposals until time τ and receives w'_j proposals after time τ (the total number of proposals $w_j + w'_j = W_{j,(n+k)d}$ equals to her degree). Under MOSM, each woman j is matched to her most preferred one among the first w_j proposals, and the rank of her matched partner under MOSM, R_j, can be determined by the number of men among the remaining (at time τ) w'_j men on her list that she prefers to her matched partner.

More specifically, fix j and let Z_t^j be the indicator that the woman j prefers her t^{th} proposal to all of her first w_j proposals for $t \in \{w_j+1, \ldots, w_j+w'_j\}$. Then, the rank R_j can be represented as

$$\begin{aligned} R_j &= 1 + \sum_{t=w_j+1}^{w_j+w'_j} \mathbb{I}\left(\text{woman } j \text{ prefers her } t^{\text{th}} \text{ proposal to all of her first } w_j \text{ proposals}\right) \\ &= 1 + \sum_{t=w_j+1}^{w_j+w'_j} Z_t^j \,. \end{aligned}$$

Note that $(Z_t^j)_{t=w_j+1}^{w_j+w'_j}$ has the same distribution as $(\mathbb{I}\{U_t^j > V_j\})_{t=w_j+1}^{w_j+w'_j}$, where $(U_t^j)_{t=w_j+1}^{w_j+w'_j}$ are i.i.d. Uniform$[0,1]$ random variables, V_j is the largest order statistic of w_j i.i.d. Uniform$[0,1]$ random variables, and V_j is independent of U_t^j's. Therefore,

$$\mathbb{E}\left[R_j\right] = 1 + w'_j \cdot \mathbb{E}[Z_{w_j+1}^j] = 1 + w'_j \cdot \mathbb{P}(U_{w_j+1}^j > V_j) = 1 + \frac{w'_j}{w_j+1}\,,$$

and

$$\mathbb{E}[R_{\text{WOMEN}}(\text{MOSM})|\vec{W}_\tau, \vec{W}_{(n+k)d}] = \frac{1}{n}\sum_{j\in\mathcal{W}} \mathbb{E}[R_j] = 1 + \bar{R}[\tau]\,,$$

which proves the first claim in Lemma C.16.

Note that $(R_j)_{j\in\mathcal{W}}$ are i.i.d. and that $R_j \in [0, w'_j]$. Applying Hoeffding's inequality, we have

$$\mathbb{P}\left(R_{\text{WOMEN}}(\text{MOSM}) \geq 1 + (1+\epsilon)\bar{R}[\tau] \,\middle|\, \vec{W}_\tau, \vec{W}_{(n+k)d}\right) = \mathbb{P}\left(\frac{1}{n}R_j \geq 1 + (1+\epsilon)\mathbb{E}[\bar{R}[\tau]]\right)$$
$$\leq \exp\left(-\frac{2\epsilon^2 n^2 \bar{R}[\tau]^2}{\sum_{j\in\mathcal{W}} W_{j,(n+k)d}^2}\right).$$

Similarly, we can show that

$$\mathbb{P}\left(R_{\text{WOMEN}}(\text{MOSM}) \leq 1 + (1-\epsilon)\bar{R}[\tau] \,\middle|\, \vec{W}_\tau, \vec{W}_{(n+k)d}\right) \leq \exp\left(-\frac{2\epsilon^2 n^2 \bar{R}[\tau]^2}{\sum_{j\in\mathcal{W}} W_{j,(n+k)d}^2}\right).$$

This concludes the proof.

□

Proof of Proposition C.5. We obtain a high probability lower bound on R_{WOMEN} by combining the results of Proposition C.1, and Lemmas C.15 and C.16. By Proposition C.1 and Lemma C.15 and by the fact that $\bar{R}[t]$ is decreasing on each sample path,

$$\begin{aligned}
&\mathbb{P}\left(\bar{R}[\tau] \leq \sqrt{d} - 2.3d^{\frac{1}{4}}\right)\\
&\leq \mathbb{P}\left(\bar{R}[\tau] \leq \sqrt{d} - 2.3d^{\frac{1}{4}}, \tau < n(\sqrt{d} + d^{\frac{1}{4}})\right) + \mathbb{P}\left(\tau \geq n(\sqrt{d} + d^{\frac{1}{4}})\right)\\
&\leq \mathbb{P}\left(\bar{R}[n(\sqrt{d} + d^{\frac{1}{4}})] \leq \sqrt{d} - 2.3d^{\frac{1}{4}}, \tau < n(\sqrt{d} + d^{\frac{1}{4}})\right) + \mathbb{P}\left(\tau \geq n(\sqrt{d} + d^{\frac{1}{4}})\right)\\
&\leq \exp\left(-\frac{n}{8}\right) + O(\exp(-\sqrt{n})) = O(\exp(-\sqrt{n}))\,. \qquad \text{(C.36)}
\end{aligned}$$

We also need a high probability upper bound on $\sum_{j\in\mathcal{W}} W_{j,(n+k)d}^2$. Since $W_{j,(n+k)d} \sim \text{Binomial}((n+k)d, \frac{1}{n})$, we have for large enough n,

$$\mathbb{E}\left[W_{j,(n+k)d}^2\right] = \mathbb{E}^2\left[W_{j,(n+k)d}\right] + \text{Var}\left[W_{j,(n+k)d}\right] = (n+k)^2 d^2 \frac{1}{n^2}\left(1 + (1-\frac{1}{n})^2\right) \leq 2d^2\,.$$

Denote $\mu \triangleq \mathbb{E}[W_{1,(n+k)d}] = \frac{(n+k)d}{n}$. Looking up the table of the central moments of Binomial distribution, we have

$$\mathbb{E}[(W_{1,(n+k)d} - \mu)^4] = (n+k)d\frac{1}{n}\left(1 - \frac{1}{n}\right)\left(1 + (3(n+k)d - 6)\frac{1}{n}\left(1 - \frac{1}{n}\right)\right).$$

Using the fact that $k = o(n)$, $d = o(n)$ and $d = \omega(1)$, we have for large enough n,

$$\mathbb{E}[(W_{1,(n+k)d} - \mu)^4] \leq \ 2d\left(1 + 3(n+k)d\frac{1}{n}\right) \leq \ 2d \cdot 4d = \ 8d^2 \,.$$

Therefore, for large enough n,

$$\begin{aligned}
\mathrm{Var}[W^2_{1,(n+k)d}] &\leq \mathbb{E}[W^4_{1,(n+k)d}] \\
&= \mathbb{E}[(\mu + (W_{1,(n+k)d} - \mu))^4] \\
&\leq 8\mu^4 + 8\mathbb{E}[(W_{1,(n+k)d} - \mu)^4] \\
&= 8\frac{(n+k)^4 d^4}{n^4} + 64d^2 \\
&\leq 10d^4 \,.
\end{aligned}$$

In the proof of Lemma C.4, we have shown that $W_{1,(n+k)d}, \ldots, W_{n,(n+k)d}$ are NA. By Lemma C.3–((iii)), $W^2_{1,(n+k)d}, \ldots, W^2_{n,(n+k)d}$ are NA, hence we have for large enough n,

$$\mathrm{Var}\left[\sum_{j\in\mathcal{W}} W^2_{j,(n+k)d}\right] \leq \ n\mathrm{Var}\left[W^2_{1,(n+k)d}\right] \leq 10nd^4 \,.$$

Applying Chebyshev's inequality, we have for large enough n,

$$\begin{aligned}
&\mathbb{P}\left(\sum_{j\in\mathcal{W}} W^2_{j,(n+k)d} \geq 4nd^2\right) \\
&\leq \mathbb{P}\left(\sum_{j\in\mathcal{W}} (W^2_{j,(n+k)d} - \mathbb{E}[W^2_{j,(n+k)d}]) \geq 2nd^2\right) \\
&= \mathbb{P}\left(\sum_{j\in\mathcal{W}} (W^2_{j,(n+k)d} - \mathbb{E}[W^2_{j,(n+k)d}]) \geq \frac{2\sqrt{n}}{\sqrt{10}}\sqrt{10nd^4}\right) \\
&\leq \mathbb{P}\left(\sum_{j\in\mathcal{W}} (W^2_{j,(n+k)d} - \mathbb{E}[W^2_{j,(n+k)d}]) \geq \frac{2\sqrt{n}}{\sqrt{10}}\sqrt{\mathrm{Var}\left[\sum_{j\in\mathcal{W}} W^2_{j,(n+k)d}\right]}\right) \\
&\leq \frac{5}{2n} \leq \ \frac{3}{n} \,.
\end{aligned} \tag{C.37}$$

Given that $\bar{R}[\tau] > \sqrt{d} - 2.3d^{\frac{1}{4}}$, by plugging $\epsilon \triangleq 0.5d^{-\frac{1}{4}}$ in (C.35) of Lemma C.16, we obtain for large enough n,

$$1 + (1-\epsilon)\bar{R}[\tau] \geq 1 + (1 - 0.5d^{-\frac{1}{4}}) \cdot \sqrt{d}(1 - 2.3d^{-\frac{1}{4}}) \geq \sqrt{d}(1 - 3d^{-\frac{1}{4}}) = \sqrt{d} - 3d^{\frac{1}{4}} \,. \tag{C.38}$$

Therefore,

$$
\begin{aligned}
&\mathbb{P}\left(R_{\text{WOMEN}} \le \sqrt{d} - 3d^{\frac{1}{4}}\right) \\
\le\ &\mathbb{P}\left(R_{\text{WOMEN}} \le \sqrt{d} - 3d^{\frac{1}{4}} \,\middle|\, \bar{R}[\tau] > \sqrt{d} - 2.3d^{\frac{1}{4}}, \sum_{j\in\mathcal{W}} W^2_{j,(n+k)d} < 4nd^2\right) \\
&+ \mathbb{P}\left(\bar{R}[\tau] \le \sqrt{d} - 2.3d^{\frac{1}{4}}\right) + \mathbb{P}\left(\sum_{j\in\mathcal{W}} W^2_{j,(n+k)d} \ge 4nd^2\right) \\
\overset{(a)}{\le}\ &\mathbb{P}\left(R_{\text{WOMEN}} \le 1 + (1-\epsilon)\bar{R}[\tau] \,\middle|\, \bar{R}[\tau] > \sqrt{d} - 2.3d^{\frac{1}{4}}, \sum_{j\in\mathcal{W}} W^2_{j,(n+k)d} < 4nd^2\right) \\
&+ O(\exp(-\sqrt{n})) + \frac{3}{n} \\
\overset{(b)}{\le}\ &\mathbb{E}\left[\exp\left(-\frac{\frac{1}{2}d^{-\frac{1}{2}}n^2\bar{R}[\tau]^2}{4nd^2}\right) \,\middle|\, \bar{R}[\tau] > \sqrt{d} - 2.3d^{\frac{1}{4}}\right] + \frac{4}{n} \\
\le\ &\exp\left(-\frac{1}{8}d^{-\frac{5}{2}}n \cdot d(1-2.3d^{-\frac{1}{4}})^2\right) + \frac{4}{n} \\
\le\ &\exp\left(-\frac{nd^{-\frac{3}{2}}}{16}\right) + \frac{4}{n} \\
\le\ &\frac{5}{n}\,.
\end{aligned}
$$

Here inequality (a) follows from (C.38), (C.36), and (C.37); inequality (b) follows from Lemma C.16. □

Proof of Proposition C.6. We obtain a high probability lower bound on R_{WOMEN} by combining the results of Proposition C.4, and Lemma C.15 and C.16. By Proposition C.4 and Lemma C.15,

$$
\begin{aligned}
&\mathbb{P}\left(\bar{R}[\tau] \ge \sqrt{d} + 7.5d^{\frac{1}{4}}\right) \\
\le\ &\mathbb{P}\left(\bar{R}[\tau] \ge \sqrt{d} + 7.5d^{\frac{1}{4}}, \tau > n(\sqrt{d} - 5d^{\frac{1}{4}})\right) + \mathbb{P}\left(\tau \le n(\sqrt{d} - 5d^{\frac{1}{4}})\right) \\
\le\ &\mathbb{P}\left(\bar{R}[n(\sqrt{d} - 5d^{\frac{1}{4}})] \ge \sqrt{d} + 7.5d^{\frac{1}{4}}, \tau > n(\sqrt{d} - 5d^{\frac{1}{4}})\right) + \mathbb{P}\left(\tau \le n(\sqrt{d} - 5d^{\frac{1}{4}})\right) \\
\le\ &\exp\left(-\frac{n}{d^4}\right) + O(\exp(-d^{\frac{1}{4}})) \\
\le\ &O(\exp(-d^{\frac{1}{4}}))\,.
\end{aligned}
$$

Given that $\bar{R}[\tau] < \sqrt{d} + 7.5d^{\frac{1}{4}}$, by plugging $\epsilon \triangleq 0.1d^{-\frac{1}{4}}$ in (C.34) of Lemma C.16, we obtain

$$1 + (1+\epsilon)\bar{R}[\tau] \le 1 + (1 + 0.1d^{-\frac{1}{4}}) \cdot \sqrt{d}(1 + 7.5d^{-\frac{1}{4}}) \le \sqrt{d}(1 + 8d^{-\frac{1}{4}}) = \sqrt{d} + 8d^{\frac{1}{4}},$$

for large enough n. Recall that we have shown in the proof of Proposition C.5 that

$$\mathbb{P}\left(\sum_{j\in\mathcal{W}} W_{j,(n+k)d}^2 \ge 4nd^2\right) \le \frac{3}{n}.$$

Therefore, similar to the proof of Proposition C.5, we have

$$\begin{aligned}
&\mathbb{P}\left(R_{\text{WOMEN}} \ge \sqrt{d} + 8d^{\frac{1}{4}}\right) \\
&\le \mathbb{P}\left(R_{\text{WOMEN}} \ge \sqrt{d} + 8d^{\frac{1}{4}} \,\middle|\, \sqrt{d} - 2.3d^{\frac{1}{4}} < \bar{R}[\tau] < \sqrt{d} + 7.5d^{\frac{1}{4}}, \sum_{j\in\mathcal{W}} W_{j,(n+k)d}^2 < 4nd^2\right) \\
&\quad + \mathbb{P}\left(\bar{R}[\tau] \le \sqrt{d} - 2.3d^{\frac{1}{4}}\right) + \mathbb{P}\left(\bar{R}[\tau] \ge \sqrt{d} + 7.5d^{\frac{1}{4}}\right) + \mathbb{P}\left(\sum_{j\in\mathcal{W}} W_{j,(n+k)d}^2 \ge 4nd^2\right) \\
&\le \mathbb{P}\left(R_{\text{WOMEN}} \ge 1 + (1+\epsilon)\bar{R}[\tau] \,\middle|\, \sqrt{d} - 2.3d^{\frac{1}{4}} < \bar{R}[\tau] < \sqrt{d} + 7.5d^{\frac{1}{4}}, \sum_{j\in\mathcal{W}} W_{j,(n+k)d}^2 < 4nd^2\right) \\
&\quad + O(\exp(-d^{\frac{1}{4}})) \\
&\le \mathbb{E}\left[\exp\left(-\frac{2\epsilon^2 n^2 \bar{R}[\tau]^2}{\sum_{j\in\mathcal{W}} W_{j,(n+k)d}^2}\right) \,\middle|\, \sqrt{d} - 2.3d^{\frac{1}{4}} < \bar{R}[\tau] < \sqrt{d} + 7.5d^{\frac{1}{4}}, \sum_{j\in\mathcal{W}} W_{j,(n+k)d}^2 < 4nd^2\right] \\
&\quad + O(\exp(-d^{\frac{1}{4}})) \\
&\le \exp\left(-\frac{1}{200} d^{-\frac{5}{2}} n \cdot d(1 - 2.3d^{-\frac{1}{4}})^2\right) + O(\exp(-d^{\frac{1}{4}})) \\
&\le \exp\left(-\frac{nd^{-\frac{3}{2}}}{300}\right) + O(\exp(-d^{\frac{1}{4}})) = O(\exp(-d^{\frac{1}{4}})).
\end{aligned}$$

□

C.2.4 Proof of Theorem C.1

Theorem C.1 immediately follows from Propositions C.1, C.2, C.3, C.4, C.5, and C.6.

C.3 Proof for Large Sized d: the Case of $d = \omega(\log^2 n)$, $d = o(n)$

In this section, we consider the case such that $d = \omega(\log^2 n)$ and $d = o(n)$. We will prove a quantitative version of Theorem 3.2.

Theorem C.2 (Quantitative version of Theorem 3.2). *Consider a sequence of random matching markets indexed by n, with $n + k$ men and n women ($k = k(n)$ is negative), and the men's degrees are $d = d(n)$. If $|k| = o(n)$, $d = \omega(\log^2 n)$ and $d = o(n)$, we have the following results.*

1. Men's average rank of wives. *With probability* $1 - \exp(-\sqrt{\log n})$*, we have*

$$R_{\text{MEN}}(\text{MOSM}) \le \left(1 + 2\frac{|k|}{n} + 2\frac{1}{\sqrt{\log n}}\right) \log n \,.$$

2. Women's average rank of husbands. *With probability* $1 - O(\exp(-\sqrt{\log n}))$*, we have*

$$R_{\text{WOMEN}}(\text{MOSM}) \ge \left(1 - 1.1\left(\frac{|k|}{n} + \frac{3}{\sqrt{\log n}} + \frac{d}{n/\log n}\right)\right)\frac{d}{\log n} \,.$$

Proof of Theorem C.2. **Proof of Theorem C.2 part 1.** Recall that τ is the *the total number of proposals* that are made until the end of MPDA, i.e., the time at which the men-optimal stable matching (MOSM) is found. We introduce an extended process (which is *different* from the one defined in Appendix C.1.5) as a natural continuation of the MPDA procedure that continues to evolve even after the MOSM is found (i.e., the extended process continues for $t > \tau$). To define the extended process, we start by defining an extended market, which has the same n women and $n + k$ men, but each man has a complete preference list, i.e. each man ranks all n women. We call the first d women of a man's preference list his "real" preferences and the last $n - d$ women his "fake" preferences. The distribution of preferences in the extended market is again as described in Section 3.2. We then define the *extended process* as tracking the progress of Algorithm 3.1 on the extended market: the $n + k$ men enter first in Algorithm 3.1

with only their real preferences, as before. After time τ, we let the men see their fake preferences and continue Algorithm 3.1 until the MOSM with full preferences is found. We denote by τ' the total number of proposals to find the MOSM with full preferences. It is easy to see that τ is stochastically dominated by τ'.

Note that τ' is the total number of proposals needed to find the MOSM in a completely-connected market, which has been studied in previous works including [132, 133]. It is well-known that τ' is stochastically dominated by the number of draws in a coupon collector's problem, in which one coupon is chosen out of n coupons uniformly at random at a time and it runs until n distinct coupons are collected. Let X be the number of draws in the coupon collector's problem. A widely used tail bound of X is the following: for $\beta > 1$, $\mathbb{P}(X \geq \beta n \log n) \leq n^{-\beta+1}$. By taking $\beta = 1 + \frac{1}{\sqrt{\log n}}$, we have

$$\mathbb{P}\left(X \geq n \log n + n\sqrt{\log n}\right) \leq n^{-\frac{1}{\sqrt{\log n}}} = e^{-\sqrt{\log n}} = o(1)\,.$$

Hence with probability $1 - e^{-\sqrt{\log n}}$, we have $\tau \leq n(\log n + \sqrt{\log n})$. Because X stochastically dominates τ, we have, with probability $1 - e^{-\sqrt{\log n}}$,

$$R_{\text{MEN}}(\text{MOSM}) \leq \frac{n}{n+k}(\log n + \sqrt{\log n}) + 1\,.$$

Because $k = o(n)$ and $k < 0$, for large enough n we have $\frac{n}{n+k} \leq 1 + \frac{2|k|}{n}$, $\frac{|k|}{n} < \frac{1}{3}$, $1 \leq \frac{1}{3}\sqrt{\log n}$, hence

$$\begin{aligned}
&\frac{n}{n+k}(\log n + \sqrt{\log n}) + 1 \\
\leq\ & \left(1 + 2\frac{|k|}{n}\right)\log n + (1 + \frac{2}{3})\sqrt{\log n} + \frac{1}{3}\sqrt{\log n} \\
=\ & \left(1 + 2\frac{|k|}{n} + 2\frac{1}{\sqrt{\log n}}\right)\log n\,.
\end{aligned}$$

This concludes the proof.

Proof of Theorem C.2 part 2.

The proof is similar to that of Proposition C.5. Recall that the proof of Proposition C.5 relies on Proposition C.1, Lemma C.15, and Lemma C.16. In the following, we first establish the counterparts of these results in dense markets.

Counterpart of Proposition C.1 in dense markets. We have shown in the proof of Theorem C.2(1) that with probability $1 - \exp(-\sqrt{\log n})$, we have

$$\tau \leq n \left(\log n + \sqrt{\log n} \right) . \tag{C.39}$$

Counterpart of Lemma C.15 in dense markets. Fix $t = n\left(\log n + \sqrt{\log n}\right)$. Given the asymptotic condition, we have for large enough n,

$$\frac{t}{nd} = \frac{\log n + \sqrt{\log n}}{d} \leq 0.1(\log n)^{-1} .$$

By examining the proof of Lemma C.14, we can see that we have proved the following result (see the statement of Lemma C.14 for the definition of the notations), which is stronger than than (C.28):

$$\mathbb{P}\left(Y_{t,T} \leq (1-\epsilon)\frac{T-t}{n-d}\sum_{j=1}^{n-d} c_j \,\middle|\, \vec{W}_t \right) \leq \exp\left(-\frac{1}{4}\epsilon^2 \times \frac{T-t}{n-d}\sum_{j=1}^{n-d} c_j \right) \tag{C.40}$$

Let $c_j \triangleq \frac{1}{W_{j,t}+1}$ where $W_{1,t} \geq W_{2,t} \geq \cdots \geq W_{n,T}$, and $T \triangleq (n+k)d$. Due to the convexity of $f(x) \triangleq \frac{1}{x+1}$, we have for large enough n,

$$\begin{aligned} \frac{1}{n-d}\sum_{j=1}^{n-d} c_j &= \frac{1}{n-d}\sum_{j=1}^{n-d} f(W_{j,t}) \geq\ f\left(\frac{1}{n-d}\sum_{j=1}^{n-d} W_{j,t}\right) \geq\ f\left(\frac{t}{n-d}\right) \\ &= \frac{1}{t/(n-d)+1} \geq \frac{1}{\log n\left(1 + 1.05\frac{1}{\sqrt{\log n}} + 1.05\frac{d}{n}\right)} . \end{aligned} \tag{C.41}$$

Therefore, for large enough n,

$$\begin{aligned} \frac{T-t}{n-d}\sum_{j=1}^{n-d} c_j &\geq n\frac{(n+k)d - t}{n}\frac{1}{\log n\left(1 + 1.05\frac{1}{\sqrt{\log n}} + 1.05\frac{d}{n}\right)} \\ &\geq nd\left(1 - \frac{|k|}{n} - \frac{t}{nd}\right)\frac{1}{\log n\left(1 + 1.05\frac{1}{\sqrt{\log n}} + 1.05\frac{d}{n}\right)} \\ &\geq nd\left(1 - \frac{|k|}{n} - \frac{0.1}{\log n}\right)\frac{1}{\log n}\left(1 - 1.05\frac{1}{\sqrt{\log n}} - 1.05\frac{d}{n}\right) \\ &\geq \frac{nd}{\log n}\left(1 - 1.1\frac{|k|}{n} - 1.1\frac{1}{\sqrt{\log n}} - 1.1\frac{d}{n}\right) . \end{aligned} \tag{C.42}$$

Utilizing Lemma C.14 (which does not use assumptions on d) with $\epsilon \triangleq \frac{1}{\sqrt{\log n}}$, we obtain

$$
\begin{aligned}
&\mathbb{P}\left(\bar{R}[t] \le \frac{d}{\log n}\left(1-\left(1.1\frac{|k|}{n}+\frac{2.1}{\sqrt{\log n}}+1.1\frac{d}{n}\right)\right)\middle| \vec{W}_t\right)\\
&=\mathbb{P}\left(n\bar{R}[t] \le \frac{nd}{\log n}\left(1-\left(1.1\frac{|k|}{n}+\frac{2.1}{\sqrt{\log n}}+1.1\frac{d}{n}\right)\right)\middle| \vec{W}_t\right)\\
&\le \mathbb{P}\left(n\bar{R}[t] \le (1-\epsilon)\times\frac{nd}{\log n}\left(1-1.1\left(\frac{|k|}{n}+\frac{1}{\sqrt{\log n}}+\frac{d}{n}\right)\right)\middle| \vec{W}_t\right)\\
&\overset{(a)}{\le} \mathbb{P}\left(n\bar{R}[t] \le (1-\epsilon)\times\frac{(n+k)d-t}{n-d}\sum_{j=1}^{n-d} c_j\middle| \vec{W}_t\right)\\
&\overset{(b)}{\le} \exp\left(-\frac{1}{4}\epsilon^2\times\frac{(n+k)d-t}{n-d}\sum_{j=1}^{n-d} c_j\right)\\
&\overset{(c)}{\le} \exp\left(-\frac{1}{4}\frac{1}{\log n}\times\frac{nd}{\log n}\left(1-1.1\left(\frac{|k|}{n}+\frac{1}{\sqrt{\log n}}+\frac{d}{n}\right)\right)\right)\\
&\le \exp(-n/8)\,.
\end{aligned}
$$

Here inequalities (a) and (c) follow from (C.42), inequality (b) follows from Lemma C.14, and the last inequality follows from the fact that $d=\omega(\log^2 n)$. Since the above result holds for any realization of $\vec{W}_t$, we have

$$
\mathbb{P}\left(\bar{R}\left[n(\log n+\sqrt{\log n})\right] \le \frac{d}{\log n}\left(1-1.1\left(\frac{|k|}{n}+\frac{2}{\sqrt{\log n}}+\frac{d}{n}\right)\right)\right) \le \exp\left(-\frac{n}{8}\right). \tag{C.43}
$$

Counterpart of Lemma C.16 in dense markets. Note that the proof of Lemma C.16 does not make any assumption on d, hence (C.35) still holds.

Proof of Theorem C.2 part 2. Using (C.39) and (C.43), and the fact that $\bar{R}[t]$ is decreasing on each sample path,

$$
\mathbb{P}\left(\bar{R}[\tau] \le \frac{d}{\log n}\left(1-1.1\left(\frac{|k|}{n}+\frac{2}{\sqrt{\log n}}+\frac{d}{n}\right)\right)\right)
$$

$$\leq \mathbb{P}\left(\bar{R}[\tau] \leq \frac{d}{\log n}\left(1 - 1.1\left(\frac{|k|}{n} + \frac{2}{\sqrt{\log n}} + \frac{d}{n}\right)\right), \tau < n(\log n + \sqrt{\log n})\right)$$
$$+ \mathbb{P}\left(\tau \geq n(\log n + \sqrt{\log n})\right)$$
$$\leq \mathbb{P}\left(\bar{R}\left[n(\log n + \sqrt{\log n})\right] \leq \frac{d}{\log n}\left(1 - 1.1\left(\frac{|k|}{n} + \frac{2}{\sqrt{\log n}} + \frac{d}{n}\right)\right)\right)$$
$$+ \mathbb{P}\left(\tau \geq n(\log n + \sqrt{\log n})\right)$$
$$\stackrel{(a)}{\leq} \exp\left(-\frac{n}{8}\right) + O(\exp(-\sqrt{\log n})) = O(\exp(-\sqrt{\log n}))\,. \tag{C.44}$$

Here inequality (a) follows from (C.43). Recall inequality (C.37): for large enough n

$$\mathbb{P}\left(\sum_{j\in\mathcal{W}} W^2_{j,(n+k)d} \geq 4nd^2\right) \leq \frac{3}{n}\,.$$

In the derivation of the above inequality, we only used the fact that $d = \omega(1), d = o(n)$ and $k = o(n)$, which also holds in dense markets.

Given that $\bar{R}[\tau] > \frac{d}{\log n}\left(1 - 1.1\left(\frac{|k|}{n} + \frac{2}{\sqrt{\log n}} + \frac{d}{n}\right)\right)$, by plugging $\epsilon \triangleq 0.5\frac{1}{\sqrt{\log n}}$ in (C.35) of Lemma C.16, we obtain for large enough n,

$$1 + (1-\epsilon)\bar{R}[\tau] \geq 1 + \left(1 - 0.5\frac{1}{\sqrt{\log n}}\right)\frac{d}{\log n}\left(1 - 1.1\left(\frac{|k|}{n} + \frac{2}{\sqrt{\log n}} + \frac{d}{n}\right)\right)$$
$$\geq \frac{d}{\log n}\left(1 - 1.1\left(\frac{|k|}{n} + \frac{3}{\sqrt{\log n}} + \frac{d}{n}\right)\right)\,. \tag{C.45}$$

Therefore,

$$\mathbb{P}\left(R_{\text{WOMEN}} \leq \frac{d}{\log n}\left(1 - 1.1\left(\frac{|k|}{n} + \frac{3}{\sqrt{\log n}} + \frac{d}{n}\right)\right)\right)$$
$$\leq \mathbb{P}\left(R_{\text{WOMEN}} \leq \frac{d}{\log n}\left(1 - 1.1\left(\frac{|k|}{n} + \frac{3}{\sqrt{\log n}} + \frac{d}{n}\right)\right)\right|$$
$$\left.\bar{R}[\tau] > \frac{d}{\log n}\left(1 - 1.1\left(\frac{|k|}{n} + \frac{2}{\sqrt{\log n}} + \frac{d}{n}\right)\right), \sum_{j\in\mathcal{W}} W^2_{j,(n+k)d} < 4nd^2\right)$$
$$+ \mathbb{P}\left(\bar{R}[\tau] \leq \frac{d}{\log n}\left(1 - 1.1\left(\frac{|k|}{n} + \frac{2}{\sqrt{\log n}} + \frac{d}{n}\right)\right)\right) + \mathbb{P}\left(\sum_{j\in\mathcal{W}} W^2_{j,(n+k)d} \geq 4nd^2\right)$$

$$
\begin{aligned}
&\overset{\text{(a)}}{\leq} \mathbb{P}\left(R_{\text{WOMEN}} \leq 1 + (1-\epsilon)\bar{R}[\tau] \middle| \bar{R}[\tau] > \frac{d}{\log n}\left(1 - 1.1\left(\frac{|k|}{n} + \frac{2}{\sqrt{\log n}} + \frac{d}{n}\right)\right), \right.\\
&\qquad \left. \sum_{j\in\mathcal{W}} W_{j,(n+k)d}^2 < 4nd^2 \right) + O(\exp(-\sqrt{\log n})) + \frac{3}{n}\\
&\overset{\text{(b)}}{\leq} \mathbb{E}\left[\exp\left(-\frac{\frac{1}{2\log n} n^2 \bar{R}[\tau]^2}{4nd^2} \right) \middle| \bar{R}[\tau] > \frac{d}{\log n}\left(1 - 1.1\left(\frac{|k|}{n} + \frac{2}{\sqrt{\log n}} + \frac{d}{n}\right)\right) \right]\\
&\qquad + O(\exp(-\sqrt{\log n}))\\
&\leq \exp\left(-\frac{1}{8}\frac{n}{d^2\log n} \cdot \frac{d^2}{2\log^2 n} \right) + O(\exp(-\sqrt{\log n}))\\
&\leq \exp\left(-\frac{n(\log n)^{-3}}{16} \right) + O(\exp(-\sqrt{\log n}))\\
&\leq O(\exp(-\sqrt{\log n}))\,.
\end{aligned}
$$

Here inequality (a) follows from (C.45), (C.44), and (C.37); inequality (b) follows from Lemma C.16. This concludes the proof. □

APPENDIX D

Proofs in "Price Discovery and Efficiency in Waiting Lists: A Connection to Stochastic Gradient Descent"

Organization of the Mathematical Appendices. The appendix is organized as follows.

1. In Appendix D.1, we prove Proposition 4.1, which shows that the optimal allocative efficiency equals the value of the static allocation problem.
2. In Appendix D.2, we prove Theorem 4.1.
3. In Appendix D.3, we bound the price change granularity Δ using the property of the waiting cost functions.
4. In Appendix D.4, we establish the genericity of Assumption 4.1, and prove Theorem 4.2, which shows that for generic instances with finite agent types, the efficiency loss of the queueing mechanism is exponentially small in N as market size N increases.

D.1 Optimal Allocative Efficiency: Proof of Proposition 4.1

In this section we prove Proposition 4.1.

Proof of Proposition 4.1. We first show that $W^{\mathrm{OPT}} \leq W^*$. This part of the proof mostly consists of a careful treatment of expectations and limits.

Let $\eta \in H$ be any no-Ponzi allocation. Recall that $\eta_t \in \mathcal{J}_\emptyset$ is the kind of item assigned under η to the agent that arrived at epoch t. For each $j \in \mathcal{J}$ and $\theta \in \Theta$, define $\hat{G}_j^T(\theta)$ as

$$\hat{G}_j^T(\theta) \triangleq \frac{1}{A_T} \sum_{t \leq T} \xi_t \mathbb{1}_{\{\theta_t \leq \theta, \eta_t = j\}} .$$

Recall that A_T is the number of agents that arrived in the first T epochs. Therefore, $\hat{G}_j^T(\theta)$ is proportional to the empirical cumulative distribution function of the types of the agents in A_T who are assigned a type j item. When $A_T = 0$, we set $\hat{G}_j^T(\theta) = 0$ for all $j \in \mathcal{J}$ and $\theta \in \Theta$. By definition, the allocative efficiency under η is defined as

$$W(\eta) = \liminf_{T \to \infty} \frac{1}{A_T} \sum_{j \in \mathcal{J}} \sum_{t \leq T} \xi_t v(\theta_t, j) = \liminf_{T \to \infty} \sum_{j \in \mathcal{J}} \int_{\theta \in \Theta} v(\theta, j) d\hat{G}_j^T(\theta) . \tag{D.1}$$

Note that for any T, $\hat{G}_j^T(\theta)$ satisfies

$$\sum_{j \in \mathcal{J}} \hat{G}_j^T(\theta) \leq \frac{1}{A_T} \sum_{t \leq T} \xi_t \mathbb{1}_{\{\theta_t \leq \theta\}} , \qquad \forall \theta \in \Theta \tag{D.2}$$

$$\hat{G}_j^T(0) = 0 \, , \hat{G}_j^T(1) \leq \frac{1}{A_T} \left(\sum_{t \leq T} \mathbb{1}_{\{\eta_t = j\}} + M \right) \qquad \forall j \in \mathcal{J} \tag{D.3}$$

$$\hat{G}_j^T(\theta) \text{ is non-decreasing and right-continuous.} \qquad \forall j \in \mathcal{J} \tag{D.4}$$

for some $M \in \mathbb{R}$. Here (D.2) and (D.4) are trivial. (D.3) is satisfied by any no-Ponzi assignment for the following reason: The agents in A_T who are assigned a type $j \in \mathcal{J}$ item are either assigned before the T-th epoch or after the T-th epoch. The number of those who are assigned before T cannot exceed the the total number of type j items that arrive before T. The number of those who are assigned after T is bounded by some $M \in \mathbb{R}$ by the definition of no-Ponzi assignments.

Combining the above, we have

$$\mathbb{E} \left[\sum_{j \in \mathcal{J}} \int_{\theta \in \Theta} v(\theta, j) d\hat{G}_j^T(\theta) \right] \leq \mathbb{E} \left[\max_{\hat{G}_j(\theta) \text{ satisfying (D.2)(D.3)(D.4)}} \sum_{j \in \mathcal{J}} \int_{\theta \in \Theta} v(\theta, j) d\hat{G}_j(\theta) \right] .$$

It is easy to check that the optimal value of the inner maximization problem above is concave and non-decreasing in the RHS of (D.2) and (D.3). Note that

$$\begin{aligned}
&\text{Expectation of RHS of (D.2)} = F(\theta)\,, && \forall \theta \in \Theta\,,\\
&\text{Expectation of RHS of (D.3)} = \frac{\mu_j}{\lambda} + \frac{(1+\lambda)M}{\lambda T}\,, && \forall j \in \mathcal{J}\,.
\end{aligned}$$

It follows from Fatou's Lemma that

$$\mathbb{E}[W(\eta)] = \mathbb{E}\left[\liminf_{T\to\infty} \sum_{j\in\mathcal{J}} \int_{\theta\in\Theta} v(\theta, j) d\hat{G}_j^T(\theta)\right] \leq \liminf_{T\to\infty} \mathbb{E}\left[\sum_{j\in\mathcal{J}} \int_{\theta\in\Theta} v(\theta, j) d\hat{G}_j^T(\theta)\right] .$$

Applying Jensen's inequality, we have

$$\liminf_{T\to\infty} \mathbb{E}\left[\sum_{j\in\mathcal{J}} \int_{\theta\in\Theta} v(\theta, j) d\hat{G}_j^T(\theta)\right] \leq W^*\,,$$

where W^* is defined in (4.4). Therefore $\mathbb{E}[W(\eta)] \leq W^*$ for any $\eta \in H$. Since $W(\eta)$ is uniformly bounded above by $v_{\max}$, by Bounded Convergence Theorem we have $W^{\mathrm{OPT}} = \mathbb{E}[\sup_{\eta\in H} W(\eta)] = \sup_{\eta\in H} \mathbb{E}[W(\eta)] \leq W^*$. This concludes the proof.

Next we prove that $W^{\mathrm{OPT}} \geq W^*$. We explicitly construct a sequence of randomized policies that can achieve allocative efficiencies which are arbitrarily close to W^*. Note that the constructed policies are more of technical devices used to prove the desired bound, rather than practical policies.

Denote the optimal solution of the optimization problem (4.4) by $\mathbf{x}^*$. Consider the following randomized policy: Maintain a separate First-Come-First-Served queue for each item. An arriving agent will be assigned to one of the queues or rejected, based on a coin-toss (to be specified later). An agent who joins a queue will wait in that queue until receiving an item. When an item arrives, it is assigned to the agent at the head of its queue, if there is any; and the item is discarded if the item's queue is empty. The coin-toss is defined as follows: Fix $M \in \mathbb{Z}_+$. If the arriving agent is of type θ, it is assigned to queue j with probability $x^*_{\theta j}$, or rejected with probability $1 - \sum_{j\in\mathcal{J}} x^*_{\theta j}$. If the length of the queue to which the agent is assigned exceeds M, the agent is also rejected.

Denote the match value collected by the randomized policy in epoch t by v_t^{RD}. Then by definition of the policy, we have

$$\mathbb{E}[v_t^{\mathrm{RD}} | q_{j,t} < M, \forall j \in \mathcal{J}] = \frac{\lambda}{1+\lambda} \sum_{j \in \mathcal{J}} \int_\Theta x^*_{\theta j} \; v(\theta, j) dF(\theta) = \frac{\lambda}{1+\lambda} W^* \,.$$

It follows that

$$\mathbb{E}[v_t^{\mathrm{RD}}] \geq \frac{\lambda}{1+\lambda} W^* \cdot \mathbb{P}(q_{j,t} < M, \forall j \in \mathcal{J}) \,.$$

Let W^{RD} be the allocative efficiency of the randomized policy, we therefore have

$$W^{\mathrm{RD}} \geq W^* \cdot \mathbb{P}(q_{j,\infty} < M, \forall j \in \mathcal{J}) \,,$$

where $\mathbf{q}_\infty$ is the steady-state queue length distribution. The allocative efficiency loss of the randomized policy can be bounded as:

$$\begin{aligned} W^* - W^{\mathrm{RD}} &\leq W^* - W^* \cdot \mathbb{P}(q_{j,\infty} < M, \forall j \in \mathcal{J}) \\ &\leq v_{\max} \cdot \mathbb{P}(q_{j,\infty} = M, \exists j \in \mathcal{J}) \\ &\leq v_{\max} \cdot \sum_{j \in \mathcal{J}} \mathbb{P}(q_{j,\infty} = M) \,, \end{aligned} \tag{D.5}$$

where the second inequality follows from the fact that $W^* \leq v_{\max}$, and the last inequality comes from the union bound. It remains to bound $\mathbb{P}(q_{j,\infty} = M)$ for any $j \in \mathcal{J}$ under the randomized policy. Fix $j \in \mathcal{J}$, then $q_{j,t}$ is a birth-death process on $\{0, 1, \cdots, M\}$ with death rate μ_j and birth rate

$$\lambda_j^* = \; \lambda \int_\Theta x^*_{\theta j} dF(\theta) \,.$$

It follows from the constraint in (4.4) that $\lambda_j^* \leq \mu_j$. As a result, $\mathbb{P}(q_{j,\infty} = M) \leq \mathbb{P}(q_{j,\infty} = M-1) \leq \cdots \leq \mathbb{P}(q_{j,\infty} = 0)$, hence $\mathbb{P}(q_{j,\infty} = M) \leq \frac{1}{M+1}$. Plugging in the bound on $\mathbb{P}(q_{j,\infty} = M)$ to (D.5), we have

$$W^* - W^{\mathrm{RD}} \leq \frac{v_{\max} |\mathcal{J}|}{M+1} \,.$$

Notice that by definition, $W^{\mathrm{RD}} \leq W^{\mathrm{OPT}}$, hence

$$W^* - W^{\mathrm{OPT}} \leq \frac{v_{\max} |\mathcal{J}|}{M+1} \,.$$

Since M can be chosen arbitrarily, it must be true that $W^* - W^{\text{OPT}} \leq 0$. This concludes the proof. □

D.2 Upper Bound on the Allocative Efficiency Loss of the Queueing Mechanism

In this section we prove Theorem 4.1.

Proof of Theorem 4.1. The following proof generalizes the one described in the main paper which focuses on the special case of linear waiting cost.

Recall that $W_T(\eta^{\text{WL}})$ is the total value of items assigned to agents that arrive before epoch T, that is

$$W_T(\eta^{\text{WL}}) = \sum_{t=1}^{T} \xi_t \cdot v(\theta_t, a(\theta_t, \mathbf{q}_t)) \ ,$$

and that

$$W^{\text{WL}} = \frac{1+\lambda}{\lambda} \mathbb{E}\left[\liminf_{T\to\infty} \frac{W_T(\eta^{\text{WL}})}{T}\right] \ .$$

Similar to the proof in the main paper, we use the Lyapunov analysis to bound the allocative efficiency. It turns out for general waiting costs, using the Lyapunov function of *queue lengths* is notationally simpler than using the Lyapunov function of *waiting costs*. As a result, we use the following Lyapunov function: let the Lyapunov function $L(\mathbf{q})$ be such that $\nabla L(\mathbf{q}) = \mathbf{p}(\mathbf{q})$. The analysis uses the Bregman divergence generated by $L(\mathbf{q})$ as the notion of proximity, which is defined as follows

$$D_L(\mathbf{q}_1, \mathbf{q}_2) \triangleq L(\mathbf{q}_1) - L(\mathbf{q}_2) - \left\langle \nabla L(\mathbf{q}_2), \mathbf{q}_1 - \mathbf{q}_2 \right\rangle \ .$$

Let $\mathbf{a}_t$ and $\mathbf{d}_t$ be the vectors representing the arriving agent and item at time t, respectively:

$$\mathbf{a}_t \triangleq \mathbf{e}_{a(\theta_t, \mathbf{q}_t)} \xi_t \ , \qquad \mathbf{d}_t \triangleq \mathbf{e}_{j_t}(1-\xi_t) \ ,$$

and let $u_{j,t} \triangleq \max\{0, d_{j,t} - q_{j,t} - a_{j,t}\}$ denote the number of discarded items of type j at time t. The evolution of the length of queue j is governed by

$$q_{j,t+1} = [q_{j,t} + a_{j,t} - d_{j,t}]^+ = q_{j,t} + a_{j,t} - d_{j,t} + u_{j,t}, \quad \text{for each } j \in \mathcal{J}.$$

By Lemma D.2 we have that

$$\mathbb{E}[v(\theta_t, a(\theta_t, \mathbf{q}_t))\xi_t|\mathbf{q}_t] \geq \frac{\lambda}{1+\lambda}h(\mathbf{p}_t) - \big(L(\mathbf{q}_t) - \mathbb{E}[L(\mathbf{q}_{t+1}) \mid \mathbf{q}_t]\big) - \frac{\Delta}{2(1+\lambda)} - \mathbb{E}[D_L(\mathbf{q}_t + \mathbf{a}_t - \mathbf{d}_t, \mathbf{q}_t) \mid \mathbf{q}_t]. \tag{D.6}$$

By Lemma D.3 we have that $D_L(\mathbf{q}_t + \mathbf{a}_t - \mathbf{d}_t, \mathbf{q}_t) \leq \Delta/2$. By Lemma 4.1 we have that $h(\mathbf{p}_t) \geq W^*$. Together, we have that

$$\mathbb{E}[v(\theta_t, a(\theta_t, \mathbf{q}_t))\xi_t|\mathbf{q}_t] \geq \frac{\lambda}{1+\lambda}W^* - \big(L(\mathbf{q}_t) - \mathbb{E}[L(\mathbf{q}_{t+1}) \mid \mathbf{q}_t]\big) - \left(\frac{\Delta}{2(1+\lambda)} + \frac{\Delta}{2}\right).$$

Therefore, we have that

$$\begin{aligned}
\mathbb{E}\left[W_T(\eta^{\mathrm{WL}})\right] &= \mathbb{E}\left[\sum_{t=1}^{T} \xi_t \cdot v(\theta_t, a(\theta_t, \mathbf{q}_t))\right] \\
&= \mathbb{E}\left[\mathbb{E}\left[\sum_{t=1}^{T} \xi_t \cdot v(\theta_t, a(\theta_t, \mathbf{q}_t)) \mid \mathbf{q}_t\right]\right] \\
&= \mathbb{E}\left[\sum_{t=1}^{T} \mathbb{E}\left[\xi_t \cdot v(\theta_t, a(\theta_t, \mathbf{q}_t)) \mid \mathbf{q}_t\right]\right] \\
&\geq \mathbb{E}\left[\sum_{t=1}^{T} \frac{\lambda}{1+\lambda}W^* - \big(L(\mathbf{p}_t) - \mathbb{E}[L(\mathbf{q}_{t+1}) \mid \mathbf{q}_t]\big) - \frac{2+\lambda}{2(1+\lambda)}\Delta\right] \\
&= T\frac{\lambda}{1+\lambda}W^* - \big(L(\mathbf{p}_1) - \mathbb{E}[L(\mathbf{q}_{T+1})]\big) - T\frac{2+\lambda}{2(1+\lambda)}\Delta.
\end{aligned} \tag{D.7}$$

By Lemma 4.5, we have

$$W^{\mathrm{WL}} = \frac{1+\lambda}{\lambda} \lim_{T\to\infty} \frac{1}{T}\mathbb{E}\left[W_T(\eta^{\mathrm{WL}})\right].$$

Plugging in (D.7) to the above equality, we have

$$W^{\mathrm{WL}} \geq W^* - \frac{2+\lambda}{2\lambda}\Delta.$$

This concludes the proof. ∎

Lemma D.1. *We have that*

$$L(\mathbf{q}_{t+1}) \leq L(\mathbf{q}_t + \mathbf{a}_t - \mathbf{d}_t) + \frac{\Delta}{2} \cdot \mathbb{1}\{\xi_t = 0\}\,.$$

Proof. By definition of Bregman divergence, we have

$$\begin{aligned} D_L(\mathbf{q}_t + \mathbf{a}_t - \mathbf{d}_t, \mathbf{q}_{t+1}) &= L(\mathbf{q}_t + \mathbf{a}_t - \mathbf{d}_t) - L(\mathbf{q}_{t+1}) - \langle \mathbf{p}(\mathbf{q}_{t+1}), \mathbf{q}_t + \mathbf{a}_t - \mathbf{d}_t - \mathbf{q}_{t+1} \rangle \\ &= L(\mathbf{q}_t + \mathbf{a}_t - \mathbf{d}_t) - L(\mathbf{q}_{t+1}) + \langle \mathbf{p}(\mathbf{q}_{t+1}), \mathbf{u}_t \rangle\,. \end{aligned}$$

Therefore

$$L(\mathbf{q}_{t+1}) = L(\mathbf{q}_t + \mathbf{a}_t - \mathbf{d}_t) + \langle \mathbf{p}(\mathbf{q}_{t+1}), \mathbf{u}_t \rangle - D_L(\mathbf{q}_t + \mathbf{a}_t - \mathbf{d}_t, \mathbf{q}_{t+1})\,. \tag{D.8}$$

To bound the RHS of (D.8), we consider two cases:

Case 1. If $\exists j \in \mathcal{J}$ such that $d_{j,t} = 1$ and $q_{j,t} = 0$, we have $q_{j,t+1} = 0$ and $u_{j,t} = 1$. Note that in this case $\xi_t = 0$. Let $P_j(q)$ be an anti-derivative of $p_j(q)$, then $L(\mathbf{q}) = \sum_{j \in \mathcal{J}} P_j(q)$ is a Lyapunov function because it satisfies $\nabla L(\mathbf{q}) = \mathbf{p}(\mathbf{q})$. We have

$$\begin{aligned} &\langle \mathbf{p}(\mathbf{q}_{t+1}), \mathbf{u}_t \rangle - D_L(\mathbf{q}_t + \mathbf{a}_t - \mathbf{d}_t, \mathbf{q}_{t+1}) \\ &= p_j(0) - \big(P_j(-1) - P_j(0) - p_j(0) \cdot (-1)\big) \\ &= P_j(0) - P_j(-1)\,. \end{aligned} \tag{D.9}$$

Since $p_j(\cdot)$ is non-negative and Δ-Lipshitz, we have

$$P_j(0) \leq P_j(-1) + \int_0^1 \Delta \cdot x dx = P_j(-1) + \Delta/2\,.$$

Plugging in the above equality to (D.9), we have

$$\langle \mathbf{p}(\mathbf{q}_{t+1}), \mathbf{u}_t \rangle - D_L(\mathbf{q}_t + \mathbf{a}_t - \mathbf{d}_t, \mathbf{q}_{t+1}) \leq \frac{\Delta}{2} \cdot \mathbb{1}\{\xi_t = 0\}\,. \tag{D.10}$$

Case 2. If the condition in Case 1 does not hold, we have $\mathbf{u}_t = \mathbf{0}$ and $\mathbf{q}_t + \mathbf{a}_t - \mathbf{d}_t = \mathbf{q}_{t+1}$, hence

$$\langle \mathbf{p}(\mathbf{q}_{t+1}), \mathbf{u}_t \rangle - D_L(\mathbf{q}_t + \mathbf{a}_t - \mathbf{d}_t, \mathbf{q}_{t+1}) = 0\,.$$

Therefore, plugging in the above two cases to (D.8), we have

$$L(\mathbf{q}_{t+1}) \leq L(\mathbf{q}_t + \mathbf{a}_t - \mathbf{d}_t) + \frac{\Delta}{2} \cdot \mathbb{1}\{\xi_t = 0\}\,.$$

□

Lemma D.2. *For the model defined in Section 4.3, we have that*

$$\mathbb{E}[v(\theta_t, a(\theta_t, \mathbf{q}_t))\xi_t | \mathbf{q}_t] \geq \frac{\lambda}{1+\lambda} h(\mathbf{p}_t) - \big(L(\mathbf{q}_t) - \mathbb{E}[L(\mathbf{q}_{t+1}) \mid \mathbf{q}_t]\big) - \frac{\Delta}{2(1+\lambda)} - \mathbb{E}[D_L(\mathbf{q}_t + \mathbf{a}_t - \mathbf{d}_t, \mathbf{q}_t) \mid \mathbf{q}_t]\,.$$

Proof. We have that the drift of Lyapunov function $L(\mathbf{q})$ in one period is

$$\begin{aligned} & L(\mathbf{q}_t) - L(\mathbf{q}_{t+1}) \\ \geq\ & L(\mathbf{q}_t) - L(\mathbf{q}_t + \mathbf{a}_t - \mathbf{d}_t) - \frac{\Delta}{2} \cdot \mathbb{1}\{\xi_t = 0\} \\ =\ & - \langle \mathbf{p}(\mathbf{q}_t), \mathbf{a}_t - \mathbf{d}_t \rangle - D_L(\mathbf{q}_t + \mathbf{a}_t - \mathbf{d}_t, \mathbf{q}_t) - \frac{\Delta}{2} \cdot \mathbb{1}\{\xi_t = 0\}\,, \end{aligned} \tag{D.11}$$

where the inequality follows from Lemma D.1, and the equality comes from the definition of Bregman divergence.

We expanded the expected value of the next arrival $\mathbb{E}\Big[v(\theta_t, a(\theta_t, \mathbf{q}_t))\xi_t \mid \mathbf{q}_t\Big]$ plus the expected value of the term $\langle \mathbf{p}_t, \mathbf{a}_t - \mathbf{d}_t \rangle$ to show these are related to the dual objective given in (4.5).

$$\begin{aligned} & \mathbb{E}\Big[v(\theta_t, a(\theta_t, \mathbf{q}_t))\xi_t - \langle \mathbf{p}_t, \mathbf{a}_t - \mathbf{d}_t \rangle \mid \mathbf{q}_t\Big] \\ =\ & \mathbb{E}\Big[v(\theta_t, a(\theta_t, \mathbf{q}_t))\xi_t - \sum_{j \in \mathcal{J}} p_{j,t}(a_{j,t} - d_{j,t}) \mid \mathbf{q}_t\Big] \\ =\ & \mathbb{E}\Big[\max_{j \in \mathcal{J}_\emptyset} \big[v(\theta_t, j) - p_{j,t}\big]\, \xi_t + \sum_{j \in \mathcal{J}} p_{j,t} d_{j,t} \mid \mathbf{q}_t\Big] \\ =\ & \mathbb{E}[\xi_t]\mathbb{E}\Big[\max_{j \in \mathcal{J}_\emptyset} \big[v(\theta_t, j) - p_{j,t}\big] \mid \mathbf{q}_t\Big] + \mathbb{E}\Big[\sum_{j \in \mathcal{J}} p_{j,t} d_{j,t} \mid \mathbf{q}_t\Big] \\ =\ & \frac{\lambda}{1+\lambda} \int_\Theta \max_{j \in \mathcal{J}_\emptyset} \big[v(\theta_t, j) - p_{j,t}\big]\, dF(\theta) + \frac{1}{1+\lambda} \sum_{j \in \mathcal{J}} \mu_j p_{j,t} \\ =\ & \frac{\lambda}{1+\lambda} h(\mathbf{p}_t)\,. \end{aligned} \tag{D.12}$$

Adding $v(\theta_t, a(\theta_t, \mathbf{q}_t))\xi_t$ to both sides of equation (D.11), we have

$$
\begin{aligned}
& v(\theta_t, a(\theta_t, \mathbf{q}_t))\xi_t + L(\mathbf{q}_t) - L(\mathbf{q}_{t+1}) \\
\geq\ & v(\theta_t, a(\theta_t, \mathbf{q}_t))\xi_t - \langle \mathbf{p}(\mathbf{q}_t), \mathbf{a}_t - \mathbf{d}_t \rangle - D_L(\mathbf{q}_t + \mathbf{a}_t - \mathbf{d}_t, \mathbf{q}_t) - \frac{\Delta}{2} \cdot \mathbb{1}\{\xi_t = 0\}\,.
\end{aligned}
$$

Taking expectation conditional on $\mathbf{q}_t$ and applying equation (D.12), we have

$$
\begin{aligned}
& \mathbb{E}[v(\theta_t, a(\theta_t, \mathbf{q}_t))\xi_t | \mathbf{q}_t] + \big(L(\mathbf{q}_t) - \mathbb{E}[L(\mathbf{q}_{t+1}) \mid \mathbf{q}_t]\big) \\
\geq\ & \frac{\lambda}{1+\lambda} h(\mathbf{p}_t) - \mathbb{E}[D_L(\mathbf{q}_t + \mathbf{a}_t - \mathbf{d}_t, \mathbf{q}_t) \mid \mathbf{q}_t] - \frac{\Delta}{2(1+\lambda)}\,.
\end{aligned}
$$

Rearranging the terms, and we obtain the desired inequality. □

Lemma D.3. *For the model defined in Section 4.3, we have that for any* $\mathbf{q}_t$*,*

$$
\mathbb{E}[D_L(\mathbf{q}_t + \mathbf{a}_t - \mathbf{d}_t, \mathbf{q}_t) \mid \mathbf{q}_t] \leq \frac{\Delta}{2}\,.
$$

Proof. Note that $L(\mathbf{q})$ is convex because its gradient $\nabla L(\mathbf{q}) = \mathbf{p}(\mathbf{q})$ is increasing in each coordinate. Also note that $L(\mathbf{q})$ has Δ-Lipschitz gradient, because for queue lengths $\mathbf{q}_1, \mathbf{q}_2$,

$$
||\nabla L(\mathbf{q}_1) - \nabla L(\mathbf{q}_2)|| = ||\mathbf{p}(\mathbf{q}_1) - \mathbf{p}(\mathbf{q}_2)|| \leq \Delta ||\mathbf{q}_1 - \mathbf{q}_2||\,.
$$

Equivalently, $L(\mathbf{q})$ is Δ-strongly smooth, i.e.,

$$
L(\mathbf{q}_2) - L(\mathbf{q}_1) \leq \langle \nabla L(\mathbf{q}_1), \mathbf{q}_2 - \mathbf{q}_1 \rangle + \frac{\Delta}{2} ||\mathbf{p}(\mathbf{q}_2) - \mathbf{p}(\mathbf{q}_1)||^2\,.
$$

By definition of Bregman divergence, we have

$$
\begin{aligned}
& D_L(\mathbf{q}_t + \mathbf{a}_t - \mathbf{d}_t, \mathbf{q}_t) \\
=\ & L(\mathbf{q}_t + \mathbf{a}_t - \mathbf{d}_t) - L(\mathbf{q}_t) - \langle \nabla L(\mathbf{q}_t), \mathbf{a}_t - \mathbf{d}_t \rangle \\
\leq\ & \frac{\Delta}{2} ||\mathbf{a}_t - \mathbf{d}_t||^2 \\
=\ & \frac{\Delta}{2}\,.
\end{aligned}
$$

Here the second last inequality follows from the strong smoothness of $L(\mathbf{q})$. This concludes the proof. □

D.3 Price Change Granularity of Nonlinear Waiting Cost Functions

We stated our main result (Theorem 4.1) in terms of the price change granularity Δ. However, for nonlinear waiting costs, it remains to be shown how Δ is related to the waiting cost function $c(w)$. In this section, we focus on the waiting costs that satisfy the assumption below.

Assumption D.1. *We consider the following classes of waiting cost functions.*

- Convex waiting costs. *$c(w)$ is convex, twice-differentiable for $w \geq 0$, and that $c'(w)$ and $c''(w)$ are subexponential, i.e., there exists α such that $c'(w)$, $c''(w) \leq e^{\alpha w}$ for all $w \geq 0$.*

- Concave waiting costs. *$c(w)$ is concave and twice-differentiable for $w \geq 0$.*

Proposition D.1. *Consider the asymptotic regime in Corollary 4.2 and waiting cost functions satisfying Assumption D.1. The following holds:*

1. *For convex $c(w)$, there exists $\ell_0 < \infty$ such that for $\ell \geq \ell_0$, $\Delta \leq \frac{2c'(c^{-1}(v_{\max}))}{\ell\mu_{\min}}$.*

2. *For concave $c(w)$, for any $\ell > 0$, $\Delta \leq \frac{c'(0)}{\ell\mu_{\min}}$.*

Proof of Proposition D.1. Consider the system with index ℓ. Let X_t be the interarrival time between the t-th type j item and the $(t+1)$-th type j item, hence X_t is an exponential random variable with rate $\ell\mu_j$, and $\{X_t\}_{t=1}^{\infty}$ are i.i.d. Let $S_n \triangleq \sum_{t=1}^{n} X_t$.

Let $q_{\max,\ell}$ be the threshold queue length above which no arriving agent will join that queue. Then approximately

$$v_{\max} = p_j(q_{\max,\ell}) = \mathbb{E}[c(S_{q_{\max,\ell}})] .$$

For convex cost function $c(w)$, by Jensen's inequality we have

$$\mathbb{E}[c(S_{q_{\max,\ell}})] \geq c\left(\mathbb{E}[S_{q_{\max,\ell}}]\right) = c\left(\frac{q_{\max,\ell}}{\ell\mu_j}\right) .$$

Compare the above two inequalities, we have $q_{\max,\ell} \le \ell\mu_j c^{-1}(v_{\max})$. Notice that

$$p_j(q_j+1) - p_j(q_j) = \mathbb{E}[c(S_{q_j} + X_{q_j+1}) - c(S_{q_j})] \le \mathbb{E}[c'(S_{q_j} + X_{q_j+1}) \cdot X_{q_j+1}]\,,$$

where the inequality follows from the convexity of $c(w)$. Take supremum over all $0 \le q_j \le q_{\max,\ell}$ on both sides of the above inequality. Because $c(w)$ is convex, $c'(w)$ must be non-decreasing, hence

$$\Delta = \sup_{0\le q_j \le q_{\max,\ell}} \big(p_j(q_j+1) - p_j(q_j)\big) \le \mathbb{E}[c'(S_{q_{\max,\ell}} + X_{q_{\max,\ell}+1}) \cdot X_{q_{\max,\ell}+1}]\,.$$

Using Holder's inequality, we have

$$\mathbb{E}[c'(S_{q_{\max,\ell}} + X_{q_{\max,\ell}+1}) \cdot X_{q_{\max,\ell}+1}] \le \sqrt[\alpha]{\mathbb{E}[(c'(S_{q_{\max,\ell}+1}))^{\alpha}]} \cdot \sqrt[\beta]{\mathbb{E}[X_1^{\beta}]}\,,$$

where $\alpha, \beta \in (1,\infty)$ and $\frac{1}{\alpha} + \frac{1}{\beta} = 1$. Because $c(w)$ satisfies Assumption D.1, for any $\alpha \in (1,\infty)$ we can apply Lemma D.4 and it follows that

$$\lim_{\ell\to\infty} \sqrt[\alpha]{\mathbb{E}[(c'(S_{q_{\max,\ell}}))^{\alpha}]} = \sqrt[\alpha]{c'\left(\mathbb{E}[S_{q_{\max,\ell}}]\right)^{\alpha}} = c'\left(\frac{q_{\max,\ell}}{\ell\mu_j}\right).$$

Therefore

$$\lim_{\ell\to\infty} \mathbb{E}[c'(S_{q_{\max,\ell}} + X_{q_{\max,\ell}+1}) \cdot X_{q_{\max,\ell}+1}] \le c'\left(\frac{q_{\max,\ell}}{\ell\mu_j}\right) \inf_{\beta>1} \sqrt[\beta]{\mathbb{E}[X_1^{\beta}]}$$
$$\le c'\left(\frac{q_{\max,\ell}}{\ell\mu_j}\right)\mathbb{E}[X_1] = \frac{c'\left(\frac{q_{\max,\ell}}{\ell\mu_j}\right)}{\ell\mu_j}.$$

As a result, there exists $\ell_0 > 0$ such that for $\ell \ge \ell_0$, it holds that

$$\Delta \le \frac{2c'\left(\frac{q_{\max,\ell}}{\ell\mu_j}\right)}{\ell\mu_j} \le \frac{2c'\left(c^{-1}(v_{\max})\right)}{\ell\mu_j}\,. \tag{D.13}$$

For concave $c(w)$, using its concavity we have for any $0 \le q_j \le q_{\max,\ell}$,

$$p_j(q_j+1) - p_j(q_j) = \mathbb{E}[c(S_{q_j} + X_{q_j+1}) - c(S_{q_j})] \le \mathbb{E}[c'(0)X_1] = \frac{c'(0)}{\ell\mu_j}\,. \tag{D.14}$$

Combine (D.13) and (D.14), we conclude the proof. □

Lemma D.4. *Let $\{X_i\}_{i=1}^{\infty}$ be i.i.d. exponential random variables with rate 1, $\bar{X}_n \triangleq \frac{1}{n}\sum_{i=1}^{n} X_i$, and $\alpha \in (0,\infty)$. For any continuously differentiable function $f(x)$ defined* *on $\mathbb{R}^+$ such that there exists $C_1, C_2 \in (0, \infty)$ and $f(x) \le C_1 e^{C_2 x}$, $f'(x) \le C_1 e^{C_2 x}$ for all $x \in \mathbb{R}^+$, we have*

$$\lim_{N\to\infty} \mathbb{E}\left[f(\alpha \cdot \bar{X}_N)\right] = f(\alpha)\,.$$

Proof of Lemma D.4. The result simply follows the proof of Theorem 1(c) in [134], therefore we omit the details. □

D.4 Proof of Exponentially Small Loss

In this section, we prove the results in Section 4.5.

D.4.1 Rate of price adjustment

Proof of Lemma 4.7. We proceed in two steps.

Step 1. We first show that we can lower bound $h(\mathbf{p}) - h(\mathbf{p}^*)$ by a support function:

$$h(\mathbf{p}) - h(\mathbf{p}^*) \ge \sup_{s\in\mathcal{S}} \langle \mathbf{p}^* - \mathbf{p}, \mathbf{s}\rangle\,.$$

for some convex set $\mathcal{S}$.

For each agent type $\theta \in \Theta$, define

$$\Delta_\theta \triangleq \left\{\mathbf{x} \in \mathbb{R}_+^{|\mathcal{J}|} : \sum_{j\in\mathcal{J}} x_j = 1\,, x_j = 0 \text{ for } j \notin \mathcal{J}_\theta^*\right\}\,,$$
$$\tilde{\Delta}_\theta \triangleq \left\{\mathbf{x} \in \mathbb{R}_+^{|\mathcal{J}|} : \sum_{j\in\mathcal{J}} x_j \le 1\,, x_j = 0 \text{ for } j \notin \mathcal{J}_\theta^*\right\}\,,$$

$$\Delta \triangleq \left\{\mathbf{x} \in \mathbb{R}_+^{|\mathcal{J}|} : \sum_{j\in\mathcal{J}} x_j = 1\right\}\,, \quad \tilde{\Delta} \triangleq \left\{\mathbf{x} \in \mathbb{R}_+^{|\mathcal{J}|} : \sum_{j\in\mathcal{J}} x_j \le 1\right\}\,.$$

Using the definitions above, we can rewrite the dual function 4.5 as

$$h(\mathbf{p}) = \frac{1}{\lambda}\sum_{\theta\in\Theta}\lambda_\theta\left(\max_{\mathbf{x}_\theta\in\tilde{\Delta}}\sum_{j\in\mathcal{J}}(v(\theta,j) - p_j)x_{\theta,j}\right) + \frac{1}{\lambda}\sum_{j\in\mathcal{J}}\mu_j p_j\,.$$

Let $\mathbf{x}_\theta^*$ be a maximizer of the inner maximization problem above. Define $\mathbf{s} \in \mathbb{R}^{|\mathcal{J}|}$ where

$$s_j = \sum_{\theta \in \Theta} \lambda_\theta \cdot x_{\theta,j}^* - \mu_j \,,$$

then it is easy to see that $-\frac{1}{\lambda}\mathbf{s}$ is a subgradient of $h(\mathbf{p})$ at $\mathbf{p}$, denoted by $-\frac{1}{\lambda}\mathbf{s} \in \partial h(\mathbf{p})$.

For $\theta \in \Theta^*$, let

$$\mathbf{x}'_\theta \triangleq \text{argmax}_{\mathbf{x}_\theta \in \Delta_\theta} \sum_{j \in \mathcal{J}} (p_j^* - p_j) x_{\theta,j} \,,$$

for $\theta \in \Theta \backslash \Theta^*$, let

$$\mathbf{x}'_\theta \triangleq \text{argmax}_{\mathbf{x}_\theta \in \tilde{\Delta}_\theta} \sum_{j \in \mathcal{J}} (p_j^* - p_j) x_{\theta,j} \,.$$

The interpretation of $\mathbf{x}'_\theta$ is as follows. Consider a type θ agent. If the current price is exactly $\mathbf{p}^*$, then the agent is indifferent between the items in $\mathcal{J}_\theta^*$, and strictly prefers these items to other items. If the price deviates a little from $\mathbf{p}^*$: (1) if $\theta \in \Theta^*$, the agent will prefer the item in $\mathcal{J}_\theta^*$ that is the cheapest; (2) if $\theta \notin \Theta^*$, the agent's optimal utility is zero, hence she will choose an item in $\mathcal{J}_\theta^*$ that is the cheapest only if the price is lower than the optimal price, otherwise she will not choose any item. $\mathbf{x}'_\theta$ characterizes the choice of an agent when $\mathbf{p}$ is sufficiently close to $\mathbf{p}^*$.

A key observation is that when $\mathbf{p} \in \mathcal{P}$, the above observation for "sufficiently close" $\mathbf{p}$ holds. Therefore for $\mathbf{s}' \in \mathbb{R}^{|\mathcal{J}|}$ defined as

$$s'_j \triangleq \sum_{\theta \in \Theta} \lambda_\theta \cdot x'_{\theta,j} - \mu_j \,,$$

we have $-\frac{1}{\lambda}\mathbf{s}' \in \partial h(\mathbf{p})$, and that for $\mathbf{p} \in \mathcal{P}^*$,

$$\begin{aligned}
\lambda h(\mathbf{p}) = & \sum_{\theta \in \Theta} \lambda_\theta \sum_{j \in \mathcal{J}} (v(\theta, j) - p_j) x'_{\theta,j} + \sum_{j \in \mathcal{J}} \mu_j p_j \\
= & \sum_{\theta \in \Theta} \lambda_\theta \sum_{j \in \mathcal{J}} (p_j^* - p_j) x'_{\theta,j} + \sum_{j \in \mathcal{J}} \mu_j (p_j - p_j^*) \\
& + \sum_{\theta \in \Theta} \lambda_\theta \sum_{j \in \mathcal{J}} (v(\theta, j) - p_j^*) x'_{\theta,j} + \sum_{j \in \mathcal{J}} \mu_j p_j^*
\end{aligned}$$

Note that the sum of the terms in the last row is exactly $\lambda h(\mathbf{p}^*)$. This is because for agent of type $\theta \in \Theta^*$, under price $\mathbf{p}$ she must choose an item from $\mathcal{J}_\theta^*$, hence

$$\sum_{j\in\mathcal{J}}(v(\theta,j)-p_j^*)x'_{\theta,j} = \max_{j\in\mathcal{J}}\,(v(\theta,j)-p_j^*)\,,$$

whereas for agent of type $\theta \notin \Theta^*$, she either chooses an item from $\mathcal{J}_\theta^*$, or she balks. Hence

$$\sum_{j\in\mathcal{J}}(v(\theta,j)-p_j^*)x'_{\theta,j} = 0 = \max_{j\in\mathcal{J}}\,(v(\theta,j)-p_j^*)\,.$$

Thus we have

$$\begin{aligned}\lambda(h(\mathbf{p})-h(\mathbf{p}^*)) = &\sum_{\theta\in\Theta^*}\lambda_\theta\left(\max_{\mathbf{x}_\theta\in\Delta_\theta}\sum_{j\in\mathcal{J}}(p_j^*-p_j)x_{\theta,j}\right)+\sum_{\theta\in\Theta\backslash\Theta^*}\lambda_\theta\left(\max_{\mathbf{x}_\theta\in\tilde{\Delta}_\theta}\sum_{j\in\mathcal{J}}(p_j^*-p_j)x_{\theta,j}\right)\\ &+\sum_{j\in\mathcal{J}}\mu_j(p_j-p_j^*)\,. \qquad \text{(D.15)}\end{aligned}$$

Define the *rate region* $\mathcal{S}$ as:

$$\mathcal{S} \triangleq \left\{\sum_{\theta\in\Theta}\lambda_\theta\mathbf{x}_\theta-\boldsymbol{\mu} : \mathbf{x}_\theta\in\Delta_\theta \text{ for } \theta\in\Theta^*\,, \mathbf{x}_\theta\in\tilde{\Delta}_\theta \text{ for } \theta\in\Theta\backslash\Theta^*\right\}\,,$$

which is the set of possible rates of change of dual prices when $\mathbf{p}\in\mathcal{P}$. Therefore we can rewrite the RHS of (D.15) as

$$\sup_{\mathbf{s}\in\mathcal{S}}\,\langle\mathbf{p}^*-\mathbf{p},\mathbf{s}\rangle\,.$$

Using the fact that $h(\mathbf{p})$ is convex, we have for any $\mathbf{p}$,

$$h(\mathbf{p})-h(\mathbf{p}^*) \geq \frac{1}{\lambda}\sup_{\mathbf{s}\in\mathcal{S}}\,\langle\mathbf{p}^*-\mathbf{p},\mathbf{s}\rangle\,.$$

This concludes step 1.

Step 2. Characterizing the set $\mathcal{S}$. Note that $\mathcal{S}$ is the Minkowski sum of simplices shifted by $\boldsymbol{\mu}$, which is known as the generalized permutohedron [see, e.g., 135]. Using Proposition 6.3 from [135], we have the following defining inequalities of $\mathcal{S}$:

$$\begin{aligned}&\sum_{j\in\mathcal{I}}s_j \leq \lambda-\sum_{j\in\mathcal{I}}\mu_j\,, && \forall\mathcal{I}\subset\mathcal{J}\,,\\ &\sum_{j\in\mathcal{I}}s_j \geq \sum_{\theta\in\Theta^*}\lambda_\theta-\sum_{j\in\mathcal{I}}\mu_j\,, && \forall\mathcal{I}:\mathcal{I}\subset\mathcal{J}\,,\mathcal{I}\supset\cup_{\theta\in\Theta^*}\mathcal{J}_\theta^*\,,\end{aligned}$$

$$s_j \geq \; -\mu_j \, , \qquad\qquad \forall j \in \mathcal{J} \, .$$

We first argue that there exists $\epsilon > 0$ such that the ball $B(\mathbf{0}, \epsilon)$ is contained in $\mathcal{S}$. This can be proved by contradiction: if it is not true, then using (D.15), we can show that the minimizer of $h(\mathbf{p})$ is non-unique, leading to contradiction with Assumption 4.1. Note that this already leads to a lower bound of $h(\mathbf{p}) - h(\mathbf{p}^*)$: we have $\epsilon \frac{\mathbf{p}^* - \mathbf{p}}{||\mathbf{p}^* - \mathbf{p}||_2} \subset \mathcal{S}$, hence $\lambda(h(\mathbf{p}) - h(\mathbf{p}^*)) \geq \epsilon ||\mathbf{p}^* - \mathbf{p}||_2$.

It remains to quantitatively characterize ϵ. To simplify the notation, we consider the centered version of $\mathbf{p}$, defined as $\tilde{\mathbf{p}} \triangleq \mathbf{p}^* - \mathbf{p}$; let $\tilde{h}(\tilde{\mathbf{p}}) \triangleq h(\mathbf{p}) - h(\mathbf{p}^*)$.

Since $\mathcal{S}$ is defined "locally" (i.e., for $\mathbf{p} \in \mathcal{P}$), all the arguments below assume that $\mathbf{p} \in \mathcal{P}$. We have derived that $\tilde{h}(\tilde{\mathbf{p}}) = \sup_{\mathbf{s} \in \mathcal{S}} \langle \tilde{\mathbf{p}}, \mathbf{s} \rangle$. Define the level sets of $\tilde{h}(\tilde{\mathbf{p}})$:

$$\mathcal{L} \triangleq \left\{ \tilde{\mathbf{p}} \in \mathbb{R}^{|\mathcal{J}|} : \tilde{p}_j \leq 0 \text{ for } j \neq \mathcal{J}^*, \tilde{h}(\tilde{\mathbf{p}}) \leq 1 \right\} .$$

Here the constraints $\tilde{p}_j \leq 0$ for $j \neq \mathcal{J}^*$ come from the fact that $\mathbf{p} \geq \mathbf{0}$. Using the theory of polar duality [see, e.g., 136], since the ball $B(\mathbf{0}, \epsilon)$ is contained in $\mathcal{S}$, we have that

$$\left(\mathcal{S} \cap \{ \tilde{\mathbf{p}} : \tilde{p}_j \leq 0 \text{ for } j \neq \mathcal{J}^* \} \right)^* \subset \left(B(\mathbf{0}, \epsilon) \cap \{ \tilde{\mathbf{p}} : \tilde{p}_j \leq 0 \text{ for } j \neq \mathcal{J}^* \} \right)^* .$$

Here the asterisk outside stands for polar set. Denote the LHS set as $\mathcal{B}^*$, the RHS set as $\mathcal{L}^*$. We have

$$\begin{aligned} \mathcal{B}^* &= \left\{ \mathbf{s} + \sum_{j \notin \mathcal{J}^*} \gamma_j \mathbf{e}_j : ||\mathbf{s}||_2 \leq \epsilon \, , \gamma_j \geq 0 \right\} , \\ \mathcal{L}^* &= \left\{ \mathbf{s} + \sum_{j \notin \mathcal{J}^*} \gamma_j \mathbf{e}_j : \mathbf{s} \in \mathcal{S} \, , \gamma_j \geq 0 \right\} . \end{aligned}$$

Because $\mathcal{B}^* \subset \mathcal{L}^*$, ϵ can take value up to the inradius of $\mathcal{S}$, which is larger than the minimum of the distances between $\mathbf{0}$ and the defining hyperplanes of $\mathcal{S}$. It follows that

$$\epsilon \geq \left\{ \min_{\mathcal{I} \subset \mathcal{J}} \frac{\lambda - \sum_{j \in \mathcal{I}} \mu_j}{\sqrt{|\mathcal{I}|}} \, , \min_{\{\mathcal{I} : \mathcal{I} \subset \mathcal{J} \, , \mathcal{I} \supset \cup_{\theta \in \Theta^*} \mathcal{J}_\theta^*\}} \frac{\sum_{j \in \mathcal{I}} \mu_j - \sum_{\theta \in \Theta^*} \lambda_\theta}{\sqrt{|\mathcal{I}|}} \, , \min_{j \in \mathcal{J}} \mu_j \right\} .$$

□

D.4.2 Proof of Theorem 4.2

Building on the observations obtained in the last two sections, we first prove a lemma that establishes the exponential concentration of $\mathbf{p}$ around $\mathbf{p}^*$, then complete the proof of Theorem 4.2.

Lemma D.5 (Concentration of dual prices)**.** *Suppose Assumption 4.1 holds. Then for any* $c \leq \frac{\mu_{\min}\delta\gamma}{36}$*, we have*

$$\mathbb{P}(\mathbf{p}_\infty \notin \mathcal{P}) \leq \exp\left(-\log\left(1+\frac{\gamma\mu_{\min}}{4}\right)\left(\frac{\delta}{12\Delta}\right)\right).$$

Proof. We prove the result using the Lyapunov functions:

$$\bar{L}(\mathbf{p}) = \frac{1}{2}\sum_{j\in\mathcal{J}}\mu_j(p_j-p_j^*)^2 \qquad \text{and} \qquad V(\mathbf{p}) \triangleq \sqrt{L(\mathbf{p})}\,.$$

Using (4.12), we have

$$\mathbb{E}[\bar{L}(\mathbf{p}_{t+1})|\mathbf{q}_t] - \bar{L}(\mathbf{p}_t) \leq -c\left(h(\mathbf{p}_t)-h(\mathbf{p}^*)\right) + c\Delta\,.$$

Plugging in the result in Lemma 4.7, we have

$$\mathbb{E}[\bar{L}(\mathbf{p}_{t+1})|\mathbf{q}_t] - \bar{L}(\mathbf{p}_t) \leq -c\gamma||\mathbf{p}_t-\mathbf{p}^*||_2 + c\Delta\,. \tag{D.16}$$

Use the fact that $f(x)=\sqrt{x}$ is concave for $x \geq 0$ so that for $y > x > 0$, $f(y)-f(x) \leq (y-x)f'(x) = \frac{y-x}{2\sqrt{x}}$, we have

$$V(\mathbf{p}_{t+1}) - V(\mathbf{p}_t) \leq \frac{L(\mathbf{p}_{t+1})-L(\mathbf{p}_t)}{2V(\mathbf{p}_t)}\,. \tag{D.17}$$

Take conditional expectation given $\mathbf{q}_t$ on both sides of (D.17) and plug in (D.16), we have current progress, mind the constants.

$$\begin{aligned}\mathbb{E}[V(\mathbf{p}_{t+1})|\mathbf{q}_t] - V(\mathbf{p}_t) &\leq \frac{\mathbb{E}[L(\mathbf{p}_{t+1})|\mathbf{q}_t]-L(\mathbf{p}_t)}{2V(\mathbf{p}_t)}\\ &\leq -\frac{c\gamma}{2} + \frac{c^2}{2\mu_{\min}V(\mathbf{p}_t)}\,.\end{aligned}$$

Now we use a concentration bound from [57] to prove the desired result. Let $\mathbf{p}_\infty$ be the steady-state distribution of the prices. For $\mathbf{p}_t$ such that

$$V(\mathbf{p}_t) \geq \frac{2c}{\mu_{\min}\gamma}\,,$$

we have

$$\mathbb{E}[V(\mathbf{p}_{t+1})|\mathbf{q}_t] - V(\mathbf{p}_t) \leq \ -\frac{c\gamma}{4}\,.$$

As a result, $V(\cdot)$ is a Lyapunov function with exception parameter $\frac{2c}{\mu_{\min}\gamma}$ and negative drift $\frac{c\gamma}{4}$. Note that in each step, the Lyapunov function can increase by at most $\frac{c}{\mu_{\min}}$. Using Theorem 1 in [57], we have for any $r = 0, 1, \cdots,$

$$\mathbb{P}\left(V(\mathbf{p}_\infty) > \frac{2c}{\mu_{\min}\gamma} + 2r\frac{c}{\mu_{\min}}\right) \leq \left(\frac{\frac{c}{\mu_{\min}}}{\frac{c}{\mu_{\min}} + \frac{c\gamma}{4}}\right)^{r+1} = \left(\frac{1}{1 + \frac{\gamma\mu_{\min}}{4}}\right)^{r+1}.$$

Note that

$$\left\{\mathbf{p} : V(\mathbf{p}) \leq \frac{\delta}{3}\right\} \subset \left\{\mathbf{p} : ||\mathbf{p} - \mathbf{p}^*||_\infty \leq \frac{\delta}{2}\right\}.$$

As a result, for $c \leq \frac{\mu_{\min}\delta\gamma}{36}$, plugging in $r = \frac{\mu_{\min}\delta}{12c}$, we have

$$\begin{aligned}\mathbb{P}(\mathbf{p}_\infty \notin \mathcal{P}) \leq \mathbb{P}\left(V(\mathbf{p}_\infty) > \frac{\delta}{3}\right) &\leq \left(\frac{1}{1 + \frac{\gamma\mu_{\min}}{4}}\right)^{\frac{\mu_{\min}\delta}{12c}} \\ &= \exp\left(-\log\left(1 + \frac{\gamma\mu_{\min}}{4}\right)\left(\frac{\delta}{12\Delta}\right)\right).\end{aligned}$$

□