In Which Matching Markets do Costly Compatibility Inspections Lead to a Deadlock?*

Nicole Immorlica†, Yash Kanoria‡, Jiaqi Lu§

Abstract

A key feature of many real-world matching markets is congestion, i.e., market participants struggle to find match partners. We characterize congestion in a model of random matching markets where an agent pair must perform a mutual inspection to verify compatibility prior to matching with each other. Motivated by the notion of regret-free stability, we assume agents are only willing to inspect their current favorite agent and will do so only if, upon a successful inspection, that match is guaranteed. We ask when, in large random two-sided markets, will *information deadlocks* arise in which many agents delay inspections indefinitely awaiting a match guarantee. The market consists of N women and αN men. We characterize the existence and size of information deadlock as a function of the men-to-women ratio α, women's average size K of the consideration set, and an inspection's success probability p, as the number of women N grows. We find a phase transition from a deadlock-free regime (where a vanishingly small fraction of agents are stuck waiting) to the information deadlock regime as we increase K, decrease α or decrease p. A number of market design insights emerge from our characterization, for example, the market connectivity K which maximizes the number of matches formed is that which causes the market to be at the phase boundary between the deadlock-free regime and the deadlock regime. Vertical differentiation between agents reduces deadlock, as does a willingness by agents to perform parallel inspections. Our analysis is inspired by the machinery of message passing and density evolution from statistical physics, and the emergence of deadlock corresponds to a certain branching process being supercritical.

Keywords: matching market; search frictions; information deadlock; message passing algorithm; phase transition.

*We provide an e-companion for this paper at https://tinyurl.com/infodeadlock. A preliminary version of this paper appeared in the WINE 2021 conference as an extended abstract. Y. Kanoria and J. Lu gratefully acknowledge the support of the National Science Foundation via grant CMMI-1653477. J. Lu also gratefully acknowledges financial support from the National Natural Science Foundation of China (72401245, 72192805); and the Guangdong Key Lab of Mathematical Foundations for Artificial Intelligence.

†Microsoft Research, New York — nicimm@gmail.com

‡Columbia Business School, Columbia University — ykanoria@gsb.columbia.edu

§The Chinese University of Hong Kong (Shenzhen) — lujiaqi@cuhk.edu.cn

1 Introduction

In matching markets, agents seek out one-another, looking to form a relationship for mutual benefit. Men and women seek out romantic partners in marriage markets, workers and firms look to fill job vacancies in labor markets, landlords and prospective tenants look to fill vacant apartments in housing markets. However, in several matching markets, many participants fail to find good matches, and even when they do, the process can take a very long time. This phenomenon is often referred to as market *congestion* [48]. For example, according to the General Social Survey and Pew Research Center [11], in 2019 more than half of young people (ages 18 to 34) in the US did not have a steady romantic partner, and among those who were seeking relationships, 36% had been on no date in the previous three months. From 2006 to 2018, the top 50 US colleges' average acceptance rate plummeted from 35.9% to 22.6%. This induced more applications per person, which caused longer waiting lists and further delays in the admission process, making it increasingly difficult for colleges to predict their yield.

A key factor contributing to market congestion is informational frictions [4, 28]. Even if participants know their preference ordering over potential matches – e.g., for Alice an MIT education is clearly more desirable than a CalTech one – they often need to gather information to determine the feasibility of the match. Maybe MIT is prohibitively expensive, or Boston is prohibitively cold. This information-gathering exercise is typically costly, making participants reluctant to gather information about potentially infeasible partners. In the extreme, participants may be unwilling to investigate the feasibility of a match unless it is guaranteed should they reach a positive conclusion. This unwillingness can result in congestion, as illustrated by the following (stylized) example: Each of Caltech and MIT can accommodate one of Alice and Bob, but not both. Alice is admitted to Caltech but prefers MIT and is waiting for an MIT admission decision prior to inspecting either university, while Bob similarly is admitted to MIT but prefers Caltech and is awaiting a Caltech admission decision prior to inspecting either university.

We call this type of congestion an *information deadlock*. The concept of information deadlocks was recently formalized by Immorlica, Leshno, Lo and Lucier [30] in conjunction with a new solution concept for incomplete-information matching markets called *regret-free stability*. In a regret-free stable outcome, participants collect information as if they were "last to market," i.e., as if all other uncertainty in the market were already resolved and hence they know their options before they perform inspections. In the above example, it would be a regret-free stable outcome if Alice successfully inspects and is matched to CalTech and Bob successfully inspects and is matched to MIT because in this case MIT was never an option for Alice. It would also be regret-free stable for Alice to inspect MIT if Bob's MIT inspection were unsuccessful. In fact, if Bob's MIT inspection is unsuccessful, Alice *must* inspect MIT in any stable outcome. However, it would not be regret-free stable for Alice to inspect MIT if Bob's inspection were successful – she would regret doing so because Bob had higher priority at MIT and so she never had a chance at an MIT admission. So whether Alice can inspect MIT depends on Bob's inspection results. The situation is, of course, symmetric, creating an information deadlock in which each participant is waiting for the other to perform an inspection before doing so her/himself; see Figure 1. Immorlica et al show that regret-free stable outcomes always exist, but markets may be unable to find one due to these information deadlocks. That is, there is no sequence of inspections and subsequent assignment (called a *communication process* in Immorlica et al) that guarantees the outcome is regret-free stable for any underlying economy. They show markets with deadlocks arising due to short cycles like the one above in the example, and leave unresolved several important questions. In particular, how prevalent are deadlocks as function of market characteristics? Are deadlocks primarily driven by short cycles? In this paper, we quantitatively characterize the prevalence of deadlock in random matching markets as a function of market characteristics. Even though random markets lack short cycles, we find large deadlocks in a range of markets.

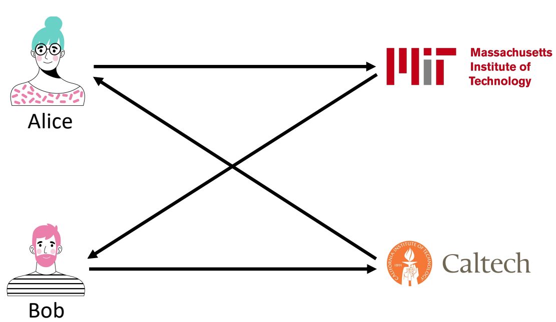

Figure 1: Stylized representation of a deadlock in school admission. Directed edges show with whom agents want to inspect next, e.g., Alice wants to inspect MIT, while MIT hopes to inspect Bob.

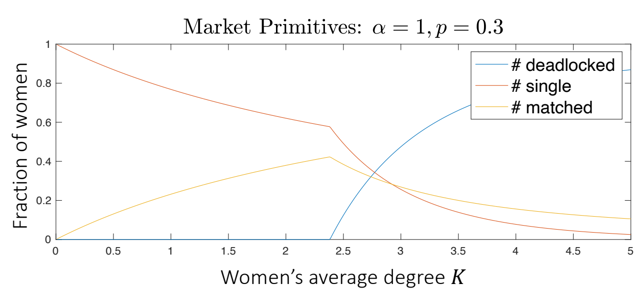

Figure 2: Fraction of women in information deadlock (deadlock), fraction of women exhausting consideration set without finding a match (single), and fraction of women matched (match) as a function of the women's average degree. We fix men-to-women ratio $\alpha = 1$ and inspection success probability $p = 0.3$.

Model. We consider a two-sided matching market with N women and αN men. Agent *considera-*

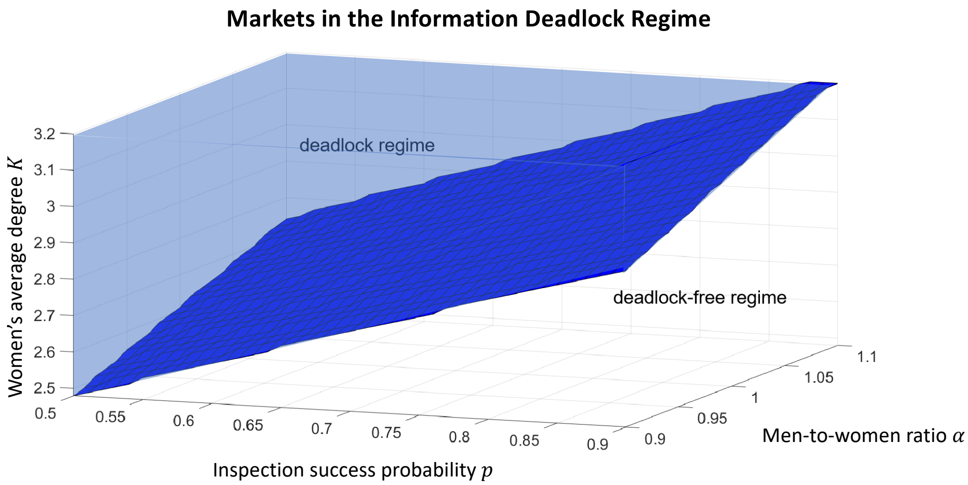

Figure 3: Phase transition from the deadlock-free regime and the deadlock regime. The area above [below] the plane are market primitives in the deadlock [deadlock-free] regime.

tion sets are captured by an undirected consideration graph, where each man [woman] is interested in his [her] neighbors under a preference ordering. In order for a match to occur, however, the pair must perform a costly inspection to determine if they are actually compatible. We consider *satisficing search* [52], where agents search sequentially according to their preference order until a successful inspection is found. We assume that an agent only inspects if the match is *guaranteed* once the inspection succeeds (the search behavior is termed *partner search* in this paper). We focus attention on random markets consisting of N women and $M = \alpha N$ men where each man is in a woman's consideration set independently with probability K/M, hence women have an average degree K, whereas men have average degree K/α. Men and women's preference lists are independent uniformly random orders. Each inspection succeeds independently with probability p.

Main results. We find that despite the absence of short cycles in our random markets, a broad range of markets suffer from information deadlock involving many agents. We find a sizeable set Θ of market primitives such that holding fixed $(\alpha, K, p) \in \Theta$ and growing the size of the market N (and hence M), $\Omega(N)$ agents are stuck in an information deadlock (up to at least $o(\log N)$ time) with high probability; see Figure 3. We characterize that fraction of agents on each side of the market in Theorem 1. For markets with $(\alpha, K, p) \notin \Theta$, with high probability no information deadlock occurs (see Figure 2). Thus, iteratively occurring guaranteed inspections suffice to clear the market in the deadlock-free regime $\{(\alpha, K, p) \notin \Theta\}$, but fail to in the information deadlock regime $\{(\alpha, K, p) \in \Theta\}$. For the special case $p = 0$, we show that in the information deadlock regime $\{(\alpha, K, p) \in \Theta\}$, deadlock persists *forever* (with $\Omega(1)$ probability); this is formalized in Theorem 2. Our main theorems, together with numerical evaluation of our characterization, shows that there is a sharp phase transition between the deadlock-free regime and the deadlock regime, in the market primitives (see Figure 3). For example, when fixing the men-to-women ratio α and the inspection success probability p, deadlock occurs if the average degree of women K is above a certain threshold. We consider two extensions of the original model in Appendix H, allowing for vertical differentiation between agents and parallel inspections, and find that the phenomenon of information deadlock is reduced in both cases.

We highlight that the emergence of a deadlock structure involving $\Omega(N)$ nodes is somewhat counterintuitive. Consider a directed graph where each node points to its favorite node in its remaining consideration set (Figure 1 provides an example). Initially, each agent has an $\Omega(1)$ probability of being able to proceed with a guaranteed inspection (because their favorite option points back at them with $\Omega(1)$ probability[1]). The directed graph lacks short cycles and consists primarily of short directed paths which end in a guaranteed inspection. If an inspection succeeds, the pair matches and leaves, whereas failure prompts each agent to point to their next favorite partner. The directed paths would appear susceptible to disappearing (rather than closing into directed cycles) as agents iteratively conduct inspections, causing the market to clear. While disappearance of directed paths, i.e., market clearing, indeed occurs for $(\alpha, K, p) \notin \Theta$, we find that, surprisingly, for $(\alpha, K, p) \in \Theta$ these directed paths link up with other paths as the process progresses, and a deadlock involving $\Omega(N)$ nodes emerges[2]. The analysis leading to Theorem 2 shows that the formation of information deadlock can be intuitively explained by a certain branching process involving preferred agents becoming supercritical (i.e., branching factor exceeding one).

Market design insights. Our results show that in a wide range of markets, a nontrivial fraction of agents will get stuck in information deadlock. Moreover, deadlocks become more prevalent as the market gets more connected (larger K), as the men-to-women ratio α decreases (holding womens' average consideration set size K fixed), and as inspections are more likely to fail (smaller p). Our findings suggest that platforms can carefully restrict the market's connectivity to reduce deadlocks. In practice, online dating platforms such as Coffee Meets Bagel indeed restrict market connectivity

[1] For example, when each agent has a degree K, the probability of pointing back is $1/K$.

[2] Another way to consider the likelihood of deadlock is through a lower bound of $(1 - 1/K)^N$ probability (suppose each agent has a degree K and $N = M$) that the market starts with a deadlock. When the degree K is linear in N, say, $K = \theta N$ for some $\theta > 0$, then this bound become $(1/e)^{1/\theta} > 0$. However, as $1/\theta \to \infty$ (e.g., when K is a constant that does not depend on N as considered in this paper), this probability shrinks to zero.

by limiting the agent's total number of profile views or likes in a day. While our results imply that this may have a positive effect on reducing deadlocks, it may, on the other hand, lead to a higher chance of exhausting all options without finding a compatible match (see Figure 2). We find that *the level of market connectivity that maximizes match rate* (the fraction of agents matched) *is that level which causes the market to be at the phase boundary between the deadlock-free regime and the deadlock regime.* The platform can also reduce the fraction of women stuck in deadlock by incorporating more men in the market (while holding womens' average consideration set size K fixed), thus reducing the competition for men. Finally, we observe that in markets where deadlock is prevalent, agents who are still waiting after a few rounds of guaranteed inspections are very likely to get stuck in an information deadlock (see Figure 12); i.e., allowing more time does not help clear the market.

In the extensions (see Appendix H), we show that introducing vertical differentiation between agents and allowing parallel inspections can help mitigate deadlocks. This suggests that, for example, since introducing vertical differentiation among agents reduces deadlocks, hiring platforms can make certain distinguishing features of workers such as educational background or work experience more conspicuous if currently experiencing deadlocks. Platforms can also encourage more inspections to reduce deadlocks.

Technical contribution. Our primary technical contribution is to show how statistical physics tools, particularly *message passing* (see, e.g., [36, Chapter 14]), can be augmented to understand market-wide phenomena in marketplaces and to inform market design. Statistical physicists study the macroscopic phenomena that emerge when a large number of particles interact locally in simple ways. The underlying theory helps researchers explain, for example, the spontaneous magnetization of a magnetic material below a certain temperature threshold. The theory has been extended to other fields as well, helping explain and engineer macroscopic phenomena in artificial intelligence [40, 41], computer science including solving satisfiability problems [37], computer vision [18], digital communication including high performing error correcting codes [35, 19, 46], and neurobiology [57, 20]. In contrast, market design research is conspicuous in that connections with statistical physics have been little exploited so far.[3] Marketplaces bring together agents, who interact with each other, often in relatively simple ways, and yet these interactions can produce unexpected

[3]However, note work in the area of "Econophysics" `https://en.wikipedia.org/wiki/Econophysics` which has attempted to explain phenomena and observational data in financial markets using statistical physics ideas.

market-wide or "macroscopic" outcomes, e.g., see [7]. We show how a statistical physics approach can throw light on emergent macroscopic phenomena in marketplaces.

To augment and apply the statistical physics tool of message passing in our problem requires us to overcome technical challenges resulting from the cyclic dependence inherent in matching processes, which will be explained next. In message passing algorithms (e.g., belief propagation), nodes in a graph pass "messages" to their neighbours. Each message from i to j in the graph is iteratively updated (according to some predefined rule) based on the (recent) messages received by i from its neighbours *except* for j itself. Henceforce we call this the *self-exclusion* property of a message passing algorithm. The self-exclusion property of message updates facilitates direct (iterative) computation of the distribution of messages on tree graphs, simply based on the message passing rule and the (distribution of) node degrees. This computation is called *density evolution*, and also applies asymptotically to large random graphs (e.g., Erdos-Renyi graphs), since they are locally treelike. The limiting message densities as the number of iterations grows throw light on "macro" phenomena such as whether a graph contains a "giant" connected component (see Appendix G in the e-companion[4] for a heuristic derivation of the giant component regime using message passing) and whether a graphical error correcting code can be successfully decoded [36]. In our setting, whether and when a match (i, j) is formed under the partner search and matching process depends on the history of *both* potential partners; and in turn the history of i and j depends on whether inspection (i, j) was already performed. Given this apparent cyclical dependence, it is a priori unclear that our partner search process can be viewed through the lens of message passing. Somewhat surprisingly, we are able to construct a message passing algorithm which tracks the partner search process *and* possesses the key self-exclusion property. This algorithm facilitates our characterization of whether information deadlock occurs, and its size in markets which suffer from deadlock.

Interchange of limits. A direct application of the above methodology leads to a formal guarantee that in the deadlock regime, $\Omega(N)$ agents remain waiting after $t = o(\log N)$ rounds of partner search, whereas $o(N)$ agents remain waiting in the deadlock-free regime. This result is dissatisfying in that it leaves open the theoretical possibility that if one conducts additional rounds of partner search, deadlock eventually goes away even in the so-called deadlock regime (whereas numerics suggest that deadlock persists forever). The main technical barrier is that to track partner

[4]All appendices of this paper are in the e-companion.

search for t steps requires to investigate the neighborhood of radius t, and for $t > \Theta(\log N)$ this neighborhood resists analysis because it involves $\Omega(N)$ nodes, is no longer locally treelike and has strong correlations. We discover a new way to view deadlock by uncovering iteratively the ultimate chain of waiting (as $t \to \infty$ when partner search converges for the given market) starting from a given focal node, while carefully avoiding revealing parts of the neighborhood which are not relevant to deadlock to preserve sufficient independence in the revealed structure up to a depth $\sqrt{N}$. Interestingly, we relate this chain of waiting to a branching process, and find an elegant "signature" for which markets suffer from deadlock: a market suffers from deadlock if and only if this branching process is supercritical (has branching factor exceeding 1). We formalize this connection for the special case of markets with inspection success probability $p = 0$.

Related work. The phenomenon of information deadlock under our stylized model of partner search may be viewed as a manifestation of "market congestion", where a matching market fails to clear efficiently despite efforts by participants to find a partner. The study of congestion in matching markets was initiated by Roth and Xing [49]. When firms must wait for their offer to be rejected before they can issue a new offer, which takes time, some workers may experience long delays before they receive the first offer. This is worsened when the firms' preferences are correlated, since they would all initially make offers to the most desirable workers. This source of congestion is mainly due to competition, and may be alleviated by carefully restricting choice, cf. our insight that the market connectivity K should be carefully chosen. For example, Halaburda et al. [23] show in a stylized model that a matching platform can reduce congestion by restricting agent choice, and such a restricted choice platform is preferred by agents with low outside options. Similarly, Arnosti et al. [4] show that in a dynamic matching market, with costly screening and uncertain applicant availability, restricting the number of applications can effectively alleviate congestion. Congestion in matching markets is also identified in several empirical works, including Fradkin [17] and Horton [28], and related interventions such as enabling costly capacity signaling have been found to be effective [16]. Our work contributes to the literature on congestion by capturing quantitatively which random matching markets suffer from congestion (information deadlock) and suggests ways to alleviate it.

More broadly, there is a rich literature studying the behavior and design of matching markets, as a function of various market primitives, such as market imbalance [8], the cost of screening potential partners [33], whether preferences are easy to describe [51], and heterogeneity in agent quality [8]. Various design levers have been suggested to improve market performance in different contexts, for example appropriate match recommendation [51], [6], signalling [12], communication between agents [2], design of the tie breaking rule [8], [9], [3], choosing the meeting rate between different agent types [31], and "cheap talk" as a way for agents to share their vertical "type" and promote sorting and market clearing [29] (cf. our finding that vertical differentation between agents reduces deadlock). Our work provides a new lens (information deadlock in random markets) for understanding market clearing and designing matching markets; specific findings capture the effect of market imbalance, broader search for a partner (see our extension to parallel inspections), the success probability of matches, and heterogeneity in agent quality on market performance.

There is a substantial literature in economics on games between agents (or markets) in which information acquisition is costly; a classical example is the work of Grossman and Stiglitz [22] which shows informationally efficient asset markets are impossible, a finding which bears some spiritual resemblance to the information deadlock phenomenon we study. Another example of work that is loosely related to our market design insights is the literature on contest design [21, 58, 50], where the quality of submissions depends on effort put in by participants (this effort could be towards information acquisition or other aspects), and the designer aims to encourage higher effort by participants. Hatfield et al. [26] discusses designing markets to incentivize optimal investment by participants, while Pakzad-Hurson [39] brings together crowdsourcing and market design (for matching and other markets).

Two recent works by Peski [42, 43] have identified intriguing connections between matching markets and statistical physics concepts: roughly, under certain assumptions, equilibria in matching markets minimize a certain "free energy" function, or, equivalently, maximize a certain utility-plus-entropy function. This characterization is pleasantly surprising, in that individual agents are each concerned with maximizing their own utility and yet the market equilibrium is shown to maximize a global function.

The partner search process in our paper is related to the Karp-Sipser algorithm (e.g., [5]), in the sense that they are both some kind of peeling algorithm that iteratively removes edges from the graph. However, there are important difference such that readily available analysis for the Karp-Sipser algorithm does not apply here. For example, unlike the Karp-Sipser algorithm, an edge where one end node (she) has a degree of one cannot be directly removed under our Partner

Search process, since the other end node (he) does not necessarily consider her as his remaining favorite match. In fact, the agents' dynamically changing their most preferred match partner due to updates in the residual graph is the major obstacle and complicating factor in adopting or translating existing results. Therefore we need to develop a different technique, which utilizes the message passing machinery.

Organization of the paper. Section 2 introduces our model of a random matching market, the partner search process, and the information deadlock phenomenon. Section 3 gives our main result characterizing which markets incur information deadlock and the size of the deadlock, as well as our result showing formally how deadlocks persist overtime for a special case of our model, and the connection between deadlock and a certain branching process. Section 4 provides the key methods used to prove the first main result, including the design of a message passing algorithm which tracks the partner search process in the market. Section 5 outlines the analysis establishing that deadlock persists by coupling with a branching process. Section 6 describes our numerical results. We conclude in Section 7.

2 Model

Consider a matching market that consists of N women and $M = \alpha N$ men. Denote by $\mathcal{I}$ the set of women and $\mathcal{J}$ the set of men. We use an undirected bipartite graph $G = (\mathcal{I}, \mathcal{J}, \mathcal{E})$ to describe the consideration sets of all agents, where the edge set $\mathcal{E}$ includes all woman-man pairs (i, j) that consider each other. (In our setup, whenever a woman $i \in \mathcal{I}$ considers a man $j \in \mathcal{J}$, man j also considers woman i.)[5] We denote by $\mathcal{N}(i) := \{j \in \mathcal{J} : (i, j) \in \mathcal{E}\}$ the consideration set of woman i, and $\mathcal{N}(j) := \{i \in \mathcal{I} : (i, j) \in \mathcal{E}\}$ the consideration set of man j. Each woman [man] wants to match with at most one man [woman] from her [his] consideration set. Prior to matching and inspections, a woman i has preference ranking $\succ_i$ over men in her consideration set $\mathcal{N}(i)$. Similarly, a man j has prior preference ranking $\succ_j$ over his consideration set $\mathcal{N}(j)$. Therefore, we can describe the initial state of the market via $(G, (\succ_i, \forall i \in \mathcal{I}), (\succ_j, \forall j \in \mathcal{J}))$. As the search for partners progresses, the state we describe below will evolve due to elimination of some edges and some agents (preference rankings remain unchanged). We will write $G_t = (\mathcal{I}_t, \mathcal{J}_t, \mathcal{E}_t)$ for the state at time t and $\mathcal{N}_t(i)$ and

[5]This restriction is without loss of generality under the assumption that if any agent i's interest in any agent j is unreciprocated, then agent i can learn this fact without cost or delay, and hence eliminate j from their consideration set at the outset.

$\mathcal{N}_t(j)$ for the corresponding consideration sets.

A woman-man pair (i, j) needs to inspect each other to determine compatibility before the match (i, j) can be formed. Inspections are costly and succeed independently with probability p (else they fail). We assume agents only perform *guaranteed inspections*.

Definition 1 (Guaranteed inspection). *An inspection between woman-man pair (i, j) is a* guaranteed inspection *at time t if*

- *woman i has no preferred man in her consideration set: for all $j' \in \mathcal{N}(i)$ such that $j' \succ_i j$, $j' \notin \mathcal{N}_t(i)$*
- *and man j has no preferred woman in his consideration set: for all $i' \in \mathcal{N}(j)$ such that $i' \succ_j i$, $i' \notin \mathcal{N}_t(j)$.*

Since guaranteed inspections require the pair to be each other's favorite remaining partner, we assume that if the inspection succeeds, the pair indeed matches and leaves the market.

We describe the market dynamics by a process called *partner search* (PS), formally specified in Figure 4, which models a market where agents iteratively conduct guaranteed inspections and updates consideration sets to remove matched pairs of agents or edges that fail inspections. At the end of the t-th round of partner search, $\mathcal{I}_t$ is the set of remaining women with nonempty consideration sets. We say that women in $\mathcal{I}_t$ are *waiting* (to find a partner) in the residual market at time t. For notational convenience, if the algorithm terminates at time t, define $\mathcal{I}_{t'} = \mathcal{I}_t$ for all $t' > t$. The sets $(\mathcal{I}_t)_{t=1,2,\ldots}$ then form a (weakly) decreasing sequence. The market has not fully cleared at time t as long as $\mathcal{I}_t \neq \emptyset$, i.e, $\Lambda^t > 0$. We are concerned with the size of this group $\Lambda^t = |\mathcal{I}_t|$ at time[6] t.

Our analysis studies the number of waiting women in a sequence of Erdos-Renyi random markets defined in Definition 2. The parameters α, K and p are held fixed while letting $N \to \infty$.

Definition 2 (Random market). *For any $(\alpha, K, p) \in (0, \infty) \times (0, \infty) \times [0, 1]$, let $G_N(\alpha, K, p)$ be the* random market *$(\mathcal{I}, \mathcal{J}, \mathcal{E}, (\succ_i, \forall i \in \mathcal{I}), (\succ_j, \forall j \in \mathcal{J}))$ with $|\mathcal{I}| = N$ women and $|\mathcal{J}| = M = \alpha N$ men, where*

- *each woman-man pair (i, j) is in $\mathcal{E}$ independently with probability K/M;*

[6] The same analysis can be done for men without loss of generality.

Partner Search (PS)

```
Initialize the market: women 𝓘_0 ← 𝓘; men 𝓙_0 ← 𝓙; potential pairs 𝓔_0 ← 𝓔; deadlock ← FALSE; time t ← 1.
while |𝓘_{t−1}| > 0 AND NO deadlock do
    Initialization: 𝓘_t ← 𝓘_{t−1}; 𝓙_t ← 𝓙_{t−1}; 𝓔_t ← 𝓔_{t−1}
    if 𝓔_{t−1} has no guaranteed inspection then
        deadlock ← TRUE.
    end if
    while ∃ a guaranteed inspection (i, j) ∈ 𝓔_{t−1} do
        Perform inspection (i, j).
        if inspection (i, j) succeeds then
            The pair matches and leaves the market: 𝓘_t ← 𝓘_t\{i}, 𝓙_t ← 𝓙_t\{j}, 𝓔_t ← 𝓔_t\{(i′, j′) : i′ = i OR j′ = j}.
        else
            The pair is removed from the set of edges, but the agents remain: 𝓔_t ← 𝓔_t \ {(i, j)}.
        end if
    end while
    Eliminate any woman i and man j with empty consideration sets: 𝓘_t ← 𝓘_t \ {i}, 𝓙_t ← 𝓙_t \ {j}.
    The number of women who have not cleared Λ^t ← |I_t|.
    Advance time: t ← t + 1.
end while
```

Figure 4: The Partner Search process. Preference rankings remain unchanged throughout.

- *preference rankings $\succ_i$ for each $i \in \mathcal{I}$ and preference rankings $\succ_j$ for each $j \in \mathcal{J}$ are independent uniformly random permutations of elements in $\mathcal{N}(i) := \{j \in \mathcal{J} : (i,j) \in \mathcal{E}\}$ and $\mathcal{N}(j) := \{i \in \mathcal{I} : (i,j) \in \mathcal{E}\}$, respectively;*
- *each inspection's latent outcome is a success independently with probability p.*

Also, for a random market $G_N(\alpha, K, p)$, following the PS process in Figure 4, define $\Lambda_N^t(\alpha, K, p) = |\mathcal{I}_t|$ to be the number of women who have not cleared at time t.

From Definition 2, in a random market $G_N(\alpha, K, p)$, each woman $i \in \mathcal{I}$ has a consideration set of size $|\mathcal{N}(i)|$ distributed according to[7] $\text{Binomial}(M, K/M) \xrightarrow{N\to\infty} \text{Poisson}(K)$. Also, for each man j, $|\mathcal{N}(j)|$ is distributed according to $\text{Binomial}(N, K/M) \xrightarrow{N\to\infty} \text{Poisson}(K/\alpha)$.

In Appendix A, we discuss and formalize the connection between our partner search process where agents are assumed to perform satisficing search, and the model of Immorlica et al [30] which involves cardinal utilities and introduces the notions of information deadlock and regret-free stability. In the same appendix section, we also discuss our assumptions of i.i.d. uniform agent preferences and independent inspection successes with probability p.

We aim to understand which markets suffer from an information deadlock, namely, a large fraction of women waiting for a long time. Clearly, $\Lambda_N^t(\alpha, K, p)$ — the number of women who have

[7] This convergence is in distribution.

not cleared at time t — is a random variable. We say a market suffers from information deadlock if the following holds.

Definition 3. *A sequence of markets $G_N(\alpha, K, p)$ as defined in Definition 2, parameterized by (α, K, p) and indexed by N, suffers from an* information deadlock *of size $\tilde{\lambda} > 0$ if, there exists a sequence of times $t = t_N = \omega(1)$ such that*

$$\lim_{N\to\infty} \Pr\left(|\Lambda_N^t(\alpha, K, p) - \tilde{\lambda}N| \leq f(N)\right) = 1 \tag{1}$$

for some $f(N) = o(N)$.

We will find that whenever (1) holds for some $\tilde{\lambda}$ and some sequence of times t which is $\omega(1)$ and $o(\log N)$, it holds with the same $\tilde{\lambda}$ for *all* sequences of times which are $\omega(1)$ and $o(\log N)$. Our analysis leading to our first main result needs $t = o(\log N)$ for technical reasons. In particular, this assumption allows us to focus on the local neighborhood of a representative agent up to a radius $o(\log N)$. This neighborhood is a tree with high probability for radius $o(\log N)$, but not for radius $\Omega(\log N)$. (Lemma 1 in Section 4.1 shows that the local neighborhood of a representative agent up to radius $o(\log N)$ converges in distribution to a specific Galton-Watson tree.) Simulation results strongly suggest that the size of information deadlock remains the same for $t = \Omega(\log N)$ as well; see Section 6, and we moreover formally establish that information deadlock persists forever for the special case of inspection success probability $p = 0$.

In the next section, we will characterize $\lim_{N\to\infty} \Lambda_N^t(\alpha, K, p)/N$ for large t. In particular, we identify a broad range of markets where information deadlock occurs with high probability.

3 Main Results

In this section, we characterize which markets suffer from information deadlock and which do not. In particular, Section 3.1 establishes that information deadlock (for $t = \omega(1)$ and $o(\log N)$) vanishes in some cases, but is prevalent in a wide range of random markets, corresponding to a large region in the space of primitives (α, K, p), and quantifies the size of deadlock as a function of (α, K, p). Section 3.2 investigates how deadlock persists in the deadlock regime when $t = \omega(\log N)$.

3.1 Main Theorem

We identify a set Θ of values of (α, K, p) that characterizes in which markets is there an information deadlock up to $o(\log N)$ time as $N \to \infty$: if $(\alpha, K, p) \in \Theta$, then with high probability (w.h.p), a non-vanishing fraction of women wait in the residual market; on the other hand, if $(\alpha, K, p) \notin \Theta$, then w.h.p., only a vanishing fraction of women wait. To facilitate the statement of our main result, we start by defining a certain vector sequence. We do not expect the reader to understand where these vectors come from at this point. We will later obtain these vectors in Section 4 via an analysis method from statistical physics termed density evolution.

Definition 4 (Density evolution). *For any $(\alpha, K, p) \in (0,\infty) \times (0,\infty) \times [0,1]$, let $u_0 = 1, y_0 = \nu_0 = 0$ and define sequence $(a_t, r_t, w_t, q_t, y_t, \nu_t, u_t)_{t\geq 1}$ as follows. For all $t = 1, 2, \dots,$ let*

$$a_t = \frac{e^{\frac{K}{\alpha}(\nu_{t-1}-1)} - 1}{\frac{K}{\alpha}(\nu_{t-1} - 1)}; \quad r_t = \frac{y_{t-1}}{1-\nu_{t-1}}(1-a_t); \quad w_t = \frac{u_{t-1}}{1-\nu_{t-1}}(1-a_t);$$
$$q_t = \frac{1-e^{-K(a_t p + w_t)}}{K\,(a_t p + w_t)}; \quad y_t = q_t p; \quad \nu_t = q_t(1-p) + \frac{a_t p(1-q_t)}{a_t p + w_t}; \quad u_t = \frac{w_t(1-q_t)}{a_t p + w_t}.$$

Also define[8] $\lambda(\alpha, K, p) := \lim_{t\to\infty} K w_t q_t$.

The sequence $(a_t, r_t, w_t, q_t, y_t, \nu_t, u_t)_{t\geq 1}$ in Definition 4 is closely related to the probabilistic description of a representative agent's "local information" at time t in the large market limit. For example, reveal (only) the consideration set of woman i and pick any man j in a woman i's initial consideration set. At time t, a_t is the probability that j can no longer possibly be matched with any preferred woman, r_t is the probability that j has been matched with some other preferred woman and is not available to woman i, and w_t is the probability that j is still considering i but does not yet see i as most preferred. The interpretations of (q_t, y_t, ν_t, u_t) are slightly more involved and will be discussed in detail in Section 4. We will show that the fraction of women who remain waiting in the large market limit will be $\lambda(\alpha, K, p)$. The characterization of $\lambda(\alpha, K, p)$ is unique for any choice of (α, K, p).

The next proposition characterizes markets where $\lambda(\alpha, K, p) > 0$ and establishes existence of such markets.

Proposition 1. *Consider the value $\lambda(\alpha, K, p)$ defined in Definition 4. Define set*

[8] By Corollary 1 in Section 4.1, the limit $\lim_{t\to\infty} K w_t q_t$ exists.

$$\Theta := \{(\alpha, K, p) : \alpha > 0, K > 0, p \in [0,1], \lambda(\alpha, K, p) > 0\}.$$

Then Θ is a nonempty set with infinite volume in $\mathbb{R}^3$.

In the proof, we show the set Θ is nonempty by showing that Θ contains a non-empty subset. This non-empty subset corresponds to markets where the women's average degree K is sufficiently large and an inspection's success probability is sufficiently low. In particular, if we fix any men-to-women ratio α and women's average number of compatible potential man partners Kp, we will get $\lambda(\alpha, K, p) > 0$ as the women's degree K exceeds a certain threshold, i.e., as the inspections' success probability p falls below a certain threshold. The full proof can be found in Appendix C in the e-companion.

Now we are ready to state our first main result.

Theorem 1. *Consider the sequence of markets $G_N(\alpha, K, p)$ indexed by N and number of women who are still waiting at time t, $\Lambda_N^t(\alpha, K, p)$, both defined in Definition 2, for some sequence of t also implicitly indexed by N. Consider set Θ in Proposition 1 and $\lambda(\alpha, K, p)$ in Definition 4. Then the following statements are true.*

- *(Information Deadlock Regime) If $(\alpha, K, p) \in \Theta$, i.e., $\lambda(\alpha, K, p) > 0$, then for any sequence of times $t = w(1)$ that is also $o(\log N)$, we have*

$$\lim_{N\to\infty} \mathbb{P}\left(|\Lambda_N^t(\alpha, K, p) - \lambda(\alpha, K, p)N| \le f(N)\right) = 1$$

  for some $f(N) = o(N)$.

- *(Deadlock-free Regime) If $(\alpha, K, p) \notin \Theta$, then for any sequence of times $t = w(1)$, we have*

$$\lim_{N\to\infty} \mathbb{P}\left(\Lambda_N^t(\alpha, K, p) \le f(N)\right) = 1$$

  for some $f(N) = o(N)$.

The information deadlock regime is where, with high probability (w.h.p.), a positive fraction of women[9] wait in the market for a long time, and Theorem 1 moreover identifies this fraction to be

[9] We can similarly characterize the asymptotic fraction of men waiting in the market to be $\lim_{t\to\infty} \frac{K a_t w_t(1-q_t)}{\alpha(a_t p + w_t)}$. Note that the mens' fraction is zero if and only if the womens' fraction $\lambda(\alpha, K, p) = \lim_{t\to\infty} K w_t q_t$ is zero, which occurs if and only if $\lim_{t\to\infty} w_t = 0$.

nearly $\lambda(\alpha, K, p)$. The deadlock-free regime is where, w.h.p., only $o(N)$ women get stuck waiting in the market for a long time. The most surprising implication of Proposition 1 and Theorem 1 is the existence of a sizeable information deadlock regime $(\alpha, K, p) \in \Theta$. As discussed in Section 1, this *a priori* seemed unlikely in random markets, given the apparent implausibility of forming the structure required for information deadlock to occur ("cycles" of agents waiting for each other and "directed paths" of waiting agents which end in such cycles) given the absence of short cycles in the consideration graph. In Section 4 we provide an overview of the proof of Theorem 3.1 and explain why information deadlock nevertheless arises for $(\alpha, K, p) \in \Theta$: roughly, directed paths of waiting agents join each other as partner search progresses and become longer, causing deadlock (see Figure 5). In Section 6, we numerically evaluate the characterization in Theorem 1, which point towards a phase transition between the two regimes and show the monotonicities of $\lambda(\alpha, K, p)$ and projections of Θ.

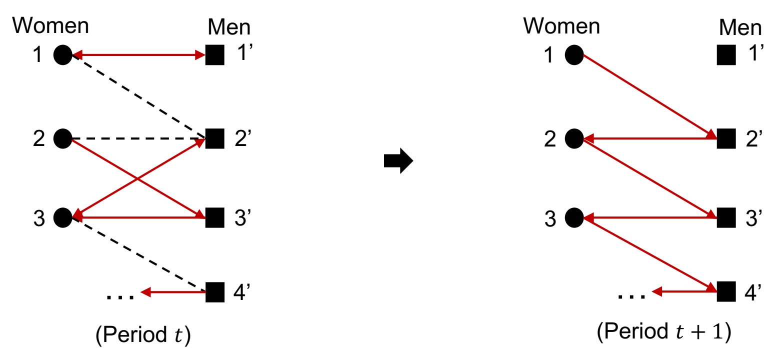

Figure 5: A graphical illustration of directed paths of waiting agents joining each other and becoming longer. Directed edges indicate agents' favorite remaining partner, and dashed edges indicate other remaining neighbors in the agents' consideration set. Suppose inspections all failed in period t. Then the directed paths in the left figure can join each other to form a longer path in period $t + 1$ as shown in the right figure.

3.2 The Deadlock Regime

From Theorem 1, the set Θ characterizes the information deadlock regime. That is, after $\omega(1)$ and $o(\log N)$ rounds of partner search, w.h.p., there remains a nontrivial fraction of agents waiting if and only if $(\alpha, K, p) \in \Theta$. The requirement of $o(\log N)$ rounds of partner search is due to technical reasons. This does not pose a problem in the deadlock-free regime, since the fraction of agents waiting is already vanishing at that point. However, in the deadlock regime, one may wonder what happens after $o(\log N)$ rounds of partner search. For example, is it possible that the fraction of agents who are waiting further decreases, possibly to zero? Numerical evidence strongly suggests that this is not the case. Later, in Section 6, we provide simulation results for finite-sized markets, and find that the fraction of agents who are still waiting after the partner search process converges is well captured by the characterization in Theorem 1. In this section, we investigate this issue theoretically and show (under certain parameter regimes) that information deadlock persists forever if it exists after $o(\log N)$ and $\omega(1)$ rounds of partner search.

To manage the intricacy of the analysis, we restrict our theoretical development to the special case of $p = 0$. When $p = 0$, any inspection will fail if conducted. Nevertheless, the notion of information deadlock remains applicable to this case, since agents who are not aware of their actual compatibilities could still get stuck waiting for others, hoping that something might work out eventually. The fact that the information deadlock phenomenon is qualitatively the same at $p = 0$ as at $p > 0$ makes us optimistic that the persistence of deadlock in the deadlock regime for $p = 0$ generalizes to $p > 0$, though we do not pursue the latter formally.

We now present the formal result.

Theorem 2. *If $(\alpha, K, 0) \in \Theta$, i.e, $\lambda(\alpha, K, 0) > 0$, then there exists $\epsilon = \epsilon(\alpha, K) > 0$ such that $\mathbb{E}[\Lambda_N^\infty(\alpha, K, 0)] > \epsilon N$ and $\liminf_{N\to\infty} \mathbb{P}(\Lambda_N^\infty(\alpha, K, 0) > \epsilon N) \geq \epsilon$. Moreover, $(\alpha, K, 0) \in \Theta$ if and only if $\frac{K^2}{4\alpha} > 1$.*

We highlight a few observations from Theorem 2. First, this result establishes the same deadlock phase transition as predicted by Theorem 1; in particular, for any primitives in the deadlock regime $(\alpha, K, 0) \in \Theta$, a linear sized deadlock persists forever with positive probability. This supplements the numerical evidence (see Figure 13) that the deadlock characterization in Theorem 1 is exact, with a formal proof of a linear lower bound on the eventual deadlock size everywhere in the deadlock regime. Notably, the analysis leading to Theorem 2 enables us to understand the formation of deadlock through the lens of a branching process of preferred neighbors which is supercritical in the deadlock regime. Specifically, by a coupling argument, the network structure supporting a representative agent being in deadlock is uncovered by a carefully designed tree revelation procedure, through which the event of being in deadlock or not is captured by the associated branching process of preferred neighbors living forever or not. The underlying branching process is difficult to track in general due to a complicated correlation structure, but it reduces to a standard branching process in the special case of $p = 0$, and we can see that the deadlock phase transition predicted by Theorem 1 exactly coincides with the branching factor of the underlying tree process exceeding one (therefore the probability of the tree living forever becomes positive).

The branching process of preferred partners. The condition $\frac{K^2}{4\alpha}$ in Theorem 2 can be intuitively explained as follows. Consider a representative woman who receives an offer from some man. She would like to wait for men that she prefers to this offer. In a large market, the number of preferred men is distributed like Poisson(Ku), where $u \sim$ Uniform$(0,1)$, and the mean is $\frac{K}{2}$. Likewise, independently for each man, the number of preferred women is distributed like Poisson($\frac{K\hat{u}}{\alpha}$), where $\hat{u} \sim$ Uniform$(0,1)$, and the mean is $\frac{K}{2\alpha}$. Therefore, we can approximate the representative woman's local neighborhood structure that induces her waiting using a branching process of preferred partners, where each woman node has Poisson(Ku) men offspring and each man node has Poisson($\frac{K\hat{u}}{\alpha}$) offspring, independently. The branching factor of this process is $\frac{K^2}{4\alpha}$. Intuitively, the representative woman will wait forever if and only if this branching process is nonextinctive. It is well known in the literature (e.g. [10, 24]) that the latter occurs with positive probability if and only if the branching factor exceeds 1.

Heuristic branching process perspective for $p > 0$. The above heuristic argument relies on the assumption of $p = 0$, i.e., inspections never succeed, hence an agent will always proceed to wait for the next (according to the preference ranking) preferred agent after an inspection. As a consequence, in the associated branching process of preferred partners, we can include for each node all its preferred agents as offsprings, independently. This is no longer the case when $p > 0$. If an inspection succeeds, an agent is matched and does not proceed to wait for the next preferred agent. This means that the branching process of preferred partners being supercritical no longer (asymptotically) implies deadlock for $p > 0$. Instead, supercriticality only provides a loose necessary condition for deadlock, in fact too loose to identify the phase transition between the deadlock regime and the deadlock-free regime. To construct the deadlock-relevant branching process of preferred partners, when we generate offspring for each node, we need to consider, according to the preference ranking, whether the preferred agent will inspect with the parent node and succeed, since if so, we should not include additional offsprings for the parent node. In particular, the associated deadlock-relevant branching process has a nontrivial correlation structure, where the number of offspring for a node depends on future offspring generations.

Although the associated deadlock-relevant branching process lacks a simple description for $p > 0$, we can still guess its branching factor using a heuristic argument (see Appendix F), and conjecture that as in the case of $p = 0$, even when $p > 0$, the condition for a nontrivial fraction of agents to wait forever is captured by this branching factor exceeding one (see Figure 14 in Section 6). Moreover, the condition for the guessed branching factor to exceed one matches the characterization of the deadlock regime Θ in Theorem 1 (see Appendix F).

Conjecture 1. *Consider $p > 0$. If $(\alpha, K, p) \in \Theta$, then $\mathbb{E}[\Lambda_N^\infty(\alpha, K, p)] > \epsilon N$ for some $\epsilon > 0$. Moreover, $(\alpha, K, p) \in \Theta$ if and only if $\left(\frac{1}{a^*p} - \frac{y^*}{a^*p^2}\right) \cdot \left(\frac{1}{y^*} - \frac{a^*}{y^*}\right) > 1$, where y^* and a^* are the nonzero solution to*

$$Ky^*a^* = 1 - e^{-Ka^*p}, \quad \frac{K}{\alpha}y^*a^* = 1 - e^{-\frac{K}{\alpha}y^*}.$$

In Conjecture 1, the value $\left(\frac{1}{a^*p} - \frac{y^*}{a^*p^2}\right) \cdot \left(\frac{1}{y^*} - \frac{a^*}{y^*}\right)$ is the conjectured branching factor of the deadlock-relevant branching process, and nontrivial deadlock which persists occurs if and only if this value exceeds one, just as in the $p = 0$ case. We provide a heuristic argument of its derivation in Appendix F. However, a formal proof requires careful analysis of the deadlock-relevant branching process much beyond its (conjectured) branching factor. Due to the additional nontrivial correlation structure when $p > 0$ mentioned above, techniques used in the proof of Theorem 2 would not suffice. Therefore it is beyond the scope of this paper to prove Conjecture 1, and we leave it for future study.

4 Proof Idea of Theorem 1

Theorem 1 characterizes the probability that a representative agent will remain waiting in the large market limit. Our proof involves three mainsteps. First, we show that the representative agent's neighborhood in the market (up to any $o(\log N)$ depth) converges in distribution to a tree generated from a marked "bipartite" branching process. In particular, a representative agent's neighborhood as "locally tree-like" (up to any depth which is $o(\log N)$). Second, we construct a *message passing* (MP) algorithm that passes messages between women and men iteratively, such that its progress on a marked random tree (the branching process in the first step) "tracks" what happens in the partner search process in a random market. Inspections, matches, and whether an agent is still waiting can be inferred from the messages passed between women and men at any given time. The third and final step studies the evolution of message distributions on the random tree, yielding the probability the root agent is waiting at any given time in the large market limit; an analysis called *density evolution.*

4.1 Local Structure

In this section, we establish that the local neighborhood of any agent in the large market limit converges in distribution to a tree generated from a Poisson branching process, suitably augmented with random preferences and latent inspection outcomes. Classic random graph theory already establishes that a bipartite Erdos-Renyi graph, which captures the consideration graph of the market, converges locally weakly to a corresponding bipartite Galton-Watson tree. We extend this result by, for both structures, embedding participant preference (fitness) rankings as well as "latent" inspection outcomes on edges, so that the trajectory of partner search in the corresponding market is uniquely determined, and can be compared across the two markets.

We define the distance between any two nodes as the number of men on the shortest path connecting them. The distance-r neighborhood of an agent $i \in G_N(\alpha, K, p)$ is the subgraph (including the preference/fitness rankings and latent inspection outcomes) spanned by all nodes within distance r from i. We denote by $\mathcal{B}_r(i)$ the distance-r neighborhood of agent i. We compare this neighborhood to a marked Galton-Watson tree, defined below, up to the same depth r.

Definition 5. *A* marked Galton-Watson (GW) tree $\mathcal{T} \equiv \mathcal{T}(\alpha, K, p)$ *is a random tree with woman and man nodes constructed as follows:*

- *The root is a woman node.*
- *Each woman node has, independently, a Poisson*(K) *number of man offspring nodes.*
- *Each man node has, independently, a Poisson*(K/α) *number of woman offspring nodes.*
- *Each woman (man) node ranks its neighboring man (woman) nodes according to an independent uniformly random permutations.*
- *Each edge has a latent inspection outcome associated with it, which is Bernoulli*(p)*, independently drawn.*

The depth-r tree $\mathcal{T}_r$ is the distance-r neighborhood of the root woman node in $\mathcal{T}$. The next lemma establishes that in large markets, $\mathcal{B}_r(i)$ converges to $\mathcal{T}_r$ in distribution.

Lemma 1. *Fix any (α, K, p). Take a sequence of markets $G_N(\alpha, K, p)$ and a sequence of radii $r = r_N = o(\log N)$. Also take a depth-r tree $\mathcal{T}_r$ generated by the Poisson branching process in Definition 5. Consider an agent $i \in G_N(K, \alpha, p)$ and its distance-r neighborhood[10] $\mathcal{B}_r(i)$. Then there exists a coupling between $\mathcal{B}_r(i)$ and $\mathcal{T}_r$ such that $\mathbb{P}(\mathcal{B}_r(i) \neq \mathcal{T}_r) \leq o(1)$.*

To illustrate the importance of Lemma 1, we now present an informal argument showing the absence of deadlock in markets with $\alpha > K^2$. If a number of agents end up waiting forever, then the final deadlock structure must consist of cycles of nodes pointing to their most preferred neighbors (who remain consideration) and paths pointing to such cycles, cf. the example in Figure 7. Note that the cycles in the deadlock structure must be cycles in the original consideration graph G. That is to say, if the original consideration graph G lacks cycles, information deadlock should also be trivial. Now when the branching factor K^2/α of the corresponding branching process $\mathcal{T}$ is less than one, i.e., when $\alpha > K^2$, classic branching process theory (e.g., [10, 24]) tells us that the branching process becomes extinct almost surely in this case. Therefore, from Lemma 1, w.h.p. all but $o(N)$ nodes in the graph G belong to small-sized trees. Hence, with $\alpha > K^2$, w.h.p. at most $o(N)$ nodes can end up in a deadlock.

4.2 Message Passing

To investigate what happens when $\alpha < K^2$, the description of the initial market at time 0 alone is not enough. Namely, we need to also understand the residual market after several rounds of the partner search process. For this purpose, we construct a message passing (MP) algorithm which tracks the PS process, and derive the corresponding *density evolution* equations to gain a macroscopic view of the residual market.

As its name suggests, the MP algorithm passes messages, iteratively over time, between women and men that are connected in the prior consideration graph $\mathcal{E}$. Messages are directed: for any $(i, j) \in \mathcal{E}$ and time t, we denote by $m^{(t)}_{i\to j}$ the message passed from woman i to man j and $\hat{m}^{(t)}_{j\to i}$ the message from man j to woman i. As will become clear, the key property of the MP algorithm[11] is that the message i passes to j at time t does not depend on the messages that i has received from j, but is only a function of the messages i received most recently from all other neighboring nodes in $\mathcal{N}(i) \setminus \{j\}$.

[10]We suppress notational dependence of $\mathcal{B}_r(i)$ on N.

[11]This key property leads to the messages received by a node from its different neighbors being independent of

Message Passing (MP)

Input: $(\mathcal{I}, \mathcal{J}, \mathcal{E}, (\succ_i, \forall i \in \mathcal{I}), (\succ_j, \forall j \in \mathcal{J}))$ with latent inspection results $\epsilon_{ij} \sim \text{Bernoulli}(p), \forall (i,j) \in \mathcal{E}$.
Initialize the messages: $m^{(0)}_{i\to j} = \text{U}, \hat{m}^{(0)}_{j\to i} = \text{W}$ for all $(i,j) \in \mathcal{E}$.
$t \leftarrow 1$
do
Update man-to-woman messages: For all $(i,j) \in \mathcal{E}$, let

$$\hat{m}^{(t)}_{j\to i} = \begin{cases} \text{A if for all } i' \succ_j i, m^{(t-1)}_{i'\to j} = \text{N}; \\ \text{R if for some } i' \succ_j i, m^{(t-1)}_{i'\to j} = \text{Y, and } m^{(t-1)}_{i''\to j} = \text{ N for all } i'' \succ_j i'; \\ \text{W otherwise.} \end{cases} \tag{2}$$

Inspection ready: For all $(i,j) \in \mathcal{E}$, let

$$I^{(t)}_{ij} = \begin{cases} 1 \text{ if for all } j' \succ_i j, \text{ either } \hat{m}^{(t)}_{j'\to i} = \text{A and } \epsilon_{ij'} = 0; \text{ or } \hat{m}^{(t)}_{j'\to i} = \text{R}. \\ 0 \text{ otherwise.} \end{cases} \tag{3}$$

Update woman-to-man messages: For all $(i,j) \in \mathcal{E}$, let

$$m^{(t)}_{i\to j} = \begin{cases} \text{Y if } I^{(t)}_{ij} = 1 \text{ and } \epsilon_{ij} = 1; \\ \text{N if } I^{(t)}_{ij} = 1 \text{ and } \epsilon_{ij} = 0, \text{ or if for some } j' \succ_i j, \\ \quad \hat{m}^{(t)}_{j'\to i} = \text{A and } I^{(t)}_{ij'} = 1 \text{ and } \epsilon_{ij'} = 1; \\ \text{U otherwise.} \end{cases} \tag{4}$$

$t \leftarrow t + 1$
while some message changed

Figure 6: The Message Passing algorithm.

Our MP algorithm is specified in Figure 6. It employs a ternary alphabet for both woman-to-man messages, Y, N and U (interpreted as Yes, No and Unsure, resp.), and man-to-woman messages, A, R and W (interpreted as Accept, Reject and Waitlist, resp.). The message a node sends to each neighboring node is updated over time, based on the messages it receives from all other neighboring nodes. All messages are initialized to U (or W). Then some messages change to Y or N from U (or A or R from W) at some point in time, indicating a change in the sender's availability or in their top preference for the receiver. In particular, consider a man-to-woman message $\hat{m}^{(t)}_{j\to i} \in \{\text{A, R, W}\}$. $\hat{m}^{(t)}_{j\to i}$ =A means man j ruled out all women they prefer to i. $\hat{m}^{(t)}_{j\to i}$ =R means j has matched with woman he prefers to i and is not available to i. $\hat{m}^{(t)}_{j\to i}$ =W means j is still considering i but does not yet see i as his favorite. Then based on the messages woman i receives during the t-th iteration of MP, she decides which of her connected men are *inspection ready*. Inspection ready ($I^{(t)}_{ij} = 1$) means woman i has ruled out men she prefers to j. Women send messages to neighboring men $m^{(t)}_{i\to j} \in \{\text{Y, N, U}\}$ with the following interpretations. Y means woman i will match (will inspect and find compatible) with j if j also sees her as most fit. N means one of the following two cases is

each other, and hence plays a pivotal role in defining and legitimizing the density evolution analysis in Section 4.3.

true. In case one, i is matched with man she prefers over j and is not available to j. In case two, i sees j as the most preferred and will inspect j and find incompatibility, if j also sees her as most preferred. U means i is still considering j but does not yet see j as most preferred. Note that once a message is changed to Y, N, A, or R, it stays the same thereafter.

A reader might notice that the progress of "time" in MP is different from that in PS. In fact, MP advances faster than PS. Also, although it might appear that some messages inherit future information such as "will inspect and will fail", we show next that the MP algorithm tracks exactly the progress of an *accelerated* partner search (APS) which does not see into the future. To manage the length of the paper, we provide the precise description of APS in Appendix B (Algorithm 1). The main difference between APS and PS is that under APS, if an agent experiences an unsuccessful inspection, they immediately (in the same round) proceed to look into their next most preferred neighbor, and try to inspect them. Specifically, under APS in each round, first all the men look into potential partners sequentially as described, and then all the women do the same.

Lemma 2 (MP tracks APS)**.** *The MP algorithm tracks the accelerated partner search (APS, Algorithm 1 in Appendix B) process in the following sense. Conduct MP and APS on the same initial market* $(\mathcal{I}, \mathcal{J}, \mathcal{E}, (\succ_i, i \in \mathcal{I}), (\succ_j, j \in \mathcal{J}))$ *with latent inspection outcomes* $(\epsilon_{ij}, (i,j) \in \mathcal{E})$*. Consider any woman-man pair* $(i,j) \in \mathcal{E}$ *after* t' *rounds of APS for any* t'*. Then* i *inspected* j *iff* $\hat{m}^{(t')}_{j \to i} = \mathrm{A}$ *and* $I^{(t')}_{ij} = 1$*. Also* i *is waiting iff there exists some* $j \in \mathcal{N}(i)$ *such that* $\hat{m}^{(t')}_{j \to i} = \mathrm{W}$ *and* $I^{(t')}_{ij} = 1$*.*

In fact, from Lemma 2, other than inspections and waiting, the actual correspondence between APS and MP extends to matches formed and agents who exhaust their preference lists. For example, (i,j) are matched iff they inspected and the latent outcome $\epsilon_{ij} = 1$. This is equivalent to $\hat{m}^{(t)}_{j \to i} = A, m^{(t)}_{i \to j} = Y$. On the other hand, i exhausted her consideration set iff she is neither matched nor waiting.

The order of the message updates in the MP algorithm in Figure 6 is without loss of generality. The next lemma establishes that the outcome of MP does not depend on the sequence in which inspections are conducted. This implies that PS and APS must converge to the same final result.

Lemma 3. *Fix any* $(\alpha, K, p) \in \mathbb{R}_+ \times \mathbb{R}_+ \times [0,1]$*. Given any realization of random market* $G_N(\alpha, K, p)$*, the MP algorithm in Figure 6 will converge to a unique fixed point that does not depend on the sequence of message updates.*

Lemma 3 is shown by defining a partial ordering over message vectors and utilizing Tarski's fixed point theorem. Observe that there is a natural ordering of the man-to-woman messages $\{\mathrm{A}, \mathrm{R}, \mathrm{W}\}$, since the MP algorithm will only update the W message to either A or R, and will not update the A and R messages. Likewise, the woman-to-man messages will only possibly be updated from U to either Y or N, and will not update the N and Y messages. We can extend the ordering of individual messages to a partial ordering over message vectors. By applying that the MP update rule is a monotone function with respect to this partial ordering, we can use the standard technique of Tarski's fixed point theorem as in Hatfield & Milgrom [27], Quint [45], among others.

Next we establish a technical lemma that brings together PS and APS and formalizes that (even though PS is not as fast as APS) PS progresses quickly. Recall that $\mathcal{N}_t(i)$ is woman i's consideration set at time t in the PS process. We add superscript PS and write $\mathcal{N}_t^{\mathrm{PS}}(i)$ to differentiate with her consideration set $\mathcal{N}_t^{\mathrm{APS}}(i)$ under the APS process after t periods. Define $\mathcal{N}_t^{\mathrm{APS}}(i) = \emptyset$ (resp. $\mathcal{N}_t^{\mathrm{PS}}(i) = \emptyset$) if i is matched or removed by t under APS (resp. PS).

Lemma 4. *For any sequence of times $t = t_N = \omega(1)$, there is a sequence of times $t' = t'_N = \omega(1)$ such that, if we conduct both PS and APS on the same initial market with the same latent inspection outcomes, then, w.h.p., $\left|\left\{i \in \mathcal{I} : \mathcal{N}_t^{\mathrm{PS}}(i) \subseteq \mathcal{N}_{t'}^{\mathrm{APS}}(i)\right\}\right| \geq N - o(N)$. Also, the converged outcomes of PS and APS are identical.*

Lemma 4 states that the PS process is fast. Note that APS is an accelerated version of PS since under APS an agent can conduct multiple inspections. Therefore, we must run fewer rounds of APS than PS in order for APS to be running "behind" PS in most of the market (such that only $o(N)$ women in APS proceed faster than in PS). Lemma 4 shows that this can be achieved while still running "many" rounds of APS. In particular, for any $t = \omega(1)$ rounds in PS and its residual market, it must also take $\omega(1)$ rounds in APS to achieve a close residual market where only $o(N)$ women proceed faster than in PS. On the other hand, again, since APS is faster than PS, hence for any $t = o(\log N)$ in PS, we must also have $t' = o(\log N)$ in order to bring the residual market under t' rounds of APS close to that under PS.

4.3 Density Evolution

Next we establish that we can calculate the marginal distribution of messages after any number of rounds of MP on a marked GW tree $\mathcal{T}(\alpha, K, p)$ (Definition 5), and that the distribution is given by Definition 4. To understand the evolution of MP on the marked GW tree, we note that the messages i receives come from independent and identically distributed subtrees. This is because, in the MP algorithm, the message i receives from j only depends on the messages that j previously received from his *other* neighbors $\mathcal{N}\backslash\{i\}$, and the messages are iteratively updated. Similarly, the messages $j \in \mathcal{N}(i)$ receives also come from independent and identically distributed subtrees. Therefore, we can iteratively compute the marginal distribution of messages that the root node in the marked GW tree receives after $1, 2, ...$ rounds of message updates. This technique is called *density evolution*. The next lemma formalizes this result.

Lemma 5 (Definition 4 captures density evolution). *Conduct MP on a random marked GW tree $\mathcal{T}(\alpha, K, p)$. Then the messages the root woman node receives from neighboring men at any $t \geq 1$ are ex-ante i.i.d. with the following distribution:* A *w.p.* a_t*,* R *w.p.* r_t*, and* W *w.p.* w_t*, where (a_t, r_t, w_t) are as in Definition 4.*

Similary to the meaning of (a_t, r_t, w_t), the tuple (y_t, ν_t, u_t) gives the marginal distribution of message (Y, N, U) that a man node in the marked GW tree receives after t rounds of message updates. The value of q_t is the marginal probability of an edge in the marked GW tree being inspection ready after t rounds of message updates. An immediate corollary of Lemma 5, together with the observation that messages are monotone in time (see Lemma 10 in Appendix B in the e-companion) is:

Corollary 1. *Consider w_t and a_t as defined in Definition 4. Fix any (α, K, p). Then w_t is decreasing in t, a_t is increasing in t, and hence both $\lim_{t\to\infty} w_t$ and $\lim_{t\to\infty} a_t$ exist. Consequently, the limit $\lim\limits_{t\to\infty} K w_t q_t$ exists.*

This completes our description of the three key components of the proof of Theorem 1. All proofs can be found in Appendices B and D in the e-companion, including the proof of Theorem 1.

5 Proof Idea of Theorem 2

So far, to analyze an agent's state after some rounds of partner search, we relied on one technical lemma (Lemma 1) to establish the local weak convergence between the agent's local neighborhood and some independent branching process. However, the established local weak convergence result can at most reach an agent's $o(\log N)$-depth neighborhood, and does not suffice for us to identify her final deadlock structure. That is to say, an agent's $o(\log N)$-depth neighborhood is not deep enough for the chain of waiting to form a cycle, hence causing her to wait forever. In fact, a typical deadlock (which lasts forever) requires tracking a chain of waiting of length in the order of $\sqrt{N}$. The intuition is roughly similar to the birthday paradox, where a group of $\Theta(\sqrt{N})$ people are needed to have at least one pair of people with the same birthday. In this case, we need the chain of waiting to close on itself forming a cycle, i.e., we need a member of the chain to be waiting for an earlier member of the chain, as illustrated in Figure 7. To tackle this challenge, we must use an entirely different approach. In particular, we will try to directly identify an agent's final deadlock structure at the end of the partner search process, and we will do so via a depth-first-search-like neighborhood revelation procedure. Moreover, by revealing only some partial information of her neighborhood structure instead of the full information, we can introduce less correlation to the revealed structure. This enables us to analytically track an agent's relevant neighborhood up to depth $\Theta(\sqrt{N})$ by coupling it with a branching process which we call the branching process of preferred partners.

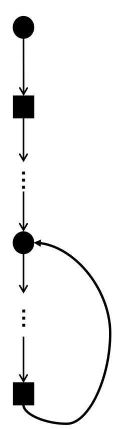

Figure 7: A chain of waiting that closes on a cycle. Directed edges point to whom each node is waiting for. Circles and squares represent women and men.

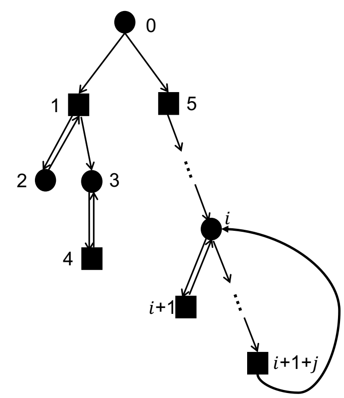

Figure 8: A focal node's final deadlock structure revealed by ALG-G at $t \to \infty$. A node's offsprings from left to right are ranked from high to low in its preference list. The number besides the node marks the time of its revelation, where $i > 5, j > 0$.

We now present the key steps towards the proof of Theorem 2.

A neighborhood revelation procedure. We first describe a neighborhood revelation procedure ALG-G, which selectively reveals information about the local neighborhood of a focal agent to identify the agent's final deadlock structure, if any. The idea is that ALG-G iteratively reveals the chain of agents that are being waited for when the partner search process terminates.

We introduce some additional notation to facilitate our definition of ALG-G. For a node i, let $\theta_i \in \{w, m\}$ be its type where $\theta_i = w$ indicates that node i is a woman node and $\theta_i = m$ indicates that node i is a man node; let $d_i \in \{0, 1, 2, ...\}$ be its actual degree; and for all $j \in \mathcal{N}(i)$ let $r_i(j) \in \{1, 2, ..., d_i\}$ be the rank of node j in the preference list of i, with rank 1 denoting the most preferred neighbor. Let $r_i(\emptyset) = d_i + 1$. Let $A_l(i)$ be the level l ancestor of i, with $A_1(i)$ being i's parent. We denote the degree sequence of the nodes in the graph by $(d_i)_{i=1}^N$ and $(d'_j)_{j=1}^{\alpha N}$, where the former (latter) are the degrees of all women (men) nodes.

Prior to the start of ALG-G, only a single piece of information about G is revealed: the degree sequence. Figure 9 describes the ALG-G process stepwise. Observe that ALG-G visits a subset of nodes in the focal node's neighborhood like a depth-first search. Its route follows the chain of waiting (see Figure 8) from the focal node at each step of partner search, while the current chain of waiting is always given by the set $\mathcal{D}_t$. Whenever the chain of waiting forms a cycle, ALG-G terminates by revealing the same node in $\mathcal{D}_t$ twice.

Definition 6 (Deadlock event). *Let $\mathcal{A}_t$ be the event that ALG-G terminates at time t and* deadlock = TRUE.

We are interested in the probability of event $\bigcup_{t\geq 0} \mathcal{A}_t$, which is the probability of the focal node waiting forever. In what follows, we will lower bound this probability using $\bigcup_{t\in[0,\tau]} \mathcal{A}_t$ for some $\tau = \Theta(\sqrt{N})$.

Coupling with a tree process. To understand what happens in ALG-G which runs on a bipartite graph $G_N(\alpha, K, p)$, we couple it with its counterpart ALG-T that runs on a tree closely related to $G_N(\alpha, K, p)$. We define ALG-T and the associated tree (generated by a branching process) below. We slightly modify Definition 5 for a marked Galton-Watson tree to generate an associated tree. In particular, instead of using Poisson distributions to generate offsprings for each node, we use the empirical distribution η_w and η_m from the degree sequence $(d_i)_{i=1}^N$ and $(d'_j)_{j=1}^{\alpha N}$ of $G_N(\alpha, K, p)$,

$$\eta_w(l) = \frac{\sum_{i=1}^N \mathbb{1}\{d_i = l\}}{N} \quad \text{and} \quad \eta_m(l) = \frac{\sum_{j=1}^{\alpha N} \mathbb{1}\{d'_j = l\}}{\alpha N}. \tag{5}$$

We first generalize Definition 5 as follows.

Definition 7 (Marked GW tree with degree distributions $(\eta_w(l))_{l=0}^\infty$ and $(\eta_m(l))_{l=0}^\infty$). *A marked GW tree $\tilde{\mathcal{T}} \equiv \tilde{\mathcal{T}}(\eta_w, \eta_m, p)$ is a random tree with woman and man nodes constructed as follows.*

- *The root is a woman node. The number of its man offspring nodes is drawn from the degree*

ALG-G on $G_N(\alpha, K, p)$ with degree sequence $(d_i)_{i=1}^N$ and $(d'_j)_{j=1}^{\alpha N}$

Inputs: Realized degree sequence $(d_i)_{i=1}^N$ and $(d'_j)_{j=1}^{\alpha N}$ of $G_N(\alpha, K, p)$.
A woman node i drawn uniformly at random (with known degree d_i)

Outputs:
Result $\in$ {DEADLOCK, CLEAR, INTERFERENCE}
Revealed marked neighborhood $(\mathcal{D}_t, \mathcal{U}_t, \mathcal{E}_t)$ of i including degree and rank of parent for revealed nodes, and inspection outcome for inspected edges

Initializations:
For convenience, set the rank of i's parent in its preference list $r_i(\emptyset) \leftarrow d_i + 1$
Active node $v(1) \leftarrow i$; active set $\mathcal{D}_0 \leftarrow \{i\}$; inactive set $\mathcal{U}_0 \leftarrow \emptyset$
Set of revealed directed edges $\mathcal{E}_0 \leftarrow \emptyset$
$A_1(i) \leftarrow \emptyset$, $r_i(\emptyset) \leftarrow d_i + 1$

while True **do**
    Initialization: $\mathcal{D}_t \leftarrow \mathcal{D}_{t-1}$; $\mathcal{U}_t \leftarrow \mathcal{U}_{t-1}$; $\mathcal{E}_t \leftarrow \mathcal{E}_{t-1}$
    while $\mathcal{D}_t \neq \emptyset$ and $|\{j : (v(t), j) \in \mathcal{E}_t\}| = r_{v(t)}(A_1(v(t))) - 1$ **do**
        ▷ Repeat while the active node $v(t) \neq i$ has no offspring that it prefers to its parent
        if $v(t) = i$ **then** ▷ Focal node i has reached the end of their list $|\{j : (v(t), j) \in \mathcal{E}_t\}| = d_i$
            Result $\leftarrow$ CLEAR
            break outer while
        end if
        The pair $(v(t), A_1(v(t)))$ inspects and the result is revealed
        Add edge $\mathcal{E}_t \leftarrow \mathcal{E}_t \cup \{(v(t), A_1(v(t)), \text{Inspection outcome})\}$ ▷ Note that the inspection outcome is included
        if inspection $(v(t), A_1(v(t)))$ succeeds **then**
            The pair matches and both nodes become inactive: $\mathcal{D}_t \leftarrow \mathcal{D}_t \setminus \{v(t), A_1(v(t))\}, \mathcal{U}_t \leftarrow \mathcal{U}_t \cup \{v(t), A_1(v(t))\}$
            The grandparent becomes the active node $v(t) \leftarrow A_2(v(t))$
        else
            The child becomes inactive: $\mathcal{D}_t \leftarrow \mathcal{D}_t \setminus \{v(t)\}, \mathcal{U}_t \leftarrow \mathcal{U}_t \cup \{v(t)\}$, and the parent node becomes the new active node: $v(t) \leftarrow A_1(v(t))$
        end if
    end while
    if $\mathcal{D}_t = \emptyset$ **then**
        Result $\leftarrow$ CLEAR
        break
    end if
    Consider j that satisfies $r_{v(t)}(j) = |\{j' : (v(t), j') \in \mathcal{E}_t\}| + 1$ ▷ Reveal the active node's next most preferred offspring. Note that such a j must exist since $\mathcal{D}_t \neq \emptyset$ and $|\{j : (v(t), j) \in \mathcal{E}_t\}| < r_{v(t)}(A_1(v(t))) - 1$.
    if $j \in \mathcal{D}_t$ **then**
        $\mathcal{E}_t \leftarrow \mathcal{E}_t \cup \{(v(t), j)\}$
        Result $\leftarrow$ DEADLOCK
        break
    else if $j \in \mathcal{U}_t$ **then**
        $\mathcal{E}_t \leftarrow \mathcal{E}_t \cup \{(v(t), j)\}$
        Result $\leftarrow$ INTERFERENCE
        break
    else
        Reveal j (including its type θ_j, its degree d_j, the rank of its parent in its preference list $r_j(A_1(j))$)
        $\mathcal{D}_t \leftarrow \mathcal{D}_t \cup \{j\}$; $\mathcal{E}_t \leftarrow \mathcal{E}_t \cup \{(v(t), j)\}$; $v(t+1) \leftarrow j$
        Advance time: $t \leftarrow t + 1$
    end if
end while
return Result and $(\mathcal{D}_t, \mathcal{U}_t, \mathcal{E}_t)$

Figure 9: The ALG-G process.

distribution η_w: it has l man offspring nodes with probability $\eta_w(l)$, for $l = 0, 1, 2, \ldots$.

- *Each non-root woman node, independently, draws offsprings from the women's edge-perspective degree distribution: it has l man offsprings with probability $\frac{(l+1)\eta_w(l+1)}{\sum_{j=0}^{\infty}(j+1)\eta_w(j+1)}$, for $l = 0, 1, 2, \ldots$.*
- *Each man node, independently, draws offsprings from the men's edge-perspective degree distribution: it has l woman offsprings with probability $\frac{(l+1)\eta_m(l+1)}{\sum_{j=0}^{\infty}(j+1)\eta_m(j+1)}$, for $l = 0, 1, 2, \ldots$.*
- *Each woman (man) node ranks its neighboring man (woman) nodes according to an independent uniformly random permutation.*
- *Each edge has a latent inspection outcome, which is Bernoulli(p), independently drawn.*

Then the *associated tree* of $G_N(\alpha, K, p)$ with empirical degree distributions η_w and η_m given by (5) is the marked GW tree with degree distributions η_w and η_m.

We define ALG-T to be the same process as ALG-G, but on the associated tree of $G_N(\alpha, K, p)$. We add to the notation $\mathcal{D}_t$ a superscript G or T to indicate whether it comes from ALG-G or ALG-T. Note that nodes no longer have the possibility of appearing twice in the associated tree, so the conditions in line 36 and 40 of ALG-G are never true, and so ALG-T never suffers DEADLOCK or INTERFERENCE. It either terminates with the result CLEAR (i.e., the root node gets matched or reaches the end of their list) or the active set remains non-empty and the tree continues forever, revealing an infinite chain of waiting.

ALG-G and ALG-T each produce a stochastic process of information revealed with t: Let $\mathcal{P}_t^G$ track all the information revealed up to the end of time t of ALG-G, with $\mathcal{P}_0^G$ including the degree sequences $(d_i)_{i=1}^N$ and $(d'_j)_{j=1}^{\alpha N}$. We let $\mathcal{P}_{t+1}^G = \mathcal{P}_t^G$ if ALG-G terminates before t. Similarly, let $\mathcal{P}_t^T$ track all the information revealed up to the end of time t of ALG-T, again with $\mathcal{P}_0^T$ including the degree distribution information η_w, η_m. We find it useful to couple the two stochastic processes maximally as follows: Note that $\mathcal{P}_t^G$ includes the identity information of each revealed node up to time t, according to which the deadlock (or interference) event can be identified when the same node in the active (inactive) set is revealed twice. In contrast, the revealed nodes in $\mathcal{P}_t^T$ do not have identities (every node revealed is distinct from previous nodes revealed). To define the proper coupling between ALG-G and ALG-T, we hence define a condensed version of $\mathcal{P}_t^G$, called $\tilde{\mathcal{P}}_t^G$, which erases the identity of each revealed node from $\mathcal{P}_t^G$. Note that conditional on the history $\{\mathcal{P}_s^G\}_{0\leq s\leq t}$ and $\{\mathcal{P}_s^T\}_{1\leq s\leq t}$, the realizations of $\mathcal{P}_{t+1}^G$ and $\mathcal{P}_{t+1}^T$ each follow some probability distribution. We create a coupling between the two processes, in a sequential manner starting from $t = 0$, such that the next-step realization of $\tilde{\mathcal{P}}^G_{t+1}$ and $\mathcal{P}^T_{t+1}$ are maximally coupled for each t and each possible history such that $\tilde{\mathcal{P}}^G_t = \mathcal{P}^T_t$. Define the natural filtration $\mathcal{F} = \{\mathcal{F}_t\}_{t\geq 0}$ with respect to the coupled process $\{\mathcal{P}^G_t, \mathcal{P}^T_t\}_{t\geq 0}$.

In order for the coupling between ALG-G and ALG-T to persist up to $\Theta(\sqrt{N})$ time with nontrivial probability, we also need some regularity conditions on the degree sequence $(d_i)_{i=1}^N$ and $(d'_j)_{j=1}^{\alpha N}$, which control properties of the empirical degree distributions such as the average degree, the probability generating function of the degree distributions, the extinction probability of the associated tree, etc. Details can be found in Definition 8 in Appendix E. We show that the degree sequence of $G_N(\alpha, K, p)$ is regular with high probability as $N \to \infty$ (Lemma 13). Therefore we can focus on only degree sequence realizations that are regular. Let $\mathcal{R}$ be the event that the degree sequence of $G_N(\alpha, K, p)$ is regular.

We next establish a coupling result in each step t, conditional on successful coupling in all previous steps, event $\mathcal{R}$, and a regularity condition on $\mathcal{P}^T_t$ that is satisfied with high probability conditional on $\mathcal{R}$. In the rest of the proof, we assume inspection success probability $p = 0$.

Lemma 6 (One-step coupling). *Let $z^T(t)$ be the sum of the actual degree of all revealed nodes in $\mathcal{P}^T_t$. Define event*

$$\mathcal{H}_t := \{z^T(t) \leq (t+1)\overline{W}\}$$

where $\overline{W} \in (0, \infty)$ is a constant. There exists a constant $\epsilon_c \equiv \epsilon_c(K, \alpha, \overline{W}) \in (0, \infty)$ such that, for any constant C, there exists $N_0 \equiv N_0(C, K, \alpha) \in (0, \infty)$ such that for any $N \geq N_0$ and $0 \leq t \leq C\sqrt{N}$,

$$\mathbb{P}(\tilde{\mathcal{P}}^G_0 \neq \mathcal{P}^T_0 | \mathcal{R}) = 0\,, \quad \mathbb{P}\left(\tilde{\mathcal{P}}^G_{t+1} \neq \mathcal{P}^T_{t+1} \middle| \mathcal{F}_t \text{ such that } \tilde{\mathcal{P}}^G_t = \mathcal{P}^T_t, \mathcal{H}_t, \mathcal{R}\right) \leq \frac{\epsilon_c t}{N}.$$

Lower-bounding the probability of waiting forever. Recall that we are interested in establishing a lower bound on the size of the final deadlock. This requires us to lower-bound the probability of a focal node waiting forever. A key step to show this is to establish, at each time $t = o(N)$, a lower bound of $\frac{\delta_1 t}{N}$, for some $\delta_1 > 0$, on the probability that ALG-G will identify the final deadlock structure involving the focal node at time $t+1$, conditional on that the following has occurred up to time t: (i) the coupling between ALG-G and ALG-T has held up, (ii) ALG-T has not terminated yet, and (iii) $\mathcal{P}_t^T$ satisfies a certain regularity condition. This result is stated in the next lemma.

Lemma 7 (One-step deadlock). *Consider $\mathcal{P}_t^T$ and the node s in it who will reveal the next offspring. Let $d_A^{T,\text{unre}}(t)$ be the total number of unrevealed edges of s's ancestor (excluding the parent) nodes with the opposite type to s (recall that node types are binary: there are man nodes and woman nodes) in the active set in $\mathcal{P}_t^T$. Define two events*

$$\mathcal{C}_t := \{\mathcal{D}_t^T \neq \emptyset\}, \quad \mathcal{J}_t := \{d_A^{T,\text{unre}}(t) \geq \delta_0 t\},$$

where $\delta_0 > 0$ is a constant, and $\mathcal{D}_t^T$ is $\mathcal{D}_t$ under ALG-T at the end of time t. Then, there exists a constant $\delta_1 \equiv \delta_1(\delta_0, K, \alpha) > 0$, such that for any $t \geq 0$,

$$\mathbb{P}(\mathcal{A}_{t+1} | \mathcal{F}_t \text{ such that } \tilde{\mathcal{P}}_t^G = \mathcal{P}_t^T, \mathcal{J}_t, \mathcal{C}_t, \mathcal{R}) \geq \frac{\delta_1 t}{N}.$$

The crucial event in the events conditioned on is $\mathcal{C}_t$. In order for the chain of waiting to be long enough to form a cycle, we need $\mathcal{C}_t$ to happen with a nonvanishing probability for $t = \omega(1)$. This crucially depends on a GW subtree of the associated GW tree in Definition 7 being supercritical or not. In this GW subtree, only each node's preferred offsprings (compared to its parent) are included, hence we referred to the subtree as the *branching process of preferred partners* in our informal summary in Section 3.2. As $N \to \infty$, the branching factor of this GW subtree converges to $\frac{K^2}{4\alpha}$. Therefore, the GW subtree is supercritical with high probability if and only if $\frac{K^2}{4\alpha} > 1$, in which case there is a positive probability that $\mathcal{C}_t$ happens for all $t \geq 0$, i.e., the chain of waiting goes on forever, providing opportunities for a deadlock to form via cycle formation in ALG-G, as long as the coupling between ALG-G and ALG-T lasts. Event $\mathcal{J}_t$ ensures that nodes in the active set have enough unrevealed edges to provide opportunities for cycle formation. We later show (see Lemma 19) that the chance of $\overline{\mathcal{J}_t}$ is small, provided that the degree sequence is regular and that ALG-T does not terminate for a long time.

The next lemma establishes the key condition on model primitives for deadlock formation.

Lemma 8. *If $\frac{K^2}{4\alpha} > 1$, then $\lim_{N\to\infty} \mathbb{P}(\mathcal{R}) = 1$ and $\mathbb{P}(\mathcal{C}_t | \mathcal{R}) \geq q$, $\forall t \geq 1$ for some $q \equiv q(\alpha, K) > 0$.*

We argued above that in order to create enough opportunities in ALG-G for deadlock formation, it is sufficient for the chain of waiting in the coupled ALG-T process to go on forever. Lemma 8 shows that the latter happens with positive probability when the effective branching factor of the GW subtree is above one. We include this requirement of the effective branching factor larger than one in the regularity conditions. Recall that we are using the empirical degree distributions for the GW tree. When $\frac{K^2}{4\alpha} > 1$, with high probability, the branching factor of the GW subtree is larger than one.

Lemmas 6–8 together lead to the next key lemma, which establishes the probability lower bound of the focal node waiting forever.

Lemma 9. *If $\frac{K^2}{4\alpha} > 1$, then there exist constants $C \equiv C(K, \alpha) < \infty, \varrho \equiv \varrho(K, \alpha) > 0, \underline{N} \equiv \underline{N}(K, \alpha) < \infty$, such that, for all $N \geq \underline{N}$,*

$$\mathbb{P}\left(\bigcup_{t\geq 0} \mathcal{A}_t\right) \geq \mathbb{P}(\exists 0 \leq t \leq C\sqrt{N} : \tilde{\mathcal{P}}_t^G = \mathcal{P}_t^T, \mathcal{A}_{t+1} = 1) \geq \varrho.$$

By Lemma 8, the assumption that $\frac{K^2}{4\alpha} > 1$ ensures that the regularity conditions $\mathcal{R}$ is satisfied with high probability. Therefore we can apply Lemmas 6 and 7 to infer that, for each $t \leq C\sqrt{N}$, the probability of ALG-G terminating with a deadlock is $\Omega(1)$ times the probability of it terminating due to a breakdown in coupling if the tree process does not terminate. Moreover, if the tree process does not terminate $\mathcal{C}_\infty$ and the active stack has sufficiently large unrevealed degree $\mathcal{J}_t$, the total probability of terminating due to deadlock before time $C\sqrt{N}$ is $\sum_{t=1}^{C\sqrt{N}} \Omega(t/N) = \Omega(1)$ by Lemma 7. Other than using Lemmas 6, 7 and 8, the full proof of Lemma 9 involves additional details involving concentration results to regulate the structure of $\mathcal{P}_t^T$. We present the full proof in Appendix E.

With Lemma 9 in place, the proof of Theorem 2 is fairly straightforward. We provide the formal proof in Appendix E.4.

6 Numerics

In this section, we present various numerics to supplement the theoretical results in Section 3.

DE and information deadlock. As Theorem 1 establishes, the fraction of women in information deadlock in random markets is described by $\lambda(\alpha, K, p) = \lim_{t\to\infty} K w_t q_t$ in Definition 4. Given any market primitives (α, K, p), we can compute this limit $\lim_{t\to\infty} K w_t q_t$ easily using Matlab. The two figures in Figure 10 show the time evolution of three message densities, the accept message density a_t, the reject message density r_t and the wait message density w_t, one figure each for a point in the deadlock-free regime and a point in the deadlock regime, respectively. The two cases differ in the wait message's density in the time limit. In particular, for $(\alpha = 1, K = 2, p = 3) \notin \Theta$ in the deadlock-free regime (the left figure), the wait message has density $w_\infty = 0$ in the limit. On the other hand, for $(\alpha = 1, K = 3, p = 3) \in \Theta$ in the deadlock regime, the wait message has density $w_\infty > 0$ in the limit. In fact, the value of w_∞ crucially indicates whether the market is in the deadlock-free regime or the deadlock regime (note that $\lambda(\alpha, K, p) = \lim_{t\to\infty} K w_t q_t$).

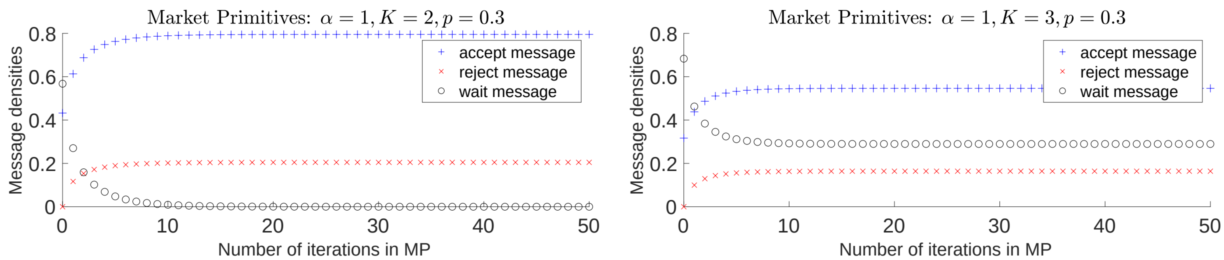

Figure 10: The evolution of message densities for the accept, reject and wait messages under different primitives. In the left figure we fix the men-to-women ratio $\alpha = 1$, the women's average degree $K = 2$, and the inspection's success probability $p = 0.3$. In the right figure we fix the men-to-women ratio $\alpha = 1$, the women's average degree $K = 3$, and the inspection's success probability $p = 0.3$.

The value of $\lambda(\alpha, K, p)$ is plotted in Figure 11. One can clearly see its monotonicity from the plots. In particular, the left figure plots the size of the deadlock $\lambda(\alpha, K, p)$ versus the men-to-women ratio $\alpha \in (0, 2]$ for different values of women's average degree K while fixing the inspection success probability $p = 0.3$. As the men-to-women ratio α increases, there are more men in the market, and hence the women face less competition. As a result, the size of the deadlock $\lambda(\alpha, K, p)$ decreases. The figure also shows that the size of the deadlock is increasing in the women's average degree K; as women's consideration sets grow larger, the market is more likely to have congestion since fewer guaranteed inspections are available. The right figure plots $\lambda(\alpha, K, p)$ versus $\alpha \in (0, 2]$ for different values of p, while fixing $K = 3$. It shows that the size of the deadlock decreases in the inspection's success probability p. This is because compared with a failed inspection, a successful inspection "peels" the market to a greater extent, and hence reduces market congestion more. Also, both figures in Figure 11 show that the size of the deadlock decreases with the men-to-women ratio α. The results suggest a phase transition with a sharp threshold $\alpha^*(K, p)$ between the information deadlock regime and the deadlock-free regime. Finally, notice that the information deadlock regime is prevalent under various reasonable market primitives. For example, consider a market where the number of women equals the number of men. Each woman, on average, connects to three men

$K = 3$. In this case, if each inspection has equal chance to succeed or fail, then w.h.p., more than 30% women will be stuck in an information deadlock if they only wish to perform guaranteed inspections. These women still have men in their consideration set who they have not ruled out, but are unwilling to inspect them. The size of the deadlock is larger still if women on average are connected to more men, or if the inspection is more likely to fail, or if there are less men in the market.

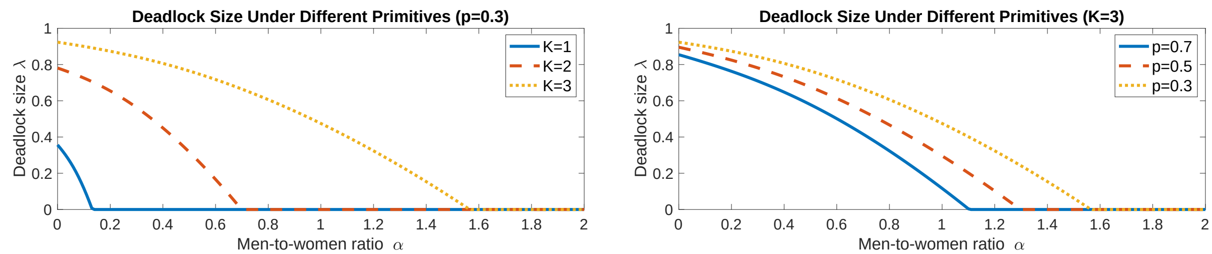

Figure 11: Deadlock size λ^* under different primitives. In the left figure we fix the inspection's success probability $p = 0.3$. In the right figure we fix the women's average degree $K = 3$.

We state these monotonicity observations in the next conjecture[12].

Conjecture 2. *The value* $\lambda(\alpha, K, p)$ *in Definition 4 is a decreasing function of* α *and* p*, and an increasing function of* K*. For any given* $K > 0$ *and* $p \in (0, 1]$, $\lim_{\alpha \to 0^+} \lambda(\alpha, K, p) > 0$ *whereas* $\lambda(\alpha_0, K, p) = 0$ *for some* $\alpha_0 < \infty$*. Also, for any given* $\alpha > 0$ *and* $p \in (0, 1]$, $\lim_{K \to \infty} \lambda(\alpha, K, p) > 0$ *whereas* $\lambda(\alpha, K_0, p) = 0$ *for some* $K_0 > 0$*.*

As α grows, there are more men per woman, and the same consideration set size K for women, so less women end up stuck. As p grows, guaranteed inspections succeed more often and hence clear the market more effectively, resulting in less women being stuck. As K grows, women consider more men, but the the total number of men are unchanged, hence the market gets more congested[13].

The phase transition between the deadlock regime and the deadlock-free regime in the 3-dimensional space of market primitives can be better observed in Figure 3. Specifically, in Figure 3, the three axes correspond to the three market primitives α, K and p. There is a separating surface such that the region above it corresponds to the deadlock regime, and the region below it corresponds to the deadlock-free regime. One can see from the plot that if a market suffers

[12] One might expect that the conjecture can be proved using some comparison theorems. However, some key quantities turn out to be non-monotonic in the market primitive. For example, ν_{t-1} first increases then decreases in K. This in turn makes the inductive argument of a_t's monotonicity difficult, even though we numerically observe that a_t is decreasing in K.

[13] On the other hand, if both K and α decreases such that K/α is unchanged, the effect on deadlock is similar to when K is fixed and α increases, i.e., information deadlock decreases.

from information deadlock, then the market designer can try to mitigate the situation by adjusting the market primitives. For example, for a market with given inspection success probability p, the market designer can promote market clearing by restricting the size of womens' consideration sets K (hence also restricting the size of mens' consideration sets), or by increasing the men to women ratio α (while holding K fixed).[14].

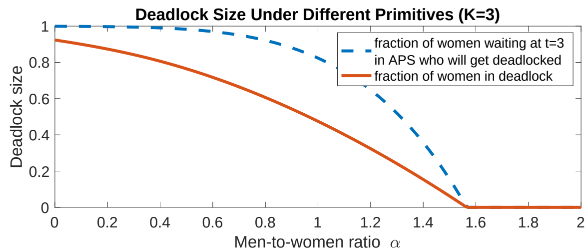

Figure 12: Deadlock size and the conditional probability of being in deadlock for women waiting after three rounds of APS, respectively, as a function of the men-to-women ratio α. We fix the inspection success probability $p = 0.3$ and the women's average degree $K = 3$.

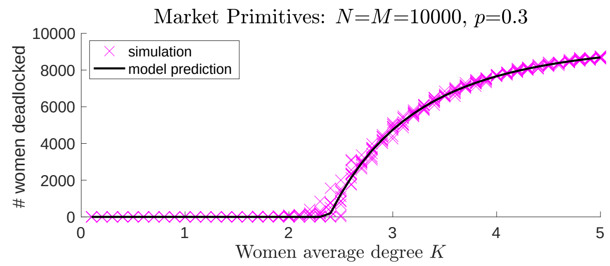

Figure 13: Number of women who get stuck after PS terminates in randomly generated markets versus $\lambda(\alpha, K, p)N$, under varying women's average degree K. We fix the total number of women and men to be $N = M = 10,000$ and an inspection's success probability to be $p = 0.3$.

We also compare the number of women waiting after a short number of periods only, to the number of women who remain waiting as time $\to \infty$. This gives the conditional probability of waiting indefinitely for a woman who is waiting after, for example, three rounds of an accelerated version of partner search (termed APS, see Algorithm 1 in Appendix B) that allows multiple consecutive guaranteed inspections to be conducted in the same period until one succeeds. We plot this conditional probability in the blue curve in Figure 12. (The fraction of women waiting in period t of APS is Kw_tq_t, where w_t and q_t are as defined in Definition 4.) This figure shows that, in a market with the same number of men and women, where an agent's average degree is three and the inspection success probability is 30%, if a woman is waiting after three rounds of APS, then there is an 82% chance she will get stuck in deadlock. Based on such conditional probability estimates, a platform can encourage agents to inspect inferior options if they are still waiting after a certain number of periods.

Lastly, we are interested in comparing the asymptotic quantities computed using message densities with quantities observed in simulations of finite random markets. As an illustrative example, we present results comparing $N\lambda(\alpha, K, p)$ with the number of women who get stuck after running

[14]Note that if we hold fixed the *mens'* average degree instead, the size of the deadlock will increase with α.

the Partner Search process until convergence in randomly generated markets in Figure 13. We can see that in random markets with 10,000 women and men, the simulated number of women who remain waiting after the PS process terminates in randomly generated markets concentrates tightly around the predicted size of information deadlock. This strongly suggests that the characterized deadlock size at $\omega(1)$ and $o(\log N)$ time of PS in Theorem 1 does not resolve over time but persists till $t \to \infty$.

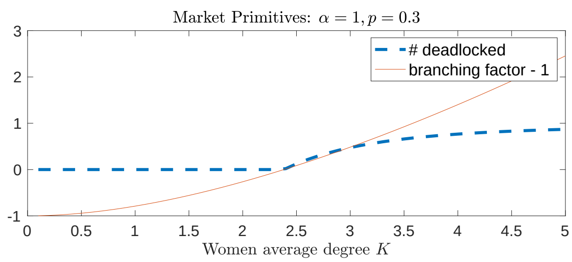

Figure 14: Deadlock versus the branching factor, as functions of the women's average degree K. We fix the inspection success probability $p = 0.3$ and the men-to-women ratio $\alpha = 1$. The blue curve plots the deadlock size. The red curve plots the branching factor (in Conjecture 1) minus 1.

Branching factor and information deadlock. We also compare the branching factor in Conjecture 1 with the deadlock size characterization in Theorem 1. As conjectured in Conjecture 1, the market exhibits large-scale deadlock if and only if the branching factor of some associated branching process exceeds 1. In Figure 14, we plot both the fraction of women in deadlock and the branching factor (as in Conjecture 1) minus 1 for different market primitives. One can see that the phase transition between the deadlock-free regime and the deadlock regime happens at the point where the branching factor exceeds one, which numerically validates Conjecture 1.

Additional asymptotics from DE. Our framework allows us to similarly obtain asymptotic estimates of *any* market-level ("macroscopic") quantity of interest, after any number of time steps, as a function of market primitives, using the message densities in Definition 4 and our characterization leading to Theorem 1. For example, we can similarly calculate the fraction of women who successfully get matched to be $\lim_{t\to\infty} Kpa_tq_t$, where a_t and q_t are as defined in Definition 4. Note that a woman does not get matched for one of two reasons (i) she is stuck in information deadlock, or (ii) her consideration set has been depleted after unsuccessful inspections (we call such an agent "single"). We present these results in Figure 15. Observe that the phase boundary between the deadlock regime and the deadlock-free regime maximizes the match rate.[15] This critical value of the

[15]Unfortunately we are unable to prove this due to the complicated form of the match rate. For example, we

women's average degree K balances between providing more match opportunities (increasing K) to women, and preventing deadlock which increases as the market gets more connected (decreasing K).

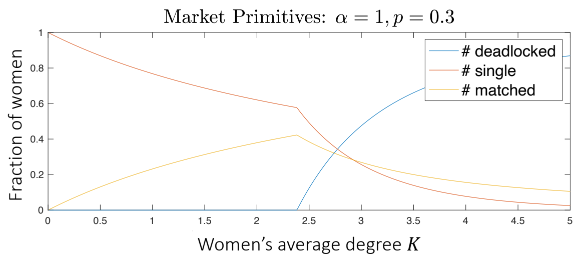

Figure 15: Fraction of women in information deadlock, fraction of women who exhaust their consideration set (singleton), and fraction of matched women as functions of women's average degree K, respectively. We fix the men-to-women ratio to be $\alpha = 1$ and an inspection's success probability to be $p = 0.3$.

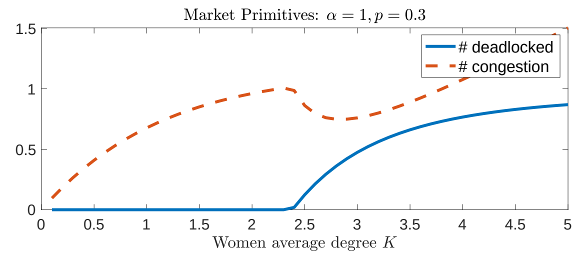

Figure 16: Deadlock versus congestion, as functions of the women's average degree K. We fix the inspection success probability $p = 0.3$ and the men-to-women ratio $\alpha = 1$. The blue curve plots the deadlock size. The red curve plots the congestion level (as measured by the average number of inspections of a woman).

Consider another quantity of interest: the congestion level. We measure the congestion level of the market by the expected number of inspections (including both failures and successes) experienced by a woman until some $o(\log N)$ and $\omega(1)$ time of the Partner Search process (as argued before, we have reason to strongly believe that this is also the measure at any $t = \omega(\log N)$, in particular, the end of the Partner Search). With some algebra work using the message densities, one gets

$$\mathbb{E}[\# \text{ inspections}] = \frac{a_\infty(1-p)(a_\infty p + w_\infty)}{(a_\infty(1-p) + r_\infty)^2} + \frac{a_\infty p}{a_\infty p + w_\infty} - \left(\frac{a_\infty(1-p)(a_\infty p + w_\infty)}{(a_\infty(1-p) + r_\infty)^2} + a_\infty p \right) e^{-(a_\infty(1-p)+r_\infty)K}.$$

We plot in Figure 16 how this congestion varies under different market primitives, and compare it with the size of deadlock in the market. From the figure, one can see that as the women's average degree K increases, the congestion level increases until deadlock emerges, after which congestion first decreases and then increases. The behavior of the congestion level after deadlock emerges is interesting. In particular, a small to moderate deadlock reduces the amount of inspections in the market, since people wait for each other instead of performing inspections. However, as deadlock becomes more severe, people also inspect more. This is bad for the market since participants inspect a lot but still eventually get stuck in the market.

numerically observe that a_t decreases in K, and Kq_t first increases (until K hits the deadlock phase transition threshold) then decreases in K. However, it is difficult to show, when K is small in the deadlock-free regime, that the force from Kq_t increasing in K dominates the opposite force from a_t decreasing in K.

Extensions of our model. In Appendix H, we discuss two extensions of the base model in Section 2. First, we consider vertical differentiation between agents, i.e., instead of assuming all agents are ex ante homogeneous, we assume a certain fraction of them are more preferred than the rest. In the second extension, we model agents on one side of the market as being willing to conduct parallel inspections with their current top two potential partners, thus expanding the set of inspections which can be conducted. We show that our analysis framework can still be applied to these extensions with proper modifications, and that we can still characterize the size of information deadlock in these markets. Moreover, we show that introducing vertical differentiation and allowing parallel inspections help reduce deadlock.

7 Discussion

In this paper, we investigated in which markets there is an information deadlock, where a large number of agents wait for each other to acquire information first. We looked into random markets with three key primitives: the men-to-women ratio α, the womens' average size of consideration set K, and an inspection's success probability p. We also imposed a behavioral assumption on agents' inspection decisions that only guaranteed inspections are performed. An inspection becomes guaranteed when the pair both considers each other as their best available option. We study the limit as the number of women N goes to infinity, and give specific characterization of the fraction of women stuck in information deadlock. This characterization is made possible via a technique called message passing. We show that information deadlock is prevalent in a wide range of markets. Furthermore, numerical evidence demonstrates a phase transition in each of the three market primitives that separates the information deadlock regime and the deadlock-free regime. A number of market design insights emerge from our characterization, for example, the market connectivity K which maximizes the number of matches formed is that which causes the market to be at the phase boundary between the deadlock-free regime and the deadlock regime. Vertical differentiation between agents reduces deadlock, as does a willingess by agents to perform parallel inspections.

References

[1] S Nageeb Ali and Ran I Shorrer. Hedging when applying: Simultaneous search with correlation. *American Economic Review*, 2024.

[2] Gad Allon, Achal Bassamboo, and Eren B Cil. Large-scale service marketplaces: The role of the moderating firm. *Management Science*, 58(10):1854–1872, 2012.

[3] Nick Arnosti. Lottery design for school choice. *Management Science*, 69(1):244–259, 2023.

[4] Nick Arnosti, Ramesh Johari, and Yash Kanoria. Managing congestion in matching markets. *Manufacturing & Service Operations Management*, 23(3):620–636, 2021.

[5] Jonathan Aronson, Alan Frieze, and Boris G Pittel. Maximum matchings in sparse random graphs: Karp–sipser revisited. *Random Structures & Algorithms*, 12(2):111–177, 1998.

[6] Itai Ashlagi, Mark Braverman, Yash Kanoria, and Peng Shi. Clearing matching markets efficiently: informative signals and match recommendations. *Management Science*, 66(5):2163–2193, 2020.

[7] Itai Ashlagi, Yash Kanoria, and Jacob D Leshno. Unbalanced random matching markets: The stark effect of competition. *Journal of Political Economy*, 125(1):69–98, 2017.

[8] Itai Ashlagi and Afshin Nikzad. What matters in school choice tie-breaking? how competition guides design. *Journal of Economic Theory*, 190:105120, 2020.

[9] Itai Ashlagi and Peng Shi. Improving community cohesion in school choice via correlated-lottery implementation. *Operations Research*, 62(6):1247–1264, 2014.

[10] Krishna B Athreya and Peter E Ney. *Branching processes*. Courier Corporation, 2004.

[11] A Brown. Nearly half of us adults say dating has gotten harder for most people in the last 10 years. *Washington, DC: Pew Research Center*, 2020.

[12] Peter Coles, Alexey Kushnir, and Muriel Niederle. Preference signaling in matching markets. *American Economic Journal: Microeconomics*, 5(2):99–134, 2013.

[13] Colin Cooper, Alan Frieze, Michael Molloy, and Bruce Reed. Perfect matchings in random r-regular, s-uniform hypergraphs. *Combinatorics, Probability and Computing*, 5(1):1–14, 1996.

[14] Bnaya Dreyfuss, Ori Heffetz, and Matthew Rabin. Expectations-based loss aversion may help explain seemingly dominated choices in strategy-proof mechanisms. *American Economic Journal: Microeconomics*, 14(4):515–555, 2022.

[15] Richard Durrett. *Random graph dynamics*, volume 200. Cambridge university press Cambridge, 2007.

[16] Apostolos Filippas, John Horton, Prasanna Parasurama, and Diego Urraca. Costly capacity signaling increases matching efficiency: Evidence from a field experiment. In *Proceedings of the 25th ACM Conference on Economics and Computation*, pages 414–415, 2024.

[17] Andrey Fradkin. Search frictions and the design of online marketplaces. *Work. Pap., Mass. Inst. Technol*, 2015.

[18] William T Freeman, Egon C Pasztor, and Owen T Carmichael. Learning low-level vision. *International journal of computer vision*, 40(1):25–47, 2000.

[19] Brendan J Frey and David JC MacKay. A revolution: Belief propagation in graphs with cycles. In *Advances in neural information processing systems*, pages 479–485, 1998.

[20] Karl J Friston, Thomas Parr, and Bert de Vries. The graphical brain: belief propagation and active inference. *Network Neuroscience*, 1(4):381–414, 2017.

[21] Arpita Ghosh and Robert Kleinberg. Optimal contest design for simple agents. *ACM Transactions on Economics and Computation (TEAC)*, 4(4):1–41, 2016.

[22] Sanford J Grossman and Joseph E Stiglitz. On the impossibility of informationally efficient markets. *The American economic review*, 70(3):393–408, 1980.

[23] Hanna Halaburda, Mikołaj Jan Piskorski, and Pınar Yıldırım. Competing by restricting choice: The case of matching platforms. *Management Science*, 64(8):3574–3594, 2018.

[24] Theodore Edward Harris. *The theory of branching processes*, volume 6. Springer Berlin, 1963.

[25] Avinatan Hassidim, Assaf Romm, and Ran I Shorrer. The limits of incentives in economic matching procedures. *Management Science*, 67(2):951–963, 2021.

[26] John William Hatfield, Fuhito Kojima, and Scott Duke Kominers. Strategy-proofness, investment efficiency, and marginal returns: An equivalence. *Becker Friedman Institute for Research in Economics Working Paper*, 2018.

[27] John William Hatfield and Paul R Milgrom. Matching with contracts. *American Economic Review*, 95(4):913–935, 2005.

[28] John J Horton. Buyer uncertainty about seller capacity: Causes, consequences, and a partial solution. *Management Science*, 65(8):3518–3540, 2019.

[29] John J Horton, Ramesh Johari, and Philipp Kircher. Cheap talk messages for market design: Theory and evidence from a labor market with directed search. Technical report, National Bureau of Economic Research, 2021.

[30] Nicole Immorlica, Jacob D Leshno, Irene Y Lo, and Brendan J Lucier. Information acquisition in matching markets: The role of price discovery. 2020.

[31] Nicole Immorlica, Brendan Lucier, Vahideh Manshadi, and Alexander Wei. Designing approximately optimal search on matching platforms. In *Proceedings of the 22nd ACM Conference on Economics and Computation*, pages 632–633, 2021.

[32] Nicole Immorlica and Mohammad Mahdian. Marriage, honesty, and stability. In *Proceedings of the sixteenth annual ACM-SIAM symposium on Discrete algorithms*, pages 53–62, 2005.

[33] Yash Kanoria and Daniela Saban. Facilitating the search for partners on matching platforms. *Management Science*, 67(10):5990–6029, 2021.

[34] Donald E Knuth, Rajeev Motwani, and Boris Pittel. Stable husbands. *Random Structures & Algorithms*, 1(1):1–14, 1990.

[35] David JC MacKay and Radford M Neal. Near shannon limit performance of low density parity check codes. *Electronics letters*, 32(18):1645–1646, 1996.

[36] Marc Mezard and Andrea Montanari. *Information, physics, and computation*. Oxford University Press, 2009.

[37] Marc Mézard, Giorgio Parisi, and Riccardo Zecchina. Analytic and algorithmic solution of random satisfiability problems. *Science*, 297(5582):812–815, 2002.

[38] Michael Molloy. Cores in random hypergraphs and boolean formulas. *Random Structures & Algorithms*, 27(1):124–135, 2005.

[39] Bobak Pakzad-Hurson. Crowdsourcing and optimal market design. *Available at SSRN 2618837*, 2019.

[40] Judea Pearl. Fusion, propagation, and structuring in belief networks. *Artificial intelligence*, 29(3):241–288, 1986.

[41] Judea Pearl. Probabilistic reasoning in intelligent systems: Networks of plausible inference. *Morgan Kauffmann Publishers Inc*, 1988.

[42] Marcin Peski. Large roommate problem with non-transferable random utility. *Journal of Economic Theory*, 168:432–471, 2017.

[43] Marcin Peski. Tractable model of dynamic many-to-many matching. *American Economic Journal: Microeconomics*, 2020.

[44] Boris Pittel. The average number of stable matchings. *SIAM Journal on Discrete Mathematics*, 2(4):530–549, 1989.

[45] Thomas Quint. The lattice of core (sub) matchings in a two-sided matching market. *Mathematics of Operations Research*, 19(3):603–617, 1994.

[46] Tom Richardson and Ruediger Urbanke. *Modern coding theory*. Cambridge university press, 2008.

[47] Alvin E Roth and Elliott Peranson. The redesign of the matching market for american physicians: Some engineering aspects of economic design. *American economic review*, 89(4):748–780, 1999.

[48] Alvin E Roth and Xiaolin Xing. Jumping the gun: Imperfections and institutions related to the timing of market transactions. *The American Economic Review*, pages 992–1044, 1994.

[49] Alvin E Roth and Xiaolin Xing. Turnaround time and bottlenecks in market clearing: Decentralized matching in the market for clinical psychologists. *Journal of political Economy*, 105(2):284–329, 1997.

[50] Roman M Sheremeta. Contest design: An experimental investigation. *Economic Inquiry*, 49(2):573–590, 2011.

[51] Peng Shi. Optimal matchmaking strategy in two-sided marketplaces. *Management Science*, 2022.

[52] Herbert A Simon. A behavioral model of rational choice. *The quarterly journal of economics*, 69(1):99–118, 1955.

[53] Herbert A Simon. Rational choice and the structure of the environment. *Psychological review*, 63(2):129, 1956.

[54] Aad W Van der Vaart. *Asymptotic statistics*, volume 3. Cambridge university press, 2000.

[55] Martin L Weitzman. Optimal search for the best alternative. *Econometrica: Journal of the Econometric Society*, pages 641–654, 1979.

[56] Jon Wellner et al. *Weak convergence and empirical processes: with applications to statistics.* Springer Science & Business Media, 2013.

[57] Jonathan S Yedidia, William T Freeman, and Yair Weiss. Constructing free-energy approximations and generalized belief propagation algorithms. *IEEE Transactions on information theory*, 51(7):2282–2312, 2005.

[58] Haichao Zheng, Dahui Li, and Wenhua Hou. Task design, motivation, and participation in crowdsourcing contests. *International Journal of Electronic Commerce*, 15(4):57–88, 2011.

E-companion for "In Which Matching Markets do Costly Compatibility Inspections Lead to a Deadlock?"

This online appendix is an e-companion for paper "In Which Matching Markets do Costly Compatibility Inspections Lead to a Deadlock?" and provides additional proofs and materials in supplement to it. This online appendix is subdivided into several sections. Appendix A discusses our search model and our model assumptions and situates them in the context of the related literature. Appendices B – D contain proofs of the base model, including Lemma 2 – 5 in Section 4, existence of the deadlock regime (Proposition 1), and Theorem 1. In particular, Appendix B contains proofs of lemmas regarding the MP algorithm and the PS process in Section 4, as well as a description of the APS algorithm. Appendix C shows existence of the deadlock regime (Proposition 1). Appendix D proves Lemma 1 and Theorem 1. Appendix E contains proofs for Theorem 2 and related supporting lemmas. Appendix F gives the heuristic argument for Conjecture 1. We give a heuristic derivation of the giant component regime using MP in Appendix G. Appendix H discusses two model extensions, including vertically differentiated agents and parallel inspections. Appendices I and J provide supplementary materials regarding the model extensions.

A Discussion of our search model and model assumptions

In this appendix section, we describe in detail the connection between our partner search process where agents perform satisficing search and the model of Immorlica et al [30] which introduces the notions of information deadlock and regret-free stable matching. In particular, we extend our model to include cardinal utilities and inspection costs as in the model of [30] and show that the resulting model reduces to our satisficing search model of agent behavior. We also detail the connection between our notion of guaranteed inspections and the regret-free stable matching notion of [30]. Subsequently, we discuss our modeling assumptions such as i.i.d. preferences and i.i.d. inspection outcomes and situate them in the context of the theoretical literature on matching markets.

Connection between Partner Search and regret-free stable matching. Immorlica et al. [30], who introduce the notions of information deadlock and regret-free stability in matching markets, explicitly model match values and inspection costs and insist agents perform optimal search among their option set. In such a setting, the optimal search among a subset of options can be calculated using Weitzman's index policy [55]. This policy assigns each option in the subset an index based on the value distribution and inspection cost of that option alone. The agent then sorts the options in decreasing order of index and inspects options in that order until finding one whose value is greater than the index of the following option. She then selects the inspected option with highest value (which may not be the most recently inspected option). Immorlica et al. [30] define an outcome to be regret-free stable if all agents optimally inspect the subset of options in their *budget set*, i.e., the set of matches that are attainable for them fixing the choices of others in the market.

In our paper, we consider a simplified "Partner Search" model in which agents perform satisficing search [53], i.e., given a fixed option set, agents wish to inspect options in decreasing preference order until they observe a success (as is the case in the running example in Immorlica et al, reproduced in Figure 1). To explore the optimality of this behavior, we extend our ordinal model to a cardinal one in which agent i's potential match partners have values, as opposed to just a preference ordering $\succ_i$. In particular, we assume agent i gets value $\tilde{v}_{ij} = v_{ij}$ from matching to agent j if agent j is compatible, and $\tilde{v}_{ij} = 0$ otherwise, which happens with probability p.[16] To

[16]Such a value distribution is a special case of the one considered in Immorlica et al. [30], but is the natural analogue of our ordinal model in which the agent knows the preference ordering conditional on compatibility with certainty.

relate this cardinal utility model to our base model, we assume the values $(v_{ij})_{j\in\mathcal{N}(i)}$ are strictly ordered, and furthermore assume that the Weitzmann index explained below $\sigma_j = v_{ij} - c/p$ are strictly positive for all $j \in \mathcal{N}(i)$ and take this order as the ordinal preferences of agent i over their consideration set in our base model. Agent i knows v_{ij} and p, and, at a cost of $c \geq 0$, can learn whether j is compatible (i.e., whether her true value for j is v_{ij} or 0).

Proposition 2. *Within the framework of [30], under the aforementioned model of cardinal utilities and inspection cost, an agent performs satisficing search in the same order as the one the agent follows under partner search in our base model.*

Proof. In the setting with cardinal utilities, the Weitzman indices tell us the order of inspection. The index σ_j of option j is, by definition, the solution to the equation $c = E[\max(\tilde{v}_{ij} - \sigma_j, 0)]$. For our binary-valued random variable $\tilde{v}_{ij}$, this expression has a closed form, namely $\sigma_j = v_{ij} - c/p$, implying that agent i inspects according to the decreasing order of v_{ij}. Furthermore, note that for $j \succ_i j'$, we have $\sigma_{j'} \leq v_{ij'} < v_{ij}$, and hence, the agent stops inspecting upon the first successful inspection, i.e., upon finding j such that $\tilde{v}_{ij} = v_{ij}$. In other words, agents perform satisficing search as per the ordinal preferences in our base model.

□

Above, we assumed that inspection costs were identical for all $j \in \mathcal{N}(i)$. More generally, consider costs c_j which may be different for different $j \in \mathcal{N}(i)$. We now assume the resulting order of the Weitzmann indices is strict and take this order as the ordinal preference of agent i in our base model (we now *do not* assume that the order of the $(v_{ij})_{j\in\mathcal{N}(i)}$ is identical to the preference order in our base model), then, once again we recover Proposition 2 by noticing that, for $j' \prec_i j$, we have $\sigma_{j'} < \sigma_j \leq v_{ij}$. However, if the two-point distribution of $\tilde{v}_{ij}$ is such that the low value is $\epsilon_j > 0$, then it is possible that agent i stops searching further even when the inspection for j fails. For example, if the parameters satisfy $v_{ij} - c_j/p > \epsilon_j > v_{ij'} - c'_j/p > \epsilon_{j'} > 0$ for some $j \in \mathcal{N}(i)$ and all other $j \in \mathcal{N}(i)$, then, according to the Weitzman indices, option j has the highest order, with its low reward realization higher than all other options' indices. In this case, agent i will match with j regardless of the inspection outcome.

For an outcome in our setting to be regret-free stable as defined in Immorlica et al. [30], agent i must therefore inspect options in this order among those in her budget set, i.e., among those that are attainable given the choices of others. This is precisely our definition of a guaranteed inspection. From this point of view, our results identify markets where any mechanism, centralized or not, that guarantees regret-free stability would get stuck (deadlock regime); versus markets in which a mechanism as simple as asking each agent to inspect their favorite remaining option (i.e., perform guaranteed inspections) would clear the market efficiently (deadlock-free regime).

Further discussion of our modelling assumptions. Our model of a random market, including i.i.d. uniform agent preferences is key to making our problem tractable, by allowing us to reveal information sequentially (sometimes called "the principle of deferred decisions") while still ensuring that the residual market is uniformly random conditioned on what has been revealed so far (e.g., see Lemma 20 in Appendix E.4). We note a rich tradition of theoretical advances in the field of matching markets based on models of random markets with i.i.d. uniform preferences, e.g., see influential papers by Pittel [44], Knuth, Motwani and Pittel [34], Ashlagi, Kanoria and Leshno [7], Roth and Peranson [47], and others. Such preferences are an extreme of a well-studied market structure (see, e.g., Immorlica and Mahdian [32]) where deadlocks could be problematic. Indeed, if preferences were perfectly correlated even on just one side of the market, there would be no deadlocks – the most-preferred agent would receive their most-preferred match and so on. In practice, preferences are thought to lie between these extremes. Similarly, our assumption of independent inspection success with probability p is necessary for tractability, allowing the outcome of an inspection to be independent of the information revealed so far. In practice, this does not always hold, and other literature studies implications of correlated success, see, for example, recent work by Ali and Shorrer [1].

We now discuss indirect empirical backing for our assumption that agents only perform guaranteed inspections. While there hasn't been direct evidence for such behavior by participants, there is related evidence for loss averse behavior by market participants in matching and other settings: Hassidim et al [25] find that participants in a truthful matching mechanism often report preferences which indicate that they prefer to not receive funding, even though the mechanism preserves privacy and there is no downside to receiving funding. One possible explanation they suggest is that participants may be trying to avoid the disappointment of asking for funding and not receiving it. Dreyfuss et al [14] document loss averse choice behavior more generally by participants in truthful mechanisms. Connecting back to our setting, one may expect that a market participant who is unwilling to even report a preference for a potential match if there is a risk of being rejected, might more so be unwilling to spend the effort to inspect such a potential match at risk of being later rejected by them, motivating our assumption that only guaranteed inspections are performed.

B The MP Algorithm and the PS Process

This section contains proofs of Lemma 2 – 5 regarding the MP algorithm and the PS process in Section 4.

Before providing the proofs, we provide a precise description of the accelerated partner search process in Algorithm 1. We use the term "best" to capture most preferred. We say a woman i is waiting after t' rounds of APS if $i \in \mathcal{I}_{t'}^{\mathrm{APS}}$.

Next, we prove Lemma 2 (MP tracks accelerated PS). In order to prove Lemma 2, we need the next supporting lemma on MP.

Lemma 10. *We have the following monotonicity properties on MP:*

1. *If* $\hat{m}_{j\to i}^{(t)} =A$*, then* $\hat{m}_{j\to i}^{(t')} =A$ *for all* $t' \geq t$*.*
2. *If* $\hat{m}_{j\to i}^{(t)} =R$*, then* $\hat{m}_{j\to i}^{(t')} =R$ *for all* $t' \geq t$*.*
3. *If* $I_{ij}^{(t)} = 1$*, then* $I_{ij}^{(t')} = 1$ *for all* $t' \geq t$*.*
4. *If* $m_{i\to j}^{(t)} =Y$*, then* $m_{i\to j}^{(t')} =Y$ *for all* $t' \geq t$*.*
5. *If* $m_{i\to j}^{(t)} =N$*, then* $m_{i\to j}^{(t')} =N$ *for all* $t' \geq t$*.*

The proof of Lemma 10 is obvious from the MP algorithm and is omitted here. Now we are ready to prove Lemma 2.

Proof of Lemma 2. We will inductively show the following claims hold.

1. (i, j) inspected by $t(1)$ in accelerated PS iff $\hat{m}_{j\to i}^{(t)} =$A and $I_{ij}^{(t-1)} = 1$.
2. (i, j) inspected by $t(2)$ in accelerated PS iff $\hat{m}_{j\to i}^{(t)} =$A and $I_{ij}^{(t)} = 1$.
3. j is waiting for i at $t(1)$ iff $m_{i\to j}^{(t-1)} =$U and $\hat{m}_{j\to i}^{(t)} =$A.
4. i is waiting for j at $t(2)$ iff $\hat{m}_{j\to i}^{(t)} =$W and $I_{ij}^{(t)} = 1$.

Algorithm 1 Accelerated Partner Search (APS)

```
Initialize the market: $\mathcal{I}_0^{\text{APS}} \leftarrow \emptyset$; $\mathcal{J}_0^{\text{APS}} \leftarrow \emptyset$;
for all $i \in \mathcal{I}_0$ such that $\mathcal{N}(i) \neq \emptyset$ do
    $\mathcal{I}_0^{\text{APS}} \leftarrow \mathcal{I}_0^{\text{APS}} \cup \{i\}$; $\mathcal{N}_0^{\text{APS}}(i) \leftarrow \mathcal{N}(i)$;
end for
for all $j \in \mathcal{J}_0$ such that $\mathcal{N}(j) \neq \emptyset$ do
    $\mathcal{J}_0^{\text{APS}} \leftarrow \mathcal{J}_0^{\text{APS}} \cup \{j\}$; $\mathcal{N}_0^{\text{APS}}(j) \leftarrow \mathcal{N}(j)$;
end for
$t' \leftarrow 1$
while $\mathcal{J}_{t'-1}^{\text{APS}} \neq \emptyset$ do    ▷ round $t'$
    $\mathcal{J}_{t'}^{\text{APS}} \leftarrow \mathcal{J}_{t'-1}^{\text{APS}}$; $\mathcal{I}_{t'}^{\text{APS}} \leftarrow \mathcal{I}_{t'-1}^{\text{APS}}$    ▷ Initialization. Will eliminate agents who match or reach the end of their list.
    $\forall j \in \mathcal{J}_{t'-1}^{\text{APS}}$ do $\mathcal{N}_{t'}^{\text{APS}}(j) \leftarrow \mathcal{N}_{t'-1}^{\text{APS}}(j)$; $\forall i \in \mathcal{I}_{t'-1}^{\text{APS}}$ do $\mathcal{N}_{t'}^{\text{APS}}(i) \leftarrow \mathcal{N}_{t'-1}^{\text{APS}}(i)$    ▷ Initialize neighborhoods.
    Phase 1 (Men):
    for all $j \in \mathcal{J}_{t'-1}^{\text{APS}}$ do
        for all $i \in \mathcal{N}_{t'-1}^{\text{APS}}(j) \cap \mathcal{I}_{t'-1}^{\text{APS}}$ in the order $\succ_j$ do    ▷ Make offers to available women in order
            Man $j$ makes an offer to woman $i$
            if $j$ is not the favorite of $i$ in her consideration set $\mathcal{N}_{t'-1}^{\text{APS}}(i)$ then
                break inner for    ▷ $j$ must wait if $j$ is in $i$'s consideration set $\mathcal{N}_{t'-1}^{\text{APS}}(i_1)$ but not the best one.
            else
                She inspects with $j$ if he is the best in $\mathcal{N}_{t'-1}^{\text{APS}}(i_1)$.
                if the inspection succeeds then
                    She accepts the offer and the pair is matched. Remove $j$ from $\mathcal{J}_{t'}^{\text{APS}}$ and $i$ from $\mathcal{I}_{t'}^{\text{APS}}$.
                else    ▷ The inspection fails.
                    $i$ rejects the offer. Remove $i$ from $\mathcal{N}_{t'}^{\text{APS}}(j)$ and $j$ from $\mathcal{N}_{t'}^{\text{APS}}(i)$.
                end if
            end if
        end for
        if $\mathcal{N}_{t'}^{\text{APS}}(j) = \emptyset$ and $j \in \mathcal{J}_{t'}^{\text{APS}}$ then
            Remove $j$ from $\mathcal{J}_{t'}^{\text{APS}}$
        end if
    end for
    Phase 2 (Women):
    for all $i \in \mathcal{I}_{t'-1}^{\text{APS}}$ do
        if $i$ has matched already, i.e., $i \notin \mathcal{I}_{t'}^{\text{APS}}$ then
            continue
        end if
        for all $j \in \mathcal{N}_{t'-1}^{\text{APS}}(i) \cap \mathcal{J}_{t'}^{\text{APS}}$ in the order $\succ_i$ do    ▷ Inspect available men in order
            if $i$ and $j$ have already inspected then    ▷ Since they didn't match, the inspection failed.
                continue inner for
            else if $j$ has not made an offer to $i$ then
                break inner for    ▷ $i$ must wait
            else    ▷ Man $j$ made an offer to woman $i$ in the ongoing round
                $i$ and $j$ inspect
                if the inspection succeeds then
                    She accepts the offer and the pair is matched. Remove $j$ from $\mathcal{J}_{t'}^{\text{APS}}$ and $i$ from $\mathcal{I}_{t'}^{\text{APS}}$.
                else    ▷ The inspection fails.
                    $i$ rejects the offer. Remove $i$ from $\mathcal{N}_{t'}^{\text{APS}}(j)$ and $j$ from $\mathcal{N}_{t'}^{\text{APS}}(i)$.
                end if
            end if
        end for
        if $\mathcal{N}_{t'}^{\text{APS}}(i) = \emptyset$ and $i \in \mathcal{I}_{t'}^{\text{APS}}$ then
            Remove $i$ from $\mathcal{I}_{t'}^{\text{APS}}$
        end if
    end for
    $t' \leftarrow t' + 1$
end while
```

It's easy to check that they are true when $t = 0$. Suppose they hold for $t-1$ and consider t.

First we show Claim 1. If (i, j) inspected by $t(1)$, then it must be that (1) i sees j as the best by $(t-1)(2)$ (we'll show this implies $I_{ij}^{(t-1)} = 1$) and that (2) j sees i as the best by $t(1)$ (we'll show this implies $\hat{m}_{j\to i}^{(t)} = A$). First consider condition (1). if i likes j the best by $(t-1)(2)$, then it must be true that for all $j' \succ_i j$, either (i, j') inspected and failed by $(t-1)(2)$, or j' matched with a better agent than i by $(t-1)(1)$. By assumption together with the MP update rule, this implies for all $j' \succ_i j$, either $\hat{m}_{j'\to i}^{(t-1)}$ =A and $\epsilon_{ij'} = 0$, or $\hat{m}_{j'\to i}^{(t-1)}$ =R, which again by the MP update rule implies $I_{ij}^{(t-1)} = 1$. Now consider condition (2). If j likes i the best by $t(1)$, then it must be that for all $i' \succ_j i$, either (i', j) inspected and failed by $(t-1)(2)$, or (i', j) inspected and failed during $t(1)$, or i' matched with a better opportunity than j by $(t-1)(2)$. If (i', j) inspected and failed by $(t-1)(2)$, then by assumption and MP update rule this implies $m_{i'\to j}^{(t-1)}$ =N. If i' matched with a better opportunity than j by $(t-1)(2)$, then by assumption and MP update rule together with Lemma 10 this implies $m_{i'\to j}^{(t-1)}$ =N. Finally consider the case where (i', j) inspected and failed during $t(1)$, then it must be true that i' sees j as the best by $(t-1)(2)$, which we have already shown implies $I_{i'j}^{(t-1)} = 1$. Also since (i', j) inspection failed, it must be true that $\epsilon_{ij'} = 0$, which by the MP update rule implies $m_{i'\to j}^{(t-1)}$ =N. Putting it together we have shown $m_{i'\to j}^{(t-1)}$ =N for all $i' \succ_j i$, which by the MP update rule implies $\hat{m}_{j\to i}^{(t)}$ =A.

Now we consider the other direction of Claim 1. If $\hat{m}_{j\to i}^{(t)} = A$ and $I_{ij}^{(t-1)} = 1$, then it must be true that (i, j) inspected by $t(1)$. Suppose this is not true, then it must be one of the following four cases. Case 1: i matched with a better opportunity than j by $t(1)$. Case 2: j matched with a better agent than i by $t(1)$. Case 3: i (still in the market) does not yet see j as the best by $(t-1)(2)$. Case 4: j (still in the market) does not yet see i as the best by $t(1)$. Case 1 cannot be true since by the previous argument this implies $\hat{m}_{j'\to i}^{(t)}$ =A and $\epsilon_{ij'} = 1$ for some $j' \succ_i j$, which by the MP updated rule implies $I_{ij}^{(t)} = 0$ and hence by Lemma 10 contradicts with $I_{ij}^{(t-1)} = 1$. Similarly, case 2 implies $\hat{m}_{j\to i'}^{(t)}$ =A and $m_{i'\to j}^{(t-1)}$ =Y for some $i' \succ_j i$, which by the MP update rule means $\hat{m}_{j\to i}^{(t)}$ =R and hence by Lemma 10 contradicts with $\hat{m}_{j\to i}^{(t)}$ =A. Case 3 implies i is waiting for some better opportunity than j at $(t-1)(2)$, which by assumption implies $\hat{m}_{j'\to i}^{(t-1)}$ =W for some $j' \succ_i j$, and hence by the MP update rule implies $I_{ij}^{(t-1)} = 0$, contradicting with $I_{ij}^{(t-1)} = 1$. Finally, case 4 implies j is waiting for some better agent i' than i at $t(1)$, which implies that i' (still in the market) does not yet see j as the best by $(t-1)(2)$. As we just argued, this implies $I_{i'j}^{(t-1)} = 0$. Then it must be true that (by the MP update rule) $m_{i'\to j}^{(t-1)}$ =U, since if it's not true then it must be that $\hat{m}_{j'\to i'}^{(t-1)}$ and $I_{i'j'}^{(t-1)} = 1$ and $\epsilon_{i'j'} = 1$ for some $j' \succ_{i'} j$, which by assumption implies (i', j') are matched by $(t-1)(2)$ and hence contracting with i' still in the market by $(t-1)(2)$. Now that we know $m_{i'\to j}^{(t-1)} = X$ for some $i' \succ_j i$. This contradicts with $\hat{m}_{j\to i}^{(t)}$ =A (see the MP update rule).

Secondly we show Claim 3. If j is waiting for i at $t(1)$, then it must be that j sees i as the best by $t(1)$ (we already know this implies $\hat{m}_{j\to i}^{(t)}$ =A) and that i (still in the market) does not yet see j as the best by $(t-1)(2)$ (we also know this implies $m_{i\to j}^{(t-1)}$ =U from the earlier argument). Now consider the other direction of Claim 3. If $m_{i\to j}^{(t-1)}$ =U and $\hat{m}_{j\to i}^{(t)}$ =A, then it must be that j is waiting for i at $t(1)$. Suppose this is not true, then at least one of the following cases is true. Case 1: j (still in the market) does not yet see i as the best by $t(1)$. Case 2: j is matched with a better agent than i by $t(1)$. Case 3: i sees j as the best by $(t-1)(2)$. Case 4: i matched with a better opportunity than j by $(t-1)(2)$. We already know that both case 1 and case 2 contradict with $\hat{m}_{j\to i}^{(t)}$ =A from earlier arguments. We also know that case 3 implies $I_{ij}^{(t-1)} = 1$ which contradicts with $m_{i\to j}^{(t-1)}$ =X by the MP update rule. Lastly, case 4 implies that $\hat{m}_{j'\to i}^{(t-1)}$ =A and $I_{ij'}^{(t-1)} = 1$ and $\epsilon_{ij'} = 1$ for some $j' \succ_i j$, which by the MP update rule implies $m_{i\to j}^{(t-1)}$ =N and hence contracting $m_{i\to j}^{(t-1)}$ =U.

Next we show Claim 2 and Claim 4. First we show one direction of Claim 2. If (i, j) inspected by $t(2)$, then it must be that i sees j as the best by $t(2)$ (we'll show this implies $I_{ij}^{(t)} = 1$) and that j sees i as the best by $t(1)$ (we already know this implies $\hat{m}_{j\to i}^{(t)}$ =A). If i sees j as the best by $t(2)$, then it must be true that for all $j' \succ_i j$, either (i, j') inspected and failed by $t(1)$ (by assumption and the MP update rule, we know this implies $\hat{m}_{j'\to i}^{(t)}$ =A and $\epsilon_{ij'} = 0$), or (i, j') inspected and failed during $t(2)$, or j' matched with a better agent than i by $t(1)$ (by assumption and the MP update rule, we know this implies $\hat{m}_{j'\to i}^{(t)}$ =R). If (i, j') inspected and failed during $t(2)$, then it must be that j' sees i as the best by $t(1)$ (which we already know implies $\hat{m}_{j'\to i}^{(t)}$ =A) and that $\epsilon_{ij'} = 0$. Putting it together, by the MP update rule, we have shown that $I_{ij}^{(t)} = 1$.

Next we show one direction of Claim 4. If i is waiting for j at $t(2)$, then it must be that i sees j as the best by $t(2)$ (we already know this implies $I_{ij}^{(t)} = 1$) and that j (still in the market) does not yet see i as the best by $t(1)$. The latter implies that j is waiting for some better opportunity i' than i by $t(1)$, which by assumption implies $m_{i'\to j}^{(t-1)}$ =U and $\hat{m}_{j\to i'}^{(t)}$ =A. In fact, this also implies $\hat{m}_{j\to i}^{(t)}$ =W, since if otherwise ($\hat{m}_{j\to i}^{(t)}$ =R), then for some $i'' \succ_j i$, $m_{i''\to j}^{(t-1)}$ =Y and $\hat{m}_{j\to i''}^{(t)}$ =A. If $i'' \succ_j i'$, then $\hat{m}_{j\to i'}^{(t)}$ =R by the MP update rule, which contradicts with $\hat{m}_{j\to i'}^{(t)}$ =A. If $i' \succ_j i''$, then by the MP update rule, $m_{i'\to j}^{(t-1)}$ =U contradicts with $\hat{m}_{j\to i''}^{(t)}$ =A. Also it obviously cannot be that $i' = i''$ since the messages they receive from j are different.

Now consider the other direction of Claim 2 and Claim 4. First start with Claim 2. If $\hat{m}_{j\to i}^{(t)}$ =A and $I_{ij}^{(t)} = 1$, the it must be true that (i, j) inspected by $t(2)$. Suppose this is not true, then it must be one of the following four cases. Case 1: i matched with a better opportunity than j by $t(2)$. Case 2: j matched with a better agent than i by $t(2)$. Case 3: i (still in the market) does not yet see j as the best by $t(2)$. Case 4: j (still in the market) does not yet see i as the best by $t(1)$. By assumption, case 1 implies there is some $j' \succ_i j$ such that $\hat{m}_{j'\to i}^{(t)}$ =A and $\epsilon_{ij'} = 1$, which contradicts with $I_{ij}^{(t)} = 1$. Case 2 implies (by assumption and the MP update rule) there is some $i' \succ_j i$ such that $\hat{m}_{j\to i'}^{(t)}$ =A and $m_{i'\to j}^{(t)}$ =Y, which again by the MP update rule together with Lemma 10 implies $\hat{m}_{j\to i}^{(t+1)}$ =R, hence contradicting $\hat{m}_{j\to i}^{(t)}$ =A by Lemma 10. Case 3 implies that i is waiting for some better opportunity j' than j by $t(2)$. We have already shown that this implies $\hat{m}_{j'\to i}^{(t)}$ =W, which by the MP update rule contradicts with $I_{ij}^{(t)} = 1$. Finally, case 4 implies j is waiting for some better agent i' than i by $t(1)$. We have shown this implies $m_{i'\to j}^{(t-1)}$ =U, which by the MP update rule contradicts with $\hat{m}_{j\to i}^{(t)}$ =A.

It now only remains to the other direction of Claim 4. If $\hat{m}_{j\to i}^{(t)}$ =W and $I_{ij}^{(t)} = 1$, the it must be that i is waiting for j at $t(2)$. Suppose this is not true. Then at least one of the following cases must be true. Case 1: i (still in the market) does not yet see j as the best by $t(2)$. Case 2: i matched with a better opportunity than j by $t(2)$. Case 3: j sees i as the best by $t(1)$. Case 4: j matched with a better agent than i by $t(1)$. We have shown case 1 and case 2 contradicts with $I_{ij}^{(t)} = 1$. We have also shown that case 3 implies $\hat{m}_{j\to i}^{(t)}$ =A, which contradicts with $\hat{m}_{j\to i}^{(t)}$ =W. Lastly, we have also shown that case 4 implies $\hat{m}_{j\to i}^{(t)}$ =R which contradicts with $\hat{m}_{j\to i}^{(t)}$ =W.

Therefore we have completed the proof. □

Next we prove Lemma 3.

Proof of Lemma 3. We define partial ordering $\geq^m$ on the set $\{A, R, W\}$ of man-to-woman messages by $W \geq^m A \geq^m R$. Similarly, we define partial ordering $\geq^w$ on the set $\{Y, N, U\}$ of woman-to-man messages by $U \geq^w N \geq^w Y$. Then, we can define the partial ordering $\geq^v$ of message vectors on $\{A, R, W\}^{|\mathcal{E}|} \times \{Y, N, U\}^{|\mathcal{E}|}$, where $(\hat{m}, m) \geq^v (\hat{m}', m')$ means $\hat{m}_{j\to i} \geq^m \hat{m}'_{j\to i}$ and $m_{i\to j} \geq^w m'_{i\to j}$ for all $(i, j) \in \mathcal{E}$. Note that $\{A, R, W\}^{|\mathcal{E}|} \times \{Y, N, U\}^{|\mathcal{E}|}$ is a complete lattice. Now consider an arbitrary sequence of message updates as defined in the MP algorithm in Figure 6, represented by a permutation of $(\{j \to i\}_{\forall(i,j)\in\mathcal{E}}, \{i \to j\}_{\forall(i,j)\in\mathcal{E}})$. Let s_k be the kth element in this permutation. For any realization of random market $G_N(\alpha, K, p)$, this resequenced MP algorithm can be viewed as iteratively performing $F(\cdot) : \{A, R, W\}^{|\mathcal{E}|} \times \{Y, N, U\}^{|\mathcal{E}|} \to \{A, R, W\}^{|\mathcal{E}|} \times \{Y, N, U\}^{|\mathcal{E}|}$ as defined below (F depends on the realization of random market $G_N(\alpha, K, p)$):

First denote by $\tilde{m}$ the permuted vector of $(\hat{m}, m)$ according to the above defined order $(s_k)_{k=1,\ldots,2|\mathcal{E}|}$. For each $k = 1, ..., 2|\mathcal{E}|$, s_k is either $i \to j$ or $j \to i$ for some $(i, j) \in \mathcal{E}$. For each of the two cases, we define $f_{s_k}(\cdot) : \{A, R, W\}^{|\mathcal{E}|} \times \{Y, N, U\}^{|\mathcal{E}|}$ as follows: if s_k is $j \to i$,

$$f_{j\to i}(\tilde{m}) = \begin{cases} \text{A if for all } i' \succ_j i, m_{i'\to j} = \text{N}; \\ \text{R if for some } i' \succ_j i, m_{i'\to j} = \text{Y, and } m_{i''\to j} = \text{ N for all } i'' \succ_j i'; \\ \text{W otherwise,} \end{cases}$$

or if s_k is $i \to j$,

$$g_{ij}(\tilde{m}) = \begin{cases} 1 \text{ if for all } j' \succ_i j, \text{ either } \hat{m}_{j'\to i} = \text{A and } \epsilon_{ij'} = 0; \text{ or } \hat{m}_{j'\to i} = \text{R.} \\ 0 \text{ otherwise,} \end{cases}$$

$$f_{i\to j}(\tilde{m}) = \begin{cases} \text{Y if } g_{ij}(\hat{m}) = 1 \text{ and } \epsilon_{ij} = 1; \\ \text{N if } g_{ij}(\hat{m}) = 1 \text{ and } \epsilon_{ij} = 0, \text{ or if for some } j' \succ_i j, \\ \quad \hat{m}_{j'\to i} = \text{A and } g_{ij'}(\hat{m}) = 1 \text{ and } \epsilon_{ij'} = 1; \\ \text{U otherwise.} \end{cases}$$

Then, define a sequence of vectors $\tilde{m}_1, ..., \tilde{m}_{2|\mathcal{E}|}$ in$\{A, R, W\}^{|\mathcal{E}|} \times \{Y, N, U\}^{|\mathcal{E}|}$, where each vector's elements are permutated according to $(s_k)_{k=1,\ldots,2|\mathcal{E}|}$. This sequence of vectors represents the updated messages following the resequenced MP update according to $(s_k)_{k=1,\ldots,2|\mathcal{E}|}$. For each $\tilde{m}_k$, the only element that is changed from $\tilde{m}_{k-1}$ (let $\tilde{m}_0 = \tilde{m}$) is the kth element, and the update is according to $f_{s_k}(\tilde{m}_{k-1})$ as defined above.

Finally, $F(\tilde{m}) = \tilde{m}_{2|\mathcal{E}|}$.

From the definition, it's not difficult to see that each $f_{s_k}(\cdot)$ is isotone, hence F is isotone from a finite lattice into itself. By Tarski's fixed point theorem, the set of fixed points of F is a nonempty lattice, and iterated applications of F starting from the maximum point $(W^{|\mathcal{E}|} \times U^{|\mathcal{E}|})$ in $\{A, R, W\}^{|\mathcal{E}|} \times \{Y, N, U\}^{|\mathcal{E}|}$ converge to the maximum fixed point. Observe that the permutation of the $(\hat{m}, m)$ vector, i.e., the order of the message updates in the MP algorithm considered here is arbitrary. Also, for any different permutation of the $(\hat{m}, m)$ vector, we can similarly define the corresponding F function (just apply the above procedure for the new permutation), and any fixed point of the new F function is also a fixed point of the old F function (with proper reordering of the indices) and vice versa. Therefore, iterated applications of F starting from the maximum point $(W^{|\mathcal{E}|} \times U^{|\mathcal{E}|})$ in $\{A, R, W\}^{|\mathcal{E}|} \times \{Y, N, U\}^{|\mathcal{E}|}$ converge to the same maximum fixed point, irrespective of the permutation $(s_k)_{k=1,\ldots,2|\mathcal{E}|}$. We have thus proved that the MP algorithm converges to a unique fixed point that does not depend on the sequence of message updates. □

Next we prove Lemma 4. To prove this lemma, we first establish the following claim.

Claim 1. *Fix a sequence of times $t = t_N = \omega(1)$. Then for any $L < \infty$, the following holds. For any agent i, we have*

$$\lim_{N\to\infty} \mathbb{P}(\mathcal{N}_t^{PS}(i) \nsubseteq \mathcal{N}_L^{APS}(i)) = 0$$

Proof. The probability of there being more than t_N nodes in the $(L+1)$-neighborhood of agent i is vanishing as $N \to \infty$. (The number of nodes in the $(L+1)$-neighborhood has expectation bounded above by $\Theta(1)$ and the assertion follows from Markov's inequality.) It suffices to show $\mathcal{N}_{l(L+1)}^{\mathrm{PS}}(i) \subseteq \mathcal{N}_L^{\mathrm{APS}}(i)$, where $l(L+1)$ is the number of nodes in the $(L+1)$-neighborhood of agent i. Consider a subset of $\mathcal{N}(i)$, denoted by $\tilde{\mathcal{N}}_L(i)$, where we remove from $\mathcal{N}(i)$ all neighboring opportunities that are inspected by i but failed as well as all neighboring opportunities that are matched with a better agent than i during the first L rounds of APS. If i is matched during the first L rounds of APS then we let $\tilde{\mathcal{N}}_L(i) = \emptyset$. Obviously $\tilde{\mathcal{N}}_L(i) \subseteq \mathcal{N}_L^{\mathrm{APS}}(i)$. Therefore if we can show $\mathcal{N}_{l(L+1)}^{\mathrm{PS}}(i) \subseteq \tilde{\mathcal{N}}_L(i)$ then we are done. If fact, it suffices to show that if (i,j) inspected by L in APS, then (i,j) must also inspected by $l(L+1)$ in PS; and to show that if a neighboring opportunity of i, say j, matched with a better agent i' than i by L in APS, then (i',j) must also matched by $l(L+1)$ in PS.

We first show the former. It suffices to show that in order to make the (i,j) inspection happen by L in APS, all the prerequisite inspections needed are on edges within the $(L+1)$-neighborhood of agent i. Suppose this is not true, i.e., we need an inspection on edge (i',j') outside the $(L+1)$-neighborhood of agent i to happen as a prerequisite to make the (i,j) inspection happen by L in APS. By Lemma 2, this means we need $\hat{m}_{j'\to i'}^{(L')}$ and $I_{i'j'}^{(L')}$ for some $L' \le L$ to determine $\hat{m}_{j\to i}^{(L)}$ and $I_{ij}^{(L)}$. By the MP rule, this is not true, hence a contradiction.

Now we only need show the latter. In fact it follows easily from the previous argument. The claim hence follows. □

Now we are ready to prove Lemma 4.

Proof of Lemma 4. The fact that PS and APS converge to identical outcomes is immediated from Lemma 3. Take any agent i. Consider her consideration set just after the first t rounds of PS. We carefully define the sequence t' as follows. For each $L \in \mathbb{N}$, let N_L be the smallest number s.t., for all $N \ge N_L$, it holds that

$$\mathbb{P}(\mathcal{N}_t^{\mathrm{PS}}(i) \nsubseteq \mathcal{N}_L^{\mathrm{APS}}(i)) \le \frac{1}{L}\,. \tag{6}$$

By Claim 1, we know that $N_L < \infty$ for all $L \in \mathbb{N}$. Note that $N_2, N_3, N_4, \ldots$ is a monotone non-decreasing sequence. For all N, define $t'_N \triangleq \max\{L : N_L \le N\}$. Clearly, $t'_N = \omega(1)$. By definition of t'_N, in the N-th market we have

$$\mathbb{P}(\mathcal{N}_t^{\mathrm{PS}}(i) \nsubseteq \mathcal{N}_{t'_N}^{\mathrm{APS}}(i)) \le \frac{1}{t'_N}\,. \tag{7}$$

Since this holds for all agents, by Markov's inequality,

$$\mathbb{P}\left(\left|\left\{i \in \mathcal{I} : \mathcal{N}_t^{\mathrm{PS}}(i) \nsubseteq \mathcal{N}_{t'_N}^{\mathrm{APS}}(i)\right\}\right| \ge \frac{N}{\sqrt{t'_N}}\right) \le \frac{1}{\sqrt{t'_N}}\,. \tag{8}$$

The lemma follows. □

Finally we prove Lemma 5 (DE is exact on trees).

Proof of Lemma 5. We prove by induction. First consider $t = 1$. Obviously the messages all distance-1 opportunities from the root node i receive from their children agent nodes at time 0 are X. Also, ex-ante, the messages i receive from her neighboring opportunity nodes come from i.i.d. subtrees (by the MP update rule, the message i receives from her neighboring node j at t only depends on the messages j receives from its other neighboring agent nodes at $t-1$), and hence we only need to show the probability distribution of $\hat{m}^{(1)}_{j\to i}$ is A w.p. a_1, R w.p. r_1, and W w.p. w_1. By the MP update rule, $\hat{m}^{(1)}_{j\to i}$ can only possibly be A or W, and $\hat{m}^{(1)}_{j\to i}$ =A iff j sees i as most fit ex-ante. Conditioning on $(i, j) \in \mathcal{E}$, the probability of the number d of better than i agents for j has distribution Poisson(Ku/α), where u ~Uniform$[0, 1]$. Therefore the probability of j seeing i as most fit ex-ante can be computed by

$$\begin{aligned}\mathbb{P}(d = 0) &= \mathbb{E}_u \mathbb{P}(d = 0|u)\\ &= \mathbb{E}e^{-Ku/\alpha}\\ &= \frac{e^{-K/\alpha-1} - 1}{-K/\alpha}.\end{aligned}$$

One can easily verify that this expression matches with a_1 in Definition 4, and that $r_1 = 0$ and $w_1 = 1 - a_1$.

Now we know that Lemma 5 is true for $t = 1$. Next assume it holds for any $1, ..., t-1$ rounds of MP and consider the messages the root node i receives after t rounds of MP. We start by examine the messages all distance-1 opportunities receive from their children agent nodes. Take all distance-1 agents and consider the subtrees starting from these agents. Clearly, ex-ante, these subtree are i.i.d. GW trees with mark, hence by assumption, the messages each of these distance-1 agent nodes receives from their children are i.i.d. A w.p. a_{t-1}, R w.p. r_{t-1}, and W w.p. w_{t-1}. By the MP update rule, this implies that the messages all distance-1 opportunities receive from their children agents are i.i.d.. Next we will show that if we take any opportunity j_1 and its child agent i_1 that are both of distance-1 to the root agent node i, then the probability distribution of $m^{(t-1)}_{i_1\to j_1}$ is Y w.p. y_{t-1}, N w.p. ν_{t-1}, and U w.p. u_{t-1}. We first compute the probability of $I^{(t-1)}_{i_1 j_1} = 1$. From the MP update rule, we know that $I^{(t-1)}_{i_1 j_1} = 1$ iff for all $j' \succ_{i_1} j_1$, either $\hat{m}^{(t-1)}_{j'\to i_1}$ =A and $\epsilon_{i_1 j'} = 0$, or $\hat{m}^{(t-1)}_{j'\to i_1}$ =R. We can also imply from the MP update rule that the message $\hat{m}^{(t-1)}_{j'\to i_1}$ does not depend on $\epsilon_{i_1 j'}$ or $\epsilon_{i_1 j''}$ for any other $j'' \succ_{i_1} j_1$. Therefore we can treat $\Pr(\epsilon_{i_1 j'} = 0) = p$ for all $j' \succ_{i_1} j_1$ independently from $\hat{m}^{(t-1)}_{j'\to i_1}$ for all $j' \succ_{i_1} j_1$ (note that the probability distributions of $\hat{m}^{(t-1)}_{j'\to i_1}$ for all $j' \succ_{i_1} j_1$ are also independent). Moreover, the probability distribution of the number d_1 of better than j_1 opportunities for i_1 is Poisson(Ku_1), where u_1 ~Uniform$[0, 1]$. Therefore, by the chain rule of conditional expectations and utilizing the moment generating functions for the Poisson random variable $d_1|u_1$ and Uniform random variable u_1, we can compute

$$\begin{aligned}\mathbb{P}\left(I^{(t-1)}_{i_1 j_1} = 1\right) &= \mathbb{P}\left(\text{for all } j' \succ_{i_1} j_1, \text{ either } \hat{m}^{(t-1)}_{j'\to i_1} = \text{A and } \epsilon_{i_1 j'} = 0, \text{ or } \hat{m}^{(t-1)}_{j'\to i_1} = \text{R}\right)\\ &= \mathbb{E}_{d_1}\left[(a_{t-1}(1-p) + r_{t-1})^{d_1}\right]\\ &= \mathbb{E}_{u_1}\left[\mathbb{E}_{d_1}\left[(a_{t-1}(1-p) + r_{t-1})^{d_1}|u_1\right]\right]\\ &= \mathbb{E}_{u_1}\left[\mathbb{E}_{d_1}\left[(a_{t-1}(1-p) + r_{t-1})^{d_1}|u_1\right]\right]\\ &= \mathbb{E}_{u_1}\left[e^{-Ku_1(a_{t-1}p + w_{t-1})}\right]\\ &= \frac{1 - e^{-K(a_{t-1}p + w_{t-1})}}{K(a_{t-1}p + w_{t-1})} = q_{t-1}. \qquad (9)\end{aligned}$$

Now, consider the probability of $m_{i_1\to j_1}^{(t-1)}$ =Y and the probability of $m_{i_1\to j_1}^{(t-1)}$ =N. By the MP update rule, it's obvious to see that

$$\mathbb{P}\left(m_{i_1\to j_1}^{(t-1)}=\text{Y}\right)=\mathbb{P}\left(I_{i_1j_1}^{(t-1)}=1\text{ and }\epsilon_{i_1j_1}=1\right)=q_{t-1}p=y_{t-1}.$$

Also, by the MP update rule we have

$$\begin{aligned}&\mathbb{P}\left(m_{i_1\to j_1}^{(t-1)}=\text{N}\right)\\=&\mathbb{P}\left(I_{i_1j_1}^{(t-1)}=1\text{ and }\epsilon_{i_1j_1}=0\text{ or for some }j'\succ_{i_1}j_1,\ \hat{m}_{j'\to i_1}^{(t-1)}=\text{A and }I_{i_1j'}^{(t-1)}=1\text{ and }\epsilon_{i_1j'}=1\right).\end{aligned}$$

Observe that by the MP update rule, if $I_{i_1j_1}^{(t-1)}=1$, then there must not $\exists j'\succ_{i_1}j_1$ s.t. $\hat{m}_{j'\to i_1}^{(t-1)}$ =A and $\epsilon_{i_1j'}=1$. Therefore the two events $\{$for some $j'\succ_{i_1}j_1$, $\hat{m}_{j'\to i_1}^{(t-1)}$ =A, $I_{i_1j'}^{(t-1)}=1$ and $\epsilon_{i_1j'}=1\}$ and $\{I_{i_1j_1}^{(t-1)}=1$ and $\epsilon_{i_1j_1}=0\}$ are disjoint, and hence

$$\begin{aligned}&\mathbb{P}\left(m_{i_1\to j_1}^{(t-1)}=\text{N}\right)\\=&\mathbb{P}\left(I_{i_1j_1}^{(t-1)}=1\text{ and }\epsilon_{i_1j_1}=0\text{ or for some }j'\succ_{i_1}j_1,\ \hat{m}_{j'\to i_1}^{(t-1)}=\text{A and }I_{i_1j'}^{(t-1)}=1\text{ and }\epsilon_{i_1j'}=1\right)\\=&\mathbb{P}\left(I_{i_1j_1}^{(t-1)}=1\text{ and }\epsilon_{i_1j_1}=0\right)+\mathbb{P}\left(\text{for some }j'\succ_{i_1}j_1,\ \hat{m}_{j'\to i_1}^{(t-1)}=\text{A and }I_{i_1j'}^{(t-1)}=1\text{ and }\epsilon_{i_1j'}=1\right).\end{aligned}\tag{10}$$

The first term above is simply

$$\mathbb{P}\left(I_{i_1j_1}^{(t-1)}=1\text{ and }\epsilon_{i_1j_1}=0\right)=q_{t-1}(1-p).\tag{11}$$

The second term can be written as the summation of a number of disjoint events' probability, where each event corresponds to $\hat{m}_{j'\to i_1}^{(t-1)}$ =A, $I_{i_1j'}^{(t-1)}=1$ and $\epsilon_{i_1j'}=1$ for the lth best opportunity j' of i_1, $l=1,2,...,d_1$. Taking the conditional expectation on d_1, we get

$$\begin{aligned}&\mathbb{P}\left(\text{for some }j'\succ_{i_1}j_1,\ \hat{m}_{j'\to i_1}^{(t-1)}=\text{A and }I_{i_1j'}^{(t-1)}=1\text{ and }\epsilon_{i_1j'}=1\right)\\=&\mathbb{E}_{d_1}\left[\sum_{l=1}^{d_1}\mathbb{P}\left(\text{for the }l\text{th best opportunity }j'\text{ of }i_1,\ \hat{m}_{j'\to i_1}^{(t-1)}=\text{A, }I_{i_1j'}^{(t-1)}=1\text{ and }\epsilon_{i_1j'}=1\middle|d_1\right)\right]\\=&\mathbb{E}_{d_1}\left[\sum_{l=1}^{d_1}(a_{t-1}(1-p)+r_{t-1})^{l-1}a_{t-1}p\right]\\=&\mathbb{E}_{d_1}\left[\frac{\left(1-(a_{t-1}(1-p)+r_{t-1})^{d_1}\right)a_{t-1}p}{a_{t-1}p+w_{t-1}}\right]\\=&\frac{a_{t-1}p(1-q_{t-1})}{a_{t-1}p+w_{t-1}},\end{aligned}\tag{12}$$

where the third equality follows by applying the geometric series formula, and the last step follows from the derivation of $\mathbb{P}\left(I_{i_1j_1}^{(t-1)}=1\right)=q_{t-1}$ in Eq. (9). Therefore, combining Eq. (10) – (12), we obtain

$$\mathbb{P}\left(m_{i_1\to j_1}^{(t-1)}=\text{N}\right)=q_{t-1}(1-p)+\frac{a_{t-1}p(1-q_{t-1})}{a_{t-1}p+w_{t-1}}=\nu_{t-1}.\tag{13}$$

One can also easily verify that $u_{t-1} = 1 - y_{t-1} - \nu_{t-1}$. Therefore we have shown that, ex-ante, the probability distribution of the message $m^{(t-1)}_{i_1 \to j_1}$ from any distance-1 agent i_1 to her parent opportunity j_1 is Y w.p. y_{t-1}, N w.p. ν_{t-1}, U w.p. u_{t-1}, and is independent from all other such pairs.

Now consider $\hat{m}^{(t)}_{j_1 \to i}$ (as before, j_1 is any distance-1 opportunity from root agent node i) and its probability distribution ex-ante. We already know that they are i.i.d for all neighboring opportunities hence it suffices to consider just the pair (i, j_1). Denote by $\hat{d}$ the number of better than i agents for j_1, which is distributed according to Poisson$(K\hat{u}/\alpha)$, where $\hat{u}$ ∼Uniform[0, 1]. By the MP update rule, we can compute

$$\begin{aligned}
\mathbb{P}\left(\hat{m}^{(t)}_{j_1 \to i} = \text{A}\right) &= \mathbb{P}\left(\text{for all } i' \succ_{j_1} i,\, m^{(t)}_{i' \to j_1} = \text{N}\right) \\
&= \mathbb{E}_{\hat{d}} \mathbb{P}\left(\text{for all } i' \succ_{j_1} i,\, m^{(t)}_{i' \to j_1} = \text{N} \middle| \hat{d}\right) \\
&= \mathbb{E}_{\hat{d}}\left(\nu^{\hat{d}}_{t-1}\right) \\
&= \mathbb{E}_{\hat{u}}\left[\mathbb{E}\left(\nu^{\hat{d}}_{t-1} | \hat{u}\right)\right] \\
&= \mathbb{E}_{\hat{u}}\left(e^{\frac{K\hat{u}}{\alpha}(\nu_{t-1}-1)}\right) \\
&= \frac{e^{\frac{K}{\alpha}(\nu_{t-1}-1)} - 1}{\frac{K}{\alpha}(\nu_{t-1} - 1)} = a_t,
\end{aligned} \tag{14}$$

where the third step utilizes Eq. (13), the fourth step applies the chain rule of conditional expectations, and the last two lines follow from the moment generating functions of Poisson and Uniform random variables.

Similarly, by the MP update rule for the R message,

$$\begin{aligned}
&\mathbb{P}\left(\hat{m}^{(t)}_{j_1 \to i} = \text{R}\right) \\
=&\mathbb{P}\left(\text{for some } i' \succ_{j_1} i,\, m^{(t)}_{i' \to j_1} = \text{Y, and } m^{(t-1)}_{i'' \to j_1} = \text{N for all } i'' \succ_{j_1} i'\right) \\
=&\mathbb{E}_{\hat{d}}\left[\sum_{l=1}^{\hat{d}} \mathbb{P}\left(\text{for the } l\text{th best agent } i' \text{ of } j_1,\, m^{(t-1)}_{i' \to j_1} = \text{Y, and } m^{(t-1)}_{i'' \to j_1} = \text{N for all } i'' \succ_{j_1} i'\right)\right] \\
=&\mathbb{E}_{\hat{d}}\left[\sum_{l=1}^{\hat{d}} (\nu^{l-1}_{t-1} y_{t-1})\right] \\
=&\mathbb{E}_{\hat{d}}\left[\frac{(1 - \nu^{\hat{d}}_{t-1}) y_{t-1}}{1 - \nu_{t-1}}\right] \\
=&\frac{y_{t-1}(1 - a_t)}{1 - \nu_{t-1}} = r_t,
\end{aligned}$$

where the second step follows from summing up the probability of $\hat{d}$ disjoint events, conditional on the value of $\hat{d}$ and taking expectation over $\hat{d}$. The fifth step results from applying $\mathbb{E}_{\hat{d}}\left(\nu^{\hat{d}}_{t-1}\right) = a_t$ in Eq. (14). One can also easily verify that $w_t = 1 - a_t - r_t$. The lemma follows. □

The proof of Corollary 1 is straightforward from Lemma 5 and Lemma 10, and is omitted here.

C Existence of the Deadlock Regime

In this section, we prove Proposition 1, which establishes that the deadlock regime Θ is nonempty. In other words, there exist market primitives $\alpha > 0, K > 0, p \in (0,1)$ for which $\lambda(\alpha, K, p) > 0$.

Recall that, by Definition 4, $a_t, r_t, w_t, q_t, y_t, \nu_t, u_t$ satisfy

$$\begin{aligned} a_t &= \frac{e^{\frac{K}{\alpha}(\nu_{t-1}-1)}-1}{\frac{K}{\alpha}(\nu_{t-1}-1)}; \quad r_t = \frac{y_{t-1}}{1-\nu_{t-1}}(1-a_t); \quad w_t = \frac{u_{t-1}}{1-\nu_{t-1}}(1-a_t); \\ q_t &= \frac{1-e^{-K(a_t p + w_t)}}{K\left(a_t p + w_t\right)}; \quad y_t = q_t p; \quad \nu_t = q_t(1-p) + \frac{a_t p(1-q_t)}{a_t p + w_t}; \quad u_t = \frac{w_t(1-q_t)}{a_t p + w_t}. \end{aligned} \tag{15}$$

By Corollary 1, w_t and a_t are monotone in t. We can similarly show that r_t, q_t, y_t, ν_t and u_t are monotone in t as well. Therefore the limits exist. Denote

$$\begin{aligned} a^* &:= \lim_{t\to\infty} a_t; \quad r^* = \lim_{t\to\infty} r_t; \quad w^* = \lim_{t\to\infty} w_t; \\ q^* &= \lim_{t\to\infty} q_t; \quad y^* = \lim_{t\to\infty} y_t; \quad \nu^* = \lim_{t\to\infty} \nu_t; \quad u^* = \lim_{t\to\infty} u_t. \end{aligned}$$

Obviously, they satisfy the following system of equations:

$$\begin{aligned} a^* &= \frac{e^{\frac{K}{\alpha}(\nu^*-1)}-1}{\frac{K}{\alpha}(\nu^*-1)}; \quad r^* = \frac{y^*}{1-\nu^*}(1-a^*); \quad w^* = \frac{u^*}{1-\nu^*}(1-a^*); \\ q^* &= \frac{1-e^{-K(a^* p + w^*)}}{K\left(a^* p + w^*\right)}; \quad y^* = q^* p; \quad \nu^* = q^*(1-p) + \frac{a^* p(1-q^*)}{a^* p + w^*}; \quad u^* = \frac{w^*(1-q^*)}{a^* p + w^*}. \end{aligned} \tag{16}$$

The next lemma characterizes the value of a^* for the deadlock-free regime.

Lemma 11. *If $(K,p,\alpha) \notin A$, then the corresponding a^* is the unique root of $e^{-Kpa} - 1 - \alpha e^{\frac{e^{-Kpa}-1}{\alpha a}} + \alpha = 0$ on $(0,1)$, denoted by $a^*(Kp,\alpha)$. Moreover, $a^*(Kp,\alpha)$ is increasing in α.*

Proof. If $(K,p,\alpha) \notin A$, then $w^* = u^* = 0$ and one can easily show that Eq. (16) reduces to

$$a^* = \frac{e^{\frac{K}{\alpha}(\nu^*-1)}-1}{\frac{K}{\alpha}(\nu^*-1)}; \quad r^* = 1 - a^*; \quad y^* = \frac{1-e^{-Ka^*p}}{Ka^*}; \quad \nu^* = 1 - y^*; \quad w^* = u^* = 0. \tag{17}$$

In fact, by combining the three expressions for a^*, y^* and ν^*, we can obtain the following equation:

$$e^{-Ka^*p} - 1 - \alpha e^{\frac{e^{-Ka^*p}-1}{\alpha a^*}} + \alpha = 0. \tag{18}$$

Let $c = Kp$ and $F(c,\alpha,a) := e^{-ca} - 1 - \alpha e^{\frac{e^{-ca}-1}{\alpha a}} + \alpha$. Therefore Eq. (18) implies $F(c,\alpha,a^*) = 0$. Since $c = Kp > 0$ and $\alpha > 0$, we have $e^{ac} > 1 + ac$ and hence

$$\frac{\partial F(c,\alpha,a)}{\partial a} = -ce^{-ac} + \frac{e^{-\frac{1-e^{-ac}+a^2 c\alpha}{a\alpha}}(1 + ac - e^{ac})}{a^2} < 0.$$

By L'Hôpital's rule, $\lim_{a\to 0} \frac{e^{-ca}-1}{\alpha a} = \lim_{a\to 0} \frac{-ce^{-ca}}{\alpha} = -\frac{c}{\alpha} < 0$, and hence $\lim_{a\to 0} F(c,\alpha,a) = -\alpha e^{-\frac{c}{\alpha}} + \alpha > 0$. Also, $F(c,\alpha,1) = e^{-c} - 1 - \alpha e^{\frac{e^{-c}-1}{\alpha}} + \alpha < e^{-c} - 1 - \alpha\left(\frac{e^{-c}-1}{\alpha} + 1\right) + \alpha = 0$ since $e^{\frac{e^{-c}-1}{\alpha}} > \frac{e^{-c}-1}{\alpha} + 1$. Therefore, there exists a unique $a^*(c,\alpha) \in (0,1)$ such that $F(c,\alpha,a^*(c,\alpha)) = 0$. Also note that $F(c,\alpha,a^*) = 0$, i.e., $a^* = a^*(c,\alpha)$.

To prove that $a^*(c,\alpha)$ is increasing in α, compute

$$\frac{\partial^2 F(c,\alpha,a)}{\partial \alpha^2} = -\frac{e^{\frac{-1+e^{-ac}}{a\alpha}}\left(e^{-ac}-1\right)^2}{a^2\alpha^3} < 0.$$

Moreover,

$$\frac{\partial F(c,\alpha,a)}{\partial \alpha} = 1 + \frac{e^{\frac{-1+e^{-ac}}{a\alpha}}\left(e^{-ac}-1-a\alpha\right)}{a\alpha}.$$

One can easily see that $\lim_{\alpha\to\infty} \frac{\partial F(c,\alpha,a)}{\partial\alpha} = 0$, which implies that $\frac{\partial F(c,\alpha,a)}{\partial\alpha} > 0$. By the implicit function Theorem, we have $\frac{\partial a^*(c,\alpha)}{\partial\alpha} > 0$, i.e., $a^*(c,\alpha)$ is increasing in α. □

Next we show that the deadlock regime as characterized by Θ is nonempty by showing that it cannot be the case that $(\alpha, K, p) \notin \Theta$ for all $\alpha \geq 0$, $K \geq 0$ and $p \in [0,1]$.

Proof of Proposition 1. Recall from Eq. (15) that

$$w_t = \frac{(1-q_{t-1})(1-a_t)w_{t-1}}{(1-\nu_{t-1})(a_{t-1}p + w_{t-1})}.$$

Since w_t is decreasing in t by Corollary 1, we must have $\frac{(1-q_{t-1})(1-a_t)}{(1-\nu_{t-1})(a_{t-1}p+w_{t-1})} \leq 1$ for all $t = 1, 2, ...$ In particular, $\frac{(1-q^*)(1-a^*)}{(1-\nu^*)(a^*p+w^*)} \leq 1$. If $w^* = 0$, this reduces to $\frac{(1-q^*)(1-a^*)}{(1-\nu^*)a^*p} \leq 1$. Plugging in Eq. (17), we can rewrite $\frac{(1-q^*)(1-a^*)}{(1-\nu^*)a^*p} \leq 1$ as

$$\frac{\left(Ka^*p - 1 + e^{-Ka^*p}\right)(1-a^*)}{a^*p^2\left(1-e^{-Ka^*p}\right)} \leq 1,$$

or equivalently,

$$p^2 \geq \frac{\left(Ka^*p - 1 + e^{-Ka^*p}\right)(1-a^*)}{a^*\left(1-e^{-Ka^*p}\right)}. \tag{19}$$

Let $c = Kp$ and note that we have proved $a^* = a^*(c,\alpha)$ in Lemma 11. Observe that the RHS of Eq. (19) is only a function of c and α, hence we denote it by $p^*(c,\alpha)$. Since $c > 0$ and $a^*(c,\alpha) \in (0,1)$ by Lemma 11, we have $p^*(c,\alpha) > 0$. Therefore fix any $c > 0$ and $\alpha > 0$, then for $\forall p \in (0, \sqrt{p^*(c,\alpha)}) \neq \emptyset$, Eq. (19) is not satisfied. In other words

$$\left\{(K,p,\alpha) : p < \sqrt{p^*(Kp,\alpha)}\right\} \subseteq \Theta$$

and

$$\left\{(K,p,\alpha) : p < \sqrt{p^*(Kp,\alpha)}\right\} \neq \emptyset,$$

which means $\Theta \neq \emptyset$. □

D Proof of Theorem 1

In this section, we prove Lemma 1 and Theorem 1.

We first prove Lemma 1 (random markets are locally tree-like).

Proof of Lemma 1. Note that conditional on the topology of $\mathcal{B}_r(i)$ and $\mathcal{T}_r$ being identical, we can easily construct a coupling of the preference/fitness rankings and latent inspection outcomes such that $\mathcal{B}_r(i) = \mathcal{T}_r$. Therefore it suffices to show the coupling without markings, i.e., without preference/fitness rankings and without latent inspection outcomes.

We do a breadth-first search on $G_N(\alpha, K, p)$ starting from node i. Put node i into a queue and then iteratively do the following, starting from $s = 1$. In FIFO order, take one node from the queue and count all nodes connected to it that have not been added to the queue. Denote this number Z_s and add these nodes to the end of the queue in arbitrary order. Then update $s = s + 1$. Repeat until the queue is emptied. Denote by S the number of total iterations. Take any positive integer l. Consider the first $l \wedge S$ nodes (including both agents and opportunities) examined from the queue. Next we show that the probability of the subgraph (in the original economy) spanned by these nodes not being a tree is bounded above by $\Theta(\frac{l^2}{N})$. For integer $s \geq 0$, define set C_s that contains all nodes in the queue after the sth iteration in the above breadth-first search ($C_0 = \{i\}$). Also let v_s be the node taken out from the queue in the sth iteration ($v_1 = i$). Let $U_s = \{u \in C_s : \{u, v_{s+1}\} \in \mathcal{E} \text{ or } \{v_{s+1}, u\} \in \mathcal{E}\}$. Observe that $|U_s| \sim$Binomial$(|C_s| - 1, K/(\alpha N))$. Therefore,

$$\begin{aligned}
\mathbb{P}\left(\exists 1 \leq s \leq l \wedge S : |U_s| \neq 0\right) &\leq \mathbb{E}\left[\sum_{s=1}^{l} \mathbb{P}(|U_s| \neq 0 | C_s)\right] \\
&\leq \mathbb{E}\left[\sum_{s=1}^{l} \left(1 - \left(1 - \frac{K}{\alpha N}\right)^{|C_s|-1}\right)\right] \\
&\leq \Theta\left(\frac{l^2}{N}\right),
\end{aligned}$$

where the first inequality is an application of the union bound, the second inequality follows from the distribution of $|U_s|$, and the last step follows from $|C_l| = 1 + \sum_{s=1}^{l}(Z_s - 1)$, Jensen's inequality, and $\mathbb{E}Z_s \leq (K/\alpha) \wedge K$.

Similarly, we can also conduct a breadth-first search on the tree $\mathcal{T}(\alpha, K, p)$ in the same way as above. Similar to Z_s, denote by Z'_s the number of untouched nodes connected to the sth node being examined from the queue. Denote by S' the number of total iterations before the queue is emptied. Consider the first $l \wedge S'$ nodes. Next we show that there exists coupling between Z_s and Z'_s such that $\mathbb{P}((Z_1, ..., Z_{l\wedge S}) \neq (Z'_1, ..., Z'_{l\wedge S'})) \leq \Theta(\frac{l^2}{N})$. Define I_s and J_s to be the set of agent nodes and the set of opportunity nodes, respectively, that have not been added to the queue nor taken out from the queue by (just after) the sth iteration in the breath-first search in $G_N(\alpha, K, p)$ ($I_0 = \mathcal{I} - \{i\}$ and $J_0 = \mathcal{J}$). Clearly we have $Z_{s+1} \sim$Binomial$(|I_s|, K/(\alpha N))$ if v_{s+1} is an opportunity node, and $Z_{s+1} \sim$Binomial$(|J_s|, K/(\alpha N))$ if v_{s+1} is an agent node. Let ξ_{s+1} be independent $\sim$Binomial$(N - |I_s|, K/(\alpha N))$ if v_{s+1} is an opportunity node, and $\sim$Binomial$(\alpha N - |J_s|, K/(\alpha N))$ if v_{s+1} is an agent node. Then $Z_{s+1} + \xi_{s+1}$ is $\sim$Binomial$(N, K/(\alpha N))$ if v_{s+1} is an opportunity node, and $\sim$Binomial$(\alpha N, K/(\alpha N))$ if v_{s+1} is an agent node. We have

$$\begin{aligned}
\mathbb{P}((Z_1, ..., Z_{l\wedge S}) \neq (Z_1 + \xi_1, ..., Z_{l\wedge S} + \xi_{l\wedge S})) &\leq \mathbb{E}\left[\sum_{s=1}^{l} \mathbb{P}(\xi_s \neq 0 \big| |I_{s-1}|, |J_{s-1}|, v_s)\right] \\
&\leq \mathbb{E}\left[\sum_{s=1}^{l} \left(1 - \left(1 - \frac{K}{\alpha N}\right)^{N+\alpha N - |I_{s-1}| - |J_{s-1}|}\right)\right] \\
&\leq \mathbb{E}\left[\sum_{s=1}^{l} \left(1 - \left(1 - \frac{K}{\alpha N}\right)^{s-1+|C_{s-1}|}\right)\right]
\end{aligned}$$

$$\leq \Theta(\frac{l^2}{N}). \quad (20)$$

Note that Eq. (20) still holds if we use l instead of $l \wedge S$ on the L.H.S subscript and draw i.i.d. additional Binomial($\alpha N, K/(\alpha N)$) samples for Z_i and additional 0s for ξ_i if $S < l$. On the other hand, it is known that the total variation distance between Binomial($\alpha N, K/(\alpha N)$) and Poisson(K) as well as the total variation distance between Binomial($N, K/(\alpha N)$) and Poisson(K/α) are both no bigger than $K/(\alpha N)$. Thus, by maximal coupling, there exists a coupling between $Z_1 + \xi_1$ and Z_1' such that $\mathbb{P}(Z_1 + \xi_1 \neq Z_1') \leq \frac{K}{\alpha N}$. We can hence inductively show that there exists a coupling between $(Z_1 + \xi_1, ..., Z_l + \xi_l)$ and $(Z_1', ..., Z_l')$ such that $\mathbb{P}((Z_1 + \xi_1, ..., Z_l + \xi_l) \neq (Z_1', ..., Z_l')) \leq \frac{lK}{\alpha N}$ provided it is true for vectors of length $l - 1$. This combined with Eq. (20) imply that $\mathbb{P}((Z_1, ..., Z_l) \neq (Z_1', ..., Z_l')) \leq \Theta(\frac{l^2}{N})$. Since the values of S and S' would be identical provided that $S < l$ and that $(Z_1, ..., Z_l) = (Z_1', ..., Z_l')$, we have the desired coupling between Z_s and Z_s' such that $\mathbb{P}((Z_1, ..., Z_{l\wedge S}) \neq (Z_1', ..., Z_{l\wedge S'}')) \leq \Theta(\frac{l^2}{N})$.

Combine the results above implies that there exists coupling such that the probability of the subgraph of $G_N(\alpha, K, p)$ spanned by the first $l \wedge S$ nodes being different from the subtree in $\mathcal{T}(\alpha, K, p)$ spanned by the first $l \wedge S'$ nodes is bounded above by $\Theta(\frac{l^2}{N})$.

Clearly, the number of breadth-first search iterations on $\mathcal{T}(\alpha, K, p)$ up to depth r is $|\mathcal{T}_r|$. One can check that $\mathbb{E}(|\mathcal{T}_r|) = e^{Cr}$ where $C = \log(K^2/\alpha)$. Fix any sequence of $r = o(\log N)$, we can find some $\Delta = o(N)$ such that $\frac{\Delta}{e^{2Cr}} \to \infty$. Define $M(r) := \sqrt{(}e^{2Cr} + \Delta)$ so that $\mathbb{P}(|\mathcal{T}_r| > M(r)) \leq \frac{e^{Cr}}{M(r)} \to 0$ as $n \to \infty$. We have the following result:

$$\begin{aligned}
&\mathbb{P}(\mathcal{B}_r(i) \neq \mathcal{T}_r) \\
=&\mathbb{P}(\mathcal{B}_r(i) \neq \mathcal{T}_r \&\& |\mathcal{T}_r| \leq M(r)) + \mathbb{P}(\mathcal{B}_r(i) \neq \mathcal{T}_r \&\& |\mathcal{T}_r| > M(r)) \\
\leq&\mathbb{P}(\mathcal{B}_r(i) \neq \mathcal{T}_r \big| |\mathcal{T}_r| \leq M(r)) + \mathbb{P}(|\mathcal{T}_r| > M(r)) \\
\leq&\Theta(\frac{M(r)^2}{N}) + o(1) \\
=&o(1).
\end{aligned}$$

The lemma follows. □

Take any agent $i \in G_N(\alpha, K, p)$ and consider the probability of her waiting after t rounds of accelerated PS. Apply the Message Passing (MP) algorithm on both $G_N(\alpha, K, p)$ and $\mathcal{T}(\alpha, K, p)$, Lemma 1 together with Lemma 2 and Lemma 5 imply the following corollary.

Corollary 2. *Fix any α, K and p and suppress the notations. Take any agent $i \in G_N$ and define $\lambda_N^{\text{APS}}(t) = \Pr(i \in \mathcal{I}_t^{\text{APS}})$, where $\mathcal{I}_t^{\text{APS}}$ is defined in accelerated PS in Algorithm 1. Let w_t and q_t be as defined in Definition 4. Then for any sequence of $t = o(\log N)$, $|\lambda_N^{\text{APS}}(t) - K w_t q_t| \leq o(1)$.*

Proof. Define $X_N^G(t) = \mathbb{1}(i \in \mathcal{I}_t^{\text{APS}})$. Also consider a marked GW tree $\mathcal{T}(\alpha, K, p)$ that satisfies the coupling $\Pr(\mathcal{B}_t(i) \neq \mathcal{T}_t) \leq o(1)$ in Lemma 1. Conduct MP on the tree and denote by m. By Lemma 2 and the MP update rule, we know that the probability of the root agent node waiting after t rounds of APS is fully determined by $\mathcal{T}_t$. Denote the indicator function of this event happening by $X_N^T(t)$. By Lemma 2 and Lemma 5, let $d \sim$Poisson(K), we can compute (by the MP update rule)

$$\begin{aligned}
\mathbb{P}(X_N^T(t) = 1) &= \mathbb{P}(\text{the root agent node } \tilde{i} \text{ of } \mathcal{T}_r \text{ is waiting after } t \text{ rounds of APS}) \\
&= \mathbb{P}(\exists j \in \mathcal{T}_1 : \hat{m}_{j\to\tilde{i}}^{(t)} = \text{W and } I_{\tilde{i}j}^{(t)} = 1) \\
&= \mathbb{P}(\exists j \in \mathcal{T}_1 : \hat{m}_{j\to\tilde{i}}^{(t)} = \text{W and } \forall j' \succ_{\tilde{i}} j, \text{ either } \hat{m}_{j'\to\tilde{i}}^{(t)} = \text{A and } \epsilon_{\tilde{i}j'} = 0; \text{ or } \tilde{m}_{j'\to\tilde{i}}^{(t)} = \text{R})
\end{aligned}$$

$$
\begin{aligned}
&= \mathbb{E}\left[\sum_{l=1}^{d} w_t(a_t(1-p)+r_t)^{l-1}\right] \\
&= w_t\mathbb{E}\left[\frac{1-(a_t(1-p)+r_t)^d}{a_t p + w_t}\right] \\
&= \frac{w_t}{a_t p + w_t}\left(1-e^{-K(a_t p + w_t)}\right) \\
&= K w_t q_t,
\end{aligned}
$$

where the second step applies Lemma 2, the third step utilizes the MP update rule, the fourth step follows from Lemma 5, and the last three steps results from applying the geometric series formula, the moment generating function for Poisson random variables, and the definition of q_t in Definition 4, respectively.

Therefore,

$$
\begin{aligned}
\mathbb{P}(X_N^G(t)=1) &= \mathbb{P}(X_N^G(t)=1\&\&\mathcal{B}_t(i)=\mathcal{T}_t) + \mathbb{P}(X_N^G(t)=1\&\&\mathcal{B}_t(i)\neq\mathcal{T}_t) \\
&= \mathbb{P}(X_N^T(t)=1\&\&\mathcal{B}_t(i)=\mathcal{T}_t) + \mathbb{P}(X_N^G(t)=1\&\&\mathcal{B}_t(i)\neq\mathcal{T}_t) \\
&\leq \mathbb{P}(X_N^T(t)=1) + \mathbb{P}(\mathcal{B}_t(i)\neq\mathcal{T}_t) \\
&\leq K w_t q_t + o(1),
\end{aligned}
$$

where the last step follows from Lemma 1. On the other hand, we also have

$$
\begin{aligned}
\mathbb{P}(X_N^G(t)=1) &= \mathbb{P}(X_N^G(t)=1\&\&\mathcal{B}_t(i)=\mathcal{T}_t) + \mathbb{P}(X_N^G(t)=1\&\&\mathcal{B}_t(i)\neq\mathcal{T}_t) \\
&\geq \mathbb{P}(X_N^G(t)=1\&\&\mathcal{B}_t(i)=\mathcal{T}_t) \\
&= \mathbb{P}(X_N^T(t)=1\&\&\mathcal{B}_t(i)=\mathcal{T}_t) \\
&= \mathbb{P}(X_N^T(t)=1) - \mathbb{P}(X_N^T(t)=1\&\&\mathcal{B}_t(i)\neq\mathcal{T}_t) \\
&\geq \mathbb{P}(X_N^T(t)=1) - \mathbb{P}(\mathcal{B}_t(i)\neq\mathcal{T}_t) \\
&\geq K w_t q_t - o(1).
\end{aligned}
$$

The lemma follows. □

We also have the following corollary that establishes that the local neighborhoods of any two agents in the market are two independent trees.

Corollary 3. *Fix any α, K and p and suppress the notations. Take any two agents v_1 and v_2 in G_N and consider their distance t neighborhood $\mathcal{B}_t(v_1)$ and $\mathcal{B}_t(v_2)$, respectively. Denote this subgraph by $\mathcal{B}_t(v_1, v_2)$. Also denote by $\mathcal{T}^2$ the graph of two independent GW trees with marks (under primitive (α, K, p)). Then for any sequence of $t = o(\log N)$, there exists a coupling between G_N and $\mathcal{T}^2$ such that $\mathbb{P}(\mathcal{B}_t(v_1, v_2) \neq \mathcal{T}_t^2) \leq o(1)$.*

Proof. Consider the following exploration process. Denote by $V = \mathcal{I} \cup \mathcal{J}$ the set of agent nodes and opportunity nodes in G_N. Maintain four sets $V_s^{(1)}, V_s^{(2)}, U_s, C_s$ which are a partition of V in each iteration s. Start with $V_0^{(1)} = \{i_1\}$, $V_0^{(2)} = \{i_2\}$, $U_0 = \emptyset$, $C_0 = \mathcal{I} \setminus \{i_1, i_2\}$. In each odd (even) round of iteration $s = 1, 3, 5, ...$ ($s = 2, 4, 6, ...$), take a vertex from set $V_{s-1}^{(1)}$ ($V_{s-1}^{(2)}$), denote it by v_s. Define $I_i = \{u \in U_{s-1} : \{u, v_s\} \in \mathcal{E} \text{ or } \{v_s, u\} \in \mathcal{E}\}$. If either $V_{s-1}^{(1)}$ or $V_{s-1}^{(2)}$ is empty, we let $\{v_s\} = I_s = \emptyset$. Now we update $V_{s+1}^{(1)} = V_s^{(1)} \setminus \{v_s\} \cup I_s$ ($V_{s+1}^{(1)} = V_s^{(1)}$), $V_{s+1}^{(2)} = V_s^{(2)}$ ($V_{s+1}^{(2)} = V_s^{(2)} \setminus \{v_s\} \cup I_s$), $U_{s+1} = U_s \setminus I_s$, $C_{s+1} = C_s \cup \{v_s\}$. In fact, in each round s we can define set $V_s = V_s^{(1)} \cup V_s^{(2)}$, then applying the same technique as when proving Lemma 1, we get a similar result. Take any integer l.

There exists coupling such that the probability of the subgraph of G_N spanned by at most the first (according to the sequence of breath-first search) l nodes in the neighborhood of i_1 and at most the first l nodes in the neighborhood of i_2 being different from the subgraph of $\mathcal{T}^2$ spanned by at most the first l nodes in tree-1 and the first l nodes in tree-2 is bounded above by $\Theta(\frac{l^2}{N})$.

Now we need to consider the depth-t neighborhood of i_1 and i_2. Denote by $\mathcal{T}_t^{(1)}$ and $\mathcal{T}_t^{(2)}$ the two independent GW trees up to depth-t in $\mathcal{T}^2$. We know that $\mathbb{E}(|\mathcal{T}_t^{(1)}|) = \mathbb{E}(|\mathcal{T}_t^{(2)}|) = e^{Ct}$ where $C = \log(K^2/\alpha)$. Fix any sequence of $t = o(\log N)$, we can find some $\Delta = o(n)$ such that $\frac{\Delta}{e^{2Ct}} \to \infty$. Define $M(t) := \sqrt{(e^{2Ct} + \Delta)}$ so that $\mathbb{P}(|\mathcal{T}_t^{(1)}| > M(t)) = \mathbb{P}(|\mathcal{T}_t^{(2)}| > M(t)) \leq \frac{e^{Ct}}{M(t)} \to 0$ as $N \to \infty$. We have the following result:

$$
\begin{aligned}
&\mathbb{P}(\mathcal{B}_t(v_1, v_2) \neq \mathcal{T}_t^2) \\
\leq &\mathbb{P}(\mathcal{B}_t(v_1, v_2) \neq \mathcal{T}_t^2 \&\& |\mathcal{T}_t^{(1)}| \leq M(t) \&\& |\mathcal{T}_t^{(2)}| \leq M(t)) + \mathbb{P}(|\mathcal{T}_t^{(1)}| > M(t)) + \mathbb{P}(|\mathcal{T}_t^{(2)}| > M(t)) \\
= &\Theta(\frac{M(t)^2}{N}) = o(1).
\end{aligned}
$$

The lemma follows. □

Now we are ready to prove Theorem 1.

Proof of Theorem 1. We first prove the deadlock-free regime. It is obvious that for any fixed economy, the number of agents who wait is decreasing over time. Therefore, it suffices to consider sequence of $t = \omega(1)$ that is also $o(\log N)$. By Lemma 4, there is a sequence of times $t' = \omega(1)$ (that is also $o(\log N)$ since APS obviously proceeds faster than PS) such that for any $\delta > 0$, we have

$$
\mathbb{P}\left(|\{i \in \mathcal{I} : \mathcal{N}_t^{\mathrm{PS}}(i) \nsubseteq \mathcal{N}_{t'}^{\mathrm{APS}}(i)\}| < \delta N/2\right) \geq 1 - o(1). \tag{21}
$$

Take any $\epsilon > 0$. Note that $q_s \geq (1 - e^{-K})/K > 0$ for any $s \geq 0$ from Definition 4. Therefore if $(\alpha, K, p) \notin \Theta$, then it must that $\lim_{s\to\infty} w_s = 0$, and hence there exists some finite $\tau > 0$ such that $w_\tau < \epsilon/K$. Also by Corollary 2, for any $t' = o(\log N)$ we have $|\lambda_N^{\mathrm{APS}}(t') - K w_t q_t| \to 0$ as $N \to \infty$, where $\lambda_N^{\mathrm{APS}}(t')$ is the probability of agent i waiting after t' rounds of APS. Therefore, for any $t' = \omega(1)$ that is also $o(\log N)$, when N is sufficiently large, we have $t' > \tau$ and hence $\lambda_N^{\mathrm{APS}}(t') < \epsilon + o(1)$. Since the choice of ϵ is arbitrary, this implies $\lambda_N^{\mathrm{APS}}(t') < o(1)$. Take any $\delta > 0$. By Markov's Inequality,

$$
\mathbb{P}(|\{i \in \mathcal{I} : \mathcal{N}_{t'}^{\mathrm{APS}}(i) \neq \emptyset\}| \geq \delta N/2) \leq \frac{\lambda_N^{\mathrm{APS}}(t')N}{\delta N/2} = o(1). \tag{22}
$$

This implies $\mathbb{P}(|\{i \in \mathcal{I} : \mathcal{N}_{t'}^{\mathrm{APS}}(i) \neq \emptyset\}| < \delta N/2) \geq 1 - o(1)$. Therefore combining Eqs. (21) – (22), we get

$$
\begin{aligned}
&\mathbb{P}(\Lambda_N^t(\alpha, K, p) \leq \delta N) \\
=&\mathbb{P}(|\{i \in \mathcal{I} : \mathcal{N}_t^{\mathrm{PS}}(i) \neq \emptyset\}| \leq \delta N) \\
=&\mathbb{P}(|\{i \in \mathcal{I} : \mathcal{N}_t^{\mathrm{PS}}(i) \neq \emptyset \wedge \mathcal{N}_t^{PS}(i) \nsubseteq \mathcal{N}_{t'}^{\mathrm{APS}}(i)\}| + |\{i \in \mathcal{I} : \mathcal{N}_t^{\mathrm{PS}}(i) \neq \emptyset \wedge \mathcal{N}_t^{PS}(i) \subseteq \mathcal{N}_{t'}^{\mathrm{APS}}(i)\}| \leq \delta N) \\
\geq&\mathbb{P}(|\{i \in \mathcal{I} : \mathcal{N}_t^{PS}(i) \nsubseteq \mathcal{N}_{t'}^{\mathrm{APS}}(i)\}| + |\{i \in \mathcal{I} : \mathcal{N}_{t'}^{\mathrm{APS}}(i) \neq \emptyset\}| \leq \delta N) \\
\geq&\mathbb{P}((|\{i \in \mathcal{I} : \mathcal{N}_t^{PS}(i) \nsubseteq \mathcal{N}_{t'}^{\mathrm{APS}}(i)\}| < \delta N/2) \wedge (|\{i \in \mathcal{I} : \mathcal{N}_{t'}^{\mathrm{APS}}(i) \neq \emptyset\}| < \delta N/2)) \\
\geq&1 - o(1).
\end{aligned}
$$

Since the choice of $\delta > 0$ is arbitrary, we have the desired result.

It now only remains to show the information deadlock regime. If $(\alpha, K, p) \in \Theta$, then for any $\epsilon > 0$, there exists τ such that for any $s > \tau$ we have $|Kw_s q_s - \lambda| \leq \epsilon$. By Corollary 2, since $t = o(\log N)$, we have $|\lambda_N^{\text{APS}}(t) - Kw_t q_t| \to 0$ as $N \to \infty$, where $\lambda_N^{\text{APS}}(t)$ is the ex-ante probability of an agent i waiting after t rounds of APS. Therefore, when N is sufficiently large, we have $t > \tau$ and hence $\lambda - \epsilon - o(1) < \lambda_N^{\text{APS}}(t') < \lambda + \epsilon + o(1)$. Since the choice of ϵ is arbitrary, this implies $\lambda - o(1) < \lambda_N^{\text{APS}}(t') < \lambda + o(1)$.

For any agent i, define $X_N^t(i)$ to be the indicator function of agent i in G_N waiting at time t in APS. Let $\tilde{\Lambda}_N^t = \sum_{i \in \mathcal{I}} X_N^t(i)$. Then $\mathbb{E}\tilde{\Lambda}_N^t = N\lambda_N^{\text{APS}}(t)$. Obviously $\Lambda_N^t \geq \tilde{\Lambda}_N^t$ (since the set of waiting agents is decreasing over time and APS proceeds faster than PS). Take any $\delta > 0$, we have

$$\begin{aligned}
\mathbb{P}(\Lambda_N^t \geq \lambda N - \delta N) &\geq \mathbb{P}(\tilde{\Lambda}_N^t \geq \lambda N - \delta N) \\
&\geq \mathbb{P}(|\tilde{\Lambda}_N^t - \lambda N| \leq \delta N) \\
&= \mathbb{P}(|\tilde{\Lambda}_N^t - \lambda_N^{\text{APS}}(t)N + \lambda_N^{\text{APS}}(t)N - \lambda N| \leq \delta N) \\
&\geq \mathbb{P}(|\tilde{\Lambda}_N^t - \lambda_N^{\text{APS}}(t)N| + |\lambda_N^{\text{APS}}(t)N - \lambda N| \leq \delta N) \\
&\geq \mathbb{P}(|\tilde{\Lambda}_N^t - \lambda_N^{\text{APS}}(t)N| \leq \delta N/2 \quad \wedge \quad |\lambda_N^{\text{APS}}(t)N - \lambda N| \leq \delta N/2) \\
&\xrightarrow{N \to \infty} \mathbb{P}(|\tilde{\Lambda}_N^t - \lambda_N^{\text{APS}}(t)N| \leq \delta N/2),
\end{aligned}$$

where the last step follows from $\lambda - o(1) < \lambda_N^{\text{APS}}(t) < \lambda + o(1)$. Therefore, to show the desired result, it suffices to show $\mathbb{P}(|\tilde{\Lambda}_N^t - \lambda_N^{\text{APS}}(t)N| \leq \delta N/2) \geq 1 - o(1)$, or equivalently, $\mathbb{P}(|\tilde{\Lambda}_N^t - \lambda_N^{\text{APS}}(t)N| > \delta N/2) \leq o(1)$

Note that $\mathbb{E}\tilde{\Lambda}_N^t = \lambda_N^{\text{APS}}(t)N$. By Chebyshev's inequality, we have

$$\mathbb{P}(|\tilde{\Lambda}_N^t - \lambda_N^{\text{APS}}(t)N| > \delta N/2) \leq \frac{\sigma(\tilde{\Lambda}_N^t)^2}{\delta^2 N^2/4},$$

where $\sigma(\tilde{\Lambda}_N^t)^2$ is the variance of $\tilde{\Lambda}_N^t$:

$$\sigma(\tilde{\Lambda}_N^t)^2 = N\sigma(X_N^t(1))^2 + N(N-1)\text{Cov}(X_N^t(i_1), X_N^t(i_2)).$$

It's easy to check that $\lambda(1-\lambda) - o(1) < \sigma(X_N^t(i_1))^2 < \lambda(1-\lambda) + o(1)$. Now it only remains to compute $\text{Cov}(X_N^t(i_1), X_N^t(i_2))$. Note that $\text{Cov}(X_N^t(i_1), X_N^t(i_2)) = \mathbb{E}[X_N^t(i_1)X_N^t(i_2)] - \mathbb{E}X_N^t(i_1)\mathbb{E}X_N^t(i_2)$, where

$$\begin{aligned}
\mathbb{E}[X_N^t(i_1)X_N^t(i_2)] &= \mathbb{P}\left(X_N^t(i_1) = X_N^t(i_2) = 1\right) \\
&\leq \mathbb{P}\left(X_N^t(i_1) = X_N^t(i_2) = 1 \middle| \mathcal{B}_t(i_1, i_2) = \mathcal{T}_r^2\right) + \mathbb{P}(\mathcal{B}_t(i_1, i_2) \neq \mathcal{T}_r^2) \\
&= \lambda^2 + o(1).
\end{aligned}$$

The last step above follows from Corollary 3 and the fact that $t = \omega(1)$. Now, we can compute

$$\sigma(\tilde{\Lambda}_N^t)^2 \leq N\lambda(1-\lambda) + N(N-1)(\lambda^2 - \lambda^2) + o(N^2) = o(N^2),$$

and hence

$$\mathbb{P}(|\tilde{\Lambda}_N^t - \lambda_N^{\text{APS}}N| > \delta N/2) \leq \frac{\sigma(\tilde{\Lambda}_N^t)^2}{\delta^2 N^2/4} = o(1).$$

Therefore we have shown $\mathbb{P}(\Lambda_N^t \geq \lambda N - \delta N) \geq 1 - o(1)$.

Next we show $\mathbb{P}(\Lambda_N^t \leq \lambda N + \delta N) \geq 1 - o(1)$. By Lemma 4, there is a sequence of $t' = \omega(1)$

that is also $o(\log N)$ such that for any $\delta > 0$, we have

$$\mathbb{P}\left(|\{i \in \mathcal{I} : \mathcal{N}_t^{\mathrm{PS}}(i) \not\subseteq \mathcal{N}_{t'}^{\mathrm{APS}}(i)\}| < \delta N/2\right) \geq 1 - o(1). \tag{23}$$

Therefore

$$\begin{aligned}
&\mathbb{P}(\Lambda_N^t \leq \lambda N + \delta N) \\
=&\mathbb{P}(|\{i \in \mathcal{I} : \mathcal{N}_t^{\mathrm{PS}}(i) \neq \emptyset\}| \leq \lambda N + \delta N) \\
=&\mathbb{P}(|\{i \in \mathcal{I} : \mathcal{N}_t^{\mathrm{PS}}(i) \neq \emptyset \wedge \mathcal{N}_t^{\mathrm{PS}}(i) \not\subseteq \mathcal{N}_{t'}^{\mathrm{APS}}(i)\}| + |\{i \in \mathcal{I} : \mathcal{N}_t^{\mathrm{PS}}(i) \neq \emptyset \wedge \mathcal{N}_t^{\mathrm{PS}}(i) \subseteq \mathcal{N}_{t'}^{\mathrm{APS}}(i)\}| \leq \lambda N + \delta N) \\
\geq&\mathbb{P}(|\{i \in \mathcal{I} : \mathcal{N}_t^{\mathrm{PS}}(i) \not\subseteq \mathcal{N}_{t'}^{\mathrm{APS}}(i)\}| + |\{i \in \mathcal{I} : \mathcal{N}_{t'}^{\mathrm{APS}}(i) \neq \emptyset\}| \leq \lambda N + \delta N) \\
\geq&\mathbb{P}((|\{i \in \mathcal{I} : \mathcal{N}_t^{\mathrm{PS}}(i) \not\subseteq \mathcal{N}_{t'}^{\mathrm{APS}}(i)\}| \leq \delta N/2) \wedge (|\{i \in \mathcal{I} : \mathcal{N}_{t'}^{\mathrm{APS}}(i) \neq \emptyset\}| \leq \lambda N + \delta N/2)).
\end{aligned}$$

Observe that

$$\begin{aligned}
&\mathbb{P}((|\{i \in \mathcal{I} : \mathcal{N}_t^{\mathrm{PS}}(i) \not\subseteq \mathcal{N}_{t'}^{\mathrm{APS}}(i)\}| \leq \delta N/2) \wedge (|\{i \in \mathcal{I} : \mathcal{N}_{t'}^{\mathrm{APS}}(i) \neq \emptyset\}| \leq \lambda N + \delta N/2)) \\
=&1 - \mathbb{P}((|\{i \in \mathcal{I} : \mathcal{N}_t^{\mathrm{PS}}(i) \not\subseteq \mathcal{N}_{t'}^{\mathrm{APS}}(i)\}| > \delta N/2) \vee (|\{i \in \mathcal{I} : \mathcal{N}_{t'}^{\mathrm{APS}}(i) \neq \emptyset\}| > \lambda N + \delta N/2)) \\
\geq&1 - \mathbb{P}(|\{i \in \mathcal{I} : \mathcal{N}_t^{\mathrm{PS}}(i) \not\subseteq \mathcal{N}_{t'}^{\mathrm{APS}}(i)\}| > \delta N/2) - \mathbb{P}(|\{i \in \mathcal{I} : \mathcal{N}_{t'}^{\mathrm{APS}}(i) \neq \emptyset\}| > \lambda N + \delta N/2).
\end{aligned}$$

Therefore

$$\begin{aligned}
&\mathbb{P}(\Lambda_N^t \leq \lambda N + \delta N) \\
\geq&1 - \mathbb{P}(|\{i \in \mathcal{I} : \mathcal{N}_t^{\mathrm{PS}}(i) \not\subseteq \mathcal{N}_{t'}^{\mathrm{APS}}(i)\}| > \delta N/2) - \mathbb{P}(|\{i \in \mathcal{I} : \mathcal{N}_{t'}^{\mathrm{APS}}(i) \neq \emptyset\}| > \lambda N + \delta N/2) \\
\geq&1 - \mathbb{P}(\tilde{\Lambda}_N^{t'} > \lambda N + \delta N/2) - o(1) \\
\geq&1 - \mathbb{P}(|\tilde{\Lambda}_N^{t'} - \lambda N| > \delta N/2) - o(1),
\end{aligned}$$

where the second inequality utilizes Eq. (23).

We just showed for $t' = \omega(1)$ that is also $o(\log N)$, $\mathbb{P}(|\tilde{\Lambda}_N^{t'} - \lambda N| \leq \delta N/2) \geq 1 - o(1)$. Therefore, we have the desired result that $\mathbb{P}(\Lambda_N^t \leq \lambda N + \delta N|) \geq 1 - o(1)$. Finally, combine both the lower bound and the upper bound, we get

$$\mathbb{P}(|\Lambda_N^t - \lambda N| \leq \delta N) = \mathbb{P}(\Lambda_N^t \leq \lambda N + \delta N \wedge \Lambda_N^t \geq \lambda N - \delta N) \geq 1 - o(1).$$

Since the choice of $\delta > 0$ is arbitrary, the theorem follows.

□

E Proof of Section 5

In this appendix we prove the key lemmas in Section 5 for proving Theorem 2. As outlined in Section 5, we will prove the lower bound by identifying, in some cases, the final deadlock structure that makes a focal node wait forever. To uncover those deadlock structures of interest, we define a key neighborhood revelation process ALG-G (see Figure 9) and couple it with an analogous process ALG-T on the associated tree of $G_N(\alpha, K, p)$, so that we can make full use of the latter process's independence property to facilitate the analysis. In Section E.1, we define and discuss some regularity conditions needed on the empirical degree distributions of $G_N(\alpha, K, p)$, as well as the implied nice properties about ALG-T that hold with high probability, including Lemma 8. Section E.2 proves the coupling result in Lemma 6. Section E.3 proves Lemma 7, i.e., the chance of deadlock formation in each step of ALG-G is on the order of $\frac{t}{N}$, up to some $\Theta\sqrt{N}$ time. Lemma 9 is also proved in Section E.3. Finally we prove Theorem 2 in Section E.4.

To facilitate the analysis, we first introduce some preliminaries and auxiliary lemmas.

E.1 Preliminaries and auxiliary lemmas

For convenience, we will use a random configuration model to describe the bipartite graph. This is introduced in Section E.1.1. We then present several useful probability bounds in Section E.1.2. These bounds are useful in the later proofs.

E.1.1 A random configuration model

Consider $G_N(\alpha, K, 0)$, a random bipartite graph with N women nodes and $M = \alpha N$ men nodes, where each edge (connecting one woman and one man) is present independently with probability $\frac{K}{M}$, and each node independently has a random preference ranking over connected nodes. We now give an alternative way to generate $G_N(\alpha, K, 0)$ using a random configuration model.

We first fix a degree sequence of $G_N(\alpha, K, 0)$ specified by $\vec{d} = ((d_i)_{i=1}^N, (d'_j)_{j=1}^{\alpha N})$. Given the degree sequence $\vec{d}$, we can draw a random bipartite graph from the random configuration model as follows. For each woman node $i = 1, ..., N$, generate d_i half-edges with a (independent) random permutation of ranking among them. Similary, for each man node $j = 1, ..., \alpha N$, generate d'_j half edges with a (independent) random permutation of ranking among them. Then, connect women half-edges and men half-edges uniformly at random. The relationship between this random configuration model and $G_N(\alpha, K, 0)$ is given in the next lemma.

Lemma 12 (The random configuration model v.s. the Erdos-Renyi graph). *Fix any degree sequence $\hat{d}$ in $G_N(\alpha, K, 0)$. A bipartite graph drawn from the random configuration model with that degree sequence, conditional on it being simple, has the same distribution as that of $G_N(\alpha, K, 0)$ conditional on that degree sequence.*

Proof. Standard result in random graph theory, see, e.g., [13] and [38] for more details. □

We call the degree sequence of $G_N(\alpha, K, 0)$ regular if it satisfies a few conditions in the next definition. All results about ALG-G will be established under regular degree sequences, and we will later show that the degree sequence of $G_N(\alpha, K, 0)$ is w.h.p. regular if $K^2/(4\alpha) > 1$.

Definition 8 (Regular degree sequence). *Consider the degree sequence $\vec{d} = ((d_i)_{i=1}^N, (d'_j)_{j=1}^{\alpha N})$ of an Erdos-Renyi graph $G_N(\alpha, K, p)$. Let*

$$g_N^w(s) \triangleq \sum_{l=0}^{\alpha N-1} \frac{(l+1)\sum_{i=1}^N \mathbb{1}\{d_i = l+1\}}{\sum_{i=1}^N d_i} s^l, \quad g_N^m(s) \triangleq \sum_{l=0}^{N-1} \frac{(l+1)\sum_{j=1}^{\alpha N} \mathbb{1}\{d'_j = l+1\}}{\sum_{j=1}^{\alpha N} d'_j} s^l$$

be the empirical probability generating functions of the (edge-perspective) woman nodes' degree minus 1 and the (edge-perspective) man nodes' degree minus 1, respectively. Also let

$$f_N^w(s) \triangleq \sum_{l=0}^{\alpha N-1} \frac{\sum_{i=1}^N \mathbb{1}\{d_i \geq l+1\}}{\sum_{i=1}^N d_i} s^l, \quad f_N^m(s) \triangleq \sum_{l=0}^{N-1} \frac{\sum_{j=1}^{\alpha N} \mathbb{1}\{d'_j \geq l+1\}}{\sum_{j=1}^{\alpha N} d'_j} s^l$$

be the empirical probability generating functions of the (edge-perspective) number of better edges of a woman node and of a man node, respectively. Let q_N^w be the smallest nonnegative root of $s = f_N^w(f_N^m(s))$, and q_N^m be the smallest nonnegative root of $s = f_N^m(f_N^w(s))$.

Then the degree sequence is called $(a, \beta, \gamma, \kappa, \xi, \mu)$-regular, where $a > 1, \beta > 1, \gamma \in (0,1), \kappa \in (0,1), \xi \in (0,1), \mu \in (0,1)$, if the following conditions hold:

1. $g_N^w(a) \leq \beta$, $g_N^m(a) \leq \beta$;

2. $f_N^w(\gamma) \le \kappa$, $f_N^m(\gamma) \le \kappa$;

3. $\left|\frac{\sum_{i=1}^N d_i}{N} - K\right| \le \left|\sqrt{\frac{4\alpha+K^2}{2}} - K\right|$; $\left|\frac{\sum_{j=1}^{\alpha N} d'_j}{\alpha N} - \frac{K}{\alpha}\right| \le \left|\sqrt{\frac{4\alpha+K^2}{2\alpha^2}} - \frac{K}{\alpha}\right|$; $\left|(g_N^w)'(1) - K\right| \le \left|\sqrt{\frac{4\alpha+K^2}{2}} - K\right|$, $\left|(g_N^m)'(1) - \frac{K}{\alpha}\right| \le \left|\sqrt{\frac{4\alpha+K^2}{2\alpha^2}} - \frac{K}{\alpha}\right|$;

4. $q_N^w \le \xi$, $q_N^m \le \xi$;

5. $f_N^{w\prime}(f_N^m(q_N^w))f_N^{m\prime}(q_N^w) \le \mu$, $f_N^{m\prime}(f_N^w(q_N^m))f_N^{w\prime}(q_N^m) \le \mu$.

Condition 1 ensures that the average actual degree of all revealed nodes in $\mathcal{P}_t^T$ is not too large, and serves to guarantee that the coupling between ALG-G and ALG-T persists for sufficiently long (probabilistically speaking). Condition 2, 3, 4 and 5 serve to ensure that the one step deadlock probability is bounded below by a sufficiently large probability at each time step in ALG-G. In particular, condition 2 is used to show that the unrevealed degree of nodes in the active set cannot be too small, so that there are many opportunities for cycle formation during each step of ALG-G, conditional on the coupling between ALG-G and ALG-T. In order for the chances of cycle formation to continue for a sufficiently long time, we also need a positive probability of the ALG-G process to go on forever. This is guaranteed by Conditions 3 and 4. Condition 3 requires that the branching factor of the branching process under the empirical degree distribution is close to the branching factor $\frac{K^2}{4\alpha}$ under the limit degree distribution, close enough such that the former is above one whenever the latter is; quantitatively $g_N^{w\prime}(1)g_N^{m\prime}(1) \ge 1 + (K^2/(4\alpha) - 1)/2$. This implies Condition 4, which ensures that ALG-T has a positive probability of going on forever. We state both conditions separately for convenience in later proofs. Lastly, Condition 5 requires that the effective branching factor of the branching process under the empirical degree distribution, rooted at either a woman node or a man node, conditional on extinction, is bounded above by a constant strictly less than one that only depends on K, α. Conditions 2–5 together guarantee that the size of the active set, and the total number of unrevealed edges of nodes in it grow linearly in time with a high probability, conditional on ALG-T not terminating for a long time, thus providing enough different possibilities for deadlock to form.

Lemma 13 (Degree sequence is regular w.h.p.). *Suppose $K^2 > 4\alpha$. There exist constants $a \equiv a(\alpha, K) > 1, \beta \equiv \beta(\alpha, K) > 1, \gamma \equiv \gamma(\alpha, K) \in (0,1), \kappa \equiv \kappa(\alpha, K) \in (0,1), \xi \equiv \xi(\alpha, K) \in (0,1), \mu \equiv \mu(\alpha, K) \in (0,1)$, such that,*

$$\lim_{N\to\infty} \mathbb{P}(\textit{the empirical degree sequence of } G_N(\alpha, K, p) \textit{ is } (a, \beta, \gamma, \kappa, \xi, \mu)\textit{-regular}) = 1.$$

Proof. Consider $G_N(\alpha, K, p)$ where $K^2 > 4\alpha$ and a realization of its degree sequence $(N_l^w)_{l=0}^{\alpha N}$, $(N_l^m)_{l=0}^N$. Note that it suffices to establish that, separately for each condition in Definition 8, the probability that the condition is satisfied converges to one as $N \to \infty$.

First consider condition 1. Note that

$$\begin{aligned} g_N^w(s) &= \sum_{l=0}^{\alpha N-1} \frac{(l+1)\sum_{i=1}^N \mathbb{1}\{d_i = l+1\}}{\sum_{i=1}^N d_i} s^l \\ &= \frac{\sum_{l=0}^{\alpha N-1}(l+1)\sum_{i=1}^N \mathbb{1}\{d_i = l+1\} s^l}{N} \cdot \frac{N}{\sum_{i=1}^N d_i} \\ &= \frac{\sum_{i=1}^N \sum_{l=0}^{\alpha N-1}(l+1)\mathbb{1}\{d_i = l+1\} s^l}{N} \cdot \frac{N}{\sum_{i=1}^N d_i} \\ &= \frac{\sum_{i=1}^N d_i s^{d_i-1}}{N} \cdot \frac{N}{\sum_{i=1}^N d_i} \end{aligned} \tag{24}$$

Observe that $\frac{1}{N}\sum_{i=1}^{N} d_i s^{d_i-1}$ is the empirical mean of $d_1 s^{d_1-1}, ..., d_N s^{d_N-1}$, where $d_i \overset{iid}{\sim}$ Binomial$(\alpha N, \frac{K}{\alpha N})$. Take any $a > 1$, say, $a = 2$. By the Law of Large Numbers, for any $\epsilon_1 > 0$,

$$\lim_{N\to\infty} \mathbb{P}\left(\frac{1}{N}\sum_{i=1}^{N} d_i a^{d_i-1} \le \mathbb{E}\left[d_1 a^{d_1-1}\right] + \epsilon_1\right) = 1. \tag{25}$$

Similarly, by the Law of Large Numbers, for any $\epsilon_2 > 0$,

$$\lim_{N\to\infty} \mathbb{P}\left(\frac{\sum_{i=1}^{N} d_i}{N} \ge \mathbb{E}\left[d_1\right] - \epsilon_2\right) = 1,$$

or equivalently (suppose $\epsilon_2 < \mathbb{E}[d_1] = K$),

$$\lim_{N\to\infty} \mathbb{P}\left(\frac{N}{\sum_{i=1}^{N} d_i} \le \frac{1}{\mathbb{E}\left[d_1\right] - \epsilon_2}\right) = 1. \tag{26}$$

Eq. (24)–(26) together yield

$$\lim_{N\to\infty} \mathbb{P}\left(g_N^w(a) \le \frac{\mathbb{E}\left[d_1 a^{d_1-1}\right] + \epsilon_1}{\mathbb{E}\left[d_1\right] - \epsilon_2}\right) = 1$$

Since $d_1 \sim$Binomial$(\alpha N, \frac{K}{\alpha N})$, one can verify that

$$\lim_{N\to\infty} \mathbb{E}[d_1] = K, \quad \lim_{N\to\infty} \frac{\mathbb{E}\left[d_1 a^{d_1-1}\right]}{\mathbb{E}\left[d_1\right]} = c_1$$

for some $c_1 \equiv c_1(\alpha, K) > 1$. Therefore, there must exist $\beta_1 \equiv \beta_1(\alpha, K) \in (1, \infty)$ such that

$$\lim_{N\to\infty} \mathbb{P}(g_N^w(a) \le \beta_1) = 1.$$

Similarly, there must exist $\beta_2 \equiv \beta_2(\alpha, K) \in (1, \infty)$ such that

$$\lim_{N\to\infty} \mathbb{P}(g_N^m(a) \le \beta_2) = 1.$$

Take $\beta = \max\{\beta_1, \beta_2\}$, condition 1 is satisfied w.h.p..

Condition 2 can be treated similarly with the observation that

$$\begin{aligned}
f_N^w(s) &= \sum_{l=0}^{\alpha N-1} \frac{\sum_{i=1}^{N} \mathbb{1}\{d_i \ge l+1\}}{\sum_{i=1}^{N} d_i} s^l \\
&= \frac{\sum_{l=0}^{\alpha N-1}\sum_{i=1}^{N} \mathbb{1}\{d_i \ge l+1\} s^l}{N} \cdot \frac{N}{\sum_{i=1}^{N} d_i} \\
&= \frac{\sum_{i=1}^{N}\sum_{l=0}^{\alpha N-1} \mathbb{1}\{d_i \ge l+1\} s^l}{N} \cdot \frac{N}{\sum_{i=1}^{N} d_i} \\
&= \frac{\sum_{i=1}^{N}\sum_{l=0}^{d_i-1} s^l}{N} \cdot \frac{N}{\sum_{i=1}^{N} d_i} \\
&= \frac{\sum_{i=1}^{N} \frac{1-s^{d_i}}{1-s}}{N} \cdot \frac{N}{\sum_{i=1}^{N} d_i}.
\end{aligned}$$

The rest of the proof is similar to the proof of condition 1 and we omit the details.

Now consider condition 3. Note that, by the Law of Large Numbers, for any $\tilde{\epsilon}_1 > 0, \tilde{\epsilon}_2 > 0$,

$$\lim_{N\to\infty} \mathbb{P}\left(\mathbb{E}[d_1] - \tilde{\epsilon}_1 \le \frac{\sum_{i=1}^N d_i}{N} \le \mathbb{E}[d_1] + \tilde{\epsilon}_2\right) = 1.$$

One can verify that $\lim_{N\to\infty} \mathbb{E}[d_1] = K$. Also, $\left|\sqrt{\frac{4\alpha+K^2}{2}} - K\right| > 0$ since $K^2 > 4\alpha$. Therefore we can choose $\tilde{\epsilon}_1 > 0, \tilde{\epsilon}_2 > 0$ appropriately such that

$$\lim_{N\to\infty} \mathbb{P}\left(K - \left|\sqrt{\frac{4\alpha+K^2}{2}} - K\right| \le \frac{\sum_{i=1}^N d_i}{N} \le K + \left|\sqrt{\frac{4\alpha+K^2}{2}} - K\right|\right) = 1.$$

The second inequality in condition 3 can be treated similarly. Now consider the third inequality. From Eq. (24),

$$(g_N^w)'(1) = \frac{\sum_{i=1}^N d_i(d_i-1)}{N} \cdot \frac{N}{\sum_{i=1}^N d_i}.$$

Similar to the previous argument, by the Law of Large Numbers, for any $\epsilon_1 > 0$, $\epsilon_2 \ge 0$, $\epsilon_3 > 0$ and $\epsilon_4 \in (0, K)$,

$$\lim_{N\to\infty} \mathbb{P}\left(\frac{\mathbb{E}[d_1(d_1-1)] - \epsilon_1}{\mathbb{E}[d_1] + \epsilon_2} \le (g_N^w)'(1) \le \frac{\mathbb{E}[d_1(d_1-1)] + \epsilon_3}{\mathbb{E}[d_1] - \epsilon_4}\right) = 1.$$

One can easily verify that

$$\lim_{N\to\infty} \mathbb{E}[d_1] = K, \quad \lim_{N\to\infty} \frac{\mathbb{E}[d_1(d_1-1)]}{\mathbb{E}[d_1]} = K.$$

Therefore, we can choose $\epsilon_1, \epsilon_2, \epsilon_3, \epsilon_4$ appropriately such that (note that $\left|\sqrt{\frac{4\alpha+K^2}{2}} - K\right| > 0$ since $K^2 > 4\alpha$)

$$\lim_{N\to\infty} \mathbb{P}\left(\left|(g_N^w)'(1) - K\right| \le \left|\sqrt{\frac{4\alpha+K^2}{2}} - K\right|\right) = 1.$$

We can similarly obtain

$$\lim_{N\to\infty} \mathbb{P}\left(\left|(g_N^m)'(1) - \frac{K}{\alpha}\right| \le \left|\sqrt{\frac{4\alpha+K^2}{2\alpha^2}} - \frac{K}{\alpha}\right|\right) = 1.$$

Now consider condition 4. Recall that q_N^w is defined as the smallest nonnegative root of $s = f_N^w(f_N^m(s))$. One can easily verify that $1 = f_N^w(f_N^m(1))$, hence $q_N^w \in (0, 1]$. Using the Uniform Law of Large Numbers (ULLN) from empirical process theory (see, e.g., [54][56]), we can get that $f_N^m(\cdot)$ uniformly converges to some function $f^m(\cdot)$ on $[0, 2]$, where $f^m(\cdot)$ is the probability generating function of Poisson(Ku/α) where $u\sim$Uniform$(0, 1)$. Denote $\bar{u} = f^m(2)$. Also, by ULLN, $f_N^w(\cdot)$ uniformly converges to some function $f^w(\cdot)$ on $[0, \bar{u}]$, where $f^w(\cdot)$ is the probability generating function of Poisson(Ku') where $u' \sim$Uniform$(0, 1)$. By the branching process theory (e.g., see [10][24]), the smallest nonnegative root of $s = f^w(f^m(s))$, denoted by $q^w \equiv q^w(\alpha, K)$, is strictly less than one if the branching factor $\left.\frac{df^w(f^m(s))}{ds}\right|_{s=1} = \frac{K^2}{4\alpha} > 1$. In this case, since $f_N^w(f_N^m(\cdot))$ is continuous, by the continuous mapping theorem (see, e.g., [54]), q_N^w converges to $q^w \in (0,1)$. The first part of condition 4 hence follows. The case with q_N^m is similar, and condition 4 hence follows.

Finally, consider condition 5. Note that $f_N^{w\prime}(f_N^m(q_N^w))f_N^{m\prime}(q_N^w)$ is the derivative of $f_N^w(f_N^m(\cdot))$ evaluated at q_N^w. Since $f_N^w(f_N^m(\cdot))$ uniformly converges to $f^w(f^m(\cdot))$ on $[0,1]$, q_N^w converges to $q^w \in (0,1)$, and $f^w(f^m(\cdot))$ is continuous, we have that $f_N^{w\prime}(f_N^m(q_N^w))f_N^{m\prime}(q_N^w)$ converges to $f^{w\prime}(f^m(q^w))f^{m\prime}(q^w)$. By the branching process theory (e.g., see Theorem 3 in Chapter I, Part D, Section 12 of [10]), $\frac{d(f^w(f^m(q^w s))/q^w)}{ds}\Big|_{s=1}$ is the branching factor of a subcritical branching process, denoted by $\tilde{w} \equiv \tilde{w}(\alpha, K) \in (0,1)$. The result hence follows. The second part of condition 5 is similar.

□

E.1.2 Probabilistic bounds

We prove several probabilistic bounds when the degree sequence is regular. Recall constants $a > 1, \beta > 1, \gamma \in (0,1), \kappa \in (0,1), \xi \in (0,1), \mu \in (0,1)$ that depend on (K,α) from Lemma 13. Define event

$$\mathcal{R} \triangleq \{\text{the degree sequence of } G_N(\alpha, K, p) \text{ is } (a,\beta,\gamma,\kappa,\xi,\mu)\text{-regular}\}. \tag{27}$$

From Lemma 13, we know that $\mathbb{P}(\mathcal{R}) \to 1$ as $N \to \infty$.

The first result upper-bounds the average actual degree of all revealed nodes in $\mathcal{P}_t^T$.

Lemma 14 (Upper bound on $d^T(t)$). *Let $d^T(t)$ be the sum of the actual degree of all revealed nodes in $\mathcal{P}_t^T$. Also recall constants a, β from the definition of $\mathcal{R}$ in Eq. (27). Then, for any $\epsilon \in (0,1)$, there exists some constant $\overline{W} \equiv \overline{W}(\epsilon, a, \beta) \in (0,\infty)$ such that, for event*

$$\mathcal{H}_t := \{d^T(t) \le (t+1)\overline{W}\},$$

we have, for all $t \ge 0$ and any degree sequence $\vec{d} \in \mathcal{R}$,

$$\mathbb{P}(\overline{\mathcal{H}_t}|\vec{d} \in \mathcal{R}) \le \epsilon^{t+1}.$$

In particular,

$$\mathbb{P}(\overline{\mathcal{H}_t}|\mathcal{R}) \le \epsilon^{t+1}.$$

Proof of Lemma 14. We will show that the same result holds for any empirical degree sequence $\vec{d} \in \mathcal{R}$.

Consider an empirical degree sequence $(d_i)_{i=1}^N, (d'_j)_{j=1}^{\alpha N}$ of $G_N(\alpha, K, p)$ satisfying $\mathcal{R}$ and its corresponding empirical (edge-perspective) woman degree distribution F_w and (edge-perspective) man degree distribution F_m. Observe that there are at most $t+1$ revealed nodes in $\mathcal{P}_t^T$, and the actual degree of each of these nodes (unconditional on the information of $\mathcal{P}_t^T$) is independently drawn from either F_w or F_m. Let d_i be the actual degree of the ith revealed node in $\mathcal{P}_t^T$ if it is a woman node, and d'_i be the actual degree of the ith revealed node in $\mathcal{P}_t^T$ if it is a man node. If the ith revealed node in $\mathcal{P}_t^T$ is a woman node, we also draw an independent sample d'_i from F_m. Similarly, if the ith revealed node in $\mathcal{P}_t^T$ is a man node, we also draw an independent sample d_i from F_w. Furthermore, if there are less than $t+1$ revealed nodes (i.e., ALG-G stops before t), say, $s < t+1$ nodes, we also draw independent samples $d_{s+1}, ..., d_{t+1}$ from F_w and independent samples $d'_{s+1}, ..., d'_{t+1}$ from F_m. Therefore, $d_1, ..., d_{t+1}$ are i.i.d. random variables following F_w, and $d'_1, ..., d'_{t+1}$ are i.i.d. random variables (also independent from $d_1, ..., d_t$) following F_m. Moreover, we have $d^T(t) \le \sum_{i=1}^{t+1}(d_i + d'_i)$.

By independence, the probability generating function of $\sum_{i=1}^{t+1}(d_i + d'_i)$ evaluated at s is equal to $s^{2t+2} g_N^w(s)^{t+1} g_N^m(s)^{t+1}$, where $g_N^w(\cdot), g_N^m(\cdot)$ are the probability generating functions of $d_1 - 1, d'_1 - 1$, respectively (see Definifion 8). Conditional on the event $\mathcal{R}$, we know that $g_N^w(a) \le \beta$ and $g_N^m(a) \le \beta$, where $a > 1, \beta > 1$. Therefore, by the Chernoff bound,

$$\begin{aligned}\mathbb{P}(\overline{\mathcal{H}_t}|\vec{d}\in\mathcal{R}) &\le \mathbb{P}(\sum_{i=1}^{t+1}(d_i + d'_i) > \overline{W}(t+1)|\vec{d}\in\mathcal{R}) \\ &\le \frac{a^{2t+2} g_N^w(a)^{t+1} g_N^m(a)^{t+1}}{a^{(t+1)\overline{W}}} \\ &\le \frac{a^{2t+2}\beta^{2t+2}}{a^{\overline{W}(t+1)}} = (\beta^2 a^{-(\overline{W}-2)})^{t+1}.\end{aligned}$$

Since $a > 1$, $\beta^2 a^{-(\overline{W}-2)}$ can be made arbitrarily small by choosing $\overline{W}$ big enough. The choice of $\overline{W}$ does not depend on the degree sequence $(d_i)_{i=1}^N, (d'_j)_{j=1}^{\alpha N}$ except that it satisfies $\mathcal{R}$. The result hence follows.

□

We next present an upper bound for the total progeny of the branching process conditional on it being extinctive. Note that the total progeny of a tree is defined as the total number of nodes in that tree.

Lemma 15 (Upper bound on the extinctive tree size). *Consider an empirical degree sequence of $G_N(\alpha, K, p)$ and its resulting empirical distribution F_w on the (edge-perspective) number of better edges of a woman node, as well as the empirical distribution F_m on the (edge-perspective) number of better edges of a man node. Then consider the branching process whose offspring distribution for woman nodes and man nodes independently follow F_w and F_m, respectively. Let $\tau_1, \tau_2, \ldots$ be independent draws of the total progeny of this branching process rooted at a woman node, and $\tilde{\tau}_1, \tilde{\tau}_2, \ldots$ be independent draws of the total progeny of this branching process rooted at a man node. Recall constant μ from the definition of $\mathcal{R}$ in Eq. (27). For any $t \ge 1$, we have for any degree sequence $\vec{d} \in \mathcal{R}$,*

$$\mathbb{P}(\tau_1 > t|\tau_1 < \infty, \vec{d}\in\mathcal{R}) \le \frac{1}{(1-\mu)t}.$$

Moreover, for any $\epsilon_\tau \in (0,1)$, there exists some large constant $\overline{M}_\tau \equiv \overline{M}_\tau(\epsilon_\tau, \mu) < \infty$, such that, for any $n \ge 1$ and any degree sequence $\vec{d} \in \mathcal{R}$,

$$\mathbb{P}\left(\sum_{i=1}^{n}(\tau_i + \tilde{\tau}_i) > \overline{M}_\tau n \,\middle|\, \sum_{i=1}^{n}(\tau_i + \tilde{\tau}_i) < \infty, \vec{d}\in\mathcal{R}\right) \le \epsilon_\tau.$$

Proof. We will show that the same result holds for any empirical degree sequence $\vec{d} \in \mathcal{R}$.

We first consider $\mathbb{E}[\tau_1|\tau_1 < \infty, \vec{d}\in\mathcal{R}]$ and $\mathbb{E}[\tilde{\tau}_1|\tilde{\tau}_1 < \infty, \vec{d}\in\mathcal{R}]$. Let $f_w(\cdot)$ and $f_m(\cdot)$ be the probability generating functions of the woman nodes' offspring distribution and the man nodes' offspring distribution, respectively. From classic branching process theory, e.g., see [10][24], we know that

$$\mathbb{E}[\tau_1|\tau_1 < \infty, \vec{d}\in\mathcal{R}] = \frac{1}{1 - \tilde{f_w}'(1)},$$

where $\tilde{f_w}(\cdot)$ is defined as

$$\tilde{f_w}(s) \triangleq \frac{f_w(f_m(sq_w))}{q_w},$$

with q_w being the smallest nonnegative root of $s = f_w(f_m(s))$. Conditional on $\mathcal{R}$, we have $f_w'(f_m(q_w))f_m'(q_w) \le \mu < 1$. Therefore,

$$\mathbb{E}[\tau_1|\tau_1 < \infty, \vec{d} \in \mathcal{R}] = \frac{1}{1 - \tilde{f_w}'(1)} = \frac{1}{1 - f_w'(f_m(q_w))f_m'(q_w)} \le \frac{1}{1-\mu}. \tag{28}$$

Similarly,

$$\mathbb{E}[\tilde{\tau}_1|\tilde{\tau}_1 < \infty, \vec{d} \in \mathcal{R}] = \frac{1}{1 - \tilde{f_m}'(1)},$$

where $\tilde{f}_m(\cdot)$ is defined as

$$\tilde{f}_m(s) \triangleq \frac{f_m(f_w(sq_m))}{q_m},$$

with q_m being the smallest nonnegative root of $s = f_m(f_w(s))$. Conditional on event $\mathcal{R}$, we know that $f_m'(f_w(q_w))f_w'(q_w) \le \mu < 1$. Therefore,

$$\mathbb{E}[\tilde{\tau}_1|\tilde{\tau}_1 < \infty, \vec{d} \in \mathcal{R}] = \frac{1}{1 - \tilde{f_m}'(1)} = \frac{1}{1 - f_m'(f_w(q_m))f_w'(q_m)} \le \frac{1}{1-\mu}. \tag{29}$$

By Markov's inequality and Eq. (28),

$$\begin{aligned}\mathbb{P}(\tau_1 > t|\tau_1 \le \infty, \vec{d} \in \mathcal{R}) &\le \frac{\mathbb{E}[\tau_1|\tau_1 < \infty, \vec{d} \in \mathcal{R}]}{t} \\ &\le \frac{1}{(1-\mu)t}.\end{aligned}$$

Similarly, by Markov's inequality and Eq. (28)–(29), and since $\tau_1, \ldots$ and $\tilde{\tau}_1, \ldots$ are all independent,

$$\begin{aligned}&\mathbb{P}\left(\sum_{i=1}^{n}(\tau_i + \tilde{\tau}_i) > \overline{M}_\tau n \,\middle|\, \sum_{i=1}^{n}(\tau_i + \tilde{\tau}_i) < \infty, \vec{d} \in \mathcal{R}\right) \\ &\le \frac{\mathbb{E}\left[\sum_{i=1}^{n}(\tau_i + \tilde{\tau}_i) \,\middle|\, \sum_{i=1}^{n}(\tau_i + \tilde{\tau}_i) < \infty, \vec{d} \in \mathcal{R}\right]}{\overline{M}_\tau n} \\ &= \frac{\sum_{i=1}^{n}\mathbb{E}\left[\tau_i|\tau_i < \infty, \vec{d} \in \mathcal{R}\right] + \sum_{i=1}^{n}\mathbb{E}\left[\tilde{\tau}_i|\tilde{\tau}_i < \infty, \vec{d} \in \mathcal{R}\right]}{\overline{M}_\tau n} \\ &\le \frac{2n}{(1-\mu)\overline{M}_\tau n} = \frac{2}{(1-\mu)\overline{M}_\tau}\end{aligned}$$

can be made arbitrarily small if we choose $\overline{M}_\tau$ big enough.

□

We need the following bound as well.

Lemma 16 (Upper bound on the number of siblings). *Suppose ALG-T does not terminate, i.e., $\mathcal{D}_\infty^T \neq \emptyset$. Consider the i-th node (according to the sequence of being revealed) in $\mathcal{D}_\infty^T$. Let s_i be the number of its higher-ranked (by the parent's preference) siblings that share the same parent. Recall ξ from the definition of $\mathcal{R}$ in Eq. (27). Then, for any $\epsilon_s \in (0,1)$, there exists some large* *constant* $\overline{M}_s \equiv \overline{M}_s(\epsilon_s, K, \alpha, \xi) \in (0,\infty)$ *such that, for any* $n \geq 1$ *and any degree sequence* $\vec{d} \in \mathcal{R}$,

$$\mathbb{P}\left(\sum_{i=1}^{n} s_i > \overline{M}_s n \middle| \mathcal{D}_\infty^T \neq \emptyset, \vec{d} \in \mathcal{R}\right) \leq \epsilon_s.$$

Proof. We will show that the same result holds for any empirical degree sequence $\vec{d} \in$.

Obviously $s_1 = 0$ since the 1st node (according to the sequence of being revealed) in the active set is the root. We now consider $\mathbb{E}[s_i|\mathcal{D}_\infty^T \neq \emptyset, \vec{d} \in \mathcal{R}]$ for $i \geq 2$. Let X_i be the total number of offsprings (including both revealed and unrevealed) of the $(i-1)$-th node in the active set. Obviously $s_i < X_i$. In particular, $\mathbb{E}[s_i|\mathcal{D}_\infty^T \neq \emptyset, \vec{d} \in \mathcal{R}] < \mathbb{E}[X_i|\mathcal{D}_\infty^T \neq \emptyset, \vec{d} \in \mathcal{R}]$. Moreover,

$$\begin{aligned}\mathbb{E}[X_i|\mathcal{D}_\infty^T \neq \emptyset, \vec{d} \in \mathcal{R}] &= \frac{\mathbb{E}\left[X_i \mathbb{1}\{\mathcal{D}_\infty^T \neq \emptyset\}|\vec{d} \in \mathcal{R}\right]}{\mathbb{P}(\mathcal{D}_\infty^T \neq \emptyset|\vec{d} \in \mathcal{R})} \\ &\leq \frac{\mathbb{E}[X_i|\vec{d} \in \mathcal{R}]}{\mathbb{P}(\mathcal{D}_\infty^T \neq \emptyset|\vec{d} \in \mathcal{R})}.\end{aligned}$$

From the classic branching process theory, e.g., see [10][24], we know that the non-extinction probability

$$\mathbb{P}(\mathcal{D}_\infty^T \neq \emptyset|\vec{d} \in \mathcal{R}) = 1 - q_w,$$

where q_w is the smallest nonnegative root of $s = f_w(f_m(s))$, and $f_w(\cdot), f_m(\cdot)$ are the empirical probability generating functions of the (edge-perspective) number of better edges of a woman node and of a man node, respectively. Conditional on the event $\mathcal{R}$, we have $q_w \leq \xi < 1$, hence

$$\mathbb{P}(\mathcal{D}_\infty^T \neq \emptyset|\vec{d} \in \mathcal{R}) = 1 - q_w \geq 1 - \xi.$$

Also, note that

$$\mathbb{E}[X_i|\vec{d} \in \mathcal{R}] \leq \max\{(g_w)'(1), (g_m)'(1)\},$$

where $g_w(\cdot), g_m(\cdot)$ are the empirical probability generating functions of the (edge-perspective) woman nodes' degree minus 1 and of the (edge-perspective) man nodes' degree minus 1, respectively. Again, conditional on event $\mathcal{R}$, we have $(g_w)'(1) \leq K + \left|\sqrt{\frac{4\alpha+K^2}{2}} - K\right|$ and $(g_m)'(1) \leq \frac{K}{\alpha} + \left|\sqrt{\frac{4\alpha+K^2}{2\alpha^2}} - \frac{K}{\alpha}\right|$. Combine the above, we have that for all $i \geq 2$,

$$\begin{aligned}\mathbb{E}[s_i|\mathcal{D}_\infty^T \neq \emptyset, \vec{d} \in \mathcal{R}] &< \mathbb{E}[X_i|\mathcal{D}_\infty^T \neq \emptyset, \vec{d} \in \mathcal{R}] \\ &\leq \frac{\max\left\{K + \left|\sqrt{\frac{4\alpha+K^2}{2}} - K\right|, \frac{K}{\alpha} + \left|\sqrt{\frac{4\alpha+K^2}{2\alpha^2}} - \frac{K}{\alpha}\right|\right\}}{1-\xi}.\end{aligned}$$

We now prove the lemma statement. By Markov's inequality,

$$\begin{aligned}\mathbb{P}\left(\sum_{i=1}^{n} s_i > \overline{M}_s n \middle| \mathcal{D}_\infty^T \neq \emptyset, \vec{d} \in \mathcal{R}\right) &\leq \frac{\mathbb{E}\left[\sum_{i=1}^n s_i \middle| \mathcal{D}_\infty^T \neq \emptyset, \vec{d} \in \mathcal{R}\right]}{\overline{M}_s n} \\ &= \frac{\sum_{i=1}^n \mathbb{E}[s_i|\mathcal{D}_\infty^T \neq \emptyset, \vec{d} \in \mathcal{R}]}{\overline{M}_s n}\end{aligned}$$

$$\leq \frac{n \cdot \max\left\{K + \left|\sqrt{\frac{4\alpha+K^2}{2}} - K\right|, \frac{K}{\alpha} + \left|\sqrt{\frac{4\alpha+K^2}{2\alpha^2}} - \frac{K}{\alpha}\right|, 1\right\}}{(1-\xi)\overline{M}_s n}$$
$$= \frac{\max\left\{K + \left|\sqrt{\frac{4\alpha+K^2}{2}} - K\right|, \frac{K}{\alpha} + \left|\sqrt{\frac{4\alpha+K^2}{2\alpha^2}} - \frac{K}{\alpha}\right|, 1\right\}}{(1-\xi)\overline{M}_s}.$$

Since $K, \alpha > 0, \xi < 1$ are all constants, the above probability upper bound can be made arbitrarily small if we choose $\overline{M}_s$ sufficiently big. □

Lemma 15–16 can be used to show the next lower-bound on the length of the active set $|\mathcal{D}_t^T|$. Lemma 8 is also part of the following statement.

Lemma 17 **($|\mathcal{D}_t^T|$ is linear in t).** *Recall constants $\mu \in (0,1), \xi \in (0,1)$ from the definition of $\mathcal{R}$ in Eq. (27). Define event*

$$\mathcal{C}_t \triangleq \{\mathcal{D}_t^T \neq \emptyset\}.$$

Then, for any $t \geq 1$ and any degree sequence $\vec{d} \in \mathcal{R}$,

$$\mathbb{P}(\mathcal{C}_t | \vec{d} \in \mathcal{R}) \geq 1 - \xi > 0, \quad \mathbb{P}(\overline{\mathcal{C}_\infty} | \mathcal{C}_t, \vec{d} \in \mathcal{R}) \leq \frac{\xi}{(1-\xi)(1-\mu)t}.$$

In particular,

$$\mathbb{P}(\mathcal{C}_t | \mathcal{R}) \geq 1 - \xi > 0, \quad \mathbb{P}(\overline{\mathcal{C}_\infty} | \mathcal{C}_t, \mathcal{R}) \leq \frac{\xi}{(1-\xi)(1-\mu)t}.$$

Also, for any $\epsilon_l \in (0,1)$, there exists $\delta \equiv \delta(\epsilon_l, K, \alpha, \mu, \xi) \in (0,1)$, and $\underline{t}_0 \equiv \underline{t}_0(\epsilon_l, \xi, \mu) \in (0, \infty)$ such that, for event

$$\mathcal{L}_t \triangleq \{|\mathcal{D}_t^T| \geq \delta t\},$$

we have, for any $t \geq t_0 \geq \underline{t}_0$ and any degree sequence $\vec{d} \in \mathcal{R}$,

$$\mathbb{P}(\overline{\mathcal{L}_t} | \mathcal{C}_{t_0}, \vec{d} \in \mathcal{R}) \leq \epsilon_l.$$

In particular,

$$\mathbb{P}(\overline{\mathcal{L}_t} | \mathcal{C}_{t_0}, \mathcal{R}) \leq \epsilon_l.$$

Proof of Lemma 17. We will show the same result holds for any empirical degree sequence $\vec{d} \in \mathcal{R}$.

Note that $\mathcal{C}_\infty \subseteq \mathcal{C}_t$ for any $t \geq 0$, hence $\mathbb{P}(\mathcal{C}_t | \vec{d} \in \mathcal{R}) \geq \mathbb{P}(\mathcal{C}_\infty | \vec{d} \in \mathcal{R})$. From branching process theory, e.g., see [10][24], the probability of event $\mathcal{C}_\infty$ is given by $1 - q_N^w$, which is the smallest nonnegative root of $s = f_N^w(f_N^m(s))$, where $f_N^w(\cdot), f_N^m(\cdot)$ are the empirical probability generating functions of the (edge-perspective) number of better edges of a woman node and of a man node, respectively. Conditional on the event $\mathcal{R}$, we have $q_N^w \leq \xi < 1$, hence

$$\mathbb{P}(\mathcal{C}_t | \vec{d} \in \mathcal{R}) \geq \mathbb{P}(\mathcal{C}_\infty | \vec{d} \in \mathcal{R}) \geq 1 - \xi > 0.$$

Now consider $\mathbb{P}(\overline{\mathcal{C}_\infty}|\mathcal{C}_t, \vec{d} \in \mathcal{R})$. Note that

$$\begin{aligned}
\mathbb{P}(\overline{\mathcal{C}_\infty}|\mathcal{C}_t, \vec{d} \in \mathcal{R}) &= \frac{\mathbb{P}(\overline{\mathcal{C}_\infty}, \mathcal{C}_t, \vec{d} \in \mathcal{R})}{\mathbb{P}(\mathcal{C}_t, \vec{d} \in \mathcal{R})} \\
&\leq \frac{\mathbb{P}(\overline{\mathcal{C}_\infty}, \mathcal{C}_t, \vec{d} \in \mathcal{R})}{\mathbb{P}(\mathcal{C}_\infty, \vec{d} \in \mathcal{R})} \\
&= \frac{\mathbb{P}(\mathcal{C}_t|\overline{\mathcal{C}_\infty}, \vec{d} \in \mathcal{R})\mathbb{P}(\overline{\mathcal{C}_\infty}|\vec{d} \in \mathcal{R})\mathbb{P}(\vec{d} \in \mathcal{R})}{\mathbb{P}(\mathcal{C}_\infty|\vec{d} \in \mathcal{R})\mathbb{P}((d_i)_{i=1}^N, (d'_j)_{j=1}^{\alpha N}, \mathcal{R})} \\
&= \frac{\mathbb{P}(\mathcal{C}_t|\overline{\mathcal{C}_\infty}, \vec{d} \in \mathcal{R})\mathbb{P}(\overline{\mathcal{C}_\infty}|\vec{d} \in \mathcal{R})}{\mathbb{P}(\mathcal{C}_\infty|\vec{d} \in \mathcal{R})}.
\end{aligned}$$

From the previous argument, we know that

$$\mathbb{P}(\overline{\mathcal{C}_\infty}|\vec{d} \in \mathcal{R}) \leq \xi, \quad \mathbb{P}(\mathcal{C}_\infty|\vec{d} \in \mathcal{R}) \geq 1 - \xi > 0, \quad \frac{\mathbb{P}(\overline{\mathcal{C}_\infty}|\vec{d} \in \mathcal{R})}{\mathbb{P}(\mathcal{C}_\infty|\vec{d} \in \mathcal{R})} \leq \frac{\xi}{1-\xi}.$$

Therefore it remains to bound $\mathbb{P}(\mathcal{C}_t|\overline{\mathcal{C}_\infty}, \vec{d} \in \mathcal{R})$. Observe that conditional on $\overline{\mathcal{C}_\infty}$, i.e., extinction, the event $\mathcal{C}_t$ is equivalent to ALG-T not yet terminated at t after revealing $t+1$ nodes in total, i.e., the final tree size τ_1 rooted at the focal node is at least $t+1$. Therefore,

$$\begin{aligned}
\mathbb{P}(\mathcal{C}_t|\overline{\mathcal{C}_\infty}, \vec{d} \in \mathcal{R}) &= \mathbb{P}(\tau_1 > t|\overline{\mathcal{C}_\infty}, \vec{d} \in \mathcal{R}) \\
&= \mathbb{P}(\tau_1 > t|\tau_1 < \infty, \vec{d} \in \mathcal{R}) \\
&\text{(by Lemma 15)} \\
&\leq \frac{1}{(1-\mu)t},
\end{aligned}$$

where $\mu \in (0,1)$ is a constant defined in $\mathcal{R}$ in Eq. (27). Thus

$$\mathbb{P}(\overline{\mathcal{C}_\infty}|\mathcal{C}_t, \vec{d} \in \mathcal{R}) \leq \frac{\xi}{(1-\xi)(1-\mu)t}.$$

Finally, consider

$$\begin{aligned}
\mathbb{P}(\overline{\mathcal{L}_t}|\mathcal{C}_{t_0}, \vec{d} \in \mathcal{R}) &= \mathbb{P}(\overline{\mathcal{L}_t}, \mathcal{C}_\infty|\mathcal{C}_{t_0}, \vec{d} \in \mathcal{R}) + \mathbb{P}(\overline{\mathcal{L}_t}, \overline{\mathcal{C}_\infty}|\mathcal{C}_{t_0}, \vec{d} \in \mathcal{R}) \\
&\leq \frac{\mathbb{P}(\overline{\mathcal{L}_t}, \mathcal{C}_\infty, \mathcal{C}_{t_0}, \vec{d} \in \mathcal{R})}{\mathbb{P}(\mathcal{C}_{t_0}, \vec{d} \in \mathcal{R})} + \mathbb{P}(\overline{\mathcal{C}_\infty}|\mathcal{C}_{t_0}, \vec{d} \in \mathcal{R}) \\
&\leq \frac{\mathbb{P}(\overline{\mathcal{L}_t}, \mathcal{C}_\infty, \vec{d} \in \mathcal{R})}{\mathbb{P}(\mathcal{C}_\infty, \vec{d} \in \mathcal{R})} + \mathbb{P}(\overline{\mathcal{C}_\infty}|\mathcal{C}_{t_0}, \vec{d} \in \mathcal{R}) \\
&\leq \mathbb{P}(\overline{\mathcal{L}_t}|\mathcal{C}_\infty, \vec{d} \in \mathcal{R}) + \mathbb{P}(\overline{\mathcal{C}_\infty}|\mathcal{C}_{t_0}, \vec{d} \in \mathcal{R}).
\end{aligned}$$

We already know that $\mathbb{P}(\overline{\mathcal{C}_\infty}|\mathcal{C}_{t_0}, \vec{d} \in \mathcal{R}) \leq \frac{\xi}{(1-\xi)(1-\mu)t_0}$. It remains to bound $\mathbb{P}(\overline{\mathcal{L}_t}|\mathcal{C}_\infty, \vec{d} \in \mathcal{R})$. First consider time ∞ of ALG-T. Since we are conditioning on event $\mathcal{C}_\infty$, there must be a non-empty active set in $\mathcal{P}_\infty^T$. Consider the ith node (according to the sequence of being revealed) in $\mathcal{D}_\infty^T$ and its number of revealed siblings s_i, for $i = 1, ..., |\mathcal{D}_t^T|$. Denote $n_t = \sum_{i=1}^{|\mathcal{D}_t^T|} s_i$ the total number of revealed siblings of these nodes and index the sibling by $j = 1, 2, ..., n_t$. For each sibling node j, let τ_j be the size of the final extinctive subtree rooted at j. Now we go back to time t of ALG-T. If $|\mathcal{D}_t^T| < \delta t$ and $\mathcal{C}_\infty = 1$, then it must be that $t - |\mathcal{D}_t^T| > (1-\delta)t$. Note that $t - |\mathcal{D}_t^T|$ is the total number of inactive nodes revealed at time t of ALG-T. These inactive nodes must belong to one of the extinctive subtrees whose roots are siblings of nodes in $\mathcal{D}_t^T$. In fact, an important observation is that the total number of inactive nodes revealed in $\mathcal{P}_t^T$ must be no more than the total number of inactive nodes revealed until the $|\mathcal{D}_t^T|$-th last node (according to the sequence of being revealed) in $\mathcal{D}_\infty^T$ is revealed. That is,

$$t - |\mathcal{D}_t^T| \leq \sum_{j=1}^{n_t} \tau_j, \text{ where } n_t = \sum_{i=1}^{|\mathcal{D}_t^T|} s_i.$$

Pick any constant $\rho \in (0,1)$, then

$$\begin{aligned}
\mathbb{P}(\overline{\mathcal{L}_t}|\mathcal{C}_\infty, \vec{d} \in \mathcal{R}) &= \mathbb{P}(|\mathcal{D}_t^T| < \delta t|\mathcal{C}_\infty, \vec{d} \in \mathcal{R}) \\
&= \mathbb{P}(t - |\mathcal{D}_t^T| > (1-\delta)t, \quad |\mathcal{D}_t^T| < \delta t|\mathcal{C}_\infty, \vec{d} \in \mathcal{R}) \\
&\leq \mathbb{P}(\sum_{j=1}^{n_t} \tau_j > (1-\delta)t, \quad |\mathcal{D}_t^T| < \delta t|\mathcal{C}_\infty, \vec{d} \in \mathcal{R}) \\
&\leq \mathbb{P}(\sum_{j=1}^{n_t} \tau_j > (1-\delta)t, \quad n_t \leq \rho t, \quad |\mathcal{D}_t^T| < \delta t|\mathcal{C}_\infty, \vec{d} \in \mathcal{R}) \\
&\quad + \mathbb{P}(\sum_{j=1}^{n_t} \tau_j > (1-\delta)t, \quad n_t > \rho t, \quad |\mathcal{D}_t^T| < \delta t|\mathcal{C}_\infty, \vec{d} \in \mathcal{R}).
\end{aligned}$$

We now bound each term on the RHS. Observe that conditional on $\mathcal{C}_\infty$ and $\mathcal{R}$, $\tau_1, ...$ are independent draws from a branching process rooted at either a woman node or a man node, conditional on being extinctive. Therefore we can apply Lemma 15 to bound the first term:

$$\mathbb{P}(\sum_{j=1}^{n_t} \tau_j > (1-\delta)t, \quad n_t \leq \rho t, \quad |\mathcal{D}_t^T| < \delta t|\mathcal{C}_\infty, \vec{d} \in \mathcal{R})$$

$$\leq \mathbb{P}(\sum_{j=1}^{\rho t} \tau_j > (1-\delta)t, \quad n_t \leq \rho t, \quad |\mathcal{D}_t^T| < \delta t|\mathcal{C}_\infty, \vec{d} \in \mathcal{R})$$

(If $\rho t > n_t$, we sample additional independent extinctive tree sizes from the woman-root branching process.)

$$\leq \mathbb{P}(\sum_{j=1}^{\rho t} \tau_j > \frac{1-\delta}{\rho} \rho t|\mathcal{C}_\infty, \vec{d} \in \mathcal{R}).$$

By Lemma 15, for any $\epsilon_\tau \in (0,1)$, there exists some large constant $\overline{M}_\tau > 0$ depending on ϵ_τ, μ, such that the above probability is smaller than ϵ_τ if $\frac{1-\delta}{\rho} \geq \overline{M}_\tau$.

Now consider the second term:

$$\begin{aligned}
&\mathbb{P}(\sum_{j=1}^{n_t} \tau_j > (1-\delta)t, \quad n_t > \rho t, \quad |\mathcal{D}_t^T| < \delta t|\mathcal{C}_\infty, \vec{d} \in \mathcal{R}) \\
&= \mathbb{P}(\sum_{j=1}^{n_t} \tau_j > (1-\delta)t, \quad \sum_{i=1}^{|\mathcal{D}_t^T|} s_i > \rho t, \quad |\mathcal{D}_t^T| < \delta t|\mathcal{C}_\infty, \vec{d} \in \mathcal{R}) \\
&\leq \mathbb{P}(\sum_{i=1}^{\delta t} s_i > \frac{\rho}{\delta} \delta t|\mathcal{C}_\infty, \vec{d} \in \mathcal{R}).
\end{aligned}$$

By Lemma 16, for any $\epsilon_s \in (0,1)$, there exists some large constant $\overline{M}_s > 0$ depending on $\epsilon_s, K, \alpha, \xi$, such that the above probability is smaller than ϵ_s if $\frac{\rho}{\delta} \geq \overline{M}_s$.

By straightforward algebra, one can verify that both $\frac{1-\delta}{\rho} \geq \overline{M}_\tau$ and $\frac{\rho}{\delta} \geq \overline{M}_s$ are satisfied if $\rho \leq \frac{\overline{M}_s}{\overline{M}_\tau \overline{M}_s+1}$ and $\delta \leq \frac{1}{\overline{M}_\tau \overline{M}_s+1}$. This means that for any $\epsilon_l \in (0,1)$, there exists $\delta > 0$ depending on $\epsilon_l, K, \alpha, \mu, \xi$, such that

$$\mathbb{P}(|\mathcal{D}_t^T| < \delta t | \mathcal{C}_\infty, \vec{d} \in \mathcal{R}) \leq \frac{\epsilon_l}{2},$$

and hence

$$\mathbb{P}(|\mathcal{D}_t^T| < \delta t | \mathcal{C}_{t_0}, \vec{d} \in \mathcal{R}) \leq \frac{\epsilon_l}{2} + \frac{\xi}{(1-\xi)(1-\mu)t_0}.$$

Furthermore, if $t_0 \geq \frac{2\xi}{(1-\xi)(1-\mu)\epsilon_l}$, then

$$\mathbb{P}(|\mathcal{D}_t^T| < \delta t | \mathcal{C}_{t_0}, \vec{d} \in \mathcal{R}) \leq \epsilon_l.$$

□

Finally, we bound the average unrevealed degree of nodes in the active set in $\mathcal{P}_t^T$. We need the following bound.

Lemma 18 (**Upper bound on the number of inferior edges).** *Consider an empirical degree sequence of $G_N(\alpha, K, p)$ and its resulting empirical distribution F_w on the (edge-perspective) number of inferior edges of a woman node, as well as the empirical distribution F_m on the (edge-perspective) number of inferior edges of a man node. Let $d_1^{\inf}, d_2^{\inf}, \ldots$ be independent draws from F_w and $\tilde{d}_1^{\inf}, \tilde{d}_2^{\inf}, \ldots$ be independent draws from F_m. Recall constant $\kappa \in (0,1)$ from the definition of $\mathcal{R}$ in Eq. (27). Then there exists a constant $\underline{W} \equiv \underline{W}(\kappa) \in (0, \infty)$ such that for any degree sequence $\vec{d} \in \mathcal{R}$,*

$$\mathbb{P}\left(\sum_{i=1}^n d_i^{\inf} < \underline{W} n \middle| \vec{d} \in \mathcal{R}\right) \leq \left(\frac{\kappa+1}{2}\right)^n, \quad \mathbb{P}\left(\sum_{i=1}^n \tilde{d}_i^{\inf} < \underline{W} n \middle| \vec{d} \in \mathcal{R}\right) \leq \left(\frac{\kappa+1}{2}\right)^n.$$

Proof. We will show that the same result holds for any empirical degree sequence $\vec{d} \in \mathcal{R}$.

Recall the functions $f_N^w(\cdot)$ and $f_N^m(\cdot)$ from Definition 8. It can be easily verified that $f_N^w(\cdot)$ and $f_N^m(\cdot)$ are the probability generating functions of F_w and F_m (suppose the empirical degree sequence of $G_N(\alpha, K, p)$ is denoted by $(d_i)_{i=1}^N, (d'_j)_{j=1}^{\alpha N}$) respectively. Conditional on the event $\mathcal{R}$, we know that $f_N^w(\gamma) \leq \kappa$ and $f_N^m(\gamma) \leq \kappa$, where $\gamma \in (0,1)$, $\kappa \in (0,1)$. Therefore, since $d_1^{\inf}, d_2^{\inf}, \ldots$ and $\tilde{d}_1^{\inf}, \tilde{d}_2^{\inf}, \ldots$ are all independent, we can apply the Chernoff bound to get

$$\mathbb{P}\left(\sum_{i=1}^n d_i^{\inf} < \underline{W} n | \vec{d} \in \mathcal{R}\right) \leq \frac{f_N^w(\gamma)^n}{\gamma^{\underline{W} n}} \leq \frac{\kappa^n}{\gamma^{\underline{W} n}} = \left(\kappa \gamma^{-\underline{W}}\right)^n$$

and

$$\mathbb{P}\left(\sum_{i=1}^n \tilde{d}_i^{\inf} < \underline{W} n | \vec{d} \in \mathcal{R}\right) \leq \frac{f_N^m(\gamma)^n}{\gamma^{\underline{W} n}} \leq \frac{\kappa^n}{\gamma^{\underline{W} n}} = \left(\kappa \gamma^{-\underline{W}}\right)^n.$$

Since $\gamma \in (0,1)$, by choosing $\underline{W} > 0$ small enough, $\kappa \gamma^{-\underline{W}}$ can be made arbitrarily close to κ, in particular, smaller than $\frac{\kappa+1}{2} < 1$. □

Lemma 19 (Upper bound on the number of unrevealed edges). *Consider $\mathcal{P}_t^T$ and the node s in it who will reveal the next offspring. Let $d_A^{T,\text{unre}}(t)$ be the total number of unrevealed edges of s's ancestor (excluding the parent) nodes with the opposite type to s (recall that node types are binary: there are man nodes and woman nodes) in the active set in $\mathcal{P}_t^T$. Recall constants κ, μ, ξ from the definition of $\mathcal{R}$ in Eq. (27). Then, for any $\epsilon_a \in (0,1)$, there exists $\delta_0 \equiv \delta_0(\epsilon_a, K, \alpha, \kappa, \mu, \xi) \in (0, \infty)$ and $\underline{t}_0 \equiv \underline{t}_0(\epsilon_a, K, \alpha, \kappa, \mu, \xi)$ such that, for event*

$$\mathcal{J}_t := \{d_A^{T,\text{unre}}(t) \geq \delta_0 t\},$$

we have, for any $t \geq t_0 \geq \underline{t}_0$ and any degree sequence $\vec{d} \in \mathcal{R}$,

$$\mathbb{P}(\overline{\mathcal{J}_t} | \mathcal{C}_{t_0}, \vec{d} \in \mathcal{R}) \leq \epsilon_a.$$

In particular,

$$\mathbb{P}(\overline{\mathcal{J}_t} | \mathcal{C}_{t_0}, \mathcal{R}) \leq \epsilon_a.$$

Proof of Lemma 19. We will show that the same result holds for any empirical degree sequence $\vec{d} \in \mathcal{R}$.

Note that

$$\begin{aligned}\mathbb{P}(\overline{\mathcal{J}_t} | \mathcal{C}_{t_0}, \vec{d} \in \mathcal{R}) &= \mathbb{P}(\overline{\mathcal{J}_t}, \overline{\mathcal{L}_{t+1}} | \mathcal{C}_{t_0}, \vec{d} \in \mathcal{R}) + \mathbb{P}(\overline{\mathcal{J}_t}, \mathcal{L}_{t+1} | \mathcal{C}_{t_0}, \vec{d} \in \mathcal{R}) \\ &\leq \mathbb{P}(\overline{\mathcal{L}_{t+1}} | \mathcal{C}_{t_0}, \vec{d} \in \mathcal{R}) + \frac{\mathbb{P}(\overline{\mathcal{J}_t}, \mathcal{L}_{t+1} | \vec{d} \in \mathcal{R})}{\mathbb{P}(\mathcal{C}_{t_0} | \vec{d} \in \mathcal{R})}.\end{aligned}$$

We bound each term on the RHS. Recall that $\mathcal{L}_t = \{|\mathcal{D}_t^T| \geq \delta t\}$. By Lemma 17, for any $\epsilon_a \in (0,1)$, there exists $\delta \in (0,1)$ depending on $\epsilon_a, K, \alpha, \mu, \xi$, and t_0 larger than some constant depending on ϵ_a, ξ, μ, such that for any $t \geq t_0$, $\mathbb{P}(\overline{\mathcal{L}_t} | \mathcal{C}_{t_0}, \vec{d} \in \mathcal{R}) \leq \epsilon_a/2$.

Now consider the second term. By Lemma 17, we know that $\mathbb{P}(\mathcal{C}_{t_0} | \vec{d} \in \mathcal{R}) \geq 1 - \xi > 0$ for any $t_0 \geq 1$. We are left to bound $\mathbb{P}(\overline{\mathcal{J}_t}, \mathcal{L}_{t+1} | \vec{d} \in \mathcal{R})$.

Recall the node s in $\mathcal{P}_t^T$ who will reveal the next offspring. Let $\mathcal{I}_A^T(t)$ be the set of s's ancestor (excluding the parent) nodes with the opposite type to s in $\mathcal{D}_t^T$. Consider the set $\mathcal{I}_A^T(t)$ if not empty. Observe that it is either the set of all the man nodes or the set of all the woman nodes in $\mathcal{D}_{t+1}^T$ (excluding the node in $\mathcal{D}_{t+1}^T - \mathcal{D}_t^T$ and its first two ancestors). Denote the set of all the man nodes in $\mathcal{D}_{t+1}^T$ (excluding the node in $\mathcal{D}_{t+1}^T - \mathcal{D}_t^T$ and its first two ancestors) by $\mathcal{M}_t$, and the set of all the woman nodes in $\mathcal{D}_{t+1}^T$ (excluding the node in $\mathcal{D}_{t+1}^T - \mathcal{D}_t^T$ and its first two ancestors) by $\mathcal{W}_t$. Therefore event $\overline{\mathcal{J}_t}$ must be a subset of the union of two events $\overline{\mathcal{J}_t^m}$ and $\overline{\mathcal{J}_t^w}$, where $\overline{\mathcal{J}_t^m}$ is the event that the total unrevealed degree of nodes in $\mathcal{M}_t$ is less than $\delta_0 t$, and $\overline{\mathcal{J}_t^w}$ is the event that the total unrevealed degree of nodes in $\mathcal{W}_t$ is less than $\delta_0 t$. In other words,

$$\begin{aligned}\mathbb{P}(\overline{\mathcal{J}_t}, \mathcal{L}_{t+1} | \vec{d} \in \mathcal{R}) &\leq \mathbb{P}(\overline{\mathcal{J}_t^m} \cup \overline{\mathcal{J}_t^w}, \mathcal{L}_{t+1} | \vec{d} \in \mathcal{R}) \\ &\leq \mathbb{P}(\overline{\mathcal{J}_t^m}, \mathcal{L}_{t+1} | \vec{d} \in \mathcal{R}) + \mathbb{P}(\overline{\mathcal{J}_t^w}, \mathcal{L}_{t+1} | \vec{d} \in \mathcal{R}),\end{aligned}$$

and it suffices to upper-bound each of the two probabilities on the RHS.

For each node i in $\mathcal{W}_t$, let d_i^{inf} denote the number of its unrevealed offsprings that are less preferred by node i than its parent (call it the inferior degree). Obviously, since the unrevealed degree of i must be no smaller than the inferior degree of i, $\overline{\mathcal{J}_t^w} \subseteq \{\sum_{i=1}^{|\mathcal{W}_t|} d_i^{\text{inf}} < \delta_0 t\}$. Similarly, for each node $j \in \mathcal{M}_t$, let $\tilde{d}_j^{\text{inf}}$ be the number of its unrevealed offsprings that are less preferred by node j than its parent, and we have $\overline{\mathcal{J}_t^m} \subseteq \{\sum_{j=1}^{|\mathcal{M}_t|} \tilde{d}_j^{\text{inf}} < \delta_0 t\}$. Also observe that $|\mathcal{W}_t| \geq |\mathcal{D}_t^T|/2 - 2$

and $|\mathcal{M}_t| \geq |\mathcal{D}_t^T|/2 - 2$. Therefore,

$$\begin{aligned}
\mathbb{P}(\overline{\mathcal{J}_t^w}, \mathcal{L}_{t+1} | \vec{d} \in \mathcal{R}) &\leq \mathbb{P}\left(\sum_{i=1}^{|\mathcal{W}_t|} d_i^{\inf} < \delta_0 t, \quad |\mathcal{D}_t^T| \geq (t+1)\delta \middle| \vec{d} \in \mathcal{R} \right) \\
&\leq \mathbb{P}\left(\sum_{i=1}^{|\mathcal{W}_t|} d_i^{\inf} < \delta_0 t, \quad |\mathcal{W}_t| \geq \frac{(t+1)\delta}{2} - 2 \middle| \vec{d} \in \mathcal{R} \right) \\
&\leq \mathbb{P}\left(\sum_{i=1}^{|\mathcal{W}_t|} d_i^{\inf} < \delta_0 t, \quad |\mathcal{W}_t| \geq \frac{\delta t}{2} - 2 \middle| \vec{d} \in \mathcal{R} \right) \\
&\leq \mathbb{P}\left(\sum_{i=1}^{\frac{\delta t}{2}-2} d_i^{\inf} < \delta_0 t \middle| \vec{d} \in \mathcal{R} \right) \\
&\quad \text{(If } \frac{\delta t}{2} - 2 > |\mathcal{W}_t| \text{, we generate additional iid samples from} \\
&\quad \text{distribution } F_w \text{ in Lemma 18)} \\
&= \mathbb{P}\left(\sum_{i=1}^{\frac{\delta t}{2}-2} d_i^{\inf} < \frac{2\delta_0}{\delta - \frac{4}{t}} \cdot \left(\frac{\delta t}{2} - 2 \right) \middle| \vec{d} \in \mathcal{R} \right).
\end{aligned}$$

Observe that when conditioning on $\mathcal{R}$ only, $d_1^{\inf}, ..., d_{\frac{\delta t}{2}-2}^{\inf}$ are i.i.d. samples from distribution F_w in Lemma 18. Therefore, by Lemma 18, there exists a constant $\underline{W} > 0$ depending on κ (see the definition of $\kappa \in (0,1)$ in Eq. (27)), such that the above probability is smaller than $(\frac{\kappa+1}{2})^{\frac{\delta t}{2}-2}$ if $\frac{2\delta_0}{\delta - \frac{4}{t}} \leq \underline{W}$, or equivalently, if $\delta_0 \leq \frac{\underline{W}(\delta - \frac{4}{t})}{2}$. Similarly, we have

$$\mathbb{P}(\overline{\mathcal{J}_t^m}, \mathcal{L}_{t+1} | \vec{d} \in \mathcal{R}) \leq \left(\frac{\kappa+1}{2} \right)^{\frac{\delta t}{2}-2}$$

under the same condition of $\delta_0 \leq \frac{\underline{W}(\delta - \frac{4}{t})}{2}$. This condition reduces to $\delta_0 \leq \frac{\underline{W}\delta}{4}$ when t is sufficiently large, e.g., when $t \geq \frac{8}{\delta}$. Furthermore, since $\kappa < 1$, the probability upper bound $(\frac{\kappa+1}{2})^{\frac{\delta t}{2}-2}$ can be smaller than $\frac{\epsilon_a(1-\xi)}{4}$ if t is sufficiently large, in particular, larger than a constant depending on $\epsilon_a, \kappa, \xi, \delta$.

Recall that $\underline{W}$ is a constant depending on κ from Lemma 18 and δ is a constant depending on $\epsilon_a, K, \alpha, \mu, \xi$ from the first part of this proof. Therefore, combine the above, there exists $\delta_0 > 0$ depending on $\epsilon_a, K, \alpha, \kappa, \mu, \xi$, such that, when t is larger than some constant depending on $\epsilon_a, K, \alpha, \kappa, \mu, \xi$,

$$\frac{\mathbb{P}(\overline{\mathcal{J}_t}, \mathcal{L}_{t+1} | \vec{d} \in \mathcal{R})}{\mathbb{P}(\mathcal{C}_{t_0} | \vec{d} \in \mathcal{R})} \leq \frac{\epsilon_a}{2}.$$

Thus we have the desired result.

□

E.2 Proof of Lemma 21

To couple $\mathcal{P}_t^G$ with $\mathcal{P}_t^T$, we show that each step of ALG-G and ALG-T can be coupled. We first need the following lemma.

Lemma 20 (Uniform randomness). *The bipartite graph after t-rounds of ALG-G is uniformly random given $\mathcal{P}_t^G$ and the degree sequence.*

Proof of Lemma 20. The proof is similar to the proof of Claim 1 in [38]. Given $\mathcal{P}_t^G$ and the degree sequence, consider any two feasible bipartite graphs G_1, G_2 and node 1 which are consistent with the degree sequence and $\mathcal{P}_t^G$ after t rounds of ALG-G. Since G_1 and G_2 have the same number of edges, the probability that the initial graph is G_1 is equal to the probability that the initial graph is G_2. This implies the desired result. $\square$

Given the uniform randomness, we show that the probability distribution from which the next step of ALG-G is drawn closely resembles that of ALG-T, given that the revealed nodes' total actual degree is not too large (event $\mathcal{H}_t$ in Lemma 14). This result is stated in Lemma 6, and we now provide its proof.

Proof of Lemma 6. The initial coupling such that $\mathbb{P}(\tilde{\mathcal{P}}_0^G \neq \mathcal{P}_0^T | \mathcal{R}) = 0$ is trivial.

Consider $t \geq 0$. Let $(d_i)_{i=1}^N, (d'_j)_{j=1}^{\alpha N}$ be the empirical degree sequence. Suppose the next revealed node according to $\mathcal{P}_t^T$ is a man node (this event is measurable with respect to $\mathcal{F}_t$). (We treat the woman node case similarly later.) For any $l, l' \geq 0$ such that $1 + l + l' \leq N$,

$$\begin{aligned}
\nu^G(l,l') &:= \mathbb{P}(\text{The new revealed node in } \mathcal{P}_{t+1}^G \text{ has } l \text{ better (than parent) edges and } l' \text{ worse edges} \\
&\qquad |\mathcal{F}_t \text{ such that } \tilde{\mathcal{P}}_t^G = \mathcal{P}_t^T, (d_i)_{i=1}^N, (d'_j)_{j=1}^{\alpha N}, \text{the next revealed node is a man node}) \\
&\text{(by lemma 20)} \\
&= \frac{\text{number of men nodes not in } \mathcal{P}_t^G \text{ whose degree is } 1+l+l'}{\text{number of remaining unrevealed men half edges at } t} \\
&\text{(denote } \mathcal{I}^G(t) \text{ the set of nodes in } \mathcal{P}_t^G \text{ whose type is the same as the next revealed node,} \\
&d_j^{G,\mathrm{re}} \text{ the revealed degree of node } j \text{ in } \mathcal{P}_t^G) \\
&= \frac{\sum_{j=1}^{\alpha N} \mathbb{1}\{d'_j = 1+l+l'\} - \sum_{j \in \mathcal{I}^G(t)} \mathbb{1}\{d'_j = 1+l+l'\}}{\sum_{j=1}^{\alpha N} d'_j - \sum_{j \in \mathcal{I}^G(t)} d_j^{G,\mathrm{re}}} \\
&= \frac{\sum_{j=1}^{\alpha N} \mathbb{1}\{d'_j = 1+l+l'\}}{\sum_{j=1}^{\alpha N} d'_j} \\
&\quad + \frac{\sum_{j=1}^{\alpha N} \mathbb{1}\{d'_j = 1+l+l'\} - \sum_{j \in \mathcal{I}^G(t)} \mathbb{1}\{d'_j = 1+l+l'\}}{\sum_{j=1}^{\alpha N} d'_j - \sum_{j \in \mathcal{I}^G(t)} d_j^{G,\mathrm{re}}} - \frac{\sum_{j=1}^{\alpha N} \mathbb{1}\{d'_j = 1+l+l'\}}{\sum_{j=1}^{\alpha N} d'_j}.
\end{aligned}$$

In ALG-T, this probability is $\nu^T(l,l') := \frac{\sum_{j=1}^{\alpha N} \mathbb{1}\{d'_j = 1+l+l'\}}{\sum_{j=1}^{\alpha N} d'_j}$ for any $l, l' \geq 0$. We now bound the total variation distance between the two measures. We have

$$\begin{aligned}
\nu^G(l,l') - \nu^T(l,l') &= \frac{\sum_{j=1}^{\alpha N} \mathbb{1}\{d'_j = 1+l+l'\} - \sum_{j \in \mathcal{I}^G(t)} \mathbb{1}\{d'_j = 1+l+l'\}}{\sum_{j=1}^{\alpha N} d'_j - \sum_{j \in \mathcal{I}^G(t)} d_j^{G,\mathrm{re}}} - \frac{\sum_{j=1}^{\alpha N} \mathbb{1}\{d'_j = 1+l+l'\}}{\sum_{j=1}^{\alpha N} d'_j} \\
&\geq \frac{\sum_{j=1}^{\alpha N} \mathbb{1}\{d'_j = 1+l+l'\} - \sum_{j \in \mathcal{I}^G(t)} \mathbb{1}\{d'_j = 1+l+l'\}}{\sum_{j=1}^{\alpha N} d'_j} - \frac{\sum_{j=1}^{\alpha N} \mathbb{1}\{d'_j = 1+l+l'\}}{\sum_{j=1}^{\alpha N} d'_j} \\
&= -\frac{\sum_{j \in \mathcal{I}^G(t)} \mathbb{1}\{d'_j = 1+l+l'\}}{\sum_{j=1}^{\alpha N} d'_j}.
\end{aligned}$$

Also observe that $0 \le \sum_{j\in\mathcal{I}^G(t)} d_j^{G,\mathrm{re}} \le 2t$ since ALG-G reveals at most one edge in one time step. Therefore,

$$\begin{aligned}\nu^G(l,l') - \nu^T(l,l') &= \frac{\sum_{j=1}^{\alpha N} \mathbb{1}\{d'_j = 1+l+l'\} - \sum_{j\in\mathcal{I}^G(t)} \mathbb{1}\{d'_j = 1+l+l'\}}{\sum_{j=1}^{\alpha N} d'_j - \sum_{j\in\mathcal{I}^G(t)} d_j^{G,\mathrm{re}}} - \frac{\sum_{j=1}^{\alpha N} \mathbb{1}\{d'_j = 1+l+l'\}}{\sum_{j=1}^{\alpha N} d'_j} \\ &\le \frac{\sum_{j=1}^{\alpha N} \mathbb{1}\{d'_j = 1+l+l'\} - \sum_{j\in\mathcal{I}^G(t)} \mathbb{1}\{d'_j = 1+l+l'\}}{\sum_{j=1}^{\alpha N} d'_j - 2t} - \frac{\sum_{j=1}^{\alpha N} \mathbb{1}\{d'_j = 1+l+l'\}}{\sum_{j=1}^{\alpha N} d'_j} \\ &\le \frac{\sum_{j=1}^{\alpha N} \mathbb{1}\{d'_j = 1+l+l'\}}{\sum_{j=1}^{\alpha N} d'_j - 2t} - \frac{\sum_{j=1}^{\alpha N} \mathbb{1}\{d'_j = 1+l+l'\}}{\sum_{j=1}^{\alpha N} d'_j}.\end{aligned}$$

By the above two inequalities, we have, for any $l, l' \ge 0$ such that $1+l+l' \le N$,

$$|\nu^G(l,l') - \nu^T(l,l')| \le \frac{\sum_{j=1}^{\alpha N} \mathbb{1}\{d'_j = 1+l+l'\}}{\sum_{j=1}^{\alpha N} d'_j - 2t} - \frac{\sum_{j=1}^{\alpha N} \mathbb{1}\{d'_j = 1+l+l'\}}{\sum_{j=1}^{\alpha N} d'_j} + \frac{\sum_{j\in\mathcal{I}^G(t)} \mathbb{1}\{d'_j = 1+l+l'\}}{\sum_{j=1}^{\alpha N} d'_j}.$$

We now examine the total variation distance $\sum_{l=0}^{N-1}\sum_{l'=0}^{N-1-l} |\nu^G(l,l') - \nu^T(l,l')|$. Observe that

$$\sum_{l=0}^{N-1}\sum_{l'=0}^{N-1-l} \mathbb{1}\{d'_j = 1+l+l'\} = d'_j$$

and

$$\sum_{l=0}^{N-1}\sum_{l'=0}^{N-1-l}\sum_{j=1}^{\alpha N} \mathbb{1}\{d'_j = 1+l+l'\} = \sum_{j=1}^{\alpha N} d'_j.$$

Therefore,

$$\begin{aligned}\sum_{l=0}^{N-1}\sum_{l'=0}^{N-1-l} |\nu^G(l,l') - \nu^T(l,l')| &\le \sum_{l=0}^{N-1}\sum_{l'=0}^{N-1-l}\sum_{j=1}^{\alpha N} \mathbb{1}\{d'_j = 1+l+l'\}\left(\frac{1}{\sum_{j=1}^{\alpha N} d'_j - 2t} - \frac{1}{\sum_{j=1}^{\alpha N} d'_j}\right) \\ &\quad + \frac{\sum_{l=0}^{N-1}\sum_{l'=0}^{N-1-l}\sum_{j\in\mathcal{I}^G(t)} \mathbb{1}\{d'_j = 1+l+l'\}}{\sum_{j=1}^{\alpha N} d'_j} \\ &= \frac{2t}{\sum_{j=1}^{\alpha N} d'_j - 2t} + \frac{\sum_{j\in\mathcal{I}^G(t)} d'_j}{\sum_{j=1}^{\alpha N} d'_j} \\ &\quad \text{(let } d^G(t) \text{ be the total actual degree of all revealed nodes in } \mathcal{P}_t^G) \\ &\le \frac{2t}{\sum_{j=1}^{\alpha N} d'_j - 2t} + \frac{d^G(t)}{\sum_{j=1}^{\alpha N} d'_j}.\end{aligned}$$

Conditional on $\tilde{\mathcal{P}}_t^G = \mathcal{P}_t^T$, we have $d^G(t) = d^T(t)$, where the latter is the total actual degree of all revealed nodes in $\mathcal{P}_t^T$. Moreover, conditional on $\mathcal{H}_t$ (see Lemma 14), we have $d^T(t) \le (t+1)\overline{W}$. Therefore,

$$\sum_{l=0}^{N-1}\sum_{l'=0}^{N-1-l} |\nu^G(l,l') - \nu^T(l,l')| \le \frac{2t}{\sum_{j=1}^{\alpha N} d'_j - 2t} + \frac{(t+1)\overline{W}}{\sum_{j=1}^{\alpha N} d'_j}.$$

Note that $\sum_{j=1}^{\alpha N} d'_j = \alpha N \bar{k}_m$, where $\bar{k}_m$ is the empirical mean of the man node's degree in the bipartite graph. Conditional on $\mathcal{R}$, we know that $\bar{k}_m \geq \frac{K}{\alpha} - \left|\sqrt{\frac{4\alpha+K^2}{2\alpha^2}} - \frac{K}{\alpha}\right|$, hence the above can be rewritten as

$$\sum_{l=0}^{N-1}\sum_{l'=0}^{N-1-l} |\nu^G(l,l') - \nu^T(l,l')| \leq \frac{2t}{\alpha N \bar{k}_m - 2t} + \frac{(t+1)\overline{W}}{\alpha N \bar{k}_m}$$
$$\leq \frac{2t}{\alpha N\left(\frac{K}{\alpha} - \left|\sqrt{\frac{4\alpha+K^2}{2\alpha^2}} - \frac{K}{\alpha}\right|\right) - 2t} + \frac{(t+1)\overline{W}}{\alpha N\left(\frac{K}{\alpha} - \left|\sqrt{\frac{4\alpha+K^2}{2\alpha^2}} - \frac{K}{\alpha}\right|\right)}.$$

Therefore, there exists a constant $\epsilon_c > 0$ that depends $K, \alpha, \overline{W}$, such that for any constant C, there exists a constant N_0 that depends on C, K, α, such that, for any $N \geq N_0$ and $t \leq C\sqrt{N}$, we have

$$\sum_{l=0}^{N-1}\sum_{l'=0}^{N-1-l} |\nu^G(l,l') - \nu^T(l,l')| \leq \frac{\epsilon_c t}{N}.$$

Thus, by maximal coupling, there exists a coupling between $\{\mathcal{P}_t^G\}_{t\geq 0}$ and $\{\mathcal{P}_t^T\}_{t\geq 0}$ such that

$$\mathbb{P}(\tilde{\mathcal{P}}_{t+1}^G \neq \mathcal{P}_{t+1}^T | \mathcal{F}_t \text{ such that } \tilde{\mathcal{P}}_t^G = \mathcal{P}_t^T, \text{the next revealed node in } \mathcal{P}_t^T \text{ is a man node}, \mathcal{H}_t, \mathcal{R}) \leq \frac{\epsilon_c t}{N}.$$

The case when the next revealed node in $\mathcal{P}_t^T$ is a woman node is analogous and yields the same bound. The result hence follows.

□

We now establish the coupling up to time $\Theta\sqrt{N}$.

Lemma 21 (**Coupling between ALG-G and ALG-T up to time $\Theta(\sqrt{N})$**). *Recall constants a, β in the definition of $\mathcal{R}$ in Eq. (27). There exists a coupling between $\{\mathcal{P}_t^G\}_{t\geq 0}$ and $\{\mathcal{P}_t^T\}_{t\geq 0}$ such that, fix any $\epsilon_0 > 0$ (e.g., $\epsilon_0 = 0.5$), there exists a constant $C \equiv C(\epsilon_0, K, \alpha, a, \beta) \in (0, \infty)$ and $N_0 \equiv N_0(C, K, \alpha) \equiv N_0(\epsilon_0, K, \alpha, a, \beta) \in (0, \infty)$ such that, for any $N \geq N_0$ and $t \leq C\sqrt{N}$,*

$$\mathbb{P}(\tilde{\mathcal{P}}_t^G = \mathcal{P}_t^T | \mathcal{R}) \geq 1 - \epsilon_0.$$

Proof of Lemma 21. We have

$$\mathbb{P}(\tilde{\mathcal{P}}_t^G = \mathcal{P}_t^T | \mathcal{R}) \geq \mathbb{P}(\tilde{\mathcal{P}}_t^G = \mathcal{P}_t^T, \cap_{0\leq s\leq t}\mathcal{H}_s | \mathcal{R})$$
$$= 1 - \mathbb{P}(\cup_{0\leq s\leq t}\overline{\mathcal{H}_s} \text{ OR } \tilde{\mathcal{P}}_t^G \neq \mathcal{P}_t^T | \mathcal{R})$$
$$= 1 - \mathbb{P}(\cup_{0\leq s\leq t}\overline{\mathcal{H}_s} | \mathcal{R}) - \mathbb{P}(\tilde{\mathcal{P}}_t^G \neq \mathcal{P}_t^T, \cap_{0\leq s\leq t}\mathcal{H}_s | \mathcal{R}).$$

Now we upper bound each of the two probabilities on the RHS. By Lemma 14, for any $\epsilon_0 \in (0,1)$, there exists some large constant $\overline{W}$ depending on ϵ_0, a, β, such that $\mathbb{P}(\overline{\mathcal{H}_t} | \mathcal{R}) \leq (\frac{\epsilon_0}{2+\epsilon_0})^{t+1}$ for all $t \geq 0$. Fix this choice of $\overline{W}$ for the rest of the proof. Therefore

$$\mathbb{P}(\cup_{0\leq s\leq t}\overline{\mathcal{H}_s} | \mathcal{R}) \leq \sum_{s=0}^{t} P(\overline{\mathcal{H}_s} | \mathcal{R})$$
$$\leq \sum_{s=0}^{t} (\frac{\epsilon_0}{2+\epsilon_0})^{s+1}$$

$$\leq \frac{\frac{\epsilon_0}{2+\epsilon_0}}{1-\frac{\epsilon_0}{2+\epsilon_0}} = \frac{\epsilon_0}{2}.$$

The other term

$$\begin{aligned}
&\mathbb{P}(\tilde{\mathcal{P}}_t^G \neq \mathcal{P}_t^T, \cap_{0\leq s\leq t}\mathcal{H}_s|\mathcal{R}) \\
&= \mathbb{P}(\tilde{\mathcal{P}}_0^G \neq \mathcal{P}_0^T, \cap_{0\leq s\leq t}\mathcal{H}_s|\mathcal{R}) + \sum_{s=0}^{t-1}\mathbb{P}(\tilde{\mathcal{P}}_{s+1}^G \neq \mathcal{P}_{s+1}^T, \tilde{\mathcal{P}}_s^G = \mathcal{P}_s^T, \cap_{0\leq s\leq t}\mathcal{H}_s|\mathcal{R}) \\
&\leq \mathbb{P}(\tilde{\mathcal{P}}_0^G \neq \mathcal{P}_0^T|\mathcal{R}) + \sum_{s=0}^{t-1}\mathbb{P}(\tilde{\mathcal{P}}_{s+1}^G \neq \mathcal{P}_{s+1}^T, \tilde{\mathcal{P}}_s^G = \mathcal{P}_s^T, \mathcal{H}_s|\mathcal{R}) \\
&\text{(by Lemma 6, } \mathbb{P}(\tilde{\mathcal{P}}_0^G \neq \mathcal{P}_0^T|\mathcal{R}) = 0) \\
&= \sum_{s=0}^{t-1}\mathbb{P}(\tilde{\mathcal{P}}_{s+1}^G \neq \mathcal{P}_{s+1}^T|\tilde{\mathcal{P}}_s^G = \mathcal{P}_s^T, \mathcal{H}_s, \mathcal{R}) \cdot \mathbb{P}(\tilde{\mathcal{P}}_s^G = \mathcal{P}_s^T, \mathcal{H}_s|\mathcal{R}) \\
&\leq \sum_{s=0}^{t-1}\mathbb{P}(\tilde{\mathcal{P}}_{s+1}^G \neq \mathcal{P}_{s+1}^T|\tilde{\mathcal{P}}_s^G = \mathcal{P}_s^T, \mathcal{H}_s, \mathcal{R}).
\end{aligned}$$

By Lemma 6, there exists a coupling between $\{\mathcal{P}_t^G\}_{t\geq 0}$ and $\{\mathcal{P}_t^T\}_{t\geq 0}$, and a constant ϵ_c that depends on $K, \alpha, \overline{W}$, such that, for any constant C, there exists N_0 large enough that depends on C, K, α, such that for any $N \geq N_0$ and $0 \leq s \leq C\sqrt{N}$,

$$\mathbb{P}(\tilde{\mathcal{P}}_{s+1}^G \neq \mathcal{P}_{s+1}^T|\tilde{\mathcal{P}}_s^G = \mathcal{P}_s^T, \mathcal{H}_s, \mathcal{R}) \leq \frac{\epsilon_c s}{N}$$

and

$$\mathbb{P}(\tilde{\mathcal{P}}_0^G \neq \mathcal{P}_0^T|\mathcal{R}) = 0.$$

Choose $C = \sqrt{\frac{\epsilon_0}{\epsilon_c}}$. Therefore, C is a constant depending on $\epsilon_0, K, \alpha, \overline{W}$, and there exists a constant N_0 that depends on C, K, α, such that for any $N \geq N_0$ and $t \leq C\sqrt{N}$,

$$\sum_{s=0}^{t-1}\mathbb{P}(\tilde{\mathcal{P}}_{s+1}^G \neq \mathcal{P}_{s+1}^T|\tilde{\mathcal{P}}_s^G = \mathcal{P}_s^T, \mathcal{H}_s, \mathcal{R}) \leq \sum_{s=0}^{t-1}\frac{\epsilon_c s}{N} = \frac{\epsilon_c t(t-1)}{2N} \leq \frac{\epsilon_c C^2}{2} = \frac{\epsilon_0}{2}$$

The result hence follows.

□

E.3 Proof of Lemma 7–9

We first restate the coupling result conditional on $\mathcal{C}_{t_0}$.

Lemma 22 (Conditional coupling). *Recall constants a, β, ξ from the definition of $\mathcal{R}$ in Eq.* (27)*. There exists a coupling between $\{\mathcal{P}_t^G\}_{t\geq 0}$ and $\{\mathcal{P}_t^T\}_{t\geq 0}$ such that, fix any $\epsilon_0 \geq 0$ (e.g., $\epsilon_0 = 0.5$) and any $t_0 \geq 1$, there exists a constant $C \equiv C(\epsilon_0, K, \alpha, a, \beta, \xi) \in (0, \infty)$ and $N_0 \equiv N_0(C, K, \alpha) \in (0, \infty)$ such that, for any $N \geq N_0$ and $t_0 \leq t \leq C\sqrt{N}$,*

$$\mathbb{P}(\tilde{\mathcal{P}}_t^G = \mathcal{P}_t^T|\mathcal{C}_{t_0}, \mathcal{R}) \geq 1 - \epsilon_0.$$

Proof of Lemma 22. By Lemma 17, we know that $\mathbb{P}(\mathcal{C}_{t_0}|\mathcal{R}) \geq 1 - \xi > 0$ for any $t_0 \geq 1$. Also by Lemma 21, there exits a constant C that depends on $\epsilon_0, \xi, K, \alpha, a, \beta$, such that for any N larger than some N_0 that depends on C, K, α, we have $\mathbb{P}(\tilde{\mathcal{P}}_t^G \neq \mathcal{P}_t^T | \mathcal{R}) \leq \epsilon_0(1-\xi)$ for all $t \leq C\sqrt{N}$. Therefore, for this choice of C and N_0, for any $N \geq N_0$ and $t \leq C\sqrt{N}$,

$$\begin{aligned}
\mathbb{P}(\tilde{\mathcal{P}}_t^G = \mathcal{P}_t^T | \mathcal{C}_{t_0}, \mathcal{R}) &= 1 - \mathbb{P}(\tilde{\mathcal{P}}_t^G \neq \mathcal{P}_t^T | \mathcal{C}_{t_0}, \mathcal{R}) \\
&= 1 - \frac{\mathbb{P}(\tilde{\mathcal{P}}_t^G \neq \mathcal{P}_t^T, \mathcal{C}_{t_0}, \mathcal{R})}{\mathbb{P}(\mathcal{C}_{t_0}, \mathcal{R})} \\
&\geq 1 - \frac{\mathbb{P}(\tilde{\mathcal{P}}_t^G \neq \mathcal{P}_t^T, \mathcal{R})}{\mathbb{P}(\mathcal{R})\mathbb{P}(\mathcal{C}_{t_0} | \mathcal{R})} \\
&= 1 - \frac{\mathbb{P}(\tilde{\mathcal{P}}_t^G \neq \mathcal{P}_t^T | \mathcal{R})}{\mathbb{P}(\mathcal{C}_{t_0} | \mathcal{R})} \\
&\geq 1 - \frac{\epsilon_0(1-\xi)}{1-\xi} = 1 - \epsilon_0.
\end{aligned}$$

□

Now we lower bound the probability of deadlock in one step of ALG-G (Lemma 7).

Proof of Lemma 7. Suppose $\tilde{\mathcal{P}}_t^G = \mathcal{P}_t^T$ and $\mathcal{C}_t = 1$, i.e., ALG-G is not terminated at t and will reveal a next node s. Let $(d_i)_{i=1}^N, (d'_j)_{j=1}^{\alpha N}$ be the degree sequence used by ALG-G and ALG-T. Also suppose $\mathcal{J}_t = 1$, i.e., $d_A^{T,\text{unre}}(t) \geq \delta_0 t$ (recall that $d_A^{T,\text{unre}}(t)$ is the total number of unrevealed edges of s's ancestor (excluding the parent) nodes with the opposite type to s in $\mathcal{D}_t^T$. Given all the information in $\mathcal{P}_t^G$ and suppose the next revealed node is a man node, then the conditional probability of $\mathcal{A}_{t+1}$ is

$$\begin{aligned}
&\mathbb{P}(\mathcal{A}_{t+1} | \mathcal{F}_t \text{ such that } \tilde{\mathcal{P}}_t^G = \mathcal{P}_t^T, \mathcal{J}_t, \mathcal{C}_t, \mathcal{R}, \text{the next revealed node is a man node}) \\
&= \frac{\text{total number of unrevealed edges of nodes in } \mathcal{I}_A^G(t)}{\text{total number of remaining unrevealed men half edges at } t \text{ of ALG-G}},
\end{aligned}$$

where $\mathcal{I}_A^G(t)$ is the set of s's ancestor (excluding the parent) nodes with the opposite type to s in $\mathcal{D}_t^T$. Note that the information of $\mathcal{I}_A^G$ only depends on $\mathcal{P}_t^G$. Therefore, conditional on $\mathcal{P}_t^G = \mathcal{P}_t^T$, we have $\mathcal{I}_A^G(t) = \mathcal{I}_A^T(t)$, where $\mathcal{I}_A^T(t)$ is the analogy of $\mathcal{I}_A^G(t)$ defined on $\mathcal{P}_t^T$. The numerator above is hence equal to the total number of unrevealed edges of nodes in $\mathcal{I}_A^T(t)$, denoted by $d_A^{T,\text{unre}}(t)$. Also, observe that the denominator above is no larger than the total number of men half edges in the bipartite graph, which is $\sum_{j=1}^{\alpha N} d'_j$. Therefore, the above probability

$$\begin{aligned}
&\mathbb{P}(\mathcal{A}_{t+1} | \mathcal{F}_t \text{ such that } \tilde{\mathcal{P}}_t^G = \mathcal{P}_t^T, \mathcal{J}_t, \mathcal{C}_t, \mathcal{R}, \text{the next revealed node is a man node}) \\
&\geq \frac{d_A^{T,\text{unre}}(t)}{\sum_{j=1}^{\alpha N} d'_j} \\
&\text{(since } \mathcal{J}_t = 1\text{, see Lemma 19)} \\
&\geq \frac{\delta_0 t}{\sum_{j=1}^{\alpha N} d'_j}.
\end{aligned}$$

Note that $\sum_{j=1}^{\alpha N} d'_j = \alpha N \bar{k}_m$, where $\bar{k}_m$ is the empirical mean of the man node's degree in the bipartite graph. Conditional on $\mathcal{R}$, we know that $\bar{k}_m \geq \frac{K}{\alpha} - \left|\sqrt{\frac{4\alpha+K^2}{2\alpha^2}} - \frac{K}{\alpha}\right|$, hence the above can be rewritten as

$$\mathbb{P}(\mathcal{A}_{t+1} | \mathcal{F}_t \text{ such that } \tilde{\mathcal{P}}_t^G = \mathcal{P}_t^T, \mathcal{J}_t, \mathcal{C}_t, \mathcal{R}, \text{the next revealed node is a man node})$$

$$\geq \frac{\delta_0 t}{\alpha N \left(\frac{K}{\alpha} - \left| \sqrt{\frac{4\alpha + K^2}{2\alpha^2}} - \frac{K}{\alpha} \right| \right)}.$$

Therefore, there exists a constant $\delta_1 > 0$ that depends on δ_0, K, α, such that for any $t \geq 0$, we have

$$\mathbb{P}(\mathcal{A}_{t+1} | \mathcal{F}_t \text{ such that } \tilde{\mathcal{P}}_t^G = \mathcal{P}_t^T, \mathcal{J}_t, \mathcal{C}_t, \mathcal{R}, \text{the next revealed node is a man node}) \geq \frac{\delta_1 t}{N}.$$

Note that the same bound works for any degree sequence $(d_i)_{i=1}^N, (d'_j)_{j=1}^{\alpha N}$ that satisfies $\mathcal{R}$. The case when the next revealed node is a woman node is analogous and yields the same bound. The result hence follows.

□

We are now ready to prove Lemma 9.

Proof of Lemma 9. Fix any K, α such that $K^2 > 4\alpha$. By Lemma 13, there exist constants $a > 1, \beta > 1, \gamma \in (0,1), \kappa \in (0,1), \xi \in (0,1), \mu \in (0,1)$ that depend on (K, α), such that, for event

$$\mathcal{R} = \{\text{the degree sequence of } G_N(\alpha, K, p) \text{ is } (a, \beta, \gamma, \kappa, \xi, \mu)\text{-regular}\}, \tag{30}$$

we have, for all N larger than some constant N_R depending on K, α,

$$\mathbb{P}(\mathcal{R}) \geq \frac{1}{2}.$$

Fix these constants $a, \beta, \gamma, \kappa, \xi, \mu, N_R$, the event $\mathcal{R}$, and $N \geq N_R$ for the rest of the proof.

For any $t_0 \geq 1, \tau \geq 1$ such that $t_0 < \tau$, we have

$$\begin{aligned}
\mathbb{P}\left(\bigcup_{t \geq 0} \mathcal{A}_t \right) &\geq \mathbb{P}(\exists t \geq 0 : \tilde{\mathcal{P}}_t^G = \mathcal{P}_t^T, \mathcal{A}_{t+1}) \\
&\geq \mathbb{P}(\exists t_0 \leq t \leq \tau : \tilde{\mathcal{P}}_t^G = \mathcal{P}_t^T, \mathcal{A}_{t+1}) \\
&\geq \mathbb{P}(\mathcal{C}_{t_0}, \mathcal{R}, \exists t_0 \leq t \leq \tau : \tilde{\mathcal{P}}_t^G = \mathcal{P}_t^T, \mathcal{A}_{t+1}) \\
&= \mathbb{P}(\mathcal{R}) \cdot \mathbb{P}(\mathcal{C}_{t_0} | \mathcal{R}) \cdot \mathbb{P}(\exists t_0 \leq t \leq \tau : \tilde{\mathcal{P}}_t^G = \mathcal{P}_t^T, \mathcal{A}_{t+1} \big| \mathcal{C}_{t_0}, \mathcal{R}) \\
&\geq \frac{1}{2} \mathbb{P}(\mathcal{C}_{t_0} | \mathcal{R}) \cdot \mathbb{P}(\exists t_0 \leq t \leq \tau : \tilde{\mathcal{P}}_t^G = \mathcal{P}_t^T, \mathcal{A}_{t+1} \big| \mathcal{C}_{t_0}, \mathcal{R}).
\end{aligned}$$

By Lemma 17, $\mathbb{P}(\mathcal{C}_{t_0} | \mathcal{R}) \geq 1 - \xi > 0$ for any $t_0 \geq 1$. Thus we only need to lower-bound $\mathbb{P}(\exists t_0 \leq t \leq \tau : \tilde{\mathcal{P}}_t^G = \mathcal{P}_t^T, \mathcal{A}_{t+1} \big| \mathcal{C}_{t_0}, \mathcal{R})$. Observe that the events $\{\tilde{\mathcal{P}}_t^G = \mathcal{P}_t^T, \mathcal{A}_{t+1} | \mathcal{C}_{t_0}, \mathcal{R}\}$ for different t are mutually exclusive, since if $\mathcal{A}_{t+1} = 1$, then it must be that $\tilde{\mathcal{P}}_{t'}^G \neq \mathcal{P}_{t'}^T$ for all $t' > t$. Therefore

$$\begin{aligned}
&\mathbb{P}(\exists t_0 \leq t \leq \tau : \tilde{\mathcal{P}}_t^G = \mathcal{P}_t^T, \mathcal{A}_{t+1} \big| \mathcal{C}_{t_0}, \mathcal{R}) &(31) \\
&= \sum_{t=t_0}^{\tau} \mathbb{P}(\tilde{\mathcal{P}}_t^G = \mathcal{P}_t^T, \mathcal{A}_{t+1} \big| \mathcal{C}_{t_0}, \mathcal{R}) \\
&= \sum_{t=t_0}^{\tau} \mathbb{P}(\tilde{\mathcal{P}}_t^G = \mathcal{P}_t^T | \mathcal{C}_{t_0}, \mathcal{R}) \cdot \mathbb{P}(\mathcal{A}_{t+1} | \tilde{\mathcal{P}}_t^G = \mathcal{P}_t^T, \mathcal{C}_{t_0}, \mathcal{R}). &(32)
\end{aligned}$$

Recall that the argument so far holds for any $1 \leq t_0 < \tau$. By Lemma 22, we can choose $\tau = C\sqrt{N}$ where C is a constant that depends on K, α, a, β, ξ, such that for any N that is larger than some constant N_0 depending on C, K, α, and any $t_0 \geq 1$,

$$\mathbb{P}(\tilde{\mathcal{P}}_t^G = \mathcal{P}_t^T | \mathcal{C}_{t_0}, \mathcal{R}) \geq \frac{1}{2} \tag{33}$$

for all $t_0 \leq t \leq \tau$. Fix these constants C, τ, N_0, and $N \geq \max\{N_R, N_0\}$ for the rest of the proof.

It now only remains to lower-bound $\sum_{t=t_0}^{\tau} \mathbb{P}(\mathcal{A}_{t+1} | \tilde{\mathcal{P}}_t^G = \mathcal{P}_t^T, \mathcal{C}_{t_0}, \mathcal{R})$. We bound each term in this summation. Recall that (see Lemma 7, and we will specify the choice of δ_0 later)

$$\mathcal{J}_t := \{d_A^{T,\text{unre}}(t) \geq \delta_0 t\}.$$

Also, for any $t_0 \leq t \leq \tau$, $\mathcal{C}_t = \{L_t^T > 0\} \subseteq \mathcal{C}_{t_0}$, hence (for any choice of δ_0)

$$\begin{aligned}
&\mathbb{P}(\mathcal{A}_{t+1} | \tilde{\mathcal{P}}_t^G = \mathcal{P}_t^T, \mathcal{C}_{t_0}, \mathcal{R}) && (34)\\
&\geq \mathbb{P}(\mathcal{A}_{t+1} \cap \mathcal{J}_t \cap \mathcal{C}_t | \tilde{\mathcal{P}}_t^G = \mathcal{P}_t^T, \mathcal{C}_{t_0}, \mathcal{R})\\
&= \mathbb{P}(\mathcal{A}_{t+1} | \tilde{\mathcal{P}}_t^G = \mathcal{P}_t^T, \mathcal{J}_t, \mathcal{C}_t, \mathcal{C}_{t_0}, \mathcal{R}) \cdot \mathbb{P}(\mathcal{J}_t \cap \mathcal{C}_t | \tilde{\mathcal{P}}_t^G = \mathcal{P}_t^T, \mathcal{C}_{t_0}, \mathcal{R})\\
&= \mathbb{P}(\mathcal{A}_{t+1} | \tilde{\mathcal{P}}_t^G = \mathcal{P}_t^T, \mathcal{J}_t, \mathcal{C}_t, \mathcal{R}) \cdot \mathbb{P}(\mathcal{J}_t \cap \mathcal{C}_t | \tilde{\mathcal{P}}_t^G = \mathcal{P}_t^T, \mathcal{C}_{t_0}, \mathcal{R}). && (35)
\end{aligned}$$

By Lemma 7, for any choice of $\delta_0 > 0$ in $\mathcal{J}_t$, there exists a constant $\delta_1 > 0$ that depends on δ_0, K, α, such that, for any $t \geq 0$,

$$\mathbb{P}(\mathcal{A}_{t+1} | \mathcal{F}_t \text{ s.t. } \tilde{\mathcal{P}}_t^G = \mathcal{P}_t^T, \mathcal{J}_t, \mathcal{C}_t, \mathcal{R}) \geq \frac{\delta_1 t}{N},$$

in particular,

$$\mathbb{P}(\mathcal{A}_{t+1} | \tilde{\mathcal{P}}_t^G = \mathcal{P}_t^T, \mathcal{J}_t, \mathcal{C}_t, \mathcal{R}) \geq \frac{\delta_1 t}{N}. \tag{36}$$

On the other hand,

$$\begin{aligned}
\mathbb{P}(\mathcal{J}_t \cap \mathcal{C}_t | \tilde{\mathcal{P}}_t^G = \mathcal{P}_t^T, \mathcal{C}_{t_0}, \mathcal{R}) &= 1 - \mathbb{P}(\overline{\mathcal{J}_t} \cup \overline{\mathcal{C}_t} | \tilde{\mathcal{P}}_t^G = \mathcal{P}_t^T, \mathcal{C}_{t_0}, \mathcal{R})\\
&\geq 1 - \mathbb{P}(\overline{\mathcal{J}_t} | \tilde{\mathcal{P}}_t^G = \mathcal{P}_t^T, \mathcal{C}_{t_0}, \mathcal{R}) - \mathbb{P}(\overline{\mathcal{C}_t} | \tilde{\mathcal{P}}_t^G = \mathcal{P}_t^T, \mathcal{C}_{t_0}, \mathcal{R}). && (37)
\end{aligned}$$

We now upper-bound the probabilities $\mathbb{P}(\overline{\mathcal{J}_t} | \tilde{\mathcal{P}}_t^G = \mathcal{P}_t^T, \mathcal{C}_{t_0}, \mathcal{R})$ and $\mathbb{P}(\overline{\mathcal{C}_t} | \tilde{\mathcal{P}}_t^G = \mathcal{P}_t^T, \mathcal{C}_{t_0}, \mathcal{R})$. First consider

$$\begin{aligned}
\mathbb{P}(\overline{\mathcal{J}_t} | \tilde{\mathcal{P}}_t^G = \mathcal{P}_t^T, \mathcal{C}_{t_0}, \mathcal{R}) &= \frac{\mathbb{P}(\overline{\mathcal{J}_t}, \tilde{\mathcal{P}}_t^G = \mathcal{P}_t^T, \mathcal{C}_{t_0}, \mathcal{R})}{\mathbb{P}(\tilde{\mathcal{P}}_t^G = \mathcal{P}_t^T, \mathcal{C}_{t_0}, \mathcal{R})}\\
&\leq \frac{\mathbb{P}(\overline{\mathcal{J}_t}, \mathcal{C}_{t_0}, \mathcal{R})}{\mathbb{P}(\mathcal{C}_{t_0}, \mathcal{R}) \cdot \mathbb{P}(\tilde{\mathcal{P}}_t^G = \mathcal{P}_t^T | \mathcal{C}_{t_0}, \mathcal{R})}\\
&= \frac{\mathbb{P}(\overline{\mathcal{J}_t} | \mathcal{C}_{t_0}, \mathcal{R})}{\mathbb{P}(\tilde{\mathcal{P}}_t^G = \mathcal{P}_t^T | \mathcal{C}_{t_0}, \mathcal{R})}.
\end{aligned}$$

From Eq. (33), we already know that $\mathbb{P}(\tilde{\mathcal{P}}_t^G = \mathcal{P}_t^T | \mathcal{C}_{t_0}, \mathcal{R}) \geq \frac{1}{2}$ for any $1 \leq t_0 \leq t \leq \tau$ (where $t_0 < \tau$). Moreover, by Lemma 19, there exists a choice of δ_0 in $\mathcal{J}_t$, that depends on $K, \alpha, \kappa, \mu, \xi$, and a choice of t_0' that depends on $K, \alpha, \kappa, \mu, \xi$, such that $\mathbb{P}(\overline{\mathcal{J}_t} | \mathcal{C}_{t_0}, \mathcal{R}) \leq \frac{1}{6}$ for any $t_0' \leq t_0 \leq t \leq \tau$ (where $t_0 < \tau$). Recall that we have chosen $\tau = C\sqrt{N}$. Let N_c be a large constant that $C\sqrt{N_c} > t_0'$. We fix this choice of δ_0, t_0', N_c, and $N \geq \max\{N_R, N_0, N_c\}$ for the rest of the proof. As a result, for any $t_0' \le t_0 \le t \le \tau$,

$$\mathbb{P}(\overline{\mathcal{J}_t} | \tilde{\mathcal{P}}_t^G = \mathcal{P}_t^T, \mathcal{C}_{t_0}, \mathcal{R}) \le \frac{1}{3}. \tag{38}$$

We now upper-bound $\mathbb{P}(\overline{\mathcal{C}_t} | \mathcal{P}_t^G = \mathcal{P}_t^T, \mathcal{C}_{t_0}, \mathcal{R})$ for any $t_0' \le t_0 \le t \le \tau$. In particular,

$$\begin{aligned}
\mathbb{P}(\overline{\mathcal{C}_t} | \tilde{\mathcal{P}}_t^G = \mathcal{P}_t^T, \mathcal{C}_{t_0}, \mathcal{R}) &= \frac{\mathbb{P}(\overline{\mathcal{C}_t}, \tilde{\mathcal{P}}_t^G = \mathcal{P}_t^T, \mathcal{C}_{t_0}, \mathcal{R})}{\mathbb{P}(\tilde{\mathcal{P}}_t^G = \mathcal{P}_t^T, \mathcal{C}_{t_0}, \mathcal{R})} \\
&\le \frac{\mathbb{P}(\overline{\mathcal{C}_t}, \mathcal{C}_{t_0}, \mathcal{R})}{\mathbb{P}(\mathcal{C}_{t_0}, \mathcal{R}) \cdot \mathbb{P}(\tilde{\mathcal{P}}_t^G = \mathcal{P}_t^T | \mathcal{C}_{t_0}, \mathcal{R})} \\
&\le \frac{\mathbb{P}(\overline{\mathcal{C}_\infty}, \mathcal{C}_{t_0}, \mathcal{R})}{\mathbb{P}(\mathcal{C}_{t_0}, \mathcal{R}) \cdot \mathbb{P}(\tilde{\mathcal{P}}_t^G = \mathcal{P}_t^T | \mathcal{C}_{t_0}, \mathcal{R})} \\
&= \frac{\mathbb{P}(\overline{\mathcal{C}_\infty} | \mathcal{C}_{t_0}, \mathcal{R})}{\mathbb{P}(\tilde{\mathcal{P}}_t^G = \mathcal{P}_t^T | \mathcal{C}_{t_0}, \mathcal{R})}.
\end{aligned}$$

Again, we already know from Eq. (33) that $\mathbb{P}(\tilde{\mathcal{P}}_t^G = \mathcal{P}_t^T | \mathcal{C}_{t_0}, \mathcal{R}) \ge \frac{1}{2}$ for any $t_0' \le t_0 \le t \le \tau$. Moreover, by Lemma 17, for any $t_0 \ge 1$, $\mathbb{P}(\overline{\mathcal{C}_\infty} | \mathcal{C}_{t_0}, \mathcal{R}) \le \frac{\xi}{(1-\xi)(1-\mu)t_0}$. Thus there exists a constant t_0'' that depends on ξ, μ such that for any $t_0 \ge t_0''$,

$$\begin{aligned}
\mathbb{P}(\overline{\mathcal{C}_\infty} | \mathcal{C}_{t_0}, \mathcal{R}) &\le \frac{\xi}{(1-\xi)(1-\mu)t_0} \\
&\le \frac{\xi}{(1-\xi)(1-\mu)t_0''} \le \frac{1}{6}.
\end{aligned}$$

Fix $t_0 = \max\{t_0', t_0''\}$ that depends on $K, \alpha, \kappa, \mu, \xi$ for the rest of the proof. Also let N_τ be a large constant such that $C\sqrt{N_\tau} > t_0$. Therefore, for any $t_0 \le t \le \tau$ and $N > \max\{N_R, N_0, N_c, N_\tau\}$ (so that $\tau = C\sqrt{N} > t_0$),

$$\mathbb{P}(\overline{\mathcal{C}_t} | \tilde{\mathcal{P}}_t^G = \mathcal{P}_t^T, \mathcal{C}_{t_0}, \mathcal{R}) \le \frac{1}{3}. \tag{39}$$

Combine Eq. (37)–(39), we have (with the chosen parameters)

$$\mathbb{P}(\mathcal{J}_t \cap \mathcal{C}_t | \tilde{\mathcal{P}}_t^G = \mathcal{P}_t^T, \mathcal{C}_{t_0}, \mathcal{R}) \ge 1 - \mathbb{P}(\overline{\mathcal{J}_t} | \tilde{\mathcal{P}}_t^G = \mathcal{P}_t^T, \mathcal{C}_{t_0}, \mathcal{R}) - \mathbb{P}(\overline{\mathcal{C}_t} | \tilde{\mathcal{P}}_t^G = \mathcal{P}_t^T, \mathcal{C}_{t_0}, \mathcal{R}) \ge \frac{1}{3}. \tag{40}$$

Eqs. (31), (34), (36) and (40) together yield

$$\begin{aligned}
\mathbb{P}(\exists t_0 \le t \le \tau : \tilde{\mathcal{P}}_t^G = \mathcal{P}_t^T, \mathcal{A}_{t+1} \big| \mathcal{C}_{t_0}, \mathcal{R}) &\ge \frac{1}{6} \sum_{t=t_0}^{\tau} \frac{\delta_1 t}{N} \\
&\ge \frac{\delta_1(\tau^2 - t_0^2)}{12N}.
\end{aligned}$$

Recall that $\delta_1 > 0$ is a constant that depends on δ_0, K, α, t_0 is a constant that depends on $K, \alpha, \kappa, \mu, \xi$, and $\tau = C\sqrt{N}$ where C is a constant that depends on K, α, ξ. Therefore, there must exist a constant N_1 depending on $K, \alpha, \kappa, \mu, \xi$, such that for any $N \ge \max\{N_R, N_0, N_c, N_\tau, N_1\}$, the above RHS is large than $\frac{\delta_1 C^2}{20} > 0$.

We thus complete the proof.

□

E.4 Proof of Theorem 2

Proof of Theorem 2. We will show that $\mathbb{P}\big(\#\text{Deadlocked women} > \varrho^2 N\big) \geq \varrho/(1+\varrho)$. To establish this, we will use Markov's inequality on the number of women who are *not* deadlocked. By Lemma 9, we know that

$$\mathbb{E}[\#\text{Deadlock-free women}] \leq (1-\varrho)N\,. \tag{41}$$

By Markov's inequality, we infer that

$$\mathbb{P}\big(\#\text{Deadlock-free women} > (1-\varrho^2)N\big) \leq \frac{1-\varrho}{1-\varrho^2} = \frac{1}{1+\varrho}\,. \tag{42}$$

Hence,

$$\mathbb{P}\big(\#\text{Deadlocked women} > \varrho^2 N\big) = \mathbb{P}\big(\#\text{Deadlock-free women} \leq (1-\varrho^2)N\big) \geq \frac{\varrho}{1+\varrho}\,. \tag{43}$$

Choosing $\epsilon = \min(\varrho^2, \frac{\varrho}{1+\varrho})$ we get the claim in the theorem.

To see that $(\alpha, K, 0) \in \Theta$ if and only if $\frac{K^2}{4\alpha} > 1$, notice that the system of equations in Definition 4 reduces to

$$w_t = 1 - a_t = \frac{\frac{K}{\alpha}(1-\nu_{t-1}) - 1 + e^{-\frac{K}{\alpha}(1-\nu_{t-1})}}{\frac{K}{\alpha}(1-\nu_{t-1})}, \quad u_t = 1-\nu_t = \frac{K(1-a_t) - 1 + e^{-K(1-a_t)}}{K(1-a_t)}$$

if we restrict $p = 0$. From this we can equivalently write

$$w_t = f_1(f_2(w_{t-1}))$$

where

$$f_1(x) = \frac{\frac{K}{\alpha}x - 1 + e^{-\frac{K}{\alpha}x}}{\frac{K}{\alpha}x}, \quad f_2(x) = \frac{Kx - 1 + e^{-Kx}}{Kx}.$$

By straightforward algebra, one can verify that $f_1(f_2(\cdot))$ is concave on $[0,1]$ with

$$\left.\frac{df_1(f_2(w))}{dw}\right|_{w=0} = \frac{K^2}{4\alpha}, \quad \left.\frac{df_1(f_2(w))}{dw}\right|_{w=1} < 1.$$

Since $w_0 = 1$, we must have that $\lim_{t\to\infty} w_t$ is the largest root of $w = f_1(f_2(w))$ on $[0,1]$, and that $\lim_{t\to\infty} w_t > 0$ if and only if $\left.\frac{df_1(f_2(w))}{dw}\right|_{w=0} > 1$, or equivalently, $\frac{K^2}{4\alpha} > 1$. □

F A Heuristic Argument for Conjecture 1

A heuristic argument of the branching factor. We first explain the branching factor in the special case of $p = 0$ and extend the argument (heuristically) for the general case of $p > 0$.

In standard branching process theory (e.g. [10][24]), it is well known that the branching factor of a branching process, given the offspring distribution being η, can be derived as follows. Let u^* be the probability that this branching process does not extinct. Then, a key observation is that the tree rooted at the root node extincts if and only if all subtrees rooted at one of the root node's offsprings extinct. Since each subtree is an i.i.d. sample of the branching process with offspring distribution η, we have

$$u^* = F(u^*) \triangleq 1 - \mathbb{E}_{d\sim\eta}[(1-u^*)^d]. \tag{44}$$

One can observe that

$$\left.\frac{dF(u)}{du}\right|_{u=0} = \mathbb{E}_{d\sim\eta}[d].$$

Therefore, even if the offspring distribution η of the branching process is hard to express, we can still hope to obtain the effective branching factor $\mathbb{E}_{d\sim\eta}[d]$ of the branching process by deriving the function F and evaluating its derivative at zero.

For example, consider the GW tree in Definition 5. Let u^* denote the (edge-perspective) probability that a woman node in the GW tree waits forever, and w^* denote the (edge-perspective) probability that a man node in the GW tree waits forever. A key observation here is that a woman node in the GW tree waits forever if and only if one of the better-ranked offspring man nodes (better than her parent in her preference list), say, j, waits forever, and all the even better men (better than j in her preference list) neither waits forever nor inspect with her and succeed. Similarly, a man node in the GW tree waits forever if and only if one of his better-ranked offspring woman nodes, say, i, waits forever, and all the even better women neither wait forever nor inspect with him and succeed. Moreover, a man only wishes to inspect the parent woman (denote this probability by a^*) if all his better-ranked offspring woman nodes neither wait forever nor inspect with him and succeed; and a woman only wishes to inspect with the parent man if all her better-ranked offspring man nodes neither wait forever nor inspect with her and succeed. Denote y^* to be the probability that a woman wishes to inspect with the parent man and that the inspection will be successful if conducted. Note that the terms y^*, a^*, u^*, w^* defined above coincide with $y_\infty, a_\infty, u_\infty, w_\infty$ defined in Definition 4. We formalize this connection in Proposition 3 later in this section.

From the above description, we can derive the mapping function between u and itself, as in Eq. (44). Note that the number of better-edge degree of a woman node is Poisson(Kz), where $z \sim$ Uniform$(0,1)$; and the number of better-edge degree of a man node is Poisson$(K\tilde{z}/\alpha)$, where $\tilde{z} \sim$ Uniform$(0,1)$. Therefore,

$$\begin{aligned}
y^* &= f_1(w^*, a^*) \triangleq \mathbb{E}_{l\sim\text{Poisson}(Kz)}\left[p(1-w^*-a^*p)^l\right], \\
a^* &= \hat{f}_1(u^*, y^*) \triangleq \mathbb{E}_{l\sim\text{Poisson}(K\tilde{z}/\alpha)}\left[(1-u^*-y^*)^l\right], \\
u^* &= f_2(w^*, a^*) \triangleq \mathbb{E}_{\sim\text{Poisson}(Kz)}\left[\sum_{i=0}^{l-1} w^*(1-w^*-a^*p)^i\right], \\
w^* &= \hat{f}_2(u^*, y^*) \triangleq \mathbb{E}_{\sim\text{Poisson}(K\tilde{z}/\alpha)}\left[\sum_{i=0}^{l-1} u^*(1-u^*-y^*)^i\right].
\end{aligned}$$

Equivalently, we write

$$u^* = G_1(u^*, y^*) \triangleq f_2(\hat{f}_2(u^*, y^*), \hat{f}_1(u^*, y^*)), \quad y^* = G_2(u^*, y^*) \triangleq f_1(\hat{f}_2(u^*, y^*), \hat{f}_1(u^*, y^*)). \tag{45}$$

We first examine Eq. (45) when $p = 0$. In this case, Eq. (45) reduces to

$$u^* = f(w^*) \triangleq \mathbb{E}_{\sim\text{Poisson}(Kz)}\left[\sum_{i=0}^{l-1} w^*(1-w^*)^i\right], \quad w^* = \hat{f}(u^*) \triangleq \mathbb{E}_{\sim\text{Poisson}(K\tilde{z}/\alpha)}\left[\sum_{i=0}^{l-1} u^*(1-u^*)^i\right],$$

or equivalently,

$$u^* = G(u^*) \triangleq f(\hat{f}(u^*)). \tag{46}$$

Thus Eq. (46) resembles Eq. (44) in the standard branching process case. Moreover, one can verify easily that $G'(0) = \frac{K^2}{4\alpha}$, which, as Theorem 2 states, identifies the deadlock regime if and only if $G'(0) > 1$. In fact, another direct way of verifying that $\frac{K^2}{4\alpha}$ is indeed the branching factor is from the offspring distribution, which is the product of the woman node's better-edge degree $\frac{K}{2}$ and the man node's better-edge degree $\frac{K}{2\alpha}$.

Now consider Eq. (47) when $p > 0$. We are interested in the partial derivative $\frac{\partial G_1(u,y)}{\partial u}$ evaluated at $u = 0$ and the corresponding value of $y = \tilde{y}$. One can solve the value of $\tilde{y}$ by finding the (nonnegative) root for $\tilde{y} = G_2(0, \tilde{y})$ (see Definition 47). This gives

$$\begin{aligned}\left.\frac{\partial G_1(u,y)}{\partial u}\right|_{u=0,y=\tilde{y}} &= \left.\frac{df_2(w^*,a^*)}{dw^*}\right|_{w=0,a=\tilde{a}} \cdot \left.\frac{d\hat{f}_2(u^*,y^*)}{du^*}\right|_{u=0,y=\tilde{y}} + \left.\frac{df_2(w^*,a^*)}{da^*}\right|_{w=0,a=\tilde{a}} \cdot \left.\frac{d\hat{f}_1(u^*,y^*)}{du^*}\right|_{u=0,y=\tilde{y}} \\ &= \left(\frac{1}{\tilde{y}} - \frac{\tilde{a}}{\tilde{y}}\right) \cdot \left(\frac{1}{\tilde{a}p} - \frac{\tilde{y}}{\tilde{a}p^2}\right),\end{aligned} \tag{47}$$

where $\tilde{y}$ and $\tilde{a}$ are the unique nonzero solution (we show in Proposition 3 below that there exists a unique nonzero solution) to

$$Kya = 1 - e^{-Kap}, \quad \frac{K}{\alpha}ay = 1 - e^{-\frac{K}{\alpha}y}.$$

This is the expression given by Conjecture 1.

The connection between the branching factor and the density evolution equations. In the above we heuristically conjecture a branching factor for the associated tree process of $G_N(\alpha, K, p)$. We now formally establish the connection between this conjectured branching factor with the density evolution equations in Definition 4.

Proposition 3 (**Connection between density evolution and branching factor in Conjecture 1**)**.** *The system of equations in Definition 4 can be equivalently written as*

$$u_t = G_1(u_{t-1}, y_{t-1}), \quad y_t = G_2(u_{t-1}, y_{t-1})$$

where $G_1(\cdot,\cdot)$ and $G_2(\cdot,\cdot)$ are as defined in Eq. (47). Moreover, $y = G_2(0, y)$ has a unique nonzero root, denoted by $\tilde{y}$, and

$$\left.\frac{\partial G_1(u,y)}{\partial u}\right|_{u=0,y=\tilde{y}}$$

is equal to the branching factor in Conjecture 1.

Proof of Proposition 3. The first part of the proposition can be verified using straightforward algebra. Now consider the equation $y = G_2(0, y)$, which gives

$$Kya = 1 - e^{-Kap}, \quad \frac{K}{\alpha}ay = 1 - e^{-\frac{K}{\alpha}y}.$$

Suppose $y > 0$, then this gives

$$\alpha - \alpha e^{-\frac{Ky}{\alpha}} = 1 - e^{-p\frac{\alpha - \alpha e^{-\frac{Ky}{\alpha}}}{y}}. \tag{48}$$

Note that the LHS increases in $y > 0$. To examine the RHS, let

$$g(y) \triangleq \frac{1 - e^{-\frac{Ky}{\alpha}}}{y}.$$

Take its derivative, we get

$$g'(y) = \frac{e^{-\frac{Ky}{\alpha}}\left(1 + \frac{Ky}{\alpha} - e^{\frac{Ky}{\alpha}}\right)}{y^2}.$$

Note that $e^{\frac{Ky}{\alpha}} \geq 1 + \frac{Ky}{\alpha}$, hence $g'(y) \leq 0$ for all $y > 0$, which implies that the RHS of Eq. (48) decreases in $y > 0$. Since

$$\lim_{y \to 0^+} \alpha - \alpha e^{-\frac{Ky}{\alpha}} = 0, \quad \lim_{y \to \infty} \alpha - \alpha e^{-\frac{Ky}{\alpha}} = \alpha;$$

$$\lim_{y \to 0^+} 1 - e^{-p\frac{\alpha - \alpha e^{-\frac{Ky}{\alpha}}}{y}} = 1 - e^{-pK}, \quad \lim_{y \to \infty} 1 - e^{-p\frac{\alpha - \alpha e^{-\frac{Ky}{\alpha}}}{y}} = 0,$$

we have that Eq. (48) has a unique nonzero root. Therefore we have the desired result stated in Proposition 3. □

G A Heuristic Derivation of the Giant Component Regime Using MP

As a simple illustrative example, we now heuristically derive the classical result that for any $c > 1$, with high probability (w.h.p.) there is a "giant" connected component involving $\Omega(n)$ nodes in an Erdos-Renyi (ER) graph with edge probability c/n, whereas for any $c < 1$, whp there is no giant component: As $n \to \infty$, for any node i, the (random) local neighborhood of i (up to any fixed depth) converges in distribution to a Galton-Watson (GW) process with offspring distribution Poisson(c), i.e., a tree rooted at i where i has Poisson(c) children, each of these children, independently, has Poisson(c) children, and so on. Hence, one would expect there to be a giant component in the ER graph if and only if, with positive probability the GW process does not suffer extinction (see [15] for a formalization of this intuition). Whether the GW process suffers extinction can be captured via the following simple message passing algorithm with binary messages in $\{0, 1\}$: initialize $m_{j \to k}^{(0)} = 1$ for all edges (j, k). Thereafter for all $t > 0$, update the messages as follows: $m_{j \to k}^{(t)} = 1$ if there exists $l \in \mathcal{N}(j) \backslash k$ such that $m_{l \to j}^{(t-1)} = 1$, else $m_{j \to k}^{(t)} = 0$. (Note that $m_{j \to k}^{(t)}$ does not depend on $m_{k \to j}^{(t-1)}$, i.e., the self-exclusion property holds.) Observe that if k is the parent of j then $m_{j \to k}^{(t)} = 1$ if and only if the subtree rooted at j has depth at least t; we are interested in $\lim_{t \to \infty} m_{j \to k}^{(t)}$ which is 1 if and only the subtree has infinite depth. Fixing a node j (and not revealing its subtree), the probability $\pi^{(t)} = \Pr(m_{j \to k}^{(t)} = 1)$ can be iteratively computed to be $\pi^{(0)} = 1$, and $\pi^{(t)} =$ probability that j has no children who sent a 1 to it at $t - 1$. Since each child independently sent a 1 with probability $\pi^{(t-1)}$), the number of children who sent it a 1 is Poisson($p\pi^{(t-1)}$), and we obtain $\pi^{(t)} = 1 - \exp(-p\pi^{(t-1)})$. It easy to verify that $\pi^* = \lim_{t \to \infty} \pi^{(t)}$ exists and is strictly positive if and only if $c > 1$. Thus, if $c > 1$, the GW process avoids extinction with positive probability π^*, and correspondingly a giant component of size nearly $\pi^* n$ occurs in the ER graph. On the other hand if $c < 1$, we have $\pi^{(t)} < c^t \xrightarrow{t \to \infty} 0$, i.e., the GW process suffers a quick extinction, and w.h.p. there is no giant component in the ER graph.

H Model Extensions

In this appendix, we discuss two extensions of the base model presented in Section 2. First, we consider vertical differentiation between agents, i.e., instead of assuming all agents are ex ante homogeneous, we assume a certain fraction of them are more preferred than the rest. In the second extension, we model agents on one side of the market as being willing to conduct parallel inspections with their current top two potential partners, thus expanding the set of inspections which can be conducted. We show that our analysis framework can still be applied to these extensions with proper modifications, and that we can still characterize the size of information deadlock in these markets. Moreover, we show that introducing vertical differentiation and allowing parallel inspections help reduce deadlock.

H.1 Vertical Differentiation Between Agents

Assume there are two types of women, popular (P) and ordinary (O). The random consideration graph is generated as in the base case, as well as the women's preference rankings. The men's preference rankings are generated as follows: any P-type woman in a man's consideration set is ranked higher than any O-type woman in that consideration set, while within each type the rankings are independent uniformly random permutations. We denote by $G_N(\alpha, \beta, K, p)$ a random market with βN P-type women, $(1-\beta)N$ O-type women, and αN men, where each woman-man pair considers each other independently with probability $\frac{K}{\alpha N}$, and any inspection succeeds independently with probability p. Given such a randomly generated market, following the PS process defined in Figure 4, we can similarly count the number of P-type and O-type women who have not cleared at time t, denoted by $\Lambda^t_{\text{P},N}(\alpha, \beta, K, p)$ and $\Lambda^t_{\text{O},N}(\alpha, \beta, K, p)$, respectively. Let $\Lambda^t_{B,N}(\alpha, \beta, K, p)$ be their sum. Also, we can probabilistically describe the residual market for $N \to \infty$ by the MP algorithm in Figure 6 and associated density evolution (DE) equations. Note that the DE equations come with an amendment that there are two densities for each message, one for each type of woman (characterized by a superscript). For example, the densities for accept message are denoted by a^P_t and a^O_t, where a^P_t is for when the recipient woman is of P-type, and a^O_t is for when the recipient woman is of O-type. We give the corresponding DE equations in Definition 9 in Appendix I in the e-companion. Similar to $\lambda(\alpha, K, p) := \lim_{t\to\infty} K w_t q_t$ in Definition 4 capturing the fraction of women in deadlock in large random markets in the base model, we define here $\lambda^P(\alpha, \beta, K, p) := \lim_{t\to\infty} K w^P_t q^P_t$ and $\lambda^O(\alpha, \beta, K, p) := \lim_{t\to\infty} K w^O_t q^O_t$ to capture the fraction of P-type and O-type women in deadlock, respectively, in large random markets where β fraction of women are of P-type. Moreover, $\lambda^B(\alpha, \beta, K, p) := \beta\lambda^P(\alpha, \beta, K, p) + (1-\beta)\lambda^O(\alpha, \beta, K, p)$ is the fraction of both types of women in deadlock in such a market. This result, which is analogous to Theorem 1 in the base model, is formalized in Theorem 3 in Appendix I in the e-companion.

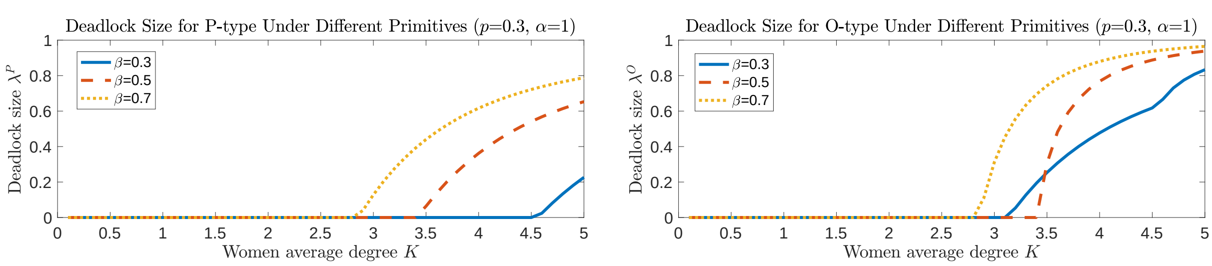

Figure 17: Deadlock size λ^P and λ^O for P-type and O-type women, respectively, as a function of the women' average degree K. We fix the inspection success probability $p = 0.3$ and the men-to-women ratio $\alpha = 1$. In the left figure we plot the deadlock size for P-type. In the right figure we plot the deadlock size for O-type. We compare values for different H-type proportion β.

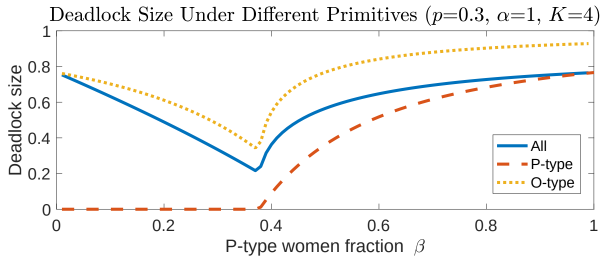

Figure 18: Deadlock size λ^P, λ^O and λ^B for P-type, O-type, and All women, respectively, as a function of the P-type women fraction β. We fix the inspection success probability $p = 0.3$, the men-to-women ratio $\alpha = 1$ and the women's average degree $K = 4$. The yellow curve plots the deadlock size for O-type. The blue curve plots the deadlock size for all type. The red curve plots the deadlock size for P-type.

The outcomes of P-type women are not affected by the presence of O-type women. The sub-market consisting only P-type women and men is equivalent to the base model with parameters $(K, \alpha/\beta, p)$. Therefore the size of deadlock for P-type women increases with β, since in the base model the size of deadlock decreases with α. For O-type women, their outcomes are affected by both O-type and P-type women. There are two competing forces that drive O-type women into/out of deadlocks. If the fraction of P-type women is larger, since they are systematically more popular than O-type women, the effective number of match opportunities for O-type women is smaller, causing more O-women to be stuck in a deadlock. On the other hand, since the number of O-type women decreases, the level of competition from within the group (O-type women) decreases, alleviating deadlock. Figure 17 shows that the size of deadlock for both types increases with women's average degree K, and P-type women experience the same outcome as in markets with primitives $(K, \alpha/\beta, p)$. The fraction of O-type women in deadlock, however, is not monotone in β. Figure 18 shows the size of deadlock as a function of the P-type fraction β. Note that the size of O-type deadlock is first decreasing then increasing in β. For small values of β, the outcomes of O-type women are mainly driven by the in-group competition, hence deadlock decreases with β since the number of O-type women decreases. For large values of β, the competition effect from P-type women increases and dominates the effect of other O-type women. Therefore the size of deadlock increases for O-type women as it increases for P-type women. Holding (K, α, p) fixed, the total size of deadlock for women is minimized for β lying at the threshold β^* between the deadlock versus deadlock-free regime for P-type women.

H.2 Parallel Inspections

Assume that while women are only willing to inspect with their top one most preferred men, men are willing to inspect with their top two most preferred women. Therefore the inspections conducted now include any woman-man pair (i, j) such that woman i has no preferred man in her consideration set and man j has at most one preferred woman in his consideration set. We call this *top-2 mutual inspections*. However, since every agent only has match capacity one, not all successful inspections are guaranteed to match. For example, a man j may conduct a top-2 mutual inspection with his second most preferred woman i' and succeed, hence (i', j) are temporarily matched. Later, the same man j may perform another top-2 mutual inspection with his top most preferred woman i because she now sees him as the most desirable in the remaining consideration set. If this later inspection succeeds, j will break up with i' to form a new match with i. What's more, the formation and breaking-up of temporary matches may propagate in the network. In the previous example, woman i', after her temporary match with man j breaks up, can then go to her next favorite man j' for a mutual inspection, which, if it succeeds, may cause j' to break-up with his previously formed match, and so on. The Partner Search process for this setting is described in Figure 19.

Partner Search (PS) following top-2 mutual inspections

Initialize the market: agents $\mathcal{I}_0 \leftarrow \mathcal{I}$; opportunities $\mathcal{J}_0 \leftarrow \mathcal{J}$; potential pairs $\mathcal{E}_0 \leftarrow \mathcal{E}$; temporary matched pairs $\tilde{\mathcal{E}}_0 \leftarrow \emptyset$; time $t \leftarrow 1$.
while $|\mathcal{I}_{t-1}| > 0$ **do**
  Initialization: $\mathcal{I}_t \leftarrow \mathcal{I}_{t-1}$; $\mathcal{J}_t \leftarrow \mathcal{J}_{t-1}$; $\mathcal{E}_t \leftarrow \mathcal{E}_{t-1}$; $\tilde{\mathcal{E}}_t \leftarrow \tilde{\mathcal{E}}_{t-1}$
  while $\exists$ a top-2 mutual inspection $(i, j) \in \mathcal{E}_{t-1} \setminus \{(i', j') : \exists j'' \text{ s.t. } (i', j'') \in \tilde{\mathcal{E}}_{t-1}\}$ **do**
    Perform inspection (i, j).
    if inspection (i, j) succeeds and $\forall i' \succ_j i$, $(i', j) \notin \mathcal{E}_t$ **then**
      The pair matches and leaves the market: $\mathcal{I}_t \leftarrow \mathcal{I}_t \backslash \{i\}, \mathcal{J}_t \leftarrow \mathcal{J}_t \backslash \{j\}, \mathcal{E}_t \leftarrow \mathcal{E}_t \backslash \{(i', j') : i' = i \text{ OR } j' = j\}$.
    else if inspection (i, j) succeeds and $\exists i' \succ_j i$ such that $(i', j) \in \mathcal{E}_t$ **then**
      The pair temporarily matches: $\tilde{\mathcal{E}}_t \leftarrow \tilde{\mathcal{E}}_t \cup \{(i, j)\}$, $\mathcal{E}_t \leftarrow \mathcal{E}_t \setminus \{(i', j) : i' \prec_j i\}$.
      Any inferior temporary matches get unmatched: $\tilde{\mathcal{E}}_t \leftarrow \tilde{\mathcal{E}}_t \setminus \{(i', j) : i' \prec_j i\}$
    else ▷ The inspection fails
      The pair is removed from the set of edges, but the agents remain: $\mathcal{E}_t \leftarrow \mathcal{E}_t \setminus \{(i, j)\}$.
    end if
  end while
  Eliminate any woman i and man j with empty consideration sets: $\mathcal{I}_t \leftarrow \mathcal{I}_t \setminus \{i\}, \mathcal{J}_t \leftarrow \mathcal{J}_t \setminus \{j\}$.
  Advance time: $t \leftarrow t + 1$.
end while

Figure 19: The Partner Search process under top-2 mutual inspections. Preference/fitness rankings remain unchanged throughout.

Note that although matches can be formed and then broken-up, the same match cannot be reformed again after it's broken-up, and the PS process will eventually converge (we show this in

Message Passing (MP) following top-2 mutual inspections

Input: $(\mathcal{I}, \mathcal{J}, \mathcal{E}, (\succ_i, \forall i \in \mathcal{I}), (\succ_j, \forall j \in \mathcal{J}))$ with latent inspection results $\epsilon_{ij} \sim \text{Bernoulli}(p), \forall (i, j) \in \mathcal{E}$.
Initialize the messages: $m^{(0)}_{i \to j} = \text{U}, \hat{m}^{(0)}_{j \to i} = \text{W}$ for all $(i, j) \in \mathcal{E}$.
$t \leftarrow 1$
do
  Update man-to-woman messages: For all $(i, j) \in \mathcal{E}$, let

$$\hat{m}^{(t)}_{j \to i} = \begin{cases} \text{A if for all } i' \succ_j i, m^{(t-1)}_{i' \to j} = \text{N except for at most one } i'' \succ_j i \text{ such that } m^{(t-1)}_{i'' \to j} = \text{U}; \\ \text{R if for some } i' \succ_j i, m^{(t-1)}_{i' \to j} = \text{Y, and } m^{(t-1)}_{i'' \to j} = \text{ N for all } i'' \succ_j i' \text{ except} \\ \qquad \text{for at most one } i''' \succ_j i' \text{ such that } m^{(t-1)}_{i''' \to j} = \text{U}; \\ \text{W otherwise.} \end{cases} \tag{49}$$

  Inspection ready: For all $(i, j) \in \mathcal{E}$, let

$$I^{(t)}_{ij} = \begin{cases} 1 \text{ if for all } j' \succ_i j, \text{ either } \hat{m}^{(t)}_{j' \to i} = \text{A and } \epsilon_{ij'} = 0; \text{ or } \hat{m}^{(t)}_{j' \to i} = \text{R}. \\ 0 \text{ otherwise.} \end{cases} \tag{50}$$

  Update woman-to-man messages: For all $(i, j) \in \mathcal{E}$, let

$$m^{(t)}_{i \to j} = \begin{cases} \text{Y if } I^{(t)}_{ij} = 1 \text{ and } \epsilon_{ij} = 1; \\ \text{N if } I^{(t)}_{ij} = 1 \text{ and } \epsilon_{ij} = 0, \text{ or if for some } j' \succ_i j, \\ \quad \hat{m}^{(t)}_{j' \to i} = \text{A and } I^{(t)}_{ij'} = 1 \text{ and } \epsilon_{ij'} = 1; \\ \text{U otherwise.} \end{cases} \tag{51}$$

  $t \leftarrow t + 1$
while some message changed

Figure 20: The Message Passing algorithm following top-2 mutual inspections.

the next paragraph). Define $\tilde{\Lambda}_N^t := \left|\mathcal{I}_t \setminus \left\{i' : \exists j \text{ s.t. } (i', j) \in \tilde{\mathcal{E}}_t\right\}\right|$, which is the number of women waiting at time[17] t. Intuitively, in the final market outcome, the size of deadlock should be smaller than that in the base model since a man previously stuck waiting for his favorite woman can now go to his next favorite. We show that this is indeed the case, and we characterize the market outcome using a modified Message Passing algorithm in Figure 20.

Similar to the base model, the modified MP algorithm tracks an accelerated version of the modified PS process, described in Figure 22 in Appendix J in the e-companion. With the modified PS, MP and APS, Lemma 2–4 follow as in the base case. Note that similar to Lemma 3 and its proof (in Appendix B in the e-companion), we can also show that the outcome of the modified MP algorithm does not depend on the sequence of message updates, by defining a partial ordering over message vectors and utilizing Tarski's fixed point theorem. Observe that there is a natural ordering of the man-to-woman messages $\{\mathrm{A}, \mathrm{R}, \mathrm{W}\}$ here, since the MP algorithm will only update the W message to either A or R, and update the A message to R. Likewise, the woman-to-man messages will only possibly be updated from U to either Y or N, and from N to Y. The rest of the argument is analogous to that in the proof of Lemma 3 in Appendix B. The result of Theorem 1 also follows in this modified case, with a set of modified DE equations and $\tilde{\lambda}$ given in Definition 10 in Appendix J.

Figure 21 shows the value of $\tilde{\lambda}$, together with the fraction of women who exhaust their consideration sets (termed "single" as before) and the fraction of women matched. By comparing the blue curve with that in Figure 15, we can easily see that the deadlock regime indeed reduces as agents are willing to perform more inspections. For example, in a market with the same number of men and women, where women's average degree is five and the inspection success probability is 0.3, if men are willing to inspect their top two candidates this transforms a situation with 87% deadlock into a deadlock-free market.

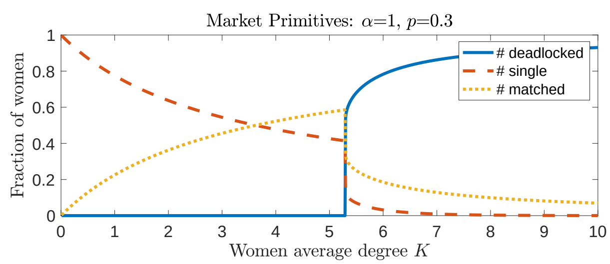

Figure 21: Size of information deadlock, singleton, and matched women as functions of women's average degree K when men are willing to inspection their top two most preferred women. We fix the men-to-women ratio to be $\alpha = 1$ and an inspection's success probability to be $p = 0.3$.

Call the modified PS process which includes top-2 mutual inspections as *top-2 PS*. We formalize in the next lemma that top-2 PS clears the market better than the vanilla PS.

Lemma 23. *If a woman-man pair (i, j) match and leave by time t under standard PS, then they also match and leave by time t under top-2 PS. Also, the residual market at time t under top-2 PS is always a subset of that under standard PS, i.e., $\mathcal{E}_t^{top\text{-}2\ PS} \subseteq \mathcal{E}_t^{PS}$.*

I Appendix for Section H.1

In this section, we provide supplementary materials for Section H.1, including the modified density evolution in Definition 9 and the modified main result in Theorem 3.

[17] We do not count women who are temporarily matched, since if the process ends then, they can leave the market with their temporariry partner, as opposed to waiting for the next available guaranteed inspection.

We first give the derivation of density evolution. The derivation is similar to the proof of Lemma 5 (DE is exact on trees). By induction, assume that DE is exact on trees for up to iteration $t-1$. Then, at iteration t, by the MP update rule,

$$\begin{aligned}
a_t^P &= \mathbb{P}\,(\text{man } j \text{ sends A to neighboring P-type woman } i \text{ in the } t\text{th iteration})\\
&= \mathbb{E}_{d_P}\{(\nu_{t-1}^P)^{d_P} | \text{P-type woman } i \text{ ranks number } d_P+1 \text{ for man } j\}\\
&\quad (\text{note that } d_P \sim\text{Poisson}\left(\frac{\beta K u}{\alpha}\right) \text{ where } u \sim\text{Uniform}[0,1])\\
&= \mathbb{E}_u e^{\frac{\beta K u}{\alpha}(\nu_{t-1}^P-1)}\\
&= \frac{e^{\frac{\beta K}{\alpha}(\nu_{t-1}^P-1)}-1}{\frac{\beta K}{\alpha}(\nu_{t-1}^P-1)},
\end{aligned}$$

$$\begin{aligned}
a_t^O &= \mathbb{P}\,(\text{man } j \text{ sends A to neighboring O-type woman } i \text{ in the } t\text{th iteration})\\
&= \mathbb{E}_{d_P,d_O}\{(\nu_{t-1}^P)^{d_P}(\nu_{t-1}^O)^{d_O} | \text{O-type woman } i \text{ ranks number } d_P+d_O+1 \text{ for man } j\}\\
&\quad (\text{note that } d_P \sim\text{Poisson}\left(\frac{\beta K}{\alpha}\right) \text{ and } d_O \sim\text{Poisson}\left(\frac{(1-\beta)uK}{\alpha}\right) \text{ where } u \sim\text{Uniform}[0,1])\\
&= \mathbb{E}_u e^{\frac{\beta K}{\alpha}(\nu_{t-1}^P-1)} e^{\frac{(1-\beta)Ku}{\alpha}(\nu_{t-1}^O-1)}\\
&= e^{\frac{\beta K}{\alpha}(\nu_{t-1}^P-1)} \cdot \frac{e^{\frac{(1-\beta)K}{\alpha}(\nu_{t-1}^O-1)}-1}{\frac{(1-\beta)K}{\alpha}(\nu_{t-1}^O-1)},
\end{aligned}$$

$$\begin{aligned}
r_t^P =& \mathbb{P}\,(\text{man } j \text{ sends R to neighboring P-type woman } i \text{ in the } t\text{th iteration})\\
=& \mathbb{E}_{d_P}\left[\sum_{l=1}^{d_P}((\nu_{t-1}^P)^{l-1} y_{t-1}^P) \middle| \text{P-type woman } i \text{ ranks number } d_P+1 \text{ for man } j\right]\\
=& \mathbb{E}_{d_P}\left[\frac{(1-(\nu_{t-1}^P)^{d_P}) y_{t-1}^P}{1-\nu_{t-1}^P} \middle| \text{P-type woman } i \text{ ranks number } d_P+1 \text{ for man } j\right]\\
=& \frac{y_{t-1}^P(1-a_t^P)}{1-\nu_{t-1}^P},
\end{aligned}$$

$$\begin{aligned}
r_t^O =& \mathbb{P}\,(\text{man } j \text{ sends R to neighboring O-type woman } i \text{ in the } t\text{th iteration})\\
=& \mathbb{E}_{d_P,d_O}\Bigg[\sum_{l=1}^{d_P}(\nu_{t-1}^P)^{l-1} y_{t-1}^P + \sum_{l=d_P+1}^{d_P+d_O}((\nu_{t-1}^P)^{d_P}(\nu_{t-1}^O)^{l-d_P-1} y_{t-1}^O)\\
&\qquad \Bigg| \text{O-type woman } i \text{ ranks number } d_P+d_O+1 \text{ for man } j\Bigg]\\
=& \mathbb{E}_{d_P,d_O}\Bigg[\frac{(1-(\nu_{t-1}^P)^{d_P}) y_{t-1}^P}{1-\nu_{t-1}^P} + \frac{(1-(\nu_{t-1}^O)^{d_O})(\nu_{t-1}^P)^{d_P} y_{t-1}^O}{1-\nu_{t-1}^O}\\
&\qquad \Bigg| \text{O-type woman } i \text{ ranks number } d_P+d_O+1 \text{ for man } j\Bigg]
\end{aligned}$$

$$=\frac{y^P_{t-1}\left(1-e^{\frac{\beta K}{\alpha}(\nu^P_{t-1}-1)}\right)}{1-\nu^P_{t-1}}+\frac{\left(e^{\frac{\beta K}{\alpha}(\nu^P_{t-1}-1)}-a^O_t\right)y^O_{t-1}}{1-\nu^O_{t-1}},$$

$$w^P_t=1-a^P_t-r^P_t,\qquad w^O_t=1-a^O_t-r^O_t,$$

$$\begin{aligned}
q^P_t &= \mathbb{P}\left(\text{man } j \text{ is inspection ready for neighboring P-type woman } i \text{ in the } t\text{th iteration}\right)\\
&= \mathbb{E}_d\left[\left(a^P_t(1-p)+r^P_t\right)^d \middle| \text{man } j \text{ ranks number } d+1 \text{ for P-type woman } i\right]\\
&\quad (\text{note that } d \sim\text{Poisson}(Ku) \text{ where } u \sim\text{Uniform}[0,1])\\
&= \frac{1-e^{-K(a^P_t p+w^P_t)}}{K(a^P_t p+w^P_t)},
\end{aligned}$$

$$\begin{aligned}
q^O_t &= \mathbb{P}\left(\text{man } j \text{ is inspection ready for neighboring O-type woman } i \text{ in the } t\text{th iteration}\right)\\
&= \mathbb{E}_d\left[\left(a^O_t(1-p)+r^O_t\right)^d \middle| \text{man } j \text{ ranks number } d+1 \text{ for O-type woman } i\right]\\
&\quad (\text{note that } d \sim\text{Poisson}(Ku) \text{ where } u \sim\text{Uniform}[0,1])\\
&= \frac{1-e^{-K(a^O_t p+w^O_t)}}{K(a^O_t p+w^O_t)},
\end{aligned}$$

$$y^P_t=q^P_t p,\qquad y^O_t=q^O_t p,$$

$$\begin{aligned}
\nu^P_t &= \mathbb{P}\left(\text{P-type woman } i \text{ sends N to neighboring man } j \text{ in the } t\text{th iteration}\right)\\
&= q^P_t(1-p)+\mathbb{E}_d\left[\sum_{l=1}^{d}\left(a^P_t(1-p)+r^P_t\right)^{l-1}a^P_t p \middle| \text{man } j \text{ ranks number } d+1 \text{ for P-type woman } i\right]\\
&= q^P_t(1-p)+\frac{a^P_t p(1-q^P_t)}{a^P_t p+w^P_t},
\end{aligned}$$

$$\begin{aligned}
\nu^O_t &= \mathbb{P}\left(\text{O-type woman } i \text{ sends N to neighboring man } j \text{ in the } t\text{th iteration}\right)\\
&= q^O_t(1-p)+\mathbb{E}_d\left[\sum_{l=1}^{d}\left(a^O_t(1-p)+r^L_t\right)^{l-1}a^O_t p \middle| \text{man } j \text{ ranks number } d+1 \text{ for L type woman } i\right]\\
&= q^O_t(1-p)+\frac{a^O_t p(1-q^O_t)}{a^O_t p+w^O_t},
\end{aligned}$$

$$u^P_t=1-y^P_t-\nu^P_t,\qquad u^O_t=1-y^O_t-\nu^O_t.$$

The DE equations for this case is summarized in the following definition.

Definition 9. *For any* $(\alpha,\beta,K,p)\in\mathbb{R}_+\times(0,1)\times\mathbb{R}_+\times(0,1)$*, let* $u^P_0=u^O_0=1, y^P_0=y^O_0=\nu^P_0=\nu^O_0=0$ *and define sequence* $(a^P_t,a^O_t,r^P_t,r^O_t,w^P_t,w^O_t,q^P_t,q^O_t,y^P_t,y^O_t,\nu^P_t,\nu^O_t,u^P_t,u^O_t)_{t\geq1}$ *as follows.*

For all $t = 1, 2, \ldots,$ *let*

$$a_t^P = \frac{e^{\frac{\beta K}{\alpha}(\nu_{t-1}^P - 1)} - 1}{\frac{\beta K}{\alpha}(\nu_{t-1}^P - 1)}; \quad r_t^P = \frac{y_{t-1}^P(1 - a_t^P)}{1 - \nu_{t-1}^P}; \quad w_t^P = 1 - a_t^P - r_t^P;$$

$$q_t^P = \frac{1 - e^{-K(a_t^P p + w_t^P)}}{K(a_t^P p + w_t^P)}; \quad y_t^P = q_t^P p; \quad \nu_t = q_t^P(1-p) + \frac{a_t^P p(1 - q_t^P)}{a_t^P p + w_t^P}; \quad u_t^P = 1 - y_t^P - \nu_t^P.$$

$$a_t^O = e^{\frac{\beta K}{\alpha}(\nu_{t-1}^P - 1)} \cdot \frac{e^{\frac{(1-\beta)K}{\alpha}(\nu_{t-1}^O - 1)} - 1}{\frac{(1-\beta)K}{\alpha}(\nu_{t-1}^O - 1)}; \quad r_t^O = \frac{y_{t-1}^P\left(1 - e^{\frac{\beta K}{\alpha}(\nu_{t-1}^P - 1)}\right)}{1 - \nu_{t-1}^P} + \frac{\left(e^{\frac{\beta K}{\alpha}(\nu_{t-1}^P - 1)} - a_t^O\right) y_{t-1}^O}{1 - \nu_{t-1}^O};$$

$$w_t^O = 1 - a_t^O - r_t^O; \quad q_t^O = \frac{1 - e^{-K(a_t^L p + w_t^O)}}{K(a_t^O p + w_t^O)}; \quad y_t^O = q_t^O p; \quad \nu_t^O = q_t^O(1-p) + \frac{a_t^O p(1 - q_t^O)}{a_t^O p + w_t^O};$$

$$u_t^O = 1 - y_t^O - \nu_t^O.$$

Also define $\lambda^P(\alpha, \beta, K, p) := \lim_{t\to\infty} K w_t^P q_t^P$*,* $\lambda^O(\alpha, \beta, K, p) := \lim_{t\to\infty} K w_t^O q_t^O$*,* $\lambda^B(\alpha, \beta, K, p) := \beta\lambda^P(\alpha, \beta, K, p) + (1-\beta)\lambda^O(\alpha, \beta, K, p)$*, and*

$$\Theta^r = \{(\alpha, \beta, K, p) : \alpha \geq 0, \beta \in [0,1], K \geq 0, p \in [0,1], \lambda^r(\alpha, \beta, K, p) > 0\}$$

for $r = P, O, B$.

Next we state the main theorem of this extension, which is analogous to Theorem 1 in the base model. Recall that in a random market $G_N(\alpha, \beta, K, p)$ with βN P-type women, following the PS process defined in Figure 4, the number of P-type and O-type women who have not cleared at time t are denoted by $\Lambda^t_{\mathrm{P},N}(\alpha, \beta, K, p)$ and $\Lambda^t_{\mathrm{O},N}(\alpha, \beta, K, p)$, respectively. Also, $\Lambda^t_{B,N}(\alpha, \beta, K, p)$ is their sum.

Theorem 3. *Consider a sequence of markets* $G_N(\alpha, \beta, K, p)$ *indexed by* N *and the corresponding* $\Lambda^t_{P,N}(\alpha, \beta, K, p)$*,* $\Lambda^t_{O,N}(\alpha, \beta, K, p)$ *and* $\Lambda^t_{B,N}(\alpha, \beta, K, p)$*, for some sequence of* t *also implicitly indexed by* N*. Consider set* Θ^r *and* $\lambda^r(\alpha, \beta, K, p)$ *for* $r = P, O, B$ *in Definition 9. Then the following statements are true for* $r = P, O, B$*.*

- *If* $(\alpha, \beta, K, p) \in \Theta^r$*, i.e.,* $\lambda^r(\alpha, \beta, K, p) > 0$*, then for any sequence of times* $t = w(1)$ *that is also* $o(\log N)$*, we have*

  $$\lim_{N\to\infty} \mathbb{P}\left(|\Lambda^t_{r,N}(\alpha, \beta, K, p) - \lambda^r(\alpha, \beta, K, p)N| \leq f(N)\right) = 1$$

  for some $f(N) = o(N)$*.*

- *If* $(\alpha, \beta, K, p) \notin \Theta^r$*, then for any sequence of* $t = w(1)$*, we have*

  $$\lim_{N\to\infty} \mathbb{P}\left(\Lambda^t_{r,N}(\alpha, \beta, K, p) \leq f(N)\right) = 1$$

  for some $f(N) = o(N)$*.*

Proof of Theorem 3. Analogous to the proof of Theorem 1. □

J Appendix for Section H.2

In this section, we provide supplementary materials for Section H.2, including the modified APS in Figure 22, the DE equations in Definition 10, and the proof of Lemma 23.

Accelerated Partner Search (APS) following top-2 mutual inspections

Initialize the market: $\tilde{\mathcal{E}}_0^{\text{APS}} = \emptyset$; $\mathcal{I}_0^{\text{APS}} \leftarrow \mathcal{I}$; $\mathcal{J}_0^{\text{APS}} \leftarrow \mathcal{J}$; $\mathcal{N}_0^{\text{APS}}(i) \leftarrow \mathcal{N}(i), \forall i \in \mathcal{I}_0$; $\mathcal{N}_0^{\text{APS}}(j) \leftarrow \mathcal{N}(j), \forall j \in \mathcal{J}_0$. Eliminate women and men with empty consideration set.

$t' \leftarrow 1$

while $\exists$ a woman i with a nonempty consideration set $\mathcal{N}_{t'-1}^{\text{APS}}(i)$ **do** ▷ round t'

[Phase 1.] Each man $j \in \mathcal{J}_{t'-1}^{\text{APS}}$ makes an offer to his best woman i_1 in the consideration set $\mathcal{N}_{t'-1}^{\text{APS}}(j)$. Woman i_1 (case 1) rejects the offer if $i_1 \notin \mathcal{I}_{t'-1}^{\text{APS}}$. She (case 2) does nothing if j is in her consideration set $\mathcal{N}_{t'-1}^{\text{APS}}(i_1)$ but not the best one. She agrees to inspect with j if he is the best in $\mathcal{N}_{t'-1}^{\text{APS}}(i_1)$. If the inspection succeeds (case 3), she accepts the offer and the pair is temporarily matched. If the inspection fails (case 4), i_2 rejects the offer. Under case 1, 2 and 4, j makes a second offer to his second best woman i_2 in $\mathcal{N}_{t'-1}^{\text{APS}}(j)$. i_2 then acts according to the same rules as did i_1. j stops only when his offer is accepted, or when he does not hear back from two women, or when j reaches the end of his consideration set. Then we update the consideration set of j by removing all women who rejected his offer and call it $\mathcal{N}_{t'}^{\text{APS}}(j)$. Also update $\mathcal{J}_{t'}^{\text{APS}}$ by removing all men with an empty consideration set. If a man j is temporarily matched with i during this phase, update $\tilde{\mathcal{E}}_{t'}^{\text{APS}}$ by including (i, j) and removing any pair (i', j) where $i' \neq i$. If (i', j) is removed from $\tilde{\mathcal{E}}_{t'}^{\text{APS}}$, then move i' to $\mathcal{I}_{t'-1}^{\text{APS}}$.

[Phase 2.] Each woman $i \in \mathcal{I}_{t'-1}^{\text{APS}}$ finds her best man $j_1 \in \mathcal{J}_{t'}^{\text{APS}}$ in her consideration set $\mathcal{N}_{t'-1}^{\text{APS}}(i)$ and inspects with j_1 if she received an offer from him. If the inspection succeeds, the pair is temporarily matched. If the inspection fails, i goes to her second best man $j_2 \in \mathcal{J}_{t'}^{\text{APS}}$ in $\mathcal{N}_{t'-1}^{\text{APS}}(i)$. i stops searching only when she is temporarily matched, or when her next best man in $\mathcal{J}_{t'}^{\text{APS}}$ did not make her an offer. Then we update the consideration set of i by removing all men that failed the inspection and all men that are not in $\mathcal{J}_{t'}^{\text{APS}}$ and call it $\mathcal{N}_{t'}^{\text{APS}}(i)$. Also update $\mathcal{I}_{t'}^{\text{APS}}$ by removing all women who are temporarily matched and all women with an empty consideration set. Update $\tilde{\mathcal{E}}_{t'}^{\text{APS}}$ by including all temporarily matched woman-man pairs (i, j) in the current phase, and remove (i', j) where $i' \neq i$.

$t' \leftarrow t' + 1$

end while

Figure 22: The Accelerated Partner Search process following top-2 mutual inspections. Preference rankings remain unchanged throughout.

We first give the modified APS definition in Figure 22.

We now give the derivation of density evolution for this extension. Note that in the modified MP algorithm in Figure 20, the only part that is different from that in the base model (Figure 6) is the message updates for A, R, and W. Therefore we only give the derivations for $\tilde{a}_t$ and $\tilde{r}_t$ below, while $\tilde{w}_t = 1 - \tilde{a}_t - \tilde{r}_t$ and the other parts are the same as before. By the MP update rule in Figure 20,

$$
\begin{aligned}
\tilde{a}_t &= \mathbb{P}\left(\text{man } j \text{ sends A to neighboring woman } i \text{ in the } t\text{th iteration}\right) \\
&= \mathbb{E}\left[\mathbb{P}\left(\forall i' \succ_j i, i' \text{ sends N to } j \text{ except for at most one } i'' \succ_j i \text{ such that } i'' \text{ sends U to } j\right)\right] \\
&= \mathbb{E}_d\left[\tilde{\nu}_{t-1}^d\right] + \mathbb{E}_d\left[d\tilde{\nu}_{t-1}^{d-1}\tilde{u}_{t-1}\right] \qquad (\ d \sim \text{Poisson}(Ku/\alpha), \text{ where } u \sim \text{Uniform}(0,1).) \\
&= \frac{e^{\frac{K}{\alpha}(\tilde{\nu}_{t-1}-1)} - 1}{\frac{K}{\alpha}(\nu_{t-1}-1)} + \frac{\frac{K}{\alpha}(\tilde{\nu}_{t-1}-1)e^{\frac{K}{\alpha}(\tilde{\nu}_{t-1}-1)} - e^{\frac{K}{\alpha}(\tilde{\nu}_{t-1}-1)} + 1}{\frac{K}{\alpha}(\tilde{\nu}_{t-1}-1)^2}\tilde{u}_{t-1}, \\
\tilde{r}_t &= \mathbb{P}\left(\text{man } j \text{ sends R to neighboring woman } i \text{ in the } t\text{th iteration}\right) \\
&= \mathbb{E}[\mathbb{P}(\exists i' \succ_j i, i' \text{ sends Y to } j \text{ and } \forall i'' \succ_j i', i'' \text{ sends N to } j \text{ for at most one } i''' \succ_j i' \\
&\qquad\qquad \text{such that } i''' \text{ sends U to } j)] \\
&= \mathbb{E}_d\left[\sum_{l=1}^{d}\left(\tilde{\nu}_{t-1}^{l-1} + (l-1)\tilde{\nu}_{t-1}^{l-2}\tilde{\nu}_{t-1}\right)\tilde{y}_{t-1}\right] \\
&= \frac{1 - \tilde{\nu}_{t-1} + \tilde{u}_{t-1} + \frac{\alpha(\tilde{\nu}_{t-1}-1-2\tilde{u}_{t-1})}{K(\tilde{\nu}_{t-1}-1)}\left(e^{\frac{K}{\alpha}(\tilde{\nu}_{t-1}-1)} - 1\right) + e^{\frac{K}{\alpha}(\tilde{\nu}_{t-1}-1)}\tilde{u}_{t-1}}{(1-\tilde{\nu}_{t-1})^2}\tilde{y}_{t-1}.
\end{aligned}
$$

We summarize the DE equations in Definition 10 below.

Definition 10. *For any* $(\alpha, K, p) \in \mathbb{R}_+ \times \mathbb{R}_+ \times (0,1)$, *let* $\tilde{u}_0 = 1, \tilde{y}_0 = \tilde{\nu}_0 = 0$, *and define sequence* $(\tilde{a}_t, \tilde{r}_t, \tilde{w}_t, \tilde{q}_t, \tilde{y}_t, \tilde{\nu}_t, \tilde{u}_t)_{t\geq 1}$ *as follows. For all* $t = 1, 2, ...,$ *let*

$$\begin{aligned}
\tilde{a}_t &= \frac{e^{\frac{K}{\alpha}(\tilde{\nu}_{t-1}-1)} - 1}{\frac{K}{\alpha}(\tilde{\nu}_{t-1}-1)} + \frac{\frac{K}{\alpha}(\tilde{\nu}_{t-1}-1)e^{\frac{K}{\alpha}(\tilde{\nu}_{t-1}-1)} - e^{\frac{K}{\alpha}(\tilde{\nu}_{t-1}-1)} + 1}{\frac{K}{\alpha}(\tilde{\nu}_{t-1}-1)^2} u_{t-1};\\
r_t &= \frac{1 - \tilde{\nu}_{t-1} + \tilde{u}_{t-1} + \frac{\alpha(\tilde{\nu}_{t-1}-1-2\tilde{u}_{t-1})}{K(\tilde{\nu}_{t-1}-1)}\left(e^{\frac{K}{\alpha}(\tilde{\nu}_{t-1}-1)} - 1\right) + e^{\frac{K}{\alpha}(\tilde{\nu}_{t-1}-1)}\tilde{u}_{t-1}}{(1-\tilde{\nu}_{t-1})^2}\tilde{y}_{t-1};\\
\tilde{w}_t &= 1 - \tilde{a}_t - \tilde{r}_t;\\
\tilde{q}_t &= \frac{1 - e^{-K(\tilde{a}_t p + \tilde{w}_t)}}{K\left(\tilde{a}_t p + \tilde{w}_t\right)}; \quad \tilde{y}_t = \tilde{q}_t p; \quad \tilde{\nu}_t = \tilde{q}_t(1-p) + \frac{\tilde{a}_t p(1-\tilde{q}_t)}{\tilde{a}_t p + \tilde{w}_t}; \quad \tilde{u}_t = \frac{\tilde{w}_t(1-\tilde{q}_t)}{\tilde{a}_t p + \tilde{w}_t}.
\end{aligned}$$

Also define $\tilde{\lambda}(\alpha, K, p) := \lim_{t\to\infty} K\tilde{w}_t\tilde{q}_t$.

Next we prove Lemma 23.

Proof of Lemma 23. We prove this lemma by induction. Obviously the statement is true at time $t = 1$, since the initial market at $t = 0$ under either PS or top-2 PS is identical, and any guaranteed inspection under standard PS must also be a top-1 guaranteed inspection under top-2 PS (so this pair must be permanently matched as opposed to temporarily matched). This implies that the residual market at time $t = 1$ under top-2 PS must be a subset of that under standard PS, i.e., $\mathcal{E}_1^{\text{top-2 PS}} \subseteq \mathcal{E}_1^{\text{PS}}$. Assume the statement is true by time $t-1$, i.e., assume that if a woman-man pair (i, j) matched and left by time $t-1$ under standard PS, then they must also matched and left by time $t-1$ under top-2 PS. Also assume that the residual market under top-2 PS is a subset of the that under standard PS at time $t-1$, i.e., $\mathcal{E}_{t-1}^{\text{top-2 PS}} \subseteq \mathcal{E}_{t-1}^{\text{PS}}$. Consider time t. We will show that if a woman-man pair (i, j) matched and left by time t under standard PS, then they must also matched and left by time t under top-2 PS. We will also show that the residual market under top-2 PS is a subset of the residual market under standard PS at time t, i.e., $\mathcal{E}_t^{\text{top-2 PS}} \subseteq \mathcal{E}_t^{\text{PS}}$.

We first show that if a woman-man pair (i, j) matched and left by time t under standard PS, then they must alsohave matched and left by time t under top-2 PS. If (i, j) matched and left before time t under PS, then by assumption they must also matched and left before t under top-2 PS. It only remains to consider the case that (i, j) match and leave at time t under PS. If (i, j) match and leave at time t under standard PS, then there are two cases for the same pair at time t under top-2 PS. In case one, $(i, j) \in \mathcal{E}_{t-1}^{\text{top-2 PS}}$, meaning (i, j) are still considering each other at the beginning of the t-th iteration. In case two, $(i, j) \notin \mathcal{E}_{t-1}^{\text{top-2 PS}}$, meaning (i, j) are no longer considering each other at the beginning of the t-th iteration. Next we discuss the two cases respectively.

In case one, $(i, j) \in \mathcal{E}_{t-1}^{\text{top-2 PS}}$. Since (i, j) is a guaranteed inspection in $\mathcal{E}_{t-1}^{\text{PS}}$ and since $\mathcal{E}_{t-1}^{\text{top-2 PS}} \subseteq \mathcal{E}_{t-1}^{\text{PS}}$ by assumption, (i, j) must also be a top-1 guaranteed inspection in $\mathcal{E}_{t-1}^{\text{top-2 PS}}$. Therefore, (i, j) must inspect, match and leave at time t under top-2 PS (note that their latent inspection outcome is a success by assumption).

In case two, $(i, j) \notin \mathcal{E}_{t-1}^{\text{top-2 PS}}$. This can be further classified into two subcases. In subcase one, woman i is already matched and left (note that this excludes the possibility of i being temporarity matched) with a preferred man by time $t-1$ under top-2 PS. In subcase two, man j is already matched (possibility temporarily matched) with a preferred woman by time $t-1$ under top-2 PS. The possibility of (i, j) inspected and failed is excluded since their latent inspection outcome must be a success. The possibility of i or j matched and left with a less preferred partner is also excluded since this means that j or i was ruled out by i or j, i.e., j or i matched with some preferred partner. This goes back to one of the two subcases above. Next we discuss the two subcases.

First consider subcase one. If i is matched and left with j, then we are done. Otherwise, i is matched and left with another man $j' \succ_i j$, say, at time t' under top-2 PS, for some $t' < t$. This implies that (i, j') must be considering each other at the beginning of time t' under top-2 PS, i.e., $(i, j') \in \mathcal{E}_{t'-1}^{\text{top-2 PS}}$. Since $\mathcal{E}_{t'-1}^{\text{top-2 PS}} \subseteq \mathcal{E}_{t'-1}^{\text{PS}}$ by assumption, (i, j') must also be considering each other at the beginning of time t' under PS. In order for (i, j) to inspect at time $t > t'$ under PS, i must have ruled out all preferred men before t, including j'. Denote by t'' the time i ruled out j' under PS. We must have $t' \leq t'' < t$. Note that (i, j') cannot inspect and fail, since their latent inspection outcome is a success. Therefore j' must be matched and left with some other woman $i' \neq i$ at time t'' under PS in order to be ruled out by i. By the induction assumption, this indicates that (i', j') must also be matched and left by time t'' under top-2 PS, contradicting the presumed fact that j' is matched and left with woman $i \neq i'$ at time t' under top-2 PS.

Now consider subcase two. Suppose j is matched (possibly temporarily) with another woman $i' \succ_j i$, say, at time t' under top-2 PS, for some $t' < t$. Since $\mathcal{E}_{t'-1}^{\text{top-2 PS}} \subseteq \mathcal{E}_{t'-1}^{\text{PS}}$ by assumption, it must be that $(i', j) \in \mathcal{E}_{t'-1}^{\text{PS}}$. In order for (i, j) to be able to match and leave at $t > t'$ under PS, j must have ruled out i' before t and after t'. This implies that i' matched and left with a preferred man $j' \succ_{i'} j$, say, at time $t'' \in [t', t)$ under PS. By the induction assumption, (i', j') also matched and left by time t'' under top-2 PS, making it impossible for (i', j) to also match and leave. The only remaining possibility is that (i', j) temporarily matched at t', and later (i', j') matched and left by $t'' > t'$ under top-2 PS, indicating $(i', j') \in \mathcal{E}_{t'-1}^{\text{top-2 PS}}$. However this is also impossible since $j' \succ_{i'} j$ and hence (i', j) cannot be a top-2 guaranteed inspection in $\mathcal{E}_{t'-1}^{\text{top-2 PS}}$.

We have thus proved that if a woman-man pair (i, j) matched and left by time t under standard PS, then they must also matched and left by time t under top-2 PS. It remains to show that $\mathcal{E}_t^{\text{top-2 PS}} \subseteq \mathcal{E}_t^{\text{PS}}$. It suffices to show that if (i, j) inspected and failed at time t under PS, then $(i, j) \notin \mathcal{E}_t^{\text{top-2 PS}}$. If $(i, j) \notin \mathcal{E}_{t-1}^{\text{top-2 PS}}$, then we are done. Suppose $(i, j) \in \mathcal{E}_{t-1}^{\text{top-2 PS}}$. Note that (i, j) inspected and failed at time t under PS implies that (i, j) is a guaranteed inspection in $\mathcal{E}_{t-1}^{\text{PS}}$. Then since $\mathcal{E}_{t-1}^{\text{top-2 PS}} \subseteq \mathcal{E}_{t-1}^{\text{PS}}$ by assumption, (i, j) must also be a guaranteed inspection in $\mathcal{E}_{t-1}^{\text{top-2 PS}}$. Thus (i, j) must inspect and fail at t under top-2 PS, and hence $(i, j) \notin \mathcal{E}_t^{\text{top-2 PS}}$.

□