On Matching and Thickness in Heterogeneous Dynamic Markets*

Itai Ashlagi‡ Maximillien Burq Patrick Jaillet§
Vahideh Manshadi

June 14, 2016

Abstract

We study dynamic matching in an infinite-horizon stochastic networked market, in which some agents are a priori more difficult to match than others. Agents have compatibility-based preferences and can match either bilaterally or indirectly through chains. We study the effects matching technologies and matching policies have on efficiency in markets with different compositions of hard and easy-to-match agents.

First, we analyze myopic matching policies and identify a strong connection between the market thickness and the efficiency driven by the matching technology. We show that when hard-to-match agents join the market more frequently than easy-to-match ones, moving from bilateral matchings to chains significantly increases efficiency. Otherwise, the difference between matching bilaterally or through a chain is negligible. Second, we show that the lack of thickness cannot be compensated by non-myopic matching policies implying that the only way to thicken the market is by attracting more agents.

1 Introduction

Many matching markets are naturally dynamic. Every year thousands of incompatible patient-donor pairs register to kidney exchange clearinghouses that search periodically to

*First Draft: Feb 2015.

‡Ashlagi acknowledges the research support of the National Science Foundation grant SES-1254768.

§Jaillet acknowledges the research support of the National Science Foundation grant OR-1029603 and the Office of Naval Research grants N00014-12-1-0033 and N00014-15-1-2083.

match these pairs. Online platforms (dating, online workplace, etc.), labor markets, and even housing markets can be viewed as dynamic matching markets.

The matching policy, which determines when and who to match, plays an important role in the efficiency of the marketplace. A myopic policy, which attempts to match agents upon arrival, may have short run benefits but could harm future arriving agents. So a centralized clearinghouse may wait to thicken the market before identifying matches.[1] Interestingly, kidney exchange clearinghouses typically search for matches very frequently.[2]

In various marketplaces, the matching technology is also instrumental for efficiency. While kidney exchanges were first conducted in 2-way cycles (bilateral matches), most transplants are now conducted through chains initiated by an altruistic donor (Anderson et al. (2015)).[3] Transactions in housing markets may also be viewed as conducted through chains.[4] In some markets, such as dating, only bilateral matches take place.

This paper is concerned with the effects matching technologies and matching policies have on the efficiency of markets with different thickness levels. We are not the first to address these issues. However, previous studies focus on either very thick (Ünver (2010)) or very thin marketplaces (Anderson et al. (2013), Akbarpour et al. (2014)). In the former, chains (or exchanges through long cycles) do not improve efficiency in contrast to empirical evidence. The latter works restrict attention to markets in which all agents are ex ante symmetric. Instead, we look at markets that are comprised of both thick and thin components, and we explore the effects under different compositions of the marketplace.

We study a stylized infinite horizon model with two types of agents distinguished by their difficulty to match. Every period a single agent arrives to the market whose type is either *hard-to-match* (H), or *easy-to-match* (E), with probability θ and $1-\theta$, respectively. Easy and hard-to-match agents can be matched by any other agent independently with probability p_E and p_H, respectively, where $p_H \ll p_E$. So the fraction of hard-to-match agents joining the market, θ, can be viewed as a measure for the market thickness, with a higher fraction interpreted as a thinner market.

Agents in our model prefer to match as early as possible and indifferent between acceptable matches. We adopt the average waiting time of agents in steady-state as a measure for efficiency. Agents leave the market only after they match (but see Section 4 for a discus-

[1]How to attract agents to the marketplace is a complementary issue (see Ashlagi and Roth (2014) who study how to incentivize hospitals to fully participate in kidney exchange).

[2]Matching is done on a daily basis at the Alliance for Paired Donation and the National Kidney Registry, weekly in the United Network for Organ Sharing and monthly in the Netherlands national program.

[3]A chain consists of a sequence of patient-donor pairs with the donor of each pair donating to the patient of the next pair. The chain is initiated by an altruistic donor.

[4]Typically an agent will sell a house to some agent and buy from another.

sion about a model with departures). Two types of matchings are considered, bilateral and chains, and we assume that only one of these take place. In a bilateral exchange a pair of agents match (with) each other and in a chain an agent is matched by one agent but matches another. We are interested in the behavior of the average waiting time for small values of p_H.

We remark that Anderson et al. (2013) and Akbarpour et al. (2014) analyze optimal policies in a market with only hard-to-match agents (in fact Anderson et al. (2013) is a special case of our model where $\theta = 1$). Markets, however, typically have more heterogeneity. Ashlagi et al. (2013) document that kidney exchange pools most patients' sensitization level is either very low or very high.[5]

Our main contribution is identifying a tight connection between the fraction of hard-to-match agents joining the market and the effect the matching technology has on efficiency. When easy-to-match agents join the market more frequently than hard-to-match ones ($\theta < 0.5$), myopic matching policies that use a chain, or just bilateral matches, are both approximately optimal. Otherwise, any bilateral matching policy is highly sub-optimal.

More formally, we first analyze a myopic policy which conducts bilateral matches. Since easy-to-match agents in our model can be matched almost instantaneously (relatively to hard-to-match agents), we focus on the average waiting time of hard-to-match agents. For a fixed p_E, and sufficiently small values of p_H, the average waiting time of hard-to-match agents is of order $\Theta(\frac{1}{p_E p_H})$ for any $\theta < 0.5$, and $\Theta(\frac{1}{p_H^2})$ for any $\theta \geq 0.5$.[6]

Intuitively, the waiting time for agents is inversely proportional to the probability for a bilateral match to occur. When hard-to-match agents are on the long side ($\theta \geq 0.5$), many hard-to-match agents must match with each other, but otherwise almost all hard-to-match agents match with easy-to-match ones. [7]

Next, we analyze a policy that conducts matches by constructing a single chain myopically. Whenever an agent is matched, the next agent in the chain is instantaneously selected arbitrarily among all feasible agents, and so forth. Once the chain cannot be continued, the last agent in the chain waits for the next arriving agent that it can match. Note that at every period we identify a maximal chain, and not necessarily the longest possible chain. To simplify the analysis we assume that easy-to-match agents can be matched by any other agent. For any $\theta < 1$, hard-to-match agents will wait on average $\Theta(\frac{1}{p_H})$ periods, and interestingly,

[5]The more sensitized a patient , the harder it can be matched.

[6]We write $f = \Theta(g)$ if there exist constants a and b such that $ag(x) \leq f(x) \leq bg(x)$ for every sufficiently large x.

[7]The myopic policy that we analyze assigns priority to H agents in the presence of ties. However, as we show, the average waiting time of H agents remains the same order of magnitude even when ties are broken in favor of E agents.

when $\theta = 1$, their average waiting time is $\Theta(\frac{\log(1/p_H)}{p_H})$. Again, hard-to-match agents' average waiting is inversely proportional to their chance to be matched by a random agent.

We further ask whether non-myopic matching policies can improve hard-to-match agents' waiting times. For example, by making easy-to-match agents wait to thicken the market, or by searching for the longest chain (rather than just advancing it myopically) one may potentially increase overall efficiency. We show, however, that no matching policy can reduce the average waiting time of hard-to-match agents without significantly harming easy-to-match agents and thus any attempt to thicken the pool (subject to its composition) is artificial. Putting together, we find that when hard-to-match agents arrive more frequently than easy-to-match ones, matching through a chain results in much lower waiting times than matching only bilaterally. Otherwise, bilateral matchings are comparable to matching through chains.

Although we carry out our analysis in the asymptotic regime of p_H, using simulations, we find the results to hold even when p_H is not too small (see Appendix A). For example, in kidney exchange many patients have $1 - 5\%$ chance of being compatible with a random donor.

Finally, in Section 4 we briefly discuss several extensions to our model. In Section 4.1 we relax the assumption that arrival rates of easy-to-match and hard-to-match agents are perfectly correlated, and using computational experiments, we explore the effects independent arrival rates have on waiting times. In a homogeneous market, increased arrival rate always means lower waiting times. However, we find that in the heterogeneous case, under bilateral matching, there is a large regime for which the average waiting time increases significantly as the arrival rate of hard to match agents increases. This phenomenon is specific to bilateral matching. Intuitively, as hard-to-match agents' arrival rate increases, the competition to match with easy-to-match agents also increases up to the point they can match with each other quickly. In Section 4.2, we allow agents to depart without being matched and discuss the relation between minimizing the number of unmatched agents and minimizing the average waiting time in the market. In Section 4.3 we further extend our discussion to two-sided markets with asymmetric arrival rates between both sides. In Section 4.4, we discuss how our analysis could be extended to decentralized markets.

Related work

The first strand of research, related directly to our work, is the growing area of dynamic matching markets in networks, where agents are matched to other agents. Several papers in this strand focused on compatibility-based preferences. These studies focus on markets that are either very thick or very sparse and provide different recommendations regarding the best matching technology. Our work aims to bridge the gap between the models by varying the thickness level. Thus we provide a unified model to study efficiency of technologies as thickness varies.

Ünver (2010) analyzes a dynamic kidney exchange model in which preferences are based on compatibilities determined only by blood types. He finds that an "almost" myopic mechanism that uses only 2 and 3-way cycles is optimal.[8] The finite number of types together with deterministic compatibility essentially creates a thick marketplace which cannot explain why chains increase efficiency in practice.[9] Our work deviates by studying a sparse marketplace with many agents that have a small probability of being matched, where chains become an important factor.

Closest to our paper are Ashlagi et al. (2013), Anderson et al. (2013), and Akbarpour et al. (2014) who study mechanisms in dynamic matching markets with a stochastic underlying network where preferences are based on compatibilities.[10] Ashlagi et al. (2013) study a class of batching policies in a finite-horizon model with hard and easy-to-match agents. They analyze the number of matches throughout the horizon. They find that when matching through 2-way and 3-way cycles, there is essentially no benefit in small batches over greedy policies, but chains lead to many more matches compared to 2-way or 3-way cycles.[11] This paper focuses on agents' average waiting time in an infinite horizon and we derive scaling laws for waiting time under various matching policies and thickness levels.

Akbarpour et al. (2014) study an infinite-horizon model with only hard-to-match agents and focus on bilateral matching policies.[12] Agents in their model can depart and they focus on minimizing the loss rate. When agents departure times are known to the clearinghouse, a greedy policy will result in a significant loss. However, when agents' departure times are unknown to the clearinghouse they find the greedy policy to be almost optimal, consistent with our results for the case of bilateral matchings and a market with only hard-to-match agents. This is due to the close relationship between their model and ours and the fact that in both formulations the objective translates to minimizing the expected number of agents

[8]His results are thus closely related to findings in static large markets (see, e.g. Roth et al. (2007)).

[9]See also Gurvich and Ward (2012), who study a similar compatibility-based inventory control model.

[10]See also Dickerson et al. (2012) who also demonstrate the benefit of chains using simulations in dynamic kidney exchange pools.

[11]See also Ding et al. (2015) who study matching using a random walk in a static random graph with hard and easy-to-match agents.

[12]Their scaling is different than ours but closely related: every period a large number of agents arrive to the market and the probability to match is inversely proportional to the arrival rate. We model small arrival rate and hard-to-match agents match with a small probability.

waiting (see Section 4.2 for further discussion).[13]

Anderson et al. (2013) study a special case of our model with only hard-to-match agents. They consider three settings of exchanges, 2-ways, 2 and 3-ways, and chains. In each of these settings, they find a greedy policy to be asymptotically optimal, and that moving from 2-ways or 3-ways to chains reduces significantly the average waiting time. There are two main differences between their work and ours. First, in their greedy policy that matches through a chain, they always search for the longest possible chain and find the average waiting time to be $\Theta(\frac{1}{p_H})$. We find a myopic policy that advances the chain using only local information on the compatibility graph, which still leads to the same average waiting time (except for the case in which all agents are hard-to-match, in which case the waiting time increases by a small logarithmic factor).[14] Second, and more importantly, our model allows us to study how the presence of heterogeneous (both hard-to-match and easy-to-match) agents impacts the efficiency and design of the market, and how much thickness is necessary to create an efficient market.

The second strand of research incorporates preferences that are not only based on compatibility. Baccara et al. (2015) study a dynamic two-sided matching market with two types of agents on each side. They find that in an optimal mechanism, agents wait for some period of time to be matched by their preferred type. They further find that a decentralized market is inefficient. Fershtman and Pavan (2015) characterize the optimal mechanism in a many-to-many two-sided matching market in which agents' preferences change endogenously over time. Doval (2014) and Kadam and Kotowski (2014) study stability in dynamic matching markets. Like Baccara et al. (2015), Doval (2014) finds that a clearinghouse should restrict agents' waiting in order to achieve efficiency.[15]

Another strand of research is about allocation in dynamic one-sided markets, and some papers in this thread study how to organize queues. Zenios (1999) analyzes a queueing model for the allocation of cadaver organs for patients with different types. Leshno (2014) designs queueing policies to minimize the mismatches of allocated objects, which translates into how to incentivize agents to wait for their preferred objects.[16] Note that a dynamic networked market can be viewed as a queue that serves itself.[17]

[13] This relationship is further formalized by Anderson et al. (2013) in a model with only hard-to-match agents.

[14] Finding the longest chain is computationally hard, and requires global information on the whole compatibility graph.

[15] There is also a large literature concerned with search in labor markets, where firms and workers are randomly matched and decide whether to wait for better matches. See the survey by Rogerson et al. (2005).

[16] See also Bloch and Cantala (2014) for a related study on a queueing system.

[17] Some papers that study optimal allocation mechanisms in static settings identify when organizing a queue is optimal (e.g., Hoppe et al. (2009) and Chakravarty and Kaplan (2013)).

Our work is also closely related to the problem of matching multi-class customers to multi-class servers studied in queueing literature (for example, Caldentey et al. (2009)): An infinite sequence of customers and servers arrive over time; each customer can only be severed by a certain types of servers. A few papers study such settings Caldentey et al. (2009); Adan and Weiss (2012) where the underlying graph structure is deterministic, and compute the matching rates under FCFS policy. Further, Gurvich and Ward (2012) studied an optimal control problem in a generalized setting where items of different types arrive to their queues and wait to be matched with items of (possibly more than one) other types. Similar to Caldentey et al. (2009); Adan and Weiss (2012), in the work of Gurvich and Ward (2012), the structure of possible matches among finite number of types is deterministically given. In our model, an agent can be thought as a pair of customer-server, and the compatibility between any two agents is probabilistic, thus we will not have a finite number of queues.

Online bipartite matching has been studied extensively in the past few decades motivated by matching advertisers to users on the web. Initially, Karp et al. (1990) studied online matching in an adversarial setting with no probabilistic information about the structure of the input. Several follow up papers considered settings with various level of stochastic information about the arriving sequence and their corresponding underlying graph Goel and Mehta (2008); Feldman et al. (2009); Manshadi et al. (2011); Jaillet and Lu (2013). The main factor distinguishing our work from this line of research is that we assume that an agent who is not matched upon arrival waits in the pool and do not leave immediately. This allows us to study the timing of matching in addition to the question of how to match.

Finally, several papers study the timing of exchanges in markets with buyers and sellers. For example Mendelson (1982) studies the behavior of prices and quantities resulting from periodic trading. Budish et al. (2013) analyzes the tradeoff between the trading frequency and efficiency in continuous time financial markets.

2 Setup

We study an infinite-horizon dynamic matching market. For every two agents a and a', either a' is *acceptable* to a or not. We say that agent a can be *matched* by a' only if a' is acceptable to a, in which case we also say that a' can *match* a. There are two types of agents, *hard-to-match* and *easy-to-match* denoted by H and E[18]. Any agent is acceptable to an H agent with probability p_H and to an E agent with probability p_E, independently, where $p_H < p_E$. We will be interested in the case where p_H is significantly smaller than p_E. In the context of

[18]Note that in the Kidney exchange application, the patient and the donor together are considered to be one agent.

kidney exchange, medical data determines whether the donor of a patient-donor pair a' can donate (or is acceptable) to the patient of the pair a, and the two types may correspond to low and high sensitized patients, while abstracting away from blood types. In practice, the chance that a patient is incompatible with a random donor based on her antibodies follows approximately a bimodal distribution in kidney exchange programs (Ashlagi et al. (2012)).[19]

Every period a new agent arrives to the market, whose type is either H or E with probability θ and $1-\theta$, respectively. We assume that every agent is indifferent between acceptable matches and prefers to be matched as soon as possible and do not discount the future. Agents in our model leave the market only after they are matched. We therefore adopt the *average waiting time* of agents as a measure for efficiency.

We study matching policies in two different settings distinguished by how agents can get matched: *bilaterally* or indirectly through *chains*. In a bilateral match two agents match each other (which is the common form in the matching and search literatures). In a chain an agent a who gets matched by some agent a', matches a different agent a''. It is worth noting that in kidney exchange, chains initiated by an altruistic donor (who does not expect a kidney in return), account for the majority of transplants (Anderson et al. (2015)).

Assuming that the market reaches a steady-state distribution, we denote by w_H and w_E the average waiting times of H and E agents, respectively. In our analysis we use the observation that $w_H = n_H/\theta$, where n_H is the average number of H agents in steady-state (Little's law).

Several assumptions require some discussion. Indifference between acceptable matches removes agents' incentives to wait for preferable matches. This allows us, however, to focus on the tradeoffs between thickness, efficiency, and the matching technology. Assuming agents have a noisy set of potential matches is realistic. In kidney exchange, for example, there is large heterogeneity among sensitized patients' antibodies, which determine their compatibility with donors. The stochastic structure can further capture agents' tastes. Assuming that match probabilities are independent allows us to track the state of the market and dealing with correlated matches is very challenging. We believe most of our insights would hold in markets with more than two-types but a rigorous analysis is challenging.

Assuming that agents do not leave without matching ignores some interesting tradeoffs but enables us to focus on average waiting times. However, as discussed in Section 4.2, improving the average waiting time results in increasing the chance of matching before departure. Roughly speaking, this suggests that in a market with departures, the average waiting time and the match rate are aligned. Finally, fixing the arrival rate to be deterministic

[19]This probability is called the Panel Reactive Antibody (PRA). A significant fraction of the population have either PRA above 97% or below 20%.

simplifies the analysis, and we believe our results still hold for Poisson arrival rates.

We use the following notations throughout the paper. Consider any two functions $f : \mathbb{R} \to \mathbb{R}$ and $g : \mathbb{R} \to \mathbb{R}_{>0}$. We write $f = o(g)$ if $\lim_{x\to\infty} \frac{f(x)}{g(x)} = 0$; $f = O(g)$ if there exist a constant a such that $f(x) \leq ag(x)$ for every sufficiently large x; $f = \Omega(g)$ if there exist a constant b such that $bg(x) \leq f(x)$ for every sufficiently large x, and $f = \Theta(g)$ if $f = \Omega(g)$ and $f = O(g)$.

3 Myopic matching and beyond

In this section, we first analyze myopic matching policies under bilateral matchings and matching through chains. Next we discuss the optimality of our findings in a broader context of non-myopic matching policies.

3.1 Bilateral matching

We study here myopic matching policies that allow only bilateral matches between pairs of agents. Our benchmark matching policy attempts to match each agent upon arrival bilaterally with another agent in the market while breaking ties in favor of hard-to-match agents. Formally, call *BilateralMatch* the policy that attempts to match an arriving agent a upon arrival with an H agent if possible (breaking ties randomly). If no H agent can form a bilateral match with a, then a is matched to an E agent if possible, and otherwise a remains in the market and waits for an eventual match.

The dynamic system that results from both the stochastic arrivals and the *BilateralMatch* policy can be represented as a Markov chain. Intuitively, agents that are waiting in the market never match each other. Therefore the number of H and E agents uniquely defines the state of the market at any time.

Next, we quantify the average waiting times in steady-state of H and E agents under *BilateralMatch* and find that $\theta = 0.5$ is a sharp threshold for the behavior of hard-to-match agents' average waiting time.

Theorem 1. *Under BilateralMatch, the market reaches steady-state. Furthermore, there exist positive constants A_θ, B_θ, and C_θ such that:*

1. *If $\theta < 1/2$, then*

$$w_H \leq \frac{A_\theta}{p_E p_H}. \tag{1}$$

2. *If $\theta \geq 1/2$, then*

$$w_H \leq \frac{B_\theta}{p_H^2}. \tag{2}$$

3. *For $0 \leq \theta < 1$*

$$w_E \leq \frac{C_\theta}{p_E^2}. \tag{3}$$

Propositions 3 and 4 in Section 3.3 will show that (1) and (2) are in fact asymptotically tight.[20] This shows that H agents wait on average significantly less when they are on the short side ($\theta < 0.5$) than when they are on the long side ($\theta > 0.5$) of the market. Hence, despite the fact that p_H is significantly smaller than p_E, the balance between H and E agents in the market plays a key role in the average waiting times of H agents. The proof of Theorem 1 provides numerical values for the constants A_θ, B_θ, and C_θ.

Intuitively, agents' average waiting time is inversely proportional to the probability for a match to occur. Under a myopic bilateral policy, no existing pair of agents in the market can match each other (otherwise they would match and leave the market). An arriving E agent forms a bilateral match with an existing E agent with probability p_E^2, implying that E agents experience relatively small waiting times, compared to what H agents experience. For an H agent, however, the probability of matching to an existing E agent is $p_E p_H$ and to an existing H agent is p_H^2. When $\theta < 0.5$, we show that almost all H agents are matched with E agents, and thus their average waiting time is $\Theta(1/p_E p_H)$. When H agents arrive more frequently than E agents, there are simply not enough E agents to match with H ones. This means that a non-negligible fraction of H agents must match with each other, thus increasing their average waiting time to $\Theta(1/p_H^2)$.

Somewhat surprisingly, we find that when $p_E = 1$, Theorem 1 remains true even if the *BilateralMatch* policy is modified to break ties in favor of easy-to-match agents. We call this policy *BilateralMatch(E)*.

Proposition 1. *Let $p_E = 1$. Under the BilateralMatch(E), the market reaches steady-state, and there exist positive constants $\tilde{B}_\theta$, $\tilde{C}_\theta$ such that:*

1. *if $\theta < 1/2$, then*

$$w_H \leq \frac{\tilde{A}_\theta}{p_H}.$$

2. *if $\theta \geq 1/2$, then*

$$w_H \leq \frac{\tilde{B}_\theta}{p_H^2}.$$

[20] See also Appendix A.1 where we further discuss tightness, and provide simulations results which show that we can find exact values for A_θ and B_θ.

The proofs of Theorem 1 and Proposition 1 rely on analyzing a Markov chain whose state space can almost be reduced to a single dimension that accounts for the number of H agents in the market. In fact, when $p_E = 1$, the state space is two dimensional, where one dimension has two possible states (the number of L agents in the market - zero or one). For the general case where $p_E < 1$ within the *BilateralMatch(E)* setting, or alternatively when ties are broken randomly among all agents, analyzing the state space is technically challenging, yet we conjecture that the same result holds with perhaps different constants.

3.2 Matching through chains

When only bilateral matchings can be formed we found that when H agents are on the long side their average waiting time is large due to the rare coincidence of wants. In some marketplaces, however, agents match indirectly through chains. In kidney exchange a patient may receive a transplant from a donor of an incompatible pair a, while her intended donor will give a kidney to a patient of a different incompatible pair a'. In housing markets, agents typically purchase from one agent and sell to another or vice versa.[21]

For our chain model, we introduce the notion of a *bridge* agent, who has already been matched by some other agent, but is still waiting to match a different agent. We restrict our attention to the case in which a single bridge agent arrives to the market at time $t = 0$ and therefore only a single chain exists at any given time.

Denote by *ChainMatch* the myopic policy under which the bridge agent b attempts to match the incoming agent, a_1. If such a match is possible a chain segment is initiated in the current period as follows. We pick an agent a_2 uniformly at random among all other agents that a_1 can match (while prioritizing H agents), and if such an agent exists, we pick uniformly at random an agent a_3 among the remaining agents that a_2 can match, and so on and so forth. This process is repeated until the first time an agent, say a_k, cannot match any other agent in the market outside of the set $\{b, a_1, a_2, ...a_{k-1}\}$. We implement these matches (i.e., agents $b, a_1, ..., a_{k-1}$ leave the market), and a_k becomes the new bridge agent. Note that under this matching policy there is always a single bridge agent in the market. Further notice that since matches are formed myopically, the policy does not necessarily implement the longest chain segment in each period.

For our next results, to simplify notations and technical challenges, we restrict our attention to the case in which $p_E = 1$, implying that E agents never remain in the market.[22] We believe that dropping this assumption is not crucial for the next result.[23]

[21] The timing of transactions may depend on the supply and demand in these markets.

[22] This assumption allows us to analyze a single dimension Markov Chain.

[23] This belief is backed by the simulations reported in Appendix A.2.

Theorem 2. *Suppose that $p_E = 1$ and $0 < \theta \leq 1$. Under the ChainMatch policy, the market reaches steady-state and:*

1. *If $\theta < 1$, there exists a constant K_θ such that:*

$$w_H \leq \frac{K_\theta}{p_H}. \tag{4}$$

2. *If $\theta = 1$:*

$$w_H \leq \frac{2\ln(1/p_H)}{p_H}. \tag{5}$$

The formal proof of Theorem 2 is given in Appendix E. The main idea is to analyze the length of a new chain segment that is formed by *ChainMatch.* In Proposition 2 below we use the steady-state property of the system to compute the average length of chain segments. Further, we show that if the number of agents in the system becomes large (of the order of $\Theta\left(\frac{\ln(1/p_H)}{p_H}\right)$), the new chain segment will match a large fraction of agents. Together with Proposition 2 this allows us to upper-bound the number of agents in steady-state.

Figure 1 plots simulation results that compare the average waiting time of hard-to-match agents under the *BilateralMatch* and *ChainMatch* policies for a variety of thickness levels when $p_H = 0.02$ and $p_E = 1$.[24] Observe that when θ takes values below 0.5 the difference between *BilateralMatch* and *ChainMatch* is small. Furthermore, note the threshold effect around $\theta = 0.5$, above which the average waiting time increases rapidly when bilateral matches are conducted.

Anderson et al. (2013) study a special case of our model with only hard-to-match agents (i.e., $\theta = 1$). They show that with a greedy chain policy that always selects the longest chain segment in the market, the average waiting time is $\Theta(\frac{1}{p_H})$. The disadvantages of such a policy is that finding the longest chain segment is computationally hard and is only meaningful when it is implemented by a centralized planner. Our myopic approach is not only computationally straightforward, it can be implemented in a decentralized market as well. Theorem 2 together with Proposition 3 of Section 3.3 show that using the myopic *ChainMatch* policy (as opposed to finding the longest chain segment) has little impact on H agents' average waiting time. When $\theta < 1$, H agents will also wait on average $\Theta(\frac{1}{p_H})$ periods, and when all agents are hard-to-match, their average waiting time increases at most by a logarithmic factor.

The next result gives comparative statics for the expected length of a chain segment.

[24] For all the empirical results, we run 5 independent simulations, each for 100,000 time periods and to avoid transient states we only compute the average waiting times for agents that arrive in the last 20,000 periods.

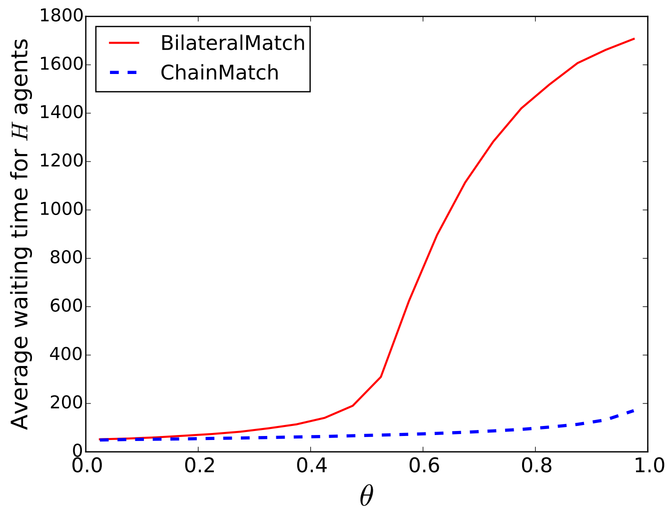

Figure 1: The average waiting time for H agents under the *BilateralMatch* and *ChainMatch* policies for $p_H = 0.02$ and $p_E = 1$.

Proposition 2. *Let L_∞ be the length of a new chain segment formed by ChainMatch policy in steady-state, and suppose $p_E = 1$. Then we have:*

$$\mathbb{E}[L_\infty] = \frac{\theta(1 - p_H)}{\theta p_H + (1 - \theta)} + 1.$$

Proposition 2 implies that increasing arrival rates of easy-to-match agents decreases the expected length of a chain segment, which may reduce some logistical difficulties that arise with long chain segments. For example, in kidney exchange crossmatch (tissue-type) tests should pass prior to actual transplants, and in housing markets, buyers or sellers may wait for earlier transactions to take place.

The proof is given in Appendix E.2. Intuitively, if a policy reaches steady-state, the number of agents matched in a chain segment corresponds to the number of agents that have arrived since the previous chain segment. Therefore, the expected length of a chain segment is the inverse of the frequency at which it occurs. Under the *ChainMatch* policy a chain segment begins when an arriving agent can be matched by the bridge agent, which happens with probability $\theta p_H + (1 - \theta)$, and in the meantime $\theta(1 - p_H)$ agents accumulate in the market.[25]

[25]The proof reveals that Proposition 2 holds true for any greedy chain policy that reaches steady-state.

3.3 Beyond myopic matching

So far we have analyzed myopic matching policies. It is natural to ask whether one can decrease agents' average waiting times by using more sophisticated matching policies.

The next result gives a lower bound on the average waiting time for any matching policy regardless of the matching technology.

Proposition 3. *For any $p_E \leq 1$ and under any matching policy that reaches steady-state, for any $\theta \in (0,1)$ and any $p_H > 0$, there exists a constant c_θ such that $w_H + w_E \geq \frac{c_\theta}{p_H}$.*

The proof uses similar ideas as in Anderson et al. (2013), who study the same model with only hard-to-match agents. We assume that there is at most one bridge donor in the system. The main intuition behind this proof is the following: Suppose that the pool size is too small, then an arriving agent has a small probability of being matched immediately, and therefore must wait a "long" time to obtain at least one incoming edge. This long waiting time contradicts the small pool size (with Little's law).[26]

Proposition 3 provides a bound on the sum of average waiting times. Therefore one may possibly decrease the average waiting time of hard-to-match agents beyond $\frac{1}{p_H}$ (for example, by letting easy-to-match agents wait and artificially thicken the market, one could increase H agents' chances to match and thus reduce their waiting times). This, however, comes at a cost of imposing high waiting times for easy-to-match agents. In particular, Proposition 3 together with results from Sections 3.2 and 3.1 imply that when $\theta < \frac{1}{2}$, a myopic policy, whether using chains or bilateral matches, is optimal up to constant factors. When $\frac{1}{2} < \theta < 1$, matching myopically using chains is optimal up to constant factors.

It is natural to ask whether there exist a bilateral matching policy that decreases average waiting of hard-to-match agents when $\theta > 1/2$? The next proposition shows that the answer to this question is essentially negative.

Proposition 4. *Let $\theta > 1/2$. For any bilateral matching policy that reaches steady-state, under a mild regular condition, there exists a positive constant c'_θ such that $w_H \geq \frac{c'_\theta}{p_H^2}$.*[27]

4 Extensions

In this section, we consider four independent extensions to our model. In 4.1 the assumption that one agent, H or E, arrives every period is relaxed. In 4.2 we discuss the implications

[26] Therefore, having many chains running at the same time is the only way to reduce waiting times beyond $1/p_H$. Suppose that we have $\Omega(\frac{1}{p_H})$ chains in the system. Then the probability of being matched immediately is very large, and waiting times become very small.

[27] The regular condition states that for every H agent that arrives at steady-state the probability to eventually being matching to an E agent is the same.

from adding departures to our model. In 4.3, we discuss two-sided matching markets. Finally, in 4.4 we discuss our model in a decentralized environment.

4.1 The impact of thickening the pool

In our model every period a single agent arrives to the market, implying that the arrival rates of hard and easy-to-match agents are perfectly correlated. We relax this assumption in order to explore the effect that thickening the market has on efficiency.

We conduct simulations assuming independent arrival rates θ_H and θ_E for H and E agents, respectively. In particular arrivals of H and E agents are modeled by two independent Bernoulli processes with rates θ_H and θ_E, respectively. For all our simulations we run 5 independent trials, each for 100k time periods and to avoid transient states we only compute the average waiting times for agents that arrive in the last 20,000 periods.

Simulation results suggest that our theoretical findings, which were derived under a slightly different arrivals model, hold to a large extent. Figures 2 and 3 illustrate the behavior of the marketplace under the *ChainMatch(H)* and *BilateralMatch(H)* matching policies, respectively, under a variety of arrival rates.[28]

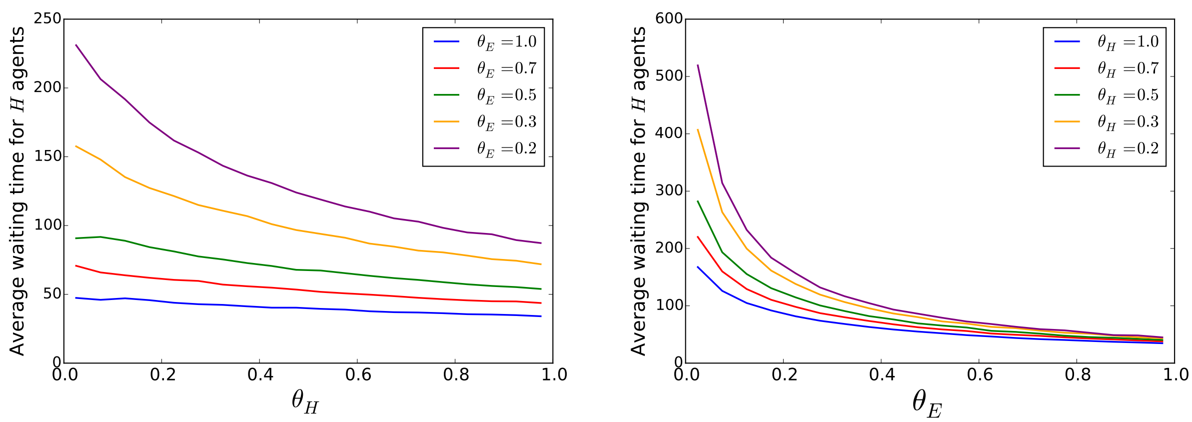

Figure 2: Simulations of a pool with independent arrivals of H and E agents, using *ChainMatch* and parameters $p_H = 0.02$, $p_E = 0.8$. The left panel plots the waiting times of H agents as a function of θ_H, for various values of θ_E. The right panel plots waiting times for H agents as a function of θ_E, for various values of θ_H.

Observe that under the *ChainMatch* policy (Figure 2), the average waiting time for H agents always decreases with both θ_H and θ_E. This means that attracting agents to join the market is always beneficial.

[28] We find similar results when *ChainMatch(E)* and *BilateralMatch(E)* are tested.

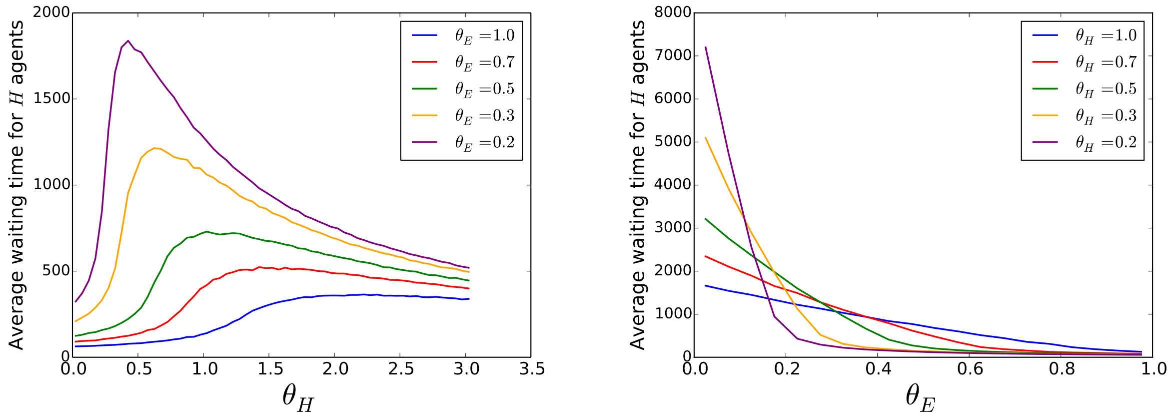

Figure 3: Simulations of a pool with independent arrivals of H and E agents, using *BilateralMatch* and parameters $p_H = 0.02$, $p_E = 0.8$. The left panel plots the waiting times of H agents as a function of θ_H, for varying θ_E. The right panel plots waiting times for H agents as a function of θ_E, for varying θ_H .

When bilateral matches are conducted, the average waiting time of H agents decreases with θ_E, as the right panel in Figure 3 shows (this decrease is sharp as long as $\theta_E \leq \theta_H$).

Interestingly, the average waiting time for H agents is *is not monotone* in θ_H under the *Bilateral Match* policy (left panel in Figure 3). Observe that when θ_H is much smaller than θ_E, the average waiting time for H increases slowly with θ_H. In line with our theoretical findings, this increase becomes steeper when $\theta_H \geq \theta_E$. However, this phenomenon stops after θ_H becomes large enough and the average waiting time decreases again.

Intuitively, when H agents have a small arrival rate they easily get matched to E agents. As θ_H increases, the competition between H agents becomes harsher, and their waiting increases. But when θ_H is large enough, H agents can also match to each other quickly, and the competition is replaced by complementarity.

Following our intuition from the theoretical results, we can look at a stylized computation to quantify what happens: Let $\frac{\theta_E}{\theta_H}$ be the fraction of H agents who match to E agents. Intuitively, we expect them to wait on average $\frac{c_1}{\theta_H p_E p_H}$[29]. Let $\frac{\theta_H - \theta_E}{\theta_H}$ be the fraction of H

[29]This relies on the fact that the waiting of H agents, conditional on the (a posteriori) information that they eventually match to an E agent, is roughly the same as the waiting when $\theta_H < \theta_E$ i.e. when almost all the H agents match to E agents. However, formally proving this result is challenging open problem that is not addressed by the theory we have developed.

agents who match to H agents. Intuitively, they will wait on average $\frac{c_2}{\theta_H p_H^2}$[30]. Therefore

$$w_H \approx \frac{\theta_E}{\theta_H}\frac{c_1}{\theta_H p_E p_H} + \frac{\theta_H - \theta_E}{\theta_H}\frac{c_2}{\theta_H p_H^2} = \frac{c_2}{\theta_H p_H^2} + \frac{\theta_E}{\theta_H^2 p_H}\left(\frac{c_1}{p_E} - \frac{c_2}{p_H}\right).$$

The first order condition[31] yields that the maximum waiting is attained for $\theta_H^* = 2\theta_E\left(1 - \frac{c_1 p_H}{c_2 p_E}\right)$. When p_H is much smaller than p_E, the constants have little influence and the maximum waiting is obtained for $\theta_H \approx \theta_E$.

An interesting question is under which conditions this non-monotonous behavior exists. Note that the slope is maximal for $\theta_H = \theta_E$. Therefore, waiting increases with thickness if and only if $\frac{\partial w}{\partial \theta_H}(\theta_H = \theta_E) > 0$. This condition is satisfied whenever $p_H < \frac{c_2 p_E}{2c_1}$. [32]

So when matches are formed through chains, thickening the market by increasing either θ_E and θ_H always reduces waiting times. Under bilateral matching, however, two phenomena arise. If possible, increasing the arrival of easy-to-match agents to the level where $\theta_E \geq \theta_H$ will significantly reduce waiting times of H agents. Otherwise, when $\theta_H >> \theta_E$, one can reduce average waiting times by increasing the arrivals of hard-to-match agents to the market.

We remark here that the average waiting time of E agents is orders of magnitude lower than that of H agents (we report these average waiting times in Appendix G).

4.2 Departures

In our model every agent leaves the market matched. We briefly discuss how the results change when departure rates are incorporated into the model, allowing agents to leave the market unmatched. We argue heuristically that minimizing the average waiting time in our model with no departures is closely related to minimizing the likelihood that an agent will depart before being matched in a model with departures. Roughly speaking the idea is that minimizing the expected waiting time in our model is the same as minimizing the expected number of agents in the market (by Little's law), which in turn reduces the rate of agents that depart the market without being matched.[33]

Consider a dynamic matching market with $\theta < 1$ and suppose that at each period every

[30] Because we are now computing the waiting of H agents conditional on being matched to another H, it is possible to show that $\frac{c_2}{\theta_H p_H^2}$ is indeed an upper bound on their waiting time. Again, proving tightness is a difficult open question.

[31] $\frac{\partial w}{\partial \theta_H} = \frac{2\theta_E}{\theta_H^3 p_H}\left(\frac{c_2}{p_H} - \frac{c_1}{p_E}\right) - \frac{c_2}{\theta_H^2 p_H^2}$

[32] Intuitively, the difference between p_H and p_E measures how "attractive" E agents are compared to H agents. If they are too close, then the increase in potential H matches outweighs the decrease in potential E matches due to competition.

[33] This relation is formalized by Anderson et al. (2013) in a model with only hard-to-match agents.

unmatched H agent departs the market independently with probability q_H.[34] In steady-state the equilibrium between arrivals and departure rates yields

$$\theta = \mu + q_H n_H, \tag{6}$$

where n_H is the expected number of H agents in the market and μ is the expected match rate. By Little's law, the average total waiting time for H agents (over both matched and unmatched agents) is $w_H = \frac{n_H}{\theta}$, implying that in steady-state

$$w_H = \frac{1 - \mu/\theta}{q_H} = O(1/q_H). \tag{7}$$

Therefore reducing agents' average total waiting times is aligned with reducing departures.[35]

We can now discuss the effects different departures rates have on our results. Hard-to-match agents' overall average waiting time (both matched and unmatched) will remain the same as in the original model under myopic policies if their departure rate scales similar to their average waiting time in the model without departures. Under this regime the only difference is that a constant fraction of hard-to-match agents will leave the market unmatched.

Recall that in our original model agents' average waiting time is $\Theta(1/p_H)$ both under the *ChainMatch* policy (for any $\theta < 1$) and under the *BilateralMatch* policy for $\theta < 0.5$ (note that in these settings, we need to have $q_H = \Theta(p_H)$ in order to insure that the likelihood that an agent will depart before being matched is bounded away from 0 and 1).

An interesting effect happens when $\theta > 0.5$ and agents match bilaterally. Note that under the *BilateralMatch* policy, the average number of H agents in the market is given by $\min\{\Theta(1/p_H^2), O(1/q_H)\}$. When the departure rate is $q_H = o(p_H^2)$ (but at least the order of p_H),[36] the market behaves essentially as a bipartite market: since the probability that two H agents match each other is very small, almost all H agents are either matched with E agents, or depart before they have a chance to match with other H agents.

[34]Since easy-to-match agents rarely wait to be matched, we simplify the discussion by ignoring their departure rate.

[35]Analyzing the tradeoff between match rates and the waiting times for matched agents remains an interesting open question.

[36]Note that $q_H = \Theta(p_h^2)$ scales similar to the waiting time in the model without departures and any other scaling leads to trivial probabilities for departing prior to being matched.

4.3 Two-sided matching markets with departures

Many matching markets are naturally two-sided and we discuss how to extend our analysis to such markets with departure rates. Consider a two-sided market in which every period one agent arrives to the market. Denote the two sides by S and L and let $\theta > 0.5$ be the probability that an arriving agent belongs to L. Since agents on the long side will accumulate (they cannot match with each other), it is natural to incorporate a small departure rate q_L for L agents which will allow the system to reach a steady-state.

Denote by p_{match} the probability that a given a pair of agents, one from each side, can form a bilateral match. One can show that under the *BilateralMatch* policy, the overall average waiting time for L agents (both matched and not) is $\Theta(1/q_L)$, independently of p_{match}. The idea is that for all p_{match} and q_L, a significant fraction of L agents will depart the market without being matched and these agents' waiting time is of order $\frac{1}{q_L}$.

Let N_L be the number of agents in the market at a given time and suppose $N_L = \Theta(1/q_L)$. Then an incoming S agent matches immediately with probability $1 - (1 - p_{match})^{N_L} \approx 1 - e^{-p_{match}/q_L} + o(p_{match}/q_L)$.[37] So for q_L that is significantly smaller than p_{match}, S agents match upon arrival with high probability and a fraction $\frac{1-\theta}{\theta}$ of L agents leave the market matched.[38] One interesting implication is that a chain in a two-sided market would not lead to better waiting times than bilateral matchings since agents on the short side match immediately upon arrival and agents on the long side cannot match each other.

Observe that the two-sided model has striking similarities to the one-sided model we studied earlier. In both models, there is a parameter which allows the long side to self-regulate: q_L through departures in two-sided markets, or p_H^2 for H to H matches in one-sided markets. When $\theta > 0.5$, this parameter determines the H agents' waiting times. We therefore believe that similar results to our one-sided model will hold in two-sided markets.[39]

[37]For example when $q_L < p_{match}/5$, the chance that an S agent matches immediately is approximately $1 - e^{-5} \approx 99.4\%$.

[38]Thus when q_L is small, departure rates for S agents may be ignored as they do not accumulate in the market. Note, however, that whether agents are on the long or short side depends on departure and arrival rates of both sides.

[39]One interesting application is a sub-market in kidney exchange with O-A (O patients and A donors) and A-O incompatible pairs. A-O pairs are on the short side in this market since many A patients are compatible to their live O intended donor and thus do not enter this market. A-O pairs in kidney exchange are likely to be much more sensitized than patients in O-A pairs since these A-O pairs must be tissue type incompatible. Our findings show that pairs on the short side are easy-to-match in practice despite their patients' high sensitivity while O-A pairs are under-demanded and will find it difficult to match.

4.4 Decentralized matching

Consider a decentralized version of our model, in which agents that join the market form matches selfishly. Note that since preferences are based on compatibility, agents will behave myopically in equilibrium. Therefore, with some caution, one may apply our findings regarding the average waiting times under myopic matching policies to decentralized markets.

Suppose first that only bilateral matches can be formed and $p_E = 1$. Since the average waiting time of H agents is approximately the same under *BilateralMatch(H)* (Theorem 1) and *BilateralMatch(E)* (Proposition 1), their average waiting time in a decentralized market is also the same. One shortcoming of this conclusion is that in a decentralized market agents are likely to break ties arbitrarily among compatible matches without assigning priority to either type. We believe that breaking ties randomly will not change this conclusion.

Figure 4 reports simulation results for the *BilateralMatch* policy when ties are broken either in favor of E agents or in favor of H agents. While the results are given only for $\theta = 0.4$ and $\theta = 0.6$, we observed a small gap in the average waiting times for $\theta < 0.5$, and otherwise there is essentially no gap. This is intuitive since when $\theta > 0.5$ almost all E agents match to H agents upon arrival.

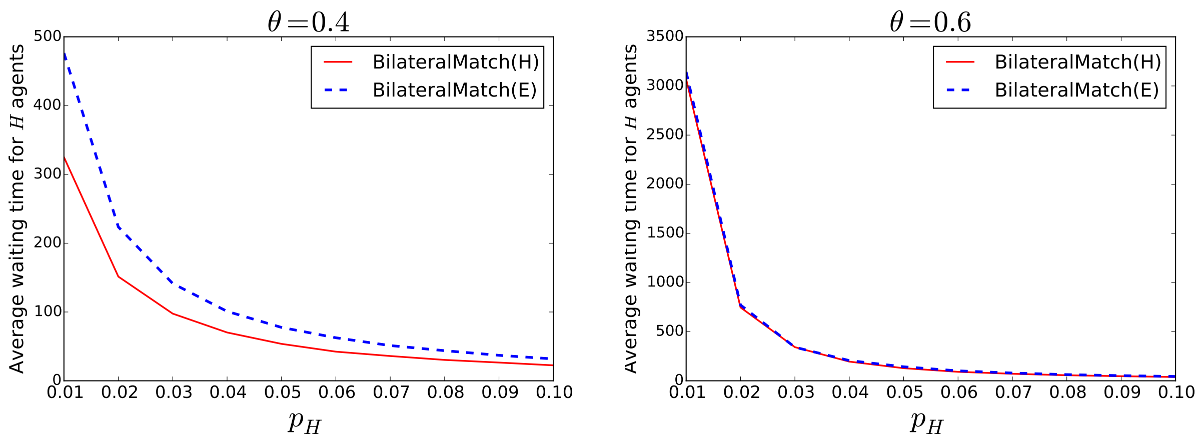

Figure 4: Average waiting times for H agents under *BilateralMatch* policies with parameters $p_H = 0.02$ and $p_E = 0.8$. We compare policies that either break ties in favor of E agents or in favor of H agents.

When matches are formed indirectly through a chain, Theorem 2 provides the equilibrium average waiting time of hard-to-match agents in a decentralized market when $p_E = 1$. Figure 5 plots simulation results that compare the *ChainMatch* policy which break ties in favor of H agents and a similar policy that break ties in favor of E agents. Note that there is essentially no difference, since E agents never remain in the market.

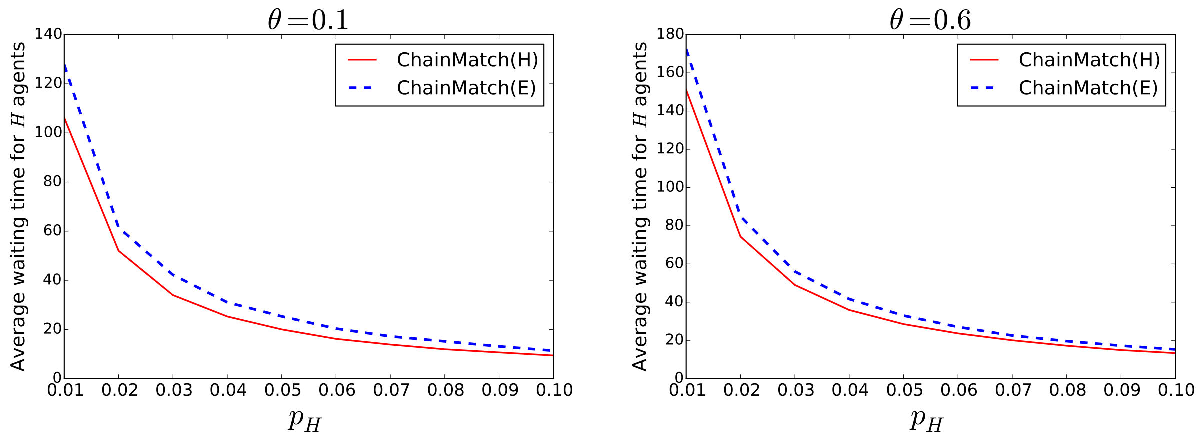

Figure 5: Average waiting times for H agents under *ChainMatch* policies with parameters $p_H = 0.02$ and $p_E = 0.8$. We compare policies that either break ties in favor of E agents or in favor of H agents.

5 Conclusion

We studied a dynamic matching market and analyzed how the thickness and the forms of matchings (bilateral or through a chain) affect efficiency. We identified that the balance between hard and easy-to-match agents in the market plays a crucial role on the desired matching technology. In particular, matching through chains (even a single one) leads to significantly better waiting times than bilateral matching *only* when hard-to-match agents arrive more frequently to the market.[40] Moreover there is almost no harm from adopting myopic matching policies.

Our work has several implications. First, in a decentralized market, the equilibrium in which agents attempt to match as soon as possible, is asymptotically efficient. Second, the lack of thickness cannot be compensated by better matching mechanisms and therefore the only way to thicken the market fruitfully is by attracting more agents. Our findings depend strongly on the assumption that agents preferences are based on compatibility. Indeed Baccara et al. (2015) and Doval (2014) find that with more preference structure optimal matching policies are not myopic.

Our results imply that agents' average waiting time depends crucially on whether they are on the long or short side of the market. In two-sided markets, agents on the short side of the market will match quickly even if their a priori chance to be matched by an arbitrary agent is very small. Therefore, while chains play a crucial role in one-sided markets, they

[40]Our findings hold for a single chain and one can further reduce waiting times by having multiple chains. We believe, however, that any constant number of chains will have no asymptotic advantage over a single chain.

have almost no benefit over bilateral matchings in two-sided markets.

This work raises several questions. Thickening the marketplace, by having more easy-to-match agents naturally increases the match rate. However, this rate increase depends on the matching technology and the composition of the market and is highly non-linear. Therefore the planner should be aware of the composition of the market in order to use the right matching mechanism. Furthermore, understanding the positive externality provided by easy-to-match agents is a key step in providing incentives that will attract them to the marketplace[41] Finally, what forms of exchanges will arise in a decentralized market is a question that relates to research on the origin of money (Jones (1976) and Kiyotaki and Wright (1989)).[42]

References

I. Adan and G. Weiss. Exact FCFS Matching Rates for Two Infinite Multitype Sequences. *Operations Research*, 60(2):475–489, 2012.

M. Akbarpour, S. Li, and S. Oveis Gharan. Dynamic matching market design. mimeo, 2014.

R. Anderson, I. Ashlagi, Y. Kanoria, and D. Gamarnik. Efficient dynamic barter exchange. mimeo, 2013.

R. Anderson, I. Ashlagi, D. Gamarnik, M. Rees, A.E. Roth, T. Sönmez, and M.U. Ünver. Kidney exchange and the alliance for paired donation: Operations research changes the way kidneys are transplanted. *Interfaces*, 2015.

I. Ashlagi, D. Gamarnik, Michael A. Rees, and A. E. Roth. The need for (long) NEAD chains. Working paper, 2012.

I. Ashlagi, P. Jaillet, and V.H. Manshadi. Kidney exchange in dynamic sparse heterogenous pools. *arXiv preprint arXiv:1301.3509*, 2013.

Itai Ashlagi and Alvin E Roth. Free riding and participation in large scale, multi-hospital kidney exchange. *Theoretical Economics*, 9(3):817–863, 2014.

M. Baccara, S. Lee, and L. Yariv. Optimal dynamic matching. mimeo, 2015.

[41] For instance, hospitals in the U.S. often conduct internal exchanges with easy-to-match pairs rather than enrolling them into kidney exchange clearinghouses (Roth et al. (2007); Ashlagi and Roth (2014)).

[42] These papers study how patterns of trade, such as direct barter and using currency, arise in equilibrium.

F. Bloch and D. Cantala. Dynamic allocation of objects to queuing agents: The discrete model. 2014.

E.B. Budish, P. Cramton, and J.J Shim. The high-frequency trading arms race: Frequent batch auctions as a market design response. *Fama-Miller Working Paper*, pages 14–03, 2013.

R. Caldentey, E. H. Kaplan, and G. Weiss. Fcfs infinite bipartite matching of servers and customers. *Adv. Appl. Prob*, 41(3):695–730, 2009.

S. Chakravarty and T.R. Kaplan. Optimal allocation without transfer payments. *Games and Economic Behavior*, 77(1):1–20, 2013.

J. P. Dickerson, A. D. Procaccia, and T. Sandholm. Optimizing Kidney Exchange with Transplant Chains: Theory and Reality. *Proc of the eleventh international conference on autonomous agents and multiagent systems*, 2012.

Y. Ding, D. Ge, S. He, and C. T. Ryan. A non-asymptotic approach to analyzing kidney exchange graphs. In *Proceedings of the Sixteenth ACM Conference on Economics and Computation*, pages 257–258. ACM, 2015.

L. Doval. A theory of stability in dynamic matching markets. Technical report, mimeo, 2014.

J. Feldman, A. Mehta, V. S. Mirrokni, and S. Muthukrishnan. Online stochastic matching: Beating 1-1/e. In *Proceedings of the 50th Annual IEEE Symposium on Foundations of Computer Science (FOCS)*, pages 117–126, 2009.

D. Fershtman and A. Pavan. Re-matching, experimentation, and cross-subsidization. mimeo, 2015.

F.G. Foster. On the stochastic matrices associated with certain queuing processes. *The Annals of Mathematical Statistics*, pages 355–360, 1953.

G. Goel and A. Mehta. Online budgeted matching in random input models with applications to adwords. In *Proceedings of the nineteenth annual ACM-SIAM symposium on Discrete algorithms (SODA)*, pages 982–991, 2008.

I. Gurvich and A. Ward. On the dynamic control of matching queues. Working paper, 2012.

H.C. Hoppe, B. Moldovanu, and A. Sela. The theory of assortative matching based on costly signals. *The Review of Economic Studies*, 76(1):253–281, 2009.

P. Jaillet and X. Lu. Online stochastic matching: New algorithms with better bounds. *Mathematics of Operations Research*, 2013. Forthcoming.

R.A. Jones. The origin and development of media of exchange. *The Journal of Political Economy*, pages 757–775, 1976.

S.V. Kadam and M.H. Kotowski. Multi-period matching. Technical report, Mimeo, 2014.

R. M Karp, U. V. Vazirani, and V. V. Vazirani. An optimal algorithm for on-line bipartite matching. In *Proceedings of the twenty-second annual ACM symposium on Theory of computing (STOC)*, pages 352–358, 1990.

N. Kiyotaki and R. Wright. On money as a medium of exchange. *The Journal of Political Economy*, pages 927–954, 1989.

J. Leshno. Dynamic Matching in Overloaded Waiting Lists. Working paper, 2014.

V. H. Manshadi, S. Oveis-Gharan, and A. Saberi. Online stochastic matching: online actions based on offline statistics. In *Proceedings of the Twenty-Second Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 1285–1294, 2011.

H. Mendelson. Market behavior in a clearing house. *Econometrica: Journal of the Econometric Society*, pages 1505–1524, 1982.

R. Rogerson, R. Shimer, and R. Wright. Search-theoretic models of the labor market: A survey. *Journal of Economic Literature*, 43(4):959–988, 2005.

A. E. Roth, T. Sönmez, and M. U. Ünver. Efficient kidney exchange: coincidence of wants in markets with compatibility-based preferences. *American Economic Review*, 97:828–851, 2007.

M. U. Ünver. Dynamic Kidney Exchange. *Review of Economic Studies*, 77(1):372–414, 2010.

S.A. Zenios. Modeling the transplant waiting list: A queueing model with reneging. *Queueing systems*, 31(3-4):239–251, 1999.

A Tightness of the theoretical bounds

In the next following subsections, we provide simulation results under the myopic *BilateralMatch* and the *ChainMatch* policies and examine how tight are the theoretical upper bounds on the waiting times we derived.

A.1 Bilateral matching

Theorem 1 states that under the *BilateralMatch* policy, there exist constants A_θ and B_θ such that (i) $w_H \leq \frac{A_\theta}{p_E p_H}$ if $\theta < 1/2$, and (ii) $w_H \leq \frac{B_\theta}{p_H^2}$ if $\theta \geq 1/2$. The proof (Appendix D) yields the following numerical values for these constants:

$$A_\theta = \frac{\ln\left(\frac{1-\theta}{1-2\theta}\right)}{\theta} + \frac{2\theta}{\theta(1-2\theta)\ln\left(\frac{1-\theta}{1-2\theta}\right)} + 1,$$

$$B_\theta = \frac{\ln(2\theta) + 2/\ln(2\theta)}{\theta}.$$

Simulations show that the upper bounds with these constants are not very tight and we believe that this is mainly due to an additional term needed for technical reasons. We conjecture that the average waiting time of hard-to-match agents in the market at steady-state are bounded according to (8) and (9).

- If $\theta < 1/2$, then

$$w_H \leq \frac{\ln\left(\frac{1-\theta}{1-2\theta}\right)}{\theta p_E p_H}. \tag{8}$$

- If $\theta \geq 1/2$, then

$$w_H \leq \frac{\ln(2\theta)}{\theta p_H^2}. \tag{9}$$

Simulations show that these conjectured upper bounds are indeed tight for values of θ that are sufficiently far enough from 0.5. However, as θ approaches 0.5, these bounds are tight only for sufficiently small p_H. Figure 6 shows simulation results alongside the conjectured bounds for $\theta = 0.1$ and $\theta = 0.6$.

A.2 Matching through chains

Theorem 2 states that under the *ChainMatch* policy there exists a constant K_θ such that (i) $w_H \leq \frac{K_\theta}{\theta p_H}$ when $\theta < 1$, and (ii) $w_H \leq \frac{2\ln(1/p_H)}{\theta p_H}$ when $\theta = 1$.

The proof (Appendix E) yields the following numerical value for K_θ:

$$K_\theta = \frac{2}{(1-p_H)} + \frac{\theta(1-p_H)}{\theta p_H + (1-\theta)} + o(1).$$

Simulations show that the upper bounds with this constant are not tight and we conjecture that the tighter inequalities (10) and (11) hold true.

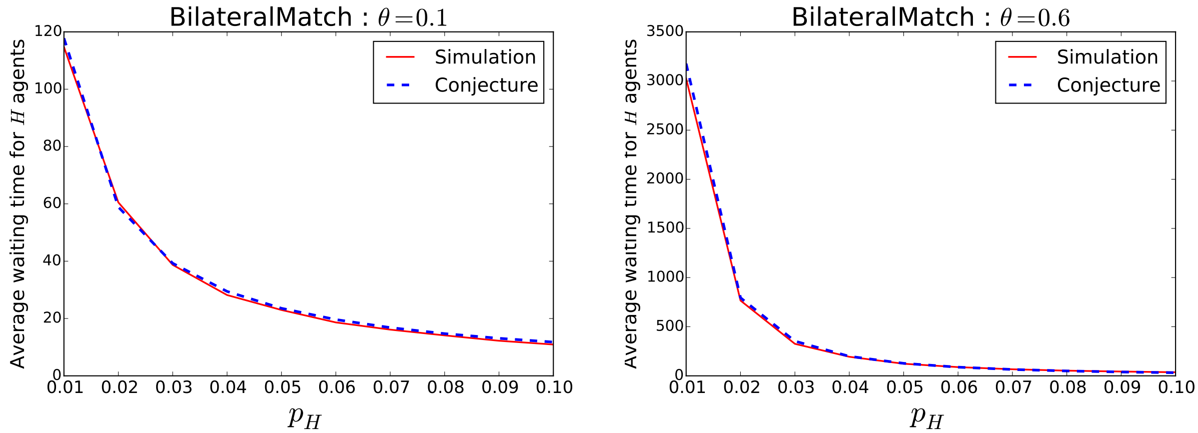

Figure 6: Numerical results (solid red) and the conjecture bound (dashed blue) for $\theta = 0.1$ and $\theta = 0.6$ under the *BilateralMatch* policy.

- If $\theta < 1$,

$$w_H \leq \frac{1}{2\theta p_H}\left(\frac{\theta}{(1-p_H)} + \frac{\theta(1-p_H)}{\theta p_H + (1-\theta)}\right). \tag{10}$$

- If $\theta = 1$,

$$w_H \leq \frac{\ln(1/p_H)}{\theta p_H}. \tag{11}$$

Figure 7 shows simulation results under the *ChainMatch* policy alongside the conjectured bounds. We note that the conjecture is tight when $\theta = 1$ or $\theta < 1/2$. However, for $1/2 < \theta < 1$, it seems that our conjectured bounds are not perfectly tight.

B Proof Notations

Throughout the appendices below we denote by $\mathbf{B}(\lambda)$ a random variable drawn from a Bernoulli distribution with parameter λ. Similarly, $\mathbf{Bin}(n, p)$ corresponds to a random variable drawn from a Binomial distribution, with parameters n and p.

All the proofs for the existence of steady-state of the Markov chains that we consider rely on the following result:

Proposition 5 (Foster (1953))**.** *Suppose $\{X_k\}$ is a discrete time, irreducible Markov Chain on a countable state space $\mathcal{X}$. If there exists a function $V : \mathcal{X} \mapsto R$, > 0, and a finite set $B \subset \mathcal{X}$ such that for all $x \in B$,*

$$\mathbb{E}_x[V(X_1) - V(X_0)] < \infty \tag{12}$$

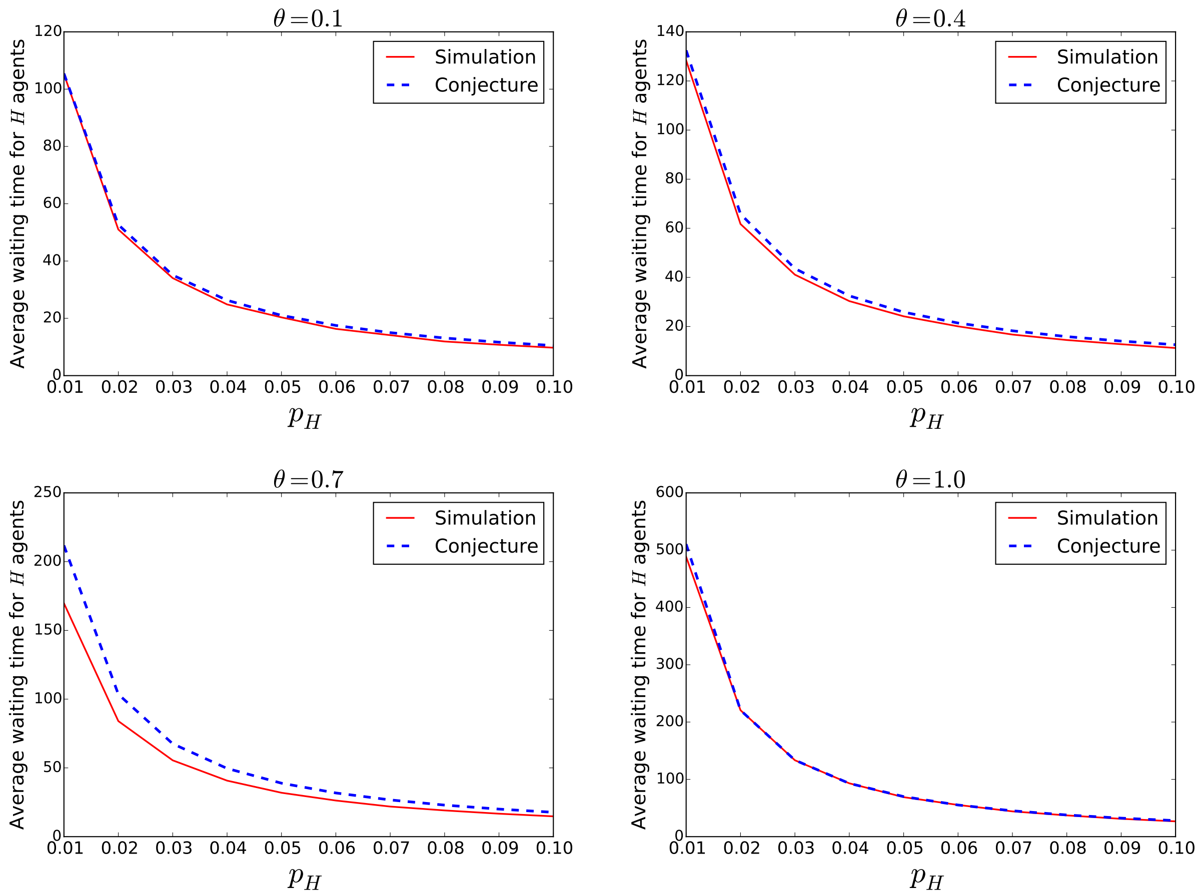

Figure 7: Numerical results (solid red) and the conjecture bound (dashed blue) under the *ChainMatch* policy.

and for all $x \in X \backslash B$,

$$\mathbb{E}_x[V(X_1) - V(X_0)] < -\gamma, \tag{13}$$

then X_k *is positive recurrent.*

C Lower bounds: proofs of Propositions 3 and 4

We prove the lower bound results stated in Section 3.3.

Proposition 3. *For any* $p_E \leq 1$ *and under any matching policy that reaches steady-state, for any* $\theta \in (0, 1)$ *and any* $p_H > 0$*, there exists a constant* c_θ *such that* $w_H + w_E \geq \frac{c_\theta}{p_H}$.

Proof of Proposition 3. The main intuition behind this proof is the following: Suppose that the pool size is too small, then an arriving agent has to wait a "long" time to obtain at least one incoming edge. This long waiting time contradicts the small pool size (with Little's law).

Let n_H and n_E be respectively the expected numbers of H and E agents in the market in steady-state, and let $n = n_H + n_E$. Little's law implies that $w_H = n_H/\theta$ and $w_E = n_E/(1-\theta)$. Therefore, it is enough to prove that there exists a constant k_θ such that $n \geq k_\theta/p_H$ (we then choose $c_\theta = \frac{k_\theta}{\max(\theta, 1-\theta)}$).

Let k_θ be a constant to be defined later. Assume for contradiction that there exists p_H such that $n < k_\theta/p_H$. Let i be an H agent entering the market at steady-state, and let W_i be her waiting time until she is matched. Let $\mathcal{V}$ be the set of agents in the market when agent i arrives. Note that $\mathbb{E}[|\mathcal{V}|] = n \leq k_\theta/p_H$. Define the event $E_1 = \{|\mathcal{V}| \leq 3n/\theta\}$. By Markov's inequality and Little's law, $\mathbb{P}[E_1] \geq 1 - \frac{\mathbb{E}[|\mathcal{V}|]\theta}{3n} \geq 1 - \theta/3$.

Let $\mathcal{A}$ be the first $3n/\theta$ arrivals after i, and let E_2 be the event that at least one agent from $\mathcal{V} \cup \mathcal{A}$ has an outgoing edge towards i. We have

$$\mathbb{P}[E_2] = \mathbb{P}[\mathbf{Bin}(|\mathcal{V}| + |\mathcal{A}|, p_H) \geq 1].^{43}$$

Therefore we get:

$$\mathbb{P}[E_2|E_1] \leq \mathbb{P}[\mathbf{Bin}(6n/\theta, p_H) \geq 1] \leq \mathbb{P}[\mathrm{Bin}(6k_\theta/\theta p_H, p_H) \geq 1] \leq 6k_\theta/\theta.$$

Where the first inequality derives from the definition of E_1, the second uses the fact that $n \leq k_\theta/p_H$ and the third is Markov's inequality.

We now use the fact that if i doesn't have any edge from either $\mathcal{V}$ or $\mathcal{A}$, then she must wait longer than $3n/\theta$ time steps. Together with the bounds derived above, we get

$$w_H = \mathbb{E}[W_i] \geq \frac{3n}{\theta}\mathbb{P}[E_2^c] \geq \frac{3n}{\theta}\mathbb{P}[E_2^c|E_1]\mathbb{P}[E_1] \geq \frac{3n}{\theta}(1-6k_\theta/\theta)(1-\theta/3) \geq \frac{3n}{\theta}(1-6k_\theta/\theta)(2/3).$$

Thus we get:

$$n \geq n_H = w_H\theta \geq 2n(1 - \frac{6k_\theta}{\theta}).$$

Therefore for $k_\theta = \frac{\theta}{24}$, we obtain a contradiction.

□

Proposition 4. *Let $\theta > 1/2$. For any bilateral matching policy that reaches steady-state, under a mild regular condition, there exists a positive constant c'_θ such that $w_H \geq \frac{c'_\theta}{p_H^2}$.*[44]

[43] Note that we abuse the notation of **Bin(n,p)** by allowing its parameters to be random variables. In this case, conditional on the event $|\mathcal{V}|+|\mathcal{A}| = k$, the random variable $\mathbf{Bin}(|\mathcal{V}|+|\mathcal{A}|, p)$ has a binomial distribution with parameters k and p.

[44] The regular condition states that for every H agent that arrives at steady-state the probability to eventually being matching to an E agent is the same.

Proof of Proposition 4. Consider a bilateral matching policy $\mathcal{P}$, i.e. a policy that can only match pairs of agents. We allow any possibility for $\mathcal{P}$ regarding the choice of whether to match agents and which agents to match. The main idea is to show that a significant fraction of H agents have to match to each other as a necessary condition for steady-state. Similarly to the proof of Proposition 3, we focus on finding a lower bound on the number of H agents in steady-state.

Let c_θ be a constant to be defined later, and assume for contradiction that there exists p_H such that $n_H \leq c_\theta/p_H^2$. Let us consider an H agent entering the system at steady-state and let W_i be its waiting time. Let $\mathcal{V}_\mathcal{H}$ be the set of H agents in the pool upon its arrival, which implies $\mathbb{E}[|\mathcal{V}_H|] = n_H \leq c_\theta/p_H^2$. We define the event $E_1 = \{|\mathcal{V}_H| \leq u_\theta^1 n_H\}$ where u_θ^1 is a function of θ to be defined later. By Markov's inequality, $\mathbb{P}[E_1] \geq 1 - \frac{1}{u_\theta^1}$. Let $\mathcal{A}_H$ be the next $u_\theta^2 n_H$ arrivals of H agents, where u_θ^2 is a function of θ to be defined later. We consider different matching scenarios.

1. The agent eventually gets matched to an E agent. We call this event E_2.
2. The agent is matched to an agent from $\mathcal{V}_H \cup \mathcal{A}_H$. We call this event E_3.
3. The agent is matched to a later H agent outside $\mathcal{V}_H \cup \mathcal{A}_H$. We call this event E_4.

We obtain:

$$\mathbb{P}[E_3] \leq \mathbb{P}[\mathrm{Bin}(|\mathcal{V}_H| + |\mathcal{A}_H|, p_H^2) \geq 1], \tag{14}$$

We compute $\mathbb{P}[E_3]$ as follows:

$$\mathbb{P}[E_3|E_1] \leq \mathbb{P}[\mathrm{Bin}((u_\theta^1 + u_\theta^2)n_H, p_H^2) \geq 1] \leq \mathbb{P}[\mathrm{Bin}((u_\theta^1 + u_\theta^2)c_\theta/p_H^2, p_H^2) \geq 1] \leq (u_\theta^1 + u_\theta^2)c_\theta, \tag{15}$$

where we used the definition of E_1, the fact that $n_H \leq c/p_H^2$, and Markov's inequality.

Note that because the system is in steady-state, any H agent will be matched with probability 1. Furthermore, because the system has the *fairness* property, the probability of any H agent to eventually match an E is the same, which we denote p_{H-E}. Because the system is in steady-state, we have $\theta p_{H-E} = 1 - \theta$. Therefore:

$$\mathbb{P}[E_2] \leq \frac{1-\theta}{\theta}. \tag{16}$$

Putting everything together, we get:

$$\begin{aligned}\frac{n_H}{\theta} &= \mathbb{E}[W_i] \geq u_\theta^2 n_H \mathbb{P}[E_4] \geq u_\theta^2 n_H (\mathbb{P}[E_3^c] - \mathbb{P}[E_2]) \\ &\geq u_\theta^2 n_H \left(\mathbb{P}[E_3^c | E_1]\mathbb{P}[E_1] - \frac{1-\theta}{\theta} \right) \\ &\geq u_\theta^2 n_H \left((1 - (u_\theta^1 + u_\theta^2)c_\theta)(1 - \frac{1}{u_\theta^1}) - \frac{1-\theta}{\theta} \right),\end{aligned} \tag{17}$$

where we used Little's law, then the fact that the expected waiting time conditional on E_4 is greater than $u_\theta^2 n_H$, then the fact that $1 = \mathbb{P}[E_2] + \mathbb{P}[E_3] + \mathbb{P}[E_4]$, and (16).

Note that here it becomes clear that the assumption that $\theta > 1/2$ is necessary in order to obtain a contradiction (otherwise the right hand term is negative).

Here, because $\theta > 1/2$, we have: $\frac{1-\theta}{\theta} < 1$, and we can consider the mid-point $\alpha_\theta = \frac{1}{2}\left(\frac{1-\theta}{\theta} + 1\right)$. Therefore taking for instance u^1 such that $1 - \frac{1}{u_\theta^1} = \sqrt{\alpha_\theta}$, $u_\theta^2 = \frac{4}{2\theta - 1}$, and c_θ such that $(1 - (u_\theta^1 + u_\theta^2)c_\theta = \sqrt{\alpha_\theta}$, we have that

$$\begin{aligned}\frac{n_H}{\theta} &\geq u_\theta^2 n_H \left((1 - (u_\theta^1 + u_\theta^2)c_\theta)(1 - \frac{1}{u_\theta^1}) - \frac{1-\theta}{\theta} \right) \\ &= u_\theta^2 n_H \left(\sqrt{\alpha_\theta}\sqrt{\alpha_\theta} - \frac{1-\theta}{\theta} \right) \\ &= u_\theta^2 n_H \frac{1}{2} \left(1 - \frac{1-\theta}{\theta} \right) \\ &= \frac{4}{2\theta - 1} n_H \frac{2\theta - 1}{2\theta}. \\ &= \frac{2n_H}{\theta}.\end{aligned} \tag{18}$$

Therefore we obtain a contradiction. □

D Bilateral matching: proof of Theorem 1

Theorem 1. *Under BilateralMatch, the market reaches steady-state. Furthermore, there exist positive constants A_θ, B_θ, and C_θ such that:*

1. *If $\theta < 1/2$, then*

$$w_H \leq \frac{A_\theta}{p_E p_H}. \tag{1}$$

2. If $\theta \geq 1/2$, then

$$w_H \leq \frac{B_\theta}{p_H^2}. \tag{2}$$

3. For $0 \leq \theta < 1$

$$w_E \leq \frac{C_\theta}{p_E^2}. \tag{3}$$

The proof involves several steps, which we will state using intermediate lemmas for the important results. The outline of the proof is as follows:

- We represent *BilateralMatch* as a two dimensional stochastic process $(N_{H,t}, N_{E,t})$ where $N_{H,t}$ ($N_{E,t}$) is the number of H (E) agents in the system at time t. We prove that this process is a Markov chain, which admits a steady-state distribution$(N_{H,\infty}, N_{E,\infty})$. Cf. Lemma 1 in Appendix D.1.

- We define formally the waiting times w_H and w_E, and use Little's law to prove the results by focusing on the average steady-state number of agents n_H and n_E.

- We couple our Markov chain with an auximary 1-dimensional Markov chain M_t and show that $\mathbb{E}[N_{H,\infty}] \leq \mathbb{E}[M_\infty] + 1$. This allows us to focus on proving bounds on M_t.

- We provide three proofs using coupling techniques to obtain the bounds announced in equations (1), (2) and (3). Proofs can be found in appendix D.4 D.3 D.5

D.1 Markov chain representation

Observe that because the algorithm matches agents in a myopic way (i.e. matches are conducted whenever they are possible), there can never be two agents in the system that have the possibility to match to each other. Therefore, it is enough to only keep track of the number of H and E agents to completely characterize the system. At any given time t, $N_{H,t}$ ($N_{E,t}$) represents the number of H (E) agents in the pool.

Lemma 1. *For any $0 < p_H < 1$, $0 < p_L \leq 1$ and $0 < \theta \leq 1$, the number of agents in the pool under BilateralMatch $(N_{H,t}, N_{E,t})_{t\in\mathbb{N}}$ is a positive recurrent Markov chain and it converges to a steady-state distribution $(N_{H,\infty}, N_{E,\infty})$.*

The proof is defered to appendix D.6. This results allows us to apply Little's law to our system, which yields $w_H = \mathbb{E}[N_{H,\infty}]/\theta$ and $w_E = \mathbb{E}[N_{E,\infty}]/(1-\theta)$. In all the proof that follows, we will therefore focus on providing upper bounds for $\mathbb{E}[N_{H,\infty}]$ and $\mathbb{E}[N_{E,\infty}]$.

The structure of this Markov chain is shown in figure 8, and the transitions probabilities are in equation (28). Analyzing 2-dimension Markov chains can be very difficult, therefore we use a coupling technique to gain analytic tractability.

D.2 Coupling to a 1-dimensional Birth-Death process.

We now introduce an auxiliary Markov chain M_t defined by:

$$M_{t+1} = \begin{cases} (M_t + 1) \text{ with probability } \theta(1-p_H^2)^{M_t}, \\ (M_t - 1) \text{ with probability } \theta(1-(1-p_H^2)^{M_t}) + (1-\theta)(1-(1-p_E p_H)^{M_t}), \\ (M_t) \text{ with probability } (1-\theta)(1-p_E p_H)^{M_t}. \end{cases} \tag{19}$$

Remark. *This corresponds to the number of H agents in the pool under the following "simplified" algorithm: For each incoming agent v to the pool, match v to one of the H agents in the pool if such a match exists. If no match is found, add v to the pool if it is of type H, otherwise discard v. Note that we do not keep unmatched E agents, which is why we get a 1-dimensional Markov Chain, and why we expect the number of unmatched H agents to be larger in this auxiliary chain.*

The next lemma states that with an appropriate coupling, $N_{H,t}$ can be upper bounded by M_t. This allows us to focus on analyzing for M_t.

Lemma 2. *The Markov chain M_t defined in (19) reaches a steady-state distribution M_∞, and there exists a way to couple the Markov chains $(N_{H,t}, N_{H,t})$ and M_t such that for every $t > 0, N_{H,t} \leq M_t + 1$. Therefore $\mathbb{E}[N_{H,\infty}] \leq \mathbb{E}[M_\infty] + 1$.*

The proof of lemma 2 is defered to Appendix D.6.

D.3 Proof of of Theorem 1 part 1

Proof. In this regime there are more E agents than H agents, and therefore we expect most of the H agents to eventually get matched to E agents. Therefore, we couple M_t to a simple Birth-Death process that corresponds to the following algorithm: We never try to match arriving H agents (this is unlikely anyways because we never keep E agents in the system). We only try to match incoming E agents to a *fixed* number of H agents, equal to $\frac{\eta}{p_E p_H}$ agents for $\eta = \ln(\frac{1-\theta}{1-2\theta})$. This will lead to a new Markov chain S_t for which we can effectively compute the expected number of agents.

We can compute the transition states:

$$S_{t+1} = \begin{cases} (S_t + 1) \text{ with probability } \theta, \\ (S_t - 1) \text{ with probability } (1-\theta)(1-(1-p_E p_H)^{\frac{\eta}{p_E p_H}}), \\ (S_t) \text{ with probability } (1-\theta)(1-p_E p_H)^{\frac{\eta}{p_E p_H}}. \end{cases} \tag{20}$$

In the following lemma, we show that the chain S_t reaches steady-state and we provide a bound on the expected number of H agents in steady-state.

Lemma 3. *S_t reaches steady-state S_∞ and we have*

$$\mathbb{E}[S_\infty] = \frac{2\theta}{(1-2\theta)\ln\frac{1-\theta}{1-2\theta} p_E p_H} + o(1/p_H). \tag{21}$$

The proof can be found in Appendix D.6.3

We now wish to couple the random walk S_t with the Markov chain M_t defined in (19).

Lemma 4. *There exists a coupling of M_t and S_t such that for all $M_t > \frac{\eta}{p_E p_H}$ we have*

$$M_t - M_{t-1} \leq S_t - S_{t-1}.$$

The proof of Lemma 4 is deferred to the Appendix D.6, but the intuition is that if there are more than $\frac{\eta}{p_E p_H}$ agents in the system, then S_t grows faster than M_t because it only matches agents "up to" the first $\frac{\eta}{p_E p_H}$. Using this lemma, for all t let us denote $t^*(t) = \max\{u \in \mathbb{R} \text{ s.t. } M_{u-1} \leq \frac{\eta}{p_E p_H}\}$ we get

$$M_t \leq M_{t^*-1} + (S_t - S_{t^*-1}) \leq \frac{\eta}{p_E p_H} + S_t.$$

Therefore

$$\begin{aligned} \mathbb{E}[N_{H,\infty}] &\leq \mathbb{E}[M_\infty] + 1 \leq \frac{\eta}{p_E p_H} + \mathbb{E}[S_\infty] + 1 \\ &\leq \frac{\ln\left(\frac{1-\theta}{1-2\theta}\right)}{p_E p_H} + \frac{2\theta}{(1-2\theta)\ln\left(\frac{1-\theta}{1-2\theta}\right) p_E p_H} + 1. \\ &\leq \frac{A_\theta}{p_E p_H}, \end{aligned} \tag{22}$$

$\square$

where

$$A_\theta = \ln\left(\frac{1-\theta}{1-2\theta}\right) + \frac{2\theta}{(1-2\theta)\ln\left(\frac{1-\theta}{1-2\theta}\right)} + 1. \tag{23}$$

Remark. *We note that using the exact same proof it is possible to derive another bound, which in some cases is tighter. By letting* $\eta = \ln(\frac{1-\theta}{1-2\theta-\delta})$ *with* $0 < \delta < \theta$ *we get another constant:*

$$A_\theta = \ln\left(\frac{1-\theta}{1-2\theta-\delta}\right) + \frac{2\theta}{\delta + o(1)} + 1 = \ln\left(\frac{1-\theta}{1-2\theta-\delta}\right) + o(1).$$

In Section A.1, simulations show that $\ln\left(\frac{1-\theta}{1-2\theta}\right)/p_E p_H$ *seems to be a tight constant. However, proving tightness remains an open question.*

D.4 Proof of Theorem 1 part 2

Proof. The proof is similar to the previous case. Again, we couple the Markov chain M_t with a simpler random walk that is the result of the following algorithm: we only try to match incoming agents (E and H) to a "virtual" set of $\ln(2\theta)/p_H^2$ H agents. This leads to a random walk S'_t that provides an upper bound on M_t when the number of agents becomes large:

$$S'_{t+1} = \begin{cases} (S'_t + 1) \text{ with probability } \theta g(p_H), \\ (S'_t - 1) \text{ w.p. } \theta(1-g(p_H)) + (1-\theta)(1-f(p_H)) = 1 - \theta g(p_H) - (1-\theta) f(p_H), \\ (S'_t) \text{ w.p. } (1-\theta) f(p_H), \end{cases}$$

where

$$f(p_H) = (1 - p_H p_E)^{\ln(2\theta)/p_H^2} \quad \text{and} \quad g(p_H) = (1 - p_H^2)^{\ln(2\theta)/p_H^2}.$$

Lemma 5. *For* p_H *small enough, we get* $\mathbb{E}[S'_\infty] \leq \frac{2}{\ln(2\theta)p_H^2}$.

Lemma 6. *There exists a coupling of the random walks* S_t *and* M_t *as defined in (19) such that if* $M_t \geq \frac{\ln(2\theta)}{p_H^2}$ *we get:*

$$M_{t+1} - M_t \leq S'_{t+1} - S'_t.$$

The proof of Lemma 6 is deferred to the Appendix D.6. Again the intuition is that if there are more than $\frac{\ln(2\theta)}{p_H^2}$ agents in the system, then S_t grows faster than M_t because it only matches agents "up to" the first $\frac{\ln(2\theta)}{p_H^2}$. We get that for every t,

$$M_t \leq \frac{\ln(2\theta)}{p_H^2} + S'_t. \tag{24}$$

Therefore

$$\begin{aligned}\mathbb{E}[N_{H,\infty}] &\leq \frac{\ln(2\theta)}{p_H^2} + \mathbb{E}[S'_\infty] + 1 \\ &\leq \frac{\ln(2\theta) + 2/\ln(2\theta)}{p_H^2}. \\ &\leq \frac{B_\theta}{p_H^2},\end{aligned} \tag{25}$$

where $B_\theta = \ln(2\theta) + 2/\ln(2\theta)$. $\square$

Remark. *Similarly to the proof of case (2), one can derive a similar bound by considering* $n = \frac{\ln(\frac{2\theta}{1-\delta})}{p_H^2}$ *for* $0 < \delta$*. This leads to:*

$$\mathbb{E}[N_{H,\infty}] \leq \frac{\ln(\frac{2\theta}{1-\delta})}{p_H^2} + \frac{(1-\delta)/2}{\delta + o(p_H^2)}.$$

In Section A.1, simulations show that $\ln(2\theta)/(\theta p_H^2)$ *seems to be a tight constant. However here again, proving tightness remains an open question.*

D.5 Proof of Theorem 1 part 3

Proof. This proof follows the same procedure as the proof for H agents. To study the E agents, we study the Markov chain T_t obtained with a simplified algorithm, that both disregards H agents when looking for matches, and only considers the first $\frac{\ln(4)}{p_E^2}$ E agents. Assuming that $T_t \geq \frac{\ln(4)}{p_E}$ and using the above notations, we get:

$$T_{t+1} = \begin{cases} (T_t + 1) \text{ with probability } (1-\theta)(1-p_E^2)^{\frac{\ln(4)}{p_E^2}}, \\ (T_t - 1) \text{ when } \text{ with probability } (1-\theta)(1-(1-p_E^2)^{\frac{\ln(4)}{p_E^2}}). \end{cases}$$

Lemma 7. *Using a coupling argument, we show that if* $N_{E,t} \geq \frac{\ln(4)}{p_E^2}$ *then* $N_{E,t+1} - N_{E,t} \leq T_{t+1} - T_t$*, and therefore* $N_{E,t} \leq T_t + \frac{\ln(4)}{p_E^2}$.

Using the notation $x = \frac{(1-p_E^2)^{\ln(4)/p_E^2}}{1-(1-p_E^2)^{\ln(4)/p_E^2}}$, the expected value of T_∞ is given by,

$$\mathbb{E}[T_\infty] = \frac{x}{1-x} = \frac{(1-p_E^2)^{n/p_E^2}}{1-2(1-p_E^2)^{n/p_E^2}} \leq \frac{e^{-\ln(4}}{1-2e^{-\ln(4)}} = \frac{1}{2}, \tag{26}$$

and

$$\mathbb{E}[N_{E,\infty}] \leq \frac{\ln(4)}{p_E^2} + 1/2.$$

Therefore taking $C_\theta = \ln(4)$. And this concludes the proof of Theorem 1 □

D.6 Proofs of the technical lemmas

D.6.1 Proof of Lemma 1

We provides a formal proof that the random process $(N_{H,t}, N_{E,t})$ is indeed a positive recurrent Markov chain, and therefore has a steady-state distribution.

Lemma 1. *For any $0 < p_H < 1$, $0 < p_L \leq 1$ and $0 < \theta \leq 1$, the number of agents in the pool under BilateralMatch $(N_{H,t}, N_{E,t})_{t\in\mathbb{N}}$ is a positive recurrent Markov chain and it converges to a steady-state distribution $(N_{H,\infty}, N_{E,\infty})$.*

Proof. Because the algorithm is online and only allows for 2-ways, an incoming agent v will either be matched immediately or join the pool. In the latter case, v will never match in a 2-way to an agent arrived before v. Therefore, the system does not add any memory in terms of edge realizations between agents of the pool. Thus, keeping track of only the number of agents $(N_{H,t}, N_{E,t})_{t\in\mathbb{N}}$ is enough to characterize the system. For any time t, we consider the following events:

- $A_H(t)$ is the event of an H agent arriving to the pool. We have $A_H(t) = \mathbf{B}(\theta)$.
- $A_E(t)$ is the event of an E agent arriving, correlated with $A_H(t)$ so that the two events are mutually exclusive. We have $A_E(t) = 1 - A_H(t)$.
- $H_H(t)$ is the event that an incoming H agent matches an H agent already present in the pool. We have $H_H(t) = \mathbf{B}(1-(1-p_H^2)^{N_{H,t}})$.
- $H_E(t)$ is the event that an incoming H matches an E. We have $H_E(t) = (1 - H_H(t))\mathbf{B}(1 - (1-p_H p_E)^{N_{E,t}})$.
- $H_\emptyset(t)$ is the event that an incoming H doesn't match immediately, and is therefore added to the pool. We have $H_\emptyset(t) = 1 - H_H(t) - H_E(t)$.

- $E_H(t)$ is the event that an incoming E matches an H. We have $E_H(t) = \mathbf{B}(1-(1-p_Ep_H)^{N_{H,t}})$.

- $E_E(t)$ is the event that an incoming E matches an E. We have $E_E(t) = (1-E_H(t))\mathbf{B}(1-(1-p_E^2)^{N_{E,t}})$.

- $E_\emptyset(t)$ the event that an incoming E doesn't match immediately. We have $E_\emptyset(t) = 1 - E_H(t) - E_E(t)$.

We obtain the following equations:

$$\begin{cases} N_{H,t+1} = N_{H,t} + A_H(t)(H_\emptyset(t) - H_H(t)) - A_E(t)E_H(t), \\ N_{E,t+1} = N_{E,t} - A_H(t)H_E(t) + A_E(t)(E_\emptyset(t) - E_E(t)). \end{cases} \tag{27}$$

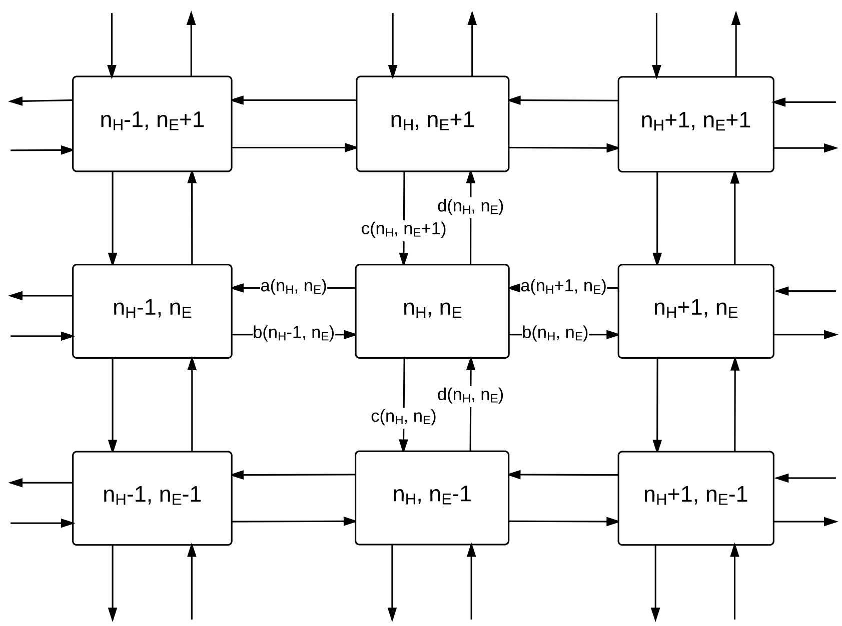

Figure 8: Transitions the Markov chain representation of the system.

In Figure 8, we give the transitions for the Markov chain that represents our dynamic system. We use the notations:

$$\begin{aligned}
a(n_H, n_E) &= \theta(1-(1-p_H^2)^{N_{H,t}}) + (1-\theta)(1-(1-p_E p_H)^{N_{H,t}}),\\
b(n_H, n_E) &= \theta(1-p_H^2)^{N_{H,t}}(1-p_E p_H)^{N_{E,t}},\\
c(n_H, n_E) &= \theta(1-p_H^2)^{N_{H,t}}(1-(1-p_E p_H)^{N_{E,t}}) + (1-\theta)(1-p_E p_H)^{N_{H,t}}(1-(1-p_E^2)^{N_{E,t}}),\\
d(n_H, n_E) &= (1-\theta)(1-p_E p_H)^{N_{H,t}}(1-p_E^2)^{N_{E,t}}.
\end{aligned} \tag{28}$$

Let (n_H, n_E) be the state corresponding to a system with n_H (n_E) H (E) agents. Let $S_{\alpha,\beta} = \{(n_H, n_E) | n_H \leq \alpha, n_E \leq \beta\}$. Consider the Lyapunov function $f : (n_H, n_E) \to n_H + n_E$. We observe that:

$$\begin{aligned}
\mathbb{E}[f(N_{N,t+1}, N_{E,t+1}) - f(N_{H,t}, N_{E,t})] = 2[&\theta(1-p_H^2)^{N_{H,t}}(1-p_E p_H)^{N_{E,t}}\\
&+ (1-\theta)(1-p_E p_H)^{N_{H,t}}(1-p_E^2)^{N_{E,t}}] - 1,
\end{aligned} \tag{29}$$

and that for large enough α and β, there exists $\gamma > 0$ such that

$$\mathbb{E}[f(N_{H,t+1}, N_{E,t+1}) - f(N_{H,t}, N_{E,t}) | (N_{H,t}, N_{E,t}) \notin S_{\alpha,\beta}] \leq -\gamma.$$

By Proposition 5 (Foster (1953)) we conclude that (N_H, N_E) is a positive recurrent Markov chain, and therefore it admits a unique steady-state distribution. □

D.6.2 Proof of Lemma 2

Lemma 2 is composed of two parts. First we show in Claim 1 the existence of a steady-state distribution for M_t. Second we derive the existence of a suitable coupling.

Lemma 2. *The Markov chain M_t defined in (19) reaches a steady-state distribution M_∞, and there exists a way to couple the Markov chains $(N_{H,t}, N_{H,t})$ and M_t such that for every $t > 0, N_{H,t} \leq M_t + 1$. Therefore $\mathbb{E}[N_{H,\infty}] \leq \mathbb{E}[M_\infty] + 1$.*

Proof. Recall that in Section D.2 we defined M_t by:

$$M_{t+1} = \begin{cases} (M_t + 1) \text{ with probability } \theta(1-p_H^2)^{M_t}, \\ (M_t - 1) \text{ when with probability } \theta(1-(1-p_H^2)^{M_t}) + (1-\theta)(1-(1-p_E p_H)^{M_t}), \\ (M_t) \text{ with probability } (1-\theta)(1-p_E p_H)^{M_t}. \end{cases}$$

Claim 1. *M_t is a positive recurrent Markov chain. Therefore there exists a steady-state distribution M_∞.*

Let us now give a more precise definition in terms of random variables. We define:

- $A_H(t)$ and $A_E(t)$ are arrivals of H and E agents as defined as in section D.6.1.
- $\widetilde{H}_H(t)$ is the event that the incoming H agent matches with another H in the "simplified system" $\widetilde{H}_H(t) = \mathbf{B}(1-(1-p_H^2)^{M_t})$.
- $\widetilde{E}_H(t)$ is the event that the incoming E agent matches with another H in the "simplified system" $\widetilde{E}_H(t) = \mathbf{B}(1-(1-p_E p_H)^{M_t})$.

Note that these differ from $H_H(t)$ and $E_H(t)$ defined in section D.6.1 because they depend on M_t instead of $N_{H,t}$. This leads us to an equation for M_t:

$$M_{t+1} = M_t + A_H(t)(1-2\widetilde{H}_H(t)) - A_E(t)\widetilde{E}_H(t). \tag{30}$$

Notice that this is very similar to (27), which gives

$$N_{H,t+1} = N_{H,t} + A_H(t)(1-2H_H(t)-H_E(t)) - A_E(t)E_H(t). \tag{31}$$

Now we define the coupling of the Markov chain $(N_{H,t}, N_{E,t})_{t\geq 0}$ with $(M_t)_{t\geq 0}$ we obtained from the simplified algorithm. We start with $M_t = 0$ and $N_{H,t} = N_{E,t} = 0$. We let the two chains evolve independently, except for times where $N_{H,t} = M_t$ or $N_{H,t} = M_t + 1$. In these cases, we correlate H_H with $\widetilde{H}_H$ in the following way:

- If $\widetilde{H}_H = 1$ then $H_H = 1$, else $\widetilde{H}_H = 0$ and $H_H = \mathbf{B}(1-(1-p_H^2)^{N_{H,t}-M_t})$.
- If $\widetilde{E}_H = 1$ then $E_H = 1$, else $\widetilde{E}_H = 0$ and $E_H = \mathbf{B}(1-(1-p_H^2)^{N_{H,t}-M_t})$.

Notice that this does not modify the marginal distribution of H_H (E_H) because when $N_{H,t} \geq M_t$, we have:

$$\mathbb{P}[\widetilde{H}_H = 1] = 1-(1-p_H^2)^{M_t} \leq 1-(1-p_H^2)^{N_{H,t}} = \mathbb{P}[H_H = 1],$$

and

$$\mathbb{P}[\widetilde{E}_H = 1] = 1-(1-p_E p_H)^{M_t} \leq 1-(1-p_E p_H)^{N_{H,t}} = \mathbb{P}[E_H = 1].$$

This coupling implies that with probability 1, we get for every realization of the random variables, $H_H \geq \widetilde{H}_H$.

We next prove by induction that for all $t \geq 0, N_{H,t} \leq M_t + 1$. Note that $M_0 = N_{H,0} = 0$. Further, if $N_{H,t} < M_t$ then $N_{H,t+1} \leq M_{t+1} + 1$. The only interesting case is when $N_{H,t} = M_t$ or $N_{H,t} = M_t + 1$. Then we obtain:

$$\begin{aligned} N_{H,t+1} - N_{H,t} &= A_H(t)(1 - 2H_H(t) - H_E(t)) - A_E(t)E_H(t) \\ &\leq A_H(t)(1 - 2H_H(t)) - A_E(t)E_H(t) \\ &\leq A_H(t)(1 - 2\widetilde{H}_H(t)) - A_E(t)\widetilde{E}_H(t) \\ &= M_{t+1} - M_t. \end{aligned} \tag{32}$$

□

Proof of Claim 1. Let $S_\alpha = \{M_t | M_t \leq \alpha\}$ then for any $\epsilon > 0$ and for α large enough, we get for all $M_t > \alpha$ we have $(1 - p_H^2)^{M_t} < \epsilon$ and $(1 - p_E p_H)^{M_t} < \epsilon$. Therefore

$$\mathbb{E}[M_{t+1} - M_t | M_t \notin S_\alpha] \leq \theta\epsilon - (1 - \epsilon) < -1/2.$$

□

D.6.3 Proof of Lemma 3

Lemma 3. *S_t reaches steady-state S_∞ and we have*

$$\mathbb{E}[S_\infty] = \frac{2\theta}{(1 - 2\theta) \ln \frac{1-\theta}{1-2\theta} p_E p_H} + o(1/p_H). \tag{21}$$

Proof.

$$\begin{aligned} \theta - (1-\theta)(1 - (1 - p_E p_H)^{\frac{n}{p_E p_H}}) &\leq \theta - (1-\theta)(1 - e^{\frac{n}{p_E p_H} \ln(1 - p_E p_H)}) \\ &= \theta - (1-\theta)(1 - e^{-n(1 + \frac{p_E p_H}{2} + o(p_E p_H))}) \\ &= \theta - (1-\theta)\left(1 - e^{-n}\left(1 - \frac{n p_E p_H}{2} + o(p_E p_H)\right)\right). \end{aligned} \tag{33}$$

Notice that the quantity we wish to bound is decreasing with n and therefore it is only needed to find a bound for $n = \ln(\frac{1-\theta}{1-2\theta})$. We have $e^{-n} = \frac{1-2\theta}{1-\theta}$, therefore

$$\theta - (1-\theta)\left(1 - e^{-n}\left(1 - \frac{n p_E p_H}{2} + o(p_E p_H)\right)\right) = -(1 - 2\theta)\frac{n p_E p_H}{2} + o(p_E p_H). \tag{34}$$

□

D.6.4 Proof of Lemma 4

Here we provide a rigorous derivation of the coupling for chains M_t and S_t.

Lemma 4. *There exists a coupling of M_t and S_t such that for all $M_t > \frac{\eta}{p_E p_H}$ we have*

$$M_t - M_{t-1} \leq S_t - S_{t-1}.$$

Proof. In line with the notations introduced in in section D.6.1, we define:

- $A_H(t)$ is the event of an H agent arriving to the pool.
- $A_E(t)$ is the event of an E agent arriving.
- $\bar{E}_H(t) = \mathbf{B}((1-(1-p_E p_H)^{\frac{n}{p_E p_H}})$ is the probability to match if there were exactly $\frac{n}{p_E p_H}$ agents.

Let S_t be the random walk defined by:

$$S_{t+1} = S_t + A_H(t) - A_E(t)\bar{E}'_H(t). \tag{35}$$

We can verify that the transition states are what we expected:

$$S_{t+1} = \begin{cases} (S_t+1) \text{ with probability } \theta, \\ (S_t-1) \text{ with probability } (1-\theta)(1-(1-p_E p_H)^{\frac{n}{p_E p_H}}), \\ (S_t) \text{ with probability } (1-\theta)(1-p_E p_H)^{\frac{n}{p_E p_H}}. \end{cases}$$

Like previously, the main idea is to couple M_t with S_t such that when $M_t \geq \frac{n}{p_E p_H}$ (i.e. when it is easy to match in chain M_t), every time there is a match for S_t, there is also a match for M_t. This is only possible because $\mathbb{P}[\widetilde{H}_H = 1] \geq \mathbb{P}[\bar{H}_H = 1]$.

Formally, suppose that $M_t > \frac{n}{p_E p_H}$. We modify the definition of $\widetilde{E}_H(t) = \mathbf{B}(1-(1-p_E p_H)^{M_t})$ into $1-(1-\bar{E}_H(t)) * \mathbf{B}((1-p_E p_H)^{M_t - \frac{n}{p_E p_H}})$ which does not change its distribution. This gives us:

$$M_{t+1} - M_t \leq A_H(t) - A_E(t)\bar{E}_H(t) = S_{t+1} - S_t.$$

$\square$

D.6.5 Proof of Lemma 5

Lemma 5. *For p_H small enough, we get $\mathbb{E}[S'_\infty] \leq \frac{2}{\ln(2\theta) p_H^2}$.*

Proof. Note that

$$f(p_H) = e^{-\ln(2\theta)\frac{p_H p_E + o(p_H)}{p_H^2}} = o(p_H^2). \tag{36}$$

$$g(p_H) \leq e^{\frac{-\ln(2\theta)}{p_H^2}(p_H^2 + p_H^4/2)} = \frac{1}{2\theta}\left(1 - \frac{\ln(2\theta)p_H^2}{2}\right) + o(p_H^2). \tag{37}$$

Note that $g(p_H) < 1$ for $p_H < 1$.

With $r = \frac{\theta g(p_H)}{1-\theta g(p_H)+(1-\theta)f(p_H)} < 1$ for $f(p_H)$ small enough, we get in steady-state $\mathbb{P}[S'_\infty = i] = r^i(1-r)$. Thus using Equations 39, 40 we get:

$$\begin{aligned}
\mathbb{E}[S'_\infty] &= \frac{r}{1-r} \\
&= \frac{\theta g(p_H)}{1 - 2\theta g(p_H) + (1-\theta)f(p_H)} \\
&= \frac{(1/2 + o(1))}{(\frac{\ln(2\theta)p_H^2}{2} + o(p_H^2)) + (1-\theta)o(p_H^2)} \\
&= \frac{1}{\ln(2\theta)p_H^2} + o(1/p_H^2).
\end{aligned} \tag{38}$$

□

D.6.6 Proof of Lemma 6

Here we provide a rigorous derivation of the coupling for chains M_t and S'_t.

Lemma 6. *There exists a coupling of the random walks S_t and M_t as defined in (19) such that if $M_t \geq \frac{\ln(2\theta)}{p_H^2}$ we get:*

$$M_{t+1} - M_t \leq S'_{t+1} - S'_t.$$

Proof. Recall that we defined:

$$f(p_H) := (1 - p_H p_E)^{\ln(2\theta)/p_H^2}. \tag{39}$$

$$g(p_H) := (1 - p_H^2)^{\ln(2\theta)/p_H^2}. \tag{40}$$

Let:

- $Q(t) := \mathcal{B}(1 - g(p_H))$.

- $P(t) := \mathcal{B}(1 - f(p_H))$.

Let S'_t be the random walk defined by:

$$S'_{t+1} = S'_t + A_H(t)(1 - 2Q(t)) - A_E(t)P(t),$$

It is easy to verify that this leads to the Markov transitions we expected:

$$S'_{t+1} = \begin{cases} (S'_t + 1) \text{ with probability } \theta g(p_H), \\ (S'_t - 1) \text{ w.p. } \theta(1 - g(p_H)) + (1 - \theta)(1 - f(p_H)) = 1 - \theta g(p_H) + (1 - \theta)f(p_H), \\ (S'_t) \text{ w.p. } (1 - \theta)f(p_H). \end{cases}$$

We now wish to update the definition of M_t to introduce some correlation with S_t without modifying its marginal distribution. To do this we redefine $\widetilde{H}_H(t) := 1 - (1 - Q(t)) * \mathcal{B}((1 - p_H^2)^{M_t}/g(p_H))$, and $\widetilde{E}_H(t) := 1 - (1 - P(t)) * \mathcal{B}((1 - p_H p_E)^{M_t}/f(p_H))$ which does not change their marginal distributions.

Then if $M_t \geq \frac{\ln(2\theta)}{p_H^2}$, we get:

$$M_{t+1} - M_t \leq A_H(t)(1 - Q(t)) - A_E(t)P(t) = S'_{t+1} - S'_t.$$

□

D.6.7 Proof of Lemma 7

Lemma 7. *Using a coupling argument, we show that if $N_{E,t} \geq \frac{\ln(4)}{p_E^2}$ then $N_{E,t+1} - N_{E,t} \leq T_{t+1} - T_t$, and therefore $N_{E,t} \leq T_t + \frac{\ln(4)}{p_E^2}$.*

Proof. Using notations introduced in Section D.6.4, recall that $N_{E,t}$ follows:

$$\begin{cases} N_{H,t+1} = N_{H,t} + A_H(t)(H_\emptyset(t) - H_H(t)) - A_E(t)E_H(t), \\ N_{E,t+1} = N_{E,t} - A_H(t)H_E(t) + A_E(t)(E_\emptyset(t) - E_E(t)). \end{cases}$$

To study the E agents, we study the Markov chain T_t obtained with a simplified algorithm, that both disregards H agents when looking for matches, and only considers the first $\frac{n}{p_E^2}$ E agents for n suitably chosen. Assuming that $E_t \geq \frac{n}{p_E}$ and using the above notations, we get:

$$T_{t+1} = T_t + A_E(t)(1 - 2\widetilde{E}_E(t)),$$

Where $\widetilde{E}_E(t)$ is the indicator that the incoming E agent is able to match to one of the first $\frac{n}{p_E^2}$ E agents in the pool.

We now introduce a correlated version of $E_E(t) = 1 - (1 - \widetilde{E}_E(t))\mathcal{B}((1 - p_E^2)^{N_{E,t} - \frac{n}{p_E^2}}) \geq$

$\widetilde{E}_E(t)$, which does not impact the marginal distribution of $E_E(t)$. We now have:

$$\begin{aligned} N_{E,t+1} - N_{E,t} &= -A_H(t)H_E(t) + A_E(t)(1 - E_H(t) - 2E_E(t)) \\ &\leq A_E(t)(1 - 2E_E(t)) \\ &\leq A_E(t)(1 - 2\widetilde{E}_E(t)) = T_{t+1} - T_t. \end{aligned} \tag{41}$$

Therefore $N_{E,t} \leq T_t + \frac{n}{p_E^2}$.

□

E Chain matching: proof of Theorem 2

Theorem 2. *Suppose that $p_E = 1$ and $0 < \theta \leq 1$. Under the ChainMatch policy, the market reaches steady-state and:*

1. *If $\theta < 1$, there exists a constant K_θ such that:*

$$w_H \leq \frac{K_\theta}{p_H}. \tag{4}$$

2. *If $\theta = 1$:*

$$w_H \leq \frac{2\ln(1/p_H)}{p_H}. \tag{5}$$

The derivation of the proof involves the following steps:

- First we define formally our system as a Markov chain, and we show that it is positive recurrent, and therefore reaches steady-state.

- We prove that the chain segments that are conducted in *ChainMatch* have known expected length (Proposition 2).

- Then we prove theorem 2. We separate the case $\theta = 1$ and the case $\theta < 1$.

- Finally we prove all the intermediate technical lemmas used.

E.1 Markov Chain representation of *ChainMatch*

Because we consider only the case in which $p_E = 1$, there will never be any E agent in the system. Because there is no ambiguity, we will simplify the notations and write N_t instead of $N_{H,t}$. Therefore we can count only the number of H agents in the pool. We consider the

1-dimensional Markov Chain $(N_t)_{t\geq 0}$ that represents the number of agents in the pool at time t. From state (N_t) all the states between 0 and N_t+1 are reachable.

Note that the number of H agents only decreases when there arrives an agent (H or E) that is able to recieve from the bridge donor. This happens with probability $\theta p_H + (1-\theta)$ and a chain is then started. Suppose that there are n agents in the system, and we are currently running a chain. The probability that the bridge donor is able to match at least one of the agents is $1-(1-p_H)^n$. This allows us to compute the probability that a chain is of length $k \leq N_t$ conditional on the fact that there were N_t agents just before the start of the chain.

$$N_{t+1} = \begin{cases} (N_t+1) \text{ with probability } \theta(1-p_H), \\ N_t - k \text{ with probability } (\theta p_H + (1-\theta)) * (1-p_H)^{N_t-k} \prod_{i=0}^{k-1}(1-(1-p_H)^{N_t-i}), \forall k \leq N_t. \end{cases}$$

Note that the process N_t so defined has the Markov property.

E.2 Proof of Proposition 2

Proposition 2. *Let L_∞ be the length of a new chain segment formed by ChainMatch policy in steady-state, and suppose $p_E = 1$. Then we have:*

$$\mathbb{E}[L_\infty] = \frac{\theta(1-p_H)}{\theta p_H + (1-\theta)} + 1.$$

Note that L_t is only defined if a chain was started at time t. The idea of the proof is that the expected number of H agents must stay constant. Therefore, the length of chain segments is inversely proportional with the frequency at which those chain segments are started.

Proof. In steady-state,

$$\begin{aligned} 0 = \mathbb{E}[N_t - N_{t+1}] &= \mathbb{P}[N_{t+1} > N_t]\mathbb{E}[N_t - N_{t+1}|N_{t+1} > N_t] \\ &+ \mathbb{P}[N_{t+1} \leq N_t]\mathbb{E}[N_t - N_{t+1}|N_{t+1} \leq N_t] \\ &= \theta(1-p_H)\mathbb{E}[N_t - N_{t+1}|N_{t+1} > N_t] \\ &+ (\theta p_H + (1-\theta))\mathbb{E}[(N_t - N_{t+1})|N_{t+1} \leq N_t]. \end{aligned} \tag{42}$$

Where we used the law of total expectation, the fact that $\mathbb{P}[N_{t+1} > N_t] = \mathbb{P}[N_{t+1} = N_t + 1] = \theta(1-p_H)$, the fact that $\mathbb{P}[N_{t+1} \leq N_t] = \theta p_H + (1-\theta)$. Lastly note that if $N_{t+1} \leq N_t$ then a new chain segment has been formed with expected length of $\mathbb{E}[L_t] =$

$\mathbb{E}[(N_t - N_{t+1})|N_{t+1} \leq N_t]$. Therefore:

$$\theta(1-p_H) = (\theta p_H + (1-\theta))\mathbb{E}[L_t].$$

$\square$

E.3 Proof of Theorem 2 : case $\theta = 1$

To prove Theorem 2 we will need the following technical claim (the proof is deferred to E.5.2).

Claim 2. *For $0 \leq n \leq \infty$ let*

$$z_{n,i,p} = \prod_{j=0}^{i-1}(1-(1-p)^{n-j}).$$

Then for any $\alpha \in (0,1)$ and $\epsilon > 0$ there exists p_0 such that for all $p < p_0$, for all $n \geq \frac{2\ln(1/p_H)}{p_H}$

$$\sum_{i=0}^{n} z_{n,i,p} \geq n\alpha(1 - \frac{\epsilon}{1-\epsilon}).$$

As we shall see later, $z_{n,i,p}$ represents the probability that a chain that is started when there are n agents in the pool reaches a length of more than i. Therefore, $\sum_{i=0}^{n} z_{n,i,p}$ is the expected length of such a chain for a given parameter p. This lemma states that if there are "enough" agents in the pool (in particular at least a log factor more than $\frac{1}{p}$), then the chain length is of the order of the number of agents in the pool.

Remark. *The constant 2 in $\frac{2\ln(1/p_H)}{p_H}$ in claim 2 could be lowered to any $1+\delta$ with $\delta > 0$ without any alterations to the proof. This is also true for the theorem 2.*

Proof of Theorem 2. The idea of the proof is to cut the expectation in two parts. Observe that

$$\mathbb{E}[N_t] = \sum_{n=0}^{\frac{2\ln(1/p_H)}{p_H}} n\mathbb{P}[N_t = n] + \sum_{n=\frac{2\ln(1/p_H)}{p_H}}^{\infty} n\mathbb{P}[N_t = n].$$

We bound the first part by

$$\sum_{n=0}^{\frac{2\ln(1/p_H)}{p_H}} n\mathbb{P}[N_t = n] \leq \frac{2\ln(1/p_H)}{p_H}\mathbb{P}[N_t = n] \leq \frac{2\ln(1/p_H)}{p_H}.$$

It is sufficient to show that there exists a constant C independent of p such that

$$\sum_{n=\frac{2\ln(1/p_H)}{p_H}}^{\infty} n\mathbb{P}[N_t = n] \leq C.$$

By the definition of the Chains removal algorithm we have that

$$\mathbb{P}[L_t \geq x | N_t = n] = \mathbb{I}_{x \leq n} \prod_{i=0}^{x-1} (1 - (1-p_H)^{n-i}) = z_{n,x,p_H} \mathbb{I}_{x \leq n}.$$

Therefore

$$\begin{aligned}
\mathbb{E}[L_t] &= \mathbb{E}_{N_t}[\mathbb{E}[L_t | N_t]] \\
&= \sum_{n=0}^{\infty} \left(\sum_{i=0}^{n} \mathbb{P}[L_t > i | N_t = n] \right) \mathbb{P}[N_t = n] \\
&= \sum_{n=0}^{\infty} \left(\sum_{i=0}^{n} z_{n,i,p_H} \right) \mathbb{P}[N_t = n] \\
&\geq \sum_{n=\frac{2\ln(1/p_H)}{p_H}}^{\infty} \left(\sum_{i=0}^{n} z_{n,i,p_H} \right) \mathbb{P}[N_t = n].
\end{aligned} \tag{43}$$

By Claim 2, we obtain that for any $\alpha \in (0,1)$ and $\epsilon \in (0,1)$ there exists p_0 such that for all $p < p_0$:

$$\mathbb{E}[L_t] \geq \sum_{n=\frac{2\ln(1/p_H)}{p_H}}^{\infty} \alpha(1 - \frac{\epsilon}{1-\epsilon}) n\mathbb{P}[N_t = n].$$

Therefore

$$\frac{\mathbb{E}[L_t]}{\alpha(1 - \frac{\epsilon}{1-\epsilon})} \geq \sum_{n=\frac{2\ln(1/p_H)}{p_H}}^{\infty} n\mathbb{P}[N_t = n].$$

We prove in Lemma 8 (cf appendix E.5.1) that the algorithm leads to a market with steady-state behavior, therefore by Proposition 2

$$\mathbb{E}[L_t] = \frac{\theta(1-p_H)}{\theta p_H + (1-\theta)}.$$

The following concluses the proof:

$$
\begin{aligned}
\mathbb{E}[N_t] &= \sum_{n=0}^{\frac{2\ln(1/p_H)}{p_H}} n\mathbb{P}[N_t = n] + \sum_{n=\frac{2\ln(1/p_H)}{p_H}}^{\infty} n\mathbb{P}[N_t = n] \\
&\leq \frac{2\ln(1/p_H)}{p_H} + \frac{\mathbb{E}[L_t]}{\alpha(1-\frac{\epsilon}{1-\epsilon})} \\
&\leq \frac{2\ln(1/p_H)}{p_H} + \frac{1}{p_H\alpha(1-\frac{\epsilon}{1-\epsilon})}.
\end{aligned} \tag{44}
$$

□

Remark. *Note that this proof also applies to the case where we have $\theta < 1$:*

$$\mathbb{E}[L_t] = \frac{\theta(1-p_H)}{\theta p_H + (1-\theta)} \leq \frac{\theta}{(1-\theta)}.$$

And by taking $\alpha = 1/2$ and $\epsilon = 1/4$ we get:

$$\mathbb{E}[N_t] \leq \frac{2\ln(1/p_H)}{p_H} + \frac{3\theta}{(1-\theta)}.$$

However, this bound is weaker than the bound that we prove in the special case $\theta < 1$. Put simply, this stems from the fact that the cut of the expectation in two parts is not optimal.

E.4 Proof of Theorem 2 : case $\theta < 1$

Proof. This proof follows the outline of the proof for the case $\theta = 1$. Here we change where we cut the expectation in two parts. Let c be function of p_H to be determined later. We get:

$$\mathbb{E}[N_t] = \sum_{n=0}^{\frac{c}{p_H}} n\mathbb{P}[N_t = n] + \sum_{n=\frac{c}{p_H}}^{\infty} n\mathbb{P}[N_t = n].$$

Similarly to the previous proof, we have

$$\sum_{n=0}^{\frac{c}{p_H}} n\mathbb{P}[N_t = n] \leq \frac{c}{p_H}.$$

And

$$
\begin{aligned}
\widetilde{l} &= \mathbb{E}[\mathbb{E}[L_t|N_t]] \\
&\geq \sum_{n=\frac{c}{p_H}}^{\infty} \left(\sum_{i=0}^{n} z_{n,i,p_H} \right) \mathbb{P}[N_t = n].
\end{aligned}
\tag{45}
$$

We now use another technical claim:

Claim 3. *There exists $k < 1/2$ such that for any $n > 0$, there exists $p_0 > 0$ such that for all $p_H < p_0$, for all $n \geq \frac{2}{(1-p_H)p_H}$:*

$$
\sum_{i=0}^{n} z_{n,i,p_H} \geq np_H(1 - \frac{k}{1-k}).
$$

Therefore, choosing $c = \frac{2}{1-p_H}$ we get

$$
\sum_{n=\frac{2}{(1-p_H)p_H}}^{\infty} n\mathbb{P}[N_t = n] \leq \frac{\widetilde{l}}{p_H(1 - \frac{k}{1-k})}.
$$

Given that the market reaches steady-state and $\theta < 1$ we obtain that

$$
\widetilde{l} = l^* = \frac{\theta(1-p_H)}{\theta p_H + (1-\theta)}.
$$

Finally, we get that

$$
\mathbb{E}[N_t] \leq \frac{2}{(1-p_H)p_H} + \frac{\theta(1-p_H)}{(\theta p_H + 1 - \theta)(p_H(1 - \frac{k}{1-k}))} = \frac{K_\theta}{p_H} + o(\frac{1}{p_H})
$$

Where $C_\theta = 2 + \frac{\theta}{(1-\theta)} \frac{1-k}{(1-2k)}$. $\square$

Remark. *Notice that the assumption that $\theta < 1$ is crucial in order for this bound to be tight, otherwise, it goes to the order of $\frac{1}{p_H^2}$. This means that this bound gets very bad as θ gets closer to 1, and in practice it is observed that this is not tight when $\theta \geq 0.7$. A search for a tighter bound could be achieved through a splitting of the expectation in a manner that depends on θ. The exact value of the tightest bound remains an open question.*

E.5 Proof of technical claims

E.5.1 Existence of steady-state

Lemma 8. *For any arrival rate $0 \leq \theta \leq 1$ of H agents and p sufficiently small, the chain algorithm leads to steady-state.*

Proof. The system has the Markov Property (it is memoryless by construction), and its transitions are given by:

$$\begin{cases} N_{t+1} = N_t + 1 \text{ w.p. } \theta(1 - p_H) \\ \forall k \in [0; N_t], N_{t+1} = (N_t - k) \text{ w.p.} \\ (\theta p_H + (1 - \theta)) * (1 - p_H)^{N_t - k} \prod_{i=0}^{k-1}(1 - (1 - p_H)^{N_t - i}). \end{cases}$$

The first item corresponds to the case where an H agent arrives (θ) and isn't able to receive from the bridge donor $(1-p)$. The second corresponds to the case where the incoming agent (either E with probability $(1-\theta)$ or H with probability θp) is able to receive from the bridge donor. In that case we consider that a chain is started (of initial length 0). Then for any k between 0 and N_t, $\prod_{i=0}^{k-1}(1 - (1 - p_H)^{N_t - i})$ is the probability that the chain is of length at least k. $(1 - p_H)^{N_t - k}$ corresponds to the fact that the chain isn't able to continue beyond k.

We wish to show that there exists $\widetilde{n}$ and $\gamma > 0$ such that

$$\mathbb{E}[N_{t+1} - N_t | N_t > \widetilde{n}] < -\gamma.$$

Note that

$$\mathbb{E}[N_{t+1} - N_t | N_t > \widetilde{n}]\mathbb{P}[N_t > \widetilde{n}] = \sum_{n=\widetilde{n}}^{\infty} \left(\theta(1 - p_H) - (\theta p_H + (1 - \theta)) \sum_{i=0}^{n} z_{n,i,p} \right) \mathbb{P}[N_t = n] \tag{46}$$

As a reminder, $z_{n,i,p} = \prod_{j=0}^{i-1}(1 - (1 - p)^{n-j})$ for $0 \leq n \leq \infty$, and we have the property that for all $n \geq m$, $z_{n,i,p} \geq z_{m,i,p}$.

Let us fix $\alpha = 1/2$, $\epsilon = 1/4$. From the technical lemma there exists p_0 such that if $p < p_0$ and $n > \frac{2\ln(1/p)}{p}$ then we get:

$$\sum_{i=0}^{n} z_{n,i,p} \geq \frac{n}{3}.$$

Let us consider $\widetilde{n} = \max\left(\frac{2\ln(1/p)}{p}, \frac{3(\theta(1-p)+\gamma)}{(1-\theta)+\theta p}\right)$. For $n \geq \widetilde{n}$:

$$\sum_{i=0}^{n} z_{n,i,p} \geq n/4 \geq \frac{(\theta(1-p)+\gamma)}{(1-\theta)+\theta p}.$$

Morevoer

$$\theta(1-p_H) - (\theta p_H + (1-\theta))\sum_{i=0}^{n} z_{n,i,p} \leq -\gamma$$

Therefore there exists p_0 such that for $p < p_0$,

$$\mathbb{E}[N_{t+1} - N_t | N_t > \frac{2\ln(1/p)}{p}]\mathbb{P}[N_t > \frac{2\ln(1/p)}{p}] \leq -\gamma \mathbb{P}[N_t \geq \frac{2\ln(1/p)}{p}],$$

and

$$\mathbb{E}[N_{t+1} - N_t | N_t > \frac{2\ln(1/p)}{p}] \leq -\gamma.$$

Therefore N_t is a recurring Markov Chain, and it has a steady-state distribution. □

E.5.2 Proof of Claim 2

Claim 2. *For $0 \leq n \leq \infty$ let*

$$z_{n,i,p} = \prod_{j=0}^{i-1}(1-(1-p)^{n-j}).$$

Then for any $\alpha \in (0,1)$ and $\epsilon > 0$ there exists p_0 such that for all $p < p_0$, for all $n \geq \frac{2\ln(1/p_H)}{p_H}$

$$\sum_{i=0}^{n} z_{n,i,p} \geq n\alpha(1 - \frac{\epsilon}{1-\epsilon}).$$

Proof. Let us consider any $0 < \alpha < 1$ independent of n and p. In order to be able to get a lower bound of the term inside the product, we only consider the truncated sum:

$$\sum_{i=0}^{n} z_{n,i,p} \geq \sum_{i=0}^{n\alpha}\prod_{j=0}^{i-1}(1-(1-p)^{n-j}).$$

Using the fact that for all $j \leq \alpha n$, we get $(1-(1-p)^{n-j}) \geq (1-(1-p)^{n(1-\alpha)})$, we can simplify:

$$\sum_{i=0}^{n} z_{n,i,p} \geq \sum_{i=0}^{n\alpha}(1-(1-p)^{n(1-\alpha)})^{i}.$$

This is a geometric sum with parameter $y_n = 1 - (1-p)^{n(1-\alpha)}$ therefore we that

$$\sum_{i=0}^{n} z_{n,i,p} \geq \frac{1 - y_n^{n\alpha+1}}{1 - y_n} \geq \frac{1 - y_n^{n\alpha}}{1 - y_n}. \tag{47}$$

Let us now expand y_n:

$$\begin{aligned} y_n^{\alpha n} &= \left(1 - (1-p)^{n(1-\alpha)}\right)^{\alpha n} \\ &= 1 - n\alpha(1-p)^{n(1-\alpha)} + \sum_{k=2}^{n\alpha} \binom{\alpha n}{k} \left(-(1-p)^{n(1-\alpha)}\right)^k. \end{aligned} \tag{48}$$

And

$$1 - y_n^{\alpha n} \geq n\alpha(1-p)^{n(1-\alpha)} - \sum_{k=2}^{n\alpha} \binom{\alpha n}{k} (1-p)^{kn(1-\alpha)}. \tag{49}$$

Claim 4. *For all $\epsilon > 0$ there exists p_0 such that for all $p < p_0$, for all $n \geq \frac{2\ln(1/p)}{p}$,*

$$(1-p)^{n(1-\alpha)} \leq \frac{\epsilon}{n\alpha}.$$

Using Claim 4 together with the fact that $\binom{\alpha n}{k} \leq (\alpha n)^k$ we get that for p small enough and $n \geq \frac{2\ln(1/p)}{p}$,

$$\begin{aligned} \sum_{k=2}^{n\alpha} \binom{\alpha n}{k} (1-p)^{kn(1-\alpha)} &\leq \sum_{k=2}^{n\alpha} (\alpha n(1-p)^{n(1-\alpha)})^k \\ &= (\alpha n(1-p)^{n(1-\alpha)})^2 \sum_{k=0}^{n\alpha-2} (\alpha n(1-p)^{n(1-\alpha)})^k \\ &\leq (\alpha n(1-p)^{n(1-\alpha)})\epsilon \sum_{k=0}^{n\alpha-2} \epsilon^k \\ &\leq (\alpha n(1-p)^{n(1-\alpha)})\frac{\epsilon}{1-\epsilon}. \end{aligned} \tag{50}$$

Putting together (49) and (50), this proves the claim:

$$\begin{aligned}\sum_{i=0}^{n} z_{n,i,p} &\geq \frac{1-y_n^{n\alpha}}{1-y_n} \\ &\geq \frac{n\alpha(1-p)^{n(1-\alpha)}(1-\frac{\epsilon}{1-\epsilon})}{(1-p)^{n(1-\alpha)}} \\ &= n\alpha(1-\frac{\epsilon}{1-\epsilon}).\end{aligned} \tag{51}$$

□

Proof of Claim 4. Let us pose $\widetilde{n}_p = \frac{2\ln(1/p)}{p}$ We notice that the function $\phi : n \mapsto n\alpha(1-p)^{n(1-\alpha)}$ is decreasing after $n^* = \frac{1}{(1-\alpha)\ln(1/(1-p))}$. We have $\frac{1}{(1-\alpha)\ln(1/(1-p))} \leq \frac{1}{(1-\alpha)p}$ and therefore for p small enough $\widetilde{n}_p \geq \frac{-1}{(1-\alpha)\ln(1-p)}$ and for $n \geq \widetilde{n}_p$ we have $\phi(n) \leq \phi(\widetilde{n}_p)$:

$$\begin{aligned}n\alpha(1-p)^{n(1-\alpha)} &\leq \widetilde{n}_p\alpha(1-p)^{\widetilde{n}_p(1-\alpha)} \\ &\leq \frac{2\ln(1/p)}{p}\alpha(1-p)^{\frac{2\ln(1/p)}{p}(1-\alpha)} \\ &= \frac{2\ln(1/p)}{p}\alpha e^{\frac{2\ln(1/p)}{p}(1-\alpha)\ln(1-p)} \\ &\leq \frac{2\ln(1/p)}{p}\alpha e^{-2\ln(1/p)(1-\alpha)} \\ &= \frac{2\ln(1/p)}{p}\alpha p^2 e^{-\alpha} \\ &\mapsto 0.\end{aligned} \tag{52}$$

Therefore for all $\epsilon > 0$ there exists p_0 such that for all $p < p_0$, for all $n \geq \frac{2\ln(1/p)}{p}$,

$$(1-p)^{n(1-\alpha)} \leq \frac{\epsilon}{n\alpha}.$$

□

E.5.3 Proof of Claim 3

The proof of Claim 3 follows the same ideas that were used in proving Claim 2. Notice that we only ask n to grow linearly in $\frac{1}{p}$. This is at the expense of a slightly weaker statement (there exists k instead of for all ϵ).

Claim 3. *There exists $k < 1/2$ such that for any $n > 0$, there exists $p_0 > 0$ such that for all*

$p_H < p_0$, for all $n \geq \frac{2}{(1-p_H)p_H}$:

$$\sum_{i=0}^{n} z_{n,i,p_H} \geq np_H(1 - \frac{k}{1-k}).$$

Proof. As a reminder, for $0 \leq n \leq \infty$, we have $z_{n,i,p_H} = \prod_{j=0}^{i-1}(1-(1-p_H)^{n-j})$.

Let us consider any $p_0 > 0$, and $p_H < p_0$. Using the fact that $(1-(1-p_H)^{n-j}) \geq (1-(1-p_H)^{n(1-p_H)})$ for all $j \leq p_H n$, we can truncate the sum:

$$\sum_{i=0}^{n} z_{n,i,p_H} \geq \sum_{i=0}^{np_H}\prod_{j=0}^{i-1}(1-(1-p_H)^{n-j}) \geq \sum_{i=0}^{np_H}(1-(1-p_H)^{n(1-p_H)})^i.$$

Writing $y_n = 1-(1-p_H)^{n(1-p_H)}$ we get

$$\sum_{i=0}^{n} z_{n,i,p_H} \geq \frac{1-y_n^{np_H+1}}{1-y_n} \geq \frac{1-y_n^{np_H}}{1-y_n}. \tag{53}$$

Claim 5. *For $n \geq \widetilde{n}_p$, there exists $k < 1/2$ such that:*

$$np_H(1-p_H)^{n(1-p_H)} \leq k.$$

This yields

$$\begin{aligned} y_n^{p_H n} &= \left(1-(1-p_H)^{n(1-p_H)}\right)^{p_H n} \\ &= 1 - np_H(1-p_H)^{n(1-p_H)} + \sum_{i=2}^{np_H}\binom{p_H n}{i}\left(-(1-p_H)^{n(1-p_H)}\right)^i. \end{aligned} \tag{54}$$

And

$$1 - y_n^{p_H n} \geq np_H(1-p_H)^{n(1-p_H)} - \sum_{i=2}^{np_H}\binom{p_H n}{i}(1-p_H)^{in(1-p_H)}. \tag{55}$$

Using the fact that $\binom{p_H n}{i} \leq (p_H n)^i$ we get that

$$\begin{aligned}
\sum_{i=2}^{np_H} \binom{p_H n}{i} (1-p_H)^{in(1-p_H)} &\leq \sum_{i=2}^{np_H} (p_H n(1-p_H)^{n(1-p_H)})^i \\
&= (p_H n(1-p_H)^{n(1-p_H)})^2 \sum_{i=0}^{np_H-2} (p_H n(1-p_H)^{n(1-p_H)})^i \\
&\leq (p_H n(1-p_H)^{n(1-p_H)}) k \sum_{i=0}^{np_H-2} k^i \\
&\leq (p_H n(1-p_H)^{n(1-p_H)}) \frac{k}{1-k}.
\end{aligned} \tag{56}$$

Therefore, putting together (55) and (56), we get that

$$\begin{aligned}
\sum_{i=0}^{n} z_{n,i,p_H} &\geq \frac{1-y_n^{np_H}}{1-y_n} \\
&\geq \frac{np_H(1-p_H)^{n(1-p_H)}(1-\frac{k}{1-k})}{(1-p_H)^{n(1-p_H)}} \\
&\geq np_H(1-\frac{k}{1-k}).
\end{aligned} \tag{57}$$

□

Proof. Let us pose $\widetilde{n}_p = \frac{c}{p_H}$ We notice that $np_H(1-p_H)^{n(1-p_H)}$ is maximal for $n = \frac{-1}{(1-p_H)\ln(1-p_H)}$, and is decreasing for larger values of n. We have $\frac{-1}{(1-p_H)\ln(1-p_H)} \leq \frac{1}{(1-p_H)p_H}$ and therefore if $c > \frac{1}{1-p_H}$ then for p_H small enough $\widetilde{n}_p \geq \frac{-1}{(1-p_H)\ln(1-p_H)}$. [45] We also fix $\widetilde{n}_p = \frac{2}{(1-p_H)p_H}$

For $n \geq \widetilde{n}_p$ we have

$$\begin{aligned}
np_H(1-p_H)^{n(1-p_H)} &\leq \widetilde{n}_p p_H(1-p_H)^{\widetilde{n}_p(1-p_H)} \\
&\leq \frac{2}{(1-p_H)p_H} p_H(1-p_H)^{\frac{2}{(1-p_H)p_H}(1-p_H)} \\
&\leq \frac{2}{(1-p_H)} e^{\frac{2}{p_H}\ln(1-p_H)} \\
&\leq \frac{2}{(1-p_H)} e^{-2+o(1)} \\
&\leq 2e^{-2} + o(1).
\end{aligned} \tag{58}$$

Using the fact that $2e^{-2} < \frac{1}{2}$ we get that there exists $k < 1/2$, and $p_0 > 0$ such that

[45] In everything that follows we consider $c = \frac{2}{1-p_H}$, but any constant $1+\delta$ could work instead of 2.

$\frac{2e^{2\ln(1-p_H)/p_H}}{1-p_H} < k$ for all $p_H < p_0$. Therefore for $n \geq \widetilde{n}_p$, we have :

$$np_H(1-p_H)^{n(1-p_H)} \leq k.$$

$\square$

F Proof of Proposition 1

We now modify the *BilateralMatch(H)* algorithm into *BilateralMatch(E)* which gives priority to E nodes instead of H nodes. We show that when $p_E = 1$ [46], Theorem 1 still holds. Namely:

Theorem 3. *Under BilateralMatch(E), the market reaches steady-state. Furthermore, there exist positive constants A'_θ, B'_θ, and C'_θ such that:*

- *If $\theta < 1/2$, then*

$$w_H \leq \frac{A'_\theta}{p_E p_H}. \tag{59}$$

- *If $\theta \geq 1/2$, then*

$$w_H \leq \frac{B'_\theta}{p_H^2}. \tag{60}$$

- *For $0 \leq \theta < 1$*

$$w_E \leq \frac{C'_\theta}{p_E^2}. \tag{61}$$

The proof of Theorem 3 is organised as follows:

First in Subsection F.1, we provide a Markov chain representation of the dynamic system under *BilateralMatch(E)*. Next, in Subsection F.2. we couple the chain with a simplified Markov chain for which the computation of the steady-state equilibrium is tractable. In Subsection F.3, we analyze the simplified chain, and show that bounding its eigenvalues is sufficient to provide a bound on the steady-state expected number of agents.

F.1 Markov chain representation

Because we assume $p_E = 1$, the E agents match to each other with probability 1, there can never be more than 1 E agent in the system at any given time. Therefore the system can

[46] From the simulations we conducted, it seems that this result extends to the case where $p_E < 1$. However due to an increased dimensionality of the problem, we haven't been able to prove the Theorem 3 in this general case.

be represented by a Markov chain $(N_{H,t}, N_{E,t})$ where $N_{N,t} \geq 0$ and $n_{E,t} \in \{0, 1\}$. Let π_{n_H,n_E} be the steady-state probability of the chain being in state (n_H, n_E). We have:

$$\mathbb{E}[N_{H,\infty}] = \sum_{n_H=0}^{\infty} n_H \left(\pi_{n_H,0} + \pi_{n_H,1}\right).$$

Because we removed the n_E dimension of the problem, and in order to simplify the notations, we will use n instead of n_H in the equations that follow.

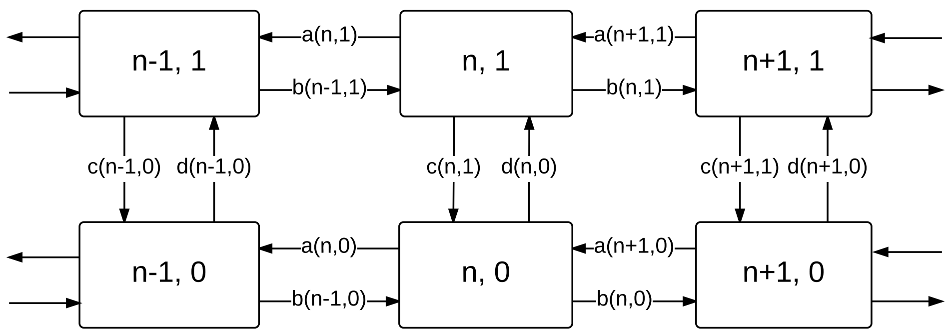

Figure 9: Transitions Markov chain resulting from *Bilateral(E)* in the case $p_E = 1$.

In Figure 9, we show the transitions for the Markov chain that represents our dynamic system.

$$\begin{aligned}
a(n,0) &= \theta(1-(1-p_H^2)^n) + (1-\theta)(1-(1-p_H)^n), \\
b(n,0) &= \theta(1-p_H^2)^n, \\
d(n,0) &= (1-\theta)(1-p_H)^n, \\
a(n,1) &= \theta(1-p_H)(1-(1-p_H^2)^n), \\
b(n,1) &= \theta(1-p_H)(1-p_H^2)^n, \\
c(n,1) &= \theta(1-(1-p_H)) + (1-\theta).
\end{aligned} \tag{62}$$

$a(n,0)$ corresponds to the case where an E or an H agent arrives and matches an existing H, $a(n,1)$ corresponds to the event where an H arrives to the system, doesn't match the existing E (to which we give priority), and matches one of the existing H agents. Similarly, $b(n,0)$ corresponds to the case where an H agent arrives to the system and is not able to match to the existing H, and $b(n,1)$ corresponds to the case where the incoming H is not able to match to either the E or another H. Finally, $d(n,0)$ corresponds to the case where an incoming E is unable to match any of the existing H agents, and $c(n,1)$ corresponds to the event where either an H or an incoming E agent matches the existing E.

F.2 Coupling to a simplified Markov Chain

We now wish to simplify the Markov chain in order to provide a tractable computation of the steady-state distribution. To do this, we notice that it suffice to upper bound the "forward transitions" of the upper states (corresponding to 1 E agent) using $b'(n,1) = \theta(1-p_H^2)^n \geq b(n,1)$ and to lower bound the "backwards transitions" of the upper states using $a'(n,1) = 0 \leq a(n,1)$. Note that this new Markov Chain does not satisfy the conservation of flow, therefore we add self loops (to the upper states) (not shown in Figure 10 so that the outgoing probabilities out of every node sum up to 1. This yields the chain $(N'_{H,t}, N'_{E,t})$ for which we plot the transitions in Figure 10:

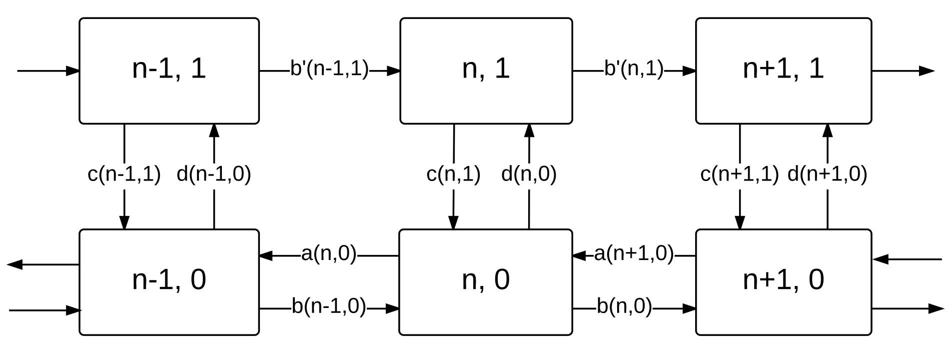

Figure 10: Transitions for the simplified Markov Chain

Using similar coupling ideas we develop in Appendix D and E, we can prove that for all t, $N_{H,t} \leq N'_{H,t} + 1$.

F.3 Analysing the simplified Markov Chain

Flow conservation for the coupled chain between n and $n+1$ yields:

$$\begin{aligned}&b(n,0)\pi_{n,0}+b'(n,1)\pi_{n,1}=a(n+1,0)\pi_{n+1,0}\\&(\pi_{n,0}+\pi_{n,1})\theta(1-p_H^2)^n=\pi_{n+1,0}\left(1-\theta(1-p_H^2)^{n+1}-(1-\theta)(1-p_H)^{n+1}\right)\end{aligned}\tag{63}$$

Flow conservation for the coupled chain out of node $(n,1)$ yields:

$$\begin{aligned}&(b'(n,1)+c(n,1))\pi_{n,1}=b'(n-1,1)\pi_{n-1,1}+d(n,0)\pi_{n,0},\\&\pi_{n,1}\left(\theta(1-p_H^2)^n+\theta p_H+(1-\theta)\right)=\theta(1-p_H^2)^{n-1}\pi_{n-1,1}+(1-\theta)(1-p_H)^n\pi_{n,0}.\end{aligned}\tag{64}$$

To simplify notations, we introduce

$$u_n=\frac{\theta(1-p_H^2)^{n-1}}{\theta(1-p_H^2)^n+(\theta p_H+1-\theta)},\tag{65}$$

$$v_n=\frac{(1-\theta)(1-p_H)^n}{\theta(1-p_H^2)^n+(\theta p_H+1-\theta)},\tag{66}$$

$$g_n=\frac{1}{1-\theta(1-p_H^2)^n-(1-\theta)(1-p_H)^n}.\tag{67}$$

Equation (64) simplifies into:

$$\pi_{n,1}=u_n\pi_{n-1,1}+v_n\pi_{n,0},$$

and (63) simplifies into:

$$\begin{aligned}\pi_{n+1,0}=&g_{n+1}\theta(1-p_H^2)^n\left(\pi_{n,0}+\pi_{n,1}\right)\\=&g_{n+1}\theta(1-p_H^2)^n\left((1+v_n)\pi_{n,0}+u_n\pi_{n-1,1}\right).\end{aligned}\tag{68}$$

Putting everything together, we get:

$$\begin{pmatrix}\pi_{n,1}\\\pi_{n,0}\\\pi_{n+1,0}\end{pmatrix}=\begin{pmatrix}u_n&0&v_n\\0&0&1\\g_{n+1}\theta(1-p_H^2)^nu_n&0&g_{n+1}\theta(1-p_H^2)^n(1+v_n)\end{pmatrix}\begin{pmatrix}\pi_{n-1,1}\\\pi_{n-1,0}\\\pi_{n,0}\end{pmatrix}\tag{69}$$

We call this matrix A_n. In Section F.3.1, we show that for n large enough, the eigenvalues of A_n are bounded by a constant strictly less than 1. Using a Lyapunov argument, this justifies that the system reaches steady-state. In Section F.3.2, we use our bound on the eigenvalues of the matrix to upper-bound $\mathbb{E}[N'_{H,\infty}]$. This gives an upper-bound on $\mathbb{E}[N_{H,\infty}]$ which completes the proof.

F.3.1 Bounding the matrix eigenvalues

Let $\lambda_1^{(n)}$, $\lambda_2^{(n)}$ and $\lambda_3^{(n)}$ be its eigenvalues, and let $\lambda_n = \max(|\lambda_1^{(n)}|, |\lambda_2^{(n)}|, |\lambda_3^{(n)}|)$. We wish to prove that there exists $\lambda < 1$ such that for all n, $\lambda_n < \lambda$. To do this, we will look separately into the cases $\theta > 1/2$ and $\theta < 1/2$.

Case $\theta > 1/2$ **:** Let k_ϵ be a constant (We allow this k_ϵ to depend on θ, but it is independent of p_H and n) to be fixed later. Let us consider $n^* = \frac{k_\epsilon}{p_H^2}$. For any $n \geq n^*$ we have $u_n \leq \frac{\theta e^{-k_\epsilon}}{\theta p_H + 1 - \theta}$ and $v_n \leq \frac{(1-\theta)e^{-k_\epsilon/p_H}}{\theta p_H + 1 - \theta} = o(p_H)$. Therefore for any $\epsilon > 0$ there exists a k_ϵ large enough such that $u_n \leq \epsilon$, $v_n \leq \epsilon$, $g_{n+1}\theta(1-p_H^2)^n u_n \leq \epsilon$ and $g_{n+1}\theta(1-p_H^2)^n(1+v_n) \leq \epsilon$.

$$A = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$$

All the eigenvalues of A are 0, therefore there exists a small perturbation ϵ such that for $n \geq n^* = k_\epsilon / p_H^2$, we get that $\lambda_i < 1/2$.

Case $\theta < 1/2$**:** Let us consider $n^* = \frac{k'_\epsilon}{p_H}$. For any $n \geq n^*$ we have $(1 - p_H 2)^n = 1 + o(1)$ $u_n = \frac{\theta}{\theta + \theta p_H + 1 - \theta} + o(\theta) = \frac{\theta}{1 + \theta p_H} + o(\theta)$, $g_n = \frac{1}{1 - \theta - (1-\theta)e^{-k'_\epsilon}}$ and $v_n \leq \frac{(1-\theta)e^{-k'_\epsilon}}{\theta + \theta p_H + 1 - \theta} = \frac{(1-\theta)e^{-k'_\epsilon}}{1 + \theta p_H}$
Furthermore:

$$g_{n+1}\theta(1 - p_H^2)^n u_n = \frac{\theta^2}{(1 - \theta - (1-\theta)e^{-k'_\epsilon})(1 + \theta p_H)} + o(1),$$

and

$$g_{n+1}\theta(1 - p_H^2)^n(1 + v_n) = \frac{\theta(1 + v_n)}{(1 - \theta - (1-\theta)e^{-k'_\epsilon})}.$$

Therefore for any ϵ there exists a k'_ϵ large enough such that:

$$v_n \leq \epsilon$$

$$g_{n+1}\theta(1 - p_H^2)^n u_n \leq \frac{\theta^2}{(1-\theta)(1 + \theta p_H)} + \epsilon.$$

$$g_{n+1}\theta(1 - p_H^2)^n(1 + v_n) \leq \frac{\theta}{(1-\theta)} + \epsilon.$$

Let B be the matrix defined by:

$$B = \begin{pmatrix} \frac{\theta}{1+\theta p_H} & 0 & 0 \\ 0 & 0 & 1 \\ \frac{\theta^2}{(1-\theta)(1+\theta p_H)} & 0 & \frac{\theta}{(1-\theta)} \end{pmatrix}$$

Its eigenvalues are $\frac{\theta}{(1-\theta)}$, $\frac{\theta}{1+\theta p_H}$ and 0, which are all strictly less than 1. By the same perturbation argument as before, there exists an ϵ small enough such all the eigenvalues of A_n are also strictly less than $\lambda = \left(\max(\frac{\theta}{(1-\theta)}, \frac{\theta}{1+\theta p_H}) + 1\right)/2 < 1$.

Therefore for both cases, there exists a λ that depends only on θ and not p_H, an ϵ such that for $n^* = k'_\epsilon$ and for all $n > n^*$, the maximum eigenvalue λ_n of A_n is such that $|\lambda_n| \leq |\lambda| < 1$

F.3.2 Bounding the expected waiting time for H agents.

Let us consider $x_n = ||(\pi_{n,1}, \pi_{n,0}, \pi_{n+1,0})||_2^2 = \pi_{n,1}^2 + \pi_{n,0}^2 + \pi_{n+1,0}^2$. We have

$$(\pi_{n,0} + \pi_{n,1}) = \sqrt{\pi_{n,0}^2} + \sqrt{\pi_{n,1}^2} \leq \sqrt{\pi_{n,1}^2 + \pi_{n,0}^2 + \pi_{n+1,0}^2} + \sqrt{\pi_{n,1}^2 + \pi_{n,0}^2 + \pi_{n+1,0}^2} = 2\sqrt{x_n}. \tag{70}$$

Furthermore, for $n > n^*$, we have:

$$\begin{aligned} x_n &\leq |||A_n|||_2^2 x_{n-1} \\ &\leq \lambda^2 x_{n-1} \\ &\leq \lambda^{2(n-n^*)} x_{n^*} \end{aligned} \tag{71}$$

Therefore :

$$
\begin{aligned}
\mathbb{E}[N_{H,\infty}] &= \sum_{n_H=0}^{\infty} n_H \left(\pi_{n_H,0} + \pi_{n_H,1}\right) \\
&\leq \sum_{n_H=0}^{n^*} n_H \left(\pi_{n_H,0} + \pi_{n_H,1}\right) + \sum_{n_H=n^*}^{\infty} n_H \left(\pi_{n_H,0} + \pi_{n_H,1}\right) \\
&\leq n^* + \sum_{n=n^*}^{\infty} n 2\sqrt{x_n} \\
&\leq n^* + 2\sqrt{x_{n^*}} \sum_{n=n^*}^{\infty} n\lambda^{n-n^*} \\
&\leq n^* + 2\frac{\lambda}{\lambda^{n^*}} \left(\frac{\lambda^{n^*-1}(n^*(1-\lambda)+\lambda)}{(1-\lambda)^2}\right) \\
&\leq n^* \left(1 + \frac{2}{(1-\lambda)}\right) + o(n^*).
\end{aligned}
\tag{72}
$$

Where we successively used the splitting of an infinite sum, equation (70), equation (71), and the fact that $x_n^* \leq 1$. Note that λ only depends on θ and not on n^* or p_H. Using Little's law to derive the bound for w_H, this concludes our proof. Note that our n^* depends on the value of θ: for $\theta > 0.5$, we have $n^* = \frac{k_\epsilon}{p_H^2}$ whereas for $\theta < 0.5$ we have $n^* = \frac{k'_\epsilon}{p_H}$.

G Easy-to-match agents' waiting times

We report here simulation results for the average waiting times of easy-to-match agents under *BilateralMatch* and *ChainMatch* for uncorrelated arrivals as modeled in Section 4.1. Observe that the average waiting time of E agents are orders of magnitude lower than that of H agents. Under *ChainMatch* the average waiting of E agents increases with θ_H but has almost not effect from increasing θ_E. For the *BilateralMatch* however, while waiting times still seem to decrease approximately like $1/\theta_H$, we observe a sharp jump in waiting times when $\theta_E > \theta_H$.

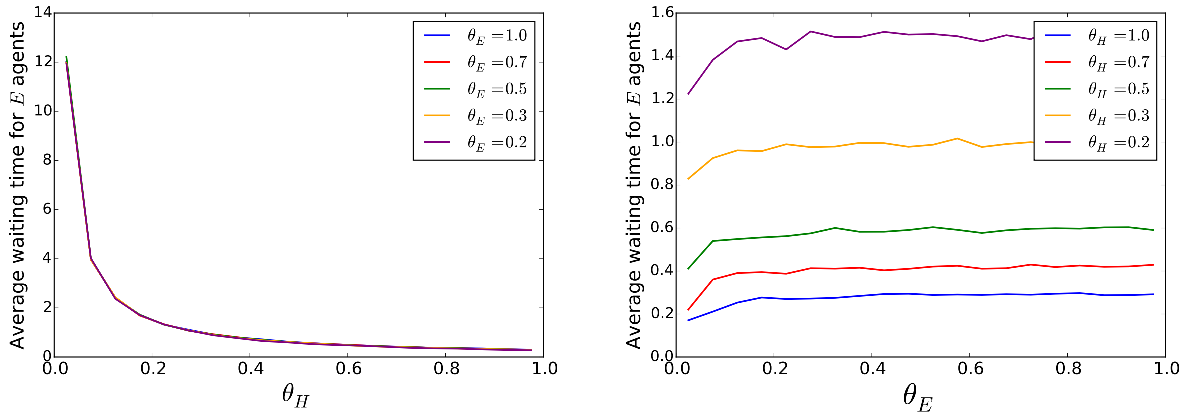

Figure 11: Simulations of a pool with independent arrivals of H and E agents, using *Chain-Match* and parameters $p_H = 0.02$, $p_E = 0.8$. Left panel shows the waiting times of E agents as a function of θ_H, for varying θ_E. Right panel shows waiting times for E agents as a function of θ_E, for varying θ_H .

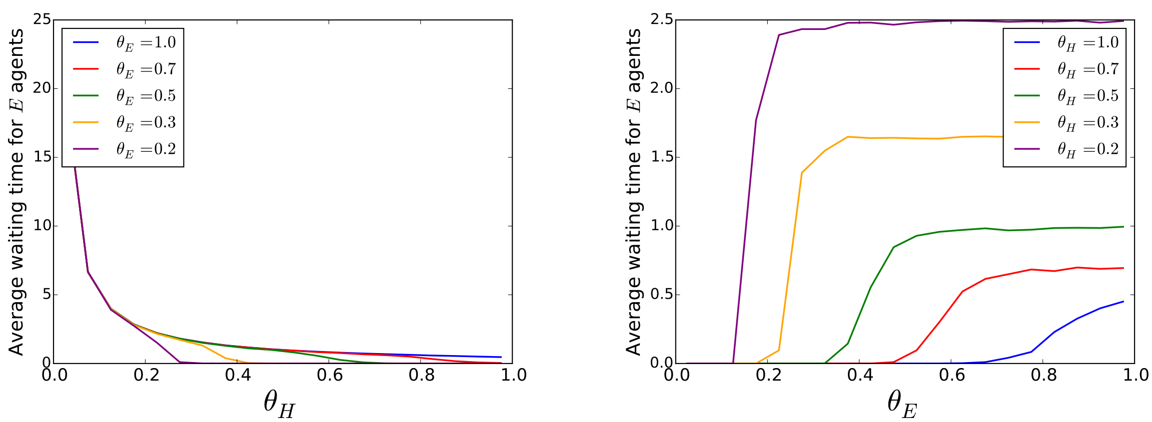

Figure 12: Simulations of a pool with independent arrivals of H and E agents, using *BilateralMatch* and parameters $p_H = 0.02$, $p_E = 0.8$. Left panel shows the waiting times of E agents as a function of θ_H, for varying θ_E. Right panel shows waiting times for E agents as a function of θ_E, for varying θ_H .