

Kidney Exchange in Dynamic Sparse Heterogenous Pools ^{*}

Itai Ashlagi Patrick Jaillet Vahideh H. Manshadi Michael A. Rees^{†‡}

January 10, 2014

Abstract

Kidney exchange pools are currently thin and sparse consisting of many highly sensitized patients. One way to increase matching opportunities for patients is by waiting for many incompatible pairs to join the pool. We analyze an algorithm that periodically finds allocations and study how the period length, as well as the technology, affects the number of matched patients. We find that if only 2-way cycles are conducted, matching online (as a new pair arrives) results in almost the same number of matches as the algorithm finds when the period length is “short”. However, if 3-way cycles are also allowed, we find regimes in which waiting for a short period considerably increases the number of matches that the algorithm finds. Finally, we show that when matching online, by allowing just a single non-simultaneous chain the number of matches will increase significantly. Our theoretical findings are supported by data-driven computational experiments.

1 Introduction

More than 630,000 people in the U.S currently suffer from kidney failure, known as End-Stage Renal Disease (ESRD) USRDS [2012], and more than 100,000 patients are on the waiting list for cadaver kidneys (see UNOS [2013]). Transplantation is not only the preferred treatment for patients (see e.g. Suthanthiran and Strom [1994]), but it is also much cheaper. Indeed, Medicare cost for a transplanted patient is approximately \$30,000 per year per patient while dialysis cost is approximately \$66,000.

The need for kidney exchange arises when a healthy person wishes to donate a kidney but is incompatible with her intended recipient (for a patient to receive a kidney, the donor and the patient should be blood type compatible and tissue-type compatible). Two or more incompatible

^{*}Research was partially funded by NSF (grant #1029603) and ONR (grant N00014-12-1-0033).

[†]Ashlagi: iashlagi@mit.edu. Jaillet: jaillet@mit.edu. Manshadi: manshadi@mit.edu. Rees: Michael.Rees2@utoledo.edu.

[‡]Dr. Rees is the medical director of the Alliance for Paired Donation. We thank participants in the NBER workshop on market design and the ACM Conference on Electronic Commerce (EC) for valuable suggestions and comments.

pairs can form a cyclic exchange so that each patient can receive a kidney from a compatible donor. In addition, a chain of exchanges can be initiated by a non-directed donor (an altruistic donor who does not designate a particular intended patient), in which case the sequence of exchanges does not close a cycle. Logistical constraints require that cycles involve no more than 3 patient-donor pairs since these should be conducted simultaneously (Roth et al. [2005b,a, 2004]).¹ However, chains can be done non-simultaneously as suggested by Roth et al. [2006], which led to the first non-simultaneous chain Rees et al. [2009].

Exchange pools currently experience a low arrival rate and further contain a large fraction of highly sensitized patients (Ashlagi et al. [2012]), who are very unlikely to find a tissue-type compatible donor.² One major decision clearinghouses are facing is how often to search for an allocation, i.e., a set of disjoint exchanges. On one hand, waiting is costly, and on the other hand matching too frequently may reduce the number of matched pairs, especially highly sensitized ones. Clearinghouses for kidney exchange currently use matching algorithms that periodically search for allocations that result in a (weighted) maximum number of matches.³ For example, the exchange program by the United Network for Organ Sharing (UNOS) searches for matches weekly while the National Kidney Registry (NKR) and the Alliance for Paired Donation (APD) search for matches almost daily. In this paper we study how the period length between finding allocations as well as the length of cycles and chains impact the number of matched patients over time. In particular, we analyze an algorithm, hereafter called *Chunk Matching (CM)*, which “waits” for a given number of incompatible pairs to arrive (a chunk) before it searches for an allocation in the existing pool. We analyze the performance of the algorithm for different chunk sizes.

In current exchange pools, most patients are either very highly sensitized or very low sensitized, and a large fraction of them are very highly sensitized. We focus on these two types and abstract away from blood-types⁴. We consider a discrete time model with n pairs that arrive sequentially, one pair at each time period. Each arriving pair is sampled from a bi-modal distribution independently. One way to think about this model is to view n as the number of pairs in a “relevant” horizon, considered to be the longest reasonable period of waiting. Since waiting is costly for patients, we mainly focus on analyzing the *CM* algorithm when the chunk size is small. We compare the performance of the *CM* algorithm with “short waiting” periods/small chunks to the online scenario where chunks are of size one. In other words, our goal is to find whether the chance a typical pair to match will increase by using short waiting periods rather than matching online. In Subsection 5 we discuss how to generalize our model to incorporate various blood-types as well as how to extend it to an infinite horizon framework.

¹Cycles are conducted simultaneously since it is required that a donor does not donate her kidney before her associated patient receives a kidney.

²This is partially since hospitals match internally their easy-to-match pairs and only register their hard-to-match ones Ashlagi and Roth [2012].

³Weights typically favor highly sensitized patients.

⁴Equivalently we assume that all pairs in the pool are blood-type compatible.

We first study the setting in which allocations include only cycles of length 2. We show that if the waiting period between two subsequent match runs is a sublinear function of n , then waiting with easy-to-match pairs (or waiting with all pairs) results in approximately the same number of matches as when matching in an online fashion (Theorem 3.1 and Corollary 3.2). This implies that “short” waiting period is not effective when only 2-way cycles are conducted.

We then analyze *CM* when cycles of length 2 and 3 are allowed. We show that for some regimes, sub-linear waiting with easy-to-match pairs will result in a linear addition of matches comparing to the online scenario (Theorem 3.3 part (a)). However, if in addition we wait with the highly sensitize patients, we will not gain significantly (Theorem 3.3 part (b)). This result highlights the importance of modeling heterogeneity and using different strategies for different types. From a design perspective, it implies that some waiting with the easy-to-match pairs can increase the number of matches. Particular, when an easy-to-match pair arrives that is compatible with a hard-to-match patient in the pool, some waiting increases the likelihood that these two pairs can be part of a 3-way cycle.

In recent years, there have been a growing number of non-directed donors (NDDs) who registered in the kidney exchange programs without having intended patients. This allows kidney exchange programs to organize chains of exchanges that start with NDDs. In practice, chains proved to be quite effective and only in the two data sets that we have access to, more than 180 chains have been arranged in the past few years and around 700 patients have been matched through chains. A chain can be organized non-simultaneously; when an NDD donates a kidney it creates a sequence of exchanges; the last pair in the sequence only receives a kidney without donating one. Thus the donor of the last pair becomes a bridge donor who continues the chain in a later period. Even though bridge donors can renege, very few cases of withdrawal have been recorded (3 cases in our data sets).

As chains have become very effective, it is important to study their benefit and analyze the efficiency of dynamic matching with chains. A major difficulty with chains is that they can be of arbitrary length.⁵ We compare online matching with or without chains and show that adding one non-directed donor will increase linearly the number of matches that the *CM* algorithm will find over the number of matches it will find without a chain (Theorem 3.5). This complements a result by Ashlagi et al. [2011], who showed that chains are powerful in static pools. From a design perspective, our analysis highlights the important role of easy-to-match pairs in continuing the chain. Intuitively, because easy-to-match pairs are more likely to connect to the bridge donor they should be used to start a new sequence of exchanges.

Waiting for “long” periods of time, may not be practical for kidney exchange. For completeness, however, we analyze *CM* when the chunk size is of linear size (in n). We focus on matching using cycles of length 2 and compare large waiting periods with the online solution as well as the hindsight

⁵Note that a chain can be conducted non-simultaneously while keeping the restriction that every patient receives a kidney before her associated donor gives one.

one. We first show that linear waiting (with either easy-to-match pairs or all pairs) results in a linear increase in the number of matches compared to the online scenario (See Theorem 4.1). In addition, we compare *CM* with large chunk sizes to the hindsight solution, in which we only run the matching once after all the pairs arrived, and show that the gap between any chunk matching with chunk size less than n and the hindsight solution is linear in n (Theorem 4.2).

In all our results in which waiting (or adding chains) proves to be effective, the additional matches correspond to pairs with highly sensitized patients. Almost all pairs with low sensitized patients will be matched regardless of the size of the chunk (and the technology that we use) in each match run. These findings explain computational simulations using clinical data (Figures 3 and 4).

Our results may be of independent interest to the literature on dynamic matching in random graphs. Kidney exchange serves well as an example for which we have distributional information on the underlying graphs, thus we can exploit this information to make analysis and prediction far more accurate than the worst-case analysis can do. We believe our average-case analysis can have implications beyond the kidney exchange and can be applied to other dynamic allocation problems with such distributional information.

While this paper focuses on kidney exchange, there are many dynamic markets for barter exchange for which our findings apply. There is a growing number of websites that accommodate a marketplace for exchange of goods (often more than 2 goods), e.g. ReadItSwapIt.com and Swap.com. In these markets, the demand for goods, cycle lengths and waiting times play a significant role in efficiency.

1.1 Related work

Roth et al. [2004] first proposed a way to organize kidney exchange integrating cycles and chains. Logistical constraints required that cycles involve no more than 2 patient-donor pairs. Subsequent work suggested that a modest expand of infrastructure, that is allowing only slightly larger, 3- and 4-way exchanges would be efficient (Roth et al. [2007], Ashlagi and Roth [2013]) in large static pools. These studies assume either implicitly or explicitly that no tissue-type incompatibilities exist.

Ünver [2010] initiates the study of dynamic kidney exchange. He shows a closely related result to the static case under the assumption that no tissue-type incompatibilities exist. In particular, waiting is not an issue when only blood type compatibilities are considered. Our work deviates from his model significantly by abstracting away from blood types and focusing on the tissue-type compatibility and the sparseness of the compatibility graph. Further, our approach to study dynamic kidney exchange is combinatorial and is based on the structure of the underlying random graph while Ünver [2010] takes a dynamic programming approach.

Awasthi and Sandholm [2009] use online stochastic optimization methods to computationally

study the allocation in kidney exchange using short cycles in a dynamic setting. Dickerson et al. [2012a] conduct computational simulations in the dynamic settings to understand the benefit of chains. Dickerson et al. [2012b] study dynamic optimization and propose an algorithm that assigns weights to different matches using future stochastic sampling. These studies use however dense compatibility graphs (they generate graphs according to Saidman et al. [2006] which have been shown in Ashlagi et al. [2012] to be very dense as opposed to clinical data).

In a related line of work, Zenios [2002] studied a dynamic kidney allocation model where both direct (2-way) and indirect kidney exchange are allowed.⁶ The objective is to study the trade-off between the (lower) quality of a deceased donor (that is assumed to be available immediately) and the cost of waiting for a direct exchange. The exchange model studied in this work ignores the tissue-type compatibility.

Another line of work in kidney exchange studies matching in the presence of failure. Many of the match offers given by the exchange programs fail to proceed to actual transplant for various logistical and medical reasons. A few recent papers study kidney exchange with failures, Dickerson et al. [2013], Blum et al. [2013], Molinaro and Ravi [2013]. They use a probabilistic model for failure and assume that each compatible match is only viable with a given probability. They design algorithms to maximize the expected number of successful matches in static settings. Further, Dickerson et al. [2013] computationally study this problem in a dynamic setting using dense graph models.

The problem of online matching (equivalent to our online scenario with only two-ways) arises naturally in information technology applications such as online advertising in which advertisements need to be assigned instantly to queries searched or webpages viewed by users. The study of online matching was initiated by Karp et al. [1990], in which they analyze the problem in adversarial settings with no probabilistic information about the graph. Several follow up papers, studied the problem in settings that limit the power of the adversary. Goel and Mehta [2008] studied the model in which the underlying graph has unknown distribution. Feldman et al. [2009] noticed that in applications such as online advertising there is information about the graph structure, and they analyzed a model where the graph distribution belongs to a certain class. Manshadi et al. [2011], Jaillet and Lu [2013] studied the same problem with a general known distribution. Note that here we focus on one special class of distributions; however, unlike the computer science literature, we consider various regimes of waiting (and not just the online scenario).

Mendelson [1982] analyzed the behavior of a clearinghouse in a dynamic market with prices in which sellers and buyers arrive over time according to a given stochastic process. Similar to our work, he considers a mechanism in which the clearing prices are computed periodically, and he studies the market behavior for different time (period) scales.

⁶In an indirect exchange, the donor gives a kidney to the highest priority patient in deceased donor waiting list, and her intended patient gets the highest priority in the waiting list.

2 Dynamic compatibility graphs and empirical findings

In a kidney exchange pool there are patients with kidney failure, each associated with an incompatible living donor, and non-directed donors (NDDs).⁷ The set of incompatible pairs and NDDs in the pool, V , induces a **compatibility graph** where a directed edge from v_1 to v_2 exists if and only if the donor of pair v_1 is compatible with the patient of pair v_2 .⁸

A k -way cycle is a directed cycle in the graph involving k pairs. A chain is a directed path starting from an NDD. A k -way allocation or a k -way matching is a set of disjoint cycles each of size at most k . In practice, cycles of size at most 3 are considered due to incentive and logistic reasons. On the other hand, chains can be arbitrarily long.

In a *dynamic compatibility graph* the pairs (nodes) arrive sequentially one at a time, and at each time step a centralized program can decide on an allocation and remove the participating nodes in that allocation from the graph. In this paper, we analyze an algorithm that finds a maximum allocation every given number of periods (the algorithm is described in detail in Section 3). Such an algorithm is used in practice and a question faced by centralized programs is how often to search for an allocation. In the next section, we discuss empirical findings that will motivate our modeling assumptions.

2.1 Empirical findings

As opposed to earlier studies that focused on blood types and ignored market size and sensitization of patients, Ashlagi et al. [2012] have shown, using historical data from the Alliance for Paired Donation (APD), that sensitization of patients plays a crucial role in the number of matches that can be achieved. Each patient has a level of percentage reactive antibodies (PRA) that captures how likely it is that the patient will not match a random blood-type compatible donor in the population. The lower the PRA of a patient, the more likely he will match a random donor. Ashlagi et al. [2012] find that the percentage of high PRA (PRA above 80%) in the pool is significantly higher than what previous studies have assumed to support earlier theoretical findings (see also Saidman et al. [2006] and Roth et al. [2007] for such simulations). They further find that among patients that have high PRA the average PRA is above 95.

In this paper, we also study the historical data from the APD in a dynamic setting. Over a period of 3 years, 756 pairs have enrolled in the APD program; thus we can approximate the daily arrival rate as 0.7 pair per day. Figure 1 shows the PRA distribution of these 756 pairs. Note that most pairs have either very high PRA (above 95) or relatively low PRA (in fact more than 40% of the patients with PRA above 95 have PRA above 99). This motivates us to use a

⁷Pairs that are compatible would presently go directly to transplantation and not join the exchange pool although Roth et al. [2005a] and Sönmez and Ünver [2011] study the advantage of adding such pairs to the pool.

⁸In practice a minority of patients enroll with multiple donors. One can extend the model appropriately to capture this multiplicity.

bimodal distribution for the PRA, in which we categorize the pairs into two types: type H (L) which represents pairs with high (low) PRA patients.

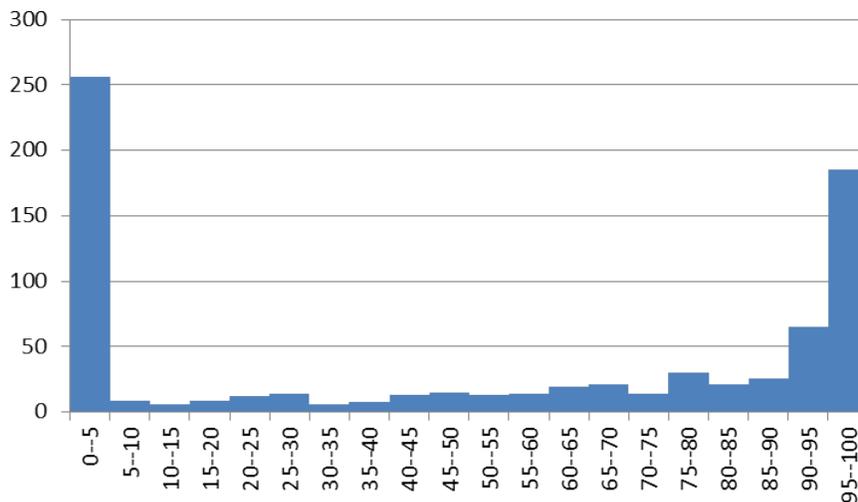


Figure 1: PRA histogram of the patients in the APD pool.

We note that the PRA distribution in the historical data, is significantly different from the one in the general population, because the pairs joining the pool are not uniformly random samples of the population. Most blood-type compatible pairs (like a pair with O-patient and O-donor) that join the pool have highly sensitized patients. Another reason that accounts for this biased sampling is that many hospitals conduct internal exchanges, which often involves easy-to-match patients and thus enter mostly their hard-to-match patients into the exchange program (see Ashlagi and Roth [2012] for more details).

In Figure 2, we present the “historical” in-degree distribution of the A-patients (and O-patients) that have high PRA ($\text{PRA} > 95$). For every donor and every patient compatibility can be determined using their medical characteristics (blood type, antibodies, antigens) even if they have not been present in the pool simultaneously. Observe that most pairs have very small in-degree. Note that we assume here that all pairs are present in the pool simultaneously, which is certainly not the case. For instance if pair v_1 arrives in the first month of the first year, it may get matched and leave the pool before the end of the first year, and thus a pair v_2 that joins in the second year will not see v_1 . Therefore, in the “actual” pool, the in-degree of these pairs are even smaller. Finally note that if a pair with PRA of 99 does not see any compatible donor at the time that it joins the pool, in expectation, it needs to wait for 100 new blood-type compatible pairs to “see” one tissue-type compatible donor. If the arrival rate is roughly 0.7 pairs per day, this will take 6 months to a year (depending on the blood type).

Before we proceed to our empirical results in the dynamic setting, we note that pairs do not stay in the program indefinitely; they may leave the pool for several reasons such as getting a

transplant elsewhere (another program or a cadaver), getting too sick to be transplanted, or dying. Currently the estimated departure rate (based on the APD data) is around 0.02% per month. In what follows, we limit the time horizon to 2 years, and study the effect of waiting in this time horizon.

We provide here initial computational experiments when matching over time using the historical data. We test how many matches are obtained when an allocation is found after every x pairs join the pool. For each scenario we conduct 100 trials, in which we permute the order in which the pairs arrive. Figure 3 plots the average number of pairs matched under different waiting periods for 2-way and 3-way matchings. Further, in our last scenario (3-ways and chain), we add an NDD to the pool at time 0, and after x periods, we greedily search for cycles of length at most 3 as well as a directed path that connects to the NDD (i.e, the first patient in the path is compatible with the NDD).⁹ If the solution consists of such a path (which we call a segment of the chain), the donor of the last pair in the directed path serves as the *bridge donor* to continue the chain in future matching. We repeat the same at time $2x$, and so forth. Figure 4 is similar only counting the number of highly sensitized patients that were matched.

Note that the x-axis of these figures is logarithmic. Also, note that x is the number of arrivals (and not the time) between any two consecutive matching. For instance, when x is 32, it means that the estimated waiting time between two match runs is about 45 days.

First note that in the course of around 2 years, the total number of pairs that the online algorithm matches is significantly less than what can be matched in the offline scenario (when waiting for 2 years). This is mainly due to the high sensitization of a large fraction of the patients in the pool, which makes it very unlikely to match with a blood-type compatible donor. Most of these patients will see only a few donors in a “relevant” time horizon they can wait for a kidney. In online matching, we may allocated “mistakenly” the sole compatible donor currently available for a highly sensitized patient (say a patient with PRA of 0.99) to another patient. Such a highly sensitized patient may not see any other compatible donor in this time horizon.

Further, note that when matching in an online fashion, allowing cycles of length 3 is quite effective, and increases the number of matches significantly compared to the case when we only allow 2-way exchanges. Further, if we add even a single non-simultaneous chain (by adding one altruistic donor at the beginning), we will match many more pairs.

Even though a significant increase in the number of matches occurs when the waiting period is “large”, such long periods of waiting do not seem practical. However, it might be possible to make some pairs wait for a “short” period of time. Our experiments suggest that short waiting is not effective in 2-way matching, but it does improve the number of allocations when using 3-ways as well. This is again another benefit of using 3-way cycles.

Observe from Figures 3 and 4 that most of the gains come from matching more high PRA patients. This is because matching low PRA patients is easy and all of these schemes match almost

⁹In breaking ties, we give priority to the pair with the highest PRA.

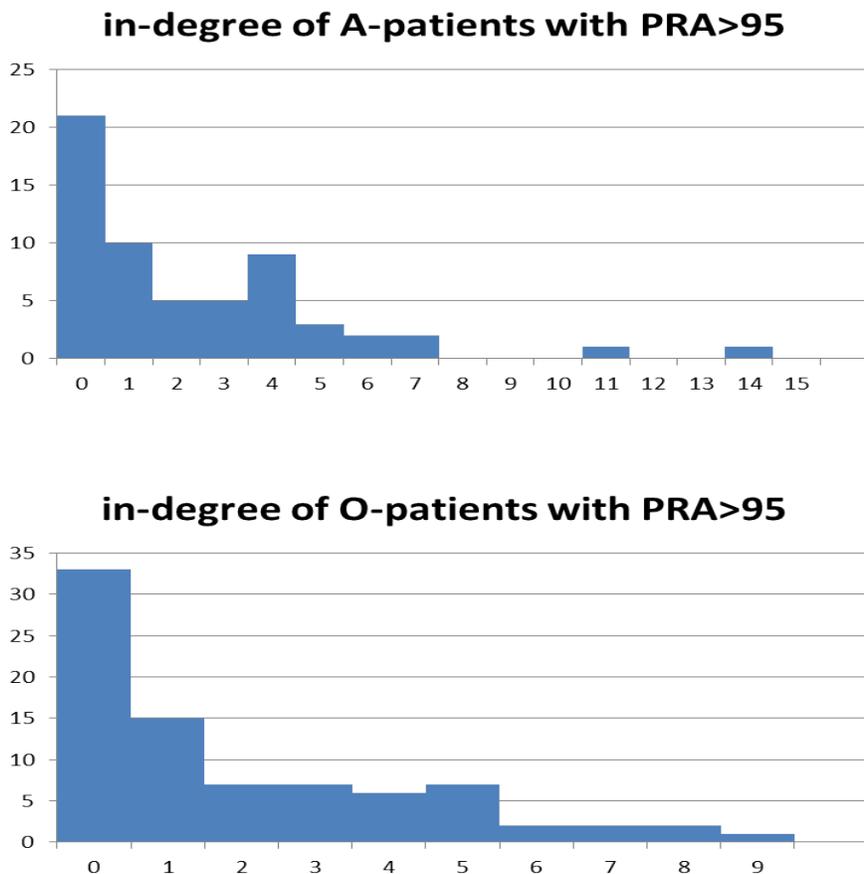


Figure 2: Top: In-degree histogram of patients with blood type A and PRA > 95; Bottom: In-degree histogram of patients with blood type O and PRA > 95.

all low PRA patients.

In this paper, we provide theoretical foundations that explain these behaviors in dynamic matchings in sparse compatibility graphs. We first focus on the results that have policy implications. In particular, we first prove that short waiting is not helpful in 2-ways, but using 3-ways, in some regimes, helps matching more high PRA patients. Also, we prove that even when using 3-ways, short waiting with the high PRA pairs is not effective. This implies that when only 2-way cycles are possible matching online would be better than matching after waiting for a short period in the sense that it matches almost the same number, but its average waiting is smaller. On the other hand, if 3-ways are possible ¹⁰, making low PRA pairs wait for a moderately short period would crucially help high PRA patients and increase the number of matchings. Further, under online schemes, we prove that using even a single chain is quite effective and results in matching significantly more high PRA patients. Finally, for the sake of completeness, we prove that waiting for a “long” period

¹⁰Note that in some parts of the pool 3-way matching is not possible because of the blood-type compatibility.

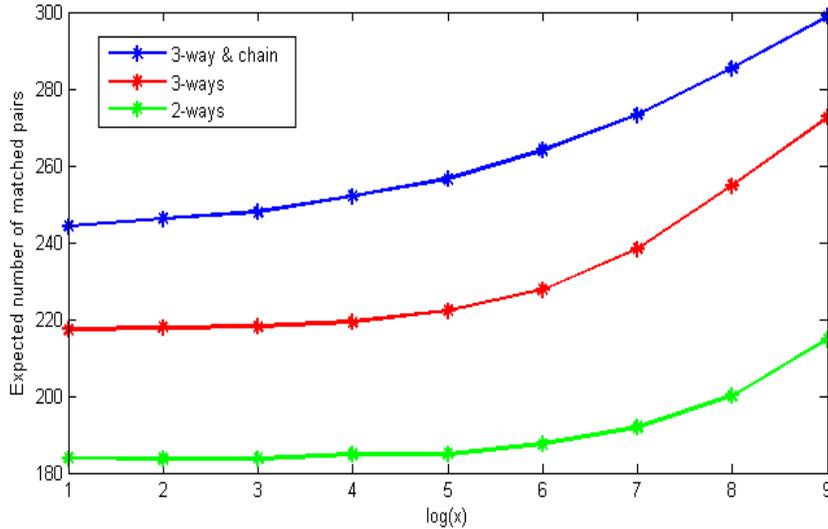


Figure 3: Number of matched pairs vs. waiting for x new patients to arrive.

of time increases the number of matches that we can perform in a fixed time horizon.

In the rest of this section, we describe a dynamic model for the current kidney exchange motivated by the findings above. We use this model to prove theoretical results that explain the above observations.

2.2 A dynamic random compatibility graph

In a *dynamic kidney exchange graph*, there are n patient/donor incompatible pairs which arrive sequentially at times $t = 1, 2, \dots, n$.¹¹ Each pair corresponds to a node in the graph. Each node is one of two types, L (low PRA) or H (high PRA) capturing whether the patient of that node is easy- or hard-to-match. Each arriving node is of type H independently and with probability $0 \leq \rho \leq 1$. When joining the pool, the arriving node i forms directed edges to the existing nodes. If node i is of type H (L), it forms an incoming directed edge with any of the existing nodes independently with probability p_H (p_L). Further, it forms outgoing directed edges to each L-node (H-node) independently with probability p_L (p_H) (See Figure 5). At each time step, there is an underlying compatibility graph and the centralized program can find an allocation and remove the participating nodes in that allocation from the graph. For the case of $k = 2$ it will be convenient to reduce cycles of length two to undirected edges and remove the rest of the directed edges from the graph. Allocations with $k = 2$ are just matchings in the reduced graph.

¹¹Without loss of generality, assume n is a power of 2.

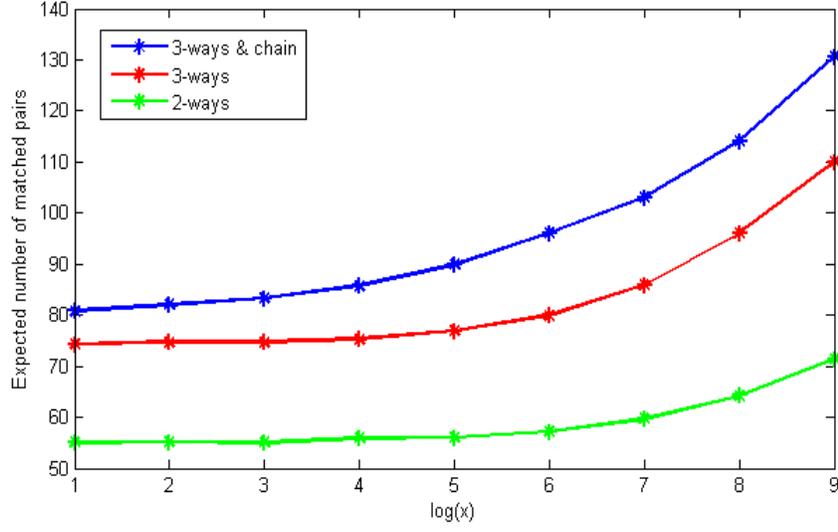


Figure 4: Number of highly sensitized patients (PRA > 80) matched vs. waiting for x new patients to arrive

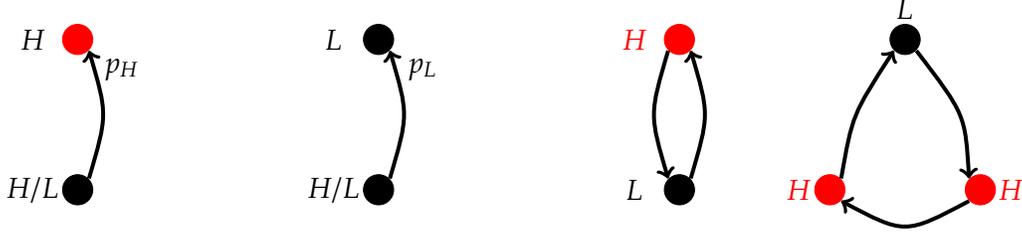


Figure 5: Directed edge formation in the dynamic kidney exchange; An H-L 2-way cycle (or, alternatively, an undirected edge); An H-H-L 3-way cycle.

2.3 Pool thinness

To further develop our model we use the following lemma which provides lower/upper bounds on the expected size of the matching obtained using an online greedy scheme.

Lemma 2.1. *Let M be the size of matching obtained using an online greedy scheme, which finds a maximum number of matches (2-way cycles) after each node's arrival.*

(a) **Lower bound:** $\frac{\mathbb{E}[M]}{n} \geq 1 - \frac{1}{np_H^4}$.

(b) **Upper bound:** $\frac{\mathbb{E}[M]}{n} \leq 1 - \rho(1 - p_H p_L)^{(1-\rho)n} (1 - p_H^2)^{\rho n - 1}$.

Lemma 2.1 is proven in Appendix A. Let us elaborate on the implications of this lemma by putting some numbers. Let $p_H = 0.02$, $p_L = 0.98$, and $\rho = 0.5$, which are reasonably consistent with our data on PRA distribution (Figure 1). Now suppose we have a pool of size 3×10^9 , then

the lower bound tells us that, in expectation, more than 99.7% of the pairs will be matched in online matching. On the other hand, for a pool of size 300, the upper bound tells us that in expectation online matching cannot match more than 97.6% of the pairs. However, for the same pool size if p_H was 0.9, then online matching would match at least 99.4% of the pairs. Thus if we had a large pool or a moderate-size pool consisting of low PRA patients, then online matching would be “optimal.” This is related to the result by Ünver [2010] which assumes no tissue-type incompatibilities. However, the parameters in the current pools are closer to those of our second example; we have pools of size around 300,¹² with a large fraction of pairs with PRA around 0.99 (i.e., $n = \Theta(1/p_H)$). Thus as our experiments and the above lemma suggest, current pools are too “thin” in the sense that online matching cannot match most of the pairs in a relevant time horizon (say 2 years), and waiting longer will result in considerably more matches. In addition, random graph results imply that in a large dense graph all blood-type compatible pairs can be matched to each other using only 2-way cycles. As has been seen in Ashlagi et al. [2012], this is not the case (see for example Figure 2 in their paper).

To incorporate these findings, and for ease of presentation we model the pool size by letting $n = \frac{c}{p_H}$ where $c > 0$ or equivalently setting $p_H = c/n$.

Remarks:

1. In practice p_H should only depend on medical characteristics of the patient regardless of the population size. Setting p_H to be a small number may seem to be a reasonable assumption. However, in our “relevant” horizon we observe only a small number of pairs to arrive, approximately $\Theta(1/p_H)$, which brings us to the proposed model, linking n and $\Theta(1/p_H)$. For further discussion see Ashlagi et al. [2012] from which we adopt these probabilistic assumptions. Our model also abstracts away from blood type compatibilities and focuses on the sensitivity of patients as the sensitivity of patients is of first order importance in maximizing the number of matches in current pools.
2. We do not attempt to model the behavior of how and which pairs join the pool. Sparsity is assumed simply since the pool is not large enough.

3 Main Results

We analyze a simple greedy algorithm termed *k-way Chunk Matching* (CM^k) which finds allocations using cycles of length at most k each time a given number of new pairs have joined the pool. CM^k receives as input two chunk sizes, S_H and S_L , that determine the waiting times before making decisions. In particular, after the arrival of S_H new nodes, it finds a maximum k -way allocation in the graph ignoring the directed edges between L-nodes. After receiving S_L/S_H chunks each of size

¹²This estimate is the typical size of a blood-type compatible subgraph and not the whole pool.

S_H , it also finds a maximum k -way allocation including the L-L directed edges. Note that even if $S_H = S_L$, we still slightly favor the H-nodes (we first find matching in the graph without directed edges between L-nodes and then we consider the entire graph), trying to compensate for the fact that they have fewer options and so are harder to match. Also, note that once we include the L-L directed edges, most of the unmatched L-nodes will be matched because they can easily form 2-ways or 3-ways with each other. Thus allowing $S_H < S_L$ does not only give priority to H-nodes, but also provides H-nodes more matching opportunities by letting L-nodes (that can be easily matched fast) wait longer.

We next formalize CM^k as Algorithm 1. At any time t , let G_t denote the residual graph with only unmatched pairs.

Algorithm 1 k -way Chunk Matching (CM^k)

1: Let S_L be a divisor of n and S_H be a divisor of S_L ;

For $\theta = S_L, 2S_L, \dots, n$:

For $z = (\theta - S_L) + S_H, (\theta - S_L) + 2S_H, \dots, \theta$:

2: Find the k -way matching on graph G_z , **ignoring** the L-L directed edges and breaking ties arbitrarily.

3: Remove the matched nodes.

End for

4: Find the k -way matching on graph G_θ breaking ties arbitrarily.

5: Remove the matched nodes.

End for

Let $M^k(S_H, S_L)$ denote the number of matches obtained by CM^k with chunk sizes of S_H and S_L . We will call the *online scenario* the case in which $S_L = S_H = 1$.

We analyze CM^k for different regimes of waiting with H- and L-nodes. First, in the next subsection, we focus on the moderate waiting regime which is of practical interest. For completeness in Section 4, we study CM^k with large chunk sizes, and compare it to the online solution and to the hindsight one.

In Subsection 3.2, we compare the online matching defined above with an online scheme that uses both cycles and chains and prove that using even a single chain results in a significant increase in the number of allocations.

3.1 Moderate waiting with short cycles

In Theorem 3.1, we show that when restricting to 2-way cycles, waiting for chunks of sublinear size will not improve significantly the number of matches over the online scenario. This result holds when either waiting with only L-nodes (part a), or waiting with both L- and H-nodes (part b). As we shall see in Theorem 3.3, this is not the case when 3-way cycles are allowed.

Theorem 3.1 (Sublinear waiting with 2-ways). *Suppose $S = n^{1-\epsilon}$ for some $0 < \epsilon < 1$ and S is a divisor of n .*

(a) **Waiting with L-nodes:** $\mathbb{E} [M^2(1, S)] \leq \mathbb{E} [M^2(1, 1)] + o(n)$.

(b) **Waiting with H-nodes:** $\mathbb{E} [M^2(S, S)] \leq \mathbb{E} [M^2(1, S)] + o(n)$.

The detailed proof of this theorem is presented in Appendix B. Here we only describe the high level ideas of the proof.

Proof idea of Theorem 3.1. The proof of part (a) follows this simple observation: the only difference of online chunk matching ($S_H = S_L = 1$) and the one with $S_H = 1$ and $S_L = S$ is that, in a middle of a chunk, the former may match an arriving L-node v_1 that cannot match any existing H-node to another L-node. However, the latter that sees the whole chunk before considering L-L edges, may see an H-node v_2 that arrives after v_1 in the same chunk who can be matched to v_1 . In such a case, online matching may lose the matching opportunity of v_2 . However, the likelihood of such an event is quite small, (because the probability that v_1 is compatible with v_2 is very small), and thus the total loss due to these mistakes is $o(n)$.

The intuition of the proof of part (b) is as follows: after each chunk arrives, and after removing the matched nodes, the residual graph (before the next chunk arrives) has no remaining undirected edges (2-ways)¹³. Now suppose S new nodes arrive and form undirected edges. Excluding the L-L edges, the resulting graph after these arrivals will contain at most $O(S) = o(n)$ undirected edges and thus is extremely sparse and disconnected. Further, it is very unlikely that two H-nodes form a 2-way, and thus almost all of these edges are H-L 2-ways. This graph consists of $O(S)$ connected components each of size $O(1)$; we show that with high probability, each of these components is a tree with depth one with an L-node as a root (See Figure 6). The maximum matching in a disconnected graph is the union of the maximum matching of each of its connected components. Thus without loss of generality, the CM^2 with $S_H = 1$, and $S_L = S$ will also find the maximum matching in each of these components separately. For instance consider the example of Figure 6; when r_1 arrives, it forms its three undirected edges. Now since, w.h.p., nodes c_1 , c_2 , and c_3 will not have any other neighbors in this arriving chunk (the filled nodes in Figure 6), the decision of an CM^2 with $S_H = 1$, and $S_L = S$ would be the same as that of CM^2 with $S_H = S_L = S$. Further, after including the L-L edges, with high probability, both algorithms will match all the remaining L-nodes to each other. \square

Part (a) and (b) of Theorem 3.1 together imply that:

Corollary 3.2. *Suppose $S = n^{1-\epsilon}$ for some $0 < \epsilon < 1$ and S is a divisor of n .*

$$\mathbb{E} [M^2(S, S)] \leq \mathbb{E} [M^2(1, 1)] + o(n).$$

¹³Since only 2-ways are allowed, we can ignore any directed edge that does not form a 2-way, and the allocation problem is simply a matching problem.

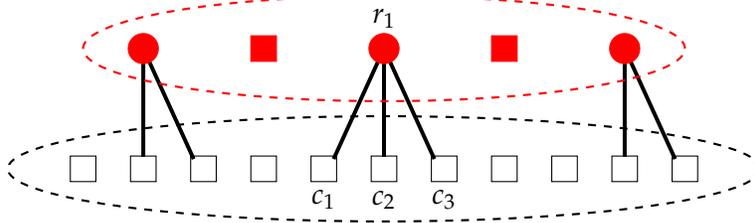


Figure 6: The typical connected components (excluding the L-L edges) when the chunk is of a sublinear size; the circle nodes are of type L and the rectangle ones are of type H; the filled nodes are the ones which arrived in the last chunk, and the not-filled nodes are those which arrived in the previous chunks, but which have not been matched yet.

Observe that even though the online scenario has the worst performance, it still matches $\Theta(n)$ nodes; indeed it finds a maximal matching, and the size of a maximal matching is at least half of the maximum. Formally, there exists $\delta > 0$ such that: $\mathbb{E}[M^2(1, 1)] \geq \delta n$.

We next analyze the CM^3 algorithm that allows for 3-way cycles as well. In particular, we state the counterpart of the above theorem when 3-way cycles are allowed, and show that unlike the 2-way matching (where waiting moderately with L-nodes is not effective), if L-nodes wait sublinearly before matching (to other L-nodes), then the number of allocations increases significantly compare to the online 3-way matching.

Theorem 3.3 (Sublinear waiting with 3-ways). *Suppose $S = n^{1-\epsilon}$ for some $0 < \epsilon < 1$ and S is a divisor of n .*

(a) **Waiting with L-nodes:** *If the parameters $0 < \rho < 1$, p , and c satisfy the following condition:*

$$(1-p)(1-\rho)ce^{-c(1+2\rho)} - p(1-e^{-c\rho})(1-c(1-\rho)e^{-c} - e^{-c(1-\rho)}) \geq 2\delta, \quad (1)$$

where $\delta > 0$ is a constant, then:

$$\mathbb{E}[M^3(1, S)] \geq \mathbb{E}[M^3(1, 1)] + \delta n.$$

(b) **Waiting with H-nodes:** $\mathbb{E}[M^3(S, S)] \leq \mathbb{E}[M^3(1, S)] + o(n)$.

The main challenge in the analysis of CM^3 is that now we have a directed graph, and the residual graph does not consist of only isolated nodes anymore (as opposed to the analysis of CM^2); it can contain many edges and even cycles of length greater than 3. However, the residual graph contains mainly H-nodes; similar to the proof of Lemma 2.1, we can show that at the beginning of each chunk the expected number of L-nodes in the residual graph is $O(1)$. Thus to compare the number of allocations of the two schemes it suffices to compare the number of H-nodes they match. The proof for Theorem 3.3 is given in Appendix B. Here, we only outline the main ideas.

Proof idea of Theorem 3.3. The intuition for part (a) is as follows: In the online scenario, in many occasions, there will be a directed edge from an (arriving) L-node v to an (existing) H-node u , but u is not part of a cycle at that time. In fact, there are (linearly) many such v and u nodes such that v does not have a directed edge to any other H-node in the graph. Since v is easy-to-match, the online scenario will “quickly” find another cycle for the L-node v and the H-node u node will remain unmatched. However, under chunk matching, v will have to wait and since it is an L-node, it will be relatively “easy” to close a 3-way cycle with u, v and another L-node arriving in the same chunk. The proof deals with various subtleties such as “harming” other H-nodes by matching node u too early. Figure 7 shows a sample of the set of (c, ρ) parameters that satisfies Condition (1) for $p = 0.8$ and $\delta = 0.01$.

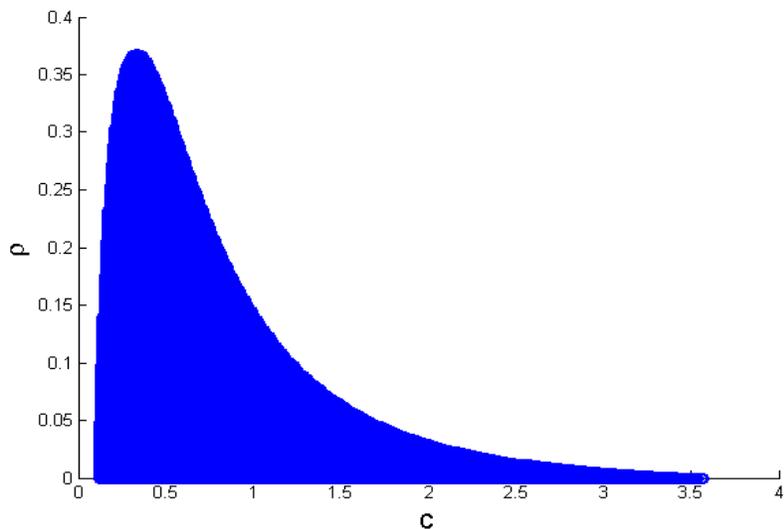


Figure 7: The (c, ρ) region satisfying Condition (1) for $p = 0.8$ and $\delta = 0.005$.

To prove part (b), similar to proof of part (b) of Theorem 3.1, we show that if we exclude the L-L edges, the graph after a new chunk arrives is very disconnected; thus the decisions of CM^3 with $S_H = S$ and those of CM^3 with $S_H = 1$ result in almost the same number of matchings. A formal proof is given in Appendix B. \square

The following corollary is a direct implication of Theorem 3.3. Similar to corollary 3.2, we compare the CM^3 with equal chunk sizes $S_H = S_L = S$ to the online scenario $S_H = S_L = 1$; however, here, for a certain set of parameters (ρ , p , and c), the gain of waiting for S steps in 3-way matching is linear in n as opposed to the 2-way matching where the gain is $o(n)$.

Corollary 3.4. *Let $S = n^{1-\epsilon}$ for some $0 < \epsilon < 1$ where S is a divisor of n . If parameters $0 < \rho < 1$,*

p , and c satisfy Condition (1), then, there exists constant $\delta' > 0$ such that:

$$\mathbb{E} [M^3(S, S)] \geq \mathbb{E} [M^3(1, 1)] + \delta' n.$$

3.2 Online matching with chains

In this section, we add to the model an altruistic donor to the pool at time $t = 0$ and analyze how the addition of a single non-simultaneous chain will affect the number of matches. In dynamic matching with a chain, each time period that a chain segment is found, it will be matched and removed from the pool. Furthermore, the donor of the last pair in the chain becomes a *bridge donor* (BD) for the next segment.

Here, we only consider the online scenario and analyze the following scheme: after each new node arrives, we try to match it through a cycle of length at most k or by adding it to the chain according to the following rules: (i) the bridge donor (last pair in the chain) must be of type H and (ii) if the arriving node can form a k -way cycle with at least $k - 1$ nodes of type H and can also form a path (of any length) connected to the BD, we break the tie in favor of the k -way. We denote such an online scheme by $CM^{k,c}(1, 1)$ and we denote the online k -way matching without chain by $CM^k(1, 1)$.

Observe that under $CM^{k,c}(1, 1)$ and $CM^k(1, 1)$ the residual graph does not contain any cycles with length smaller than or equal to k (otherwise we would have performed such a cycle). Our main results compare the performance of the greedy online matching with or without a chain ($CM^{k,c}(1, 1)$ and $CM^k(1, 1)$):

Theorem 3.5 (Online allocation with or without a chain). *Consider the model given in 2.2; suppose we have one altruistic donor at time 0;*

- (a) *Suppose $\rho = 1$, i.e. all nodes are of type H, and $k = 3$; in expectation, $CM^{3,c}(1, 1)$ matches $\Theta(n)$ more nodes than $CM^3(1, 1)$.*
- (b) *Suppose $0 < \rho \leq 1$, but $k = 2$. In expectation, $CM^{2,c}(1, 1)$ matches $\Theta(n)$ more nodes than $CM^2(1, 1)$.*

The proof of Theorem 3.5 is presented in Appendix C. Interestingly, the proof with L-nodes (part (b)) and without L-nodes (part (a)) are very different. For example, in the first part without any L-nodes the bridge donor is essentially “forced” to wait for “many” H-nodes before it connects to one and by that time a long path has been formed which allows the bridge donor to match many nodes at once. On the other hand, with L-nodes, those long chains are not formed; when an L-node arrives it can form a cycle or a short chain relatively quickly preventing long paths with only H-nodes to be formed. However, with L-nodes, we can construct a solution with many short chain segments; the idea is to show that, after enough nodes have arrived, each time an L-node arrives, with a constant probability, the bridge donor will initiate a small chain segment by connecting to

the newly arrived L-node, and continuing to an isolated path containing only H-nodes and of length at least 2; note that the H-nodes in such a path have no other incoming edges, and thus can never be matched when allowing only 2-way cycles.

Part (b) of Theorem 3.5 is the online version of the main result of Ashlagi et al. [2011], who show that in a static large sparse pool (equivalent to our offline solution) chains add significantly to the number of matched pairs.

Note that Theorem 3.5 does not cover the case in which we have both a mixture of H and L type nodes, and we also allow 3-way cycles. We believe a similar comparison holds for this case as well, but we were not able to prove it.

Remark 3.6. *The continuation of the dynamic chains crucially depends on the existence of the bridge donor. Because of ethical reasons, there is no formal assurance that such donor will donate in the future. Even though the historical data suggests that there have only been very few cases that the bridge donor reneged, we can extend our analysis to the model with possibility of renege. In particular, suppose at each time step t , the bridge donor (or the altruistic donor if the chain has not yet started) reneges independently with probability p_r . If $p_r = O(1/n)$, the qualitative results of both parts of Theorem 3.5 hold, implying that if renege probability is small enough we will still gain a significant (linear) number of matches when using even one dynamic chain. For such a renege probability, in expectation, it takes $\Omega(n)$ steps before a BD reneges. Thus, the proofs of parts (a) and (b) can be modified when a “small” renege rate is incorporated in the model.*

4 Long waiting with short cycles

In this section, we study the effect of long waiting on the number of matches, by comparing the CM with $\Theta(n)$ chunk sizes to the online scenario. This is done for the sake of completeness as long waiting periods might not be practical in kidney exchange. We show the positive effect of long waiting on the number of allocations. For brevity, we only present the analysis of this regime for 2-way matching. Similar qualitative results hold for 3-way matching as well.

The following theorem states that waiting linearly with L-nodes results in matching linearly more pairs compared to online scenario, and similarly, waiting linearly with both L and H-nodes, increases the number of matchings linearly compared to the case that we only wait with the L-node.

Theorem 4.1 (Linear waiting with 2-ways). *Let $0 < \beta < 1$ where βn is a divisor of n .*

(a) **Waiting with L-nodes:** *There exists $\delta_\beta > 0$ such that:*

$$\mathbb{E} \left[M^2(1, \beta n) \right] \geq \mathbb{E} \left[M^2(1, 1) \right] + \delta_\beta n.$$

(b) **Waiting with H-nodes:** *There exists $\delta'_\beta > 0$ such that:*

$$\mathbb{E} \left[M^2(\beta n, \beta n) \right] \geq \mathbb{E} \left[M^2(1, \beta n) \right] + \delta'_\beta n.$$

The proof of the above theorem is presented in Appendix C. Finally, in the following theorem, we show that even if we divide the data into a “few” chunks (or equivalently run the matching after βn steps instead of waiting until the end) we will match linearly less nodes as compared to the hindsight solution. The proof of theorem is deferred to Appendix C.

Theorem 4.2 (Upper bound on linear waiting). *Let $0 < \beta < 1$ where βn is a divisor of n . There exists $\tilde{\delta}_\beta > 0$ such that:*

$$\mathbb{E} \left[M^2(\beta n, \beta n) \right] \leq \mathbb{E} \left[M^2(n, n) \right] - \tilde{\delta}_\beta n.$$

5 Modeling extensions

In this section we point out some of the limitations of the model we studied in this paper and discuss how our main results and intuitions extend to generalized models.

5.1 Dynamic pool with blood-type compatibility

In the model presented in Section 2, we abstract away from blood-type compatibility factor and assume that all pairs in the pool are blood-type compatible. We next describe how to generalize our model: the pool consists of several subgraphs, where subgraph $X - Y$ contains all the pairs with blood-type X patient and blood-type Y donor, where $X, Y \in \{A, B, O, AB\}$.¹⁴ If X is not blood-type compatible with Y (such a pair is called under-demanded pairs) then there is no edge in subgraph $X - Y$, and directed edges from subgraph $X - Y$ to $X' - Y'$ exist if type X' is compatible with type Y . Given that donor of pair v_1 and patient of pair v_2 are blood-type compatible, there will be a directed edge from v_1 to v_2 if and only if they are tissue-type compatible. Further, the fraction of highly sensitized patients in each subgraph is different. For example the PRA of patients in the $A - A$ subgraph is expected to be higher because an $A - A$ pair only joins the pool if the patient is not tissue-type compatible with the donor. Thus one needs to model the PRA distribution of each subgraph separately. Furthermore, the size of each subgraph in the residual pool is different

¹⁴A patient with blood-type O is blood-type compatible only with donors with blood-type O . A patient with blood-type A is blood-type compatible with donors with blood-types A and O . A patient with blood-type B is blood-type compatible with donors with blood-type B and O . Finally, a patient with blood-type AB is blood-type compatible with all donors.

because of the heterogenous blood-type distribution of the population and the fact that most of blood-type compatible pairs do not join the exchange pool. In a dynamic exchange model, at any time many under-demanded pairs are waiting to be matched. In fact, even in static models, it has been shown Ashlagi and Roth [2013] that around 2/3 of each under-demanded subgraph will not be matched.

As our simulation results (Figures 3 and 4) show the qualitative behavior of chunk matching follows our theoretical results proven for the simplified model. Here we mention a few differences: 1) At any time, under-demanded subgraphs contain many low sensitized patients that do not have any blood-type compatible option (while in our analysis, we show that low sensitized patients get matched easily and do not stay in the pool.) This makes fast matching even more effective, because it increases the probability that an arriving over-demanded pairs ¹⁵with a highly sensitized patient has some incoming edges from under-demanded pairs, and can be matched quickly. 2) Union of subgraphs such as $A - B$ and $B - A$ forms random bipartite graphs; in such graphs, 3-way matching is not possible, and we are limited to match by only 2-ways or chains. Thus short waiting in 3-way matchings is not effective in such subgraphs.

5.2 Infinite horizon

Our dynamic model of kidney exchange assumes a finite horizon where only n pairs arrive, and we use the hindsight matching of these n pairs as a benchmark to evaluate the efficiency (expected number of allocations) of chunk matching with different scenarios (k-way matching and chains) and chunk sizes. In our model, n is viewed as the relevant horizon or the longest reasonable period of waiting. However, real programs run continuously and do not stop. Thus to add this to our model, we can consider an infinite horizon model in which each pair remains in the pool for at most n periods. Intuitively, in such a model the gap between different chunk sizes decreases, but our qualitative results still hold for different chunk sizes and scenarios; in most of our analysis, we compare the number of H-nodes that each scheme can match and show that one scheme is inferior if it makes a mistake on matching an H-node that does not have any other options in the remaining horizon. In a similar way, we can show that there is a constant probability that such an H-node does not have any matching option in the next n arrivals.

6 Discussion

Previous theory for kidney exchange dealt with dense graphs, finding that efficiency can be obtained via short cycles. When graphs are dense, the efficiency loss from not waiting to accumulate incompatible pairs is minor. In practice, however, pools are very sparse with a large fraction of very highly sensitized patients. In such pools, waiting will increase matching opportunities for patients,

¹⁵Pair $X - Y$ is called over-demanded if X is blood-type compatible with Y and $X \neq Y$.

specially for the highly sensitized ones. Waiting, however, may lead to a significant cost. This raises the question of the tradeoff between the amount of waiting between two consecutive match runs and the number of matches one obtains. We initiate here this direction by studying a class of algorithms that find a maximum allocation every time x new pairs arrive and performing sensitivity analysis on x (or equivalently the waiting time between two consecutive match runs).

We find that when matching dynamically, it is particularly important to use easy-to-match (low-sensitized) pairs in cycles or chains that involve highly sensitized (hard-to-match) patients as well. We show that sometimes by waiting briefly with some low-sensitized pairs we can help matching more highly sensitized ones. Such a design premise is beyond giving priority (in breaking ties) to hard-to-match patients.

It has been shown that even a single unbounded chain, beginning with a non-directed donor, increases efficiency significantly in large static sparse pools beyond just short cycles. We show here the dynamic version of this result for the online setting: we find that in the online scenario with a single non-directed donor the algorithm will match linearly many more pairs than without the non-directed donor (this result assumes that either, in both settings, cycles can be of length at most 2, or the pool only contains pairs with hard-to-match patients). We conjecture that the last result holds also when cycles of length $k > 2$ are allowed.

Our results suggest that if a centralized clearinghouse cannot afford to wait “too long” and limited to use short cycles, then even short waiting with low sensitized pairs can be effective.¹⁶ Dynamic chains can be used to reduce the disadvantage of online matching over matching with waiting. Further, our analysis also suggests that online matching is more effective if pools of different programs merge to form a thicker pool. In such a merged pool, the rate of arrivals increases, and so does the number of matching candidates for each patient in a given time horizon. This reduces the loss of online matching due to making mistakes in matching pairs.

Our work opens directions to many further studies. While waiting times are part of the matching algorithms, we do not study the average waiting time of pairs and only focus on the number of allocations. Observe, however, that there maybe a tight correspondence between the number of matches achieved and the waiting time. Thus, although with linear size chunks one obtains more matches, the average waiting time may increase. It is further interesting to study the steady state of the system. Another direction is whether non-myopic algorithms can improve both the waiting times and the number of pairs matched. As pairs wait to be matched, designing mechanisms that take into account incentives for patients (e.g., Roth et al. [2005b]) and hospitals (e.g., Ashlagi and Roth [2013]) becomes an intriguing task.

Dynamic kidney exchange in an infinite-horizon framework can be viewed as the problem of matching multi-class costumers to multi-class servers studied in queueing literature (for example, Caldentey et al. [2009]): An infinite sequence of costumers and servers arrive over time; each

¹⁶No multi-hospital kidney exchange program in the US is currently waiting more than a month before finding allocations.

costumer can only be severed by a certain types of servers. For the kidney exchange application, the additional constraint is that costumers (patients) and servers (donors) arrive in pairs (except for NDDs), and must be matched at the same time. A few papers study such settings Caldentey et al. [2009] Gurvich and Ward [2012] where the underlying graph structure is deterministic. More recently Anderson et al. [2013] study a similar setting with random homogenous graphs, and show that in a homogenous population, where the matching probabilities are identical, greedy (online) policy is almost optimum. Our analysis, in the finite horizon model, suggests that if we include heterogeneity in matching probabilities (which is closer to the reality of current kidney exchange pools) then the conclusion will be different. In particular, some waiting might be helpful.

Thickness is an important property for efficiency in market design. Kidney exchange clearing-houses can create a thick market at the cost of waiting for many pairs to arrive. Tradeoffs between unraveling (waiting before entering the market) and thickness are of practical importance in many other markets such as job markets and markets for graduate students (see e.g., Neiderle and Roth [2009]). Our theory can serve as a building block for studying such tradeoffs and for the study of implementing “efficient” outcomes in the long run when agents have preferences.

References

- R. Anderson, I. Ashlagi, Y. Kanoria, and D. Gamarnik. A dynamic model of barter exchange. Working paper, 2013.
- I. Ashlagi and A. E. Roth. Free riding and participation in large scale, multi-hospital kidney exchange. *Theoretical Economics*, 2013. Forthcoming.
- I. Ashlagi and A.E. Roth. New Challenges in Multi-hospital Kidney Exchange. *American Economic Review, Papers and Proceedings*, 102(3):354–359, 2012.
- I. Ashlagi, D. Gamarnik, and A. E. Roth. An algorithm for finding cycles and long chains in kidney exchange. Working paper, 2011.
- I. Ashlagi, D. Gamarnik, Michael A. Rees, and A. E. Roth. The need for (long) NEAD chains. Working paper, 2012.
- P. Awasthi and T. Sandholm. Online stochastic optimization in the large: application to kidney exchange. In *Proceedings of the 21st international joint conference on Artificial intelligence, IJ-CAI’09*, pages 405–411, San Francisco, CA, USA, 2009. Morgan Kaufmann Publishers Inc. URL <http://dl.acm.org/citation.cfm?id=1661445.1661510>.
- A. Blum, A. Gupta, A. Procaccia, and A. Sharma. Harnessing the power of two crossmatches. In *Proceedings of the fourteenth ACM conference on Electronic commerce, EC ’13*, pages 123–140,

- New York, NY, USA, 2013. ACM. ISBN 978-1-4503-1962-1. doi: 10.1145/2482540.2482569. URL <http://doi.acm.org/10.1145/2482540.2482569>.
- R. Caldentey, E. H. Kaplan, and G. Weiss. Fcfs infinite bipartite matching of servers and customers. *Adv. Appl. Prob.*, 41(3):695–730, 2009.
- 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*, 2012a.
- J. P. Dickerson, A. D. Procaccia, and T. Sandholm. Dynamic Matching via Weighted Myopia with Application to Kidney Exchange. *Proc of the 6th AAAI Conference on Artificial Intelligence*, pages 1340–1346, 2012b.
- J. P. Dickerson, A. D. Procaccia, and T. Sandholm. Failure-aware kidney exchange. In *ACM Conference on Electronic Commerce*, pages 323–340, 2013.
- 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.
- 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.
- P. Jaillet and X. Lu. Online stochastic matching: New algorithms with better bounds. *Mathematics of Operations Research*, 2013. Forthcoming.
- 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.
- M. Krivelevich, E. Lubetzky, and B. Sudakov. Longest cycles in sparse random digraphs. *Random Structures and Algorithms*, 2012. Forthcoming.
- 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*, 50(6):1505–1524, 1982.
- M. Molinaro and R. Ravi. Kidney Exchanges and the Query-commit Problem. Manuscript, 2013.

- M. Neiderle and A.E. Roth. Market Culture: How Rules Governing Exploding Offers Affect Market Performance. *American Economic Journal: Microeconomics*, 1(2):199–219, 2009.
- M. A. Rees, J. E. Kopke, R. P. Pelletier, D. L. Segev, M. E. Rutter, A. J. Fabrega, J. Rogers, O. G. Pankewycz, J. Hiller, A. E. Roth, T. Sandholm, M. U. Ünver, and R. A. Montgomery. A non-simultaneous extended altruistic donor chain. *New England Journal of Medicine*, 360:1096–1101, 2009.
- A. E. Roth, T. Sönmez, and M. U. Ünver. Kidney exchange. *Quarterly Journal of Economics*, 119:457–488, 2004.
- A. E. Roth, T. Sönmez, and M. U. Ünver. A kidney exchange clearinghouse in New England. *American Economic Review Papers and Proceedings*, 95(2):376–380, 2005a.
- A. E. Roth, T. Sönmez, and M. U. Ünver. Pairwise kidney exchange. *Journal of Economic Theory*, 125:151–188, 2005b.
- A. E. Roth, T. Sönmez, M. U. Ünver, F. L. Delmonico, and S. L. Saidman. Utilizing list exchange and nondirected donation through chain kidney paired donations. *American Journal of Transplantation*, 6:2694–2705, 2006.
- 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.
- S. L. Saidman, A. E. Roth, T. Sönmez, M. U. Ünver, and F. L. Delmonico. Increasing the Opportunity of Live Kidney Donation by Matching for Two and Three Way Exchanges. *Transplantation*, 81:773–782, 2006.
- T. Sönmez and M. U. Ünver. Altruistic Kidney Exchange. Working paper, 2011.
- M. Suthanthiran and T. B. Strom. Renal transplantation. *New England Journal of Medicine*, page 331:365, 1994.
- UNOS. United network for organ sharing. 2013. URL <http://www.unos.org/>.
- M. U. Ünver. Dynamic Kidney Exchange. *Review of Economic Studies*, 77(1):372–414, 2010.
- USRDS. United states renal data system. 2012. URL <http://www.usrds.org/qtr/default.aspx>.
- S. A. Zenios. Optimal control of a paired-kidney exchange program. *Manage. Sci.*, 48(3):328–342, March 2002. ISSN 0025-1909. doi: 10.1287/mnsc.48.3.328.7732. URL <http://dx.doi.org/10.1287/mnsc.48.3.328.7732>.

A Missing proofs of Section 2

Proof of Lemma 2.1. In order to prove part (a) of Lemma 2.1, we study the number of unmatched nodes at time n ; let Z be the set of unmatched nodes at time n . We show that $\mathbb{E}[|Z|] \leq \frac{1}{p_H^4}$. To do so, we use the basic property of any online greedy algorithm: if node i and j belong to Z they could not be matched to each other, thus there is no undirected edge between them (the probability of this event is at most $(1 - p_H^2)$). This fact gives us the following upper bound on the probability that $|Z| = i$, for any $2 \leq i \leq n$:

$$\mathbb{P}(|Z| = i) \leq (1 - p_H^2)^{\binom{i}{2}} \leq (1 - p_H^2)^{i-1} \quad 2 \leq i \leq n.$$

Using this bound, we compute an upper bound for $\mathbb{E}[|Z|]$:

$$\begin{aligned} \mathbb{E}[|Z|] &= \sum_{i=1}^n \mathbb{P}(|Z| \geq i) \leq 1 + \sum_{i=2}^n \sum_{j=i}^n (1 - p_H^2)^{j-1} \\ &\leq 1 + \frac{1 - p_H^2 - p_H^2(n-1)(1 - p_H^2)^n}{p_H^4} \leq \frac{1}{p_H^4}. \end{aligned} \quad (2)$$

Now we have $M = n - |Z|$, and the above inequality implies the statement. To prove part (b), we simply count the expected number of H-nodes that are not part of any 2-ways. Clearly no algorithm can match these nodes. Suppose at time n , we have n_H nodes of type H where $\mathbb{E}[n_H] = \rho n$. Conditioned on this, for each H-node v , we have:

$$\mathbb{P}(\text{node } v \text{ does not form any 2-way}) = (1 - p_H p_L)^{n - n_H} (1 - p_H^2)^{n_H - 1}$$

Thus in expectation, we have:

$$\mathbb{E}[\text{number of H-nodes with no 2-way}] = \frac{(1 - p_H p_L)^n}{1 - p_H^2} \mathbb{E}\left[n_H \left(\frac{1 - p_H^2}{1 - p_H p_L}\right)^{n_H}\right].$$

Note that the function $f(x) = x \left(\frac{1 - p_H^2}{1 - p_H p_L}\right)^x$ is a convex function, thus the Jensen's inequality implies that:

$$\begin{aligned} \mathbb{E}[\text{number of H-nodes with no 2-way}] &\geq \rho n \frac{(1 - p_H p_L)^n}{1 - p_H^2} \left(\frac{1 - p_H^2}{1 - p_H p_L}\right)^{\rho n} \\ &= \rho n (1 - p_H p_L)^{(1 - \rho)n} (1 - p_H^2)^{\rho n - 1}, \end{aligned}$$

which completes the proof. \square

B Missing proofs of Section 3

Proof of part (a) of Theorem 3.1. We focus on the first chunk and suppose we index the nodes by the time they arrive; in the online schemes (i.e., $S_H = S_L = 1$), at any time $1 \leq t \leq S$, if an L-node arrives and it gets matched to an L-node, it may cause a loss in the number of matching, because if the algorithm had waited before matching node t , this L-node might have been used to match an H-node that has arrived after time t . However, we show that the probability of this event is “small”: We can have at most $S - t$ such H-nodes, and each has an edge with node t with probability pc/n . Thus by the union bound:

$$\mathbb{P}(\text{L-node } t \text{ could be used to match an H-node arriving after } t) \leq \frac{(S-t)pc}{n}.$$

Using this upper bound, we compute an upper bound on the expected number of the mistakes that the online scheme can make:

$$\mathbb{E}[\text{\#of mistakes of online scheme}] \leq \sum_{t=1}^S \frac{(S-t)pc}{n} = \Theta\left(\frac{S^2}{n}\right).$$

The same bound holds for any later chunk as well. We have n/S chunks and for each of them the online makes at most $\Theta\left(\frac{S^2}{n}\right)$ mistakes, therefore the total number of mistakes is at most $O(S)$. This implies that chunk with $S_L = S$ and $S_H = 1$ matches at most $O(S) = o(n)$ more than the online scenario ($S_H = S_L = 1$). \square

Proof of part (b) of Theorem 3.1. We first prove the result for $S < n^{1/2}$, and then generalize it to the case where $n^{1/2} \leq S \leq n^{1-\epsilon}$.

We begin by showing that, excluding the L-L edges, the graph induced by the set of nodes in the arriving chunk S contains no edges with high probability. Denote by \mathcal{E} the set of edges induced by the most recent chunk of nodes (the filled nodes in Figure 6).

$$\begin{aligned} \mathbb{E}[|\mathcal{E}|] &= 1/2 \sum_{i,j \in S} \mathbb{P}(i \text{ is connected to } j) \\ &\leq \frac{|S|(|S|-1)pc}{2n} = O\left(\frac{S^2}{n}\right) = o(1). \end{aligned} \tag{3}$$

The first inequality in (3) holds because the probability of a 2-way that is not L-L is at most pc/n . By Markov’s inequality,

$$\mathbb{P}(|\mathcal{E}| > 1) \leq \mathbb{E}[|\mathcal{E}|] = o(1),$$

implying that w.h.p. the set \mathcal{E} is empty.

Note that after a new chunk of S nodes arrive, the graph consists of nodes from the previous residual graph (with no edges between themselves) and the new S nodes. Next we show that after the arrival of the new S nodes, w.h.p. no node from the residual graph (not-filled nodes in Figure 6) has degree larger than one. Denote by \mathcal{C} the set of nodes of the residual graph. By union bound, we have:

$$\begin{aligned} \mathbb{P}(\exists i \in \mathcal{C} \text{ with degree more than 1}) &\leq \sum_{i \in \mathcal{C}} \mathbb{P}(i \text{ has degree more than 1}) \\ &= |\mathcal{C}| [1 - \mathbb{P}(i \text{ has degree zero or one})] \\ &\leq |\mathcal{C}| \left[1 - \left(1 - \frac{pc}{n}\right)^{|\mathcal{C}|} - \frac{|\mathcal{C}|pc}{n} \left(1 - \frac{pc}{n}\right)^{|\mathcal{C}|-1} \right] \end{aligned}$$

The last inequality holds because the RHS gives the probability that node i has degree more than 1 when all the nodes in the arriving chunk are of type L. Clearly, when some of the nodes in the arriving chunk are of type H, the probability of node i having degree more than 1 will be smaller. Using the well-known approximation that for small x , $(1 - x)^y = e^{-xy} (1 + O(x^2y))$, we have:

$$\begin{aligned} \mathbb{P}(\exists i \in \mathcal{C} \text{ with degree more than 1}) &\leq |\mathcal{C}| \left(1 - e^{-pc|\mathcal{C}|/n} - \frac{|\mathcal{C}|pc}{n} e^{-pc|\mathcal{C}|/n} + O\left(\frac{|\mathcal{C}|}{n^2}\right) \right) \\ &= O\left(\frac{|\mathcal{C}||\mathcal{C}|}{n^2}\right) = o(1), \end{aligned} \tag{4}$$

where the last order equality holds because the size of \mathcal{C} is at most $\Theta(n)$. We have n/S chunks, and we showed that in each chunk, the gain of CM^2 over the online scenario is $o(S)$. Thus the total gain of CM^2 compared to online matching is $o(n)$.

Next we extend this analysis to the regime $n^{1/2} \leq S \leq n^{1-\epsilon}$. The basic intuition is the same as for $S < n^{1/2}$; the subgraph of the arrived nodes (excluding the L-L edges) is very sparse and w.h.p. there exists no nodes in the residual graph that has degree larger than one. The proof of the latter is the same as it is done for $S < n^{1/2}$; just consider (4), the $\mathbb{P}(\exists i \in \mathcal{C} \text{ with degree more than 1})$ is still $o(1)$ for $n^{1/2} \leq S \leq n^{1-\epsilon}$. However, the proof of the former is different due to the fact that when we increase S above $n^{1/2}$, the subgraph of the arrived nodes will have a few edges; in fact, Equation (3) says that it has $O\left(\frac{S^2}{n}\right) = o(S)$ edges. Suppose we ignore these edges, then similar to the case $S < n^{1/2}$, we show that in each arriving chunk, CM^2 matches at most $o(S)$ more nodes than the online does. Now since adding K edges to a graph increases the size of its maximum matching by at most K , it follows that when we add these $o(S)$ edges to the whole graph (both filled nodes and not-filled nodes), the size of its maximum matching increases by at most an $o(S)$ factor. Thus the gain of CM^2 over the online scenario in each chunk is $o(S)$, and we have n/S chunks, which implies that the overall gain of CM^2 with $S_L = S \leq n^{1-\epsilon}$ is $o(n)$. \square

Proof of part (a) of Theorem 3.3. Let A^{CM} be the set of H-nodes that CM^3 (with $S_L = S$) matches, but online can never match in the entire horizon. Further, let A^O be the set of H-nodes that online may match, but CM^3 may not be able to match in the entire horizon. First observe that $\mathbb{E}[M^3(1, S)] \geq \mathbb{E}[M^3(1, 1)] + \mathbb{E}[|A^{CM}|] - \mathbb{E}[|A^O|]$. We will show that when condition (1) holds then $\mathbb{E}[|A^{CM}|] - \mathbb{E}[|A^O|] \geq \delta n$ which implies the result, because as mentioned before both schemes match almost all L-nodes. To do so, we find a lower bound on $\mathbb{E}[|A^{CM}|]$ and an upper bound on $\mathbb{E}[|A^O|]$.

Consider the entire graph, and the set of H-nodes u with the following properties:

1. Node u has only one incoming edge that is from an L-node v .
2. Node u has no outgoing edge to any H-node.
3. Node v arrives after u .
4. Node v has no outgoing edge to any other H-node.
5. Node u does not form an edge to node v .

First let us evaluate the probability for having such H-nodes u : the above five events are asymptotically independent, and (for large enough n) respectively have probabilities: $c(1 - \rho)e^{-c}$, $e^{-c\rho}$, $1/2$, $e^{-c\rho}$, and $(1 - p)$.¹⁷ Thus conditioned on the event that a node is H, the probability that these five properties hold is $(1/2)(1 - p)(1 - \rho)ce^{-c(1+2\rho)}$.

Now we claim that any H-node that has the above properties will, w.h.p., be matched by CM^3 but not by the online scheme. Thus these H-nodes will belong to the set A^{CM} : note that node u can only be matched in a 2-way or a 3-way that includes v . Because of the last property, u and v cannot form a 2-way, and the only way to match u is by using 3-way cycles. Because node u has no outgoing edge to another H-node, it is not possible to form a H-H-L cycle with u ; thus the only possible cycle is an H-L-L cycle. We show that CM^3 can easily form this cycle, but the online scheme cannot. Consider the chunk in which node v arrives and forms the edge (v, u) . There exist $(1 - \rho)S$ other L-nodes in that chunk; each of these nodes can form a 3-way cycle with u and v with the constant probability p^2 ; thus, w.h.p., CM^3 can find such a 3-way at the end of the chunk. Also, note that in this chunk, node v can be part of other cycles as well, but since it has no other outgoing H neighbor, those cycles can be either L-L or L-L-L cycles. Since we give priority to matching more H-nodes, the H-L-L cycle including v has the priority and CM^3 will choose it. On the other hand, w.h.p., the online scheme will match v in some other cycle right after it arrives, and node u will remain unmatched.

Now we compute an upper bound on $\mathbb{E}[|A^O|]$ using the H-nodes u that have the following properties:

¹⁷Note that for n large enough, the indegree of H-nodes has a Poisson distribution with rate c . Also, the outdegree of L-/H-nodes to H-nodes has Poisson distribution with rate $c\rho$.

1. Node u has indegree at least 2, and at least one of the incoming neighbors is an L-node v .
2. Node u has at least one outgoing edge to another H-node u' .
3. Node u' has an edge to node v .

Again, let us compute the probability that these three events happen: conditioned on u being H, the probability of the intersection of the above events is $(1/2)p(1 - e^{-c\rho})\left(1 - c(1 - \rho)e^{-c} - e^{-c(1-\rho)}\right)$.

We show that any H-node that has the above properties may be used by the online scheme to match *another* H-node that CM^3 may not be able to match: suppose node u is matched by CM^3 at time t , but not by the online. At some later time t' in a later chunk, node u will be in the residual graph of the online scheme (but not in the residual graph of CM^3). Suppose that node u has an outgoing edge to another H-node u' and at time t' , node u' has not been matched by neither online nor CM^3 . Assume that at time t' an L-node arrives and it forms an outgoing edge to node u , and it has an incoming edge from u' ; now the online scenario can form the H-H-L 3-way cycle $u - u' - v$, but because CM^3 has already matched u in a previous chunk it cannot form this cycle. In the worst case, CM^3 will never be able to match u' in any future chunk, and node u' will belong to the set A^O . The above three properties are minimally required for having such $u - u' - v$ cycles, and thus this gives us an upper bound on $\mathbb{E}[|A^O|]$.

Given these two bounds, $\mathbb{E}[|A^{CM^3}|] \geq (n/2)(1 - p)(1 - \rho)ce^{-c(1+2\rho)}$ and $\mathbb{E}[|A^O|] \leq (n/2)p(1 - e^{-c\rho})\left(1 - c(1 - \rho)e^{-c} - e^{-c(1-\rho)}\right)$, condition (1) implies that CM^3 with $S_L = S$ matches at least δn more nodes than CM^3 with $S_L = 1$ (i.e., the online scheme) does. \square

Proof of part (b) of Theorem 3.3. Similar to the proof of part (b) of Theorem 3.1, we show that after a new chunk arrives, the graph (the union of the residual graph and the new chunk excluding the L-L edges) is so disconnected that the decisions that the CM^3 with $S_H = S$ makes are mostly the same as those of CM^3 with $S_H = 1$ (See Figure 8). Further, in the second phase when searching for an allocation in the entire graph (i.e., including the L-L edges), both of these schemes will “see” almost the same residual graph, and thus find nearly the same number of exchanges.

Consider the k -th chunk where $k = \Theta(n)$; the chunk consists of $\rho S + o(S)$ nodes of type H and $(1 - \rho)S + o(S)$ nodes of type L. The expected number of incoming edges to this new set of H-nodes is $\Theta(S^2/n) = o(S)$, so w.l.o.g, we can ignore these edges. However, there are $\Theta(n)$ nodes of type H in the residual graph, and thus we have $\Theta(S)$ directed edges from the new chunk to the H-nodes in the residual graph. With high probability, none of these outgoing edges will have the same endpoint in the residual graph. More precisely, let C^3 denote the set of H-nodes in the residual graph. Similar to the calculation of (4), one can show that the probability that there exists a node $i \in C^3$ with more than one incoming edge from the new chunk is $o(1)$.

Given the above observations, let us study the structure of the allocations after the new chunk arrives. We argue that the matching made by the CM^3 (in the first phase when we ignore the L-L edges) mainly consists of new L-nodes and the nodes of set C^3 (all the circle nodes in Figure 8);

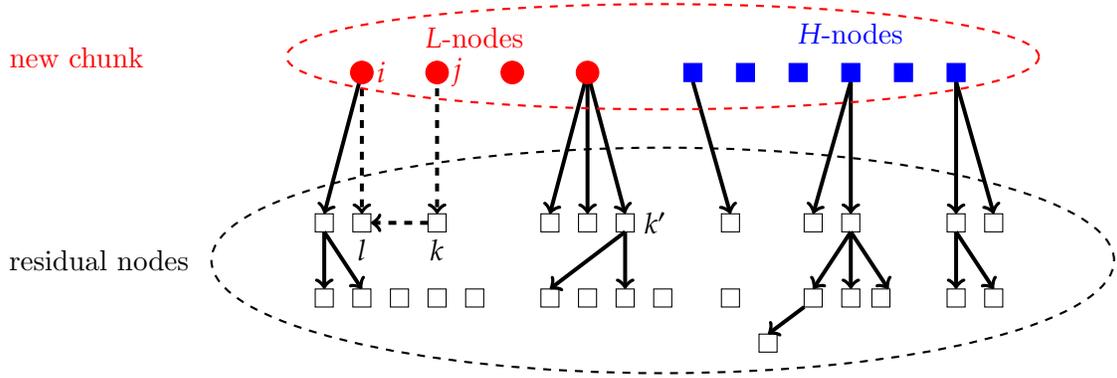


Figure 8: The typical connected components when the chunk is of a sublinear size; the filled nodes are the ones arrived in the last chunk; the circle nodes are L-type and the square ones are H-types. The not-filled nodes arrived in the previous chunks, but have not been matched yet. The incoming edges to L-nodes are not shown.

We showed that the number edges to new H-nodes (and thus 2 and 3-way cycles involving them) is $o(S)$. Further, the number of 2 and 3-way cycles involving only H-nodes is also $o(S)$ (in the entire graph there are $O(1)$ such cycles). However, there are $\Theta(S)$ edges from the new L-nodes to the residual H ones (i.e., the set \mathcal{C}^3). For each node in set \mathcal{C}^3 , we know that it is unlikely that it has more than one incoming edge from the new chunk. Fix a node $v \in \mathcal{C}^3$. We distinguish between two cases.

Suppose first that there does not exist a node $w \in \mathcal{C}^3$ for which v has an outgoing edge to such that w also has an edge from a new L-node (for example k' in Figure 8), then if CM^3 chooses a cycle that contains v it will also choose it in the online scenario.

Second, suppose v has an outgoing neighbor $d \in \mathcal{C}^3$ such that node d also receives an edge from a new L-node (for example, see nodes k, l, i, j , and the dashed edges in Figure 8). This additional information may result in matching more nodes by the CM^3 . For instance, in Figure 8, if node i arrives before node j and edges (l, i) and (l, j) exist but edge (k, i) does not (which happens with a constant probability), then the online will match l to i and node k will remain unmatched, but the CM^3 will choose the 3-way of $j-k-l$. Even though such mistakes are possible, we show that having these kind of structures is unlikely; For a connected pair of nodes $k, l \in \mathcal{C}^3$, the probability that both k and l receive edges from the new L-nodes is $\binom{S}{2} (c/n)^2$. We only have $\Theta(n)$ such connected pairs in the residual graph, thus the expected number of such patterns is $\Theta(S^2/n) = o(S)$. Thus the number of mistakes that the online scheme can make due to making early 2-way cycles is $o(S)$. This shows that in a single chunk, CM^3 can only match at most $o(S)$ more nodes than the online does. Therefore, in the entire horizon, the gap between number of matches the two schemes achieve is of order $o(n)$.

□

Proof of part (a) of Theorem 3.5. Observe that the offline graph contains only $O(1)$ cycles of length 2 or 3 since the expected number of cycles of length constant k is $\binom{n}{k} (c/n)^k = O(1)$. Thus the size of matching obtained by the online scenario without a chain (i.e., $CM^3(1,1)$) will be $O(1)$. On the other hand, we show that, in expectation, $CM^{3,c}(1,1)$ can match $\Theta(n)$ nodes thus proving the claim: Consider the arriving process and suppose we index the nodes by the time they arrive. Let i be the first arriving node that is connected to the altruistic donor. Clearly no chain has been formed before time i . Also suppose that no 2- or 3-way cycles have been performed either (this happens with a constant probability). Notice that $\mathbb{E}[i] = n/c$ and for any ϵ , we have $\mathbb{P}(i > \epsilon n/c) = (1 - c/n)^{\epsilon n/c} = e^{-\epsilon} + o(1)$. Conditioning on the two events that $\{i > \epsilon n/c\}$ and no nodes were matched before time i , we compute a lowerbound on the expected size of the matching obtained by the online matching with a single chain; The residual graph at time i is simply a directed Erdős-Rényi graph with at least $\epsilon n/c$ nodes and edge probability c/n . With high probability, such a graph has a path of length $\Theta(n)$ Krivelevich et al. [2012]. Let $p = (p_1, p_2, \dots, p_L)$ be such a path. The probability that i has an outgoing edge to at least one of the p_j 's for $1 \leq j \leq L/2$ is $1 - (1 - c/n)^{L/2}$. Now since $L = \Theta(n)$, this probability is bounded away from zero, implying that with a constant probability the chain formed at time i matches at least $L/2 = \Theta(n)$ nodes. Finally, note that the expected number of total allocations is at least the number of nodes matched at time i ; this completes the proof. \square

Proof of part (b) of Theorem 3.5. As usual one can show (similar to the proof of Lemma 2.1) that at any time t we have only $O(1)$ nodes of type L in the residual graph. Thus we know that both schemes will match almost all L-nodes, and hence it suffices to compare the number of H-nodes matched by these two schemes. In particular, let A_t^c (A_t) be the set of allocated H-nodes by $CM^{2,c}(1,1)$ ($CM^2(1,1)$) by time t (i.e., up to time t). We aim to show that

$$\mathbb{E}[|A_n^c \setminus A_n|] \geq \mathbb{E}[|A_n \setminus A_n^c|] + \Theta(n). \quad (5)$$

To do so, we study the evolution of the two sets A_t^c and A_t , and in particular their differences. The proof will follow from the next two claims that show that $\mathbb{E}[|A_n \setminus A_n^c|] = o(n)$ and that $\mathbb{E}[|A_n^c \setminus A_n|] = \Theta(n)$, implying together inequality (5).

Claim B.1. For every t , $\mathbb{E}[|A_n \setminus A_n^c|] = o(n)$.

Proof. Let $I_t = \mathbb{E}[|A_t \setminus A_t^c| - |A_{t-1} \setminus A_{t-1}^c|]$ be the expected increment in the number of nodes that $CM^2(1,1)$ matches at time t but $CM^{2,c}(1,1)$ does not. To prove the claim we show that $I_t = o(1)$ for any t . Consider the node arriving at time t and distinguish between the following cases.

- (a) Node t is an H-node and it is connected to the bridge donor. Observe that the contribution of this case to the I_t is at most $2\rho c/n$: the probability of case (a) is $\rho c/n$, and at any time t , the

maximum number of H-nodes that $CM^2(1,1)$ can match (and thus can add to $A_t \setminus A_t^c$ at the worst case) is 2.

- (b) Node t is H type and it is not connected to the bridge donor. In this case, it is possible that $CM^2(1,1)$ matches node t to a node u , but $CM^{2^c}(1,1)$ cannot do so, because node u was matched before. Thus node t will be added to $A_t \setminus A_t^c$. However, we argue that the probability of this event is $o(1)$: If node u is H type, the probability of a having a H-H two-way is $(c/n)^2$ and since we only have $O(n)$ nodes of H type in our pool, the probability that such a two-way exists is at most $O(1/n)$. Next suppose that node u is of type L. In this case the probability of a having a H-L two-way is pc/n , but in expectation, we only have $O(1)$ such L-nodes in the residual graph, thus the probability of this event is of order $o(1)$ as well. Therefore, the contribution of this case to I_t is $o(1)$.

- (c) Node t is an L-node and is connected to the bridge donor.

Note that scheme $CM^2(1,1)$ can match at most one H-node, say node u . If node u was matched already by $CM^{2^c}(1,1)$, then I_t would be zero. On the other hand, if node u was not matched by $CM^{2^c}(1,1)$ before, then $CM^{2^c}(1,1)$ can also perform the 2-way cycle (t, u) , and since we give the priority to the 2-way with one H-node, $CM^{2^c}(1,1)$ will perform such a cycle, and again I_t will be zero.

- (d) Node t is L type and it is not connected to the bridge donor.

The analysis of this case is very similar to the previous case and again we can show that $I_t = 0$.

□

Claim B.2. $\mathbb{E}[|A_n^c \setminus A_n|] = \Theta(n)$.

Proof. It is enough to show that the expected increment in each step $t = \Theta(n)$ is $\Theta(1)$. Fix a step t . With probability $(1 - \rho)p$ the arriving node is an L-node which is also connected to the bridge donor. We show that with a constant probability, $CM^{2^c}(1,1)$ can add a path of length at least 4 to the chain that contains at least 2 nodes of type H that $CM^2(1,1)$ can never match through any two-ways, thus these two nodes will surely belong to $A_n^c \setminus A_n$. Consider the entire pool (i.e., the graph that we obtain if we wait until time n and make no allocations). A constant fraction of the H type nodes in this pool have indegree one: more precisely, the probability that a node of type H has indegree one is $\frac{c(n-1)}{n}(1 - c/n)^{n-2}$. Suppose node u is such a node with the only incoming edge (v, u) . With a constant probability node v is also of type H and has only one incoming edge (w, v) where w is of type L. At time t large enough ($t = \Theta(n)$), we will have a linear number of such isolated (v, u) edges in both of our residual pools (i.e., the residual graph of $CM^{2^c}(1,1)$ and the one of $CM^2(1,1)$). Suppose at time t , case (c) happens, and there is an edge from t to one of these isolated directed edges, say edge (v, u) , (i.e., edge (t, v) exists) and node t has no other outgoing

edges to any of the H-nodes in the pool. This happens with a constant probability bounded away from zero. Also suppose there is no edge from v to t . Note that $CM^2(1, 1)$ can never match neither u nor v . However, $CM^{2^c}(1, 1)$ can add the path (BD, t, v, u, \dots) to the chain. Thus we prove that each time a new L-node arrives, and it is connected to the BD, with a constant probability we add two nodes to $A_t^c \setminus A_t$ that can never be removed from $A_j^c \setminus A_j$ for $j \geq t$. \square

\square

C Missing proofs of Section 4

Proof of part (a) of Theorem 4.1. First note that Lemma 2.1 implies that in online matching, almost all the L-nodes will be matched at the end of each chunk. Using similar calculation shows that when $S_L = \beta n$, after including the L-L edges, and matching using these edges, the number of L-nodes remaining in the system is $O(1)$. Thus we focus on comparing the number of H-nodes that these two schemes can match, and show that CM^2 with $S_L = \beta n$ will match $\Theta(n)$ more H-nodes.

Let us focus on the first chunk and suppose we index the nodes by the time they arrive; suppose that at time $t < \beta n/2$ and time $t + 1$ two successive L-nodes have arrived and they are connected to each other. Suppose node t is not connected to any H-node that has arrived before (this probability is at least $(1 - pc/n)^{t-1}$). Similarly suppose node $t + 1$ is not connected to any H-node that has arrived before (again, this probability is at least $(1 - pc/n)^{t-1}$). Now in the online matching, we would match node t to an L-node either at time t or time $t + 1$.

On the other hand, we show that, in expectation, we have $\Theta(1)$ nodes of type H that will arrive after t in this chunk and are only connected to node t in the entire graph. Clearly if we wait until time βn , we could have matched such an H-node to node t . Thus by matching the L-node t along the way and not waiting until time βn , we will decrease the size of matching by $\Theta(1)$ factor. Summing over all $1 < t < \beta n/2$, this implies that not waiting with L-nodes (online scenario) decreases the size of matching by $\Theta(n)$.

We can now start the detailed proof by introducing some notations. Denote the event that nodes t and $t + 1$ are L-nodes, connected to each other, and not connected to any available H-nodes by \mathcal{E}_t . Further, let \mathcal{H}_t be the set of H-nodes that arrive after t and are only connected to node t in the entire graph. First let us compute the probability of event \mathcal{E}_t :

$$\mathbb{P}(\mathcal{E}_t) \geq p^2(1 - \rho)^2 \left(1 - \frac{pc}{n}\right)^{2\beta n},$$

Suppose in the entire graph, n_L nodes of type L arrive. Conditioned on event \mathcal{E}_t , let us count the

set \mathcal{H}_t .

$$\begin{aligned} \mathbb{E}_{n_L} \left[\mathbb{E} \left[|\mathcal{H}_t| \mid \mathcal{E}_t \right] \right] &= \sum_{i=t+2}^{\beta n} \mathbb{E}_{n_L} \left[\mathbb{P} \left(i \text{ is H-node and only connected to } t \mid \mathcal{E}_t \right) \right] \\ &\geq \rho (\beta n - t - 2) \frac{pc}{n} \mathbb{E}_{n_L} \left[\left(1 - \frac{pc}{n} \right)^{n_L-1} \left(1 - \frac{c^2}{n^2} \right)^{n-n_L-1} \right] \\ &\geq \rho (\beta n - t - 2) \frac{pc}{n} \left(1 - \frac{pc}{n} \right)^{(1-\rho)n} (1 - o(1)), \end{aligned}$$

where the last inequality follows from Jensen's inequality. Putting these two together and summing over all $t < \beta n/2$ gives us:

$$\begin{aligned} \sum_{t=1}^{\beta n/2} \mathbb{E} [|\mathcal{H}_t| \mathbb{I}(\mathcal{E}_t)] &= \sum_{t=1}^{\beta n/2} \mathbb{E} [|\mathcal{H}_t| \mid \mathcal{E}_t] \mathbb{P}(\mathcal{E}_t) \\ &\geq p^2 (1-\rho)^2 \left(1 - \frac{pc}{n} \right)^{(2\beta+1-\rho)n} (1 - o(1)) \frac{pc (\beta n/2) (\beta n/2 - 1)}{2n} = \kappa n + o(n). \end{aligned}$$

where $\mathbb{I}(\cdot)$ is the indicator function and $\kappa > 0$ is just a constant used for the ease of presentation.

So far, we have shown that at time βn (end of the first chunk), CM^2 with $S_L = \beta n$ matches $\Theta(n)$ more nodes as compared to the online scheme. Note that any H-node that was counted above cannot be matched in any later chunk. We can find similar patterns in later chunks (i.e., L-nodes that could be matched to H-nodes that arrive later in the same chunk and only have one edge in the entire graph, but will be matched to L-nodes by the online scenario) as well and show that in any chunk, CM^2 with $S_L = \beta n$ will match $\Theta(n)$ more nodes, and this completes the proof. \square

Proof of part (b) of Theorem 4.1. Similar to the proof of part (a), we focus on comparing the number of H-nodes that these two schemes can match, and show that CM^2 with $S_L = S_H = \beta n$ will match $\Theta(n)$ more H-nodes compared to CM^2 with $S_H = 1$ and $S_L = \beta n$. Also, note that the total number of H-H 2-ways that exist in the entire graph is $\Theta(1)$, thus most of the H-nodes are matched by H-L 2-ways. Therefore, it suffices to only compare the number of H-L 2-ways that these two schemes can form.

We first consider the first chunk of nodes, and show that, after βn nodes arrive, CM^2 with $S_H = \beta n$ matches linearly more H-nodes than CM^2 with $S_H = 1$.

Let \mathcal{M} be a matching obtained by CM^2 with $S_H = 1$ (and $S_L = \beta n$) at time βn and before including the L-L edges. Matching \mathcal{M} is not a maximum matching if it has an *augmenting path*, that is an odd length path v_1, v_2, \dots, v_{2l} , where the even edges (v_{2i}, v_{2i+1}) for all $i = 1, \dots, l-1$ are in \mathcal{M} but the odd ones are not.

We show that the residual graph of the CM^2 with $S_H = 1$ at time βn contains linearly many disjoint augmenting paths with respect to \mathcal{M} .

Index the nodes by their arrival time, denote by Π the set of augmenting paths in the graph at time βn that have the following structure: there are four nodes, i, i', j', j such that i and j are of type L and i' and j' are of type H, and (a) i is matched to i' in \mathcal{M} (or $\mathcal{M}(i) = i'$), (b) $j', \mathcal{M}(i) < i$, i.e., j' and $\mathcal{M}(i)$ arrived before i , (c) $j > i$, i.e., j arrived after i and (d) j, j' are not matched in \mathcal{M} (See Figure 9).

Note that such augmenting paths exist because when the node i arrives, the CM^2 with $S_H = 1$ needs to decide whether to match it to $\mathcal{M}(i)$ or j' (or maybe other existing nodes) and cannot predict that $\mathcal{M}(i)$ is the wrong choice in this case. The following two claims prove that $\Theta(n)$ disjoint edges can be added to \mathcal{M} at the end of the first chunk. Thus CM^2 with $S_H = \beta n$ matches $\Theta(n)$ compared to CM^2 with $S_H = 1$.

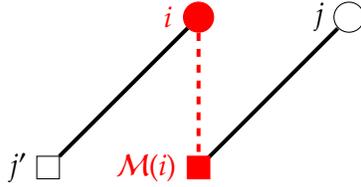


Figure 9: An augmenting path in the set Π . The circle nodes are of type L and the rectangle ones are of type H. The dashed edge belongs to matching \mathcal{M} and the solid edges do not. Nodes j' and $\mathcal{M}(i)$ arrived prior to i , and j arrived after i .

Claim C.1. $\mathbb{E}[|\Pi|] = \Theta(n)$.

Proof. First note that

$$\mathbb{E}[|\Pi| | \mathcal{M}] = \sum_{\pi_4 \text{ is a 4-tuple}} \mathbb{P}(\{\pi_4 \in \Pi\} | \mathcal{M}),$$

where π_4 is a 4-tuple, i.e., any 4 distinct nodes in the chunk. If π_4 does not include any of the pairs $(i, \mathcal{M}(i)) \in \mathcal{M}$ then the probability that it belongs to set Π will be zero. Thus by restricting summation over the 4-tuples that include at least one pair $(i, \mathcal{M}(i)) \in \mathcal{M}$, we have:

$$\mathbb{E}[|\Pi| | \mathcal{M}] = \sum_i \sum_{\pi_4: (i, \mathcal{M}(i)) \in \pi_4} \mathbb{P}(\{\pi_4 \in \Pi\} | \mathcal{M}). \quad (6)$$

We first bound from below $\mathbb{P}(\{\pi_4 \in \Pi\} | \mathcal{M})$: note that a 4-tuple $(i, \mathcal{M}(i), j, j')$ belongs to Π if and only if $j > i$, $j' < i$, edges $(\mathcal{M}(i), j)$ and (i, j') exist, and $j, j' \in Z$, where Z is the set of nodes that were not matched in \mathcal{M} . Analyzing the probability of the event $\{j, j' \in Z\}$ is difficult, so instead we compute the probability of another event that is a subset of $\{j, j' \in Z\}$ (which will provide a lower bound on the probability of the event $\{j, j' \in Z\}$); if in the entire graph, the node j is only connected

to one H-node that is $\mathcal{M}(i)$, then $j \in Z$ (note that node j may later get matched to a node of type L). Similarly, if the H-node j' has degree one in the entire graph, and its only neighbor is node i , then $j' \in Z$. Suppose in the entire graph, n_L nodes of type L arrive. Then given n_L , for $j > i$ and $j' < i$:

$$\begin{aligned} & \mathbb{P}(\{(i, \mathcal{M}(i), j, j') \in \Pi\} | \mathcal{M}) \\ & \geq \left(\frac{pc}{n}\right)^2 \mathbb{P}(\text{nodes } j \text{ has one neighbor of type H and } j' \text{ has degree one} \mid \mathcal{M} \cap \{(\mathcal{M}(i), j), (i, j') \text{ exist}\}) \\ & \geq \left(\frac{pc}{n}\right)^2 \left(1 - \frac{pc}{n}\right)^{n-n_L-1} \left(1 - \frac{pc}{n}\right)^{n_L-1} \left(1 - \left(\frac{c}{n}\right)^2\right)^{n-n_L-1} = \left(\frac{pc}{n}\right)^2 e^{-pc} (1 - o(1)). \end{aligned}$$

Plugging in (6), we obtain that

$$\begin{aligned} \mathbb{E} [|\Pi| | \mathcal{M}] &= \sum_i \sum_{j>i, j'<i, j, j' \in Z} \mathbb{P}(\{(i, \mathcal{M}(i), j, j') \in \Pi\} | \mathcal{M}) \\ &\geq \left(\frac{pc}{n}\right)^2 e^{-pc} \sum_{i=1}^{|\mathcal{M}|} (i-1)(\beta n - i) - O(1) \\ &= \left(\frac{pc}{n}\right)^2 e^{-pc} \beta n \left(\frac{|\mathcal{M}|^2}{2} + \frac{|\mathcal{M}|}{2}\right) - \left(\frac{|\mathcal{M}|^3}{3} + \frac{|\mathcal{M}|^2}{2} + \frac{|\mathcal{M}|}{6}\right) - O(1). \end{aligned}$$

The function $f(|\mathcal{M}|) := \beta n \left(\frac{|\mathcal{M}|^2}{2} + \frac{|\mathcal{M}|}{2}\right) - \left(\frac{|\mathcal{M}|^3}{3} + \frac{|\mathcal{M}|^2}{2} + \frac{|\mathcal{M}|}{6}\right)$ is convex: since $|\mathcal{M}| \leq \beta n/2$, the second derivatives of $f(\cdot)$ which is $f''(|\mathcal{M}|) := \beta n - 2|\mathcal{M}|$ is always non-negative. Thus by Jensens's inequality:

$$\mathbb{E} [|\Pi|] \geq \left(\frac{pc}{n}\right)^2 e^{-pc} f(\mathbb{E} [|\mathcal{M}|]) + O(1) = \kappa n - O(1),$$

where κ is a positive constant that is used for the ease of presentation. The last order equality holds because $\mathbb{E} [|\mathcal{M}|] = \Theta(n)$: matching \mathcal{M} is a maximal matching, thus its size is at least 1/2 of the size of the maximum matching which is, in turn, lower bounded by the number of isolated edges. A simple counting shows that the expected number of isolated edges is $\Theta(n)$. \square

Next we show that there exists $\delta_1 > 0$ such that at least $\delta_1 n$ of these paths are disjoint, and can be added to \mathcal{M} to construct a larger matching.

Claim C.2. *In expectation, the set Π consists of at least $\delta_1 n$ vertex-disjoint paths.*

Proof. Consider a path $p = (i, \mathcal{M}(i), j, j')$ that was counted in the proof of the previous claim, i.e., node j has only one neighbor of type H in the entire graph, and node j' has degree one in the entire graph (from previous claim, we know that there are $\Theta(n)$ of such paths). Clearly these nodes cannot be part of any other paths in Π . However, nodes i and $\mathcal{M}(i)$ can be in more than one of

such paths. We show that in expectation, each pair $(i, \mathcal{M}(i))$ is only included in a constant number of such paths: node $\mathcal{M}(i)$ is of type H, and thus its expected degree in the entire graph is $\Theta(1)$, thus it cannot be included in more than $\Theta(1)$ paths. Therefore, in expectation, $\Theta(n)$ of these paths include distinct $(i, \mathcal{M}(i))$ pairs, implying the claim. \square

This completes the proof for the first chunk of βn nodes. The following claim asserts a similar result for the later chunks.

Claim C.3. *For any chunk $1 \leq l \leq n/S$, there exists $\delta_l > 0$ such that, after chunk l arrives, in expectation, CM^2 with $S_H = S_L = \beta n$ matches $\delta_l n$ more H-nodes compared to CM^2 with $S_H = 1$ and $S_L = \beta n$.*

Proof. First note that we have already proved the claim for $l = 1$. Next we prove for $l = 2$; the main difference between the first two chunks is that at time βn there are some nodes left unmatched in both of these schemes. Note that after matching through the L-L edges, almost all of the nodes in the residual graph are of type H. However, the size of these two residual graphs are not the same. Suppose that before the new chunk arrives, CM^2 with $S_H = 1$ has left the set Z of nodes unmatched. Similarly let Z' be the set of unmatched nodes of the CM^2 with $S_H = \beta n$. We obtained that $|Z| \geq |Z'| + \delta_1 n$ for some $\delta_1 n$ (i.e., CM^2 with $S_H = \beta n$ matches $\delta_1 n$ more H-nodes in the first chunk).

Note that both residual graphs include no edges. Let us partition the set Z into two subsets Z_1 with size $|Z'|$ and Z_2 . Since all the nodes have degree zero and the future edge formations are independent and identical, the partition is arbitrary. We can look at the nodes in Z_2 as the nodes that were matched by CM^2 with $S_H = \beta n$ and not by CM^2 with $S_H = 1$. Thus in the second chunk, if CM^2 with $S_H = 1$ matches a node from the set Z_2 , it only reduces its previous gap with the CM^2 with $S_H = \beta n$. However, at time $2\beta n$ we can compare CM^2 with $S_H = 1$ and CM^2 with $S_H = \beta n$ on the union of the sets Z_1 and the new arrived chunk in a similar way that we compared them at the end of the first chunk, and get the same order result. Repeating similar arguments for the other chunks, and showing that in each chunk l , where $1 < l \leq n/S$, there exists δ_l such that the CM^2 with $S_H = 1$ matches at least $2\delta_l n$ less nodes than CM^2 with $S = \beta n$ does, completes the proof. \square

Let $\delta'_\beta = \sum_{l=1}^{n/S} \delta_l$; the above claim implies that at the end, the expected number of allocations of CM^2 with $S_H = \beta n$ is $\delta'_\beta n$ more than that of the CM^2 with $S_H = 1$, and this completes the proof. \square

Proof of Theorem 4.2. We first prove the theorem for $S = \frac{n}{2}$, i.e., we show that by matching twice, once after $\frac{n}{2}$ nodes arrive and once after the last node arrives, CM^2 with $S_H = S_L = n/2$ results in linearly many less matches compared CM^2 with $S_H = S_L = n$. Similar to the previous proofs, we only compare the number of H-L 2-ways matched by these two algorithms. To do this,

we show there are linearly many disjoint augmenting paths consisting of H-L edges for the union of the two matchings found by CM^2 with $S_H = S_L = n/2$.

Denote by Z_1 the set of nodes arriving up to time $n/2$ and are matched (by CM^2 with $S_H = S_L = n/2$) through an H-L 2-way at time $n/2$, and let Z_2 be the set of nodes that arrive after time $n/2$ and are not matched in the second matching (by CM^2 with $S_H = S_L = n/2$) before including the L-L edges. For any L-node $v_1 \in Z_1$, H-node $v_2 \in Z_1$, L-node $w_1 \in Z_2$, and H-node $w_2 \in Z_2$ such that v_1 is matched to v_2 and the H-L edges (v_1, w_2) and (v_2, w_1) exist, the path $p = (w_2, v_1)(v_1, v_2)(v_2, w_1)$ is an augmenting path. We call such augmenting paths *simple*. Denote by P the set of simple augmenting paths. Note that after including the L-L edges, with high probability, CM^2 with $S_H = S_L = n/2$ will match the L-node w_1 , but it cannot match the H-node w_2 .

The following two claims state that the expected cardinality of P is $\Theta(n)$, and that linearly many paths in P are disjoint.

Claim C.4. $\mathbb{E}[|P|] = \Theta(n)$.

Proof. In order to prove this claim, we show that in the entire graph, excluding the L-L edges, there exist $\Theta(n)$ paths of length 3 which are isolated and have the form H-L-H-L. More formally, let $p = (u_1, u_2, u_3, u_4)$, where u_1 and u_3 are of type H and u_2 and u_4 are of type L. Nodes u_1 and u_4 have degree one in the entire graph, and u_2 and u_3 have degree two. Further, let Q denote the set of such paths.

Note that for any $p \in Q$, with a constant probability, nodes u_2 and u_3 arrive before time $n/2$ and nodes u_1 and u_4 arrive after time $n/2$. If this happens, then edge (u_2, u_3) will belong to the first matching, but nodes u_1 and u_4 cannot be matched by the second matching, and thus such a path belongs to set P (again note that u_4 will get matched once including the L-L edges but u_1 will remain unmatched). Therefore, if we show that $\mathbb{E}[|Q|] = \Theta(n)$, it implies that $\mathbb{E}[|P|] = \Theta(n)$.

We compute the $\mathbb{E}[|Q|]$ as follows: suppose in the entire graph n_L (n_H) nodes of type L (H) arrive, where $\mathbb{E}[n_L] = (1 - \rho)n$ ($\mathbb{E}[n_H] = \rho n$). Given n_L and n_H , we have:

$$\begin{aligned} \mathbb{E}[|Q|] &= 2 \binom{n_L}{2} \binom{n_H}{2} \left(\frac{\rho c}{n}\right)^3 \left(1 - \frac{\rho c}{n}\right)^{n_L-1} \left(1 - \frac{\rho c}{n}\right)^{n_H-2} \left(1 - \frac{\rho c}{n}\right)^{n_L-2} \left(1 - \frac{\rho c}{n}\right)^{n_H-1} (1 - o(1)) \\ &= 2 \binom{n_L}{2} \binom{n_H}{2} \left(\frac{\rho c}{n}\right)^3 e^{-2\rho c} (1 - o(1)). \end{aligned} \tag{7}$$

where the factor $(1 - o(1))$ in the first line accounts for the probability that the H-nodes u_1 and u_3 are not connected to any other H nodes. We are left to compute $\mathbb{E}\left[\binom{n_L}{2}\binom{n_H}{2}\right]$. Note that n_H is a binomial random variable with mean ρn . Thus by Chernoff's inequality,

$$\mathbb{P}(n_H \leq \rho n/2) \leq e^{-\frac{\rho n}{8}}.$$

Also function $\binom{n_H}{2}$ is increasing in n_H . Therefore,

$$\mathbb{E} \left[\binom{n_L}{2} \binom{n_H}{2} \right] \geq \binom{\rho n/2}{2} \mathbb{E} \left[\binom{n_L}{2} \right] (1 - e^{-\frac{\rho n}{8}}) \geq \frac{\rho^2(1-\rho)^2 n^4}{16} (1 - o(1)), \quad (8)$$

where we used the Jensen's inequality which asserts that $\mathbb{E} \left[\binom{m_L}{2} \right] \geq \binom{\mathbb{E}[m_L]}{2}$. Putting (7) and (8) together show that $\mathbb{E}[|Q|] = \Theta(n)$ and complete the proof. \square

Claim C.5. *In expectation, the set P consists of at least $\delta_{0.5}n$ vertex-disjoint paths.*

Proof. We prove the claim using the set of paths Q defined in the proof of Claim C.2. Note that the set $Q \cap P$ contains only disjoint (isolated) paths, and we have already shown in the proof of Claim C.2 that the expected size of this set is $\Theta(n)$. \square

The above claim implies that there are $\Theta(n)$ distinct H-nodes like w_2 that can be matched by CM^2 with $S_H = S_L = n$ by not by CM^2 with $S_H = S_L = n/2$, and this complete the proof for $\beta = 1/2$.

For any $\beta < 1/2$, we can slightly modify the above proof to show the linear loss of CM^2 with $S_H = S_L = \beta n$ compared to CM^2 with $S_H = S_L = n$. Define set Z_1 to be the set of nodes matched in the first chunk (i.e, at time βn) and excluding the L-L edges, and the set Z_2 to be the set of nodes arrive after time βn and were not matched through any H-L edge in any later chunks. Again, we show that there are $\Theta(n)$ disjoint augmenting paths of the same form described above. In particular, the number of such disjoint paths is lower bounded by the number of paths $p \in Q$ (defined in the proof of Claim C.2) in which nodes u_2 and u_3 arrive before time βn and nodes u_1 and u_4 arrive after βn . For each $p \in Q$ this happens with a constant probability (more precisely with probability $\beta^2(1-\beta)^2 - o(1)$), and we showed that $\mathbb{E}[|Q|] = \Theta(n)$. This completes the proof for any β . \square