

The Stochastic Matching Problem: Beating Half with a Non-Adaptive Algorithm

SEPEHR ASSADI, University of Pennsylvania
SANJEEV KHANNA, University of Pennsylvania
YANG LI, University of Pennsylvania

In the *stochastic matching* problem, we are given a general (not necessarily bipartite) graph $G(V, E)$, where each edge in E is *realized* with some constant probability $p > 0$ and the goal is to compute a *bounded-degree* (bounded by a function depending only on p) subgraph H of G such that the expected maximum matching size in H is close to the expected maximum matching size in G . The algorithms in this setting are considered *non-adaptive* as they have to choose the subgraph H without knowing any information about the set of realized edges in G . Originally motivated by an application to *kidney exchange*, the stochastic matching problem and its variants have received significant attention in recent years.

The state-of-the-art non-adaptive algorithms for stochastic matching achieve an approximation ratio of $\frac{1}{2} - \epsilon$ for any $\epsilon > 0$, naturally raising the question that if $1/2$ is the limit of what can be achieved with a non-adaptive algorithm. In this work, we resolve this question by presenting the first algorithm for stochastic matching with an approximation guarantee that is strictly better than $1/2$: the algorithm computes a subgraph H of G with the maximum degree $O\left(\frac{\log(1/p)}{p}\right)$ such that the ratio of expected size of a maximum matching in realizations of H and G is at least $1/2 + \delta_0$ for some absolute constant $\delta_0 > 0$. The degree bound on H achieved by our algorithm is essentially the best possible (up to an $O(\log(1/p))$ factor) for any *constant factor* approximation algorithm, since an $\Omega(\frac{1}{p})$ degree in H is necessary for a vertex to acquire at least one incident edge in a realization.

Our result makes progress towards answering an open problem of Blum *et al.* (EC 2015) regarding the possibility of achieving a $(1 - \epsilon)$ -approximation for the stochastic matching problem using non-adaptive algorithms. From the technical point of view, a key ingredient of our algorithm is a structural result showing that a graph whose expected maximum matching size is OPT always contains a b -matching of size (essentially) $b \cdot \text{OPT}$, for $b = \frac{1}{p}$.

CCS Concepts: • **Theory of computation** → **Design and analysis of algorithms**;

Additional Key Words and Phrases: stochastic matching, kidney exchange

1 INTRODUCTION

We study the problem of finding a maximum matching in presence of *uncertainty* in the input graph. Specifically, we consider the *stochastic* setting where for an input graph $G(V, E)$ and a parameter $p > 0$, each edge in E is realized *independently* w.p.¹ p . We call the graph obtained from this stochastic process (which should be viewed as a random variable) a *realization* of $G(V, E)$,

¹Throughout, we use *w.p.*, *w.h.p.*, and *prob.* to abbreviate “with probability”, “with high probability”, and “probability”, respectively.

A full version of this paper is available as an online preprint at [7].

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

EC’17, June 26–30, 2017, Cambridge, Massachusetts, USA. ©2017 ACM. ISBN 978-1-4503-4527-9/17/06...\$15.00

DOI: <https://doi.org/10.1145/3033274.3085146>

denoted by $G_p(V, E_p)$. The *stochastic matching* problem can now be defined as follows. Given a general (not necessarily bipartite) graph $G(V, E)$ and an edge realization probability $p > 0$, compute a subgraph H of G such that:

- (i) The expected maximum matching size in a realization of H is close to the expected maximum matching size in a realization of G .
- (ii) The degree of each vertex in H is bounded by some function that only depends on p , independent of the size of G .

In other words, the stochastic matching problem asks if every graph G contains a subgraph H of *bounded degree* (depending only on the realization probability p) such that the expected matching size in realizations of G and H are close.

Kidney exchange. A canonical and arguably the most important application of the stochastic matching problem appears in *kidney exchange*, where patients waiting for kidney transplant can *swap* their incompatible donors to each get a compatible donor. The goal is to identify a maximum set of patient-donor pairs to perform such a swap (i.e., finds a maximum matching). However, through medical records of patients and donors, one can only filter out the patient-donor pairs where donation is *impossible*, and more costly and time consuming tests must be performed before a transplant can be performed.

The stochastic setting captures the essence of the need of extra tests for kidney exchange: an algorithm selects a set of patient-donor pairs to perform the extra tests (i.e., computes a subgraph H), while making sure that there is a large matching among the pairs that pass the extra tests. The objective that the subgraph H has small degree captures the essence of minimizing the number of (costly and time consuming) tests that each patient needs to go through. The kidney exchange problem has been extensively studied in the literature, particularly under stochastic settings (see, e.g., [2–4, 8, 16–18, 25, 30]). We remark that the the stochastic matching problem captures the simplest form of the kidney exchange, referred to as *pairwise exchange*. Modern kidney exchange programs regularly employ swaps between three or patient-donor pairs and this setting has also been studied previously in the literature; we refer the interested reader to [10] for more details.

Previous work. Our results are directly related to the results in [10] and [6] which we describe in detail below. Blum *et al.* [10] introduced the (variant of) stochastic matching problem and proposed a $(\frac{1}{2} - \epsilon)$ -approximation algorithm (for any $\epsilon > 0$) which requires the subgraph H to have maximum degree of $\frac{\log(1/\epsilon)}{p^{\Theta(1/\epsilon)}}$. The algorithm of Blum *et al.* [10] works as follows: Pick a maximum matching M_i in G and remove the edges in M_i ; repeat for $R := \frac{\log(1/\epsilon)}{p^{\Theta(1/\epsilon)}}$ times. In order to analyze this algorithm, the authors showed that, for any $i \in [R]$, if the size of the maximum matching among the realized edges in M_1, \dots, M_i is less than $\text{OPT}/2$, the matching M_{i+1} contains many augmenting paths of M of length $O(\frac{1}{\epsilon})$; since each such augmenting path is realized w.p. $p^{O(\frac{1}{\epsilon})}$, one needs to repeat this augmentation process for $\frac{1}{p^{O(\frac{1}{\epsilon})}}$ time (as is roughly the value of R) to increase the matching size to $(\frac{1}{2} - \epsilon) \cdot \text{OPT}$.

In a recent work [6], we showed that in order to obtain a $(\frac{1}{2} - \epsilon)$ -approximation algorithm, one only needs a subgraph H with max-degree of $O(\frac{\log(1/\epsilon p)}{\epsilon p})$, significantly smaller than the bounds in [10]. Interestingly, the algorithm of [6] and the one in [10] are essentially identical (modulo an extra sparsification part required in [6]) and the main difference is in the analysis. In [6], we completely bypassed the need for using augmenting paths in the analysis and instead, took advantage of structural properties of matchings in a global manner (by using *Tutte-Berge formula*; see, e.g., [24]). In particular, we showed that repeatedly picking $O(\frac{\log(1/\epsilon p)}{\epsilon p})$ maximum matchings (as described before) suffices to ensure that, among the chosen edges, a matching of size (essentially) equal to the size of the last

chosen matching would be realized (with high probability). Having this, one can show that running the aforementioned algorithm even for $R := O\left(\frac{\log(1/\varepsilon p)}{\varepsilon p}\right)$ suffices to obtain a $(\frac{1}{2} - \varepsilon)$ -approximation.

Adaptive algorithms for stochastic matching have also been studied by [6, 10]. In an adaptive algorithm, instead of a single graph H , one is allowed to pick a *small* number of bounded-degree graphs H_1, \dots, H_k where the choice of each H_i can be made after *probing* the edges in H_1, H_2, \dots, H_{i-1} to see if they are realized or not. A $(1 - \varepsilon)$ -approximation adaptive algorithm for this problem was first proposed in [10] and further refined in [6].

Beating the half approximation. This state-of-the-art highlights the following natural question:

Is half-approximation the limit for non-adaptive algorithms or is there a non-adaptive algorithm that achieves approximation guarantee of strictly better than half?

It is worth mentioning that in many variations, obtaining half approximation for the maximum matching problem is typically a relatively easy task (usually via a greedy approach), while beating half approximation turns out to be difficult. Some notable examples include, randomized greedy matching [5, 13, 19, 28], online stochastic matching [22, 26, 27], and semi-streaming matching [20, 23].

1.1 Our Contributions

We resolve the aforementioned question of obtaining an algorithm for stochastic matching with an approximation guarantee of *strictly* better than half. Formally,

THEOREM 1.1. *There exists an algorithm that given any graph $G(V, E)$ and any parameter $p > 0$, computes a subgraph $H(V, Q)$ of G with a maximum degree of $O\left(\frac{\log(1/p)}{p}\right)$ such that the ratio of the expected maximum matching size of a realization of H to a realization of G is at least:*

- (i) 0.52 when $p \leq p_0$ for an absolute constant $p_0 > 0$.
- (ii) $0.5 + \delta_0$ for any $0 < p < 1$, where $\delta_0 > 0$ is an absolute constant.

Our result in Theorem 1.1 makes progress towards an open problem posed by Blum *et al.* [10] regarding the possibility of having a non-adaptive $(1 - \varepsilon)$ -approximation algorithm for stochastic matching. We further remark that the assumption in Part (i) of Theorem 1.1 is standard in the stochastic matching literature and is referred as the case of *vanishing probabilities*, see, e.g. [26, 27].

It is worth mentioning the max-degree on H achieved in Theorem 1.1 is essentially the best possible (up to an $O(\log \frac{1}{p})$ factor) for any *constant factor* approximation algorithm: suppose G is a complete graph; in this case the expected matching size in G is $n - o(n)$ by standard results on random graphs (see, e.g., [12], Chapter 7); however, if max-degree of H is $o(\frac{1}{p})$, then the expected number of realized edges in H is $o(n)$, implying that the expected matching size in H is $o(n)$.

Our approach to proving Theorem 1.1 can be divided into two parts. In the first part, we prove a structural result showing that if a realization of G has expected maximum matching size OPT , then G itself should contain essentially $\frac{1}{p}$ edge-disjoint matchings of size OPT each. This result, established through a characterization of b -matching size in general graphs (see Section 3), sheds more light into the structure of a graph in terms of its expected maximum matching size, which may be of independent interest.

In the second part, we combine the aforementioned structural result with the $(\frac{1}{2} - \varepsilon)$ -approximation algorithm of [6] to obtain a matching of size strictly larger than $\text{OPT}/2$. In order to do this, we first find a collection of $\frac{1}{p}$ edge-disjoint matchings of size at least OPT , remove them from the graph, and then run the algorithm of [6] on the remaining edges. We show that the edges in this collection of edge-disjoint matchings must form many *length-three augmenting paths* of the matching computed

by the algorithm of [6], hence leading to a matching of size strictly larger than $\text{OPT}/2$. The analysis is separated into two steps: we first formulate the increment in the matching size (through these augmenting paths) via a (non-linear) minimization program, and then analyze the optimal solution of this minimization program and hence lower bound the increment in the matching size obtained from the augmenting paths.

Other related work. Multiple variants of stochastic matching have been considered in the literature. Blum *et al.* [11] studied a similar setting where one can only probe *two* edges incident on any vertex and the goal is to find the optimal set of edges to query. Another well studied setting is the *query-commit* model, whereby an algorithm probes one edge at a time and if an edge e is probed and realized, then the algorithm must take e as part of the matching it outputs [1, 9, 14, 15, 21]. We refer the reader to [10] for a detailed description of the related work.

2 TECHNICAL OVERVIEW

In this section, we give a more detailed overview of the main ideas used in our algorithm for stochastic matching. For clarity of exposition, throughout this section, we assume p is a sufficiently small constant (corresponding to Part (i) of Theorem 1.1) and the expected maximum matching size in G (i.e., OPT) is $n - o(n)$, or in other words, a realization of G, G_p , has a near perfect matching in expectation.

Our starting point is the following observation: In order for G_p to have a (near) perfect matching in expectation, the input graph G must have many (roughly $1/p$) edge-disjoint (near) perfect matchings. To gain some intuition why this is true, suppose for the moment that the input graph is a bipartite graph $G(L, R, E)$. Then, by *Hall's Marriage Theorem*, we know that in order for G_p to have a matching of size $n - o(n)$, for any two subsets $X \subseteq L$ and $Y \subseteq R$, with $|X| - |Y| \geq o(n)$, at least one edge from X to \bar{Y} should realize in G_p . However, this requirement implies that in G , there should be $1/p$ edges from X to \bar{Y} so that at least one of these edges appears in G_p . One can then show that a bipartite graph G with such a structure has $1/p$ edge-disjoint matchings of size at least $n - o(n)$.

In general, we need to handle graphs that are not necessarily bipartite. In order to adapt the previous strategy, we slightly relax our requirement of having $1/p$ edge-disjoint matchings to having one (simple) b -matching² of size nb for the parameter $b = \frac{1}{p}$. We show that,

b -Matching Lemma. Any graph G where G_p has a matching of size $n - o(n)$ in expectation, has a $\frac{1}{p}$ -matching of size (essentially) $\frac{n}{p}$.

Next, we combine the fact that a large $\frac{1}{p}$ -matching, denoted by B , always exists in G , with the $(\frac{1}{2} - \epsilon)$ -approximation algorithm of [6] to obtain a strictly better than $\frac{1}{2}$ -approximation algorithm.

To continue, we briefly describe the algorithm of [6], which we refer to as *MatchingCover*. *MatchingCover* works by picking a maximum matching M_i in G and removing the edges of M_i for $R := \Theta\left(\frac{\log(1/p)}{p}\right)$ times³. This collection of matchings, denoted by E_{MC} , is referred to as a *matching cover* of the original graph G . The main property of this matching cover, proved in [6], is that the set of realized edges in E_{MC} has a matching of size (essentially) $|M_R|$; note that M_R is the smallest size matchings among the matchings in E_{MC} .

We are now ready to define our main algorithm: Pick a maximum $\frac{1}{p}$ -matching B from G ; run *MatchingCover* over the edges $E \setminus B$ and obtain a matching cover E_{MC} ; return $H(V, B \cup E_{MC})$. If $|M_R| < (\frac{1}{2} - \delta_0)n$, using the fact that E_{MC} is obtained by repeatedly picking maximum matchings,

²Recall that a (simple) b -matching is simply a graph with degree of each vertex bounded by b . See Section 3 for more details.

³We remark that this algorithm has an extra *sparsification* step which is needed to handle the case where $\text{OPT} = o(n)$. However, since in this section we assume $\text{OPT} = n - o(n)$, this extra step is not required.

one can show that any matching M of size $n - o(n)$ in G has more than $(\frac{1}{2} + \delta_0)n - o(n)$ edges in $B \cup E_{MC}$. This also implies that the expected matching size in H is at least $(\frac{1}{2} + \delta_0)n - o(n)$. The more difficult case, which is where we concentrate bulk of our technical effort, is when $|M_R| \geq (\frac{1}{2} - \delta_0)n$. For simplicity, assume $|M_R| = n/2$ from here on.

As stated above, if $|M_R| = n/2$, then in almost every realization of the edges in E_{MC} , there exists a matching M of size at least $n/2$. Our strategy is to *augment* the matching M using the (realized) edges in B , so that the matching size becomes $(\frac{1}{2} + \delta_0)n$. It is important to note that the set of edges in E_{MC} and B are disjoint, and hence whether edges in E_{MC} and B are realized are independent of each other.

Let U be the set of vertices matched by M . There are two cases here to consider:

- **Case 1.** Nearly all edges in B are incident on vertices in U .
- **Case 2.** An ε -fraction of edges in B are not incident on U (for some constant $\varepsilon > 0$).

The second case is relatively easy to handle: we show that a realization of a $\frac{1}{p}$ -matching with N/p edges has a matching of size at least $N/3$ in expectation. This implies that B_p has a matching M' of size $\varepsilon \cdot \frac{n}{3} = \Theta(\varepsilon) \cdot n$ which is not incident on U . Consequently, $B \cup E_{MC}$ has a matching of size $\frac{n}{2} + \Theta(\varepsilon) \cdot n$ in expectation. The more challenging task is to tackle the first case. To convey the main idea, we make a series of simplifying assumptions here: (i) all edges in B are incident on U , (ii) each edge in B is incident on exactly one vertex in U , and (iii) every vertex in U is incident on exactly $\frac{1}{p}$ edges of B .

Our goal is to identify a large collection of length-three augmenting paths for the matching M using the edges of B . To achieve this, we consider the event that an edge (u, v) in M has a length-three augmenting path $a - u - v - b$ where u (resp. v) is the only neighbor of a (resp. b). We say such an edge (u, v) is *successful*. Since the length-three augmenting that certifies successful edges are vertex-disjoint by definition, they can all (simultaneously) augment M . Consequently, it suffices to lower bound the expected number of successful edges, or, equivalently, to lower bound the prob. that each edge is successful.

Let us further assume for the moment that G is a bipartite graph. In this case, u and v do not share a common neighbor and we can consider the neighborhood of u and v separately. The prob. that u has a neighbor w where u is the only neighbor of w (we say u is successful in this case) is not difficult to bound: enumerate all $1/p$ neighbors w of u and account for the the prob. that the edge (u, w) is realized and the prob. that no other edge incident on w is realized. A similar argument can be made for v . Now, the prob. that (u, v) is successful is simply the product of the prob. that u is successful and the prob. that v is successful.

However, in general (non-bipartite) graphs, u and v might have common neighbors which results in prob. of u being successful *not independent* of prob. of v being successful. Handling this case requires a more careful argument and analysis. Moreover, recall that in the above discussion, we made rather strong simplifying assumptions about how the edges in B are distributed across the vertices of U . In order to further remove these assumptions, in the actual analysis, we cast the probability of each edge (u, v) being successful as a function of the degrees of the vertices u and v , and formulate a (non-linear) minimization program to capture the minimum number of possible successful edges. Finally, we analyze the optimal solution of this minimization program, which allows us to achieve the target lower bound on the expected increment in the matching size.

3 PRELIMINARIES

Notation. For a graph $G(V, E)$, n denotes the number of vertices in G . For any $U \subseteq V$, we use $G[U]$ to denote the subgraph of G induced only on vertices in U , and use $E[U]$ to denote the set of edges in $G[U]$, i.e., the set of edges with both end points in U . For any two subsets U, W of V , we further use $E[U, W]$ to denote the set of edges with one end point in U and another in W . For

any $X \subseteq E$, we use $V(X)$ to denote the set of vertices incident on X . Finally, we use $\mu(E)$ to denote the maximum matching size among a set of edge E .

When sampling from a set of edges X (resp. a graph H) where each edge in X (resp. H) is sampled w.p. p , we use X_p (resp. H_p) to denote the random variable for the set of sampled edges. We use $\text{OPT}(G)$ (or shortly OPT if the graph G is clear from the context) to denote the *expected* maximum matching size of a realization of G (i.e., $G_p(V, E_p)$)⁴. For any algorithm for the stochastic matching problem, we use ALG to denote the expected matching size in a realization of H , where H is the subgraph computed by the algorithm.

b-matchings. For any graph $G(V, E)$ and any integer $b \geq 1$, a subset $M \subseteq E$ is called a *simple b-matching*, iff the number of edges M that are incident on each vertex is at most b . Throughout, we drop the word ‘simple’, and refer to M as a *b-matching*.

We use the following characterization of the maximum b -matching size in general graphs (see [29], Volume A, Chapter 33).

THEOREM 3.1. *Let $G(V, E)$ be a graph and $b \geq 1$ be any integer. The maximum size of a b -matching is equal to the minimum value of*

$$b \cdot |U| + |E[W]| + \sum_K \left[\frac{1}{2} (b \cdot |K| + |E[K, W]|) \right]$$

taken over all disjoint subsets U, W of V , where K ranges over all connected components in the graph $G[V - U - W]$.

Useful inequalities. We also use the following simple inequalities (see [7] for the proofs).

PROPOSITION 3.2. *Let $f(x) := \frac{1-e^{-x}}{x}$. Then, for any c , and any $x \in [0, c]$, $e^{-x} \leq 1 - f(c) \cdot x$.*

PROPOSITION 3.3. *For any $x \in (0, 0.43]$, $(1 - x)^{\frac{1}{x}} \geq \frac{1-x}{e}$.*

3.1 MatchingCover Algorithm

We use the $(0.5 - \varepsilon)$ -approximation algorithm of [6] (Algorithm 3) as a sub-routine. For simplicity, throughout the paper, we refer to this algorithm as *MatchingCover*. In the following lemma, we summarize the properties of *MatchingCover* that we use in this paper. The proof of this lemma immediately follows from Lemma 3.9 and Lemma 5.2 in [6].

LEMMA 3.4 ([6]). *For any graph $G(V, E)$, and any input parameter $\varepsilon > 0$, $\text{MatchingCover}(G, \varepsilon)$ outputs a collection of R matchings M_1, M_2, \dots, M_R (denote $E_{MC} = M_1 \cup M_2 \cup \dots \cup M_R$), such that, w.p. $1 - o(1)$:*

- (1) *The size of a maximum matching among realized edges in E_{MC} is at least $(1 - \varepsilon) |M_R|$.*
- (2) *$|M_1| \geq \dots \geq |M_R| \geq (1 - \varepsilon) \cdot \mu(E \setminus E_{MC})$.*
- (3) *$R = \Theta\left(\frac{\log 1/(\varepsilon p)}{\varepsilon p}\right)$.*

We can also prove the following simple claim based on the second property of the *MatchingCover* in Lemma 3.4. Roughly speaking, this claim states that if the *MatchingCover* is not able to extract any further large matching (of size essentially $\text{OPT}/2$) from G , then the set of extracted edges already provides a matching of size $\text{OPT}/2$ in any realization. We defer the proof of this claim to the full version of the paper [7]; we also note that a similar result is proven in [6] (see Lemma 5.3).

⁴We assume $\text{OPT} = \omega(1)$ to obtain the desired concentration bounds (for example in Lemma 3.4).

CLAIM 3.5. Fix $0 < \varepsilon < \delta < 1$. Let $G(V, E)$ be a graph, X be any arbitrary subset of E , and $(M_1, \dots, M_R) = \text{MatchingCover}(G(V, E \setminus X), \varepsilon)$. Define $E_{MC} = M_1 \cup \dots \cup M_R$. If $|M_R| \leq \left(\frac{1}{2} - \delta\right) \text{OPT}$, then the expected maximum matching size in a realization of $G(V, X \cup E_{MC})$ is at least $\left(\frac{1}{2} + \delta - \varepsilon\right) \text{OPT}$.

4 B-MATCHING LEMMA

Here, we develop one of the main ingredients of our algorithm, namely, any input graph G contains a b -matching of size almost $b \cdot \text{OPT}(G)$ for $b = 1/p$. Intuitively, if the expected matching size in G is OPT , then since only p fraction of edges are realized in expectation, one may hope to find up to $1/p$ edge-disjoint matchings of size OPT in G . The following lemma formalizes this intuition by using b -matchings (for $b = 1/p$) instead of a collection of edge-disjoint matchings.

LEMMA 4.1 (*b*-MATCHING LEMMA). Let $b = \left\lfloor \frac{1}{p} \right\rfloor$; any graph $G(V, E)$ has a b -matching of size at least $(b - 1) \cdot \text{OPT}(G)$.

PROOF. Suppose by contradiction that the maximum b -matching B in G is of size less than $(b - 1) \cdot \text{OPT}$. Consequently, by Theorem 3.1, there exist disjoint subsets U, W of V such that,

$$b \cdot |U| + |E[W]| + \sum_K \left\lfloor \frac{1}{2} \left(b \cdot |K| + |E[K, W]| \right) \right\rfloor < (b - 1) \cdot \text{OPT} \quad (1)$$

where K ranges over all connected components in the graph $G[V - U - W]$. Let c be the number of connected components in $G[V - U - W]$. We first note that $c < 2\text{OPT}$; otherwise,

$$\sum_K \left\lfloor \frac{1}{2} \left(b \cdot |K| + |E[K, W]| \right) \right\rfloor \geq \sum_K \left\lfloor \frac{b}{2} \right\rfloor \geq c \cdot \left(\frac{b-1}{2} \right) \geq (b-1) \cdot \text{OPT}$$

and hence the LHS in Eq (1) would be more than $(b - 1) \cdot \text{OPT}$, i.e., the RHS; a contradiction.

Additionally, we have $|U| + |W| + \sum_K |K| = n$. Hence, by multiplying each side in Eq (1) by 2 and plugging in this bound, we have,

$$\begin{aligned} 2b \cdot \text{OPT} - 2\text{OPT} &> nb - b|W| + b|U| + 2|E[W]| + \sum_K (|E[K, W]| - 1) \\ &\geq nb - b|W| + b|U| + 2|E[W]| + \sum_K |E[K, W]| - 2\text{OPT} \end{aligned}$$

Let $T := V \setminus (U \cup W)$, i.e., the set of vertices in connected components K . Using this notation, we can write the above equation simply as,

$$b \cdot |W| - b \cdot |U| - 2|E[W]| - |E[T, W]| > b \cdot (n - 2\text{OPT}) \quad (2)$$

Now consider the partition T, U, W in a *realized* graph $G(V, E_p)$. Let $E_p[W]$ and $E_p[T, W]$ denote, respectively, the set of edges in $E[W]$ and $E[T, W]$ after sampling the edges w.p. p . For any matching M in G_p , define $x(M)$ to be the number of *unmatched* vertices (by M) in W . Finally, define $x^* := \min_M x(M)$, where the minimum is taken over all matchings in G_p . Clearly, x^* is a random variable depending on the choice of edges in G_p . We have the following simple claim.

CLAIM 4.2. For any realization G_p , $x^* \geq |W| - |U| - 2|E_p[W]| - |E_p[T, W]|$.

PROOF. Consider the set of vertices in W . At most $|U|$ vertices of W can be matched to vertices in U . Additionally, any edge in $E_p[W]$ can further reduce the number of unmatched vertices in W by at most 2. Finally, any edge in $E_p[T, W]$ can reduce the number of remaining unmatched vertices in W by at most 1. \square

Using the fact that $\mathbb{E}[x^*] \leq n - 2\text{OPT}$, we have,

$$\begin{aligned}
b \cdot (n - 2\text{OPT}) &\geq b \cdot \mathbb{E}[x^*] \\
&\geq b \cdot |W| - b \cdot |U| - b \cdot \mathbb{E} \left[2 |E_p[W]| + |E_p[T, W]| \right] && \text{(by Claim 4.2)} \\
&= b \cdot |W| - b \cdot |U| - pb \cdot (2 |E[W]| + |E[T, W]|) \\
&\geq b \cdot |W| - b \cdot |U| - 2 |E[W]| - |E[T, W]| && \text{(since } pb = p \lfloor \frac{1}{p} \rfloor \leq 1) \\
&> b \cdot (n - 2\text{OPT}) && \text{(by Eq (2))}
\end{aligned}$$

a contradiction. \square

We can further show that the bound established in Lemma 4.1 is essentially tight (see the full version of the paper [7]).

CLAIM 4.3. *For any constant $0 < p < 1$, there exist bipartite graphs G where G_p has a matching of size $n - o(n)$ in expectation, but for any $b \geq \frac{2}{p}$, there is no b -matching in G with (at least) $b \cdot 0.99n$ edges; here n is the number of vertices on each side of G .*

Finally, we establish the following auxiliary lemma. The proof is deferred to the full version [7].

LEMMA 4.4. *Let B be a $\lfloor \frac{1}{p} \rfloor$ -matching with $\left(\lfloor \frac{1}{p} \rfloor \cdot N\right)$ edges; then, $\mathbb{E} \left[\mu(B_p) \right] \geq (1 - 3p) \cdot \frac{N}{3}$.*

5 MAIN ALGORITHM AND ANALYSIS

We provide our main algorithm for the stochastic matching problem (when p is sufficiently small) in this section and prove Part (i) of Theorem 1.1. We assume throughout this section that the edge realization probability $p \leq p_0$ for some sufficiently small constant p_0 . In this case, $\lfloor \frac{1}{p} \rfloor - 1 \geq (1 - O(p_0)) \cdot \frac{1}{p}$ and we use this inequality frequently in the proof. Indeed, throughout this section, one should view p_0 as a negligible constant and hence the term $(1 - O(p_0))$ can essentially be ignored.

Let $\delta_0 = 0.02$, and $\varepsilon_0 = 0.02001$. Our algorithm is stated as Algorithm 1 below:

ALGORITHM 1: A 0.52-Approximation Algorithm for Stochastic Matching

Input: A graph $G(V, E)$ and an edge realization probability $p \leq p_0$.

Output: A subgraph $H(V, Q)$ of $G(V, E)$.

- (1) Let B be a maximum $\lfloor \frac{1}{p} \rfloor$ -matching in G .
 - (2) Let $(M_1, M_2, \dots, M_R) := \text{MatchingCover}(G(V, E \setminus B), \varepsilon_1)$
for $\varepsilon_1 = (\varepsilon_0 - \delta_0)/2$, and $E_{MC} = M_1 \cup \dots \cup M_R$.
 - (3) Return $H(V, Q)$ where $Q := B \cup E_{MC}$.
-

Each vertex in H has degree $O\left(\frac{\log(1/p)}{p}\right)$ – this follows immediately from Lemma 3.4. In what follows, we prove that H has a matching of size at least $(0.5 + \delta_0) \cdot \text{OPT} = 0.52 \cdot \text{OPT}$ in expectation, which will complete the proof of Part (i) of Theorem 1.1.

First notice that if $|M_R| < \left(\frac{1}{2} - \frac{\varepsilon_0 + \delta_0}{2}\right)\text{OPT}$ where M_R is the smallest matching in the matching cover E_{MC} found by Algorithm 1, then by Claim 3.5, the expected matching size in Q is at least $\left(\frac{1}{2} + \frac{\varepsilon_0 + \delta_0}{2} - \frac{\varepsilon_0 - \delta_0}{2}\right)\text{OPT} = \left(\frac{1}{2} + \delta_0\right) \cdot \text{OPT}$. Therefore, from now on we focus on the case that $|M_R| \geq \left(\frac{1}{2} - \frac{\varepsilon_0 + \delta_0}{2}\right)\text{OPT}$.

In this case, by Lemma 3.4, w.p. $1 - o(1)$, there exists a matching M among the realized edges in E_{MC} with size at least

$$\begin{aligned} \left(1 - \frac{\varepsilon_0 - \delta_0}{2}\right) \left(\frac{1}{2} - \frac{\varepsilon_0 + \delta_0}{2}\right) \text{OPT} &\geq \left(\frac{1}{2} - \frac{\varepsilon_0 + \delta_0}{2} - \frac{\varepsilon_0 - \delta_0}{4}\right) \text{OPT} \\ &= \left(\frac{1}{2} - \frac{3\varepsilon_0}{4} - \frac{\delta_0}{4}\right) \text{OPT} \geq \left(\frac{1}{2} - \varepsilon_0\right) \text{OPT} \end{aligned}$$

In the following, we assume this event happens⁵ and prove that the set of edges realized in the $\lfloor \frac{1}{p} \rfloor$ -matching B can be used to augment the matching M to create a matching of size $(\frac{1}{2} + \delta_0) \cdot \text{OPT}$ in expectation. To simplify the analysis, we assume w.l.o.g. that $|M| = (\frac{1}{2} - \varepsilon_0) \text{OPT}$ (i.e., we only keep $(\frac{1}{2} - \varepsilon_0) \text{OPT}$ edges of M and remove any additional edges if there is any). By the b -matching lemma (Lemma 4.1), $|B| \geq \left(\lfloor \frac{1}{p} \rfloor - 1\right) \text{OPT} \geq (1 - O(p_0)) \cdot \frac{\text{OPT}}{p}$, and hence, to prove Part (1) of Theorem 1.1, it suffices to prove the following statement.

LEMMA 5.1. *Let M be a matching of size $(\frac{1}{2} - \varepsilon_0) \text{OPT}$, and B be a $\lfloor \frac{1}{p} \rfloor$ -matching of size at least $(1 - O(p_0)) \cdot \frac{\text{OPT}}{p}$; then the expected maximum matching size in $M \cup B_{\overline{M}}$ is at least $(\frac{1}{2} + \delta_0) \text{OPT}$.*

PROOF. Let B_M be the set of edges in B that are incident on the vertices in the matching M , and let $B_{\overline{M}} = B \setminus B_M$. Let $s_{\overline{M}}$ be the random variable denoting the maximum matching size of a realization of $B_{\overline{M}}$. By Lemma 4.4,

$$\mathbb{E}[s_{\overline{M}}] \geq \frac{|B_{\overline{M}}|}{\lfloor \frac{1}{p} \rfloor} \cdot \frac{1 - 3p}{3} \geq p |B_{\overline{M}}| \cdot \frac{1 - 3p}{3} \quad (3)$$

Therefore, if $|B_{\overline{M}}| \geq 6\varepsilon_0 \cdot \frac{\text{OPT}}{p}$, then

$$\mathbb{E}[s_{\overline{M}}] \geq 6\varepsilon_0 \cdot \text{OPT} \cdot \frac{1 - 3p}{3} = 2\varepsilon_0(1 - 3p) \cdot \text{OPT} \geq (\varepsilon_0 + \delta_0) \cdot \text{OPT} \quad (\text{assuming } p_0 \leq \frac{\varepsilon_0 - \delta_0}{6\varepsilon_0})$$

and since no edge in $B_{\overline{M}}$ is incident on the vertices in M , the expected matching size in $M \cup B_p$ is at least

$$\left(\frac{1}{2} - \varepsilon_0\right) \text{OPT} + (\varepsilon_0 + \delta_0) \text{OPT} = \left(\frac{1}{2} + \delta_0\right) \text{OPT}$$

as asserted by Lemma 5.1. In the following, we assume $|B_{\overline{M}}| \leq 6\varepsilon_0 \cdot \frac{\text{OPT}}{p}$. Furthermore, we fix a realization of $B_{\overline{M}}$ and fix a maximum matching M' in the realization of $B_{\overline{M}}$ (whose size is $s_{\overline{M}}$ by definition). In other words, we will lower bound the expected maximum matching size in $M \cup B_p$ conditioned on *any* realization of $B_{\overline{M}}$. The lower bound we obtain would be a linear function of $s_{\overline{M}}$, and by linearity of expectation, we can simply replace $s_{\overline{M}}$ with $\mathbb{E}[s_{\overline{M}}]$, use Eq (3) to lower bound $\mathbb{E}[s_{\overline{M}}]$, and obtain the desired lower bound of $(1/2 + \delta_0) \text{OPT}$ on the expected maximum matching size.

Denote by M^+ the matching $M \cup M'$ (since the matchings M and M' are vertex-disjoint, $M \cup M'$ is indeed a valid matching of size $|M| + s_{\overline{M}}$). We can focus the realizations of $B_{\overline{M}}$ where $s_{\overline{M}} \leq 2\varepsilon_0 \cdot \text{OPT}$ since otherwise the matching M^+ already have size $(1/2 + \varepsilon_0) \cdot \text{OPT} > (1/2 + \delta_0) \cdot \text{OPT}$. Therefore, we have $s_{\overline{M}} \leq 2\varepsilon_0 \cdot \text{OPT} = O(\text{OPT})$ and $|M^+| = O(\text{OPT})$, which will be useful in simplifying the presentation.

Now consider the edges in B_M . We further denote by C the set of edges in B_M that are incident on *exactly* one vertex in M^+ . In the following, we first show that $|C|$ must be large (Claim 5.2) and then show that many edges in C can be used to augment the matching M^+ , which leads to an

⁵This assumption can be removed while losing a negligible factor of $o(1)$ in the size of final matching.

increment on the matching size as a function of $|C|$ (Lemma 5.3 and Lemma 5.4). Combining these two statements completes the proof of Lemma 5.1.

$$\text{CLAIM 5.2. } |C| \geq 2|B_M| - \frac{2|M^+|}{p}.$$

PROOF. Let x denote the number of edges in B_M that have degree 2 to $V(M^+)$ (i.e., are incident on two vertices in M^+). By definition, every edge in B_M is incident on M , and hence every edge in B_M is also incident on $M^+ (= M \cup M')$. Consequently, there are $|B_M| - x$ edges in B_M that have degree 1 to $V(M^+)$ (i.e. belongs to C). Therefore, the total degrees of all vertices $V(M^+)$ provided by B_M is at least: $2 \cdot x + 1 \cdot (|B_M| - x) = x + |B_M|$.

On the other hand, since $|V(M^+)| = 2|M^+|$ and B (hence B_M) is a $\lfloor \frac{1}{p} \rfloor$ -matching, the total degree of the vertices $V(M^+)$ provided by B_M is at most $\frac{2|M^+|}{p}$. Therefore, $x + |B_M| \leq \frac{2|M^+|}{p}$, which implies $x \leq \frac{2|M^+|}{p} - |B_M|$. Therefore, the number of edges in B_M incident on exactly one vertex in $V(M^+)$ (i.e., $|C|$) is at least

$$|B_M| - \left(\frac{2|M^+|}{p} - |B_M| \right) = 2|B_M| - \frac{2|M^+|}{p}$$

completing the proof. \square

The following two lemmas are dedicated to showing that the edges in a realization of C , C_p , form many vertex-disjoint length-three augmenting paths for the matching M^+ in expectation, which is a lower bound on the expected increment on the matching size. We first define some notation. Let $W := V \setminus V(M^+)$, i.e., W is the set of vertices *not* matched by M^+ . Denote the edges in M^+ by $\{(u_i, v_i) \mid i \in [|M^+|]\}$, and denote by $d(u_i)$ (resp. $d(v_i)$) the number of edges in C incident on u_i (resp. v_i). Since, by definition, the edges in C are only incident on one vertex in $V(M^+)$, $d(u_i)$ (resp. $d(v_i)$) is also the number of edges in C between u_i (resp. v_i) and W . In the following, whenever we say “neighbors” or “degrees”, they are only w.r.t. the edges C . Let $f(x)$ be the function defined in Proposition 3.2.

LEMMA 5.3. For any edge $(u_i, v_i) \in M$, w.p. at least

$$(1 - O(p_0)) \frac{f(1/e) \cdot p}{e^2} \left(1 - e^{-p \cdot d(v_i)}\right) \cdot \max\{d(u_i) - 1, 0\},$$

there exists a length-three augmenting path $a_i - u_i - v_i - b_i$ in the realization C_p of C , such that a_i, b_i have no neighbors other than u_i and v_i .

Note that we can use all edges (u_i, v_i) in M^+ with such an augmenting path $a_i - u_i - v_i - b_i$ to (simultaneously) augment M^+ since these augmenting paths are vertex-disjoint (a_i and b_i are only neighbors of u_i and v_i). Therefore, the expected number of edges in M^+ that has such an augmenting path is a lower bound on the expected increment on the matching size.

PROOF OF LEMMA 5.3. We consider three disjoint subsets of edges in C one by one: (i) the edges between v_i and W , (ii) the edges incident on a specific vertex w in W (excluding the edge (v_i, w)), and (iii) the edges incident on neighbors of u_i other than w (excluding the edges incident on v_i).

First, consider the edges between v_i and W . The prob. that none of these $d(v_i)$ edges are realized is at most

$$(1 - p)^{d(v_i)} \leq e^{-p \cdot d(v_i)}$$

Therefore, w.p. at least $1 - e^{-p \cdot d(v_i)}$, at least one edge between v_i and W is realized. We condition on this event and fix any such edge, denoted by (v_i, b_i) .

Second, consider the edges incident on b_i (excluding the edge (v_i, b_i)). There are at most $1/p$ such edges, and the prob. that none of them is realized is at least

$$(1-p)^{1/p} \geq \frac{1-p}{e} \geq \frac{1-p_0}{e} \quad (p \leq p_0 \leq 0.43)$$

where the first inequality is by Proposition 3.3. In the following, we further condition on no other edges incident on w is realized.

Third, consider all neighbors of u_i other than b_i (there are at least $\max\{d(u_i) - 1, 0\}$ such neighbors) and the edges incident on these neighbors (excluding the edges incident on v_i). For each one of these neighbors w of u_i , the prob. that the edge (u_i, w) is realized (w.p. p) and w does not have any neighbor other than u_i (and possibly v_i) (w.p. at least $\frac{1-p_0}{e}$ by Proposition 3.3) is at least $p \cdot \frac{1-p_0}{e}$. Therefore, the prob. that at least one neighbor of u_i satisfies these two properties is at least

$$\begin{aligned} 1 - \left(1 - p \cdot \frac{1-p_0}{e}\right)^{\max\{d(u_i)-1, 0\}} &\geq 1 - e^{-(1-p_0) \cdot p \cdot \max\{d(u_i)-1, 0\}/e} \\ &\geq f\left(\frac{1}{e}\right) \cdot (1-p_0) \cdot p \cdot \max\{d(u_i) - 1, 0\} / e \end{aligned}$$

where $f(x) = \frac{1-e^{-x}}{x}$ and the second inequality is by Proposition 3.2, using the fact that

$$\frac{(1-p_0) \cdot p \max\{d(u_i) - 1, 0\}}{e} \leq \frac{1}{e} \quad (\text{since } d(u_i) \leq \frac{1}{p})$$

Putting the three steps together, the prob. that there is an augmenting path $a_i - u_i - v_i - b_i$ where a_i and b_i has no neighbors other than u_i and v_i is at least

$$\begin{aligned} &\left(1 - e^{-p \cdot d(v_i)}\right) \cdot \frac{1-p_0}{e} \cdot \frac{f\left(\frac{1}{e}\right) \cdot (1-p_0) \cdot p \cdot \max\{d(u_i) - 1, 0\}}{e} \\ &= (1 - O(p_0)) \cdot f\left(\frac{1}{e}\right) \cdot \frac{p}{e^2} \left(1 - e^{-p \cdot d(v_i)}\right) \cdot \max\{d(u_i) - 1, 0\} \end{aligned}$$

□

As we pointed out after the statement of Lemma 5.3, we need to lower bound the expected number of edges in M^+ that has such an augmenting path, which, by Lemma 5.3, is lower bounded by the function F defined below. For the two vectors $d_u := (d(u_1), \dots, d(u_{|M^+|}))$ and $d_v := (d(v_1), \dots, d(v_{|M^+|}))$,

$$F(d_u, d_v) := \sum_{i \in [|M^+|]} (1 - O(p_0)) \cdot f\left(\frac{1}{e}\right) \cdot \frac{p}{e^2} \left(1 - e^{-p \cdot d(v_i)}\right) \cdot \max\{d(u_i) - 1, 0\}$$

The goal now is to find the smallest value of $F(d_u, d_v)$, with the constraint on the vectors d_u and d_v formulated in the following (non-linear) minimization program (referred to as MP-(4)).

$$\begin{aligned} &\text{minimize} && F(d_u, d_v) \\ &\text{subject to} && \sum_{i \in [|M^+|]} d(u_i) + d(v_i) = |C| \\ & && d(u_i), d(v_i) \in \left[\left[\frac{1}{p}\right]\right] \quad i = 1, \dots, |M^+| \end{aligned} \quad (4)$$

The constraint on each individual $d(u_i)$ and $d(v_i)$ is because C is a $\left[\frac{1}{p}\right]$ -matching. The following lemma lower bounds the value of the objective function in MP-(4).

LEMMA 5.4. Let F^* denote the optimal value of MP-(4); then,

$$F^* \geq (p \cdot |C| - |M^+|) \cdot \eta - O(p_0) \cdot \text{OPT}$$

where $\eta := f(\frac{1}{e}) \cdot \frac{1}{e^2} (1 - e^{-1}) > 0.07157$.

The proof of the Lemma 5.4 is technical, and we defer it to Section 5.1. By Lemma 5.4 and Claim 5.2 (the lower bound on $|C|$) the expected increment (over M^+) of the matching size is at least

$$\begin{aligned} F^* &\geq (p \cdot |C| - |M^+|) \cdot \eta - O(p_0) \cdot \text{OPT} \\ &\geq (p \cdot (2|B_M| - 2|M^+|/p) - |M^+|) \cdot \eta - O(p_0) \cdot \text{OPT} && \text{(By Claim 5.2)} \\ &= (2p|B_M| - 3|M^+|) \cdot \eta - O(p_0) \cdot \text{OPT} \\ &= (2p(|B| - |B_{\overline{M}}|) - 3(|M| + s_{\overline{M}})) \cdot \eta - O(p_0) \cdot \text{OPT} \\ &= (2p|B| - 3|M| - 2p|B_{\overline{M}}| - 3s_{\overline{M}}) \cdot \eta - O(p_0) \cdot \text{OPT} \\ &= \left(2p(1 - O(p_0)) \frac{\text{OPT}}{p} - 3\left(\frac{1}{2} - \varepsilon_0\right) \text{OPT} - 2p|B_{\overline{M}}| - 3s_{\overline{M}}\right) \cdot \eta - O(p_0) \text{OPT} \\ &= \left(\left(\frac{1}{2} + 3\varepsilon_0\right) \text{OPT} - 2p|B_{\overline{M}}| - 3s_{\overline{M}}\right) \cdot \eta - O(p_0) \cdot \text{OPT} \end{aligned}$$

Since the original matching M^+ is of size $(1/2 - \varepsilon_0) \cdot \text{OPT} + s_{\overline{M}}$, the expected matching size in $M \cup B_p$, i.e., $\mu(M \cup B_p)$, is

$$\begin{aligned} \mathbb{E}[\mu(M \cup B_p)] &= \sum_{s_{\overline{M}}} \Pr(s_{\overline{M}}) \mathbb{E}[\mu(M \cup B_p) \mid s_{\overline{M}}] && (\mathbb{E}[X] = \sum_Y \Pr(Y) \mathbb{E}[X \mid Y]) \\ &\geq \sum_{s_{\overline{M}}} \Pr(s_{\overline{M}}) (|M^+| + F^*) \\ &\geq \left(\frac{1}{2} - \varepsilon_0\right) \cdot \text{OPT} + \mathbb{E}[s_{\overline{M}}] + \left(\left(\frac{1}{2} + 3\varepsilon_0\right) \text{OPT} - 2p|B_{\overline{M}}| - 3\mathbb{E}[s_{\overline{M}}]\right) \cdot \eta - O(p_0) \cdot \text{OPT} \\ &\geq \left(\frac{1}{2} + \frac{\eta}{2} - \varepsilon_0 + 3\eta\varepsilon_0\right) \cdot \text{OPT} + (1 - 3\eta)\mathbb{E}[s_{\overline{M}}] - 2p\eta|B_{\overline{M}}| - O(p_0) \cdot \text{OPT} \\ &\geq \left(\frac{1}{2} + \frac{\eta}{2} - \varepsilon_0 + 3\eta\varepsilon_0\right) \cdot \text{OPT} + (1 - 3\eta)(1 - O(p_0)) \frac{p|B_{\overline{M}}|}{3} - 2p\eta|B_{\overline{M}}| - O(p_0) \cdot \text{OPT} \\ &&& \text{(By Equation 3)} \\ &\geq \left(\frac{1}{2} + \frac{\eta}{2} - \varepsilon_0 + 3\eta\varepsilon_0\right) \cdot \text{OPT} + \left(\frac{1}{3} - 3\eta\right) p|B_{\overline{M}}| - O(p_0) \cdot \text{OPT} \end{aligned}$$

Since $\eta \approx 0.07157$, $\frac{1}{3} - 3\eta > 0$, and we have

$$\begin{aligned} &\left(\frac{1}{2} + \frac{\eta}{2} - \varepsilon_0 + 3\eta\varepsilon_0\right) \cdot \text{OPT} + \left(\frac{1}{3} - 3\eta\right) p|B_{\overline{M}}| - O(p_0) \cdot \text{OPT} \\ &\geq \left(\frac{1}{2} + \frac{\eta}{2} - \varepsilon_0 + 3\eta\varepsilon_0\right) \cdot \text{OPT} - O(p_0) \cdot \text{OPT} \\ &> 0.52 \cdot \text{OPT} && (\varepsilon_0 = 0.02001, \eta > 0.07157, \text{ and } p_0 \text{ is sufficiently small.}) \\ &= (1/2 + \delta_0) \cdot \text{OPT} \end{aligned}$$

completing the proof of Lemma 5.1.

5.1 Lower Bounding the Value of MP-(4)

In this section, we prove Lemma 5.4, i.e., the following inequality,

$$F^* (= \min F(d_u, d_v)) \geq (p \cdot |C| - |M^+|) \cdot \eta - O(p_0)_{\text{OPT}}$$

where $\eta := f(\frac{1}{e}) \cdot \frac{1}{e^2} (1 - e^{-1})$.

Recall that

$$\begin{aligned} F(d_u, d_v) &= \sum_i (1 - O(p_0)) \cdot f\left(\frac{1}{e}\right) \cdot \frac{p}{e^2} \left(1 - e^{-p \cdot d(v_i)}\right) \cdot \max\{d(u_i) - 1, 0\} \\ &= (1 - O(p_0)) \cdot f\left(\frac{1}{e}\right) \cdot \frac{p}{e^2} \sum_i \left(1 - e^{-p \cdot d(v_i)}\right) \cdot \max\{d(u_i) - 1, 0\} \end{aligned}$$

Since the term $(1 - O(p_0)) \cdot f(\frac{1}{e}) \cdot \frac{p}{e^2}$ is independent of d_u and d_v ,

$$\arg \min_{d_u, d_v} F = \arg \min_{d_u, d_v} \sum_i \left(1 - e^{-p \cdot d(v_i)}\right) \cdot \max\{d(u_i) - 1, 0\}$$

Define $d(V) := \sum_i d(v_i)$ and $d(U) := \sum_i d(u_i)$; then, $d(V) + d(U) = |C|$. We need to prove that for any choice of $d(V)$ and $d(U)$, the lemma statement holds. First of all, we can assume $d(U) \geq |M^+|$; otherwise, since $d(V) \leq \lfloor \frac{1}{p} \rfloor |M^+|$, we will have

$$|C| = d(V) + d(U) \leq \left\lfloor \frac{1}{p} \right\rfloor |M^+| + |M^+| \leq \left(\frac{1}{p} + 1\right) |M^+|$$

Therefore, for the target lower bound on F^*

$$\begin{aligned} &(p \cdot |C| - |M^+|) \cdot \eta - O(p_0)_{\text{OPT}} \\ &\leq \left(p \left(\frac{1}{p} + 1\right) |M^+| - |M^+|\right) \cdot \eta - O(p_0)_{\text{OPT}} \\ &\leq p_0 \eta |M^+| - O(p_0)_{\text{OPT}} \end{aligned}$$

which can be made negative by choosing the constant hidden in $O(p_0)$ to be 1, proving Lemma 5.4.

We further assume $d(U) - |M^+|$ is an integer multiple of $\lfloor \frac{1}{p} \rfloor - 1$ and $d(V)$ is an integer multiple of $\lfloor \frac{1}{p} \rfloor$. This can be achieved by removing at most $1/p$ edges from $d(U)$ and $d(V)$ respectively. Since F is monotonically increasing with any $d(u_i)$ or $d(v_i)$, removing edges from $d(U)$ and $d(V)$ can only make F^* even smaller. Therefore, if we show that after removing these edges, the target lower bound on F^* holds, then it definitely holds for the original $d(U)$ and $d(V)$. In the following, we fix any $d(U)$ and $d(V)$ where $d(U) - |M^+|$ is an integer multiple of $\lfloor \frac{1}{p} \rfloor - 1$, $d(V)$ is an integer multiple of $\lfloor \frac{1}{p} \rfloor$, and $d(V) + d(U) \geq |C| - \frac{2}{p}$. We prove the following key property of $F(d_u, d_v)$.

LEMMA 5.5. *There exists d_u and d_v that minimizes $F(d_u, d_v)$ where any entry in d_u is either 1 or $\lfloor \frac{1}{p} \rfloor$ and any entry in d_v is either 0 or $\lfloor \frac{1}{p} \rfloor$.*

We first show why Lemma 5.5 implies the target lower bound on F^* , and then prove Lemma 5.5. Fix $d^*(u_i)$ and $d^*(v_i)$ that satisfy the property in Lemma 5.5. Since every entry in d_u^* is either 1 or $\lfloor \frac{1}{p} \rfloor$, the number of 1's in d_u^* is $x := |M^+| - (d(U) - |M^+|) / (\lfloor \frac{1}{p} \rfloor - 1)$. Similarly, the number of 0's in d_v^* is $y := |M^+| - d(V) / \lfloor \frac{1}{p} \rfloor$. Therefore, the number of edges in M^+ where $d^*(v_i) = d^*(u_i) = \lfloor \frac{1}{p} \rfloor$

is at least

$$\begin{aligned}
|M^+| - x - y &= |M^+| - \left(|M^+| - \frac{d(U) - |M^+|}{\lfloor \frac{1}{p} \rfloor - 1} \right) - \left(|M^+| - \frac{d(V)}{\lfloor \frac{1}{p} \rfloor} \right) \\
&= \frac{d(U) - |M^+|}{\lfloor \frac{1}{p} \rfloor - 1} - |M^+| + \frac{d(V)}{\lfloor \frac{1}{p} \rfloor} \\
&\geq \frac{d(U) - |M^+|}{\frac{1}{p}} - |M^+| + d(V)p && \left(\lfloor \frac{1}{p} \rfloor - 1 \leq \lfloor \frac{1}{p} \rfloor \leq \frac{1}{p} \right) \\
&= p \cdot d(U) + p \cdot d(V) - (1+p)|M^+| \\
&\geq p \left(|C| - \frac{2}{p} \right) - (1+p)|M^+| && (d(V) + d(U) \geq |C| - \frac{2}{p}) \\
&= p|C| - |M^+| - O(p_0)_{\text{OPT}} && (|M^+| = O(\text{OPT}))
\end{aligned}$$

And just focusing on these $p|C| - |M^+| - O(p_0)_{\text{OPT}}$ edges, we have

$$\begin{aligned}
F^* &\geq (p|C| - |M^+| - O(p_0)_{\text{OPT}}) \cdot (1 - O(p_0)) \cdot f\left(\frac{1}{e}\right) \cdot \frac{p}{e^2} (1 - e^{-1}) \cdot \left(\lfloor \frac{1}{p} \rfloor - 1 \right) \\
&\geq (p|C| - |M^+| - O(p_0)_{\text{OPT}}) \cdot (1 - O(p_0)) \cdot f\left(\frac{1}{e}\right) \cdot \frac{p}{e^2} (1 - e^{-1}) \cdot \frac{1}{p} \left(\lfloor \frac{1}{p} \rfloor - 1 \geq (1 - O(p_0)) \frac{1}{p} \right) \\
&\geq (p|C| - |M^+| - O(p_0)_{\text{OPT}}) \cdot (1 - O(p_0)) \cdot \eta \\
&\geq (p|C| - |M^+|) \cdot \eta - O(p_0)_{\text{OPT}} && (|M^+| = O(\text{OPT}), p|C| \leq p|B_M| = O(\text{OPT}))
\end{aligned}$$

which proves Lemma 5.4. We now prove Lemma 5.5 which will complete the proof.

PROOF OF LEMMA 5.5. Fix any allocation d_u^* and d_v^* that minimizes $F(d_u, d_v)$. We will show that first, there exists a sequence of locally reallocating the values (i.e., degrees) in d_u^* without changing the value of $F(d_u, d_v)$ such that at the end, every entry in d_u^* is either 1 or $\lfloor \frac{1}{p} \rfloor$. After changing the vector d_u^* , we then show that there exists a sequence of locally reallocating the values in d_v^* without changing the value of $F(d_u, d_v)$ such that at the end, every entry in d_v^* is either 0 or $\lfloor \frac{1}{p} \rfloor$.

We first explain how to change d_u^* . To simplify the presentation, we define $q_i = 1 - e^{-p \cdot d^*(v_i)}$ and the target expression becomes

$$\sum_i \left(1 - e^{-p \cdot d^*(v_i)} \right) \cdot \max \{ d^*(u_i) - 1, 0 \} = \sum_i q_i \cdot \max \{ d^*(u_i) - 1, 0 \} \quad (5)$$

Recall that $d^*(u_i)$ satisfies $d^*(u_i) \in \left[\lfloor \frac{1}{p} \rfloor \right]$ (in MP-(4)) (and hence $q_i \geq 0$) and $\sum_i d^*(u_i) = d(U)$. First of all, if there exists some i_1 where $d^*(u_{i_1}) = 0$, then since $d(U) \geq |M^+|$, there must exist an index i_2 where $d^*(u_{i_2}) \geq 2$. Then, we can shift one degree from $d^*(u_{i_2})$ to $d^*(u_{i_1})$ and after the shift, (a) $\max \{ d^*(u_{i_1}) - 1, 0 \}$ remains 0 and hence $q_{i_1} \max \{ d^*(u_i) - 1, 0 \}$ remains 0, and (b) $\max \{ d^*(u_{i_2}) - 1, 0 \}$ decreases and hence $q_{i_2} \max \{ d^*(u_{i_2}) - 1, 0 \}$ does not increase. Therefore, $F(d_u, d_v)$ does not increase after the shift, and from now on, we have $d^*(u_i) \geq 1$ for all $i \in [|M^+|]$. To proceed, we need the following property of $d^*(u_i)$.

CLAIM 5.6. For any pair of indices i_1, i_2 where $q_{i_1} > q_{i_2}$, either $d^*(u_{i_2}) = \lfloor \frac{1}{p} \rfloor$ or $d^*(u_{i_1}) \leq 1$.

PROOF. Suppose not. We have $d^\star(u_{i_2}) < \lfloor \frac{1}{p} \rfloor$ and $d^\star(u_{i_1}) > 1$, for some i_1 and i_2 . We can shift one degree from $d^\star(u_{i_1})$ to $d^\star(u_{i_2})$ and still get a valid allocation. In the following, we show that this new allocation achieves a smaller value of F , which contradicts to the optimality of $d^\star(u_i)$.

Since shifting from $d^\star(u_{i_1})$ to $d^\star(u_{i_2})$ only changes the degrees for u_{i_1} and u_{i_2} , it suffices for us to prove that

$$\begin{aligned} \Delta := & (q_{i_2} \max \{d^\star(u_{i_2}) - 1, 0\} + q_{i_1} \max \{d^\star(u_{i_1}) - 1, 0\}) \\ & - (q_{i_2} \max \{d^\star(u_{i_2}), 0\} + q_{i_1} \max \{d^\star(u_{i_1}) - 2, 0\}) > 0 \end{aligned}$$

Since $d^\star(u_{i_1}) \geq 2$, $\max \{d^\star(u_{i_1}) - 1, 0\} = d^\star(u_{i_1}) - 1$, $\max \{d^\star(u_{i_1}) - 2, 0\} = d^\star(u_{i_1}) - 2$. In addition, since $d^\star(u_{i_2}) \geq 0$, $\max \{d^\star(u_{i_2}), 0\} = d^\star(u_{i_2})$. We have

$$\begin{aligned} \Delta = & (q_{i_2} \max \{d^\star(u_{i_2}) - 1, 0\} + q_{i_1} (d^\star(u_{i_1}) - 1)) - (q_{i_2} d^\star(u_{i_2}) + q_{i_1} (d^\star(u_{i_1}) - 2)) \\ = & q_{i_2} (\max \{d^\star(u_{i_2}) - 1, 0\} - d^\star(u_{i_2})) + q_{i_1} \\ \geq & q_{i_2} (d^\star(u_{i_2}) - 1 - d^\star(u_{i_2})) + q_{i_1} \\ = & q_{i_1} - q_{i_2}. \end{aligned}$$

Since we have $q_{i_1} > q_{i_2}$, the value of F decreases after the shifting according to Eq 5, a contradiction. \square

We use Claim 5.6 to prove the correctness of the following sequence of reallocation of d_u^\star . Now, as long as there exists an index i_1 , where $d^\star(u_{i_1}) \in (1, \lfloor \frac{1}{p} \rfloor]$, since $d(U) - |M^+|$ is an integer multiple of $(\lfloor \frac{1}{p} \rfloor - 1)$, there must exist some index i_2 where $d^\star(u_{i_2}) \in (1, \lfloor \frac{1}{p} \rfloor]$ (recall that $d^\star(u_i) \geq 1$), and we will shift the values between $d^\star(u_{i_1})$ and $d^\star(u_{i_2})$ such that one of them becomes either 1 or $\lfloor \frac{1}{p} \rfloor$ and both of them are still at least 1 (it is easy to see this is always possible). First of all, every step of this reallocation reduces the number of vertices with $d^\star(u_i) \in (1, \lfloor \frac{1}{p} \rfloor]$, and hence it will terminate. To see that this process never changes $F(d_u, d_v)$, (a) it cannot be that $q_{i_1} \neq q_{i_2}$, since otherwise the indices i_1 and i_2 will contradict Claim 5.6, and (b) if $q_{i_1} = q_{i_2}$, shifting the allocation between i_1 to i_2 will not change $F(d_u, d_v)$.

Therefore, we can focus on the case where the entries of $d^\star(u_i)$ are either 1 or $\lfloor \frac{1}{p} \rfloor$. We now consider $d^\star(v_i)$. Recall that $d(V) = \sum_i d(v_i)$ and $d(V)$ is an integer multiple of $\lfloor \frac{1}{p} \rfloor$. The target expression can be written as

$$\begin{aligned} & \sum_i (1 - e^{-p \cdot d^\star(v_i)}) \cdot \max \{d^\star(u_i) - 1, 0\} \\ = & \sum_{i: d^\star(u_i) = \lfloor \frac{1}{p} \rfloor} (1 - e^{-p \cdot d^\star(v_i)}) \left(\left\lfloor \frac{1}{p} \right\rfloor - 1 \right) + \sum_{i: d^\star(u_i) = 1} (1 - e^{-p \cdot d^\star(v_i)}) \cdot 0 \\ = & \sum_{i: d^\star(u_i) = \lfloor \frac{1}{p} \rfloor} (1 - e^{-p \cdot d^\star(v_i)}) \end{aligned}$$

Since F is monotonically increasing when any $d(u_i)$ increases, for the indices i where $d^\star(u_i) = 1$, ideally, one should allocate as many degrees to $d^\star(v_i)$ as possible, i.e., $d^\star(v_i) = \lfloor \frac{1}{p} \rfloor$. However, it might be the case that $d(V)$ cannot supply $\lfloor \frac{1}{p} \rfloor$ degree for all i where $d^\star(u_i) = 1$. But in this case, we are done since reallocating between different $d^\star(v_i)$ where $d^\star(u_i) = 1$ does not change the F

(in fact, F is always 0), and we can shift them such that we have as many $d^*(v_i) = \lfloor \frac{1}{p} \rfloor$ as possible and leave the rest equal to 0.

In the following, we assume $d(V)$ can supply $\lfloor \frac{1}{p} \rfloor$ degree for all i with $d^*(u_i) = 1$, and hence $d^*(v_i) = \lfloor \frac{1}{p} \rfloor$ whenever $d^*(u_i) = 1$. It suffices to only focus on $d^*(v_i)$ where $d^*(u_i) = \lfloor \frac{1}{p} \rfloor$. We need the following property of $d^*(v_i)$ to complete the argument.

CLAIM 5.7. *For any pair of indices i_1 and i_2 such that $d^*(u_{i_1}) = d^*(u_{i_2}) = \lfloor \frac{1}{p} \rfloor$, we have that either $\min \{d^*(v_{i_1}), d^*(v_{i_2})\} = 0$ or $\max \{d^*(v_{i_1}), d^*(v_{i_2})\} = \lfloor \frac{1}{p} \rfloor$.*

PROOF. Suppose not. Then, for some i_1 and i_2 , we have $\min \{d^*(v_{i_1}), d^*(v_{i_2})\} > 0$, and also $\max \{d^*(v_{i_1}), d^*(v_{i_2})\} < \lfloor \frac{1}{p} \rfloor$. Without loss of generality, assume $1 \leq d^*(v_{i_1}) \leq d^*(v_{i_2}) \leq \lfloor \frac{1}{p} \rfloor - 1$. Then, shifting one degree from $d^*(v_{i_1})$ to $d^*(v_{i_2})$ leads to a valid allocation, and we prove in the following that the new allocation decreases the objective function which contradicts the optimality of d_v^* .

Since only the indices i_1 and i_2 are affected, it suffices for us to prove that

$$\Delta := \left(1 - e^{-p \cdot d^*(v_{i_1})}\right) + \left(1 - e^{-p \cdot d^*(v_{i_2})}\right) - \left(1 - e^{-p \cdot (d^*(v_{i_1})-1)}\right) + \left(1 - e^{-p \cdot (d^*(v_{i_2})+1)}\right) > 0$$

We have

$$\Delta \geq e^{-p \cdot d^*(v_{i_1})}(e^p - 1) + e^{-p \cdot d^*(v_{i_2})}(e^{-p} - 1)$$

Since $d^*(v_{i_1}) \leq d^*(v_{i_2})$, $e^{-p \cdot d^*(v_{i_1})} \geq e^{-p \cdot d^*(v_{i_2})}$. We further have

$$\begin{aligned} e^{-p \cdot d^*(v_{i_1})}(e^p - 1) + e^{-p \cdot d^*(v_{i_2})}(e^{-p} - 1) &\geq e^{-p \cdot d^*(v_{i_2})}(e^p - 1) + e^{-p \cdot d^*(v_{i_2})}(e^{-p} - 1) \\ &\geq e^{-p \cdot d^*(v_{i_2})}(e^p - 1 + e^{-p} - 1) \\ &> e^{-p \cdot d^*(v_{i_2})}(2\sqrt{e^p \cdot e^{-p}} - 2) = 0 \end{aligned}$$

where the strict inequality is true since the two terms can only be equal when $e^p = e^{-p}$ which does not happen for $p > 0$. \square

Using Claim 5.7, we can now show that any $d^*(v_i)$ where $d^*(u_i) = \lfloor \frac{1}{p} \rfloor$, must be either 0 or $\lfloor \frac{1}{p} \rfloor$. Suppose not. If $d^*(v_{i_1}) \in (0, \lfloor \frac{1}{p} \rfloor)$, then since $d(V)$ is an integer multiple of $\lfloor \frac{1}{p} \rfloor$, there must exist some other index i_2 where $d^*(v_{i_2}) \in (0, \lfloor \frac{1}{p} \rfloor)$; hence, $0 < \min \{d^*(v_{i_1}), d^*(v_{i_2})\} < \max \{d^*(v_{i_1}), d^*(v_{i_2})\} < \lfloor \frac{1}{p} \rfloor$, a contradiction to Claim 5.7. \square

6 AN ALGORITHM FOR LARGE VALUES OF P

In this section, we present Algorithm 2 with approximation ratio $(1/2 + \Theta(p^2))$. Algorithm 2 is required to handle the case when p is not small enough for Algorithm 1 to perform well.

Let p_0 be any fixed constant independent of n , $\delta = \frac{p^2}{4}$, and $\varepsilon = \frac{p_0^2}{10^4}$. The new algorithm (i.e., Algorithm 2) is similar to Algorithm 1 with the only difference being that instead of a $\lfloor \frac{1}{p} \rfloor$ -matching, here, we simply pick a single maximum matching in G . Our algorithm is stated as Algorithm 2.

The following lemma proves the approximation ratio of Algorithm 2. The proof is a simpler version of the proof for Algorithm 1 and is deferred to the full version of the paper [7].

LEMMA 6.1. *For any constant $p_0 > 0$, any probability $p \geq p_0$, and any graph $G(V, E)$ the expected maximum matching size in the graph H computed by Algorithm 2 is at least $\left(\frac{1}{2} + \frac{p^2}{4} - \frac{p_0^2}{10^4}\right) \cdot \text{OPT}(G)$.*

ALGORITHM 2: A $(0.5 + \Theta(p^2))$ -Approximation Algorithm for Stochastic Matching

Input: A graph $G(V, E)$ and an edge realization probability $p_0 \leq p < 1$.

Output: A subgraph $H(V, Q)$ of $G(V, E)$.

- (1) Let M be a maximum matching in G .
 - (2) Let $(M_1, M_2, \dots, M_R) := \text{MatchingCover}(G(V, E \setminus M), \varepsilon)$
(recall that $\varepsilon = \frac{p_0^2}{10^4}$), and $E_{MC} = M_1 \cup \dots \cup M_R$.
 - (3) Return $H(V, Q)$ where $Q := M \cup E_{MC}$.
-

We now show how to combine Algorithm 1 and Algorithm 2 to prove Part (ii) of Theorem 1.1.

PROOF OF THEOREM 1.1, PART (ii). Let p_0 be the constant such that Algorithm 1 achieves an approximation ratio of 0.52 for any $p \leq p_0$. The algorithm for Part (ii) is simply as follows. If the realization probability $p \leq p_0$, run Algorithm 1 and otherwise run Algorithm 2. By Lemma 6.1, the approximation ratio of this algorithm is $\min \left\{ 0.52, \frac{1}{2} + \frac{p^2}{4} - \frac{p_0^2}{10^4} \right\} = 0.5 + \delta_0$ for some absolute constant δ_0 (since p_0 is an absolute constant and $p \geq p_0$).

□

7 CONCLUDING REMARKS AND OPEN PROBLEMS

We presented the first non-adaptive algorithm for stochastic matching with an approximation ratio that is strictly better than half. In particular, we showed that any graph G has a subgraph H with maximum degree $O(\frac{\log(1/p)}{p})$ such that the ratio of expected size of a maximum matching in realizations of H and G is at least 0.52 when p is sufficiently small, i.e., case of vanishing probabilities, and $0.5 + \delta_0$ (for an absolute constant $\delta_0 > 0$) for any $p \in (0, 1)$.

A main open problem is to determine the best approximation ratio achievable by a non-adaptive algorithm. In particular, can non-adaptive algorithms qualitatively match the performance of adaptive algorithms by achieving a $(1 - \varepsilon)$ -approximation for any $\varepsilon > 0$ using a subgraph with maximum degree $f(\varepsilon, p)$ for some function f ? In the full version of the paper [7], we mention some potential directions towards resolving this problem.

ACKNOWLEDGMENTS

We thank the anonymous reviewers of EC 2017 for many insightful comments and suggestions. This work was supported in part by National Science Foundation grants CCF-1552909, CCF-1617851, and IIS-1447470.

REFERENCES

- [1] Marek Adamczyk. 2011. Improved analysis of the greedy algorithm for stochastic matching. *Inf. Process. Lett.* 111, 15 (2011), 731–737.
- [2] Mohammad Akbarpour, Shengwu Li, and Shayan Oveis Gharan. 2014. Dynamic matching market design. In *ACM Conference on Economics and Computation, EC '14, Stanford, CA, USA, June 8–12, 2014*. 355.
- [3] Ross Anderson, Itai Ashlagi, David Gamarnik, and Yash Kanoria. 2015. A dynamic model of barter exchange. In *Proceedings of the Twenty-Sixth Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2015, San Diego, CA, USA, January 4–6, 2015*. 1925–1933.
- [4] Ross Anderson, Itai Ashlagi, David Gamarnik, and Alvin E. Roth. 2015. Finding long chains in kidney exchange using the traveling salesman problem. *Proceedings of the National Academy of Sciences* 112, 3 (2015), 663–668.
- [5] Jonathan Aronson, Martin E. Dyer, Alan M. Frieze, and Stephen Suen. 1995. Randomized Greedy Matching II. *Random Struct. Algorithms* 6, 1 (1995), 55–74.

- [6] Sepehr Assadi, Sanjeev Khanna, and Yang Li. 2016. The Stochastic Matching Problem with (Very) Few Queries. *Proceedings of the 2016 ACM Conference on Economics and Computation, EC '16, Maastricht, The Netherlands, July 24-28, 2016* (2016), 43–60.
- [7] Sepehr Assadi, Sanjeev Khanna, and Yang Li. 2017. The Stochastic Matching Problem: Beating Half with a Non-Adaptive Algorithm. *CoRR* abs/1705.02280 (2017).
- [8] Pranjal Awasthi and Tuomas Sandholm. 2009. Online Stochastic Optimization in the Large: Application to Kidney Exchange. In *IJCAI 2009, Proceedings of the 21st International Joint Conference on Artificial Intelligence, 2009*. 405–411.
- [9] Nikhil Bansal, Anupam Gupta, Jian Li, Julián Mestre, Viswanath Nagarajan, and Atri Rudra. 2012. When LP Is the Cure for Your Matching Woes: Improved Bounds for Stochastic Matchings. *Algorithmica* 63, 4 (2012), 733–762.
- [10] Avrim Blum, John P. Dickerson, Nika Haghtalab, Ariel D. Procaccia, Tuomas Sandholm, and Ankit Sharma. 2015. Ignorance is Almost Bliss: Near-Optimal Stochastic Matching With Few Queries. In *Proceedings of the Sixteenth ACM Conference on Economics and Computation, EC '15, Portland, OR, USA, June 15-19, 2015*. 325–342.
- [11] Avrim Blum, Anupam Gupta, Ariel D. Procaccia, and Ankit Sharma. 2013. Harnessing the power of two crossmatches. In *ACM Conference on Electronic Commerce, EC '13, Philadelphia, PA, USA, June 16-20, 2013*. 123–140.
- [12] Béla Bollobás. 2001. *Random Graphs*. Number 73. Cambridge University Press.
- [13] T.-H. Hubert Chan, Fei Chen, Xiaowei Wu, and Zhichao Zhao. 2014. Ranking on Arbitrary Graphs: Rematch via Continuous LP with Monotone and Boundary Condition Constraints. In *Proceedings of the Twenty-Fifth Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2014, Portland, Oregon, USA, January 5-7, 2014*. 1112–1122.
- [14] Ning Chen, Nicole Immorlica, Anna R. Karlin, Mohammad Mahdian, and Atri Rudra. 2009. Approximating Matches Made in Heaven. In *Automata, Languages and Programming, 36th International Colloquium, ICALP 2009, Rhodes, Greece, July 5-12, 2009, Proceedings, Part I*. 266–278.
- [15] Kevin P. Costello, Prasad Tetali, and Pushkar Tripathi. 2012. Stochastic Matching with Commitment. In *Automata, Languages, and Programming - 39th International Colloquium, ICALP 2012, Proceedings, Part I*. 822–833.
- [16] John P. Dickerson, Ariel D. Procaccia, and Tuomas Sandholm. 2012. Dynamic Matching via Weighted Myopia with Application to Kidney Exchange. In *Proceedings of the Twenty-Sixth AAAI Conference on Artificial Intelligence*.
- [17] John P. Dickerson, Ariel D. Procaccia, and Tuomas Sandholm. 2013. Failure-aware kidney exchange. In *ACM Conference on Electronic Commerce, EC '13*. 323–340.
- [18] John P. Dickerson and Tuomas Sandholm. 2015. FutureMatch: Combining Human Value Judgments and Machine Learning to Match in Dynamic Environments. In *Proceedings of the Twenty-Ninth AAAI Conference on Artificial Intelligence, January 25-30, 2015, Austin, Texas, USA*. 622–628.
- [19] Martin E. Dyer and Alan M. Frieze. 1991. Randomized Greedy Matching. *Random Struct. Algorithms* 2, 1 (1991), 29–46.
- [20] Joan Feigenbaum, Sampath Kannan, Andrew McGregor, Siddharth Suri, and Jian Zhang. 2005. On graph problems in a semi-streaming model. *Theor. Comput. Sci.* 348, 2-3 (2005), 207–216. DOI: <http://dx.doi.org/10.1016/j.tcs.2005.09.013>
- [21] Anupam Gupta and Viswanath Nagarajan. 2013. A Stochastic Probing Problem with Applications. In *Integer Programming and Combinatorial Optimization - 16th International Conference, IPCO 2013, Proceedings*. 205–216.
- [22] Richard M. Karp, Umesh V. Vazirani, and Vijay V. Vazirani. 1990. An Optimal Algorithm for On-line Bipartite Matching. In *Proceedings of the 22nd Annual ACM Symposium on Theory of Computing, May 13-17, 1990, Baltimore, Maryland, USA*. 352–358. DOI: <http://dx.doi.org/10.1145/100216.100262>
- [23] Christian Konrad, Frédéric Magniez, and Claire Mathieu. 2012. Maximum Matching in Semi-streaming with Few Passes. In *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques - 15th International Workshop, APPROX 2012, and 16th International Workshop, RANDOM 2012, Cambridge, MA, USA, August 15-17, 2012, Proceedings*. 231–242.
- [24] L. Lovász and D. Plummer. 2009. *Matching Theory*. American Mathematical Soc. <https://books.google.com/books?id=yW3WSVq8ygcC>
- [25] David F. Manlove and Gregg O'Malley. 2014. Paired and Altruistic Kidney Donation in the UK: Algorithms and Experimentation. *ACM Journal of Experimental Algorithmics* 19, 1 (2014).
- [26] Aranyak Mehta and Debmalaya Panigrahi. 2012. Online Matching with Stochastic Rewards. In *53rd Annual IEEE Symposium on Foundations of Computer Science, FOCS 2012, New Brunswick, NJ, USA, October 20-23, 2012*. 728–737.
- [27] Aranyak Mehta, Bo Waggoner, and Morteza Zadimoghaddam. 2015. Online Stochastic Matching with Unequal Probabilities. In *Proceedings of the Twenty-Sixth Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2015, San Diego, CA, USA, January 4-6, 2015*. 1388–1404.
- [28] Matthias Poloczek and Mario Szegedy. 2012. Randomized Greedy Algorithms for the Maximum Matching Problem with New Analysis. In *53rd Annual IEEE Symposium on Foundations of Computer Science, FOCS 2012, New Brunswick, NJ, USA, October 20-23, 2012*. 708–717.
- [29] Alexander Schrijver. 2003. *Combinatorial optimization: polyhedra and efficiency*. Springer Science & Business Media.
- [30] Utku Unver. 2010. Dynamic Kidney Exchange. *Review of Economic Studies* 77, 1 (2010), 372–414.