

# Opting Into Optimal Matchings\*

Avrim Blum<sup>1</sup>, Ioannis Caragiannis<sup>2</sup>, Nika Haghtalab<sup>1</sup>,  
Ariel D. Procaccia<sup>1</sup>, Eviatar B. Procaccia<sup>3</sup>, and Rohit Vaish<sup>4</sup>

<sup>1</sup>Computer Science Department, Carnegie Mellon University, {avrim,nhaghtal,arielpro}@cs.cmu.edu

<sup>2</sup>Department of Computer Engineering & Informatics, University of Patras,  
caragian@ceid.upatras.gr

<sup>3</sup>Department of Mathematics, Texas A&M University, procaccia@math.tamu.edu

<sup>4</sup>Department of Computer Science and Automation, Indian Institute of Science,  
rohit.vaish@csa.iisc.ernet.in

## Abstract

We revisit the problem of designing optimal, *individually rational* matching mechanisms (in a general sense, allowing for cycles in directed graphs), where each player—who is associated with a subset of vertices—matches as many of his own vertices when he opts into the matching mechanism as when he opts out. We offer a new perspective on this problem by considering an arbitrary graph, but assuming that vertices are associated with players at random. Our main result asserts that, under certain conditions, *any* fixed optimal matching is likely to be individually rational up to lower-order terms. We also show that a simple and practical mechanism is (fully) individually rational, and likely to be optimal up to lower-order terms. We discuss the implications of our results for market design in general, and kidney exchange in particular.

## 1 Introduction

Matching theory has made an astounding real-world impact, through the field of *market design*; it is the cornerstone of the design and analysis of widely deployed applications that match residents to hospitals [25], students to schools [1], and organ donors to patients [26, 27, 28]. But as matching markets become more prevalent, new issues arise, which potentially limit their (economic) efficiency. In this paper, we tackle one such is-

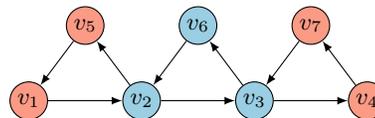


Figure 1: A compatibility graph where individual rationality fails.

sue: *individual rationality* (or the lack thereof). Specifically, we study situations where the vertices of the graph are partitioned between a set of players, and each player is interested in matching as many of *his own* vertices as possible. An *individually rational matching* is one that matches at least as many vertices of each player as he can match on his own.

Why is individual rationality a real issue? Of the examples listed earlier, *kidney exchange* provides arguably the most concrete, compelling answer. It is a medical innovation that, in its basic form, allows patients who need kidney transplant, and have willing but medically incompatible donors, to swap donors. From the matching viewpoint, the kidney exchange setting can be represented via a directed *compatibility graph*, where each vertex corresponds to an incompatible *patient-donor pair*, and there is an edge  $(u, v)$  if the donor of  $u$  is medically compatible with the patient of  $v$ . A pairwise swap corresponds to a 2-cycle in this graph, but exchanges along longer cycles—and even along chains, initiated by altruistic donors—are also important in practice (we also use the term *matching* to refer to cycles and chains in these directed graphs).

Based on their work with practitioners, Ashlagi and Roth [7] have recently raised serious concerns regarding individual rationality in kidney exchange; they convincingly argue that as kidney exchange programs outgrow

\*This work was partially supported by the NSF under grants DMS-1407558, IIS-1350598, CCF-1215883, CCF-1525932, CCF-1331175, and CCF-1525971; and by Caratheodory grant E.114 from the University of Patras, COST Action CA15210 ENCKEP, Sloan Research Fellowship, IBM Ph.D. Fellowship, and MSR Ph.D. Fellowship. Part of this work was done when RV was visiting Carnegie Mellon University as part of a student visit under the Indo-US Joint Center for Advanced Research in Machine Learning, Game Theory and Optimization supported by the Indo-US Science and Technology Forum.

their regional origins, the incentives of hospitals (the players in this case)—which have little to no interaction outside of the kidney exchange program—become misaligned. In particular, hospitals cannot be certain that if they opt into a kidney exchange program, which optimizes overall efficiency, their patients would be better off overall than under the optimal *internal* matching (which relies only on donor-patient pairs associated with the hospital). A bad example (due to Ashlagi and Roth) is given in Figure 1: the maximum cardinality matching selects the 3-cycles  $\{v_1, v_2, v_5\}$  and  $\{v_3, v_4, v_7\}$ , but the blue player can do better by internally matching the single 3-cycle  $\{v_2, v_3, v_6\}$ . That is, the maximum cardinality matching is twice as large as the unique individually rational matching.

**1.1 Our Approach** To summarize the preceding discussion, individual rationality is potentially a major obstacle to the economic efficiency of matching markets. Our goal is to analytically demonstrate that, in fact, individual rationality (or an almost perfect approximation thereof) can be achieved with nearly no loss of efficiency. Our key insight is that it suffices to assume that each vertex of the graph is owned by a *random player*.

In more detail, we consider an arbitrary graph with  $n$  vertices  $v_1, \dots, v_n$ , and a set of  $k$  players  $1, \dots, k$ . For each vertex, we draw its owner independently from the probability distribution  $p_1, \dots, p_k$  over the players, that is, each vertex is assigned to player  $i$  with probability  $p_i$ . In the kidney exchange setting, for example, the rationale is very simple: the graph represents medical compatibility information, and there is no special reason why a patient-donor pair with particular medical characteristics would belong to a particular hospital—the probability of that happening depends chiefly on the size of the hospital.

To see how randomization helps, let us revisit the example given in Figure 1, and suppose that the two players (red and blue) have probability  $1/2$  each:  $p_1 = p_2 = 1/2$ . The expected utility of a player under the maximum cardinality matching is  $6/2 = 3$ . In contrast, a straightforward upper bound on the expected cardinality of an internal matching can be derived by observing that each of the three 3-cycles is owned by a single player with probability  $1/8$  and adds at most 3 to the cardinality of the matching, leading to an upper bound of  $3 \cdot 3 \cdot (1/8) = 9/8$ . Now, suppose we made  $t$  copies of the graph of Figure 1, for a large  $t$ ; then a simple measure concentration argument would imply that it is very likely that each player is better off in the optimal solution than he is working alone.

Our goal is to establish this phenomenon in some generality. Indeed, our qualitative message (a few

technical caveats apply) is that

*... in an arbitrary graph, under a random assignment of vertices to players, it is likely that any fixed optimal matching is individually rational, up to lower order terms, for each player; and there is a practical mechanism that yields an individually rational matching that is likely to be optimal up to lower order terms.*

**1.2 Our Results and Techniques** In §2, we formalize the first part of the above statement. Specifically, we prove the following theorem:

**THEOREM 2.1 (INFORMALLY STATED).** *Let  $G$  be a directed graph, and let  $\text{opt}(G)$  be the set of vertices matched under a specific maximum cardinality matching on  $G$ . Assume that one of the following conditions holds:*

1. *Matchings are restricted to 2-cycles, and  $p_i \leq 1/2$  for each player  $i$ , or*
2. *Matchings are restricted to cycles of constant length, and for each player  $i$ ,  $1/p_i$  is an integer.*

*Then with probability at least  $1 - \delta$ , for each player  $i \in [k]$ , the difference between the size of his optimal internal matching, and his share of  $\text{opt}(G)$ , is at most  $O(\sqrt{|\text{opt}(G)|} \cdot \ln(k/\delta))$ .*

The theorem’s first case deals with 2-cycles, a common abstraction for kidney exchange in theoretical studies [27, 29, 4, 5, 15, 11, 10, 8]. Of course in this case there always *exists* an optimal and individually rational matching (find the optimal internal matchings and then add augmenting paths), but nonetheless this statement is appealing because it applies to *any* optimal solution that the exchange—which might also be optimizing some secondary objective—might produce. Also note that this case is essentially unrestrictive in terms of the probability distribution. The second case is the opposite: its assumption of constant length cycles is essentially unrestrictive, as chains can be represented as cycles by adding an edge from every patient-donor pair to every altruistic donor; and major kidney exchanges—such as the US national program, run by the United Network for Organ Sharing (UNOS)—use only cycles of length at most 3, and chains of length at most 4 [28, 7, 18, 6]. But the assumption regarding the probability distribution is, of course, somewhat restrictive. Note, however, that probabilities can be “rounded” at a cost, as we discuss later; and that the natural case of equal probabilities is captured by the second case.

The proof of Theorem 2.1 relies on two main ingredients. The first is the claim that the expected

size of the maximum internal matching of player  $i$  is at most a  $p_i$  fraction of the optimal (global) matching. This statement is almost trivial in Case 2; to establish it in Case 1, we decompose the maximum cardinality matching via the Edmonds-Gallai Decomposition [20], and show that the inequality holds for each component separately.

The second ingredient is the concentration of the cardinality of the optimal internal matching of each player around its expectation. To this end, we leverage machinery from modern probability theory including a concentration inequality for so-called *self-bounding functions* [14].

The power of Theorem 2.1 is that it applies to *any* maximum cardinality matching. In the context of kidney exchange, the theorem captures the matching algorithms currently in use (including the ones employed by UNOS); its conceptual message is that hospitals need not worry about opting into kidney exchange programs, even under the *status quo*.

By contrast, in §3 we give the designer more power in choosing the matching, with the goal of constructing a mechanism that is (perfectly) individually rational, and almost optimal. As noted earlier, this is quite trivial when only 2-cycles are allowed, as there always exists an optimal, individually rational matching. When longer cycles are allowed, we can derive the following corollary from the proof of Theorem 2.1.

**COROLLARY 3.1 (INFORMALLY STATED).** *Let  $G$  be a directed graph with  $n$  vertices, and let  $\text{opt}(G)$  be the set of vertices matched under a specific maximum cardinality matching on  $G$ . Suppose that matchings are restricted to cycles of constant length, and  $p_i = p_j$  for any two players  $i, j$ . Then, with probability  $1 - \delta$ , there exists a matching that is individually rational for each player  $i$ , and has maximum cardinality up to  $O(k\sqrt{|\text{opt}(G)|} \cdot \ln(k/\delta))$ .*

Importantly, such an individually rational and almost optimal matching can be found with a *practical*<sup>1</sup> mechanism: (i) compute a maximum cardinality matching, (ii) any player who wishes to work alone is allowed to defect.

Furthermore, we show that our results are tight. Among other things, we construct an example (Example 3.1) with two players and cycles up to length 3 such that, with constant probability, *any* individually rational matching is smaller than the optimal matching by

$$\Omega(\sqrt{|\text{opt}(G)|}).$$

**1.3 Related Work** The two papers that are most closely related to ours are the ones by Ashlagi and Roth [7] and Toulis and Parkes [29]. Ashlagi and Roth show that under some technical assumptions, and under a random graph model of kidney exchange, large random graphs admit an individually rational matching that is optimal up to a certain constant fraction of the number of vertices, with high probability. Toulis and Parkes [29] independently study a very similar random graph model (it does make different assumptions about the size of hospitals), and obtain a similar result regarding individual rationality.

These important results have inspired our own work, but—in addition to a number of significant technical advantages<sup>2</sup>—we believe our high-level approach is significantly more compelling. In a nutshell, the random graph model studied by Ashlagi and Roth [7] and Toulis and Parkes [29] draws blood types for each donor and patient from a distribution that gives each of the four blood types (O, A, B, and AB) constant probability. For each pair of blood type compatible vertices (e.g., an O donor is blood type compatible with an A patient, but a B donor is not), a directed edge exists with *constant* probability. This model clearly gives rise to very dense graphs; the key to the abovementioned results is that, with high probability, there exist matchings between blood type compatible groups (such as A patient and B donor, and B patient and A donor) that are perfect in the sense that they match all the vertices in the smaller group. Consequently, the structure of the optimal matching can be accurately predicted with high probability. This model has subsequently been employed in several other papers [18, 11].

However, more recent work by Ashlagi and Roth themselves—together with collaborators [6]—introduces a completely different random graph model of kidney exchange, which gives rise to sparse graphs, and better captures some real-world phenomena. This model was later employed by Dickerson et al. [19]. At this point it is fair to say that, on the question of whether random graph models are a valid approach for the analysis of kidney exchange, the jury is still out. But we are convinced that an analysis that holds for *arbitrary* graphs—when it is feasible, as in this paper—is the right approach.

Individual rationality limits players to two possible

<sup>1</sup>By “practical” we mean that it can be easily implemented in practice. It is not a polynomial-time algorithm, as computing a maximum cardinality matching is  $\mathcal{NP}$ -hard when 3-cycles are allowed [2]; but the problem is routinely solved via integer programming.

<sup>2</sup>In contrast to their work, we obtain optimality up to lower-order terms (instead of up to a constant fraction of  $n$ ), and our results have a good dependence on the number of players (instead of assuming a very large [7] or a very small [29] number) and on the size of the graph (instead of assuming that  $n$  goes to infinity).

strategies: work alone or participate fully. More generally, players can choose to reveal a subset of their vertices, and internally match the rest. Several papers seek to design mechanisms that incentivize players to reveal all their vertices, either as a dominant strategy [5, 23] or in equilibrium [7, 29]. These known results are quite limited; obtaining stronger results is a central open problem. Our own approach does not seem to extend beyond individual rationality.

Needless to say, matching is a major research topic in theoretical computer science. In particular, there are many papers that are directly motivated by market design applications, especially kidney exchange [17, 16, 3, 21, 9, 22]. These papers are largely orthogonal to our work.

## 2 Optimal Matchings Are Almost Individually Rational

Designing and implementing new matching mechanisms can require significant changes to current policies and deployed algorithms. In this section, we show that even without any changes to the existing (optimal) matching mechanisms—at least in the case of kidney exchange—it is likely that each player matches almost as many vertices as what he could have obtained on his own.

Consider an arbitrary directed graph  $G$  with  $n$  vertices. Recall that for each player  $i \in [k]$  with corresponding probability  $p_i$ , the player owns each vertex with probability  $p_i$ , independently. We denote by  $H_i \sim_{p_i} G$  the random subgraph of player  $i$ , which is a subgraph of  $G$  induced by assigning each vertex to  $i$  with probability  $p_i$ . We suppress  $p_i$  from this notation when it is clear from the context. We use  $\text{opt}(G)$  to denote the set of *vertices* of an arbitrary but fixed matching of  $G$ . Furthermore,  $\text{opt}(G) \upharpoonright H_i$  denotes the restriction of  $\text{opt}(G)$  to subgraph  $H_i$ , i.e., the *vertices* of  $H_i$  that are matched under  $\text{opt}(G)$ . Therefore, to compare the size of the internal matching of  $H_i$  with the number of vertices of  $H_i$  that are matched under the global matching, we compare  $|\text{opt}(H_i)|$  to  $|\text{opt}(G) \upharpoonright H_i|$ , and show that these values are within  $\tilde{O}(\sqrt{|\text{opt}(G)|})$  of one another.

Let us first describe a graph in which, with a constant probability, a player’s internal matching is larger by  $\Omega(\sqrt{|\text{opt}(G)|})$  than the player’s share of any fixed optimal matching.

**EXAMPLE 2.1.** Suppose one of the players has probability  $p = \frac{1}{2}$ , and consider a graph that consists of  $n/\log(n)$  stars, each with  $\log(n)$  vertices that are connected to the center via 2-cycles. Fix an optimal global matching,  $\text{opt}(G)$ , and note that  $|\text{opt}(G)| = 2n/\log(n)$ . We

informally argue that there is a constant  $c > 0$  such that

$$\Pr_{H \sim_p G} \left[ |\text{opt}(H)| - |\text{opt}(G) \upharpoonright H| > \Omega(\sqrt{|\text{opt}(G)|}) \right] > c.$$

Indeed, let us consider the subgraph internal to the player,  $H \sim_p G$ . While the expected number of centers in  $H$  is  $\frac{1}{4}|\text{opt}(G)|$ , it is easy to see (by looking up the standard deviation of the binomial distribution) that, with constant probability,  $H$  includes  $t = \frac{1}{4}|\text{opt}(G)| + \Theta(\sqrt{|\text{opt}(G)|})$  centers. Moreover, with probability  $\frac{1}{2}$ ,  $H$  includes no more than half of the non-center vertices matched by  $\text{opt}(G)$ . If both events occur,  $|\text{opt}(G) \upharpoonright H| \leq t + \frac{1}{4}|\text{opt}(G)|$ , where  $t$  of the matched vertices correspond to the center vertices and at most  $\frac{1}{4}|\text{opt}(G)|$  vertices correspond to non-center vertices of  $H$  that coincide with  $\text{opt}(G)$ . On the other hand, each star is large enough so that with constant probability  $H$  includes at least one non-center vertex in each star. In that case, for every center vertex in  $H$ ,  $\text{opt}(H)$  gets two matched vertices. Therefore,  $|\text{opt}(H)| \geq 2t$  internally. It follows that the player can gain an additional  $\Theta(\sqrt{|\text{opt}(G)|})$  matched vertices when deviating from a fixed global optimal matching.

Our main result shows that Example 2.1 is asymptotically tight.

**THEOREM 2.1.** *Let  $G$  be a directed graph and let  $\text{opt}(G)$  be the set of vertices matched under some fixed maximum cardinality matching on  $G$ . Assume that one of the following conditions holds:*

1. *Matchings are restricted to 2-cycles, and  $p_i \leq 1/2$  for each player  $i \in [k]$ , or*
2. *Matchings are restricted to cycles of constant length, and for each player  $i \in [k]$ ,  $1/p_i$  is an integer.*

*Then for any  $\delta > 0$ , with probability at least  $1 - \delta$  over the choice of  $H_i \sim_{p_i} G$ , for all  $i \in [k]$ ,*

$$|\text{opt}(H_i)| - |\text{opt}(G) \upharpoonright H_i| < O\left(\sqrt{|\text{opt}(G)| \ln \frac{k}{\delta}}\right).$$

The proof of Theorem 2.1 involves two main lemmas. The first shows that in expectation  $|\text{opt}(H_i)|$  is at most  $|\text{opt}(G) \upharpoonright H_i|$ . The second asserts that  $|\text{opt}(H_i)|$  is concentrated nicely around its expectation. We formally state these two lemmas without further ado, but defer their proofs to §2.1 and §2.2, respectively.

**LEMMA 2.1.** *Let  $G$  be a directed graph and let  $\text{opt}(G)$  be the set of vertices matched under some fixed maximum cardinality matching on  $G$ . Then  $\mathbb{E}_{H \sim_p G}[|\text{opt}(H)|] \leq p|\text{opt}(G)|$  if (i) matchings are restricted to 2-cycles, and  $p \leq 1/2$ , or (ii)  $1/p$  is an integer.*

LEMMA 2.2. *Let  $G$  be a directed graph and let  $\text{opt}(G)$  be the set of vertices matched under some fixed maximum cardinality matching on  $G$ . Assume matchings are restricted to cycles of length up to a constant  $L$ . Then for any  $\delta > 0$ , with probability  $1 - \delta$  over random choices of  $H_i \sim_{p_i} G$ , for all  $i \in [k]$ ,*

$$\mathbb{E}_{H_i \sim_{p_i} G} [|\text{opt}(H_i)|] - L\sqrt{2\mathbb{E}[|\text{opt}(H_i)|] \ln \frac{2k}{\delta}} < |\text{opt}(H_i)|$$

and

$$\mathbb{E}_{H_i \sim_{p_i} G} [|\text{opt}(H_i)|] + 2L\sqrt{|\text{opt}(G)| \ln \frac{2k}{\delta}} > |\text{opt}(H_i)|$$

We now easily prove our main result—Theorem 2.1—by directly leveraging the two lemmas we just stated.

*Proof of Theorem 2.1.* Since  $\text{opt}(G)$  is fixed and  $H_i$  is drawn from  $G$  independently of  $\text{opt}(G)$ , the expected number of vertices player  $i$  has in  $\text{opt}(G)$  is

$$(2.1) \quad \mathbb{E}_{H_i \sim_{p_i} G} [|\text{opt}(G) \upharpoonright H_i|] = p_i \cdot |\text{opt}(G)|.$$

Moreover,  $|\text{opt}(G) \upharpoonright H_i| = \sum_{v \in \text{opt}(G)} \mathbb{1}_{v \in H_i}$ , where  $\mathbb{1}_{v \in H_i}$  is an indicator variable with value 1 if  $v$  is owned by player  $i$  and 0 otherwise. So,  $\mathbb{1}_{v \in H_i}$  is a random variable that has value 1 with probability  $p_i$ , and value 0 otherwise. Using Hoeffding’s inequality over  $|\text{opt}(G)|$  i.i.d. variables for a fixed  $i \in [k]$ , as well as Equation (2.1), with probability at least  $1 - \frac{\delta}{2k}$  over the choice of  $H_i \sim_{p_i} G$ ,

$$(2.2) \quad |\text{opt}(G) \upharpoonright H_i| \geq p_i \cdot |\text{opt}(G)| - \sqrt{\frac{1}{2} |\text{opt}(G)| \ln \frac{2k}{\delta}}$$

Putting this together with Lemmas 2.1 and 2.2, we have that for all  $i \in [k]$ , with probability  $1 - \delta$ ,

$$\begin{aligned} |\text{opt}(H_i)| &\leq \mathbb{E}_{H_i \sim_{p_i} G} [|\text{opt}(H_i)|] + 2L\sqrt{|\text{opt}(G)| \ln \frac{4k}{\delta}} \\ &\leq p_i |\text{opt}(G)| + 2L\sqrt{|\text{opt}(G)| \ln \frac{4k}{\delta}} \\ &\leq |\text{opt}(G) \upharpoonright H_i| + (2L + 1)\sqrt{|\text{opt}(G)| \ln \frac{4k}{\delta}}, \end{aligned}$$

where the first inequality follows from Lemma 2.2 (using  $\delta' = \delta/2$ ), the second from Lemma 2.1, and the third from applying Equation (2.2) to each  $i \in [k]$ .  $\square$

Note the logarithmic dependence of Theorem 2.1 on the number of players  $k$ . A subtle point is that if the number of players is large, some will have a small  $p_i$ , which means that the expectation of  $|\text{opt}(G) \upharpoonright H_i|$  is small compared to  $|\text{opt}(G)|$ , by Equation (2.1). In that case, a gain of  $\sqrt{|\text{opt}(G)|}$  is significant. Nevertheless,

the theorem’s conceptual message—that following the global matching is individually rational up to lower order terms—holds for any  $p_i = \omega(1/\sqrt{|\text{opt}(G)|})$ .

In addition, recall that Theorem 2.1 considers two cases, (i)  $1/p_i$  is an integer and (ii)  $p_i \leq 1/2$  and  $\text{opt}(G)$  is restricted to 2-cycles. Importantly, these two assumptions are only needed for Lemma 2.1. We conjecture that indeed Lemma 2.1 holds for any  $p \leq \frac{1}{2}$ , whenever  $\text{opt}(G)$  is restricted to cycles of constant length—in which case Theorem 2.1, too, would hold under this weaker assumption. One might wonder why assuming  $p \leq \frac{1}{2}$  is even necessary. To highlight this point, we provide an example in Appendix B.1 that violates the conclusion of Theorem 2.1 for certain values of  $p > 1/2$ . Finally, we remark that if  $1/p_i$  is close to an integer but not itself an integer, one can first round down  $p_i$  to the largest  $q_i < p_i$  such that  $1/q_i$  is an integer, and then apply Theorem 2.1. This would give the same result, up to an additional *constant* fraction of  $|\text{opt}(G)|$ . As  $p_i$  becomes smaller the rounding error also diminishes.

**2.1 Proof of Lemma 2.1** First, let us address Case 2 of the lemma. Consider  $p$  such that  $1/p$  is an integer;  $\text{opt}(G)$  may include cycles of any length. Imagine there are  $\frac{1}{p}$  players, each with probability  $p$ . By symmetry between the players, the expected size of the optimal matching in all subgraphs is equal. Furthermore, the total number of vertices matched by players individually is at most  $\text{opt}(G)$ . Therefore,

$$|\text{opt}(G)| \geq \sum_{i=1}^{1/p} \mathbb{E}_{H_i \sim_p G} [|\text{opt}(H_i)|] = \frac{1}{p} \mathbb{E}_{H \sim_p G} [|\text{opt}(H)|],$$

which proves the claim.

In the remainder of this section we focus on Case 1 of Lemma 2.1, where the matchings are restricted to 2-cycles and  $p \leq 1/2$ . For ease of exposition, we treat  $G$  as an undirected graph: each directed 2-cycle corresponds to an undirected edge, and we may remove directed edges that are not involved in 2-cycles (as they are useless).

Assume there is a partition  $G = G_1 \uplus G_2 \uplus \dots \uplus G_\ell$  of (the undirected graph)  $G$  into edge-disjoint (but not necessarily vertex-disjoint) subgraphs that preserve the size of the optimal matching, i.e.,

$$(2.3) \quad |\text{opt}(G)| = \sum_{i=1}^{\ell} |\text{opt}(G_i)|.$$

Moreover, assume that each of these subgraphs has the property that

$$(2.4) \quad \mathbb{E}_{H \sim_p G_i} [|\text{opt}(H)|] \leq p |\text{opt}(G_i)|.$$

Then, the next equation proves that this property also holds for  $G$  at the global level. That is,

$$\begin{aligned} \mathbb{E}_{H \sim_p G}[|\mathbf{opt}(H)|] &\leq \sum_{i=1}^{\ell} \mathbb{E}_{H \sim_p G}[|\mathbf{opt}(H \cap G_i)|] \\ &= \sum_{i=1}^{\ell} \mathbb{E}_{H \sim_p G_i}[|\mathbf{opt}(H)|] \\ &\leq \sum_{i=1}^{\ell} p |\mathbf{opt}(G_i)| = p |\mathbf{opt}(G)|. \end{aligned}$$

For the first transition,  $H \cap G_i$  is the graph with edges that are present in both  $H$  and  $G_i$ ; the intuition behind this inequality is that we are essentially allowed to match the same vertices multiple times on the right hand side. The third and fourth transitions follow from Equations (2.3) and (2.4).

So, it remains to find a partition of  $G$  into edge-disjoint subgraphs,  $G_1 \uplus G_2 \uplus \dots \uplus G_\ell$ , which satisfies (2.3) and (2.4). We prove that the Edmonds-Gallai Decomposition [20] can be used to construct a partition satisfying these properties.

**LEMMA 2.3. (EDMONDS-GALLAI DECOMPOSITION)**  
*Let  $G = (V, E)$  be an undirected graph, let  $B$  be the set of vertices matched by every maximum cardinality matching in  $G$ , and let  $D = V \setminus B$ . Furthermore, partition  $B$  into subsets  $A$  and  $C = B \setminus A$ , where  $A$  is the set of vertices with at least one neighbor outside  $B$ . And let  $D_1, \dots, D_r$ , be the connected components of the induced subgraph  $G[D]$ . Then the following properties hold.*

1.  $\mathbf{opt}(G)$  matches each node in  $A$  to a distinct connected component of  $G[D]$ .
2. Each  $D_i$  is factor-critical, i.e., deleting any one vertex of  $D_i$  leads to a perfect matching in the remainder of  $D_i$ .

We now describe how the Edmonds-Gallai Decomposition is used to construct the desired partition of  $G$ . For the  $i$ th connected component of  $G[C]$  and  $G[D]$ , create a subgraph  $G_i$  corresponding to its edges. Furthermore, for each vertex  $i \in A$ , create a subgraph  $G_i$  corresponding to the set of edges incident on  $i$ . If there is an edge between two vertices of  $A$ ,  $i$  and  $i'$ , then include the edge in only one of  $G_i$  or  $G_{i'}$ . Since the Edmonds-Gallai Decomposition has no edges between  $C$  and  $D$ ,  $G = \uplus_i G_i$  forms a partition of the edge set of  $G$ . See Figure 2 for an example of this construction.

We argue that the foregoing partition satisfies Equation (2.3).

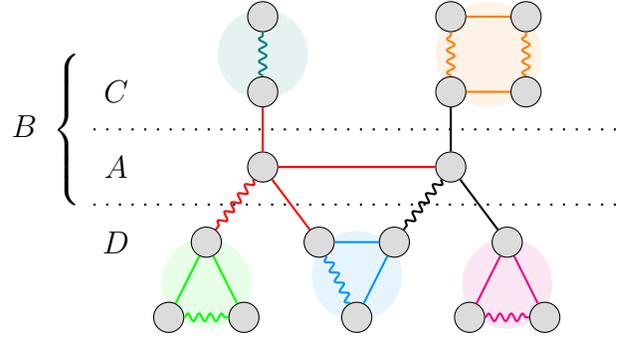


Figure 2: A graph demonstrating the Edmonds-Gallai Decomposition and the edge-disjoint graph partition for the proof of Lemma 2.1. In this graph, each color represents one  $G_i$  in the partition  $G = \uplus_i G_i$  and the wavy edges represent the matched edges in  $\mathbf{opt}(G)$ .

**CLAIM 2.1.**  $|\mathbf{opt}(G)| = \sum_i |\mathbf{opt}(G_i)|$ .

*Proof.* Let us define  $\mathbf{opt}(G) \upharpoonright^* G_i$  to be the vertices of  $\mathbf{opt}(G) \upharpoonright G_i$  that are matched by edges that lie within  $G_i$ . Since the decomposition is edge-disjoint, it holds that  $|\mathbf{opt}(G)| = \sum_i |\mathbf{opt}(G) \upharpoonright^* G_i|$ . It is therefore sufficient to show that for all  $i$ ,  $|\mathbf{opt}(G_i)| = |\mathbf{opt}(G) \upharpoonright^* G_i|$ . There are three cases:

1.  $G_i$  corresponds to a component of  $G[C]$ . Recall that  $\mathbf{opt}(G)$  matches all vertices of  $C \subseteq B$ . Moreover,  $C$  has no edges to  $D$ , and  $A$  is only matched with  $D$ , so the vertices of  $G_i$  have no matched edges outside  $G_i$ . It follows that  $\mathbf{opt}(G) \upharpoonright^* G_i$  is itself a perfect matching on  $G_i$ , and  $|\mathbf{opt}(G_i)| = |\mathbf{opt}(G) \upharpoonright^* G_i|$ .
2.  $G_i$  corresponds to a star with vertex  $i \in A$ : For each  $i \in A$ , by the first property of the Edmonds-Gallai Decomposition,  $i$  is matched to a distinct component of  $G[D]$ . Therefore,  $\mathbf{opt}(G) \upharpoonright^* G_i$  includes an edge from  $\mathbf{opt}(G)$ . Since any star can have at most one matched edge, we have that  $|\mathbf{opt}(G) \upharpoonright^* G_i| = |\mathbf{opt}(G)|$ .
3.  $G_i$  corresponds to a component of  $G[D]$ : Since such a component is factor-critical, it has an odd number of vertices, and, for any vertex, a maximum matching that covers all other vertices. Therefore, both  $\mathbf{opt}(G_i)$  and  $\mathbf{opt}(G) \upharpoonright^* G_i$  match all but one vertex of this component, and  $|\mathbf{opt}(G_i)| = |\mathbf{opt}(G) \upharpoonright^* G_i|$ .

□

Next, we show that  $G = \uplus_i G_i$  satisfies (2.4). There are three types of  $G_i$  in this partition: (i)  $G_i$  is a star,

(ii)  $G_i$  is a component of  $G[D]$  and has a matching that covers all but one vertex, and (iii)  $G_i$  is a component of  $G[C]$  and has a perfect matching.

Let us first address case (i)—that of a star. Clearly it holds that  $|\text{opt}(G_i)| = 2$ . Now,  $|\text{opt}(H)| \in \{0, 2\}$ , and for  $\text{opt}(H)$  to be non-empty,  $H$  must include the center of the star, which happens with probability  $p$ .

The following claim establishes Equation (2.4) in cases (ii) and (iii). Note that this is the only place where the assumption  $p \leq 1/2$  is used.

**CLAIM 2.2.** *For any  $p \leq \frac{1}{2}$ , and any graph  $G$  with  $n$  vertices such that  $|\text{opt}(G)| \geq n - 1$ ,*

$$\mathbb{E}_{H \sim_p G} [|\text{opt}(H)|] \leq p |\text{opt}(G)|.$$

*Proof.* Let  $t \in \mathbb{N}$  and  $p \in [0, 1/2]$ . It holds that

$$(2.5) \quad \frac{1}{2} - \frac{1}{2}(1 - 2p)^{2t+1} \geq p,$$

because the left hand side is concave and has value 0 for  $p = 0$  and  $1/2$  for  $p = 1/2$ . We also use the equalities

$$(2.6) \quad \sum_{i=0}^k \binom{k}{i} x^i = (1 + x)^k,$$

and

$$(2.7) \quad \sum_{i=1}^k i \binom{k}{i} x^{i-1} = k(1 + x)^{k-1}.$$

Assume that  $n = 2t + 1$  for some  $t \geq 0$ . By the claim's assumption, it holds that  $\text{opt}(G) = 2t$ . Any matching among a set of  $i$  vertices matches at most  $2\lfloor i/2 \rfloor$  vertices. Hence, the expected matching size of the subgraph induced by a random set of vertices when each vertex is included independently with probability  $p$  is

$$\begin{aligned} \mathbb{E}_{H \sim_p G} [|\text{opt}(H)|] &\leq \sum_{i=1}^{2t+1} 2\lfloor i/2 \rfloor \binom{2t+1}{i} p^i (1-p)^{2t+1-i} \\ &= p(1-p)^{2t} \sum_{i=1}^{2t+1} i \binom{2t+1}{i} \left(\frac{p}{1-p}\right)^{i-1} \\ &\quad - \frac{1}{2}(1-p)^{2t+1} \sum_{i=0}^{2t+1} \binom{2t+1}{i} \left(\frac{p}{1-p}\right)^i \\ &\quad + \frac{1}{2}(1-p)^{2t+1} \sum_{i=0}^{2t+1} \binom{2t+1}{i} \left(\frac{p}{1-p}\right)^i (-1)^i \\ &= (2t+1)p - \frac{1}{2} + \frac{1}{2}(1-2p)^{2t+1} \\ &\leq 2tp, \end{aligned}$$

where the penultimate transition follows by applying Equation (2.7) to the first term on the left hand side, and Equation (2.7) to the second and third terms; and the last transition follows from Equation (2.5).

If  $n = 2t$  and  $\text{opt}(G) \geq n - 1$ , it must hold that  $\text{opt}(G) = 2t$ , because each edge corresponds to two matched vertices. Moreover,

$$\begin{aligned} \mathbb{E}_{H \sim_p G} [|\text{opt}(H)|] &\leq \sum_{i=1}^{2t} 2\lfloor i/2 \rfloor \binom{2t}{i} p^i (1-p)^{2t-i} \\ &\leq p(1-p)^{2t-1} \sum_{i=1}^{2t} i \binom{2t}{i} \left(\frac{p}{1-p}\right)^{i-1} \\ &= p(1-p)^{2t-1} 2t \left(1 + \frac{p}{1-p}\right)^{2t-1} \\ &= 2tp. \end{aligned}$$

□

Having established (2.4), the proof of Lemma 2.1 is now complete.

**2.2 Proof of Lemma 2.2** We will prove the following equivalent formulation of Lemma 2.2:

$$(2.8) \quad \Pr_{H \sim G} [|\text{opt}(H)| \geq \mathbb{E}_{H \sim G} [|\text{opt}(H)|] + \epsilon] \leq \exp\left(-\frac{\epsilon^2}{4L^2 |\text{opt}(G)|}\right),$$

and

$$(2.9) \quad \Pr_{H \sim G} [|\text{opt}(H)| \leq \mathbb{E}_{H \sim G} [|\text{opt}(H)|] - \epsilon] \leq \exp\left(-\frac{\epsilon^2}{2L^2 \mathbb{E}[|\text{opt}(H)|]}\right).$$

Let us first describe a failed approach for proving the lemma, which brings to light some subtleties in the above inequalities. Consider an explicit description of  $|\text{opt}(H)|$  as a function of  $n$  random variables,  $X_1, \dots, X_n$ , where  $X_i = 1$  if vertex  $i$  is in  $H$  and 0 otherwise. Then  $|\text{opt}(H)| = f(X_1, \dots, X_n)$  is the size of the optimal matching on  $H$ . One can show that  $f(\cdot)$  is  $L$ -Lipschitz, that is, changing  $X_i$  to  $\neg X_i$ , which corresponds to adding or removing one vertex from  $H$ , changes the size of the maximum matching by at most  $L$ . Lipschitz functions are known to enjoy strong concentration guarantees, as shown by McDiarmid's inequality,

$$\Pr [|f - \mathbb{E}[f]| > \epsilon] \leq 2 \exp\left(-\frac{2\epsilon^2}{\sum_{i=1}^n c_i^2}\right),$$

where  $c_i$  is the Lipschitz constant for the  $i$ th variable, that is, for all  $i$  and for every possible input  $x_1, \dots, x_n$ ,  $|f(x_1, \dots, x_i, \dots, x_n) - f(x_1, \dots, \neg x_i, \dots, x_n)| \leq c_i$ .

While there are only  $|\text{opt}(G)|$  variables that truly participate in  $\text{opt}(G)$ , even vertices that are not in  $\text{opt}(G)$  can participate in matchings of subsets of  $G$ , and as a result have a non-zero Lipschitz constant. Therefore, using McDiarmid's inequality for the concentration of  $\text{opt}(H)$  gives an  $O(\sqrt{n})$  gap between  $\text{opt}(H)$  and  $\mathbb{E}_{H \sim G}[\text{opt}(H)]$ .

Instead, in order to prove a gap of  $\tilde{O}(\sqrt{|\text{opt}(G)|})$ , we use two alternative concentration bounds from statistical learning theory, which have recently been used to simplify and prove tight concentration and sample complexity results for learning combinatorial functions [30].

LEMMA 2.4. [12, Theorem 12] *Let  $X_1, X_2, \dots, X_n$  be independent random variables, each taking values in a set  $\mathcal{X}$ . Let  $f : \mathcal{X}^n \rightarrow \mathbb{R}$  be a measurable function. Let  $X'_1, \dots, X'_n$  be independent copies of  $X_1, \dots, X_n$  and for all  $i \in [n]$ , define  $f'_i = f(X_1, \dots, X'_i, \dots, X_n)$ . For all  $i \in [n]$  and  $\mathbf{x} \in \mathcal{X}^n$  assume that there exists  $C > 0$ , such that*

$$\mathbb{E} \left[ \sum_{i=1}^n (f - f'_i)^2 \cdot \mathbb{1}_{f > f'_i} \mid \mathbf{x} \right] \leq C,$$

then for all  $\epsilon > 0$ ,

$$\Pr[f > \mathbb{E}[f] + \epsilon] \leq e^{-\epsilon^2/4C}.$$

We show that the conditions of Lemma 2.4 hold for  $f(x_1, \dots, x_n) = \text{opt}(H)$ . Let  $\mathbf{x} = (x_1, \dots, x_n)$  and  $\mathbf{x}'_i = (x_1, \dots, x'_i, \dots, x_n)$ . For all  $\mathbf{x}$ , let  $H_{\mathbf{x}}$  be the subgraph corresponding to non-zero variables of  $\mathbf{x}$ . Note that if  $x_i$  is replaced by  $x'_i$  and the matching size is reduced, then the decrease is at most the maximum cycle length  $L$ . Furthermore, the only variables that can lead to a non-zero decrease from  $|\text{opt}(H_{\mathbf{x}})|$  to  $|\text{opt}(H_{\mathbf{x}'_i})|$  are variables that are in every optimal matching on  $H_{\mathbf{x}}$ . Therefore, there are at most  $|\text{opt}(H_{\mathbf{x}})| \leq |\text{opt}(G)|$  such variables. We conclude that for all  $\mathbf{x}$ ,

$$\mathbb{E} \left[ \sum_{i=1}^n (f - f'_i)^2 \cdot \mathbb{1}_{f > f'_i} \mid \mathbf{x} \right] \leq L^2 |\text{opt}(G)|.$$

The proof of the upper tail (2.8) follows immediately by using Lemma 2.4 with  $C = L^2 |\text{opt}(G)|$ .

Unfortunately, Lemma 2.4 and its variants for lower-tail concentration cannot be used to establish the desired lower-tail bound (2.9). Indeed, consider the condition  $\mathbb{E}[\sum_{i=1}^n (f - f'_i)^2 \cdot \mathbb{1}_{f < f'_i} \mid \mathbf{x}] \leq C$ ; while removing one of only  $|\text{opt}(G)|$  vertices can reduce the size of a matching, it may be possible that for some subgraph of  $G$ , adding any of the remaining vertices increases the

size of the matching. Instead, we use the lower-tail concentration of *self-bounding functions* [14].

DEFINITION 2.1. [14] A function  $g : \mathcal{X}^n \rightarrow \mathbb{R}$  is  $(a, b)$ -self-bounding if there exist functions  $g_{-i} : \mathcal{X}^{n-1} \rightarrow \mathbb{R}$  for all  $i \in [n]$  such that for all  $\mathbf{x} = (x_1, \dots, x_n) \in \mathcal{X}^n$  and  $i \in [n]$ ,

$$0 \leq g(\mathbf{x}) - g_{-i}(\mathbf{x}_{-i}) \leq 1,$$

and

$$\sum_{i=1}^n (g(\mathbf{x}) - g_{-i}(\mathbf{x}_{-i})) \leq a \cdot g(\mathbf{x}) + b,$$

where  $\mathbf{x}_{-i} = (x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n)$  is obtained by dropping the  $i$ th component of  $\mathbf{x}$ .

LEMMA 2.5. [14] *If  $Z = g(X_1, \dots, X_n)$ , where  $X_i \in \{0, 1\}$  are independent random variables and  $g$  is an  $(a, b)$ -self-bounding function with  $a \geq \frac{1}{3}$ , then for any  $0 < t < \mathbb{E}[Z]$ ,*

$$\Pr[Z \leq \mathbb{E}[Z] - t] \leq \exp\left(-\frac{t^2}{2a \cdot \mathbb{E}[Z] + 2b}\right).$$

Let  $g(\mathbf{x})$  be  $\frac{1}{L}$  times the size of optimal matching on the subgraph whose vertices correspond to the non-zero  $x_i$ 's, i.e.,  $g(\mathbf{x}) = \frac{1}{L} |\text{opt}(H_{\mathbf{x}})|$ . Define  $g_{-i}(\mathbf{x}_{-i}) = \min_{x_i} g(\mathbf{x})$ . We show that  $g(\cdot)$  is  $(L, 0)$ -self-bounding.

Since  $g_{-i}(\mathbf{x}_{-i})$  is the matching size in  $H_{\mathbf{x}_{-i}}$  and  $|\text{opt}(H_{\mathbf{x}})| \geq |\text{opt}(H_{\mathbf{x}_{-i}})| \geq |\text{opt}(H_{\mathbf{x}})| - L$ , we have that  $0 \leq g(\mathbf{x}) - g_{-i}(\mathbf{x}_{-i}) \leq 1$ . Furthermore,  $g(\mathbf{x}) - g_{-i}(\mathbf{x}_{-i})$  is non-zero only if vertex  $i$  was in every  $\text{opt}(H_{\mathbf{x}})$ . Since there are at most  $|\text{opt}(H_{\mathbf{x}})|$  such variables, we have

$$\sum_{i=1}^n (g(\mathbf{x}) - g_{-i}(\mathbf{x}_{-i})) \leq |\text{opt}(H_{\mathbf{x}})| = L \cdot g(\mathbf{x}).$$

Because  $L > \frac{1}{3}$ , we can use Lemma 2.5 with  $\epsilon = tL$ , and obtain

$$\begin{aligned} \Pr[|\text{opt}(H_i)| \leq \mathbb{E}[|\text{opt}(H_i)|] - \epsilon] \\ \leq \exp\left(-\frac{(\epsilon/L)^2}{2L(\frac{1}{L}\mathbb{E}[|\text{opt}(H_i)|])}\right) \\ \leq \exp\left(-\frac{\epsilon^2}{2L^2\mathbb{E}[|\text{opt}(H_i)|]}\right). \end{aligned}$$

One might wonder whether the upper-tail bound of self-bounding functions could be used similarly to achieve an improved upper bound that depends on  $|\text{opt}(H_i)|$  instead of  $|\text{opt}(G)|$ . In Appendix A, we answer this question in the negative and discuss this point further.

### 3 Individually Rational Matchings That Are Almost Optimal

In this section, we provide a *simple and practical* mechanism for kidney exchange that guarantees individual rationality, and with high probability yields a matching that is optimal up to lower-order terms. In comparison to the results of §2, its disadvantage is that it requires modifying deployed matching mechanisms, which simply return some optimal matching—our mechanism selects a *specific* matching (which may be suboptimal). However, it is only a minor modification, and therefore has the potential to inform practice.

Let us first consider the case where  $\text{opt}(G)$  is restricted to 2-cycles. In this case we can represent  $G$  as an undirected graph, as in §2.1. Let  $H_1, \dots, H_k$  be the subgraphs corresponding to the players. Consider the following matching mechanism,  $\mathcal{M}(H_1, \dots, H_k)$ : First, compute the matching  $M = \bigcup_i \text{opt}(H_i)$ ; then grow  $M$  to a globally maximum cardinality matching by repeatedly applying augmenting paths. While an augmenting path changes the structure of a matching by adding and removing edges, it strictly expands the set of matched vertices. Therefore, this mechanism leads to a maximum cardinality matching on  $G$ , with the property that  $\mathcal{M}(H_1, \dots, H_k) \supseteq \bigcup_i \text{opt}(H_i)$  for all  $i \in [k]$ . That is,  $\mathcal{M}$  is individually rational.

The performance of the above mechanism for 2-cycles holds even when the subgraphs owned by players are chosen adversarially, rather than through a random process. Furthermore, this mechanism enjoys the stronger guarantee that every vertex that is matched under  $\text{opt}(H_i)$  is also matched under  $\mathcal{M}(H_1, \dots, H_k)$ . As we discussed earlier (see Figure 1), these strong guarantees are unattainable when cycles of length 3 are allowed. But in our model for randomly generating  $H_1, \dots, H_k$ , there is a mechanism that is individually rational and almost optimal, as we show next.

**COROLLARY 3.1.** *Let  $G$  be a directed graph. Consider optimal matchings on  $G$  that are restricted to constant-length cycles. For all  $i \in [k]$ , let  $p_i = 1/k$ . Then there exists a mechanism  $\mathcal{M}$  such that  $\mathcal{M}(H_1, \dots, H_k)$  is individually rational, and for any  $\delta > 0$ , with probability at least  $1 - \delta$ ,*

$$|\mathcal{M}(H_1, \dots, H_k)| \geq |\text{opt}(G)| - O\left(k\sqrt{|\text{opt}(G)| \ln \frac{k}{\delta}}\right).$$

As advertised, the mechanism underlying Corollary 3.1 is very simple: Choose an arbitrary optimal matching  $\text{opt}(G)$ , independently of  $H_1, \dots, H_k$ . If for all players  $i \in [k]$  we have  $|\text{opt}(H_i)| \leq |\text{opt}(G) \upharpoonright H_i|$ , then  $\mathcal{M}(H_1, \dots, H_k) = \text{opt}(G)$ . Else,

let  $\mathcal{M}(H_1, \dots, H_k) = \bigcup_i \text{opt}(H_i)$ . We call this mechanism the *Veto* mechanism, as any player can veto the proposed optimal matching. Alternatively, we can let players defect if they wish, while allowing the remaining players to continue to work together; for our mathematical purposes this is the same as the Veto mechanism, but the latter interpretation may be even more appealing from a practical viewpoint.

In a nutshell, the idea behind the proof of Corollary 3.1 is that because  $|\text{opt}(H_i)|$  is concentrated around its expectation by Lemma 2.2, if some player wants to veto the proposed matching then it is likely that  $\mathbb{E}[|\text{opt}(H_i)|]$  is close to  $|\text{opt}(G) \upharpoonright H_i|$ , which is tightly concentrated around  $|\text{opt}(G)|/k$  by Hoeffding’s inequality. But due to symmetry, this is true for all players, so the players can obtain on their own almost what they can obtain by collaborating. Below we formalize these ideas.

*Proof of Corollary 3.1.* We start with an informal overview of the proof. In a nutshell, the idea is that because  $|\text{opt}(H_i)|$  is concentrated around its expectation by Lemma 2.2, if some player wants to veto the proposed matching then it is likely that  $\mathbb{E}[|\text{opt}(H_i)|]$  is close to  $|\text{opt}(G) \upharpoonright H_i|$ , which is tightly concentrated around  $|\text{opt}(G)|/k$  by Hoeffding’s inequality. But due to symmetry, this is true for all players, so the players can obtain on their own almost what they can obtain by collaborating. Details follow.

Let  $p = 1/k$ . If

$$\mathbb{E}_{H_i \sim_p G}[|\text{opt}(H_i)|] \leq p|\text{opt}(G)| - (2L+1)\sqrt{|\text{opt}(G)| \ln \frac{2k}{\delta}}$$

then

$$\begin{aligned} & \Pr_{H_i \sim G}[|\text{opt}(H_i)| > |\text{opt}(G) \upharpoonright H_i|] \\ & \leq \Pr_{H_i \sim G}\left[|\text{opt}(H_i)| - |\text{opt}(G) \upharpoonright H_i| + \left(p|\text{opt}(G)| - \mathbb{E}[|\text{opt}(H_i)|] - (2L+1)\sqrt{|\text{opt}(G)| \ln \frac{2k}{\delta}}\right) > 0\right] \\ & \leq \Pr_{H_i \sim G}\left[|\text{opt}(H_i)| - \mathbb{E}[|\text{opt}(H_i)|] > 2L\sqrt{|\text{opt}(G)| \ln \frac{2k}{\delta}}\right] \\ & \quad + \Pr_{H_i \sim G}\left[p|\text{opt}(G)| - |\text{opt}(G) \upharpoonright H_i| > \sqrt{|\text{opt}(G)| \ln \frac{2k}{\delta}}\right] \\ & \leq \frac{\delta}{k}, \end{aligned}$$

where the last inequality holds by the upper-tail bound of Lemma 2.2 and Hoeffding’s inequality. Therefore, with probability  $1 - \delta$ , no player vetoes the proposed optimal matching, and  $\mathcal{M}(H_1, \dots, H_k) = \text{opt}(G)$  is optimal.

On the other hand, if

$$\mathbb{E}_{H_i \sim_p G}[\text{opt}(H_i)] \geq p |\text{opt}(G)| - (2L+1) \sqrt{|\text{opt}(G)| \ln \frac{2k}{\delta}},$$

then the expected total size of the internal matching is large, and we can fall back to the internal matchings. Indeed, note that  $\mathbb{E}_{H_i \sim_p G}[|\text{opt}(H_i)|] \leq p|\text{opt}(G)|$  by symmetry. By the lower-tail bound of Lemma 2.2, with probability  $1 - \delta$ , for all  $i \in [k]$ ,

$$|\text{opt}(H_i)| \geq \mathbb{E}_{H_i \sim_p G}[|\text{opt}(H_i)|] - L \sqrt{2p|\text{opt}(G)| \ln \frac{2k}{\delta}}.$$

Therefore, with probability  $1 - \delta$ ,

$$\begin{aligned} & \sum_{i=1}^k |\text{opt}(H_i)| \\ & \geq k \mathbb{E}_{H_i \sim G}[|\text{opt}(H_i)|] - kL \sqrt{\frac{2}{k} |\text{opt}(G)| \ln \frac{2k}{\delta}} \\ & \geq |\text{opt}(G)| - k(2L+1) \sqrt{|\text{opt}(G)| \ln \frac{2k}{\delta}} \\ & \quad - kL \sqrt{\frac{2}{k} |\text{opt}(G)| \ln \frac{2k}{\delta}} \\ & = |\text{opt}(G)| - O\left(k \sqrt{|\text{opt}(G)| \ln \frac{k}{\delta}}\right). \end{aligned}$$

So,  $\mathcal{M}(H_1, \dots, H_k) = \bigcup_i \text{opt}(H_i)$  is a near optimal.  $\square$

We remark—without proof—that Corollary 3.1 still holds even if the probabilities, instead of being equal, are of the form  $1/s^{t_i}$  for a fixed  $s \in \mathbb{N}$ , and possibly different  $t_1, \dots, t_n$ . Similarly to Theorem 2.1, we conjecture that the statement actually holds for any  $p_1, \dots, p_n$  such that  $p_i \geq 1/2$  for all  $i \in [k]$ . To prove this, one would need to strengthen Lemma 2.1, as discussed in §2. But now there is another difficulty: One would need to show that if  $\mathbb{E}[\text{opt}(H_i)]$  is close to  $p_i |\text{opt}(G)|$  for one player, then the same is true for all players—in which case falling back to the internal matchings is almost optimal. In the symmetric case, this claim trivially holds, which is precisely why we assume that  $p_i = 1/k$  for all  $i \in [k]$ .

While we require a relatively strong assumption on the probabilities, it is satisfying that the theorem's bound is asymptotically tight. To show this, we present and analyze an example of a graph with  $n$  vertices where, with constant probability, *any* individually rational matching is smaller than the optimal matching by  $\Omega(\sqrt{n})$ .

**EXAMPLE 3.1.** Suppose that there are two players, each with probability  $p = \frac{1}{2}$ . Consider the graph  $G$

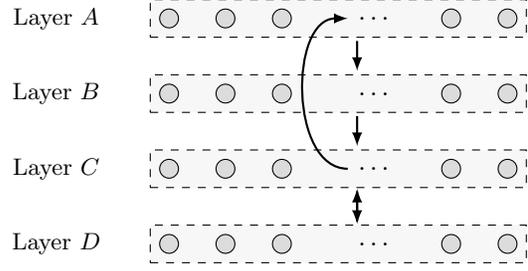


Figure 3: The graph constructed in Example 3.1.

shown in Figure 3, which consists of four layers  $A$ ,  $B$ ,  $C$  and  $D$ , each with  $n/4$  vertices. Any two layers of the graph are fully connected if there is an edge between them according to Figure 3. That is, the edge set of this graph is such that *any 3 vertices* from  $A$ ,  $B$ , and  $C$ , respectively, form a directed 3-cycle, and *any 2 vertices* from  $C$  and  $D$ , respectively, form a directed 2-cycle. It is optimal to match the vertices in  $A$ ,  $B$ , and  $C$  via 3-cycles, and therefore  $|\text{opt}(G)| = 3n/4$ .

It is easy to show, using the standard deviation of the binomial distribution, that the number of vertices a player owns in each layer deviates by  $\pm\Theta(\sqrt{n})$  from its expectation (either larger or smaller) with constant probability. Denote the number of vertices owned by player 1 in layers  $A$ ,  $B$ ,  $C$ , and  $D$  by  $a$ ,  $b$ ,  $c$ , and  $d$ , respectively. We focus on the case where  $c \in \{n/8 + \sqrt{n}, \dots, n/8 + (3/2)\sqrt{n}\}$ ,  $a$  and  $b$  are both in  $\{n/8 - (3/2)\sqrt{n}, \dots, n/8 - \sqrt{n}\}$ , and  $d \geq 3\sqrt{n}$ —which happens with constant probability. Intuitively, player 1 is doing well, because he owns significantly more than half of the vertices in layer  $C$ , which is especially important.

In the foregoing case,  $\text{opt}(H_1)$  is obtained by taking  $\min\{a, b\}$  3-cycles between  $A$ ,  $B$ , and  $C$  (as many as possible), and then  $c - \min\{a, b\}$  2-cycles between  $D$  and the unmatched vertices of  $C$ . Therefore,

$$\begin{aligned} |\text{opt}(H_1)| &= 3 \cdot \min\{a, b\} + 2(c - \min\{a, b\}) \\ &= 2 \cdot c + \min\{a, b\} = 3n/8 + \Theta(\sqrt{n}). \end{aligned}$$

On the other hand, consider some matching  $\mathcal{M}$  with  $x$  3-cycles and  $y$  2-cycles, such that (without loss of generality)  $x + y = n/4$ —as the total number is constrained by the  $n/4$  vertices in layer  $C$ . Note that under the optimal matching we have  $x = n/4$ , and

$$|\text{opt}(G) \upharpoonright H_1| = a + b + c = 3n/8 - \Theta(n).$$

More generally, we have that  $|\mathcal{M} \upharpoonright H_1| \leq a + b + c + y$ . In order to guarantee that  $\mathcal{M}$  is individually rational for player 1, we must close the gap between  $|\text{opt}(H_1)|$  and

$|\mathcal{M} \upharpoonright H_1|$ , which implies that  $y = \Omega(\sqrt{n})$ . That is, we must sacrifice  $\Omega(\sqrt{n})$  3-cycles in favor of 2-cycles. But that means that  $|\mathcal{M}| \leq |\text{opt}(G)| - \Omega(\sqrt{n})$ .

Finally, note that Corollary 3.1 assumes cycles of constant length (as does Theorem 2.1). As noted in §1, major kidney exchanges do, in fact, only use very short cycles and chains (which can also be represented as cycles) in each match run. But it is nevertheless interesting to point out that the same statement is false when long cycles are allowed. Indeed, in Appendix B.2 we present an example of a graph with long cycles, where (with high probability) every individually rational matching is smaller than the optimal matching by  $\Omega(n)$ .

## References

- [1] A. Abdulkadiroğlu, P. Pathak, and A. E. Roth. The New York City high school match. *American Economic Review*, 95(2):364–367, 2005.
- [2] D. J. Abraham, A. Blum, and T. Sandholm. Clearing algorithms for barter exchange markets: Enabling nationwide kidney exchanges. In *Proceedings of the 8th ACM Conference on Economics and Computation (EC)*, pages 295–304, 2007.
- [3] M. Adamczyk. Improved analysis of the greedy algorithm for stochastic matching. *Information Processing Letters*, 111(15):731–737, 2011.
- [4] M. Akbarpour, S. Li, and S. Oveis Gharan. Dynamic matching market design. In *Proceedings of the 15th ACM Conference on Economics and Computation (EC)*, page 355, 2014.
- [5] I. Ashlagi, F. Fischer, I. Kash, and A. D. Procaccia. Mix and match: A strategyproof mechanism for multi-hospital kidney exchange. *Games and Economic Behavior*, 91:284–296, 2015.
- [6] I. Ashlagi, D. Gamarnik, M. A. Rees, and A. E. Roth. The need for (long) chains in kidney exchange. NBER Working Paper Series No. 18202, 2012.
- [7] I. Ashlagi and A. Roth. Free riding and participation in large scale, multi-hospital kidney exchange. *Theoretical Economics*, 9:817–863, 2014.
- [8] S. Assadi, S. Khanna, and Y. Li. The stochastic matching problem with (very) few queries. In *Proceedings of the 17th ACM Conference on Economics and Computation (EC)*, 2016. Forthcoming.
- [9] N. Bansal, A. Gupta, J. Li, J. Mestre, V. Nagarajan, and A. Rudra. When LP is the cure for your matching woes: Improved bounds for stochastic matchings. *Algorithmica*, 63(4):733–762, 2012.
- [10] A. Blum, J. P. Dickerson, N. Haghtalab, A. D. Procaccia, T. Sandholm, and A. Sharma. Ignorance is almost bliss: Near-optimal stochastic matching with few queries. In *Proceedings of the 16th ACM Conference on Economics and Computation (EC)*, pages 325–342, 2015.
- [11] A. Blum, A. Gupta, A. D. Procaccia, and A. Sharma. Harnessing the power of two crossmatches. In *Proceedings of the 14th ACM Conference on Economics and Computation (EC)*, pages 123–140, 2013.
- [12] S. Boucheron, G. Lugosi, and O. Bousquet. Concentration inequalities. In O. Bousquet, U. von Luxburg, and G. Rätsch, editors, *Advanced Lectures on Machine Learning*, pages 208–240. Springer, 2004.
- [13] S. Boucheron, G. Lugosi, and P. Massart. A sharp concentration inequality with applications. *Random Structures and Algorithms*, 16(3):277–292, 2000.
- [14] S. Boucheron, G. Lugosi, and P. Massart. On concentration of self-bounding functions. *Electronic Journal of Probability*, 14(64):1884–1899, 2009.
- [15] I. Caragiannis, A. Filos-Ratsikas, and A. D. Procaccia. An improved 2-agent kidney exchange mechanism. *Theoretical Computer Science*, 589:53–60, 2015.
- [16] N. Chen, N. Immorlica, A. R. Karlin, M. Mahdian, and A. Rudra. Approximating matches made in heaven. In *Proceedings of the 36th International Colloquium on Automata, Languages and Programming (ICALP)*, pages 266–278, 2009.
- [17] K. P. Costello, P. Tetali, and P. Tripathi. Stochastic matching with commitment. In *Proceedings of the 39th International Colloquium on Automata, Languages and Programming (ICALP)*, pages 822–833, 2012.
- [18] J. P. Dickerson, A. D. Procaccia, and T. Sandholm. Optimizing kidney exchange with transplant chains: Theory and reality. In *Proceedings of the 11th International Conference on Autonomous Agents and Multi-Agent Systems (AAMAS)*, pages 711–718, 2012.
- [19] J. P. Dickerson, A. D. Procaccia, and T. Sandholm. Failure-aware kidney exchange. In *Proceedings of the 14th ACM Conference on Economics and Computation (EC)*, pages 323–340, 2013.
- [20] J. Edmonds. Paths, trees, and flowers. *Canadian Journal of Mathematics*, 17:449–467, 1965.
- [21] G. Goel and P. Tripathi. Matching with our eyes closed. In *Proceedings of the 53rd Symposium on Foundations of Computer Science (FOCS)*, pages 718–727, 2012.
- [22] A. Gupta and V. Nagarajan. A stochastic probing problem with applications. In *Proceedings of the 16th Conference on Integer Programming and Combinatorial Optimization (IPCO)*, pages 205–216, 2013.
- [23] C. Hajaj, J. P. Dickerson, A. Hassidim, T. Sandholm, and D. Sarne. Strategy-proof and efficient kidney exchange using a credit mechanism. In *Proceedings of the 29th AAAI Conference on Artificial Intelligence (AAAI)*, pages 921–928, 2015.
- [24] C. McDiarmid and B. Reed. Concentration for self-bounding functions and an inequality of Talagrand. *Random Structures and Algorithms*, 29(4):549–557, 2006.
- [25] A. E. Roth and E. Peranson. The redesign of the matching market for American physicians: Some engineering aspects of economic design. *American Economic Review*, 89(4):748–780, 1999.

- [26] A. E. Roth, T. Sönmez, and M. U. Ünver. Kidney exchange. *Quarterly Journal of Economics*, 119(2):457–488, 2004.
- [27] A. E. Roth, T. Sönmez, and M. U. Ünver. Pairwise kidney exchange. *Journal of Economic Theory*, 125:151–188, 2005.
- [28] 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(3):828–851, 2007.
- [29] P. Toulis and D. C. Parkes. Design and analysis of multi-hospital kidney exchange mechanisms using random graphs. *Games and Economic Behavior*, 91:360–382, 2015.
- [30] J. Vondrák. A note on concentration of submodular functions. arXiv:1005.2791, 2010.

## A Self-Bounding Functions Do Not Lead to Better Upper Bounds

One might wonder whether the upper-tail bound of self-bounding functions could be used similarly to achieve an improved upper bound of  $L\sqrt{2 \cdot \mathbb{E}[|\text{opt}(H_i)|]} \ln \frac{2k}{\delta}$  for Lemma 2.2—that is, a bound that depends on  $|\text{opt}(H_i)|$  instead of  $|\text{opt}(G)|$ . Here, we answer this question in the negative. The next lemma bounds the upper tail of  $(a, b)$ -self-bounding functions.

LEMMA A.1. [14] *If  $Z = g(X_1, \dots, X_n)$ , where  $X_i \in \{0, 1\}$  are independent random variables and  $g$  is an  $(a, b)$ -self-bounding function, then for any  $0 < t < \mathbb{E}[Z]$ ,*

$$\Pr[Z \geq \mathbb{E}[Z] + t] \leq \exp\left(-\frac{t^2}{2a \cdot \mathbb{E}[Z] + 2b + 2ct}\right),$$

where  $c = \max\{0, (3a - 1)/6\}$ .

Note that for  $a = L > \frac{1}{3}$ , the additional  $2ct$  term in the denominator causes the upper-tail bound to decay only as a simple exponential, and leads to significantly weaker concentration. Whether the upper-tail bound of self-bounding functions can be improved to remove this term is an open problem in probability theory, with the first bound appearing in the work of Boucheron et al. [13], and improved bounds due to McDiarmid and Reed [24] and Boucheron et al. [14]. Successfully removing the  $2ct$  term from the denominator would improve the result stated in Theorem 2.1 from  $O\left(\sqrt{|\text{opt}(G)| \ln(k/\delta)}\right)$  to  $O\left(\sqrt{p_i |\text{opt}(G)| \ln(k/\delta)}\right)$ , and Corollary 3.1 from  $O\left(k \sqrt{|\text{opt}(G)| \ln(k/\delta)}\right)$  to  $O\left(\sqrt{k |\text{opt}(G)| \ln(k/\delta)}\right)$ .

## B Additional Examples

In this section we present two examples that are referenced in the body of the paper.

**B.1 The Case of Large  $p$**  We provide examples where the conclusion of Theorem 2.1 is violated under  $p_i > 1/2$ . This happens because the examples violate Lemma 2.1, that is, they satisfy

$$\mathbb{E}_{H \sim_p G}[|\text{opt}(H)|] > \mathbb{E}_{H \sim_p G}[|\text{opt}(G) \upharpoonright H|] = p \cdot |\text{opt}(G)|.$$

First, suppose that only 2-cycles are allowed. Consider a graph with three vertices  $v_1, v_2, v_3$ , and 2-cycles between  $v_1$  and  $v_2$ ,  $v_2$  and  $v_3$ , and  $v_3$  and  $v_1$ . Suppose  $p = 2/3$ . Then  $p|\text{opt}(G)| = 2 \cdot (2/3)$ . On the other hand,  $|\text{opt}(H)| = 2$  if  $H$  contains at least two vertices (otherwise it is 0), hence

$$\begin{aligned} \mathbb{E}_{H \sim_p G}[|\text{opt}(H)|] &= 2 \left( \binom{3}{2} p^2 (1-p) + \binom{3}{3} p^3 \right) \\ &= 2 \cdot \frac{20}{27} > 2 \cdot \frac{2}{3}. \end{aligned}$$

Now consider a graph that contains many disjoint copies of the one just discussed. We have that both  $\text{opt}(H)$  and  $\text{opt}(G) \upharpoonright H$  are concentrated around their expectations (by Hoeffding’s inequality), so, with high probability,  $|\text{opt}(H)| > |\text{opt}(G) \upharpoonright H| + \Omega(n)$ .

When 3-cycles are allowed too, it is possible to show that the same phenomenon happens, for a value of  $p$  sufficiently close to 1, in a graph with five vertices  $v_1, \dots, v_5$ , and 2-cycles between  $v_i$  and  $v_{i+1}$  for  $i = 1, \dots, 4$ , as well as between  $v_5$  and  $v_1$ .

**B.2 The Case of Long Cycles** We construct an example where long cycles are allowed, and every individually rational matching is smaller than the optimal matching by  $\Omega(n)$ . Motivated by kidney exchange, the example includes an altruistic donor, and matchings may include chains initiated by the altruist. However, we can easily transform the example into one where matchings can only include cycles, by adding directed edges from every vertex in the graph to the altruistic donor.

Consider the graph  $G$  in Figure 4. The altruistic donor  $d$  is shown as a triangle (we do not count him as one of the  $n$  vertices). The vertices consist of (i) a chain of  $\frac{3n}{9}$  vertices, and (ii) a network of  $\frac{2n}{9}$  layers—shown as dashed ellipses—with each layer consisting of three vertices. All vertices in layer  $i$  have edges to all vertices in layers  $(i+1), \dots, \frac{2n}{9}$ , and there are edges from  $d$  to all vertices in each layer. Observe that since there are no cycles in this graph, all matches must happen only via a chain that originates in  $d$ . The longest chain consists of  $\frac{3n}{9}$  vertices. Thus, there is a unique optimal matching in  $G$ , and its size is  $|\text{opt}(G)| = \frac{3n}{9}$ .

Let us now assume there are two players with probability  $p = \frac{1}{2}$  each. We first observe that the

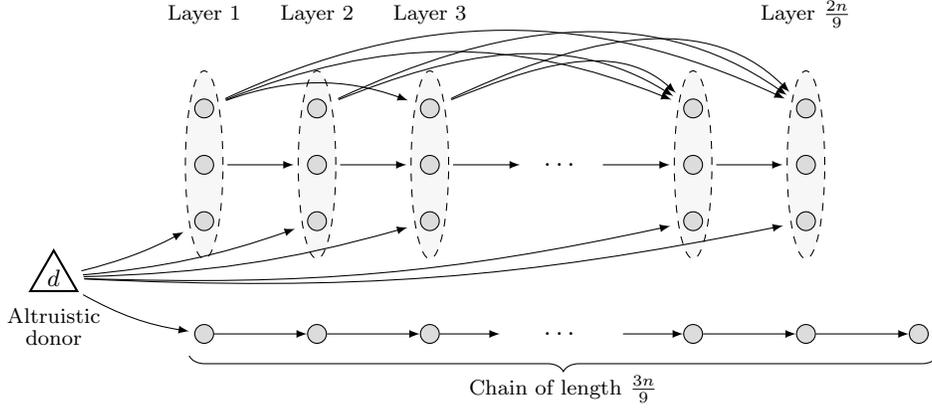


Figure 4: The graph constructed in the example of Appendix B.2.

expected share of a subgraph  $H \sim_p G$  in the optimal matching is  $\mathbb{E}_{H \sim_p G}[|\text{opt}(G) \upharpoonright H|] = \frac{3n}{18}$ .

Next we examine  $\text{opt}(H)$ . We can assume that  $d \in H$ , as this is always true for one of the two players (so we focus on that player without loss of generality). For any given assignment of the other vertices to the players, we say that a layer is *good* if at least one of the vertices in that layer is in  $H$ . It is easy to see that—under the assumption of  $d \in H$ — $\text{opt}(H)$  is at least the number of good layers (via a chain that starts at  $d$  and visits each good layer in order). Notice that each layer is good with probability  $1 - (1/2)^3 = 7/8$ . Therefore, the expected number of layers that are good is  $\frac{7}{8} \cdot \frac{2n}{9}$ . It follows that  $\mathbb{E}_{H \sim_p} [|\text{opt}(H)|] \geq \frac{14}{72}n$ .

Since both  $|\text{opt}(H)|$  and  $|\text{opt}(G) \upharpoonright H|$  are almost always within  $O(\sqrt{n})$  of their expected values, we have that with high probability,  $|\text{opt}(H)| - |\text{opt}(G) \upharpoonright H| \geq \frac{2}{72}n - O(\sqrt{n})$ . That is,  $\text{opt}(G)$  is not individually rational, and an individually rational matching would have to use  $d$  to initiate a chain into the layered network. But such a chain can have length at most  $\frac{2n}{9}$ , whereas  $\text{opt}(G) = \frac{3n}{9}$ —the difference is  $\Omega(n)$ , as desired.