An Improved 2-Agent Kidney Exchange Mechanism

Ioannis Caragiannis[1], Aris Filos-Ratsikas[1], and Ariel D. Procaccia[2,⋆]

[1] Department of Computer Engineering and Informatics, University of Patras, 26504 Rio, Greece
[2] Computer Science Department, Carnegie Mellon University, Pittsburgh, PA 15213, USA

Abstract. We study a mechanism design version of matching computation in graphs that models the game played by hospitals participating in pairwise kidney exchange programs. We present a new randomized matching mechanism for two agents which is truthful in expectation and has an approximation ratio of 3/2 to the maximum cardinality matching. This is an improvement over a recent upper bound of 2 [Ashlagi et al., EC 2010] and, furthermore, our mechanism beats for the first time the lower bound on the approximation ratio of deterministic truthful mechanisms. We complement our positive result with new lower bounds. Among other statements, we prove that the weaker incentive compatibility property of truthfulness in expectation in our mechanism is necessary; universally truthful mechanisms that have an inclusion-maximality property have an approximation ratio of at least 2.

1 Introduction

In an attempt to address the wide need for kidney transplantation and the scarcity of cadaver kidneys, several countries have launched, or are considering, national kidney exchange programs involving live donors [7,11,1,4]. Patients can enter such a program together with a member of their family or friend who is willing to donate them a kidney but cannot due to incompatibility. National kidney exchange programs aim to implement exchanges between two compatible patient-donor pairs u and v so that the donor of pair u donates her kidney to the patient of pair v and vice versa. This requires four simultaneous operations. More complicated exchanges involving more than two donor-patient pairs are also possible; however, we focus on pairwise exchanges since they are easier to perform in practice.

Donor-patient pairs approach a hospital in order to enroll into the national kidney exchange programs. In an ideal scenario, each hospital reports its donor-patient pairs to the program and a central authority runs an algorithm that decides which pairwise kidney exchanges will take place. In practice, strategic issues immediately arise. A hospital may prefer to not enroll some easy-to-match

⋆ The third author is supported by gifts from Yahoo! and Google.

N. Chen, E. Elkind, and E. Koutsoupias (Eds.): WINE 2011, LNCS 7090, pp. 37–48, 2011.

donor-patient pairs to the program and instead match them and perform the kidney exchange operations internally. This may have an impact on patients of other hospitals who could have benefited if the hospital truthfully reported all its donor-patient pairs to the program. The current paper follows the line of research that seeks to design algorithms (or mechanisms) that discourage hospitals from behaving untruthfully. The main objective is to perform as many kidney exchanges as possible under this constraint. This is a *mechanism design* [5] problem, and in particular—because paying for organs is illegal in almost all countries—it falls within the scope of approximate mechanism design without money [6].

We can model the problem as a matching problem in graphs. The input consists of a graph in which the nodes represent donor-patient pairs and an edge connects two nodes u and v when the donor of pair u and the patient of pair v are compatible, and the donor of pair v and the patient of pair u are compatible. Each node of the graph is controlled by exactly one self-interested agent (a hospital). A *mechanism* takes the graph as input and returns a matching, i.e., a disjoint pair of edges indicating which pairwise kidney exchanges will take place. The *gain* of an agent is the number of nodes under her control that are matched. Clearly, an optimal solution is easy to find by a maximum matching computation. Unfortunately, a mechanism that returns such a solution may incentivize hospitals to behave untruthfully in the following sense. A hospital could hide some of its nodes from (i.e., not enroll them into) the system so that the mechanism is essentially applied on a graph that contains neither the hidden nodes nor the edges incident to them. Then, the gain of the hospital is the number of its nodes that are matched by the mechanism plus the number of nodes it managed to match internally. Such behavior can lead to fewer matched nodes compared to the best possible solution, i.e., fewer patients who receive kidneys. So, we seek mechanisms that guarantee that no agent has any incentive to deviate from truth-telling. Our goal is to design such mechanisms that also return matchings of high cardinality, i.e., high total gain.

The mechanisms can be deterministic or randomized. Given an instance of the problem, a deterministic mechanism returns a simple matching. A randomized mechanism returns a probability distribution over matchings. In the latter case, we distinguish between *universally truthful* mechanisms and mechanisms that are *truthful in expectation*. The former are induced by a probability distribution over truthful deterministic mechanisms, whereas the latter guarantee that no agent can deviate from truth-telling in order to increase her expected gain. The efficiency of truthful mechanisms is assessed through their *approximation ratio*, i.e., the maximum ratio over all possible instances of the problem of the size of the maximum cardinality matching over the expected size of the matching returned by the mechanism.

Early work on kidney exchange problems in Economics [8,9,10] has considered the incentives of incompatible donor-patient pairs. However, as national kidney exchange programs emerged, it has become apparent that such incentives are less important compared to the incentives of the hospitals [3]. The model considered in the current paper has also been studied in [2,3,12,13]. The fact that the maximum cardinality matching mechanism is not truthful was first observed by Sönmez and Ünver [12] (see also [3]). Ashlagi et al. [2] present a universally truthful randomized 2-approximation mechanism (called Mix-and-Match) for arbitrarily many agents. Mix-and-Match is based on a simple deterministic truthful 2-approximation mechanism for two agents, henceforth called Match. Match returns a matching that contains the maximum number of *internal edges* (where the nodes on both sides are controlled by the same agent), breaking ties in favor of the matching with maximum cardinality. A nice property of Match is *inclusion-maximality*; this translates to the requirement that a donor-patient pair does not participate in any kidney exchange only when all its compatible donor-patient pairs participate in some pairwise kidney exchange. A randomized mechanism has this property when it returns a probability distribution over inclusion-maximal matchings. On the negative side, there are lower bounds of 2 and 8/7 for deterministic truthful mechanisms and randomized mechanisms that are truthful in expectation, respectively [2,3]. Ashlagi et al. [2] also propose the mechanism Flip-and-Match for two agents. Flip-and-Match equiprobably selects among the outcome of Match and a maximum cardinality matching. They prove that this mechanism has approximation ratio 4/3 and leave open the question of whether it is truthful in expectation. Ashlagi and Roth [3] and Toulis and Parkes [13] consider weaker notions of truthfulness in random graph models that reflect the compatibility frequency among donors and patients from the human population. As in [2], no such information is required in our setting.

In an attempt to better understand the potential and limitations of randomized mechanisms, we consider the case of two agents. This case is of special interest because efficient mechanisms can enable cooperation between pairs of hospitals on an *ad-hoc* basis, in countries where a national kidney exchange program is not yet in place. Our main result is a randomized mechanism called Weight-and-Match for 2-agent pairwise kidney exchange that is truthful in expectation and has a tight approximation ratio of 3/2. This establishes, for the first time, a separation between the power of randomized mechanisms and deterministic mechanisms (for which there is a lower bound of 2).

Weight-and-Match is inspired by the mechanism Flip-and-Match proposed in [2]. Unfortunately, it turns out that Flip-and-Match is not truthful due to its use of maximum cardinality matchings. This observation is our starting point for the definition of the new mechanism. Weight-and-Match first assigns weights to the edges of the input graph and then selects equiprobably among two maximum-weight matchings: one with minimum cardinality (the particular weights assigned to the edges guarantee that this matching is identical to the one returned by Match) and one with maximum cardinality (which replaces the second matching used by Flip-and-Match). Informally, this definition guarantees that the bad incentives created by the second matching are canceled out by the outcome of Match.

We complement this result with new lower bounds on the approximation ratio of randomized mechanisms that are truthful in expectation or universally truthful, distinguishing between mechanisms that are inclusion-maximal and those that are not. Here we use the same 2-agent instance as in previous work [2,3,12] but our stronger analysis leads to improved bounds for inclusion-maximal and universally truthful mechanisms. Our general lower bound is 5/4. Interestingly, we prove a lower bound of 2 for inclusion-maximal universally truthful mechanisms that indicates that the weaker notion of truthfulness satisfied by Weight-and-Match (which is inclusion-maximal) is indeed necessary.

The rest of the paper is structured as follows. We warm up by showing that Flip-and-Match is not truthful in Section 2. Our mechanism and its analysis are presented in Section 3. The lower bounds are presented in Section 4. We conclude with a short discussion of open problems in Section 5.

2 An Unsuccessful Attempt: Flip-and-Match

Throughout the paper, we refer to the two agents as agent 1 and agent 2. We also call the nodes of agents 1 and 2 *white* and *gray* nodes, respectively.

Let us warm up by considering the mechanism Flip-and-Match proposed in [2]. Flip-and-Match selects equiprobably among the matching returned by Match and a maximum cardinality matching. In the original definition of [2], ties among maximum cardinality matchings are broken in favor of matchings that maximize the number of internal edges (i.e., edges between two nodes controlled by the same agent) and then arbitrarily. In our proof, we essentially show that any modification of the tie-breaking rule violates truthfulness.

Theorem 1. Flip-and-Match *is not truthful.*

Proof. Our proof uses the instance I and subinstances I_1 and I_2 of Figure 1. When applied to instance I, Match returns the matching $M_1 = \{(v_2, v_3), (v_4, v_5), (v_7, v_8)\}$. The gain of agent 1 is 4 while the gain of agent 2 is 2. Let M_2 be a maximum cardinality matching. It leaves exactly one node unmatched; this can be either a white or a gray node, i.e., M_2 matches either 4 white nodes and 4 gray nodes or 5 white nodes and 3 gray nodes. We distinguish between these two cases and show that, in both cases, some agent has an incentive to withhold nodes.

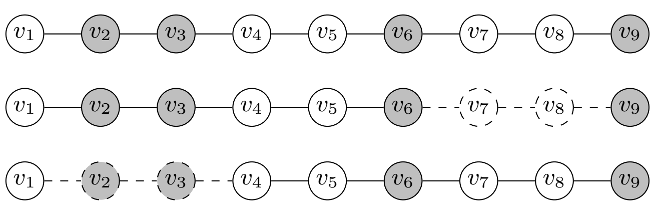

Fig. 1. The original instance I used in the proof of Lemma 1 and the two subinstances I_1 and I_2 used in cases 1 and 2, respectively. The dashed nodes and edges are not part of the instances I_1 and I_2 but are shown here in order to compare with instance I.

Case 1. M_2 matches 4 white nodes and 4 gray nodes and, hence, the expected gain of agent 1 from the application of FLIP-AND-MATCH on instance I is 4. Consider the instance I_1 in which agent 1 hides the white nodes v_7 and v_8 and matches them internally. In the new instance, MATCH returns the matching $\{(v_2, v_3), (v_4, v_5)\}$ that contains 2 matched white nodes while the maximum cardinality matching is $\{(v_1, v_2), (v_3, v_4), (v_5, v_6)\}$ that contains 3 matched white nodes. The expected gain of agent 1 (including the hidden nodes) is 4.5.

Case 2. M_2 matches 5 white nodes and 3 gray nodes and hence the expected gain of agent 2 from the application of FLIP-AND-MATCH to the original instance I is 2.5. Consider the instance I_2 in which agent 2 hides nodes v_2 and v_3 (and matches them internally). In the new instance, MATCH returns the matching $\{(v_4, v_5), (v_7, v_8)\}$ that contains no matched gray nodes while the maximum cardinality matching is $\{(v_4, v_5), (v_6, v_7), (v_8, v_9)\}$ that contains 2 matched gray nodes. The expected gain of agent 2 (including the hidden nodes) is 3. □

3 Our Mechanism: WEIGHT-AND-MATCH

In this section, we present our new mechanism for two agents, which we call WEIGHT-AND-MATCH. The main idea behind it is similar to the one that led to FLIP-AND-MATCH: we try to combine mechanism MATCH with another mechanism that yields a higher gain. However, given the negative result for FLIP-AND-MATCH presented in the previous section, we should be careful with the definition of our mechanism. We can think of the following alternative definition for MATCH. We first assign weights to the edges of the input graph as follows. Internal edges have weight 1; edges between nodes of different agents have weight 1/2. The matching returned by MATCH is then a maximum-weight matching on the weighted version of the input graph, where ties are broken in favor of the matching with minimum cardinality. Our mechanism WEIGHT-AND-MATCH also computes a maximum-weight maximum-cardinality matching on the weighted version of the input graph, and selects equiprobably among the two matchings. Note that WEIGHT-AND-MATCH is inclusion maximal. The rest of the section is devoted to proving the following statement.

Theorem 2. *Mechanism* WEIGHT-AND-MATCH *can be implemented in polynomial time, has approximation ratio* 3/2*, and is truthful in expectation.*

Due to lack of space, we omit the proof that our mechanism can be implemented efficiently; we proceed with the proof of its approximation guarantee.

Lemma 1. WEIGHT-AND-MATCH *has an approximation ratio of* 3/2*.*

Proof. Let M be a matching of maximum cardinality and let M_1 and M_2 be the maximum-weight matchings of minimum and maximum cardinality, respectively, that are used by WEIGHT-AND-MATCH. Consider the symmetric difference $M \Delta M_1 = (M \setminus M_1) \cup (M_1 \setminus M)$. It consists of several connected components which are either cycles (of even length), or paths with edges alternating between edges of M and edges of M_1. Let m_1 be the number of edges of M that either belong also to M_1 or belong to cycles or paths of $M \Delta M_1$ with even length. Let m_3 and m_5 be the edges of M that belong to paths of $M \Delta M_1$ with length exactly 3 and odd length at least 5, respectively. Clearly, $|M| = m_1 + m_3 + m_5$.

Note that the number of edges of M_1 that either belong also to M or belong to cycles or paths of $M \Delta M_1$ of even length is exactly m_1 as well. Also, since M has maximum cardinality, the first and the last edge in a path with odd length in $M \Delta M_1$ belong to M. So, M_1 contains exactly $m_3/2$ edges in paths of $M \Delta M_1$ of length 3 and at least $2m_5/3$ edges in paths of $M \Delta M_1$ of odd length at least 5. Hence, $|M_1| \geq m_1 + m_3/2 + 2m_5/3$.

We now show that M_2 (the maximum-weight matching of maximum cardinality) contains at least $m_1 + m_3 + 2m_5/3$ edges. Observe that, since M_1 is a maximum-weight matching, in any path with length 3 in $M \Delta M_1$, the edge of M_1 should have endpoints belonging to the same agent (and, hence, weight 1) and the two edges of M should have endpoints belonging to different agents (and, hence, weight 1/2). Consider the edges of M_1 that do not belong to paths of length 3 of $M \Delta M_1$ and the edges of M that belong to paths of length 3 in $M \Delta M_1$. All these edges form a matching that has the same total weight as the edges of M_1, and their cardinality is at least $m_1 + m_3 + 2m_5/3$. Clearly, this is also a lower bound on the cardinality of M_2, i.e., $|M_2| \geq m_1 + m_3 + 2m_5/3$.

Hence, the expected cardinality of the mechanism's matching is

$$\frac{1}{2}(|M_1| + |M_2|) \geq m_1 + \frac{3m_3}{4} + \frac{2m_5}{3} \geq \frac{2}{3}(m_1 + m_3 + m_5) = \frac{2}{3}|M|. \quad \square$$

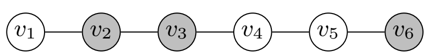

Fig. 2. An instance indicating that the analysis of Lemma 1 is tight. The maximum matching matches all 6 nodes but mechanism WEIGHT-AND-MATCH returns the matching that consists of edges (v_2, v_3) and (v_4, v_5). Note that here the symmetric difference is a path of length 5.

The bound obtained in Lemma 1 is tight through the example of Figure 2. We now turn to proving that our mechanism is truthful.

Lemma 2. WEIGHT-AND-MATCH *is truthful in expectation.*

Proof. We will show that agent 1 never has an incentive to deviate from truth-telling. The case of agent 2 is identical.

Let G be the input graph and consider the maximum-weight matchings M_1 and M_2 of minimum and maximum cardinality, respectively, that are used by WEIGHT-AND-MATCH. Also, assume that agent 1 hides some nodes and matches them internally. Then, the mechanism is applied to the subgraph G' of G which does not contain the hidden white nodes and edges incident to them. Let M_3 and

M_4 be the maximum-weight matchings of minimum and maximum cardinality computed by Weight-and-Match on input G', augmented by the edges used by agent 1 to match the hidden white nodes internally. Denote by $\texttt{gain}(M)$ the gain of agent 1 from matching M and by $\texttt{wgt}(M)$ the weight of matching M. Our proof will follow from the next two lemmas.

Lemma 3. $\texttt{gain}(M_3) = \texttt{gain}(M_1) - 2(\texttt{wgt}(M_1) - \texttt{wgt}(M_3))$.

Proof. Denote by $n_{\texttt{ww}}(M)$, $n_{\texttt{wg}}(M)$, and $n_{\texttt{gg}}(M)$ the number of edges in matching M connecting two white nodes, two nodes belonging to different agents, and two gray nodes, respectively. We will first show that $n_{\texttt{gg}}(M_1) = n_{\texttt{gg}}(M_3)$. Consider the symmetric difference of the two matchings $M_1 \Delta M_3 = (M_1 \setminus M_3) \cup (M_3 \setminus M_1)$ and the subgraph of G induced by these edges. This subgraph consists of several connected components which can be cycles or paths (see Figure 3 for an example). Consider such a connected component C and let C_1 and C_3 be the sets of edges of M_1 and M_3 it contains, respectively.

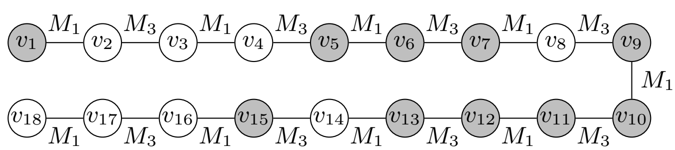

Fig. 3. A connected component of $M_1 \Delta M_3$ considered in the proof of Lemma 3. The sets of gray nodes $\{v_1\}$, $\{v_5, v_6, v_7\}$, $\{v_9, ..., v_{13}\}$, and $\{v_{15}\}$ form blocks. The main argument in the proof is that each block has an odd number of gray nodes.

In order to prove that $n_{\texttt{gg}}(M_1) = n_{\texttt{gg}}(M_3)$, it suffices to prove that $n_{\texttt{gg}}(C_1) = n_{\texttt{gg}}(C_3)$. This is clearly true if C is a cycle consisting of gray nodes only, since such a cycle should have an even number of edges, half of which belong to C_1 and half to C_3. Assume that C contains a block of t consecutive gray nodes $b_1, b_2, ..., b_t$ such that the first and the last have either degree 1 or are connected to another white node outside the block. We will show that t cannot be even. Assume that this was the case; then one of the two matchings (say M_1; the argument for M_3 is completely symmetric) would contain the $\frac{t}{2} - 1$ edges (b_2, b_3), $(b_4, b_5), \ldots, (b_{t-2}, b_{t-1})$ and the other (say M_3) would contain the $\frac{t}{2}$ edges (b_1, b_2), $(b_3, b_4), \ldots, (b_{t-1}, b_t)$. Then, by replacing the $\frac{t}{2} - 1$ edges of matching M_1 in the block as well as the edges of M_1 that are incident to nodes b_1 and b_t (if any) with the $\frac{t}{2}$ edges of M_3 in the block, we would obtain a matching that either has higher weight than M_1 (if some of nodes b_1 and b_t has degree 1) or the same weight as M_1 (recall that the edges connecting nodes b_1 and b_t to white nodes outside the block have weight 1/2) but smaller cardinality. Both cases contradict the fact that the matching M_1 is a minimum cardinality maximum-weight matching. Hence, every block has an odd number of nodes and an even number of edges between gray nodes that alternate between matchings M_1 and M_3. This implies that $n_{\texttt{gg}}(C_1) = n_{\texttt{gg}}(C_3)$. Consequently, by summing over all connected components of $M_1 \Delta M_3$ and the edges of $M_1 \cap M_3$ connecting gray nodes, we also have that $n_{\mathtt{gg}}(M_1) = n_{\mathtt{gg}}(M_3)$.

Next, observe that $\mathtt{gain}(M) = 2n_{\mathtt{ww}}(M) + n_{\mathtt{wg}}(M)$ and $\mathtt{wgt}(M) = n_{\mathtt{ww}}(M) + n_{\mathtt{gg}}(M) + n_{\mathtt{wg}}(M)/2$. Hence, since $n_{\mathtt{gg}}(M_1) = n_{\mathtt{gg}}(M_3)$, we have

$$\begin{aligned}\mathtt{gain}(M_3) &= 2n_{\mathtt{ww}}(M_3) + n_{\mathtt{wg}}(M_3)\\ &= 2n_{\mathtt{ww}}(M_3) + n_{\mathtt{wg}}(M_3) + 2n_{\mathtt{gg}}(M_3) - 2n_{\mathtt{gg}}(M_1)\\ &= \mathtt{gain}(M_1) - 2(\mathtt{wgt}(M_1) - \mathtt{wgt}(M_3)),\end{aligned}$$

as desired. □

Lemma 4. $\mathtt{gain}(M_4) \leq \mathtt{gain}(M_2) + 2(\mathtt{wgt}(M_2) - \mathtt{wgt}(M_4))$.

Proof. First consider each edge in $M_2 \cap M_4$ and observe that its contribution to $\mathtt{gain}(M_4)$ equals its contribution to $\mathtt{gain}(M_2) + 2(\mathtt{wgt}(M_2) - \mathtt{wgt}(M_4))$. We will now consider the symmetric difference of the two matchings $M_2 \Delta M_4 = (M_2 \setminus M_4) \cup (M_4 \setminus M_2)$ and the subgraph of G induced by these edges. Again, this subgraph consists of several connected components which can be cycles or paths. Consider such a connected component C and let C_2 and C_4 be the sets of edges of M_2 and M_4 it contains, respectively. We will complete the proof of the lemma by showing that

$$\mathtt{gain}(C_4) \leq \mathtt{gain}(C_2) + 2(\mathtt{wgt}(C_2) - \mathtt{wgt}(C_4). \quad (1)$$

First, observe that since M_2 is a maximum-weight matching in G, it holds that $\mathtt{wgt}(C_2) \geq \mathtt{wgt}(C_4)$ (otherwise, we could replace the edges of C_2 with the edges of C_4 in M_2 and obtain a matching with higher weight). We now use a four-letter/number notation to classify the connected components of the subgraph of G induced by $M_2 \Delta M_4$ that are paths into different types: the first and last letters are `w` or `g` and denote whether the left and right endpoint of the connected component is a white or gray node, respectively. The second and third numbers are either `2` or `4` and denote whether the first and the last edge of the connected component belong to matching M_2 or M_4, respectively. Examples of paths of type `w22w`, `w44g`, and `w44w` are depicted in Figure 4. We distinguish between three main cases:

Case 1. If C is a cycle, or a path of type `w22w`, `w24w`, `w42w`, `w22g`, `w24g`, `g22g`, `g24g`, `g42g`, or `g44g`, we have $\mathtt{gain}(C_4) \leq \mathtt{gain}(C_2)$ and inequality (1) follows easily since $\mathtt{wgt}(C_2) \geq \mathtt{wgt}(C_4)$.

Case 2. If C is a path of type `w42g` or `w44g`, we claim that $\mathtt{wgt}(C_2) + \mathtt{wgt}(C_4)$ is non-integer. Indeed, since the first and the last node in the path belong to different agents, there is an odd number of external edges (between a white and a gray node) in C, and each such edge contributes 1/2 to the sum $\mathtt{wgt}(C_2) + \mathtt{wgt}(C_4)$. Recall that $\mathtt{wgt}(C_2) \geq \mathtt{wgt}(C_4)$, and therefore $\mathtt{wgt}(C_2) - \mathtt{wgt}(C_4) \geq 1/2$. Inequality (1) follows by observing that $\mathtt{gain}(C_2) = \mathtt{gain}(C_4) - 1$ in this case.

Case 3. If C is of type `w44w`, observe that C_4 contains one more edge than C_2 and, hence, $\mathtt{wgt}(C_2) > \mathtt{wgt}(C_4)$ (otherwise, we could replace the edges of C_2 with

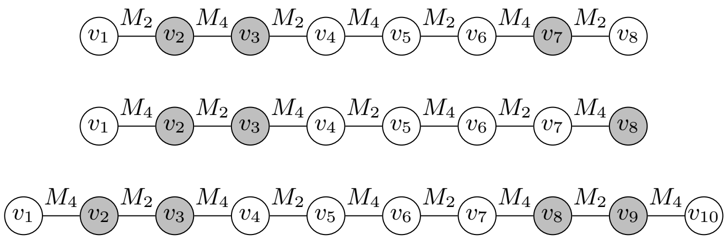

Fig. 4. Examples of connected components of $M_2 \Delta M_4$ considered in the proof of Lemma 4 (paths of type `w22w`, `w44g`, and `w44w`)

the edges of C_4 in M_2 in order to obtain a matching of the same weight but with higher cardinality). Also, observe that the number of external edges in C is even, and hence $\mathtt{wgt}(C_2) + \mathtt{wgt}(C_4)$ is integer. It follows that $\mathtt{wgt}(C_2) \geq \mathtt{wgt}(C_4) + 1$. Inequality (1) follows by further observing that $\mathtt{gain}(C_2) = \mathtt{gain}(C_4) - 2$. □

Since $\mathtt{wgt}(M_1) = \mathtt{wgt}(M_2)$ and $\mathtt{wgt}(M_3) = \mathtt{wgt}(M_4)$, by Lemmas 3 and 4 we have that the expected gain $\frac{1}{2}(\mathtt{gain}(M_3) + \mathtt{gain}(M_4))$ of agent 1 when she hides some white nodes and matches them internally is upper-bounded by the expected gain $\frac{1}{2}(\mathtt{gain}(M_1) + \mathtt{gain}(M_2))$ when she acts truthfully. □

4 Lower Bounds

Ashlagi et al. [2] and Ashlagi and Roth [3] provide a lower bound of 8/7 for truthful-in-expectation randomized mechanisms.[1] The proof of the next lemma starts with the same initial instance as [2,3] but uses a more detailed reasoning in order to prove lower bounds for randomized mechanisms that are either universally truthful or truthful in expectation, distinguishing between mechanisms that are inclusion-maximal and those that are not.

Theorem 3. *Let A be a randomized mechanism for 2-agent kidney exchange.*

(a) If A is truthful in expectation, then its approximation ratio is at least $5/4$.
(b) If A is truthful in expectation and inclusion-maximal, then its approximation ratio is at least $4/3$.
(c) If A is universally truthful, then its approximation ratio is at least $3/2$.
(d) If A is universally truthful and inclusion-maximal, then its approximation ratio is at least 2.

Proof. Our proof uses the instances depicted in Figure 5. The starting point is instance I. We denote by I_1 the instance obtained by removing the white nodes

[1] Ashlagi et al. [2] actually claim a bound of 4/3 but this is inaccurate. In fact it is not hard to design an artificial mechanism (as a probability distribution over matchings) that is truthful in expectation and has approximation ratio at most 5/4 for the instances considered in their proof.

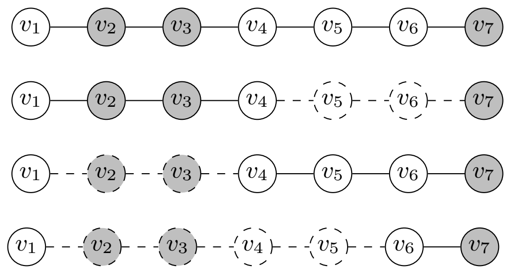

Fig. 5. The instances I, I_1, I_2, and I_3 used in the proof of Theorem 3. The dashed nodes and edges are not part of the instances I_1, I_2, and I_3 but are shown here in order to compare with instance I.

v_5 and v_6 and their incident edges from I, by I_2 the instance obtained from I by removing the nodes v_2 and v_3 and their incident edges, and by I_3 the instance obtained from I_2 by removing the nodes v_4 and v_5 and their incident edges.

(a) Consider the application of mechanism A to instance I. Observe that the maximum cardinality matching of this instance has size 3, i.e., the total gain of both agents from any matching is at most 6. So, assume that the expected gain of agents 1 and 2 from the matching returned by A is at most $4-u$ and at most $2+u$ respectively, for some $u \in [0,1]$. Then, consider the application of mechanism A to instance I_1. The expected gain of agent 1 from the matching returned by A should be at most $2-u$ (otherwise, in the original instance I, agent 1 would have an incentive to hide the white nodes v_5 and v_6 and match them internally). This means that, on input I_1, the probability that A returns a matching consisting of two edges is at most $1-u/2$. Hence, the approximation ratio of mechanism A on instance I_1 is at least $\frac{4}{4-u}$.

Also, consider the application of mechanism A to instance I_2. The expected gain of agent 2 from the matching returned by A should be at most u (otherwise, in the original instance I, agent 2 would have an incentive to hide the gray nodes v_2 and v_3 and match them internally). This means that, on input I_2, the probability that A returns a matching consisting of two edges is at most u. Hence, the expected gain of agent 1 from instance I_2 is at most $2+u$. Now, consider the application of A to instance I_3; A should return a non-empty matching with probability at most u (otherwise, agent 1 would have an incentive to hide nodes v_4 and v_5 from instance I_2 and match them internally). Hence, the approximation ratio of mechanism A on instance I_3 would be $1/u$.

We conclude that the approximation ratio of A is at least $\max\left\{\frac{4}{4-u}, \frac{1}{u}\right\}$ which is minimized to $5/4$ for $u = 4/5$.

(b) From the analysis of (a), we have that A is inclusion-maximal only when $u = 1$ (otherwise, it would return an empty matching for instance I_3 with non-zero probability). In this case, the approximation ratio of A is at least $4/3$.

(c) Since A is universally truthful, it uses a probability distribution over deterministic truthful mechanisms. We partition the set of truthful deterministic

mechanisms into two sets $\mathcal{A}_{\mathtt{w}}$ and $\mathcal{A}_{\mathtt{g}}$: the set $\mathcal{A}_{\mathtt{w}}$ (respectively, $\mathcal{A}_{\mathtt{g}}$) consists of mechanisms which, on input instance I, return a matching that leaves at least one white node (respectively, at least one gray node) unmatched. Any other truthful deterministic mechanism is arbitrarily put in one of the two sets.

Let $A_{\mathtt{w}}$ be a deterministic mechanism that belongs to $\mathcal{A}_{\mathtt{w}}$. On input instance I_1, $A_{\mathtt{w}}$ should return a matching with just one edge. Otherwise, a matching with two edges would match the two white nodes v_1 and v_4 which means that agent 1 would have an incentive to hide nodes v_5 and v_6 from instance I and match them internally; this would violate the truthfulness of mechanism $A_{\mathtt{w}}$. Hence, mechanism $A_{\mathtt{w}}$ returns matchings of size at most 1 on input instances I_1 and I_3.

Also, let $A_{\mathtt{g}}$ be a deterministic mechanism that belongs to $\mathcal{A}_{\mathtt{g}}$. Consider the application of $A_{\mathtt{g}}$ to instance I_2. The matching it returns should not match node v_7 since otherwise agent 2 would have an incentive to hide nodes v_2 and v_3 in the original instance I and match them internally. Hence, only two white nodes are matched by mechanism $A_{\mathtt{g}}$ on input instance I_2. Now consider the application of $A_{\mathtt{g}}$ to the instance I_3. It should return an empty matching otherwise agent 1 would have an incentive to hide the white nodes v_4 and v_5 from instance I_2 and match them internally. Hence, the matchings returned by mechanism $A_{\mathtt{g}}$ on input instances I_2 and I_3 have size at most 2 and 0, respectively.

Next, let p be the probability that mechanism A runs a deterministic truthful mechanism from $\mathcal{A}_{\mathtt{w}}$. Then, the expected size of the matching returned by A on input instances I_1 and I_3 is at most $2-p$ and p, respectively, and its approximation ratio is at least $\max\left\{\frac{2}{2-p}, \frac{1}{p}\right\}$ which is minimized to $3/2$ for $p = 2/3$.

(d) In the proof of (c), the mechanisms in $\mathcal{A}_{\mathtt{g}}$ are not inclusion-maximal. Hence, if A is universally truthful and inclusion-maximal, it should use only deterministic mechanisms from $\mathcal{A}_{\mathtt{w}}$, i.e., $p = 1$. Following the analysis in the previous case for instance I_1, we obtain that A has approximation ratio at least 2. □

Theorems 2 and 3(d) establish a separation between truthfulness in expectation and universal truthfulness with respect to inclusion-maximal mechanisms.

5 Discussion and Open Problems

Our work has shed some light on the efficiency of randomized truthful mechanisms for the 2-agent pairwise kidney exchange problem. Although the number of agents is restricted, we believe that this case is of special interest because 2-agent mechanisms can enable *ad-hoc* arrangements between hospitals in countries where national exchanges are not in place.

Clearly, the question of whether the upper bound of 2 of Ashlagi et al. [2] can be improved for instances with arbitrarily many agents remains wide open. Unfortunately, several extensions of Weight-and-Match that we have considered for this case have failed, and in fact it seems likely that this upper bound is tight for more than two agents. Still, the 2-agent case deserves some further investigation because there are gaps between our upper and lower bounds.

In this context, it is especially interesting to know whether a truthful in expectation, inclusion-maximal, 4/3-approximation mechanism exists. For the 2-agent case, we also believe that characterizations of truthful mechanisms would be very useful in order to complete the picture. Finally, Ashlagi et al. [2] were unable to provide a truthful deterministic mechanism for the case of more than two agents that gives any nontrivial approximation ratio. Providing such a mechanism, or proving a lower bound, remains an enigmatic open problem.

References

1. Abraham, D., Blum, A., Sandholm, T.: Clearing algorithms for barter exchange markets: enabling nationwide kidney exchanges. In: Proc. of the 8th ACM Conference on Electronic Commerce (EC), pp. 295–304 (2007)
2. Ashlagi, I., Fischer, F., Kash, I.A., Procaccia, A.D.: Mix and match. In: Proc. of the 11th ACM Conference on Electronic Commerce (EC), pp. 305–314 (2010)
3. Ashlagi, I., Roth, A.E.: Individual rationality and participation in large scale, multi-hospital kidney exchange. In: Proc. of the 12th ACM Conference on Electronic Commerce (EC), p. 321 (2011), http://web.mit.edu/iashlagi/www/papers/LargeScaleKidneyExchange_1_13.pdf
4. Biró, P., Manlove, D.F., Rizzi, R.: Maximum weight cycle packing in directed graphs, with application to kidney exchange programs. Discrete Mathematics, Algorithms, and Applications 1(4), 499–517 (2009)
5. Nisan, N.: Introduction to mechanism design for computer scientists. In: Algorithmic Game Theory, ch. 9, pp. 209–241. Cambridge University Press (2007)
6. Procaccia, A.D., Tennenholtz, M.: Approximate mechanism design without money. In: Proc. of the 10th ACM Conference on Electronic Commerce (EC), pp. 177–186 (2009)
7. Rees, M., Pelletier, R., Mulgaonkar, S., Laskow, D., Nibhanupudy, B., Kopke, J., Roth, A.E., Ünver, M.U., Sandholm, T., Rogers, J.: Report from a 60 transplant center multiregional kidney paired donation program. Transplantation 86(2S), 1 (2008)
8. Roth, A.E., Sönmez, T., Ünver, M.U.: Kidney exchange. Quarterly Journal of Economics 119, 457–488 (2004)
9. Roth, A.E., Sönmez, T., Ünver, M.U.: Pairwise kidney exchange. Journal of Economic Theory 125, 151–188 (2005)
10. Roth, A.E., Sönmez, T., Ünver, M.U.: Efficient kidney exchange: coincidence of wants in markets with compatibility-based preferences. American Economic Review 97, 828–851 (2007)
11. Saidman, S., Roth, A.E., Sönmez, T., Ünver, M.U., Delmonico, F.: Increasing the opportunity of live kidney donation by matching for two- and three-way exchanges. Transplantation 81(5), 773 (2006)
12. Sönmez, T., Ünver, M.U.: Market design for kidney exchange. In: Neeman, Z., Niederle, N., Vulkan, M. (eds.) Oxford Handbook of Market Design, Oxford University Press (to appear)
13. Toulis, P., Parkes, D.: A random graph model of kidney exchanges: efficiency, individual-rationality and incentives. In: Proc. of the 12th ACM Conference on Electronic Commerce (EC), pp. 323–332 (2011)