Online Supplement to Recourse in Kidney Exchange Programs

B.Smeulders, V.Bartier, Y.Crama, F.C.R.Spieksma

A The complexity of the selection problem: proofs

In this appendix, we provide the proofs of the theorems stated in Section 3.1 establishing the complexity of three versions of the selection problem.

Our reductions are from the satisfiability problem SAT, and from a special case of SAT, called (2,2)-3SAT. Recall that an instance of SAT is defined as follows.

Problem: SAT

Instance: A set of n Boolean variables $\{w_1, \ldots, w_n\}$ and a set of m clauses $C = \{c_1, \ldots, c_m\}$ over the variables.

Question: Is there a truth-assignment that satisfies all clauses in C?

The special case (2,2)-3SAT arises when (i) each variable occurs twice negated, and twice unnegated, and (ii) each clause contains exactly three literals. It is proven NP-complete in Berman et al. (2004).

A.1 The selection problem with edge probabilities

Problem DecSPedge is the decision version of the selection problem with edge probabilities. See Section 3.3 for a complete statement.

Theorem 1. DecSPedge is NP-complete.

Proof. **Proof:** We are going to provide a reduction from (2,2)-3SAT. Given an instance of (2,2)-3SAT, we construct an instance of DecSPedge as follows. We first build the vertex set V.

- For each clause $c_i \in C$, there are three so-called *clause vertices* $v_{c_{(i,1)}}, v_{c_{(i,2)}}$ and $v_{c_{(i,3)}}$, $i = 1, \ldots, m$.
- For each variable w_j, there are vertices $v_{w_j}, v^+_{w_j}$ and $v^-_{w_j}$, $j = 1, \ldots, n$.

Thus $V = \{v_{c_{(i,1)}}, v_{c_{(i,2)}}, v_{c_{(i,3)}} |\ i = 1, \ldots, m\} \cup \{v_{w_j}, v^+_{w_j}, v^-_{w_j} |\ j = 1, \ldots, n\}$.

We proceed by creating the edge set E. We use three sets of edges: truth-assignment edges (TA), clause-satisfying edges (CS), and dummy edges (D).

- The truth-assignment edges are defined as follows:

$$TA \equiv \{e(v_{w_j}, v^+_{w_j}), e(v_{w_j}, v^-_{w_j})|\ j = 1, \ldots, n\}.$$

  Each edge in TA has success probability 1.

- The clause-satisfying edges are defined as follows. Recall that each clause in C contains three variables; we arbitrarily index them using index k, $k = 1, 2, 3$. We have, for each $i = 1, \dots, m$:

  $CS_i^+ \equiv \cup_j \{e(v_{c_{(i,k)}}, v_{w_j}^+) |\, w_j$ occurs unnegated as k-th variable in clause $c_i, k = 1, 2, 3\}$,

  and

  $CS_i^- \equiv \cup_j \{e(v_{c_{(i,k)}}, v_{w_j}^-) |\, w_j$ occurs negated as k-th variable in clause $c_i, k = 1, 2, 3\}$.

  We set $CS^+ = \cup_i CS_i^+$, $CS^- = \cup_i CS_i^-$, and $CS = CS^+ \cup CS^-$. Each edge in CS has success probability $\frac{1}{m}$.

- The dummy edges are defined as follows:

  $$D \equiv \{e(v_{c_{(i,1)}}, v_{c_{(i,2)}}), e(v_{c_{(i,1)}}, v_{c_{(i,3)}}), e(v_{c_{(i,2)}}, v_{c_{(i,3)}})|\ i = 1, \dots, m\}.$$

  Each dummy edge has success probability 1. We set $E = TA \cup CS \cup D$.

See Figure 1 for graphical illustration.

Furthermore, we set $B := n + 2m$ and $Z := n + m + 1 - \frac{1}{2m}$, thereby completing the description of the instance of DecSPedge. We claim that there is a satisfying truth-assignment for C if and only if there exists an edge set E^* with $|E^*| \leq B$ and $\mathbb{E}((V, E^*), p) \geq Z$.

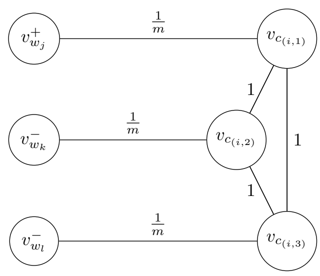


Figure 1: Clause-satisfying edges and dummy edges used in the construction of an instance of SPedge for clause $c_i = (w_j \vee \bar{w}_k \vee \bar{w}_l)$.

$\Rightarrow$ Suppose we have a satisfying truth-assignment for C. We will show how to identify an edge set E^* with $|E^*| \leq n + 2m$ such that the expected value of a maximum matching in $G = (V, E^*)$ is at least $n + m + 1 - \frac{1}{2m}$.

If, in a satisfying truth assignment for C, variable w_j is TRUE, we add edge $e(v_{w_j}, v_{w_j}^-)$ to E^*; else, if w_j is FALSE, we add edge $e(v_{w_j}, v_{w_j}^+)$ to E^*. Next, for each clause c_i, let w_j be a variable satisfying this clause, and, in fact, let w_j be the k^{th} variable in that clause, $k = 1, 2, 3$. Then, we add edge $e(v_{c_{(i,k)}, v_{w_j}^-})$ to E^* if w_j is TRUE, and we add edge $e(v_{c_{(i,k)}, v_{w_j}^+})$ to E^* if w_j is FALSE. Additionally, we add dummy edges $e(v_{c_{(i,\ell)}}, v_{c_{(i,\ell')}})$ with distinct $\ell, \ell' \neq k$ to E^*, for $i = 1, \dots, m$.

We consider the three sets of edges and their corresponding contributions to the expected size of a maximum matching.

- Each of the m dummy edges in E^* is not adjacent to any other edge in E^*. Since each dummy edge succeeds with probability 1, they collectively add m to the size of the matching.
- Each of the n truth-assignment edges in E^* is not adjacent to any other edge in E^*. Since each edge in TA succeeds with probability 1, they collectively add n to the size of the matching.
- Next, consider the clause-satisfying edges, of which m are present in E^*. A clause-satisfying edge in E^* can be adjacent to at most one other clause-satisfying edge in E^*; this follows from the fact that the number of clause-satisfying edges that are incident to the same node is bounded by 2, which, in turn, follows from the fact that each literal occurs twice negated and twice unnegated. These edges succeed with probability $\frac{1}{m}$, and since the probability of adjacent edges succeeding is independent, the probability of both edges succeeding is $\frac{1}{m^2}$. Two adjacent edges thus jointly add $\frac{2}{m} - \frac{1}{m^2}$, or $\frac{1}{m} - \frac{1}{2*m^2}$ on average, to the expected size of the matching. All m clause-satisfying edges thus collectively contribute at least $1 - \frac{1}{2m}$ to the size of the matching.

Adding all these contributions together, we conclude that the expected size of the matching constructed is at least $n + m + 1 - \frac{1}{2m}$.

$\Leftarrow$ Suppose there exists an edge set E^* with $|E^*| \leq n + 2m$ such that the expected size of a maximum matching in (V, E^*) equals at least $n+m+1-\frac{1}{2m}$. From this we will infer that (i) E^* must contain n edges from TA that prescribe a truth assignment, and contribute n, (ii) E^* must contain m edges from D contributing m, and (iii) E^* must contain m edges from CS. Moreover, this latter set of edges must contribute at least $1 - \frac{1}{2m}$, implying that none of them is adjacent to either an edge from TA in E^*, or to a dummy edge in E^*. In addition, no pair of CS edges in E^* shares a clause vertex. Thus, there is a CS edge in E^* incident to exactly one of the three vertices $v_{c_{(i,1)}}, v_{c_{(i,2)}}$ and $v_{c_{(i,3)}}$, for each $i = 1, \ldots, m$, implying that the truth assignment found satisfies C.

We now argue the validity of the claims above.

First, we establish that exactly $n + m$ of the edges in E^* are edges from $TA \cup D$, and that no pair of these edges is adjacent. This will imply the existence of a truth assignment, i.e., the edges in TA that are present in E^* prescribe how to set each variable to either TRUE or FALSE. We argue by contradiction.

- Consider the case where less than $n + m$ edges in E^* are from $TA \cup D$. Since every edge in E^* is either an edge from $TA \cup D$ or an edge from CS, it follows that E^* consists of $n+m-q$ edges from $TA \cup D$ and $m+q$ clause-satisfying edges, for some $q \geq 1$. Clearly, the resulting size of a matching in E^* is at most $(n + m - q) + (m + q)\frac{1}{m}$, as a single edge from $TA \cup D$ adds at most 1 to the solution value, while a single edge from CS adds at most $\frac{1}{m}$. For $q \geq 1$, this gives $n+m+1-q(\frac{m-1}{m}) < n+m+1-\frac{1}{2m}$ (assuming $m \geq 2$).
- On the other hand, if more than $n+m$ edges in E^* are from $TA \cup D$, they can collectively still contribute at most $n + m$ to the size of a matching, as no matching can contain more than n truth-assignment edges and m

edges from D. The maximum size of a matching is then bounded by $n+m+(m-q)\frac{1}{m} = n+m+1-\frac{q}{m} < n+m+1-\frac{1}{2m}$ for $q \geq 1$.

- Finally, two adjacent edges from $TA \cup D$ that are both in E^* collectively only add at most 1 unit to the size of a matching. If $n+m$ edges from $TA \cup D$ are in E^*, and some of these are adjacent, the maximum size of a matching cannot exceed $n+m$.

It follows that E^* contains m edges from D, as well as n edges from TA that prescribe a truth assignment as follows: if $e(w_j, w_j^+)$ ($e(w_j, w_j^-)$) is in E^*, set v_j to FALSE (TRUE), $j = 1, \ldots, n$. Since exactly $n+m$ edges from $TA \cup D$ are in E^*, we know that E^* contains m clause-satisfying edges. It remains to argue that (i) no CS edge in E^* is adjacent to a TA edge in E^*, and (ii) no pair of CS edges in E^* is adjacent.

Since the m clause-satisfying edges in E^* collectively contribute at least $1-\frac{1}{2m}$ to the expected size of a matching, we derive the following claims.

- We claim that no CS edge in E^* is adjacent to a TA edge in E^*. Indeed, even if one CS edge that is adjacent to a TA edge in E^*, is also in E^*, the contribution of the CS edges in E^* is at most $(m-1)\frac{1}{m} < 1-\frac{1}{2m}$. This observation implies that if a CS edge is in E^*, it is consistent with the truth assignment induced by the TA edges, and hence represents a literal satisfying a clause in C.

- If there is a triple of clause nodes $v_{c_{(i,1)}}, v_{c_{(i,2)}}, v_{c_{(i,3)}}$ to which no CS edge in E^* is adjacent, then the contribution from the CS edges to the expected size of the matching does not exceed $(m-1)\frac{1}{m}$. Since the expected size of a maximum matching exceeds $n+m+1-\frac{1}{2m}$, it follows that, for each triple of clause nodes, exactly one must be adjacent to an edge in CS.

These implications ensure that the set of m CS edges present in E^* is consistent with the truth-assignment edges in E^*, thereby satisfying each clause in C. The proof is complete. □

A.2 The selection problem with vertex probabilities

DecSPvertex is the decision version of the selection problem with vertex probabilities. We have stated the following theorem in Section 3.3:

THEOREM 2. DecSPvertex is NP-complete.

Proof. **Proof:** Given an instance of (2,2)-3SAT, we construct an instance of DecSPvertex as follows. Let us first build the vertex set V and the associated probabilities; we use a similar construction as in the proof of Theorem 1, but slightly simplified.

- For each clause $c_i \in C$, there is a so-called *clause vertex* $v_{c_i} \in V$ with $p(v_{c_i}) = \frac{1}{m}$, $i = 1, \ldots, m$.

- For each variable w_j, there are three vertices $v_{w_j}, v_{w_j}^+$ and $v_{w_j}^-$ in V, each with $p(v_{w_j}) = p(v_{w_j^+}) = p(v_{w_j^-}) = 1$, $j = 1, \ldots, n$.

Thus $V = \{v_{c_i} | \ i = 1, \ldots, m\} \cup \{v_{w_j}, v^+_{w_j}, v^-_{w_j} | \ j = 1, \ldots, n\}$. As in the proof of Theorem 1, we introduce truth-assignment edges:

$$TA \equiv \{e(v_{w_j}, v^+_{w_j}), e(v_{w_j}, v^-_{w_j}), j = 1, \ldots, n\}.$$

We also use clause-satisfying edges as follows. For each $i = 1, \ldots, m$ define:

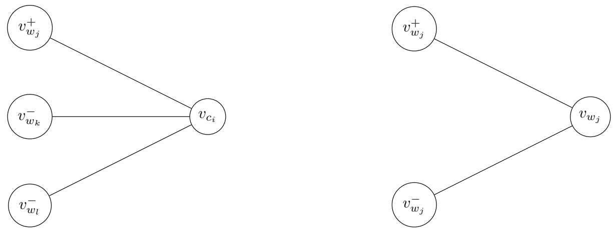
$$CS^+_i \equiv \cup_j \{e(v_{c_i}, v^+_{w_j}) | \text{ variable } w_j \text{ occurs positively in clause } c_i\},$$

and

$$CS^-_i \equiv \cup_j \{e(v_{c_i}, v^-_{w_j}) | \text{ variable } w_j \text{ occurs negatively in clause } c_i\}.$$

By setting $E = TA \cup \cup_i CS^+_i \cup \cup_i CS^-_i$, we have constructed G (see Figures 2(a) and 2(b)). Furthermore, we set $B := n + m$, and $Z := n + 1 - \frac{1}{2m}$.

We claim that there is a satisfying truth-assignment for C if and only if there exists an edge set E^* with $|E^*| \leq B$ and $\mathbb{E}((V, E^*), p) \geq Z$. The reasoning we use is similar to the reasoning used in the proof of Theorem 1.



(a) Clause-satisfying edges

(b) Truth-assignment edges

Figure 2: Edges used in the construction of an instance of SPvertex.

$\Rightarrow$ Suppose we have a satisfying truth-assignment for C. We will show how to identify an edge set E^* with $|E^*| \leq B = n + m$ such that the expected value of a maximum matching in $G = (V, E^*)$ is at least $n + 1 - \frac{1}{2m}$.

If, in a satisfying truth assignment for C, variable w_j is TRUE, we add edge $e(v_{w_j}, v^-_{w_j})$ to E^*; else we add edge $e(v_{w_j}, v^+_{w_j})$ to E^*, $j = 1, \ldots, n$. In this way, we add n truth-assignment edges to E^*. Further, since C is satisfiable, there exists for each clause c_i, a variable, say w_j, whose truth assignment realizes this clause. For each $i = 1, \ldots, m$, we do the following: if w_j occurs positively in c_i we add edge $e(v_{c_i}, v^+_{w_j})$ to E^*, and if w_j occurs negatively in c_i we add edge $e(v_{c_i}, v^-_{w_j})$ to E^*. In this way, we add m clause-satisfying edges to E^*. We have now specified E^*; observe that it contains $n + m$ edges. Since each edge in E^* is spanned by two vertices, say v and w with associated success probabilities p_v and p_w, we say that edge $\{v, w\}$ *succeeds* with probability $p_v p_w$. In particular, observe that each edge in TA succeeds with probability 1, whereas each edge in CS succeeds with probability $\frac{1}{m}$.

Consider now the graph (V, E^*). We will identify a matching in this graph, with the required expected size. The matching consists of truth-assignment edges and clause-satisfying edges.

- First, we put all n truth-assignment edges that are in E^* in the matching. Notice that (i) no pair of TA edges in E^* is adjacent (by construction), (ii) no TA edge in E^* is adjacent to a CS edge in E^* (since we have a satisfying truth assignment for C), and (iii) each edge in TA succeeds with probability 1. Thus, the set of TA edges in E^* form a partial matching and jointly contribute expected weight n to the matching.

- Next, consider the m clause-satisfying edges present in E^*. Each clause-satisfying edge is spanned by some node v_{c_i}, and one node from $\{v^+_{w_j}, v^-_{w_j}\}$, say node v. The construction of E^* allows that pairs of CS edges in E^* are adjacent; indeed, it may happen that a same literal makes different clauses true. Observe however, that the reduction from (2,2)-3SAT, where each variable occurs twice negated, and twice unnegated, guarantees that there do not exist three CS edges sharing a node. Thus, there are two cases: either node v is adjacent to a single CS edge in E^*, or node v is adjacent to two CS edges in E^*. In the first case, we simply select the corresponding edge in the matching; it contributes weight $\frac{1}{m}$. In the second case node v is adjacent to two CS edges in E^*; then, we select one of these edges. More precisely, if one such edge succeeds, or if both edges survive, we can select one of them in the matching; the corresponding probability of such an event is $\frac{2}{m} - \frac{1}{m^2}$. Thus, summing over the m edges, we can add at least $1 - \frac{1}{2m}$ to the expected value of the matching.

Concluding, the expected value of this matching equals at least $n + 1 - \frac{1}{2m}$.

$\Leftarrow$ Suppose there exists an edge set E^* with $|E^*| \leq n + m$ such that the expected size of a maximum matching (V, E^*) equals at least $n + 1 - \frac{1}{2m}$. We construct a truth-assignment satisfying C as follows.

First, we establish that exactly n of the edges in E^* must be truth-assignment edges, and that no pair of these edges is adjacent. This will imply the existence of a truth assignment, i.e., the TA edges present in E^* prescribe how to set each variable to either TRUE or FALSE. By following a similar reasoning as in the proof of Theorem 1, we conclude that E^* contains n edges from TA that prescribe a truth assignment as follows: if $e(w_j, w_j^+)$ $(e(w_j, w_j^-))$ is in E^*, set v_j to FALSE (TRUE), $j = 1, \ldots, n$. Since exactly n truth-assignment edges are in E^*, we know that E^* contains m clause-satisfying edges. It remains to argue that (i) no CS edge in E^* is adjacent to a TA edge in E^*, and (ii) no pair of CS edges in E^* are adjacent to a clause vertex.

Since the m clause-satisfying edges in E^* collectively contribute at least $1 - \frac{1}{2m}$ to the expected size of a matching, we derive the following claims.

- We claim that no CS edge in E^* is adjacent to a TA edge in E^*. Indeed, even if one CS edge that is adjacent to a TA edge in E^*, is also in E^*, the contribution of the CS edges in E^* is at most $(m-1)\frac{1}{m} < 1 - \frac{1}{2m}$. This observation implies that if a CS edge is in E^*, it is consistent with the truth assignment induced by the TA edges, and hence represents a literal satisfying a clause in C.

- If two clause-satisfying edges corresponding to the same clause are both in E^*, they jointly add at most $\frac{1}{m}$ to the expected size of a matching. This is a consequence of the fact that, in that case, these two edges must have some node v_{c_i} in common which - by construction - succeeds with probability $\frac{1}{m}$, while the other nodes spanning these edges do not fail. This ensures that both edges are either both successful or both fail. Thus, if two such edges are in E^*, the contribution of the CS edges in E^* is bounded by $(m-1)\frac{1}{m}$. It follows that no two clause-satisfying edges corresponding to the same clause are in E^*. The m tested clause-satisfying edges thus all correspond to a different clause, and hence all clauses are satisfied.

These implications ensure that the set of m CS edges present in E^* is consistent with the truth-assignment edges in E^*, thereby satisfying each clause in C. The proof is complete. □

A.3 The selection problem with explicit scenarios

We finally turn to problem DecSPscen, the decision version of the selection problem associated with a subset of scenarios (see Section 3.1).

Theorem 3. DecSPscen is NP-complete.

Proof. **Proof:** Given an instance of SAT, we construct an instance of DecSPscen as follows. Let us first build the graph $G = (V, E)$.

As in the previous proof, we set $V = \{v_{c_i} |\ i = 1, \ldots, m\} \cup \{v_{w_j}, v^+_{w_j}, v^-_{w_j} |\ j = 1, \ldots, n\}$

We now specify the edge sets E_s $(1 \leq s \leq t)$, as well as the set E.

We set $t := m$, i.e., there is a scenario (i.e, an edge set E_s) for every clause c_s. We use, as in previous proofs, the truth-assignment edges that will be present in each set E_s $(1 \leq s \leq t)$:

$$TA := \{e(v_{w_j}, v^+_{w_j}), e(v_{w_j}, v^-_{w_j}), j = 1, \ldots, n\}.$$

In addition, for each clause c_s $(s = 1, \ldots, t)$, define the edge sets

$$CS^+_s := \bigcup_j \{e(v_{c_s}, v^+_{w_j})|\text{ variable } w_j \text{ occurs positively in clause } c_s\},$$

$$CS^-_s := \bigcup_j \{e(v_{c_s}, v^-_{w_j})|\text{ variable } w_j \text{ occurs negatively in clause } c_s\}.$$

We now define, for each $s = 1, \ldots, t$:

$$E_s := TA \cup CS^+_s \cup CS^-_s.$$

Moreover, we set $E := \cup_s E_s$. Further, we set $B := n + m$ and $Z := m(n + 1)$. This completes the description of an instance of the decision version of SPscen.

We claim that there exists a satisfying truth assignment for C if and only if there exists an edge set $E^* \subseteq E$ with $|E^*| \leq B$ and $\sum_{s=1}^{t} z(V, E_s \cap E^*) \geq m(n+1)$.

$\Rightarrow$ Suppose we have a satisfying truth-assignment for C. We will show how to identify an edge set E^* with $|E^*| \leq B = n+m$ such that whatever scenario/edge set E_s realizes, at least $n+1$ edges can be selected in an optimum solution of the resulting instance of the KEP, i.e., $z(V, E_s \cap E^*) \geq n + 1$ for each $s = 1, \ldots, t$.

If, in a satisfying truth assignment for C, variable w_j is TRUE, we add edge $e(v_{w_j}, v^-_{w_j})$ to E^*; else we add edge $e(v_{w_j}, v^+_{w_j})$ to E^*, $j = 1, \ldots, n$. Further, since C is satisfiable, there exists for each clause c_s, a variable, say w_j, whose truth assignment satisfies this clause. Now, for each $s = 1, \ldots, t$, we do the following: if this w_j occurs positively in c_s, we add edge $e(v_{c_s}, v^+_{w_j})$ to E^*, and if this w_j occurs negatively in c_s, we add edge $e(v_{c_s}, v^-_{w_j})$ to E^*. We have now specified E^*; observe that it need not be a matching in G, and that it contains $n + m$ edges.

Consider now the instance of the KEP defined by the graph $(V, E_s \cap E^*)$. We claim that one can always find a matching of size $n+1$ in $(V, E_s \cap E^*)$. Indeed, there are n edges present in $TA \cap E^*$, and, by construction of E^*, there is one edge incident to v_{c_s} that is in E^*, and is not adjacent to any of the edges in $TA \cap E^*$. Hence $z(V, E_s \cap E^*) \geq n+1$ for each $s = 1, \ldots, t$, and this implies that the instance of the selection problem with explicit scenarios is a yes-instance.

$\Leftarrow$ Now, we show that if the instance of the decision version of SPscen is a yes-instance, i.e., if there exists an E^* with $|E^*| \leq n+m$ such that $\sum_{s=1}^{t} z(V, E_s \cap E^*) \geq m(n+1)$, there must be a satisfying truth assignment for C. First, consider the graph $G_s = (V, E_s)$ $(1 \leq s \leq t)$. We claim:

$$z(V, E_s \cap E^*) \leq z(G_s) \leq n + 1 \text{ for } s = 1, \ldots, t. \tag{1}$$

Indeed, in an optimum solution to the KEP instance defined by G_s, one can select at most one edge out of the two adjacent edges $\{e(w_j, w_j^+), e(w_j, w_j^-)\}$ for $j = 1, \ldots, n$, and one can select at most one edge incident to node v_{c_s}. Since no other edges exist in E_s, the upper bound in (1) is valid. Further, given that our instance of the decision version of SPscen is a yes-instance, we have:

$$\sum_{s=1}^{t} z(V, E_s \cap E^*) \geq m(n + 1). \tag{2}$$

Combining (1) and (2) implies that E^* is such that:

$$z(V, E_s \cap E^*) = n + 1 \text{ for each } s = 1, \ldots, t.$$

In words, E^* is such that, when intersected with the edges of any scenario E_s, the size of a maximum matching equals $n + 1$. Clearly, this value can only be realized by having in E^* at least one edge out of each of the n pairs $\{e(w_j, w_j^+), e(w_j, w_j^-)\}$ $(1 \leq j \leq n)$, and at least one edge incident to each v_{c_s} $(1 \leq s \leq t)$. However, since we know that $|E^*| \leq n + m$, we conclude that E^* contains exactly one edge from each pair $\{e(w_j, w_j^+), e(w_j, w_j^-)\}$ and exactly one edge incident to each v_{c_s}. The n edges in E^* from the pairs $\{e(w_j, w_j^+), e(w_j, w_j^-)\}$ determine the truth assignment of the variables: if $e(w_j, w_j^+)$ is in E^*, then we set w_j to FALSE, else we set w_j to TRUE. The m edges in E^* incident to the nodes v_{c_s} imply that this truth assignment satisfies C: for each individual clause c_s there is an edge $e(v_{c_s}, v_{w_j})$ in E^*, meaning there is a variable w_j whose truth assignment satisfies clause c_s. The proof is complete. $\square$

B Integer programming formulation

In Section 2.2, we have introduced a very generic formulation of the deterministic kidney exchange problem, which has been extended in Section 3.2 to a formulation of the selection problem. We present hereunder the specific formulation of the selection problem which has been used in our computational experiments.

Position Indexed Edge formulation

The *Position Indexed Edge* (PIE) formulation is proposed by Dickerson et al. (2016). The main feature of this formulation is the introduction of *position indices*, which denote in which position a particular arc is used in a cycle. In order to model the selection problem, we further extend this formulation by defining binary variables $\gamma_{i,j,k,\ell,s}$, where $\gamma_{i,j,k,\ell,s} = 1$ if and only if the arc (i,j) is chosen in the k^{th} position of a cycle in graph copy ℓ in scenario s. Arcs in chains are handled similarly, though the graph copies are not needed here: variable $\delta_{i,j,k,s} = 1$ if and only if arc (i,j) is chosen in the k^{th} position of a chain in scenario s. The sets $\mathcal{K}(i,j,\ell,s)$ and $\mathcal{K}'(i,j,s)$ are computed in a preprocessing step, and contain those positions in which (i,j) can be selected in scenario s. We refer to Dickerson et al. (2016) for more details on the precomputation of these sets.

In the formulation hereunder, the objective function (3) and the constraints (4), (5), (6) are directly inherited from the generic formulation. The cycle flow constraints (7) enforce that if an arc arriving at a vertex i is selected in position k in a cycle, another arc leaving that vertex i must be selected in position $k+1$. By not allowing arcs in positions larger than K, the cycle length constraints are enforced. Similarly, the chain flow constraints (8) enforce that an arc may only be chosen in position k of a chain if an arc enters the same vertex in position $k-1$. Graph copies are again used to break symmetry, as in Section 4.1.

$$\max \sum_{s\in S} q_s \sum_{(i,j)\in A_s} \Big(\sum_{\ell\in V} \sum_{k\in\mathcal{K}(i,j,\ell,s)} \gamma_{i,j,k,\ell,s} + \sum_{k\in\mathcal{K}'(i,j,s)} \delta_{i,j,k,s}\Big) \tag{3}$$

$$\text{subject to} \quad \sum_{(i,j)\in A} \beta_{i,j} \le B, \tag{4}$$

$$\sum_{\ell\in V} \sum_{j:(j,i)\in A_s^\ell} \sum_{k\in\mathcal{K}(j,i,\ell,s)} \gamma_{j,i,k,\ell,s} + \sum_{j:(j,i)\in A_s} \sum_{k\in\mathcal{K}'(j,i,s)} \delta_{j,i,k,s} \le 1 \qquad \forall s\in S, i\in V, \tag{5}$$

$$\sum_{\ell\in V} \sum_{k\in\mathcal{K}(i,j,\ell,s)} \gamma_{i,j,k,\ell,s} + \sum_{k\in\mathcal{K}'(i,j,s)} \delta_{i,j,k,s} \le \beta_{i,j} \qquad \forall s\in S, (i,j)\in A_s, \tag{6}$$

$$\sum_{j:((j,i)\in A_s^\ell \wedge k\in\mathcal{K}(j,i,\ell,s))} \gamma_{j,i,k,\ell,s} = \sum_{j:((i,j)\in A_s^\ell \wedge (k+1)\in\mathcal{K}(i,j,\ell,s))} \gamma_{i,j,k+1,\ell,s} \qquad \begin{array}{l}\forall s\in S, \ell\in V,\\ i\in\{\ell+1,\dots,n\},\\ k\in\{1,\dots,K-1\}.\end{array} \tag{7}$$

$$\sum_{j:((j,i)\in A_s \wedge k\in\mathcal{K}'(j,i,s))} \delta_{j,i,k,s} \le \sum_{j:((i,j)\in A_s \wedge k\in\mathcal{K}'(i,j,s))} \delta_{i,j,k-1,s} \quad \forall s\in S, i\in V, k\in\{2,\dots,L\}, \tag{8}$$

$$\beta_{i,j}, \gamma_{i,j,k,\ell,s}, \delta_{i,j,k',s} \in \{0,1\} \quad \forall s\in S, i,j,\ell\in V, k\in\mathcal{K}(i,j,\ell,s), k'\in\mathcal{K}'(i,j,s). \tag{9}$$

References

Piotr Berman, Marek Karpinski, and Alexander Scott. Approximation hardness of short symmetric instances of MAX-3SAT. *Report IHES-M-2004-28*, 2004.

John P Dickerson, David F Manlove, Benjamin Plaut, Tuomas Sandholm, and James Trimble. Position-indexed formulations for kidney exchange. In *Proceedings of the 2016 ACM Conference on Economics and Computation*, pages 25–42. ACM, 2016.