Linearized formulations for failure aware barter exchange

Noam Goldberg, Michael Poss

Noname manuscript No.
(will be inserted by the editor)

Linearized Formulations for Failure Aware Barter Exchange

Noam Goldberg[*] · **Michael Poss**

the date of receipt and acceptance should be inserted later

Abstract Mathematical programming formulations are developed for determining chains of organ-donation exchange pairs in a compatibility graph where pairwise exchanges may fail. The objective is to maximize the expected value where pairs are known to fail with given probabilities. In previous work, namely that of Dickerson et al. (2019) this NP-hard problem was solved heuristically or exactly only for limited path lengths. Although the problem appears highly nonlinear, we formulate it as a mixed-integer linear program (MILP). A computationally tractable layered formulation that approximately solves larger instances is also proposed and a computational study is presented for evaluating the proposed formulations.
Keywords: Longest path, MILP, layered formulation, clearing algorithms, kidney exchange, failure aware barter exchange

1 Introduction

In this paper we consider a maximum expected value partially successful (s,t)-path problem (PSPP), where each (s,v)-subpath of an (s,t)-path, up to some vertex v along the path, has a value of the product of probabilities of arcs along the (s,v)-subpath, times the sum of arc profits.

Kidney exchange market clearing problems are barter exchanges where donors and recipients interact without exchange of monetary means; pairs of recipients and donors enter the market when they are not physiologically

[*]Corresponding

Noam Goldberg
Department of Management, Bar-Ilan University, Ramat Gan 5290002, Israel
E-mail: noam.goldberg@biu.ac.il

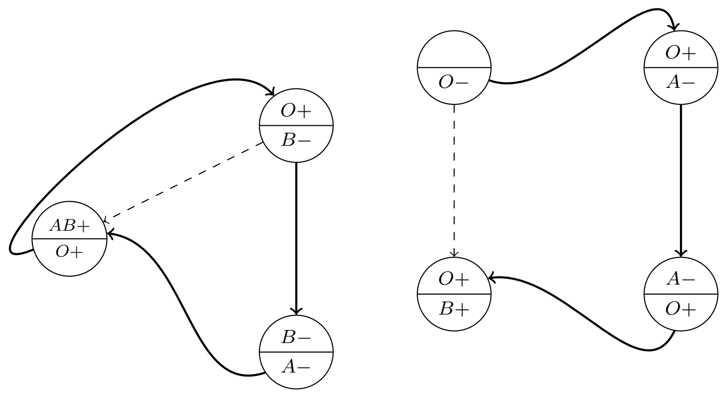
Michael Poss
LIRMM, University of Montpellier, CNRS, France
E-mail: michael.poss@lirmm.fr

compatible and a suitable exchange is sought so that the recipient can be accommodated. Paired-kidney donation graphs indicate donor-recipient compatibility using directed arcs. Each vertex in such a graph corresponds to a pair of a recipient (patient) and donor – typically a family member who is willing to donate a kidney in exchange for the related patient receiving one [1, 22, 23]. The market clearing problem is that of matching all recipients with donors via cycles [23], and paths [1]; see also later work [7,2] and references therein. Paths or chains can be formed when altruistic donor vertex is available to initiate the exchange chain without requiring any donation in return (there is no related recipient and so an altruistic vertex by definition only has outgoing arcs in the donation graph). Exchange cycles and chains are illustrated using small compatibility graph examples in Figure 1. Additional work concerning kidney-transplant exchanges using cycles and/or paths initiated by altruistic donors includes [12, 19].

There is significant uncertainty with respect to the success of a particular donation to take place as a part of the exchange: Physiological incompatibilities may be detected at later stages prior to the transplantation. In addition, a particular donor may become unavailable due to illness or simply because reneging on the agreement. Yet, chains are considered less risky than cycles because a failure of one donation does not necessarily prevent the entire chain of transplants from taking place. This setting with uncertainty gives rise to partially successful paths and the objective considered in this paper; see the expected utility of chains formulated in [9] and the future work discussion in [18]. The corresponding problem for a cycle, where all edges must succeed in order for it to be deemed successful, is also considered in [9] as well as other work including [6].

Most literature to date has considered heuristic methods for solving the pricing subproblem of barter-exchange market clearing optimization problem. Yet some exact solutions approaches have been considered in the absence of uncertainty, as [2, 12] that consider a cycle and path subproblems (i.e., matching failures). On the other hand, the pricing problem of the failure-aware market clearing problem introduced in [9], has been solved by heuristic means for both cycles and paths. In particular, it appears that exact approaches for the failure-aware path subproblem (i.e., PSPP) has not been considered in previous work. However, solving the pricing problem exactly is important for two reasons. First, exact column generation and/or branch-and-price schemes ultimately need to solve the pricing problem exactly at least in the last few iterations. Second, PSPP is arguably of interest in its own right, and more so than the corresponding cycle problem, because as altruistic donors are scarce and their offer to donate may expire, the PSPP may arise in an online setting.

Uncertainty of the planned donations and failed matches are also considered in stochastic setting with recourse where the probabilities of success are known (similar to the probabilities given as a part of the problem's input in the current setting), for example in [17], or in adjustable robust settings with finite uncertainty sets [5]. Both stochastic and robust approaches consider different types of recourse such as "back-arcs recourse" [17, 5], where exchanges



(a) Example of a cycle of recipient-donor pairs.

(b) Example of a chain initiated by an altruistic donor.

Fig. 1 Examples of a cycle and chain in organ compatibility graphs. Blood types of recipients and donors are indicated on the top part and lower part, respectively, of each vertex. Dashed and solid edges indicate unmatched or matched donations, respectively. Here compatibility is assumed to depend only on the blood types but in practice compatibility graphs may be sparser due to additional compatibility considerations.

may take place along the planned chains (or cycles) up to the failed match. This is in contrast with complete recourse that allows for re-planning of entire cycles and chains in the event of failure. The current model with kidney donation chains maximizes the expected value while allowing for such "back-arcs" recourse.

Our problem bears similarities with the probabilistic all-or-nothing problem, which is to select a subset of items so as to maximize the expected total profit – the product of probabilities of selected items times the sum of profits of these items. In particular, the all-or-nothing probabilistic path problem [13], has the subsets of items corresponding to paths in a graph, each of which has an expected value that is equal to the product of its underlying arc probabilities times the sum of its arc profits. Other all-or-nothing problems have been considered over all unrestricted subsets of a given ground set, and matchings in a graph [15]. In the context of failure-aware barter-exchange this "all-or-nothing" objective is considered more appropriate for cycles than for paths [9].

Next we formally describe PSPP and relate our mixed-integer nonlinear programming (MINLP) formulation to the optimization problem as it appears in the literature. We then propose a linearization and reformulation of the MINLP, followed by an extended formulation based on a layered graph. Finally, we conclude with computational experiments on simulated kidney exchange data to evaluate the proposed formulations along with attempts to strengthen them.

2 Problem definition

We now formally state the (PSPP) problem. Let $G = (V, E)$ be a directed graph, $s, t \in V$ be the source and the sink, and let $\delta^+(u) = \{v \in V : (u, v) \in E\}$ and $\delta^-(u) = \{v \in V : (v, u) \in E\}$ be sets of direct successors and predecessors of vertex $i \in V$, respectively. The problem with multiple sources and sinks can be also handled, by constructing a supersource s and a supersink t that are connected to the multiple sources and sinks, respectively. For each edge $e = (u, v) \in E$, we are given a positive profit $c_e = c_{uv} \geq 0$ and a probability of success $p_e = p_{uv} \in [0, 1]$. Let $m = |E|$ and $n = |V|$. Given a path $\pi \in \mathcal{P}(t)$ from s to t, where $\mathcal{P}(t) \in 2^m$ is the set of all *elementary* paths from s to t (meaning every vertex appears at most once), for $v \in \pi$ we define $\pi(v)$ as the subpath of π joining s and v. Following [9] (see also the future work discussion in [18]), the maximum expected value PSPP is to find a path $\pi \in \mathcal{P}(t)$ so that the objective function

$$z(\pi) = \sum_{(u,v)\in\pi} (1 - p_{uv}) \sum_{e\in\pi(u)} c_e \prod_{f\in\pi(u)} p_f + \sum_{e\in\pi} c_e \prod_{f\in\pi} p_f, \tag{1}$$

is maximized. First, the following proposition simplifies objective (1) to a form that can be exploited in order to propose an MILP for this problem.

Proposition 1 *The objective* (1) *is equivalent to*

$$z(\pi) = \sum_{(u,v)\in\pi} c_{uv} \prod_{f\in\pi(v)} p_f \tag{2}$$

Proof By straightforward algebra,

$$\begin{aligned}
(1) \quad &= -\sum_{(u,v)\in\pi} p_{uv} \sum_{e\in\pi(u)} c_e \prod_{f\in\pi(u)} p_f + \sum_{(u,v)\in\pi} \sum_{e\in\pi(u)} c_e \prod_{f\in\pi(u)} p_f + \sum_{e\in\pi} c_e \prod_{f\in\pi} p_f \\
&= -\sum_{(u,v)\in\pi} \sum_{e\in\pi(u)} c_e \prod_{f\in\pi(v)} p_f + \sum_{(u,v)\in\pi} \sum_{e\in\pi(v)} c_e \prod_{f\in\pi(v)} p_f \\
&= \sum_{(u,v)\in\pi} c_{uv} \prod_{f\in\pi(v)} p_f \\
&= z(\pi)
\end{aligned}$$

□

Then, the maximum expected value PSPP is

$$\max\{z(\pi) : \pi \in \mathcal{P}(t)\}. \tag{3}$$

Evidently the problem (3) with $p_e = 1$ for all $e \in E$ is a longest path problem that is strongly NP-hard in general directed graphs [11].

Note that although a particular source and destination are assumed there can be many altruistic donors and a chain may not have to end in any particular recipient. Accordingly, a super-source s is created with arcs (s, a) having

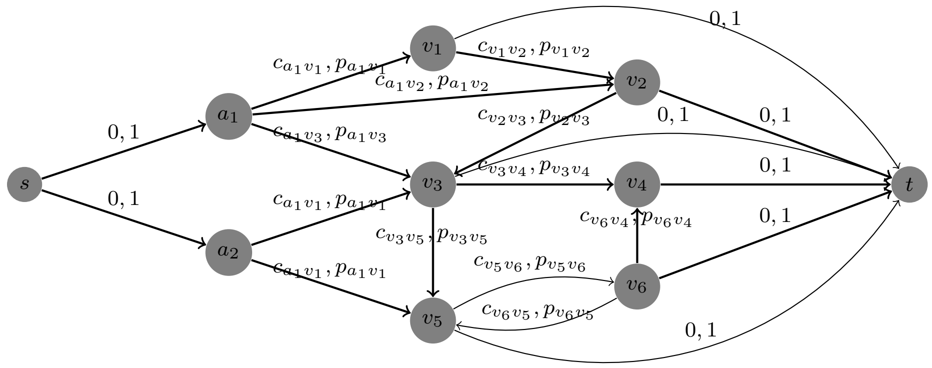


Fig. 2 Example organ compatibility graph with two altruistic donors (a_1 and a_2) and six vertices corresponding to donor-recipient pairs ($v_1, \ldots, v_6$).

probability $p_{sa} = 1$ and profit $c_{sa} = 0$ for all altruistic donors $a \in V$. Similarly the super-sink t with arcs (v, t) having probability $p_{vt} = 1$ and $c_{vt} = 0$ reflects the fact that any recipient may be last in the transplant chain. An example of such a graph with two altruistic donors and six donor-recipient pairs is shown in Figure 2. Finally, note that although altruistic donors are typically used to initiate chains, universal recipients (i.e., patients with an AB blood type) corresponding to vertices that have an in-degree of $n - 3$, may also be effectively used, since they are compatible with each and every possible donor who may terminate the exchange.

3 MILP formulations

We now present mixed-integer linear programming (MILP) formulations to exactly and approximately solve (3).

3.1 Exact formulation

Although the problem (3) appears to be highly nonlinear in fact it can be formulated as a MILP. Given a graph $G = (V, E)$, with $s, t \in V$, let the path flow polytope be denoted by

$$\mathit{\Pi} = \left\{ x \in \mathbb{R}_+^{|E|} \;\middle|\; \begin{array}{c} \sum_{u \in \delta^-(v)} x_{uv} = \sum_{u \in \delta^+(v)} x_{vu} \text{ for } v \in V \setminus \{s, t\} \\ \sum_{u \in \delta^+(s)} x_{su} = \sum_{u \in \delta^-(t)} x_{ut} = 1 \end{array} \right\}. \tag{4}$$

and consider the following MILP formulation.

$$
\begin{aligned}
\max \quad & \sum_{(u,v)\in E} c_{uv} q_{uv} && && \text{(5a)}\\
\text{subject to:} \quad & \sum_{w\in\delta^-(u)} q_{wu} = \sum_{v\in\delta^+(u)} \frac{1}{p_{uv}} q_{uv} && u \in V \setminus \{s,t\} && \text{(5b)}\\
& q_{uv} \le M_{uv} x_{uv} && (u,v) \in E && \text{(5c)}\\
& x \in \Pi && (u,v) \in E && \text{(5d)}\\
& x_{uv} \in \{0,1\}, q_{uv} \ge 0 && (u,v) \in E, && \text{(5e)}
\end{aligned}
$$

where $0 < M_{uv} \le 1$ is a sufficiently large probability constant. For each $(u,v) \in E$ the decision variable x_{uv} is used to indicate whether (u,v) is selected as a path edge, and the auxiliary variables q_{uv} is used to define $q_{uv} = \prod_{(u,v)\in\pi(v)} p_{uv}$ as in $z(\pi)$. The constraints corresponding to (5d) are the $s-t$ path-flow conservation constraints. Constraint (5c) enforces that $q_{uv} = 0$ for each $(u,v) \in E$ with $x_{uv} = 0$. Finally, (5b) implements the recursion to accumulate the probability product for each pair of selected incident edges (w,u) and (u,v) with $x_{uv} = x_{wu} = 1$, it requires that $q_{wu} p_{uv} = q_{uv}$.

To strengthen the formulation (5) it is desirable to select M_{uv} for each $(u,v) \in E$ as small as possible. It can be observed it suffices to set $M_{uv} = \bar{p}(u) p_{uv}$ where $\bar{p}(v)$ denotes the maximum probability of a path from s to v for all $v \in V \setminus \{s\}$ and $\bar{p}(s) = 1$. It is known that $\bar{p}$ can be determined in polynomial time by solving a maximum reliability problem. In particular, it can be observed by considering the shortest path problem with arc costs $-\ln p_e$ for each $e \in E$ (in an exact arithmetic setting) or replacing the addition in the Bellman equations with multiplication.

To prevent solutions that are feasible for (5) from containing cycles, cycle-breaking inequalities must be added. We consider two alternative ways of breaking the cycles. First, following [25], we consider *generalized cutset inequalities*,

$$
\sum_{(u,v)\in\delta^+(S)} x_{uv} \ge \sum_{(w,v)\in\delta^+(w)} x_{wv} \qquad S \subseteq V \setminus \{s,t\} : |S| \ge 2, w \in S. \tag{6}
$$

These constraints break any cycle by enforcing that the number of arcs leaving S be not smaller than the number of arcs outgoing from any vertex k of S.

A second type of inequalities will prevent cycles if the probabilities are strictly less than unity (specifically their product along cycles). The sufficiency of the inequalities

$$
\sum_{v\in\delta^+(u)} x_{uv} \le 1 \qquad u \in V \setminus \{t\}, \tag{7}
$$

for prevention of cycles, given the condition that success is uncertain along cycles, is established by the following proposition.

Proposition 2 *The problem* (5) *with* (7) *has an acyclic optimal solution if* $\prod_{f\in c} p_f < 1$ *for each cycle* $c \subseteq E$.

Proof Consider (x, q) that is a feasible solution of (5) with (7). Note that by the flow-conservation constraints and the integrality of x, the support of x is a union of one path π from s to t and a set of cycles $C \subset 2^E$. Suppose that there is some $u \in V$ that is incident to edges in two different cycles, or an edge in a cycle and an edge in π. Then, by the integrality of x, $\sum_{v \in \delta^+(u)} x_{uv} > 1$, thereby violating (7). It follows that the cycles $c \in C$ and the path π are all vertex disjoint. Now further suppose for the sake of deriving a contradiction that (x, q) is an optimal solution with minimal $|C| > 0$ and consider a cycle $c \in C$. If $q_e = 0$ for each $e \in c$, then a new solution (x', q) can be defined (using (x, q)) by setting $x'_e = 0$ for each $e \in c$. The objective value of (x', q) is equal to that of (x, q) and has one less cycle, thereby establishing a contradiction. Otherwise, if $q_e > 0$ for some $e \in c$, constraints (5b) written for all $(u, v) \in c$ imply that $\prod_{f \in c} p_f = 1$. □

3.2 Rounded formulation

The formulation from the previous section tends to have a weak linear programming relaxation, due the big-M coefficients involved in constraints (5c). We present in this section an alternative and stronger approach that is based on layered graphs, where multiple copies of graph vertices and edges are created and "stacked on top of each other"; see for example [16]. Specifically, we consider next a base $0 < \alpha < 1$ and define the rounded logarithms of the probabilities as $r_e = \lceil \log_\alpha p_e \rceil$. The set of rounded logarithms values is $K = \{0, \ldots, n \max_{e \in E}\{r_e\}\}$. Define $V_0 = V \setminus \{s, t\}$ and let its induced subset of edges be denoted by $E_0 = E[V_0]$. Then, consider copies of V_0, one for each $k \in K$, to be denoted by $V_0^k = \{(v, q) \in V_0 \times K \mid q = k\}$; each member $(v, k) \in V_0^k$ is the k-th copy of $v \in V_0$ in $\widehat{V}$. The layered graph is then $\widehat{G} = (\widehat{V}, \widehat{E})$, where $\widehat{V} = \{(s, 0), (t, 0)\} \cup \bigcup_{k \in K} V_0^k$ and

$$\widehat{E} = \left\{ (u, k, v, k + r_{u,v}) \;\middle|\; (u, v) \in E_0, k \in \{0, \ldots, n \max_{e \in E}\{r_e\} - r_{u,v}\} \right\}$$
$$\cup \left\{ (s, 0, v, r_{s,v}) \;\middle|\; v \in \delta^+(s) \right\} \cup \left\{ (u, k, t, 0) \;\middle|\; u \in \delta^-(t), k \in K \right\}.$$

Notice that $\widehat{G}$ is acyclic. Let $\widehat{\Pi}$ be the corresponding path flow polytope Π given by (4), defined for the layered graph $\widehat{G}$ in place of G, with $(s, 0)$ and $(t, 0)$ as the source and the sink vertices. Evidently, all paths in this graph correspond to extreme points of this polytope. Consider the formulation

$$\begin{aligned}
\max \quad & \sum_{e=(u,k,v,\bar{k})\in\widehat{E}} \alpha^{\bar{k}} c_{uv} x_e && && \text{(8a)}\\
\text{subject to:} \quad & \sum_{\substack{v\in V\setminus\{s\}:\exists\bar{k}\in K\\ e=(u,k,v,\bar{k})\in\widehat{E}}} x_e \leq 1 && u\in V\setminus\{t\} && \text{(8b)}\\
& x \in \widehat{\Pi} && && \text{(8c)}\\
& x_e \in \{0,1\} && e\in\widehat{E}. && \text{(8d)}
\end{aligned}$$

Note that while any path $\widehat{\pi} = ((u_1, q_1) = (s_1, q_1), (u_2, q_2), \ldots, (u_l, q_l) = (t_l, q_l))$ in $\widehat{G}$ is elementary, its projection onto V, $\pi = (u_1 = s, u_2, \ldots, u_l = t)$ might contain a cycle. We forbid such paths $\widehat{\pi}$ (whose projection contain one or more cycles) by limiting the number of outgoing arcs from any $u \in V \setminus \{s\}$ to at most one in constraints (8b). Also note that the rounding makes solving formulation (8) an inexact solution of (5) (and (3)). However, if α is chosen close enough to one, the worst-case performance of the optimal solution of (8) can be bounded from the theoretical viewpoint as formalized below.

Observation 1 *For $\alpha \leq (1-\epsilon)^{1/(n-1)}$, then each solution that is optimal to* (8) *is an $(1-\epsilon)$-approximate solution to* (5) *with* (7).

The correctness of this observation follows from the fact that if path $\pi = (u_1, \ldots, u_l)$ is optimal in the original graph then there is a path $\widehat{\pi}$ in $\widehat{G}$, whose projection onto V is π, and whose objective value is

$$\sum_{(u,k,v,\bar{k})\in\widehat{\pi}} \alpha^{\bar{k}} c_{uv} = \sum_{i=1}^{l-1} \alpha^{\sum_{j=1}^{i} r_{u_j u_{j+1}}} c_{u_i u_{i+1}} \geq \sum_{i=1}^{l-1} c_{u_i u_{i+1}} \alpha^i \prod_{j=1}^{i} p_{u_j u_{j+1}} \geq \alpha^{n-1} z(\pi).$$

4 Computational results

In this section we experiment with simulated kidney compatibility graphs that are publicly available [20] or randomly generated following the structure suggested in kidney exchange literature. All data used is maintained in a public repository [14].

Graph topology simulated according to [4] (graph[n]a[A] datasets). We generate a "sparse-dense" graph with the vertex connectivity generated according to the panel reactive antibodies (PRA) levels as suggested in [4], and we make use of additional data given in [3,9]. In particular, graph vertices are created with a highly sensitized recipient with probability 0.27. Such highly sensitized patients (with a high PRA value) are less likely to be compatible with a random donor. Accordingly, these recipients are then connected with donors (corresponding to an incident incoming edge) with probability 0.03. The failure probabilities are assigned also based on PRA levels and highly sensitized patients are matched with a success probability of 0.5. Otherwise, low PRA recipients are compatible with donors with probability 0.5, 57% of them are matched with a success probability of 0.75 (cross-match failure rate of 0.25) and 43% of them are matched with a success probability of 0.95 (cross-match failure rate of 0.05). The values of matching PRA recipients are generated from a normal distribution with mean 10 and standard deviation 2. The values of matching highly sensitized recipients are a factor of 1.5 greater (so the mean is 15) as suggested in [9].

Graphs based on simulated donor data instances of [8] and [9] (MD[n] datasets). These instances were generated using a simulator based on [24]. Graph arcs were constructed based on blood-type compatibility and we set the weights to be either unitary in datasets MD[n]unit or normally distributed with mean 10 and standard deviation 2 in datasets MD[n]-stoch. For patients that are considered highly sensitized with PRA levels exceeding 0.73 in the original data, the arc values considered were twice as much as initially generated. The success probabilities are set to 0.5 for highly sensitized patients (determined by the PRA levels). Otherwise, a success probability of 0.75 is set with probability 0.57 and success probability of 0.95 is set with probability 0.43 (similar to the graph[n]a[A] datasets).

4.1 Implementation details

All algorithms have been coded in Julia 1.3.1 on a platform using a (Intel i7-10510U) 1.80GHz CPU with 4 cores and an 8MB cache and 16GB of RAM. The mathematical programs are modeled using the JuMP v0.21.1 [10] and Gurobi v0.7.6 packages and solved with the Gurobi 9.0.1 solver running 4 threads.

None of our instances contain an edge with a probability equal to 1. Although, if an instance has multiple altruistic donors, we then connect supersources and supersinks to all vertices with edges having probability 1, but this may not lead to cycles c such that $\prod_{f \in c} p_f = 1$. Hence, the condition of Proposition 2 holds so inequalities (7) are enough to guarantee the existence of optimal acyclic solutions. This being said, inequalities (6) are not dominated by inequalities (7), so separation of the former through callbacks might further improve the overall performance of the formulation.

Following [25], we thus tested the separation of inequalities (6) at both fractional and integer solutions by checking whether they are satisfied on each strongly connected component of the graph induced by the positive components of x that is optimal to (a relaxation of) (5). Our results indicated a worsening of the overall performance: the relaxation improvement, often nonexistent, is too little to compensate for the time spent in the separation procedure. Therefore, we do not use inequalities (6) in the results presented next.

4.2 Description of the results

Table 1 shows for the particular kidney exchange graph instances that we experimented, statistics and characteristics including the number of vertices, edges, number of altruistic donors, length of the longest path from the supersource to any vertices in the graph, and the root relaxation gap (in %). Solution times are reported in seconds in Tables 2 and 3. Here T indicates that a time limit of 1800 seconds is reached and M indicates that the solver has exhausted the available memory. Several observations can be made from the results of these tables. First, we see that the extended formulation (8) scales reasonably well, as long as the memory limit is not exceeded. Some of the large MD instances, namely 84unit, 84stoch, 125unit, are solved faster than formulation (5). Second, the exact formulation (5) appears to be more efficiently solved with instances MD than with instances "graph". To further investigate this behavior we refer to the graph instance characteristics in Table 1. The table data shows that graph30a2 and instances MD43 – MD70 have comparable numbers of vertices and edges. However, graph30a2 cannot be solved in 1800 seconds, while these MD instances are all solved within a few seconds, except instance 66unit which requires roughly 72 seconds. A partial explanation of this behavior could lie in the strong root relaxation gap of (5) for these MD instances, ranging from 6 to 20% while for the graph30a2 instance it equals 31%. The strength of the relaxation appears to correlate with the length of paths in these graphs (directly impacting M_{uv}); for the MD instances with 34 vertices the length ranges from 17 to 22 hops, compared with 30 hops in graph30a2.

Next, the third group of gap columns of Tables 2 and 3 provide the ratio of the best solution found by the extended formulation (8) for different values of α, to the best solution determined by the exact formulation (5) (expressed as %). All instances in this table may exceed the time limit when solved exactly by formulation (5) and thus lead to negative gap values. Evidently, $\alpha < 0.8$ seems to lead to poor solutions, while setting α higher than 0.8 appears to yield only marginal improvements. Finally, we remark that for the instance MD125unit that cannot be solved to optimality, the layered formulation (8) even provides slightly better solutions than the best solution found by the exact formulation (within the 30 minute time limit). Note that the theoretical gap of Observation 1 for entries in these tables ranges from $\epsilon \leq 1-\alpha^{n-1} \leq 43\%$ for $n = 12$, $\alpha = 0.95$ to more than 99% for most table entries where $\alpha \leq 0.9$ and $n \geq 50$ (or even $n \geq 20$ when $\alpha = 0.75$). Evidently, the gap appearing in practice tends to be much smaller than the theoretical guarantee.

5 Conclusions and future work

In this paper we proposed a linear formulation for the maximum expected value partially successful path problem. For this formulation we prove and evaluate cycle prevention inequalities that outperform standard longest path

Instance	$\|V\|$	$\|E\|$	n_a	LP	Root gap
MD12stoch	18	96	1	9	9
MD12unit	18	96	1	9	10
MD19stoch	18	150	1	14	7
MD19unit	18	150	1	14	4
MD21stoch	18	114	2	10	7
MD21unit	18	114	2	10	1
MD23stoch	18	121	2	13	17
MD23unit	18	121	2	13	28
MD43stoch	34	402	1	19	14
MD43unit	34	402	1	19	13
MD44stoch	34	337	1	17	15
MD44unit	34	337	1	17	11
MD51stoch	34	405	3	19	6
MD51unit	34	405	3	19	16
MD60stoch	34	371	3	19	10
MD60unit	34	371	3	19	9
MD61stoch	34	422	4	19	16
MD61unit	34	422	4	19	17
MD66stoch	34	403	4	22	20
MD66unit	34	403	4	22	16
MD70stoch	34	332	4	18	13
MD70unit	34	332	4	18	18
MD81stoch	66	2205	3	55	15
MD81unit	66	2205	3	55	20
MD84stoch	66	1512	3	36	29
MD84unit	66	1512	3	36	27
MD125stoch	130	6026	6	73	7
MD125unit	130	6026	6	73	12

Instance	$\|V\|$	$\|E\|$	n_a	LP	Root gap
graph20a1	22	178	1	20	25
graph30a2	32	329	2	30	31
graph40a2	42	681	2	41	30
graph50a3	52	981	3	51	23
graph60a3	62	1387	3	61	17
graph70a4	72	1851	4	71	13
graph80a4	82	2247	4	81	10

Table 1 Kidney exchange graph instance main characteristics. LP stands for longest path and the root gaps are in % with respect to formulation (5). Here n_a denotes the number of altruistic donors.

	CPU Seconds						Gap (%)				
Instance \ α	0.75	0.8	0.85	0.9	0.95	(5)	0.75	0.8	0.85	0.9	0.95
graph20a1	3	6	7	10	16	11	31	4	4	1	0
graph30a2	14	32	40	79	25	T	44	3	1	0	0
graph40a2	119	345	635	115	608	T	55	1	0	-1	-1
graph50a3	1198	128	407	590	114	T	49	0	-1	-1	-1
graph60a3	233	386	534	1456	216	T	71	0	-1	-1	-1
graph70a4	T	1521	1784	729	662	T	72	0	0	-1	-1
graph80a4	T	T	T	492	T	T	74	1	20	0	–

Table 2 Solution CPU seconds and and optimality gap for approximate MILP (8) for the "graph" instances.

cycle prevention inequalities. We then evaluated a layered formulation that effectively solved the problem within a guaranteed approximation gap for larger instances. According to our experiments the approximation gap in practice usually appeared much smaller than the theoretical guarantee.

The problem considered in the current paper is strongly NP-hard in general graphs. In contrast, computing a path that maximizes the similar yet different all-or-nothing objective remains NP-hard even in directed acyclic graphs

Instance \ α	CPU Seconds 0.75	0.8	0.85	0.9	0.95	(5)	Gap (%) 0.75	0.8	0.85	0.9	0.95
MD12unit	1	1	1	1	4	1	9	0	0	0	0
MD12stoch	1	1	1	1	4	1	7	0	0	0	0
MD19unit	1	1	1	2	6	1	27	0	0	0	0
MD19stoch	1	1	1	3	18	1	26	0	0	0	0
MD21unit	0	1	1	2	6	1	1	0	1	0	0
MD21stoch	1	1	1	2	4	1	24	0	0	0	0
MD23unit	1	1	1	2	3	1	46	0	0	0	0
MD23stoch	1	1	1	1	2	1	37	2	3	2	0
MD43unit	8	28	43	57	37	2	43	0	0	0	0
MD43stoch	10	27	37	54	53	2	44	10	0	0	0
MD44unit	4	10	13	17	25	3	62	0	0	0	0
MD44stoch	4	11	13	17	13	1	52	1	1	1	0
MD51unit	15	38	49	107	46	2	37	0	0	0	0
MD51stoch	11	39	57	114	73	1	30	0	0	0	0
MD60unit	8	26	34	81	55	2	48	0	0	0	0
MD60stoch	10	21	31	59	70	1	34	2	2	0	0
MD61unit	15	33	36	72	54	1	54	0	0	0	0
MD61stoch	11	29	35	74	44	2	42	0	0	0	0
MD66unit	12	30	38	72	39	73	24	0	0	0	0
MD66stoch	18	64	121	170	62	2	59	3	3	0	0
MD70unit	7	19	28	60	27	7	58	0	0	0	0
MD70stoch	6	21	37	58	45	1	50	3	2	2	0
MD81unit	564	731	210	207	557	T	77	0	0	0	0
MD81stoch	T	T	T	T	T	T	63	100	27	8	1
MD84unit	367	1469	1524	T	T	T	54	1	0	1	1
MD84stoch	332	603	715	521	T	T	49	7	5	1	0
MD125unit	409	1200	1343	M	M	T	64	-1	-1	–	–
MD125stoch	M	M	M	M	M	T	–	–	–	–	–

Table 3 Solution CPU seconds and optimality gap for approximate MILP (8) for the MD instances.

(DAGs); see [13]. Although it is not difficult to extend the FPTAS suggested for the all-or-nothing objectives to the current problem, it is a question of theoretical interest for future work to determine whether the current problem remains NP-hard in a DAG.

While robust optimization models have been considered where the existence of edges and the value of exchanges are subject to uncertainty [5, 21], the failure-aware model considered in this paper could be extended to a distributionally robust model in which the edge success probabilities are subject to uncertainty. This is also a subject for future work. Finally, computational techniques such as the network layering that is applied to our MIP could potentially be applied to other kidney exchange formulations such as ones appearing in [7].

References

1. D. J. Abraham, A. Blum, and T. Sandholm. Clearing algorithms for barter exchange markets: enabling nationwide kidney exchanges. In J. K. MacKie-Mason, D. C. Parkes,

and P. Resnick, editors, *Proceedings 8th ACM Conference on Electronic Commerce (EC-2007), San Diego, California, USA, June 11-15, 2007*, pages 295–304. ACM, 2007.
2. R. Anderson, I. Ashlagi, D. Gamarnik, and A. E. Roth. Finding long chains in kidney exchange using the traveling salesman problem. *Proceedings of the National Academy of Sciences*, 112(3):663–668, 2015.
3. I. Ashlagi, D. Gamarnik, M. A. Rees, and A. E. Roth. The need for (long) chains in kidney exchange. Technical report, National Bureau of Economic Research, 2012.
4. I. Ashlagi, D. S. Gilchrist, A. E. Roth, and M. A. Rees. Nonsimultaneous chains and dominos in kidney- paired donation-revisited. *American Journal of Transplantation*, 11:984–994, 2011.
5. M. Carvalho, X. Klimentova, K. Glorie, A. Viana, and M. Constantino. Robust models for the kidney exchange problem. *INFORMS Journal on Computing*, 2020.
6. Y. Chen, Y. Li, J. D. Kalbfleisch, Y. Zhou, A. Leichtman, and P. X.-K. Song. Graph-based optimization algorithm and software on kidney exchanges. *IEEE Transactions on Biomedical Engineering*, 59(7):1985–1991, 2012.
7. M. Constantino, X. Klimentova, A. Viana, and A. Rais. New insights on integer-programming models for the kidney exchange problem. *European Journal of Operational Research*, 231(1):57–68, 2013.
8. 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 Multiagent Systems-Volume 2*, pages 711–718, 2012.
9. J. P. Dickerson, A. D. Procaccia, and T. Sandholm. Failure-aware kidney exchange. *Management Science*, 65(4):323–340, 2019.
10. I. Dunning, J. Huchette, and M. Lubin. JuMP: A modeling language for mathematical optimization. *SIAM Review*, 59(2):295–320, 2017.
11. M. R. Garey and D. S. Johnson. *Computers and Intractability: A Guide to the Theory of NP-Completeness*. W. H. Freeman, 1979.
12. K. M. Glorie, J. J. van de Klundert, and A. P. M. Wagelmans. Kidney exchange with long chains: An efficient pricing algorithm for clearing barter exchanges with branch-and-price. *Manufacturing & Service Operations Management*, 16:498–512, 2014.
13. N. Goldberg and M. Poss. Maximum probabilistic all-or-nothing paths. *European Journal of Operational Research*, 283:279–289, 2020.
14. N. Goldberg and M. Poss. PSPP – GitHub repository. `https://github.com/noamgold/PSPP-Data/`, 2021.
15. N. Goldberg and G. Rudolf. On the complexity and approximation of the maximum expected value all-or-nothing subset. *Discrete Applied Mathematics*, 283:1–10, 2020.
16. L. E. N. Gouveia, M. Leitner, and M. Ruthmair. Layered graph approaches for combinatorial optimization problems. *Computers and Operations Research*, 102:22–38, 2019.
17. X. Klimentova, J. P. Pedroso, and A. Viana. Maximising expectation of the number of transplants in kidney exchange programmes. *Computers & Operations Research*, 73:1–11, 2016.
18. Y. Li, P. X.-K. Song, Y. Zhou, A. B. Leichtman, M. A. Rees, and J. D. Kalbfleisch. Optimal decisions for organ exchanges in a kidney paired donation program. *Statistics in Biosciences*, 6:85–104, 2014.
19. D. F. Manlove and G. O'Malley. Paired and altruistic kidney donation in the UK: algorithms and experimentation. *ACM J. Exp. Algorithmics*, 19(1), 2014.
20. N. Mattei and T. Walsh. Preflib: A library for preferences http://www. preflib. org. In *International Conference on Algorithmic DecisionTheory*, pages 259–270. Springer, 2013.
21. D. C. McElfresh, H. Bidkhori, and J. P. Dickerson. Scalable robust kidney exchange. In *The Thirty-Third AAAI Conference on Artificial Intelligence, AAAI 2019, The Thirty-First Innovative Applications of Artificial Intelligence Conference, IAAI 2019, The Ninth AAAI Symposium on Educational Advances in Artificial Intelligence, EAAI 2019, Honolulu, Hawaii, USA, January 27 - February 1, 2019*, pages 1077–1084. AAAI Press, 2019.
22. M. A. Rees, J. E. Kopke, R. P. Pelletier, D. L. Segev, M. E. Rutter, A. J. Fabrega, J. Rogers, O. G. Pankewycz, J. Hiller, A. E. Roth, T. Sandholm, M. U. Ünver, and R. A. Montgomery. A nonsimultaneous, extended, altruistic-donor chain. *The New England Journal of Medicine*, 360:1096–1101, 2009.

23. 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.
24. S. L. Saidman, A. E. Roth, T. Sönmez, M. U. Ünver, and F. L. Delmonico. Increasing the opportunity of live kidney donation by matching for two-and three-way exchanges. *Transplantation*, 81(5):773–782, 2006.
25. L. Taccari. Integer programming formulations for the elementary shortest path problem. *European Journal of Operational Research*, 252(1):122–130, 2016.