Algorithms for Inverse Optimization Problems

Sara Ahmadian
Combinatorics and Optimization, Univ. Waterloo, Waterloo, ON N2L 3G1, Canada
sahmadian@uwaterloo.ca

Umang Bhaskar[1]
Tata Institute of Fundamental Research, Mumbai, India 400 005
umang@tifr.res.in

Laura Sanità
Combinatorics and Optimization, Univ. Waterloo, Waterloo, ON N2L 3G1, Canada
sanita@uwaterloo.ca

Chaitanya Swamy[2]
Combinatorics and Optimization, Univ. Waterloo, Waterloo, ON N2L 3G1, Canada
cswamy@uwaterloo.ca

Abstract

We study *inverse optimization problems*, wherein the goal is to map *given solutions* to an underlying optimization problem to a *cost vector* for which the given solutions are the (unique) optimal solutions. Inverse optimization problems find diverse applications and have been widely studied. A prominent problem in this field is the *inverse shortest path* (ISP) problem [9, 3, 4], which finds applications in shortest-path routing protocols used in telecommunications. Here we seek a cost vector that is positive, *integral*, induces a set of given paths as the unique shortest paths, and has minimum ℓ_∞ norm. Despite being extensively studied, very few algorithmic results are known for inverse optimization problems involving integrality constraints on the desired cost vector whose norm has to be minimized.

Motivated by ISP, we initiate a systematic study of such integral inverse optimization problems from the perspective of designing polynomial time approximation algorithms. For ISP, our main result is an *additive* 1-approximation algorithm for multicommodity ISP with node-disjoint commodities, which we show is *tight* assuming $P \neq NP$. We then consider the integral-cost inverse versions of various other fundamental combinatorial optimization problems, including min-cost flow, max/min-cost bipartite matching, and max/min-cost basis in a matroid, and obtain tight or nearly-tight approximation guarantees for these. Our guarantees for the first two problems are based on results for a broad generalization, namely *integral inverse polyhedral optimization*, for which we also give approximation guarantees. Our techniques also give similar results for variants, including ℓ_p-norm minimization of the integral cost vector, and distance-minimization from an initial cost vector.

2012 ACM Subject Classification Theory of computation → Network optimization, Theory of computation → Approximation algorithms analysis, Mathematics of computing → Network flows

Keywords and phrases Inverse optimization, Shortest paths, Approximation algorithms, Linear programming, Polyhedral theory, Combinatorial optimization

Digital Object Identifier 10.4230/LIPIcs.ESA.2018.1

[1] Funded in part by a Ramanujan Fellowship. Part of this work was done while visiting U. Waterloo.
[2] Supported in part by NSERC grant 327620-09 and an NSERC Discovery Accelerator Supplement Award.

26th Annual European Symposium on Algorithms (ESA 2018).
Editors: Yossi Azar, Hannah Bast, and Grzegorz Herman; Article No. 1; pp. 1:1–1:14
Leibniz International Proceedings in Informatics
Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

1 Introduction

Consider the following problem, adapted from [4], faced by the administrator of a telecommunication network. The administrator seeks to impose a desired routing pattern (e.g., one that distributes traffic along multiple paths to minimize congestion) under a given underlying routing protocol. Many routing protocols – OSPF (open-shortest-path-first), IS-IS etc. – use shortest-path routing, with path lengths defined as the sum of link lengths that are set by the administrator, where the link lengths must typically be positive integers that can be stored using a limited number of bits (e.g., in IS-IS, they must be at most 63). Thus, the administrator must choose small, positive, integer link lengths so that the resulting shortest paths *coincide* with the prescribed paths (thus ensuring that we utilize precisely these paths).

This is an *inverse shortest path* (ISP) problem (which also arises in seismic tomography, traffic modeling, and network tolling [9, 19, 8, 12, 13]), a prominent problem from the rich class of *inverse optimization problems*, wherein we are *given solutions* to an underlying optimization problem and we *seek a cost vector* under which the given solutions constitute the (unique) optimal solutions. Since we map solutions to a suitable cost vector, this is termed inverse optimization. Inverse optimization problems find applications in a variety of domains including telecommunication routing [3, 4], seismic and medical tomography [9, 19, 23], traffic modeling and network tolling [12, 13, 9, 8], and portfolio optimization [18]. They also arise in the domain of *revealed-preference theory* in economics [24], which seeks to understand when observations can be attributed to behavior consistent with game-theoretic models. As these examples indicate, inverse optimization problems typically have two primary objectives: (a) *parameter estimation*, where we seek to infer certain parameters of a system that are consistent with a set of observations (e.g., seismic and medical tomography, revealed preference theory); and (b) *solution imposition*, where the goal is to (minimally) perturb the system parameters so as to enforce a set of solutions (e.g., the telecommunication routing application mentioned above, and network tolling where we want to find (minimal) edge tolls imposing a given routing pattern as an equilibrium flow).

Motivated by ISP, we consider inverse optimization problems wherein the desired cost vector c is required to be positive, *integral*, and induce the given subset $\mathcal{S}$ of solutions as the *unique optimal solutions* to the underlying optimization problem; we call these problems *integral inverse optimization problems*. We primarily consider the objective of minimizing $\|c\|_\infty$, but our results also yield guarantees for the objective of minimizing the perturbation $\|c - c^{(0)}\|_\infty$ of a given "base" cost vector $c^{(0)}$, which is frequently considered in the inverse-optimization literature. Uniqueness can be important because we may want to explain/impose $\mathcal{S}$ without introducing spurious solutions (i.e., "we get precisely what we bargain for"), and integrality is, in many cases, a desirable or necessary practical consideration (as in the telecommunication-routing setting). Despite extensive literature, very few algorithmic results are known for inverse optimization problems involving integrality constraints on the desired cost vector whose norm, or deviation from a given cost vector $c^{(0)}$, is to be minimized; we only know of [3, 4] that address this, both in the context of ISP.

Our contributions and results. We initiate a systematic study of integral inverse optimization problems from the perspective of designing polynomial time (approximation) algorithms. We focus on the inverse versions of various combinatorial optimization problems, a natural starting point to investigate integral inverse optimization problems. As our results demonstrate, even for such problems, wherein the underlying optimization problem is well structured and polytime-solvable, the resulting integral inverse optimization problems are

Table 1 Summary of our main results. These are stated for the implicit model, wherein the solution-set is specified implicitly by listing its support set. Most of our guarantees also hold in the explicit model.

Problem	**Our results**
Inverse shortest path (ISP)	polytime
Node-disjoint multicommodity ISP	additive 1-approximation; problem is *NP*-hard (Previous work gives multiplicative $O(\|V\|)$-approx.)
Inverse polyhedral optimization (IOpt-Poly) with TU constraint matrix	additive 1-approximation for minimization (IMin-Poly) multiplicative 2-approx. for maximization (IMax-Poly)
IOpt-Poly with $\{0,1\}$ matrix A (r, k = row, column sparsity of A)	multiplicative $\widetilde{O}\left(\sqrt{\min\{r,k\}}\right)$-approximation additive factors of k: IMin-Poly; $(k-1)$: IMax-Poly
Inverse versions of min-cost flow and min/max-cost bipartite matching	additive 1-approx.; follows from results for IOpt-Poly
Inverse matroid-basis optimization	polytime (for minimization and maximization)

quite non-trivial and exhibit an interesting range of possibilities in terms of positive (approximation) algorithmic results and hardness of approximation results. We obtain tight or nearly-tight guarantees for a variety of integral inverse optimization problems, including the well-studied inverse shortest path (ISP) problem. Our salient contributions are as follows; Table 1 summarizes our main results.

- We begin by considering ISP (Section 3). We show that the single-commodity version (Section 3.1), wherein $\mathcal{S}$ is a subset of $s \rightsquigarrow t$ paths in a directed graph, is *polytime solvable* (Theorem 5). We then consider *multicommodity* ISP, the generalization where we have multiple commodities, each specified by an (s_i, t_i) pair of nodes and a subset $\mathcal{S}_i$ of $s_i \rightsquigarrow t_i$ paths, and we seek positive, integral edge costs that ensure that $\mathcal{S}_i$ is the unique set of shortest $s_i \rightsquigarrow t_i$ paths for each commodity i. We *resolve* the status of node-disjoint multicommodity ISP, where the $\mathcal{S}_i$s correspond to node-disjoint subgraphs (Section 3.2): we devise an *additive* 1-*approximation* algorithm (Theorem 6), which is the best possible guarantee (if $P \neq NP$) since we show that this node-disjoint version is *NP*-hard (Theorem 7). Our proof also shows that it is *NP*-hard to obtain a multiplicative $\left(\frac{3}{2} - \epsilon\right)$-approximation for multicommodity ISP, for any $\epsilon > 0$.
  Our results improve upon the previous-best multiplicative $O(|V|)$-approximation guarantees for these problems, which follow from the work of [3, 4]. The algorithms in [3, 4] are for multicommodity ISP, but they apply to the restrictive setting where $\mathcal{S}_i$ includes a *single* $s_i \rightsquigarrow t_i$-path for every commodity; moreover, they do not yield improved guarantees even for the special cases of single-commodity ISP or node-disjoint multicommodity ISP. We also improve upon the factor $\frac{9}{8}$ hardness-of-approximation guarantee in [4].
- Motivated by the fact that many combinatorial optimization problems can be cast as polyhedral optimization problems, in Section 4, we consider a broad generalization of integral inverse discrete optimization, namely *integral inverse polyhedral optimization*. Here, we are given a *polytope* $\mathcal{P} \subseteq \mathbb{R}^n$ explicitly, and the set $\mathcal{S}$ is replaced by a set X of *extreme points* of $\mathcal{P}$; we seek a positive, integral cost vector $c \in \mathbb{Z}^n$ that induces X as the unique set of extreme-point optimal solutions to the problem of optimizing (minimizing or maximizing) $c^T x$ over $x \in \mathcal{P}$. We obtain approximation guarantees for integral inverse polyhedral optimization that depend on the structure of the constraint matrix A defining $\mathcal{P}$. When A is *totally unimodular* (TU), we obtain an *additive* 1-, *or multiplicative* 2-*approximation* (see Theorem 8), and for a general $\{0,1\}$ matrix A, our approximation factor depends on the *row and/or column sparsity* of A (see Theorem 9). As *corollaries*

of these results, we obtain *additive* 1-*approximation* algorithms for the integral inverse versions of min-cost flows and max/min-cost bipartite matchings.

Similar to ISP, *integral inverse min-cost flow* (IMCF) captures the optimization problem encountered in the context of *spanning-tree protocols* (STPs) – e.g., rapid STP, multiple STP etc. – which route using a *shortest-path tree* rooted at a given node s under the assigned link weights; enforcing a prescribed routing tree rooted at s by choosing small, positive, integer link lengths is then an IMCF problem, and in fact, the special case involving a single source and infinite (or equivalently, very large) capacities. This link-weight assignment problem was studied in [15, 16], who prove upper bounds on the optimum value (in a more general setting). We show that this single-source IMCF problem is polytime solvable, which implies that *we can solve this link–weight assignment problem in polynomial time.*

It is illuminating to view integral inverse polyhedral optimization (IOpt-Poly) geometrically. The set of cost vectors that yield X as extreme-point optimal solutions in $\mathcal{P}$, form a *polyhedral cone*; a cost vector in the *interior* of this cone yields X as the unique set of extreme-point optimal solutions. Thus, the goal in IOpt-Poly is to find a *shortest* (in $\|.\|_\infty$-norm) *positive, integral vector* in the interior of this cone (if one exists). Viewed from this perspective, integral inverse polyhedral optimization can be seen as a problem in the field of *geometry of numbers* and in the same vein as the important shortest-vector-problem in lattices. We believe that this geometric connection makes IOpt-Poly an appealing problem of independent interest meriting further study.

- In Section 5, we consider integral inverse matroid-basis optimization. Here, $\mathcal{S}$ is a collection of bases of a matroid, and we seek positive, integer costs on the elements under which $\mathcal{S}$ is the unique set of optimal bases. We give a polytime algorithm for this problem (Theorem 12).

Our techniques are versatile and yield results for various variants (see Section 6), including, most notably, integral inverse optimization under two other commonly considered objectives in the literature: (1) ℓ_p*-norm minimization*, where we seek to minimize $\|c\|_p$; and (2) *distance minimization*, where we seek to minimize the perturbation $\|c - c^{(0)}\|_\infty$ of an integral "base" vector $c^{(0)}$. Our results typically also hold in an *implicit* model, where the input specifies a (potentially exponential-size) set $\mathcal{S}$ implicitly by listing the elements in terms of its support.

Most prior results on inverse optimization, with the exception of ISP, are obtained in the setting where $\mathcal{S}$ consists of a single solution $\hat{x}$ (with [26, 28] being exceptions), which is not required to be the unique optimal solution, and the objective is to minimize $\|c - c^{(0)}\|_\infty$ (or $\|c - c^{(0)}\|_p$ for some other p), with c fractional. This setting is significantly simpler than the integral inverse optimization setting we consider. In particular, it is not hard to see that, as noted in [2], even for a general inverse polyhedral optimization problem, one can: (a) utilize the complementary slackness (CS) conditions from LP theory to encode the problem of finding a suitable cost vector c as another LP (or a convex program for ℓ_p norms); or (b) use the ellipsoid method to solve the LP that directly encodes that $\hat{x}$ has optimal objective value among all $x \in \mathcal{P}$, given an optimization/separation oracle for $\mathcal{P}$. This work therefore focuses on obtaining faster algorithms for the integral inverse optimization problem.

In contrast, in the integral inverse optimization setting, two distinct sources of difficulty arise that do not appear in the above setup. First, even computing a suitable fractional cost vector is non-trivial due to the uniqueness constraint. For instance, in inverse polyhedral optimization, this entails discerning if the given solutions form the extreme points of a face of the given polytope, and determining how to encode, and separate over, the constraints enforcing uniqueness. Second, rounding a fractional cost vector poses the difficulty that we need to *coordinate* things so as to simultaneously ensure that all solutions in $\mathcal{S}$ *continue to have the same cost*, and solutions not in $\mathcal{S}$ *remain non-optimal solutions.* This creates unique challenges, and we leverage tools from optimization theory, polyhedral theory, and recent results in discrepancy theory to circumvent these difficulties and obtain our results. An interesting and notable implication of our work is that, in many cases, *imposing integral costs does not significantly impact the achievable performance guarantees.*

Our array of results allude to the richness of integral inverse optimization problems. While our work makes significant progress towards understanding these problems, it also opens up various directions for further research, such as investigating the inverse-optimization versions of *NP*-hard optimization problems.

Related work. Inverse problems were initially extensively studied in geophysics for the estimation of model parameters (see, e.g., [23]). Since then there has been a great deal of work in inverse optimization in the optimization community (see, e.g., the survey [17]). In the optimization community, Burton and Toint [9] (see also [8]) were the first to consider inverse optimization problems. They introduced the the ℓ_2-norm distance-minimization variant of ISP,where we seek to minimize $\|c - c^{(0)}\|_2$, where $c^{(0)}$ is a base vector, while allowing for fractional cost vectors, and do not require the given paths to be the unique shortest paths. They motivate ISP from applications in traffic modeling and seismic tomography, and suggest the extension to the ℓ_1 and ℓ_∞ norms. Ben-Ameur and Gordin [3] and Bley [4] study (among other problems) ISP under the constraints of positive, integral edge costs, and uniqueness of the given paths (i.e., integral ISP), motivated by its applications to shortest-path routing protocols. These give algorithms having multiplicative approximation ratios of $O\big(\min\{|V|/2, (\text{maximum length of a given path})\}\big)$, and [4] also shows that it is *NP*-hard to obtain an approximation ratio better than 9/8. Other ISP variants have also been investigated [2, 4, 7, 11, 12, 13, 25].

Following initial work on inverse shortest paths, algorithms were developed for the inverse-optimization versions of other combinatorial optimization problems, such as minimum spanning tree, min-cost flow, min-cut, matroid intersection, and general inverse polyhedral optimization (also called *inverse linear programming* [29, 30]); see [17] for details. Most of this work pertains to the distance-minimization problem when we allow fractional costs, and only a single solution is given ([26, 28] are exceptions that consider multiple solutions) that is not required to be the unique optimal solution. These papers focus on developing fast combinatorial algorithms. Ahuja and Orlin [2] unify and generalize many of these results. They note that inverse polyhedral optimization can be solved in the above setting by solving a suitable LP: a compact LP encoding this can be obtained by utilizing the CS conditions, and even the (huge) LP that directly encodes that the given solution be optimal can be solved via the ellipsoid method. They show that in various cases, the compact LP leads to an LP similar to the one for optimizing over $\mathcal{P}$, and hence one can obtain combinatorial algorithms for various inverse discrete optimization problems. Similar results were also obtained by [27].

We remark that while we also solve an LP to obtain fractional cost vectors en route to obtaining integral cost vectors, a crucial difference in our setting is that we need to devise suitable ways of encoding (and separating over) the constraint that the costs induce the given (multiple) solutions as the *unique* optimal solutions. Our algorithms for integral inverse polyhedral optimization require either *a face oracle* for $\mathcal{P}$, which determines if the given set X of extreme points forms a face of $\mathcal{P}$, or an oracle that determines if all *maximal/minimal* points on a face of $\mathcal{P}$ have the same cost under a given cost vector. Devising a face oracle is related to the problem of enumerating all vertices (i.e., extreme points) of a polyhedron, or all vertices on its optimal face (under an objective function), with each new vertex being output in polynomial delay. (For instance, we can decide if X forms a face by determining if the minimal face of $\mathcal{P}$ containing X contains at least $|X|+1$ vertices.) Such procedures are known for various polyhedra such as network-flow polyhedra [20], general 0/1 *polytopes* [10], simplicial and simple polyhedra [6, 14], but this is *NP*-hard for general 0/1 polyhedra [5].

2 Problem definitions, notation, and preliminaries

For an integer n, we use $[n]$ to denote $\{1,\ldots,n\}$. Given $z \in \mathbb{R}^E$, and $S \subseteq E$, we use $z(S)$ to denote $\sum_{e\in S} z_e$. We use $\lfloor z \rfloor$ and $\lceil z \rceil$ to denote the vectors $\big(\lfloor z_e \rfloor\big)_{e\in E}$ and $\big(\lceil z_e \rceil\big)_{e\in E}$ respectively.

Inverse discrete optimization. An inverse discrete optimization problem involves an underlying discrete optimization problem specified in terms of a ground set E and a collection $\mathcal{F} \subseteq 2^E$ of feasible solutions, and a subset $\mathcal{S} \subseteq \mathcal{F}$ of feasible solutions to the optimization problem. We seek a cost vector $c \in \mathbb{R}^E$ such that the solutions in $\mathcal{S}$ are the optimal solutions to the underlying optimization problem. Formally, in an *inverse minimization problem*, the underlying optimization problem is a minimization problem, and we seek a cost vector $c \in \mathbb{R}^E$ such that $c(S) = \min_{F\in\mathcal{F}} c(F)$ for all $S \in \mathcal{S}$. In an *inverse maximization problem*, the underlying optimization problem is a maximization problem, and we seek $c \in \mathbb{R}^E$ such that $c(S) = \max_{F\in\mathcal{F}} c(F)$ for all $S \in \mathcal{S}$. More precisely, motivated by applications of the inverse-shortest-path problem in the context of shortest-path network-routing protocols in telecommunication, we impose the following requirements on the cost vector c.

(C1) *Positive, integer costs*: $c_e \geq 1,\ c_e \in \mathbb{Z}$ for all $e \in E$;

(C2) *Unique optimal solutions*: For inverse minimization, we require $c(S) = \min_{F\in\mathcal{F}} c(F) < c(F')$ for all $S \in \mathcal{S}$ and $F' \in \mathcal{F} \setminus \mathcal{S}$; for inverse maximization, we require $c(S) = \max_{F\in\mathcal{F}} c(F) > c(F')$ for all $S \in \mathcal{S}$ and $F' \in \mathcal{F} \setminus \mathcal{S}$;

Our goal is to find a vector c satisfying 1 and 2 that minimizes $\|c\|_\infty$. We call this an *integral inverse optimization problem*; we drop "integral" when it is clear from the context.

The uniqueness condition 2 is often important in applications, where the inverse optimization problem is used to infer or perturb some system parameters so as to explain or impose a set $\mathcal{S}$ of observations, since we would like to do so without introducing spurious solutions. We impose $c \geq 1$ as a normalization requirement: this prevents one from arbitrarily scaling a vector satisfying 2 to obtain another feasible solution. Integrality is a discretization condition that ensures that we are optimizing over a *closed set* (note that 2 leads to an open feasible region). (Without an underlying objective such as minimizing $\|c\|_\infty$, 1 becomes redundant as one can always scale a rational vector c to satisfy 1.)

We allow for specifying exponentially large (in the natural input size) solution sets $\mathcal{S}$ (thus obtaining greater modeling power), by also considering the following *implicit* model for specifying $\mathcal{S}$: we specify a set U of elements, which implicitly describes the set $\mathcal{S} = \{S \in \mathcal{F} : S \subseteq U\}$ of feasible solutions. For example, in the implicit version of inverse shortest paths, U is a set of arcs and $\mathcal{S}$ comprises all $s \rightsquigarrow t$ paths contained in U; so a solution is a positive, integral cost vector such that the collection of shortest $s \rightsquigarrow t$ paths is precisely $\mathcal{S}$. Our results typically apply to both models, and the underlying arguments are similar.

Our techniques are versatile and yield results for other variants of the above integral inverse optimization problem such as, most notably,

(1) the *ℓ_p-norm version*: find a vector c satisfying 1, 2 that minimizes $\|c\|_p$

(2) the *distance-minimization version* (with ℓ_∞ norm): the input specifies a "base" vector $c^{(0)} \in \mathbb{Z}_+^E$, and we seek a cost vector c satisfying 1, 2 that minimizes $\|c - c^{(0)}\|_\infty$.

At a high level, this follows because our results are obtained by first obtaining an (near-) optimal fractional cost vector c^* satisfying 1, 2 via an LP (or, for ℓ_p-norms where $1 < p < \infty$, via a convex program) and then rounding it to a feasible integral vector $\tilde{c}$ while introducing an *additive* $O(1)$ rounding error; this rounding error easily translates to a multiplicative approximation for problems (1), (2). The following theorem makes this precise.

▶ **Theorem 1.** *Let $c^* \in \mathbb{R}^E$ be a cost vector satisfying $c^*_e \geq 1\ \forall e \in E$. Let $\tilde{c} \in \mathbb{Z}^E$ be a vector satisfying 1, 2.*

- **(i)** *Let $O^*{}_p := \min\left\{\|c\|_p : c \text{ satisfies } 1,2\right\}$. Suppose that $\|c^*\|_p \leq O^*{}_p + \epsilon$, and $\tilde{c}_e \leq \alpha c^*_e + \beta$ for all $e \in E$. Then, $\|\tilde{c}\|_p \leq (\alpha+\beta)(1+\epsilon)O^*{}_p$; if $\epsilon < \left(\frac{1}{2(\alpha+\beta)O^*{}_p}\right)^p$, this implies that $\|\tilde{c}\|_p \leq \lceil \alpha+\beta \rceil O^*{}_p$.*
- **(ii)** *Let $O^*_{\mathsf{dist}} := \min\left\{\|c - c^{(0)}\|_\infty : c \text{ satisfies } 1,2\right\}$. Suppose that $O^*_{\mathsf{dist}} > 0$, $\|c^* - c^{(0)}\|_\infty \leq O^*_{\mathsf{dist}}$, and $\|\tilde{c} - c^*\|_\infty < \beta$. Then, $\|\tilde{c} - c^{(0)}\|_\infty \leq O^*_{\mathsf{dist}} + \lceil \beta \rceil - 1 \leq \lceil \beta \rceil O^*_{\mathsf{dist}}$.*

Inverse polyhedral optimization. Many combinatorial optimization problems have convenient polyhedral descriptions and can be modeled via linear programs that have integral optimal solutions; this indeed holds for the problems whose integral inverse optimization versions we investigate. With this in mind, we consider the following general inverse polyhedral optimization problem, which is a natural abstraction of an inverse discrete optimization problem. We are given a *polytope* $\mathcal{P} \subseteq \mathbb{R}^E_+$ with explicitly specified constraints, and a collection $X \subseteq \mathcal{P}$ of *extreme points* of $\mathcal{P}$. In *integral inverse polyhedral minimization* (IMin-Poly), we seek a cost vector $c \in \mathbb{R}^E$ that minimizes $\|c\|_\infty$ and satisfies 1, and (C2): $c^T\hat{x} = \min_{x \in \mathcal{P}} c^T x < c^T x'$ for every $\hat{x} \in X$ and every extreme point x' of $\mathcal{P}$ not in X. Similarly, in *integral inverse polyhedral maximization* (IMax-Poly), we seek $c \in \mathbb{R}^E$ that minimizes $\|c\|_\infty$ and satisfies 1, and (C2'): $c^T\hat{x} = \max_{x \in \mathcal{P}} c^T x > c^T x'$ for every $\hat{x} \in X$ and every extreme point x' of $\mathcal{P}$ not in X. If the underlying discrete optimization problem is captured by the problem of optimizing over $\mathcal{P}$ (e.g., if extreme points of $\mathcal{P}$ correspond to feasible solutions to the discrete optimization problem), then this integral inverse polyhedral optimization problem captures the integral inverse discrete optimization problem defined earlier. As before, we also consider the implicit version, wherein we are given $U \subseteq E$, which implicitly specifies $X := \big\{\text{extreme points } \hat{x} \text{ of } \mathcal{P} \text{ s.t. } \{e : \hat{x}_e > 0\} \subseteq U, \hat{x} \text{ is maximal/minimal in } \mathcal{P}\big\}$. By $\hat{x}$ being maximal in $\mathcal{P}$, we mean that there is no $x \in \mathcal{P}$ such that $x \geq \hat{x}$, $x \neq \hat{x}$; minimality is similarly defined. The set X must be maximal for IMax-Poly, and minimal for IMin-Poly, as only such points can be optimal solutions since $c > 0$.

We say that X forms a face of $\mathcal{P}$, if X is precisely the set of extreme points of some face of $\mathcal{P}$. Integral inverse polyhedral optimization can be stated geometrically as: determine if X forms a face, say F, of $\mathcal{P}$, and if so, find a positive, integral vector (if one exists) of minimum ℓ_∞ norm in the interior of the polyhedral cone of vectors yielding F as the optimal face.

Difference systems. We often need to obtain a solution to a system of constraints of the following form, called a *difference system with bounds*, involving n variables $z_1, \ldots, z_n$:

$$z_i - z_j \leq d_{ij} \quad \forall (i,j) \in A, \qquad z_i \geq \ell_i \quad \forall i \in L, \qquad z_i \leq u_i \quad \forall i \in U \tag{1}$$

where $A \subseteq [n] \times [n]$, $L, U \subseteq [n]$. The d_{ij}s can be arbitrary, so (1) can also incorporate constraints of the form $z_i - z_j \geq d_{ij}$. The following useful result is well known (see, e.g., [1]).

▶ **Theorem 2.** *We can find a feasible solution to a difference system* (1)*, or detect it is infeasible, by computing a shortest path in a digraph with $|A| + |L| + |U|$ arcs, $n+1$ nodes. If the data is integral, and* (1) *is feasible, this yields an integer-valued feasible solution.*

Further, given costs $\{b_i\}_{i=1}^n$, we can solve a min-cost flow problem to find an optimal solution to the following LP: minimize $\sum_i b_i z_i$ subject to (1). If this LP has an optimal solution and the d_{ij}s, ℓ_is and u_is are integral, this yields an integer-valued optimal solution.

3 The inverse shortest path problem

In the *integral inverse shortest path* (ISP) problem, we are given a directed graph $D = (V, E)$, terminals $s, t \in V$, and a collection $\mathcal{S}$ of simple $s \rightsquigarrow t$ paths; we seek positive, integral edge costs $\{c_e\}_{e\in E}$ such that the paths in $\mathcal{S}$ are the unique shortest $s \rightsquigarrow t$ paths under these edge costs, so as to minimize $\|c\|_\infty = \max_e c_e$. In *multicommodity* ISP, we have k commodities, with each commodity $i = 1, \ldots, k$ specified by a pair $s_i, t_i \in N$ of terminals, and a collection $\mathcal{S}_i$ of $s_i \rightsquigarrow t_i$ paths. We seek positive, integral edge costs $\{c_e\}_{e\in E}$ minimizing $\|c\|_\infty$ such that for each commodity $i = 1, \ldots, k$, the paths in $\mathcal{S}_i$ are the unique $s_i \rightsquigarrow t_i$ shortest paths under these edge costs. Clearly, ISP is the special case where $k = 1$. In the implicit version of multicommodity ISP, we are given edge-sets $E^1, \ldots, E^k$, which implicitly defines $\mathcal{S}_i$ to be the collection of all $s_i \rightsquigarrow t_i$ paths in E^i.

We show that ISP is *polytime solvable* (Section 3.1). For multicommodity ISP (Section 3.2), we devise an *additive* 1-approximation algorithm in the setting where the $\mathcal{S}_i$s correspond to node-disjoint subgraphs. Our guarantee is *tight*, since we show that (even) this special case of multicommodity ISP is *NP*-hard to approximate within a factor better than $\frac{3}{2}$. Previously, only a multiplicative $O(|V|)$-approximation guarantee was known for these problems [3, 4], and a factor $\frac{9}{8}$ hardness-of-approximation was known for general multicommodity ISP [4]. In Section 6, we show that our techniques yield results for various other ISP variants including: (1) the ℓ_p-norm minimization version; (2) the distance minimization version; and (3) variants involving shortest-$s_i \rightsquigarrow t_i$-path distances in the objective or constraints.

3.1 A polynomial time exact algorithm for ISP

We may assume that every edge in D lies on some $s \rightsquigarrow t$ path, as otherwise we can assign it cost 1, and so can simply delete the edge. Let O^* denote the optimal value of the ISP instance. We utilize the following well-known properties of shortest paths.

▶ **Claim 3.** *Let $D = (N, A)$ be a digraph with nonnegative edge costs $\{c_e\}_{e\in A}$, and $s, t \in N$. Suppose that every edge of A lies on some $s \rightsquigarrow t$ path. Let $\mathcal{S}$ be a collection of $s \rightsquigarrow t$ paths.*

(i) *$\mathcal{S}$ consists of shortest $s \rightsquigarrow t$ paths (under c) iff there are* node potentials *$\{y_v\}_{v\in N}$ such that:*

$$y_v - y_u \le c_{u,v} \quad \text{for all } (u,v) \in A, \qquad y_v - y_u = c_{u,v} \quad \text{for all } (u,v) \in \bigcup_{P\in\mathcal{S}} P. \quad (2)$$

(ii) *Node potentials satisfying (2) exist iff the node potentials obtained by setting $y_v =$ (shortest-$s \rightsquigarrow v$-path distance) $\forall v$, satisfy (2).*

(iii) *$\mathcal{S}$ comprises shortest $s \rightsquigarrow t$ paths iff every $s \rightsquigarrow t$ path $Q \subseteq \bigcup_{P\in\mathcal{S}} P$ is shortest $s \rightsquigarrow t$ path.*

If the input is in the explicit model (i.e., $\mathcal{S}$ is explicitly given), define $E^1 := \bigcup_{P\in\mathcal{S}} P$. By Claim 3 (iii), an ISP instance in the explicit model is feasible only if $\mathcal{S}$ includes all $s \rightsquigarrow t$ paths contained in E^1. Also, since we seek positive edge costs, E^1 must be acyclic (a directed cycle must have cost 0 due to (2)), otherwise the ISP instance is infeasible. In the explicit model, we first check if E^1 contains an $s \rightsquigarrow t$ path not in $\mathcal{S}$. This can be checked in polynomial time in various ways: for instance, we can use topological sort to count the number of $s \rightsquigarrow t$ paths in E^1 and check if this number is $|\mathcal{S}|$. (We can also use depth-first search and backtracking to enumerate $|\mathcal{S}|+1$ distinct $s \rightsquigarrow t$ paths in polytime (if they exist); see, e.g., [21].)

In the sequel, we assume that the ISP instance meets these feasibility requirements (so the explicit and implicit models coincide). Let $G^1 = (V^1, E^1)$ be the subgraph induced by E^1. We may assume that every edge $e \in E^1$ lies on an $s \rightsquigarrow t$ path contained in E^1 (which holds by definition in the explicit model); otherwise, we can remove e from E^1 and solve the resulting ISP instance. We consider the following LP-relaxation of the problem with the c_es and node potentials $\{y_v\}_{v\in N}$ as variables. (The objective function and constraints are easily linearized.)

$$\min \quad \|c\|_\infty \tag{ISP-P}$$
$$\text{s.t.} \quad \max\{1, y_v - y_u\} \le c_{u,v} \quad \forall (u,v) \in E, \qquad y_v - y_u = c_{u,v} \quad \forall (u,v) \in E^1 \tag{3}$$
$$y_v - y_u + 1 \le c(P) \quad \forall (u,v) \in V^1 \times V^1,\ \forall u \rightsquigarrow v \text{ paths } P \subseteq E \setminus E^1. \tag{4}$$

Constraints (3) follow from Claim 3, and ensure that all $s \rightsquigarrow t$ paths in E^1 are shortest $s \rightsquigarrow t$ paths. Note that if there is no $u \rightsquigarrow v$ path in $E \setminus E^1$, then there is no constraint (4) for (u,v). We argue below that constraints (4) are valid; this follows because (4) encodes that every $s \rightsquigarrow t$ path Q not contained in E^1 has length at least $1 + \min_{s\rightsquigarrow t \text{ path } P: P\subseteq E^1} c(P)$, and with integer edge costs, this is equivalent to the condition that every $s \rightsquigarrow t$ path Q not contained in E^1 is not a shortest $s \rightsquigarrow t$ path.

▶ **Lemma 4.** (ISP-P) *is a relaxation of* ISP.

We can efficiently solve (ISP-P) via the ellipsoid method since we can efficiently separate over constraints (4) when $c \ge 0$ by solving a shortest-path problem. (We can actually avoid the ellipsoid method and obtain a much more efficient algorithm for ISP. We retain the LP-based exposition since this extends easily to multicommodity ISP and other variants of ISP.) If (ISP-P) is infeasible, then the ISP instance is infeasible. Otherwise, let (c^*, y^*) be an optimal solution to (ISP-P). Let $B^* = \|c^*\|_\infty$. Note that $O^* \ge \lceil B^* \rceil$. Our rounding algorithm is quite simple. We first round the $\{y_v^*\}$ node potentials by solving the following difference system:

$$\lfloor y_v^* - y_u^* \rfloor \le \psi_v - \psi_u \le \lceil y_v^* - y_u^* \rceil \qquad \text{for all } (u,v) \in V^1 \times V^1.$$

Notice that $\psi = y^*$ is a feasible solution to this difference system, so since the constant terms in the above inequalities are integers, it has a feasible integer solution $\tilde{y}$ (Theorem 2). We set edge costs $\tilde{c}_{u,v} = \tilde{y}_v - \tilde{y}_u$ for all $(u,v) \in E^1$, and $\tilde{c}_{u,v} = \lceil c^*_{u,v} \rceil$ for all $(u,v) \in E \setminus E^1$.

▶ **Theorem 5.** *Vector $\tilde{c}$ satisfies $\lfloor c^* \rfloor \le \tilde{c} \le \lceil c^* \rceil$, and is hence an optimal solution to* ISP.

3.2 Multicommodity ISP with node-disjoint subgraphs

We now consider multicommodity ISP, where the edges in the $\mathcal{S}_i$s induce node-disjoint subgraphs. More precisely, if the input is in the explicit model, define $E^i := \bigcup_{P\in\mathcal{S}_i} P$. Let $G^i = (V^i, E^i)$ be the subgraph induced by E^i. We consider the setting where the V^is are disjoint; we call this *node-disjoint multicommodity* ISP. As before, by Claim 3, a muticommodity ISP instance in the explicit model is feasible only if $\mathcal{S}_i$ includes all $s_i \rightsquigarrow t_i$ paths contained in E^i for all $i = 1, \ldots, k$, which can be verified efficiently. Moreover, each E^i must be acyclic, and we may assume that for every i, and every $e \in E^i$, there is some $s_i \rightsquigarrow t_i$ path contained in E^i that contains e. We prove the following results, which together resolve the complexity of node-disjoint multicommodity ISP.

▶ **Theorem 6.** *There is an additive* 1*-approximation for node-disjoint multicommodity* ISP.

▶ **Theorem 7.** *Node-disjoint multicommodity* ISP *is* NP*-hard. Moreover, it is* NP*-hard to obtain a multiplicative* $(\frac{3}{2} - \epsilon)$*-approximation for any* $\epsilon > 0$*.*

4 Inverse polyhedral optimization

Recall that in an abstract integral inverse polyhedral optimization problem, we are given a polytope $\mathcal{P} \subseteq \mathbb{R}^E_+$ with explicitly specified constraints, and a set X of extreme points of $\mathcal{P}$. We want to find a positive, integral cost vector $c \in \mathbb{R}^E$ minimizing $\|c\|_\infty$ such that: (i) in inverse polyhedral minimization (IMin-Poly), X is the set of extreme-point optimal solutions to $\min_{x\in\mathcal{P}} c^T x$; and (ii) in inverse polyhedral maximization (IMax-Poly), X is the set of extreme-point optimal solutions to $\max_{x\in\mathcal{P}} c^T x$. In the implicit version, we are given $U \subseteq E$, which defines X to be all extreme points $\hat{x}$ of $\mathcal{P}$ such that $\{e : \hat{x}_e > 0\} \subseteq U$, and $\hat{x}$ is maximal (for IMax-Poly) or minimal (for IMin-Poly) in $\mathcal{P}$.

Our approach consists of two main steps. We first find an optimal fractional cost vector $c^* \geq 1$, and then round this. While prior work also deals with obtaining such a fractional cost vector, in our case, this step is significantly more complicated due to both the existence of *multiple* solutions in X, and the requirement that these be the *unique* optimal solutions. Let $Ax \leq b$ denote the constraints of $\mathcal{P}$ (including nonnegativity). Let K be an integer such that all entries of A, b, and all extreme points of $\mathcal{P}$ are integer multiples of $\frac{1}{K}$. We can compute K with $\log K = \text{poly}(\text{input size})$. (If A is totally unimodular (TU) and b is integral, then $K = 1$.) So for any solution c to IMax-Poly or IMin-Poly, we have $|c^T\hat{x} - c^T x| \geq \frac{1}{K}$ for any $\hat{x} \in X$ and $x' \notin X$. For IMin-Poly, we solve the following LP-relaxation to find c^*. (For IMax-Poly, (6), and the arguments below, are modified appropriately.)

$$\text{(IMin-P)} \quad \min \quad \|c\|_\infty \quad \text{s.t.} \quad \begin{cases} c_e \geq 1 \quad \forall e \in E, \quad c^T\hat{x} = \lambda \quad \forall \hat{x} \in X & (5) \\ c^T\hat{x} \geq \lambda + \dfrac{1}{K} \quad \forall \hat{x} : \hat{x} \text{ is an extreme point of } \mathcal{P}, \hat{x} \notin X. & (6) \end{cases}$$

To solve (IMin-P) in the explicit model, we require a face oracle for $\mathcal{P}$. We first use this to determine if X forms a face F of $\mathcal{P}$; if not, then the inverse problem is infeasible. Otherwise, letting J be the set of constraints that are tight for all $x \in X$, the face F is given by $F = \{x \in \mathcal{P} : (Ax)_i = b_i \ \forall i \in J\}$. Further, any extreme point $x \in \mathcal{P} \setminus X$ does not lie in F, so there is some $i \in J$ such that $(Ax)_i < b_i$, and hence $(Ax)_i \leq b_i - \frac{1}{K}$. Our separation oracle for (IMin-P) is as follows. Constraints (5) can be directly checked. For (6), we consider every $i \in J$ and check that the minimum value of $c^T x$ over the set $\{x \in \mathcal{P} : (Ax)_i \leq b_i - \frac{1}{K}\}$ is at least $\lambda + \frac{1}{K}$. This can be done in polynomial time.

In the implicit setting, to separate over constraints (5), we require what we call a *minimality oracle* (or maximality oracle, for IMax-Poly), which given a set $U \subseteq E$ and a cost vector c, determines if every minimal (extreme) point of $\mathcal{P}$ whose support lies in U has the same (minimum) cost. To verify if constraints (6) hold, first note that if $x' \notin X$ is an extreme point supported on U, then there is some $\hat{x} \in X$ with $\hat{x} \leq x'$, and so $c^T x' \geq c^T\hat{x} + \frac{1}{K}$. So we only need to check if (6) holds for extreme points x' not supported on U; this can be done by the same procedure as for the explicit model, taking $J = E \setminus U$. The problem of devising a minimality/maximality oracle is itself an interesting and non-trivial problem for various combinatorial optimization problems. We show how to devise such an oracle for min-cost flow and bipartite matching (Theorems 10 and 11). Note that for a *polytope* $\mathcal{P}$ of the form $\{x : Ax = b,\ x \geq 0\}$, any two feasible points are incomparable; so the face formed by X is simply $F = \{x \in \mathcal{P} : x_e = 0 \ \forall e \notin U\}$, and we can obtain a minimality/maximality oracle by checking if the minimum and maximum values of $c^T x$ over F are equal.

The next step is to round c^* to obtain an approximately optimal integral cost vector. We show how to do this in two settings, when the constraint matrix A is TU, and when A is a sparse $\{0,1\}$-matrix. In both cases, we round an optimal solution to the dual of the problem of optimizing ${c^*}^T x$ over $\mathcal{P}$ but the details and bounds obtained differ.

▶ **Theorem 8.** *Let $\mathcal{P} = \{x \in \mathbb{R}^E : Ax \geq b,\ x \geq 0\}$, where A is TU, and b is integral. We can round an optimal fractional solution c^* to the inverse problem to obtain: (a) an additive 1-approximation for* IMin-Poly*; and (b) a multiplicative 2-approximation for* IMax-Poly*.*

▶ **Theorem 9.** *Let $A \in \{0,1\}^{m\times n}$ have row sparsity r and column sparsity k, where $n = |E|$. (Row sparsity is the maximum number of nonzero entries in a row of A; column sparsity is the maximum number of nonzero entries in a column of A.) Let A_1 and A_2 be submatrices of A with n columns, whose rows partition $[m]$. We can round an optimal fractional solution c^* to the inverse problem to obtain the following guarantees (in both implicit and explicit models).*

(a) *Additive $(k-1)$-approx. for* IMax-Poly *with $\mathcal{P} = \big\{x \in \mathbb{R}^E : A_1 x = b_1,\ A_2 x \leq b_2,\ x \geq 0\big\}$.*

(b) *Additive k-approximation for* IMin-Poly *with $\mathcal{P} = \big\{x \in \mathbb{R}^E : A_1 x = b_1,\ A_2 x \geq b_2,\ x \geq 0\big\}$.*

(c) *Multiplicative β-approximation for* IMax-Poly *and* IMin-Poly*, where we have $\beta = \min\big\{k + O(1), O(\sqrt{r\log n}), O(\sqrt{k\min(\log(kr),\log n)})\big\}$.*

4.1 Applications to inverse min-cost flow and inverse bipartite matching

Inverse min-cost flow. In the *integral inverse min-cost flow* (IMCF) problem, we are given a directed graph $D = (N, E)$, integer bounds $0 \leq \ell_e \leq u_e$ on every edge e, integer demands $\{b_v\}_{v\in N}$ (which could be arbitrary) such that $b(N) := \sum_{v\in N} b_v = 0$, and a set $E^1 \subseteq E$ of edges. A *flow* in D is a vector $x \in \mathbb{R}^E$ satisfying

$$x\big(\delta^{\text{in}}(v)\big) - x\big(\delta^{\text{out}}(v)\big) = b_v \quad \forall v \in N, \qquad \ell_e \leq x_e \leq u_e \quad \forall e \in E. \tag{7}$$

Given edge costs $\{c_e\}_{e\in E}$, the cost of a flow x is $\sum_e c_e x_e$. We seek positive, integral edge costs $\{c_e\}_{e\in E}$ minimizing $\|c\|_\infty$ so that the set of min-cost integral flows is precisely the set of acyclic integral flows supported on E^1. As with ISP, we may assume that every $e \in E^1$ is used by some feasible flow supported on E^1, and then, we may further assume that E^1 is acyclic, as otherwise the inverse problem is infeasible.

The min-cost flow problem is given by the LP: min $\sum_e c_e x_e$ subject to (7). The constraint matrix specifying (7) is TU (see, e.g., [22]), so IMCF is an instance of IMin-Poly with a TU constraint matrix. Since E^1 is acyclic, any two distinct feasible flows supported on E^1 are incomparable, so it is easy to obtain a minimality oracle and solve (IMin-P). We then obtain the following positive result in the above implicit model as a corollary of Theorem 8 (a). We discuss the explicit model in the full version. where we also show that IMCF is polytime solvable in certain cases, such as, the single-source setting with no (or equivalently, very large) capacities. As noted earlier, in the context of *spanning-tree protocols*, this implies that *in polynomial time, we can find the smallest positive integer link weights that enforce a prescribed routing tree as a shortest-path tree rooted a given node s.*

▶ **Theorem 10.** *There is an additive 1-approximation for* IMCF *in the implicit model.*

Inverse bipartite matching. In inverse bipartite matching, the input is an undirected bipartite graph $G = (V, E)$. In *integral max-cost bipartite matching* (IMax-BMat), we have a collection $M_1, \ldots, M_k$ of maximal matchings, and we seek positive, integral edge costs $\{c_e\}_{e\in E}$

minimizing $\|c\|_\infty$ so that $M_1, \ldots, M_k$ are the unique max-cost bipartite matchings in G. In the implicit model, we are given $E^1 \subseteq E$, and we require that the set of max-cost bipartite matchings be the set of maximal matchings contained in E^1. (Max-cost matchings must be maximal.) The max-cost bipartite matching LP is: max $\sum_e c_e x_e$ s.t. $x\big(\delta(v)\big) \le 1 \ \forall v \in V, \ \ x \ge 0$.

We also consider *integral min-cost bipartite matching* (IMin-BMat), where we are given *perfect matchings* $M_1, \ldots, M_k$, and we seek positive integral edge costs $\{c_e\}_{e\in E}$ minimizing $\|c\|_\infty$ such that these are the unique min-cost perfect matchings in G. In the implicit model, we are given $E^1 \subseteq E$, and the set of min-cost perfect matchings should be the set of perfect matchings contained in E^1. The min-cost perfect matching problem can be modeled by the following LP: min $\sum_e c_e x_e$ s.t. $x\big(\delta(v)\big) = 1 \ \forall v \in V, \ \ x \ge 0$. The constraint matrix in the above LPs is TU and has column sparsity 2. We devise a face oracle for IMax-BMat and IMin-BMat in the explicit setting, and a maximality oracle for IMax-BMat in the implicit setting when $E^1 = E$ by exploiting various structural properties of bipartite matchings. (A minimality oracle for IMin-BMat is easy since the corresponding polytope is defined by equations and nonnegativity constraints.) The maximality oracle for IMax-BMat determines if there exist two maximal matchings of different costs; we note that the related problem of finding a min-cost maximal matching is *NP*-hard. Theorems 8(a) and 9(a) then yield the following.

▶ **Theorem 11.** *We can obtain additive guarantees of* 1 *for* IMin-BMat*, and* IMax-BMat *in the explicit setting, and* IMax-BMat *in the implicit setting when* $E^1 = E$*.*

5 Inverse matroid-basis optimization

We consider the *integral inverse min-cost matroid basis* (IMin-Basis) and *integral inverse max-cost matroid basis* (IMax-Basis) problems. In both problems, the input is a matroid $M = (E, \mathcal{I})$ (specified by an independence oracle) and a collection $\mathcal{S}$ of bases of M. The goal is to find positive, integral costs $\{c_e\}_{e\in E}$ such that the bases in $\mathcal{S}$ are the unique optimal bases under these costs, so as to minimize $\|c\|_\infty$. More precisely, in IMin-Basis, we require that the bases in $\mathcal{S}$ be the unique *min-cost* bases under the $\{c_e\}$ costs, while in IMax-Basis, we require that the bases in $\mathcal{S}$ be the unique *max-cost* bases under the $\{c_e\}$ costs. In the implicit model, we are given $U \subseteq E$, which implicitly specifies $\mathcal{S}$ to be all bases of M contained in U.

▶ **Theorem 12.** *We can solve* IMin-Basis *and* IMax-Basis *in polynomial time.*

6 Extensions and variants

Our techniques are versatile and yield guarantees for other variants of integral inverse optimization mentioned in Section 2, including the ℓ_p-norm version (minimize $\|c\|_p$) and distance-minimization version (minimize $\|c - c^{(0)}\|_\infty$, where $c^{(0)} \in \mathbb{Z}_+^E$) problems; for these two problems our guarantees follow by simply combining our earlier results with Theorem 1.

Inverse shortest paths. We obtain multiplicative guarantees of 2 and 3 respectively for the ℓ_p-norm variant of ISP and multicommodity ISP respectively, and obtain the optimal solution and an additive guarantee of 1 for the distance-minimization variants. Bley [4] considered the ISP variant where we seek positive, integral costs so as to minimize $\max_{i=1,\ldots,k}$ (shortest-$s_i \rightsquigarrow t_i$-path distance). A related variant specifies integer upper bounds $\{D_i\}_{i=1}^k$ on the shortest-$s_i \rightsquigarrow t_i$-path distances, and seeks a positive, integral cost vector c that respects these bounds and minimizes $\|c\|_\infty$. The guarantees in Theorems 5, 6 hold for both variants.

▶ **Theorem 13.** *We obtain the following multiplicative guarantees for the ℓ_p-norm and distance-minimization versions of integral inverse polyhedral optimization.*

(a) 3*-approximation for* IMin-Poly *with* $\mathcal{P} = \{x \in \mathbb{R}^E : Ax \geq b,\ x \geq 0\}$*, where* A *is TU and* b *is integral.*

(b) $(k+1)$*-approximation for* IMax-Poly *with* $\mathcal{P} = \left\{x \in \mathbb{R}^E : A_1 x = b_1,\ A_2 x \leq b_2,\ x \geq 0\right\}$*, where* $A^T = \begin{pmatrix} A_1^T & A_2^T \end{pmatrix}$ *is a* $\{0,1\}$ *matrix, and* A *has column sparsity* k*.*

(c) $(k+2)$*-approximation for of* IMin-Poly *with* $\mathcal{P} = \left\{x \in \mathbb{R}^E : A_1 x = b_1,\ A_2 x \geq b_2,\ x \geq 0\right\}$*, where* $A^T = \begin{pmatrix} A_1^T & A_2^T \end{pmatrix}$ *is a* $\{0,1\}$ *matrix, and* A *has column sparsity* k*.*

In the explicit model, we assume that we have a face oracle for $\mathcal{P}$*. In the implicit model, we assume that we have a minimality/maximality oracle for* $\mathcal{P}$*.*

▶ **Theorem 14.** *(a) There is a multiplicative* 2*-approximation algorithm for the* ℓ_p*-norm minimization versions of* IMin-Basis *and* IMax-Basis*. (b) The distance-minimization versions of* IMin-Basis *and* IMax-Basis *can be solved exactly in polytime.*

References

1 Ravindra K Ahuja, Thomas L Magnanti, and James B Orlin. *Network flows: theory, algorithms, and applications.* Prentice hall, 1993.

2 R.K. Ahuja and J.B. Orlin. Inverse optimization. *Oper. Res.*, 49:171–783, 2001.

3 W. Ben-Ameur and E. Gourdin. Internet routing and related topology issues. *SIAM J. Discrete Math.*, 17:18–49, 2004.

4 A. Bley. Inapproximability results for the inverse shortest paths problem with integer lengths and unique shortest paths. *Networks*, 50:29–36, 2007.

5 Endre Boros, Khaled Elbassioni, Vladimir Gurvich, and Hans Raj Tiwary. The negative cycles polyhedron and hardness of checking some polyhedral properties. *Annals of Operations Research*, 188(1):63–76, 2011.

6 David Bremner, Komei Fukuda, and Ambros Marzetta. Primal-dual methods for vertex and facet enumeration. *Discrete & Computational Geometry*, 20(3):333–357, 1998.

7 D. Burton, W. Pulleyblank, and Ph. L. Toint. The inverse shortest paths problem with upper bounds on shortest paths costs. In *Network optimization*, pages 156–171. Springer, 1996.

8 D Burton and Ph L Toint. On the use of an inverse shortest paths algorithm for recovering linearly correlated costs. *Mathematical Programming*, 63(1):1–22, 1994.

9 D. Burton and Ph.L. Toint. On an instance of the inverse shortest paths problem. *Math. Program.*, 53:45–61, 1992.

10 Michael R Bussieck and Marco E Lübbecke. The vertex set of a 01-polytope is strongly p-enumerable. *Computational Geometry*, 11(2):103–109, 1998.

11 Mikael Call and Kaj Holmberg. Complexity of inverse shortest path routing. In *INOC*, pages 339–353. Springer, 2011.

12 Robert B Dial. Minimal-revenue congestion pricing part i: A fast algorithm for the single-origin case. *Transportation Research Part B: Methodological*, 33(3):189–202, 1999.

13 Robert B Dial. Minimal-revenue congestion pricing part ii: An efficient algorithm for the general case. *Transportation Research Part B: Methodological*, 34(8):645–665, 2000.

14 Martin E Dyer. The complexity of vertex enumeration methods. *Mathematics of Operations Research*, 8(3):381–402, 1983.

15 F. Grandoni, G. Nicosia, G. Oriolo, and L. Sanità. Stable routing under the spanning tree protocol. *Operations Research Letters*, 38:399–404, 2010.

16 N. Haehnle, L. Sanità, and R. Zenklusen. Stable routing and unique-max coloring on trees. *SIAM J. Discret. Math*, 27:109–125, 2013.

17 C. Heuberger. Inverse combinatorial optimization: A survey on problems, methods, and results. *J. Combin. Optim.*, 8:329–361, 2004.

18 Garud Iyengar and Wanmo Kang. Inverse conic programming with applications. *Operations Research Letters*, 33(3):319–330, 2005.

19 Gertrud Neumann-Denzau and Jörn Behrens. Inversion of seismic data using tomographical reconstruction techniques for investigations of laterally inhomogeneous media. *Geophysical Journal International*, 79(1):305–315, 1984.

20 J. Scott Provan. Efficient enumeration of the vertices of polyhedra associated with network LP's. *Math. Program.*, 63(1):47–64, 1994.

21 Robert C Read. Bounds on backtrack algorithms for listing cycles, paths and spanning trees. *Networks*, 5:237–252, 1975.

22 Alexander Schrijver. *Theory of linear and integer programming.* John Wiley & Sons, 1998.

23 Albert Tarantola. *Inverse problem theory and methods for model parameter estimation.* SIAM, 2005.

24 H.R. Varian. Revealed preference. *Samuelsonian Economics and the Twenty-first Century*, pages 99–115, 2006.

25 S. Xu and J. Zhang. An inverse problem of the weighted shortest path problem. *Japan J. Indust. Appl. Math.*, 12:47–59, 1995.

26 J. Zhang and M.C. Cai. Inverse problem of minimum cuts. *Mathematical Methods of Oper. Res.*, 47:51–58, 1998.

27 J. Zhang and Z. Liu. A general model of some inverse combinatorial optimization problems and its solution method under l-infinity norm. *J. Combin. Optim.*, 6:207–227, 2002.

28 J. Zhang and Z. Ma. Solution structure of some inverse combinatorial optimization problems. *J. Combin. Optim.*, 3:127–139, 1999.

29 Jianzhong Zhang and Zhenhong Liu. Calculating some inverse linear programming problems. *Journal of Computational and Applied Mathematics*, 72(2):261–273, 1996.

30 Jianzhong Zhang and Zhenhong Liu. A further study of inverse linear programming problems. *Journal of Computational and Applied Mathematics*, 106:345–359, 1999.