Mathematical Programming (2020) 183:525–554
https://doi.org/10.1007/s10107-019-01444-6

FULL LENGTH PAPER

Series B

Linear programming using limited-precision oracles

Ambros Gleixner[1] · **Daniel E. Steffy**[2]

Received: 15 May 2019 / Accepted: 30 October 2019 / Published online: 19 November 2019

Abstract

Since the elimination algorithm of Fourier and Motzkin, many different methods have been developed for solving linear programs. When analyzing the time complexity of LP algorithms, it is typically either assumed that calculations are performed exactly and bounds are derived on the number of elementary arithmetic operations necessary, or the cost of all arithmetic operations is considered through a bit-complexity analysis. Yet in practice, implementations typically use limited-precision arithmetic. In this paper we introduce the idea of a limited-precision LP oracle and study how such an oracle could be used within a larger framework to compute exact precision solutions to LPs. Under mild assumptions, it is shown that a polynomial number of calls to such an oracle and a polynomial number of bit operations, is sufficient to compute an exact solution to an LP. This work provides a foundation for understanding and analyzing the behavior of the methods that are currently most effective in practice for solving LPs exactly.

Keywords Linear programming · Oracle complexity · Diophantine approximation · Exact solutions · Symbolic computation · rational arithmetic · Extended-precision arithmetic · Iterative refinement

An extended abstract of this work appeared as: Gleixner A., Steffy D.E. (2019) Linear Programming Using Limited-Precision Oracles. In: Lodi A., Nagarajan V. (eds) Integer Programming and Combinatorial Optimization. IPCO 2019. Lecture Notes in Computer Science, vol 11480. Springer, Cham.
Ambros Gleixner was supported by the Research Campus MODAL *Mathematical Optimization and Data Analysis Laboratories* funded by the German Ministry of Education and Research (BMBF Grant Number 05M14ZAM).

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s10107-019-01444-6) contains supplementary material, which is available to authorized users.

✉ Daniel E. Steffy
steffy@oakland.edu

Ambros Gleixner
gleixner@zib.de

1 Konrad-Zuse-Zentrum für Informationstechnik Berlin, Takustr. 7, 14195 Berlin, Germany

2 Mathematics and Statistics, Oakland University, Rochester, Michigan 48309, USA

Mathematics Subject Classification 90C05 Linear programming · 68Q25 Analysis of algorithms and problem complexity · 11K60 Diophantine approximation · 68W30 Symbolic computation and algebraic computation · 65G30 Interval and finite arithmetic · 65F99 Numerical analysis

1 Introduction

This paper studies algorithms for solving linear programming problems (LPs) exactly over the rational numbers. The focus lies on methods that employ a limited-precision LP oracle—an oracle that is capable of providing approximate primal–dual solutions. Connections will be made to previous theoretical and practical studies. We consider linear programs of the following *standard form*:

$$\min\{c^T x \mid Ax = b,\ x \geq \ell\} \tag{P}$$

where A is an $m \times n$ matrix of full row rank with $m \leqslant n$, x is a vector of variables, c is the objective vector and ℓ is a vector of lower bounds. This form of LP is convenient for describing the algorithms, any of which can be adapted to handle alternative forms. It is assumed that all input data is rational and our goal is to compute exact solutions over the rational numbers. We note that although the assumption of rational data is common in the literature and holds for most applications, there are some applications where irrational data arises naturally (e.g. from geometric structures). In such cases where irrational data is necessary, many of the algorithms described herein may still be of use, for example to compute highly accurate approximate solutions. In the following, we survey relevant background material and previous work.

1.1 Basic terminology

We assume that the reader is familiar with fundamental results related to linear optimization, such as those presented in [17,27]. The following notation will be used throughout the paper. For a matrix A, and subsets J, K of the rows and columns, respectively, we use A_{JK} to denote the submatrix of A formed by taking entries whose row and column indices are in those sets. When J or K consists of a single element i we will use i in place of $\{i\}$, and '$\cdot$' is used to represent all rows or columns. The unit matrix of dimension n is denoted by I_n.

It is well known that if an LP has a bounded objective value then it has an optimal *basic solution* x^* of the following form. For a subset $\mathscr{B}$ of m linearly independent columns A, set $x^*_{\mathscr{B}} := A^{-1}_{\cdot\mathscr{B}}(b - \sum_{i \notin \mathscr{B}} A_{\cdot i}\ell_i)$ and $x^*_i := \ell_i$ for all $i \notin \mathscr{B}$. Generally, such a set $\mathscr{B}$ is called a *basis* and the matrix $A_{\cdot\mathscr{B}}$ is the *basis matrix* associated with $\mathscr{B}$.

We denote the *maximum norm* of a vector $x \in \mathbb{R}^n$ by $\|x\|_\infty := \max_{i=1,\ldots,n}|x_i|$. The corresponding *row sum norm* of a matrix $A \in \mathbb{R}^{m\times n}$ given by

$$\|A\|_\infty := \max_{i=1,\ldots,m} \sum_{j=1}^{n} |A_{ij}|, \tag{1}$$

is compatible with the maximum norm in the sense that $\|Ax\|_\infty \leqslant \|A\|_\infty \|x\|_\infty$ for all $A \in \mathbb{R}^{m\times n}$, $x \in \mathbb{R}^n$. Furthermore, we define the *encoding length* or *size* of an integer $n \in \mathbb{Z}$ as

$$\langle n \rangle := 1 + \lceil \log(|n| + 1) \rceil. \tag{2}$$

Unless otherwise noted, logarithms throughout the paper are base two. Encoding lengths of other objects are defined as follows. For a rational number p/q with $p \in \mathbb{Z}$ and $q \in \mathbb{Z}_{\geq 0}$, $\langle p/q \rangle := \langle p \rangle + \langle q \rangle$. For a vector $v \in \mathbb{Q}^n$, $\langle v \rangle := \sum_i \langle v_i \rangle$ and for a matrix $A \in \mathbb{Q}^{m\times n}$, $\langle A \rangle := \sum_{i,j} \langle A_{ij} \rangle$. Note that $\langle v \rangle \geqslant n$ and $\langle A \rangle \geqslant nm$. To clearly distinguish between the size and the value of numbers, we will often explicitly use the term *value* when referring to the numeric value taken by numbers.

1.2 Approximate and exact solutions

In order to compute exact rational solutions to linear programs, many algorithms rely on a methodology of first computing sufficiently accurate approximate solutions and then applying techniques to convert these solutions to exact rational solutions. This general technique is not unique to linear programming and has been applied in other areas such as exact linear algebra. In this section we will describe some results that make this possible.

First, we consider a result related to the *Diophantine approximation problem*, a problem of determining low-denominator rational approximations p/q of a given number α.

Theorem 1 ([27], Cor. 6.3b) *For $\alpha \in \mathbb{Q}$, $M > 0$ there exists at most one rational number p/q such that $|p/q - \alpha| < 1/(2Mq)$ and $1 \leqslant q \leqslant M$. There exists a polynomial-time algorithm to test whether this number exists and, if so, to compute this number.*

The proof makes use of continued fraction approximations. The resulting algorithm, which is essentially the extended Euclidean algorithm, runs in polynomial time in the encoding length of α and M and is very fast in practice. This technique is sometimes referred to as "rounding" since for $M = 1$ above it corresponds to rounding to the nearest integer, or as "numerical rational number reconstruction" (see [30]). In the remainder of the paper we will often refer to the process as simply *rational reconstruction*.

Suppose there is an unknown rational number p/q with $q \leqslant M$. If an approximation α that satisfies $|p/q - \alpha| < 1/(2M^2)$ can be computed, then Theorem 1 can be applied to determine the exact value of p/q in polynomial time. Since basic solutions of linear programs correspond to solutions of linear systems of equations, one can always derive an a priori bound on the size of the denominators of any basic solution using Cramer's rule and the Hadamard inequality as follows.

Lemma 1 *The entries of any basic primal-dual solution of a rational LP (P) have denominator bounded by*

$$H := n^{m/2} L^n \prod_{j=1}^{m} \|A_{j\cdot}\|_\infty \tag{3}$$

with L the least common multiple of the denominators of the entries in A, b, ℓ, and c.

Proof Scaling with L transforms (P) into an LP with identical primal-dual solutions $\min\{(Lc)^T x \mid (LA)x = Lb,\ x \geq L\ell\}$. Then basic solutions are uniquely determined by linear systems with integral coefficient matrix $L\tilde{B} \in \mathbb{Z}^{n\times n}$, where $\tilde{B}$ is a square matrix constructed from all rows of A plus unit vectors, see Eqs. (31) and (32) of Sect. 5.3. By Cramer's rule, the denominators are then a factor of $|\det(L\tilde{B})| = L^n |\det(\tilde{B})|$. Applying Hadamard's inequality to the rows of $\tilde{B}$ yields $|\det(\tilde{B})| \leqslant (\prod_{j=1}^{m} \|A_{j\cdot}\|_2) \cdot 1 \cdots 1$. With $\|A_{j\cdot}\|_2 \leqslant \sqrt{n}\|A_{j\cdot}\|_\infty$ we obtain the desired result $|\det(L\tilde{B})| \leqslant n^{m/2} L^n \prod_{j=1}^{m} \|A_{j\cdot}\|_\infty$. □

Therefore, upon computing an approximation of an optimal basic solution vector where each component is within $1/(2H^2)$ of the exact value, we may apply Theorem 1 componentwise to recover the exact solution vector.

A different technique for reconstructing rational vectors, one that is computationally more expensive but works under milder assumptions, is based on polynomial-time *lattice reduction* algorithms as pioneered by [22] and recently improved by [25]. It rests on the following theorem.

Theorem 2 ([22], Prop. 1.39) *There exists a polynomial-time algorithm that, given positive integers n, M and α ∈ ℚ^n, finds integers p_1, ..., p_n, q for which*

$$|p_i/q - \alpha_i| \leqslant 1/(Mq) \text{ for } i = 1, \ldots, n,$$
$$1 \leqslant q \leqslant 2^{n(n+1)/4} M^n.$$

Note that in contrast to Theorem 1, the above can be used to recover the entire solution vector at once, instead of componentwise; this process is often referred to as *simultaneous Diophantine approximation*. This has prominently been applied by [17] in the following fashion. Let the *facet complexity* of a polyhedron P be the smallest number $\varphi \geqslant n$ such that P is defined by a list of inequalities each having encoding length at most φ.

Lemma 2 ([17], Lem. 6.2.9) *Suppose we are given a polyhedron P in ℝ^n with facet complexity at most φ and a point α ∈ ℝ^n within Euclidean distance at most 2^{−6nφ} of P. If p ∈ ℤ^n and q ∈ ℤ satisfy*

$$\|p - q\alpha\|_2 \leqslant 2^{-3\varphi} \text{ and}$$
$$1 \leqslant q < 2^{4n\varphi} \tag{4}$$

then $\frac{1}{q} p$ *is in P.*

A solution $\frac{1}{q}p$ satisfying (4) can be computed in polynomial time by the lattice reduction algorithm behind Theorem 2. Current lattice reduction algorithms are, despite being polynomial-time, generally slower than the continued-fraction based techniques discussed above. When applying Theorem 1 componentwise to reconstruct a vector there exist heuristic methods to accelerate this process by taking advantage of the fact that denominators of the component vectors often share common factors [5,8]. We now provide an overview of algorithms for computing exact solutions to rational linear programs.

1.3 Methods for exact linear programming

Exact polynomial-time algorithms for solving rational LPs based on the ellipsoid method of Khachiyan [19] are described in Grötschel et al. [17]. There, a clear distinction is made between the *weak problem* of finding approximate solutions and the *strong problem* of finding exact solutions. The ellipsoid method produces smaller and smaller ellipsoids enclosing an optimal solution such that eventually simultaneous Diophantine approximation can be applied to recover an exact rational solution from the center of the ellipsoid. The same methods could equally be applied in the context of interior point algorithms in order to convert an approximate, sufficiently advanced solution along the central path to an exact rational solution. However, the original algorithm of Karmarkar [18] moves from the approximate solution to a nearby basis solution that matches as closely as possible and checks this for optimality. The variant of Renegar [26] instead identifies an optimal face of the feasible region from approximately tight inequalities and performs a projection step via solving a system of linear (normal) equations. The paper includes a detailed analysis and discussion of these ideas and also references the method developed by Edmonds [10] for solving linear systems of equations in polynomial time. (Note that if one is not careful, Gaussian elimination may require an exponential number of bit operations.) While these methods exhibit polynomial worst-case complexity, the high levels of precision required may be too high to use in any practical setting. This is illustrated by the following example.

Example 1 Consider the following small and unremarkable LP.

$$\begin{aligned} \max\ & 2x_1 + 3x_2 + 2x_3 + x_4 + 2x_5 - x_6 \\ s.t.\ & x_1 + x_2 + 2x_3 + 3x_4 + x_5 \leqslant 3 \\ & x_1 - x_2 + x_4 + 3x_5 - 2x_6 \leqslant 2 \\ & x_1 + 2x_2 + x_3 + 3x_4 + x_6 \leqslant 4 \\ & x_1, x_2, x_3, x_4, x_5, x_6 \geqslant 0 \end{aligned}$$

Theorem 6.3.2 of Grötschel et al. [17] says that an exact rational solution to an LP can be found by first calling a weak optimization oracle to find an approximate solution to the LP, and then apply simultaneous Diophantine approximation to recover the exact rational solution. The tolerance indicated for this purpose is given as $\varepsilon = \frac{2^{-18n\langle c\rangle - 24n^4\varphi}}{\|c\|_\infty}$, where φ is the facet complexity of the feasible region and c is the objective vector. For the above problem, this works out to $\varepsilon \approx 10^{-169,059}$. In comparison, double-precision

arithmetic only stores about 16 significant decimal digits. For any problem of real practical interest, ε will be even smaller. This underlines that—despite the fact that the ε is suitable for establishing polynomial running time of algorithms—it may be far beyond what is feasible for real computations in practice.

By contrast, the largest encoding length of any vertex of the above example is merely 27 and the largest denominator across all vertices is 8. Thus, a solution whose componentwise difference from a vertex was under 1/128 would be sufficient to apply Theorem 1 componentwise to recover the vertex.

Also in general, many LPs of practical interest are highly sparse and may have other special characteristics that result in their solutions having encoding length dramatically smaller than the value of derivable worst-case bounds on these values. Therefore it is of interest to work with what are often known as *output sensitive* algorithms; where the running time on a problem instance depends not only on the input length, but also on the size of the output. Many of the algorithms described in the remaining literature review, and the results derived in this paper, can be thought of in this context as they often have the chance to find an exact solution and terminate earlier than a worst-case bound might suggest.

The remainder of this subsection focuses on methods used in practice for computing exact rational solutions to linear programs. We first note that many of these practical methods are based on the simplex method. There is a trivial method of solving LPs with rational input data exactly, which is to apply a simplex algorithm and perform all computations in (exact) rational arithmetic. Espinoza [14] observed computationally that the running time of this naïve approach depends heavily on the encoding length of the rational numbers encountered during intermediate computations and is much slower than floating-point simplex implementations.

Edmonds noted that the inverse of a basis matrix can be represented as matrix of integer coefficients divided by a common denominator and that this representation can be efficiently updated when pivoting from basis to basis; this method is referred to as the Q-method and is further developed by [4,11]. Compared to computing a basis inverse with rational coefficients, the Q-method avoids the repeated GCD computations required by exact rational arithmetic. Escobedo and Moreno-Centeno [12,13] have applied similar ideas to achieve roundoff-error free matrix factorizations and updates.

The most successful methods in practice for solving LPs exactly over the rational numbers rely on combining floating-point and exact arithmetic in various ways. The key idea is to use the the inexact computation in ways that can provide useful information, but that exact computation is used for critical decisions or to validate final solutions. For example, Gärtner [15] uses floating-point arithmetic for some parts of the simplex algorithm, such as pricing, but uses exact arithmetic and ideas from Edmonds' Q-method when computing the solutions. Another way to combine inexact and exact computation relies on the observation that floating-point solvers often return optimal bases; since the basis provides a structural description of the corresponding solution, it can be recomputed using exact arithmetic and checked for optimality via LP duality [9,20,21]. This approach was systematized by Applegate et al. [3], in the QSopt_ex solver which utilizes increasing levels of arithmetic precision until an optimal basis is found, and its rational solution computed and verified. This approach of *incremental*

precision boosting is often very effective at finding exact solutions quickly, particularly when the initial double-precision LP subroutines are able to find an optimal LP basis, but becomes slower in cases where many extended-precision simplex pivots are used. In related work, Cheung and Cucker [6] have developed and analyzed exact LP algorithms that adopt variable precision strategies. A recent algorithm that has proven most effective in practice for computing high-accuracy and exact solutions to linear programs is *iterative refinement* for linear programming [16]. It will be described in detail in Sect. 2 and serves as the basis for much of the work in this paper.

1.4 Contribution and organization of the paper

This paper explores the question of how LP oracles based on limited-precision arithmetic can be used to design algorithms with polynomial running time guarantees in order to compute exact solutions for linear programs with rational input data. In contrast to classic methods from the literature that rely on ellipsoid or interior point methods executed with limited, but high levels of extended-precision arithmetic, our focus is more practical, on oracles with low levels of precision as used by standard floating-point solver implementations. Section 2 formalizes this notion of limited-precision LP oracles and revises the iterative LP refinement method from [16] in order to guarantee polynomial bounds on the encoding length of the numbers encountered. Sections 3 and 4 present two methodologically different extensions in order to construct exact solutions—basis verification and rational reconstruction. For both methods oracle-polynomial running time is established; for the latter, we bound the running time by the encoding length of the final solution, which renders it an output-sensitive algorithm. Some of the more technical proofs are collected in Sect. 5. Section 6 analyzes the properties of both methods computationally on a large test set of linear programs from public sources and compares their performance to the incremental precision boosting algorithm implemented in QSopt_ex. The code base used for the experiments is freely and publicly available for research. Concluding remarks are given in the final Sect. 7.

2 Convergence properties of iterative refinement

Our starting point is the iterative refinement method proposed in [16], which uses calls to a limited-precision LP solver in order to generate a sequence of increasingly accurate solutions. Its only assumption is that the underlying LP oracle returns solutions with absolute violations of the optimality conditions being bounded by a constant smaller than one. In the following we give a precise definition of a limited-precision LP oracle. This formal notion is necessary to evaluate the behavior of the algorithms defined in this paper, in particular to show that the number of oracle calls, the size of the numbers encountered in the intermediate calculations, and the time required for intermediate calculations are all polynomial in the encoding length of the input. It is also helpful to introduce the set $\mathbb{F}(p) := \{n/2^p \in \mathbb{Q} : n \in \mathbb{Z}, |n| \leqslant 2^{2p}\}$ for some fixed $p \in \mathbb{N}$; this can be viewed as a superset of floating-point numbers, that is easier to handle in

the subsequent proofs. Standard IEEE-754 double-precision floating-point numbers, for example, are all contained in $\mathbb{F}(1074)$.

Definition 1 We call an oracle a *limited-precision LP oracle* if there exist constants $p \in \mathbb{N}$, $0 < \eta < 1$, and $\sigma > 0$ such that for any LP

$$\min\{c^T x : Ax = b, x \geqslant \ell\} \tag{P}$$

with $A \in \mathbb{Q}^{m\times n}$, $b \in \mathbb{Q}^m$, and $c, \ell \in \mathbb{Q}^n$, the oracle either reports a "failure" or returns an approximate primal–dual solution $\bar{x} \in \mathbb{F}(p)^n$, $\bar{y} \in \mathbb{F}(p)^m$ that satisfies

$$\|A\bar{x} - b\|_\infty \leqslant \eta, \tag{5a}$$
$$\bar{x} \geqslant \ell - \eta\mathbb{1}, \tag{5b}$$
$$c - A^T\bar{y} \geqslant -\eta\mathbb{1}, \tag{5c}$$
$$|(\bar{x} - \ell)^T(c - A^T\bar{y})| \leqslant \sigma, \tag{5d}$$

when it is given the LP

$$\min\{\bar{c}^T x : \bar{A}x = \bar{b}, x \geqslant \bar{\ell}\} \tag{$\bar{P}$}$$

where $\bar{A} \in \mathbb{Q}^{m\times n}$, $\bar{c}, \bar{\ell} \in \mathbb{Q}^n$, and $\bar{b} \in \mathbb{Q}^m$ are A, c, ℓ, and b with all numbers rounded to $\mathbb{F}(p)$. We call the oracle a *limited-precision LP-basis oracle* if it additionally returns a basis $\mathscr{B} \subseteq \{1, \ldots, n\}$ satisfying

$$|\bar{x}_i - \ell_i| \leqslant \eta \quad \text{for all } i \notin \mathscr{B}, \tag{6a}$$
$$|c_i - \bar{y}^T A_{\cdot i}| \leqslant \eta \quad \text{for all } i \in \mathscr{B}, \tag{6b}$$

Relating this definition with the behavior of real-world limited-precision LP solvers, we note that real-world solvers are certainly not guaranteed to find a solution with residual errors bounded by a fixed constant.[1] However, these errors could nonetheless be computed and checked, correctly identifying the case of "failure". Algorithm 1 states the basic iterative refinement procedure introduced in [16]. For clarity of presentation, in contrast to [16], Algorithm 1 uses equal primal and dual scaling factors and tracks the maximum violation of primal feasibility, dual feasibility, and complementary slackness in the single parameter δ_k. The basic convergence result, restated here as Lemma 3, carries over from [16].

Lemma 3 *Given an LP of form (P) and a limited-precision LP oracle with constants η and σ, let x_k, y_k, Δ_k, $k = 1, 2, \ldots$, be the sequences of primal–dual solutions and*

[1] Otherwise, as noted in [16], an oracle that guarantees a solution with residual errors bounded by a fixed constant could be queried for solutions of arbitrarily high precision in a single call by first scaling the problem data by a large constant. As we will see, this type of manipulation is not used by our algorithms. Moreover, the requirement in the definition of the oracle that the input data entries are in the bounded set $\mathbb{F}(p)$ would forbid the input of arbitrarily scaled LPs.

Algorithm 1: Iterative Refinement for a Primal and Dual Feasible LP

input : rational LP data A, b, ℓ, c, termination tolerance $\tau \geqslant 0$
parameters : incremental scaling limit $\alpha \in \mathbb{N}, \alpha \geqslant 2$
output : primal–dual solution $x^* \in \mathbb{Q}^n$, $y^* \in \mathbb{Q}^m$ within tolerance τ

1 **begin**
2 $\Delta_1 \leftarrow 1$ `/* initial solve */`
3 get $(\bar{A}, \bar{b}, \bar{\ell}, \bar{c}) \approx (A, b, \ell, c)$ in working precision of the oracle
4 call oracle for $\min\{\bar{c}^T x : \bar{A}x = \bar{b}, x \geqslant \bar{\ell}\}$, abort if failure
5 $(x_1, y_1) \leftarrow$ approximate primal–dual solution returned
6 **for** $k \leftarrow 1, 2, \ldots$ **do** `/* refinement loop */`
7   $\hat{b} \leftarrow b - Ax_k$, $\hat{\ell} \leftarrow \ell - x_k$, $\hat{c} \leftarrow c - A^T y_k$ `/* compute residual error */`
8   $\delta_k \leftarrow \max\left\{\max_j |\hat{b}_j|, \max_i \hat{\ell}_i, \max_i -\hat{c}_i, |\sum_i -\hat{\ell}_i \hat{c}_i|\right\}$
9   **if** $\delta_k \leqslant \tau$ **then** return $x^* \leftarrow x_k$, $y^* \leftarrow y_k$
10   $\delta_k \leftarrow \max\{\delta_k, 1/(\alpha \Delta_k)\}$ `/* scale problem */`
11   $\Delta_{k+1} \leftarrow 2^{\lceil \log(1/\delta_k) \rceil}$ `/* round scaling factor to a power of two */`
12   get $(\bar{b}, \bar{\ell}, \bar{c}) \approx \Delta_{k+1}(\hat{b}, \hat{\ell}, \hat{c})$ in working precision of the oracle
   `/* solve for corrector solution */`
13   call oracle for $\min\{\bar{c}^T x : \bar{A}x = \bar{b}, x \geqslant \bar{\ell}\}$, abort if failure
14   $(\hat{x}, \hat{y}) \leftarrow$ approximate primal–dual solution returned
15   $(x_{k+1}, y_{k+1}) \leftarrow (x_k, y_k) + \frac{1}{\Delta_{k+1}}(\hat{x}, \hat{y})$ `/* perform correction */`

scaling factors produced by Algorithm 1 with incremental scaling limit $\alpha \geqslant 2$. *Let* $\varepsilon := \max\{\eta, 1/\alpha\}$. *Then for all iterations k,* $\Delta_{k+1} \geqslant \Delta_k/\varepsilon$, *and*

$$\|Ax_k - b\|_\infty \leqslant \varepsilon^k, \tag{7a}$$

$$x_k - \ell \geqslant -\varepsilon^k \mathbb{1}, \tag{7b}$$

$$c - A^T y_k \geqslant -\varepsilon^k \mathbb{1}, \tag{7c}$$

$$|(x_k - \ell)^T (c - A^T y_k)| \leqslant \sigma \varepsilon^{2(k-1)}. \tag{7d}$$

Hence, for any $\tau > 0$, *Algorithm 1 terminates in finite time after at most*

$$\lceil \max\{\log(\tau)/\log(\varepsilon), \log(\tau\varepsilon/\sigma)/\log(\varepsilon^2)\} \rceil \tag{8}$$

calls to the limited-precision LP oracle.

Proof Corollary 1 in [16] proves the result for a more general version of Algorithm 1 that treats primal and dual scaling factors independently, but does not include the rounding step in line 11. However, the same proof continues to hold because the upward rounding does not decrease the scaling factor Δ_{k+1}. Hence, the termination criterion $\delta_k \leqslant \tau$ in line 9 is met if $\max\{\varepsilon^k, \sigma\varepsilon^{2(k-1)}\} \leqslant \tau$, which holds when k reaches the bound in (8). Note that this bound on the number of refinements also holds in case the algorithm aborts prematurely after a failed oracle call. □

This lemma proves that the number of calls to the LP oracle before reaching a positive termination tolerance τ is linear in the encoding length of τ. However, the

correction step in line 15 could potentially cause the size of the numbers in the corrected solution to grow exponentially. The size of x_{k+1} and y_{k+1} may be as large as $2(\langle x_k\rangle + \langle\Delta_{k+1}\rangle\langle\hat{x}\rangle)$ and $2(\langle y_k\rangle + \langle\Delta_{k+1}\rangle\langle\hat{y}\rangle)$, respectively, if the numbers involved have arbitrary denominators. The following lemma shows that the rounding of the scaling factors prevents this behavior.

Lemma 4 *The size of the numbers encountered during Algorithm 1 with a limited-precision LP oracle according to Definition 1 is polynomially bounded in the size of the input A, b, ℓ, c, τ, when $\tau > 0$.*

A technical proof is provided in Sect. 5.2. The bound established on the encoding length of all numbers encountered during the course of the algorithm is $O(\langle A, b, \ell, c\rangle + (n+m)\langle\tau\rangle)$. This leads to the main result of this section.

Theorem 3 *Algorithm 1 with a limited-precision LP oracle according to Definition 1 runs in oracle-polynomial time, i.e., it requires a polynomial number of oracle calls and a polynomial number of bit operations in the size of the input A, b, ℓ, c, τ, when $\tau > 0$.*

Proof The initial setup and each iteration are $O(n+m+nnz)$ operations on numbers that, by Lemma 4, are of polynomially bounded size. Here nnz denotes the number of nonzero entries in A. By Lemma 3 we know that the number of iterations of the algorithm is polynomially bounded in the encoding length of the input. □

3 Oracle algorithms with basis verification

Iterative refinement as stated in Algorithm 1 only terminates in finite time for positive termination tolerance $\tau > 0$. The first extension, presented in this section, assumes a *limited-precision LP-basis oracle* as formalized in Definition 1 and computes exact basic solutions with zero violation in finite, oracle-polynomial running time.

3.1 Convergence of basic solutions

If the LP oracle additionally returns basic solutions for the transformed problem $\min\{\hat{c}^T x : \hat{A}x = \hat{b}, x \geqslant \hat{\ell}\}$, then Algorithm 1 produces a sequence $(x_k, y_k, \mathscr{B}_k)_{k=1,2,\ldots}$. Theorem 3.1 in [16] already states that if the corrector solution $\hat{x}, \hat{y}$ returned by the LP oracle is the exact basic solution for $\mathscr{B}_k$, then the corrected solution x_{k+1}, y_{k+1} in line 15 is guaranteed to be the unique solution to the original LP determined by $\mathscr{B}_k$. This is only of theoretical interest since the LP oracle returns only approximately basic solutions: Still, we can ask whether and under which conditions the sequence of bases is guaranteed to become optimal in a finite number of refinements. We show that properties (6a) and (6b) suffice to guarantee optimality after a number of refinements that is polynomial in the size of the input. The proof relies on the fact that there are only finitely many non-optimal basic solutions and that their infeasibilities are bounded. This is formalized by the following lemma.

Lemma 5 *Given an LP (P) with rational data $A \in \mathbb{Q}^{m\times n}$, $b \in \mathbb{Q}^m$, and $\ell, c \in \mathbb{Q}^n$, the following hold for any basic primal–dual solution x, y:*

1. *Either x is (exactly) primal feasible or its maximum primal violation has at least the value $1/2^{4\langle A,b\rangle+5\langle\ell\rangle+2n^2+4n}$.*
2. *Either y is (exactly) dual feasible or its maximum dual violation has at least the value $1/2^{4\langle A,c\rangle+2n^2+4n}$.*

A detailed proof of Lemma 5 is found in Sect. 5.3 but the basic idea is summarized as follows. Suppose x, y is a basic primal–dual solution with respect to some basis $\mathscr{B}$. By standard arguments, we show that the size of the entries in x and y is bounded by a polynomial in $\langle A, b, \ell, c\rangle$ and that all possible violations can be expressed as differences of rational numbers with bounded denominator (or zero).

The following theorem states the main convergence result.

Theorem 4 *Suppose we are given an LP (P), a fixed ε, $0 < \varepsilon < 1$, and a sequence of primal–dual solutions x_k, y_k with associated bases $\mathscr{B}_k$ such that (7a–7c) and*

$$|(x_k)_i - \ell_i| \leqslant \varepsilon^k \quad \text{for all } i \notin \mathscr{B}_k, \tag{9a}$$

$$|c_i - y_k^T A_{\cdot i}| \leqslant \varepsilon^k \quad \text{for all } i \in \mathscr{B}_k \tag{9b}$$

hold for $k = 1, 2, \ldots$. Then there exists a threshold $K = K(A, m, n, b, \ell, c, \varepsilon)$ such that the bases $\mathscr{B}_k$ are optimal for all $k \geqslant K$. The function satisfies the asymptotic bound $K(A, m, n, b, \ell, c, \varepsilon) \in O((m^2\langle A\rangle + \langle b, \ell, c\rangle + n^2)/\log(1/\varepsilon))$.

Again, the detailed proof of Theorem 4 is found in Sect. 5.3. The proof uses (9a) and (9b) to show analogs of (7a–7c) hold with right-hand side $2^{4m^2\langle A\rangle+2}\varepsilon^k$ for the solutions $\tilde{x}_k$, $\tilde{y}_k$ associated with bases $\mathscr{B}_k$. Then for $k \geqslant K$, the primal and dual violations of $\tilde{x}_k$, $\tilde{y}_k$ drop below the minimum thresholds stated in Lemma 5. From then on $\mathscr{B}_k$ must be optimal.

Conditions (7a–7c) require that the primal and dual violations of x_k, y_k converge to zero precisely as is guaranteed by Lemma 3 for the sequence of numeric solutions produced by iterative refinement. Additionally, (9a) and (9b) assume that the numeric solutions become "more and more basic" in the sense that the deviation of the non-basic variables from their bounds and the absolute value of the reduced costs of basic variables converges to zero at the same rate as the primal and dual violations. This is shown by the following lemma using properties (6a) and (6b) of Definition 1.

Lemma 6 *Suppose we are given a primal and dual feasible LP (P) and a limited-precision LP-basis oracle according to Definition 1 with constants p, η, and σ. Let x_k, y_k, $\mathscr{B}_k$, Δ_k, $k = 1, 2, \ldots$, be the sequences of primal–dual solutions, bases, and scaling factors produced by Algorithm 1 with incremental scaling limit $\alpha \geqslant 2$, and let $\varepsilon := \max\{\eta, 1/\alpha\}$. Then conditions (9a) and (9b) are satisfied for all k.*

Proof We prove both points together by induction over k. For $k = 1$, they hold directly because the defining conditions (6a) and (6b) are satisfied for the initial floating-point solution and $\eta \leqslant \varepsilon$. Suppose (9a) and (9b) hold for $k \geqslant 1$ and consider $k+1$. Let $\hat{x}, \hat{y}, \hat{\mathscr{B}}$ be the last approximate solution returned by the LP solver. Then for all $i \notin \mathscr{B}_{k+1}$

$$\begin{aligned}
|(x_{k+1})_i - \ell_i| &= \left|((x_k)_i + \frac{\hat{x}_i}{\Delta_{k+1}}) - \ell_i\right| \\
&= |\hat{x}_i + \Delta_{k+1}\underbrace{((x_k)_i - \ell_i)}_{=-\hat{\ell}_i}|/\underbrace{\Delta_{k+1}}_{\geqslant \varepsilon^{-k} \text{ by Lemma 3}} \\
&\leqslant |\hat{x}_i - \Delta_{k+1}\hat{\ell}_i|/\varepsilon^k \overset{(6a)}{\leqslant} \eta\varepsilon^k \leqslant \varepsilon^{k+1},
\end{aligned}$$

and similarly for all $i \in \mathscr{B}_{i+1}$

$$\begin{aligned}
|c_i - y_{k+1}^T A_{\cdot i}| &= \left|c_i - \left(y_k + \frac{\hat{y}}{\Delta_{k+1}}\right)^T A_{\cdot i}\right| \\
&= |\Delta_{k+1}\underbrace{(c_i - y_k^T A_{\cdot i})}_{=\hat{c}_i} - \hat{y}^T A_{\cdot i}|/\underbrace{\Delta_{k+1}}_{\geqslant \varepsilon^{-k} \text{ by Lemma 3}} \\
&\leqslant |\Delta_{k+1}\hat{c}_i - \hat{y}^T A_{\cdot i}|/\varepsilon^k \overset{(6a)}{\leqslant} \eta\varepsilon^k \leqslant \varepsilon^{k+1}.
\end{aligned}$$

This completes the induction step. □

3.2 Iterative refinement with basis verification

The bound on the number of refinements may seem surprisingly large, especially when considering that the best-known iteration complexity for interior point methods is $O(\sqrt{n+m}\langle A, b, \ell, c\rangle)$ [26] and that in each refinement round we solve an entire LP. One reason for this difference is that iterative refinement converges only linearly as proven in Lemma 3, while interior point algorithms are essentially a form of Newton's method, which allows for superlinear convergence. Additionally, the low-precision LPs solved by Algorithm 1 may be less expensive in practice than performing interior point iterations in very high-precision arithmetic. Nevertheless, the convergence results above provide an important theoretical underpinning for the following algorithm.

As already observed experimentally in [16], in practice, an optimal basis is typically reached after very few refinements, much earlier than guaranteed by the worst-case bound of Theorem 4. Hence, we do not want to rely on bounds computed a priori, but check the optimality of the basis early. A natural idea is to perform an exact rational solve as soon as the basis has not changed for a specified number of refinements. This is easily achieved by extending Algorithm 1 as follows.

Suppose the LP oracle returns a basis $\hat{\mathscr{B}}$ in line 14. First, we continue to perform the quick correction step and check the termination conditions for the corrected solution

until line 9. If they are violated, we solve the linear systems of equations associated with $\hat{\mathscr{B}}$ in order to obtain a basic solution $\tilde{x}$, $\tilde{y}$. Because it is by construction complementary slack, we only check primal and dual feasibility. If this check is successful, the algorithm terminates with $x^* = \tilde{x}$ and $y^* = \tilde{y}$ as optimal solution. Otherwise it is discarded and Algorithm 1 continues with the next refinement round.

In practice, the basis verification step can be skipped if the basis has not changed since the last LP oracle call in order to save redundant computation. Furthermore, because the linear systems solves can prove to be expensive, in practice, it can be beneficial to delay them until the LP oracle has returned the same basis information for a certain number of refinement rounds. This does not affect the following main result.

Theorem 5 *Suppose we are given a rational, primal and dual feasible LP (P) and a limited-precision LP-basis oracle according to Definition 1. Algorithm 1 interleaved with a basis verification step before Line 10 as described above terminates with an optimal solution to (P) in oracle-polynomial running time.*

Proof Lemmas 3 and 6 prove that the sequence of basic solutions x_k, y_k, $\mathscr{B}_k$ satisfies the conditions of Theorem 4, hence $\mathscr{B}_k$ is optimal after a polynomial number of refinements. According to Theorem 3, this runs in oracle-polynomial time. As proven by [10], the linear systems used to compute the basic solutions exactly over the rational numbers can be solved in polynomial time and this computation is done at most once per refinement round. □

4 Rational reconstruction algorithms

The LP algorithm developed in the previous section relies solely on the optimality of the basis information to construct an exact solution. Except for the computation of the residual vectors it does not make use of the more and more accurate numerical solutions produced. In this section, we discuss a conceptually different technique that exploits the approximate solutions as starting points in order to reconstruct from them an exact optimal solution. First we need to show that the sequence of approximate solutions actually converges.

4.1 Convergence to an optimal solution

Until now the convergence of the residual errors to zero was sufficient for our results and we did not have to address the question whether the sequence of solutions x_k, y_k itself converges to a limit point. The following example shows that this does not necessarily hold if the solutions returned by the LP oracle are not bounded.

Example 2 Consider the degenerate LP

$$\min\{x_1 - x_2 \mid x_1 - x_2 = 0, -x_1 + x_2 = 0, x_1, x_2 \geqslant 0\}.$$

One can show that Algorithm 1 may produce the sequence of primal–dual solutions

$$x_k = \left(2^k + 2^{-k-1}, 2^k - 2^{-k-1}\right)$$
$$y_k = \left(2^k + 2^{-1} + 2^{-3k-1}, 2^k - 2^{-1} - 2^{-3k-1}\right)$$

for $k = 1, 2, \ldots$. This happens if the LP oracle returns the approximate solution $\hat{x}_k = \left(2^{2k} - 1/4, 2^{2k} + 1/4\right)$ and $\hat{y}_k = \left(2^{2k} - 7 \cdot 2^{-2k-4}, 2^{2k} + 7 \cdot 2^{-2k-4}\right)$ for the k-th transformed problem

$$\min\{-1/2^{2k} x_1 + 1/2^{2k} x_2 \mid x_1 - x_2 = -1, -x_1 + x_2 = +1,$$
$$x_1 \geqslant -2^{2k-1} - 1/2, x_2 \geqslant -2^{2k-1} + 1/2\},$$

which is obtained with scaling factors $\Delta_k = 2^k$. The primal violation of x_k, y_k is 2^{-k}, the dual violation is 2^{-3k}, and the violation of complementary slackness is 2^{-4k}. Hence, all residual errors go to zero, but the iterates themselves go to infinity.

However, for corrector solutions from limited-precision LP oracles the following holds.

Lemma 7 *Given a rational, primal and dual feasible LP (P) and a limited-precision LP-basis oracle with precision p, let $(x_k, y_k, \Delta_k)_{k=1,2,\ldots}$ be the sequence of primal–dual solutions and scaling factors produced by Algorithm 1. Define $C := 2^p$. Then (x_k, y_k) converges to a rational, basic, and optimal solution $(\tilde{x}, \tilde{y})$ of (P) such that*

$$\|(\tilde{x}, \tilde{y}) - (x_k, y_k)\|_\infty \leqslant C \sum_{i=k+1}^{\infty} \Delta_i^{-1}. \tag{10}$$

Proof This result inherently relies on the boundedness of the corrector solutions returned by the oracle. Since their entries are in $\mathbb{F}(p)$, $\|(\hat{x}_k, \hat{y}_k)\|_\infty \leqslant 2^p$. Then $(x_k, y_k) = \sum_{i=1}^{k} \frac{1}{\Delta_i}(\hat{x}_i, \hat{y}_i)$ constitutes a Cauchy sequence: for any $k, k' \geqslant K$,

$$\|(x_k, y_k) - (x_{k'}, y_{k'})\|_\infty \leqslant 2^p \sum_{i=K+1}^{\infty} \varepsilon^i = 2^p \varepsilon^{K+1}/(1-\varepsilon), \tag{11}$$

where ε is the rate of convergence from Lemma 3. Thus, a unique limit point $(\tilde{x}, \tilde{y})$ exists.

The fact that $(\tilde{x}, \tilde{y})$ is basic, hence also rational, follows from Claims 1 and 3 in the proof of Theorem 4, see Sect. 5. Because of $\|(\tilde{x}, \tilde{y}) - (\tilde{x}_k, \tilde{y}_k)\|_\infty \leqslant \|(\tilde{x}, \tilde{y}) - (x_k, y_k)\|_\infty + \|(x_k, y_k) - (\tilde{x}_k, \tilde{y}_k)\|_\infty$, also the sequence of basic primal–dual solutions $(\tilde{x}_k, \tilde{y}_k)$ converges to the limit point $(\tilde{x}, \tilde{y})$. Since there are only finitely many basic solutions, this implies that $(\tilde{x}, \tilde{y})$ must be one of the $(\tilde{x}_k, \tilde{y}_k)$. □

Note that the statement holds for any upper bound C on the absolute values in the corrector solutions $\hat{x}, \hat{y}$ returned by the oracle in line 14 of Algorithm 1. In practice, this may be much smaller than the largest floating-point representable value, 2^p.

4.2 Output-sensitive reconstruction of rational limit points

Suppose we know a priori a bound M on the denominators in the limit, then we can compute $\tilde{x}, \tilde{y}$ from an approximate solution satisfying $\|(x_k, y_k) - (\tilde{x}, \tilde{y})\|_\infty < 1/(2M^2)$ by applying Theorem 1 componentwise. If the size of M is small, i.e., polynomial in the input size, then iterative refinement produces a sufficiently accurate solution after a polynomial number of refinements. This eliminates the need to use other methods such as rational matrix factorization to compute an exact solution.

However, known worst-case bounds for denominators in basic solutions are often very weak. This has been demonstrated by [1] for random square matrices and by [28] for the special case of selected basis matrices from linear programs. For the Hadamard bound H from (3) that holds for *all* basis matrices of an LP, this situation must be even more pronounced. Tighter bounds that are reasonably cheap to compute are—to the best of our knowledge—not available. Computing an approximate solution with error below $1/(2H^2)$ before applying the rounding procedure can thus be unnecessarily expensive. This motivates the following design of an output-sensitive algorithm, Algorithm 2, that attempts to reconstruct exact solution vectors during early rounds of refinement and tests the correctness of these heuristically reconstructed solutions exactly using rational arithmetic. We now give a description of it, followed by a proof of correctness and analysis of its running time.

The algorithm is an extension of the basic iterative refinement for linear programs, Algorithm 1, interleaved with attempts at rational reconstruction. For $k = 1$, the algorithm starts with the first oracle call to obtain the initial approximate solution and the corresponding residual errors. Unless the solution is exactly optimal, we enter the rounding routine. We compute a speculative bound on the denominator as $M_k := \sqrt{\Delta_{k+1}/(2\beta^k)}$. Then the value $1/(2M_k^2)$ equals $\beta^k/\Delta_{k+1} \approx \beta^k\delta_k$ and tries to estimate the error in the solution. If reconstruction attempts fail, the term β^k keeps growing exponentially such that we eventually obtain a true bound on the error. Initially, however, β^k is small in order to account for the many cases where the residual δ_k is a good proxy for the error. We first apply rational reconstruction to each entry of the primal vector x_k using denominator bound M_k, denoted by "round_to_denom($(x_k)_i$, M_k)". Then we check primal feasibility before proceeding to the dual vector. If feasibility and optimality could be verified in rational arithmetic, the rounded solution is returned as optimal. Otherwise, we compute the next refinement round after which reconstruction should be tried again. Because computing continued fraction approximations becomes increasingly expensive as the encoding length of the approximate solutions grows, rational reconstruction is executed at a geometric frequency governed by parameter f. This limits the cumulative effort that is spent on failed reconstruction attempts at the expense of possibly increasing the number of refinements by a factor of f. The following theorem shows that the algorithm computes an exactly optimal solution to a primal and dual feasible LP under the conditions guaranteed by Lemma 7.

Theorem 6 *Suppose we are given an LP (P), fixed constants $C \geqslant 1$, $0 < \varepsilon < 1$, $1 < \beta < 1/\varepsilon$, and a rational limit point $\tilde{x}, \tilde{y}$ with the denominator of each component at most $\tilde{q}$. Furthermore, suppose a sequence of primal–dual solutions x_k, y_k and*

Algorithm 2: Iterative Refinement with Rational Reconstruction

input : rational, primal and dual feasible LP data A, b, ℓ, c
parameters : incremental scaling limit $\alpha \in \mathbb{N}, \alpha \geqslant 2$, geometric reconstruction frequency $f \geqslant 1$, error correction factor $\beta > 1$
output : optimal primal–dual solution $x^* \in \mathbb{Q}^n$, $y^* \in \mathbb{Q}^m$

begin
  initialize Algorithm 1 with termination tolerance $\tau = 0$
  $k^* \leftarrow 0$
  for $k \leftarrow 1, 2, \ldots$ **do** `/* refinement loop */`
    perform next refinement step in Algorithm 1
    $x_k, y_k \leftarrow$ refined numeric solution
    $\Delta_{k+1} \leftarrow$ next scaling factor
    if $\delta_k = 0$ **then** `/* check termination */`
      return $x^* \leftarrow x_k, y^* \leftarrow y_k$
    else if $k \geqslant k^*$ **then** `/* try reconstruction */`
      $M_k \leftarrow \sqrt{\Delta_{k+1}/2\beta^k}$
      for $i \leftarrow 1, \ldots, n$ **do**
        $x_i^* \leftarrow$ round_to_denom$((x_k)_i, M_k)$
      if x^* *exactly feasible in rational arithmetic* **then**
        for $j \leftarrow 1, \ldots, m$ **do**
          $y_j^* \leftarrow$ round_to_denom$((y_k)_j, M_k)$
        if x^*, y^* *exactly optimal in rational arithmetic* **then**
          return x^*, y^*
      $k^* \leftarrow \lceil fk \rceil$ `/* counter for next reconstruction */`

scaling factors $\Delta_k \geqslant 1$ satisfies $\|(\tilde{x}, \tilde{y}) - (x_k, y_k)\|_\infty \leqslant C \sum_{i=k+1}^{\infty} \Delta_i^{-1}$ with $\Delta_1 = 1$ and $\Delta_k/\Delta_{k+1} \leqslant \varepsilon$. Let $M_k := \sqrt{\Delta_{k+1}/(2\beta^k)}$.

Then there exists $K = K(\tilde{q}, C) \in O(\max\{\langle\tilde{q}\rangle, \langle C\rangle\})$ such that

$$\|(\tilde{x}, \tilde{y}) - (x_k, y_k)\|_\infty < 1/(2M_k^2), \ 1 \leqslant \tilde{q} \leqslant M_k, \tag{12}$$

holds for all $k \geqslant K$, i.e., $\tilde{x}, \tilde{y}$ can be reconstructed from x_k, y_k componentwise in polynomial time using Theorem 1.

Proof Note that $\Delta_k/\Delta_{k+1} \leqslant \varepsilon$ for all k implies $\Delta_i \geqslant \Delta_j \varepsilon^{j-i}$ for all $j \leqslant i$. Then $M_k = \sqrt{\Delta_{k+1}/(2\beta^k)} \geqslant \tilde{q}$ holds if $\sqrt{1/(2\beta^k\varepsilon^k)} \geqslant \tilde{q}$. This holds for all

$$k \geqslant K_1 := (2\log\tilde{q} + 1)/\log(1/\beta\varepsilon) \in O(\langle\tilde{q}\rangle). \tag{13}$$

Furthermore, $C \sum_{i=k+1}^{\infty} \Delta_i^{-1} \leqslant C \sum_{i=k+1}^{\infty} \varepsilon^{i-k-1} \Delta_{k+1}^{-1} = C/((1-\varepsilon)\Delta_{k+1})$, which is less than $1/(2M_k^2) = \beta^k/\Delta_{k+1}$ for all

$$k > K_2 := (\log C - \log(1-\varepsilon))/\log\beta \in O(\langle C\rangle). \tag{14}$$

Hence (12) holds for all $k > K := \max\{K_1, K_2\}$. □

This running time result is output-sensitive as it depends on the encoding length of the solution. The value of C is a constant bound on the absolute values in the corrector solutions. Although C is independent of the input size and does not affect asymptotic running time, we include it explicitly in order to exhibit the practical dependency on the corrector solutions returned by the oracle.

We now consider the cost associated with reconstructing the solution vectors. The cost of applying the standard extended Euclidean algorithm to compute continued fraction approximations of a rational number with encoding length d is $O(d^2)$. Asymptotically faster variants of the extended Euclidean algorithm exist, and can accomplish this goal in $O(d \log^2 d \log\log d)$ time [30], but for simplicity in our discussion and analysis we use the quadratic bound given above. The following lemma shows that the total time spent on rational reconstruction within Algorithm 2 is polynomial in the number of refinement rounds and that if rational reconstruction is applied at a geometric frequency with $f > 1$ then the total time spent on this task is asymptotically dominated by the reconstruction of the final solution vector. We remark that this lemma can be used to show oracle polynomial running time of the entire algorithm even in the case that $f = 1$, but it sheds light on the fact that choosing $f > 1$ can lead to asymptotically improved performance.

Lemma 8 (Reconstruction of Solution Vectors) *The running time of applying rational reconstruction componentwise to x_k and y_k within the k-th round of Algorithm 2 is $O((n+m)k^2)$. Moreover, if $f > 1$ and Algorithm 2 terminates at round K then the cumulative time spent on rational reconstruction throughout the algorithm is $O((n+m)K^2)$.*

Proof From the proof of Lemma 4 we know that at the k-th refinement round, the encoding length of components of x_k, y_k is each bounded by $(2\alpha k + 3p + 2)$. The scaling limit α and the precision level p can be considered as constants and thus the encoding length of each component is $O(k)$. Together with the fact that the extended Euclidean algorithm can be implemented to run in quadratic time in the encoding length of its input, the first result is established.

To show the second claim we assume that $f > 1$ and consider the indices of the rounds at which reconstruction is attempted, and use K to denote the final such index. We observe that the sequence of these indices, listed in decreasing order, is term-wise bounded above by the following sequence: $S = (K, \lfloor K/f \rfloor, \lfloor K/f^2 \rfloor, \ldots, \lfloor K/f^a \rfloor)$ where $a = \log_f K$. This observation follows from the fact that line 19, of Algorithm 2 involves rounding up to the nearest integer. Since the encoding length of the components of x_k, y_k increase at each round, so does the cost of rational reconstruction, so by considering the cost to perform rational reconstruction at indices in the above sequence S we derive an upper bound on the true cost. Now, since the encoding length of each component used for reconstruction is linear in the iteration index, and the cost for reconstruction is quadratic in this value, we arrive at the following bound on the total cost, using well known properties of geometric series:

$$O\left(\sum_{i=0}^{a}(n+m)\lfloor K/f^i \rfloor^2\right) = O\left((n+m)K^2\sum_{i=0}^{a} f^{-2i}\right)$$

$$= O\left((n+m)K^2 \sum_{i=0}^{\infty} f^{-2i}\right)$$
$$= O\left((n+m)K^2 \frac{1}{1-f^{-2}}\right) = O((n+m)K^2)$$

This establishes the result. □

Now, assuming the conditions laid out in Theorem 6 hold we see that the number of refinements Algorithm 2 performs before computing an exact rational solution is polynomially bounded in the encoding length of the input. Together with this bound on the number of refinements, Lemma 8 gives a polynomial bound on the time spent on rational reconstruction. Lemma 4 and the arguments from Theorem 3 still apply and limit the growth of the numbers and cost of the other intermediate computations. Taken together, we arrive at the following.

Theorem 7 *Suppose we are given a primal and dual feasible LP (P) and a limited-precision LP-basis oracle according to Definition 1 with constants p, η, and σ. Fix a scaling limit $\alpha \geqslant 2$ and let $\varepsilon := \max\{\eta, 1/\alpha\}$. Then Algorithm 2 called with $\beta < 1/\varepsilon$ terminates with an optimal solution in oracle-polynomial running time.*

Note that the basis does not need to be known explicitly. Accordingly, Algorithm 2 may even return an optimal solution x^*, y^* that is different from the limit point $\tilde{x}$, $\tilde{y}$ if it is discovered by rational reconstruction at an early iterate x_k, y_k. In this case, x^*, y^* is not guaranteed to be a basic solution unless one explicitly discards solutions that are not basic during the optimality checks.

5 Proofs

This section collects some technical proofs for previous results and the necessary background material including well-known inequalities regarding encoding lengths, norms, and systems of equations.

5.1 Properties of encoding lengths

Lemma 9 *For $z_1, \ldots, z_n \in \mathbb{Z}$ and $r_1, \ldots, r_n \in \mathbb{Q}$,*

$$\langle z_1 + \ldots + z_n \rangle \leqslant \langle z_1 \rangle + \ldots + \langle z_n \rangle, \tag{15}$$
$$\langle r_1 \cdot \ldots \cdot r_n \rangle \leqslant \langle r_1 \rangle + \ldots + \langle r_n \rangle, \tag{16}$$
$$\langle r_1 + \ldots + r_n \rangle \leqslant 2(\langle r_1 \rangle + \ldots + \langle r_n \rangle). \tag{17}$$

For matrices $A \in \mathbb{Q}^{m \times n}$, $B \in \mathbb{Q}^{n \times p}$, $D \in \mathbb{Q}^{n \times n}$,

$$\langle AB \rangle \leqslant 2(p\langle A \rangle + m\langle B \rangle), \tag{18}$$
$$\langle \det D \rangle \leqslant 2\langle D \rangle - n^2. \tag{19}$$

Proof See [17, Lemma 1.3.3, 1.3.4, and Exercise 1.3.5]. Note that the factor 2 in (17) is best possible. □

Lemma 10 *For any matrix* $A \in \mathbb{Q}^{m\times n}$,

$$\|A\|_\infty \leqslant 2^{\langle A\rangle - mn} - mn \leqslant 2^{\langle A\rangle}. \tag{20}$$

Proof

$$\begin{aligned}\|A\|_\infty &= \max_{i=1,\ldots,m} \sum_{j=1,\ldots,n} |A_{ij}| \leqslant \sum_{i,j} |A_{ij}| = \left(\sum_{i,j} (|A_{ij}|+1)\right) - mn \\ &= 2^{\log(\sum_{i,j}(|A_{ij}|+1))} - mn \leqslant 2^{\sum_{i,j}\log(|A_{ij}|+1)} - mn \\ &= 2^{(\sum_{i,j}\langle A_{ij}\rangle)-mn} - mn = 2^{\langle A\rangle - mn} - mn.\end{aligned}$$

□

Lemma 11 *For any nonsingular matrix* $A \in \mathbb{Q}^{n\times n}$, *right-hand side* $b \in \mathbb{Q}^n$, *and solution vector* x *of* $Ax = b$,

$$\langle x_i\rangle \leqslant 4\langle A, b\rangle - 2n^2 \leqslant 4\langle A, b\rangle. \tag{21}$$

for all $i = 1, \ldots, n$. *Furthermore,*

$$\langle A^{-1}\rangle \leqslant 4n^2\langle A\rangle - 2n^4 \leqslant 4n^2\langle A\rangle. \tag{22}$$

Proof By Cramer's rule, the entries of x are quotients of determinants of submatrices of (A, b). With (19) this yields (21). For the second inequality note that the i-th column of A^{-1} is the solution of $Ax = e_i$. Using Cramer's rule again and the fact that submatrices of (A, I_n) have size at most $\langle A\rangle$ gives inequality (22). □

5.2 Results from Sect. 2

Proof of Lemma 4 Lines 10 and 11 of Algorithm 1 ensure that $\Delta_k \leqslant 2^{\lceil\log\alpha\rceil(k-1)}$ holds at iteration k. This can be shown inductively. For $k = 1$, $\Delta_k = 1 = 2^{\lceil\log\alpha\rceil(k-1)}$ holds by initialization. For $k + 1$,

$$\begin{aligned}\Delta_{k+1} = 2^{\lceil\log(1/\delta_k)\rceil} &\leqslant 2^{\lceil\log(\alpha\Delta_k)\rceil} \leqslant 2^{\lceil\log\alpha + \log\Delta_k\rceil} \leqslant 2^{\lceil\log\alpha + \lceil\log\alpha\rceil(k-1)\rceil} \\ &\leqslant 2^{\lceil\lceil\log\alpha\rceil k\rceil} = 2^{\lceil\log\alpha\rceil k}.\end{aligned}$$

As a result, $\langle\Delta_k\rangle \leqslant \lceil\log\alpha\rceil k$.

Furthermore, by induction from line 15, the entries in the refined solution vectors x_k and y_k have the form

$$\sum_{j=1}^{k} {\Delta_j}^{-1} \frac{n_j}{2^p} \tag{23}$$

with $n_j \in \mathbb{Z}$, $|n_j| \leqslant 2^{2p}$, for $j = 1, \ldots, k$. With $D_j := \log(\Delta_j)$ and $a := \lceil \log(\alpha) \rceil$ this can be rewritten as

$$\sum_{j=1}^{k} 2^{-D_j} \frac{n_j}{2^p} = \left(\sum_{j=1}^{k} n_j 2^{a(k-1)-D_j} \right) / 2^{p+a(k-1)}. \tag{24}$$

The latter is a fraction with integer numerator and denominator. The numerator is bounded as follows:

$$\begin{aligned} \Big| \sum_{j=1}^{k} n_j 2^{a(k-1)-D_j} \Big| &\leqslant \sum_{j=1}^{k} |n_j| 2^{a(k-1)-D_j} \leqslant 2^{2p} \sum_{i=0}^{a(k-1)} 2^i \\ &\leqslant 2^{2p} (2^{a(k-1)+1} - 1) \leqslant 2^{2p+ak} - 1. \end{aligned} \tag{25}$$

Hence, the size of the entries of x_k and y_k grows only linearly with the number of iterations k,

$$\begin{aligned} \langle x_k \rangle + \langle y_k \rangle &\leqslant (n+m) \left(\left\langle \sum_{j=1}^{k} n_j 2^{a(k-1)-D_j} \right\rangle + \langle 2^{p+a(k-1)} \rangle \right) \\ &\leqslant (n+m) \left(2 + \lceil \log(2^{2p+ak}) \rceil + \lceil \log(2^{p+a(k-1)} + 1) \rceil \right) \\ &\leqslant (n+m)(2ak + 3p + 2). \end{aligned} \tag{26}$$

The size of the remaining numbers set at iteration k are bounded accordingly,

$$\langle \hat{b} \rangle \leqslant 4(\langle b \rangle + \langle A \rangle + \langle x_k \rangle), \tag{27}$$

$$\langle \hat{\ell} \rangle \leqslant 2(\langle \ell \rangle + \langle x_k \rangle), \tag{28}$$

$$\langle \hat{c} \rangle \leqslant 4(\langle c \rangle + \langle A \rangle + \langle y_k \rangle), \tag{29}$$

$$\langle \delta_k \rangle \leqslant \max\{\langle \hat{b} \rangle, \langle \hat{\ell} \rangle, \langle \hat{c} \rangle, 2(\langle \hat{\ell} \rangle + \langle \hat{c} \rangle), \langle \alpha \rangle \langle \Delta_k \rangle\}. \tag{30}$$

By Lemma 3, the maximum number of iterations is $O(\log(1/\tau)) = O(\langle \tau \rangle)$. To summarize, the encoding length of any of the numbers encountered during the course of the algorithm is $O(\langle A, b, \ell, c \rangle + (n+m)\langle \tau \rangle)$. □

5.3 Results from Sect. 3

Proof of Lemma 5 Let x, y be a basic primal–dual solution with respect to some basis $\mathscr{B}$. Let $\mathscr{N} = \{1, \ldots, n\} \backslash \mathscr{B}$, let $B = A_{\cdot \mathscr{B}}$ be the corresponding (square, nonsingular) basis matrix and $N = A_{\cdot \mathscr{N}}$ the matrix formed by the nonbasic columns.

Then the primal solution is given as solution of

$$\underbrace{\begin{pmatrix} B & N \\ 0 & I_{n-m} \end{pmatrix}}_{=:\tilde{B}\in\mathbb{Q}^{n\times n}} \begin{pmatrix} x_{\mathscr{B}} \\ x_{\mathscr{N}} \end{pmatrix} = \underbrace{\begin{pmatrix} b \\ \ell_{\mathscr{N}} \end{pmatrix}}_{=:\tilde{b}\in\mathbb{Q}^{n}}. \tag{31}$$

The dual vector y is determined by

$$\tilde{B}^T \begin{pmatrix} y \\ z \end{pmatrix} = c \tag{32}$$

together with the vector $z \in \mathbb{Q}^{n-m}$ containing the dual slacks of the nonbasic variables.

First, primal infeasibilities can only stem from violations of the bound constraints on basic variables since the equality constraints and the nonbasic bounds are satisfied by construction. Hence, they are of the form $|x_i - \ell_i|$ for some $i \in \mathscr{B}$. From Lemma 11 applied to (31) we know that the entries in $x_{\mathscr{B}}$ have size at most $4\langle \tilde{B}, \tilde{b}\rangle - 2n^2$. Because

$$\begin{aligned} \langle \tilde{B}, \tilde{b}\rangle &= \langle B\rangle + \langle N\rangle + \langle 0\rangle + \langle I_{n-m}\rangle + \langle b\rangle + \langle \ell_{\mathscr{N}}\rangle \\ &= \langle A, b, \ell_{\mathscr{N}}\rangle + (n+1)(n-m) \\ &\leqslant \langle A, b, \ell\rangle + (n+1)(n-m), \end{aligned}$$

it follows that all nonzero entries of $x_{\mathscr{B}}$ are of form p/q, $p \in \mathbb{Z}$, $q \in \mathbb{Z}_{\geq 0}$, with

$$q \leqslant 2^{4\langle A,b,\ell\rangle + 4(n+1)(n-m) - 2n^2} \leqslant 2^{4\langle A,b,\ell\rangle + 2n^2 + 4n}.$$

The entries in ℓ can be written as p'/q', $p' \in \mathbb{Z}$, $q' \in \mathbb{Z}_{\geq 0}$, with $q' \leqslant 2^{\langle \ell\rangle}$. Combining this we know for all nonzero primal violations

$$\begin{aligned} |x_i - \ell_i| &= \left|\frac{p}{q} - \frac{p'}{q'}\right| = \frac{|pq' - p'q|}{qq'} \\ &\geqslant 1/(2^{4\langle A,b,\ell\rangle + 2n^2 + 4n} \cdot 2^{\langle \ell\rangle}) = 1/2^{4\langle A,b\rangle + 5\langle \ell\rangle + 2n^2 + 4n}. \end{aligned}$$

Second, note that dual infeasibilities are precisely the (absolute values of the) negative entries of z in (32). Again from Lemma 11 we have

$$\begin{aligned} \langle z_j\rangle &\leqslant 4\langle \tilde{B}, c\rangle - 2n^2 \\ &\leqslant 4\langle A, c\rangle + 4(n+1)(n-m) - 2n^2 \\ &\leqslant 4\langle A, c\rangle + 2n^2 + 4n \end{aligned}$$

for all $j \in \mathscr{N}$, and so any nonzero dual violation is at least $1/2^{4\langle A,c\rangle + 2n^2 + 4n}$. □

Proof of Theorem 4 The derivations presented in the following all refer to objects at one fixed iteration k. Hence, the iteration index is dropped for better readability, i.e.,

we write x for x_k and y for y_k, and $\mathscr{B}$ for $\mathscr{B}_k$. Furthermore, define $\tilde{x}$, $\tilde{y}$ to be the exact basic solution vectors corresponding to current basis $\mathscr{B}$, let $\mathscr{N} = \{1, \ldots, n\} \backslash \mathscr{B}$, and let $B = A_{\cdot\mathscr{B}}$ and $N = A_{\cdot\mathscr{N}}$. We will show that for a sufficiently large k, the primal and dual violations of the exact basic solution vectors drop below the minimum infeasibility thresholds from Lemma 5.

By construction, the basic solution $\tilde{x}$ satisfies the equality constraints $Ax = b$ exactly. For the violation of the lower bounds, we first show that x and $\tilde{x}$ converge towards each other.

Claim 1 At iteration $k = 1, 2, \ldots$, $\|x - \tilde{x}\|_\infty \leqslant 2^{4m^2\langle A\rangle+1}\varepsilon^k$.

For the nonbasic variables we have

$$\|x_{\mathscr{N}} - \tilde{x}_{\mathscr{N}}\|_\infty = \|x_{\mathscr{N}} - \ell_{\mathscr{N}}\|_\infty \overset{(9a)}{\leqslant} \varepsilon^k.$$

For the basic variables we have

$$\begin{aligned}
\|x_{\mathscr{B}} - \tilde{x}_{\mathscr{B}}\|_\infty &= \|B^{-1}(Bx_{\mathscr{B}} - B\tilde{x}_{\mathscr{B}})\|_\infty \\
&\leqslant \|B^{-1}\|_\infty \|Bx_{\mathscr{B}} - B\tilde{x}_{\mathscr{B}}\|_\infty \\
&= \|B^{-1}\|_\infty \|\underbrace{Bx_{\mathscr{B}}}_{=\,Ax\,-\,Nx_{\mathscr{N}}} - (b - N\ell_{\mathscr{N}})\|_\infty \\
&= \|B^{-1}\|_\infty \|Ax - b - (Nx_{\mathscr{N}} - N\ell_{\mathscr{N}})\|_\infty \\
&\leqslant \|B^{-1}\|_\infty \big(\|Ax - b\|_\infty + \|N(x_{\mathscr{N}} - \ell_{\mathscr{N}})\|_\infty\big) \\
&\leqslant \|B^{-1}\|_\infty \big(\underbrace{\|Ax - b\|_\infty}_{\leqslant \varepsilon^k \text{ by (7a)}} + \|N\|_\infty \underbrace{\|x_{\mathscr{N}} - \ell_{\mathscr{N}}\|_\infty}_{\leqslant \varepsilon^k \text{ by(9a)}}\big) \\
&\leqslant \|B^{-1}\|_\infty \big(\|N\|_\infty + 1\big)\varepsilon^k.
\end{aligned}$$

By Lemmas 10 and 11,

$$\|B^{-1}\|_\infty \leqslant 2^{\langle B^{-1}\rangle} \leqslant 2^{4m^2\langle B\rangle},$$

and

$$\|N\|_\infty + 1 \leqslant 2^{\langle N\rangle} + 1 \leqslant 2^{\langle N\rangle+1},$$

hence

$$\|B^{-1}\|_\infty \big(\|N\|_\infty + 1\big) \leqslant 2^{4m^2\langle B\rangle} \cdot 2^{\langle N\rangle+1} \leqslant 2^{4m^2\langle A\rangle+1},$$

proving the claim.

Claim 2 At iteration $k = 1, 2, \ldots$, $\tilde{x}_i - \ell_i \geqslant -2^{4m^2\langle A\rangle+2}\varepsilon^k$ for all $i \in \{1, \ldots, n\}$.

This follows from

$$\tilde{x}_i - \ell_i = \tilde{x}_i - x_i + x_i - \ell_i$$

$$\begin{aligned}
&\geqslant -\|\tilde{x}_i - x_i\|_\infty + x_i - \ell_i \\
&\geqslant -2^{4m^2\langle A\rangle+1}\varepsilon^k - \varepsilon^k \geqslant -2^{4m^2\langle A\rangle+2}\varepsilon^k.
\end{aligned}$$

For dual feasibility, we first show that the dual solutions y and $\tilde{y}$ converge towards each other.

Claim 3 At iteration $k = 1, 2, \ldots$, $\|y - \tilde{y}\|_\infty \leqslant 2^{4m^2\langle B\rangle}\varepsilon^k$.

This follows from

$$\begin{aligned}
\|y - \tilde{y}\|_\infty &= \|(B^T)^{-1}B^T(y - \tilde{y})\|_\infty \leqslant \|(B^T)^{-1}\|_\infty \|B^T(y - \tilde{y})\|_\infty \\
&= \|(B^T)^{-1}\|_\infty \underbrace{\|B^T y - c_{\mathscr{B}}\|_\infty}_{=\max\{|c_i - y^T A_{\cdot i}| : i \in \mathscr{B}\}} \overset{(9b)}{\leqslant} \|(B^T)^{-1}\|_\infty \varepsilon^k
\end{aligned}$$

and

$$\|(B^T)^{-1}\|_\infty \leqslant 2^{\langle (B^T)^{-1}\rangle} \leqslant 2^{4m^2\langle B^T\rangle} \leqslant 2^{4m^2\langle B\rangle}$$

using Lemmas 10 and 11.

Claim 4 At iteration $k = 1, 2, \ldots,$ $c_i - \tilde{y}^T A_{\cdot i} \geqslant -2^{4m^2\langle A\rangle+1}\varepsilon^k$ for all $i \in \{1, \ldots, n\}$.

For the basic variables, $c_i - \tilde{y}^T A_{\cdot i} = 0$. For $i \in \mathscr{N}$,

$$\begin{aligned}
|(y - \tilde{y})^T A_{\cdot i}| &\leqslant \|A_{\cdot i}\|_\infty \|y - \tilde{y}\|_\infty \\
&\leqslant \|N\|_\infty 2^{4m^2\langle B\rangle}\varepsilon^k \\
&\leqslant 2^{\langle N\rangle} 2^{4m^2\langle B\rangle}\varepsilon^k \\
&\leqslant 2^{4m^2\langle B\rangle + \langle N\rangle}\varepsilon^k \\
&= 2^{4m^2\langle A\rangle}\varepsilon^k.
\end{aligned}$$

This proves the claim via

$$\begin{aligned}
c_i - \tilde{y}^T A_{\cdot i} &= c_i - y^T A_{\cdot i} + (y - \tilde{y})^T A_{\cdot i} \\
&\geqslant c_i - y^T A_{\cdot i} - |(y - \tilde{y})^T A_{\cdot i}| \\
&\geqslant -\varepsilon^k - 2^{4m^2\langle A\rangle}\varepsilon^k \\
&\geqslant -2^{4m^2\langle A\rangle+1}\varepsilon^k.
\end{aligned}$$

From Claims 2 and 4, $\tilde{x}$ and $\tilde{y}$ violate primal and dual feasibility by at most $2^{4m^2\langle A\rangle+2}\varepsilon^k$ and $2^{4m^2\langle A\rangle+1}\varepsilon^k$, respectively. These values drop below the thresholds from Theorem 5 as soon as

$$2^{4m^2\langle A\rangle+2}\varepsilon^k < 1/2^{4\langle A,b\rangle+5\langle \ell\rangle+2n^2+4n}$$

$$\begin{aligned} &\Leftrightarrow \varepsilon^k < 1/2^{4m^2\langle A\rangle+2+4\langle A,b\rangle+5\langle \ell\rangle+2n^2+4n} \\ &\Leftrightarrow k > \frac{4m^2\langle A\rangle + 4\langle A, b\rangle + 5\langle \ell\rangle + 2n^2 + 4n + 2}{\log(1/\varepsilon)} =: K_P \end{aligned}$$

and

$$\begin{aligned} 2^{4m^2\langle A\rangle+1}\varepsilon^k &< 1/2^{4\langle A,c\rangle+2n^2+4n} \\ &\Leftrightarrow \varepsilon^k < 1/2^{4m^2\langle A\rangle+1+4\langle A,c\rangle+2n^2+4n} \\ &\Leftrightarrow k > \frac{4m^2\langle A\rangle + 4\langle A, c\rangle + 2n^2 + 4n + 1}{\log(1/\varepsilon)} =: K_D. \end{aligned}$$

From Theorem 5, the solution $\tilde{x}$, $\tilde{y}$ must be primal and dual feasible for $k \geqslant K := \max\{K_P, K_D\} + 1$. Since the solution is basic, $\tilde{x}$ and $\tilde{y}$ also satisfy complementary slackness, and hence they are optimal. The resulting threshold K has the order claimed. □

6 Computational experiments

Despite the theoretical analysis it remains an open question which of the proposed methods for solving linear programs exactly will perform best empirically, iterative refinement with basis verification or with rational reconstruction. We implemented both algorithms in the simplex-based LP solver SoPlex [32] in order to analyze and compare their computational performance on a large collection of LP instances from publicly available test sets. In particular, we aim to answer the following questions:

- How many instances can be solved by each algorithm and how do they compare in running time?
- How much of the solving time is consumed by rational factorization and rational reconstruction, and how often are these routines called?
- How many refinements are needed until reconstruction succeeds, and are the reconstructed solutions typically basic?

In addition, we compared both methods against the state-of-the-art solver QSopt_ex, which is based on incremental precision boosting [2,3].

6.1 Implementation and experimental setup

Both methods—basis verification and rational reconstruction—were implemented as extensions to the existing LP iterative refinement procedure of SoPlex, which is detailed in [16]. In the following, these extensions are denoted by $\text{SoPlex}_{\text{fac}}$ and $\text{SoPlex}_{\text{rec}}$, respectively.

The exact solution of the primal and dual basis systems relies on a rational LU factorization and two triangular solves for the standard, column-wise basis matrix containing slack columns for basic rows. The implementation of the rational solves is an adjusted

version of the floating-point LU code of SoPlex, removing parts specific to floating-point operations. The stalling threshold L for calling rational factorization is set to 2, i.e., factorization will only be called after two refinement steps have not updated the basis information.

The rational reconstruction routine is an adaptation of the code used in [28]. The newly introduced error correction factor β was set to 1.1. Furthermore, since nonbasic variables are held fixed at one of their bounds, the corresponding entries of the primal vector can be skipped during reconstruction. The rational reconstruction frequency f is set to 1.2, i.e., after a failed attempt at reconstructing an optimal solution, reconstruction is paused until 20% more refinement steps have been performed. We also employ the DLCM method described in [8] (see also [5]) for accelerating the reconstruction of the primal and dual solution vectors.

As test bed we use a set of 1202 primal and dual feasible LPs collected from several publicly available sources: Netlib [29], Hans Mittelmann's benchmark instances [24], Csaba Mészáros's LP collection [23], and the LP relaxations of the COR@L and MIPLIB mixed-integer linear programming libraries [7,31]. For details regarding the compilation we refer to the electronic supplement of [16].

The experiments were conducted on a cluster of 64-bit Intel Xeon X5672 CPUs at 3.2 GHz with 48 GB main memory, simultaneously running at most one job per node. SoPlex was compiled with GCC 4.8.2 and linked to the external libraries GMP 5.1.3 and EGlib 2.6.20. QSopt_ex was run in version 2.5.10. A time limit of two hours per instance was set for each SoPlex and QSopt_ex run.

6.2 Results

For the evaluation of the computational experiments we collected the following statistics: the total number of simplex iterations and running times for each solver; for QSopt_ex the maximum floating-point precision used; for the SoPlex runs the number of refinement steps and the number and execution times for basis verification and rational reconstruction, respectively. For the solutions returned by $\text{SoPlex}_{\text{fac}}$ and $\text{SoPlex}_{\text{rec}}$, we additionally computed the least common multiple of the denominators of their nonzero entries and report their order of magnitude, as an indicator for how "complicated" the representation of the exact solution is. The electronic supplement provides these statistics for each instance of the test set. Tables 1 and 2 below report an aggregated summary of these results. In the following, we discuss the main observations.

6.2.1 Overall comparison

None of the solvers dominates the others meaning that for each of QSopt_ex, $\text{SoPlex}_{\text{fac}}$, and $\text{SoPlex}_{\text{rec}}$ there exist instances that can be solved only by this one solver. Overall, however, the iterative refinement-based methods are able to solve more instances than QSopt_ex, and $\text{SoPlex}_{\text{fac}}$ exhibits significantly shorter running times than the other two methods. Of the 1202 instances, 1158 are solved by all three within the available time and memory resources. QSopt_ex solves 1163 instances, $\text{SoPlex}_{\text{rec}}$ solves 1189 instances, and $\text{SoPlex}_{\text{fac}}$ solves the largest number of instances: 1191.

Table 1 Aggregate comparison of solvers QSopt_ex and exact SoPlex with basis verification ($\text{SoPlex}_{\text{fac}}$) and rational reconstruction ($\text{SoPlex}_{\text{rec}}$) on instances that could be solved by all and where one solver took at least 2 s

QSopt_ex prec	#inst	QSopt_ex		$\text{SoPlex}_{\text{fac}}$			$\text{SoPlex}_{\text{rec}}$		
		#iter	t	#iter	t	Δt	#iter	t	Δt
Any	492	8025.7	15.6	9740.6	8.5	0.54	9740.6	24.2	1.55
64-bit	324	8368.3	16.1	11,683.7	11.3	0.70	11,683.7	14.8	0.92
128-bit	163	7217.1	13.9	6757.2	4.3	0.31	6757.2	58.5	4.21
192-bit	5	16,950.9	72.5	10,763.4	20.5	0.28	10,763.4	134.6	1.86

Columns #iter and t report shifted geometric means of simplex iterations and solving times, using a shift of 2 s and 100 simplex iterations, respectively. Column Δt reports the ratio between the mean solve times of SoPlex and QSopt_ex

Table 2 Computational results for iterative refinement with basis verification ($\text{SoPlex}_{\text{fac}}$) and rational reconstruction ($\text{SoPlex}_{\text{rec}}$)

Test set	#inst	$\text{SoPlex}_{\text{fac}}$				$\text{SoPlex}_{\text{rec}}$				
		#ref	#fac	t_{fac}	t	#ref	#rec	t_{rec}	t	Δt
All	1186	2.1	0.95	0.21	2.8	68.3	6.74	1.26	5.2	1.82
[1, 7200]	591	2.3	0.98	0.43	8.8	135.1	11.04	3.28	21.8	2.47
[10, 7200]	311	2.4	0.99	0.83	24.1	241.6	15.47	9.15	101.9	4.22
[100, 7200]	161	2.7	0.98	1.40	42.8	384.9	19.70	22.95	340.3	7.95

Columns #ref, #fac, #rec contain arithmetic means of the number of refinements, basis verifications, and reconstruction attempts, respectively. Columns t, t_{fac}, t_{rec} report shifted geometric mean times for the total solving process, the basis verifications, and rational reconstruction routines, respectively, with a shift of 2 s. Column Δt reports the ratio between the mean solve times of $\text{SoPlex}_{\text{rec}}$ and $\text{SoPlex}_{\text{fac}}$

Regarding running time, QSopt_ex is fastest 324 times, $\text{SoPlex}_{\text{rec}}$ 569 times, and $\text{SoPlex}_{\text{fac}}$ is fastest for 702 instances.[2]

For 32 of the 44 instances not solved by QSopt_ex, this is due to the time limit. On the other 12 instances, it cannot allocate enough memory on the 48 GB machine. Eight times this occurs during or after precision boosts and points to the disadvantage that keeping and solving extended-precision LPs may not only be time-consuming, but also require excessive memory. By contrast, the iterative refinement-based methods work with a more memory-efficient double-precision floating-point rounding of the LP and never reach the memory limit. However, $\text{SoPlex}_{\text{rec}}$ and $\text{SoPlex}_{\text{fac}}$ could not solve seven instances because of insufficient performance of the underlying floating-point oracle.[3]

[2] In order to account fairly for small arbitrary deviations in time measurements, an algorithm is considered "fastest" if it solves an instance in no more than 5% of the time taken by the fastest method. Hence, the numbers do not add up to the total of 1202 instances. Furthermore, note that the same picture holds also when excluding easy instances that took less than 2 s by all solvers: $\text{SoPlex}_{\text{fac}}$ wins more than twice as often as QSopt_ex.

[3] On three instances, floating-point SoPlex could not solve the first refinement LP within the time limit. For three further instances, the initial floating-point solve incorrectly claimed unboundedness and for one instance it incorrectly claimed infeasibility; in all of these cases, the incorrect claims were rejected successfully using feasibility and unboundedness tests as described in [16], but after starting to refine the original

Finally, for the 492 instances that could be solved by all three algorithms, but were sufficiently nontrivial such that one of the solvers took at least 2 s, Table 1 compares average running times and number of simplex iterations. In addition to all 492 instances, the lines starting with 64-bit, 128-bit, and 192-bit filter for the subsets of instances corresponding to the final precision level used by QSopt_ex. It can be seen that $\text{SoPlex}_{\text{fac}}$ outperforms the other two algorithms. Overall, it is a factor of 1.85 faster than QSopt_ex and even 2.85 times faster than $\text{SoPlex}_{\text{rec}}$. On the instances where QSopt_ex found the optimal solution after the double-precision solve (line 64-bit), $\text{SoPlex}_{\text{fac}}$ is 30% faster although it uses about 40% more simplex iterations than QSopt_ex. Not surprisingly, when QSopt_ex has to boost the working precision of the floating-point solver (lines 128-bit and 192-bit), the results become even more pronounced, with $\text{SoPlex}_{\text{fac}}$ being over three times faster than QSopt_ex.

The remaining analysis looks at the results of the iterative refinement-based methods in more detail. Table 2 provides a summary of the statistics in the electronic supplement for all 1186 instances that are solved by both $\text{SoPlex}_{\text{rec}}$ and $\text{SoPlex}_{\text{fac}}$. The lines starting with $[t, 7200]$ filter for the subsets of increasingly hard instances for which at least one method took $t = 1$, 10, or 100 s.

6.2.2 Rational reconstruction

The results largely confirm the predictions of Theorem 1, which expects an approximate solution with error about $10^{-2\,\text{dlcm}}$ or less. Here dlcm is the $\log_{10}$ of the least common multiple of the denominators in the solution vector as reported in the electronic supplement. Indeed, the dlcm value mostly correlates with the number of refinement rounds, though several instances exist where reconstruction succeeds with even fewer refinements than predicted. As can be seen from column "t_{rec}", the strategy of calling rational reconstruction at a geometric frequency succeeds in keeping reconstruction time low also as the number of refinements increases.

The 5 instances that could be solved by $\text{SoPlex}_{\text{fac}}$, but not by $\text{SoPlex}_{\text{rec}}$, show large dlcm value. This helps to explain the time outs and points to a potential bottleneck of $\text{SoPlex}_{\text{rec}}$. The number of refinements that could be performed within the time limit simply did not suffice to produce an approximate solution of sufficiently high accuracy. Finally, the reconstructed solution was almost always basic and showed identical dlcm value as the solution of $\text{SoPlex}_{\text{fac}}$. For 7 instances, rational reconstruction computed a non-basic solution.

6.2.3 Basis verification

Compared to $\text{SoPlex}_{\text{rec}}$, the number of refinements for $\text{SoPlex}_{\text{fac}}$ is very small, because the final, optimal basis is almost always reached by the second round. This confirms earlier results of [16]. Accordingly, for most LPs, $\text{SoPlex}_{\text{fac}}$ performs exactly one rational factorization (1123 out of 1191 solved); for 7 instances two factorizations.

Notably, there are 61 instances where no factorization is necessary because the approximate solution is exactly optimal. This is explained by the fact that the numbers

LP again, floating-point SoPlex failed to return an approximately optimal solution even when trying to run with different floating-point settings.

in the solution have small denominator. For 59 instances, the denominator is even one, i.e., the solution is integral. As a result, the average number of factorizations (column "#fac") of Table 2 is slightly below one. This situation even occurs for LPs with longer running times, since the simplicity of the solution is not necessarily correlated with short running times of the floating-point simplex.

On average, the time for rational factorization and triangular solves (column "t_{fac}") is small compared to the total solving time. Also in absolute values, t_{fac} is small for the vast majority of instances: for 901 instances it is below 0.1 s. In combination with the small number of refinements needed to reach the optimal basis, this helps to explain why $\text{SoPlex}_{\text{fac}}$ is on average between 1.82 and 7.95 times faster than $\text{SoPlex}_{\text{rec}}$. However, for 21 instances, t_{fac} exceeds 7 s and consumes more than 90% of the running time. On 3 of these instances, $\text{SoPlex}_{\text{fac}}$ times out, while they can be solved by $\text{SoPlex}_{\text{rec}}$.

7 Conclusion

This paper developed and analyzed two new algorithms for exact linear programming over the rational numbers. A notion of limited-precision LP oracles was formalized, which closely resembles modern floating-point simplex implementations. The methods extend the iterative refinement scheme of [16] in conceptually different directions: basis verification using rational linear systems solves and rational reconstruction using the extended Euclidean algorithm. Both are proven to converge to an optimal basic solution in oracle-polynomial time.

Computational experiments revealed that the rational factorization approach solved slightly more instances within a time limit and was about 46% faster on average. However, several instances were identified that were solved much faster by rational reconstruction; we also found that the reconstruction approach was slightly faster for those LPs with very short running times. This raises the question how to combine both techniques most efficiently into a hybrid algorithm. An immediate idea would be to perform a rational factorization only after rational reconstruction has failed for a fixed number of refinements. However, the critical instances on which reconstruction wins typically have solutions with large denominators and require a high number of refinements. Putting the factorization on hold in the meantime would incur a major slowdown on the majority of instances. Hence, currently the most promising hybridization seems to be a straightforward parallelization: whenever iterative refinement reaches a basis candidate that is assumed to be optimal, a rational factorization can be performed in the background while refinement and reconstruction is continued in the foreground.

Finally, we compared the iterative refinement based algorithms against the current state-of-the-art approach, the incremental precision boosting procedure implemented by the solver QSopt_ex. We found that SoPlex (using the rational factorization strategy) is 1.85 to 3 times faster on our test set and solves more instances within given time and memory restrictions. However, the advantage of incremental precision boosting is its capability to handle extremely ill-conditioned LPs by increasing the working precision of the floating-point solver when necessary. In order to harness the strengths of both approaches, incremental precision boosting can be integrated into iterative

refinement quite naturally: whenever the underlying floating-point solver encounters numerical difficulties and fails to return a satisfactory approximate solution, boost the precision of the floating-point solver to the next level. This would help to handle instances on which iterative refinement failed in our experiments because SoPlex's double-precision simplex broke down, while for the vast majority of instances no precision boosts would be necessary, retaining the significant performance benefits of iterative refinement.

Acknowledgements The authors would like to thank the anonymous reviewers for their detailed study of the manuscript and their comments, which were of exceptionally high quality.

References

1. Abbott, J., Mulders, T.: How tight is Hadamard's bound? Exp. Math. **10**(3), 331–336 (2001). http://projecteuclid.org/euclid.em/1069786341
2. Applegate, D.L., Cook, W., Dash, S., Espinoza, D.G.: QSopt_ex. http://www.dii.uchile.cl/~daespino/ESolver_doc/
3. Applegate, D.L., Cook, W., Dash, S., Espinoza, D.G.: Exact solutions to linear programming problems. Oper. Res. Lett. **35**(6), 693–699 (2007). https://doi.org/10.1016/j.orl.2006.12.010
4. Azulay, D.O., Pique, J.F.: Optimized Q-pivot for exact linear solvers. In: Maher, M., Puget, J.F. (eds.) Principles and Practice of Constraint Programming—CP98. Lecture Notes in Computer Science, vol. 1520, pp. 55–71. Springer, Berlin (1998). https://doi.org/10.1007/3-540-49481-2_6
5. Chen, Z., Storjohann, A.: A BLAS based C library for exact linear algebra on integer matrices. In: Proceedings of the 2005 International Symposium on Symbolic and Algebraic Computation, ISSAC '05, pp. 92–99 (2005). https://doi.org/10.1145/1073884.1073899
6. Cheung, D., Cucker, F.: Solving linear programs with finite precision: II. Algorithms. J. Complex. **22**(3), 305–335 (2006). https://doi.org/10.1016/j.jco.2005.10.001. http://www.sciencedirect.com/science/article/pii/S0885064X05000956
7. Computational Optimization Research at Lehigh: MIP Instances. http://coral.ie.lehigh.edu/data-sets/mixed-integer-instances/
8. Cook, W., Steffy, D.E.: Solving very sparse rational systems of equations. ACM Trans. Math. Softw. **37**(4), 1–39 (2011). https://doi.org/10.1145/1916461.1916463
9. Dhiflaoui, M., Funke, S., Kwappik, C., Mehlhorn, K., Seel, M., Schömer, E., Schulte, R., Weber, D.: Certifying and repairing solutions to large LPs: How good are LP-solvers? In: Proceedings of the 14th Annual ACM-SIAM Symposium on Discrete Algorithms, SODA '03, SIAM, pp. 255–256 (2003)
10. Edmonds, J.: Systems of distinct representatives and linear algebra. J. Res. Natl. Bur. Stand. **71B**(4), 241–245 (1967)
11. Edmonds, J., Maurras, J.F.: Note sur les Q-matrices d'Edmonds. RAIRO. Recherche Opérationnelle **31**(2), 203–209 (1997). http://www.numdam.org/item?id=RO_1997__31_2_203_0
12. Escobedo, A.R., Moreno-Centeno, E.: Roundoff-error-free algorithms for solving linear systems via Cholesky and LU factorizations. INFORMS J. Comput. **27**(4), 677–689 (2015). https://doi.org/10.1287/ijoc.2015.0653
13. Escobedo, A.R., Moreno-Centeno, E.: Roundoff-error-free basis updates of LU factorizations for the efficient validation of optimality certificates. SIAM J. Matrix Anal. Appl. **38**(3), 829–853 (2017). https://doi.org/10.1137/16M1089630
14. Espinoza, D.G.: On linear programming, integer programming and cutting planes. Ph.D. Thesis, Georgia Institute of Technology (2006). http://hdl.handle.net/1853/10482
15. Gärtner, B.: Exact arithmetic at low cost—a case study in linear programming. Comput. Geom. **13**(2), 121–139 (1999). https://doi.org/10.1016/S0925-7721(99)00012-7
16. Gleixner, A.M., Steffy, D.E., Wolter, K.: Iterative refinement for linear programming. INFORMS J. Comput. **28**(3), 449–464 (2016). https://doi.org/10.1287/ijoc.2016.0692
17. Grötschel, M., Lovász, L., Schrijver, A.: Geometric Algorithms and Combinatorial Optimization, Algorithms and Combinatorics, vol. 2. Springer, Berlin (1988)

18. Karmarkar, N.: A new polynomial-time algorithm for linear programming. Combinatorica **4**(4), 373–395 (1984). https://doi.org/10.1007/BF02579150
19. Khachiyan, L.G.: Polynomial algorithms in linear programming (in Russian). Zhurnal Vychislitel'noi Matematiki i Matematicheskoi Fiziki **20**(1), 51–68 [English translation: USSR Computational Mathematics and Mathematical Physics, 20(1):53–72, 1980] (1980). https://doi.org/10.1016/0041-5553(80)90061-0
20. Koch, T.: The final NETLIB-LP results. Oper. Res. Lett. **32**(2), 138–142 (2004). https://doi.org/10.1016/S0167-6377(03)00094-4
21. Kwappik, C.: Exact linear programming. Master's Thesis, Universität des Saarlandes (1998)
22. Lenstra, A.K., Lenstra, H.W., Lovász, L.: Factoring polynomials with rational coefficients. Math. Ann. **261**(4), 515–534 (1982). https://doi.org/10.1007/BF01457454
23. Mészáros, C.: LP test set. http://www.sztaki.hu/~meszaros/public_ftp/lptestset/
24. Mittelmann, H.: LP test set. http://plato.asu.edu/ftp/lptestset/
25. Nguyen, P.Q., Stehlé, D.: An LLL algorithm with quadratic complexity. SIAM J. Comput. **39**(3), 874–903 (2009). https://doi.org/10.1137/070705702
26. Renegar, J.: A polynomial-time algorithm based on Newton's method, for linear programming. Math. Program. **40**(1–3), 59–93 (1988). https://doi.org/10.1007/BF01580724
27. Schrijver, A.: Theory of Linear and Integer Programming. Wiley, New York (1986)
28. Steffy, D.E.: Exact solutions to linear systems of equations using output sensitive lifting. ACM Commun. Comput. Algebra **44**(3/4), 160–182 (2011). https://doi.org/10.1145/1940475.1940513
29. University of Tennessee Knoxville and Oak Ridge National Laboratory: Netlib LP Library. http://www.netlib.org/lp/
30. Wang, X., Pan, V.Y.: Acceleration of euclidean algorithm and rational number reconstruction. SIAM J. Comput. **2**(32), 548–556 (2003)
31. Zuse Institute Berlin: MIPLIB—Mixed Integer Problem Library. http://miplib.zib.de/
32. Zuse Institute Berlin: SoPlex. Sequential object-oriented simPlex. http://soplex.zib.de/