Math. Program., Ser. A (2017) 166:327–367
DOI 10.1007/s10107-017-1123-x

FULL LENGTH PAPER

Integral simplex using decomposition with primal cutting planes

Samuel Rosat[1] · **Issmail Elhallaoui**[1] · **François Soumis**[1] · **Andrea Lodi**[2]

Received: 4 May 2015 / Accepted: 9 February 2017 / Published online: 13 March 2017

Abstract This paper concentrates on the addition of cutting planes to the integral simplex using decomposition (ISUD) of Zaghrouti et al. (Oper Res 62(2):435–449, 2014). This method solves the set partitioning problem by iteratively improving an existing feasible solution. We present the algorithm in a primal language and relate it to existing augmenting methods. The resulting theoretical properties, stronger than the ones already known, simplify termination proofs and deepen the geometrical insights on ISUD in particular. We show that primal cuts, that is, cutting planes that are tight at the current feasible integer solution, can be used to improve the performance of the algorithm, and further that such cutting planes are enough to solve each augmentation problem. We propose efficient separation procedures for well-known polyhedral inequalities, namely primal clique and odd-cycle cuts. Numerical results demonstrate the effectiveness of primal cutting planes; tests are performed on small and large-scale set partitioning problems from aircrew and bus-driver scheduling instances up to 1600 constraints and 570,000 variables.

✉ Andrea Lodi
andrea.lodi@polymtl.ca

Samuel Rosat
samuel.rosat@polymtl.ca

Issmail Elhallaoui
issmail.elhallaoui@gerad.ca

François Soumis
francois.soumis@gerad.ca

[1] GERAD and École Polytechnique de Montréal, C.P. 6079, Succ. Centre-ville, Montreal, QC H3C 3A7, Canada

[2] CERC and École Polytechnique de Montréal, C.P. 6079, Succ. Centre-ville, Montreal, QC H3C 3A7, Canada

Keywords {0, 1}-Programming · Integral simplex · Primal algorithms · Set partitioning · Primal cutting-planes · Scheduling

1 Introduction

Introduced in 1969 by Garfinkel et al. [8], the Set Partitioning Problem (SPP) is a well known model of integer linear programming. Its popularity mainly comes from its simple expression, and the wide range of its applications. It can be expressed as[1]

SPP	Given a set $\mathcal{X}$ and a set of its subsets, $\mathcal{X}_1, \ldots, \mathcal{X}_n \subseteq \mathcal{X}$ of respective cost $c_1, \ldots, c_n$, determine a partition of $\mathcal{X}$, using only some of the $\mathcal{X}_i$, $1 \leq i \leq n$, and of minimum (or maximum) cost.

Applications range from aircrew scheduling [4] to vehicle routing [2] and electricity production planning [25], among others.

Primal methods for integer linear programming were introduced in the 1960s, at the same time as most classical frameworks (e.g., branch and bound). These algorithms, sometimes described as *augmenting methods* or *all-integer* algorithms, are based on the following pattern: given a starting integer solution, improve it iteratively to obtain a sequence of integer solutions with a better objective function value until optimality is reached.[2] Two of their main features are (i) to avoid the combinatorial exploration of a branching tree and (ii) to take advantage of existing starting solutions. In turn, a key feature of SPP is that it possesses the *quasi-integral* (or *Trubin*) property [34], i.e., every edge of the convex hull of the feasible set is also an edge of the polytope of the linear relaxation. Therefore, it is suitable for the implementation of all-integer algorithms in which all improvements are obtained by performing simplex pivots;[3] such algorithms are hence named *integral simplex* algorithms. One of the first attempt in this direction is the seminal work of Balas and Padberg [1], who first proposed an integral simplex algorithm specifically designed for the set partitioning problem.

One of the main drawbacks of algorithms based on simplex pivots is their inability to perform well on degenerate problems. In mathematical programming, degeneracy occurs when some basic variables are at one of their bounds, which is common in the particular case of SPP. In this case, it is very much likely that the value of variable entering the simplex basis cannot be modified without making the current solution infeasible. The resulting degenerate pivot leads to no change in the solution, and no improvement in the objective value. Recently, Zaghrouti et al. [37] proposed a new algorithm for SPP, the *Integral Simplex Using Decomposition* (ISUD) which is an offspring of recent works conducted around the *Improved Primal Simplex* in [6,17,18,33]. It is therefore designed to take advantage of degeneracy, rather than suffer from it. Combined with (i) the canonically degenerate nature of the SPP, particularly in

[1] A formulation of SPP as an integer linear programming problem is given in Sect. 2.2.

[2] The above definition includes *local search* algorithms, whose analysis is, however, outside of the scope of the present paper that is concerned with simplex-type algorithms that are able to prove optimality of the solution.

[3] More details on this feature and its implications are given in Sect. 2.2.

the industrial applications, and (ii) the observation that primal algorithms experience troubles with degeneracy, particularly when primal cutting planes are used (e.g., [16]), the previous observations on the advantageous way that ISUD copes with degeneracy show its potential to embody the next generation of primal algorithms. Furthermore, it seems natural to apply primal cutting planes techniques within an "anti-degeneracy" framework since degeneracy is reported as the main trouble experienced when adding cuts in primal methods.

From the applicative point of view, and as shown in the last part of this paper, ISUD proves highly efficient for *reoptimization*. This key feature makes it a very promising method in two particularly important cases which are reoptimization of existing solutions, and all-integer column generation. First, the need to quickly reoptimize an existing solution (a schedule) in case of unforeseen events is fundamental in many practical cases. Take for instance the example of airline crew scheduling (see [31]): The manpower planned schedule must often be modified to react to day-to-day operational constraints such as schedule disruptions, aircraft substitutions, crew absences, strikes or even volcanic eruptions! Second, consider the all-integer column generation process. Column generation is an optimization technique used to solve very large problems. When solving problems with integer variables, it is usually embedded in a branch-and-price framework. As for standard integer programming, the exploration of the branching tree can turn to be a very long process. In the same way that primal algorithms are an alternative to branch and bound, all-integer column generation is an alternative to branch and price. In all-integer column generation, a subset of the columns of the constraint matrix is generated in the beginning, and the optimal solution to this program, called *restricted master problem*, is found. Then, subproblems generate new columns to be appended to the matrix, and the new optimal solution for the extended matrix must be determined. One then iterates the process until no column can be generated, that improves the solution. Obviously, solving the extended problem from scratch (*branch and bound*) results in a loss of information and a probable lack of efficiency, while using the previous solution as a warm-start could give the algorithm a substantial advantage (*primal algorithms*). Hence, any efficient all-integer column generation code requires a good reoptimization method, since the global process is based on successive updates of an integral solution. Thus, ISUD is an excellent candidate for the reoptimization of the optimal solution of the restricted master problem every time it is extended.

This paper is organized as follows. A literature review on primal algorithms, primal cutting planes and integral simplex methods, as well as some notation, problem definitions and contribution statement are given in Sect. 2. The ISUD algorithm is presented in an innovative way, and new theoretical results are discussed in Sect. 3. Primal cutting planes are discussed in Sect. 4, new separation procedures are described, and we show that the search space of the separation can be restricted without preventing from finding these cutting planes. Numerical results are displayed in Sect. 5, which show the potential of adding primal cuts to ISUD, and conclusions are drawn in Sect. 6. An extended abstract of this work was published in [24].

2 Literature review and contribution statement

In this paper, lower-case bold symbols are used for column vectors and upper-case bold symbols denote matrices. For subsets $\mathcal{X} \subseteq \{1, \ldots, m\}$ of row indices and $\mathcal{Y} \subseteq \{1, \ldots, n\}$ of column indices, the submatrix of $\boldsymbol{A}$ with rows indexed by $\mathcal{X}$ and columns indexed by $\mathcal{Y}$ is denoted as $\boldsymbol{A}_{\mathcal{X}\mathcal{Y}}$. Similarly, $\boldsymbol{A}_{\mathcal{X}\cdot}$ is the set of rows of $\boldsymbol{A}$ indexed by $\mathcal{X}$, while $\boldsymbol{A}_{\cdot\mathcal{Y}}$ is the set of columns of $\boldsymbol{A}$ indexed by $\mathcal{Y}$, and for any vector $\boldsymbol{v} \in \mathbb{R}^n$, $\boldsymbol{v}_{\mathcal{Y}}$ is the subvector of all v_y, $y \in \mathcal{Y}$. The vector of all zeros (resp. ones) with dimension dictated by the context is denoted by $\mathbf{0}$ (resp. $\boldsymbol{e}$), and $\boldsymbol{A}^T$ is the transpose of $\boldsymbol{A}$. Finally, the linear span of all columns of any matrix $\boldsymbol{M}$, also called image of $\boldsymbol{M}$, is denoted as *Span* $(\boldsymbol{M})$.

2.1 Primal algorithms for integer linear programs

Before addressing the SPP, a general introduction on primal algorithms for ILP is given here. We consider a generic integer linear program

$$z^\star_{\text{ILP}} = \min_{\boldsymbol{x}\in\mathbb{R}^n} \left\{ \boldsymbol{c}^T\boldsymbol{x} \mid \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}\,,\ \boldsymbol{x} \geq 0 \text{ and } \boldsymbol{x} \text{ is integer} \right\}, \tag{ILP}$$

where $\boldsymbol{A} \in \mathbb{N}^{m\times n}$, $\boldsymbol{b} \in \mathbb{N}^m$ and $\boldsymbol{c} \in \mathbb{N}^n$, with $\mathbb{N}$ the set of natural integers. $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$ are called the linear constraints and $\boldsymbol{x} \geq 0$ the nonnegativity constraints. The set of all feasible solutions of ILP is denoted by $\mathcal{F}_{\text{ILP}}$. $z^\star_{\text{ILP}}$ is called the optimal value of ILP and any feasible solution $\boldsymbol{x}^\star \in \mathcal{F}_{\text{ILP}}$ such that $\boldsymbol{c}^T\boldsymbol{x}^\star = z^\star_{\text{ILP}}$ is called an optimal solution of ILP. The linear relaxation of ILP, denoted as ILP^{LR} is the linear program obtained by relaxing the integrality constraints of ILP. Its feasible domain and optimal value are resp. denoted as $\mathcal{F}_{\text{ILP}^{\text{LR}}}$ and $z^\star_{\text{ILP}^{\text{LR}}}$.

As noted by Letchford and Lodi [14], algorithms for integer linear programming can be divided into three classes: *dual fractional*, *dual integral*, and *primal* methods. *Dual fractional* algorithms maintain optimality and linear-constraint feasibility at every iteration, and they stop when integrality is reached. They are typically standard cutting plane procedures such as Gomory's algorithm [10]. The classical branch-and-bound scheme is also based on a dual-fractional approach, in particular for the determination of lower bounds. *Dual integral* methods maintain integrality and optimality, and they terminate once the (primal) linear constraints are satisfied. Letchford and Lodi give the sole example of another algorithm of Gomory [11]. Finally, *primal algorithms* maintain feasibility (including integrality) throughout the process and stop when optimality is reached. These are in fact descent algorithms for which the improving sequence $(\boldsymbol{x}_k)_{k=1\ldots K}$ satisfies the conditions

C1 $\boldsymbol{x}^k \in \mathcal{F}_{\text{ILP}}$;
C2 $\boldsymbol{x}^K$ is optimal;
C3 $\boldsymbol{c}^T\boldsymbol{x}^{k+1} < \boldsymbol{c}^T\boldsymbol{x}^k$.

Primal methods—sometimes classified as *augmenting algorithms* – were first introduced simultaneously by Ben-Israel and Charnes [3] and Young [35] and improved by Young [36] and Glover [9]. In Young's method [35,36], at iteration k, a simplex

pivot is considered: if it leads to an integer solution, it is performed; otherwise, cuts are generated and added to the problem, thereby changing the underlying structure of the constraints. Young also developed the concept of a *augmenting* (or *improving*) vector at $\boldsymbol{x}^k$, i.e., a vector $z \in \mathbb{R}^n$ such that $\boldsymbol{x}^k + z$ is integer, feasible, and of lower cost than $\boldsymbol{x}^k$. From this notion comes the *integral augmentation problem* (**iAUG**) that involves finding such a direction if it exists or asserting that $\boldsymbol{x}^k$ is optimal.

iAUG	Find an improving vector $z \in \mathbb{Z}^n$ such that $(\boldsymbol{x}^k + z) \in \mathcal{F}_{\text{ILP}}$ and $\boldsymbol{c}^T z < 0$ or assert that $\boldsymbol{x}^k$ is optimal for ILP.

Remark 1 Traditionally, papers on constraint aggregation and integral simplex algorithms deal with minimization problems, whereas authors usually present generic primal algorithms for maximization problems. We therefore draw the reader's attention to the following: to retain the usual classification, we call the improving direction problem **iAUG**, although it supplies a decreasing direction. In the same way, an improvement (general term) is, in our case, a decrease. For the same reasons, we still call it an augmentation.

For the sake of readability, in this paper, we will differentiate **iAUG** (above) from the fractional augmentation **fAUG**. The latter is the relaxation of the former in which z may be fractional and $\boldsymbol{x}^k + z$ needs only be a solution of the linear relaxation ILP^{LR}.

In the end of the 1990s, there has been a renewed interest in primal integer algorithms, inspired by Robert Weismantel as mentioned in [15]. Many recent works specifically concern {0,1}-LP (integer programming problems for which the variables can only take values 0 or 1). However, only a few papers have addressed the practical solution of **iAUG**, most of them considering it as an oracle. As a matter of fact, most of the rare computational work since 2000 on primal algorithms concerns the *primal separation problem*, defined as

P-SEP	Given a feasible solution $\boldsymbol{x}^k \in \mathcal{F}_{\text{ILP}}$ and an infeasible point $\boldsymbol{x}^\star$, find a hyperplane that separates $\boldsymbol{x}^\star$ from $\mathcal{F}_{\text{ILP}}$ and that is tight at $\boldsymbol{x}^k$ or assert that none exists.

In the general case, there is no guarantee that a primal separation hyperplane exists, and no particular conclusion can be drawn if none is found. In our case, however, $\boldsymbol{x}^\star$ is typically an adjacent vertex of the feasible domain of the linear relaxation $\mathcal{F}_{\text{ILP}^{\text{LR}}}$ obtained by performing one or several simplex pivots from $\boldsymbol{x}^k$, which guarantees that such an hyperplane exists (see Sects. 3 and 4). An example of primal separation hyperplanes is given in Fig. 1.

In 2003, Eisenbrand et al. [5] proved that the primal separation problem is, from the theoretical point of view, as difficult as the integral optimization problem for {0,1}-LP.

It is therefore expected to be a "complicated" problem because {0,1}-LP is $\mathcal{NP}$-hard. Letchford and Lodi [14,16] and Eisenbrand et al. [5] adapt well-known algorithms for the standard separation problem to primal separation. To the best of our knowledge, only few papers present computational experiments using primal methods. Best examples are those of Salkin and Koncal [26], Letchford and Lodi [14], Haus et al. [12], and Stallmann and Brglez [30]. All these papers present results on small to

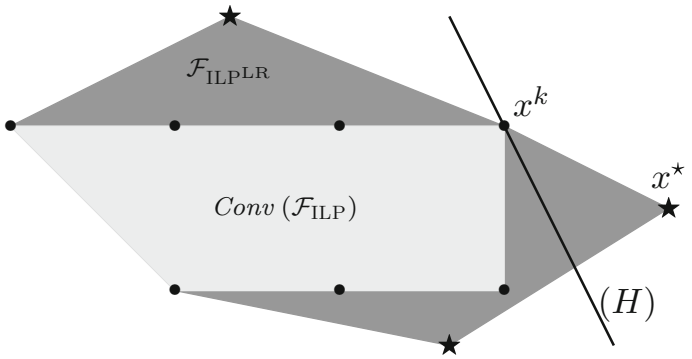


Fig. 1 Example of primal separation. (H) is a primal cut separating $\boldsymbol{x}^\star$ from *Conv* $(\mathcal{F}_{\mathrm{ILP}})$, and tight at $\boldsymbol{x}^k$. The *dark area* represents the feasible domain of the linear relaxation $\mathcal{F}_{\mathrm{ILP}^{\mathrm{LR}}}$. The feasible domain of ILP is a finite set represented by *bullets* (•) and its convex hull *Conv* $(\mathcal{F}_{\mathrm{ILP}})$ is the *light grey* polyhedron

mid-size instances. Haus et al. [12] describe a solid framework and their implementation is certainly the most complete one. Letchford and Lodi [14] present results for an algorithm using primal cutting planes and, interestingly, they stated that degeneracy prevented them from solving larger instances. As was already mentioned above, the ISUD algorithm was originally designed to cope with degeneracy and therefore seems a promising answer to these computational limits of primal algorithms.

For further information on primal algorithms, the reader is referred to the more extensive review of Spille et al. [29].

2.2 The set partitioning problem

The Set Partitioning Problem (SPP) is a particular case of {0,1}-LP, presented in the introduction. In a mathematical programming form, it reads as

$$\min_{\boldsymbol{x}\in\mathbb{R}^n} \left\{ \boldsymbol{c}^T\boldsymbol{x} \mid \boldsymbol{A}\boldsymbol{x} = \boldsymbol{e}\,,\, \boldsymbol{0} \le \boldsymbol{x} \le \boldsymbol{e} \text{ and } \boldsymbol{x} \text{ is integer} \right\}, \tag{SPP}$$

where $\boldsymbol{A} \in \{0,1\}^{m\times n}$ is a {0,1}-matrix, and $\boldsymbol{c} \in \mathbb{N}^n$ is any integer cost vector. Obviously, given the bounds ($\boldsymbol{0}$ and $\boldsymbol{e}$) and the integrality constraints, $\mathcal{F}_{\mathrm{SPP}}$ only contains {0,1}-vectors. The set of indices of the rows is denoted as $\mathcal{R} = \{1,\ldots,m\}$. Moreover, given a current solution $\boldsymbol{x}^0 \in \mathcal{F}_{\mathrm{SPP}}$, the indices $\{1,\ldots,n\}$ of its components can be partitioned into sets $\mathcal{P} = \{j | x^0_j = 1\}$, which is the set of positive-valued variables, and $\mathcal{Z} = \{j | x^0_j = 0\}$, which is the set of null variables. Submatrix $\boldsymbol{A}_{.\mathcal{P}}$ is referred to as the *working basis*, and so is $\mathcal{P}$ by extension. Its indices and the variables of $\boldsymbol{x}_{\mathcal{P}}$ are respectively referred to as *basic indices* and *basic variables*, and $p = |\mathcal{P}|$ denotes the cardinality of this working basis. The working basis is different from a standard simplex basis because it only contains linearly independent positive-valued variables (no degenerate variables). In the rest of this paper, the terms basis, basic, etc. refer to the working basis $\mathcal{P}$.

As proved in 1969 by Trubin [34], the SPP is *quasi-integral*, i.e., every edge of the convex hull of the feasible set *Conv* ($\mathcal{F}_{\text{SPP}}$) is also an edge of the polytope of the linear relaxation $\mathcal{F}_{\text{SPP}^{\text{LR}}}$. A consequence of this property is the existence of a decreasing sequence of integer solutions of SPP leading to an optimal solution, such that two consecutive solutions are adjacent vertices of $\mathcal{F}_{\text{SPP}^{\text{LR}}}$ [easily proven by applying the simplex algorithm over *Conv* ($\mathcal{F}_{\text{SPP}}$)]. When working on the SPP, the following condition `C4` can therefore be added to `C1`–`C3` to transform a primal algorithm into an *integral simplex* without preventing the procedure to reach optimality.

`C4` x^{k+1} is a neighbor of x^k in $\mathcal{F}_{\text{SPP}^{\text{LR}}}$.

These methods, first introduced in 1975 by Balas and Padberg [1], yield a sequence of improving all-integer solutions, obtained by only performing simplex pivots. Since this seminal paper, other integral simplex methods have been proposed (see Thompson [32] and Saxena [27]). Amongst contemporary work conducted parallel to ours, the most interesting one to mention is that conducted by Rönnberg and Larsson (see [21,22]). Finally, note that the SPP is by nature highly degenerate. Hence it seems relevant to apply that kind of "anti-degeneracy" techniques in this case.

2.3 Contribution statement

With the concepts introduced in Sect. 2, we can describe the contributions of Sects. 3 and 4 more clearly. In Sect. 3, we present the ISUD algorithm in a primal way, and relate it to the augmentation problems **iAUG** and **fAUG**. We formulate **fAUG** as a linear program, and **iAUG** as a nonlinear one of which **fAUG** is the linear relaxation. Each of these problems is decomposed into two subproblems. We reduce the number of constraints of both decomposed subproblems, and give a geometrical interpretation of these row-reduced problems. We also provide a simple characterization of integer directions. Section 4 addresses the improvement of the linear relaxation of **iAUG** with cutting planes. We demonstrate that every valid inequality for **iAUG** can be obtained as the linear transformation of a primal cut of SPP. We prove that such a primal cut always exists and that the characterization of integer directions given in Sect. 3 remains correct after the addition of cutting planes. Finally, we introduce two new **P-SEP** procedures for primal clique and odd-cycle cuts, and show that the search space for the cuts can be reduced to a small number of variables without changing the outcome of **P-SEP**.

3 The integral simplex using decomposition (ISUD)

This section aims to present the Integral Simplex Using Decomposition (ISUD) of Zaghrouti et al. [37] from a purely primal point of view, so as to make the link with primal algorithms straightforward, and simplify some proofs as well as the geometrical interpretation of the process. Section 3.1 concentrates on **fAUG**, while Sect. 3.2 also considers integrality and tackles **iAUG**.

Hereinafter, suppose a decreasing sequence of solutions of SPP ending at $\boldsymbol{x}^k$ is known, and **iAUG** must now be solved. For the sake of readability, we always denote the current (binary) solution as $\boldsymbol{x}^0$. We want to determine a direction $\boldsymbol{d} \in \mathbb{R}^n$ and a

step $r > 0$ such that $\boldsymbol{x}^1 = \boldsymbol{x}^0 + r\boldsymbol{d} \in \mathcal{F}_{\text{SPP}}$ and of lower cost than $\boldsymbol{x}^0$ or to assert that $\boldsymbol{x}^0$ is optimal. The set of positive-valued variables is denoted as $\mathcal{P}^0 = \{j | x_j^0 = 1\}$, and that of null variables as $\mathcal{Z}^0 = \{j | x_j^0 = 0\}$. For the sake of readability, $\mathcal{P}$, $\mathcal{Z}$, $\boldsymbol{d}$, and r (and later other objects) will not be indexed on the index of the current solution (k or 0) although they depend on $\boldsymbol{x}^0$ (or $\boldsymbol{x}^k$).

3.1 Fractional augmentation fAUG and phase decomposition

In this section, the sole problem of a fractional augmentation, **fAUG**, is considered. However, for the sake of clarity, $\boldsymbol{x}^0 \in \mathcal{F}_{\text{SPP}}$ is still supposed to be integer. This section extends the work done by Omer et al. [18], and Rosat et al. [23] and details it in the case of the SPP.

3.1.1 Generic fractional augmentation

To practically address **fAUG**, it must be formulated in such a way that it can algorithmically be solved. It consists in finding a direction $\boldsymbol{d} \in \mathbb{R}^n$ such that it is *feasible* ($\exists \rho > 0 | \boldsymbol{x}^k + \rho \boldsymbol{d} \in \mathcal{F}_{\text{SPP}^{\text{LR}}}$) and *augmenting* ($\boldsymbol{c}^T \boldsymbol{d} < 0$). The set of all feasible directions at $\boldsymbol{x}^0$ is the cone

$$\Gamma = \left\{ \boldsymbol{d} \in \mathbb{R}^n \mid \boldsymbol{A}\boldsymbol{d} = 0\,,\ \boldsymbol{d}_{\mathcal{P}} \leq \boldsymbol{0}\,,\ \boldsymbol{d}_{\mathcal{Z}} \geq \boldsymbol{0} \right\}, \tag{1}$$

from which we will only consider a section,

$$\Delta = \Gamma \cap \left\{ \boldsymbol{d} \in \mathbb{R}^n \mid \boldsymbol{e}^T \boldsymbol{d}_{\mathcal{Z}} = 1 \right\}. \tag{2}$$

The linear constraint $\boldsymbol{e}^T \boldsymbol{d}_{\mathcal{Z}} = 1$ is called the normalization constraint. A geometric interpretation of Γ and Δ is given in Fig. 2. Note that, since $\boldsymbol{A}$ is nonnegative and because of the sign constraints, any nonzero feasible direction $\boldsymbol{d} \in \Gamma$ has nonzero terms in both $\boldsymbol{d}_{\mathcal{P}}$ and $\boldsymbol{d}_{\mathcal{Z}}$. Hence, the normalization constraint defines a proper section of the cone and Δ is a (bounded) polytope.

It is easy to see that at least one feasible direction is augmenting if and only if the program

$$z^\star_{\text{MIMA}} = \min_{d \in \mathbb{R}^n} \left\{ \boldsymbol{c}^T \boldsymbol{d} \,|\boldsymbol{d} \in \Delta \right\} \tag{MIMA}$$

satisfies $z^\star_{\text{MIMA}} < 0$. On the one hand, any optimal solution of MIMA yields a solution to **fAUG**; on the other hand, if $z^\star_{\text{MIMA}}$ is nonnegative, no feasible augmenting direction exists and $\boldsymbol{x}^0$ is optimal for SPP. Finally, since Δ is a polytope, MIMA is a bounded linear program. The name MIMA stands for *Maximum Incoming Mean Augmentation*: The normalization constraint only concerns incoming variables, so the objective is a mean over a reduced subset of the variables. The choice of this normalization constraint follows that of the original algorithm of Zaghrouti et al. [37].

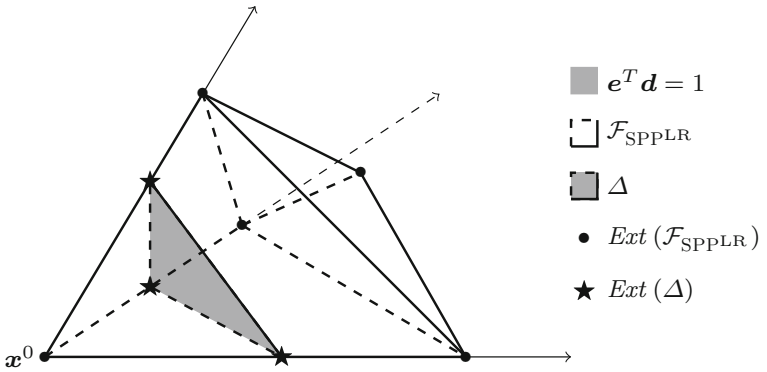


Fig. 2 Geometric description of Δ and Γ. The cone Γ of all feasible directions at $\boldsymbol{x}^0$ is defined by the *three arrowed lines*. The *grey area* represents Δ, the set of normalized feasible directions

Remark 2 The normalization constraint $\boldsymbol{e}^T \boldsymbol{d}_{\mathcal{Z}}$ can be replaced with any other equality of the form $\boldsymbol{w}^T \boldsymbol{d} = 1$, as long as Δ defines a proper section of Γ. On the theoretical side, this does not change the nature of the results presented in Sects. 3 and 4. In particular, the algorithm does not change, and neither do the separation algorithms. On the practical side, the choice of the normalization influences the performances of the algorithm. We use the aformentionned specific constraint in this paper for the following reasons: (1) it has been used in the original version of the algorithm [37] and therefore makes the comparison with it fairer and (2) most proofs are significantly more readable with this constraint than with a generic one. For a detailed study of the influence (both theroretical and practical) of the choice of the normalization in ISUD, the reader is referred to [23].

Once an optimal solution $\boldsymbol{d}^\star$ to MIMA has been found, the idea of an augmentation algorithm is to follow that direction as far as possible while remaining feasible. The *maximal feasible step* alongside direction $\boldsymbol{d}$ is thus defined as $r(\boldsymbol{d}) = \max\left\{\rho > 0 \,|\, \boldsymbol{x}^0 + \rho \boldsymbol{d} \in \mathcal{F}_{\mathrm{SPP^{LR}}}\right\}$. From $\boldsymbol{x}^0$, fractional augmentation can therefore be performed as $\boldsymbol{x}^1 = \boldsymbol{x}^0 + r(\boldsymbol{d})\boldsymbol{d}$.

3.1.2 Incompatibility degree of the nonbasic variables

To decompose MIMA into smaller problems, both in terms of number of variables and constraints, the notion of incompatibility degree introduced by Elhallaoui et al. [7] must be presented here. Given a column $\boldsymbol{A}_{\cdot j}$ of the constraint matrix, a row $r \in \mathcal{R}$ is said to be *covered* by that column if $A_{rj} = 1$. For each column $\boldsymbol{A}_{\cdot j}$ of $\boldsymbol{A}$, let $\mathcal{R}_j = \left\{r \in \mathcal{R} \,|\, A_{rj} = 1\right\}$ be the set of rows covered by that column. By definition, the sets of the rows covered by the columns of $\mathcal{P}$, i.e., $\{\mathcal{R}_l\}_{l \in \mathcal{P}}$, form a partition of $\mathcal{R}$ (see Fig. 3). Assume that a total order $\preceq$ is known over the indices of the rows $\mathcal{R}$. For any j, that order is extended to $\mathcal{R}_j$, and the elements of $\mathcal{R}_j$ are written according to $\preceq$ as

$$\boldsymbol{A} = \begin{array}{|cc|ccccc|} \multicolumn{2}{c}{\mathcal{P}} & \multicolumn{5}{c}{\mathcal{Z}} \\ \hline 1 & 0 & 1 & 1 & 1 & 1 & 1 \\ 1 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 0 & 0 & 0 & 1 \\ \hline 0 & 0 & 1 & 1 & 2 & 1 & 2 \\ \multicolumn{2}{c}{\iota_j^{\mathcal{P}}} & \multicolumn{5}{c}{} \end{array} \qquad \hat{\boldsymbol{A}} = \begin{array}{|cc|ccccc|} \multicolumn{2}{c}{\mathcal{P}} & \multicolumn{5}{c}{\mathcal{Z}} \\ \hline 0 & 1 & 0 & 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 1 & 1 & 1 & 1 \\ 1 & 0 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 1 & 1 & 1 & 1 & 1 \\ \hline 0 & 0 & 1 & 2 & 3 & 2 & 0 \\ \multicolumn{2}{c}{\iota_j^{\mathcal{P}}} & \multicolumn{5}{c}{} \end{array}$$

Fig. 3 Incompatibility degree with respect to $\mathcal{P}$ on a 5-rows example for two different ordering of the rows of the same matrix, namely (1, 2, 3, 4, 5) in matrix $\boldsymbol{A}$, and (3, 4, 2, 5, 1) in matrix $\hat{\boldsymbol{A}}$

$$\mathcal{R}_j \,:\, r_1^j \preceq r_2^j \preceq \cdots \preceq r_{|\mathcal{R}_j|}^j \,.$$

Definition 1 *(Incompatibility degree)* The *incompatibility degree* of a column $\boldsymbol{A}_{\cdot j}$, $j \in \{1, \ldots, n\}$, with respect to the working basis $\mathcal{P}$ is computed as

$$\iota_j^{\mathcal{P}} = \sum_{l \in \mathcal{P}} \sum_{t=1}^{|\mathcal{R}_l|-1} \kappa_{lj}^t, \tag{3}$$

where $\kappa_{lj}^t = 1$ if $\boldsymbol{A}_{\cdot j}$ covers r_t^j or r_{t+1}^j but not both, 2 if $\boldsymbol{A}_{\cdot j}$ covers r_t^j and r_{t+1}^j but not consecutively, 0 otherwise. Here, r_t^j and r_{t+1}^j denote two rows covered by column l that are performed consecutively within that column. An example is given in Fig. 3.

Remark 3 Any ordering of the rows can be used with these definitions. In scheduling applications, the rows correspond to tasks to carry out. It is therefore intuitive to order them by starting time and an incompatibility will correspond to the breaking of a succession of tasks that are performed consecutively in the current solution $\boldsymbol{x}^0$.

For the sake of clarity, we will always consider that the nonzero entries of a column of the current solution are consecutive within $\mathcal{R}$, i.e., given any pair of different indices $i, j \in \mathcal{P}$, either $\forall r \in \mathcal{R}_i, \forall r' \in \mathcal{R}_j, r_i \preceq r'_j$, or $\forall r \in \mathcal{R}_i, \forall r' \in \mathcal{R}_j, r'_j \preceq r_i$. Note that this is not the case of the left part of Fig. 3. With $\iota_j^{\mathcal{P}}$ as defined in Formula (3), $\boldsymbol{A}_{\cdot j}$ is said to be $\iota_j^{\mathcal{P}}$*-incompatible*. 0-incompatible columns are called *compatible* and the others are called *incompatible*. These notions extend to the index of the column and to the corresponding variable.

Observation 1 *All columns from the working basis are compatible;*

Observation 2 *An incompatible column is one that breaks the partition of the rows* $\{\mathcal{R}_l\}_{l \in \mathcal{P}}$.

Given a solution $\boldsymbol{x}^0$, we define the following subsets of nonbasic variables $\mathcal{Z}$ as

$$\mathcal{C} = \left\{ j \in \mathcal{Z} \,|\, \iota_j^{\mathcal{P}} = 0 \right\} \text{ and } \mathcal{I}^\iota = \left\{ j \in \mathcal{Z} \,|\, 1 \leq \iota_j^{\mathcal{P}} \leq \iota \right\}, \forall 1 \leq \iota \leq k_{max}. \tag{4}$$

Then, $\mathcal{C}$ is called the *compatible* set and the corresponding indices and variables are called *compatible*. Thus, $\mathcal{I}^\iota$ is called the *at most ι-incompatible* set and the corresponding indices and variables are said to be *at most ι-incompatible*. By definition, $\mathcal{I}^1 \subseteq \mathcal{I}^2 \subseteq \ldots \subseteq \mathcal{I}^{\iota max} = \mathcal{I}$. $\mathcal{I}$ is called the incompatible set, and its elements and the corresponding variables are called incompatible. Finally, $(\mathcal{P}, \mathcal{C}, \mathcal{I})$ form a partition of $\{1, \ldots, n\}$.

Lemma 1 *Given $j \in \mathcal{Z}$, the following statements are equivalent:*

(i) $\boldsymbol{A}_{\cdot j}$ *is compatible;*
(ii) $\exists \mathcal{P}_j \subseteq \mathcal{P} \mid \boldsymbol{A}_{\cdot j} = \sum_{l \in \mathcal{P}_j} \boldsymbol{A}_{\cdot l}$*;*
(iii) $\boldsymbol{A}_{\cdot j} \in Span\left(\boldsymbol{A}_{\cdot \mathcal{P}}\right)$.

Lemma 1 does not apply to the left part of the example given in Fig. 3 because the rows are not ordered properly in that case. The notion of compatible/incompatible column is extended to all vectors of $\mathbb{R}^m$, namely $\boldsymbol{w} \in \mathbb{R}^m$ is compatible if and only if $\boldsymbol{w} \in Span\left(\boldsymbol{A}_{\cdot \mathcal{P}}\right)$. In the theoretical part of this paper, we will consider the partition of $\{1, \ldots, n\}$ into $(\mathcal{P}, \mathcal{C}, \mathcal{I})$. However it is algorithmically efficient to look first for an augmenting direction over $\mathcal{C}$, and then, successively $\mathcal{I}^1, \mathcal{I}^2, \ldots$, until $\mathcal{I}$.

3.1.3 The IPS decomposition

With $\bar{\mathcal{P}} = \{1, \ldots, m\} \setminus \mathcal{P}$ and reordering the rows and columns of $\boldsymbol{A}$ so that $\mathcal{R}$ is partitioned into $(\mathcal{P}, \bar{\mathcal{P}})$, we can write

$$\boldsymbol{A} = \begin{bmatrix} \boldsymbol{I}_p & \boldsymbol{A}_{\mathcal{P}\mathcal{C}} & \boldsymbol{A}_{\mathcal{P}\mathcal{I}} \\ \boldsymbol{A}_{\bar{\mathcal{P}}\mathcal{P}} & \boldsymbol{A}_{\bar{\mathcal{P}}\mathcal{C}} & \boldsymbol{A}_{\bar{\mathcal{P}}\mathcal{I}} \end{bmatrix}, \tag{5}$$

where $\boldsymbol{I}_p$ is the $p \times p$ identity matrix. Given $\boldsymbol{d} \in \Delta$, from the constraints $\boldsymbol{A}\boldsymbol{d} = \boldsymbol{0}$, one can easily see that the aggregation of all columns corresponding to increasing variables (for which $d_j \geq 0$) $\boldsymbol{w} = \boldsymbol{A}_{\cdot \mathcal{Z}} \boldsymbol{d}_{\mathcal{Z}}$ is compatible. As in a reduced-gradient algorithm, we are in fact looking for an aggregate column $\boldsymbol{w}$ that can enter the working basis and take a positive value by lowering only some variables of $\mathcal{P}$. We introduce the following problems, respectively called *Restricted*-MIMA and *Complementary*-MIMA):

$$z^\star_{\text{R- MIMA}} = \min_{\boldsymbol{d} \in \mathbb{R}^{p+|\mathcal{C}|}} \left\{ \boldsymbol{c}_{\mathcal{P}}^T \boldsymbol{d}_{\mathcal{P}} + \boldsymbol{c}_{\mathcal{C}}^T \boldsymbol{d}_{\mathcal{C}} \mid \boldsymbol{A}_{\cdot \mathcal{P}} \boldsymbol{d}_{\mathcal{P}} + \boldsymbol{A}_{\cdot \mathcal{C}} \boldsymbol{d}_{\mathcal{C}} = 0\,,\ \boldsymbol{e}_{\mathcal{C}}^T \boldsymbol{d}_{\mathcal{C}} = 1\,,\ \boldsymbol{d}_{\mathcal{C}} \geq 0 \right\} \tag{R- MIMA}$$

$$z^\star_{\text{C- MIMA}} = \min_{\boldsymbol{d} \in \mathbb{R}^{p+|\mathcal{I}|}} \left\{ \boldsymbol{c}_{\mathcal{P}}^T \boldsymbol{d}_{\mathcal{P}} + \boldsymbol{c}_{\mathcal{I}}^T \boldsymbol{d}_{\mathcal{I}} \mid \boldsymbol{A}_{\cdot \mathcal{P}} \boldsymbol{d}_{\mathcal{P}} + \boldsymbol{A}_{\cdot \mathcal{I}} \boldsymbol{d}_{\mathcal{I}} = 0\,,\ \boldsymbol{e}_{\mathcal{I}}^T \boldsymbol{d}_{\mathcal{I}} = 1\,,\ \boldsymbol{d}_{\mathcal{I}} \geq 0 \right\}. \tag{C- MIMA}$$

Theorem 1 $z^\star_{\text{MIMA}} = \min\left\{ z^\star_{\text{R- MIMA}}, z^\star_{\text{C- MIMA}} \right\}$.

Proof Let $\boldsymbol{d} = (\boldsymbol{d}_{\mathcal{P}}, \boldsymbol{d}_{\mathcal{C}}, \boldsymbol{d}_{\mathcal{I}})$ be an optimal solution of MIMA. If $\boldsymbol{d}_{\mathcal{C}} = \boldsymbol{0}$ or $\boldsymbol{d}_{\mathcal{I}} = \boldsymbol{0}$, $\boldsymbol{d}$ is respectively a solution of C- MIMA or R- MIMA and the result clearly holds.

Suppose now that $\boldsymbol{d}_{\mathcal{C}} \neq \boldsymbol{0}$ and $\boldsymbol{d}_{\mathcal{I}} \neq \boldsymbol{0}$. We first prove that $\boldsymbol{d}$ can be written as a convex combination of $\boldsymbol{u}'$, a solution of R- MIMA, and $\boldsymbol{v}'$, a solution of C- MIMA. By Lemma 1-(ii), the surrogate column $\boldsymbol{A}_{\cdot\mathcal{C}}\boldsymbol{d}_{\mathcal{C}}$ can be written as a linear combination of columns of $\mathcal{P}$, $\boldsymbol{A}_{\cdot\mathcal{C}}\boldsymbol{d}_{\mathcal{C}} = -\boldsymbol{A}_{\cdot\mathcal{P}}\boldsymbol{u}'_{\mathcal{C}}, \boldsymbol{u}'_{\mathcal{P}} \leq \boldsymbol{0}$. Let $\boldsymbol{u} = (\boldsymbol{u}'_{\mathcal{P}}, \boldsymbol{d}_{\mathcal{C}}, \boldsymbol{0})$, and $\boldsymbol{v} = \boldsymbol{d} - \boldsymbol{u}$. Let $\alpha_{\boldsymbol{u}} = \|\boldsymbol{d}_{\mathcal{C}}\|_1 = \sum_{j\in\mathcal{C}} d_j$ and $\alpha_{\boldsymbol{v}} = \|\boldsymbol{d}_{\mathcal{I}}\|_1 = \sum_{j\in\mathcal{I}} d_j$. Moreover, let $\boldsymbol{u}' = \boldsymbol{u}/\alpha_{\boldsymbol{u}}$ and $\boldsymbol{v}' = \boldsymbol{v}/\alpha_{\boldsymbol{v}}$ be the corresponding normalized directions. Thus, $0 < \alpha_{\boldsymbol{u}}, \alpha_{\boldsymbol{v}}$ and, since $\boldsymbol{d}$ is a solution of MIMA, the normalization constraint yields $\alpha_{\boldsymbol{u}} + \alpha_{\boldsymbol{v}} = 1$ because $\boldsymbol{d} \in \mathcal{F}_{\text{MIMA}}$. Therefore, $\boldsymbol{d} = \alpha_{\boldsymbol{u}}\boldsymbol{u}' + \alpha_{\boldsymbol{v}}\boldsymbol{v}'$ is a convex combination of $\boldsymbol{u}'$ a solution of R- MIMA and $\boldsymbol{v}'$ a solution of C- MIMA.

Looking at the objective function, the convex combination reads $\boldsymbol{c}^T\boldsymbol{d} = \alpha_{\boldsymbol{u}}(\boldsymbol{c}^T\boldsymbol{u}') + \alpha_{\boldsymbol{v}}(\boldsymbol{c}^T\boldsymbol{v}')$. Either $\boldsymbol{c}^T\boldsymbol{d} \geq \boldsymbol{c}^T\boldsymbol{u}'$ or $\boldsymbol{c}^T\boldsymbol{d} \geq \boldsymbol{c}^T\boldsymbol{v}'$. However, since every solution of R- MIMA and C- MIMA is also a solution of MIMA and $\boldsymbol{d}$ is optimal for MIMA, $\boldsymbol{c}^T\boldsymbol{d} \leq \boldsymbol{c}^T\boldsymbol{u}'$ and $\boldsymbol{c}^T\boldsymbol{d} \leq \boldsymbol{c}^T\boldsymbol{v}'$. One of the two inequalities is thus an equality and the second follows from $\boldsymbol{c}^T\boldsymbol{d} = \alpha_{\boldsymbol{u}}(\boldsymbol{c}^T\boldsymbol{u}') + \alpha_{\boldsymbol{v}}(\boldsymbol{c}^T\boldsymbol{v}')$. Therefore, $\boldsymbol{c}^T\boldsymbol{d} = \boldsymbol{c}^T\boldsymbol{u}' = \boldsymbol{c}^T\boldsymbol{v}'$, and since $\boldsymbol{d}$ is an optimal solution of MIMA, $z^\star_{\text{MIMA}} = z^\star_{\text{R- MIMA}} = z^\star_{\text{C- MIMA}}$. □

In the previous works on IPS by Elhallaoui et al. [6], a weaker version of Theorem 1 stated that $z^\star_{\text{MIMA}}$ and $\min\left\{z^\star_{\text{R- MIMA}}, z^\star_{\text{C- MIMA}}\right\}$ have the same sign. Theorem 1 strengthens this result and makes the justification of most subsequent procedures more straightforward. This purely primal interpretation of their less-intuitive dual approach allows us to state a precise decomposition of MIMA into R- MIMA and C- MIMA. As a consequence of Theorem 1, we will consider the pair of problems R- MIMA and C- MIMA instead of the more complicated MIMA, and we will solve them sequentially. In particular, C- MIMA will not be solved if $z^\star_{\text{R- MIMA}} < 0$ because an improving direction is already known. In the next sections, we discuss how to reduce the number of rows in R- MIMA and C- MIMA to ease their solution. Recall that, for the moment, $\boldsymbol{x}^1 = \boldsymbol{x}^0 + r(\boldsymbol{d})\boldsymbol{d}$ may be fractional.

3.1.4 Row-reduction of R- MIMA

The practical solution of the restriction of **fAUG** to the compatible variables only is based on the following proposition. Recall that for all $i \in \mathcal{C}$, $\mathcal{P}_i \subseteq \mathcal{P}$ is defined as in Lemma 1-(ii), i.e., it is the unique subset of $\mathcal{P}$ such that $\boldsymbol{A}_{\cdot i} = \sum_{j\in\mathcal{P}_i} \boldsymbol{A}_{\cdot j}$.

Proposition 1 *Given $j \in \mathcal{C}$, the* minimal direction *associated with j is the vector $\boldsymbol{\delta}^j \in \mathbb{R}^{p+|\mathcal{C}|}$ defined as*

$$\forall i \in \mathcal{P}\cup\mathcal{C}\,,\ \delta^j_i = \begin{cases} -1 & \text{if } i \in \mathcal{P}_j, \\ 1 & \text{if } i = j, \\ 0 & \text{otherwise.} \end{cases} \tag{6}$$

For any $j \in \mathcal{C}$, $\boldsymbol{\delta}^j$ is feasible and the set of all extreme points of $\mathcal{F}_{\text{R- MIMA}}$ is exactly the set of the minimal directions associated with variables of $\mathcal{C}$, i.e., $\left\{\boldsymbol{\delta}^j \,|\, j \in \mathcal{C}\right\}$. Moreover, the maximal feasible step alongside direction $\boldsymbol{\delta}^j$ is $r(\boldsymbol{\delta}^j) = 1$.

Proof First, let $j \in \mathcal{C}$. By definition of $\boldsymbol{\delta}^j$ and $\mathcal{F}_{\text{R- MIMA}}$, $\boldsymbol{\delta}^j$ is obviously feasible for R- MIMA.

Second, let $\boldsymbol{d}^\star = (\boldsymbol{d}^\star_{\mathcal{P}}, \boldsymbol{d}^\star_{\mathcal{C}}) \in \mathcal{F}_{\text{R- MIMA}}$ be a feasible solution of R- MIMA. We will show that $\boldsymbol{d}$ is a convex combination of one or several minimal solutions. Let $\boldsymbol{u} = \sum_{j\in\mathcal{C}} d^\star_j \boldsymbol{\delta}^j$. By construction, $\boldsymbol{u}_{\mathcal{C}} = \boldsymbol{d}^\star_{\mathcal{C}}$. Because all $\boldsymbol{\delta}^j$ are in $\mathcal{F}_{\text{R- MIMA}}$, then the linearity constraints apply to their combination $\boldsymbol{u}$. Hence, $\boldsymbol{A}_{.\mathcal{P}}\boldsymbol{u}_{\mathcal{P}} = -\boldsymbol{A}_{.\mathcal{C}}\boldsymbol{u}_{\mathcal{C}} = -\boldsymbol{A}_{.\mathcal{C}}\boldsymbol{d}^\star_{\mathcal{C}} = \boldsymbol{A}_{.\mathcal{P}}\boldsymbol{d}^\star_{\mathcal{P}}$. The columns of the working basis $\boldsymbol{A}_{.\mathcal{P}}$ are linearly independent, therefore, $\boldsymbol{u}_{\mathcal{P}} = \boldsymbol{d}^\star_{\mathcal{P}}$ and $\boldsymbol{d}^\star = \boldsymbol{u}$.

Moreover, $\boldsymbol{d} \in \mathcal{F}_{\text{R- MIMA}}$ satisfies the normalization constraint $\boldsymbol{e}^T\boldsymbol{d}^\star_{\mathcal{C}} = \sum_{i\in\mathcal{C}} d^\star_i = 1$. Thus, $\boldsymbol{d}^\star = \boldsymbol{u} = \sum_{j\in\mathcal{C}} d^\star_j \boldsymbol{\delta}^j$ is a convex combination of minimal directions. Finally, given the current solution $\boldsymbol{x}^0$, for any $j \in \mathcal{C}$ and $\rho > 0$,

$$\forall i \in \mathcal{P}\cup\mathcal{C}\,,\ \left[\boldsymbol{x}^0 + \rho\boldsymbol{\delta}^j\right]_i = \begin{cases} 1-\rho & \text{if } i \in \mathcal{P}_j, \\ \rho & \text{if } i = j, \\ x^0_i & \text{otherwise.} \end{cases}$$

Therefore, considering the bounds $\mathbf{0} \le (\boldsymbol{x}^0 + \rho\boldsymbol{d}) \le \boldsymbol{e}$, the maximum possible value for ρ is $r(\boldsymbol{d}^0) = 1$. □

Corollary 1 *There exists $j \in \mathcal{P}$ such that $\boldsymbol{\delta}^j$ is an optimal solution of* R- MIMA.

As a consequence of Proposition 1 and Corollary 1, solving R- MIMA and following the optimal direction until a bound is reached is equivalent to pivoting the corresponding compatible variable into the working basis. These pivots are guaranteed to be nondegenerate, as proven in the following proposition.

Proposition 2 *Compatible variables are exactly those that yield nondegenerate pivots when inserted in the working basis.*

Proof For SPP, when inserted in the working basis, a nonbasic variable yields a nondegenerate pivot iff it can be written as a nonnegative linear combination of the variables in $\mathcal{P}$ ($\boldsymbol{x}^0_{\mathcal{P}} = \mathbf{0}$ and $\mathbf{0} \le \boldsymbol{x} \le \mathbf{1}$). By Lemma 1-$(iii)$, compatible columns are exactly those that can be written as a linear combination of the columns of $\boldsymbol{A}_{.\mathcal{P}}$. Therefore, if a column yields a nondegenerate pivot, it must be compatible. As proven in Proposition 1 pivoting a compatible variable x_j in the working basis can be interpreted as following direction $\boldsymbol{\delta}^j$. Since the corresponding maximal step is positive ($r(\boldsymbol{\delta}^j) = 1$), the pivot is nondegenerate and compatible variables all yield nondegenerate pivots. □

Proposition 3 R- MIMA *is equivalent to the minimization program*

$$z^\star_{\text{R- MIMA}} = z^\star_{\text{RP}} = \min_{j\in\mathcal{C}} \left\{\bar{c}_j\right\}, \tag{RP}$$

called reduced problem, *where $\bar{c} = \boldsymbol{c} - \boldsymbol{c}^T_{\mathcal{P}}\boldsymbol{A}_{\mathcal{P}.}$ is the reduced costs vector. Moreover, given an optimal solution $j \in \mathcal{C}$, the corresponding direction is $\boldsymbol{\delta}^j$.*

Proof Given an index $j \in \mathcal{C}$, the cost of the corresponding minimal direction in R-MIMA is $\boldsymbol{c}^T \boldsymbol{\delta}^j = \boldsymbol{c}_{\mathcal{P}}^T \boldsymbol{\delta}_{\mathcal{P}}^j + \boldsymbol{c}_{\mathcal{C}}^T \boldsymbol{\delta}_{\mathcal{C}}^j$. With Eq. (5), the linear constraints of R- MIMA become

$$\begin{cases} \boldsymbol{I}_p \boldsymbol{\delta}_{\mathcal{P}}^j + \boldsymbol{A}_{\mathcal{P}\mathcal{C}} \boldsymbol{\delta}_{\mathcal{C}}^j = 0 \\ \boldsymbol{A}_{\bar{\mathcal{P}}\mathcal{P}} \boldsymbol{\delta}_{\mathcal{P}}^j + \boldsymbol{A}_{\bar{\mathcal{P}}\mathcal{C}} \boldsymbol{\delta}_{\mathcal{C}}^j = 0 \end{cases},$$

and, with Eq. (6), the first row gives $\boldsymbol{\delta}_{\mathcal{P}}^j = -\boldsymbol{A}_{\mathcal{P}j}$, and thus, $\boldsymbol{c}^T \boldsymbol{\delta}^j = c_j - \boldsymbol{c}_{\mathcal{P}}^T \boldsymbol{A}_{\mathcal{P}j}$. Because the extreme points of $\mathcal{F}_{\text{R- MIMA}}$ are the $\boldsymbol{\delta}^j$ for $j \in \mathcal{C}$, the result holds. □

Note that the previous result can be extended for any choice of normalization weights. Namely, if the normalization constraint reads as $\boldsymbol{w}^T \boldsymbol{d} = 1$, then $\bar{c}_j$ is replaced by $\bar{c}_j / w_j$ in equation (RP) and the result holds (see, [23] for more details).

The term of row-reduction refers to the following two facts. First, the linear program becomes a simple reduced-cost determination. Second, the computation of the whole improving direction is made without any matrix multiplication since $\boldsymbol{\delta}_{\mathcal{P}}^j = -\boldsymbol{A}_{\mathcal{P}j}$. This is in the spirit of a standard simplex pivot in which a reduced-cost analysis is performed, and then a system must be solved to determine the future solution $\boldsymbol{x}^{k+1}$. As in Proposition 3, in practice, the determination of $j \in \mathcal{C}$ is made as a reduced-cost analysis and exactly reproduces what a simplex pivot would be. Namely, the nonbasic variable of lowest reduced-cost $z^\star_{\text{R- MIMA}}$ is determined, and then, if it satisfies $z^\star_{\text{R- MIMA}} < 0$, a pivot is performed by inserting that variable into the working basis. However, if $z^\star_{\text{R- MIMA}} \geq 0$, it becomes necessary to consider C- MIMA to determine an augmenting direction. This process is the topic of the following section.

3.1.5 Row-reduction of C- MIMA

In this section, we suppose that the compatible variables yield no improvement, or equivalently that $\mathcal{C} = \emptyset$. The first p constraints of C- MIMA read $\boldsymbol{d}_{\mathcal{P}} = -\boldsymbol{A}_{\mathcal{P}\mathcal{I}} \boldsymbol{d}_{\mathcal{I}}$. As in a reduced gradient algorithm (e.g., [13]), the modification of the basic variables is inferred via a linear transformation of the increasing nonbasic variables $\boldsymbol{d}_{\mathcal{Z}}$, or in the case of C- MIMA, $\boldsymbol{d}_{\mathcal{I}}$. Hence, we reduce C- MIMA to an equivalent problem over the nonbasic variables of $\mathcal{I}$ only. For the sake of clarity, denote as $\Delta_{\mathcal{I}}$ the feasible domain of C- MIMA, i.e., all elements of Δ that satisfy $\boldsymbol{d}_{\mathcal{C}} = 0$, and define $q = |\mathcal{I}|$. That domain can be defined by less linear constraints and variables than Δ as shown by Proposition 4.

Proposition 4 *With* $\bar{\Delta_{\mathcal{I}}} = \left\{ \boldsymbol{d}_{\mathcal{I}} \in \mathbb{R}^q \mid \bar{\boldsymbol{A}}_{\bar{\mathcal{P}}\mathcal{I}} \boldsymbol{d}_{\mathcal{I}} = \boldsymbol{0}\,,\ \boldsymbol{e}^T \boldsymbol{d}_{\mathcal{I}} = 1\,,\ \boldsymbol{d}_{\mathcal{I}} \geq \boldsymbol{0} \right\}$, *then*

$$\Delta_{\mathcal{I}} = \left\{ \boldsymbol{d} = \boldsymbol{T} \boldsymbol{d}_{\mathcal{I}} \mid \boldsymbol{d}_{\mathcal{I}} \in \bar{\Delta_{\mathcal{I}}} \right\}, \tag{7}$$

where $\boldsymbol{T} = \left[-\boldsymbol{A}_{\mathcal{P}\mathcal{I}} \boldsymbol{I}_q \right]$ *and* $\bar{\boldsymbol{A}}_{\bar{\mathcal{P}}\mathcal{I}} = -\boldsymbol{A}_{\bar{\mathcal{P}}\mathcal{P}} \boldsymbol{A}_{\mathcal{P}\mathcal{I}} + \boldsymbol{A}_{\bar{\mathcal{P}}\mathcal{I}} = \boldsymbol{A}_{\bar{\mathcal{P}}\cdot} \boldsymbol{T}$.

Proof Let $\boldsymbol{d} \in \mathbb{R}^n$ such that $\boldsymbol{d}_{\mathcal{C}} = \boldsymbol{0}$. Then,

$$\boldsymbol{d} \in \Delta_{\mathcal{I}} \Leftrightarrow \boldsymbol{A}_{\cdot\mathcal{P}} \boldsymbol{d}_{\mathcal{P}} + \boldsymbol{A}_{\cdot\mathcal{I}} \boldsymbol{d}_{\mathcal{I}} = 0\,,\ \boldsymbol{e}^T \boldsymbol{d}_{\mathcal{I}} = 1\,,\ \boldsymbol{d}_{\mathcal{P}} \leq \boldsymbol{0}\,,\ \boldsymbol{d}_{\mathcal{I}} \geq \boldsymbol{0}$$

$$
\begin{aligned}
&\Leftrightarrow \begin{cases} \boldsymbol{d}_{\mathcal{P}} + \boldsymbol{A}_{\mathcal{P}\mathcal{I}}\boldsymbol{d}_{\mathcal{I}} = \boldsymbol{0} \\ \boldsymbol{A}_{\bar{\mathcal{P}}\mathcal{P}}\boldsymbol{d}_{\mathcal{P}} + \boldsymbol{A}_{\bar{\mathcal{P}}\mathcal{I}}\boldsymbol{d}_{\mathcal{I}} = \boldsymbol{0} \\ \boldsymbol{e}^T\boldsymbol{d}_{\mathcal{I}} = 1 \\ \boldsymbol{d}_{\mathcal{P}} \leq 0, \boldsymbol{d}_{\mathcal{I}} \geq 0 \end{cases} \\
&\Leftrightarrow \begin{cases} \boldsymbol{d}_{\mathcal{P}} = -\boldsymbol{A}_{\mathcal{P}\mathcal{I}}\boldsymbol{d}_{\mathcal{I}} \\ \left(-\boldsymbol{A}_{\bar{\mathcal{P}}\mathcal{P}}\boldsymbol{A}_{\mathcal{P}\mathcal{I}} + \boldsymbol{A}_{\bar{\mathcal{P}}\mathcal{I}}\right)\boldsymbol{d}_{\mathcal{I}} = \boldsymbol{0} \\ \boldsymbol{e}^T\boldsymbol{d}_{\mathcal{I}} = 1 \\ \boldsymbol{d}_{\mathcal{I}} \geq 0 \end{cases} \\
&\Leftrightarrow \boldsymbol{d} = \boldsymbol{T}\boldsymbol{d}_{\mathcal{I}} \text{ and } \boldsymbol{d}_{\mathcal{I}} \in \bar{\Delta}_{\mathcal{I}}.
\end{aligned}
$$

Hence, the proposition holds. □

The cost of such a direction can be computed as $\boldsymbol{c}^T\boldsymbol{d} = \boldsymbol{c}^T\boldsymbol{T}\boldsymbol{d}_{\mathcal{I}}$. Define $\bar{\boldsymbol{c}}^T = \boldsymbol{c}^T\boldsymbol{T} = \boldsymbol{c}_{\mathcal{I}}^T - \boldsymbol{c}_{\mathcal{P}}^T\boldsymbol{A}_{\mathcal{P}\mathcal{I}}$, and C- MIMA is equivalent to the following program, called the *complementary problem*:

$$
z^{\star}_{\text{C- MIMA}} = z^{\star}_{\text{CP}} = \min\left\{\bar{c}^T\boldsymbol{d}_{\mathcal{I}} \mid \boldsymbol{d}_{\mathcal{I}} \in \bar{\Delta}_{\mathcal{I}}\right\}. \tag{CP}
$$

The cost vector of CP is the reduced-costs vector associated with all incompatible variables, i.e., their own cost ($\boldsymbol{c}_{\mathcal{I}}^T$) minus the marginal impact they have on the values of the variables of $\mathcal{P}$ if they enter the reduced-basis ($-\boldsymbol{c}_{\mathcal{P}}^T\boldsymbol{A}_{\mathcal{P}\mathcal{I}}$).

Remark 4 In Proposition 1, for $j \in \mathcal{C}$, we called $\boldsymbol{\delta}^j$ a *minimal direction*. This denomination can be extended to directions of $\bar{\Delta}_{\mathcal{I}}$ (while still applying to those of $\bar{\Delta}_{\mathcal{C}}$) as follows: $\boldsymbol{d} \in \bar{\Delta}_{\mathcal{I}}$ is a *minimal direction* iff there exists no other direction whose support is strictly contained in that of $\boldsymbol{d}$. In that sense, the set of minimal directions of $\bar{\Delta}$ is exactly $Ext\left(\bar{\Delta}_{\mathcal{C}}\right) \cup Ext\left(\bar{\Delta}_{\mathcal{I}}\right)$ (see [23] for more details on the geometrical structure of Δ and $\bar{\Delta}$).

3.2 Integral augmentation iAUG

The previous section provides a method to find a fractional augmentation. If $z^{\star}_{\text{RP}} < 0$ or $z^{\star}_{\text{CP}} < 0$, then the corresponding solution of negative reduced cost $\boldsymbol{d} \in \Delta$ yields a new solution $\boldsymbol{x}^1 = \boldsymbol{x}^0 + r(\boldsymbol{d})\boldsymbol{d} \in \mathcal{F}_{\text{SPP}^{\text{LR}}}$. However, nothing guarantees that $\boldsymbol{x}^1$ is integral. In this section, we characterize directions that lead to an integral solution $\boldsymbol{x}^1$. Such directions are called *integral directions* as opposed to *fractional directions*. The set of all integral directions within Δ is denoted as Δ^{int}. These names apply to $\boldsymbol{d}$, and its restrictions $\boldsymbol{d}_{\mathcal{Z}}$, $\boldsymbol{d}_{\mathcal{C}}$ or $\boldsymbol{d}_{\mathcal{I}}$ depending on the context. We also give some insights on the structure of the set of all integral directions, and a generic framework for our algorithm.

3.2.1 Over compatible variables

Note that for any $j \in \mathcal{C}$, the feasible solution $\boldsymbol{\delta}^j$ of RP is an integral direction. Indeed, $\boldsymbol{A}_{\cdot j} = \sum_{i \in \mathcal{P}_j} \boldsymbol{A}_{\cdot i}$ and following a step of length $r(\boldsymbol{\delta}^j) = 1$ in that direction is

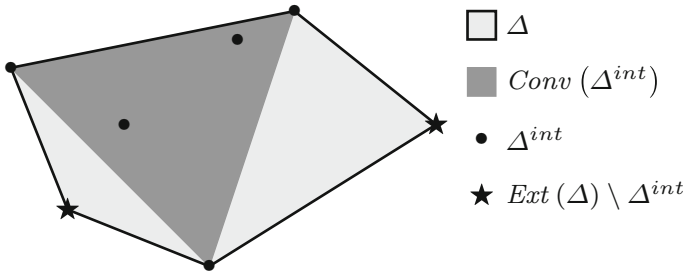


Fig. 4 Geometrical insight on the extreme points of Δ and Δ^{int}. All extreme points of $Conv\left(\Delta^{int}\right)$ ($\bullet$) are extreme points of Δ. Moreover, Δ^{int} is a finite set ($\mathcal{F}_{\text{SPP}}$ is finite)

equivalent to let j enter the working basis $\mathcal{P}$ and let $\mathcal{P}_j$ leave it. Namely,

$$\left[\boldsymbol{x}^0 + \rho\boldsymbol{\delta}^j\right]_i = \begin{cases} 1 & \text{if } i \in \mathcal{P} \setminus \mathcal{P}_j \\ 1-\rho & \text{if } i \in \mathcal{P}_j \\ \rho & \text{if } i = j \\ 0 & \text{otherwise.} \end{cases}$$

The corresponding maximal step is $\rho = r(\boldsymbol{\delta}^j) = 1$, so $\boldsymbol{x}^1 = \boldsymbol{x}^0 + \boldsymbol{\delta}^j \in \mathcal{F}_{\text{SPP}}$ and integrality is maintained.

3.2.2 Characterization of integral directions in $\Delta_{\mathcal{I}}$

Now, let us consider the problem for C- MIMA (and CP). We base our analysis on previous results of Zaghrouti et al. [37] that we first recall here. For any polyhedron P, denote as $Ext\,(P)$ the set of its extreme points (or vertices).

Definition 2 A set $\mathcal{S}$ of indices of the variables in $\{1, \ldots, n\}$ is *column-disjoint* if no pair of columns of $\boldsymbol{A}_{.\mathcal{S}}$ has a common nonzero entry, i.e., $\forall i, j \in \mathcal{S}, i \neq j, \forall k \in \{1, \ldots, m\}$, $A_{ki}A_{kj} = 0$. This definition is extended to $\boldsymbol{A}_{.\mathcal{S}}$ and $\boldsymbol{x}_{\mathcal{S}}$. Since $\boldsymbol{A}$ is binary, $\mathcal{S}$ is column-disjoint iff $\boldsymbol{A}_{.\mathcal{S}}$ is a set of orthogonal columns.

Proposition 5 *(*Propositions 6 and 7, *[37]) Given $\boldsymbol{d} \in Ext\,(\Delta)$, $\boldsymbol{d}$ is an integral direction iff the support of $\boldsymbol{d}_{\mathcal{I}}$, denoted as $S = Supp\ (\boldsymbol{d}_{\mathcal{I}})$, is column-disjoint. In this case, $r(\boldsymbol{d}) = |S|$.*

Because SPP is quasi-integral, then the edges of $Conv\ (\mathcal{F}_{\text{SPP}})$ are edges of $\mathcal{F}_{\text{SPP}^{\text{LR}}}$, and $Ext\left(\Delta_{\mathcal{I}}^{int}\right) \subseteq Ext\,(\Delta)$. A geometrical description of this is shown on Fig. 4. Since every linear program has at least one optimal solution that is also an extreme point of its feasible domain, and since the simplex algorithm guarantees to find such a solution, Proposition 5 has a strong practical interest. In the next section, we will show that it still remains valid when cutting planes or branching techniques are used. Hence, it is sufficient to test whether the solution of the relaxation (CP) is column-disjoint to determine if it is integral.

3.2.3 Algorithmic framework

Algorithm 1: Integral Simplex Using Decomposition

Input: $\boldsymbol{x}^0$, a solution of SPP; INC_MAX, maximal incompatibility degree considered.
Output: $\boldsymbol{x}^k$, a possibly better solution of SPP.

1 Compute $\mathcal{P}$ and $\mathcal{C}$ associated with $\boldsymbol{x}^0$; $k \leftarrow 0$; $\iota \leftarrow 1$;
2 **while** *true* **do**
3 If necessary, update $\mathcal{P}$, $\mathcal{C}$, and $\mathcal{I}^\iota$ associated with $\boldsymbol{x}^k$;
4 **if** $z^\star_{\text{RP}} < 0$ **then**
5 $\boldsymbol{\delta}^i \leftarrow$ an optimal solution of RP ($\bar{c}_i < 0, i \in \mathcal{C}$);
6 $\boldsymbol{x}^{k+1} \leftarrow \boldsymbol{x}^k + \boldsymbol{\delta}^i$; $k \leftarrow k+1$;
7 **else**
8 continue $\leftarrow$ true;
9 **while** *continue = true* **do**
10 If necessary, update $\mathcal{P}$, $\mathcal{C}$, and $\mathcal{I}^\iota$ associated with $\boldsymbol{x}^k$, and CP$^\iota$;
11 **if** $z^\star_{\text{CP}^\iota} < 0$ **then**
12 $\boldsymbol{d}^\star_{\mathcal{I}} \leftarrow$ an optimal solution of CP$^\iota$; $\boldsymbol{d}^\star \leftarrow \boldsymbol{T}\boldsymbol{d}^\star_{\mathcal{I}}$;
13 **if** $\boldsymbol{d}^\star_{\mathcal{I}}$ *is column-disjoint* **then**
14 $\boldsymbol{x}^{k+1} \leftarrow \boldsymbol{x}^k + r(\boldsymbol{d}^\star)\boldsymbol{d}^\star$; $k \leftarrow k+1$; continue $\leftarrow$ false;
15 **else**
16 **if** *some stopping criterion is reached* **then**
17 **return** $\boldsymbol{x}^k$;
18 **else**
• Add cutting planes to CP$^\iota$;
• or use branching strategy;
• or increase ι (only if $\iota <$ INC_MAX);
• or continue $\leftarrow$ false;
19 **else**
20 **if** $\iota <$ INC_MAX **then**
21 $\iota \leftarrow \iota + 1$;
22 **else**
23 **return** $\boldsymbol{x}^k$;

Algorithm 1 is based on successive resolutions of augmenting problems either to $\mathcal{C}$ or to $\mathcal{I}^\iota$ for a certain incompatibility degree ι. Hence, the incompatibility degree defines a neighborhood of the current solution on which the augmentation problem is solved with integer programming techniques. CP$^\iota$ describes the restriction of CP to the columns of $\mathcal{I}^\iota$ and INC_MAX describes the largest incompatibility degree considered during the execution. Whenever INC_MAX $\geq \iota_{max}$, and provided that one uses exhaustive primal cutting planes families such as Gomory-Young inequalities (line 18 of Algorithm 1), the algorithm is exact. When it is not the case, the solution returned by Algorithm 1 may not be optimal: This strongly depends on the value of INC_MAX, and on the chosen branching and cutting planes techniques. For instance, if INC_MAX $< \iota_{max}$, there may exist augmenting integral directions involving columns of incompatibility degree greater than INC_MAX. The same holds when the branching technique is nonexhaustive. This kind of heuristic stopping criterion of the algorithm makes the execution more efficient in practice.

4 Solving the augmentation problem with cutting-planes

In this section, we suppose that we know an optimal solution $\boldsymbol{d}_{\mathcal{I}}^{\star}$ of CP, but that this solution is not integral (*Supp* $\left(\boldsymbol{d}_{\mathcal{I}}^{\star}\right)$ is not column-disjoint). An idea to tighten the known relaxation of $\bar{\Delta}_{\mathcal{I}}^{int}$ is to add cutting planes to CP. Given a polyhedron P, a valid inequality for P is an inequality satisfied by all elements of P; and given a point $\boldsymbol{x}'$, the inequality *separates* $\boldsymbol{x}'$ from P if it is valid for P but violated by $\boldsymbol{x}'$. Such an inequality is called a cut (or cutting plane). The main issue is to characterize valid inequalities for $\bar{\Delta}_{\mathcal{I}}^{int}$ that cut off the current optimal solution $\boldsymbol{d}_{\mathcal{I}}^{\star}$.

Notation. Denote as $\mathcal{H}_{=}^{1} = \left\{\boldsymbol{d}_{\mathcal{I}} \in \mathbb{R}^{q} \mid \boldsymbol{e}^{T}\boldsymbol{d}_{\mathcal{I}} = 1\right\}$ the hyperplane defined by the normalization constraint of $\bar{\Delta}_{\mathcal{I}}$. Let $\bar{\boldsymbol{\alpha}} \in \mathbb{R}^{q}$, $\bar{\beta} \in \mathbb{R}$, and denote by $\left(\bar{\Gamma}\right)$ the inequality

$$\left(\bar{\Gamma}\right) \; : \; \bar{\boldsymbol{\alpha}}^{T}\boldsymbol{d}_{\mathcal{I}} \leq \bar{\beta}. \tag{8}$$

The associated hyperplane is denoted as $\mathcal{H}_{=}^{\bar{\Gamma}} = \left\{\boldsymbol{x} \in \mathbb{R}^{q} | \bar{\boldsymbol{\alpha}}^{T}\boldsymbol{d}_{\mathcal{I}} = \bar{\beta}\right\}$ and the associated half-space as $\mathcal{H}_{\leq}^{\bar{\Gamma}} = \left\{\boldsymbol{x} \in \mathbb{R}^{q} | \bar{\boldsymbol{\alpha}}^{T}\boldsymbol{d}_{\mathcal{I}} \leq \bar{\beta}\right\}$. Without loss of generality, we can assume that $\bar{\boldsymbol{\alpha}}$ is not proportional to $\boldsymbol{e}$.

Definition 3 Let $\left(\bar{\boldsymbol{\alpha}}_1, \bar{\beta}_1\right), \left(\bar{\boldsymbol{\alpha}}_2, \bar{\beta}_2\right) \in \mathbb{R}^{q} \times \mathbb{R}$, and $\left(\bar{\Gamma}_1\right)$ and $\left(\bar{\Gamma}_2\right)$ be the corresponding inequalities. $\left(\bar{\Gamma}_1\right)$ and $\left(\bar{\Gamma}_2\right)$ are *equivalent* for $\bar{\Delta}_{\mathcal{I}}^{int}$ if

$$\mathcal{H}_{=}^{1} \cap \mathcal{H}_{\leq}^{\bar{\Gamma}_1} = \mathcal{H}_{=}^{1} \cap \mathcal{H}_{\leq}^{\bar{\Gamma}_2}.$$

Since $\bar{\Delta}_{\mathcal{I}} \subseteq \mathcal{H}_{=}^{1}$, two equivalent inequalities exactly discard the same subset of $\bar{\Delta}_{\mathcal{I}}$ and are simultaneously valid. Hence, adding the one or the other to the formulation is equivalent.

Proposition 6 *There exists* $\bar{\boldsymbol{\alpha}}' \in \mathbb{R}^{q}$ *such that* $\left(\bar{\Gamma}'\right) : \left(\bar{\boldsymbol{\alpha}}'\right)^{T}\boldsymbol{d}_{\mathcal{I}} \leq 0$ *is equivalent to* $\left(\bar{\Gamma}\right)$.

Proof Consider $\mathcal{E}_{=} = \mathcal{H}_{=}^{1} \cap \mathcal{H}_{=}^{\bar{\Gamma}}$. The two hyperplanes $\mathcal{H}_{=}^{1}$ and $\mathcal{H}_{=}^{\bar{\Gamma}}$ are not parallel, so $\dim(\mathcal{E}_{=}) = q - 2$. $\mathbf{0} \notin \mathcal{H}_{=}^{1}$, therefore $\mathbf{0} \notin \mathcal{E}_{=}$ and $\mathcal{H}_{=}^{0} = Span\left(\mathcal{E}_{=} \cup \{\mathbf{0}\}\right)$ is a hyperplane containing $\mathbf{0}$. Thus, there exists $\bar{\boldsymbol{\alpha}}' \in \mathbb{R}^{q}$ such that $\mathcal{H}_{=}^{0}$ is defined by $\left(\bar{\boldsymbol{\alpha}}'\right)^{T}\boldsymbol{d}_{\mathcal{I}} = 0$. Furthermore, by construction, $\mathcal{H}_{=}^{0} \cap \mathcal{H}_{=}^{1} = \mathcal{E}_{=} = \mathcal{H}_{=}^{\bar{\Gamma}} \cap \mathcal{H}_{=}^{1}$. With $\bar{\boldsymbol{\alpha}}' = \pm\bar{\boldsymbol{\alpha}}^{0}$ (sign to be well chosen), the proposition holds. □

The geometrical interpretation of Proposition 6 is given on Fig. 5. Hereinafter, we suppose that $\bar{\beta} = 0$. We can now characterize valid inequalities for $\bar{\Delta}_{\mathcal{I}}^{int}$.

Proposition 7 *Given* $\boldsymbol{\alpha} \in \mathbb{R}^{n}$ *such that* $\bar{\boldsymbol{\alpha}} = \boldsymbol{T}^{T}\boldsymbol{\alpha}$, $\left(\bar{\Gamma}\right)$ *is a valid inequality for* $\bar{\Delta}_{\mathcal{I}}^{int}$ *if and only if*

$$(\Gamma) \; : \; \boldsymbol{\alpha}^{T}\left(\boldsymbol{x} - \boldsymbol{x}^{0}\right) \leq 0 \tag{9}$$

is a valid inequality for $\mathcal{F}_{\text{SPP}}$.

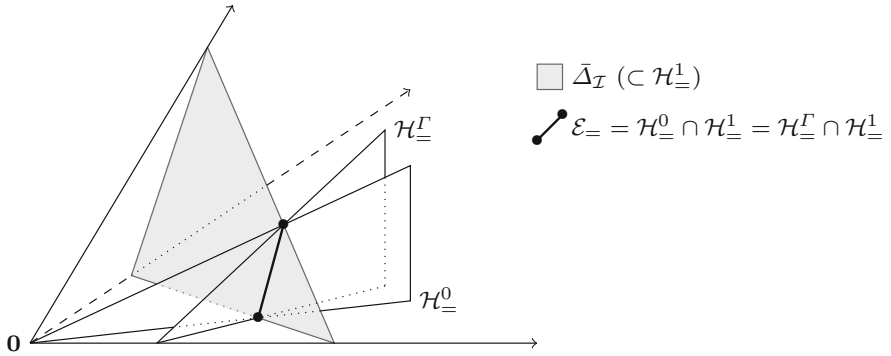


Fig. 5 Equivalent inequalities. Both $\mathcal{H}^0_=$ and $\mathcal{H}^\Gamma_=$ cut off the same part of $\bar{\Delta}_{\mathcal{I}}$. $\mathcal{H}^0_=$ is of the form $\bar{\boldsymbol{\alpha}}^T \boldsymbol{d}_{\mathcal{I}} \leq 0$ but not $\mathcal{H}^\Gamma_=$ ($\mathbf{0} \notin \mathcal{H}^\Gamma_=$)

Proof Let $\boldsymbol{\alpha} \in \mathbb{R}^n$ such that $\bar{\boldsymbol{\alpha}} = \boldsymbol{T}^T \boldsymbol{\alpha}$. Recall Proposition 4: $\Delta_{\mathcal{I}} = \{\boldsymbol{d} = \boldsymbol{T}\boldsymbol{d}_{\mathcal{I}} \mid \boldsymbol{d}_{\mathcal{I}} \in \bar{\Delta}_{\mathcal{I}}\}$. This extends to $\Delta^{int}_{\mathcal{I}}$ and $\bar{\Delta}^{int}_{\mathcal{I}}$. Hence,

$$\begin{aligned}
\left(\bar{\Gamma}\right) \text{ is valid for } \bar{\Delta}^{int}_{\mathcal{I}} &\Leftrightarrow \forall \boldsymbol{d}_{\mathcal{I}} \in \bar{\Delta}^{int}_{\mathcal{I}}, \bar{\boldsymbol{\alpha}}^T \boldsymbol{d}_{\mathcal{I}} \leq 0 \\
&\Leftrightarrow \forall \boldsymbol{d}_{\mathcal{I}} \in \bar{\Delta}^{int}_{\mathcal{I}}, \boldsymbol{\alpha}^T \boldsymbol{T}\boldsymbol{d}_{\mathcal{I}} \leq 0 \\
&\Leftrightarrow \forall \boldsymbol{d} \in \Delta^{int}_{\mathcal{I}}, \boldsymbol{\alpha}^T \boldsymbol{d} \leq 0 \\
&\Leftrightarrow \forall \boldsymbol{d} \in \Delta^{int}_{\mathcal{I}}, \boldsymbol{\alpha}^T \left(\boldsymbol{x}^0 + r(\boldsymbol{d})\boldsymbol{d}\right) \leq \boldsymbol{\alpha}^T \boldsymbol{x}^0 \\
&\Leftrightarrow \forall \boldsymbol{x}' \in \mathcal{F}_{\text{SPP}}, \boldsymbol{\alpha}^T \left(\boldsymbol{x}' - \boldsymbol{x}^0\right) \leq 0 \\
&\Leftrightarrow (\Gamma) \text{ is valid for SPP.}
\end{aligned}$$

□

Proposition 8 *Let* (Γ) *and* $\left(\bar{\Gamma}\right)$ *be valid inequalities defined as in Proposition 7, and* $\boldsymbol{d}^\star_{\mathcal{I}} \in \Delta_{\mathcal{I}}$, $\boldsymbol{d}^\star = \boldsymbol{T}\boldsymbol{d}_{\mathcal{I}}$, *and* $\boldsymbol{x}^\star = \boldsymbol{x}^0 + r(\boldsymbol{d}^\star)\boldsymbol{d}^\star$. *Then,*

$$\left(\bar{\Gamma}\right) \textit{ separates } \boldsymbol{d}^\star_{\mathcal{I}} \textit{ from } \bar{\Delta}^{int}_{\mathcal{I}} \Leftrightarrow (\Gamma) \textit{ separates } \boldsymbol{x}^\star \textit{ from } \mathcal{F}_{\text{SPP}}.$$

Proof We only need to prove that $\left(\bar{\Gamma}\right)$ is violated by $\boldsymbol{d}^\star_{\mathcal{I}}$ if and only if (Γ) is violated by $\boldsymbol{x}^\star$. We have

$$\begin{aligned}
\bar{\boldsymbol{\alpha}}^T \boldsymbol{d}^\star > 0 &\Leftrightarrow \boldsymbol{\alpha}^T \boldsymbol{d}^\star > 0 \\
&\Leftrightarrow \boldsymbol{\alpha}^T \left(\boldsymbol{x}^0 + r(\boldsymbol{d}^\star)\boldsymbol{d}^\star\right) > \boldsymbol{\alpha}^T \boldsymbol{x}^0 \\
&\Leftrightarrow \boldsymbol{\alpha}^T \left(\boldsymbol{x}^\star - \boldsymbol{x}^0\right) > 0.
\end{aligned}$$

□

What we have shown must now be seen the other way round to take advantage of previous work on primal separation. Assume that we know how to determine a primal

cut (Γ) for SPP that separates $\boldsymbol{x}^\star$ from $\mathcal{F}_{\text{SPP}}$. Then, with $\bar{\boldsymbol{\alpha}} = \boldsymbol{T}^T\boldsymbol{\alpha}$, the associated inequality $\left(\bar{\Gamma}\right)$ is a cut for CP, that separates $\boldsymbol{d}^\star_{\mathcal{I}}$ from $\bar{\Delta}^{int}_{\mathcal{I}}$. Moreover, we have shown that any cut for CP can be obtained in this way. This enables us to develop a procedure based on the primal separation problem **P-SEP** – given $\boldsymbol{x}^0$ and $x^\star$, is there a valid inequality for $\mathcal{F}_{\text{SPP}}$, tight at $\boldsymbol{x}^0$, that separates $\boldsymbol{x}^\star$ from $\mathcal{F}_{\text{SPP}}$? If it exists, it will be transferred to CP to tighten the relaxation of $\bar{\Delta}^{int}_{\mathcal{I}}$. From the theoretical point of view, if $\boldsymbol{d}^\star$ is extremal, such a cut always exists as shown on Corollary 2.

Corollary 2 *Assume $\boldsymbol{d}^\star_{\mathcal{I}} \in Ext\,(\mathcal{F}_{\text{CP}})$, and let $\boldsymbol{d}^\star = \boldsymbol{T}\boldsymbol{d}^\star_{\mathcal{I}}$ (we still assume that $\boldsymbol{d}^\star_{\mathcal{I}}$ is not an integral direction). There always exists a valid inequality (Γ) tight at $\boldsymbol{x}^0$ that separates $\boldsymbol{x}^\star = \boldsymbol{x}^0 + r(\boldsymbol{d}^\star)\boldsymbol{d}^\star$ from $\mathcal{F}_{\text{SPP}}$.*

Proof $\boldsymbol{d}^\star_{\mathcal{I}}$ is an extreme point of $\mathcal{F}_{\text{CP}}$ but does not belong to Δ^{int}. Since Δ^{int} is a finite subset of elements of $\mathcal{F}_{\text{CP}}$, and since $\mathcal{F}_{\text{CP}}$ is a polyhedron, there exists a cut $\left(\bar{\Gamma}\right)$ that separates $\boldsymbol{d}^\star_{\mathcal{I}}$ from $\bar{\Delta}^{int}_{\mathcal{I}}$. By Proposition 8, the result holds. □

It is also interesting to note that the linear transformation applied to the primal cut to transfer it to CP ($\bar{\boldsymbol{\alpha}}^T = \boldsymbol{\alpha}^T\boldsymbol{T}$) is the same as for the objective function ($\bar{\boldsymbol{c}}^T = \boldsymbol{c}^T\boldsymbol{T}$) and for the constraint matrix ($\bar{\boldsymbol{A}}_{\bar{\mathcal{P}}\mathcal{I}} = \boldsymbol{A}_{\bar{\mathcal{P}}\cdot}\boldsymbol{T}$). Finally, note that adding cuts does not prevent to use the characterization of extreme integer solutions as those having a column-disjoint support. Since no cutting plane is added to CP that cuts any integer directions, the set of extreme integral directions $Ext\left(Conv\left(\bar{\Delta}^{int}_{\mathcal{I}}\right)\right)$ is still contained in the set of extreme directions $Ext\,(\mathcal{F}_{\text{CP}})$ even after the addition of cutting planes. Geometrically, adding any valid inequality for $Conv\left(\bar{\Delta}^{int}_{\mathcal{I}}\right)$ to the relaxation cannot transform any nonextreme integral direction into a extreme one.

4.1 Specific primal separation procedures

As exposed in the introduction, there exist extensive work on the primal separation side. It has been shown that **P-SEP** is equivalent to SPP in terms of complexity [28]. Also, there exist general-purpose families of primal cuts, such as Gomory–Young's cuts [35]. Iteratively adding cuts from these families would ultimately lead to a column-disjoint solution of CP, after a finite number of separations – provided that the choice of incoming variables follow some lexicographic order (see [35]). However, these families are not polyhedral, i.e., they do not take into account the SPP structure. We chose to study two families of inequalities that are known to yield strong cutting planes for SPP, namely clique and odd-cycle inequalities.

4.1.1 Primal clique inequalities

Let us consider the conflict graph of matrix $\boldsymbol{A}$, $\mathcal{G} = (\mathcal{N}, E)$, where each node of $\mathcal{N} = \{1, \ldots, n\}$ corresponds to a column of $\boldsymbol{A}$, and E is such that $\{i, j\} \in E$ if and only if $\boldsymbol{A}^T_{\cdot i}\boldsymbol{A}_{\cdot j} \neq 0$. Two vertices are linked by an edge if the corresponding columns are not disjoint. Given a clique $\mathcal{W}$ in this graph, any feasible solution of SPP satisfies

$$(\Gamma_{\mathcal{W}}) \;:\; \sum_{i \in \mathcal{W}} x_i \leq 1, \tag{10}$$

called the clique inequality associated with $\mathcal{W}$. clique cuts form a family of generally strong cutting planes for SPP, and were first introduced by Padberg [20]. Given a fractional solution $\boldsymbol{x}^\star$ of SPP^{LR}, the standard clique separation problem consists in associating the weight $w_i = x_i^\star$ with each vertex of $\mathcal{G}$ and determine a clique of weight greater than 1 in $\mathcal{G}$ if any.

From the work of Letchford and Lodi [16], we develop here a more efficient procedure for primal clique cuts. In the primal context, let $\boldsymbol{d}^\star$ be a fractional direction, and $\boldsymbol{x}^\star = \boldsymbol{x}^0 + r(\boldsymbol{d}^\star)\boldsymbol{d}^\star$. For $(\Gamma_{\mathcal{W}})$ to be tight at $\boldsymbol{x}^0$, we need $\sum_{i\in\mathcal{W}} x_i^0 = |\mathcal{W} \cap \mathcal{P}| = 1$. Hence, exactly one variable from $\mathcal{P}$ must be part of clique $\mathcal{W}$. Denote that variable as l, and the corresponding clique as $\mathcal{W}_l$. Furthermore, for $\left(\Gamma_{\mathcal{W}_l}\right)$ to separate $\boldsymbol{x}^\star$ from $\mathcal{F}_{\text{SPP}}$, we must have $\sum_{i\in\mathcal{W}_l} x_i^\star > 1$. Since $x_i^\star = r(\boldsymbol{d}^\star)d_i^\star$ for all $i \in \mathcal{I}$, $\sum_{i\in\mathcal{W}_l} x_i^\star = x_l^\star + \sum_{i\in\mathcal{S}^+} x_i^\star$, where $\mathcal{S}^+ = \mathit{Supp}\left(\boldsymbol{d}_{\mathcal{I}}^\star\right)$. Denote also as $\mathcal{S}^- = \mathit{Supp}\left(\boldsymbol{d}_{\mathcal{P}}^\star\right)$ the set of columns of $\mathcal{P}$ that are impacted by direction $\boldsymbol{d}^\star$. Sets $(\mathcal{S}^-, \mathcal{S}^+)$ form a partition of $\mathit{Supp}\left(\boldsymbol{d}^\star\right)$ such that for all $i \in \mathit{Supp}\left(\boldsymbol{d}^\star\right)$, $i \in \mathcal{S}^+$ if $d_i^\star > 0$ and $i \in \mathcal{S}^-$ if $d_i^\star < 0$, so

$$\sum_{i\in\mathcal{W}_l} x_i^\star > 1 \Leftrightarrow x_l^\star + \sum_{i\in\mathcal{W}_l\cap\mathcal{S}^+} x_i^\star > 1.$$

Because $\mathcal{W}_l$ is a clique in $\mathcal{G}$, then $\mathcal{W}_l \cap \mathcal{S}^+ \subseteq \mathcal{W}_l$ is also a clique. The consequence of these observations is summarized in the following proposition, which shows that it is sufficient to look for a primal clique cut within $\mathcal{S}^+ \cup \mathcal{S}^-$ to solve the complete primal clique separation problem.

Proposition 9 *There exists a primal clique inequality that separates $\boldsymbol{x}^\star$ from $\mathcal{F}_{\text{SPP}}$ if and only if for some $l \in \mathcal{S}^-$, there exists a clique $\mathcal{W}_l$ that satisfies*

$$(1)\ \mathcal{W}_l \setminus \{l\} \subseteq N_l,\ \ (2)\ l \in \mathcal{W}_l,\ \ \textit{and}\ (3)\ w(\mathcal{W}_l) > 1,$$

where N_l is the set of neighbors of l in $\mathcal{G}$ that are also in $\mathcal{S}^+$ and $w(\mathcal{W}_l) = \sum_{i\in\mathcal{W}_l} x_i^\star$.

Hence, the primal separation procedure for primal clique cuts, called `CL_PSEP` and summarized in Algorithm 2, returns an empty set of primal clique cuts if and only if none exists, i.e., this separation procedure is exact. This algorithm always finds a cut if there exists one, and that will be the most violated. However, the size of the clique could be potentially increased by adding 0-weight variables. Even if Step 4 of Algorithm 2 requires solving a $\mathcal{NP}$-hard problem,[4] note that the procedure is practically very fast because the size of N_l is usually small.

[4] For the determination of the clique of maximal weight in $\mathcal{G}_l$, we use the *cliquer* open source library, available at http://users.aalto.fi/~pat/cliquer.html, based on the algorithm described in [19].

Algorithm 2: CL_PSEP

Input: $\boldsymbol{d}^\star_{\mathcal{I}} \leftarrow$ an optimal solution of CP (fractional direction).
Output: $\mathcal{K}$, a set of primal clique cuts (empty if none exists)

$\mathcal{K} \leftarrow \emptyset$; $\boldsymbol{d}^\star \leftarrow \boldsymbol{T}\boldsymbol{d}^\star_{\mathcal{I}}$; $\mathcal{S}^- \leftarrow Supp\left(\boldsymbol{d}^\star_{\mathcal{P}}\right)$; $\mathcal{S}^+ \leftarrow Supp\left(\boldsymbol{d}^\star_{\mathcal{I}}\right)$;
for $l \in \mathcal{S}^-$ **do**
  $\mathcal{G}_l = (N_l, E_l) \leftarrow$ weighted subgraph of $\mathcal{G}$ induced by $N_l \subseteq \mathcal{S}^+$ ($w_i = x^\star_i$);
  $\mathcal{W}_l \leftarrow$ clique of maximum weight in $\mathcal{G}_l$;
  $\mathcal{W}_l \leftarrow \mathcal{W}_l \cup \{l\}$;
  if $w(\mathcal{W}_l) > 1$ **then**
    $\mathcal{K} \leftarrow \mathcal{K} \cup \{\left(\Gamma_{\mathcal{W}_l}\right)\}$;
return $\mathcal{K}$;

4.1.2 Primal odd-cycle inequalities

odd-cycle inequalities form another well-known family of valid inequalities for SPP. Given a cycle $\mathcal{Q}$ of odd length in $\mathcal{G}$, the following inequality

$$\left(\Gamma_{\mathcal{Q}}\right) \; : \; \sum_{i \in \mathcal{Q}} x_i \leq \frac{|Q| - 1}{2} \tag{11}$$

is valid for SPP. Clearly, $\left(\Gamma_{\mathcal{Q}}\right)$ is tight at $\boldsymbol{x}^0$ if and only if $\sum_{i \in \mathcal{Q}} x^0_i = |\mathcal{Q} \cap \mathcal{P}| = (|\mathcal{Q}| - 1)/2$. Furthermore, since $\mathcal{P}$ is column-disjoint, there exists no edge between any pair of vertices of $\mathcal{P}$. Therefore, $\mathcal{Q}$ is an alternate cycle with vertices in $\mathcal{P}$ and $\mathcal{I}$, except for one $\mathcal{I} - \mathcal{I}$ edge (i.e., an edge that links two vertices of $\mathcal{I}$). Based on these observations and similarly to clique cuts, the search graph can be restricted to vertices in $Supp\left(\boldsymbol{d}^\star\right)$, but in a stronger manner.

Proposition 10 *Every primal odd-cycle cut* $\left(\Gamma_{\mathcal{Q}}\right)$ *that separates* $\boldsymbol{x}^\star$ *from* $\mathcal{F}_{\text{SPP}}$ *satisfies* $\mathcal{Q} \subseteq Supp\left(\boldsymbol{d}^\star\right)$.

Proof Suppose that the result is false and that there exists a cycle $\mathcal{Q} \nsubseteq Supp\left(\boldsymbol{d}^\star\right)$. Let $\mathcal{S}^- = Supp\left(\boldsymbol{d}_{\mathcal{P}}\right)$ and $\mathcal{S}^+ = Supp\left(\boldsymbol{d}_{\mathcal{I}}\right)$. Define $\mathcal{Q} = (q_1, q_2, \ldots, q_{2T+1}, q_1)$, where $q_1, q_{2t+1} \in \mathcal{I}$ and $q_{2t} \in \mathcal{P}$ for all $t \in \{1, \ldots, T\}$. $\mathcal{Q}$ is an alternate cycle but for the $\{q_{2T+1}, q_1\}$ edge. Consider the two following cases:

(i) $\underline{q_1 \notin \mathcal{S}^+}$. Then $d^\star_{q_1} = 0$ and

$$\sum_{i \in \mathcal{Q}} x^\star_i = \sum_{i=2}^{2T+1} x^\star_{q_i} = \sum_{t=1}^{T} \left(x^\star_{q_{2t}} + x^\star_{q_{2t+1}}\right).$$

Every (q_{2t}, q_{2t+1}) is an edge of $\mathcal{G}$, so for every pair of columns $\left(\boldsymbol{A}_{\cdot 2t}, \boldsymbol{A}_{\cdot 2t+1}\right)$ there exists a row i such that $A_{i(2t)} = A_{i(2t+1)} = 1$. The linear set partitioning constraint that corresponds to that conflict yields $x^\star_{q_{2t}} + x^\star_{q_{2t+1}} \leq 1$ for all $t \in \{1, \ldots T\}$. Hence, $\sum_{i \in \mathcal{Q}} x^\star_i \leq T = (|Q| - 1)/2$ and $\boldsymbol{x}^\star$ does not violate $\left(\Gamma_{\mathcal{Q}}\right)$.

(ii) $q_2 \notin \mathcal{S}^-$. From $\boldsymbol{A}\boldsymbol{d}^\star = \boldsymbol{0}$ and $\boldsymbol{d}^\star_{\mathcal{I}} \geq \boldsymbol{0}$, necessarily $q_1, q_3 \notin \mathcal{S}^+$. Case (i) applies and this concludes the proof. □

Proposition 10 suggests to distinguish between $\mathcal{S}^- - \mathcal{S}^+$ and $\mathcal{S}^+ - \mathcal{S}^+$ edges. Let $\mathcal{G}_B = (N_B, E_B)$ be a subgraph of $\mathcal{G}$ with $N_B = \mathcal{S}^- \cup \mathcal{S}^+ = Supp\left(\boldsymbol{d}^\star\right)$, and $\{i, j\} \in E_B$ if and only if $i \in \mathcal{S}^-$, $j \in \mathcal{S}^+$, and $\{i, j\} \in E$ (in conflict graph $\mathcal{G}$). $\mathcal{G}_B$ is a bipartite graph. In the case of the primal odd-cycle separation, the edges (and not the vertices) of the graph are weighted as $w_{ij} = 1 - x^\star_i - x^\star_j$ if $\{i, j\} \in E_B$. The weight of a path or a cycle is the sum of the weights of its edges. Given a cycle $\mathcal{Q}$, its weight is therefore $w(\mathcal{Q}) = |\mathcal{Q}| - 2\sum_{i\in\mathcal{Q}} x^\star_i$ and the corresponding odd-cycle inequality is violated by $\boldsymbol{x}^\star$ if and only if $w(\mathcal{Q}) < 1$. The primal odd-cycle separation procedure, `CY_PSEP` is given in Algorithm 3. This algorithm is usually fast because it consists of finding at most $|\mathcal{S}^-|$ shortest paths in a relatively small bipartite graph $\mathcal{G}_B$.

Algorithm 3: `CY_PSEP`

Input: $\boldsymbol{d}^\star_{\mathcal{I}} \leftarrow$ an optimal solution of CP (fractional direction).
Output: $\mathcal{K}$, a set of primal odd-cycle cuts (empty if none exists)

1 $\mathcal{K} \leftarrow \emptyset$; $\boldsymbol{d}^\star \leftarrow \boldsymbol{T}\boldsymbol{d}^\star_{\mathcal{I}}$; $\mathcal{S}^- \leftarrow Supp\left(\boldsymbol{d}^\star_{\mathcal{P}}\right)$; $\mathcal{S}^+ \leftarrow Supp\left(\boldsymbol{d}^\star_{\mathcal{I}}\right)$; $\boldsymbol{x}^\star \leftarrow \boldsymbol{x}^0 + r(\boldsymbol{d}^\star)\boldsymbol{d}^\star$;
2 Build weighted graph $\mathcal{G}_B$ ($w_{ij} = 1 - x^\star_i - x^\star_j$);
3 **for** $i, j \in \mathcal{S}^+$ *such that* $\{i, j\} \in E$ ($\mathcal{S}^+$–$\mathcal{S}^+$ *conflict*) **do**
4 $\quad$ $(q_1, q_2, \ldots, q_{2T}, q_{2T+1}) \leftarrow$ shortest path from i to j in $\mathcal{G}_B$ ($i = q_1$, $j = q_{2T+1}$);
5 $\quad$ $\mathcal{Q} \leftarrow (q_1, q_2, \ldots, q_{2T}, q_{2T+1}, q_1)$ (odd-cycle);
6 $\quad$ **if** $w(\mathcal{Q}) = w_{ij} + \sum_{k=1}^{2T+1} w_{q_k q_{k+1}} < 1$ **then**
7 $\quad\quad$ $\mathcal{K} \leftarrow \mathcal{K} \cup \{\left(\Gamma_{\mathcal{Q}}\right)\}$;
8 **return** $\mathcal{K}$;

4.2 Algorithm

Algorithm 4 incorporates the cutting plane results discussed in the previous sections into Algorithm 1. The different (nonexhaustive) branching strategies that we use are presented in Sect. 5.

SEP_MAX is a parameter that represents the maximum number of separation problems to be solved consecutively. Explicit values chosen for INC_MAX and SEP_MAX, as well as the methods to determine the set of variables to be fixed (line 33) are discussed in Sect. 5 before the numerical results are given. Different priority rules for choosing the separation algorithm (line 28) are studied and the corresponding results are displayed in the next section.

5 Numerical results

The results presented in this section empirically show the relevance of our theoretical results. We show that primal cuts indeed improve the performance of our algorithm and help foster integral solutions in ISUD.

Algorithm 4: ISUD_CUTS

Input: $\boldsymbol{x}^0$: a solution of SPP; INC_MAX: maximal incompatibility degree considered; SEP_MAX: maximal number of consecutive **P-SEP**;
Output: $\boldsymbol{x}^k$, a better solution of SPP.

$c \leftarrow 0$ (number of consecutive **P-SEP** solved);
$\iota \leftarrow 1$ (current maximum incompatibility degree);
Pool $\leftarrow \emptyset$ (pool of valid inequalities);
Compute $\mathcal{P}$ and $\mathcal{C}$ associated with $\boldsymbol{x}^0$; $k \leftarrow 0$;
while *true* **do**
  If necessary, update $\mathcal{P}$, $\mathcal{C}$, and $\mathcal{I}^\iota$ associated with $\boldsymbol{x}^k$;
  if $z^\star_{\mathrm{RP}} < 0$ **then**
    $\boldsymbol{\delta}^i \leftarrow$ an optimal solution of RP ($\bar{c}_i < 0, i \in \mathcal{C}^k$);
    $\boldsymbol{x}^{k+1} \leftarrow \boldsymbol{x}^k + \boldsymbol{\delta}^i$; $k \leftarrow k+1$;
  else
    Build CP$^\iota$ anew (without cutting planes nor fixed variables);
    continue $\leftarrow$ true;
    while *continue* = *true* **do**
      If necessary, update $\mathcal{P}$, $\mathcal{C}$, and $\mathcal{I}^\iota$ associated with $\boldsymbol{x}^k$, and CP$^\iota$;
      if $z^\star_{\mathrm{CP}^\iota} < 0$ **then**
        $\boldsymbol{d}^\star_{\mathcal{I}} \leftarrow$ an optimal solution of CP$^\iota$; $\boldsymbol{d}^\star \leftarrow \boldsymbol{T}\boldsymbol{d}^\star_{\mathcal{I}}$;
        if $\boldsymbol{d}^\star_{\mathcal{I}}$ *is column-disjoint* **then**
          $\boldsymbol{x}^{k+1} \leftarrow \boldsymbol{x}^k + r(\boldsymbol{d}^\star)\boldsymbol{d}^\star$;
          $k \leftarrow k+1$;
          continue $\leftarrow$ false;
        else
          $\boldsymbol{x}^\star \leftarrow \boldsymbol{x}^0 + r(\boldsymbol{d}^\star)\boldsymbol{d}^\star$;
          if $c <$ SEP_MAX **then**
            $c \leftarrow c+1$;
            if Pool *contains primal cuts that separate* $\boldsymbol{x}^\star$ *from* $\mathcal{F}_{\mathrm{SPP}}$ **then**
              Transfer all these cuts to CP$^\iota$;
            else
              $\mathcal{K} \leftarrow$ set of primal cuts generated with CL_PSEP and/or CY_PSEP;
              if $\mathcal{K} \neq \emptyset$ **then**
                Transfer every cut in $\mathcal{K}$ to CP$^\iota$;
                Pool $\leftarrow$ Pool $\cup\, \mathcal{K}$;
          else
            Fix at least one variable of CP$^\iota$ to zero;
            $c \leftarrow 0$;
      else
        if $\iota <$ INC_MAX **then**
          $\iota \leftarrow \iota + 1$;
          $c \leftarrow 0$;
        else
          return $\boldsymbol{x}^k$;

5.1 Methodology

5.1.1 Instances

The instances presented here are those used by Zaghrouti et al. [37] (SPPAA01, VCS1200, and VCS1600) to which we added another instance from the OR-Library,[5] SPPAA04. Both SPPAA01 and SPPAA04 are small flight assignment problems (respec-

[5] http://people.brunel.ac.uk/~mastjjb/jeb/info.html.

tively 823 constraints and 8904 variables, and 426 constraints and 7195 constraints) with an average of 9 nonzero entries per column. VCS1200 is a medium-size bus driver scheduling problem (1200 constraints, 133,000 variables), and VCS1600 is a large bus driver scheduling problem (1600 constraints, 571,000 variables); both have an average of 40 nonzero entries per column. These numbers of nonzero entries per column are typical of aircrew and bus driver scheduling problems and do not vary much with the number of constraints. In scheduling instances, they typically correspond to the number of tasks to be completed by a worker in a given time span. The optimal solutions of the problems are known: SPPAA01 and SPPAA04 were solved with CPLEX,[6] and VCS1200 and VCS1600 were made with all columns generated during a column-generation process by GENCOL[7] whose optimal solution is known.

5.1.2 Initial solutions

Generating meaningful initial solutions is crucial for accurately testing primal algorithms. For this reason, we took here an "application-oriented" viewpoint and we generated initial solutions (from known optimal ones) that accurately mimic the ones available in practice, especially in transportation scheduling problems, i.e., where SPP is widely used. More precisely, in industrial crew scheduling instances, one can observe the following:

(1) In the case of beforehand planning, one can use the vehicle (aircraft/bus) routes as initial crew schedules. Each column of this solution corresponds to the set of legs (flight or bus trips) covered by that vehicle. If some of these schedules are not feasible for the pilots/drivers (they may not respect the collective agreement), the corresponding columns are given prohibitive big-M costs.
(2) In the case of reoptimization, one can use the existing schedule as initial solution. If some schedules are not feasible anymore in the new situation, they are given prohibitive big-M costs.

Suppose now that the rows of $\boldsymbol{A}$ are given in chronological order of the corresponding tasks, i.e., for all $i, j \in \mathcal{R}$, $i < j$ if and only if task i starts before task j. Given a solution $\boldsymbol{x}$ and the corresponding indices of the nonzero variables $\mathcal{P} = Supp\ (x)$, a pair $(i, j) \in \mathcal{R}^2$ of tasks is called *consecutive* in $\boldsymbol{x}$ if and only if there exists $k \in \mathcal{P}$ such that: (i) $A_{ik} = A_{jk} = 1$ and (ii) $\forall j' \in \{i+1, \ldots, j-1\}$, $A_{j'k} = 0$. Consecutive tasks are therefore all the pairs of tasks performed consecutively by the same employee (see Example 1).

The main observation here is the following: the initial solutions described above share a high proportion of consecutive tasks with the optimal solution. Quantitatively, this is described by the *primal information* contained in the initial solution $\boldsymbol{x}^0$ that we define as the percentage of consecutive tasks of $\boldsymbol{x}^0$ that are also consecutive tasks in the optimal solution $\boldsymbol{x}^{\star}_{\text{SPP}}$. In bus driver scheduling problems, the initial solution

[6] CPLEX is freely available for academic and research purposes under the IBM academic initiative: http://www-03.ibm.com/ibm/university/academic. When referring to CPLEX, we always refer to the version 12.4 of this software, with single-thread settings (all other settings being default).

[7] GENCOL is a commercial software developed at the GERAD research center and now owned by the AD OPT company, a division of KRONOS.

that follows the bus routes typically contains 90% of primal information (VCS1200, VCS1600). In aircrew scheduling, the figure is generally around 75% [37] (SPPAA01, SPPAA04). These percentages are even higher in the reoptimization case. Example 1 shows how primal information is computed on an 11-task case.

Example 1 Consider an 11-task scheduling problem. The tasks (rows) are in chronological order and denoted as A, …, K. The individual schedules used in the optimal and initial solutions are respectively those of $\mathcal{P}^\star$ and $\mathcal{P}^0$ and are described here:

$\mathcal{P}^\star$			Task	$\mathcal{P}^0$		
1	0	0	A	1	0	0
0	1	0	B	0	1	0
0	1	0	C	0	1	0
1	0	0	D	1	0	0
1	0	0	E	1	0	0
1	0	0	F	0	1	0
1	0	0	G	0	1	0
0	1	0	H	0	0	1
0	0	1	I	1	0	0
0	1	0	J	0	0	1
0	0	1	K	1	0	0

The set of consecutive tasks in the optimal solution is $\mathcal{CT}^\star = \{(A, D), (D, E), (E, F), (F, G), (B, C), (C, H), (H, J), (I, K)\}$. That of the initial solution is: $\mathcal{CT}^0 = \{(A, D), (D, E), (E, I), (I, K), (B, C), (C, F), (F, G), (H, J)\}$. Therefore, the primal information contained in $\boldsymbol{x}^0$ is $|\mathcal{CT}^\star \cap \mathcal{CT}^0|/|\mathcal{CT}^0| = 5/8 = 62.5\%$.

We chose to perturb the (known) optimal solutions to generate initial solutions that contain a similar level of primal information to that experienced in practice. The details of the perturbation method can be found in [37]. These solutions are generally infeasible, so we assign the perturbed columns a high cost, as is done in companies for the schedules that follow the bus/airplane routes. In our perturbation method, the input parameter π is the percentage of columns of the optimal solution that will appear in the new initial solution. For SPPAA04 and SPPAA01, we generated initial solutions for π = 10, 15, 20, and 35%; for VCS1200 and VCS1600, we used π = 20, 35, and 50%. These parameters were chosen so that the resulting primal information (given in Table 1) is consistent with the typical values. The initial gaps range from 50 to 80%, depending on the instance. Here (and in the rest of the paper), the gap is the ratio of the difference between the current solution value and the optimal solution value divided by the latter.

Remark 5 The primal information is related to a form of consecutive-ones property in the SPP. Suppose here that the rows are re-sorted so that each column of the initial solution has consecutive ones *and* that the corresponding nonzero rows are sorted in chronological order of the tasks. The primal information is the measure of how close the optimal solution is to the consecutive-ones property with that particular ordering, namely, it is the percentage of consecutive ones of the initial solution that are also

Table 1 Percentage of primal information in the initial solutions of the benchmark

Instance	m	n	Primal information				
			$\pi = 10\%$	$\pi = 15\%$	$\pi = 20\%$	$\pi = 35\%$	$\pi = 50\%$
SPPAA01	803	8904	71.5	75.6	80.0	–	–
SPPAA04	423	7195	–	64.0	70.1	78.5	–
VCS1200	1200	133,000	–	–	87.6	91.1	93.9
VCS1600	1600	570,000	–	–	86.8	90.9	93.9

consecutive in the optimal solution. With that order, given the typical figures of primal information in practical problems (over 75%), the columns of the optimal solution are thus made of few blocks of consecutive ones ("quasi"-consecutive-ones). From the algorithmic point of view, this characteristic of the problem is only superficial in the sense that it depends on its presentation (ordering of the rows) and *not* on its intrinsic structure. Furthermore, with a pure chronological ordering (the one that we use), neither the initial nor the optimal solution have the consecutive-ones property. More interesting is the computation of the incompatibility degree (see, Sect. 3.1.2): it is a dynamic measure of the number of breaks of consecutive-ones in the nonbasic columns, based on the ordering in which the *current* solution has the consecutive-ones property with chronological order of the tasks. Rather than sticking to a static ordering of the rows corresponding to the initial solution when measuring the incompatibility degree, we compute it in terms of the current solution.

5.1.3 Cutting planes strategies

The tests were conducted for the following cutting planes strategies:

NONE: Without primal cuts;
CLIQUE: With primal clique cuts only;
CYCLE: With primal odd-cycle cuts only;
BOTH: Both aforementioned cut types are separated at every **P-SEP** step;
PRIO: Primal odd-cycle cuts are separated only if no primal clique cut is found.

Moreover, another parameter is included, which is the maximum number of separation problems solved before using another technique (such as branching). We analyzed the results for different values ranging from 40 to 40,000 (virtually infinite). In a very large majority of the cases, either no cut could be found before the 40th **P-SEP** was solved, or the cutting planes yield an integral direction in less than 40 **P-SEP**. Hence, we fixed the maximum number of consecutive **P-SEP** to 40 and only present these results here.

5.1.4 Branching strategies

We propose four different nonexhaustive branching techniques to improve the performance of the algorithm. They correspond to the decision made when no primal cut is found that cuts the current fractional direction.

NOBR: stop the algorithm;
LAST: all variables of the last fractional direction found are set to zero in CP until an augmentation is performed;
FIRST: all variables of the first direction found since the last augmentation or the last branching are set to zero in CP until an augmentation is performed;
COVER: a subproblem is solved to determine a small subset of $\mathcal{I}$ such that the support of each fractional solution found since the last augmentation or the last branching contains at least one element of this set. These variables are set to zero in CP, so that all the aforementioned fractional solutions become infeasible.

All these techniques were experimented, but we present detailed results for NOBR and FIRST only. The performance for the other two branching strategies are exposed more briefly. The variation in the results is not significant enough to make a detailed presentation of all four aforementioned strategies.

5.2 Results

All the tests were performed on a Linux PC with a processor of 3.4 GHz. For each problem and each perturbation parameter π, 10 different instances (different initial solutions and corresponding artificial columns) are generated. Column ALGO indicates the cutting strategy used; BEST is the best of the four on each instance, i.e., the one with the smallest gap, and in the event of a tie, the shortest computational time to reach the best solution. All times are given in seconds.

The commercial solver CPLEX was tested on all the instances of the benchmark, with the initial solutions given as MIP-start. Instances SPPAA01 and SPPAA04 are solved to optimality in an average of 22.1 and 14.4 s, respectively. Within a time limit of 1 h, provided the initial solutions as a warm start, CPLEX only slightly improves them on the instance VCS1200, never lowering the gap under 14% (average gap of 49.7%). Within that same time limit, it never improves any of the initial solutions given for VCS1600 at all (the average gap remains 73% as for the initial solutions). Note that, if the initial solution is not given, CPLEX only finds a feasible solution for 1 of 30 instances for both VCS1200 and VCS1600 within 1 h, and that solution is of comparable cost to that of our initial solutions. Here, one should consider SPPAA01 and SPPAA04 as benchmark instances used for a detailed analysis of our algorithm and its behavior, and VCS1200 and VCS1600 as practical instances that we aim to solve within a few minutes.

Tables 2 and 3 show the results of our algorithm for all instances generated from SPPAA01, with associated branching strategies NOBR and LAST, for parameters INC_MAX $= 10$, SEP_MAX $= 40$. Other values were tested and gave very similar results. Columns π and ALGO display the perturbation degree of the instances and the chosen separation strategy, respectively. The next three columns display the number of instances (out of 10) that were solved to optimality (0%), with a positive gap $\leq 2\%$, and with a gap $> 2\%$. The number of instances where the best solution found by ISUD still contains artificial columns from the initial solution (with big-M

Table 2 Comparison of the performances of ISUD with branching NOBR for instance SPPAA01

π (%)	ALGO	GAP				Time (s)			AUG		
		0%	≤2%	>2%	MEAN (%)	t_{ISUD}	t_{BEST}	$t_{\mathbf{AUG}}$	K	$\lvert\mathcal{S}\rvert^{\text{D}}$	$\lvert\mathcal{S}\rvert^{\text{N}}$
10	NONE	2	2	6 [6]	0.4	6.2	5.1	0.21	29	4.5	101
	CLIQUE	3	5	2 [2]	0.5	12.5	8.6	0.28	33	4.9	122
	CYCLE	3	2	5 [5]	0.3	8.0	5.6	0.22	31	4.6	143
	BOTH	3	3	4 [4]	0.5	15.9	7.4	0.25	35	4.6	142
	PRIO	3	5	2 [2]	0.5	16.1	9.3	0.30	33	4.9	113
	BEST	**3**	**5**	**2 [2]**	**0.5**	**11.7**	**7.9**	–	–	–	–
15	NONE	2	3	5 [5]	0.4	6.7	5.3	0.17	37	3.7	48
	CLIQUE	5	3	2 [2]	0.2	10.4	6.7	0.18	39	3.9	70
	CYCLE	3	4	3 [3]	0.3	8.0	5.9	0.18	37	3.8	70
	BOTH	5	4	1 [1]	0.2	15.7	7.3	0.20	39	4.0	95
	PRIO	5	4	1 [1]	0.2	15.6	7.4	0.20	39	4.0	98
	BEST	**5**	**4**	**1 [1]**	**0.2**	**9.2**	**6.7**	–	–	–	–
20	NONE	7	2	1 [1]	0.0	7.6	6.0	0.17	37	3.4	32
	CLIQUE	5	5	0 [0]	0.2	11.0	6.9	0.17	40	3.4	57
	CYCLE	6	3	1 [1]	0.0	7.7	6.0	0.16	39	3.4	43
	BOTH	6	4	0 [0]	0.2	15.0	7.0	0.18	40	3.4	76
	PRIO	6	4	0 [0]	0.2	13.3	6.9	0.17	40	3.4	74
	BEST	**9**	**1**	**0 [0]**	**0.0**	**10.4**	**6.3**	–	–	–	–

Parameters: INC_MAX = 10 and SEP_MAX = 40

cost) is indicated between brackets next to those with a gap $> 2\%$ (because of the big-M, no solution with a gap lower or equal to 2% contains any of these columns). Under label MEAN is the mean gap computed over instances where solutions contain no artificial column only. Then, the overall computation time (t_{ISUD}), the time to reach the best solution obtained (t_{BEST}) and the average time per augmentation ($t_{\mathbf{AUG}}$) are displayed. The last columns contain the mean number of augmentations K performed to reach the best solution, and the mean size of $|\mathcal{S}| = \mathit{Supp}\ (\boldsymbol{d})$ for disjoint ($|\mathcal{S}|^{\text{D}}$) and nondisjoint ($|\mathcal{S}|^{\text{N}}$) directions. Note that, while the other mean values are computed over all executions of the algorithm, the mean number of augmentations K is based on the instances for which the final gap is $\leq 2\%$. Large gaps often mean a much smaller number of iterations, and taking these instances into account would distort that mean. Tables 4 and 5 show the results of the same experiments for instances generated from SPPAA04.

Consistently with our expectations and whatever the cutting planes or branching strategy, the primal information strongly influences the results of the algorithm. Namely, the higher the percentage of primal information is, the better the algorithm performs, both in terms of quality of the solution, and average running time. The branching strategies highlight common features and global differences between the various cutting techniques. First, the performance of the algorithms can be globally

Table 3 Comparison of the performances of ISUD with branching LAST for instance SPPAA01

π (%)	ALGO	GAP				Time (s)			AUG		
		0%	≤2%	>2%	MEAN (%)	t_{ISUD}	t_{BEST}	$t_{\mathbf{AUG}}$	K	$\lvert\mathcal{S}\rvert^{\text{D}}$	$\lvert\mathcal{S}\rvert^{\text{N}}$
10	NONE	3	3	4 [4]	0.3	11.4	8.7	0.27	38	4.4	112
	CLIQUE	3	3	4 [3]	0.3	52.6	33.0	0.92	39	4.7	198
	CYCLE	3	2	5 [5]	0.4	18.0	12.0	0.35	37	4.3	151
	BOTH	4	1	5 [4]	0.4	76.2	32.2	0.91	40	4.6	208
	PRIO	5	1	4 [3]	0.3	54.6	24.1	0.68	39	4.7	177
	BEST	**7**	**1**	**2 [2]**	**0.2**	**33.2**	**19.4**	–	–	–	–
15	NONE	2	4	4 [4]	0.1	14.9	12.2	0.30	44	4.0	87
	CLIQUE	7	2	1 [1]	0.0	28.2	9.6	0.24	41	4.0	142
	CYCLE	5	3	2 [2]	0.2	16.0	10.6	0.27	41	3.9	129
	BOTH	8	2	0 [0]	0.0	30.7	16.3	0.39	42	4.1	111
	PRIO	9	1	0 [0]	0.0	34.6	17.6	0.43	41	4.1	120
	BEST	**9**	**1**	**0 [0]**	**0.0**	**25.0**	**14.8**	–	–	–	–
20	NONE	7	2	1 [1]	0.1	9.9	7.0	0.17	42	3.3	67
	CLIQUE	6	4	0 [0]	0.0	22.2	9.2	0.21	43	3.5	104
	CYCLE	8	2	0 [0]	0.1	12.7	8.5	0.20	42	3.4	86
	BOTH	9	1	0 [0]	0.0	23.6	9.8	0.22	45	3.6	91
	PRIO	9	1	0 [0]	0.0	24.4	9.7	0.21	45	3.6	102
	BEST	**9**	**1**	**0 [0]**	**0.0**	**16.1**	**9.1**	–	–	–	–

Parameters: INC_MAX = 10 and SEP_MAX = 40

ranked from worst to best as follows: NONE $\preceq$ CYCLE $\preceq$ CLIQUE $\preceq$ PRIO $\preceq$ BOTH. Moreover, the execution time is significantly shorter for NONE and CYCLE than for the others. In the case of NONE, no primal separation problem is solved, and either no branching, or very simple fixing rules are applied, hence the high speed of the algorithm. In the case of CYCLE, the number of primal cycle cuts found is quite small, so the computation time is also smaller. The same holds for the time spent to reach the best solution (t_{BEST}) and the time per augmentation ($t_{\mathbf{AUG}}$). Note that the difference is much bigger for t_{ISUD} than for t_{BEST} because the time spent at the optimal solution to generate cutting planes is important. The number of augmentation steps is higher for the algorithms that generate more primal cuts (column K). Generally, cutting planes allow the algorithm to find more disjoint solutions within the same phase, hence yielding a larger number of steps for each incompatibility degree. The algorithms using cuts tend to generate disjoint and nondisjoint combinations of larger size (columns $\lvert\mathcal{S}\rvert^{\text{D}}$ and $\lvert\mathcal{S}\rvert^{\text{N}}$). Cutting planes tend to make the problem more complex, add constraints, and the size of the support of a basic solution of CP therefore increases with the number of cuts inserted, hence the larger size of the combinations.

Furthermore, the algorithm is significantly better if fixing variables is available. This is particularly true when no cut is applied. Both nonexhaustive branching and cutting planes improve the performance of the algorithm. If we analyze the performance over

Table 4 Comparison of the performances of ISUD with branching NOBR for instance SPPAA04

π (%)	ALGO	GAP				Time (s)			AUG		
		0%	≤2%	>2%	MEAN (%)	t_{ISUD}	t_{BEST}	$t_{\textbf{AUG}}$	K	$\lvert\mathcal{S}\rvert^{\text{D}}$	$\lvert\mathcal{S}\rvert^{\text{N}}$
15	NONE	1	0	9 [8]	1.6	2.6	1.8	0.10	29	3.4	86
	CLIQUE	3	2	5 [3]	1.1	8.1	3.3	0.14	26	4.0	159
	CYCLE	2	1	7 [6]	0.5	3.7	2.3	0.12	26	3.9	102
	BOTH	5	2	3 [2]	0.4	12.4	4.1	0.16	30	4.1	154
	PRIO	5	2	3 [2]	0.3	9.8	4.0	0.16	30	4.1	148
	BEST	**5**	**2**	**3 [1]**	**0.1**	**6.9**	**3.8**	–	–	–	–
20	NONE	4	0	6 [4]	1.0	3.2	2.2	0.10	28	3.4	60
	CLIQUE	5	2	3 [1]	0.9	6.6	3.0	0.12	28	3.7	138
	CYCLE	4	0	6 [4]	1.0	3.5	2.2	0.10	28	3.4	93
	BOTH	7	2	1 [0]	0.3	11.6	3.7	0.14	27	3.8	150
	PRIO	7	2	1 [0]	0.3	9.1	3.6	0.13	27	3.8	143
	BEST	**8**	**1**	**1 [0]**	**0.0**	**5.9**	**3.2**	–	–	–	–
35	NONE	9	0	1 [0]	0.0	3.3	2.2	0.9	24	3.0	73
	CLIQUE	9	0	1 [0]	0.0	7.1	2.2	0.9	25	2.9	153
	CYCLE	9	0	1 [0]	0.0	3.8	2.2	0.9	25	2.9	117
	BOTH	9	0	1 [0]	0.0	10.9	2.2	0.9	25	2.9	162
	PRIO	9	0	1 [0]	0.0	8.1	2.2	0.9	25	2.9	155
	BEST	**9**	**0**	**1 [0]**	**0.0**	**4.2**	**2.1**	–	–	–	–

Parameters: INC_MAX = 10 and SEP_MAX = 40

SPPAA01 for $\pi = 15\%$ (closest to real-life problems) we see that (1) when no branching is made, PRIO solves 5/10 problems to optimality (against 2/10 for NONE), and 9/10 within 2% of the optimum (5/10 for NONE); and (2) when variables are fixed, PRIO solves 9/10 problems to optimality (2/10 for NONE) and all of them within 2% of the optimum (6/10 for NONE). Cutting planes therefore solve 7 out of the 8 problems that ISUD did not solve previously, hence yielding an improvement of 87.5% over SPPAA01 for $\pi = 15\%$.

Figure 6 displays the four performance diagrams of the algorithms for branching strategies NOBR, LAST, FIRST and COVER, over all SPPAA04 and SPPAA01 instances (70). Here again, whatever the branching strategy, cutting planes allow to solve many more instances. However, the improvement factor is much higher when no variable is fixed (less that 50% of the instances are solved without cuts; against more than 80% for PRIO or BOTH separation strategies). Interestingly, the time-increase factor (x-axis) never exceeds 10, and is most of the time lower than 4. Hence, the addition of cutting planes does not slow down the algorithm too much. Detailed computing times for each of the two problems are displayed in Figs. 7 and 8, respectively. The COVER branching strategy shows slightly better performances, but not significantly enough to draw any conclusion yet. Moreover, more instances are solved by ISUD faster than by CPLEX.

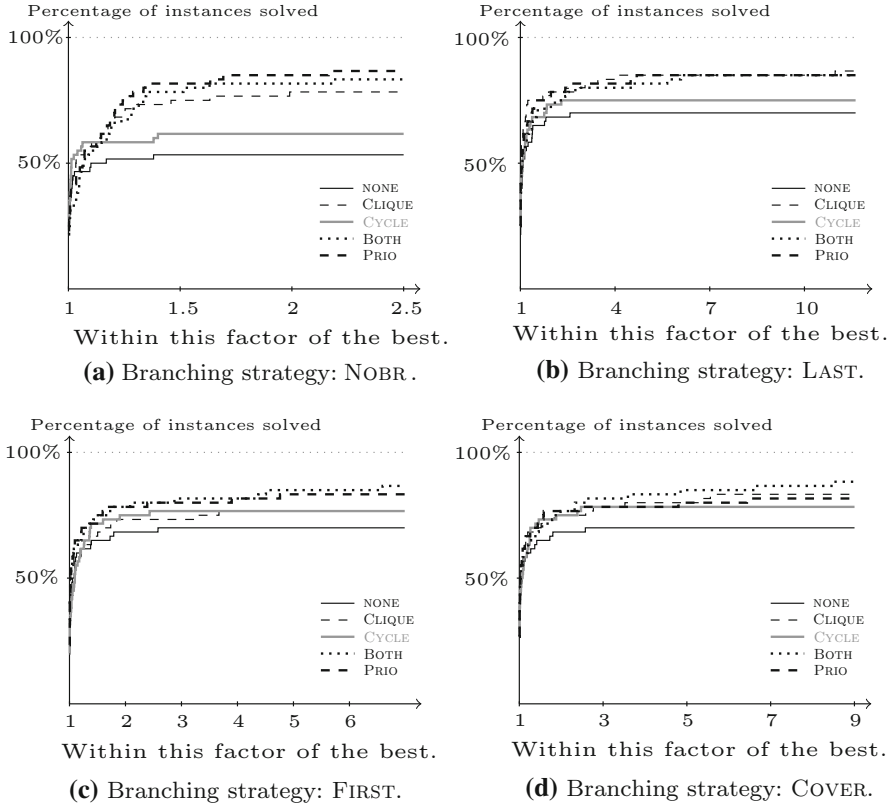
Tables 6 and 7 display specific characteristics concerning the cutting planes generated during the process. For each branching strategy (NOBR and LAST), the number

Table 5 Comparison of the performances of ISUD with branching LAST for instance SPPAA04

π (%)	ALGO	GAP				Time (s)			AUG		
		0%	≤2%	>2%	MEAN (%)	t_{ISUD}	t_{BEST}	$t_{\textbf{AUG}}$	K	$\lvert\mathcal{S}\rvert^{\text{D}}$	$\lvert\mathcal{S}\rvert^{\text{N}}$
15	NONE	3	3	4 [4]	0.6	6.2	4.9	0.18	31	4.0	72
	CLIQUE	4	3	3 [3]	0.1	34.9	20.5	0.72	31	4.1	180
	CYCLE	3	3	4 [3]	0.7	7.2	4.9	0.18	30	3.9	102
	BOTH	5	2	3 [2]	0.1	34.3	13.9	0.48	31	4.1	177
	PRIO	5	2	3 [2]	0.0	24.3	10.9	0.38	31	4.1	168
	BEST	**6**	**1**	**3 [2]**	**0.0**	**24.3**	**15.0**	–	–	–	–
20	NONE	3	2	5 [4]	0.5	5.4	3.6	0.13	33	3.4	76
	CLIQUE	7	3	0 [0]	0.2	15.6	9.7	0.30	32	4.0	146
	CYCLE	3	3	4 [4]	0.2	9.0	5.2	0.18	35	3.4	110
	BOTH	9	0	1 [1]	0.0	28.7	10.8	0.35	32	3.9	150
	PRIO	7	1	2 [1]	0.4	27.8	14.0	0.47	31	3.8	142
	BEST	**10**	**0**	**0 [0]**	**0.0**	**15.1**	**7.9**	–	–	–	–
35	NONE	9	1	0 [0]	0.0	3.5	2.4	0.9	25	2.9	65
	CLIQUE	9	1	0 [0]	0.1	8.9	3.0	0.11	27	3.0	140
	CYCLE	9	1	0 [0]	0.0	4.3	2.5	0.10	25	2.9	108
	BOTH	10	0	0 [0]	0.0	12.9	2.5	0.10	27	2.9	158
	PRIO	10	0	0 [0]	0.0	9.8	2.5	0.9	26	2.9	150
	BEST	**10**	**0**	**0 [0]**	**0.0**	**5.3**	**2.4**	–	–	–	–

Parameters: INC_MAX = 10 and SEP_MAX = 40

of instances solved within 2% of the optimum is shown under label INST. The mean time spent in the primal separation and cut management process including the cut pool management is given (t_{SEP}), as well as the mean number of separation problems solved (n_{SEP}), the mean number of primal clique cuts (n_{CL}), and primal odd-cycle cuts (n_{CY}). We can see that the time spent in the separation process is much higher when branching is applied. Indeed, in our algorithm, between two branching stages, SEP_MAX separation problems are solved. When no branching is performed, at most one sequence of SEP_MAX separation problems are solved and in case of failure to find a new primal cut, the algorithm instantly stops. This can also be seen in the n_{SEP} column, for which the number of separation problems solved with branching is significantly higher than without branching. Moreover, the time t_{SEP}, and number of **P-SEP** n_{SEP} are significantly higher when the algorithm fails to find a good solution (gap $>$ 2%). This reflects the struggling of the algorithm to improve the solution and the associated running time increase. As it is often the case for the SPP, clique cuts are more efficient, and easier to find. Furthermore, as seen in Sect. 4.1, the requirements for an odd-cycle cut to be a primal cut are harder to meet (alternated cycles) than those that apply to clique cuts (only one vertex of the clique must be in the current solution); this also make them rarer.



(a) Branching strategy: NOBR. **(b)** Branching strategy: LAST.

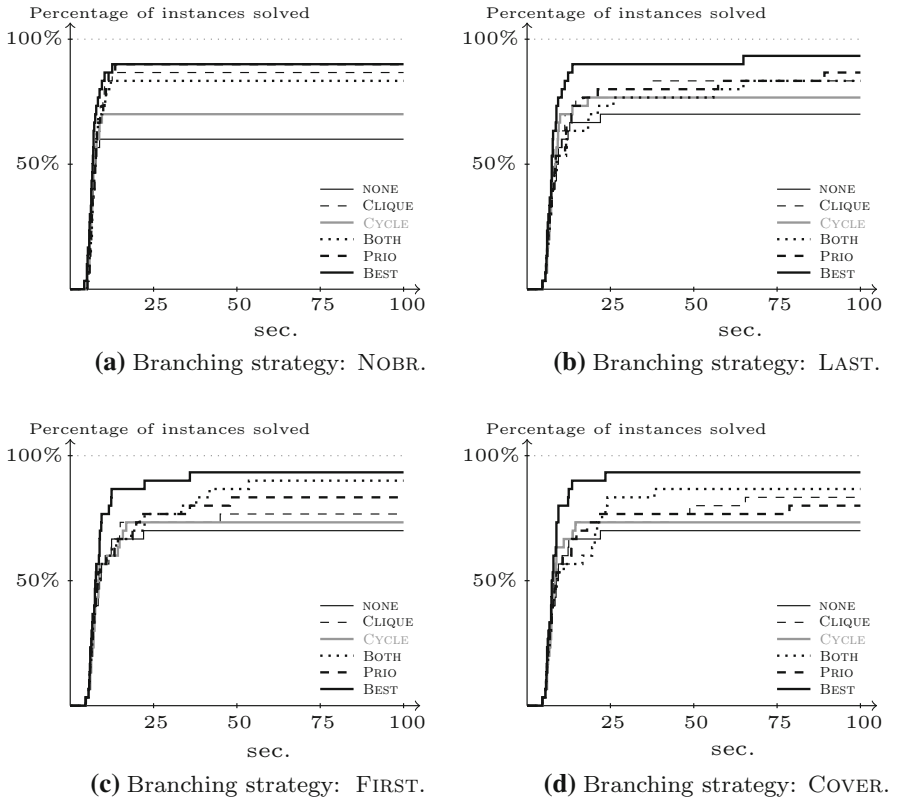
(c) Branching strategy: FIRST. **(d)** Branching strategy: COVER.

Fig. 6 Performance diagrams over all SPPAA04 and SPPAA01 instances for the four branching strategies (NOBR (*top-left*), LAST (*top-right*), FIRST (*bottom-left*) and COVER (*bottom-right*). An instance is considered as solved if the gap is lower than 2%

Definition 4 A solution $\boldsymbol{x}^0$ is *ι-optimal* if it is optimal for the restriction of SPP to $\mathcal{P} \cup \mathcal{I}^\iota$, i.e., for the set of all at most ι-incompatible columns of $\boldsymbol{A}$ computed at $\boldsymbol{x}^0$.

Detailed results of ISUD on VCS1200 and VCS1600 are displayed in Table 8 and 9. Branching did not change the results, hence the figures shown here correspond to the NOBR strategy and the time limit of half an hour was never reached. The first column shows the perturbation degree π, and the second the cutting plane strategy ALGO. The next four display the number of instances solved to optimality (0%), with a positive gap $\leq 2\%$ and with a gap $> 2\%$, as well as the mean gap computed over instances where solutions contained no more artificial columns (MEAN). Columns t_{ISUD} and t_{BEST} indicate the total running time and the time spent before reaching the best solution found, respectively. Then, the number of instances for which ι-optimality has been proved is given for $\iota = 7$ and $\iota = 8$, and finally the mean size of the disjoint ($|\mathcal{S}|^D$) and nondisjoint ($|\mathcal{S}|^N$) combination are shown.

In the experiments on VCS1200 and VCS1600, we chose a lower value for the maximum incompatibility number (INC_MAX $= 8$) than for the smaller problems;

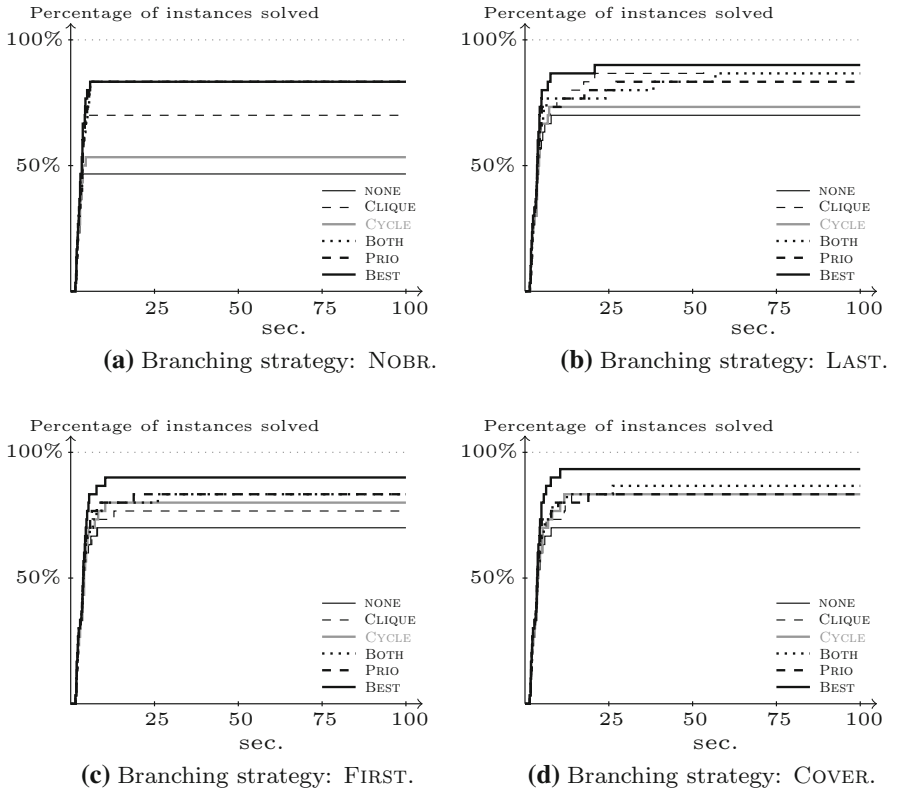


(a) Branching strategy: NOBR. **(b)** Branching strategy: LAST.

(c) Branching strategy: FIRST. **(d)** Branching strategy: COVER.

Fig. 7 Percentage of instances solved over solution time for all SPPAA01 instances for the four branching strategies (NOBR (*top-left*), LAST (*top-right*), FIRST (*bottom-left*) and COVER (*bottom-right*). An instance is considered as solved if the gap is lower than 2%

this choice is motivated by several reasons. First, it limits the running time of the algorithm. Second, this number is already high compared to what swap heuristics can consider. As explained in Sect. 3.1.2, the incompatibility degree of $\boldsymbol{A}_{\cdot j}$ is proportional to the number of sequences of consecutive tasks from $\boldsymbol{A}_{\cdot j}$ that are not performed consecutively in the current solution. Hence, it is a kind of measure of the primal distance between a column and the current solution. This number can be compared to the typical parameter of a swap heuristic that tries to swap parts of the current solution to form new individual schedules which is the number of times an individual schedule of the current solution may be split. Nonbasic columns for which $\iota = 8$ are made of at least 4 separate sequences of tasks performed consecutively in the current solution (there is no maximum number); in practice, this number usually ranges between 6 and 7. For an average of 40 nonzero entries per columns, this number therefore seems reasonable. It is in particular much higher than the typical number of splits allowed in the existing schedules in a swap heuristic, and the neighborhood explored by our algorithm is hence significantly wider than that of a swap heuristic.



(a) Branching strategy: NOBR. **(b)** Branching strategy: LAST.

(c) Branching strategy: FIRST. **(d)** Branching strategy: COVER.

Fig. 8 Percentage of instances solved over solution time for all SPPAA04 instances for the four branching strategies (NOBR (*top-left*), LAST (*top-right*), FIRST (*bottom-left*) and COVER (*bottom-right*). An instance is considered as solved if the gap is lower than 2%

Furthermore, the numerical results show that, even though primal cuts yield no improvement, they prove the ι-optimality of the solution in many cases, i.e., they show that the solution returned by ISUD is optimal for the restriction of SPP to the set of at most ι-incompatible columns ($\mathcal{I}^{\iota}$). This is especially true in the case of VCS1200, where 3 instances are proven 7-optimal without cutting planes, whereas 26 are with cutting planes.

Finally, the failure of the cutting planes to improve the results of ISUD on VCS1200 and VCS1600 can be partly explained by the nature of these instances. Although they describe real-world problems, they are in fact large instances obtained from column generation solution of bus drivers scheduling instances. All columns generated throughout the Branch-and-Price process are stored and put together to form a large scale instance, the solution of which is known. However, due to the intrinsic nature of that solution process, the density of the vertices of the resulting relaxed polyhedron ($\mathcal{F}_{\text{SPP}^{\text{LR}}}$) tends to be higher near the optimal solution. Hence, if a wrong decision is made in the beginning, there is a high probability of going to a region of the polyhedron where very few extreme integer neighbors, and

Table 6 Cutting planes behavior of ISUD on SPPAA01

GAP	ALGO	INST		t_{SEP} (s)		MEAN					
		NOBR	LAST	NOBR	LAST	NOBR			LAST		
						n_{SEP}	n_{CL}	n_{CY}	n_{SEP}	n_{CL}	n_{CY}
≤2%	NONE	18	21	0.0	0.0	0	0	0	0	0	0
	CLIQUE	26	25	0.5	2.8	56	240	0	164	853	0
	CYCLE	21	23	0.3	0.7	12	0	28	44	0	125
	BOTH	25	25	1.7	7.1	74	289	95	159	808	233
	PRIO	27	26	1.2	4.9	95	356	22	219	1073	48
>2%	NONE	12	9	0.0	0.0	0	0	0	0	0	0
	CLIQUE	4	5	1.5	29.9	68	454	0	290	3383	0
	CYCLE	9	7	1.1	6.8	21	0	122	90	0	424
	BOTH	5	5	7.1	51.8	59	653	202	241	2957	567
	PRIO	3	4	4.8	27.3	141	998	79	336	3259	153

Comparison of branching strategies NOBR and LAST

Table 7 Cutting planes behavior of ISUD on SPPAA04

GAP	ALGO	INST		t_{SEP} (s)		MEAN					
		NOBR	LAST	NOBR	LAST	NOBR			LAST		
						n_{SEP}	n_{CL}	n_{CY}	n_{SEP}	n_{CL}	n_{CY}
≤2%	NONE	14	21	0.0	0.0	0	0	0	0	0	0
	CLIQUE	21	27	1.5	3.1	83	446	0	149	734	0
	CYCLE	16	22	0.3	0.5	15	0	52	29	0	100
	BOTH	25	26	5.2	8.4	91	438	103	138	665	165
	PRIO	25	25	2.2	3.2	106	480	31	144	658	44
>2%	NONE	16	9	0.0	0.0	0	0	0	0	0	0
	CLIQUE	9	3	1.6	34.9	58	366	0	540	3997	0
	CYCLE	14	8	0.4	3.3	17	0	39	101	0	514
	BOTH	5	4	5.2	41.5	74	426	88	375	2458	580
	PRIO	5	5	2.4	25.3	95	453	30	576	3033	216

Comparison of branching strategies NOBR and LAST

thus integer directions, exist. Getting back to an area where the density of integer neighbors is higher, only by traveling alongside integer-to-integer edges, may then become extremely complicated. In the context of all-integer column generation, i.e., alternatively generate columns and improve the integer solution with ISUD as evoked in the introduction, this issue would disappear, because the dynamically generated columns increase the density of extreme points around the current solution, hence providing many more potential directions to the complementary problem.

Table 8 Results for instance VCS1200

π (%)	ALGO	GAP				Time (s)		ι-OPT		**AUG**	
		0%	≤2%	>2%	MEAN (%)	t_{ISUD}	t_{BEST}	7-OPT	8-OPT	$\lvert\mathcal{S}\rvert^{\text{D}}$	$\lvert\mathcal{S}\rvert^{\text{N}}$
20	NONE	4	0	6 [6]	0.0	53	46	3	2	2.5	198
	CLIQUE	4	0	6 [6]	0.0	112	46	8	2	2.5	480
	CYCLE	4	0	6 [6]	0.0	55	46	3	2	2.5	213
	BOTH	4	0	6 [6]	0.0	164	46	8	2	2.5	476
	PRIO	4	0	6 [6]	0.0	112	46	8	2	2.5	480
35	NONE	8	0	2 [2]	0.0	50	43	0	0	2.3	151
	CLIQUE	8	0	2 [2]	0.0	142	55	8	0	2.3	361
	CYCLE	8	0	2 [2]	0.0	53	43	0	0	2.3	176
	BOTH	8	0	2 [2]	0.0	166	49	8	0	2.3	371
	PRIO	8	0	2 [2]	0.0	142	55	8	0	2.3	361
50	NONE	10	0	0 [0]	0.0	37	28	0	0	2.2	132
	CLIQUE	10	0	0 [0]	0.0	81	29	10	0	2.2	315
	CYCLE	10	0	0 [0]	0.0	39	29	0	0	2.2	156
	BOTH	10	0	0 [0]	0.0	130	30	10	0	2.2	354
	PRIO	10	0	0 [0]	0.0	81	29	10	0	2.2	315

Parameters: INC_MAX = 8 and SEP_MAX = 40; Branching NOBR

Table 9 Results for instance VCS1600

π (%)	ALGO	GAP				Time (s)		ι-OPT		**AUG**	
		0%	≤2%	>2%	MEAN (%)	t_{ISUD}	t_{BEST}	7-OPT	8-OPT	$\lvert\mathcal{S}\rvert^{\text{D}}$	$\lvert\mathcal{S}\rvert^{\text{N}}$
20	NONE	5	0	5 [5]	0.0	462	324	6	6	2.8	528
	CLIQUE	5	0	5 [5]	0.0	646	325	6	6	2.8	831
	CYCLE	5	0	5 [5]	0.0	465	324	6	6	2.8	534
	BOTH	5	0	5 [5]	0.0	723	355	6	6	2.8	806
	PRIO	5	0	5 [5]	0.0	647	325	6	6	2.8	831
35	NONE	6	0	4 [4]	0.0	433	244	7	6	2.4	519
	CLIQUE	6	0	4 [4]	0.0	750	244	9	6	2.4	753
	CYCLE	6	0	4 [4]	0.0	459	244	7	6	2.4	539
	BOTH	6	0	4 [4]	0.0	813	244	9	6	2.4	761
	PRIO	6	0	4 [4]	0.0	750	244	9	6	2.4	753
50	NONE	8	0	2 [2]	0.0	404	156	9	8	2.2	517
	CLIQUE	8	0	2 [2]	0.0	492	156	9	8	2.2	691
	CYCLE	8	0	2 [2]	0.0	405	156	9	8	2.2	521
	BOTH	8	0	2 [2]	0.0	521	156	9	8	2.2	693
	PRIO	8	0	2 [2]	0.0	492	156	9	8	2.2	691

Parameters: INC_MAX = 8 and SEP_MAX = 40; Branching NOBR

6 Conclusions

In this work, we proposed a purely primal formulation of ISUD and introduced cutting planes in the complementary problem CP. This method is guaranteed to reach optimality whenever an appropriate set of cutting planes is used (Gomory-Young, for example). We showed that the incompatibility degree defines a neighborhood of the current solution on which the augmentation problem is solved with integer programming techniques. In order to improve the practical numerical performances, we compromise by both heuristically increasing the neighborhood and by stopping the algorithm prematurely. The efficiency of linear programming and of our separation algorithms (Algorithms 2 and 3) allows us to explore a larger neighborhood than what a typical exchange heuristic would consider.

The work conducted on the mathematical formulation led us to characterize the set of cuts that can be transferred to CP as a nonempty subset of primal cuts tight at $\boldsymbol{x}^0$. We showed that, given a fractional direction $\boldsymbol{d}^\star$, a primal clique (or odd-cycle) cut exists if and only if there exists one that only involves variables of $Supp\left(\boldsymbol{d}^\star\right)$. Furthermore, in the case of primal odd-cycles, the separation over all variables is strictly equivalent to that over $Supp\left(\boldsymbol{d}^\star\right)$. Tests conducted over a benchmark of practical-like instances proved the potential of our method and highlight the importance of primal cutting planes in ISUD.

This work should be extended threefold. First, to gain better understanding of our algorithm, a larger benchmark is to be taken in consideration and a combination of cutting planes and other techniques (such as those proposed in [23] for instance) must be tested over that larger set of problems. Second, other cutting planes families should be taken in consideration, and the corresponding primal separation procedures must be developed. Last, the algorithm need to be extended to more general {0,1}-programs, and not only to SPP instances. This last point seems to be the most important if the method is to be used in practice, since supplementary constraints are often added to set partitioning problems. All these extensions are active research subjects and the present work will have extensions in a near future. Of course, further extensions to mixed {0,1}-problems and to general mixed-integer ones would be very interesting but also harder, especially for the latter class. This is because the property of any feasible solution being a vertex of the polyhedron of the continuous relaxation is not generally true.

Acknowledgements This work was supported by a Collaborative Research and Development Grant from the Natural Sciences and Engineering Research Council of Canada (NSERC) and Kronos Inc. Samuel Rosat benefitted of a grant from the International Internship Program of the Fonds de Recherche du Québec Nature et Technologies (FRQNT).

References

1. Balas, E., Padberg, M.: On the set-covering problem: 2—an algorithm for set partitioning. Oper. Res. **23**(1), 74–90 (1975)
2. Baldacci, R., Mingozzi, A.: A unified exact method for solving different classes of vehicle routing problems. Math. Program. **120**, 347–380 (2009)

3. Ben-Israel, A., Charnes, A.: On some problems of diophantine programming. Cahiers du Centre d'Études de Recherche Opérationnelle **4**, 215–280 (1962)
4. Desrosiers, J., Dumas, Y., Solomon, M.M., Soumis, F.: Time constrained routing and scheduling. Handb. Oper. Res. Manag. Sci. **8**, 35–139 (1995)
5. Eisenbrand, F., Rinaldi, G., Ventura, P.: Primal separation for 0/1 polytopes. Math. Program. **95**(3), 475–491 (2003)
6. Elhallaoui, I., Metrane, A., Desaulniers, G., Soumis, F.: An improved primal simplex algorithm for degenerate linear programs. INFORMS J. Comput. **23**(4), 569–577 (2011)
7. Elhallaoui, I., Metrane, A., Soumis, F., Desaulniers, G.: Multi-phase dynamic constraint aggregation for set partitioning type problems. Math. Program. **123**(2), 345–370 (2010)
8. Garfinkel, R.S., Nemhauser, G.L.: The set-partitioning problem: set covering with equality constraints. Oper. Res. **17**(5), 848–856 (1969)
9. Glover, F.: A new foundation for a simplified primal integer programming algorithm. Oper. Res. **16**, 727–740 (1968)
10. Gomory, R.E.: Outline of an algorithm for integer solutions to linear program. Bull. Am. Math. Soc. **64**(5), 275–278 (1958)
11. Gomory, R.E.: All-integer integer programming algorithm. Industrial Scheduling, pp. 193–206 (1963)
12. Haus, U.U., Köppe, M., Weismantel, R.: A primal all-integer algorithm based on irreducible solutions. Math. Program. **96**(2), 205–246 (2003)
13. Kallio, M.J., Porteus, E.L.: A class of methods for linear programming. Math. Program. **14**(1), 161–169 (1978)
14. Letchford, A.N., Lodi, A.: Primal cutting plane algorithms revisited. Math. Methods Oper. Res. **56**(1), 67–81 (2002)
15. Letchford, A.N., Lodi, A.: An augment-and-branch-and-cut framework for mixed 0–1 programming. In: Jünger, M., Reinelt, G., Rinaldi, G. (eds.) Combinatorial Optimization - Eureka, You Shrink!. Lecture Notes in Computer Science, vol. 2570, pp. 119–133. Springer, Berlin (2003)
16. Letchford, A.N., Lodi, A.: Primal separation algorithms. Q. J. Belg Fr. Ital. Oper. Res. Soc. **1**(3), 209–224 (2003)
17. Metrane, A., Soumis, F., Elhallaoui, I.: Column generation decomposition with the degenerate constraints in the subproblem. Eur. J. Oper. Res. **207**(1), 37–44 (2010)
18. Omer, J., Rosat, S., Raymond, V., Soumis, F.: Improved Primal Simplex: A More General Theoretical Framework and an Extended Experimental Analysis. Les Cahiers du GERAD G-2014-13, HEC Montréal, Canada. Submitted to Informs Journal on Computing (2014)
19. Östergård, P.R.J.: A new algorithm for the maximum-weight clique problem. Nord. J. Comput. **8**(4), 424–436 (2001)
20. Padberg, M.W.: On the facial structure of set packing polyhedra. Math. Program. **5**(1), 199–215 (1973)
21. Rönnberg, E., Larsson, T.: Column generation in the integral simplex method. Eur. J. Oper. Res. **192**(1), 333–342 (2009)
22. Rönnberg, E., Larsson, T.: All-integer column generation for set partitioning: basic principles and extensions. Eur. J. Oper. Res. **233**(3), 529–538 (2014)
23. Rosat, S., Elhallaoui, I., Soumis, F., Chakour, D.: Influence of the normalization constraint on the integral simplex using decomposition. Discret. Appl. Math. **217, Part 1**, 53–70 (2017)
24. Rosat, S., Elhallaoui, I., Soumis, F., Lodi, A.: Integral simplex using decomposition with primal cuts. In: Gudmundsson, J., Katajainen, J. (eds.) Experimental Algorithms, *Lecture Notes in Computer Science*, vol. 8504, pp. 22–33. Springer, New York (2014)
25. Rozenknop, A., Wolfler Calvo, R., Alfandari, L., Chemla, D., Létocart, L.: Solving the electricity production planning problem by a column generation based heuristic. J. Sched. **16**(6), 585–604 (2013)
26. Salkin, H.M., Koncal, R.D.: Set covering by an all-integer algorithm: computational experience. J. ACM **20**(2), 189–193 (1973)
27. Saxena, A.: Set-partitioning via integral simplex method. Unpublished manuscript, OR Group, Carnegie-Mellon University, Pittsburgh (2003)
28. Schulz, A.S., Weismantel, R., Ziegler, G.M.: 0/1-integer programming: Optimization and augmentation are equivalent. In: Spirakis, P. (ed.) ESA '95, LNCS, vol. 979, pp. 473–483. Springer, Berlin (1995)
29. Spille, B., Weismantel, R.: Primal integer programming. In: Aardal, K., Nemhauser, G., Weismantel, R. (eds.) Discrete Optimization, Handbooks in Operations Research and Management Science, vol. 12, pp. 245–276. Elsevier, Amsterdam (2005)

30. Stallmann, M.F., Brglez, F.: High-contrast algorithm behavior: observation, hypothesis, and experimental design. In: Proceedings of the 2007 Workshop on Experimental Computer Science, ExpCS '07. ACM, New York, NY, USA (2007)
31. Stojković, M., Soumis, F., Desrosiers, J.: The operational airline crew scheduling problem. Transp. Sci. **32**(3), 232–245 (1998)
32. Thompson, G.L.: An integral simplex algorithm for solving combinatorial optimization problems. Comput. Optim. Appl. **22**(3), 351–367 (2002)
33. Towhidi, M., Desrosiers, J., Soumis, F.: The positive edge criterion within COIN-OR's CLP. Comput. Oper. Res. **49**, 41–46 (2014)
34. Trubin, V.: On a method of solution of integer linear programming problems of a special kind. Sov. Math. Dokl. **10**, 1544–1546 (1969)
35. Young, R.D.: A primal (all-integer) integer programming algorithm. J. Res. Natl. Bur. Stand. B Math. Math. Phys. **69B**, 213–250 (1965)
36. Young, R.D.: A simplified primal (all-integer) integer programming algorithm. Oper. Res. **16**(4), 750–782 (1968)
37. Zaghrouti, A., Soumis, F., El Hallaoui, I.: Integral simplex using decomposition for the set partitioning problem. Oper. Res. **62**(2), 435–449 (2014)