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

...

...

Received: date / Accepted: date

Abstract Keywords augmented Lagrangian method · proximal bundle method · nonlinear block Gauss-Seidel method · simplicial decomposition method · parallel computing

Mathematics Subject Classification (2000) 90-08, 90C06, 90C11, 90C15, 90C25, 90C26, 90C30, 90C46

1 Introduction and Background

The problem of interest has the form

$$f^* := \min_{x,y,z} \left\{ \sum_{i=1}^{m} p_i \left[c^\top x_i + d_i^\top y_i \right] : (x_i, y_i) \in K_i, x_i = z \text{ for } i = 1, \ldots, m,\ z \in \mathbb{R}^{n_1} \right\}, \tag{1}$$

where K is a bounded polyhedral set intersected with mixed-integrality constraints.

We develop a branch-and-bound approach to solving (1) based on solving Lagrangian relaxations due to relaxing $Qx = z$. At each node, we use the recently developed algorithm SDM-GS-ALM to solve two equivalent characterizations of the Lagrangian dual problem. The Lagrangian dual function is defined

$$\phi(\omega) := \min_{x,y,z} \left\{ p_i \left[c^\top x_i + \omega_i^\top (x_i - z) + d_i^\top y_i \right] : (x_i, y_i) \in K_i, z \in \mathbb{R}^{n_1} \right\}.$$

Under the dual feasibility condition $\sum_{i=1}^{m} p_i \omega_i = 0$, we have

$$\phi(\omega) = \min_{x,y,z} \left\{ p_i \left[(c + \omega_i)^\top x_i + d_i^\top y_i \right] : (x_i, y_i) \in K_i, z \in \mathbb{R}^{n_1} \right\}$$

The Lagrangian dual problem is:

$$\phi^* := \max_{\omega \in \Omega} \phi(\omega). \tag{2}$$

Its well-known primal characterization takes the form

$$f_C^* := \min_{x,y,z} \left\{ \sum_{i=1}^{m} p_i \left[c^\top x_i + d_i^\top y_i \right] : (x_i, y_i) \in \text{conv}(K_i), x_i = z \text{ for } i = 1, \ldots, m,\ z \in \mathbb{R}^{n_1} \right\}. \tag{3}$$

It is well-known that $f_C^* = \phi^*$, although in general, $\phi^* \leqslant f^*$, with $\phi^* < f^*$ being typical with the presence of mixed-integrality constraints on x and y.

Dual optimal solutions to (2), when they exist, are generically denoted by ω^*, and and the set of corresponding optimal primal solutions verifying the value $\phi(\omega^*)$ is denoted by $K^*(\omega^*)$. The optimal solutions to the second characterization (3), when they exist, are denoted (x^*, y^*, z^*). We shall denote the set $Z := \{(x, y) \in K : \exists\ z \in \mathbb{R}^n \text{ s.t. } x_i = z \text{ for } i = 1, \ldots, m\}$.

We shall use both characterizations (2) and (3) to inform different ways of branching on variables and branching values.

1. Viewing the node subproblems as instances of problem (3), we have optimal solutions (x^*, y^*, z^*) to problem (3) that satisfy

   $$(x^*, y^*, z^*) \in \{(x, y, z) : (x, y) \in \text{conv}(K), x_i = z,\ i = 1, \ldots, m\},$$

   but $(x^*, y^*) \notin K$ in the case of strict duality gap $\phi_C^* < f^*$. Thus we would project (x^*, y^*) onto K, denoting the projection by $\text{P}_K((x^*, y^*))$. (Given that K is defined by polyhedral and mixed-integer constraints, this projection can just take the form of rounding the components of (x, y) with integer restrictions to their nearest integer values.) We would consider the discrepancies $(x^*, y^*) - \text{P}_K((x^*, y^*))$ and their dispersions.
2. Viewing the node subproblems as instances of problem (2), assuming we have a dual optimal solution ω^*, we have corresponding sets of primal certificate solutions $K^*(\omega^*)$. For these solutions, we have $K^*(\omega^*) \cap Z = \varnothing$ in the case of duality gap $\phi_C^* < f^*$. We consider one of two projections:
   (a) Take the projection $\text{P}_Z(K(\omega^*))$. Conceptually, we want to accumulate the dispersions in some set-wise or point-wise fashion. Ideally, we do not consider the dispersion of only one $\widehat{x} \in K^*(\omega^*)$, but the entire set $K^*(\omega^*)$, which is contained in the highest dimensional face of $\text{conv}(K)$ for which (x^*, y^*) is in its relative interior.
   (b) Rather than projection, we use the information from x^*, z^* and ω^*. (Or, more practically, x^k, z^k and ω^k.) We still work with the set $K^*(\omega^k)$, whose vertex description is available from computing $\widehat{x}^t \in K^*(\omega^t)$ at each iteration $t = 1, \ldots, k$.

Due to the large number of variables that would be branched on under the first approach, branching is not based on the first approach. Although we do use the dispersions from the first approach discrepancies to inform branching based on the second approach. Branching occurs on components $x_{i,j}$ of each x_i

that either violate integrality constraints, or that show dispersion from some consensus value associated with other violated constraints. The constraints added due to branching are denoted $x \in B$. Based on this constraint, define $K_B := \{(x, y) \in K : x \in B\}$, so that we define, respectively,

$$f_B^* := \min_{x,y,z} \left\{ \sum_{i=1}^{m} p_i \left[c^\top x_i + d_i^\top y_i \right] : (x_i, y_i) \in K_i,\ x_i \in B_i,\ x_i = z \text{ for } i = 1, \ldots, m,\ z \in \mathbb{R}^{n_1} \right\}, \tag{4}$$

$$\phi_B^* := \max_{\omega \in \Omega} \phi_B(\omega). \tag{5}$$

with

$$\phi_B(\omega) = \min_{x,y,z} \left\{ p_i \left[(c + \omega_i)^\top x_i + d_i^\top y_i \right] : (x_i, y_i) \in K_i,\ x_i \in B_i,\ z \in \mathbb{R}^{n_1} \right\} \tag{6}$$

and

$$\phi_B^* = \min_{x,y,z} \left\{ \sum_{i=1}^{m} p_i \left[c^\top x_i + d_i^\top y_i \right] : \begin{array}{c} (x_i, y_i) \in \operatorname{conv}(K_i \cap B_i), \\ x_i = z \text{ for } i = 1, \ldots, m, \\ z \in \mathbb{R}^{n_1} \end{array} \right\}. \tag{7}$$

During each node processing of the branch-and-bound, an attempt is made to find a feasible solution to the node-specific instance of (4), which provides a finite upper bound, which we shall refer to as the *incumbent value* and denote as $\upsilon \in \mathbb{R}$. Furthermore, we use an iterative approach, referred to as the *oracle*, to solving each instance of (5) and (7) simultaneously, that is guaranteed to converge optimally. During the processing of each node, one of three things will happen:

1. The oracle terminates due to the objective value exceeding the incumbent value (fathoming due to bound), or due to subproblem infeasibility.
2. The oracle terminates due to optimality of (4) within pre-specified precision. That is $\|(x^*, y^*) - \mathrm{P}_K(x^*, y^*)\| + \sum_{i=1}^{m} \|x_i^* - z^*\| < \epsilon$. The node is fathomed after testing whether $f_B^* < \upsilon$ and updating $\upsilon \leftarrow f_B^*$ if that is the case. (In Caroe and Schultz, it is the vertex solutions $(\widehat{x}, \widehat{y})$ that are tested only, so that the criterion is $\|\widehat{x}_i - \widehat{z}\| < \epsilon$, where $\widehat{z}$ is some averaging of $\widehat{x}_i$, $i = 1, \ldots, m$.)
3. Otherwise, branchings need to be determined.

Definitions:
$\mathcal{S} := \{f, K, B\}$, $\mathcal{P} := \{\rho, \gamma, \epsilon, t_{max}, k_{max}\}$
υ is an incumbent value, and
Precondition:
It is assumed that the value of the current node has been checked to not exceed the incumbent value υ
function PROCESS($\mathcal{S}$, $\mathcal{P}$, ω^0, υ, $(\widetilde{x}, \widetilde{y}, \widetilde{z})$)
$(K^*(\omega^*), (x^*, y^*, z^*), \omega^*, \phi^*) \leftarrow$ **bound**($\mathcal{S}$, $\mathcal{P}$ ω^0)
if $\breve{\phi}^* > \upsilon$ **or** $\breve{\phi}^* = \infty$ **then**
fathom node
$\mathcal{B} \leftarrow \varnothing$
return $(\upsilon, \widetilde{z}, \mathcal{B})$
end if
$\mathcal{B} \leftarrow$ **findBranchings**$(K^*(\omega^*), (x^*, y^*, z^*), \omega^*)$
$(\widetilde{x}, \widetilde{y}, \widetilde{z}, \upsilon) \leftarrow$ **findFeasibleSolution**$(z^*, \omega^*, \mathcal{S}, \mathcal{P})$
if $\mathcal{B} \neq \varnothing$ **then**
node needs to branch
else
fathom node (Due to optimality)
end if
return $(\widetilde{x}, \widetilde{y}, \widetilde{z}, \upsilon), \mathcal{B})$
end function
function BOUND($\mathcal{S}$, $\mathcal{P}$ ω^0)
$(D^0, x^0, z^0, \omega^0, \breve{\phi}^0) \leftarrow$ Initialize$(\mathcal{S}, \mathcal{P}, \omega^0)$
$k \leftarrow 0$, **term** $\leftarrow$ **false**
while $\neg(\breve{\phi}^k > \upsilon$ **or term**) **do**
$(K^*(\omega^{k+1}), (x^{k+1}, y^{k+1}, z^{k+1}), \omega^{k+1}, \phi^{k+1}, \textbf{term}) \leftarrow$ PSCG$(\mathcal{S}, \mathcal{P}, D^k, x^k, z^k, \omega^k, \breve{\phi}^k)$
$k \leftarrow k + 1$
end while
return $\left(K^*(\omega^k), (x^k, y^k, z^k), (\widetilde{x}, \widetilde{y}, \widetilde{z}), \omega^k, \phi^k, \upsilon\right)$
end function

1.1 Parallel Stabilized Column Generation (PSCG)

In applying the AL method to problem (**??**), the continuous master problem for fixed $\omega \in Z^{\perp}$ takes the form

$$\phi_{\rho}^{AL}(\omega) := \min_{x,z} \{L_{\rho}(x, z, \omega), x \in \text{conv}(X), z \in Z\} \tag{8}$$

where the augmented Lagrangian (AL) relaxes $Qx = z$ and is defined by

$$L_{\rho}(x, z, \omega) := f(x) + \omega^{\top} Qx + \frac{\rho}{2} \|Qx - z\|_2^2 . \tag{9}$$

In the algorithm that follows, we use the following approximation $\widehat{\phi} : \mathbb{R}^q \times \mathbb{R}^n \times \mathbb{R}^q \mapsto \mathbb{R}$ of ϕ^C centered at (x^k, z^k), $k \geqslant 0$, in place the cutting plane model:

$$\widehat{\phi}(\omega, x^k, z^k) := L_{\rho}(x^k, z^k, \omega) + \frac{\rho}{2} \|Qx^k - z^k\|_2^2 .$$

The convex hull $\text{conv}(X)$ is not known explicitly, and so ϕ^C cannot be evaluated directly. Consequently, we additionally make use of the following

minorization $\breve{\phi}$ of ϕ^C that can be evaluated. For $x^k \in \text{conv}(X)$, $k \geqslant 0$, define $\breve{\phi}(\omega, x^k)$ as follows:

$$\breve{\phi}(\omega, x^k) := \min_x \left\{ f(x^k) + \nabla_x f(x^k)(x - x^k) + \omega^\top Q x : x \in X \right\}. \tag{10}$$

Observe that, due to the linearity of the objective function with respect to x in (10), the use of constraint sets X and $\text{conv}(X)$ are interchangeable, and so in evaluating $\breve{\phi}$, an explicit description of $\text{conv}(X)$ is not required. Furthermore, from the definition of ϕ^C, the convexity of f over $\mathbb{R}^n$, and the interchangeability of X and $\text{conv}(X)$ in (10), it is clear that for all $x^k \in \mathbb{R}^n$, $k \geqslant 0$, we have $\phi^C(\omega) \geqslant \breve{\phi}(\omega, x^k)$. Furthermore, when f is linear, we have $\phi^C(\omega) \equiv \breve{\phi}(\omega, x^k)$ for all x^k, $k \geqslant 0$; the two functions collapse into the same function with the centering at x^k of the latter function now irrelevant.

In [**?**], we developed an efficiently parallelizable iterative procedure to solving problem (**??**). In this paper, we incorporate this procedure as part of the bounding mechanism within a branch-and-bound approach....

For review purposes, the iterative dual procedure is given as follows.

Algorithm 1 A regular iteration of PSCG.

Preconditions:
$x^k \in \text{conv}(X)$, $z^k \in \text{argmin}_z \left\{ \|Qx - z\|^2 : z \in Z \cap B \right\}$, $\omega^k \in Z^\perp$,
$\breve{\phi}^k = \breve{\phi}(\omega^k, x^k)$, $\left\{ x^k + \alpha(\hat{x} - x^k) : \alpha \in [0,1] \right\} \subseteq D \subseteq \text{conv}(X)$ where
$\hat{x} \in \text{argmin}_x \left\{ \nabla_x L_\rho(x^k, z^k, \omega^k)(x - x^k) : x \in X \right\}$.
$\mathcal{S} := \{f, Q, X, Z, B\}$, $\mathcal{P} := \{\rho, \gamma, \epsilon, t_{max}, k_{max}\}$

function PSCG($\mathcal{S}$, $\mathcal{P}$, D^k, x^k, z^k, ω^k, $\breve{\phi}^k$)
  for $k = 1, 2, \ldots, k_{max}$ **do**
    Initialize $\omega^{k+1} \leftarrow \omega^k$, $\breve{\phi}^{k+1} \leftarrow \breve{\phi}^k$ ▷ (Default, null-step updates)
    $(\hat{x}^{k+1}, x^{k+1}, z^{k+1}, D^{k+1}, \Gamma) \leftarrow$ SDM-GS($L_\rho(\cdot, \cdot, \omega^k)$, X, $Z \cap B$, D^k, x^k, z^k, t_{max})
    if $L_\rho(x^{k+1}, z^{k+1}, \omega^k) + \frac{\rho}{2} \left\|Qx^{k+1} - z^{k+1}\right\|_2^2 - \breve{\phi}^k \leqslant \epsilon$ **then**
      return $(x^{k+1}, z^{k+1}, \omega^{k+1}, \breve{\phi}^{k+1}, D^{k+1}, \textbf{true})$
    end if
    $\tilde{\phi} \leftarrow L_\rho(x^{k+1}, z^{k+1}, \omega^k) + \frac{\rho}{2} \left\|Qx^{k+1} - z^{k+1}\right\|_2^2 - \Gamma$
    $\tilde{\gamma} \leftarrow \frac{\tilde{\phi} - \breve{\phi}^k}{L_\rho(x^{k+1}, z^{k+1}, \omega^k) + \frac{\rho}{2} \left\|Qx^{k+1} - z^{k+1}\right\|_2^2 - \breve{\phi}^k}$
    if $\tilde{\gamma} \geqslant \gamma$ **then**
      set $\omega^{k+1} \leftarrow \omega^k + \rho(Qx^{k+1} - z^{k+1})$, $\breve{\phi}^{k+1} \leftarrow \tilde{\phi}$
    end if
    Possibly update ρ, e.g., $\rho \leftarrow \frac{1}{\min\left\{\max\left\{(2/\rho)(1-\tilde{\gamma}), 1/(10\rho)\right\}, 10^{-4}\right\}, 10/\rho\}}$ as in [1]
  end for
  return $(\hat{x}^{k+1}, x^{k+1}, z^{k+1}, \omega^{k+1}, \breve{\phi}^{k+1}, D^{k+1}, \textbf{false})$
end function

Preconditions: $\tilde{x} \in \text{conv}(X)$, $\tilde{z} \in \text{argmin}_z \{F(\tilde{x}, z) : z \in Z\}$, $D \subseteq \text{conv}(X)$
function SDM-GS(F, X, Z, D, $\tilde{x}$, $\tilde{z}$, t_{max})
  for $t = 1, \ldots, t_{max}$ **do**
    $\tilde{x} \leftarrow \text{argmin}_x \{F(x, \tilde{z}) : x \in D\}$
    $\tilde{z} \leftarrow \text{argmin}_z \{F(\tilde{x}, z) : z \in Z\}$
  end for
  $\hat{x} \in \text{argmin}_x \{\nabla_x F(\tilde{x}, \tilde{z})(x - \tilde{x}) : x \in X\}$
  Reconstruct D to be any set such that
    $\{\tilde{x} + \alpha(\hat{x} - \tilde{x}) : \alpha \in [0,1]\} \subseteq D \subseteq \text{conv}(X)$
  Set $\Gamma \leftarrow -\nabla_x F(\tilde{x}, \tilde{z})(\hat{x} - \tilde{x})$
  return $(\hat{x}, \tilde{x}, \tilde{z}, D, \Gamma)$
end function

Algorithm PSCG addresses the solution to an alternative dual problem which is equivalent to (**??**) when f is linear, but in general provides a weaker dual bound otherwise. This dual problem is used to address the more general setting where f is convex but possibly nonlinear.

Proposition 1 *Let* $\left\{(x^k, z^k, \omega^k)\right\}$ *be a sequence generated by Algorithm 1 applied to problem* (1) *with* X *compact,* Z *a linear subspace,* $\omega^0 \in Z^\perp$, B *closed and convex,* $\rho > 0$, $\gamma \in (0,1)$, $\epsilon = 0$ *and* $k_{max} = \infty$. *If there exists a dual optimal solution* ω^* *to the dual problem* (**??**), *then either*

1. $\omega^k = \overline{\omega}$ *is fixed and optimal for* (**??**) *for* $k \geqslant \bar{k}$ *for some finite* $\bar{k}$; *or*
2. ω^k *is never optimal for* (**??**) *for any finite* $k \geqslant 1$, *but* $\lim_{k \to \infty} \omega^k = \overline{\omega}$ *is optimal,*

and the sequence $\{(x^k, z^k)\}$ *has limit points* $(\overline{x}, \overline{z})$*, each of which are optimal for problem* (**??**).

1.2 Parallelization and workload

The opportunities for parallelization and distribution of the computational workload in PSCG, as stated in Algorithm 1, are not immediately apparent. This subsection explicitly indicates which update problems may be solved in parallel, and the nature of the required communication between the parallel computational nodes.

The bulk of computational work, parallelization, and parallel communication occurs within the SDM-GS method stated in Algorithm **??**, where for the problems of interest, the following decomposable structures apply: $X = \prod_{i=1}^m X_i$, $D = \prod_{i=1}^m D_i$, and $F(x, z) = \sum_{i=1}^m F(x_i, z)$. In the larger context of Algorithm 1, the subproblem of Line 3 in Algorithm **??** can be solved in parallel given fixed $\tilde{z} \in Z$ and $\omega \in Z^\perp$ along the block indices $i = 1, \ldots, m$ as

$$\min_x \left\{ f_i(x) + (\omega_i)^\top Q_i x + \frac{\rho}{2} \|Q_i x - \tilde{z}_i\|_2^2 : x \in D_i \right\}, \tag{11}$$

while the subproblem of Line 6 is solved as

$$\min_x \left\{ \nabla_x f_i(\tilde{x}_i) + (\omega_i + \rho(Q_i \tilde{x}_i - \tilde{z}_i))^\top Q_i x : x \in X_i \right\}.$$

Remark 1 In the setting where problem (1) is a large-scale mixed-integer linear optimization problem, the subproblems of Line 3 are continuous convex quadratic optimization problems for each block $i = 1, \ldots, m$, which can be solved independently of one another and in parallel. In the same setting, the Line 6 subproblems are mixed-integer optimization problems for each block $i = 1, \ldots, m$, which can also be solved independently of one another and in parallel. Additionally, the reconstruction of D occurring in Line 8 can be done in parallel for each D_i along the indices $i = 1, \ldots, m$.

Parallel communication is needed for the computation of the z update. In the larger context of Algorithm 1, this takes the form of solving

$$\min_z \left\{ \sum_{i=1}^m \|Q_i \tilde{x}_i - z_i\|_2^2 : z \in Z \right\}.$$

This is solved as an averaging that requires the reduce-sum type parallel communication. The computation of values required to compute γ^k in Line 9 in Algorithm 1 also requires a reduce-sum type parallel communication. For implementation purposes, the computation of these values, including the computation of Γ from the SDM-GS call, can be combined into one reduce-sum communication. In total, each iteration of Algorithm 1 requires two reduce-sum type communications, one for computing the z-update, and one combined reduce-sum communication to compute scalars associated with the Lagrangian

bounds and the critical values for the termination conditions. The storage and updates of x^k and ω^k and D can also be done in parallel, while z^k and γ^k need to be computed and stored by every processor at each iteration k.

2 Branch-and-bound with SCG-GS-ALM as bounding procedure

In order to inform branching decisions, we need a sense of dispersion based on some violation of feasibility, and a value with which to bound.

2.1 Measure of feasibility

A criterion for fathoming by optimality may be given by

$$\|x^* - \mathrm{P}_X(x^*)\| + \|Qx^* - z^*\| < \epsilon. \tag{12}$$

for some suitably small tolerance $\epsilon > 0$. If the feasibility criterion (12) is not met, then branching must occur, and the branching should be informed by the dispersions measured in the two norm terms of (12). It will be convenient to to refine the dispersions, where the vector

$$\sigma_X := [\|x_i^* - \mathrm{P}_{X_i}(x_i^*)\|]_{i=1,\dots,m}$$

is defined. The diagonal matrix whose diagonal entries are taken from σ_X will be denoted Σ_X.

2.2 Branching decisions

We have the following approaches at measuring dispersion from the constraint $Qx = z$ at iteration k:

1. Given a dual feasible solution ω^k, for some $\widehat{x}^k \in \widehat{x}(\omega^k)$, setting $\widehat{x}^k = \left[\widehat{x}_i^k\right]_{i=1,\dots,m}$, we take an average $\widehat{z}^k = \mathrm{P}_Z(Q\widehat{x}^k)$. Based on this average, we compute a component-wise dispersion

$$\sigma_j^k := \sum_{i=1,\dots,m} \left| \left[Q_i \widehat{x}_i^k - \widehat{z}_i^k\right]_j \right|$$

for each component index $j = 1, \dots, n_x$, and denote $\sigma := [\sigma_j]_{j=1,\dots,n_x}$. This is the approach [in Caroe and Schultz]. One quirk with this approach is that, at the optimal dual solution ω^*, any of the previously generated solutions $\widehat{x} \in \widehat{x}(\omega^k)$ may also satisfy $\widehat{x} \in \widehat{\omega}^*$. This is especially true in the typically satisfied case that $c \in Z$. Also, the approach does not use information from the dispersions $\|x^* - \mathrm{P}_X(x^*)\|$. The latter issue is addressed with the following modification to σ_j^k:

$$\sigma_j^k := \left\| \left[([\sigma_X]_i + \epsilon)[Q_i \widehat{x}_i^k) - \widehat{z}_i^k]_j \right]_{i=1,\dots,m} \right\| + \|Qx^k - z^k\|$$

2. Let $\widehat{x}^t \in \widehat{x}(\omega^t)$ for $t = 0, 1, \ldots, k$. We have for each $i = 1, \ldots, m$ that $x_i^k = \sum_t \alpha^t \widehat{x}^t$. Then another form of dispersion is given by:

$$\sigma_j := \sum_{i=1,\ldots,m} \beta_i \sum_{t=1,\ldots,k} \alpha_t \left| Q_i \widehat{x}_{i,j}^t - z_{i,j}^k \right|$$

$$\sigma_j := \frac{1}{\sum_{i=1} [\sigma_X]_i + \epsilon} \sum_{i=1,\ldots,m} [\sigma_X + \epsilon]_i \sum_{t=1,\ldots,k} \alpha_t \left| Q_i \left(\widehat{x}_{i,j}^t - x_{i,j}^k \right) \right| + \rho_{i,j} \left| Q_i x_{i,j}^k - z_j^k \right|,$$

where $0 \leqslant \alpha_t \leqslant 1$ for $t = 1, \ldots, T_i$ and $\sum_{t=1,\ldots,T_i} \alpha_t = 1$ are the convex multipliers used to form a convex combination $x^k = \sum_{t=1,\ldots,k} \alpha_t \widehat{x}^t$.

Remark 2 We also considered using a measure of the dispersion of the dual solutions ω^k, taking

$$\sigma_j := \left\| [\omega_{i,j}]_{i=1,\ldots,m} \right\|.$$

This approach is (somewhat) sensible for when restricting the branching to variables with integrality restrictions, or when the disjunctions are otherwise strictly separated. It is otherwise an ill-defined approach, in that optimal dual solutions are typically non-unique, especially when bounding constraints are being added. Any further thoughts on this matter must take into account the set of dual optimal solutions.

3 Computational results

4 Conclusion and future work

References

1. Kiwiel, K.C.: Approximations in proximal bundle methods and decomposition of convex programs. Journal of Optimization Theory and Applications **84**(3), 529–548 (1995)