Double-Perspective Functions for Mixed-Integer Fractional Programs with Indicator Variables

Adam N. Letchford* Qiang Ni† Zhaoyu Zhong*

September 2017

Abstract

Perspective functions have long been used to convert fractional programs into convex programs. More recently, they have been used to form tight relaxations of mixed-integer nonlinear programs with so-called indicator variables. Motivated by a practical application (maximising energy efficiency in an OFDMA system), we consider problems that have a fractional objective and indicator variables simultaneously. To obtain a tight relaxation of such problems, one must consider what we call a “double-perspective” (DP) function. An analysis of DP functions leads to the derivation of a new kind of cutting planes, which we call “DP-cuts”. Computational results indicate that DP-cuts typically close a substantial proportion of the integrality gap.

Keywords: fractional programming, mixed-integer nonlinear programming, mobile wireless communications, OFDMA systems

1 Introduction

Let $y$ be a vector of $n$ continuous variables, and let $f(y)$ be a real function of $y$ that is defined over a convex domain $C \subseteq \mathbb{R}^n_+$. The perspective function of $f$ takes the form $tf(y/t)$, where $t$ is a new continuous variable, and is defined over the domain $t \in \mathbb{R}_+, y \in tC$ [7, 16]. (By convention, the perspective function takes the value zero when $t = 0$ and $y$ is the origin.) It is known that the perspective function of $f(y)$ is convex (or concave) if and only if $f(y)$ is convex (or concave) [16].

Perspective functions have a wide range of uses in, e.g., convex analysis, optimisation, statistics, signal processing and machine learning (see [2] for a recent survey). In this paper, we focus on two uses in optimisation. The first is to convert fractional programs into convex optimisation problems,
and thereby render them easier to solve [1, 18]. The second is to reformulate certain mixed-integer nonlinear programs (MINLPs), in such a way that the continuous relaxation is strengthened [4, 5]. The MINLPs in question are those with so-called indicator variables.

Recently, while studying certain problems arising in the context of mobile wireless communications, we encountered an MINLP that exhibited both of these features (i.e., a fractional objective and indicator variables) simultaneously. It turns out that, in order to obtain tight convex relaxations of such problems, one needs to study a new kind of function, which we call a double-perspective (DP) function. These functions are the topic of this paper.

Unfortunately, it turns out that the DP function of a concave function is not concave in general. To deal with this, we characterise the concave envelope of a DP function over a rectangular domain. We then derive a family of linear inequalities, which we call DP cuts, that completely describe the concave envelope. We also show how to generalise the DP cuts when there are “multiple-choice” constraints, stating that two or more indicator variables cannot take the value 1 simultaneously. Finally, we report the results of some computational experiments. It turns out that, for our particular problem, the new cuts typically close over 95% of the integrality gap.

The paper is structured as follows. The relevant literature is reviewed in Section 2. In Section 3, we present the theoretical results concerning DP functions and DP cuts. In Section 4, we consider the multiple-choice case. In Section 5, we give details of our practical application. Computational results are given in Section 6, and concluding remarks are made in Section 7.

2 Literature Review

Now we briefly review the relevant literature. We cover perspective functions and perspective cuts in Subsections 2.1 and 2.2, respectively. In Subsection 2.3, we give some background on optimisation in mobile wireless communications.

2.1 Perspective functions

As mentioned in the introduction, two particular uses of perspective functions will turn out to be of importance to us. These are as follows.

1. Consider a fractional program of the form:

   \[ \max \left\{ \frac{f(y)}{g(y)} : x \in C \right\} , \]

   where \( C \subseteq \mathbb{R}^n \) is convex, \( f(y) \) is non-negative and concave over the domain \( C \), and \( g(y) \) is positive and convex over the domain \( C \). It is
known [1,18] that such a problem can be reformulated as
\[
\max \left\{ tf(y'/t) : tg(y'/t) \leq 1, \ y' \in tC, \ t > 0 \right\},
\]
where \( t \) is a new non-negative variable representing \( 1/g(y) \), and \( y' \) is a new vector of variables representing \( y/g(y) \). The reformulated problem can often be solved efficiently, since the objective function is concave and the feasible region is convex.

2. Consider an MINLP of minimisation type, in which the cost function contains a term \( f(y) \), where \( y \) is a vector of continuous variables and \( f \) is a convex function. Suppose that the MINLP also contains an indicator variable, i.e., a binary variable \( x \) with the property that, if \( x \) takes the value zero, then all of the components of \( y \) must also take the value zero. Then the continuous relaxation of the MINLP is strengthened, while maintaining convexity, if we replace \( f(y) \) with the perspective function \( xf(y/x) \) (see, e.g., [4,5]).

2.2 Perspective cuts

One problem with perspective functions is that they are non-differentiable at the origin. Moreover, they become increasingly ill-conditioned as \( t \) approaches zero from above. This can cause algorithms for convex optimisation (or indeed convex MINLP) to run into serious numerical difficulties. To get around this, Frangioni & Gentile [4] proposed to approximate perspective functions using linear inequalities. They show that imposing a non-linear constraint of the form \( z \geq tf(y/t) \), where \( f \) is a convex function and \( t \geq 0 \), is equivalent to imposing the linear constraints

\[
z \geq \nabla f(\bar{y}) \cdot y + \left( f(\bar{y}) - \nabla f(\bar{y}) \cdot \bar{y} \right) t
\]

for all \( \bar{y} \) in the domain of \( y \). The constraints (1) are called perspective cuts. Although the perspective cuts are infinite in number, they can be very useful as cutting planes within an exact algorithm for convex MINLPs with indicator variables (see again [4,5]).

Note that the classical Kelley cuts [8] for the function \( f(y) \) take the form

\[
z \geq f(\bar{y}) + \nabla f(\bar{y}) \cdot (y - \bar{y}).
\]

Thus, when \( t \) is upper-bounded by 1, as is the case when \( t \) is an indicator variable, the perspective cuts can be viewed as strengthened Kelley cuts.

2.3 Optimisation in mobile wireless communications

In mobile wireless communications, mobile devices (such as smartphones or tablets) communicate with one another via so-called base stations. Each
base station must periodically allocate its available resources (time, power and bandwidth) in order to receive and transmit data efficiently (see, e.g., [3, 6, 15]).

These days, many base stations follow an Orthogonal Frequency-Division Multiple Access (OFDMA) architecture. In an OFDMA system, we have a set $I$ of communication channels, called subcarriers, and a set $J$ of users. Each subcarrier can be assigned to at most one user, but a user may be assigned to more than one subcarrier. If we let $p_i$ denote the power (in watts) assigned to subcarrier $i$, the classical Shannon–Hartley theorem [20] states that the maximum data rate (in bits per second) that can be transmitted from subcarrier $i$ is:

$$f_i(p_i) = B_i \log_2 (1 + p_i/N_i),$$

where $B_i$ is the bandwidth of subcarrier $i$ (in hertz), and $N_i$ is the level of noise in channel $i$ (in watts). We remark that $f_i(p_i)$ is concave over the domain $p_i \geq 0$. Moreover, if we let $S_j \subset I$ denote the set of subcarriers allocated to user $j$, the total data rate for user $j$ is just $\sum_{i \in S_j} f_i(p_i)$.

A wide variety of optimisation problems concerned with OFDMA systems have been considered, with various objectives and side-constraints (see, e.g., [9–12, 19, 21–25]). Recently, driven by environmental considerations, some authors working on OFDMA systems have focused on maximising energy efficiency, which is defined as total data rate divided by total power (e.g., [21–23, 25]). This leads immediately to a fractional objective function, which is what led us to the present paper.

3 Double Perspective Functions and Cuts

This section is concerned with double perspective (DP) functions and cuts. In Subsection 3.1, we define DP functions and point out that they are neither convex nor concave in general. In Subsection 3.2, we show how to compute concave over-estimators of DP functions. Then, in Subsection 3.3, we present the DP cuts.

3.1 DP functions

To begin, we give a formal definition of DP functions.

**Definition 1** Let $y$ be a vector of $n$ continuous variables, and let $f(y)$ be a real function of $y$ that is defined over a convex domain $C \subseteq \mathbb{R}^n_+$. The function

$$tx f \left( \frac{y}{tx} \right),$$

with domain $t, x \in \mathbb{R}_+, y \in t x C$, will be called the “double perspective” (DP) function of $f(y)$. (By convention, the DP function takes the value zero when $tx = 0$ and $y$ is the origin.)
Whereas standard perspective functions preserve convexity and/or concavity, the same is not true for DP functions.

**Lemma 1** Even if \( n = 1 \), \( f(y) \) is linear and \( C = \mathbb{R}_+ \), the DP function of \( f(y) \) may be neither convex nor concave over its domain.

**Proof.** Just let \( f(y) = 1 \) for all \( y \in \mathbb{R}_+ \). The DP function is \( tx \). Since it is an indefinite quadratic function, it is neither convex nor concave over the domain \( y, t, x \in \mathbb{R}_+ \) (or indeed over any convex domain with non-empty interior). \( \square \)

### 3.2 Concave envelope

Now suppose that \( f(y) \) is concave over the convex domain \( C \). Since the DP function is not guaranteed to be concave, it is natural to seek the strongest possible concave over-estimating function (sometimes called the concave envelope). The following theorem gives this function for the case in which the domain of \( (t, x) \) is a square.

**Theorem 1** Let \( y, C \) and \( f(y) \) be as in Definition 1, and suppose that \( f(y) \) is concave over \( C \). Also suppose that the domain of \( t \) is \([a_1, b_1]\), where \( 0 \leq a_1 < b_1 \), and the domain of \( x \) is \([a_2, b_2]\), where \( 0 \leq a_2 < b_2 \). Define the following two auxiliary functions:

\[
g^1(t, x) = a_2 t + b_1 (x - a_2) \\
g^2(t, x) = a_1 x + b_2 (t - a_1).
\]

Then the concave envelope of the DP function of \( f(y) \), over the domain \( t \in [a_1, b_1], x \in [a_2, b_2] \) and \( y \in t x C \), is:

\[
\min \left\{ g^1(t, x) f \left( \frac{y}{g^1(t, x)} \right), g^2(t, x) f \left( \frac{y}{g^2(t, x)} \right) \right\}.
\]

**Proof.** Let \( f^D(y, t, x) \) denote the DP function, and consider the hypograph of the DP function, i.e., the set of all quadruples \( (y, t, x, z) \) such that \( t \in [a_1, b_1], x \in [a_2, b_2], y \in t x C \) and \( z \leq f^D(y, t, x) \). Note that, if the point \( (y, t, x, z) \) lies on the boundary of the hypograph, then so does the point \( (\lambda y, \lambda t, x, \lambda z) \) for all \( \lambda \in \mathbb{R}_+ \) such that \( \lambda t \in [a_1, b_1] \). This implies that all extreme points of the boundary of the hypograph satisfy \( t \in \{a_1, b_1\} \). A similar argument shows they also satisfy \( x \in \{a_2, b_2\} \). Thus, the extreme points are of four types:

- Type A: \( t = a_1, x = a_2 \);
- Type B: \( t = a_1, x = b_2 \);
- Type C: \( t = b_1, x = a_2 \);
• Type D: \( t = b_1, x = b_2 \).

One can check that:

• \( g^1(t, x) \) is equal to \( tx \) at all extreme points of type A, C and D, but larger than \( tx \) at all extreme points of type B;

• \( g^2(t, x) \) is equal to \( tx \) at all extreme points of type A, B and D, but larger than \( tx \) at all extreme points of type C.

This implies that:

• The function \( g^1(t, x) f \left( \frac{y}{g^1(t, x)} \right) \) is equal to \( f^D(y, t, x) \) at all extreme points of type A, C and D, and larger than \( f^D(y, t, x) \) at all extreme points of type B.

• The function \( g^2(t, x) f \left( \frac{y}{g^2(t, x)} \right) \) is equal to \( f^D(y, t, x) \) at all extreme points of type A, B and D, and larger than \( f^D(y, t, x) \) at all extreme points of type C.

Thus, the desired concave envelope must be the minimum of those two functions.

Theorem 1 is a generalisation of the classical result of McCormick [13], which (in our notation) states that the concave envelope of the quadratic function \( tx \) over the domain \( t \in [a_1, b_1] \), \( x \in [a_2, b_2] \) is given by \( \min \{ g^1(t, x), g^2(t, x) \} \).

Remark 1 When \( x \) is an indicator variable, we have \( a_2 = 0 \) and \( b_2 = 1 \). Thus, in this case, \( g^1(x, t) \) and \( g^2(x, t) \) reduce to \( b_1x \) and \( t - a_1(1 - x) \), respectively.

### 3.3 DP cuts

The nonlinear function describing the concave envelope is non-differentiable when \( g^1(t, x) = g^2(t, x) \), i.e., when \( (t, x) \) is a convex combination of \( (a_1, a_2) \) and \( (b_1, b_2) \). Moreover, there are situations in which one might prefer to use an LP (or MILP) solver rather than an NLP (or MINLP) solver. So, following Frangioni & Gentile [4], we describe the concave envelope by linear inequalities. This is explained in the following proposition.

**Proposition 1** The hypograph of the concave envelope of a DP function is described by the linear inequalities

\[
 z \leq \nabla f(\bar{y}) y + \left( f(\bar{y}) - \nabla f(\bar{y}) \bar{y} \right) g^k(t, x)
\]

for \( \bar{y} \in C \) and \( k = 1, 2 \), where the functions \( g^k \) are as defined in Theorem 1.
Proof. We prove the result for $k = 1$. (The proof for $k = 2$ is similar.) First, note that $g^1(t,x) \in [0,b_1 b_2]$. We are therefore interested in the hypograph of the function $g^1 f(y/g^1)$ over the domain $g^1 \in [0,b_1 b_2]$, $y \in tx C$. Unfortunately, the domain of $y$ depends on $tx$ rather than on $g^1$. Recall however that $g^1$ is an over-estimator of $tx$. So it suffices to enlarge the domain of $y$ to $g^1 C$. Now we have the hypograph of a standard perspective function over a standard domain. It is described by the inequalities (3) with $k = 1$. □

We call the inequalities (3) DP cuts. Note that, when $k = 1$, the DP cuts pass through all points satisfying $a_1 \leq t \leq b_1$, $x = a_2$, $y = t a_2 \bar{y}$ and $z = t a_2 f(\frac{y}{t a_2})$. They also pass through the point $t = b_1$, $x = b_2$, $y = b_1 b_2 \bar{y}$ and $z = b_1 b_2 f(\frac{y}{b_1 b_2})$. Thus, they define maximal faces of the hypograph of the concave envelope whenever the point $(\bar{y}, f(\bar{y}))$ lies on a maximal face of the hypograph of the original function $f$. A similar argument applies when $k = 2$.

We now make some remarks about DP cuts.

**Remark 2** If we reduce the domains of $t$ and/or $x$, then the DP cuts become stronger. Thus, if one wishes to make the cuts as tight as possible, it may be worthwhile applying “domain reduction” techniques (see, e.g., [17]) to $t$ and/or $x$.

**Remark 3** Let $t^* \in [a_1,b_1]$, $x^* \in [a_2,b_2]$, $y^* \in C$ and $z^* \in \mathbb{R}$ be given. To solve the separation problem for DP cuts, it suffices to compute the quantity (2) for the given $t^*$, $x^*$ and $y^*$. If this quantity is less than $z^*$, then a DP cut is violated. The vector $\bar{y}$ yielding the cut is either $y^*/g^1(t^*,x^*)$ or $y^*/g^2(t^*,x^*)$, according to which term in (2) is the minimum.

**Remark 4** When $x$ is an indicator variable, the DP cuts reduce to:

\[ z \leq \nabla f(\bar{y}) y + \left( f(\bar{y}) - \nabla f(\bar{y}) \bar{y} \right) b_1 x \tag{4} \]
\[ z \leq \nabla f(\bar{y}) y + \left( f(\bar{y}) - \nabla f(\bar{y}) \bar{y} \right) \left( t - a_1 (1 - x) \right). \tag{5} \]

We will call constraints (4) and (5) type-1 and type-2 DP cuts, respectively. Note that the type-2 DP cuts (5) dominate the standard perspective cuts (1).

### 4 DP cuts and Multiple-Choice Constraints

It is very common in integer programming to encounter constraints which state that at most one of a set of binary variables can take a positive value. Such constraints go by various names, such as multiple-choice constraints,
clique constraints or generalised upper bounds. In this subsection, we consider the case in which each of the variables in the given set is an indicator variable.

More precisely, suppose we have:

- a non-negative continuous variable $t$;
- constants $a, b$ with $b > a \geq 0$;
- an integer $m \geq 2$ and positive integers $n_1, \ldots, n_m$;
- binary indicator variables $x_1, \ldots, x_m$;
- a vector of continuous variables, $y^j \in \mathbb{R}^{n_j}$, for $j = 1, \ldots, m$;
- continuous variables $z_1, \ldots, z_m$;
- convex domains $C_1, \ldots, C_m$;
- a concave function $f_j(y^j)$ for $j = 1, \ldots, m$.

Suppose we also know that all feasible quadruples $(t, x, y, z)$ must satisfy the following constraints:

$$
t \in [a, b] \tag{6}
$$

$$
\sum_{j=1}^{m} x_j \leq 1 \tag{7}
$$

$$
y^j \in C_j \quad (j = 1, \ldots, m) \tag{8}
$$

$$
z_j \leq t f_j(y^j/t) \quad (j = 1, \ldots, m) \tag{9}
$$

$$
x_j = 0 \Rightarrow y^j_i = 0 \quad (j = 1, \ldots, m; i = 1, \ldots, n_j). \tag{10}
$$

The following example shows that, when $m$ is large, the relaxation obtained using DP cuts (both types) can be surprisingly weak.

**Example:** Suppose that (i) $a = 0.01$ and $b = 1$; (ii) for all $j$, we have $n_j = 1$, $C_j = [0, 1]$ and $f_j(y^j) = \sqrt{y^j}$; (iii) the constraints (10) are modelled by the linear constraints $y^j_1 \leq x_j$ for all $j$; and (iv) the goal is to maximise the sum of the $z$ variables. One can check that an optimal solution is obtained by setting $t$ to 1 and exactly one $x$, $y$ and $z$ variable to 1, giving a profit of 1.

Now consider the fractional point obtained by setting $t$ to 1, all $x$ and $y$ variables to $1/m$, and all $z$ variables to $0.99/\sqrt{m}$. One can check that this point satisfies all DP cuts (and therefore all perspective cuts). The corresponding upper bound on the profit is $0.99\sqrt{m}$. Thus, the ratio between the upper bound and the optimum tends to infinity as $m$ increases. \qed

Now, for any $j \in \{1, \ldots, m\}$ and any $\bar{y}^j \in C^j$, let $||\bar{y}^j||$ denote $f_j(\bar{y}^j) - \nabla_j(\bar{y}^j) \cdot \bar{y}^j$. The following theorem presents a huge family of valid inequalities, which generalise the type-2 DP cuts (5).
Theorem 2 Let $S$ be any non-empty subset of $\{1,\ldots,m\}$. For each $j \in S$, let $\bar{y}^j$ be any point in $C^j$ such that $||\bar{y}^j|| > 0$. Then the linear inequality

$$\sum_{j \in S} \frac{z_j}{||\bar{y}^j||} \leq \sum_{j \in S} \frac{\nabla_j(\bar{y}^j) \cdot y^j}{||\bar{y}^j||} + t - a \left(1 - \sum_{j \in S} x_j\right)$$

(11)

is satisfied by all feasible quadruples $(t, x, y, z)$.

Proof. We consider two cases:

Case 1: $x_j = 0$ for all $j \in S$. This forces $y^j_k$ to be zero for all $j \in S$ and for $i = 1,\ldots,n_j$. This in turn forces $z_j$ to be zero for all $j \in S$. Thus, the inequality reduces to $t \geq a$, which is trivially valid.

Case 2: $x_j = 1$ for some $j \in S$. This forces $x_k$ to be zero for all $k \in S \setminus \{j\}$, along with the associated $y$ and $z$ variables. Thus, the inequality reduces to

$$\frac{z_j}{||\bar{y}^j||} \leq \frac{\nabla_j(\bar{y}^j) \cdot y^j}{||\bar{y}^j||} + t.$$

Multiplying this by $||\bar{y}^j||$ we obtain:

$$z_j \leq \nabla_j(\bar{y}^j) \cdot y^j + ||\bar{y}^j|| t.$$

This is equivalent to a standard perspective cut (1). □

We call the constraints (11) multiple-choice DP cuts. They reduce to type-2 DP cuts when $|S| = 1$. It turns out that they can be very useful when $m$ is large.

Example (cont): Let $S = \{1,\ldots,m\}$ and let $\bar{y}^j_1 = 1/m$ for all $j$. For all $j$, we have $f(\bar{y}^j_1) = 1/\sqrt{m}$ and $f'(\bar{y}^j_1) = ||\bar{y}^j_1|| = 1/2\sqrt{m}$. Thus, we obtain the following multiple-choice DP cut:

$$\sum_{j=1}^m 2\sqrt{m} z_j \leq \sum_{j=1}^m y^j_1 + t.$$

At the “bad” fractional point mentioned above, the left-hand side evaluates to $1.98 \ m$ and the right-hand side evaluates to $2$. Thus, the cut is violated for $m > 2$, and the amount of violation increases as $m$ increases. □

It can be shown that each multiple-choice DP cut defines a face of maximal dimension of the convex hull of feasible quadruples $(t, x, y, z)$. It can also be shown that the separation problem for the cuts can be solved in polynomial time (assuming that the functions $f_j$ and their partial derivatives can be computed in polynomial time). We omit details, for brevity.
5 Application to OFDMA Systems

We now apply the theoretical results in the last section to an optimisation problem associated with OFDMA systems. In Subsection 5.1, we define our problem formally and model it as a mixed 0-1 fractional program with indicator variables. In Subsection 5.2, we use standard perspective cuts to reformulate the problem as a semi-infinite mixed 0-1 linear program. In Subsection 5.3, we show how to strengthen the semi-infinite formulation using DP cuts.

5.1 The problem

Let $I, J, B_i, N_i$ and $f_i$ be defined as in Subsection 2.3. Let $P > 0$ denote the maximum power available, and let $\sigma \in (0, P)$ denote the system power, which is the amount of power needed by the OFDMA system regardless of actual data rates. Finally, suppose that each user $j \in J$ has a non-negative demand $d_j$. The task is to maximise the energy efficiency, subject to the constraint that the total data rate for each user $j$ is at least $d_j$.

We call this problem the fractional subcarrier and power allocation problem with rate constraints (F-SPARC). (A related problem, called the SPARC, was studied in [10]. The difference is that the objective in the SPARC was simply to maximise the total data rate.)

A natural formulation of the F-SPARC is obtained as follows. For all $i \in I$ and $j \in J$, let $x_{ij}$ be a binary variable, indicating whether user $j$ is assigned to subcarrier $i$. Also let $p_{ij}$ be a non-negative continuous variable, which represents the amount of power supplied to subcarrier $i$ (if $x_{ij} = 1$), or zero otherwise. We then have:

$$\max \quad \frac{\sum_{i \in I} \sum_{j \in J} f_i(p_{ij})}{\sigma + \sum_{i \in I} \sum_{j \in J} p_{ij}} \quad (12)$$

s.t.

$$\sum_{j \in J} x_{ij} \leq 1 \quad (\forall i \in I) \quad (14)$$

$$\sum_{i \in I} f_i(p_{ij}) \geq d_j \quad (\forall j \in J) \quad (15)$$

$$p_{ij} \leq Px_{ij} \quad (\forall i \in I, j \in J) \quad (16)$$

$$p_{ij} \in \mathbb{R}_+ \quad (\forall i \in I, j \in J) \quad (17)$$

$$x_{ij} \in \{0, 1\} \quad (\forall i \in I, j \in J). \quad (18)$$

The objective function (12) represents the total data rate divided by the total power used (including system power). The constraint (13) ensures that the total power used does not exceed the amount available. Constraints (14) ensure that each subcarrier is assigned to at most one user. Constraints (15) ensure that the user demands are met. Constraints (16)–(18) are self-explanatory.
The problem (12)–(18) is a mixed 0-1 fractional program. Moreover, the $x$ variables are clearly indicator variables, since setting any $x$ variable to zero forces the corresponding $p$ variable to zero.

5.2 Reformulation

We now reformulate the problem to make it easier to solve. This is done in three steps.

The first step is to convert the fractional objective function into a concave function. To do this, we use the first “trick” mentioned in Subsection 2.1. Let $t$ be a non-negative continuous variable, representing $1/(\sigma + \sum_{i \in I} \sum_{j \in J} p_{ij})$, i.e., the reciprocal of the total power used. Also, for all $i$ and $j$, let $\tilde{p}_{ij}$ be a non-negative continuous variable, representing $tp_{ij}$. The problem is then equivalent to the following:

$$\max \sum_{i \in I} \sum_{j \in J} tf_i(\tilde{p}_{ij}/t)$$

s.t. (14), (18)

$$\sigma t + \sum_{i \in I} \sum_{j \in J} \tilde{p}_{ij} = 1$$

$$1/P \leq t \leq 1/\sigma$$

$$\sum_{i \in I} f_i(\tilde{p}_{ij}/t) \geq d_j \quad (\forall j \in J)$$

$$0 \leq \tilde{p}_{ij} \leq x_{ij} \quad (\forall i \in I, j \in J).$$

The objective function (19) is now concave, and all constraints are linear, apart from (22), which are convex.

The second step is the following. For $i \in I$ and $j \in J$, define a new variable, say $z_{ij}$, representing the quantity $tf_i(\tilde{p}_{ij}/t)$. The problem then becomes:

$$\max \sum_{i \in I} \sum_{j \in J} z_{ij}$$

s.t. (14), (18), (20), (21), (23)

$$\sum_{i \in I} z_{ij} \geq d_j t \quad (\forall j \in J)$$

$$z_{ij} \leq f_i(\tilde{p}_{ij}/t) \quad (\forall i \in I, j \in J)$$

$$z_{ij} \in \mathbb{R}^+ \quad (\forall i \in I, j \in J).$$

Now, all of the nonlinearity has been “concentrated” in the (convex) constraints (25).

Finally, following Frangioni & Gentile [4], we replace the constraints (25) with the following (standard) perspective cuts:

$$z_{ij} \leq f'_i(\bar{p}) \tilde{p}_{ij} + (f_i(\bar{p}) - f'_i(\bar{p}) \bar{p}) t \quad (\forall i \in I, j \in J, \bar{p} \in [0, P - \sigma]).$$

Since the perspective cuts are linear, but infinite in number, we have formulated the F-SPARC as a semi-infinite mixed 0-1 linear program. This means that it can be solved (to arbitrary fixed accuracy) with an LP-based (or MILP-based) approach.
5.3 DP cuts

We now use the results in Section 3 to strengthen our semi-infinite formulation of the F-SPARC.

We start by reducing the domain of $t$. Let $D = \sum_{j \in J} d_j$ denote the total user demand. We compute the following lower bound on the total amount of power used:

$$P_{\text{min}} = \sigma + \min \left\{ \sum_{i \in I} \frac{f_i(p_i)}{P_{\text{min}}} : \sum_{i \in I} f_i(p_i) \geq D, p_i \geq 0 (i \in I) \right\}.$$ 

The upper bound on $t$ in (21) can then be reduced from $1/\sigma$ to $1/P_{\text{min}}$.

Now, let $i$ and $j$ be fixed, and consider the constraints (18), (23) and (25). It is clear that all feasible triples $(x_{ij}, \tilde{p}_{ij}, r_{ij})$ satisfy

$$z_{ij} \leq t x_{ij} f_i \left( \frac{\tilde{p}_{ij}}{tx_{ij}} \right),$$

where, as usual, the convention is that the function on the right-hand side takes the value zero when $\tilde{p}_{ij} = tx_{ij} = 0$. The function is clearly a DP function, in which $t$ plays the role of $t_1$ and $x_{ij}$ plays the role of $t_2$. The only tricky part is that the natural domain of $\tilde{p}_{ij}$ is $[0, 1]$, which is not equal to $tx_{ij}C$ for some convex domain $C$. Despite this complication, we have the following result:

**Proposition 2** For all $i \in I$, $j \in J$ and $\tilde{p} \in [0, P - \sigma]$, the following DP cuts are valid for the F-SPARC:

$$z_{ij} \leq f_i'(\tilde{p}) \tilde{p}_{ij} + \left( f_i(\tilde{p}) - f_i'(\tilde{p}) \tilde{p} \right) x_{ij} \quad (28)$$

$$z_{ij} \leq f_i'(\tilde{p}) \tilde{p}_{ij} + \left( f_i(\tilde{p}) - f_i'(\tilde{p}) \tilde{p} \right) \left( t - \frac{1 - x_{ij}}{P} \right). \quad (29)$$

**Proof.** We already know that $t$ must lie in the interval $[1/P, 1/P_{\text{min}}]$. Also, $p_{ij} \leq P - \sigma$, or, equivalently, $\tilde{p}_{ij} \in [0, P - \sigma]$. In fact, given that $\tilde{p}_{ij}$ must be zero when $x_{ij}$ is zero, we can conclude that $\tilde{p}_{ij} \in tx_{ij}[0, P - \sigma]$. We can now apply Remark 4. 

Finally, observe that (14) is a multiple-choice constraint involving indicator variables. Thus, Theorem 2 is applicable. The resulting multiple-choice DP cuts take the form:

$$\sum_{j \in S} \frac{z_{ij}}{f_i(\tilde{p}_{ij}) - f_i'(\tilde{p}_{ij})\tilde{p}_{ij}} \leq \sum_{j \in S} \frac{f_i'(\tilde{p}_{ij})p_{ij}}{f_i(\tilde{p}_{ij}) - f_i'(\tilde{p}_{ij})\tilde{p}_{ij}} + t - \frac{1 - \sum_{j \in S} x_{ij}}{P}.$$
For simplicity, we concentrate on the special case obtained when (a) \( S = J \) and (b) \( \bar{p}_{ij} \) takes the same (positive) value for all \( j \in J \). In this case, the cuts take the following simpler form:

\[
\sum_{j \in J} z_{ij} \leq \sum_{j \in J} f_i'(\bar{p}) p_{ij} + \left( f_i(\bar{p}) - f_i'(\bar{p}) \bar{p} \right) t. \tag{30}
\]

They are valid for all \( i \in I \) and all \( \bar{p} \in [0, P - \sigma] \).

## 6 Computational Experiments

To explore the potential of DP cuts, we performed some computational experiments on F-SPARC instances.

### 6.1 Experimental setting

To compute upper bounds, we use a standard LP-based cutting-plane algorithm, based on the semi-infinite mixed 0-1 LP formulation described in Subsection 5.2. To compute exact solutions (within some specified tolerance), we use a standard “LP/NLP-based branch-and-bound” algorithm [14], based on the mixed 0-1 convex program formulation described in the same subsection. These algorithms enable us to compare the performance of the different families of cuts described in the previous sections.

The algorithms were coded in Julia v0.5 and run on a virtual machine cluster with 16 CPUs (ranging from Sandy Bridge to Haswell architectures) and 16GB of RAM, under Ubuntu 16.04.1 LTS. The program calls on MOSEK 7.1 (with default settings) to solve NLPs, and on the mixed-integer solver from the CPLEX 12.6.3 Callable Library (again with default settings) to solve LP relaxations.

### 6.2 Test instances

To construct our test instances, we used the procedure described in [10], which is designed to produce instances typical of a small base station. These instances have \(|I| = 72\), \(|J| \in \{4, 6, 8\}\) and \(P\) set to 36 watts. The noise powers \(N_i\) are random numbers distributed uniformly in \((0, 10^{-11})\), and the bandwidths \(B_i\) are all set to 1.25MHz. The user demands follow a lognormal distribution. Let \(D\) be the sum of the user demands, and let \(M\) be the maximum possible data rate of the system. As in [10], we call the quantity \(D/M\) the demand ratio (DR) of the given instance. The user demands are scaled so that the DR takes values in \(\{0.75, 0.8, 0.85, 0.9, 0.95, 0.98, 0.99\}\). (The closer the DR is to 1, the harder the instance tends to be.) For each combination of \(|J|\) and DR, we constructed 500 random instances. This makes \(3 \times 7 \times 500 = 11,500\) instances in total.
Table 1: Average percentage gap with perspective cuts

| Demand Ratio | 0.75 | 0.80 | 0.85 | 0.90 | 0.95 | 0.98 | 0.99 |
|--------------|------|------|------|------|------|------|------|
| | 240  | 240  | 242  | 243  | 248  | 250  | 250  |
| 4 | 383  | 383  | 386  | 393  | 399  | 403  | 404  |
| 6 | 520  | 520  | 525  | 535  | 546  | 552  | 553  |

Table 2: Average percentage gap with type-2 DP cuts

| Demand Ratio | 0.75 | 0.80 | 0.85 | 0.90 | 0.95 | 0.98 | 0.99 |
|--------------|------|------|------|------|------|------|------|
| | 99.3 | 98.0 | 92.6 | 77.6 | 50.2 | 24.2 | 12.9 |
| 4 | 162  | 161  | 151  | 126  | 82.4 | 39.9 | 21.5 |
| 6 | 219  | 223  | 205  | 174  | 116  | 56.1 | 30.2 |

6.3 Results

We consider four versions of the cutting-plane algorithm. In the first, the standard perspective cuts (27) are used. In the second, the stronger type-2 DP cuts (29) are used instead. In the third, the type-1 DP cuts (28) are used as well. Finally, in the fourth, the special multiple-choice DP cuts (30) are used as well.

Tables 1–4 show, for each set of 50 instances, the average gap between the upper bound and the optimum, expressed as a percentage of the optimum, under the four algorithmic settings.

From Table 1 we see that classical perspective cuts lead to extremely poor upper bounds for this problem, with gaps of over 200% in all cases. The gaps get worse as the number of users $|J|$ increases, but they do not vary with the DR. Table 2 shows that type-2 DP cuts do a lot better in all cases, but the gaps are still very large. The gaps again get worse as $|J|$ increases, but, interestingly, they decrease as the DR increases.

Table 3 reveals that the type-1 and type-2 DP cuts in combination do
Table 4: Average percentage gap with multiple-choice DP cuts added

| Demand Ratio | 0.75 | 0.80 | 0.85 | 0.90 | 0.95 | 0.98 | 0.99 |
|--------------|------|------|------|------|------|------|------|
| 4            | 1.13 | 0.68 | 0.42 | 0.31 | 0.44 | 0.76 | 1.04 |
| 6            | 1.18 | 0.85 | 0.42 | 0.44 | 0.64 | 0.90 | 1.36 |
| 8            | 1.14 | 0.76 | 0.46 | 0.61 | 0.96 | 1.44 | 1.55 |

a lot better, with all gaps being below 25%. Interestingly, the gaps no longer get worse as $|J|$ increases, but they still decrease as the DR increases. Finally, Table 4 shows that the addition of multiple-choice DP cuts leads to a dramatic improvement, with all gaps now being less than 1.6%. The gaps do not get worse as $|J|$ increases but, strangely, they decrease and then increase again as the DR increases.

Further experiments, not reported here, revealed a clear “hierarchy” of cuts, with multiple-choice DP cuts being a lot more useful than type-1 cuts, which in turn are a lot more useful than type-2 DP cuts. We also found the multiple-choice DP cuts essential within the exact algorithm. As for running times, in all cases the time taken by the cutting-plane algorithm was negligible in comparison with the time taken by the exact algorithm.

7 Concluding Remarks

Perspective reformulations and cuts are an invaluable tool for both fractional programming and MINLP with indicator variables. We have shown that, when one is dealing with a mixed-integer fractional program with indicator variables, one needs to use “double perspective” reformulations and cuts in order to obtain bounds that are useful within an exact solution algorithm. We believe that extensions of perspective reformulations and cuts to other classes of problems would be a valuable topic for future research.

As for our specific application, to optimisation in OFDMA systems, we plan to look next at stochastic dynamic variants of the problem, in which users arrive and depart at random over time.

References

[1] A. Charnes & W. Cooper (1962) Programming with linear fractional functionals. *Nav. Res. Log. Quart.*, 9, 181–186.

[2] P.L. Combettes (2017) Perspective functions: properties, constructions, and examples. *Set-Valued Var. Anal.*, to appear.
[3] K. Fazel & S. Kaiser (2008) *Multi-Carrier and Spread Spectrum Systems (2nd edn.).* Chichester: Wiley.

[4] A. Frangioni & C. Gentile (2006) Perspective cuts for a class of convex 0-1 mixed-integer programs. *Math. Program.*, 106, 225–236.

[5] O. Günlük & J. Linderoth (2011) Perspective reformulation and applications. In: J. Lee & S. Leyffer (eds.) *Mixed Integer Nonlinear Programming*, pp. 61–89. Berlin: Springer.

[6] S. Haykin (1994) *Communication Systems*. New York: Wiley.

[7] J.-B. Hiriart-Urruty & C. Lémarechal (1993) *Convex Analysis and Minimization Algorithms*. New York: Springer US.

[8] J.E. Kelley, Jr. (1960) The cutting-plane method for solving convex programs. *SIAM J.*, 8, 703–712.

[9] K. Kim, Y. Han & S.L. Kim (2005) Joint subcarrier and power allocation in uplink OFDMA systems. *IEEE Commun. Lett.*, 9, 526–528.

[10] A.N. Letchford, Q. Ni & Z. Zhong (2017) An exact algorithm for a resource allocation problem in mobile wireless communications. *Comput. Optim. Appl.*, to appear.

[11] Y.F. Liu & Y.H. Dai (2014) On the complexity of joint subcarrier and power allocation for multi-user OFDMA systems. *IEEE Trans. Signal Process.*, 62, 583–596.

[12] Z.Q. Luo & S. Zhang (2008) Dynamic spectrum management: complexity and duality. *IEEE J. Sel. Topics Signal Process.*, 2, 57–73.

[13] G.P. McCormick (1976) Computability of global solutions to factorable nonconvex programs. Part I: convex underestimating problems. *Math. Program.*, 10, 147–175.

[14] I. Quesada & I.E. Grossmann (1992) An LP/NLP based branch and bound algorithm for convex MINLP optimization problems. *Comput. & Chem. Eng.*, 16, 937–947.

[15] M. Resende & P. Pardalos (eds.) (2007) *Handbook of Optimization in Telecommunications*. New York: Springer US.

[16] R.T. Rockafellar (1970) *Convex Analysis*. Princeton: Princeton University Press.

[17] H.S. Ryoo & N.V. Sahinidis (1996) A branch-and-reduce approach to global optimization. *J. Glob. Optim.*, 8, 107–138.
[18] S. Schaible (1983) Fractional programming. Zeitschrift für Oper. Res., 27, 39–54.

[19] K. Seong, M. Mohseni & J. Cioffi (2006) Optimal resource allocation for OFDMA downlink systems. In Proc. 2006 IEEE Inter. Symp. Inform. Th., pp. 1394–1398.

[20] C. Shannon (1949) Communication in the presence of noise. Proc. IRE, 1, 10–21.

[21] T. Ting, S.F. Chien, X.S. Yang & S. Lee (2014) Analysis of quality-of-service aware orthogonal frequency division multiple access system considering energy efficiency. IET Commun., 8, 1947–1954.

[22] X. Xiao, X. Tao & J. Lu (2013) QoS-aware energy-efficient radio resource scheduling in multi-user OFDMA systems. IEEE Commun. Lett., 17, 75–78.

[23] C. Xiong, G. Li, S. Zhang, Y. Chen & S. Xu (2011) Energy-and spectral-efficiency tradeoff in downlink OFDMA networks. IEEE Trans. Wirel. Commun., 10, 3874–3886.

[24] W. Yu & R. Lui (2006) Dual methods for nonconvex spectrum optimization of multicarrier systems. IEEE Trans. Commun., 54, 1310–1322.

[25] C.C. Zarakovitis C & Q. Ni (2016) Maximising energy efficiency in multi-user multi-carrier broadband wireless systems: convex relaxation and global optimisation techniques. IEEE Trans. Veh. Technol., 65, 5275–5286.