A Natural Gas Cash-Out Problem: A Bilevel Programming Framework and a Penalty Function Method

Vyacheslav V. Kalashnikov
Department of Mathematics
Instituto Tecnológico y de Estudios Superiores de Monterrey
Ave. Eugenio Garza Sada 2501 Sur
Monterrey, NL 64849, México
E-mail: *kalash@itesm.mx*

Roger Z. Ríos-Mercado
Graduate Program in Systems Engineering
Universidad Autónoma de Nuevo León
AP 111 – F, Cd. Universitaria
San Nicolás de los Garza, NL 66450, México
E-mail: *roger@uanl.mx*

April 2003
Revised July 2004

Abstract

One of the many complex problems that arise from the transmission and marketing of natural gas is when a shipper draws a contract with a pipeline company to deliver a certain amount of gas among several points. What is actually delivered is often different from the amount that had been originally agreed upon. This phenomenon is called *an imbalance*. When an imbalance occurs, the pipeline penalizes the shipper by imposing a cash-out penalty policy. Since this penalty is a function of the operating daily imbalances, an important decision-making problem for the shippers is how to carry out their daily imbalances so as to minimize their incurred penalty.

In this paper, we introduce the problem of minimizing the cash-out penalty costs from the point of view of a natural gas shipping party. We present a mixed integer bilevel linear programming model and discuss its underlying assumptions. To solve it efficiently, we reformulate it as a standard mathematical program and describe a penalty-function algorithm functions for its solution. The algorithm is well-founded and its convergence is proved. Results of numerical experiments support the algorithm's robustness providing a valuable solution technique for this very important and complex problem in the natural gas market.

Keywords: mixed integer bilevel programming, penalty function, natural gas market, cash-out penalty policy

1 Introduction

In many decision processes there is a hierarchy of decision makers, and decisions are made at different levels in this hierarchy. One way to handle such hierarchies is to focus on one level and include other levels' behaviors as assumptions. Multilevel programming is the research area that focuses on the whole hierarchy structure. In terms of modeling, the constraint domain associated with a multilevel programming problem is implicitly determined by a series of optimization problems which must be solved in a predetermined sequence. If only two levels are considered, we have one leader (associated with the upper level) and one follower (associated with the lower one), and call the problem as a *bilevel programming problem*. If the constraints at both levels are all linear, we have a bilevel linear programming problem (BLP). In [1, 11], one can find both the essential fundamentals of the multilevel optimization and its applications to solution of real systems.

The field of multilevel optimization today is a well-known and important research field. Hierarchical structures can be found in diverse scientific disciplines including environmental studies, classification theory, databases, network design, transportation, game theory, and economics; and new applications (like the above described gas cash-out problem) are constantly being introduced. This is, in turn, positive for the development of new theory and efficient algorithms.

A particular case of the bilevel programming problem is presented by the following mixed-integer model arising from the problem of minimization of cash-out penalty costs of a natural gas shipping company. In countries like the United States of America, for instance, the natural gas industry has been going through a deregulation process since the mid-1980s leading to significant market changes. Now the decision making procedure of gas buying, selling, storing, transporting, etc., is immersed into a very complex world in which producers, pipelines (transporters), and brokers, all play quite important roles in the chain. This chain becomes even more complex if we take into account the network of pipelines transporting gas and oil throughout the Latin America, which reaches Canada at one edge and Argentina at the other edge, passing through the USA, Mexico, the Central and South Americas. Regarding Mexico, it is very important to study and understand this complex phenomenon and moreover, to develop the supporting techniques that permits one to make decisions well-grounded when one needs to interact with the foreign counter-partners in the processes of buying/selling/transporting the gas.

The problem in question arises when a shipper draws a contract with a pipeline company to deliver a certain amount of gas among several points. What is actually delivered may be more or less of the amount that had been originally agreed upon (this phenomenon is called *an imbalance*). When an imbalance occurs, the pipeline penalizes the shipper by imposing a cash-out penalty policy. As this penalty is a function of the operating daily imbalances, an important problem for the shippers is how to carry out their daily imbalances so as to minimize their incurred penalty. To the best of our knowledge, this problem has not appeared in the literature before.

One of the main goals of this paper is to introduce a mathematical model for this problem and to present a detailed discussion of the underlying assumptions. The problem is modeled as a mixed-integer bilevel linear programming problem (BLP) where the shipper is the leader (upper level) and the pipeline represents the follower (lower level). Even the simplest version of a multilevel optimization problem, a linear problem with two levels, is known to be very hard to solve. Mixed-integer BLP possess even a higher degree of difficulty as the typical concepts for fathoming in traditional branch-and-bound algorithms for mixed-integer programming (MIP) cannot be directly applied to mixed-integer BLP.

Another goal of this work is to develop an efficient method for addressing the above-described problem of minimizing the cash-out penalties of the shipper (the leader) subject to the lower level problem reflecting the aims of the pipeline (the follower). We consider a hierarchical system where a leader incorporates into its strategy the reaction of the follower to its decision. The follower's reaction is quite generally represented as the solution set to a monotone variational inequality (e.g., [1, 3, 7]). For the solution of this nonconvex mathematical program, a penalty approach, based on the formulation of the lower level variational inequality as a mathematical program and using a linear function in the leader's (gas shipper) objective, was first attempted. As the numerical experiments showed, such a problem had a quite predictable solution: the gas shipper is forced to increase its positive imbalances day after day, to minimize its penalties by the pipeline. Thus, somewhat implausible pseudo-stores of the natural gas may be created which could lead to unjustified storage costs. Therefore, we introduce a sort of self-retaining penalties for the shipper in its objective function. As we demonstrate, that does not spoil the convergence properties of the inexact penalty function method.

The paper is organized as follows. In Section 2, the problem is introduced, including a discussion of the modeling assumptions and the model. The penalty function method, including its convergence results, is described in Section 3. Section 4 describes a direct algorithm to be compared with the penalty function method. Section 5 contains a summary of results of numerical experiments, including the details of a small illustrative test instance. This is followed by our conclusions in Section 6.

2 Problem Description

Assume that a shipper has entered into a contract (with other customers) to deliver a given amount of natural gas from a receipt to a delivery meter in a given time frame. (From now on, we treat "natural gas" and "gas" as synonyms). The shipper must stipulate title transfer agreements with the meter operators and a transportation agreement with the pipeline. Under such agreements, the shipper nominates a daily amount of gas to be injected by the receipt meter operator into the pipeline and to be withdrawn by the delivery meter operator from the pipeline. The pipeline transports the gas from the receipt meter to the delivery meter.

Due to the nature of the natural gas industry, what is actually transported is inevitably different from what is nominated. Such a difference constitutes an imbalance. There exist operational and transportation imbalances. The first type of imbalance refers to differences between nominated and actual flows, while the latter involves differences between net receipts (receipts minus fuel) and deliveries. While pipelines allow for small imbalances, they issue penalties for higher (both operational and transportation) imbalances to the other parties. In the following, we assume that the shippers are held responsible for imbalance penalties, and we analyze the cash-out penalties associated with operational imbalances.

On the shipper side, an operational imbalance can be either positive or negative. A positive (negative) imbalance arises when the shipper leaves (takes) gas in (from) the pipeline. Alternatively, a positive (negative) imbalance means that the actual flow is smaller (greater) than the nominated amount of gas. A positive [negative] end-of-the-month imbalance implies a cash transaction from the pipeline (shipper) to the shipper (pipeline). Cash-out prices are set in a way that whenever a shipper sells [buys] gas to (from) the pipeline, he does that at a very low (high) price. The relation between cash-out price and imbalance position depends non-linearly on the average, maximum and minimum gas spot price for the past month.

While it is true that the imbalances are due to external phenomena, the shipper may play a roll in the imbalances decisions as well. So from this perspective the total imbalance has two components, one due to the shipper and one due to the external phenomena. Since one of our goals is to investigate how the shipper should attempt to carry the imbalances and how it would fare at the end of the process we are assuming this imbalance decision is due totally to the shipper. This is a first step towards understanding the nature of this complex decision process. Modeling the external causes (as random variables) would require incorporating a stochastic programming approach, which is beyond the scope of this work at this point.

Shippers daily nominate gas flows taking into account the constraints deriving from their buy/sell activity, their contractual constraints, and future market opportunities. The gas price is one of the major factors affecting their decisions. In the absence of cash-out provisions, historically shippers would take out high cost gas in the winter from the pipeline (causing negative imbalances), and pay the transporter back with low cost gas in the summer. This corresponds to a speculative behavior by the shippers, whereby imbalances are created and managed as pseudo-storage in order to take advantage of movements in the gas price. Cash-out penalties were designed in order to avoid such pricing arbitrages. In the framework below, shippers are concerned with minimizing the cash-out penalties.

Now, depending on the actual values of the penalty cost parameters for both negative and positive imbalances, it can happen that, in some cases, the total cash-out penalty may turn out to be negative. The interpreation of this is that the shipper receives a cash transaction (credit) from the shipper. Thus, another issue we investigate in this work is precisely the circumstances under which this may occur and how this may affect the decision-making process.

2.1 Notation

As stated previously, the decision making process for the shipper (leader) is to determine how to carry out its daily imbalances so as to minimize the penalty that will be imposed by the pipeline (follower). The following notation is used to describe the model.

Indices and sets

i, j, k	zone pool indices; $i, j, k \in J = \{1, 2, \ldots, P\}$;
t	time index; $t \in T = \{1, 2, \ldots, N\}$.

Parameters

x_{ti}^L, x_{ti}^U	bounds on daily imbalances at (end of) day t in zone i; $t \in T$, $i \in J$;
x_t^L, x_t^U	bounds on total daily imbalances at (end of) day t; $t \in T$;
s_{ti}^L, s_{ti}^U	bounds on balance swings during day t in zone i; $t \in T, i \in J$;
e_{ij}	percentage of fuel retained for moving one dekatherm (dt) of gas from zone i to j; $i, j \in J$;
f_{ij}	transportation charge for moving one dt of gas from zone i to j; $i, j \in J$, $i < j$;
b_{ij}	backward haul credit for moving one dt of gas from zone j to i; $i, j \in J$, $i < j$;
x_{0j}	initial imbalance (start of day 1) in zone j; $j \in J$;
r_i, δ_i	the imbalance penalization parameters, $i \in J$.

Decision variables (first level)

x_{ti} imbalance at (end of) day t in zone i; $t \in T$, $i \in J$;

s_{ti} imbalance swing during day t in zone i; $t \in T$, $i \in J$;

Decision variables (second level)

y_i final imbalance at zone i; $i \in J$;

u_{ij} forward haul volume moved from zone i to j; $i, j \in J$, $i < j$;

v_{ij} backward haul volume moved from zone j to i; $i, j \in J$, $i < j$;

z total cash-out cost for shipper.

Auxiliary variables

q binary variable equal to 1 (0) if final imbalances are non-negative (non-positive). In the special case when all the final imbalances are zeros, we accept $q = 1$.

2.2 Model

Here we provide the set of constraints involved in both the upper and lower levels of the problem.

Upper Level Model:

Objective: Shipper's cost.

$$\text{Minimize} \quad h_1(x, s, y, u, v, z) = z \tag{1a}$$

subject to

$$x_{ti}^L \leq x_{ti} \leq x_{ti}^U \quad t \in T,\ i \in J \tag{1b}$$

$$s_{ti}^L \leq s_{ti} \leq s_{ti}^U \quad t \in T,\ i \in J \tag{1c}$$

$$x_t^L \leq \sum_{i \in J} x_{ti} \leq x_t^U \quad t \in T \tag{1d}$$

$$x_{ti} = x_{t-1,i} + s_{ti} \quad t \in T,\ i \in J \tag{1e}$$

Lower Level Model:

Objective: The penalty is determined by minimizing the amount of cash transactions. What happens in many cases in practice, is that both shipper and pipeline agree in a policy that represents a compromise between the two. Thus, rather than maximizing the cash-out penalty for the shipper (which would be the incentive of the pipeline alone), it is agreed to minimize deviations from zero. By acting this way, the shipper (pipeline) will have a protection against negative (positive) values of the shipper's total penalty (represented by (2i) below). Hence, the objective is given by

$$\text{minimize} \qquad h_2(x,s,y,u,v,z) = |z|, \tag{2a}$$

subject to the constraints below.

Balance constraints: This constraint identifies the relationship between the imbalance at day $N = |T|$, forward and backward haul volumes, retained fuel,and final imbalance at zone j:

$$y_j = x_{N,j} + \sum_{i:i<j}(1-e_{ij})u_{ij} + \sum_{k:k>j} v_{jk} - \sum_{k:k>j} u_{jk} - \sum_{i:i<j} v_{ij},\ j \in J; \tag{2b}$$

Gas conservation: This constraint ensures no gas loss occurs. Although it follows directly from (2b) after summation with respect to all $j \in J$, we keep it on to make the problem clearer to non-technical users.

$$\sum_{i\in J} y_i + \sum_{(i,j):i<j} e_{ij}u_{ij} = \sum_{i\in J} x_{N,i}. \tag{2c}$$

Note that $\sum_{(ij)} e_{ij}u_{ij} \geq 0$, hence $\sum_i y_i \leq \sum_i x_{N,i}$.

Zone upper bounds: This constraint prevents cyclic movements of gas. It simply states that, from any given zone, we cannot move more than any initial positive imbalance.

$$\sum_{j:i<j} u_{ij} + \sum_{k:k<i} v_{ki} \leq \max\{0, x_{N,i}\}, \quad i \in J. \tag{2d}$$

Forward haul upper bounds: These bounds prevent positive-to-positive and negative forward movement of imbalances.

$$u_{ij} \leq \begin{cases} x_{N,i} & \text{if } x_{N,i} > 0 \text{ and } x_{N,j} < 0; \\ 0 & \text{otherwise.} \end{cases} \tag{2e}$$

Backward haul upper bounds: These bounds prevent positive-to-positive and negative backward movement of imbalances.

$$v_{ij} \leq \begin{cases} x_{N,j} & \text{if } x_{N,j} > 0 \text{ and } x_{N,i} < 0; \\ 0 & \text{otherwise.} \end{cases} \tag{2f}$$

Bounds on final imbalances: These bounds ensure that all final imbalances have the "right" sign, i.e. an imbalance must not change sign.

$$\min\{0, x_{N,i}\} \leq y_i \leq \max\{0, x_{N,i}\}, \quad i \in J. \tag{2g}$$

Sign of final imbalances: This is a business rule that states that final imbalances for all zones must have the same "sign" (i.e. all non-positive or non-negative); that means that the imbalances must not change sign from zone to zone:

$$-M(1-q) \leq y_i \leq Mq, \qquad i \in J, \tag{2h}$$

where M is a large number and q is a binary 0-1 variable.

Shipper's cash-out cost: This equation represents the penalty cost from the shipper's point of view:

$$z = \sum_{i \in J} \left[\delta_i \left(y_i\right)_+^2 - r_i y_i \right] - \sum_{(i,j):i<j} b_{ij} v_{ij} + \sum_{(i,j):i<j} f_{ij}(1 - e_{ij}) u_{ij}. \tag{2i}$$

where as usual, $(y_i)_+ = \max\{0, y_i\}$. The second and third terms reflect the haul credit for moving gas backwards and the transportation charge for moving gas forward, respectively. The first term reflects the cash-out policy as follows. When, in a given zone i, there is a negative imbalance ($y_i < 0$), there is a cash transaction from the shipper to the pipeline (equal to $|r_i y_i|$ because $r_i > 0$). In the other hand, if a postive imbalance occurs ($y_i > 0$), then the pipeline buys the gas from the shipper at a price of $-\left[\delta_i (y_i)^2 - r_i y_i\right]$ (i.e., a "negative" cost for the shipper). Note that the cash-out penalty for the shipper is higher than buying price for the pipeline. The quadratic term is an implicit penalty dictated to the shipper to prevent unjustified shipper's pseudo-storages.

Variable types:

$$y_i, z \qquad \text{free}; \tag{2j}$$

$$u_{ij}, v_{ij} \geq 0; \tag{2k}$$

$$q \in \{0, 1\}. \tag{2l}$$

In the following section, we describe our proposed solution technique, which is motivated by the penalty function method by Marcotte and Zhu [10].

3 A Penalty Method

When the last day imbalance $\sum_i x_{N,i}$ is either non-positive or positive and large enough, then the bilevel problem (1a)–(2l) solutions sets will not change if we move the integrality constraint (2h) from the lower level to the upper one. So given this is the only integer variable, it suffices to solve the following two bilevel programs:

$$\begin{aligned} \min \quad & h_1(x, s, y, u, v, z) = z \\ \text{subject to} \quad & \text{(1b)–(1e)} \\ & \begin{aligned} \min \quad & h_3(x, s, y, u, v, z) = z^2 \\ \text{subject to} \quad & \text{(2b)–(2k)} \\ & q = 0 \end{aligned} \end{aligned}$$

and

$$
\begin{array}{lll}
\min & h_1(x,s,y,u,v,z) = z & \\
\text{subject to} & \text{(1b)–(1e)} & \\
& \min \quad h_3(x,s,y,u,v,z) = z^2 & \\
& \text{subject to} \quad \text{(2b)–(2k)} & \\
& \qquad\qquad q = 1 &
\end{array}
$$

These are bilevel hierarchical systems in the Euclidean space where the upper level decision maker (hereafter the leader) controls a vector of variables $w_1 = (x,s) \in R^{NP} \times R^{NP}$, and the lower level (hereafter the follower) controls a vector of continuous variables $w_2 = (y,u,v) \in R^P \times R^{P(P-1)/2} \times R^{P(P-1)/2}$, and a binary variable $q = \beta$. The leader makes its decision first, taking into account the reaction of the follower to its course of action. If we introduce the function

$$
F(y,u,v) = \sum_{i \in J} \left[r_i y_i - \delta_i (y_i)_+^2 \right] + \sum_{(i,j):i<j} b_{ij} v_{ij} - \sum_{(i,j):i<j} f_{ij}(1 - e_{ij}) u_{ij}
$$

and the set $W_2^\beta = W_2^\beta(x,s) = \{w_2^\beta = (y,u,v) \ : \text{(2b)-(2h),(2j)-(2k)}, \ q = \beta\}$, then the reaction $w_2^\beta = w_2^\beta(x,s) = (y,u,v)$ of the follower to the leaders's decision $w_1 = (x,s)$ is a solution to an equilibrium problem represented by the following variational inequality:

$$
\langle F(w_2^\beta) \nabla F(w_2^\beta), w_2 - w_2^\beta \rangle \geq 0, \quad \text{for all } w_2 \in W_2^\beta(x,s), \tag{3}
$$

with $\nabla F(\cdot) = \left\{ [r_i - 2\delta_i (y_i)_+]_{i \in J}, (-f_{ij}(1 - e_{ij}))_{(i,j):i<j}, (b_{ij})_{(i,j):i<j} \right\}^T$. Note that this function $F(y,u,v)$ is the negative of the original shipper's penalty function, so we then obtain the following two equivalent generalized bilevel programs, or GBLP (β), for $\beta = 0,1$, in maximization form:

$$
\text{GBLP}(\beta): \quad F^{\beta,*} = \max_{(x,s) \in W_1, (y,u,v) \in W_2^\beta(x,s)} F(y,u,v),
$$

subject to (3), where the leader is implicitly restricted to the set W_1 of (x,s)-vectors such that the lower level constraint set $W_2(x,s)^\beta$ is nonempty. The optimal solution can be obtained by solving both problems GBLP(0) and GBLP(1).

Now, let us formulate the lower level variational inequality as a parameterized equation related to the duality gap of the lower level problem. We then use this "gap" function as a penalty term for the upper level problem. Since the gap function characterizing the lower level is non-negative over the feasible domain, the penalty term reduces to a very simple form. This is the gap function we associate with the lower level variational inequality (3):

$$
G_\alpha^\beta(w_1, w_2^\beta) = \max_{w_2 \in W_2^\beta(x,s)} \phi(w_1, w_2^\beta, w_2), \tag{4}
$$

where

$$\phi(w_1, w_2^\beta, w_2) = -\langle F(w_2^\beta)\nabla F(w_2^\beta), w_2 - w_2^\beta\rangle - \frac{1}{2}\alpha\|w_2 - w_2^\beta\|^2,$$

and α is a non-negative number. The gap function has been used to construct descent methods for solving variational inequalities. We refer in particular to Marcotte [8] and Marcotte and Dussault [9] for the linear gap function ($\alpha = 0$), and to Fukushima [3] for the quadratic, differentiable function ($\alpha > 0$).

The function ϕ is concave in w_2 (strongly concave if α is positive). Also, since G_α^β is non-negative over $S_\beta = \bigcup_{w_1 \in W_1} w_1 \times W_2^\beta(w_1)$, and $G_\alpha^\beta(w_1, w_2^\beta) = 0$ if and only if w_2^β is a solution to the lower level variational inequality parameterized by (x, s), the variational inequality (3) can be rewritten as the following nonlinear equation

$$G_\alpha^\beta(w_1, w_2^\beta) = \max_{w_2 \in W_2^\beta(x,s)} \phi(w_1, w_2^\beta, w_2) = \phi(w_1, w_2^\beta, p_\alpha(w_1, w_2^\beta)) = 0,$$

where $p_\alpha(w_1, w_2^\beta)$ is any solution of (4). Finally, this leads to a reformulation of the GBLP(β) as a standard mathematical programming problem

$$\text{PR1}(\beta): \qquad \max_{w_1=(x,s)\in W_1, w_2^\beta=(y,u,v)\in W_2^\beta(x,s)} F(y, u, v),$$

$$\text{subject to} \quad G_\alpha^\beta(w_1, w_2^\beta) = 0.$$

3.1 Inexact Penalization Algorithms

In order to implement the penalty approach, we approximate PR1(β) by the penalized problem (see Marcotte and Zhu [10]):

$$\text{PR2}(\beta): \qquad \min_{(w_1, w_2^\beta)\in C_\beta} Q_\alpha(w_1, w_2^\beta, \mu),$$

where $Q_\alpha(w_1, w_2^\beta, \mu) = -F(w_2^\beta) + \mu G_\alpha^\beta(w_1, w_2^\beta)$, μ is a positive number, and the subset C_β in R^{P^2+2NP} is given by

$$C_\beta = \left\{(w_1, w_2^\beta) \;:\; (1\text{b}) - (1\text{e}), (2\text{b}) - (2\text{h}), (2\text{j}) - (2\text{k}), q = \beta\right\}.$$

Lemma 1 *The subset C_β is compact for each $\beta = 0, 1$.*

Proof: It is easy to see that the constraints (1b)–(1c) imply that the values of the variables (x, s) are bounded, whereas the values of y belong to the cube $[-M, M]^P$ (cf. (2h)). The values of variables u and v are bounded since they satisfy (2e), (2f) and (2k), and values of $x_{N,j}$ are also bounded. Therefore, the subset C_β is bounded, and being closed it is compact. ■

Therefore, problem PR2(β) is nonlinear, and its objective function is continuously differentiable when α is positive. For each value of the weight μ we denote by $\left(w_1(\mu), w_2^\beta(\mu)\right)$ a global optimal solution of PR2(β) which always exists due to the compactness of the feasible set C_β and continuity of the objective function Q_α. An inexact penalty function algorithm is obtained by specifying a sequence of increasing (unbounded) positive weights $\{\mu_k\}$ and the associated sequence of iterates $\left\{\left(w_1(\mu_k), w_2^\beta(\mu_k)\right)\right\}$.

Below we obtain the convergence results for the penalty function algorithm, making use of the corresponding techniques from [10]. First we prove the following lemma.

Lemma 2 *Let $\left\{\left(w_1^k = w_1(\mu_k), w_2^k = w_2^\beta(\mu_k)\right)\right\}_{k=1}^{\infty}$ be a sequence of iterates generated by a penalty function algorithm based upon the function Q_α from program PR2(β). Let $\left(w_1^*, w_2^{\beta,*}\right)$ be an optimal solution of GBPL(β), and $F^{\beta,*} = F\left(w_2^{\beta,*}\right)$. Then the following inequalities are valid:*

$$
\begin{aligned}
Q_\alpha\left(w_1^k, w_2^k, \mu_k\right) &\le Q_\alpha\left(w_1^{k+1}, w_2^{k+1}, \mu_{k+1}\right) \qquad (5)\\
G_\alpha^\beta\left(w_1^k, w_2^k\right) &\ge G_\alpha^\beta\left(w_1^{k+1}, w_2^{k+1}\right) \qquad (6)\\
F\left(w_2^k\right) &\ge F\left(w_2^{k+1}\right) \qquad (7)\\
-Q_\alpha\left(w_1^k, w_2^k, \mu_k\right) &\ge F^{\beta,*}. \qquad (8)
\end{aligned}
$$

Proof: Since $w_1^k, w_1^{k+1} \in W_1$, $w_2^k \in W_2^\beta\left(w_1^k\right)$ and $w_2^{k+1} \in W_2^\beta\left(w_1^{k+1}\right)$, we obtain

$$-F\left(w_2^{k+1}\right) + \mu_{k+1} G_\alpha^\beta\left(w_1^{k+1}, w_2^{k+1}\right) \le -F\left(w_2^k\right) + \mu_{k+1} G_\alpha^\beta\left(w_1^k, w_2^k\right),$$

and

$$
\begin{aligned}
Q_\alpha\left(w_1^{k+1}, w_2^{k+1}, \mu_{k+1}\right) &= -F\left(w_2^{k+1}\right) + \mu_{k+1} G_\alpha^\beta\left(w_1^{k+1}, w_2^{k+1}\right)\\
\ge -F\left(w_2^{k+1}\right) + \mu_k G_\alpha^\beta\left(w_1^{k+1}, w_2^{k+1}\right) &\ge -F\left(w_2^k\right) + \mu_k G_\alpha^\beta\left(w_1^k, w_2^k\right)
\end{aligned}
$$

(since $\left(w_1^k, w_2^k\right)$ is an optimal solution)

$$= Q_\alpha\left(w_1^k, w_2^k, \mu_k\right),$$

which proves (5). Combining the first and the last inequalities obtained above, we deduce that

$$\left(\mu_{k+1} - \mu_k\right) G_\alpha^\beta\left(w_1^{k+1}, w_2^{k+1}\right) \le \left(\mu_{k+1} - \mu_k\right) G_\alpha^\beta\left(w_1^k, w_2^k\right),$$

which implies (6) by dividing both parts by the positive number $(\mu_{k+1} - \mu_k)$. In its turn, (6) together with the inequality

$$-F\left(w_2^{k+1}\right) + \mu_k G_\alpha^\beta\left(w_1^{k+1}, w_2^{k+1}\right) \ge -F\left(w_2^k\right) + \mu_k G_\alpha^\beta\left(w_1^k, w_2^k\right),$$

give us (7). At last, observe that

$$-F^{\beta,*} = -F\left(w_2^{\beta,*}\right) + \mu_k G_\alpha^\beta\left(w_1^*, w_2^{\beta,*}\right)$$

$$\geq -F\left(w_2^k\right) + \mu_k G_\alpha^\beta\left(w_1^k, w_2^k\right) = Q_\alpha\left(w_1^k, w_2^k, \mu_k\right) \geq -F\left(w_2^k\right),$$

which implies (8) and completes the proof. ■

Theorem 1 *Let* $\left\{\left(w_1^k, w_2^k\right)\right\}_{k=1}^{\infty}$ *be a sequence of iterates generated by a penalty function algorithm based upon the function* Q_α *from program PR2(*β*).*

Then every limit point of the sequence $\left\{\left(w_1^k, w_2^k\right)\right\}_{k=1}^{\infty}$ *is a solution of the bilevel program PR1(*β*).*

Proof: As the subset C_β is compact, there is a convergent subsequence $\left\{\left(w_1^k, w_2^k\right)\right\}_{k\in K}$ that converges to a limit point $\left(\bar{w}_1^\beta, \bar{w}_2^\beta\right)$. From the continuity of F we can conclude that

$$\lim_{k\in K, k\to\infty} F\left(w_2^k\right) = F\left(\bar{w}_2^\beta\right). \tag{9}$$

Lemma 2 implies that

$$-Q_\alpha^* = \lim_{k\in K, k\to\infty}\left[-Q_\alpha\left(w_1^k, w_2^k, \mu_k\right)\right] \geq F^{\beta,*}. \tag{10}$$

Subtracting (9) from (10) gives us

$$-\lim_{k\in K, k\to\infty} \mu_k G_\alpha^\beta\left(w_1^k, w_2^k\right) = -Q_\alpha^* - F\left(\bar{w}_2^\beta\right) \geq F^{\beta,*} - F\left(\bar{w}_2^\beta\right) \geq 0.$$

Since G_α^β assumes non-negative values, the last inequality implies

$$\lim_{k\in K, k\to\infty} G_\alpha^\beta\left(w_1^k, w_2^k\right) = 0,$$

which combined with the continuity of G_α^β leads to the equality $G_\alpha^\beta\left(\bar{w}_1^\beta, \bar{w}_2^\beta\right) = 0$, that means that $\left(\bar{w}_1^\beta, \bar{w}_2^\beta\right)$ is a feasible point for problem P1(β). Moreover, inequality (8) from Lemma 2 allows us to write

$$F\left(\bar{w}_2^\beta\right) = \lim_{k\in K, k\to\infty} F\left(w_2^k\right) \geq F^{\beta,*},$$

which proves that $\left(\bar{w}_1^\beta, \bar{w}_2^\beta\right)$ is an optimal solution to problem P1(β), hence to problem GBLP(β) as well. The proof is complete. ■

Remark 1. After having obtained two optimal solutions $\left(\bar{w}_1^\beta, \bar{w}_2^\beta\right)$ to the problems PR1(β), $\beta = 0, 1$, we can find the optimal solution of the initial problem (1a)–(1e), (2a)–(2l) as follows:

$$\left(w_1^*, w_2^*\right) = \begin{cases} \left(w_1^0, w_2^0\right), & \text{if } F\left(w_2^0\right) \geq F\left(w_2^1\right), \\ \left(w_1^1, w_2^1\right), & \text{otherwise.} \end{cases}$$

4 A Direct Algorithm

To evaluate and verify robustness of our inexact penalty function algorithm, we implemented a direct search algorithm for the gas cash-out problem and compared its performance with the penalty function scheme applied to the same test problems. This direct algorithm takes into account the problem's structure which allows one to split the computation into two almost independent stages: (i) search for the pipeline's optimal response to the gas shipper's strategy and, based upon that, the gas shipper's optimal distribution of the last day imbalances; (ii) verification of feasibility of thus obtained optimal last day imbalances distribution (which involves solution of a quadratic programming problem). This idea of splitting leads to a computational scheme that may be valid for a wide class of bilevel problems. Another important feature of the direct algorithm is transforming the mixed integer programming problem to a standard nonlinear programming problem, to get rid of the inconveniences of the presence of integer variables.

Now we describe the direct algorithm's steps.

Data Input. Input parameter values:

(i) natural values $P > 0, N > 0$ and x_{0j}, $j \in J = \{1, 2, \ldots, P\}$;

(ii) $x_{tj}^L, x_{tj}^U, s_{tj}^L, s_{tj}^U, x_t^L, x_t^U$, $j \in J$, $t \in T = \{1, 2, \ldots, N\}$;

(iii) $e_{ij} > 0, b_{ij} > 0, f_{ij} > 0$, $i, j \in J, i < j$;

(iv) $r_j > 0, \delta_j > 0$, $j \in J$;

(v) $\varepsilon > 0$.

Step 0 (Initialization). Set $k := 0$ and solve the standard non-smooth quadratic programming problem (1a)-(1d), (2b), (2i), with $u_{ij} = v_{ij} = 0$ for all $i, j \in J$. If the original problem is solvable, so is the latter one. If the obtained optimal vector (x, s) has the last day imbalance components all non-positive or all non-negative, i.e., if either $x_{N,j} \geq 0$, $j \in J$, or $x_{N,j} \leq 0$, $j \in J$, then the vector (w_1, w_2) with $w_1 = (x, s)$, $w_2 = (x, 0, 0)$ is obviously an optimal solution of the original problem as well. Otherwise, (i.e., if the components $x_{N,j}$ have different signs), set $x_N^{(k)} = \left(x_{N,1}^{(k)}, \ldots, x_{N,P}^{(k)}\right)^T := (x_{N,1}, \ldots, x_{N,P})^T$ and GOTO Step 1.

Step 1 (Solving Bi-Level Problem). Apply the Nelder-Mead simplex algorithm [6] to solve the upper level problem:

$$z = z(x, s, y, u, v) \to \min, \tag{11}$$

subject to:

$$s_{tj}^L \leq s_{tj} \leq s_{tj}^U, \qquad j \in J, t \in T; \tag{12}$$

$$x_{1j}^L \leq x_{0j} + s_{1j} \leq x_{1j}^U, \quad j \in J; \tag{13}$$

$$x_1^L \leq \sum_{j \in J} (x_{0j} + s_{1j}) \leq x_1^U; \tag{14}$$

$$\vdots$$

$$x_{N-1,j}^L \leq x_{0j} + \sum_{t=1}^{N-1} s_{tj} \leq x_{N-1,j}^U, \; j \in J; \tag{15}$$

$$x_{N-1}^L \leq \sum_{j \in J} \left(x_{0j} + \sum_{t=1}^{N-1} s_{tj} \right) \leq x_{N-1}^U; \tag{16}$$

$$x_{N,j}^L \leq x_{0j} + \sum_{t \in T} s_{tj} = \tilde{x}_{N,j}^{(k+1)} \leq x_{N,j}^U, \; j \in J; \tag{17}$$

$$x_N^L \leq \sum_{j \in J} \left(x_{0j} + \sum_{t \in T} s_{tj} \right) \leq x_N^U. \tag{18}$$

Here the value of the objective function $z = z(x, s, y, u, v)$ at each step of the Nelder-Mead simplex algorithm is determined as a solution of the lower level problem:

$$|z| \equiv |F(w_2)| \equiv |F(y, u, v)| \to \min_{w_2 \in W_2}, \tag{19}$$

where $W_2 = \{w_2 = (y, u, v) \text{ satisfying } (2b) - (2j)\}$. Here, to avoid dealing with the absolute value in the objective function of the lower level problem, we introduce an additional variable d and minimize d instead of $|z|$ after having added two auxiliary constraints $-d \leq z \leq d$.

5 Computational Work

To test both the inexact penalty function algorithm and the direct algorithm, we made use of GAMS [2], version 2.25 in a Sun Ultra 10 running Solaris 7. GAMS, a widely known algebraic modeling software with interface to several optimization algorithms, was applied to some test problems of various sizes, the maximal dimensions being $N = 30$ and $P = 9$. These test problems were built by using real-world data provided by industry.

The performance of both algorithms was stable and successful as they found good approximate solutions with the tolerance $\varepsilon = 10^{-3}$ and $\varepsilon = 10^{-4}$ quite fast (less than 10 seconds of computing time). However, in contrast to the penalty function algorithm that always needed no more than two restarts with the penalty parameter values $10^2 \leq \mu \leq 10^6$, the direct algorithm often jammed, and more than 4-5 restarts were required to reach a reasonable approximation. The jamming is probably explained by the standard non-convexity of the follower's optimal responses set in the bilevel problems. A comparison table is given below, while a small test example is described in details in Section 5.1. Table 1 shows the results when the penalty and direct algorithm are applied to four problem instances.

		Penalty Algorithm		Direct Algorithm	
P	N	CPU time (sec)	Number of restarts	CPU time (sec)	Number of restarts
4	2	0.006	0	0.020	3
9	7	0.020	1	0.050	4
6	30	0.030	2	0.060	5
9	30	0.045	2	0.072	5

Table 1. Test results

While performing the numerical experiments we noticed that the shipper's revenue function z defined by (2i) with $\delta_j = 0$, $j \in J$, leads not only to a very fast computation (5-6 times faster than for the non-linear case) but also to a very simple strategy by the shipper at the upper level: he tends to increase his imbalance day after day trying to maximize his (positive) last day imbalance. As this strategy leads to an unjustified pseudo-storage of gas, we had to introduce the (non-smooth) quadratic terms $\delta_j\,(y_j)_+^2$ with $\delta_j > 0$, $j \in J$, into the shipper's revenue function (2i) to prevent this phenomenon. This remark is illustrated in the example below.

5.1 An Illustrative Example

Here we present an example of a small test model with four pools and a two days account (i.e., $N = 2$ and $P = 4$) processed by the algorithms. The input data are given by the Tables 2 to 8.

Pools j	Initial imbalance x_{0j}	Penalty factor r_j	Coefficient δ_j
Pool 1	-10.0	10.0	0.1
Pool 2	-4.0	8.0	0.2
Pool 3	3.0	6.0	0.3
Pool 4	6.0	4.0	0.4

Table 2. Initial imbalances and coefficients

Time intervals t	Lower bounds x_t^L	Upper bounds x_t^U
Day 1	-11.0	1.0
Day 2	-17.0	7.0

Table 3. Lower and upper bounds on total daily imblances

	Pool 1	Pool 2	Pool 3	Pool 4
Day 1	-15.0	-9.0	0.0	1.0
Day 2	-20.0	-13.0	-5.0	-4.0

Table 4. Lower limits of imbalances at pool j on day t

	Pool 1	Pool 2	Pool 3	Pool 4
Day 1	-5.0	1.0	6.0	10.0
Day 2	0.0	5.0	10.0	15.0

Table 5. Upper limits of imbalances at pool j on day t

	Pool 1	Pool 2	Pool 3	Pool 4
Pool 1		2.0	4.0	4.0
Pool 2			2.0	2.0
Pool 3				1.0
Pool 4				

Table 6. Transportation costs between pools f_{ij}

	Pool 1	Pool 2	Pool 3	Pool 4
Pool 1		0.1	0.2	0.2
Pool 2			0.1	0.1
Pool 3				0.05
Pool 4				

Table 7. Percentage of gas retained as fuel when transported between pools e_{ij}

	Pool 1	Pool 2	Pool 3	Pool 4
Pool 1		4.0	2.0	2.0
Pool 2			4.0	4.0
Pool 3				2.0
Pool 4				

Table 8. Bones for shipper for backward haul b_{ij}

The lower and upper bounds for imbalance swings are: $s_{ti}^L = -3, s_{ti}^U = 3$ for all $t \in T$, $i \in J$. After the preliminary step starting with the initial imbalances $x_0 = (-10.0, -4.0, 3.0, 6.0)^T$, the last day imbalances $x_N = (-4.0, 2.0, 9.0, 5.0)^T$ with the shipper's objective function value $z = 14.9$ have been obtained. As the latter vector does not have all its entries of the same sign, we proceed with the steps of the algorithm. To compare the numerical results for two cases: a) without penalties for gas pseudo-storage and b) with the penalties for pseudo-storage, we put first $\delta_j = 0,\ j = 1, \ldots, 4$. In this case, with the final tolerance $\varepsilon = 0.001$, Step 1 obtains the accepted approximate of the last day imbalances as $x_N = (-4.001, -9.998, 9.0, 11.999)^T$ with the shipper's penalty cost value $z = -75.994$. The pipeline's optimal response to this shipper's strategy is the final imbalance $y = (0.0, 0.0, 0.0, 7.0)^T$ with the same pipeline's objective function value $z = -75.994$. All the forward hauls u_{ij} are zero, whereas the optimal backward hauls are: $v_{13} = 0.0$; $v_{14} = 4.001$; $v_{23} = 9.0$; $v_{24} = 0.998$, with all other $v_{ij} = 0$. As the last day imbalance $x_N = (-4.001, -9.998, 9.0, 11.999)^T$ satisfies the feasibility test, that is, it is obtained from the initial imbalances $x_0 = (-10.0, -4.0, 3.0, 6.0)^T$ by the daily swings indicated in Table 9 below, then the gas cash-out problem is solved, the approximate solution being presented in Table 10. We note that the total amount of the gas pseudo-storage in this case is 7.002.

	Pool 1	Pool 2	Pool 3	Pool 4
Day 1	3.000	-3.000	3.000	3.000
Day 2	2.999	-2.998	3.000	2.999

Table 9. Optimal swings (s_{ti}) of imbalance at pool j on day t when $\delta_j = 0, j = 1, \cdots, 4$

	Pool 1	Pool 2	Pool 3	Pool 4
Day 1	-7.000	-7.000	6.000	9.000
Day 2	-4.001	-9.998	9.000	11.999

Table 10. Optimal imbalances (x_{ti}) at pool j on day t when $\delta_j = 0, j = 1, \ldots, 4$

However, if we use the nonlinear values of penalty coefficients δ_j from Table 1, we come to another approximate optimal solution $x_N = (-9.8889, -10.0, 9.0, 12.0)$ with the lower absolute value of the shipper's cash-out cost of $z = -64.2222$ and with the much lower pseudo- storage amount of 1.1111. Table 11 and Table 12 provide the daily swings and the approximate solution, respectively, for this nonlinear case.

	Pool 1	Pool 2	Pool 3	Pool 4
Day 1	0.0000	-3.000	3.000	3.000
Day 2	0.1111	-3.000	3.000	3.000

Table 11. Optimal swings (s_{ti}) of imbalance at pool j on day t when δ_j are taken from Table 1

	Pool 1	Pool 2	Pool 3	Pool 4
Day 1	-10.000	-7.000	6.000	9.000
Day 2	-9.8889	-10.000	9.000	12.000

Table 12. Optimal imbalances (x_{ti}) at pool j on day t when δ_j are taken from Table 1

6 Closing Remarks

In this paper, we introduced the problem of minimizing the cash-out penalties from the point of view of a natural gas shipper. The problem is modeled as a mixed integer bilevel non-linear program. To solve it efficiently, we reformulate it as a standard mathematical programming problem and describe a penalty function algorithm for its solution. The algorithm is well-founded and its convergence is proved. We derived the convergence results for the inexact penalty function algorithm, when there is no finite value of the parameter μ for which the solution of the penalized problem agrees with the optimal solution of GBLP(β).

Another contribution was to show that actually, the initial problem can be splitted into two almost independent subproblems. The solution set of the first subproblem consists of all feasible imbalances on the last day of the term, each of which, in its turn, can be chosen as an initial iterate for the second subproblem, the bilevel leader-follower one. The first subproblem can be solved by general methods of non-linear programming. This motivated the implementation of a direct algorithm. In our computational experience, we have observed that the performance of both approaches was stable when tested over a set of instances provided by industry.

For our future research we are working on developing results concerning the exactness property of the penalization scheme based on the "classical" gap function introduced by Hearn [5] in an optimization setting and corresponding to the choice $\alpha = 0$ in P1(β), in both cases of separable and non-separable constraints. Results of this kind were obtained in [10] but only for linear constraints. Since not all the constraints in our initial problem are linear, those results of [10] are not applicable directly to our problem.

We recall that in this work we are assuming that imbalances' decisions are due to the shipper. As we stated in Section 2, external phenomena may also play a roll on these imbalances. Thus, a natural extension of this work is to investigate the effect that external causes could have in the decision-making process. This will require modeling this phenomena as a random process, and could lead to an important application of stochastic programming.

For many years, the only problems of (limited) interest in industry where the ones involving the optimization of fuel consumption in natural gas pipeline networks (see [12] for a survey), that is, focused on the transportation side. However, since the industry was deregulated in the US, the nature and complexity of the decision-making problems has grown significantly. So, there is a tremendous opportunity for both academic research and employing operations research techniques for tackling many of the complex problems arising in the marketing side of the natural gas industry.

Acknowledgements: We are very grateful to Nicola Secomandi and Viroj Buraparate, from PROS Revenue Management, a Houston-based firm that provides scientific support on revenue management problems. They both had a major role on working and developing the first version of the mathematical model for this application. Their assistance, suggestions, and discussion is highly appreciated. The research of Vyacheslav Kalashnikov was financially supported by Universidad Autónoma de Nuevo León under its Scientific and Technological Research Support Program (UANL-PAICYT award CA562-01), and that of Roger Ríos-Mercado was financially supported by the Mexican National Council for Science and Technology (CONACyT award J33187–A) and by Universidad Autónoma de Nuevo León (UANL-PAICYT awards CA555–01 and CA763–02).

We also would like to thank two anonymous referees whose valuable comments and remarks have helped us improve the quality of the paper.

References

[1] J. F. Bard. *Practical Bilevel Programming: Algorithms and Applications*. Kluwer, Dordrecht, 1998.

[2] A. Brooke, D. Kendrick, and A. Meeraus. *GAMS: A User's Guide*. The Scientific Press, San Francisco, 1992.

[3] M. Fukushima. Equivalent differentiable optimization problems and descent methods for asymmetric variational inequality problems. *Mathematical Programming*, 53:99–110, 1992.

[4] M. Gendreau, P. Marcotte, and G. Savard. A hybrid tabu-ascent algorithm for the linear bilevel programming problem. *Journal of Global Optimization*, 9:1–14, 1996.

[5] D. W. Hearn. The gap function of a convex program. *Operations Research Letters*, 1:67–71, 1981.

[6] D. M. Himmelblau. *Applied Nonlinear Programming.* McGraw-Hill, New York, 1972.

[7] V. V. Kalashnikov and N. I. Kalashnikova. Solving two-level variational inequalities. *Journal of Global Optimization*, 8:289–294, 1996.

[8] P. Marcotte. A new algorithm for solving variational inequalities with application to the traffic assignment problem. *Mathematical Programming*, 33:339–351, 1985.

[9] P. Marcotte and J.-P. Dussault. A sequential linear programming algorithm for solving monotone variational inequalities. *SIAM Journal on Control Optimization*, 27:1260–1278, 1989.

[10] P. Marcotte and D. L. Zhu. Exact and inexact penalty methods for the generalized bilevel programming problem. *Mathematical Programming*, 74:141–157, 1996.

[11] A. Migdalas, P. Pardalos, and P. Värbrand, editors. *Multilevel Optimization: Algorithms and Applications.* Kluwer, Dordrecht, 1998.

[12] R. Z. Ríos-Mercado. Natural Gas Pipeline Optimization. In P. M. Pardalos and M. G. C. Resende, editors, *Handbook of Applied Optimization*, chapter 18.8.3, pages 813–825. Oxford University Press, New York, 2002.