Uncertainty Feature Optimization: an Implicit Paradigm for Problems with Noisy Data

Niklaus Eggenberg, Transport and Mobility Laboratory, EPFL

Matteo Salani, Transport and Mobility Laboratory, EPFL

Michel Bierlaire, Transport and Mobility Laboratory, EPFL

Conference paper STRC 2008

STRC

8th Swiss Transport Research Conference

Monte Verità / Ascona, October 15-17, 2008

Uncertainty Feature Optimization: an Implicit Paradigm for Problems with Noisy Data[1]

Niklaus Eggenberg
TRANSP-OR, EPFL
Lausanne

+41216932432
niklaus.eggenberg@epfl.ch

Matteo Salani
TRANSP-OR, EPFL
Lausanne

+41216932432
matteo.salani@epfl.ch

Michel Bierlaire
TRANSP-OR, EPFL
Lausanne

+41216932537
michel.bierlaire@epfl.ch

Abstract

Optimization problems due to noisy data are usually solved using stochastic programming or robust optimization approaches. Both requiring the explicit characterization of an *uncertainty set* that models the nature of the noise. Such approaches tightly depend on the modeling of the uncertainty set.

In this paper, we introduce a framework that implicitly models the uncertain data. We define the general concept of *Uncertainty Features* (UF) which are structural properties of a solution. We show how to formulate an uncertain problem using the *Uncertainty Feature Optimization* (UFO) framework as a multi-objective problem. We prove that stochastic programming and robust optimization are particular cases of the UFO framework. We present computational results for the Multi-Dimensional Knapsack Problem (MDKP) and discuss the application to the airline scheduling problem. Computational results show a stability of the solutions in variations of the noise's nature, unlike methods based on an explicit uncertainty set.

Keywords

Robust Optimization – Uncertainty

[1]This work was first presented at the CTW 2008 conference in Gargnano, Italy, in May 2008. See `http://ctw08.dti.unimi.it`

1 Introduction

Nowadays, Operations Research tools are widely used to optimize real world problems (without loss of generality, we suppose we only deal with minimization problems). The major difficulty the modelers are faced with is the noisy nature of the data most of the problems are due to. As shown by Birge and Louveaux, 1997, Herroelen and Leus, 2005 and Sahinidis, 2004 and references therein, a greedy solution neglecting the uncertain nature of the data leads to unstable solutions: either we loose feasibility or the average performance of the solution in reality is poor. The ability of a solution to remain feasible with respect to data changes is called the *robustness* of the solution. In the case the solution is not robust, we define the *recoverability* of the solution as the average performance of the solution including both original costs and the costs incurred when modifying the solution to retrieve feasibility, which are called the *recovery costs*. The operation of repairing a solution is the *recovery* algorithm.

The existing methods for solving problems due to noisy data are divided in two distinct classes: the *reactive* and the *proactive* methods. Reactive methods are also called *on-line* algorithms. Such algorithms re-compute a new solution each time the revealed data requires the solution to be updated (mainly when feasibility is lost). The proactive approaches compute the solution before any data is revealed. Such methods require some predictions on the future data outcome to perform better than the greedy deterministic approach. The possible set of outcomes is modeled by an *uncertainty set*. We distinguish two different proactive methods according to way the uncertainty set is defined: the *expected-mean* and the *worst-case* methods. On the one hand, expected-mean methods aim at finding the solution performing best in average and thus require an explicit characterization of all possible data outcomes. We suppose that expected-mean methods require a probabilistic distribution on the uncertainty set. On the other hand, worst-case approaches are conservative methods seeking the stability of the solution even in the worst possible configuration. The requirements of such methods are the characterization of the worst case for any solution. The modeling of the uncertainty sets is the key to a proactive method's efficiency. Unfortunately, it is a difficult task and, as we show in this paper, errors in the estimation of the uncertainty set might have dramatic consequences, making the solution even worse than the greedy deterministic solution.

The concept of *Uncertainty Feature Optimization* (UFO) is different from on-line, expected-mean and worst-case approaches: it aims at finding a proactive solution but without the explicit characterization of an uncertainty set. The fact that the problem is due to noisy data is considered *implicitly* using *Uncertainty Features* (UF), which are structural properties of the solution improving its robustness or recoverability.

The initial motivation for UFO comes from the airline scheduling problem, see Kohl et al., 2004 for a general survey. Airline scheduling requires a proactive method, because of the early publication deadlines of the schedule. In addition, due to many unpredictable influencing factors, modeling the uncertainty set is a difficult task. Several works in the literature attempt to model an uncertainty set, see for example Lan et al., 2006, Chebalov and Klabjan, 2004, Policella, 2004. The main conclusion of the works using robust approaches are that the obtained solutions exhibit a particular property such as the number of plane crossings (Klabjan et al., 2002, Bian et al., 2004), a reduced length of plane rotations Rosenberger et al., 2004 or increased idle time Al-Fawzana and Haouari, 2005. Remarkably, this also holds for models aiming at an increase of the solution's recoverability. The stochastic model with recourse of Yen and Brige, 2006 addresses the crew scheduling problem. Their solutions exhibit pairings with a reduced number of plane changes. The UFs corresponding to these structural properties are the maximization of idle time, number of plane crossings and number of plane rotations satisfying the crew pairing constraints.

The contributions of this work are that UFO is a general framework for optimization under uncertainty, where the implicit handling does not require the modeling effort of characterizing an uncertainty set. The consequence is a gain of stability of the solutions' behavior when the noise's nature changes. Furthermore, we prove that UFO is a generalization of existing proactive methods for a particular choice of UFs.

The structure of the paper is as follows: section 2 summarizes the literature on methods for optimization under uncertainty and discusses their benefits and drawbacks. Section 3 presents the Uncertainty Feature Optimization (UFO) framework and section 4 demonstrates how to derive existing proactive methods from the UFO framework. In section 5, we show practical examples of UFO: we present simulation results on the Multi-Dimensional Knapsack Problem (MDKP) and we discuss the application of UFO to airline scheduling in section (6). Finally, section 7 concludes the paper with some future research issues.

2 Optimization under Uncertainty

For general surveys on optimization under uncertainty we refer to Herroelen and Leus, 2005 and Sahinidis, 2004 and references therein. Their main conclusions are that uncertainty should not be neglected when solving an optimization problem due to uncertainty.

Notice that for the remaining part of the paper, we suppose, without loss of generality, that we are faced with a problem where the objective is to minimize a cost function. Moreover, we decompose the survey into three parts corresponding to the three classes we identify: reactive, stochastic and worst-case approaches.

Reactive Algorithms Reactive algorithms are also known as *on-line* algorithms. The concept of an on-line algorithm is based on the wait-and-see strategy: there is, in general, no baseline solution computed a priori, the solution being built iteratively according to a decision process based on the revealed data; decisions are (potentially) taken each time new information is gathered. The clear benefit is that, if it exists the policy eventually provides a globally feasible solution once all data is revealed. There are however several drawback. The first is the lack of stability of the solution, since it depends on the data realization. Moreover, the approach does not allow for deriving a baseline schedule, and suffers from the real time requirements, implying that the decision process must be determined in real time, excluding thereby sophisticated decision processes for large-scale problems. Finally, it is difficult to derive a measure of performance for such algorithms: the *competitivity ratio* is an a posteriori measure comparing the obtained solution against the optimal solution with known data. In real world applications, on-line algorithms perform at acceptable ranges in terms of optimality deviation, but one can usually find scenarios for which the algorithms perform poorly. For a survey on reactive algorithms, we refer to Albers, 2003 and Grötschel et al., 2002.

Stochastic Programing Stochastic optimization is a widely studied field and a standard approach to deal with uncertainty, see Birge and Louveaux, 1997. The main objective is to optimize the *expected* value of the objective over the whole set of uncertain data, i.e. the uncertainty set: this implies the knowledge of a probabilistic measure on the *uncertainty set*. The clear benefit of the approach is that the obtained solution is the one that performs best in average: if the solution is carried out many times under the same conditions, then the average cost of the solution tends to the expected cost. The drawback is that stochastic programming (with or without recourse) needs an explicit uncertainty set provided with a probabilistic measure. In addition, the approach requires the evaluation of the solution on the whole uncertainty set in order to determine it's expected cost, which is, in general, computationally hard. Finally, the computed expected cost is only an estimator on the possible solution's outcome: one cannot guarantee the real cost matches the expected cost: the expected cost is a good indicator only when the obtained solution is implemented many times under the same conditions, as then, the average cost converges almost surely to the expected cost.

In *stochastic optimization with recourse* or *multi-stage stochastic optimization* (Birge and Louveaux, 1997, Kall and Wallace, 1994, Herroelen and Leus, 2005), a *recourse* strategy that refines the reaction to take when information on a scenario is revealed is considered. The major advantages of this approach is that we consider two levels of information, namely the a priory knowledge and the possible data outcomes along time: the solution thus provides the action to take in case of significant information gain. The benefit of the approach is that the two decisional levels lead to the best expected solution, *including recourse costs*, which is a much better approximation on the real cost than the only expected cost (without the recourse costs). The drawbacks are again the needs of the probabilistic uncertainty set. Moreover, the computational complexity is increased by several orders, since one needs to solve the recourse problem for each scenario in order to get only one solution's expected recourse cost, and to consider all scenarios in order to determine the one minimizing the total expected cost (the sum of first level and recourse costs). In fact, for large scale problems where evaluation is not realistic, the method needs either a closed form for any solution's recourse costs or to formulate the recourse problem as an underlying problem of the same complexity than the original problem. Note that even in the case of a discrete uncertainty set for which the recourse problem can be expressed as a set of m linear functions given a solution, we get a problem with at least $n \times m$ constraints, where n is the number of decisional stages at which recourse has to be taken.

Wort-Case Based Approaches The class of worst-case based approaches is mainly composed of methods leading to *robust* solutions, i.e. solutions that are feasible even in the worst possible scenario. Many works use robust optimization; Soyster, 1973 was the first to introduce a formal approach of robustness, and Bertsimas and Sim, 2004 and Ben-Tal and Nemirovski, 2001 give a more formal framework for different classes of problems. The main advantage of a robust solution is that, if the uncertainty set is exhaustive and a robust solution exists, then the methodology provides an upper bound to the cost. Moreover, as it is a worst case based method, it doesn't need a probability distribution on the uncertainty set. The drawback is that an exhaustive uncertainty characterization is still needed, although no probability distribution is required. In fact, the considered uncertainty set plays a crucial role, since it determines the level of protection of the solution. But this is a major drawback, since if all scenarios are considered, the solution might be way too conservative and lead to a solution with high costs for most of the possible outcomes; neglecting part of the possible outcomes leaves the possibility for the solution to become unfeasible. In this case, the cost of the solution is no longer an upper bound, and then the question arises whether the additional costs on the considered outcomes are worth it.

This leads to another type of worst-case based approach, namely the *risk management* methods, see Kall and Mayer, 2005. For these methods, a probabilistic measure on the uncertainty set is required, and the optimal solution is the one that has the best trade-off between expected cost and probability to be infeasible. The probability to be infeasible is modeled using quantile functions, and the optimal solution is the one with lowest expected cost given a specific value of the quantile function. The benefit of the approach is to find the solution with lowest expected cost and provide a probabilistic measure of infeasibility. The method suffers, however, from the needs of a probabilistic uncertainty set as does stochastic programming. Moreover, the obtained problem is computationally hard, such that only particular problems are solvable. Note that risk management also fits into the class of stochastic methods.

Lately, Fischetti and Monaci, 2008 introduce the concept of *light robustness*, which can be seen as an extension of Bertsimas and Sim, 2004. The aim of a light robust solution is to minimize the constraint violation within a determined maximal deviation from the deterministic optimal solution. The quality of a solution is defined as the worst violation in the basic Light Robustness (LR) and the deviation from the average violation in the Heuristic Light Robustness (HLR) approach. The originality of this work is that the authors fix a maximal optimality deviation within which the LR or HLR measures of robustness have to be optimized. The study limits to integer linear problems with the uncertainty set defined by Bertsimas and Sim, 2004.

In both the (light) robust and the risk management methods, the user invests some additional costs in order to gain feasibility within a determined set of outcomes. Bertsimas and Sim, 2004 calls it the *price of robustness*.

We learn from the literature that all existing methods have some drawbacks: deriving an uncertainty set is a difficult problem; erroneous uncertainty sets may dramatically impact the solution's performance in reality; only few a priori information is known about the real outcome. Additionally, stochastic programming approaches lead to computationally hard problems Birge and Louveaux, 1997 and robust solutions might be too conservative.

The aim of the Uncertainty Feature Optimization (UFO) framework is to overcome the main drawback of the existing a priori approaches: no uncertainty set required. This reduces the modeling effort of the uncertainty characterization, makes the approach stable against errors in the noise's nature estimation and does not significantly increase the complexity of the original problem. The inconvenient is that no a priori guarantee about future outcome is possible: only simulation allows to test the approach's performance.

3 UFO Framework

The general idea of Uncertainty Feature Optimization (UFO) is to save the modeling effort to derive an uncertainty set, modeling the uncertainty implicitly with Uncertainty Features (UF). An UF is a structural property of the solution that is proven to ameliorate the solution's *robustness* (capacity to remain feasible) or *recoverability* (reduction of recovery costs when solution is infeasible). Without loss of generality, we suppose the UF has to be maximized in order to increase the solution's robustness or recoverability.

Consider the general optimization problem (P) that is prone to noise in the data, whose nature is unknown:

$$z_P = \min f(\mathbf{x}) \tag{1}$$
$$\alpha(\mathbf{x}) \leq \mathbf{b} \tag{2}$$
$$\mathbf{x} \in X \tag{3}$$

An *Uncertainty Feature* (UF) is a function $\mu : \mathbb{R}^n \rightarrow \mathbb{R}$ that maps $\mathbf{x}$ into a scalar $\mu(\mathbf{x})$. We suppose that an increase of $\mu(\mathbf{x})$ implies a better performance of the solution $\mathbf{x}$ under noisy data: there is a significant (inverse) correlation between $\mu(\mathbf{x})$ and $f(\mathbf{x})$. Let $\mathbb{M}$ be the number of considered uncertainty features.

We reformulate (P) as a multi-objective optimization problem by adding the uncertainty features $\mu(\mathbf{x})$. Objective (1) becomes:

$$[z_P, z_1, \cdots, z_{\mathbb{M}}] = [\min f(\mathbf{x}), \max \mu_1(\mathbf{x}), \cdots, \max \mu_{\mathbb{M}}(\mathbf{x})]. \tag{4}$$

The obtained problem is then transformed into the following problem (P'):

$$z_{P'} = [\max \mu_1(\mathbf{x}), \cdots, \mu_{\mathbb{M}}(\mathbf{x})] \tag{5}$$
$$\alpha(\mathbf{x}) \leq \mathbf{b} \tag{6}$$
$$f(\mathbf{x}) \leq (1+\rho) f^* \tag{7}$$
$$\mathbf{x} \in X \tag{8}$$

where f^* is the optimal solution of the deterministic problem (P), and $\rho \geq 0$ is a scalar called the *budget ratio*. We call constraint (7) the *budget constraint*. It limits the optimality gap with respect to the deterministic optimal solution f^*.

Remarkably, the feasibility of solution $\mathbf{x}$ according to (P) remains: any feasible solution of (P') is also feasible for (P). Additionally, the noisy data the problem is prone to is implicitly consider when maximizing the UFs.

We choose here to solve the multi-objective optimization problem (4) using the initial objective relaxation with the budget constraint. The other possibilities of solving such a problem are the exploration of the Pareto frontier or to optimize a weighted combination of the different objectives. Although the choice seems arbitrary at this point, we show in the next section that the budget constraint is particularly convenient: first of all, it is an intuitive approach, and it allows to derive existing a priori methods as particular cases of the UFO framework.

The main difficulty of the UFO framework is to derive the UFs. We think there is no a priori way to define them, only simulation reveals an UF's efficiency: an efficient UF must be inversely correlated with the initial objective $f(\mathbf{x})$.

When using several UFs, then (P') is still a multi-objective optimization problem. We suggest, at this stage, to solve a weighted combination of the UFs, normalizing them according to their respective correlation with the original objective.

4 UFO as a Generalization

The aim of this section is to show that when an uncertainty set is provided, then existing methods can be derived from the UFO framework using particular UFs: no assumption is made on the nature of the UF, so it is possible to use an UF relying on the provided uncertainty set, call it U. In addition, we suppose U is provided with a probabilistic measure. Moreover, when not specified, we consider a deterministic problem of the form (1)-(3).

Light Robustness The formulation of Fischetti and Monaci, 2008 uses the same budget constraint. Their objective, however, is based on an uncertainty characterization: they seek for the solution with lowest constraint violation in the worst case. UFO is clearly a generalization of the approach, as the proposed LR and HLR violation measures can be regarded as UFs.

Stochastic Programing Consider the following uncertainty feature:

$$\mu_{\texttt{Stoc}}(\mathbf{x}) = -\mathbb{E}_U(f(\mathbf{x})),$$

where $\mathbb{E}_U(f(\mathbf{x}))$ is the expected value of $f(\mathbf{x})$ over the uncertainty set U. Applying the UFO framework, we get the following problem:

$$z_{\texttt{Stoc}} = \min \mathbb{E}_U(f(\mathbf{x})) \tag{9}$$

$$\alpha(\mathbf{x}) \leq \mathbf{b} \tag{10}$$

$$f(\mathbf{x}) \leq (1+\rho)f^* \tag{11}$$

$$\mathbf{x} \in X \tag{12}$$

When $\rho = 0$, the solution space reduces to the deterministic optimal solutions only, and the value $z^*_{\texttt{Stoch}}$ is the expected cost of the deterministic solution. When $\rho \to \infty$, all feasible solutions are considered: the solution is the one minimizing the expected cost, i.e. the solution of a the corresponding stochastic expected cost minimization problem.

Suppose that we are provided with a recovery (or *recourse*) strategy: for each solution $\mathbf{x}$, let $g(\mathbf{x}, \xi)$ be the recovery (fixed recourse) costs for solution $\mathbf{x}$ when the observed data outcome is $\xi \in U$. The corresponding deterministic equivalent program (D.E.P) Birge and Louveaux, 1997 formulation of a two-stage stochastic program with fixed recourse is:

$$z_{\texttt{Rec}} = \min f(\mathbf{x}) + \mathbb{E}_U(g(\mathbf{x}, \xi)) \tag{13}$$

$$\alpha(\mathbf{x}) \leq \mathbf{b} \tag{14}$$

$$\mathbf{x} \in X \tag{15}$$

We define the following UF:

$$\mu_{\texttt{Rec}}(\mathbf{x}) = -\left[f(\mathbf{x}) + \mathbb{E}_U(g(\mathbf{x}, \xi))\right],$$

Applying UFO framework, we obtain formulation (13)-(15) with the additional budget constraint $f(\mathbf{x}) \leq (1+\rho)f^*$. Again, $\rho = 0$ means only deterministic optimal are considered, whereas $\rho \to \infty$ finds the solution of the D.E.P. Birge and Louveaux, 1997.

Robust Optimization Consider the approach of Bertsimas and Sim, 2004[2] for linear robust optimization:

$$z^*_{ROB} = \min \mathbf{c}^T\mathbf{x} \tag{16}$$
$$\sum_j a_{ij}x_j + \beta_i(\mathbf{x}, \Gamma_i) \leq b_i \tag{17}$$
$$\mathbf{x} \in X \tag{18}$$

In this problem, only the matrix coefficients A vary. The uncertainty set U is characterized by the set J_i containing the indexed of the uncertain coefficients for each row $i = 1, \cdots, n$. Each coefficient satisfies $a_{ij} \in [a_{ij} - \hat{a}_{ij}, a_{ij} + \hat{a}_{ij}]$.

Given a solution $\mathbf{x}$, the worst coefficient realization at row i is given by

$$\beta_i(x, \Gamma_i) = \max_{\{S_i \cup \{t_i\} | S_i \in J_i, |S_i| = \lfloor \Gamma_i \rfloor, t_i \in J_i \setminus S_i\}} \left\{ \sum_{j \in S_i} \hat{a}_{ij} \mid x_j \mid + (\Gamma_i - \lfloor \Gamma_i \rfloor) \hat{a}_{it_i} \mid x_{t_i} \mid \right\}, \tag{19}$$

where Γ_i is an upper bound on the number of coefficient allowed to vary simultaneously. We define the complementary function of $\beta_i(x, \Gamma_i)$:

$$\overline{\beta}_i(\mathbf{x}, \Gamma_i) = \min_{\{S_i \cup \{t_i\} | S_i \in J_i, |S_i| = \lfloor J_i - \Gamma_i \rfloor, t_i \in J_i \setminus S_i\}} \left\{ \sum_{j \in S_i} \hat{a}_{ij} \mid x_j \mid + (1 - \Gamma_i + \lfloor \Gamma_i \rfloor) \hat{a}_{it_i} \mid x_{t_i} \mid \right\}.$$

Theorem (Complementarity)
If β_i and $\overline{\beta}_i$ are defined as above, then

$$\beta_i(\mathbf{x}, J_i) = \overline{\beta}_i(\mathbf{x}, \Gamma_i) + \beta_i(\mathbf{x}, \Gamma_i).$$

The proof of the theorem is left in Appendix A. Note also that both β_i and $\overline{\beta}_i$ are positive valued functions.

As robust optimization aims at feasible solutions, we begin with the original feasibility problem (F) as:

$$\begin{aligned} (F) \quad z^*_F &= \min \{f(\mathbf{x})\} \\ &= \min \{\max_i (f_i(\mathbf{x}))\} \\ &= \min \left\{\max_i \left(\sum_{j=1}^n a_{ij}x_j + \beta_i(\mathbf{x}, J_i) - b_i\right)\right\} \end{aligned}$$

[2] REMARK: Bertsimas and Sim, 2004 use a maximisation problem; we transform it to a minimisation problem to match our framework.

The uncertain set of (F) is U, i.e.

$$a_{ij} \in [a_{ij} - \hat{a}_{ij}, a_{ij} + \hat{a}_{ij}], \forall j \in J_i, i = 1, \cdots, n.$$

$f(\mathbf{x})$ is to the value of the most violated constraint in the worst scenario when all the J_i coefficients of row i vary, i.e. the unbounded worst case. A solution with $f(\mathbf{x}) \leq 0$ is a solution that is feasible on the whole uncertainty set; (F) is thus the feasibility problem that seeks the solution $\mathbf{x}$ that is closest to be feasible for all scenarios or, i.e. a robust solution.

We define the UF $\mu(\mathbf{x}) = -\mathbf{c}^T\mathbf{x}$, the original cost function with negative sign. Clearly, μ and f are inversely correlated because of the *price of robustness* Bertsimas and Sim, 2004. Additionally, the UF increases the performance of the solution: the cost is decreased.

If a solution of (F) such that $z_F^* \leq 0$ exists, at least one robust solution exists. Using the UFO framework with budget constraint $f(\mathbf{x}) \leq 0$ ($\rho = \frac{z_F^*}{|z_F^*|}$ is $z_F^* \neq 0$ or $\rho = 0$ if $z_F^* = 0$) leads to the robust solution that has lowest cost, which is what is sought.

Suppose thus that $z_F^* > 0$, i.e. no robust solution exist on U. We define the budget ratio ρ as:

$$\rho = \max_i \left\{ \frac{\rho_i f_i(\mathbf{x}^*)}{z_F^*} - 1 \right\},$$

where ρ_i is defined as the ratio:

$$\rho_i = \begin{cases} \frac{\overline{\beta}_i(\mathbf{x},\Gamma_i)}{f_i(\mathbf{x}^*)} & \text{if } f_i(\mathbf{x}^*) \neq 0 \\ 0 & \text{otherwise.} \end{cases}$$

$f_i(\mathbf{x}^*) = a_{ij}x_j^* + \beta_i(\mathbf{x}^*, J_i) - b_i$ is the deviation of constraint i in the optimal solution $\mathbf{x}^*$ of problem (F), which is a deterministically known value when (F) is solved. The additional budget constraint when applying the UFO framework is thus

$$f(\mathbf{x}) \leq (1+\rho)f^* = \min_i \left\{ \overline{\beta}_i(\mathbf{x}, \Gamma_i) \right\}.$$

The ratio $(1+\rho)$ corresponds to the maximal proportion the $J_i - \Gamma_i$ least varying coefficients represent at the optimal solution of (F): it is capturing how far the solution is from being robust when at most Γ_i coefficients vary.

A negative value of ρ implies $f_i^* < 0, \forall i$, i.e. the solution is robust, which contradicts the hypothesis $z_F^* > 0$.

If the ratio is $\rho > 0$, then $\overline{\beta}_i(\mathbf{x}, \Gamma_i) > f_i^* \geq 0$. This implies that the solution space is extended: the solutions are no longer required to be feasible on the whole uncertainty set U. Robustness is required on a subset of U, for which the number of simultaneously varying coefficients is bounded.

(F') is obtained by re-ordering the budget constraint:

$$z_{F'}^* = \min \mathbf{c}^T\mathbf{x} \tag{20}$$

$$\sum_j a_{ij}x_j + \beta_i(\mathbf{x}, J_i) - \min_i \overline{\beta}_i(\mathbf{x}, \Gamma_i) \leq b_i \tag{21}$$

$$\mathbf{x} \in X \tag{22}$$

Due to the complementarity theorem, we have that every solution satisfying constraint (21) also satisfy constraints (17). (F') is a tighter formulation than problem (16)-(18), as it is robust for more data variation.

Applying the procedure iteratively leads either to the robust solution of formulation (16)-(18) or proves that the problem is unfeasible, i.e. no robust solution with the proposed Γ_i exists.

At each iteration k we consider the problem

$$\Gamma_i^{(k+1)} = \Gamma_i^{(k)} - \inf_{\Gamma \geq 0} \left\{ \Gamma \mid \overline{\beta}_i(\mathbf{x}_\mathbf{k}^*, J_i - \Gamma_i^{(k)} - \Gamma) \geq f_i^{(k)}(\mathbf{x}_k^*), \Gamma \leq J_i - \Gamma_i^{(k)} \right\}.$$

where $\mathbf{x}_k^*$ is the solution minimizing

$$f^{(k)}(\mathbf{x}) = \max_{i,\mathbf{x}\in X} f_i^{(k)}(\mathbf{x}) = \max_{i,\mathbf{x}\in X} \left\{ \sum_j a_{ij}x_j + \beta_i(\mathbf{x}, \Gamma_i^{(k-1)}) - b_i \right\}.$$

$\Gamma_i^{(k)}$ is, at each iteration, the maximal number of varying coefficients at row i. It is bounded by the number of varying coefficients $\Gamma_i^{(k-1)}$ from the previous iteration, i.e. $\Gamma_i^{(k)} \leq \Gamma_i^{(k-1)}$. This procedure is repeated until either $\rho^k \leq 0$ or all $\Gamma_i^{(k)} = 0$. In the first case, we find a robust solution, and the maximal allowed values of Γ_i such that a robust solution exists. In the latter, we prove that the problem has no feasible solution, i.e. the initial problem (16)-(18) is infeasible. The proof of the convergence of the method is left in Appendix B. In the case Γ_i are integer, the method converges in at most $m \times n$ iterations.

Remark that the methodology we develop here has many similitudes with the light robustness of Fischetti and Monaci, 2008. The main difference is the approach: Fischetti and Monaci, 2008 start from the original cost minimization problem, aiming at minimizing the constraint violation, whereas the current approach starts from violation minimization and aims at cost minimization. Additionally, Fischetti and Monaci, 2008 measure characterize the worst violation according to the deterministic optimum; this is not correct, since the worst case is solution dependent.

Summary The UFO framework can be seen as a generalization of existing methods using particular uncertainty features. The point of interest is that an UF is *any* feature expected to improve the solution's performance in reality, and it is left to the user to decide the complexity and computational effort to invest in the estimation of the future outcome. Additionally, we show that the UFO framework leads to an algorithm determining upper bounds on the Γ_i coefficients in Bertsimas and Sim, 2004; no such consideration was found in the literature so far.

5 Illustration on the Multi-Dimensional Knapsack Problem

Consider the Multi-Dimensional Knapsack Problem (MDKP):

$$\begin{aligned} z^*_{MDKP} = &\max \mathbf{c}^T\mathbf{x} \\ &\sum_{j=1}^{m} a_{ij}x_j \leq b_i && \forall i = 1, \cdots, n \\ &x_i \in \mathbb{Z}_+ && \forall i = 1, \cdots, n \end{aligned}$$

We suppose that all coefficients a_{ij} may vary. In the context of robust optimization, the set varying coefficients is $J_i = \{1, \cdots, m\}$. No assumption is made on the uncertainty set, i.e. the possible outcomes of the coefficients a_{ij} is made.

The corresponding UFO formulation is:

$$
\begin{aligned}
z_{MDKP'} = &\max \mu(\mathbf{x}) \\
&\sum_{j=1}^{m} a_{ij} x_j \leq b_i && \forall i = 1, \cdots, n \\
&\mathbf{c}^T \mathbf{x} \geq (1-\rho) z^*_{MDKP} \\
&x_i \in \mathbb{Z}_+ && \forall i = 1, \cdots, n
\end{aligned}
$$

The x_i variables are the number of times object i is taken in the solution. Since we have a maximisation problem, the budget constraint is multiplied by $1-\rho$: the optimality deviation is a loss of revenue instead of a direct cost.

We derive four different UFs for the problem:

μ_{MTk}	$= 1 - \max_{i=1,\cdots,n}\{\frac{x_i}{u^*}\}$,	the Maximal Taken object;,
μ_{Div}	$= \sum_{i=1,\cdots,n}(\frac{\min\{x_i,1}{n}\})$,	the Diversification of the taken objects;
μ_{IR}	$= 1 - \max_{i=1,\cdots,n}\{\frac{a_{ij}x_i}{b_i}\}$,	the maximal Impact Ratio of a taken object;
μ_{2Sum}	$= 1 - \max_{i,j\neq k}\{\frac{a_{ij}x_i + a_{ik}x_k}{b_i}\}$,	the maximal *size* of two objects in a same constraint;

The robust formulation of Bertsimas and Sim, 2004 is referred to as the UF μ_{Rob}.

The derived UFs follow intuition: taking many times a same object i is risky, as if any of its coefficient a_{ij} increases, the solution becomes more likely to be infeasible. The negative sign of μ_{MTk} ensures the maximal taken object is minimized. Having a diversified solution is another potentially improving property: we do not expect that all coefficients increase at the same time, and the increase of some coefficients might be compensated by the decrease by some others. Finally, the IR and 2Sum capture the fact that it is risky to take objects with high coefficients in the constraints.

Note that the definition of the UFs ensure their optimal (i.e. maximal) value is 1. We also normalize the robust problem in order to set its optimal value to 1. As all UFs are normalized, the combination of multiple UFs is handled using the arithmetic mean of the considered UFs. The value of $z_{MDKP'}$ is thus the average value of the (potentially) different considered UFs.

5.1 Simulation Description

The used instances are generated by the MDKP-simulator[3]. The generated instances can be solved by any combination of the UFs proposed in the previous section. The user the chooses the simulation parameters to test the solution.

The generation parameters we use are similar to the ones of Pisinger, 1995, which also uses cost-weight correlation and different densities for the marginal costs (the ratio $\frac{b_i}{c_i}$). We refer to the marginal cost density as the *degeneration* of the solution.

[3]A beta version of the simulator will be available soon. See http://transp-or.epfl.ch for updates.

We generate a total of 3240 instances with 50 objects. Each instance is defined by the cost and right hand side vectors $\mathbf{c}, \mathbf{b}$, the constraint matrix A and the standard deviation matrix $\hat{A}$, which characterizes the *real* uncertainty set the instance is due to. 5 instances are generated for each combination of the following parameters:

Number of constraints	1, 5 and 10;
Marginal Cost Correlation	cost and weights are/are not correlated;
Data Distribution	uniform or gaussian distribution;
Degeneration	clustered, medium and wide marginal cost distribution;
Deviation Matrix	average coefficient $\hat{a}_{ij}$ is proportional to ρa_{ij} with $\rho \in \{0.2, 0.5, 0.8\}$;
Deviation Distribution	deterministic, gaussian or uniform distribution of $\hat{A}$;
Γ	$\Gamma_i = 2, \forall i$ or $\Gamma_i = 50, \forall i$.

The marginal cost degeneration refers to the degeneration of the solution. A clustered instance, for example, is such that the marginal costs $\frac{\hat{a}_{ij}}{b_i}$ have a low standard deviation. The instances with $\Gamma_i = 50$ are instances where, in the robust model of Bertsimas and Sim, 2004, all coefficients are allowed to vary: robustness is guaranteed. In the case $\Gamma_i = 2$, the robust approach is underestimating the number of varying coefficients: infeasibility might occur. The choice of $\Gamma_i = 2$ is motivate by the fact that in the knapsack problem, only few objects are taken in the optimal solution.

The used UFs are the ones described in the previous section. The name of a solution is followed by the allowed budget ratio: `MTk_0.1` is the solution obtained with the Maximal Taken object UF and a budget $\rho = 0.1$. `Det` refers to the deterministic problem. In the case the robust model of Bertsimas and Sim, 2004 is used, we differentiate which uncertainty characterization is used when solving the problem: `Rob_Â` uses the real uncertainty set. `Rob_A_r` uses an approximated uncertainty characterization derived from the original constraint matrix A by multiplying each coefficient by `r`.

When we compare different UFs, their value is normalized such that the optimal value (when $\rho = 1$) is 1. When values of combined UFs are presented, we alway show the average value.

Each solution is then evaluated on a certain number of *scenarios*, which is defined by one realization of the constraint matrix $\tilde{A}$. The solution $\mathbf{x}^*$ is feasible for the scenario if $\tilde{A}\mathbf{x}^* \leq \mathbf{b}$ and the optimality gap is the gap between the solution's cost (which is constant) and the optimal value of the scenario.

For each of the following simulations, we randomly generate 5 scenarios with uniform distribution, for each instance:

$\hat{A}$ 75, 100	$\tilde{A}$ has mean proportional to 75% or 100% times $\hat{A}$;
A 10, 25, 50	$\tilde{A}$ has mean proportional to 10%, 25% or 50% times A;
R 10, 20, 30	$\tilde{A}$ is randomly generated with mean value of 10, 20 and 30.

We thus have a total of 3240 instances and $129,600$ scenarios; we test a total of 22 different UF combinations.

5.2 Computation Results for MDKP

In this section, we detail only the most relevant results. The remaining results are described qualitatively at the end of the section.

Selected Detailed Results The detailed results show the average performance of one solution over a set of simulations. It shows the normalized UF value (when meaningful), the number of infeasible scenarios, the percentage of feasibility failure, the average and maximal optimality gaps (between solution and the scenario's optimal value) and the maximal number of violated constraints.

Tables 1-3 summarize the average results of the 180 instances with cost-correlation clustered degeneration and 10 constraints.

		Det	Rob_Â	Rob_A_0.1	MTk_0.2	Div_0.1	IR_0.3	2Sum_0.1
Â 75, 100 (1800 Scen.)	UF value	-	-	-	0.974	0.610	0.962	0.932
	# Infeas.	1642	166	1014	914	1199	85	1174
	Infeas [%]	91.22	9.22	56.33	50.78	66.61	4.72	65.22
	Avg Opt Gap [%]	0.56	20.93	4.72	10.53	5.29	31.04	5.56
	Max Opt Gap [%]	25.21	68.36	38.91	49.81	50.47	59.53	41.99
	Max # Violated	9	3	7	4	5	1	5
A 10, 25, 50 (2700 Scen.)	UF value	-	-	-	0.974	0.610	0.962	0.932
	# Infeas.	2404	489	1232	1141	1544	76	1566
	Infeas [%]	89.04	18.11	45.63	42.26	57.19	2.81	58.00
	Avg Opt Gap [%]	0.56	18.13	5.09	11.38	6.03	30.04	6.03
	Max Opt Gap [%]	34.03	57.07	40.47	47.12	40.39	52.16	40.19
	Max # Violated	8	6	7	3	5	2	5
R 10, 20, 30 (2700 Scen.)	UF value	-	-	-	0.974	0.610	0.962	0.932
	# Infeas.	2616	1079	2100	1506	1974	171	1962
	Infeas [%]	96.89	39.96	77.78	55.78	73.11	6.33	72.67
	Avg Opt Gap [%]	0.45	17.32	3.36	11.24	5.36	33.33	5.53
	Max Opt Gap [%]	46.67	62.71	51.87	57.25	51.81	61.33	51.65
	Max # Violated	8	7	8	4	6	2	6

Table 1: Simulation results for instances with 10 constraints.

The UF value is the same for the different simulations: indeed, the UF value is the one of the computed solution, which does not change for different scenarios. The deterministic model performs extremely bad: the solution is infeasible for more than 91% of the scenarios. The rare scenarios for which the solution is feasible are the ones equivalent to the deterministic instance, for which the solution is optimal, explaining the low optimality gaps.

The robust solution Rob_$\hat{A}$ is performing best when the scenario is generated according to the real deviation matrix $\hat{A}$. Infeasibility for the `Rob_Â$` is due to the instances where $\Gamma_i = 2$ is too low to guarantee complete robustness. The robust solution is, however, the only one with decreasing performance when the noise's characterization is erroneous: `Rob_Â$` is the only solution having less feasibility success in the A tests than the $\hat{A}$ tests. The `Rob_A_0.1` solutions use the same matrix that is used to generate the scenario. The results are not accordingly better: the randomness of the scenarios makes the estimation erroneous. In both `Rob_Â` and `Rob_A_0.1`, the solution significantly looses in feasibility success when the scenario is generated randomly (R simulations).

The IR_30 model clearly outperforms the robust model in terms of feasibility even for the case where the exact deviation matrix is used: this is due to the choices of Γ_i. This shows that even with the correct uncertainty characterization, the estimation of the Γ_i strongly influences the solution's performance.

As expected, the UF solutions are less sensitive to the noise's nature. Remarkably, the budget ratio seems a decent estimator of the average optimality gap. This becomes more relevant when looking at Table 2, that shows the simulation results on the same instances than Table 1 for the globally best two UFs in our tests.

		2Sum_0.1	2Sum_0.2	2Sum_0.3	IR_0.1	IR_0.2	IR_0.3
Â 75, 100 (1800 Scen.)	UF value	0.932	0.958	0.962	0.932	0.958	0.962
	# Infeas.	1174	528	90	1220	528	85
	Infeas [%]	65.22	29.33	5.00	67.78	29.33	4.72
	Avg Opt Gap [%]	5.56	16.17	30.60	5.13	16.24	31.04
	Max Opt Gap [%]	41.99	52.36	59.53	42.18	48.58	59.53
	Max # Violated	5	3	1	5	3	1
A 10, 25, 50 (2700 Scen.)	UF value	0.932	0.958	0.962	0.932	0.958	0.962
	# Infeas.	1566	579	84	1671	592	76
	Infeas [%]	58.00	21.44	3.11	61.89	21.93	2.81
	Avg Opt Gap [%]	6.03	16.85	29.57	5.4	16.83	30.04
	Max Opt Gap [%]	40.19	46.95	46.75	40.39	46.99	52.16
	Max # Violated	5	3	2	5	3	2
R 10, 20, 30 (2700 Scen.)	UF value	0.932	0.958	0.962	0.932	0.958	0.962
	# Infeas.	1962	997	174	2001	996	171
	Infeas [%]	72.67	36.93	6.44	74.11	36.89	6.33
	Avg Opt Gap [%]	5.53	16.73	32.93	5.33	16.81	33.33
	Max Opt Gap [%]	51.65	57.11	60.71	51.81	57.15	61.33
	Max # Violated	6	3	2	5	4	2

Table 2: Sentibility to increasing budget ρ.

An increase of ρ clearly ameliorates the performance of the shown UFs in terms of robustness. The results are coherent with the statement of Bertsimas and Sim, 2004: an increase of robustness comes at a certain price. The differences of the UF values for the two UFs seem identical: they are not, but it is not clear because of rounding effect.

Table 3 shows the effect of combining the different models. We use a complete budget relaxation ($\rho = 1$) for the model combining `Rob_A_Div`. The reason is that the budget constraint may contradict with the revenue maximization objective of the `Rob` model.

The table shows that the combination `MTk_2Sum` is very efficient, leading to robust solutions with low optimality gaps, outperforming all other methods for the shown simulations. In addition, combining the Rob model with the Div UFO shows an impressive improvement: the Rob_0.2 model is clearly outperformed by the combined models, both in terms of feasibility and optimality gaps.

Qualitative Analysis Tables 1-3 show the results for only one out of 18 classes. We describe here in a qualitative way the content of the remaining tests[4].

[4]The complete results are available on written demand to niklaus.eggenberg@epfl.ch

		MTK_2Sum_0.3	_Rob_A_0.1 Div_1.0	Rob_A_0.1	Rob_A_0.2
Â 75, 100 (1800 Scen.)	UF value	0.969	1.475	-	-
	# Infeas.	102	514	1014	643
	Infeas [%]	5.67	28.56	56.33	35.72
	Avg Opt Gap [%]	33.02	16.47	4.72	10.93
	Max Opt Gap [%]	80.03	53.41	38.91	44.15
	Max # Violated	2	5	7	5
A 10, 25, 50 (2700 Scen.)	UF value	0.969	1.475	-	-
	# Infeas.	111	542	1232	700
	Infeas [%]	4.11	20.07	45.63	25.93
	Avg Opt Gap [%]	31.90	16.93	5.09	11.59
	Max Opt Gap [%]	76.44	51.97	40.47	45.34
	Max # Violated	2	5	7	5
R 10, 20, 30 (2700 Scen.)	UF value	0.969	1.475	-	-
	# Infeas.	209	969	2100	1503
	Infeas [%]	7.74	35.89	77.78	55.67
	Avg Opt Gap [%]	34.94	17.16	3.36	9.18
	Max Opt Gap [%]	80.20	61.17	51.87	55.81
	Max # Violated	3	5	8	6

Table 3: Combination effect of different UFs.

Interestingly, the results show that the number of constraints is a crucial parameter, especially for the performance of the `Det` solutions: feasibility is lost in only 37.01% of all the scenarios with 1 constraint, for $83.72\%$for 5 constraints and more than 91% with 10 constraints, with the cost-correlated and clustered degeneration instances. The other solutions show a similar behavior. The reason is that the more a problem is constrained, the more it is sensitive to variations.

The next observation is that the `Rob` model performs much better than the UFs when the degeneration becomes lower. The global statistics show that the robust solution is, in average, feasible around twice as often as the solutions of the best UFs (22.6% of infeasible scenarios for the robust model for around 50.3% for the `MTk_2Sum`). The main difference comes from the instances with low and medium degeneration (i.e. 66% of the instances), where the robust solutions clearly outperforms the UF ones. The reason is that there is no optimality deviation restriction for the `Rob` solution. In the medum and low degeneration instances, the budget ratio ρ might not be sufficient to significantly extend the solution space, leading to solutions very similar to the deterministic one. No clear pattern can be identified for the robust solution according to variation of degeneration; for the UF solutions, a significant increase in the number of infeasible scenarios occurs when decreasing the degeneration. For `IR_0.3` for example, 29.01% scenarios are infeasible in the clustered instances, for 54.93% in the medium and 62.67% in the low degeneration case. The same remark holds for the cost correlation. The robust solutions are less sensitive to cost correlation than the UF solutions, for the same reasons.

Remarkably, unlike the robust or deterministic approaches, the UF solutions tend to be stable for the different simulation types, whatever the instance type. Surprisingly, this also holds for the number of constraints. The reason is that the robust approach is based on an uncertainty characterization. The robust solutions are only better when their information is sufficient and when the solution is degenerated enough for the UF to have increased solution space.

Synthesis The simulations show that the UF are competitive (sometimes even much better) for problems with clustered degeneration and cost-correlation, which is the case of most of the complex real world optimization problems. Additionally, UF solutions are not sensitive to changes in the noise's nature, unlike the robust approach. The simulations also show that even if a robust method benefits from the exact uncertainty characterization, the method might still lead to inefficient solutions because of it's parameters. In the case of low or medium degeneration, UF perform worse when the budget ration is too small to allow significant extension of the search space.

6 Extension to Airline Scheduling

The Airline Scheduling Problem (ASP) is a huge problem involving many complex regulations, see Kohl et al., 2004 for a survey. The many facets of the problem (route choice, fleet assignment, tail assignment, crew pairing and crew roistering) represent a combinatorial challenge for operations research scientists Clausen et al., 2004.The additional problem is that the computed schedules have to be carried out in a rapidly varying environment influenced by many factors such as weather, human factors (strikes, illness, ...) and economical factors. The complexity of the environment makes it extremely difficult, if not impossible, to derive a complete and correct characterization of its behavior.

Being already a hard problem in its deterministic form, it seems not realistic to use proactive methods for solving the ASP: it is a good candidate for the UFO framework. This does not hold uniquely for the ASP problem: Fischetti and Monaci, 2008 successfully applies light robustness, which is computationally similar to UFO, to the train tabling problem, showing impressing computational time savings, in addition to competitive solutions in terms of robustness.

As discussed in section 1, some possible uncertainty features to increase robustness of an airline schedule are idle time, plane crossings or number of plane routes matching the worker's union constraints. As it is unlikely that a robust solution exists, it is appealing to search for increased recoverability as well.

In Eggenberg et al., 2008, the authors present a Column Generation (CG) algorithm to solve the Aircraft Recovery Problem. The advantage of the technique is that it is flexible enough to be applied for crew and a combination of aircraft routing, crew and passenger recovery. The algorithm is based on *recovery networks*, encoding each *unit*'s (aircrafts, crew or passengers) feasible route. The performance of the recovery algorithm is directly linked with the structure of the recovery networks. This can be exploited at the ASP phase, using UFs based on the recovery networks' structure in order to increase recoverability.

As an example, a promising UF is to minimize the number of successive airports where no maintenance can be performed: a plane requiring an unpredicted maintenance at an airport that does not support maintenance operations must be re-routed to an airport where the maintenance can be done; the number of potentially canceled flights is linked with the number of successive flights visiting unequipped airports. Similarly, if maintenance is performed only at a base airport, the UF is equivalent to minimize the length of the rotations.

UFO is a promising framework for computationally hard problems due to uncertain data such as the ASP for two reasons: the first is that, as long the used UFs are of the same nature than the original objective, then the computational difficulty is equivalent to solving twice a problem of same difficulty than the initial objective: once to get the lower bound f^* and once to solve the UFO problem (5)-(8). The second reason is that the characterization of uncertainty sets for such problems is a crucial but hard problem that is not required for general UFs. The major difficulty is the validation of an UF: it requires large simulations. We believe, however, that any scheduling approach should be validated by simulation.

7 Conclusion

In this paper, we address the problem of optimization prone to noisy data. Unlike most of the existing methods, the Uncertainty Feature Optimization framework does not require the explicit characterization of an uncertainty set, i.e. the possible outcomes of the data: the UFO framework models the uncertainty implicitly.

We show that existing methods such as stochastic optimization or robust optimization are special cases of UFs, supposing the uncertainty set provided. The proof of the generalization for the robust approach of Bertsimas and Sim, 2004 leads to an algorithm computing upper bounds on the method's parameter to guarantee a robust solution exists.

Computational results on the Multi Dimensional Knapsack Problem (MDKP) show that the UFO approach is competitive against the robust approach. The results show the stable behavior of UFO with respect to variations on the noise's nature, unlike the robust approach: the exact knowledge of the noise's nature is a benefit, but when the nature is erroneously approximated, it might annihilate a method's efficiency. Additionally, as show our results, the only knowledge of the noise's nature is not sufficient for the robust approach: the parameters of the method clearly influence the performance of a robust solution.

The future research directions are to test the approach on more complex problems. Indeed, the results show that the performance of UF solutions increases for more constraining problems. The airline scheduling problem seems an appropriated candidate: the problem is computationally hard and well studied with stochastic and robust methods to benchmark the performance of UFO, and simulation tools already exist for this problem.

Another research is to derive an UF generation framework, enabling the elaboration of an uncertainty feature based on a problem's structure. The generation framework may require a problem classification, where UF's would depend on the class a problem belongs to. It is not clear though whether a classification uniquely based on mathematical properties (such as number of constraints, variables, ...) is possible, or if the nature of the problem is relevant. The underlying question is whether it is possible to classify problems according to their difficulty with respect to uncertain data.

References

Al-Fawzana, M. and Haouari, M. (2005). A bi-objective model for robust resource-constrained project scheduling, *International Journal of Production Economics* **96**: 175Ű187.

Albers, S. (2003). Online algorithms: A survey, *Mathematical Programming* **97**: 3–26. Invited paper at ISMP 2003.

Ben-Tal, A. and Nemirovski, A. (2001). *Lectures on Modern Convex Optimization, Analysis, Algorithma, and Engineering Applications*, MPS-SIAM Series on Optimization.

Bertsimas, D. and Sim, M. (2004). The price of robustness, *Operations Research* **52**: 35–53.

Bian, F., Burke, E., Jain, S., Kendall, G., Koole, G., Mulder, J. L. S. J., Paelinck, M., Reeves, C. and Suleman, I. R. M. (2004). Measuring the robustness of airline fleet schedules.
URL: *http://www.cs.nott.ac.uk/ gxk/papers/jdsmista2003sel.pdf*

Birge, J. R. and Louveaux, F. (1997). *Introduction to Stochastic Programming*, Springer.

Chebalov, S. and Klabjan, D. (2004). Robust airline scheduling: Move-up crews, *Technical report*, Department of Mechanical and Industrial Engineering, University of Illinois at Urbana-Champaign.

Clausen, J., Larsen, A. and Larsen, J. (2004). Disruption management in the airline industry – review of models and methods, *Technical report*, Department of Informatics and Mathematical Modelling, The Technical University of Denmark.

Eggenberg, N., Salani, M. and Bierlaire, M. (2008). Constraint specific recovery networks for solving airline recovery problems, *Technical Report TRANSP-OR 080828*, Ecole Polytechnique Fédérale de Lausanne, Switzerland.

Fischetti, M. and Monaci, M. (2008). Light robustness, *Technical report*, DEI, Università di Padova, Italy.
URL: *http://www.dei.unipd.it/∼fisch/papers/light_robustness.pdf*

Grötschel, M., Krumke, S., Rambau, J. and Torres, L. (2002). Making the yellow angels fly: Online dispatching of service vehicles in real time, *Technical report*, Konrad-Zuse-Zentrum für Informationstechnik Berlin.

Herroelen, W. and Leus, R. (2005). Project scheduling under uncertainty: Survey and research potentialy, *EJOR* **165**: 289–306.

Kall, P. and Mayer, J. (2005). *Stochastic Linear Programming, Models, Theory and Computation*, Springer.

Kall, P. and Wallace, S. (eds) (1994). *Stochastic Programming*, John Wiley & Sons, New York, N.Y.

Klabjan, D., Johnson, E., Nemhauser, G., Gelman, E. and Ramaswamy, S. (2002). Airline crew scheduling with time windows and plane-count constraints, *Transportation Science* **36**(3): 337–348.

Kohl, N., Larsen, A., Larsen, J., Ross, A. and Tiourine, S. (2004). Airline disruption management - perspectives, experiences and outlook, *Technical report*, Informatics and Mathematical Modelling, Technical University of Denmark.

Lan, S., Clarke, J.-P. and Barnhart, C. (2006). Planning for robust airline operations: Optimizing aircraft routings and flight departure times to minimize passenger disruptions, *Transportation Science* **40**: 15–28.

Pisinger, D. (1995). An expanding-core algorithm for the exact 0-1 knapsack problem, *EJOR* **87**: 175–187.

Policella, N. (2004). *Robust Scheduling: Analysis and Synthesis of Flexible Solutions*, PhD thesis, Universita di Roma "la Sapienza".

Rosenberger, J., Johnson, E. and Nemhauser, G. (2004). A robust fleet assignment model with hub isolation and short cycles, *Transportation Science* **38**(3): 357–368.

Sahinidis, N. V. (2004). Optimization under uncertainty: state-of-the-art and opportunities, *Computers and Chemical Engineering* **28**: 971–983.

Soyster, A. (1973). Convex programming with set-inclusive constraints and application to inexact linear programming, *Operations Research* **21**(5): 1154–1157.

Yen, J. W. and Brige, J. R. (2006). A stochastic programming approach to the airline crew scheduling problem, *Transportation Science* **40**: 3–14.

A Complementarity Theorem

Theorem (Complementarity)
If β_i and $\overline{\beta}_i$ are defined as in section 4, then

$$\beta_i(\mathbf{x}, J_i) = \overline{\beta}_i(\mathbf{x}, \Gamma_i) + \beta_i(\mathbf{x}, \Gamma_i).$$

Proof:

For a given $\mathbf{x}$, let $S_i^* \cup \{t_i\}$ be the optimal set maximizing $\beta_i(\mathbf{x}, \Gamma_i)$ and $\overline{S}_i^* \cup \{\overline{t}_i\}$ the optimal set minimizing $\overline{\beta}_i(\mathbf{x}, \Gamma_i)$.

Suppose that we order the J_i changing coefficients in increasing order of the value $\hat{a}_{1j} \mid x_j \mid$. Then, the $\lfloor J - \Gamma_i \rfloor$ first ones are in $\overline{S}_i^*$. Similarly, the $\lfloor \Gamma_i \rfloor$ biggest ones are in S_i^*. Remains to check the fractional part of variable x_{t_i}: first of all, if Γ_i is integer, then there is no fractional part, so suppose Γ_i is non-integer. Clearly, $t_i = \overline{t}_i$, namely t_i is the variable in position $\lfloor J_i - \Gamma_i \rfloor + 1$.

Let us sum all terms of $\overline{S}_i^*$ and S_i^*, recalling that, as J_i is integer and Γ_i in non-integer, then $\lfloor J_i - \Gamma_i \rfloor = J_i - 1 - \lfloor \Gamma_i \rfloor$.

$$\begin{aligned}\overline{\beta}_i(\mathbf{x},\Gamma_i)+\beta_i(\mathbf{x},\Gamma_i) &= \sum_{j\in\overline{S}_i^*}\hat{a}_{1j}\mid x_j\mid+(J_i-\Gamma_i-\lfloor J_i-\Gamma_i\rfloor)\mid x_{t_i}+\sum_{j\in S_i^*}\hat{a}_{1j}\mid x_j\mid+(\Gamma_i-\lfloor\Gamma_i\rfloor)\mid x_{t_i}\\ &= \sum_{j\neq t_i}\hat{a}_{1j}\mid x_j\mid+\mid x_j\mid+(J_i-\Gamma_i-(J_i-1-\lfloor\Gamma_i\rfloor)+\Gamma_i-\lfloor\Gamma_i\rfloor)\mid x_{t_i}\mid\\ &= \sum_{j\neq t_i}\hat{a}_{1j}\mid x_j\mid+\mid x_j\mid+x_{t_i}\mid\\ &= \sum_{j}\hat{a}_{1j}\mid x_j\mid\\ &= \beta_i(\mathbf{x},J_i)\end{aligned}$$

□

B Convergence Proof

Proposition 1
Using the definitions in section 4, we have that:

$$\beta_i(\mathbf{x},J_i)\leq\beta_i(\mathbf{x},J_i-\Gamma_i)+\beta_i(\mathbf{x},\Gamma_i).$$

Proof:

By definition,

$$\begin{aligned}\beta_i(\mathbf{x},J_i-\Gamma_i) &= \max_{\{S_i\cup\{t_i\}|S_i\in J_i,|S_i|=\lfloor J_i-\Gamma_i\rfloor,t_i\in J_i\setminus S_i\}}\left\{\sum_{j\in S_i}\hat{a}_{ij}\mid x_j\mid+(\Gamma_i-\lfloor\Gamma_i\rfloor)\hat{a}_{it_i}\mid x_{t_i}\mid\right\}\\ &\geq \min_{\{S_i\cup\{t_i\}|S_i\in J_i,|S_i|=\lfloor J_i-\Gamma_i\rfloor,t_i\in J_i\setminus S_i\}}\left\{\sum_{j\in S_i}\hat{a}_{ij}\mid x_j\mid+(\Gamma_i-\lfloor\Gamma_i\rfloor)\hat{a}_{it_i}\mid x_{t_i}\mid\right\}\\ &\geq \overline{\beta}_i(\mathbf{x},\Gamma_i)\end{aligned}$$

As by definition $\overline{\beta}_i(\mathbf{x},\Gamma_i)=\min_{\{S_i\cup\{t_i\}|S_i\in J_i,|S_i|=\lfloor J_i-\Gamma_i\rfloor,t_i\in J_i\setminus S_i\}}\left\{\sum_{j\in S_i}\hat{a}_{ij}\mid x_j\mid+(\Gamma_i-\lfloor\Gamma_i\rfloor)\hat{a}_{it_i}\mid x_{t_i}\mid\right\}$.

Invoking the complementarity theorem we get:

$$\beta_i(\mathbf{x},J_i-\Gamma_i)+\beta_i(\mathbf{x},\Gamma_i)\geq\overline{\beta}_i(\mathbf{x},\Gamma_i)+\beta_i(\mathbf{x},\Gamma_i)=\beta_i(\mathbf{x},J_i),$$

which proves the proposition.

□

Theorem (Convergence)
Consider the sequence of problems for increasing k defined in section 4:

$$\Gamma_i^{(k+1)}=\Gamma_i^{(k)}-\inf_{\Gamma\geq0}\left\{\Gamma\mid\overline{\beta}_i(\mathbf{x}_\mathbf{k}^*,J_i-\Gamma_i^{(k)}-\Gamma)\geq f_i^{(k)}(\mathbf{x}_k^*),\Gamma\leq J_i-\Gamma_i^{(k)}\right\}.$$

where $\mathbf{x}_k^*$ is the solution minimizing

$$f^{(k)}(\mathbf{x}) = \max_{i,\mathbf{x}\in X} f_i^{(k)}(\mathbf{x}) = \max_{i,\mathbf{x}\in X} \left\{ \sum_j a_{ij}x_j + \beta_i(\mathbf{x}, \Gamma_i^{(k-1)}) - b_i \right\}.$$

Then the sequence either converges to a problem with $f^{(k)*} \leq 0$ or proves that there is no solution satisfying $\sum_j a_{ij}x_j \leq b_i$ for all i.

Proof:

First of all, notice that $\overline{\beta}_i(\mathbf{x}_\mathbf{k}^*, \Gamma)$ is a decreasing function for increasing Γ (see the complementarity theorem). Thus, if no solution exist for

$$\overline{\beta}_i(\mathbf{x}_\mathbf{k}^*, J_i - \Gamma_i^{(k-1)} - \Gamma) \geq f_i^{(k-1)}(\mathbf{x}_k^*),$$

this holds in particular for $\overline{\beta}_i(\mathbf{x}_\mathbf{k}^*, 0) = \beta(\mathbf{x}_k^*, J)$, again using the complementarity theorem.

This leads to

$$\sum_j a_{ij}(x_j)_k^* + \beta(\mathbf{x}_k^*, \Gamma^{(k-1)}) - b_i > \beta(\mathbf{x}_k^*, J).$$

As $\beta(\mathbf{x}_k^*, J) \geq \beta(\mathbf{x}_k^*, \Gamma^{(k-1)})$, this means $\sum_j a_{ij}(x_j)_k^* + \beta(\mathbf{x}_k^*, \Gamma^{(k-1)}) > b_i$, i.e. $\mathbf{x}^*$ is infeasible for the original problem.

We first prove that the sequence is not stationary; suppose it is, i.e. that $\Gamma_i^{(k+1)} = \Gamma_i^{(k)}$ for all i. In particular, this is also true for i^*, which is the index of the maximal valued function $f_i^{(k)}(\mathbf{x}_k^*)$. In this case, clearly, we have that

$$\overline{\beta}_{i^*}(\mathbf{x}_k^*, J_{i^*} - \Gamma_{i^*}^{(k)}) \geq f_{i^*}^{(k)}(\mathbf{x}_k^*).$$

Invoking Proposition 1, we end up with

$$f^{(k)}(\mathbf{x}_k^*) = f_{i^*}^{(k)}(\mathbf{x}_k^*) = \sum_j a_{i^*j}(x_j)_k^* - b_{i^*} \leq 0$$

i.e. we have converged.

For a non stationary solution, we have that $\Gamma_i^{(k+1)} < \Gamma_i^{(k)}$ for at least i^*. Moreover, we know that a solution of $\Gamma_{i^*}^{(k+1)}$ exists, otherwise we would have proved that no solution satisfying the set of equations $A\mathbf{x} \leq \mathbf{b}$ exist.

Now, at iteration $k+1$, all functions satisfy $f^{(k+1)}(\mathbf{x}) \leq f^{(k)}(\mathbf{x})$ for all $\mathbf{x} \in X$, the inequality being strict at least for i^*. The function $f^{(k)}(\mathbf{x})$ is thus strictly decreasing as well.

Thus, the method eventually converges either to a solution with $f^{(k)*} \leq 0$, or we have $\Gamma_i^{(k)} = 0$ for all i, meaning no solution for $A\mathbf{x} \leq \mathbf{b}$ exist, which proves the theorem.

□