

# A new robust optimization approach for scheduling under uncertainty

## II. Uncertainty with known probability distribution

Stacy L. Janak, Xiaoxia Lin, Christodoulos A. Floudas\*

*Department of Chemical Engineering, Princeton University,  
Princeton, NJ 08544-5263, USA*

Received 1 April 2005; received in revised form 3 May 2006; accepted 22 May 2006

Available online 1 August 2006

---

### Abstract

In this work, we consider the problem of scheduling under uncertainty where the uncertain problem parameters can be described by a known probability distribution function. A novel robust optimization methodology, originally proposed by Lin, Janak, and Floudas [Lin, X., Janak, S. L., & Floudas, C. A. (2004). A new robust optimization approach for scheduling under uncertainty: I. Bounded uncertainty. *Computers and Chemical Engineering*, 28, 1069–1085], is extended in order to consider uncertainty described by a known probability distribution. This robust optimization formulation is based on a min–max framework and when applied to mixed-integer linear programming (MILP) problems, produces “robust” solutions that are immune against data uncertainty. Uncertainty is considered in the coefficients of the objective function, as well as the coefficients and right-hand-side parameters of the inequality constraints in MILP problems. Robust optimization techniques are developed for uncertain data described by several known distributions including a uniform distribution, a normal distribution, the difference of two normal distributions, a general discrete distribution, a binomial distribution, and a poisson distribution. The robust optimization formulation introduces a small number of auxiliary variables and additional constraints into the original MILP problem, generating a deterministic robust counterpart problem which provides the optimal/feasible solution given the (relative) magnitude of the uncertain data, a feasibility tolerance, and a reliability level. The robust optimization approach is then applied to the problem of short-term scheduling under uncertainty. Using the continuous-time model for short-term scheduling developed by Floudas and co-workers [Ierapetritou, M. G. & Floudas, C. A. (1998a). Effective continuous-time formulation for short-term scheduling: 1. Multipurpose batch plants processes. *Ind. Eng. Chem. Res.*, 37, 4341–4359; Lin, X. & Floudas, C. A. (2001). Design, synthesis and scheduling of multipurpose batch plants via an effective continuous-time formulation. *Comp. Chem. Engng.*, 25, 665–674], three of the most common sources of uncertainty in scheduling problems are explored including processing times of tasks, market demands for products, and prices of products and raw materials. Computational results on several examples and an industrial case study are presented to demonstrate the effectiveness of the proposed approach.

© 2006 Published by Elsevier Ltd.

**Keywords:** Process scheduling; Uncertainty; Robust optimization; MILP; Probability distribution

---

### 1. Introduction

A significant body of work has appeared in the literature over the past decade in the research area of production scheduling. Most of the existing work assumes that all data are of known, constant values. However, in reality, uncertainty is common in many scheduling problems due to the lack of accurate process models and variability of process and environmental data. Thus, an emerging area of

---

\* Corresponding author. Tel.: +1 609 258 4595; fax: +1 609 258 0211.

E-mail address: [floudas@titan.princeton.edu](mailto:floudas@titan.princeton.edu) (C.A. Floudas).

research aims at developing methods to address the problem of scheduling under uncertainty, in order to create reliable schedules which remain feasible in the presence of parameter uncertainty (see the recent reviews by Floudas (2005) and Floudas and Lin (2004, 2005)). Different methodologies can be used for the problem of scheduling under uncertainty including stochastic, probabilistic, and fuzzy programming methods (Sahinidis, 2004).

Stochastic formulations incorporate uncertainty by taking advantage of the fact that probability distributions governing the data are known or can be estimated. The goal then is to find an optimal solution that is feasible for all, or almost all, of the instances of the uncertain parameters while maximizing the expectation of some function of the problem variables and the random variables. The most commonly studied stochastic programming models for scheduling problems are two-stage programs. For two-stage programs, in the first stage, all data are assumed to be known. Then, a random event occurs that affects the outcome of the first stage and a recourse decision is made in the second stage to compensate for any negative effects that might have been experienced as a result of the first stage decision. The result of two-stage programming models is a single first stage policy followed by a collection of recourse decisions that indicate which second stage policy should be implemented in response to each random outcome. Thus, as the number of uncertain parameters increases, more scenarios must be considered for the recourse decisions, resulting in a much larger problem. In fact, the number of scenarios increases exponentially with the number of uncertain parameters. This main drawback limits the application of these approaches to solve practical problems with a large number of uncertain parameters.

Bassett, Pekny, and Reklaitis (1997) employed stochastic methods to consider uncertainty in processing times, equipment reliability/availability, process yields, demands, and manpower changes of scheduling problems. They used Monte Carlo sampling to generate random instances of the uncertainty and determined a schedule for each instance. Then, distributions of aggregated properties are generated using the instances to infer operating policies that are robust. However, the formulation does not produce a robust schedule. Vin and Ierapetritou (2001) considered demand uncertainty in the short-term scheduling of multiproduct and multipurpose batch plants. They introduced several metrics to evaluate the robustness of a schedule and proposed a multiperiod programming model using extreme points of the demand range as scenarios to generate a single sequence of tasks with the minimal average makespan over all scenarios. Balasubramanian and Grossmann (2002) proposed a multiperiod mixed-integer linear programming (MILP) model for scheduling multistage flowshop plants with uncertain processing times described by discrete or continuous probability distributions. The objective is the minimization of the expected makespan and a special branch and bound algorithm is employed based on a lower bound generated by an aggregated probability model. Balasubramanian and Grossmann (2004) also considered the problem of scheduling under demand uncertainty. They utilized a multistage stochastic MILP model where some decisions are made without respect to the uncertainty and others are made upon realization of the uncertainty. They proposed an approximation strategy which solves a series of two-stage models within a shrinking-horizon approach. In addition, Jia and Ierapetritou (2004) used the idea of inference-based sensitivity analysis for MILP problems to determine the importance of different parameters and constraints in their scheduling model. The formulation provides a set of alternative schedules for the range of uncertain parameters under consideration. Bonfill, Bagajewicz, Espuña, and Puigjaner (2004) presented an approach to manage risk for scheduling with uncertain demands. They employed a two-stage stochastic optimization model which maximizes the expected profit and manages risk explicitly by considering a new objective as a control measure, leading to a multiobjective optimization formulation. Subsequently, Bonfill, Espuña, and Puigjaner (2005) extended this formulation to consider uncertain processing times. A two-stage stochastic approach is again utilized where a weighted sum of the expected makespan and the expected wait times is minimized and risk is measured using different robustness criteria. Ostrovsky, Datskov, Achenie, and Volin (2004) discussed extensions of the two-stage optimization model to take into account the possibility of accurately estimating some of the uncertain parameters. A split and bound approach is utilized to solve the problem and is based on a partitioning of the uncertain region and estimation of bounds on the objective function.

There have also been attempts to transform a stochastic model to a direct deterministic equivalent representation. Bertsimas and Sim (2004) defined a robust optimization formulation that allows for a degree of conservatism in every uncertain constraint. Their formulation is developed for uncertainty that is bounded and symmetric and a large number of parameters must be uncertain in each constraint for the resulting formulation to be desirably tight, or achieve solutions that are not too conservative. In addition, Bertsimas and Sim (2003) also used this robust optimization formulation to incorporate data uncertainty for discrete optimization and network flow problems where cost coefficients and the data in the constraints can be uncertain. The formulation allows for control over the degree of conservatism of the solution in terms of the probabilistic bounds on constraint violation.

Probabilistic or chance constraints have also been used to incorporate uncertainty into scheduling problems. Chance constraints do not require all decisions to be feasible for every outcome of the random variables, but require feasibility with at least some specified probability distribution. Orçun, Altinel, and Hortaçsu (1996) considered uncertain processing times in batch processes and employed chance constraints to account for the risk of violation of timing constraints under certain conditions such as uniform distribution functions. In addition, Petkov and Maranas (1997) proposed a stochastic model for scheduling with uncertain product demands which involves the maximization of the expected profit where single or multiple product demands are satisfied with a prespecified probability level using chance constraints.

Wang (2004) presented a robust optimization methodology based on fuzzy set theory for uncertain product development projects where the uncertain parameters in the model are represented by fuzzy sets. A genetic algorithm approach is used to solve the

problem where the measure of schedule robustness is based on qualitative possibility theory. Also, Balasubramanian and Grossmann (2003) used the concept of fuzzy set theory to describe the uncertainty in scheduling problems. They proposed an MILP model for uncertainty in processing times for flowshop scheduling problems and new product development process scheduling problems.

An alternative approach for scheduling under uncertainty is reactive scheduling in which an existing schedule is adjusted upon realization of the uncertain parameters or the occurrence of unexpected events. Due to the “on-line” nature of reactive scheduling, it is necessary to generate updated schedules in a timely manner and often, heuristic approaches are utilized (e.g., Cott & Macchietto, 1989; Honkomp, Mockus, & Reklaitis, 1999; Kanakamedala, Reklaitis, & Venkatasubramanian, 1994; Sanmarti, Huercio, Espuña, & Puigjaner, 1996; Rodrigues, Gimeno, Passos, & Campos, 1996; Vin & Ierapetritou, 2000). A recent review on scheduling approaches that includes reactive scheduling can be found in Floudas and Lin (2004).

In this work, we extend the robust optimization approach presented in Lin, Janak, and Floudas (2004) for scheduling under bounded uncertainty and bounded and symmetric uncertainty to consider uncertainty described by a known probability distribution function. The underlying mathematical framework is based on a robust optimization methodology first introduced for linear programming (LP) problems by and extended in Lin et al. (2004) and this work for mixed-integer linear programming problems. The approach produces “robust” solutions for uncertainties in both the coefficients and right-hand-side parameters of the linear inequality constraints and can be applied to address the problem of production scheduling with uncertain parameters. The rest of this paper is organized as follows. We will first review the problem of short-term scheduling under parameter uncertainty. Then the robust optimization approach is extended for the case of general MILP problems with uncertain parameters in the inequality constraints described by known probability distributions. Next, this approach is applied to the problem of short-term scheduling with uncertainty in processing times of tasks, product demands, and market prices. Finally, computational results are presented, followed by concluding remarks.

## 2. Robust optimization for MILP problems

Consider the following generic mixed-integer linear programming problem

$$\begin{aligned} \text{Min/Max}_{x,y} \quad & c^T x + d^T y \\ \text{s.t.} \quad & Ex + Fy = e \\ & Ax + By \leq p \\ & \underline{x} \leq x \leq \bar{x} \\ & y = 0, 1. \end{aligned} \quad (1)$$

Assume that the uncertainty arises from both the coefficients and the right-hand-side parameters of the inequality constraints, namely,  $a_{lm}$ ,  $b_{lk}$  and  $p_l$  where  $l$  is the index of the uncertain inequality,  $m$  is the index of the continuous terms, and  $k$  is the index of the binary terms. Thus, we are concerned about the feasibility of the following inequality

$$\sum_m \tilde{a}_{lm} x_m + \sum_k \tilde{b}_{lk} y_k \leq \tilde{p}_l \quad (2)$$

where  $a_{lm}$ ,  $b_{lk}$ , and  $p_l$  are the nominal values of the uncertain parameters and  $\tilde{a}_{lm}$ ,  $\tilde{b}_{lk}$ , and  $\tilde{p}_l$  are the “true” values of the uncertain parameters.

Assume that for inequality constraint  $l$ , the true values of the uncertain parameters are obtained from their nominal values by random perturbations

$$\begin{aligned} \tilde{a}_{lm} &= (1 + \epsilon \xi_{lm}) a_{lm} \\ \tilde{b}_{lk} &= (1 + \epsilon \xi_{lk}) b_{lk} \\ \tilde{p}_l &= (1 + \epsilon \xi_l) p_l \end{aligned} \quad (3)$$

where  $\xi_{lm}$ ,  $\xi_{lk}$  and  $\xi_l$  are independent random variables and  $\epsilon > 0$  is a given (relative) uncertainty level.

In this situation, we call a solution  $(x, y)$  robust if it satisfies the following

- (i)  $(x, y)$  is feasible for the nominal problem, and
- (ii) for every inequality  $l$ , the probability of violation of the uncertain inequality in Eq. (2) (i.e., the left-hand-side exceeds the right-hand-side) is at most  $\kappa$ ,

$$\Pr \left\{ \sum_m \tilde{a}_{lm} x_m + \sum_k \tilde{b}_{lk} y_k > \tilde{p}_l + \delta \max[1, |p_l|] \right\} \leq \kappa \quad (4)$$

where  $\delta > 0$  is a given feasibility tolerance and is introduced to allow a small amount of infeasibility in the uncertain inequality, and  $\kappa > 0$  is a given reliability level. Thus,  $\kappa$  represents the probability of violation of constraint  $l$  where  $\kappa = 0\%$  indicates that there is no chance of constraint violation, yielding the most conservative solution.

As shown in the first part of this work (Lin et al., 2004), the optimal solution of an MILP program may become severely infeasible, that is, one or more constraints are violated substantially, if the nominal data are slightly perturbed. This makes the “nominal” optimal solution questionable. Thus, in this work, our objective is to extend this robust optimization framework to be able to generate “reliable” solutions to the MILP program in Eq. (1), which are immune against uncertainty that can be described by a known probability distribution. This robust optimization methodology was first introduced for linear programming problems with uncertain linear coefficients by Ben-Tal and Nemirovski (2000) and was extended in Lin et al. (2004) to consider uncertainty in MILP problems.

It is important to emphasize that the robust optimization approach is different than the chance-constrained programming approach. Chance-constrained programming problems use constraints of the following type

$$\Pr\{g_i(x, \xi) \geq 0, \forall i\} \geq p \quad \text{or} \quad \Pr\{g_i(x, \xi) \geq 0\} \geq p_i, \quad \forall i \quad (5)$$

where  $x$  is a decision variable,  $\xi$  is a random variable,  $i$  indicates a finite set of inequalities,  $Pr$  is a probability measure, and  $p$  is the given probability level such that  $0 \leq p \leq 1$ . The first form describes the case where multiple constraints hold with a single probability while the second form, termed individual chance constraints, describes the case where multiple probabilistic constraints hold separately. Chance-constrained programming was first defined by Charnes and Cooper (1959, 1962, 1963) where the authors consider individual chance constraints and random variations are confined to the right-hand-side vector, or constant term. In contrast, robust optimization techniques, which were originally proposed by Ben-Tal, Nemirovski, and co-workers (Ben-Tal & Nemirovski, 1998, 1999, 2000; Ben-Tal, Goryashko, Guslitzer, & Nemirovski, 2004) and independently by El Ghaoui and co-workers (El Ghaoui, 1997; El Ghaoui, Oustry, & Lebret, 1998) and later extended by Bertsimas and Sim (2003, 2004, in press) and Floudas and co-workers (Lin et al., 2004), seek to determine a robust feasible/optimal solution to an uncertain problem. This means that the optimal solution should provide the best possible value of the original objection function and also be guaranteed to remain feasible in the range of the uncertainty set considered. Using this notion, the concept of robust optimization, as defined in these works, involves the following uncertain linear programming problem

$$\begin{aligned} \text{Min}_x \quad & c^T x \\ \text{s.t.} \quad & Ax \leq b \\ & \xi \equiv [A, b, c] \in Z \end{aligned} \quad (6)$$

where the data set  $\xi \equiv [A, b, c]$  varies in a given uncertainty set  $Z$ . Then the optimal solution which gives the best possible value of the original objection is defined as

$$\sup_{\xi \in Z} c^T x \quad (7)$$

where sup is the supremum of a function. Note that if the maximum of a function exists, then it must be finite and it occurs at the supremum of the function. Thus, the robust feasible/optimal solution can be found by solving the following problem

$$\begin{aligned} \text{Min}_x \quad & \text{Max}_{\xi \in Z} [c^T x] \\ \text{s.t.} \quad & A^T x - b \leq 0 \\ & \xi \equiv [A, b, c] \in Z \end{aligned} \quad (8)$$

called the robust counterpart problem. Note that this is a min–max problem. The uncertainty set  $Z$  is selected so as to balance the robustness and optimality of the solution. Thus, although the proposed robust optimization formulation uses individual probabilistic constraints, as seen in constraint (4), these constraints are employed to guarantee the reliability of the solution (or robustness of the solution) and not as initial problem constraints. As can be seen in the subsequent section, the probabilistic constraints are used in the proofs of the individual theorems deriving the deterministic formulations for each probability distribution and not in the formulations themselves as in individual chance-constrained programming.

### 2.1. Uncertainty with known probability distribution

If the probability distributions of the random variables  $\xi_{lm}$ ,  $\xi_{lk}$  and  $\xi_l$  in the uncertain parameters are known, it is possible to obtain a more accurate estimation of the probability measures involved. The MILP from (1) can be rewritten as an uncertain MILP

as follows

$$\begin{aligned}
 & \text{Min}_{x, y} \quad \text{Max}_{\xi \in Z} [c^T x + d^T y] \\
 & \text{s.t.} \quad Ex + Fy = e \\
 & \quad \quad \tilde{A}x + \tilde{B}y - \tilde{p} \leq \delta \cdot \max[1, |p|] \\
 & \quad \quad \underline{x} \leq x \leq \bar{x} \\
 & \quad \quad y = 0, 1 \\
 & \quad \quad \xi \equiv [A, B, p] \in Z
 \end{aligned} \tag{9}$$

where the data set  $\xi \equiv [A, B, p]$  varies in a given uncertainty set  $Z$ ,  $\tilde{A}$ ,  $\tilde{B}$ , and  $\tilde{p}$  represent the “true” values of the uncertain coefficients, and  $\delta \geq 0$  is an infeasibility tolerance introduced to allow a certain amount of infeasibility into the inequality constraint. The inequality can be written in expanded form as

$$\sum_m \tilde{a}_{lm} x_m + \sum_k \tilde{b}_{lk} y_k - \tilde{p}_l \leq \delta \cdot \max[1, |p_l|] \tag{10}$$

for every constraint  $l$  where  $\tilde{a}_{lm}$ ,  $\tilde{b}_{lk}$ , and  $\tilde{p}_l$  are again the true values of the uncertain coefficients. Substituting the expressions for the true values of the uncertain coefficients given in constraint (3) (i.e.,  $\tilde{a}_{lm}$ ,  $\tilde{b}_{lk}$ , and  $\tilde{p}_l$ ), the uncertain inequality in (10) can be rewritten as follows

$$\sum_m (1 + \epsilon \xi_{lm}) a_{lm} x_m + \sum_k (1 + \epsilon \xi_{lk}) b_{lk} y_k - (1 + \epsilon \xi_l) p_l \leq \delta \cdot \max[1, |p_l|]. \tag{11}$$

Rearranging terms, we get

$$\sum_m a_{lm} x_m + \sum_k b_{lk} y_k - p_l + \epsilon \left( \sum_{m \in M_l} \xi_{lm} a_{lm} x_m + \sum_{k \in K_l} \xi_{lk} b_{lk} y_k - \xi_l p_l \right) \leq \delta \cdot \max[1, |p_l|] \tag{12}$$

where  $M_l$  and  $K_l$  define the sets of uncertain parameters  $a_{lm}$  and  $b_{lk}$ , respectively, for constraint  $l$ . Then, a solution  $(x, y)$  to the original uncertain MILP given in Eq. (9) which satisfies this constraint is called “reliable” because it takes into account the maximum amount of uncertainty  $\xi \in Z$  and allows an amount of infeasibility  $\delta$ . Now, to transform the constraint into a deterministic form, we instead consider the following formulation

$$\Pr \left\{ \sum_m a_{lm} x_m + \sum_k b_{lk} y_k - p_l + \epsilon \left( \sum_{m \in M_l} \xi_{lm} a_{lm} x_m + \sum_{k \in K_l} \xi_{lk} b_{lk} y_k - \xi_l p_l \right) > \delta \cdot \max[1, |p_l|] \right\} \leq \kappa. \tag{13}$$

This constraint enforces that the probability of violation of the uncertain inequality is at most  $\kappa$ , where  $\delta \geq 0$  is a given feasibility tolerance (i.e., amount of error allowed in the feasibility of constraint  $l$ ) and  $\kappa \geq 0$  is a given reliability level (i.e., the probability of violation of constraint  $l$  where  $\kappa = 0$  indicates that there is no chance of constraint violation). Thus, if we know a probability distribution function for the sum of the random variables,

$$\xi = \sum_{m \in M_l} \xi_{lm} a_{lm} x_m + \sum_{k \in K_l} \xi_{lk} b_{lk} y_k - \xi_l p_l \tag{14}$$

we can use this information in the probabilistic constraint (13) to write a deterministic form for the uncertain constraint which is “almost reliable”, depending on the value of  $\kappa$ . This is done using the definition of a probability distribution function and the following relationship

$$F_\xi(\lambda) = \Pr\{\xi \leq \lambda\} = 1 - \Pr\{\xi > \lambda\} = 1 - \kappa \tag{15}$$

to replace the stochastic elements in constraint (12), generating a deterministic constraint that is “almost reliable” for the given uncertainty level,  $\epsilon$ , infeasibility tolerance,  $\delta$ , and reliability level,  $\kappa$ . The final form of the deterministic constraint (or robust counterpart problem) is simply determined using the inverse distribution function (quantile) of the random variable  $\xi$

$$F_\xi^{-1}(1 - \kappa) = f(\lambda, |a_{lm}|x_m, |b_{lk}|y_k, |p_l|). \tag{16}$$

Thus, the additional constraints in the robust counterpart (RC) problem can be written as

$$\sum_m a_{lm} x_m + \sum_k b_{lk} y_k + \epsilon f(\lambda, |a_{lm}|x_m, |b_{lk}|y_k, |p_l|) \leq p_l + \delta \max[1, |p_l|], \quad \forall l \tag{17}$$

where  $\epsilon$  is a given uncertainty level,  $\delta$  is a given infeasibility tolerance, and  $\lambda$  is determined from  $\kappa$  using constraint (15) and the probability distribution function for  $\xi$ .

In the following sections, robust optimization methods are developed for several of the most common probability distributions including the uniform distribution, normal distribution, difference of normal distributions, general discrete distribution, binomial distribution, and poisson distribution.

### 2.1.1. Uncertainty with uniform probability distribution

Assume that there is only one uncertain parameter in each constraint, which can be one of the following

$$\begin{aligned} \tilde{a}_{lm'} &= (1 + \epsilon \xi_l) a_{lm'} \\ \text{or } \tilde{b}_{lk'} &= (1 + \epsilon \xi_l) b_{lk'} \\ \text{or } \tilde{p}_l &= (1 + \epsilon \xi_l) p_l \end{aligned} \quad (18)$$

where  $\xi_l$  is a random variable with uniform distribution in the interval  $[-1, 1]$ .

**Theorem 1.** Given an uncertainty level ( $\epsilon$ ), an infeasibility tolerance ( $\delta$ ), and a reliability level ( $\kappa$ ), to generate robust solutions, the following ( $\epsilon, \delta, \kappa$ )-robust counterpart (RC[ $\epsilon, \delta, \kappa$ ]) of the original uncertain MILP problem can be derived.

$$\begin{aligned} \text{Min/Max} \quad & c^T x + d^T y \\ \text{s.t.} \quad & Ex + Fy = e \\ & Ax + By \leq p \\ & \sum_m a_{lm} x_m + \sum_k b_{lk} y_k + \epsilon(1 - 2\kappa) |a_{lm'}| u_{m'} \leq p_l + \delta \max[1, |p_l|] \quad \forall l \in L_c \\ & -u_{m'} \leq x_{m'} \leq u_{m'} \quad \forall l \in L_c, m' \in M_l \\ & \sum_m a_{lm} x_m + \sum_k b_{lk} y_k + \epsilon(1 - 2\kappa) |b_{lk'}| y_{k'} \leq p_l + \delta \max[1, |p_l|] \quad \forall l \in L_b \\ & \sum_m a_{lm} x_m + \sum_k b_{lk} y_k \leq p_l - \epsilon(1 - 2\kappa) |p_l| + \delta \max[1, |p_l|] \quad \forall l \in L_r \\ & \underline{x} \leq x \leq \bar{x} \\ & y_k = 0, 1 \quad \forall k \end{aligned} \quad (19)$$

where  $L_c, L_b, L_r$  are the set of inequalities with uncertainty in the coefficients of continuous variables, the coefficients of binary variables, and the right-hand-side parameters, respectively.

**Proof.** Let  $(x, y)$  satisfy the following:

If  $a_{lm'}$  is the uncertain parameter, then

$$\sum_m a_{lm} x_m + \sum_k b_{lk} y_k + \epsilon(1 - 2\kappa) |a_{lm'}| u_{m'} \leq p_l + \delta \max[1, |p_l|], \quad (20)$$

$$-u_{m'} \leq x_{m'} \leq u_{m'}, \quad (21)$$

or if  $b_{lk'}$  is the uncertain parameter, then

$$\sum_m a_{lm} x_m + \sum_k b_{lk} y_k + \epsilon(1 - 2\kappa) |b_{lk'}| y_{k'} \leq p_l + \delta \max[1, |p_l|], \quad (22)$$

or if  $p_l$  is the uncertain parameter, then

$$\sum_m a_{lm} x_m + \sum_k b_{lk} y_k \leq p_l - \epsilon(1 - 2\kappa) |p_l| + \delta \max[1, |p_l|]. \quad (23)$$

It follows that if  $a_{lm'}$  is the uncertain parameter,

$$\begin{aligned} & \Pr \left\{ \sum_{m \neq m'} a_{lm} x_m + \tilde{a}_{lm'} x_{m'} + \sum_k b_{lk} y_k > p_l + \delta \max[1, |p_l|] \right\} \\ &= \Pr \left\{ \sum_m a_{lm} x_m + \epsilon \xi_l |a_{lm'}| x_{m'} + \sum_k b_{lk} y_k > p_l + \delta \max[1, |p_l|] \right\} \leq \Pr \{ \xi_l x_{m'} > (1 - 2\kappa) u_{m'} \} \\ &\leq \Pr \{ \xi_l x_{m'} > (1 - 2\kappa) |x_{m'}| \} \quad (\text{note: we only need to consider } \kappa \leq 0.5) \\ &= 1 - \Pr \{ \xi_l x_{m'} \leq (1 - 2\kappa) |x_{m'}| \} = 1 - (1 - \kappa) = \kappa, \end{aligned}$$

or if  $b_{lk'}$  is the uncertain parameter,

$$\begin{aligned} & \Pr \left\{ \sum_m a_{lm} x_m + \sum_{k \neq k'} b_{lk} y_k + \tilde{b}_{lk'} y_{k'} > p_l + \delta \max[1, |p_l|] \right\} \\ &= \Pr \left\{ \sum_m a_{lm} x_m + \sum_k b_{lk} y_k + \epsilon \xi_l |b_{lk'}| y_{k'} > p_l + \delta \max[1, |p_l|] \right\} \leq \Pr \{ \xi_l > (1 - 2\kappa) \} = 1 - \Pr \{ \xi_l \leq (1 - 2\kappa) \} = \kappa, \end{aligned}$$

or if  $p_l$  is the uncertain parameter,

$$\begin{aligned} & \Pr \left\{ \sum_m a_{lm} x_m + \sum_k b_{lk} y_k > \tilde{p}_l + \delta \max[1, |p_l|] \right\} \\ &= \Pr \left\{ \sum_m a_{lm} x_m + \sum_k b_{lk} y_k > p_l + \epsilon \xi_l |p_l| + \delta \max[1, |p_l|] \right\} \leq \Pr \{ -\xi_l > (1 - 2\kappa) \} = \kappa. \quad \square \end{aligned}$$

Note that this formulation for uncertainty described by a uniform probability distribution results in an MILP problem and includes a set of auxiliary variables,  $u_m$ .

### 2.1.2. Uncertainty with normal probability distribution

Suppose that the distributions of the random variables  $\xi_{lm}$ ,  $\xi_{lk}$  and  $\xi_l$  in (3) are all standardized normal distributions with zero as the mean and one as the standard deviation. Then, the distribution of  $\xi$  defined in (14) is also a normal distribution, with zero as the mean and  $\sqrt{\sum_{m \in M_l} a_{lm}^2 x_m^2 + \sum_{k \in K_l} b_{lk}^2 y_k + p_l^2}$  as the standard deviation.

**Theorem 2.** Given an uncertainty level ( $\epsilon$ ), an infeasibility tolerance ( $\delta$ ), and a reliability level ( $\kappa$ ), to generate robust solutions, the following ( $\epsilon, \delta, \kappa$ )-robust counterpart (RC[ $\epsilon, \delta, \kappa$ ]) of the original uncertain MILP problem can be derived.

$$\begin{aligned} & \text{Min/Max}_{x,y} \quad c^T x + d^T y \\ & \text{s.t.} \quad Ex + Fy = e \\ & \quad \quad Ax + By \leq p \\ & \quad \quad \sum_m a_{lm} x_m + \sum_k b_{lk} y_k + \epsilon \lambda \sqrt{\sum_{m \in M_l} a_{lm}^2 x_m^2 + \sum_{k \in K_l} b_{lk}^2 y_k + p_l^2} \leq p_l + \delta \max[1, |p_l|] \quad \forall l \\ & \quad \quad \underline{x} \leq x \leq \bar{x} \\ & \quad \quad y_k = 0, 1 \quad \forall k \end{aligned} \tag{24}$$

where  $\lambda = F_n^{-1}(1 - \kappa)$  and  $F_n^{-1}$  is the inverse distribution function of a random variable with standardized normal distribution. Thus,  $\lambda$  and  $\kappa$  are related as follows

$$\begin{aligned} \kappa &= 1 - F_n(\lambda) \\ \kappa &= 1 - \Pr\{\xi \leq \lambda\} \\ \kappa &= 1 - \int_{-\infty}^{\lambda} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right) dx \end{aligned}$$

where  $\xi$  is a random variable with standardized normal distribution.

**Proof.** Let  $(x, y)$  satisfy the following

$$\sum_m a_{lm} x_m + \sum_k b_{lk} y_k + \epsilon \lambda \sqrt{\sum_{m \in M_l} a_{lm}^2 x_m^2 + \sum_{k \in K_l} b_{lk}^2 y_k + p_l^2} \leq p_l + \delta \max[1, |p_l|] \tag{25}$$

where  $\lambda = F_n^{-1}(1 - \kappa)$  and  $F_n^{-1}$  is the inverse distribution function of a random variable with standardized normal distribution. Then

$$\begin{aligned} & \Pr \left\{ \sum_m \tilde{a}_{lm} x_m + \sum_k \tilde{b}_{lk} y_k > \tilde{p}_l + \delta \max[1, |p_l|] \right\} \\ &= \Pr \left\{ \sum_m a_{lm} x_m + \epsilon \sum_{m \in M_l} \xi_{lm} |a_{lm}| x_m + \sum_k b_{lk} y_k + \epsilon \sum_{k \in K_l} \xi_{lk} |b_{lk}| y_k > p_l + \epsilon \xi_l |p_l| + \delta \max[1, |p_l|] \right\} \end{aligned}$$



$$\begin{aligned}
&\leq \Pr \left\{ \left( \sum_{m \in M_l} \xi_{lm} |a_{lm}| x_m + \sum_{k \in K_l} \xi_{lk} |b_{lk}| y_k - \xi_l |p_l| \right) / \sqrt{\sum_{m \in M_l} a_{lm}^2 x_m^2 + \sum_{k \in K_l} b_{lk}^2 y_k + p_l^2} > \lambda \right\} \\
&= 1 - \Pr \left\{ \left( \sum_{m \in M_l} \xi_{lm} |a_{lm}| x_m + \sum_{k \in K_l} \xi_{lk} |b_{lk}| y_k - \xi_l |p_l| \right) / \sqrt{\sum_{m \in M_l} a_{lm}^2 x_m^2 + \sum_{k \in K_l} b_{lk}^2 y_k + p_l^2} \leq \lambda \right\} \\
&= 1 - F_n(\lambda) = 1 - (1 - \kappa) = \kappa. \quad \square
\end{aligned}$$

Note that

$$\frac{\sum_{m \in M_l} \xi_{lm} |a_{lm}| x_m + \sum_{k \in K_l} \xi_{lk} |b_{lk}| y_k - \xi_l |p_l|}{\sqrt{\sum_{m \in M_l} a_{lm}^2 x_m^2 + \sum_{k \in K_l} b_{lk}^2 y_k + p_l^2}}$$

is also a random variable with standardized normal distribution. This formulation results in a convex MINLP problem, but can still be solved efficiently using a mixed-integer nonlinear solver (e.g., DICOPT (Viswanathan & Grossmann, 1990), MINOPT (Schweiger & Floudas, 1998)).

### 2.1.3. Uncertainty with difference of normal probability distributions

Suppose that the distributions of random variables  $\xi_{lm}$ ,  $\xi_{lk}$  and  $\xi_l$  in (3) are all represented as the difference of two normal random variables, each with a given mean,  $\mu$ , and a given standard deviation,  $\sigma$ . Then, the random variables can be represented by

$$\begin{aligned}
\xi_{lm} &= A_{lm} \xi_{lmx} - B_{lm} \xi_{lmy}, \\
\xi_{lk} &= A_{lk} \xi_{lkx} - B_{lk} \xi_{lky}, \\
\xi_l &= A_l \xi_{lx} - B_l \xi_{ly}
\end{aligned} \tag{26}$$

where each has a mean  $\mu$  and a standard deviation  $\sigma$ . Furthermore, the overall random variable  $\xi$  defined in (14) is also given by a normal distribution and has a mean and standard deviation of

$$\mu = \sum_{m \in M_l} |a_{lm}| x_m (A_{lm} \mu_{lmx} - B_{lm} \mu_{lmy}) + \sum_{k \in K_l} |b_{lk}| y_k (A_{lk} \mu_{lkx} - B_{lk} \mu_{lky}) - p_l (A_l \mu_{lx} - B_l \mu_{ly}) \tag{27}$$

$$\sigma = \sum_{m \in M_l} a_{lm}^2 x_m^2 (A_{lm}^2 \sigma_{lmx}^2 + B_{lm}^2 \sigma_{lmy}^2) + \sum_{k \in K_l} b_{lk}^2 y_k^2 (A_{lk}^2 \sigma_{lkx}^2 + B_{lk}^2 \sigma_{lky}^2) + p_l^2 (A_l^2 \sigma_{lx}^2 + B_l^2 \sigma_{ly}^2). \tag{28}$$

We can then standardize the random variable  $\xi$  so that it has zero as the mean and one as the standard deviation as follows

$$\xi^{\text{std}} = \frac{(\xi - \mu)}{\sqrt{\sigma}}. \tag{29}$$

**Theorem 3.** Given an uncertainty level ( $\epsilon$ ), an infeasibility tolerance ( $\delta$ ), and a reliability level ( $\kappa$ ), to generate robust solutions, the following ( $\epsilon, \delta, \kappa$ )-robust counterpart (RC[ $\epsilon, \delta, \kappa$ ]) of the original uncertain MILP problem can be derived.

$$\begin{aligned}
&\text{Min/Max}_{x,y} \quad c^T x + d^T y \\
&\text{s.t.} \quad Ex + Fy = e \\
&\quad \quad Ax + By \leq p \\
&\quad \quad \sum_m a_{lm} x_m + \sum_k b_{lk} y_k + \epsilon(\lambda \sqrt{\sigma} + \mu) \leq p_l + \delta \max[1, |p_l|] \quad \forall l \\
&\quad \quad \underline{x} \leq x \leq \bar{x} \\
&\quad \quad y_k = 0, 1 \quad \forall k.
\end{aligned} \tag{30}$$

where  $\lambda = F_n^{-1}(1 - \kappa)$  and  $F_n^{-1}$  is the inverse distribution function of a random variable with standardized normal distribution and  $\mu$  and  $\sigma$  are described as given in Eqs. (27) and (28).

**Proof.** Let  $(x, y)$  satisfy the following

$$\sum_m a_{lm} x_m + \sum_k b_{lk} y_k + \epsilon(\lambda \sqrt{\sigma} + \mu) \leq p_l + \delta \max[1, |p_l|] \tag{31}$$



where  $\lambda = F_n^{-1}(1 - \kappa)$  and  $F_n^{-1}$  is the inverse distribution function of a random variable with a standardized normal distribution. Then

$$\begin{aligned}
 & \Pr \left\{ \sum_m \tilde{a}_{lm} x_m + \sum_k \tilde{b}_{lk} y_k > \tilde{p}_l + \delta \max[1, |p_l|] \right\} \\
 &= \Pr \left\{ \sum_m a_{lm} x_m + \epsilon \sum_{m \in M_l} \xi_{lm} |a_{lm}| x_m + \sum_k b_{lk} y_k + \epsilon \sum_{k \in K_l} \xi_{lk} |b_{lk}| y_k > p_l + \epsilon \xi_l |p_l| + \delta \max[1, |p_l|] \right\} \\
 &\leq \Pr \left\{ \left( \sum_{m \in M_l} \xi_{lm} |a_{lm}| x_m + \sum_{k \in K_l} \xi_{lk} |b_{lk}| y_k - \xi_l |p_l| - \mu \right) / \sqrt{\sigma} > \lambda \right\} \\
 &= 1 - \Pr \left\{ \left( \sum_{m \in M_l} \xi_{lm} |a_{lm}| x_m + \sum_{k \in K_l} \xi_{lk} |b_{lk}| y_k - \xi_l |p_l| - \mu \right) / \sqrt{\sigma} \leq \lambda \right\} = 1 - F_n(\lambda) = 1 - (1 - \kappa) = \kappa. \quad \square
 \end{aligned}$$

This formulation results in a convex MINLP problem, but can still be solved efficiently using a mixed-integer nonlinear solver (e.g., DICOPT (Viswanathan & Grossmann, 1990), MINOPT (Schweiger & Floudas, 1998)).

#### 2.1.4. Uncertainty with general discrete probability distribution

Suppose that the distributions of the random variables  $\xi_{lm}$ ,  $\xi_{lk}$  and  $\xi_l$  in (3) are all given by general discrete distributions. Then the distribution of  $\xi$  defined in (14) is also a discrete distribution.

**Theorem 4.** Given an uncertainty level ( $\epsilon$ ), an infeasibility tolerance ( $\delta$ ), and a reliability level ( $\kappa$ ), to generate robust solutions, the following ( $\epsilon, \delta, \kappa$ )-robust counterpart (RC[ $\epsilon, \delta, \kappa$ ]) of the original uncertain MILP problem can be derived.

$$\begin{aligned}
 & \text{Min/Max}_{x,y} \quad c^T x + d^T y \\
 & \text{s.t.} \quad Ex + Fy = e \\
 & \quad \quad Ax + By \leq p \\
 & \quad \quad \sum_m a_{lm} x_m + \sum_k b_{lk} y_k + \epsilon \left( \sum_{m \in M_l} \lambda |a_{lm}| u_m + \sum_{k \in K_l} \lambda |b_{lk}| y_k - \lambda |p_l| \right) \leq p_l + \delta \max[1, |p_l|] \quad \forall l \\
 & \quad \quad -u_m \leq x_m \leq u_m \quad \forall m \\
 & \quad \quad \underline{x} \leq x \leq \bar{x} \\
 & \quad \quad y_k = 0, 1 \quad \forall k
 \end{aligned} \tag{32}$$

where  $\kappa = 1 - F(\lambda)$  and where  $F$  is the distribution function of a general discrete random variable.

**Proof.** Let  $(x, y)$  satisfy the following

$$\sum_m a_{lm} x_m + \epsilon \sum_{m \in M_l} \lambda |a_{lm}| u_m + \sum_k b_{lk} y_k + \epsilon \sum_{k \in K_l} \lambda |b_{lk}| y_k \leq p_l + \epsilon \lambda |p_l| + \delta \max[1, |p_l|], \tag{33}$$

$$-u_m \leq x_m \leq u_m \tag{34}$$

where  $M_l$  and  $K_l$  are the set of indices of the  $x$  and  $y$  variables, respectively, with uncertain coefficients in the  $l$ th inequality constraint and  $F(\lambda) = 1 - \kappa$  and  $\kappa > 0$  is a given reliability level. Then

$$\begin{aligned}
 & \Pr \left\{ \sum_m \tilde{a}_{lm} x_m + \sum_k \tilde{b}_{lk} y_k > \tilde{p}_l + \delta \max[1, |p_l|] \right\} \\
 &= \Pr \left\{ \sum_m a_{lm} x_m + \epsilon \sum_{m \in M_l} \xi_{lm} |a_{lm}| x_m + \sum_k b_{lk} y_k + \epsilon \sum_{k \in K_l} \xi_{lk} |b_{lk}| y_k > p_l + \epsilon \xi_l |p_l| + \delta \max[1, |p_l|] \right\} \\
 &\leq \Pr \left\{ \sum_{m \in M_l} \xi_{lm} |a_{lm}| x_m + \sum_{k \in K_l} \xi_{lk} |b_{lk}| y_k - \xi_l |p_l| > \sum_{m \in M_l} \lambda |a_{lm}| u_m + \sum_{k \in K_l} \lambda |b_{lk}| y_k - \lambda |p_l| \right\}
 \end{aligned}$$

$$\begin{aligned}
&= 1 - \Pr \left\{ \sum_{m \in M_l} \xi_{lm} |a_{lm}| x_m + \sum_{k \in K_l} \xi_{lk} |b_{lk}| y_k - \xi_l |p_l| \leq \sum_{m \in M_l} \lambda |a_{lm}| u_m + \sum_{k \in K_l} \lambda |b_{lk}| y_k - \lambda |p_l| \right\} \\
&= 1 - F(\lambda) = 1 - (1 - \kappa) = \kappa. \quad \square
\end{aligned}$$

Note that this formulation for uncertainty described by a general discrete distribution results in an MILP problem and includes a set of auxiliary variables,  $u_m$ .

### 2.1.5. Uncertainty with binomial probability distribution

Assume that in each inequality constraint  $l$ , there is at most one uncertain parameter, which can be the coefficient of a continuous variable, or the coefficient of a binary variable, or the right-hand-side parameter. The true value of the uncertain parameter is obtained from its nominal value by random perturbation

$$\begin{aligned}
&\tilde{a}_{lm'} = (1 + \epsilon \xi_l) a_{lm'} \\
&\text{or } \tilde{b}_{lk'} = (1 + \epsilon \xi_l) b_{lk'} \\
&\text{or } \tilde{p}_l = (1 + \epsilon \xi_l) p_l
\end{aligned} \tag{35}$$

where  $\xi_l$  is a discrete random variable with a binomial distribution.

**Theorem 5.** Given an uncertainty level ( $\epsilon$ ), an infeasibility tolerance ( $\delta$ ), and a reliability level ( $\kappa$ ), to generate robust solutions, the following ( $\epsilon, \delta, \kappa$ )-robust counterpart (RC[ $\epsilon, \delta, \kappa$ ]) of the original uncertain MILP problem can be derived.

$$\begin{aligned}
&\text{Min/Max}_{x,y,u} \quad c^T x + d^T y \\
&\text{s.t.} \quad Ex + Fy = e \\
&\quad \quad Ax + By \leq p \\
&\quad \quad \sum_m a_{lm} x_m + \sum_k b_{lk} y_k + \epsilon \lambda |a_{lm'}| u_{m'} \leq p_l + \delta \max[1, |p_l|] \quad \forall l \in L_c \\
&\quad \quad -u_{m'} \leq x_{m'} \leq u_{m'} \quad \forall l \in L_c, m' \in M_l \\
&\quad \quad \sum_m a_{lm} x_m + \sum_k b_{lk} y_k + \epsilon \lambda |b_{lk'}| y_{k'} \leq p_l + \delta \max[1, |p_l|] \quad \forall l \in L_b \\
&\quad \quad \sum_m a_{lm} x_m + \sum_k b_{lk} y_k \leq p_l - \epsilon \lambda |p_l| + \delta \max[1, |p_l|] \quad \forall l \in L_r \\
&\quad \quad \underline{x} \leq x \leq \bar{x} \\
&\quad \quad y_k = 0, 1 \quad \forall k
\end{aligned} \tag{36}$$

where  $L_c, L_b, L_r$  are the set of inequalities with uncertainty in the coefficients of continuous variables, the coefficients of binary variables, and the right-hand-side parameters, respectively.

**Proof.** Let  $(x, y)$  satisfy the following: If  $a_{lm'}$  is the uncertain parameter, then

$$\sum_m a_{lm} x_m + \sum_k b_{lk} y_k + \epsilon \lambda |a_{lm'}| u_{m'} \leq p_l + \delta \max[1, |p_l|], \tag{37}$$

$$-u_{m'} \leq x_{m'} \leq u_{m'}, \tag{38}$$

or if  $b_{lk'}$  is the uncertain parameter, then

$$\sum_m a_{lm} x_m + \sum_k b_{lk} y_k + \epsilon \lambda |b_{lk'}| y_{k'} \leq p_l + \delta \max[1, |p_l|], \tag{39}$$

or if  $p_l$  is the uncertain parameter, then

$$\sum_m a_{lm} x_m + \sum_k b_{lk} y_k \leq p_l - \epsilon \lambda |p_l| + \delta \max[1, |p_l|]. \tag{40}$$

where  $\lambda = F^{-1}(1 - \kappa)$  and  $F^{-1}$  is the inverse distribution function of a discrete random variable with a binomial distribution.

It follows that if  $a_{lm'}$  is the uncertain parameter,

$$\begin{aligned}
&\Pr \left\{ \sum_{m \neq m'} a_{lm} x_m + \tilde{a}_{lm'} x_{m'} + \sum_k b_{lk} y_k > p_l + \delta \max[1, |p_l|] \right\} \\
&= \Pr \left\{ \sum_m a_{lm} x_m + \epsilon \xi_l |a_{lm'}| x_{m'} + \sum_k b_{lk} y_k > p_l + \delta \max[1, |p_l|] \right\} \leq \Pr \{ \xi_l x_{m'} > \lambda u_{m'} \} \\
&= 1 - \Pr \{ \xi_l x_{m'} \leq \lambda u_{m'} \} = 1 - (1 - \kappa) = \kappa,
\end{aligned}$$

or if  $b_{lk'}$  is the uncertain parameter,

$$\begin{aligned} & \Pr \left\{ \sum_m a_{lm} x_m + \sum_{k \neq k'} b_{lk} y_k + \tilde{b}_{lk'} y_{k'} > p_l + \delta \max[1, |p_l|] \right\} \\ &= \Pr \left\{ \sum_m a_{lm} x_m + \sum_k b_{lk} y_k + \epsilon \xi_l |b_{lk'}| y_{k'} > p_l + \delta \max[1, |p_l|] \right\} \leq \Pr \{ \xi_l > \lambda \} = 1 - \Pr \{ \xi_l \leq \lambda \} = \kappa, \end{aligned}$$

or if  $p_l$  is the uncertain parameter,

$$\begin{aligned} & \Pr \left\{ \sum_m a_{lm} x_m + \sum_k b_{lk} y_k > \tilde{p}_l + \delta \max[1, |p_l|] \right\} \\ &= \Pr \left\{ \sum_m a_{lm} x_m + \sum_k b_{lk} y_k > p_l + \epsilon \xi_l |p_l| + \delta \max[1, |p_l|] \right\} \leq \Pr \{ -\xi_l > \lambda \} = \kappa. \quad \square \end{aligned}$$

Note that this formulation for uncertainty described by a binomial distribution results in an MILP problem and includes a set of auxiliary variables,  $u_m$ .

### 2.1.6. Uncertainty with poisson probability distribution

The derivation for uncertainty with a poisson distribution is analogous to that of uncertainty with a binomial distribution, as given in Eq. (36) with the only difference being that  $F^{-1}$  is the inverse distribution function of a discrete random variable with a poisson distribution.

In the discussion above, for simplicity, we have assumed that there is a single common uncertainty level ( $\epsilon$ ), infeasibility tolerance ( $\delta$ ), and reliability level ( $\kappa$ ) in each MILP or convex MINLP problem with uncertain parameters. However, the proposed robust optimization techniques can easily be extended to account for the more general case in which the uncertainty level varies from one parameter to another and the infeasibility tolerance and reliability level are dependent on the constraint of interest. Furthermore, note that for each type of uncertainty addressed above, one additional constraint is introduced for each inequality constraint with uncertain parameter(s) and auxiliary variables are added if needed. Because the transformation is carried out at the level of constraints, in principle, the various robust optimization techniques presented can be applied to a single MILP or convex MINLP problem involving different types of uncertainties. More specifically, for each inequality constraint, as long as all of its uncertain parameters are of the same type, an additional constraint that corresponds to the uncertainty type can be introduced to obtain the deterministic robust counterpart problem.

It should be pointed out that the aforementioned robust optimization methodology circumvents any need for explicit or implicit discretization or sampling of the uncertain data, avoiding an undesirable increase in the size of the problem. Thus, the proposed methodology is potentially capable of handling problems with a large number of uncertain parameters. In addition, the resulting conservatism of a robust solution is determined by the values of the uncertainty level ( $\epsilon$ ), infeasibility tolerance ( $\delta$ ), and reliability level ( $\kappa$ ) used in the uncertain inequalities. As the uncertainty level increases, the uncertain inequalities become more difficult to satisfy resulting in a more conservative robust solution with a worse objective function value. In contrast, as the infeasibility tolerance increases, the uncertain inequalities tolerate a higher level of infeasibility, generating less conservative solutions with better objective function values. Similarly, as the reliability level increases (or the probability of violation increases), the amount of violation allowed in the uncertain inequalities increases, generating less conservative solutions with better objective function values. In this way, parametric studies of the effects of each of these parameters can be carried out, determining the tradeoffs between the level uncertainty and reliability with the quality and conservatism of the robust solution.

## 3. Robust optimization for scheduling under uncertainty

The robust optimization methodology proposed in the previous section can now be used to study the effects of uncertainty in short-term scheduling problems. In order to model this scheduling problem, we use the continuous-time formulation developed by Floudas and co-workers (Floudas & Lin, 2004, 2005; Ierapetritou & Floudas, 1998a, 1998b; Ierapetritou, Hené, & Floudas, 1999; Janak, Lin, & Floudas, 2004; Lin & Floudas, 2001; Lin, Floudas, Modi, & Juhasz, 2002; Lin, Chajakis, & Floudas, 2003), which results in MILP models (see Appendix A for the complete scheduling formulation). This continuous-time formulation has also been extended to determine the medium-term production scheduling of a large-scale, industrial batch plant utilizing actual plant operating data (Janak, Floudas, Kallrath, & Vormbrock, 2006a, 2006b). In addition, a comparative study of several continuous-time formulations for short-term scheduling can be found in Shaik, Janak, and Floudas (in press). Let us consider uncertainty in three classes of parameters that participate in short-term scheduling problems, namely: (i) uncertainty in processing

times/rates of tasks, (ii) uncertainty in market demands for products, and (iii) uncertainty in market prices of products and/or raw materials.

### 3.1. Problem statement for short-term scheduling

The problem of short-term scheduling for chemical processes under uncertainty is defined as follows. Given (i) the production recipes (i.e., the processing times for each task in the suitable units and the amount of material required for the production of each product), (ii) the available equipment and the ranges of their capacities, (iii) the material storage policy, (iv) the production requirements, and (v) the time horizon under consideration, we want to determine (i) the optimal sequence of tasks taking place in each unit, (ii) the amount of material being processed at each time in each unit, and (iii) the processing time of each task in each unit, so as to optimize a performance criterion, for example, to minimize the makespan or to maximize the overall profit while taking into account uncertainty inherent in some of the process parameters.

The most common sources of uncertainty in scheduling problems are (i) the processing times of tasks, (ii) the market demands for products, and (iii) the prices of products and/or raw materials. An uncertain parameter can be described using a discrete or continuous distribution and in some cases, only limited knowledge about the distribution is available. In the best situation, the distribution function for the uncertain parameter is given, for instance, as a normal distribution with known mean and standard deviation or as a uniform distribution in a given range. In this work, we will focus on uncertainty characterized by a known probability distribution function including combinations of probability distributions within a single problem.

### 3.2. Uncertainty in processing times

The parameters of processing times/rates of tasks appear in the duration constraint and appear as linear coefficients of the binary variable (i.e.,  $\alpha_{ij}$ ) and the continuous variable (i.e.,  $\beta_{ij}$ ) as follows

$$T^f(i, j, n) - T^s(i, j, n) = \alpha_{ij} \cdot wv(i, n) + \beta_{ij} \cdot B(i, j, n) \quad (41)$$

where  $wv(i, n)$  is a binary variable indicating whether or not task ( $i$ ) starts at event point ( $n$ ),  $B(i, j, n)$  is a continuous variable determining the batch-size of the task, and  $T^s(i, j, n)$  and  $T^f(i, j, n)$  are continuous variables representing the starting and finishing time of the task, respectively. Note that this is an equality constraint. Thus, in order to apply the robust optimization techniques proposed in Section 2 for inequality constraints with uncertain parameters, two separate approaches are developed.

#### Approach 1

In the first approach, the duration constraint is relaxed to an inequality constraint

$$T^f(i, j, n) - T^s(i, j, n) \geq \alpha_{ij} \cdot wv(i, n) + \beta_{ij} \cdot B(i, j, n). \quad (42)$$

Consequently, the variable  $T^f(i, j, n)$  represents the lower bound on the finishing time of the task, instead of the exact finishing time as determined by the original duration constraint. Using this modified duration constraint, the various robust optimization techniques can be readily applied to consider uncertainty in the parameters  $\alpha_{ij}$  and  $\beta_{ij}$ .

For example, consider a batch task with fixed processing time represented by parameter  $\alpha_{ij}$ . Then the true value of the processing time can be represented in terms of the nominal processing time as follows

$$\tilde{\alpha}_{ij} = (1 + \epsilon \xi_{\alpha_{ij}}) \alpha_{ij} \quad (43)$$

where  $\xi_{\alpha_{ij}}$  is a random variable with known distribution.

For the case where the uncertainty is uniform in the interval  $[-1, 1]$ , then according to [Theorem 1](#), to obtain the deterministic robust counterpart problem, the following constraint is added to the original scheduling model

$$T^s(i, j, n) - T^f(i, j, n) + [1 + \epsilon(1 - 2\kappa)] \alpha_{ij} \cdot wv(i, n) \leq \delta. \quad (44)$$

As another example, consider the case where the uncertainty is characterized by a standardized normal distribution. Then, according to [Theorem 2](#), to obtain the deterministic robust counterpart problem, the following constraint is added to the original scheduling model

$$T^s(i, j, n) - T^f(i, j, n) + [1 + \epsilon\lambda] \alpha_{ij} \cdot wv(i, n) \leq \delta \quad (45)$$

where  $\lambda = F^{-1}(1 - \kappa)$  and  $F_n^{-1}$  is the inverse distribution function of a random variable with a standardized normal distribution.

In addition, for the case where the uncertainty is characterized by a discrete binomial distribution, then according to [Theorem 5](#), to obtain the deterministic robust counterpart problem, the following constraint is added to the original scheduling model

$$T^s(i, j, n) - T^f(i, j, n) + [1 + \epsilon\lambda] \alpha_{ij} \cdot wv(i, n) \leq \delta \quad (46)$$

where  $\lambda = F^{-1}(1 - \kappa)$  and  $F_n^{-1}$  is the inverse distribution function of a random variable with a binomial distribution.

### Approach 2

In the second approach, the original duration constraint (41) is eliminated from the scheduling model and the variable  $T^f(i, j, n)$  is substituted into the other timing constraints using the following expression

$$T^f(i, j, n) = T^s(i, j, n) + \alpha_{ij} \cdot wv(i, n) + \beta_{ij} \cdot B(i, j, n). \quad (47)$$

The timing constraints which include the variable  $T^f(i, j, n)$  are the sequencing constraints and the time horizon constraints

$$T^s(i, j, n+1) \geq T^f(i, j, n) \quad \forall i \in I, j \in J_i, n \in N, n \neq N, \quad (48)$$

$$T^s(i, j, n+1) \geq T^f(i', j, n) - H[1 - wv(i', n)] \quad \forall j \in J, i, i' \in I_j, i \neq i', n \in N, n \neq N, \quad (49)$$

$$T^s(i, j, n+1) \geq T^f(i', j', n) - H[1 - wv(i', n)] \quad \forall j, j' \in J, i \in I_j, i' \in I_{j'}, i \neq i', n \in N, n \neq N, \quad (50)$$

$$T^f(i, j, n) \leq H, \quad \forall i \in I, j \in J_i, n \in N. \quad (51)$$

After the substitution, they become

$$T^s(i, j, n+1) \geq T^s(i, j, n) + \alpha_{ij} \cdot wv(i, n) + \beta_{ij} \cdot B(i, j, n) \quad \forall i \in I, j \in J_i, n \in N, n \neq N, \quad (52)$$

$$T^s(i, j, n+1) \geq T^s(i', j, n) + \alpha_{i'j} \cdot wv(i', n) + \beta_{i'j} \cdot B(i', j, n) - H[1 - wv(i', n)] \quad \forall j \in J, i, i' \in I_j, i \neq i', n \in N, n \neq N, \quad (53)$$

$$T^s(i, j, n+1) \geq T^s(i', j', n) + \alpha_{i'j'} \cdot wv(i', n) + \beta_{i'j'} \cdot B(i', j', n) - H[1 - wv(i', n)] \quad \forall j, j' \in J, i \in I_j, i' \in I_{j'}, i \neq i', n \in N, n \neq N, \quad (54)$$

$$T^s(i, j, n) + \alpha_{ij} \cdot wv(i, n) + \beta_{ij} \cdot B(i, j, n) \leq H, \quad \forall i \in I, j \in J_i, n \in N. \quad (55)$$

Now the uncertain parameters  $\alpha_{ij}$  and  $\beta_{ij}$  participate only in these inequality constraints and the robust optimization techniques can be readily applied as in approach 1.

### 3.3. Uncertainty in product demands

The product demands (i.e.,  $\text{dem}_s$ ) appear as the right-hand-side parameters in the demand constraints

$$\text{STF}(s) \geq \text{dem}_s, \quad \forall s \in S_p \quad (56)$$

where  $\text{STF}(s)$  is a continuous variable representing the amount of state ( $s$ ) accumulated at the end of the time horizon and  $S_p$  is the set of final products.

If we consider an uncertain product demand represented by parameter  $\text{dem}_s$ , then the true value of the product demand can be represented in terms of the nominal product demand as follows

$$\tilde{\text{dem}}_s = (1 + \epsilon \xi_s) \text{dem}_s \quad (57)$$

where  $\xi_s$  is a random variable with known distribution.

For the case of uniform uncertainty in the product demands, then according to [Theorem 1](#), the constraint to be added to the original scheduling model to derive the deterministic robust counterpart problem is

$$\text{STF}(s) \geq \text{dem}_s(1 + \epsilon(1 - 2\kappa) - \delta). \quad (58)$$

In addition, in the case of normal uncertainty, according to [Theorem 2](#), the constraint to be added to the original scheduling model to derive the deterministic robust counterpart problem is

$$\text{STF}(s) \geq \text{dem}_s(1 + \epsilon\lambda - \delta) \quad (59)$$

where  $\lambda = F_n^{-1}(1 - \kappa)$  and  $F_n^{-1}$  is the inverse distribution function of a random variable with standardized normal distribution.

Moreover, in the case of discrete uncertainty, according to [Theorem 4](#), the constraint to be added to the original scheduling model to derive the deterministic robust counterpart problem is

$$\text{STF}(s) \geq \text{dem}_s(1 + \epsilon\lambda - \delta) \quad (60)$$

where  $\lambda = F_n^{-1}(1 - \kappa)$  and  $F_n^{-1}$  is the inverse distribution function of a random variable with a general discrete distribution.

### 3.4. Uncertainty in market prices

The market prices (i.e.,  $p_s$ ) participate in the objective function for the calculation of the overall profit

$$\text{Maximize Profit} = \sum_{s \in S_p} p_s \cdot \text{STF}(s) - \sum_{s \in S_r} p_s \cdot \text{STI}(s) \quad (61)$$

where  $S_p$  and  $S_r$  are the sets of final products and raw materials, respectively, and  $\text{STI}(s)$  and  $\text{STF}(s)$  are continuous variables representing the initial amount of state ( $s$ ) at the beginning and the final amount of state ( $s$ ) at the end, respectively. The objective function can be expressed in an equivalent way as follows

$$\begin{aligned} &\text{Maximize Profit} \\ \text{s.t.} \quad &\text{Profit} \leq \sum_{s \in S_p} p_s \cdot \text{STF}(s) - \sum_{s \in S_r} p_s \cdot \text{STI}(s). \end{aligned} \quad (62)$$

Now the uncertain parameters  $p_s$  appear as linear coefficients multiplying the continuous variables  $\text{STF}(s)$  and  $\text{STI}(s)$  in an inequality constraint and the robust optimization techniques can be readily applied. For example, if the uncertainty is normally distributed

$$\tilde{p}_s = (1 + \epsilon \xi_s) p_s \quad (63)$$

where  $\xi_s$  is a standardized normal random variable, then, according to [Theorem 2](#), the deterministic robust counterpart problem can be obtained by introducing the following constraint to the original scheduling model

$$\text{Profit} \leq \sum_{s \in S_p} p_s \cdot \text{STF}(s) - \sum_{s \in S_r} p_s \cdot \text{STI}(s) - \epsilon \lambda \sqrt{\sum_{s \in S_p} p_s^2 \text{STF}(s)^2 + \sum_{s \in S_r} p_s^2 \text{STI}(s)^2} + \delta \quad (64)$$

where  $\lambda = F^{-1}(1 - \kappa)$  and  $F_n^{-1}$  is the inverse distribution function of a random variable with standardized normal distribution.

## 4. Computational studies

In this section, the robust optimization formulation is applied to four example problems. All the examples are implemented with GAMS 2.50 ([Brooke, Kendrick, Meeraus, & Raman, 2003](#)) on a 3.20 GHz Linux workstation. The MILP problems are solved using CPLEX 8.1 while the MINLP problems are solved using DICOPT ([Viswanathan & Grossmann, 1990](#)).

### 4.1. Example 1: Uncertainty with a poisson distribution in the processing times

Consider the following example process that was first presented by [Kondili, Pantelides, and Sargent \(1993\)](#) and was also used as the motivating example in the Part 1 paper ([Lin et al., 2004](#)) on bounded uncertainty. Two products are produced from three feeds according to the State-Task Network shown in [Fig. 1](#). The STN utilizes three different types of tasks which can be performed in four different units. The corresponding data for the example including suitabilities, capacities, processing times, and storage limitations are given in [Table 1](#). The objective is to maximize the profit from sales of products manufactured in a time horizon of 12 h.

Assume that the uncertainty in the processing times has a poisson distribution with a parameter value of 5 and let us consider an uncertainty level ( $\epsilon$ ) of 5%, an infeasibility tolerance level ( $\delta$ ) of 20%, and a reliability level ( $\kappa$ ) of 24% (corresponding to a  $\lambda$  value of 6). By solving the  $\text{RC}[\epsilon, \delta, \kappa]$  problem, a “robust” schedule is obtained, as shown in [Fig. 3](#), which takes into account uncertainty in the processing times. The nominal schedule can be seen in [Fig. 2](#).

Compared to the nominal solution which is obtained at the nominal values of the processing times, the robust solution exhibits very different scheduling strategies. For example, even the sequences of tasks in the two reactors in [Fig. 3](#) deviates significantly from

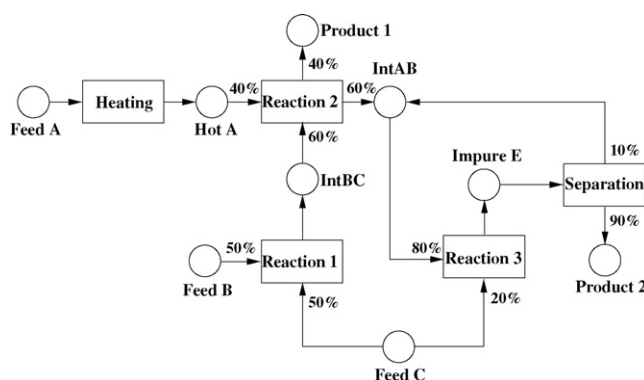


Fig. 1. State-task network for example 1.

Table 1  
Data for example 1

Units	Capacity	Suitability	Processing time
Heater	100	Heating	1.0
Reactor 1	50	Reaction1, 2, 3	2.0, 2.0, 1.0
Reactor 2	80	Reaction1, 2, 3	2.0, 2.0, 1.0
Separator	200	Separation	2.0
States	Storage capacity	Initial amount	Price
Feed A	Unlimited	Unlimited	0.0
Feed B	Unlimited	Unlimited	0.0
Feed C	Unlimited	Unlimited	0.0
Hot A	100	0.0	0.0
IntAB	200	0.0	0.0
IntBC	150	0.0	0.0
ImpureE	200	0.0	0.0
Product 1	Unlimited	0.0	10.0
Product 2	Unlimited	0.0	10.0

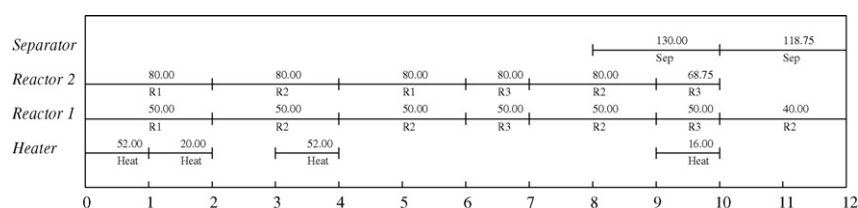


Fig. 2. Optimal solution with nominal processing times (profit = 3638.75).

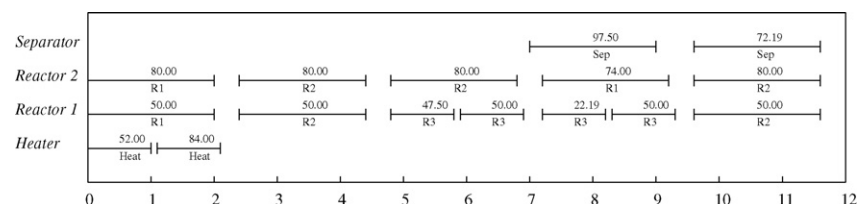


Fig. 3. Robust solution with uncertain processing times (profit = 2887.19).

those in the nominal solution in Fig. 2. The robust solution ensures that the robust schedule obtained is feasible with the specified uncertainty level, infeasibility tolerance, and reliability level. However, the resulting profit is reduced, from 3638.75 to 2887.19, which reflects the effect of uncertainty on overall production. A comparison of the model and solution statistics for the nominal and robust solutions can be found in Table 2.

Fig. 4 summarizes the results of the RC problem with several different combinations of levels of uncertainty and infeasibility at increasing values of the reliability level. It is shown that at a given reliability level, the maximal profit that can be achieved decreases as the uncertainty level increases, which indicates more conservative scheduling decisions because of the existence of uncertainty. Also, at a given reliability level, the maximal profit increases as the infeasibility tolerance level increases, which means more aggressive scheduling arrangements can be incorporated if violations of related timing constraints can be tolerated to a larger extent. In addition, at a given uncertainty level and infeasibility tolerance, the profit increases as the reliability level increases, meaning that the probability of violation of the uncertain constraint allows for more aggressive scheduling. These results are consistent with intuition and other approaches; however, with the robust optimization approach, the effects of uncertainty and the trade-offs between conflicting objectives are quantified rigorously and efficiently.

Table 2  
Model and solution statistics for example 1

	Nominal solution	Robust solution
Profit	3638.75	2887.19
CPU time (s)	0.46	11.33
Binary variables	96	96
Continuous variables	442	442
Constraints	553	777



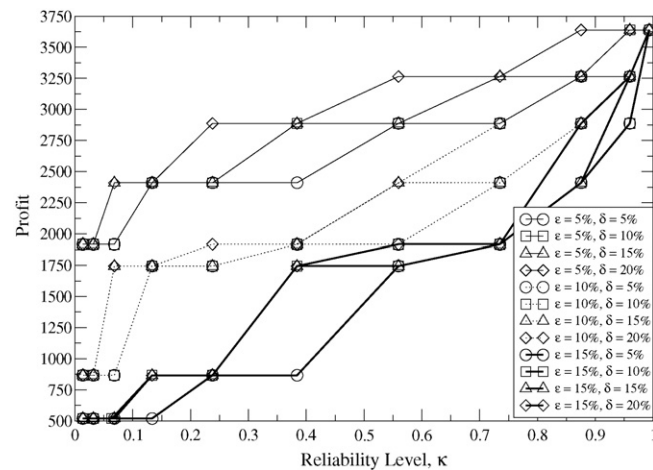


Fig. 4. Profit vs. reliability level at different uncertainty and infeasibility levels in example 1.

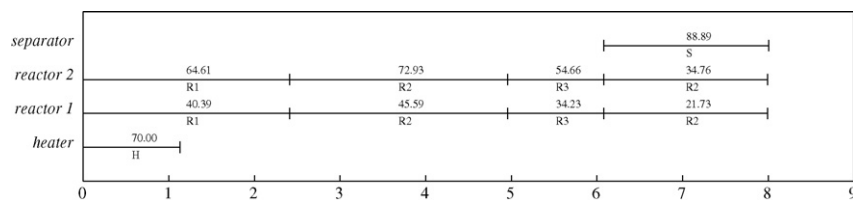


Fig. 5. Optimal solution with nominal product demands (makespan = 8.007).

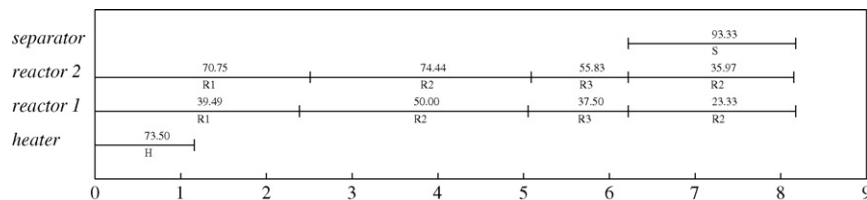


Fig. 6. Robust solution with uncertain product demands (makespan = 8.174).

#### 4.2. Example 2: Uncertainty with a uniform distribution in the product demands

In this example, we consider uncertainty with a uniform distribution in the product demands for the same process given in example 1. However, the objective function is the minimization of the makespan for a given demand of 70 for Product 1 and 80 for Product 2. The uncertainty level ( $\epsilon$ ) is 10%, the infeasibility tolerance ( $\delta$ ) is 5%, and the reliability level ( $\kappa$ ) is 0%. The nominal schedule is shown in Fig. 5 with a makespan of 8.007. The robust schedule is obtained by solving the robust counterpart problem, as shown in Fig. 6, and the corresponding makespan is 8.174. By executing this schedule, the makespan is guaranteed to be at most 8.174 with a probability of 100% in the presence of the 10% uncertainty in the demands of the products. A comparison of the model and solution statistics for the nominal and robust solutions can be found in Table 3.

Fig. 7 summarizes the results of the RC problem with several different combinations of levels of uncertainty and infeasibility at increasing values of the reliability level. It is shown that at a given reliability level, the minimum makespan increases as the uncertainty level increases, which indicates more conservative scheduling decisions that take more time because of the existence of uncertainty in the demands. Also, at a fixed reliability level, the minimum makespan decreases as the infeasibility tolerance level increases, which means more aggressive scheduling arrangements can be incorporated if violations of related demand constraints

Table 3  
Model and solution statistics for example 2

	Nominal solution	Robust solution
Makespan	8.007	8.174
CPU time (s)	0.02	0.02
Binary variables	60	60
Continuous variables	280	280
Constraints	375	409

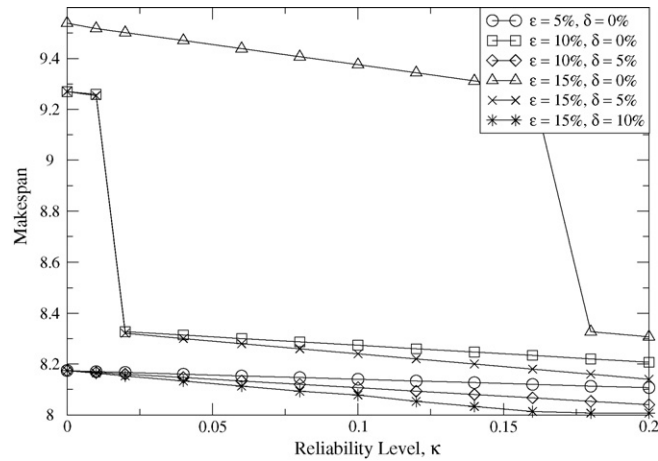


Fig. 7. Makespan vs. reliability level at different uncertainty and infeasibility levels for example 2.

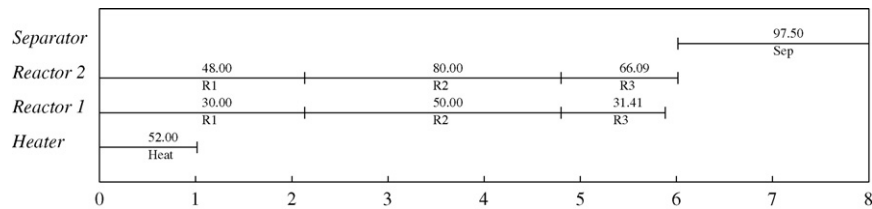


Fig. 8. Optimal solution with nominal market prices (profit = 1088.75).

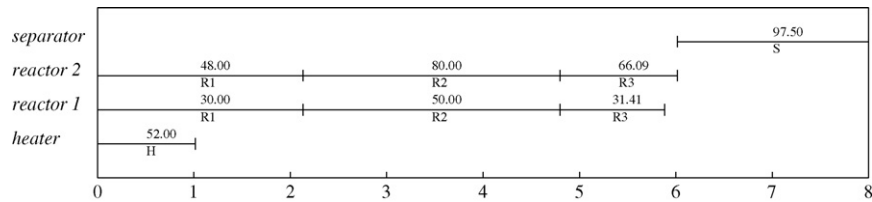


Fig. 9. Robust solution with uncertain market prices (profit = 966.97).

can be tolerated to a larger extent. On the other hand, at a given uncertainty level and infeasibility tolerance, the makespan decreases as the reliability level increases, meaning that the probability of violation of the uncertain constraint allows for more aggressive scheduling. By applying the robust optimization methodology, the effect of the uncertainty on the schedule is clearly quantified.

#### 4.3. Example 3: Uncertainty with a normal distribution in the market prices

In this example, we consider uncertainty with a normal distribution in the market prices for the same process given in examples 1 and 2. However, the objective function is the maximization of profit in a time horizon of 8 h. The uncertainty level ( $\epsilon$ ) is 5%, the infeasibility tolerance ( $\delta$ ) is 5%, and the reliability level ( $\kappa$ ) is 5%. The nominal schedule is shown in Fig. 8 with a profit of 1088.75. The robust schedule is obtained by solving the robust counterpart problem, as shown in Fig. 9, and the corresponding profit is 966.97. By executing this schedule, the profit is guaranteed to be least 966.97 with a probability of 95% in the presence of the 5% uncertainty in the prices of the products and raw materials. A comparison of the model and solution statistics for the nominal and robust solutions can be found in Table 4.

Table 4  
Model and solution statistics for example 3

	Nominal solution	Robust solution
Profit	1088.75	966.97
CPU time (s)	0.02	0.05
Binary variables	60	60
Continuous variables	280	280
Constraints	334	334

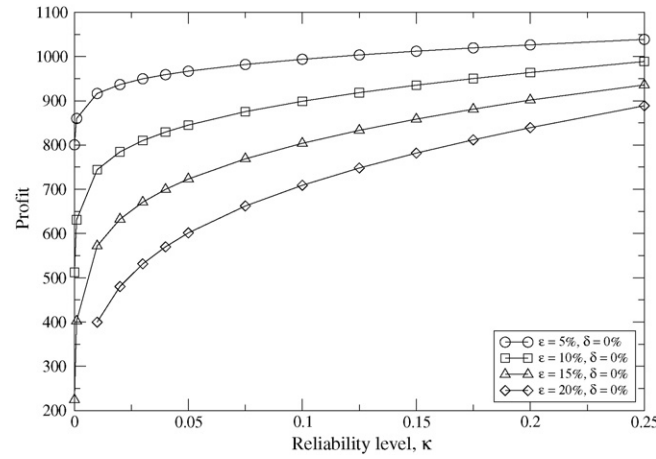


Fig. 10. Profit vs. reliability level at different uncertainty and infeasibility levels for example 3.

Fig. 10 summarizes the results of the RC problem at several different levels of uncertainty and an infeasibility tolerance of 0% at increasing values of the reliability level. It is shown that at a given reliability level, the maximal profit that can be achieved decreases as the uncertainty level increases, which indicates more conservative scheduling decisions because of the existence of uncertainty. Also, at a given uncertainty level and infeasibility tolerance, the profit increases as the reliability level increases, meaning that as the probability of violation of the uncertain constraint, or  $\kappa$ , increases, then  $\lambda$  decreases and according to Eq. (64), the profit takes on a larger value.

#### 4.4. A large case study

##### 4.4.1. Problem description

This example considers uncertainty in the processing times of tasks for an industrial case study originally presented by Lin et al. (2002). Actual plant data is used to determine the form and level of uncertainty in the processing times. The industrial plant is a multiproduct chemical plant that manufactures tens of different products according to a basic three-stage recipe and its variations using 10 pieces of equipment. We consider the first subhorizon in the original case study from Lin et al. (2002) which consists of five days and involves eight different products. The objective function is the maximization of overall production defined as the weighted sum of materials accumulated at the end of the subhorizon minus a penalty term for not meeting demands at the intermediate due dates. For each of the eight products, one of the processing recipes shown in Fig. 11 is applied.

The plant is made up of three types of units, each corresponding to one the three basic processing operations. Four type 1 units (units 1–4) are used for operation 1, three type 2 units (units 5–7) are used for operation 2, and three type 3 units (units 8–10) are used for operation 3. Type 1 units and type 3 units are utilized in batch mode, while type 2 units are operated in a continuous mode. The nominal processing time or processing rate of each task in its suitable units is shown in Table 5.

To determine the form of the uncertainties in the processing times/rates, actual plant data was analyzed. Two different types of uncertainty were chosen based on the data, namely bounded uncertainty and uncertainty with a normal distribution. For bounded uncertainty, the range for the uncertain parameters are given and for normal uncertainty, the mean and standard deviation for the uncertain parameters are given. There were a total of 23 uncertain parameters identified. Eight in units 1–4, five in units 5–7, and 10 in units 8–10. The summary of the typical nominal values, ranges, means and standard deviations for each uncertain parameter is presented in Table 6.

Approach 2 for uncertainty in processing times/rates in Section 3.2 is applied for this case study. In addition to the basic sequencing constraints, the processing time also appears in two additional constraints related to the timing of operation 1 tasks

$$T^s(i, j, n+1) \leq T^f(i, j, n) + H(2 - wv(i, j, n) - wv(i, j, n+1)) \quad \forall i \in I_r, j \in J_i, n \in N, n \neq N, \quad (65)$$

$$T^s(i, j, n+1) \leq T^f(i', j, n) + tcl_{i'} + H(2 - wv(i, j, n) - wv(i', j, n+1)) \quad \forall j \in J_r, i, i' \in I_j, i \neq i', n \in N, n \neq N \quad (66)$$

where  $I_r$  is the set of operation 1 tasks and  $J_r$  is the set of type 1 units suitable for operation 1 tasks. Upon substitution of the  $T^f(i, j, n)$  variables, the following additional constraints are introduced for the parameters with bounded uncertainty to obtain the robust counterpart problem

$$T^s(i, j, n+1) - T^s(i, j, n) \leq \alpha_{ij}^L \cdot wv(i, j, n) + \beta_{ij}^L \cdot B(i, j, n) + H(2 - wv(i, j, n) - wv(i, j, n+1)) + \delta_2, \quad (67)$$

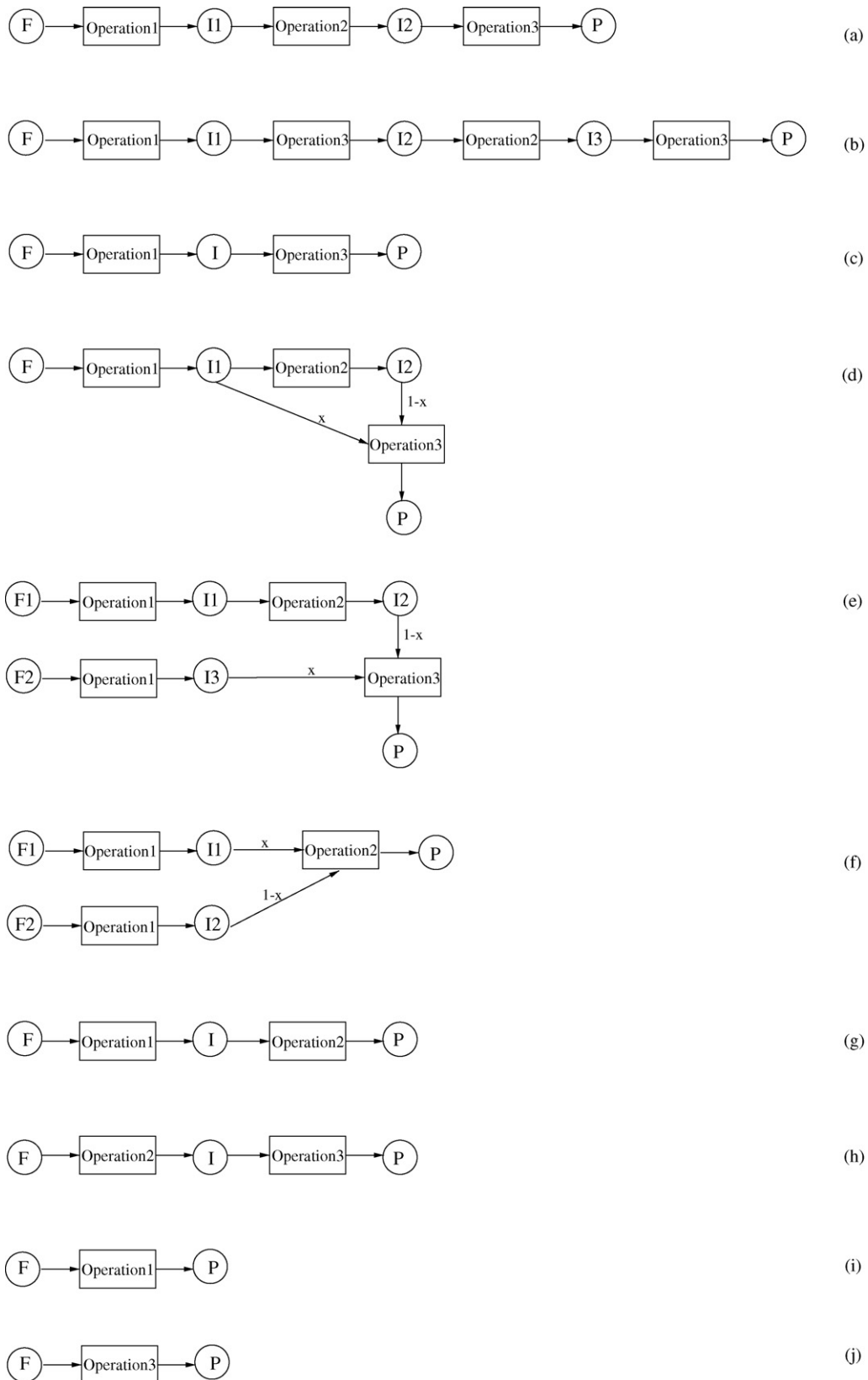


Fig. 11. State-task network of production recipes in the industrial case study.

Table 5  
Nominal processing times and rates in the industrial case study

Task	Unit 1	Unit 2	Unit 3	Unit 4	Unit 5	Unit 6	Unit 7	Unit 8	Unit 9	Unit 10
1	0	0	0	9.5	–	–	–	–	–	–
2	–	–	–	–	0	0	0.95	–	–	–
3	–	–	–	–	–	–	–	12	12.8	12.5
4	0	10	10	10	–	–	–	–	–	–
5	–	–	–	–	0.575	0.575	0.725	–	–	–
6	–	–	–	–	–	–	–	12	12.8	12.5
7	6.09	6.09	6.09	11.1	–	–	–	–	–	–
8	–	–	–	–	0.6	0.6	0.8	–	–	–
9	–	–	–	–	–	–	–	12.5	13.8	12.9
10	6.09	6.09	6.09	11.1	–	–	–	–	–	–
11	–	–	–	–	0.6	0.6	0.8	–	–	–
12	–	–	–	–	–	–	–	12.5	13.8	12.9
13	6.09	6.09	6.09	11.1	–	–	–	–	–	–
14	–	–	–	–	0.6	0.6	0.8	–	–	–
15	–	–	–	–	–	–	–	12.5	13.8	12.9
16	–	–	–	–	0.6	0.6	0.8	–	–	–
17	–	–	–	–	–	–	–	12.5	13.8	12.9
18	0	8.5	8.5	0	–	–	–	–	–	–
19	–	–	–	–	–	–	–	0	15	16
20	0	0	8.38	9.5	–	–	–	–	–	–

$$T^s(i, j, n+1) - T^s(i', j, n) \leq \alpha_{i'j}^L \cdot wv(i', j, n) + \beta_{i'j}^L \cdot B(i', j, n) + tcl_{i'j} + H(2 - wv(i, j, n) - wv(i', j, n+1)) + \delta 2 \quad (68)$$

where  $\alpha_{ij}^L = (1 - \epsilon)\alpha_{ij}$ ,  $\beta_{ij}^L = (1 - \epsilon)\beta_{ij}$ , and  $\delta 2$  is defined as a variable and correlates as follows with parameter  $\delta$  that participates in the additional constraints corresponding to the basic sequencing constraints

$$\begin{aligned} \delta + \delta 2 &= \alpha^U - \alpha^L = 2 \cdot \epsilon \cdot \alpha, \\ \text{or } \delta + \delta 2 &= \beta^U - \beta^L = 2 \cdot \epsilon \cdot \beta. \end{aligned} \quad (69)$$

Similarly, the following additional constraints are introduced for the parameters with normal uncertainty

$$\begin{aligned} T^s(i, j, n+1) - T^s(i, j, n) &\leq [1 - \epsilon(\lambda_{ij}^\alpha \cdot \sqrt{\sigma_{ij}^\alpha} - \mu_{ij}^\alpha)] \cdot \alpha_{ij} \cdot wv(i, j, n) + [1 - \epsilon(\lambda_{ij}^\beta \cdot \sqrt{\sigma_{ij}^\beta} - \mu_{ij}^\beta)] \cdot \beta_{ij} \cdot B(i, j, n) \\ &+ H(2 - wv(i, j, n) - wv(i, j, n+1)) + \delta 2, \end{aligned} \quad (70)$$

$$\begin{aligned} T^s(i, j, n+1) - T^s(i', j, n) &\leq [1 - \epsilon(\lambda_{i'j}^\alpha \cdot \sqrt{\sigma_{i'j}^\alpha} - \mu_{i'j}^\alpha)] \alpha_{i'j} \cdot wv(i', j, n) + [1 - \epsilon(\lambda_{i'j}^\beta \cdot \sqrt{\sigma_{i'j}^\beta} - \mu_{i'j}^\beta)] \beta_{i'j} \cdot B(i', j, n) \\ &+ tcl_{i'j} + H(2 - wv(i, j, n) - wv(i', j, n+1)) + \delta 2 \end{aligned} \quad (71)$$

where  $\delta 2$  is defined as

$$\delta + \delta 2 = 2 \cdot \epsilon(\lambda \cdot \sqrt{\sigma} - \mu). \quad (72)$$

The objective function for this problem is the maximization of production in terms of the relative value of all states minus a penalty term for not meeting demands at the intermediate due dates

$$\gamma \sum_s \text{vald}_s \text{valp}_s \text{valm}_s \text{STF}(s) - \sum_s \sum_n \text{pri}_{sn} \text{SL}(s, n) \quad \forall s \in S, n \in N \quad (73)$$

Table 6  
Bounded and normal uncertainty in processing times/rates for the case study

Task	Unit	Uncertainty	Nominal value	Range	Mean	Standard deviation
1	4	Normal	9.5	–	9.912	0.523
7,10,13	1–3	Normal	6.09	–	6.153	0.152
7,10,13	4	Bounded	11.1	10.1–11.3	–	–
20	3	Bounded	8.38	8.00–10.42	–	–
2	7	Normal	0.95	–	0.9611	0.112
8,11,14,16	5–6	Bounded	0.60	0.344–0.853	–	–
3,6	9	Bounded	12.8	10.5–19.3	–	–
9,12,15,17	9	Bounded	13.8	12.0–16.3	–	–
9,12,15,17	10	Normal	12.9	–	12.100	0.760

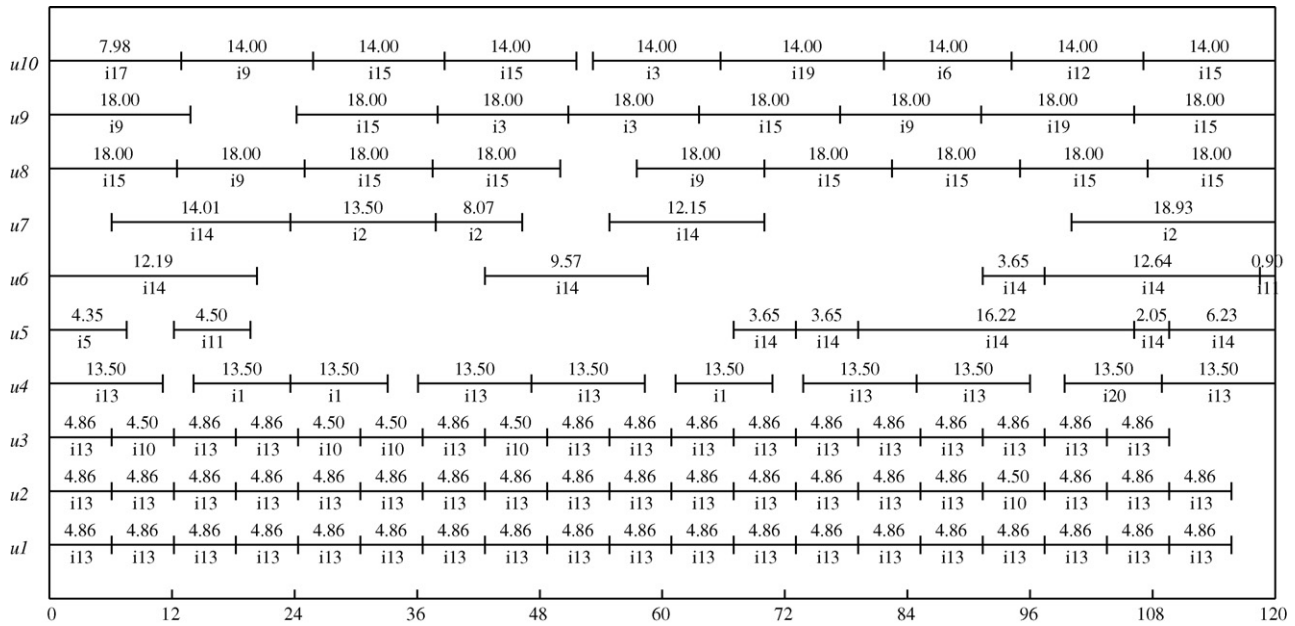


Fig. 12. Nominal schedule in the case study (Obj = 121.37, U1–U10:units, I1–I19:tasks).

where  $\text{vald}_s$  is the relative value of the corresponding product indicating its importance to fulfill future demands,  $\text{valp}_s$  is the relative value of the corresponding product indicating its priority,  $\text{valm}_s$  is the relative value of state ( $s$ ) in the sequence of materials for the corresponding product,  $\text{STF}(s)$  is the amount of state ( $s$ ) at the end of the horizon,  $\text{pri}_{sn}$  is the priority of demand for state ( $s$ ) at event point ( $n$ ),  $\text{SL}(s, n)$  is a slack variable for the amount of state ( $s$ ) not meeting the demand at event point ( $n$ ), and  $\gamma$  is a constant coefficient used to balance the relative value of the two terms in the objective function.

The resulting mathematical model for the industrial case study problem uses the sequencing constraints (52)–(54) described under approach 2 in Section 3.2 for uncertainty in the processing times of tasks. This problem also requires the additional sequencing constraints (67)–(71) for tight timing of the operation 1 tasks. Using these constraints to describe the problem uncertainty results in a mixed-integer linear programming problem. This is due to the fact that these uncertain constraints each contain only one uncertain parameter, resulting in linear deterministic forms for both the bounded uncertain constraints and the normal uncertain constraints.

#### 4.4.2. Computational results and discussion

The nominal solution to this problem using the continuous-time formulation is shown in Fig. 12 and the objective function value is 121.37. At a (relative) infeasibility tolerance level ( $\delta$ ) of 10% for the bounded uncertain parameters and 20% for the normal uncertain parameters and an uncertainty level ( $\epsilon$ ) of 5% and a reliability level ( $\kappa$ ) of 5% for the normal uncertain parameters, meaning the constraints are violated only 5% of the time, the solution to the robust counterpart problem with all twenty-three uncertain parameters is shown in Fig. 13 and the objective function value is 105.76. It can be seen that the processing time of each uncertain task is extended to ensure that the schedule is feasible within the specified uncertainty level, infeasibility tolerance, and reliability level; however, the objective function value has decreased. A closer examination of the terms involved in the objective function indicates that the objective function value for the robust solution decreased because the relative values of the violations of the intermediate due dates increased while the overall production decreased. A comparison of the model and solution statistics for the nominal and robust solutions of the industrial case study can be found in Table 7. Note that the reported cpu times indicate the time it took to obtain the best solution possible within a time limit of two hours.

Table 7  
Model and solution statistics for case study

	Nominal solution	Robust solution
Objective	121.37	105.76
Binary variables	930	930
Continuous variables	6005	6161
Constraints	18907	22931
CPU time (s)	3880	5910
Nodes	15230	35640

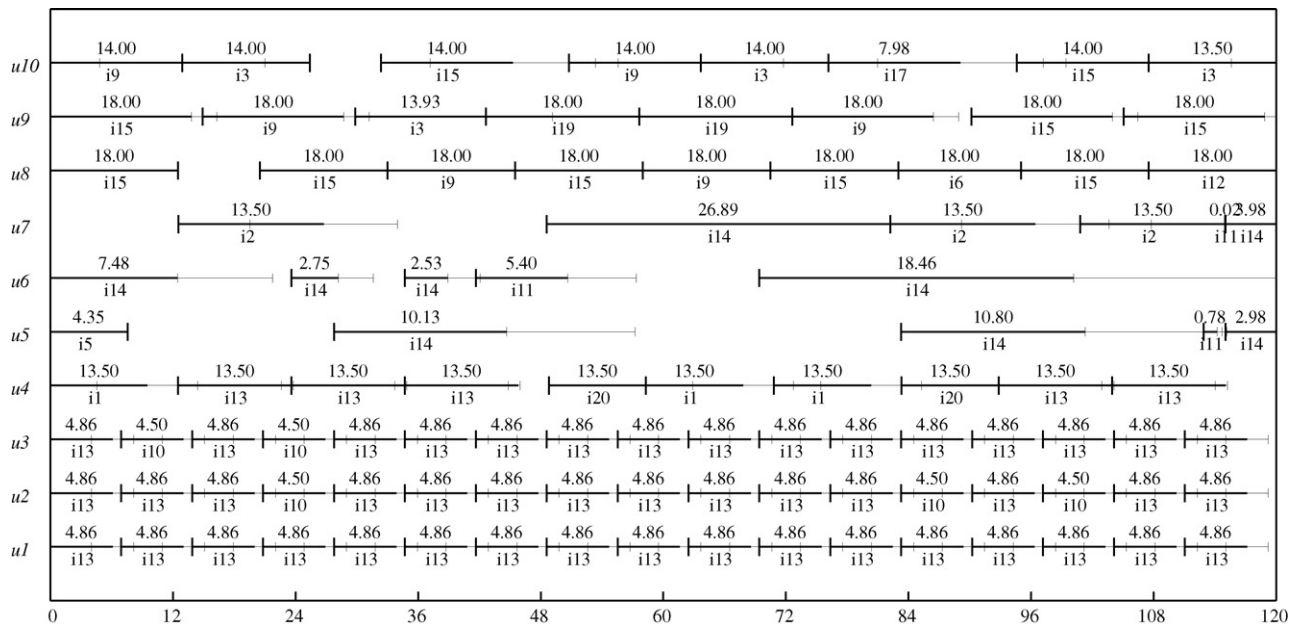


Fig. 13. Robust schedule in the case study (Obj = 105.76,  $\delta = 10\%$ , 20%, U1–U10:units, I1–I19:tasks, *Note: The endpoint of each thick horizontal line indicates the finishing time of the task based on the nominal processing time while the thin vertical lines represent the range of the finishing time due to the uncertainty present in the processing time.*)

## 5. Discussion

For the short-term scheduling problem considered in Section 4.1, the level of conservatism experienced in the resulting production schedule varies with the form the probability distribution function used to describe the uncertain parameters. Comparisons of each type of uncertainty on all three uncertain constraints over a variety of problem parameters for this example problem indicate the following: (i) the general discrete distribution is always the least conservative, giving the best objective function values, (ii) the binomial and poisson distributions always give the most conservative results, yielding the worst objective function values, (iii) of the other types of distributions (i.e., bounded, uniform, normal), in general, the normal distribution is usually more conservative than the other two and the uniform distribution is usually less conservative than the other two. However, the relative results for these three distributions are highly dependent on the value of the reliability level,  $\kappa$ . As  $\kappa$  increases, indicating a higher reliability, the normal distribution becomes less conservative compared with the other two. In fact, for large  $\kappa$ , the normal distribution is less conservative than the bounded distribution.

In addition, although the robust optimization formulation can be used to model uncertainty in a wide variety of MILP problems, there are some limitations of the proposed approach. For instance, some of the probability distribution functions are only applicable to constraints that contain a single uncertain parameter (i.e., uniform, binomial, poisson). This is due to limits in probability theory and not the proposed formulation. Also, the robust optimization formulation cannot address dependent uncertain parameters which are related through general nonlinear expressions, but it is applicable to linearly dependent uncertain parameters. Finally, the current formulation is only capable of handling uncertainty present in linear constraints. Future work will aim at addressing these points.

## 6. Conclusions

In this work, we propose a new approach to address the scheduling under uncertainty problem based on a robust optimization methodology, which when applied to mixed-integer (MILP) problems produces “robust” solutions that are, in a sense, immune against uncertainties in both the coefficients and right-hand-side parameters of the inequality constraints. The approach can be applied to address the problem of production scheduling with uncertain processing times, market demands, and/or prices of products and raw materials. Our computational results show that this approach provides an effective way to address scheduling problems under uncertainty, producing reliable schedules and generating helpful insights on the tradeoffs between conflicting objectives. Furthermore, due to its efficient transformation, the approach is capable of solving real-world problems with a large number of uncertain parameters.

## Acknowledgments

The authors gratefully acknowledge support from the National Science Foundation, Arkema Inc., and the BASF Corporation.



## Appendix A. Continuous-time process scheduling formulation

This formulation was proposed by Floudas and co-workers (Ierapetritou & Floudas, 1998a, 1998b; Ierapetritou et al., 1999; Janak et al., 2004, 2006a, 2006b; Lin & Floudas, 2001; Lin et al., 2002, 2003).

### Nomenclature:

$\alpha_{ij}$	constant term of processing time of task ( $i$ ) in unit ( $j$ );
$\beta_{ij}$	variable term of processing time of task ( $i$ ) in unit ( $j$ ) expressing the time required by the unit to process one unit of material performing task ( $i$ );
$\rho_{si}^p, \rho_{si}^c$	proportion of state ( $s$ ) produced, consumed by task ( $i$ ), respectively;
$B(i, j, n)$	continuous, amount of material undertaking task ( $i$ ) in unit ( $j$ ) at event point ( $n$ );
$dem_s$	market demand for state ( $s$ ) at the end of the time horizon;
$H$	time horizon;
$i \in I$	tasks;
$I_j$	tasks which can be performed in unit ( $j$ );
$I_s$	tasks which produce or consume state ( $s$ );
$j \in J$	units;
$J_i$	units which are suitable for performing task ( $i$ );
$n \in N$	event points representing the beginning of a task;
$p_s$	price of state ( $s$ );
Profit	continuous, overall profit;
$s \in S$	states;
$S_p$	states corresponding to final products;
$S_r$	states corresponding to raw materials;
$ST(s, n)$	continuous, amount of state ( $s$ ) at event point ( $n$ );
$ST_s^{\max}$	available maximum storage capacity for state ( $s$ );
$STF(s)$	continuous, final amount of state ( $s$ ) at the end of the time horizon;
$STI(s)$	continuous, initial amount of state ( $s$ ) at the beginning of the time horizon;
$T^s(i, j, n)$	continuous, time at which task ( $i$ ) starts in unit ( $j$ ) at event point ( $n$ );
$T^f(i, j, n)$	continuous, time at which task ( $i$ ) finishes in unit ( $j$ ) while it starts at event point ( $n$ );
$V_{ij}^{\min}$	minimum amount of material processed by task ( $i$ ) required to start operating unit ( $j$ );
$V_{ij}^{\max}$	maximum capacity of unit ( $j$ ) when processing task ( $i$ );
$wv(i, n)$	binary, whether or not task ( $i$ ) starts at event point ( $n$ );

### Constraints:

#### Allocation constraints

$$\sum_{i \in I_j} wv(i, n) \leq 1, \quad \forall j \in J, n \in N$$

#### Capacity constraints

$$V_{ij}^{\min} wv(i, n) \leq B(i, j, n) \leq V_{ij}^{\max} wv(i, n), \quad \forall i \in I, j \in J_i, n \in N$$

#### Storage constraints

$$ST(s, n) \leq ST_s^{\max}, \quad \forall s \in S, n \in N$$

#### Material balances

$$ST(s, n) = STI(s) + \sum_{i \in I_s} \rho_{si}^c \sum_{j \in J_i} B(i, j, n), \quad \forall s \in S, n \in N, n = 1$$

$$ST(s, n) = ST(s, n-1) + \sum_{i \in I_s} \rho_{si}^p \sum_{j \in J_i} B(i, j, n-1) + \sum_{i \in I_s} \rho_{si}^c \sum_{j \in J_i} B(i, j, n), \quad \forall s \in S, n \in N$$

$$STF(s) = ST(s, n) + \sum_{i \in I_s} \rho_{si}^p \sum_{j \in J_i} B(i, j, n), \quad \forall s \in S, n \in N, n = N$$

#### Demand constraints

$$STF(s) \geq dem_s, \quad \forall s \in S$$

#### Duration constraints

$$T^f(i, j, n) = T^s(i, j, n) + \alpha_{ij} wv(i, n) + \beta_{ij} B(i, j, n), \quad \forall i \in I, j \in J_i, n \in N$$

Sequence constraints: same task in the same unit

$$T^s(i, j, n + 1) \geq T^f(i, j, n), \quad \forall i \in I, j \in J_i, n \in N, n \neq N$$

Sequence constraints: different tasks in the same unit

$$T^s(i, j, n + 1) \geq T^f(i', j, n) - H[1 - wv(i', n)], \quad \forall j \in J, i, i' \in I_j, i \neq i', n \in N, n \neq N$$

Sequence constraints: different tasks in different units

$$T^s(i, j, n + 1) \geq T^f(i', j', n) - H[1 - wv(i', n)], \quad \forall j, j' \in J, i \in I_j, i' \in I_{j'}, i \neq i', n \in N, n \neq N$$

time horizon constraints

$$T^f(i, j, n) \leq H, \quad \forall i \in I, j \in J_i, n \in N$$

$$T^s(i, j, n) \leq H, \quad \forall i \in I, j \in J_i, n \in N$$

Objective function:

$$\text{Max Profit} = \sum_{s \in S_p} p_s \cdot \text{STF}(s) - \sum_{s \in S_r} p_s \cdot \text{STI}(s)$$

## References

- Balasubramanian, J., & Grossmann, I. E. (2002). A novel branch and bound algorithm for scheduling flowshop plants with uncertain processing times. *Computers and Chemical Engineering*, 26, 41–57.
- Balasubramanian, J., & Grossmann, I. E. (2003). Scheduling optimization under uncertainty—An alternative approach. *Computers and Chemical Engineering*, 27, 469–490.
- Balasubramanian, J., & Grossmann, I. E. (2004). Approximation to multistage stochastic optimization in multiperiod batch plant scheduling under demand uncertainty. *Industrial and Engineering Chemistry Research*, 43, 3695–3713.
- Bassett, M. H., Pekny, J. F., & Reklaitis, G. V. (1997). Using detailed scheduling to obtain realistic operating policies for a batch processing facility. *Industrial and Engineering Chemistry Research*, 36, 1717–1726.
- Ben-Tal, A., Goryashko, A., Guslitzer, E., & Nemirovski, A. (2004). Adjustable robust solutions of uncertain linear programs. *Mathematical Programming A*, 99, 351–376.
- Ben-Tal, A., & Nemirovski, A. (1998). Robust convex optimization. *Mathematical Operations Research*, 23, 769–805.
- Ben-Tal, A., & Nemirovski, A. (1999). Robust solutions of uncertain linear programs. *Operations Research Letters*, 25, 1–13.
- Ben-Tal, A., & Nemirovski, A. (2000). Robust solutions of linear programming problems contaminated with uncertain data. *Mathematical Programming A*, 88, 411–424.
- Bertsimas, D., & Sim, M. (2003). Robust discrete optimization and network flows. *Mathematical Programming B*, 98, 49–71.
- Bertsimas, D., & Sim, M. (2004). The price of robustness. *Operations Research*, 52, 35–53.
- Bertsimas, D., & Sim, M. Robust conic optimization, *Mathematical Programming*, submitted for publication.
- Bonfill, A., Bagajewicz, M., Espuña, A., & Puigjaner, L. (2004). Risk management in the scheduling of batch plants under uncertain market demand. *Industrial and Engineering Chemistry Research*, 43, 741–750.
- Bonfill, A., Espuña, A., & Puigjaner, L. (2005). Addressing robustness in scheduling batch processes with uncertain operation times. *Industrial and Engineering Chemistry Research*, 44, 1524–1534.
- Brooke, A., Kendrick, D., Meeraus, A., & Raman, R. (2003). *GAMS: A user's guide*. South San Francisco, CA.
- Charnes, A., & Cooper, W. W. (1959). Chance-constrained programming. *Management Science*, 6, 73–79.
- Charnes, A., & Cooper, W. W. (1962). Chance constraints and normal deviates. *Journal of the American Statistical Association*, 57, 134–148.
- Charnes, A., & Cooper, W. W. (1963). Deterministic equivalents for optimizing and satisficing under chance constraints. *Operations Research*, 11, 18–39.
- Cott, B. J., & Macchietto, S. (1989). Minimizing the effects of batch process variability using online schedule modification. *Computers and Chemical Engineering*, 13, 105–113.
- El Ghaoui, L. H. (1997). Robust solutions to least-square problems with uncertain data. *SIAM Journal on Matrix Analysis and Applications*, 18, 1035–1064.
- El Ghaoui, L. H., Oustry, F., & Lebret, H. (1998). Robust solutions to uncertain semidefinite programs. *SIAM Journal on Optimization*, 9, 33–52.
- Floudas, C. A. (2005). Research challenges, opportunities and synergism in systems engineering and computational biology. *Annals of Operations Research*, 51, 1872–1884.
- Floudas, C. A., & Lin, X. (2004). Continuous-time versus discrete-time approaches for scheduling of chemical processes: A review. *Computers and Chemical Engineering*, 28, 2109–2129.
- Floudas, C. A., & Lin, X. (2005). Mixed integer linear programming in process scheduling: Modeling, algorithms, and applications. *Annals of Operations Research*, 139, 131–162.
- Honkomp, S. J., Mockus, L., & Reklaitis, G. V. (1999). A framework for schedule evaluation with processing uncertainty. *Computers and Chemical Engineering*, 23, 595–609.
- Ierapetritou, M. G., & Floudas, C. A. (1998a). Effective continuous-time formulation for short-term scheduling: 1. Multipurpose batch processes. *Industrial and Engineering Chemistry Research*, 37, 4341–4359.
- Ierapetritou, M. G., & Floudas, C. A. (1998b). Effective continuous-time formulation for short-term scheduling: 2. Continuous and semi-continuous processes. *Industrial and Engineering Chemistry Research*, 37, 4360–4374.
- Ierapetritou, M. G., Hené, T. S., & Floudas, C. A. (1999). Effective continuous-time formulation for short-term scheduling: 3. Multiple intermediate due dates. *Industrial and Engineering Chemistry Research*, 38, 3446–3461.

- Janak, S. L., Floudas, C. A., Kallrath, J., & Vormbrock, N. (2006a). Production scheduling of a large-scale industrial batch plant. I. Short-term and medium-term scheduling. *Industrial and Engineering Chemistry Research*, in press.
- Janak, S. L., Floudas, C. A., Kallrath, J., & Vormbrock, N. (2006b). Production scheduling of a large-scale industrial batch plant. II. Reactive scheduling. *Industrial and Engineering Chemistry Research*, in press.
- Janak, S. L., Lin, X., & Floudas, C. A. (2004). Enhanced continuous-time unit-specific event-based formulation for short-term scheduling of multipurpose batch processes: Resource constraints and mixed storage policies. *Industrial and Engineering Chemistry Research*, 43, 2516–2533.
- Jia, Z., & Ierapetritou, M. G. (2004). Short-term scheduling under uncertainty using MILP sensitivity analysis. *Industrial and Engineering Chemistry Research*, 43, 3782–3791.
- Kanakamedala, K. B., Reklaitis, G. V., & Venkatasubramanian, V. (1994). Reactive schedule modification in multipurpose batch chemical plants. *Industrial and Engineering Chemistry Research*, 33, 77–90.
- Kondili, E., Pantelides, C. C., & Sargent, R. W. H. (1993). A general algorithm for short-term scheduling of batch operations—I. MILP formulation. *Computers and Chemical Engineering*, 17, 211–227.
- Lin, X., Chajakis, E. D., & Floudas, C. A. (2003). Scheduling of tanker lightering via a novel continuous-time optimization framework. *Industrial and Engineering Chemistry Research*, 42, 4441–4451.
- Lin, X., & Floudas, C. A. (2001). Design, synthesis and scheduling of multipurpose batch plants via an effective continuous-time formulation. *Computers and Chemical Engineering*, 25, 665–674.
- Lin, X., Floudas, C. A., Modi, S., & Juhasz, N. M. (2002). Continuous-time optimization approach for medium-range production scheduling of a multiproduct batch plant. *Industrial and Engineering Chemistry Research*, 41, 3884–3906.
- Lin, X., Janak, S. L., & Floudas, C. A. (2004). A New robust optimization approach for scheduling under uncertainty: I. Bounded uncertainty. *Computers and Chemical Engineering*, 28, 1069–1085.
- Orçun, S., Altinel, I. K., & Hortaçsu, Ö. (1996). Scheduling of batch processes with operational uncertainties. *Computers and Chemical Engineering*, 20, S1191–S1196.
- Ostrovsky, G. M., Datskov, I. V., Achenie, L. E. K., & Volin, Y. M. (2004). Process uncertainty: Case of insufficient process data at the operation stage. *AIChE Journal*, 49, 1216–1232.
- Petkov, S. B., & Maranas, C. D. (1997). Multiperiod planning and scheduling of multiproduct batch plants under demand uncertainty. *Industrial and Engineering Chemistry Research*, 36, 4864–4881.
- Rodrigues, M. T. M., Gimeno, L., Passos, C. A. S., & Campos, M. D. (1996). Reactive scheduling approach for multipurpose chemical batch plants. *Computers and Chemical Engineering*, 20, S1215–S1220.
- Sahinidis, N. V. (2004). Optimization under uncertainty: State-of-the-art and opportunities. *Computers and Chemical Engineering*, 128, 971–983.
- Sanmarti, E., Huercio, A., Espuña, A., & Puigjaner, L. (1996). A combined scheduling/reactive scheduling strategy to minimize the effect of process operations uncertainty in batch plants. *Computers and Chemical Engineering*, 20, S1263–S1268.
- Schweiger, C., & Floudas, C. (1998). *MINOPT: A software package for mixed-integer nonlinear optimization, user's guide*. Princeton, NJ: Computer-Aided Systems Laboratory, Dept. of Chemical Engineering, Princeton University.
- Shaik, M., Janak, S. L., & Floudas, C. A. Comparative study of continuous-time models for short-term scheduling of multipurpose batch plants. *Industrial and Engineering Chemistry Research*, in press.
- Vin, J. P., & Ierapetritou, M. G. (2000). A new approach for efficient rescheduling of multiproduct batch plants. *Industrial and Engineering Chemistry Research*, 39, 4228–4238.
- Vin, J. P., & Ierapetritou, M. G. (2001). Robust short-term scheduling of multiproduct batch plants under demand uncertainty. *Industrial and Engineering Chemistry Research*, 40, 4543–4554.
- Viswanathan, J., & Grossmann, I. E. (1990). *DICOPT ++: A program for mixed integer nonlinear optimization, user's guide*. Pittsburgh, Pennsylvania: Engineering Design Research Center, Carnegie Mellon University.
- Wang, J. (2004). A fuzzy robust scheduling approach for product development projects. *European Journal of Operational Research*, 152, 180–194.