An approximation framework for two-stage ambiguous stochastic integer programs under mean-MAD information

Postek, Krzysztof; Romeijnders, Ward; den Hertog, Dick; van der Vlerk, M. H.

Published in:
European Journal of Operational Research

DOI:
10.1016/j.ejor.2018.10.008

European Journal of Operational Research 274 (2019) 432–444

Contents lists available at ScienceDirect

European Journal of Operational Research

journal homepage: www.elsevier.com/locate/ejor

Discrete Optimization

An approximation framework for two-stage ambiguous stochastic integer programs under mean-MAD information

Krzysztof Postek[a,*], Ward Romeijnders[b], Dick den Hertog[c], Maarten H. van der Vlerk[b,1]

[a] Department of Econometrics, Erasmus School of Economics, Erasmus University Rotterdam, The Netherlands
[b] Department of Operations, Faculty of Economics and Business, University of Groningen, The Netherlands
[c] CentER and Department of Econometrics and Operations Research, Tilburg University, The Netherlands

ARTICLE INFO

Article history:
Received 9 March 2017
Accepted 5 October 2018
Available online 11 October 2018

Keywords:
Stochastic programming
Robust
Ambiguous
Integer
Recourse

ABSTRACT

We consider two-stage recourse models in which only limited information is available on the probability distributions of the random parameters in the model. If all decision variables are continuous, then we are able to derive the worst-case and best-case probability distributions under the assumption that only the means and mean absolute deviations of the random parameters are known. Contrary to most existing results in the literature, these probability distributions are the same for every first-stage decision. The ambiguity set that we use in this paper also turns out to be particularly suitable for ambiguous recourse models involving integer decisions variables. For such problems, we develop a general approximation framework and derive error bounds for using these approximatons. We apply this approximation framework to mixed-ambiguous mixed-integer recourse models in which some of the probability distributions of the random parameters are known and others are ambiguous. To illustrate these results we carry out numerical experiments on a surgery block allocation problem.

1. Introduction

Many practical decisions are made while key information is uncertain. Consider, for example, customer demand in production planning, supply of renewable energy in unit commitment problems, the precision of physical devices in engineering design, and the return on investment in finance. All these problems can be modelled as stochastic programming (SP) problems, or recourse models, in which the uncertain information is represented by random parameters; see, e.g., the textbooks Birge and Louveaux (1997), Prékopa (1995) and Shapiro, Dentcheva, and Ruszczyński (2009). Practical applications may add two difficulties to traditional or standard recourse models. First, instead of the probability distributions of the random parameters being known, only limited information may be available so that it is more realistic to assume that their distributions are ambiguous, i.e., only partly known (Knight, 1921). Second, some of the recourse (later-stage) decision variables can be integer. In this paper we consider the even more difficult situation when both of these difficulties are present, that is, we consider ambiguous two-stage recourse models, possibly involving integer decision variables. For such problems, we derive an approximation framework so that we can efficiently obtain good approximating solutions for them, even for large-scale instances. By deriving error bounds for the approximations we guarantee the performance of the approximating solutions.

Ambiguous recourse models were first considered by Scarf (1958) and Žáčková (1966). In the SP literature they are called *minimax* problems (see also Kemperman, 1968; Shapiro & Kleywegt, 2002; Shapiro & Ahmed, 2004), whereas in the robust optimization (RO) literature they are called *distributionally robust optimization* problems (see, e.g., Delage & Ye, 2010 and Wiesemann, Kuhn, & Sim, 2014). The above references all minimize *worst-case* expected costs over all admissible probability distributions, whereas in this paper in addition also the *best-case* expected costs will be minimized. In fact, we will determine worst-case and best-case probability distributions (distinct from each other) that are, contrary to most of these references, the same for all first-stage decisions. This is more intuitive for practitioners and also convenient for stress testing.

The ambiguity set we use contains information on the supports, means, and mean absolute deviations (MADs) of the random parameters, and the probabilities that they are greater than or equal to their mean. These values are easy to estimate using, e.g., the procedures given in Postek, Ben-Tal, Hertog, and Melenberg (2018). Under this information on the random parameters, we can

* Corresponding author.
E-mail addresses: postek@ese.eur.nl (K. Postek), w.romeijnders@rug.nl (W. Romeijnders), d.denhertog@tilburguniversity.edu (D. den Hertog).
[1] 1961–2016

https://doi.org/10.1016/j.ejor.2018.10.008
0377-2217/

use a result of Ben-Tal and Hochman (1972), referred to as BTH72, to prove that the worst-case and best-case probability distributions are discrete with at most three possible realizations per random parameter if all decision variables in the model are continuous. The difference between the worst-case and best-case expected costs gives an easy-to-calculate upper bound on the *value of distributional information* (VDI); see, e.g., Delage, Arroyo, and Ye (2015). The VDI is particularly relevant in a data-driven environment where it can be used to assess the costs of gathering more data.

The ambiguity set also turns out to be useful for ambiguous recourse models involving integer decision variables. Such problems are extremely challenging, since they combine the difficulties of having (i) integer decision variables, and (ii) random parameters with (iii) probability distributions that are partly known. Such problems have been studied only limitedly in the literature; see, e.g., Hanasusanto, Kuhn, and Wiesemann (2016) and Xie and Ahmed (2017) for recent contributions in RO and SP, respectively. For example, the latter consider distributionally robust simple integer recourse models in which only the means and supports of the random parameters are known. Ambiguous recourse models generalize standard integer recourse models that have been studied by, e.g., Laporte and Louveaux (1993), Carœ and Schultz (1999), Ahmed, Tawarmalani, and Sahinidis (2004), Sen and Higle (2005), and Gade, , and Küçükyavuz (2014) (see also the surveys by Schultz, 2003, Klein Haneveld & van der Vlerk, 1999, and Sen, 2005). We refer to Bertsimas and Georghiou (2015), Hanasusanto, Kuhn, and Wiesemann (2015), and Postek and Hertog (2016) for studies in adjustable RO involving integer decision variables.

The reason why even these standard mixed-integer recourse models are so hard to solve is that they are generally non-convex. For this reason, van der Vlerk (2004), Klein Haneveld, Stougie, and van der Vlerk (2006), Romeijnders, van der Vlerk, and Haneveld (2015), Romeijnders, van der Vlerk, and Klein Haneveld (2016b), and Romeijnders, Schultz, van der Vlerk, and Haneveld (2016a) have proposed convex approximations for several classes of mixed-integer recourse models. For these approximations error bounds have been derived that depend on the total variations of the probability density functions of the random parameters in the model. Inspired by these results we derive an approximation framework for *ambiguous* mixed-integer recourse models. This is the main contribution of this paper.

We derive error bounds for using convex approximations for ambiguous mixed-integer recourse models, minimizing both worst-case and best-case expected costs. For the convex approximating models, we can apply the results of BTH72, obtaining the same worst-case and best-case probability distributions as for continuous recourse models. This explains why our ambiguity set is suitable when using these approximations. Interestingly, we are also able to derive error bounds for incorrectly assuming that the worst-case and best-case probability distributions are, respectively, the same for ambiguous mixed-integer recourse models as for continuous recourse models.

We apply the approximation framework to two-stage *mixed-ambiguous* mixed-integer recourse models in which some distributions of the random parameters are known and others are ambiguous. For such models we can use the existing convex approximations of Romeijnders et al. (2016a, 2016b) and Romeijnders, Morton, and van der Vlerk (2017) with corresponding error bounds for standard mixed-integer recourse models. We apply these convex approximations to a surgery block allocation problem. Using numerical experiments we illustrate that these convex approximations are indeed good approximations. In fact, we obtain surprisingly good performance guarantees for the approximating solutions given that ambiguous mixed-integer recourse models are extremely hard to solve and we are unable to obtain the exact optimal solution. The performance guarantees are obtained by combining the multiple replications procedure (MRP) of Bayraksan and Morton (2006) with new tighter error bounds for convex approximations of simple integer recourse models.

Summarizing, the main contributions of our paper are

- Introducing a mean-MAD ambiguity set for continuous recourse models, for which the worst-case and best-case probability distributions are the same for every first-stage decision;
- Deriving a general approximation framework for ambiguous mixed-integer recourse models;
- Applying this approximation framework to mixed-ambiguous mixed-integer recourse models, and using numerical experiments to show that it yields good solutions.

The structure of our paper is as follows. In Section 2 we introduce our approach for two-stage ambiguous recourse models with continuous decision variables. Section 3 includes our new approximation framework for two-stage ambiguous recourse models with integer decision variables. In Section 4 we apply this framework to a surgery block allocation problem and we carry out numerical experiments. For reasons of space, several proofs of propositions and theorems are relegated to Appendix A.

2. Two-stage ambiguous continuous recourse models

In this section we describe our approach for solving ambiguous recourse models in case all decision variables are continuous. We only consider two-stage models here. However, the results can easily be generalized to a multi-stage setting. Although the results in this section appear to be known in the SP literature (Ben-Tal & Hochman, 1976), we are the first – to our knowledge – to make these results explicit in a two-stage setting. Moreover, in this section we will set the stage for our new results on ambiguous mixed-integer recourse models.

The ambiguous recourse model that we consider is

$$\inf_{x \in X} \sup_{\mathbb{P}_z \in \mathcal{P}_z} \mathbb{E}_{\mathbb{P}_z}[c^\top x + v(x,z)], \tag{1}$$

where $X = \{x \in \mathbb{R}^{n_1}_+ : Ax = b\}$ represents the set of feasible first-stage solutions, $\mathcal{P}_z$ is the ambiguity set for probability distributions, and $v(x,z)$ is the second-stage value function defined as a function of the first-stage variables x and the random parameters $z = (\xi, \omega)$:

$$v(x,z) = \inf_{y \in Y}\{q(\xi)^\top y : Wy = h(\omega) - T(\omega)x\}. \tag{2}$$

Here, y are the second-stage (or recourse) variables and $Y \subset \mathbb{R}^{n_2}_+$ is a polyhedral set. The second-stage costs $q(\xi)$, the technology matrix $T(\omega)$, and the right-hand side $h(\omega)$ depend on the random vector $z = (\xi, \omega)$. Moreover, since the recourse matrix W is deterministic, we say that the problem has *fixed recourse* (see, e.g., Shapiro et al., 2009). Throughout the paper, we take the following assumption.

Assumption 1. We assume that q, T, and h are affine functions of z and that *all components of z are independent*. Thus, in particular, $q(\xi)$ is independent from $T(\omega)$ and $h(\omega)$.

In problem (1), the here-and-now decisions x have to be made while the parameter z is unknown, and after the uncertain parameter z is revealed we are allowed to take recourse actions y to compensate for possible violations of the constraints $T(\omega)x = h(\omega)$. The objective is to minimize the sum of the direct costs $c^\top x$ and the worst-case expected costs $\sup_{\mathbb{P}_z \in \mathcal{P}_z} \mathbb{E}_{\mathbb{P}_z}[v(x,z)]$.

Here, the ambiguity set $\mathcal{P}_z$ is defined as

$$\mathcal{P}_z = \{\mathbb{P}_z : \text{supp}(z_i) \subseteq [a_i, b_i], \quad \mathbb{E}_{\mathbb{P}_z}[z_i] = \mu_i, \quad \mathbb{E}_{\mathbb{P}_z}|z_i - \mu_i| = d_i, \\ \mathbb{P}_z\{z_i \geq \mu_i\} = \beta_i, \quad z_i \perp z_j, \ i \neq j\}, \tag{3}$$

where $z_i \perp z_j$ means that z_i and z_j are stochastically independent. Postek et al. (2018) explain procedures to estimate these parameters from historical data. Moreover, BTH72 show that the ambiguity set $\mathcal{P}_z$ is non-empty if for all i we have $a_i < \mu_i < b_i$, $0 \leq d_i \leq \frac{2(b_i-\mu_i)(\mu_i-a_i)}{b_i-a_i}$, and

$$\frac{d_i}{2(b_i - \mu_i)} \leq \beta_i \leq 1 - \frac{d_i}{2(\mu_i - a_i)}.$$

Throughout this paper we refer to the ambiguity set $\mathcal{P}_z$ in (3) as a (μ, d, β) ambiguity set.

2.1. Worst-case expectation

In general in problem (1), the worst-case probability distribution $\mathbb{P}_z \in \mathcal{P}_z$ will differ for a different first-stage decision $x \in X$. However, for the (μ, d, β) ambiguity set $\mathcal{P}_z$ in (3), the worst-case distribution $\mathbb{P}_{\bar{z}}$ turns out to be the same for every first-stage decision so that the ambiguous recourse model in (1) reduces to

$$\inf_{x\in X} \mathbb{E}_{\mathbb{P}_{\bar{z}}}[c^\top x + v(x, \bar{z})],$$

where each component of $\bar{z}$ follows a known discrete distribution with at most three realizations. This result is summarized in Proposition 1 below. Its proof combines the fact that the second-stage value function $v(x, z)$ is convex in ω and concave in ξ (see, e.g., Fiacco & Kyparisis, 1986) with results from BTH72, who provide closed-form expressions for the worst-case expectations maximizing and minimizing the expectations of convex and concave functions.

Proposition 1. *The two-stage ambiguous continuous recourse model*

$$\inf_{x\in X} \sup_{\mathbb{P}_z\in\mathcal{P}_z} \mathbb{E}_{\mathbb{P}_z}\left[c^\top x + \inf_{y\in Y}\left\{q(\xi)^\top y : Wy = h(\omega) - T(\omega)x\right\}\right]$$

with (μ, d, β) ambiguity set $\mathcal{P}_z$ for $z = (\xi, \omega) \in \mathbb{R}^{n_\xi} \times \mathbb{R}^{n_\omega}$ as defined in (3) is equivalent to

$$\inf_{x\in X} \mathbb{E}_{\mathbb{P}_{\bar{z}}}\left[c^\top x + \inf_{y\in Y}\left\{q(\bar{\xi})^\top y : Wy = h(\bar{\omega}) - T(\bar{\omega})x\right\}\right], \tag{4}$$

where the worst-case random vector $\bar{z} = (\bar{\xi}, \bar{\omega}) \in \mathbb{R}^{n_\xi} \times \mathbb{R}^{n_\omega}$ has independent components with marginal distributions

$$\mathbb{P}\left\{\bar{\xi}_i = \mu_i - \frac{d_i}{2(1-\beta_i)}\right\} = 1 - \beta_i, \quad \textit{and}$$
$$\mathbb{P}\left\{\bar{\xi}_i = \mu_i + \frac{d_i}{2\beta_i}\right\} = \beta_i, \ \ i = 1, \ldots, n_\xi,$$

and

$$\mathbb{P}\{\bar{\omega}_i = a_{n_\xi+i}\} = \frac{d_{n_\xi+i}}{2(\mu_{n_\xi+i} - a_{n_\xi+i})}, \quad \mathbb{P}\left\{\bar{\omega}_i = b_{n_\xi+i}\right\}$$
$$= \frac{d_{n_\xi+i}}{2(b_{n_\xi+i} - \mu_{n_\xi+i})}, \mathbb{P}\left\{\bar{\omega}_i = \mu_{n_\xi+i}\right\}$$
$$= 1 - \frac{d_{n_\xi+i}}{2(\mu_{n_\xi+i} - a_{n_\xi+i})} - \frac{d_{n_\xi+i}}{2(b_{n_\xi+i} - \mu_{n_\xi+i})}$$

for $i = 1, \ldots, n_\omega$.

Proof. Follows directly from BTH72 since the second-stage value function $v(x, z)$ defined in (2) is convex in ω and concave in ξ for every feasible first-stage solution $x \in X$. □

Remark 1. Note that the worst-case distribution $\bar{\omega}$ does not depend on the parameter β. This means that for the random parameters ω we do not have to estimate the probability that ω_i exceeds its mean to obtain the worst-case expectation. In Proposition 2, we show that we do require β to obtain the best-case distribution $\underline{\omega}$ of ω.

Remark 2. Observe that the worst-case probability distribution $\bar{\omega}$ is not necessarily feasible. Indeed, we may have $\mathbb{P}\{\bar{\omega}_i \geq \mu_{n_\xi+i}\} < \beta_{n_\xi+i}$ for some $1 \leq i \leq n_\omega$. However, by slightly adjusting the probability distribution of $\bar{\omega}_i$ we may obtain arbitrarily close approximations of $\bar{\omega}_i$ that are feasible, see BTH72. Hence, it makes sense to refer to $\bar{\omega}$ as the worst-case distribution.

2.2. Best-case expectation

Similar as for the worst-case expectation we can obtain the best-case expectation over all probability distributions in the (μ, d, β) ambiguity set $\mathcal{P}_z$ by using results of BTH72. Again, the best-case distribution $P_{\underline{z}}$ is a discrete distribution with at most three realizations per component that does not depend on the first-stage decision x.

Proposition 2. *The two-stage ambiguous continuous recourse model*

$$\inf_{x\in X} \inf_{\mathbb{P}_z\in\mathcal{P}_z} \mathbb{E}_{\mathbb{P}_z}\left[c^\top x + \inf_{y\in Y}\left\{q(\xi)^\top y : Wy = h(\omega) - T(\omega)x\right\}\right]$$

with (μ, d, β) ambiguity set $\mathcal{P}_z$ for $z = (\xi, \omega) \in \mathbb{R}^{n_\xi} \times \mathbb{R}^{n_\omega}$ as defined in (3) is equivalent to

$$\inf_{x\in X} \mathbb{E}_{\mathbb{P}_{\underline{z}}}\left[c^\top x + \inf_{y\in Y}\left\{q(\underline{\xi})^\top y : Wy = h(\underline{\omega}) - T(\underline{\omega})x\right\}\right], \tag{5}$$

where the best-case random vector $\underline{z} = (\underline{\xi}, \underline{\omega}) \in \mathbb{R}^{n_\xi} \times \mathbb{R}^{n_\omega}$ has independent components with marginal distributions

$$\mathbb{P}\left\{\underline{\xi}_i = a_i\right\} = \frac{d_i}{2(\mu_i - a_i)}, \quad \mathbb{P}\left\{\underline{\xi}_i = b_i\right\} = \frac{d_i}{2(b_i - \mu_i)},$$
$$\mathbb{P}\left\{\underline{\xi}_i = \mu_i\right\} = 1 - \frac{d_i}{2(\mu_i - a_i)} - \frac{d_i}{2(b_i - \mu_i)}$$

for $i = 1, \ldots, n_\xi$ and

$$\mathbb{P}\left\{\underline{\omega}_i = \mu_{n_\xi+i} - \frac{d_{n_\xi+i}}{2(1-\beta_{n_\xi+i})}\right\} = 1 - \beta_{n_\xi+i},$$
$$\mathbb{P}\left\{\underline{\omega}_i = \mu_{n_\xi+i} + \frac{d_{n_\xi+i}}{2\beta_{n_\xi+i}}\right\} = \beta_{n_\xi+i}.$$

for $i = 1, \ldots, n_\omega$.

Proof. Follows directly from BTH72 since the second-stage value function $v(x, z)$ defined in (2) is convex in ω and concave in ξ for every feasible first-stage solution $x \in X$. □

Notice that since $v(x, z)$ is concave in ξ and convex in ω the worst-case distribution of ξ has the same structure as the best-case distribution of ω.

Remark 3. One can argue that the distribution parameters a, b, μ, d and β of the (μ, d, β) uncertainty set $\mathcal{P}_z$ in Propositions 1 and 2 are subject to estimation and thus, uncertainty, and that it should be accounted for. On this note, we state that the dependence of the worst- and best-case expectations on parameters a, b, and μ depends on the problem at hand and hence, the only way to accommodate for this uncertainty is to try multiple values. With respect to dependence on d, the worst-case expectation of a convex function is nondecreasing in d and hence, the maximum value of the expectation in Proposition 1 is attained at the largest possible value of d. With respect to uncertainty in β, the worst-case expectation of the recourse function of Proposition 1 (best-case expectation of Proposition 2) is concave (convex) w.r.t. to each component β_i separately. Thus, in case of uncertainty about β, one can use this fact to obtain more (less) conservative upper (lower) bounds on the worst-case (best-case) expectations by alternatingly optimizing the decisions and maximizing (minimizing) the recourse function w.r.t. β_i.

2.3. Value of distributional information

Best-case expectation is a useful complement to the worse-case expectation, since the difference between the two can be interpreted as an upper bound on the *value of distributional information* (VDI, Delage et al., 2015). The VDI is the price one pays for not knowing the true probability distribution $\mathbb{P}_z^\star$ of z. Thus, it is the difference in expected costs between implementing the worst-case expectation solution and the optimal solution when the probability distribution $\mathbb{P}_z^\star$ is known. At the same time, it can also be interpreted as the amount we are willing to pay for complete knowledge of the probability distribution of z, and thus a maximum on the amount we are willing to invest in gathering more information on this probability distribution.

Definition 1. Consider the ambiguous recourse model

$$\inf_{x\in X}\sup_{\mathbb{P}_z\in\mathcal{P}_z}\mathbb{E}_{\mathbb{P}_z}[c^\top x+v(x,z)], \tag{6}$$

and assume that $\mathbb{P}_z^\star\in\mathcal{P}_z$ is the true probability distribution of the random parameters z. Then, the *value of distributional information* is defined as

$$\text{VDI} := \mathbb{E}_{\mathbb{P}_z^\star}[c^\top\bar{x}+v(\bar{x},z)]-\inf_{x\in X}\mathbb{E}_{\mathbb{P}_z^\star}[c^\top x+v(x,z)],$$

where $\bar{x}\in X$ denotes the optimal solution to (6).

In general, the VDI can only be determined when the true probability distribution $\mathbb{P}_z^\star$ is known. However, for (μ, d, β) ambiguity sets $\mathcal{P}_z$ it is possible to upper bound the VDI *without knowing* $\mathbb{P}_z^\star$, since the worst-case and best-case distributions are the same for every first-stage decision x.

Lemma 1. *Consider the ambiguous recourse model*

$$\inf_{x\in X}\sup_{\mathbb{P}_z\in\mathcal{P}_z}\mathbb{E}_{\mathbb{P}_z}[c^\top x+v(x,z)],$$

with $\mathcal{P}_z$, denoting the (μ, d, β) ambiguity set as defined in (3). Then,

$$\text{VDI} \le \mathbb{E}_{\mathbb{P}_{\bar{z}}}[c^\top\bar{x}+v(\bar{x},\bar{z})]-\mathbb{E}_{\mathbb{P}_{\underline{z}}}[c^\top\underline{x}+v(\underline{x},\underline{z})],$$

where $\bar{x}$ and $\underline{x}$ are the optimal worst-case and best-case solutions of (4) and (5), respectively.

Proof. Follows directly from the definition of VDI and the inequalities

$$\mathbb{E}_{\mathbb{P}_z^\star}[c^\top\bar{x}+v(\bar{x},z)]\le\mathbb{E}_{\mathbb{P}_{\bar{z}}}[c^\top\bar{x}+v(\bar{x},\bar{z})],$$

and

$$\inf_{x\in X}\mathbb{E}_{\mathbb{P}_z^\star}[c^\top x+v(x,z)]\ge\inf_{x\in X}\mathbb{E}_{\mathbb{P}_{\underline{z}}}[c^\top x+v(x,\underline{z})]=\mathbb{E}_{\mathbb{P}_{\underline{z}}}[c^\top\underline{x}+v(\underline{x},\underline{z})],$$

that hold since the worst-case and best-case distributions $\mathbb{P}_{\bar{z}}$ and $\mathbb{P}_{\underline{z}}$ are the same for every first-stage decision x. □

2.4. Solution methods for continuous recourse models

Propositions 1 and 2 show that we can obtain the worst- and best-case expectation, respectively, by solving a standard continuous recourse model. That is why we review techniques for solving such models in this section. We consider the standard recourse model

$$\inf_{x\in X}\left\{c^\top x+\mathbb{E}_{\mathbb{P}_z}[v(x,z)]\right\}, \tag{7}$$

where all decision variables are continuous and the probability distribution $\mathbb{P}_z$ has finite support. Enumerating all K scenarios of z, we can rewrite the problem in (7) as

$$\inf_{x\in X}\mathbb{E}_{\mathbb{P}_z}[c^\top x+v(x,z)]=\inf_{x\in X}\sum_{k=1}^K p_k[c^\top x+v(x,z^k)],$$

where p_k denotes the probability of scenario z^k, $k=1,\dots,K$. The latter problem can be rewritten in its deterministic equivalent form, yielding

$$\inf_{x\in X, y_k\in Y}\left\{c^\top x+\sum_{k=1}^K p_k q(\xi^k)y_k : Wy_k=h(\omega^k)-T(\omega^k)x,\ k=1,\dots,K\right\}.$$

Observe that the worst-case and best-case probability distributions $\mathbb{P}_{\bar{z}}$ and $\mathbb{P}_{\underline{z}}$ have $K:=2^{n_\xi}\times 3^{n_\omega}$ and $K:=3^{n_\xi}\times 2^{n_\omega}$ scenarios, respectively. Hence, the number of scenarios is exponential in the number of random parameters. From a robust optimization point of view this means that the problem in (7) is intractable. Indeed, Dyer and Stougie (2006) show that these recourse models are #*P*-hard. Nevertheless, there has been a vast amount of work in the SP literature that deals with this kind of problems, yielding efficient (approximate) solution methods to these recourse models. The fact that the size of the problem grows exponentially in the number of random parameters is common in SP, and many SP approaches are aimed at reducing the number of scenarios.

One of the most frequently used solution methods is the sample average approximation (SAA), discussed in, e.g., Shapiro et al. (2009). The idea of this method is to replace the original distribution of z in (7) by a sample z^s, $s=1,\dots,N_s$, where N_s is much smaller than the number of scenarios of z, yielding

$$\inf_{x\in X}\left\{c^\top x+\frac{1}{N_s}\sum_{s=1}^{N_s}v(x,z^s)\right\}. \tag{8}$$

If the sample size N_s is small, then the approximation in (8) is easier to solve than the original model in (7). We may solve (8) for several different samples of z yielding (possibly) different first-stage solutions x, and use an out-of-sample test to determine the best among them (or average them, see Sen & Liu (2016)).

Alternatively, we may use other approaches to reduce the number of scenarios. For example, Dupačová, Gröwe-Kuska, and Römisch (2003) and Heitsch and Römisch (2003) do so by combining similar scenarios. Pflug (2001) uses the Wasserstein metric to construct a discrete probability distribution (with few scenarios) that minimizes the distance between the original and approximating distribution. His method can also be applied to multi-stage recourse models. Approximations relying on a reduced scenario set are justified by stability results of, e.g., Römisch (2003) which shows that a small change in the distributions of the random parameters only result in a small change in the optimal first-stage solutions.

For two-stage recourse models with only a modest number of scenarios efficient solution methods are available. Most of them rely on decomposition of the problem and are variants of the L-shaped algorithm of van Slyke and Wets (1969); see, e.g., Ruszczyński (1986) and Higle and (1991) for well-known examples. We refer to Zverovich, Fábián, Ellison, and Mitra (2012) for a recent survey comparing several decomposition methods.

So far we have only discussed how to obtain a first-stage solution. However, when this solution is obtained by solving an approximation of the original recourse model, then we may use sampling to assess the quality of the solution; see, e.g., the Multiple Replications Procedure (MRP) of Bayraksan and Morton (2009). Different sampling methods, such as Latin Hypercube sampling, may be used to reduce the bias and sample variance of the optimality gap of the approximating solution. We use the MRP to assess the quality of a surgery-to-OR assignment in the surgery block allocation problem of Section 4.

3. Two-stage ambiguous mixed-integer recourse models

In this section we consider the two-stage ambiguous mixed-integer recourse model

$$\eta^* := \inf_{x\in X} \sup_{\mathbb{P}_z\in\mathcal{P}_z} \mathbb{E}_{\mathbb{P}_z}[c^\top x + v(x,z)]. \tag{9}$$

Similar as in Section 2, the set $\mathcal{P}_z$ represents the (μ, d, β) ambiguity set defined in (3), and $v(x,z)$ the second-stage value function defined in (2). The difference with the models in Section 2 is that here the feasible sets X and Y may also impose integrality restrictions on the first- and second-stage decision variables x and y, respectively.

The difficulty of having integer decision variables in the second-stage is that the second-stage value function $v(x,z)$ is generally not convex in ω so that the results of BTH72 cannot be applied. To deal with this difficulty we develop a general approximation framework for two-stage ambiguous mixed-integer recourse models. The key idea in developing this framework is to approximate $v(x,z)$ by a value function $\hat{v}(x,z)$ that is concave in ξ and convex in ω so that the results of BTH72 do apply. We derive error bounds for two types of approximations, one with $v(x,z)$ replaced by $\hat{v}(x,z)$ in (9), and one in which we keep the $v(x,z)$ but incorrectly assume that $\mathbb{P}_{\bar{z}}$ is the worst-case probability distribution in (9).

The fact that we are also able to (approximately) obtain the best-case expectation for ambiguous mixed-integer recourse models illustrates that our approximation framework is very suitable for combination with the (μ, d, β) ambiguity set. We apply the approximation framework to mixed-ambiguity models in which some of the distributions of the random parameters in the model are known and others are ambiguous. In this setting we can use existing convex approximations for standard mixed-integer recourse models from the literature.

In Section 3.1 we introduce the approximation framework for two-stage ambiguous mixed-integer recourse models and in Section 3.2 we apply this framework to the mixed-ambiguous setting.

3.1. General approximation framework

Similar as in Section 2, we consider the worst-case and best-case expectation separately in Sections 3.1.1 and 3.1.2, respectively.

3.1.1. Worst-case mixed-integer expectation

We develop a general approximation framework for ambiguous mixed-integer recourse models by approximating $v(x,z)$ by a function $\hat{v}(x,z)$ which satisfies the assumptions for applying BTH72. We call such functions *convex approximations*.

Definition 2. We call $\hat{v}(x,z)$ a *convex approximation* of the second-stage value function v if

(i) $\hat{v}(x,z)$ is convex in x for every given $z = (\xi, \omega)$,
(ii) $\hat{v}(x,z)$ is convex in ω for every given ξ and $x\in X$,
(iii) $\hat{v}(x,z)$ is concave in ξ for every given ω and $x\in X$.

Remark 4. Observe that the continuous second-stage value function defined in Section 2, satisfies properties (i)–(iii) of Definition 2.

Remark 5. To be able to apply BTH72, we do not need property (i) in Definition 2. However, for optimization purposes it is highly desirable that $\hat{v}(x,z)$ is convex in the first-stage decision vector x.

If $\hat{v}(x,z)$ is a convex approximation of $v(x,z)$, then, we may approximate (9) by replacing $v(x,z)$ by $\hat{v}(x,z)$, obtaining

$$\hat{\eta} := \inf_{x\in X} \sup_{\mathbb{P}_z\in\mathcal{P}_z} \mathbb{E}_{\mathbb{P}_z}[c^\top x + \hat{v}(x,z)] \tag{10}$$

$$= \inf_{x\in X} \mathbb{E}_{\mathbb{P}_{\bar{z}}}[c^\top x + \hat{v}(x,\bar{z})], \tag{11}$$

where the equality in (11) follows from applying the result of BTH72 to (10). The approximating problem is an optimization problem for which the distributions of the random parameters are known. Since $\hat{v}(x,z)$ is convex in x for every z, the optimization problem can be solved efficiently using existing solution methods from convex optimization. To guarantee the quality of the approximate solution $\hat{x}$ obtained from solving the optimization problem in (11), we derive an error bound on the optimality gap $G(\hat{x}) - \eta^*$, where $G(\hat{x})$ represents the true objective value of the solution $\hat{x}$:

$$G(x) := \sup_{\mathbb{P}_z\in\mathcal{P}_z} \mathbb{E}_{\mathbb{P}_z}[c^\top x + v(x,z)], \quad x\in X. \tag{12}$$

In fact, we show in Theorem 1 below that $|\hat{\eta} - \eta^*| \le \|v - \hat{v}\|_\infty$ and $G(\hat{x}) - \eta^* \le 2\|v - \hat{v}\|_\infty$, where

$$\|v - \hat{v}\|_\infty := \sup_{x,z} \left\{ |v(x,z) - \hat{v}(x,z)| : x\in X \right\}.$$

Interestingly, we may approximate the optimization model in (11) by replacing $\hat{v}(x,z)$ by the original mixed-integer recourse function $v(x,z)$ to obtain the approximating model

$$\tilde{\eta} := \inf_{x\in X} \mathbb{E}_{\mathbb{P}_{\bar{z}}}[c^\top x + v(x,\bar{z})]. \tag{13}$$

This model indirectly approximates the original mixed-integer recourse model (9), but it can also be derived directly from (9) by assuming that $\mathbb{P}_{\bar{z}}$ is the worst-case distribution in that model. However, using the interpretation of an indirect approximation via the convex approximating model in (11), we can derive an error bound for the approximate solution $\tilde{x}$ obtained from solving (13).

Theorem 1. *Consider the two-stage ambiguous mixed-integer recourse model defined in (9) and let $\hat{v}(x,z)$ be a convex approximation, conform Definition 2, of the second-stage value function $v(x,z)$ defined in (2). Let $\hat{x}$ and $\tilde{x}$ denote optimal solutions of the approximating models defined in (11) and (13), respectively. Then,*

(i) $|\hat{\eta} - \eta^*| \le \|v - \hat{v}\|_\infty$ *and* $G(\hat{x}) - \eta^* \le 2\|v - \hat{v}\|_\infty$,
(ii) $0 \le \eta^* - \tilde{\eta} \le 2\|v - \hat{v}\|_\infty$ *and* $G(\tilde{x}) - \eta^* \le 2\|v - \hat{v}\|_\infty$.

Furthermore, since the upper bound on $G(\tilde{x}) - \eta^*$ holds for every approximation $\hat{v}$, it actually holds for the best convex approximation:

$$G(\tilde{x}) - \eta^* \le 2\inf_{\hat{v}} \Big\{ \|v - \hat{v}\|_\infty : \ \hat{v}(x,z) \text{ is a convex approximation of } v(x,z) \Big\}.$$

Proof. See Appendix A. □

Remark 6. The error bounds in Theorem 1 do not only hold for the mixed-integer second-stage value function $v(x,z)$ defined in (2), but for all functions v of x and z in general. We use this in Section 3.2 for mixed-ambiguity models where we replace $v(x,z)$ by an *expected* value function $Q(x, z)$.

The error bounds in Theorem 1 depend on the maximum difference between $v(x,z)$ and $\hat{v}(x,z)$ over all feasible first-stage solutions $x\in X$ and random parameters z. However, as can easily be induced from the proof of Theorem 1, the error bound actually only depends on the maximum difference between $v(x,z)$ and $\hat{v}(x,z)$ over the random parameters z in the optimal solution x^* and the approximating solution $\hat{x}$. Since we generally are not able to obtain the exact optimal solution x^*, the error bound in Theorem 1 is typically more convenient. However, we may obtain tighter bounds than in Theorem 1, since we can compute the approximating solution $\hat{x}$ and evaluate the maximum difference between $v(\hat{x}, z)$ and $\hat{v}(\hat{x}, z)$ over z.

Corollary 1. *Consider the setting of Theorem 1, and let $\hat{x}$ be a feasible first-stage solution. Then,*

$$G(\hat{x}) - \eta^* \le \hat{G}(\hat{x}) - \tilde{\eta} + \sup_{\mathbb{P}_z \in \mathcal{P}_z} \mathbb{E}_{\mathbb{P}_z}\Big[v(\hat{x}, z) - \hat{v}(\hat{x}, z)\Big],$$

where $\hat{G}(\hat{x})$ equals $G(\hat{x})$ with $v(x, z)$ replaced by $\hat{v}(x, z)$.

Proof. Follows directly from the fact that

$$G(\hat{x}) \le \hat{G}(\hat{x}) + \sup_{\mathbb{P}_z \in \mathcal{P}_z} \mathbb{E}_{\mathbb{P}_z}[v(\hat{x}, z) - \hat{v}(\hat{x}, z)]$$

and $\tilde{\eta} \le \eta^*$ by Theorem 1. □

Remark 7. The bound in Corollary 1 holds for all first-stage solutions $\hat{x}$, and thus also for the optimal solutions $\hat{x}$ and $\tilde{x}$ of the approximating models in (11) and (13), respectively.

To bound the optimality gap in Corollary 1 we need to obtain an upper bound on $\hat{G}(\hat{x}) - \tilde{\eta}$. The advantage of dealing with this difference, rather than $G(\hat{x}) - \eta^*$, is that $\hat{G}(\hat{x})$ and $\tilde{\eta}$ correspond to the objective value at $\hat{x}$ and the optimal objective value of (13), respectively, of *standard* mixed-integer recourse models in which the distributions of all random parameters are known. In the numerical experiments of Section 4 we will use the MRP to obtain an upper bound on $\hat{G}(\hat{x}) - \tilde{\eta}$. Moreover, in these experiments we will show that we may obtain a very tight bound on $\sup_{\mathbb{P}_z \in \mathcal{P}_z} \mathbb{E}_{\mathbb{P}_\zeta}[v(\hat{x}, z) - \hat{v}(\hat{x}, z)]$.

3.1.2. Best-case mixed-integer expectation

So far we have only discussed how to approximate the worst-case expectation of mixed-integer recourse models. However, in an analoguous way we can deal with the best-case expectation

$$\eta^* := \inf_{x \in X} \underline{G}(x), \tag{14}$$

where

$$\underline{G}(x) := \inf_{\mathbb{P}_z \in \mathcal{P}_z} \mathbb{E}_{\mathbb{P}_z}\Big[c^\top x + v(x, z)\Big].$$

Similar as for the worst-case expectation, we approximate (14) by replacing $v(x, z)$ by a convex approximation $\hat{v}(x, z)$, yielding

$$\hat{\eta} := \inf_{x \in X} \inf_{\mathbb{P}_z \in \mathcal{P}_z} \mathbb{E}_{\mathbb{P}_z}[c^\top x + \hat{v}(x, z)] \tag{15}$$

$$= \inf_{x \in X} \mathbb{E}_{\mathbb{P}_{\underline{z}}}[c^\top x + \hat{v}(x, \underline{z})], \tag{16}$$

where again the equality in (16) follows from applying the results of BTH72 to (15). Moreover, we can also approximate (14) by assuming that $\mathbb{P}_{\underline{z}}$ is the best-case distribution:

$$\tilde{\eta} := \inf_{x \in X} \mathbb{E}_{\mathbb{P}_{\underline{z}}}[c^\top x + v(x, \underline{z})]. \tag{17}$$

In Theorem 2 we give error bounds for using the convex approximations in (16) and (17).

Theorem 2. *Consider the two-stage ambiguous mixed-integer recourse model defined in (14) and let $\hat{v}(x, z)$ be a convex approximation, conform Definition 2, of the second-stage value function $v(x, z)$ defined in (2). Let $\hat{x}$ and $\tilde{x}$ denote optimal solutions of the approximating models defined in (16) and (17), respectively. Then,*

(i) $|\hat{\eta} - \eta^*| \le \|v - \hat{v}\|_\infty$ *and* $\underline{G}(\hat{x}) - \eta^* \le 2\|v - \hat{v}\|_\infty$,
(ii) $0 \le \eta^* - \tilde{\eta} \le 2\|v - \hat{v}\|_\infty$ *and* $\underline{G}(\tilde{x}) - \eta^* \le 4\|v - \hat{v}\|_\infty$.

Furthermore, since the upper bound on $\underline{G}(\tilde{x}) - \eta^*$ holds for every approximation $\hat{v}$, it actually holds for the best convex approximation:

$$\underline{G}(\tilde{x}) - \eta^* \le 4 \inf_{\hat{v}} \Big\{ \|v - \hat{v}\|_\infty : \ \hat{v}(x, z) \text{ is a convex approximation of } v(x, z) \Big\}.$$

Proof. See Appendix A. □

The computational complexity of the approximating models in (16) and (17) for the best-case expectation are similar to those for the worst-case expectation. The approximating models in (11) and (16), obtained by replacing $v(x, z)$ by $\hat{v}(x, z)$, are the easiest to solve since they are convex optimization problems. In contrast, the approximating models in (13) and (17), obtained by assuming that $\mathbb{P}_{\bar{z}}$ and $\mathbb{P}_{\underline{z}}$ are the worst- and best-case distributions, respectively, are non-convex standard two-stage mixed-integer recourse model for which the distributions of the random parameters are known. These models are significantly harder to solve than convex optimization problems, but at the same time easier to solve than their ambiguous counterparts.

In addition, the error bounds for these latter approximating models seem twice as large as for the convex approximating models in (11) and (16). However, this is only true for the *best* convex approximation $\hat{v}(x, z)$. If no good convex approximation $\hat{v}(x, z)$ of $v(x, z)$ is known, then we can still approximate the ambiguous mixed-integer recourse models by (13) and (17), and obtain a good first-stage solution $\tilde{x}$ as long as there *exists* a good convex approximation $\hat{v}(x, z)$. In the numerical experiment of Section 4 we show that indeed a good convex approximation $\hat{v}(x, z)$ of $v(x, z)$ may exist.

3.2. Mixed-ambiguity stochastic mixed-integer programs

In this section we consider so-called mixed-ambiguity problems. In these problems, the distributions of some of the random variables are ambiguous and others are known. This is useful to model practical problems in which there are different sources of uncertainty in the problem. For example, in multi-product production and inventory problems there may be products for which a long history of demand data is available, so that their demand distributions can be accurately estimated, whereas other products may be relatively new, so that limited demand data for these products is available.

Similar as in the previous sections we let $z = (\xi, \omega)$ denote the ambiguous random variables. However, we also introduce random variables ζ of which the distributions are known.

Assumption 2. We assume that q, T, and h are affine functions of (z, ζ) and that all components of z and ζ are independent.

The mixed-ambiguous mixed-integer recourse model that we consider is

$$\eta^* := \inf_{x \in X} \sup_{\mathbb{P}_z \in \mathcal{P}_z} \mathbb{E}_{\mathbb{P}_z}[c^\top x + Q(x, z)], \tag{18}$$

where $Q(x, z)$ is defined as

$$Q(x, z) = \mathbb{E}_{\mathbb{P}_\zeta}\left[\inf_{y \in Y}\Big\{q(\xi, \zeta)^\top y : Wy = h(\omega, \zeta) - T(\omega, \zeta)x\Big\}\right]. \tag{19}$$

The problem in (18) is similar to the ambiguous mixed-integer recourse model (9) of Section 3.1, but with v replaced by Q. We use Q here instead of v to emphasize that it is not just a second-stage value function, but an *expected* value function, where the expectation is taken over all random variables ζ with a known distribution.

This expected value function Q is key to solving (18), since if Q is convex in ω and concave in ξ, then we may apply the result of BTH72 to obtain the worst-case distribution $\mathbb{P}_{\bar{z}}$ of z. For example, Klein Haneveld et al. (2006) show that this may be true for *simple integer recourse* models. In general, however, $Q(x, z)$ is not convex in ω, and we have to resort to the approximation framework of Section 3.1. This is possible, since there exist convex approximations $\hat{Q}(x,z)$ of $Q(x, z)$ in the literature with corresponding error bounds on $\|Q-\hat{Q}\|_\infty$.

In Section 3.2.1 we discuss the case where the simple integer recourse function $Q(x, z)$ is convex in ω and concave in ξ, and in Section 3.2.2 we discuss existing and new error bounds for convex approximations $\hat{Q}(x,z)$ of this simple integer recourse function $Q(x, z)$. Although we only discuss simple integer recourse models in Sections 3.2.1 and 3.2.2, we want to stress that there are other convex approximations for more general standard mixed-integer recourse models in the literature that can easily be embedded in the mixed-ambiguous setting of this section using the approximation framework of Section 3.1. We refer to Romeijnders, Stougie, and van der Vlerk (2014) for an overview on these convex approximations for standard mixed-integer recourse models.

3.2.1. Mixed-ambiguous simple integer recourse models

The one-sided simple integer recourse model, introduced in Louveaux and van der Vlerk (1993), is a special case of (18) for which a closed-from expression for the second-stage value function can be obtained. Assuming that q and T are ambiguous and the distribution of h is known, so that we can write $h(\zeta)=\zeta$, the expected value function $Q(x, z)$ is given by

$$Q(x,z)=\sum_{i=1}^{m}\mathbb{E}_{\mathbb{P}_{\zeta_i}}\left[q_i(\xi)\lceil\zeta_i-T_i(\omega)x\rceil^+\right],\quad x\in\mathbb{R}^{n_1},\tag{20}$$

where $\lceil s\rceil^+:=\max\{0,\lceil s\rceil\}$, $s\in\mathbb{R}$ and $T_i(\omega)$ is the ith row of the matrix $T(\omega)$. It is not hard to verify that Q is concave in ξ. Interestingly, however, Klein Haneveld et al. (2006) show that this simple integer recourse function Q may also be convex in the tender variables $u=T(\omega)x$, and thus in ω, if the underlying random vector ζ is continuously distributed and every marginal probability density function f_i of ζ_i can be expressed as

$$f_i(s)=H_i(s+1)-H_i(s),\quad s\in\mathbb{R},\tag{21}$$

for some cumulative distribution function H_i with finite mean. This implies that under these conditions the worst-case distribution $\mathbb{P}_{\bar{z}}$ of z can be derived using the results of BTH72 (this worst-case distribution is the same for every first-stage decision x).

Proposition 3. *Consider the mixed-ambiguous simple integer recourse model*

$$\inf_{x\in X}\sup_{\mathbb{P}_z\in\mathcal{P}_z}\mathbb{E}_{\mathbb{P}_z}[c^\top x+Q(x,z)],\tag{22}$$

where Q is defined in (20) and the ambiguity set $\mathcal{P}_z$ for the distributions $\mathbb{P}_z$ of z is defined as in (3). If each random variable ζ_i has a probability density function f_i satisfying (21), then the optimization problem in (22) is equivalent to

$$\inf_{x\in X}\mathbb{E}_{\mathbb{P}_{\bar{z}}}[c^\top x+Q(x,\bar{z})],$$

where the worst-case distribution $\mathbb{P}_{\bar{z}}$ of z is defined as in Proposition 1.

Remark 8. We can derive an analogue result for the best-case expectation of a mixed-ambiguous simple integer recourse model if each random variable ζ_i has a pdf f_i satisfying (21).

If a probability density function f_i of some random parameter ζ_i does not satisfy (21), then a natural approach is to approximate it by a density function $\hat{f}_i$ that is approximately the same as f_i, but does satisfy (21), yielding a convex approximation $\hat{Q}(x,z)$ of $Q(x, z)$. This is the main idea behind the so-called α-approximations derived in Klein Haneveld et al. (2006), and their generalization to complete integer recourse models by van der Vlerk (2004). We describe this convex approximation and its corresponding error bound in the next section.

3.2.2. Convex approximations and error bounds

The α-approximations of Klein Haneveld et al. (2006) can be obtained using $\hat{H}_i(s):=F_i(\lceil s-\alpha_i\rceil+\alpha_i)$ in (21) to generate approximate probability density functions $\hat{f}_i$. For every $\alpha\in\mathbb{R}^m$, this yields the approximating expected value function

$$\hat{Q}_\alpha(x,z)=\sum_{i=1}^{m}\mathbb{E}_{\mathbb{P}_{\zeta_i}}\left[q_i(\xi)(\lceil\zeta_i-\alpha_i\rceil+\alpha_i-T_i(\omega)x)^+\right],\quad x\in X.\tag{23}$$

Similarly, the so-called shifted LP-relaxation approximation of Romeijnders et al. (2016b)

$$\hat{Q}(x,z)=\sum_{i=1}^{m}\mathbb{E}_{\mathbb{P}_{\zeta_i}}\left[q_i(\xi)(\zeta_i+1/2-T_i(\omega)x)^+\right],\quad x\in X,\tag{24}$$

can be derived using $\hat{H}_i(s):=F_i(s-1/2)$ in (21). The result of both approximations is that we simultaneously remove the round-up operator in the expression for $Q(x, z)$ in (20) and adjust the random parameters ζ_i. For the shifted LP-relaxation approximation of (24), we add 1/2 to the random parameter ζ_i since on average this is (approximately) the effect of rounding.

For these convex approximations upper bounds have been derived on $\|Q-\hat{Q}_\alpha\|_\infty$ and $\|Q-\hat{Q}\|_\infty$ using the *total variations* of the probability density functions of the random parameters ζ.

Definition 3. Let $f:\mathbb{R}\to\mathbb{R}$ be a real-valued function, and let $I\in\mathbb{R}$ be an interval. Let $\Pi(I)$ denote the set of all finite ordered sets $P=\{x_0,x_1,\dots,x_{N+1}\}$ with $x_0<x_1<\cdots<x_{N+1}\in I$. Then, the total variation of f on I, denoted $|\Delta|f(I)$, is defined as

$$|\Delta|f(I)=\sup_{P\in\Pi(I)}\sum_{i=0}^{N}|f(x_{i+1})-f(x_i)|.$$

We write $|\Delta|f=|\Delta|f(\mathbb{R})$.

Proposition 4. *Consider the simple integer expected value function $Q(x, z)$ defined in (20) and the convex approximations $\hat{Q}_\alpha(x,z)$ and $\hat{Q}(x,z)$ defined in (23) and (24), respectively. Then,*

$$\|Q-\hat{Q}_\alpha\|_\infty\le\sup_\xi\left\{\sum_{i=1}^{m}q_i(\xi)h(|\Delta|f_i)\right\}\qquad\text{and}$$

$$\|Q-\hat{Q}\|_\infty\le\frac{1}{2}\sup_\xi\left\{\sum_{i=1}^{m}q_i(\xi)h(|\Delta|f_i)\right\},$$

where $h:[0,\infty)\mapsto\mathbb{R}$ is defined as

$$h(t)=\begin{cases}t/8, & t\le 4,\\ 1-2/t, & t\ge 4,\end{cases}\tag{25}$$

and where $|\Delta|f_i$ are the total variations of the marginal density functions f_i of ζ_i for $i=1,\dots,m$.

Proof. See Romeijnders et al. (2016b). □

The error bounds on $\|Q-\hat{Q}_\alpha\|_\infty$ and $\|Q-\hat{Q}\|_\infty$ in Proposition 4 can be used in Theorem 1 to obtain an error bound for when $\hat{Q}_\alpha(x,z)$ and $\hat{Q}(x,z)$ are used as convex approximations for the mixed-ambiguous model in (18). The error bounds are small, and thus the convex approximations are good, if the total variations $|\Delta|f_i$ of the probability density functions f_i of the random parameters ζ in the model are small. This is for example the case if $\zeta_i, = 1, \ldots, m$ are normally distributed with large variances.

The error bounds in Proposition 4 may be tightened for a fixed first-stage solution $\hat{x} \in X$, conform Corollary 1. Below in Proposition 5 we present such a tighter bound for the shifted LP-relaxation approximation $\hat{Q}(x, z)$ only, since the error bound for this approximation was already twice smaller than for the α-approximations.

Proposition 5. *Consider the simple integer expected value function* $Q(x, z)$ *defined in (20) and the convex approximation* $\hat{Q}(x, z)$ *defined in (24). Then for every feasible first-stage solution* $\hat{x} \in X$, *we have*

$$\sup_{\mathbb{P}_z \in \mathcal{P}_z} \mathbb{E}_{\mathbb{P}_z}\left[Q(\hat{x}, z) - \hat{Q}(\hat{x}, z)\right] \le \frac{1}{2} \sup_{\mathbb{P}_z \in \mathcal{P}_z} \mathbb{E}_{\mathbb{P}_z} \times \left[\sum_{i=1}^{m} q_i(\xi) h\Big(2|\Delta| f_i([T_i(\omega)\hat{x} - 1/2, +\infty))\Big)\right],$$

where h is defined in (25).

Proof. See Appendix A. □

Contrary to the error bounds in Proposition 4, the bound in Proposition 5 is new and may also be used to improve existing error bounds for convex approximations of standard two-stage mixed-integer recourse models.

4. Surgery block allocation

In this section we apply the approximation framework that we have derived in Section 3 for ambiguous mixed-integer recourse models to an adapted version of the surgery block allocation problem introduced by Denton, Miller, Balasubramanian, and Huschka (2010). We are able to obtain surgery-to-OR allocations with limited computational effort but with surprisingly good performance guarantees, given the fact that ambiguous mixed-integer recourse models are extremely difficult to solve exact.

4.1. Problem formulation

In the surgery block allocation problem, several surgeries with random durations have to be assigned to ORs before the durations of these surgeries are known. The problem can be formulated as a two-stage recourse model, where in the first stage we have to determine how many ORs to open and we have to assign the surgeries to the ORs. With N denoting the number of surgeries that have to be performed, we define π_{ij} for every $i, j = 1, \ldots, N$, as a binary variable equal to 1 if surgery j is assigned to the ith OR, and 0 otherwise. Thus, we assume that there are N ORs available. Accordingly, we define θ_i for every $i = 1, \ldots, N$, as a binary variable equal to 1 if the ith OR is opened, and 0 otherwise. Furthermore, for every opened OR we incur fixed costs c_f and for every hour of overtime exceeding a regular workday of T hours we incur variable costs c_v per OR. Let ζ represent the random vector of surgery durations and y_i the hours of overtime in the ith OR. Then, in case the surgery durations ζ would be deterministic, the surgery block allocation problem reads

$$\min_{\theta, \pi, y} \quad \sum_{i=1}^{N} c_f \theta_i + \sum_{i=1}^{N} c_v y_i$$

$$\text{s.t.} \quad \sum_{i=1}^{N} \pi_{ij} = 1, \qquad j = 1, \ldots, N, \tag{26}$$

$$\pi_{ij} \le \theta_i, \qquad i, j = 1, \ldots, N, \tag{27}$$

$$y_i \ge \sum_{j=1}^{N} \zeta_j \pi_{ij} - T\theta_i, \qquad i = 1, \ldots, N, \tag{28}$$

$$\theta_i \in \{0, 1\}, \quad \pi_{ij} \in \{0, 1\}, \quad y_i \in \mathbb{Z}_+, \qquad i, j = 1, \ldots, N. \tag{29}$$

Constraint (26) means that every surgery j is assigned to exactly one OR, constraint (27) models that surgery j can only be assigned to the ith OR if it is opened, and constraint (28) defines y_i as the hours of overtime for the ith OR. Notice that we assume y_i to be integer, meaning that we have to pay overtime in full hours even if it actually was a few minutes. This is one of the small differences compared to the model of Denton et al. (2010).

We let X denote the set of feasible first-stage decisions $x = (\theta, \pi)$ satisfying (26), (27), and (29). In addition, we assume that X includes several symmetry breaking constraints introduced in Denton et al. (2010). For example, we assume without loss of generality that $\theta_1 \ge \cdots \ge \theta_N$.

Similar as Denton et al. (2010) we assume that the random surgery durations ζ are unknown when the surgery-to-OR assignment has to be made, and that we know the probability distribution of ζ. Contrary to this reference, however, we assume that there is also uncertainty in the regular work day duration T_i, denoted $T_i(\omega_i)$ of the ith OR. This duration may be interpreted as the effective time spent on performing surgeries and may be smaller (or larger) than the targeted 8 h due to inefficiency (or efficiency) of the OR staff. We assume that the distribution $\mathbb{P}_z$ of the random vector $z = \omega$ is unknown and belongs to a (μ, d, β) ambiguity set $\mathcal{P}_z$ as defined in (3). The objective is to find a surgery-to-OR assignment, i.e., to determine $x = (\theta, \pi) \in X$, that minimizes the worst-case expected total costs:

$$\eta^* := \inf_{x \in X} \sup_{\mathbb{P}_z \in \mathcal{P}_z} \mathbb{E}_{\mathbb{P}_z}\left[\sum_{i=1}^{N} c_f \theta_i + Q(x, z)\right], \tag{30}$$

where

$$Q(x, z) := \mathbb{E}_{\mathbb{P}_\zeta}\left[\inf_{y \in \mathbb{Z}_+^N}\left\{\sum_{i=1}^{N} c_v y_i : \; y_i \ge \sum_{j=1}^{N} \zeta_j \pi_{ij} - T_i(\omega_i)\theta_i, \;\; i = 1, \ldots, N\right\}\right]. \tag{31}$$

4.1.1. Convex approximation for the surgery allocation problem

We can obtain an exact expression for the expected value function $Q(x, z)$ given by

$$Q(x, z) = \mathbb{E}_{\mathbb{P}_\zeta}\left[\sum_{i=1}^{N} c_v \left\lceil \sum_{j=1}^{N} \zeta_j \pi_{ij} - T_i(\omega_i)\theta_i \right\rceil^+\right].$$

Here, the round-up operator ensures that overtime wages are paid in full hours. Observe that $Q(x, z)$ is the same as the simple integer expected value function in (20) of Section 3.2.1 with $m := N$, $q_i(\xi) := c_v$, $\zeta_i := \sum_{j=1}^{N} \zeta_j \pi_{ij}$, and $x := \theta$. In fact, the only difference between the surgery allocation problem and the simple integer recourse model of Section 3.2.1 is that in the latter model we assume that the distributions of right-hand side random vector $h(\zeta)$ were known and the technology matrix $T(\omega)$ ambiguous, whereas in this problem the right-hand side random vector equals zero and the distributions of the technology matrix corresponding to the first-stage variables π and θ, are partly known and partly ambiguous. Nevertheless, we can use a similar reasoning as in Section 3.2.2 for the shifted LP-relaxation approximation to obtain the convex approximation

$$\hat{Q}(x, z) = \mathbb{E}_{\mathbb{P}_\zeta}\left[\sum_{i=1}^{N} c_v \left(\sum_{j=1}^{N} \zeta_j \pi_{ij} - (T_i(\omega_i) - 1/2)\theta_i\right)^+\right]. \tag{32}$$

Here, we simultaneously relax the integrality of the overtime hours and subtract half an hour from the work day duration (if the ith OR is opened). Again, the rationale of doing so is that on average we have to pay approximately half an hour of additional overtime if overtime is paid in full hours.

The convex approximating model with $Q(x, z)$ replaced by $\hat{Q}(x,z)$ equals

$$\inf_{x\in X}\left\{\sum_{i=1}^{N} c_f\theta_i + \mathbb{E}_{\mathbb{P}_{\bar{z}}}[\hat{Q}(x,\bar{z})]\right\}$$
$$= \inf_{x\in X}\left\{\sum_{i=1}^{N} c_f\theta_i + \mathbb{E}_{\mathbb{P}_{\bar{z}}}\left[\mathbb{E}_{\mathbb{P}_{\zeta}}\left[\sum_{i=1}^{N} c_v\left(\sum_{j=1}^{N}\zeta_j\pi_{ij} - (T_i(\omega_i) - 1/2)\theta_i\right)^{+}\right]\right]\right\}. \quad (33)$$

This model can be solved, e.g., using SAA yielding an approximating surgery-to-OR assignment $\hat{x} = (\hat{\theta}, \hat{\pi})$.

4.1.2. Error bounds for the convex approximation of the surgery allocation problem

In this section we apply the error bounds of Section 3 to the approximating surgery-to-OR allocation $\hat{x} = (\hat{\theta}, \hat{\pi})$. We will use the resulting error bounds in the numerical experiments of Section 4.2.

We derive three error bounds. The first is an upper bound on $\|Q - \hat{Q}\|_\infty$ which can be used in combination with Theorem 1. To derive this bound, we fix x and z, and bound $|Q(x,z) - \hat{Q}(x,z)|$ using Proposition 4. Interestingly, the resulting bound depends significantly on the surgery-to-OR assignment x. For example, if every surgery is carried out in a separate OR then the bound reduces to

$$|Q(x,z) - \hat{Q}(x,z)| \le \frac{1}{2}\sum_{j=1}^{N} c_v h(|\Delta|f_j), \quad (34)$$

where f_j is the marginal density of the random surgery duration ζ_j. In contrast, if all surgeries are carried out in a single OR, then the error bound reduces to

$$|Q(x,z) - \hat{Q}(x,z)| \le \frac{1}{2}c_v h(|\Delta|\bar{g}), \quad (35)$$

where $\bar{g}$ is the probability density function of *the sum* of all surgery durations. It turns out that the bound in (34) is actually the largest bound over all surgery-to-OR allocations x.

Lemma 2. *Let $Q(x, z)$ denote the expected value function of the surgery allocation problem defined in (31), and $\hat{Q}(x,z)$ its convex approximation defined in (32). Then,*

$$\|Q - \hat{Q}\|_\infty \le \frac{1}{2}\sum_{j=1}^{N} c_v h(|\Delta|f_j),$$

where f_j is the marginal density function of the random surgery duration ζ_j for $j = 1, \ldots, N$.

The bound in (34), and thus in Lemma 2, is much larger than the bound in (35). In fact, in the numerical experiments in Section 4.2 the error bound of Lemma 2 turns out to be too large for practical purposes. However, the actual error will only be so large if either the optimal or approximating surgery-to-OR allocation, x^* or $\hat{x}$, respectively, is to open *all* ORs. In practice we do not expect such an extreme surgery-to-OR allocation to be optimal. However, since we cannot verify this because we cannot compute the optimal surgery-to-OR allocation x^*, we will also use Corollary 1 to compute an upper bound on the optimality gap $G(\hat{x}) - \eta^*$ using

$$G(\hat{x}) - \eta^* \le \hat{G}(\hat{x}) - \tilde{\eta} + \sup_{\mathbb{P}_z\in\mathcal{P}_z}\mathbb{E}_{\mathbb{P}_z}\left[Q(\hat{x},z) - \hat{Q}(\hat{x},z)\right].$$

To bound the first term $\hat{G}(\hat{x}) - \tilde{\eta}$ on the right-hand side, we use the MRP. For the second term, we use Proposition 5 to derive tighter error bounds than in Lemma 2 since we only need to bound the difference between $Q(\hat{x},z) - \hat{Q}(\hat{x},z)$ for a *fixed* surgery-to-OR allocation $\hat{x}$.

Lemma 3. *Let $Q(x, z)$ denote the expected value function of the surgery allocation problem defined in (31), and $\hat{Q}(x,z)$ its convex approximation defined in (32). Then, for a fixed surgery-to-OR allocation $\hat{x}$, we have*

$$\sup_{\mathbb{P}_z\in\mathcal{P}_z}\mathbb{E}_{\mathbb{P}_z}\left[Q(\hat{x},z) - \hat{Q}(\hat{x},z)\right] \le \frac{1}{2}\sup_{\mathbb{P}_z\in\mathcal{P}_z}\mathbb{E}_{\mathbb{P}_z}\left[\sum_{i=1}^{N} c_v h(2|\Delta|g_i([T_i(\omega_i) - 1/2, +\infty)))\hat{\theta}_i\right], \quad (36)$$

where g_i is the marginal density function of the total surgery duration $\sum_{j=1}^{N}\hat{\pi}_{ij}\zeta_j$ in the ith OR.

The right-hand side of (36) may be much smaller than the error bound in Lemma 2. For one, since $\hat{\theta}_i$ may be zero for many ORs. In addition, since g_i is the probability density function of the sum of several independent random variables ζ_j, and its total variation is decreasing in the number of surgeries in the ith OR. The final reason why the error bound in Lemma 3 is tighter than that of Lemma 2 is that it only considers the total variation of g_i on the interval $[T_i(\omega_i) - 1/2, +\infty)$. The intuition is that if the total surgery duration in the ith OR does not exceed $T_i(\omega_i) - 1/2$, then both the original and approximating model have zero overtime costs in this OR.

A difficulty of the error bound in (36) of Lemma 3 is that we have to take the supremum over all probability distributions $\mathbb{P}_z \in \mathcal{P}_z$. To avoid this difficulty we can replace $T_i(\omega_i)$ by $\inf_{\omega_i\in[a_i,b_i]} T_i(\omega_i)$ for $i = 1, \ldots, N$. For notational convenience we assume that $T_i(\omega_i) = \omega_i$ so that the infimum is attained at a_i. An alternative error bound to Lemma 3 is then

$$\sup_{\mathbb{P}_z\in\mathcal{P}_z}\mathbb{E}_{\mathbb{P}_z}\left[Q(\hat{x},z) - \hat{Q}(\hat{x},z)\right] \le \frac{1}{2}\sum_{i=1}^{N} c_v h(2|\Delta|g_i([T_i(a_i) - 1/2, +\infty)))\hat{\theta}_i. \quad (37)$$

Interestingly, we may obtain a tighter bound if the error bound in (36) of Lemma 3 is convex in ω_i for $\omega_i \in [a_i, b_i]$ since it allows us to apply the result of Ben-Tal and Hochman (1972) in a surprising way. Indeed, if the bound is convex in $z_i = \omega_i \in [a_i, b_i]$, for all $i = 1, \ldots, N$, then

$$\sup_{\mathbb{P}_z\in\mathcal{P}_z}\mathbb{E}_{\mathbb{P}_z}\left[Q(\hat{x},z) - \hat{Q}(\hat{x},z)\right] \le \frac{1}{2}\mathbb{E}_{\mathbb{P}_{\bar{z}}}\left[\sum_{i=1}^{N} c_v h(2|\Delta|g_i([T_i(\bar{\omega}_i) - 1/2, +\infty)))\hat{\theta}_i\right]. \quad (38)$$

Of course, the bound $h(2|\Delta|g_i([T_i(\omega_i) - 1/2, +\infty))$ is in general not convex in ω_i, but it may be in special cases. Notice, for example, that h is linear on [0,4] so that the bound is convex if $|\Delta|g_i([t - 1/2, +\infty))$ is convex in $t \in [a_i, b_i]$ and this total variation is small enough. In our numerical experiments, these requirements might be satisfied since g_i is the pdf of the sum of several independent lognormal random variables. By the Central Limit Theorem this is asymptotically a normal pdf which has a convex decreasing right tail. Since this argument only holds asymptotically, we will check numerically for every opened OR i whether convexity holds in the numerical experiments of Section 4.2. If not, then we will replace $\bar{\omega}_i$ by a_i in the error bound of (38).

4.2. Numerical experiments

We carry out numerical experiments on problem instances of similar size as in Denton et al. (2010), i.e., with $N = 10$ and

Table 1
Means and standard deviations (in minutes) of the lognormal surgery durations (Gul et al., 2011).

Surgical group	μ	σ
Oral Maxillofacial procedure	36.00	33.88
Pain Medicine	20.93	15.08
Ophthalmology	41.63	16.43
Urology	138.16	56.77

$N = 15$. In all experiments we assume that $c_f = 1$ and $c_v = 2$ or $c_v = 0.5$, similar as in Denton et al. (2010), who provide estimates of surgery duration distributions from Gul, Denton, Fowler, and Huschka (2011). In this reference, estimates of surgery duration distributions for several types of surgeries are given. Table 1 shows the parameters of the four types of surgeries that we consider in our experiments; all surgery durations are lognormally distributed. Moreover, we set $T_i(\omega_i) = \omega_i$ where ω_i is contained in a (μ, d, β) ambiguity set with $a_i = 7$, $b_i = 9$, $\mu_i = 8$, $d_i = 0.5$, and $\beta_i = 0.5$. This means that the regular work day duration $T_i(\omega_i)$ will be between 7 and 9 h.

For all four combinations of N and c_v we generate 10 problem instances by randomly selecting N surgery types from Table 1. For each problem instance, we compute six different approximating surgery-to-OR allocations. The first three are obtained by solving the large-scale deterministic equivalent formulation (LSDE) of an SAA of the convex approximating model in (11) with $\hat{v}(x,z) := \hat{Q}(x,z)$ as defined in (32) for sample sizes $\hat{N}_s = 10, 100$, and 1000, respectively. The next three are obtained by solving the LSDE of an SAA of the approximating model in (13) with $v(x,z) := Q(x,z)$ as defined in (31) for sample sizes $\tilde{N}_s = 10, 100$, and 1000. In fact, for each value of $\hat{N}_s$ and $\tilde{N}_s$ we repeatedly solve an SAA model $N_{rep} = 10$ times, yielding N_{rep} approximating solutions and we take the best among these solutions using an out-of-sample test with sample size 10,000. Next, we apply the MRP with $N_{MRP} = 30$ replications using a sample size $N_s = 1000$ to obtain an upper bound on $\hat{G}(\hat{x}) - \tilde{\eta}$ that holds with 95% confidence.

The experiments were carried out on 15 Intel Xeon 2.5 GHz cores of the Peregrine HPC cluster of the University of Groningen. Multiple replications of the SAAs of the approximating models and the MRP were carried out in parallel.

For all six approximating solutions, we report the following performance criteria:

- The fixed costs (FC) of an approximating solution, i.e., how many ORs are opened,
- An expected lower bound $\mathbb{E}[\tilde{\eta}_{N_s}]$, obtained by the MRP, on $\tilde{\eta}$ and thus on η^*,
- An upper bound on the objective value $G(\hat{x})$ of $\hat{x}$ that holds with 95% confidence,
- A 95% upper bound on the relative optimality gap $\frac{G(\hat{x})-\eta^*}{\eta^*} \times$ 100%,
- The contribution to this 95% upper bound by applying the MRP to $\hat{G}(\hat{x}) - \tilde{\eta}$,
- The contribution to this 95% upper bound by the total variation bound in (38),
- The error bound (EB-2) of (37) relative to $\mathbb{E}[\tilde{\eta}_{N_s}]$,
- The error bound (EB-3) of Lemma 2 relative to $\mathbb{E}[\tilde{\eta}_{N_s}]$,
- The average run time (RT) of a single SAA run over all N_{rep} runs.

We only report results for $N = 15$ in Tables 2 and 3, since results for $N = 10$ are similar. From these tables we observe that on average we open between 2 and 3 ORs. Moreover, the 95% upper bounds on the objective values $G(\hat{x})$ of the approximating surgery-to-OR allocations are close to the expected lower bound $\mathbb{E}[\tilde{\eta}_{N_s}]$ on the optimal objective value η^*. This indicates that $\hat{Q}(x,z)$ is a good convex approximation of $Q(x, z)$, and thus both the surgery-to-OR allocations $\hat{x}$ and $\tilde{x}$, obtained by replacing $Q(x, z)$ by $\hat{Q}(x,z)$ and by assuming that $\mathbb{P}_{\tilde{z}}$ is the worst-case distribution, respectively, are close to optimal. Indeed, the 95% upper bounds on the optimality gap are surprisingly small, i.e., between 2% and 3% for $c_v = 2$ and around 1% for $c_v = 0.5$, given that these ambiguous mixed-integer recourse models are extremely hard to solve and we are not able to calculate the exact optimal objective values. In fact, it is not unlikely that the actual optimality gaps of the approximating solutions are even smaller than the values presented in Tables 2 and 3.

The difference in solution quality over the six approximating surgery-to-OR allocations is very small. As expected, the solutions obtained by using a smaller sample size are slightly worse. This is typically not because we open a different number of ORs in these solutions, but because we divide the surgeries over the same number of ORs in a slightly worse way. There is however a large difference in the running times required to obtain the six approximating surgery-to-OR allocations. For small sample sizes of $\hat{N}_s$ and $\tilde{N}_s$ equal to 10 both approximations models run within seconds, whereas solving the second approximating model for $\tilde{N}_s = 1000$ may take more than half an hour. The first approximating model, on the other hand, only requires 3 min of computation time. This illustrates the difference in nature between the two approximating models since the second is a standard mixed-integer recourse model with integer second-stage variables whereas the first approximating model is a standard mixed-integer recourse model with *continuous* second-stage variables. Given that both approximating models yield very similar solutions we prefer the first approximating model for this application.

The most computational effort for obtaining Tables 2 and 3 does not go in obtaining the approximating surgery-to-OR allocations but in assessing their quality. This is because to apply the MRP we effectively have to solve the second approximating model $N_{rep} =$ 30 times with a sample size of $N_s = 1000$, requiring on average more than half an hour per replication. The contribution of the MRP, applied to $\hat{G}(\hat{x}) - \tilde{\eta}$, on the 95% upper bound on the optimality gap $\frac{G(\hat{x})-\eta^*}{\eta^*} \times 100\%$ is also given in Tables 2 and 3, together with the contribution of the total variation error bound on $\sup_{\mathbb{P}_z \in \mathcal{P}_z} \mathbb{E}_{\mathbb{P}_z}[Q(\hat{x},z) - \hat{Q}(\hat{x},z)]$. Here, we have used the error bound in (38) where we replace $\bar{\omega}_i$ with a_i if the bound in Lemma 3 is not convex in ω_i on $[a_i, b_i]$. In Tables 2 and 3 we also show what the contribution to the optimality gap would have been if we would have used alternative total variation error bounds. Here, EB-2 refers to the bound in (37), where $\bar{\omega}_i$ is always replaced by a_i, and EB-3 to the bound on $\|Q - \hat{Q}\|_\infty$ from Lemma 2. We observe that EB-2 is only slightly larger than the tightest bound, whereas EB-3 is much larger. This is because the bound on $\|Q - \hat{Q}\|_\infty$ from Lemma 2 takes into account the extreme surgery-to-OR allocation in which all 15 ORs are opened and every surgery is carried out in a separate OR. However, from Tables 2 and 3 we conclude that for these experiments the optimal number of ORs to open seems to be at most three.

4.3. Out of sample tests

In this section we compare our approximation with several alternative solution methods using out of sample tests. Since we do not compute optimality gaps as in Section 4.2, we are able to solve problems with up to $N = 30$ surgeries.

We compare four different surgery-to-OR allocations. The first is denoted WC and is the same as in Section 4.2, obtained by solving the LSDE of an SAA of the convex approximating model in (11) with $\hat{v}(x,z) := \hat{Q}(x,z)$ as defined in (32) for sample size $\hat{N}_s = 100$. We use $\hat{N}_s = 100$ since similar results were obtained for $\hat{N}_s = 100$ as for $\hat{N}_s = 1000$, but the computation times with $\hat{N}_s =$

Table 2
Average results for the surgery allocations problem over 10 random problem instances with $N = 15$ surgeries and $c_v = 2$. The surgery types for each problem instance are randomly selected from Table 1.

Sample size	FC	$\mathbb{E}[\tilde{\eta}_{N_s}]$	95% UB $G(\hat{x})$	95% Opt. gap	MRP contr.	EB contr.	EB-2	EB-3	RT (in sec)
$\hat{N}_s = 10$	2.6	3.09	3.19	2.9%	1.2%	1.7%	3.3%	217%	0.6
$\hat{N}_s = 100$	2.6	3.09	3.17	2.5%	0.9%	1.7%	3.2%	217%	7.1
$\hat{N}_s = 1000$	2.7	3.09	3.15	2.0%	0.5%	1.5%	2.9%	217%	183
$\tilde{N}_s = 10$	2.6	3.09	3.21	3.7%	2.0%	1.7%	3.2%	217%	0.6
$\tilde{N}_s = 100$	2.6	3.09	3.17	2.5%	0.8%	1.7%	3.1%	217%	51
$\tilde{N}_s = 1000$	2.6	3.09	3.17	2.6%	1.0%	1.6%	3.2%	217%	1831

Table 3
Average results for the surgery allocations problem over 10 random problem instances with $N = 15$ surgeries and $c_v = 0.5$. The surgery types for each problem instance are randomly selected from Table 1.

Sample size	FC	$\mathbb{E}[\tilde{\eta}_{N_s}]$	95% UB $G(\hat{x})$	95% Opt. gap	MRP contr.	EB contr.	EB-2	EB-3	RT (in sec)
$\hat{N}_s = 10$	2.2	2.57	2.60	1.3%	0.5%	0.8%	1.1%	64%	0.7
$\hat{N}_s = 100$	2.2	2.57	2.60	1.1%	0.3%	0.8%	1.1%	64%	6
$\hat{N}_s = 1000$	2.2	2.57	2.60	1.1%	0.3%	0.8%	1.1%	64%	172
$\tilde{N}_s = 10$	2.2	2.57	2.60	1.3%	0.5%	0.8%	1.1%	64%	1.0
$\tilde{N}_s = 100$	2.2	2.57	2.60	1.1%	0.3%	0.8%	1.1%	64%	59
$\tilde{N}_s = 1000$	2.2	2.57	2.59	1.0%	0.2%	0.8%	1.1%	64%	2210

100 are significantly smaller. For the same reason, we use the first type of approximation from Section 4.2 and not the second.

Secondly, we also compute the best-case solution, referred to as BC, by solving the LSDE of an SAA with sample size $\hat{N}_s = 100$ of the convex approximating model in (16) with $\hat{v}(x, z) := \hat{Q}(x, z)$ as defined in (32).

Third, we incorrectly assume that

$$\mathcal{P}_z = \left\{ \mathbb{P}_z : \text{supp}(z_i) \subseteq [a_i, b_i], \quad \mathbb{E}_{\mathbb{P}_z}[z_i] = \mu_i, \quad z_i \perp z_j, \; i \neq j \right\}.$$

That is, only the mean and support of each random parameter is known. This is the same ambiguity set as in Xie and Ahmed (2017) for ambiguous simple integer recourse models with fixed technology matrix. We use this ambiguity set in combination with the convex approximation $\hat{v}(x, z)$ so that the worst-case distribution can be computed exactly, conform Edmundson (1956) and Madansky (1959). Based on these references, we refer to this solution as EM.

Finally, we also consider the LPT heuristic from Denton et al. (2010) that iteratively assigns the surgery with the longest mean duration to the OR with the current lowest total mean surgery duration. Since the number of ORs is prespecified in this heuristic, we carry out the heuristic for all possible values for the number of opened ORs, and we select the value leading to lowest costs under the assumption that the surgery durations and work day duration attain their mean values.

The remaining parameters are similar as in Section 4.2. We use $c_v = 0.5$ and $c_v = 2$, randomly select $N = 30$ surgery duration distributions from Table 1, and let $a_i = 6, b_i = 10, \mu_i = 8, d_i = 1$, and $\beta_i = 0.5$. In our out of sample test we use $N_{oos} = 100,000$ scenarios from the worst-case (WC) distribution $\mathbb{P}_{\bar{z}}$, the uniform (U) distribution on $[a, b]$, and the best-case (BC) distribution $\mathbb{P}_{\underline{z}}$. We expect the WC and BC solutions to perform best under the distributions $\mathbb{P}_{\bar{z}}$ and $\mathbb{P}_{\underline{z}}$, respectively. However, this is not necessary true, since the WC and BC solutions are computed using a convex approximation for the expected overtime costs, whereas in this out of sample test we report the real expected overtime costs. Moreover, for the original integer problem, $\mathbb{P}_{\bar{z}}$ and $\mathbb{P}_{\underline{z}}$ are not necessarily the worst-case and best-case distribution.

For all four approximating solutions WC, BC, EM, and LPT, we report the following performance criteria:

- The fixed costs (FC) of an approximating solution, i.e., how many ORs are opened,
- The expected overtime costs (EOC) of an approximating solution under the BC, U, and WC distribution, respectively,
- The expected total costs (ETC), similar as for the EOC,
- The maximum total costs (MTC), similar as for the EOC and ETC,
- The average run time (RT) of a solution method.

Table 4 shows the results of our out of sample test. Notice that we do not report fixed costs (FC) for every out of sample distribution. This is because the fixed costs are directly determined by the number of opened ORs, and do not depend on the expected overtime costs determined by the out of sample distribution. Moreover, observe that the differences in fixed costs, and also in EOC, ETC, and MTC, between the four solution methods are larger for $c_v = 2$ than for $c_v = 0.5$. This makes sense intuitively, since the solution methods deal with expected overtime costs in different ways, e.g., considering worst-case or best-case expected overtime costs, and thus the differences between these methods are larger if these costs are relatively larger, i.e., if $c_v = 2$.

The conclusions, however, are similar for $c_v = 2$ and $c_v = 0.5$: our solution method WC outperforms EM and LPT in terms of expected total costs (ETC). As can be observed from the average number of ORs that are opened, the LPT heuristic is too optimistic and opens too few ORs leading to larger expected overtime costs (EOC), whereas the EM method is too conservative and opens too many ORs leading to larger fixed costs (FC). The LPT heuristic does so because it ignores the uncertainty, and thus variability, in both the surgery durations and the work day durations. The EM method, on the other hand, ignores the restriction on the MAD of the random parameters, and thus has a larger ambiguity set to take into account. For this reason, it is not surprising that the EM method generally performs best in terms of maximum total costs (MTC). Comparing the EM method and the WC method, we observe that for $c_v = 2$ there is considerable value in knowing the MAD.

The differences in expected total costs (ETC) between the worst-case (WC) and best-case (BC) approximating solutions are small. As expected, the BC solution is more optimistic and typically opens fewer ORs, in particular for $c_v = 2$. In this way, its fixed costs (FC) are smaller but its expected overtime costs (EOC) are larger than for the WC solution, and these effects approximately cancel out. Interestingly, the WC solution performs best for $c_v = 2$, also under the BC distribution. This is possible since both WC and BC are *approximate* solutions and $\mathbb{P}_{\bar{z}}$ and $\mathbb{P}_{\underline{z}}$ are not necessarily the exact worst-case and best-case distributions, respectively. The same applies to the BC solution under the WC distribution for $c_v = 0.5$.

Table 4
Out of sample test for the surgery allocation problem with $N = 30$ surgeries and both $c_v = 0.5$ and $c_v = 2$. Averages over 10 random problem instances are reported.

			Out of sample distribution									
			BC			U			WC			
Method	c_v	FC	EOC	ETC	MTC	EOC	ETC	MTC	EOC	ETC	MTC	RT
WC	0.5	4.5	0.874	5.374	15.10	0.943	5.443	15.40	1.080	5.580	16.25	2160
BC	0.5	4.5	0.873	5.373	15.15	0.940	5.440	15.35	1.077	5.577	15.85	2860
EM	0.5	4.6	0.796	5.396	14.85	0.861	5.461	14.50	0.996	5.596	16.10	688
LPT	0.5	4.2	1.246	5.446	15.70	1.312	5.512	15.60	1.449	5.649	16.15	0.002
WC	2	5.9	0.891	6.791	41.46	1.031	6.931	45.46	1.328	7.228	43.26	1657
BC	2	5.5	1.296	6.796	44.06	1.475	6.975	44.26	1.844	7.344	43.06	1933
EM	2	6.3	0.650	6.950	41.66	0.762	7.062	44.26	0.997	7.297	41.46	1954
LPT	2	4.4	3.965	8.365	48.56	4.236	8.636	51.15	4.788	9.188	49.36	0.005

5. Conclusion

We have considered two-stage ambiguous recourse models and we have shown that under mean-MAD information, continuous ambiguous recourse models admit a closed-form reformulation as a standard recourse model with a worst- or best-case distributions consisting of 2 or 3 points per random parameter in the model. These worst- and best-case distributions are the same for every first-stage decision. The mean-MAD ambiguity set that we use in this paper also turns out to be particularly suitable for ambiguous recourse models involving integer decisions variables. For such problems, we develop a general approximation framework and derive corresponding error bounds. We apply this approximation framework to mixed-ambiguous mixed-integer recourse models in which some of the distributions of the random parameters are known and others are ambiguous. We illustrate the developed theory by applying it to a surgery block allocation problem. The numerical experiments show that good approximations for ambiguous mixed-integer recourse models exist for which we can obtain surprisingly good performance guarantees.

Acknowledgments

We would like to thank three anonymous referees for their constructive comments. Moreover, we would like to thank Ruben van Beesten (University of Groningen) for the idea behind Proposition 5, and we would like to thank the Center for Information Technology of the University of Groningen for their support and for providing access to the Peregrine high performance computing cluster. The first author's Ph.D. position was financed by NWO Research Talent grant 406-12-101.

Supplementary material

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.ejor.2018.10.008.

References

Ahmed, S., Tawarmalani, M., & Sahinidis, N. V. (2004). A finite branch-and-bound algorithm for two-stage stochastic integer programs. *Mathematical Programming, 100*, 355–377.

Bayraksan, G., & Morton, D. P. (2006). Assessing solution quality in stochastic programs. *Mathematical Programming, 108*, 495–514.

Bayraksan, G., & Morton, D. P. (2009). Assessing solution quality in stochastic programs via sampling. *INFORMS Tutorials in Operations Research.*, 102–122.

Ben-Tal, A., & Hochman, E. (1972). More bounds on the expectation of a convex function of a random variable. *Journal of Applied Probability, 9*, 803–812.

Ben-Tal, A., & Hochman, E. (1976). Stochastic programs with incomplete information. *Operations Research, 24*(2), 336–347.

Bertsimas, D., & Georghiou, A. (2015). Design of near optimal decision rules in multistage adaptive mixed-integer optimization. *Operations Research, 63*(3), 610–627.

Birge, J. R., & Louveaux, F. (1997). *Introduction to stochastic programming*. New York: Springer.

Carœ, C. C., & Schultz, R. (1999). Dual decomposition in stochastic integer programming. *Operations Research Letters, 24*, 37–45.

Delage, E., Arroyo, S., & Ye, Y. (2015). The value of distribution information in distributionally robust optimization. http://web.hec.ca/pages/erick.delage/ValueDistributionInfo.pdf.

Delage, E., & Ye, Y. (2010). Distributionally robust optimization under moment uncertainty with application to data-driven problems. *Operations Research, 58*(3), 595–612.

Denton, B. T., Miller, A. J., Balasubramanian, H. J., & Huschka, T. R. (2010). Optimal allocation of surgery blocks to operating rooms under uncertainty. *Operations Research, 58*, 802–816.

Dupačová, J., Gröwe-Kuska, N., & Römisch, W. (2003). Scenario reduction in stochastic programming: An approach using probability metrics. *Mathematical Programming, 95*, 493–511.

Dyer, M., & Stougie, L. (2006). Computational complexity of stochastic programming problems. *Mathematical Programming, 106*, 423–432.

Edmundson, H. P. (1956). Bounds on the expectation of a convex function of a random variable. *Technical Report Paper 982*.

Fiacco, A. V., & Kyparisis, J. (1986). Convexity and concavity properties of the optimal value function in parametric nonlinear programming. *Journal of Optimization Theory and Applications, 48*(1), 95–126.

Gade, D., Sen, S., & Küçükyavuz, S. (2014). Decomposition algorithms with parametric Gomory cuts for two-stage stochastic integer programs. *Mathematical Programming, 144*, 39–64.

Gul, S., Denton, B. T., Fowler, J. W., & Huschka, T. (2011). Bi-criteria scheduling of surgical services for an outpatient procedure center. *Production and Operations Management, 20*, 406–417.

Hanasusanto, G. A., Kuhn, D., & Wiesemann, W. (2015). *K*-adaptability in two-stage robust binary programming. *Operations Research, 63*(4), 877–891.

Hanasusanto, G. A., Kuhn, D., & Wiesemann, W. (2016). *K*-adaptability in two-stage distributionally robust binary programming. *Operations Research Letters, 44*(1), 6–11.

Heitsch, H., & Römisch, W. (2003). Scenario reduction algorithms in stochastic programming. *Computational Optimization and Applications, 24*, 187–206.

Higle, J. L., & Sen, S. (1991). Stochastic decomposition: An algorithm for two-stage linear programs with recourse. *Mathematics of Operations Research, 16*, 650–669.

Kemperman, J. H. B. (1968). The general moment problem, a geometric approach. *Annals of Mathematical Statistics, 39*, 93–122.

Klein Haneveld, W. K., Stougie, L., & van der Vlerk, M. H. (2006). Simple integer recourse models: convexity and convex approximations. *Mathematical Programming, 108*, 435–473.

Klein Haneveld, W. K., & van der Vlerk, M. H. (1999). Stochastic integer programming: General models and algorithms. *Annals of operations research, 85*, 39–57.

Knight F. (1921). Risk, uncertainty and profit, Sentry PressNew York.

Laporte, G., & Louveaux, F. V. (1993). The integer L-shaped method for stochastic integer programs with complete recourse. *Operations Research Letters, 13*, 133–142.

Louveaux, F. V., & van der Vlerk, M. H. (1993). Stochastic programming with simple integer recourse. *Mathematical Programming, 61*, 301–325.

Madansky, A. (1959). Bounds on the expectation of a convex function of a multivariate random variable. *The Annals of Mathematical Statistics, 30*, 743–746.

Pflug, G. C. (2001). Scenario tree generation for multiperiod financial optimization by optimal discretization. *Mathematical Programming, 89*, 251–271.

Postek, K., Ben-Tal, A., Hertog, D. d., & Melenberg, B. (2018). Robust optimization with ambiguous stochastic constraints under mean and dispersion information. *Operations Research, 66*(3), 814–833.

Postek, K., & Hertog, D. d. (2016). Multi-stage adjustable robust mixed-integer optimization via iterative splitting of the uncertainty set. *INFORMS Journal on Computing, 28*(3), 553–574.

Prékopa, A. (1995). *Stochastic programming*. Dordrecht: Kluwer Academic Publishers.

Romeijnders, W., Morton, D. P., & van der Vlerk, M. H. (2017). Assessing the quality of convex approximations for two-stage totally unimodular integer recourse models. *INFORMS Journal on Computing, 29*, 211–231.

Romeijnders, W., Schultz, R., van der Vlerk, M. H., & Haneveld, W. K. K. (2016a). A convex approximation for two-stage mixed-integer recourse models with a uniform error bound. *SIAM Journal on Optimization, 26*, 426–447.
Romeijnders, W., Stougie, L., & van der Vlerk, M. H. (2014). Approximation in two-stage stochastic integer programming. *Surveys in Operations Research and Management Science, 19*, 17–33.
Romeijnders, W., van der Vlerk, M. H., & Haneveld, W. K. K. (2015). Convex approximations of totally unimodular integer recourse models: A uniform error bound. *SIAM Journal on Optimization, 25*, 130–158.
Romeijnders, W., van der Vlerk, M. H., & Klein Haneveld, W. K. (2016b). Total variation bounds on the expectation of periodic functions with applications to recourse approximations. *Mathematical Programming, 157*, 3–46.
Römisch, W. (2003). Stability of stochastic programming problems. In A. Ruszczyński, & A. Shapiro (Eds.), *Stochastic programming: Handbooks in operations research and management science: 10* (pp. 483–554). Elsevier.
Ruszczyński, A. (1986). A regularized decomposition method for minimizing a sum of polyhedral functions. *Mathematical Programming, 35*, 309–333.
Scarf, H. E. (1958). A min-max solution of an inventory problem. In K. J. Arrow, S. Karlin, & H. E. Scarf (Eds.), *Studies in the mathematical theory of inventory and production* (pp. 201–209). Stanford University Press.
Schultz, R. (2003). Stochastic programming with integer variables. *Mathematical Programming, 97*, 285–309.
Sen, S. (2005). Algorithms for stochastic mixed-integer programming models. In K. Aardal, G. L. Nemhauser, & R. Weismantel (Eds.), *Discrete optimization: Handbooks in operations research and management science: 12* (pp. 515–558). Elsevier.
Sen, S., & Higle, J. L. (2005). The C^3 theorem and a D^2 algorithm for large scale stochastic mixed-integer programming: Set convexification. *Mathematical Programming, 104*, 1–20.
Sen, S., & Liu, Y. (2016). Mitigating uncertainty via compromise decisions in two-stage stochastic linear programming: Variance reduction. *Operations Research, 64*, 1422–1437.
Shapiro, A., & Ahmed, S. (2004). On a class of minimax stochastic programs. *SIAM Journal on Optimization, 14*(4), 1237–1249.
Shapiro, A., Dentcheva, D., & Ruszczyński, A. (2009). *Lectures on stochastic programming: Modeling and theory*. Philadelphia: SIAM.
Shapiro, A., & Kleywegt, A. (2002). Minimax analysis of stochastic problems. *Optimization Methods and Software, 17*(3), 523–542.
van Slyke, R. M., & Wets, R. (1969). *L*-shaped linear programs with applications to optimal control and stochastic programming. *SIAM Journal on Applied Mathematics, 17*, 638–663.
van der Vlerk, M. H. (2004). Convex approximations for complete integer recourse models. *Mathematical Programming, 99*, 297–310.
Wiesemann, W., Kuhn, D., & Sim, M. (2014). Distributionally robust convex optimization. *Operations Research, 62*(6), 1358–1376.
Xie, W., & Ahmed, S. (2017). Distributionally robust simple integer recourse, Optimization online.
Žáčková, J. (1966). On minimax solutions of stochastic linear programming problems. *Časopis Pro Pětováni Matematiky, 91*, 423–430.
Zverovich, V., Fábián, C. I., Ellison, E. F. D., & Mitra, G. (2012). A computational study of a solver system for processing two-stage stochastic LPs with enhanced Benders decomposition. *Mathematical Programming Computation, 4*, 211–238.