2018015-OPERA

Generalized alpha-approximations for two-stage mixed-integer recourse models

December 2018

Niels van der Laan
Ward Romeijnders

SOM is the research institute of the Faculty of Economics & Business at the University of Groningen. SOM has six programmes:
- Economics, Econometrics and Finance
- Global Economics & Management
- Innovation & Organization
- Marketing
- Operations Management & Operations Research
- Organizational Behaviour

Research Institute SOM
Faculty of Economics & Business
University of Groningen

Visiting address:
Nettelbosje 2
9747 AE Groningen
The Netherlands

Postal address:
P.O. Box 800
9700 AV   Groningen
The Netherlands

T +31 50 363 7068/3815

www.rug.nl/feb/research

Generalized alpha-approximations for two-stage mixed-integer recourse models

Niels van der Laan
University of Groningen, Faculty of Economics and Business
n.van.der.laan@rug.nl

Ward Romeijnders
University of Groningen, Faculty of Economics and Business

Generalized alpha-approximations for two-stage mixed-integer recourse models

Niels van der Laan   Ward Romeijnders

December 3, 2018

Abstract

We propose a new class of convex approximations for two-stage mixed-integer recourse models, the so-called generalized alpha-approximations. To guarantee the performance of these approximations we derive corresponding error bounds that depend on the total variations of the probability density functions of the random variables in the model. The error bounds converge to zero if these total variations converge to zero. The advantage of our convex approximations over existing ones is that they are more suitable for efficient computations. Indeed, we use these approximations to construct an algorithm that is able to solve large problem instances in reasonable time. We empirically assess our solution method on nurse scheduling test instances and randomly generated test instances, and show that our method finds near-optimal solutions if the variability of the random elements in the model is large. Moreover, our method outperforms existing methods in terms of computation time, especially for large problem instances.

1 Introduction

Consider the two-stage mixed-integer recourse model with random right-hand side

$$\eta^* := \min_x \left\{ cx + Q(x) : \quad Ax = b, \quad x \in X \subseteq \mathbb{R}_+^{n_1} \right\}, \tag{1}$$

where the recourse function Q is defined as

$$Q(x) := \mathbb{E}_\omega \left[\min_y \left\{ qy : \quad Wy = \omega - Tx, \quad y \in Y \subseteq \mathbb{R}_+^{n_2} \right\} \right], \quad x \in \mathbb{R}_+^{n_1}. \tag{2}$$

This model represents a two-stage decision problem under uncertainty. In the first stage, a decision x has to be made here-and-now, subject to deterministic constraints $Ax = b$ and random goal constraints $Tx = \omega$. Here, ω is a random vector whose probability distribution is known. In the second stage, the realization of ω becomes known and any infeasibilities with respect to $Tx = \omega$ have to be repaired. This is modelled by the second-stage problem

$$v(\omega, x) := \min_y \left\{ qy : \quad Wy = \omega - Tx, \quad y \in Y \subseteq \mathbb{R}_+^{n_2} \right\}. \tag{3}$$

The objective in this two-stage recourse model is to minimize the sum of immediate costs cx and expected second-stage costs $Q(x) = \mathbb{E}_\omega[v(\omega, x)]$, $x \in X$.

Frequently, integrality restrictions are imposed on the first- and second-stage decisions. That is, X and Y are of the form $X = \mathbb{Z}_+^{p_1} \times \mathbb{R}_+^{n_1-p_1}$ and $Y = \mathbb{Z}_+^{p_2} \times \mathbb{R}_+^{n_2-p_2}$. Such restrictions arise naturally when modelling real-life problems, for example to model on/off decisions or batch size restrictions. The resulting model is called a mixed-integer recourse (MIR) model. Such models have many practical applications in for example energy, telecommunication, production planning, and environmental control, see e.g. Wallace and Ziemba [2005] and Gassmann and Ziemba [2013].

While MIR models are highly relevant in practice, they are notoriously difficult to solve. The reason is that Q is in general non-convex if integrality restrictions are imposed on the second-stage decision variables y [Rinnooy Kan and Stougie, 1988]. Therefore, standard techniques for convex optimization cannot be used to solve these models. In contrast, if $Y = \mathbb{R}^{n_2}_+$, then Q is convex and efficient solution methods are available, most notably the L-shaped method by Van Slyke and Wets [1969] and variants thereof.

Because of the non-convexity of Q, traditional solution methods for MIR models typically combine ideas from deterministic mixed-integer programming and stochastic continuous programming, see e.g. Laporte and Louveaux [1993], Schultz et al. [1998], Carøe and Schultz [1999], Sen and Higle [2005], Ntaimo [2013], Gade et al. [2014], Zhang and Küçükyavuz [2014], Bansal et al. [2018], and the survey papers by Schultz [2003], Sen [2005], and Romeijnders et al. [2014]. In general, however, these solution methods typically have difficulties solving large problem instances in reasonable time.

We, however, will use a fundamentally different approach to deal with the non-convex recourse function Q. Instead of solving the original MIR model in (1), we solve an approximating problem where Q is replaced by a *convex approximation* $\hat{Q}$ of Q, yielding the approximating convex optimization problem

$$\hat{\eta} := \min_x \left\{ cx + \hat{Q}(x) : \quad Ax = b, \quad x \in \mathbb{R}^{n_1}_+ \right\}. \tag{4}$$

Because $\hat{Q}$ is convex, we can solve the optimization problem in (4) much faster than the original MIR model in (1) using techniques from convex optimization. Thus, we expect to solve similar-sized problems much faster than traditional solution methods, and we also expect to be able to solve larger problem instances than traditional methods. In fact, this is what we show in our numerical experiments on nurse scheduling and randomly generated problem instances.

Obviously, the optimal solution $\hat{x}$ of the approximating problem in (4) is not necessarily optimal for the original MIR model in (1). That is why we guarantee the quality of the approximating solution $\hat{x}$, by deriving an error bound on

$$||Q - \hat{Q}||_\infty := \sup_x |Q(x) - \hat{Q}(x)|.$$

This error bound directly gives us an upper bound on the optimality gap of $\hat{x}$:

$$c\hat{x} + Q(\hat{x}) - \eta^* \leq 2||Q - \hat{Q}||_\infty,$$

see Romeijnders et al. [2015].

Convex approximations and corresponding error bounds have been derived for many different classes of models. The idea to solve convex approximations $\hat{Q}$ for the non-convex mixed-integer recourse function Q dates back to Van der Vlerk [1995], who proposed to use α-approximations for the special case of simper integer recourse (SIR) models. These α-approximations are obtained by perturbing the probability distribution of the random vector ω. Klein Haneveld et al. [2006] derive an error bound for the α-approximations that depends on the total variations $|\Delta|f_i$ of the marginal density functions f_i of the random variables in the SIR model. More convex approximations and corresponding error bounds have been described for more general classes of problems than SIR models. For example, Van der Vlerk [2004] generalizes the α-approximations to integer recourse models with totally unimodular (TU) recourse matrix W and Romeijnders et al. [2016b] derive the so-called shifted LP-relaxation approximation for the same class of problems. The latter approximation is generalized to two-stage MIR models in general in Romeijnders et al. [2016a]. Corresponding error bounds for these approximations are derived in Romeijnders et al. [2015, 2016a,b], respectively.

The quality of the convex approximations for TU integer recourse models is assessed empirically in Romeijnders et al. [2017], and it turns out that the convex approximations perform well if the variability of the random parameters in the models is large enough. For general MIR models, however, the performance of the shifted LP-relaxation approximation has not been investigated,

since it is difficult to efficiently solve the approximating problem with this convex approximation, as discussed in Romeijnders et al. [2016a].

That is why we propose an alternative class of convex approximations for general two-stage MIR models that can be considered the generalization of the first α-approximations of Van der Vlerk [1995] for SIR models. Hence, we call our convex approximations generalized α-approximations.

We will show that, contrary to the shifted LP-relaxation, the generalized α-approximations, denoted by $\tilde{Q}_\alpha$, can be used to efficiently obtain approximating solutions $\tilde{x}_\alpha$. In fact, we will develop a loose Benders decomposition algorithm to solve the approximating model in (4) with $\tilde{Q}_\alpha$. In this algorithm, we do not compute $\tilde{Q}_\alpha(\bar{x})$ and subgradients of $\tilde{Q}_\alpha$ at $\bar{x}$ for current first-stage solutions $\bar{x}$ exactly, but we efficiently obtain a lower bound for $\tilde{Q}_\alpha(\bar{x})$, yielding loose optimality cuts. This does not necessarily yield the same solution as solving the approximating model in (4) exactly, but we will prove that these loose optimality cuts will be tight enough and show that a similar performance guarantee applies to the solution obtained by the loose Benders decomposition as to $\tilde{x}_\alpha$. Moreover, since we actually apply the loose Benders decomposition to a sample average approximation (SAA) of the approximating model in (4), we prove that our performance guarantee applies as the sample size $S \to \infty$.

Summarizing, our main contributions are as follows.

- We propose a new class of convex approximations, called generalized α-approximations for general two-stage MIR models. These approximations generalize the original α-approximations of Van der Vlerk [1995] for SIR models.

- We derive an error bound for the generalized α-approximations that depends on the total variations of the probability density functions of the random parameters in the model. The error bound converges to zero if these total variations converge to zero.

- We derive a loose Benders decomposition algorithm to (approximately) solve the approximating model with the generalized α-approximations. This is the first efficient algorithm for solving non-trivial convex approximations of general two-stage MIR models.

- We prove that the solution obtained by the loose Benders decomposition algorithm has a similar performance guarantee as the exact solution to the generalized α-approximations.

- We carry out numerical experiments on a nurse scheduling problem and randomly generated test instances, and show that using our loose Benders decomposition algorithm we obtain good solutions within reasonable time, also for large problem instances, in particular when the variability of the random parameters in the model is large.

The remainder of this paper is organized as follows. In Section 2, we discuss preliminaries of known convex approximations of MIR models and corresponding error bounds. In Section 3, we present the generalized α-approximations and an efficient algorithm for solving the corresponding approximating problem. Section 4 contains the proof of the error bound for the generalized α-approximations. In Section 5, we report on numerical experiments to evaluate the performance of our algorithm. Finally, Section 6 concludes.

Throughout, we make the following assumptions. Assumptions (A2)-(A4) guarantee that $Q(x)$ is finite for all $x \in X$ such that $Ax = b$.

(A1) The first-stage feasible region $\mathcal{X} = \{x \in X : Ax = b\}$ is bounded.

(A2) The recourse is relatively complete: for all $\omega \in \mathbb{R}^m$ and $x \in \mathcal{X}$, there exists a $y \in Y$ such that $Wy = \omega - Tx$, so that $v(\omega, x) < \infty$.

(A3) The recourse is sufficiently expensive: $v(\omega, x) > -\infty$ for all $\omega \in \mathbb{R}^m$ and $x \in \mathcal{X}$.

(A4) $\mathbb{E}[|\omega_i|]$ is finite for all $i = 1, \ldots, m$.

(A5) The recourse matrix W is integer.

2 Literature review on convex approximations of mixed-integer recourse functions

In this section, we review existing convex approximations of MIR models and corresponding error bounds. We distinguish two types of convex approximations, namely α-approximations and shifted-LP relaxations. The latter have been derived for general two-stage MIR models, whereas α-approximations are only known for SIR and TU integer recourse models. Table 1 provides an overview of the approximating value functions $\hat{v}(\omega, x)$ of the shifted LP-relaxation approximation and the α-approximations for the different model classes that we will discuss in this section. The generalized α-approximations that we develop in this paper can be considered as the generalization of the α-approximations to general MIR models, and as the "α-approximations equivalent" of the existing shifted LP-relaxation approximation for general MIR models.

In Sections 2.1 and 2.2, we discuss both types of approximations and their corresponding error bound for SIR models and TU integer recourse models, respectively. In these sections we also gradually introduce the notation required to state the error bound for the shifted LP-relaxation for general MIR models in Section 2.3. In this section, we also discuss results by Romeijnders et al. [2016a] regarding asymptotic periodicity results in mixed-integer linear programming, since they are used to derive the shifted-LP relaxation, and since we use them in a similar way in Section 3.1 to derive the new generalized α-approximations for general two-stage MIR models.

Table 1: Value functions $\hat{v}(\omega, x)$ of convex approximations in the literature, where $\hat{Q}(x) = \mathbb{E}_\omega[\hat{v}(\omega, x)]$.

Section	Model class	Shifted LP-relaxation	α-approximations
2.1	SIR	$(\omega + 1/2 - Tx)^+$	$(\lceil \omega - \alpha \rceil + \alpha - Tx)^+$
2.2	TU integer recourse	$\max\limits_{k=1,\ldots,K} \lambda^k(\omega + 1/2e_m - Tx)$	$\max\limits_{k=1,\ldots,K} \lambda^k(\lceil \omega - \alpha \rceil + \alpha - Tx)$
2.3	General MIR	$\max\limits_{k=1,\ldots,K} \left\{ q_{B^k}(B^k)^{-1}(\omega - Tx) + \Gamma_k \right\}$	

2.1 Convex approximations of simple integer recourse functions

Consider the one-dimensional SIR function $\mathcal{Q}$ defined as

$$\mathcal{Q}(x) := \mathbb{E}_\omega \lceil \omega - Tx \rceil^+, \quad x \in \mathbb{R}^{n_1},$$

where ω represents a random variable, where $\lceil s \rceil$ denotes the round-up function, and where $\lceil s \rceil^+ := \max\{0, \lceil s \rceil\}$, $s \in \mathbb{R}$. Note that the technology matrix T is a row vector. The SIR function is a special case of the general MIR function Q defined in (2), see Louveaux and Van der Vlerk [1993].

In the literature, two types of convex approximations of $\mathcal{Q}$ are known for this SIR function, namely the α-approximations $\tilde{\mathcal{Q}}_\alpha$, due to Van der Vlerk [1995], and the shifted LP-relaxation approximation $\hat{\mathcal{Q}}$, due to Romeijnders et al. [2016b]. They are defined in Definitions 1 below.

Definition 1. For any $\alpha \in \mathbb{R}$, we define the α-approximation $\tilde{\mathcal{Q}}_\alpha$ of the SIR function $\mathcal{Q}$ as

$$\tilde{\mathcal{Q}}_\alpha(x) := \mathbb{E}_\omega(\lceil \omega - \alpha \rceil + \alpha - Tx)^+, \quad x \in \mathbb{R}^{n_1},$$

and the *shifted LP-relaxation approximation* $\hat{\mathcal{Q}}$ as

$$\hat{\mathcal{Q}}(x) := \mathbb{E}_\omega(\omega + 1/2 - Tx)^+, \quad x \in \mathbb{R}^{n_1},$$

where $(s)^+ := \max\{s, 0\}$.

4

An interesting observation is that the α-approximations are exact on a grid of points. Indeed, it follows from the definition of $\tilde{\mathcal{Q}}_\alpha$ that if $Tx \in \alpha + \mathbb{Z}$, then $\tilde{\mathcal{Q}}_\alpha(x) = \mathcal{Q}(x)$. In contrast, the shifted LP-relaxation approximation is a good approximation of $\mathcal{Q}$ on average.

Klein Haneveld et al. [2006] and Romeijnders et al. [2016b] derive error bounds for the α-approximations and the shifted-LP relaxation approximation, respectively. That is, they derive bounds on $||\mathcal{Q} - \tilde{\mathcal{Q}}_\alpha||_\infty$ and on $||\mathcal{Q} - \hat{\mathcal{Q}}||_\infty$, respectively. These error bounds are expressed in terms of the *total variation* of the probability density function (pdf) f of the random variable ω.

Definition 2. Let $f : \mathbb{R} \to \mathbb{R}$ be a real-valued function, and let $I \subset \mathbb{R}$ be an interval. Let $\Pi(I)$ denote the set of all finite ordered sets $P = \{x_1, \ldots, x_{N+1}\}$ with $x_1 < \cdots < x_{N+1}$ in I. Then, the *total variation* of f on I, denoted $|\Delta|f(I)$, is defined as

$$|\Delta|f(I) = \sup_{P \in \Pi(I)} V_f(P),$$

where

$$V_f(P) = \sum_{i=1}^N |f(x_{i+1}) - f(x_i)|.$$

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

Theorem 1. *Consider the SIR function $\mathcal{Q}(x) = \mathbb{E}_\omega[\omega - Tx]^+$, $x \in \mathbb{R}^{n_1}$, its α-approximation $\tilde{\mathcal{Q}}_\alpha(x) = \mathbb{E}_\omega(\lceil \omega - \alpha \rceil + \alpha - Tx)^+$ for $\alpha \in \mathbb{R}$, and its shifted LP-relaxation approximation $\hat{\mathcal{Q}}(x) = \mathbb{E}_\omega(\omega + 1/2 - Tx)^+$, $x \in \mathbb{R}^{n_1}$. Then, for every continuous random variable ω with probability density function f of bounded variation, we have*

$$||\mathcal{Q} - \tilde{\mathcal{Q}}_\alpha||_\infty \le h(|\Delta|f) \text{ and } ||\mathcal{Q} - \hat{\mathcal{Q}}||_\infty \le \frac{1}{2}h(|\Delta|f),$$

where $h : [0, \infty) \mapsto \mathbb{R}$ is defined for $t \in \mathbb{R}$ with $t \ge 0$ as

$$h(t) := \begin{cases} t/8, & t \le 4, \\ 1 - 2/t, & t \ge 4. \end{cases} \tag{5}$$

Proof. See Romeijnders et al. [2016b]. $\square$

Observe that the error bound for $\hat{\mathcal{Q}}$ in Theorem 1 improves on the error bound for $\tilde{\mathcal{Q}}_\alpha$ by a factor two. Furthermore, both error bounds are decreasing in the total variation $|\Delta|f$ of the pdf f of ω. For unimodal distributions this means that the error bounds are smaller, and thus the convex approximations are better, if the variance of the random variable ω is larger. We say that the approximation becomes better as the variability in the model increases. In Example 1 below we derive expressions for the error bound for $\omega \sim \mathcal{N}(\mu, \sigma^2)$ in terms of the standard deviation σ of ω.

Example 1. Suppose that ω follows a normal distribution with mean μ and standard deviation σ. Since the corresponding pdf f is unimodal with mode μ and $f(\mu) = \frac{1}{\sqrt{2\pi}\sigma}$, we have $|\Delta|f = 2f(\mu) = \sqrt{\frac{2}{\pi}\sigma}$, and thus

$$||\mathcal{Q} - \tilde{\mathcal{Q}}_\alpha||_\infty \le h(|\Delta|f) = \begin{cases} \frac{1}{8}\sqrt{\frac{2}{\pi}\sigma}, & \sigma \ge \frac{1}{8\pi}, \\ 1 - \sqrt{2\pi}\sigma, & \sigma \le \frac{1}{8\pi}. \end{cases}$$

Because $|\Delta|f$ is decreasing in the standard deviation σ, the error bounds for $\tilde{\mathcal{Q}}_\alpha$ and $\hat{\mathcal{Q}}$ are decreasing in σ. $\diamond$

Remark 1. Van der Laan et al. [2018] improve the error bound in Theorem 1 for the shifted-LP relaxation approximation $\hat{\mathcal{Q}}$ using the total variations of higher-order derivatives of f.

2.2 Convex approximations of TU integer recourse functions

Next, we consider the pure integer recourse function

$$Q(x) := \mathbb{E}_\omega \left[\min_y \left\{ qy : \quad Wy \geq \omega - Tx, \quad y \in \mathbb{Z}_+^{n_2} \right\} \right], \quad x \in \mathbb{R}^{n_1}, \tag{6}$$

where the recourse matrix W is TU. Such TU integer recourse functions are a generalization of SIR functions, but a special case of the MIR function Q defined in (2). Both the α-approximations and the shifted LP-relaxation approximation have been generalized to TU integer recourse models. The first by Van der Vlerk [2004] and the second by Romeijnders et al. [2016b].

Definition 3. For any $\alpha \in \mathbb{R}^m$, we define the α-*approximation* $\tilde{Q}_\alpha$ of the TU integer recourse function Q as

$$\tilde{Q}_\alpha(x) := \mathbb{E}_\omega \left[\min_y \left\{ qy : \quad Wy \geq \lceil \omega - \alpha \rceil + \alpha - Tx, \quad y \in \mathbb{R}_+^{n_2} \right\} \right], \quad x \in \mathbb{R}^{n_1},$$

where $\lceil \cdot \rceil$ denotes the component-wise round-up function, and we define the *shifted LP-relaxation approximation* as

$$\hat{Q}(x) := \mathbb{E}_\omega \left[\min_y \left\{ qy : \quad Wy \geq \omega - Tx + 1/2e_m, \quad y \in \mathbb{R}_+^{n_2} \right\} \right], \quad x \in \mathbb{R}^{n_1},$$

where e_m denotes the m-dimensional all-one vector.

Both the TU integer recourse function Q and its approximations have a dual representation, since by strong LP duality,

$$\min_y \{ qy : \quad Wy \geq s \} = \max_\lambda \{ \lambda s : \quad \lambda W \leq q, \quad \lambda \in \mathbb{R}_+^{n_1} \},$$

for every $s \in \mathbb{R}^m$. Here strong LP duality holds by Assumptions (A2) and (A3). These assumptions also imply that the dual feasible region is non-empty and bounded, and thus has finitely many extreme points λ^k, $k = 1, \ldots, K$. Since at least one of the extreme points is an optimal dual solution we can rewrite the convex approximations as

$$\tilde{Q}_\alpha(x) = \mathbb{E}_\omega \left[\max_{k=1,\ldots,K} \lambda^k (\lceil \omega - \alpha \rceil + \alpha - Tx) \right], \quad x \in \mathbb{R}^{n_1}, \tag{7}$$

and

$$\hat{Q}(x) = \mathbb{E}_\omega \left[\max_{k=1,\ldots,K} \lambda^k (\omega + 1/2e_m - Tx) \right], \quad x \in \mathbb{R}^{n_1}.$$

When ω is a random variable and there are only two dual vertices $\lambda^1 = 0$ and $\lambda^2 = 1$, then the above approximations reduce to their SIR counterparts. Moreover, since

$$Q(x) = \mathbb{E}_\omega \left[\max_{k=1,\ldots,K} \lambda^k \lceil \omega - Tx \rceil \right], \quad x \in \mathbb{R}^{n_1},$$

the approximations have similar interpretations as for SIR models. In the shifted LP-relaxation approximation $\hat{Q}$, the round-up operator is removed and a value $1/2$ is added to each component of the random vector ω. The α-approximations on the other hand are exact for $x \in \mathbb{R}^{n_1}$ such that $Tx \in \alpha + \mathbb{Z}^m$.

Romeijnders et al. [2016b] derive error bounds for both the shifted LP-relaxation and the α-approximations, see Theorem 2. Similar as in the SIR case, the error bound of the shifted LP-relaxation improves the error bound of the α-approximations by a factor two. Moreover, the error bounds depend on the extreme points of the dual feasible region and the total variations of the one-dimensional conditional density functions of the random variables in the model.

6

Definition 4. For every $i = 1, \ldots, m$ and $x_{-i} \in \mathbb{R}^{m-1}$, define the i-th conditional density function $f_i(\cdot|x_{-i})$ of the m-dimensional joint pdf as

$$f_i(x_i|x_{-i}) = \frac{f(x)}{f_{-i}(x_{-i})},$$

where x_{-i} denotes x without its i-th component. Define $\mathcal{H}^m$ as the set of all m-dimensional joint pdf f such that $f_i(\cdot|x_{-i})$ is of bounded variation for all $i = 1, \ldots, m$ and $x_{-i} \in \mathbb{R}^{m-1}$.

Theorem 2. *Consider the TU integer recourse function of* (6), *the shifted LP-relaxation approximation, and the α-approximations of Definition 3. Then, for every continuous random vector ω with probability density function $f \in \mathcal{H}^m$, we have*

$$||Q - \tilde{Q}_\alpha||_\infty \leq \sum_{i=1}^{m} \lambda_i^* \mathbb{E}_{\omega_{(-i)}} \left[h\left(|\Delta| f_i(\cdot|\omega_{(-i)}) \right) \right]$$

and

$$|||Q - \hat{Q}||_\infty \leq \frac{1}{2} \sum_{i=1}^{m} \lambda_i^* \mathbb{E}_{\omega_{(-i)}} \left[h\left(|\Delta| f_i(\cdot|\omega_{(-i)}) \right) \right],$$

where $\lambda_i^* = \max_{k=1,\ldots,K} \lambda_i^k$ and h is defined in (5).

Theorem 2 applies to general pdf $f \in \mathcal{H}^m$ of the random vector ω. If the components of ω are independent, then the expressions for the error bounds simplify to

$$||Q - \tilde{Q}_\alpha||_\infty \leq \sum_{i=1}^{m} \lambda_i^* h(|\Delta| f_i), \text{ and } ||Q - \hat{Q}||_\infty \leq \frac{1}{2} \sum_{i=1}^{m} \lambda_i^* h(|\Delta| f_i),$$

where f_i denotes the marginal pdf of the i-th component of ω, $i = 1, \ldots, m$.

2.3 Convex approximations of general mixed-integer recourse functions

In this section, we consider the general two-stage MIR case. Before we describe the shifted LP-relaxation of Romeijnders et al. [2016a] in Section 2.3.2, we first discuss asymptotic periodicity properties for parametric mixed-integer linear programming problems in Section 2.3.1. These properties are not only used to derive the shifted LP-relaxation, but we also use them to derive our new generalized α-approximations in Section 3.1, and to prove a corresponding error bound in Section 4.1.

2.3.1 Asymptotic periodicity in mixed-integer linear programming

Consider the second-stage value function $v(\omega, x)$, defined in (3), with $Y = \mathbb{Z}_+^{p_2} \times \mathbb{R}_+^{n_2 - p_2}$:

$$v(\omega, x) = \min_y \left\{ qy : \quad Wy = \omega - Tx, \quad y \in \mathbb{Z}_+^{p_2} \times \mathbb{R}_+^{n_2 - p_2} \right\}. \tag{8}$$

Using LP-duality, we know that the LP-relaxation $v_{\text{LP}}(\omega, x)$ of $v(\omega, x)$ is polyhedral in the right-hand side vector $\omega - Tx$:

$$v_{\text{LP}}(\omega, x) = \max_{k=1,\ldots,K} \lambda^k(\omega - Tx), \tag{9}$$

where λ^k, $k = 1 \ldots, K$, are the extreme points of the dual feasible region $\{\lambda : \lambda W \leq q\}$. Romeijnders et al. [2016a] derive a similar characterization of $v(\omega, x)$ in terms of linear and periodic functions, see Lemma 1. The proof of this lemma is based on the Gomory relaxation of v.

We will briefly discuss the Gomory relaxation before we state Lemma 1. The Gomory relaxation is defined for any dual feasible basis matrix of the LP-relaxation v_{LP} of v. Let B denote such

a matrix and let N be such that $W \equiv \begin{pmatrix} B & N \end{pmatrix}$, meaning equality up to a permutation of the columns. Let y_B and y_N denote the second-stage variables corresponding to the columns in B and N, respectively, and q_B and q_N their corresponding cost parameters. Since B is a basis matrix, it is invertible, and thus $Wy = \omega - Tx$ is equivalent to $y_B = B^{-1}(\omega - Tx - Ny_N)$. We obtain

$$v(\omega, x) = q_B B^{-1}(\omega - Tx)$$
$$+ \min \left\{ \bar{q}_N y_N : \quad B^{-1}(\omega - Tx - Ny_N) \in \mathbb{Z}_+^{p_B} \times \mathbb{R}_+^{n_B}, \quad y_N \in \mathbb{Z}_+^{p_N} \times \mathbb{R}_+^{n_N} \right\},$$

where $\bar{q}_N := q_N - q_B B^{-1} N$ denotes the reduced costs. Note that $\bar{q}_N \geq 0$, because B is a dual feasible basis matrix. The Gomory relaxation v_B is obtained by relaxing non-negativity of y_B:

$$v_B(\omega, x) = q_B B^{-1}(\omega - Tx) + \psi_B(\omega - Tx),$$

where

$$\psi_B(s) := \min \left\{ \bar{q}_N y_N, \quad B^{-1}(s - Ny_N) \in \mathbb{Z}^{p_B} \times \mathbb{R}^{n_B}, \quad y_N \in \mathbb{Z}_+^{p_N} \times \mathbb{R}_+^{n_N} \right\}, \quad s \in \mathbb{R}^m. \quad (10)$$

Romeijnders et al. [2016a] prove properties of $v_B(\omega, x)$ and $\psi_B(s)$. In particular, they derive conditions on ω and x which guarantee that $v(\omega, x) = v_B(\omega, x)$. Under these conditions, the value function v is characterized by

$$v(\omega, x) = q_B B^{-1}(\omega - Tx) + \psi_B(\omega - Tx). \quad (11)$$

It turns out that $v(\omega, x) = v_B(\omega, x)$ if $\omega - Tx$ is in the closed convex cone $\Lambda := \{t : B^{-1}t \geq 0\}$ and the distance of $\omega - Tx$ to the boundary of Λ is sufficiently large. This latter condition is formalized in Definition 5. Moreover, Romeijnders et al. [2016a] prove that ψ_B is B-periodic, meaning that

$$\psi_B(s + Bl) = \psi_B(s)$$

for all $l \in \mathbb{Z}^m$. The notion of B-periodicity is formalized in Definition 6.

Definition 5. Let $\Lambda \subset \mathbb{R}^m$ be a closed convex cone and let $d \in \mathbb{R}$ with $d > 0$ be given. Then, we define $\Lambda(d)$ as

$$\Lambda(d) := \{s \in \Lambda : \mathcal{B}(s, d) \subset \Lambda\},$$

where $\mathcal{B}(s, d) := \{t \in \mathbb{R}^m : ||t - s||_2 \leq d\}$ is the closed ball centered at s with radius d. We can interpret $\Lambda(d)$ as the set of points in Λ with at least Euclidean distance d to the boundary of Λ.

Definition 6. Let the function $g : \mathbb{R}^m \mapsto \mathbb{R}^n$ be given and let B be an $m \times m$ matrix. Then, the function g is called B-periodic if and only if for every $x \in \mathbb{R}^m$ and $l \in \mathbb{Z}^m$

$$g(x) = g(x + Bl).$$

We are now ready to characterize $v(\omega, x)$.

Lemma 1. Consider the mixed-integer programming problem

$$v(\omega, x) := \min_y \left\{ qy : \quad Wy = \omega - Tx, \quad y \in \mathbb{Z}_+^{p_2} \times \mathbb{R}_+^{n_2 - p_2} \right\},$$

where W is an integer matrix, and $v(\omega, x)$ is finite for all $\omega \in \mathbb{R}^m$ and $x \in \mathbb{R}^n$. Then, there exist dual feasible basis matrices B^k of v_{LP}, $k = 1, \ldots, K$, closed convex polyhedral cones $\Lambda^k := \{t \in \mathbb{R}^m : (B^k)^{-1}t \geq 0\}$, distances d_k, B^k-periodic functions π^k and ψ^k, and constants w_k such that we have the following:

(i) $\cup_{k=1}^K \Lambda^k = \mathbb{R}^m$.

(ii) $(int \, \Lambda^k) \cap (int \, \Lambda^l) = \emptyset$ for every $k, l \in \{1, \ldots, K\}$ with $k \neq l$.

(iii) For all $\omega \in \mathbb{R}^m$ and $x \in \mathbb{R}_+^{n_1}$,

$$0 \le \psi_k(\omega - Tx) \le w_k, \quad k = 1, \ldots, K.$$

(iv) If $\omega - Tx \in \Lambda^k(d_k)$, then

$$y_{B^k}(\omega, x) = (B^k)^{-1}(\omega - Tx - N^k \pi^k(\omega - Tx))$$

$$y_{N^k}(\omega, x) = \pi^k(\omega - Tx)$$

is optimal for $v(\omega, x)$.

(v) If $\omega - Tx \in \Lambda^k(d_k)$, then

$$v(\omega, x) = v_{\text{LP}}(\omega, x) + \psi^k(\omega - Tx) = q_{B^k}(B^k)^{-1}(\omega - Tx) + \psi^k(\omega - Tx),$$

where $\psi^l \equiv \psi^k$ if $q_{B^k}(B^k)^{-1} = q_{B^l}(B^l)^{-1}$.

Proof. See Romeijnders et al. [2016a]. □

Lemma 1 (v) shows that if $\omega - Tx \in \Lambda^k(d_k)$ for some $k = 1, \ldots, K$, then $v(\omega, x)$ is equal to the sum of the LP-relaxation $v_{\text{LP}}(\omega, x)$ and $\psi^k(\omega - Tx)$. Hence, $\psi^k(\omega - Tx)$ can be interpreted as the additional costs resulting from the integrality restrictions on the decision variables y.

2.3.2 Convex approximation and error bound

Lemma 1 also shows why the second-stage value function is not convex in x. On regions of its domain it is the sum of a linear function $q_{B^k}(B^k)^{-1}(\omega - Tx)$ and a periodic function $\psi^k(\omega - Tx)$. Clearly the periodic part is causing v to be non-convex. That is why the shifted LP-relaxation is obtained by replacing this periodic part $\psi^k(\omega - Tx)$ by a constant Γ_k for every $k = 1, \ldots, K$, with Γ_k defined as

$$\Gamma_k := p_k^{-m} \int_0^{p_k} \cdots \int_0^{p_k} \psi^k(x) dx_1 \ldots dx_m, \tag{12}$$

where $p_k = |\det B_k|$. The K constants Γ_k can be interpreted as the averages of the periodic functions ψ^k. The shifted LP-relaxation approximation is obtained by taking the pointwise maximum over all dual feasible basis matrices B^k, $k = 1, \ldots, K$.

Definition 7. Define the *shifted LP-relaxation approximation* $\hat{Q}$ of the MIR function Q as $\hat{Q}(x) = \mathbb{E}_\omega[\hat{v}(\omega, x)]$, where

$$\hat{v}(\omega, x) := \max_{k=1,\ldots,K} \{q_{B^k}(B^k)^{-1}(\omega - Tx) + \Gamma_k\}, \tag{13}$$

where B^k, $k = 1, \ldots, K$, are the dual feasible basis matrices of Lemma 1, and Γ_k is defined in (12).

For TU integer recourse models it turns out that $\Gamma_k = 1/2 q_{B^k}(B^k)^{-1} e_m$, $k = 1, \ldots, K$, and thus by defining $\lambda^k := q_{B^k}(B^k)^{-1}$, the shifted LP-relaxation above reduces to that of Definition 3. Thus, the shifted LP-relaxation of Definition 7 can be considered as the natural generalization of those for SIR and TU integer recourse models.

Romeijnders et al. [2016a] derive a total variation error bound for the shifted LP-relaxation $\hat{Q}$ of Definition 7. We state the error bound below and we briefly sketch their line of proof.

Theorem 3. *Consider the mixed-integer recourse function Q defined as $Q(x) = \mathbb{E}_\omega[v(\omega, x)]$, where the value function v is defined in (3). Moreover, consider the shifted LP-relaxation approximation $\hat{Q}$ of Definition 7. Then, there exists a constant $C > 0$ such that for every continuous random vector ω with joint pdf $f \in \mathcal{H}^m$,*

$$\sup_x |Q(x) - \hat{Q}(x)| \le C \sum_{i=1}^m \mathbb{E}_{\omega_{-i}} \left[|\Delta| f_i(\cdot | \omega_{-i}) \right]. \tag{14}$$

Proof. See Romeijnders et al. [2016a]. □

In general, the error bound in Theorem 3 may be large, in particular since the constant $C > 0$ may be large. Nevertheless, the theorem shows that if the variability in the model is large enough, then $\hat{Q}$ is a good approximation of Q. Moreover, this holds for any two-stage mixed-integer recourse model.

The line of proof by Romeijnders et al. [2016a] for Theorem 3 is to consider the difference $v(\omega, x) - \hat{v}(\omega, x)$ for a given x. They show that there exist $\sigma_k \in \Lambda^k(d_k)$ such that

$$\omega - Tx \in \sigma_k + \Lambda^k \implies v(\omega, x) - \hat{v}(\omega, x) = \psi^k(\omega - Tx) - \Gamma_k,$$

which is zero-mean B^k-periodic in ω. Moreover, the difference $v(\omega, x) - \hat{v}(\omega, x)$ is uniformly bounded. Romeijnders et al. [2016a] derive a total variation bound on the expectation of B-periodic functions and a bound on the probability

$$\mathbb{P}\left[\omega - Tx \notin \bigcup_{k=1}^{K}(\sigma_k + \Lambda^k)\right].$$

By combining these results, they prove (14). In Section 4.1, we will also use these total variation bounds to derive an error bound for the generalized α-approximations, similar to Theorem 3.

3 Generalized α-approximations for two-stage mixed-integer recourse models

In this section, we introduce our new convex approximations for two-stage MIR models. These generalized α-approximations derived in Section 3.1, turn out to be more suitable for optimization than the shifted-LP relaxation. In fact, the remainder of this section is devoted to deriving a loose Benders decomposition algorithm that solves the convex approximation model corresponding to the generalized α-approximation to near optimality. In Sections 3.2 and 3.3, we discuss Benders decompositions, and we introduce loose optimality cuts, respectively, for the generalized α-approximations. Section 3.4 contains our loose Benders decomposition algorithm for the generalized α-approximations, called LBDA(α). We conclude this section with Proposition 1, where we state a total variation error bound on the optimality gap of the solution generated by LBDA(α). The proof of Proposition 1 is postponed to Section 4.

3.1 Generalized α-approximations

To derive the generalized α-approximations $\tilde{Q}_\alpha$, we first derive a convex approximation $\tilde{v}_\alpha$ of the second-stage value function v defined in (8). By taking expectations, we obtain the convex approximation $\tilde{Q}_\alpha(x) := \mathbb{E}_\omega[\tilde{v}_\alpha(\omega, x)]$. Similar as for the shifted LP-relaxation discussed in Section 2.3, we exploit asymptotic periodicity properties of mixed-integer linear programming problems. In particular, we use that by Lemma 1, for $\omega - Tx \in \Lambda^k(d_k)$,

$$v(\omega, x) = q_{B^k}(B^k)^{-1}(\omega - Tx) + \psi^k(\omega - Tx).$$

Instead of replacing $\psi^k(\omega - Tx)$ by its average Γ_k, as is done for the shifted LP-relaxation, we replace $\psi^k(\omega - Tx)$ by $\psi^k(\omega - \alpha)$ for some $\alpha \in \mathbb{R}^m$ to obtain the convex approximation $\tilde{v}_\alpha^k$ defined as

$$\tilde{v}_\alpha^k(\omega, x) = q_{B^k}(B^k)^{-1}(\omega - Tx) + \psi^k(\omega - \alpha).$$

Since ψ^k is B^k-periodic, the approximation $\tilde{v}_\alpha^k$ is exact for the Gomory relaxation v_{B^k} if $Tx \in \alpha + B^k \mathbb{Z}^m$. Hence, if in addition $\omega - Tx \in \Lambda^k(d_k)$, then $\tilde{v}_\alpha^k$ is exact for v. Moreover, the difference between $v_{B^k}(\omega, x)$ and $\tilde{v}_\alpha^k(\omega, x)$ for $\omega - Tx \in \Lambda^k(d_k)$ equals $\psi^k(\omega - Tx) - \psi^k(\omega - \alpha)$. This difference

10

is zero-mean periodic in ω, similar as for the shifted LP-relaxation, so that a similar total variation error bound can be derived for these generalized α-approximations, see Theorem 4 in Section 4.1. The generalized α-approximations are derived by taking the pointwise maximum over all dual feasible basis matrices B^k, $k = 1, \ldots, K$, yielding

$$\hat{v}_\alpha(\omega, x) = \max_{k=1,\ldots,K} \tilde{v}_\alpha^k(\omega, x),$$

and by taking the expected value.

Definition 8. For $\alpha \in \mathbb{R}^m$, we define the *generalized α-approximation* $\tilde{Q}_\alpha$ of Q as

$$\tilde{Q}_\alpha(x) := \mathbb{E}_\omega \left[\max_{k=1,\ldots,K} \{\lambda^k(\omega - Tx) + \psi^k(\omega - \alpha)\} \right], \quad x \in \mathbb{R}^{n_1},$$

with $\lambda^k := q_{B^k}(B^k)^{-1}$ and $\psi^k := \psi_{B^k}$, where B^k, $k = 1, \ldots, K$, are the dual feasible basis matrices of Lemma 1.

In the definition of $\tilde{Q}_\alpha$, the *shift parameter* α can take on any value. Indeed, the error bound we derive in Section 4.1 is valid independent of the value of α. That is, $\tilde{Q}_\alpha$ is a good approximation of the true MIR function in (2) for any choice of α.

An interesting difference between the generalized α-approximations and the shifted LP-relaxation of Section 2.3 is that the approximating value function $\tilde{v}_\alpha$ is not convex in ω for fixed $x \in \mathbb{R}^{n_1}$. The function is only convex in x for every fixed ω, but this is sufficient to guarantee that the generalized α-approximation $\tilde{Q}_\alpha$ is convex. In contrast, the value function $\hat{v}$ of the shifted LP-relaxation is convex in both ω and x. Example 2 illustrates these properties of $\tilde{v}_\alpha$ for a one-dimensional example.

Example 2. Consider the second-stage mixed-integer value function v defined as

$$v(\omega, x) := \min\{y_1 + 2y_2 + 2y_3 : \quad y_1 + y_2 - y_3 = \omega - x, \quad y_1 \in \mathbb{Z}_+, \quad y_2, y_3 \in \mathbb{R}_+\},$$

see also Example 3.3 in Romeijnders et al. [2016a]. The LP-relaxation of v has two dual feasible basis matrices $B^1 = [-1]$ and $B^2 = [1]$. Thus, $K = 2$, and straightforward computations yields $\lambda^1 = -2$, $\lambda^2 = 1$, $\psi^1 \equiv 0$, and for every $s \in \mathbb{R}$,

$$\psi^2(s) = \begin{cases} s - \lfloor s \rfloor, & \text{if } s - \lfloor s \rfloor \leq 3/4, \\ 3 - 3(s - \lfloor s \rfloor), & \text{if } s - \lfloor s \rfloor \geq 3/4. \end{cases}$$

For every $\alpha \in \mathbb{R}$, the approximating value function is thus

$$\tilde{v}_\alpha(\omega, x) = \max\{-2(\omega - x), \omega - x + \psi^2(\omega - \alpha).\}$$

In contrast, the value function $\hat{v}$ of the shifted LP-relaxation equals

$$\hat{v}(\omega, x) = \max\{-2(\omega - x), \omega - x + 3/8\},$$

since

$$\Gamma^2 := \int_0^1 \psi^2(s)ds = 3/8.$$

Figure 1 shows both v and $\tilde{v}_\alpha$. Observe that $\tilde{v}_\alpha$ is convex in x for fixed ω, but not convex in ω for fixed x. Moreover, Figure 1d illustrates that for this example $\tilde{v}_\alpha(\omega, x) = v(\omega, x)$ for all ω if $x \in \alpha + \mathbb{Z}$. $\diamond$

The difference between the shifted LP-relaxation of Definition 7 and the generalized α-approximations of Definition 8 seems small: the constants Γ_k are replaced by $\psi^k(\omega - \alpha)$. From a computational point of view, however, this difference is significant. This is because the constants Γ_k are the
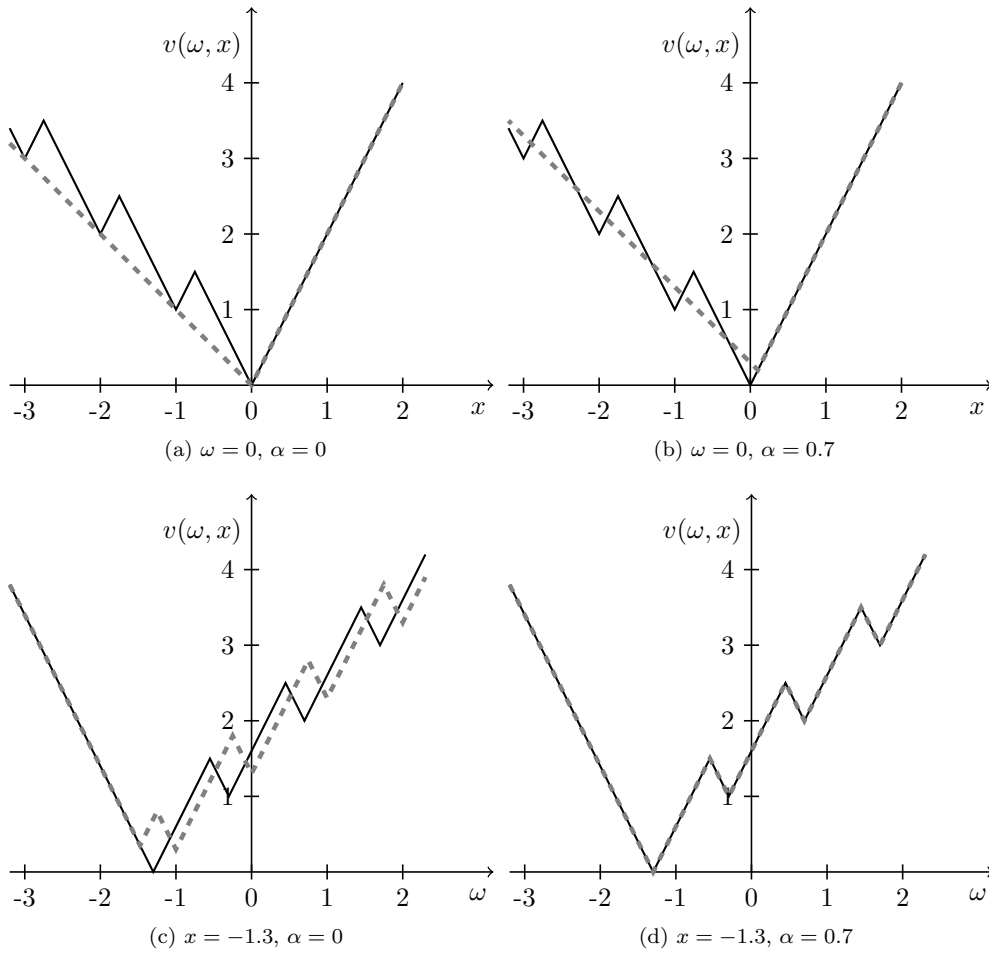
11

Figure 1: The value function $v(\omega, x)$ (solid) and approximating value function $\tilde{v}_\alpha(\omega, x)$ (dashed) of Example 2 as a function of x (Figures 1a and 1b) and as a function of ω (Figures 1c and 1d), where $\alpha \in \{0, 0.7\}$.

averages over ψ^k, and in general need to be obtained by computing a multi-dimensional integral of a mixed-integer value function. For a fixed ω and α, however, the value of $\psi^k(\omega - \alpha)$ is obtained by solving *a single* mixed-integer programming problem of the same size as the second-stage value function $v(\omega, x)$. In fact, we need to solve the Gomory relaxation discussed in Section 2.3.1, which can be done in polynomial time if all second-stage variables are integer [Gomory, 1958].

We conclude this section with Example 3 in which we show that for TU integer recourse models the generalized α-approximations reduce to the α-approximations of Van der Vlerk [2004] in Definition 3.

Example 3. Consider the TU integer recourse function Q discussed in Section 2.2. The second-stage value function v is defined as

$$v(\omega, x) := \min_y \left\{ qy : \quad Wy \geq \omega - Tx, \quad y \in \mathbb{Z}_+^{n_2} \right\},$$

where the recourse matrix W is TU. To obtain the generalized α-approximation of Q, rewrite v as

$$v(\omega, x) = \min_y \{ qy : \quad Wy - u = \omega - Tx, \quad y \in \mathbb{Z}_+^{p_2}, \quad u \in \mathbb{R}_+^m \},$$

where the decision vector u contains slack variables.

Let B^k, $k = 1, \ldots, K$, denote the dual feasible basis matrices of $\begin{pmatrix} W & -I \end{pmatrix}$, and let $\lambda^k = q_{B^k}(B^k)^{-1}$. Note that if $\lceil \omega - Tx \rceil \in \Lambda^k$, then

$$v(\omega, x) = \max_{k=1,\ldots,K} \lambda^k \lceil \omega - Tx \rceil = \lambda^k \lceil \omega - Tx \rceil.$$

Moreover, by Lemma 1, if $\omega - Tx \in \Lambda^k(d_k)$, then

$$v(\omega, x) = \lambda^k(\omega - Tx) + \psi^k(\omega - Tx).$$

It follows that $\psi^k(s) = \lambda^k(\lceil s \rceil - s)$, $s \in \mathbb{R}^m$.

Using these expressions, we obtain that the generalized α-approximation $\tilde{Q}_\alpha$ is given by

$$\tilde{Q}_\alpha(x) = \mathbb{E}_\omega \left[\max_{k=1,\ldots,K} \lambda^k(\lceil \omega - \alpha \rceil + \alpha - Tx) \right].$$

It follows from (7) that the generalized α-approximations equal the α-approximations of Van der Vlerk [2004] for TU integer recourse models. $\diamond$

3.2 Benders decomposition for the generalized α-approximations

To obtain an approximating solution $\tilde{x}_\alpha$, using the generalized α-approximations $\tilde{Q}_\alpha$, we solve

$$\tilde{\eta}_\alpha := \min_x \{ cx + \tilde{Q}_\alpha(x) : Ax = b, x \in \mathbb{R}_+^{n_1} \},$$

using a Benders decomposition [Benders, 1962] on an SAA of $\tilde{Q}_\alpha$. The challenge is that this involves computing $\tilde{Q}_\alpha(x)$ and a subgradient $u \in \partial \tilde{Q}_\alpha(x)$ at every iteration, which in turn requires evaluating $\psi^k(\omega - \alpha)$ for all ω and all $k = 1, \ldots, K$. This is inhibitive, since the number of dual feasible basis matrices K is exponentially large in the input size of the second-stage problem. That is why we use approximations for $\tilde{Q}_\alpha(x)$ and u that can be computed efficiently, see Section 3.3. This is justified, since $\tilde{Q}_\alpha$ is already an approximation of the true MIR function Q. Moreover, we show that the approximation error of this *double approximation* goes to zero as the total variations of the one-dimensional conditional pdf of the random vector ω go to zero.

Because we can only use a Benders decomposition on $\tilde{Q}_\alpha$ if ω follows a discrete distribution, we use an SAA of $\tilde{Q}_\alpha$. That is, we draw a sample $\omega^{(1)}, \ldots, \omega^{(S)}$ of size S from the distribution of ω, and we define the random vector $\tilde{\omega}$ as follows:

$$\mathbb{P}[\tilde{\omega} = \omega^{(s)}] = \frac{1}{S}, \quad s = 1, \ldots, S.$$

By replacing the distribution of ω by that of $\tilde{\omega}$ in the definition of $\tilde{Q}_\alpha$, we obtain the *sample average approximation* $\tilde{Q}_\alpha^S$ of $\tilde{Q}_\alpha$.

Definition 9. Let $\alpha \in \mathbb{R}^m$ be given and let $\omega^{(1)}, \ldots, \omega^{(S)}$ be a sample of size S from the distribution of ω. We define the *sample average approximation* $\tilde{Q}_\alpha^S$ of $\tilde{Q}_\alpha$ as

$$\tilde{Q}_\alpha^S(x) := \frac{1}{S} \sum_{s=1}^{S} \max_{k=1,\ldots,K} \{\lambda^k(\omega^{(s)} - Tx) + \psi^k(\omega^{(s)} - \alpha)\}. \tag{15}$$

Note that $\tilde{Q}_\alpha^S$ is a convex polyhedral function, so that we can construct an outer approximation of $\tilde{Q}_\alpha^S$ using a Benders decomposition in the spirit of the L-shaped algorithm [Van Slyke and Wets, 1969]. The idea is to rewrite

$$\min_x \{cx + \tilde{Q}_\alpha^S(x) : \quad Ax = b, \quad x \in \mathbb{R}_+^{n_1}\} \tag{16}$$

as

$$\min_x \{cx + \theta : \quad \theta \geq \tilde{Q}_\alpha^S(x), \quad Ax = b, \quad x \in \mathbb{R}_+^{n_1}, \quad \theta \in \mathbb{R}\}.$$

Since $\tilde{Q}_\alpha^S$ is convex polyhedral, the constraint $\theta \geq \tilde{Q}_\alpha^S(x)$ can be represented by finitely many linear constraints of the form

$$\theta \geq \beta_r x + \delta_r, \quad r = 1, \ldots, R. \tag{17}$$

In a Benders decomposition, these so-called *optimality cuts* are added sequentially in order to obtain an improved description of $\tilde{Q}_\alpha^S$. Moreover, not all constraints in (17) are included, only those necessary to find an optimal solution. At iteration τ, we first solve the *master problem*

$$\min_x \{cx + \theta : \quad Ax = b, \quad \theta \geq \alpha_r + \beta_r x, \quad r = 1, \ldots, \tau, \quad x \in \mathbb{R}_+^{n_1}, \quad \theta \in \mathbb{R}\},$$

where $\tau \leq R$ is the number of optimality cuts that have been added to the master problem. That is, only a subset of all the constraints in (17) have been added to the master problem. Next, we check if the *current solution* $(x^{(\tau)}, \theta^{(\tau)})$, which is an optimal solution in the master problem, satisfies the optimality criterion $\theta^{(\tau)} \geq \tilde{Q}_\alpha^S(x^{(\tau)})$. If so, the algorithm terminates. Else, we exploit convexity of $\tilde{Q}_\alpha^S$ to derive a new optimality cut, based on the subgradient inequality

$$\theta \geq \tilde{Q}_\alpha^S(x) \geq \tilde{Q}_\alpha^S(x^{(\tau)}) + u^{(\tau)}(x - x^{(\tau)}), \tag{18}$$

where $u^{(\tau)} \in \partial \tilde{Q}_\alpha^S(x^{(\tau)})$ is a subgradient of $\tilde{Q}_\alpha^S$ at $x^{(\tau)}$. Note that this optimality cut cuts away the current solution $(x^{(\tau)}, \theta^{(\tau)})$. The master problem is then updated by adding this optimality cut. This procedure always terminates, because $\tilde{Q}_\alpha^S$ is convex polyhedral and can therefore be described by finitely many linear inequalities.

The problem is that $\tilde{Q}_\alpha^s(x^{(\tau)})$ and the subgradient $u^{(\tau)}$ need to be computed using the expressions

$$\tilde{Q}_\alpha^S(x^{(\tau)}) = \frac{1}{S} \sum_{s=1}^{S} \max_{k=1,\ldots,K} \{\lambda^k(\omega^{(s)} - Tx^{(\tau)}) + \psi^k(\omega^{(s)} - \alpha)\},$$

and

$$u^{(\tau)} := -\frac{1}{S} \sum_{s=1}^{S} \lambda^{k_s} Tx^{(\tau)} \in \partial \tilde{Q}_\alpha^S(x^{(\tau)}),$$

where

$$k_s \in \arg\max_{k=1,\ldots,K} \{\lambda^k(\omega^{(s)} - Tx^{(\tau)}) + \psi^k(\omega^{(s)} - \alpha)\}, \quad s = 1, \ldots, S. \tag{19}$$

This is computationally expensive, since to determine k_s, we need to compute $\psi^k(\omega^{(s)} - \alpha)$ for all $k = 1, \ldots, K$, and K grows exponentially in the input size of the second-stage problem. That is why, in Section 3.3, we derive loose optimality cuts, which can be computed much faster than the tight optimality cuts in (18). We use them to develop LBDA(α), our loose Benders decomposition algorithm, which solves the approximating model in (16). The loose optimality cuts we propose are tight enough in the sense that the solution $\tilde{x}_\alpha$ obtained by LBDA(α) is of high quality. We formalize this in Proposition 1, which contains a performance guarantee on $\tilde{x}_\alpha$.

3.3 Loose optimality cuts for the generalized α-approximations

In this section, we derive loose optimality cuts for the SAA of the generalized α-approximation $\tilde{Q}_\alpha^S$, defined in (15). In other words, we derive inequalities of the form $\tilde{Q}_\alpha^S(x) \geq \beta x + \delta$, but we do not require that this inequality is tight at the current solution $x^{(\tau)}$. We say that the inequality is tight at $x^{(\tau)}$ if $\tilde{Q}_\alpha^S(x^{(\tau)}) = \beta x^{(\tau)} + \delta$.

It follows directly from the definition of $\tilde{Q}_\alpha^S$ in (15) that if we pick any $\tilde{k}_s \in \{1, \ldots, K\}$ for $s = 1, \ldots, S$, then

$$\tilde{Q}_\alpha^S(x) \geq \frac{1}{S} \sum_{s=1}^{S} \lambda^{k_s}(\omega^{(s)} - Tx) + \psi^{k_s}(\omega^{(s)} - \alpha). \tag{20}$$

The inequality in (20) is tight if we select $\tilde{k}_s = k_s$, $s = 1, \ldots, S$, with k_s defined in (19). In all other cases, the inequality defines a loose optimality cut of the form $\beta x + \delta$. In order to guarantee that the loose optimality cut is nearly tight, we consider the inner maximization problem

$$\max_{k=1,\ldots,K} \{\lambda^k(\omega^{(s)} - Tx^{(\tau)}) + \psi^k(\omega^{(s)} - \alpha)\}, \tag{21}$$

$s = 1, \ldots, S$, in the definition of $\tilde{Q}_\alpha^S$. We will identify a basis matrix index $\tilde{k}_s$ that is optimal in (21) with high probability by setting $\psi^k \equiv 0$ for all $k = 1, \ldots, K$. Then, we define $\tilde{k}_s$ as an optimal index in

$$\max_{k=1,\ldots,K} \lambda^k(\omega^{(s)} - Tx^{(\tau)}). \tag{22}$$

By LP-duality, this choice of $\tilde{k}_s$ coincides with the optimal basis matrix index in the LP-relaxation

$$v_{\text{LP}}(\omega^{(s)}, x^{(\tau)}) = \min_y \left\{ qy : \quad Wy = \omega^{(s)} - Tx^{(\tau)}, \quad y \in \mathbb{R}_+^{n_2} \right\} \tag{23}$$

of v. This means that to obtain $\tilde{k}_s$, we merely have to solve the LP in (23), which can be done fast, and find the optimal basis matrix $B^{\tilde{k}_s}$. Next, using (20), we construct our loose optimality cut.

Definition 10. Let $x^{(\tau)}$ be given and let $\tilde{k}_s$ be the optimal basis matrix index in the LP problem in (23). We define the *loose optimality cut* at $x^{(\tau)}$ as

$$\beta x + \delta = \frac{1}{S} \sum_{s=1}^{S} \lambda^{\tilde{k}_s}(\omega^{(s)} - Tx) + \psi^{\tilde{k}_s}(\omega^{(s)} - \alpha), \quad x \in \mathbb{R}^{n_1}.$$

To compute this cut we need to compute $\psi^{\tilde{k}_s}(\omega^{(s)} - \alpha)$ for every $s = 1, \ldots, S$, by solving a mixed-integer programming problem of similar size as the second-stage problem. In Section 4.2, we show that the loose optimality cut at $x^{(\tau)}$ is nearly tight at $x^{(\tau)}$ by proving a total variation bound on the difference $\tilde{Q}_\alpha^S(x^{(\tau)}) - (\beta x^{(\tau)} + \delta)$, see Theorem 5.

3.4 Loose Benders decomposition for the generalized α-approximations

We are now ready to describe LBDA(α), the loose Benders decomposition algorithm for the generalized α-approximations. We solve the generalized α-approximations by applying the Benders decomposition described in Section 3.2, using the loose optimality cuts in (20). The resulting solution $\tilde{x}_\alpha$ is near-optimal in the original MIR problem (1). Indeed, Proposition 1 contains a total variation bound on the optimality gap of the resulting solution $\tilde{x}_\alpha$. That is, $\tilde{x}_\alpha$ is not necessarily optimal in (1), but we derive a total variation bound on

$$c\tilde{x}_\alpha + Q(\tilde{x}_\alpha) - \eta^*.$$

This bound holds with probability 1 as the sample size S of the sample average approximation approaches infinity.

LBDA(α) is developed for the case $X = \mathbb{R}_+^{n_1}$ and $Y = \mathbb{Z}_+^{p_2} \times \mathbb{R}_+^{n_2 - p_2}$. Furthermore, L denotes a lower bound on $\tilde{Q}_\alpha^S$, that is, $\tilde{Q}_\alpha^S(x) \geq L$ for all $x \in \mathcal{X} = \{x \in X : Ax = b\}$. Such an L exists because of Assumptions (A1)-(A4).

Loose Benders decomposition algorithm (LBDA(α))

1: **Inputs** Parameters: A, b, c, T, q, W. Distribution of ω. Lower bound L on $\tilde{Q}_\alpha^S$. Shift parameter α. Tolerance ε. Sample size S.
2: **Output** Near-optimal solution $\hat{x}$.
3: **Initialization**
4:     Initialize $\tau = 0$.
5:     Obtain a sample $\omega^{(1)}, \ldots, \omega^{(S)}$ of size S from the distribution of ω.
6: **Iteration step**
7:     Solve $\min_x \{cx + \theta : \quad Ax = b, \quad \theta \geq \beta_r x + \delta_r, \quad r = 1, \ldots, \tau, \quad x \in \mathbb{R}_+^{n_1}, \quad \theta \geq L\}$,
8:     denote the optimal solution by $x^{(\tau)}, \theta^{(\tau)}$.
9:     **for** $s = 1, \ldots, S$ **do**
10:         Solve $\min_y \{qy : \quad Wy = \omega^{(s)} - Tx^{(\tau)}, \quad y \in \mathbb{R}_+^{n_2}\}$,
11:         denote the optimal basis matrix index by k_s.
12:         Evaluate $\psi^{k_s}(\omega^{(s)} - \alpha)$.
13:     **end for**
14:     $\beta_{\tau+1} \leftarrow -\frac{1}{S} \sum_{s=1}^{S} \lambda^{k_s} T$.
15:     $\delta_{\tau+1} \leftarrow \frac{1}{S} \sum_{s=1}^{S} \lambda^{k_s} \omega^{(s)} + \psi^{k_s}(\omega^{(s)} - \alpha)$.
16: **Stopping criterion**
17:     **if** $\theta^{(\tau)} \geq \beta_{\tau+1} x^{(\tau)} + \delta_{\tau+1} - \varepsilon$ **then**
18:         **return** $\tilde{x}_\alpha := x^{(\tau)}$.
19:         **stop**.
20:     **else**
21:         $r \leftarrow r + 1$. Go to line 7.
22:     **end if**

In the iteration step of LBDA(α), we solve the master problem and we generate the loose optimality cut coefficients $\beta_{\tau+1}$ and $\delta_{\tau+1}$. We use these coefficients to check if the stopping criterion

$$\theta^{(r)} \geq \beta_{\tau+1} x^{(\tau)} + \delta_{\tau+1} - \varepsilon$$

is satisfied. If so, then the algorithm terminates: $x^{(\tau)}$ is near-optimal. Else, we add the loose optimality cut $\theta \geq \beta_{\tau+1} x + \delta_{\tau+1}$, which cuts away the current solution $(x^{(\tau)}, \theta^{(\tau)})$.

Note that $x^{(\tau)}$ is not necessarily optimal in the approximating problem

$$\min_x \left\{ cx + \tilde{Q}_\alpha^S(x) : \quad Ax = b, \quad x \in \mathbb{R}_+^{n_1} \right\}. \tag{24}$$

This is because we use loose optimality cuts to construct an outer approximation of $\tilde{Q}_\alpha^S$. To ensure optimality of $x^{(\tau)}$ in (24), we should have used the stopping criterion $\theta^{(\tau)} \geq \tilde{Q}_\alpha^S(x^{(\tau)}) - \varepsilon$. However, evaluating $\tilde{Q}_\alpha^S(x^{(\tau)})$ is computationally expensive, which is why we used loose optimality cuts in the first place. Fortunately, this turns out not to be a problem, as (24) is itself an approximation of the true MIR problem in (1). Moreover, the loose optimality cuts we use are nearly tight, implying that $x^{(\tau)}$ is near-optimal in (24), and consequently in (1). Indeed, in Proposition 1, we ensure the quality of $x^{(\tau)}$ in (1) by proving a bound on the optimality gap of $x^{(\tau)}$.

LBDA(α) can be implemented efficiently if the input size of the second-stage problem $v(\omega, x)$ is moderate. During each iteration, we have to solve the LP relaxation $v_{\text{LP}}(\omega, x)$ and a Gomory

relaxation $v_B(\omega, x)$ of $v(\omega, x)$ S times in order to generate an optimality cut. The Gomory relaxation $v_B(\omega, x)$ can be solved in reasonable time if the input size of $v(\omega, x)$ is not too large. Moreover, $v_{LP}(\omega, x)$ as well as the master problem can be solved efficiently using standard LP solvers.

Improved implementations of LBDA(α) using a multicut approach [see Birge and Louveaux, 1988] and regularization techniques [see Ruszczyński, 1986] are possible. Furthermore, parallelization can be implemented by simultaneously evaluating $\psi^{k_s}(\omega^{(s)} - \alpha)$, $s = 1, \ldots, S$. Finally, during the initial phase of the algorithm, $\psi^{k_s}(\omega^{(s)} - \alpha)$ can be solved inexactly to speed up computations. Indeed, a lower bound on $\psi^{k_s}(\omega^{(s)} - \alpha)$ suffices to generate a valid optimality cut. In this case, the termination criterion should be disregarded and the iteration step should be repeated. During the final iteration of the algorithm, all subproblems $\psi^{k_s}(\omega^{(s)} - \alpha)$, $s = 1, \ldots, S$, should be solved exactly, to ensure that the stopping criterion is valid.

Proposition 1. *Consider the two-stage mixed-integer recourse model*

$$\eta^* = \min_x \{cx + Q(x) : \quad Ax = b, \quad x \in \mathbb{R}^{n_1}_+\}. \tag{25}$$

Let $\tilde{x}_\alpha$ denote the solution by LBDA(α) with tolerance ε and sample size S. Then, there exists a constant $C > 0$ such that for every continuous random vector ω with joint pdf $f \in \mathcal{H}^m$

$$\hat{\tilde{c}}x_\alpha + Q(\tilde{x}_\alpha) - \eta^* \leq \varepsilon + C \sum_{i=1}^{m} \mathbb{E}_{\omega_{-i}} \left[|\Delta| f_i(\cdot | \omega_{-i}) \right],$$

with probability 1 as $S \to \infty$.

Proposition 1 states that the optimality gap of the solution $\tilde{x}_\alpha$ returned by LBDA(α) converges to the pre-specified tolerance ε as the total variations of the underlying one-dimensional conditional pdf go to zero. In other words, LBDA(α) performs well under these conditions.

4 Performance guarantees of LBDA(α) and the generalized α-approximations

In this section, we prove Proposition 1. First, however, we derive several intermediate results. In particular, we derive an error bound for the generalized α-approximations in Section 4.1. In Section 4.2, we prove that $\tilde{Q}_\alpha^S$, the SAA of the generalized α-approximation $\tilde{Q}_\alpha$, converges uniformly to $\tilde{Q}_\alpha$ as $S \to \infty$. In addition, we show that the loose optimality cuts discussed in Section 3.3 are nearly tight as $S \to \infty$. In Section 4.3, we combine these results to prove Proposition 1.

4.1 Error bound for the generalized α-approximations

Consider the generalized α-approximation $\tilde{Q}_\alpha$, defined in Definition 8 as $\tilde{Q}_\alpha(x) = \mathbb{E}_\omega[\tilde{v}_\alpha(\omega, x)]$, where the approximating value function $\tilde{v}_\alpha$ is defined as

$$\tilde{v}_\alpha(w, x) := \max_{k=1,\ldots,K} \{\lambda^k(\omega - Tx) + \psi^k(\omega - \alpha)\}. \tag{26}$$

In Theorem 4, we derive an error bound on $||Q - \tilde{Q}_\alpha||_\infty$ by considering the difference between $\tilde{v}_\alpha$ and the MIR value function v, defined in (8). Recall from Lemma 1 that v is asymptotically periodic, i.e. $v(\omega, x) = \lambda^k(\omega - Tx) + \psi^k(\omega - Tx)$ if $\omega - Tx \in \Lambda^k(d_k)$. In Lemma 2, we prove a similar result for $\tilde{v}_\alpha$. Together, this implies that the difference $\Delta(\omega, x) := v(\omega, x) - \tilde{v}_\alpha(\omega, x)$ is zero-mean periodic in ω on large parts of the domain. Moreover, we show that $\Delta(\omega, x)$ is uniformly bounded. Applying the total variation bounds discussed in Section 2.3.2 yields the error bound for $\tilde{Q}_\alpha$.

Lemma 2. Consider the mixed-integer value function v defined in (3) and the value function $\tilde{v}_\alpha$ of $\tilde{Q}_\alpha$ defined in (26). Let B^k and Λ^k, $k = 1, \ldots, K$, denote the basis matrices and closed convex cones from Lemma 1. Then, there exist vectors $\sigma_k \in \Lambda^k$ and a constant $R > 0$ such that

(i) for all $\omega \in \mathbb{R}^m$ and $x \in \mathbb{R}^{n_1}$,

$$|v(\omega, x) - \tilde{v}_\alpha(\omega, x)| \leq R,$$

(ii) $\omega - Tx \in \sigma_k + \Lambda^k \implies \tilde{v}_\alpha(\omega, x) = \lambda^k(\omega - Tx) + \psi^k(\omega - \alpha).$

Proof. Consider the dual representation of the LP-relaxation of v, denoted v_{LP}, in (9). There exists a constant $R_1 > 0$ such that for all $\omega \in \mathbb{R}^m$ and $x \in \mathbb{R}^{n_1}$,

$$|v(\omega, x) - v_{\mathrm{LP}}(\omega, x)| \leq R_1,$$

see e.g. Blair and Jeroslow [1979]. Furthermore, there exists a constant $R_2 > 0$ such that

$$|v_{\mathrm{LP}}(\omega, x) - \tilde{v}_\alpha(\omega, x)| \leq R_2, \tag{27}$$

so that (i) follows from the triangle inequality by setting $R = R_1 + R_2$. Such an R_2 exists, since by Lemma 1, there exist w_k such that $0 \leq \psi^k(\omega - \alpha) \leq w_k$, $k = 1, \ldots, K$. Thus, we can take $R_2 = \max_k w_k$, so that (27) follows from (9) and (26).

To prove (ii), let B^k be the dual feasible basis matrices from Lemma 1. Fix arbitrary $l \in \{1, \ldots, K\}$. It suffices to show that there exist $\sigma_{kl} \in \Lambda^k(d_k)$ such that $\omega - Tx \in \sigma_{kl} + \Lambda^k$ implies that

$$\lambda^k(\omega - Tx) + \psi^k(\omega - \alpha) \geq \lambda^l(\omega - Tx) + \psi^l(\omega - \alpha). \tag{28}$$

This is because

$$\bigcap_{l=1}^{K}(\sigma_{kl} + \Lambda^k) = \sigma_k + \Lambda^k$$

for some $\sigma_k \in \Lambda^k$. Hence, if $\omega - Tx \in \sigma_k + \Lambda^k$, then

$$\tilde{v}_\alpha(\omega, x) = \max_{k=1,\ldots,K}\left\{\lambda^k(\omega - Tx) + \psi^k(\omega - \alpha)\right\} = \lambda^k(\omega - Tx) + \psi^k(\omega - \alpha).$$

To prove (28), note that if $\lambda^k = \lambda^l$, then by Lemma 1 (v), $\psi^k(\omega - \alpha) = \psi^l(\omega - \alpha)$, so that (28) holds with equality. If $\lambda^k \neq \lambda^l$, then $\lambda^k s > \lambda^l s$ for any $s \in \mathrm{int}(\Lambda^k)$. For sufficiently large $\gamma > 0$, we thus have

$$\gamma(\lambda^k s - \lambda^l s) \geq w_l.$$

If we take $\sigma_{kl} = \gamma s$, then (28) holds by observing that $\psi^k(\omega - \alpha) \geq 0$ and $\psi^l(\omega - \alpha) \leq w_l$. $\qquad \square$

Theorem 4. *Consider the mixed-integer recourse function Q defined as $Q(x) = \mathbb{E}_\omega[v(\omega, x)]$, where the value function v is defined in (3). Consider the generalized α-approximation $\tilde{Q}_\alpha$ of Definition 8. Then, there exists a constant $C > 0$ such that for every continuous random vector ω with joint pdf $f \in \mathcal{H}^m$,*

$$\sup_x |Q(x) - \tilde{Q}_\alpha(x)| \leq C \sum_{i=1}^{m} \mathbb{E}_{\omega_{-i}}\left[|\Delta|f_i(\cdot|\omega_{-i})\right].$$

Proof. Fix $x \in \mathbb{R}^{n_1}$ and $f \in \mathcal{H}^m$, and let $\tilde{\omega} = \omega - Tx$. Note that $\tilde{\omega}$ has pdf $g \in \mathcal{H}^m$ given by $g(s) = f(s + Tx)$, $s \in \mathbb{R}^m$. Define $\Delta(\omega, x) := v(\omega, x) - \tilde{v}_\alpha(\omega, x)$, and observe that $\Delta(\omega, x) = \Delta(\hat{\omega}.0)$. We use this to rewrite the difference $|Q(x) - \tilde{Q}_\alpha(x)|$ as

$$|Q(x) - \tilde{Q}_\alpha(x)| = \left|\mathbb{E}_\omega\left[\Delta(\omega, x)\right]\right| = \left|\mathbb{E}_{\tilde{\omega}}\left[\Delta(\tilde{\omega}, 0)\right]\right|.$$

Let Λ^k and d_k be the closed convex cones and constants from Lemma 1. Let $\sigma_k \in \Lambda^k$ be the vectors from Lemma 2. Define $\bar{\Lambda}^k := \Lambda^k(d_k) \cap (\sigma_k + \Lambda^k)$. It follows from Lemmas 1 and 2 that if $\tilde{\omega} \in \bar{\Lambda}^k$, then

$$\Delta(\tilde{\omega}, 0) := \psi^k(\tilde{\omega}) - \psi^k(\tilde{\omega} - \alpha),$$

which is zero-mean B^k-periodic in $\tilde{\omega}$. Define $\mathcal{M} := \cup_{k=1}^K \bar{\Lambda}^k$. Note that

$$|Q(x) - \tilde{Q}_\alpha(x)| = \left| \mathbb{E}_\omega [\Delta(\tilde{\omega}, 0)] \right|$$

$$\leq \left| \int_{\mathcal{M}} \Delta(\tilde{\omega}, 0) g(\tilde{\omega}) d\tilde{\omega} \right| + \left| \int_{\mathbb{R}^m \setminus \mathcal{M}} \Delta(\tilde{\omega}, 0) g(\tilde{\omega}) d\tilde{\omega} \right|. \tag{29}$$

For the first term in (29) we have

$$\left| \int_{\mathcal{M}} \Delta(\tilde{\omega}, 0) g(\tilde{\omega}) d\tilde{\omega} \right| \leq \sum_{k=1}^K \left| \int_{\bar{\Lambda}^k} [\psi^k(\tilde{\omega}) - \psi^k(\tilde{\omega} - \alpha)] g(\tilde{\omega}) d\tilde{\omega} \right|.$$

Applying the total variation bound on the expectation of periodic functions in Theorem 4.13 of Romeijnders et al. [2016a], yields that there exists a constant $C_1 > 0$, independent of g, such that

$$\left| \int_{\mathcal{M}} \Delta(\tilde{\omega}, 0) g(\tilde{\omega}) d\tilde{\omega} \right| \leq C_1 \sum_{i=1}^m \mathbb{E}_{\tilde{\omega}_{-i}} \left[|\Delta| g_i(\cdot | \tilde{\omega}_{-i}) \right].$$

For the second term in (29), we use a result in Romeijnders et al. [2016a], which states that there exist hyperslices $H_j := \{x : 0 \leq a_j^T x \leq \delta_j\}$, $j = 1, \dots, J$, such that

$$\mathbb{R}^m \setminus \mathcal{M} \subset \bigcup_{j=1}^J H_j.$$

We thus obtain

$$\left| \int_{\mathbb{R}^m \setminus \mathcal{M}} \Delta(\tilde{\omega}, 0) g(\tilde{\omega}) d\tilde{\omega} \right| \leq \sum_{j=1}^J \int_{H_j} |\Delta(\tilde{\omega}, 0)| g(\tilde{\omega}) d\tilde{\omega}$$

$$\leq R \sum_{j=1}^J \mathbb{P}[\tilde{\omega} \in H_j],$$

where R is the upper bound on $|v(\omega, x) - \tilde{v}_\alpha(\omega, x)|$ from Lemma 2. By the total variation probability bound in Theorem 4.6 of Romeijnders et al. [2016a], there exists a constant $C_2 > 0$, independent of g, such that

$$\left| \int_{\mathbb{R}^m \setminus \mathcal{M}} \Delta(\tilde{\omega}, 0) g(\tilde{\omega}) d\tilde{\omega} \right| \leq C_2 \sum_{i=1}^m \mathbb{E}_{\tilde{\omega}_{-i}} \left[|\Delta| g_i(\cdot | \tilde{\omega}_{-i}) \right].$$

The result now follows from (29) by setting $C = C_1 + C_2$ and by noting that $\mathbb{E}_{\tilde{\omega}_{-i}} [|\Delta| g_i(\cdot | \tilde{\omega}_{-i})] = \mathbb{E}_{\omega_{-i}} [|\Delta| f_i(\cdot | \omega_{-i})]$. $\qquad \square$

Theorem 4 states that the maximum difference between the MIR function Q and the generalized α-approximation $\tilde{Q}_\alpha$ is bounded. Moreover, this maximum difference goes to zero as the total variations $\mathbb{E}_{\omega_{-i}} [|\Delta| f_i(\cdot | \omega_{-i})]$, $i = 1, \dots, m$, all go to zero. As we will show, LBDA(α) generates a near-optimal solution $\tilde{x}_\alpha$ to the generalized α-approximation, so that we can use Theorem 4 to prove Proposition 1. Indeed, Theorem 4 then implies that $\tilde{x}_\alpha$ is also near-optimal in the original MIR problem (1).

4.2 Convergence of the sample average approximation $\tilde{Q}_\alpha^S$ and loose optimality cuts

In LBDA(α), we sample from the distribution of the random vector ω in order to obtain a discrete approximation of the true distribution of ω. Intuitively, as the sample size S increases, this approximation becomes better. In this section, we formalize this intuition. Lemma 3 states that the SAA $\tilde{Q}_\alpha^S$ of $\tilde{Q}_\alpha$ converges uniformly to $\tilde{Q}_\alpha$ on finite sets as $S \to \infty$. Lemma 4 extends this to infinite bounded sets. Furthermore, we consider the limiting behaviour of the loose optimality cuts discussed in Section 3.3 as $S \to \infty$. In particular, we show that our loose optimality cuts are nearly tight with probability 1 as $S \to \infty$, see Lemma 5.

We say that events such as convergence and inequalities happen with probability 1 as $S \to \infty$ if for almost every realization $\bar{\omega} := \{\bar{\omega}^{(1)}, \bar{\omega}^{(2)}, \dots\}$ of the random sample drawn from the distribution of ω, there exists an integer $S(\bar{\omega})$ such that the event occurs for all samples $\{\bar{\omega}^{(1)}, \dots, \bar{\omega}^{(s)}\}$ from $\bar{\omega}$ with $s \geq S(\bar{\omega})$ [Kleywegt et al., 2002].

Lemma 3. Consider the generalized α-approximation $\tilde{Q}_\alpha$ and its sample average approximation $\tilde{Q}_\alpha^S$. Suppose that $\mathcal{X} \subset \mathbb{R}^{n_1}$ is a finite set. Then,

$$\max_{x \in \mathcal{X}} |\tilde{Q}_\alpha(x) - \tilde{Q}_\alpha^S(x)| \to 0$$

with probability 1 as $S \to \infty$.

Proof. Fix any $x \in \mathcal{X}$. Consider the random variable

$$\xi := \max_{k=1,\dots,K} \{\lambda^k(\omega^{(s)} - Tx) + \psi^k(\omega^{(s)} - \alpha)\}.$$

By Assumptions (A2)-(A4), $\mathbb{E}[\xi]$ exists and is finite. Thus, by the strong law of large numbers (SLLN), $\tilde{Q}_\alpha^S(x) \to \tilde{Q}_\alpha(x)$ with probability 1 as $S \to \infty$. The result follows because $\mathcal{X}$ is finite. $\square$

Lemma 4. Consider the generalized α-approximation $\tilde{Q}_\alpha$ and its sample average approximation $\tilde{Q}_\alpha^S$. Then,

$$\sup_{x \in \mathcal{X}} \left| \tilde{Q}_\alpha(x) - \tilde{Q}_\alpha^S(x) \right| \to 0$$

with probability 1 as $S \to \infty$, where $\mathcal{X} = \{x \in \mathbb{R}_+^{n_1} : Ax \leq b\}$.

Proof. Our line of proof is based on Ahmed and Shapiro [2002]. For any $\nu > 0$, consider a finite set $\mathcal{X}_\nu$ such that for all $x \in \mathcal{X}$, there exists an $x' \in \mathcal{X}_\nu$ such that $||x - x'|| \leq \nu$. Such a set $\mathcal{X}_\nu$ exist due to Assumption (A1). Let $x \in \mathcal{X}$ be given and let $x' \in \mathcal{X}_\nu$ be such that $||x - x'|| \leq \nu$. Note that

$$|\tilde{Q}_\alpha(x) - \tilde{Q}_\alpha^S(x)| \leq |\tilde{Q}_\alpha(x) - \tilde{Q}_\alpha(x')| + |\tilde{Q}_\alpha(x') - \tilde{Q}_\alpha^S(x')| + |\tilde{Q}_\alpha^S(x') - \tilde{Q}_\alpha^S(x)|. \tag{30}$$

The first and third term on the right-hand side of (30) can be bounded by noting that both $\tilde{Q}_\alpha$ and $\tilde{Q}_\alpha^S$ are Lipschitz continuous. Denote Lipschitz constants of $\tilde{Q}_\alpha$ and $\tilde{Q}_\alpha^S$ by L_1 and L_2, respectively. We obtain

$$|\tilde{Q}_\alpha(x) - \tilde{Q}_\alpha^S(x)| \leq (L_1 + L_2)\nu + |\tilde{Q}_\alpha(x') - \tilde{Q}_\alpha^S(x')|,$$

which gives

$$\sup_{x \in X} \left| \tilde{Q}_\alpha(x) - \tilde{Q}_\alpha^S(x) \right| \leq (L_1 + L_2)\nu + \sup_{x \in X_\nu} \left| \tilde{Q}_\alpha(x') - \tilde{Q}_\alpha^S(x') \right|.$$

The first term $(L_1 + L_2)\nu$ can be made arbitrarily small by letting $\nu \to 0$. The result follows, because for fixed ν, the second term $\sup_{x \in X_\nu} \left| \tilde{Q}_\alpha(x') - \tilde{Q}_\alpha^S(x') \right|$ goes to zero with probability 1 as $S \to \infty$ due to Lemma 3. $\square$

In Lemma 5, we consider the loose optimality cuts discussed in Section 3.3, which are of the form

$$\tilde{Q}_\alpha^S(x) \geq \beta x + \delta = \frac{1}{S} \sum_{s=1}^S \lambda^{k_s}(\omega^{(s)} - Tx) + \psi^{k_s}(\omega^{(s)} - \alpha),$$

where $k_s \in \arg \max \lambda^k(\omega^{(s)} - T\bar{x})$. In particular, we show that these are nearly tight at the current solution $\bar{x}$, that is, we show that the difference $\tilde{Q}_\alpha^S(\bar{x}) - (\beta\bar{x} + \delta)$ is small. To this end, we define an intermediate function $\tilde{\mathcal{Q}}_\alpha^S$, which we will also use in the proof of Proposition 1. Define $\tilde{\mathcal{Q}}_\alpha^S$ as

$$\tilde{\mathcal{Q}}_\alpha^S(x) := \frac{1}{S} \sum_{s=1}^S \hat{v}_\alpha(\omega^{(s)}, x), \quad x \in \mathbb{R}^{n_1}, \tag{31}$$

where the value function $\hat{v}_\alpha$ is defined as

$$\hat{v}_\alpha(\omega, x) = \lambda^{k(\omega,x)}(\omega - Tx) + \psi^{k(\omega,x)}(\omega - \alpha),$$

and $k(\omega, x) \in \arg \max_{k=1,\dots,K} \lambda^k(\omega - Tx)$. Recall that $\tilde{Q}_\alpha^S$ is defined as

$$\tilde{Q}_\alpha^S(x) = \frac{1}{S} \sum_{s=1}^S \tilde{v}_\alpha(\omega^{(s)}, x),$$

where $\tilde{v}_\alpha$ is defined is (26). It follows immediately that $\tilde{\mathcal{Q}}_\alpha^S$ is a lower bound of $\tilde{Q}_\alpha^S$. Note that $k(\omega^{(s)}, \bar{x}) = k_s$, and thus $\tilde{\mathcal{Q}}_\alpha^S(\bar{x}) = \beta\bar{x} + \delta$. Hence, by proving a total variation bound on $||\tilde{Q}_\alpha^S - \tilde{\mathcal{Q}}_\alpha^S||_\infty$, as we do in Lemma 5, we ensure that our loose optimality cuts are nearly tight.

The intuition behind Lemma 5 is that $k(\omega, x)$ is optimal in $\tilde{v}_\alpha$ with high probability, and thus $\hat{v}_\alpha(\omega, x) = \tilde{v}_\alpha(\omega, x)$. In addition, the difference $\tilde{v}_\alpha(\omega, x) - \hat{v}_\alpha(\omega, x)$ is bounded. Applying the total variation probability bound discussed in Section 2.3.2 completes the proof.

Lemma 5. Consider the SAA of the generalized α-approximation $\tilde{Q}_\alpha^S$ and its lower bound $\tilde{\mathcal{Q}}_\alpha^S$, defined in (31). There exists a constant $C > 0$ such that for every continuous random vector ω with joint pdf $f \in \mathcal{H}^m$,

$$\sup_{x \in X} \left| \tilde{Q}_\alpha^S(x) - \tilde{\mathcal{Q}}_\alpha^S(x) \right| \leq C \sum_{i=1}^m \mathbb{E}_{\omega_{-i}} \left[|\Delta| f_i(\cdot | \omega_{-i}) \right]$$

with probability 1 as $S \to \infty$.

Proof. Define $\Delta(\omega, x) := \tilde{v}_\alpha(\omega, x) - \hat{v}_\alpha(\omega, x)$, so that

$$\tilde{Q}_\alpha^S(x) - \tilde{\mathcal{Q}}_\alpha^S(x) = \frac{1}{S} \sum_{s=1}^S \Delta(\omega^{(s)}, x).$$

We derive an upper bound on $\Delta(\omega, x)$, independent of x. We then apply the SLLN to obtain the desired result.

We first show that for a given x, with high probability $\Delta(\omega, x) = 0$. More formally, we show that for some $\mathcal{M} \subset \mathbb{R}^m$,

$$\omega - Tx \in \mathcal{M} \implies \tilde{v}_\alpha(\omega, x) = \hat{v}_\alpha(\omega, x) \tag{32}$$

and we derive an upper bound on

$$\mathbb{P}[\exists x \in \mathcal{X} : \omega - Tx \notin \mathcal{M}].$$

By the Basis Decomposition Theorem of Walkup and Wets [1969], there exist basis matrices B^k, $k = 1, \ldots, K$, and closed convex cones $\Lambda^k = \{t : (B^k)^{-1}t \geq 0\}$ such that $\omega - Tx \in \Lambda^k$ implies $k(\omega, x) = k$, so that

$$\hat{v}_\alpha(\omega, x) = \lambda^k(\omega - Tx) + \psi^k(\omega - \alpha). \tag{33}$$

Moreover, by Lemma 2, there exist $\sigma_k \in \Lambda^k$ such that if $\omega - Tx \in \sigma_k + \Lambda^k$, then

$$\tilde{v}_\alpha(\omega, x) = \lambda^k(\omega - Tx) + \psi^k(\omega - \alpha).$$

Hence, if we define $\mathcal{M} := \bigcup_{k=1}^K (\sigma_k + \Lambda^k)$, then the implication in (32) holds. Note that an upper bound on $\tilde{v}_\alpha(\omega, x) - \hat{v}_\alpha(\omega, x)$ is given by $R := \max_{k=1,\ldots,K} w_k$, where w_k denotes the upper bound on $\psi^k(\omega - \alpha)$ from Lemma 1 (iii). We thus have the following for the difference $\Delta(\omega, x) = \tilde{v}_\alpha(\omega, x) - \hat{v}_\alpha(\omega, x)$,

$$\Delta(\omega, x) \leq \xi(\omega, x) := \begin{cases} R, & \text{if } \omega - Tx \notin \mathcal{M}, \\ 0, & \text{if } \omega - Tx \in \mathcal{M}. \end{cases}$$

Unfortunately, $\xi(\omega, x)$ still depends on x. Therefore, we cannot apply the SLLN to

$$\sup_{x \in X} \frac{1}{S} \sum_{s=1}^S \xi(\omega^{(s)}, x)$$

if $\mathcal{X}$ is infinite. To resolve this, we use that $\mathcal{X}$ is bounded. Let D denote the diameter of TX, i.e., $\|Tx - Tx'\| \leq D$ for all $x, x' \in X$. Define $\mathcal{M}' \subset \mathcal{M}$ as $\mathcal{M}' := \bigcup_{k=1}^K (\sigma_k + \Lambda^k)(D)$. Fix an arbitrary $x^* \in X$. Note that for all $x \in X$,

$$\omega - Tx^* \in \mathcal{M}' \implies \exists k : \omega - Tx^* \in (\sigma_k + \Lambda^k)(D) \implies \omega - Tx \in (\sigma_k + \Lambda^k) \implies \omega - Tx \in \mathcal{M}.$$

We obtain

$$\Delta(\omega, x) \leq \xi(\omega) := \begin{cases} R & \text{if } \omega - Tx^* \notin \mathcal{M}' \\ 0 & \text{if } \omega - Tx^* \in \mathcal{M}' \end{cases}$$

Note that $\bar{\xi}(\omega)$ only depends on a fixed $x^* \in X$ and is independent of x. By the SLLN,

$$\frac{1}{S} \sum_{s=1}^S \bar{\xi}(\omega^{(s)}) \to R\mathbb{P}[\omega - Tx^* \notin \mathcal{M}'],$$

with probability 1, as $S \to \infty$.

By a result in Romeijnders et al. [2016a], $\mathbb{R}^m \setminus \mathcal{M}'$ can be covered by finitely many hyperslices, that is,

$$\mathbb{R}^m \setminus \mathcal{M}' \subset \bigcup_{j=1}^J H_j,$$

where the hyperslices H_j are defined as

$$H_j := \{x \in \mathbb{R}^m : 0 \leq a_j^T x \leq \delta_j\},$$

for some a_j^T and δ_j. It follows from the total variation probability bound in Romeijnders et al. [2016a] that there exists a constant $\beta > 0$ such that

$$\mathbb{P}[\omega - Tx^* \notin \mathcal{M}'] \leq \beta \sum_{i=1}^m \mathbb{E}_{\omega_{-i}} \left[|\Delta| f_i(\cdot | \omega_{-i}) \right].$$

The result now follows from

$$\sup_{x \in X} \left| \tilde{Q}_\alpha^S(x) - \tilde{\mathcal{Q}}_\alpha^S(x) \right| = \sup_{x \in X} \left\{ \tilde{Q}_\alpha^S(x) - \tilde{\mathcal{Q}}_\alpha^S(x) \right\}$$

$$= \sup_{x \in X} \frac{1}{S} \sum_{s=1}^{S} \Delta(\omega^{(s)}, x)$$

$$\leq \frac{1}{S} \sum_{s=1}^{S} \bar{\xi}(\omega^{(s)})$$

$$\to R\mathbb{P}[\omega - Tx^* \notin \mathcal{M}']$$

$$\leq C \sum_{i=1}^{m} \mathbb{E}_{\omega_{-i}} \left[|\Delta| f_i(\cdot | \omega_{-i}) \right],$$

with probability 1 as $S \to \infty$, where $C = R\beta$. $\qquad\square$

4.3 Error bound on the optimality gap of LBDA(α)

We are now ready to prove Proposition 1. We prove that the solution $\tilde{x}_\alpha$ generated by LBDA(α) is near-optimal with respect to the generalized α-approximation $\tilde{Q}_\alpha$. The error bound on $||Q - \tilde{Q}_\alpha||_\infty$ in Theorem 4 then implies that $\tilde{x}_\alpha$ is also near optimal in the original MIR problem (1). To prove that $\tilde{x}_\alpha$ is near-optimal with respect to $\tilde{Q}_\alpha$, note that $\tilde{x}_\alpha$ is optimal with respect to an outer approximation $\hat{\mathcal{Q}}$ of $\tilde{Q}_\alpha^S$, the SAA of $\tilde{Q}_\alpha$. If this outer approximation were exact at $\tilde{x}_\alpha$, that is, if $\hat{\mathcal{Q}}(\tilde{x}_\alpha) = \tilde{Q}_\alpha^S(\tilde{x}_\alpha)$, then $\tilde{x}_\alpha$ would also be optimal with respect to $\tilde{Q}_\alpha^S$. Because we used loose optimality cuts to construct $\hat{\mathcal{Q}}$, this is not necessarily the case. However, the termination criterion of LBDA(α) ensures that

$$\tilde{\mathcal{Q}}_\alpha^S(\tilde{x}_\alpha) \leq \hat{\mathcal{Q}}(\tilde{x}_\alpha) \leq \tilde{Q}_\alpha^S(\tilde{x}_\alpha).$$

Together with the upper bound on $||\tilde{Q}_\alpha^S - \tilde{\mathcal{Q}}_\alpha^S||_\infty$ in Lemma 5, this implies that $\tilde{x}_\alpha$ is near-optimal with respect to $\tilde{Q}_\alpha^S$. Finally, $\tilde{x}_\alpha$ is also near-optimal with respect to $\tilde{Q}_\alpha$, since $\tilde{Q}_\alpha^S$ converges to $\tilde{Q}_\alpha$, see Lemma 4.

Theorem 5. *Proposition 1 is correct.*

Proof. Let $\tilde{x}_\alpha$ denote the solution returned by LBDA(α), and let τ be the iteration index of the final iteration. Consider the $\tau + 1$-st master problem defined as

$$\hat{\eta}_{\tau+1} := \min_x \{ cx + \tilde{\mathcal{Q}}_{\tau+1}(x) : \quad Ax = b, \quad x \in \mathbb{R}_+^{n_1} \}, \tag{34}$$

where $\tilde{\mathcal{Q}}_{\tau+1}$ denotes the outer approximation of $\tilde{Q}_\alpha^S$ constructed during the algorithm:

$$\tilde{\mathcal{Q}}_{\tau+1}(x) := \max_{r=1,\ldots,\tau+1} \{ \beta_r x + \delta_r \}.$$

The candidate solution $\tilde{x}_\alpha$ is ε-optimal in (34), that is, $\hat{\eta}_{\tau+1} \geq c\tilde{x}_\alpha + \tilde{\mathcal{Q}}_{\tau+1}(\tilde{x}_\alpha) - \varepsilon$. To see this, note that the termination criterion implies that $\tilde{x}_\alpha$ is optimal in the following relaxation of (34):

$$\hat{\eta}_{\tau+1}^\varepsilon := \min_x \{ cx + \tilde{\mathcal{Q}}_{\tau+1}^\varepsilon(x) : \quad Ax = b, \quad x \in \mathbb{R}_+^{n_1} \}, \tag{35}$$

where

$$\tilde{\mathcal{Q}}_{\tau+1}^\varepsilon(x) := \max_{r=1,\ldots,\tau} \{ \beta_r x + \delta_r, \quad \beta_{\tau+1} x + \delta_{\tau+1} - \varepsilon \}.$$

Hence,

$$\hat{\eta}_{\tau+1} \geq \hat{\eta}_{\tau+1}^\varepsilon = c\tilde{x}_\alpha + \tilde{\mathcal{Q}}_{\tau+1}^\varepsilon(\tilde{x}_\alpha) \geq c\tilde{x}_\alpha + \tilde{\mathcal{Q}}_{\tau+1}(\tilde{x}_\alpha) - \varepsilon,$$

23

where the first inequality is because (35) is a relaxation of (34), the equality follows from optimality of $\tilde{x}_\alpha$ in (35), and the final inequality from the definitions of $\tilde{\mathcal{Q}}_{\tau+1}$ and $\tilde{\mathcal{Q}}_{\tau+1}^\varepsilon$.

Consider the following problem

$$\tilde{\eta}_{\tau+1} := \min_x \{cx + \tilde{\mathcal{Q}}_{\alpha,\tau+1}^S(x) : \quad Ax = b, \quad x \in \mathbb{R}_+^{n_1}\}, \tag{36}$$

where $\tilde{\mathcal{Q}}_{\alpha,\tau+1}^S(x) := \max\{\tilde{\mathcal{Q}}_{\tau+1}(x), \tilde{\mathcal{Q}}_\alpha^S(x)\}$. The candidate solution $\tilde{x}_\alpha$ is also ε-optimal in (36). To see this, note that

$$\tilde{\eta}_{\tau+1} \geq \hat{\eta}_{\tau+1} \geq c\tilde{x}_\alpha + \tilde{\mathcal{Q}}_{\tau+1}(\tilde{x}_\alpha) - \varepsilon = c\tilde{x}_\alpha + \tilde{\mathcal{Q}}_{\alpha,\tau+1}^S(\tilde{x}_\alpha) - \varepsilon,$$

where the latter equality follows from

$$\tilde{\mathcal{Q}}_{\tau+1}(\tilde{x}_\alpha) \geq \delta_{\tau+1} + \beta_{\tau+1}\tilde{x}_\alpha = \tilde{\mathcal{Q}}_\alpha^S(\tilde{x}_\alpha),$$

implying that $\tilde{\mathcal{Q}}_{\alpha,\tau+1}^S(\tilde{x}_\alpha) = \tilde{\mathcal{Q}}_{\tau+1}(\tilde{x}_\alpha)$.

Let x^* denote an optimal solution of (25), so that $\eta^* = cx^* + Q(x^*)$. Note that

$$\begin{aligned}
c\tilde{x}_\alpha + Q(\tilde{x}_\alpha) - \eta^* &= c\tilde{x}_\alpha + \tilde{\mathcal{Q}}_{\alpha,\tau+1}^S(\tilde{x}_\alpha) + Q(\tilde{x}_\alpha) - \tilde{\mathcal{Q}}_{\alpha,\tau+1}^S(\tilde{x}_\alpha) - \eta^* \\
&\leq \varepsilon + cx^* + \tilde{\mathcal{Q}}_{\alpha,\tau+1}^S(x^*) + Q(\tilde{x}_\alpha) - \tilde{\mathcal{Q}}_{\alpha,\tau+1}^S(\tilde{x}_\alpha) - \eta^* \\
&= \varepsilon + \tilde{\mathcal{Q}}_{\alpha,\tau+1}^S(x^*) - Q(x^*) + Q(\tilde{x}_\alpha) - \tilde{\mathcal{Q}}_{\alpha,\tau+1}^S(\tilde{x}_\alpha),
\end{aligned} \tag{37}$$

where the inequality follows from sub-optimality of x^* and ε-optimality of $\tilde{x}_\alpha$ in (36). We consider the two terms $\tilde{\mathcal{Q}}_{\alpha,\tau+1}^S(x^*) - Q(x^*)$ and $Q(\tilde{x}_\alpha) - \tilde{\mathcal{Q}}_{\alpha,\tau+1}^S(\tilde{x}_\alpha)$ separately.

For the first term note that

$$\begin{aligned}
\tilde{\mathcal{Q}}_{\alpha,\tau+1}^S(x^*) - Q(x^*) &= \tilde{\mathcal{Q}}_{\alpha,\tau+1}^S(x^*) - \tilde{Q}_\alpha^S(x^*) + \tilde{Q}_\alpha^S(x^*) - \tilde{Q}_\alpha(x^*) + \tilde{Q}_\alpha(x^*) - Q(x^*) \\
&\leq [\tilde{Q}_\alpha^S(x^*) - \tilde{Q}_\alpha(x^*)] + [\tilde{Q}_\alpha(x^*) - Q(x^*)],
\end{aligned} \tag{38}$$

where the inequality follows from $\tilde{\mathcal{Q}}_{\alpha,\tau+1}^S(x^*) \leq \tilde{Q}_\alpha^S(x^*)$. To see this, note that (i) $\tilde{\mathcal{Q}}_{\tau+1}(x^*) \leq \tilde{Q}_\alpha^S(x^*)$, as $\tilde{\mathcal{Q}}_{\tau+1}$ is an outer approximation of $\tilde{Q}_\alpha^S(x^*)$ and (ii) $\tilde{\mathcal{Q}}_\alpha^S(x^*) \leq \tilde{Q}_\alpha^S(x^*)$, by their definitions. Applying Theorem 4 and Lemma 3 to the first and second term in (38), we conclude that there exists a constant $C_1 > 0$ such that

$$\tilde{\mathcal{Q}}_{\alpha,\tau+1}^S(x^*) - Q(x^*) \leq C_1 \sum_{i=1}^m \mathbb{E}_{\omega_{-i}} \left[|\Delta| f_i(\cdot | \omega_{-i}) \right],$$

with probability 1 as $S \to \infty$.

For the second term, we have

$$\begin{aligned}
Q(\tilde{x}_\alpha) - \tilde{\mathcal{Q}}_{\alpha,\tau+1}^S(\tilde{x}_\alpha) &\leq Q(\tilde{x}_\alpha) - \tilde{\mathcal{Q}}_\alpha^S(\tilde{x}_\alpha) \\
&= [Q(\tilde{x}_\alpha) - \tilde{Q}_\alpha(\tilde{x}_\alpha)] + [\tilde{Q}_\alpha(\tilde{x}_\alpha) - \tilde{Q}_\alpha^S(\tilde{x}_\alpha)] + [\tilde{Q}_\alpha^S(\tilde{x}_\alpha) - \tilde{\mathcal{Q}}_\alpha^S(\tilde{x}_\alpha)]. \tag{39}
\end{aligned}$$

Applying Theorem 4, and Lemmas 4 and 5 to the first, second, and third term in (39), we conclude that there exists a constant $C_2 > 0$ such that

$$Q(\tilde{x}_\alpha) - \tilde{\mathcal{Q}}_{\alpha,\tau+1}^S(\tilde{x}_\alpha) \leq C_2 \sum_{i=1}^m \mathbb{E}_{\omega_{-i}} \left[|\Delta| f_i(\cdot | \omega_{-i}) \right],$$

with probability 1 as $S \to \infty$.

Proposition 1 now follows from (37) by setting $C = C_1 + C_2$. $\qquad \square$

Proposition 1 states a theoretical bound on the optimality gap of the solution returned by LBDA(α). Moreover, as the total variations of the one-dimensional conditional pdf of the random vector ω go to zero, the optimality gap goes to zero. However, the error bound in Proposition 1 is a worst-case bound. For many problem instances, the actual performance may be much better. In Section 5, we assess the performance of LBDA(α) empirically using Monte Carlo sampling.

5 Numerical experiments

We test the performance of LBDA(α) on two types of problem instances. The first are based on a nurse scheduling problem of Kim and Mehrotra [2015], and the second are randomly generated test instances. The results are discussed in Sections 5.2 and 5.3, respectively. First, however, we describe the set-up of our numerical experiments in Section 5.1.

5.1 Set-up numerical experiments

In our numerical experiments, we compare seven candidate solutions, described below, in terms of costs, relative optimality gaps, and computation times. The expected costs $cx + Q(x)$ of a candidate solution x are estimated using out-of-sample estimation with a sample size of 10^5, which is chosen such that the standard errors are sufficiently small. The relative optimality gaps are determined using the multiple replications procedure (MRP) by Mak et al. [1999] with Latin hypercube sampling [Bayraksan and Morton, 2006]. Using the MRP, we obtain an upper bound on the relative optimality gap

$$\frac{cx + Q(x) - \eta^*}{\eta^*} \times 100\%$$

of a candidate solution x, where η^* denotes the optimal value of the true MIR problem in (1). The upper bound generated by the MRP has a confidence level equal to $1 - \gamma$. We use $\gamma = 0.05$ in our experiments. For both types of problems, we construct small and large problem instances. Because the MRP is not tractable for the large problem instances, we use it only for the small ones. For the large instances, we compare the candidate solutions in terms of out-of-sample estimated expected costs.

The seven candidate solutions we obtain are generated using a sample $\tilde{\omega} = \{\omega^{(1)}, \ldots, \omega^{(S)}\}$ of size $S = 1000$ from the distribution of ω. To ensure a fair comparison, we use common random numbers where possible and we limit the computation time of each algorithm to two hours. If a candidate solution could not be computed within two hours, we report out of time (OOT).

First, we consider the solution generated by LBDA(α), denoted $\tilde{x}_\alpha$, where $\alpha = 0$. Second, we solve the α-approximations exactly, that is, we find the optimal solution x_α^* of the approximating problem (4), with $\hat{Q} = \tilde{Q}_\alpha$, where again $\alpha = 0$. We do so by solving the approximating second-stage problems

$$\max_{k=1,\ldots,K} \{\lambda^k(\omega - Tx) + \psi^k(\omega - \alpha)\}$$

by enumeration over $k = 1, \ldots, K$. For this reason x_α^* can only be computed in reasonable time for small problem instances.

Third, we apply LBDA(α) multiple times using 100 different values of α, drawn from a multivariate uniform distribution on $[0, 100]^m$. We then select the best candidate solution, denoted $\tilde{x}_\alpha^+$, using out-of-sample evaluation with a sample size 10^4. Note that this procedure can be efficiently parallelized, by running LBDA(α) in parallel. For this reason, we report the maximum computation time of LBDA(α) over all 100 values of α.

The fourth candidate solution $\hat{x}_S$ is obtained by solving the deterministic equivalent formulation (DEF), in which the distribution of ω is replaced by the empirical distribution of the sample $\tilde{\omega}$:

$$\min_x \left\{ cx + \hat{Q}^S(x) : \quad Ax = b, \quad x \in \mathbb{R}_+^{n_1} \right\},$$

where

$$\hat{Q}^S(x) = \frac{1}{S} \sum_{s=1}^{S} \min_y \left\{ qy : \quad Wy = \omega^{(s)} - Tx, \quad y \in \mathbb{Z}_+^{p_2} \times \mathbb{R}_+^{n_2 - p_2} \right\}.$$

The DEF is a large-scale MIP, which, typically, cannot be solved in reasonable time by standard MIP solvers. Hence, we also solve the DEF using a smaller sample size $S' = 100$. We denote the resulting candidate solution by $\hat{x}_{S'}$.

Finally, we consider two trivial benchmark solutions, which we expect LBDA(α) to outperform significantly. First, we solve the LP-relaxation of the DEF to obtain the candidate solution x_{LP}. Second, we consider the Jensen approximation, which replaces the distribution of ω by a degenerate distribution at $\mu = \mathbb{E}_\omega[\omega]$. The resulting problem is a small scale MIP, which can be solved efficiently using standard MIP solvers. We denote its optimal solution by x_μ.

For easy reference, we summarize the definitions of the candidate solutions in Table 2.

Table 2: Definitions candidate solutions.

Solution	Definition
$\tilde{x}_\alpha$	LBDA(α)
x_α^*	Optimal solution α-approximations
$\tilde{x}_\alpha^+$	Best out of 100 LBDA(α) runs for different α's
$\hat{x}_S$	Optimal solution DEF
$\hat{x}_{S'}$	Optimal solution DEF (small sample size)
x_{LP}	Optimal solution LP-relaxation DEF
x_μ	Optimal solution Jensen approximation

5.2 Nurse scheduling

We consider an adapted version of the integrated staffing and scheduling problem for nurses introduced by Kim and Mehrotra [2015]. In this problem, nurses are scheduled according to a set of predefined shifts. These are the first-stage decisions. In the second stage, demand for nurses is revealed. If there are insufficient nurses, additional nurses can be scheduled according to a set of adjustment shifts. Additionally, there are penalties for over- and understaffing. Kim and Mehrotra [2015] formulate this problem as a two-stage mixed-integer recourse model with a TU recourse matrix.

Let T denote the length of the planning horizon, e.g. $T = 24$, where each time period represents one hour. Let I denote the set of shifts available during the planning horizon. For every shift $i \in I$, let $a_{it} = 1$ if i contains hour t, and 0 otherwise, $t = 1, \ldots, T$. The first-stage decision x_i is the number of nurses working according to shift i. The number of working nurses at time t, denoted z_t, is given by

$$z_t = \sum_{i \in I} a_{it} x_i.$$

For the second-stage problem, let ω_t denote random demand in period t. Let J denote the set of adjustment shifts available to cover the difference between demand and supply of nurses $\omega_t - z_t$. Any adjustment shift $j \in J$ can be added or cancelled, which comes at cost q_j^+ and q_j^-, respectively. Let y_j^+ and y_j^- denote the number of added and cancelled adjustment shifts of type j, respectively. Analogous to a_{it}, let $w_{jt} = 1$ if adjustment shift j contains hour t, and 0 otherwise. Let u_t and v_t denote the amount of over- and understaffing at hour t. We then face for each time period t the constraint

$$\sum_{j \in J} w_{jt}(y_j^+ - y_j^-) - u_t + v_t = \omega_t - z_t. \tag{40}$$

Over- and understaffing comes at penalty costs r_u and r_v, respectively. This leads to the two-stage mixed-integer recourse problem

$$\min_x \left\{ \sum_{i \in I} c_i x_i + \mathbb{E}_\omega[v(\omega, z)] : \quad z_t = \sum_{i \in I} a_{it} x_i, \quad x_i \in \mathbb{Z}_+ \right\},$$

where

$$v(\omega, z) = \min \left\{ \sum_{t=1}^{T}(r_u u_t + r_v v_t) + \sum_{j \in J} q_j^+ y_j^+ + q_j^- y_j^- : \right.$$

$$\sum_{j \in J} w_{jt}(y_j^+ - y_j^-) + u_t - v_t = \omega_t - z_t, \quad t = 1, \ldots, T,$$

$$\left. y_j^+, y_j^- \in \mathbb{Z}_+, \quad u_t, v_t \in \mathbb{R}_+ \right\}.$$

The main difference between our version of the problem and that of Kim and Mehrotra [2015] is that we relax the integrality restrictions on x. Furthermore, we set $r_u = 0$ and $r_v = +\infty$, so that the second-stage constraints in (40) reduce to

$$\sum_{j \in J} w_{jt}(y_j^+ - y_j^-) \geq \omega_t - z_t.$$

In our test instances, we set $T = 8$ and $T = 48$, corresponding to a small and large version of the problem, respectively. The set of first-stage scheduling patterns I consists of all consecutive three-hour shifts. It follows that the number of first stage-decisions is given by $|I| = T - 2$. The set of second-stage adjustment patterns J consists of non-overlapping consecutive four hour shifts, so that the number of integer second stage variables equals $|J| = T/4$. The cost parameters are given by $c_i = 1$ for all $i \in I$, and $q_j^+ = 5$ and $q_j^- = 0$ for all $j \in J$. The random demand for nurses in period t, denoted ω_t, $t = 1, \ldots, T$, follow independent normal distributions with mean 10 and standard deviation σ, where $\sigma \in \{0.1, 0.5, 1, 2, 4, 10\}$.

The reason that we consider multiple values of σ is that Proposition 1 informs us that LBDA(α) performs well if the variability in the model is large. Here, variability can be measured by σ, see Example 1. Indeed, the total variation of the one-dimensional conditional pdf of ω goes to zero as σ increases and thus the error bound on the optimality gap goes to zero as well. We expect to see this reflected in the results.

In Tables 3 and 4, we provide the results of our numerical experiments for the small and large nurse scheduling test instances, respectively. We report OOT if an instance could not be solved in under two hours.

Table 3: Nurse scheduling: $T = 8$.

	Computation time (seconds)							Relative optimality gap (%)						
σ	$\tilde{x}_\alpha$	$\tilde{x}_\alpha^+$	x_α^*	$\hat{x}_S$	$\hat{x}_{S'}$	x_μ	x_{LP}	$\tilde{x}_\alpha$	$\tilde{x}_\alpha^+$	x_α^*	$\hat{x}_S$	$\hat{x}_{S'}$	x_μ	x_{LP}
0.1	7.8	10.5	39.7	321.11	0.1	0.0	0.1	7.05	1.77	7.14	**0.56**	1.03	26.87	8.76
0.5	8.6	12.6	38.1	OOT	0.4	0.0	0.1	1.34	1.28	**0.56**	OOT	1.19	19.00	4.60
1.0	10.3	12.8	40.4	OOT	3.7	0.0	0.1	1.94	**1.19**	1.38	OOT	1.48	22.75	2.70
2.0	10.0	13.8	43.5	OOT	3.4	0.0	0.1	1.43	1.22	**0.87**	OOT	1.29	30.56	1.79
4.0	12.9	16.3	44.7	OOT	3.7	0.0	0.1	1.32	**0.98**	1.13	OOT	1.67	41.72	1.41
10.0	16.6	18.4	49.7	OOT	10.9	0.0	0.1	0.96	0.82	**0.79**	OOT	1.67	60.30	0.91

We make several observations based on Table 3. First, LBDA(α) is able to solve every test instance within a minute, whereas the sampling solution $\hat{x}_S$ cannot be computed within two hours except for $\sigma = 0.1$. Hence, LBDA(α) is much faster than solving the DEF. The performance of LBDA(α) is also very good. In particular, the relative optimality gap of $\tilde{x}_\alpha^+$ is always below 2%, and $\tilde{x}_\alpha^+$ outperforms both sampling solutions for $\sigma \geq 1.0$. For smaller values of σ, performing multiple LBDA(α) runs leads to significant improvements compared to using a single value of α. In particular, for $\sigma = 0.1$, the difference in optimality gap between $\tilde{x}_\alpha$ and $\tilde{x}_\alpha^+$ is large.

We observe that the performance of x_α^* is not the same as that of the LBDA(α)-solution $\tilde{x}_\alpha$. In fact, the relative optimality gaps corresponding to x_α^* are slightly smaller. This is not surprising

since LBDA(α) is used to obtain an approximation of x_α^*. However, by running LBDA(α) multiple times we obtain a performance similar to x_α^*. Moreover, we will see in Section 5.3 that x_α^* does not necessarily outperform $\tilde{x}_\alpha$.

Finally, we observe that x_{LP} and x_μ have the largest optimality gaps. However, x_{LP} performs better than expected with optimality gaps below 5% for $\sigma \geq 0.5$. We suspect that this is because of the TU structure in this nurse scheduling problem. In the next section we confirm our suspicion by showing that x_{LP} performs worse for randomly generated problem instances.

Table 4: Nurse scheduling: $T = 48$.

	Computation time (seconds)						Costs (lowest costs normalized to 100)					
σ	$\tilde{x}_\alpha$	$\tilde{x}_\alpha^+$	$\hat{x}_S$	$\hat{x}_{S'}$	x_μ	x_{LP}	$\tilde{x}_\alpha$	$\tilde{x}_\alpha^+$	$\hat{x}_S$	$\hat{x}_{S'}$	x_μ	x_{LP}
0.1	496.6	501.8	OOT	5.9	0.0	0.7	104.90	102.17	OOT	**100**	130.10	106.99
0.5	685.9	742.7	OOT	96.7	0.0	0.6	101.26	100.99	OOT	**100**	121.51	102.72
1.0	719.4	844.7	OOT	892.3	0.0	0.6	100.34	100.46	OOT	**100**	126.66	100.92
2.0	801.9	929.8	OOT	OOT	0.0	0.6	100.33	**100**	OOT	OOT	136.61	100.10
4.0	1053.7	1140.3	OOT	OOT	0.0	0.6	100.10	**100**	OOT	OOT	151.08	100.05
10.0	973.9	1158.5	OOT	OOT	0.0	0.6	100.00	**100**	OOT	OOT	171.90	100.01

Table 4 displays the results for the large nurse scheduling test instances. Again, we observe a clear difference in computation times between LBDA(α) and the sampling solutions. The sampling solutions can only be computed with a small sample size $S' = 100$ for $\sigma \leq 1.0$, whereas LBDA(α) solves every instance within 20 minutes. Also note the dramatic increase in computation times for x'_S compared to the small instances, especially for $\sigma \geq 2.0$. In comparison, the computation times of LBDA(α) scale favourably in the size of the instance, measured by the number of times periods T.

In terms of performance, we observe that both LBDA(α)-solutions $\tilde{x}_\alpha$ and $\tilde{x}_\alpha^+$ outperform both sampling solutions if $\sigma \geq 2.0$. In addition, $\tilde{x}_\alpha^+$ has the lowest expected costs out of all candidate solutions for these values of σ. Furthermore, $\tilde{x}_\alpha^+$ performs significantly better than $\tilde{x}_\alpha$ for $\sigma = 0.1$ and $\sigma = 0.5$. In other words, there is a significant benefit to running LBDA(α) using multiple values of α for small values of σ, similar as for the small nurse scheduling test instances.

5.3 Randomly generated test instances

We generate random MIR problems with the following structure

$$\min_x \{cx + Q(x) : x \in \mathbb{R}_+^{n_1}\},$$

where

$$Q(x) := \mathbb{E}_\omega \left[\min_y \{qy : \quad Wy \geq \omega - Tx, \quad y \in \mathbb{Z}_+^p\} \right].$$

Here, $\omega \in \mathbb{R}^m$ is a random vector, whose elements follow independent normal distributions with mean 10 and standard deviation $\sigma \in \{0.1, 0.5, 1, 2, 4, 10\}$. The parameters c, q, T, and W are fixed, and their elements are drawn from discrete uniform distributions whose supports can be found in Table 5.

Table 5: Random instances: parameter supports.

Parameter	Support
c	$\{1,\ldots,5\}^{n_1}$
q	$\{5,\ldots,10\}^{p}$
T	$\{1,\ldots,6\}^{m \times n_1}$
W	$\{1,\ldots,6\}^{m \times n_2}$

To prevent noise in the outcomes of our experiments, we compute the average optimality gaps, costs, and computation times over 20 randomly generated test instances for each value of σ. We report OOT if at least one of these instances could not be solved in under two hours.

We consider test instances of two different sizes, namely $n_1 = 10$, $p = 5$, $m = 5$ (small), and $n_1 = 100$, $p = 40$, $m = 20$ (large). Tables 6 and 7 display the results for the small and large versions, respectively.

Table 6: Randomly generated test instances ($n_1 = 10$, $p = 5$, $m = 5$).

	Running time (seconds)							Relative optimality gap (%)						
σ	$\tilde{x}_\alpha$	$\tilde{x}_\alpha^+$	x_α^*	$\hat{x}_S$	$\hat{x}_{S'}$	x_μ	x_{LP}	$\tilde{x}_\alpha$	$\tilde{x}_\alpha^+$	x_α^*	$\hat{x}_S$	$\hat{x}_{S'}$	x_μ	x_{LP}
0.1	2.6	5.8	93.9	447.5	0.1	0.0	0.3	24.45	3.20	28.34	**0.34**	2.56	102.55	36.83
0.5	4.5	8.3	87.9	OOT	0.3	0.0	0.2	12.13	**1.72**	16.98	OOT	2.26	84.83	28.06
1.0	3.9	6.9	91.7	OOT	0.7	0.0	0.2	8.98	**2.28**	10.20	OOT	2.83	71.88	21.24
2.0	4.1	7.5	91.1	OOT	OOT	0.0	0.2	7.66	**3.60**	5.57	OOT	OOT	60.01	13.76
4.0	3.8	6.4	87.0	OOT	OOT	0.0	0.2	6.73	**4.18**	4.82	OOT	OOT	65.22	8.35
10.0	3.6	7.0	83.8	OOT	OOT	0.0	0.2	3.74	**3.24**	3.32	OOT	OOT	87.59	4.23

Table 7: Randomly generated test instances ($n_1 = 100$, $p = 40$, $m = 20$).

	Running time (seconds)						Costs (lowest costs normalized to 100)					
σ	$\tilde{x}_\alpha$	$\tilde{x}_\alpha^+$	$\hat{x}_S$	$\hat{x}_{S'}$	x_μ	x_{LP}	$\tilde{x}_\alpha$	$\tilde{x}_\alpha^+$	$\hat{x}_S$	$\hat{x}_{S'}$	x_μ	x_{LP}
0.1	35.7	183.5	1174.4	2.0	0.0	21.4	131.66	108.18	**100**	112.96	242.78	142.93
0.5	34.5	59.3	OOT	4.9	0.0	21.5	116.19	**100**	OOT	104.36	205.78	126.68
1.0	32.9	41.3	OOT	6.0	0.0	21.3	107.06	**100**	OOT	103.74	185.38	120.18
2.0	32.6	38.3	OOT	54.0	0.0	22.1	103.24	**100**	OOT	102.80	180.51	111.66
4.0	31.3	36.5	OOT	184.6	0.0	22.1	101.66	**100**	OOT	102.52	202.42	104.76
10.0	29.9	36.7	OOT	OOT	0.0	43.9	100.66	**100**	OOT	OOT	231.82	100.94

From Tables 6 and 7, we observe that LBDA(α) clearly outperforms the sampling solutions in terms of computation time and scalability to larger problem instances, similar as in the nurse scheduling test instances in Section 5.2. In particular, we observe that the computation time of LBDA(α) is of the same order of magnitude as that of x_{LP}, while it performs significantly better in terms of optimality gaps and out-of-sample estimated expected costs. Strikingly, for the large problem instances with $\sigma = 10.0$, LBDA(α) is faster than solving the LP-relaxation of the DEF. This is explained by the fact that LBDA(α) is an augmented version of the L-shaped algorithm, which is designed to efficiently compute x_{LP}. Undeniably, our results indicate that LBDA(α) can be implemented very efficiently and that we can handle large MIR problem instances.

In terms of performance, the solution generated by multiple LBDA(α) runs outperforms every other candidate solution for $\sigma \geq 0.5$. Only if the variability of the random parameters in the model is very small ($\sigma = 0.1$), then LBDA(α) is outperformed by the sampling solutions. This is

in line with Proposition 1, which states that the solution generated by LBDA(α) performs better if the variability of the random parameters in the model is large. Similar to the nurse scheduling test instances, we observe that especially for small values of σ, $\tilde{x}_\alpha^+$ significantly outperforms $\tilde{x}_\alpha$. This implies that there is a considerable benefit to using multiple values of α in LBDA(α) instead of just one.

Compared to the nurse scheduling test instances in Section 5.2, the optimality gaps of $\tilde{x}_\alpha^+$ in Table 6 are relatively large, though all well below 5%. Moreover, the optimality gaps do not decrease as σ increases, as we would expect based on Proposition 1. This is explained by the fact that the MRP relies on solving multiple DEF's in order to derive sharp bounds on the optimality gap. As is clear from our results, standard solvers have difficulties with solving DEF's in reasonable time, especially for larger values of σ. Because we stop the solver after one hour in the MRP, the DEF's are not solved to optimality. As a result, the bounds on the optimality gap generated by the MRP are not sharp. Therefore, we expect that the true optimality gaps of $\tilde{x}_\alpha^+$ for $\sigma \geq 0.5$ are much smaller than the numbers we present in Table 6.

In Section 5.2, we observed that for the nurse scheduling test problems, x_{LP} performed relatively well in terms of optimality gaps. The results in this section support the explanation that this is due to the TU property of the nurse scheduling test problems. Indeed, the models in this section do not have this property, and x_{LP} performs much worse.

Finally, we compare the exact solution x_α^* of the generalized α-approximations and the solution $\tilde{x}_\alpha$ generated by LBDA(α). Recall that LBDA(α) approximates x_α^*, which itself approximates the optimal solution of the true problem. However, we observe that neither solution clearly outperforms the other. In other words, the additional approximation step made by LBDA(α) has no negative effect on the solution quality.

6   Conclusion

We consider two-stage mixed-integer recourse models with random right-hand side. Due to non-convexity of the recourse function, such models are extremely difficult to solve. We develop a tractable approximating model by using convex approximations of the recourse function. In particular, we propose a new class of convex approximations, the so-called generalized α-approximations, and we derive a corresponding error bound on the difference between these approximations and the true recourse function. In addition, we show that this error bound is small if the variability of the random parameters in the model is large. More precisely, the error bound for the generalized α-approximations goes to zero as the total variations of the one-dimensional conditional probability density functions of the random right-hand side vector in the model go to zero.

The advantage of the generalized α-approximations over existing convex approximations is that it can be solved efficiently. In fact, we describe a loose Benders decomposition algorithm, LBDA(α), which efficiently solves the corresponding approximating model. The quality of the candidate solution $\hat{x}$ generated by LBDA(α) in the original model is guaranteed by Proposition 1, which states an upper bound on the optimality gap of $\hat{x}$. This performance guarantee is similar to the error bound we prove for the generalized α-approximations. Indeed, we show that the optimality gap of $\hat{x}$ is small if the variability of the random parameters in the model is large.

In addition to this theoretical guarantee on the solution quality, we use Monte Carlo sampling to assess LBDA(α) empirically. In our numerical experiments, we consider test instances based on a nurse scheduling application and randomly generated test instances. We find that LBDA(α) performs well in terms of computation times, scalability to larger problem instances, and solution quality. In particular, LBDA(α) is able to solve larger instances than traditional sampling techniques and its computation times scale more favourably in the input size of the instances. Moreover, LBDA(α) outperforms traditional sampling techniques in terms of optimality gaps and total expected costs if the variability of the random parameters in the model is medium to large.

One avenue for future research is to derive sharper theoretical error bounds for the generalized α-approximations. While Proposition 1 provides conditions under which our solution method

performs well, the quantitative error bound cannot be computed, as it depends on an unknown and potentially large constant C. A sharp tractable error bound would be an improvement over our current results. Another avenue is the extension of our solution method to more general mixed-integer recourse models, for example by allowing for randomness in the second-stage cost coefficients q, technology matrix T, or recourse matrix W.

Acknowledgements

References

S. Ahmed and A. Shapiro. The sample average approximation method for stochastic programs with integer recourse. *Preprint available from www.optimization-online.org*, 2002.

M. Bansal, K.-L. Huang, and S. Mehrotra. Tight second stage formulations in two-stage stochastic mixed integer programs. *SIAM Journal on Optimization*, 28(1):788–819, 2018.

G. Bayraksan and D. P. Morton. Assessing solution quality in stochastic programs. *Mathematical Programming*, 108(2-3):495–514, 2006.

J. F. Benders. Partitioning procedures for solving mixed-variables programming problems. *Numerische mathematik*, 4(1):238–252, 1962.

J. R. Birge and F. V. Louveaux. A multicut algorithm for two-stage stochastic linear programs. *European Journal of Operational Research*, 34(3):384–392, 1988.

C. E. Blair and R. G. Jeroslow. The value function of a mixed integer program II. *Discrete Mathematics*, 25(1):7–19, 1979.

C. C. Carøe and R. Schultz. Dual decomposition in stochastic integer programming. *Operations Research Letters*, 24(1-2):37–45, 1999.

D. Gade, S. Küçükyavuz, and S. Sen. Decomposition algorithms with parametric Gomory cuts for two-stage stochastic integer programs. *Mathematical Programming*, 144(1-2):39–64, 2014.

H. I. Gassmann and W. T. Ziemba, editors. *Stochastic Programming: Applications in Finance, Energy, Planning and Logistics*, volume 4 of *World Scientific Series in Finance*. World Scientific, 2013.

R. E. Gomory. Outline of an algorithm for integer solutions to linear programs. *Bulletin of the American Mathematical Society*, 64(5):275–278, 1958.

K. Kim and S. Mehrotra. A two-stage stochastic integer programming approach to integrated staffing and scheduling with application to nurse management. *Operations Research*, 63(6): 1431–1451, 2015.

W. K. Klein Haneveld, L. Stougie, and M. H. van der Vlerk. Simple integer recourse models: convexity and convex approximations. *Mathematical Programming*, 108(2-3):435–473, 2006.

A. J. Kleywegt, A. Shapiro, and T. Homem-de Mello. The sample average approximation method for stochastic discrete optimization. *SIAM Journal on Optimization*, 12(2):479–502, 2002.

N. van der Laan, W. Romeijnders, and M. H. van der Vlerk. Higher-order total variation bounds for expectations of periodic functions and simple integer recourse approximations. *Computational Management Science*, 15(3-4):325–349, 2018.

G. Laporte and F. V. Louveaux. The integer L-shaped method for stochastic integer programs with complete recourse. *Operations Research Letters*, 13(3):133–142, 1993.

F. V. Louveaux and M. H. van der Vlerk. Stochastic programming with simple integer recourse. *Mathematical programming*, 61(1-3):301–325, 1993.

W. K. Mak, D. P. Morton, and R. K. Wood. Monte Carlo bounding techniques for determining solution quality in stochastic programs. *Operations Research Letters*, 24(1-2):47–56, 1999.

L. Ntaimo. Fenchel decomposition for stochastic mixed-integer programming. *Journal of Global Optimization*, 55(1):141–163, 2013.

A. H. G. Rinnooy Kan and L. Stougie. Stochastic integer programming. In Y. Ermoliev and R. J.-B. Wets, editors, *Numerical Techniques for Stochastic Optimization*, volume 10 of *Springer Series in Computational Mathematics*, pages 201–213. Springer, 1988.

W. Romeijnders, L. Stougie, and M. H. van der Vlerk. Approximation in two-stage stochastic integer programming. *Surveys in Operations Research and Management Science*, 19(1):17–33, 2014.

W. Romeijnders, M. H. van der Vlerk, and W. K. Klein Haneveld. Convex approximations of totally unimodular integer recourse models: A uniform error bound. *SIAM Journal on Optimization*, 25(1):130–158, 2015.

W. Romeijnders, R. Schultz, M. H. van der Vlerk, and W. K. Klein Haneveld. A convex approximation for two-stage mixed-integer recourse models with a uniform error bound. *SIAM Journal on Optimization*, 26(1):426–447, 2016a.

W. Romeijnders, M. H. van der Vlerk, and W. K. Klein Haneveld. Total variation bounds on the expectation of periodic functions with applications to recourse approximations. *Mathematical Programming*, 157(1):3–46, 2016b.

W. Romeijnders, D. P. Morton, and M. H. van der Vlerk. Assessing the quality of convex approximations for two-stage totally unimodular integer recourse models. *INFORMS Journal on Computing*, 29(2):211–231, 2017.

A. Ruszczyński. A regularized decomposition method for minimizing a sum of polyhedral functions. *Mathematical programming*, 35(3):309–333, 1986.

R. Schultz. Stochastic programming with integer variables. *Mathematical Programming*, 97(1-2): 285–309, 2003.

R. Schultz, L. Stougie, and M. H. van der Vlerk. Solving stochastic programs with integer recourse by enumeration: A framework using Gröbner basis reductions. *Mathematical Programming*, 83 (1-3):229–252, 1998.

S. Sen. Algorithms for stochastic mixed-integer programming models. *Handbooks in Operations Research and Management Science*, 12:515–558, 2005.

S. Sen and J. L. Higle. The C^3 theorem and a D^2 algorithm for large scale stochastic mixed-integer programming: Set convexification. *Mathematical Programming*, 104(1):1–20, 2005.

R. M. Van Slyke and R. Wets. L-shaped linear programs with applications to optimal control and stochastic programming. *SIAM Journal on Applied Mathematics*, 17(4):638–663, 1969.

M. H. van der Vlerk. *Stochastic Programming with Integer Recourse*. PhD thesis, University of Groningen, The Netherlands, 1995.

M. H. van der Vlerk. Convex approximations for complete integer recourse models. *Mathematical Programming*, 99(2):297–310, 2004.

D. W. Walkup and R. J.-B. Wets. Lifting projections of convex polyhedra. *Pacific Journal of Mathematics*, 28(2):465–475, 1969.

S. W. Wallace and W. T. Ziemba, editors. *Applications of Stochastic Programming*. MPS-SIAM Series on Optimization. 2005.

M. Zhang and S. Küçükyavuz. Finitely convergent decomposition algorithms for two-stage stochastic pure integer programs. *SIAM Journal on Optimization*, 24(4):1933–1951, 2014.

List of research reports

13001-EEF: Kuper, G.H. and M. Mulder, Cross-border infrastructure constraints, regulatory measures and economic integration of the Dutch – German gas market

13002-EEF: Klein Goldewijk, G.M. and J.P.A.M. Jacobs, The relation between stature and long bone length in the Roman Empire

13003-EEF: Mulder, M. and L. Schoonbeek, Decomposing changes in competition in the Dutch electricity market through the Residual Supply Index

13004-EEF: Kuper, G.H. and M. Mulder, Cross-border constraints, institutional changes and integration of the Dutch – German gas market

13005-EEF: Wiese, R., Do political or economic factors drive healthcare financing privatisations? Empirical evidence from OECD countries

13006-EEF: Elhorst, J.P., P. Heijnen, A. Samarina and J.P.A.M. Jacobs, State transfers at different moments in time: A spatial probit approach

13007-EEF: Mierau, J.O., The activity and lethality of militant groups: Ideology, capacity, and environment

13008-EEF: Dijkstra, P.T., M.A. Haan and M. Mulder, The effect of industry structure and yardstick design on strategic behavior with yardstick competition: an experimental study

13009-GEM: Hoorn, A.A.J. van, Values of financial services professionals and the global financial crisis as a crisis of ethics

13010-EEF: Boonman, T.M., Sovereign defaults, business cycles and economic growth in Latin America, 1870-2012

13011-EEF: He, X., J.P.A.M Jacobs, G.H. Kuper and J.E. Ligthart, On the impact of the global financial crisis on the euro area

13012-GEM: Hoorn, A.A.J. van, Generational shifts in managerial values and the coming of a global business culture

13013-EEF: Samarina, A. and J.E. Sturm, Factors leading to inflation targeting – The impact of adoption

13014-EEF: Allers, M.A. and E. Merkus, Soft budget constraint but no moral hazard? The Dutch local government bailout puzzle

13015-GEM: Hoorn, A.A.J. van, Trust and management: Explaining cross-national differences in work autonomy

13016-EEF: Boonman, T.M., J.P.A.M. Jacobs and G.H. Kuper, Sovereign debt crises in Latin America: A market pressure approach

13017-GEM: Oosterhaven, J., M.C. Bouwmeester and M. Nozaki, The impact of production and infrastructure shocks: A non-linear input-output programming approach, tested on an hypothetical economy

13018-EEF: Cavapozzi, D., W. Han and R. Miniaci, Alternative weighting structures for multidimensional poverty assessment

14001-OPERA: Germs, R. and N.D. van Foreest, Optimal control of production-inventory systems with constant and compound poisson demand

14002-EEF: Bao, T. and J. Duffy, Adaptive vs. eductive learning: Theory and evidence

14003-OPERA: Syntetos, A.A. and R.H. Teunter, On the calculation of safety stocks

14004-EEF: Bouwmeester, M.C., J. Oosterhaven and J.M. Rueda-Cantuche, Measuring the EU value added embodied in EU foreign exports by consolidating 27 national supply and use tables for 2000-2007

14005-OPERA: Prak, D.R.J., R.H. Teunter and J. Riezebos, Periodic review and continuous ordering

14006-EEF: Reijnders, L.S.M., The college gender gap reversal: Insights from a life-cycle perspective

14007-EEF: Reijnders, L.S.M., Child care subsidies with endogenous education and fertility

14008-EEF: Otter, P.W., J.P.A.M. Jacobs and A.H.J. den Reijer, A criterion for the number of factors in a data-rich environment

14009-EEF: Mierau, J.O. and E. Suari Andreu, Fiscal rules and government size in the European Union

14010-EEF: Dijkstra, P.T., M.A. Haan and M. Mulder, Industry structure and collusion with uniform yardstick competition: theory and experiments

14011-EEF: Huizingh, E. and M. Mulder, Effectiveness of regulatory interventions on firm behavior: a randomized field experiment with e-commerce firms

14012-GEM: Bressand, A., Proving the old spell wrong: New African hydrocarbon producers and the 'resource curse'

14013-EEF: Dijkstra P.T., Price leadership and unequal market sharing: Collusion in experimental markets

14014-EEF: Angelini, V., M. Bertoni, and L. Corazzini, Unpacking the determinants of life satisfaction: A survey experiment

14015-EEF: Heijdra, B.J., J.O. Mierau, and T. Trimborn, Stimulating annuity markets

14016-GEM: Bezemer, D., M. Grydaki, and L. Zhang, Is financial development bad for growth?

14017-EEF: De Cao, E. and C. Lutz, Sensitive survey questions: measuring attitudes regarding female circumcision through a list experiment

14018-EEF: De Cao, E., The height production function from birth to maturity

14019-EEF: Allers, M.A. and J.B. Geertsema, The effects of local government amalgamation on public spending and service levels. Evidence from 15 years of municipal boundary reform

14020-EEF: Kuper, G.H. and J.H. Veurink, Central bank independence and political pressure in the Greenspan era

14021-GEM: Samarina, A. and D. Bezemer, Do Capital Flows Change Domestic Credit Allocation?

14022-EEF: Soetevent, A.R. and L. Zhou, Loss Modification Incentives for Insurers Under ExpectedUtility and Loss Aversion
14023-EEF: Allers, M.A. and W. Vermeulen, Fiscal Equalization, Capitalization and the Flypaper Effect.

14024-GEM: Hoorn, A.A.J. van, Trust, Workplace Organization, and Comparative Economic Development.

14025-GEM: Bezemer, D., and L. Zhang, From Boom to Bust in de Credit Cycle: The Role of Mortgage Credit.

14026-GEM: Zhang, L., and D. Bezemer, How the Credit Cycle Affects Growth: The Role of Bank Balance Sheets.

14027-EEF: Bružikas, T., and A.R. Soetevent, Detailed Data and Changes in Market Structure: The Move to Unmanned Gasoline Service Stations.

14028-EEF: Bouwmeester, M.C., and B. Scholtens, Cross-border Spillovers from European Gas Infrastructure Investments.

14029-EEF: Lestano, and G.H. Kuper, Correlation Dynamics in East Asian Financial Markets.

14030-GEM: Bezemer, D.J., and M. Grydaki, Nonfinancial Sectors Debt and the U.S. Great Moderation.

14031-EEF: Hermes, N., and R. Lensink, Financial Liberalization and Capital Flight: Evidence from the African Continent.

14032-OPERA: Blok, C. de, A. Seepma, I. Roukema, D.P. van Donk, B. Keulen, and R. Otte, Digitalisering in Strafrechtketens: Ervaringen in Denemarken, Engeland, Oostenrijk en Estland vanuit een Supply Chain Perspectief.

14033-OPERA: Olde Keizer, M.C.A., and R.H. Teunter, Opportunistic condition-based maintenance and aperiodic inspections for a two-unit series system.

14034-EEF: Kuper, G.H., G. Sierksma, and F.C.R. Spieksma, Using Tennis Rankings to Predict Performance in Upcoming Tournaments

15001-EEF: Bao, T., X. Tian, X. Yu, Dictator Game with Indivisibility of Money

15002-GEM: Chen, Q., E. Dietzenbacher, and B. Los, The Effects of Ageing and Urbanization on China's Future Population and Labor Force

15003-EEF: Allers, M., B. van Ommeren, and B. Geertsema, Does intermunicipal cooperation create inefficiency? A comparison of interest rates paid by intermunicipal organizations, amalgamated municipalities and not recently amalgamated municipalities

15004-EEF: Dijkstra, P.T., M.A. Haan, and M. Mulder, Design of Yardstick Competition and Consumer Prices: Experimental Evidence

15005-EEF: Dijkstra, P.T., Price Leadership and Unequal Market Sharing: Collusion in Experimental Markets

15006-EEF: Anufriev, M., T. Bao, A. Sutin, and J. Tuinstra, Fee Structure, Return Chasing and Mutual Fund Choice: An Experiment

15007-EEF: Lamers, M., Depositor Discipline and Bank Failures in Local Markets During the Financial Crisis

15008-EEF: Oosterhaven, J., On de Doubtful Usability of the Inoperability IO Model

15009-GEM: Zhang, L. and D. Bezemer, A Global House of Debt Effect? Mortgages and Post-Crisis Recessions in Fifty Economies

15010-I&O: Hooghiemstra, R., N. Hermes, L. Oxelheim, and T. Randøy, The Impact of Board Internationalization on Earnings Management

15011-EEF: Haan, M.A., and W.H. Siekman, Winning Back the Unfaithful while Exploiting the Loyal: Retention Offers and Heterogeneous Switching Costs

15012-EEF: Haan, M.A., J.L. Moraga-González, and V. Petrikaite, Price and Match-Value Advertising with Directed Consumer Search

15013-EEF: Wiese, R., and S. Eriksen, Do Healthcare Financing Privatisations Curb Total Healthcare Expenditures? Evidence from OECD Countries

15014-EEF: Siekman, W.H., Directed Consumer Search

15015-GEM: Hoorn, A.A.J. van, Organizational Culture in the Financial Sector: Evidence from a Cross-Industry Analysis of Employee Personal Values and Career Success

15016-EEF: Te Bao, and C. Hommes, When Speculators Meet Constructors: Positive and Negative Feedback in Experimental Housing Markets

15017-EEF: Te Bao, and Xiaohua Yu, Memory and Discounting: Theory and Evidence

15018-EEF: Suari-Andreu, E., The Effect of House Price Changes on Household Saving Behaviour: A Theoretical and Empirical Study of the Dutch Case

15019-EEF: Bijlsma, M., J. Boone, and G. Zwart, Community Rating in Health Insurance: Trade-off between Coverage and Selection

15020-EEF: Mulder, M., and B. Scholtens, A Plant-level Analysis of the Spill-over Effects of the German *Energiewende*

15021-GEM: Samarina, A., L. Zhang, and D. Bezemer, Mortgages and Credit Cycle Divergence in Eurozone Economies

16001-GEM: Hoorn, A. van, How Are Migrant Employees Manages? An Integrated Analysis

16002-EEF: Soetevent, A.R., Te Bao, A.L. Schippers, A Commercial Gift for Charity

16003-GEM: Bouwmeerster, M.C., and J. Oosterhaven, Economic Impacts of Natural Gas Flow Disruptions

16004-MARK: Holtrop, N., J.E. Wieringa, M.J. Gijsenberg, and P. Stern, Competitive Reactions to Personal Selling: The Difference between Strategic and Tactical Actions

16005-EEF: Plantinga, A. and B. Scholtens, The Financial Impact of Divestment from Fossil Fuels

16006-GEM: Hoorn, A. van, Trust and Signals in Workplace Organization: Evidence from Job Autonomy Differentials between Immigrant Groups

16007-EEF: Willems, B. and G. Zwart, Regulatory Holidays and Optimal Network Expansion

16008-GEF: Hoorn, A. van, Reliability and Validity of the Happiness Approach to Measuring Preferences

16009-EEF: Hinloopen, J., and A.R. Soetevent, (Non-)Insurance Markets, Loss Size Manipulation and Competition: Experimental Evidence

16010-EEF: Bekker, P.A., A Generalized Dynamic Arbitrage Free Yield Model

16011-EEF: Mierau, J.A., and M. Mink, A Descriptive Model of Banking and Aggregate Demand

16012-EEF: Mulder, M. and B. Willems, Competition in Retail Electricity Markets: An Assessment of Ten Year Dutch Experience

16013-GEM: Rozite, K., D.J. Bezemer, and J.P.A.M. Jacobs, Towards a Financial Cycle for the US, 1873-2014

16014-EEF: Neuteleers, S., M. Mulder, and F. Hindriks, Assessing Fairness of Dynamic Grid Tariffs

16015-EEF: Soetevent, A.R., and T. Bružikas, Risk and Loss Aversion, Price Uncertainty and the Implications for Consumer Search

16016-HRM&OB: Meer, P.H. van der, and R. Wielers, Happiness, Unemployment and Self-esteem

16017-EEF: Mulder, M., and M. Pangan, Influence of Environmental Policy and Market Forces on Coal-fired Power Plants: Evidence on the Dutch Market over 2006-2014

16018-EEF: Zeng,Y., and M. Mulder, Exploring Interaction Effects of Climate Policies: A Model Analysis of the Power Market

16019-EEF: Ma, Yiqun, Demand Response Potential of Electricity End-users Facing Real Time Pricing

16020-GEM: Bezemer, D., and A. Samarina, Debt Shift, Financial Development and Income Inequality in Europe

16021-EEF: Elkhuizen, L, N. Hermes, and J. Jacobs, Financial Development, Financial Liberalization and Social Capital

16022-GEM: Gerritse, M., Does Trade Cause Institutional Change? Evidence from Countries South of the Suez Canal

16023-EEF: Rook, M., and M. Mulder, Implicit Premiums in Renewable-Energy Support Schemes

17001-EEF: Trinks, A., B. Scholtens, M. Mulder, and L. Dam, Divesting Fossil Fuels: The Implications for Investment Portfolios

17002-EEF: Angelini, V., and J.O. Mierau, Late-life Health Effects of Teenage Motherhood

17003-EEF: Jong-A-Pin, R., M. Laméris, and H. Garretsen, Political Preferences of (Un)happy Voters: Evidence Based on New Ideological Measures

17004-EEF: Jiang, X., N. Hermes, and A. Meesters, Financial Liberalization, the Institutional Environment and Bank Efficiency

17005-EEF: Kwaak, C. van der, Financial Fragility and Unconventional Central Bank Lending Operations

17006-EEF: Postelnicu, L. and N. Hermes, The Economic Value of Social Capital

17007-EEF: Ommeren, B.J.F. van, M.A. Allers, and M.H. Vellekoop, Choosing the Optimal Moment to Arrange a Loan

17008-EEF: Bekker, P.A., and K.E. Bouwman, A Unified Approach to Dynamic Mean-Variance Analysis in Discrete and Continuous Time

17009-EEF: Bekker, P.A., Interpretable Parsimonious Arbitrage-free Modeling of the Yield Curve

17010-GEM: Schasfoort, J., A. Godin, D. Bezemer, A. Caiani, and S. Kinsella, Monetary Policy Transmission in a Macroeconomic Agent-Based Model

17011-I&O: Bogt, H. ter, Accountability, Transparency and Control of Outsourced Public Sector Activities

17012-GEM: Bezemer, D., A. Samarina, and L. Zhang, The Shift in Bank Credit Allocation: New Data and New Findings

17013-EEF: Boer, W.I.J. de, R.H. Koning, and J.O. Mierau, Ex-ante and Ex-post Willingness-to-pay for Hosting a Major Cycling Event

17014-OPERA: Laan, N. van der, W. Romeijnders, and M.H. van der Vlerk, Higher-order Total Variation Bounds for Expectations of Periodic Functions and Simple Integer Recourse Approximations

17015-GEM: Oosterhaven, J., Key Sector Analysis: A Note on the Other Side of the Coin

17016-EEF: Romensen, G.J., A.R. Soetevent: Tailored Feedback and Worker Green Behavior: Field Evidence from Bus Drivers

17017-EEF: Trinks, A., G. Ibikunle, M. Mulder, and B. Scholtens, Greenhouse Gas Emissions Intensity and the Cost of Capital

17018-GEM: Qian, X. and A. Steiner, The Reinforcement Effect of International Reserves for Financial Stability

17019-GEM/EEF: Klasing, M.J. and P. Milionis, The International Epidemiological Transition and the Education Gender Gap

2018001-EEF: Keller, J.T., G.H. Kuper, and M. Mulder, Mergers of Gas Markets Areas and Competition amongst Transmission System Operators: Evidence on Booking Behaviour in the German Markets

2018002-EEF: Soetevent, A.R. and S. Adikyan, The Impact of Short-Term Goals on Long-Term Objectives: Evidence from Running Data

2018003-MARK: Gijsenberg, M.J. and P.C. Verhoef, Moving Forward: The Role of Marketing in Fostering Public Transport Usage

2018004-MARK: Gijsenberg, M.J. and V.R. Nijs, Advertising Timing: In-Phase or Out-of-Phase with Competitors?

2018005-EEF: Hulshof, D., C. Jepma, and M. Mulder, Performance of Markets for European Renewable Energy Certificates

2018006-EEF: Fosgaard, T.R., and A.R. Soetevent, Promises Undone: How Committed Pledges Impact Donations to Charity

2018007-EEF: Durán, N. and J.P. Elhorst, A Spatio-temporal-similarity and Common Factor Approach of Individual Housing Prices: The Impact of Many Small Earthquakes in the North of Netherlands

2018008-EEF: Hermes, N., and M. Hudon, Determinants of the Performance of Microfinance Institutions: A Systematic Review

2018009-EEF: Katz, M., and C. van der Kwaak, The Macroeconomic Effectiveness of Bank Bail-ins

2018010-OPERA: Prak, D., R.H. Teunter, M.Z. Babai, A.A. Syntetos, and J.E. Boylan, Forecasting and Inventory Control with Compound Poisson Demand Using Periodic Demand Data

2018011-EEF: Brock, B. de, Converting a Non-trivial Use Case into an SSD: An Exercise

2018012-EEF: Harvey, L.A., J.O. Mierau, and J. Rockey, Inequality in an Equal Society

2018013-OPERA: Romeijnders, W., and N. van der Laan, Inexact cutting planes for two-stage mixed-integer stochastic programs

2018014-EEF: Green, C.P., and S. Homroy, Bringing Connections Onboard: The Value of Political Influence

2018015-OPERA: Laan, N. van der, and W. Romeijnders, Generalized aplha-approximations for two-stage mixed-integer recourse models