Convex approximations for mixed-integer mean-risk recourse models with conditional value-at-risk

Ruben van Beesten
July 7, 2017

Supervisors:
Dr W. Romeijnders
Prof. M.H. van der Vlerk

Co-assessor:
Prof. K.J. Roodbergen
In memory of
Maarten van der Vlerk
Convex approximations for mixed-integer mean-risk recourse models with conditional value-at-risk

Ruben van Beesten
July 7, 2017

Abstract

In traditional two-stage mixed-integer recourse models, the expected value of the total costs is minimized. In order to address risk-averse attitudes of decision makers, we consider a weighted mean-risk objective. Conditional value-at-risk is used as our risk measure. Integrality conditions on decision variables make the model non-convex and hence, hard to solve. To tackle this problem, we derive a convex approximation model and a corresponding uniform error bound, which depends on the total variations of the density functions of the random variables in the model. We show that the error bound converges to zero if these total variations go to zero. In addition, for the special case of simple integer recourse, we derive a convex approximation with an error bound which is tight if the random variable in the model has a fat-tailed distribution.

Key words:

stochastic programming, mean-risk models, conditional value-at-risk, integer recourse, convex approximations

1 Introduction. In many practical situations, decisions have to be made under uncertainty. For instance, production firms need to decide on their production level without knowing the exact demand level beforehand. For such problems, recourse models have been developed within the field of stochastic programming. These models can be used to find optimal solutions for decision problems under uncertainty and have a wide range of applications (for examples of applications, see e.g. Birge and Louveaux, 2011). Often, decision makers face certain indivisibilities or on/off decisions, for example if production can only take place in batches of a fixed size. To allow for such problem characteristics, we restrict some decision variables in the recourse model to be integer. Finally, empirical evidence suggests that most people are risk-averse (for empirical evidence, see e.g. Dohmen et al., 2011). In order to capture this risk-averseness, we use a mean-risk objective function in our recourse model.

We consider the two-stage mixed-integer mean-risk recourse model (with random right-hand side only):

$$\min_{x \in X} \{cx + Q_{\beta}^p(x)\},$$

with $X = \{x \in \mathbb{R}^{n_1} : Ax = b\}$ and $Q_{\beta}^p$ is the recourse function

$$Q_{\beta}^p(x) := E_\omega [v(\omega - Tx)] + \rho \text{CVaR}_\beta [v(\omega - Tx)], \quad x \in \mathbb{R}^{n_1},$$

where $v$ is the second-stage value function

$$v(s) := \min_y \{qy : Wy = s, \; y \in \mathbb{Z}_{+}^{n_2} \times \mathbb{R}_{+}^{n_3}\}, \quad s \in \mathbb{R}^{m},$$
and CVaR$\beta$ denotes the $\beta$-Conditional Value-at-Risk, defined as
\[
\text{CVaR}_\beta[v(\omega - Tx)] = \min_{\alpha \in \mathbb{R}} \left\{ \alpha + \frac{1}{\beta} \mathbb{E}_\omega \left[ (v(\omega - Tx) - \alpha)^+ \right] \right\},
\]
for probability $\beta \in (0, 1)$. The second-stage decision variables $y$ represent the recourse actions that compensate for infeasibilities with respect to the constraint $Tx = \omega$. Throughout we assume that $W$ is an integer matrix and $\omega$ is a random vector with joint density function $f$. In order to model indivisibilities or on/off decisions, we require some of the second-stage decision variables to be integer. This integrality causes the recourse function to be generally non-convex and hence, the model is not easily solvable. In this paper we try to overcome this difficulty of non-convexity. For ease of exposition, we assume that the first-stage decision variables $x$ are continuous, but this goes without loss of generality of the results found in this paper.

To reflect the fact that most decision makers are risk-averse, we explicitly incorporate a risk measure into the recourse function $Q^p\rho$. The risk measure we use is Conditional Value-at-Risk (CVaR). The reason for using this specific risk measure is twofold. Firstly, CVaR is a so-called coherent risk measure. Artzner et al. (1999) define a risk measure to be coherent if it adheres to a list of four axioms. Specifically, these axioms are positive homogeneity, monotonicity, subadditivity, and translation invariance. These axioms are to be regarded as a set of minimum requirements that a risk measure should satisfy. Although the set of requirements is minimal, it should be noted that for instance the popular risk measure Value-at-Risk is not coherent (Acerbi and Tasche, 2002).

Secondly, CVaR is consistent with second-order stochastic dominance (Ogryczak and Ruszczyński, 2002). Levy (1992) shows that all utility maximizing risk-averse decision makers prefer a second-order stochastically dominating risk to a dominated one. Thus, in order for a risk measure to model risk averseness consistently, it is necessary that the risk measure used is consistent with second-order stochastic dominance.

The setting of two-stage mixed-integer mean-risk recourse models using CVaR as a risk measure (referred to as mean-CVaR models in the remainder) has been considered by researchers before. Schultz and Tiedemann (2006) address structure, stability and algorithms for this class of models. Specifically, they show that despite the fact that such problems are generally non-convex, they still have certain continuity properties with respect to the first-stage decision variables and the integrating probability measure. Furthermore, it is shown that under mild technical conditions, the mean-CVaR model has an optimal solution. In addition, the authors propose a scenario decomposition algorithm for the case that there is a finite number of scenarios for the random parameters in the model.

Most other solution methods proposed in literature assume a finite number of scenarios as well. Some authors use the mixed-integer linear programming (MILP) formulation of mean-CVaR models due to Schultz and Tiedemann (2006) and solve these MILPs using off-the-shelf software (Soleimani et al., 2014; Yau et al., 2011). Others use decomposition methods, based on the L-shaped method introduced by Van Slyke and Wets (1969) (Ahmed, 2006; Noyan, 2012). In the rare cases that models with continuously distributed random variables are considered, authors resort to simulation to find approximate solutions (Alem and Morabito, 2013; Farzan et al., 2015).

The fact that existing solution techniques do not allow for continuously distributed random variables is a drawback, since many uncertainties in real life have a continuum of possible
 outcomes (e.g. stock prices). Moreover, in the context of risk-averseness, we have a particular interest in extreme events that occur only rarely. By allowing for continuously distributed random variables, we solve this issue, at the expense of making the model harder to solve. Furthermore, it turns out that using continuously distributed random variables is particularly useful for the solution method that we propose.

In order to overcome the difficulty of non-convexity, we construct a convex approximation \( \hat{Q}_\rho \) for the recourse function \( Q_\rho \). The idea is that we construct a model that (1) is convex and thus easy to solve, and (2) is a close approximation of the original model, such that the resulting first-stage decisions \( x \) are near-optimal in the original model. In order to assess this closeness of the approximation, we will derive an upper bound for \( ||Q_\rho - \hat{Q}_\rho||_\infty \). The derivation of this upper bound is the main focus of this paper.

Convex approximations of mixed-integer recourse models were first developed by Klein Hanевeld et al. (2006) for the special case of simple integer recourse with a separable recourse matrix \( W \), and were later extended to totally unimodular integer (van der Vlerk, 2004) and mixed-integer recourse models (van der Vlerk, 2010).

Recently, substantial progress has been made by deriving error bounds for convex approximations for recourse models with a non-separable recourse matrix \( W \). For the totally unimodular integer case, Romeijnders et al. (2015) derive an error bound which depends on the total variations of the density functions of all random variables in the model. This result is improved upon by Romeijnders, van der Vlerk et al. (2016), leading to a convex approximation that is tight in a worst-case sense. The main building block in the derivations is a total variation upper bound on the expectation of periodic functions. The results for the TU integer case are generalized to the general mixed-integer case in Romeijnders, Schultz et al. (2016). Here, asymptotic periodicity of the second-stage mixed-integer value function is used in combination with the total variation upper bounds for expectations of periodic functions mentioned above.

The main contribution of this paper is that we generalize the convex approximation technique to recourse models with a mean-CVaR objective function. Specifically, we will construct convex approximations for mixed-integer mean-CVaR recourse models and derive total variation error bounds.

The remainder of this paper is structured as follows. First, we discuss how the error bounds derived in this paper can be applied in Section 2. This will clarify some of the choices made in the remainder of the paper. Next, in Section 3 we consider the general mixed-integer case and derive a convex approximation with a uniform error bound. This error bound is shown to converge to zero as the total variations of the density functions of the random variables in the model go to zero. In Section 4 we consider the special case of simple integer recourse and again derive a convex approximation with a uniform error bound. The bound for this special case is tighter than for the general mixed-integer case and is shown to be especially useful if the density function of the random variable in the model has fat tails. Finally, we discuss our results and directions for future research in Section 5.

Throughout the paper we make the following assumptions.

(i) The recourse is complete; i.e. for every \( s \in \mathbb{R}^m \) there exists a feasible recourse action \( y \), such that \( v(s) < +\infty \).

(ii) The recourse is sufficiently expensive; i.e. \( v(s) > -\infty \) for all \( s \in \mathbb{R}^m \).
(iii) $E_\omega [||\omega||_2]$ is finite.

Assumptions (i)-(iii) imply that the recourse function $Q^p_\beta$ is finite everywhere.

2 Preliminaries: application of error bounds. Before deriving error bounds, we first discuss how these bounds can be applied. This discussion will explain some of the choices made in the rest of the paper. Specifically, it will clarify why we focus on pure CVaR models in the remainder of this paper, rather than on mean-CVaR models.

In this paper we consider mean-CVaR models in which the recourse function $Q_\rho^\beta$ consists of an expected value term and a CVaR term. It should be noted that for expected value models, error bounds have already been derived in literature (see Romeijnders, van der Vlerk et al. (2016) and Romeijnders, Schultz et al. (2016)). The novelty in this paper is the inclusion of CVaR in the recourse function. To simplify derivation, we will restrict our attention to pure CVaR problems. That is, we define the CVaR recourse function
\[
\bar{Q}_\beta(x) = \text{CVaR}_{\beta}[v(\omega - Tx)], \quad x \in \mathbb{R}^n, \tag{4}
\]
and use this in the remainder of our analysis. The resulting error bounds can be generalized to error bounds for mean-CVaR models by simply adding the obtained error bounds for CVaR models to error bounds for expected value models from literature. Lemma 1 shows how this can be done. Since these generalizations are straightforward and somewhat cumbersome, we have decided to leave them to the reader and present error bounds for CVaR models only.

**Lemma 1** Let $Q^p_\beta$ and $\bar{Q}_\beta$ be the recourse functions defined as in (2) and (4), respectively. Moreover, define the expected value recourse function $Q^E$ as
\[
Q^E(x) = E_\omega [v(\omega - Tx)].
\]
Let $v$ be the corresponding second-stage value function with an approximation $\hat{v}$. Let $\hat{Q}^p_\beta$, $\hat{Q}_\beta$ and $\hat{Q}^E$ be the approximating recourse functions obtained by replacing $v$ by $\hat{v}$ in their corresponding definitions. Suppose that we have the uniform error bounds $||\bar{Q}_\beta - \hat{Q}_\beta||_\infty \leq \bar{B}$ and $||Q^p_\beta - \hat{Q}^p_\beta||_\infty \leq B^E$ for the CVaR and the expected value model, respectively. Then, for the mean-CVaR model, we have the error bound
\[
||Q^p_\beta - \hat{Q}^p_\beta||_\infty \leq B^E + \rho \bar{B}.
\]

**Proof.** The proof follows immediately from the triangle inequality. \qed

If an upper bound on $||Q^p_\beta - \hat{Q}^p_\beta||_\infty$ has been obtained, this error bound can be translated to an error bound for the optimal value of the mean-CVaR optimization problem. The following lemma, which is proven by Romeijnders et al. (2015), shows how this can be done.

**Lemma 2** Consider the minimization problem
\[
\xi^* := \min_{x \in X} \left\{ cx + Q^p_\beta(x) \right\}, \tag{5}
\]
and its approximation
\[
\hat{\xi} := \min_{x \in X} \left\{ cx + \hat{Q}^p_\beta(x) \right\}, \tag{6}
\]
with optimal solutions $x^*$ and $\hat{x}$, respectively. Then,
\[
||\xi^* - \hat{\xi}|| \leq ||Q^p_\beta - \hat{Q}^p_\beta||_\infty \quad \text{and} \quad ||\xi^* - cx^* - \hat{Q}^p_\beta(\hat{x})|| \leq 2||Q^p_\beta - \hat{Q}^p_\beta||_\infty.
\]
So the optimal value in the approximate model differs from the actual optimal value by at most $||Q_\beta^p - \hat{Q}_\beta^p||_\infty$.

3 General mixed-integer CVaR models. In this section we derive a convex approximation for the general mixed-integer CVaR recourse model. That is, we consider the model described by equations (1)-(3), but with the mean-CVaR recourse function $Q_\beta^p$ replaced by the CVaR recourse function $\bar{Q}_\beta$ from (4). By definition,

$$Q_\beta(x) = \min_{\alpha \in \mathbb{R}} \{ \alpha + \frac{1}{\beta} \mathbb{E}_\omega \left[ (v(\omega - Tx) - \alpha)^+ \right] \}.$$ 

The main idea in this section is that we will interpret $\alpha$ in the minimization problem above as a first-stage decision variable. The possibility to do so is in fact one of the nice properties of CVaR and it has been suggested by Rockafellar and Uryasev (2002). We introduce a new second-stage value function $\bar{v}_\alpha$, defined as

$$\bar{v}_\alpha(s) = (v(s) - \alpha)^+.$$ 

Now we can rewrite the stochastic programming problem as

$$\min_{x \in \mathcal{X}} \{ cx + \bar{Q}_\beta(x) \} = \min_{x \in \mathcal{X}} \min_{\alpha \in \mathbb{R}} \{ cx + \alpha + \hat{Q}_\beta(x, \alpha) \},$$

where

$$\hat{Q}_\beta(x, \alpha) = \frac{1}{\beta} \mathbb{E}_\omega \left[ \bar{v}_\alpha(\omega - Tx) \right].$$

The value function $\bar{v}_\alpha$ can be rewritten as a mixed-integer linear program, as we will show later on. Hence, the optimization problem above is again a risk-neutral two-stage mixed-integer recourse problem. Romeijnders, Schultz et al. (2016) derive a convex approximation with a uniform error bound for such problems under the condition that the right-hand side variables of the constraints in the MILP formulation of $\bar{v}_\alpha$ are stochastic. We will show that this condition does not hold for the problem described above. Hence, the results cannot be applied to our pure CVaR model. We will construct an alternative convex approximation of $\hat{Q}_\beta$ and use ideas from Romeijnders, Schultz et al. (2016) to derive a uniform error bound.

3.1 Asymptotic semi-periodicity of $\bar{v}_\alpha$. The first step in our analysis is proving that the mixed-integer value function $\bar{v}_\alpha$ is asymptotically semi-periodic. By this short-hand term we mean that on shifted polyhedral cones, $\bar{v}_\alpha$ is the sum of a linear function and a function that exhibits periodicity with respect to certain basis matrices.

To understand the cause of this periodicity, consider the LP-relaxation of the mixed-integer value function $v$. By the basis decomposition theorem by Walkup and Wets (1969), we can identify basis matrices $B^k$ and corresponding polyhedral cones $\Lambda^k \subseteq \mathbb{R}^m$, $k = 1, \ldots, K$, such that for all $s \in \Lambda^k$, $v_{\text{LP}}(s)$ attains its value through the basis matrix $B^k$, i.e. $v_{\text{LP}}(s) = q_B(B^k)^{-1}s$. We will show that a similar result holds for $v$, but only on shifted versions of the cones $\Lambda^k$, $k = 1, \ldots, K$. 

5
Definition 1 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

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

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

From Romeijnders, Schultz et al. (2016), we know that there exist constants $d^k > 0$, $k = 1, \ldots, K$ such that for all $s \in \Lambda^k(d^k)$, the mixed-integer value function $v(s)$ attains its value through the basis matrix $B^k$. That is, $v(s) = q_{B^k}(B^k)^{-1}s + \psi^k(s)$. Here $\psi^k(s)$ represents the “penalty” incurred from having non-zero non-basic variables in order to satisfy integrality restrictions on basic variables. Romeijnders, Schultz et al. (2016) show that if $s^1$ and $s^2$ are distinct elements of $\Lambda^k(d^k)$ such that $(B^k)^{-1}s^1$ and $(B^k)^{-1}s^2$ have the same fractional values, then $\psi^k(s^1) = \psi^k(s^2)$. We say that $\psi^k$ is $B^k$-periodic on $\Lambda^k(d^k)$. It turns out that $\bar{v}_\alpha$ exhibits the same type of periodicity as $v$.

Definition 2 Let the function $g : \mathbb{R}^m \rightarrow \mathbb{R}^n$ be given and let $B$ be an $m \times m$ matrix. Then, $g$ is called $B$-periodic if for every $x \in \mathbb{R}^m$ and $l \in \mathbb{Z}^m$,

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

Proposition 1 Consider the mixed-integer value function $v$, defined as

$$v(s) = \min_y \{qy :Wy = s, y \in \mathbb{Z}^{n_2}_+ \times \mathbb{R}^{n_3}_+\}, \quad s \in \mathbb{R}^m,$$

and its LP-relaxation $v^{LP}$, where $W$ is an integer matrix and $v(s)$ is finite for all $s \in \mathbb{R}^m$. Moreover, consider the mixed-integer value function $\bar{v}_\alpha(s)$, defined by

$$\bar{v}_\alpha(s) = (v(s) - \alpha)^+.$$

Then, there exist dual feasible basis matrices $B^k$ of $v^{LP}$, closed convex polyhedral cones $\Lambda^k := \{t \in \mathbb{R}^m : (B^k)^{-1}t \geq 0\}$, distances $d^k := ||\det(B^k)|| \sum_{j=1}^n ||N^k_j||_{2o}$, where $N^k_j$ is the $j$th column of $N^k$, $B^k$-periodic functions $\psi^k$ and non-negative constants $r^k$, $k = 1, \ldots, K$, such that we have the following:

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

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

(iii) For every $k = 1, \ldots, K$,

$$\bar{v}_\alpha(s) = (q_{B^k}(B^k)^{-1}s + \psi^k(s) - \alpha)^+, \quad s \in \Lambda^k(d^k), \quad \alpha \in \mathbb{R},$$

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

(iv) For every $k = 1, \ldots, K$,

$$0 \leq \psi^k(s) \leq r^k, \quad s \in \mathbb{R}^m.$$
Proof. Part (i), (ii), (iv) and the fact that $\psi^k = \psi^l$ if $q_{B_k}(B_k)^{-1} = q_{B_l}(B_l)^{-1}$ follow directly from Lemma 2.1 in Romeijnders, Schultz et al. (2016). By the same result we can write
\[ v(s) = q_{B_k}(B_k)^{-1}s + \psi^k(s), \quad s \in \Lambda^k(d^k), \]
for every $k = 1, \ldots, K$. Fix $k$ and let $s \in \Lambda^k(d^k)$ be given. Then,
\[ \bar{v}_\alpha(s) = (v(s) - \alpha)^+ = (q_{B_k}(B_k)^{-1}s + \psi^k(s) - \alpha)^+. \]

Proposition 1 tells us that on shifted convex cones $\Lambda^k(d^k)$, the approximating value function $\bar{v}_\alpha(s)$ is the positive part of the sum of a linear function and a periodic function, hence the term asymptotic semi-periodicity.

We can derive an upper bound on the probability that $s = \omega - Tx$ is not on one of these shifted convex cones. Define the set $\mathcal{N}$ as the subset of $\mathbb{R}^m$ on which $\bar{v}_\alpha$ does not exhibit this kind of periodicity. That is, $\mathcal{N} = \mathbb{R}^m \setminus \bigcup_{k=1}^K \Lambda^k(d^k)$. Romeijnders, Schultz et al. (2016) show that $\mathcal{N}$ can be covered by finitely many hyperslices.

Definition 3 A hyperslice is a set $H \subset \mathbb{R}^m$ of the form
\[ H := \{ x \in \mathbb{R}^m \mid \gamma \leq a^T x \leq \gamma + \delta \}, \]
where $a \in \mathbb{R}^m \setminus \{0\}$, $\gamma \in \mathbb{R}$, and $\delta \in \mathbb{R}^+$. 

Moreover, Romeijnders, Schultz et al. (2016) prove that $\mathbb{P}\{\omega - Tx \in \mathcal{N}\}$ has an asymptotically converging upper bound. That is, as the total variations of the univariate conditional densities $f_i(\cdot | x - i)$ go to zero, $\mathbb{P}\{\omega - Tx \in \mathcal{N}\}$ converges to zero. So the probability that $\omega - Tx$ is on a subset of the domain of $\bar{v}_\alpha$ on which it exhibits semi-periodicity converges to one. This asymptotic semi-periodicity can be exploited to derive a convex approximation of $\hat{Q}_\beta$ with an asymptotically converging error bound.

### 3.2 Convex approximation of the recourse function.
Recall that the recourse function $\hat{Q}_\beta$ is defined as
\[ \hat{Q}_\beta(x, \alpha) = \frac{1}{\beta} \mathbb{E}_\omega [\bar{v}_\alpha(\omega - Tx)]. \]

This function is not convex in general. In this subsection we construct a convex approximation $\hat{\hat{Q}}_\beta$ of $\hat{Q}_\beta$ using the results from Proposition 1. Our approximating recourse function will be of the form $\hat{\hat{Q}}_\beta(x, \alpha) = \frac{1}{\beta} \mathbb{E}_\omega [\hat{\bar{v}}_\alpha(\omega - Tx)]$. The main idea is to approximate the $B_k$-periodic functions $\psi^k(s)$ in the characterization of $\bar{v}_\alpha$ by their average values. The resulting approximating value function $\hat{\bar{v}}_\alpha$ is convex and hence, the approximating recourse function $\hat{\hat{Q}}_\beta$ is convex as well.

Before defining our convex approximation $\hat{\hat{Q}}_\beta$, we give a short outline of our approach in this section. The reason for this is that we have to take our procedure for deriving an error bound into account when constructing the convex approximation. Fix $x \in X$ and $\alpha \in \mathbb{R}$ and
write \( \hat{\omega} := \omega - Tx \), with corresponding joint density function \( g \). Our goal is to find an upper bound on

\[
|\hat{Q}_\beta(x, \alpha) - \hat{Q}_\beta(x, \alpha)| = \frac{1}{\beta} \left| \int_{\mathbb{R}^m} (\hat{v}_\alpha(s) - \hat{v}_\alpha(s))g(s)ds \right|.
\]

We take the following approach. We define the approximating value function \( \hat{v}_\alpha \) in such a way that there exist disjoint subsets \( B^j_\alpha \) of \( \mathbb{R}^m \) and corresponding periodic functions \( \phi^j_\alpha \) with mean values \( \nu^j_\alpha \), \( j = 1, \ldots, J \), such that we have \( \hat{v}_\alpha(s) - \hat{v}_\alpha(s) = \phi^j_\alpha(s) - \nu^j_\alpha \), for \( s \in B^j_\alpha \). Then we have

\[
|\hat{Q}_\beta(x, \alpha) - \hat{Q}_\beta(x, \alpha)| \leq \frac{1}{\beta} \sum_{j=1}^{J} \left| \int_{B^j_\alpha} (\phi^j_\alpha(s) - \nu^j_\alpha)g(s)ds \right| + \frac{1}{\beta} \left| \int_{\mathcal{N}_\alpha} (\hat{v}_\alpha(s) - \hat{v}_\alpha(s))g(s)ds \right|,
\]

where \( \mathcal{N}_\alpha := \mathbb{R}^m \setminus \bigcup_{j=1}^{J} B^j_\alpha \). Using a result from Romeijnders, Schultz et al. (2016), we can find an asymptotically converging error bound on \( |\int_{B^j_\alpha} (\phi^j_\alpha(s) - \nu^j_\alpha)g(s)ds| \), \( j = 1, \ldots, J \). Moreover, it can be shown that \( \mathcal{N}_\alpha \) can be covered by finitely many hyperslices, which will be used to show that \( \mathbb{P}\{\hat{\omega} \in \mathcal{N}_\alpha\} \) has an asymptotically converging upper bound. After proving that \( ||\hat{v} - \hat{v}_\alpha||_{\infty} \leq \gamma \) for some positive scalar \( \gamma \), it follows that \( |\int_{\mathcal{N}_\alpha} (\hat{v}_\alpha(s) - \hat{v}_\alpha(s))g(s)ds| \leq |\gamma \int_{\mathcal{N}_\alpha} g(s)ds| = \gamma \mathbb{P}\{\hat{\omega} \in \mathcal{N}_\alpha\} \). Combining all findings yields an asymptotically converging upper bound on \( |\hat{Q}_\beta(x, \alpha) - \hat{Q}_\beta(x, \alpha)| \).

### 3.2.1 Convex approximation \( \hat{v}_\alpha \) of the value function \( \hat{v}_\alpha \)

We now introduce our convex approximation \( \hat{v}_\alpha \) of the value function \( \hat{v}_\alpha \). Fix the value of \( \alpha \in \mathbb{R} \). From Proposition 1 we know that for \( s \in \Lambda^k(d^k) \),

\[
\hat{v}_\alpha(s) = (q_{B^k}(B^k)^{-1}s + \psi^k(s) - \alpha)^+,
\]

where \( \psi^k \) is \( B^k \)-periodic. This function is not convex due to the periodicity of \( \psi^k(s) \). A possible way to obtain a convex approximation of \( \hat{v}_\alpha \) is to replace \( \psi^k \) by its mean value. From Romeijnders, Schultz et al. (2016) we know that the \( B^k \)-periodicity of \( \psi^k \) implies that \( \psi^k \) is also \( p_k I_m \)-periodic, where \( p_k := |\det(B^k)| \) and \( I_m \) is the identity matrix of size \( m \). Hence, we can characterize the mean value of \( \psi^k \) as \( \Gamma^k := p_k^{-m} \int_0^{p_k} \cdots \int_0^{p_k} \psi^k(s)ds_1 \cdots ds_m \).

Interestingly, we can and will use this mean value for all \( k = 1, \ldots, K \), except when \( q_{B^k} = 0 \). For this exception, we will define an average value \( \Gamma^k_\alpha \) that depends on \( \alpha \). This is the only way to make sure that we can to cover \( \mathcal{N}_\alpha \) by finitely many hyperslices.

**Definition 4** Consider the mixed-integer value function \( \hat{v}_\alpha(s) := (v(s) - \alpha)^+ \), where \( v(s) \) is defined as in (3). Let \( B^k \), \( q_{B^k} \), and \( \psi^k \), \( k = 1, \ldots, K \) be the basis matrices, corresponding cost vectors, and \( B^k \)-periodic functions, respectively, from Proposition 1. Then, for every \( k = 1, \ldots, K \), and \( \alpha \in \mathbb{R} \) we define

\[
\Gamma^k_\alpha := \begin{cases} 
\Gamma^k, & \text{if } q_{B^k} \neq 0, \\
p_k^{-m} \int_0^{p_k} \cdots \int_0^{p_k} (\psi^k(s) - \alpha)^+ ds_1 \cdots ds_m + \alpha, & \text{if } q_{B^k} = 0,
\end{cases}
\]

where

\[
\Gamma^k := p_k^{-m} \int_0^{p_k} \cdots \int_0^{p_k} \psi^k(s)ds_1 \cdots ds_m,
\]

and \( p_k := |\det(B^k)| \).
Replacing \( \psi^k(s) \) by \( \Gamma^k_\alpha \) now yields an approximation of \( \bar{v}_\alpha(s) \) for \( s \in \Lambda^k(d^k) \) for each \( k = 1, \ldots, K \). We combine these approximations by taking the pointwise maximum, yielding our convex approximating value function \( \hat{v}_\alpha \).

**Definition 5** Consider the mixed-integer value function \( \bar{v}_\alpha(s) := (v(s) - \alpha)^+ \), where \( v(s) \) is defined in (3). Let \( B^k \) and \( q_{B^k} \) be the basis matrices and corresponding cost vectors from Proposition 1. We define the approximation \( \hat{v}_\alpha(s) \) of \( \bar{v}_\alpha(s) \) by

\[
\hat{v}_\alpha(s) = \max_{k=1,\ldots,K} \left\{ \left( q_{B^k}(B^k)^{-1} s + \Gamma^k_\alpha - \alpha \right)^+ \right\}, \quad s \in \mathbb{R}^m, \alpha \in \mathbb{R},
\]

where \( \Gamma^k_\alpha \), \( k = 1, \ldots, K \), are defined as in Definition 4.

The fact that \( \Gamma^k_\alpha \), \( k = 1, \ldots, K \), in the definition of \( \hat{v}_\alpha(s) \) depend on the decision variable \( \alpha \) is a deviation from literature. It is due to a fundamental difference in the structure of the second-stage value function \( \bar{v}_\alpha \), compared to the value function \( v \), which is considered in e.g. Romeijnders, Schultz et al. (2016). The MILP-formulation of \( v \) is given by

\[
v(s) = \min_y \left\{ qy : Wy = s, y \in \mathbb{Z}^{n_2}_+ \times \mathbb{R}^{n_3}_+ \right\}, \quad s \in \mathbb{R}^m.
\]

Stochasticity is incorporated in the model by taking \( s = \omega - Tx \) as argument for \( v \). Hence, the right-hand side of the constraints in the MILP-formulation of \( v \) is completely stochastic. The MILP-formulation of \( \bar{v}_\alpha \), on the other hand, is given by

\[
\bar{v}_\alpha(s) = \min_y \left\{ y_0 : Wy = s, y_0 - qy - z = -\alpha, y \in \mathbb{Z}^{n_2}_+ \times \mathbb{R}^{n_3}_+, y_0 \in \mathbb{R}_+, z \in \mathbb{R}_+ \right\}, \quad s \in \mathbb{R}^m,
\]

where \( y_0 \) and \( z \) are artificial variables, used to model the relationship \( \bar{v}_\alpha(s) = (v(s) - \alpha)^+ \). The right-hand side of the constraint \( y_0 - qy - z = -\alpha \), is non-stochastic. Instead, it only depends on a first-stage decision variable. Literature on solving mixed-integer recourse models using convex approximations has so far only considered problems with a fully random right-hand side in the mixed-integer linear program defining the second-stage value function. For instance, see Romeijnders, van der Vlerk et al. (2016) and Romeijnders, Schultz et al. (2016). The fact that we deviate from this results in \( \alpha \) entering in \( \Gamma^k_\alpha \).

We provide a useful characterization of \( \Gamma^k_\alpha \) and prove convexity of \( \Gamma^k_\alpha \) in \( \alpha \). This result is used to prove convexity of \( \hat{v}_\alpha \) in \( (s, \alpha) \), which will in turn suffice to prove convexity of \( \hat{Q}_\beta \).

**Lemma 3** Let \( \Gamma^k_\alpha \), \( k = 1, \ldots, K \) be defined as in Definition 4. Furthermore, let \( r^k \), \( k = 1, \ldots, K \) be defined as in Proposition 1. Then, for \( k \) with \( q_{B^k} = 0 \), we can write

\[
\Gamma^k_\alpha = \begin{cases} 
\Gamma^k, & \text{if } \alpha \leq 0, \\
p_k^{-m} \int_0^{p_k} \cdots \int_0^{p_k} (\psi^k(s) - \alpha)^+ ds_1 \cdots ds_m + \alpha, & \text{if } 0 < \alpha < r^k, \\
\alpha, & \text{if } \alpha \geq r^k.
\end{cases}
\]

Moreover, for all \( k = 1, \ldots, K \), \( \Gamma^k_\alpha \) is convex as a function of \( \alpha \).

**Proof.** We first prove the characterization of \( \Gamma^k_\alpha \). Let \( k \) be such that \( q_{B^k} = 0 \). Then, by Definition 4, we can write \( \Gamma^k_\alpha = p_k^{-m} \int_0^{p_k} \cdots \int_0^{p_k} (\psi^k(s) - \alpha)^+ ds_1 \cdots ds_m + \alpha \). This immediately proves the characterization of \( \Gamma^k_\alpha \) for \( 0 < \alpha < r^k \). Now suppose that \( \alpha \leq 0 \). Note that for all \( s \in \mathbb{R}^m \) we have \( \psi^k(s) \geq 0 \). Hence, \( \psi^k(s) - \alpha \geq -\alpha \geq 0 \) and thus \( (\psi^k(s) - \alpha)^+ = \psi^k(s) - \alpha \). This yields \( \Gamma^k_\alpha = p_k^{-m} \int_0^{p_k} \cdots \int_0^{p_k} \psi^k(s) ds_1 \cdots ds_m = \Gamma^k \). Next, suppose that \( \alpha \geq r^k \). Note
that for all $s \in \mathbb{R}^m$ we have $\psi^k(s) \leq r^k$. Hence, $\psi^k(s) - \alpha \leq r^k - r^k = 0$, which implies $(\psi^k(s) - \alpha)^+ = 0$. Hence, $\Gamma^{k}_{\alpha} = \int_{0}^{p_{k}^{-m}} \cdots \int_{0}^{p_{k}} (\psi^k(s) - \alpha)^+ ds_1 \cdots ds_m + \alpha = \alpha$.

Finally, we need to prove convexity of $\Gamma^{k}_{\alpha}$. For $k$ with $q_{B^k} \neq 0$, we have $\Gamma^{k}_{\alpha} = \Gamma^{k}$, which is constant and thus convex in $\alpha$. Next, fix $k$ such that $q_{B^k} = 0$. Then, we need to show that $p_{k}^{-m} \int_{0}^{p_{k}} \cdots \int_{0}^{p_{k}} (\psi^k(s) - \alpha)^+ ds_1 \cdots ds_m + \alpha$ is convex in $\alpha$. Let $\alpha_0$ and $\alpha_1$ be two given scalars with $\alpha_0 < \alpha_1$ and let $\lambda \in [0, 1]$ be given. Define $\tilde{\alpha} := \lambda \alpha_0 + (1 - \lambda) \alpha_1$. Then,

$$
\begin{align*}
\Gamma^{k}_{\tilde{\alpha}} - \tilde{\alpha} &= p_{k}^{-m} \int_{0}^{p_{k}} \cdots \int_{0}^{p_{k}} (\psi^k(s) - \tilde{\alpha})^+ ds_1 \cdots ds_m \\
&= p_{k}^{m} \int_{0}^{p_{k}} \cdots \int_{0}^{p_{k}} (\lambda (\psi^k(s) - \alpha_0) + (1 - \lambda) (\psi^k(s) - \alpha_1))^+ ds_1 \cdots ds_m \\
&\leq \lambda (\Gamma^{k}_{\alpha_0} - \alpha_0) + (1 - \lambda)(\Gamma^{k}_{\alpha_1} - \alpha_1) \\
&= \lambda \Gamma^{k}_{\alpha_0} + (1 - \lambda) \Gamma^{k}_{\alpha_1} - \tilde{\alpha},
\end{align*}
$$

where the inequality follows from the convexity of $(\cdot)^+$. This implies that $\Gamma^{k}_{\tilde{\alpha}} \leq \lambda \Gamma^{k}_{\alpha_0} + (1 - \lambda) \Gamma^{k}_{\alpha_1}$ and by definition, $\Gamma^{k}_{\alpha}$ is convex as a function of $\alpha$.

**Lemma 4** Let $\hat{\alpha}$ be the approximating value function from Definition 5. Then, $\hat{\alpha}$ is convex in $(s, \alpha)$.

**Proof.** By definition, $\hat{\alpha}(s) = \max_{k=1, \ldots, K} \{ q_{B^k} (B^k)^{-1} s + \Gamma^{k}_{\alpha} - \alpha \}^+$, $s \in \mathbb{R}^m$, $\alpha \in \mathbb{R}$. For each $k = 1, \ldots, K$, the function $(s, \alpha) \mapsto q_{B^k} (B^k)^{-1} s - \alpha$ is linear and thus convex. Furthermore, by Lemma 3, $\Gamma^{k}_{\alpha}$ is convex as a function of $\alpha$. Using the fact that $(\cdot)^+$ and the point-wise maximum of a finite number of convex functions are convex, the result follows.

Finally, we conclude this paragraph by showing that $\|\hat{\alpha} - \bar{\alpha}\|_{\infty}$ has a finite upper bound, not depending on $\alpha$. This result will be used later on to find an upper bound on the probability that $\hat{\omega} := \omega - T \bar{x}$ is an element of $\bar{N}$.

**Lemma 5** Consider the mixed-integer value function $\bar{\alpha}(s) := (v(s) - \alpha)^+$, where $v(s)$ is defined as in (3), and its approximating value function $\hat{\alpha}(s)$ defined in Definition 5. Then, there exists a constant $\gamma > 0$ such that for every $\alpha \in \mathbb{R}$,

$$
\|\bar{\alpha} - \hat{\alpha}\|_{\infty} \leq \gamma,
$$

where $\|\bar{\alpha} - \hat{\alpha}\|_{\infty} := \sup_{s \in \mathbb{R}^m} |\bar{\alpha}(s) - \hat{\alpha}(s)|$.

**Proof.** Let $\hat{\alpha}_{\alpha}^{LP}$ be the LP-relaxation of $\bar{\alpha}_\alpha$. Then, by e.g. Blair and Jeroslav (1979) and Cook et al. (1986), there exists a constant $\gamma^{*} > 0$ such that $\|\bar{\alpha}_\alpha - \hat{\alpha}_\alpha^{LP}\|_{\infty} \leq \gamma^{*}$ for all $\alpha \in \mathbb{R}$. Next, let $s \in \mathbb{R}^m$ and $\alpha \in \mathbb{R}$ be arbitrarily given. We show that $|\hat{\alpha}_\alpha^{LP}(s) - \bar{\alpha}_\alpha(s)| \leq \max_{k=1, \ldots, K} |r^k|$. By definition of $\hat{\alpha}_\alpha^{LP}$ and $\bar{\alpha}_\alpha$, we have

$$
|\hat{\alpha}_\alpha^{LP}(s) - \bar{\alpha}_\alpha(s)| = \left| \left( \max_{k=1, \ldots, K} \{ q_{B^k} (B^k)^{-1} s - \alpha \}^+ \right) - \left( \max_{k=1, \ldots, K} \{ q_{B^k} (B^k)^{-1} s + \Gamma^{k}_{\alpha} - \alpha \}^+ \right) \right|.
$$

Let $k_1$ and $k_2$ be the maximizing indices of the first and second maximization problem in the equation above, respectively. Since $k_1$ and $k_2$ are maximizing indices, we can write

$$
|\hat{\alpha}_\alpha^{LP}(s) - \bar{\alpha}_\alpha(s)| = \left| \left( q_{B^{k_1}} (B^{k_1})^{-1} s - \alpha \right)^+ - \left( q_{B^{k_2}} (B^{k_2})^{-1} s + \Gamma^{k_2}_{\alpha} - \alpha \right)^+ \right| \leq \Gamma^{k_2}_{\alpha}.
$$

10
Suppose that \( q_{B^k_2} \neq 0 \). Then, by definition of \( \Gamma^{k_2}_\alpha \), we have \( \Gamma^{k_2}_\alpha = \Gamma^{k_2} \leq r^{k_2} \leq \max_{k=1,...,K} r^k \) and hence, \( |v^{\alpha}_L(s) - \hat{v}_\alpha(s)| \leq \max_{k=1,...,K} r^k \).

Next, suppose that \( q_{B^k_2} = 0 \). We consider two cases. First, suppose that \( \alpha < r^{k_2} \). Then, by definition of \( \Gamma^{k_2}_\alpha \), we have \( \Gamma^{k_2}_\alpha \leq r^{k_2} \leq \max_{k=1,...,K} r^k \) and thus, \( |v^{\alpha}_L(s) - \hat{v}_\alpha(s)| \leq \max_{k=1,...,K} r^k \). Second, suppose that \( \alpha \geq r^{k_2} \). Then, by Lemma 3, \( \Gamma^{k_2}_\alpha = \alpha \). Optimality of \( k_1 \) and \( k_2 \) and non-negativity of \( \Gamma^{k_2}_\alpha \), \( k=1,\ldots,K \), implies that \( q_{B^k_2}(B^{k_2})^{-1}s \leq q_{B^k_2}(B^{k_2})^{-1}s + \Gamma^{k_2}_\alpha \). Using \( q_{B^k_2} = 0 \), we obtain

\[
\left(q_{B^k_2}(B^{k_2})^{-1}s - \alpha\right)^+ = \left(q_{B^k_2}(B^{k_2})^{-1}s + \Gamma^{k_2}_\alpha - \alpha\right)^+ = (0 + \alpha - \alpha)^+ = 0.
\]

We conclude that \( v^{\alpha}_L(s) = \hat{v}_\alpha(s) = 0 \) and hence, \( |v^{\alpha}_L(s) - \hat{v}_\alpha(s)| \leq \max_{k=1,...,K} r^k \). Since \( s \in \mathbb{R}^m \) and \( \alpha \in \mathbb{R} \) were arbitrarily given, we obtain \( \|v^{\alpha}_L - \hat{v}\|_{\infty} \leq \max_{k=1,...,K} r^k \). Defining \( \gamma := \gamma^* + \max_{k=1,...,K} r^k \), we find that

\[
\|v - \hat{v}\|_{\infty} \leq \|v^{\alpha}_L - \hat{v}\|_{\infty} + \|v^{\alpha}_L - \hat{v}_\alpha\|_{\infty} \leq \gamma^* + \max_{k=1,...,K} r^k = \gamma.
\]

3.2.2 Periodic characterization of \( \tilde{v}_\alpha - \hat{v}_\alpha \). We now give a characterization of the difference between the original value function \( v_\alpha \) and its approximation \( \hat{v}_\alpha \). Using Proposition 1, we identify subsets of \( \mathbb{R}^m \) on which both \( v_\alpha \) and \( \hat{v}_\alpha \) attain their value through the same basis matrix \( B^k \). We show that on these subsets, \( v_\alpha - \hat{v}_\alpha \) is a periodic function with mean zero.

Fix \( k = 1,\ldots,K \). From Proposition 1, we know that for \( s \in \Lambda^k(d^k) \), we have

\[
\tilde{v}_\alpha(s) = (q_{B^k}(B^{k})^{-1}s + \psi^k(s) - \alpha)^+, \quad \alpha \in \mathbb{R}.
\]

That is, \( \tilde{v}_\alpha(s) \) attains its value through the basis matrix \( B^k \). Note that it is not necessarily the case that \( \tilde{v}_\alpha(s) \) attains its value through \( B^k \) as well, since a large value for \( \Gamma^{k_1}_l \), \( l \neq k \), may dominate the maximum defining \( \tilde{v}_\alpha(s) \). Nevertheless, we can find a vector \( \sigma^k \in \Lambda^k(d^k) \) such that \( \tilde{v}_\alpha(s) \) attains its value through \( B^k \), for all \( s \in \sigma^k + \Lambda^k \subset \Lambda^k(d^k) \).

**Lemma 6** Consider the mixed-integer value function \( \hat{v}_\alpha(s) := (v(s) - \alpha)^+ \), with \( v(s) \) defined in (3), and its approximation \( \tilde{v}_\alpha(s) \), defined in Definition 5. Moreover, let \( B^k \), \( \psi^k \), \( \Lambda^k \), and \( d^k \), \( k = 1,\ldots,K \), denote the basis matrices, \( B^k \)-periodic functions, closed convex polyhedral cones, and distances, respectively, of Proposition 1. Then, for every \( k = 1,\ldots,K \), there exists \( \sigma^k \in \Lambda^k(d^k) \) such that for all \( s \in \sigma^k + \Lambda^k \subset \Lambda^k(d^k) \),

\[
\hat{v}_\alpha(s) = \left(q_{B^k}(B^{k})^{-1}s + \Gamma^{k}_\alpha - \alpha\right)^+, \quad \alpha \in \mathbb{R},
\]

Moreover, there exists \( b^k \in \mathbb{R}^m_+ \) such that \( \sigma^k + \Lambda^k = \{ t \in \mathbb{R}^m : (B^k)^{-1}t \geq b^k \} \).

**Proof.** From duality theory we know that for the LP-relaxation \( v^{LP} \) of the mixed-integer value function \( v \) defined in (3), we have \( v^{LP}(s) = \max_{k=1,...,K} \{ q_{B^k}(B^{k})^{-1}s \} \), \( s \in \mathbb{R}^m \) and that for every \( k = 1,\ldots,K \), we have \( v^{LP}(s) = q_{B^k}(B^{k})^{-1}s \), \( s \in \Lambda^k \). Fix \( k = 1,\ldots,K \). Then, the above implies that for all \( l \neq k \), we have \( q_{B^k}(B^{k})^{-1}s \geq q_{B^l}(B^{l})^{-1}s \), \( s \in \Lambda^k \).

Fix \( l \neq k \). Suppose that \( q_{B^k}(B^{k})^{-1} = q_{B^l}(B^{l})^{-1} \). Then from Proposition 1 we know that \( \psi^k = \psi^l \). Hence, we also have \( \Gamma^{k}_\alpha = \Gamma^{l}_\alpha \) for all \( \alpha \in \mathbb{R} \) and thus

\[
(q_{B^k}(B^{k})^{-1}s + \Gamma^{k}_\alpha - \alpha)^+ = (q_{B^l}(B^{l})^{-1}s + \Gamma^{l}_\alpha - \alpha)^+, \quad s \in \Lambda^k, \quad \alpha \in \mathbb{R}.
\]
Next, suppose that \( q_{B^k}(B^k)^{-1} \neq q_{B^l}(B^l)^{-1} \). This implies that there exists some \( s^* \in \Lambda^k(d^k) \) such that \( q_{B^k}(B^k)^{-1}s^* > q_{B^l}(B^l)^{-1}s^* \). Fix such an \( s^* \). We now distinguish two cases.

First, suppose that \( q_{B^l} \neq 0 \). Then, \( \Gamma^l_\alpha = \Gamma^l \leq r^l \) for all \( \alpha \in \mathbb{R} \). For a large enough scalar \( \gamma \geq 1 \), we find \( q_{B^k}(B^k)^{-1}(\gamma s^*) > q_{B^l}(B^l)^{-1}(\gamma s^*) + r^l \). Observing that \( \Gamma^k_\alpha \geq 0 \) for all \( \alpha \in \mathbb{R} \), this implies that

\[
(q_{B^k}(B^k)^{-1}(\gamma s^*) + \Gamma^k_\alpha - \alpha)^+ \geq (q_{B^l}(B^l)^{-1}(\gamma s^*) + \Gamma^l_\alpha - \alpha)^+, \quad \alpha \in \mathbb{R}. \tag{8}
\]

Second, suppose that \( q_{B^l} = 0 \). If \( \alpha < r^l \), then by definition, \( \Gamma^l_\alpha \leq r^l \) holds for all \( \alpha \in \mathbb{R} \) and hence, (8) holds true. Conversely, if \( \alpha \geq r^l \), then \( \Gamma^l_\alpha = \alpha \) and hence, \( (q_{B^l}(B^l)^{-1}(\gamma s^*) + \Gamma^l_\alpha - \alpha)^+ = (0 + \alpha - \alpha)^+ = 0 \). It follows that (8) holds.

Now, define \( \sigma^{kl} := \gamma s^* \). We have \( \sigma^{kl} \in \Lambda^k(d^k) \) since \( s^* \in \Lambda^k(d^k) \) and \( \gamma \geq 1 \). Using (7), (8), and the inequality \( q_{B^k}(B^k)^{-1}s \geq q_{B^l}(B^l)^{-1}s \) for \( s \in \Lambda^k \), we obtain

\[
(q_{B^k}(B^k)^{-1}s + \Gamma^k_\alpha - \alpha)^+ \geq (q_{B^l}(B^l)^{-1}s + \Gamma^l_\alpha - \alpha)^+, \quad s \in \sigma^{kl} + \Lambda^k, \quad \alpha \in \mathbb{R}.
\]

Since this inequality holds for any \( l \neq k \), it follows that for any \( s \in \bigcap_{l \neq k} (\sigma^{kl} + \Lambda^k) \) and \( \alpha \in \mathbb{R} \), we have \( \hat{\nu}_\alpha(s) = (q_{B^k}(B^k)^{-1}s + \Gamma^k_\alpha - \alpha)^+ \). Note that we can write \( \sigma^{kl} + \Lambda^k = \{ t \in \mathbb{R}^n : (B^k)^{-1}t \geq b^{kl} \} \), with \( b^{kl} := (B^k)^{-1}\sigma^{kl} \). It follows that for \( b^k \) the componentwise maximum of \( b^{kl} , k \neq l \), and \( \sigma^k := B^k b^k \), we have

\[
\bigcap_{l \neq k} (\sigma^{kl} + \Lambda^k) = \{ t \in \mathbb{R}^n : (B^k)^{-1}t \geq b^k \} = \sigma^k + \Lambda^k.
\]

We conclude that

\[
\hat{\nu}_\alpha(s) = (q_{B^k}(B^k)^{-1}s + \Gamma^k_\alpha - \alpha)^+, \quad s \in \sigma^k + \Lambda^k.
\]

Using Lemma 6, we derive a characterization of \( \nu_\alpha(s) - \hat{\nu}_\alpha(s) \) on subsets of its domain in terms of periodic functions minus their mean value.

**Proposition 2** Consider the setting of Lemma 6. Then, for all \( s \in \sigma^k + \Lambda^k \), and \( \alpha \in \mathbb{R} \),

\[
\bar{v}_\alpha(s) - \hat{\nu}_\alpha(s) = \phi^k_\alpha(s) - \nu_{\phi^k_\alpha}, \quad \text{if } q_{B^k} = 0,
\]

\[
\bar{v}_\alpha(s) - \hat{\nu}_\alpha(s) = \psi^k(s) - \Gamma^k_\alpha, \quad \text{if } q_{B^k} \neq 0 \text{ and } q_{B^k}(B^k)^{-1}s \geq \alpha,
\]

\[
\bar{v}_\alpha(s) - \hat{\nu}_\alpha(s) = 0, \quad \text{if } q_{B^k} \neq 0 \text{ and } q_{B^k}(B^k)^{-1}s \leq \alpha - r^k.
\]

where for every \( \alpha \in \mathbb{R} \), \( \phi^k_\alpha(s) = (\psi^k(s) - \alpha)^+ \) is a \( B^k \)-periodic function with “average value” \( \nu_{\phi^k_\alpha} \) defined as

\[
\nu_{\phi^k_\alpha} = \frac{1}{p_k} \int_0^{p_k} \cdots \int_0^{p_k} (\psi^k(s) - \alpha)^+ ds_1 \cdots ds_m,
\]

where \( p_k := \lfloor \det(B^k) \rfloor \).

**Proof.** Fix \( k \) and suppose that \( q_{B^k} \neq 0 \). Using the fact that \( \sigma^k + \Lambda^k \subseteq \Lambda^k(d^k) \), it immediately follows from Proposition 1 and Lemma 6 that (10) and (11) hold. To prove (9), suppose that \( q_{B^k} = 0 \) and let \( s \in \sigma^k + \Lambda^k \). It follows from Proposition 1 that

\[
\bar{v}_\alpha(s) = (q_{B^k}(B^k)^{-1}s + \psi^k(s) - \alpha)^+ = (\psi^k(s) - \alpha)^+ = \phi^k_\alpha(s).
\]
Moreover, by Lemma 6 it holds that

\[
\hat{\alpha}(s) = \left( q_{B^k}(B^k)^{-1}s + \Gamma^k - \alpha \right)^+ = \left( \Gamma^k - \alpha \right)^+ = \nu_{\phi_k}.
\]

Together, this yields \( \hat{\alpha}(s) - \hat{\alpha}(s) = \phi^k_\alpha(s) - \nu_{\phi_\alpha} \).

\[
\square
\]

In the proof above, consider the case with \( q_{B^k} = 0 \). By definition of \( \Gamma^k_\alpha \), \( k = 1, \ldots, K \), we can write \( \hat{\alpha}(s) - \hat{\alpha}(s) = \phi^k_\alpha(s) - \nu_{\phi_\alpha} \), for all \( s \in \sigma^k + \Lambda^k \). Now suppose that we had naively defined \( \Gamma^k_\alpha := \Gamma^k \). Then, such a periodic characterization would only hold for \( q_{B^k}(B^k)^{-1}s \geq \alpha \) and \( q_{B^k}(B^k)^{-1}s \leq \alpha - r^k \), just like in the case that \( q_{B^k} \neq 0 \). As will be seen shortly, we need to cover the remaining case, i.e., \( \alpha - r^k < q_{B^k}(B^k)^{-1}s < \alpha \) by a hyperslice. Since \( q_{B^k} = 0 \), the inequalities reduce to \( \alpha - r^k < 0 < \alpha \). If \( \alpha \) is such that these inequalities hold, then this implies that we need to cover \( \mathbb{R}^m \) by finitely many hyperslices, which is impossible. This problem is the reason why we let \( \Gamma^k_\alpha \) depend on \( \alpha \) for the case \( q_{B^k} = 0 \).

### 3.2.3 Covering \( \mathcal{N}_\alpha \) by hyperslices.

As has been discussed before, in order to derive asymptotically converging error bounds for \( ||\hat{Q}_\beta - \hat{Q}_\beta||_\infty \) we need to cover \( \mathcal{N}_\alpha \) by finitely many hyperslices. Here, \( \mathcal{N}_\alpha \) is the subset of the domain of \( \hat{\alpha} \) and \( \hat{\alpha} \) on which the difference \( \hat{\alpha}(s) - \hat{\alpha}(s) \) cannot be written as a periodic function minus its mean value. That is, \( \mathcal{N}_\alpha \) is the complement of the subsets considered in Proposition 2.

**Definition 6** Consider the mixed-integer value function \( \hat{\alpha}(s) := (v(s) - \alpha)^+ \), where \( v(s) \) is defined in (3), and let \( B^k, q_{B^k}, r^k, \) and \( \Lambda^k, k = 1, \ldots, K \), denote the dual feasible basis matrices, second-stage cost vectors, constants, and polyhedral cones of Proposition 1, respectively. Moreover, let \( \sigma^k, k = 1, \ldots, K \), be the vectors from Lemma 6. For every \( \alpha \in \mathbb{R} \), we define the set \( \mathcal{N}_\alpha \subseteq \mathbb{R}^m \) by \( \mathcal{N}_\alpha := \mathcal{N}_\alpha^0 \cup \mathcal{N}_\alpha^1 \), where

\[
\mathcal{N}_\alpha^0 := \mathbb{R}^m \backslash \bigcup_{k=1}^K \left( \sigma^k + \Lambda^k \right),
\]

\[
\mathcal{N}_\alpha^1 := \bigcup_{k:q_{B^k} \neq 0} \left\{ s \in \sigma^k + \Lambda^k : \alpha - r^k < q_{B^k}(B^k)^{-1}s < \alpha \right\}.
\]

From Romeijnders, Schultz et al. (2016) we know that we can cover \( \mathcal{N}_\alpha^0 \) by finitely many hyperslices. Moreover, we prove that we can cover \( \mathcal{N}_\alpha^1 \) by finitely many hyperslices as well.

**Definition 7** Consider the mixed-integer value function \( \hat{\alpha}(s) := (v(s) - \alpha)^+ \), where \( v(s) \) is defined in (3), and let \( B^k, q_{B^k}, \) and \( r^k, k = 1, \ldots, K \), denote the dual feasible basis matrices, second-stage cost vectors, and constants of Proposition 1, respectively, and \( b^k, k = 1, \ldots, K \), the translation vectors of Lemma 6. Then, for every \( k = 1, \ldots, K \) and \( j = 1, \ldots, m \), let \( a_{jk} \) denote the \( j \)th row of \( (B^k)^{-1} \) and \( \delta_{jk} \) the \( j \)th component of \( b^k \). For every \( \alpha \in \mathbb{R} \), we define the hyperslices \( H^j_k \) and \( \hat{H}_\alpha^k \) as

\[
H^j_k := \{ t \in \mathbb{R}^m : 0 \leq a_{jk}t \leq \delta_{jk} \},
\]

\[
\hat{H}_\alpha^k := \{ t \in \mathbb{R}^m : \alpha - r^k \leq q_{B^k}(B^k)^{-1}t \leq \alpha \}.
\]

**Lemma 7** Consider the set \( \mathcal{N}_\alpha \) from Definition 6 and the hyperslices from Definition 7.
We show that \( \hat{\tilde{\beta}} \) solved using convex optimization techniques.

**3.2.4 Convex approximation \( \hat{\tilde{Q}}_\beta \) of the recourse function \( \tilde{Q}_\beta \).** Using the approximating value function \( \hat{\tilde{\beta}} \), we will define an approximation \( \hat{\tilde{Q}}_\beta \) of the recourse function \( \tilde{Q}_\beta \).

We show that \( \hat{\tilde{Q}}_\beta \) is convex, such that the resulting approximating recourse model can be solved using convex optimization techniques.

**Definition 8** Consider the recourse function \( \tilde{Q}_\beta \), defined as

\[
\tilde{Q}_\beta(x, \alpha) = \frac{1}{\beta} \mathbb{E}_\omega \left[ \tilde{v}_\alpha(\omega - T \beta x) \right],
\]

where the mixed-integer value function \( \tilde{v}_\alpha \) is defined as \( \tilde{v}_\alpha(s) := (v(s) - \alpha)^+ \), with \( v(s) \) defined as in (3). Furthermore, let \( B^k \), \( q_{B^k} \), and \( \psi^k \) be the basis matrices, cost vectors, and \( B^k \)-periodic functions from Proposition 1. Then, we define the approximating recourse function \( \hat{\tilde{Q}}_\beta \) as

\[
\hat{\tilde{Q}}_\beta(x, \alpha) := \frac{1}{\beta} \mathbb{E}_\omega \left[ \hat{v}_\alpha(\omega - T \beta x) \right],
\]

where the approximating mixed-integer value function \( \hat{v}_\alpha(s) \) is defined in Definition 5.

**Lemma 8** The function \( \hat{\tilde{Q}}_\beta \), defined in Definition 8, is convex in \( (x, \alpha) \).

**Proof.** The result follows immediately from convexity of \( \hat{v}_\alpha \), which is proven in Lemma 4. \( \square \)

**3.3 Total variation bounds.** In this subsection, we derive a uniform upper bound on \( \| \hat{\tilde{Q}}_\beta - \tilde{Q}_\beta \|_{\infty} \). This error bound will depend on the total variations of the density functions of the random variables in the model. We first define this notion of total variations, along with some related concepts.

**Definition 9** 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 write \(|\Delta f| := |\Delta f(\mathbb{R})|\).

To ensure validity of the total variation error bound derived in this section, we need that the conditional density functions of the random variables in the model are of bounded variation.

**Definition 10** A function \( f : \mathbb{R} \to \mathbb{R} \) is of bounded variation if \(|\Delta f| < +\infty\). We let \( \mathcal{F} \) denote the set of one-dimensional probability density functions \( f \) of bounded variation.

**Definition 11** For a vector \( x \in \mathbb{R}^n \), we define \( x_{-i} \) to be the vector \( x \) without the \( i \)th element. That is,

\[ x_{-i} = (x_1, x_2, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n) \in \mathbb{R}^{n-1}. \]

**Definition 12** For every \( i = 1, \ldots, m \) and \( x_{-i} \in \mathbb{R}^{m-1} \), define the \( i \)th conditional density function \( f_i(x_{-i}) \) of the \( m \)-dimensional joint pdf \( f \) as

\[ f_i(x_{-i}) = \frac{f(x)}{f_i(x_{-i})}, \quad \text{if } f_{-i}(x_{-i}) > 0. \]

Furthermore, \( f_i(x_{-i}) \) is undefined if \( f_{-i}(x_{-i}) = 0 \).

**Definition 13** Let \( \mathcal{H}^m \) denote the set of all \( m \)-dimensional joint pdfs \( f \) whose conditional density functions \( f_i(x_{-i}) \) are of bounded variation for all values of \( x_{-i} \) for which they are defined.

### 3.3.1 A total variation error bound.

The main strategy for deriving our upper bound on \(||\hat{Q}_\beta - \hat{Q}_\beta||_\infty||\) is to consider several subsets of the set of values that \( \hat{\omega} = \omega - Tx \) can take separately. On one type of subsets, we have that \( \hat{v}_\alpha - \tilde{v}_\alpha \) is a periodic function with mean zero. This periodicity can be exploited to bound the expectation of \( \hat{v}_\alpha(\hat{\omega}) - \tilde{v}_\alpha(\hat{\omega}) \), given that \( \hat{\omega} \) is in such a subset. The complement of these subsets is given by \( \mathcal{N}_\alpha \). We show that the total variation error derived in this section can be bounded. Together, this yields an upper bound on \(||\hat{Q}_\beta - \hat{Q}_\beta||_\infty||\).

Let \( f \in \mathcal{H}^m \), fix \( x \in X \) and \( \alpha \in \mathcal{R} \) and define \( \hat{\omega} := \omega - Tx \). Then, \( \hat{\omega} \) has joint pdf \( g \in \mathcal{H}^m \) defined as \( g(s) = f(s + Tx) \). We consider \( |\hat{Q}_\beta(\hat{\omega}, \alpha) - \hat{Q}_\beta(\hat{\omega}, \alpha)| \) and observe that by Proposition 2 and the triangle inequality,

\[
|\hat{Q}_\beta(x, \alpha) - \hat{Q}_\beta(x, \alpha)| = \frac{1}{\beta} \left| \mathbb{E}_\omega \left[ \hat{v}_\alpha(\hat{\omega}) - \tilde{v}_\alpha(\hat{\omega}) \right] \right|
\]

\[
= \frac{1}{\beta} \left| \int_{\mathbb{R}^m} (\hat{v}_\alpha(s) - \tilde{v}_\alpha(s)) g(s) ds \right|
\]

\[
\leq \frac{1}{\beta} \sum_{k : \sigma_k \neq \emptyset} \left( \left| \int_{A_k^+} (\psi^k(s) - \Gamma^k) g(s) ds \right| + \left| \int_{A_k^-} 0 \cdot g(s) ds \right| \right) \tag{12}
\]

\[
+ \frac{1}{\beta} \sum_{k : \sigma_k = \emptyset} \left| \int_{\sigma_k + A_k} (\phi^k(\hat{s}) - \nu_{\phi^k} g(s) ds \right| \tag{13}
\]
where $\mathcal{A}_\alpha^{k+} := \{ s \in \sigma^k + \Lambda^k \mid q_{B_k}(B^{k})^{-1} s \geq \alpha \}$ and $\mathcal{A}_\alpha^{k-} := \{ s \in \sigma^k + \Lambda^k \mid q_{B_k}(B^{k})^{-1} s \leq \alpha - \nu \}$, such that $\left( \bigcup_{j \in \mathbb{N}_k} \mathcal{A}_\alpha^{k+} \cup \mathcal{A}_\alpha^{k-}\right) \cup \left( \bigcup_{j \in \mathbb{N}_k} \sigma^k + \Lambda^k \right) \cap \mathcal{N}_\alpha = \mathbb{R}^m$. We will show that we can apply Lemma 9, which is proven in Romeijnders, Schultz et al. (2016), to obtain upper bounds on (12) and (13).

**Lemma 9** Let $\psi : \mathbb{R}^m \to \mathbb{R}$ be a $B$-periodic function with finite mean value $\nu := p^{-m} \int_0^1 \cdots \int_0^1 \psi(x)dx_1 \cdots dx_m$, where $B \in \mathbb{Z}^{m \times m}$ is nonsingular and $p = \det(B)$. Assume that there exists $\gamma > 0$ such that $|\psi(x) - \nu| \leq \gamma$ for all $x \in \mathbb{R}^m$. Then, for every convex set $\Lambda \subset \mathbb{R}^m$ and every continuous random vector $\omega$ with joint probability density function $f \in \mathcal{H}^m$, we have

$$\left| \int_\Lambda (\psi(x) - \nu) f(x)dx \right| \leq \frac{1}{2} \gamma |\det(B)| \sum_{i=1}^m \mathbb{E}_{\omega^{-i}} |\Delta| f_i(\cdot |\omega_{-i})|.$$ 

Considering (14), observe that $|\hat{v}_\alpha(s) - \hat{v}_\alpha(s)|$ is bounded by Lemma 5 and that $\mathcal{N}_\alpha$ can be covered by finitely many hyperslices by Lemma 7. Using these facts, we can apply Lemma 10 to obtain an upper bound on (14).

**Lemma 10** Let $\eta \in \mathbb{R}$, $\delta > 0$, and $a \in \mathbb{R}^m \setminus \{0\}$ be given. Then, there exists $D > 0$, which does not depend on $\eta$, such that for every continuous random vector $\omega$ with joint pdf $f \in \mathcal{H}^m$, we have

$$\mathbb{P} \left\{ \eta \leq a^T \omega \leq \eta + \delta \right\} \leq D \sum_{i=1}^m \mathbb{E}_{\omega^{-i}} |\Delta| f_i(\cdot |\omega_{-i})|.$$ 

**Proof.** Define $H := \{ x \in \mathbb{R}^m : \eta \leq a^T x \leq \eta + \delta \}$ such that $\mathbb{P}\{\eta \leq a^T \omega \leq \eta + \delta\} = \int_H f(x)dx$. Since $a \neq 0$, there exists $j = 1, \ldots, m$ such that $a_j \neq 0$. By conditioning on $\omega_{-j}$, we have

$$\mathbb{P} \left\{ \eta \leq a^T \omega \leq \eta + \delta \right\} = \int_{\mathbb{R}^{m-1}} \int_{H_j(x_{-j})} f_j(x_j | x_{-j})dx_j f_{-j}(x_{-j})dx_{-j},$$

where

$$H_j(x_{-j}) = \left\{ x_j \in \mathbb{R} : \frac{\eta}{a_j} - \frac{a_j^T x_{-j}}{a_j} \leq x_j \leq \frac{\eta + \delta}{a_j} - \frac{a_j^T x_{-j}}{a_j} \right\}, \quad \text{if } a_j > 0,$$

$$H_j(x_{-j}) = \left\{ x_j \in \mathbb{R} : \frac{\eta + \delta}{a_j} - \frac{a_j^T x_{-j}}{a_j} \leq x_j \leq \frac{\eta}{a_j} - \frac{a_j^T x_{-j}}{a_j} \right\}, \quad \text{if } a_j < 0.$$ 

Note that for every $x_{-j} \in \mathbb{R}^{m-1}$ and $a_j \neq 0$, the interval length $|H_j(x_{-j})|$ equals $|a_j|^{-1} \delta$. Since $f_j(x_j | x_{-j}) \leq \frac{1}{2} |\Delta| f_j(\cdot | x_{-j})$, it follows that

$$\mathbb{P} \left\{ \eta \leq a^T \omega \leq \eta + \delta \right\} \leq \int_{\mathbb{R}^{m-1}} \frac{1}{2} |a_j|^{-1} \delta |\Delta| f_j(\cdot | x_{-j}) f_{-j}(x_{-j})dx_{-j}$$

$$= \frac{1}{2} |a_j|^{-1} \delta \mathbb{E}_{\omega_{-j}} |\Delta| f_j(\cdot | x_{-j})|.$$ 

Defining $D := \frac{1}{2} |a_j|^{-1} \delta$, which does not depend on $\eta$, and observing that $\mathbb{E}_{\omega_{-i}} [\Delta f_i(\cdot | x_{-i})] \geq 0$ for $i = 1, \ldots, m$, we conclude that

$$\mathbb{P} \left\{ \eta \leq a^T \omega \leq \eta + \delta \right\} \leq D \sum_{i=1}^m \mathbb{E}_{\omega_{-i}} [\Delta f_i(\cdot | \omega_{-i})].$$
Remark 1  
Observe that the bound in Lemma 10 can be improved by minimizing \(\mathbb{E}_{\omega_{-i}}[\|\Delta f_{i}(\cdot | \omega_{-i})]\) over all \(i\) with \(a_{i} \neq 0\), rather than taking the summation over \(i = 1, \ldots, m\). However, we choose to use \(\sum_{i=1}^{m} \mathbb{E}_{\omega_{-i}}[\|\Delta f_{i}(\cdot | \omega_{-i})]\), since this expression is also present in the bound in Lemma 9. This allows us to write the bound in Theorem 1 more conveniently. Since our focus in this section is on deriving an asymptotically converging bound, i.e. we are mainly interested in the asymptotic behavior of the bound, we prefer larger but more concisely written bounds.

Combining the upper bounds on (12), (13), and (14), we can derive an upper bound on \(\bar{Q}_{\beta}(x, \alpha) - \hat{Q}_{\beta}(x, \alpha)\).

Theorem 1  
Consider the mixed-integer recourse function

\[
\bar{Q}_{\beta}(x, \alpha) = \frac{1}{\beta} \mathbb{E}_{\omega} [\bar{v}_{\alpha}(\omega - Tx)], \quad (x, \alpha) \in \mathbb{R}^{n_1} \times \mathbb{R},
\]

where \(\bar{v}_{\alpha}(s) = (v(s) - \alpha)^{+}\), for all \((s, \alpha) \in \mathbb{R}^{m} \times \mathbb{R}\), with \(v\) the mixed-integer second-stage value function

\[
v(s) := \min_y \{qy : Wy = s, \; y \in \mathbb{Z}^{n_2} \times \mathbb{R}^{n_3}\}, \quad s \in \mathbb{R}^{m}.
\]

Let \(\hat{Q}_{\beta}\) denote the convex approximation of \(\bar{Q}_{\beta}\), defined as

\[
\hat{Q}_{\beta}(x, \alpha) = \frac{1}{\beta} \mathbb{E}_{\omega} [\hat{v}_{\alpha}(\omega - Tx)], \quad (x, \alpha) \in \mathbb{R}^{n_1} \times \mathbb{R},
\]

where \(\hat{v}_{\alpha}\) is the convex approximation of \(\bar{v}_{\alpha}\) from Definition 5. Then, there exists a constant \(C > 0\) such that for every continuous random vector \(\omega\) with joint probability density function \(f \in \mathcal{H}^{m}\),

\[
\|\bar{Q}_{\beta} - \hat{Q}_{\beta}\|_{\infty} \leq \frac{1}{\beta} C \sum_{i=1}^{m} \mathbb{E}_{\omega_{-i}}[\|\Delta f_{i}(\cdot | \omega_{-i})]\],
\]

where \(\|\bar{Q}_{\beta} - \hat{Q}_{\beta}\|_{\infty} := \sup_{x \in \mathbb{R}^{n_1}, \alpha \in \mathbb{R}} |\bar{Q}_{\beta}(x, \alpha) - \hat{Q}_{\beta}(x, \alpha)|\).

Proof. Let \(B^{k}, q_{B^{k}}, \psi^{k}, \Gamma^{k},\) and \(\Gamma_{\alpha}^{k}\) be the basis matrices, cost vectors, \(B^{k}\)-periodic functions, mean values, and approximation functions, respectively, from Definition 4. We fix \(x \in \mathbb{R}^{n_1}\) and \(\alpha \in \mathbb{R}\) and define \(\hat{\omega} := \omega - Tx\), which is a random vector with joint pdf \(g \in \mathcal{H}^{m}\), defined by \(g(s) = f(s + Tx)\). By (12)-(14), we have

\[
\left|\bar{Q}_{\beta}(x, \alpha) - \hat{Q}_{\beta}(x, \alpha)\right| \leq \frac{1}{\beta} \sum_{k : q_{B^{k}} \neq 0} \left| \int_{A_{\alpha}^{k}} (\psi^{k}(s) - \Gamma^{k})g(s)ds \right|
\]

\[
+ \frac{1}{\beta} \sum_{k : q_{B^{k}} = 0} \left| \int_{A_{\alpha}^{k}} (\phi^{k}_{\alpha}(s) - \nu_{\phi^{k}_{\alpha}})g(s)ds \right|
\]

\[
+ \frac{1}{\beta} \left| \int_{\mathcal{N}_{\alpha}} (\bar{v}_{\alpha}(s) - \hat{v}_{\alpha}(s))g(s)ds \right|,
\]

where \(A_{\alpha}^{k}\) is the set defined as \(A_{\alpha}^{k} := \{s \in \sigma^{k} + \Lambda^{k} \mid q_{B^{k}}(B^{k})^{-1}s \geq \alpha\}\), \(\sigma^{k}\) and \(\Lambda^{k}\), \(k = 1, \ldots, K\), are the vectors and convex polyhedral cones from Lemma 6, \(\phi^{k}_{\alpha}\) and \(\nu_{\phi^{k}_{\alpha}}\) are the
periodic function and its mean value from Proposition 2, and \(N_\alpha\) is the set from Definition 6.

We first analyze (15). Fix \(k\) with \(q_{B^k} \neq 0\). Note that by Proposition 2, \(\psi^k(s) - \Gamma^k\) is a \(B^k\)-periodic function with mean zero on the corresponding set \(A^k_\alpha\). Since this set is the intersection of two convex sets (a shifted convex cone and a half-space), it is itself convex. Furthermore, from Proposition 1(iv) and the definition of \(\Gamma^k\) it follows that \(\Gamma^k\) is finite and that \(|\psi^k(s) - \Gamma^k| \leq r^k\) for all \(s \in \mathbb{R}^m\). All assumptions of Lemma 9 are satisfied and we use the lemma to obtain

\[
\left| \int_{A^k_\alpha} (\psi^k(s) - \Gamma^k)g(s)ds \right| \leq \frac{1}{2} r^k \left| \det(B^k) \right| \sum_{i=1}^{m} \mathbb{E}_{\omega_{-i}} [||\Delta g_i(\cdot|\omega_{-i})||] = \frac{1}{2} r^k \left| \det(B^k) \right| \sum_{i=1}^{m} \mathbb{E}_{\omega_{-i}} [||\Delta f_i(\cdot|\omega_{-i})||],
\]

where the equality follows from the observation that \(g_i(\omega_{-i}) = g_i(\omega_{-i}(T x)) = g_i(\omega_{-i}(T x)_{-i}) = f_i(\omega_{-i})\).

We treat \(k\) with \(q_{B^k} = 0\) and observe that \(\sigma^k + \Lambda^k\) is a convex set. Note that \(\phi^k_\alpha\) is a \(B^k\)-periodic function with mean value \(\nu_{\phi^k_\alpha}\). From Proposition 1(iv), we have \(0 \leq \psi^k(s) \leq r^k\), \(s \in \mathbb{R}^m\). Hence, for every \(s \in \mathbb{R}^m\), we have \(0 \leq \phi^k_\alpha(s) = (\psi^k(s) - \alpha)^+ \leq (r^k - \alpha)^+\). Since \(\nu_{\phi^k_\alpha}\) is the mean value of \(\phi^k_\alpha\), this implies that \(0 \leq \nu_{\phi^k_\alpha} \leq (r^k - \alpha)^+\), and hence, \(\nu_{\phi^k_\alpha}\) is finite. Moreover, these final inequalities imply that \(\nu_{\phi^k_\alpha} - \phi^k_\alpha(s) \leq (r^k - \alpha)^+ - (\psi^k(s) - \alpha)^+ \leq r^k\) and \(\nu_{\phi^k_\alpha} - \phi^k_\alpha(s) \geq (0 - \alpha)^+ - (\psi^k(s) - \alpha)^+ \geq -r^k\) for all \(s \in \mathbb{R}^m\). Hence, \(|\phi^k_\alpha(s) - \nu_{\phi^k_\alpha}| \leq r^k\).

We have shown that all assumptions of Lemma 9 are satisfied and applying it yields an upper bound on (16):

\[
\left| \int_{\sigma^k + \Lambda^k} (\phi^k_\alpha(s) - \nu_{\phi^k_\alpha})g(s)ds \right| \leq \frac{1}{2} r^k \left| \det(B^k) \right| \sum_{i=1}^{m} \mathbb{E}_{\omega_{-i}} [||\Delta g_i(\cdot|\omega_{-i})||] = \frac{1}{2} r^k \left| \det(B^k) \right| \sum_{i=1}^{m} \mathbb{E}_{\omega_{-i}} [||\Delta f_i(\cdot|\omega_{-i})||].
\]

Finally, we derive an upper bound on (17). By Lemma 5, there exists some \(\gamma > 0\) such that \(||\hat{v}_\alpha - \tilde{v}_\alpha||_{\infty} \leq \gamma\) for all \(\alpha \in \mathbb{R}\). Substituting this into (17) yields

\[
\left| \int_{N_\alpha} (\tilde{v}_\alpha(s) - \hat{v}_\alpha(s))g(s)ds \right| \leq \gamma \int_{N_\alpha} g(s)ds = \gamma \mathbb{P}\{\tilde{\omega} \in N_\alpha\}.
\]

Let \(H_j^k\) and \(\tilde{H}_j^k\), \(j = 1, \ldots, m, k = 1, \ldots, K\) be the hyperslices from Definition 7. From Lemma 7 we know that

\[
N_\alpha \subset \left( \bigcup_{k=1}^{K} \bigcup_{j=1}^{m} H_j^k \right) \cup \left( \bigcup_{k:q_{B^k} \neq 0} \tilde{H}_j^k \right).
\]

Hence,

\[
\mathbb{P}\{\tilde{\omega} \in N_\alpha\} \leq \mathbb{P}\{\tilde{\omega} \in \left( \bigcup_{k=1}^{K} \bigcup_{j=1}^{m} H_j^k \right) \cup \left( \bigcup_{k:q_{B^k} \neq 0} \tilde{H}_j^k \right)\} \leq \sum_{k=1}^{K} \sum_{j=1}^{m} \mathbb{P}\{\tilde{\omega} \in H_j^k\} + \sum_{k:q_{B^k} \neq 0} \mathbb{P}\{\tilde{\omega} \in \tilde{H}_j^k\}.
\]
For each \( k = 1, \ldots, K \), and \( j = 1, \ldots, m \), we bound \( P\{\hat{\omega} \in H^{jk}\} \) and \( P\{\hat{\omega} \in \tilde{H}^k_\alpha\} \) using Lemma 10. Hence, there exist constants \( D^{jk} > 0 \) and \( \tilde{D}^k > 0 \), such that for every \( \alpha \in \mathbb{R} \),

\[
P\{\hat{\omega} \in H^{jk}\} \leq D^{jk} \sum_{i=1}^{m} \mathbb{E}_{\omega_i} \left[ |\Delta|g_i(\cdot |\hat{\omega}_{-i})\right] = D^{jk} \sum_{i=1}^{m} \mathbb{E}_{\omega_i} \left[ |\Delta|f_i(\cdot |\omega_{-i})\right],
\]

\[
P\{\hat{\omega} \in \tilde{H}^k_\alpha\} \leq \tilde{D}^k \sum_{i=1}^{m} \mathbb{E}_{\omega_i} \left[ |\Delta|g_i(\cdot |\hat{\omega}_{-i})\right] = \tilde{D}^k \sum_{i=1}^{m} \mathbb{E}_{\omega_i} \left[ |\Delta|f_i(\cdot |\omega_{-i})\right].
\]

Hence, we obtain the upper bound

\[
\left| \int_{N_\alpha} (\tilde{v}_\alpha(s) - \hat{v}_\alpha(s))g(s)ds \right| \leq \gamma \left( \sum_{k=1}^{K} \sum_{j=1}^{m} D^{jk} + \sum_{k:q_{jk} \neq 0} \tilde{D}^k \right) \sum_{i=1}^{m} \mathbb{E}_{\omega_i} \left[ |\Delta|f_i(\cdot |\omega_{-i})\right].
\]

Collecting terms now yields

\[
\left| \hat{Q}_\beta(x, \alpha) - \tilde{Q}_\beta(x, \alpha) \right| \leq \gamma \left( \sum_{k=1}^{K} \sum_{j=1}^{m} D^{jk} + \sum_{k:q_{jk} \neq 0} \tilde{D}^k \right) \sum_{i=1}^{m} \mathbb{E}_{\omega_i} \left[ |\Delta|f_i(\cdot |\omega_{-i})\right].
\]

Defining \( C := \sum_{k=1}^{K} \left( \frac{1}{2} r^k |\det(B^k)| \right) \gamma \left( \sum_{k=1}^{K} \sum_{j=1}^{m} D^{jk} + \sum_{k:q_{jk} \neq 0} \tilde{D}^k \right) \sum_{i=1}^{m} \mathbb{E}_{\omega_i} \left[ |\Delta|f_i(\cdot |\omega_{-i})\right] \) completes the proof.

The error bound from Theorem 1 is asymptotically converging. That is, if the total variations of the one-dimensional conditional densities \( f_i(\cdot |x_{-i}) \) go to zero, then the error bound from Theorem 1 converges to zero and hence, \( \|\hat{Q}_\beta - \tilde{Q}_\beta\| \to 0 \). This result implies that any mixed-integer CVaR recourse function \( \hat{Q}_\beta \) can be approximated reasonably well by a convex approximation \( \tilde{Q}_\beta \) if the total variations of the densities of the random variables in the model are small enough.

The error bound from Theorem 1 holds for all mixed-integer CVaR recourse models. However, for special cases we may perform specialized analysis yielding tighter error bounds. We will do this in the next section for CVaR recourse models with simple integer recourse. It turns out that in these cases the error bound does not necessarily go to \( +\infty \) if \( \beta \downarrow 0 \), as in the bound of Theorem 1.

4 Two-sided simple integer CVaR models. In this section, we consider a special case of the CVaR recourse models considered in the previous section, namely, (one-dimensional) simple integer CVaR recourse models. That is, we consider the optimization problem

\[
\min_{x \in X} \left\{ cx + \hat{Q}_\beta(z), \ z = Tx \right\},
\]

where \( z \in \mathbb{R} \) are tender variables, \( \hat{Q}_\beta \) is the CVaR recourse function

\[
\hat{Q}_\beta(z) = \text{CVaR}_{\beta}[v(\omega - z)], \quad z \in \mathbb{R},
\] (18)
$\beta \in (0, 1)$ is a probability level, $\omega$ is a random variable with density function $f \in \mathcal{F}$, and $v$ is the simple integer value function

$$v(s) = \min\{q^+ y^+ + q^- y^- :$$
$$y^+ \geq \omega - s,$$
$$y^- \geq s - \omega,$$
$$y^+ \in \mathbb{Z}_+, y^- \in \mathbb{Z}_+\}$$

$$= q^+ \lceil s \rceil^+ + q^- \lfloor s \rfloor^-, \quad s \in \mathbb{R},$$

with $q^+ > 0$ and $q^- > 0$. Here, $\lceil \cdot \rceil$ and $\lfloor \cdot \rfloor$ represent the ceiling and floor function, respectively.

The motivation for considering the simple integer case is twofold. Firstly, the special structure in the model facilitates the derivation of a tighter error bound than for the general mixed-integer case in the previous section. This error bound will turn out to behave nicely with respect to the probability parameter $\beta$ in the definition of CVaR. Secondly, this well-known special case has some straightforward practical applications. For instance, production planning problems in which production only takes place in batches can be modeled using simple integer recourse models.

Convexity of the recourse function $\bar{Q}_\beta$ is not guaranteed, due to the integrality in the definition of $v$. Hence, convex optimization techniques cannot be used and the model is hard to solve. Analogously to Section 3, we will take the approach of approximating $\bar{Q}_\beta$ by a convex function $\hat{Q}_\beta$, defined as

$$\hat{Q}_\beta(z) = \text{CVaR}_\beta[\hat{v}(\omega - z)], \quad z \in \mathbb{R},$$

where $\hat{v}$ is the convex approximation of the second-stage value function used by Romeijnders, van der Vlerk et al. (2016). It is defined as

$$\hat{v}(s) = q^+ (s + 1/2)^+ + q^- (s - 1/2)^-, \quad s \in \mathbb{R}.$$  

We use the fact that CVaR is a coherent risk measure to prove that the function $\hat{Q}_\beta$ is indeed convex.

**Proposition 3** Let the function $\hat{Q}_\beta$ be defined as

$$\hat{Q}_\beta(z) = \text{CVaR}_\beta[\hat{v}(\omega - z)], \quad z \in \mathbb{R},$$

where $\beta \in (0, 1)$, $\omega$ is a random variable, and $\hat{v}$ is a convex function. Then, $\hat{Q}_\beta$ is convex in $z$.

**Proof.** From Artzner et al. (1999) we know that CVaR is a coherent risk measure. This implies that CVaR adheres to monotonicity, subadditivity and positive homogeneity. Let $z_0, z_1 \in \mathbb{R}$ and $\lambda \in [0, 1]$ be given arbitrarily and define $\bar{z} = \lambda z_0 + (1 - \lambda) z_1$. Then, using the abovementioned properties of CVaR and convexity of $\hat{v}$, we obtain

$$\hat{Q}_\beta(\bar{z}) = \text{CVaR}_\beta[\hat{v}(\omega - \bar{z})]$$

$$\leq \text{CVaR}_\beta[\lambda \hat{v}(\omega - z_0) + (1 - \lambda) \hat{v}(\omega - z_1)]$$

$$\leq \text{CVaR}_\beta[\lambda \hat{v}(\omega - z_0)] + \text{CVaR}_\beta[(1 - \lambda) \hat{v}(\omega - z_1)]$$

$$\leq \lambda \text{CVaR}_\beta[\hat{v}(\omega - z_0)] + (1 - \lambda) \text{CVaR}_\beta[\hat{v}(\omega - z_1)]$$

$$= \lambda \hat{Q}_\beta(z_0) + (1 - \lambda) \hat{Q}_\beta(z_1).$$

$\square$
To assess the quality of the convex approximation, we will derive an upper bound on the error of the approximation.

4.1 An error bound depending on the first-stage decision. In this subsection, we derive an upper bound on the approximation error \(|\hat{Q}_\beta(z) - \hat{Q}_\beta(z)|\), depending on the value of \(z = Tx\), the tender variable depending on the first-stage decision variables \(x\). In the next subsection, this bound will be generalized to a uniform error bound.

The main steps in the derivation are the following. We first give a characterization of \(\hat{Q}_\beta(z)\) and \(\hat{Q}_\beta(z)\) in terms of expected values. Next, we apply some manipulations on the error \(|\hat{Q}_\beta(z) - \hat{Q}_\beta(z)|\) to obtain an upper bound on this difference in terms of the expected value of a periodic function. Finally, we apply a result from Romeijnders, van der Vlerk et al. (2016) that provides an upper bound for the expected value of a periodic function.

We start by characterizing \(\hat{Q}_\beta\) and \(\hat{Q}_\beta\) in terms of expected values. Let \(z \in \mathbb{R}\) be given. From Rockafellar and Uryasev (2002) we know that

\[
\hat{Q}_\beta(z) = \text{CVaR}_\beta[v(\omega - z)] = \min_{\alpha \in \mathbb{R}} \left\{ \alpha + \frac{1}{\beta} \mathbb{E}_\omega [(v(\omega - z) - \alpha)^+] \right\},
\]

and

\[
\hat{Q}_\beta(z) = \text{CVaR}_\beta[\hat{v}(\omega - z)] = \min_{\alpha \in \mathbb{R}} \left\{ \hat{\alpha} + \frac{1}{\beta} \mathbb{E}_\omega [(\hat{v}(\omega - z) - \hat{\alpha})^+] \right\}.
\]

We define \(\alpha^*_\beta\) and \(\hat{\alpha}^*_\beta\) as the \(\beta\)-VaR (Value at Risk) of \(v(\omega - z)\) and \(\hat{v}(\omega - z)\), respectively. That is,

\[
\alpha^*_\beta := \min \{ \alpha \mid \mathbb{P}\{v(\omega - z) \leq \alpha\} \geq 1 - \beta \}, \tag{22}
\]

\[
\hat{\alpha}^*_\beta := \min \{ \hat{\alpha} \mid \mathbb{P}\{\hat{v}(\omega - z) \leq \hat{\alpha}\} \geq 1 - \beta \}. \tag{23}
\]

From Rockafellar and Uryasev (2002) we know that \(\alpha^*_\beta\) and \(\hat{\alpha}^*_\beta\) are the minimal optimal arguments for the minimization problems corresponding to \(\hat{Q}_\beta\) and \(\hat{Q}_\beta\), respectively. Thus, we can write \(\hat{Q}_\beta(z) = \alpha^*_\beta + \frac{1}{\beta} \mathbb{E}_\omega [(v(\omega - z) - \alpha^*_\beta)^+]\) and \(\hat{Q}_\beta(z) = \hat{\alpha}^*_\beta + \frac{1}{\beta} \mathbb{E}_\omega [(\hat{v}(\omega - z) - \hat{\alpha}^*_\beta)^+]\). Using the fact that \(\alpha^*_\beta\) and \(\hat{\alpha}^*_\beta\) are not necessarily optimal in the minimization problem for \(\hat{Q}_\beta(z)\) and \(\hat{Q}_\beta(z)\), respectively, we can find an upper bound on the difference between \(\hat{Q}_\beta(z)\) and \(\hat{Q}_\beta(z)\).

Lemma 11 Let \(\hat{Q}_\beta, \hat{Q}_\beta, v, \) and \(\hat{v}\) be the recourse function, its convex approximation, the value function, and its approximation from (18), (20), (19), and (21), respectively. Moreover, for all \(z \in \mathbb{R}\), let \(\alpha^*_\beta\) and \(\hat{\alpha}^*_\beta\) be the \(\beta\)-VaR of \(v(\omega - z)\) and \(\hat{v}(\omega - z)\) from (22) and (23), respectively. Then, for all \(z \in \mathbb{R}\) we have

\[
\hat{Q}_\beta(z) - \hat{Q}_\beta(z) \leq \frac{1}{\beta} \mathbb{P}\{\omega \in \Omega_{\beta,z}\} \mathbb{E}_\omega [\hat{v}(\omega - z) - v(\omega - z) \mid \omega \in \Omega_{\beta,z}],
\]

\[
\hat{Q}_\beta(z) - \hat{Q}_\beta(z) \leq \frac{1}{\beta} \mathbb{P}\{\omega \in \hat{\Omega}_{\beta,z}\} \mathbb{E}_\omega [\hat{v}(\omega - z) - v(\omega - z) \mid \omega \in \hat{\Omega}_{\beta,z}],
\]

where \(\Omega_{\beta,z} := \{ \omega \mid \hat{v}(\omega - z) \geq \alpha^*_\beta \}\) and \(\hat{\Omega}_{\beta,z} := \{ \omega \mid v(\omega - z) \geq \hat{\alpha}^*_\beta \}\).

Proof. We prove the first inequality. The proof for the second inequality is completely analogous. Since \(\alpha^*_\beta\) is not necessarily optimal in the minimization problem for \(\hat{Q}_\beta(z)\), it
follows that
\[
\hat{Q}_\beta(z) - Q_\beta(z) \leq \alpha_\beta^+ + \frac{1}{\beta} E_\omega[(\hat{v}(\omega - z) - \alpha_\beta^+)] - \alpha_\beta^- - \frac{1}{\beta} E_\omega[(v(\omega - z) - \alpha_\beta^-)] \\
= \frac{1}{\beta} E_\omega[(\hat{v}(\omega - z) - \alpha_\beta^+ - (v(\omega - z) - \alpha_\beta^-)].
\] (24)

We consider two cases. First, suppose that \( \hat{v}(\omega - z) < \alpha_\beta^- \). Then, \((\hat{v}(\omega - z) - \alpha_\beta^-) = 0 - (v(\omega - z) - \alpha_\beta^-) \leq 0 \). So the argument in the expectation in equation (24) is at most zero in this case. Second, suppose that \( \hat{v}(\omega - z) \geq \alpha_\beta^- \). Then, \((\hat{v}(\omega - z) - \alpha_\beta^-) = (v(\omega - z) - \alpha_\beta^-) \leq \hat{v}(\omega - z) - v(\omega - z) \). Using these facts, we can bound (24) by conditioning on \( \hat{v}(\omega - z) < \alpha_\beta^- \) and \( \hat{v}(\omega - z) \geq \alpha_\beta^- \). We obtain
\[
\hat{Q}_\beta(z) - Q_\beta(z) \leq \frac{1}{\beta} P\{\hat{v}(\omega - z) < \alpha_\beta^-\} E_\omega[\hat{v}(\omega - z) - v(\omega - z)] \\
+ \frac{1}{\beta} P\{\hat{v}(\omega - z) \geq \alpha_\beta^-\} E_\omega[\hat{v}(\omega - z) - v(\omega - z) | \hat{v}(\omega - z) \geq \alpha_\beta^-] \\
= \frac{1}{\beta} P\{\omega \in \Omega_{\beta,z}^\alpha \} E_\omega[\hat{v}(\omega - z) - v(\omega - z) | \omega \in \Omega_{\beta,z}].
\]

Our goal is to write the conditional expectations from Lemma 11 as (non-conditional) expectations of periodic functions. Note that the function \( \hat{v} - v \) is periodic on the subsets \((-\infty, -1/2] \) and \([1/2, \infty) \) of its domain. If we can show that \( \Omega_{\beta,z} \cap (z - 1/2, z + 1/2) = \emptyset \), then we can “ignore” the possibility that \( \omega - z \in (-1/2, 2/1) \) and we can split up each of the conditional expectations into two conditional expectations of a periodic function: one corresponding to \( \omega - z \leq -1/2 \) and one to \( \omega - z \geq 1/2 \). We will show that this is indeed the case under the following assumption, which is assumed to hold true in the remainder of this section.

**Assumption 1** We assume that \( \alpha_\beta^- > \max\{q^+, q^-\} \) and \( \alpha_\beta^+ > \max\{q^+, q^-\} \).

**Remark 2** A sufficient condition for Assumption 1 to hold is that \( F^{-1}(1/2) = \beta > (1/4 + 1/2) \max\{q^+, q^-\} \), as will be proven at a later stage (Lemma 16). Thus, if the support of the density function \( f \) of the random variable \( \omega \) is the entire real line, then Assumption 1 is guaranteed to hold if \( \beta \) is small enough.

**Lemma 12** Suppose that Assumption 1 holds. Then, \( \Omega_{\beta,z} = \Omega_{\beta,z}^+ \cup \Omega_{\beta,z}^- \) and \( \hat{\Omega}_{\beta,z} = \hat{\Omega}_{\beta,z}^+ \cup \hat{\Omega}_{\beta,z}^- \), where
\[
\Omega_{\beta,z}^+ := \{\omega \in \Omega_{\beta,z} | \omega - z \leq -1/2\}, \quad \Omega_{\beta,z}^- := \{\omega \in \Omega_{\beta,z} | \omega - z \geq 1/2\}, \\
\hat{\Omega}_{\beta,z}^+ := \{\omega \in \hat{\Omega}_{\beta,z} | \omega - z \leq -1/2\}, \quad \hat{\Omega}_{\beta,z}^- := \{\omega \in \hat{\Omega}_{\beta,z} | \omega - z \geq 1/2\},
\]
and \( \Omega_{\beta,z}^+ \cap \Omega_{\beta,z}^- = \emptyset \). \( \hat{\Omega}_{\beta,z}^+ \cap \hat{\Omega}_{\beta,z}^- = \emptyset \).

**Proof.** We prove the result for \( \Omega_{\beta,z} \). The proof for \( \hat{\Omega}_{\beta,z} \) is completely analogous. Suppose that the inequality \( \hat{v}(\omega - z) \geq \alpha_\beta^- \) and Assumption 1 hold. Then, it follows that \( q^+(\omega - z + 1/2)^+ + q_-(\omega - z - 1/2)^- \geq \max\{q^+, q^-\} \), which implies that either \( \omega - z \leq -1/2 \) or \( \omega - z \geq 1/2 \). Hence, \( \Omega_{\beta,z} \cap (z - 1/2, z + 1/2) = \emptyset \) and we can split up \( \Omega_{\beta,z} \) into two disjoint subsets \( \Omega_{\beta,z}^+ \) and \( \Omega_{\beta,z}^- \) as defined above.
We restrict our attention to finding an upper bound on \( \hat{Q}_\beta(z) - Q_\beta(z) \). By Lemma 11 and 12,

\[
\hat{Q}_\beta(z) - Q_\beta(z) \leq \frac{1}{\beta} \mathbb{P}(\omega \in \Omega_{\beta,z}^+) \mathbb{E}_\omega[q^+(\omega - z + 1/2) - q^+ [\omega - z] | \omega \in \Omega_{\beta,z}^+] \\
+ \frac{1}{\beta} \mathbb{P}(\omega \in \Omega_{\beta,z}^-) \mathbb{E}_\omega[q^- [\omega - z] - q^- (\omega - z - 1/2) | \omega \in \Omega_{\beta,z}^-].
\]

This can be written in terms of (non-conditional) expectations of periodic functions by introducing the random variables \( \omega_{\beta,z}^+ \sim \omega | \omega \in \Omega_{\beta,z}^+ \) and \( \omega_{\beta,z}^- \sim \omega | \omega \in \Omega_{\beta,z}^- \), with corresponding density functions \( f_{\beta,z}^+ \) and \( f_{\beta,z}^- \), respectively. We obtain

\[
\hat{Q}_\beta(z) - Q_\beta(z) \leq \frac{q^+}{\beta} \mathbb{P}(\omega \in \Omega_{\beta,z}^+) \mathbb{E}_{\omega_{\beta,z}^+}[(\omega_{\beta,z}^+ - z + 1/2) - [\omega_{\beta,z}^+ - z]] \\
+ \frac{q^-}{\beta} \mathbb{P}(\omega \in \Omega_{\beta,z}^-) \mathbb{E}_{\omega_{\beta,z}^-}[[\omega_{\beta,z}^- - z] - (\omega_{\beta,z}^- - z - 1/2)].
\]

We now have an expression in terms of expectations of periodic functions. For such expressions, Romeijnders, van der Vlerk et al. (2016) derive a total variations error bound (see Theorem 6 in their paper). Applying this result yields

\[
\hat{Q}_\beta(z) - Q_\beta(z) \leq \frac{q^+}{2\beta} \mathbb{P}(\omega \in \Omega_{\beta,z}^+) h(|\Delta f_{\beta,z}^+|) + \frac{q^-}{2\beta} \mathbb{P}(\omega \in \Omega_{\beta,z}^-) h(|\Delta f_{\beta,z}^-|), \tag{25}
\]

where the function \( h \) is defined as follows.

**Definition 14** The function \( h : \mathbb{R}_{++} \rightarrow \mathbb{R} \) is defined as

\[
h(x) = \begin{cases} 
  x/8, & 0 < x \leq 4, \\
  1 - 2/x, & x \geq 4.
\end{cases}
\]

We would like to express \( |\Delta f_{\beta,z}^+| \) and \( |\Delta f_{\beta,z}^-| \) in terms of the density function \( f \) of the original random variable \( \omega \). We have the following characterization of \( f_{\beta,z}^+ \) and \( f_{\beta,z}^- \).

**Lemma 13** Let the sets \( \Omega_{\beta,z}^+ \) and \( \Omega_{\beta,z}^- \) be defined as in Lemma 12. Furthermore, let the random variables \( \omega_{\beta,z}^+ \) and \( \omega_{\beta,z}^- \) be defined as \( \omega_{\beta,z}^+ \sim \omega | \omega \in \Omega_{\beta,z}^+ \) and \( \omega_{\beta,z}^- \sim \omega | \omega \in \Omega_{\beta,z}^- \), with corresponding probability density functions \( f_{\beta,z}^+ \) and \( f_{\beta,z}^- \), respectively. Then, the following statements hold.

1. We can write \( \Omega_{\beta,z}^+ = [r_{\beta,z}^+, \infty) \) and \( \Omega_{\beta,z}^- = (-\infty, r_{\beta,z}^-] \), where

\[
r_{\beta,z}^+ := z + \frac{\alpha_3^\beta}{q^+} - 1/2 \quad \text{and} \quad r_{\beta,z}^- := z - \frac{\alpha_3^\beta}{q^-} + 1/2.
\]

2. We can write the probability density functions \( f_{\beta,z}^+ \) and \( f_{\beta,z}^- \) as

\[
f_{\beta,z}^+(x) = \begin{cases} 
  0, & \text{if } x < r_{\beta,z}^+, \\
  \mathbb{P}(\omega \in \Omega_{\beta,z}^+)^{-1} f(x), & \text{if } x \geq r_{\beta,z}^+.
\end{cases}
\]

and

\[
f_{\beta,z}^-(x) = \begin{cases} 
  \mathbb{P}(\omega \in \Omega_{\beta,z}^-)^{-1} f(x), & \text{if } x \leq r_{\beta,z}^-, \\
  0, & \text{if } x > r_{\beta,z}^-.
\end{cases}
\]
Proof. To prove the first part, we restrict our attention to $\Omega^+_{\beta,z}$. The proof for $\Omega^-_{\beta,z}$ is completely analogous. By definition, $\Omega^+_{\beta,z} := \{ \omega \mid \hat{v}(\omega - z) \geq \alpha^+_\beta \wedge \omega - z \geq 1/2 \}$. Note that if $\omega - z \geq 1/2$, then $\hat{v}(\omega - z) = q^+(\omega - z + 1/2)$. This function is monotonically increasing in $\omega$. Hence, the condition $\hat{v}(\omega - z) \geq \alpha^+_\beta$ reduces to $\omega \geq r^+_{\beta,z}$, where $r^+_{\beta,z}$ is the value of $\omega$ that solves $q^+(\omega - z + 1/2) = \alpha^+_\beta$. We find that $r^+_{\beta,z} = z - 1/2 + \frac{\alpha^+_\beta}{q^+}$ and we conclude that $\Omega^+_{\beta,z} = [r^+_{\beta,z}, \infty)$.

The second part follows trivially from the first part and the definitions of $\omega^+_{\beta,z}$ and $\omega^-_{\beta,z}$. □

Using (25) and Lemma 13, we can derive an upper bound on $\hat{Q}_\beta(z) - Q_\beta(z)$ in terms of total variations of the density function $f$. Similarly, we can derive an upper bound on the reverse difference. Combining the two yields an upper bound on $|\hat{Q}_\beta(z) - Q_\beta(z)|$.

**Proposition 4** Let the functions $\bar{Q}_\beta$ and $\hat{Q}_\beta$ be defined as

$$
\bar{Q}_\beta(z) = \text{CVaR}_\beta[v(\omega - z)] = \min_{\alpha \in \mathbb{R}} \left\{ \alpha + \frac{1}{\beta} E_{\omega}[v(\omega - z) - \alpha]^+ \right\},
$$

$$
\hat{Q}_\beta(z) = \text{CVaR}_\beta[\hat{v}(\omega - z)] = \min_{\alpha \in \mathbb{R}} \left\{ \alpha + \frac{1}{\beta} E_{\omega}[\hat{v}(\omega - z) - \alpha]^+ \right\},
$$

$z \in \mathbb{R}$, where $v(s) = q^+|s|^+ + q^-|s|^{-}$ and $\hat{v}(s) = q^+(s + 1/2)^+ + q^-(s - 1/2)^-$, $s \in \mathbb{R}$, $\omega$ is a random variable with probability density function $f$, and $\beta$ is a constant in $(0,1)$.

Furthermore, let $\alpha^+_\beta$ and $\hat{\alpha}^+_\beta$, be the $\beta$-VaR of $v(\omega - z)$ and $\hat{v}(\omega - z)$, respectively, $z \in \mathbb{R}$, defined in (22) and (23). Finally, suppose that Assumption 1 holds. Then, for all $z \in \mathbb{R}$,

$$
|\bar{Q}_\beta(z) - \hat{Q}_\beta(z)| \leq \frac{q^+}{8\beta}|\Delta f([\hat{r}^+_{\beta,z}, \infty))| + \frac{q^-}{8\beta}|\Delta f((-\infty, \hat{r}^-_{\beta,z})],
$$

where

$$
\hat{r}^+_{\beta,z} = \min\{r^+_{\beta,z}, \hat{r}^+_{\beta,z}\},
$$

$$
\hat{r}^-_{\beta,z} = \max\{\hat{r}^-_{\beta,z}, r^-_{\beta,z}\},
$$

with

$$
r^+_{\beta,z} = z + \frac{\alpha^+_\beta}{q^+} - 1/2, \quad r^-_{\beta,z} = z - \frac{\alpha^+_\beta}{q^-} + 1/2, \quad \hat{r}^+_{\beta,z} = z + \left\lceil \frac{\hat{\alpha}^+_\beta}{q^+} \right\rceil - 1, \quad \hat{r}^-_{\beta,z} = z - \left\lfloor \frac{\hat{\alpha}^+_\beta}{q^-} \right\rfloor + 1.
$$

Proof. Consider the upper bound in (25) and let $z \in \mathbb{R}$ be given. Using Lemma 13, we can express the total variation of $f_{\beta,z}^+$ as $|\Delta|f_{\beta,z}^+| \leq |\Delta|f_{\beta,z}^+((-\infty, r_{\beta,z}^+))| + |\Delta|f_{\beta,z}^+([r_{\beta,z}^+, \infty))| \leq 2\mathbb{P}\{\omega \in \Omega^+_{\beta,z}\}^{-1}|\Delta f((r_{\beta,z}^+, \infty))$. Notice the occurrence of the probability $\mathbb{P}\{\omega \in \Omega^+_{\beta,z}\}$ in both (25) and the expression for $|\Delta|f_{\beta,z}^-$. We can make these probabilities cancel out by using the fact that $h(x) \leq x/8$, for all $x \geq 0$, which is clear from the definition of $h$. Using the analysis above and an analogous analysis for the expression $|\Delta|f_{\beta,z}^-$, we obtain $\hat{Q}_\beta(z) - Q_\beta(z) \leq \frac{q^+}{8\beta}|\Delta f((r_{\beta,z}^+, \infty))| + \frac{q^-}{8\beta}|\Delta f((-\infty, r_{\beta,z}^-))|$. From the fact that $|\Delta f((r_{\beta,z}^+, \infty))|$ is non-increasing in $r_{\beta,z}^+$ and $|\Delta f((-\infty, r_{\beta,z}^-))|$ is non-decreasing in $r_{\beta,z}^-$, it follows that $\hat{Q}_\beta(z) - Q_\beta(z) \leq \frac{q^+}{8\beta}|\Delta f((r_{\beta,z}^+, \infty))| + \frac{q^-}{8\beta}|\Delta f((-\infty, r_{\beta,z}^-))|

By a completely analogous analysis we can show that the last inequality above also holds if we interchange $\hat{Q}_\beta(z)$ and $Q_\beta(z)$. The analysis for this reverse case is omitted to avoid repetition of arguments. Together, these two inequalities prove the result. □

24
The error bound from Proposition 4 depends on the tender variable \( z = Tx \), which in turn depends on the first-stage decisions \( x \). This dependence on \( z \) is troublesome for two reasons. Firstly, note that the bound depends on \( z \) in a rather intricate way. This hinders interpretation of the result and makes it hard to predict whether the error bound will be tight in practical problem settings.

Secondly, the dependency on \( z \) reduces the applicability of the error bound. Suppose we obtain an optimal solution \( \hat{x}^* \) to the convex approximation model with optimal value \( \hat{\zeta}^* = \hat{Q}_\beta(\hat{x}^*) \). Then, we can find an upper bound on the error \(|\hat{\zeta}^* - \zeta(Tx^*)|\) at the approximation solution \( \hat{x}^* \). However, we want an upper bound on the difference between \( \hat{\zeta}^* \) and \( \zeta^* = Q_\beta(Tx^*) \), the optimal value of the original model, with corresponding optimal solution \( x^* \). In order to obtain an upper bound on \(|\hat{\zeta}^* - \zeta^*|\), we need an upper bound on \(|\hat{\zeta}^* - \hat{Q}_\beta(\hat{x}^*)|\) as well, the approximation error at the optimal solution \( x^* \). The problem is that our convex approximation strategy does not provide us with this optimal solution \( x^* \).

For the abovementioned reasons, we will use the result from Proposition 4 to derive a uniform error bound, i.e. an upper bound on \(|Q_\beta - \hat{Q}_\beta|_\infty \). This uniform error bound will be more easily interpretable than the \( z \)-dependent bound from Proposition 4. Moreover, the uniform error bound will solve the second problem sketched above, since we no longer need to know the optimal solution \( x^* \).

### 4.2 A uniform error bound

Having derived an error bound in Proposition 4 that depends on \( z \), we next use this bound to derive a uniform error bound, i.e. a constant \( D_\beta \) such that \(|Q_\beta - \hat{Q}_\beta|_\infty \leq D_\beta \). The approach we will take is to derive a lower bound on both \( \alpha_{\bar{\gamma}} \) and \( \hat{\alpha}_{\bar{\gamma}} \). By substituting this lower bound into the error bound from Proposition 4, we will see that \( z \) drops out of the resulting error bound and hence, this constitutes a uniform error bound.

#### 4.2.1 A lower bound on \( \alpha_{\bar{\gamma}} \) and \( \hat{\alpha}_{\bar{\gamma}} \)

The goal of this paragraph is to derive a lower bound \( \hat{\alpha}_{\bar{\gamma}} \) on \( \alpha_{\bar{\gamma}} \) and \( \hat{\alpha}_{\bar{\gamma}} \) such that substituting this lower bound into the error bound of Proposition 4 yields a uniform error bound. We consider \( v^{\text{LP}} \), the LP-relaxation of \( v \), defined as

\[
v^{\text{LP}}(s) = q^+(s)^+ + q^-(s)^-, \quad s \in \mathbb{R},
\]

which acts as a lower bound to the functions \( v \) and \( \hat{v} \). That is, \( v^{\text{LP}}(s) \leq v(s) \) and \( v^{\text{LP}}(s) \leq \hat{v}(s) \) for all \( s \in \mathbb{R} \). We also define the corresponding \( \beta \)-VaR:

\[
\tilde{\alpha}_{\bar{\gamma}} := \min \{ \tilde{\alpha} \mid \mathbb{P}\{v^{\text{LP}}(\omega - z) \leq \tilde{\alpha} \} \geq 1 - \beta \}.
\]

Observe that the probabilities \( \mathbb{P}\{v(\omega - z) \leq \alpha \} \) and \( \mathbb{P}\{v^{\text{LP}}(\omega - z) \leq \tilde{\alpha} \} \) are non-decreasing as a function of \( \alpha \) and \( \tilde{\alpha} \), respectively. Together with \( v^{\text{LP}}(s) \leq v(s) \), \( s \in \mathbb{R} \), this implies that \( \tilde{\alpha}_{\bar{\gamma}} \leq \alpha_{\bar{\gamma}} \). Similarly, we have \( \tilde{\alpha}_{\bar{\gamma}} \leq \hat{\alpha}_{\bar{\gamma}} \). So any lower bound on \( \tilde{\alpha}_{\bar{\gamma}} \) is a lower bound on \( \alpha_{\bar{\gamma}} \) and \( \hat{\alpha}_{\bar{\gamma}} \) as well. Suppose we can find a scalar function \( \alpha_{\bar{\gamma}}^{\text{min}}(z) \) such that \( \mathbb{P}\{v^{\text{LP}}(\omega - z) \leq \alpha_{\bar{\gamma}}^{\text{min}}(z) \} \leq 1 - \beta \), for all \( z \in \mathbb{R} \) and \( \beta \in (0, 1) \). Then it follows from the definition of \( \tilde{\alpha}_{\bar{\gamma}} \) that this function constitutes a lower bound to \( \tilde{\alpha}_{\bar{\gamma}} \). In Lemma 14 we show that we can indeed find such a function.
Lemma 14 Let $\beta \in (0,1)$ be given and let $\alpha_{\beta}^z$ and $\hat{\alpha}_{\beta}^z$ be the $\beta$-VaR of the value functions $v(\omega-z) = q^+[\omega-z]^+ + q^- [\omega-z]^-$ and $\hat{v}(\omega-z) = q^+(\omega-z+1/2)^+ + q^- (\omega-z-1/2)^-$, respectively. That is,
\[
\alpha_{\beta}^z := \min \{ \alpha \mid \mathbb{P} \{ v(\omega-z) \leq \alpha \} \geq 1 - \beta \}, \\
\hat{\alpha}_{\beta}^z := \min \{ \hat{\alpha} \mid \mathbb{P} \{ \hat{v}(\omega-z) \leq \hat{\alpha} \} \geq 1 - \beta \}.
\]
Furthermore, let $\alpha_{\beta}^{\text{min}}(z)$ be defined as
\[
\alpha_{\beta}^{\text{min}}(z) = \begin{cases} 
q^+ (F^{-1}(1-\beta) - z), & \text{if } z \leq \bar{z}_{\beta}, \\
q^- (z - F^{-1}(\beta)), & \text{if } z > \bar{z}_{\beta},
\end{cases}
\]
where
\[
\bar{z}_{\beta} = \frac{q^+ F^{-1}(1-\beta) + q^- F^{-1}(\beta)}{q^+ + q^-}.
\]
Then, $\alpha_{\beta}^{\text{min}}(z)$ constitutes a lower bound to both $\alpha_{\beta}^z$ and $\hat{\alpha}_{\beta}^z$. That is, for all $z \in \mathbb{R}$, we have
\[
\alpha_{\beta}^{\text{min}}(z) \leq \alpha_{\beta}^z, \quad \text{and} \quad \alpha_{\beta}^{\text{min}}(z) \leq \hat{\alpha}_{\beta}^z.
\]

Proof. As discussed above, it suffices to show that $\mathbb{P} \{ v^{\text{LP}}(\omega-z) \leq \alpha_{\beta}^{\text{min}}(z) \} \leq 1 - \beta$ holds for all $z \in \mathbb{R}$ and $\beta \in (0,1)$ to prove the lemma. For some $\tilde{\alpha}$ to satisfy $\mathbb{P} \{ v^{\text{LP}}(\omega-z) \leq \tilde{\alpha} \} \leq 1 - \beta$, continuity of $\omega$ and the function $v^{\text{LP}}$ imply that it suffices to show that $\mathbb{P} \{ v^{\text{LP}}(\omega-z) \leq \tilde{\alpha} \} \geq \beta$. We take the approach to first consider the “right part” of the function $v^{\text{LP}}$ only, i.e. the function $s \mapsto q^+ s$. Observe that $v^{\text{LP}}(s) \geq q^+(s) \geq q^+s$, for all $s \in \mathbb{R}$, and hence, $\mathbb{P} \{ v^{\text{LP}}(\omega-z) \geq \tilde{\alpha} \} \geq \mathbb{P} \{ q^+(\omega-z) \geq \tilde{\alpha} \}$. Then, the value of $\tilde{\alpha}$ for which $\mathbb{P} \{ q^+(\omega-z) \geq \tilde{\alpha} \} = \beta$ is obtained by
\[
\mathbb{P} \{ q^+(\omega-z) \geq \tilde{\alpha} \} = \beta \\
\mathbb{P} \{ \omega \leq z + \frac{\tilde{\alpha}}{q^+} \} = 1 - \beta \\
z + \frac{\tilde{\alpha}}{q^+} = F^{-1}(1-\beta) \\
\tilde{\alpha} = q^+ (F^{-1}(1-\beta) - z).
\]
That is, $\tilde{\alpha}$ is the value of the “right part” of the function $v^{\text{LP}}$ at the point $\omega = F^{-1}(1-\beta)$. Similarly, for the “left part” of the function $v^{\text{LP}}$, we obtain $\tilde{\alpha} = q^- (z - F^{-1}(\beta))$.

Both values for $\tilde{\alpha}$ found above constitute a lower bound to $\alpha_{\beta}^z$. Since we want a lower bound that is as tight as possible, we take the maximum of the two values found above. That is, we define the lower bound $\alpha_{\beta}^{\text{min}}(z) := \max \{ q^+ (F^{-1}(1-\beta) - z), q^- (z - F^{-1}(\beta)) \}$. Some elementary calculations make clear that this definition of $\alpha_{\beta}^{\text{min}}(z)$ is equivalent to the definition in the lemma. We have shown that $\mathbb{P} \{ v^{\text{LP}}(\omega-z) \leq \alpha_{\beta}^{\text{min}}(z) \} \leq 1 - \beta$ holds true and hence, $\alpha_{\beta}^{\text{min}}(z)$ constitutes a lower bound to both $\alpha_{\beta}^z$ and $\hat{\alpha}_{\beta}^z$. \qed
4.2.2 Deriving a uniform error bound. Using the lower bound \( \tilde{\alpha}^{\text{min}}(z) \), we now work towards a uniform error bound for \( \hat{Q}_\beta \). We take the error bound from Proposition 4 as a starting point, which is of the form

\[
|\bar{Q}_\beta(z) - \hat{Q}_\beta(z)| \leq \frac{q^+}{8\beta} |\Delta| f([\bar{r}^+_{\beta,z}, \infty)) + \frac{q^-}{8\beta} |\Delta| f((-\infty, \bar{r}^-_{\beta,z})),
\]

where \( \bar{r}^+_{\beta,z} = \min \{r^+_{\beta,z}, \bar{r}^+_{\beta,z}\} \) and \( \bar{r}^-_{\beta,z} = \max \{r^-_{\beta,z}, \bar{r}^-_{\beta,z}\} \). We consider the first term on the right-hand side of (26) and observe that this depends on \( z \) through the total variation \( |\Delta| f([\bar{r}^+_{\beta,z}, \infty]) \). If we can find a lower bound on \( \bar{r}^+_{\beta,z} \), then substituting this lower bound into the total variation \( |\Delta| f([\bar{r}^+_{\beta,z}, \infty]) \) constitutes an upper bound for this total variation. Moreover, if this lower bound on \( \bar{r}^+_{\beta,z} \) does not depend on \( z \), then we have found a uniform upper bound on the total variation \( |\Delta| f([\bar{r}^+_{\beta,z}, \infty]) \). By a similar analysis for \( |\Delta| f((-\infty, \bar{r}^-_{\beta,z})) \) this yields a uniform upper bound on \( |\bar{Q}_\beta(z) - \hat{Q}_\beta(z)| \), which is what we are aiming at.

We now show that we can indeed find such a lower bound on \( \bar{r}^+_{\beta,z} = \min \{r^+_{\beta,z}, \bar{r}^+_{\beta,z}\} \) by substituting \( \tilde{\alpha}^{\text{min}}(z) \) from Lemma 14 for \( \alpha^*_\beta \) and \( \alpha^*_{\beta,z} \) in the definition of \( r^+_{\beta,z} \) and \( \bar{r}^+_{\beta,z} \), respectively.

**Lemma 15** Consider the setting of Proposition 4 and let \( F \) be the cdf of \( \omega \). Then,

\[
\bar{r}^+_{\beta,z} \geq F^{-1}(1 - \beta) - 1 \quad \text{and} \quad \bar{r}^-_{\beta,z} \leq F^{-1}(1 + \beta) + 1.
\]

**Proof.** We prove the first inequality; the proof for the second inequality is completely analogous. Fix \( \beta \in (0, 1) \) and \( z \in \mathbb{R} \). By definition, we have \( \bar{r}^+_{\beta,z} = \min \{r^+_{\beta,z}, \bar{r}^+_{\beta,z}\} \). We first consider \( r^+_{\beta,z} \). By definition of \( r^+_{\beta,z} \) and the inequality \( \alpha^*_\beta \geq \tilde{\alpha}^{\text{min}}(z) \), we obtain

\[
r^+_{\beta,z} = z - 1/2 + \frac{2\alpha^*_\beta}{q^+} \geq z - 1/2 + \frac{\tilde{\alpha}^{\text{min}}(z)}{q^+}.
\]

We consider two cases. First, suppose that \( z \leq \bar{z}_\beta \). Then,

\[
r^+_{\beta,z} \geq z - 1/2 + \frac{q^+(F^{-1}(1-\beta)-z)}{q^+} = F^{-1}(1-\beta) - 1/2.
\]

Second, suppose that \( z > \bar{z}_\beta \). Then,

\[
r^+_{\beta,z} \geq z - 1/2 + \frac{q^-(z - F^{-1}(\beta))}{q^+} = \left(1 + \frac{q^-}{q^+}\right)z - \frac{q^-}{q^+}F^{-1}(\beta) - 1/2 \geq \left(1 + \frac{q^-}{q^+}\right)\bar{z}_\beta - \frac{q^-}{q^+}F^{-1}(\beta) - 1/2 = F^{-1}(1-\beta) + \frac{q^-}{q^+}F^{-1}(\beta) - \frac{q^-}{q^+}F^{-1}(\beta) - 1/2 = F^{-1}(1-\beta) - 1/2.
\]

So we have shown that \( r^+_{\beta,z} \geq F^{-1}(1-\beta) - 1/2 \). By an analogous argument, we obtain the inequality \( \bar{r}^+_{\beta,z} \geq F^{-1}(1-\beta) - 1 \). It follows that \( \bar{r}^+_{\beta,z} \geq F^{-1}(1-\beta) - 1 \). \( \square \)

Note that these bounds on \( \bar{r}^+_{\beta,z} \) and \( \bar{r}^-_{\beta,z} \) indeed do not depend on \( z \). Now, substituting the corresponding lower and upper bound for \( \bar{r}^+_{\beta,z} \) and \( \bar{r}^-_{\beta,z} \), respectively, into (26) yields a uniform upper bound on \( |\bar{Q}_\beta(z) - \hat{Q}_\beta(z)| \). We obtain

\[
|\bar{Q}_\beta(z) - \hat{Q}_\beta(z)| \leq \frac{q^+}{8\beta} |\Delta| f([-F^{-1}(1-\beta) - 1, \infty)) + \frac{q^-}{8\beta} |\Delta| f((-\infty, F^{-1}(\beta) + 1]) \quad (27)
\]
Observe that the uniform error bound above is based on Proposition 4, in which it is assumed that Assumption 1 holds. Hence, validity of (27) is only ensured for values of $z$ for which Assumption 1 holds. We make the following assumption, which is a sufficient condition for Assumption 1 to hold for all $z \in \mathbb{R}$.

**Assumption 2** We assume that $F^{-1}(1 - \beta) - F^{-1}(\beta) > \left(\frac{1}{q^+} + \frac{1}{q^-}\right) \max\{q^+, q^-\}$.

Assumption 2 states that for a given level of $\beta$, the “variability” of the random variable $\omega$, i.e. the extent to which the probability mass is spread out over the real line, should be large enough. On the other hand, for a given probability distribution, the assumption states that $\beta$ should be small enough. Specifically, we cannot have $\beta \geq 1/2$.

**Lemma 16** If Assumption 2 holds, then Assumption 1 holds for all $z \in \mathbb{R}$.

**Proof.** Suppose that Assumption 2 holds and define $\bar{q} := \max\{q^+, q^-\}$. Let $z \in \mathbb{R}$ be given arbitrarily. We need to show that $\alpha^{\bar{q}} > \bar{q}$ and $\hat{\alpha}^{\bar{q}} > \bar{q}$. By Lemma 14, we know that $\alpha^{\bar{q}} \geq \bar{\alpha}^{\min}(z)$ and $\hat{\alpha}^{\bar{q}} \geq \bar{\alpha}^{\min}(z)$. Hence, it is sufficient to show that $\bar{\alpha}^{\min}(z) > \bar{q}$. We distinguish two cases.

First, suppose that $z \leq \bar{z}_\beta$, with $\bar{z}_\beta$ defined as in Lemma 14. We have assumed that Assumption 2 holds, i.e. $F^{-1}(1 - \beta) - F^{-1}(\beta) > \left(\frac{1}{q^+} + \frac{1}{q^-}\right) \bar{q}$. Some elementary calculations make clear that this is equivalent to $\bar{z}_\beta < F^{-1}(1 - \beta) - \frac{\bar{q}}{q^+}$. Using this inequality and the assumption $z \leq \bar{z}_\beta$, we obtain

$$\bar{\alpha}^{\min}(z) = q^+(F^{-1}(1 - \beta) - z) \geq q^+(F^{-1}(1 - \beta) - \bar{z}_\beta) > q^+(F^{-1}(1 - \beta) - F^{-1}(1 - \beta) + \frac{\bar{q}}{q^+}) = \bar{q}.$$ 

Second, suppose that $z > \bar{z}_\beta$. Similarly to above, it can be shown that Assumption 2 is equivalent to $\bar{z}_\beta > F^{-1}(\beta) + \frac{\bar{q}}{q^-}$. Using this inequality and the assumption $z > \bar{z}_\beta$, we obtain

$$\bar{\alpha}^{\min}(z) = q^-(z - F^{-1}(\beta)) > q^-(\bar{z}_\beta - F^{-1}(\beta)) > q^-(F^{-1}(\beta) + \frac{\bar{q}}{q^-} - F^{-1}(\beta)) = \bar{q}.$$ 

Hence, we have shown that $\bar{\alpha}^{\min}(z) > \bar{q}$ for all $z \in \mathbb{R}$.

We are now ready to present a uniform upper bound on the approximation error $||Q_\beta - \hat{Q}_\beta||_\infty$.

**Theorem 2** Let the functions $Q_\beta$ and $\hat{Q}_\beta$ be defined as

$$Q_\beta(z) = \text{CVaR}_\beta[v(\omega - z)] = \min_{\alpha \in \mathbb{R}} \left\{ \alpha + \frac{1}{\beta} \mathbb{E}_\omega[(v(\omega - z) - \alpha)^+] \right\},$$

$$\hat{Q}_\beta(z) = \text{CVaR}_\beta[\hat{v}(\omega - z)] = \min_{\hat{\alpha} \in \mathbb{R}} \left\{ \hat{\alpha} + \frac{1}{\beta} \mathbb{E}_\omega[(\hat{v}(\omega - z) - \hat{\alpha})^+] \right\},$$

where $v(s) = q^+[s] + q^-[s]$ and $\hat{v}(s) = q^+(s + 1/2)^+ + q^-(s - 1/2)^-$, $s \in \mathbb{R}$; $\omega$ is a random variable with probability density function $f$ and cumulative distribution function...
\( F \); and \( \beta \) is a constant in \((0, 1)\). Furthermore, let \( \alpha^\circ_\beta \) and \( \alpha^\circ_\beta \) be the \( \beta \)-VaR of \( v(\omega - z) \) and \( \hat{v}(\omega - z) \), respectively. Finally, suppose that Assumption 2 holds. Then,

\[
||\bar{Q}_\beta - \hat{Q}_\beta||_{\infty} \leq \frac{q^+}{8\beta} |\Delta| f([-1/(1 - \beta) - 1, \infty]) + \frac{q^-}{8\beta} |\Delta| f((-\infty, F^{-1}(\beta) + 1])
\]

Proof. Suppose Assumption 2 holds. Then, by Lemma 16, Assumption 1 holds for all \( z \in \mathbb{R} \). By Proposition 4 this implies that for all \( z \in \mathbb{R} \),

\[
|\bar{Q}_\beta(z) - \hat{Q}_\beta(z)| \leq \frac{q^+}{8\beta} |\Delta| f([\tilde{r}^+_\beta, \infty]) + \frac{q^-}{8\beta} |\Delta| f((-\infty, \tilde{r}^-_\beta]),
\]

with \( \tilde{r}^+_\beta,z \) and \( \tilde{r}^-_\beta,z \) defined in Proposition 4. Note that the right-hand side of the inequality above is non-increasing in \( \tilde{r}^+_\beta,z \) and non-decreasing in \( \tilde{r}^-_\beta,z \). Hence, by Lemma 15 it follows that for all \( z \in \mathbb{R} \),

\[
|\bar{Q}_\beta(z) - \hat{Q}_\beta(z)| \leq \frac{q^+}{8\beta} |\Delta| f([-1/(1 - \beta) - 1, \infty]) + \frac{q^-}{8\beta} |\Delta| f((-\infty, F^{-1}(\beta) + 1])
\]

In the next paragraph, we will investigate this uniform error bound more closely. In particular, we will investigate its behavior with respect to the probability parameter \( \beta \) in the definition of CVaR.

### 4.2.3 Asymptotic behavior of the uniform error bound.

In practical decision making, the reason to use mean-risk models instead of expected value models is often to take into account the possibility of extreme adverse events (e.g. a large financial crash). Such extreme events usually have a low probability of occurring. Translated to our mean-CVaR model, this means that we are interested in models with small values for \( \beta \), say \( \beta = 0.01 \). If the uniform error bound from Theorem 2 is small for small values of \( \beta \), then the mean-CVaR model is suitable for practical situations as described above. To investigate this matter, we will look at the asymptotic behavior of the error bound as \( \beta \) goes to zero.

In this subsection we show that under some conditions on the distribution of \( \omega \), the uniform error bound from Theorem 2 converges to zero as \( \beta \) goes to zero. Specifically, we assume that the pdf \( f \) of \( \omega \) is unimodal, positive everywhere and has fat tails. Note that the fat tails assumption fits well within the practical context sketched above, in which the possibility of extreme events is not negligible. We formalize the assumptions below.

**Assumption 3** Consider the setting of Theorem 2. We assume that

1. \( f \) is unimodal. That is, there exists \( x_0 \in \mathbb{R} \) such that \( f \) is monotonically non-decreasing on \((-\infty, x_0] \) and monotonically non-increasing on \([x_0, \infty) \).
2. \( f(x) > 0 \) for all \( x \in \mathbb{R} \).

Before defining Assumption 4, we need to define the concept of asymptotic equivalence of functions.

**Definition 15** Two functions \( f : \mathbb{R} \to \mathbb{R} \) and \( g : \mathbb{R} \to \mathbb{R} \) are asymptotically equivalent as \( x \to l \), with \( l \in \{-\infty, \infty\} \), if

\[
\lim_{x \to l} \frac{f(x)}{g(x)} = 1.
\]

We write \( f(x) \sim g(x) \) as \( x \to l \).
Assumption 4 Let $\omega$ be the random variable from Theorem 2 with cdf $F$ and pdf $f$. Moreover, define $\bar{F}(x) = 1 - F(x), \ x \in \mathbb{R}$. We assume that the left and right tail of the distribution of $\omega$ are fat. That is, we assume that there exist positive constants $c_1, c_2, \alpha_1$, and $\alpha_2$ such that

$$\bar{F}(x) \sim \bar{G}(x), \text{ as } x \to \infty,$$

and

$$F(x) \sim G(x), \text{ as } x \to -\infty,$$

where $\bar{G}(x) = c_1 x^{-\alpha_1}, \ x \in \mathbb{R}, G(x) = -c_2 x^{-\alpha_2}, \ x \in \mathbb{R},$ and where the symbol “$\sim$” denotes asymptotic equivalence of functions (see Definition 15).

Remark 3 Note that by L’Hôpital’s rule, Assumption 4 implies that for the derivatives $f$, $\bar{f}$, $g$, and $\bar{g}$ of $F, \bar{F}, G,$ and $\bar{G}$, respectively,

$$\bar{f}(x) \sim \bar{g}(x), \text{ as } x \to \infty,$$

$$f(x) \sim g(x), \text{ as } x \to -\infty,$$

where $\bar{f}(x) = -f(x)$ for all $x \in \mathbb{R}$, $\bar{g}(x) = -\alpha_1 c_1 x^{-(\alpha_1 + 1)}$ for all $x > 0$, and $g(x) = \alpha_2 c_2 x^{-(\alpha_2 + 1)}$ for all $x < 0$.

Consider the uniform error bound from Theorem 2. To show that this error bound converges to zero as $\beta \downarrow 0$, we focus our attention on the expression $\frac{1}{\beta} |\Delta| f\left((\infty, F^{-1}(\beta) + 1]\right)$. Assumption 3 implies that for $\beta$ small enough, we have $F^{-1}(\beta) + 1 \leq x_0$, such that $f$ is monotonically non-decreasing on $(-\infty, F^{-1}(\beta) + 1]$. This implies that $\frac{1}{\beta} |\Delta| f\left((\infty, F^{-1}(\beta) + 1]\right) = \frac{1}{\beta} f(F^{-1}(\beta) + 1), \text{ for } \beta \text{ small enough.}$

We can write

$$\lim_{\beta \downarrow 0} \frac{1}{\beta} f(F^{-1}(\beta) + 1) = \lim_{\beta \downarrow 0} \frac{f(F^{-1}(\beta) + 1)}{F(F^{-1}(\beta))} = \lim_{x \to -\infty} \frac{f(x + 1)}{F(x)}, \quad (28)$$

since $\lim_{\beta \downarrow 0} F^{-1}(\beta) = -\infty$. Using Assumption 4 and Remark 3, we can express the last limit in (28) in terms of $g$ and $G$ and it can be shown that this limit equals zero and hence, $\lim_{\beta \downarrow 0} \frac{1}{\beta} |\Delta| f\left((\infty, F^{-1}(\beta) + 1]\right) = 0$. By an analogous argument, it can be shown that $\lim_{\beta \downarrow 0} \frac{1}{\beta} |\Delta| f\left([F^{-1}(1 - \beta) - 1, \infty)\right) = 0$. By substituting these two limits, we can show that the error bound from Theorem 2 converges to zero as $\beta \downarrow 0$.

Proposition 5 Consider the setting of Theorem 2. Suppose that $F^{-1}(1 - \beta) - F^{-1}(\beta) > \left(\frac{1}{q^+} + \frac{1}{q^-}\right) \max\{q^+, q^-\}$ and that $f$ is unimodal, positive everywhere, and has fat tails, that is, Assumption 2, 3, and 4 hold. Let $D^\beta$ denote the error bound from Theorem 2, i.e.

$$D^\beta = \frac{q^+}{8\beta} |\Delta| f\left([F^{-1}(1 - \beta) - 1, \infty)\right) + \frac{q^-}{8\beta} |\Delta| f\left((\infty, F^{-1}(\beta) + 1]\right).$$

Then,

$$\lim_{\beta \downarrow 0} D^\beta = 0.$$
Proof. Suppose Assumption 2, 3, and 4 hold. Then, by (28), Assumption 4 Remark 3, and the Algebraic Limit Theorem,

$$
\lim_{\beta \to 0} \frac{1}{\beta} |\Delta f ((-\infty, F^{-1}(\beta) + 1]) = \lim_{x \to -\infty} \frac{f(x + 1)}{F(x)}
$$

$$
= \lim_{x \to -\infty} \left( \frac{g(x + 1)}{G(x)} \cdot \frac{f(x + 1)}{g(x + 1)} \cdot \frac{G(x)}{F(x)} \right)
$$

$$
= \lim_{x \to -\infty} \frac{g(x + 1)}{G(x)} \cdot \lim_{x \to -\infty} \frac{f(x + 1)}{g(x + 1)} \cdot \lim_{x \to -\infty} \frac{G(x)}{F(x)}
$$

$$
= \lim_{x \to -\infty} \frac{\alpha_2 c_2 (x + 1)^{-\alpha_2}}{-c_2 x^{-\alpha_2}} \cdot 1 \cdot 1
$$

$$
= -\alpha_2 \lim_{x \to -\infty} (x + 1)^{-1} \left( \frac{x + 1}{x} \right)^{-\alpha_2}
$$

$$
= -\alpha_2 \lim_{x \to -\infty} (x + 1)^{-1} \lim_{x \to -\infty} \left( \frac{x + 1}{x} \right)^{-\alpha_2}
$$

$$
= -\alpha_2 \cdot 0 \cdot 1
$$

$$
= 0.
$$

Similarly, it can be shown that \( \lim_{\beta \to 0} \frac{1}{\beta} |\Delta f ([F^{-1}(1 - \beta) - 1, \infty)) = 0 \). The result follows from substituting these two limits into the expression for \( \lim_{\beta \to 0} D^\beta \).

Proposition 5 implies that for models with a fat-tailed distribution, the convex approximation technique works relatively well if we choose a low value for \( \beta \). As mentioned earlier, if we have fat-tailed distributions, we are especially interested in CVaR recourse models, since these can capture the risks caused by extreme events with non-negligible probabilities. Moreover, to capture these extreme events, we are generally interested in low values for \( \beta \). We conclude that CVaR recourse models can be approximated well using the technique described in this section in a situation in which it is desirable to use these models.

5 Discussion. We considered mixed-integer mean-risk recourse models using Conditional Value-at-Risk as our risk measure. Integrality conditions on decision variables cause non-convexity of the recourse function, which makes these mean-CVaR models hard to solve. Contrarily to most other researchers, we considered continuously distributed random variables. In order to overcome the problem caused by the non-convexity, we constructed convex approximations of the mean-CVaR models. The resulting approximate models can be solved using convex optimization techniques. In order to assess the closeness of the approximations to the original models, we derived uniform error bounds. For ease of calculation, rather than mean-CVaR models, pure CVaR models were considered. Using existing error bounds for the corresponding expected value models, results can be generalized to the mean-CVaR setting.

For the general mixed-integer case, a total variation error bound was derived. This bound is asymptotically converging in the sense that it converges to zero when the total variations of the densities of the random variables in the model go to zero. Intuitively, this means that the approximation becomes arbitrarily good as the variability of the random variables increases. High variability mitigates the distorting effect of the integrality conditions and makes the original model closer to being convex. We would like to stress that this result
holds for any mixed-integer CVaR recourse model adhering to a few technical conditions. So any such CVaR model becomes arbitrarily close to a convex model if the variability in the model increases.

The results for general mixed-integer CVaR recourse models in this paper are the first of this kind to be found. Therefore, it is not surprising that the error bound from Theorem 1 is of an asymptotic nature. We merely show that there exists a constant $C > 0$ such that $||\tilde{Q}_\beta - \hat{Q}_\beta||_\infty \leq C\theta(f)$, for any model with continuously distributed random variables with a joint density function $f \in \mathcal{H}^m$, where $\theta(f)$ is a function depending on the total variations of $f$. It should be noted that this constant $C$ depends heavily on $K$, the number of dual feasible basis matrices for the LP-relaxation of the second-stage value function $v$. In general, this number grows exponentially in the dimension of the problem. This implies that the error bound from Theorem 1 can be too large for practical purposes, even if the actual error is small. Nevertheless, its asymptotic properties provide a valuable insight into the characteristics of these models.

In order to derive a practically useful, sharper error bound, we consider the special case of simple integer CVaR recourse models. This well-known special case has various practical applications and its simple structure facilitates derivation of a tighter error bound. This bound has an additional desirable property compared to the bound for the general mixed-integer case.

It was shown that under the condition that the pdf of the random variable in the model has fat tails (and some regularity assumptions hold), the error bound converges to zero as $\beta$, the probability parameter in the definition of CVaR, goes to zero. This is an interesting result, for the following reason. We are especially interested in modeling risk-averseness explicitly in the situation that extreme risks occur with a non-negligible probability. This is the case when we have fat-tailed distributions. Moreover, we are generally interested in small values for $\beta$, e.g. $\beta = 0.01$. Precisely in this combination of conditions, we find that the convex approximation technique described in this paper works well. That is, our model behaves well in a setting in which one would like to use it.

It would be interesting to apply our results for simple integer CVaR recourse models to a practical problem. This would provide insight into the numerical behavior of the error bound and, more generally, into the practical performance of our convex approximation approach as a whole. Other directions for future research include considering TU integer mean-CVaR models. For the expected value case, Romeijnders, van der Vlerk et al. (2016) obtain total variation error bounds which are not merely of an asymptotic nature. It may be possible to extend these results to the mean-CVaR case. Furthermore, multi-stage models or models with random cost parameters $q(\omega)$ and random technology matrices $T(\omega)$ may be considered.

References


