Total variation error bounds for convex approximations of two-stage mixed-integer recourse models

Romeijnders, Ward

Publication date:
2015

Chapter 3

Convex approximations for TU integer recourse models

Abstract. *We consider a class of convex approximations for totally unimodular (TU) integer recourse models and derive a uniform error bound by exploiting properties of the total variation of the probability density functions involved. For simple integer recourse models this error bound is tight and improves the existing one by a factor 2, whereas for TU integer recourse models this is the first nontrivial error bound available. The bound ensures that the performance of the approximations is good as long as the total variations of the densities of all random variables in the model are small enough.*

3.1 Introduction

We consider the two-stage integer recourse problem

$$\eta^* := \min_x \Big\{ cx + Q(z) : Ax \geq b,\ z = Tx,\ x \in \mathbb{R}^{n_1}_+ \Big\}, \tag{3.1}$$

where z are tender variables, Q is the recourse (expected value) function

$$Q(z) := \mathbb{E}_\omega \Big[v(\omega - z) \Big], \quad z \in \mathbb{R}^m,$$

This chapter is based on the journal publication [63].

and v is the second-stage value function

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

The second-stage decision variables y represent the so-called recourse actions that compensate for infeasibilities with respect to the random goal constraints $Tx \geq \omega$. Here, there is only randomness in the right-hand side ω, which is a random vector with known distribution. The functions Q and v represent the (expected) recourse cost associated with the recourse actions y.

Modeling indivisibilities or on/off decisions typically requires integer (or binary) decision variables. For this reason, introducing such integer variables to the model is highly relevant for practice but at the same time makes the model considerably more difficult to solve. Most exact solution methods combine ideas behind algorithms designed for either stochastic continuous or deterministic integer programs (see e.g. [1, 11, 22, 30, 44, 45, 46, 53, 71, 73, 75] and the survey papers by Klein Haneveld and Van der Vlerk [43], Louveaux and Schultz [47], Romeijnders et al. [60], Schultz [70], and Sen [72]). Although substantial progress has been made, in general these algorithms have difficulties solving large real-life problem instances.

The main reason that integer recourse models are considerably more difficult to solve than continuous recourse models is that the integer recourse function Q is generally *non-convex* [56]. A possible approach to dealing with this difficulty is to construct convex approximations of the recourse function Q by modifying the recourse data (MRD) [82], which comprises the parameters and structure of the model, and the distributions of the random variables involved. The rationale for doing this is that convex optimization problems are computationally much more tractable than non-convex problems, and as long as we only make small changes in the recourse data we expect to obtain close approximations.

Using MRD a class of convex approximations of Q has been developed, first for the special case of simple integer recourse (SIR) models (when $W = I_m$) [42] and later for general complete integer recourse models [83] and mixed-integer recourse models with a single recourse constraint [85]. The recurring idea in these so-called α-approximations is to simultaneously relax the integrality constraints and perturb the distribution of the right-hand side random vector ω. In this way, a difficult-to-solve integer recourse problem is approximated by a continuous recourse problem for which efficient algorithms exist such as (variants of) the L-shaped algorithm [87], regularized decomposition [66], and stochastic decomposition [32]; see e.g. [94] for a recent computational study comparing various algorithms.

Although a uniform error bound for these approximations is available for models with a simple recourse structure [42], such an error bound is lacking for integer recourse models in general. We derive a uniform error bound for integer recourse models with a totally unimodular (TU) recourse matrix W by exploiting properties of the total variation of probability density functions. This error bound is tight for SIR models and improves the existing error bound by a factor 2. Moreover, the error bound ensures that the convex approximations are good as long as the total variations of the densities of all random variables in the model are small enough. Furthermore, due to this error bound the convex approximations can be used as an approximate lower bound for complete integer recourse models.

The remainder of this chapter is organized as follows. We introduce α-approximations of integer recourse models in Section 3.2. To set the stage for our analysis, we discuss properties of the total variation of probability density functions in Section 3.3, and we solve a simplified one-dimensional bounding problem in Section 3.4. In Sections 3.5 and 3.6 we derive a uniform error bound for α-approximations of TU integer recourse models with independent and dependent random variables, respectively. The approximate lower bound for complete integer recourse models is discussed in Section 3.7, and we end with a summary and conclusions in Section 3.8.

3.2 Convex approximations and literature review

Throughout this chapter we use the following assumptions.

(i) W is a complete recourse matrix; that is, for every $s \in \mathbb{R}^m$ there exists $y \in \mathbb{Z}_+^{n_2}$ such that $Wy \geq s$, and thus $v(s) < +\infty$.

(ii) The recourse structure is sufficiently expensive; that is, $v(s) > -\infty$ for all $s \in \mathbb{R}^m$.

(iii) $\mathbb{E}_\omega[|\omega|]$ is finite.

As a result the recourse function Q is finite everywhere.

We consider so-called *α-approximations* of Q, which is a class of convex approximations of Q studied in Van der Vlerk [83] and related work. These α-approximations are an example of MRD, as discussed earlier.

Definition 3.1. For every $\alpha \in \mathbb{R}^m$, the *α-approximation of Q* is given by

$$Q_\alpha(z) := \mathbb{E}_\omega \left[\min_y \left\{ qy : Wy \geq \lceil \omega \rceil_\alpha - z,\ y \in \mathbb{R}^{n_2}_+ \right\} \right], \quad z \in \mathbb{R}^m,$$

where $\lceil \omega \rceil_\alpha := \lceil \omega - \alpha \rceil + \alpha$ is the round-up of ω with respect to $\alpha + \mathbb{Z}^m$.

Remark 3.1. Note that the definition of α-approximations is given for $\alpha \in \mathbb{R}^m$, but since $Q_\alpha \equiv Q_{\alpha'}$ if $\alpha - \alpha' \in \mathbb{Z}^m$, we could have restricted the definition to $\alpha \in [0,1)^m$.

For every $\alpha \in \mathbb{R}^m$, the random vector $\lceil \omega \rceil_\alpha$ is discretely distributed with support in $\alpha + \mathbb{Z}^m$. Hence, the α-approximation Q_α is the recourse function of a *continuous* recourse model with *discrete* random right-hand side vector $\lceil \omega \rceil_\alpha$, and thus Q_α is a convex polyhedral function. Although Dyer and Stougie [18] show that from a theoretical complexity point of view these problems are hard to solve in general, there exist algorithms that can solve such recourse problems involving discrete distributions within reasonable time limits. This implies that if the difference between $Q(z)$ and its approximation $Q_\alpha(z)$ is small enough for all $z \in \mathbb{R}^m$, then the approximating model

$$\hat{\eta}_\alpha := \min_x \left\{ cx + Q_\alpha(z) : Ax \geq b,\ z = Tx,\ x \in \mathbb{R}^{n_1}_+ \right\}, \tag{3.2}$$

not only is computationally tractable but also leads to (near-)optimal solutions; see Lemma 3.1. For this reason, we use the supremum norm to measure the error of the approximations:

$$\|Q - Q_\alpha\|_\infty := \sup_{z \in \mathbb{R}^m} |Q(z) - Q_\alpha(z)|, \quad \alpha \in \mathbb{R}^m.$$

Lemma 3.1. *Let $\alpha \in \mathbb{R}^m$ be given, and consider the optimization problem in (3.1) and its approximation in (3.2) with optimal solutions (x^*, z^*) and $(\hat{x}_\alpha, \hat{z}_\alpha)$, respectively. Then,*

(*i*) $|\eta^* - \hat{\eta}_\alpha| \leq \|Q - Q_\alpha\|_\infty$

and

$$(ii) \quad |\eta^* - c\hat{x}_\alpha - Q(\hat{z}_\alpha)| \leq 2\|Q - Q_\alpha\|_\infty.$$

Proof. Using that (x^*, z^*) is optimal in (3.1) and $(\hat{x}_\alpha, \hat{z}_\alpha)$ is optimal in (3.2), we have

$$\eta^* \leq c\hat{x}_\alpha + Q(\hat{z}_\alpha) \leq c\hat{x}_\alpha + Q_\alpha(\hat{z}_\alpha) + \|Q - Q_\alpha\|_\infty = \hat{\eta}_\alpha + \|Q - Q_\alpha\|_\infty, \tag{3.3}$$

and

$$\hat{\eta}_\alpha \leq cx^* + Q_\alpha(z^*) \leq cx^* + Q(z^*) + \|Q - Q_\alpha\|_\infty = \eta^* + \|Q - Q_\alpha\|_\infty. \tag{3.4}$$

Combining (3.3) and (3.4) yields (i). The inequality in (ii) follows from (3.3) and (3.4) as well, since

$$\eta^* \leq c\hat{x}_\alpha + Q(\hat{z}_\alpha) \leq \hat{\eta}_\alpha + \|Q - Q_\alpha\|_\infty \leq \eta^* + 2\|Q - Q_\alpha\|_\infty. \qquad \square$$

That is to say, $\|Q - Q_\alpha\|_\infty$ is an upper bound for the difference $\eta^* - \hat{\eta}_\alpha$ in objective values of the original and approximating models, and the objective value of the *approximate* solution $(\hat{x}_\alpha, \hat{z}_\alpha)$ in the *original* model differs at most $2\|Q - Q_\alpha\|_\infty$ from the optimal objective value η^*.

The main contribution of this chapter is the derivation of nontrivial upper bounds of $\|Q - Q_\alpha\|_\infty$ for integer recourse models with TU recourse matrix W. In the remaining part of this section we review the existing literature on such upper bounds.

First, consider the case $W = I_m$. Then, problem (3.1) reduces to a one-sided SIR problem [48]. This problem is called simple because the recourse function $Q(z)$ is separable in the components of z, so that

$$Q(z) = \mathbb{E}_\omega\Big[\min_y \Big\{qy : y \geq \omega - z,\, y \in \mathbb{Z}_+^m\Big\}\Big] = \sum_{i=1}^m q_i Q_i(z_i), \quad z \in \mathbb{R}^m, \tag{3.5}$$

where $Q_i(z_i) := \mathbb{E}_{\omega_i}[\lceil \omega_i - z_i \rceil^+]$, and similarly

$$Q_\alpha(z) = \sum_{i=1}^m q_i \mathbb{E}_{\omega_i}\Big[(\lceil \omega_i \rceil_{\alpha_i} - z_i)^+\Big], \quad z \in \mathbb{R}^m.$$

Here, $(x)^+ := \max\{0, x\}$ denotes the positive part of $x \in \mathbb{R}$ (also, componentwise for $x \in \mathbb{R}^m$), and we conveniently write $\lceil x \rceil^+$ to denote $\max\{0, \lceil x \rceil\}$.

The properties of the m-dimensional SIR function Q follow directly from those of the generic one-dimensional SIR function

$$\mathcal{Q}(z) := \mathbb{E}_\omega[\lceil \omega - z \rceil^+], \quad z \in \mathbb{R}. \tag{3.6}$$

If the one-dimensional random variable ω is discretely distributed, then efficient algorithms are available to construct the convex hull of $\mathcal{Q}$ [40, 41]. If ω is continuously distributed with probability density function (pdf) f of bounded variation, then Klein Haneveld et al. [42] show that for every $\alpha \in \mathbb{R}$,

$$\|\mathcal{Q} - \mathcal{Q}_\alpha\|_\infty \leq \min\left\{\frac{|\Delta| f}{4}, 1\right\}, \tag{3.7}$$

where $\mathcal{Q}_\alpha$ denotes the α-approximation of $\mathcal{Q}$ and $|\Delta| f := |\Delta| f(\mathbb{R})$ the total variation of f on $\mathbb{R}$; see Definition 3.2. This result leads to the following uniform upper bound on the error in the case of SIR:

$$\sup_{z \in \mathbb{R}^m} |Q(z) - Q_\alpha(z)| \leq \sum_{i=1}^m q_i \min\left\{\frac{|\Delta| f_i}{4}, 1\right\}, \quad \alpha \in \mathbb{R}^m, \tag{3.8}$$

where f_i is the marginal pdf of ω_i.

Let us now consider the more general case, where the recourse matrix W is TU. The second-stage value function v can be rewritten in a more convenient form. Since the recourse is complete and sufficiently expensive, we have for all $s \in \mathbb{R}^m$,

$$\begin{aligned} v(s) &= \min_y \{qy : Wy \geq s,\ y \in \mathbb{Z}_+^{n_2}\} \\ &= \min_y \{qy : Wy \geq \lceil s \rceil,\ y \in \mathbb{R}_+^{n_2}\} && (3.9) \\ &= \max_\lambda \{\lambda \lceil s \rceil : \lambda W \leq q,\ \lambda \in \mathbb{R}_+^m\}, && (3.10) \end{aligned}$$

where the equality in (3.9) follows from the fact that W is TU, and the equality in (3.10) holds by strong linear programming (LP) duality. Assumptions (i) and (ii) also imply that the dual feasible region $\{\lambda W \leq q, \lambda \geq 0\}$ is non-empty and bounded.

Thus it is spanned by finitely many extreme points λ^k, $k = 1, \dots, K$. Hence,

$$v(s) = \max_{k=1,\dots,K} \lambda^k \lceil s \rceil, \quad s \in \mathbb{R}^m,$$

and thus

$$Q(z) = \mathbb{E}_\omega \left[\max_{k=1,\dots,K} \lambda^k \lceil \omega - z \rceil \right], \quad z \in \mathbb{R}^m. \tag{3.11}$$

Correspondingly, for every $\alpha \in \mathbb{R}^m$ the α-approximation Q_α can be written as

$$Q_\alpha(z) = \mathbb{E}_\omega \left[\max_{k=1,\dots,K} \lambda^k \left(\lceil \omega \rceil_\alpha - z \right) \right], \quad z \in \mathbb{R}^m. \tag{3.12}$$

Now it is easy to observe that Q is the expectation of the pointwise maximum of finitely many round-up functions, so that Q is generally non-convex, whereas Q_α is a convex polyhedral function.

Van der Vlerk [83] claims that there exists $\alpha^* \in \mathbb{R}^m$ such that Q_{α^*} is the convex hull of Q, so that the approximation model in (3.2) yields exact results if the matrix T is of full row rank and the optimal solution x^* is an interior point of the deterministic constraint set $\{x \in \mathbb{R}^{n_1}_+ : Ax \geq b\}$. These two conditions, especially the latter, may be very restrictive from a practical point of view. Moreover, if one of these conditions does not hold, then there is no performance guarantee at all for the approximate solution $\hat{x}_{\alpha^*}$ that is obtained.

An upper bound of $\|Q - Q_\alpha\|_\infty$ is not subject to these drawbacks and provides a performance guarantee irrespective of whether the above-mentioned conditions are satisfied or not. For this reason, deriving such an error bound is important. In addition, we have shown in Chapter 2 that the claim of Van der Vlerk [83] does not hold in general; see also [61] and [86]. In fact, the claim holds only if all random variables in the model are independently and uniformly distributed, underlining the relevance for practical purposes of the error bound we derive in this chapter.

3.3 Piecewise flattening of density functions without increasing total variation

The error bound for SIR models in (3.8) shows that the total variations of the densities of the random variables in the model are the main determinants of the magnitude of the error $\|Q - Q_\alpha\|_\infty$. Since these total variations will play an important role in the derivation of an error bound for TU integer recourse models as well, we first give a formal definition of total variation.

Definition 3.2 (Total variation)**.** Let $f : \mathbb{R} \mapsto \mathbb{R}$ 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 $P = \{x_1, \dots, x_{N+1}\}$ with $x_1 < \cdots < x_{N+1}$ in I. Then, the total variation of f on I, denoted $|\Delta|f(I)$, is defined as

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

where

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

We will write $|\Delta|f := |\Delta|f(\mathbb{R})$. Moreover, f is of bounded variation if and only if $|\Delta|f < +\infty$.

In this section we introduce several lemmas based on properties of the total variation of one-dimensional density functions. We use these lemmas extensively to solve a simplified one-dimensional bounding problem in Section 3.4, and to derive a bound for $\|Q - Q_\alpha\|_\infty$ for TU integer recourse models in Sections 3.5 and 3.6. In order to avoid technicalities, we only consider density functions f that are well behaved in the following sense. (The obvious generalization to (in)dependent pdf on $\mathbb{R}^m$ is given in Sections 3.5 and 3.6.)

Definition 3.3. Let $\mathcal{F}$ denote the set of one-dimensional probability density functions f of bounded variation that have finitely many discontinuity points on any bounded interval.

Remark 3.2. Note that for every $f \in \mathcal{F}$ there exists a left-continuous version $\hat{f} \in \mathcal{F}$ that is practically equivalent to f with $|\Delta|\hat{f} \leq |\Delta|f$.

The first lemma says that the total variation does not increase when we *flatten* a density function on some bounded interval I in such a way that the probability of the event $\{\omega \in I\}$ does not change. The intuition behind this lemma is that a constant function has lower total variation than a varying one.

Lemma 3.2. *Let $f \in \mathcal{F}$ be given, and let $I \subset \mathbb{R}$ denote a bounded interval with positive length $|I|$. Define $g \in \mathcal{F}$ as*

$$g(x) = \begin{cases} f(x), & x \notin I, \\ K_I, & x \in I, \end{cases} \tag{3.13}$$

with $K_I := |I|^{-1} \int_I f(u)du$. Then $|\Delta|g \leq |\Delta|f$.

Proof. Let $f \in \mathcal{F}$ be given and assume for the moment that I is open, so that $I = (a, b)$ for some $a < b$. Since $g(x) = f(x)$ for $x \notin (a, b)$, it follows that $|\Delta|g \leq |\Delta|f$ if and only if $|\Delta|g([a, b]) \leq |\Delta|f([a, b])$. Since g has the constant value K_I on the interval (a, b), it follows that

$$|\Delta|g([a, b]) = |K_I - f(a)| + |f(b) - K_I|.$$

In particular, if $\min\{f(a), f(b)\} \leq K_I \leq \max\{f(a), f(b)\}$, we have

$$|\Delta|g([a, b]) = |f(b) - f(a)| \leq |\Delta|f([a, b]).$$

For larger or smaller values of K_I we use that

$$|\Delta|f([a, b]) \geq |f(d) - f(a)| + |f(b) - f(d)| \quad \text{for all } d \in (a, b).$$

Note that there exists $d_1 \in (a, b)$ with $f(d_1) \leq K_I$. Otherwise, $\int_I f(u)du > \int_I K_I du = |I|K_I = \int_I f(u)du$ yields a contradiction. Similarly, there exists $d_2 \in (a, b)$ with $f(d_2) \geq K_I$.

Now suppose $K_I < \min\{f(a), f(b)\}$. Then

$$\begin{aligned} |\Delta|f([a, b]) &\geq |f(d_1) - f(a)| + |f(b) - f(d_1)| \\ &\geq |K_I - f(a)| + |f(b) - K_I| = |\Delta|g([a, b]), \end{aligned}$$

the latter inequality being true since $f(d_1) \leq K_I < \min\{f(a), f(b)\}$.

Analogously, if $K_I > \max\{f(a), f(b)\}$,

$$
\begin{aligned}
|\Delta|f([a,b]) &\geq |f(d_2) - f(a)| + |f(b) - f(d_2)| \\
&\geq |K_I - f(a)| + |f(b) - K_I| = |\Delta|g([a,b]).
\end{aligned}
$$

We conclude that $|\Delta|g([a,b]) \leq |\Delta|f([a,b])$ and thus $|\Delta|g \leq |\Delta|f$.

When I is not open, the proof is more technical but follows the same line of argument as above; therefore, we omit this part of the proof. $\square$

The next two lemmas use the result from Lemma 3.2 and are designed with deriving an upper bound for $\|Q - Q_\alpha\|_\infty$ in mind. Assuming the same properties as those of the functions involved in deriving this upper bound, we show in Lemma 3.3 that flattening a density function leads to an expected value of zero for 'average-zero' functions, and in Lemma 3.4 we show that this operation can be carried out in such a way that the expected value of *piecewise constant* functions does not change.

Lemma 3.3. *Let φ be a bounded function with the property that $\int_I \varphi(x)dx = 0$ for some bounded interval I. Then for every $f \in \mathcal{F}$, there exists $g \in \mathcal{F}$ such that*

(i) $|\Delta|g \leq |\Delta|f$,

(ii) $g(x) = f(x)$ *for* $x \notin I$,

(iii) $\int_I \varphi(x)g(x)dx = 0$, *and*

(iv) $\int \varphi(x)f(x)dx - \int \varphi(x)g(x)dx = \int_I \varphi(x)f(x)dx$.

For example, the pdf g defined in (3.13) satisfies these four properties.

Proof. Let $f \in \mathcal{F}$ be given. Since φ is bounded it follows that $|\int \varphi(x)f(x)dx| < +\infty$. Define $g \in \mathcal{F}$ as in (3.13); hence by Lemma 3.2, properties (i) and (ii) follow. Because of (ii), $\int_{\mathbb{R}\setminus I} \varphi(x)g(x)dx = \int_{\mathbb{R}\setminus I} \varphi(x)f(x)dx$. Moreover, since g has constant value K_I on I, (iii) $\int_I \varphi(x)g(x)dx = K_I \int_I \varphi(x)dx = 0$, and (iv) follows immediately. $\square$

Lemma 3.4. *Let $\varphi : \mathbb{R} \mapsto \mathbb{R}$ be a bounded* piecewise constant *function such that*

$$\varphi(x) := \sum_{j \in J} \varphi_j \mathbb{1}_{I_j}(x),$$

where $\mathbb{1}_I$ is the indicator function of interval I, $\{I_j\}_{j\in J}$ is a countable collection of disjoint bounded intervals of positive length such that $\cup_{j\in J} I_j = \mathbb{R}$, and $\varphi_j \in \mathbb{R}$, $j \in J$.

Let V_φ denote the set of discontinuity points of φ, and assume that V_φ coincides with the endpoints of I_j, $j \in J$, and that $|V_\varphi \cap I|$ is finite for any bounded interval I. Then, for every $f \in \mathcal{F}$ there exists a $g \in \mathcal{F}$ that is piecewise constant *with*

(i) $V_g \subseteq V_\varphi$,

(ii) $|\Delta|g \leq |\Delta|f$, *and*

(iii) $\int \varphi(x)g(x)dx = \int \varphi(x)f(x)dx$.

For example,

$$g(x) := |I_j|^{-1} \int_{I_j} f(u)du \quad \text{for } x \in I_j,\ j \in J \tag{3.14}$$

satisfies these properties.

Proof. Let g be defined as in (3.14) so that g is a piecewise constant density function in $\mathcal{F}$ with (i) $V_g \subseteq V_\varphi$. Moreover, since $\int_{I_j} g(x)dx = \int_{I_j} f(x)dx$ for all $j \in J$, we have that

$$\begin{aligned}
(iii) \int \varphi(x)f(x)dx &= \sum_{j\in J} \int_{I_j} \varphi(x)f(x)dx \\
&= \sum_{j\in J} \varphi_j \int_{I_j} f(x)dx \\
&= \sum_{j\in J} \varphi_j \int_{I_j} g(x)dx \\
&= \int \varphi(x)g(x)dx.
\end{aligned}$$

By applying Lemma 3.2 repeatedly, we also have that (ii) $|\Delta|g \leq |\Delta|f$. □

Remark 3.3. Equivalently to $\int \varphi(x)g(x)dx = \int \varphi(x)f(x)dx$, we can write $\mathbb{E}_g[\varphi(\omega)] = \mathbb{E}_f[\varphi(\omega)]$, where $\mathbb{E}_g$ and $\mathbb{E}_f$ indicate that the expectation is with respect to g and f, respectively.

3.4 Uniform error bound for one-dimensional round-up functions

In the next sections we derive an error bound for the α-approximation Q_α of the TU integer recourse function Q. One of the main difficulties in calculating this error bound is that the maximizing dual vertices λ in (3.11) and (3.12) depend on ω and are possibly different. If it were true that a deterministic $\hat{\lambda}$ exists such that

$$Q(z) = \mathbb{E}_\omega\left[\max_{k=1,\dots,K} \lambda^k \lceil \omega - z \rceil\right] \leq \mathbb{E}_\omega\left[\hat{\lambda} \lceil \omega - z \rceil\right]$$

and

$$Q_\alpha(z) = \mathbb{E}_\omega\left[\max_{k=1,\dots,K} \lambda^k (\lceil \omega \rceil_\alpha - z)\right] \geq \mathbb{E}_\omega\left[\hat{\lambda}(\lceil \omega \rceil_\alpha - z)\right],$$

then

$$Q(z) - Q_\alpha(z) \leq \mathbb{E}_\omega\left[\hat{\lambda}(\lceil \omega \rceil_z - \lceil \omega \rceil_\alpha)\right] = \sum_{i=1}^m \hat{\lambda}_i \mathbb{E}_{\omega_i}\left[\lceil \omega_i \rceil_{z_i} - \lceil \omega_i \rceil_{\alpha_i}\right],$$

so that we obtain an error bound if we derive a bound on each component of $\mathbb{E}_\omega[\lceil \omega \rceil_z - \lceil \omega \rceil_\alpha]$. In this section we analyze this simplified one-dimensional bounding problem. It can be solved by clever application of flattening of densities, using the special properties of the underlying difference function. Surprisingly, it turns out that the uniform upper bound of this hypothesized α-approximation is very useful for the TU model, to be discussed in the next section. As we will show then, a suitable relaxation of the set of dual vertices λ to a set with deterministic pointwise supremum λ^* is possible, and together with suitable flattening of the densities involved an error bound will be derived.

Definition 3.4 (Difference function)**.** For every $\alpha \in \mathbb{R}$, $z \in \mathbb{R}$, define the *difference function* $\varphi_{\alpha,z}$ as

$$\varphi_{\alpha,z}(x) := \lceil x \rceil_z - \lceil x \rceil_\alpha = \lceil x - z \rceil + z - \lceil x - \alpha \rceil - \alpha, \quad x \in \mathbb{R}.$$

Moreover, for every $\alpha \in \mathbb{R}$, $z \in \mathbb{R}$, define the *expected difference function* $D_{\alpha,z} : \mathcal{F} \mapsto$

$\mathbb{R}$ as

$$D_{\alpha,z}(f) := \mathbb{E}_f[\varphi_{\alpha,z}(\omega)], \quad f \in \mathcal{F}.$$

Remark 3.4. For fixed $\alpha \in \mathbb{R}$ and $f \in \mathcal{F}$, the expected difference function $D_{\alpha,z}(f)$ can be interpreted as the difference between the round-up function $R(z) := \mathbb{E}_\omega[\lceil \omega - z \rceil]$, $z \in \mathbb{R}$, and its α-approximation $\mathbb{E}_\omega[(\lceil \omega \rceil_\alpha - z)]$, where the expectations are with respect to the pdf f.

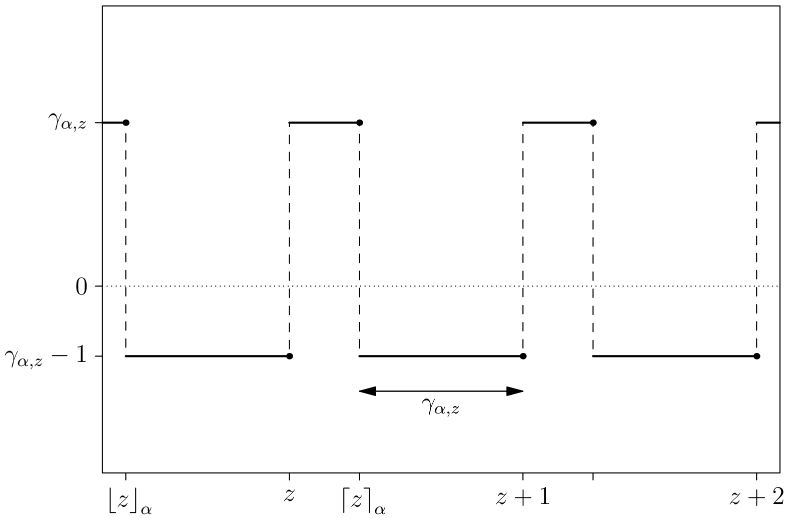

Figure 3.1: The difference function $\varphi_{\alpha,z}$ from Definition 3.4.

The following properties of the difference function $\varphi_{\alpha,z}$ are illustrated in Figure 3.1.

Lemma 3.5 (Properties of the difference function)**.** *Consider the difference function* $\varphi_{\alpha,z}(x) := \lceil x \rceil_z - \lceil x \rceil_\alpha$, $x \in \mathbb{R}$.

(i) $\varphi_{\alpha,z}$ *is periodic in* $x, \alpha,$ *and* z *with period 1, and moreover* $\varphi_{\alpha,z}(x) = -\varphi_{z,\alpha}(x)$.

(ii) If $\alpha - z \in \mathbb{Z}$, *then* $\varphi_{\alpha,z} \equiv 0$.

(iii) If $\alpha - z \notin \mathbb{Z}$*, then* $\varphi_{\alpha,z}$ *is a two-valued function*

$$\varphi_{\alpha,z}(x) = \begin{cases} \gamma_{\alpha,z}, & x \in \cup_{l\in\mathbb{Z}}(z+l, \lceil z \rceil_\alpha + l], \\ \gamma_{\alpha,z} - 1, & x \in \cup_{l\in\mathbb{Z}}(\lfloor z \rfloor_\alpha + l, z+l], \end{cases} \tag{3.15}$$

with

$$\gamma_{\alpha,z} := z - \lfloor z \rfloor_\alpha = z + 1 - \lceil z \rceil_\alpha \in (0,1).$$

Thus, $\varphi_{\alpha,z}$ *has jumps of size* $+1$ *on* $z+\mathbb{Z}$ *and jumps of size* -1 *on* $\alpha + \mathbb{Z}$*, and it is left-continuous.*

(iv) $\int_I \varphi_{\alpha,z}(x)dx = 0$ *for any interval* I *of length* $|I| = 1$*.*

Proof. Properties (i) and (ii) are obvious. (iii) Since $\lceil x - y \rceil + y$ is a piecewise constant (left-continuous) function with jumps of size $+1$ on $y + \mathbb{Z}$, it follows that $\varphi_{\alpha,z}$ is piecewise constant (left-continuous) with jumps of size $+1$ on $z+\mathbb{Z}$ and jumps of size -1 on $\alpha + \mathbb{Z}$.

Note that for $x \in (z, \lceil z \rceil_\alpha]$,

$$\varphi_{\alpha,z}(x) = z + 1 - \lceil z - \alpha \rceil - \alpha = z + 1 - \lceil z \rceil_\alpha = z - \lfloor z \rfloor_\alpha = \gamma_{\alpha,z} \in (0,1).$$

Since $\varphi_{\alpha,z}$ has jumps of size -1 on $\alpha + \mathbb{Z}$, it follows that

$$\varphi_{\alpha,z}(x) = \gamma_{\alpha,z} - 1 \quad \text{for } x \in (\lceil z \rceil_\alpha, z+1].$$

Since $\varphi_{\alpha,z}$ is periodic with period 1, (3.15) holds. Moreover, we have

$$\int_{\lfloor z \rfloor_\alpha}^{\lceil z \rceil_\alpha} \varphi_{\alpha,z}(x)dx = \int_{\lfloor z \rfloor_\alpha}^{z} \varphi_{\alpha,z}(x)dx + \int_{z}^{\lceil z \rceil_\alpha} \varphi_{\alpha,z}(x)dx = 0,$$

since

$$\int_{\lfloor z \rfloor_\alpha}^{z} \varphi_{\alpha,z}(x)dx = (z - \lfloor z \rfloor_\alpha)(\gamma_{\alpha,z} - 1) = -\gamma_{\alpha,z}(1 - \gamma_{\alpha,z}) \tag{3.16}$$

and

$$\int_{z}^{\lceil z \rceil_\alpha} \varphi_{\alpha,z}(x)dx = (\lceil z \rceil_\alpha - z)\gamma_{\alpha,z} = (1-\gamma_{\alpha,z})\gamma_{\alpha,z}. \tag{3.17}$$

From the periodicity of $\varphi_{\alpha,z}$ it now follows that (iv) $\int_I \varphi_{\alpha,z}(x)dx = 0$ for any interval I of length $|I| = 1$. □

Lemma 3.5 directly implies the following properties of the expected difference function $D_{\alpha,z}$.

Corollary 3.1. *For every $f \in \mathcal{F}$,*

(i) $D_{\alpha,z}(f)$ is periodic in both α and z with period 1,

(ii) $D_{\alpha,z}(f) = -D_{z,\alpha}(f)$,

(iii) $D_{\alpha,z}(f) = 0$ if $\alpha - z \in \mathbb{Z}$, and

(iv) $|D_{\alpha,z}(f)| \leq 1$.

After these technical preparations we are ready to derive a nontrivial upper bound for $|D_{\alpha,z}(f)|$. Obviously, for any given $f_0 \in \mathcal{F}$ and any $\alpha \in \mathbb{R}$ the sharpest upper bound is

$$\mathcal{M}(\alpha, f_0) := \sup_{z\in\mathbb{R}} |D_{\alpha,z}(f_0)|. \tag{3.18}$$

However, it is practically impossible to calculate this bound. Surprisingly, a kind of worst-case analysis appears to be very helpful. Instead of considering f_0 which has $|\Delta| f_0 = B_0$, we will solve, for all $B > 0$, the optimization problem

$$M(B) := \sup_{\alpha\in\mathbb{R}} \sup_{f\in\mathcal{F}} \Big\{ \mathcal{M}(\alpha, f) : |\Delta| f \leq B \Big\},$$

so that $M(B_0)$ is an upper bound for $\mathcal{M}(\alpha, f_0)$. This key result is contained in Theorem 3.1, concluding this section. Observe that $M(B)$ exists since $|D_{\alpha,z}(f)| \leq 1$ for all $f \in \mathcal{F}$ by Corollary 3.1 (iv).

We first explain why the worst-case approach works. By interchanging supremizations and using $D_{\alpha,z}(f) = -D_{z,\alpha}(f)$, it follows that

$$M(B) = \sup_{\alpha\in\mathbb{R}} \sup_{z\in\mathbb{R}} \sup_{f\in\mathcal{F}} \Big\{ |D_{\alpha,z}(f)| : |\Delta| f \le B \Big\}$$
$$= \sup_{\alpha\in\mathbb{R}} \sup_{z\in\mathbb{R}} \sup_{f\in\mathcal{F}} \Big\{ D_{\alpha,z}(f) : |\Delta| f \le B \Big\}. \tag{3.19}$$

We will show that the inner supremization,

$$(\mathcal{P}) \quad \sup_{f\in\mathcal{F}} \Big\{ D_{\alpha,z}(f) : |\Delta| f \le B \Big\},$$

with fixed α and z, can be solved explicitly, using the tools of Section 3.3.

Proposition 3.1. *Let $\alpha, z \in \mathbb{R}$ be given. Then, for every $B > 0$,*

$$\sup_{f\in\mathcal{F}} \Big\{ D_{\alpha,z}(f) : |\Delta| f \le B \Big\} = \min \Big\{ \gamma_{\alpha,z}, \gamma_{\alpha,z}(1-\gamma_{\alpha,z})\frac{B}{2} \Big\}, \tag{3.20}$$

with $\gamma_{\alpha,z} := z - \lfloor z \rfloor_\alpha$.

Proof. If $\alpha - z \in \mathbb{Z}$ so that $\gamma_{\alpha,z} = 0$, then Corollary 3.1 (iii) shows that $D_{\alpha,z}(f) = 0$ for all $f \in \mathcal{F}$ so that $\sup_{f\in\mathcal{F}}\{D_{\alpha,z}(f) : |\Delta| f \le B\} = 0$ and thus (3.20) holds.

If $\alpha - z \notin \mathbb{Z}$, then the difference function $\varphi_{\alpha,z}$ is piecewise constant with $V_{\varphi_{\alpha,z}} = (\alpha + \mathbb{Z}) \cup (z + \mathbb{Z})$ so that it satisfies the conditions of Lemma 3.4. Application of this lemma shows that for every feasible g of maximization problem $(\mathcal{P})$ there exists a *piecewise constant* feasible solution f with the same objective value, and with $V_f \subset V_{\varphi_{\alpha,z}}$. Hence, we can (and will) restrict the feasible region of $(\mathcal{P})$ to piecewise constant density functions f with $V_f \subset (\alpha + \mathbb{Z}) \cup (z + \mathbb{Z})$. We will denote its function values to the left of $z + l$ by f_l^- and those to the right of $z + l$ by f_l^+; that is

$$f(x) = \begin{cases} f_l^-, & \text{for } x \in (\lfloor z \rfloor_\alpha + l, z + l], \quad l \in \mathbb{Z}, \\ f_l^+, & \text{for } x \in (z + l, \lceil z \rceil_\alpha + l], \quad l \in \mathbb{Z}. \end{cases}$$

Consider such feasible $f \in \mathcal{F}$. We will derive necessary optimality conditions on its function values by applying Lemma 3.3 with $\varphi = \varphi_{\alpha,z}$ and I an arbitrary interval with $|I| = 1$. Lemma 3.5 (iv) shows that the conditions of Lemma 3.3 are satisfied. Lemma 3.3 (i, iv) shows that a feasible g exists such that $D_{\alpha,z}(f) - D_{\alpha,z}(g) =$

$\int_I \varphi_{\alpha,z}(x)f(x)dx$. If the right-hand side happens to be negative, f cannot be optimal for $(\mathcal{P})$ since g has a better objective value. Hence, for each interval I with $|I| = 1$ we have the following necessary optimality condition for f in $(\mathcal{P})$:

$$\int_I \varphi_{\alpha,z}(x)f(x)dx \geq 0.$$

In particular, for $I = (z+l-1, z+l]$ and $I = (\lfloor z \rfloor_\alpha + l, \lceil z \rceil_\alpha + l]$, $l \in \mathbb{Z}$, it can be derived from (3.16) and (3.17) that

$$\int_{z+l-1}^{z+l} \varphi_{\alpha,z}(x)f(x)dx = \gamma_{\alpha,z}(1-\gamma_{\alpha,z})\{f_{l-1}^+ - f_l^-\},$$

and

$$\int_{\lfloor z \rfloor_\alpha + l}^{\lceil z \rceil_\alpha + l} \varphi_{\alpha,z}(x)f(x)dx = \gamma_{\alpha,z}(1-\gamma_{\alpha,z})\{f_l^+ - f_l^-\},$$

yielding the optimality conditions

$$f_{l-1}^+ \geq f_l^-, \quad l \in \mathbb{Z},$$

and

$$f_l^+ \geq f_l^-, \quad l \in \mathbb{Z}.$$

Under these restrictions f is a piecewise constant density function whose value alternately increases and decreases. For such density functions the total variation can be expressed as $|\Delta|f = 2\sum_{l\in\mathbb{Z}}\{f_l^+ - f_l^-\}$, i.e., as the sum of its total increase and total decrease. Moreover, using (3.16), (3.17), and the periodicity of $\varphi_{\alpha,z}$, we have that

$$\begin{aligned} D_{\alpha,z}(f) &= \int \varphi_{\alpha,z}(x)f(x)dx \\ &= \sum_{l\in\mathbb{Z}} \left\{ f_l^- \int_{\lfloor z \rfloor_\alpha + l}^{z+l} \varphi_{\alpha,z}(x)dx + f_l^+ \int_{z+l}^{\lceil z \rceil_\alpha + l} \varphi_{\alpha,z}(x)dx \right\} \\ &= \gamma_{\alpha,z}(1-\gamma_{\alpha,z}) \sum_{l\in\mathbb{Z}} \left\{ f_l^+ - f_l^- \right\}. \end{aligned}$$

Hence, problem $(\mathcal{P})$ reduces to the optimization problem

$$\begin{aligned}
\sup_{f_l^+, f_l^-} \quad & D_{\alpha,z}(f) = \gamma_{\alpha,z}(1-\gamma_{\alpha,z}) \sum_{l\in\mathbb{Z}} \left\{ f_l^+ - f_l^- \right\} \\
\text{s.t.} \quad & \sum_{l\in\mathbb{Z}} \left\{ (1-\gamma_{\alpha,z}) f_l^+ + \gamma_{\alpha,z} f_l^- \right\} = 1 && (3.21) \\
& \sum_{l\in\mathbb{Z}} \left\{ f_l^+ - f_l^- \right\} \leq \frac{B}{2} && (3.22) \\
& f_l^+ \geq f_l^-, \quad f_{l-1}^+ \geq f_l^-, \quad l \in \mathbb{Z} && (3.23) \\
& f_l^+ \geq 0, \quad f_l^- \geq 0, \qquad l \in \mathbb{Z} && (3.24)
\end{aligned}$$

Here, (3.24) ensures that f is non-negative, (3.21) that f integrates to 1, and (3.22) that $|\Delta| f \leq B$, whereas the inequalities in (3.23) represent the necessary optimality conditions derived above. Notice that the variables f_l^+ have a positive coefficient in the objective and f_l^- a negative one.

We solve this reduced version of $(\mathcal{P})$ by providing an upper bound which we subsequently prove to be tight. On one hand (3.22) implies that

$$D_{\alpha,z}(f) \leq \gamma_{\alpha,z}(1-\gamma_{\alpha,z}) \frac{B}{2}, \tag{3.25}$$

and on the other hand, since (3.21) is equivalent to

$$(1-\gamma_{\alpha,z}) \sum_{l\in\mathbb{Z}} \{ f_l^+ - f_l^- \} = 1 - \sum_{l\in\mathbb{Z}} f_l^-,$$

we have

$$\begin{aligned}
D_{\alpha,z}(f) &= \gamma_{\alpha,z}(1-\gamma_{\alpha,z}) \sum_{l\in\mathbb{Z}} \{ f_l^+ - f_l^- \} \\
&= \gamma_{\alpha,z} \left(1 - \sum_{l\in\mathbb{Z}} f_l^- \right) \\
&\leq \gamma_{\alpha,z}, && (3.26)
\end{aligned}$$

since $\sum_{l\in\mathbb{Z}} f_l^- \geq 0$. Combining the upper bounds in (3.25) and (3.26) yields, for every $f \in \mathcal{F}$ with $|\Delta| f \leq B$,

$$\begin{aligned} D_{\alpha,z}(f) &\leq \min\{\gamma_{\alpha,z}, \gamma_{\alpha,z}(1-\gamma_{\alpha,z})B/2\} \\ &= \begin{cases} \gamma_{\alpha,z}, & \text{if } \gamma_{\alpha,z} \leq 1 - 2/B, \\ \gamma_{\alpha,z}(1-\gamma_{\alpha,z})B/2, & \text{if } \gamma_{\alpha,z} \geq 1 - 2/B. \end{cases} \end{aligned}$$

Consider first the case $0 < \gamma_{\alpha,z} \leq 1 - 2/B$ (i.e., $(1-\gamma_{\alpha,z})^{-1} \leq B/2$). Then the density $\hat{f}$ with

$$\hat{f}_0^- = 0,\ \hat{f}_0^+ = c,\ \hat{f}_l^- = \hat{f}_l^+ = 0 \quad \text{for all } l \in \mathbb{Z}\backslash\{0\}$$

satisfies all constraints (3.21)–(3.24) if $c := (1-\gamma_{\alpha,z})^{-1}$, and the objective value $D_{\alpha,z}(\hat{f})$ equals $\gamma_{\alpha,z}$, indeed.

Consider next the case $1 - 2/B < \gamma_{\alpha,z} < 1$ (so that $(1-\gamma_{\alpha,z})B/2 < 1$). Then the density $\bar{f}$ with

$$\begin{aligned} \bar{f}_0^- = 0,\ \bar{f}_0^+ = B/2,\ \bar{f}_l^- &= \bar{f}_l^+ = c \quad l = 1, \dots, k \\ \bar{f}_l^- &= \bar{f}_l^+ = 0 \quad l < 0,\ l > k \end{aligned}$$

satisfies all constraints (3.21)–(3.24) if

$$\begin{aligned} (1-\gamma_{\alpha,z})B/2 + kc = 1 & \qquad \text{(from (3.21))} \\ 0 \leq c \leq B/2 & \qquad (\text{from } 0 \leq \bar{f}_1^- \leq \bar{f}_0^+) \end{aligned}$$

and these are satisfied by $k = k^*$, $c = c^*$ given by

$$k^* := \min_{k \in \mathbb{Z}}\{k : (1-\gamma_{\alpha,z})B/2 + kB/2 \geq 1\} = \lceil \gamma_{\alpha,z} - (1 - 2/B) \rceil \tag{3.27}$$

$$c^* := (1 - (1-\gamma_{\alpha,z})B/2)/k^*. \tag{3.28}$$

The objective value $D_{\alpha,z}(\bar{f})$ equals $\gamma_{\alpha,z}(1-\gamma_{\alpha,z})B/2$, indeed. □

It is interesting to picture the optimal densities $\hat{f}$ and $\bar{f}$ from the proof of Proposition 3.1 because for these densities the error of the α-approximation is largest. Obviously, the shape of such an optimal density will depend on the value of B.

For large values of B, the constraint on the total variation of f is not very restrictive. Therefore, it is not hard to imagine that (since $\varphi_{\alpha,z}$ is two-valued with maximum value $\gamma_{\alpha,z}$) it might be possible to attain the upper bound $\gamma_{\alpha,z}$ by setting $f(x) > 0$ if and only if $\varphi_{\alpha,z}(x) = \gamma_{\alpha,z} > 0$. It turns out that this is indeed possible if $\gamma_{\alpha,z} \leq 1 - 2/B$. For example, the pdf $\hat{f}$ defined as

$$\hat{f}(x) = \begin{cases} (1-\gamma_{\alpha,z})^{-1}, & z < x \leq \lceil z \rceil_\alpha \\ 0, & \text{otherwise,} \end{cases} \tag{3.29}$$

has objective value $D_{\alpha,z}(\hat{f}) = \gamma_{\alpha,z}$.

For smaller values of B for which $1-2/B < \gamma_{\alpha,z} < 1$, the pdf $\hat{f}$ is infeasible because it violates the total variation constraint. In fact, any pdf f with $D_{\alpha,z}(f) = \gamma_{\alpha,z}$ now violates this constraint, so that intuitively any optimal pdf f must satisfy $|\Delta|f = B$. An example of such an optimal density is given by the pdf $\bar{f}$ in Figure 3.2, defined as

$$\bar{f}(x) = \begin{cases} B/2, & x \in (z, \lceil z \rceil_\alpha] \\ c^*, & x \in (\lceil z \rceil_\alpha, \lceil z \rceil_\alpha + k^*] \\ 0, & \text{otherwise,} \end{cases} \tag{3.30}$$

with k^* and c^* as defined in (3.27) and (3.28), respectively. Indeed, it can be shown that any pdf f that is piecewise constant with $V_f \subset (\alpha+\mathbb{Z}) \cup (z+\mathbb{Z})$ satisfying (3.21), (3.23), (3.24), and $|\Delta|f = B$ is optimal with objective value $D_{\alpha,z}(f) = D_{\alpha,z}(\bar{f}) = \gamma_{\alpha,z}(1-\gamma_{\alpha,z})B/2$.

Now that we have solved the inner optimization problem ($\mathcal{P}$) explicitly, it is easy to find an upper bound for $\mathcal{M}(\alpha, f)$.

Theorem 3.1 (Error bound for the expected difference function)**.** *For every $\alpha \in \mathbb{R}$ and every random variable ω with pdf $f \in \mathcal{F}$,*

$$\mathcal{M}(\alpha, f) := \sup_{z\in\mathbb{R}} |D_{\alpha,z}(f)| \leq h(|\Delta|f),$$

where $h : (0,\infty) \mapsto \mathbb{R}$ is given by

$$h(x) = \begin{cases} x/8, & 0 < x \leq 4, \\ 1 - 2/x, & x \geq 4. \end{cases} \tag{3.31}$$

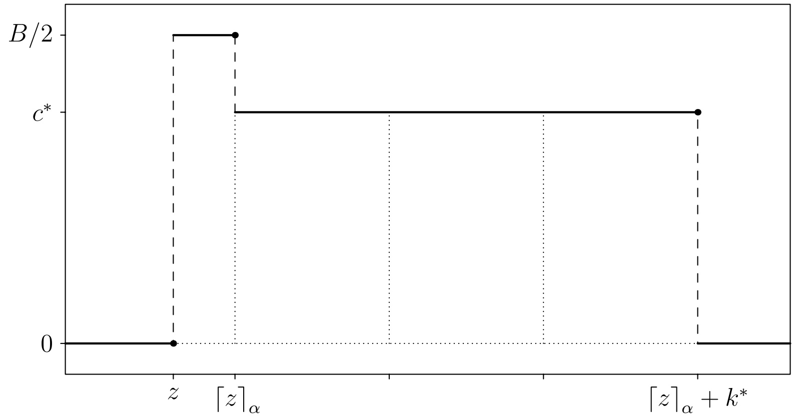

Figure 3.2: The pdf $\bar{f}$ defined in (3.30) with $k^* = 3$.

Proof. Let $f_0 \in \mathcal{F}$ with $|\Delta| f_0 = B_0$ be given. Then, $M(B_0)$ with M as in (3.19) is an upper bound of $\mathcal{M}(\alpha, f_0)$. Using Proposition 3.1, we have that

$$M(B_0) = \sup_{\alpha \in \mathbb{R}} \sup_{z \in \mathbb{R}} \min \left\{ \gamma_{\alpha,z}, \gamma_{\alpha,z}(1 - \gamma_{\alpha,z}) \frac{B_0}{2} \right\},$$

with $\gamma_{\alpha,z} := z - \lfloor z \rfloor_\alpha \in [0,1)$. Hence, it follows that

$$M(B_0) = \sup_{\gamma \in [0,1)} \min \left\{ \gamma, \gamma(1-\gamma) \frac{B_0}{2} \right\}.$$

In this optimization problem we have to maximize the minimum of a linear and a quadratic function over the domain $[0,1)$. Elementary analysis shows that the optimal solution is given by $\gamma_{B_0} := \max\{1/2, 1 - 2/B_0\}$, whereas the optimal value is equal to $h(B_0)$, where h is as defined in (3.31). □

As argued before, for every $\alpha \in \mathbb{R}$, $z \in \mathbb{R}$, and $f \in \mathcal{F}$, the expected difference function $D_{\alpha,z}(f)$ satisfies $|D_{\alpha,z}(f)| \leq 1$. Theorem 3.1 shows that only for density functions $f \in \mathcal{F}$ with a large total variation $|\Delta| f$ the expected difference function $|D_{\alpha,z}(f)|$ may be close to this trivial upper bound. It turns out that these are exactly the type of density functions for which $\|Q - Q_\alpha\|_\infty$ and its upper bound, to be derived in the next sections, are large, suggesting that (depending on the problem and desired accuracy) the error bound of the approximation may be too large for the practical problem at hand.

However, the next example shows that for many density functions $f \in \mathcal{F}$ that are expected to arise in practice, the value of $h(|\Delta|f)$ and thus also of $|D_{\alpha,z}(f)|$ is actually much smaller than 1, suggesting that in these cases $\|Q-Q_\alpha\|_\infty$ and its upper bound may be small enough for practical purposes.

Example 3.1. Let $f \in \mathcal{F}$ be the density function corresponding to a normal distribution with mean μ and variance σ^2; that is, $f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\{-\frac{1}{2\sigma^2}(x-\mu)^2\}, x \in \mathbb{R}$. This density function is unimodal with maximum value $\frac{1}{\sqrt{2\pi\sigma^2}}$ at $x = \mu$, so that $|\Delta|f = 2\frac{1}{\sqrt{2\pi\sigma^2}} = \sigma^{-1}\sqrt{2/\pi}$. For example, if $\sigma = 1$, then $|\Delta|f \approx 0.789$ and $h(|\Delta|f) \approx 0.0997$. $\triangleleft$

In general, there is no one-to-one correspondence between the standard deviation σ and the total variation $|\Delta|f$ as in the case of the normal distribution. In fact, $|\Delta|f$ is not even a measure of dispersion of the distribution of ω. For example, for unimodal density functions it is uniquely determined by the mode of f and does not depend on the shape of f. Thus, $|\Delta|f$ is small if f resembles a uniform distribution with a large support, and $|\Delta|f$ is large if f has one or more high peaks (with $|\Delta|f \to +\infty$ as f approximates a discrete distribution).

Table 3.1 specifies values of $|\Delta|f$ and $h(|\Delta|f)$ for various instances of well-known density functions. It is good to keep in mind that $h(|\Delta|f)$ represents a worst-case bound for $|D_{\alpha,z}(f)|$ so that in practice the actual value of $|D_{\alpha,z}(f)|$ may be much lower than $h(|\Delta|f)$. For example, for the two uniform density functions with integer length support in Table 3.1, $D_{\alpha,z}(f) = 0$ for all $z \in \mathbb{R}$ if $\alpha = 0$.

Table 3.1: Values of $|\Delta|f$ and $h(|\Delta|f)$ for several well-known density functions.

Distribution	$f(x)$	Parameter value(s)	$\|\Delta\|f$	$h(\|\Delta\|f)$
Normal	$\frac{1}{\sqrt{2\pi\sigma^2}} \exp\{-\frac{1}{2\sigma^2}(x-\mu)^2\}$	$\sigma = 0.1$	7.98	0.749
–	–	$\sigma = 1$	0.798	0.0997
–	–	$\sigma = 10$	0.0798	0.00997
Exponential	$\lambda \exp\{-\lambda x\}\mathbb{1}_{(0,\infty)}(x)$	$\lambda = 1$	2	0.25
–	–	$\lambda = 0.1$	0.2	0.025
Uniform	$\frac{1}{b-a}\mathbb{1}_{[a,b]}(x)$	$a = 0, b = 1$	2	0.25
–	–	$a = 0, b = 10$	0.2	0.025

3.5 TU integer recourse models with independent random variables

Now we have set the stage for the analysis of TU integer recourse models. To avoid obscuring technicalities we first assume that the components of the m-dimensional random right-hand side vector ω are independently distributed and that the joint density function f of ω is contained in $\mathcal{F}^m$ defined below. We will deal with dependent distributions in the next section.

Definition 3.5. Let $\mathcal{F}^m$ denote the set of m-dimensional joint density functions f whose marginal densities $f_i, i = 1, \ldots, m$, are contained in $\mathcal{F}$, and for which

$$f(x) = \prod_{i=1}^{m} f_i(x_i), \quad x \in \mathbb{R}^m.$$

We will derive an error bound for the α-approximation Q_α of the TU integer recourse function Q given by (3.12) and (3.11), respectively. Similar to the expected difference function in Section 3.4, for almost any given $f \in \mathcal{F}^m$ with $|\Delta| f_i = B_i$ and $\alpha \in \mathbb{R}^m$, direct calculation of the sharpest upper bound

$$\mathcal{N}(\alpha, f) := \sup_{z \in \mathbb{R}^m} |Q(z) - Q_\alpha(z)|$$

is too demanding. As already mentioned, the main difficulty in calculating this bound is that the maximizing dual vertices λ in (3.11) and (3.12) depend on ω, and are possibly different for a given ω. In order to overcome this difficulty, we relax the set of possible dual vertices and use a worst-case analysis over this relaxed set. As we will see, this approach, combined with the analysis of the one-dimensional expected difference function, yields the desired upper bound.

Consider, therefore, the TU integer expected value function Q, and pick for every $z \in \mathbb{R}^m$ a function $\lambda_Q^z : \mathbb{R}^m \mapsto \mathbb{R}^m$ such that

$$\lambda_Q^z(x) \in \operatorname*{argmax}_{k=1,\ldots,K} \lambda^k \lceil x - z \rceil, \quad x \in \mathbb{R}^m, \tag{3.32}$$

and λ_Q^z is constant on

$$C_z^l := \prod_{i=1}^{m} C_{z_i}^{l_i} := \prod_{i=1}^{m} (z_i + l_i - 1, z_i + l_i]$$

for every $l \in \mathbb{Z}^m$. This is indeed possible since $\lceil x - z \rceil$ is constant on C_z^l. Analogously, associated with Q_α, pick for every $\alpha \in \mathbb{R}^m$ and $z \in \mathbb{R}^m$, a function

$$\lambda_{Q_\alpha}^z(x) \in \operatorname*{argmax}_{k=1,\dots,K} \lambda^k(\lceil x \rceil_\alpha - z), \quad x \in \mathbb{R}^m,$$

such that $\lambda_{Q_\alpha}^z$ is constant on C_α^l for every $l \in \mathbb{Z}^m$. Now we can rewrite Q and Q_α as $Q(z) = \mathbb{E}_\omega[\lambda_Q^z(\omega) \lceil \omega - z \rceil]$ and $Q_\alpha(z) = \mathbb{E}_\omega[\lambda_{Q_\alpha}^z(\omega)(\lceil \omega \rceil_\alpha - z)]$, respectively.

Note that λ_Q^z and $\lambda_{Q_\alpha}^z$ have three important properties in common. First, both functions are nonnegative. Second, both functions are bounded by $\lambda^* \in \mathbb{R}^m$ defined as

$$\lambda_i^* := \max_{k=1,\dots,K} \lambda_i^k, \quad i = 1, \dots, m, \tag{3.33}$$

and third, for both functions there exists $\beta \in \mathbb{R}^m$ such that the function is constant on C_β^l for every $l \in \mathbb{Z}^m$. These three properties are paramount to obtaining an upper bound for $\mathcal{N}(\alpha, f)$, as we show now.

Definition 3.6. Let Λ^m denote the set of functions $\lambda : \mathbb{R}^m \mapsto \mathbb{R}^m$ for which

(i) $0 \le \lambda(x) \le \lambda^*$ for every $x \in \mathbb{R}^m$, and

(ii) there exists $\beta \in \mathbb{R}^m$ such that λ is constant on C_β^l for every $l \in \mathbb{Z}^m$.

Definition 3.7. For every $\alpha \in \mathbb{R}^m, z \in \mathbb{R}^m$, define $G_{\alpha,z} : \Lambda^m \times \mathcal{F}^m \mapsto \mathbb{R}$ as

$$G_{\alpha,z}(\lambda, f) := \mathbb{E}_f\Big[\lambda(\omega)\Big(\lceil \omega \rceil_z - \lceil \omega \rceil_\alpha\Big)\Big],$$

where $\lambda \in \Lambda^m$ and $f \in \mathcal{F}^m$.

Lemma 3.6. *For every $\hat{\alpha} \in \mathbb{R}^m$ and every $f \in \mathcal{F}^m$,*

$$\mathcal{N}(\hat{\alpha}, f) \le \mathcal{N}^*(f) := \sup_{\alpha \in \mathbb{R}^m} \sup_{z \in \mathbb{R}^m} \sup_{\lambda \in \Lambda^m} G_{\alpha,z}(\lambda, f).$$

Proof. Let $\hat{\alpha} \in \mathbb{R}^m$ and $f \in \mathcal{F}^m$ be given. We will show that for every $z \in \mathbb{R}^m$,

$$Q(z) - Q_{\hat{\alpha}}(z) \leq \sup_{\lambda \in \Lambda^m} G_{\hat{\alpha},z}(\lambda, f),$$

and

$$Q_{\hat{\alpha}}(z) - Q(z) \leq \sup_{\lambda \in \Lambda^m} G_{z,\hat{\alpha}}(\lambda, f),$$

implying that

$$\sup_{z \in \mathbb{R}^m} |Q(z) - Q_{\hat{\alpha}}(z)| \leq \sup_{\alpha \in \mathbb{R}^m} \sup_{z \in \mathbb{R}^m} \sup_{\lambda \in \Lambda^m} G_{\alpha,z}(\lambda, f)$$

as postulated.

To prove the first inequality, let $z \in \mathbb{R}^m$ be given, and consider the function λ_Q^z as defined in (3.32). Note that $\lambda_Q^z(x)$ is a maximizer of $\max_{k=1,\dots,K} \lambda^k \lceil x - z \rceil$ for every $x \in \mathbb{R}^m$, but not necessarily of $\max_{k=1,\dots,K} \lambda^k (\lceil x \rceil_{\hat{\alpha}} - z)$. Thus,

$$Q(z) - Q_{\hat{\alpha}}(z) \leq \mathbb{E}_\omega \Big[\lambda_Q^z(\omega) \Big\{ \lceil \omega \rceil_z - \lceil \omega \rceil_{\hat{\alpha}} \Big\} \Big] = G_{\hat{\alpha},z}(\lambda_Q^z, f).$$

Since $\lambda_Q^z \in \Lambda^m$, the first inequality follows. Analogously, the second inequality follows from

$$Q_{\hat{\alpha}}(z) - Q(z) \leq \mathbb{E}_\omega \Big[\lambda_{Q_{\hat{\alpha}}}^z(\omega) \Big\{ \lceil \omega \rceil_{\hat{\alpha}} - \lceil \omega \rceil_z \Big\} \Big] = G_{z,\hat{\alpha}}(\lambda_{Q_{\hat{\alpha}}}^z, f). \qquad \square$$

The final step in our analysis comprises a similar worst-case analysis as carried out for the one-dimensional case in the previous section. For all $B \in \mathbb{R}^m$ with $B > 0$ we consider the optimization problem

$$\begin{aligned} N(B) &:= \sup_{f \in \mathcal{F}^m} \Big\{ \mathcal{N}^*(f) : |\Delta| f_i \leq B_i,\ i = 1, \dots, m \Big\} \qquad (3.34) \\ &= \sup_{\alpha \in \mathbb{R}^m} \sup_{z \in \mathbb{R}^m} \sup_{f \in \mathcal{F}^m} \sup_{\lambda \in \Lambda^m} \Big\{ G_{\alpha,z}(\lambda, f) : |\Delta| f_i \leq B_i,\ i = 1, \dots, m \Big\}. \end{aligned}$$

The following proposition allows us to reduce the problem to one involving the constant function $\lambda \equiv \lambda^*$, with λ^* as defined in (3.33).

Proposition 3.2. *For every $\alpha \in \mathbb{R}^m, z \in \mathbb{R}^m, \lambda \in \Lambda^m$, and $f \in \mathcal{F}^m$, there exists $g \in \mathcal{F}^m$ with $|\Delta| g_i \leq |\Delta| f_i, i = 1, \dots, m$, such that $G_{\alpha,z}(\lambda, f) \leq G_{\alpha,z}(\lambda, g) \leq G_{\alpha,z}(\lambda^*, g)$.*

Proof. Let $\alpha \in \mathbb{R}^m, z \in \mathbb{R}^m, \lambda \in \Lambda^m$, and $f \in \mathcal{F}^m$ be given with λ constant on every C^l_β for some $\beta \in \mathbb{R}^m$. Observe that

$$\begin{aligned} G_{\alpha,z}(\lambda, f) &:= \mathbb{E}_\omega \left[\lambda(\omega)\Big(\lceil \omega \rceil_z - \lceil \omega \rceil_\alpha \Big)\right] \\ &= \mathbb{E}_\omega \left[\sum_{i=1}^m \lambda_i(\omega)\varphi_{\alpha_i,z_i}(\omega_i)\right] \\ &= \sum_{i=1}^m \int_{\mathbb{R}^m} \lambda_i(x)\varphi_{\alpha_i,z_i}(x_i)f(x)dx, \end{aligned}$$

where φ_{α_i,z_i} is the one-dimensional difference function introduced in Definition 3.4. Here, the last equality is obtained by writing the expectation $\mathbb{E}_\omega$ as an integral, and by interchanging summation and integration. Since λ is constant on C^l_β for every l, we can calculate the expected value on each C^l_β separately:

$$\begin{aligned} G_{\alpha,z}(\lambda, f) &= \sum_{i=1}^m \sum_{l\in\mathbb{Z}^m} \int_{C^l_\beta} \lambda_i(x)\varphi_{\alpha_i,z_i}(x_i)f(x)dx \\ &= \sum_{i=1}^m \sum_{l\in\mathbb{Z}^m} \lambda_i(l+\beta) \int_{C^l_\beta} \varphi_{\alpha_i,z_i}(x_i)f(x)dx. \end{aligned}$$

Moreover, since $C^l_\beta = \prod_{j=1}^m C^{l_j}_{\beta_j}$ and $f(x) = \prod_{j=1}^m f_j(x_j)$, we obtain

$$\int_{C^l_\beta} \varphi_{\alpha_i,z_i}(x_i)f(x)dx = \left(\int_{C^{l_i}_{\beta_i}} \varphi_{\alpha_i,z_i}(x_i)f_i(x_i)dx_i\right) \prod_{j\neq i} \int_{C^{l_j}_{\beta_j}} f_j(x_j)dx_j.$$

Writing $l_{(i)} := (l_1, \ldots, l_{i-1}, l_{i+1}, \ldots, l_m)$, we replace $\sum_{l\in\mathbb{Z}^m}$ by $\sum_{l_i\in\mathbb{Z}} \sum_{l_{(i)}\in\mathbb{Z}^{m-1}}$ and get

$$G_{\alpha,z}(\lambda, f) = \sum_{i=1}^m \sum_{l_i\in\mathbb{Z}} \psi_{\alpha,z,\lambda,f}(i, l_i) \int_{C^{l_i}_{\beta_i}} \varphi_{\alpha_i,z_i}(x_i)f_i(x_i)dx_i \tag{3.35}$$

with

$$\psi_{\alpha,z,\lambda,f}(i, l_i) := \sum_{l_{(i)}\in\mathbb{Z}^{m-1}} \lambda_i(l+\beta) \prod_{j\neq i} \int_{C^{l_j}_{\beta_j}} f_j(x_j)dx_j. \tag{3.36}$$

Observe that $\psi_{\alpha,z,\lambda,f}(i,l_i) \geq 0$ for every $i = 1,\dots,m$, $l_i \in \mathbb{Z}$. Thus, if we adapt f such that the integrals in (3.35) and (3.36) do not decrease, then an upper bound for $G_{\alpha,z}(\lambda, f)$ is obtained. To this end, we construct the joint density function $g \in \mathcal{F}^m$ as follows. Let

$$g(x) := \prod_{i=1}^{m} g_i(x_i), \quad x \in \mathbb{R}^m,$$

where for every $i = 1,\dots,m$, the marginal density function g_i is a particular flattened version of f_i: the function f_i is only flattened over those intervals $C^{l_i}_{\beta_i}$ for which $\int_{C^{l_i}_{\beta_i}} \varphi_{\alpha_i,z_i}(u) f_i(u) du < 0$. That is, for every $l_i \in \mathbb{Z}$, and $x_i \in C^{l_i}_{\beta_i}$,

$$g_i(x_i) := \begin{cases} f_i(x_i), & \text{if } \int_{C^{l_i}_{\beta_i}} \varphi_{\alpha_i,z_i}(u) f_i(u) du \geq 0 \\ \int_{C^{l_i}_{\beta_i}} f_i(u) du & \text{otherwise.} \end{cases}$$

Obviously, because of Lemma 3.2, we have $|\Delta| g_i \leq |\Delta| f_i$ for every $i = 1,\dots,m$. In order to show that $G_{\alpha,z}(\lambda, f) \leq G_{\alpha,z}(\lambda, g) \leq G_{\alpha,z}(\lambda^*, g)$, notice that for every $l_i \in \mathbb{Z}$ and every $i = 1,\dots,m$,

(i) $\int_{C^{l_i}_{\beta_i}} g_i(u) du = \int_{C^{l_i}_{\beta_i}} f_i(u) du$,

(ii) $\int_{C^{l_i}_{\beta_i}} \varphi_{\alpha_i,z_i}(u) g_i(u) du \geq \int_{C^{l_i}_{\beta_i}} \varphi_{\alpha_i,z_i}(u) f_i(u) du$,

(iii) $\int_{C^{l_i}_{\beta_i}} \varphi_{\alpha_i,z_i}(u) g_i(u) du \geq 0$.

These properties follow directly from the construction. Indeed, if $g_i(x_i) = f_i(x_i)$ on $C^{l_i}_{\beta_i}$, nothing has to be shown. Otherwise, (i) is obvious, and

$$0 = \int_{C^{l_i}_{\beta_i}} \varphi_{\alpha_i,z_i}(u) g_i(u) du > \int_{C^{l_i}_{\beta_i}} \varphi_{\alpha_i,z_i}(u) f_i(u) du,$$

where the equality follows from Lemma 3.3 (iii) using $|C^{l_i}_{\beta_i}| = 1$ and Lemma 3.5 (iv).

From (i) it follows immediately that

$$\psi_{\alpha,z,\lambda,g}(i,l_i) = \psi_{\alpha,z,\lambda,f}(i,l_i), \quad l_i \in \mathbb{Z},\ i = 1,\dots,m,$$

which together with (ii) implies

$$G_{\alpha,z}(\lambda, f) \leq G_{\alpha,z}(\lambda, g).$$

In addition,

$$\begin{aligned} G_{\alpha,z}(\lambda, g) &= \sum_{i=1}^{m} \sum_{l\in\mathbb{Z}^m} \lambda_i(l+\beta) \int_{C^{l_i}_{\beta_i}} \varphi_{\alpha_i,z_i}(x_i) g_i(x_i) dx_i \prod_{j\neq i} \int_{C^{l_j}_{\beta_j}} g_j(x_j) dx_j \\ &\leq \sum_{i=1}^{m} \sum_{l\in\mathbb{Z}^m} \lambda_i^* \int_{C^{l_i}_{\beta_i}} \varphi_{\alpha_i,z_i}(x_i) g_i(x_i) dx_i \prod_{j\neq i} \int_{C^{l_j}_{\beta_j}} g_j(x_j) dx_j \\ &= G_{\alpha,z}(\lambda^*, g), \end{aligned}$$

where the inequality is true, since the coefficient of each $\lambda_i(l+\beta)$ is nonnegative because of (iii). □

Next we state an upper bound for the relaxed optimization problem $N(B)$ defined in (3.34).

Proposition 3.3. *For every $B \in \mathbb{R}^m$ with $B > 0$,*

$$N(B) \leq \sum_{i=1}^{m} \lambda_i^* h(B_i),$$

with N defined in (3.34), λ_i^* defined in (3.33), and h defined in (3.31).

Proof. Using Proposition 3.2, we have that

$$\begin{aligned} N(B) &= \sup_{\alpha\in\mathbb{R}^m} \sup_{z\in\mathbb{R}^m} \sup_{f\in\mathcal{F}^m} \sup_{\lambda\in\Lambda^m} \Big\{ G_{\alpha,z}(\lambda, f) : |\Delta| f_i \leq B_i,\ i = 1, \ldots, m \Big\} \\ &\leq \sup_{\alpha\in\mathbb{R}^m} \sup_{z\in\mathbb{R}^m} \sup_{f\in\mathcal{F}^m} \Big\{ G_{\alpha,z}(\lambda^*, f) : |\Delta| f_i \leq B_i,\ i = 1, \ldots, m \Big\}. \end{aligned}$$

Note that for every $\alpha \in \mathbb{R}^m$, $z \in \mathbb{R}^m$, and $f \in \mathcal{F}^m$ with $|\Delta|f_i = B_i$,

$$\begin{aligned} G_{\alpha,z}(\lambda^*, f) &= \mathbb{E}_\omega\Big[\lambda^*(\omega)\Big\{ \lceil \omega \rceil_z - \lceil \omega \rceil_\alpha \Big\}\Big] \\ &= \sum_{i=1}^m \lambda_i^* \mathbb{E}_{\omega_i}\Big[\Big\{ \lceil \omega_i \rceil_{z_i} - \lceil \omega_i \rceil_{\alpha_i} \Big\}\Big] \\ &= \sum_{i=1}^m \lambda_i^* D_{\alpha_i,z_i}(f_i) \\ &\le \sum_{i=1}^m \lambda_i^* \mathcal{M}(\alpha_i, f_i), \end{aligned}$$

where D_{α_i,z_i} is as defined in Definition 3.4 and $\mathcal{M}$ is as defined in (3.18). The result now follows from Theorem 3.1. □

We are now ready to state our main result on the independent case.

Theorem 3.2 (Error bound for TU integer recourse models with independent random variables)**.** *Consider the TU integer recourse function Q defined as*

$$Q(z) = \mathbb{E}_\omega \left[\min_y \left\{ qy : Wy \ge \omega - z,\, y \in \mathbb{Z}_+^{n_2} \right\} \right], \quad z \in \mathbb{R}^m,$$

and for every $\alpha \in \mathbb{R}^m$ its α-approximation Q_α defined as

$$Q_\alpha(z) = \mathbb{E}_\omega \left[\min_y \left\{ qy : Wy \ge \lceil \omega \rceil_\alpha - z,\, y \in \mathbb{R}_+^{n_2} \right\} \right], \quad z \in \mathbb{R}^m.$$

Under assumptions (i)–(iii) introduced in Section 3.2, we have for every $\alpha \in \mathbb{R}^m$ and every random right-hand side vector ω with independently distributed components *and with joint density function $f \in \mathcal{F}^m$ that*

$$\sup_{z \in \mathbb{R}^m} |Q(z) - Q_\alpha(z)| \le \sum_{i=1}^m \lambda_i^* h(|\Delta| f_i),$$

where λ_i^* is as defined in (3.33) and h is as defined in (3.31).

Proof. Let $\alpha \in \mathbb{R}^m$ and $f \in \mathcal{F}^m$ with $|\Delta|f_i = B_i$, $i = 1, \dots, m$, be given. Then,

$$\sup_{z \in \mathbb{R}^m} |Q(z) - Q_\alpha(z)| =: \mathcal{N}(\alpha, f) \le \mathcal{N}^*(f) \le N(B),$$

where the first inequality follows from Lemma 3.6 and the second from the definition of N in (3.34). Now the result follows directly from Proposition 3.3. □

Remark 3.5. In order to obtain λ^*, it suffices to solve m linear programming problems, since $\lambda_i^* = \max_\lambda \{\lambda_i : \lambda W \leq q,\ \lambda \in \mathbb{R}_+^m\}$.

The error bound in Theorem 3.2 is a function of both λ_i^* and $h(|\Delta|f_i)$ for $i = 1, \ldots, m$. The values of λ_i^* depend only on the parameters q and W of the second-stage value function v; in general, the higher the cost q, the higher the values of λ_i^* and thus the error bound. This need not be problematic in practice, since a larger error may be acceptable if the recourse costs are higher.

The values of $h(|\Delta|f_i)$, however, depend only on the total variations $|\Delta|f_i$ of the densities of the random variables in the model. In some sense the effects on the error bound of the randomness in the model and the parameters of the second-stage value function are thus separated. Recalling that the values of $h(|\Delta|f_i)$ are small for many practically relevant density functions f_i (see Table 3.1), we conclude that α-approximations perform well as soon as the total variations of the densities of all random variables in the model are small enough.

In Section 3.5.2 we compare $\|Q - Q_\alpha\|_\infty$ and its upper bound $\sum_{i=1}^m \lambda_i^* h(|\Delta|f_i)$ in a numerical study. Using several examples, we indicate how tight the upper bound actually is.

3.5.1 Tight bounds for simple integer recourse models

Interestingly, the generic one-dimensional SIR function $\mathcal{Q}$ defined in (3.6) is a special case of the TU integer recourse function Q of Theorem 3.2 with $m = 1$, $q = 1$, and $W = I_1$. Thus, Theorem 3.2 yields an upper bound for $\|\mathcal{Q} - \mathcal{Q}_\alpha\|_\infty$.

Corollary 3.2. *Consider the generic one-dimensional simple integer recourse function* $\mathcal{Q}(z) := \mathbb{E}_\omega[\lceil \omega - z \rceil^+]$, $z \in \mathbb{R}$, *and its* α*-approximation* $\mathcal{Q}_\alpha(z) := \mathbb{E}_\omega[(\lceil \omega \rceil_\alpha - z)^+]$, $z \in \mathbb{R}$. *Then, for every* $\alpha \in \mathbb{R}$ *and random variable* ω *with density function* $f \in \mathcal{F}$ *we have*

$$\|\mathcal{Q} - \mathcal{Q}_\alpha\|_\infty \leq h(|\Delta|f) = \begin{cases} |\Delta|f/8, & |\Delta|f \leq 4, \\ 1 - 2/|\Delta|f, & |\Delta|f \geq 4. \end{cases}$$

Proof. Apply Theorem 3.2, and observe that $\lambda^* = 1$. □

Comparing this error bound with that of Klein Haneveld et al. [42] given in (3.7), we observe that for $|\Delta|f \leq 4$ we improve this error bound by a factor 2. Moreover, for $|\Delta|f \geq 4$ the error bound in Corollary 3.2 increases hyperbolically to the trivial bound 1 as $|\Delta|f$ increases, whereas the old bound is equal to 1 for all $|\Delta|f \geq 4$.

We will show (for the m-dimensional case) that the error bound in Theorem 3.2 is tight for SIR models implying that the bound in Corollary 3.2 cannot be improved further.

Corollary 3.3. *Consider the m-dimensional SIR function*

$$Q(z) = \mathbb{E}_\omega\Big[\min_y \Big\{qy : y \geq \omega - z,\ y \in \mathbb{Z}_+^{n_2}\Big\}\Big], \quad z \in \mathbb{R}^m,$$

and let $B \in \mathbb{R}^m$ with $B > 0$ be given. Assume that $q \geq 0$ so that the recourse is sufficiently expensive. Then, for every $\alpha \in \mathbb{R}^m$ there exists $f \in \mathcal{F}^m$ such that $|\Delta|f_i = B_i$, $i = 1, \ldots, m$, and

$$\sup_{z \in \mathbb{R}^m} |Q(z) - Q_\alpha(z)| = \sum_{i=1}^m \lambda_i^* h(|\Delta|f_i).$$

Proof. For SIR models, the dual feasible region is given by $\{\lambda \in \mathbb{R}_+^{n_2} : \lambda \leq q\}$ so that $\lambda_i^* = q_i \geq 0$. Hence, by Theorem 3.2, the bound reads

$$\sup_{z \in \mathbb{R}^m} |Q(z) - Q_\alpha(z)| \leq \sum_{i=1}^m \lambda_i^* h(|\Delta|f_i) = \sum_{i=1}^m q_i h(|\Delta|f_i).$$

On the other hand, since for SIR models Q and Q_α are separable, see (3.5), we have

$$Q(z) - Q_\alpha(z) = \sum_{i=1}^m q_i \mathbb{E}_{f_i}\left[\lceil \omega_i - z_i \rceil^+ - \left(\lceil \omega_i \rceil_{\alpha_i} - z_i\right)^+\right], \quad z \in \mathbb{R}^m.$$

It is convenient to restrict our attention to pdf f_i and real numbers z_i such that f_i vanishes on $(-\infty, z_i]$. Then the '+' operations in the last formula are superfluous, so that

$$
\begin{aligned}
Q(z) - Q_\alpha(z) &= \sum_{i=1}^{m} q_i \mathbb{E}_{f_i}\Big[\lceil \omega_i - z_i \rceil - \lceil \omega_i \rceil_{\alpha_i} + z_i \Big] \\
&= \sum_{i=1}^{m} q_i D_{\alpha_i, z_i}(f_i),
\end{aligned}
$$

as mentioned in Remark 3.4. Consequently, in order to show that the bound of Theorem 3.2 is tight, it is sufficient to show that for all $i \in \{1, \ldots, m\}$, $\alpha_i \in \mathbb{R}$ and $B_i \in \mathbb{R}$ with $B_i > 0$ there exist $z_i \in \mathbb{R}$ and $f_i \in \mathcal{F}$ with $f_i(x_i) = 0$ for $x_i \leq z_i$ and $|\Delta| f_i = B_i$ such that

$$
D_{\alpha_i, z_i}(f_i) = h(B_i) = \begin{cases} B_i/8, & 0 < B_i \leq 4, \\ 1 - 2/B_i, & B_i \geq 4, \end{cases}
$$

which can be shown using the pdf $\hat{f}$ and $\bar{f}$ introduced in (3.29) and (3.30). Indeed, if $B_i \in (0, 4]$, then choose $z_i = \alpha_i - 1/2$, so that $\gamma_{\alpha_i, z_i} = 1/2$ and thus $\gamma_{\alpha_i, z_i} \geq 1 - 2/B_i$, and $f_i = \bar{f}$ with parameters $z := z_i$ and $\alpha := \alpha_i$. Then,

$$
D_{\alpha_i, z_i}(f_i) = \gamma_{\alpha_i, z_i}(1 - \gamma_{\alpha_i, z_i}) B_i/2 = B_i/8.
$$

If $B_i \geq 4$, then choose $z_i = \alpha_i - 2/B_i$, so that $\gamma_{\alpha_i, z_i} = 1 - 2/B_i$, and $f_i = \hat{f}$ with parameters $z := z_i$ and $\alpha := \alpha_i$. Then,

$$
D_{\alpha_i, z_i}(f_i) = \gamma_{\alpha_i, z_i} = 1 - 2/B_i. \qquad \square
$$

3.5.2 Numerical study of $\|Q - Q_\alpha\|_\infty$ and its upper bound

In this section we compare $\|Q - Q_\alpha\|_\infty$ and its upper bound in a numerical study. As already mentioned, and indeed the motivation of deriving the upper bound in Theorem 3.2, it is not possible to calculate $\|Q - Q_\alpha\|_\infty$ for large problem instances, so we restrict our attention to SIR models and a small TU integer recourse example.

Example 3.2. Consider the generic one-dimensional SIR function $\mathcal{Q}$ defined in (3.6), and let ω be a normally distributed random variable with mean μ and variance σ^2.

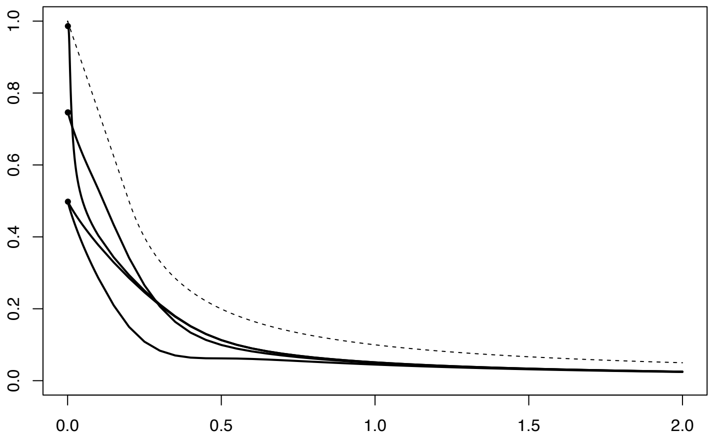

Figure 3.3: The supremum norm $\|\mathcal{Q} - \mathcal{Q}_\alpha\|_\infty$ and its upper bound $h(|\Delta|f)$ of Example 3.2 as a function of σ, the standard deviation of the random variable $\omega \sim N(0, \sigma^2)$. The dashed line corresponds to $h(|\Delta|f)$ and the solid lines to $\|\mathcal{Q} - \mathcal{Q}_\alpha\|_\infty$ for $\alpha = 0, 0.5, 0.75, 0.99$.

From Example 3.1 and Corollary 3.2 it follows that for all $\alpha \in \mathbb{R}$,

$$\|\mathcal{Q} - \mathcal{Q}_\alpha\|_\infty \leq h(|\Delta|f) = \begin{cases} 1 - \sigma\sqrt{2\pi}, & \sigma \leq \frac{1}{4}\sqrt{2/\pi}, \\ (8\sigma)^{-1}\sqrt{2/\pi}, & \sigma \geq \frac{1}{4}\sqrt{2/\pi}. \end{cases}$$

Notice that the upper bound $h(|\Delta|f)$ converges to the trivial upper bound 1 as $\sigma \to 0$. Moreover, $h(|\Delta|f)$ decreases linearly for $\sigma \leq \frac{1}{4}\sqrt{2/\pi}$ and hyperbolically for $\sigma \geq \frac{1}{4}\sqrt{2/\pi}$ with limit 0 as $\sigma \to +\infty$. This can also be observed in Figure 3.3, where both $\|\mathcal{Q} - \mathcal{Q}_\alpha\|_\infty$ and its upper bound $h(|\Delta|f)$ are given as a function of σ for various values of α; the mean μ equals 0 in all cases.

Clearly, the difference between $\mathcal{Q}$ and $\mathcal{Q}_\alpha$ decreases as the standard deviation σ increases (and thus the total variation $|\Delta|f$ decreases). Moreover, we observe that for larger values of σ, the value of $\|\mathcal{Q} - \mathcal{Q}_\alpha\|_\infty$ is approximately 50% of the upper bound $h(|\Delta|f)$ for all α. For smaller values of σ, i.e. as $\sigma \to 0$, the value of $\|\mathcal{Q} - \mathcal{Q}_\alpha\|_\infty$ may converge to any value between 1/2 and 1 depending on the value of α. For example, for $\alpha = 0$, it converges to $1/2$ and for $\alpha = 1 - \epsilon$ with $\epsilon > 0$ very small, its limit is $1 - \epsilon$, which converges to the trivial upper bound 1 as $\epsilon \to 0$. $\lhd$

It is not surprising that $\|\mathcal{Q} - \mathcal{Q}_\alpha\|_\infty$ is large for small values of σ, since in these cases the random variable ω is approximately a degenerate random variable with $\mathbb{P}\{\omega = \mu\} = 1$ so that $\mathcal{Q}(z) \approx v(\mu - z)$. This latter function is highly non-convex and even discontinuous with jumps of size 1 at $z \in \mu - \mathbb{Z}_+$, and thus for *any* convex approximation $\bar{\mathcal{Q}}$ (including α-approximations $\mathcal{Q}_\alpha$), we have $\lim_{\sigma \to 0} \|\mathcal{Q} - \bar{\mathcal{Q}}\|_\infty \geq 1/2$.

This result illustrates the counterintuitive nature of the α-approximations: in case ω resembles a discrete random variable, which corresponds to a large total variation $|\Delta|f$ of the pdf f, then α-approximations perform badly, and if $|\Delta|f$ is small, which is, for example, the case if ω resembles a uniform random variable with a large support, then α-approximations perform well. This contrasts strongly with most approximations in the stochastic programming literature for which typically the quality of the solutions is better if the random vector ω is discretely distributed with only a small number of scenarios.

Example 3.3. Again, consider the generic one-dimensional SIR function $\mathcal{Q}$, but now assume that ω is uniformly distributed on $[0, b]$ with $b > 0$. Observing that $|\Delta|f = 2/b$, it follows that

$$\|\mathcal{Q} - \mathcal{Q}_\alpha\|_\infty \leq h(|\Delta|f) = \begin{cases} 1 - b, & b \leq 1/2, \\ (4b)^{-1}, & b \geq 1/2. \end{cases}$$

Similar to Example 3.2, the upper bound $h(|\Delta|f)$ converges to the trivial upper bound 1 as $b \to 0$, $h(|\Delta|f)$ decreases linearly in b for $b \leq 1/2$, and $h(|\Delta|f)$ decreases hyperbolically in b for $b \geq 1/2$ with limit 0 as $b \to +\infty$. Interestingly, $\|\mathcal{Q} - \mathcal{Q}_\alpha\|_\infty$ does not decrease monotonically in b, as can be observed in Figure 3.4, where both $\|\mathcal{Q} - \mathcal{Q}_\alpha\|_\infty$ and its upper bound $h(|\Delta|f)$ are given as a function of b for $\alpha = 0$. In fact, for every $b \in \mathbb{Z}$ with $b \geq 1$, we have $\|\mathcal{Q} - \mathcal{Q}_\alpha\|_\infty = 0$ since in these cases $\mathcal{Q}$ and $\mathcal{Q}_\alpha$ coincide, whereas $\|\mathcal{Q} - \mathcal{Q}_\alpha\|_\infty = h(|\Delta|f)$ for $b \in (0, 1/2]$ and $b \in 1/2 + \mathbb{Z}_+$. Thus, for some values of b the recourse function $\mathcal{Q}$ is convex and can be approximated exactly by $\mathcal{Q}_\alpha$ and for other values of b the worst-case bound $h(|\Delta|f)$ is actually sharp. $\lhd$

Based on these examples, we conclude that for SIR models the upper bound $h(|\Delta|f)$ of $\|\mathcal{Q} - \mathcal{Q}_\alpha\|_\infty$ is reasonably tight for several well-known distributions of ω, especially taking into account that the bound holds for all $f \in \mathcal{F}$.

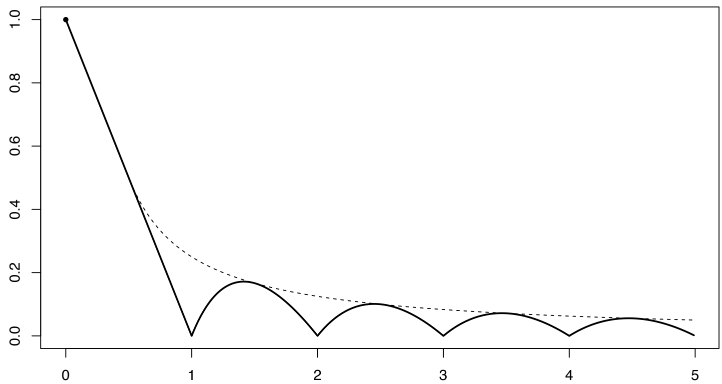

Figure 3.4: The supremum norm $\|\mathcal{Q} - \mathcal{Q}_\alpha\|_\infty$ and its upper bound $h(|\Delta|f)$ of Example 3.3 as a function of b, the right endpoint of the support of ω which is uniformly distributed on $[0, b]$. The dashed line corresponds to $h(|\Delta|f)$ and the solid line to $\|\mathcal{Q} - \mathcal{Q}_\alpha\|_\infty$ with $\alpha = 0$.

Next, we discuss a more general TU integer recourse example.

Example 3.4. Consider a TU integer recourse model with $m = 2$, $q = (3, 2, 2)$ and

$$W = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix},$$

and assume that ω is normally distributed with mean $\mu = (0, 0)$ and covariance matrix $V = \sigma^2 I_2$. The dual feasible region $\Lambda := \{\lambda \in \mathbb{R}^2_+ : \lambda W \leq q\}$ is given by

$$\Lambda = \{\lambda \in \mathbb{R}^2_+ : \lambda_1 + \lambda_2 \leq 3,\ \lambda_1 \leq 2,\ \lambda_2 \leq 2\}.$$

Straightforward computation shows that $\lambda_1^* = \lambda_2^* = 2$, and thus combining Theorem 3.2 and the expression for $|\Delta|f_i$ in Example 3.1,

$$\|Q - Q_\alpha\|_\infty \leq \sum_{i=1}^{2} \lambda_i^* h(|\Delta|f_i) = 4h(\sigma^{-1}\sqrt{2/\pi}).$$

Figure 3.5 shows $\|Q - Q_\alpha\|_\infty$ with $\alpha = (0, 0)$ and its upper bound as functions of

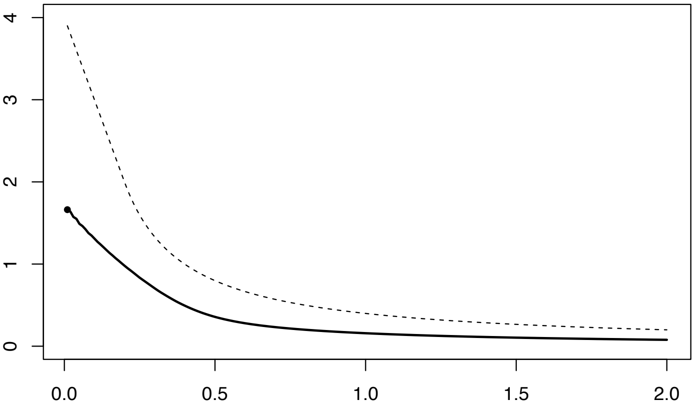

Figure 3.5: The supremum norm $\|Q - Q_\alpha\|_\infty$ with $\alpha = (0, 0)$ (solid line) and its upper bound (dashed line) of Example 3.4 as a function of σ, with ω normally distributed with mean $\mu = (0, 0)$ and covariance matrix $V = \sigma^2 I_2$.

σ. The results are very similar to those in Example 3.2. However, in this case the values of $\|Q - Q_\alpha\|_\infty$ are 40% of its upper bound instead of 50%. The increased gap can be attributed to λ^*, which is obtained as the componentwise maximum of the dual vertices $\lambda^k \in \Lambda$. In fact, in the current TU example it holds that $\lambda^* \notin \Lambda$, contrary to the previous SIR examples. We conclude that the quality of the error bound in Theorem 3.2 depends on q and W but that generally the bound appears to be reasonably tight. $\lhd$

3.6 TU integer recourse models with dependent random right-hand side parameters

In this section we again assume that ω is continuously distributed, but now we assume that the joint density function f is contained in a larger set $\mathcal{H}$, allowing for dependency.

Definition 3.8. Let $\mathcal{H}$ denote the set of m-dimensional joint density functions f

whose conditional density functions $f_i(\cdot|x_{(i)})$ defined as

$$f_i(x_i|x_{(i)}) = f(x)/f_{(i)}(x_{(i)})$$

are contained in $\mathcal{F}$ for all $i = 1, \ldots, m$, and $x_{(i)} \in \mathbb{R}^{m-1}$. (As before, we use the notation $x_{(i)}$ for the vector x without its i-th component.)

Of course, this definition only makes sense for those i and $x_{(i)}$ for which $f_{(i)}(x_{(i)}) > 0$. If $f_{(i)}(x_{(i)}) = 0$, any definition of $f_i(x_i|x_{(i)})$ is good but irrelevant, since in calculating expectations via conditioning its contribution is multiplied by $f_{(i)}(x_{(i)})$, that is, by 0.

Using the results from the previous sections, we are able to derive an error bound in this case as well.

Theorem 3.3 (Error bound for TU integer recourse models)**.** *Consider the TU integer recourse function Q defined as*

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

and for every $\alpha \in \mathbb{R}^m$ its α-approximation Q_α defined as

$$Q_\alpha(z) = \mathbb{E}_\omega \left[\min_y \left\{ qy : Wy \geq \lceil \omega \rceil_\alpha - z,\, y \in \mathbb{R}_+^{n_2} \right\} \right], \quad z \in \mathbb{R}^m.$$

Under assumptions (i)–(iii) introduced in Section 3.2, we have for every $\alpha \in \mathbb{R}^m$ and every random right-hand side vector ω with joint density function $f \in \mathcal{H}$ that

$$\sup_{z \in \mathbb{R}^m} |Q(z) - Q_\alpha(z)| \leq \sum_{i=1}^m \lambda_i^* \mathbb{E}_{\omega_{(i)}} \left[h\left(|\Delta| f_i(\cdot|\omega_{(i)}) \right) \right],$$

where λ_i^* is as defined in (3.33) and h is as defined in (3.31).

Proof. We follow the line of proof of the previous section, using the same notation. Obviously, Lemma 3.6 also holds for $f \in \mathcal{H}$ so that

$$\sup_{z \in \mathbb{R}^m} |Q(z) - Q_\alpha(z)| \leq \sup_{\alpha \in \mathbb{R}^m} \sup_{z \in \mathbb{R}^m} \sup_{\lambda \in \Lambda^m} G_{\alpha,z}(\lambda, f),$$

and similar to the proof of Proposition 3.2, we have

$$\begin{aligned} G_{\alpha,z}(\lambda, f) &= \mathbb{E}_f\Big[\lambda(\omega)\Big(\lceil\omega\rceil_z - \lceil\omega\rceil_\alpha\Big)\Big] \\ &= \sum_{i=1}^{m} \int_{\mathbb{R}^m} \lambda_i(x)\varphi_{\alpha_i,z_i}(x_i)f(x)dx. \end{aligned}$$

However, now we apply conditioning using $f(x) = f_i(x_i|x_{(i)})f_{(i)}(x_{(i)})$ to obtain

$$\begin{aligned} G_{\alpha,z}(\lambda, f) &= \sum_{i=1}^{m} \int_{\mathbb{R}^{m-1}} \Big\{ \int_{\mathbb{R}} \lambda_i(x)\varphi_{\alpha_i,z_i}(x_i)f_i(x_i|x_{(i)})dx_i \Big\} f_{(i)}(x_{(i)})dx_{(i)} \\ &= \sum_{i=1}^{m} \int_{\mathbb{R}^{m-1}} G^1_{\alpha_i,z_i}\Big(\hat{\lambda}_i(\cdot|x_{(i)}), f_i(\cdot|x_{(i)})\Big) f_{(i)}(x_{(i)})dx_{(i)}, \end{aligned}$$

where $G^1_{\alpha_i,z_i}$ denotes the case $m = 1$ in the definition of $G_{\alpha,z}$ and $\hat{\lambda}_i(\cdot|x_{(i)}) : \mathbb{R} \mapsto \mathbb{R}$ is defined as $\hat{\lambda}_i(x_i|x_{(i)}) = \lambda_i(x)$. Since this function $\hat{\lambda}_i(\cdot|x_{(i)}) \in \Lambda^1$ for all $x_{(i)} \in \mathbb{R}^{m-1}$, we can apply Proposition 3.3 with $m = 1$, $\alpha = \alpha_i$, $z = z_i$, $\lambda = \hat{\lambda}_i(\cdot|x_{(i)})$, and $f = f_i(\cdot|x_{(i)})$, yielding

$$\begin{aligned} G_{\alpha,z}(\lambda, f) &\le \sum_{i=1}^{m} \int_{\mathbb{R}^{m-1}} \lambda_i^* h\left(|\Delta|f_i(\cdot|x_{(i)})\right) f_{(i)}(x_{(i)})dx_{(i)} \\ &= \sum_{i=1}^{m} \lambda_i^* \mathbb{E}_{\omega_{(i)}}\Big[h\left(|\Delta|f_i(\cdot|\omega_{(i)})\right)\Big]. \qquad \square \end{aligned}$$

Theorem 3.3 generalizes Theorem 3.2 since $\mathcal{F}^m \subset \mathcal{H}$. If $f \in \mathcal{F}^m$, then the conditional density $f_i(x_i|x_{(i)}) = f_i(x_i)$ for all $x \in \mathbb{R}^m$, and thus the error bound in Theorem 3.3 reduces to that in Theorem 3.2.

The following example illustrates the increase in value of the error bound by introducing dependency.

Example 3.5. Let $f \in \mathcal{H}$ be the joint density function of a bivariate normal random vector ω with correlation coefficient ρ. It is well known that $\omega_1|\omega_2 = x_2$ follows a normal distribution with variance $(1-\rho^2)\sigma_1^2$. Hence, using Example 3.1, for $i = 1, 2$, and $x_{(i)} \in \mathbb{R}$,

$$|\Delta|f_i(\cdot|x_{(i)}) = \frac{1}{\sqrt{1-\rho^2}}\sigma_i^{-1}\sqrt{2/\pi}.$$

This implies that the error bound in Theorem 3.3 for this particular joint density function equals

$$\sum_{i=1}^{2} \lambda_i^* \mathbb{E}_{\omega_{(i)}} \left[h\left(|\Delta| f_i(\cdot|\omega_{(i)}) \right) \right] = \sum_{i=1}^{2} \lambda_i^* h \left(\frac{1}{\sqrt{1-\rho^2}} \sigma_i^{-1} \sqrt{2/\pi} \right).$$

Compared to the independent case, the total variations increase by a factor $1/\sqrt{1-\rho^2}$ with equivalence if $\rho = 0$ (see Example 3.4 on the independent case). For example, if $|\rho| \leq 0.4$, then this factor is smaller than 1.1, and thus the total variations in the dependent case are less than 10% higher than in the independent case. We conclude that only for high correlation values $|\rho|$ the error bound in the dependent case increases substantially compared to the independent case. This is also confirmed by numerical experiments similar to those in Example 3.4. $\triangleleft$

3.7 Complete integer recourse models

If the recourse matrix W is not TU but a general integer-valued matrix, then the error bounds for the α-approximation Q_α in Theorem 3.2 and 3.3 are no longer valid. This is because the equality in (3.9) now holds with inequality, implying that

$$Q(z) \geq \mathbb{E}_\omega \Big[\max_{k=1,\ldots,K} \lambda^k \lceil \omega - z \rceil \Big], \quad z \in \mathbb{R}^m. \tag{3.38}$$

Nonetheless, the α-approximation Q_α may be useful as an approximate lower bound for Q, to be used in several special-purpose algorithms; see Van der Vlerk [83]. In fact, if the random variables in the model are independently and uniformly distributed, then there exists an $\alpha^* \in \mathbb{R}^m$ such that the α^*-approximation Q_{α^*} is the convex hull of $\mathbb{E}_\omega[\max_{k=1,\ldots,K} \lambda^k \lceil \omega - z \rceil]$ (see Chapter 2) and thus a lower bound of Q. In all other cases, Q_α is not necessarily a lower bound for Q, but a one-sided error bound is available.

Corollary 3.4. *Consider the* complete integer *recourse function Q defined as*

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

and for every $\alpha \in \mathbb{R}^m$ its α-approximation Q_α defined as

$$Q_\alpha(z) = \mathbb{E}_\omega \left[\min_y \left\{ qy : Wy \geq \lceil \omega \rceil_\alpha - z,\, y \in \mathbb{R}_+^{n_2} \right\} \right], \quad z \in \mathbb{R}^m.$$

Under assumptions (i)–(iii) introduced in Section 3.2, we have for every $\alpha \in \mathbb{R}^m$, $z \in \mathbb{R}^m$, and every random right-hand side vector ω with joint density function $f \in \mathcal{H}$ that

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

where λ_i^* is as defined in (3.33) and h is as defined in (3.31).

Proof. Combine (3.38) and Theorem 3.3. □

Both $Q_\alpha - \sum_{i=1}^m \lambda_i^* \mathbb{E}_{\omega_{(i)}} \left[h\left(|\Delta| f_i(\cdot|\omega_{(i)}) \right) \right]$ and Q_α can be used as exact lower bound and approximate lower bound, respectively. Using the latter yields approximate solutions with only a small error if the total variations of the densities of the random variables in the model are small enough, whereas using the first yields exact solutions but is computationally more demanding since the lower bound is weaker.

Both approaches are of interest, the latter in particular since for every $\alpha \in \mathbb{R}^m$, we have $Q_\alpha \geq Q^{LP}$, where Q^{LP} is the recourse function obtained by using the LP relaxation of the second-stage integer program defining the value function v [83]. Moreover, if $Q_\alpha(z) > 0$, then $Q_\alpha(z) > Q^{LP}(z)$; see [86].

As already observed in [83], this implies that Q_α is a *strictly better* (approximate) lower bound than Q^{LP}, and, moreover, Q_α is *computationally more tractable* than Q^{LP} since Q_α corresponds to a continuous recourse model with discrete random variables, whereas Q^{LP} corresponds to a continuous recourse model with continuous random variables.

3.8 Summary and conclusions

We consider a class of convex approximations for totally unimodular (TU) integer recourse models. Using piecewise flattening of density functions, we derive a uniform error bound for these approximations that depends on the total variations of the probability density functions involved. For simple integer recourse models this error bound is tight and improves the existing one by a factor 2. Moreover, for TU integer recourse models this is the first nontrivial error bound available. Due to this error bound the convex approximations can also be used as an approximate lower bound for complete integer recourse models.

As illustrated by several numerical examples, we show that the approximations are good if all total variations of the probability density functions of the random variables in the model are small enough. For example, for normally distributed random variables ω this implies that the convex approximations are good if the standard deviations σ are large and the approximations are bad if the σ are small. This result contrasts strongly with other approximations in the literature, where typically approximations perform better for small values of σ, i.e. if ω can be better approximated by a discrete random vector.

A future research direction is to apply the idea of modifying the recourse data to pure integer and mixed-integer recourse models. Alternatively, for the convex approximations in this chapter, an error bound may be obtained that depends on characteristics of the joint pdf f and not only on its one-dimensional conditional densities.