Convex approximations for complete integer recourse models

Maarten H. van der Vlerk*

3rd April 2002

Abstract

We consider convex approximations of the expected value function of a two-stage integer recourse problem. The convex approximations are obtained by perturbing the distribution of the random right-hand side vector. It is shown that the approximation is optimal for the class of problems with totally unimodular recourse matrices. For problems not in this class, the result is a convex lower bound that is strictly better than the one obtained from the LP relaxation.

1

1.   Introduction

We consider the two-stage integer recourse problem

$$\min_{x} \quad cx + \mathcal{Q}(x)$$
$$\text{s.t.} \quad Ax \leq b$$
$$\quad x \in \mathbb{R}_+^{n_1},$$

where

$$\mathcal{Q}(x) := \mathbb{E}_\omega \left[v(\omega - Tx) \right], \quad x \in \mathbb{R}^{n_1},$$

and, for $s \in \mathbb{R}^m$,

$$v(s) := \min_{y} \quad qy \qquad\qquad (1)$$
$$\text{s.t.} \quad Wy \geq s$$
$$\quad y \in \mathbb{Z}_+^{n_2}.$$

The function v is the second-stage value function, and the function $\mathcal{Q}$ is called the expected value function. These functions model the (expected) costs of recourse actions to compensate for infeasibilities associated with the random goal constraints $Tx \geq \omega$. The right-hand side vector ω is a random vector with known cumulative distribution function (cdf) F_ω.

In addition to difficulties met in the continuous recourse case (see e.g. [2, 4, 14]), this problem is hard to solve because – due to integrality of the second-stage variables – the expected value function $\mathcal{Q}$ is non-convex in general, see e.g. [8]. For the special case of simple integer recourse models, associated with $W = I$ in (1), we gave a complete description of the class of (continuous) distributions of ω that result in convexity of the function $\mathcal{Q}$, see [7], allowing the construction of convex approximations of $\mathcal{Q}$ by approximation of the distribution of ω, see [6]. In this paper we extend these results to obtain convex approximations for the general complete recourse case (see e.g. [15]), but first we concentrate on a special case.

2.   Assumptions

We will derive our results for problems with *totally unimodular* (TU) recourse matrix W. In that case, the extreme points of the polyhedral set $\{y \in \mathbb{R}^{n_2} : Wy \geq h\}$ are integral for any *integer* right-hand side h (see e.g. [12]), so that solving an LP problem

over this set results in an integer optimal solution. However, note that in our recourse problem the right-hand side is $\omega - Tx$, which is not an integer vector in general.

A TU matrix has integer components necessarily (in fact, all components are either -1, 0, or 1). In Section 6, we drop the TU assumption on W but still assume that W is integer (or rational, so that integrality of W can be obtained by scaling). We will show that in this case our approach leads to a convex lower bound for $\mathcal{Q}$ that is strictly better than the one obtained from the LP relaxation of (1).

Throughout, we assume that the recourse structure is

   (i) *complete*, i.e., $v < +\infty$, and
   (ii) *sufficiently expensive*, i.e., $v > -\infty$,

so that v is finite for all $s \in \mathbb{R}^m$. Together with the assumption that $\mathbb{E}_\omega\left[|\omega|\right]$ is finite, this implies that $\mathcal{Q}$ is finite for all $x \in \mathbb{R}^{n_1}$.

REMARK.   Actually, it is sufficient to require *relatively complete* recourse, so that v is finite for all relevant $s \in \mathbb{R}^m$.

3.   The value function

Assuming that W is an $m \times n_2$ TU matrix, and given that v is finite, we have for all $s \in \mathbb{R}^m$

$$
\begin{aligned}
v(s) \quad := \quad & \min_y \quad qy \\
& \text{s.t.} \quad Wy \geq s, \ y \in \mathbb{Z}_+^{n_2} \\
= \quad & \min_y \quad qy \\
& \text{s.t.} \quad Wy \geq \lceil s \rceil, \ y \in \mathbb{R}_+^{n_2} \\
= \quad & \max_\lambda \quad \lambda \lceil s \rceil \\
& \text{s.t.} \quad \lambda W \leq q, \ \lambda \in \mathbb{R}_+^m,
\end{aligned}
\tag{2, 3}
$$

where $\lceil \cdot \rceil$ denotes component-wise round up. Equality (2) holds because Wy is integral for all $y \in \mathbb{Z}^{n_2}$, so that no feasible solutions are cut off if we replace the right-hand side s by its round up; subsequently, we may relax the integrality of y because W is a TU matrix. The equality (3) follows by (strong) LP duality.

Since the recourse structure is complete and sufficiently expensive, it follows that the

3

dual feasible region $\{\lambda \in \mathbb{R}_+^m : \lambda W \le q\}$ is a non-empty polyhedron. Hence,

$$v(s) = \max_{k=1,\dots,K} \lambda^k \lceil s \rceil, \quad s \in \mathbb{R}^m, \tag{4}$$

where λ^k, $k = 1, \dots, K$, are the finitely many extreme points of the dual feasible set.

Thus, v is the maximum of finitely many weighted round up functions, and hence non-convex. However, we are not interested in convexity of v per se, but in convexity of its expectation $\mathcal{Q}(x) := \mathbb{E}_\omega [v(\omega - Tx)]$, $x \in \mathbb{R}^{n_1}$. Properties of the one-dimensional expected round up function follow from results obtained in the context of simple integer recourse (SIR) models [10, 18, 7]. In particular, in [7] a complete description is given of a class of probability density functions (pdf) for $\omega \in \mathbb{R}$, such that the SIR expected value function is convex if and only if ω has a pdf from this class. In the next section, we summarize and adapt these results for the current situation.

4.   Properties of the expected round up function

For any fixed $\lambda \in \mathbb{R}_+^m$, we define the m-dimensional round up function

$$R(z) := \lambda \mathbb{E}_\omega [\lceil \omega - z \rceil], \quad z \in \mathbb{R}^m.$$

It is easy to see that $R(z+k) = R(z) - \lambda k$, $k \in \mathbb{Z}^m$, $z \in \mathbb{R}^m$, so that R is linear on every grid $\alpha + \mathbb{Z}^m$ with $\alpha = (\alpha_1, \dots, \alpha_m) \in [0, 1)^m$. It follows that any reasonable convex approximation of R is an affine function with gradient $-\lambda$. Moreover, if such a convex approximation is equal to $R(\bar{z})$ for some $\bar{z}$, then the same is true for all $z \in \bar{z} + \mathbb{Z}^m$.

In this section we will show how to construct convex approximations of R using certain approximations of the distribution of ω. To this end, we first consider the one-dimensional round up function, for which results follow from those for simple integer recourse models.

In the literature on simple integer recourse models [10, 18, 19], results are presented in terms of the one-dimensional functions $G(z) := \mathbb{E}_\xi \left[(\lceil \xi - z \rceil)^+ \right]$ and $H(z) := \mathbb{E}_\xi \left[(\lfloor \xi - z \rfloor)^- \right]$, $z \in \mathbb{R}$, where $(s)^+ := \max\{0, s\}$ and $(s)^- := \max\{0, -s\}$ are the positive and negative part of $s \in \mathbb{R}$, respectively, and $\lfloor s \rfloor$ is the round down of s.

Using straightforward computation, we find that the expected round up function satisfies

$$\mathbb{E}_\xi [\lceil \xi - z \rceil] = G(z) - H(z-1) - \sum_{k=1}^{\infty} \Pr\{\xi = z - k\}, \quad z \in \mathbb{R}. \tag{5}$$

4

Note that the last term vanishes if ξ follows a continuous distribution.

This leads to the following result for the one-dimensional expected round up function.

Lemma 4.1 *Let $\xi \in \mathbb{R}$ be a continuous random variable with finite mean value and pdf f_ξ that is constant on every interval $(\alpha + k - 1, \alpha + k]$, $k \in \mathbb{Z}$, for an arbitrary but fixed $\alpha \in [0, 1)$. Then*

$$\mathbb{E}_\xi \left[\lceil \xi - z \rceil \right] = \mathbb{E}_{\varphi_\alpha} \left[\varphi_\alpha - z \right] = \mu_\alpha - z, \quad z \in \mathbb{R},$$

where $\varphi_\alpha = \lceil \xi - \alpha \rceil + \alpha$ is a discrete random variable with mean value μ_α and support in $\alpha + \mathbb{Z}$, with

$$
\begin{aligned}
\Pr\{\varphi_\alpha = \alpha + k\} &= F_\xi(\alpha + k) - F_\xi(\alpha + k - 1) \\
&= \Pr\{\xi \in (\alpha + k - 1, \alpha + k]\}, \quad k \in \mathbb{Z},
\end{aligned}
$$

where F_ξ is the cdf of ξ.
Hence, in this case $\mathbb{E}_\xi \left[\lceil \xi - z \rceil \right]$, $z \in \mathbb{R}$, is an affine function with slope -1.

PROOF. To prove the first claim we use Corollary 5.2 in [6], which states that $G(z) = \mathbb{E}_{\varphi_\alpha} \left[(\varphi_\alpha - z)^+ \right]$ and $H(z - 1) = \mathbb{E}_{\varphi_\alpha} \left[(\varphi_\alpha - z)^- \right]$, $z \in \mathbb{R}$. The result now follows from (5), since $(s)^+ - (s)^- = s$, $s \in \mathbb{R}$.

The second claim follows trivially from the first one. $\square$

Next we generalize this result to m-dimensional round up functions.

Lemma 4.2 *Let $\omega \in \mathbb{R}^m$ be a continuous random vector with joint pdf f_ω that is constant on every hypercube $C_\alpha^k := \prod_{i=1}^m (\alpha_i + k_i - 1, \alpha_i + k_i]$, $k \in \mathbb{Z}^m$, for an arbitrary but fixed $\alpha = (\alpha_1, \dots, \alpha_m) \in [0, 1)^m$. Then*

$$\mathbb{E}_\omega \left[\lceil \omega - z \rceil \right] = \mathbb{E}_{\varphi_\alpha} \left[\varphi_\alpha - z \right] = \mu_\alpha - z, \quad z \in \mathbb{R}^m,$$

where $\varphi_\alpha = \lceil \omega - \alpha \rceil + \alpha$ is a discrete random vector with mean value μ_α and support in $\alpha + \mathbb{Z}^m$, with

$$\Pr\{\varphi_\alpha = \alpha + k\} = \Pr\{\omega \in C_\alpha^k\}, \quad k \in \mathbb{Z}^m.$$

Hence, in this case the round up function $R(z) = \lambda \mathbb{E}_\omega \left[\lceil \omega - z \rceil \right]$, $z \in \mathbb{R}^m$, is affine with gradient $-\lambda$.

PROOF.   We use that

$$\mathbb{E}_\omega\left[\lceil \omega - z \rceil\right] = \sum_{k \in \mathbb{Z}^m} \Pr\{\omega \in C_\alpha^k\} \mathbb{E}_\omega\left[\lceil \omega - z \rceil \mid \omega \in C_\alpha^k\right], \quad z \in \mathbb{R}^m. \tag{6}$$

For each fixed $k \in \mathbb{Z}^m$, $\Pr\{\omega \in C_\alpha^k\}$ is either zero or the conditional distribution of ω given $\omega \in C_\alpha^k$ is uniform on C_α^k. In that case, the components of the vector ω are independent random variables on C_α^k, with each ω_i uniformly distributed on $(\alpha_i + k_i - 1, \alpha_i + k_i]$, $i = 1, \ldots, m$, so that

$$\mathbb{E}_\omega\left[\lceil \omega - z \rceil \mid \omega \in C_\alpha^k\right] = \alpha + k - z, \quad z \in \mathbb{R}^m, \tag{7}$$

by Lemma 4.1. Substitution of (7) in (6) proves the first claim.

The second claim follows trivially from the first one.   □

As in the simple integer recourse case, the main use of Lemma 4.2 is that it allows to construct *convex approximations* of the generally non-convex expected round up function by *modifying the distribution* of the right-hand side vector ω.

Definition 4.1   *Let $\omega \in \mathbb{R}^m$ be a random vector with arbitrary continuous or discrete distribution, and choose $\alpha = (\alpha_1, \ldots, \alpha_m) \in [0, 1)^m$. Define the α-approximation ω_α as the random vector with joint pdf f_α that is constant on every hypercube $C_\alpha^k := \prod_{i=1}^m (\alpha_i + k_i - 1, \alpha_i + k_i]$, $k \in \mathbb{Z}^m$, such that $\Pr\{\omega_\alpha \in C_\alpha^k\} = \Pr\{\omega \in C_\alpha^k\}$, $k \in \mathbb{Z}^m$.*

That is, for each $k \in \mathbb{Z}^m$, the probability mass assigned to C_α^k is the same under ω and ω_α; the conditional distribution of ω_α is uniform on C_α^k.

Accordingly, we will call the function R_α, defined for each $\alpha \in [0, 1)^m$,

$$R_\alpha(z) := \lambda \mathbb{E}_{\omega_\alpha}\left[\lceil \omega_\alpha - z \rceil\right], \quad z \in \mathbb{R}^m,$$

the *α-approximation of $R(z) = \lambda \mathbb{E}_\omega\left[\lceil \omega - z \rceil\right]$*. In general, an α-approximation is neither a lower bound nor an upper bound.

However, since $R(z + k) = R(z) - \lambda k$, $k \in \mathbb{Z}^m$, for every z, we see that $R(z) + \lambda z$ is a periodic function, which repeats itself on every set C_α^k. Thus, defining

$$\alpha^\star \in \operatorname{argmin}\left\{R(z) + \lambda z : z \in [0, 1)^m\right\}, \tag{8}$$

$R_{\alpha^\star}$ is a lower bound for R, which is sharp at every $z \in \alpha^\star + \mathbb{Z}^m$. Moreover, since $R_{\alpha^\star}$ is affine by Lemma 4.2, it is the pointwise largest convex lower bound of R. In other words, $R_{\alpha^\star}$ is the convex hull of R.

6

Lemma 4.3 *Consider the convex hull $R_{\alpha^\star}$ of the function R. The components $\alpha_i^\star$, $i = 1, \ldots, m$, of the parameter vector $\alpha^\star$ can be determined analytically in the following cases.*

  (i) *If the marginal distribution of ω_i is continuous with pdf f_i (assumed to be of bounded variation), then $\alpha_i^\star$ is one of the solutions of*

$$-1 \in \left[-\sum_{k=-\infty}^{\infty} f_i^-(z_i + k), -\sum_{k=-\infty}^{\infty} f_i^+(z_i + k) \right], \quad z_i \in [0, 1),$$

  with f_i^- and f_i^+ the left and right continuous version of the pdf f_i.

  (ii) *Let the marginal distribution of ω_i be discrete with finitely many different fractional values in its support. That is, the support is contained in $\bigcup_{s=1}^{S_i} \{\bar{\omega}_i^{\,s} + \mathbb{Z}\}$, where $0 \le \bar{\omega}_i^{\,1} < \bar{\omega}_i^{\,2} < \ldots < \bar{\omega}_i^{\,S_i} < 1$. Then $\alpha_i^\star$ can be chosen equal to any $\bar{\omega}_i^{\,k}$ such that*

$$k \in \operatorname*{argmin}_{n=1,\ldots,S_i} \left\{ \sum_{s=n+1}^{S_i} p_i^s + \bar{\omega}_i^{\,n} \right\},$$

  where $p_i^s := \Pr\{\omega_i \in \bar{\omega}_i^{\,s} + \mathbb{Z}\}$, $s = 1, \ldots, S_i$.

PROOF.   According to (8), $\alpha^\star$ is a solution of $\min\{R(z) + \lambda z : z \in [0, 1)^m\}$, where $\lambda \ge 0$ is fixed. Since

$$\min_{z \in [0,1)^m} \left\{ R(z) + \lambda z \right\} = \min_{z_i \in [0,1) \forall i} \sum_{i=1}^{m} \lambda_i \left(R_i(z_i) + z_i \right)$$

$$= \sum_{i=1}^{m} \lambda_i \min_{z_i \in [0,1)} \left\{ R_i(z_i) + z_i \right\},$$

we see that

$$\alpha_i^\star \in \operatorname{argmin}\left\{ R_i(z_i) + z_i : z_i \in [0, 1) \right\}, \tag{9}$$

so that the components of $\alpha^\star$ can be determined independently.

  (i) If the marginal distribution of ω_i is continuous, we can resort to results known for simple integer recourse models. Using (5), it follows from the differentiability properties of the functions G and H, discussed in [18], that the left derivative R_i^- and right derivative R_i^+ of the function R_i exist everywhere. Thus, being a global minimizer of (9), $\alpha_i^\star$ satisfies $-1 \in \left[R_i^-(\alpha_i^\star), R_i^+(\alpha_i^\star) \right]$, with

$$R_i^-(z_i) = -\sum_{k=-\infty}^{\infty} f_i^-(z_i + k) \quad \text{and} \quad R_i^+(z_i) = -\sum_{k=-\infty}^{\infty} f_i^+(z_i + k),$$

where f_i^- and f_i^+ denote the left and right continuous versions of the pdf f_i of ω_i, respectively.

(ii) If the marginal distribution of ω_i is discrete, then the function R_i is lower semi-continuous with a discontinuity at every point z_i such that $\Pr\{\omega_i \in z_i + \mathbb{Z}\} > 0$. Moreover, R_i is constant in between discontinuity points. Hence, if the discrete distribution satisfies the conditions, the function $R_i(z_i) + z_i$ attains its minimum value on $[0, 1)$ in one of the fractional values $\bar{\omega}_i^s$, $s = 1, \ldots, S_i$.

Since $\lceil t \rceil = \lceil t - l \rceil + l, l \in \mathbb{Z}$, for all $t \in \mathbb{R}$, it holds

$$
\begin{aligned}
R_i(\bar{\omega}_i^k) &= \mathbb{E}_{\omega_i}\left[\lceil \omega_i - \bar{\omega}_i^k \rceil\right] \\
&= \mathbb{E}_{\omega_i}\left[\lceil \omega_i - \lfloor \omega_i \rfloor - \bar{\omega}_i^k \rceil\right] + \mathbb{E}_{\omega_i}\left[\lfloor \omega_i \rfloor\right] \\
&= \sum_{s=1}^{S_i} p_i^s \lceil \bar{\omega}_i^s - \bar{\omega}_i^k \rceil + \mathbb{E}_{\omega_i}\left[\lfloor \omega_i \rfloor\right].
\end{aligned}
$$

Using that $\lceil \bar{\omega}_i^s - \bar{\omega}_i^k \rceil = 0$ if $s \le k$ and $\lceil \bar{\omega}_i^s - \bar{\omega}_i^k \rceil = 1$ if $s > k$, we see that

$$
R_i(\bar{\omega}_i^k) + \bar{\omega}_i^k = \sum_{s=k+1}^{S_i} p_i^s + \mathbb{E}_{\omega_i}\left[\lfloor \omega_i \rfloor\right] + \bar{\omega}_i^k. \tag{10}
$$

Since the second term of the right-hand side of (10) is a constant, this completes the proof.

$\square$

Example 1. Let the random variable ω_i be uniformly distributed on $(0, u_i)$, with $0 < u_i < 1$. Then, for $z_i \in [0, 1)$,

$$
-\sum_{k=-\infty}^{\infty} f_i^-(z_i + k) = \begin{cases} -1/u_i, & z_i \in (0, u_i] \\ 0, & \text{otherwise}, \end{cases}
$$

and

$$
-\sum_{k=-\infty}^{\infty} f_i^+(z_i + k) = \begin{cases} -1/u_i, & z_i \in [0, u_i) \\ 0, & \text{otherwise}, \end{cases}
$$

so that $\alpha_i^\star = u_i$. See Figure 4.1 for the corresponding functions R_i and $R_{\alpha_i^\star}$. $\blacksquare$

5. Convex approximation of the expected value function

We now return to the analysis of the expected value function $\mathcal{Q}$. Instead of studying $\mathcal{Q}$ as a function of the first-stage variables $x \in \mathbb{R}_+^{n_1}$, we first consider it as a function of the
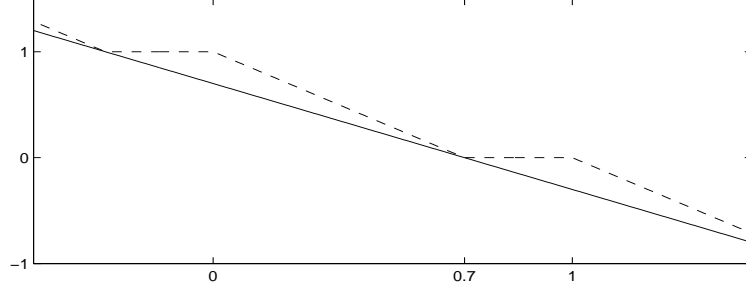
8

Figure 4.1: The functions R_i (dashed) and its convex hull $R_{\alpha_i^\star}$ of Example 1.

tender variables $z := Tx \in \mathbb{R}^m$. To avoid confusion, we will denote the latter function by Q.

Theorem 5.1   *Consider the integer recourse expected value function Q, defined as*

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

Under the assumptions of Section 2, in particular that W is totally unimodular, the convex hull of Q is the function $Q_{\alpha^\star}$, defined as

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

where $\alpha^\star$ is defined by (8), and $\varphi_{\alpha^\star}$ is the discrete random vector $\varphi_{\alpha^\star} = \lceil \omega - \alpha^\star \rceil + \alpha^\star$ with support in $\alpha^\star + \mathbb{Z}^m$, and

$$\Pr\{\varphi_{\alpha^\star} = \alpha^\star + k\} = \Pr\{\omega \in C_{\alpha^\star}^k\}, \quad k \in \mathbb{Z}^m,$$

with $C_{\alpha^\star}^k := \prod_{i=1}^m \left(\alpha_i^\star + k_i - 1, \alpha_i^\star + k_i \right], k \in \mathbb{Z}^m.$

PROOF.   First we will prove that $Q_{\alpha^\star}$ is a lower bound for Q, and subsequently that $Q_{\alpha^\star}(z) = Q(z)$ for all $z \in \alpha^\star + \mathbb{Z}^m$. This completes the proof, since all vertices of the polyhedral function $Q_{\alpha^\star}$ are contained in $\alpha^\star + \mathbb{Z}^m$.

Using the dual representation (4) of the value function v, we have

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

9

and, analogously,

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

Conditioning on the events $\omega \in C_{\alpha^\star}^l$, $l \in \mathbb{Z}^m$, we obtain, for $z \in \mathbb{R}^m$,

$$
\begin{aligned}
Q(z) &= \sum_{l \in \mathbb{Z}^m} \Pr\{\omega \in C_{\alpha^\star}^l\}\, \mathbb{E}_\omega\left[\max_{k=1,\dots,K} \lambda^k \lceil \omega - z \rceil \,\big|\, \omega \in C_{\alpha^\star}^l\right] \\
&\geq \sum_{l \in \mathbb{Z}^m} \Pr\{\omega \in C_{\alpha^\star}^l\} \max_{k=1,\dots,K} \lambda^k \mathbb{E}_\omega\left[\lceil \omega - z \rceil \,\big|\, \omega \in C_{\alpha^\star}^l\right] \\
&\geq \sum_{l \in \mathbb{Z}^m} \Pr\{\omega \in C_{\alpha^\star}^l\} \max_{k=1,\dots,K} \lambda^k \mathbb{E}_{\omega_{\alpha^\star}}\left[\lceil \omega_{\alpha^\star} - z \rceil \,\big|\, \omega_{\alpha^\star} \in C_{\alpha^\star}^l\right] \\
&= \sum_{l \in \mathbb{Z}^m} \Pr\{\omega \in C_{\alpha^\star}^l\} \max_{k=1,\dots,K} \lambda^k \left(\alpha^\star + l - z\right) \\
&= \sum_{l \in \mathbb{Z}^m} \Pr\{\varphi_{\alpha^\star} = \alpha^\star + l\} \max_{k=1,\dots,K} \lambda^k \left(\alpha^\star + l - z\right) = Q_{\alpha^\star}(z).
\end{aligned}
$$

The second inequality is valid because each λ^k is nonnegative, so that the α-approximation $\lambda^k \mathbb{E}_{\omega_{\alpha^\star}}\left[\lceil \omega_{\alpha^\star} - z \rceil \,\big|\, \omega_{\alpha^\star} \in C_{\alpha^\star}^l\right]$ is a lower bound for $\lambda^k \mathbb{E}_\omega\left[\lceil \omega - z \rceil \,\big|\, \omega \in C_{\alpha^\star}^l\right]$ by the choice of $\alpha^\star$. The subsequent equality holds by Lemma 4.2.

It remains to prove that $Q_{\alpha^\star} = Q$ on $\alpha^\star + \mathbb{Z}^m$. Consider a fixed $\bar{z} \in \alpha^\star + \mathbb{Z}^m$ and a fixed $l \in \mathbb{Z}^m$. Then $\lceil \omega - \bar{z} \rceil = l - \lfloor \bar{z} \rfloor$ is constant for all $\omega \in C_\alpha^l$, so that there exists a $\lambda(\bar{z}, l)$ satisfying

$$\lambda(\bar{z}, l) \in \operatorname*{argmax}_{k=1,\dots,K} \lambda^k \lceil \omega - \bar{z} \rceil \quad \forall \omega \in C_\alpha^l.$$

Since this is true for every $\bar{z} \in \alpha^\star + \mathbb{Z}^m$ and $l \in \mathbb{Z}^m$, it follows that, for $z \in \alpha^\star + \mathbb{Z}^m$,

$$
\begin{aligned}
Q(z) &= \sum_{l \in \mathbb{Z}^m} \Pr\{\omega \in C_{\alpha^\star}^l\}\, \lambda(z, l)\, \mathbb{E}_\omega\left[\lceil \omega - z \rceil \,\big|\, \omega \in C_{\alpha^\star}^l\right] \\
&= \sum_{l \in \mathbb{Z}^m} \Pr\{\omega \in C_{\alpha^\star}^l\}\, \lambda(z, l)\, \mathbb{E}_{\omega_{\alpha^\star}}\left[\lceil \omega_{\alpha^\star} - z \rceil \,\big|\, \omega_{\alpha^\star} \in C_{\alpha^\star}^l\right] \\
&= \sum_{l \in \mathbb{Z}^m} \Pr\{\omega \in C_{\alpha^\star}^l\}\, \lambda(z, l)\, \left(\alpha^\star + l - z\right) \\
&= \sum_{l \in \mathbb{Z}^m} \Pr\{\omega \in C_{\alpha^\star}^l\} \max_{k=1,\dots,K} \lambda^k \left(\alpha^\star + l - z\right).
\end{aligned}
$$

10

The second equality follows from the fact that each α-approximation is sharp on $\alpha+\mathbb{Z}^n$. The last equality follows from the definition of $\lambda(z,l)$ and $\varphi^l_{\alpha^\star} - z = l - \lfloor z \rfloor$, $z \in \alpha^\star + \mathbb{Z}^m$. $\qquad\square$

We now return to the function $\mathcal{Q}$, which gives the expected recourse costs for $x \in \mathbb{R}^{n_1}$. It is intuitively clear that the construction of the convex hull of $\mathcal{Q}$ needs to be based on a complete description of $\mathcal{Q}$. This idea is formalized in the following corollary.

Corollary 5.1 *Consider the integer recourse expected value function $\mathcal{Q}$, defined as*

$$\mathcal{Q}(x) = \mathbb{E}_\omega \left[\min_y qy : Wy \geq \omega - Tx, \ y \in \mathbb{Z}_+^{n_2} \right], \quad x \in \mathbb{R}^{n_1}, \tag{11}$$

and the continuous recourse expected value function $\mathcal{Q}_{\alpha^\star}$, defined as

$$\mathcal{Q}_{\alpha^\star}(x) = \mathbb{E}_{\varphi_{\alpha^\star}} \left[\min_y qy : Wy \geq \varphi_{\alpha^\star} - Tx, \ y \in \mathbb{R}_+^{n_2} \right], \quad x \in \mathbb{R}^{n_1}, \tag{12}$$

with $\alpha^\star$ and $\varphi_{\alpha^\star}$ as defined in Theorem 5.1.

Under the assumptions of Section 2, in particular that W is totally unimodular, the function $\mathcal{Q}_{\alpha^\star}$ is the convex hull of $\mathcal{Q}$ if the matrix T is of full row rank. If not, then $\mathcal{Q}_{\alpha^\star}$ is a lower bound for $\mathcal{Q}$.

PROOF.   The result follows from Theorem 5.1 and Theorem 2.2 in [5]. $\qquad\square$

Example 2.   Consider the value function

$$
\begin{aligned}
v(s) \ &= \ \min\{y : y \geq s_1, \ y \geq s_2, \ y \in \mathbb{Z}_+\} \\
&= \ \max\{\lceil s_1 \rceil, \lceil s_2 \rceil, 0\}, \qquad\qquad s \in \mathbb{R}^2,
\end{aligned}
$$

and the expected value function $\mathcal{Q}(x) = \mathbb{E}_\omega \left[v(\omega - x) \right]$, $x \in \mathbb{R}^2$. Note that the recourse matrix $W = (1\ 1)'$ is totally unimodular.
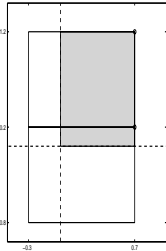
Let ω be uniformly distributed on $(0, 0.7) \times (0, 1.2)$. Using Lemma 4.3 (i) we find that $\alpha^\star = (0.7, 0.2)$, so that by Theorem 5.1 the distribution of the discrete random vector $\varphi_{\alpha^\star}$ is given by $\Pr\{\varphi_{\alpha^\star} = (0.7, 0.2)\} = 1/6$, $\Pr\{\varphi_{\alpha^\star} = (0.7, 1.2)\} = 5/6$.

Since T is the identity matrix, it is of full row rank. Hence, by Corollary 5.1 the function $\mathcal{Q}_{\alpha^\star}$ is the convex hull of $\mathcal{Q}$, see Figure 5.1.
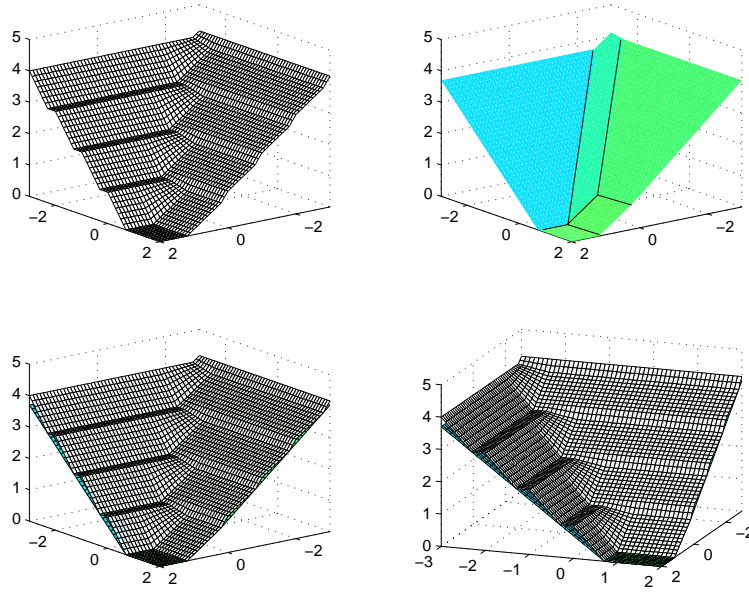
11

Figure 5.1: The functions $\mathcal{Q}$ (top left) and its convex hull $\mathcal{Q}_{\alpha^\star}$ (top right) of Example 2; both functions are shown in the bottom two figures.

∎

We conclude that if W is totally unimodular and if T is of full row rank, then solving the *unrestricted* integer recourse problem $\min_x cx + \mathcal{Q}(x)$ is equivalent to solving $\min_x cx + \mathcal{Q}_{\alpha^\star}(x)$. The latter is a continuous recourse problem with discretely distributed right-hand side vector, for which several good algorithms exist, see e.g. [11].

6.   General complete recourse

In this section we present results for the case that the recourse matrix is not totally unimodular; we still assume, however, that W is integral (which can be obtained by scaling if W is rational) and such that the recourse structure is complete.

In this case the convex function $\mathcal{Q}_{\alpha^\star}$ is still a lower bound for $\mathcal{Q}$, since (2) now holds

12

with inequality, so that

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

but $\mathcal{Q}_{\alpha^\star}$ can not be expected to equal to the convex hull of $\mathcal{Q}$. However, we will prove that $\mathcal{Q}_{\alpha^\star}$ is a strictly better convex approximation than the one obtained using the LP relaxation of the second-stage problem. The latter convex function will be denoted by $\mathcal{Q}^{LP}$, defined as

$$\mathcal{Q}^{LP}(x) := \mathbb{E}_\omega \left[\min_y \left\{ qy : Wy \geq \omega - Tx, \ y \in \mathbb{R}_+^{n_2} \right\} \right], \quad x \in \mathbb{R}^{n_1}. \tag{13}$$

Theorem 6.1 *Consider the functions $\mathcal{Q}_{\alpha^\star}$ and $\mathcal{Q}^{LP}$, defined by (12) and (13) respectively, which both are convex lower bounds for the integer recourse expected value function $\mathcal{Q}$, defined by (11).*

(a) $\mathcal{Q}_{\alpha^\star} \geq \mathcal{Q}^{LP}$

(b) Assume

   (i) $q \geq 0$, so that 0 is a trivial lower bound for v and $\mathcal{Q}$;

   (ii) there exists a subset L of $\mathbb{Z}^m$ such that the support $\Omega \subset \bigcup_{l \in L} \{ \omega : \omega \leq \alpha^\star + l \}$ and $\Pr\{ \omega < \alpha^\star + l \mid \omega \in C_{\alpha^\star}^l \} > 0$ for all $l \in L$.

   Then the function $\mathcal{Q}_{\alpha^\star}$ is a strictly better convex approximation of $\mathcal{Q}$ than $\mathcal{Q}^{LP}$, in the sense that $\mathcal{Q}(x) > 0$ implies $\mathcal{Q}_{\alpha^\star}(x) > \mathcal{Q}^{LP}(x)$.

PROOF.   As before, we condition on the events $\omega \in C_{\alpha^\star}^l$, $l \in \mathbb{Z}^m$, to obtain, for $x \in \mathbb{R}^{n_1}$,

$$\mathcal{Q}_{\alpha^\star}(x) = \sum_{l \in \mathbb{Z}^m} \Pr\{\omega \in C_{\alpha^\star}^l\} \max_{k=1,\ldots,K} \lambda^k \left(\alpha^\star + l - Tx \right) \tag{14}$$

and

$$\mathcal{Q}^{LP}(x) = \sum_{l \in \mathbb{Z}^m} \Pr\{\omega \in C_{\alpha^\star}^l\} \mathbb{E}_\omega \left[\max_{k=1,\ldots,K} \lambda^k (\omega - Tx) \mid \omega \in C_{\alpha^\star}^l \right]. \tag{15}$$

For each $l \in \mathbb{Z}^m$ it follows from the definition of $C_{\alpha^\star}^l = \prod_{i=1}^m \left(\alpha_i^\star + l_i - 1, \alpha_i^\star + l_i \right]$ that $\alpha^\star + l \geq \omega$ for all $\omega \in C_{\alpha^\star}^l$. Thus, for $k = 1, \ldots, K$, $\lambda^k (\alpha^\star + l - Tx) \geq \lambda^k (\omega - Tx)$ for all $\omega \in C_{\alpha^\star}^l$, since $\lambda^k \geq 0$. Substitution in (14) and (15) proves that $\mathcal{Q}_{\alpha^\star} \geq \mathcal{Q}^{LP}$.

To prove (b), we first show that $\mathcal{Q}(x) > 0$ implies $\mathcal{Q}_{\alpha^\star}(x) > 0$. To this end, define

$$N(x) := \left\{ t \in \mathbb{R}^m : v(t - Tx) > 0 \right\}, \quad x \in \mathbb{R}^{n_1}.$$

13

Then $\mathcal{Q}(x) > 0$ if and only if $\Pr\{\omega \in N(x)\} > 0$, which is equivalent to $\Pr\{\omega \in \text{int } N(x)\} > 0$ since $N(x)$ is an open set. By Definition 4.1, it follows that then also $\Pr\{\omega_\alpha \in N(x)\} > 0$, which implies $\mathcal{Q}_\alpha(x) > 0$ for all $\alpha \in [0, 1)^m$.

Let x be such that $\mathcal{Q}(x) > 0$, implying $\mathcal{Q}_{\alpha^\star}(x) > 0$. Then, since each term of (14) is non-negative by assumption (i), there exists an $\bar{l} \in L$ such that $\max_{k=1,\ldots,K} \lambda^k \left(\alpha^\star + \bar{l} - Tx \right) > 0$; obviously, any optimal solution $\bar\lambda$ of this problem satisfies $\bar\lambda \neq 0$. For an arbitrary but fixed $\bar\omega \in C_{\alpha^\star}^{\bar{l}}$ such that $\bar\omega < \alpha^\star + \bar{l}$, it holds

$$\lambda \left(\bar\omega - Tx \right) \leq \lambda \left(\alpha^\star + \bar{l} - Tx \right) \quad \forall \lambda \geq 0,$$

with strict inequality unless $\lambda = 0$. Let $\hat\lambda$ be an optimal solution of $\max_{k=1,\ldots,K} \lambda^k \left(\bar\omega - Tx \right)$. Then there are two possibilities:

(i) $\hat\lambda = 0$, so that $\hat\lambda \left(\bar\omega - Tx \right) = 0 < \bar\lambda \left(\alpha^\star + \bar{l} - Tx \right)$;

(ii) $\hat\lambda \neq 0$, so that $\hat\lambda \left(\bar\omega - Tx \right) < \hat\lambda \left(\alpha^\star + \bar{l} - Tx \right) \leq \bar\lambda \left(\alpha^\star + \bar{l} - Tx \right)$.

We conclude that, for all $\bar\omega \in C_{\alpha^\star}^{\bar{l}}$ with $\bar\omega < \alpha^\star + \bar{l}$,

$$\max_{k=1,\ldots,K} \lambda^k \left(\alpha^\star + \bar{l} - Tx \right) > \max_{k=1,\ldots,K} \lambda^k \left(\bar\omega - Tx \right). \tag{16}$$

Since $\Pr\{\omega < \alpha^\star + \bar{l} \mid \omega \in C_{\alpha^\star}^{\bar{l}}\} > 0$ by assumption (ii), and (16) holds with weak inequality for all $\omega \in C_{\alpha^\star}^{\bar{l}}$, it follows that

$$\max_{k=1,\ldots,K} \lambda^k \left(\alpha^\star + \bar{l} - Tx \right) > \mathbb{E}_\omega \left[\max_{k=1,\ldots,K} \lambda^k \left(\omega - Tx \right) \,\middle|\, \omega \in C_{\alpha^\star}^{\bar{l}} \right]. \tag{17}$$

Finally, using that (17) holds with weak inequality for all $l \in \mathbb{Z}^m$, we see from (14) and (15) that $\mathcal{Q}_{\alpha^\star}(x) > \mathcal{Q}^{LP}(x)$. $\square$

Corollary 6.1 *Assume that ω follows a non-degenerated continuous distribution and $q \geq 0$. Then $\mathcal{Q}_{\alpha^\star}$ is strictly better convex approximation of $\mathcal{Q}$ than $\mathcal{Q}^{LP}$ in the sense of Theorem 6.1.*

PROOF. Immediate from the observation that condition (ii) of Theorem 6.1 is trivially satisfied in this case. $\square$

REMARK. If the support of ω is contained in $\alpha^\star + \mathbb{Z}^m$, then assumption (ii) of Theorem 6.1 is not satisfied. In this case $\mathcal{Q}_{\alpha^\star} = \mathcal{Q}^{LP}$ since $\varphi_{\alpha^\star} = \omega$, as can be readily verified.
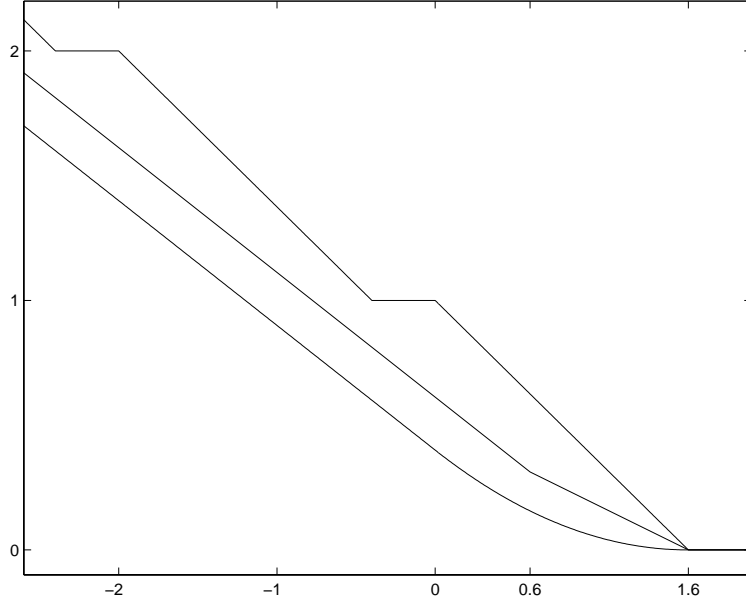
Figure 6.1: The functions $\mathcal{Q}$, $\mathcal{Q}_{\alpha^\star}$, and $\mathcal{Q}^{LP}$ of Example 3.

Example 3. Consider the value function

$$v(s) = \min_{y}\{y : 2y \geq s, \ y \in \mathbb{Z}_+\}, \quad s \in \mathbb{R},$$

and the expected value function $\mathcal{Q}(x) = \mathbb{E}_\omega[v(\omega - x)]$, $x \in \mathbb{R}$, with ω uniformly distributed on $(0, 1.6)$. Note that the recourse structure is complete but that the recourse matrix $W = 2$ is not totally unimodular.

Using Lemma 4.3 (i) we find that $\alpha^\star = 0.6$, so that by Theorem 5.1 the distribution of the discrete random variable $\varphi_{\alpha^\star}$ is given by $\Pr\{\varphi_{\alpha^\star} = 0.6\} = 3/8$, $\Pr\{\varphi_{\alpha^\star} = 1.6\} = 5/8$.

The assumptions of Theorem 6.1 are satisfied (with $q = 1$ and $L = \{1\}$), so that $\mathcal{Q}_{\alpha^\star}(x) > \mathcal{Q}^{LP}(x)$ for all x such that $\mathcal{Q}(x) > 0$, see Figure 6.1.

∎

Moreover, $\mathcal{Q}_{\alpha^\star}$ is *computationally more tractable* than $\mathcal{Q}^{LP}$. Indeed, $\mathcal{Q}_{\alpha^\star}$ is the expectation of an LP value function with respect to the random vector $\varphi_{\alpha^\star}$, which is discretely distributed by Theorem 5.1. On the other hand, $\mathcal{Q}^{LP}$ is the expectation of the same

15

function with respect to the arbitrarily distributed random vector ω. In particular, if the distribution of ω is continuous, each evaluation of $\mathcal{Q}^{LP}$ amounts to computing an m-dimensional integral. Even if ω is a discrete random vector, computation of $\mathcal{Q}_{\alpha^\star}$ is possibly easier than that of $\mathcal{Q}^{LP}$, since the number of realizations of $\varphi_{\alpha^\star}$ is not larger than that of ω.

Finally, we consider the practical use of the convex approximation $\mathcal{Q}_{\alpha^\star}$ for the current case, i.e., if the recourse matrix W is not totally unimodular.

The function $\mathcal{Q}_{\alpha^\star}$ can of course be used instead of $\mathcal{Q}$ in order to solve the corresponding integer recourse problem approximately. It should be stressed, however, that no (non-trivial) bound on the approximation error is known. On the other hand, this approach is easy to implement given the availability of continuous recourse solvers (see e.g. [11]).

We believe, however, that the convex lower bound $\mathcal{Q}_{\alpha^\star}$ is most useful as a building block in a range of special-purpose algorithms for complete integer recourse models. Several of these algorithms, such as *integer L-shaped* [9], *stochastic branch-and-bound* [13], and *structured enumeration* [16] (see also [1]), use the function $\mathcal{Q}^{LP}$ (corresponding to the LP relaxation) for bounding purposes. As explained above, the function $\mathcal{Q}_{\alpha^\star}$ provides an alternative convex lower bound which in many cases is both better as well as easier to compute.

7. Summary and conclusions

We have shown that optimal α-approximations, which are obtained through a specific modification of the distribution of the right-hand side parameters, provide good convex approximations of the integer recourse expected value function. Indeed, if the recourse matrix is totally unimodular then the α-approximation equals the convex hull. For general complete recourse structures, the result is a convex lower bound which is strictly better than the one corresponding to the LP relaxation. In all cases, the convex approximation can be represented as the expected value function of a continuous recourse problem, whose right-hand side parameters follow an explicitly given discrete distribution.

The results of this paper generalize those obtained for models with simple integer recourse [6] and multiple simple (integer) recourse [20]. In particular, they do not depend on separability of the second-stage value function, which so far appeared to be an indispensable condition for this approach. In this sense the current results open up a wide range of possibilities for constructing convex approximations for integer recourse

16

models, both for general and specific recourse structures. In future research we will combine this approach with the use of certain classes of valid inequalities (see also [3, 17]); similar approaches for mixed-integer models will also be investigated.

Acknowledgement

References

[1] S. Ahmed, M. Tawarmalani, and N.V. Sahinidis. A finite branch and bound algorithm for two-stage stochastic integer programs. Stochastic Programming E-Print Series, `http://dochost.rz.hu-berlin.de/speps/`, 2000.

[2] J.R. Birge and F.V. Louveaux. *Introduction to Stochastic Programming.* Springer Verlag, 1997.

[3] C.C. Carøe and J. Tind. L-shaped decomposition of two-stage stochastic programs with integer recourse. *Math. Program.*, 83(3):451–464, 1998.

[4] P. Kall and S.W. Wallace. *Stochastic Programming.* Wiley, Chichester, 1994.

[5] W.K. Klein Haneveld, L. Stougie, and M.H. van der Vlerk. On the convex hull of the composition of a separable and a linear function. Discussion Paper 9570, CORE, Louvain-la-Neuve, Belgium, 1995.

[6] W.K. Klein Haneveld, L. Stougie, and M.H. van der Vlerk. Convex approximations for simple integer recourse models by perturbing the underlying distribution. Research Report 97A19, SOM, University of Groningen, `http://som.rug.nl`, 1997.

[7] W.K. Klein Haneveld, L. Stougie, and M.H. van der Vlerk. Convex simple integer recourse problems. Research Report 97A10, SOM, University of Groningen, `http://som.rug.nl`, 1997.

[8] W.K. Klein Haneveld and M.H. van der Vlerk. Stochastic integer programming: General models and algorithms. *Ann. Oper. Res.*, 85:39–57, 1999.

[9] G. Laporte and F.V. Louveaux. The integer L-shaped method for stochastic integer programs with complete recourse. *Oper. Res. Lett.*, 13:133–142, 1993.

[10] F.V. Louveaux and M.H. van der Vlerk. Stochastic programming with simple integer recourse. *Math. Program.*, 61:301–325, 1993.

[11] J. Mayer. *Stochastic Linear Programming Algorithms: A Comparison Based on a Model Management System.* Optimization theory and applications; v. 1. Gordon and Breach Science Publishers, OPA Amsterdam, The Netherlands, 1998.

17

[12] G.L. Nemhauser and L.A. Wolsey. *Integer and Combinatorial Optimization.* Wiley, New York, 1988.

[13] V.I. Norkin, Yu.M. Ermoliev, and A. Ruszczyński. On optimal allocation of indivisibles under uncertainty. *Oper. Res.*, 46(3):381–395, 1998.

[14] A. Prékopa. *Stochastic Programming.* Kluwer Academic Publishers, Dordrecht, 1995.

[15] R. Schultz. Continuity properties of expectation functions in stochastic integer programming. *Math. Oper. Res.*, 18:578–589, 1993.

[16] R. Schultz, L. Stougie, and M.H. van der Vlerk. Solving stochastic programs with complete integer recourse: A framework using Gröbner Bases. *Math. Program.*, 83(2):229–252, 1998.

[17] S. Sen and J.L. Higle. The C^3 theorem and a D^2 algorithm for large scale stochastic integer programming: Set convexification. Stochastic Programming E-Print Series, `http://dochost.rz.hu-berlin.de/speps/`, 2000.

[18] M.H. van der Vlerk. *Stochastic programming with integer recourse.* PhD thesis, University of Groningen, The Netherlands, 1995.

[19] M.H. van der Vlerk. Convex approximations for stochastic programs with simple integer recourse. In W.K. Klein Haneveld, O.J. Vrieze, and L.C.M. Kallenberg, editors, *Ten years LNMB: Ph.D. research and graduate courses of the Dutch Network for Operations Research, CWI Tract 122*, pages 357–365. CWI, Amsterdam, 1997.

[20] M.H. van der Vlerk. On multiple simple recourse models. Research Report 02A06, SOM, University of Groningen, `http://som.rug.nl`, 2002.