Total variation bounds on the expectation of periodic functions with applications to recourse approximations

Romeijnders, Ward; van der Vlerk, Maarten H.; Klein Haneveld, Willem K.

Published in:
Mathematical Programming

DOI:
10.1007/s10107-014-0829-2

Math. Program., Ser. B (2016) 157:3–46
DOI 10.1007/s10107-014-0829-2

FULL LENGTH PAPER

Total variation bounds on the expectation of periodic functions with applications to recourse approximations

Ward Romeijnders · Maarten H. van der Vlerk · Willem K. Klein Haneveld

Received: 24 January 2014 / Accepted: 10 October 2014 / Published online: 30 October 2014

Abstract We derive a lower and upper bound for the expectation of periodic functions, depending on the total variation of the probability density function of the underlying random variable. Using worst-case analysis we derive tighter bounds for functions that are periodically monotone. These bounds can be used to evaluate the performance of approximations for both continuous and integer recourse models. In this paper, we introduce a new convex approximation for totally unimodular recourse models, and we show that this convex approximation has the best worst-case error bound possible, improving previous bounds with a factor 2. Moreover, we use similar analysis to derive error bounds for two types of discrete approximations of continuous recourse models with continuous random variables. Furthermore, we derive a tractable Lipschitz constant for pure integer recourse models.

Keywords Periodic functions · Total variation · Stochastic Programming · Integer recourse · Convex approximations · Discrete approximations

Mathematics Subject Classification 90C15 · 90C10

This research is supported by The Netherlands Organisation for Scientific Research (NWO).

W. Romeijnders (✉) · M. H. van der Vlerk · W. K. Klein Haneveld
Department of Operations, Faculty of Economics and Business,
University of Groningen, PO Box 800, 9700 AV, Groningen,
The Netherlands
e-mail: w.romeijnders@rug.nl

M. H. van der Vlerk
e-mail: m.h.van.der.vlerk@rug.nl

W. K. Klein Haneveld
e-mail: w.k.klein.haneveld@rug.nl

1 Introduction

Let φ be a real-valued, periodic function, and let ω denote a continuously distributed random variable with probability density function (pdf) f of bounded variation. We consider the expected value $\mathbb{E}_f[\varphi(\omega)]$. Usually, it is not possible to obtain a closed form expression for this expectation. That is why we will derive lower and upper bounds for $\mathbb{E}_f[\varphi(\omega)]$, both depending on the *total variation* of f.

Periodic functions arise naturally in real life. Consider for example the motion of the tides, household voltage, or the blood flow through an artery. Also in mathematics and physics, periodic functions such as the trigonometric functions play an important role. Motivating the current research, such functions may also arise as a result of rounding: the difference function $\lceil x \rceil - x$, $x \in \mathbb{R}$, is periodic.

In [27] we show that the error of a class of convex approximations of totally unimodular (TU) integer recourse models can be bounded by the expectation of a periodic function φ. By deriving bounds on $\mathbb{E}_f[\varphi(\omega)]$ in terms of the total variation of f, we obtain a uniform error bound for this class of approximations. However, for this application the analysis is restricted to a special class of two-valued piecewise constant periodic functions.

In this paper we derive bounds for $\mathbb{E}_f[\varphi(\omega)]$ that hold for all periodic functions φ. These bounds are the result of a worst-case analysis, where for every $B \in \mathbb{R}$ with $B > 0$, we consider

$$M(\varphi, B) := \sup_{f \in \mathcal{F}} \Big\{ \mathbb{E}_f[\varphi(\omega)] : |\Delta| f \leq B \Big\},$$

and

$$N(\varphi, B) := \inf_{f \in \mathcal{F}} \Big\{ \mathbb{E}_f[\varphi(\omega)] : |\Delta| f \leq B \Big\},$$

with $\mathcal{F}$ denoting the set of density functions of bounded variation and $|\Delta| f$ the total variation of f, both defined in Sect. 2. Surprisingly, it is possible to derive closed form expressions for $M(\varphi, B)$ and $N(\varphi, B)$ when φ is periodically monotone. In all other cases we obtain an upper and lower bound, respectively, yielding Proposition 1 in Sect. 2.

Although the results are valid for one-dimensional periodic functions only, it is possible to apply them in a higher dimensional setting. Using additional analysis we are able to derive error bounds for convex approximations of TU integer recourse models. We introduce a new convex approximation, improving the best known error bound of [27] by a factor 2, and prove that this approximation has the best worst-case error bound possible. Moreover, we show that the same analysis can be used to obtain error bounds for so-called discrete approximations of continuous recourse models, providing a link between two seemingly unrelated areas of research. Furthermore, we derive a tractable Lipschitz constant for the expected value function of pure integer recourse models.

The remainder of this paper is organized as follows. In Sect. 1.1 we first discuss properties of $M(\varphi, B)$ and $N(\varphi, B)$. Next, in Sect. 2 we introduce the concepts of total variation and *packed densities*, and we derive an upper and lower bound on $M(\varphi, B)$

and $N(\varphi, B)$, respectively. Section 3 introduces lemmas on the *flattening of density functions*, which are used to derive exact expressions for $M(\varphi, B)$ and $N(\varphi, B)$ in case φ is periodically monotone. Finally, in Sect. 4 we derive the indicated results on approximations of recourse models, and in Sect. 5 we give a summary and conclusions. Readers only interested in the results on approximations of recourse models, and not on their derivation, may skip Sects. 1–3 and proceed directly to Sect. 4.

1.1 Properties of $M(\varphi, B)$ and $N(\varphi, B)$

In this subsection we discuss properties of $M(\varphi, B)$ and $N(\varphi, B)$. We collect them here for easy reference, as they are used frequently in the remainder of this paper. The first set of properties deals with shifting and scaling of periodic functions φ.

Lemma 1 *Let $\varphi : \mathbb{R} \mapsto \mathbb{R}$ be a periodic function with period p and finite mean value $\nu := p^{-1} \int_0^p \varphi(x)dx$. Then, for every $B \in \mathbb{R}$ with $B > 0$,*

(i) $M(\varphi, B) = -N(-\varphi, B)$ *and* $N(\varphi, B) = -M(-\varphi, B)$.
(ii) *for every $r > 0$,*

$$M(r\varphi, B) = rM(\varphi, B) \quad \textit{and} \quad N(r\varphi, B) = rN(\varphi, B).$$

(iii) *if $\bar{\varphi}_r(x) := \varphi(x/r),\ x \in \mathbb{R}$, for some $r > 0$, then*

$$M(\bar{\varphi}_r, B) = M(\varphi, rB) \quad \textit{and} \quad N(\bar{\varphi}_r, B) = N(\varphi, rB).$$

(iv) *if $\hat{\varphi}(x) := \varphi(-x),\ x \in \mathbb{R}$, then*

$$M(\hat{\varphi}, B) = M(\varphi, B) \quad \textit{and} \quad N(\hat{\varphi}, B) = N(\varphi, B).$$

(V) *if $\tilde{\varphi}_\beta(x) := \varphi(x - \beta),\ x \in \mathbb{R}$, for some $\beta \in \mathbb{R}$, then*

$$M(\tilde{\varphi}_\beta, B) = M(\varphi, B) \quad \textit{and} \quad N(\tilde{\varphi}_\beta, B) = N(\varphi, B).$$

Proof See the Appendix. □

The next lemma shows convexity properties of $M(\varphi, B)$ and $N(\varphi, B)$. These properties hold because $M(\varphi, B)$ and $N(\varphi, B)$ can be considered as convex optimization problems. The results are presented here for periodic functions φ, but hold, in fact, for a much more general class of functions.

Lemma 2 *Let $\varphi : \mathbb{R} \mapsto \mathbb{R}$ be a periodic function with period p and finite mean value $\nu := p^{-1} \int_0^p \varphi(x)dx$. Then,*

(i) *$M(\varphi, B)$ is concave in B, and*
(ii) *$N(\varphi, B)$ is convex in B.*

Proof See the Appendix. □

2 Packed densities and total variation

Before we derive bounds on $\mathbb{E}_f[\varphi(\omega)]$, we first introduce some notation. Let $p \in \mathbb{R}$ with $p > 0$ denote the period of φ and $\nu_\varphi := p^{-1} \int_0^p \varphi(x)dx$ its mean value over any interval of length p. It is assumed throughout that this mean value ν_φ exists and is finite. We drop the index φ in case it can easily be deduced from the context. Thus, the function $\varphi_\nu : \mathbb{R} \mapsto \mathbb{R}$ defined as

$$\varphi_\nu(x) = \varphi(x) - \nu, \quad x \in \mathbb{R},$$

has the same shape and period as φ, but its mean equals zero.

Throughout this paper we only consider pdf f of bounded variation, and we let $\mathcal{F}$ denote the set containing these functions. Note that for practical purposes this definition of $\mathcal{F}$ is not very restrictive.

Definition 1 A function $f : \mathbb{R} \mapsto \mathbb{R}$ is of bounded variation if and only if there exist bounded monotone non-decreasing functions f_1 and f_2 such that $f = f_1 - f_2$.

Remark 1 Equivalently, we say that f is of bounded variation if and only if the total variation of f, to be defined in Definition 4, is finite [29].

Definition 2 Let $\mathcal{F}$ denote the set of one-dimensional probability density functions f of bounded variation.

Since φ is periodic with period p, for our purposes it is possible to summarize all relevant information of a pdf f in a so-called packed density f_p with support contained in $[0, p]$. This packed density is defined such that $\mathbb{E}_{f_p}[\varphi(\omega)] = \mathbb{E}_f[\varphi(\omega)]$, see Lemma 3.

Definition 3 For every $f \in \mathcal{F}$ and $p \in \mathbb{R}$ with $p > 0$, we define the packed density $f_p : \mathbb{R} \mapsto \mathbb{R}$ of f with support contained in $[0, p]$ as

$$f_p(x) := \begin{cases} \sum_{k\in\mathbb{Z}} f(x + pk), & x \in [0, p]; \\ 0, & \text{otherwise.} \end{cases}$$

Remark 2 Note that we define packed densities f_p on $[0, p]$ instead of $[0, p)$ so that $f_p(p) = f_p(0)$. This choice is more convenient in Lemma 4, where we compare the total variations of f_p and f, and it does not change the expectation $\mathbb{E}_{f_p}[\varphi(\omega)]$.

Lemma 3 *For every $f \in \mathcal{F}$ and periodic function φ with period p,*

$$\mathbb{E}_{f_p}[\varphi(\omega)] = \mathbb{E}_f[\varphi(\omega)].$$

Proof By definition of f_p and using the periodicity of φ, we have

$$\mathbb{E}_{f_p}[\varphi(\omega)] = \int_0^p \varphi(x) \sum_{k\in\mathbb{Z}} f(x + pk)dx = \int_0^p \sum_{k\in\mathbb{Z}} \varphi(x + pk) f(x + pk)dx.$$

Moreover, by interchanging summation and integration we obtain

$$\mathbb{E}_{f_p}[\varphi(\omega)] = \sum_{k\in\mathbb{Z}} \int_0^p \varphi(x+pk) f(x+pk)dx = \int_{-\infty}^{\infty} \varphi(x) f(x)dx = \mathbb{E}_f[\varphi(\omega)].$$

□

Lemma 3 implies that knowledge of f_p is sufficient to compute $\mathbb{E}_f[\varphi(\omega)]$. However, obtaining a useful closed-form expression for f_p may actually be as difficult as finding such an expression for $\mathbb{E}_f[\varphi(\omega)]$. That is why it is interesting to observe that the total variation of f_p on $[0, p]$ can be bounded by the total variation of f.

Definition 4 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 $|\Delta| f := |\Delta| f(\mathbb{R})$.

Lemma 4 *For every $f \in \mathcal{F}$ and $p \in \mathbb{R}$ with $p > 0$, the corresponding packed density f_p of f with support contained in $[0, p]$ satisfies*

$$|\Delta| f_p([0, p]) \le |\Delta| f.$$

Proof Let $f \in \mathcal{F}$ be given. We will show that for every $\epsilon > 0$ and $P \in \Pi([0, p])$ there exists $\bar{P} \in \Pi(\mathbb{R})$ such that $V_{f_p}(P) \le V_f(\bar{P}) + \epsilon$. Then,

$$|\Delta| f_p([0, p]) = \sup_{P\in\Pi([0,p])} V_{f_p}(P) \le \sup_{\bar{P}\in\Pi(\mathbb{R})} V_f(\bar{P}) + \epsilon = |\Delta| f + \epsilon \tag{1}$$

for every $\epsilon > 0$, and thus $|\Delta| f_p([0, p]) \le |\Delta| f$.

In order to prove (1) let $\epsilon > 0$ be given. For every $P \in \Pi([0, p])$ corresponding to f_p we will construct $\bar{P} \in \Pi(\mathbb{R})$ corresponding to f by repeating P on any interval $[kp, (k+1)p]$, $k \in \mathbb{Z}$. So let $P = \{x_1, \dots, x_{N+1}\} \in \Pi([0, p])$ be given. We assume without loss of generality that $x_1 = 0$ and $x_{N+1} = p$. Since $f \in \mathcal{F}$ it follows that $f_p(x_i) := \sum_{k\in\mathbb{Z}} f(x_i + pk)$ is finite for all $i = 1, \dots, N$, and thus there exists $K \in \mathbb{N}$ such that

$$\left| f_p(x_i) - \sum_{k=-K}^{K} f(x_i + pk) \right| < \frac{\epsilon}{2N} \quad \text{for all } i = 1, \dots, N+1. \tag{2}$$

Conveniently writing $\gamma_i := \sum_{k=-K}^{K} f(x_i + pk)$ for $i = 1, \ldots, N+1$, we have

$$\begin{aligned} V_{f_p}(P) &= \sum_{i=1}^{N} \Big| f_p(x_{i+1}) - f_p(x_i) \Big| \\ &= \sum_{i=1}^{N} \Big| f_p(x_{i+1}) - \gamma_{i+1} + \gamma_{i+1} - \gamma_i + \gamma_i - f_p(x_i) \Big|. \end{aligned}$$

By applying the triangle inequality and using (2) we have

$$\begin{aligned} V_{f_p}(P) &\leq \frac{\epsilon}{2} + \sum_{i=1}^{N} |\gamma_{i+1} - \gamma_i| + \frac{\epsilon}{2} \\ &= \epsilon + \sum_{i=1}^{N} \left| \sum_{k=-K}^{K} \Big(f(x_{i+1} + pk) - f(x_i + pk) \Big) \right|. \end{aligned}$$

Once more applying the triangle inequality and interchanging summations we obtain

$$V_{f_p}(P) \leq \epsilon + \sum_{k=-K}^{K} \sum_{i=1}^{N} \Big| f(x_{i+1} + pk) - f(x_i + pk) \Big|.$$

Now define $\bar{P}$ as the ordered set containing all elements in $P + p\{-K, \ldots, K\}$ and observe that $\bar{P} \in \Pi(\mathbb{R})$. Moreover, since $x_{N+1} + pk = x_1 + p(k+1)$ for every $k \in \mathbb{Z}$ it follows that

$$V_f(\bar{P}) = \sum_{k=-K}^{K} \sum_{i=1}^{N} \Big| f(x_{i+1} + pk) - f(x_i + pk) \Big|,$$

and thus $V_{f_p}(P) \leq V_f(\bar{P}) + \epsilon$, which completes the proof. $\square$

Now we are ready to prove one of our main results.

Proposition 1 *Let $\varphi : \mathbb{R} \mapsto \mathbb{R}$ be a periodic function with period p and finite mean value $\nu := p^{-1} \int_0^p \varphi(x)dx$. Then,*

$$M(\varphi, B) := \sup_{f \in \mathcal{F}} \Big\{ \mathbb{E}_f[\varphi(\omega)] : |\Delta| f \leq B \Big\} \leq \nu + \frac{B}{4} \int_0^p |\varphi_\nu(x)| dx, \tag{3}$$

and

$$N(\varphi, B) := \inf_{f \in \mathcal{F}} \Big\{ \mathbb{E}_f[\varphi(\omega)] : |\Delta| f \leq B \Big\} \geq \nu - \frac{B}{4} \int_0^p |\varphi_\nu(x)| dx, \tag{4}$$

where $\varphi_\nu(x) := \varphi(x) - \nu$, as before.

Proof It follows immediately from Lemmas 3 and 4 that for every $f \in \mathcal{F}$ with $|\Delta| f \leq B$ its packed density f_p satisfies $|\Delta| f_p([0, p]) \leq B$ and $\mathbb{E}_{f_p}[\varphi(\omega)] = \mathbb{E}_f[\varphi(\omega)]$. Define $\overline{f_p} := \sup\{f_p(x) : x \in [0, p]\}$ and $\underline{f}_p := \inf\{f_p(x) : x \in [0, p]\}$. Since $f_p(0) = f_p(p)$ and $|\Delta| f_p([0, p]) \leq B$, it follows that $\overline{f_p} - \underline{f}_p \leq B/2$. Moreover, for all $x \in \mathbb{R}$ we have

$$\begin{aligned}\varphi(x) f_p(x) &= \nu f_p(x) + (\varphi(x) - \nu)^+ f_p(x) - (\varphi(x) - \nu)^- f_p(x) \\ &\leq \nu f_p(x) + (\varphi_\nu(x))^+ \overline{f_p} - (\varphi_\nu(x))^- \underline{f}_p.\end{aligned}$$

Hence,

$$\begin{aligned}\mathbb{E}_f[\varphi(\omega)] &= \mathbb{E}_{f_p}[\varphi(\omega)] \\ &= \int_0^p \varphi(x) f_p(x) dx \\ &\leq \int_0^p \nu f_p(x) dx + \int_0^p (\varphi_\nu(x))^+ \overline{f_p} dx - \int_0^p (\varphi_\nu(x))^- \underline{f}_p dx \\ &= \nu + (\overline{f_p} - \underline{f}_p) \int_0^p (\varphi_\nu(x))^+ dx \qquad (5) \\ &\leq \nu + \frac{B}{4} \int_0^p |\varphi_\nu(x)| dx. \qquad (6)\end{aligned}$$

Here, the equality in (5) holds since

$$\int_0^p (\varphi_\nu(x))^+ dx - \int_0^p (\varphi_\nu(x))^- dx = \int_0^p (\varphi_\nu(x)) dx = 0$$

by definition of ν, and (6) follows from the fact that $\overline{f_p} - \underline{f}_p \leq B/2$ and

$$\int_0^p (\varphi_\nu(x))^+ dx = \int_0^p (\varphi_\nu(x))^- dx = \frac{1}{2} \int_0^p |\varphi_\nu(x)| dx. \qquad (7)$$

Since the above holds for every pdf $f \in \mathcal{F}$ with $|\Delta| f \leq B$, we conclude that (3) holds. The bound for $N(\varphi, B)$ follows from the observations that $N(\varphi, B) = -M(-\varphi, B)$ by Lemma 1 (i), and that $\nu_{-\varphi} = -\nu_\varphi$ is the mean value of $-\varphi$. □

Theorem 1 *Let $\varphi : \mathbb{R} \mapsto \mathbb{R}$ be a periodic function with period p and finite mean value $\nu := p^{-1} \int_0^p \varphi(x) dx$. Then, for every $f \in \mathcal{F}$,*

$$\left| \mathbb{E}_f[\varphi(\omega)] - \nu \right| \leq \frac{|\Delta| f}{4} \int_0^p |\varphi_\nu(x)| dx.$$

Proof This result follows directly from Proposition 1, using the observation that every $f \in \mathcal{F}$ is a feasible solution of the optimization problem in $M(\varphi, |\Delta| f)$ and $N(\varphi, |\Delta| f)$. □

It follows from Theorem 1 that if the total variation of f is small then the expected value $\mathbb{E}_f[\varphi(\omega)]$ will be close to ν. For unimodal density functions such as the normal density function, the total variation decreases as the variance of the corresponding random variable ω increases. Hence, for normal random variables ω with a large variance, $\mathbb{E}_f[\varphi(\omega)]$ will be close to ν.

Clearly, the difference between $\mathbb{E}_f[\varphi(\omega)]$ and ν depends on both the pdf f and the periodic function φ. In the bound of Theorem 1 both factors are separated which shows that for every periodic function φ the expectation $\mathbb{E}_f[\varphi(\omega)]$ will be arbitrarily close to ν for pdf f with sufficiently small total variation.

In case φ is Lipschitz continuous on $[0, p)$ we can bound $\int_0^p |\varphi_\nu(x)|dx$, yielding bounds for $\mathbb{E}_f[\varphi(\omega)]$ not involving this possibly intractable integral:

Corollary 1 *Let $\varphi : \mathbb{R} \mapsto \mathbb{R}$ be a periodic function with period p and finite mean value $\nu := p^{-1} \int_0^p \varphi(x)dx$ such that φ is Lipschitz continuous with Lipschitz constant L on $(0, p)$. Then, for every random variable ω with pdf $f \in \mathcal{F}$,*

$$|\mathbb{E}_f[\varphi(\omega)] - \nu| \le \frac{Lp^2}{16}|\Delta|f.$$

The proof of Corollary 1 is based on the following lemma.

Lemma 5 *Let $\varphi : \mathbb{R} \mapsto \mathbb{R}$ be a periodic function with period p and finite mean value $\nu := p^{-1} \int_0^p \varphi(x)dx$, and absolute deviation $\delta := \int_0^p |\varphi_\nu(x)|dx$, such that φ is Lipschitz continuous with Lipschitz constant L on $(0, p)$. Then, there exists a function $\psi : \mathbb{R} \mapsto \mathbb{R}$ with the same properties, and such that ψ is* monotone non-decreasing *on $(0, p)$.*

Proof of Lemma 5 Basically, we construct ψ as a 'sorted version' of φ: on $(0, p)$, and similarly on $(kp, (k+1)p)$ for $k \in \mathbb{Z}$, the function values of φ are rearranged, from small to large. The main building block in this construction is the function $H : \mathbb{R} \mapsto \mathbb{R}$ defined as

$$H(y) := \mathcal{L}\Big(S(y)\Big), \quad y \in \mathbb{R},$$

where $\mathcal{L}$ denotes the Lebesgue measure, and

$$S(y) = \{u \in (0, p) : \varphi(u) \le y\}, \quad y \in \mathbb{R}.$$

Defining $\overline{y} := \sup_{u\in(0,p)} \varphi(u)$ and $\underline{y} := \inf_{u\in(0,p)} \varphi(u)$, the function H can be considered as a cdf of a random variable on $[\underline{y}, \overline{y}]$ with total probability mass p, since

(i) H is non-decreasing
(ii) $H(y) = 0$ for $y < \underline{y}$, and $H(y) = p$ for $y \ge \overline{y}$
(iii) H is right-continuous
(iv) H has a jump of size $d > 0$ at y if and only if $\mathcal{L}(\{u \in (0, p) : \varphi(u) = y\}) = d$
(v) if $\underline{y} < \overline{y}$, then H is strictly increasing on $[\underline{y}, \overline{y}]$
(vi) if $\underline{y} < \overline{y}$, then the right-derivative $H'_+(y)$ exists and satisfies $H'_+(y) \ge 1/L$ for $y \in [\underline{y}, \overline{y})$

We do not give a proof of these properties, since they can be verified easily.

The function H specifies a distribution on the function values of φ, and thus it can be used to derive properties of φ. For example,

$$\int_{-\infty}^{\infty} y dH(y) = \int_0^p \varphi(x)dx = \nu,$$

and

$$\int_{-\infty}^{\infty} |y-\nu| dH(y) = \int_0^p |\varphi(x)-\nu|dx.$$

Using the ordering of the function values of φ, given in H, we derive a *sorted version* ψ *of* φ:

$$\psi(x):=H^{-1}(x) := \inf\{y \in \mathbb{R} : H(y) \geq x\}, \quad x \in (0,p).$$

Using (i)–(vi) it is not hard to verify that ψ satisfies the desired properties.

Proof of Corollary 1 By applying Theorem 1 it suffices to show that

$$\int_0^p |\varphi_\nu(x)|dx \leq \frac{1}{4}Lp^2.$$

Consider $\psi : \mathbb{R} \mapsto \mathbb{R}$ that is periodic with period p, mean value ν and $\int_0^p |\psi_\nu(x)|dx = \int_0^p |\varphi_\nu(x)|dx$, and assume that ψ is also Lipschitz continuous with Lipschitz constant L and *monotone non-decreasing* on $(0,p)$. Such a function exists by Lemma 5. Moreover, since ψ satisfies these properties there exists $y \in (0,p)$ such that $\psi_\nu(y) = 0$, and $\psi_\nu(x) \leq L(x-y)$ for $x \geq y$ and $\psi_\nu(x) \geq L(x-y)$ for $x \leq y$, so that

$$\int_0^p (\psi_\nu(x))^+ dx \leq L \int_y^p (x-y)dx = \frac{1}{2}L(p-y)^2, \tag{8}$$

and

$$\int_0^p (\psi_\nu(x))^- dx \leq -L \int_0^y (x-y)dx = \frac{1}{2}Ly^2. \tag{9}$$

Combining (8) and (9) with (7) yields

$$\int_0^p |\psi_\nu(x)|dx \leq \min\left\{L(p-y)^2, Ly^2\right\} \leq \frac{1}{4}Lp^2.$$

The claim follows since $\int_0^p |\psi_\nu(x)|dx = \int_0^p |\varphi_\nu(x)|dx$. □

3 Periodically monotone functions

In this section we restrict our attention to functions φ that are periodically monotone.

Definition 5 Let $\varphi : \mathbb{R} \mapsto \mathbb{R}$ be a periodic function with period p. Then, φ is *periodically monotone* if there exists $\beta \in \mathbb{R}$ such that φ is either non-increasing or non-decreasing on $(\beta, \beta + p)$. We say that φ is *periodically non-increasing*, or *periodically non-decreasing*, respectively.

Under this additional assumption we are able to derive exact expressions for $M(\varphi, B)$ and $N(\varphi, B)$, and thus tighter bounds on $\mathbb{E}_f[\varphi(\omega)]$ than (3) and (4). In order to derive these expressions we use the concept of *flattening of density functions*, first introduced in [27]. Observing that a constant function has lower total variation than a varying one, we can restrict the optimization in $M(\varphi, B)$ and $N(\varphi, B)$ to *piecewise constant* density functions, allowing to find sharp bounds.

In Sect. 3.1 we first discuss two alternative ways to flatten bounded monotone non-decreasing functions. The results will be used to flatten density functions in Sect. 3.2. We derive exact expressions for $M(\varphi, B)$ and $N(\varphi, B)$ in Sect. 3.3, and we discuss an extension of these results in Sect. 3.4, which is used to derive error bounds for recourse approximations in Sect. 4.

3.1 Flattening of monotone functions

In this subsection we only consider bounded non-decreasing functions, bearing in mind that every pdf $f \in \mathcal{F}$ can be written as the difference of such functions. We show two alternative ways to flatten non-decreasing functions in Lemmas 6 and 7, respectively.

Lemma 6 *Let $f : \mathbb{R} \mapsto \mathbb{R}$ be a bounded monotone non-decreasing function, and let $I \subset \mathbb{R}$ denote a bounded interval of positive length. Define the function $g : \mathbb{R} \mapsto \mathbb{R}$ as*

$$g(x) = \begin{cases} f(x), & x \notin I, \\ K_I, & x \in I, \end{cases} \tag{10}$$

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

(i) $\int_I g(x)dx = \int_I f(x)dx$,
(ii) *g is a bounded monotone non-decreasing function, and*
(iii) $|\Delta|g = |\Delta|f$.

Proof Equation (i) follows from the definition of K_I. Moreover, it is obvious that g is bounded since f is bounded. To show that g is non-decreasing, let $x_1, x_2 \in \mathbb{R}$ with $x_1 < x_2$ be given. Then, using that f is non-decreasing,

(1) if $x_1 \notin I$ and $x_2 \notin I$, then $g(x_1) = f(x_1) \leq f(x_2) = g(x_2)$.
(2) if $x_1 \notin I$ and $x_2 \in I$, then $x_1 < y$ for all $y \in I$ so that

$$g(x_1) = f(x_1) = |I|^{-1} \int_I f(x_1)dy \leq |I|^{-1} \int_I f(y)dy = K_I = g(x_2).$$

(3) if $x_1 \in I$ and $x_2 \notin I$, then $x_2 > y$ for all $y \in I$ so that

$$g(x_1) = K_I = |I|^{-1} \int_I f(y)dy \leq |I|^{-1} \int_I f(x_2)dy = f(x_2) = g(x_2).$$

(4) if $x_1 \in I$ and $x_2 \in I$, then $g(x_1) = K_I = g(x_2)$.

Thus, (ii) g is a bounded monotone non-decreasing function. Hence,

$$|\Delta|g = \sup g(\mathbb{R}) - \inf g(\mathbb{R}),$$

and since $\sup g(\mathbb{R}) = \sup f(\mathbb{R})$ and $\inf g(\mathbb{R}) = \inf f(\mathbb{R})$, we conclude that (iii) $|\Delta|g = |\Delta|f$. □

Lemma 7 *Let $f : \mathbb{R} \mapsto \mathbb{R}$ be a bounded monotone non-decreasing function, and let $I \subset \mathbb{R}$ denote a bounded interval of positive length. Write $\overline{f}_I := \sup\{f(x) : x \in I\}$ and $\underline{f}_I := \inf\{f(x) : x \in I\}$. Then, there exists $z \in I$ such that the function $g_z : \mathbb{R} \to \mathbb{R}$ defined as*

$$g_z(x) = \begin{cases} f(x), & x \notin I, \\ \overline{f}_I, & x \in I \text{ and } x \geq z, \\ \underline{f}_I, & x \in I \text{ and } x < z, \end{cases}$$

satisfies

(i) $\int_I g_z(x)dx = \int_I f(x)dx$,
(ii) *g_z is a bounded monotone non-decreasing function, and*
(iii) $|\Delta|g_z = |\Delta|f$.

In fact, for all $z \in I$ both (ii) and (iii) hold.

Proof It is easy to observe that g_z is bounded for every $z \in I$. Moreover, for every $x_1, x_2 \in \mathbb{R}$ with $x_1 < x_2$ and for every $z \in \mathbb{R}$ we have

(1) if $x_1 \notin I$ and $x_2 \notin I$, then $g_z(x_1) = f(x_1) \leq f(x_2) = g_z(x_2)$.
(2) if $x_1 \notin I$ and $x_2 \in I$, then $x_1 < y$ for all $y \in I$ so that

$$g_z(x_1) = f(x_1) \leq \inf\{f(y) : y \in I\} = \underline{f}_I \leq g_z(x_2).$$

(3) if $x_1 \in I$ and $x_2 \notin I$, then $x_2 > y$ for all $y \in I$ so that

$$g_z(x_1) \leq \overline{f}_I = \sup\{f(y) : y \in I\} \leq f(x_2) = g_z(x_2).$$

(4) if $x_1 \in I$ and $x_2 \in I$, then $g_z(x_1) \leq g_z(x_2)$ since g_z is non-decreasing on I.

Thus, for every $z \in I$, the proof of (ii) is complete, and using similar arguments as in the proof of Lemma 6, (iii) is true, too.

In order to show (i), define $D(z) = \int_I g_z(x)dx$ for every $z \in I$, and observe that D is linear hence continuous on I. Moreover, $\inf D(I) = |I|\underline{f}_I$ and $\sup D(I) = |I|\overline{f}_I$. Since

$$|I|\underline{f}_I = \int_I \underline{f}_I dx \leq \int_I f(x)dx \leq \int_I \overline{f}_I dx = |I|\overline{f}_I, \tag{11}$$

it follows from the intermediate value theorem that if both inequalities in (11) are strict, then there exists $z^* \in I$ such that

$$D(z^*) = \int_I f(x)dx,$$

and thus g_{z^*} satisfies (i)–(iii). It is not difficult to verify that this conclusion also holds if at least one inequality in (11) is an equality. □

3.2 Flattening of density functions

Since every pdf $f \in \mathcal{F}$ can be written as the difference of two bounded non-decreasing functions f_1 and f_2, the results in Lemmas 6 and 7 can be used to flatten density functions as well. For example, applying Lemma 6 to both f_1 and f_2 yields Lemma 1 in [27]. This lemma is stated here without proof.

Lemma 8 *Let $f \in \mathcal{F}$ be given and let $I \subset \mathbb{R}$ denote a bounded interval with positive length. Define $g \in \mathcal{F}$ as*

$$g(x) = \begin{cases} f(x), & x \notin I \\ K_I, & x \in I, \end{cases} \tag{12}$$

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

The next lemma is derived with expectations $\mathbb{E}_f[\varphi(\omega)]$ of periodically non-increasing functions φ in mind. It is used to show that for such functions φ, the optimization in $M(\varphi, B)$ can be restricted to piecewise constant density functions. The main idea is that given any feasible $f \in \mathcal{F}$, which can be written as $f = f_1 - f_2$, we apply Lemma 6 to f_1 yielding g_1 and Lemma 7 to f_2 yielding g_2 so that $g \in \mathcal{F}$ defined as $g := g_1 - g_2$ is feasible in $M(\varphi, B)$ and has an objective value at least as large as f. Similar results can be obtained independently for non-decreasing functions, but we will derive them directly from the non-increasing case instead.

Lemma 9 *Let $I \subset \mathbb{R}$ be a bounded interval of positive length, and let $\varphi : \mathbb{R} \mapsto \mathbb{R}$ be a real-valued function that is non-increasing on I. Then, for every $f \in \mathcal{F}$ there exists $g \in \mathcal{F}$ such that*

(i) *g is non-increasing piecewise constant and at most two-valued on I,*
(ii) *$\int_I g(x)dx = \int_I f(x)dx$,*
(iii) *$|\Delta|g \leq |\Delta|f$, and*
(iv) *$\mathbb{E}_g[\varphi(\omega)] \geq \mathbb{E}_f[\varphi(\omega)]$.*

Proof Let $f \in \mathcal{F}$ be given. Since f is a pdf of bounded variation, there exist bounded monotone non-decreasing functions f_1 and f_2 such that $f = f_1 - f_2$ and $|\Delta|f_1 =$

$|\Delta| f_2 = \frac{1}{2}|\Delta| f$. Apply Lemma 6 to f_1 and Lemma 7 to f_2 to obtain g_1 and g_2, respectively, defined for every $x \in \mathbb{R}$ as

$$g_1(x) = \begin{cases} f_1(x), & x \notin I \\ K_I^1, & x \in I, \end{cases}$$

with $K_I^1 := |I|^{-1} \int_I f_1(u)du$, and

$$g_2(x) = \begin{cases} f_2(x), & x \notin I, \\ \overline{f}_I^2, & x \in I \text{ and } x \geq z, \\ \underline{f}_I^2, & x \in I \text{ and } x < z, \end{cases}$$

where $\overline{f}_I^2 := \sup\{f_2(x) : x \in I\}$, $\underline{f}_I^2 := \inf\{f_2(x) : x \in I\}$, and $z \in I$ is chosen such that $\int_I g_2(x)dx = \int_I f_2(x)dx$. We will show that $g := g_1 - g_2$ satisfies (i)–(iv).

First of all observe that Lemmas 6 and 7 imply that for $j = 1, 2$, g_j is a bounded non-decreasing function with $|\Delta| g_j = |\Delta| f_j$ and $\int_I g_j(x)dx = \int_I f_j(x)dx$. Since g is the difference of two bounded monotone non-decreasing functions it follows that g is of bounded variation. Moreover, since

$$(iii) \int_I g(x)dx = \int_I g_1(x)dx - \int_I g_2(x)dx = \int_I f_1(x)dx - \int_I f_2(x)dx = \int_I f(x)dx,$$

holds for all bounded I and $\int_{-\infty}^{\infty} g(x)dx = \int_{-\infty}^{\infty} f(x)dx = 1$, we have $g \in \mathcal{F}$. Furthermore, since $|\Delta| g_1 = \frac{1}{2}|\Delta| f$ and $|\Delta| g_2 = \frac{1}{2}|\Delta| f$, we conclude that (ii) $|\Delta| g \leq |\Delta| g_1 + |\Delta| g_2 = |\Delta| f$, and by definition of g_1 and g_2 it follows immediately that (i) g is non-increasing piecewise constant and at most two-valued on I. Finally, we will show that $\int_I \varphi(x)g_1(x)dx \geq \int_I \varphi(x) f_1(x)dx$ and $\int_I \varphi(x)g_2(x)dx \leq \int_I \varphi(x) f_2(x)dx$ so that $\int_I \varphi(x)g(x)dx \geq \int_I \varphi(x) f(x)dx$. Then, (iv) $\mathbb{E}_g[\varphi(\omega)] \geq \mathbb{E}_f[\varphi(\omega)]$ follows from the above and the observation that $g(x) = f(x)$ for $x \notin I$.

First consider g_1 and f_1. Observing that f_1 is non-decreasing on I, g_1 is constant on I and $\int_I g_1(x)dx = \int_I f_1(x)dx$ [by Lemma 6 (i)], it follows that there exists $y \in I$ such that for every $x \in I$,

(A) $g_1(x) \geq f_1(x)$ and $\varphi(x) \geq \varphi(y)$ if $x < y$, and
(B) $g_1(x) \leq f_1(x)$ and $\varphi(x) \leq \varphi(y)$ if $x > y$.

Here we use that φ is non-increasing on I, too. The inequalities in (A) and (B) imply that

$$\Big(\varphi(x) - \varphi(y)\Big)\Big(g_1(x) - f_1(x)\Big) \geq 0 \quad \text{for all } x \in I.$$

Hence,

$$\int_I \Big(\varphi(x) - \varphi(y)\Big)\Big(g_1(x) - f_1(x)\Big)dx \geq 0,$$

and $\int_I \varphi(x)g_1(x)dx \geq \int_I \varphi(x)f_1(x)dx$ follows immediately since $\varphi(y)\int_I(g_1(x) - f_1(x))dx = 0$ by Lemma 6 (i).

Next consider g_2 and f_2. By definition of g_2 and since φ is non-increasing on I, it follows that for all $x \in I$, and with z as defined above,

(A) $g_2(x) \leq f_2(x)$ and $\varphi(x) \geq \varphi(z)$ if $x < z$, and
(B) $g_2(x) \geq f_2(x)$ and $\varphi(x) \leq \varphi(z)$ if $x > z$.

Hence,

$$\Big(\varphi(x) - \varphi(z)\Big)\Big(g_2(x) - f_2(x)\Big) \leq 0 \quad \text{for all } x \in I.$$

This implies that

$$\int_I \Big(\varphi(x) - \varphi(z)\Big)\Big(g_2(x) - f_2(x)\Big)dx \leq 0,$$

and thus $\int_I \varphi(x)g_2(x)dx \leq \int_I \varphi(x)f_2(x)dx$, completing the proof. □

Remark 3 Conversely, applying Lemmas 7 to f_1 and Lemma 6 to f_2 yields a pdf $g \in \mathcal{F}$ that is *non-decreasing* piecewise constant and at most two-valued on I, satisfying (ii) and (iii), and (iv) with the reverse inequality sign.

3.3 Exact worst-case bounds for periodically monotone functions

We first restrict the analysis to the upper bound $M(\varphi, B)$ for functions φ that are periodically non-increasing, as in Sect. 3.2. Results for the lower bound $N(\varphi, B)$ and periodically non-decreasing φ can easily be derived from this single case, see Corollary 2. Building on Lemma 9 we show that it suffices to consider a class of two-valued piecewise constant packed densities in the optimization of $M(\varphi, B)$, so that the remaining optimization problem is straightforward to solve.

Lemma 10 *Let $\varphi : \mathbb{R} \mapsto \mathbb{R}$ be a periodic function with period p and finite mean value $\nu := p^{-1}\int_0^p \varphi(x)dx$. Assume that φ is non-increasing on* $(0, p)$. *Then, for every $f \in \mathcal{F}$ there exists a packed density $g_p \in \mathcal{F}$ such that*

(i) *g_p is non-increasing piecewise constant and at most two-valued on* $(0, p)$,
(ii) *$|\Delta|g_p([0, p]) \leq |\Delta|f$, and*
(iii) *$\mathbb{E}_{g_p}[\varphi(\omega)] \geq \mathbb{E}_f[\varphi(\omega)]$.*

Proof Let $f \in \mathcal{F}$ be given and consider the packed density f_p of f with support contained in $[0, p]$. From Lemma 4 it follows that $|\Delta|f_p([0, p]) \leq |\Delta|f$, so that $f_p \in \mathcal{F}$, and from Lemma 3 it follows that $\mathbb{E}_{f_p}[\varphi(\omega)] = \mathbb{E}_f[\varphi(\omega)]$. Now apply Lemma 9 to f_p with $I = (0, p)$ yielding g_p, satisfying (i) by construction and (iii) since $\mathbb{E}_{g_p}[\varphi(\omega)] \geq \mathbb{E}_{f_p}[\varphi(\omega)] = \mathbb{E}_f[\varphi(\omega)]$. It remains to prove (ii) using $|\Delta|g_p \leq |\Delta|f_p$, which holds by Lemma 9 (iii). The inequality in (ii) does not follow immediately since $|\Delta| f_p$ may actually be much larger than $|\Delta| f$, but it holds since for any packed density h_p we have

$$\begin{aligned}|\Delta| h_p &= |\Delta| h_p((-\infty, 0]) + |\Delta| h_p([0, p]) + |\Delta| h_p([p, \infty)) \\ &= h_p(0) + |\Delta| h_p([0, p]) + h_p(p),\end{aligned}$$

and moreover, by construction $g_p(0) = f_p(0) = g_p(p) = f_p(p)$ so that

$$\begin{aligned}(ii)\ \ |\Delta| g_p([0, p]) &= |\Delta| g_p - g_p(0) - g_p(p) \\ &\leq |\Delta| f_p - f_p(0) - f_p(p) \\ &= |\Delta| f_p([0, p]) \\ &\leq |\Delta| f.\end{aligned}$$

□

Now we are ready to prove an exact expression for $M(\varphi, B)$ for periodically non-increasing functions φ.

Proposition 2 *Let $\varphi : \mathbb{R} \mapsto \mathbb{R}$ be a periodic function with period p and finite mean value $\nu = p^{-1} \int_0^p \varphi(x)dx$. Assume that φ is non-increasing on $(\beta, \beta + p)$ for some $\beta \in \mathbb{R}$. Then, for every $B > 0$,*

$$M(\varphi, B) = \nu + \frac{B}{2} \int_{\beta}^{\beta+\min\{p,2/B\}} (\varphi_\nu(x))^+ dx. \tag{13}$$

Proof We will derive an expression for $M(\hat{\varphi}, B)$ for $\hat{\varphi} : \mathbb{R} \mapsto \mathbb{R}$ defined as $\hat{\varphi}(x) = \varphi(x - \beta)$, $x \in \mathbb{R}$, so that $\hat{\varphi}$ is non-increasing on $(0, p)$. Since $M(\varphi, B) = M(\hat{\varphi}, B)$ by Lemma 1 (v), the desired result for $M(\varphi, B)$ follows by straightforward transformation.

Due to Lemma 10 we can restrict the optimization in $M(\hat{\varphi}, B)$ to packed densities g_p of the form

$$g_p(x) = \begin{cases} g_0, & x = 0 \text{ or } x = p \\ g^+, & 0 < x < z, \\ g^-, & z \leq x < p, \\ 0, & \text{otherwise,} \end{cases}$$

for some $z \in (0, p)$ with $g^+ \geq g^- \geq 0$. By definition of ν it follows immediately that $\int_z^p \hat{\varphi}_\nu(x)dx = -\int_0^z \hat{\varphi}_\nu(x)dx$, and thus

$$\begin{aligned}\mathbb{E}_{g_p}[\hat{\varphi}(\omega)] &= \nu + \int_0^p (\hat{\varphi}(x) - \nu) g_p(x)dx \\ &= \nu + g^+ \int_0^z (\hat{\varphi}(x) - \nu)dx + g^- \int_z^p (\hat{\varphi}(x) - \nu)dx \\ &= \nu + (g^+ - g^-) \int_0^z \hat{\varphi}_\nu(x)dx.\end{aligned}$$

Moreover, since

$$|\Delta| g_p([0, p]) = |g^+ - g_0| + (g^+ - g^-) + |g^- - g_0|,$$

it follows that

$$\min_{g_0 \in \mathbb{R}_+} |\Delta| g_p([0, p]) = 2(g^+ - g^-)$$

where the minimum is attained by any $g_0 \in [g^-, g^+]$. Since the variable g_0 does only occur in the expression for $|\Delta| g_p([0, p])$, we assume without loss that g_0 attains such a minimizing value. Hence, the optimization problem that we have to solve reads

$$\begin{aligned} M(\varphi, B) = \nu + \sup_{g^+, g^-, z} \; & (g^+ - g^-) \int_0^z \hat{\varphi}_\nu(x) dx \\ \text{s.t. } & z g^+ + (p - z) g^- = 1 && (14) \\ & g^+ - g^- \leq B/2 && (15) \\ & g^+ \geq g^- \geq 0, \end{aligned}$$

where $g^+ \geq 0$, $g^- \geq 0$, and (14) guarantee that g_p is a pdf and (15) captures the bound on $|\Delta| g_p[0, p]$.

In order to solve this problem, first fix $z \in (0, p)$, and consider the resulting optimization problem with optimal value denoted by $M_z(\hat{\varphi}, B)$. Note that any feasible solution with $g^+ - g^- = B/2$ is optimal yielding $M_z(\hat{\varphi}, B) = \nu + B/2 \int_0^z \hat{\varphi}_\nu(x) dx$. Such a feasible solution exists if (14) holds, so if $z \leq 2/B$. At the same time, since (14) can be rewritten as $g^+ - g^- = (1 - p g^-)/z$ and by substituting this equality in the objective, the solution $g^- = 0$ and $g^+ = 1/z$ is optimal with $M_z(\hat{\varphi}, B) = \nu + z^{-1} \int_0^z \hat{\varphi}_\nu(x) dx$ if it is feasible. This is the case if (15) holds, that is, if $z \geq 2/B$. Hence, for every $z \in (0, p)$,

$$M_z(\hat{\varphi}, B) = \begin{cases} \nu + \dfrac{B}{2} \displaystyle\int_0^z \hat{\varphi}_\nu(x) dx, & z \leq \dfrac{2}{B}, \\ \nu + \dfrac{1}{z} \displaystyle\int_0^z \hat{\varphi}_\nu(x) dx, & z \geq \dfrac{2}{B}, \end{cases} \tag{16}$$

where for $z = 2/B$ both formulas give the same result, of course. Now we will calculate

$$M(\hat{\varphi}, B) = \sup_z \left\{ M_z(\hat{\varphi}, B) : z \in (0, p) \right\}. \tag{17}$$

First consider the second case in (16). Since $\hat{\varphi}$, and thus $\hat{\varphi}_\nu$, is non-increasing on $(0, p)$, it follows that the mean value $1/z \int_0^z \hat{\varphi}_\nu(x) dx$ of $\hat{\varphi}_\nu$ over $(0, z)$ is non-increasing in z on $(0, p)$, so that $z^* \leq 2/B$, where z^* is the optimal solution in (17). So, without loss we may restrict ourselves to the first case in (16). Its objective is concave in z on $[0, p]$, non-decreasing for small values of z and non-increasing for large values of z, since its right derivative is $\frac{B}{2}\hat{\varphi}_\nu(z)$ which is non-increasing with mean value $\int_0^p \hat{\varphi}_\nu(z)dz = 0$. Hence,

$$z^* = \min\{x_0, 2/B\},$$

where x_0 is any global (i.e. on $(0, p)$) maximizer of the objective in the first case, characterized by $\hat{\varphi}_\nu(x) \geq 0$ for $x \in (0, x_0)$ and $\hat{\varphi}_\nu(x) \leq 0$ for $x \in (x_0, p)$. We conclude that

$$\begin{aligned} M(\hat{\varphi}, B) = M_{z^*}(\hat{\varphi}, B) &= \nu + \frac{B}{2}\int_0^{\min\{x_0,2/B\}} \hat{\varphi}_\nu(x)dx \\ &= \nu + \frac{B}{2}\int_0^{\min\{p,2/B\}} (\hat{\varphi}_\nu(x))^+ dx, \end{aligned}$$

where the last equality is true since

$$(\hat{\varphi}_\nu(x))^+ = \begin{cases} \hat{\varphi}_\nu(x), & 0 < x < x_0, \\ 0, & x_0 < x < p. \end{cases}$$

□

Corollary 2 *Let $\varphi : \mathbb{R} \mapsto \mathbb{R}$ be a periodic function with period p and finite mean value $\nu := p^{-1} \int_0^p \varphi(x)dx$. Then,*

(i) *if φ is non-increasing on $(\beta, \beta + p)$ for some $\beta \in \mathbb{R}$, then*

$$M(\varphi, B) = \nu + \frac{B}{2}\int_\beta^{\beta+\min\{p,2/B\}} (\varphi_\nu(x))^+ dx, \tag{18}$$

and

$$N(\varphi, B) = \nu - \frac{B}{2}\int_{\beta+p-\min\{p,2/B\}}^{\beta+p} (\varphi_\nu(x))^- dx. \tag{19}$$

(ii) *if φ is non-decreasing on $(\beta, \beta + p)$ for some $\beta \in \mathbb{R}$, then*

$$M(\varphi, B) = \nu + \frac{B}{2}\int_{\beta+p-\min\{p,2/B\}}^{\beta+p} (\varphi_\nu(x))^+ dx, \tag{20}$$

and

$$N(\varphi, B) = \nu - \frac{B}{2}\int_\beta^{\beta+\min\{p,2/B\}} (\varphi_\nu(x))^- dx. \tag{21}$$

Proof For completeness (18) repeats Proposition 2. Next, continuing the analysis of $M(\varphi, B)$, suppose that φ is non-decreasing on $(\beta, \beta + p)$. Then, $\bar{\varphi} : \mathbb{R} \mapsto \mathbb{R}$ defined as $\bar{\varphi}(x) = \varphi(-x)$ is periodic with period p and mean value $\nu_{\bar{\varphi}} = \nu_\varphi$, and $\bar{\varphi}$ is *non-increasing* on $(-\beta, p - \beta)$. Thus, using (18) we have

$$M(\bar{\varphi}, B) = \nu + \frac{B}{2}\int_{-\beta}^{-\beta+\min\{p,2/B\}} (\bar{\varphi}_\nu(x))^+ dx.$$

Since $M(\varphi, B) = M(\bar{\varphi}, B)$ by Lemma 1 (iv), simple computation using $\bar{\varphi}(x) = \varphi(-x)$ yields (20).

The expressions for $N(\varphi, B)$ in (19) and (21) are obtained using $N(\varphi, B) = -M(\varphi, B)$, which holds by Lemma 1 (i), and the observation that if φ is non-increasing (non-decreasing) on $(\beta, \beta+p)$, then $-\varphi$ is non-decreasing (non-increasing) on $(\beta, \beta + p)$. Hence, (19) and (21) can be obtained by applying (20) and (18) to $-M(-\varphi, B)$, respectively. □

Below we give two examples of periodically monotone functions φ and corresponding expressions for $M(\varphi, B)$ and $N(\varphi, B)$. Both examples will be used in Sect. 4 to derive error bounds for approximations of recourse models.

Example 1 Let $\alpha, \beta \in \mathbb{R}$ be given, and consider the function $\varphi_{\alpha,\beta} : \mathbb{R} \mapsto \mathbb{R}$ defined as

$$\varphi_{\alpha,\beta}(x) = \lceil x \rceil_\alpha - \lceil x \rceil_\beta := \lceil x - \alpha \rceil + \alpha - \lceil x - \beta \rceil - \beta, \quad x \in \mathbb{R}, \tag{22}$$

see also Fig. 1. This class of functions has been studied in [27] to derive a uniform error bound for a class of convex approximations of TU integer recourse models. In Lemma 4 of this reference we show that for every $\alpha, \beta \in \mathbb{R}$,

(i) $\varphi_{\alpha,\beta}$ is periodic with period $p = 1$ and mean value $\nu = 0$.
(ii) If $\alpha - \beta \in \mathbb{Z}$, then $\varphi_{\alpha,\beta} \equiv 0$.
(iii) If $\alpha - \beta \notin \mathbb{Z}$, then $\varphi_{\alpha,\beta}$ is two-valued piecewise constant with

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

with $\gamma_{\alpha,\beta} := \alpha + 1 - \lceil \alpha \rceil_\beta \in (0, 1)$.

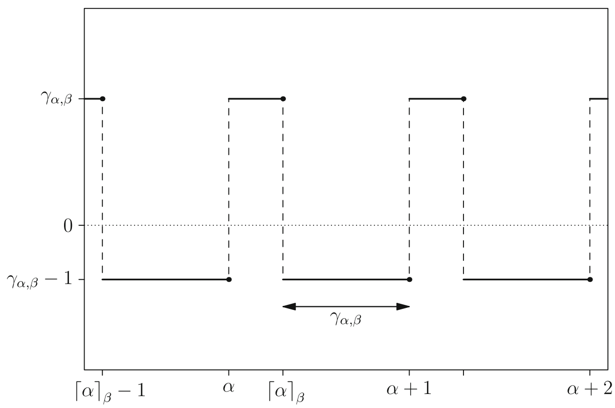

Fig. 1 The periodic function $\varphi_{\alpha,\beta}$ defined in (22)

Since $\varphi_{\alpha,\beta}$ is non-increasing on $(\alpha, \alpha+1)$, using (18) in Corollary 2 gives

$$M(\varphi_{\alpha,\beta}, B) = \frac{B}{2}\int_{\alpha}^{\alpha+\min\{1,2/B\}} (\varphi_{\alpha,\beta}(x))^{+} dx.$$

After some calculations it follows that

$$M(\varphi_{\alpha,\beta}, B) = \min\left\{\gamma_{\alpha,\beta}, \gamma_{\alpha,\beta}(1-\gamma_{\alpha,\beta})\frac{B}{2}\right\}, \tag{23}$$

and

$$N(\varphi_{\alpha,\beta}, B) = -\min\left\{1-\gamma_{\alpha,\beta}, \gamma_{\alpha,\beta}(1-\gamma_{\alpha,\beta})\frac{B}{2}\right\}.$$

Furthermore, $M(\varphi_{\alpha,\beta}, B)$ and $N(\varphi_{\alpha,\beta}, B)$ can be maximized and minimized, respectively, over $\alpha, \beta \in \mathbb{R}$ for every $B > 0$, yielding

$$\sup_{\alpha,\beta} M(\varphi_{\alpha,\beta}, B) = -\inf_{\alpha,\beta} N(\varphi_{\alpha,\beta}, B) = h(B), \tag{24}$$

where $h : \mathbb{R}_{++} \mapsto \mathbb{R}$ is defined as

$$h(x) = \begin{cases} x/8, & 0 < x \le 4, \\ 1 - 2/x, & x \ge 4. \end{cases} \tag{25}$$

Example 2 Let $\alpha \in \mathbb{R}$ be given, and consider the function $\bar{\varphi}_\alpha : \mathbb{R} \mapsto \mathbb{R}$ defined as

$$\bar{\varphi}_\alpha(x) := x + 1/2 - \lceil x \rceil_\alpha = x + 1/2 - \lceil x - \alpha \rceil - \alpha, \quad x \in \mathbb{R}. \tag{26}$$

For every $\alpha \in \mathbb{R}$, this function is periodic with period $p = 1$ and mean value $\nu = 0$, see also Fig. 2. Moreover, $\bar{\varphi}_\alpha(\mathbb{R}) = [-1/2, 1/2)$, and $\bar{\varphi}_\alpha$ is non-decreasing on $(\alpha, \alpha+1)$. Since

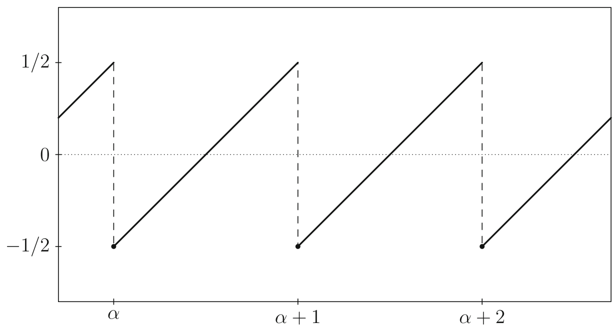

Fig. 2 The periodic function $\bar{\varphi}_\alpha$ defined in (26)

$$\int_{\alpha+1-y}^{\alpha+1} \left(\bar{\varphi}_\alpha(x)\right)^+ dx = \begin{cases} \frac{1}{2}y(1-y), & 0 \le y \le 1/2, \\ \frac{1}{8}, & 1/2 \le y \le 1, \end{cases}$$

it follows from (20) in Corollary 2, and using $y = \min\{1, 2/B\}$, that

$$M(\bar{\varphi}_\alpha, B) = \frac{1}{2}h(B) = \begin{cases} B/16, & 0 < B \le 4, \\ 1/2 - 1/B, & B \ge 4. \end{cases} \tag{27}$$

By symmetry we have $N(\bar{\varphi}_\alpha, B) = -M(\bar{\varphi}_\alpha, B)$.

3.4 Periodic functions with monotone amplitude

In this subsection we consider an extension of the results obtained so far. This extension is necessary to derive error bounds of recourse approximations in Sect. 4. Instead of considering $\mathbb{E}_f[\varphi(\omega)]$, we consider $\mathbb{E}_f[\lambda(\omega)\varphi(\omega)]$, where $\lambda : \mathbb{R} \mapsto \mathbb{R}$ is a non-negative bounded monotone function. The function $\lambda\varphi$ can be interpreted as a function that is periodic in shape but has non-decreasing or non-increasing amplitude.

We show in Proposition 3 that for periodic functions with zero mean value ν, we can derive bounds for $\mathbb{E}_f[\lambda(\omega)\varphi(\omega)]$ by rounding λ to its supremum, denoted λ^*. This is not a trivial result since φ attains both positive and negative values.

Proposition 3 *Let $\lambda : \mathbb{R} \mapsto \mathbb{R}$ be a real-valued monotone function such that $0 \le \lambda(x) \le \lambda^*$ for all $x \in \mathbb{R}$, and let $\varphi : \mathbb{R} \mapsto \mathbb{R}$ be a bounded periodic function with period p and finite mean value $\nu := p^{-1}\int_0^p \varphi(x)dx$. Then, for every $f \in \mathcal{F}$,*

$$\lambda^* N(\varphi_\nu, |\Delta| f) \le \mathbb{E}_f[\lambda(\omega)\varphi_\nu(w)] \le \lambda^* M(\varphi_\nu, |\Delta| f). \tag{28}$$

Proof First suppose that λ is monotone non-decreasing and define $L := \int_{-\infty}^{\infty} \lambda(x) f(x) dx$. Then, $0 \le L \le \lambda^*$. If $L = 0$, then $\mathbb{E}_f[\lambda(\omega)\varphi_\nu(\omega)] = 0$, and (28) holds trivially. If $L > 0$, then define $g : \mathbb{R} \mapsto \mathbb{R}$ as $g := L^{-1}\lambda f$. Observe that g is non-negative since both λ and f are non-negative, and that g integrates to 1 by definition of L, so that g is a pdf. Moreover,

$$\begin{aligned} \mathbb{E}_f[\lambda(\omega)\varphi_\nu(\omega)] &= \int_{-\infty}^{\infty} \lambda(x)\varphi_\nu(x) f(x) dx \\ &= L \int_{-\infty}^{\infty} \varphi_\nu(x) g(x) dx \\ &= L\mathbb{E}_g[\varphi_\nu(\omega)]. \end{aligned}$$

Since $f \in \mathcal{F}$, there exist bounded non-negative monotone non-decreasing functions f_1 and f_2 such that $f = f_1 - f_2$, $|\Delta| f_1 = |\Delta| f_2 = \frac{1}{2}|\Delta| f$, and

$$\lim_{x \downarrow -\infty} f_1(x) = \lim_{x \downarrow -\infty} f_2(x) = 0.$$

Then, $g_1 := L^{-1}\lambda f_1$ and $g_2 := L^{-1}\lambda f_2$ are both bounded monotone non-decreasing with $|\Delta| g_1 = |\Delta| g_2 \leq \frac{1}{2} L^{-1}\lambda^* |\Delta| f$, so that for $g = g_1 - g_2$, we have $|\Delta| g \leq L^{-1}\lambda^* |\Delta| f$. We conclude that $g \in \mathcal{F}$, and

$$LN(\varphi_\nu, L^{-1}\lambda^*|\Delta|f) \leq \mathbb{E}_f[\lambda(\omega)\varphi_\nu(\omega)] \leq LM(\varphi_\nu, L^{-1}\lambda^*|\Delta|f). \tag{29}$$

Since $M(\varphi, B)$ is concave in B by Lemma 2, and $\lim_{B\downarrow 0} M(\varphi_\nu, B) = 0$, it follows that $M(\varphi_\nu, t|\Delta|f)$ as a function of t lies below the line through (0, 0) and $(1, M(\varphi, |\Delta|f))$ for $t \geq 1$. That is, for $t \geq 1$ we have $M(\varphi_\nu, t|\Delta|f) \leq tM(\varphi_\nu, |\Delta|f)$. Similarly, using the convexity of $N(\varphi, B)$ we can show that $tN(\varphi_\nu, |\Delta|f) \leq N(\varphi_\nu, t|\Delta|f)$ for $t \geq 1$. Substituting these inequalities into (29) with $t = L^{-1}\lambda^* \geq 1$ yields the desired bounds.

If λ is monotone non-increasing, then the result can be obtained from the non-decreasing case since $\hat{\lambda}(x) := \lambda(-x)$ is non-decreasing and $\hat{f}(x) := f(-x)$ satisfies $|\Delta|\hat{f} = |\Delta| f$. □

Remark 4 The bounds in (29) are actually tighter than the bounds in (28). However, in order to derive error bounds for recourse approximations in Sect. 4 we will use the bounds in (28) since we are not able to compute L in that setting.

4 Approximations of two-stage recourse models

In previous sections we derived total variation bounds for the expectation of periodic functions. The results can be applied in problems involving both uncertainty and periodicity. Here, we apply the total variation bounds to approximation of two-stage recourse models. The presence of uncertainty is inherent in these models, whereas periodicity will appear as a result of rounding.

Recourse models can be used for many practical problems that deal with decision making under uncertainty, including problems in energy, logistics, and finance (see, e.g., [12,35]). We consider the two-stage recourse model with random right-hand side (only) defined as

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

where $z \in \mathbb{R}^m$ are so-called tender variables, and

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

with

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

Here, Q and v are called the (expected) recourse function and second-stage value function, respectively. We assume that the distribution of the random vector ω is known.

In this model, the first-stage decisions x have to be made here-and-now before the realization of the random vector ω is observed. After the realization of ω is revealed, we are allowed to take so-called recourse actions y to compensate for possible infeasibilities of the random constraint $Tx \geq \omega$. The second-stage value function v specifies the possible recourse actions and their cost. The objective (and challenge) is to find feasible first-stage decisions x as to minimize the costs cx incurred now *and* the expected recourse costs $Q(z)$.

In general, it is very hard to solve two-stage recourse models, since $Y \subset \mathbb{R}_+^{n_2}$ may impose integrality restrictions on the recourse actions y, and since the random vector ω may be continuously distributed so that evaluating $Q(z)$ is equivalent to computing a multi-dimensional integral. However, in order to model many practical problems realistically, we have to be able to deal with these two difficulties.

A typical approach to overcome these difficulties is to approximate the recourse function Q with a recourse function $\hat{Q}$ such that the approximating model can be solved within reasonable time limits. For example, for continuous recourse functions $\hat{Q}$ (with $Y = \mathbb{R}_+^{n_2}$) with a discretely distributed random vector ω, efficient solution methods are available, most of them based on the L-shaped method of [34]. In this section, we consider several approximations $\hat{Q}$ dealing with the above mentioned difficulties, and, in particular, we derive error bounds for these approximations which are necessary to guarantee the quality of the resulting approximating solutions.

To solve continuous recourse models with continuous random variables typically discrete approximations are used, that is, the continuous random variables in the model are approximated by discrete ones. We discuss two types of discrete approximations in Sect. 4.3, one of them a lower bound of Q discussed in e.g. [16], and we derive error bounds for these approximations showing that the approximations are good as long as the total variations of the densities of the random variables in the model are small enough.

The main difficulty solving integer recourse models (with $Y = \mathbb{Z}_+^{n_2}$) is that generally the integer recourse function is *non-convex* [26]. Several algorithms are available for mixed-integer and pure integer recourse models (see [20,21,31] for an overview), but these algorithms are restricted to special cases or have difficulties solving very large real-life problem instances. An alternative approach to obtain good first-stage decisions x is to use convex approximations of Q. For example, for totally unimodular (TU) integer recourse models convex approximations are derived in [33], building on earlier work in [19], and in [27] we derive a uniform error bound for this class of convex approximations. In Sect. 4.4 we show that this error bound can be obtained using the results in Corollary 2 of Sect. 3.3. Moreover, we construct a new convex approximation improving the error bound of [27] by a factor 2, and we show that this convex approximation has the best worst-case error bound possible.

In Sect. 4.5 we use the total variation bounds of Sect. 3.3 to derive a tractable Lipschitz constant for pure integer recourse models.

Throughout the remainder of this paper we assume that

(A1) the recourse is complete, i.e. $v(s) < +\infty$ for all $s \in \mathbb{R}^m$,
(A2) the recourse is sufficiently expensive, i.e. $v(s) > -\infty$ for all $s \in \mathbb{R}^m$, and
(A3) $\mathbb{E}_\omega[|\omega_i|] < +\infty$, $i = 1, \ldots, m$.

As a consequence the recourse function Q is finite for all $z \in \mathbb{R}^m$.

4.1 Dual representations

For continuous recourse models with $Y = \mathbb{R}^{n2}_+$, a common approach is to use strong LP-duality to rewrite the second-stage value function v, see e.g. [2] or [32]. We have for every $s \in \mathbb{R}^m$,

$$\begin{aligned} v(s) &= \min\{qy : Wy \geq s,\ y \in \mathbb{R}^{n2}_+\} \\ &= \max\{\lambda s : \lambda W \leq q,\ \lambda \in \mathbb{R}^m_+\}. \end{aligned}$$

Assumptions (A1) and (A2) imply that the dual feasible region $\Lambda := \{\lambda \in \mathbb{R}^m : \lambda W \leq q\}$ is non-empty and bounded, so that it is spanned by finitely many extreme points denoted $\lambda^1, \ldots, \lambda^K$. Since v attains its optimum in one of these extreme points, we have

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

For totally unimodular (TU) integer recourse models with $Y = \mathbb{Z}^{n2}_+$ and TU recourse matrix W we can obtain a similar dual representation for v. Since W is TU, and thus integral, we have

$$\begin{aligned} v(s) &= \min\{qy : Wy \geq \lceil s \rceil,\ y \in \mathbb{Z}^{n2}_+\} \\ &= \min\{qy : Wy \geq \lceil s \rceil,\ y \in \mathbb{R}^{n2}_+\}. \end{aligned}$$

Again using strong LP-duality we find in this case that

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

Using these expressions we also obtain dual representations for the expected recourse function Q. For continuous recourse models we have

$$Q(z) = \mathbb{E}_\omega\Big[\max_{k=1,\ldots,K} \lambda^k(\omega - z)\Big], \quad z \in \mathbb{R}^m, \tag{30}$$

and for TU integer recourse

$$Q(z) = \mathbb{E}_\omega\Big[\max_{k=1,\ldots,K} \lambda^k \lceil \omega - z \rceil\Big], \quad z \in \mathbb{R}^m. \tag{31}$$

4.2 Extending total variation bounds to a multi-dimensional setting

From the dual representation of Q in (30) and (31) it is still unclear how we can use our one-dimensional results to derive error bounds for approximations of Q. One of the difficulties is the maximum operator, since different dual vertices λ^k may be optimal for different values of ω. Nevertheless, by analyzing properties of the function

$$\lambda(s) \in \operatorname*{argmax}_{k=1,\dots,K} \lambda^k s, \quad s \in \mathbb{R}^m,$$

in Lemma 11, we are able to prove Theorem 2, the main result of this subsection. It allows to derive error bounds for discrete approximations of continuous recourse models in Sect. 4.3 and for convex approximations of TU integer recourse models in Sect. 4.4.

Lemma 11 *Let $\lambda^1, \dots, \lambda^K \in \mathbb{R}^m_+$ be given, and let $H : \mathbb{R}^m \mapsto \mathbb{R}^m$ be a separable function defined as*

$$H(x) = (H_1(x_1), \dots, H_m(x_m)),$$

where each H_i is non-decreasing. Then, there exists a function $\lambda : \mathbb{R}^m \mapsto \mathbb{R}^m$ such that

(i) *$\lambda(x) \in \operatorname{argmax}_{k=1,\dots,K} \lambda^k H(x)$ for all $x \in \mathbb{R}^m$, and*
(ii) *$\lambda_i(\cdot|x_{(i)}) : \mathbb{R} \mapsto \mathbb{R}$ defined as $\lambda_i(x_i|x_{(i)}) = \lambda_i(x)$ is non-decreasing for every $i = 1, \dots, m$, and $x_{(i)} \in \mathbb{R}^{m-1}$.*

Here, $x_{(i)}$ denotes x without its i-th component.

Proof Obviously, there exists $\lambda : \mathbb{R}^m \mapsto \mathbb{R}^m$ such that (i) holds. However, such λ is not necessarily unique because for some $x \in \mathbb{R}^m$, the set $\operatorname{argmax}_{k=1,\dots,K} \lambda^k H(x)$ may contain more than one element. Note that in order to prove that there exists λ satisfying both (i) and (ii) it is of no restriction to assume, as we do, that $\lambda(x^1) = \lambda(x^2)$ if $H(x^1) = H(x^2)$.

Now suppose for contradiction that for every λ satisfying (i) there exists $i = 1, \dots, m$, and $x_{(i)} \in \mathbb{R}^{m-1}$ such that $\lambda_i(\cdot|x_{(i)})$ is not non-decreasing. That is, there exist $x^1, x^2 \in \mathbb{R}^m$ with $x^1_{(i)} = x^2_{(i)}$ and $x^1_i < x^2_i$ such that $\lambda_i(x^1) > \lambda_i(x^2)$. Note that $\lambda_i(x^1) > \lambda_i(x^2)$ implies that $H_i(x^1_i) \neq H_i(x^2_i)$ and thus $H_i(x^1_i) < H_i(x^2_i)$ since H_i is non-decreasing.

Since λ satisfies (i), we have

$$\lambda(x^1)H(x^1) \geq \lambda(x^2)H(x^1) \tag{32}$$

and

$$\lambda(x^2)H(x^2) \geq \lambda(x^1)H(x^2). \tag{33}$$

By adding the inequalities in (32) and (33) it follows that

$$\left[\lambda(x^1) - \lambda(x^2)\right]\left[H(x^1) - H(x^2)\right] \geq 0,$$

and since $x^1_{(i)} = x^2_{(i)}$, implying $H_j(x^1_j) = H_j(x^2_j)$ for $j \neq i$, we have

$$\left[\lambda_i(x^1) - \lambda_i(x^2)\right]\left[H_i(x^1_i) - H_i(x^2_i)\right] \geq 0.$$

This contradicts the assertion that $\lambda_i(x^1) > \lambda_i(x^2)$ and $H_i(x^1_i) < H_i(x^2_i)$. $\square$

Now we are almost ready to prove the main result of this subsection. First, however, we need to define a sufficiently rich set $\mathcal{H}^m$ of 'nice' m-dimensional joint density functions.

Definition 6 Let $\mathcal{H}^m$ denote the set of all m-dimensional joint pdf f whose conditional density functions $f_i(\cdot|x_{(i)})$ are of bounded variation. That is, $f_i(\cdot|x_{(i)}) \in \mathcal{F}$ for all $i = 1, \ldots, m$, and $x_{(i)} \in \mathbb{R}^{m-1}$.

An example of such a joint density function is the multivariate normal joint pdf f since its conditional density functions $f_i(\cdot|x_{(i)})$ are (one-dimensional) normal pdf as well, and thus of bounded variation. Moreover, a joint pdf f belongs to $\mathcal{H}^m$ if f corresponds to a continuous random vector with independent components ω_i, whose density functions f_i are of bounded variation.

Theorem 2 *Let ω be a random vector with joint pdf $f \in \mathcal{H}^m$, and let $\lambda^1, \ldots, \lambda^K$ be a collection of vectors in $\mathbb{R}^m_+$. Define*

$$Q^j := \mathbb{E}_\omega\Big[\max_{k=1,\ldots,K} \lambda^k H^j(\omega)\Big], \quad j = 1, 2,$$

where $H^1, H^2 : \mathbb{R}^m \mapsto \mathbb{R}^m$ are separable functions defined as $H^j(x) = (H^j_1(x_1), \ldots, H^j_m(x_m))$ with H^j_i non-decreasing for $j = 1, 2$, and $i = 1, \ldots, m$. Assume that $\varphi_i(x_i) := H^1_i(x_i) - H^2_i(x_i)$, $x_i \in \mathbb{R}$, is a periodic function with period p_i and finite mean value $\nu_i = p_i^{-1} \int_0^{p_i} \varphi_i(u)du = 0$ for every $i = 1, \ldots, m$. Then,

$$\sum_{i=1}^m \lambda^*_i \mathbb{E}_{\omega_{(i)}}\left[N\Big(\varphi_i, |\Delta| f_i(\cdot|\omega_{(i)})\Big)\right] \le Q^1 - Q^2 \le \sum_{i=1}^m \lambda^*_i \mathbb{E}_{\omega_{(i)}}\left[M\Big(\varphi_i, |\Delta| f_i(\cdot|\omega_{(i)})\Big)\right], \tag{34}$$

where $\lambda^*_i := \max_{k=1,\ldots,K} \lambda^k_i$.

Proof Since H^1 is separable and its components are non-decreasing there exists a function $\lambda : \mathbb{R} \mapsto \mathbb{R}$ satisfying (i) and (ii) of Lemma 11. For this function λ, we have $Q^1 = \mathbb{E}_\omega[\lambda(\omega)H^1(\omega)]$ and $Q^2 \ge \mathbb{E}_\omega[\lambda(\omega)H^2(\omega)]$. Thus,

$$\begin{aligned} Q^1 - Q^2 &\le \mathbb{E}_\omega\Big[\lambda(\omega)\Big(H^1(\omega) - H^2(\omega)\Big)\Big] \\ &= \sum_{i=1}^m \mathbb{E}_\omega[\lambda_i(\omega)\varphi_i(\omega_i)] \\ &= \sum_{i=1}^m \int_{\mathbb{R}^m} \lambda_i(x)\varphi_i(x_i) f(x)dx. \end{aligned}$$

For every $i = 1, \ldots, m$, we condition on $x_{(i)} \in \mathbb{R}^{m-1}$ to obtain

$$Q^1 - Q^2 \le \sum_{i=1}^m \int_{\mathbb{R}^{m-1}} \left\{\int_{\mathbb{R}} \lambda_i(x_i|x_{(i)})\varphi_i(x_i) f_i(x_i|x_{(i)})dx_i\right\} f_{(i)}(x_{(i)})dx_{(i)}.$$

Observe that the inner integral can be bounded using Proposition 3, since φ_i is a periodic function with period p_i and mean value $\nu_i = 0$, $f_i(\cdot|x_{(i)}) \in \mathcal{F}$, and $\lambda_i(\cdot|x_{(i)})$ is a non-decreasing function such that $0 \leq \lambda_i(x_i|x_{(i)}) \leq \lambda_i^*$ for all $x_i \in \mathbb{R}$. Hence, the second inequality in (34) follows from

$$Q^1 - Q^2 \leq \sum_{i=1}^{m} \int_{\mathbb{R}^{m-1}} \lambda_i^* M\Big(\varphi_i, |\Delta| f_i(\cdot|x_{(i)})\Big) f_{(i)}(x_{(i)}) dx_{(i)}$$
$$= \sum_{i=1}^{m} \lambda_i^* \mathbb{E}_{\omega_{(i)}} \Big[M\Big(\varphi_i, |\Delta| f_i(\cdot|\omega_{(i)})\Big)\Big].$$

The first inequality in (34) holds since by symmetry

$$Q^2 - Q^1 \leq \sum_{i=1}^{m} \lambda_i^* \mathbb{E}_{\omega_{(i)}} \Big[M\Big(-\varphi_i, |\Delta| f_i(\cdot|\omega_{(i)})\Big)\Big],$$

and since $M\Big(-\varphi_i, |\Delta| f_i(\cdot|\omega_{(i)})\Big) = -N\Big(\varphi_i, |\Delta| f_i(\cdot|\omega_{(i)})\Big)$ by Lemma 1 (i). □

Corollary 3 *Consider the same setting as in Theorem 2. Then,*

$$|Q^1 - Q^2| \leq \frac{1}{4} \sum_{i=1}^{m} \lambda_i^* \mathbb{E}_{\omega_{(i)}} \Big[|\Delta| f_i(\cdot|\omega_{(i)})\Big] \int_0^{p_i} |\varphi_i(x)| dx, \tag{35}$$

where $\lambda_i^* := \max_{k=1,\ldots,K} \lambda_i^k$.

Proof The inequality in (35) follows immediately from applying Proposition 1 to (34). □

In the following sections we will rely on Theorem 2 to prove a variety of approximation results.

4.3 Discrete approximations for continuous recourse models

Consider the continuous recourse function

$$Q(z) = \mathbb{E}_f\Big[\min\{qy : Wy \geq \omega - z,\ y \in \mathbb{R}_+^{n_2}\}\Big], \quad z \in \mathbb{R}^m, \tag{36}$$

where ω is a continuous random vector with joint pdf $f \in \mathcal{H}^m$. Solving the corresponding recourse model is difficult, since evaluation of Q requires the computation of a multi-dimensional integral. This is why the continuous random vector ω is typically approximated by a *discrete* random vector ξ. The approximating model can be solved efficiently using existing methods, most of them inspired by the L-shaped method of [34].

There has been a vast amount of work in the literature on how to select a discrete random vector ξ and how to guarantee the quality of the approximating solutions. The construction of approximating distributions goes back to Jensen [15] and Edmundson-Madansky [9,23]. Examples of sampling methods are the sample average approximation, discussed in e.g. [32], and stochastic decomposition [14]. Scenario generation for multistage recourse problems is surveyed in [6], see also [25], and scenario reduction techniques, based on stability results (see e.g. [28]), are discussed in e.g. [13] and [7]. In this section, we do not provide a detailed (numerical) comparison with these existing methods but merely show initial results and indicate why they might be useful.

We consider two types of discrete approximations and derive error bounds for both, showing that the approximations are good as long as the total variations of the densities in the model are small enough. The first so-called mid-point approximation, is rather simple but an error bound can be derived directly from Theorem 2. The second discrete approximation, the lower bound denoted as the Jensen approximation [15], is more sophisticated so that deriving an error bound is not straightforward. Nevertheless, we do so by comparing the Jensen approximation with the mid-point approximation.

The obvious advantage of an error bound for the Jensen approximation is that it obviates the need for an upper bound for Q to guarantee the quality of the approximation, as is often the case in approximation schemes (see e.g. [4]) where iteratively better lower and upper bounds are calculated until the gap is sufficiently small. This might considerably speed up computations, especially in case this error bound is small, since computing a tight upper bound is usually much more demanding than computing a lower bound. Precisely this latter observation motivated several studies deriving tight upper bounds within reasonable time limits. For example, Frauendorfer [10], Kall [17], Gassmann and Ziemba [11], Edirisinghe and Ziemba [8], and Birge and Wets [5] use moment problems with varying conditions to derive such upper bounds, Birge and Wallace [3] construct a separable piecewise linear upper bound that requires a polynomial number of iterations in the number of random variables, and Birge [1] derives bounds for aggregating scenarios and time stages in multistage problems.

4.3.1 Mid-point approximation

The mid-point approximation can be interpreted as follows. Partition $\mathbb{R}^m$ into equally sized hyperrectangles $C^l(\alpha, \rho)$, defined as

$$C^l(\alpha, \rho) := \prod_{i=1}^{m} \rho_i \Big(\alpha_i + l_i - 1, \alpha_i + l_i\Big], \quad l \in \mathbb{Z}^m.$$

Here, $\rho \in \mathbb{R}^m$ with $\rho > 0$ represents the size of the hyperrectangles and $\alpha \in \mathbb{R}^m$ is a shift parameter. Let $\zeta^l(\alpha, \rho)$ denote the mid-point of $C^l(\alpha, \rho)$ defined as

$$\zeta^l(\alpha,\rho) := P\Big(\alpha + l - \frac{1}{2}e_m\Big), \quad l \in \mathbb{Z}^m,$$

with $P = \text{diag}\{\rho_1,\dots,\rho_m\}$ and e_m the all-one vector. Then, the mid-point approximation is obtained by concentrating all probability mass corresponding to ω on $C^l(\alpha,\rho)$ in its mid-point $\zeta^l(\alpha,\rho)$ for every $l \in \mathbb{Z}^m$.

Definition 7 For every $\alpha \in \mathbb{R}^m$ and $\rho \in \mathbb{R}^m$ with $\rho > 0$, let the *ρ-size mid-point approximation* $\hat{Q}^\rho_\alpha$ of the continuous recourse function Q be defined as

$$\hat{Q}^\rho_\alpha(z) := \mathbb{E}_f\Big[\min\{qy : Wy \geq \xi(\omega;\alpha,\rho) - z,\ y \in \mathbb{R}^{n_2}_+\}\Big], \quad z \in \mathbb{R}^m, \tag{37}$$

where $\xi(\omega;\alpha,\rho) := P\left(\left\lceil P^{-1}\omega\right\rceil_\alpha - 1/2e_m\right)$, with $P = \text{diag}\{\rho_1,\dots,\rho_m\}$, is a discrete random vector with support contained in $\{\zeta^l(\alpha,\rho) : l \in \mathbb{Z}^m\}$ such that

$$\mathbb{P}\Big\{\xi(\omega;\alpha,\rho) = \zeta^l(\alpha,\rho)\Big\} = \mathbb{P}\Big\{\omega \in C^l(\alpha,\rho)\Big\}, \quad \text{for all} \quad l \in \mathbb{Z}^m.$$

Intuitively, it is clear that if the components of ρ are small enough, and thus the size of the hyperrectangles is small, then $\hat{Q}^\rho_\alpha$ will be a good approximation of Q. The shift parameter α does not influence the worst-case error bound for the ρ-size mid-point approximation.

Theorem 3 *Consider the continuous recourse function*

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

where ω is a continuous random vector with joint pdf $f \in \mathcal{H}^m$, and consider its ρ-size mid-point approximation *with shift parameter $\alpha \in \mathbb{R}^m$, defined (according to Definition 7) as*

$$\hat{Q}^\rho_\alpha(z) := \mathbb{E}_f\Big[\min\Big\{qy : Wy \geq P\left(\left\lceil P^{-1}\omega\right\rceil_\alpha - 1/2e_m\right) - z,\ y \in \mathbb{R}^{n_2}_+\Big\}\Big], \quad z \in \mathbb{R}^m.$$

Then,

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

where $\lambda_i^* := \max_{k=1,\dots,K}\lambda_i^k$ and h is defined in (25).

Proof Using the dual representation for continuous recourse models in (30) we can write

$$Q(z) = \mathbb{E}_f\Big[\max_{k=1,\dots,K}\lambda^k(\omega - z)\Big], \quad z \in \mathbb{R}^m,$$

and

$$\hat{Q}_{\alpha}^{\rho}(z) = \mathbb{E}_f \Big[\max_{k=1,\dots,K} \lambda^k \Big(P \Big(\big\lceil P^{-1}\omega \big\rceil_{\alpha} - 1/2e_m \Big) - z \Big) \Big], \quad z \in \mathbb{R}^m.$$

For every $z \in \mathbb{R}^m$, we define $H^1(\omega) = \omega - z$ and $H^2(\omega) = P(\lceil P^{-1}\omega \rceil_{\alpha} - 1/2e_m) - z$, and observe that H^1 and H^2 are separable functions, the latter since P is a diagonal matrix. Moreover, the components of H^1 and H^2 are non-decreasing, and the functions

$$\varphi_i(x_i) = H_i^1(x_i) - H_i^2(x_i) = x_i + \rho_i/2 - \rho_i \lceil x_i/\rho_i \rceil_{\alpha_i}, \quad x_i \in \mathbb{R},$$

are periodic with period $p_i = \rho_i$ and mean value $\nu_i = 0$. Thus, all conditions of Theorem 2 are satisfied, and since the functions φ_i do not depend on z, we conclude from (34) that

$$\sup_{z \in \mathbb{R}^m} \Big\{ Q(z) - \hat{Q}_{\alpha}^{\rho}(z) \Big\} \leq \sum_{i=1}^{m} \lambda_i^* \mathbb{E}_{\omega_{(i)}} \Big[M(\varphi_i, |\Delta| f_i(\cdot|\omega_{(i)})) \Big],$$

and

$$\sup_{z \in \mathbb{R}^m} \Big\{ \hat{Q}_{\alpha}^{\rho}(z) - Q(z) \Big\} \leq - \sum_{i=1}^{m} \lambda_i^* \mathbb{E}_{\omega_{(i)}} \Big[N(\varphi_i, |\Delta| f_i(\cdot|\omega_{(i)})) \Big].$$

We can compute these upper bounds using special properties of the functions φ_i involved. In fact, the φ_i are scaled versions of the function $\bar{\varphi}_{\alpha}$ in (26) in Example 2 of Sect. 3.3, since $\varphi_i(x_i) = \rho_i \bar{\varphi}_{\alpha_i}(x_i/\rho_i)$ for all $x_i \in \mathbb{R}$ and all $i = 1, \dots, m$. Hence we will conclude the proof by showing that

$$M(\varphi_i, B) = \rho_i M(\bar{\varphi}_{\alpha_i}, \rho_i B) = \frac{\rho_i}{2} h(\rho_i B), \tag{38}$$

and

$$N(\varphi_i, B) = \rho_i N(\bar{\varphi}_{\alpha_i}, \rho_i B) = -\frac{\rho_i}{2} h(\rho_i B). \tag{39}$$

In both (38) and (39) the first equality follows from Lemma 1 (ii) and (iii), and the second one from Example 2. □

Theorem 3 shows that the error of the ρ-sized mid-point approximation decreases as either the components of ρ or the total variations of the conditional densities decrease. The first is common for discrete approximations since finer discretizations generally reduce the 'distance' between the probability distributions of ω and $\xi(\omega; \alpha, \rho)$ (see e.g. Römisch [28] for a formal definition), leading to better approximations. The second, however, does not follow from a reduction in the distance between probability distributions, since in Example 3 we show that the total variations may decrease whereas the distance between the probability distribution and its approximation remains the same. The example illustrates the unique nature of the error bound in Theorem 3, which is not just obtained by comparing probability distributions but by exploiting both the direct relation between ω and $\xi(\omega; \alpha, \rho)$ *and* the structure of the underlying second-stage value function v. Indeed, e.g. Kaut and Wallace [18] advocate the use of scenario generation methods taking the latter into account.

Example 3 Consider the continuous recourse function Q and its ρ-sized mid-point approximation $\hat{Q}_\alpha^\rho$ with $\rho = e_m$ and $\alpha = 0$, as defined in (36) and (37), respectively. For every $n \in \mathbb{N}$, let the continuous random vector ω^n be uniformly distributed on $[0, n]^m$, so that its components ω_i^n are independent and have pdf f_i^n defined as $f_i^n(x) = 1/n$ if $x \in [0, n]$ and $f_i^n(x) = 0$, otherwise. Since $|\Delta| f_i^n = 2/n$ for every $i = 1, \dots, m$, it follows from Theorem 3 that for $\omega := \omega_n$,

$$\|Q - \hat{Q}_\alpha^\rho\|_\infty \leq (8n)^{-1} \sum_{i=1}^m \lambda_i^*,$$

which converges to zero as $n \to \infty$.

Interestingly, the Kantorovich distance between the probability distributions of ω_n and its discrete approximation $\xi(\omega_n; \alpha, \rho) = \lceil \omega_n \rceil - 1/2$ is constant for all $n \in \mathbb{N}$, implying that the analysis of e.g. Römisch [28] cannot be used to derive the convergence result above.

4.3.2 Jensen approximation

The Jensen approximation is more sophisticated than the mid-point approximation: instead of concentrating all probability mass of ω on a hyperrectangle $C^l(\alpha, \rho)$ in its mid-point $\zeta^l(\alpha, \rho)$, the Jensen approximation puts it in its conditional mean $\mu^l(\alpha, \rho)$ defined as

$$\mu^l(\alpha, \rho) := \mathbb{E}_f\Big[\omega | \omega \in C^l(\alpha, \rho)\Big], \quad l \in \mathbb{Z}^m.$$

Although the Jensen approximation can be defined using any partition of $\mathbb{R}^m$, we use equally sized hyperrectangles $C^l(\alpha, \rho)$ to be able to compare the Jensen approximation with the mid-point approximation.

Definition 8 For every $\alpha \in \mathbb{R}^m$ and $\rho \in \mathbb{R}^m$ with $\rho > 0$, let the ρ-size Jensen approximation $\tilde{Q}_\alpha^\rho$ of the continuous recourse function Q be defined as

$$\tilde{Q}_\alpha^\rho(z) := \mathbb{E}_\omega\Big[\min\{qy : Wy \geq \xi^f(\omega; \alpha, \rho) - z,\ y \in \mathbb{R}_+^{n_2}\}\Big], \quad z \in \mathbb{R}^m,$$

where $\xi^f : \mathbb{R}^m \mapsto \mathbb{R}^m$ is defined as $\xi^f(x; \alpha, \rho) := \mu^l(\alpha, \rho)$ for $x \in C^l(\alpha, \rho)$, $l \in \mathbb{Z}^m$, and thus $\xi^f(\omega; \alpha, \rho)$ is a discrete random vector with

$$\mathbb{P}\Big\{\xi^f(\omega; \alpha, \rho) = \mu^l(\alpha, \rho)\Big\} = \mathbb{P}\Big\{\omega \in C^l(\alpha, \rho)\Big\}, \quad l \in \mathbb{Z}^m.$$

We will derive an upper bound for $\|Q - \tilde{Q}_\alpha^\rho\|_\infty$ by comparing the Jensen approximation $\tilde{Q}_\alpha^\rho$ and the mid-point approximation $\hat{Q}_\alpha^\rho$. In fact, the main purpose is to find an upper bound for

$$\sup_{z \in \mathbb{R}^m} \hat{Q}_\alpha^\rho(z) - \tilde{Q}_\alpha^\rho(z).$$

Indeed, this bound yields an upper bound for $Q(z) - \tilde{Q}_\alpha^\rho(z) = Q(z) - \hat{Q}_\alpha^\rho(z) + \hat{Q}_\alpha^\rho(z) - \tilde{Q}_\alpha^\rho(z)$ for all $z \in \mathbb{R}^m$, since $\| Q - \hat{Q}_\alpha^\rho \|_\infty$ is bounded by Theorem 3. Since the Jensen approximation is a lower bound for Q, it holds $Q(z) - \tilde{Q}_\alpha^\rho(z) \geq 0, z \in \mathbb{R}^m$.

The difference $\hat{Q}_\alpha^\rho(z) - \tilde{Q}_\alpha^\rho(z)$ can not be bounded using Theorem 2, since the underlying difference function is not periodic. Moreover, this function will depend on the pdf f, since the conditional means $\mu^l(\alpha, \rho)$ depend on f. Nevertheless, we are able to find a uniform upper bound for the one-sided difference $\hat{Q}_\alpha^\rho(z) - \tilde{Q}_\alpha^\rho(z)$ using additional analysis. The key observation is that for f constant on $C^l(\alpha, \rho)$, the conditional mean $\mu^l(\alpha, \rho)$ coincides with the mid-point $\zeta^l(\alpha, \rho)$ of $C^l(\alpha, \rho)$.

Proposition 4 *Let $\alpha \in \mathbb{R}^m$ and $\rho \in \mathbb{R}^m$ be given. Consider the ρ-size mid-point approximation $\hat{Q}_\alpha^\rho$ and the ρ-size Jensen approximation $\tilde{Q}_\alpha^\rho$. Then, for every $z \in \mathbb{R}^m$,*

$$\hat{Q}_\alpha^\rho(z) - \tilde{Q}_\alpha^\rho(z) \leq \frac{1}{2} \sum_{i=1}^m \rho_i \lambda_i^* \mathbb{E}_{\omega_{(i)}} \Big[h\Big(\rho_i |\Delta| f_i(\cdot|\omega_{(i)}) \Big) \Big],$$

where $\lambda_i^* := \max_{k=1,\ldots,K} \lambda_i^k$ and h is defined in (25).

Proof Let $z \in \mathbb{R}^m$ be given and consider the dual representations of $\hat{Q}_\alpha^\rho(z)$ and $\tilde{Q}_\alpha^\rho(z)$ given by

$$\hat{Q}_\alpha^\rho(z) = \mathbb{E}_f \Big[\max_{k=1,\ldots,K} \lambda^k \Big(P \left(\left\lceil P^{-1}\omega \right\rceil_\alpha - \frac{1}{2} e_m \right) - z \Big) \Big],$$

and

$$\tilde{Q}_\alpha^\rho(z) = \mathbb{E}_f \Big[\max_{k=1,\ldots,K} \lambda^k \Big(\xi^f(\omega; \alpha, \rho) - z \Big) \Big],$$

respectively. Observe that for every $l \in \mathbb{Z}^m$, both $P\left(\left\lceil P^{-1}\omega \right\rceil_\alpha - \frac{1}{2} e_m \right) - z$ and $\xi^f(\omega; \alpha, \rho) - z$ are constant for $\omega \in C^l(\alpha, \rho)$. Because of the first, there exists $\lambda^z : \mathbb{R}^m \mapsto \mathbb{R}^m$ such that λ^z is constant on $C^l(\alpha, \rho)$ for every $l \in \mathbb{Z}^m$, and

$$\lambda^z(x) \in \operatorname*{argmax}_{k=1,\ldots,K} \lambda^k \left(P \left(\left\lceil P^{-1} x \right\rceil_\alpha - \frac{1}{2} e_m \right) - z \right), \quad x \in \mathbb{R}^m.$$

Thus,

$$\hat{Q}_\alpha^\rho(z) - \tilde{Q}_\alpha^\rho(z) \leq \mathbb{E}_f \left[\lambda^z(\omega) \Big(P \left(\left\lceil P^{-1}\omega \right\rceil_\alpha - \frac{1}{2} e_m \right) - \xi^f(\omega; \alpha, \rho) \Big) \right].$$

We can rewrite this expression as

$$\begin{aligned}
\hat{Q}_\alpha^\rho(z) - \tilde{Q}_\alpha^\rho(z) &\leq \sum_{i=1}^m \mathbb{E}_f\Big[\lambda_i^z(\omega)\Big(\rho_i(\lceil \omega_i/\rho_i \rceil_{\alpha_i} - 1/2) - \xi_i^f(\omega;\alpha,\rho)\Big)\Big] \\
&= \sum_{i=1}^m \int_{\mathbb{R}^m} \lambda_i^z(x)\Big(\rho_i(\lceil x_i/\rho_i \rceil_{\alpha_i} - 1/2) - \xi_i^f(x;\alpha,\rho)\Big) f(x)dx \\
&= \sum_{i=1}^m \sum_{l\in\mathbb{Z}^m} \int_{C^l(\alpha,\rho)} \lambda_i^z(x)\Big(\rho_i(\lceil x_i/\rho_i \rceil_{\alpha_i} - 1/2) - \xi_i^f(x;\alpha,\rho)\Big) f(x)dx.
\end{aligned}$$

Observing that the integrand is constant in x on $C^l(\alpha,\rho)$ and writing it as $\lambda_i^z(l)(\zeta_i^l(\alpha,\rho) - \mu_i^l(\alpha,\rho))$, it follows that

$$\hat{Q}_\alpha^\rho(z) - \tilde{Q}_\alpha^\rho(z) \leq \sum_{i=1}^m \theta_i^f(\alpha,\rho), \tag{40}$$

where $\theta_i^f(\alpha,\rho)$ is defined for every $i = 1,\dots,m$ as

$$\theta_i^f(\alpha,\rho) := \sum_{l\in\mathbb{Z}^m} \lambda_i^z(l)\Big(\zeta_i^l(\alpha,\rho) - \mu_i^l(\alpha,\rho)\Big) \int_{C^l(\alpha,\rho)} f(x)dx.$$

Consider first the special case that $\zeta_i^l(\alpha,\rho) \geq \mu_i^l(\alpha,\rho)$ for all $l \in \mathbb{Z}^m$ for some $i = 1,\dots,m$. Then,

$$\begin{aligned}
\theta_i^f(\alpha,\rho) &\leq \lambda_i^* \sum_{l\in\mathbb{Z}^m} \Big(\zeta_i^l(\alpha,\rho) - \mu_i^l(\alpha,\rho)\Big) \int_{C^l(\alpha,\rho)} f(x)dx \\
&= \lambda_i^* \int_{\mathbb{R}^m} (\rho_i(\lceil x_i/\rho_i \rceil_{\alpha_i} - 1/2) - x_i) f(x)dx,
\end{aligned}$$

since for every $l \in \mathbb{Z}^m$,

$$\mu_i^l(\alpha,\rho) := \mathbb{E}_f\Big[\omega_i | \omega \in C^l(\alpha,\rho)\Big] = \frac{\int_{C^l(\alpha,\rho)} x_i f(x)dx}{\int_{C^l(\alpha,\rho)} f(x)dx}$$

and thus

$$\sum_{l\in\mathbb{Z}^m} \mu_i^l(\alpha,\rho) \int_{C^l(\alpha,\rho)} f(x)dx = \int_{\mathbb{R}^m} x_i f(x)dx.$$

Since $\rho_i(\lceil x_i/\rho_i\rceil_{\alpha_i}-1/2)-x_i = -\rho_i\bar{\varphi}_{\alpha_i}(x_i/\rho_i)$ with $\bar{\varphi}_{\alpha_i}$ the periodic function defined in (26) of Example 2, proceeding as before to obtain

$$\begin{aligned}
-\int_{\mathbb{R}^m}\rho_i\bar{\varphi}_{\alpha_i}(x_i/\rho_i)f(x)dx_i &= -\int_{\mathbb{R}^{m-1}}\int_{\mathbb{R}}\rho_i\bar{\varphi}_{\alpha_i}(x_i/\rho_i)f_i(x_i)dx_i\,f_{(i)}(x_{(i)})dx_{(i)}\\
&\le \int_{\mathbb{R}^{m-1}}\frac{1}{2}\rho_i h\Big(\rho_i|\Delta|f_i(\cdot|x_{(i)})\Big)f_{(i)}(x_{(i)})dx_{(i)}\\
&= \frac{1}{2}\rho_i\mathbb{E}_{\omega_{(i)}}\Big[h\Big(\rho_i|\Delta|f_i(\cdot|\omega_{(i)})\Big)\Big],
\end{aligned}$$

with h defined in (25), it follows that

$$\theta_i^f(\alpha,\rho) \le \frac{1}{2}\rho_i\lambda_i^*\mathbb{E}_{\omega_{(i)}}\Big[h\Big(\rho_i|\Delta|f_i(\cdot|\omega_{(i)})\Big)\Big]. \tag{41}$$

Hence, if indeed $\zeta_i^l(\alpha,\rho)\ge\mu_i^l(\alpha,\rho)$ holds for every $i=1,\dots,m$, and $l\in\mathbb{Z}^m$, then

$$\hat{Q}_\alpha^\rho(z)-\tilde{Q}_\alpha^\rho(z) \le \frac{1}{2}\sum_{i=1}^m\rho_i\lambda_i^*\mathbb{E}_{\omega_{(i)}}\Big[h\Big(\rho_i|\Delta|f_i(\cdot|\omega_{(i)})\Big)\Big].$$

Now consider the general case for some fixed $i\in\{1,\dots,m\}$. Obviously, $\zeta_i^l(\alpha,\rho)\ge\mu_i^l(\alpha,\rho)$ does not hold for every $l\in\mathbb{Z}^m$ and $f\in\mathcal{H}^m$. We will show that nevertheless (41) is also true in this case, by constructing a pdf g for which the above condition holds together with $\theta_i^f(\alpha,\rho)\le\theta_i^g(\alpha,\rho)$ and $|\Delta|g_i(\cdot|x_{(i)})\le|\Delta|f_i(\cdot|x_{(i)})$ for every $x_{(i)}\in\mathbb{R}^{m-1}$, so that

$$\begin{aligned}
\theta_i^f(\alpha,\rho)\le\theta_i^g(\alpha,\rho) &\le \frac{1}{2}\rho_i\lambda_i^*\mathbb{E}_{\omega_{(i)}}\Big[h\Big(\rho_i|\Delta|g_i(\cdot|\omega_{(i)})\Big)\Big]\\
&\le \frac{1}{2}\rho_i\lambda_i^*\mathbb{E}_{\omega_{(i)}}\Big[h\Big(\rho_i|\Delta|f_i(\cdot|\omega_{(i)})\Big)\Big].
\end{aligned} \tag{42}$$

This pdf g is obtained by flattening the ith conditional densities of f, and is defined as

$$g(x)=g_i(x_i|x_{(i)})g_{(i)}(x_{(i)}),\qquad \text{for all } x\in\mathbb{R}^m,$$

where $g_{(i)}(x_{(i)}):=f_{(i)}(x_{(i)})$ for all $x_{(i)}\in\mathbb{R}^{m-1}$, and for every $l\in\mathbb{Z}^m$ and $x\in C^l(\alpha,\rho)$ we define

$$g_i(x_i|x_{(i)})=\begin{cases} f_i(x_i|x_{(i)}), & \text{if } \zeta_i^l(\alpha,\rho)\ge\mu_i^l(\alpha,\rho),\\ K_i^l(x_{(i)};\alpha,\rho), & \text{otherwise},\end{cases}$$

with

$$K_i^l(x_{(i)};\alpha,\rho):=|C_i^l(\alpha,\rho)|^{-1}\int_{C_i^l(\alpha,\rho)}f_i(x_i|x_{(i)})dx_i=\rho_i^{-1}\int_{C_i^l(\alpha,\rho)}f_i(x_i|x_{(i)})dx_i.$$

Here, $C_i^l(\alpha, \rho) \subset \mathbb{R}$ is defined as $C_i^l(\alpha, \rho) := \rho_i(\alpha_i + l_i - 1, \alpha_i + l_i]$, and similarly, we define $C_{(i)}^l(\alpha, \rho) \subset \mathbb{R}^{m-1}$ as $C_{(i)}^l(\alpha, \rho) := \prod_{j \neq i} C_j^l(\alpha, \rho)$. In order to prove (42), we will show that the pdf g satisfies the properties

(i) $g_{(i)}(x_{(i)}) = f_{(i)}(x_{(i)})$ for all $x_{(i)} \in \mathbb{R}^{m-1}$,
(ii) $\int_{C_i^l(\alpha,\rho)} g_i(x_i|x_{(i)})dx_i = \int_{C_i^l(\alpha,\rho)} f_i(x_i|x_{(i)})dx_i$ for all $l \in \mathbb{Z}^m$ and $x_{(i)} \in C_{(i)}^l(\alpha, \rho)$,
(iii) $\int_{C^l(\alpha,\rho)} f(x)dx = \int_{C^l(\alpha,\rho)} g(x)dx$ for every $l \in \mathbb{Z}^m$,
(iv) $g_i(\cdot|x_{(i)}) \in \mathcal{F}$ and $|\Delta|g_i(\cdot|x_{(i)}) \leq |\Delta| f_i(\cdot|x_{(i)})$ for all $x_{(i)} \in \mathbb{R}^{m-1}$,
(v) the conditional means $\hat{\mu}_i^l(\alpha, \rho)$ corresponding to g satisfy $\zeta_i^l(\alpha, \rho) \geq \hat{\mu}_i^l(\alpha, \rho)$ for all $l \in \mathbb{Z}^m$,
(vi) $\hat{\mu}_i^l(\alpha, \rho) \leq \mu_i^l(\alpha, \rho)$ for all $l \in \mathbb{Z}^m$.

Note that (i) holds by definition of $g_{(i)}$, (ii) holds either trivially or

$$\int_{C_i^l(\alpha,\rho)} g_i(x_i|x_{(i)})dx_i = \int_{C_i^l(\alpha,\rho)} K_i^l(x_{(i)}; \alpha, \rho)dx_i = \int_{C_i^l(\alpha,\rho)} f_i(x_i|x_{(i)})dx_i,$$

and (iii) follows from (i) and (ii) by construction of g. Property (iv) holds since for every $x_{(i)} \in \mathbb{R}^{m-1}$, $g(\cdot|x_{(i)})$ is a flattened version of $f(\cdot|x_{(i)})$ obtained by applying Lemma 8 repeatedly. Moreover, if for some $l \in \mathbb{Z}^m$, $\zeta_i^l(\alpha, \rho) \geq \mu_i^l(\alpha, \rho)$, then $g_i(x_i|x_{(i)}) = f(x_i|x_{(i)})$ for $x \in C^l(\alpha, \rho)$ so that $\hat{\mu}_i^l(\alpha, \rho) = \mu_i^l(\alpha, \rho)$, and if $\zeta_i^l(\alpha, \rho) < \mu_i^l(\alpha, \rho)$, then $g_i(\cdot|\cdot)$ is constant on $C^l(\alpha, \rho)$ so that the conditional mean $\hat{\mu}_i(\alpha, \rho)$ coincides with the midpoint $\zeta_i^l(\alpha, \rho)$. From these observations (v) and (vi) follow.

Combining (iii) and (vi) we immediately have

$$\begin{aligned}
\theta_i^f(\alpha, \rho) &:= \sum_{l \in \mathbb{Z}^m} \lambda_i^z(l)\Big(\zeta_i^l(\alpha, \rho) - \mu_i^l(\alpha, \rho)\Big) \int_{C^l(\alpha,\rho)} f(x)dx \\
&\leq \sum_{l \in \mathbb{Z}^m} \lambda_i^z(l)\Big(\zeta_i^l(\alpha, \rho) - \hat{\mu}_i^l(\alpha, \rho)\Big) \int_{C^l(\alpha,\rho)} g(x)dx \\
&= \theta_i^g(\alpha, \rho).
\end{aligned}$$

Moreover, since $\zeta_i^l(\alpha, \rho) \geq \hat{\mu}_i^l(\alpha, \rho)$ for every $l \in \mathbb{Z}^m$ by (v), the inequality in (41) holds for g, yielding

$$\theta_i^g(\alpha, \rho) \leq \frac{1}{2}\rho_i \lambda_i^* \mathbb{E}_{\omega_{(i)}}\Big[h\Big(\rho_i|\Delta|g_i(\cdot|\omega_{(i)})\Big)\Big].$$

Since h is non-decreasing, it now follows from (iv) that

$$\frac{1}{2}\rho_i \lambda_i^* \mathbb{E}_{\omega_{(i)}}\Big[h\Big(\rho_i|\Delta|g_i(\cdot|\omega_{(i)})\Big)\Big] \leq \frac{1}{2}\rho_i \lambda_i^* \mathbb{E}_{\omega_{(i)}}\Big[h\Big(\rho_i|\Delta|f_i(\cdot|\omega_{(i)})\Big)\Big],$$

and we conclude that (42) holds, indeed. Hence, for every $z \in \mathbb{R}^m$, (40) yields

$$\hat{Q}^\rho_\alpha(z) - \tilde{Q}^\rho_\alpha(z) \leq \frac{1}{2} \sum_{i=1}^m \rho_i \lambda_i^* \mathbb{E}_{\omega_{(i)}} \Big[h\Big(\rho_i |\Delta| f_i(\cdot|\omega_{(i)}) \Big) \Big].$$

□

Using this proposition we derive an error bound for the Jensen approximation.

Theorem 4 *Consider the continuous recourse function*

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

where ω is a continuous random vector with joint pdf $f \in \mathcal{H}^m$, and its ρ-size Jensen approximation $\tilde{Q}^\rho_\alpha$ with shift parameter α defined in Definition 8. Then,

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

where $\lambda_i^* := \max_{k=1,\dots,K} \lambda_i^k$ and h is defined in (25).

Proof Since every Jensen approximation is a lower bound for Q, we have $Q(z) - \tilde{Q}^\rho_\alpha(z) \geq 0$ for every $z \in \mathbb{R}^m$. The claim follows from combining Theorem 3 and Proposition 4. □

Remark 5 The value of the error bound for the Jensen approximation is precisely twice the value of the bound for the corresponding mid-point approximation in Theorem 3. We suspect that the first bound is not tight, leaving room for improvement. However, to our knowledge it is the first non-trivial a priori error bound for the Jensen approximation.

4.4 Convex approximations for totally unimodular integer recourse models

In this subsection we consider the totally unimodular integer recourse function

$$Q(z) := \mathbb{E}_f \Big[\min\{qy : Wy \geq \omega - z, \ y \in \mathbb{Z}^{n_2}_+\} \Big], \quad z \in \mathbb{R}^m, \tag{43}$$

where ω is a continuous random vector with joint pdf $f \in \mathcal{H}^m$ and W is a TU matrix. In general, Q is a non-convex function. We discuss two types of convex approximations for Q. The first type of so-called α-approximations has been developed by van der Vlerk [33], and in [27] we derived a uniform error bound for these approximations. The second type is a new convex approximation for which we can derive an error bound using Theorem 2, improving on α-approximations by a factor 2. Interestingly, the underlying periodic functions in the derivation of this error bound are the same as for the mid-point approximation of the previous subsection, providing a connection between two seemingly unrelated research areas. Another interesting observation is that the α-approximations can be obtained by applying the mid-point approximation to the new convex approximation. We conclude this subsection by showing that the new convex approximation has the best worst-case error bound possible.

The α-approximations of [33] are defined for every $\alpha \in \mathbb{R}^m$ as

$$\bar{Q}_\alpha(z) := \mathbb{E}_f\Big[\min\{qy : Wy \geq \lceil\omega\rceil_\alpha - z,\ y \in \mathbb{R}^{n_1}_+\}\Big], \quad z \in \mathbb{R}^m, \tag{44}$$

where $\lceil\omega\rceil_\alpha := \lceil\omega - \alpha\rceil + \alpha$ is a discrete random vector with support contained in $\alpha + \mathbb{Z}^m$. Since $\bar{Q}_\alpha$ is the recourse function of a *continuous* recourse model, it is convex, and efficient solution methods are available to solve such models. Moreover, we can use Theorem 2 to derive the same uniform error bound as in [27].

Theorem 5 *Consider the totally unimodular integer recourse function Q defined in* (43) *with ω a continuous random vector with joint pdf $f \in \mathcal{H}^m$, and let $\bar{Q}_\alpha$ denote its* α-approximation *defined in* (44). *Then, for every $\alpha \in \mathbb{R}^m$,*

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

where $\lambda_i^* := \max_{k=1,\ldots,K} \lambda_i^k$ and h is defined in (25).

Proof Since $\bar{Q}_\alpha$ is a continuous recourse function we can use the dual representation in (30) to obtain for every $\alpha \in \mathbb{R}^m$,

$$\bar{Q}_\alpha(z) = \mathbb{E}_f\Big[\max_{k=1,\ldots,K} \lambda^k(\lceil\omega\rceil_\alpha - z)\Big], \quad z \in \mathbb{R}^m.$$

The dual representation of Q is given in (31). For every $z \in \mathbb{R}^m$, we define $H^1(\omega) = \lceil\omega\rceil_\alpha - z$ and $H^2(\omega) = \lceil\omega - z\rceil$ so that both H^1 and H^2 are separable and have non-decreasing components. Moreover, for every $i = 1, \ldots, m$, the function

$$\varphi_i(x_i) = H_i^1(x_i) - H_i^2(x_i) = \lceil x_i\rceil_{\alpha_i} - \lceil x_i\rceil_{z_i}, \quad x_i \in \mathbb{R},$$

is the same as $\varphi_{\alpha,\beta}$ in (22) in Example 1 with $\alpha = \alpha_i$ and $\beta = z_i$, and thus φ_i is periodic with period $p_i = 1$ and mean value $\nu_i = 0$. Since all conditions of Theorem 2 are satisfied, we conclude from (34), and from (24) in Example 1, that for every $\alpha \in \mathbb{R}^m$ and $z \in \mathbb{R}^m$,

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

□

Next, we consider a new convex approximation which we denote as the *shifted LP-relaxation approximation*. This approximation is obtained by shifting the right-hand side in the LP-relaxation of Q by $\frac{1}{2}e_m$, and its error bound, which we will derive using Theorem 2, yields an improvement over α-approximations by a factor 2.

Definition 9 Let the *shifted LP-relaxation approximation* $\tilde{Q}$ of the TU integer recourse function Q be defined as

$$\tilde{Q}(z) := \mathbb{E}_f\Big[\min\Big\{qy : Wy \geq \omega + \frac{1}{2}e_m - z,\ y \in \mathbb{R}_+^{n_2}\Big\}\Big], \quad z \in \mathbb{R}^m. \tag{45}$$

Theorem 6 *Consider the totally unimodular integer recourse function*

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

where ω is a continuous random vector with joint pdf $f \in \mathcal{H}^m$, and consider its shifted LP-relaxation approximation $\tilde{Q}$ *defined in* (45). *Then,*

$$\sup_{z\in\mathbb{R}^m} |Q(z) - \tilde{Q}(z)| \leq \frac{1}{2}\sum_{i=1}^m \lambda_i^* \mathbb{E}_{\omega_{(i)}}\Big[h\Big(|\Delta| f_i(\cdot|\omega_{(i)})\Big)\Big],$$

where $\lambda_i^* := \max_{k=1,\ldots,K} \lambda_i^k$ and h is defined in (25).

Proof Using the dual representation

$$\tilde{Q}(z) = \mathbb{E}_f\Big[\max_{k=1,\ldots,K} \lambda^k(\omega + \frac{1}{2}e_m - z)\Big], \quad z \in \mathbb{R}^m,$$

for $\tilde{Q}$, and (31) for Q, we can apply Theorem 2 for every $z \in \mathbb{R}^m$ with $H^1(\omega) = \omega + \frac{1}{2}e_m - z$ and $H^2(\omega) = \lceil \omega - z \rceil$ yielding

$$\varphi_i(x_i) = H_i^1(x_i) - H_i^2(x_i) = x_i + 1/2 - \lceil x_i \rceil_{z_i}, \quad x_i \in \mathbb{R},$$

for every $i = 1, \ldots, m$. The functions φ_i are the same as $\bar{\varphi}_\alpha$ in (26) of Example 2 with $\alpha := z_i$, so that using (27) and $N(\bar{\varphi}_\alpha, B) = -N(\bar{\varphi}_\alpha, B)$ it follows immediately that for every $z \in \mathbb{R}^m$,

$$|Q(z) - \tilde{Q}(z)| \leq \frac{1}{2}\sum_{i=1}^m \lambda_i^* \mathbb{E}_{\omega_{(i)}}\Big[h\Big(|\Delta| f_i(\cdot|\omega_{(i)})\Big)\Big].$$

□

Although the error bound of $\tilde{Q}$ is a factor 2 better than that of $\bar{Q}_\alpha$, we do not necessarily prefer $\tilde{Q}$ to $\bar{Q}_\alpha$. Both being continuous recourse functions, $\tilde{Q}$ involves continuous random variables whereas those of $\bar{Q}_\alpha$ are discrete. Indeed, this implies that solving the approximating problem with $\tilde{Q}$ is much more demanding than that with $\bar{Q}_\alpha$. In order to solve the former problem, discrete approximations of the distribution are required (except in special cases).

Surprisingly, $\bar{Q}_\alpha$ may also be obtained by applying the ρ-size mid-point approximation of Sect. 4.3.1 to $\tilde{Q}$ with $\rho = e_m$ and $\alpha \in \mathbb{R}^m$. Hence, half of the (worst-case) error bound of the α-approximations can be contributed to approximating a non-convex TU integer recourse model by a convex continuous recourse model and the remainder to applying a discrete approximation to a continuous distribution.

Note that it is also possible to apply a more refined discrete approximation to $\tilde{Q}$, such as the ρ-size mid-point approximation where the components of ρ are small, or the Jensen approximation, yielding a better approximation than $\bar{Q}_\alpha$. Based on a tradeoff between available computation time and required quality of the approximation one may decide which discrete approximation to use.

4.4.1 Shifted LP-relaxation: best worst-case error bound

In Theorem 7 we show for the special case of (one-dimensional) simple integer recourse, i.e., for the case $W = I_m$ as introduced in [22], that surprisingly the error bound of the shifted LP-relaxation approximation is the best worst-case bound possible for *any* convex approximation. This result can easily be generalized to the m-dimensional simple integer case (similar to the proof of Corollary 2 in [27], where we show that the error bound of α-approximations in Theorem 5 is tight in the simple integer recourse case) showing that indeed the shifted LP-relaxation approximation is the best convex approximation in a worst-case sense.

Theorem 7 *Let $Q : \mathbb{R} \mapsto \mathbb{R}$ denote the simple integer recourse function defined as $Q(z) = \mathbb{E}_f[\lceil \omega - z \rceil^+]$, $z \in \mathbb{R}$. Then, for every $B \in \mathbb{R}$ with $B > 0$, there exists $f \in \mathcal{F}$ with $|\Delta| f = B$ such that*

$$\sup_{z \in \mathbb{R}} |Q(z) - \hat{Q}(z)| \geq \frac{1}{2} h(B) \tag{46}$$

for every convex function $\hat{Q} : \mathbb{R} \mapsto \mathbb{R}$, where h is defined in (25). Thus, the shifted LP-relaxation approximation gives the best worst-case error bound in simple integer recourse.

Proof Let $\hat{Q} : \mathbb{R} \mapsto \mathbb{R}$ be a convex function. Since the simple integer recourse function Q restricted to $\mathbb{Z}$ is convex (see Lemma 3.1 in [19]) with $Q(z) - Q(z-1) \in [-1, 0]$ for every $z \in \mathbb{Z}$, it follows immediately that $\|Q - \hat{Q}\|_\infty = +\infty$ unless the right-derivative $\hat{Q}'_+(z)$ of $\hat{Q}$ is contained in $[-1, 0]$ for every $z \in \mathbb{R}$, which we assume without loss. Under this condition we will construct for every $B > 0$ a pdf $f \in \mathcal{F}$ with $|\Delta| f = B$ such that there exists $z \in \mathbb{R}$ with $|Q(z) - \hat{Q}(z)| \geq \frac{1}{2} h(B)$. We will consider the cases $B \leq 4$ and $B \geq 4$ separately.

First, let $B \geq 4$ be given, and let ω be uniformly distributed on $[0, 2/B]$ with density $\hat{f} : \mathbb{R} \mapsto \mathbb{R}$ defined as

$$\hat{f}(x) = \begin{cases} B/2, & 0 < x < 2/B, \\ 0, & \text{otherwise}, \end{cases}$$

with $|\Delta|\hat{f} = B$. Moreover, simple computation based on this density shows that $Q(0) = 1$ and $Q(2/B) = 0$. Since $\hat{Q}'_+(z) \in [-1, 0]$ for every $z \in \mathbb{R}$, it follows that $\hat{Q}(0) - \hat{Q}(2/B) \in [0, 2/B]$. Thus,

$$
\begin{aligned}
&|Q(0) - \hat{Q}(0)| + |\hat{Q}(2/B) - Q(2/B)| \\
&\quad \geq \Big(Q(0) - Q(2/B)\Big) - \Big(\hat{Q}(0) - \hat{Q}(2/B)\Big) \geq h(B),
\end{aligned}
$$

since $h(B) = 1-2/B$ for $B \geq 4$, and we conclude that either $|Q(0)-\hat{Q}(0)| \geq \frac{1}{2}h(B)$, or $|\hat{Q}(2/B) - Q(2/B)| \geq \frac{1}{2}h(B)$, or both hold with equality.

Next, consider the case $B \leq 4$, and define $\hat{f} : \mathbb{R} \mapsto \mathbb{R}$ as

$$
\hat{f}(x) = \begin{cases} B/2, & 1/2 < x \leq k^*, \\ c^*, & k^* < x \leq k^* + 1 \\ 0, & \text{otherwise}, \end{cases}
$$

where $k^* \in \mathbb{Z}$ with $k^* \geq 1$, and $0 \leq c^* \leq B/2$ are defined such that

$$
\frac{B}{2}\left(k^* - \frac{1}{2}\right) + c^* = 1. \tag{47}
$$

Such a k^* and c^* exist for every $B \leq 4$, with $0 \leq c^* \leq B/2$ ensuring that $|\Delta|\hat{f} = B$ and (47) that $\hat{f}$ is indeed a pdf. Moreover, straightforward computation based on $\hat{f}$ shows that $Q(-1/2) = \mu+1+B/16$ and $Q(0) = \mu+1/2-B/16$, where $\mu := \mathbb{E}_{\hat{f}}[\omega]$. Hence, $Q(-1/2) - Q(0) = 1/2 + B/8$. Since $\hat{Q}'_+(z) \in [-1, 0]$ for every $z \in \mathbb{R}$, it follows that $\hat{Q}(-1/2) - \hat{Q}(0) \in [0, 1/2]$, so that

$$
\begin{aligned}
|Q(-1/2) - \hat{Q}(-1/2)| + |\hat{Q}(0) - Q(0)| &\geq \Big(Q(-1/2) - Q(0)\Big) \\
&\quad -\Big(\hat{Q}(-1/2) - \hat{Q}(0)\Big) \geq h(B),
\end{aligned}
$$

since $h(B) = B/8$ for $B \leq 4$. Thus, either $|Q(-1/2) - \hat{Q}(-1/2)| \geq \frac{1}{2}h(B)$, or $|\hat{Q}(0) - Q(0)| \geq \frac{1}{2}h(B)$, or both hold with equality.

Since the error bound for the shifted LP-relaxation approximation in Theorem 6 reduces to $\frac{1}{2}h(|\Delta|f)$ for this special case, it follows immediately that this approximation gives the best worst-case error bound. □

4.5 Lipschitz constant for pure integer recourse models

Continuity properties of mixed-integer expected value functions Q have been studied in [30]. By carefully analyzing properties of the underlying mixed-integer value function Schultz shows that under certain conditions, the recourse function Q is Lipschitz continuous. However, in the proof it is only shown that a finite-valued Lipschitz constant exists. Here, we derive a Lipschitz constant for the pure integer recourse function

Q, depending on the underlying second-stage value function v and the total variations of the marginal densities of f. Once more we rely on the total variation bounds on the expectation of periodic functions, complemented with subadditivity of v and related properties.

Lemma 12 *Let v be the value function of a pure integer program defined as*

$$v(s) = \min\{qy : Wy \geq s,\ y \in \mathbb{Z}_+^{n_2}\}, \quad s \in \mathbb{R}^m. \tag{48}$$

Then,

(i) $v(s+t) \leq v(s) + v(t)$ *for every* $s, t \in \mathbb{R}^m$,
(ii) $v(rs) \leq rv(s)$ *for every* $r \in \mathbb{Z}_+$ *and* $s \in \mathbb{R}^m$,
(iii) $v(\sum_{i=1}^N r_i s_i) \leq \sum_{i=1}^N r_i v(s_i)$ *for every* $r_1, \ldots, r_N \in \mathbb{Z}_+$ *and* $s_1, \ldots, s_N \in \mathbb{R}^m$.

Proof See Proposition 2.3 in Chapter II.3 of [24] for a proof of (i). Results (ii) and (iii) follow immediately from (i). □

Corollary 4 *Consider v as defined in (48). Then, for every $s, t \in \mathbb{Z}^m$,*

$$v(s) - v(t) \leq \sum_{i=1}^m (s_i - t_i)^+ v(e_i),$$

where e_i denotes the ith unit vector.

Proof Let $s, t \in \mathbb{Z}^m$ be given. By subadditivity of v in Lemma 12 (i), we have

$$v(s) \leq v(t) + v(s-t). \tag{49}$$

For $i = 1, \ldots, m$, let $r_i = s_i - t_i$ and $s^i = e_i$ if $s_i \geq t_i$, and let $r_i = t_i - s_i$ and $s^i = -e_i$ if $s_i < t_i$. Observe that $r_i \in \mathbb{Z}_+$ for all $i = 1, \ldots, m$, and that $s - t = \sum_{i=1}^m r_i s^i$ so that

$$v(s-t) \leq \sum_{i=1}^m r_i v(s^i) = \sum_{i=1}^m \Big((s_i - t_i)^+ v(e_i) + (t_i - s_i)^+ v(-e_i) \Big) \tag{50}$$

by Lemma 12 (iii). The claim now follows from substituting (50) into (49) and by observing that $v(-e_i) \leq 0$ for all $i = 1, \ldots, m$, since $y = 0$ is a feasible solution in $v(-e_i)$. □

Theorem 8 *Consider the pure integer recourse function*

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

where ω is a continuous random vector with joint pdf $f \in \mathcal{H}^m$, and assume that $W \in \mathbb{Z}^{m \times n_2}$. Under assumptions (A1) and (A2) it holds for every $z^1, z^2 \in \mathbb{R}^m$ that

$$|Q(z^1) - Q(z^2)| \leq \sum_{i=1}^m v(e_i) \left(1 + \frac{|\Delta| f_i}{2} \right) |z_i^1 - z_i^2|, \tag{51}$$

and thus $|Q(z^1) - Q(z^2)| \leq L \|z^1 - z^2\|_1$, with Lipschitz constant

$$L := \max_{i=1,\ldots,m} \left\{ v(e_i) \left(1 + \frac{|\Delta| f_i}{2} \right) \right\}.$$

Proof First, we show that the inequality (51) holds if we omit the absolute value on the left-hand side. The claim then follows by symmetry.

Let $z^1, z^2 \in \mathbb{R}^m$ be given. Since $W \in \mathbb{Z}^{m \times n_2}$, it follows that for all $s \in \mathbb{R}^m$, $v(s) = v(\lceil s \rceil)$. Thus, by applying Corollary 4 we have

$$\begin{aligned} Q(z^1) - Q(z^2) &= \mathbb{E}_f \left[v\left(\left\lceil \omega - z^1 \right\rceil \right) - v\left(\left\lceil \omega - z^2 \right\rceil \right) \right] \\ &\leq \mathbb{E}_f \left[\sum_{i=1}^m \left(\left\lceil \omega_i - z_i^1 \right\rceil - \left\lceil \omega_i - z_i^2 \right\rceil \right)^+ v(e_i) \right] \\ &= \sum_{i=1}^m v(e_i) \mathbb{E}_{f_i} \left[\left(\left\lceil \omega_i - z_i^1 \right\rceil - \left\lceil \omega_i - z_i^2 \right\rceil \right)^+ \right]. \end{aligned}$$

Observe that

$$\mathbb{E}_{f_i} \left[(\left\lceil \omega_i - z_i^1 \right\rceil - \left\lceil \omega_i - z_i^2 \right\rceil)^+ \right] = \begin{cases} \mathbb{E}_{f_i} \left[(\lceil \omega_i - z_i^1 \rceil - \lceil \omega_i - z_i^2 \rceil) \right], & z_i^1 \leq z_i^2, \\ 0, & z_i^1 \geq z_i^2. \end{cases}$$

Thus, we have

$$Q(z^1) - Q(z^2) \leq \sum_{i=1}^m \mathbb{1}_{\{z_i^1 \leq z_i^2\}} v(e_i) \mathbb{E}_{f_i} \left[\left(\left\lceil \omega_i - z_i^1 \right\rceil - \left\lceil \omega_i - z_i^2 \right\rceil \right) \right],$$

where $\mathbb{1}_{\{z_i^1 \leq z_i^2\}}$ is an indicator function equal to 1 if $z_i^1 \leq z_i^2$, and 0 otherwise. It remains to show that $\mathbb{E}_{f_i} \left[(\lceil \omega_i - z_i^1 \rceil - \lceil \omega_i - z_i^2 \rceil) \right] \leq (1 + \frac{|\Delta| f_i}{2}) |z_i^1 - z_i^2|$ if $z_i^1 \leq z_i^2$. In order to do so, observe that $\mathbb{E}_{f_i} \left[(\lceil \omega_i - z_i^1 \rceil - \lceil \omega_i - z_i^2 \rceil) \right] = \mathbb{E}_{f_i} \left[\varphi_{\alpha,\beta}(\omega_i) \right] + \beta - \alpha$, with $\varphi_{\alpha,\beta}$ defined in (22) of Example 1 with $\alpha := z_i^1$ and $\beta := z_i^2$. It follows from (23) that

$$\begin{aligned} \mathbb{E}_{f_i} \left[\left(\left\lceil \omega_i - z_i^1 \right\rceil - \left\lceil \omega_i - z_i^2 \right\rceil \right) \right] &\leq \min \left\{ \gamma_{z_i^1, z_i^2}, \gamma_{z_i^1, z_i^2} (1 - \gamma_{z_i^1, z_i^2}) \frac{|\Delta| f_i}{2} \right\} + z_i^2 - z_i^1 \\ &\leq (1 - \gamma_{z_i^1, z_i^2}) \frac{|\Delta| f_i}{2} + z_i^2 - z_i^1, \end{aligned}$$

where the last inequality is true since $\gamma_{z_i^1, z_i^2} := z_i^1 + 1 - \left\lceil z_i^1 \right\rceil_{z_i^2} \in (0, 1]$. Since $z_i^1 \leq z_i^2$,

$$\begin{aligned} 1 - \gamma_{z_i^1, z_i^2} &= \left\lceil z_i^1 \right\rceil_{z_i^2} - z_i^1 = \left\lceil z_i^1 - z_i^2 \right\rceil + z_i^2 - z_i^1 \\ &\leq z_i^2 - z_i^1, \end{aligned}$$

and thus

$$\mathbb{E}_{f_i}\left[\left(\left\lceil \omega_i - z_i^1 \right\rceil - \left\lceil \omega_i - z_i^2 \right\rceil\right)\right] \leq (z_i^2 - z_i^1)\frac{|\Delta| f_i}{2} + z_i^2 - z_i^1$$
$$\leq \left(1 + \frac{|\Delta| f_i}{2}\right) |z_i^1 - z_i^2|.$$

□

5 Summary and conclusions

We use worst-case analysis to derive upper and lower bounds for the expectation $\mathbb{E}_f[\varphi(\omega)]$ of periodic functions φ depending on the total variation $|\Delta|f$ of the probability density function f of the underlying random variable ω. For periodically monotone functions we obtain exact worst-case bounds. The bounds show that if the total variation $|\Delta|f$ of the pdf f is small, then the expectation $\mathbb{E}_f[\varphi(\omega)]$ will be close to the mean value $\nu := p^{-1} \int_0^p \varphi(x)dx$ of the periodic function φ with period p. These results are derived in a general setting so that they can readily be applied in problems involving both uncertainty and periodicity.

We apply these total variation bounds to approximations of recourse models, using additional analysis to be able to apply these one-dimensional bounds in a multi-dimensional setting. Interestingly, the same analysis is used to obtain error bounds for approximations of both continuous and integer recourse models.

For continuous recourse models with continuous random variables we derive error bounds for two types of discrete approximations: the so-called mid-point approximation and the Jensen approximation. In general, the smaller the total variations of the densities of the random variables in the model are, the smaller the error bounds.

For totally unimodular integer recourse models we introduce the so-called shifted LP-relaxation approximation. Its error bound improves the bound for α-approximations of [33]—the best convex approximations known so far—by a factor 2. Moreover, we show that, in a worst-case sense, this shifted LP-relaxation provides the best convex approximation possible. Interestingly, the α-approximations can be obtained by applying the mid-point approximation to the shifted LP-relaxation approximation. This implies that half of the (worst-case) α-approximation error can be contributed to coping with non-convexity, whereas the other half accounts for approximating continuous distributions by discrete ones.

Finally, we use total variation bounds to derive a tractable Lipschitz constant for the pure integer recourse function Q.

Extensions to multistage models may be considered for future research. Another research direction is to assess the actual performance of the approximations (compared to their error bounds) in a numerical study.

Appendix

In this appendix we give the proofs of Lemmas 1 and 2 of Sect. 1.1.

Proof of Lemma 1 Properties (i) and (ii) follow trivially from the definition of M and N.

To show (iii) let $r > 0$ be given, and consider $\bar{\varphi}_r(x) := \varphi(x/r),\ x \in \mathbb{R}$. We will show that for every $f \in \mathcal{F}$ with $|\Delta| f \le B$, there exists $g \in \mathcal{F}$ with $|\Delta| g \le rB$ such that $\mathbb{E}_f[\bar{\varphi}_r(\omega)] = \mathbb{E}_g[\varphi(\omega)]$, and vice versa, that for every $f \in \mathcal{F}$ with $|\Delta| f \le rB$ there exists $g \in \mathcal{F}$ with $|\Delta| g \le B$ such that $\mathbb{E}_f[\varphi(\omega)] = \mathbb{E}_g[\bar{\varphi}_r(\omega)]$. Together these results imply that (iii) holds. Observe that for $f \in \mathcal{F}$ with $|\Delta| f \le B$ the pdf g defined as $g(x) := rf(rx),\ x \in \mathbb{R}$ satisfies the first conditions and for $f \in \mathcal{F}$ with $|\Delta| f \le rB$ the pdf g defined as $g(x) := r^{-1} f(x/r)$ satisfies the latter.

Similarly, for $\hat{\varphi}(x) := \varphi(-x),\ x \in \mathbb{R}$, let $f \in \mathcal{F}$ with $|\Delta| f \le B$ be given. Then, $g(x) := f(-x),\ x \in \mathbb{R}$, satisfies $g \in \mathcal{F}$ with $|\Delta| g = |\Delta| f \le B$ and $\mathbb{E}_f[\hat{\varphi}(\omega)] = \mathbb{E}_g[\varphi(\omega)]$, implying

$$M(\hat{\varphi}, B) \ge M(\varphi, B) \quad \text{and} \quad N(\hat{\varphi}, B) \le N(\varphi, B).$$

Since $\varphi(x) = \hat{\varphi}(-x),\ x \in \mathbb{R}$, the reverse inequalities hold as well, proving (iv).

Finally, property (v) can be proven in a similar way, using that for every $f \in \mathcal{F}$ with $|\Delta| f \le B$, and $\beta \in \mathbb{R}$, the pdf $g(x) := f(x+\beta),\ x \in \mathbb{R}$, satisfies $|\Delta| g = |\Delta| f \le B$. □

Proof of Lemma 2 Since for every $f^0, f^1 \in \mathcal{F}$ with $|\Delta| f^0 \le B$ and $|\Delta| f^1 \le B$, and $0 \le t \le 1$, the pdf $f := (1-t) f^0 + t f^1$ satisfies

$$|\Delta| f \le (1-t)|\Delta| f^0 + t|\Delta| f^1 \le B,$$

it follows that the constraint $|\Delta| f \le B$ in $M(\varphi, B)$ and $N(\varphi, B)$ is convex. Since, in addition, the objective $\mathbb{E}_f[\varphi(\omega)]$ is linear in f, we conclude that both $M(\varphi, B)$ and $N(\varphi, B)$ are convex optimization problems with a linear objective. Since $M(\varphi, B)$ is a maximization problem and $N(\varphi, B)$ is a minimization problem, it follows that $M(\varphi, B)$ is concave in B, and $N(\varphi, B)$ is convex in B, respectively. □

References

1. Birge, J.R.: Aggregation bounds in stochastic linear programming. Math. Program. **31**, 25–41 (1985)
2. Birge, J.R., Louveaux, F.V.: Introduction to Stochastic Programming. Springer, New York (1997)
3. Birge, J.R., Wallace, S.W.: A separable piecewise linear upper bound for stochastic linear programs. SIAM J. Control Optim. **26**, 725–739 (1988)
4. Birge, J.R., Wets, R.J.-B.: Designing approximation schemes for stochastic optimization problems, in particular for stochastic programs with recourse. Math. Program. Study **27**, 54–102 (1986)
5. Birge, J.R., Wets, R.J.-B.: Computing bounds for stochastic programming problems by means of a generalized moment problem. Math. Oper. Res. **12**, 149–162 (1987)
6. Dupačová, J., Consigli, G., Wallace, S.W.: Scenarios for multistage stochastic programs. Ann. Oper. Res. **100**, 25–53 (2000)
7. Dupačová, J., Gröwe-Kuska, N., Römisch, W.: Scenario reduction in stochastic programming: an approach using probability metrics. Math. Program. **95**, 493–511 (2003)
8. Edirisinghe, N.C.P., Ziemba, W.T.: Bounds for two-stage stochastic programs with fixed recourse. Math. Oper. Res. **19**, 292–313 (1994)
9. Edmundson, H.P.: Bounds on the Expectation of a Convex Function of a Random Variable. Technical Report Paper 982, The RAND Corporation, Santa Monica (1956)
10. Frauendorfer, K.: Solving SLP recourse problems with arbitrary multivariate distributions—the dependent case. Math. Oper. Res. **13**, 377–394 (1988)

11. Gassmann, H., Ziemba, W.T.: A tight upper bound for the expectation of a convex function of a multivariate random variable. Math. Program. Study **27**, 39–53 (1986)
12. Gassmann, H.I., Ziemba, W.T. (eds.): Stochastic Programming: Applications in Finance, Energy, Planning and Logistics. World Scientific Series in Finance: Volume 4, World Scientific, Singapore (2013)
13. Heitsch, H., Römisch, W.: Scenario reduction algorithms in stochastic programming. Comput. Optim. Appl. **24**, 187–206 (2003)
14. Higle, J.L., Sen, S.: Stochastic Decomposition: A Statistical Method for Large Scale Stochastic Linear Programming. Kluwer, Dordrecht (1996)
15. Jensen, J.L.: Sur les fonctions convexes et les inégalités entre les valeurs moyennes. Acta Math. **30**, 175–193 (1906)
16. Kall, P.: On approximations and stability in stochastic programming. In: Guddat, J., Jongen, H.T., Kummer, B., Nožička, F. (eds.) Parametric Optimization and Related Topics, pp 387–407. Akademie-Verlag, Berlin (1987)
17. Kall, P.: An upper bound for SLP using first and total second moments. Ann. Oper. Res. **30**, 267–276 (1991)
18. Kaut, M., Wallace, S.W.: Evaluation of scenario-generation methods for stochastic programming. Pac. J. Optim. **3**, 257–271 (2007)
19. Klein Haneveld, W.K., Stougie, L., van der Vlerk, M.H.: Simple integer recourse models: convexity and convex approximations. Math. Program. **108**, 435–473 (2006)
20. Klein Haneveld, W.K., van der Vlerk, M.H.: Stochastic integer programming: general models and algorithms. Ann. Oper. Res. **85**, 39–57 (1999)
21. Louveaux, F.V., Schultz, R.: Stochastic integer programming. In: Ruszczyński, A., Shapiro, A. (eds.) Handbooks in Operations Research and Management Science, Volume 10: Stochastic Programming, pp. 213–266. Elsevier, Amsterdam (2003)
22. Louveaux, F.V., van der Vlerk, M.H.: Stochastic programming with simple integer recourse. Math. Program. **61**, 301–325 (1993)
23. Madansky, A.: Bounds on the expectation of a convex function of a multivariate random variable. Ann. Math. Stat. **30**, 743–746 (1959)
24. Nemhauser, G.L., Wolsey, L.A.: Integer and Combinatorial Optimization. Wiley, New York (1988)
25. Pflug, G.Ch.: Scenario tree generation for multiperiod financial optimization by optimal discretization. Math. Program. **89**, 251–271 (2001)
26. Rinnooy Kan, A.H.G., Stougie, L.: Stochastic integer programming. In: Ermoliev, Yu., Wets, R.J.-B. (eds.) Numerical Techniques for Stochastic Optimization, Volume 10 of Springer Ser. Comput. Math, pp. 201–213. Springer, Berlin (1988)
27. Romeijnders, W., van der Vlerk, M.H., Klein Haneveld, W.K.: Convex approximations of totally unimodular integer recourse models: a uniform error bound (submitted). See also SPePS (2013)
28. Römisch, W.: Stability of stochastic programming problems. In: Ruszczyński, A., Shapiro, A. (eds.) Handbooks in Operations Research and Management Science, Volume 10: Stochastic Programming, pp. 483–554. Elsevier, Amsterdam (2003)
29. Rudin, W.: Real and Complex Analysis. McGraw-Hill, New York (1987)
30. Schultz, R.: Continuity properties of expectation functions in stochastic integer programming. Math. Oper. Res. **18**, 578–589 (1993)
31. Sen, S.: Algorithms for stochastic mixed-integer programming models. In: Aardal, K., Nemhauser, G.L., Weismantel, R. (eds.) Handbooks in Operations Research and Management Science, Volume 12: Discrete Optimization, pp. 515–558. Elsevier, Amsterdam (2005)
32. Shapiro, A., Dentcheva, D., Ruszczyński, A.: Lectures on Stochastic Programming: Modeling and Theory. SIAM, Philadelphia (2009)
33. van der Vlerk, M.H.: Convex approximations for complete integer recourse models. Math. Program. **99**, 297–310 (2004)
34. van Slyke, R.M., Wets, R.: L-shaped linear programs with applications to optimal control and stochastic programming. SIAM J. Appl. Math. **17**, 638–663 (1969)
35. Wallace, S.W., Ziemba, W.T. (eds.).: Applications of Stochastic Programming. MPS-SIAM Series on Optimization (2005)