DEPARTMENT OF MATHEMATICS,
STATISTICS AND ACTUARIAL SCIENCES

Optimal Moment Bounds under Multiple Shape Constraints

Geert Wouters & Ann De Schepper

UNIVERSITY OF ANTWERP
Faculty of Applied Economics
Stadscampus
Prinsstraat 13, B.213
BE-2000 Antwerpen
Tel. +32 (0)3 265 40 32
Fax +32 (0)3 265 47 99
http://www.ua.ac.be/tew

FACULTY OF APPLIED ECONOMICS

DEPARTMENT OF MATHEMATICS,
STATISTICS AND ACTUARIAL SCIENCES

Optimal Moment Bounds under Multiple Shape Constraints

Geert Wouters & Ann De Schepper

RESEARCH PAPER 2009-005
JUNE 2009

University of Antwerp, City Campus, Prinsstraat 13, B-2000 Antwerp, Belgium
Research Administration – room B.213
phone: (32) 3 265 40 32
fax: (32) 3 265 47 99
e-mail: joeri.nys@ua.ac.be

The papers can be also found at our website:
www.ua.ac.be/tew
(research > working papers)

D/2009/1169/005

Optimal Moment Bounds under Multiple Shape Constraints

Geert Wouters
University of Antwerp, Faculty of Applied Economics,
Department of Mathematics, Statistics and Actuarial Sciences, Antwerp, Belgium
email: geert.wouters2@ua.ac.be

Ann De Schepper
University of Antwerp, Faculty of Applied Economics,
Department of Mathematics, Statistics and Actuarial Sciences, Antwerp, Belgium
email: ann.deschepper@ua.ac.be

Consider the problem of computing the optimal lower and upper bound for the expected value $E[\phi(X)]$, where X is an uncertain random probability variable. This paper studies the case in which the density of X is restricted by multiple shape constraints, each imposed on a different subset of the domain. We derive (closed) convex hull representations that allow us to reduce the optimization problem to a class of generating measures that are composed of convex sums of local probability measures. Furthermore, the notion of mass constraints is introduced to spread out the probability mass over the entire domain. A generalization to mass uncertainty is discussed as well.

Key words: probability bounds; shape constraints; convex optimization

MSC2000 Subject Classification: Primary: 60E15; Secondary: 90C34, 28A99

OR/MS subject classification: Primary: Probability Distributions; Secondary: Mathematics Convexity, Infinite Dimensional Programming

---

1. Introduction Consider a multivariate random variable X on a closed domain $\Omega \subseteq \mathbb{R}^M$. Suppose that X is partially known by moment information $E[f(X)] = \mu$. We call $f = (f_0, \ldots, f_N)'$ the vector of moment functions and $\mu = (\mu_0, \ldots, \mu_N)'$ the corresponding moments. It is tacitly assumed that $f_0 = 1_\Omega$ and $\mu_0 = 1$ in order to normalize X. Popular choices for f are the power moment functions $f_n(x) = x^n$ or the piecewise linear functions $f_n(x) = (x - k_n)_+$, with $0 = k_1 < \ldots < k_N$. The main goal is to find an optimal lower and upper bound on $E[\phi(X)]$, given the incomplete information on X. This can be stated as the following optimization program

$$\text{(P1)} \qquad \max_{X \in \mathcal{P}} \quad E[\phi(X)] \qquad s.t. \quad E[f(X)] = \mu,$$

where X is optimized over random probability variables belonging to an appropriate class $\mathcal{P}$. Note that this program calculates the upper bound; to obtain the lower bound the max-operator should be replaced by the min-operator.

This formulation as an optimization program covers a wide spectrum of specific instances. In finance, the function ϕ can be interpreted as the payoff function of a path-independent option. The solution of (P1) provides a safe bound for the option price (see e.g. Bertsimas and Popescu [1]). Boyle and Lin [3] computed bounds on an option with multivariate payoff function under moment information. In Han et al. [10] options on multi-assets were considered. In risk management the setting $\phi(x) = 1_{[0,k]}(x)$ is used to calculate conservative bounds for the value-at-risk (see Cox et al. [5], El Ghaoui et al. [8]).

For particular choices of ϕ and f it is possible to express the bounds analytically. Early results were achieved by Chebychev and Markov, who derived the so-called Chebychev bounds (or Chebychev type inequalities), in the late 1800s. Their results are situated in the field of probability. The first attempt to obtain a closed-form formula for the bound on a call option price under moment constraints is due to Lo [12]. More recent results that generalize Lo's are for instance given in Zuluaga et al. [26] and De Schepper and Heijnen [6]. In actuarial literature, we refer to De Vylder and Goovaerts [7], where analytic expressions for bounds on stop-loss premiums were derived under various types of information. In Goovaerts et al. [9] a link with optimization and duality theory is established in an attempt to solve complicated instances of (P1) in a numerical way.

Bounds under power moment information have received the most attention in the literature. However, working under plain moment constraints has several drawbacks. First, deriving analytic expressions is only feasible for a small number of moments, as the formulas become extremely complicated in the general case. Second, such bounds are often not sharp enough for financial applications, mostly providing an overconservative estimate. The reason for that lays in the fact that they are reached by discrete distributions (see e.g. Rogosinsky [20]) which are typically far from realistic. Third, when dealing with heavy-tailed distributions, one cannot take for granted that all power moments exist.

This motivates the study to take into account structural properties of the underlying density in order to sharpen the bounds. Concretely, this paper explores the combination of mass and shape constraints on the probability density function of X. Considering (P1), this additional information is contained in the set $\mathcal{P}$ and will be imposed jointly with the moment information.

This paper is not the first one to take shape information into consideration. Some of the forementioned papers [6, 7] reported explicit bounds under unimodality (with fixed mode m). A more general concept is that of bell-shapedness, explored by Mallows [13, 14], Mulholland and Rogers [15] and Winkler [25]. Unfortunately, an efficient computational approach lacks in their work.

At the end of the twentieth century, Bertsimas and Popescu proposed a semidefinite programming (SDP) approach that could solve several interesting instances of the generalized moment problem in polynomial time [1, 2]. Popescu extended these results in [17] to a framework for optimizing over a general shape constraint. Using conic duality theory, she succesfully derives bounds under unimodality, symmetry or second order information, such as convexity or concavity. The main idea is to describe a shape-constrained measure set with a convex hull or integral representation. In simple cases the generating measure set can be indexed by a single, one-dimensional parameter; in such a case, we call this measure set elementary.

The contribution of this paper is to provide a general approach to optimize over a number of specific shape constraints, each one imposed on a different subset $\Omega^{(j)}$ of Ω. We will refer to $\Omega^{(j)}$ as a subdomain. A shape constraint imposed on a subdomain is called local. This generalization is particularly interesting for applications, because in practice densities have a more complicated structure than can be described with a single, elementary shape constraint. In finance indeed we deal with densities that are partly increasing, decreasing, convex and concave. Furthermore, we will show that it is convenient to impose a mass constraint on each of the subdomains, to guarantee that the probability mass is spread accordingly over all subdomains.

It is shown by Popescu [17] that, under certain conditions, the optimization over an elementary shape constraint can be performed by a single SDP. The use of a semidefinite solver (such as SeDuMi [22] or SDPT3 [23]) enables one to compute the bounds numerically in an efficient manner. As will be pointed out, the extension to multiple shape constraints entails optimization over measure sets indexed by multiple parameters. In future papers we will show that, under certain conditions, it is still possible to reduce such programs to a single SDP.

This paper is organized as follows. Section 2 will guide the reader through some important results achieved by Popescu [17], which will be built upon in later sections. Section 3 contains our main results on the combination of multiple and local shape constraints. Section 4 describes the generalization to mass uncertainty.

2. Preliminary Results Let us first introduce some notations. The cardinality of an arbitrary set W is denoted by $|W|$. For $W \subseteq \Omega$, the indicator function 1_W takes the value 1 if $x \in W$ and 0 elsewhere. Supposing that W is a subset of a topological space, denote its closure and interior by $cl(W)$ and $int(W)$ respectively. The boundary of W is defined as $\partial(W) = cl(W) \backslash int(W)$. In real sets we work under the standard topology and in measure sets we work under the weak topology. Convergence of reals is denoted by $x_n \to x$, whereas weak convergence of measures is denoted by $\pi_n \overset{w}{\to} \pi$.

We adopt the usual brackets $\lceil . \rceil$ for the cap function. The notation x_+ stands for $\max\{0, x\}$. The transpose of a matrix M is denoted with a prime (M').

A cone C, belonging to a real vector space $\mathcal{W}$, is a set that is closed under positive scalar multiplication, that is $\lambda w \in C$ if $\lambda > 0$ and $w \in C$. Note that the origin is not necessarily included. For two vector spaces $\mathcal{W}$ and $\mathcal{W}^*$, paired with a bilinear form $\langle \cdot, \cdot \rangle : \mathcal{W} \times \mathcal{W}^* \to \mathbb{R}$, the (positive) dual of C is defined as

$$C^* = \{w^* \in \mathcal{W}^* \mid \langle w, w^* \rangle \geq 0, \, \forall w \in C\}.$$

Let $w_1, \ldots, w_R \in \mathcal{W}$ and $\lambda_1, \ldots, \lambda_R \in \mathbb{R}$. The sum $\lambda_1 w_1 + \ldots + \lambda_R w_R$ is called a linear combination. If

$\lambda_1 + \ldots + \lambda_R = 1$, it is said to be an affine combination; if $\lambda_1, \ldots, \lambda_R \geq 0$ it is a conic combination. A convex combination is a linear combination that is conic and affine. For $W \subseteq \mathcal{W}$, the convex hull $\mathcal{H}(W)$ is defined as the set of all convex combinations of elements of W, the conic hull $\mathcal{C}(W)$ is defined in an analogous way. The sets $\mathcal{H}(W)$ and $\mathcal{C}(W)$ are respectively the smallest convex set and convex cone in $\mathcal{W}$ that include W. We will say that W is a generating set for $\mathcal{H}(W)$ and $\mathcal{C}(W)$ respectively. If w belongs to a convex set $W \subseteq \mathcal{W}$, it is said to be an extreme point if it is not in the interior of a line segment in W. Denote $ex(W)$ for the set of the extreme points of W.

Let X be a probability random variable defined on a closed domain $\Omega \subseteq \mathbb{R}^M$. The set of all probability measures on Ω is denoted as $\mathcal{P}_\Omega$. Let $\mathcal{B}_\Omega$ be the Borel sigma algebra of Ω. For $W \in \mathcal{B}_\Omega$, define $\mathcal{P}_W$ as the set of probability measures with domain W, meaning $\mathcal{P}_W = \{\pi \in \mathcal{P}_\Omega \mid \pi(W) = 1\}$. The symbol $\mathcal{P}$ is used for a convex subset of $\mathcal{P}_\Omega$, containing probability measures with a special property, for instance having a unimodal density. The restriction of a measure set $\mathcal{P}$ to a set W is defined as $\mathcal{P}|_W = \mathcal{P} \cap \mathcal{P}_W$. The Minkowski sum of two measure sets $\mathcal{P}_1$ and $\mathcal{P}_2$ is defined in the usual way as $\mathcal{P}_1 + \mathcal{P}_2 = \{\pi_1 + \pi_2 \mid \pi_1 \in \mathcal{P}_1,\ \pi_2 \in \mathcal{P}_2\}$.

The random variable X defines a probability measure $\pi : X \longrightarrow [0,1]$ by $\pi(B) = P(X \in B)$ for each $B \in \mathcal{B}_\Omega$. Let $\phi, f_0, \ldots, f_N$ be measurable functions on $(\Omega, \mathcal{B}_\Omega)$ that are π-integrable for all $\pi \in \mathcal{P}$. Let $\mathcal{X}$ be the linear space of signed measures generated by $\mathcal{P}_\Omega$ and let $\mathcal{X}^*$ be the linear span of $\phi, f_0, \ldots, f_N$. The vector spaces $\mathcal{X}^*$ and $\mathcal{X}$ are paired by the bilinear form given by the integral operator $\langle \phi, \pi \rangle := \int_\Omega \phi \,\mathrm{d}\pi$. For probability measures, this corresponds to the expectation operator $E_\pi[\phi(X)] = \int_\Omega \phi \,\mathrm{d}\pi$. The space $\mathbb{R}^N$ is paired with itself with respect to the scalar product denoted as $\langle x, y \rangle := x \cdot y = \sum_{n=1}^N x_n y_n$.

Following Shapiro [21], program (P1) can be casted as the following linear conic program

(P2)
$$\max_{\pi \in \mathcal{C}(\mathcal{P})} \langle \phi, \pi \rangle$$
$$\text{s.t. } A(\pi) - \mu \in K,$$

where $\mathcal{C}(\mathcal{P})$ is the cone of (positive) measures generated by a convex set $\mathcal{P}$ of probability measures, A is a linear map from $\mathcal{X}$ to $\mathbb{R}^{N+1}$ defined by $A(\pi) = (\langle f_0, \pi \rangle, \ldots, \langle f_N, \pi \rangle)'$ and K is a closed convex cone in $\mathbb{R}^{N+1}$. Remark that the decision variable π ranges over a cone, and therefore the probability mass constraint should be explicitly included in the moment constraints ($f_0 = 1_\Omega$ and $\mu_0 = 1$). Remark as well that in (P1) all moments are fixed, whereas the construction in (P2) is more general and allows for moment uncertainty. For example, the moment inequalities $\underline{\mu}_n \leq \mu_n \leq \overline{\mu}_n$, $n = 0, \ldots, N$ can be imposed by the constraints

$$\begin{cases} A(\pi) - \underline{\mu} \in K^+ \\ A(\pi) - \overline{\mu} \in -K^+, \end{cases}$$

where $K^+ = [0, +\infty)^{N+1}$.

A first step in solving (P2), is to convert it to its conic dual

(P3)
$$\min_{y \in -K^*} \quad \mu \cdot y$$
$$s.t. \quad y'f - \phi \in (\mathcal{C}(\mathcal{P}))^*,$$

where $y'f = \sum_{n=0}^N y_n f_n$. In Shapiro [21] it is proved that strong duality holds between (P2) and (P3) under certain mild conditions (e.g. $\mu \in int(Q_f - K)$, where Q_f is the set of all feasible moments, i.e. $Q_f = \{\mu \in \mathbb{R}^{N+1} \mid \exists\, \pi \in \mathcal{P}_\Omega :\ \mu = E_\pi[f(X)]\}$). Strong duality means that the optimal values of the primal program and its dual equal each other if the optimal value can be attained. We will assume throughout the paper that the conditions for strong duality are satisfied, so we only need to consider (P3).

Remark that by the definition of the dual cone, program (P3) can be rewritten as

(P4)
$$\min_{y \in -K^*} \quad \mu \cdot y$$
$$s.t. \quad \int_\Omega (y'f - \phi) \,\mathrm{d}\pi \geq 0 \quad \forall \pi \in \mathcal{P}.$$

Here we clearly observe that the dual program is semi-infinite, meaning it has a finite number of decision variables and an infinite number of constraints. Next, we discuss a technique that reduces the number of constraints significantly for certain sets $\mathcal{P}$. The proposed technique is due to Popescu [17] and assumes that $\mathcal{P}$ admits a convex hull representation.

ASSUMPTION 2.1 *There exists a measure set $\mathcal{T} \subseteq \mathcal{P}$ that admits one of the following representations*[1]*:*

[A1] $\mathcal{P} = \mathcal{H}(\mathcal{T})$.

[A2] $\mathcal{P} = cl(\mathcal{H}(\mathcal{T}))$, *with* $\mathcal{T}$ *closed.*

[A3] $\mathcal{P} = cl(\mathcal{H}(\mathcal{T}))$, *with* f *and* ϕ *continuous and bounded.*

We call $\mathcal{T}$ a generating set for $\mathcal{P}$, or equivalently, we say that $\mathcal{T}$ generates $\mathcal{P}$. The requirement [A3] demands that f and ϕ should be bounded and continuous, which is in fact a restriction on the function space $\mathcal{X}^*$ and the cones $\mathcal{P}^*$ and $\mathcal{T}^*$. Remark that we do not require $\mathcal{T}$ to coincide with the extremal measures of $\mathcal{P}$. If Assumption 2.1 is met, Popescu shows that the constraints of program (P4) only need to be checked for the generating set, in other words (P4) reduces to

$$\text{(P5)} \qquad \begin{aligned} \min_{y \in -K^*} \quad & \mu \cdot y \\ s.t. \quad & \int_\Omega (y' f - \phi)\, \mathrm{d}\tau \geq 0 \quad \forall \tau \in \mathcal{T}. \end{aligned}$$

Next, we describe some elementary generating sets $\mathcal{T}$. In the following section we will compose them on different subdomains to build more complicated generating sets. The examples are mainly inspired by Popescu [17].

EXAMPLE 2.1 *The most basic example is given by $\mathcal{T} = \Delta_\Omega = \{\delta_t \mid t \in \Omega\}$, where δ_t denotes the Dirac measure at t. It is well-known that $\mathcal{P}_\Omega$ admits the representation $\mathcal{P}_\Omega = cl(\mathcal{H}(\Delta_\Omega))$. Remark that the closedness of Ω ensures that Δ_Ω is closed. Indexing the generating set Δ_Ω with an M-dimensional parameter t, we get the semi-infinite program*

$$\text{(P6)} \qquad \begin{aligned} \min_{y \in -K^*} \quad & \mu \cdot y \\ s.t. \quad & y' f(t) - \phi(t) \geq 0 \quad \forall t \in \Omega. \end{aligned}$$

Remark that Δ_Ω is elementary in the univariate case ($M = 1$).

EXAMPLE 2.2 *Let $\Omega = [\underline{a}, \overline{a}]$ be a real interval, where $\underline{a}$ is allowed to attain $-\infty$. Let $\mathcal{P}^{inc}_{\underline{a},\overline{a}}$ be the class of probability measures composed of a non-decreasing density and a Dirac at $\overline{a}$, where each of these two components is optional. It can be shown that $\mathcal{P}^{inc}_{\underline{a},\overline{a}}$ is a closed and convex set, which can be generated as $\mathcal{P}^{inc}_{\underline{a},\overline{a}} = cl(\mathcal{H}(\mathcal{T}^{inc}_{\underline{a},\overline{a}} \cup \{\delta_{\overline{a}}\}))$, where $\mathcal{T}^{inc}_{\underline{a},\overline{a}} = \{\tau_t \mid t \in [\underline{a}, \overline{a})\}$ and τ_t is a probability measure corresponding to the rectangular density*

$$\zeta_t(x) = \frac{1}{\overline{a} - t} 1_{[t,\overline{a}]}(x).$$

By an analogous reasoning, one can show that $\mathcal{P}^{dec}_{\underline{a},\overline{a}}$, the class of probability measures composed of a non-increasing density and a Dirac at $\underline{a}$ (each of the two components is optional), can be generated by the union of $\delta_{\underline{a}}$ and $\mathcal{T}^{dec}_{\underline{a},\overline{a}} = \{\tau_t \mid t \in (\underline{a}, \overline{a}]\}$, where τ_t is a measure corresponding to the rectangular density

$$\zeta_t(x) = \frac{1}{t - \underline{a}} 1_{[\underline{a},t]}(x).$$

EXAMPLE 2.3 *Let $\Omega = [\underline{a}, \overline{a}]$ and allow $\underline{a}$ to attain $-\infty$. Let $\mathcal{P}^{inc,cx}_{\underline{a},\overline{a}}$ be the class of probability measures composed of a non-decreasing and convex density and a Dirac at $\overline{a}$, where each of these two components is optional. It can be shown that $\mathcal{P}^{inc,cx}_{\underline{a},\overline{a}}$ is a closed and convex set, which can be generated as $\mathcal{P}^{inc,cx}_{\underline{a},\overline{a}} =$*

[1] Popescu includes a mixture (or integral) representation as well, which we will not consider in this paper.

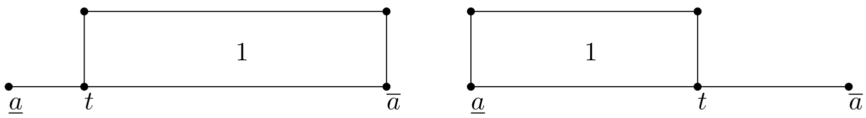

Figure 1: Illustration of the rectangular densities in Example 2.2 that generate non-decreasing densities (left) and non-increasing densities (right).

$cl(\mathcal{H}(\mathcal{T}_{\underline{a},\overline{a}}^{inc,cx} \cup \{\delta_{\overline{a}}\}))$, where $\mathcal{T}_{\underline{a},\overline{a}}^{inc,cx} = \{\tau_t \mid t \in (0, \overline{a} - \underline{a}]\}$ and τ_t is a probability measure corresponding to the left-triangular density

$$\zeta_t(x) = \frac{1}{t^2}(x - \overline{a} + t)_+ \, 1_{[\underline{a},\overline{a}]}(x).$$

If Ω is bounded, the probability measure with a rectangular density function on $[\underline{a}, \overline{a}]$ should be added to $\mathcal{T}_{\underline{a},\overline{a}}^{inc,cx}$. As Popescu in [17] further points out, it is possible to limit the slope of the underlying density at the point $x = \overline{a}$ to values in $[\alpha, \beta]$, by restricting t to $[\sqrt{2/\beta}, \sqrt{2/\alpha}]$.

By an analogous reasoning, one can show that $\mathcal{P}_{\underline{a},\overline{a}}^{dec,cx}$, the class of probability measures composed of a non-increasing and convex density and a Dirac at $\underline{a}$ (each of the two components is optional), can be generated by the union of $\delta_{\underline{a}}$ and the set $\mathcal{T}_{\underline{a},\overline{a}}^{dec,cx} = \{\tau_t \mid t \in (0, \overline{a} - \underline{a}]\}$, where τ_t is a measure corresponding to the right-triangular density

$$\zeta_t(x) = \frac{1}{t^2}(\underline{a} + t - x)_+ \, 1_{[\underline{a},\overline{a}]}(x).$$

Again, if Ω is bounded, the probability measure with a rectangular density function on $[\underline{a}, \overline{a}]$ should be added to $\mathcal{T}_{\underline{a},\overline{a}}^{dec,cx}$. A restriction on the (absolute) slope of the density function can be built in by a similar restriction on t.

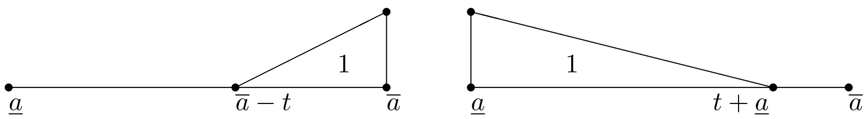

Figure 2: Illustration of the triangular densities in Example 2.3 that generate convex non-decreasing densities (left) and convex non-increasing densities (right).

Example 2.4 *Let $\Omega = [\underline{a}, \overline{a}]$ be a bounded interval. Let $\mathcal{P}_{\underline{a},\overline{a}}^{inc,cv}$ be the class of probability measures having a non-decreasing and concave density. It can be shown that $\mathcal{P}_{\underline{a},\overline{a}}^{inc,cv}$ is a closed and convex set, which can be generated as $\mathcal{P}_{\underline{a},\overline{a}}^{inc,cv} = cl(\mathcal{H}(\mathcal{T}_{\underline{a},\overline{a}}^{inc,cv}))$, where $\mathcal{T}_{\underline{a},\overline{a}}^{inc,cv} = \{\tau_t \mid t \in [0, \overline{a} - \underline{a}]\}$ with τ_t the probability measure corresponding to the left-trapezoidal density*

$$\zeta_t(x) = \frac{2}{\overline{a} - \underline{a} + t} \min\left(1, \frac{x - \underline{a}}{\overline{a} - \underline{a} - t}\right) 1_{[\underline{a},\overline{a}]}(x).$$

Again, we can limit the slope of the underlying density at the point $x = \underline{a}$ to values in $[\alpha, \beta]$, by restricting t to $[\sqrt{(\overline{a} - \underline{a})^2 - 2/\beta}, \sqrt{(\overline{a} - \underline{a})^2 - 2/\alpha}]$.

By an analogous reasoning, one can show that $\mathcal{P}_{\underline{a},\overline{a}}^{dec,cv}$, the class of probability measures having a non-increasing and concave density, can be generated by a set $\mathcal{T}_{\underline{a},\overline{a}}^{dec,cv} = \{\tau_t \mid t \in [0, \overline{a} - \underline{a}]\}$, where τ_t is a measure corresponding to the right-trapezoidal density

$$\zeta_t(x) = \frac{2}{\overline{a} - \underline{a} + t} \min\left(1, \frac{\overline{a} - x}{\overline{a} - \underline{a} - t}\right) 1_{[\underline{a},\overline{a}]}(x).$$

A restriction on the (absolute) slope of the density function can be built in by a similar restriction on t.

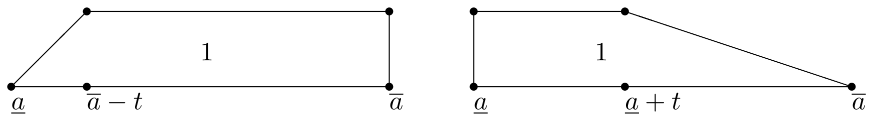

Figure 3: Illustration of the trapezoidal densities in Example 2.4 that generate non-decreasing densities (left) and concave non-increasing densities (right).

3. Combining Shape Constraints Let the subdomains $(\Omega^{(j)})_{j\in\mathcal{J}}$ be a collection of subsets of Ω and let J be the cardinality of the index set $\mathcal{J}$. In practical situations the collection of subdomains reflects the structure of a finite tesselation.

Definition 3.1 *A finite tesselation (or tiling) $(\Omega^{(j)})_{j\in\mathcal{J}}$ on Ω consists of a finite family of subdomains $\Omega^{(j)}$ that cover Ω, have non-empty and pairwise disjoint interiors and are closed and connected.*

A tesselation can be seen as a cover of Ω with the additional restrictions that its elements (subdomains) are closed and that the only overlappings occur at the boundaries of adjacent subdomains. In the univariate case a tesselation is a set of closed intervals, which are only allowed to overlap at their endpoints. In the multivariate case subdomains can adopt complicated shapes, depending on the nature of the shape constraint imposed on it.

Although in practice the tesselation structure is the most logical to work with, we do not require it throughout the paper. Most of our results are valid in a more general context. We only make two hard requirements: The number of subdomains is finite and every subdomain is Borel-measurable. Additional assumptions will be mentioned in the forthcoming propositions when necessary.

Next, we assign a fixed mass to each subdomain. The vector containing the masses is called the mass vector.

Definition 3.2 *A mass vector on a collection of subdomains $(\Omega^{(j)})_{j\in\mathcal{J}}$ is a vector $\eta = (\eta^{(j)})_{j\in\mathcal{J}}$ in $\mathbb{R}^J$ satisfying the conditions*

$$\sum_{j\in\mathcal{J}} \eta^{(j)} = 1, \qquad 0 \le \eta^{(j)} \le 1, \quad \forall j \in \mathcal{J}.$$

Subdomains with zero mass are allowed, but they have no impact and can in fact be removed from the collection. The extension to uncertain mass vectors is handled in the next section.

Suppose that $\mathcal{P}^{(j)}$ is a convex set of probability measures in $\mathcal{P}_{\Omega^{(j)}}$, representing a local shape constraint. Suppose further that $\mathcal{P}^{(j)}$ can be generated by a set of local generating measures $\mathcal{T}^{(j)}$ in the sense of Assumption 2.1. We can compose these local measure sets by taking weighted sums with respect to a mass vector η.

Definition 3.3 *Let $(\Omega^{(j)})_{j\in\mathcal{J}}$ be a collection of subdomains with mass vector $(\eta^{(j)})_{j\in\mathcal{J}}$. Let $\mathcal{T}^{(j)}$ and $\mathcal{P}^{(j)}$ be local measure sets in $\mathcal{P}_{\Omega^{(j)}}$ for each $j \in \mathcal{J}$. Define the measure sets*

$$\mathcal{P}^\eta = \sum_{j\in\mathcal{J}} \eta^{(j)}\mathcal{P}^{(j)}, \quad \mathcal{G}^\eta = \sum_{j\in\mathcal{J}} \eta^{(j)}\mathcal{P}^{(j)}|_{int(\Omega^{(j)})}, \quad \mathcal{T}^\eta = \sum_{j\in\mathcal{J}} \eta^{(j)}\mathcal{T}^{(j)}.$$

In this way we construct a measure set $\mathcal{P}^\eta$, which contains the probability measures satisfying all local shape constraints simultaneously. We will see in the forthcoming propositions that, under certain conditions, this set is generated by $\mathcal{T}^\eta$. The measure set $\mathcal{G}^\eta$ will play a role in some additional results.

Proposition 3.1 *Let $(\Omega^{(j)})_{j\in\mathcal{J}}$ be a collection of subdomains with mass vector $\eta \in \mathbb{R}^J$. Let $\mathcal{P}^{(j)}$ be a measure set in $\mathcal{P}_{\Omega^{(j)}}$ for all $j \in \mathcal{J}$. Then the following holds:*

(i) If $\mathcal{P}^{(j)}$ is convex for all $j \in \mathcal{J}$, then $\mathcal{P}^\eta$ is convex.

(ii) If $\mathcal{P}^{(j)}$ and $\Omega^{(j)}$ are closed for all $j \in \mathcal{J}$, then $\mathcal{P}^\eta$ is closed.

Proof.

(i) This is straightforward. Let $\theta \in [0,1]$ and $\pi_1, \pi_2 \in \mathcal{P}^\eta$, then the convexity of $\mathcal{P}^\eta$ is shown if

$$\pi = \theta\pi_1 + (1-\theta)\pi_2$$

belongs to $\mathcal{P}^\eta$. There exist local measures $\pi_1^{(j)}$ and $\pi_2^{(j)}$ in $\mathcal{P}^{(j)}$ such that

$$\pi_1 = \sum_{j\in\mathcal{J}} \eta^{(j)}\pi_1^{(j)}, \quad \pi_2 = \sum_{j\in\mathcal{J}} \eta^{(j)}\pi_2^{(j)}.$$

This enables us to rewrite π as

$$\pi = \sum_{j\in\mathcal{J}} \eta^{(j)}(\theta\pi_1^{(j)} + (1-\theta)\pi_2^{(j)}),$$

and by the convexity of $\mathcal{P}^{(j)}$ it follows that $\pi \in \mathcal{P}^\eta$.

(ii) This requires some more work. Let $(\pi_p)_p$ be a sequence of measures in $\mathcal{P}^\eta$ weakly converging to π. This sequence is relatively compact[2]. By Prohorov's theorem it is tight[3] if the underlying metric space is separable and complete (see e.g. Capasso [4], Theorem B.71), which is satisfied because $\Omega \subseteq \mathbb{R}^M$ is closed. We can verify with the definition of tightness that the measure sets $(\pi_p^{(j)})_p$, for all $j \in \mathcal{J}$, are tight as well by a transposition argument. Indeed, suppose that there exists $j_0 \in \mathcal{J}$ such that $(\pi_p^{(j_0)})_p$ is not tight, i.e. there exists $\epsilon > 0$ such that for all $G^{(j_0)}$ compact in $\Omega^{(j_0)}$ there exists p_0 with $\pi_{p_0}^{(j_0)}(G^{(j_0)}) \leq 1-\epsilon$. Define $\epsilon_0 = \epsilon\eta^{(j_0)}$ and let $G \subseteq \Omega$ be compact. Then $G^{(j)} = \Omega^{(j)} \cap G$ is compact because $\Omega^{(j)}$ is closed. It holds that

$$\begin{aligned}
\pi_{p_0}(G) &= \eta^{(j_0)}\pi_{p_0}^{(j_0)}(G^{(j_0)}) + \sum_{j\in\mathcal{J}\setminus\{j_0\}} \eta^{(j)}\pi_{p_0}^{(j)}(G^{(j)}) \\
&\leq \eta^{(j_0)}(1-\epsilon) + \sum_{j\in\mathcal{J}\setminus\{j_0\}} \eta^{(j)} \\
&= 1-\epsilon_0,
\end{aligned}$$

which shows that $(\pi_p)_p$ is not tight.

By Prohorov's theorem the tight measure sets $(\pi_p^{(j)})_p$ are relatively compact for all $j \in \mathcal{J}$. This allows us to construct a subsequence $(\pi_{p_1(q)}^{(1)})_q$ weakly converging to a measure $\pi^{(1)}$, which belongs to $\mathcal{P}^{(1)}$ because this set is closed. The subsequence $(\pi_{p_1(q)}^{(2)})_q$ is relatively compact and therefore has a subsequence $(\pi_{p_2(q)}^{(2)})_q$ weakly converging to $\pi^{(2)} \in \mathcal{P}^{(2)}$. By repeating this procedure, we construct a subsequence $(\pi_{p_J(q)}^{(J)})_q$ of $(\pi_p^{(J)})_p$ weakly converging to $\pi^{(J)} \in \mathcal{P}^{(J)}$. But this implies that for all $j \in \mathcal{J}$ the subsequences $(\pi_{p_J(q)}^{(j)})_q$ weakly converge to $\pi^{(j)}$. By one of the equivalences in Portmanteau's theorem (see e.g. Capasso [4], Theorem B.63), it can be checked that

$$\left(\pi_{p_J(q)}\right)_q = \left(\sum_{j\in\mathcal{J}} \eta^{(j)}\pi_{p_J(q)}^{(j)}\right)_q \xrightarrow{w} \sum_{j\in\mathcal{J}} \eta^{(j)}\pi^{(j)} \in \mathcal{P}^\eta.$$

This shows that a subsequence of the original sequence $(\pi_p)_p$ converges to an element of $\mathcal{P}^\eta$, hence $\pi = \sum_{j\in\mathcal{J}} \eta^{(j)}\pi^{(j)}$, which concludes the proof. □

The next proposition states that $\mathcal{T}^\eta$ is a generating set for $\mathcal{P}^\eta$, if the convex hull representation is of the same type on each subdomain.

[2] A measure set $\mathcal{P}$ is said to be relatively compact if every sequence in $\mathcal{P}$ has a (weakly) convergent subsequence.

[3] A measure set $\mathcal{P}$ is said to be tight if for all $\epsilon > 0$ there exists a compact set G such that for all $\pi \in \mathcal{P}$ it holds that $\pi(G) > 1-\epsilon$.

Proposition 3.2 *Let $(\Omega^{(j)})_{j\in\mathcal{J}}$ be a collection of subdomains with mass vector $\eta \in \mathbb{R}^J$. Let $\mathcal{P}^{(j)}$ be a measure set in $\mathcal{P}_{\Omega^{(j)}}$ for all $j \in \mathcal{J}$. Then the following holds:*

(i) If $\mathcal{P}^{(j)} = \mathcal{H}(\mathcal{T}^{(j)})$ for all $j \in \mathcal{J}$, then $\mathcal{P}^\eta = \mathcal{H}(\mathcal{T}^\eta)$.

(ii) If $\mathcal{P}^{(j)} = cl(\mathcal{H}(\mathcal{T}^{(j)}))$ and $\Omega^{(j)}$ is closed for all $j \in \mathcal{J}$, then $\mathcal{P}^\eta = cl(\mathcal{H}(\mathcal{T}^\eta))$.

Proof.

Let us first prove the identity

$$\sum_{j\in\mathcal{J}} \eta^{(j)}\mathcal{H}(\mathcal{T}^{(j)}) = \mathcal{H}(\mathcal{T}^\eta).$$

Its proof is similar to that of Lemaréchal [11], Proposition A.3. Let τ be a measure in $\sum_{j\in\mathcal{J}} \eta^{(j)}\mathcal{H}(\mathcal{T}^{(j)})$, then for each $j \in \mathcal{J}$ there exist an integer T_j and $\lambda^{(j)}_{t_j} \in [0,1]$, for $t_j = 1,\ldots,T_j$, with $\lambda^{(j)}_1 + \ldots + \lambda^{(j)}_{T_j} = 1$ and $\tau^{(j)}_{t_j} \in \mathcal{T}^{(j)}$ such that

$$\tau = \sum_{j\in\mathcal{J}} \eta^{(j)} \sum_{t_j=1}^{T_j} \lambda^{(j)}_{t_j}\tau^{(j)}_{t_j}.$$

For a fixed $j_0 \in \mathcal{J}$ and $t_{j_0} \in \{1,\ldots,T_{j_0}\}$, we can express $\tau^{(j_0)}_{t_{j_0}}$ as the convex combination

$$\tau^{(j_0)}_{t_{j_0}} = \sum_{t_1=1}^{T_1} \lambda^{(1)}_{t_1} \cdots \widehat{\sum_{t_{j_0}=1}^{T_{j_0}} \lambda^{(j_0)}_{t_{j_0}}} \cdots \sum_{t_J=1}^{T_J} \lambda^{(J)}_{t_J}\tau^{(j_0)}_{t_{j_0}},$$

where the hat indicates missing terms. We substitute this in the expression for τ and rearrange terms as follows

$$\begin{aligned}\tau &= \sum_{j\in\mathcal{J}} \eta^{(j)} \sum_{t_1=1}^{T_1} \lambda^{(1)}_{t_1} \cdots \sum_{t_J=1}^{T_J} \lambda^{(J)}_{t_J}\tau^{(j)}_{t_j}\\ &= \sum_{t_1=1}^{T_1} \lambda^{(1)}_{t_1} \cdots \sum_{t_J=1}^{T_J} \lambda^{(J)}_{t_J} \sum_{j\in\mathcal{J}} \eta^{(j)}\tau^{(j)}_{t_j}.\end{aligned}$$

In this way we can write τ as a convex combination of measures in $\mathcal{T}^\eta$, which proves one inclusion of the identity (the opposite inclusion is straightforward).

The proof of (i) is given by the inclusion sequence

$$\mathcal{P}^\eta = \sum_{j\in\mathcal{J}} \eta^{(j)}\mathcal{H}(\mathcal{T}^{(j)}) = \mathcal{H}(\mathcal{T}^\eta) \subseteq \mathcal{H}(\mathcal{P}^\eta) = \mathcal{P}^\eta,$$

whereas (ii) follows by

$$\begin{aligned}\mathcal{P}^\eta &= \sum_{j\in\mathcal{J}} \eta^{(j)} cl(\mathcal{H}(\mathcal{T}^{(j)})) \subseteq cl\Big(\sum_{j\in\mathcal{J}} \eta^{(j)}\mathcal{H}(\mathcal{T}^{(j)})\Big)\\ &= cl(\mathcal{H}(\mathcal{T}^\eta)) \subseteq cl(\mathcal{H}(\mathcal{P}^\eta)) = \mathcal{P}^\eta.\end{aligned}$$

□

A direct consequence of this proposition is that in the case $\mathcal{P} = \mathcal{P}^\eta$ we can reformulate program (P5) in terms of the local generators,

$$\text{(P7)} \qquad \begin{aligned}\min_{y\in -K^*} \quad & \mu \cdot y\\ s.t. \quad & \sum_{j\in\mathcal{J}} \eta^{(j)} \int_{\Omega^{(j)}} (y'f - \phi)\, \mathrm{d}\tau^{(j)} \geq 0 \qquad \forall \tau^{(j)} \in \mathcal{T}^{(j)},\end{aligned}$$

provided that the following assumption is satisfied.

Assumption 3.1 *Given a collection of subdomains* $(\Omega^{(j)})_{j\in\mathcal{J}}$ *and measure sets* $\mathcal{P}^{(j)} \subseteq \mathcal{P}_{\Omega^{(j)}}$. *One of the following statements holds:*

[B1] For all $j \in \mathcal{J}$*, the set* $\mathcal{P}^{(j)}$ *admits a representation* $\mathcal{P}^{(j)} = \mathcal{H}(\mathcal{T}^{(j)})$.

[B2] For all $j \in \mathcal{J}$*, the set* $\mathcal{P}^{(j)}$ *admits a representation* $\mathcal{P}^{(j)} = cl(\mathcal{H}(\mathcal{T}^{(j)}))$*, with* $\Omega^{(j)}$ *and* $\mathcal{T}^{(j)}$ *closed.*

[B3] For all $j \in \mathcal{J}$*, the set* $\mathcal{P}^{(j)}$ *admits a representation* $\mathcal{P}^{(j)} = cl(\mathcal{H}(\mathcal{T}^{(j)}))$*, with* $\Omega^{(j)}$ *closed,* f *and* ϕ *continuous and bounded.*

As we already mentioned, we have not required the subdomains to constitute a tesselation structure. The derived results are valid under very loose assumptions concerning the subdomains; only closedness is necessary in the case of the closed convex hull representation. In reality we aim to spread the mass over the different subdomains. Clearly, this can be achieved when there are no overlappings between the different subdomains. However, in the case of closed subdomains, the possibility remains that some mass is located on the boundary of two adjacent subdomains. We will show that, if the local generating measure sets are tight on the interior of their subdomain, the mass of the generated measures is concentrated on the interior of the subdomains and hence a proper mass dispersion is attained.

Lemma 3.1 *Let* O *be an open subset of* Ω*, and let* $\mathcal{T} = \mathcal{T}|_O$ *be a family of probability measures. Then the following assertions are equivalent:*

(i) The family $\mathcal{T}$ *is tight in* $\mathcal{P}_O$.

(ii) For every sequence of bounded continuous functions $(g_q)_{q\in\mathbb{N}}$ *which decreases pointwise to* $1_{\Omega\setminus O}$ *the following equality holds:*

$$\inf_{q\in\mathbb{N}} \sup_{\tau\in\mathcal{T}} \int_O g_q \, \mathrm{d}\tau = 0.$$

(iii) For every sequence of bounded continuous functions $(g_q)_{q\in\mathbb{N}}$ *which decreases pointwise to* $1_{\Omega\setminus O}$ *the following equality holds:*

$$\inf_{q\in\mathbb{N}} \sup_{\pi\in\mathcal{P}} \int_O g_q \, \mathrm{d}\pi = 0,$$

where $\mathcal{P} = cl(\mathcal{H}(\mathcal{T}))$.

The proof of the lemma is deferred to the appendix.

Proposition 3.3 *Let* $(\Omega^{(j)})_{j\in\mathcal{J}}$ *be a collection of subdomains with mass vector* $\eta \in \mathbb{R}^J$*. Let* $\mathcal{P}^{(j)}$ *be a measure set in* $\mathcal{P}_{\Omega^{(j)}}$ *for all* $j \in \mathcal{J}$*. Then the following holds:*

(i) If $\mathcal{T}^{(j)} = \mathcal{T}^{(j)}|_{int(\Omega^{(j)})}$ *and* $\mathcal{P}^{(j)} = \mathcal{H}(\mathcal{T}^{(j)})$ *for all* $j \in \mathcal{J}$*, then* $\mathcal{P}^\eta = \mathcal{G}^\eta$.

(ii) If $\mathcal{T}^{(j)} = \mathcal{T}^{(j)}|_{int(\Omega^{(j)})}$ *is tight in* $\mathcal{P}_{int(\Omega^{(j)})}$ *and* $\mathcal{P}^{(j)} = cl(\mathcal{H}(\mathcal{T}^{(j)}))$ *for all* $j \in \mathcal{J}$*, then* $\mathcal{P}^\eta = \mathcal{G}^\eta$.

Proof. (i) For all $j \in \mathcal{J}$ we have that

$$\mathcal{P}^{(j)} = \mathcal{H}(\mathcal{T}^{(j)}|_{int(\Omega^{(j)})}) \subseteq \mathcal{H}(\mathcal{P}_{int(\Omega^{(j)})}) = \mathcal{P}_{int(\Omega^{(j)})},$$

which implies $\mathcal{P}^{(j)} = \mathcal{P}^{(j)}|_{int(\Omega^{(j)})}$ and proves the desired result.

(ii) Fix $j \in \mathcal{J}$ and let $\pi^{(j)} \in \mathcal{P}^{(j)}$. Let $(g_q)_{q\in\mathbb{N}}$ be a sequence of bounded and continuous functions which decrease pointwise to $1_{\Omega\setminus int(\Omega^{(j)})}$. Approximate $\pi^{(j)}$ with a weakly convergent sequence of measures

$(\pi_p)_{p\in\mathbb{N}}$ in $\mathcal{H}(\mathcal{T}^{(j)}) \subset \mathcal{P}_{int(\Omega^{(j)})}$. We now apply Lemma 3.1 and Portmanteau's theorem:

$$\begin{aligned} 0 &= \inf_{q\in\mathbb{N}} \sup_{\pi\in\mathcal{P}^{(j)}} \int_{int(\Omega^{(j)})} g_q \, \mathrm{d}\pi \\ &\geq \inf_{q\in\mathbb{N}} \sup_{p\in\mathbb{N}} \int_{int(\Omega^{(j)})} g_q \, \mathrm{d}\pi_{\mathrm{p}} \\ &\geq \inf_{q\in\mathbb{N}} \int_{\Omega^{(j)}} g_q \, \mathrm{d}\pi^{(\mathrm{j})} \\ &= \pi^{(j)}(\Omega^{(j)} \backslash int(\Omega^{(j)})) \\ &= 1 - \pi^{(j)}(int(\Omega^{(j)})). \end{aligned}$$

Hence $\pi^{(j)} \in \mathcal{P}_{int(\Omega^{(j)})}$ which concludes the proof. □

PROPOSITION 3.4 *If $\mathcal{P}^\eta = \mathcal{G}^\eta$ and the collection $(\Omega^{(j)})_{j\in\mathcal{J}}$ has pairwise disjoint interiors, then $\pi(cl(\Omega^{(j)})) = \pi(int(\Omega^{(j)})) = \eta^{(j)}$ for all $j \in \mathcal{J}$ and $\pi \in \mathcal{P}^\eta$.*

PROOF.

We have for every $\pi^{(j)} \in \mathcal{P}^{(j)}|_{int(\Omega^{(j)})}$ that

$$\begin{aligned} \pi^{(j)}(int(\Omega^{(j)})) &= \pi^{(j)}(cl(\Omega^{(j)})) = 1 \qquad \forall j \in \mathcal{J}, \\ \pi^{(j)}(int(\Omega^{(j_0)})) &= \pi^{(j)}(cl(\Omega^{(j_0)})) = 0 \qquad \forall j \in \mathcal{J} \backslash \{j_0\}. \end{aligned}$$

□

If the conditions of the above proposition are satisfied, it is guaranteed that the probability mass is concentrated on the interior of all subdomains. This leads to easy bounds in the particular case that $\phi = 1_{\Omega^{(j)}}$ for a fixed $j \in \mathcal{J}$. The optimal value of (P2) equals the mass on the subdomain,

$$\langle 1_{\Omega^{(j)}}, \pi \rangle = \int_{\Omega^{(j)}} \mathrm{d}\pi = \pi(\Omega^{(j)}) = \eta^{(j)},$$

provided that the program is feasible.

EXAMPLE 3.1 *A first possibility is to set $\mathcal{T}^{(j)} = \Delta_{\Omega^{(j)}}$ for all $j \in \mathcal{J}$. If $\Omega^{(j)}$ is closed, then $\Delta_{\Omega^{(j)}}$ is closed and program (P7) becomes*

$$\text{(P8)} \qquad \begin{aligned} &\min_{y\in -K^*} \quad \mu \cdot y \\ &\quad s.t. \quad \sum_{j\in\mathcal{J}} \eta^{(j)} \left(y' f(t^{(j)}) - \phi(t^{(j)}) \right) \geq 0, \qquad \forall t^{(j)} \in \Omega^{(j)}. \end{aligned}$$

The generating set $\Delta_{\Omega^{(j)}}$ assigns a strictly positive mass to the boundary of all subdomains, implying that the probability mass on a subdomain may exceed the value given by the mass vector. Hence Proposition 3.4 is not applicable here.

EXAMPLE 3.2 *Let $\Omega = [\underline{a}, \overline{a}]$, where Ω may be unbounded in both directions. Let m be in the interior of Ω and define $\Omega^{(1)} = [\underline{a}, m]$ and $\Omega^{(2)} = [m, \overline{a}]$. Let $\eta = (\eta^{(1)}, \eta^{(2)})$ be a feasible mass vector. Consider the generating set $\mathcal{T} = \eta^{(1)}\mathcal{T}^{(1)} + \eta^{(2)}\mathcal{T}^{(2)}$, where*

$$\begin{aligned} \mathcal{T}^{(1)} &= \{\tau^{(1)}_{t_1} \mid t_1 \in [\underline{a}, m)\} \cup \{\delta_m\} \\ \mathcal{T}^{(2)} &= \{\tau^{(2)}_{t_2} \mid t_2 \in (m, \overline{a}]\} \cup \{\delta_m\} \end{aligned}$$

and the probability measures $\tau^{(1)}_{t_1}$ and $\tau^{(2)}_{t_2}$ correspond to the density functions

$$\zeta^{(1)}_{t_1}(x) = \frac{1}{m - t_1} 1_{[t_1, m]}(x), \quad \zeta^{(2)}_{t_2}(x) = \frac{1}{t_2 - m} 1_{[m, t_2]}(x)$$

respectively. By Proposition 3.2 it follows that the generated probability measures are m-unimodal, i.e. they have a non-decreasing density at the left of m, a non-increasing density at the right of m and possibly include a Dirac component at m. Furthermore, they obey the mass constraint, which means that the mass on $\Omega^{(1)}$ and $\Omega^{(2)}$ equals $\eta^{(1)}$ and $\eta^{(2)}$ respectively, where possibly some mass is shared at m.

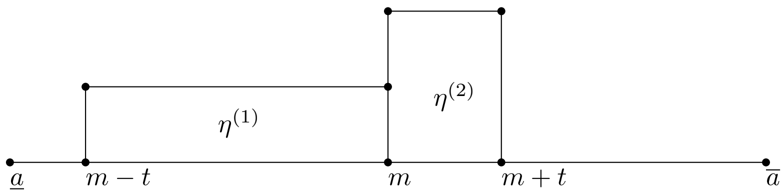

Figure 4: Illustration of the densities in Example 3.2 that generate the mass-constrained m-unimodal measures.

Example 3.3 *We can extend the class of mass-constrained m-unimodal measures to the class of the mass-constrained $[m_1, m_2]$-unimodal measures as follows. Let $\Omega = [\underline{a}, \overline{a}]$, where Ω may be unbounded in both directions, and let $m_1 < m_2$ be two reals in the interior of Ω. Let $\eta = (\eta^{(1)}, \eta^{(2)}, \eta^{(3)})$ be a feasible mass vector. Let*

$$\mathcal{T}^{(1)} = \mathcal{T}^{inc}_{\underline{a},m_1}, \quad \mathcal{T}^{(2)} = \Delta_{[m_1,m_2]}, \quad \mathcal{T}^{(3)} = \mathcal{T}^{dec}_{m_2,\overline{a}}.$$

By Proposition 3.2 it follows that $\mathcal{T} = \eta^{(1)}\mathcal{T}^{(1)} + \eta^{(2)}\mathcal{T}^{(2)} + \eta^{(3)}\mathcal{T}^{(3)}$ generates a class of probability measures which are mass-constrained $[m_1, m_2]$-unimodal, i.e. they have a non-decreasing density at the left of m_1, a non-increasing density at the right of m_2 and are arbitrary in the interval $[m_1, m_2]$. Remark that the generated measures are not required to be unimodal on $[m_1, m_2]$.

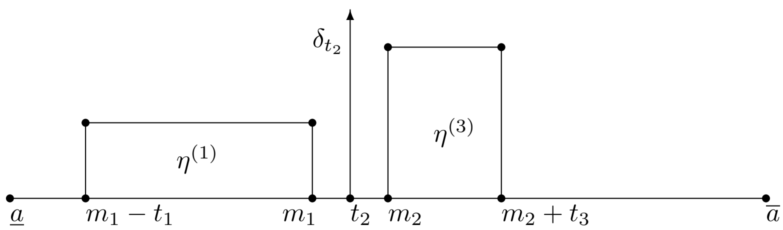

Figure 5: Illustration of the densities in Example 3.3 that generate the mass-constrained $[m_1, m_2]$-unimodal measures.

Example 3.4 *We can also combine the convex and concave generators to optimize over more complex shapes. For example in finance one may be interested in distributions which are convex and non-decreasing on $[0, b_1]$, concave and non-decreasing on $[b_1, m]$, concave and non-increasing on $[m, b_2]$ and convex and non-increasing on $[b_2, +\infty)$, assuming that the mode m and the inflection points b_1 and b_2 are fixed. Note that it is possible to introduce mode uncertainty by adding an interval with Dirac measures in the neighbourhood of m and inflection point uncertainty by adding an interval with rectangular measures in the neighbourhood of an inflection point.*

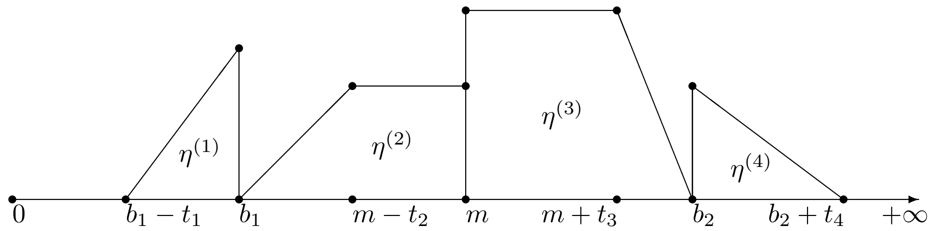

Figure 6: The measures in Example 3.4 that generate second-order constrained measures.

4. Mass Uncertainty We will generalize the propositions of the previous section by allowing the mass vector η to range in an uncertainty set U.

Definition 4.1 *A non-empty set $U \subset \mathbb{R}^J$ is called a mass uncertainty set if it is the intersection of a closed, convex subset in the J-dimensional unit hypercube $\{(\eta^{(j)})_{j\in\mathcal{J}} \in \mathbb{R}^J \,|\, 0 \le \eta^{(j)} \le 1,\, j \in \mathcal{J}\}$ with the hyperplane $H \leftrightarrow \sum_{j\in\mathcal{J}} \eta^{(j)} = 1$.*

Each mass uncertainty set is a convex and compact set, implying it is the convex hull of its extreme points.

An interesting setting is that of box uncertainty, where a lower and upper mass is imposed on each subdomain
$$0 \leq \underline{\eta^{(j)}} \leq \eta^{(j)} \leq \overline{\eta^{(j)}} \leq 1 \qquad \forall j \in \mathcal{J}.$$
In this case U is the intersection of H with a J-dimensional box, parallel to the coordinate axes and fully determined by a couple of points, the lower and upper mass vector pair $(\underline{\eta}, \overline{\eta})$. In order to be sure that the box intersects with H, it suffices that
$$\sum_{j \in \mathcal{J}} \underline{\eta^{(j)}} \leq 1 \leq \sum_{j \in \mathcal{J}} \overline{\eta^{(j)}}.$$
Let S denote the number of extreme points in U, thus $S = |ex(U)|$. Naturally, S is not fixed and depends on the location of the mass vector pair $(\underline{\eta}, \overline{\eta})$. For an upper bound on S we refer to the work of O'Neil [16] and Reza [18], who introduced the hypercube slicing number as the maximal number of edges that a single hyperplane can cut in a hypercube. Following their reasoning, the optimal upper bound for S equals $\lceil \frac{J}{2} \rceil \binom{J}{\lceil \frac{J}{2} \rceil}$.

Consider a general mass uncertainty set U and a set of local measures $\mathcal{T}^{(j)} \subseteq \mathcal{P}_{\Omega^{(j)}}$, generating $\mathcal{P}^{(j)}$ in the sense of Assumption 2.1. Again, we combine the local measures by taking convex combinations over the different subdomains, but now taking into account the uncertainty of the mass vector.

Definition 4.2 *Let $\mathcal{T}^{(j)}$ and $\mathcal{P}^{(j)}$ be local measure sets defined on a collection of subdomains $(\Omega^{(j)})_{j \in \mathcal{J}}$. Let U be a mass uncertainty set. Then define the following measure sets*

$$\begin{aligned}
\mathcal{P}^U &= \bigcup_{\eta \in U} \mathcal{P}^\eta = \Big\{ \sum_{j \in \mathcal{J}} \eta^{(j)} \pi^{(j)} \mid \eta \in U,\ \pi^{(j)} \in \mathcal{P}^{(j)},\ j \in \mathcal{J} \Big\}, \\
\mathcal{G}^U &= \bigcup_{\eta \in U} \mathcal{G}^\eta = \Big\{ \sum_{j \in \mathcal{J}} \eta^{(j)} \pi^{(j)} \mid \eta \in U,\ \pi^{(j)} \in \mathcal{P}^{(j)}|_{int(\Omega^{(j)})},\ j \in \mathcal{J} \Big\}, \\
\mathcal{T}^U &= \bigcup_{\eta \in U} \mathcal{T}^\eta = \Big\{ \sum_{j \in \mathcal{J}} \eta^{(j)} \tau^{(j)} \mid \eta \in U,\ \tau^{(j)} \in \mathcal{T}^{(j)},\ j \in \mathcal{J} \Big\}, \\
\mathring{\mathcal{T}}^U &= \bigcup_{\eta \in ex(U)} \mathcal{T}^\eta = \Big\{ \sum_{j \in \mathcal{J}} \eta^{(j)} \tau^{(j)} \mid \eta \in ex(U),\ \tau^{(j)} \in \mathcal{T}^{(j)},\ j \in \mathcal{J} \Big\}.
\end{aligned}$$

We are now in position to formulate a generalization of Proposition 3.1 and 3.2.

Proposition 4.1 *Let $(\Omega^{(j)})_{j \in \mathcal{J}}$ be a collection of subdomains with mass uncertainty set $U \subset \mathbb{R}^J$. Let $\mathcal{P}^{(j)}$ be a measure set in $\mathcal{P}_{\Omega^{(j)}}$ for all $j \in \mathcal{J}$. Then the following holds:*

(i) If $\mathcal{P}^{(j)}$ is convex for all $j \in \mathcal{J}$, then $\mathcal{P}^U$ is convex.

(ii) If $\mathcal{P}^{(j)}$ and $\Omega^{(j)}$ are closed for all $j \in \mathcal{J}$, then $\mathcal{P}^U$ is closed.

Proof. (i) Let the probability measure π be the convex combination $\pi = \theta \pi_1 + (1-\theta)\pi_2$, with $\pi_1, \pi_2 \in \mathcal{P}^U$ and $\theta \in [0,1]$. Then there exist two mass vectors $(\eta_1^{(1)}, \ldots, \eta_1^{(J)}), (\eta_2^{(1)}, \ldots, \eta_2^{(J)}) \in U$ and local probability measures $\pi_1^{(j)}, \pi_2^{(j)} \in \mathcal{P}^{(j)}$ for all $j \in \mathcal{J}$, such that π can be written as
$$\pi = \theta \sum_{j \in \mathcal{J}} \eta_1^{(j)} \pi_1^{(j)} + (1-\theta) \sum_{j \in \mathcal{J}} \eta_2^{(j)} \pi_2^{(j)}.$$

Defining the weights

$$\begin{aligned}
\eta^{(j)} &= \theta \eta_1^{(j)} + (1-\theta)\eta_2^{(j)} \\
\lambda_1^{(j)} &= \theta \eta_1^{(j)} / \eta^{(j)} \\
\lambda_2^{(j)} &= (1-\theta)\eta_2^{(j)} / \eta^{(j)}
\end{aligned}$$

for all $j \in \mathcal{J}$, we can rearrange the terms in π as

$$\pi = \sum_{j\in\mathcal{J}} \eta^{(j)} \left(\lambda_1^{(j)}\pi_1^{(j)} + \lambda_2^{(j)}\pi_2^{(j)}\right),$$

which obviously shows that $\pi \in \mathcal{P}^U$.

(ii) Let $(\pi_p)_p$ be a sequence of measures in $\mathcal{P}^U$ weakly converging to π. There exists a sequence of mass vectors $(\eta_p^{(1)}, \dots, \eta_p^{(J)})_p$ in U and sequences of probability measures $(\pi_p^{(j)})_p$ in $\mathcal{P}^{(j)}$ for each $j \in \mathcal{J}$ such that $\pi_p = \sum_{j\in\mathcal{J}} \eta_p^{(j)}\pi_p^{(j)}$ for all p. Because U is compact, there exists a subsequence of mass vectors $(\eta_{p(q)}^{(1)}, \dots, \eta_{p(q)}^{(J)})_q$ converging to a mass vector $\eta = (\eta^{(1)}, \dots, \eta^{(J)}) \in U$. Consider now the subsequence $(\pi_{p(q)})_q$, which is convergent and therefore relatively compact. By the same reasoning as in the proof of Proposition 3.2, we construct subsequences $(\pi_{p_J(q)}^{(j)})_q$ weakly converging to $\pi^{(j)} \in \mathcal{P}^{(j)}$ for all $j \in \mathcal{J}$. Let $g : \Omega \longrightarrow \mathbb{R}$ be a continuous and bounded function. By Portmanteau's theorem it follows that

$$\left(\int_{\Omega^{(j)}} g \,\mathrm{d}\pi_{p_J(q)}^{(j)}\right)_q \to \int_{\Omega^{(j)}} g \,\mathrm{d}\pi^{(j)}, \quad \forall j \in \mathcal{J}.$$

If we combine these converging sequences with the converging mass vector subsequence $(\eta_{p_J(q)}^{(1)}, \dots, \eta_{p_J(q)}^{(J)})_q$, we obtain the converging sequence

$$\left(\sum_{j\in\mathcal{J}} \eta_{p_J(q)}^{(j)} \int_{\Omega^{(j)}} g \,\mathrm{d}\pi_{p_J(q)}^{(j)}\right)_q \to \sum_{j\in\mathcal{J}} \eta^{(j)} \int_{\Omega^{(j)}} g \,\mathrm{d}\pi^{(j)},$$

or equivalently

$$\left(\pi_{p_J(q)}\right)_q \xrightarrow{w} \sum_{j\in\mathcal{J}} \eta^{(j)}\pi^{(j)}.$$

This shows that a subsequence of the original sequence $(\pi_p)_p$ converges to $\pi = \sum_{j\in\mathcal{J}} \eta^{(j)}\pi^{(j)}$, which obviously belongs to $\mathcal{P}^U$. This concludes the proof. □

Proposition 4.2 *Let $(\Omega^{(j)})_{j\in\mathcal{J}}$ be a collection of subdomains with mass uncertainty set $U \subset \mathbb{R}^J$. Let $\mathcal{P}^{(j)}$ be a measure set in $\mathcal{P}_{\Omega^{(j)}}$ for all $j \in \mathcal{J}$. Then the following holds:*

(i) If $\mathcal{P}^{(j)} = \mathcal{H}(\mathcal{T}^{(j)})$ for all $j \in \mathcal{J}$, then $\mathcal{P}^U = \mathcal{H}(\mathring{\mathcal{T}}^U)$.

(ii) If $\mathcal{P}^{(j)} = cl(\mathcal{H}(\mathcal{T}^{(j)}))$ and $\Omega^{(j)}$ is closed for all $j \in \mathcal{J}$, then $\mathcal{P}^U = cl(\mathcal{H}(\mathring{\mathcal{T}}^U))$.

Proof. Let us first show that the identity

$$\mathcal{H}(\mathcal{T}^U) = \mathcal{H}(\mathring{\mathcal{T}}^U)$$

holds. It clearly suffices to prove the inclusion $\subseteq$. Any $\tau \in \mathcal{H}(\bigcup_{\eta\in U} \mathcal{T}^\eta)$ allows the representation

$$\tau = \theta \sum_{j\in\mathcal{J}} \eta_1^{(j)}\tau_1^{(j)} + (1-\theta) \sum_{j\in\mathcal{J}} \eta_2^{(j)}\tau_2^{(j)},$$

with $\theta \in [0,1]$ and $\tau_1^{(j)}, \tau_2^{(j)} \in \mathcal{T}^{(j)}$. The mass vectors η_1 and η_2 belong to U and are each a convex combination of at most $J+1$ extremal points $(\eta_{1r}^{(j)})_{r=1}^{J+1}$ and $(\eta_{2r}^{(j)})_{r=1}^{J+1}$ respectively (Carathéodory's theorem, see e.g. in Rockafellar [19]). It follows that there exist two sets of non-negative coefficients $(\lambda_{1r})_{r=1}^{J+1}$ and $(\lambda_{2r})_{r=1}^{J+1}$, each summing up to 1, such that

$$\tau = \sum_{r=1}^{J+1} \theta\lambda_{1r} \sum_{j\in\mathcal{J}} \eta_{1r}^{(j)}\tau_1^{(j)} + \sum_{r=1}^{J+1} (1-\theta)\lambda_{2r} \sum_{j\in\mathcal{J}} \eta_{2r}^{(j)}\tau_2^{(j)},$$

which shows that τ is a convex combination of elements in $\mathcal{T}^U$ with extremal mass vectors. This proves the above identity.

The proof of (i) is then given by the following inclusion sequence:

$$\mathcal{P}^U = \bigcup_{\eta \in U} \mathcal{H}(\mathcal{T}^\eta) \subseteq \mathcal{H}(\mathcal{T}^U) = \mathcal{H}(\mathring{\mathcal{T}}^U) \subseteq \mathcal{H}(\mathcal{P}^U) = \mathcal{P}^U.$$

The proof of (ii) is similar and is given by

$$\mathcal{P}^U = \bigcup_{\eta \in U} cl(\mathcal{H}(\mathcal{T}^\eta)) \subseteq cl(\mathcal{H}(\mathcal{T}^U)) = cl(\mathcal{H}(\mathring{\mathcal{T}}^U)) \subseteq cl(\mathcal{H}(\mathcal{P}^U)) = \mathcal{P}^U.$$

□

Again, we can use this proposition to reformulate (P5) in the case $\mathcal{P} = \mathcal{P}^U$ in terms of the local generating measures to

(P9)
$$\begin{aligned} \min_{y \in -K^*} \quad & \mu \cdot y \\ s.t. \quad & \sum_{j \in \mathcal{J}} \eta^{(j)} \int_{\Omega^{(j)}} (y'f - \phi)\, \mathrm{d}\tau^{(j)} \geq 0 \qquad \forall \eta \in ex(U),\ \forall \tau^{(j)} \in \mathcal{T}^{(j)}, \end{aligned}$$

provided that Assumption 3.1 holds. Remark that if U is polyhedral, then S is finite and (P9) becomes a variant of (P7) with multiple integral constraints. Next, we extend Proposition 3.3 and 3.4 to mass uncertainty as well.

Proposition 4.3 *Let $(\Omega^{(j)})_{j \in \mathcal{J}}$ be a collection of subdomains with mass uncertainty set $U \subset \mathbb{R}^J$. Let $\mathcal{P}^{(j)}$ be a measure set in $\mathcal{P}_{\Omega^{(j)}}$ for all $j \in \mathcal{J}$. Then the following holds:*

(i) If $\mathcal{T}^{(j)} = \mathcal{T}^{(j)}|_{int(\Omega^{(j)})}$ and $\mathcal{P}^{(j)} = \mathcal{H}(\mathcal{T}^{(j)})$ for all $j \in \mathcal{J}$, then $\mathcal{P}^U = \mathcal{G}^U$.

(ii) If $\mathcal{T}^{(j)} = \mathcal{T}^{(j)}|_{int(\Omega^{(j)})}$ is tight in $\mathcal{P}_{int(\Omega^{(j)})}$ and $\mathcal{P}^{(j)} = cl(\mathcal{H}(\mathcal{T}^{(j)}))$ for all $j \in \mathcal{J}$, then $\mathcal{P}^U = \mathcal{G}^U$.

Proof. This follows directly from Proposition 3.3. □

Proposition 4.4 *If $\mathcal{P}^U = \mathcal{G}^U$ and the collection $(\Omega^{(j)})_{j \in \mathcal{J}}$ has pairwise disjoint interiors, then*

$$\min_{\eta \in ex(U)} \eta^{(j)} \leq \pi(int(\Omega^{(j)})) = \pi(cl(\Omega^{(j)})) \leq \max_{\eta \in ex(U)} \eta^{(j)}$$

for all $j \in \mathcal{J}$ and $\pi \in \mathcal{P}^U$.

Proof. This is basically the same reasoning as in the proof of Proposition 3.4. □

Again, this leads to easy bounds in the particular case that $\phi = 1_{\Omega^{(j)}}$ for a fixed $j \in \mathcal{J}$. The optimal value of (P2) equals the maximal value of the mass vector component of the corresponding subdomain

$$\max_{\pi \in \mathcal{P}} \langle 1_{\Omega^{(j)}}, \pi \rangle = \max_{\pi \in \mathcal{G}^U} \int_{\Omega^{(j)}} \mathrm{d}\pi = \max_{\pi \in \mathcal{G}^U} \pi(\Omega^{(j)}) = \max_{\eta \in ex(U)} \eta^{(j)}$$

provided that the program is feasible.

Example 4.1 *If no mass information is available, it is still possible to impose local shape constraints as follows. Define U as the intersection of H with the J-dimensional unit hypercube. The extremal points of U are precisely the canonical unit vectors of $\mathbb{R}^J$. In this case, problem (P9) becomes*

(P10)
$$\begin{aligned} \min_{y \in -K^*} \quad & \mu \cdot y \\ s.t. \quad & \int_{\Omega^{(j)}} (y'f - \phi)\, \mathrm{d}\tau^{(j)} \geq 0 \qquad \forall \tau^{(j)} \in \mathcal{T}^{(j)},\ \forall j \in \mathcal{J}. \end{aligned}$$

Example 4.2 *Let us apply the setting of the previous example to that of m-unimodality (Example 3.2). It follows that*

$$\mathcal{T} = \mathcal{T}^{inc}_{\underline{a},m} \cup \mathcal{T}^{dec}_{m,\overline{a}} \cup \{\delta_m\}.$$

Moreover, it is possible to describe $\mathcal{T}$ with a single parameter as $\{\tau_t \mid t \in [\underline{a}, \overline{a}]\}$, where τ_t is a rectangular measure on $[t, m]$ if $t \in [\underline{a}, m)$, $\tau_t = \delta_m$ if $t = m$ and τ_t is a rectangular measure on $[m, t]$ if $t \in (m, \overline{a}]$.

Example 4.3 *In an analogous way, we can construct a single-parameter generating set for the* $[m_1, m_2]$*-unimodal measures of Example 3.3. Here* $\mathcal{T} = \{\tau_t \mid t \in [\underline{a}, \overline{a}]\}$*, where* τ_t *is the rectangular measure on* $[t, m_1]$ *if* $t \in [\underline{a}, m_1)$*,* $\tau_t = \delta_t$ *if* $t \in [m_1, m_2]$ *and* τ_t *is the rectangular measure on* $[m_2, t]$ *if* $t \in (m_2, \overline{a}]$*.*

Example 4.4 *Another particular situation occurs if only mass constraints are imposed, in other words there is no shape information available. Then* $\mathcal{T}^{(j)} = \Delta_{\Omega^{(j)}}$ *for all* $j \in \mathcal{J}$*, and program (P9) becomes*

$$\text{(P}\mathit{11}\text{)} \qquad \begin{aligned} \min_{y \in -K^*} \quad & \mu \cdot y \\ s.t. \quad & \sum_{j \in \mathcal{J}} \eta^{(j)} \left(y' f(t^{(j)}) - \phi(t^{(j)}) \right) \geq 0, \qquad \forall \eta \in ex(U),\ \forall t^{(j)} \in \Omega^{(j)}. \end{aligned}$$

5. Conclusion In this paper we gave an explicit characterization for measure sets that are constrained by shape information on different subdomains. We showed the existence of (closed) convex hull representations, which makes it possible to restrict the optimization over a multi-indexed measure set. This generalization is an important step for applications, since realistic distributions obey different shape laws at different locations.

In our future work we will propose a suitable numerical approach to optimize over these multi-indexed measure sets. This will be necessary, since the present technique, proposed by Popescu [17], is only applicable in the case of elementary generating measure sets.

Appendix A. Proof of Lemma 3.1. The following proofs are based on results obtained by Van Casteren (to be published, cfr. [24]). We start by proving (i) $\Rightarrow$ (ii). Choose $\epsilon > 0$ and choose a sequence of bounded and continuous functions $(g_q)_{q \in \mathbb{N}}$ pointwise decreasing to the indicator function $1_{\Omega \setminus O}$. Then there exists a compact set K_ϵ in O such that

$$\tau(O \setminus K_\epsilon) \leq \frac{\epsilon}{2 \sup_{x \in O \setminus K_\epsilon} g_0(x)}, \qquad \forall \tau \in \mathcal{T}.$$

The sequence $(g_q)_q$ converges pointwise to zero on K_ϵ, hence it converges uniformly on K_ϵ. This means that there exists $q_0 \in \mathbb{N}$ such that $\sup_{x \in K_\epsilon} g_{q_0}(x) \leq \frac{\epsilon}{2}$. Thus, we have for all $\tau \in \mathcal{T}$ that

$$\begin{aligned} \int_O g_{q_0} \, \mathrm{d}\tau &= \int_{K_\epsilon} g_{q_0} \, \mathrm{d}\tau + \int_{O \setminus K_\epsilon} g_{q_0} \, \mathrm{d}\tau \\ &\leq \tau(K_\epsilon) \sup_{K_\epsilon} g_{q_0}(x) + \tau(O \setminus K_\epsilon) \sup_{x \in O \setminus K_\epsilon} g_0(x) \\ &\leq \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon. \end{aligned}$$

Next, we prove the reverse implication (ii) $\Rightarrow$ (i). First, remark that any open set O in $\mathbb{R}^M$ is Polish. This means that we can choose a metric d such that O is complete and contains a dense sequence $(x_p)_{p \in \mathbb{N}}$. Put $B_{p,l} = \{x \in O | d(x_p, x) \leq 2^{-l}\}$ and denote its complement by $B^c_{p,l}$. Choose continuous (and bounded) functions $g_{p,l}$ on Ω such that $1_{B^c_{p,l}} \leq g_{p,l} \leq 1_{B^c_{p,l+1}}$. Put $h_{q,l} = \min_{p=1}^q g_{p,l}$. It can be checked that, for every $l \in \mathbb{N}$, the sequence $(h_{q,l})_q$ decreases pointwise to $1_{\Omega \setminus O}$. So for given $\epsilon > 0$ and for given $l \in \mathbb{N}$ there exists $q_{l,\epsilon} \in \mathbb{N}$ such that

$$\tau\left(\bigcap_{p=1}^{q_{l,\epsilon}} B^c_{p,l} \right) \leq \int_O h_{q_{l,\epsilon},l} \, \mathrm{d}\tau \leq \frac{\epsilon}{2^l}, \qquad \forall \tau \in \mathcal{T}.$$

By summing over l we obtain

$$\tau\left(\bigcup_{l=1}^{\infty} \bigcap_{p=1}^{q_{l,\epsilon}} B^c_{p,l} \right) \leq \epsilon, \qquad \forall \tau \in \mathcal{T}.$$

Put

$$K_\epsilon = \bigcap_{l=1}^{\infty} \bigcup_{p=1}^{q_{l,\epsilon}} B_{p,l} = O \setminus \bigcup_{l=1}^{\infty} \bigcap_{p=1}^{q_{l,\epsilon}} B^c_{p,l}.$$

Then K_ϵ is closed and d-bounded, hence compact. Moreover, $\tau(\Omega\backslash K_\epsilon) = \tau(O\backslash K_\epsilon) \le \epsilon$ for all $\tau \in \mathcal{T}$.

Finally, the equivalence (iii) $\Leftrightarrow$ (ii) is a direct consequence of the fact that for all bounded and continuous functions g_q

$$\sup_{\tau\in\mathcal{T}} \int_O g_q \,\mathrm{d}\tau = \sup_{\pi\in\mathcal{P}} \int_O g_q \,\mathrm{d}\pi.$$

Indeed, for any $\pi \in \mathcal{P}$, there exists a sequence $(\pi_p)_{p\in\mathbb{N}}$ in $\mathcal{H}(\mathcal{T})$ weakly converging to π. By Portmanteau's theorem it follows that

$$\int_O g_q \,\mathrm{d}\pi \le \sup_{p\in\mathbb{N}} \int_O g_q \,\mathrm{d}\pi, \qquad \forall \pi \in \mathcal{P},$$

which implies that

$$\begin{aligned}
\sup_{\pi\in\mathcal{P}} \int_O g_q \,\mathrm{d}\pi &\le \sup_{p\in\mathbb{N}} \int_O g_q \,\mathrm{d}\pi_p \\
&\le \sup_{\pi\in\mathcal{H}(\mathcal{T})} \int_O g_q \,\mathrm{d}\pi \\
&= \sup_{R\in\mathbb{N}} \sup_{\substack{\lambda_r\in[0,1] \\ \tau_r\in\mathcal{T}}} \sum_{r=1}^{R} \lambda_r \int_O g_q \,\mathrm{d}\tau_r \\
&\le \sup_{R\in\mathbb{N}} \sup_{\tau_r\in\mathcal{T}} \max \left\{ \int_O g_q \,\mathrm{d}\tau_1, \dots, \int_O g_q \,\mathrm{d}\tau_R \right\} \\
&\le \sup_{\tau\in\mathcal{T}} \int_O g_q \,\mathrm{d}\tau.
\end{aligned}$$

□

Acknowledgments. The authors would like to thank Jan Van Casteren for the results concerning Lemma 3.1.

References

[1] D. Bertsimas and I. Popescu, *On the relation between option and stock prices: A convex optimization approach*, Oper. Res. **50** (2002), no. 2, 358–374.

[2] ______, *Optimal inequalities in probability theory: A convex optimization approach*, SIAM J. Optim. **15** (2005), no. 3, 780–804.

[3] P. P. Boyle and X. S. Lin, *Bounds on contingent claims based on several assets*, Journal of Financial Economics **46** (1997), no. 3, 383–400.

[4] V. Capasso and D. Bakstein, *An introduction to continuous-time stochastic processes*, Birkhäuser Boston, 2004.

[5] S. H. Cox, Y. Lin, R. Tian, and L. F. Zuluaga, *Bounds for ruin probabilities and value at risk*, Actuarial Research Clearing House Newsletters **1** (2008), http://www.soa.org/library/proceedings/arch/2008/arch-2008-iss1-cox-lin-tian.pdf.

[6] A. De Schepper and B. Heijnen, *Distribution-free option pricing*, Insurance Math. Econom. **40** (2007), no. 2, 179–199.

[7] F. De Vylder and M. J. Goovaerts, *Analytical best upper bounds on stop-loss premiums*, Insurance Math. Econom. **1** (1982), 197–211.

[8] L. El Ghaoui, F. Oustry, and M. Oks, *Worst-case value-at-risk and robust portfolio optimization: A conic programming approach*, Oper. Res. **51** (2003), no. 4, 543–556.

[9] M. J. Goovaerts, J. Haezendonck, and F. De Vylder, *Numerical best bounds on stop-loss premiums*, Insurance Math. and Econom. **1** (1982), 287–302.

[10] D. Han, X. Li, D. Sun, and J. Sun, *Bounding option prices of multi-assets: A semidefinite programming approach*, Pac. J. Optim. **1** (2005), no. 1, 59–79.

[11] C. Lemaréchal and A. Renaud, *A geometric study of duality gaps, with applications*, Math. Program. Ser A **90** (2001), 399–427.

[12] A. Lo, *Semiparametric upper bounds for option prices and expected payoffs*, Journal of Financial Economics **19** (1987), 373–388.

[13] C. L. Mallows, *Generalizations of Tchebycheff's inequalities*, J. Roy. Statist. Soc. Ser B. **18** (1956), no. 2, 139–176.

[14] ______, *A generalization of the Chebyshev inequalities*, Proc. London Math. Soc. **3** (1963), no. 13, 385–412.

[15] H. P. Mulholland and C. A. Rogers, *Representation theorems for distribution functions*, Proc. London Math. Soc. **3** (1958), no. 8, 177–223.

[16] P. E. O'Neil, *Hyperplane cuts of an n-cube*, Discrete Math. **1** (1971), no. 2, 193–195.

[17] I. Popescu, *A semidefinite programming approach to optimal moment bounds for convex classes of distributions*, Math. Oper. Res. **30** (2005), no. 3, 632–657.

[18] M. Reza Emamy-Khansary and M. Ziegler, *New bounds for hypercube slicing numbers*, Discrete Math. Theor. Comput. Sci. **AA** (2001), 155–164.

[19] R. T. Rockafellar, *Convex analysis*, Princeton University Press, Princeton, 1970.

[20] W. W. Rogosinsky, *Moment of non-negative mass*, Proc. Roy. Soc. London Ser. A. **245** (1958), 1–27.

[21] A. Shapiro, *On duality theory of conic linear problems*, Semi-infinite Programming: recent advances (M.A. Goberna and M.A. Lopez, eds) (2001), 135–165.

[22] J. F. Sturm, *Using SeDuMi 1.02, a Matlab toolbox for optimization over symmetric cones*, Optim. Methods Softw. **11-12** (1999), 625–653.

[23] K. C. Toh, M. J. Todd, and R. H. Tütüncü, *SDPT3 - a MATLAB software package for semidefinite programming*, Optim. Methods Softw. **11** (1998), 545–581.

[24] J. Van Casteren, *Markov processes, Feller semigroups and evolution equations (forthcoming monography).*

[25] G. Winkler, *Moment sets of bell-shaped distributions: Extreme points, extremal decomposition and Chebysheff inequalities*, Math. Nach. **215** (2000), no. 1, 161–184.

[26] L. F. Zuluaga, J. F. Peña, and D. Du, *Third-order extensions of Lo's semiparametric bound for European call options*, European J. Oper. Res. **198** (2009), no. 2, 557–570.