SAND2019-2008J

Noname manuscript No.
(will be inserted by the editor)

Higher-Moment Buffered Probability

D. P. Kouri

Received: date / Accepted: date

Abstract In stochastic optimization, probabilities naturally arise as cost functionals and chance constraints. Unfortunately, these functions are difficult to handle both theoretically and computationally. The buffered probability of failure and its subsequent extensions were developed as numerically tractable, conservative surrogates for probabilistic computations. In this manuscript, we introduce the higher-moment buffered probability. Whereas the buffered probability is defined using the Conditional Value-at-Risk, the higher-moment buffered probability is defined using Higher-Moment Coherent Risk measures. In this way, the higher-moment buffered probability encodes information about the magnitude of tail moments, not simply the tail average. We prove that the higher-moment buffered probability is closed, monotonic, quasi-convex and can be computed by solving a smooth one-dimensional convex optimization problem. These properties enable smooth reformulations of both higher-moment buffered probability cost functionals and constraints.

Keywords Risk Measures, Chance Constraints, Stochastic Programming, Conditional Value-at-Risk

Mathematics Subject Classification (2000) 62C05, 65K05, 65K10, 90C15

This research was sponsored by DARPA EQUiPS grant SNL 014150709.

D. P. Kouri
Optimization and Uncertainty Quantification, Sandia National Laboratories
P.O. Box 5800, MS-1320, Albuquerque, NM 87185-1320
E-mail: dpkouri@sandia.gov
Sandia National Laboratories is a multimission laboratory managed and operated by National Technology and Engineering Solutions of Sandia, LLC., a wholly owned subsidiary of Honeywell International, Inc., for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525. This paper describes objective technical results and analysis. Any subjective views or opinions that might be expressed in the paper do not necessarily represent the views of the U.S. Department of Energy or the United States Government.

1 Introduction

Probabilistic functions of the form $P(X; x) := \text{prob}(X > x)$, where X is a random variable representing loss and x is a threshold defining "acceptable" losses, naturally arise in a multitude of optimization applications including engineering design and control under uncertainty [4,11,15] and portfolio optimization [5,18]. In these applications, $P(X; x)$ enters in the form of cost functionals and chance constraints. Unfortunately, $P(X; x)$ may not be very informative if there are important, high-magnitude events that occur with low probability. For example, critical electrical circuits such as static random access memories (SRAMs) often consist of identical cells replicated across millions of instances. For SRAMs, a rare event in a single memory cell may induce a failure for the entire system [22]. In this and similar applications, $P(X; x)$ falls short in quantifying system reliability because it does not encode information about the magnitude of rare events (e.g., system failure). Such application requirements as well as concerns regarding the mathematical and approximation properties of $P(X; x)$ led Rockafellar and Royset to develop the buffered probability of failure in [16], which was subsequently extended to the lower and upper buffered probabilities of exceedance (bPOE) in [12].

In [12], Mafusalov and Uryasev show that the lower and upper bPOEs have numerous desirable mathematical properties. For example, they are monotonic and quasi-convex. In fact, the lower bPOE is closed and is the minimal upper bound for $P(X; x)$ over all closed, quasi-convex, and monotonic functions. Additionally, one can compute the upper bPOE by solving the one-dimensional, convex optimization problem

$$\min_{a \geqslant 0} \; \mathbb{E}[(a(X - x) + 1)_+] \tag{1}$$

where $\mathbb{E}$ denotes the expected value and $(\cdot)_+ = \max\{0, \cdot\}$. Aside from these properties, bPOE also encodes a measure of "tail weight." By definition, the lower bPOE is $(1 - \beta)$ where $\beta \in [0, 1)$ solves

$$x = \text{CVaR}_\beta[X].$$

Here, $\text{CVaR}_\beta[X]$ denotes the Conditional Value-at-Risk (CVaR) of X at confidence level β [12]. Informally, the lower bPOE is the probability that the tail weight is equal to x, i.e., $P(X; \gamma)$ where γ solves

$$\mathbb{E}[X | X > \gamma] = x.$$

Owing to the properties of CVaR, incorporating the lower bPOE as a constraint in an optimization problem results in a constraint on CVaR and hence provides a convex feasible set. In fact, lower bPOE (or equivalently CVaR) constraints provide the best convex approximation for chance constraints [14]. Moreover, when X represents a linear or convex loss, then minimizing bPOE results in a linear or convex program, respectively.

Incorporating the bPOE into nonconvex optimization problems often introduces nonsmoothness that can be circumvented with the addition of slack variables. However, this smooth reformulation introduces an enormous, even infinite, number of variables and constraints (i.e., proportional to the number of scenarios). On the other hand, one can directly solve the nonsmooth, nonconvex optimization problem using, e.g., subgradient or bundle methods [13]. Unfortunately, such methods are often computationally infeasible for applications with expensive objective function evaluations such as optimization problems constrained by partial differential equations (see, e.g., [2,8,20]). Furthermore, in many applications it is important to quantify tail moments, not simply tail averages. To address these concerns, we introduce the higher-moment buffered probability. Analogous to bPOE, we define the higher-moment counterpart using the Higher-Moment Coherent Risk (HMCR) measures [9]. The HMCR measures generalize CVaR by accounting for higher-order tail moments.

In the upcoming sections, we introduce notation and discuss the fundamental properties of HMCR. Using these properties, we then define the higher-moment bPOE and prove that it is closed, monotonic and quasi-convex. To conclude, we derive a one-dimensional, convex optimization formula for computing higher-moment bPOE analogous to (1).

2 Notation and Definitions

Let $(\Omega, \mathcal{F}, \mathbb{P})$ denote a probability space; Ω is the set of outcomes, $\mathcal{F} \subseteq 2^{\Omega}$ is a σ-algebra of events and $\mathbb{P} : \mathcal{F} \to [0,1]$ is a probability measure. Throughout, we denote by $L^p(\Omega) = L^p(\Omega, \mathcal{F}, \mathbb{P})$ with $1 \leqslant p < \infty$ the usual Banach space of $\mathbb{P}$-integrable random variables X satisfying $\|X\|_p := \mathbb{E}[|X|^p]^{1/p} < \infty$. Here, $\mathbb{E}[X]$ denotes the expected value of the random variable X and $\|X\|_p$ is the $L^p(\Omega)$-norm of X. We further denote the probability that X exceeds the threshold $x \in \mathbb{R}$ by

$$P(X; x) := \mathbb{P}(\{\omega \in \Omega \,:\, X(\omega) > x\}),$$

the essential supremum of X by

$$\operatorname{ess\,sup}[X] = \inf\{x \in \mathbb{R} \,:\, X < x \ \mathbb{P}\text{-a.e.}\},$$

and the quantile of X with confidence level $\beta \in (0,1)$ by

$$Q(X; \beta) := \inf\{x \in \mathbb{R} \,:\, P(X; x) \leqslant 1 - \beta\}.$$

In financial applications, the quantile $Q(\cdot; \beta)$ is often called the Value-at-Risk.

A functional $\mathcal{R} : L^p(\Omega) \to \mathbb{R} \cup \{+\infty\}$ is a *coherent* measure of risk in the sense of [1] if it satisfies: For X, $X' \in L^p(\Omega)$ and $t \in \mathbb{R}$,

(R1) *Subadditivity*: $\mathcal{R}[X + X'] \leqslant \mathcal{R}[X] + \mathcal{R}[X']$;
(R2) *Monotonicity*: If $X \leqslant X'$ $\mathbb{P}$-a.e., then $\mathcal{R}[X] \leqslant \mathcal{R}[X']$;
(R3) *Translation Equivariance*: $\mathcal{R}[X + t] = \mathcal{R}[X] + t$;

(R4) *Positive Homogeneity*: If $t \geqslant 0$, then $\mathcal{R}[tX] = t\mathcal{R}[X]$.

The CVaR of a random loss X with confidence level $\beta \in [0,1)$,

$$\bar{Q}(X;\beta) := \frac{1}{1-\beta}\int_\beta^1 Q(X;\alpha)\,\mathrm{d}\alpha = \inf_{\eta\in\mathbb{R}}\left\{\eta + \frac{1}{1-\beta}\mathbb{E}[(X-\eta)_+]\right\},$$

and the HMCR measure of X with order $p \geqslant 1$ and confidence level $\beta \in [0,1)$,

$$\bar{Q}_p(X;\beta) := \inf_{\eta\in\mathbb{R}}\left\{\eta + \frac{1}{1-\beta}\|\,(X-\eta)_+\,\|_p\right\},$$

are coherent measures of risk. To simplify notation, we denote

$$\mathcal{S}_p(X;\beta) := \arg\min\left\{\eta + \frac{1}{1-\beta}\|\,(X-\eta)_+\,\|_p \;:\; \eta\in\mathbb{R}\right\}$$

and note that $\mathcal{S}_p(X;\beta)$ is nonempty when $\beta \in [0,1)$ (see Theorem 2 in [9]). Finally, we recall the biconjugate representation of HMCR,

$$\bar{Q}_p(X;\beta) = \sup\left\{\mathbb{E}[\theta X] \;:\; \theta \in \mathfrak{A}_p(\beta)\right\} \tag{2a}$$

where

$$\mathfrak{A}_p(\beta) := \left\{\theta \in L^q(\Omega) \;:\; \theta \geqslant 0 \;\; \mathbb{P}\text{-a.e.},\; \mathbb{E}[\theta] = 1,\; \|\theta\|_q \leqslant \frac{1}{1-\beta}\right\} \tag{2b}$$

with $\frac{1}{p}+\frac{1}{q} = 1$. See [3, S. 5.3.1] for more details. With these definitions, we note that $\bar{Q}_1(X;\beta) = \bar{Q}(X;\beta)$ and $Q(X;\beta) \in \mathcal{S}_1(X;\beta)$ if $\beta \in (0,1)$. Additionally, we note that if the random variable $X \in L^p(\Omega)$, $p \in [1,\infty)$, is degenerate, i.e., there exists $C \in \mathbb{R}$ such that $X = C$ $\mathbb{P}$-a.e., then $\bar{Q}_p(X;\beta) = C$ for all $\beta \in [0,1)$.

3 Higher-Moment Coherent Risk Measures

The HMCR measures were introduced in [9] as a generalization of CVaR. Both HMCR and CVaR are relevant in risk and reliability analysis because they quantify risks associated with tail events. The HMCR measures are law invariant, finite-valued coherent risk measures on $L^p(\Omega)$. In fact, Dentcheva et al. derive a Kusuoka representation for HMCR in [6].

In an analogous manner to bPOE, we will define our higher-moment extension by "inverting" HMCR. To do this, we must understand certain properties of $\bar{Q}_p(X;\cdot)$. For CVaR, one can show that $\bar{Q}(X;\cdot)$ is continuous, nondecreasing and satisfies the bounds

$$\mathbb{E}[X] \leqslant \bar{Q}(X;\beta) \leqslant \operatorname{ess\,sup}[X] \quad \forall\, \beta \in (0,1).$$

In fact, one has that

$$\bar{Q}(X;0) = \mathbb{E}[X] \quad \text{and} \quad \lim_{\beta\uparrow 1}\bar{Q}(X;\beta) = \operatorname{ess\,sup}[X].$$

We extend these properties to $\bar{Q}_p(X;\cdot)$ for $p > 1$.

Proposition 1 *Let $X \in L^p(\Omega)$ with $1 \leqslant p < \infty$. Then, the mapping $\beta \mapsto \bar{Q}_p(X;\beta)$ is continuous on $[0,1)$.*

Proof This proof generalizes the proof of Proposition 13 in [17]. To begin, we define the functions

$$\theta_a(\gamma) := a + \gamma \| (X-a)_+ \|_p \quad \text{and} \quad \theta(\gamma) := \inf_{a\in\mathbb{R}} \theta_a(\gamma).$$

Since $\theta_a(\cdot)$ is affine for each $a \in \mathbb{R}$, we have

$$\begin{aligned}\theta(\lambda\gamma + (1-\lambda)\gamma') &= \inf_{a\in\mathbb{R}} \left\{\lambda\theta_a(\gamma) + (1-\lambda)\theta_a(\gamma')\right\} \\ &\geqslant \lambda\theta(\gamma) + (1-\lambda)\theta(\gamma')\end{aligned}$$

for all γ, $\gamma' \in \mathbb{R}$ and $\lambda \in [0,1]$, i.e., θ is concave. Since θ is a finite, concave function on $\mathbb{R}$, it is continuous. Composing θ with the continuous function $\beta \mapsto (1-\beta)^{-1} : [0,1) \to \mathbb{R}$ gives the desired result. □

In the subsequent proposition, we characterize the behavior of $\bar{Q}_p(X;\beta)$ at the end points $\beta \in \{0,1\}$.

Proposition 2 *Let $X \in L^p(\Omega)$ with $1 < p < \infty$. Then*

$$\bar{Q}_p(X;0) = \mathbb{E}[X] \quad \textit{and} \quad \lim_{\beta\uparrow 1} \bar{Q}_p(X;\beta) = \operatorname{ess\,sup}[X].$$

Proof The $\beta \uparrow 1$ case was proved in [9], see equation (39) in the Appendix. To prove the $\beta = 0$ case, we note that Jensen's inequality with $\psi(x) = x^q$ ensures that for any $\theta \in \mathfrak{A}_p(\beta)$, we have

$$\frac{1}{1-\beta} \geqslant \|\theta\|_q \geqslant \mathbb{E}[|\theta|] = 1.$$

Therefore, if $\beta = 0$, then any $\theta \in \mathfrak{A}_p(0)$ satisfies

$$\theta \geqslant 0 \ \ \mathbb{P}\text{-a.e.} \quad \text{and} \quad \mathbb{E}[\theta] = \|\theta\|_q = 1.$$

By the equality case in Hölder's inequality ($p, q > 1$), the only such θ is $\theta^\star = 1$ $\mathbb{P}$-a.e., which implies $\mathfrak{A}_p(0) = \{\theta^\star\}$ and $\bar{Q}_p(X;0) = \mathbb{E}[X]$ by (2) as desired. □

Building on Proposition 2, we further investigate the behavior of $\bar{Q}_p(X;\beta)$ for β in the vicinity of 1.

Proposition 3 *Suppose $X \in L^p(\Omega)$ with $1 \leqslant p < \infty$ is not degenerate and $\operatorname{ess\,sup}[X] < \infty$. Then, $\bar{Q}_p(X;\beta) = \operatorname{ess\,sup}[X]$ and $\operatorname{ess\,sup}[X] \in \mathcal{S}_p(X;\beta)$ if and only if $\beta \in [1-\pi_\infty^{1/p}, 1]$ where*

$$\pi_\infty = \pi_\infty[X] := \mathbb{P}(\{\omega \in \Omega \,:\, X(\omega) = \operatorname{ess\,sup}[X]\}).$$

Proof First note that if $\pi_\infty \in (0,1)$, then

$$\| (X-\eta)_+ \|_p \geqslant \pi_\infty^{1/p} (\operatorname{ess\,sup}[X] - \eta)_+ \quad \forall\, \eta \in \mathbb{R},$$

which implies

$$\bar{Q}_p(X;\beta) \geqslant \inf_{\eta\in\mathbb{R}} \left\{ \eta + \frac{\pi_\infty^{1/p}}{1-\beta} (\operatorname{ess\,sup}[X]-\eta)_+ \right\} \quad \forall\, 0 \leqslant \beta < 1.$$

The lower bound is finite if and only if $\beta \in [1-\pi_\infty^{1/p}, 1)$. Moreover, the minimum value is $\operatorname{ess\,sup}[X]$, which is attained at $\eta = \operatorname{ess\,sup}[X]$. This implies that $\bar{Q}_p(X;\beta) = \operatorname{ess\,sup}[X]$ and $\operatorname{ess\,sup}[X] \in \mathcal{S}_p(X;\beta)$ if $\beta \in [1-\pi_\infty^{1/p}, 1]$. We now consider the case that $\beta \in [0, 1-\pi_\infty^{1/p})$. Since the function

$$\gamma(\eta) := \mathbb{P}(\{\omega \in \Omega \,:\, \eta \leqslant X(\omega)\})$$

is left continuous and $\gamma(\eta) \to \pi_\infty$ as $\eta \uparrow \operatorname{ess\,sup}[X]$, we can select $\eta^\star < \operatorname{ess\,sup}[X]$ such that $\gamma(\eta^\star) \leqslant \pi_\infty < (1-\beta)^p$. It then follows that

$$\|(X-\eta^\star)_+\|_p \leqslant \gamma(\eta^\star)^{1/p}(\operatorname{ess\,sup}[X]-\eta^\star) < (1-\beta)(\operatorname{ess\,sup}[X]-\eta^\star).$$

This ensures that $\bar{Q}_p(X;\beta) < \operatorname{ess\,sup}[X]$ for $\beta \in [0, 1-\pi_\infty^{1/p})$ as desired. □

Up to this point, we have demonstrated that HMCR continuously maps $[0,1)$ into $[\mathbb{E}[X], \operatorname{ess\,sup}[X])$. To conclude our HMCR discussion, we prove the following monotonicity results.

Proposition 4 *Suppose $X \in L^p(\Omega)$ with $1 \leqslant p < \infty$ is not degenerate. Then the map $\beta \mapsto \bar{Q}_p(X;\beta)$ is nondecreasing on $[0,1)$ and strictly increasing on $[0, 1-\pi_\infty^{1/p}]$. Moreover, the map $q \mapsto \bar{Q}_q(X;\beta)$ is nondecreasing on $[1,p]$ for fixed $\beta \in [0,1)$.*

Proof The result that $\beta \mapsto \bar{Q}_p(X;\beta)$ is nondecreasing follows from (2). In particular, since $\beta \mapsto (1-\beta)^{-1}$ is strictly increasing on $[0,1)$, we have that $\mathfrak{A}_p(\beta) \subseteq \mathfrak{A}_p(\alpha)$ for any $0 < \beta \leqslant \alpha < 1$, which ensures that $\bar{Q}_p(X;\beta) \leqslant \bar{Q}_p(X;\alpha)$. To prove that $\beta \mapsto \bar{Q}_p(X;\beta)$ is strictly increasing on $[0, 1-\pi_\infty^{1/p}]$, we first note that $\pi_\infty^{1/p} < 1$ since X is not degenerate. Now, let $0 \leqslant \beta < \alpha < 1 - \pi_\infty^{1/p}$. For all $\eta_{p,\alpha} \in \mathcal{S}_p(X;\alpha) \neq \varnothing$, we have $\eta_{p,\alpha} < \operatorname{ess\,sup}[X]$, $\| (X-\eta_{p,\alpha})_+ \|_p > 0$ and

$$\begin{aligned}\bar{Q}_p(X;\alpha) &= \eta_{p,\alpha} + \frac{1}{1-\alpha}\| (X-\eta_{p,\alpha})_+ \|_p \\ &> \eta_{p,\alpha} + \frac{1}{1-\beta}\| (X-\eta_{p,\alpha})_+ \|_p \geqslant \bar{Q}_p(X;\beta).\end{aligned}$$

Hence, $\bar{Q}_p(X;\cdot)$ is strictly increasing on $[0, 1-\pi_\infty^{1/p})$, which extends to $[0, 1-\pi_\infty^{1/p}]$ by continuity. To complete the proof, $q \mapsto \bar{Q}_q(X;\beta)$, $\beta \in (0,1)$, is nondecreasing by [9, Eq. 5]. □

4 Higher-Moment Buffered Probability

Given the threshold x, the lower bPOE is defined using CVaR (see, e.g., [12, Def. 2.1]) as:

$$\bar{P}(X;x) := \begin{cases} 0 & \text{if } x \geqslant \operatorname{ess\,sup}[X]; \\ 1 - \bar{Q}(X;\cdot)^{-1}(x) & \text{if } \mathbb{E}[X] < x < \operatorname{ess\,sup}[X]; \\ 1 & \text{otherwise.} \end{cases}$$

Here, $\beta = \bar{Q}(X;\cdot)^{-1}(x)$ solves the equation

$$x = \bar{Q}(X;\beta) = \mathrm{CVaR}_{\beta}[X]$$

which has a unique solution for $\mathbb{E}[X] < x < \operatorname{ess\,sup}[X]$ since $\bar{Q}(X;\cdot)$ is continuous and strictly increasing on $(0, 1-\pi_\infty)$ where π_∞ is defined in Proposition 3. The upper bPOE is defined analogously to the lower bPOE with the modification that the upper bPOE is equal to π_∞ when $x = \operatorname{ess\,sup}[X]$ [12, Def. 3.9]. Similarly, since $\bar{Q}_p(X;\cdot)$, with fixed $p > 1$, is continuous and strictly increasing on $(0, 1-\pi_\infty^{1/p})$, it also has a well-defined, continuous and strictly increasing inverse function $\bar{Q}_p(X;\cdot)^{-1} : (\mathbb{E}[X], \operatorname{ess\,sup}[X]) \to (0, 1-\pi_\infty^{1/p})$. Thus, we define the lower p^{th} moment bPOE as

$$\bar{P}_p(X;x) := \begin{cases} 0 & \text{if } x \geqslant \operatorname{ess\,sup}[X]; \\ 1 - \bar{Q}_p(X;\cdot)^{-1}(x) & \text{if } \mathbb{E}[X] < x < \operatorname{ess\,sup}[X]; \\ 1 & \text{otherwise.} \end{cases}$$

We define the upper p^{th} moment bPOE analogously to the lower p^{th} moment bPOE with the modification that the upper p^{th} moment bPOE is equal to $\pi_\infty^{1/p}$ when $x = \operatorname{ess\,sup}[X]$. Note that the upper and lower p^{th} moment bPOEs coincide if $\pi_\infty = 0$. If not otherwise specified, we refer to the lower p^{th} moment bPOE as simply the p^{th} moment bPOE. As with lower bPOE, $x \mapsto \bar{P}_p(X;x)$ is nonincreasing on $\mathbb{R}$, strictly decreasing on $(\mathbb{E}[X], \operatorname{ess\,sup}[X])$ and continuous everywhere except at $x = \operatorname{ess\,sup}[X]$ when $\pi_\infty > 0$.

Using the properties of the inverse function $\bar{Q}_p(X;\cdot)^{-1}$, we have the following equivalence:

$$\bar{P}_p(X;x) \leqslant p_0 \iff \bar{Q}_p(X;1-p_0) \leqslant x. \tag{3}$$

Therefore, in the context of optimization, incorporating constraints on the higher-moment buffered probability is equivalent to enforcing constraints on HMCR. On the other hand, when minimizing the higher-moment buffered probability, relationship (3) ensures the following important properties hold.

Proposition 5 *For any $x \in \mathbb{R}$, $\bar{P}_p(\cdot;x)$ is closed, quasi-convex and satisfies the monotonicity property: if X, $X' \in L^p(\Omega)$ with $X \leqslant X'$ $\mathbb{P}$-a.e., then $\bar{P}_p(X;x) \leqslant \bar{P}_p(X';x)$.*

Proof This proof is a generalization of the proof of Proposition 3.5 in [12]. Since $\bar{Q}_p(\cdot;\beta)$ is convex, continuous and finite, the lower level sets

$$\{X \in L^p(\Omega) \,:\, \bar{Q}_p(X;\beta) \leqslant x\}$$

for $x \in \mathbb{R}$ are closed and convex. This combined with the equivalence (3) ensures that the lower level sets of $\bar{P}_p(\cdot;x)$ are also closed and convex, i.e., $\bar{P}_p(\cdot;x)$ is closed and quasi-convex. To conclude, the monotonicity property follows since $\bar{Q}_p(\cdot;\beta)$ is coherent (i.e., satisfies (R2)). In particular, if $X \leqslant X'$ $\mathbb{P}$-a.e., then $\bar{Q}_p(X';\beta) \leqslant x$ implies $\bar{Q}_p(X;\beta) \leqslant x$ and therefore (3) ensures $\bar{P}_p(X';x) \leqslant 1-\beta$ implies $\bar{P}_p(X;x) \leqslant 1-\beta$. □

As with $\bar{P}(X;x)$, we can compute $\bar{P}_p(X;x)$ by solving a one-dimensional, convex optimization problem similar to (1).

Proposition 6 *Fix $X \in L^p(\Omega)$ with $1 \leqslant p < \infty$ and suppose X is not degenerate. Then, for any $\mathbb{E}[X] \leqslant x \leqslant \operatorname{ess\,sup}[X]$, we have that*

$$\bar{P}_p(X;x) = \begin{cases} 0 & \text{if } x = \operatorname{ess\,sup}[X] \\ \min_{a \geqslant 0} \| (a(X-x)+1)_+ \|_p & \text{otherwise.} \end{cases}$$

Proof This proof is a generalization of the proof of Proposition 2.1 in [12]. Suppose without loss of generality that $0 = x < \operatorname{ess\,sup}[X]$. Owing to the properties of $\bar{Q}_p(X;\cdot)$, we can write $\bar{P}_p(X;x)$ as the optimal value

$$\inf_{0 \leqslant \pi \leqslant 1} \pi \quad \text{subject to} \quad \bar{Q}_p(X;1-\pi) \leqslant 0.$$

This minimization problem is equivalently reformulated as

$$\inf_{0 \leqslant \pi \leqslant 1,\, \eta \in \mathbb{R}} \pi \quad \text{subject to} \quad \eta + \frac{1}{\pi}\| (X-\eta)_+ \|_p \leqslant 0.$$

The constraint $\eta + \frac{1}{\pi}\| (X-\eta)_+ \|_p \leqslant 0$ ensures that any feasible η satisfies $\eta < 0$. Otherwise, $\| (X-\eta)_+ \|_p \leqslant -\pi\eta \leqslant 0$ which cannot happen since $\operatorname{ess\,sup}[X] > 0$. Therefore, we can divide the constraint by $|\eta|$. This yields

$$\inf_{0 \leqslant \pi \leqslant 1,\, \eta \in \mathbb{R}} \pi \quad \text{subject to} \quad \| \left(|\eta|^{-1}X + 1\right)_+ \|_p \leqslant \pi$$

or equivalently

$$\inf_{a > 0} \| (aX+1)_+ \|_p.$$

Since $\| (aX+1)_+ \|_p = 1$ when $a = 0$, we can replace the constraint $a > 0$ by $a \geqslant 0$. Moreover, since we assumed $\operatorname{ess\,sup}[X] > 0$, we have

$$\lim_{a \uparrow \infty} \| (aX+1)_+ \|_p = \lim_{a \uparrow \infty} a \| \left(X + a^{-1}\right)_+ \|_p = \infty.$$

This combined with the continuity and convexity of $a \mapsto \| (aX+1)_+ \|_p$ ensures a minimizer exists in $[0,\infty)$ and we can replace the 'inf' with 'min'. To conclude, since HMCR is a coherent risk measure (in particular, it satisfies (R3)), we have that $\bar{Q}_p(X-x;\beta) = \bar{Q}_p(X;\beta) - x$ and thus, $\bar{P}_p(X-x;0) = \bar{P}_p(X;x)$ which proves the desired result. □

Remark 1 The upper p^{th} moment bPOE is equal to

$$\min_{a \geqslant 0} \| (a(X - x) + 1)_+ \|_p \tag{4}$$

for all $x \in \mathbb{R}$ since $\pi_\infty^{1/p} = \min_{a \geqslant 0} \| (a(X - \operatorname{ess\,sup}[X]) + 1)_+ \|_p$. Very recently and independent of this work, the scalar optimization problem (4) was discovered by Zhitlukhin [21] in the context of Sharpe ratios. □

One potential benefit of higher-moment bPOE is differentiability. Differentiability is extremely important for optimization problems in which the objective function is expensive to evaluate. For such problems nonsmooth optimization algorithms are often computationally infeasible due to their typically slow (e.g., (sub)linear) convergence rates [13]. We have the following result regarding differentiability.

Proposition 7 *Fix $X \in L^p(\Omega)$ with $1 < p < \infty$ and suppose*

$$P(X; 0) = \mathbb{P}(\{\omega \in \Omega \,:\, X(\omega) > 0\}) > 0.$$

Then, the function $\| (\cdot)_+ \|_p$ is Fréchet differentiable at X with gradient

$$\frac{(X)_+^{p-1}}{\| (X)_+ \|_p^{p-1}}.$$

Proof First note that $f(x) = (x)_+^p$ is continuously differentiable on $\mathbb{R}$ with derivative $f'(x) = p\,(x)_+^{p-1}$. Moreover, the assumptions of Theorem 7 in [7] hold for f and therefore, $(\cdot)_+^p$ is continuously Fréchet differentiable as a function from $L^p(\Omega)$ into $L^1(\Omega)$ with derivative $p\,(\cdot)_+^{p-1}$. Since $\mathbb{E}[\cdot]$ is a bounded linear functional from $L^1(\Omega)$ into $\mathbb{R}$, it is also continuously Fréchet differentiable and thus $\mathbb{E}[(\cdot)_+^p]$ is continuously Fréchet differentiable from $L^p(\Omega)$ into $\mathbb{R}$ with gradient $p\,(\cdot)_+^{p-1}$. To conclude, we note that $g(x) = x^{1/p}$ is continuously differentiable from $(0, \infty)$ into $\mathbb{R}$ with derivative $g'(x) = p^{-1}x^{1/p-1}$. Since $P(X; 0) > 0$ by assumption, we have that $\| (X)_+ \|_p > 0$ and the result follows from the chain rule. □

In addition to differentiability, there exists a specific ordering of HMCR that leads to an ordering of the higher-moment bPOE. In [9], Krokhmal showed that the minimizers $\mathcal{S}_p(X; \beta)$ associated with $\bar{Q}_p(X; \beta)$, $\beta > 0$, satisfy

$$\eta \in \mathcal{S}_p(X; \beta) \quad \Longrightarrow \quad \| (X - \eta)_+ \|_{p-1} = (1 - \beta)^{\frac{1}{p-1}} \| (X - \eta)_+ \|_p.$$

From this relationship it is easy to see that for any $\eta \in \mathcal{S}_p(X; \beta)$ we have

$$\bar{Q}_p(X; \beta) = \eta + \frac{1}{(1 - \beta_0)^{\frac{p}{p-1}}} \| (X - \eta)_+ \|_{p-1} \geqslant \bar{Q}_{p-1}(X; \beta_1)$$

where $\beta_0 = \beta$ and $\beta_1 = 1 - (1 - \beta)^{\frac{p}{p-1}}$. Continuing in this fashion yields

$$\bar{Q}_p(X; \beta) \geqslant \bar{Q}_{p-k}(X; \beta_k) \tag{5}$$

where $\beta_k = 1-(1-\beta_{k-1})^{\frac{p-(k-1)}{p-k}} = 1-(1-\beta)^{\frac{p}{p-k}}$ for $k = 1,\dots,\lfloor p-1\rfloor$. Now, for any $\mathbb{E}[X] < x < \operatorname{ess\,sup}[X]$, Propositions 1 and 4 guarantee the existence of a unique $\beta \in (0, 1-\pi_\infty^{1/p})$ satisfying $\bar{Q}_p(X;\beta) = x$, or equivalently, $\bar{P}_p(X;x) = 1-\beta$. Therefore, for $p \in \mathbb{N}$ with $p \geqslant 2$, inequality (5) ensures $\bar{Q}_{p-k}(X;\beta_k) \leqslant x$ and the equivalence in (3) produces the bound

$$\bar{P}_{p-k}(X;x) \leqslant 1-\beta_k = (1-\beta)^{\frac{p}{p-k}} = \bar{P}_p(X;x)^{\frac{p}{p-k}}, \tag{6}$$

for any $k = 1,\dots,p-1$. In particular, (6) yields the sequence of bounds

$$P(X;x) \leqslant \bar{P}(X;x) \leqslant \bar{P}_2(X;x)^2 \leqslant \dots \leqslant \bar{P}_p(X;x)^p, \quad p > 2.$$

Hence, $\bar{P}_p(X;x)^p \leqslant \bar{P}_p(X;x)$ provides a conservative, smooth surrogate for estimating $\bar{P}(X;x)$ and $P(X;x)$.

5 Example

In this section, we investigate the first- and second-order bPOE for a standard normal random variable. In this case, we can analytically represent the objective functions for the bPOE computation using the standard normal distribution and density functions denoted by Φ and ϕ, respectively. Suppose X is normally distributed with zero mean and unit variance. Let $x \geqslant 0$, then the random variable

$$Z := (a(X-x)+1)$$

with fixed $a > 0$ is normally distributed with mean $\mu := 1-ax$ and standard deviation $\sigma := a$. Using the moments of the truncated normal distribution [19], the buffered probability that X exceeds x is

$$\bar{P}(X;x) = \min_{a\geqslant 0}\{(1-ax)(1-\Phi(x-1/a)) + a\phi(x-1/a)\}$$

and the second-order buffered probability that X exceeds x is

$$\bar{P}_2(X;x)^2 = \min_{a\geqslant 0}\{((1-ax)^2+a^2)(1-\Phi(x-1/a)) + a(1-ax)\phi(x-1/a)\}.$$

Figure 1 depicts the objective functions for $\bar{P}(X;x)$ (blue) and $\bar{P}_2(X;x)^2$ (red) for $x \in \{1,2,3\}$ and the table in Figure 2 lists the values of the $P(X;x)$, $\bar{P}(X;x)$, $\bar{P}_2(X;x)^2$, and $\bar{P}_2(X;x)$ for varying values of threshold x. These values are plotted in the right image in Figure 2. The second-order bPOE is considerably more conservative than the bPOE and the bPOE is considerably more conservative than the probability of exceedance. However, $\bar{P}_2(X;x)^2$ provides a modest upper bound for $\bar{P}(X;x)$ for this example, which suggests that $\bar{P}_2(X;x)^2$ may be used in place $\bar{P}(X;x)$ when higher moments or differentiability are desired.

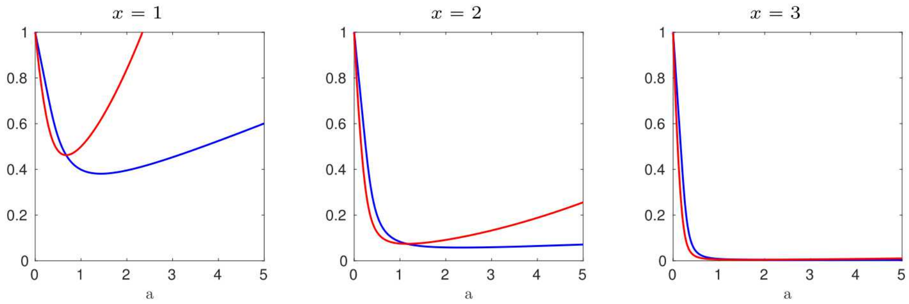

Fig. 1: This figure depicts the objective functions for the first- (blue) and second-order (red) bPOE with $x = 1$ (left), $x = 2$ (center), and $x = 3$ (right).

x	$P(X;x)$	$\bar{P}(X;x)$	$\bar{P}_2(X;x)^2$	$\bar{P}_2(X;x)$
0	0.5	1	1	1
1	0.15866	0.38109	0.46235	0.67996
2	0.02275	0.05799	0.07423	0.27246
3	0.00135	0.00353	0.00463	0.06806
4	0.00003	0.00008	0.00011	0.01056

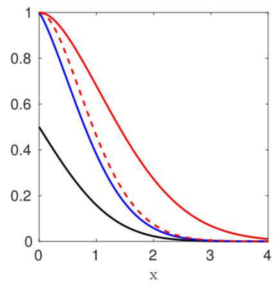

Fig. 2: This table lists the probability of exceedance and the first- and second-order bPOEs for varying thresholds x. The table is plotted in the right image. Here, the black curve is $P(X;x)$, the blue curve is $\bar{P}(X;x)$, the dashed red curve is $\bar{P}_2(X;x)^2$ and the solid red curve is $\bar{P}_2(X;x)$.

6 Application to Optimal Design

In this section, we demonstrate the use of higher-moment bPOE on a three dimensional topological design problem. Let $D := (0,2) \times (0,1) \times (0,1)$ denote the physical domain and consider the volume minimization problem

$$\min_{0 \leqslant \varrho \leqslant 1} \left\{ \int_D \varrho \, \mathrm{d}x =: \mathrm{vol}(\varrho) \right\} \tag{7a}$$

$$\text{subject to} \quad \bar{P}_p \left(\int_D f \cdot S(\varrho) \, \mathrm{d}x; \, c_0 \right) \leqslant p_0 \tag{7b}$$

where the optimization variables $\varrho : D \to [0,1]$ represents the material density and

$$c_0 = 2\mathbb{E}\left[\int_D f \cdot S(\varrho_1) \, \mathrm{d}x \right], \quad \varrho_1 \equiv 1,$$

is the *acceptable* compliance threshold. That is, no feasible design should result in a compliance that is more than two times the average compliance of the solid beam with buffered probability larger than $p_0 = 0.25$. The *uncertain* displacement field $[S(z)](\omega) = u(\omega) : D \to \mathbb{R}^3$, $\omega \in \Omega$, solves the linear elasticity equation

$$\begin{aligned} -\nabla \cdot (\mathbf{E}(\varrho) : \epsilon u(\omega)) &= f(\omega), && \text{in } D, \text{ a.s.} \\ \epsilon u(\omega) &= \frac{1}{2}(\nabla u(\omega) + \nabla u(\omega)^\top), && \text{in } D, \text{ a.s.} \\ \epsilon u(\omega) n &= 0 && \text{on } \Gamma_t, \text{ a.s.} \\ u(\omega) &= 0, && \text{on } \Gamma_d, \text{ a.s.} \end{aligned}$$

Here, $\mathbf{E}(\varrho)$ is the density-dependent material tensor, $\Gamma_d := \{0\} \times [0,1] \times [0,1]$, and $\Gamma_t := \partial\Omega \backslash \Gamma_d$. To obtain a material density, ϱ, that is near 0 (no material) or 1 (material) and that respects a minimal length scale, we employ the SIMP model and density filter described in [10]. The random volumetric load, $f(\omega)$, is centered at $\bar{x} = (2, 0.5, 0.5)$ and has the spherical coordinate form

$$f(\omega) = m(\omega) \begin{bmatrix} \cos(\theta(\omega)) \sin(\phi(\omega)) \\ \sin(\theta(\omega)) \sin(\phi(\omega)) \\ \cos(\phi(\omega)) \end{bmatrix} \delta_{\bar{x}}$$

where the magnitude m is uniformly distributed on $(1, 2)$, the polar angle θ is uniformly distributed on $(225, 315)$ degrees, the azimuthal angle ϕ is uniformly distributed on $(45, 135)$ degrees, and $\delta_{\bar{x}}$ is a Gaussian function centered at $\bar{x}$ with small variance parameter. The left image in Figure 3 depicts the problem setup.

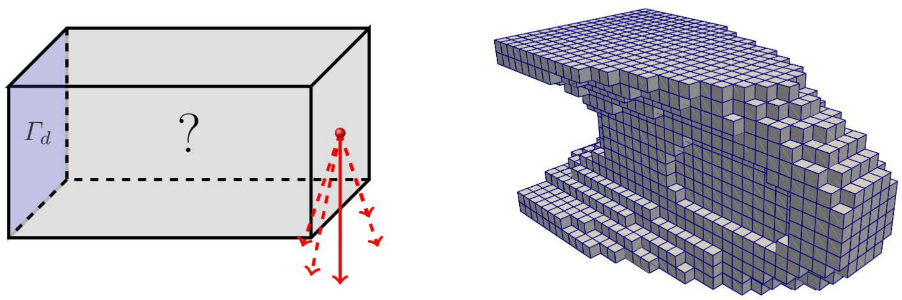

Fig. 3: (Left) The physical domain for the 3D topology optimization problem. The blue face denotes the fixed boundary and the red ball and arrows denote the uncertain load. (Right) The optimal design for the deterministic problem in which the random inputs (m, θ, ϕ) are set to their expected values.

We discretize (7) using continuous piecewise linear finite elements on a uniform quadrilateral mesh of $32 \times 16 \times 16$ elements and we sample $f(\omega)$ with 120 Monte Carlo samples. The right image in Figure 3 depicts the optimal design for the deterministic problem in which (m, θ, ϕ) are replaced by their mean values (MV), whereas Figure 4 depicts the optimal designs for the bPOE formulation using $p \in \{1, 2, 3\}$. By including uncertainty, the optimal design changes from a solid beam (MV) to a hollow shell, which indicates that the shell is more resilient to the uncertain load $f(\omega)$. We list the volume fractions for the optimal topologies in Table 1. As one might expect, the volume of the optimal design increases with p.

	MV	Order 1	Order 2	Order 3
Volume Fraction	49.061%	67.204%	77.369%	80.075%

Table 1: Volume fraction for the optimal designs. Note that the volume increases for the higher-order bPOE formulations.

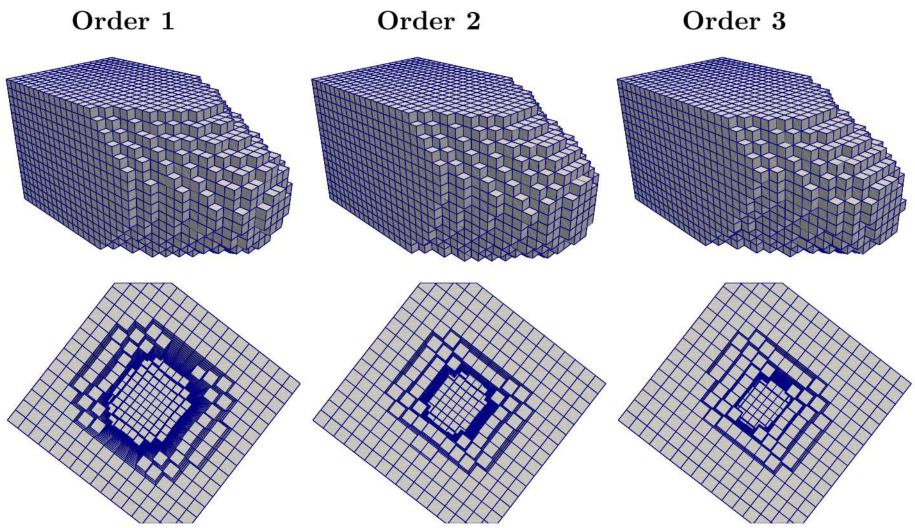

Fig. 4: The optimal designs for first, second and third-order buffered probability.

7 Conclusions

Asking the question "What is the probability that a loss exceeds a prescribed threshold?" is natural in many applications. Unfortunately, such a probability function often does not define an appropriate reliability metric because it does not quantify the magnitude of tail events. To address this concern, we introduced a new probability function called the higher-moment buffered probability. This function is a generalization of the buffered probability [12,16] that accounts for the magnitude of tail moments, not simply the tail average. The higher-moment buffered probability enjoys many desirable properties, similar to those of the buffered probability, with the added benefit of differentiability when incorporated in optimization. Differentiability enables the use of rapidly converging, derivative-based optimization algorithms, which is critical when solving large-scale nonconvex stochastic programs with, e.g., expensive objective function evaluations.

Acknowledgement

The author gratefully acknowledges the comments and suggestions from the reviewers. Their comments helped to simplify the original results in Section 3.

References

1. Artzner, P., Delbaen, F., Eber, J.M., Heath, D.: Coherent measures of risk. Math. Finance **9**(3), 203–228 (1999)
2. Bonfiglio, L., Perdikaris, P., Brizzolara, S., Karniadakis, G.: A multi-fidelity framework for investigating the performance of super-cavitating hydrofoils under uncertain flow conditions. In: 19th AIAA Non-Deterministic Approaches Conference. American Institute of Aeronautics and Astronautics (2017)
3. Cheridito, P., Li, T.: Risk measures on Orlicz hearts. Mathematical Finance **19**(2), 189–214 (2009)
4. Choi, S.K., Grandhi, R., Canfield, R.A.: Reliability-based Structural Design. Springer London (2010)
5. Dempster, M.A.H.: Risk Management: Value at Risk and Beyond. Cambridge University Press (2002)
6. Dentcheva, D., Penev, S., Ruszczyński, A.: Kusuoka representation of higher order dual risk measures. Annals of Operations Research **181**(1), 325–335 (2010)
7. Goldberg, H., Kampowsky, W., Tröltzsch, F.: On Neymtskij operators in Lp-spaces of abstract functions. Mathematische Nachrichten **155**(1), 127–140 (1992)
8. Kouri, D.P., Heinkenschloss, M., Ridzal, D., van Bloemen Waanders, B.G.: Inexact objective function evaluations in a trust-region algorithm for PDE-constrained optimization under uncertainty. SIAM Journal on Scientific Computing **36**(6), A3011–A3029 (2014)
9. Krokhmal, P.A.: Higher moment coherent risk measures. Quantitative Finance **7**(4), 373–387 (2007)
10. Lazarov, B.S., Sigmund, O.: Filters in topology optimization based on Helmholtz-type differential equations. International Journal for Numerical Methods in Engineering **86**(6), 765–781 (2010)
11. Lind, N.C.: Formulation of probabilistic design. Journal of Engineering Mechanics **103**(ASCE 12861 Proceeding) (1977)
12. Mafusalov, A., Uryasev, S.: Buffered probability of Exceedance: Mathematical Properties and Optimization Algorithms. SIAM Journal on Optimization **28**(2), 1077–1103 (2018)
13. Mäkelä, M.M., Neittaanmäki, P.: Nonsmooth optimization. World Scientific Publishing Co. Inc., River Edge, NJ (1992)
14. Nemirovski, A., Shapiro, A.: Convex approximations of chance constrained programs. SIAM Journal on Optimization **17**(4), 969–996 (2007)
15. Rackwitz, R.: Practical Probabilistic Approach to Design. Bulletin d'information. Technical University of Munich, Institut für Bauingenieurwesen III (1976)
16. Rockafellar, R.T., Royset, J.O.: On buffered failure probability in design and optimization of structures. Reliability Engineering & System Safety **95**(5), 499 – 510 (2010)

17. Rockafellar, R.T., Uryasev, S.: Conditional value-at-risk for general loss distributions. Journal of Banking & Finance **26**(7), 1443 – 1471 (2002)
18. Shang, D., Kuzmenko, V., Uryasev, S.: Cash flow matching with risks controlled by buffered probability of exceedance and conditional value-at-risk. Annals of Operations Research **260**(1-2), 501–514 (2018)
19. Tallis, G.M.: The moment generating function of the truncated multi-normal distribution. Journal of the Royal Statistical Society. Series B (Methodological) pp. 223–229 (1961)
20. Tartakovsky, D.M., Guadagnini, A., Riva, M.: Stochastic averaging of nonlinear flows in heterogeneous porous media. Journal of Fluid Mechanics **492**, 47–62 (2003)
21. Zhitlukhin, M.V.: On maximization of the expectation-to-deviation ratio of a random variable. Russian Mathematical Surveys **72**(4), 765–766 (2017)
22. Zhou, Q., Choudhury, M.R., Mohanram, K.: Design optimization for robustness to single event upsets. 2014 IEEE 32nd VLSI Test Symposium (VTS) **0**, 202–207 (2006)