Submitted to *Mathematics of Operations Research*
manuscript MOR-2013-340

Proper Efficiency and Tradeoffs in Multiple Criteria and Stochastic Expected-Value Optimization

Alexander Engau
Associate Professor, Mathematical and Statistical Sciences, University of Colorado Denver, United States
Visiting Professor, Rowe School of Business, Faculty of Management, Dalhousie University, Canada
aengau@alumni.clemson.edu

The mathematical equivalence between linear scalarizations in multiobjective programming and expected-value functions in stochastic optimization suggests to investigate and establish further conceptual analogies between these two areas. In this paper, we focus on the notion of proper efficiency that allows us to provide a first comprehensive analysis of solution and scenario tradeoffs in stochastic optimization. In generalization of two standard characterizations of properly efficient solutions using weighted sums and augmented weighted Tchebycheff norms for finitely many criteria, we show that these results are generally false for infinitely many criteria. In particular, these observations motivate a slightly modified definition to prove that expected-value optimization over continuous random variables still yields bounded tradeoffs almost everywhere in general. Further consequences and practical implications of these results for decision-making under uncertainty and its related theory and methodology of multiple criteria, stochastic and robust optimization are discussed.

Key words: proper efficiency; tradeoffs; multicriteria optimization; multiobjective programming; stochastic optimization; stochastic programming; robust optimization; linear scalarization; weighted sum method; augmented Tchebycheff norm; expected value function; decision-making under uncertainty
MSC2000 subject classification: Primary: 90C29; secondary: 90C15, 90B50

1. Introduction Optimization under uncertainty has been studied since the foundational work by Beale [3], Charnes and Cooper [12] and Dantzig [15] and remains one of the most active research areas in optimization and operations research in general. It traditionally includes methods from stochastic programming which assume that uncertainties are probabilistic and can be quantified using probability distributions or other statistical techniques [7, 35, 61], the more recent but already similarly well-established paradigm of robust optimization which considers general uncertainty sets without such assumption [4, 6, 25, 45], and a growing number of contributions to fuzzy optimization which is based on the membership concept for fuzzy sets and further generalized notions of uncertainty [47]. Despite their significant differences in formulation and solution concepts, many of these approaches have in common that an original stochastic problem is eventually replaced by some (in a certain sense equivalent) deterministic counterpart which can be solved and analyzed using one of the many methods from deterministic optimization or decision making.

Recently, there also has been increased interest in using the theory and methodology of multiple criteria optimization (MCO) for decision making under uncertainty and robust optimization in particular. Based on the scenario-interpretation of random realizations in a discrete uncertainty set by Kouvelis and Yu [45], Kouvelis and Sayın [43, 44] use an equivalent minimax formulation for MCO to develop a new two-stage, robust optimization algorithm. Hu and Mehrotra [32] also propose a minimax approach for risk-averse MCO that leads to the new concepts of a robust value, robust

Pareto optimality and robust stochastic Pareto optimality. A different notion of Pareto robustness is proposed by Iancu and Trichakis [33] who use the standard concept of Pareto optimality to further distinguish between solutions that are equally robust in a worst-case scenario but still allow for significant tradeoffs in other scenarios, which is usually ignored by the classical robust formulation. Bonnel and Collonge [8] use a sample average approximation method to minimize the expected value of a real-valued random function which is defined over the Pareto set of a general stochastic MCO. Schöbel [60] also uses Pareto solutions for a generalization of the concept of light robustness which was originally introduced by Fischetti and Monaci [23] to relax strict robustness in the context of railway timetabling. To illustrate these new methodological contributions, some of the other papers also discuss relevant practical applications to project selection, revenue management and crop planting [32] or portfolio optimization, inventory management and project management [33]. Some other recent applications of robust MCO include the optimal power flow and dispatch of distributed energy resources in microgrids [16], internet routing in telecommunications [17], robust data classification in supervised learning [21], logistics and supply chain management in exploration and production [22], and the optimal design of water distribution networks [54].

From an alternative theoretical perspective, few other authors have addressed the more fundamental similarities between MCO and stochastic or robust optimization. The strong link between the mathematics of MCO and the theory of decisions under uncertainty as developed by Fishburn [24] and Keeney and Raiffa [40] is pointed out clearly by Ogryczak [52] who states that "most of the classical solution concepts commonly used in MCO have their roots (or equivalents) in some approaches to handle uncertainty in the decision analysis." He then continues to raise the rhetorical question whether new advances in MCO may in response provide insightful feedback to support decision-making under risk and shows affirmatively how certain symmetric and equitable risk aversion preferences can be modeled successfully with new MCO methodology. Restricted to the case of finitely many criteria, he also mentions the mathematical equivalence between expected values and maximum regrets in relation to weighted sums and achievement functions respectively. Following his observation, more recent concepts of robustness and stochastic programming have been shown to be special cases of certain other linear or generally nonlinear scalarization functions by Klamroth *et al.* [42], and to be related also to the more general framework of set-valued optimization by Ide *et al.* [34]. Again under the assumption that the set of criteria or scenarios is finite, especially the first of these two related papers shows that each of the considered uncertain optimization problems still has a deterministic multicriteria counterpart whose solutions can provide relevant tradeoff information to facilitate the choice of a most preferred decision.

Despite these recognized similarities in problem formulation in addition to the central role played by a decision maker due to the inadequacy of a unique solution concept besides the natural dominance relation, other authors have questioned the applicability of multiple criteria decision analysis and optimization for decisions under risk [31] and robustness in general [57]. In their qualitative critiques, these authors argue that significant conceptual differences remain and thus promote a more careful distinction between these different problems. Among these differences, Hites *et al.* [31] remark that while many multicriteria methods assign weights and an associated notion of importance to different criteria, a corresponding concept in robustness is not well-defined. Arguably, this concern does not apply for stochastic optimization in which uncertainty is modeled using probability distributions so that such weights are provided in a very natural way. Another existing criticism states that these weights were absolutely given probabilities and thus nonnegotiable so that any consideration of alternative weights or tradeoffs was unnecessary, but this seems to ignore that in all but the simplest cases the statistical inference of a suitable probability distribution is far from trivial and subject to data errors and general inaccuracies itself. In consequence, the consideration of different distributions including weights or densities with their own inherent tradeoffs is often especially important due to significant effects on a decision's overall stability and sensitivity also in practice [1, 53, 56]. More substantially, Hites *et al.* [31] also highlight that the number of scenarios in both robust and stochastic optimization may be very large or even infinite whereas multicriteria methods typically require that the number of criteria is reasonably small and finite in particular. Indeed, the possibility of continuous uncertainty sets goes back to the early work on inexact linear and convex programming by Soyster [62] and is very common both for robust optimization [4, 6] as well as for stochastic programs involving random variables with continuous or discrete probability distributions over compact or unbounded supports respectively. Thus motivated, to address and further explore this current limitation is one of the main objectives in this present paper.

1.1. Motivation and Scope of Paper Based on the mathematical equivalence between weighted sums and expected values in multiple criteria and stochastic optimization respectively, in this paper we make use of this analogy specifically to investigate the role of proper efficiency and its implications to solution or scenario tradeoffs. The definition and characterization of these two general concepts are of significant theoretical and computational interest in optimization, decision making and economics to prevent solutions with unbounded marginal rates of substitution. The notion of proper efficiency was first defined by Kuhn and Tucker [46] for a vector-valued criterion function $f(x) = (f_1(x), \dots, f_p(x))$ from a subset $X \subseteq \mathbb{R}^n$ to $\mathbb{R}^p$ under the assumption that all component functions are continuously differentiable, and later refined by Geoffrion [26] without this condition. In each case, the authors show that solutions that maximize linear scalarizations of the form $\sum_{i=1}^p w_i f_i(x)$ with strictly positive parameters ("weights") $w_i > 0$ are properly efficient so that all their remaining tradeoffs are bounded, and that the reverse statement remains true under suitable convexity conditions on the criterion functions and the underlying feasible set.

In view of our previous discussion and review of the literature, such results are of high interest and bear immediate relevance also for applications in stochastic optimization and decision-making under uncertainty where a single real-valued criterion function $f(x, \tau)$ depends on both a decision vector x and a random vector τ. If this random vector has a discrete probability distribution with a finite number of possible scenario realizations t_i and nonzero probability masses $w_i > 0$ for all $i = 1, \dots, p$, then we may set $f_i(x) = f(x, t_i)$ and replace the criterion vector of possible outcomes $(f_1(x), \dots, f_p(x))$ with its expectation $E(x) = \sum_{i=1}^p w_i f_i(x)$. Hence, in this case it follows that a solution that maximizes an expected value is properly efficient also in a stochastic sense, namely that there is no other feasible decision whose gain-to-loss ratio from one possible outcome or scenario to another outcome or scenario is still infinitely large.

It now seems natural to ask whether this property is also satisfied if τ has a general discrete or continuous probability distribution whose support T of realizations of τ, or analogously, whose number of associated criterion functions $f(x, t)$ is either countably or uncountably infinite respectively. Mathematically, this is equivalent to the question whether the definition and characterization of properly efficient solutions can be generalized from weighted sums $\sum_{i=1}^p w_i f_i(x)$ to weighted infinite series $\sum_{i=1}^\infty w_i f_i(x)$ and to proper or improper integrals of the form $\int_T w(t) y(t)\, \mathrm{d}\, t$, where the strictly positive probability mass or weight vector $w > 0$ is replaced by a probability density function $w(t)$ and each scenario outcome $y(t) = f(x, t)$ now corresponds to a real-valued function. An affirmative answer to this question would give further theoretical support for expected-value optimization in stochastic programming to yield optimal solutions that are always guaranteed to also be properly efficient in the sense of bounded scenario tradeoffs.

Maybe surprisingly but complementary to the recent observation that classical robust solutions also are not necessarily efficient in the sense of Pareto optimality [33], it turns out that the similarly desirable result of properly efficient outcomes in stochastic optimization cannot be guaranteed in general. To the best of the author's knowledge, a related finding for general vector optimization was first made by Winkler [66] who notes that in the context of continuous real-valued functions, "a point received by positive linear scalarization need not be proper efficient in the sense of Geoffrion."

Independent of this earlier paper, we had made a similar observation for the ℓ^∞ sequence space [20] but could still prove that for multiobjective programs with countably many criteria, or analogously, for stochastic programs with discrete random variables, we can maintain existing characterizations of proper efficiency after a slight modification to Geoffrion's original definition [26].

The comparison of these definitions in the sense of Geoffrion [20, 26, 66] to several other notions of proper efficiency including those by Benson [5], Borwein [9], Borwein and Zhuang [10], Hartley [28] and Henig [30] is discussed in these former papers and several other review articles or monographs [27, 37, 39, 59] that are also included in the author's previous bibliographic survey [18]. In addition, the interesting relationships between existence and density results for proper efficiency based on the Arrow-Barankin-Blackwell theorem [2] and related more recent results [14] has also already been analyzed in some detail by Truong [65]. Moreover, it has been shown that the definition and characterization of some of these other concepts that are based on more general ordering cones are often rather straightforward to extend or generalize regardless of whether the number of criteria is finite or infinite [20]. Hence, and in contrast, our new results based on Geoffrion's pairwise definition are particularly interesting due to (a) their surprising negation of expected results and some natural conjecture by Winkler [66], and (b) their specific insights into the more common economic interpretation of pairwise rates of substitution ("ceteris paribus") rather than alternative cone-based tradeoff directions [48, 49, 50, 51] that have clear theoretical meaning but remain without a similarly practical implication for (infinite-dimensional) applications in practice.

1.2. Main Contributions and Outline The new results and their consequences that are summarized in this present paper make several contributions to the existing theory and methodology of both multiple criteria optimization and decision-making under uncertainty. While the strong link between these two areas has been well recognized and is now subject both to increasingly active research [8, 32, 33, 34, 42, 60] and ongoing discussion [31, 57], a substantial criticism remains concerning the frequent limitation of methods from multiobjective programming and decision making to only finitely many criteria. Using the more general framework of vector optimization which is well-established also over infinite-dimensional spaces [37], however, we can show that conceptual analogies can be established in a much broader context and thereby provide valuable insight specifically for the further characterization of those solution and scenario tradeoffs that may result from stochastic expected-value optimization and related approaches in robust optimization.

In particular, in this paper we provide a first comprehensive analysis of such tradeoffs based on the notion of proper efficiency and its relationship to marginal rates of substitutions both of which have received significant attention in optimization, decision-making and economics but so far seem without corresponding analog in the study of optimization problems under uncertainty. In fact, whereas the proper efficiency of solutions that maximize expected values of finitely discrete random variables is an immediate consequence of a standard result in multiobjective programming, corresponding results for generally discrete or continuous random variables require a more careful analysis and turn out to be false in general. The mathematical "proofs" of these negative results are naturally presented in the form of counterexamples for which we have chosen simplicity over alternative but much more complex practical applications to keep this paper relatively shorter in length and focused primarily on its theoretical contribution. However, we suspect that the existence and generation of improperly efficient solutions with remaining unbounded tradeoffs are equally present or even more likely especially in realistic and more complicated settings in practice.

Moreover, and although these generalized characterizations of proper efficiency from finitely to countably and uncountably many criteria do not continue to hold anymore using weighting methods and augmented Tchebycheff norms, we can prove that optimal solutions to these scalar maximum problems are still properly efficient almost everywhere. These new results are based on suitable modifications of the original formulation of proper efficiency by Geoffrion [26] and its later reformulation in the context of real-valued continuous functions by Winkler [66]. In addition, we are able to show that the latter's conjecture that all properly efficient solutions can still be computed as solutions using weighting methods is false for the standard weighting method that we consider in this paper, even under the usual and otherwise sufficient convexity assumptions. Nevertheless, and despite these rather surprising and unexpected findings, our remaining results are somewhat relieving at least from a decision-theoretic point of view by confirming that the typical solutions to stochastic programming approaches may not prevent the existence of unbounded tradeoffs in general but can reduce their likelihood to be of probability zero.

The remaining paper is structured as follows. Section 2 reviews the vector maximum problem together with the standard definition and characterization of proper efficiency for finitely many criteria by Geoffrion [26] and their extension to countably many criteria from our earlier paper [20]. Together, these results provide a complete analysis of tradeoffs for stochastic expected-value optimization with finite or infinite discrete random variables. In addition, based on these former results we can now offer some further insight into the possible advantages to combine or relax stochastic programming by an additional contribution based on robustness or maximum regret. Section 3 gives the analogous problem formulation in the context discussed by Winkler [66] that we can associate with continuous random variables over compact supports. Our main contributions in this section are several new results that demonstrate by proof or counterexample which characterizations of proper efficiency remain valid or become invalid for uncountably many criteria in general. The other main contributions of this paper are given in Section 4 and includes our new definition and two theorems with sufficient conditions for properly efficient solutions almost everywhere using weighting methods and augmented weighted Tchebycheff norms respectively. This section also offers a brief discussion to extend this definition and our results to continuous random variables over other and generally unbounded supports. Finally, a few other consequences and interesting implications of our results as well as avenues for further work are summarized in Section 5.

2. Proper Efficiency for Countably Many Criteria We begin to consider the original problem statement by Geoffrion [26] in a slightly generalized form. Given a finite or countably infinite index set I, an associated criterion function $f(x) = (f_i(x) \colon i \in I)$ and a set of feasible points $X \subseteq \mathbb{R}^n$, the vector maximum problem

$$\max f(x) \text{ subject to } x \in X \qquad \text{(VMP)}$$

is the problem of finding all points that are efficient: a point $\bar{x}$ is said to be *efficient* if $\bar{x} \in X$ is feasible and if there exists no other feasible point $x \in X$ such that $f_i(x) \geq f_i(\bar{x})$ for all $i \in I$ and $f_{i_0}(x) > f_{i_0}(\bar{x})$ for some $i_0 \in I$. Because this notion of efficiency uses the natural ordering cone which corresponds to the (finite or infinite-dimensional closed convex pointed) nonnegative orthant, the idea of bounded tradeoffs leads naturally to the following componentwise definition.

Definition 1 (Geoffrion [26]). A point $\bar{x} \in X$ is said to be a *properly efficient solution* in (VMP) if it is efficient and if there exists a scalar $M > 0$ such that, for each $i \in I$, we have

$$\frac{f_i(x) - f_i(\bar{x})}{f_j(\bar{x}) - f_j(x)} \leq M$$

for some $j \in I$ such that $f_j(x) < f_j(\bar{x})$ whenever $x \in X$ and $f_i(x) > f_i(\bar{x})$.

In addition, Geoffrion calls a point $\bar{x}$ *improperly efficient* if it is efficient but not properly efficient, thus meaning that to every scalar $M > 0$ (no matter how large) there is a point $x \in X$ and an index $i \in I$ such that $f_i(x) - f_i(\bar{x}) > M\,(f_j(\bar{x}) - f_j(x))$ for all $j \in J$ such that $f_j(x) < f_j(\bar{x})$. Under the assumption that I is finite, Geoffrion continues to remark that because "there is but a finite number of criteria we see that for some criterion i_0, the marginal gain in f_{i_0} can be made arbitrarily large relative to each of the marginal losses in other criteria" and further, that $\bar{x}$ "certainly seems undesirable." This intention that proper efficiency shall prevent the existence of unbounded tradeoffs is maintained for a general countable set I by a slightly reworded definition.

DEFINITION 2 (ENGAU [20]). A point $\bar{x} \in X$ is said to be *properly efficient in the sense of Geoffrion* if it is efficient and if, for each $i \in I$, there exists a scalar $M_i > 0$ such that

$$\frac{f_i(x) - f_i(\bar{x})}{f_j(\bar{x}) - f_j(x)} \le M_i$$

for some $j \in I$ such that $f_j(x) < f_j(\bar{x})$ whenever $x \in X$ and $f_i(x) > f_i(\bar{x})$.

It is clear that for the special case that I is finite, Definitions 1 and 2 are equivalent by choosing $M = \max\{M_i \colon i \in I\}$ or $M_i = M$ for all $i \in I$. Similarly, if I is infinite and a point $\bar{x}$ satisfies Definition 1 with a scalar $M > 0$, then it is also properly efficient by Definition 2 with $M_i = M$ for all $i \in I$. These observations are summarized in the following proposition.

PROPOSITION 1. *Geoffrion proper efficiency by Definition 1 is sufficient for proper efficiency in the sense of Geoffrion by Definition 2 and necessary if the number of criteria is finite.*

Definition 2 is less restrictive in general, however, and seems to better agree with the original intention by Geoffrion to "propose a slightly restricted definition of efficiency that (a) eliminates efficient points of a certain anomalous type; and (b) lends itself to more satisfactory characterization" also if I is infinite. In particular, together with the result in Proposition 1 the following Example 1 suggests that Geoffrion's original definition may be too strong for infinitely many criteria and eliminate too many efficient points despite any clear anomalies. Similarly, the investigation of a satisfactory theoretical characterizations is addressed by Proposition 2 and Example 2 in the following Section 2.1 and will then be discussed throughout the rest of this paper.

EXAMPLE 1. Let $I = \{0, 1, 2, \ldots\}$ and consider the linear functions $f_0(x) = x$ and $f_i(x) = 1 - 2^i x$ for all $i \ge 1$ on $X = [0, 1]$. Because f_0 is strictly increasing whereas all other f_i are strictly decreasing, all $x \in X$ are efficient but arguably do not exhibit any anomalous behavior. However, it is not difficult to see that only $x = 0$ is properly efficient with respect to Definition 1 whereas all x are still properly efficient only with respect to Definition 2. $\diamondsuit$

In addition, this example also demonstrates that a related corollary statement by Isermann [36], "that each efficient solution of a linear vector maximum problem is also properly efficient" is generally false with respect to Definition 1 if the number of criteria is infinite but remains true using Definition 2. Other implications of this improved definition for countably many criteria are briefly summarized in the two following sections which at the same time provide further preliminaries for our new results in Sections 3 and 4.

2.1. Characterization Using Weighting Methods Let $w = (w_i \colon i \in I)$ be a vector (if I is finite) or sequence (if I is countable) of nonnegative parameters ("weights") normalized according to $\sum_{i \in I} w_i = 1$ and consider the following scalar maximum problem:

$$\max \sum_{i \in I} w_i f_i(x) \text{ subject to } x \in X. \tag{P$_w$}$$

Note that this normalization is without loss of generality if I is finite but restricts the possible choice of w if I is infinite. The fundamental results characterizing properly efficient solutions in (VMP) in terms of the optimal solutions in (P$_w$) are given in the following theorem [26].

THEOREM 1. *Consider the vector maximum problem with a finite index set I.*

(i) Let $w_i > 0$ for all $i \in I$. If $\bar{x}$ is optimal in (P$_w$), then $\bar{x}$ is properly efficient in (VMP).

(ii) Let X be a convex set and let each f_i be concave on X. Then $\bar{x}$ is properly efficient in (VMP) if and only if $\bar{x}$ is optimal in (P$_w$) for some w with strictly positive components.

Based on the equivalence in definition if the index set I is finite, it is clear that Theorem 1 holds both for proper efficiency with respect to Definition 1 and with respect to Definition 2. Moreover, for the case that I is infinite, we have proven the following result [20].

PROPOSITION 2. *If the index set I is countably infinite, then the statements in Theorem 1 remain true for proper efficiency in the sense of Geoffrion by Definition 2.*

The following example shows that the results in Theorem 1 are generally false for Geoffrion proper efficiency by Definition 1, however, using the same linear functions already defined in Example 1.

EXAMPLE 2. In Example 1, let $w_i = 3/4^{i+1} > 0$ for all $i \geq 0$ so that

$$\sum_{i=0}^{\infty} w_i = \frac{3}{4}\left(\sum_{i=0}^{\infty} \frac{1}{4^i}\right) = \frac{3}{4}\left(\frac{1}{1-1/4}\right) = 1.$$

For this particular choice of w, it follows that

$$\begin{aligned}\sum_{i=0}^{\infty} w_i f_i(x) &= \frac{3}{4}\left(x + \sum_{i=1}^{\infty} 4^{-i}(1-2^i x)\right)\\ &= \frac{3}{4}\left(x + \left(\frac{1}{1-1/4} - 1\right) - \left(\frac{1}{1-1/2} - 1\right)x\right)\\ &= \frac{3}{4}\left(x + \frac{1}{3} - x\right) = \frac{1}{4}\end{aligned}$$

so that each x that is feasible is also optimal in (P_w) with strictly positive $w_i > 0$ for all $i \in I$. However, from Example 1, only $x = 0$ is properly efficient with respect to Definition 1 whereas all x are still properly efficient with respect to Definition 2. $\Diamond$

Hence, Proposition 2 and Example 2 demonstrate that Definition 2 is a quite natural modification of proper efficiency and improves Geoffrion's original definition also with respect to intention (b).

2.2. Characterization Using Augmented Tchebycheff Norms Based on the original definition by Geoffrion [26], several other authors have shown that the characterization of properly efficient solutions for convex problems can be extended to generally nonconvex problems using combinations of linear scalarizations and weighted l_∞ or Tchebycheff norms [11] including the modified weighted Tchebycheff norm [13, 38] and the augmented weighted Tchebycheff norm [63, 64]. The definition of these "norms" depends on the existence of a utopia point u with components

$$u_i = \sup_{x \in X} \{f_i(x)\} + \epsilon \text{ for all } i \in I \tag{1}$$

where ϵ is some arbitrarily small positive number. Extending this line of work to the general countable case, the next result remains valid for proper efficiency in the sense of Geoffrion by Definition 2 regardless of whether I is finite or countably infinite [20].

THEOREM 2. *Let $w_i > 0$ and u_i be defined as in (1) for all $i \in I$. A feasible point $\bar{x} \in X$ is properly efficient in (VMP) if and only if there exist scalars $\alpha > 0$ and $v_i > 0$ for all $i \in I$ such that $\bar{x}$ is optimal for the scalar maximum problem:*

$$\max \sum_{i \in I} w_i f_i(x) - \alpha \sup_{i \in I}\{v_i(u_i - f_i(x))\} \text{ subject to } x \in X. \tag{P_∞}$$

While the original definitions of modified and augmented Tchebycheff norms take on a slightly different form [13, 38, 63, 64], the above formulation is particularly interesting in the context of stochastic and robust optimization for decision-making under uncertainty. Specifically, here note that the objective itself has the form of a two-factor weighted sum that combines or relaxes the standard expected value (to be maximized) by an additional contribution based on robustness or weighted maximum regret (to be minimized and thus subtracted). While this interpretation of the second term as worst-case regret is particularly true if u is utopian, however, in the proof of Theorem 2 this condition on u is only required for the necessary condition of proper efficiency. In particular, the same proof shows that an optimal solution to (P_∞) is still properly efficient in (VMP) even if u is chosen arbitrarily or non-utopian in which case the terms $u_i - f_i(x)$ may not only represent regret, if $f_i(x) < u_i$, but a certain level of achievement, if $f_i(x) > u_i$ [20]. Hence, from a decision-theoretic point of view, this implies that any properly efficient solution can always compromise between its expected value and its worst-case regret and that any solution that augments expected-value optimization by an arbitrary achievement level criterion will be properly efficient; in fact, these achievement levels could also be replaced by a suitable utopia point, in principle.

3. Proper Efficiency for Uncountably Many Criteria For uncountably many criteria, we now consider the following problem formulation that was similarly studied by Winkler [66]. Given a compact subset T of a separable Banach space and a subset $Y \subseteq \mathcal{C}(T)$ of the space of continuous real-valued functions on T, the problem

$$\max y(t) \text{ subject to } y \in Y \tag{VMP$'$}$$

is the problem of finding all functions that are efficient: a function $\bar{y}$ is said to be *efficient* if $\bar{y} \in Y$ is feasible and if there exists no other feasible function $y \in Y$ such that $y(t) \geq \bar{y}(t)$ for all $t \in T$ and $y(t_0) > \bar{y}(t_0)$ for some $t_0 \in T$. Hence, similar to (VMP) for countable many criteria, here (VMP$'$) is again based on the natural ordering cone of nonnegative functions which is closed convex and pointed and has nonempty interior but in general is neither well-based nor nuclear and does not imply the Daniell property [41, 66]. The corresponding definition of proper efficiency formulated by Winkler is rephrased here for maximization analogously to Definition 1.

Definition 3 (Winkler [66]). A function $\bar{y} \in Y$ is said to be a *properly efficient solution* in (VMP$'$) if it is efficient and if there exists a scalar $\delta > 0$ such that, for each $t \in T$ we have

$$\frac{y(t) - \bar{y}(t)}{\bar{y}(t_0) - y(t_0)} \leq \delta$$

for some $t_0 \in T$ such that $y(t_0) < \bar{y}(t_0)$ whenever $y \in Y$ and $y(t) > \bar{y}(t)$.

In particular, because it is clear that this inequality is satisfied for any $\delta > 0$ and any $t_0 \in T$ such that $y(t_0) < \bar{y}(t_0)$ also if $y(t) \leq \bar{y}(t)$, an equivalent statement of Definition 3 is that there exists a scalar $\delta > 0$ such that, for each $y \in Y$, we have

$$y(t) - \bar{y}(t) \leq \delta\left(\bar{y}(t_0) - y(t_0)\right) \text{ for all } t \in T$$

for some $t_0 \in T$. It follows that this inequality must be satisfied especially for

$$t_0 = t_0(y, \bar{y}) = \arg\max_{t \in T} \{\bar{y}(t) - y(t)\}$$

which is well-defined because all $y \in Y$ are continuous and because T is compact, and which satisfies $y(t_0) < \bar{y}(t_0)$ whenever $y(t) > \bar{y}(t)$ because $\bar{y}$ is efficient. Hence, under the stated assumptions, we have the following result.

Proposition 3. *An efficient function $\bar{y}$ is properly efficient in (VMP$'$) if and only if there exists a scalar $\delta > 0$ such that, for each $y \in Y$, $t \in T$ and $t_0 = \arg\max_{t \in T}\{\bar{y}(t) - y(t)\}$ we have*

$$y(t) - \bar{y}(t) \leq \delta\left(\bar{y}(t_0) - y(t_0)\right).$$

Similar to the limitation of Definition 1 in Examples 1 and 2 for countably many criteria, however, it turns out that Definition 3 also does not maintain Geoffrion's original intention to only eliminate anomalous solutions and still allow for a satisfactory characterization using the generalized scalar maximum problem:

$$\max \int_T w(t)y(t)\,\mathrm{d}\,t \text{ subject to } y \in Y \tag{P'_w}$$

where analogously to (P_w) we assume that $w \geq 0$ is finitely integrable or, without loss of generality, normalized according to $\int_T w(t)\,\mathrm{d}\,t = 1$.

Example 3. Let $T = [0,1]$ and consider the set $Y = \{y_x \in \mathcal{C}(T) \colon x \geq 1\}$ of piecewise linear functions defined by

$$y_x(t) = \max\{2x(1-xt), 0\} \text{ for all } t \in T.$$

It is not difficult to see that for each $\bar{x} \geq 1$, the function $y_{\bar{x}}$ is efficient but not properly efficient with respect to Definition 3: for every $x > \bar{x}$ we have

$$\max_{t\in T}\{y_{\bar{x}}(t) - y_x(t)\} = y_{\bar{x}}(1/x) < y_{\bar{x}}(0) = 2\bar{x}$$

whereas $y_x(0) - y_{\bar{x}}(0) = 2(x - \bar{x})$ can be made arbitrarily large by letting x go to infinity. However, with $w(t) = 1$ for all $t \in T$ we also have

$$\int_T w(t)y_x(t)\,\mathrm{d}\,t = 2x\int_0^{1/x}(1-xt)\,\mathrm{d}\,t = 2x\left(\frac{1}{2x}\right) = 1$$

so that each y_x is optimal in (P'_w) with strictly positive $w(t) > 0$ for all $t \in T$. $\diamondsuit$

In further comparison to the countable case in Example 2 and its resolution by Definition 2, Example 3 also illustrates that an analogous definition of proper efficiency using a criterion-dependent upper bound $\delta(t)$ for each $t \in T$ will not resolve the above problem at $t = 0$ and thus not be sufficient for the uncountable case in general. Before we address this observation in Section 4 in more detail, in the remaining section we first consider the other previous characterizations and show that only the augmented Tchebycheff norm method still provides a necessary condition for proper efficiency by Definition 3 if the number of criterion functions is uncountable. Likewise, we can disprove each of the respective necessary and sufficient conditions using the weighting and augmented Tchebycheff norm method using proofs by suitably constructed counterexamples.

3.1. Characterization Using Weighting Methods Unlike the standard result in Theorem 1 and similar to Example 2 for the countable case, Example 3 shows that the weighting method generally fails to generate only properly efficient solutions if the number of criteria is uncountable, even if the weighting function is chosen to be strictly positive. To the best of the author's knowledge this observation was first made by Winkler [66] who further conjectures that the reverse statement, however, that weighting methods can still generate all properly efficient solutions remains true under the usual convexity assumptions on criterion function and its underlying feasible set.

To prove this conjecture following the classical argument used in the proof of Theorem 1 which relies on the Hahn-Banach Separation Theorem to establish the existence of a supporting hyperplane whose normal vector provides the desired parameter w, one may apply an analogous separation theorem to obtain a linear functional $\ell \colon \mathcal{C}(T) \to \mathbb{R}$ from the corresponding dual space $\mathcal{C}(T)^*$. The Riesz representation theorem [55] gives a characterization of $\mathcal{C}(T)^*$ as the space of regular Borel (or Radon) measures [58] which are also used for the weighting method formulated in the original paper by Winkler [66] but for which an applicable separation theorem is not known in general. In fact, only if $T = [a,b] \subseteq \mathbb{R}$ is a subset of the real line, then Helly [29] has proven a special case of the Hahn-Banach Separation Theorem with the result that a function $\ell \in \mathcal{C}([a,b])^*$ if and only if there exists a function $\rho\colon [a,b] \to \mathbb{R}$ of bounded variation (BV) such that

$$\ell(y) = \int_a^b y(t)\,\mathrm{d}\,\rho(t) \text{ for all } y \in \mathcal{C}([a,b]).$$

While BV functions have only jump-discontinuities and thus are continuous except on at most a countable set with derivatives almost everywhere, they are typically not absolutely continuous and thus may not admit a Radon-Nikodym derivative $w(t) = \mathrm{d}\,\rho(t)/\,\mathrm{d}\,t$ that would allow us to write

$$\int_a^b y(t)\,\mathrm{d}\,\rho(t) = \int_a^b w(t)y(t)\,\mathrm{d}\,t \text{ for all } t \in T.$$

The fact that we cannot expect a necessary condition for properly efficient solutions in (VMP′) using problem (P'_w) even in the convex case is also demonstrated by the following example.

Example 4. Let $T = [0,1]$ and consider the set $Y \subseteq \mathcal{C}(T)$ of nondecreasing continuous functions $y\colon T \to \mathbb{R}$ with $y(0) + y(1) = 0$. This set Y is convex because convex combinations of nondecreasing continuous functions are still continuous and nondecreasing and clearly

$$\lambda(y_1(0) + y_1(1)) + (1-\lambda)(y_2(0) + y_2(1)) = 0 \text{ for all } y_1, y_2 \in Y \text{ and (any) } \lambda.$$

In particular, the function $\bar{y} = 0$ is in Y and properly efficient: it is clear that it is efficient and, for every $y \neq 0$ with $\delta = 1$ and $t_0 = 0$, that

$$y(t) - \bar{y}(t) = y(t) \leq y(1) = 0 - y(0) = \delta(\bar{y}(t_0) - y(t_0)).$$

Now let $w\colon T \to \mathbb{R}$ be any integrable function such that $\int_0^1 w(t)\,\mathrm{d}\,t = 1$ and $w(t) > 0$ for all $t \in T$. Let $m \in (0,1)$ such that $\int_0^m w(t)\,\mathrm{d}\,t = \int_m^1 w(t)\,\mathrm{d}\,t = 1/2$ and consider the nondecreasing continuous function $y(t) = \min\{2t/m - 1, 1\}$ which satisfies $y(0) + y(1) = -1 + 1 = 0$. It follows that

$$\int_0^1 w(t)y(t)\,\mathrm{d}\,t = \int_0^m w(t)\left(\frac{2t}{m} - 1\right)\mathrm{d}\,t + \int_m^1 w(t)\,\mathrm{d}\,t = \frac{2}{m}\int_0^m w(t)t\,\mathrm{d}\,t > 0$$

and thus $\bar{y} = 0$ is never an optimal solution in (P'_w) for any finitely integrable $w > 0$. $\diamond$

Note that in the above example we do not need to make any other assumption on w other than that it is finitely integrable and strictly positive. Hence, we cannot expect a satisfactory characterization of Geoffrion proper efficiency using the scalar maximum problem (P'_w) in general.

3.2. Characterizations Using Augmented Tchebycheff Norms The significant limitation of the weighting method to neither provide necessary nor sufficient conditions for uncountably many criteria is at least partially resolved by an augmented Tchebycheff norm method which we define analogously to Theorem 2.

Theorem 3. *Let $w\colon T \to \mathbb{R}$ be finitely integrable and $u\colon T \to R$ be a continuous utopia function such that $w(t) > 0$ and $u(t) > \sup\{y(t)\colon y \in Y\}$ for all $t \in T$. If a function $\bar{y} \in Y$ is properly efficient in* (VMP′)*, then there exist scalars $\alpha > 0$ and $v(t) > 0$ for all $t \in T$ such that $\bar{y}$ is optimal for the scalar maximum problem:*

$$\max \int_T w(t)y(t)\,\mathrm{d}\,t - \alpha \sup_{t \in T}\{v(t)(u(t) - y(t))\} \text{ subject to } y \in Y. \qquad (\mathrm{P}'_\infty)$$

Proof. Let $\bar{y}$ be properly efficient in (VMP′) with $\delta > 0$ and let

$$\mu = \max_{t\in T}\{u(t) - \bar{y}(t)\}$$

which is well-defined because u and $\bar{y}$ are continuous and because T is compact. Let

$$\omega = \int_T w(t)\,\mathrm{d}\,t < \infty$$

and set $\alpha \geq \delta\omega\mu$. Because u is also utopia there exists $\epsilon > 0$ such that $u(t) - y(t) \geq \epsilon$ for all $t \in T$ and $y \in Y$ and thus we can define

$$v(t) = (u(t) - \bar{y}(t))^{-1} > 0 \tag{2}$$

for all $t \in T$. It follows that $\mu \geq v(t)^{-1}$ and thus $\alpha \geq \delta\omega v(t)^{-1}$ for all $t \in T$. To show that $\bar{y}$ solves (P'_∞) now consider any other $y \in Y \setminus \{\bar{y}\}$. Because $\bar{y}$ is efficient there exists $t \in T$ such that $\bar{y}(t) > y(t)$ and thus

$$\sup_{t\in T}\{v(t)(u(t) - \bar{y}(t))\} = 1 < \sup_{t\in T}\{v(t)(u(t) - y(t))\}. \tag{3}$$

Also, because $\bar{y}$ is properly efficient there exists $t_0 \in T$ such that $y(t) - \bar{y}(t) \leq \delta(\bar{y}(t_0) - y(t_0))$ for all $t \in T$ and thus

$$\int_T w(t)(y(t) - \bar{y}(t))\,\mathrm{d}\,t \leq \left(\int_T w(t)\,\mathrm{d}\,t\right)\delta(\bar{y}(t_0) - y(t_0)) = \delta\omega(\bar{y}(t_0) - y(t_0)). \tag{4}$$

Now use (2) and (3) to observe that

$$\begin{aligned} \bar{y}(t_0) - y(t_0) &= (u(t_0) - y(t_0)) - (u(t_0) - \bar{y}(t_0)) \\ &\leq v(t_0)^{-1}\left(\sup_{t\in T}\{v(t)(u(t) - y(t))\} - \sup_{t\in T}\{v(t)(u(t) - \bar{y}(t))\}\right) \end{aligned} \tag{5}$$

where we used that $v(t_0)(u(t_0) - y(t_0)) \leq \sup\{v(t)(u(t) - y(t))$ and $v(t)(u(t) - \bar{y}(t)) = 1$ for all $t \in T$. Combining (4) and (5), it follows further that

$$\begin{aligned} \int_T w(t)(y(t) - \bar{y}(t))\,\mathrm{d}\,t &\leq \delta\omega(\bar{y}(t_0) - y(t_0)) \\ &\leq \delta\omega v(t_0)^{-1}\left(\sup_{t\in T}\{v(t)(u(t) - y(t))\} - \sup_{t\in T}\{v(t)(u(t) - \bar{y}(t))\}\right) \\ &\leq \alpha\left(\sup_{t\in T}\{v(t)(u(t) - y(t))\} - \sup_{t\in T}\{v(t)(u(t) - \bar{y}(t))\}\right) \end{aligned}$$

or equivalently, after rearranging terms, that

$$\int_T w(t)y(t)\,\mathrm{d}\,t - \alpha\sup_{t\in T}\{v(t)(u(t) - y(t))\} \leq \int_T w(t)\bar{y}(t)\,\mathrm{d}\,t - \alpha\sup_{t\in T}\{v(t)(u(t) - \bar{y}(t))\}.$$

Hence, because $y \in Y \setminus \{\bar{y}\}$ is chosen arbitrarily this shows that $\bar{y}$ is optimal in (P'_∞). $\square$

Unlike Theorem 2 which provides both necessary and sufficient conditions for a complete characterization of proper efficiency using augmented Tchebycheff norms in the countable case, however, and similar to the other characterizations using weighting methods in the uncountable case, the reverse statement of Theorem 3 turns out to be generally false in the uncountable case as well.

EXAMPLE 5. Let $T=[0,1]$ and consider the set $Y=\{y_x\in\mathcal{C}(T)\colon x\geq 1\}\cup\{0\}$ where the piecewise linear functions $y_x\colon T\to\mathbb{R}$ are depicted in Figure 1 and defined by

$$y_x(t)=\max\left\{1-\frac{(x+1)^2}{2x}t,-\frac{1}{x}\right\}\text{ for all }t\in T.$$

Let $\alpha>0$ be a scalar and define $u(t)=1+\epsilon$ for some $\epsilon>0$ and $v(t)=w(t)=1$ for all $t\in T$. It is not overly difficult to show that

$$\int_T w(t)y_x(t)\,\mathrm{d}\,t=\int_0^{2/(x+1)}\left(1-\frac{(x+1)^2}{2x}t\right)\mathrm{d}\,t-\int_{2/(x+1)}^1\left(\frac{1}{x}\right)\mathrm{d}\,t=0$$

for all $x\geq 1$ so that (P'_∞) can be simplified to the expression

$$\max_{y\in Y}\inf_{t\in T}\{y(t)\}$$

for which the function $y_0=0$ is the unique optimal solution. However, whereas all $y_x\in Y\setminus\{0\}$ are properly efficient, the function $y_0=0$ is only efficient: for each $x\geq 1$, we have $y_x(0)-y_0(0)=1$ whereas $y_0(t)-y_x(t)\leq 1/x$ can be kept positive but made arbitrarily small by letting x go to infinity. This yields the unbounded tradeoff ratio

$$(y_x(0)-y_0(0))\,/\,(y_0(t)-y_x(t))\geq x\text{ for all }t\in T\text{ such that }y_0(t)>y_x(t)$$

and shows that optimal solutions in (P'_∞) need not be properly efficient in (VMP′) in general. ◇

This example and Figure 1 might suggest that improperness in the uncountable case may be a consequence of the existence of a sequence of functions that lacks uniform convergence so that a stronger result may continue to hold under more restrictive assumptions on Y or T, or u, v, w or α in particular. Based on Example 5 which only uses uniform weights and an arbitrary utopia point, however, it follows that an analogous result to Theorem 2 can fail in the uncountable case even for these quite general and rather natural choices. Moreover, the example continues to fail even if $u(t)$ is non-utopian with the same analysis as before as long as $u(t)$ is chosen to be constant.

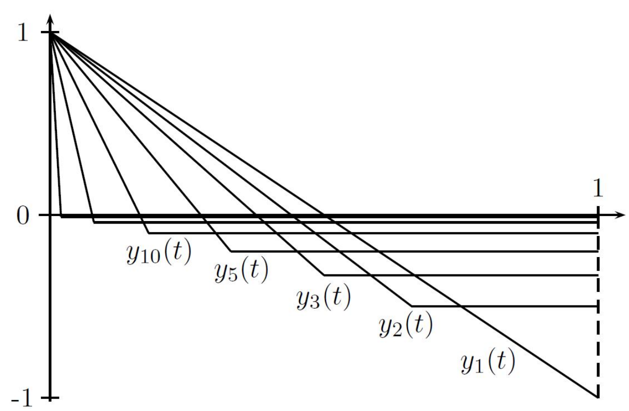

FIGURE 1. Supporting illustration of Example 5: All piecewise linear functions $y_x(t)=\max\left\{1-\frac{(x+1)^2}{2x}t,-\frac{1}{x}\right\}$ are properly efficient in (VMP′) whereas the function $y_0=0$ is only efficient but the unique optimal solution in (P'_∞).

4. Proper Efficiency Almost Everywhere In view of our results so far, Examples 3, 4 and 5 establish that the standard characterizations of proper efficiency using both weighting and augmented Tchebycheff norm methods do not continue to hold anymore for uncountably many criteria and that only the necessary condition using the latter in Theorem 3 remains valid regardless of whether the number of criteria is finite or countably or uncountably infinite. However, in addition to our discussion at the beginning of Section 3.1 these examples suggest that those unbounded tradeoffs that may still be present in weighting or Tchebycheff solutions can only be of a certain type and occur with respect to a subset $U \subseteq T$ of criteria with (Lebesgue) measure $\mathcal{L}(U) = 0$. The following statement formulates this idea analogously to the previous Definitions 1, 2 and 3.

Definition 4. A function $\bar{y} \in Y$ is said to be *properly efficient almost everywhere* in (VMP$'$) if it is efficient and if for every $\epsilon > 0$ there exists a scalar $\delta > 0$ such that, for each $t \in T \setminus U$ and some $U \subseteq T$ with $\mathcal{L}(U) < \epsilon$, we have

$$\frac{y(t) - \bar{y}(t)}{\bar{y}(t_0) - y(t_0)} \leq \delta$$

for some $t_0 \in T$ such that $y(t_0) < \bar{y}(t_0)$ whenever $y \in Y$ and $y(t) > \bar{y}(t)$.

Similar to our discussion following Definition 3, an equivalent and arguably more natural statement of Definition 4 under the assumptions that each $y \in Y$ is continuous and that T is compact is that for every $\epsilon > 0$ there exists $\delta > 0$ such that for each $y \in Y$ and $t_0 = \arg\max_{t \in T}\{\bar{y}(t) - y(t)\}$ we have

$$y(t) - \bar{y}(t) \leq \delta(\bar{y}(t_0) - y(t_0)) \text{ for all } t \in T \setminus U$$

for some subset $U \subseteq T$ of Lebesgue measure $\mathcal{L}(U) < \epsilon$ which may generally depend on both ϵ and the comparison function y. In particular, if $\bar{y}$ is properly efficient so that all tradeoffs are bounded, then $U = \emptyset$ satisfies Definition 4 and the following result is immediate.

Proposition 4. *Proper efficiency by Definition 3 implies proper efficiency almost everywhere by Definition 4.*

Moreover, based on Example 4 which shows that the set of weighting solutions do not contain all properly efficient solutions even in the case of convexity it is also clear that solutions that are properly efficient almost everywhere need not be solutions to weighting methods in general. However, using Definition 4 we now are able to provide necessary conditions and prove that all weighting and Tchebycheff solutions are still properly efficient almost everywhere which thereby provides the satisfactory characterization of their remaining unbounded tradeoffs to be at most a Lebesgue null set. In particular in view of our previous results, these proofs are somewhat relieving at least from a decision-theoretic point of view by confirming that the typical solutions to stochastic programming approaches with or without additional regret criterion may not prevent the existence of unbounded tradeoffs in general but can reduce their likelihood to be of probability zero.

A schematic overview of all relationships between solutions that are properly efficient or properly efficient almost everywhere in (VMP$'$) and optimal weighting or Tchebycheff solutions in (P$'_w$) or (P$'_\infty$) is given in Figure 2. For completeness, this already includes the two results and the counterexample for the case of continuous real-valued functions over compact supports from Section 4.1 whose extensions to other and generally unbounded supports is briefly discussed in Section 4.2.

4.1. Characterization for Compact Criterion Sets To establish the still missing results in Figure 2, we begin to show that all optimal solutions to strictly positive weighting methods are still guaranteed to be properly efficient almost everywhere.

Theorem 4. *Let $w\colon T \to \mathbb{R}$ be finitely integrable such that* $\inf\{w(t)\colon t \in T\} > 0$. *If $\bar{y}$ is optimal in* (P$'_w$), *then $\bar{y}$ is properly efficient almost everywhere in* (VMP$'$).

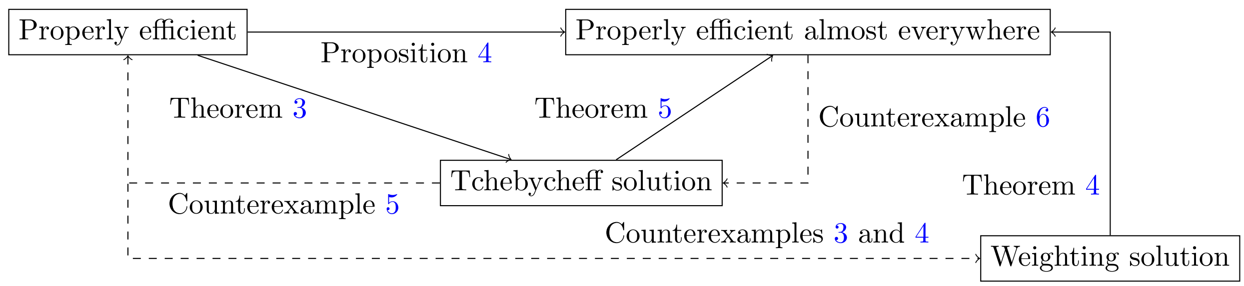

FIGURE 2. Schematic overview of the relationships between proper efficiency and proper efficiency almost everywhere in (VMP′) and solutions to weighting and augmented Tchebycheff norm methods in (P'_w) and (P'_∞) respectively.

Proof. Let $\bar{y}$ be optimal in (P'_w) and without loss of generality assume that $\int_T w(t)\,\mathrm{d}\,t = 1$. A standard argument by contradiction suffices to show that $\bar{y}$ is efficient in (VMP′), so let

$$\mu = \inf_{t\in T}\{w(t)\} > 0$$

and only suppose to the contrary that $\bar{y}$ is not properly efficient almost everywhere. It follows that there exists $\epsilon > 0$ such that

$$y(t) - \bar{y}(t) > \delta(\bar{y}(t_0) - y(t_0)) \text{ for all } t \in U$$

for some $y \in Y$ and $U \subseteq T$ with $\mathcal{L}(U) \geq \epsilon$ whenever $\delta > 0$ and $t_0 \in T$. Specifically, choose $\delta = 1/(\mu\epsilon)$ and first integrate this inequality over U to find

$$\int_U w(t)\,(y(t) - \bar{y}(t))\,\mathrm{d}\,t > \mu\epsilon\delta\,(\bar{y}(t_0) - y(t_0)) = \bar{y}(t_0) - y(t_0).$$

Next, multiply these inequalities for each $t_0 \in T \setminus U$ with $w(t_0) > 0$ and integrate over $V = T \setminus U$, immediately factoring out the constant on the left-hand side:

$$\left(\int_V w(t_0)\,\mathrm{d}\,t_0\right)\left(\int_U w(t)\,(y(t) - \bar{y}(t))\,\mathrm{d}\,t\right) > \int_V w(t_0)(\bar{y}(t_0) - y(t_0))\,\mathrm{d}\,t_0$$

where $\int_V w(t_0)\,\mathrm{d}\,t_0 = 1 - \int_U w(t_0)\,\mathrm{d}\,t_0 \leq 1 - \mu\epsilon < 1$. Hence, we have

$$\int_U w(t)\,(y(t) - \bar{y}(t))\,\mathrm{d}\,t > \left(\int_V w(t_0)\,\mathrm{d}\,t_0\right)\left(\int_U \mu\,(y(t) - \bar{y}(t))\,\mathrm{d}\,t\right) > \int_V w(t_0)(\bar{y}(t_0)) - y(t_0))\,\mathrm{d}\,t_0$$

and thus, using $T = U \cup V$ and rearranging terms, that

$$\int_T w(t)y(t)\,\mathrm{d}\,t > \int_T w(t)\bar{y}(t)\,\mathrm{d}\,t.$$

This contradicts that $\bar{y}$ is optimal in (P'_w) and thus $\bar{y}$ is also properly efficient almost everywhere in (VMP′). □

Without the need of much further discussion we continue to show the analogous result for optimal solutions with respect to the augmented Tchebycheff norm method. Fully analogous to the situation in the (finite or infinite) countable case, however, we only highlight that the following statement and its proof again require no specific assumption regarding the (utopia) function u that for the sufficient condition does not need to be a utopia function in general.

THEOREM 5. *Let $w\colon T \to \mathbb{R}$ be finitely integrable and $u(t)\colon T \to \mathbb{R}$ and $v(t)\colon T \to \mathbb{R}$ be arbitrary such that* $\inf\{w(t)\colon t \in T\} > 0$, $\sup\{w(t)\colon t \in T\} < \infty$ *and* $\sup\{v(t)\colon t \in T\} < \infty$. *If $\bar{y}$ is optimal in (P'_∞), then $\bar{y}$ is properly efficient almost everywhere in (VMP′).*

Proof. Let $\bar{y}$ be optimal in (P'_∞), without loss of generality assume $\int_T w(t)\,\mathrm{d}\,t = 1$ and denote $\mu = \inf\{w(t)\colon t \in T\} > 0$, $\omega = \sup\{w(t)\colon t \in T\} < \infty$ and $\nu = \sup\{v(t)\colon t \in T\} < \infty$. It follows that

$$\begin{aligned}\int_T w(t)\,(y(t) - \bar{y}(t))\,\mathrm{d}\,t &\leq \alpha\left(\sup_{t\in T}\{v(t)(u(t) - y(t))\} - \sup_{t\in T}\{v(t)(u(t) - \bar{y}(t))\}\right) \\ &\leq \alpha\left(\sup_{t\in T}\{v(t)(\bar{y}(t) - y(t))\}\right) \leq \alpha\nu(\bar{y}(t_0) - y(t_0))\end{aligned} \tag{6}$$

for all $y \in Y$ and $t_0 = t_0(y, \bar{y}) = \arg\max_{t\in T}\{\bar{y}(t) - y(t)\}$, and a standard argument by contradiction suffices to show that $\bar{y}$ is again efficient in (VMP′). We also use a contradiction to show that $\bar{y}$ is properly efficient almost everywhere, for if it is not there exists $\epsilon > 0$ such that

$$y(t) - \bar{y}(t) > \delta(\bar{y}(t_0) - y(t_0)) \text{ for all } t \in U \tag{7}$$

for some $y \in Y$ and $U \subseteq T$ with $\mathcal{L}(U) \geq \epsilon$ whenever $\delta > 0$ and $t_0 \in T$. Specifically, choose

$$\delta \geq \frac{\alpha\nu + (1-\epsilon)\omega}{\epsilon\mu}$$

and let $t_0 = \arg\max_{t\in T}\{\bar{y}(t) - y(t)\}$. Partitioning $T = U \cup (T \setminus U)$ and using (7) it follows that

$$\begin{aligned}\int_T w(t)\,(y(t) - \bar{y}(t))\,\mathrm{d}\,t &= \int_U w(t)\,(y(t) - \bar{y}(t))\,\mathrm{d}\,t - \int_{T\setminus U} w(t)\,(\bar{y}(t) - y(t))\,\mathrm{d}\,t \\ &> \left(\delta\int_U w(t)\,\mathrm{d}\,t - \int_{T\setminus U} w(t)\,\mathrm{d}\,t\right)(\bar{y}(t_0) - y(t_0)) \\ &\geq (\delta\epsilon\mu - (1-\epsilon)\omega)(\bar{y}(t_0) - y(t_0)) \geq \alpha\nu(\bar{y}(t_0) - y(t_0))\end{aligned}$$

in contradiction to (6). Hence, $\bar{y}$ is also properly efficient almost everywhere in (VMP′). □

Finally, unlike the result in Theorem 3 but similar to the observation already made for the weighting method after Proposition 4 we can also show that the reverse statement of Theorem 5 is generally false. Our counterexample uses a set of functions that are similar to those that we defined in Example 5 and whose general behavior was depicted for ease of discussion in Figure 1.

Example 6. Similar to the situation used in Example 5 let $T = [0, 1]$ and consider the set $Y = \{y_x \in \mathcal{C}(T)\colon x \geq 1\} \cup \{0\}$ where the piecewise linear functions $y_x\colon T \to \mathbb{R}$ are defined by

$$y_x(t) = \max\left\{1 - \frac{(x+1)^2}{2x}t, -\frac{1}{x^2}\right\}.$$

In comparison to Example 5 and Figure 1, however, here note that the second piece of this function definition is different and for all $x > 1$ starts the flat stretch earlier so that the new (negative) "areas under the curve" are generally smaller, or equivalently, that the signed difference in area remains strictly positive. In particular with $w(t) = 1$ for all $t \in T$ we can show that

$$\begin{aligned}\int_T w(t) y_x(t)\,\mathrm{d}\,t &= \int_0^{2(x^2+1)/x(x+1)^2}\left(1 - \frac{(x+1)^2}{2x}t\right)\mathrm{d}\,t - \int_{2(x^2+1)/x(x+1)^2}^1\left(\frac{1}{x^2}\right)\mathrm{d}\,t \\ &= \frac{(x-1)^2(x^2+x+1)}{x^3(x+1)^2} = \frac{x^4 - x^3 - x^2 + 1}{x^5 + 2x^4 + x^3} \geq \frac{1}{2x}\end{aligned}$$

for all $x \geq 5$. Otherwise as before all $y \in Y \setminus \{0\}$ are still properly efficient whereas the zero function $y_0 = 0$ is only properly efficient almost everywhere: for any $0 < \epsilon < 1/2$ and $U = [0, \epsilon/2)$ we have $L(U) < \epsilon$ so that we can exclude a semi-neighborhood around 0 to prevent the unbounded tradeoff observed in Example 5 when x goes to infinity. Now let $u(t) = \mu > 1$ be a constant and $v(t) > 0$ be arbitrary for all $t \in T$ such that $\nu = \sup\{v(t) \colon t \in T\} < \infty$. It follows that

$$\sup_{t\in T}\{v(t)\,(u(t) - y_x(t))\} - \sup_{t\in T}\{v(t)\,(u(t) - y_0(t))\} \le \nu \sup_{t\in T}\{y_0(t) - y_x(t)\} = \nu/x^2.$$

Hence, for any $\alpha > 0$ we can choose $x > \max\{5, 2\alpha\nu\}$ such that

$$\begin{aligned}&\left(\int_T w(t)y_x(t)\,\mathrm{d}\,t - \alpha\sup_{t\in T}\{v(t)(u(t) - y_x(t))\}\right)\\ -&\left(\int_T w(t)y_0(t)\,\mathrm{d}\,t - \alpha\sup_{t\in T}\{v(t)(u(t) - y_0(t))\}\right) \ge \frac{1}{2x} - \frac{\alpha\nu}{x^2} > \frac{1}{2x} - \frac{1}{2x} = 0\end{aligned}$$

and thus y_0 is never optimal in (P'_∞) for any choice of $\alpha > 0$ and finitely bounded $v > 0$. $\Diamond$

4.2. Extension to General Criterion Sets Following the original problem formulation (VMP) by Geoffrion [26] in Section 2 and its extension (VMP′) by Winkler [66] in Sections 3 and 4, all of our results so far consider either the countable case or the uncountable case in which all criterion functions are real-valued continuous and defined over some compact support. Whereas both the definition of efficiency and proper efficiency in Definitions 3 and 4 are independent of such assumptions and similarly apply to an unbounded or more general criterion set that is not compact, however, some of our former proofs have utilized the compactness to replace certain suprema with their (finite) maxima. To extend these results to the general case we therefore have two possibilities and may either make additional assumptions on the finiteness of these suprema or generalize our definition of proper efficiency almost everywhere in the following suitable way.

Definition 5. Given a set Y of real-valued functions defined on a subset S of a separable Banach space, a function $\bar{y} \in Y$ is said to be *properly efficient almost everywhere* if it is efficient and if, for every $\epsilon > 0$ and any compact subset $T \subseteq S$ there exists a scalar $\delta > 0$ such that, for each $t \in T \setminus U$ and some $U \subseteq T$ with $\mathcal{L}(U) < \epsilon$, we have

$$\frac{y(t) - \bar{y}(t)}{\bar{y}(t_0) - y(t_0)} \le \delta$$

for some $t_0 \in T$ such that $y(t_0) < \bar{y}(t_0)$ whenever $y \in Y$ and $y(t) > \bar{y}(t)$.

Using this reduction from a general criterion set S to compact subsets $T \subseteq S$ for the particular situation of a subset $Y \subseteq \mathcal{C}(S)$ of continuous real-valued functions, it directly follows that all of our formerly proven results remain true also when these functions are defined over a generally infinite and possibly unbounded support S. The repeated statements and proofs of these results can therefore be omitted.

5. Conclusion The solution and analysis of optimization and decision-making problems in the presence of uncertainty is both conceptually and practically challenging. Due to the lack of a unique solution concept it lends itself to a variety of different approaches that most prominently include stochastic programming or robust optimization and, more recently, the theory and methodology of multiple criteria optimization (MCO) and decision analysis [8, 32, 33, 34, 42, 52, 60]. Among this work, especially the recent paper by Iancu and Trichakis [33] highlights that solution concepts from MCO can make a significant difference and offer additional insights and improvements over the classical stochastic or robust solution paradigm. However, one of several substantial concerns brought forward earlier by Hites *et al.* [31] remains regarding the frequent limitation of methods from MCO to a sufficiently small and typically finite number of criteria which may prevent their use for continuous and more general uncertainty sets that require to consider an infinite number of scenarios or criterion functions in its corresponding deterministic multicriteria formulation.

Thus motivated, in this present paper we build on the mathematical equivalence between linear scalarizations in MCO and expected value functions in stochastic programming to provide a first comprehensive analysis of solution tradeoffs based on the notion of proper efficiency and its relationship to marginal rates of substitution for both finitely and infinitely many criteria. The respective implications and impact of this analogy and our new results are at least four-fold.

1. We extend one of the most natural solution concepts from MCO, decision-making and economics to stochastic programming, namely the concept of proper efficiency with the goal to exclude anomalous solutions that still permit unbounded tradeoffs from consideration for a final decision.

2. We emphasize the importance to combine standard stochastic programming with an additional contribution of worst-case regret or some achievement level to prevent the generation of solutions that are optimal in a stochastic sense but for which unbounded tradeoffs may still be possible.

3. We report and demonstrate a few other shortcomings of stochastic expected-value optimization including the somewhat arbitrary exclusion of certain efficient but unsupported solutions using weighting methods if the problem is not convex and the more subtle consequence of remaining unbounded tradeoffs in a Lebesgue null set if the number of random realizations is uncountable.

4. We address the former critique of MCO methods to be often limited to finitely many criteria: although it is conceptually straightforward to extend many methods to an either countably infinite or uncountable number of criteria, it is correct that such extensions may have additional theoretical and practical implications that require modifications or further work to maintain standard results or existing characterization of solutions that are Pareto optimal or properly efficient in general.

Specifically, whereas the generalization of existing characterizations of proper efficiency using weighting methods and augmented Tchebycheff norms from finitely to countably many criteria is still possible after a slight modification to the original definition which we briefly reviewed in Section 2 based on our earlier work [20], almost all standard results turn out to be false for the more general uncountable case. This bears the consequence that a complete characterization of properly efficient solutions is possible only for stochastic programs with discrete random variables but generally not for those involving continuous uncertainty sets. However, based on our analysis in Sections 3 and 4 whose results and findings are collected and summarized in Figure 2 we can now conclude that solutions to these scalar problems are still properly efficient almost everywhere and that especially the augmented Tchebycheff norm method is still capable to compute the full set of properly efficient solutions in principle. These results are somewhat relieving at least from a decision-theoretic point of view by confirming that the typical solutions to stochastic programming approaches may not prevent the existence of unbounded tradeoffs in general but can reduce their likelihood to be of probability zero. Moreover, it is interesting to observe that this method can be interpreted to compromise between stochastic and robust optimization by including one contribution from an expected value and another from a suitably weighted worst case regret or achievement criterion. This implies that stochastic and robust optimization can be seen as two complementary extremes of a unified multicriteria framework that uses a more general compromise programming approach for optimization and decision-making in the presence of uncertainty [19].

This last observation especially opens a plethora of new research questions. On the one hand and despite the existing debate whether MCO methods are suitable to address such problems at all, current literature and our own results suggest that their extension provides stronger results, additional insight and better interpretability than approaches that optimize either expected values or the worst case alone. Further research may therefore continue to investigate whether similar approaches or reformulations may be used to handle recourse or uncertainties in the constraints, for example by converting chance-constrained problems into corresponding goal programs or similar other approaches. On the other hand, the potential impact of MCO onto stochastic and robust optimization promotes a more careful study of its methods for more general settings which may also include infinite dimensions to be applicable for both discrete and continuous random variables over finite or generally infinite supports. Hence, in summary, this new research has high potential to make an equally novel contribution to multiple criteria and stochastic or robust optimization and thereby provide both theoretical insights and new methodological advantages using a consistent and unified further mathematical analysis of these two important areas of operations research.

Acknowledgments The author would like to thank the area editor, an associate editor and three anonymous referees for their suggestions and helpful comments on an earlier manuscripts of this paper which have led to its improved presentation and content.

References

[1] Adida, E., G. Perakis. 2010. Dynamic pricing and inventory control: robust vs. stochastic uncertainty models—a computational study. *Annals of Operations Research* **181** 125–157.

[2] Arrow, K. J., E. W. Barankin, D. Blackwell. 1953. Admissible points of convex sets. *Contributions to the Theory of Games, vol. 2*. Annals of Mathematics Studies, no. 28, Princeton University Press, Princeton, NJ, 87–91.

[3] Beale, E. M. L. 1955. On minimizing a convex function subject to linear inequalities. *Journal of the Royal Statistical Society Series B* **17** 173–184.

[4] Ben-Tal, A., L. El Ghaoui, A. Nemirovski. 2009. *Robust optimization*. Princeton Series in Applied Mathematics, Princeton University Press, Princeton, NJ.

[5] Benson, H. P. 1979. An improved definition of proper efficiency for vector maximization with respect to cones. *Journal of Mathematical Analysis and Applications* **71**(1) 232–241.

[6] Bertsimas, D., D. Brown, C. Caramanis. 2011. Theory and applications of robust optimization. *SIAM Review* **53** 464–501.

[7] Birge, J. R., F. Louveaux. 1997. *Introduction to stochastic programming*. Springer Series in Operations Research, Springer-Verlag, New York, NY.

[8] Bonnel, H., J. Collonge. 2014. Stochastic optimization over a Pareto set associated with a stochastic multi-objective optimization problem. *Journal of Optimization Theory and Applications* **162**(2) 405–427.

[9] Borwein, J. M. 1977. Proper efficient points for maximizations with respect to cones. *SIAM Journal on Control and Optimization* **15**(1) 57–63.

[10] Borwein, J. M., D. Zhuang. 1993. Super efficiency in vector optimization. *Transactions of the American Mathematical Society* **338**(1) 105–122.

[11] Bowman, V. J. 1976. On the relationship of the Tchebycheff norm and the efficient frontier of multiple-criteria objectives. H. Thirez, S. Zionts, eds., *Multiple Criteria Decision Making*, *Lecture Notes in Economics and Mathematical Systems*, vol. 130. Springer, 76–86.

[12] Charnes, A., W. W. Cooper. 1959. Chance-constrained programming. *Management Science* **6** 73–79.

[13] Choo, E. U., D. R. Atkins. 1983. Proper efficiency in nonconvex multicriteria programming. *Mathematics of Operations Research* **8**(3) 467–470.

[14] Daniilidis, A. 2000. Arrow-Barankin-Blackwell theorems and related results in cone duality: a survey. *Optimization (Namur, 1998)*, *Lecture Notes in Economic and Mathematical Systems*, vol. 481. Springer, Berlin, 119–131.

[15] Dantzig, G. B. 1955. Linear programming under uncertainty. *Management Science* **1** 197–206.

[16] Di Silvestre, M. L., G. Graditi, M. G. Ippolito, E. R. Sanseverino, G. Zizzo. 2011. Robust multi-objective optimal dispatch of distributed energy resources in micro-grids. *PowerTech, 2011 IEEE Trondheim*. 1–5.

[17] Doolittle, E. K., Hervé L. M. Kerivin, M. M. Wiecek. 2012. A robust multiobjective optimization problem with application to internet routing. Tech. rep., Department of Mathematical Sciences, Clemson University, SC, USA.

[18] Engau, A. 2011. Nonlinear multiobjective programming. J. Cochran, ed., *Wiley Encyclopedia of Operations Research and Management Science*, vol. 5. Wiley-Blackwell, Hoboken, NJ, 3582–3612.

[19] Engau, A. 2012/13. Multicriteria framework for robust-stochastic formulations of optimization under uncertainty. Presentations delivered at the *10th International Conference on Multiple Objective Programming & Goal Programming* held jointly with the *54th Annual Conference of the Canadian Operational Research Society* in Niagara Falls, ON, Canada, June 11-13, 2012, and the *INFORMS Annual Meeting* in Minneapolis, MN, USA, October 6-9, 2013.

[20] Engau, A. 2015. Definition and characterization of Geoffrion proper efficiency for real vector optimization with infinitely many criteria. *Journal of Optimization Theory and Applications* **165**(2) 439–457.

[21] Engau, A. 2016. A comparative study of new multicriteria models for robust data classification in supervised learning. Tech. rep., University of Colorado Denver / Dalhousie University.

[22] Engau, A., C. Moffatt, W. Dyk. 2015. Multicriteria modeling and tradeoff analysis for oil load dispatch and hauling operations at Noble Energy. *Optimization and Engineering* **16**(1) 73–101.

[23] Fischetti, M., M. Monaci. 2009. Light robustness. *Robust and Online Large-Scale Optimization*, *Lecture Notes in Computer Science*, vol. 5868. Springer, Berlin, 61–84.

[24] Fishburn, P. C. 1964. *Decision and Value Theory*. Publications in Operations Research, No. 10, John Wiley & Sons, New York, NY.

[25] Gabrel, V., C. Murat, A. Thiele. 2014. Recent advances in robust optimization: an overview. *European Journal of Operational Research* **235**(3) 471–483.

[26] Geoffrion, A. M. 1968. Proper efficiency and the theory of vector maximization. *Journal of Mathematical Analysis and Applications* **22** 618–630.

[27] Guerraggio, A., E. Molho, A. Zaffaroni. 1994. On the notion of proper efficiency in vector optimization. *Journal of Optimization Theory and Applications* **82**(1) 1–21.

[28] Hartley, R. 1978. On cone-efficiency, cone-convexity and cone-compactness. *SIAM Journal on Applied Mathematics* **34** 211–222.

[29] Helly, E. 1912. Über lineare Funktionaloperationen. *Sitzungsberichte der Kaiserlichen Akademie der Wissenschaften in Wien. Mathematisch-Naturwissenschaftliche Klasse* **121** 265–297.

[30] Henig, M. I. 1982. Proper efficiency with respect to cones. *Journal of Optimization Theory and Applications* **36**(3) 387–407.

[31] Hites, R., Y. De Smet, N. Risse, M. Salazar-Neumann, P. Vincke. 2006. About the applicability of MCDA to some robustness problems. *European Journal of Operational Research* **174** 322–332.

[32] Hu, J., S. Mehrotra. 2012. Robust and stochastically weighted multiobjective optimization models and reformulations. *Operations Research* **60**(4) 936–953.

[33] Iancu, D. A., N. Trichakis. 2014. Pareto efficiency in robust optimization. *Management Science* **60**(1) 130–147.

[34] Ide, J., E. Köbis, D. Kuroiwa, A. Schöbel, C. Tammer. 2014. The relationship between multi-objective robustness concepts and set-valued optimization. *Fixed Point Theory and Applications* **83** 20p.

[35] Infanger, G., ed. 2011. *Stochastic programming*. International Series in Operations Research & Management Science, 150, Springer, New York, NY.

[36] Isermann, H. 1974. Proper efficiency and the linear vector maximum problem. *Operations Research* **22** 189–191.

[37] Jahn, J. 2011. *Vector optimization*. Second ed. Springer-Verlag, Berlin.

[38] Kaliszewski, I. 1987. A modified weighted Tchebycheff metric for multiple objective programming. *Computers & Operations Research* **14**(4) 315–323.

[39] Kaliszewski, I. 1994. *Quantitative Pareto analysis by cone separation technique*. Kluwer Academic Publishers, Boston, MA.

[40] Keeney, R. L., H. Raiffa. 1976. *Decisions with multiple objectives: preferences and value tradeoffs*. Wiley Series in Probability and Mathematical Statistics, John Wiley & Sons, New York, NY.

[41] Khan, A. A., C. Tammer, C. and Zălinescu. 2015. *Set-valued optimization*. Vector Optimization, Springer, Heidelberg.

[42] Klamroth, K., E. Köbis, A. Schöbel, C. Tammer. 2013. A unified approach for different concepts of robustness and stochastic programming via non-linear scalarizing functionals. *Optimization* **62**(5) 649–671.

[43] Kouvelis, P., S. Sayın. 2005. The multiobjective discrete optimization problem: a weighted min-max two-stage optimization approach and a bicriteria algorithm. *Management Science* **51** 1572–1581.

[44] Kouvelis, P., S. Sayın. 2006. Algorithm robust for the bicriteria discrete optimization problem: heuristic variations and computational evidence. *Annals of Operations Research* **147** 71–85.

[45] Kouvelis, P., G. Yu. 1997. *Robust discrete optimization and its applications*, *Nonconvex Optimization and its Applications*, vol. 14. Kluwer Academic Publishers, Dordrecht.

[46] Kuhn, H. W., A. W. Tucker. 1951. Nonlinear programming. *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability, 1950*. University of California Press, Berkeley and Los Angeles, 481–492.

[47] Lodwick, W. A., J. Kacprzyk, eds. 2010. *Fuzzy optimization. Recent advances and applications*, *Studies in Fuzziness and Soft Computing*, vol. 254. Springer-Verlag, Berlin.

[48] Mäkelä, M. M., Y. Nikulin, J. Mezei. 2012. A note on extended characterization of generalized trade-off directions in multiobjective optimization. *Journal of Convex Analysis* **19**(1) 91–111.

[49] Miettinen, K., M. M. Mäkelä. 2001. On cone characterizations of weak, proper and Pareto optimality in multiobjective optimization. *Mathematical Methods of Operations Research* **53**(2) 233–245.

[50] Miettinen, K., M. M. Mäkelä. 2002. On generalized trade-off directions in nonconvex multiobjective optimization. *Mathematical Programming Ser. A* **92**(1) 141–151.

[51] Miettinen, K., M. M. Mäkelä. 2003. Characterizing generalized trade-off directions. *Mathematical Methods of Operations Research* **57**(1) 89–100.

[52] Ogryczak, W. 2002. Multiple criteria optimization and decisions under risk. *Control and Cybernetics* **31** 975–1003.

[53] Rachev, S. T., W. Römisch. 2002. Quantitative stability in stochastic programming: the method of probability metrics. *Mathematics of Operations Research* **27**(4) 792–818.

[54] Reehuis, E. 2010. Multiobjective robust optimization of water distribution networks. Master's thesis, Leiden Institute of Advanced Computer Science (LIACS), Universiteit Leiden, The Netherlands.

[55] Riesz, F. 1909. Sur les opérations fonctionnelles linéaires. *Comptes Rendus de l'Acadḿie des Sciences Paris* **149** 974–977.

[56] Römisch, W., R. Schultz. 1991. Distribution sensitivity in stochastic programming. *Mathematical Programming Ser. A* **50**(2) 197–226.

[57] Roy, B. 2010. Robustness in operational research and decision aiding: A multi-faceted issue. *European Journal of Operational Research* **200** 629–638.

[58] Rudin, W. 1987. *Real and complex analysis*. Third ed. International Series in Pure and Applied Mathematics, McGraw-Hill Book Co., New York, NY.

[59] Sawaragi, Y., H. Nakayama, T. Tanino. 1985. *Theory of multiobjective optimization*, *Mathematics in Science and Engineering*, vol. 176. Academic Press Inc., Orlando, FL.

[60] Schöbel, A. 2014. Generalized light robustness and the trade-off between robustness and nominal quality. *Mathematical Methods of Operations Research* **80**(2) 161–191.

[61] Shapiro, A., D. Dentcheva, A. Ruszczyński. 2009. *Lectures on stochastic programming. Modeling and theory*, *MPS/SIAM Series on Optimization*, vol. 9. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA.

[62] Soyster, A. L. 1973. Convex programming with set-inclusive constraints and applications to inexact linear programming. *Operations Research* **21** 1154–1157.

[63] Steuer, R. E. 1986. *Multiple Criteria Optimization. Theory, Computation, and Application*. Wiley Series in Probability and Mathematical Statistics, John Wiley & Sons, Inc., New York, NY.

[64] Steuer, R. E., E. U. Choo. 1983. An interactive weighted Tchebycheff procedure for multiple objective programming. *Mathematical Programming* **26**(3) 326–344.

[65] Truong, X. D. H. 2001. Existence and density results for proper efficiency in cone compact sets. *Journal of Optimization Theory and Applications* **111**(1) 173–194.

[66] Winkler, K. 2004. Geoffrion proper efficiency in an infinite dimensional space. *Optimization* **53**(4) 355–368.