

Subdifferentials of Convex Operators Valued in the Space of Integrable Functions with Applications to Risk-Averse Optimization

Darinka Dentcheva

*Stevens Institute of Technology, Dept. of Mathematical Sciences, Hoboken, U.S.A.
darinka.dentcheva@stevens.edu*

Andrzej Ruszczyński

*Dept. of Management Science and Information Systems, Rutgers University, Piscataway, U.S.A.
rusz@rutgers.edu*

Received: January 20, 2025

Accepted: December 9, 2025

We study differentiability properties of convex operators defined on a Banach space with values in an $\mathcal{L}_p$ space and of their compositions with monotonic convex functionals on this space. We develop new tools for operators enjoying an additional feature known as the local property. The new approach and results go beyond the classical theory of normal integrands and lattice-valued operators. We further describe the subdifferentials of compositions of such operators with convex monotonic functionals. The new results are applied to obtain novel optimality conditions in the subdifferential form for a broad class of risk-averse stochastic optimization problems with risk functionals as objectives, with partial information, and with stochastic dominance constraints. While our analysis is motivated by the theory and methods of risk-averse optimization, it addresses problems of a more general structure and has a potential for further applications.

Keywords: Risk operators, composite functionals, stochastic dominance constraints.

2020 Mathematics Subject Classification: 46G05, 49J52, 49K27.

Introduction

Our study is motivated by stochastic optimization models with non-linear operators expressing risk aversion. The area of optimization integrating risk models has developed very rapidly in the recent years due to its mathematical challenges and practical relevance. Among the theoretical questions posed by risk models, a key question pertains to the subdifferentiability of operators valued in $\mathcal{L}_p$ spaces and their compositions with non-linear functionals. A lot of work containing fundamental results on subdifferentiation of integral functionals is available in the area of convex analysis. We refer to the pioneering results in [41, 42, 48, 49, 50] and to the thorough exposition in the monographs [7, 34]. In this paper, we undertake the challenge to address subdifferentiability of convex operators $F : \mathfrak{X} \rightarrow \mathcal{L}_p(\Omega, \mathcal{F}, P)$ defined on a Banach space $\mathfrak{X}$ and composed further with nonlinear functionals. These structures arise in the modern directions of stochastic optimization, where further technical difficulties are posed by the non-linearity in probability.

This work was supported by the Air Force Office of Scientific Research award FA9550-24-1-0284.

Our goal is to establish subdifferentiability properties of such operators, describe their subdifferentials, and obtain optimality conditions in subdifferential form for stochastic optimization problems incorporating risk models as objectives and constraints.

More precisely, we consider the following general optimization problem:

$$\min \varrho(F(x)), \quad \rho_t(G(x)) \leq 0, \quad \forall t \in J, \quad x \in \mathcal{Y}. \quad (1)$$

Here, $\mathcal{Y}$ is a subset of a Banach space $\mathfrak{X}$, and the operators $F : \mathfrak{X} \rightarrow \mathcal{L}_p(\Omega, \mathcal{F}, P)$ and $G : \mathfrak{X} \rightarrow \mathcal{L}_p(\Omega, \mathcal{F}, P)$ assign to each decision $x \in \mathfrak{X}$ the random variables $Z_F = F(x)$ and $Z_G = G(x)$, respectively. The functionals $\varrho : \mathcal{L}_p(\Omega, \mathcal{F}, P) \rightarrow \mathbb{R}$ and $\rho_t : \mathcal{L}_p(\Omega, \mathcal{F}, P) \rightarrow \mathbb{R}$ for all t in an interval $J \subset \mathbb{R}$ are convex and nondecreasing with respect to the almost sure order.

We assume that the decision space $\mathfrak{X}$ is a Banach space. Frequently, $\mathfrak{X} = \mathbb{R}^n$, but it is convenient to consider more general cases in view of applications to dynamic models. We pay special attention to the case when $\mathfrak{X} = \mathcal{L}_{p'}(\Omega, \mathcal{G}, P; \mathfrak{Y})$, where $p' \in [p, \infty]$ and $\mathfrak{Y}$ is a separable Banach space. The σ -subalgebra $\mathcal{G} \subseteq \mathcal{F}$ models the information available when the decision is made. If $\mathcal{G} = \{\emptyset, \Omega\}$ then $\mathfrak{X} = \mathfrak{Y}$.

The mapping $F : \mathfrak{X} \rightarrow \mathcal{L}_p(\Omega, \mathcal{F}, P)$, where $p \geq 1$, describes the dependence of a random “cost” in a stochastic system on the decision variables $x \in \mathfrak{X}$, that is, $[F(x)](\omega)$ is the cost associated with decision x and elementary event $\omega \in \Omega$. Usually, the mapping $F(\cdot)$ results from a composition of a cost function $c : \mathfrak{Y} \times \mathfrak{D} \rightarrow \mathbb{R}$ and a random data vector $D : \Omega \rightarrow \mathfrak{D}$, where $\mathfrak{D}$ is a space of the data vector realizations. Then $[F(x)](\omega) \triangleq c(x(\omega), D(\omega))$, $\omega \in \Omega$. The mapping $G : \mathfrak{X} \rightarrow \mathcal{L}_p(\Omega, \mathcal{F}, P)$ with $p \geq 1$ may represent a performance functional so that $[G(x)](\omega)$ measures a relevant system’s feature for the elementary event $\omega \in \Omega$.

We call the functionals $\varrho : \mathcal{L}_p(\Omega, \mathcal{F}, P) \rightarrow \mathbb{R}$ and $\rho_t : \mathcal{L}_p(\Omega, \mathcal{F}, P) \rightarrow \mathbb{R}$, which are convex and monotonic, *risk functionals*. We point out that the mathematical expectation is a special case of a risk functional and our results generalize the results on subdifferentiability involving normal integrands.

While our results may be applicable to many areas, we note that the adopted structure accommodates optimization of coherent or convex measures of risk and risk constraints in the form of stochastic orders, which are of particular interest to us. Modern theory of mathematical models of risk starts with the first axiomatic proposals due to [29] in the context of mean-risk models. That set of axioms does not include the monotonicity axiom. Monotonicity with respect to stochastic dominance was first proposed in [38]. The complete sets of axioms are due to [3] for measures defined on finite probability spaces. It was extended to general spaces in [10, 24, 32, 45]. Risk measures on the space of the quantile functions were considered in [16]. Further extensions include systemic measures of risk (see, e.g., [1, 6, 8, 21, 30, 44]) and measures on a probability space with a variable probability measure, called risk forms [17, 18].

Risk constraints based on stochastic orders were introduced in [11] for integer orders and further analyzed in [12]. Optimality conditions in Lagrangian form for a basic problem formulation were established first in [11]; for the inverse formulation of the second order stochastic dominance optimality conditions are provided in [13].

Optimization with stochastic dominance constraints relates also to risk functionals; another form of optimality conditions showcasing this relation is established in [15]. For relations to chance constraints or Average (Conditional) Value-at-Risk constraints, we refer to [14]. In this paper, we shall provide a subdifferential form of optimality conditions for a general setting with inverse stochastic dominance constraints.

In successive sections, we address questions associated with the subdifferentiability of the compositions $\varrho \circ F$ and $\rho_t \circ G$ as well as and their implications for the optimality conditions for problem (1). In the first part of our paper, we focus on the calculation of the subdifferential of the composite risk function $\varphi(\cdot) = \varrho(F(\cdot))$, where $\varrho : \mathcal{L}_p(\Omega, \mathcal{F}, P) \rightarrow \mathbb{R}$ is a measure of risk. In the second part of the paper, we analyze the compositions arising in the risk constraints in (1).

In what follows we understand all equations and inequalities between elements of the space $\mathcal{L}_p(\Omega, \mathcal{F}, P)$, with $p \geq 1$, in the almost sure sense.

1. Composition of a risk functional with a convex operator

First, we analyze the differentiability of the objective function and the constraint functions in problem (1). Convex lattice-valued operators and their subgradients were already considered by [50], and a thorough exposition can be found in the book [34]. We recall the basic definitions and elementary properties.

Definition 1.1. A mapping $F : \mathfrak{X} \rightarrow \mathcal{L}_p(\Omega, \mathcal{F}, P)$ is *convex* if for all $x, y \in \mathfrak{X}$ and all $\alpha \in [0, 1]$ we have

$$F(\alpha x + (1 - \alpha)y) \leq \alpha F(x) + (1 - \alpha)F(y). \quad (2)$$

For a convex mapping $F : \mathfrak{X} \rightarrow \mathcal{L}_p(\Omega, \mathcal{F}, P)$, at each $x \in \mathfrak{X}$ and for any $h \in \mathfrak{X}$, we can construct the differential quotients:

$$Q_t(x; h) \triangleq \frac{1}{t} [F(x + th) - F(x)], \quad t > 0. \quad (3)$$

They are nondecreasing (in the lattice sense) functions of t , and satisfy for $t \in (0, 1]$ the inequalities

$$F(x) - F(x - h) \leq Q_t(x; h) \leq F(x + h) - F(x).$$

Thus the directional derivative

$$F'(x; h) = \lim_{t \downarrow 0} Q_t(x; h), \quad h \in \mathfrak{X}, \quad (4)$$

is well-defined and belongs to $\mathcal{L}_p(\Omega, \mathcal{F}, P)$, as the limit in the strong and the order sense. The mapping $F'(x; \cdot)$ is convex and positively homogeneous by construction. If $F(\cdot)$ is continuous at x , then $F'(x; \cdot)$ is continuous at 0.

Analogously to the scalar case, we define subgradients of vector-valued convex mappings.

Definition 1.2. Suppose $F : \mathfrak{X} \rightarrow \mathcal{L}_p(\Omega, \mathcal{F}, P)$ is a convex mapping. A continuous linear operator $S : \mathfrak{X} \rightarrow \mathcal{L}_p(\Omega, \mathcal{F}, P)$ such that for all $h \in \mathfrak{X}$

$$Sh \leq F(x + h) - F(x) \quad (5)$$

is called the *subgradient* of $F(\cdot)$ at x . The set of all such operators is called the *subdifferential* of $F(\cdot)$ at x and is denoted by $\partial F(x)$. $\square$

It is evident that $S \in \partial F(x)$ if and only if for all $h \in \mathfrak{X}$

$$Sh \leq F'(x; h). \quad (6)$$

This characterization implies that the set $\partial F(x)$ is convex and compact in the weak operator topology; see [34, Lem. 1.1]. The fact that $\mathcal{L}_p(\Omega, \mathcal{F}, P)$ is a topological complete vector lattice plays a key role.

The following fundamental fact is due to [50, Thm. 6].

Theorem 1.3. Suppose $F : \mathfrak{X} \rightarrow \mathcal{L}_p(\Omega, \mathcal{F}, P)$ is convex and continuous at x . Then $\partial F(x) \neq \emptyset$ and for every $h \in \mathfrak{X}$ we can find $S \in \partial F(x)$ such that $F'(x; h) = Sh$.

We proceed now to new results in the setting of $\mathcal{L}_p$ -spaces.

For an operator $S : \mathfrak{X} \rightarrow \mathcal{L}_p(\Omega, \mathcal{F}, P)$ and a function $\alpha \in \mathcal{L}_\infty(\Omega, \mathcal{F}, P)$ we define the operator $\alpha S : \mathfrak{X} \rightarrow \mathcal{L}_p(\Omega, \mathcal{F}, P)$ by $[(\alpha S)h](\omega) = \alpha(\omega)[(Sh)(\omega)]$, for $h \in \mathfrak{X}$, $\omega \in \Omega$. If S is continuous, so is αS .

Lemma 1.4. For all $x \in \mathfrak{X}$, the set $\partial F(x)$ is convex in the following generalized sense: if $S_1, S_2 \in \partial F(x)$ and $\alpha : \Omega \rightarrow [0, 1]$ is measurable, then $\alpha S_1 + (1 - \alpha)S_2 \in \partial F(x)$ as well.

Proof. Using (6), for all $h \in \mathfrak{X}$ we obtain

$$\begin{aligned} [\alpha S_1 + (1 - \alpha)S_2]h &= [\alpha S_1]h + [(1 - \alpha)S_2]h = \alpha[S_1h] + (1 - \alpha)[S_2h] \\ &\leq \alpha F'(x; h) + (1 - \alpha)F'(x; h) = F'(x; h), \end{aligned}$$

and thus $\alpha S_1 + (1 - \alpha)S_2 \in \partial F(x)$. $\square$

We pass now to the analysis of the composition

$$\varphi(x) \triangleq \varrho(F(x)), \quad x \in \mathfrak{X}, \quad (7)$$

where $F : \mathfrak{X} \rightarrow \mathcal{L}_p(\Omega, \mathcal{F}, P)$ is a convex operator, and $\varrho : \mathcal{L}_p(\Omega, \mathcal{F}, P) \rightarrow \mathbb{R}$ is a risk functional. We say that a risk functional is nondecreasing, if $Z \leq V$ implies that $\varrho(Z) \leq \varrho(V)$. By [45, Prop. 3.1], a convex and nondecreasing risk functional on $\mathcal{L}_p(\Omega, \mathcal{F}, P)$, where $p \in [1, \infty]$, is subdifferentiable everywhere. It can be easily verified that if the mapping $F : \mathfrak{X} \rightarrow \mathcal{L}_p(\Omega, \mathcal{F}, P)$ is convex and $\varrho : \mathcal{L}_p(\Omega, \mathcal{F}, P) \rightarrow \mathbb{R}$ is convex and nondecreasing, then the composite function $\varrho \circ F$ is convex. To calculate its subdifferential, we first consider its directional derivatives.

Lemma 1.5. Suppose the operator $F : \mathfrak{X} \rightarrow \mathcal{L}_p(\Omega, \mathcal{F}, P)$ is convex and continuous at x , and the risk functional $\varrho : \mathcal{L}_p(\Omega, \mathcal{F}, P) \rightarrow \mathbb{R}$ is convex and nondecreasing. Then for all $x, h \in \mathfrak{X}$

$$\varphi'(x; h) = \varrho'(F(x); F'(x; h)). \quad (8)$$

Proof. The proof uses a standard argument, but we provide it for completeness. Recalling the definition (4), for all $\tau \geq 0$ we have

$$F(x + \tau h) = F(x) + \tau F'(x; h) + \tau r(\tau, h),$$

with $r(\tau, h) \downarrow 0$ when $\tau \downarrow 0$. Furthermore, by virtue of Theorem [45, Prop. 3.1], $\varrho(\cdot)$ is continuous and subdifferentiable everywhere, and for any $d \in \mathcal{L}_p(\Omega, \mathcal{F}, P)$ we have

$$\varrho(F(x) + \tau d) = \varrho(F(x)) + \tau \varrho'(F(x); d) + \tau R(\tau, d),$$

where $R(\tau, d) \downarrow 0$ when $\tau \downarrow 0$. Suppose $0 < \tau \leq \bar{\tau}$. Since $r(\cdot, h)$ is nondecreasing, the two expansions yield

$$\begin{aligned} \varphi(x + \tau h) - \varphi(x) &= \varrho(F(x + \tau h)) - \varrho(F(x)) \\ &= \varrho(F(x) + \tau F'(x; h) + \tau r(\tau, h)) - \varrho(F(x)) \\ &\leq \varrho(F(x) + \tau F'(x; h) + \tau r(\bar{\tau}, h)) - \varrho(F(x)) \\ &= \tau \varrho'(F(x); F'(x; h) + r(\bar{\tau}, h)) + \tau R(\tau, F'(x; h) + r(\bar{\tau}, h)). \end{aligned}$$

The directional derivative $\varrho'(F(x); \cdot)$ is convex and positively homogeneous, hence subadditive, and thus

$$\begin{aligned} \varphi(x + \tau h) - \varphi(x) &\leq \tau \varrho'(F(x); F'(x; h)) + \tau [\varrho'(F(x); r(\bar{\tau}, h)) + R(\tau, F'(x; h) + r(\bar{\tau}, h))]. \end{aligned}$$

On the other hand, the convexity of both functions and the monotonicity of $\varrho(\cdot)$ imply that

$$\varphi(x + \tau h) - \varphi(x) \geq \varrho(F(x) + \tau F'(x; h)) - \varrho(F(x)) \geq \tau \varrho'(F(x); F'(x; h)).$$

Combining the last two inequalities and dividing by τ we obtain

$$\begin{aligned} 0 &\leq \frac{1}{\tau} [\varphi(x + \tau h) - \varphi(x)] - \varrho'(F(x); F'(x; h)) \\ &\leq \varrho'(F(x); r(\bar{\tau}, h)) + R(\tau, F'(x; h) + r(\bar{\tau}, h)). \end{aligned}$$

After passing to the limit with $\tau \downarrow 0$, we get

$$0 \leq \varphi'(x; h) - \varrho'(F(x); F'(x; h)) \leq \varrho'(F(x); r(\bar{\tau}, h)).$$

As $\varrho'(F(x); r(\bar{\tau}, h)) \downarrow 0$, when $\bar{\tau} \downarrow 0$, equation (8) is true. $\square$

This allows us to calculate the subdifferential of the composition. The following result is known; we may refer here to [33], [34, Thm. 1.9], and [9]. We provide a version suitable for our risk measure applications with a simple proof, which improves on [45, Thm. 3.3] and [47, Thm 6.14].

Theorem 1.6. *Suppose that the operator $F : \mathfrak{X} \rightarrow \mathcal{L}_p(\Omega, \mathcal{F}, P)$, where $\mathfrak{X}$ is a Banach space and $p \in [1, \infty)$, is convex and continuous at x , and that the risk functional $\varrho : \mathcal{L}_p(\Omega, \mathcal{F}, P) \rightarrow \mathbb{R}$ is convex and nondecreasing. Then the composite function $\varphi = \varrho \circ F$ is subdifferentiable at x and*

$$\partial \varphi(x) = \bigcup_{\substack{\zeta \in \partial \varrho(F(x)) \\ S \in \partial F(x)}} S^* \zeta. \quad (9)$$

Proof. Using Lemma 1.5 and Theorem 1.3, we obtain the equation

$$\varphi'(x; h) = \max_{\zeta \in \partial \varrho(F(x))} \mathbb{E}[\zeta F'(x; h)] = \max_{\zeta \in \partial \varrho(F(x))} \mathbb{E}[\zeta \operatorname{ess\,sup}_{S \in \partial F(x)} Sh].$$

Since every $\zeta \in \partial \varrho(F(x))$ is nonnegative, and the “ess sup” above is attained, we can continue this chain of equations as follows:

$$\begin{aligned} \varphi'(x; h) &= \max_{\zeta \in \partial \varrho(F(x))} \max_{S \in \partial F(x)} \mathbb{E}[\zeta Sh] \\ &= \max_{\zeta \in \partial \varrho(F(x))} \max_{S \in \partial F(x)} \langle S^* \zeta, h \rangle = \max_{g \in \mathcal{D}} \langle g, h \rangle, \end{aligned} \quad (10)$$

where $\mathcal{D}$ is the set on the right hand side of (9).

We shall verify that $\mathcal{D}$ is convex. Therefore suppose that $g_1 = S_1^* \zeta_1$, $g_2 = S_2^* \zeta_2$, where $\zeta_1, \zeta_2 \in \partial \varrho(F(x))$, and $S_1, S_2 \in \partial F(x)$. For $0 < \alpha < 1$ we verify whether $g = \alpha g_1 + (1 - \alpha)g_2 \in \mathcal{D}$. By the convexity of $\partial \varrho(\cdot)$, we have

$$\zeta = \alpha \zeta_1 + (1 - \alpha) \zeta_2 \in \partial \varrho(F(x)).$$

Furthermore, $g = (\theta S_1 + (1 - \theta)S_2)^* \zeta$, with (for $\zeta(\omega) > 0$)

$$\theta(\omega) = \frac{\alpha \zeta_1(\omega)}{\zeta(\omega)}, \quad \omega \in \Omega.$$

If $\zeta(\omega) = 0$, then both $\zeta_1(\omega) = 0$ and $\zeta_2(\omega) = 0$, and we can set $\theta(\omega) = 0$. In any case, since $\zeta_1, \zeta_2 \geq 0$, we have $\theta \in [0, 1]$ a.s.. The generalized convexity of $\partial F(x)$ (see Lemma 1.4) entails

$$S = \theta S_1 + (1 - \theta)S_2 \in \partial F(x),$$

and thus $g \in \mathcal{D}$. Therefore, the set $\mathcal{D}$ is convex.

The closedness of $\mathcal{D}$ follows from the fact that for every $h \in \mathfrak{X}$ the maximum of $\langle g, h \rangle$ over $g \in \mathcal{D}$ is attained in (10). Denote a maximizer by $g(h)$. If a limit $\bar{g}$ of a sequence $\{g^k\}_{k \in \mathbb{N}} \subset \mathcal{D}$ was not an element of $\mathcal{D}$, then it could be separated from $\mathcal{D}$ by a linear functional $\bar{h}$:

$$\langle \bar{g}, \bar{h} \rangle > \langle g, \bar{h} \rangle, \quad \forall g \in \mathcal{D}.$$

This would lead to a contradiction:

$$\langle \bar{g}, \bar{h} \rangle > \langle g(\bar{h}), \bar{h} \rangle = \max_{g \in \mathcal{D}} \langle g, \bar{h} \rangle \geq \lim_{k \rightarrow \infty} \langle g^k, \bar{h} \rangle = \langle \bar{g}, \bar{h} \rangle.$$

Summing up, it follows from (10) that $\varphi'(x; \cdot)$ is the support function of a convex closed set $\mathcal{D}$, and thus $\mathcal{D} = \partial \varphi'(x; 0) = \partial \varphi(x)$. $\square$

Further analysis of the composition $\rho \circ F$ is dependent on the structure of the mapping $F(\cdot)$. We address this issue in the next subsection.

2. Composition of a risk functional with a convex integrand

In this section, $\mathfrak{X} = \mathcal{L}_{p'}(\Omega, \mathcal{F}, P; \mathfrak{Y})$, where $\mathfrak{Y}$ is a separable Banach space, and $p \in [1, \infty)$, $p' \in [p, \infty]$. We include the case of $p' = \infty$ in view of the application to the analysis in the next section. In some intermediate results, we also allow $p = \infty$.

We assume that the mapping $F : \mathcal{L}_{p'}(\Omega, \mathcal{F}, P; \mathfrak{Y}) \rightarrow \mathcal{L}_p(\Omega, \mathcal{F}, P)$ has the following structure:

$$[F(x)](\omega) = f(x(\omega), \omega), \quad \omega \in \Omega, \quad (11)$$

where $f : \mathfrak{Y} \times \Omega \rightarrow \mathbb{R}$ is a convex and continuous function of the first argument, for almost all values of the second argument, and measurable with respect to the second argument, for all values of the first argument. Such functions are called *convex integrands*.

The mapping $F(\cdot)$ is convex in the sense of Definition 1.1, but our setting is more specific because it excludes a null set *before* all $x, y \in \mathfrak{X}$ are considered. This allows for an explicit description of the subgradients of $F(\cdot)$.

The differential quotients (3) of the mapping (11) have a specific form: for almost all $\omega \in \Omega$

$$[Q_t(x; h)](\omega) = \frac{1}{t} [f(x(\omega) + th(\omega), \omega) - f(x(\omega), \omega)], \quad t > 0.$$

By the convexity of $f(\cdot, \omega)$ we conclude that for all $h \in \mathfrak{X}$ and for almost all $\omega \in \Omega$

$$[F'(x; h)](\omega) = \lim_{t \downarrow 0} [Q_t(x; h)](\omega) = f'(x(\omega), \omega; h(\omega)). \quad (12)$$

We can now establish the decomposable structure of the subdifferential.

Lemma 2.1. *If $L \in \partial F(x)$ then L is local in the following sense: for all $B \in \mathcal{F}$ and all $h \in \mathfrak{X}$ we have*

$$L(\mathbb{1}_B h) = \mathbb{1}_B Lh.$$

Proof. For all events $B \in \mathcal{F}$ and all $h \in \mathfrak{X}$ we have $F'(x; \mathbb{1}_B h) = \mathbb{1}_B F'(x; h)$, and thus every $L \in \partial F(x)$ satisfies the inequalities:

$$-\mathbb{1}_B F'(x; -h) \leq L(\mathbb{1}_B h) \leq \mathbb{1}_B F'(x; h). \quad (13)$$

Consequently, $L(\mathbb{1}_B h)$ vanishes outside of B . As L is linear,

$$\mathbb{1}_B L(\mathbb{1}_B h) = \mathbb{1}_B Lh - \mathbb{1}_B L(\mathbb{1}_{B^c} h) = \mathbb{1}_B Lh,$$

because the second term disappears due to (13) for B^c . Therefore, $L(\mathbb{1}_B h)$ coincides with Lh on B . $\square$

The local property is essential in the following theorem. We draw the attention of the Reader that our technique using the Radon-Nikodym theorem for vector measures, covers the notoriously difficult case of $p' = \infty$. The latter case is less difficult than usual due to the continuity assumption about the integrand $f(\cdot, \omega)$, which does not take the value $+\infty$.

Theorem 2.2. Suppose that $F : \mathcal{L}_{p'}(\Omega, \mathcal{F}, P; \mathfrak{Y}) \rightarrow \mathcal{L}_p(\Omega, \mathcal{F}, P)$ is given by (11), $1 \leq p \leq p' \leq \infty$, $\mathfrak{Y}$ a separable Banach space, $f(\cdot, \omega)$ convex and continuous at $x(\omega)$ for almost all $\omega \in \Omega$. Then $S \in \partial F(x)$ if and only if a function $s \in \mathcal{L}_r(\Omega, \mathcal{F}, P; \mathfrak{Y}^*)$ exists, with $1/p' + 1/r = 1/p$, such that we have $s(\omega) \in \partial f(x(\omega), \omega)$ for almost all $\omega \in \Omega$, and for all $h \in \mathcal{L}_{p'}(\Omega, \mathcal{F}, P; \mathfrak{Y})$

$$[Sh](\cdot) = \langle s(\cdot), h(\cdot) \rangle, \quad a.s.. \quad (14)$$

Proof. Every operator of the form (14), with a measurable selector $s(\cdot) \in \partial f(x(\cdot), \cdot)$, is a subgradient of F at x , because for all $h \in \mathfrak{X}$

$$\langle s(\omega), h(\omega) \rangle \leq f'(x(\omega), \omega; h(\omega)) = [F'(x; h)](\omega),$$

for almost all $\omega \in \Omega$, and thus Definition 1.2 is satisfied. Since $s \in \mathcal{L}_r(\Omega, \mathcal{F}, P; \mathfrak{Y}^*)$, in view of the Hölder inequality, the operator (14) is a continuous mapping from $\mathcal{L}_{p'}(\Omega, \mathcal{F}, P; \mathfrak{Y})$ to $\mathcal{L}_p(\Omega, \mathcal{F}, P)$.

The difficult part is to prove that no other subgradients exist. Suppose $L \in \partial F(x)$. Consider the functional $\ell : \mathfrak{X} \rightarrow \mathbb{R}$ defined as

$$\ell(h) = \mathbb{E}[Lh], \quad h \in \mathfrak{X}.$$

As L is linear and continuous, so is $\ell(\cdot)$. Due to the local property of L established in Lemma 2.1,

$$\ell(\mathbb{1}_B h) = \mathbb{E}[L(\mathbb{1}_B h)] = \mathbb{E}[\mathbb{1}_B Lh], \quad \forall B \in \mathcal{F}. \quad (15)$$

Take any $y \in \mathfrak{Y}$ and set a constant $h = \mathbb{1}_\Omega y$. Then

$$\ell(\mathbb{1}_B y) = \mathbb{E}[\mathbb{1}_B L(\mathbb{1}_\Omega y)].$$

Since the functional $y \mapsto \ell(\mathbb{1}_B y)$ is continuous, then $M(B) \in \mathfrak{Y}^*$ exists, such that

$$\mathbb{E}[\mathbb{1}_B L(\mathbb{1}_\Omega y)] = \langle M(B), y \rangle, \quad \forall y \in \mathfrak{Y}.$$

It follows from the last equation that the function $B \mapsto M(B)$ is a $\mathfrak{Y}^*$ -valued, finite, and countably additive vector measure on $\mathcal{F}$, which is absolutely continuous with respect to P . By virtue of the Radon-Nikodym theorem for vector measures [20, Ch. III], a measurable function $v : \Omega \rightarrow \mathfrak{Y}^*$ exists, such that

$$M(B) = \int_B v(\omega) P(d\omega), \quad \forall B \in \mathcal{F}.$$

Combining the last three equations, we obtain

$$\ell(\mathbb{1}_B y) = \int_\Omega \langle v(\omega), \mathbb{1}_B(\omega) y \rangle P(d\omega), \quad \forall B \in \mathcal{F}.$$

For any simple function we have $h = \sum_{n=1}^N \mathbb{1}_{B_n} y_n$, and using the linearity of $\ell(\cdot)$ we get

$$\ell(h) = \sum_{n=1}^M \int_\Omega \langle v(\omega), \mathbb{1}_{B_n}(\omega) y_n \rangle P(d\omega) = \int_\Omega \langle v(\omega), h(\omega) \rangle P(d\omega).$$

Therefore, $\ell(h) = \mathbb{E}[\langle v(\cdot), h(\cdot) \rangle], \quad \forall h \in \mathfrak{X}.$

This, combined with (15), entails the equation

$$\mathbb{E}[\mathbb{1}_B Lh] = \mathbb{E}[\mathbb{1}_B \langle v(\cdot), h(\cdot) \rangle], \quad \forall B \in \mathcal{F}.$$

Thus, for all $h \in \mathfrak{X}$, we have $[Lh](\cdot) = \langle v(\cdot), h(\cdot) \rangle$, a.s..

This representation implies that the function $\langle v(\cdot), h(\cdot) \rangle$ is an element of $\mathcal{L}_p(\Omega, \mathcal{F}, P)$, for all $h \in \mathcal{L}_{p'}(\Omega, \mathcal{F}, P; \mathfrak{Y})$. Therefore, $v \in \mathcal{L}_r(\Omega, \mathcal{F}, P; \mathfrak{Y}^*)$.

Let $\mathcal{S} = \{s \in \mathcal{L}_r(\Omega, \mathcal{F}, P; \mathfrak{Y}^*) : s(\cdot) \in \partial f(x(\cdot), \cdot)\}.$

Suppose $v \notin \mathcal{S}$. Because $\mathcal{S}$ is convex and closed, by virtue of the separation theorem, a function $\bar{h} \in \mathfrak{X}$ and a real number $\varepsilon > 0$ exist, such that for all $s \in \mathcal{S}$

$$\mathbb{E}[\langle v, \bar{h} \rangle] \geq \mathbb{E}[\langle s, \bar{h} \rangle] + \varepsilon. \quad (16)$$

Let $\bar{s}$ be an element of $\mathcal{S}$ such that

$$\langle \bar{s}(\cdot), \bar{h}(\cdot) \rangle = f'(x(\cdot), \cdot; \bar{h}(\cdot)); \quad (17)$$

such a function can be found, owing to the measurable selection theorem. Indeed, for almost all $\omega \in \Omega$ we have

$$f'(x(\omega), \omega; \bar{h}(\omega)) = \max_{s(\omega) \in \partial f(\bar{x}(\omega), \omega)} \langle s(\omega), \bar{h}(\omega) \rangle;$$

the maximum is attained at some $\bar{s}(\omega)$, due to the weak* compactness of $\partial f(\bar{x}(\omega), \omega)$. The relations (16)–(17) imply that

$$\mathbb{E}[L\bar{h}] = \mathbb{E}[\langle v, \bar{h} \rangle] \geq \mathbb{E}[F'(x; \bar{h})] + \varepsilon.$$

which contradicts (6). Consequently, $v \in \mathcal{S}$. □

We can now calculate the subdifferential of the composition $\varphi(\cdot) = \varrho(F(\cdot))$, where $\varrho(\cdot)$ is a risk functional.

Theorem 2.3. *Suppose $F : \mathcal{L}_{p'}(\Omega, \mathcal{F}, P; \mathfrak{Y}) \rightarrow \mathcal{L}_p(\Omega, \mathcal{F}, P)$ is given by (11), with $p \in [1, \infty)$ and $p' \in [p, \infty]$, $f(\cdot, \omega)$ is convex and continuous at $x(\omega)$ for almost all $\omega \in \Omega$, $F(\cdot)$ is continuous at x , and $\mathfrak{Y}$ is a separable Banach space. Further, suppose $\varrho : \mathcal{L}_p(\Omega, \mathcal{F}, P) \rightarrow \mathbb{R}$ is convex and nondecreasing. Then the composite function $\varphi = \varrho \circ F$ is subdifferentiable at x and*

$$\partial \varphi(x) = \bigcup_{\zeta \in \partial \varrho(F(x))} \{g \in \mathcal{L}_{q'}(\Omega, \mathcal{F}, P; \mathfrak{Y}^*) : g(\omega) \in \zeta(\omega) \partial f(x(\omega), \omega), \text{ a.s.}\}, \quad (18)$$

where $1/p' + 1/q' = 1$.

Proof. We use the representation of the subdifferential established in Theorem 1.6 and the characterizations of the subgradients of $F(\cdot)$ established in Theorem 2.2. For any $h \in \mathfrak{X}$, any $\zeta \in \partial \varrho(F(x))$ and any $S \in \partial F(x)$ we have

$$\langle S^* \zeta, h \rangle = \langle \zeta, Sh \rangle = \int_{\Omega} \langle \zeta(\omega) s(\omega), h(\omega) \rangle P(d\omega),$$

where $s(\cdot) \in \partial f(x(\cdot), \cdot)$. This verifies (18). □

3. The Case of partial information

We now assume that $\mathfrak{X} = \mathcal{L}_{p'}(\Omega, \mathcal{G}, P; \mathfrak{Y})$, where $\mathfrak{Y}$ is a separable Banach space, and $\mathcal{G} \subseteq \mathcal{F}$ represents the information available when the decision x is made. As before, $p \in [1, \infty)$, $p' \in [p, \infty]$. The special case of $p' = \infty$ and $\mathcal{G} = \{\emptyset, \Omega\}$ covers the case when x is a deterministic element of $\mathfrak{Y}$.

The mapping $F : \mathfrak{X} \rightarrow \mathcal{L}_p(\Omega, \mathcal{F}, P)$ is defined by a convex integrand (11). Again, the mapping $F(\cdot)$ is convex and we plan to analyze the structure of its subdifferential. The directional derivatives of $F(\cdot)$ still have the form (12).

The following result extends Theorem 2.2 to the case of partial information.

Theorem 3.1. *Suppose $F : \mathfrak{X} \rightarrow \mathcal{L}_p(\Omega, \mathcal{F}, P)$, with $p \in [1, \infty]$, is given by (11), $\mathfrak{Y}$ is a separable Banach space, $f(\cdot, \omega)$ is convex and continuous at $x(\omega)$ for almost all $\omega \in \Omega$, and $F(\cdot)$ is continuous at x as well. Then $S \in \partial F(x)$ if and only if a function $s \in \mathcal{L}_r(\Omega, \mathcal{F}, P; \mathfrak{Y}^*)$ exists, with $1/p' + 1/r = 1/p$, such that $s(\omega) \in \partial f(x(\omega), \omega)$ for almost all $\omega \in \Omega$, and for all $h \in \mathfrak{X}$*

$$[Sh](\cdot) = \langle s(\cdot), h(\cdot) \rangle, \quad \text{a.s.} \quad (19)$$

Proof. As in Theorem 2.2, we can verify that every operator S of the form (19) is a subgradient of $F(\cdot)$. We need to prove that no other subgradients exist.

Our idea is to allow x and h to be $\mathcal{F}$ -measurable and to leverage Theorem 2.2. Define the space $\mathbb{X} = \mathcal{L}_{p'}(\Omega, \mathcal{F}, P; \mathfrak{Y})$ and denote by $\mathfrak{x}$ and $\mathfrak{h}$ its elements. Let the mapping $\mathbb{F} : \mathbb{X} \rightarrow \mathcal{L}_p(\Omega, \mathcal{F}, P)$ be still defined as in (11).

Suppose $\mathfrak{x} \in \mathbb{X}$ is equal to x a.s.. The function $\mathbb{F}'(\mathfrak{x}; \cdot)$ is convex, positively homogeneous, and continuous on $\mathbb{X}$. Suppose $S \in \partial F(x)$. Then $Sh \leq F'(x; h)$ for all $h \in \mathfrak{X}$. This means that $S(\cdot)$ is a minorant of $\mathbb{F}'(\mathfrak{x}; \cdot)$ on the subspace of $\mathcal{G}$ -measurable functions $\mathcal{N} = \{\mathfrak{h} \in \mathbb{X} : \mathbb{E}[\mathfrak{h}|\mathcal{G}] = \mathfrak{h}\}$. By the generalization of the Hahn-Banach theorem to lattice-valued operators (see, [28, Thm. X.5.7]), S can be extended to a linear functional $\mathbb{S}$ on $\mathbb{X}$ which is a minorant of $\mathbb{F}'(\mathfrak{x}; \cdot)$ and which coincides with S on $\mathcal{N}$. As $\mathbb{F}'(\mathfrak{x}; \cdot)$ is continuous, $\mathbb{S} \in \partial \mathbb{F}(\mathfrak{x})$. By Theorem 2.2 at the point $\mathfrak{x}$, a function $s \in \mathcal{L}_r(\Omega, \mathcal{F}, P; \mathfrak{Y}^*)$ exists, such that for almost all $\omega \in \Omega$ we have $s(\omega) \in \partial f(x(\omega), \omega)$, and for all $\mathfrak{h} \in \mathbb{X}$

$$\mathbb{S}\mathfrak{h} = \langle s, \mathfrak{h} \rangle, \quad \text{a.s.}$$

On the subspace $\mathcal{N}$, the above formula reduces to (19). □

The subdifferential of $\varphi(x) = \varrho(F(x))$ can be calculated in a similar way as in Theorem 2.3.

Theorem 3.2. *Suppose $\mathfrak{Y}$ is a separable Banach space, $F : \mathfrak{X} \rightarrow \mathcal{L}_p(\Omega, \mathcal{F}, P)$ is given by (11), $p \in [1, \infty)$, $f(\cdot, \omega)$ is convex and continuous at x for almost all $\omega \in \Omega$, $F(\cdot)$ is continuous at x as well, and $\varrho : \mathcal{L}_p(\Omega, \mathcal{F}, P) \rightarrow \mathbb{R}$ is convex and nondecreasing. Then the composite function $\varphi = \varrho \circ F$ is subdifferentiable at x and¹*

$$\partial \varphi(x) = \bigcup_{\substack{\zeta \in \partial \varrho(F(x)) \\ s(\cdot) \leq \partial f(x(\cdot), \cdot)}} \mathbb{E}[\zeta s | \mathcal{G}]. \quad (20)$$

¹ The symbol “ $\leq$ ” means “is a measurable selector of.”

Proof. We use the representation of the subdifferential established in Theorem 1.6 and the characterizations of the subgradients of $F(\cdot)$ established in Theorem 3.2. For all $h \in \mathfrak{X}$, $\zeta \in \partial \varrho(F(x))$, and $S \in \partial F(x)$, we have

$$\langle S^* \zeta, h \rangle = \langle \zeta, Sh \rangle = \mathbb{E}[\zeta \langle s, h \rangle] = \mathbb{E}[\langle \mathbb{E}[\zeta s | \mathcal{G}], h \rangle],$$

where $s(\omega) \in \partial f(x(\omega), \omega)$. This verifies (20). $\square$

Our results, in a special case, allow to obtain the famous subdifferential disintegration result, which is known as the *Strassen Theorem*. In the case of a positively homogeneous $f(\cdot, \omega)$ and $\varrho(\cdot) \equiv \mathbb{E}[\cdot]$, it was discovered in [48]. Generalizations to convex functions and convex integrands are provided in [27, 41, 42] and the books [7, 34].

Theorem 3.3. *Suppose $\mathfrak{X}$ is a separable Banach space, $F : \mathfrak{X} \rightarrow \mathcal{L}_1(\Omega, \mathcal{F}, P)$ is given by $F(x, \omega) = f(x, \omega)$, $\omega \in \Omega$, $f(\cdot, \omega)$ is convex and continuous at x for almost all $\omega \in \Omega$, and $F(\cdot)$ is continuous at x as well. Then the function $\mathbb{E}[F(\cdot)]$ is subdifferentiable at x and*

$$\partial \mathbb{E}[F(x)] = \mathbb{E}[\partial F(x)] = \int_{\Omega} \partial f(x, \omega) P(d\omega).$$

The result follows from setting $p = 1$, $p' = \infty$, $\mathcal{G} = \{\emptyset, \Omega\}$, and $\varrho(\cdot) \equiv \mathbb{E}[\cdot]$ in Theorem 3.2.

4. Stochastic dominance operators

We now turn to convex operators associated with stochastic dominance relations.

To the best of our knowledge the first general notion of a stochastic order was introduced in [35]. It is related to the theory of weak majorization, as discussed in [2]. A comparison of random variables with respect to the first-order stochastic dominance is used in the context of statistical tests in the early works [5, 31, 36]. The stochastic dominance relations have received considerable attention after the works [4, 22, 25, 26, 40, 43]. The characterization of the first and second-order dominance via families of (differentiable) utility functions was stated in [25] and rigorously proved in [4], where also the third-order dominance is characterized. Stochastic dominance of fractional order $p \geq 1$ was introduced for non-negative bounded random variables in [23], where also the monotonicity of the relation was proved. A thorough survey and analysis of stochastic order relations is contained in the monographs [37] and [46]. Below, we provide the information about fractional orders that is relevant to our research.

Consider a random variable $Z \in \mathcal{L}_1(\Omega, \mathcal{F}, P)$ with $p \in [1, \infty]$ and its distribution function $H_Z(\eta) = P[Z \leq \eta]$ for $\eta \in \mathbb{R}$. We define the integrated distribution function as follows:

$$H_Z^{(2)}(\eta) = \int_{-\infty}^{\eta} H_Z(\alpha) d\alpha = \mathbb{E}[\max(\eta - Z, 0)] \quad \text{for } \eta \in \mathbb{R}. \quad (21)$$

Recall that the left-continuous inverse of the cumulative distribution function $H_Z(\cdot)$ is defined as follows:

$$H_Z^{(-1)}(p) = \inf \{ \eta : H_Z(\eta) \geq p \} \quad \text{for all } 0 < p < 1.$$

Given $p \in [0, 1]$, the number $q = q_Z(p)$ is called a p -quantile of the random variable X if

$$P(Z < q) \leq p \leq P(Z \leq q).$$

For $p \in (0, 1)$, the set of p -quantiles is a closed interval and $H_Z^{(-1)}(p)$ represents its left end. We adopt the convention that $H_Z^{(-1)}(1) = +\infty$, if the 1-quantile does not exist.

The integrated quantile function $H_Z^{(-2)} : \mathbb{R} \rightarrow \mathbb{R}$ is defined as

$$H_Z^{(-2)}(p) = \int_0^p H_Z^{(-1)}(\alpha) d\alpha \quad \text{for } 0 < p \leq 1, \quad (22)$$

We also define $H_Z^{(-2)}(0) = 0$ and $H_Z^{(-2)}(p) = +\infty$ for all real numbers $p \notin [0, 1]$. The function $H_Z^{(-2)}(\cdot)$ is the absolute Lorenz function introduced in [35]. It is well defined for any random variable $Z \in \mathcal{L}_1(\Omega, \mathcal{F}, P)$. It is convex as an integral of a non-decreasing function.

The Fenchel duality relation between the integrated quantile function $H_Z^{(-2)}(\cdot)$ and the integrated distribution function $H_Z^{(2)}(\cdot)$ has been established in [39].

Theorem 4.1. *For every integrable random variable Z , we have*

$$H_Z^{(-2)} = [H_Z^{(2)}]^* \quad \text{and} \quad H_Z^{(2)} = [H_Z^{(-2)}]^*.$$

Definition 4.2. A random variable X dominates in the first order the random variable Y if

$$H_X(\eta) \leq H_Y(\eta) \quad \forall \eta \in \mathbb{R} \Leftrightarrow H_X^{(-1)}(p) \geq H_Y^{(-1)}(p) \quad \forall p \in (0, 1). \quad (23)$$

For $X, Y \in \mathcal{L}_1(\Omega, \mathcal{F}, P)$, X dominates in the second order the random variable Y if

$$H_X^{(2)}(\eta) \leq H_Y^{(2)}(\eta) \quad \forall \eta \in \mathbb{R} \Leftrightarrow H_X^{(-2)}(p) \geq H_Y^{(-2)}(p) \quad \forall p \in (0, 1]. \quad (24)$$

The equivalence of the two requirements in (23) is obvious while the equivalence of the two conditions in (24) follows from Theorem 4.1. In our analysis, we shall focus on the second order dominance relation expressed as a relation between the Lorenz functions of the random variable in question.

Let $J = [\alpha, \beta] \subset (0, 1]$. We shall relax the second-order stochastic dominance relation to the interval J , i.e., for two integrable random variables X and Y , it is said that X is larger than the random variable Y w.r.to the Lorenz dominance on J if

$$H_X^{(-2)}(\eta) \geq H_Y^{(-2)}(\eta) \quad \text{for all } p \in J. \quad (25)$$

We define the set

$$\mathbb{B}(Y, J) = \{X \in \mathcal{L}_1(\Omega, \mathcal{F}, P) : H_X^{(-2)}(p) \geq H_Y^{(-2)}(p) \quad \forall p \in J\}. \quad (26)$$

The following theorem establishes relevant properties of the set $B(Y, J)$. Statements (a) and (c) were shown in [19]; we provide the whole proof for the sake of clarity and convenience of the reader.

Lemma 4.3. (a) *For every $p \in (0, 1)$, the mapping $Z \mapsto H_Z^{(-2)}(p)$ is continuous, concave and positively homogeneous on $\mathcal{L}_1(\Omega, \mathcal{F}, P)$.*

(b) *The concave subdifferential of the mapping $Z \mapsto H_Z^{(-2)}(p)$ has the form*

$$\partial_Z H_Z^{(-2)}(p) = \left\{ p\zeta \in \mathcal{L}_\infty(\Omega, \mathcal{F}, P) : \begin{array}{l} \mathbb{E}[\zeta] = 1, \ 0 \leq \zeta \leq \frac{1}{p}, \\ \mathbb{E}[\zeta Z] = \frac{1}{p} H_Z^{(-2)}(p) \end{array} \right\}. \quad (27)$$

(c) *For any interval $J \subseteq (0, 1]$, the multifunction $Y \rightarrow \mathbb{B}(Y, J)$ is convex-valued and has a closed graph in $\mathcal{L}_1(\Omega, \mathcal{F}, P)$.*

Proof. The mapping $Z \mapsto H_Z^{(-2)}(p)$ is identical to $Z \mapsto -p \text{AVaR}_p^-(Z)$, where $\text{AVaR}_p^-(Z)$ is the Average Value-at-Risk representing profits (see [19, Remark 2.20]). Hence, the concavity and the positive homogeneity in statement (a) follow from the coherence of $\text{AVaR}_p^-(\cdot)$. Using the formula for the subdifferential of $\text{AVaR}_p^-(\cdot)$, we obtain statement (b).

Further, the continuity of $Z \mapsto \text{AVaR}_p^-(Z)$ implies the continuity of the mapping $Z \mapsto H_Z^{(-2)}(p)$ in $\mathcal{L}_1(\Omega, \mathcal{F}, P)$, which in turn implies that the graph of the multifunction $Y \rightarrow \mathbb{B}(Y, J)$ is closed. If $X, Z \in \mathbb{B}(Y, J)$, then the concavity of $Z \mapsto H_Z^{(-2)}(p)$ implies that $tX + (1 - t)Z \in \mathbb{B}(Y, J)$ entailing the convexity of the set $\mathbb{B}(Y, J)$. $\square$

5. Optimization with stochastic dominance constraints

Optimization problems with stochastic dominance constraints were introduced in [11] for integer orders and further analyzed in [12]; a generalization to fractional orders is given in [19]. Optimality conditions in Lagrangian form as well as in subdifferential form show the role of utility functions as Lagrange multipliers. Our objective here is to derive new optimality conditions in subdifferential form for the inverse formulation of stochastic dominance. Optimality conditions in Lagrangian form established in [13, 15, 19] show relations to distortion functionals and coherent measures of risk.

We analyze the following optimization problem:

$$\min \varphi(x) \quad (28)$$

$$\text{subject to } H_{G(x)}^{(-2)}(p) \geq H_Y^{(-2)}(p) \quad \forall p \in J = [\alpha, \beta] \subset (0, 1], \quad (29)$$

$$x \in \mathcal{Y}. \quad (30)$$

The formulation fits the setting of (1) with $\rho_p(G(x)) = H_Y^{(-2)}(p) - H_{G(x)}^{(-2)}(p)$.

The function $\varphi(\cdot)$ may be the composition (7); below we just use its general form for simplicity. In this section, we focus on the constraint operators and adopt the following general assumptions.

- Assumption 5.1.** (i) The function $\varphi(\cdot)$ is convex and continuous.
(ii) The operator $G : \mathfrak{X} \rightarrow \mathcal{L}_p(\Omega, \mathcal{F}, P)$ is norm-to-norm continuous and has the structure $[G(x)](\omega) = g(x, \omega)$, where $g : \mathfrak{X} \times \Omega \rightarrow \mathbb{R}$ is a Carathéodory function and $g(\cdot, \omega)$ is concave for all $\omega \in \Omega$.
(iii) The set $\mathcal{Y}$ is a closed and convex subset of a Banach space $\mathfrak{X}$.

The following requirement is our constraint qualification condition.

Definition 5.2. Problem (28)–(30) satisfies the *uniform inverse dominance condition* if a point $\tilde{x} \in \mathcal{Y}$ exists such that

$$\inf_{p \in J} [H_{G(x)}^{(-2)}(p) - H_Y^{(-2)}(p)] > 0.$$

The following result is known; see [19, Theorem 5.51].

Theorem 5.3. Suppose Assumption 5.1 and the uniform inverse dominance condition are satisfied. If $\hat{x}$ is an optimal solution of (28)–(30) then a spectral risk measure $\hat{\rho}(\cdot)$ and a constant $\kappa \geq 0$ exist such that $\hat{x}$ is also an optimal solution of the problem

$$\min_{x \in \mathcal{Y}} \{f(x) + \kappa \hat{\rho}(G(x))\}, \quad (31)$$

$$\kappa \hat{\rho}(G(\hat{x})) = \kappa \hat{\rho}(Y). \quad (32)$$

Conversely, if for some spectral risk measure $\hat{\rho}(\cdot)$ and some $\kappa \geq 0$ an optimal solution $\hat{x}$ of (31) satisfies (29) with $[\alpha, \beta] = [0, 1]$ and (32), then $\hat{x}$ is an optimal solution of (28)–(30) with $[\alpha, \beta] = [0, 1]$.

Recall that the normal cone to the convex set $\mathcal{Y} \subset \mathfrak{X}$ at a point $x \in \mathcal{Y}$ is given by the formula:

$$\mathcal{N}_{\mathcal{Y}}(x) = \{d \in \mathfrak{X}^* : \langle d, y - x \rangle \leq 0 \quad \forall y \in \mathcal{Y}\}$$

with $\mathfrak{X}^*$ being the topological dual to $\mathfrak{X}$.

For a fixed ω , the concave subdifferential $\partial g(\hat{x}, \omega)$ at $\hat{x}$ is as follows

$$\partial g(\hat{x}, \omega) = \{s \in \mathfrak{X}^* : g(x, \omega) \leq g(\hat{x}, \omega) + \langle s(\omega), x - \hat{x} \rangle \quad \forall x \in \mathfrak{X}\}.$$

In the theorem below, we use Definition 1.2 of a subgradient of a lattice-valued convex operator adapted to the concave case discussed here.

Theorem 5.4. Suppose Assumption 5.1 and the uniform dominance condition of Definition 5.2 are satisfied for problem (28)–(30). If $\hat{x}$ is an optimal solution to that problem, then a subgradient $S \in \partial[G(\hat{x})]$, a probability measure $\hat{\mu}$ supported on $[\alpha, \beta]$, and a constant $\kappa \geq 0$ exist such that the following inclusion

$$0 \in \partial\varphi(\hat{x}) + \kappa S^* \int_{\alpha}^{\beta} \partial \text{AVaR}_p^-(G(\hat{x})) \hat{\mu}(dp) + \mathcal{N}_{\mathcal{X}}(\hat{x}), \quad (33)$$

along with the complementarity condition (32) are satisfied. Conversely, if for some probability measure $\hat{\mu}(\cdot)$ on $[0, 1]$ and some $\kappa \geq 0$ the inclusion (33) is satisfied with a subgradient $S \in \partial[G(\hat{x})]$, as well as (32) and (29) hold with $[\alpha, \beta] = [0, 1]$, then $\hat{x}$ is an optimal solution of (28)–(30) with $[\alpha, \beta] = [0, 1]$.

Proof. In view of Theorem 5.3, we only need show the equivalence of conditions (33) and (31). Notice that if $\kappa = 0$, then condition (31) means that $\hat{x}$ minimizes $f(\cdot)$ over the set $\mathcal{Y}$ and constraint (29) is not active. In that case, (31) is equivalent to (33) trivially.

Assuming that $\kappa > 0$, the spectral risk measure in condition (31) has the form

$$\hat{\varrho}(Z) = \int_{\alpha}^{\beta} \text{AVaR}_p^-(Z) \hat{\mu}(dp)$$

with some probability measure $\hat{\mu}$ supported on the interval $[\alpha, \beta]$.

Theorem 3.1 implies that the operator $G : \mathfrak{X} \rightarrow \mathcal{L}_p(\Omega, \mathcal{F}, P)$ is subdifferentiable at $\hat{x}$ and its subdifferential is given by

$$\partial[G(\hat{x})] = \left\{ S : \mathfrak{X} \rightarrow \mathcal{L}_p(\Omega, \mathcal{F}, P) : \begin{array}{l} [Sd](\omega) = \langle s(\omega), d \rangle \text{ for some} \\ s \in \mathcal{L}_p(\Omega, \mathcal{F}, P; \mathfrak{X}^*); s(\omega) \in \partial g(\hat{x}, \omega) \end{array} \right\}.$$

Under the assumptions of the theorem, the function

$$\mathfrak{L}(x) = f(x) + \kappa \hat{\varrho}(G(x)) = f(x) + \kappa \int_{\alpha}^{\beta} \text{AVaR}_p^-(G(x)) \hat{\mu}(dp)$$

is convex. Hence $\hat{x} \in \mathcal{Y}$ minimizes $\mathfrak{L}(x)$ if and only if for some $S \in \partial G(\hat{x})$

$$0 \in \partial f(\hat{x}) + \kappa \partial_x \hat{\varrho}(G(\hat{x})) = \partial f(\hat{x}) + \kappa \int_{\alpha}^{\beta} S^* \partial \text{AVaR}_p^-(G(\hat{x})) \hat{\mu}(dp).$$

The right-hand side is equivalent to the one in (33). $\square$

We can obtain another equivalent form of the integral term in condition (33) for the non-trivial case of $\kappa > 0$:

$$\begin{aligned} & \kappa \int_{\alpha}^{\beta} S^* \partial \text{AVaR}_p^-(G(\hat{x})) \hat{\mu}(dp) \\ &= \kappa \left\{ S^* \int_{\alpha}^{\beta} \zeta(\hat{x}, p, \omega) \hat{\mu}(dp) : \zeta(\hat{x}, p, \omega) \in \partial \text{AVaR}_p^-(G(\hat{x})) \right\} \\ &= \left\{ S^* \int_{\alpha}^{\beta} \xi(\hat{x}, p, \omega) \hat{\nu}(dp) : \xi(\hat{x}, p, \omega) \in \partial H_{G(\hat{x})}^{(-2)}(p) \right\}. \end{aligned}$$

In the last formula $\hat{\nu}(dp) = \frac{\kappa}{p} \hat{\mu}(dp)$.

Conclusions

We summarize our main contributions as follows. We provide new results on subdifferentiability of convex operators and their compositions which go beyond the theory of normal integrands. The key technical advance is the description of the subdifferential of a local convex operator between $\mathcal{L}_p$ spaces in Theorem 2.2. We offer a new proof technique using the Radon-Nikodym theorem for vector measures. The key statement in Theorem 2.2 allows for dealing with the new model with partial information in section 3. Additionally, we offer new optimality conditions in subdifferential form for optimization problems with inverse stochastic dominance constraints.

References

- [1] A. Almen, D. Dentcheva: *On risk evaluation and control of distributed multi-agent systems*, J. Optimization Theory Appl. 203 (2024) 2025–2054.
- [2] B. C. Arnold: *Majorization and the Lorenz Order: A Brief Introduction*, Lecture Notes in Statistics 43, Springer, Berlin (1987).
- [3] P. Artzner, F. Delbaen, J.-M. Eber, D. Heath: *Coherent measures of risk*, Math. Finance 9 (1999) 203–228.
- [4] V. S. Bawa: *Optimal rules for ordering uncertain prospects*, J. Financial Econ. 2 (1975) 95–121.
- [5] D. Blackwell: *Comparison of experiments*, in: *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability*, vol. 2, University of California Press, Berkeley (1951) 93–103.
- [6] C. Burgert, L. Rüschendorf: *Consistent risk measures for portfolio vectors*, Insurance: Math. Econ. 38/2 (2006) 289–297.
- [7] C. Castaing, M. Valadier: *Convex Analysis and Measurable Multifunctions*, Lecture Notes in Mathematics vol. 580, Springer, Berlin (1977).
- [8] C. Chen, G. Iyengar, C. C. Moallemi: *An axiomatic approach to systemic risk*, Management Science 59/6 (2013) 1373–1388.
- [9] C. Combari, M. Laghdir, L. Thibault: *Sous-différentiels de fonctions convexes composées*, Ann. Sci. Math. Québec 18/2 (1994) 119–148.
- [10] P. Delbaen: *Coherent risk measures on general probability spaces*, in: *Advances in Finance and Stochastics. Essays in Honour of Dieter Sondermann*, K. Sandmann et al. (eds.), Springer, Berlin (2002) 1–37.
- [11] D. Dentcheva, A. Ruszczyński: *Optimization with stochastic dominance constraints*, SIAM J. Optimization 14 (2003) 548–566.
- [12] D. Dentcheva, A. Ruszczyński: *Optimality and duality theory for stochastic optimization problems with nonlinear dominance constraints*, Math. Programming 99 (2004) 329–350.
- [13] D. Dentcheva, A. Ruszczyński: *Inverse stochastic dominance constraints and rank dependent expected utility theory*, Math. Programming B 108 (2006) 297–311.
- [14] D. Dentcheva, A. Ruszczyński: *Portfolio optimization under stochastic dominance constraints*, J. Banking Finance 30 (2006) 433–451.
- [15] D. Dentcheva, A. Ruszczyński: *Duality between coherent risk measures and stochastic dominance constraints in risk-averse optimization*, Pacific J. Optimization 4/3 (2008) 433–446.
- [16] D. Dentcheva, A. Ruszczyński: *Risk preferences on the space of quantile functions*, Math. Programming B 148/1-2 (2014) 181–200.
- [17] D. Dentcheva, A. Ruszczyński: *Risk forms: representation, disintegration, and application to partially observable two-stage systems*, Math. Programming B 181 (2020) 297–317.
- [18] D. Dentcheva, A. Ruszczyński: *Mini-batch risk forms*, SIAM J. Optimization 33/2 (2023) 615–637.
- [19] D. Dentcheva, A. Ruszczyński: *Risk-Averse Optimization and Control: Theory and Methods*, Springer Series in Operations Research and Financial Engineering, Springer, Cham (2024).

- [20] J. Diestel, J. Uhl: *Vector Measures*, American Mathematical Society, Providence (1977).
- [21] I. Ekeland, W. Schachermayer: *Law-invariant risk measures on $L_\infty(\mathbb{R}^d)$* , Statistics Risk Modeling 28/3 (2011) 195–225.
- [22] P. C. Fishburn: *Decision and Value Theory*, Publications in Operations Research Vol. 43, Operations Research Society of America, John Wiley and Sons, New York (1964).
- [23] P. C. Fishburn: *Continua of stochastic dominance relations for bounded probability distributions*, J. Math. Econ. 3 (1976) 295–311.
- [24] H. Föllmer, A. Schied: *Convex measures of risk and trading constraints*, Finance Stochastics 6 (2002) 429–447.
- [25] J. Hadar, W. R. Russell: *Rules for ordering uncertain prospects*, Amer. Econ. Review 59/1 (1969) 25–34.
- [26] G. Hanoch, H. Levy: *The efficiency analysis of choices involving risk*, Rev. Econ. Studies 36/3 (1969) 335–346.
- [27] A. D. Ioffe, V. L. Levin: *Subdifferentials of convex functions*, Trudy Mosk. Mat. Obsh. 26 (1972) 3–73.
- [28] L. V. Kantorovich, G. P. Akilov: *Functional Analysis*, 3rd ed., Nauka, Moscow (1984).
- [29] M. Kijima, M. Ohnishi: *Mean-risk analysis of risk aversion and wealth effects on optimal portfolios with multiple investment opportunities*, Ann. Oper. Res. 45 (1993) 147–163.
- [30] E. Kromer, L. Overbeck, K. Zilch: *Systemic risk measures on general measurable spaces*, Math. Methods Oper. Res. 84 (2016) 323–357.
- [31] E. L. Lehmann: *Testing multiparameter hypotheses*, Ann. Math. Statistics 23 (1952) 541–552.
- [32] J. Leitner: *A short note on second-order stochastic dominance preserving coherent risk measures*, Math. Finance 15 (2005) 649–651.
- [33] B. Lemaire: *Application of a subdifferential of a convex composite functional to optimal control in variational inequalities*, in: *Nondifferentiable Optimization: Motivations and Applications*, Proc. IIASA Workshop Sopron 1984, Lecture Notes in Economics and Mathematical Systems 255, Springer, Berlin (1985) 103–117.
- [34] V. L. Levin. *Convex Analysis in Spaces of Measurable Functions and its Applications in Economics (Russian)*, Nauka, Moskva (1985).
- [35] M. O. Lorenz: *Methods of measuring the concentration of wealth*, Pub. Amer. Stat. Assoc. 9/70 (1905) 209–219.
- [36] H. Mann, D. Whitney: *On a test of whether one of two random variables is stochastically larger than the other*, Ann. Math. Stat. 18 (1947) 50–60.
- [37] A. Müller, D. Stoyan; *Comparison Methods for Stochastic Models and Risks*, John Wiley and Sons, Chichester (2002).
- [38] W. Ogryczak, A. Ruszczyński: *From stochastic dominance to mean-risk models: semideviations as risk measures*, Eur. J. Oper. Res. 116 (1999) 33–50.
- [39] W. Ogryczak, A. Ruszczyński: *Dual stochastic dominance and related mean risk models*, SIAM J. Optimization 13 (2002) 60–78.

- [40] J. Quirk, R. Saposnik: *Admissibility and measurable utility functions*, Rev. Econ. Studies 29 (1962) 140–146.
- [41] R. T. Rockafellar: *Integrals which are convex functionals*, Pacific J. Math. 24/3 (1968) 525–539.
- [42] R. T. Rockafellar: *Integrals which are convex functionals. II*, Pacific J. Math. 39/2 (1971) 439–469.
- [43] M. Rothschild, J. E. Stiglitz: *Increasing risk. I: A definition*, in: *Uncertainty in Economics*, P. Diamond et al. (eds.), Academic Press, New York (1978) 99–121.
- [44] L. Rüschendorf: *Mathematical Risk Analysis*, Springer, Berlin (2013).
- [45] A. Ruszczyński, A. Shapiro: *Optimization of convex risk functions*, Math. Oper. Research 31 (2006) 433–452.
- [46] M. Shaked, J. G. Shanthikumar: *Stochastic Orders and Their Applications*, Academic Press, Boston (1994).
- [47] A. Shapiro, D. Dentcheva, A. Ruszczyński: *Lectures on Stochastic Programming: Modeling and Theory*, Society of Industrial and Applied Mathematics, Philadelphia (2021).
- [48] V. Strassen: *The existence of probability measures with given marginals*, Ann. Math. Statistics 38 (1965) 423–439.
- [49] M. Valadier: *Sous-différentiels d’une borne supérieure et d’une somme continue de fonctions convexes*, C. R. Acad. Sci. Paris, Sér. A 268 (1969) 39–42.
- [50] M. Valadier: *Sous-différentiabilité de fonctions convexes à valeurs dans un espace vectoriel ordonné*, Math. Scandinavica 30/1 (1972) 65–74.