Regular Castaing Representations of Multifunctions with Applications to Stochastic Programming*

Darinka Dentcheva
Institut für Mathematik,
Humboldt Universität Berlin, 10099 Berlin, Germany,
darina@mathematik.hu-berlin.de

May 29, 1998

Abstract

We consider set-valued mappings defined on a topological space with convex closed images in $\mathbb{R}^n$. The measurability of a multifunction is characterized by the existence of a Castaing representation for it: a countable set of measurable selections that pointwise fill up the graph of the multifunction. Our aim is to construct a Castaing representation which inherits the regularity properties of the multifunction. The construction uses Steiner points. A notion of a generalized Steiner point is introduced. A Castaing representation called regular is defined by using generalized Steiner selections. All selections are measurable, continuous, resp. Hölder-continuous, or directionally differentiable, if the multifunction has the corresponding properties. The results are applied to various multifunctions arising in stochastic programming. In particular, statements about the asymptotic behavior of measurable selections of solution sets via the delta-method are obtained.

Keywords: Steiner center, selections, Castaing representation, stochastic programs, random sets, delta-theorems

1 Introduction

Analysis of the behavior of multifunctions includes questions on existence of selections with some regularity properties. When measurability plays a role, one of the most celebrated results is the Castaing representation theorem ([6]). It is known (see [20]), that a closed-valued measurable multifunction in a Polish target space admits a measurable selection.

*Supported by the HSP III Program of Humboldt-University Berlin, Germany

Furthermore, for a multifunction F with nonempty closed values in a Polish target space (in our case this will be $I\!R^n$), we can choose a countable family of measurable selections $\{f_n\}$ that pointwise fills up the values of the multifunction:

$$\text{for each } x \in X,\ \ F(x) = \mathrm{cl}\left(\cup_{n=1}^{\infty} f_n(x)\right).$$

Such a countable family is called *a Castaing representation* for the multifunction. The existence of such a representation characterizes measurability.
Our aim is to construct a Castaing representation of a multifunction $F : X \rightrightarrows I\!R^n$, defined on a linear metric space X, which inherits its regularity properties.
An overview of the basic facts how selections inherit measurability, Lipschitz-continuity etc. is given in [3]. The reader can find there also a presentation of some special selections and their properties which are widely studied in the literature. Althogh the well-known Steiner selection inherits measurability and continuity properties of the multifunction, its definition does not provide tools for constructing a Castaing representation.
We shall generalize the definition of a Steiner center by using arbitrary probability measure with smooth density on the unit ball. We will obtain different Steiner points with respect to different measures which will be the basis of our construction. All generalized Steiner selections will preserve measurability, continuity, Hölder- or Lipschitz-continuity, and some kind of differentiability.
Several concepts of differentiability of set-valued mappings have been developed in the literature (see, e.g., [3], [4], [25]). We shall work with the notion semi-differentiability, which was introduced by Penot [23] and corresponds to the concept of tangential approximation due to Shapiro [32, 33]. Semi-differentiability plays an important role in the delta-method, which provides information about the asymptotic behavior of stochastic processes. In particular, mappings containing feasible and optimal solutions of stochastic programs are of this kind.
The existence of a differentiable selection has been treated in [13, 8, 10].
In [8] also another construction of a Castaing representation is developed suitable for applications to delta-theorems. The construction is based on metric projections and it is sufficiently good while working with the delta-method, but the selection of that Castaing representation do not preserve the regularity properties of the multifunction.

Our results have a specific application to stochastic programming. We shall demonstrate the existence of a regular Castaing representations for various multifunctions arising in stochastic programming.

2 Generalized Steiner Points

In this section, the notions of a generalized Steiner point for a convex compact set is introduced. The notion of Steiner center can be generalized also for some unbounded sets, as it is shown in [8]. We restrict our investigations to the case of compact sets in order to simplify the presentation, moreover, this corresponds to all applications we have in mind.

Let $C \subseteq I\!\!R^n$ be a compact convex set. Furthermore, let the Lebesgue measure of the unit ball $I\!\!B$ in $I\!\!R^n$ be denoted by $\mathcal{V}$ and its surface area (computed with the n-dimensional spherical Lebesgue measure) by $\mathcal{S}$, i.e.,

$$\mathcal{V} = \frac{\pi^{n/2}}{\Gamma(1+\frac{n}{2})}, \quad \mathcal{S} = n\mathcal{V} = \frac{2\pi^{n/2}}{\Gamma(\frac{n}{2})}.$$

Definition 2.1 *The Steiner center* $\mathrm{s}(C)$ *is defined in the following way:*

$$\mathrm{s}(C) = \frac{1}{\mathcal{V}} \int_{\Sigma} p\sigma(p, C)\ \omega(dp), \tag{1}$$

where Σ denotes the unit sphere in $I\!\!R^n$, ω is the Lebesgue measure on Σ, $\sigma(\cdot, C)$ is the support function of C.
Recall that the support function $\sigma(\cdot, C) : I\!\!R^n \to I\!\!R$ of a closed convex set $C \subseteq I\!\!R^n$ is defined by $\sigma(p, C) = \sup_{y \in C} \langle p, y \rangle$.
This point was first introduced by Steiner [36] in 1840 for a C^2-convex plane curve as the barycenter of the curvature measure. A definition using normalized isometry-invariant measure was introduced by Shepard [35]. The properties of the Steiner center have been widely investigated in the literature. We refer to the monograph [31], where the interested reader can find several facts and references on this topic.

It is easy to see that changing the measure in the formula above could easily lead to obtaining points that do not belong to the set C. However, there is another representation of the Steiner point, which we shall use. Following [3], we use the notations $\partial\sigma(p, C) = \{y \in C : \langle p, y \rangle = \sigma(p, C)\}$ for the subdifferential of the support function and $m(\partial\sigma(p, C))$ for the norm-minimal element in it. The Steiner center can be expressed equivalently as follows:

$$\mathrm{s}(C) = \frac{1}{\mathcal{V}} \int_{I\!\!B} m(\partial\sigma(p, C))dp. \tag{2}$$

Let μ denote the normalized Lebesgue measure on $I\!\!B$, i.e., $d\mu = \frac{dp}{\mathcal{V}}$. We define the set

$$\mathcal{M} = \{\alpha : \text{ probability measure on } I\!\!B \text{ having } C^1 - \text{density with respect to } \mu\}.$$

Definition 2.2 *A generalized Steiner center* $\mathrm{St}_\alpha(C)$ *of a compact convex set* $C \subseteq I\!\!R^n$ *with respect to the measure* $\alpha \in \mathcal{M}$ *is defined as follows:*

$$\mathrm{St}_\alpha(C) = \int_{I\!\!B} m(\partial\sigma(p, C))\alpha(dp). \tag{3}$$

It is well-known that:

1. $\mathrm{s}(C) \in C$ for all compact convex sets $C \subseteq I\!\!R^n$.

2. $\mathrm{s}(aA + bB) = a\mathrm{s}(A) + b\mathrm{s}(B)$ for any real number a and b and any compact convex sets A and B.

We shall show that this is true for the generalized Steiner points, too. Let $\nabla f(x)$ denote the gradient of f calculated at x. In order to show some regularity of the generalized Steiner points a representation using only the values of the support function instead of its subdifferential is of interest. The equivalence of the two representations (1) and (2) known for the Steiner center holds only for uniform measures and, therefore, we cannot simply change the measure in equation (1).

Theorem 2.3 *It holds for any convex compact set C and probability measure $\alpha \in \mathcal{M}$ with a density $\theta(\cdot)$:*

$$\mathrm{St}_\alpha(C) = \frac{1}{\mathcal{V}}\Big[\int\limits_{\Sigma} p\sigma(p,C)\theta(p)\omega(dp) - \int\limits_{I\!B} \sigma(p,C)\nabla\theta(p)dp\Big]. \tag{4}$$

The point $\mathrm{St}_\alpha(C)$ belongs to C and it holds $\mathrm{St}_\alpha(aA+bB) = a\mathrm{St}_\alpha(A) + b\mathrm{St}_\alpha(B)$ for any real numbers a and b and any compact convex sets A and B.

Proof: Consider the Moreau-Yosida approximation $\sigma_\lambda(p,C)$ of the support function $\sigma(p,C)$ of C. It is continuously differentiable and we may apply the Stoke's formula to the product $\sigma(\cdot,C)\theta(\cdot)$. We obtain:

$$\int\limits_{\Sigma} p\sigma_\lambda(p,C)\theta(p)\omega(dp) = \int\limits_{I\!B} \Big[\sigma_\lambda(p,C)\nabla\theta(p) + \nabla\sigma_\lambda(p,C)\theta(p)\Big]dp \tag{5}$$

We recall also that $\sigma_\lambda(\cdot,C)$ satisfies the inequalities

$$-\sup_{y\in C}\|y\| \le \inf_{p\in\Sigma}\sigma(p,C) \le \sigma_\lambda(p,C) \le \sigma(p,C) \le \sup_{y\in C}\|y\|$$

and it converges pointwise to $\sigma(\cdot,C)$. Therefore

$$\lim_{\lambda\to 0}\int\limits_{\Sigma} p\sigma_\lambda(p,C)\theta(p)\omega(dp) = \int\limits_{\Sigma} p\sigma(p,C)\theta(p)\omega(dp)$$

On the other hand, it is shown in [3] that $\nabla\sigma_\lambda(p,C) \in C$ and converges to $m(\partial\sigma(p,C))$. Thus, having in mind that the maps $\nabla\sigma_\lambda(\cdot,C)$ are measurable and bounded by $\sup\limits_{y\in C}\|y\|$, it holds:

$$\lim_{\lambda\to 0}\int\limits_{I\!B} \nabla\sigma_\lambda(p,C)\theta(p)dp = \int\limits_{I\!B} m(\partial\sigma(p,C))\theta(p)dp$$

We pass to the limit $\lambda \to 0$ in (5) and use that $\theta(p)dp = \mathcal{V}\alpha(dp)$. This implies the equivalent representation of the generalized Steiner point.
For each $a \in I\!R^n$ one has:

$$\langle a, \mathrm{St}_\alpha(C)\rangle = \int\limits_{I\!B}\langle a, m(\partial\sigma(p,C))\rangle\alpha(dp) \le \int\limits_{I\!B}\sigma(a,C)\alpha(dp) = \sigma(a,C)$$

The latter inequality proves that $\mathrm{St}_\alpha(C) \in C$.
The affine character of the mapping $C \mapsto \mathrm{St}_\alpha(C)$ is determined by the affine character of the integral and of the support function. Therefore, the last assertion of the theorem holds true, too. $\square$

3 Measurability and Castaing Representations

Let the space X be equipped with a σ-algebra $\mathcal{A}$. We use the following definition of measurability (see also [3], [7]):

Definition 3.1 *A mapping $f : (X, \mathcal{A}) \to \mathbb{R}^n$ is* measurable *if for any open set $C \subseteq \mathbb{R}^n$ the preimage $f^{-1}(C) = \{x \in X \;:\; f(x) \in C\}$ belongs to $\mathcal{A}$.*
A multifunction $F : (X, \mathcal{A}) \rightrightarrows \mathbb{R}^n$ is measurable *if for any open set $C \subseteq \mathbb{R}^n$ the preimages $F^{-1}(C) = \{x \in X \;:\; F(x) \cap C \neq \emptyset\} \in \mathcal{A}$.*

Recall that $f : (X, \mathcal{A}) \to \mathbb{R}^n$ is called *a measurable selection* of F if f is measurable and $f(x) \in F(x)$ almost surely. It is known (see [20]), that a closed-valued measurable multifunction in a Polish target space admits a measurable selection. Furthermore, for a multifunction F with nonempty closed values in a Polish target space, we can choose a Castaing representation of it: a countable family of measurable selections $\{f_n\}$ such that:

$$\text{for each } x \in X, \;\; F(x) = \mathrm{cl}\,\Big(\cup_{n=1}^{\infty} f_n(x) \Big).$$

The existence of such a representation characterizes measurability (cf., e.g. [7]).
Our aim is to construct Castaing representations of a multifunction $F : X \rightrightarrows \mathbb{R}^n$ with convex compact images, which preserves regularity properties of F. The construction will be based on generalized Steiner selections.

Lemma 3.2 *Let C be a convex compact set. The set of generalized Steiner points $D = \{\mathrm{St}_\alpha(C) : \alpha \in \mathcal{M}\}$ is dense in C.*

Proof: Suppose the opposite, i.e., $\mathrm{cl}\, D \neq C$. Given $y, z \in D$ obtained via the measures with densities θ_y, θ_z, any convex combination $\lambda y + (1-\lambda)z$ can be obtained by the measure with density $\lambda\theta_y + (1-\lambda)\theta_z$. Therefore, D is a convex set and $\mathrm{cl}D$ is convex and compact. Consequently, there exists a point $u \in \mathrm{ri}\, C$, where $\mathrm{ri}\, C$ is the relative interior of C, that does not belong to $\mathrm{cl}\, D$. Thus, if L is the linear subspace such that $\mathrm{ri}\, C \subset u + L$, then there exists a closed ball $B(u, \beta), \beta > 0$ such that $B(u, \beta) \cap (u + L) \subset C$ and $B(u, \beta) \cap \mathrm{cl}\, D = \emptyset$. Consequently, the latter two sets can be separated by a hyperplane $\langle a, y\rangle = \gamma$, i.e., $\langle a, y\rangle > \gamma$ for $y \in \mathrm{cl}\, D$ and $\langle a, z\rangle \leq \gamma$ for $z \in B(u, \beta)$. We consider the set S of all vectors $p \in \mathbb{R}^n$ such $\langle a, z\rangle \leq \gamma$ for $z \in \partial\sigma(p, C)$. Observe that this is a convex cone with nonempty interior since it contains a translation of the ball $B(u, \beta)$. Consequently, $S \cap \mathbb{B}$ has a nonempty interior, too. Therefore, there exists a smooth function $0 \leq \tilde{\theta}(\cdot) \leq 1$, which has a nonempty support included in $S \cap I\!B$. We define

$$\theta(p) = \frac{\tilde{\theta}(p)}{\int_{I\!B} \tilde{\theta}(p) dp}.$$

Consider the Steiner point $\bar{y}$ with respect to the measure with density θ. We have by construction of θ:

$$\langle a, \bar{y} \rangle = \int_{I\!B} \langle a, m(\partial\sigma(p, C)) \rangle \theta(p) dp \leq \int_{I\!B} \gamma\theta(p) dp = \gamma.$$

Thus $\bar{y} \notin \operatorname{cl} D$ and this is a contradiction to the definition of D. $\square$

Definition 3.3 *The function $f_\alpha : X \to I\!R^n$, defined by $f(x) = \mathrm{St}_\alpha(F(x))$ is said to be a generalized Steiner selection of F with respect to the measure α.*

Theorem 3.4 *Let $F : X \rightrightarrows I\!R^n$ be a measurable multifunction with nonempty compact convex images. Then F admits a representation by countably many generalized Steiner selections $\{f_n\}$ such that:*

$$\textit{for each } x \in X, \;\; F(x) = \operatorname{cl}\Big(\cup_{n=1}^{\infty} f_n(x) \Big).$$

Proof: We consider the set of functions $C_d^1 = \{f \in C^1(I\!B, I\!R_+) : \int_{I\!B} f \mu(dp) = 1\}$. By modification of standard arguments, it can be shown that there is a countable set, which is dense in C_d^1 with respect to the supremum-norm. We include the proof for completeness. Let $\{y_i\}$ build a countable dense set in $I\!B$. We define $U_i = \{y \in I\!B : d(y, y_i) < 1/m\}$. For each m, there are finitely many sets U_i, $i = 1, 2, \ldots, k(m)$ that cover $I\!B$. Let $\{g_i\}_{i=1}^{k(m)}$ be the smooth partition of the unity that corresponds to this covering. Recall that this means the following: $g_i \geq 0$, $\operatorname{supp} g_i \subset U_i$ and $\sum_{i=1}^{k(m)} g_i(y) = 1$ for all $y \in I\!B$. We set

$$\hat{\theta}_i = \sum_{i=1}^{k(m)} r_i g_i(y) \quad \text{and} \quad \theta_i = \hat{\theta}_i / \int_{I\!B} \hat{\theta}_i \mu(dp), \quad (m = 1, 2, \ldots),$$

where r_i are positive rational numbers.
We shall show that the latter set of functions is dense in C_d^1. Let some function $\theta \in C_d^1$ and $\varepsilon > 0$ be given. Let $\kappa = \max_{y \in I\!B} \theta(y)$ and $\varrho = \varepsilon/(2\kappa + 4)$. There is $\delta > 0$ such that $|\theta(y_1) - \theta(y_2)| < \varrho$ whenever $\|y_1 - y_2\| < \delta$. Consequently, taking $1/m < \delta$, we may choose rational numbers r_i such that $|\theta(y) - r_i| < 2\varrho$ for all $y \in U_i$. Consider the function

$$h(y) = \sum_{i=1}^{k(m)} r_i g_i(y).$$

We have the following estimations:

$$|\theta(y) - h(y)| = |\sum_{i=1}^{k(m)} \theta(y) g_i(y) - \sum_{i=1}^{k(m)} r_i g_i(y)| \leq \sum_{i=1}^{k(m)} |\theta(y) - r_i| g_i(y) \leq 2\varrho,$$

and

$$-2\varrho+1 \leq \int_{I\!B} h(y)\mu(dp) = \int_{I\!B} (h(y)-\theta(y))\mu(dp) + \int_{I\!B} \theta(y)\mu(dp) \leq 2\varrho+1.$$

The function $\hat{h}(\cdot) = h(\cdot)/\int_{I\!B} h(y)\mu(dp)$ belong to the considered set. It is a routine check to see that $|\theta(y)-\hat{h}(y)| < \varepsilon$ for ε small enough. This proves the density of the set $\{\theta_i\}$ in C_d^1.

Consider the probability measures $\{\alpha_i\}_{i=1}^\infty$ with densities $\{\theta_i\}_{i=1}^\infty$ on $I\!B$. We denote the Steiner selection with respect to the measure α_i by f_i. We shall show that the union of selections $\{f_i\}_{i=1}^\infty$ is the Castaing representation we are looking for.
Let a point $(x,y) \in \operatorname{graph} F$ and $\delta > 0$ be given. By virtue of the previous lemma, there is a measure $\alpha \in \mathcal{M}$ such that $\|\mathrm{St}_\alpha(F(x)) - y\| \leq \frac{1}{2}\delta$. Let θ be the density of this measure. Further we set $\kappa := \max_{y\in F(x)} \|y\|$. There exists a density θ_δ such that $\sup_{y\in I\!B} |\theta(y)-\theta_\delta(y)| \leq \frac{\delta}{2\kappa}$. Taking the Steiner point with respect to the measure α_δ with this density, we obtain

$$\begin{aligned} \|\mathrm{St}_\alpha(F(x)) - \mathrm{St}_{\alpha_\delta}(F(x))\| &\leq \|\int_{I\!B} m(\partial\sigma(p,C))(\theta(p)-\theta_\delta(p))\mu(dp)\| \\ &\leq \frac{\delta}{2\kappa}\kappa \int_{I\!B} \mu(dp) = \frac{1}{2}\delta \end{aligned}$$

Consequently, $\|\mathrm{St}_{\alpha_\delta}(F(x)) - y\| \leq \delta$ and this proves the assertion since δ is arbitrary. □

4 Regularity Properties of Multifunctions and Their Generalized Steiner Selections

The main goal of this section is to show that the representation constructed in Theorem 3.4 preserves regularity properties of the multifunction. We shall show that all selections are measurable, continuous, Hölder- or Lipschitz-continuous, or directionally differentiable whenever the multifunction is so.
Throughout the paper we denote the Hausdorff distance between two sets $A, B \subseteq I\!R^n$ by

$$d_{\mathrm{H}}(A,B) = \max\{e(A,B), e(B,A)\}, \text{ where } e(A,B) = \sup_{y\in A} d(y,B),$$

where $d(\cdot, A)$ denotes the distance function associated with a closed set $A \subseteq I\!R^n$.
Suppose that X is a metric space with a metric ρ. We shall use the following notions of continuity for multifunctions.
A multifunction F is called *continuous* at a point $\bar{x}$, if

$$\text{for all } \varepsilon, \text{ there is a } \delta > 0 \text{ such that: } d_{\mathrm{H}}(F(x), F(\bar{x})) \leq \varepsilon \text{ for all } x : \rho(\bar{x}, x) < \delta.$$

Furthermore, a multifunction is called *Hölder-continuous* of order k at $\bar{x} \in X$ if there exist a constant L and a neighborhood U of $\bar{x}$ such that

$$d_{\mathrm{H}}(F(x_1), F(x_2)) \leq L\rho(x_1,x_2)^k \text{ for all } x_1, x_2 \in U,$$

If $k = 1$ then the multifunction is called *Lipschitz-continuous* at this point.

A multifunction will be called *Hölder***-continuous* of order k at $\bar{x} \in X$, if there exist a constant L and a neighborhood U of $\bar{x}$ such that

$$d_{\mathrm{H}}(F(x), F(\bar{x})) \leq L\rho(x, \bar{x})^k \text{ for all } x \in U,$$

If $k = 1$ then such a multifunction is called *Lipschitz***-continuous* at that point.

¿From now on we assume that the multifunction F under consideration has nonempty compact convex images.

Theorem 4.1 *Let a multifunction F be continuous, resp. Hölder-continuous, or Hölder*-continuous of order k at a point $\bar{x}$ with a constant L. Then each generalized Steiner selection f_α is continuous, resp. Hölder-continuous, or Hölder*-continuous of order k at this point with a constant:*

$$\hat{L} = (n \max_{p \in \Sigma} \theta(p) + \max_{p \in I\!\!B} \|\nabla\theta(p)\|)L,$$

where θ is the density of the measure α.

Proof: Let us recall that for every $p \in \Sigma$ it holds

$$\text{for all A,B -nonempty, convex, compact sets } |\sigma(p, A) - \sigma(p, B)| \leq d_{\mathrm{H}}(A, B).$$

We consider a generalized Steiner selection f_α, where the measure α has a density θ. The following chain of inequalities holds true:

$$\begin{aligned}
&\|f_\alpha(x) - f_\alpha(\bar{x})\| = \\
&\|\tfrac{1}{\mathcal{V}}\Big[\int_\Sigma p\sigma(p, F(x))\theta(p)\omega(dp) - \int_{I\!\!B} \sigma(p, F(x))\nabla\theta(p)dp\Big] - \\
&\tfrac{1}{\mathcal{V}}\Big[\int_\Sigma p\sigma(p, F(\bar{x}))\theta(p)\omega(dp) - \int_{I\!\!B} \sigma(p, F(\bar{x}))\nabla\theta(p)dp\Big]\| \leq \\
&\tfrac{1}{\mathcal{V}}\Big[\|\int_\Sigma p\sigma(p, F(x))\theta(p)\omega(dp) - \int_\Sigma p\sigma(p, F(\bar{x}))\theta(p)\omega(dp)\| \\
&+\|\int_{I\!\!B} \sigma(p, F(x))\nabla\theta(p)dp - \int_{I\!\!B} \sigma(p, F(\bar{x}))\nabla\theta(p)dp\|\Big] \leq \\
&\tfrac{1}{\mathcal{V}}\Big[\int_\Sigma p|\sigma(p, F(x)) - \sigma(p, F(\bar{x}))|\theta(p)\omega(dp) + \int_{I\!\!B} |\sigma(p, F(x)) - \sigma(p, F(\bar{x}))|\nabla\theta(p)dp\Big]
\end{aligned}$$

Given a positive number ε, the continuity of F implies that there is a neighborhood $B(\bar{x}, \delta)$ such that $d_{\mathrm{H}}(F(x), F(\bar{x})) \leq \varepsilon$. Consequently, $|\sigma(p, F(x)) - \sigma(p, F(\bar{x}))| \leq \varepsilon$ and we obtain:

$$\begin{aligned}
&\|f_\alpha(x) - f_\alpha(\bar{x})\| \leq \tfrac{1}{\mathcal{V}}\Big[\int_\Sigma p\varepsilon\theta(p)\omega(dp) + \int_{I\!\!B} \varepsilon\nabla\theta(p)dp\Big] \leq \\
&\varepsilon\tfrac{1}{\mathcal{V}}(\max_{p\in\Sigma} \theta(p)\mathcal{S} + \max_{p\in I\!\!B} \|\nabla\theta(p)\|\mathcal{V}) = \varepsilon(n \max_{p\in\Sigma} \theta(p) + \max_{p\in I\!\!B} \|\nabla\theta(p)\|) = \varepsilon\hat{L}
\end{aligned}$$

This proves the continuity of the generalized Steiner selection.
In order to prove Hölder*-continuity with the constant $\hat{L}$ we only need to observe that we can substitute ε by $L\rho(x, \bar{x})^k$ in the above inequalities. Consequently any order of Hölder-continuity* will be preserved. In particular, Lipschitz*-continuity will be implied by the

Lipschitz*-continuity of the multifunction. The Hölder-continuity can be shown by similar chain of inequalities. We only need to substitute x_n by x_1 and $\bar{x}$ by x_2. $\square$

Results about existence of Lipschitz-continuous selections are given in [2, 3, 8, 10], including the case of $F(x)$ being unbounded sets. An interesting result on existence of a Lipschitz-continuous selection through any given point of the graph of the multifunction is contained in [10].
The Hölder-continuity of the generalized Steiner selections can be extended to multifunctions with unbounded images in the same way as [3], or [8]. We do not provide those consideration in order to concentrate on the main goal of this paper: the existence of a regular Castaing representation.

Remark 4.2 *The observations in the proof of the previous theorem show that the mapping $C \mapsto \mathrm{St}_\alpha(C)$, defined for all nonempty convex compact sets is Lipschitz-continuous. Consequently, all generalized Steiner selections are measurable whenever F is measurable as a composition $\mathrm{St}_\alpha \circ F$ of a measurable and a continuous mapping.*

Let us now discuss the relation between differentiability of a multifunction and its generalized Steiner selections. For the purpose of this investigation we need to assume that X is a linear metric space. We denote the graph of F by $\operatorname{graph} F$.

The following notions of differentiability of set-valued mappings will be used.

Definition 4.3 *A mapping $F : X \rightrightarrows I\!\!R^n$ is called directionally differentiable at a point $(\bar{x}, \bar{y}) \in \operatorname{graph} F$ in a direction $h \in X$, if the limit*

$$F'(\bar{x}, \bar{y}; h) = \lim_{t_n \downarrow 0} t_n^{-1}[F(\bar{x} + t_n h) - \bar{y}]$$

exists in the sense of Kuratowski-Painlevé convergence.

Recall that

$$\liminf_{n\to\infty} A_n = \{z \ : \ \limsup_{n\to\infty} d(z, A_n) = 0\}, \quad \limsup_{n\to\infty} A_n = \{z \ : \ \liminf_{n\to\infty} d(z, A_n) = 0\}$$

A sequence of closed sets $\{A_n\}$, $A_n \subseteq I\!\!R^n$ converges to some $A \subseteq I\!\!R^n$ in the sense of Kuratowski-Painlevé if and only if the sequence of distance functions converges pointwise (cf.[3]), i.e.,

$$A = \lim_{n\to\infty} A_n \text{ if and only if } d(y, A) = \lim_{n\to\infty} A_n$$

or, equivalently,

$$\liminf_{n\to\infty} A_n = A = \limsup_{n\to\infty} A_n$$

Definition 4.4 ([23]) *A mapping $F : X \rightrightarrows I\!\!R^n$ is called semi-differentiable at a point $(\bar{x}, \bar{y}) \in \operatorname{graph} F$, if the limit*

$$DF(\bar{x}, \bar{y}; h_0) = \lim_{t_n \downarrow 0, h_n \to h_0} t_n^{-1}[F(\bar{x} + t_n h_n) - \bar{y}]$$

exists for all $h_0 \in X$, in the sense of Kuratowski-Painlevé.

Various differentiability concepts are compared in [4, 25]. The semi-differentiability generates a derivative that forms a continuous multifunction with respect to the direction (see [4]), i.e., $\lim_{h_n \to h} DF(x, y; h_n) = DF(x, y; h)$ where the limit is taken with respect to the Kuratowski-Painlevé convergence. The derivatives above build some cone-approximation of the graph of the multifunction.
Continuous tangential approximations of set-valued mappings are considered also in [32, 33]. It has been shown in [4] that such tangential approximations, if they exist, coincide with the semiderivatives.

Theorem 4.5 ([8]) *Suppose that a multifunction $F : X \rightrightarrows I\!R^n$ is Lipschitzian* at all $x \in X$ and semi-differentiable at all points (x, y) such that $y \in \operatorname{bd} F_s(x)$. Here* bd *stands for the boundary of $F_s(x)$. Let $F(x)$ be polyhedra for all $x \in X$. Then the generalized Steiner selection f of F is Hadamard-directionally differentiable at all points $x \in X$. Moreover, the directional derivative of f is given by the following formula:*

$$f'(x; h) = \frac{1}{\mathcal{V}} \Big[\int_{\Sigma} p\sigma(p, DF_s(x, y_p; h))\theta(p)\omega(dp) - \int_{I\!B} \sigma(p, DF_s(x, y_p; h))\nabla\theta(p)dp \Big] , \qquad (6)$$

where $y_p \in \partial\sigma(p, F_s(x))$.

Differentiability properties of the classical Steiner selection are investigated in [8, 10, 13].

Corollary 4.6 *Let $F : X \rightrightarrows I\!R^n$ be Lipschitzian*, semi-differentiable at any point (x, y) with $y \in \operatorname{bd} F(x)$, and let $F(x)$ be polyhedra for all $x \in X$. Then F admits a Castaing representation by Hadamard-directionally differentiable Steiner selections $\{f_n\}$. Moreover, if F is semi-differentiable at $(x, f_n(x))$ then $f_n'(x; h) \in DF(x, f_n(x); h)$, for all $h \in X$.*

Proof: The statement follows from Theorem 3.4 and Theorem 4.5, having in mind, that all generalized Steiner selections are measurable by their continuity. The inclusion $f_n'(x; h) \in DF(x, f_n(x); h)$ in case F is semi-differentiable at $(x, f_n(x))$ follows from the definition of semiderivative. $\square$

Now, we would like to formulate a statement relating the directional differentiability of a set-valued mapping with the existence of a Castaing representation with directionally differentiable selections.

Corollary 4.7 *Suppose that a multifunction $F : X \rightrightarrows I\!R^n$ is directionally differentiable into a direction h at all points $(\bar{x}, y) \in \operatorname{graph} F : \ y \in \operatorname{bd} F(\bar{x})$, $F(x)$ are polyhedra for all $x \in X$, and it satisfies the following condition on Lipschitz behaviour:*
there exist constants $L > 0$ and $\delta > 0$ such that

$$(LB) \quad d_{\mathrm{H}}(F(\bar{x}), F(\bar{x} + th)) \leq Lt \qquad \textit{whenever } t \in (0, \delta).$$

Then F admits a Castaing representation by generalized Steiner selections $\{f_n\}$ which are directionally differentiable into the direction h at $\bar{x}$. Moreover, if F is directionally differentiable at $(\bar{x}, f_n(\bar{x}))$ then $f_n'(\bar{x}; h) \in F'(\bar{x}, f_n(\bar{x}); h)$, and the directional derivative satisfies formula (6). If F is Lipschitzian at $\bar{x}$ and directionally differentiable into all directions, then f_n are Hadamard-directionally differentiable at $\bar{x}$.

Proof: Under the assumption (LB), we follow the same line of argument as in the proof of Theorem 4.5, considering all limits for the fixed direction h. In this way, we obtain directional differentiability of all generalized Steiner selections into the direction h at the point $\bar{x}$. Under the assumption that F is Lipschitzian, the proof is the same as the previous Corollary. We have to take into account that directional differentiability together with Lipschitz-continuity imply semi-differentiability ([25]). The formula and the inclusion of the directional derivative follow analogously. $\square$

These statements are of interest when dealing with the delta-method as we shall see in the last section.

5 Feasible and Optimal Solutions of Stochastic Programs

In this sections we shall discuss some nontrivial applications for the existence of a regular Castaing representations. We apply the results of the previous section to mappings expressing optimal solutions of stochastic programs subjected to perturbations.

While working with stochastic optimization models, one assumes that the underlying probability measure is given. In practical situations this is rarely the case; one usually works with some approximations, or statistical estimates. These circumstances motivate the stability investigations of stochastic programs with respect to perturbations of the probability distributions. We shall consider two basic types of stochastic models: stochastic programs with recourse and stochastic programs with probabilistic constraints.

In order to discuss stability with respect to the probability measure, we need to work with a suitable metric space. Let (X, d) be a separable linear normed space and $\mathcal{P}(X)$ be the set of all Borel probability measures on X. We denote:

$$\begin{aligned} \mathcal{M}_p(X) &:= \Big\{\mu \in \mathcal{P}(X) : \textstyle\int_X d(x,y)\mu(dx) < \infty\Big\}, \\ \mathcal{D}(\mu,\nu) &:= \Big\{\eta \in \mathcal{P}(X\times X) : \eta\circ\pi_1^{-1} = \mu, \eta\circ\pi_2^{-1} = \nu\Big\}, \end{aligned}$$

using π_1 and π_2 as the canonical first and second projections, resp. The L_p-Wasserstein metric W_p, $(p \geq 1)$ is defined as follows:

$$W_p(\mu,\nu) := \Big[\inf\Big\{\int_{X\times X} d^p(x,y)\eta(dx,dy) :\ \eta\in\mathcal{D}(\mu,\nu)\Big\}\Big]^{1/p} \qquad \text{for all } \mu,\nu\in\mathcal{M}_p(X)$$

Furthermore, let $\|f\|_L$ be the usual Lipschitz-norm:

$$\|f\|_L = \|f\|_\infty + \sup_{x,y\in Y}\frac{|f(x)-f(y)|}{\|x-y\|}.$$

It is known (cf.[14]) that $(\mathcal{M}_p(X), W_p)$ is a metric space. Quantitative stability of stochastic programs with respect to perturbations of probability measures is investigated in [15, 16, 26, 27, 28, 29]. We shall utilize some of the results presented in those papers.

5.1 Stochastic Recourse Programs

Let us consider a two-stage stochastic program with linear recourse and random right-hand side:

$$\min\{g(x) + Q_\mu(Ax) : x \in C\} \tag{7}$$

$$Q_\mu(\chi) = \int_{I\!R} \tilde{Q}(\theta - \chi)\mu(d\theta), \tag{8}$$

$$\tilde{Q}(z) = \min\{q^\top y : Wy = z, y \geq 0\} \tag{9}$$

where $g : I\!R^n \to I\!R$ is a convex function, $C \subseteq I\!R^n$ is a non-empty closed convex set and μ is a Borel probability measure on $I\!R^m$. Furthermore, $q \in I\!R^s$ and A is an $n \times m$ matrix, W is an $s \times m$ matrix. We make use of the general assumptions (A1)-(A3), which are common in the literature, in order to make the problem well-defined.

(A1) $\quad W(I\!R^s_+) = I\!R^m$ $\quad$ (complete recourse),

(A2) $\quad M_D := \{u \in I\!R : W^\top u \leq q\} \neq \emptyset$ $\quad$ (dual feasibility),

(A3) $\quad \int_{I\!R} \|z\|\mu(\mathrm{d}z) < +\infty$ $\quad$ (finite first moment).

Having in mind linear programming theory, observe that (A1) and (A2) imply $\tilde{Q}(z)$ to be finite for all $z \in I\!R$. Due to (A3) also the integral of $\tilde{Q}(z)$ is finite ([17, 37]).
The model is derived from an optimization problem with uncertain data, where some statistical information about the random data is available. The decision x of the first stage has to be made *here and now* before observing some realization of θ. It is supposed to solve the problem:

$$\inf\{g(x) : x \in C, Ax = \theta\}$$

After observing a realization of θ we fix a second-stage decision y (called recourse action) in order to overcome the deviation $\theta - Ax$. The matrix W determines the rule to react and q the costs of our reaction. (A1) means that we are able to overcome any deviation. To choose y properly, we minimize its costs. To choose x properly, we minimizes the sum of the first-stage costs and the expected second-stage costs, caused by the corrective action y. Further details and fundamental properties of (two-stage) stochastic programs can be found in [17, 24, 37].
Two-stage stochastic programs hardly have a unique solution. This fact has motivated the attempt to avoid the assumption on the multifunction to be a singleton at certain points in our investigations. The next example give an impression on how restrictive this assumption is.

Example 5.1 ([28]) $g(x) = 0,\ A = (1,0),\ C = [0,1] \times [0,1],\ q = (1,1),\ W = (1,-1).$ *Let μ be the uniform distribution on [-1/2,1/2]. Then*

$$\begin{aligned}
&\psi(Q_\mu) \\
&= \operatorname{argmin}\{Q_\mu(Ax) :\ x \in [0,1] \times [0,1]\} \\
&= \operatorname{argmin}\{\int_{I\!R} |\omega - x_1| \mu(d\omega) :\ x \in [0,1] \times [0,1]\} \\
&= \{(0, x_2) : x_2 \in [0,1]\} = ker A \cap C
\end{aligned}$$

One can see that even for very simple examples the solution set is not a singleton. Under an assumption that Q_μ is a strictly, respectively strongly convex function we have the uniqueness of $A\psi(Q_\mu)$, but we cannot expect that $ker A = \{0\}$.
We consider the mapping assigning to each probability measure μ the set of optimal solutions of the problem 7, i.e.,

$$\psi(\mu) = \operatorname{argmin}\{g(x) + Q_\mu(Ax) :\ x \in C\}.$$

Proposition 5.2 *Let g be a convex quadratic function and C a polyhedron. Given $\mu \in \mathcal{M}_1(I\!R)$, let $\psi(\mu)$ be nonempty and let the function Q_μ be strongly convex on a open neighborhood of the set $A(\psi(\mu))$. Then the mapping ψ admits a Castaing representation of generalized Steiner selections which are Hölder*-continuous of order $1/2$ at the point $\mu \in (\mathcal{M}_1(X), W_1)$.*

Proof: According to Theorem 2.7, [26], under the assumption of the theorem, there are constants $L > 0$ and $\delta > 0$ such that:

$$d_{\mathrm{H}}(\psi(\mu), \psi(\nu)) \leq L\ W_1(\mu, \nu)^{1/2}$$

whenever $\nu \in \mathcal{M}_1(I\!R), W_1(\mu, \nu) < \delta$. Hence, we can apply Theorem 4.1 and conclude that each generalized Steiner selection is Hölder*-continuous of order $1/2$ at the point μ. Consequently, our construction of Theorem 3.4 yields a Castaing representation of ψ with the stated property. $\square$

We consider also general perturbation of the recourse function without referring to metrics for probability measures. The following setting of a perturbed problem is relevant:

$$\inf\{g(x) + Q(Ax) : x \in C\},$$

where $Q : I\!R^m \to I\!R$ is a convex function, considered to be a perturbation (resp. approximation) of the expected recourse function Q_μ. Resorting to convex perturbations is motivated by the fact that, given (A1) and (A2), Q_μ is convex for any probability measure with finite first moment (cf. [17, 37]).
Then the definition space of the mapping ψ changes to a functional space:

$$\psi(Q) = \operatorname{argmin}\{g(x) + Q(Ax) :\ x \in C\}.$$

Setting $Y = A(C)$, we consider two functional spaces. The space $C^{1,1}(Y, R)$ of all real-valued continuously differentiable functions with locally Lipschitz-continuous derivative, defined on Y, and the space $C^{0,1}(Y, I\!R)$ of all real valued locally Lipschitz-continuous functions, defined on Y. Both spaces are metrizable (cf. [9]). We suppose here that the set C is bounded and endow the space $C^{0,1}(Y, I\!R)$ with the usual Lipschitz-norm. We work with the corresponding norm-convergence in $C^{1,1}(Y, I\!R)$.
In the following, we always consider the restriction of the solution set mapping ψ to the cone of convex functions in one of the spaces above. One more piece of notation:

$$\phi(y) = \text{argmin}\{g(x) : x \in C, Ax = y\}, \quad (y \in Y).$$

Proposition 5.3 *Let $\psi(Q_\mu)$ be nonempty, and Q_μ be strongly convex on some open neighborhood of $A\psi(Q_\mu)$. Assume, in addition, that there is a constant $L > 0$ and a neighborhood U of $\bar{y}$ with $\bar{y} = A\psi(Q_\mu)$ such that*

$$(i) \qquad d(\phi(\bar{y}), \phi(y)) \leq L\|\bar{y} - y\|, \text{ for all } y \in Y \cap U.$$

Then ψ admits a Castaing representation by generalized Steiner selections which are Lipschitzian at the point $Q_\mu \in C^{0,1}(Y, I\!R)$ Moreover, if g is linear or convex quadratic and C a polyhedron, then the assumption (i) is satisfied.*

Proof: We refer here to Theorem 2.3 and Remark 2.4 in [9]. Under the assumption of the proposition, there are constants $\hat{L} > 0$ and $\delta > 0$ such that:

$$d_H(\psi(Q_\mu), \psi(Q)) \leq \hat{L}\|Q_\mu - Q\|_L$$

for any convex function $Q \in C^{0,1}(Y, I\!R)$ such that $\|Q - Q_\mu\|_L < \delta$, which means that the mapping ψ is locally Lipschitzian* at Q_μ. Consequently, according to Theorem 4.1 each generalized Steiner selection is Lipschitz*-continuous at that point. Applying the construction of a Castaing representation by Steiner selections according to Theorem 3.4 we accomplish the goal of the proposition. □

Similar result as Theorem 2.3 in [9] is shown in [28]. We can use it and obtain a similar statement as the above proposition. Here we have chosen to present only one of them to illustrate existence of a Castaing representation for the solutions set mapping, which has Lipschitz behavior.
Restricting the solution set mapping ψ to the cone K of convex functions in one of the spaces above, has an impact on the notions of differentiability. Considering the semiderivative at a point (Q_μ, x) in a certain direction $\bar{v}$, we assume that the arguments of ψ lie in K. Hence, $Q_\mu + tv \in K \; \forall v : v \to \bar{v}$. Consequently, v are elements of the closure of the radial tangent cone to K at the point Q_μ. We denote the radial tangent cone to K at the point Q_μ by

$$T_K^r(Q_\mu) = \{\lambda(Q - Q_\mu) : \; \lambda \geq 0, \; Q : I\!R^n \to I\!R \text{ convex}\}.$$

Proposition 5.4 *Assume $\psi(Q_\mu)$ to be non-empty. Let Q_μ be strictly convex on some open neighbourhoods of $A(\psi(Q_\mu))$ and twice continuously differentiable at $\chi_* : \ A(\psi(Q_\mu)) = \{\chi_*\}$. Let g be twice continuously differentiable and $v \in T_K^r(Q_\mu)$.*
Assume additionally that:

(i) there exists $\delta > 0$ such that for all $t \in [0, \delta)$
$d_H(\psi(Q_\mu), \psi(Q_\mu + tv)) = O(t)$,

(ii) there exists a neighborhood U of $\psi(Q_\mu)$ and constants $c > 0, \ \delta > 0$ such that:

$$g(x) + (Q_\mu + tv)(Ax) \geq \varphi(Q_\mu + tv) + \mathrm{cd}(x, \psi(Q_\mu + tv))^2, \quad \forall x \in U \cap C, \ \forall t \in [0, \delta].$$

(iii) for all $x \in \psi(Q_\mu)$, for all $y \in S(x)$, where

$$S(x) = \{y \in T_C(x) : \nabla g(x)y + \nabla Q_\mu(Ax)Ay = 0\}$$

the corresponding second-order set $S^2(x, y)$ is non-empty,

$$S^2(x, y) = \{z \in T_C^2(x, y) : \nabla g(x)z + \nabla Q_\mu(Ax)Az = 0\}$$

Then ψ is directionally differentiable at $(Q_\mu, x) \in \operatorname{graph} F$ into the direction v and

$$\begin{aligned}
\psi'(Q_\mu; x)(v) &= \lim_{t \to 0+} \frac{1}{t}(\psi(Q_\mu + tv) - x) \\
&= \operatorname{argmin} \left\{ \frac{1}{2}\langle \nabla^2 g(x)y, y\rangle + \frac{1}{2}\langle \nabla^2 Q_\mu(Ax)Ay, Ay\rangle + v'(Ax; Ay) : y \in S(x) \right\}.
\end{aligned}$$

Moreover, ψ admits a Castaing representation by Steiner selections f_i which are directionally differentiable at Q_μ into the direction v and it holds:

$$f_i'(Q_\mu; v) \in \psi'(Q_\mu; f_i(Q_\mu))(v).$$

The consideration at this place is in the space $C^{0,1}(Y, R)$.
The condition (i) requires a Lipschitz behavior for the solution set-mapping, (ii) is a second-order growth-condition, and (iii) imposes a restriction on the second order tangent set to C at the optimal points with respect to some elements of the tangent cone. These conditions are verified in [9] for the particular case where g is, in addition, a quadratic function and C is a polyhedral set.

Proof: The first statement of the proposition, i.e., the directional differentiability of ψ and the formula of the derivative, is proved by Theorem 4.1 in [9]. The second statement follows from the first by virtue of the Corollary 4.7. $\square$

Now we come to the semi-differentiability of the solution set-mapping and its consequences. We consider the restriction of ψ on the space $C^{1,1}(Y, R)$.

Proposition 5.5 *Assume $\psi(Q_\mu)$ to be non-empty, g be a quadratic function and C be a polyhedron. Let Q_μ be strongly convex on some open neighborhood of $A(\psi(Q_\mu))$ and twice continuously differentiable at $\chi_* :\ A(\psi(Q_\mu)) = \{\chi_*\}$. Let $x \in \psi(Q_\mu)$.*
Then ψ admits a Castaing representation by Steiner selections. All selections are Hadamard-directionally differentiable at (Q_μ, x), and the directional derivatives of the selections belong to the semiderivative of ψ, which is given by the formula of the previous proposition.

Proof: The semi-differentiability of ψ and the formula for the semiderivative, is proved by Theorem 4.7 in [9]. As in the proof of Proposition 5.3 we obtain that ψ is also Lipschitzian* at Q_μ. Thus, we can apply Corollary 4.6, which states the existence of the Castaing representation with the desired differentiability property. □

5.2 Stochastic Programs with Probabilistic Constraints

We shall be concerned with the following stochastic problem:

$$\min\{g(x) : x \in \mathbb{R}^n,\ \mu(\{z \in \mathbb{R}^s : x \in H(z)\}) \geq p\}, \tag{10}$$

where $g : \mathbb{R}^n \to \mathbb{R}$ is a convex function, $p \in (0,1)$ is a probability (or reliability) level, and $H : \mathbb{R}^s \to \mathbb{R}^n$ is a measurable mapping. It is assumed that the constraint $x \in H(z)$ is satisfied with a probability p.

Let $\mathcal{B}$ be a subset of the Borel σ-algebra on $\mathbb{R}^s$. The $\mathcal{B}$-discrepancy of two measures is defined by:

$$\alpha_{\mathcal{B}}(\mu, \nu) = \sup_{B \in \mathcal{B}} |\mu(B) - \nu(B)|, \quad \mu, \nu \in \mathcal{P}(\mathbb{R}^s).$$

The preimages $H^{-1}(x) = \{z \in \mathbb{R}^s : x \in H(z)\}$ are Bores sets because H is measurable. Consequently, we can use the subset $\mathcal{B}_H = \{H^{-1}(x), x \in \mathbb{R}^n\}$ as a subset of descrepancy and denote $\bar{\alpha} := \alpha_{\mathcal{B}_H}$.
A special case of the $\mathcal{B}_H$-discrepancy is the Kolmogorov distance on $\mathcal{P}(\mathbb{R}^s)$ defined by

$$\alpha(\mu, \nu) = \sup_{y \in \mathbb{R}^s} |F_\mu(y) - F_\nu(y)|, \quad \mu, \nu \in \mathcal{P}(\mathbb{R}^s).$$

where F_μ is the distribution function of μ.
In the setting of the previous section recourse problems preserve the same set of feasible points when the measure is subjected to perturbations. In the models with probabilistic constraint the solution changes because the feasible set changes when the measure is perturbed. Stability investigations of probabilistically constrained models are mainly concerned with changes that affect the feasible set. The feasible set can be expressed in the following way:

$$\{x \in \mathbb{R}^n : \mu(H^{-1}(x)) \geq p\} \tag{11}$$

Mostly investigated is the case of a mapping H given by linear inequalities, i.e.,

$$H(z) = \{x \in C\ :\ Ax \geq z\}, \quad z \in \mathbb{R}^s,$$

where A is an $s \times n$-matrix and $C \subseteq I\!R^n$ is a closed set, often supposed to be a polyhedron. Then we deal with the problem:

$$\min\{g(x) : x \in C, \; F_\mu(Ax) \geq p\}, \tag{12}$$

where F_μ is the distribution function of the probability measure $\mu \in \mathcal{P}(I\!R^s)$.
We assume μ to be r-concave for some $r \in (-\infty, 0)$. Recall that r-concavity is introduced in the following way. Let the generalized mean function m_r be defined on $I\!R_+ \times I\!R_+ \times [0,1]$ as:

$$m_r(a,b,\lambda) = \begin{cases} (\lambda a^r + (1-\lambda)b^r)^{1/r} & \text{if } r \neq 0, ab > 0 \\ 0 & \text{if } r \in (-\infty, 0), ab = 0 \\ a^\lambda b^{1-\lambda} & \text{if } r = 0 \\ \max\{a,b\} & \text{if } r = \infty \\ \min\{a,b\} & \text{if } r = -\infty \end{cases}$$

The measure $\mu \in \mathcal{P}(I\!R^s)$ is called *r-concave*, if the inequality $\mu(\lambda B_1 + (1-\lambda)B_2) \geq m_r(\mu(B_1), \mu(B_2), \lambda)$ holds for all $\lambda \in [0,1]$ and all Borel subsets B_1, B_2 of $I\!R^s$ such that $\lambda B_1 + (1-\lambda)B_2$ is a Borel set.
Due to r-concavity of μ, the problem (12) represents a convex program.
We shall consider the following mapping $\Phi : \mathcal{P}(I\!R^s) \times (0,1) \to I\!R^n$ defined by setting

$$\Phi(\mu, p) := \{x \in C : p - F_\mu(Ax) \leq 0\}.$$

Proposition 5.6 *Assume that μ is r-concave and C is a convex compact set. Suppose that the mapping $\Phi(\mu, \cdot)$ is Lipschitzian at p_0. Then Φ has a Castaing representation by generalized Steiner selection f_i such that there exist constants $\delta > 0$ and $L_i > 0$, and it holds:*

$$|f_i(\nu, p_0) - f_i(\mu, p_0)| \leq L_i \bar{\alpha}(\nu, \mu) \tag{13}$$

whenever $\bar{\alpha}(\nu, \mu) \leq \delta$.

Proof: The set of feasible points is convex and compact under the assumptions of the proposition. Hence, the Steiner points are well-defined. In Proposition 5.3, [27] a kind of pseudo-Lipschitzian behavior is shown for Φ under local assumptions on $\Phi(\mu, \cdot)$. Applying this result we obtain that

for all $x \in \Phi(\mu, p_0)$ there is a neigbourhood V_x and $\delta_x > 0$, $L_x > O$ such that:

$$d_{\mathrm{H}}(\Phi(\nu, p_0) \cap V_x, \Phi(\mu, p_0) \cap V_x) \leq L_x \bar{\alpha}(\nu, \mu) \text{ for all } \bar{\alpha}(\nu, \mu) \leq \delta_x$$

The set $\Phi(\mu, p_0)$ is compact, therefore, we can choose a finite number of those neigbourhoods that cover the whole feasible set $\Phi(\mu, p_0)$. Let us denote these neighbourhoods by $V_1, V_2, \ldots, V_k$, and the corresponding constants by $\delta_1, \delta_2, \ldots, \delta_k$, resp. $\bar{L}_1, \bar{L}_2, \ldots, \bar{L}_k$. We set $L = \max_i \bar{L}_i$ and $\delta = \min_i \delta_i$ for $i = 1, \ldots k$. Then for each $x \in \Phi(\mu, p_0)$ let $x \in V_j$ for some $j \in \{1, 2, \ldots, k\}$. We have:

$$d(x, \Phi(\nu, p_0)) \leq d(x, \Phi(\nu, p_0) \cap V_j) \leq d_{\mathrm{H}}(\Phi(\nu, p_0) \cap V_j, \Phi(\mu, p_0) \cap V_j) \leq \bar{L} \bar{\alpha}(\nu, \mu)$$

whenever $\bar{\alpha}(\nu,\mu) \le \delta$. In the same way we obtain, that for all $x \in \Phi(\nu, p_0)$ it holds:

$$d(x, \Phi(\mu, p_0)) \le \bar{L}\bar{\alpha}(\nu,\mu)$$

whenever $\bar{\alpha}(\nu,\mu) \le \delta$. The latter two inequalities imply that

$$d_{\mathrm{H}}(\Phi(\nu, p_0), \Phi(\mu, p_0)) \le \bar{L}\bar{\alpha}(\nu,\mu)$$

whenever $\bar{\alpha}(\nu,\mu) \le \delta$. Then, following the proof of Theorem 4.1, we can show that the relation (13) is satisfied for each generalized Steiner selection. Applying our usual technique of Theorem 3.4 we obtain the assertion. □

Determining the probability level p is a significant modeling decision. Therefore, it is natural to investigate changes of the feasible set when this level changes.

Proposition 5.7 *Let μ be r-concave and its distribution function F_μ be locally Lipschitzian. Let, furthermore, p_0 be a given probability level, C be a convex compact set Assume that for all $x \in \Phi(\mu, p_0)$ it holds that if $F_\mu(Ax) = p_0$, then the Clarke subdifferential of $F_\mu(A\cdot)$ at x and the normal cone to C at x have an empty intersection.*
Then $\Phi(\mu,\cdot)$ has a Castaing representation by generalized Steiner selections which are Lipschitzian at p_0.

Proof: The set of feasible points is convex and compact under the assumptions of the proposition. Therefore, the Steiner points are well-defined. Furthermore, we can apply Proposition 2.1 in [29] and obtain, that $\Phi(\mu,\cdot)$ is pseudo-Lipschizian at (x, p_0) for any $x \in \Phi(\mu, p_0)$. Since the images $\Phi(\mu, p)$ are compact, it follows as in the proof of the previous proposition that $\Phi(\mu,\cdot)$ is Lipschizian at those points. Consequently, according to Theorem 4.1 each generalized Steiner selection is Lipschitz-continuous at p_0. Applying the construction of a Castaing representation by Steiner selections according to Theorem 3.4 we accomplish our goal. □

Now, we focus our attention to sets of optimal solutions. Following the notations of the previous section, we understand that $\psi(\mu)$ designate the set of global solutions to 12, and $\psi_U(\nu)$ refers to the localized solution set of this problem, where $\nu \in \mathcal{P}(I\!R^s)$ is a perturbation of μ and $U \subseteq I\!R^n$ is a neighborhood of $\psi(\mu)$.

Proposition 5.8 *Assume that*

(i) $\psi(\mu)$ *is nonempty and bounded;*

(ii) $\psi(\mu) \cap \operatorname{argmin}\{g(x) : x \in C\} = \emptyset$*;*

(iii) there is $\bar{x} \in C : F_\mu(A\bar{x}) > p$ *(Slater condition);*

(iv) F_μ^r *is strongly convex on some open convex neighborhood* V *of* $A\psi(\mu)$*, where* $r \in (-\infty, 0)$ *is chosen such that* μ *is* r*-concave.*

Then there exist a neighbourhood U of $\psi(\mu)$ and $\delta > 0$ such that setting $\hat{\psi} : \mathcal{U} \to I\!R^n$ as $\hat{\psi}(\nu) = \psi_U(\nu)$, where $\mathcal{U} = \{\nu \in \mathcal{P}(I\!R^s) : \alpha(\mu, \nu) < \delta\}$, it holds that the mapping $\hat{\psi}$ admits a Castaing representation by Steiner selections which are Hölder-continuous of order $1/2$ at μ.*

Proof: We apply Theorem 4.3 in [16]. Under the assumption of the proposition, there are constants $L > 0$ such that:

$$d_{\mathrm{H}}(\psi(\mu), \psi_U(\nu)) \leq L\alpha(\mu, \nu)^{1/2} \tag{14}$$

for any probability measure $\nu \in \mathcal{U}$. Using the notation $\hat{\psi}$ for the restriction of the solution set mapping to the mapping of local minimizers the above inequality means that $\hat{\psi}$ is locally Hölder*-continuous of order $1/2$ at μ. Consequently, according to Theorem 4.1 each generalized Steiner selection is locally Hölder*-continuous of order $1/2$ at that point. Applying the construction of a Castaing representation by Steiner selections we obtain the result. $\square$

The assumptions of the above proposition are commented in [15] and illustrated by examples. Condition (i) is satisfied, for example, if C is a polytope. The conditions (ii) and (iii) mean that the probability level p is not chosen too low and too high, respectively.From the modeling point of view both conditions show the significance of the choice of the reliability level p. Assumption (iv) is decisive for obtaining a growth condition of the objective function around the original solution set.
As a conclusion, following [15], we formulate a large deviation result for the selections of the constructed Castaing representation when estimating μ by empirical measures. Let $\xi_1, \xi_2, \ldots, \xi_n, \ldots$ be independent identically distributed $I\!R^s$-valued random variables having common distribution μ, and let $\mu_n = \frac{1}{n}\sum_{i=1}^{n} \delta_{\xi_i}$ denote the empirical measure of $\xi_1, \xi_2, \ldots, \xi_n$.

Corollary 5.9 *Under the conditions of the previous proposition, let L and δ be the constants involved in the inequality (14) and the statement. Let $\hat{L}_j = (n \max_{p\in I\!B} \theta_j(p) + \max_{p\in I\!B} \|\nabla\theta_j(p)\|)L$, where θ_j designates the density of the j-th measure applied to calculate the generalized Steiner points.*
Then for all selection f_j of the Castaing representation of $\hat{\psi}$ and all $\varepsilon > 0$ it holds:

$$\limsup_{n\to\infty} \frac{1}{n} \log P(\|f_j(\mu) - f_j(\mu_n)\| \geq \varepsilon) \leq -2\min\{\delta^2, \varepsilon^4 \hat{L}_j^{-4}\}.$$

Proof: According to Theorem 4.1, the generalized Steiner selections f_j will have a constant $\hat{L}_j = (n \max_{p\in I\!B} \theta_j(p) + \max_{p\in I\!B} \|\nabla\theta_j(p)\|)L$, of Hölder*-continuity. The assertion follows from Corollary 4.29 in [15], the construction of the Castaing representation and Theorem 4.1.

6 Asymptotic Behavior of Random Sets

One way to obtain information about the asymptotic behavior of random elements is the so-called delta-method. Delta-theorems are concerned with the asymptotic distribution of functions of random elements, when those elements satisfy a central limit formula:

Theorem 6.1 ([34]) *Let $f : (X, \mathcal{B}(X)) \to I\!R^n$ be measurable and Hadamard-directionally differentiable at some point $\bar{x} \in X$. Suppose that X is a Banach space and $t_n(x_n - \bar{x})$ are some random elements of X converging in distribution to some element h, written*

$$t_n^{-1}(x_n - \bar{x}) \xrightarrow{D} h,$$

while $t_n \downarrow 0$ and h is a random element in some separable subspace of X. Then

$$t_n^{-1}(f(x_n) - f(\bar{x})) \xrightarrow{D} f'(\bar{x}; h).$$

Here $\xrightarrow{D}$ denotes convergence in distribution.
Recall that convergence in distribution of a sequence of random elements $x_n : (\Omega, \mathcal{A}, P) \to X$, means the weak* convergence of the measures $\mu_n = P \circ x_n^{-1}$ that these elements induce on the space X. A sequence of probability measures μ_n on a metric space X weakly* converges to μ (cf.[5]) if

$$\lim_{n\to\infty} \int g(x)\ \mu_n(dx) = \int g(x)\ \mu(dx)$$

for all bounded continuous functions $g : X \to I\!R$.
Convergence in distribution of set-valued mappings is considered in [30]. The first generalized delta-theorem for set-valued mappings was formulated by King [18].

Theorem 6.2 ([18]) *Let $F : (X, \mathcal{B}(X)) \rightrightarrows I\!R^n$ be closed-valued measurable multifunction defined on a separable complete metric space X. Suppose that x_n satisfy a generalized central limit formula with limit $\bar{x}$, i.e., there is a sequence $\{t_n\}, t_n \geq 0$ monotonically decreasing to 0 and a limit element h such that*

$$t_n^{-1}(x_n - \bar{x}) \xrightarrow{D} h$$

as random variables in X.
Assume, additionally, that F is almost surely semi-differentiable at $(\bar{x}, \bar{y})$ for some $\bar{y} \in F(\bar{x})$ with respect to the measure μ induced by h. Then $F(x_n)$ satisfy the generalized central limit formula

$$t_n^{-1}(F(x_n) - \bar{y}) \xrightarrow{D} DF(\bar{x}, \bar{y}; h)$$

as random closed sets in $I\!R^n$ or, equivalently,

$$d(\cdot, t_n^{-1}[F(x_n) - \bar{y}]) \xrightarrow{D} d(\cdot, DF(\bar{x}, \bar{y}; h))$$

as stochastic processes on $I\!R^n$.

Here semi-differentiability almost surely means that the convergence of the differential quotients holds for all directions, except for a set of μ-measure 0.

In general, the distribution of a random set does not determine the distributions of its measurable selections (cf, e.g. [1]). The results of this section will contribute to the investigations of this matter.

Corollary 6.3 *Under the assumptions of Theorem 3.4, assume that the random elements $x_n \in X$ satisfy a generalized central limit formula with limit $\bar{x}$, i.e.,*

$$t_n^{-1}(x_n - \bar{x}) \xrightarrow{D} h$$

as random variables in X, where $t_n \downarrow 0$.
Then for any point $\bar{y} \in F(\bar{x})$, the random sets $F(x_n)$ satisfy the generalized central limit formula

$$t_n^{-1}(F(x_n) - \bar{y}) \xrightarrow{D} DF(\bar{x}, \bar{y}; h)$$

and F admits a Castaing representation $\{f_k\}$ by generalized Steiner selections such that all f_k satisfy the generalized central limit formula

$$t_n^{-1}[f_k(x_n) - f_k(\bar{x})] \xrightarrow{D} f_k'(\bar{x}; h) \in DF(\bar{x}, f_k(\bar{x}); h).$$

Proof: The proof follows from the Theorem 6.2, Theorem 4.6 and Theorem 6.1. □

Let us return again to the solution set mapping of the recourse problem, which assigns to each approximation $Q \in C^{1,1}(Y, R)$ of the recourse function the set of optimal solution of the approximate problem.
Supposed we have some approximations (resp. estimates) $Q_n,\ n = 1, 2, \ldots$ of Q_μ that satisfy a generalized central limit formula in the above functional space. The application of our investigations lead to the following consequences for the delta-method:

Corollary 6.4 *Under the conditions of the previous theorem suppose that $Q_n,\ n = 1, 2, \ldots$ satisfy the functional central limit formula*

$$t_n^{-1}[Q_n - Q_\mu] \xrightarrow{D} \zeta \quad in \quad C^{1,1}(D, R)$$

for some monotonically decreasing sequence $t_n \downarrow 0$. Given a point $\bar{x} \in \psi(Q_\mu)$, then ψ satisfy the generalized central limit formula

$$t_n^{-1}[\psi(Q_n) - \bar{x}] \xrightarrow{D} D\psi(Q_\mu, \bar{x}; \zeta)$$

as random sets in $\mathcal{F}(I\!R^n)$.
Moreover, ψ admits a Castaing representation $\{f_i\}_{i=1}^{\infty}$ of Steiner selections such that all f_i satisfy the central limit formula

$$t_n^{-1}[f_i(Q_n) - f_i(Q_\mu)] \xrightarrow{D} f_i'(Q_\mu; \zeta) \in D\psi(Q_\mu, f_i(Q_\mu); \zeta)$$

as random variables on $I\!R^n$.

Proof: The assertion follows from the previous Corollary and Proposition 5.5. □

Investigating the asymptotic behavior of solution sets of stochastic programs is beyond the scope of this paper. The last statements have been included for the sake of giving an application of the results of this paper and yielding non-trivial statements. For investigations on the asymptotic behavior of stochastic programs the interested reader is referred to [12, 22, 19, 34] and the references therein.

Let us mention some of the results published on convergence in distribution of measurable selections of multifunctions. Interesting results are given in [1] by Artstein in a different setting. Relevant results are given by King [18] and Lachout [21]. In Theorem 4.3 in [18] a generalized central limit formula for all measurable selections is established under the assumption that the multifunction is upper Lipschitzian, $F(\bar{x}) = \bar{y}$, and $DF(\bar{x}, \bar{y}; h)$ is single-valued almost everywhere. In [21], the values $F(x)$ are supposed to be compact and $F(\bar{x}) = \bar{y}$ to be a singleton. The statement is that the measurable selections f of F do not satisfy the central limit formula themselves, but there are subsequences for which the formula holds. Those assumptions, in particular the assumption about $F(x)$ being singleton, are too strong for the applications we were aiming at. As it was mentioned stochastic programs have very seldom unique solutions and, therefore, we are interested in statements that are applicable to solution sets.

References

[1] Z. Artstein, Distributions of random sets and random selections, *Israel J. Mathematics* 46 (1984) 313-324

[2] J.-P. Aubin and A.Cellina, Differential Inclusions, Springer, Grundlehren der math. Wiss., 1984, Vol.264.

[3] J.-P. Aubin and H. Frankowska, Set-Valued Analysis, Birkhäuser, Boston 1990.

[4] A. Auslender and R. Cominetti, A comparative study of multifunction differentiability with applications in mathematical programming, *Mathematics of Operations Research*, 16 (1991), 240-258.

[5] P.Billingsley, Convergence of probability measures, Wiley, New York, 1968.

[6] Ch. Castaing, Sur les multi-application measurables, *Inf. Rech. Op.* 1 (1967) 91-126

[7] Ch. Castaing and M. Valadier Convex analysis and measurable multifunctions, Springer, Lecture Notes in Math., 1977, Vol.580.

[8] D.Dentcheva, Differentiable Selections and Castaing Representations of Multifunctions, to appear in *Journal of Math. Analysis and Applications.*

[9] D.Dentcheva and W.Römisch, Differential stability of two-stage stochastic programs, Preprint no.96-36 of the Institute of Mathematics, Humboldt University, Berlin, 1996

[10] G.Dommisch, On the existence of Lipschitz-continuous and differentiable selections for multifunctions, *in* J.Guddat et al.(eds.): Parametric Optimization and related topics, Math. Res. Vol. 35, Akademie-Verlag Berlin, 1987, 60 - 73.

[11] R.M. Dudley, Probabilities and Metrics, Lecture Notes Series No.45 (Aarhus Universitet, Aarhus, 1976).

[12] J. Dupačová and R.J.-B. Wets, Asymptotic behaviour of statistical estimators and of optimal solutions of stochastic optimization, *Ann. Statist.* 16 (1988) 1517-1549.

[13] S. Gautier and R. Morchadi, A selection of convex-compact-valued multifunctions with remarkable properties: the Steiner selection, *Numer. Funct. Anal. and Optimiz.* 13(5& 6) (1992) 513-522

[14] C.R. Givens ans R.M. Shortt, A class of Wasserstein metrics for probability distributions, Michigan Mathematical Journal 31 (1984) 231-240.

[15] R.Henrion, The approximate subdifferential and parametric optimization, Habilitationschrift, Humboldt-Universität zu Berlin, Berlin, Germany, 1997.

[16] R.Henrion and W.Römisch, Metric regularity and quantitative stability in stochastic programs with probabilistic constraints, *Mathematical Programming*, to appear.

[17] P. Kall, Stochastic Linear Programming, Springer-Verlag, Berlin 1976.

[18] A.J.King, Generalized delta theorems for multivalued mappings and measurable selections, *Mathematics of Operations Research* 14 (1989) 4, 720-736

[19] A.J.King and R.T.Rockafellar, Asymptotic theory for solutions in statistical estimation and stochastic programming, *Mathematics of Operations Research* 18 (1993), 148-162.

[20] K. Kuratowski and C. Ryll-Nardzewski, A general theorem on selectors, *Bull. Acad. Pol. Sc.* 13, 1965, 397-403.

[21] P. Lachout, On multifunction transforms of probability measures, *Annals of Operations Research*, 56 (1995) 241-250.

[22] G.Pflug, Asymptotic stochastic programs, *Mathematics of Operations Research*, 20 (1996) 4, 769-789

[23] J.-P. Penot, Differentiability of relations and differential stability of perturbed optimization problems, *SIAM J. Control and Optimization* 22 (1984) 4, 529-551

[24] A. Prekopa Stochastic Programming, Kluwer, Math. and its Applications, Vol. 324, 1995

[25] R. T. Rockafellar, Proto-differentiability of set-valued mappings and its application in optimization, *in:* H. Attouch et al. (eds.), Analyse Non Lineaire, Gauthier-Villars, Paris, 1989, 449-482.

[26] W.Römisch and R.Schultz, Stability analysis for stochastic programs, *Annals of Operations research* 30 (1991) 241-266.

[27] W.Römisch and R.Schultz, Distribution sensitivity in stochastic programming, *Mathematical Programming* 50 (1991) 197-226.

[28] W.Römisch and R.Schultz, Lipschitz stability for stochastic programs with complete recourse, *SIAM J. Optimization* 6 (1996) 2, 531-547.

[29] W.Römisch and R.Schultz, Distribution sensitivity for certain classes of chance constrained models with application to power dispach, *Journal of Optimization Theory and Applications* 71 (1991) 569-588.

[30] G.Salinetti and R.J.-B.Wets, On the convergence in distribution of measurable multifunctions (random sets), normal integrands, stochastic processes and stochastic infima, *Mathematics of Operations Research* 11 (1986) 3, 385-419.

[31] R.Schneider, Convex bodies: The Brunn-Minkowski theory, Encyclopedia of Mathematics and its Applications, Cambridge University Press, 1993.

[32] A. Shapiro, Existence and differentiability of metric projections in Hilbert spaces, *SIAM J. Optimization* 4 (1994) 1, 130-141.

[33] A. Shapiro, On differentiability of metric projections in $\mathbb{R}^n$: Boundary Case, *Proc.Amer. Math. Sic.* 99, 123-128.

[34] A. Shapiro Asymptotic analysis of stochastic programs, *Annals of Operations Research* 30 (1991), 169-186.

[35] G.C.Shephard, A uniqueness theorem for the Steiner point of a convex region, *J. London Mat. Soc.* 43 (1968) 439-444.

[36] J.Steiner, Von dem Krümmungsschwerpunkte ebener Kurven, *Zeitschrift für Reine Angew. Math.* 21 (1840) 33-63, 101-122.

[37] R.J.-B. Wets, Stochastic programs with fixed recourse: the equivalent deterministic program, *SIAM Review* 16 (1974), 309-339.