SIAM J. OPTIM.
Vol. 29, No. 3, pp. 2179–2210

GENERALIZED DIFFERENTIATION OF PROBABILITY FUNCTIONS ACTING ON AN INFINITE SYSTEM OF CONSTRAINTS*

WIM VAN ACKOOIJ† AND PEDRO PÉREZ-AROS‡

Abstract. In decision-making problems under uncertainty, probability constraints are a valuable tool for expressing safety of decisions. They result from taking the probability measure of a given set of random inequalities depending on the decision vector. When uncertainty results from two different sources, unequally known, it becomes intuitively appealing to consider the probability of the worst case with respect to the part of the uncertainty vector for which little information was available. This and other models lead to probability functions acting on infinite systems of constraints. In this paper we study generalized (sub)differentiation of such probability functions. We also develop explicit formulae for the subdifferentials that should prove useful in first-order methods or in formulating optimality conditions.

Key words. stochastic optimization, probability functions, probabilistic constraints, generalized differentiation, chance constraints

AMS subject classification. 90C15

DOI. 10.1137/18M1181262

1. Introduction.

A probability constraint is built up from several ingredients. One requires for t in some index set T a family of maps $g_t : X \times \mathbb{R}^m \to \mathbb{R}$, where X is a (reflexive) Banach space, a random vector $\xi \in \mathbb{R}^m$ defined on an appropriate probability space, and a user-defined safety level $p \in [0, 1]$. The probability constraint then reads

$$\varphi(x) := \mathbb{P}(g_t(x, \xi) \leq 0 \ \forall t \in T) \geq p, \tag{1.1}$$

where $\varphi : X \to [0, 1]$ is a map from X to $[0, 1]$. The interpretation of (1.1) is simple: the decision vector x is required to be such that the random inequality system $g_t(x, \xi) \leq 0$, $t \in T$, holds with probability at least p. In this work, the map $g_t \colon X \times \mathbb{R}^n \to \mathbb{R}$ is a continuously differentiable function, convex in the second argument, and the set $t \in T$ is some index set, whereas ξ will be an elliptical symmetric random vector. Even though usually the index set T is a finite set, there are compelling reasons to consider arbitrary index sets T. Such situations arise for instance in "probust" constraints (see, e.g., [1]) or hybrid robust/chance constraints (see, e.g., [67] and also [26]). These "probust" constraints arise whenever in a regular optimization problem one wishes to give a meaning to the random inequality system $g_t(x, \xi) \leq 0$, $t \in T$ (with T finite), and the decision x has to be taken prior to observing ξ. If ξ consists of two distinct sources of uncertainty, say $\xi = (\xi_1, \xi_2)$, and it is only reasonable to assume availability of distributional information for ξ_1, but little is known for ξ_2, then it is reasonable to combine ideas from robust optimization with those from probabilistic

*Received by the editors April 16, 2018; accepted for publication (in revised form) June 4, 2019; published electronically September 5, 2019.
https://doi.org/10.1137/18M1181262
Funding: The second author was partially supported by CONICYT grant Fondecyt 1190110.
†EDF R&D, Department OSIRIS, 7 Boulevard Gaspard Monge, 91120 Palaiseau, France (wim.van.ackooij@gmail.com).
‡Instituto de Ciencias de la Ingeniería, Universidad de O'Higgins, Av. Libertador Bernardo O'Higgins 611, Rancagua, Chile (pedro.perez@uoh.cl).

optimization. Hence, ξ_2 will be assumed to belong to a user-defined uncertainty set $\mathcal{U}$ and feasibility of x interpreted as

$$\mathbb{P}[g_t(x, (\xi_1, u)) \leq 0 \ \ \forall t \in T, \ u \in \mathcal{U}] \geq p.$$

Since uncertainty sets could be general, and popular forms are (unbounded) ellipsoids or polyhedra, the motivation to look at (1.1) with the least possible assumptions on T arises. Further applications result from considering situations wherein some physical system is subject to constraints at each moment in time and considering time to be continuous appears reasonable. An example of this form would be setting up a good/optimal design for offshore wind turbines, which have to withstand environmental conditions over their lifespan with high probability. Further, recent applications in PDE constrained optimization (e.g., [23]) also make a strong case for extending considerations to the infinite-dimensional Banach space setting as far as the decision vector is concerned.

Prior to continuing the discussion on the investigations of this paper, let us first recall some recent work related to probability constraints in general. First, introductions and appropriate references to applications (e.g., [26]) can be found in [16, 28, 57, 70]. Popular methods for dealing with finite-dimensional probability constraints are sample-based approximations (see, e.g., [42, 43, 44, 45, 46, 53, 67]) with various strengthening procedures (see, e.g., [35]); boolean approaches (see, e.g., [38, 39]); p-efficient point-based concepts (see, e.g., [17, 18, 19, 20, 40, 41, 64]); robust optimization [3, 4]; penalty approaches [21]; scenario approximations [9, 10]; convex approximations [52], and yet other approximations [25, 30].

Although this rich literature may yield the impression that the nonlinear constraint (1.1) can not necessarily be dealt with as such, this is not so. Much progress has also been made in "exact" nonlinear programming methods. Here the word exact is essentially meant to distinguish against methods which replace the probability function by an approximation, potentially through sampling. The here thus called "exact" methods deal with a probability constraint as a nonlinear constraint without seeking to replace the probability function. Vital ingredients are thus methods for computing (generalized) derivatives of probability functions. Let us mention, briefly, the efficient nonlinear programming methods based on SQP [8] or the promising bundle methods [65, 66, 74] (where convexity of the feasible set of (1.1) is required; see [73] for recent investigations).

One main cornerstone of these latter approaches is a better understanding of (generalized) differentiability of (1.1). The consequence of these studies is not only theoretical understanding of smoothness (or lack thereof) of probability functions, but also the provision of readily implementable formulae for gradients. Such formulae can be further improved by using well-known "variance reduction" techniques, such as quasi-Monte Carlo methods (see, e.g., [7]) or importance sampling (see, e.g., [2]). The study of differentiability of probability functions has attracted much attention in the past (see, e.g., [24, 29, 32, 47, 54, 58, 59, 60, 61, 62, 71, 72]). Indeed, historically a first formula for the gradient of a probability function in the form of an involved surface integral was derived in [58]. This formula was later generalized into a combination of a surface and volume integral in [61, 62]. Here it is also worthwhile mentioning that Marti [47] first suggested transforming the probability function to a volume integral in order to ensure that the integration domain does not depend on the decision vector.

Many of these papers provide gradient formulae for fairly general classes of distributions, for instance in the form of surface and/or volume integrals associated with the feasible set $K := \{z \in \mathbb{R}^m : g(\bar{x}, z) \leq 0\}$, where $\bar{x}$ is the point at which the derivative $\nabla\varphi$ is supposed to be computed. This generality comes with two drawbacks: first, the mentioned surface/volume integrals may be difficult to deal with numerically, at least for nonlinear g (see, e.g., [56, Page 207] and [60, Page 3]). Second, a principal assumption made in order to derive differentiability of φ at all is the compactness of the set K (see, e.g., [61, Assumption (A2), Page 200], [60, Assumption 2.2(i)], and [54, Page 902]). Not to mention the fact that g is assumed to have finitely many components. Indeed, without compactness, one cannot expect differentiability of φ even with the nicest data. Things are even worse though: one can not expect that φ is locally Lipschitzian. We refer the reader to [68, Proposition 2.2] and [27, Example 1] for counterexamples. Similarly, φ may fail to be continuously differentiable when the index set T in (1.1) consists of a finite set of cardinality more than 2, as illustrated in [69]. The extension of previous results to arbitrary index sets T is our main motivation in this work. This brings about quite some additional technical difficulties, for instance, upper semicontinuity of supremum functions can not be taken for granted. Consequently a nontrivial analysis results.

Our most notable contributions are as follows.

- We extend the previous results of [69] to the situation wherein X is a (separable) Banach space and T is an arbitrary index set. We also provide estimates for the more general Kruger–Mordukhovich or limiting subdifferentials rather than just the Clarke one.
- We extend the results of [27] to arbitrary index sets T with cardinality more than 1.
- We consider random vectors ξ belonging to the large class of elliptical symmetric random vectors. As such we extend [63] to the situation wherein X is a (separable) Banach space and T is an arbitrary index set. Again, we also provide estimates for the more general Kruger–Mordukhovich or limiting subdifferentials rather than just the Clarke one.

Throughout the paper, we will also illustrate with small examples some of the difficulties occurring when T is an arbitrary index set.

The paper is organized as follows: in section 2 we provide an account of the notation, concepts, and background information used. We also provide a set of technical results which form the workhorses for further derivations in the paper. Section 3 is dedicated to providing a first careful characterization of the various subdifferentials of φ in (1.1) under the most general assumptions on the maps g_t and no assumptions on T. We also show how a concrete assumption on the nominal data g_t entails the required conditions. Subsequently, section 4 is devoted to providing stronger estimates for the subdifferentials of φ. These estimates come about from additional assumptions on g_t in section 4.1 and from compactness-related assumptions on T in section 4.2.

2. Preliminaries.

2.1. Notation. Throughout the paper $(X, \|\cdot\|)$ will be a separable reflexive Banach space and X^* its dual. For a point $x \in X$ and $r \geq 0$ the closed ball of radius r centered at x is denoted by $\mathbb{B}_r(x)$. For $x^* \in X^*$, $\mathbb{B}_r(x^*)$ denotes a similar ball in the dual space and with $\mathbb{B}$ and $\mathbb{B}^*$ we respectively denote the unit balls in X and X^*. The symmetric bilinear form $\langle\cdot,\cdot\rangle : X^* \times X \to \mathbb{R}$ is given by $\langle x^*, x\rangle := \langle x, x^*\rangle =: x^*(x)$. The weak*-topology on X^* is denoted by $w(X^*, X)$ (w^*, for short) and the dual norm on X^* is simply denoted by $\|\cdot\|_*$.

For a matrix $A \in \mathcal{M}_{m,n}(\mathbb{R})$ the notation $\|A\|$ indicates the canonical norm of linear continuous maps with respect to the Euclidean norm, that is, $\|A\| :=$

$\sup_{z\in\mathbb{B}_1(0)}\|Az\|$. The negative polar of some closed cone $C\subseteq X$ is the closed convex cone

$$C^* := \{x^*\in X^*\colon\ \langle x^*,h\rangle\le 0\ \forall h\in C\}.$$

For a set $A\subseteq X$ (or $\subseteq X^*$), we denote by $\operatorname{int}(A)$, $\overline{A}$, $\operatorname{co}(A)$, and $\overline{\operatorname{co}}(A)$ the interior, the closure, the *convex hull*, and the *closed convex hull* of A.

Let U be an open set of X. The set $C^1(U,\mathbb{R})$ (C^1 for short) is defined as the set of all continuous differentiable functions from U to $\mathbb{R}$ and for any $\varphi\in C^1$ we denote by $\nabla\varphi$ its (Fréchet) derivative.

For a given function $f\colon X\to\mathbb{R}\cup\{+\infty\}$, the *domain* of f is

$$\mathcal{D}\mathrm{om}(f) := \{x\in X \mid f(x)<+\infty\}.$$

We say that f is *proper* if $\mathcal{D}\mathrm{om}(f)\neq\emptyset$ and $f>-\infty$. Now let $x\in\mathcal{D}\mathrm{om}(f)$. Then the set

$$\partial^{\mathsf{F}} f(x) := \left\{x^*\in X^* \;\middle|\; \liminf_{h\to 0}\frac{f(x+h)-f(x)-\langle x^*,h\rangle}{\|h\|}\ge 0\right\}$$

is called the *Fréchet subdifferential* of f at x. For convenience we set $\partial^{\mathsf{F}} f(x)=\emptyset$ for all $x\notin\mathcal{D}\mathrm{om}(f)$. We refer the reader to [33, 34] for an extensive discussion of properties of this subdifferential. It is well known that in a reflexive Banach space $x^*\in\partial^{\mathsf{F}} f(x)$ if and only if there is a neighborhood U of x and a continuous differentiable function $h\colon U\to\mathbb{R}$ such that $\nabla h(x)=x^*$ and $f-h$ attains its local minimum at x (see [6, Page 40] and recall that reflexive Banach spaces are Fréchet smooth Banach spaces).

The *limiting/Kruger–Mordukhovich subdifferential* and the *singular limiting/Kruger–Mordukhovich subdifferential* can be defined as (see, e.g., [6, 48])

$$\begin{aligned}
\partial^{\mathsf{M}} f(x) &:= \left\{w^*\text{-}\lim x_n^*\colon x_n^*\in\partial^{\mathsf{F}} f(x_n) \text{ and } x_n\xrightarrow{f} x\right\},\\
\partial^{\infty} f(x) &:= \left\{w^*\text{-}\lim \lambda_n x_n^*\colon x_n^*\in\partial^{\mathsf{F}} f(x_n),\ x_n\xrightarrow{f} x, \text{ and } \lambda_n\to 0^+\right\},
\end{aligned}$$

where $x_n\xrightarrow{f} x$ denotes $x_n\to x$ such that $f(x_n)\to f(x)$.

2.2. Technical tools from variational analysis. The following lemma establishes a variational characterization of the subgradients for the Fréchet subdifferential of the infimum function of an arbitrary family of functions in terms of the original members. Although we provide a direct proof in the appendix, we refer the reader to Theorem 4.1 of [37] for a similar statement. We also refer the reader to [51] for a deeper discussion of the variational principles which form the backbone of our proof.

LEMMA 2.1. *Let $t\in T$ be arbitrary and $f_t\colon X\to\mathbb{R}\cup\{+\infty\}$ be lower semicontinuous (l.s.c.) functions. Define $f: X\to\mathbb{R}\cup\{\infty\}$ as $f(x):=\inf_{t\in T} f_t(x)$. Let $\bar{x}\in X$ be given and assume f is l.s.c. at $\bar{x}$ and that $\bar{x}^*\in\partial^{\mathsf{F}} f(\bar{x})$ is arbitrary. Then for every $\varepsilon>0$ there are $t\in T$, $x_t\in X$, $x_t^*\in\partial^{\mathsf{F}} f_t(x_t)\subseteq X^*$ satisfying*

- $\|\bar{x}-x_t\|\le\varepsilon$,
- $\|\bar{x}^*-x_t^*\|\le\varepsilon$,
- $|f_t(x_t)-f(\bar{x})|\le\varepsilon$,
- $|f_t(\bar{x})-f(\bar{x})|\le\varepsilon$.

Moreover, $\partial^{\mathsf{F}} f(\bar{x})\subseteq\partial^{\mathsf{F}} f_t(\bar{x})$ for all $t\in T$ with $f_t(\bar{x})=f(\bar{x})$.

Let us end this section with the following observation gathered in the form of a lemma.

LEMMA 2.2. *Let $\mathfrak{S}\colon X \rightrightarrows Y$ be a set-valued map between Banach spaces X and Y. Then the following statements are equivalent:*

(a) *$\mathfrak{S}$ is locally bounded.*

(b) *$\mathfrak{S}$ is bounded on compact sets.*

2.3. Supremum function: Alternative forms of the probability function. We define the pointwise supremum function $g\colon X \times \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$ as follows:

$$g(x, z) := \sup_{t \in T} g_t(x, z). \tag{2.1}$$

As a result (1.1) can also be written as $\varphi(x) := \mathbb{P}[g(x, \xi) \leq 0]$. This latter form will become useful for the study of properties of φ as soon as further assumptions are made. The first and most impactful of such assumptions is one regarding the form of ξ. Indeed when assuming ξ to be elliptical-symmetric, a convenient representation of φ, which we shall call a spherical-radial representation, becomes available. This latter representation allows us to tie properties of g, actually g_t, much more easily to those of φ. We observe nonetheless that g is convex in the second argument and that the framework thus covers any convex transform of elliptical-symmetric random vectors, and hence in particular "log-normal" distributions, which are exponentials of Gaussian random vectors. In order to make this precise, let us first introduce elliptical distributions and subsequently the spherical-radial representation of φ.

2.4. Elliptical distributions. In this work we will consider elliptical symmetrically distributed random vectors having a density. The formal definition is as follows.

DEFINITION 2.3. *We say that the random vector $\xi \in \mathbb{R}^m$ is elliptical symmetrically distributed with mean μ, positive definite covariance matrix Σ, and generator $\theta\colon \mathbb{R}_+ \to \mathbb{R}_+$, which is denoted by $\xi \sim \mathcal{E}(\mu, \Sigma, \theta)$, if and only if its density $f_\xi\colon \mathbb{R}^m \to \mathbb{R}_+$ is given by*

$$f_\xi(z) = \big(\det \Sigma\big)^{-1/2} \theta\Big((z-\mu)^\top \Sigma^{-1}(z-\mu)\Big), \tag{2.2}$$

where the generator function $\theta : \mathbb{R}_+ \to \mathbb{R}_+$ must satisfy

$$\int_0^\infty t^{\frac{m}{2}} \theta(t) dt < \infty.$$

Now, consider L as the matrix arising from the Choleski decomposition of Σ, i.e., $\Sigma = LL^\mathsf{T}$. It can be shown that ξ admits a representation as

$$\xi = \mu + \mathcal{R} L \zeta,$$

where ζ has a uniform distribution over the Euclidean m-dimensional unit sphere $\mathbb{S}^{m-1} := \{z \in \mathbb{R}^m\colon \sum_{i=1}^m z_i^2 = 1\}$ and $\mathcal{R}$ possesses a density, which is given by

$$f_\mathcal{R}(r) := \frac{2\pi^{\frac{m}{2}}}{\Gamma(\frac{m}{2})} r^{m-1} \theta(r^2). \tag{2.3}$$

The family of elliptical random vectors includes many classical families. For instance, Gaussian random vectors and Student random vectors are elliptical with the respective generators

$$\theta^{\text{Gauss}}(t) = \exp(-t/2)/(2\pi)^{m/2}, \tag{2.4}$$

$$\theta^{\text{Student}}(t) = \frac{\Gamma\left(\frac{m+\nu}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)}(\pi\nu)^{-m/2}\left(1+\frac{t}{\nu}\right)^{-\frac{m+\nu}{2}}, \tag{2.5}$$

where Γ is the usual gamma function. Other examples, such as logistic or exponential power random vectors, have been considered in the literature; see, e.g., [22] and [36].

2.5. Spherical-radial representation of the probability function. Let us consider the probability function (1.1) wherein ξ is an elliptical symmetrically distributed random vector. Then, when recalling (2.1) as well as the spherical-radial decomposition of ξ, we can provide the following highly convenient representation of φ:

$$\varphi(x) = \int_{v\in\mathbb{S}^{m-1}} \mu_{\mathcal{R}}(\{r \geq 0\colon g(x, \mu + rLv) \leq 0\})d\mu_\zeta(v), \tag{2.6}$$

where μ_ζ refers to the uniform measure on the Euclidian sphere $\mathbb{S}^{m-1}$ and $\mu_{\mathcal{R}}$ the one-dimensional measure induced by the radial component of ξ, i.e., $\mu_{\mathcal{R}}$ is a measure absolutely continuous with respect to the Lebesgue measure on $\mathbb{R}$ and has density given by (2.3). For this derivation, we also refer the reader to [27, 63, 68, 69, 73].

The study of the inner measure will be of great importance and hence we introduce the map $e : X \times \mathbb{S}^{m-1} \to [0, 1]$ defined as

$$e(x, v) = \mu_{\mathcal{R}}(\{r \geq 0\colon g(x, \mu + rLv) \leq 0\}). \tag{2.7}$$

Now, by redefining g_t through its composition with an affine transform in the second argument not altering convexity (see [63, section 2.3] or [68, Remark 3.2] for details), without loss of generality $\mu = 0$ may be taken. Likewise, Σ can be considered to be a correlation matrix. We will therefore make this assumption from now on.

For specific elements $x \in X$, actually all those at which we are usually interested in studying φ, under the current assumptions, we can provide a precise characterization of the set $\{r \geq 0\colon g(x, rLv) \leq 0\}$. This characterization relies on the characterization of the sets $\{r \geq 0\colon g_t(x, rLv) \leq 0\}$ involving the mappings g_t, which are assumed to be continuously differentiable. The following earlier established result, through the use of convexity and the implicit function theorem (see, e.g., [11, Theorem 4.7.1]), will be helpful in this regard.

LEMMA 2.4 (see Lemma 3.2 in [68]). *Let $\hat{g} : X \times \mathbb{R}^m \to \mathbb{R}$ be a continuously differentiable map that is convex in the second argument. Then for any $x \in \hat{D} := \{x \in X\colon \hat{g}(x, 0) < 0\}$ and $v \in \mathbb{S}^{m-1}$, define $\hat{\rho} : \hat{D} \times \mathbb{S}^{m-1} \to \mathbb{R}_+ \cup \{\infty\}$ as*

$$\hat{\rho}(x, v) = \begin{cases} \sup_{r\geq 0} & r \\ s.t. & \hat{g}(x, rLv) \leq 0. \end{cases} \tag{2.8}$$

Then for any fixed $(x, v) \in \hat{D} \times \mathcal{D}\text{om}(\hat{\rho}(x, \cdot))$, $r = \hat{\rho}(x, v) < \infty$ is the unique solution to $\hat{g}(x, rLv) = 0$ and at such (x, v), $\hat{\rho}$ is Fréchet differentiable on an appropriate neighborhood $U \times V$ of (x, v) with gradient formula

$$\nabla_x\hat{\rho}(x, v) = -\frac{1}{\langle\nabla_z\hat{g}(x, \hat{\rho}(x, v)Lv), Lv\rangle}\nabla_x\hat{g}(x, \hat{\rho}(x, v)Lv).$$

For any $t \in T$, we define

$$D_t := \{x \in X \colon g_t(x,0) < 0\}, \tag{2.9a}$$

$$F_t(x) := \mathcal{D}\mathrm{om}(\rho_t(x,\cdot)) = \left\{v \in \mathbb{S}^{m-1} \colon \ \exists r > 0 \text{ s.t. } g_t(x, rLv) = 0\right\}, \tag{2.9b}$$

$$I_t(x) := \left\{v \in \mathbb{S}^{m-1} \colon \text{s.t. } g_t(x, rLv) < 0 \ \forall r > 0\right\}, \tag{2.9c}$$

where $x \in D_t$, $F_t, I_t : X \rightrightarrows \mathbb{S}^{m-1}$, and ρ_t refers to the solution map resulting from the application of Lemma 2.4 with $\hat{D} = D_t$ and $\hat{g} = g_t$.

The set-valued mappings I_t, F_t partition the sphere.

Lemma 2.5 (see Lemma 3.1 of [68]). *For any $t \in T$ and $x \in D_t$, we have*

1. $F_t(x) \cup I_t(x) = \mathbb{S}^{m-1}$,
2. $F_t(x)$ *is open in* $\mathbb{S}^{m-1}$,

where F_t and I_t are as defined in (2.9b) *and* (2.9c), *respectively.*

Proof. To see the last item, observe that for $x \in D_t$, $v \in F_t(x)$, $g_t(x,0) < 0$, and convexity of g_t in the second argument imply that the $r > 0$ for which $g_t(x, rLv) = 0$ must be unique. In particular, for $r' > r$, $g_t(x, r'Lv) > 0$ must hold true. By continuity of g_t, we may thus find neighborhoods U of x, W of $r'Lv$ (in $\mathbb{R}^m$) such that $g_t(x', z') > 0$ for all $(x', z') \in U \times W$ and $g_t(x', 0) < 0$ for all $x' \in U$. Consequently, by continuity of g_t, $V := \frac{1}{r'} W \cap \mathbb{S}^{m-1}$ is a neighborhood of v in the relative topology of $\mathbb{S}^{m-1}$ contained in $F_t(x)$. ∎

It thus follows that for $x \in D_t$, the set $\{r \geq 0 \colon g_t(x, rLv) \leq 0\}$ is simply the (potentially infinite) interval $[0, \rho_t(x,v)]$. However, since g is an "arbitrary" supremum function, a fine representation of e is more intricate, and this is the topic of the following proposition.

Proposition 2.6. *Let D be defined as $D := \bigcap_{t \in T} D_t$, where D_t is as in* (2.9a). *Then we define the map $\rho : D \times \mathbb{S}^{m-1} \to \mathbb{R}_+$ as*

$$\rho(x,v) := \inf_{t \in T} \rho_t(x,v), \tag{2.10}$$

where $\rho_t : D_t \times \mathbb{S}^{m-1} \to \mathbb{R}_+$ and D_t are defined as in Lemma 2.4. Then for any $x \in D$, $v \in \mathbb{S}^{m-1}$, it holds that

$$\{r \geq 0 \colon g(x, rLv) \leq 0\} = [0, \rho(x,v)], \tag{2.11}$$

where $[0,\infty] = [0,\infty)$ is meant.

Moreover, for $x \in D^\circ := \{x \in X \colon g(x,0) < 0\}$, one has that if there exists $r > 0$ such that $g(x, rLv) = 0$, then $r = \rho(x,v)$. In particular, if g° denoting the restriction of g to $D^\circ \times \mathbb{R}^m$ is finite valued, the function ρ has the following alternative representation:

$$\rho(x,v) = \begin{cases} r & \text{such that } g(x, rLv) = 0, \\ +\infty & \text{otherwise.} \end{cases} \tag{2.12}$$

Proof. Let $x \in D$ be given and consider an arbitrary $v \in \mathbb{S}^{m-1}$ and $r < \rho(x,v)$. Let $t \in T$ be arbitrary. Then by convexity of $g_t(x,\cdot)$ and uniqueness of a potential solution (see Lemma 2.4) to $g_t(x, r'Lv) = 0$, we have that $g_t(x, rLv) < 0$. Since t was arbitrary, we obtain $g(x, rLv) \leq 0$. Consequently, since $r < \rho(x,v)$ was arbitrary too, and $g(x,\cdot)$ is l.s.c., we have $g(x, \rho(x,v)Lv) \leq 0$, which implies that

$$[0, \rho(x,v)] \subseteq \{r \geq 0 \colon g(x, rLv) \leq 0\}.$$

Now, consider $r > 0$ such that $g(x, rLv) > 0$. Then there exists $t \in T$ such that $g_t(x, rLv) > 0$. By the already invoked uniqueness of the solution $g_t(x, r'Lv) = 0$, we have that $\rho(x, v) \leq \rho_t(x, v) < r$, and consequently $r \notin [0, \rho(x, v)]$. This establishes (2.11).

Let $x \in D^\circ$ be given and assume that there exists $r > 0$ such that $g(x, rLv) = 0$. Then on the one hand, by (2.11), $r \leq \rho(x, v)$, and on the other hand, if $r < \rho(x, v)$ the convexity of $g(x, \cdot)$ implies that $g(x, rLv) < 0$, which is a contradiction.

Now let us assume that g° has finite values and that x is such that $g(x, 0) < 0$. Then for arbitrary $v \in \mathbb{S}^{m-1}$, it follows that the convex function $r \mapsto g(x, rLv)$ contains $\mathbb{R}_+$ inside its domain. Since convex functions are continuous inside their domain, it follows that should there exist an $r' > 0$ at which $g(x, r'Lv) > 0$, then there must exist some $r > 0$ such that $g(x, rLv) = 0$, and such an r is unique in $\mathbb{R}_+$. If no $r > 0$ can be found such that $g(x, rLv) = 0$, we must by these arguments have $g(x, rLv) < 0$ for all $r > 0$. Now the previous item concludes the identification of (2.12). ∎

The hypothesis related to g taking on finite values is important for ρ to actually be the solution to $g(x, rLv) = 0$ in r. This is illustrated by the following example.

Example 2.7. Consider the functions $g_n \colon \mathbb{R} \times \mathbb{R}^2 \to \mathbb{R}$ given by

$$g_n(x, z) = \begin{cases} x^2 - 1 & \text{if } z_1^2 + z_2^2 \leq 1, \\ x^2 + n(z_1^2 + z_2^2 - 1)^2 - 1 & \text{if } z_1^2 + z_2^2 > 1. \end{cases}$$

The supremum of this family is

$$g(x, z) = \begin{cases} x^2 - 1 & \text{if } z_1^2 + z_2^2 \leq 1, \\ +\infty & \text{if } z_1^2 + z_2^2 > 1. \end{cases}$$

Consequently, for any $v \in \mathbb{S}^{m-1}$, $\{r \colon g(0, rv) \leq 0\} = [0, 1]$ and there is no $r > 0$ such that $g(0, rv) = 0$. Moreover, for any $x \in [0, 1)$ and $v \in \mathbb{S}^{m-1}$, we can compute

$$\rho_n(x, v) = \sqrt{1 + \sqrt{\frac{1 - x^2}{n}}}$$

and establish $\rho(x, v) = 1$.

As a consequence of Proposition 2.6, and particularly (2.11), it becomes possible to provide a convenient and nearly explicit expression for e introduced in (2.7). Indeed at any $x \in D$, $v \in \mathbb{S}^{m-1}$, it follows that

$$e(x, v) = F_{\mathcal{R}}(\rho(x, v)), \tag{2.13}$$

where $F_{\mathcal{R}}$ is the radial distribution function associated with $\mathcal{R}$. Hence for any $x \in D$ we have the representation

$$\varphi(x) = \int_{v \in \mathbb{S}^{m-1}} e(x, v) d\mu_\zeta(v). \tag{2.14}$$

Subsequently our endeavour of providing formulae for generalized differentiation of φ will consist in carefully justifying the interchange of subdifferentiation and integration. Second, we will seek to express subdifferentials of e in terms of nominal data. One noteworthy difficulty already resides in the fact that even if $\bar{x} \in D$ holds true, no neighborhood of $\bar{x}$ entirely contained in D may exist.

We end this section by introducing the sets of nearly active indices. To this end, let $\epsilon \geq 0$ be given. We then define

$$T_\epsilon^g(x,z) := \{t \in T \colon g(x,z) \leq g_t(x,z) + \epsilon\} \tag{2.15a}$$

for any $x \in X$ and $z \in \mathbb{R}^m$ as well as

$$T_\epsilon^\rho(x,v) := \{t \in T \colon \rho_t(x,v) \leq \rho(x,v) + \epsilon\}\,, \tag{2.15b}$$

for any $x \in D$ and $v \in \mathcal{D}\mathrm{om}(\rho(x,\cdot))$. It is clear that the $\varepsilon = 0$ case refers to the set of active indices, which we will also denote by $T^g(x,z)$ and $T^\rho(x,v)$ for simplicity.

3. A characterization of the subdifferential of the probability function in terms of nominal data.

3.1. Lower semicontinuity of the resolvant map.

One of the first technical difficulties is related to understanding continuity properties of the map ρ exhibited in Proposition 2.6. Indeed, although this map is the infimum over a family of continuously differentiable maps, it is already unclear if ρ is l.s.c. This difficulty is illustrated by the following example.

Example 3.1. Consider the family of real functions defined for $n \in \mathbb{N}$ as $f_n(t) = 1$ if $t \leq 0$ and $f_n(t) = \max\{0, 1 - nt\}$ if $t \geq 0$. Then $f = \inf_{n\in\mathbb{N}} f_n$ is the characteristic function of the negative reals, and hence not l.s.c.

The second difficulty is in trying to convey this property from the original nominal data specified in terms of g_t. To this end, consider the following result.

PROPOSITION 3.2. *Let x_0 be a point in X such that there exists a neighborhood U of x_0 such that*

(a) *the function ρ defined in* (2.10) *admits the representation* (2.12)*,*
(b) *$g(x,0) < 0$ for all $x \in U$,*
(c) *the set $\mathfrak{K} := \{(x,z) \in U \times \mathbb{R}^m \colon g(x,z) = 0\}$ is closed.*

Then $\rho(x_n, v_n) \to \rho(x,v)$ for every sequence $U \times \mathbb{S}^{m-1} \ni (x_n, v_n) \to (x,v) \in U \times \mathbb{S}^{m-1}$. Consequently, the radial function e defined in (2.13) *is continuous over $U \times \mathbb{S}^{m-1}$.*

Proof. First, the set $U \subseteq D^\circ$, and so ρ, is well defined for every $(x,v) \in U \times \mathbb{S}^{m-1}$. On the one hand ρ is upper semicontinuous (u.s.c.), because by Proposition 2.6 this function is the infimum of an arbitrary family of continuous functions. On the other hand, consider a sequence $U \times \mathbb{S}^{m-1} \ni (x_n, v_n) \to (x,v)$. By the fact that ρ is u.s.c., we have

$$\rho(x,v) \geq \limsup \rho(x_n, v_n) \geq \liminf \rho(x_n, v_n) \geq 0.$$

If $\liminf \rho(x_n, v_n) = +\infty$ the conclusion holds. Otherwise, if $\liminf \rho(x_n, v_n) < +\infty$, then there exists a subsequence $\rho(x_{n_k}, v_{n_k}) \to r \in \mathbb{R}$. Since

$$\mathfrak{K} := \{(x,z) \in U \times \mathbb{R}^m \colon g(x,z) = 0\}$$

is closed and $(x_n, \rho(x_n, v_n)Lv_n) \in \mathfrak{K}$, necessarily $(x, rLv) \in \mathfrak{K}$, which means that $g(x, rLv) = 0$. Then by (2.12) we have $\rho(x,v) = r$. Finally, the continuity of e follows straightforwardly from the last convergence and the representation of e as the potentially improper integral $e(x,v) = \int_0^{\rho(x,v)} f_{\mathcal{R}}(r)dr$. ∎

Remark 3.3. It is important to mention that Assumption 3.2 in Proposition 3.2 is equivalent to the existence of some neighborhood $\tilde{U}$ of x and $l > 0$ such that for every $x \in \tilde{U}$, $g(x,0) \leq -\frac{1}{l}$. Indeed, consider the increasing sequence of sets $U_n := \{x \in U \colon g(x,0) \leq -\frac{1}{n}\}$. Then U_n is closed (because g is l.s.c.) and $U = \bigcup_{n\in\mathbb{N}} U_n$. Hence, by *Baire's theorem* one of the U_n has nonempty interior.

We end this section with another way to ensure that ρ is l.s.c.

PROPOSITION 3.4. *Assume that the mapping $g : X \times \mathbb{R}^m \to \mathbb{R}$ is u.s.c. Then for a fixed $v \in \mathbb{S}^{m-1}$, the map $\rho : D \times \mathbb{S}^{m-1} \to \mathbb{R}_+$ defined in Proposition* 2.6 *is l.s.c. at any $x \in D$ for which it is finite.*

Proof. Let us first note that ρ is the solution to the following convex optimization problem:

$$\rho(x,v) = \begin{cases} \sup_r & r \\ \text{s.t.} & g_t(x, rLv) \leq 0 \ \forall t \in T. \end{cases}$$

If ρ is not l.s.c. at a given $x_0 \in D$, we may find some $\varepsilon > 0$ and sequence $x_n \to x_0$ such that $\rho(x_n, v) < \rho(x_0, v) - \varepsilon$. Moreover, by convexity (and the arguments exhibited in the proof of Proposition 2.6), since $\rho(x_0, v) < \infty$, we must have $\rho(x_n, v) < \infty$, and consequently, by convexity, $g(x_n, (\rho(x_0, v) - \varepsilon)Lv) > 0$. Now, by using similar arguments once more, we must also have $g(x_0, (\rho(x_0, v) - \varepsilon)Lv) < 0$, and let us pick $\delta > 0$ such that $g(x_0, (\rho(x_0, v) - \varepsilon)Lv) < -\delta$. By upper semicontinuity of the map g in the first argument, at $(x_0, (\rho(x_0, v) - \varepsilon)Lv)$ we can find some neighborhood U of x_0 such that

$$g(x', (\rho(x_0, v) - \varepsilon)Lv) \leq g(x_0, (\rho(x_0, v) - \varepsilon)Lv) + \delta$$

for all $x' \in U$. This entails in particular $g(x', (\rho(x_0, v) - \varepsilon)Lv) \leq 0$, but $x_n \in U$ for sufficiently large n, and this then contradicts $g(x_n, (\rho(x_0, v) - \varepsilon)Lv) > 0$. ∎

3.2. Subdifferential estimates for the radial function. Now that we have studied conditions under which ρ, the resolvant map, and e, the radial function, are l.s.c., we will study estimates of their subdifferentials. For the mapping e, given in (2.13), as the composition of $F_{\mathcal{R}} \circ \rho$, we recall that $F_{\mathcal{R}}$ is differentiable and typically even continuously differentiable. Hence the major part of the study boils down to the study of ρ and a careful application of the chain rule.

Our starting point will be a natural outer estimate of the Fréchet subdifferential of ρ resulting from Lemma 2.1. Let us introduce this set-valued mapping $\mathcal{S} \colon D^\circ \times \mathbb{S}^{m-1} \rightrightarrows X^*$ as follows:

(3.1)
$$\mathcal{S}(x,v) := \begin{cases} \bigcap\limits_{t\in T^\rho(x,v)} \left\{ x^* = \frac{\nabla_x g_t(x,\rho(x,v)Lv)}{\langle \nabla_z g_t(x,\rho(x,v)Lv), Lv\rangle} \right\} & \text{if } T^\rho(x,v) \neq \emptyset, \\ \bigcap\limits_{\varepsilon>0} \left\{ x^* \in X^* \colon \begin{array}{c} \exists t_n \in T^\rho_\varepsilon(x,v), \ x_n \to x \\ \text{s.t. } \rho_{t_n}(x_n, v) \to \rho(x,v), \\ x^* = \lim\limits_{n\to\infty} \frac{\nabla_x g_{t_n}(x_n, \rho_{t_n}(x_n,v)Lv)}{\langle \nabla_z g_{t_n}(x_n,\rho_{t_n}(x_n,v)Lv), Lv\rangle} \end{array} \right\} & \text{otherwise.} \end{cases}$$

We can observe that $\partial^{\mathsf{F}}_x \rho(x,v) \subseteq \mathcal{S}(x,v)$ for every (x,v) with $\rho(x,v) < +\infty$.

With the help of these ingredients we can produce the following convenient outer estimate of the limiting and singular subdifferentials of the resolvant map ρ.

Proposition 3.5. *Under the assumptions of Proposition 3.2, for every $x \in U$ the (regular) partial Kruger–Mordukhovich subdifferential of ρ satisfies*

$$\partial_x^{\mathtt{M}}\rho(\bar{x},v) \subseteq \mathcal{S}^{\mathtt{M}}(\bar{x},v) := \limsup_{\substack{x_n\to x,\\ \rho(x_n,v)\to\rho(\bar{x},v)}} \mathcal{S}(x_n,v)$$

$$\subseteq \left\{ x^*\in X^*\colon \begin{array}{c} \exists\varepsilon_n\to 0^+,\ x_n\to\bar{x},\ \ \exists t_n\in T,\\ s.t.\ \rho_{t_n}(x_n,v)\to\rho(\bar{x},v),\\ x^* = w^*\text{-}\lim\limits_{n\to\infty} -\frac{\nabla_x g_{t_n}(x_n,\rho_{t_n}(x_n,v)Lv)}{\langle\nabla_z g_{t_n}(x_n,\rho_{t_n}(x_n,v)Lv),Lv\rangle} \end{array}\right\} \tag{3.2a}$$

and the singular partial Kruger–Mordukhovich subdifferential satisfies

$$\partial_x^{\infty}\rho(\bar{x},v) \subseteq \mathcal{S}^{\infty}(\bar{x},v) := \limsup_{\substack{x_n\to x,\\ \lambda_n\to 0^+,\\ \rho(x_n,v)\to\rho(\bar{x},v)}} \lambda_n\mathcal{S}(x_n,v)$$

$$\subseteq \left\{ x^*\in X^*\colon \begin{array}{c} \exists\varepsilon_n,\lambda_n\to 0^+,\ x_n\to\bar{x},\ \ \exists t_n\in T,\\ s.t.\ \rho_{t_n}(x_n,v)\to\rho(\bar{x},v),\\ x^* = w^*\text{-}\lim\limits_{n\to\infty} -\lambda_n\frac{\nabla_x g_{t_n}(x_n,\rho_{t_n}(x_n,v)Lv)}{\langle\nabla_z g_{t_n}(x_n,\rho_{t_n}(x_n,v)Lv),Lv\rangle} \end{array}\right\}. \tag{3.2b}$$

Proof. Let $x^* \in \partial_x^{\mathtt{M}}\rho(\bar{x},v)$ be arbitrary, so that the first inclusion in (3.2a) and (3.2b) follows from the definition and the fact that $\partial_x^{\mathtt{F}}\rho(x,v) \subseteq \mathcal{S}(x,v)$. Then, by construction, we can find a sequence $x_n \to \bar{x}$, $x_n^* \in \partial_x^{\mathtt{F}}\rho(x_n,v)$ such that $\rho(x_n,v) \to \rho(\bar{x},v)$ and $x_n^* \xrightarrow{w^*} x^*$. Now, by Proposition 3.2, ρ is l.s.c. at these points for n large enough. Hence, we may invoke Lemma 2.1 to establish the existence of a sequence $\varepsilon_n \downarrow 0$, $t_n \in T$, as well as $x_{t_n} \in D$, $x_{t_n}^* \in \partial_x^{\mathtt{F}}\rho_{t_n}(x_{t_n},v)$ with the following properties:

- $\|x_{t_n} - x_n\| \le \varepsilon_n$,
- $\|x_{t_n}^* - x_n^*\| \le \varepsilon_n$,
- $|\rho_{t_n}(x_{t_n},v) - \rho(x_n,v)| \le \varepsilon_n$,
- $|\rho_{t_n}(x_n,v) - \rho(x_n,v)| \le \varepsilon_n$.

By employing the triangle inequality several times we get

$$\begin{aligned} \|\bar{x} - x_{t_n}\| &\le \|\bar{x} - x_n\| + \|x_n - x_{t_n}\|, \\ |\rho(\bar{x},v) - \rho_{t_n}(x_{t_n},v)| &\le |\rho(\bar{x},v) - \rho(x_n,v)| + |\rho(x_n,v) - \rho_{t_n}(x_{t_n},v)|, \\ \|x^* - x_{t_n}^*\| &\le \|x^* - x_n^*\| + \|x_n^* - x_{t_n}^*\|. \end{aligned}$$

This means that $x_{t_n} \to \bar{x}$, $x_{t_n}^* \to x^*$, and $\rho_{t_n}(x_{t_n},v) \to \rho(\bar{x},v)$, which together with Lemma 2.4 gives

$$x_{t_n}^* = -\frac{\nabla_x g_{t_n}(x_{t_n},\rho_{t_n}(x_{t_n},v)Lv)}{\langle\nabla_z g_{t_n}(x_{t_n},\rho_{t_n}(x_{t_n},v)Lv),Lv\rangle}.$$

This thus shows that $x^* \in \partial_x^{\mathtt{M}}\rho(\bar{x},v)$ belongs to the rightmost set in (3.2a). The arguments for (3.2b) are identical upon incorporating an additional sequence $\lambda_n \to 0$ by the very definition of the singular limiting subdifferential. ∎

We can now provide vital estimates for the radial function.

Lemma 3.6 (chain rule for radial function). *Under the assumptions of Proposition 3.2, assume that the set-valued map $S(\cdot,v)$ is locally bounded at every $x \in U$ with $v \in F(x)$ and the generator function θ is continuous. Then for every $x \in U$,*

$$\partial_x^{\mathtt{M}} e(x,v) \subseteq \begin{cases} f_{\mathcal{R}}(\rho(x,v))\partial_x^{\mathtt{M}}\rho(x,v) & \text{if } v\in F(x),\\ w^*\text{-}\limsup\limits_{\substack{x_n\to x,\\ x_n\in F(x_n)}} \left(f_{\mathcal{R}}(\rho(x_n,v))\partial_x^{\mathtt{F}}\rho(x_n,v)\right)\cup\{0\} & \text{if } v\in I(x), \end{cases} \tag{3.3}$$

where $F(x) = \mathcal{D}\mathrm{om}(\rho(x,\cdot))$, $I(x) = \mathbb{S}^{m-1} \setminus F(x)$, *and* $f_{\mathcal{R}}$ *is the density function given in* (2.3).

Proof. Consider $v \in F(x)$. If $S(\cdot, v)$ is locally bounded at $\bar{x}$, then by the inclusion $\partial_x^{\mathtt{F}} \rho(x,v) \subseteq \mathcal{S}(x,v)$ we get that $\partial_x^{\mathtt{F}} \rho(\cdot, v)$ is locally bounded at $\bar{x}$; consequently, by [48, Theorem 3.52], $\rho(\cdot, v)$ is locally Lipschitzian at $\bar{x}$. So by [48, Corollary 3.43] we have that $\partial_x^{\mathtt{M}} e(\bar{x}, v) \subseteq f_{\mathcal{R}}(\rho(\bar{x}, v)) \partial_x^{\mathtt{M}} \rho(\bar{x}, v)$. Observe that this chain rule can be applied: since ρ is locally Lipschitzian, $F_{\mathcal{R}}$ is continuously differentiable and so $\partial^\infty F_{\mathcal{R}}(\bar{r}) = \{0\}$ at all $\bar{r} \geq 0$, immediately entailing the required qualification condition. Moreover, $F_{\mathcal{R}}$ is also sequential normal epi-compact (SNEC), as noted in [48, Page 121].

Now, if $v \in I(x)$, $e(\cdot, v)$ attains a maximum at x. Hence, by Corollary 1(ii) of [27], it follows that $\partial_x^{\mathtt{F}} e(x,v) \subseteq \{0\}$. Indeed, for an arbitrary $x^* \in \partial_x^{\mathtt{F}} e(x,v)$ and $u \in X \setminus \{0\}$,

$$
\begin{aligned}
-\left\langle x^*, \|u\|^{-1} u \right\rangle &= \liminf_{n\to\infty} -\frac{\langle x^*, n^{-1}u\rangle}{\|n^{-1}u\|} \\
&\geq \liminf_{n\to\infty} \frac{e(x+n^{-1}u, v) - e(x,v) - \langle x^*, n^{-1}u\rangle}{\|n^{-1}u\|} \\
&\geq \liminf_{h\to 0} \frac{e(x+h,v) - e(x,v) - \langle x^*, h\rangle}{\|h\|} \geq 0,
\end{aligned}
$$

i.e., $\langle x^*, u\rangle \leq 0$ for all $u \in X$, entailing $x^* = 0$.

Now by definition,

$$
\partial_x^{\mathtt{M}} e(x,v) = w^*\text{-} \limsup_{\substack{x_n \to x,\\ e(x_n,v)\to e(x,v)}} \partial_x^{\mathtt{F}} e(x_n, v) \subseteq w^*\text{-} \limsup_{\substack{x_n \to x,\\ e(x_n,v)\to e(x,v)}} \partial_x^{\mathtt{M}} e(x_n, v).
$$

Moreover, by the previous part, $\partial_x^{\mathtt{M}} e(x_n, v) \subseteq f_{\mathcal{R}}(\rho(x,v)) \partial_x^{\mathtt{M}} \rho(x,v)$ for each x_n close to x with $v \in F(x_n)$, so

$$
\partial_x^{\mathtt{M}} e(x,v) \subseteq w^*\text{-} \limsup_{\substack{x_n \to x,\\ x_n \in F(x_n)}} \left(f_{\mathcal{R}}(\rho(x_n,v)) \partial_x^{\mathtt{M}} \rho(x_n,v)\right) \cup \{0\}.
$$

Finally, using a diagonal argument in the definition of $\partial_x^{\mathtt{M}} \rho(x_n, v)$ and the continuity of $f_{\mathcal{R}}$, we can conclude that

$$
\partial_x^{\mathtt{M}} e(x,v) \subseteq w^*\text{-} \limsup_{\substack{x_n \to x,\\ x_n \in F(x_n)}} \left(f_{\mathcal{R}}(\rho(x_n,v)) \partial_x^{\mathtt{F}} \rho(x_n,v)\right) \cup \{0\}. \qquad \square
$$

3.3. Generalized subdifferentiation of probability functions: First formulae. Now that estimates of the subdifferentials of ρ and e have been produced, it would be convenient to simply "pull" the subdifferential operation over the integral. Although the resulting formulae would be the expected ones, in general such an operation is not justified. This was already observed in the case when T was reduced to the singleton and ρ was found to be continuously differentiable. Then φ may fail to be smooth and this lack of smoothness can be tied to growth of $\nabla_x g_{\bar{t}}$ for large z. See, for instance, the counterexamples [68, Proposition 2.2] and [27, Example 1]. In general, that the allowed growth depends on the distribution of ξ is not very restrictive. This discussion thus motivates the introduction of the following "growth" cone.

DEFINITION 3.7. *For any $x \in X$ and $l > 0$, we define*

$$C_l(x) := \Big\{ h \in X \colon \langle \nabla_x g_t(x', z), h \rangle \leq l \left\| L^{-1} z \right\|^{-m} \theta^{-1}\Big(\left\| L^{-1} z \right\|^2 \Big) \|h\| \quad \forall x' \in \mathbb{B}_{1/l}(x), \ \left\| L^{-1} z \right\| \geq l \ \ \forall t \in T \Big\} \tag{3.4}$$

as the uniform l-cone of nice directions at x. Here θ^{-1}, related to the function θ in (2.2), *is such that $\theta^{-1}(t) = \frac{1}{\theta(t)}$, with $\theta^{-1}(t) = \infty$ for $\theta(t) = 0$. Moreover, we recall that its polar cone is denoted by $C_l^*(x)$.*

We can observe that whenever θ has bounded support, $C_l(x)$ is automatically the entire space.

PROPOSITION 3.8. *Let $x_0 \in D^\circ$ be given together with a neighborhood $U \subseteq D^\circ$ of x_0. Moreover, assume that $g_{|_{U \times \mathbb{R}^m}}$ is finite valued and the set*

$$\mathfrak{K} = \{(x, z) \in U \times \mathbb{R}^m \colon g(x, z) = 0\}$$

is closed in $X \times \mathbb{R}^m$. Consider the map $e : U \times \mathbb{S}^{m-1} \to \mathbb{R}_+$ defined as $e(x, v) = F_{\mathcal{R}}(\rho(x, v))$, with ρ defined as in Proposition 2.6 *and $F_{\mathcal{R}}$ the radial distribution function related to the elliptical symmetrically distributed random vector ξ with continuous generator.*

Let us also assume that

(i) *for every $v \in \mathcal{D}\mathrm{om}(\rho(x_0, \cdot))$ the mapping $\mathcal{S}$ is locally bounded at (x_0, v), which means that there are neighborhoods U_v of x_0, V_v of v, and a constant $K_v > 0$ such that, for all $(x', v') \in U_v \times V_v$ and $x^* \in \mathcal{S}(x', v')$, $\|x^*\| \leq K_v$.*

Then, for every $l > 0$, there exists some neighborhood U_l of x_0 and constant $K_l > 0$ such that

$$\partial_x^{\mathsf{F}} e(x, v) \subseteq K_l \mathbb{B}^* - C_l^*(x_0) \quad \forall x \in U_l, \tag{3.5}$$

where ∂_x^{F} refers to the partial Fréchet subdifferential of e w.r.t. x. Moreover, if $F(x) = \mathbb{S}^{m-1}$, we have that there exists some neighborhood U' of x_0 and constant $K > 0$ such that

$$\partial_x^{\mathsf{F}} e(x, v) \subseteq K \mathbb{B}^* \quad \forall x \in U'. \tag{3.6}$$

Here $\mathbb{B}^*$ denotes the unit ball of x^*.

Proof. Let us pick an arbitrary but fixed $v \in \mathbb{S}^{m-1}$ and $l > 0$. If v is such that $\rho(x_0, v) < \infty$, then we may by assumption (i) obtain a constant $K_v > 0$ and neighborhoods U_v of x_0 and V_v of v such that any $(x', v') \in U_v \times V_v$ and $x^* \in \mathcal{S}(x', v')$ satisfies $\|x^*\| \leq K_v$. Note in particular that by moving to the neighborhood $U_v \cap U$, $U_v \subseteq U$ may be assumed. By Proposition 3.2 and the continuity of the generator of ξ we may assume that

$$C := \sup_{(x', v') \in U_v \times V_v} f_{\mathcal{R}}(\rho(x', v')) < +\infty.$$

Moreover, by Lemma 3.6,

$$\partial_x^{\mathsf{F}} e(x', v') \subseteq f_{\mathcal{R}}(\rho(x', v')) K_v \mathbb{B}^* \subseteq \tilde{K}_v \mathbb{B}^* \quad \forall (x', v') \in U_v \times V_v,$$

where $\tilde{K}_v = C K_v$. We have, thus, shown that

$$\partial_x^{\mathsf{F}} e(x', v') \subseteq \tilde{K}_v \mathbb{B}^* \quad \forall x', v' \in U_v \times V_v. \tag{3.7}$$

Next, let us assume instead that v is such that $\rho(x_0, v) = \infty$, i.e., $e(x_0, v) = 1$. We now claim that we can find neighborhoods U_v of x_0 and V_v of v such that $\rho(x', v') \geq l$ for all $(x', v') \in U_v \times V_v$. Indeed, should this not hold, we can find a sequence $x_n \to x_0$ and $v_n \to v$ such that $\rho(x_n, v_n) \leq l$. However, by Proposition 3.2, the map ρ is l.s.c. and thus $\rho(x_0, v) \leq \lim_{n\to\infty} \rho(x_n, v_n) \leq l$, thus leading to a contradiction and establishing the claim. We may by shrinking the neighborhoods if needed assume that $U_v \subseteq U$ and $U_v \subseteq \mathbb{B}_{1/l}(x_0)$, $V_v \subseteq \mathbb{B}_{1/l}(v)$, and even that we can find some $\eta > 0$ such that $g(x, 0) \leq -\eta$ for all $x \in U_v$ (consider Baire's theorem; see, e.g., Remark 3.3).

Now for an arbitrary $(x', v') \in U_v \times V_v$, we have the following.

1. If v' is such that $\rho(x', v') = \infty$, then $e(x', v') = 1$ and $\partial_x^{\mathsf{F}} e(x', v') \subseteq \{0\}$ as a result of [27, Corollary 1(ii)] using only maximality of e at (x', v').

2. If v' is such that $\rho(x', v') < \infty$, let $x^* \in \partial_x^{\mathsf{F}} e(x', v')$ be arbitrary but fixed. By Proposition 3.5 and the chain rule, we may identify a sequence $t_n \in T$, $x_n \to x'$ such that $\rho_{t_n}(x_n, v') \to \rho(x', v')$ and such that

$$x^* = -f_{\mathcal{R}}(\rho(x', v')) \lim_{n\to\infty} \frac{\nabla_x g_{t_n}(x_n, \rho_{t_n}(x_n, v')Lv')}{\langle \nabla_z g_{t_n}(x_n, \rho_{t_n}(x_n, v')Lv'), Lv' \rangle}.$$

Now let us first recall that for each fixed n,

$$\begin{aligned}\langle \nabla_z g_{t_n}(x_n, \rho_{t_n}(x_n, v')Lv'), Lv' \rangle &\geq -\frac{g_{t_n}(x_n, 0)}{\rho_{t_n}(x_n, v')} \geq -\frac{g(x_n, 0)}{\rho_{t_n}(x_n, v')} \\ &\geq \frac{\eta}{\rho_{t_n}(x_n, v')}\end{aligned}$$

as a result of [68, Lemma 3.1(3)] (see also [27, Lemma 3]).

For an arbitrary $h \in C_l(x_0)$ (see Definition 3.7), we thus obtain, by employing (2.3),

$$\langle x^*, -h \rangle = \frac{2\pi^{\frac{m}{2}}}{\Gamma(\frac{m}{2})} \rho(x', v')^{m-1} \theta(\rho(x', v')^2) \lim_{n\to\infty} \frac{\langle \nabla_x g_{t_n}(x_n, \rho_{t_n}(x_n, v')Lv'), h \rangle}{\langle \nabla_z g_{t_n}(x_n, \rho_{t_n}(x_n, v')Lv'), Lv' \rangle},$$

so in the particular case that $\theta(\rho(x', v')^2) = 0$, we have $\langle x^*, -h \rangle \leq 0$. Otherwise, if $\theta(\rho(x', v')^2) > 0$ we can continue using the definition of $C_l(x_0)$ and

$$\begin{aligned}\langle x^*, -h \rangle &= \frac{2\pi^{\frac{m}{2}}}{\Gamma(\frac{m}{2})} \rho(x', v')^{m-1} \theta(\rho(x', v')^2) \lim_{n\to\infty} \frac{\langle \nabla_x g_{t_n}(x_n, \rho_{t_n}(x_n, v')Lv'), h \rangle}{\langle \nabla_z g_{t_n}(x_n, \rho_{t_n}(x_n, v')Lv'), Lv' \rangle} \\ &\leq \eta^{-1} \frac{2\pi^{\frac{m}{2}}}{\Gamma(\frac{m}{2})} \rho(x', v')^{m-1} \theta(\rho(x', v')^2) \\ &\qquad \times \lim_{n\to\infty} \rho_{t_n}(x_n, v') \langle \nabla_x g_{t_n}(x_n, \rho_{t_n}(x_n, v')Lv'), h \rangle \\ &\leq \eta^{-1} \frac{2\pi^{\frac{m}{2}}}{\Gamma(\frac{m}{2})} \rho(x', v')^{m-1} \theta(\rho(x', v')^2) \\ &\qquad \times \lim_{n\to\infty} \rho_{t_n}(x_n, v') l \rho_{t_n}(x_n, v')^{-m} \theta^{-1}(\rho_{t_n}(x_n, v')^2) \left\| h \right\| \\ &= \eta^{-1} l \frac{2\pi^{\frac{m}{2}}}{\Gamma(\frac{m}{2})} \left\| h \right\|,\end{aligned}$$

where we used the continuity of $r \mapsto \theta(r)$, and thus the resulting continuity of θ^{-1} (recall that $\theta(\rho(x', v')^2) \neq 0$).

By defining $\tilde{K}_v := l\frac{2\pi^{\frac{m}{2}}}{\eta\Gamma(\frac{m}{2})}$, it follows that

$$\partial^{\mathtt{F}}_x e(x',v') \subseteq \tilde{K}_v\mathbb{B}_* - C_l^*(x_0) \quad \forall (x',v') \in U_v \times V_v. \tag{3.8}$$

Finally, since $\mathbb{S}^{m-1}$ is compact and the family of open sets $\{V_v\}_{v\in\mathbb{S}^{m-1}}$ covers $\mathbb{S}^{m-1}$, we can pick a finite subcover, i.e., a finite number of elements v_i, with $i = 1,\dots,k_l$ such that $\mathbb{S}^{m-1} \subseteq \bigcup_{i=1,\dots,k_l} V_{v_i}$, and define $U_l := \bigcap_{i=1,\dots,k_l} U_{v_i}$ and $K_l := \max\{\tilde{K}_{v_i}, i=1,\dots,k_l\}$, so that by (3.7) and (3.8), the inclusion (3.5) holds. ◻

We are now ready to provide the first main result in the study of the subdifferential of the probability function (1.1).

THEOREM 3.9. *Let $\xi \in \mathbb{R}^m$ be an elliptical symmetrically distributed random vector with mean 0, correlation matrix $R = LL^{\mathsf{T}}$ and continuous generator. Consider the probability function $\varphi : X \to [0,1]$, where X is a (separable) reflexive Banach space defined as*

$$\varphi(x) = \mathbb{P}[g_t(x,\xi) \le 0 \ \ \forall t \in T], \tag{3.9}$$

where $g_t : X \times \mathbb{R}^m \to \mathbb{R}$ are continuously differentiable maps convex in the second argument and T is an arbitrary index set.

Then let $\bar{x} \in X$ be such that

1. *a neighborhood U of $\bar{x}$ can be found such that $g_{|U\times\mathbb{R}^m}$ is finite valued and $\sup_{t\in T} g_t(x',0) < 0$ for all $x' \in U$;*
2. *the set $\{(x,z)\colon g(x,z) = 0\}$ is closed in $U \times \mathbb{R}^m$;*
3. *the set-valued map $\mathcal{S}$ of (3.1) is locally bounded at $\bar{x}$, $v \in \mathbb{S}^{m-1}$ such that $\rho(\bar{x},v) < \infty$, where ρ is as in Proposition 2.6;*
4. *either there exists $l > 0$ such that $C_l(\bar{x})$ has nonempty interior, or $M(\bar{x}) := \{z \in \mathbb{R}^m\colon g(\bar{x},z) \le 0\}$ is bounded.*

Then the following formulae hold true:

$$\partial^{\mathtt{M}}\varphi(\bar{x}) \subseteq \mathrm{cl}^* \left\{ \int_{v\in\mathbb{S}^{m-1}} \partial^{\mathtt{M}}_x e(\bar{x},v)\,d\mu_\zeta(v) - C_l^*(\bar{x}) \right\}, \tag{3.10a}$$

$$\partial^{\infty}\varphi(\bar{x}) \subseteq -C_l^*(\bar{x}), \tag{3.10b}$$

$$\partial^{\mathtt{C}}\varphi(\bar{x}) \subseteq \overline{\mathrm{co}} \left\{ \int_{v\in\mathbb{S}^{m-1}} \partial^{\mathtt{M}}_x e(\bar{x},v)\,d\mu_\zeta(v) - C_l^*(\bar{x}) \right\}. \tag{3.10c}$$

Provided that X is finite dimensional,

$$\partial^{\mathtt{M}}\varphi(\bar{x}) \subseteq \int_{v\in\mathbb{S}^{m-1}} \partial^{\mathtt{M}}_x e(\bar{x},v)\,d\mu_\zeta(v) - C_l^*(\bar{x}), \tag{3.10d}$$

where $\partial^{\mathtt{M}}$, $\partial^{\mathtt{C}}$, and ∂^{∞} respectively refer to the limiting (or Kruger–Mordukhovich), the Clarke, and (limiting) singular subdifferential sets of a map. Moreover, the set $C_l^*(\bar{x})$ can be replaced by $\{0\}$ if $M(\bar{x})$ is bounded.

Finally, for every $v \in F(\bar{x}) = \mathcal{D}\mathrm{om}(\rho(x,\cdot))$,

$$\partial^{\mathtt{M}}_x e(\bar{x},v) \subseteq f_{\mathcal{R}}(\rho(\bar{x},v))\mathcal{S}^{\mathtt{M}}(\bar{x},v).$$

Proof. At any $x \in U \subseteq D^{\circ}$, as a result of Proposition 2.6, it follows that $\varphi(x) = \int_{v \in \mathbb{S}^{m-1}} e(x,v) d\mu_{\zeta}(v)$ with $e(x,v) = F_{\mathcal{R}}(\rho(x,v))$. The application of Proposition 3.8 entails (3.5).

We may now invoke Proposition 3 from [27] (see also [15, Corollary 4.9]), to establish the asserted formulae (3.10), which can be made precise by recalling the chain rule of Lemma 3.6. ∎

Let us provide a careful discussion of the various hypotheses of Theorem 3.9. First, let us remark that the general assumption $\bar{x} \in D$ is not very restrictive.

Remark 3.10. Let us discuss the first hypothesis of Theorem 3.9. It follows from [73, Corollary 2.1] that whenever $\varphi(\bar{x}) > \frac{1}{2}$, 0 belongs to the topological interior of the convex set $\{z \in \mathbb{R}^m \colon \sup_{t \in T} g_t(\bar{x}, z) \leq 0\}$. For such $\bar{x}$, we thus have $g(\bar{x}, 0) \leq 0$ and even $\bar{x} \in D$. This follows since $\mathbb{P}[g_t(x,\xi) \leq 0] \geq \mathbb{P}[\sup_{t \in T} g_t(x, \xi) \leq 0]$ so that $g_t(\bar{x}, 0) < 0$ must hold if each g_t is sufficiently regular for $\{z \in \mathbb{R}^m \colon g_t(x,z) = 0\}$ to indeed be the boundary of the set $\{z \in \mathbb{R}^m \colon g_t(x,z) \leq 0\}$.

Now if the active index set $T_0^g(\bar{x})$ is nonempty, it follows too that $\bar{x} \in D^{\circ}$. If g is moreover u.s.c. (and thus continuous, since it is automatically l.s.c. as the supremum of a family of continuous maps), $g(x,0) < 0$ can be assumed to hold in a neighborhood of $\bar{x}$.

3.4. Concrete conditions through uniform local Lipschitzian families. In what follows, we will show that when *the family of functions g_t is uniformly locally Lipschitzian at $\bar{x}$*, the various assumptions of Theorem 3.9 hold true. We will first provide a formal definition and some characterizations before focussing on showing how conditions 1–3 of Theorem 3.9 are entailed.

DEFINITION 3.11. *The family of functions $\{g_t\}_{t \in T}$ is said to be uniformly locally Lipschitzian at $\bar{x}$ provided that there exists a neighborhood U of $\bar{x}$ such that for every $\bar{z} \in \mathbb{R}^m$ there exists $\gamma > 0$ and $K \geq 0$ such that*

$$|g_t(x,w) - g_t(y,z)| \leq K\left(\|x-y\| + \|w-z\|\right) \quad \forall (x,w),(y,z) \in U \times \mathbb{B}_{\gamma}(\bar{z}),\ \forall t \in T. \tag{3.11}$$

Let us establish the following characterization of the above property in terms of boundedness of the derivatives $\nabla g_t(x,z)$.

PROPOSITION 3.12. *The family of functions $\{g_t\}_{t \in T}$ is uniformly locally Lipschitzian at $\bar{x}$ if and only if there exists a neighborhood U of $\bar{x}$ such that for every $\bar{z} \in \mathbb{R}^m$, the following limit holds:*

$$\limsup_{z \to \bar{z}} \sup\{\|\nabla g_t(x,z)\| \mid x \in U,\ t \in T\} < \infty. \tag{3.12}$$

Proof. Let us assume that $\{g_t\}_{t \in T}$ is uniformly locally Lipschitzian at $\bar{x}$. Then (3.11) implies that each g_t is Lipschitz continuous on $U \times \mathbb{B}_{\gamma}(\bar{z})$ with constant K, so $\|\nabla g_t(x,w)\| \leq K$ for all $t \in T$ and $(x,w) \in \operatorname{int} U \times \mathbb{B}_{\gamma/2}(\bar{z})$ (see, e.g., [48, Theorem 3.52]), which implies (3.12). Now assuming that (3.12) holds, we can take $K > 0$ and $\gamma > 0$ such that $\sup_{t \in T} \|\nabla g_t(x,z)\| \leq K$ for all $(x,z) \in U \times \mathbb{B}_{\gamma}(\bar{z})$, so by [48, Theorem 3.52] we have that (3.11) holds over $\operatorname{int} U \times \mathbb{B}_{\gamma/2}(\bar{z})$. ∎

Using a classical argument of compactness we can establish the following proposition about the uniformity of the Lipschitz constant in (3.11) over compact convex sets.

PROPOSITION 3.13. *Assume that the family of functions $\{g_t\}_{t \in T}$ is uniformly locally Lipschitzian at $\bar{x}$. Then there exists a neighborhood U of $\bar{x}$ such that for every* *convex compact set $W \subseteq \mathbb{R}^m$ there is $K \geq 0$ such that*

(3.13)
$$|g_t(x,w) - g_t(y,z)| \leq K\left(\|x-y\| + \|w-z\|\right) \quad \forall (x,w),(y,z) \in U \times W,\ \forall t \in T.$$

Proof. Consider U as in Definition 3.11. If needed, we may shrink U, so that U could be assumed convex to begin with. Indeed the (reflexive) Banach space is a locally convex topological vector space. For each $\bar{z} \in W$ pick $K_{\bar{z}}$ and $\gamma_{\bar{z}}$ such that (3.11) holds. Then, the family of sets $\operatorname{int}(\mathbb{B}_{\gamma_{\bar{z}}}(\bar{z}))$ covers W, and from its compactness we can take a finite number of $\bar{z}_i$ with $i = 1, \ldots, p$ such that $W \subseteq \bigcup_{i=1}^p \operatorname{int}(\mathbb{B}_{\gamma_{\bar{z}_i}}(\bar{z}_i))$. Now consider $K := \max\{K_{\bar{z}_i} \colon i = 1, \ldots, p\}$. Fixing $t \in T$, we claim that g_t is Lipschitz continuous over $U \times W$ with constant K, and to prove this we simply emulate [12, Theorem 7.3]. Indeed, consider $(x_1,z_1),(x_2,z_2) \in U \times W$. Since the line $[z_1,z_2]$ is contained in $W \subseteq \bigcup_{i=1}^p \operatorname{int}(\mathbb{B}_{\gamma_{\bar{z}_i}}(\bar{z}_i))$ we can find points $t_i \in [0,1]$ for $j = 0, \ldots, \tau$ with $t_0 = 0$, $t_\tau = 1$, and $t_i < t_{i+1}$ such that $z_j, z_{j+1} \in \operatorname{int}(\mathbb{B}_{\gamma_{\bar{z}_{i_j}}}(\bar{z}_{i_j}))$ for some $\bar{z}_{i_j}$, where $z_j := z_1 + t_j(z_2 - z_1)$. Then defining $x_j := x_1 + t_j(x_2 - x_1) \in U$ (recall that U is convex),

$$\begin{aligned}
|g_t(x_1,z_1) - g_t(x_2,z_2)| &\leq \sum_{j=0}^{\tau-1} |g_t(x_j,z_j) - g_t(x_{j+1},z_{j+1})| \\
&\leq \sum_{j=0}^{\tau-1} K\big(\|x_j - x_{j+1}\| + \|z_j - z_{j+1}\|\big) \\
&\leq \sum_{j=0}^{\tau-1} (t_{j+1} - t_j) K\big(\|x_1 - x_2\| + \|z_1 - z_2\|\big) \\
&= K\big(\|x_1 - x_2\| + \|z_1 - z_2\|\big). \qquad \square
\end{aligned}$$

With the help of this last result we can readily identify a concrete situation wherein the family is readily seen to be uniformly locally Lipschitzian.

COROLLARY 3.14. *For an arbitrary index set T, consider for t the maps $g_t(x,z) = h_1(x) + h_t(z)$, where h_1 and h_t are continuously differentiable functions and each h_t is convex. If the function $z \to \sup_{t \to T} \|\nabla h_t(z)\|$ is bounded on compact sets, it follows that $\{g_t\}_{t \in T}$ is uniformly locally Lipschitzian at each $\bar{x} \in X$.*

Let us now show how condition 2 of Theorem 3.9 follows from $\{g_t\}_{t \in T}$ being uniformly locally Lipschitzian at a given point.

PROPOSITION 3.15. *Let us assume that $\{g_t\}_{t \in T}$ is uniformly locally Lipschitzian at $\bar{x}$. Then there exists a neighborhood U of $\bar{x}$ such that the set*

$$\mathfrak{K} = \{(x,z) \in U \times \mathbb{R}^m \colon g(x,z) = 0\}$$

is closed.

Proof. Consider U as in Definition 3.11. Then pick any sequence

$$\{(x,z) \in U \times \mathbb{R}^m \colon g(x,z) = 0\} \ni (x_n,z_n) \to (x,\bar{z}) \in U \times \mathbb{R}^m.$$

On the one hand, by the lower semicontinuity of g we have that $g(x,\bar{z}) \leq 0$. On the other hand, take $\gamma > 0$ and $K > 0$ such that (3.11) holds, and consider $t_n \in T$ such that $g_{t_n}(x_n, z_n) \geq g(x_n, z_n) - \frac{1}{n} = -\frac{1}{n}$. Then, for n large enough,

$$\begin{aligned} g(x, \bar{z}) \geq g_{t_n}(x, \bar{z}) &\geq -K(\|x - x_n\| + \|z_n - \bar{z}\|) + g_{t_n}(x_n, z_n) \\ &\geq -K(\|x - x_n\| + \|z_n - \bar{z}\|) - \frac{1}{n} + g(x_n, z_n) \\ &= -K(\|x - x_n\| + \|z_n - \bar{z}\|) - \frac{1}{n}, \end{aligned}$$

since $(x_n, z_n) \in \mathfrak{K}$. So taking the limit we get that $g(x, \bar{z}) \geq 0$, i.e., $(x, \bar{z}) \in \mathfrak{K}$. ∎

The two conditions of item 1 in Theorem 3.9 also follow, as shown in the subsequent pair of propositions.

Proposition 3.16. *Assume that $\{g_t\}_{t\in T}$ is uniformly locally Lipschitzian at $\bar{x}$ and suppose that $g(\bar{x}, 0)$ is finite. Then there exists $U \in \mathcal{N}_{\bar{x}}$ such that for every $r \geq 0$ there exists K such that*

$$|g(x, w) - g(y, z)| \leq K (\|x - y\| + \|w - z\|) \quad \forall (x, w), (y, z) \in U \times W,\ \forall t \in T. \tag{3.14}$$

Consequently $g_{U\times\mathbb{R}^m}$ is finite valued.

Proof. Consider $W := r\mathbb{B}$. Then by Proposition 3.13 there exists $K \geq 0$ such that (3.13) holds. In particular,

$$\begin{aligned} g_t(x, w) &\leq K (\|x - \bar{x}\| + \|w\|) + g(\bar{x}, 0) \quad \forall (x, w) \in U \times W,\ \forall t \in T, \\ g_t(y, z) &\leq K (\|x - y\| + \|w - z\|) + g_t(x, w) \quad \forall (x, w), (x, z) \in U \times W,\ \forall t \in T. \end{aligned}$$

Then, taking the supremum over T we get from the first inequality that g is finite valued and from the second that (3.14) holds. ∎

Proposition 3.17. *Assume that $\{g_t\}_{t\in T}$ is uniformly locally Lipschitzian at $\bar{x}$ and $g(\bar{x}, 0) < 0$. Then there exists $U \in \mathcal{N}_{\bar{x}}$ such that $g(x, 0) < 0$ for all $x \in U$.*

Proof. By Proposition 3.16 we get that g is continuous at $(\bar{x}, 0)$ and thus the result follows. ∎

Key condition 3 of Theorem 3.9 can also be shown to hold whenever the family $\{g_t\}_{t\in T}$ is uniformly locally Lipschitzian.

Proposition 3.18. *Assume that $\{g_t\}_{t\in T}$ is uniformly locally Lipschitzian at $\bar{x}$ and $g(\bar{x}, 0) < 0$. Then, for every $\bar{v} \in \mathcal{D}\mathrm{om}(\rho(\bar{x}, \cdot))$ the mapping $\mathcal{S}$ is locally bounded at $(\bar{x}, \bar{v})$. Consequently, for every $\bar{v} \in \mathcal{D}\mathrm{om}(\rho(\bar{x}, \cdot))$ there exist $\varepsilon > 0$ and $K \geq 0$ such that*

$$|\rho(x, \bar{v}) - \rho(y, \bar{v})| \leq K\|x - y\| \quad \forall x, y \in \mathbb{B}_\varepsilon(\bar{x}),\ \forall v \in \mathbb{B}_\varepsilon(\bar{v}) \tag{3.15}$$

Proof. Consider $\bar{z} = \rho(\bar{x}, \bar{v})L\bar{v}$. Then consider U, $K > 0$, and γ such that

$$|g_t(x, w) - g_t(y, z)| \leq K (\|x - y\| + \|w - z\|) \quad \forall (x, w), (y, z) \in U \times \mathbb{B}_\gamma(\bar{z}),\ \forall t \in T. \tag{3.16}$$

By Remark 3.3, $g(x, 0) \leq -\nu < 0$ for every $x \in U$ can be assumed. By Propositions 3.15, 3.16, 3.17, and Proposition 2.6, the hypotheses of Proposition 3.2 are satisfied. Hence, there exists a neighborhood V of $\bar{v}$ such that $\rho\colon U \times V \to \mathbb{R}$ is continuous, and shrinking the neighborhoods U and V if needed such that U and V are open, it holds that

$$\rho(x,v)Lv \in \mathbb{B}_\gamma(\bar{z}) \text{ for every } (x,v) \in U \times V \text{ and } M := \sup_{(x',v')\in U\times V} \rho(x,v) < +\infty.$$

Now pick $(x,v) \in U \times V$ and $x^* \in \mathcal{S}(x,v)$. Then by the definition of $\mathcal{S}$ there exist $t_n \in T$ and $x_n \to x$ with $\rho_{t_n}(x_n,v) \to \rho(x,v)$ such that

$$x^* = -\lim_{n\to\infty} \frac{\nabla_x g_{t_n}(x_n, \rho_{t_n}(x_n,v)Lv)}{\langle \nabla_z g_{t_n}(x_n, \rho_{t_n}(x_n,v)Lv), Lv\rangle},$$

and by [27, Lemma 3] we have that

$$\frac{1}{\langle \nabla_z g_{t_n}(x_n, \rho_{t_n}(x_n,v)Lv), Lv\rangle} \le -\frac{\rho_{t_n}(x_n,v)}{g_{t_n}(x_n,0)} \le \frac{\rho_{t_n}(x_n,v)}{\nu} \le \frac{M+1}{\nu},$$

where (w.l.o.g.) we have assumed that $\rho_{t_n}(x_n,v) \le M+1$ for all n. Moreover, by (3.16) and Proposition 3.12 we have

$$\|\nabla_x g_{t_n}(x_n, \rho_{t_n}(x_n,v)Lv)\| \le K.$$

Consequently $\|x^*\| \le \frac{K(M+1)}{\nu}$.

Finally, since $\mathcal{S}$ is an upper estimation for the Fréchet subdifferential of ρ we also have that it is bounded locally. Then by [48, Theorem 3.52] we have that (3.15) holds. ∎

4. Generalized subdifferentiation: Tighter formulae. In this section we will improve the formulae given in Theorem 3.9 by providing a, perhaps, smaller outer estimate of $\partial^{\mathtt{M}}_x \rho(x,v)$. The first results given in section 4.1 exploit a notion known as equicontinuous subdifferentiability and are related to requesting additional structure on the family $\{g_t\}_{t\in T}$, whereas the second set of results given in section 4.2 require further regularity on the set T.

4.1. Improved formulae through equicontinuous subdifferentiability. The working tool in this section will be a notion called *strong equicontinuous subdifferentiability*. Let us first provide the original definition of *equicontinuous subdifferentiability* from [50], which motivates our notion.

DEFINITION 4.1 (Definition 3.4 of [50]). *Let $f_t : X \to \mathbb{R}\cup\{\infty\}$ be a family of l.s.c. functions indexed by $t \in T$. The family is called equicontinuously subdifferentiable at $\bar{x} \in X$ if for any weak-* neighborhood V^* of the origin in X^* there is some $\varepsilon > 0$ such that*

$$\partial^{\mathtt{M}} f_t(x) \subseteq \partial^{\mathtt{M}} f_t(\bar{x}) + V^* \tag{4.1}$$

for all $t \in T_\varepsilon(\bar{x})$, $x \in \mathbb{B}_\varepsilon(\bar{x})$, where T_ε refers to the ε-active index set related to the supremum function of the family f_t.

Roughly speaking, this notion involves some *uniform continuity* over the subgradients of the data functions f_t for points close to the active index set. This definition fits perfectly with the work [50], because there the authors gave *fuzzy calculus rules* to characterize the Kruger–Mordukhovich subdifferential of the supremum function f at $\bar{x}$ involving only indexes $t \in T_\varepsilon(\bar{x})$ for arbitrary small $\varepsilon > 0$. Nevertheless, in our *fuzzy calculus* (see, e.g., Proposition 3.5) we obtain an upper estimation which considers index $t \in T_\varepsilon(x, \bar{v})$ for points arbitrarily close to $\bar{x}$ and $\varepsilon > 0$ arbitrarily small. The contrast about the index set in the upper estimation between [50] and our work is due to the fact that in [50] the authors assume that the functions f_t are uniformly locally Lipschitzian, but here the functions ρ_t (at least not directly) are not necessarily uniformly locally Lipschitzian *even when the functions g_t are locally Lipschitzian.*

Let us introduce our definition of *strong equicontinuous subdifferentiability.*

DEFINITION 4.2. *Let $f_t : X \to \mathbb{R} \cup \{\infty\}$ be a family of l.s.c. functions indexed by $t \in T$. The supremum function f of the family $\{f_t\}_{t\in T}$ is called strongly equicontinuously subdifferentiable at $\bar{x} \in X$ if for any weak-* neighborhood V^* of the origin in X^* there is some $\varepsilon > 0$ such that*

$$\partial^{\mathtt{M}} f_t(x) \subseteq \partial^{\mathtt{M}} f_t(\bar{x}) + V^* \tag{4.2}$$

for all $t \in T_\varepsilon(x)$, $x \in \mathbb{B}_\varepsilon(\bar{x})$ with $|f_t(x) - f(\bar{x})| \le \varepsilon$, where $T_\varepsilon(x)$ refers to the ε-active index set related to the supremum function of the family f_t.

Remark 4.3. It is worth mentioning that under the uniform locally Lipschitzian continuity of f_t at $\bar{x}$, there exist $K \ge 0$ and $\eta > 0$ such that

$$|f_t(y) - f_t(z)| \le K\|y - z\| \quad \forall y, z \in \mathbb{B}(\bar{x}, \eta),\ \forall t \in T.$$

Definitions 4.1 and 4.2 are equivalent. Since evidently Definition 4.2 implies Definition 4.1, let us turn our attention to the opposite case. Indeed, let V^* be given, and ε' be such that (4.2) holds for all $t \in T_{\varepsilon'}(\bar{x})$ and $x \in \mathbb{B}_{\varepsilon'}(\bar{x})$. Let $\eta \le \varepsilon'$ (shrunken if needed) be as above and $\varepsilon > 0$ such that $(K + 1)\varepsilon \le \varepsilon'$. Then for all $x \in \mathbb{B}_\eta(\bar{x})$, any $t \in T_\varepsilon(x)$ with $|f_t(x) - f(\bar{x})| \le \varepsilon$ satisfies

$$|f(\bar{x}) - f_t(\bar{x})| \le |f_t(\bar{x}) - f_t(x)| + |f_t(x) - f(\bar{x})| \le K\varepsilon + \varepsilon = (K + 1)\varepsilon,$$

i.e., $t \in T_{(K+1)\varepsilon}(\bar{x}) \subseteq T_{\varepsilon'}(\bar{x})$. Hence, (4.2) holds for all $t \in T_\varepsilon(x)$, $x \in \mathbb{B}_\varepsilon(\bar{x})$ with $|f_t(x) - f(\bar{x})| \le \varepsilon$.

We will require this strong equicontinuous subdifferentiability (see also [49, 50]) for the family of mappings ρ_t defined in Lemma 2.4 and Proposition 2.6. The following proposition intends to provide conditions entailing strong equicontinuous subdifferentiability.

PROPOSITION 4.4. *Let the family $\{f_t\}_{t\in T}$ be given and assume that each member $f_t : X \to \mathbb{R}$ is continuously differentiable. Then $\{f_t\}_{t\in T}$ is* strongly equicontinuously subdifferentiable *at $\bar{x}$ if and only if, for every $h \in X$,*

$$\inf_{\varepsilon>0} \sup \left\{ |\langle \nabla f_t(x) - \nabla f_t(\bar{x}), h\rangle| \;\middle|\; \begin{array}{c} x \in \mathbb{B}(\bar{x}, \varepsilon),\ t \in T_\varepsilon(x) \\ \textit{and } |f_t(x) - f(\bar{x})| \le \varepsilon \end{array} \right\} = 0. \tag{4.3}$$

Proof. Assume that the family is strongly equicontinuously subdifferentiable, and consider $\nu > 0$ and $h \in X$. Then, define the weak-* neighborhood V^* as

$$V^* := \{x^* \in X^* \mid |\langle x^*, h\rangle| \le \varepsilon\}.$$

By Definition 4.2 there exists $\varepsilon > 0$ such that for all $x \in \mathbb{B}(\bar{x}, \varepsilon)$, $t \in T_\varepsilon(x)$ with $|f_t(x) - f(\bar{x})| \le \varepsilon$ we have (recall (4.2)) $\nabla f_t(x) - \nabla f_t(\bar{x}) \in V^*$, which means

$$\sup \left\{ |\langle \nabla f_t(x) - \nabla f_t(\bar{x}), h\rangle| \;\middle|\; \begin{array}{c} x \in \mathbb{B}(\bar{x}, \varepsilon),\ t \in T_\varepsilon(x) \\ \text{and } |f_t(x) - f(\bar{x})| \le \varepsilon \end{array} \right\} \le \nu.$$

From the arbitrariness of $\nu > 0$ we get (4.3).

Now assume that (4.3) holds and consider an arbitrary weak-* neighborhood V^* of the origin in X^*. In order to prove (4.2), we may also assume that

$$V^* \supseteq \{x^* \mid |\langle x^*, h_i \rangle\rangle| \leq \eta \ \forall i = 1, \ldots, p\}$$

with $h_i \in X$ and $p \in \mathbb{N}$. Then using (4.3) we can take $\varepsilon > 0$ such that, for all $i = 1, \ldots, p$,

$$\sup\left\{ |\langle \nabla f_t(x) - \nabla f_t(\bar{x}), h_i \rangle| \;\middle|\; \begin{array}{c} x \in \mathbb{B}(\bar{x}, \varepsilon),\ t \in T_\varepsilon(x) \\ \text{and } |f_t(x) - f(\bar{x})| \leq \varepsilon \end{array} \right\} \leq \eta,$$

meaning that $\nabla f_t(x) \in \nabla f_t(\bar{x}) + V^*$ for all $x \in \mathbb{B}(\bar{x}, \varepsilon),\ t \in T(x)$ with $|f_t(x) - f(\bar{x})| \leq \varepsilon$, and so (4.2) holds. ∎

PROPOSITION 4.5. *Consider the mappings* $\rho_t : D_t \times \mathbb{S}^{m-1} \to \mathbb{R}_+ \cup \{\infty\}$ *as defined in Proposition* 2.6 *for* $t \in T$, *let* $(\bar{x}, \bar{v}) \in D \times \mathcal{D}\mathrm{om}(\rho(\bar{x}, \cdot))$ *be given, and assume that* $\bar{x} \in \mathrm{int}(D)$. *Then the family of maps* $\{\rho_t(\cdot, \bar{v})\}_{t \in T}$ *is strongly equicontinuously subdifferentiable at* $\bar{x}$ *if either one of the following conditions holds:*

(sc1) *The functions* $\rho_t(\cdot, \bar{v})$ *are uniformly strictly differentiable at* $\bar{x}$, *that is,* $\lim_{\eta \downarrow 0} r(\eta) = 0$, *where*

$$r(\eta) := \sup\left\{ \frac{|\rho_t(x, \bar{v}) - \rho_t(x', \bar{v}) - \langle \nabla_x \rho_t(\bar{x}, \bar{v}), x - x' \rangle|}{\|x - x'\|} : \begin{array}{c} x, x' \in \mathbb{B}_\eta(\bar{x}), \\ x \neq x',\ t \in T_\eta(x, \bar{v}) \\ \textit{with } |\rho_t(x, \bar{v}) - \rho(\bar{x}, \bar{v})| \leq \eta \end{array} \right\}.$$

(sc2) *The set* T *is a compact metric space and the function* $(t, x) \in T \times (X, \|\cdot\|) \to \nabla_x \rho_t(x, \bar{v}) \in (X^*, \|\cdot\|)$ *is continuous.*

Proof. Take $\gamma > 0$ such that $\mathbb{B}_\gamma(\bar{x}) \subseteq U$. Then $x \in D_t$ for each $t \in T$, and by Lemma 2.4 we know that ρ_t is continuously differentiable at x provided that $\rho_t(x, \bar{v}) < +\infty$. Then fix $h \in X \backslash \{0\}$. Consider $\nu > 0$ and take $\eta \in (0, \gamma)$ such that $r(\eta) \leq \nu$. Pick $x \in \mathbb{B}(\bar{x}, \eta/2)$ and $t \in T_{\frac{\eta}{2}}(x, \bar{v})$ with $|\rho_t(x, \bar{v}) - \rho(\bar{x}, \bar{v})| \leq \eta/2$ so that we have

$$\begin{aligned} |\langle \nabla \rho_t(x, \bar{v}) - \nabla \rho_t(\bar{x}, \bar{v}), h \rangle| &= \|h\| \lim_{\substack{s \downarrow 0, \\ s \in (0, \frac{\eta \|h\|}{2})}} \frac{\rho_t(x + sh, \bar{v}) - \rho_t(x, \bar{v}) - \langle \nabla \rho_t(\bar{x}, \bar{v}), sh \rangle}{s\|h\|} \\ &\leq \|h\| r(\eta) \leq \nu \|h\|. \end{aligned} \tag{4.4}$$

From the arbitrariness of ν we get that (4.3) holds for $\{\rho_t(\cdot, \bar{v})\}_{t \in T}$.

In the second situation, we will show that the condition entails (sc1) and consequently the result as already established. Let us show this by contradiction. If (sc1) does not hold, then for any $\tau > 0$, we can find some $\gamma > 0$, and sequences $t_n \in T$, $\eta_n \downarrow 0$, $x_n \neq x'_n \in \mathbb{B}_{\eta_n}(\bar{x})$ such that

$$\frac{|\rho_{t_n}(x_n, \bar{v}) - \rho_{t_n}(x'_n, \bar{v}) - \langle \nabla_x \rho_{t_n}(\bar{x}, \bar{v}), x_n - x'_n \rangle|}{\|x_n - x'_n\|} \geq \gamma - \frac{1}{n}$$

for n sufficiently large. Hence, by applying the classical mean value theorem, we can find $\theta_n \in [x_n, x_n']$ such that

$$\begin{aligned}\frac{\gamma}{2} &< \frac{|\langle \nabla_x \rho_{t_n}(\theta_n, \bar{v}), x_n - x_n'\rangle - \langle \nabla_x \rho_{t_n}(\bar{x}, \bar{v}), x_n - x_n'\rangle|}{\|x_n - x_n'\|} \\ &\leq \|\nabla \rho_{t_n}(\theta_n, \bar{v}) - \nabla_x \rho_{t_n}(\bar{x}, \bar{v})\|\end{aligned}$$

for n sufficiently large. But this contradicts the continuity of the map $(t, x) \in T \times (X, \|\cdot\|) \to \nabla_x \rho_t(x, \bar{v}) \in (X^*, \|\cdot\|_*)$. ∎

Thanks to the tool of strong equicontinuous subdifferentiability, we can provide an improved version of Proposition 3.5.

PROPOSITION 4.6. *Consider the mapping $\rho : D \times \mathbb{S}^{m-1} \to \mathbb{R}_+ \cup \{\infty\}$ of Proposition* 3.5. *Assume moreover that the set $\{(x, z) \in X \times \mathbb{R}^m : g(x, z) = 0\}$ is closed in $X \times \mathbb{R}^m$. Let $\bar{x} \in U \subseteq D^\circ$ with neighborhood U of $\bar{x}$ be given together with $v \in \mathcal{D}\mathrm{om}(\rho(\bar{x}, \cdot))$. Assume that the family of mappings $\{\rho_t(\cdot, v)\}_{t \in T}$ is strongly equicontinuously subdifferentiable at $\bar{x}$. Then the following inclusion is true:*

$$\partial_x^{\mathtt{M}} \rho(\bar{x}, v) \subseteq \bigcap_{\varepsilon > 0} \mathrm{cl}^{\mathrm{w}^*} \left\{ -\frac{\nabla_x g_t(\bar{x}, \rho_t(\bar{x}, v)Lv)}{\langle \nabla_z g_t(\bar{x}, \rho_t(\bar{x}, v)Lv), Lv\rangle} : \begin{array}{c} x \in \mathbb{B}(\bar{x}, v),\ t \in T_\varepsilon^\rho(x, v) \\ \text{with } |\rho_t(x, v) - \rho(\bar{x}, v)| \leq \varepsilon \end{array} \right\}. \tag{4.5}$$

Proof. Under the given assumptions, Proposition 3.5 yields (3.2a). Therefore for some $x^* \in \partial_x^{\mathtt{M}} \rho(\bar{x}, v)$ we can find a sequence $t_n \in T$ and a sequence $x_n \to \bar{x}$ such that $\rho_{t_n}(x_n, v) \to \rho(\bar{x}, v)$ and

$$x^* = \lim_{n \to \infty} \nabla_x \rho_{t_n}(x_n, v) = \lim_{n \to \infty} -\frac{\nabla_x g_{t_n}(x_n, \rho_{t_n}(x_n, v)Lv)}{\langle \nabla_z g_{t_n}(x_n, \rho_{t_n}(x_n, v)Lv), Lv\rangle}. \tag{4.6}$$

Let us first show that for each $\varepsilon > 0$, we may find N such that $t_n \in T_\varepsilon^\rho(\bar{x}, v)$ for all $n \geq N$. To this end let us recall that the sequence t_n came about by an application of the Borwein–Preiss variational principle (see [5]), i.e., through Lemma 2.1, so $\rho_{t_n}(\bar{x}, v) \to \rho(\bar{x}, v)$ also holds, and this immediately entails the claim as $\rho(\bar{x}, v) \leq \rho_t(\bar{x}, v)$ is evident.

By strong equicontinuous subdifferentiability of the resolvant family, we can find for $\eta > 0$ an $\varepsilon > 0$ such that $|\langle \nabla_x \rho_t(x, v) - \nabla_x \rho_t(\bar{x}, v), h\rangle| \leq \eta \|h\|$ for all $t \in T_\varepsilon^\rho(\bar{x}, v)$ and all $x \in \mathbb{B}_\varepsilon(\bar{x})$ as a result of (4.4). By the above claim, for n large enough $t_n \in T_\varepsilon(\bar{x}, v)$ and $x_n \in B_\varepsilon(\bar{x})$ so that $|\langle \nabla_x \rho_{t_n}(x_n, v) - \nabla_x \rho_{t_n}(\bar{x}, v), h\rangle| \leq \eta \|h\|$. Moreover, for n sufficiently large, $|\langle \nabla_x \rho_{t_n}(x_n, v) - x^*, h\rangle| \leq \eta \|h\|$ as a result of (4.6). Combining these relationships through the triangle inequality, we get $|\langle \nabla_x \rho_{t_n}(\bar{x}, v) - x^*, h\rangle| \leq 2\eta \|h\|$ with $t_n \in T_\varepsilon^\rho(\bar{x}, v)$. Since η and $\varepsilon > 0$ are arbitrary, the inclusion (4.5) follows. ∎

THEOREM 4.7. *Let $\xi \in \mathbb{R}^m$ be an elliptical symmetrically distributed random vector with mean* 0, *correlation matrix $R = LL^\mathsf{T}$ and continuous generator. Consider the probability function $\varphi : X \to [0, 1]$, where X is a reflexive Banach space defined as*

$$\varphi(x) = \mathbb{P}[g_t(x, \xi) \leq 0 \ \ \forall t \in T], \tag{4.7}$$

where $g_t : X \times \mathbb{R}^m \to \mathbb{R}$ are continuously differentiable maps convex in the second argument and T is an arbitrary index set.

Then let $\bar{x} \in X$ be such that assumptions 1–4 *of Theorem* 3.9 *hold and in addition*

1. *that at any $v \in \mathbb{S}^{m-1}$, the family of resolvant mappings $\{\rho_t(\cdot, v)\}_{t\in T}$ is strongly equicontinuously subdifferentiable at $\bar{x}$.*

Then the formulae (3.10) *for the subdifferentials of φ are valid, wherein we may consider the outer estimate*

$$\partial_x^{\mathtt{M}} e(\bar{x}, v) \subseteq f_{\mathcal{R}}(\rho(\bar{x}, v)) \bigcap_{\varepsilon>0} \mathrm{cl}^{\mathrm{w}^*} \left\{ -\frac{\nabla_x g_t(\bar{x}, \rho_t(\bar{x}, v)Lv)}{\langle \nabla_z g_t(\bar{x}, \rho_t(\bar{x}, v)Lv), Lv\rangle} : \begin{array}{l} x \in \mathbb{B}(\bar{x}, v),\ t \in T_\varepsilon^\rho(x, v) \\ \text{with } |\rho_t(x, v) - \rho(\bar{x}, v)| \leq \varepsilon \end{array} \right\}$$

at $v \in F(\bar{x}) = \mathcal{D}\mathrm{om}(\rho(\bar{x}, v))$.

4.2. Improved formulae through some degree of compactness. In this last part of the work we are going to assume the following.

Assumption 4.8. Let T be a metric space and let there exist a neighborhood U of $\bar{x}$ such that

(a) $g_{|_{U\times\mathbb{R}^m}}$ is finite valued;
(b) $g(x, 0) < 0$ for all $x \in U$;
(c) the function $G\colon T \times U \times \mathbb{R}^m \to X \times X^* \times \mathbb{R}^m$ given by

$$G(t, x, z) = (g_t(x, v), \nabla_x g_t(x, z), \nabla_z g_t(x, z))$$

is continuous;
(d) the active index set $T^g(x, z)$ is nonempty for every

$$(x, z) \in \mathfrak{K} = \{(x, z) \in U \times \mathbb{R}^m \colon g(x, z) = 0\};$$

(e) the set $\bigcup_{(x,z)\in\mathfrak{K}} T^g(x, z)$ is relatively compact.

Remark 4.9. It is worth mentioning that all the above hypotheses are satisfied provided that T is compact, G is continuous, and $g(\bar{x}, 0) < 0$.

Lemma 4.10. *Under Assumption* 4.8 *one has that*

(i) *the set $\mathfrak{K}$ is closed;*
(ii) *for every $T\times U\times\mathbb{S}^{m-1} \ni (t_n, x_n, v_n) \to (t, x, v) \in T\times U\times\mathbb{S}^{m-1}$, $\rho_{t_n}(x_n, v_n) \to \rho_t(x, v)$;*
(iii) *the set $T_\varepsilon^\rho(x, v)$ is closed for every $(x, v) \in U \times \mathbb{S}^{m-1}$.*

Proof. (i) Consider $(x_n, z_n) \in \mathfrak{K}$ such that $(x_n, z_n) \to (x_0, z_0)$. First, by the lower semicontinuity of g we have that $g(x_0, z_0) \leq 0$. Now take $t_n \in T^g(x_n, z_n)$ such that $g(x_n, z_n) = g_{t_n}(x_n, z_n)$. Then up to a subsequence, we may assume that $t_n \to t_0 \in T$. So by the continuity of $(t, x, z) \to g_t(x, z)$ we have that $g_{t_0}(x_0, z_0) = 0$, which implies that $g(x_0, z_0) \geq 0$, and we conclude that $\mathfrak{K}$ is closed.

(ii) Take $(t_n, x_n, v_n) \to (t, x, v)$. Let us first assume that $\{\rho_{t_n}(x_n, v_n)\}$ diverges along some subsequence. Then for some $r > 0$ and n large enough, it holds that $g_{t_n}(x_n, rLv_n) < 0$ by convexity of g_{t_n} in the second argument and $g_{t_n}(x_n, 0) < 0$. By continuity of G, i.e., Assumption 4.8(c), it follows that $g_t(x, rLv) \leq 0$. However, if $g_t(x, rLv) = 0$, it follows from uniqueness of the solution (see Lemma 2.4) that $r = \rho_t(x, v)$. Hence for some $r' > r$, $g_t(x, r'Lv) > 0$ must hold. Yet for n sufficiently large, $g_{t_n}(x_n, r'Lv_n) < 0$ must hold true too. We have now established a contradiction to the continuity of G. Hence $g_t(x, rLv) < 0$ must hold and since r was arbitrary, $\rho_t(x, v) = \infty$.

We may thus assume that the sequence $\{\rho_{t_n}(x_n, v_n)\}$ remains bounded and hence admits a cluster point. By moving to a subsequence if needed, let $r := \lim \rho_{t_n}(x_n, v_n)$ be this cluster point. By continuity of G, $0 = g_{t_n}(x_n, \rho_{t_n}(x_n, v_n)) \to g_t(x, rLv) = 0$, and by the uniqueness of the solution $r = \rho_t(x, v)$. Now for $\underline{r} = \liminf \rho_{t_n}(x_n, v_n)$ we can find a subsequence such that $\underline{r} := \lim \rho_{t_n}(x_n, v_n)$, and hence in particular $g_t(x, \underline{r}Lv) = 0$. Yet convexity in the second argument and $g_t(x, 0) < 0$ imply that $\rho_t(x, v) = \underline{r} \leq \liminf \rho_{t_n}(x_n, v_n)$.

Finally, consider $\bar{r} = \limsup \rho_{t_n}(x_n, v_n)$ and let $\gamma \in \mathbb{R}$ with $\gamma < \bar{r}$. Then there exists $n_\gamma \in \mathbb{N}$ such that $\gamma < \rho_{t_n}(x_n, v_n)\ \ \forall n \geq n_\gamma$. Then by convexity $g_{t_n}(x_n, \gamma Lv_n) < g_{t_n}(x_n, \rho_{t_n}(x_n, v_n)Lv_n)$ for all $n \geq n_\gamma$, and consequently by continuity of the data function $g_t(x, \gamma Lv) \leq 0$, which implies $\rho_t(x, v) \geq \gamma$. From the arbitrariness of γ we get the result.

(iii) By the previously established continuity of $(t, x, v) \mapsto \rho_t(x, v)$ it is easy to see that $T^\rho_\varepsilon(x, v)$ is closed. $\square$

LEMMA 4.11 (relation between active index sets of ρ and g). *Let $\bar{x} \in D$ be given and consider $v \in F(\bar{x}) = \mathcal{D}\mathrm{om}(\rho(\bar{x}, \cdot))$. Then, for every $t \in T$ with $v \in F_t(\bar{x}) = \mathcal{D}\mathrm{om}(\rho_t(\bar{x}, \cdot))$ it holds that*

$$g_t(\bar{x}, \rho(\bar{x}, v)Lv) \geq \big(\rho(\bar{x}, v) - \rho_t(\bar{x}, v)\big)\langle \nabla_z g_t(\bar{x}, \rho_t(\bar{x}, v)Lv), Lv\rangle. \tag{4.8}$$

Let $\gamma > 0$ be arbitrary and define $\eta := g(\bar{x}, (\rho(\bar{x}, v) + 2\gamma)Lv)$. Then $T^\rho_\gamma(\bar{x}, v) \subseteq T^g_\eta(\bar{x}, \rho(\bar{x}, v)Lv)$. Moreover, it always holds that $T^g(\bar{x}, \rho(\bar{x}, v)Lv) = T^\rho(\bar{x}, v)$.

In addition, if the family is uniformly locally Lipschitzian at $\bar{x}$ there exist neighborhoods U of $\bar{x}$ and V of v, and $K \geq 0$ such that $T^\rho_\gamma(x', v') \subseteq T^g_{\eta K}(x', \rho(x', v')Lv')$ for all $x' \in U$ and all $v' \in V$.

More precisely, whenever $x' \in \mathbb{B}_\gamma(\bar{x}) \subseteq U$, $v' \in V$, and $t \in T^\rho_\gamma(x', v')$ with $|\rho_t(x', v') - \rho(\bar{x}, v')| \leq \varepsilon$ we have that $t \in T^g_{K(2+K\|L\|)\gamma}(\bar{x}, \rho(\bar{x}, v')Lv')$.

Proof. Let $t \in T$ and $v \in F_t(\bar{x})$ be arbitrary. Then by convexity of $g_t(\bar{x}, \cdot)$,

$$\begin{aligned} g_t(\bar{x}, \rho(\bar{x}, v)Lv) &\geq \big(\rho(\bar{x}, v) - \rho_t(\bar{x}, v)\big)\langle \nabla_z g_t(\bar{x}, \rho_t(\bar{x}, v)Lv), Lv\rangle + g_t(\bar{x}, \rho_t(\bar{x}, v)Lv) \\ &= \big(\rho(\bar{x}, v) - \rho_t(\bar{x}, v)\big)\langle \nabla_z g_t(\bar{x}, \rho_t(\bar{x}, v)Lv), Lv\rangle, \end{aligned}$$

which proves (4.8). Furthermore, again by convexity of g_t,

$$\begin{aligned} \gamma\langle \nabla_z g_t(\bar{x}, \rho_t(\bar{x}, v)Lv), Lv\rangle &\leq g_t(\bar{x}, (\rho_t(\bar{x}, v) + \gamma)Lv) - g_t(\bar{x}, \rho_t(\bar{x}, v)Lv) \\ &\leq g_t(\bar{x}, (\rho_t(\bar{x}, v) + \gamma)Lv). \end{aligned}$$

Moreover, if t is such that $\rho_t(\bar{x}, v) \leq \rho(\bar{x}, v) + \gamma$, i.e., $t \in T^\rho_\gamma(\bar{x}, v)$, then by definition of g (2.1), we may continue the last estimate as

$$g_t(\bar{x}, (\rho_t(\bar{x}, v) + \gamma)Lv) \leq g(\bar{x}, (\rho_t(\bar{x}, v) + \gamma)Lv) \leq g(\bar{x}, (\rho(\bar{x}, v) + 2\gamma)Lv),$$

since $\rho(\bar{x}, v) \leq \rho_t(\bar{x}, v) \leq \rho_t(\bar{x}, v) + \gamma \leq \rho(\bar{x}, v) + 2\gamma$ and the map $r \mapsto g(\bar{x}, rLv)$ is increasing by convexity in the second argument on $[\rho(\bar{x}, v), \infty)$.

Therefore,

$$-\gamma\langle \nabla_z g_t(\bar{x}, \rho_t(\bar{x}, v)Lv), Lv\rangle \geq -g(\bar{x}, (\rho(\bar{x}, v) + 2\gamma)Lv). \tag{4.9}$$

So (4.8) and (4.9) give for such $t \in T^\rho_\gamma(\bar{x}, v)$, since $g(\bar{x}, \rho(\bar{x}, v)Lv) = 0$,

$$g_t(\bar{x}, \rho(\bar{x}, v)Lv) \geq -g(\bar{x}, (\rho(\bar{x}, v) + 2\gamma)Lv) = g(\bar{x}, \rho(\bar{x}, v)Lv) - \eta,$$

i.e., $t \in T^g_\eta(\bar{x}, \rho(\bar{x}, v)Lv)$, as was to be shown.

Now,

$$\begin{aligned} t \in T^g(\bar{x}, \rho(\bar{x}, v)Lv) &\Leftrightarrow g_t(\bar{x}, \rho(\bar{x}, v)Lv) = 0 = g(\bar{x}, \rho(\bar{x}, v)Lv) \\ \text{(by the uniqueness of the solution for } g_t) &\Leftrightarrow \rho_t(\bar{x}, v) = \rho(\bar{x}, v) \\ &\Leftrightarrow t \in T^\rho(\bar{x}, v). \end{aligned}$$

Assuming that the family is uniformly locally Lipschitzian at $(\bar{x}, \rho(\bar{x}, v)Lv)$, in addition $g(\bar{x}, 0) < 0$ and by Proposition 3.13 we have that the assumptions of Proposition 3.18 hold; in particular, ρ is continuous at $(\bar{x}, v)$. Then, there exist $U \in \mathcal{N}_{\bar{x}}$, $V \in \mathcal{N}_v$, and K_1 such that for all $(x', v'), (x'', v'') \in U \times V$,

$$\begin{aligned} &|g_t(x'', \rho(x'', v'')Lv'') - g_t(x', \rho(x', v')Lv')| \\ &\qquad \le K(\|x' - x''\| + \|\rho(x'', v'')Lv'' - \rho(x', v')Lv'\|) \quad \forall t \in T. \end{aligned} \tag{4.10}$$

In particular the same holds for the function g (see Proposition 3.16), and consequently

$$-g(\bar{x}, (\rho(\bar{x}, v) + 2\gamma)) \ge -2\gamma K_1 \|Lv\| - g(\bar{x}, (\rho(\bar{x}, v)Lv)) \ge -2\gamma K_1 \|L\|. \tag{4.11}$$

Then, defining $K := 2K_1\|L\|$, we have $T^\rho_\gamma(\bar{x}, v) \subseteq T^g_\eta(\bar{x}, \rho(\bar{x}, v)Lv)$ and for every $t \in T^\rho_\gamma(\bar{x}, v)$ it holds that $g_t(x', \rho(x', v')Lv') \ge -g(\bar{x}, (\rho(\bar{x}, v) + 2\gamma)) \ge -\gamma K$.

Finally, by Proposition 3.18, ρ satisfies

$$|\rho(x', v') - \rho(x'', v')| \le K\|\bar{x} - x''\| \quad \forall x', x'' \in U,\ v' \in V, \tag{4.12}$$

upon shrinking the neighborhoods and increasing the constant K if needed.

Then using (4.10) and (4.12) we have for each $x' \in \mathbb{B}_\gamma(\bar{x}) \subseteq U$, $v' \in V$, and $t \in T^\rho_\gamma(x', v')$ with $|\rho_t(x', v') - \rho(x, v')| \le \gamma$ (by the previous part $t \in T^g_{\gamma K}(x', \rho(x', v')Lv')$),

$$\begin{aligned} g_t(x', \rho(x', v')Lv') &\ge -K(\|\bar{x} - x'\| + \|\rho(x', v')Lv' - \rho(x', v')Lv'\|) \\ &\qquad + g_t(\bar{x}, \rho(x, v')Lv') \\ &\ge -K(\|\bar{x} - x'\| + |\rho(x', v') - \rho(x', v')|\|L\|) \\ &\qquad + g_t(\bar{x}, \rho(\bar{x}, v')Lv') \\ &\ge -K(\|\bar{x} - x'\| + K\|\bar{x} - x'\|\|L\|) \\ &\qquad + g_t(\bar{x}, \rho(\bar{x}, v')Lv') \\ \text{(recall } t \in T^g_{\eta K}(x', \rho(x', v')Lv')) &\ge -K(\gamma + K\|L\|\gamma) - \gamma K = -K(2 + K\|L\|)\gamma. \qquad \square \end{aligned}$$

The above lemma allows us to link the ε-active index of the function ρ to the ε-active index of the nominal data g_t. More precisely we have the following result.

PROPOSITION 4.12. *Consider the mapping $\rho : D \times \mathbb{S}^{m-1} \to \mathbb{R}_+ \cup \{\infty\}$ of Proposition 3.5. Assume that the family is uniformly locally Lipschitzian at $\bar{x}$ and $g(\bar{x}, 0) < 0$. Then*

$$\partial^{\mathtt{M}}_x \rho(\bar{x}, v) \subseteq \bigcap_{\varepsilon > 0} \mathrm{cl}^{\mathrm{w}^*} \left\{ -\frac{\nabla_x g_t(x, \rho_t(x, v)Lv)}{\langle \nabla_z g_t(x, \rho_t(x, v)Lv), Lv\rangle} : \begin{array}{c} x \in \mathbb{B}(\bar{x}, v),\ t \in T^g_\varepsilon(\bar{x}, \rho(\bar{x}, v)Lv) \\ \textit{with } |\rho_t(x, v) - \rho(\bar{x}, v)| \le \varepsilon \end{array} \right\}. \tag{4.13}$$

If in addition $\{\rho_t(\cdot, v)\}_{t \in T}$ is strongly equicontinuously subdifferentiable at $\bar{x}$,

$$\partial^{\mathtt{M}}_x \rho(\bar{x}, v) \subseteq \bigcap_{\varepsilon > 0} \mathrm{cl}^{\mathrm{w}^*} \left\{ -\frac{\nabla_x g_t(\bar{x}, \rho_t(\bar{x}, v)Lv)}{\langle \nabla_z g_t(x, \rho_t(\bar{x}, v)Lv), Lv\rangle} : t \in T^g_\varepsilon(\bar{x}, \rho(\bar{x}, v)Lv) \right\}. \tag{4.14}$$

Proof. By our assumptions we have that the hypotheses of Proposition 3.5 hold, and so $\partial^{\mathtt{M}}_x \rho(\bar{x}, v) \subseteq \mathcal{S}^{\mathtt{M}}(\bar{x}, v)$ and

$$\mathcal{S}^{\mathtt{M}}(\bar{x}, v) \subseteq \bigcap_{\varepsilon > 0} \mathrm{cl}^{\mathrm{w}^*} \left\{ -\frac{\nabla_x g_t(x, \rho_t(x,v)Lv)}{\langle \nabla_z g_t(x, \rho_t(x,v)Lv), Lv\rangle} : \begin{array}{c} x \in \mathbb{B}(\bar{x}, v),\ t \in T^{\rho}_{\varepsilon}(x, \rho(x,v)Lv) \\ \text{with } |\rho_t(x,v) - \rho(\bar{x}, v)| \le \varepsilon \end{array} \right\}.$$

By Proposition 4.6 we have that the right-hand side of the above inclusion is contained in the right-hand side of (4.13).

Finally, under the assumption of strongly equicontinuously subdifferentiable we have that by Lemma 4.11 the right-hand side of (4.5) must be included in the right-hand side of (4.14). ∎

PROPOSITION 4.13. *Under the above assumption one has that $\mathcal{S}$ is locally bounded at every $(x, v) \in U \times \mathbb{S}^{m-1}$ with $\rho(x,v) < +\infty$ and*

$$\partial^{\mathtt{M}}_x e(x, v) = \left\{ -f_{\mathcal{R}}(\rho(x,v)) \frac{\nabla_x g_t(x, \rho(x,v)Lv)}{\langle \nabla_z g_t(x, \rho(x,v)Lv), Lv\rangle} : t \in T^{\rho}(x,v) \right\}. \tag{4.15}$$

Proof. First, under our assumptions the set-valued map $\mathcal{S}$ has the following representation at every $(x,v) \in U \times \mathbb{S}^{m-1}$ with $\rho(x,v) < +\infty$:

$$\mathcal{S}(x,v) = \bigcap_{t \in T^{\rho}(x,v)} \left\{ -\frac{\nabla_x g_t(x, \rho_t(x,v)Lv)}{\langle \nabla_z g_t(x, \rho_t(x,v)Lv), Lv\rangle} \right\}. \tag{4.16}$$

Assume by contradiction that $\mathcal{S}$ is not locally bounded at some fixed $(x,v) \in U \times \mathbb{S}^{m-1}$ with $0 < \rho(x,v) < +\infty$. So let $(x_n, v_n) \to (x, v)$ such that we have

$$\left\| \frac{\nabla_x g_{t_n}(x, \rho_{t_n}(x_n, v_n)Lv_n)}{\langle \nabla_z g_{t_n}(x_n, \rho_{t_n}(x_n, v_n)Lv_n), Lv_n\rangle} \right\| \to +\infty$$

with $t_n \in T^{\rho}(x,v)$. Then by Lemma 4.11 $t_n \in T^{g}(x, \rho(x,v)Lv)$, and up to a subsequence, we have that $t_{n_k} \to t \in T$ by Assumption 4.8(e). Then by continuity of G we have that

$$\frac{\nabla_x g_{t_{n_k}}(x, \rho_{t_{n_k}}(x,v)Lv)}{\langle \nabla_z g_{t_{n_k}}(x, \rho_{t_{n_k}}(x,v)Lv), Lv\rangle} \to \frac{\nabla_x g_t(x, \rho_t(x,v)Lv)}{\langle \nabla_z g_t(x, \rho_t(x,v)Lv), Lv\rangle} = -\nabla_x \rho_t(x,v),$$

where here we have used that

$$\inf_{n \in \mathbb{N}} \langle \nabla_z g_{t_n}(x_n, \rho_{t_n}(x_n, v_n)Lv_n), Lv_n\rangle > \inf_{n \in \mathbb{N}} \frac{-g_{t_n}(x_n, 0)}{\rho_{t_n}(x_n, v_n)} > 0,$$

since $g(x, 0) < 0$ and x_n is sufficiently close to x, and also that $\rho_{t_n}(x_n, v_n) \to \rho_t(x,v) \ge \rho(x,v) > 0$ by Lemma 4.10. This is a contradiction, so $\mathcal{S}$ must be locally bounded at (x,v).

Finally, by Lemma 3.6 and Proposition 3.5 we have that

$$\partial^{\mathtt{M}}_x e(x,v) \subseteq f_{\mathcal{R}}(\rho(x,v)) \limsup_{x_n \to x} \mathcal{S}(x_n, v).$$

Let us pick $x^* \in \partial_x^{\mathtt{M}} e(x,v)$. By the above inclusion there are $t_n \in T^\rho(x_n, v) = T^g(x_n, \rho(x_n,v)Lv)$ such that

$$-\frac{\nabla_x g_{t_n}(x, \rho_{t_n}(x_n, v_n)Lv_n)}{\langle \nabla_z g_{t_n}(x_n, \rho_{t_n}(x_n, v_n)Lv_n), Lv_n\rangle} \to x^*.$$

By passing to a subsequence we can assume that $t_n \to t$ (recall once more Assumption 4.8(e)), and then it is not difficult to see that $t \in T^\rho(x,v)$. Indeed, $g_{t_n}(x_n, \rho(x_n,v)Lv) = g(x_n, \rho(x_n,v)Lv) = 0$ by assumption on t_n. By Proposition 3.2, $\rho(x_n,v) \to \rho(x,v)$ so that by continuity of G we get $g_t(x, \rho(x,v)Lv) = 0$, i.e., $\rho(x,v) = \rho_t(x,v)$. Now recall Lemma 4.10(ii) and the continuity of G to see that necessarily $x^* = -\frac{\nabla_x g_t(x,\rho_t(x,v)Lv)}{\langle \nabla_z g_t(x,\rho_t(x,v)Lv), Lv\rangle}$. This shows (4.15). ∎

Taking into account (2.3) it is convenient to introduce the following growth condition.

Definition 4.14. *Let $\theta_{\mathcal{R}} \colon \mathbb{R}_+ \to \mathbb{R}_+$ be an increasing mapping such*

$$\lim_{r\to+\infty} r f_{\mathcal{R}}(r)\theta_{\mathcal{R}}(r) = 0. \tag{4.17}$$

We say that $\{g_t \colon t \in T\}$ satisfies the $\theta_{\mathcal{R}}$-growth condition uniformly on T at $\bar{x}$ if, for some $l > 0$,

$$\|\nabla_x g_t(x,z)\| \le l\theta_{\mathcal{R}}\left(\frac{\|z\|}{\|L\|}\right) \quad \text{for all } x \in \mathbb{B}_{1/l}(\bar{x}),\ \forall z\colon \|z\| \ge l;\ \forall t \in T. \tag{4.18}$$

Again, when $f_{\mathcal{R}}$ has bounded support, one can set $\theta_{\mathcal{R}} \equiv \infty$ and no particular restriction is implied by (4.18).

We can now provide the following theorem resulting immediately from the sharper estimate (4.15) and the already established Theorem 3.9.

Theorem 4.15. *Let $\xi \in \mathbb{R}^m$ be an elliptical symmetrically distributed random vector with mean 0, correlation matrix $R = LL^\mathsf{T}$, and continuous generator. Consider the probability function $\varphi : X \to [0,1]$, where X is a reflexive Banach space defined as*

$$\varphi(x) = \mathbb{P}[g_t(x,\xi) \le 0 \ \ \forall t \in T], \tag{4.19}$$

where $g_t : X \times \mathbb{R}^m \to \mathbb{R}$ are continuously differentiable maps convex in the second argument and T is a metric space.

Then let $\bar{x} \in X$ be such that

1. *a neighborhood U of $\bar{x}$ can be found such that $g_{|_{U\times\mathbb{R}^m}}$ is finite valued and $\sup_{t\in T} g_t(x',0) < 0$ for all $x' \in U$;*
2. *the function $G\colon T \times U \times \mathbb{R}^m \to X \times X^* \times \mathbb{R}^m$ given by*

$$G(t,x,z) = (g_t(x,v), \nabla_x g_t(x,z), \nabla_z g_t(x,z))$$

is continuous;

3. *the active index set $T^g(x,z)$ is nonempty for every*

$$(x,z) \in \mathfrak{K} = \{(x,z) \in U \times \mathbb{R}^m \colon g(x,z) = 0\};$$

4. *the set* $\bigcup_{(x,z)\in\mathfrak{K}} T^g(x,z)$ *is relatively compact;*
5. *either* $\{g_t : t \in T\}$ *satisfies the* $\theta_{\mathcal{R}}$*-growth condition uniformly on* T *at* $\bar{x}$*, or* $M(\bar{x}) := \{z \in \mathbb{R}^m : g(\bar{x}, z) \leq 0\}$ *is bounded.*

Then φ *is locally Lipschitzian at* $\bar{x}$ *and the following formulae hold true:*

(4.20a)
$$\partial^{\mathtt{M}}\varphi(\bar{x}) \subseteq \mathrm{cl}^{\mathrm{w}^*} \int_{v\in F(\bar{x})} \left\{ -f_{\mathcal{R}}(\rho(x,v)) \frac{\nabla_x g_t(x,\rho(x,v)Lv)}{\langle \nabla_z g_t(x,\rho(x,v)Lv), Lv\rangle} : t \in T^{\rho}(x,v) \right\} d\mu_\zeta(v),$$

(4.20b)
$$\partial^{\mathtt{C}}\varphi(\bar{x}) \subseteq \int_{v\in F(\bar{x})} \mathrm{Co} \left\{ -f_{\mathcal{R}}(\rho(x,v)) \frac{\nabla_x g_t(x,\rho(x,v)Lv)}{\langle \nabla_z g_t(x,\rho(x,v)Lv), Lv\rangle} : t \in T^{\rho}(x,v) \right\} d\mu_\zeta(v),$$

where $\partial^{\mathtt{M}}$ *refers to the limiting (or Kruger–Mordukhovich) subdifferential and* $\partial^{\mathtt{C}}$ *to the Clarke subdifferential.*

Concluding remarks. In this paper we have provided insights on generalized differentiation of probability constraints acting on infinite systems of constraints. Such results will be useful in the analysis of so-called probust constraints arising in applications such as gas network optimization. As future research paths we will investigate dependency of the index set on the decision and/or random vectors.

Appendix A. Some proofs.

Proof of Lemma 2.1. We begin by remarking that if $\partial^{\mathtt{F}} f(\bar{x}) = \emptyset$, the statements of the lemma hold trivially for any $\bar{x}^* \in \partial^{\mathtt{F}} f(\bar{x})$ since there is no such element.

According to the assumption that $\bar{x}^* \in \partial^{\mathtt{F}} f(\bar{x})$ can be found (and hence the latter set is nonempty), we can identify a neighborhood $\mathrm{int}\,\mathbb{B}_r(\bar{x})$ of $\bar{x}$ and a mapping $h \in C^1(U;\mathbb{R})$ such that $\nabla h(\bar{x}) = \bar{x}^*$ and $f - h$ attains its minimum over $\mathbb{B}_r(\bar{x})$ at $\bar{x}$. By continuous differentiability of h we may further shrink $\mathbb{B}_r(\bar{x})$ so that

$$|h(x) - h(y)| + \|\nabla h(x) - \nabla h(y)\| \leq \varepsilon/4$$

for every $x, y \in \mathbb{B}_r(\bar{x})$ also holds true.

Now should some $t \in T$ exist with $f_t(\bar{x}) = f(\bar{x})$, then $\bar{x}$ is also a minimum of $f_t - h + \delta_{\mathbb{B}_r(\bar{x})}$, and hence $x^* \in \partial^{\mathtt{F}} f_t(\bar{x})$. The requested elements of the Lemma can now be immediately identified. Otherwise, consider $0 < \gamma < \min\{\varepsilon^2, r^2, r, \varepsilon\}/8$, then pick $t \in T$ such that $f_t(\bar{x}) \leq f(\bar{x}) + \gamma$. Then it is easy to see that $\bar{x}$ is a γ-minimum of $f_t - h + \delta_{\mathbb{B}_r(\bar{x})}$. Now let $\lambda := \sqrt{\gamma}$, so by the *Borwein–Preiss variational principle* (see, e.g., [6, Theorem 2.5.3]) there exists $x_t \in X$, a sequence $x_i \in X$ with $x_1 = \bar{x}$, and a function $\ell\colon X \to \mathbb{R}$ given by $\ell(x) = \sum_{i=1}^{\infty} \mu_i \|x - x_i\|^2$, where $\mu_i > 0$ for all $i = 1, 2, \ldots$ and $\sum_{i=1}^{\infty} \mu_i = 1$ such that

(BP.1) $\|x_i - x_t\| \leq \lambda(< r)$, $i = 1, 2, \ldots$,

(BP.2) $f_t(x_t) - h(x_t) + \frac{\gamma}{\lambda^2}\ell(x_t) \leq f_t(\bar{x}) - h(\bar{x})$,

(BP.3) $f_t(w) - h(w) + \frac{\gamma}{\lambda^2}\ell(w) > f_t(x_t) - h(x_t) + \frac{\gamma}{\lambda^2}\ell(x_t)$ for all $w \in \mathbb{B}_r(\bar{x}) \setminus \{x_t\}$.

We first note that ℓ is smooth and recall that h is too. Consequently, so is $h - \frac{\gamma}{\lambda^2}\ell$. Moreover, by (BP.3), it is clear that $f_t - (h - \frac{\gamma}{\lambda^2}\ell)$ attains a local minimum at x_t. Therefore $x_t^* := \nabla h(x_t) - \frac{\gamma}{\lambda^2}\nabla\ell(x_t)$ belongs to the viscosity Fréchet subdifferential of f_t and by [6, Proposition 3.1.3] also to the Fréchet subdifferential of f_t. It is now our claim that x_t and x_t^* satisfy the requested requirements.

Indeed, by picking $i = 1$ in (BP.1), it follows that $\|\bar{x} - x_t\| \leq \lambda < \min\{\varepsilon, r\}$ holds true. Moreover, by definition of t, $|f_t(\bar{x}) - f(\bar{x})| \leq \gamma \leq \varepsilon$.

Next, we recall that $\bar{x}$ is the γ-minimum of $f_t - h + \delta_{\mathbb{B}_r(\bar{x})}$. Hence, $f_t(x_t) - h(x_t) - f_t(\bar{x}) + h(\bar{x}) \geq -\gamma$. By reorganizing terms this gives

$$f_t(\bar{x}) - f_t(x_t) \leq h(\bar{x}) - h(x_t) + \gamma. \tag{A.1}$$

We also recall that $\ell(x) \geq 0$ and that by (BP.1), $\ell(x_t) \leq \lambda$. Consequently (BP.2) gives

$$f_t(x_t) - h(x_t) \leq f_t(x_t) - h(x_t) + \ell(x_t) \leq f_t(\bar{x}) - h(\bar{x}),$$

which yields

$$f_t(x_t) - f_t(\bar{x}) \leq h(x_t) - h(\bar{x}) \leq h(x_t) - h(\bar{x}) + \gamma \tag{A.2}$$

since $\gamma > 0$. By combining (A.1) and (A.2) we can take the absolute value on both the left- and right-hand sides to obtain

$$|f_t(x_t) - f_t(\bar{x})| \leq |h(x_t) - h(\bar{x}) + \gamma| \leq |h(x_t) - h(\bar{x})| + \gamma \leq \gamma + \frac{\varepsilon}{4}.$$

Now by the triangle inequality and the specific choice of t, we get $|f(\bar{x}) - f_t(x_t)| \leq |f(\bar{x}) - f_t(\bar{x})| + |f_t(\bar{x}) - f_t(x_t)| \leq \gamma + \epsilon/4 + \gamma \leq \varepsilon/2 < \varepsilon$.

It remains to show that $\|x_t^* - \bar{x}^*\| < \varepsilon$. To this end, we observe that the map $x \mapsto \|x\|^2$ is convex and its subdifferential at $x \in X$ is the set

$$\{2x^* \colon x^* \in X^* \text{ s.t.: } \langle x^*, x\rangle = \|x\| \, \|x^*\| \, , \ \|x\| = \|x^*\|\}$$

(see [55, Example 2.26]). By construction the map ℓ is convex, and as a consequence of (BP.1), $\ell(x_t) \leq \lambda < \infty$. Hence any $s \in \partial\ell(x_t)$, must satisfy by the above discussion $\|s\| \leq 2\sum_{i=1}^{\infty} \mu_i \lambda$, by employing (BP.1) once more. We recall here that we may interchange subdifferentiation and infinite summation as a classical result (see, e.g., [31, Theorem 1] and also [13, 14, 15] or [27, Proposition 3]). However, as we noted above, ℓ is actually a smooth map. Consequently,

$$\|x_t^* - \bar{x}^*\| = \|\nabla\ell(x_t)\| + \|\nabla h(x_t) - \nabla h(\bar{x})\| \leq 2\lambda + \varepsilon/4 \leq \varepsilon,$$

which was the last property to establish, and hence we conclude the proof. ∎

Acknowledgments. The authors are particularly grateful to two anonymous referees that have greatly helped to improve this manuscript.

REFERENCES

[1] D. Adelhütte, D. Assmann, T. Gonzàlez-Gradòn, M. Gugat, H. Heitsch, R. Henrion, F. Liers, S. Nitsche, R. Schultz, M. Stingl, and D. Wintergerst, *Joint Model of Probabilistic/Robust (Probust) Constraints with Application to Gas Network Optimization*, preprint, 2017; available at https://opus4.kobv.de/opus4-trr154/frontdoor/index/index/docId/215.

[2] J. Barrera, T. Homem de Mello, E. Moreno, and B. K. Pagnoncelli, *Chance-constrained problems and rare events: An importance sampling approach*, Math. Program., 157 (2016), pp. 153–189.

[3] A. Ben-Tal, L. E. Ghaoui, and A. Nemirovski, *Robust Optimization*, Princeton University Press, Princeton, NJ, 2009.

[4] A. Ben-Tal and A. Nemirovski, *Robust solutions of linear programming problems contaminated with uncertain data*, Math. Program., 88 (2000), pp. 411–424.

[5] J. M. Borwein and D. Preiss, *A smooth variational principle with applications to subdifferentiability and differentiability of convex functions*, Trans. Amer. Math. Soc., 303 (1987), pp. 517–527.

[6] J. M. Borwein and Q. J. Zhu, *Techniques of Variational Analysis*, CMS Books Math./Ouvrages Math. SMC 20, Springer, New York, 2005.

[7] J. S. Brauchart, E. B. Saff, I. H. Sloan, and R. S. Womersley, *QMC designs: Optimal order Quasi Monte Carlo integration schemes on the sphere*, Math. Comp., 83 (2014), pp. 2821–2851, https://doi.org/10.1090/S0025-5718-2014-02839-1.

[8] I. Bremer, R. Henrion, and A. Möller, *Probabilistic constraints via SQP solver: Application to a renewable energy management problem*, Comput. Manag. Sci., 12 (2015), pp. 435–459.

[9] G. C. Calafiore and M. C. Campi, *The scenario approach to robust control design*, IEEE Trans. Automat. Control, 51 (2006), pp. 742–753.

[10] M. C. Campi and S. Garatti, *A sampling-and-discarding approach to chance-constrained optimization: Feasibility and optimality*, J. Optim. Theory Appl., 148 (2011), pp. 257–280.

[11] H. Cartan, *Differential Calculus*, Hermann, Paris, 1971.

[12] F. H. Clarke, Y. S. Ledyaev, R. J. Stern, and P. R. Wolenski, *Nonsmooth Analysis and Control Theory*, Grad. Texts in Math. 178, Springer, New York, 1998.

[13] R. Correa, A. Hantoute, and P. Pérez-Aros, *Qualification conditions-free characterizations of the ε-subdifferential of convex integral functions*, Appl. Math. Optim., to appear.

[14] R. Correa, A. Hantoute, and P. Pérez-Aros, *Characterizations of the subdifferential of convex integral functions under qualification conditions*, J. Funct. Anal., 277 (2019), pp. 227–254, https://doi.org/10.1016/j.jfa.2019.02.007.

[15] R. Correa, A. Hantoute, and P. Pérez-Aros, *Sequential and Exact Formulae for the Subdifferential of Nonconvex Integral Functionals*, preprint, https://arxiv.org/abs/1803.05521, 2018.

[16] D. Dentcheva, *Optimisation Models with Probabilistic Constraints*, in Lectures on Stochastic Programming: Modeling and Theory, A. Shapiro, D. Dentcheva, and A. Ruszczyński, eds., MPS-SIAM Ser. Optim. 9, SIAM, MPS, Philadelphia, PA, 2009, pp. 87–154.

[17] D. Dentcheva, B. Lai, and A. Ruszczyński, *Dual methods for probabilistic optimization problems*, Math. Meth. Oper. Res., 60 (2004), pp. 331–346.

[18] D. Dentcheva and G. Martinez, *Augmented Lagrangian method for probabilistic optimization*, Ann. Oper. Res., 200 (2012), pp. 109–130.

[19] D. Dentcheva and G. Martinez, *Regularization methods for optimization problems with probabilistic constraints*, Math. Program., 138 (2013), pp. 223–251.

[20] D. Dentcheva, A. Prékopa, and A. Ruszczyński, *Concavity and efficient points for discrete distributions in stochastic programming*, Math. Program., 89 (2000), pp. 55–77.

[21] Y. Ermoliev, T. Ermolieva, G. Macdonald, and V. Norkin, *Stochastic optimization of insurance portfolios for managing exposure to catastrophic risk*, Ann. Oper. Res., 99 (2000), pp. 207–225.

[22] K. Fang, S. Kotz, and K. W. Ng, *Symmetric Multivariate and Related Distributions*, 1st ed., Monogr. Statist. Appl. Probab. 36 , Springer, Cham, 1990.

[23] M. H. Farshbaf-Shaker, R. Henrion, and D. Hömberg, *Properties of chance constraints in infinite dimensions with an application to PDE constrained optimization*, Set-Valued Var. Anal., 26 (2018), pp. 821–841, https://doi.org/10.1007/s11228-017-0452-5.

[24] J. Garnier, A. Omrane, and Y. Rouchdy, *Asymptotic formulas for the derivatives of probability functions and their Monte Carlo estimations*, Eur. J. Oper. Res., 198 (2009), pp. 848–858, https://doi.org/10.1016/j.ejor.2008.09.026.

[25] A. Geletu, A. Hoffmann, M. Klöppel, and P. Li, *A tractable approximation of non-convex chance constrained optimization with non-Gaussian uncertainties*, Eng. Optim., 47 (2015), pp. 495–520.

[26] T. González Gradón, H. Heitsch, and R. Henrion, *A joint model of probabilistic/robust constraints for gas transport management in stationary networks*, Comput. Manag. Sci., 14 (2017), pp. 443–460.

[27] A. Hantoute, R. Henrion, and P. Pérez-Aros, *Subdifferential characterization of continuous probability functions under Gaussian distribution*, Math. Program., 174 (2019), pp. 167–194, https://doi.org/10.1007/s10107-018-1237-9.

[28] R. Henrion, *Introduction to Chance-Constrained Programming*, Tutorial paper for the Stochastic Programming Community Home Page; available at http://www.wias-berlin.de/people/henrion/publikat.html, 2004.

[29] R. Henrion and A. Möller, *A gradient formula for linear chance constraints under Gaussian distribution*, Math. Oper. Res., 37 (2012), pp. 475–488, https://doi.org/10.1287/moor.1120.0544.

[30] L. Hong, Y. Yang, and L. Zhang, *Sequential convex approximations to joint chance constrained programed: A Monte Carlo approach*, Oper. Res., 3 (2011), pp. 617–630.
[31] A. D. Ioffe and V. L. Levin, *Subdifferentials of convex functions*, Tr. Mosk. Mat. Obščestva, 26 (1972), pp. 3–73.
[32] A. Kibzun and S. Uryas'ev, *Differentiability of probability function*, Stoch. Anal. Appl., 16 (1998), pp. 1101–1128, https://doi.org/10.1080/07362999808809581.
[33] A. Y. Kruger, *Strict (ϵ, δ)-subdifferentials and extremality conditions*, Optimization, 51 (2002), pp. 539–554.
[34] A. Y. Kruger, *On Fréchet subdifferentials*, J. Math Sci., 116 (2003), pp. 3325–3358.
[35] S. Küçükyavuz, *On mixing sets arising in chance-constrained programming*, Math. Program., 132 (2012), pp. 31–56.
[36] Z. M. Landsman and E. A. Valdez, *Tail conditional expectations for elliptical distributions*, N. Am. Actuar. J., 7 (2013), pp. 55–71.
[37] Y. S. Ledyaev and J. S. Treiman, *Sub- and supergradients of envelopes, semicontinuous closures, and limits of sequences of functions*, Russian Math. Surveys, 67 (2012), pp. 345–373.
[38] M. Lejeune and F. Margot, *Solving chance-constrained optimization problems with stochastic quadratic inequalities*, Oper. Res., 64 (2016), pp. 939–957.
[39] M. A. Lejeune, *Pattern-based modeling and solution of probabilistically constrained optimization problems*, Oper. Res., 60 (2012), pp. 1356–1372.
[40] M. A. Lejeune, *Pattern definition of the p-efficiency concept*, Ann. Oper. Res., 200 (2012), pp. 23–36.
[41] M. A. Lejeune and N. Noyan, *Mathematical programming approaches for generating p-efficient points*, Eur. J. Oper. Res., 207 (2010), pp. 590–600.
[42] X. Liu, S. Küçükyavuz, and J. Luedtke, *Decomposition algorithm for two-stage chance constrained programs*, Math. Program., 157 (2016), pp. 219–243, https://doi.org/10.1007/s10107-014-0832-7.
[43] J. Luedtke, *An Integer Programming and Decomposition Approach to General Chance-Constrained Mathematical Programs*, in Integer Programming and Combinatorial Optimization, F. Eisenbrand and F. B. Shepherd, eds., Lecture Notes in Comput. Sci. 6080, Springer, Cham, 2010, pp. 271–284.
[44] J. Luedtke, *A branch-and-cut decomposition algorithm for solving chance-constrained mathematical programs with finite support*, Math. Program., 146 (2014), pp. 219–244.
[45] J. Luedtke and S. Ahmed, *A sample approximation approach for optimization with probabilistic constraints*, SIAM J. Optim., 19 (2008), pp. 674–699.
[46] J. Luedtke, S. Ahmed, and G. Nemhauser, *An integer programming approach for linear programs with probabilistic constraints*, Math. Program., 122 (2010), pp. 247–272.
[47] K. Marti, *Differentiation of probability functions: The transformation method*, Comput. Math. Appl., 30 (1995), pp. 361–382, https://doi.org/10.1016/0898-1221(95)00113-1.
[48] B. S. Mordukhovich, *Variational Analysis and Generalized Differentiation* I. *Basic Theory*, Grundlehren Math. Wiss. 330, Springer, Heidelberg, 2006.
[49] B. S. Mordukhovich and T. T. A. Nghia, *Constraint qualifications and optimality conditions for nonconvex semi-infinite and infinite programs*, Math. Program., 139 (2013), pp. 271–300.
[50] B. S. Mordukhovich and T. T. A. Nghia, *Subdifferentials of nonconvex supremum functions and their applications to semi-infinite and infinite programs with Lipschitzian data*, SIAM J. Optim., 23 (2013), pp. 406–431.
[51] B. S. Mordukhovich and B. Wang, *Necessary suboptimality and optimality conditions via variational principles*, SIAM J. Control Optim., 41 (2002), pp. 623–640.
[52] A. Nemirovski and A. Shapiro, *Convex approximations of chance constrained programs*, SIAM J. Optim., 17 (2006), pp. 969–996.
[53] B. Pagnoncelli, S. Ahmed, and A. Shapiro, *Sample average approximation method for chance constrained programming: Theory and applications*, J. Optim. Theory Appl, 142 (2009), pp. 399–416.
[54] G. Pflug and H. Weisshaupt, *Probability gradient estimation by set-valued calculus and applications in network design*, SIAM J. Optim., 15 (2005), pp. 898–914.
[55] R. R. Phelps, *Convex Functions, Monotone Operators and Differentiability*, 2nd ed., Lecture Notes in Math. 1364, Springer, Berlin, 1993.
[56] A. Prékopa, *Stochastic Programming*, Kluwer, Dordrecht, 1995.
[57] A. Prékopa, *Probabilistic Programming*, in Stochastic Programming, A. Ruszczyński and A. Shapiro, eds., Handb. Oper. Res. Manag. Sci. 10, Elsevier, Amsterdam, 2003, pp. 267–351.

[58] E. Raik, *The differentiability in the parameter of the probability function and optimization of the probability function via the stochastic pseudogradient method (Russian)*, Izvestiya Akad. Nayk Est. SSR, Phis. Math., 24 (1975), pp. 3–6.

[59] J. Royset and E. Polak, *Implementable algorithm for stochastic optimization using sample average approximations*, J. Optim. Theory Appl., 122 (2004), pp. 157–184.

[60] J. Royset and E. Polak, *Extensions of stochastic optimization results to problems with system failure probability functions*, J. Optim. Theory Appl., 133 (2007), pp. 1–18.

[61] S. Uryas'ev, *Derivatives of probability functions and integrals over sets given by inequalities*, J. Comput. Appl. Math., 56 (1994), pp. 197–223, https://doi.org/10.1016/0377-0427(94)90388-3.

[62] S. Uryas'ev, *Derivatives of probability functions and some applications*, Ann. Oper. Res., 56 (1995), pp. 287–311.

[63] W. van Ackooij, I. Aleksovska, and M. M. Zuniga, *(Sub-)differentiability of probability functions with elliptical distributions*, Set-Valued Var. Anal., 26 (2018), pp. 887–910, https://doi.org/10.1007/s11228-017-0454-3.

[64] W. van Ackooij, V. Berge, W. de Oliveira, and C. Sagastizábal, *Probabilistic optimization via approximate p-efficient points and bundle methods*, Comput. Oper. Res., 77 (2017), pp. 177–193, https://doi.org/10.1016/j.cor.2016.08.002.

[65] W. van Ackooij and W. de Oliveira, *Level bundle methods for constrained convex optimization with various oracles*, Comput. Optim. Appl., 57 (2014), pp. 555–597.

[66] W. van Ackooij and W. de Oliveira, *Convexity and optimization with copulae structured probabilistic constraints*, Optimization, 65 (2016), pp. 1349–1376, https://doi.org/10.1080/02331934.2016.1179302.

[67] W. van Ackooij, A. Frangioni, and W. de Oliveira, *Inexact stabilized Benders' decomposition approaches: With application to chance-constrained problems with finite support*, Comput. Optim. Appl., 65 (2016), pp. 637–669, https://doi.org/10.1007/s10589-016-9851-z.

[68] W. van Ackooij and R. Henrion, *Gradient formulae for nonlinear probabilistic constraints with Gaussian and Gaussian-like distributions*, SIAM J. Optim., 24 (2014), pp. 1864–1889.

[69] W. van Ackooij and R. Henrion, *(Sub-)gradient formulae for probability functions of random inequality systems under Gaussian distribution*, SIAM/ASA J. Uncertainty Quantification, 5 (2017), pp. 63–87.

[70] W. van Ackooij, R. Henrion, A. Möller, and R. Zorgati, *Chance constrained programming and its applications to energy management*, in Stochastic Optimization – Seeing the Optimal for the Uncertain, I. Dritsas, ed., IntechOpen, London, 2011, pp. 291–320.

[71] W. van Ackooij, R. Henrion, A. Möller, and R. Zorgati, *On joint probabilistic constraints with Gaussian coefficient matrix*, Oper. Res. Lett., 39 (2011), pp. 99–102, https://doi.org/10.1016/j.orl.2011.01.005.

[72] W. van Ackooij and J. Malick, *Second-order differentiability of probability functions*, Optim. Lett., 11 (2017), pp. 179–194, https://doi.org/10.1007/s11590-016-1015-7.

[73] W. van Ackooij and J. Malick, *Eventual convexity of probability constraints with elliptical distributions*, Math. Program., 175 (2019), pp. 1–27, https://doi.org/10.1007/s10107-018-1230-3.

[74] W. van Ackooij and C. Sagastizábal, *Constrained bundle methods for upper inexact oracles with application to joint chance constrained energy problems*, SIAM J. Optim., 24 (2014), pp. 733–765.