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

Subdifferential characterization of probability functions under Gaussian distribution

Abderrahim Hantoute · René Henrion · Pedro Pérez-Aros

July 8, 2021

Abstract Probability functions figure prominently in optimization problems of engineering. They may be nonsmooth even if all input data are smooth. This fact motivates the consideration of subdifferentials for such typically just continuous functions. The aim of this paper is to provide subdifferential formulae of such functions in the case of Gaussian distributions for possibly infinite-dimensional decision variables and nonsmooth (locally Lipschitzian) input data. These formulae are based on the spheric-radial decomposition of Gaussian random vectors on the one hand and on a cone of directions of moderate growth on the other. By successively adding additional hypotheses, conditions are satisfied under which the probability function is locally Lipschitzian or even differentiable.

Keywords probability functions, probabilistic constraint, stochastic optimization, multivariate Gaussian distribution, spheric-radial decomposition, Clarke subdifferential, Mordukhovich subdifferential

Mathematics Subject Classification (2010) 90C15 · 90C30 · 49J52 · 49J53

This work is partially supported by CONICYT grants: Fondecyt 1151003, Fondecyt 1150909, Basal PFB-03 and Basal FB0003, CONICYT-PCHA/doctorado Nacional/2014-21140621. The second author thanks the Deutsche Forschungsgemeinschaft for their support within project B04 in the Sonderforschungsbereich / Transregio 154 Mathematical Modelling, Simulation and Optimization using the Example of Gas Networks.

A. Hantoute
Center for Mathematical Modeling, Universidad de Chile, Santiago, Chile
E-mail: ahantoute@dim.uchile.cl

R. Henrion
Weierstrass Institute for Applied Analysis and Stochastics, Berlin, Germany
E-mail: henrion@wias-berlin.de

P. Pérez-Aros
Department of Mathematical Engineering, Universidad de Chile, Santiago, Chile
E-mail: pperez@dim.uchile.cl

1 Introduction

The aim of this paper is to investigate subdifferential properties of Gaussian probability functions induced by nonnecessarily smooth initial data. This topic combines aspects of stochastic programming with arguments from variational analysis, two areas which have been crucially influenced by the fundamental work of Prof. Roger J-B Wets (see, e.g., [17], [21] and many other references). The motivation to study analytical properties of probability functions comes from their importance in the context of engineering problems affected by random parameters. They are at the core of probabilistic programming (i.e., optimization problems subject to probabilistic constraints) (e.g., [15], [18]) or of reliability maximization (e.g., [7]).

A probability function assigns to a control or decision variable the probability that a certain random inequality system induced by this decision variable be satisfied (see (1) below). Since such functions are typical constituents of optimization problems under uncertainty, it is natural to ask for their analytical properties, first of all differentiability. Roughly speaking, this can be guaranteed under three assumptions: the differentiability of the input data, an appropriate constraint qualification for the given random inequality system and the compactness of the set of realizations of the random vector for the fixed decision vector (e.g., [11], [14], [19]). While the first two assumptions are quite natural, the last one appears to be restrictive in problems involving random vectors with unbounded support. Failure of the compactness condition, however, may result in general in nonsmoothness of the probability function despite the fact that all input data are smooth and a standard constraint qualification is satisfied (see [1, Prop. 2.2]). In order to keep the differentiability while doing without the compactness assumption, one may restrict to special distributions such as Gaussian or Gaussian-like as in [1], [2]. The working horse for deriving differentiability and gradient formulae in these cases is the so-called *spheric-radial decomposition* of Gaussian random vectors [8, p. 29]. The resulting formulae for the gradient of the probability function are represented - similar to the formulae for the probability values themselves - as integrals over the unit sphere with respect to the uniform measure. The latter can be efficiently approximated by QMC methods tailored to this specific measure (e.g., [3]). Such approach, by exploiting special properties of the distribution, promises more efficiency in the solution of probabilistic programs than general gradient formulae in terms of possibly complicated surface or volume integrals. Successful applications of this methodology in the context of probabilistic programming in gas network optimization is demonstrated in [9], [10].

The aim of this paper is to substantially extend the earlier results in [1], [2] in two directions: first, decisions will be allowed to be infinite-dimensional and second, the random inequality may be just locally Lipschitzian rather than smooth. As the resulting probability function can be expected to be continuous only (rather than locally Lipschitzian or even smooth), appropriate tools (subdifferentials) from variational analysis will be employed for an analytic characterization.

We consider a probability function $\varphi : X \to \mathbb{R}$ defined by

$$\varphi(x) := \mathbb{P}\left(g\left(x,\xi\right) \leq 0\right), \tag{1}$$

where X is a Banach space, $g : X \times \mathbb{R}^m \to \mathbb{R}$ is a function depending on the realizations of an m-dimensional random vector ξ. Such probability functions are important in many optimization problems dealing with reliability maximization or probabilistic constraints. The latter one refers to an inequality $\varphi(x) \geq p$ constraining the set of feasible decisions in an optimization problem, in order to guarantee that the underlying random inequality $g\left(x,\xi\right) \leq 0$ is satisfied under decision x with probability at least $p \in (0,1]$, referred to as a a probability level (or safety level). Since we allow in our paper the function g to be locally Lipschitzian, there is no loss of generality in considering a single random inequality only because in a finite system of such inequalities one could pass to the maximum of components.

Throughout the paper, we shall make the following basic assumptions on the data of (1):

(H)

1. X is a reflexive and separable Banach space.
2. Function g is locally Lipschitzian as a function of both arguments simultaneously, and convex as a function of the second argument.
3. The random vector ξ is Gaussian of type $\xi \sim \mathcal{N}\left(0,R\right),$ where R is a correlation matrix.

A brief discussion of these assumptions is in order here: reflexivity of X is imposed in order to work with the limiting (Mordukhovich) subdifferential as introduced in Definition 2 below (actually, one could consider the more general case of Asplund spaces). The separability of X is needed in order to make use of an interchange formula for the limiting subdifferential and integration sign (see Proposition 3 below). For the same reason, g is required to be locally Lipschitzian. As already mentioned above, considering just one inequality rather than a system is no more restriction then. In particular, the single inequality $g\left(x,z\right) \leq 0$ could represent a finite or (compactly indexed) infinite system of smooth inequalities. Considering a Gaussian random vector ξ allows one to pass to a whole class of Gaussian-like multivariate distributions (e.g., Student, Log-normal, truncated Gaussian, χ^2 etc.) upon shifting their nonlinear transformations to a Gaussian one into a modified function $\tilde{g}$ satisfying the same assumptions as required for g here (e.g. [1, Section 4.3]). Moreover, assuming a centered Gaussian distribution with unit variances isn't a restriction either, because in the general case $\xi \sim \mathcal{N}\left(\mu,\Sigma\right)$, we may pass to the standardized vector $\tilde{\xi} := D(\xi - \mu)$, where D is the diagonal matrix with elements $D_{ii} := 1/\sqrt{\Sigma_{ii}}$. Then, as required above, $\tilde{\xi} \sim \mathcal{N}\left(0,R\right)$, with R being the correlation matrix associated with Σ and so

$$\varphi(x) = \mathbb{P}\left(g\left(x,\xi\right) \leq 0\right) = \mathbb{P}\left(\tilde{g}\left(x,\tilde{\xi}\right) \leq 0\right); \quad \tilde{g}\left(x,z\right) := g\left(x,D^{-1}z+\mu\right).$$

Clearly, $\tilde{g}$ is locally Lipschitzian and is convex in the second argument if g is so. Hence, there is no loss of generality in assuming that $\xi \sim \mathcal{N}(0, R)$ from the very beginning.

Our first observation is that our basic assumptions above do not guarantee the continuity of φ even if g is continuously differentiable. A simple two-dimensional example is given by $g(r,s) := r \cdot s$ (which is convex in the second argument) and $\xi \sim \mathcal{N}(0,1)$. Then, $\varphi(r) = 0.5$ for $r \neq 0$ and $\varphi(0) = 1$. Since we want to have the continuity as a minimum initial property of φ in our analysis, we will add the additional assumption that $g(\bar{x}, 0) < 0$ holds true at a point of interest $\bar{x}$ (at which a subdifferential of φ is computed). In other words, given the convexity of g in the second argument, zero is a Slater point for the inequality $g(x,z) \leq 0$, $z \in \mathbb{R}^m$. As shown in [1, Proposition 3.11], the opposite case would entail that $\varphi(\bar{x}) \leq 0.5$. Since one deals in typical applications like probabilistic programming or reliability maximization with probabilities close to one, it follows that the assumption $g(\bar{x}, 0) < 0$ can be made without any practical loss of generality.

The paper is organized as follows: In Section 3 and 4, we provide all the auxiliary results (continuity and partial subdifferential of the radial probability function) which are needed for the derivation of the main subdifferential formula presented in Section 5. This main result which is valid for general continuous probability functions will be specified then by adding additional hypotheses to the locally Lipschitzian and differentiable case. An application to probability functions induced by a finite system of smooth inequalities is given in Subsection 5.4.

2 Preliminaries

2.1 Spheric-radial decomposition of Gaussian random vectors

We recall the fact that any Gaussian random vector $\xi \sim \mathcal{N}(0, R)$ has a so-called spheric-radial decomposition, which means that the probability of ξ taking values in an arbitrary Borel subset M of $\mathbb{R}^m$ can be represented as (e.g., [6, p. 105])

$$\mathbb{P}(\xi \in M) = \int\limits_{v \in \mathbb{S}^{m-1}} \mu_\eta \left(\{ r \geq 0 \mid rLv \in M \} \right) d\mu_\zeta(v),$$

where $\mathbb{S}^{m-1} := \left\{ v \in \mathbb{R}^m \mid \|v\|^2 = 1 \right\}$ denotes the unit sphere in $\mathbb{R}^m$, μ_η is the one-dimensional Chi-distribution with m degrees of freedom, and μ_ζ refers to the uniform distribution on $\mathbb{S}^{m-1}$. Moreover, the (non-singular) matrix L is supposed to be a factor in a decomposition $R = LL^T$ of the positive definite correlation matrix R (e.g. Cholesky decomposition).

The spheric-radial decomposition allows us to rewrite the probability function (1) in the form

$$\varphi(x) = \int_{\mathbb{S}^{m-1}} e(x,v)d\mu_\zeta(v) \quad \forall x \in X, \tag{2}$$

where $e : X \times \mathbb{S}^{m-1} \to \mathbb{R}$ refers to the *radial probability function* defined by

$$e(x,v) := \mu_\eta\left(\{r \geq 0 \mid g(x,rLv) \leq 0\}\right). \tag{3}$$

With any $x \in X$ satisfying $g(x,0) < 0$, we will associate the finite and infinite directions defined respectively as

$$\begin{aligned} F(x) &:= \{v \in \mathbb{S}^{m-1} \mid \exists r \geq 0 : g(x,rLv) = 0\}, \\ I(x) &:= \{v \in \mathbb{S}^{m-1} \mid \forall r \geq 0 : g(x,rLv) < 0\}. \end{aligned}$$

It is easily observed that $F(x) \cap I(x) = \emptyset$ and that $F(x) \cup I(x) = \mathbb{S}^{m-1}$ by continuity of g. Moreover, the number $r \geq 0$ satisfying $g(x,rLv) = 0$ in the case of $v \in F(x)$ is uniquely defined, due to the convexity of g in the second argument. This leads us to define the following radius function for any x with $g(x,0) < 0$ and any $v \in \mathbb{S}^{m-1}$:

$$\rho\left(x,v\right) := \begin{cases} r \text{ such that } g(x,rLv) = 0 & \text{if } v \in F(x) \\ +\infty & \text{if } v \in I(x). \end{cases} \tag{4}$$

This definition allows us to rewrite the radial probability function e from (3) in the form

$$e(x,v) = \mu_\eta\left([0,\rho\left(x,v\right)]\right) = F_\eta\left(\rho\left(x,v\right)\right) \tag{5}$$

whenever $g(x,0) < 0$. Here, F_η refers to the distribution function of the Chi-distribution with m degrees of freedom, so that $F'_\eta(t) = \chi(t)$, where χ is the corresponding density:

$$\chi\left(t\right) := Kt^{m-1}e^{-t^2/2} \quad \forall t \geq 0. \tag{6}$$

The second equation in (5) follows from $F_\eta(0) = 0$. We formally put $F_\eta(\infty) := 1$ which translates the limiting property $F_\eta(t) \to_{t\to+\infty} 1$ of cumulative distribution functions.

2.2 Notation and tools from variational analysis

Our notation will be standard. By X and X^* we will denote a real reflexive and separable Banach space and its dual, with corresponding norms $\|\ \|$ and $\|\ \|_*$, and with corresponding balls $\mathbb{B}_r\left(x\right)$, $\mathbb{B}^*_r\left(x^*\right)$ of radius r around $x \in X$ and $x^* \in X^*$. We denote by $\langle x, x^*\rangle$, $x \in X, x^* \in X^*$ the corresponding duality product, and by $\rightharpoonup$ the weak convergence in both X and X^*. The negative polar of some closed cone $C \subseteq X$ is the closed convex cone

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

The notations $\operatorname{cl} C$, $\operatorname{cl}^* C$, $\operatorname{co} C$, and $\overline{\operatorname{co}}\, C$ will refer to the (strong or norm) closure, the weak* closure, the convex hull, and the closed convex hull of $C \subseteq X$ (or $C \subseteq X^*$), respectively.

The indicator and the support functions of a set $C \subseteq X$ (or $C \subseteq X^*$) are respectively defined as

$$i_C(x) := 0 \text{ if } x \in C \text{ and } +\infty \text{ otherwise},$$

$$\sigma_C(x^*) := \sup_{x \in C} \langle x, x^* \rangle.$$

Definition 1 Let $C \subseteq X$ be a closed subset. Then the Fréchet, the Mordukhovich, and the Clarke normal cones to C at $\bar{x} \in C$ are respectively defined as

$$N^F(\bar{x}; C) := \left\{ x^* \in X^* \mid \limsup_{x \to \bar{x}, x \in C} \frac{\langle x^*, x - \bar{x} \rangle}{\|x - \bar{x}\|} \leq 0 \right\},$$

$$N^M(\bar{x}; C) := \left\{ x^* \in X^* \mid \exists x_n \to \bar{x},\, x_n \in C,\, \exists x_n^* \rightharpoonup x^* : x_n^* \in N^F(C; x_n) \right\},$$

$$N^C(\bar{x}; C) := \overline{\operatorname{co}}\, N^M(\bar{x}; C).$$

We note that the definition of N^C is not the original but a derived one. The normal cones induce subdifferentials of functions $f : X \to \mathbb{R}$ via their epigraphs

$$\operatorname{epi} f := \{(x, t) \in X \times \mathbb{R} \mid f(x) \leq t\},$$

which are closed whenever f is lower semicontinuous (lsc, for short).

Definition 2 Let $f : X \to \mathbb{R}$ be a lsc function. Then the Fréchet, the Mordukhovich (limiting), and the Clarke subdifferentials of f at $\bar{x} \in X$, are respectively defined as

$$\partial^{F/M/C} f(\bar{x}) := \left\{ x^* \in X^* \mid (x^*, -1) \in N^{F/M/C}((\bar{x}, f(\bar{x})); \operatorname{epi} f) \right\}.$$

The singular subdifferential of f at $\bar{x}$ is defined as

$$\partial^\infty f(\bar{x}) = \left\{ x^* \in X^* \mid (x^*, 0) \in N^M((\bar{x}, f(\bar{x})); \operatorname{epi} f) \right\}.$$

We recall that the Fréchet subdifferential has the explicit representation

$$\partial^F f(\bar{x}) = \left\{ x^* \in X^* \mid \liminf_{x \to \bar{x}} \frac{f(x) - f(\bar{x}) - \langle x^*, x - \bar{x} \rangle}{\|x - \bar{x}\|} \geq 0 \right\}. \tag{7}$$

In the current setting of reflexive Banach spaces, the following representation holds true for Clarke's subdifferential [13, Theorem 3.57]:

$$\partial^C f(\bar{x}) = \overline{\operatorname{co}} \left\{ \partial^M f(\bar{x}) + \partial^\infty f(\bar{x}) \right\}. \tag{8}$$

For locally Lipschitzian functions, the following classical definition of Clarke's subdifferential applies:

$$\partial^C f(\bar{x}) = \{x^* \in X^* \mid \langle x^*, h \rangle \leq f^\circ(\bar{x}; h),\, \forall h \in X\}, \tag{9}$$

where

$$f^{\circ}(\bar{x};h) := \limsup_{x\to\bar{x},t\downarrow 0}\frac{f(x+th)-f(x)}{t}$$

denotes Clarke's directional derivative of f at $\bar{x}$ in the direction h.

In case that f happens to be convex, all the subdifferentials above coincide with the ordinary subdifferential in the sense of convex analysis:

$$\partial f(\bar{x}) := \{x^* \in X^* \mid f(x) \geq f(\bar{x}) + \langle x^*, x-\bar{x}\rangle, \forall x \in X\}.$$

For a function $f(x,y)$ of two variables, we will refer to its partial subdifferentials at a point $(\bar{x},\bar{y})$ as the corresponding subdifferentials of the partial functions:

$$\partial_x^{F/M/C} f(\bar{x},\bar{y}) := \partial^{F/M/C} f(\cdot,\bar{y})(\bar{x}); \quad \partial_y^{F/M/C} f(\bar{x},\bar{y}) := \partial^{F/M/C} f(\bar{x},\cdot)(\bar{y}).$$

3 Continuity properties

In this section, we investigate continuous properties of the radial probability and the radius functions, defined respectively in (3) and (4), which are the basis for deriving in Section 5 subdifferential formulae for probability function (1).

For all the following results, the basic assumption (H) formulated in the Introduction is tacitly required to hold; namely, function g is locally Lipschitzian as a function of both arguments simultaneously, and convex as a function of the second argument.

Lemma 1 *Define* $U := \{x \in X \mid g(x,0) < 0\}$.

1. *The radius function ρ is continuous at (x,v) for any $x \in U$ and any $v \in F(x)$.*
2. *For $x \in U$ and $v \in I(x)$ it holds that $\lim_{k\to\infty} \rho(x_k,v_k) = \infty$ for any sequence $(x_k,v_k) \to (x,v)$ such that $v_k \in F(x_k)$.*

Proof Observe first, that ρ is defined (possibly extended-valued) on $U \times \mathbb{S}^{m-1}$. To verify 1., consider any sequence $(x_k,v_k) \to_k (x,v)$ with $v_k \in \mathbb{S}^{m-1}$. We show first that the sequence $\rho(x_k,v_k)$ is bounded. Indeed, otherwise there would exist a subsequence with $\rho(x_{k_l},v_{k_l}) \to_l \infty$. Clearly $g(x_{k_l},0) < 0$ for l large enough, because of $g(x,0) < 0$. Fix an arbitrary $r \geq 0$. Then $\rho(x_{k_l},v_{k_l}) > r$. We claim that $g(x_{k_l},rLv_{k_l}) < 0$ for these l's. This is obvious in case that $v_{k_l} \in I(x_{k_l})$. If $v_{k_l} \in F(x_{k_l})$, then the relations

$$g(x_{k_l},0) < 0, \quad g(x_{k_l},\rho(x_{k_l},v_{k_l})Lv_{k_l}) = 0, \quad \rho(x_{k_l},v_{k_l}) > r,$$

and

$$g(x_{k_l},rLv_{k_l}) \geq 0,$$

would contradict the convexity of g in the second argument. Hence, for l sufficiently large,

$$g(x_{k_l}, rLv_{k_l}) < 0,$$

and passing to the limit yields that $g(x, rLv) \leq 0$, which holds true for all $r \geq 0$ because the latter was chosen arbitrary. But then, $g(x, rLv) < 0$ for all $r \geq 0$, because otherwise once more a contradiction with convexity of g in the second argument would arise from $g(x, 0) < 0$. This, however, amounts to $v \in I(x)$ contradicting our assumption $v \in F(x)$. Summarizing, we have shown that $\rho(x_k, v_k)$ is bounded and, in particular, $v_k \in F(x_k)$ for all k. Let $\rho(x_{k_l}, v_{k_l}) \to_l r_0$ be an arbitrary convergent subsequence. Then, we may pass to the limit in the relation $g(x_{k_l}, \rho(x_{k_l}, v_{k_l}) Lv_{k_l}) = 0$ in order to derive that $g(x, r_0 Lv) = 0$, which in turn implies that $r_0 = \rho(x, v)$. Hence, all convergent subsequences of $\rho(x_k, v_k)$ have the same limit $\rho(x, v)$. This implies that $\rho(x_k, v_k) \to_k \rho(x, v)$ and altogether that ρ is continuous at (x, v).

As for 2., observe that if $\rho(x_k, v_k)$ would not tend to infinity, then there would exist a converging subsequence $\rho(x_{k_l}, v_{k_l}) \to_l r_1$ for some $r_1 \geq 0$. Since $\rho(x_{k_l}, v_{k_l}) < \infty$ and $g(x_{k_l}, 0) < 0$ for l large enough, we infer that $v_{k_l} \in F(x_{k_l})$ and, hence, $g(x_{k_l}, \rho(x_{k_l}, v_{k_l}) Lv_{k_l}) = 0$ for all these l's. Now, passing to the limit yields that $g(x, r_1 Lv) = 0$, whence $v \in F(x)$, a contradiction.

Lemma 2 *If $g(x, 0) < 0$ and $v \in F(x)$, then there exist neighborhoods U and V of x and v, respectively, such that $v' \in F(x')$ for all $x' \in U$ and $v' \in V \cap \mathbb{S}^{m-1}$.*

Proof If the statement wasn't true, then there existed a sequence $(x_k, v_k) \to (x, v)$ with $g(x_k, 0) < 0$, $v_k \in \mathbb{S}^{m-1}$ and $v_k \in I(x_k)$. Hence, $\rho(x_k, v_k) = \infty$ and so $\rho(x, v) = \infty$ by 1. in Lemma 1. This yields the contradiction $v \in I(x)$.

Lemma 3 *Let $x \in X$ and $r \geq 0$ be such that $g(x, 0) < 0$ and $g(x, rLv) = 0$. Then*

$$\langle z^*, Lv \rangle \geq -\frac{g(x, 0)}{r} > 0 \quad \forall z^* \in \partial_z g(x, rLv).$$

Proof By convexity of g in the second variable and by definition of the convex subdifferential, one has that

$$\begin{aligned} -\frac{r}{2} \langle z^*, Lv \rangle &= \left\langle z^*, \frac{r}{2} Lv - rLv \right\rangle \leq g\left(x, \frac{r}{2} Lv\right) - g(x, rLv) \\ &= g\left(x, \frac{r}{2} Lv\right) \leq \frac{1}{2} g(x, 0) + \frac{1}{2} g(x, rLv) = \frac{1}{2} g(x, 0). \end{aligned}$$

Since our assumptions imply that $r > 0$, the assertion follows.

We get in the following proposition the desired continuity of the radial probability function e defined in (3).

Proposition 1 *The radial probability function is continuous at any $(x, v) \in X \times \mathbb{S}^{m-1}$ with $g(x, 0) < 0$.*

Proof Fix a point $(x, v) \in X \times \mathbb{S}^{m-1}$ with $g(x, 0) < 0$. Consider any sequence $(x_k, v_k) \to (x, v)$ with $v_k \in \mathbb{S}^{m-1}$ and assume first that $v \in F(x)$. Then, $\rho(x_k, v_k) \to_k \rho(x, v)$ by 1. in Lemma 1, and $v_k \in F(x_k)$ for k large, by Lemma 2. Hence, by (5) it follows that

$$e(x_k, v_k) = F_\eta(\rho(x_k, v_k)) \to_k F_\eta(\rho(x, v)) = e(x, v),$$

where the convergence follows from the continuity of the Chi-distribution function F_η.

If in contrast $v \in I(x)$, then, by (3), $e(x, v) = \mu_\eta(\mathbb{R}_+) = 1$. We'll be done if we can show that $e(x_k, v_k) \to_k 1$. If this did not hold true, then there would exist a subsequence and some $\varepsilon > 0$ such that

$$|e(x_{k_l}, v_{k_l}) - 1| > \varepsilon \quad \forall\, l. \tag{10}$$

Since $v_{k_l} \in I(x_{k_l})$ would imply as above that $e(x_{k_l}, v_{k_l}) = \mu_\eta(\mathbb{R}_+) = 1$, a contradiction, we conclude that $v_{k_l} \in F(x_{k_l})$ for all l. Now, 2. in Lemma 1 guarantees that $\rho(x_{k_l}, v_{k_l}) \to_l \infty$. Then, by (5), we arrive at the convergence

$$e(x_{k_l}, v_{k_l}) = F_\eta(\rho(x_{k_l}, v_{k_l})) \to_l 1,$$

where we exploited the property $\lim\limits_{t\to\infty} F_\eta(t) = 1$, following from F_η being a cumulative distribution function. This is a contradiction with (10), and the desired conclusion follows.

Consequently, we obtain the continuity of the probability function φ, defined in (1).

Theorem 1 *The probability function is continuous at any point $x \in X$ with $g(x, 0) < 0$.*

Proof For any sequence $x_n \to x$ one has by Proposition 1 that

$$e(x_n, v) \to_n e(x, v) \le 1 \quad \forall v \in \mathbb{S}^{m-1},$$

where the inequality follows from e being a probability. Since the constant function 1 is integrable on $\mathbb{S}^{m-1}$, the assertion follows from Lebesgue's dominated convergence theorem.

4 Subdifferential of the radial probability function

In this section, we provide characterizations of the Fréchet subdifferential of the radial probability function $e(\cdot, v)$, defined in (3), for arbitrarily fixed directions $v \in \mathbb{S}^{m-1}$. As before, we also consider in this section our standard assumption (H).

We need first to estimate the set $\partial_x^F \rho(x, v)$:

Proposition 2 *Let $x \in X$ with $g(x,0) < 0$ and $v \in F(x)$ be arbitrary. Then, for every $y^* \in \partial_x^F \rho(x,v)$ and every $w \in X$, there exist $x^* \in \partial_x^C g(x, \rho(x,v)Lv)$ and $z^* \in \partial_z g\left(x, \rho(x,v)Lv\right)$ such that $\langle z^*, Lv\rangle > 0$ and*

$$\langle y^*, w\rangle \leq \frac{-1}{\langle z^*, Lv\rangle} \langle x^*, w\rangle .$$

Proof Fix $y^* \in \partial_x^F \rho(x,v)$ and $w \in X$; hence, $\rho(x,v) < \infty$ (because by assumption $v \in F(x)$). Let $M > 0$ be a Lipschitz constant of g at $(x, \rho(x,v)Lv)$. Then, there exists a neighborhood U of x such that the function $g(\cdot, \rho(x,v)Lv)$ is locally Lipschitzian with Lipschitz constant M at each $x' \in U$, and such that the functions $g(x', \cdot)$, $x' \in U$, are locally Lipschitzian with the same Lipschitz constant M at $\rho(x,v)Lv$. As a consequence of [4, Proposition 2.1.2], for all $x' \in U$ one has that

$$\|x^*\|, \|z^*\| \leq M \quad \forall x^* \in \partial_x^C g(x', \rho(x,v)Lv),\ \forall z^* \in \partial_z g(x', \rho(x,v)Lv). \tag{11}$$

Consider an arbitrary sequence $t_n \downarrow 0$ so that, by Lemma 2, we may assume $v \in F(x + t_n w)$ for all n. By convexity and continuity of the function g with respect to the second variable, the set $\partial g\left(x + t_n w, \cdot\right)\left(\rho(x,v)Lv\right)$ is nonempty for all n, and so we may select a sequence

$$z_n^* \in \partial_z g\left(x + t_n w, \cdot\right)\left(\rho(x,v)Lv\right); \tag{12}$$

hence, taking into account, from the definition of function ρ, that $g(x + t_n w, \rho(x + t_n w, v)Lv) = 0$ and $g(x, \rho(x,v)Lv) = 0$,

$$\begin{aligned}
\left(\rho(x + t_n w, v) - \rho(x,v)\right)\langle z_n^*, Lv\rangle &= \langle z_n^*, \rho(x + t_n w, v)Lv - \rho(x,v)Lv\rangle \\
&\leq g(x + t_n w, \rho(x + t_n w, v)Lv) \\
&\qquad - g(x + t_n w, \rho(x,v)Lv) \\
&= -g(x + t_n w, \rho(x,v)Lv) \\
&= g(x, \rho(x,v)Lv) - g(x + t_n w, \rho(x,v)Lv).
\end{aligned} \tag{13}$$

Next, Lebourg's mean value Theorem for Clarke's subdifferential [4, Theorem 2.3.7] yields some $\tau_n \in [0,1]$ and

$$x_n^* \in \partial_x^C g(x + \tau_n t_n w, \rho(x,v)Lv) \tag{14}$$

such that

$$g(x, \rho(x,v)Lv) - g(x + t_n w, \rho(x,v)Lv) \leq -t_n \langle x_n^*, w\rangle , \tag{15}$$

and, consequently, from (13),

$$\left(\rho(x + t_n w, v) - \rho(x,v)\right)\langle z_n^*, Lv\rangle \leq -t_n \langle x_n^*, w\rangle . \tag{16}$$

Since X is reflexive and $\|z_n^*\|$, $\|x_n^*\| \leq M$, by (11), there exists a subsequence $\left(x_{n_k}^*, z_{n_k}^*\right)$ and some $(x^*, z^*) \in X \times \mathbb{R}^m$ such that $x_{n_k}^* \rightharpoonup x^*$ and $z_{n_k}^* \to z^*$.

The weak*-closedness of the graph of Clarke's subdifferential [4, Proposition 2.1.5] along with (14) and (12) implies that

$$x^* \in \partial_x^C g(x, \rho(x,v)Lv),\ \ z^* \in \partial_z g\left(x, \rho(x,v)Lv\right). \tag{17}$$

Now, Lemma 3 implies that

$$\langle z^*, Lv\rangle \geq \frac{-g(x,0)}{\rho(x,v)} > 0,$$

and, so, by passing to the (inferior) limit in (16), we arrive at

$$\langle z^*, Lv\rangle \liminf_{n\to\infty} t_n^{-1}\left(\rho(x+t_n w, v) - \rho(x,v)\right) \leq -\left\langle x^*, w\right\rangle. \tag{18}$$

Therefore, since $y^* \in \partial_x^F \rho(x,v)$,

$$\langle y^*, w\rangle \leq \liminf_{n\to\infty} t_n^{-1}\left(\rho(x+t_n w, v) - \rho(x,v)\right) \leq \frac{-1}{\langle z^*, Lv\rangle}\left\langle x^*, w\right\rangle,$$

as we wanted to prove.

Next, we give the desired estimate of the set $\partial_x^F e(x,v)$. Recall that χ is the density of the one-dimensional Chi-distribution with m degrees of freedom (see (6)).

Theorem 2 *Let $x \in X$ with $g(x,0) < 0$ and $v \in F(x)$ be arbitrary. Then, for every $y^* \in \partial_x^F e(x,v)$ and every $w \in X$, there exist $x^* \in \partial_x^C g(x, \rho(x,v)Lv)$ and $z^* \in \partial_z g\left(x, \rho(x,v)Lv\right)$ such that*

$$\langle y^*, w\rangle \leq \frac{-\chi\left(\rho(x,v)\right)}{\langle z^*, Lv\rangle}\left\langle x^*, w\right\rangle.$$

Consequently, if $M_{x,v}$ denotes a Lipschitz constant of $g(\cdot, \rho(x,v)Lv)$ at x, then

$$\|y^*\| \leq \frac{\rho(x,v)\cdot\chi\left(\rho(x,v)\right)}{|g(x,0)|} M_{x,v} \quad \forall y^* \in \partial_x^F e(x,v).$$

Proof By (5), for all y close to x we may write $e\left(y,v\right) = F_\eta\left(\rho(y,v)\right)$, with $\rho(y,v) < \infty$, as a consequence of Lemma 2. Since F_η is continuously differentiable and nondecreasing (as a distribution function), $F_\eta'\left(t\right) \geq 0$ for all $t \in \mathbb{R}$ and, from the calculus of Fréchet subdifferentials (e.g., [12, Corollary 1.14.1 and Proposition 1.11]), we obtain that

$$\begin{aligned}\partial_x^F e(x,v) &= \partial^F\left(F_\eta'\left(\rho(x,v)\right)\rho(\cdot,v)\right)(x)\\ &= F_\eta'(\rho(x,v))\partial^F\rho(\cdot,v)(x) = \chi\left(\rho(x,v)\right)\partial_x^F\rho(x,v).\end{aligned}$$

Combination with Proposition 2 yields the first assertion.

To prove the second assertion, from the first part of the proposition we choose elements $x^* \in \partial_x^C g(x, \rho(x,v)Lv)$ and $z^* \in \partial_z g\left(x, \rho(x,v)Lv\right)$ such that

$$\langle y^*, w\rangle \leq \left|\frac{-\chi\left(\rho(x,v)\right)}{\langle z^*, Lv\rangle}\right| \|x^*\|\,\|w\|,$$

and so, since $\langle z^*, Lv\rangle \geq \frac{-g(x,0)}{\rho(x,v)} > 0$ by Lemma 3,

$$\langle y^*, w\rangle \leq \frac{\rho(x,v)\cdot\chi\left(\rho(x,v)\right)}{|g(x,0)|} M_{x,v}\left\|w\right\|,$$

yielding the desired conclusion.

We shall also need the following result.

Corollary 1 (i) *For every $x_0 \in X$ with $g(x_0,0) < 0$ and every $v_0 \in F(x_0)$ there exist neighborhoods $\tilde{U}$ of x_0 and $\tilde{V}$ of v_0 as well as some $\alpha > 0$ such that*

$$\partial_x^F e(x,v) \subseteq \mathbb{B}_\alpha^*\left(0\right) \quad \forall\,(x,v) \in \tilde{U} \times \left(\tilde{V} \cap \mathbb{S}^{m-1}\right). \tag{19}$$

(ii) *For all $x \in X$ with $g(x,0) < 0$ and for all $v \in I(x)$ one has that $\partial_x^F e(x,v) \subseteq \{0\}$.*

Proof (i) Let $M > 0$ and define open neighborhoods $\tilde{U}$ of x_0 and $\tilde{V}$ of v_0 such that M is a Lipschitz constant of g on $\tilde{U} \times \tilde{V}$ and, for all $(x,v) \in \tilde{U} \times \left(\tilde{V} \cap \mathbb{S}^{m-1}\right)$ (recall Lemma 2),

$$g(x,0) < 0,\ \ \rho(x,v) < \infty.$$

Hence, by Theorem 2,

$$\partial_x^F e(x,v) \subseteq \mathbb{B}^*_{\alpha(x,v)}\left(0\right),$$

where

$$\alpha(x,v) := \frac{\rho(x,v)\cdot\chi\left(\rho(x,v)\right)}{|g(x,0)|} M_{x,v}.$$

Taking into account the continuity of ρ (see Lemma 1), we may suppose for all $(x,v) \in \tilde{U} \times \left(\tilde{V} \cap \mathbb{S}^{m-1}\right)$ that M is a Lipschitz constant for $g(\cdot, \rho(x,v)Lv)$ at the point x ($\in \tilde{U}$). Thus, we can replace $M_{x,v}$ by M in the definition of α above. Moreover, since g is continuous (also by Lemma 1), as well as the Chi-density χ, we deduce that α is continuous on $\tilde{U} \times \left(\tilde{V} \cap \mathbb{S}^{m-1}\right)$. Then, after shrinking $\tilde{U} \times \tilde{V}$ if necessary, we may assume that for some $\alpha > 0$

$$\alpha(x,v) \leq \alpha \qquad \forall\,(x,v) \in \tilde{U} \times \left(\tilde{V} \cap \mathbb{S}^{m-1}\right).$$

This proves (19).

(ii) As already observed in the proof of Proposition 1, $v \in I(x)$ implies that $e(x,v) = 1$. Consequently, the function $e(\cdot, v)$ (as the value of a probability) reaches a global maximum at x. Let $x^* \in \partial_x^F e(x,v)$ and $u \in X\backslash\{0\}$ be arbitrary. Then,

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

Hence $\langle x^*, u\rangle \leq 0$ for all $u \in X$, and so $x^* = 0$ as desired.

Definition 3 For $x \in X$ and $l > 0$, we call

$$C_l(x) := \{h \in X \mid g^\circ(\cdot, z)(y; h) \leq l \left\|z\right\|^{-m} e^{\frac{\|z\|^2}{2\|L\|^2}} \left\|h\right\| \;\; \forall y \in \mathbb{B}_{1/l}\left(x\right), \left\|z\right\| \geq l\}$$

the l-cone of nice directions at $x \in X$. We denote the polar cone to $C_l(x)$ as $C_l^*(x)$.

Note that, by positive homogeneity of Clarke's directional derivative, $\{C_l\}_{l\in\mathbb{N}}$ defines a nondecreasing sequence of closed cones.

We give in the following theorem another estimate for $\partial_x^F e(x, v)$, which will be useful in the sequel.

Theorem 3 *Fix $x_0 \in X$ such that $g(x_0, 0) < 0$. Then, for every $l > 0$, there exists some neighborhood U of x_0 and some $R > 0$ such that*

$$\partial_x^F e(x, v) \subseteq \mathbb{B}_R^*\left(0\right) - C_l^*(x_0) \quad \forall x \in U, \, v \in \mathbb{S}^{m-1}.$$

Proof Let $l > 0$ be arbitrarily fixed. It will be sufficient to show that for every $v_0 \in \mathbb{S}^{m-1}$ there are neighborhoods $\bar{U}$ of x_0 and $\bar{V}$ of v_0 and some $R > 0$ such that

$$\partial_x^F e(x, v) \subseteq \mathbb{B}_R^*\left(0\right) - C_l^*(x_0) \quad \forall\, (x, v) \in \bar{U} \times (\bar{V} \cap \mathbb{S}^{m-1}). \tag{20}$$

If this holds true, then the global inclusion in the statement of this proposition will follow from the local ones above by a standard compactness argument with respect to $\mathbb{S}^{m-1}$.

In order to prove (20), fix an arbitrary $v_0 \in \mathbb{S}^{m-1}$. Assume first that $v_0 \in I(x_0)$. Then, define open neighborhoods U^* of x_0 and V^* of v_0 such that $U^* \subseteq \mathbb{B}_{1/l}\left(x_0\right)$ (with $l > 0$ as fixed above) and, for all $x \in U^*$ and $v \in V^* \cap F(x)$,

$$g(x, 0) \leq \frac{1}{2} g(x_0, 0) < 0, \;\; \rho(x, v)\|Lv\| \geq l.$$

Note, that the last inequality is possible by virtue of 2. in Lemma 1 and by L being nonsingular and $\mathbb{S}^{m-1}$ being compact (therefore $\|Lv\| \geq \delta$ for all $v \in \mathbb{S}^{m-1}$ and some $\delta > 0$). From Corollary 1(ii) we derive that

$$\partial_x^F e(x, v) \subseteq \{0\} \quad \forall x \in U^*, \;\; v \in I(x). \tag{21}$$

Now, consider an arbitrary $(x, v) \in U^* \times V^*$ such that $v \in F(x)$. Let also $y^* \in \partial_x^F e(x, v)$ and $h \in -C_l(x_0)$ be arbitrarily given. Then, by Theorem 2, there exist $x^* \in \partial_x^C g(x, \rho(x, v)Lv)$ and $z^* \in \partial_z g\left(x, \rho(x, v)Lv\right)$ such that

$$\langle y^*, h\rangle \leq \frac{\chi\left(\rho(x, v)\right)}{\langle z^*, Lv\rangle} \langle x^*, -h\rangle \leq \frac{\chi\left(\rho(x, v)\right)}{\langle z^*, Lv\rangle} g^\circ(\cdot, \rho(x, v)Lv)(x; -h), \tag{22}$$

where the last inequality relies on (9) and on the fact that both the density function χ and $\langle z^*, Lv\rangle$ are positive (see Lemma 3). Since $-h \in C_l(x_0)$, our conditions on the neighborhoods U^* and V^* stated above guarantee that

$$\begin{aligned} g^\circ(\cdot, \rho(x,v)Lv)(x; -h) &\le l \, \|\rho(x,v)Lv\|^{-m} \, e^{\frac{\|\rho(x,v)Lv\|^2}{2\|L\|^2}} \, \|h\| \\ &\le l \, \|\rho(x,v)Lv\|^{-m} \, e^{\frac{\rho(x,v)^2}{2}} \, \|h\| \, . \end{aligned}$$

This allows us to continue (22) as

$$\begin{aligned} \langle y^*, h\rangle &\le \frac{\chi\left(\rho(x,v)\right)\rho(x,v) l}{|g(x,0)|} \, \|\rho(x,v)Lv\|^{-m} \, e^{\frac{\rho(x,v)^2}{2}} \, \|h\| \\ &= \frac{lK}{|g(x,0)|} \, \|Lv\|^{-m} \, \|h\| \, , \end{aligned}$$

where we used Lemma 3 and the definition of the Chi-density with m degrees of freedom (see (6)). Owing to $g(x,0) \le \frac{1}{2} g(x_0,0) < 0$, we may continue as

$$\langle y^*, h\rangle \le \frac{2lKK^*}{|g(x_0,0)|} \, \|h\| \, , \tag{23}$$

where (recall that L is nonsingular)

$$K^* := \max_{v \in \mathbb{S}^{m-1}} \|Lv\|^{-m} \in \mathbb{R}_+.$$

Consequently, we have shown that for some $\tilde{K} > 0$, which is independent of x and v,

$$\langle y^*, h\rangle \le \tilde{K} \, \|h\| \quad \forall y^* \in \partial_x^F e(x,v), \, h \in -C_l(x_0).$$

Using indicator and support functions, respectively, this relation is rewritten as, for all $h \in X$,

$$\begin{aligned} \langle y^*, h\rangle &\le \tilde{K} \, \|h\| + i_{-\overline{\text{co}}\, C_l(x_0)}(h) \\ &= \sigma_{\mathbb{B}^*_{\tilde{K}}(0)}(h) + \sigma_{-C_l^*(x_0)}(h) \\ &= \sigma_{\left(\mathbb{B}^*_{\tilde{K}}(0) - C_l^*(x_0)\right)}(h). \end{aligned}$$

Consequently, we get

$$\sigma_{\partial_x^F e(x,v)}(h) \le \sigma_{\left(\mathbb{B}^*_{\tilde{K}}(0) - C_l^*(x_0)\right)}(h) \quad \forall h \in X,$$

which entails the inclusion

$$\partial_x^F e(x,v) \subseteq \mathbb{B}^*_{\tilde{K}}(0) - C_l^*(x_0).$$

Since $(x,v) \in U^* \times V^*$ with $v \in F(x)$ were chosen arbitrarily, we may combine this with (21) to derive that

$$\partial_x^F e(x,v) \subseteq \mathbb{B}^*_{\tilde{K}}(0) - C_l^*(x_0) \quad \forall \, (x,v) \in U^* \times \left(V^* \cap \mathbb{S}^{m-1}\right).$$

Now, we suppose that $v_0 \in F(x_0)$. Then Corollary 1(i) guarantees the existence of neighborhoods $\tilde{U}$ of x_0 and $\tilde{V}$ of v_0 as well as some $\alpha > 0$ such that relation (19) holds true. Consequently, we end up with the claimed relation (20) upon putting

$$\bar{U} := \tilde{U} \cap U^*, \ \bar{V} := \tilde{V} \cap V^*, \ R := \max\{\alpha, \tilde{K}\}.$$

Corollary 2 *Fix $x_0 \in X$ such that $g(x_0, 0) < 0$, and assume one of the following alternative conditions:*

$$\{z \in \mathbb{R}^m \mid g(x_0, z) \leq 0\} \text{ is a bounded set,} \tag{24}$$

or

$$\exists\, l > 0 \text{ such that } C_l(x_0) = X. \tag{25}$$

Then the partial radial probability functions $e(\cdot, v)$, $v \in \mathbb{S}^{m-1}$, are uniformly locally Lipschitzian around x_0 with some common Lipschitz constant independent of v.

Proof In the case of (24), one has that $I(x_0) = \emptyset$, whence $F(x_0) = \mathbb{S}^{m-1}$. Then, by Corollary 1(i), for every $v_0 \in \mathbb{S}^{m-1}$ there exist neighborhoods $\tilde{U}_{v_0}$ of x_0 and $\tilde{V}_{v_0}$ of v_0 as well as some $\alpha_{v_0} > 0$ such that

$$\partial_x^F e(x, v) \subseteq \mathbb{B}^*_{\alpha_{v_0}}(0) \quad \forall (x, v) \in \tilde{U}_{v_0} \times \left(\tilde{V}_{v_0} \cap \mathbb{S}^{m-1}\right).$$

Then, by the evident compactness argument with respect to the sphere $\mathbb{S}^{m-1}$ already alluded to in the beginning of the proof of Theorem 3, we derive the existence of a neighborhood $\tilde{U}$ of x_0 and of some $\alpha > 0$ such that

$$\partial_x^F e(x, v) \subseteq \mathbb{B}^*_{\alpha}(0) \quad \forall (x, v) \in \tilde{U} \times \mathbb{S}^{m-1}.$$

In the case of (25), the same relation (with $\alpha := R$) is a direct consequence of Theorem 3 upon taking into account that $C_l(x_0) = X$ entails that $-C_l^*(x_0) = \{0\}$. Now, the claimed statement on uniform Lipschitz continuity follows from [13, Theorem 3.5.2].

5 Subdifferential of the Gaussian probability function φ

In this section, we provide the required formulae for the Fréchet, the Mordukhovich, and the Clarke subdifferentials of the Gaussian probability function φ, defined in (1). These results are next illustrated in Example 1, and in Subsection 5.3 to discuss the Lipschitz continuity and differentiability of φ. Finally, we study in this section, Subsection 5.4, the special and interesting setting of probability functions given by means of finite systems of smooth inequalities. In this case, formulae of the subdifferentials of φ are expressed in terms of the initial data in (1), i.e., in terms of the function g. All this is done under our standard assumption (H).

5.1 Main Result

We start by recalling the following result on the interchange of Mordukhovich subdifferentials and the integration sign when dealing with the integral functions of the form

$$I_f(x) := \int\limits_{\omega\in\Omega} f(\omega, x)d\mu.$$

Here, $(\Omega, \mathcal{A}, \nu)$ a σ-finite complete measure space, and $f : \Omega \times X \to [0, +\infty]$ is a normal integrand; that is,
(i) f is $\mathcal{A} \otimes \mathcal{B}(X)$-measurable,
(ii) $f(\omega, \cdot)$ is lsc for all $\omega \in \Omega$.

We assume that $I_f(x_0) < +\infty$ for some $x_0 \in X$. Then we have the following result in which the integral $\int\limits_{\omega\in\Omega} \partial^M f(\omega, x_0)d\nu$ is to be understood in the Aumann's sense; that is, the set of Bochner integrals over all measurable selections of the multivalued mapping $\partial^M f(\cdot, x_0)$ (see, e.g., [20]).

Proposition 3 ([5]) *Assume that for some $\delta > 0$ and $\mathcal{K} \in L^1(\Omega, \mathbb{R})$ we have*

$$\partial_x^F f(\omega, x) \subseteq \mathcal{K}(\omega)B_1^*(0) + C,\ \forall x \in B_\delta(x_0),\ \omega \in \Omega, \tag{26}$$

where $C \subseteq X^*$ is a closed convex cone with polar cone having a nonempty interior. Then

(i) $\partial^M I_f(x_0) \subseteq \mathrm{cl}^* \left\{ \int\limits_{\omega\in\Omega} \partial^M f(\omega, x_0)d\nu\,(\omega) + C \right\}$.

(ii) *Provided that X is finite-dimensional,*

$$\partial^M I_f(x_0) \subseteq \int\limits_{\omega\in\Omega} \partial^M f(\omega, x_0)d\nu\,(\omega) + C.$$

(iii) $\partial^\infty I_f(x_0) \subseteq C$.

(vi) $\partial^C I_f(x_0) \subseteq \overline{\mathrm{co}} \left\{ \int\limits_{\omega\in\Omega} \partial^M f(\omega, x_0)d\nu\,(\omega) + C \right\}$.

Now, we are in a position to prove the main result of our paper.

Theorem 4 *Let $x_0 \in X$ be such that $g(x_0, 0) < 0$. Assume that the cone $C_l(x_0)$ has a non-empty interior for some $l > 0$. Then,*

(i) $\partial^M \varphi(x_0) \subseteq \mathrm{cl}^* \left\{ \int\limits_{v\in\mathbb{S}^{m-1}} \partial_x^M e(x_0, v)d\mu_\zeta(v) - C_l^*(x_0) \right\}$

(ii) *Provided that X is finite-dimensional,*

$$\partial^M \varphi(x_0) \subseteq \int\limits_{v\in\mathbb{S}^{m-1}} \partial_x^M e(x_0, v)d\mu_\zeta(v) - C_l^*(x_0).$$

(iii) $\partial^\infty \varphi(x_0) \subseteq -C_l^*(x_0)$.

(vi) $\partial^C \varphi(x_0) \subseteq \overline{\text{co}} \left\{ \int\limits_{v\in\mathbb{S}^{m-1}} \partial_x^M e(x_0, v) d\mu_\zeta(v) - C_l^*(x_0) \right\}$.

Proof We apply Proposition 3 by putting

$$f(\omega, x) := e(x, \omega), \; C := -C_l^*(x_0),$$

and using the measurable space $(\mathbb{S}^{m-1}, \mathcal{A}, \mu_\zeta)$, with $\mathcal{A}$ being the σ-Algebra of measurable sets with respect to μ_ζ. It is known that μ_ζ is σ-finite and complete. The measurability property of f and the lower semicontinuity of $f(\omega, \cdot)$ are consequences of the continuity of e (see Proposition 1). The cone $C^* = \overline{\text{co}}\, C_l(x_0)$ has a non-empty interior, by the current assumption. Condition (26) is a consequence of Theorem 3 upon defining $\mathcal{K}(\omega) := R$ for all $\omega \in \Omega = \mathbb{S}^{m-1}$, and observing that $\mathcal{K} \in L^1(\mathbb{S}^{m-1}, \mathbb{R})$, due to $\mathbb{S}^{m-1}$ having finite (μ_ζ-) measure. Now, the claimed result follows from Proposition 3 by taking into account that $I_f = \varphi$ thanks to (2).

Our main result motivates some investigation about the impact of the parameter $l > 0$ in the definition of the cones $C_l^*(x_0)$, $x_0 \in X$. From Definition 3, it follows immediately that $(C_l(x_0))_{l\geq 0}$ forms a non-decreasing family of closed cones, and hence

$$C_k(x_0) \subseteq C_{k+1}(x_0); \quad C_k^*(x_0) \supseteqq C_{k+1}^*(x_0) \quad \forall k \in \mathbb{N}. \tag{27}$$

Moreover, $C_k(x_0)$ having a non-empty interior as required in Theorem 4, implies that $C_{k+1}(x_0)$ does so too. This means that the upper estimates in the results of Theorem 4 become increasingly precise for $k \to \infty$. This immediately raises the question if we may pass to the limit in this result. Let us then introduce the *limiting cone of nice directions*

$$C_\infty(x_0) := \bigcup_{k\in\mathbb{N}} C_k(x_0) =$$

$$\{h \in X \mid \exists k \in \mathbb{N} : g^\circ(\cdot, z)(y; h) \leq k \|z\|^{-m} e^{\frac{\|z\|^2}{2\|L\|^2}} \|h\|, \forall y \in \mathbb{B}_{\frac{1}{k}}(x), \|z\| \geq k\}.$$

The reader can simply notice (through Baire's Theorem) the non-emptiness of the interior of $C_\infty(x_0)$ is equivalent to the non-emptiness of the interior of $C_l(x_0)$ for some $l > 0$. As far as the singular subdifferential is concerned, we may immediately pass to the limit:

Proposition 4 *Fix $x_0 \in X$ with $g(x_0, 0) < 0$, and assume that $C_l(x_0)$ has a non-empty interior for some $l > 0$. Then $\partial^\infty \varphi(x_0) \subseteq -C_\infty^*(x_0)$.*

Proof By Theorem 4(iii) we have that $\partial^\infty \varphi(x_0) \subseteq -C_l^*(x_0)$. Since along with $C_l(x_0)$ the larger cones $C_k(x_0)$ for $k \in \mathbb{N}, k \geq l$, have non-empty interiors too, it follows that

$$\partial^\infty \varphi(x_0) \subseteq \bigcap_{k\in\mathbb{N}, k\geq l} -C_k^*(x_0) = -\left(\bigcup_{k\in\mathbb{N}} C_k(x_0)\right)^* = -C_\infty^*(x_0),$$

where the first equality relies on (27).

In order to formulate a corresponding result for the Mordukhovich and Clarke subdifferentials, we need an additional boundedness assumption:

Proposition 5 *Fix $x_0 \in X$ with $g(x_0, 0) < 0$, and assume that $C_l(x_0)$ has a non-empty interior for some $l > 0$. Moreover, suppose that $\partial_x^M e(x_0, v)$ is integrably bounded; i.e., there exists some function $\mathcal{R} : \mathbb{S}^{m-1} \to \mathbb{R}_+$ with $\int_{\mathbb{S}^{m-1}} \mathcal{R}(v) d\mu_\zeta(v) < \infty$ such that*

$$\partial_x^M e(x_0, v) \subseteq \mathbb{B}^*_{\mathcal{R}(v)}(0) \quad \mu_\zeta - a.e.\ v \in \mathbb{S}^{m-1}.$$

Then

$$\partial^M \varphi(x_0) \subseteq \partial^C \varphi(x_0) \subseteq \mathrm{cl} \left\{ \int_{v \in \mathbb{S}^{m-1}} \partial_x^M e(x_0, v) d\mu_\zeta(v) \right\} - C^*_\infty(x_0).$$

Proof For the purpose of abbreviation, put

$$\mathcal{I} := \int_{v \in \mathbb{S}^{m-1}} \partial_x^M e(x_0, v) d\mu_\zeta(v).$$

From our assumption on $\partial_x^M e(x_0, v)$, being integrably bounded, it follows that $\mathcal{I}$ is bounded too. Consequently, $\mathrm{cl}^* \mathcal{I}$ is w^*-compact. With $C_l(x_0)$ having a non-empty interior, for all $k \in \mathbb{N}$ with $k \geq l$, from Theorem 4(i) it follows that

$$\partial^M \varphi(x_0) \subseteq \mathrm{cl}^* \left\{ \mathcal{I} - C_k^*(x_0) \right\} = \mathrm{cl}^* \mathcal{I} - C_k^*(x_0) \quad \forall k \geq l.$$

Due to (27), we may continue as

$$\partial^M \varphi(x_0) \subseteq \bigcap_{k \in \mathbb{N}} \left\{ \mathrm{cl}^* \mathcal{I} - C_k^*(x_0) \right\}, \tag{28}$$

which in turn, using again the w^*-compacity of $\mathrm{cl}^* \mathcal{I}$, gives us

$$\partial^M \varphi(x_0) \subseteq \mathrm{cl}^* \mathcal{I} - \bigcap_{k \in \mathbb{N}} C_k^*(x_0) = \mathrm{cl}^* \mathcal{I} - \left(\bigcup_{k \in \mathbb{N}} C_k(x_0) \right)^* = \mathrm{cl}^* \mathcal{I} - C^*_\infty(x_0).$$

Now, by [13, Theorem 3.57], by Proposition 4, and by convexity of $C^*_\infty(x_0)$, we arrive at

$$\begin{aligned} \partial^C \varphi(x_0) &= \overline{\mathrm{co}} \left\{ \partial^M \varphi(x_0) + \partial^\infty \varphi(x_0) \right\} \\ &\subseteq \overline{\mathrm{co}} \left\{ \mathrm{cl}^* \mathcal{I} - C^*_\infty(x_0) - C^*_\infty(x_0) \right\} \\ &= \overline{\mathrm{co}} \left\{ \mathrm{cl}^* \mathcal{I} - C^*_\infty(x_0) \right\}. \end{aligned}$$

Now, as a consequence of [16, Theorem 3.1], the strong closure $\mathrm{cl}\, \mathcal{I}$ is convex (the measure μ_ζ being nonatomic), so that $\mathrm{cl}^* \mathcal{I} = \mathrm{cl}\, \mathcal{I}$ is convex, and the last inclusion above reads

$$\partial^C \varphi(x_0) \subseteq \mathrm{cl} \mathcal{I} - C^*_\infty(x_0).$$

This finishes the proof of our proposition.

5.2 An illustrating example

In the following, we provide an example which, on the one hand, serves as an illustration of our main result Theorem 4 and, on the other hand, shows that even for a continuously differentiable inequality $g(x,\xi) \leq 0$, satisfying a basic constraint qualification, the associated probability function φ may fail to be differentiable, actually even to be locally Lipschitzian (though it is continuous due to the constraint qualification).

Example 1 Define the function $g : \mathbb{R} \times \mathbb{R}^2 \to \mathbb{R}$ by

$$g(x, z_1, z_2) := \alpha(x)e^{h(z_1)} + z_2 - 1,$$

where

$$\alpha(x) := \begin{cases} x^2 & x \geq 0 \\ 0 & x < 0, \end{cases}$$

$$h(t) := -1 - 4\log(1 - \Phi(t))\,; \quad \Phi(t) := \frac{1}{\sqrt{2\pi}} \int\limits_{-\infty}^{t} e^{-\tau^2/2} d\tau,$$

i.e., Φ is the distribution function of the one-dimensional standard normal distribution. Moreover, let ξ have a bivariate standard normal distribution, i.e.,

$$\xi = (\xi_1, \xi_2) \sim \mathcal{N}\left((0,0), \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\right).$$

The following properties are shown in the Appendix:

1. g is continuously differentiable.
2. g is convex in (z_1, z_2).
3. $g(0,0,0) < 0$.
4. $C_1(0) = (-\infty, 0]$.
5. $\int\limits_{v \in \mathbb{S}^1} \partial_x^M e(0,v) d\mu_\zeta(v) \subseteq (-\infty, 0]$.
6. φ fails to be locally Lipschitzian in 0.

Observe that, by 1. and 2., g satisfies our basic data assumptions, (H), and that 3. forces the probability function φ to be continuous. On the other hand, by 6., φ is not locally Lipschitzian -much less differentiable - in 0 despite the continuous differentiability of g and the satisfaction of Slater's condition. Now, Theorem 4(ii), along with 4. and 5. provides that

$$\partial^M \varphi(0) \subseteq (-\infty, 0] - [0, \infty) = (-\infty, 0]\,, \quad \partial^\infty \varphi(0) \subseteq (-\infty, 0]\,.$$

On the other hand, analytical verification along with the formula for φ provided in the Appendix (or alternatively visual inspection of the graph of φ) yields that $\partial^M \varphi(0) = \{0\}$ and $\partial^\infty \varphi(0) = (-\infty, 0]$, so that the upper estimate for the singular subdifferential is strict, while the one for the basic subdifferential is not (nevertheless this upper estimate is nontrivial due to being smaller than the whole space).

5.3 Lipschitz continuity and differentiability of φ

The following result on Lipschitz continuity of the probability function φ is an immediate consequence of Clarke's Theorem on interchanging subdifferentiation and integration [4, Theorem 2.7.2] and of Corollary 2:

Theorem 5 *Fix $x \in X$ such that $g(x,0) < 0$. Under one of the alternative conditions* (24) *or* (25)*, the probability function φ is locally Lipschitz near x and the following estimate holds true*:

$$\partial^C \varphi(x) \subseteq \int\limits_{v \in \mathbb{S}^{m-1}} \partial_x^C e(x,v) d\mu_\zeta(v). \tag{29}$$

The next result provides conditions for differentiability of the probability function φ; recall that $\#A$ denotes the cardinal of a set A.

Proposition 6 *In addition to the assumptions of Theorem* 5*, assume that*

$$\#\partial_x^C e(x,v) = 1 \quad \mu_\zeta\text{-}a.e.\, v \in \mathbb{S}^{m-1}. \tag{30}$$

Then φ is strictly differentiable at x and

$$\nabla \varphi(x) = \int\limits_{v \in \mathbb{S}^{m-1}} \nabla_x e(x,v) d\mu_\zeta(v).$$

Consequently, if X is finite-dimensional and (30) *holds true in some neighborhood of x, then φ is even continuously differentiable at x.*

Proof Assumption (30) entails that the integral in (29) reduces to a singleton. On the other hand, the subdifferential on the left-hand side of (29) is nonempty, since φ is locally Lipschitz near x (see [4, Proposition 2.1.2]). Hence, the inclusion (29) yields the single-valuedness of $\partial^C \varphi(x)$ as well as the equality

$$\partial^C \varphi(x) = \int\limits_{v \in \mathbb{S}^{m-1}} \partial_x^C e(x,v) d\mu_\zeta(v).$$

Now, since a locally Lipschitzian function reducing to a singleton at some point is strictly differential at this point with gradient equal to the (single-valued) subdifferential (see [4, Proposition 2.2.4]), it follows that φ is strictly differentiable at x_0 and $\partial^C \varphi(x_0) = \{\nabla \varphi(x_0)\}$. Likewise, the local Lipschitz continuity of $e(\cdot, v)$ around x_0 for all $v \in \mathbb{S}^{m-1}$ (see Corollary 2) yields along with (30) that

$$\partial_x^C e(x_0, v) = \{\nabla_x e(x_0, v)\} \quad \mu_\zeta - a.e.\, v \in \mathbb{S}^{m-1}.$$

Altogether, we have shown the first assertion of our Proposition. The second assertion on continuous differentiability follows from [4, Corollary to Prop. 2.2.4].

5.4 Application to a finite system of smooth inequalities

In order to benefit from Theorem 4, one has to be able to express the integrand $\partial_x^M e(x_0, v)$ in terms of the initial data in (1), i.e., in terms of the function g. We will illustrate this for the case of a probability function defined over a finite system of continuously differentiable inequalities which are convex in their second argument:

$$\varphi(x) := \mathbb{P}\left(g_i\left(x, \xi\right) \leq 0,\ i = 1, \ldots, p\right),\ x \in X. \tag{31}$$

Clearly, this can be recast in the form of (1) upon defining

$$g := \max_{i=1,\ldots,p} g_i, \tag{32}$$

where g is locally Lipschitz as required and convex in the second argument because the g_i's are supposed to be so. Since $g\left(x, 0\right) < 0$ implies that $g_i\left(x, 0\right) < 0$ for all $i = 1, \ldots, p$, we may associate with each component a function ρ_i satisfying the relation $g_i\left(x, \rho_i\left(x, v\right) Lv\right) = 0$, as we did in (4). The relation between ρ associated via (4) with g in (32) is, clearly,

$$\rho\left(x, v\right) = \min_{i=1,\ldots,p} \rho_i\left(x, v\right) \quad \forall x : g\left(x, 0\right) < 0,\ \forall v \in F(x). \tag{33}$$

Note, however, that unlike ρ, the functions ρ_i are continuously differentiable because the g_i's are so. This is a consequence of the Implicit Function Theorem (see [1, Lemma 3.2]), which moreover yields the gradient formulae, for all x with $g\left(x, 0\right) < 0$ and all $v \in F(x)$,

$$\nabla_x \rho_i\left(x, v\right) = -\frac{1}{\left\langle \nabla_z g_i\left(x, \rho\left(x, v\right) Lv\right), Lv\right\rangle} \nabla_x g_i\left(x, \rho\left(x, v\right) Lv\right),\ i = 1, \ldots, p.$$

In the following proposition, we provide an explicit upper estimate of the subdifferential set $\partial_x^M e(x_0, v)$ in terms of the initial data, which can be used in the formula of Theorem 4 to get an upper estimate for the subdifferential of the probability function (31):

Proposition 7 *Fix $x \in X$ such that $g_i\left(x, 0\right) < 0$ for $i = 1, \ldots, p$. Then, for every $l > 0$, there exists some $R > 0$ such that the radial probability function associated with g in* (32) *via* (3) *satisfies*

$$\partial_x^M e(x, v) \subseteq \begin{cases} -\bigcup\limits_{i \in T(v)} \left\{ \frac{\chi(\rho(x,v))}{\langle \nabla_z g_i(x, \rho(x,v)Lv), Lv\rangle} \nabla_x g_i\left(x, \rho\left(x, v\right) Lv\right) \right\} & v \in F\left(x\right) \\ \mathbb{B}_R^*\left(0\right) - C_l^*\left(x\right) & v \in I\left(x\right). \end{cases}$$

Here, $T(v) := \{i \in \{1, \ldots, p\} \mid \rho_i\left(x, v\right) = \rho\left(x, v\right)\}$.

Proof Fix an arbitrary $v \in \mathbb{S}^{m-1}$. Given the continuity of e, we exploit the following representation [13, Theorem 2.34] of the Mordukhovich subdifferential in terms of the Fréchet subdifferential, which holds true in Asplund spaces (hence, in particular for reflexive Banach spaces)

$$x^* \in \partial_x^M e(x,v) \Longleftrightarrow \exists\, x_n \to_n x \text{ and } \exists\, x_n^* \rightharpoonup_n x^* : \; x_n^* \in \partial_x^F e(x_n, v).$$

Then, the inclusion $\partial_x^M e(x,v) \subseteq \mathbb{B}_R^*(0) - C_l^*(x)$ follows from Theorem 3, since $\mathbb{B}_R^*(0)$ is weak*-compact and $C_l^*(x)$ is weak*-closed, entailing that $\mathbb{B}_R^*(0) - C_l^*(x)$ is weak*-closed. This yields the desired estimate of $\partial_x^M e(x,v)$ when $v \in I(x)$.

Suppose now in addition that $v \in F(x)$, and, according to Lemma 2, let U be a neighborhood of x such that, for all $y \in U$,

$$g(y,0) < 0, \; v \in F(y).$$

From the proof of Theorem 2 we have seen that

$$\partial_x^F e(y,v) = \chi(\rho(y,v))\, \partial_x^F \rho(y,v), \quad \forall y \in U,$$

which, by continuity of χ and by 1. in Lemma 1, immediately entails that

$$\partial_x^M e(x,v) = \chi(\rho(x,v))\, \partial_x^M \rho(x,v).$$

From (33) and the calculus rule for minimum functions [13, Proposition 1.113] we conclude that

$$\partial_x^M \rho(x,v) \subseteq \bigcup_{i \in T(v)} \nabla_x \rho_i(x,v).$$

with $T(v)$ being defined as in the statement of the Proposition. Now, the assertion follows from (34).

We provide next a concrete characterization for the local Lipschitz continuity/differentiability of the probability function φ, defined in (31), along with an explicit subdifferential/gradient formula:

Theorem 6 *Fix $x_0 \in X$ with $g(x_0,0) < 0$, and assume that for some $l > 0$ it holds, for $i = 1, \ldots, p$,*

$$\|\nabla_x g_i(x,z)\| \le l\, \|z\|^{-m}\, e^{\frac{\|z\|^2}{2\|L\|^2}} \quad \forall x \in \mathbb{B}_{1/l}(x_0),\; \|z\| \ge l. \tag{34}$$

Then the probability function (31) *is locally Lipschitz near x_0 and there exists a nonnegative number $R \le \sup\{\|x^*\| \mid x^* \in \partial_x^M e(x_0,v)$ and $v \in I(x_0)\}$ such that*

$$\partial^C \varphi(x_0) \subseteq -\int_{v \in F(x_0)} \mathrm{co}\left\{ \bigcup_{i \in T(v)} \frac{\chi(\rho(x_0,v))\, \nabla_x g_i(x_0, \rho(x_0,v) Lv)}{\langle \nabla_z g_i(x_0, \rho(x_0,v) Lv), Lv \rangle} \right\} d\mu_\zeta(v)$$

$$+\mu_\zeta(I(x_0))\mathbb{B}_R^*(0).$$

Proof As a maximum of finitely many smooth functions, g is Clarke-regular, so that Clarke's directional derivative of g coincides with its usual directional derivative. Hence, by Danskin's Theorem and by (34), we get the following estimate, for all $h \in X$, $x \in \mathbb{B}_{1/l}(x_0)$ and $\|z\| \geq l$,

$$\begin{aligned} g^{\circ}(\cdot, z)(x; h) &= \langle \nabla_x g(x, z), h \rangle \\ &= \max \{ \langle \nabla_x g_i(x, z), h \rangle \, g_i(x, z) = g(x, z) \} \\ &\leq \max_{i=1,\dots,p} \langle \nabla_x g_i(x, z), h \rangle \leq l \, \|z\|^{-m} \, e^{\frac{\|z\|^2}{2\|L\|^2}} \, \|h\| \,. \end{aligned}$$

Hence, $C_l(x_0) = X$ and, so, Theorem 5 guarantees that φ in (31) is locally Lipschitz near x_0 and that

$$\partial^C \varphi(x_0) \subseteq \int\limits_{v \in F(x_0)} \partial_x^C e(x_0, v) d\mu_\zeta(v) + \int\limits_{v \in I(x_0)} \partial_x^C e(x_0, v) d\mu_\zeta(v). \tag{35}$$

Since $e(\cdot, v)$ is locally Lipschitzian for all $v \in \mathbb{S}^{m-1}$, it follows from [13, Theorem 3.57] and from Proposition 7 that

$$\begin{aligned} \partial_x^C e(x_0, v) &= \overline{\text{co}} \left\{ \partial_x^M e(x_0, v) \right\} \\ &= -\text{co} \left\{ \bigcup_{i \in T(v)} \frac{\chi(\rho(x_0, v)) \nabla_x g_i(x_0, \rho(x_0, v) Lv)}{\langle \nabla_z g_i(x_0, \rho(x_0, v) Lv), Lv \rangle} \right\}. \end{aligned}$$

Hence, the first term on the right-hand side of (35) coincides with the integral term in the asserted formula above. As for the second term, observe that $\partial_x^C e(x_0, v) \subseteq \mathbb{B}_R^*(0)$ for some $R > 0$ by Theorem 3, which yields the second term in the upper estimate of this theorem.

From Theorem 6 and Proposition 6, we immediately derive the following:

Corollary 3 *If in the setting of Theorem* 6 *one has that* $\mu_\zeta(I(x_0)) = 0$ *(in particular, under assumption* (24)*), or the constant* R *in Theorem 6 is zero, then*

$$\partial^C \varphi(x_0) \subseteq - \int\limits_{v \in \mathbb{S}^{m-1}} \text{co} \left\{ \bigcup_{i \in T(v)} \frac{\chi(\rho(x_0, v)) \nabla_x g_i(x_0, \rho(x_0, v) Lv)}{\langle \nabla_z g_i(x_0, \rho(x_0, v) Lv), Lv \rangle} \right\} d\mu_\zeta(v).$$

If, in addition, for μ_ζ*-a.e.* $v \in \mathbb{S}^{m-1}$ *we have that* $\#T(v) = 1$ *(say:* $T(v) = \{i^*(v)\}$*), then the probability function* (31) *is strictly differentiable with gradient*

$$\nabla \varphi(x_0) = - \int\limits_{v \in \mathbb{S}^{m-1}} \frac{\chi(\rho(x_0, v)) \nabla_x g_{i^*(v)}(x_0, \rho(x_0, v) Lv)}{\langle \nabla_z g_{i^*(v)}(x_0, \rho(x_0, v) Lv), Lv \rangle} d\mu_\zeta(v).$$

Remark 1 It is worth mentioning that under the strengthened (compared with (34)) growth condition

$$\exists l > 0 : \|\nabla_x g_i(x,z)\| \leq l e^{\|z\|} \quad \forall x \in \mathbb{B}_{1/l}(x_0), \ \|z\| \geq l, \ i = 1, \ldots, p$$

the constant R in Theorem 6 and Corollary above is zero, as it can be seen in (23) (see also [2, Theorem 3.6 and Theorem 4.1]).

6 Appendix

We verify in this Appendix properties 1.-6. in Example 1.

The continuous differentiability of g stated in 1. is obvious from the corresponding property of α and h. For h, this relies on the smoothness of the distribution function of the one-dimensional standard normal distribution Φ and on the fact that the argument $1 - \Phi(t)$ of the logarithm is always strictly positive.

By nonnegativity of α it is sufficient to check that $e^{h(t)}$ is convex in order to verify 2. To do so, it is sufficient to show that h itself is convex, which by definition would follow from the concavity of $\log(1 - \Phi(t))$. This, however, is a consequence of $\log \Phi$ being concave, which in turn implies that $\log(1 - \Phi)$ is concave (see [15, Theorem 4.2.4]).

Statement 3. follows immediately from the definition of the functions.

As for 4., observe first that, by continuous differentiability of g,

$$g^\circ(\cdot, z)(x; -1) = \nabla_x g(x, z_1, z_2) \cdot (-1) = -\alpha'(x) e^{h(z_1)} \leq 0 \quad \forall x, z_1, z_2 \in \mathbb{R},$$

whence $-1 \in C_1(0)$ by Definition 3. On the other hand, putting $x := 1$ and $z := (1, 0)$, we have that $x \in \mathbb{B}_1(0)$, $\|z\| = 1$ and

$$g^\circ(\cdot, z)(x; 1) = \nabla_x g(1, 1, 0) \cdot 1 = \alpha'(1) e^{h(1)} = 2e^{h(1)} \approx 1161,$$

whereas, due to $m = 2$ in this example,

$$\|z\|^{-m} e^{\frac{\|z\|^2}{2\|L\|^2}} = \sqrt{e} \approx 1.65.$$

Therefore, by Definition 3, $1 \notin C_1(0)$. Since $C_1(0)$ is a closed cone, this together with $-1 \in C_1(0)$ yields $C_1(0) = (-\infty, 0]$.

For proving 5., it is sufficient to show that

$$\partial_x^M e(0, v) \subseteq (-\infty, 0] \quad \forall v \in \mathbb{S}^1. \tag{36}$$

In order to calculate $\partial_x^M e(0, v)$ for an arbitrarily fixed $v \in \mathbb{S}^1$, we have to compute first the partial Fréchet subdifferentials $\partial_x^F e(x, v)$ for x in a neighborhood U of 0. Define U such that $g(x, 0, 0) < 0$ for all $x \in U$ (as a consequence of the already shown relation $g(0, 0, 0) < 0$). If $x < 0$, then, by definition of e and g,

$$e(x, v) = \mu_\eta(\{r \geq 0 \mid g(x, rLv) \leq 0\}) = \mu_\eta(\{r \geq 0 \mid rLv_2 \leq 1\}).$$

Hence, for $x < 0$, $e(x, v)$ does not depend on its first argument locally around x. Therefore, $\partial_x^F e(x, v) = \{0\}$ for all $x < 0$. Now, consider some $x \in U$ with $x \geq 0$ and $x^* \in \partial_x^F e(x, v)$. If $v \in I(x)$, then $\partial_x^F e(x, v) \subseteq \{0\}$ (see Corollary 1(ii)). If, in contrast, $v \in F(x)$, then, by Theorem 2 (putting $w := \pm 1$ there and observing that, by continuous differentiability of g, the partial Clarke subdifferentials reduce to partial gradients),

$$x^* = \frac{-\chi\left(\rho(x,v)\right)\nabla_x g\left(x, \rho(x,v)Lv\right)}{\left\langle \nabla_z g\left(x, \rho(x,v)Lv\right), Lv\right\rangle} = \frac{-2xe^{h(\rho(x,v)v_1)}\chi\left(\rho(x,v)\right)}{\left\langle \nabla_z g\left(x, \rho(x,v)Lv\right), Lv\right\rangle} \leq 0.$$

Here, the inequality relies on $x \geq 0$, on χ being positive as a density and on

$$\left\langle \nabla_z g\left(x, \rho(x,v)Lv\right), Lv\right\rangle \geq \frac{-g(x,0,0)}{\rho(x,v)} > 0$$

by Lemma 3. Altogether, we have shown that $\partial_x^F e(x, v) \subseteq (-\infty, 0]$ for all $x \in U$. This entails that also $\partial_x^M e(x, 0) \subseteq (-\infty, 0]$. Since $v \in \mathbb{S}^1$ has been fixed arbitrarily, the desired relation (36) follows.

In order to show 6. we provide first a formula for the probability function φ. If $t \leq 0$, then, by definition of g,

$$\varphi(t) = \mathbb{P}\left(g\left(x, \xi_1, \xi_2\right) \leq 0\right) = \mathbb{P}\left(\xi_2 \leq 1\right) = \Phi(1)$$

because $\xi_2 \sim \mathcal{N}\left(0,1\right)$ by the distribution assumption on ξ in Example 1. If $t > 0$, then, again by the assumed distribution of ξ,

$$\begin{aligned}
\varphi(t) &= \mathbb{P}\left(\xi_2 \leq 1 - t^2 e^{h(\xi_1)}\right) = \frac{1}{2\pi}\int_{-\infty}^{\infty}\left(\int_{-\infty}^{1-t^2e^{h(z_1)}} e^{-\left(z_1^2+z_2^2\right)/2} dz_2\right) dz_1 \\
&= \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-z_1^2/2}\cdot\frac{1}{\sqrt{2\pi}}\left(\int_{-\infty}^{1-t^2e^{h(z_1)}} e^{-z_2^2/2} dz_2\right) dz_1 \\
&= \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-s^2/2}\cdot\Phi\left(1 - t^2e^{h(s)}\right) ds.
\end{aligned}$$

Now, we are going to show that φ fails to be locally Lipschitz around 0. Observe first that, since Φ is increasing as a distribution function, h is increasing too by its definition. Then, for any s, t satisfying $s \geq \Phi^{-1}\left(1 - \sqrt{t}\right)$ (recall that Φ is strictly increasing and so its inverse exists) it holds that

$$h(s) \geq h\left(\Phi^{-1}\left(1 - \sqrt{t}\right)\right) = -1 - \log t^2.$$

Therefore, $t^2 e^{h(s)} \geq e^{-1}$. Thus, we have shown that

$$\Phi\left(1\right) - \Phi\left(1 - t^2e^{h(s)}\right) \geq \Phi\left(1\right) - \Phi\left(1 - e^{-1}\right) =: \varepsilon \quad \forall s, t : s \geq \Phi^{-1}\left(1 - \sqrt{t}\right).$$

With Φ being strictly increasing, we have that $\varepsilon > 0$. Now, for any $t > 0$, we calculate

$$\begin{aligned}
\varphi(0) - \varphi(t) &= \Phi(1) - \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-s^2/2} \cdot \Phi\left(1 - t^2 e^{h(s)}\right) ds \\
&= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-s^2/2} \cdot \left(\Phi(1) - \Phi\left(1 - t^2 e^{h(s)}\right)\right) ds \\
&\geq \varepsilon \frac{1}{\sqrt{2\pi}} \int_{\Phi^{-1}\left(1-\sqrt{t}\right)}^{\infty} e^{-s^2/2} ds = \varepsilon \left(1 - \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\Phi^{-1}\left(1-\sqrt{t}\right)} e^{-s^2/2} ds\right) \\
&= \varepsilon \left(1 - \Phi\left(\Phi^{-1}\left(1 - \sqrt{t}\right)\right)\right) = \varepsilon\sqrt{t}.
\end{aligned}$$

Since $\varepsilon > 0$, φ fails to be locally Lipschitz around 0, which finally shows 6.

References

1. W. Van Ackooij and R. Henrion, *Gradient formulae for nonlinear probabilistic constraints with Gaussian and Gaussian-like distributions*. SIAM J. Optim., 24 (2014), 1864-1889.
2. 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) 63-87.
3. 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. Comput., 83 (2014), pp. 2821–2851.
4. F.H. Clarke, Optimization and Nonsmooth Analysis, Classics in Applied Mathematics, vol.5. Society for Industrial and Applied Mathematics, 1987.
5. R. Correa, A. Hantoute and P. Pérez-Aros, *Sequential and exact formulae for the subdifferential of nonconvex integrand functionals*, submitted.
6. I. Deák, *Subroutines for Computing Normal Probabilities of Sets - Computer Experiences*, Ann. Oper. Res., 100 (2000), 103-122.
7. O. Ditlevsen and H. O. Madsen, Structural Reliability Method, Wiley, 1996.
8. A. Genz and F. Bretz, Computation of Multivariate Normal and t Probabilities, Lecture Notes in Statistics 195, Springer, Dordrecht, The Netherlands, 2009.
9. C. Gotzes, H. Heitsch, R. Henrion and R. Schultz, *Feasibility of nominations in stationary gas networks with random load*, Mathematical Methods of Operations Research, 84 (2016) 427-457.
10. T. González Grandón, H. Heitsch and R. Henrion, *A joint model of probabilistic/robust constraints with application to stationary gas networks*, submitted.
11. A. I. Kibzun and S. Uryas'ev, *Differentiability of probability function*, Stoch. Anal. Appl., 16 (1998), 1101-1128.
12. A.Y. Kruger, *On Fréchet subdifferentials*, J. Math. Sci, 116 (2003), 3325-3358.
13. B.S. Mordukhovich, Variational Analysis and Generalized Differentiation. Vol. 1: Basic Theory, Vol. 2: Applications. Springer, Berlin (2006).
14. G. Ch. Pflug and H. Weisshaupt, *Probability gradient estimation by set-valued calculus and applications in network design*, SIAM J. Optim., 15 (2005), 898–914.
15. A. Prékopa, Stochastic Programming, Kluwer, Dordrecht, 1995.
16. P. Pucci and G. Vitillaro, *A representation theorem for Aumann integrals*, J. Math. Anal. Appl., 102 (1984), 86-101.
17. R.T. Rockafellar and R. J.-B. Wets, Variational Analysis, Springer, Berlin, 1997.

18. A. Shapiro, D. Dentcheva and A. Ruszczyński, Lectures on Stochastic Programming, MPS-SIAM series on optimization Vol. 9, 2009.
19. S. Uryas'ev, *Derivatives of probability functions and integrals over sets given by inequalities*, J. Comput. Appl. Math., 56 (1994), 197–223.
20. C. Castaing and M. Valadier, Convex analysis and measurable multifunctions. Lecture Notes in Mathematics, Vol. 580. Springer-Verlag, Berlin-New York (1977).
21. R. J.-B. Wets, Stochastic Programming, in G. Nemhauser, A. Rinnooy Kan and M. Todd (eds.): Handbook for Operatins Research and Management Sciences, Vol. 1, pp. 573-629, Elsevier, 1989.