Computing the CDF of the function of a real-valued random vector with elliptical distribution: an original level-set based method

CAYET Pierre

Abstract

This paper proposes an original approach for the computation of the CDF of a continuous function h applied to a real-valued random vector $\mathbf{X}$, when this vector follows an elliptical distribution. Instead of sampling directly from the density of $\mathbf{X}$ and computing the proportion of draws for which $h(\mathbf{X})$ exceeds a given threshold $\alpha \in \mathbb{R}$, we propose an original level-set approach for calculating the CDF of $h(\mathbf{X})$, by expressing it as a weighted integral computed over the level-sets of its probability density function. Using Lebesgue integration and the fact that the level-sets of the density function of an elliptical distribution form a family of homothetic hyperellipsoids, we reformulate the probability $\mathbb{P}(h(\mathbf{X}) \geq \alpha)$ as a weighted integral taken over this family of hyperellipsoids, which Lebesgue measure can be computed under closed-form. We further generalize our expression to mixtures of elliptical distributions, and establish potentially fruitful connexions with differential forms by reformulating our integral as a weighted integral over the surface our family of hyperellipsoids using vector fields and the generalized Stokes theorem.

Finally, we show that the probability $\mathbb{P}(h(\mathbf{X}) \geq \alpha)$ can be reformulated in terms of the weighted expectation of a specific random function $\eta(\mathbf{U})$, where $\mathbf{U}$ follows a uniform distribution on the unit sphere. We provide a consistent estimator and assess the performances of our methods in terms of empirical bias and variance of the estimator.

Keywords: Elliptical distribution, Probability, Cumulative distribution function, Differential geometry

1 Introduction

This paper proposes an original approach for the computation of the cumulative distribution function (CDF) of a continuous real-valued multivariate function $h : \mathbb{R}^n \mapsto \mathbb{R}$ applied to a real-valued random vector $\mathbf{X} \in \mathbb{R}^n$, with $n > 1$. Estimating the CDF of a given random variable is a common issue in actuarial science and risk analysis. In the context of energy applications, computing the fair price of a Power Purchase Agreement (PPA) may for instance require calculating the distribution of profit for a portfolio of renewables energy sources production units, which generation is sold on the electricity spot market. However, there seem to exist no unified framework for estimating the CDF of $h(\mathbf{X})$.

Drawing insights from interval analysis, [**1**] provide a method for estimating the distribution of a function of random variables, which creates a sequence of bounds that convergence to the true distribution function for functions of independent variables or random variables of known dependencies. Their methods is decomposed into three steps: first, they construct a partition of the domain of each random variable, where each subintervals are equally spaced in probability mass relative to their marginal distributoins. Second, they construct upper and lower bounds for the conditional CDF of $h(\mathbf{X})$ for each "box" in the partition, before combining them in a final step to obtain the final estimate. However, convergence to the true CDF of $h(\mathbf{X})$ requires that $h(\mathbf{X})$ is continous and monotonically increasing for each individual component of the random vector $\mathbf{X}$. The case of non-monotonic functions requires first discovering subregions over which the function is monotonic. Global optimization techniques would be required to compute the upper and lower bounds over boxes without prior knowledge of the regions of monotonicity. This approach may rapidly become intractable when the dimension of $\mathbf{X}$ is large or when the partial derivatives of $h(\mathbf{X})$ rapidly change sign.

The CDF of a random variable may be recovered by inverting its characteristic function, through the Gil-Pelaez formula (see [**2**]), which relates the characteristic function and CDF of a random variable through the complex Fourier transform. This provides a powerful tool as it admits closed-form expressions for linear combinations of independent random variables or generalized quadratic forms for $\mathbf{X}$ following a multivariate normal distribution as shown in [**3**]. However, when the characteristic function of $h(\mathbf{X})$ is unknown or has no analytic formulation, the empirical characteristic function can be approximated from observed data before applying numerical integration methods, typically using grid evaluations through trapezoidal quadrature. Numerous heuristics are available to adjust the grid size to satisfy specific error tolerance levels, manage oscillatory integral issues or smooth the empirical characteristic function. Yet, this approach suffers severe limitations, especially when the characteristic function exhibits discontinuities or insufficient decay.

The saddlepoint approximation provides an alternative computation method for the CDF of a

function of random variables. [**4**] use this technique to derive a uniform asymptotic series approximating the CDF of the sum of a large number of independent random variables. Their method is illustrated with applications to the uniform and exponential distributions. In a similar fashion, [**5**] propose an extansion of the saddlepoint approximation method to compute the CDF of a finite sum of real-valued independent and identically distributed random vectors.

Introduced by [**6**] for Hermite polynomials and generalized by [**7**], Polynomial Chaos Expansion (PCE) uses spectral decomposition to approximate a random variable in terms of a linear combination of orthogonal polynomials of random variables, which is then propagated to compute its moments and density function. PCE provides a theoretically sound tool to quantify uncertainty in models which are parameterized by independent random variabkles. The generalized Polynomial Chaos (gPC) series expansion can be adapted to scalar functions of several random variables, by using a multivariate polynomial basis built from the tensor product of orthogonal univariate polynomials. This approach, known as functional chaos approximation (see [**8**]), allows us to directly sample from the gPC series expansion as it converges in the mean square sense to the target PDF. Yet, an additional integration step is required to approximate the CDF of the target function of random variables. The case of dependent random variables is however more intricate as it requires measure transformation methods that map the dependent random variables to a set of independent ones. [**9**] propose a discussion of the pros and cons of three methods that can be used to constructed the PCE of models with dependent variables, including a novel one using Gram-Schmidt orthogonalization to numerically compute orthonormal polynomials for the dependent random variables.

The machine learning literature also provides methodological tools for computing the CDF of $h(\mathbf{X})$. The framework of normalizing flows was formally introduced by [**10**] and [**11**], before being applied in the context of density estimation by [**12**]. The method of normalizing flows uses a sequence of transformations, noted $T = T_1 \circ \cdots \circ T_{N-1} \circ T_N$, with $N > 0$, to map simple distributions defined over continuous random variables, with known closed-form densities, to more complex ones. The composition of invertible mappings is itself invertible and allows the representation of transformations of arbitrary complexity. Normalizing flows rely on the change of variables formula to compute the change in volume associated to each transformation from its Jacobian. However, it assumes that each transformation is a diffeomorphism, such that it is bijective, differentiable, and its inverse is differentiable as well.

This paper proposes an original approach to the problem of computing the cumulative distribution function of a real-valued multivariate function applied to a random vector following an elliptical distribution. Elliptical distributions encompass a wide variety of well-known probability distributions, including the (multivariate) Gaussian distribution and $t-$distribution, but also the symmetric multivariate Laplace distribution and the multivariate logistic distribution. Elliptical distributions are thus ubiquitous and find various applications in statistics, especially in generalized multivariate analysis and robust analysis, and finance for optimal portfolio analysis and Value at Risk (VaR)

calculations. Formally, a random vector $\mathbf{X}$ follows an elliptical distribution if its density function can be expressed as a function of the quadratic form $\mathbf{x}^T\mathbf{\Sigma}\mathbf{x}$, where $\mathbf{\Sigma}$ is the variance-covariance matrix associated to $\mathbf{X}$. More generally, a random vector is said to follow an elliptical distribution if its characteristic function can be written as a function of $\mathbf{x}^T\mathbf{\Sigma}\mathbf{x}$. We focus on the first definition in the present article and leave formulations in terms of the characteristic function for further research.

For a given random vector $\mathbf{X}$ and scalar $\epsilon > 0$, the iso-density hypersurface associated to ϵ corresponds to the set of points for which the probability density function of $\mathbf{X}$ is equal to ϵ. In the case of elliptical distributions, each iso-density hypersurface geometrically corresponds to an hyperellipsoid or a union of hyperellipsoids. More generally, the level-sets of the density function form a family of homothetic hyperellipsoids with center corresponding to the mean of the random vector $\mathbf{X}$. Using tools from measure theory and calculus, we propose an original level-set approach for calculating the probability $\mathbb{P}\left(h\left(\mathbf{X}\right) \geq \alpha\right)$ for a given threshold $\alpha \in \mathbb{R}$, by expressing it as a weighted integral computed over the level-sets of its probability density function. We establish some potentially fruitful connexions with differential forms, by using the generalized Stokes theorem to express the target CDF using vector fields. Finally, we show that the probability $\mathbb{P}\left(h\left(\mathbf{X}\right) \geq \alpha\right)$ can be reformulated as the expectation of a weighted random function following a uniform distribution on the unit sphere. We provide a consistent estimator and assess the performances of our methods in terms of empirical bias and variance of the estimator.

2 Definitions and main theoretical results

2.1 Notations

Let $\mathbf{X}$ be a real random vector in $\mathbb{R}^n$, $n \in \mathbb{N}^*$, with mean vector $\boldsymbol{\mu} \in \mathbb{R}^n$,variance-covariance matrix $\mathbf{\Sigma} \in \mathbb{R}^{n\times n}$ and defined on the probability space $(\Omega_{\mathbf{X}}, \mathcal{F}, \mathbb{P})$. Without loss of generality, we assume $\mathbf{X}$ has zero mean, i.e. $\boldsymbol{\mu} = 0$. We further assume that $\mathbf{X}$ follows an elliptical distribution and its density function, noted $f_{\mathbf{X}}(\mathbf{x})$, $\mathbf{x} \in \Omega_{\mathbf{X}}$, can be expressed as an explicit function of $\boldsymbol{\mu}$ and $\mathbf{\Sigma}$. More specifically, we can write $f_{\mathbf{X}}(\mathbf{x}) = \Omega g\left((\mathbf{x}-\boldsymbol{\mu})^T\mathbf{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu})\right)$, $\Omega \in \mathbb{R}$ and where $g\colon \mathbb{R} \mapsto \mathbb{R}^+$. As $\boldsymbol{\mu} = 0$ by assumption, we further simplify to $f_{\mathbf{X}}(\mathbf{x}) = \Omega g\left(\mathbf{x}^T\mathbf{\Sigma}^{-1}\mathbf{x}\right)$.

As a reminder, we define the real-valued multivariate function $h : \mathbb{R}^n \mapsto \mathbb{R}$. The solution set of the equation $h(\mathbf{x}) = \alpha$, noted $\Phi_\alpha = \{\mathbf{x} \in \mathbb{R}^n \mid h(\mathbf{x}) = \alpha\}$, defines an implicit hypersurface associated to an equation of the form $\varphi(\mathbf{x}, \alpha) = 0$, where $\varphi(\mathbf{x}, \alpha) = h(\mathbf{x}) - \alpha$ and $\alpha \in \mathbb{R}$. We make the assumption that the hypersurface Φ_α is orientable. We can then define the complementary subspaces $\Phi_\alpha^+ = \{\mathbf{x} \in \mathbb{R}^n \mid \varphi(\mathbf{x}, \alpha) \geq 0\}$ and $\Phi_\alpha^- = \{\mathbf{x} \in \mathbb{R}^n \mid \varphi(\mathbf{x}, \alpha) < 0\}$, such that $\Phi_\alpha^+ \bigcap \Phi_\alpha^- = \emptyset$ and $\Phi_\alpha^+ \bigcup \Phi_\alpha^- = \mathbb{R}^n$. A significant advantage of the above formulation in terms of implicit hypersurface

is that it allows the computation of quantities for any graphical relationship between components of random vector $\mathbf{X}$ that can be expressed implicitely. Given two topological sets X and Y, a function verifies the restriction that $\forall x \in X, \exists! y \in Y$ such that $y = f(x)$, and thus belongs to a subset of the more general family of curves.

The probability of the event $(h(\mathbf{x}) \geq \alpha)$, or equivalently $(\varphi(\mathbf{x}, \alpha) \geq 0)$, is equal to the integral of the density function $f_{\mathbf{X}}(\mathbf{x})$ computed on the subspace Φ_k^+ :

$$\mathbb{P}\left(\varphi(\mathbf{x}, \alpha) \geq 0\right) = \int_{\Phi_\alpha^+} f_{\mathbf{X}}(\mathbf{z}) \mathrm{d}\mu\left(\mathbf{z}\right) = 1 - \int_{\Phi_\alpha^-} f_{\mathbf{X}}(\mathbf{z}) \mathrm{d}\mu\left(\mathbf{z}\right) \tag{1}$$

Directly computing the above expression under closed-form is generally not possible, which justifies the use of approximation methods.

The present paper introduces an original approach using the level-sets of the function $f_{\mathbf{X}}(\mathbf{x})$ to compute expression **(1)** efficiently. For $\epsilon \in \mathbb{R}$, $\epsilon \geq 0$, we define the level-set associated to ϵ as $\partial\mathcal{E}_\epsilon = \{\mathbf{x} \in \mathbb{R}^n \mid f_{\mathbf{X}}(\mathbf{x}) = \epsilon\}$ and the closed set $\mathcal{E}_\epsilon = \{\mathbf{x} \in \mathbb{R}^n \mid f_{\mathbf{X}}(\mathbf{x}) \geq \epsilon\}$. As by assumption the random vector $\mathbf{X}$ follows an elliptical distribution, $\mathcal{E}_\epsilon$ thus corresponds to an hyper-ellipsoid parameterized by the couple $(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, with boundary defined as $\partial\mathcal{E}_\epsilon$. We further define $\epsilon^* = \max_{\epsilon \in \mathbb{R}^{+*}} f_{\mathbf{X}}(\mathbf{x})$ corresponding to the maximum value taken by the density function, with $\epsilon^* < +\infty$ by definition. Finally, $\forall \epsilon \in [0, \epsilon^*]$, we define $\mu\left(\mathcal{C}_\epsilon\right)$ as the Lebesgue measure of the set $\mathcal{C}_\epsilon$.

2.2 Main theoretical results

2.2.1 Preliminaries

Before developing our approach in details, we start by proving the following preliminary result:

Proposition 1: Let $\Gamma_\tau = \tau\epsilon^*$, $0 \leq \tau \leq 1$. For any random vector $\mathbf{X}$ with density function $f_{\mathbf{X}}(\mathbf{x})$, the integral of the Lebesgue measure $\mu\left(\mathcal{E}_{\tau\epsilon^*}\right)$ along the path Γ_τ sums to one, i.e.:

$$\int_\Gamma \mu\left(\mathcal{E}_\Gamma\right) \mathrm{d}s = \epsilon^* \int_0^1 \mu\left(\mathcal{E}_{\tau\epsilon^*}\right) \mathrm{d}\tau = 1 \tag{2}$$

Proof: The proof of the above proposition is a simple consequence of the properties of density functions. We note $\mu(A)$ the Lebesgue measure of the set A. For any $\mathbf{x} \in \mathbb{R}^n$, $f_{\mathbf{X}}(\mathbf{x}) \geq 0$, while $f_{\mathbf{X}}(\mathbf{x}) > 0$ for $\mathbf{x} \in \Omega_{\mathbf{X}} \subseteq \mathbb{R}^n$. We use the convention that $\infty \cdot 0 = 0$. We have:

$$\int_{\mathbb{R}^n} f_{\mathbf{X}}(\mathbf{x}) \mathrm{d}\mu(\mathbf{x}) = \int_0^{+\infty} \mu\left(\{\mathbf{x} : (f_{\mathbf{X}}(\mathbf{x}) > t) \cup (f_{\mathbf{X}}(\mathbf{x}) = t)\}\right) \mathrm{d}t$$

Using the fact that $\mu\left(\{\mathbf{x} : (f_{\mathbf{X}}(\mathbf{x}) = t\}\right) = 0$ and $\mu\left(\{\mathbf{x} : (f_{\mathbf{X}}(\mathbf{x}) > \epsilon^*\}\right) = 0$ by definition, we can simplify the above expression:

$$\int_{\mathbb{R}^n} f_{\mathbf{X}}(\mathbf{x}) \mathrm{d}\mu(\mathbf{x}) = \int_0^{\epsilon^*} \mu\left(\{\mathbf{x} : (f_{\mathbf{X}}(\mathbf{x}) > t\}\right) \mathrm{d}t$$

Finally, by introducing the change of variable $t = \tau\epsilon^*$ and using the fact that the integral of $f_{\mathbf{X}}(\mathbf{x})$ sums to 1, we obtain the result sought-after:

$$\begin{aligned}
\int_{\mathbb{R}^n} f_{\mathbf{X}}(\mathbf{x}) \mathrm{d}\mu(\mathbf{x}) &= \int_0^{\epsilon^*} \mu\left(\{\mathbf{x} : (f_{\mathbf{X}}(\mathbf{x}) > t\}\right) \mathrm{d}t \\
&= \int_0^1 \mu\left(\{\mathbf{x} : (f_{\mathbf{X}}(\mathbf{x}) > \tau\epsilon^*\}\right) \epsilon^* \mathrm{d}\tau \\
&= \epsilon^* \int_0^1 \mu\left(\mathcal{E}_{\tau\epsilon^*}\right) \mathrm{d}\tau \\
&= 1
\end{aligned}$$

It is straightforward to show that we have $\epsilon^* = \Omega$.

2.2.2 Level-set reformulation of the CDF

Proposition 1 allows us to compute the probability of $\varphi(\mathbf{x}, \alpha) \geq 0$ by taking the integral of the Lebesgue measure of the hyper-ellipsoid $\mathcal{E}_\epsilon$. Let us consider the following decomposition:

$$\mu\left(\mathcal{E}_\epsilon\right) = \mu\left(\mathcal{E}_\epsilon \cap \Phi_\alpha^+\right) + \mu\left(\mathcal{E}_\epsilon \cap \Phi_\alpha^-\right)$$

We assume that there exists a function $\eta^+(\mathbf{x}, \alpha)$ with the following properties: $\mathbf{x} \in \Phi_\alpha^+ \Leftrightarrow \eta^+(\mathbf{x}, \alpha) = 1$, $\mathbf{x} \in \Phi_\alpha^- \Leftrightarrow \eta^+(\mathbf{x}, \alpha) = -1$. Then, it follows that:

$$\int_{\mathcal{E}_\epsilon} \eta^+(\mathbf{x}, \alpha) \mathrm{d}\mu(\mathbf{x}) = \int_{\mathcal{E}_\epsilon \bigcap \Phi_\alpha^+} \mathrm{d}\mu(\mathbf{x}) - \int_{\mathcal{E}_\epsilon \bigcap \Phi_\alpha^-} \mathrm{d}\mu(\mathbf{x}) \tag{3}$$

$$\iff \int_{\mathcal{E}_\epsilon} \mathrm{d}\mu\left(\mathbf{x}\right) + \int_{\mathcal{E}_\epsilon} \eta^+(\mathbf{x},\alpha)\mathrm{d}\mu\left(\mathbf{x}\right) = 2\int_{\mathcal{E}_\epsilon \bigcap \Phi_\alpha^+} \mathrm{d}\mu\left(\mathbf{x}\right) \tag{4}$$

$$\Rightarrow \mu\left(\mathcal{E}_\epsilon \cap \Phi_\alpha^+\right) = \int_{\mathcal{E}_\epsilon \bigcap \Phi_\alpha^+} \mathrm{d}\mu\left(\mathbf{x}\right) = \frac{1}{2}\left[\int_{\mathcal{E}_\epsilon} (1+\eta^+(\mathbf{x},\alpha))\mathrm{d}\mu\left(\mathbf{x}\right)\right] \tag{5}$$

The probability that $\varphi(\mathbf{x},\alpha) \geq 0$ can eventually be reformulated as follows:

$$\mathbb{P}\left(\varphi(\mathbf{x},\alpha) \geq 0\right) = \frac{\epsilon^*}{2}\int_0^1 \left[\int_{\mathcal{E}_{\tau\epsilon^*}} (1+\eta^+(\mathbf{z},\alpha))\mathrm{d}\mu\left(\mathbf{z}\right)\right]\mathrm{d}\tau$$

, where for all τ such that $0 \leq \tau \leq 1$, $\mathcal{E}_{\tau\epsilon^*}$ is a compact subset of $\mathbb{R}^n$, with piecewise smooth boundary $\partial\mathcal{E}_{\tau\epsilon^*}$. The above equation can be further developped as follows:

$$\mathbb{P}\left(\varphi(\mathbf{x},\alpha) \geq 0\right) = \frac{\epsilon^*}{2}\int_0^1 \left[\mu(\mathcal{E}_{\tau\epsilon^*}) + \int_{\mathcal{E}_{\tau\epsilon^*}} \eta^+(\mathbf{z},\alpha)\mathrm{d}\mu\left(\mathbf{z}\right)\right]d\tau = \frac{\epsilon^*}{2}\left[\frac{1}{\epsilon^*} + \int_0^1\int_{\mathcal{E}_{\tau\epsilon^*}} \eta^+(\mathbf{z},\alpha)\mathrm{d}\mu\left(\mathbf{z}\right)\mathrm{d}\tau\right]$$

$$\Longrightarrow \mathbb{P}\left(\varphi(\mathbf{x},\alpha) \geq 0\right) = \frac{1}{2} + \frac{\epsilon^*}{2}\int_0^1\int_{\mathcal{E}_{\tau\epsilon^*}} \left[\eta^+(\mathbf{z},\alpha)\mathrm{d}\mu\left(\mathbf{z}\right)\right]\mathrm{d}\tau \tag{6}$$

If $\varphi(\boldsymbol{\mu},\alpha) \neq 0$, there exists $\epsilon_0 > 0$ such that $\forall \mathbf{x} \in \mathcal{E}_{\epsilon_0} = \{\mathbf{x} \in \mathbb{R}^n | f_{\mathbf{X}}(\mathbf{x}) \geq \epsilon_0\}$, $\eta^+(\mathbf{x},\alpha) = 1$ if $\varphi(\boldsymbol{\mu},\alpha) > 0$ or $\eta^+(\mathbf{x},\alpha) = -1$ if $\varphi(\boldsymbol{\mu},\alpha) < 0$. For $\tau_0 > 0$ and $\epsilon_0 = \tau_0\epsilon^*$, we get the following expression:

$$\mathbb{P}\left(\varphi(\mathbf{x},\alpha) \geq 0\right) = \frac{1}{2} + \frac{\epsilon^*}{2}\left(\frac{\varphi(\boldsymbol{\mu},\alpha)}{|\varphi(\boldsymbol{\mu},\alpha)|}\int_0^{\tau_0} \mu\left(\mathcal{E}_{(1-\tau)\epsilon^*}\right)\mathrm{d}\tau + \int_{\tau_0}^1 \left[\int_{\mathcal{E}_{(1-\tau)\epsilon^*}} \eta^+(\mathbf{z},\alpha)\mathrm{d}\mu\left(\mathbf{z}\right)\right]\mathrm{d}\tau\right)$$

For any $\tau \in [0,1]$, the surface area of $\mathcal{E}_{(1-\tau)\epsilon^*}$ corresponds to the derivative of $\mu\left(\mathcal{E}_{(1-\tau)\epsilon^*}\right)$, such that for an infitesimal variation $d\tau$, the Lebesgue measure the hyper-ellipsoid $\mathcal{E}_{(1-\tau)\epsilon^*}$ can be computed recursively as follows:

$$\mu\left(\mathcal{E}_{(1-\tau)\epsilon^*}\right) + \frac{\partial}{\partial\tau}\mu\left(\mathcal{E}_{(1-\tau)\epsilon^*}\right)d\tau = \mu\left(\mathcal{E}_{(1-(\tau+d\tau))\epsilon^*}\right)$$

In words, the Lebesgue measure of the hyper-ellipsoid $\mu\left(\mathcal{E}_{(1-(\tau+d\tau))\epsilon^*}\right)$ is equal to the sum of the Lebesgue measure of $\mu\left(\mathcal{E}_{(1-\tau)\epsilon^*}\right)$ and of its surface area:

$$\mu\left(\mathcal{E}_{(1-\tau)\epsilon^*}\right) = \int_0^\tau \frac{\partial}{\partial t}\mu\left(\mathcal{E}_{(1-t)\epsilon^*}\right)\mathrm{d}t = \int_0^\tau \left[\int_{\partial\mathcal{E}_{(1-t)\epsilon^*}} \mathrm{d}\mu\left(\mathbf{z}\right)\right]\mathrm{d}t$$

Using a simple trick, the inner integral in **(6)** can be reformulated as follows:

$$\begin{aligned}\implies \int_{\mathcal{E}_{(1-\tau)\epsilon^*}} \eta^+(\mathbf{z},\alpha)\mathrm{d}\mu\left(\mathbf{z}\right) &= \int_0^\tau \left[\int_{\partial\mathcal{E}_{(1-t)\epsilon^*}} \eta^+(\mathbf{z},\alpha)\mathrm{d}\mu\left(\mathbf{z}\right)\right]\mathrm{d}t \\ &= \int_0^\tau \left(\frac{\partial}{\partial t}\mu\left(\mathcal{E}_{(1-t)\epsilon^*}\right) \times \left[\int_{\partial\mathcal{E}_{(1-t)\epsilon^*}} \eta^+(\mathbf{z},\alpha)\mathrm{d}\mu\left(\mathbf{z}\right)\right] \times \left[\int_{\partial\mathcal{E}_{(1-t)\epsilon^*}} \mathrm{d}\mu\left(\mathbf{z}\right)\right]^{-1}\right)\mathrm{d}t\end{aligned}$$

For a given $\tau \in [0,1]$, the above integral can be interepreted as the Lebesgue measure of the fraction of the hyper-ellipsoid $\mathcal{E}_{(1-\tau)\epsilon^*}$ lying inside the subspace Φ_α^+. However, we note that the definition of $\eta^+(\mathbf{z},\alpha)$ implies that the integral $\int_{\partial\mathcal{E}_{(1-t)\epsilon^*}} \eta^+(\mathbf{z},\alpha)d\mathbf{z}$ is positive if and only if at least half the surface of $\partial\mathcal{E}_{(1-\tau)\epsilon^*}$, as evaluated by the Lebesgue measure, belongs to the subspace Φ_α^+.

Gathering all elements developped above and plugging into **(6)** yields the following expression:

$$\mathbb{P}\left(\varphi(\mathbf{x},\alpha) \geq 0\right) = \frac{1}{2} + \frac{\epsilon^*}{2}\left(\Psi_1(\tau_0|\alpha,\epsilon^*,\boldsymbol{\mu}) + \int_{\tau_0}^1 \Psi_2(\tau|\alpha,\epsilon^*,\boldsymbol{\mu})\mathrm{d}\tau\right) \tag{7}$$

, where we use the auxiliary functions Ψ_1 and Ψ_2 defined as follows to lighten the notations:

$$\Psi_1\left(\tau_0|\alpha,\epsilon^*,\boldsymbol{\mu}\right) = \frac{\varphi(\boldsymbol{\mu},\alpha)}{|\varphi(\boldsymbol{\mu},\alpha)|}\int_0^{\tau_0} \mu\left(\mathcal{E}_{(1-\tau)\epsilon^*}\right)\mathrm{d}\tau \tag{8}$$

$$\Psi_2\left(\tau|\alpha,\epsilon^*,\boldsymbol{\mu}\right) = \int_0^\tau \left[\int_{\partial\mathcal{E}_{(1-t)\epsilon^*}} \eta^+(\mathbf{z},\alpha)\mathrm{d}\mu\left(\mathbf{z}\right)\right]\mathrm{d}t \tag{9}$$

Note that $\Psi_2(\tau|\alpha,\epsilon^*,\boldsymbol{\mu})$ can be expressed recursively through the following equation:

$$\Psi_2\left(\tau+d\tau|\alpha,\epsilon^*,\boldsymbol{\mu}\right) = \Psi_2(\tau|\alpha,\epsilon^*,\boldsymbol{\mu}) + \left[\int_{\partial\mathcal{E}_{(1-(\tau+d\tau))\epsilon^*}} \eta^+(\mathbf{z},\alpha)\mathrm{d}\mu\left(\mathbf{z}\right)\right]\mathrm{d}\tau$$

2.2.3 Extension to mixture distributions

The above approach can easily be extended to mixtures of elliptical distributions. A mixture distribution is a linear combination of component probability distributions. Let $K \in \mathbb{N}^*$. The weight given to the $k-$th component, also called mixing probability, is noted $\pi_k, 1 \leq k \leq K$, such that $0 \leq \pi_k \leq 1$. If $\sum_k \pi_k = 1$, then the function $q_{\mathbf{X}}(\mathbf{x})$ defined as follows is a valid probability density function:

$$q_{\mathbf{X}}(\mathbf{x}) = \sum_{k=1}^{K} \pi_k \times \Omega_k g\left((\mathbf{x}-\boldsymbol{\mu}_k)^T \boldsymbol{\Sigma}_k^{-1}(\mathbf{x}-\boldsymbol{\mu}_k)\right) = \sum_{k=1}^{K} \pi_k \times g_k\left(\mathbf{x}\right)$$

, where $\Omega_k \in \mathbb{R}$, $\boldsymbol{\mu}_k \in \mathbb{R}^n$ and $\boldsymbol{\Sigma}_k \in \mathbb{R}^{n\times n}$. Consistently with previous notations, we define for $\epsilon \in \mathbb{R}^+$ the hyper-ellipsoid as $\partial\mathcal{E}_{k,\epsilon} = \{\mathbf{x} \in \mathbb{R}^n \mid g_k(\mathbf{x}) = \epsilon\}$, with interior corresponding to the closed set $\mathcal{E}_{k,\epsilon} = \{\mathbf{x} \in \mathbb{R}^n \mid g_k(\mathbf{x}) \geq \epsilon\}$, for $1 \leq k \leq K$. Following equation **(7)**, the probability that $\varphi(\mathbf{x}, \alpha)$ is greater or equal to 0 when $\mathbf{X}$ follows a mixture of elliptical distributions is:

$$\mathbb{P}\left(\varphi(\mathbf{x},\alpha) \geq 0\right) = \frac{1}{2} + \sum_{k=1}^{K} \frac{\pi_k \epsilon_k^*}{2}\left(\Psi_1^k(\tau_{k,0}|\alpha, \epsilon_k^*, \boldsymbol{\mu}_k) + \int_{\tau_{k,0}}^{1} \Psi_2^k(\tau|\alpha, \epsilon_k^*, \boldsymbol{\mu}_k)\mathrm{d}\tau\right) \tag{10}$$

, where the specific parameters ϵ_k^* and $\tau_{k,0}$ are computed in the same manner to their general counterparts and the auxiliary functions Ψ_1^k and Ψ_2^k are defined as follows:

$$\Psi_1^k\left(\tau_{k,0}|\alpha, \epsilon_k^*, \boldsymbol{\mu}_k\right) = \frac{\varphi(\boldsymbol{\mu}_k, \alpha)}{|\varphi(\boldsymbol{\mu}_k, \alpha)|}\int_0^{\tau_{k,0}} \mu\left(\mathcal{E}_{k,(1-\tau)\epsilon_k^*}\right)\mathrm{d}\tau \tag{11}$$

$$\Psi_2^k(\tau|\alpha, \epsilon_k^*, \boldsymbol{\mu}_k) = \int_0^{\tau}\left[\int_{\partial\mathcal{E}_{k,(1-t)\epsilon_k^*}} \eta^+(\mathbf{z}, \alpha)\mathrm{d}\mu\left(\mathbf{z}\right)\right]\mathrm{d}t \tag{12}$$

2.3 Closed-form expression for the Lebesgue measure of $\mathcal{E}_\epsilon$

From the definition of an hyper-ellipsoid, we can derive the exact Lebesgue measure of $\mathcal{E}_\epsilon$, noted $\mu\left(\mathcal{E}_\epsilon\right)$, by recovering its canonical form equation. If the precision matrix $\boldsymbol{\Sigma}^{-1}$ is not diagonal, an additional diagonalisation step is required. Every semi-positive matrix $\boldsymbol{\Sigma}^{-1}$ can be decomposed as $\boldsymbol{\Sigma}^{-1} = \mathbf{V}\boldsymbol{\Lambda}^{-1}\mathbf{V}^T$, where the diagonal matrix $\boldsymbol{\Lambda}^{-1} = \mathrm{diag}(\lambda_1^{-1}, \cdots, \lambda_n^{-1}) \in \mathbb{R}^{n\times n}$ contains the eigenvalues of $\boldsymbol{\Sigma}^{-1}$, and $\mathbf{V} \in \mathbb{R}^{n\times n}$ is an orthogonal matrix, which satisfies $\mathbf{V}^T\mathbf{V} = 1$. The hyper-ellipsoid corresponding to $\Omega g\left(\mathbf{x}^T\boldsymbol{\Sigma}^{-1}\mathbf{x}\right) = \epsilon \Leftrightarrow \mathbf{x}^T\boldsymbol{\Sigma}^{-1}\mathbf{x} = |g^{-1}\left(\epsilon\Omega^{-1}\right)|$ can be written in canonical

form as follows:

$$\mathbf{x}^T\mathbf{\Sigma}^{-1}\mathbf{x} = \mathbf{x}^T\mathbf{V}\mathbf{\Lambda}^{-1}\mathbf{V}^T\mathbf{x} = (\mathbf{V}^T\mathbf{x})^T\mathbf{\Lambda}^{-1}(\mathbf{V}^T\mathbf{x}) = \mathbf{y}^T\mathbf{\Lambda}^{-1}\mathbf{y} = |g^{-1}\left(\epsilon\Omega^{-1}\right)|$$

, where the absolute value ensures the right-hand side of the above expression is positive. Finally, we obtain the desired expression:

$$\frac{1}{|g^{-1}\left(\epsilon\Omega^{-1}\right)|}\mathbf{y}^T\mathbf{\Lambda}^{-1}\mathbf{y} = 1 \iff \sum_{i=1}^{n}\left[\frac{y_i}{\sqrt{\lambda_i|g^{-1}\left(\epsilon\Omega^{-1}\right)|}}\right]^2 = 1$$

The volume of the hyper-ellipsoid $\mathcal{E}_\epsilon$ is then given directly given by computing the following closed-form expression:

$$\mu\left(\mathcal{E}_{(1-\tau)\epsilon^*}\right) = \left(\prod_{i=1}^{n}\lambda_i\right)^{1/2} \times \frac{\pi^{n/2}|g^{-1}\left((1-\tau)\epsilon^*\Omega^{-1}\right)|^{n/2}}{\Gamma\left(\frac{n}{2}+1\right)}$$

While the volume of $\mathcal{E}_\epsilon$ can be calculated exactly, the integral taken on the boundary $\partial\mathcal{E}_\epsilon$ in the expression for Ψ_2 must however be computed by approximation methods. It is however possible, when the derivatives of the function $h(\mathbf{x})$ satisfy specific conditions, to simplify expression **(6)**.

2.4 Specific case using the generalized Stokes theorem

The computation of Ψ_2 can be greatly simplified for a subset of functions $h(\mathbf{x})$. This requires the use of differential forms, which provide a generalization of integrands such as surfaces and volumes to higher-dimensional oriented manifolds. We refer to **[13]** for a comprehensive introduction to differential forms. For $k \in \mathbb{N}^*$, a differential k-form is a tensor of rank k that is antisymmetric under permutation of any pair of its indices. For instance in $\mathbb{R}^3$, the expression $v(x,y,z) = v_1(x,y,z)\mathrm{d}y \wedge \mathrm{d}z + v_2(x,y,z)\mathrm{d}x \wedge \mathrm{d}z + v_3(x,y,z)\mathrm{d}x \wedge \mathrm{d}y$ is a 2-form. The multiplication in the algebra of differential forms is called the exterior or wedge product $\wedge$. It is skew-symmetric, such that $\mathrm{d}x_i \wedge \mathrm{d}x_j = -\mathrm{d}x_j \wedge \mathrm{d}x_i$, which implies that $\mathrm{d}x_i \wedge \mathrm{d}x_i = 0$.

We can now introduce the differential $(n-1)$-form $\mathbf{F}_\psi(\mathbf{x},\alpha)$, which is defined as follows:

$$\begin{aligned}\mathbf{F}_\psi(\mathbf{x},\alpha) &= F_{1,\psi}(\mathbf{x},\alpha)\mathrm{d}x_2 \wedge \cdots \wedge \mathrm{d}x_n + \cdots + F_{n,\psi}(\mathbf{x},\alpha)\mathrm{d}x_1 \wedge \cdots \wedge \mathrm{d}x_{n-1} \\ &= \sum_{i=1}^{n} F_{i,\psi}(\mathbf{x},\alpha)\mathrm{d}x_1 \wedge \cdots \wedge \mathrm{d}x_{-i} \wedge \cdots \wedge \mathrm{d}x_n\end{aligned}$$

, where $\mathrm{d}x_{-i}$ signifies that the element $\mathrm{d}x_i$ is excluded from the wedge product, and $F_{i,\psi}(\mathbf{x},\alpha)\colon \mathbb{R}^n \mapsto \mathbb{R}$, $\forall i \in \{1,\cdots,n\}$. The exterior derivative of a differential form of degree k is a differential form of degree $(k+1)$ and generalizes the concept of differential for functions to differential forms. The exterior derivative of $\mathbf{F}_\psi(\mathbf{x},\alpha)$, noted $\mathrm{d}\mathbf{F}_\psi(\mathbf{x},\alpha)$, is thus a n-form. By using the alternating property of the exterior product, we can get the following expression :

$$\begin{aligned}\mathrm{d}\mathbf{F}_\psi(\mathbf{x},\alpha) &= \sum_{i=1}^{n} \mathrm{d}F_{i,\psi}(\mathbf{x},\alpha)\mathrm{d}x_1 \wedge \cdots \wedge \mathrm{d}x_{-i} \wedge \cdots \wedge \mathrm{d}x_n \\ &= \left(\frac{\partial F_{1,\psi}(\mathbf{x},\alpha)}{\partial x_1} + \cdots + \frac{\partial F_{n,\psi}(\mathbf{x},\alpha)}{\partial x_n}\right)\mathrm{d}x_1 \wedge \cdots \wedge \mathrm{d}x_n \\ &= \nabla \cdot \mathbf{F}_\psi\end{aligned}$$

As by definition $\eta^+(\mathbf{x},\alpha)$ cannot have continuous derivatives over its whole domain, we will approximate it by means of a smooth function $\eta^+_\psi(\mathbf{x},\alpha)$, where $\psi \in \mathbb{R}^+$, such that $\lim_{\psi\to+\infty}\eta^+_\psi(\mathbf{x},\alpha) = \eta^+(\mathbf{x},\alpha)$. Then, assuming there exists $\mathbf{F}_\psi(\mathbf{x},\alpha)$ that verifies $(\nabla\cdot\mathbf{F}_\psi)(\mathbf{x},\alpha) = \eta^+_\psi(\mathbf{x},\alpha)$ and since the hyperellipsoid is a smooth manifold, it follows from the generalized Stokes theorem that:

$$\int_{\partial\mathcal{E}_\epsilon} \mathbf{F}_\psi = \int_{\mathcal{E}_\epsilon} \mathrm{d}\mathbf{F}_\psi = \int_{\mathcal{E}_\epsilon} \eta^+_\psi \tag{13}$$

The expression in **(10)** can then be reformulated as follows:

$$\mathbb{P}\left(\varphi(\mathbf{x},\alpha) \geq 0\right) = \frac{1}{2} + \frac{\epsilon^*}{2}\left(\Psi_1(\tau|\tau_0,\alpha,\epsilon^*,\boldsymbol{\mu}) + \int_{\tau_0}^{1} \Psi_3(\tau|\alpha,\psi,\epsilon^*)\mathrm{d}\tau\right) \tag{14}$$

, where:

$$\Psi_3(\tau|\alpha,\psi,\epsilon^*) = \int_{\partial\mathcal{E}_{(1-\tau)\epsilon^*}} \mathbf{F}_\psi \tag{15}$$

For any $\alpha \in \mathbb{R}$, let us consider the couple of real-valued functions $\eta^+_\psi(\cdot,\alpha) : \mathbb{R}^n \mapsto \mathbb{R}$ and $E_\psi(\cdot,\alpha) : \mathbb{R}^n \mapsto \mathbb{R}$, such that $\eta^+_\psi(\mathbf{x},\alpha) = \tanh\left(\psi \times \varphi(\mathbf{x},\alpha)\right)$ and $E_\psi(\mathbf{x},\alpha) = \log\circ\cosh\left(\psi\times\varphi(\mathbf{x},\alpha)\right)$.

Our definition of $\eta_\psi(\cdot,\alpha)$ satisfies both the necessary conditions required for the application of the generalized Stokes theorem and $\lim_{\psi\to+\infty}\eta^+_\psi(\mathbf{x},\alpha) = \eta^+(\mathbf{x},\alpha)$. Then, we can easily prove that if the function $h(\mathbf{x})$ is a multilinear polynomial, there always at least one differential $(n-1)$-form $\mathbf{F}_\psi(\mathbf{x},\alpha)$ such that the necessary conditions for the application of the generalized Stokes theorem

are satisfied, with components defined as follows:

$$F_{i,\psi}(\mathbf{x},\alpha) = \left[\beta_i \left(\psi \times \frac{\partial\varphi(\mathbf{x},\alpha)}{\partial x_i}\right)^{-1} E_\psi(\mathbf{x},\alpha)\right] \mathrm{d}x_1 \wedge \cdots \wedge \mathrm{d}x_{-i} \wedge \cdots \wedge \mathrm{d}x_n$$

, with $\beta_i \in \mathbb{R}$ and $\sum_{i=1}^n \beta_i = 1$,

Proof: By definition of a multilinear polynomial, all repeated second derivatives of $h(\mathbf{x})$ are null, so by linearity we have the following implication: $\frac{\partial^2 h(\mathbf{x})}{\partial x_i^2} = 0 \Rightarrow \frac{\partial^2 \varphi(\mathbf{x},\alpha)}{\partial x_i^2} = 0$. Then, using the definition of $F_{i,\psi}(\mathbf{x},\alpha)$, we can develop its partial derivative with respect to x_i (we ommit the $\mathrm{d}x_1 \wedge \cdots \wedge \mathrm{d}x_{-i} \wedge \cdots \wedge \mathrm{d}x_n$ term to make the proof more readable):

$$\begin{aligned}
\frac{\partial F_{i,\psi}(\mathbf{x},\alpha)}{\partial x_i} &= \beta_i \left[\psi^{-1} \times \frac{\partial}{\partial x_i}\left(\frac{\partial\varphi(\mathbf{x},\alpha)}{\partial x_i}\right)^{-1} E_\psi(\mathbf{x},\alpha) + \left(\psi \times \frac{\partial\varphi(\mathbf{x},\alpha)}{\partial x_i}\right)^{-1} \frac{\partial E_\psi(\mathbf{x},\alpha)}{\partial x_i}\right] \\
&= \beta_i \left(\psi \times \frac{\partial\varphi(\mathbf{x},\alpha)}{\partial x_i}\right)^{-1} \times \frac{\partial E_\psi(\mathbf{x},\alpha)}{\partial x_i} \\
&= \beta_i \left(\psi \times \frac{\partial\varphi(\mathbf{x},\alpha)}{\partial x_i}\right)^{-1} \times \left(\psi \times \frac{\partial\varphi(\mathbf{x},\alpha)}{\partial x_i}\right) \times \eta_\psi(\mathbf{x},\alpha) \\
&= \beta_i \times \eta_\psi(\mathbf{x},\alpha)
\end{aligned}$$

It follows that:

$$(\nabla \cdot \mathbf{F}_\psi)(\mathbf{x},\alpha) = \sum_{i=1}^n \frac{\partial F_{i,\psi}(\mathbf{x},\alpha)}{\partial x_i} = \sum_{i=1}^n \beta_i \times \eta_\psi(\mathbf{x},\alpha) = \eta_\psi(\mathbf{x},\alpha)$$

It immediately follows from the above result that, for any real-valued function $h(\mathbf{x})$, if there exists at least one $i^* \in \{1, \cdots, n\}$ such that $\frac{\partial^2 h(\mathbf{x})}{\partial x_{i^*}^2} = 0$, the differential $(n-1)$-form $\mathbf{F}_\psi(\mathbf{x},\alpha)$ with components defined as $F_{i^*,\psi}(\mathbf{x},\alpha) = \left[\left(\psi \times \frac{\partial\varphi(\mathbf{x},\alpha)}{\partial x_{i^*}}\right)^{-1} E_\psi(\mathbf{x},\alpha)\right] \mathrm{d}x_1 \wedge \cdots \wedge \mathrm{d}x_{-i^*} \wedge \cdots \wedge \mathrm{d}x_n$ and $F_{i,\psi}(\mathbf{x},\alpha) = 0$ for $i \neq i^*$ satisfies the necessary conditions for the application of **(12)**.

3 Computation strategy

For each $\epsilon > 0$, we propose to approximate the integral in Ψ_2 by uniformly drawing points on the boundary $\partial\mathcal{E}_\epsilon$. Indeed, the integral of the function $\eta^+(\mathbf{x}, \alpha)$ on $\partial\mathcal{E}_\epsilon$ can be interpreted as the expectation of the random variable $\eta^+(\mathbf{U}, \alpha)$, where $\mathbf{U}$ is a random vector with uniform distribution on the unit sphere.

3.1 Approximation via uniform sampling on the unit sphere

Before going further, let us first introduce a few additional notations. Let the positive scalar $\|\lambda\|_\infty$ correspond to the largest element of the eigenvalue matrix $\mathbf{\Lambda}$. For $\epsilon > 0$, the length of the principal semi-axis of the hyper-ellipsoid $\partial\mathcal{E}_\epsilon$ is then equal to $\sqrt{\|\lambda\|_\infty |g^{-1}\left(\epsilon\Omega^{-1}\right)|}$. We finally define the unit sphere in $\mathbb{R}^n$ centered on zero as $\mathbb{S}(0,1) = \{\mathbf{x} \in \mathbb{R}^n | \|\mathbf{x}\|_2^2 = 1\}$.

Let $\mathbf{x} \in \mathbb{S}(0,1)$. Then, we have by definition:

$$\begin{aligned}\mathbf{x}^T\mathbf{x} = 1 &\Longrightarrow \left(\mathbf{\Sigma}^{1/2}\mathbf{x}\right)^T \mathbf{\Sigma}^{-1}\left(\mathbf{\Sigma}^{1/2}\mathbf{x}\right) = 1 \\ &\Longrightarrow \left(\mathbf{z}_\mathcal{E}(\mathbf{x},\epsilon) - \boldsymbol{\mu}\right)^T \mathbf{\Sigma}^{-1}\left(\mathbf{z}_\mathcal{E}(\mathbf{x},\epsilon) - \boldsymbol{\mu}\right) = |g^{-1}\left(\epsilon\Omega^{-1}\right)|\end{aligned}$$

, where $\mathbf{z}_{\partial\mathcal{E}}(\mathbf{x},\epsilon) = \sqrt{\|\lambda\|_\infty |g^{-1}\left(\epsilon\Omega^{-1}\right)|}\mathbf{\Sigma}^{1/2}\mathbf{x} + \boldsymbol{\mu}$. Thus, $\mathbf{x} \in \mathbb{S}(0,1) \Longrightarrow \mathbf{z}_{\partial\mathcal{E}}(\mathbf{x},\epsilon) \in \partial\mathcal{E}_\epsilon$. For $\epsilon > 0$ and because the variance-covariance matrix $\mathbf{\Sigma}$ is invertible, the linear transformation $\mathbf{z}_{\partial\mathcal{E}}(\cdot,\epsilon)\colon \mathbb{R}^n \mapsto \mathbb{R}^n$ admits an inverse $\mathbf{z}_{\partial\mathcal{E}}^{-1}(\cdot,\epsilon)\colon \mathbb{R}^n \mapsto \mathbb{R}^n$, defined as follows:

$$\mathbf{z}_{\partial\mathcal{E}}^{-1}(\mathbf{x},\epsilon) = \frac{1}{\sqrt{\|\lambda\|_\infty |g^{-1}\left(\epsilon\Omega^{-1}\right)|}}\mathbf{\Sigma}^{-1/2}\left(\mathbf{x} - \boldsymbol{\mu}\right)$$

Using tools from measure theory, we can use the following result to reformulate the integral in expression **(9)** as a simpler integral on the unit sphere $\mathbb{S}(0,1)$.

Proof: See Appendix.

Proposition 2: Let $(\mathbb{S}, \mathcal{S}_\mathbb{S})$ be a measure space defined on the surface of the unit sphere $\mathbb{S} \doteq \mathbb{S}(0,1)$, where $\mathcal{S}_\mathbb{S}$ is the $\sigma-$ algebra of subsets of $\mathbb{S}(0,1)$. It is equipped with the Lebesgue measure $\mu_{\mathcal{S}_\mathbb{S}}$ on $\mathcal{S}_\mathbb{S}$. We further define the measure space $(\partial\mathcal{E}_\epsilon, \mathcal{S}_{\partial\mathcal{E}_\epsilon})$ corresponding to the hyper-ellipsoid $\partial\mathcal{E}_\epsilon$, equipped with the measure $\mu_{\partial\mathcal{E}_\epsilon}$. Finally, for $\epsilon > 0$, let $\phi\colon \mathbb{S}(0,1) \mapsto \partial\mathcal{E}_\epsilon$ be a bijective map such that its inverse ϕ^{-1} exists.

Then, we have the following result:

$$\left(\int_{\partial\mathcal{E}_\epsilon} \eta^+ \mathrm{d}\mu_{\mathcal{E}_\epsilon}\right)\left(\int_{\partial\mathcal{E}_\epsilon} \mathrm{d}\mu_{\mathcal{E}_\epsilon}\right)^{-1} = \left(\int_{\mathbb{S}(0,1)} \left(\eta^+ \circ \phi\right) \mathrm{d}\mu_{\mathcal{S}_\mathbb{S}}\right)\left(\int_{\mathbb{S}(0,1)} \mathrm{d}\mu_{\mathcal{S}_\mathbb{S}}\right)^{-1}$$

Proof:

Given measure spaces $(X, \mathcal{M}_X)$ and $(Y, \mathcal{M}_Y)$, a measurable function $\phi\colon X \mapsto Y$ and a measure function $\mu\colon \mathcal{M}_X \mapsto [0, +\infty[$, the pushforward of measure μ by ϕ is defined as the measure $\phi_*\mu = \mu(\phi^{-1}(A'))$, $\forall A' \in \mathcal{M}_Y$. If ϕ is bijective, then there exists a measure function $\omega\colon \mathcal{M}_Y \mapsto [0, +\infty[$ associated to the measurable space $(Y, \mathcal{M}_Y)$ such that $\omega = \phi^*\mu = \left(\phi^{-1}\right)_*\mu$. The measure ω is defined as the pullback measure and corresponds to the adjoint of the pushfoward measure $\phi_*\mu$.

Let $\phi \triangleq \mathbf{z}_{\partial\mathcal{E}}(\cdot, \epsilon)$, such that we have $\mathcal{E}_\epsilon = \phi\left(\mathbb{S}(0,1)\right)$. Using the above notations and noting $\mu \triangleq \mu_{\mathcal{E}_\epsilon}$, we obtain the following series of equalities:

$$\begin{aligned}
\int_{\partial\mathcal{E}_\epsilon} \eta^+ \mathrm{d}\mu &= \int_{\phi(\mathbb{S}(0,1))} \eta^+ \mathrm{d}\mu \\
&= \int_{\mathbb{S}(0,1)} \left(\eta^+ \circ \phi\right) \mathrm{d}\left(\phi^*\mu\right) \\
&= \int_{\mathbb{S}(0,1)} \left(\eta^+ \circ \phi\right) |\det \mathbf{J}_\phi| \mathrm{d}\mu \\
&= \left|\det\left(\sqrt{\|\lambda\|_\infty |g^{-1}\left(\epsilon\Omega^{-1}\right)|}\boldsymbol{\Sigma}^{1/2}\right)\right| \times \int_{\mathbb{S}(0,1)} \left(\eta^+ \circ \phi\right) \mathrm{d}\mu \\
&= \|\lambda\|_\infty^{n/2} \times |g^{-1}\left((1-\tau)\epsilon^*\Omega^{-1}\right)|^{n/2} \times |\det \mathbf{V}\boldsymbol{\Lambda}^{1/2}| \times \int_{\mathbb{S}(0,1)} \left(\eta^+ \circ \phi\right) \mathrm{d}\mu \\
&= \|\lambda\|_\infty^{n/2} \times |g^{-1}\left((1-\tau)\epsilon^*\Omega^{-1}\right)|^{n/2} \times \det \boldsymbol{\Lambda}^{1/2} \times \int_{\mathbb{S}(0,1)} \left(\eta^+ \circ \phi\right) \mathrm{d}\mu
\end{aligned}$$

, where we use the property that the determinant of an orthogonal matrix is either equal to 1 or -1 and the fact that $\det \mathbf{V}\boldsymbol{\Lambda}^{1/2} = \det \mathbf{V} \times \det \boldsymbol{\Lambda}^{1/2}$. The matrix $\mathbf{J}_\phi$ corresponds to the Jacobian of the function ϕ. Similarly, we have:

$$\begin{aligned}
\int_{\partial\mathcal{E}_\epsilon} \mathrm{d}\mu &= \int_{\mathbb{S}(0,1)} \mathrm{d}\left(\phi^*\mu\right) \\
&= \int_{\mathbb{S}(0,1)} |\det \mathbf{J}_\phi| \mathrm{d}\mu \\
&= \|\lambda\|_\infty^{n/2} \times |g^{-1}\left((1-\tau)\epsilon^*\Omega^{-1}\right)|^{n/2} \times \det \boldsymbol{\Lambda}^{1/2} \times \int_{\mathbb{S}(0,1)} \mathrm{d}\mu
\end{aligned}$$

Gathering the above elements, we obtain the desired result.

For a given $\alpha \in \mathbb{R}$ and $\epsilon > 0$, we use the shorthand $\eta^+ \circ \mathbf{z}_{\partial\mathcal{E}} \triangleq \eta^+ \left(\mathbf{z}_{\partial\mathcal{E}}(\mathbf{x}, \epsilon), \alpha\right)$ in the absence of ambiguity in notations. It is then possible to rewrite expression **(9)** uniquely in terms of integrals computed on the unit sphere $\mathbb{S}(0,1)$:

$$\Psi_2\left(\tau|\alpha, \epsilon^*, \boldsymbol{\mu}\right) = \int_0^\tau \left(\frac{\partial}{\partial t} \mu\left(\mathcal{E}_{(1-t)\epsilon^*}\right) \times \left[\int_{\mathbb{S}(0,1)} \left(\eta^+ \circ \mathbf{z}_{\partial\mathcal{E}}\right) \mathrm{d}\mu_{\mathcal{S}_\mathbb{S}} \right] \times \left[\int_{\mathbb{S}(0,1)} \mathrm{d}\mu_{\mathcal{S}_\mathbb{S}} \right]^{-1} \right) \mathrm{d}t \tag{16}$$

We can finally formulate $\Psi_2\left(\tau|\tau_0, \alpha, \epsilon^*, \boldsymbol{\mu}\right)$ as the expected value of the function $\eta^+\left(\mathbf{z}_{\partial\mathcal{E}}(\mathbf{U}, \epsilon), \alpha\right)$, where the random vector $\mathbf{U}$ follows a uniform distribution on the unit sphere $\mathbb{S}(0,1)$. Given the measurable space $(\mathbb{S}, \mathcal{S}_\mathbb{S}, \mu_{\mathcal{S}_\mathbb{S}})$ defined on the surface of the unit sphere $\mathbb{S}(0,1)$, the random vector $\mathbf{U}$ follows a uniform distribution on $\mathbb{S}(0,1)$ with respect to $\mu_{\mathcal{S}_\mathbb{S}}$ if for any subset $A \in \mathcal{S}_\mathbb{S}$, we have $\mathbb{P}(\mathbf{U} \in A) = \mu_{\mathcal{S}_\mathbb{S}}(A) \times \mu_{\mathcal{S}_\mathbb{S}}(\mathbb{S}(0,1))^{-1}$. It follows that the probability density function p of $\mathbf{U}$ with respect to $\mu_{\mathcal{S}_\mathbb{S}}$ is:

$$p(\mathbf{u}) = \left(\int_{\mathbb{S}(0,1)} \mathrm{d}\mu_{\mathcal{S}_\mathbb{S}}(\mathbf{u}') \right)^{-1} \tag{17}$$

, where $\mathbf{u} \in \mathbb{S}(0,1)$. Gathering all elements, **(9)** can finally be expressed as a function of the expectation of $\eta^+ \circ \mathbf{z}_{\partial\mathcal{E}}$ over the unit sphere $\mathbb{S}(0,1)$:

$$\begin{aligned}
\mathbb{E}\left[\eta^+\left(\mathbf{z}_{\mathcal{E}}(\mathbf{U}, \epsilon)\right)\right] &= \int_{\mathbb{S}(0,1)} \left[\eta^+(\mathbf{z}_{\mathcal{E}}(\mathbf{u}, \epsilon), \alpha) p(\mathbf{u}')\right] \mathrm{d}\mu_{\mathcal{S}_\mathbb{S}}(\mathbf{u}') \\
&= \int_{\mathbb{S}(0,1)} \eta^+(\mathbf{z}_{\mathcal{E}}(\mathbf{u}, \epsilon), \alpha) \mathrm{d}\mu_{\mathcal{S}_\mathbb{S}}(\mathbf{u}') \times \left(\int_{\mathbb{S}(0,1)} \mathrm{d}\mu_{\mathcal{S}_\mathbb{S}}(\mathbf{u}') \right)^{-1}
\end{aligned}$$

$$\begin{aligned}
\implies \Psi_2\left(\tau|\tau_0, \alpha, \epsilon^*, \boldsymbol{\mu}\right) &= \int_0^\tau \left(\frac{\partial}{\partial t} \mu\left(\mathcal{E}_{(1-t)\epsilon^*}\right) \times \mathbb{E}\left[\eta^+\left(\mathbf{z}_{\mathcal{E}}\left(\mathbf{U}, (1-t)\epsilon^*\right), \alpha\right)\right] \right) \mathrm{d}t \\
&= \mathbb{E}\left[\int_0^\tau \left(\frac{\partial}{\partial t} \mu\left(\mathcal{E}_{(1-t)\epsilon^*}\right) \times \eta^+\left(\mathbf{z}_{\mathcal{E}}\left(\mathbf{U}, (1-t)\epsilon^*\right), \alpha\right) \right) \mathrm{d}t \right]
\end{aligned}$$

3.2 Sampling method

The following subsection proposes an uniform sampling method to approximate the inner integral in expression **(16)**. We use the fact that due to the rotational invariance of the spherical Gaussian distribution $\mathcal{N}(0, \mathbf{I}_n)$, where $\mathbf{I}_n \in \mathbb{R}^{n\times n}$ is the identity matrix, the vector $\frac{\mathbf{x}}{\|\mathbf{x}\|_2}$ is uniformly distributed on the unit sphere, $\mathbf{x} \sim \mathcal{N}(0, \mathbf{I}_n)$.

We may further exploit the multiple symmetries of the hyperellipsoid by defining the vector of ones $\mathbf{J}_n = (1, \cdots, 1) \in \mathbb{R}^{1\times n}$ and the vector $\mathbf{L} = (1, -1) \in \mathbb{R}^{1\times 2}$. We introduce the matrix $\mathbf{W} \in \mathbb{R}^{n\times 2^n}$, with i-th row $\mathbf{W}_{i,\cdot}$ defined as follows:

$$\mathbf{W}_{i,\cdot} = \mathbf{J}_{2^{n-i}} \otimes \mathbf{L} \otimes \mathbf{J}_{2^{i-1}} \tag{18}$$

, where $\otimes$ is the Kronecker product. We observe that $\forall j \in \{1, \cdots, 2^n\}$, if $\mathbf{x} \in \mathbb{S}(0, 1)$, it follows that $\mathbf{z}_{\mathcal{E}}\left(\mathbf{x} \circ \mathbf{W}_{\cdot,j}, \epsilon\right)^T \mathbf{\Sigma}^{-1} \mathbf{z}_{\mathcal{E}}\left(\mathbf{x} \circ \mathbf{W}_{\cdot,j}, \epsilon\right) = \mathbf{z}_{\mathcal{E}}\left(\mathbf{x}, \epsilon\right)^T \mathbf{\Sigma}^{-1} \mathbf{z}_{\mathcal{E}}\left(\mathbf{x}, \epsilon\right) = |g^{-1}\left(\epsilon\Omega^{-1}\right)|$, so $\mathbf{z}_{\mathcal{E}}\left(\mathbf{x} \circ \mathbf{W}_{\cdot,j}, \epsilon\right) \in \partial\mathcal{E}_\epsilon$. From a geometrical perspective, for each $\forall j \in \{1, \cdots, 2^n\}$, the vector $\mathbf{z}_{\mathcal{E}}\left(\mathbf{x} \circ \mathbf{W}_{\cdot,j}, \epsilon\right)$ corresponds to a given (non-unique) composition of reflections through the hyperplanes respectively spanned by the eigenvectors of $\mathbf{\Sigma}^{-1}$. For instance, for $n = 2$, the vector $\mathbf{z}_{\mathcal{E}}\left(\mathbf{x} \circ \mathbf{W}_{\cdot,2}, \epsilon\right)$ (resp. $\mathbf{z}_{\mathcal{E}}\left(\mathbf{x} \circ \mathbf{W}_{\cdot,3}, \epsilon\right)$) corresponds to the reflection of $\mathbf{z}_{\mathcal{E}}\left(\mathbf{x}, \epsilon\right)$ through the hyperplane spanned by the eigenvector associated to λ_1 (resp. λ_2). Thus, it would only be required to sample $\tilde{N} \in \mathbb{N}$ points on the unit sphere to obtain a sample of size $N = 2^n \times \tilde{N}$.

3.3 Estimators formulation

Let $(\mathbf{u}_1, \cdots, \mathbf{u}_N)$ be a sample of size N of iid random vectors following a uniform distribution on the unit sphere, i.e. $\forall i \in \{1, \cdots, N\}$, $\mathbf{u}_i$ is drawn from $\mathbf{U}$, where $\mathbf{U} \sim \mathcal{U}(\mathbb{S}(0, 1))$. For $\tau \in [0, 1]$, we define the following estimator to approximate the inner integral in expression **(9)**:

$$T\left(\tau|\mathbf{u}_1, \cdots, \mathbf{u}_N, \epsilon^*, \alpha\right) = \frac{1}{N} \times \mu\left(\partial\mathcal{E}_{(1-\tau)\epsilon^*}\right) \times \sum_{i=1}^{N} \eta^+(\mathbf{z}_{\mathcal{E}}(\mathbf{u}_i, (1-\tau)\epsilon^*))$$

Proposition 3: The estimator $T_N(\tau|\epsilon^*, \alpha) \triangleq T\left(\tau|\mathbf{u}_1, \cdots, \mathbf{u}_N, \epsilon^*, \alpha\right)$ is consistent in probability, that is:

$$\forall \varepsilon > 0, \lim_{N\to+\infty} \mathbb{P}\left(|T_N(\tau|\epsilon^*, \alpha) - \mathbb{E}\left[\eta^+(\mathbf{z}_{\mathcal{E}}(\mathbf{U}, \epsilon), \alpha)\right]| > \varepsilon\right) = 0$$

Proof:

We start by showing that $T_N(\tau|\epsilon^*,\alpha)$ is an unbiased estimator of **(9)**. As $\mathbf{u}_1,\cdots,\mathbf{u}_N$ are independent and identically distributed random variables, it follows from the weak Law of Large Numbers that:

$$\frac{1}{N}\times\sum_{i=1}^{N}\eta^+(\mathbf{z}_{\mathcal{E}}(\mathbf{u}_i,(1-\tau)\epsilon^*))\xrightarrow{\mathbb{P}}\mathbb{E}\left[\eta^+(\mathbf{z}_{\mathcal{E}}(\mathbf{U},\epsilon),\alpha)\right]$$

Proving unbiasedness requires in addition proving that the variance of the estimator is uniformly bounded. First, it follows from the definition of η^+ that $-1\leq\mathbb{E}\left[\eta^+(\mathbf{z}_{\mathcal{E}}(\mathbf{U},\epsilon),\alpha)\right]\leq 1\Longrightarrow 0\leq\mathbb{E}\left[\eta^+(\mathbf{z}_{\mathcal{E}}(\mathbf{U},\epsilon),\alpha)\right]^2\leq 1$. Similarly, it is straighforward to show that $\mathbb{E}\left[\left(\eta^+(\mathbf{z}_{\mathcal{E}}(\mathbf{U},\epsilon),\alpha)\right)^2\right]=1$. For clarity and without loss of generality, let us assume that $\mu\left(\partial\mathcal{E}_{(1-\tau)\epsilon^*}\right)=1$. Then, we can decompose the variance of our estimator as follows:

$$\begin{aligned}
\mathbb{V}\left(T_N(\tau|\epsilon^*,\alpha)\right)&=\mathbb{V}\left(\frac{1}{N}\times\sum_{i=1}^{N}\eta^+(\mathbf{z}_{\mathcal{E}}(\mathbf{u}_i,(1-\tau)\epsilon^*))\right)\\
&=\frac{1}{N^2}\times\sum_{i=1}^{N}\mathbb{V}\left(\eta^+(\mathbf{z}_{\mathcal{E}}(\mathbf{u}_i,(1-\tau)\epsilon^*))\right)\\
&=\frac{1}{N}\times\mathbb{V}\left(\eta^+(\mathbf{z}_{\mathcal{E}}(\mathbf{u}_1,(1-\tau)\epsilon^*))\right)\\
&=\frac{1}{N}\times\left(\mathbb{E}\left[\left(\eta^+(\mathbf{z}_{\mathcal{E}}(\mathbf{u}_1,\epsilon),\alpha)\right)^2\right]-\mathbb{E}\left[\eta^+(\mathbf{z}_{\mathcal{E}}(\mathbf{u}_1,\epsilon),\alpha)\right]^2\right)\\
&\leq\frac{1}{N}
\end{aligned}$$

It follows that for any $N>0$, $\mathbb{V}\left(T_N(\tau|\epsilon^*,\alpha)\right)\leq 1$ so our estimator is unbiased. Furthermore, as $\lim_{N\to+\infty}\mathbb{V}\left(T_N(\tau|\epsilon^*,\alpha)\right)=0$, our estimator is mean-square consistent, which finally implies consistency in probability.

Let $J\in\mathbb{N},J>0$ and define $\Delta\tau=J^{-1}$. We further note τ_0^J (we assume $\tau_0^J<J$ to avoid trivial results) as the largest integer such that $\forall\mathbf{x}\in\mathcal{E}_{(1-\tau_0^J)\epsilon^*}$, we have $\eta^+(\mathbf{x},\alpha)=1$ if $\varphi(\boldsymbol{\mu},\alpha)>0$ (resp. $\eta^+(\mathbf{x},\alpha)=-1$ if $\varphi(\boldsymbol{\mu},\alpha)<0$). Then, the auxiliary function Ψ_1 defined in **(8)** can be approximated by the following function:

$$\Psi_1^J(\tau_0^J|\alpha,\epsilon^*,\boldsymbol{\mu}) = \frac{\varphi(\boldsymbol{\mu},\alpha)}{|\varphi(\boldsymbol{\mu},\alpha)|} \times \sum_{j=1}^{\tau_0^J} \mu\left(\mathcal{E}_{(1-j\Delta\tau)\epsilon^*}\right)\Delta\tau$$
$$= \frac{1}{J} \times \frac{\varphi(\boldsymbol{\mu},\alpha)}{|\varphi(\boldsymbol{\mu},\alpha)|} \times \frac{\det\Lambda^{1/2}\times\pi^{n/2}}{\Gamma\left(\frac{n}{2}+1\right)} \times \sum_{j=1}^{\tau_0^J} |g^{-1}\left((1-\tau)\epsilon^*\Omega^{-1}\right)|^{n/2}$$

In a similar fashion, the auxiliary function Ψ_2 defined in equation **(16)** is approximated as follows:

$$\Psi_2^J(j|N,\alpha,\epsilon^*,\boldsymbol{\mu}) = \sum_{j'=1}^{j} T\left(j'\Delta\tau|\mathbf{u}_1,\cdots,\mathbf{u}_N,\epsilon^*,\alpha\right)\Delta\tau$$
$$= \frac{1}{J}\sum_{j'=1}^{j} T\left(j'\Delta\tau|\mathbf{u}_1,\cdots,\mathbf{u}_N,\epsilon^*,\alpha\right)$$

For $j < J$, the expression $\mu\left(\partial\mathcal{E}_{(1-\tau)\epsilon^*}\right)$ is approximated as $\left(\mu\left(\mathcal{E}_{(1-(j+1)\Delta\tau)\epsilon^*}\right) - \mu\left(\mathcal{E}_{(1-j\Delta\tau)\epsilon^*}\right)\right)\Delta\tau^{-1}$. Finally, the expression for the probability of the event $(\varphi(\mathbf{x},\alpha)\geq 0)$ in **(10)** is approximated using the following expression:

$$\mathbb{P}_{N,J}\left(\varphi(\mathbf{x},\alpha)\geq 0\right) = \frac{1}{2} + \frac{\epsilon^*}{2}\left(\Psi_1^J(\tau_0^J|\alpha,\epsilon^*,\boldsymbol{\mu}) + \frac{1}{J}\sum_{j=1}^{J}\Psi_2^J(j|N,\alpha,\epsilon^*,\boldsymbol{\mu})\right) \tag{19}$$

, where we have $\lim_{N,J\to\infty}\mathbb{P}_{N,J}\left(\varphi(\mathbf{x},\alpha)\geq 0\right) = \mathbb{P}\left(\varphi(\mathbf{x},\alpha)\geq 0\right)$.

However, as the density function of the random vector $\mathbf{X}$ takes strictly positive values for any vector in $\mathbb{R}^n$, it implies that $\lim_{\epsilon\to 0}\mu\left(\mathcal{E}_\epsilon\right) = +\infty \Longrightarrow \lim_{\tau\to 1} T\left(\tau|\mathbf{u}_1,\cdots,\mathbf{u}_N,\epsilon^*,\alpha\right) = +\infty$. Thus, the above expression is numerically undefined and needs to be expressed under a suitable form to allow numerical calculations. First, we note that $\mathbb{P}\left(\varphi(\mathbf{x},\alpha)\geq 0\right)$ can be rewritten under the following form:

$$\mathbb{P}\left(\varphi(\mathbf{x},\alpha)\geq 0\right) = \frac{1}{2} + \frac{\epsilon^*}{2}\left(\Psi_1^J(\tau_0^J|\alpha,\epsilon^*,\boldsymbol{\mu}) + \frac{1}{J}\sum_{j=1}^{J-1}\Psi_2^J(j|N,\alpha,\epsilon^*,\boldsymbol{\mu}) + \int_{(J-1)\Delta\tau}^{1}\Psi_2(\tau|\alpha,\epsilon^*,\boldsymbol{\mu})\mathrm{d}\tau\right) + \varepsilon_{\alpha,N,J}$$

, where $\varepsilon_{\alpha,N,J}$ is a random variable corresponding to the approximation error, with $\mathbb{E}\left[\varepsilon_{\alpha,N,J}\right] = 0$ because of the unbiasedness of our estimator, and $\lim_{N,J\to+\infty}\mathbb{V}\left[\varepsilon_{\alpha,N,J}\right] = 0$. While the random term

vanishes by taking the expectation, the integral term is deterministic and thus introduces a systemic error if not properly accounted for. Recalling the definitions of η^+ and Ψ_2, we derive the following double inequality:

$$-\int_0^\tau \left[\int_{\partial\mathcal{E}_{(1-t)\epsilon^*}} \mathrm{d}\mu\left(\mathbf{z}\right)\right]\mathrm{d}t \leq \Psi_2\left(\tau|\alpha,\epsilon^*,\boldsymbol{\mu}\right) \leq \int_0^\tau \left[\int_{\partial\mathcal{E}_{(1-t)\epsilon^*}} \mathrm{d}\mu\left(\mathbf{z}\right)\right]\mathrm{d}t$$

$$\implies -\int_{(J-1)\Delta\tau}^1 \mu\left(\mathcal{E}_{(1-\tau)\epsilon^*}\right)\mathrm{d}\tau \leq \int_{(J-1)\Delta\tau}^1 \Psi_2\left(\tau|\alpha,\epsilon^*,\boldsymbol{\mu}\right)\mathrm{d}\tau \leq \int_{(J-1)\Delta\tau}^1 \mu\left(\mathcal{E}_{(1-\tau)\epsilon^*}\right)\mathrm{d}\tau$$

$$\implies -\left(1-\epsilon^*\int_0^{(J-1)\Delta\tau} \mu\left(\mathcal{E}_{(1-\tau)\epsilon^*}\right)\mathrm{d}\tau\right) \leq \epsilon^*\int_{(J-1)\Delta\tau}^1 \Psi_2\left(\tau|\alpha,\epsilon^*,\boldsymbol{\mu}\right)\mathrm{d}\tau \leq 1-\epsilon^*\int_0^{(J-1)\Delta\tau} \mu\left(\mathcal{E}_{(1-\tau)\epsilon^*}\right)\mathrm{d}\tau$$

The term of the left-hand and right-hand sides of the last inequality corresponds to the remaining probability mass that is not captured when summing $\Psi_2^J(j|N,\alpha,\epsilon^*,\boldsymbol{\mu})$ for $j \in \{1,\cdots,J-1\}$. As all terms are finite, there exists a constant $\zeta \in \mathbb{R}$, $-1 \leq \zeta \leq 1$, such that:

$$\epsilon^*\int_{(J-1)\Delta\tau}^1 \Psi_2\left(\tau|\alpha,\epsilon^*,\boldsymbol{\mu}\right)\mathrm{d}\tau = \zeta\left(1-\epsilon^*\int_0^{(J-1)\Delta\tau} \mu\left(\mathcal{E}_{(1-\tau)\epsilon^*}\right)\mathrm{d}\tau\right)$$

$$\implies \mathbb{P}\left(\varphi(\mathbf{x},\alpha)\geq 0\right) = \frac{1}{2}+\frac{\epsilon^*}{2}\left(\Psi_1^J(\tau_0^J|\alpha,\epsilon^*,\boldsymbol{\mu})+\frac{1}{J}\sum_{j=1}^{J-1}\Psi_2^J(j|N,\alpha,\epsilon^*,\boldsymbol{\mu})\right)+\frac{\zeta}{2}\left(1-\epsilon^*\int_0^{(J-1)\Delta\tau} \mu\left(\mathcal{E}_{(1-\tau)\epsilon^*}\right)\mathrm{d}\tau\right)+\varepsilon_{\alpha,N,J}$$

Finally, we can reformulate **(19)** into the following well-defined expression:

$$\mathbb{P}_{N,J}\left(\varphi(\mathbf{x},\alpha)\geq 0\right) = \frac{1}{2}+\frac{\epsilon^*}{2}\left(\Psi_1^J(\tau_0^J|\alpha,\epsilon^*,\boldsymbol{\mu})+\frac{1}{J}\sum_{j=1}^{J-1}\Psi_2^J(j|N,\alpha,\epsilon^*,\boldsymbol{\mu})\right)+\frac{\zeta}{2}\left(1-\epsilon^*\int_0^{(J-1)\Delta\tau} \mu\left(\mathcal{E}_{(1-\tau)\epsilon^*}\right)\mathrm{d}\tau\right) \tag{20}$$

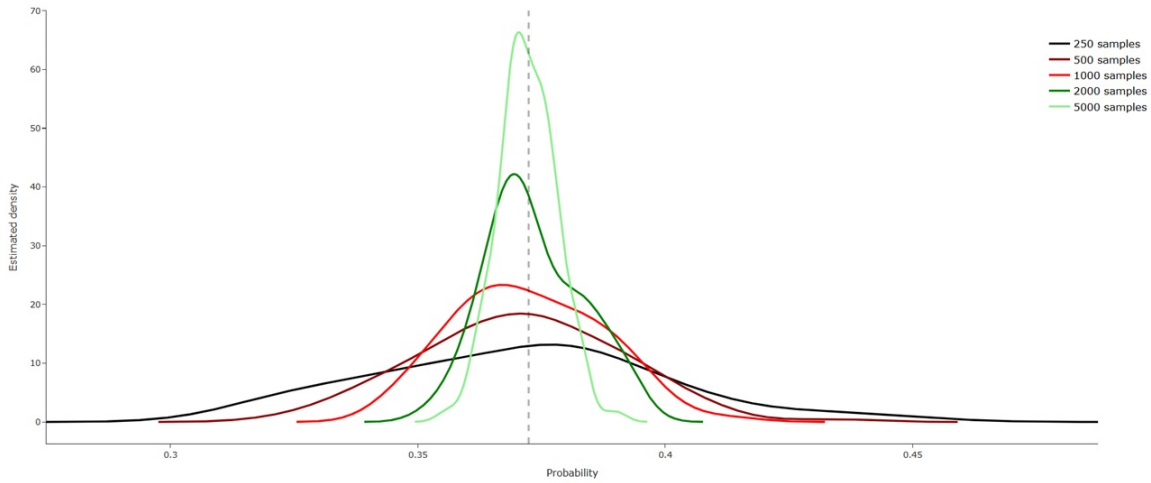
, where we choose $\zeta = T\left((J-1)\,\Delta\tau|\mathbf{u}_1,\cdots,\mathbf{u}_N,\epsilon^*,\alpha\right)\times\mu\left(\partial\mathcal{E}_{(1-(J-1)\Delta\tau)\epsilon^*}\right)^{-1}$. We leave the investigation of more adequate approximation strategies for ζ open for further research.

4 Numerical application

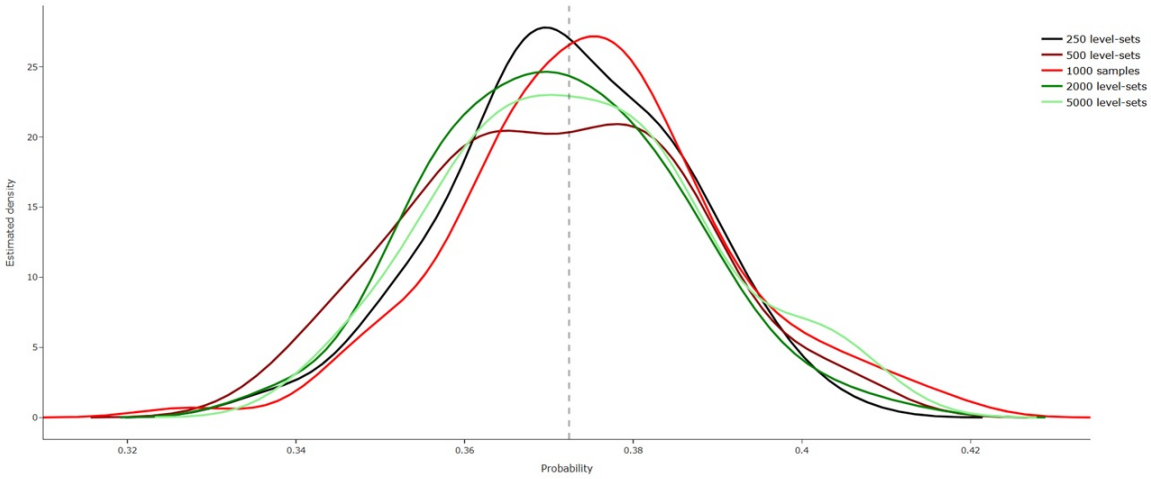
In this section, we perform a series of tests using univariate polynomials of (dependent) random variables, where the random vector $\mathbf{X}$ has dimension n equal to 3, 6 and 10. We assume $\mathbf{X}$ follows a multivariate Gaussian distribution, which allows us to directly and simply express the probability density of $h(\mathbf{X})$ under closed-form and compute the exact value of $\mathbb{P}(\varphi(\mathbf{x}, \alpha) \geq 0)$. We assume the random vector $\mathbf{X}$ has zero mean and draw a random symmetrical matrix corresponding to the variance-covariance matrix for each $n \in \{3, 6, 10\}$. As a benchmark, we directly draw samples of $(\mathbf{x}_1, \cdots, \mathbf{x}_N)$ from the density function of $\mathbf{X}$ and compute the proportion of draws $\mathbf{x}_i$, $1 \leq i \leq N$ such that $\varphi(\mathbf{x}_i, \alpha) \geq 0$. We will subsequently refer to this approach as the *reference method*.

We vary N from 2^5 to 2^{10}, and J from 250 to 10000, using step sizes equal to 32 and 250 respectively. For each combination of N and J and $\alpha \in \{-16, -8, -4, 0, 4, 8, 16\}$, we execute 100 iterations for the calculation of $\mathbb{P}(\varphi(\mathbf{x}, \alpha) \geq 0)$ using our method and the reference method. This allows us to calculate the empirical bias (defined, for a given sample size N, as the average difference between the estimated value and the true value) and variance associated to each estimator and ultimately compare their performances, by computing their respective ratios.

Before performing comparisons between the two methods, we verify that the estimator corresponding to expression **(20)** is indeed consistent. For $n = 6$ and $\alpha = 4$, **Figure 1** below displays the distribution of estimates for $\mathbb{P}(\varphi(\mathbf{x}, \alpha) \geq 0)$, by respectively holding constant the number of level-sets J and the number of samples N.



(a) J fixed ($J = 5000$)



(b) N fixed ($N = 1000$)

Figure 1: Distribution of estimates for $\mathbb{P}\left(\varphi(\mathbf{x}, \alpha) \geq 0\right)$ ($n = 6$, $\alpha = 4$)

As obseved in **Figure 2**, our method consistenly exhibits median bias ratios below 1, for all α and n. We find that the bias associated to our method is on average 14.2 % and 2.7 % lower for n equal to 3 and 6 respectively across the different values of α. For $n = 10$, although more than 50

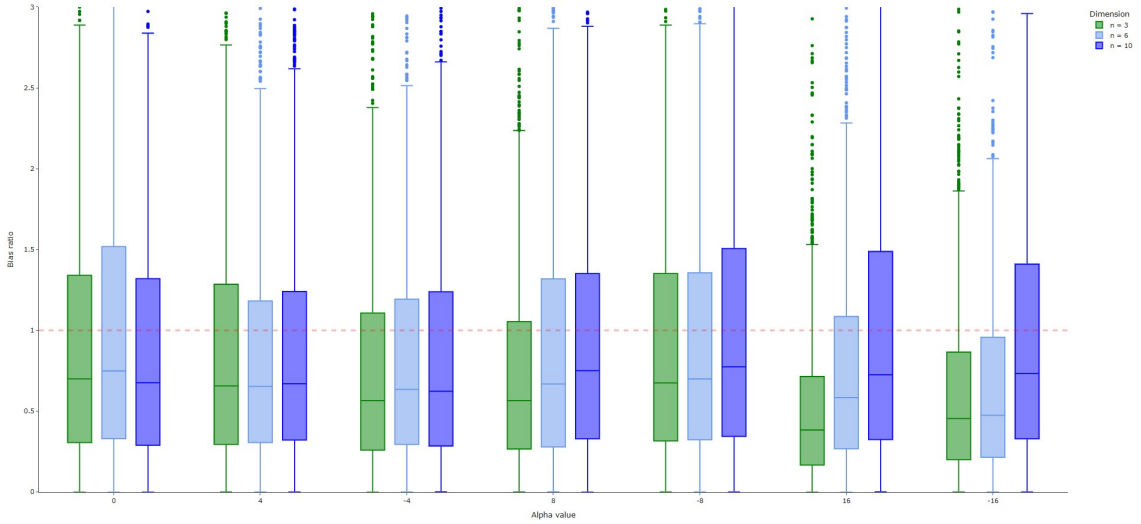


Figure 2: Distribution of the bias ratios (own vs. reference method)

% of simulations exhibit lower bias, the presence of outliers skews the distributions of ratio to the right and yield higher bias on average.

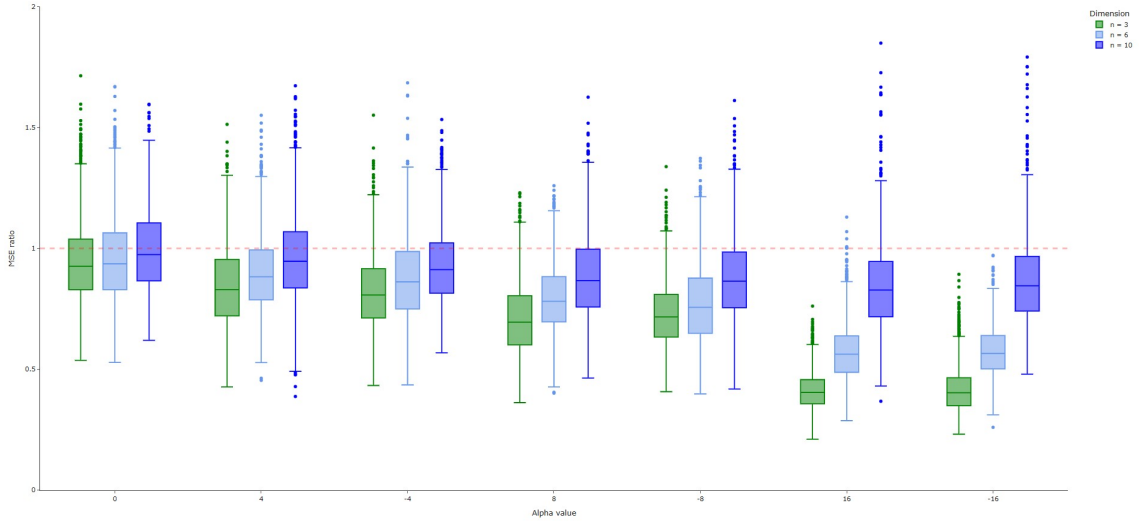


Figure 3: Distribution of the MSE ratios (own vs. reference method)

Figure 3 displays the distribution of MSE ratios. Quite strikingly, our method exhibits lower MSE values than the reference method for at least 75 % of (N, J) combination when $|\alpha| > 0$. Actually, we observe the MSE differential increases with higher values of α in absolute terms. For n

equal to 3, 6 and 10, the MSE is on average 30.4 %, 21.2 % and 9.8 % lower accross all values of α.

We should note that the bias and mean squared error are likely to be overestimated as we execute a single draw for the sample $(\mathbf{u}_1, \cdots, \mathbf{u}_N)$, which is used for the computation of $T\left(j\Delta\tau|\mathbf{u}_1, \cdots, \mathbf{u}_N, \epsilon^*, \alpha\right)$, $\forall j \in \{1, \cdots, J\}$. Indeed, for any couple $(j, j') \in \{1, \cdots, J\}$, the estimation errors for $T\left(j\Delta\tau|\mathbf{u}_1, \cdots, \mathbf{u}_N, \epsilon^*, \alpha\right)$ and $T\left(j'\Delta\tau|\mathbf{u}_1, \cdots, \mathbf{u}_N, \epsilon^*, \alpha\right)$ will be positively correlated, such that the aggregate estimation error will be greater compared to a situation where we would use a distinct sample for each separate estimate. However, resampling from $\mathbf{U}$ for each j may significantly increase the computation time of our algorithm, which would require code optimization and the development of decomposition methods using symmetries of the hyperellipsoid to decrease resolution time. We leave these issues for further research.

5 Conclusion

As illustrated above, this novel method provides significantly better CDF approximation performances compared to the direct approach, both in terms of bias and MSE. These results are especially encouraging as we are likely to overestimate the bias and estimation error because of a non-optimized sampling procedure. Although reasonable, the computational time may also be greatly reduced, by taking advantage of the multiple symmetries of the hyperellipsoid which allow for decomposition approaches.

The presented methodology may actually be considered as a specific application to elliptical distributions with explicit density functions, and may be generalized to any continuous probability distribution for which the Lebesgue measure of their iso-density hypersurfaces, for any $\epsilon > 0$, can be computed under closed-form. The specific interest of elliptical distributions lies in their ubiquitous use in robust statistics and risk analysis. The expression in **(7)** may likewise be generalized to any continous multivariate distribution which iso-density hypersurfaces correspond to an injective mapping from the unit sphere. However, the application of non-linear mappings may require the evaluation of difficult integrals on the unit sphere, which do not necessarily admit a closed-form solution.

An obvious link may be established with level-set methods (see [**14**]) from the fields of computer vision and 3D image reconstruction. [**15**] propose a specialized reparameterization technique for differentiable optimization that augments the ray tracing method, which is then combined with signed distance functions to formulate an original physically-based differentiable rendering method. The authors further establish an interesting connection between their reparameterization method and the application of the divergence theorem for integrals over the unit sphere in 3D (see [**16**]).

Surface reconstruction from oriented point clouds is a fundamental research problem in computer graphics. It strongly relies on implicit surface representations, where a shape is implicitly defined as the zero level-set of a specific scalar field. [**17**] propose to incorporate the normal vector data from sampled points to directly approximate the implicit function defining a surface S within a bounded volume V as a smooth approximation of the signed distance function to S. The vector field is then constrained to be the gradient of the implicit function.

Our approach uses the function η to assign an orientation to points sampled on the boundary of the hyperellipsoid $\mathcal{E}_\epsilon$. We neglect valuable information on their signed distance with respect to Φ_α as well as their normal vector data, coupled with prior knowledge of $h(\mathbf{x})$. Our future work will thus seek to integrate these insights from computer graphics and computational geometry to provide a generalized and more efficient formulation of our algorithm.

6 Bibliography

[**1**] LODWICK Weldon A. & JAMISON K. David, "Estimating and Validating the Cumulative Distribution of a Function of Random Variables: Toward the Development of Distribution Arithmetic", *Reliable Computing*, vol. 9, 2003, pp. 127-141

[**2**] GIL-PELEAZ J., "Note on the Inversion Theorem", *Biometrika*, vol. 38, n°3/4, 1951, pp. 481-482

[**3**] PAOLELLA Marc S., *Linear Models and Time-Series Analysis: Regression, ANOVA, ARMA and GARCH*, Wiley, First Edition, 2018, 896 p.

[**4**] LUGANNANI Robert & RICE Stephen, "Saddle Point Approximation for the Distribution of the Sum of Independent Random Variables", *Advances in Applied Probability*, vol. 12, n°2, 1980, pp. 475-490

[**5**] ANADE Dadja, GORCE Jean-Marie, MARY Philippe, PERLAZA Samir M., "Saddlepoint Approximations of Cumulative Distribution Functions of Sums of Random Vectors", *2021 IEEE International Symposium on Information Theory*, 2021, pp. 778-783

[**6**] WIENER N., "The Homogeneous Chaos", *American Journal of Mathematics*, vol. 60, n°4, 1938, pp. 897-936

[**7**] XIU Dongbin & KARNIADAKIS George Em, "The Wiener–Askey Polynomial Chaos for Stochastic Differential Equations", *SIAM Journal on Scientific Computing*, vol. 24, n°2, 2002, pp.

619-644

[**8**] LE MAITRE O. P. & KNIO O. M., "Spectral Extansions", *Spectral Methods for Uncertainty Quantification*, Springer, 2010, pp. 17-44

[**9**] JAKEMAN John D., FRANZELIN Fabian, NARAYAN Akil, ELDRED Michael, PFLÜGER Dirk, "Polynomial chaos expansions for dependent random variables", *Computer Methods in Applied Mechanics and Engineering*, vol. 351, 2019, pp. 643-666

[**10**] TABAK Esteban G. & VANDEN-EIJNDEN Eric, "Density estimation by dual ascent of the log-likelihood", *Communications in Mathematical Sciences*, vol. 8, n°1, 2010, pp. 217-233

[**11**] TABAK Esteban G. & TURNER Cristina V., "A family of nonparametric density estimation algorithms", *Communications on Pure and Applied Mathematics*, vol. 66, n°2, 2013, pp. 145-164

[**12**] DINH Laurent, KRUEGER David, BENGIO Yoshua, "NICE: Non-Linear Independent Components Estimation", ICLR Workshop, 2015

[**13**] GUILLEMIN Victor & HAINE Peter, *Differential Forms*, World Scientific Publishing Company, 1st Edition, 2019, 272 p.

[**14**] OSHER Stanley & SETHIAN James A., "Fronts propagating with curvature-dependent speed: Algorithms based on Hamilton-Jacobi formulations", *Journal of Computational Physics*, vol. 79, n°1, 1988, pp.12-49

[**15**] VICINI Delio, SPEIERER Sébastien, JAKOB Wenzel, "Differentiable Signed Distance Function Rendering", *ACM Transactions on Graphics*, vol. 41, n°4, 2022, pp. 1-18

[**16**] BANGARU Sai Praveen, LI Tzu-Mao, DURAND Frédo, "Unbiased warped-area sampling for differentiable rendering", *ACM Transactions on Graphics*, vol. 39, n°6, 2020, pp. 1-18

[**17**] CALAKLI F. & TAUBIN G., "SSD: Smooth Signed Distance Surface Reconstruction", *Computer Graphics Forum*, vol. 30, n°7, 2011, pp. 1993-2002