MDPI

Article

Characteristic Function of the Tsallis q-Gaussian and Its Applications in Measurement and Metrology

Viktor Witkovský

Institute of Measurement Science, Slovak Academy of Sciences, Dúbravská Cesta 9, 841 04 Bratislava, Slovakia; witkovsky@savba.sk

Abstract: The Tsallis q-Gaussian distribution is a powerful generalization of the standard Gaussian distribution and is commonly used in various fields, including non-extensive statistical mechanics, financial markets and image processing. It belongs to the q-distribution family, which is characterized by a non-additive entropy. Due to their versatility and practicality, q-Gaussians are a natural choice for modeling input quantities in measurement models. This paper presents the characteristic function of a linear combination of independent q-Gaussian random variables and proposes a numerical method for its inversion. The proposed technique makes it possible to determine the exact probability distribution of the output quantity in linear measurement models, with the input quantities modeled as independent q-Gaussian random variables. It provides an alternative computational procedure to the Monte Carlo method for uncertainty analysis through the propagation of distributions.

1. Introduction

According to Supplement 1 [1] of the Guide to the Expression of Uncertainty in Measurement (GUM) [2], a method for evaluating measurement uncertainty consists of three main stages: formulation, propagation and summarization. The first stage, formulation, is particularly crucial, as it involves a series of important steps that must be taken in order to accurately determine the measurement output quantity or measurand. These steps include defining the measurand, identifying the quantities that influence it, developing a model or measurement equation that relates the output to the inputs and assigning probability density functions (PDFs) to the input quantities based on available knowledge. Typically, this knowledge is derived from direct measurements and expert knowledge. In addition to the common PDFs such as Gaussian (normal) and rectangular (uniform) distributions, other distributions based on reasonable principles and available information may also be used.

In some cases, it may be necessary to assign PDFs to quantities that have not been explicitly measured or for which only partial information is available. The principle of maximum entropy (Entropy is a fundamental concept that finds applications in various scientific and engineering fields such as measurement, probability, statistics and information theory. It was introduced by William Rankine, Rudolf Clausius, Ludwig Boltzmann, Josiah Willard Gibbs, James Clerk Maxwell and other scientists in the second half of the 19th century. Claude Shannon later expanded on the concept of entropy in information theory. Entropy provides a useful tool for characterizing uncertainty, randomness and information content in different systems. The higher the entropy, the more disordered or uncertain the system is. In measuring uncertainty, entropy quantifies the uncertainty in a measurement by characterizing the probability distribution of the measurement result. The principle of maximum entropy states that the probability distribution that best represents the current state of knowledge about a system is the one with the largest entropy. This principle was

first articulated by Edwin Thompson Jaynes in 1957) is a valuable tool for this task, as it allows us to construct a PDF that accurately characterizes our incomplete knowledge of a quantity. This involves maximizing the traditional entropy, as defined, e.g., by Shannon, subject to constraints imposed by the available information. The principle of maximum entropy is particularly useful in situations where there are no indications available and we must rely solely on the available information to represent the PDF of a given quantity. To learn more about the principle of maximum entropy and its application in measurement and uncertainty evaluation, consult the Supplements to the GUM [1,3].

The Tsallis q-Gaussian distribution is a probability distribution introduced by Tsallis [4] as a generalization of the standard normal (Gaussian) distribution based on maximizing the Tsallis entropy under appropriate constraints. It belongs to a larger family of probability distributions known as q-distributions, characterized by a non-additive entropy that generalizes the Boltzmann–Gibbs entropy used in statistical mechanics; for more details, see, e.g., [5]. Non-additive entropy refers to a family of entropy measures that do not satisfy the additivity property of the traditional additive entropy measures. Additivity means that the entropy of a joint system can be obtained by adding the entropy of the individual systems. However, in some cases, additivity may not hold, particularly for complex systems with non-linear interactions between their components.

Non-additive entropy measures are used in various fields, such as physics, information theory and economics, to quantify the degree of uncertainty or disorder in a system that involves non-linear interactions. Examples of non-additive entropy measures include Tsallis entropy and Renyi entropy, which are widely used in the study of complex systems and statistical mechanics. These measures have been found to be more suitable for modeling complex systems than the traditional additive entropy measures.

Compared to the standard Gaussian distribution the spread of the q-Gaussian distribution depends on both the scale parameter and the q-index parameter which leads to different behavior in the tails of the distribution. Depending on the value of the q-index parameter, the q-Gaussian distribution may have unique properties that make it useful for modeling specific data. When the value of the q-index parameter is less than one, the distribution has a finite support, meaning that it is bounded on both sides. This can be useful for modeling data that are known to have a specific range, such as time intervals or distances. On the other hand, when the value of the q-index parameter is greater than or equal to 5/3, the variance of the q-Gaussian distribution does not exist, meaning that the distribution has infinite variance. Moreover, when $2 \leq q < 3$, other moments of the distribution may also not exist or may not be well-defined and the q-Gaussian distribution is located in a region of extremely heavy tails. This can be useful for modeling data that has extreme values, such as data with outliers or natural phenomena with rare but significant events.

The q-Gaussian distribution is a versatile and practical tool that finds potential applications in various fields such as physics, astronomy, geology, anatomy, economics and finance, molecular biology and engineering. Its ability to model complex systems with long-range interactions, memory effects or non-equilibrium dynamics makes it a useful distribution for real industry applications and as a modeling distribution of input quantities in measurement models.

In statistical mechanics, q-Gaussian distributions can describe the velocity distribution of particles in non-extensive systems such as turbulent fluids, plasmas and granular gases. In geology, q-Gaussian distributions can fit the frequency-magnitude distribution of earthquakes and volcanic eruptions. In anatomy, q-Gaussian distributions can model the distribution of human inter-beat intervals and brain activity. In astronomy, q-Gaussian distributions can represent the distribution of velocities of stars in globular clusters and galaxies. In economics, q-Gaussian distributions can capture the fat-tailed behavior of financial returns and volatility. In machine learning, q-Gaussian distributions can be used as activation functions or basis functions for neural networks and radial basis function networks.

Vignat and Plastino [6] explore the possible reasons why q-Gaussian distributions are frequently observed in various natural and artificial phenomena. They argue that the detection of q-Gaussian behavior may be influenced by the normalization process performed by the measurement device. If the incoming data have elliptical symmetry, a common property of many distributions, then the normalized data will always follow a q-Gaussian distribution, with a parameter q that depends on the normalization technique.

Beck et al. [7] show that the velocity distribution of particles in a turbulent fluid can be modeled via a q-Gaussian distribution, with a parameter q that depends on the Reynolds number, and propose a superstatistical framework to explain the origin of q-Gaussian behavior in complex systems. Several studies, e.g., [8], have found q-Gaussian behavior in the returns and volatility of financial assets, with different values of q for different markets and time scales. Carpena et al. [9] analyzed the distribution of distances between consecutive occurrences of a given nucleotide in DNA sequences and found that it follows a q-Gaussian distribution with q close to 2. Burlaga et al. [10] studied the distribution of fluctuations in the magnetic field and plasma density of the solar wind and observed that it is well-fitted by a q-Gaussian distribution with q around 1.6 and relates the value of q to the nonextensive entropy parameter. Anteneodo in [11] investigates the nature of random variables whose sums lead to q-Gaussian distributions and proposes a simple statistical mechanism based on non-extensive random walks. The paper also discusses some examples of applications of q-Gaussian distributions and non-extensive random walks in physics, biology and finance.

In measurement, it is common to encounter the challenge of having only a small number of measurement repetitions. When using the Type A evaluation method with $n \leq 3$ measurements that have Gaussian errors, we need to use the Student's t distribution with $\nu \leq 2$ degrees of freedom to model the distribution of the input variable. However, in this case the variance of the distribution does not exist, which requires caution when evaluating the associated measurement uncertainty. Nonetheless, as will be explained later, this distribution can also be represented as a q-Gaussian distribution with $q \geq 5/3$.

Combining independent q-Gaussian random variables is a powerful tool for modeling the behavior of a measurand that involves diverse and complex input quantities. The q-Gaussians offer a flexible family of distributions that can capture a wide range of statistical behaviors, making it a suitable choice for such modeling tasks.

In this paper, we propose an approach to assess the probability distribution of output quantities in linear measurement models using the Characteristic Function Approach (CFA), as described in [12]. An alternative approach is through the propagation of distributions using Monte Carlo methods, as outlined in the Supplements of the GUM [1,3].

The paper is arranged as follows. Section 2 presents an overview of the characteristic function approach for assessing the measurement uncertainty. In Section 3, we derive the characteristic function of the q-Gaussian random variables and present a numerical method for inverting the characteristic function (CF) of a linear combination of independent q-Gaussian random variables, say $Y = \sum_{k=1}^{n} c_k X_k$. Section 4 describes the basic functionality of the MATLAB toolbox **CharFunTool** which provides a collection of characteristic functions of various probability distributions and algorithms for their manipulation and inversion. Section 5 presents some illustrative numerical examples. Section 6 includes discussion and conclusions.

2. Characteristic Function Approach for Assessing the Measurement Uncertainty

There are various methods and approaches for assessing the measurement uncertainty of a quantity that depends on multiple input quantities characterized with different probability distributions. The most commonly used methods are those advocated in [1–3]. However, in certain cases, when dealing with distributions such as the Tsallis q-Gaussian distribution with a q-index greater than or equal to 5/3, the standard uncertainty of some of the considered input quantities may not exist. In such cases, the approach to evaluating the uncertainty using the law of propagation of uncertainty, as specified in [2], may not be

applicable. Nonetheless, alternative approaches can be used to assess the uncertainty by evaluating the coverage intervals in these situations and the expanded uncertainty can be formally stated as the half-width of the interval.

Here we adopt the approach of GUM Supplements [1,3] which consider the output distribution as a posterior distribution for the measurand obtained by propagating the distributions of the input quantities through the measurement model. This posterior distribution is then utilized to calculate an interval or region that encompasses a large fraction of the distribution of values that could reasonably be attributed to the measured quantity based on available information. Based on using this paradigm, an ideal method for evaluating and expressing measurement uncertainty should efficiently provide such an interval (or region), possibly under additional constraints, with a coverage probability or level of confidence that corresponds to the required level of certainty, typically with a specified coverage level, such as 95%. Generally, increasing the coverage level results in a wider coverage interval or a larger coverage region.

One of the alternative methods to the Monte Carlo method (MCM) proposed in [1,3] for deriving the probability distribution of the measured quantity and assessing measurement uncertainty is the Characteristic Function Approach [12]. This approach uses characteristic functions to fully determine the behavior and properties of probability distributions of random variables. Although characteristic functions have a fundamental role in several fields of mathematics, probability theory, statistics and engineering, they can be challenging to work with, particularly when it comes to inverting them to obtain the cumulative distribution function (CDF) or the probability distribution function (PDF) of the corresponding random variable and subsequently deriving the corresponding coverage interval with a specified level of confidence. To address these challenges, various tools and methods have been developed for evaluating, combining and inverting characteristic functions.

The CFA is primarily used to assess the probability distribution of a quantity that depends linearly on multiple input quantities with different probability distributions. However, it can also be applied to nonlinear measurement models if the characteristic function of the output quantity can be derived or by approximating the nonlinear function linearly to derive the characteristic function of the output quantity. This linear approximation is based on the first-order Taylor series expansion of the nonlinear function around the expected value or the best estimate of the input quantities. The CFA is capable of handling any type of probability distribution, including non-Gaussian and asymmetric ones, as long as the characteristic functions of the input quantities are known. The CFA obtains the characteristic function of the output quantity by applying a transformation rule to the characteristic functions of the input quantities.

In fact, the characteristic function of a weighted sum of independent random variables is simple to derive if the measurement model is linear, i.e.,

$$Y = f(X_1, \ldots, X_n) = c_1 X_1 + \cdots + c_n X_n, \tag{1}$$

for some known constants $c_1, \ldots, c_n$ (the sensitivity coefficients) and the input quantities X_i are mutually independent random variables with known distributions. In such situations, the CF of the output quantity Y is given as

$$\mathrm{cf}_Y(t) = \mathrm{cf}_{X_1}(c_1 t) \times \cdots \times \mathrm{cf}_{X_n}(c_n t), \tag{2}$$

where by $\mathrm{cf}_{X_i}(t)$ we denote the (known) CFs of the input quantities X_i.

The probability distribution function and the cumulative distribution function of the output quantity Y can be obtained by numerically inverting its characteristic function, which can be efficiently calculated, for example, using a simple trapezoidal quadrature or other advanced quadrature rules as is the adaptive Gauss–Kronod quadrature, depending on the complexity and oscilatory properties of the integrand function; see [13,14]. From the cumulative distribution function, the quantile function and coverage intervals with a specified coverage level can be derived by selecting appropriate quantiles of the distribution.

The CFA compares favorably with the MCM, which estimates the probability distribution of the output quantity by generating random samples from the input quantities and applying the measurement model and the law of propagation of distributions [1,3]. It offers several advantages, including accuracy (The proper error analysis of the approaches used for numerical inversion has been studied in the literature for several particular cases. In general, it is still an open problem which depends on the specific inversion algorithm and the properties of the characteristic function. Shepard in [15] analyzed the error caused by approximating the required inversion integral under specific assumptions. The author suggested that using the trapezium rule to approximate the integral makes it challenging to manage the induced integration error. Therefore, a simpler Riemann sum has been employed instead, resulting in a much simpler numerical integration error. This formulation has been successfully applied to solve the problem of finding the exact distribution function of a quadratic form in normal variables with adequate control of the numerical error, as demonstrated by Davies in [16].), flexibility, efficiency and repeatability, making it a valuable method for assessing measurement uncertainty. However, it faces challenges in finding an explicit expression for the characteristic function of the output quantity and inverting it to obtain the CDF, PDF or other statistics. It is worth noting that deriving the joint characteristic function of a multivariate distribution with stochastically dependent input quantities and numerically inverting such a characteristic function can be challenging.

3. Characteristic Function of q-Gaussian Distribution

Here we present the derivation of the characteristic function for the q-Gaussian distribution as well as the characteristic function for a linear combination of independent q-Gaussian random variables, $Y = \sum_{k=1}^{n} c_k X_k$, where X_k are independent q-Gaussian random variables possibly with different parameters and c_k are known constant coefficients.

The probability density function of the Tsallis q-Gaussian distribution is given by:

$$f(x) = C_{q,\sigma} \left[1 - (1-q) \frac{1}{2} \left(\frac{x-\mu}{\sigma} \right)^2 \right]_+^{\frac{1}{1-q}}, \tag{3}$$

where $C_{q,\sigma}$ is a normalization constant, μ (real) and $\sigma > 0$ are the location and scale parameters of the distribution and $[z]_+ = \max(0, z)$. The shape parameter $q < 3$ controls the degree of non-extensivity of the system and the distribution reduces to the Gaussian distribution when $q = 1$.

The parametrization used in this paper is consistent with Wolfram Mathematica, where $\sigma > 0$ represents the scale parameter. It is important to note that this differs from the parametrization defined by Tsallis—for details see also Wikipedia—where

$$f(x) = C_{q,\beta} \left[1 - (1-q)\beta(x-\mu)^2 \right]_+^{\frac{1}{1-q}}, \tag{4}$$

which uses $\beta > 0$ as the rate parameter, such that $\beta = 1/(2\sigma^2)$ or $\sigma = \sqrt{1/(2\beta)}$.

We use $X \sim TQG(\mu, \sigma, q)$ to denote a random variable with a Tsallis q-Gaussian distribution, where $\mu \in \mathbb{R}$ is the location parameter, $\sigma > 0$ is the scale parameter and $q < 3$ is the shape parameter known as the Tsallis q-index. The distribution $TQG(\mu, \sigma, q)$ is parametrized such that it is a normal distribution with mean μ and variance σ^2 when $q = 1$, i.e., $X \sim N(\mu, \sigma^2)$, or

$$X = \mu + \sigma Z, \tag{5}$$

where $Z \sim N(0, 1)$ is a random variable with the standard normal distribution.

Moreover, for $q < 1$, X is a bounded random variable with its distribution proportional to a scaled and shifted symmetric beta distribution. In particular, we obtain

$$X = \mu + \sigma\sqrt{\frac{2}{1-q}}(2B - 1), \tag{6}$$

where $B \sim Beta(\theta, \theta)$ is a beta distributed random variable with both shape parameters equal to $\theta = \frac{2-q}{1-q}$. The support of this distribution is limited to

$$\left\langle \mu - \sigma\sqrt{\frac{2}{1-q}}, \mu + \sigma\sqrt{\frac{2}{1-q}} \right\rangle. \tag{7}$$

For $1 < q < 3$, the distribution is formally proportional to a scaled and shifted Student t-distribution, which is often preferred over the Gaussian distribution due to its heavy tails, meaning that it has a higher probability of extreme events. In particular, we obtain

$$X = \mu + \sigma\sqrt{\frac{2}{3-q}}T, \tag{8}$$

where $T \sim t(\nu)$ is a Student t-distributed random variable with $\nu = \frac{3-q}{q-1}$ degrees of freedom.

It is worth noting that as q approaches 1 from above, ν approaches infinity (i.e., normal distribution). For $q \geq \frac{5}{3}$, the second moments are infinite or do not exist. For $2 \leq q < 3$, the q-Gaussian distribution is located in a region of extremely heavy tails where the moments of the distribution may not exist or may not be well-defined. If $q = 2$, the q-Gaussian distribution is proportional to a scaled and shifted Cauchy distribution; otherwise, it is a t-distribution with fractional degrees of freedom, $0 < \nu < 1$. Specifically, as q approaches 3, $\nu \to 0$ and as q approaches 2, $\nu \to 1$.

The main result of the paper is presented in Theorem 1 and was obtained through symbolic computation using Wolfram Mathematica and validated through comparison with the characteristic functions of the Student t and symmetric beta distributions; see [14,17].

Theorem 1 (Characteristic function of the Tsallis q-Gaussian). *The characteristic function of the Tsallis q-Gaussian distribution with the parameters $\mu \in \mathbb{R}$, $\sigma > 0$ and $q < 3$ is defined as*

$$\mathrm{cf}_{TQG(\mu,\sigma,q)}(t) = \exp(it\mu) \times \mathrm{cf}_{TQG(0,1,q)}(\sigma t), \tag{9}$$

where $\mathrm{i} = \sqrt{-1}$ *and* $\mathrm{cf}_{TQG(0,1,q)}(t)$ *is CF of the standard Tsallis q-Gaussian distribution,*

- *for $q < 1$ defined as*

$$\mathrm{cf}_{TQG(0,1,q)}(t) = {}_0\mathrm{F}_1\left(\theta + \frac{1}{2}, -\frac{1}{4}(at)^2\right) = 2^{\theta - \frac{1}{2}}\Gamma\left(\theta + \frac{1}{2}\right)(at)^{-(\theta + \frac{1}{2})}\mathrm{J}_{\theta - \frac{1}{2}}(at)$$

$$= \mathrm{cf}_{BetaSymmetric(\theta)}(at), \tag{10}$$

where $a = \sqrt{\frac{2}{1-q}}$, $\theta = \frac{2-q}{1-q}$, ${}_0\mathrm{F}_1(b,z)$ is a confluent hypergeometric function, $\Gamma(z)$ is a gamma function, $\mathrm{J}_\nu(z)$ is a Bessel function of the first kind and $\mathrm{cf}_{BetaSymmetric(\theta)}(t)$ denotes CF of the symmetric beta distribution with the parameter θ and the support on the interval $\langle -1, 1 \rangle$,

- *for $q = 1$ defined as*

$$\mathrm{cf}_{TQG(0,1,q)}(t) = \exp\left(-\frac{t^2}{2}\right) = \mathrm{cf}_{Normal(0,1)}(t). \tag{11}$$

- *for $1 < q < 3$ defined as*

$$\text{cf}_{TQG(0,1,q)}(t) = \frac{\left(b\sqrt{\nu}|t|\right)^{\frac{\nu}{2}} \text{K}_{\frac{\nu}{2}}\left(b\sqrt{\nu}|t|\right)}{2^{\frac{\nu}{2}-1}\Gamma\left(\frac{\nu}{2}\right)} = \text{cf}_{Student(\nu)}(bt),$$ (12)

where $b = \sqrt{\frac{2}{3-q}}$, $\nu = \frac{3-q}{q-1}$, $\text{K}_\nu(z)$ is a Bessel function of the second kind and $\text{cf}_{Student(\nu)}(t)$ denotes CF of the Student t-distribution with ν degrees of freedom with $\nu > 0$.

Corollary 1 (Characteristic function of a linear combination of the Tsallis *q*-Gaussians). *Let $\text{cf}_Y(t)$ denote the characteristic function of a linear combination of independent random variables $Y = \sum_{k=1}^{n} c_k X_k$, where c_k are real coefficients and X_k are independent Tsallis q-Gaussians with characteristic functions $\text{cf}_{X_k} = \text{cf}_{TQG(\mu_k,\sigma_k,q_k)}(t)$ for $k = 1, \ldots, n$. Then, the characteristic function of Y can be expressed as follows:*

$$\text{cf}_Y(t) = \prod_{k=1}^{n} \text{cf}_{X_k}(c_k t) = \prod_{k=1}^{n} \text{cf}_{TQG(\mu_k,\sigma_k,q_k)}(c_k t).$$ (13)

This formula allows us to compute the characteristic function of Y by taking the product of the characteristic functions of each X_k, evaluated at $c_k t$.

4. CharFunTool: The Characteristic Functions Toolbox

While the literature presents various techniques and tools for evaluating, combining and inverting characteristic functions, few reliable and efficient software implementations exist. One notable implementation, still under continuous development, is the MATLAB toolbox **CharFunTool** [14], which offers a range of characteristic functions for various probability distributions and algorithms for their manipulation and inversion. The toolbox can evaluate, combine and invert characteristic functions of different types of probability distributions, including continuous, discrete, circular and bivariate distributions, as well as mixture and empirical distributions. With over 60 characteristic functions available, the toolbox also provides a wide variety of other symmetric and non-negative probability distributions, such as Beta, Chi-Square, Exponential, Gamma, Normal, Poisson, Rayleigh, Student, Weibull and many others.

The algorithm **cf_TsallisQGaussian** was implemented into the repository of the characteristic functions of the toolbox. The algorithm evaluates the characteristic function of the random variable $Y = \sum_{k=1}^{n} c_k X_k$, where c_k are real coefficients and X_k are independent Tsallis *q*-Gaussians with arbitrary parameters for $k = 1, \ldots, n$. In general, the toolbox offers a unique and straightforward approach to evaluating, combining and inverting combined characteristic functions.

The well-known method for inverting characteristic functions that satisfies conditions specified in [18,19] is the Gil-Pelaez formula, which was proposed by Gil-Pelaez in 1951 [20]. This formula allows one to invert the characteristic function by using numerical integration techniques. The formula states that for a univariate random variable X, if x is a continuity point of its cumulative distribution function, then

$$\text{cdf}_X(x) = \frac{1}{2} - \frac{1}{\pi} \int_0^\infty \frac{\text{Im}\left[e^{-itx}\,\text{cf}_X(t)\right]}{t}\,dt.$$ (14)

Moreover,

$$\text{pdf}_X(x) = \frac{1}{\pi} \int_0^\infty \text{Re}\left[e^{-itx}\,\text{cf}_X(t)\right]dt,$$ (15)

where cdf_X is the CDF of X, pdf_X is the PDF of X, cf_X is the characteristic function of X and Re and Im denote the real and imaginary parts of complex numbers, respectively. However, the use of the formulae may be affected by numerical instability and slow

convergence in cases where the characteristic function or the integrands exhibits oscillatory or singular behavior.

There are many alternative methods and improvements available for computing the numerical inversion of the characteristic function. Several of these methods have been directly implemented or have inspired the development of inversion algorithms in the **CharFunTool**, namely [13,15,18,21–32].

The toolbox offers a range of algorithms for numerically inverting characteristic functions of univariate and bivariate probability distributions, which can be categorized into two categories: univariate inversion algorithms and bivariate inversion algorithms.

Inversion Algorithms

Univariate inversion algorithms are used to obtain the probability distribution function, cumulative distribution function, quantile function (QF) or probability mass function (PMF) of a univariate random variable, given its characteristic function. Various algorithms for univariate inversion include:

- **cf2DistGP**: This algorithm uses the Gil-Pelaez inversion formulae (14) and (15) to return the PDF, CDF or QF of a univariate distribution. It also provides an option to use different integration algorithms such as Riemann sum, trapezoidal rule or adaptive Gauss–Kronrod quadrature rule for more efficient calculations.
- **cf2DistGPR**: This algorithm returns the PDF, CDF or QF of a univariate distribution using the Gil-Pelaez inversion formulae with the Riemann sum integration method. For more details, see, e.g., [18].
- **cf2DistGPT**: This algorithm returns the PDF, CDF or QF of a univariate distribution using the Gil-Pelaez inversion formulae with the trapezoidal rule integration method. For more details, see, e.g., [13,18,25,32].
- **cf2DistGPA**: This algorithm returns the PMF of a discrete univariate distribution using the Gil-Pelaez inversion formulae with the adaptive Gauss–Kronrod quadrature rule integration method. For more details, see, e.g., [24,30].
- **cf2DistBTAV**: This algorithm uses the Bromwich–Talbot–Abate–Valko (BTAV) method to return the CDF, PDF or QF of a non-negative univariate distribution specified by a given characteristic function. For more details, see, e.g., [21,26].
- **cf2DistBV**: This algorithm returns the CDF, PDF or QF of a univariate distribution using the Gil-Pelaez inversion formulae and the Bakhvalov–Vasileva method [22].
- **cf2DistFFT**: This algorithm returns the CDF, PDF or QF of a univariate distribution using the Fast Fourier Transform (FFT) algorithm. For more details, see, e.g., [23,27].

The bivariate inversion algorithm takes a characteristic function of a bivariate random vector and returns the joint PDF or CDF of the corresponding random vector:

- **cf2Dist2D**: This returns the CDF, PDF, quantile function or random numbers of a bivariate distribution using the Gil-Pelaez inversion formulae with Fourier integrals calculated using the simple Riemann sum quadrature method suggested by Shephard in 1991 [15]. See also [29].

Note that the choice of algorithm may depend on the properties of the input characteristic function, such as its smoothness, oscillatory behavior and integrability.

5. Numerical Examples

The examples in this section do not include specific measurement uncertainty problems. Instead, they illustrate the application and implementation of the CFA as an alternative tool for mathematical computation of the probability distribution of the output quantity in linear measurement models with q-Gaussian input variables, with potential applications for measurement uncertainty analysis.

For illustration, here we present the application of the **CharFunTool** algorithms for evaluating PDF, CDF and specified quantiles of the distribution of an output quantity. For each of those four examples, we consider a linear combination of independent q-Gaussian

random variables with different parameters specified as $Y = c_1 X_1 + c \cdots + c_n X_n$, where the coefficients c_i and the characteristic functions cf_{X_i} are given and this demonstrates how the CFA can be used to obtain the PDF, CDF and QF of the output quantity. The examples cover a range of scenarios, including situations where all q-index parameters are less than 1, where the q-index parameters are a mixture of positive and negative values, where the Tsallis parametrization is used and where extreme values of q are considered.

5.1. Example 1: PDF/CDF/QF of a Linear Combination of q-Gaussian RVs with Small q-Indices, $q \leq 0$

Here we consider a linear combination of three independent q-Gaussian random variables,

$$Y = 0.8X_1 + 0.15X_2 + 0.05X_3, \tag{16}$$

where $X_1 \sim TQG(0, 3, -100)$, $X_2 \sim TQG(0, 2, -10)$ and $X_3 \sim TQG(0, 1, 0)$ are independent random variables. We used the **CharFunTool** and the characteristic function of Y to compute its PDF, CDF and QF. Figure 1 shows the graphs of the PDF and CDF along with the MATLAB code used to evaluate the result. The CFA was used to derive a 95% coverage interval, which is given by $\text{CI}_{CFA} = [-0.3751, 0.3751]$.
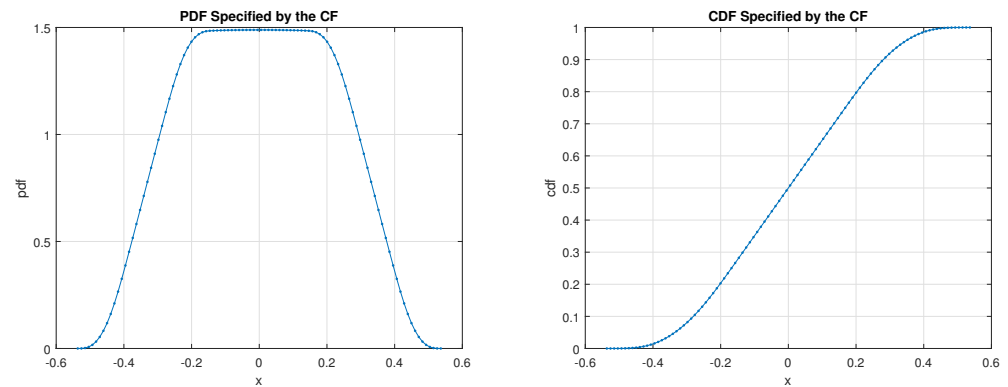

```
%% EXAMPLE 1: PDF/CDF/QF of a linear combination of independent q-Gaussians
%  RVs with all q < 1

%  Characteristic Function Approach (CFA):
mu      = [0 0 0];
sigma   = [3 2 1];
q       = [-100 -10 0];
coef    = [0.8 0.15 0.05];
cf      = @(t) cf_TsallisQGaussian(t,mu,sigma,q,coef);
clear options
options.N    = 2^10;
options.xMin = sum(mu - sigma.*sqrt(2./(1-q)) .* coef);
options.xMax = sum(mu + sigma.*sqrt(2./(1-q)) .* coef);
x       = linspace(options.xMin,options.xMax)';
prob    = [0.025 0.975];
result = cf2DistGP(cf,x,prob,options);
disp(result)
```

Figure 1. Characteristic function approach using the MATLAB algorithms from **Char-FunTool**. PDF, CDF and QF computed from the characteristic function of a linear combination of q-Gaussian RVs, $Y = 0.8X_1 + 0.15X_2 + 0.05X_3$, where $X_1 \sim TQG(0, 3, -100)$, $X_2 \sim TQG(0, 2, -10)$ and $X_3 \sim TQG(0, 1, 0)$ are independent random variables. The 95% coverage interval derived by CFA is $\text{CI}_{CFA} = [-0.3751, 0.3751]$.

Alternatively, we can use MCM with (6) to derive the result by generating $N = 100{,}000$ realizations of the X_i to obtain N realizations of Y. By sorting them, we can then derive the required coverage interval as $\text{CI}_{MCM} = [Y_{sort}(\lfloor N \times 0.025 \rfloor), Y_{sort}(\lceil N \times (1 - 0.025) \rceil)]$.

Using this method, we obtain $\text{CI}_{MCM} = [-0.3762, 0.3753]$ for comparison. Here, we can see that the coverage intervals derived using CFA and MCM are very close, with only a small difference between them.

5.2. Example 2: PDF/CDF/QF of a Linear Combination of q-Gaussian RVs with Different Types of q-Indices, $q \leq 1.5$

Consider a linear combination

$$Y = \frac{1}{3}X_1 + \frac{1}{3}X_2 + \frac{1}{3}X_3, \tag{17}$$

where $X_1 \sim TQG(0,1,-1)$, $X_2 \sim TQG(1,1,0.5)$ and $X_3 \sim TQG(2,1,1.5)$ are independent random variables with different types of q-indices. Figure 2 shows the PDF and CDF graphs along with the MATLAB code used to evaluate the result. The CFA was used to derive a 95% coverage interval, which is given by $\text{CI}_{CFA} = [-0.3409, 2.3409]$.
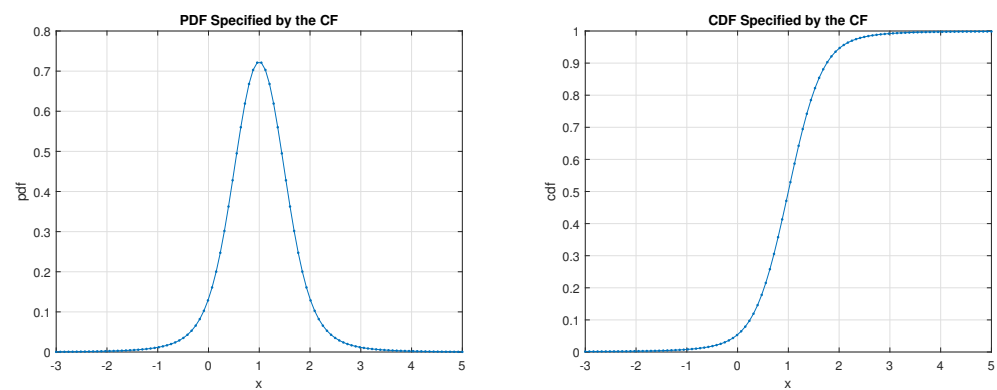

```
%% EXAMPLE 2: PDF/CDF/QF of a linear combination of independent q-Gaussians
%  RVs with different types of q-indices: q < 1 and q > 1

%  Characteristic Function Approach (CFA):
mu      = [0 1 2];
sigma   = [1 1 1];
q       = [-1 0.5 1.5];
coef    = [1 1 1]/3;
cf      = @(t) cf_TsallisQGaussian(t,mu,sigma,q,coef);
clear options
options.N = 2^10;
x       = linspace(-3,5)';
prob    = [0.025 0.975];
result = cf2DistGP(cf,x,prob,options);
disp(result)
```

Figure 2. Characteristic function approach using the MATLAB algorithms from **CharFun-Tool**. PDF, CDF and QF computed from the characteristic function of a linear combination of q-Gaussian RVs, $Y = \frac{1}{3}X_1 + \frac{1}{3}X_2 + \frac{1}{3}X_3$, where $X_1 \sim TQG(0,1,-1)$, $X_2 \sim TQG(1,1,0.5)$ and $X_3 \sim TQG(2,1,1.5)$ are independent random variables. The 95% coverage interval derived via CFA is $\text{CI}_{CFA} = [-0.3409, 2.3409]$.

We used MCM with (6) and (8) to derive the result by generating $N = 100,000$ realizations of X_i to obtain N realizations of Y. By sorting the realizations of Y, we obtained $\text{CI}_{MCM} = [-0.3248, 2.3375]$ for comparison. As before, we can see that the coverage intervals derived using CFA and MCM are very close, with only a small difference between them.

5.3. Example 3: PDF/CDF/QF of a Linear Combination of q-Gaussian RVs with Different Types of q-Indices, q ≤ 2 and Using the Tsallis Parametrization

Consider a linear combination

$$Y = \frac{1}{5}X_1 + \frac{1}{5}X_2 + \frac{1}{5}X_3 + \frac{1}{5}X_4 + \frac{1}{5}X_5, \tag{18}$$

where $X_1 \sim TQG(0, \sqrt{1/10}, -5)$, $X_2 \sim TQG(0, \sqrt{1/8}, -1)$, $X_3 \sim TQG(0, \sqrt{1/6}, 0)$, $X_4 \sim TQG(0, \sqrt{1/4}, 1)$ and $X_5 \sim TQG(0, \sqrt{1/2}, 2)$ are independent random variables defined using the Tsallis parametrization. Figure 3 shows the PDF and CDF graphs along with the MATLAB code used to evaluate the result. The CFA was used to derive a 95% coverage interval, which is given by $\text{CI}_{CFA} = [-2.5469, 2.5469]$.
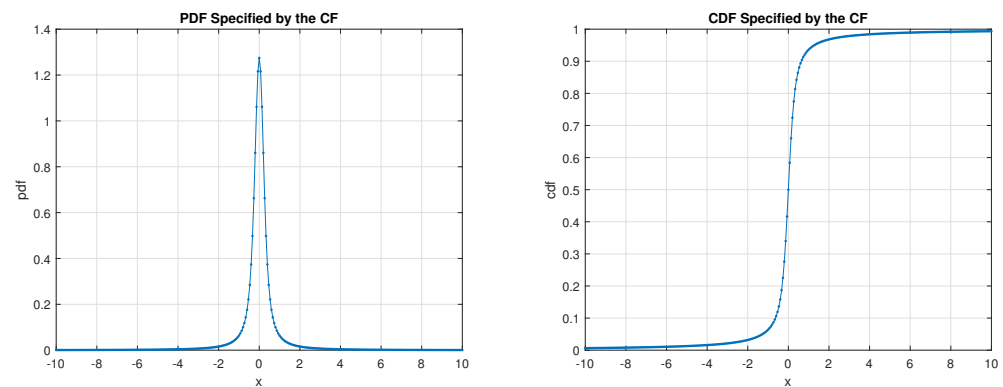

```
%% EXAMPLE 3: PDF/CDF/QF of a linear combination of independent q-Gaussians
%   RVs with different types of q-indices: q < 1 and q > 1
%   with using the Tsallis parametrization

%   Characteristic Function Approach (CFA):
mu      = [0 0 0 0 0];
beta    = [5 4 3 2 1];
sigma   = sqrt(1./(2*beta));
q       = [-5 -1 0 1 2];
coef    = [1 1 1 1 1]/5;
cf      = @(t) cf_TsallisQGaussian(t,mu,sigma,q,coef);
clear options
options.N = 2^14;
x       = linspace(-10,10,301);
prob    = [0.025 0.975];
result = cf2DistGP(cf,x,prob,options);
disp(result)
```

Figure 3. Characteristic function approach using the MATLAB algorithms from **CharFun-Tool**. PDF, CDF and QF computed from the characteristic function of a linear combination of *q*-Gaussian RVs, $Y = \frac{1}{5}X_1 + \frac{1}{5}X_2 + \frac{1}{5}X_3 + \frac{1}{5}X_4 + \frac{1}{5}X_5$, where $X_1 \sim TQG(0, \sqrt{1/10}, -5)$, $X_2 \sim TQG(0, \sqrt{1/8}, -1)$, $X_3 \sim TQG(0, \sqrt{1/6}, 0)$, $X_4 \sim TQG(0, \sqrt{1/4}, 1)$ and $X_5 \sim TQG(0, \sqrt{1/2}, 2)$ are independent random variables specified from the given Tsallis parametrization. The 95% coverage interval derived by CFA is $\text{CI}_{CFA} = [-2.5469, 2.5469]$.

We used MCM with (6), (5) and (8) to derive the result by generating $N = 100,000$ realizations of the X_i to obtain N realizations of Y. By sorting the realizations of Y, we obtained $\text{CI}_{MCM} = [-2.5482, 2.5018]$ for comparison. As before, we can see that the coverage intervals derived using CFA and MCM are very close, with only a small difference between them.

5.4. Example 4: CDF of a Linear Combination of q-Gaussian RVs Including Large q-Indices with max $q = 2.9$

We now consider an extreme case, where we have a linear combination of independent random variables X_1, X_2 and X_3 given by:

$$Y = \frac{1}{3}X_1 + \frac{1}{3}X_2 + \frac{1}{3}X_3, \tag{19}$$

where $X_1 \sim TQG(0,1,0)$, $X_2 \sim TQG(0,0.5,1)$ and $X_3 \sim TQG(0,1,2.9)$ are independent random variables.

We used MATLAB code to evaluate the specified CDF values and the 95% coverage interval for Y, as shown in Figure 4. The exact 95% coverage interval derived by CFA is $CI_{CFA} = [-9.1540 \times 10^{22}, 9.1540 \times 10^{22}]$.

```
%% EXAMPLE 4: CDF and 95% coverage interval
%  of a linear combination of independent q-Gaussians
%  RVs with different types of q-indices
%  including extreme values of q (here max q = 2.9)
%  computed with using cf2CDF_GPA and cf2QF_GPA

mu     = [0 0 0];
sigma  = [1 0.5 0.1];
q      = [ 0 1 2.9];
coef   = [1 1 1]/3;
cf     = @(t) cf_TsallisQGaussian(t,mu,sigma,q,coef);
clear options
options.isAccelerated = true;

% 95% Coverage interval CI
CI_Low = cf2QF_GPA(cf,0.025,options);
CI_Upp = cf2QF_GPA(cf,0.975,options);
CI     = [CI_Low, CI_Upp];
disp (CI)

% CDF values at specified arguments of y
y      = [1e+10 1e+20 1e+30 1e+40 1e+50 1e+60 1e+70 1e+80 1e+90]';
cdf    = cf2CDF_GPA(cf,y,options);
Table  = table(y,cdf)

      y          cdf
   --------   ---------
    1e+10      0.87974
    1e+20      0.96421
    1e+30      0.98935
    1e+40      0.99683
    1e+50      0.99906
    1e+60      0.99972
    1e+70      0.99992
    1e+80      0.99998
    1e+90      0.99999
```

Figure 4. Characteristic Function Approach based on using the MATLAB algorithms from **CharFunTool**. CDF values and the 95% coverage interval computed from the characteristic function of a linear combination of q-Gaussian RVs, $Y = \frac{1}{3}X_1 + \frac{1}{3}X_2 + \frac{1}{3}X_3$, where $X_1 \sim TQG(0,1,0)$, $X_2 \sim TQG(0,0.5,1)$ and $X_3 \sim TQG(0,1,2.9)$ are independent random variables. The exact 95% coverage interval derived via CFA is $CI_{CFA} = [-9.1540 \times 10^{22}, 9.1540 \times 10^{22}]$.

This example highlights the importance of having an exact method such as CFA for extreme cases, where MCM may require an impractically large number of realizations to achieve reasonable accuracy. Additionally, the computational cost of using Monte Carlo simulations can be significantly higher compared to the CFA. The accuracy of the Monte Carlo method is highly dependent on the number of realizations, making it computationally

infeasible in cases with extremely heavy tailed distributions of the input quantities to obtain a comparably accurate result in a reasonable amount of time. In contrast, the CFA in combination with appropriate numerical algorithms can provide exact results in a numerical sense with relatively low computational cost, making it a preferred choice for many practical applications.

To compare the results, we used MCM with (6), (5) and (8) to derive the result by generating $N = 10^5$ realizations of the X_i to obtain N realizations of Y. By sorting the realizations of Y, we obtained $\text{CI}_{MCM} = [-1.6820 \times 10^{23}, 1.1835 \times 10^{23}]$. These results indicate that $N = 10^5$ realizations of the random variable Y are insufficient to ensure accuracy. To obtain a result comparably accurate to CFA, a larger number of realizations is required.

For illustration, we increased the number of Monte Carlo simulations one thousand times to $N = 10^8$ realizations and we obtained $\text{CI}_{MCM} = [-9.1671 \times 10^{22}, 9.2231 \times 10^{22}]$. Now, this result is comparable to the CFA, but still different and the computation took 67.5 s, whereas it took 0.006 s to obtain the CFA result using the same standard laptop computer.

Using the CFA in combination with the **cf2CDF_GPA** and **cf2QF_GPA** algorithms provides exact results in a numerical sense. This means that the values obtained are numerically correct and the numerical error can be controlled by appropriately setting the control parameters in the options settings. By using these algorithms, the CDF values and quantiles required to construct a 95% coverage interval can be accurately calculated.

6. Conclusions

The Tsallis q-Gaussian distribution is a flexible and versatile generalization of the standard Gaussian distribution that can effectively model input quantities in a wide range of applications and measurement models. Its unique properties, such as a finite support for $q < 1$ and an infinite variance for $q \geq 5/3$, make it particularly useful for modeling specific quantities with a known range or extreme values.

It is important to note that in such situations (with $q \geq 5/3$), the standard uncertainty analysis as specified in GUM may not be applicable, as it implicitly assumes the existence of at least two first moments (mean and variance) of the distributions associated with the input quantities. Therefore, alternative methods are needed to adequately characterize the uncertainty of the output quantities.

This paper presents a new and original contribution by providing the explicit form of the characteristic function of q-Gaussian random variables and their linear combinations, which can be used in combination with the Characteristic Function Approach for efficient and exact uncertainty analysis.

The CFA is a powerful tool for deriving the CDF, PDF and QF of the measured quantity, which are essential for specifying measurement uncertainty, in a sense as specified in the Supplements of the GUM and constructing coverage intervals. The CFA offers several advantages, including accuracy, flexibility, efficiency and repeatability.

To perform these calculations, the MATLAB toolbox **CharFunTool** is suggested, which includes implemented inversion algorithms such as **cf2DistGP** or **cf2DistGPA**. These algorithms employ the Gil-Pelaez inversion formulae and the adaptive Gauss–Kronrod quadrature rule for numerical integration of the oscillatory integrand function. Additionally, convergence acceleration techniques are employed to compute the limit of the alternating series.

We note that finding an explicit expression for the characteristic function of the output quantity and numerically inverting it can be challenging, especially when deriving the joint characteristic function of a multivariate distribution with stochastically dependent input quantities. In such situations, we suggest considering alternative approaches, such as those based on applying the proper Bayesian approach.

References

1. JCGM 101:2008 (GUM S1). Evaluation of Measurement Data–Supplement 1 to the Guide to the Expression of Uncertainty in Measurement——Propagation of Distributions Using a Monte Carlo Method. ISO, BIPM, IEC, IFCC, ILAC, IUPAC, IUPAP and OIML. 2008. Available online: https://www.bipm.org/documents/20126/2071204/JCGM_101_2008_E.pdf (accessed on 1 April 2023).
2. JCGM 100:2008 (GUM). Evaluation of Measurement Data–Guide to the Expression of Uncertainty in Measurement (GUM 1995 with Minor Corrections), ISO, BIPM, IEC, IFCC, ILAC, IUPAC, IUPAP and OIML. 2008. Available online: https://ncc.nesdis.noaa.gov/documents/documentation/JCGM_100_2008_E.pdf (accessed on 1 April 2023).
3. JCGM 102:2011 (GUM S2). Evaluation of Measurement Data—Supplement 2 to the Guide to the Expression of Uncertainty in Measurement–Extension to Any Number of Output Quantities. ISO, BIPM, IEC, IFCC, ILAC, IUPAC, IUPAP and OIML. 2011. Available online: https://www.bipm.org/documents/20126/2071204/JCGM_102_2011_E.pdf (accessed on 1 April 2023).
4. Tsallis, C. Possible generalization of Boltzmann-Gibbs statistics. *J. Stat. Phys.* **1988**, *52*, 479–487. [CrossRef]
5. Tsallis, C. *Introduction to Nonextensive Statistical Mechanics: Approaching a Complex World*; Springer Science & Business Media: Berlin/Heidelberg, Germany, 2009; Volume 1.
6. Vignat, C.; Plastino, A. Why is the detection of q-Gaussian behavior such a common occurrence? *Phys. A Stat. Mech. Its Appl.* **2009**, *388*, 601–608. [CrossRef]
7. Beck, C.; Cohen, E.G.; Swinney, H.L. From time series to superstatistics. *Phys. Rev. E* **2005**, *72*, 056133. [CrossRef] [PubMed]
8. Borl, L.; Bouchaud, J.P.; Muzy, J.F.; Zumbach, G. The dynamics of financial markets—Mandelbrot's multifractal cascades and beyond. *Wilmott Mag.* **2005**, *3*, 40–53.
9. Carpena, P.; Bernaola-Galván, P.; Hackenberg, M.; Coronado, A.V.; Oliver, J.L. Level statistics of words: Finding keywords in literary texts and DNA. *Phys. Rev.* **2009**, *79*, 035102.
10. Burlaga, L.F.; Viñas, A.F. Triangle for the entropic index q of nonextensive statistical mechanics observed by Voyager 1 in the distant heliosphere. *Phys. Rev. Lett.* **2004**, *93*, 251101.
11. Anteneodo, C. Non-extensive random walks. *Phys. A Stat. Mech. Its Appl.* **2005**, *358*, 289–298. [CrossRef]
12. Witkovský, V.; Wimmer, G.; Ďurišová, Z.; Durixsx, S.; Palenčár, R. Brief overview of methods for measurement uncertainty analysis: GUM uncertainty framework, Monte Carlo method, characteristic function approach. In Proceedings of the 11th International Conference on Measurement, Smolenice, Slovakia, 29–31 May 2017; pp. 35–38.
13. Witkovský, V. Numerical inversion of a characteristic function: An alternative tool to form the probability distribution of output quantity in linear measurement models. *Acta IMEKO* **2016**, *5*, 32–44. [CrossRef]
14. Witkovský, V. CharFunTool: The Characteristic Functions Toolbox. 2023. Available online: https://github.com/witkovsky/CharFunTool (accessed on 1 April 2023).
15. Shephard, N.G. Numerical integration rules for multivariate inversions. *J. Stat. Comput. Simul.* **1991**, *39*, 37–46. [CrossRef]
16. Davies, R.B. The distribution of a linear combination of χ^2 random variables. *J. R. Stat. Soc. Ser. Appl. Stat.* **1980**, *29*, 323–333.
17. Witkovský, V. On the exact computation of the density and of the quantiles of linear combinations of t and F random variables. *J. Stat. Plan. Inference* **2001**, *94*, 1–13. [CrossRef]
18. Shephard, N.G. From characteristic function to distribution function: A simple framework for the theory. *Econom. Theory* **1991**, *7*, 519–529. [CrossRef]
19. Wendel, J.G. The non-absolute convergence of Gil-Pelaez inversion integral. *Ann. Math. Stat.* **1961**, *32*, 338–339. [CrossRef]
20. Gil-Pelaez, J. Note on the inversion theorem. *Biometrika* **1951**, *38*, 481–482. [CrossRef]
21. Abate, J.; Valkó, P.P. Multi-precision Laplace transform inversion. *Int. J. Numer. Methods Eng.* **2004**, *60*, 979–993. [CrossRef]
22. Bakhvalov, N.S.; Vasileva, L.G. Evaluation of the integrals of oscillating functions by interpolation at nodes of Gaussian quadratures. *USSR Comput. Math. Math. Phys.* **1968**, *8*, 241–249. [CrossRef]
23. Chourdakis, K. Option pricing using the fractional FFT. *J. Comput. Financ.* **2005**, *8*, 1–18. [CrossRef]
24. Cohen, H.; Villegas, F.R.; Zagier, D. Convergence acceleration of alternating series. *Exp. Math.* **2000**, *9*, 3–12. [CrossRef]
25. Davies, R.B. Numerical inversion of a characteristic function. *Biometrika* **1973**, *60*, 415–417. [CrossRef]
26. Evans, G.A.; Webster, J.R. A comparison of some methods for the evaluation of highly oscillatory integrals. *J. Comput. Appl. Math.* **1999**, *112*, 55–69. [CrossRef]
27. Hürlimann, W. Improved FFT approximations of probability functions based on modified quadrature rules. *Int. Math. Forum* **2013**, *8*, 829–840. [CrossRef]
28. Imhof, J.P. Computing the distribution of quadratic forms in normal variables. *Biometrika* **1961**, *48*, 419–426. [CrossRef]

29. Mijanović, A.; Popović, B.V.; Witkovský, V. A numerical inversion of the bivariate characteristic function. *Appl. Math. Comput.* **2023**, *443*, 127807. [CrossRef]
30. Sidi, A. A user-friendly extrapolation method for computing infinite range integrals of products of oscillatory functions. *IMA J. Numer. Anal.* **2012**, *32*, 602–631. [CrossRef]
31. Talbot, A. The accurate numerical inversion of Laplace transforms. *IMA J. Appl. Math.* **1979**, *23*, 97–120. [CrossRef]
32. Waller, L.A.; Turnbull, B.W.; Hardin, J.M. Obtaining distribution functions by numerical inversion of characteristic functions with applications. *Am. Stat.* **1995**, *49*, 346–350.