IOPscience iopscience.iop.org

Home Search Collections Journals About Contact us My IOPscience

Central limit theorem and deformed exponentials

This article has been downloaded from IOPscience. Please scroll down to see the full text article.

2007 J. Phys. A: Math. Theor. 40 F969

(http://iopscience.iop.org/1751-8121/40/45/F02)

View the table of contents for this issue, or go to the journal homepage for more

Download details:
IP Address: 171.66.16.146
The article was downloaded on 03/06/2010 at 06:24

Please note that terms and conditions apply.

IOP PUBLISHING JOURNAL OF PHYSICS A: MATHEMATICAL AND THEORETICAL

J. Phys. A: Math. Theor. **40** (2007) F969–F978 doi:10.1088/1751-8113/40/45/F02

FAST TRACK COMMUNICATION

Central limit theorem and deformed exponentials

C Vignat[1] **and A Plastino**[2]

[1] LTHI, EPFL, Lausanne, Switzerland
[2] Facultad de Ciencias Exactas, Universidad Nacional de La Plata and CONICET, CC 727, 1900 La Plata, Argentina

E-mail: vignat@univ-mlv.fr and plastino@uolsinectis.com.ar

Received 11 September 2007, in final form 27 September 2007
Published 23 October 2007
Online at stacks.iop.org/JPhysA/40/F969

Abstract
The central limit theorem (CLT) can be ranked among the most important ones in probability theory and statistics and plays an essential role in several basic and applied disciplines, notably in statistical thermodynamics. We show that there exists a natural extension of the CLT from exponentials to so-called deformed exponentials (also denoted as q-Gaussians). Our proposal applies exactly in the usual conditions in which the classical CLT is used.

PACS numbers: 05.40.−a, 05.20.Gg, 02.50.−r

1. Introduction

The central limit theorems (CLT) can be ranked among the most important ones in probability theory and statistics and play an essential role in several basic and applied disciplines, notably in statistical mechanics. Pioneers like A de Moivre, P S de Laplace, S D Poisson and C F Gauss have shown that the Gaussian function is the attractor of independent systems with a finite second variance. Distinguished authors like Chebyshev, Markov, Liapounov, Feller, Lindeberg and Levy have also made essential contributions to the CLT development. As far as physics is concerned one can state that, starting from any system, with any distribution function (for some measurable quantity x), and combining a sufficiently large number of such systems together, the resultant distribution function (for x) is always Gaussian. This proposition derives from the central limit theorem. Thus, for physics the CLT is one of the most important theorems in the whole of mathematics since it guarantees that the probability distribution (PD) of any measurable quantity is Gaussian, provided that a sufficiently large number of statistically independent observations are made. One can, therefore, confidently predict that Gaussian distributions are going to crop up all over the place in statistical thermodynamics. An interesting physical question, to be addressed *here*, emerges naturally: what if for some *other* PD a CLT ensues? Will such a distribution crop up all over the place as well? The answer is a resounding yes. Instead of exponentials we will have q-exponentials [1], widely

© 2007 IOP Publishing Ltd Printed in the UK

employed nowadays in several scientific disciplines. The form of these functions is given below.

2. Preliminaries

We posed a question above that was first successfully addressed by Umarov *et al* [2]. However, they needed complicated conditions, such as q-independence and nonlinear Fourier transforms, that we will avoid in what follows by using a quite different approach.

The random variables to which the classical CLT refers are required to be independent. Subsequent efforts along CLT lines have established corresponding theorems for weakly dependent random variables as well (see some pertinent and important references in [2–4]). The CLT does not hold if correlations between far-ranging random variables are not negligible (see [5]).

Recent developments in statistical mechanics that have attracted the attention of many researches do deal with strongly correlated random variables ([1] and references therein). These correlations do not rapidly decrease with any increasing *distance* between random variables and are often referred to as global correlations (see [6] for a definition). Is there an attractor that would replace the Gaussians in such a case?.

The answer is in the affirmative, as shown in [2–4], with the deformed or q-Gaussian playing the starring role. It is asserted in [3] that such a theorem cannot be obtained if we rely on classic algebra. It needs a construction based on a special algebra, which is called q-mathematics. *The goal of this communication* is to show that a q-generalization of the central limit theorem becomes indeed possible (and is relatively simple) without recourse to q-mathematics.

2.1. Systems that are q-distributed

Consider a system $\mathcal{S}$ described by a d-component vector X whose covariance matrix reads

$$K = \langle XX^t \rangle \equiv EXX^t, \tag{1}$$

the superscript t indicating transposition. We say that X is q-Gaussian (or deformed Gaussian) distributed if its probability distribution function writes as described by equations (2) or (7):

- In the case $1 < q < \frac{d+4}{d+2}$

  $$f_{X,q}(X) = A_q (1 + X^t \Lambda^{-1} X)^{\frac{1}{1-q}}, \tag{2}$$

  with matrix Λ being related to K in the fashion [7]

  $$\Lambda = (m-2)K, \tag{3}$$

  where the number of degrees of freedom m is defined in terms of the number of X-components d as [7]

  $$m = \frac{2}{q-1} - d. \tag{4}$$

  Moreover, the so-called partition function $Z_q = 1/A_q$ reads [7]

  $$Z_q = \frac{\Gamma\left(\frac{1}{q-1} - \frac{d}{2}\right) |\pi\Lambda|^{1/2}}{\Gamma\left(\frac{1}{q-1}\right)}, \tag{5}$$

  and the characteristic function is

  $$\varphi_X(U) = \frac{2^{1-\frac{m}{2}}}{\Gamma\left(\frac{m}{2}\right)} z^{\frac{m}{2}} K_{\frac{m}{2}}(z), \tag{6}$$

with $z = \sqrt{U^t \Lambda U}$ and $K_{m/2}$ the modified Bessel function of the second kind.

- In the case $q < 1$

$$f_{X,q}(X) = A_q(1 - X^t \Sigma^{-1} X)_+^{\frac{1}{1-q}}, \tag{7}$$

where the matrix Σ is related to the covariance one (1) via $\Sigma = dK$. We need here a parameter p (see below) defined as

$$p = 2\frac{2-q}{1-q} + d, \tag{8}$$

so that the partition reads

$$Z_q = \frac{\Gamma\left(\frac{2-q}{1-q}\right)|\pi\Sigma|^{1/2}}{\Gamma\left(\frac{2-q}{q-1} + \frac{d}{2}\right)}, \tag{9}$$

and the characteristic function is

$$\varphi_X(U) = 2^{\frac{p}{2}-1}\Gamma\left(\frac{p}{2}\right)\frac{J_{\frac{p}{2}-1}(z)}{z^{\frac{p}{2}-1}}, \tag{10}$$

where $z = \sqrt{U^t \Sigma U}$ and $J_{p/2-1}$ is the Bessel function of the first kind.

3. Our particular road towards a new CLT

As stated above, several attempts to generalize the central limit theorem (CLT) so that the Gaussian is replaced as the attractor by the q-Gaussian have been published recently [2–4]. We recall here a basic multivariate version of the conventional CLT.

Theorem 3.1. *Let $X_1, X_2, \ldots$ be independent and identically distributed (i.i.d.) random vectors in $\mathbb{R}^d$ with $E[X_i] = 0$ and $E\left[X_i X_i^t\right] = K$ and let*

$$W_n = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} X_i. \tag{11}$$

One has[3]

$$\forall\, \mathbf{t} \in \mathbb{R}^d, \qquad \lim_{n\to+\infty} \Pr\{W_n \leqslant \mathbf{t}\} = \Phi_1(\mathbf{t}) = \frac{1}{|2\pi K|^{1/2}} \int_{-\infty}^{t_1} \cdots \int_{-\infty}^{t_d} \mathrm{e}^{-\frac{X^t K^{-1} X}{2}}\, \mathrm{d}X.$$

The basic idea towards non-conventional CLTs is to find conditions under which convergence to the usual normal distribution Φ_1 with covariance matrix K is replaced by convergence to a q-Gaussian distribution

$$\Phi_q(\mathbf{t}) = \int_{-\infty}^{t_1} \cdots \int_{-\infty}^{t_d} f_{X,q}(X)\, \mathrm{d}X_1 \cdots \mathrm{d}X_d \tag{12}$$

with, for $q > 1$, $f_{X,q}$ as defined in (2) and parameter m defined by (4) or, for $q < 1$, $f_{X,q}$ as defined in (7) and parameter p defined by (8), which play the CLT-attractor role of Gaussian distributions. We note that both cases $m \to +\infty$ and $p \to +\infty$ correspond to convergence $q \to 1$ to the Gaussian case.

In two recent contributions [2, 3], S Umarov and C Tsallis highlight the existence of such a multivariate central limit theorem, provided there exists a certain kind of dependence, called

[3] Note that inequality between vectors $W_n \leqslant \mathbf{t}$ denotes the set of d componentwise inequalities $\{W_n(k) \leqslant t_k;\, 1 \leqslant k \leqslant d\}$.

q-independence, between random vectors X_i. This q-independence can be characterized in terms of the q-Fourier transform F_q and of the q-product $\otimes_q$ as

$$F_q[X_1 + X_2] = F_q[X_1] \otimes_q F_q[X_2]$$

and reduces to the classical independence for $q = 1$.

We recall that the q-product of x and y is defined, for $x^{1-q} + y^{1-q} \geqslant 1$, as

$$x \otimes_q y = (x^{1-q} + y^{1-q} - 1)^{\frac{1}{1-q}}$$

and the q-Fourier transform of a function $f(x)$, $x \in \mathbb{R}^d$, is

$$F_q[f](\xi) = \int_{\mathbb{R}^d} (f^{1-q}(x) + (1-q)\mathrm{i}x^t\xi)^{\frac{1}{1-q}}\,\mathrm{d}X.$$

For the conditions of existence of this q-Fourier transform, see [3, corollary 2.4]. This approach suffers from the lack of physical interpretation for such special dependence; moreover, the q-Fourier transform is a nonlinear transform (unless $q = 1$) what makes it rather awkward to use.

Another approach, as described in [8], consists in keeping the independence assumption between vectors X_i while replacing the n terms in (11) by a random number $N(n)$ of terms. That is, if the random variable $N(n)$ follows a negative binomial then convergence to a q-Gaussian distribution occurs whenever convergence occurs in the usual sense.

In the present contribution we show that there exists a much more natural way (that applies for instance to the case of fluctuating temperatures) to extend the CLT based on the Beck–Cohen notion of superstatistics [10] (see the discussion in [11]). Our 'starting point' is the same as that in Umarov's approach (i.e., assuming some kind of dependence between the summed terms). However, the manner in which we introduce this dependence among data is a natural one that can be given several physical interpretations.

4. Present results

Our present results can be conveniently condensed by stating two theorems, according to the q-value. The distinction before these two cases is usual in the theory of nonextensivity; it is due to the fundamentally different behaviour of the q-Gaussian distributions, which are Gaussian scale mixtures with unbounded support for $q > 1$, contrarily to the case $q < 1$. The extended central limit theorems we are here advancing are given below.

The essential idea in our approach is that of suitably introducing a random variable a that is chi distributed with m degrees of freedom and then constructing the 'scale mixtures' (typical of superstatistics [11])

$$Z_n = \frac{1}{a\sqrt{n}} \sum_{i=1}^{n} X_i. \tag{13}$$

4.1. The case $q > 1$

Theorem 4.1. *If $X_1, X_2, \ldots$ are i.i.d. random vectors in $\mathbb{R}^d$ with zero mean and covariance matrix K, and if a denotes a random variable chi distributed with m degrees of freedom, scale parameter $(m-2)^{-1/2}$, and independent of X_i,* then *random vectors*

$$Z_n = \frac{1}{a\sqrt{n}} \sum_{i=1}^{n} X_i \tag{14}$$

converge weakly to a multivariate q-Gaussian vector Z with covariance matrix K. Equivalently stated

$$\forall \mathbf{t} \in \mathbb{R}^d, \qquad \lim_{n\to+\infty} \Pr\{Z_n \leqslant \mathbf{t}\} = \Phi_q(\mathbf{t}); \tag{15}$$

with characteristic distribution function (cdf) $\Phi_q(\mathbf{t})$ defined as in (12). Moreover,

$$q = \frac{m+d+2}{m+d}. \tag{16}$$

Proof. First we note that the χ-density with m degrees of freedom and scale parameter $\frac{1}{\sqrt{m-2}}$ is

$$f_a(a) = \frac{2^{1-\frac{m}{2}}(m-2)^{\frac{m}{2}}}{\Gamma\left(\frac{m}{2}\right)} a^{m-1}\, \mathrm{e}^{-\frac{a^2(m-2)}{2}}.$$

Now, by the multivariate central limit theorem that we have remembered above (note that, below, symbol $\Rightarrow$ denotes weak convergence)

$$\frac{1}{\sqrt{n}} \sum_{i=1}^{n} X_i \Rightarrow N$$

where N is a normal vector in $\mathbb{R}^d$ with covariance matrix K. Applying from [12] its result (theorem 2.8) we immediately deduce that

$$Z_n \Rightarrow \frac{N}{a}$$

follows a q-Gaussian distribution with covariance matrix K and parameter q defined by (4). □

4.2. The case $q < 1$

Extension to the case $q < 1$ proceeds as follows.

Theorem 4.2. *If $X_1, X_2, \ldots$ are i.i.d. random vectors in $\mathbb{R}^d$ with zero mean and covariance matrix K, and if a is a random variable independent of X_i that is chi distributed with m degrees of freedom and scale parameter $\sqrt{m-2}$,* then *the random vectors*

$$Y_n = \frac{\frac{1}{\sqrt{n}} \sum_{i=1}^{n} X_i}{\sqrt{a^2 + \left(\frac{1}{\sqrt{n}} \sum_{i=1}^{n} X_i\right)^t K^{-1} \left(\frac{1}{\sqrt{n}} \sum_{i=1}^{n} X_i\right)}} \tag{17}$$

converge weakly to a multivariate q-Gaussian vector Y with covariance matrix K and distribution function given by (12). Moreover,

$$q = \frac{m-4}{m-2} < 1. \tag{18}$$

Proof. If Z has cdf given by (12), then [8]

$$Y = \phi(Z) = \frac{Z}{\sqrt{1 + Z^t K^{-1} Z}}$$

has cdf given by (12). Since the function $\phi = \mathbb{R}^d \to \mathbb{R}^d$ is continuous, the desired result is deduced by the application of the continuous mapping theorem (see [13], theorem 2.3, p 7). □

Remark. We note that Y_n in (17) is a normalized version of sum (14); however, the random fluctuation term a is replaced by a fluctuating term

$$b = \sqrt{a^2 + \left(\frac{1}{\sqrt{n}}\sum_{i=1}^{n} X_i\right)^t K^{-1} \left(\frac{1}{\sqrt{n}}\sum_{i=1}^{n} X_i\right)}$$

that involves the value of the sum itself—and thus is not independent of this sum anymore. Thus, the case $q < 1$ can be considered as a fluctuating version of the usual CLT for which the fluctuation depends on the state of the system. Moreover, it is intuitive that as n grows, the distribution of the fluctuation b gets closer to a chi distribution with $m + n$ degrees of freedom. At last, a geometric interpretation of this fluctuation term can be given, see [9].

4.3. Link with q-independence

Although the extensions of the CLT proposed above differ from those developed by [2], a link can be established between both approaches, *for large values of* n, as follows.

Theorem 4.3 (linking theorem). *Consider $n = 2n_0$ with n_0 large and divide the sum Z_n in (14) into two parts as*

$$Z_n = \frac{1}{a\sqrt{n}}\left(\sum_{i=1}^{n_0} X_i + \sum_{i=n_0+1}^{n} X_i\right) = Z_n^{(1)} + Z_n^{(2)}. \quad (19)$$

If the characteristic function ϕ of X_i is such that $\int_{\mathbb{R}^d} |\phi|^\nu \mathrm{d}t < \infty$ for some $\nu \geqslant 1$, then random vectors $Z_n^{(1)}$ and $Z_n^{(2)}$ are almost q-independent in the sense that

$$\forall \epsilon > 0, \qquad \exists N \; s.t. \; n > N \Rightarrow \left\| F_q\left(Z_n^{(1)} + Z_n^{(2)}\right) - F_q\left(Z_n^{(1)}\right) \otimes_q F_q\left(Z_n^{(2)}\right)\right\|_\infty < \epsilon.$$

For didactic reasons we postpone the proof of this result to the next section. We deduce from it that, asymptotically, the CLT theorem 4.1 exactly generates the q-independence condition required for the application of the particular CLT version proposed in [2, 3].

5. Proof of the linking theorem

5.1. Introduction

In order to simplify the proof we will assume that vectors X_i verify a stronger CLT version than that stated in theorem 3.1, which might be called a 'CLT in total variation'. Now, the 'total variation' divergence between two probability densities f and g is

$$d_{\mathrm{TV}}(f, g) = \frac{1}{2}\int_{\mathbb{R}^d} |f - g|. \quad (20)$$

If U and V are random vectors distributed according to f and g, respectively, we will denote

$$d_{\mathrm{TV}}(U, V) = d_{\mathrm{TV}}(f, g).$$

The total variation version of the CLT writes as follows (see [13], theorem 2.31).

Theorem 5.1. *Assume that $X_1, X_2, \ldots$ are i.i.d random vectors of $\mathbb{R}^d$ with zero mean, finite covariance matrix K and characteristic function ϕ such that $\int |\phi|^\nu \, \mathrm{d}t < \infty$ for some $\nu \geqslant 1$. If $Z_n = \frac{1}{\sqrt{n}}\sum_{i=1}^{n} X_i$ and Z is a normal vector in $\mathbb{R}^d$ with covariance matrix K then*

$$\lim_{n\to+\infty} d_{\mathrm{TV}}(Z_n, Z) = 0. \quad (21)$$

Let us introduce the following notations: $\tilde{Z}_n$ denotes a version of sum (14) where all X_i are replaced by i.i.d. Gaussian vectors $N_i \in \mathbb{R}^d$ with covariance matrix K:

$$\tilde{Z}_n = \frac{1}{a\sqrt{n}}\left(\sum_{i=1}^{n_0} N_i + \sum_{i=n_0+1}^{n} N_i\right) = \tilde{Z}_n^{(1)} + \tilde{Z}_n^{(2)}.$$

The proof of theorem 4.3 is based on the fact that $\tilde{Z}_n^{(1)}$ and $\tilde{Z}_n^{(2)}$ are exactly q-independent (see theorem 5.2 below). Since n is large, according to the above 'total variation' CLT, $Z_n^{(1)}$ and $Z_n^{(2)}$ are close to their q-Gaussian counterparts $\tilde{Z}_n^{(1)}$ and $\tilde{Z}_n^{(2)}$, respectively. It remains to check that the closeness between these vectors can be stated in terms of their q-transforms. We proceed in five steps that invoke technical lemmas that are the subject of subsection 5.3:

- *Step* 1. Components $\tilde{Z}_n^{(1)}$ and $\tilde{Z}_n^{(2)}$ are exactly q-independent, as will be proved in subsection 5.2.
- *Step* 2. Let us fix $\epsilon > 0$ and consider
$$\left\|F_q\left(Z_n^{(1)} + Z_n^{(2)}\right) - F_q\left(Z_n^{(1)}\right) \otimes_q F_q\left(Z_n^{(2)}\right)\right\|_\infty \leqslant \left\|F_q\left(Z_n^{(1)} + Z_n^{(2)}\right) - F_q\left(\tilde{Z}_n^{(1)} + \tilde{Z}_n^{(2)}\right)\right\|_\infty + \left\|F_q\left(\tilde{Z}_n^{(1)}\right) \otimes_q F_q\left(\tilde{Z}_n^{(2)}\right) - F_q\left(Z_n^{(1)}\right) \otimes_q F_q\left(Z_n^{(2)}\right)\right\|_\infty.$$
- *Step* 3. The first term $\left\|F_q\left[Z_n^{(1)} + Z_n^{(2)}\right] - F_q\left[\tilde{Z}_n^{(1)} + \tilde{Z}_n^{(2)}\right]\right\|_\infty = \|F_q[Z_n] - F_q[\tilde{Z}_n]\|_\infty$ can be bounded as follows
$$\|F_q[Z_n] - F_q[\tilde{Z}_n]\|_\infty \leqslant 2d_{\mathrm{TV}}(Z_n, \tilde{Z}_n) \leqslant 2d_{\mathrm{TV}}(X_n, \tilde{X}_n)$$
where the first inequality follows from lemma 5.3 and the second one from lemma 5.1 below. Thus a value N_1 can be chosen so that $n_0 > N_1$ and $n_1 > N_1$ ensure that this term is smaller than $\frac{\epsilon}{2}$.
- *Step* 4. The second term $\left\|F_q\left[\tilde{Z}_n^{(1)}\right] \otimes_{q_1} F_q\left[\tilde{Z}_n^{(2)}\right] - F_q\left[Z_n^{(1)}\right] \otimes_{q_1} F_q\left[Z_n^{(2)}\right]\right\|_\infty$ can be bounded by applying lemma 5.4 for a large enough value of $n = n_0 + n_1$, say $n > N_2$, we have
$$\left\|F_q\left[\tilde{Z}_n^{(1)}\right] \otimes_{q_1} F_q\left[\tilde{Z}_n^{(2)}\right] - F_q\left[Z_n^{(1)}\right] \otimes_{q_1} F_q\left[Z_n^{(2)}\right]\right\|_\infty \leqslant 2d_{\mathrm{TV}}\left(Z_n^{(1)}, \tilde{Z}_n^{(1)}\right) + 2d_{\mathrm{TV}}\left(Z_n^{(2)}, \tilde{Z}_n^{(2)}\right).$$
Finally, from the total variation CLT, there exists a value N_3 such that $n_0 > N_3$ and $n_1 > N_3$ implies that each of both total variation divergences is smaller than $\frac{\epsilon}{4}$.
- *Step* 5. The consideration of $N = \max(N_1, N_2, N_3)$ is then seen to prove the linking theorem 5.3.

5.2. Components of q-Gaussian vectors are q-independent

We first begin to check that 'sub-vectors' extracted from q-Gaussian vectors are exactly q-independent; this result is obvious from the fact that, by the CLT given in [2] (theorem 3.1), these sub-vectors can be considered as limit cases of sequences of q-independent sequences. However, the mathematical verification of this property is of an instructive nature and we proceed to give it. For readability, we will say that $X \sim (q, d)$ if X is a q-Gaussian vector of dimension d and nonextensivity parameter q.

Theorem 5.2. *If* $1 < q_0 < 1 + \frac{2}{d}$ *and vector* $X = \left[X_1^t, X_2^t\right]^t \sim (q_0, 2d)$ *with parameter* $q_0 > 1$ *then vectors* $X_1 \sim (q, d)$ *and* $X_2 \sim (q, d)$ *and they are q-independent:*

$$F_q[X_1 + X_2] = F_q[X_1] \otimes_{q_1} F_q[X_2] \tag{22}$$

with $q = z(q_0) = \frac{2q_0 + d(1-q_0)}{2 + d(1-q_0)} > 1$ *and* $q_1 = z(q) > 1$.

Proof. Since $X_1 \sim (q, d)$, we know from corollary 2.3 of [2] that $F_q[X_1] \sim (q_1, d)$. Moreover, since X_1 and X_2 are the components of the same q-Gaussian vector, from [8] we deduce that $X_1 + X_2 \sim (q, d)$ so that $F_q[X_1 + X_2] \sim (q_1, d)$. Finally, it is easy to check that since $F_q[X_1] \sim (q_1, d)$ and $F_q[X_2] \sim (q_1, d)$ then $F_q[X_1] \otimes_{q_1} F_q[X_2] \sim (q_1, d)$. The fact that both terms have same covariance matrices is straightforward, what proves the result. □

We note that q-correlation (22) corresponds to q-independence of the third kind as listed in table 1 of [2]. We pass now to the consideration of the four lemmas invoked in the proof of the linking theorem.

5.3. Technical lemmas

As we are concerned with scale mixtures of Gaussian vectors, we need the following lemmas.

Lemma 5.1. *If U and V are random vectors in $\mathbb{R}^d$ and a is a random variable independent of U and V then*

$$d_{\mathrm{TV}}\left(\frac{U}{a}, \frac{V}{a}\right) \leqslant d_{\mathrm{TV}}(U, V). \tag{23}$$

Proof. The distributions of scale mixtures U/a and V/a write, in terms of the distributions of U and V, according to

$$f_{U/a}(x) = \int_{\mathbb{R}^+} \frac{1}{a^d} f_a(a) f_U\left(\frac{x}{a}\right) \mathrm{d}a, \qquad g_{V/a}(x) = \int_{\mathbb{R}^+} \frac{1}{a^d} f_a(a) f_V\left(\frac{x}{a}\right) \mathrm{d}a.$$

It follows that

$$\begin{aligned}
d_{\mathrm{TV}}\left(\frac{U}{a}, \frac{V}{a}\right) &= \frac{1}{2}\int_{\mathbb{R}^d} |f_{U/a}(x) - f_{V/a}(x)|\,\mathrm{d}X \\
&= \frac{1}{2}\int_{\mathbb{R}^d} \left|\int_{\mathbb{R}^+} \frac{1}{a^d} f_a(a)\left(f_U\left(\frac{x}{a}\right) - f_V\left(\frac{x}{a}\right)\right)\mathrm{d}a\right| \mathrm{d}X \\
&\leqslant \frac{1}{2}\int_{\mathbb{R}^d}\int_{\mathbb{R}^+} \frac{1}{a^d} f_a(a)\left|f_U\left(\frac{x}{a}\right) - f_V\left(\frac{x}{a}\right)\right| \mathrm{d}a\,\mathrm{d}X \\
&= \frac{1}{2}\int_{\mathbb{R}^+} \frac{1}{a^d} f_a(a)\,\mathrm{d}a \int_{\mathbb{R}^d} |f_U(z) - f_V(z)| a^d\,\mathrm{d}z \\
&= \frac{1}{2}\int_{\mathbb{R}^+} f_a(a)\,\mathrm{d}a \int_{\mathbb{R}^d} |f_U(z) - f_V(z)|\,\mathrm{d}z \\
&= \frac{1}{2}\int_{\mathbb{R}^d} |f_U - f_V| = \mathrm{d}_{\mathrm{TV}}(U, V).
\end{aligned}$$

□

We will also need the following

Lemma 5.2. *For $q > 1$, the function*

$$\begin{aligned}
&\psi_{q,z} : \mathbb{R}^+ \to \mathbb{C} \\
&x \mapsto (x^{1-q} + z)^{\frac{1}{1-q}}
\end{aligned}$$

is Lipschitz, with unit constant if $\mathrm{Re}(z) \geqslant 0$.

Proof. We have

$$|\psi_{q,z}(x_1) - \psi_{q,z}(x_0)| \leqslant \sup_{x_0 \leqslant x \leqslant x_1} |\psi'_{q,z}(x)||x_1 - x_0|, \tag{24}$$

where

$$\psi'_{q,z}(x) = \frac{1}{(1 + zx^{q-1})^{\frac{q}{q-1}}}, \tag{25}$$

with $\frac{q}{q-1} > 0$, so that, for $x > 0$,

$$|\psi'_{q,z}(x)| = \frac{1}{|1 + zx^{q-1}|^{\frac{q}{q-1}}} \leqslant 1, \tag{26}$$

if $\mathrm{Re}(z) \geqslant 0$. □

Two straight consequences of this inequality are the following lemmas.

Lemma 5.3. *For any random vectors U and V, if $q \geqslant 1$, the following inequality holds:*

$$\|F_q[U] - F_q[V]\|_\infty \leqslant 2d_{\mathrm{TV}}(U, V). \tag{27}$$

Proof. Denote by f_U and f_V the respective probability densities of U and V: then $\forall\, \xi \in \mathbb{R}^d$,

$$|F_q[U](\xi) - F_q[V](\xi)| \leqslant \int_{\mathbb{R}^d} \big| \big(f_U^{1-q}(x) + (1-q)\mathrm{i}x^t\xi\big)^{\frac{1}{1-q}} - \big(f_V^{1-q}(x) + (1-q)\mathrm{i}x^t\xi\big)^{\frac{1}{1-q}} \big| \, \mathrm{d}X.$$

As $\mathrm{Re}((1-q)\mathrm{i}x^t\xi) = 0$, by lemma 5.2, the integrand is bounded by $|f_U(x) - f_V(x)|$; and since this holds $\forall \xi \in \mathbb{R}^d$, the desired result follows. □

We remark here that inequality (27) is a simple generalization of the well-known $q = 1$ case, in which $F_{q=1}$ corresponds to the classical Fourier transform. Thus a well-known result of the Fourier theory is reproduced, namely,

$$\|F_1[U] - F_1[V]\|_\infty \leqslant 2d_{\mathrm{TV}}(U, V).$$

For notational simplicity, let us denote as $Z_1 = Z_n^{(1)}$, $Z_2 = Z_n^{(2)}$, $\tilde{Z}_1 = \tilde{Z}_n^{(1)}$ and $\tilde{Z}_2 = \tilde{Z}_n^{(2)}$ those random vectors defined in part IV.A. Then, for n large enough,

$$\|F_q[Z_1](\xi) \otimes_{q_1} F_q[Z_2](\xi) - F_q[\tilde{Z}_1](\xi) \otimes_{q_1} F_q[\tilde{Z}_2](\xi)\|_\infty \leqslant 2d_{\mathrm{TV}}(Z_1, \tilde{Z}_1) + 2d_{\mathrm{TV}}(Z_2, \tilde{Z}_2).$$

Proof. For any $\xi \in \mathbb{R}^d$,

$$\begin{aligned}
&|F_q[Z_1](\xi) \otimes_{q_1} F_q[Z_2](\xi) - F_q[\tilde{Z}_1](\xi) \otimes_{q_1} F_q[\tilde{Z}_2](\xi)| \\
&\quad \leqslant \big|F_q[Z_1](\xi) \otimes_{q_1} F_q[Z_2](\xi) - F_q[\tilde{Z}_1](\xi) \otimes_{q_1} F_q[Z_2](\xi)\big| \\
&\qquad + \big|F_q[\tilde{Z}_1](\xi) \otimes_{q_1} F_q[Z_2](\xi) - F_q[\tilde{Z}_1](\xi) \otimes_{q_1} F_q[\tilde{Z}_2](\xi)\big| \\
&\quad = \big|\psi_{q_1, F_q^{1-q_1}[Z_2](\xi)-1}(F_q[Z_1](\xi)) - \psi_{q_1, F_q^{1-q_1}[Z_2](\xi)-1}(F_q[\tilde{Z}_1](\xi))\big| \\
&\qquad + \big|\psi_{q_1, F_q^{1-q_1}[\tilde{Z}_1](\xi)-1}(F_q[Z_2](\xi)) - \psi_{q_1, F_q^{1-q_1}[\tilde{Z}_1](\xi)-1}(F_q[\tilde{Z}_2](\xi))\big|.
\end{aligned}$$

Since $\tilde{Z}_2$ is q-Gaussian, and since $1 < q < 1 + \frac{2}{d}$, there exists an $\alpha_2 \geqslant 0$ (as given in equation (15) of reference [2]) such that $F_q^{1-q_1}[\tilde{Z}_2](\xi) - 1 = \alpha_2(q_1 - 1)\xi^2$ so that, since $q_1 > 0$, it follows that $F_q^{1-q_1}[\tilde{Z}_2](\xi) \geqslant 1$. From the CLT in total variation, we can choose n large enough so that $d_{\mathrm{TV}}(F_q[Z_2], F_q[\tilde{Z}_2])$ is arbitrarily small, which in turns implies, by lemma 5.3 that $|F_q[Z_2](\xi) - F_q[\tilde{Z}_2](\xi)|$ is arbitrarily small as well. By continuity of the function $x \mapsto x^{1-q_1} - 1$, and since $F_q[Z_2]$ is real-valued by the symmetry of the data, this ensures that $F_q^{1-q_1}[Z_2](\xi) - 1 \geqslant 0$. Thus, the first term can be bounded using lemma 5.3 in the fashion

$$\big|\psi_{q_1, F_q^{1-q_1}[Z_2](\xi)-1}(F_q[Z_1](\xi)) - \psi_{q_1, F_q^{1-q_1}[Z_2](\xi)-1}(F_q[\tilde{Z}_1](\xi))\big| \leqslant |F_q[\tilde{Z}_1](\xi) - F_q[Z_1](\xi)|.$$

Accordingly, since $\tilde{Z}_1$ is q-Gaussian, there exists $\alpha_1 \geqslant 0$ such that $F_q^{1-q_1}[\tilde{Z}_1](\xi) - 1 = \alpha_1(q_1 - 1)\xi^2$, hence $F_q^{1-q_1}[\tilde{Z}_1](\xi) \geqslant 1$. Recourse again to lemma 5.3 yields

$$\left|\psi_{q_1, F_q^{1-q_1}[\tilde{Z}_1](\xi)-1}(F_q[Z_2](\xi)) - \psi_{q_1, F_q^{1-q_1}[\tilde{Z}_1](\xi)-1}(F_q[\tilde{Z}_2](\xi))\right| \leqslant |F_q[\tilde{Z}_2](\xi) - F_q[Z_2](\xi)|.$$

applying now lemma 5.3 to each of both terms above yields

$$|F_q[Z_1](\xi) \otimes_{q_1} F_q[Z_2](\xi) - F_q[\tilde{Z}_1](\xi) \otimes_{q_1} F_q[\tilde{Z}_2](\xi)| \leqslant 2d_{\mathrm{TV}}(Z_1, \tilde{Z}_1) + 2d_{\mathrm{TV}}(Z_2, \tilde{Z}_2).$$

As this holds for any value of $\xi \in \mathbb{C}$, the result follows.

6. Conclusions

We have here dealt with non-conventional central limit theorems, whose attractor is a deformed or q-Gaussian. Based on the Beck–Cohen notion of superstatistics [10], with scale mixtures relating random variables à la equation (13), it has been conclusively shown, in the form of theorems, that there exists an extension of the central limit theorem (CLT) to these deformed exponentials that is derived via a quite different method than that provided by Umarov and Tsallis [2]. The latter requires a special 'q-independence condition on the data' and nonlinear Fourier transforms. We avoid theses complications entirely. Our CLT proposal applies exactly in the usual conditions in which the classical CLT is used. However, links between ours and the Umarov–Tsallis treatment have been established. Finally, we note that, in the spirit of superstatistics, convergence to distributions different from q-Gaussians can be obtained using the proposed approach by properly selecting the distribution of the fluctuation variable a in (13).

References

[1] Among literally hundreds of references see, for instance, Gell-Mann M and Tsallis C (ed) 2004 *Nonextensive Entropy: Interdisciplinary Applications* (New York: Oxford University Press)
Plastino A and Plastino A R 1999 *Braz. J. Phys.* **29** 50
Vignat C and Plastino A 2007 *Phys. Lett.* A **365** 370
[2] Umarov S and Tsallis C 2007 *Phys. Rev.* E **75** 040106 (*Preprint* cond-mat/0703533)
[3] Umarov S, Steinberg S and Tsallis C 2006 *Preprint* condmat/0603593
[4] Umarov S, Steinberg S and Tsallis C 2006 *Preprint* condmat/0606040
Umarov S, Steinberg S and Tsallis C 2006 *Preprint* condmat/0606038
[5] Dehling H G, Mikosch T and Sorensen M (ed) 2002 *Empirical Process Techniques for Dependent Data* (Boston: Birkhauser)
[6] Tsallis C, Gell-Mann M and Sato Y 2005 *Proc. Natl Acad. Sci. USA* **102** 15377
[7] Vignat C and Plastino A 2005 *Phys. Lett.* A **343** 411
[8] Vignat C and Plastino A 2007 *Phys. Lett.* A **360** 415
[9] Vignat C and Plastino A 2006 *Physica* A **365** 167
[10] Beck C and Cohen E G D 2003 *Physica* A **322** 267
[11] Vignat C, Plastino A and Plastino A R 2005 *Nuovo Cimento* B **120** 951
[12] Billingsley P 1999 *Convergence of Probability Measures (Wiley Series in Probability and Statistics)* 2nd edn
[13] van der Vaart A W 1998 *Asymptotic Statistics (Cambridge Series in Statistical and Probabilistic Mathematics)*