

This item is the archived peer-reviewed author-version of:

Stein's method and a quantitative Lindeberg CLT for the Fourier transforms of random vectors

Reference:

Berckmoes Ben, Lowen Robert, van Casteren Johannes A..- Stein's method and a quantitative Lindeberg CLT for the Fourier transforms of random vectors
Journal of mathematical analysis and applications - ISSN 0022-247X - (2015), p. 1-18
DOI: <http://dx.doi.org/doi:10.1016/j.jmaa.2015.08.040>

STEIN'S METHOD AND A QUANTITATIVE LINDBERG CLT FOR THE FOURIER TRANSFORMS OF RANDOM VECTORS

B. BERCKMOES, R. LOWEN, AND J. VAN CASTEREN

ABSTRACT. We use a multivariate version of Stein's method to establish a quantitative Lindeberg CLT for the Fourier transforms of random N -vectors. We achieve this, conceptually mainly by constructing a natural approach structure on N -random vectors overlying the topology of weak convergence, and technically mainly by deducing a specific integral representation for the Hessian matrix of a solution to the Stein equation with test function $e_t(x) = \exp\left(-i \sum_{k=1}^N t_k x_k\right)$, where $t, x \in \mathbb{R}^N$.

1. INTRODUCTION AND PRELIMINARIES

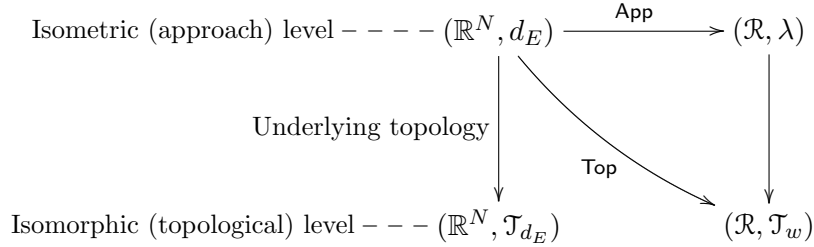
Before we start with the actual material of the paper we of course have to explain precisely what we mean when we say a “quantitative Lindeberg CLT”. A classical CLT is a limit theorem for the weak topology, and as such it is a deterministic result saying that, under given conditions, a certain sequence converges weakly to a certain limit point. Classically, what happens if the required conditions are not met is very simple: namely, we do not know that there will be convergence, and if the given conditions are both necessary and sufficient we know that there definitely will not be convergence, and the matter ends there. However, there are situations wherein the deviation from the given conditions can in a very natural way be measured numerically, in such a way that the smaller the measure of deviation is the better the conditions are approximated. This still will not help us to say anything more about the (topological) limit process and this is where index analysis in the context of approach theory comes into play ([L97], [L15]). We replace the weak topology by a canonical “weak approach structure” and apply the full machinery of approach theory. Where a topological space allows for a notion of convergence, in exactly the same structural and canonical way, an approach space allows for a notion of limit operator, whereby for any sequence and any point the limit operator gives a numerical value indicating “how far the point is away from being a limit point” of the sequence, an “index of convergence”. Of course, in order to be meaningful all this has to be such that there are natural relations between, in our case, on the one hand both the approach structure replacing the weak topology and the “measure

2000 *Mathematics Subject Classification.* 60F05.

Key words and phrases. CLT, Fourier transform, Lindeberg's condition, Stein's method, approach theory, limit operator.

- (1) Ben Berckmoes is PhD fellow at the Fund for Scientific Research of Flanders (FWO).
- (2) Corresponding author: R. Lowen, email address: bob.lowen@uantwerpen.be.

of deviation from conditions” and on the other hand both the weak topology and the conditions. The following diagram makes things more precise. Herein (\mathbb{R}^N, d_E) stands for the N -dimensional Euclidean space and metric, (\mathcal{R}, λ) stands for the derived space of \mathbb{R}^N -valued random variables defined on some probability space $(\Omega, \mathcal{A}, \mathbb{P})$ with some, yet to be defined, approach limit operator λ . The usual way would be to go straight to the topological space $(\mathcal{R}, \mathcal{T}_w)$, but instead of doing this, we equip it with an approach structure which is such that (1) the underlying topology is the weak topology \mathcal{T}_w and (2) it is obtained in a natural canonical way, as is the weak topology. Then the smaller the value of the limit operator (the index of convergence) is, the more convergence in the weak topology is being “approximated”, or the better a “virtual limit point” approximates being a real limit point in the weak topology.



The precise definition of the approach structure which we use is given in the second section.

The following simple example illustrates the meaning of limit operator and also serves to make clear to the reader that the notion of a limit operator, the index of convergence and the inequalities proven in the sequel are totally different from concepts and formulas related to the notion of rate or speed of convergence. Suppose we consider \mathbb{R} not with its usual topology but with its usual metric. This in fact is a (metric) approach space and from the general theory it follows that the associated limit operator takes the form

$$\lambda(x_n \rightarrow x) = \limsup_{n \rightarrow \infty} |x_n - x|,$$

which is a well-known expression in approximation theory (see e.g. [E72], [AMS82]). For instance if we take as sequence $x_n := (-1)^n \varepsilon$ for an arbitrary strictly positive ε then the formula yields $\lambda(x_n \rightarrow x) = |x| + \varepsilon$ and thus in this case 0 is the point which best approximates being a limit of this non-convergent sequence, with index of convergence equal to ε .

Now let us turn to the actual content of the present paper. Let ξ be a standard normally distributed random variable and $\{\xi_{n,k}\}$ a *1-dimensional standard triangular array* (1-STA), i.e. a triangular array of real random variables

$$\begin{array}{ccc}
 \xi_{1,1} & & \\
 \xi_{2,1} & \xi_{2,2} & \\
 \xi_{3,1} & \xi_{3,2} & \xi_{3,3} \\
 & \vdots & \\
 & &
 \end{array}$$

with the following properties.

(1) $\forall n : \xi_{n,1}, \dots, \xi_{n,n}$ are independent.

(2) $\forall n, k : \mathbb{E} [\xi_{n,k}] = 0$.

(3) $\forall n : \mathbb{E} [S_n^2] = 1$ with $S_n = \sum_{k=1}^n \xi_{n,k}$.

The Lindeberg CLT ([F71]) provides a useful condition under which the rowwise sums of $\{\xi_{n,k}\}$ are asymptotically normally distributed. As usual, \xrightarrow{w} stands for weak convergence.

Theorem 1.1. (Lindeberg CLT) *Suppose that $\{\xi_{n,k}\}$ satisfies Lindeberg's condition in the sense that*

$$\forall \epsilon > 0 : \sum_{k=1}^n \mathbb{E} [\xi_{n,k}^2; |\xi_{n,k}| > \epsilon] \rightarrow 0.$$

Then

$$S_n \xrightarrow{w} \xi.$$

Recall that the *Kolmogorov distance* K between random variables η and η' is defined as

$$\sup_{x \in \mathbb{R}} |F_\eta(x) - F_{\eta'}(x)|$$

where

$$F_\zeta(x) = \mathbb{E} [1_{]-\infty, x}] (\zeta) = \mathbb{P} [\zeta \leq x]$$

represents the cumulative distribution function of the random variable ζ . It is well known that K metrizes weak convergence to a continuously distributed random variable.

The following powerful result was obtained by Feller in [F68].

Theorem 1.2. *There exists a universal constant $C > 0$ such that*

$$\begin{aligned} & \sup_{x \in \mathbb{R}} |F_\xi(x) - F_{S_n}(x)| \\ & \leq C \left(\sum_{k=1}^n \mathbb{E} [\xi_{n,k}^2; |\xi_{n,k}| > 1] + \sum_{k=1}^n \mathbb{E} [|\xi_{n,k}|^3; |\xi_{n,k}| \leq 1] \right). \end{aligned} \quad (1)$$

It was shown in [F68] that the constant C in (1) can be taken equal to 6. The first proof of (1) based on Stein's method was given by Barbour and Hall in [BH84]. More recently, the result was improved by Chen and Shao in [CS01], where it was shown that C can be taken equal to 4.1. The proof in [CS01] is based on Chen's concentration inequality approach in combination with Stein's method.

Theorem 1.2 has two important consequences.

The first is immediate. It is known as the Berry-Esseen inequality.

Theorem 1.3. (Berry-Esseen inequality) *There exists a universal constant $C > 0$ such that*

$$\sup_{x \in \mathbb{R}} |F_\xi(x) - F_{S_n}(x)| \leq C \sum_{k=1}^n \mathbb{E} [|\xi_{n,k}|^3]. \quad (2)$$

It was shown by Shevtsova in [Sh10] that the constant C in (2) can be taken equal to 0.56.

For the second consequence, we recall that it was pointed out by Loh in [L75] that the truncation at 1 in (1) is optimal in the sense that

$$\begin{aligned} & \sum_{k=1}^n \mathbb{E} [\xi_{n,k}^2; |\xi_{n,k}| > 1] + \sum_{k=1}^n \mathbb{E} [|\xi_{n,k}|^3; |\xi_{n,k}| \leq 1] \\ & \leq \inf_A \left(\sum_{k=1}^n \mathbb{E} [\xi_{n,k}^2; A] + \sum_{k=1}^n \mathbb{E} [|\xi_{n,k}|^3; \mathbb{R} \setminus A] \right), \end{aligned} \quad (3)$$

the infimum being taken over all Borel subsets A of the real line. Thus, applying (1) and (3), we get, for each $\epsilon > 0$,

$$\begin{aligned} & \sup_{x \in \mathbb{R}} |F_\xi(x) - F_{S_n}(x)| \\ & \leq C \left(\sum_{k=1}^n \mathbb{E} [\xi_{n,k}^2; |\xi_{n,k}| > 1] + \sum_{k=1}^n \mathbb{E} [|\xi_{n,k}|^3; |\xi_{n,k}| \leq 1] \right) \\ & \leq C \left(\sum_{k=1}^n \mathbb{E} [\xi_{n,k}^2; |\xi_{n,k}| > \epsilon] + \sum_{k=1}^n \mathbb{E} [|\xi_{n,k}|^3; |\xi_{n,k}| \leq \epsilon] \right) \\ & \leq C \left(\sum_{k=1}^n \mathbb{E} [\xi_{n,k}^2; |\xi_{n,k}| > \epsilon] + \epsilon \sum_{k=1}^n \mathbb{E} [|\xi_{n,k}|^2] \right) \\ & = C \left(\sum_{k=1}^n \mathbb{E} [\xi_{n,k}^2; |\xi_{n,k}| > \epsilon] + \epsilon \right) \end{aligned}$$

which, after calculating the superior limit of both sides and letting $\epsilon \downarrow 0$, yields

$$\begin{aligned} & \limsup_{n \rightarrow \infty} \sup_{x \in \mathbb{R}} |F_\xi(x) - F_{S_n}(x)| \\ & \leq C \sup_{\epsilon > 0} \limsup_{n \rightarrow \infty} \sum_{k=1}^n \mathbb{E} [\xi_{n,k}^2; |\xi_{n,k}| > \epsilon]. \end{aligned} \quad (4)$$

Inspired by (4), the *Lindeberg index* of $\{\xi_{n,k}\}$ was introduced by the authors in [BLV13] as

$$\text{Lin}(\{\xi_{n,k}\}) = \sup_{\epsilon > 0} \limsup_{n \rightarrow \infty} \sum_{k=1}^n \mathbb{E} [\xi_{n,k}^2; |\xi_{n,k}| > \epsilon].$$

It is clear that $0 \leq \text{Lin}(\{\xi_{n,k}\}) \leq 1$ and that $\{\xi_{n,k}\}$ satisfies Lindeberg's condition if and only if $\text{Lin}(\{\xi_{n,k}\}) = 0$.

The following example, taken from [BLV13], provides some insight into how the Lindeberg index behaves.

Let $0 < \alpha < 1$, $\beta = \frac{\alpha}{1-\alpha}$ and set

$$s_n^2 = (1 + \beta)n - \beta \sum_{k=1}^n k^{-1} = n + \beta \sum_{k=1}^n (1 - k^{-1}). \quad (5)$$

Notice that $s_n^2 \rightarrow \infty$. Now consider the 1-STA $\{\eta_{\alpha,n,k}\}$ such that

$$\mathbb{P}[\eta_{\alpha,n,k} = -1/s_n] = \mathbb{P}[\eta_{\alpha,n,k} = 1/s_n] = \frac{1}{2}(1 - \beta k^{-1}) \quad (6)$$

and

$$\mathbb{P}[\eta_{\alpha,n,k} = -\sqrt{k}/s_n] = \mathbb{P}[\eta_{\alpha,n,k} = \sqrt{k}/s_n] = \frac{1}{2}\beta k^{-1}. \quad (7)$$

Then it was shown in [BLV13] (Proposition 2.2) that

$$\text{Lin}(\{\eta_{\alpha,n,k}\}) = \alpha$$

and that $\{\eta_{\alpha,n,k}\}$ is *infinitesimal* in the sense that

$$\forall \epsilon > 0 : \max_{k=1}^n \mathbb{P}[|\eta_{\alpha,n,k}| > \epsilon] \rightarrow 0.$$

Now, as a second consequence of Theorem 1.2, the following quantitative version of the Lindeberg CLT is yielded by (4).

Theorem 1.4. (Quantitative Lindeberg CLT) *There exists a universal constant $C > 0$ such that*

$$\limsup_{n \rightarrow \infty} \sup_{x \in \mathbb{R}} |F_{\xi}(x) - F_{S_n}(x)| \leq C \text{Lin}(\{\xi_{n,k}\}). \quad (8)$$

Using an asymptotic smoothing technique and Stein's method, it was shown in [BLV13] that under the mild assumption that $\{\xi_{n,k}\}$ be infinitesimal, the constant C in (8) can be taken equal to 1. Moreover in [BLV13] it was also shown that the expression on the left side is actually an index of convergence for a natural approach structure and thus, if we denote the associated limit operator simply λ the result reads

$$\lambda(S_n \rightarrow \xi) \leq \text{Lin}(\{\xi_{n,k}\}).$$

This was the situation in the one-dimensional case. We now turn to the multivariate case and see how the techniques and inequalities in the previous discussion can be extended. Throughout, we keep $N \in \mathbb{N}_0$ fixed and we let $|\cdot|$ stand for the norm and $\langle \cdot, \cdot \rangle$ for the inner product in Euclidean N -space \mathbb{R}^N . By a *random N -vector* we mean an \mathbb{R}^N -valued random variable. Furthermore, Ξ is a standard normally distributed random N -vector and $\{\Xi_{n,k}\}$ an N -dimensional standard triangular array (N -STA), i.e. a triangular array of random N -vectors

$$\begin{array}{ccc} \Xi_{1,1} & & \\ \Xi_{2,1} & \Xi_{2,2} & \\ \Xi_{3,1} & \Xi_{3,2} & \Xi_{3,3} \\ & \vdots & \end{array}$$

with the following properties.

$$(1) \forall n : \Xi_{n,1}, \dots, \Xi_{n,n} \text{ are independent.}$$

$$(2) \forall n, k : \mathbb{E}[\Xi_{n,k}] = 0.$$

$$(3) \forall n : \text{cov}(\Sigma_n) = I_{N \times N} \text{ with } \Sigma_n = \sum_{k=1}^n \Xi_{n,k}.$$

Note that the notion of N -STA coincides with the earlier introduced notion of 1-STA in the case where $N = 1$.

The Lindeberg CLT is now extended as follows ([S11]).

Theorem 1.5. (Lindeberg CLT for random N -vectors) *Suppose that $\{\Xi_{n,k}\}$ satisfies Lindeberg's condition in the sense that*

$$\forall \epsilon > 0 : \sum_{k=1}^n \mathbb{E} \left[|\Xi_{n,k}|^2 ; |\Xi_{n,k}| > \epsilon \right] \rightarrow 0.$$

Then

$$\Sigma_n \xrightarrow{w} \Xi.$$

It is customary to consider the distance

$$\sup_{A \in \mathcal{C}} |\mathbb{P}[\mathbf{H} \in A] - \mathbb{P}[\mathbf{H}' \in A]|,$$

\mathcal{C} being the collection of all convex Borel subsets of Euclidean N -space, between random N -vectors \mathbf{H} and \mathbf{H}' . Note that this distance is stronger than the earlier introduced Kolmogorov distance in the case where $N = 1$.

The question whether Theorem 1.2 can be extended to the multivariate setting is still open. However, multivariate versions of Stein's method (Barbour [B90], Götze [G91], Goldstein and Rinott [GR96], Chatterjee and Meckes [CM08], Meckes [M09], Reinert and Röllin [RR09], Nourdin, Peccati and Réveillac [NPR10] and of the Berry-Esseen inequality (Götze [G91], Rinott and Rotar [RR96], Bentkus [B03], Bhattacharya and Holmes [BH10], Chen and Fang [CF]) have been the object of extensive study. In this spirit, Chen and Fang have recently obtained the following result in [CF].

Theorem 1.6. (Berry-Esseen inequality for random N -vectors) *There exists a universal constant $C > 0$ such that*

$$\sup_{A \in \mathcal{C}} |\mathbb{P}[\Xi \in A] - \mathbb{P}[\Sigma_n \in A]| \leq C\sqrt{N} \sum_{k=1}^n \mathbb{E} \left[|\Xi_{n,k}|^3 \right]. \quad (9)$$

It was shown in [CF] that the constant C in (9) can be taken equal to 115. An issue of importance is the fact that the upper bound in (9) is of order $O(\sqrt{N})$, the sharpest obtained so far. We also notice that Bentkus has established in [B03] an inequality of the type (9) with an upper bound of order $O(\sqrt[4]{N})$ under the additional assumption that $\Xi_{n,1}, \dots, \Xi_{n,n}$ be identically distributed.

At this point it is natural to ask for a version of Theorem 1.4 for random N -vectors, but, even with a multivariate version of Stein's method at hand, there seem to be some intrinsic obstructions towards obtaining such a result. However, if, in the spirit of e.g. [GJT02], we consider ϕ_Ξ and ϕ_{Σ_n} , where

$$\phi_{\mathbf{H}}(t) = \mathbb{E} [\exp(-i \langle t, \mathbf{H} \rangle)], \quad t \in \mathbb{R}^N,$$

represents the Fourier transform of the random N -vector \mathbf{H} , instead of the cumulative distribution functions F_Ξ and F_{Σ_n} , then we can show that Stein's method as outlined in e.g. [M09], [NPR10] and [CF] becomes applicable to get our main results, Theorem 2.5 and Corollary 2.6. The latter is a

quantitative multivariate Lindeberg CLT of the same taste as Theorem 1.4. The crux of the matter consists in deriving an explicit integral representation for the Hessian matrix of a solution to the Stein equation with test function $e_t(x) = \exp(-i \langle t, x \rangle)$, where $t, x \in \mathbb{R}^N$ (Proposition 3.5).

2. FORMULATION OF THE MAIN RESULTS

We keep the terminology and the notation of the previous section.

Let $\phi_{\mathbf{H}}$ be the Fourier transform of the random N -vector \mathbf{H} . That is, for $t, x \in \mathbb{R}^N$,

$$\phi_{\mathbf{H}}(t) = \mathbb{E} [e_t(\mathbf{H})]$$

with

$$e_t(x) = \exp(-i \langle t, x \rangle).$$

It is well-known that if we put \mathcal{R} the space of all random N -vectors then

$$(\phi_{\mathbf{H}} : \mathcal{R}, \mathcal{T}_w \rightarrow \mathbb{R}, \mathcal{T}_{d_E} : H \mapsto \mathbb{E} [e_t(\mathbf{H})])_{t \in \mathbb{R}^N}$$

is an initial source, i.e. the topology of weak convergence is the weakest topology making all these maps continuous. We now lift this canonical way to obtain the weak topology from the topological to the approach level. In order to do this we simply replace the Euclidean topology on \mathbb{R} by the Euclidean metric and take the initial approach structure (here simply denoted by its limit operator λ) rather than the initial topology

$$(\phi_{\mathbf{H}} : \mathcal{R}, \lambda \rightarrow \mathbb{R}, d_E : H \mapsto \mathbb{E} [e_t(\mathbf{H})])_{t \in \mathbb{R}^N}$$

Note that an initial metric does not exist, it makes no sense to ask for a “weakest” metric, it is absolutely required to go to the realm of approach spaces in order to find a solution.

It then follows from general results in approach theory (see e.g. [L15]) that the limit operator in this space is given by the following formula, where $\Sigma_n, n \in \mathbb{N}$ and Ξ are random variables:

$$\lambda(\Sigma_n \rightarrow \Xi) = \sup_{t \in \mathbb{R}^N} \limsup_{n \rightarrow \infty} |\phi_{\Xi}(t) - \phi_{\Sigma_n}(t)|.$$

We refer the reader interested in the fundamentals of approach theory to [L97], [L15], [BLV11] and [BLV11’]. For the sake of this paper, the following result, which reveals that the number $\lambda(\Sigma_n \rightarrow \Xi)$ indeed measures how far the sequence $(\Sigma_n)_n$ deviates from being weakly convergent to Ξ , suffices.

Proposition 2.1.

$$0 \leq \lambda(\Sigma_n \rightarrow \Xi) \leq 2 \tag{10}$$

and

$$\lambda(\Sigma_n \rightarrow \Xi) = 0 \Leftrightarrow \Sigma_n \xrightarrow{w} \Xi. \tag{11}$$

Proof. (10) is trivial. (11) follows from Lévy’s Continuity Theorem, which states that weak convergence of random vectors is equivalent to pointwise convergence of their Fourier transforms. \square

Lemma 2.2.

$$\sum_{k=1}^n \mathbb{E} \left[|\Xi_{n,k}|^2 \right] = N. \quad (12)$$

Proof. The calculation

$$\begin{aligned} \sum_{k=1}^n \mathbb{E} \left[|\Xi_{n,k}|^2 \right] &= \sum_{k=1}^n \mathbb{E} \left[\sum_{l=1}^N \Xi_{n,k,l}^2 \right] \\ &= \sum_{l=1}^N \mathbb{E} \left[\sum_{k=1}^n \Xi_{n,k,l}^2 \right] \\ &= \sum_{l=1}^N \mathbb{E} \left[\left(\sum_{k=1}^n \Xi_{n,k,l} \right)^2 \right] \\ &= \sum_{l=1}^N \text{cov} \left(\sum_{k=1}^n \Xi_{n,k}, \sum_{k=1}^n \Xi_{n,k} \right)_{l,l} \\ &\quad \left(\text{cov} \left(\sum_{k=1}^n \Xi_{n,k} \right) = I_{N \times N} \right) = N \end{aligned}$$

where $\Xi_{n,k}$ and $\Xi_{n,j}$ are independent if $k \neq j$ and $\mathbb{E} [\Xi_{n,k}] = 0$ for all k finishes the proof. \square

We say that $\{\Xi_{n,k}\}$ is *infinitesimal* iff

$$\forall \epsilon > 0 : \max_{k=1}^n \mathbb{P} [|\Xi_{n,k}| > \epsilon] \rightarrow 0$$

and we extend the notion of *Lindeberg index* by putting

$$\text{Lin}(\{\Xi_{n,k}\}) = \sup_{\epsilon > 0} \limsup_{n \rightarrow \infty} \sum_{k=1}^n \mathbb{E} \left[|\Xi_{n,k}|^2 ; |\Xi_{n,k}| > \epsilon \right].$$

It follows from Lemma 2.2 that $0 \leq \text{Lin}(\{\Xi_{n,k}\}) \leq N$ and it is clear that $\{\Xi_{n,k}\}$ satisfies Lindeberg's condition if and only if $\text{Lin}(\{\Xi_{n,k}\}) = 0$.

Proposition 2.3. *If $\{\Xi_{n,k}\}$ satisfies Lindeberg's condition, then it is infinitesimal.*

Proof. For $\epsilon > 0$, Chebyshev's Inequality gives

$$\begin{aligned} &\max_{k=1}^n \mathbb{P} [|\Xi_{n,k}| > \epsilon] \\ &\leq \epsilon^{-2} \max_{k=1}^n \mathbb{E} \left[|\Xi_{n,k}|^2 \right] \\ &= \epsilon^{-2} \max_{k=1}^n \mathbb{E} \left[|\Xi_{n,k}|^2 ; |\Xi_{n,k}| > \epsilon^2 \right] + \epsilon^{-2} \max_{k=1}^n \mathbb{E} \left[|\Xi_{n,k}|^2 ; |\Xi_{n,k}| \leq \epsilon^2 \right] \\ &\leq \epsilon^{-2} \sum_{k=1}^n \mathbb{E} \left[|\Xi_{n,k}|^2 ; |\Xi_{n,k}| > \epsilon^2 \right] + \epsilon^2 \end{aligned}$$

from which the proposition easily follows. \square

For an N -STA $\{\mathbf{H}_{n,k}\}$, we define the auxiliary number

$$L(\{\Xi_{n,k}\}, \{\mathbf{H}_{n,k}\}) = \sup_{t \in \mathbb{R}^N} \limsup_{n \rightarrow \infty} \sum_{k=1}^n \mathbb{E} \left[|\Xi_{n,k}|^2 ; |\langle \mathbf{H}_{n,k}, t \rangle| > 1 \right].$$

Proposition 2.4 below shows how $L(\{\Xi_{n,k}\}, \{\mathbf{H}_{n,k}\})$ is linked to both the Lindeberg index and the condition of being infinitesimal.

Proposition 2.4.

$$0 \leq L(\{\Xi_{n,k}\}, \{\mathbf{H}_{n,k}\}) \leq N. \quad (13)$$

Also,

$$L(\{\Xi_{n,k}\}, \{\Xi_{n,k}\}) \leq \text{Lin}(\{\Xi_{n,k}\}) \quad (14)$$

and the inequality in (14) becomes an equality if $N = 1$. Finally, let $\{\Xi_{n,k}^0\}$ be any independent copy of $\{\Xi_{n,k}\}$. Then

$$\{\Xi_{n,k}\} \text{ is infinitesimal} \Rightarrow L(\{\Xi_{n,k}\}, \{\Xi_{n,k}^0\}) = 0. \quad (15)$$

Proof. (12) entails (13). Furthermore, by the Cauchy-Schwarz Inequality, for $t \in \mathbb{R}^N \setminus \{0\}$,

$$\sum_{k=1}^n \mathbb{E} \left[|\Xi_{n,k}|^2 ; |\langle \Xi_{n,k}, t \rangle| > 1 \right] \leq \sum_{k=1}^n \mathbb{E} \left[|\Xi_{n,k}|^2 ; |\Xi_{n,k}| > |t|^{-1} \right]$$

proving (14). If $N = 1$, then the inequality in (14) trivially becomes an equality. Finally, suppose that $\{\Xi_{n,k}\}$ is infinitesimal and let $\{\Xi_{n,k}^0\}$ be an independent copy of $\{\Xi_{n,k}\}$. Then, by the Cauchy-Schwarz Inequality and (12), for $t \in \mathbb{R}^N \setminus \{0\}$,

$$\begin{aligned} & \sum_{k=1}^n \mathbb{E} \left[|\Xi_{n,k}|^2 ; |\langle \Xi_{n,k}^0, t \rangle| > 1 \right] \\ & \leq \sum_{k=1}^n \mathbb{E} \left[|\Xi_{n,k}|^2 \right] \mathbb{P} \left[|\Xi_{n,k}^0| > |t|^{-1} \right] \\ & \leq \max_{k=1}^n \mathbb{P} \left[|\Xi_{n,k}^0| > |t|^{-1} \right] \sum_{k=1}^n \mathbb{E} \left[|\Xi_{n,k}|^2 \right] \\ & = N \max_{k=1}^n \mathbb{P} \left[|\Xi_{n,k}^0| > |t|^{-1} \right] \end{aligned}$$

which establishes (15). \square

We are now in a position to state our main results. The proof of Theorem 2.5 is deferred to the next section.

Theorem 2.5. Let $\{\Xi_{n,k}^0\}$ be an independent copy of $\{\Xi_{n,k}\}$. Then

$$\lambda(\Sigma_n \rightarrow \Xi) \leq 2 \left(L(\{\Xi_{n,k}\}, \{\Xi_{n,k}\}) + L(\{\Xi_{n,k}\}, \{\Xi_{n,k}^0\}) \right). \quad (16)$$

In particular, for $t \in \mathbb{R}^N$,

$$\limsup_{n \rightarrow \infty} |\phi_\Xi(t) - \phi_{\Sigma_n}(t)| \quad (17)$$

$$\leq 2 \left(1 - \exp \left(-\frac{1}{2} |t|^2 \right) \right) (L(\{\Xi_{n,k}\}, \{\Xi_{n,k}\}) + L(\{\Xi_{n,k}\}, \{\Xi_{n,k}^0\})).$$

Theorem 2.5 has the following corollary, which is a multivariate quantitative Lindeberg CLT of the same taste as Theorem 1.4. The proof of this theorem requires several steps and intermediate results, therefore we defer it to the third section.

Corollary 2.6. (Quantitative Lindeberg CLT for the Fourier transforms of random N -vectors) *Suppose that $\{\Xi_{n,k}\}$ is infinitesimal. Then*

$$\lambda(\Sigma_n \rightarrow \Xi) \leq 2\text{Lin}(\{\Xi_{n,k}\}). \quad (18)$$

More explicitly,

$$\sup_{t \in \mathbb{R}^N} \limsup_{n \rightarrow \infty} |\phi_\Xi(t) - \phi_{\Sigma_n}(t)| \leq 2\text{Lin}(\{\Xi_{n,k}\}). \quad (19)$$

Proof. Recall that Proposition 2.4 entails $L(\{\Xi_{n,k}\}, \{\Xi_{n,k}\}) \leq \text{Lin}(\{\Xi_{n,k}\})$ and that, $\{\Xi_{n,k}\}$ being infinitesimal, $L(\{\Xi_{n,k}\}, \{\Xi_{n,k}^0\}) = 0$. Thus (16) immediately gives (18). \square

Remarks (1) Corollary 2.6 is stronger than Theorem 1.5. Indeed, suppose that $\{\Xi_{n,k}\}$ satisfies Lindeberg's condition. Then, by Proposition 2.3, $\{\Xi_{n,k}\}$ is also infinitesimal. But then Corollary 2.6 implies that $\lambda(\Sigma_n \rightarrow \Xi) = 0$ and thus, by Proposition 2.1, $\Sigma_n \xrightarrow{w} \Xi$. The advantage of Theorem 2.5 is that it continues to be informative for STA's such as $\{\eta_{\alpha,n,k}\}$, defined by (5), (6) and (7), for which Lindeberg's condition is not satisfied, whereas Theorem 1.5 fails to be applicable for such STA's.

(2) For the large class of infinitesimal N -STA's, (19) yields an upper bound which does not depend on the dimension N . This suggests the possibility of extending the result to an infinite dimensional setting. Such extensions will be discussed elsewhere.

(3) The left-hand side in (19) is optimal in the sense that it is impossible to get similar upper bounds for $\limsup_{n \rightarrow \infty} \sup_{t \in \mathbb{R}^N} |\phi_\Xi(t) - \phi_{\Sigma_n}(t)|$. Indeed, let $N = 1$ and consider i.i.d. random variables ξ_1, ξ_2, \dots with $\mathbb{P}[\xi_k = -1] = \mathbb{P}[\xi_k = 1] = 1/2$ and put $\xi_{n,k} = \xi_k / \sqrt{n}$ and $S_n = \sum_{k=1}^n \xi_{n,k}$. Then $\{\xi_{n,k}\}$ is a 1-STA such that $\text{Lin}(\{\xi_{n,k}\}) = 0$, but, for each n , it holds that $\sup_{t \in \mathbb{R}} |\phi_\xi(t) - \phi_{S_n}(t)| = \sup_{t \in \mathbb{R}} |\exp(-t^2/2) - \cos^n(t/\sqrt{n})| = 1$.

3. PROOF OF THEOREM 2.5

We keep the terminology and the notation of the previous sections.

The proof of Theorem 2.5 heavily depends on a multivariate version of Stein's method as outlined in e.g. [M09], [NPR10] and [CF].

Let $h : \mathbb{R}^N \rightarrow \mathbb{C}$ be bounded and twice continuously differentiable with bounded first order and second order partial derivatives and let $f_h : \mathbb{R}^N \rightarrow \mathbb{C}$ be the solution to the Stein equation

$$\langle x, \nabla f(x) \rangle - \Delta f(x) = \mathbb{E}[h(\Xi)] - h(x) \quad (20)$$

given by

$$f_h(x) = - \int_0^1 \frac{1}{2s} \mathbb{E} [h(\Xi) - h(\sqrt{s}x + \sqrt{1-s}\Xi)] ds, \quad (21)$$

see [M09] or [NPR10]. Furthermore, let $\text{Hess}f_h(x)$ stand for the Hessian matrix of f_h at x and put

$$D_{\text{Hess}f_h}(x, y) = \text{Hess}f_h(x) - \text{Hess}f_h(y).$$

Finally, let $\{\Xi_{n,k}^0\}$ be an independent copy of $\{\Xi_{n,k}\}$.

The following proposition follows from the explicit structure of the Stein equation.

Proposition 3.1.

$$\begin{aligned} & \mathbb{E} [h(\Xi) - h(\Sigma_n)] \quad (22) \\ &= \sum_{k=1}^n \int_0^1 \mathbb{E} \left[\left\langle \Xi_{n,k}, D_{\text{Hess}f_h} \left(\sum_{j \neq k} \Xi_{n,j} + r\Xi_{n,k}, \sum_{j \neq k} \Xi_{n,j} \right) \Xi_{n,k} \right\rangle \right] dr \\ & \quad - \sum_{k=1}^n \mathbb{E} \left[\left\langle \Xi_{n,k}, D_{\text{Hess}f_h} \left(\sum_{k=1}^n \Xi_{n,k}^0, \sum_{j \neq k} \Xi_{n,j}^0 \right) \Xi_{n,k} \right\rangle \right]. \end{aligned}$$

Proof. The fact that f_h is a solution to the Stein equation (20) leads to

$$\begin{aligned} & \mathbb{E} [h(\Xi) - h(\Sigma_n)] \\ &= \sum_{k=1}^n \mathbb{E} \left[\left\langle \Xi_{n,k}, \nabla f_h \left(\sum_{k=1}^n \Xi_{n,k} \right) \right\rangle - \Delta f_h \left(\sum_{k=1}^n \Xi_{n,k} \right) \right]. \quad (23) \end{aligned}$$

Furthermore,

$$\begin{aligned} & \sum_{k=1}^n \mathbb{E} \left[\left\langle \Xi_{n,k}, \nabla f_h \left(\sum_{k=1}^n \Xi_{n,k} \right) \right\rangle - \Delta f_h \left(\sum_{k=1}^n \Xi_{n,k} \right) \right] \\ &= \sum_{k=1}^n \mathbb{E} \left[\left\langle \Xi_{n,k}, \right. \right. \quad (24) \\ & \quad \left. \nabla f_h \left(\sum_{k=1}^n \Xi_{n,k} \right) - \nabla f_h \left(\sum_{j \neq k} \Xi_{n,j} \right) - \text{Hess}f_h \left(\sum_{j \neq k} \Xi_{n,j} \right) \Xi_{n,k} \right\rangle \\ & \quad \left. - \sum_{k=1}^n \mathbb{E} \left[\left\langle \Xi_{n,k}, \left(\text{Hess}f_h \left(\sum_{k=1}^n \Xi_{n,k}^0 \right) - \text{Hess}f_h \left(\sum_{j \neq k} \Xi_{n,j}^0 \right) \right) \Xi_{n,k} \right\rangle \right] \right] \end{aligned}$$

which is seen by calculating the right-hand side and noticing the following three facts. Firstly,

$$\sum_{k=1}^n \mathbb{E} \left[\left\langle \Xi_{n,k}, \nabla f_h \left(\sum_{j \neq k} \Xi_{n,j} \right) \right\rangle \right]$$

$$\begin{aligned}
&= \sum_{k=1}^n \sum_{l=1}^N \mathbb{E} \left[\Xi_{n,k,l} \frac{\partial f_h}{\partial x_l} \left(\sum_{j \neq k} \Xi_{n,j} \right) \right] \\
&\quad (\Xi_{n,k} \text{ and } \sum_{j \neq k} \Xi_{n,j} \text{ are independent}) \\
&= \sum_{k=1}^n \sum_{l=1}^N \mathbb{E} [\Xi_{n,k,l}] \mathbb{E} \left[\frac{\partial f_h}{\partial x_l} \left(\sum_{j \neq k} \Xi_{n,j} \right) \right] \\
&\quad (\mathbb{E} [\Xi_{n,k}] = 0) \\
&= 0.
\end{aligned}$$

Secondly,

$$\begin{aligned}
&\sum_{k=1}^n \mathbb{E} \left[\left\langle \Xi_{n,k}, \text{Hess} f_h \left(\sum_{j \neq k} \Xi_{n,j} \right) \Xi_{n,k} \right\rangle \right] \\
&= \sum_{k=1}^n \sum_{l=1}^N \sum_{m=1}^N \mathbb{E} \left[\frac{\partial^2 f_h}{\partial x_l \partial x_m} \left(\sum_{j \neq k} \Xi_{n,j} \right) \Xi_{n,k,l} \Xi_{n,k,m} \right] \\
&\quad (\Xi_{n,k} \text{ and } \sum_{j \neq k} \Xi_{n,j} \text{ are independent}) \\
&= \sum_{k=1}^n \sum_{l=1}^N \sum_{m=1}^N \mathbb{E} \left[\frac{\partial^2 f_h}{\partial x_l \partial x_m} \left(\sum_{j \neq k} \Xi_{n,j} \right) \text{cov}(\Xi_{n,k})_{l,m} \right] \\
&\quad (\{\Xi_{n,k}^0\} \text{ is an independent copy of } \{\Xi_{n,k}\}) \\
&= \sum_{k=1}^n \sum_{l=1}^N \sum_{m=1}^N \mathbb{E} \left[\frac{\partial^2 f_h}{\partial x_l \partial x_m} \left(\sum_{j \neq k} \Xi_{n,j}^0 \right) \text{cov}(\Xi_{n,k})_{l,m} \right] \\
&\quad (\Xi_{n,k}^0 \text{ and } \sum_{j \neq k} \Xi_{n,j} \text{ are independent}) \\
&= \sum_{k=1}^n \sum_{l=1}^N \sum_{m=1}^N \mathbb{E} \left[\frac{\partial^2 f_h}{\partial x_l \partial x_m} \left(\sum_{j \neq k} \Xi_{n,j}^0 \right) \Xi_{n,k,l} \Xi_{n,k,m} \right] \\
&= \sum_{k=1}^n \mathbb{E} \left[\left\langle \Xi_{n,k}, \text{Hess} f_h \left(\sum_{j \neq k} \Xi_{n,j}^0 \right) \Xi_{n,k} \right\rangle \right].
\end{aligned}$$

Thirdly,

$$\begin{aligned}
&\sum_{k=1}^n \mathbb{E} \left[\left\langle \Xi_{n,k}, \text{Hess} f_h \left(\sum_{k=1}^n \Xi_{n,k}^0 \right) \Xi_{n,k} \right\rangle \right] \\
&= \sum_{k=1}^n \sum_{l=1}^N \sum_{m=1}^N \mathbb{E} \left[\frac{\partial^2 f_h}{\partial x_l \partial x_m} \left(\sum_{k=1}^n \Xi_{n,k}^0 \right) \Xi_{n,k,l} \Xi_{n,k,m} \right]
\end{aligned}$$

$$\begin{aligned}
& (\{\Xi_{n,k}^0\} \text{ and } \{\Xi_{n,k}\} \text{ are independent}) \\
&= \sum_{k=1}^n \sum_{l=1}^N \sum_{m=1}^N \mathbb{E} \left[\frac{\partial^2 f_h}{\partial x_l \partial x_m} \left(\sum_{k=1}^n \Xi_{n,k}^0 \right) \right] \text{cov}(\Xi_{n,k})_{l,m} \\
&= \sum_{l=1}^N \sum_{m=1}^N \mathbb{E} \left[\frac{\partial^2 f_h}{\partial x_l \partial x_m} \left(\sum_{k=1}^n \Xi_{n,k}^0 \right) \right] \left(\sum_{k=1}^n \text{cov}(\Xi_{n,k})_{l,m} \right) \\
&\quad (\Xi_{n,1}, \dots, \Xi_{n,n} \text{ are independent and } \mathbb{E}[\Xi_{n,k}] = 0) \\
&= \sum_{l=1}^N \sum_{m=1}^N \mathbb{E} \left[\frac{\partial^2 f_h}{\partial x_l \partial x_m} \left(\sum_{k=1}^n \Xi_{n,k}^0 \right) \right] \left(\text{cov} \left(\sum_{k=1}^n \Xi_{n,k} \right)_{l,m} \right) \\
&\quad \left(\text{cov} \left(\sum_{k=1}^n \Xi_{n,k} \right) = I_{N \times N} \right) \\
&= \mathbb{E} \left[\Delta f_h \left(\sum_{k=1}^n \Xi_{n,k}^0 \right) \right] \\
&\quad (\{\Xi_{n,k}^0\} \text{ is a copy of } \{\Xi_{n,k}\}) \\
&= \mathbb{E} \left[\Delta f_h \left(\sum_{k=1}^n \Xi_{n,k} \right) \right].
\end{aligned}$$

Finally, the Fundamental Theorem of Calculus reveals that

$$\begin{aligned}
& \nabla f_h \left(\sum_{k=1}^n \Xi_{n,k} \right) - \nabla f_h \left(\sum_{j \neq k} \Xi_{n,k} \right) \\
&= \int_0^1 \text{Hess} f_h \left(\sum_{j \neq k} \Xi_{n,j} + r \Xi_{n,k} \right) \Xi_{n,k} dr. \tag{25}
\end{aligned}$$

Combining (23), (24) and (25) proves (22) and we are done. \square

Proposition 3.1 highlights the role of the Hessian matrix of f_h in the search for an upper bound for expressions of the type $|\mathbb{E}[h(\Xi) - h(\Sigma_n)]|$. In the following proposition we establish an explicit integral representation for $\text{Hess} f_h$. We consider an N -vector $z \in \mathbb{C}^N$ as a $1 \times N$ -matrix and we denote its transpose as z^T .

Proposition 3.2.

$$\nabla f_h(x) = - \int_0^1 \frac{1}{2\sqrt{s(1-s)}} \mathbb{E} [h(\sqrt{s}x + \sqrt{1-s}\Xi) \Xi] ds \tag{26}$$

and

$$\text{Hess} f_h(x) = - \int_0^1 \frac{1}{2(1-s)} \mathbb{E} [h(\sqrt{s}x + \sqrt{1-s}\Xi) (\Xi \Xi^T - I_{N \times N})] ds. \tag{27}$$

Proof. Using (21) and performing an integration by parts on the Gaussian expectation gives

$$\frac{\partial f_h}{\partial x_l}(a) = \int_0^1 \frac{1}{2\sqrt{s(1-s)}} \mathbb{E} [h(\sqrt{sa} + \sqrt{1-s}\Xi) \Xi_l] ds$$

and (26) follows. Using (26) and again performing an integration by parts on the Gaussian expectation gives

$$\frac{\partial^2 f_h}{\partial x_l \partial x_m}(a) = \int_0^1 \frac{1}{2(1-s)} \mathbb{E} [h(\sqrt{sa} + \sqrt{1-s}\Xi) (\Xi_l \Xi_m - \delta_{lm})] ds,$$

with δ_{lm} the Kronecker delta, and (27) follows. \square

The singularity at 1 of the integrand in (27) makes it hard to control $\text{Hess}f_h(x)$ for general h . However, we establish in Proposition 3.5 that for the specific choice

$$h(x) = e_t(x) = \exp(-i\langle t, x \rangle), \quad t \in \mathbb{R}^N,$$

the integral representation of $\text{Hess}f_h(x)$ does not contain the factor $\frac{1}{1-s}$ anymore. We first need two lemmas.

Lemma 3.3. *Fix $y, t \in \mathbb{R}^N$ and $s \in [0, 1]$. Put*

$$\alpha_{y,t,s} = y + i\sqrt{1-st}.$$

Then

$$-i\sqrt{1-s}\langle t, y \rangle - \frac{1}{2}|y|^2 = -\frac{1}{2}(1-s)|t|^2 - \frac{1}{2}\alpha_{y,t,s}^\tau \alpha_{y,t,s}. \quad (28)$$

Furthermore,

$$yy^\tau - I_{N \times N} = \alpha_{y,t,s} \alpha_{y,t,s}^\tau - i\sqrt{1-st} \alpha_{y,t,s}^\tau - i\sqrt{1-s} \alpha_{y,t,s} t^\tau - (1-s)tt^\tau - I_{N \times N}. \quad (29)$$

Proof. This is elementary. \square

Lemma 3.4.

$$\begin{aligned} & \mathbb{E} [e_t(\sqrt{sx} + \sqrt{1-s}\Xi) (\Xi \Xi^\tau - I_{N \times N})] \\ &= -(1-s)tt^\tau \exp\left(-i\sqrt{s}\langle t, x \rangle - \frac{1}{2}(1-s)|t|^2\right). \end{aligned} \quad (30)$$

Proof. Put

$$\alpha_{y,t,s} = y + i\sqrt{1-st}.$$

From (28) and (29) we learn that

$$\begin{aligned} & \mathbb{E} [e_t(\sqrt{sx} + \sqrt{1-s}\Xi) (\Xi \Xi^\tau - I_{N \times N})] \\ &= \frac{1}{(2\pi)^{N/2}} \int_{\mathbb{R}^N} \exp\left(-i\sqrt{s}\langle t, x \rangle - i\sqrt{1-s}\langle t, y \rangle - \frac{1}{2}|y|^2\right) (yy^\tau - I_{N \times N}) dy \\ &= \exp\left(-i\sqrt{s}\langle t, x \rangle - \frac{1}{2}(1-s)|t|^2\right) \\ & \quad \left(\frac{1}{(2\pi)^{N/2}} \int_{\mathbb{R}^N} (\alpha_{y,t,s} \alpha_{y,t,s}^\tau - I_{N \times N}) \exp\left(-\frac{1}{2}\alpha_{y,t,s}^\tau \alpha_{y,t,s}\right) dy\right) \end{aligned}$$

$$\begin{aligned}
& -i\sqrt{1-s} \frac{1}{(2\pi)^{N/2}} \int_{\mathbb{R}^N} t\alpha_{y,t,s}^\tau \exp\left(-\frac{1}{2}\alpha_{y,t,s}^\tau \alpha_{y,t,s}\right) dy \\
& -i\sqrt{1-s} \frac{1}{(2\pi)^{N/2}} \int_{\mathbb{R}^N} \alpha_{y,t,s} t^\tau \exp\left(-\frac{1}{2}\alpha_{y,t,s}^\tau \alpha_{y,t,s}\right) dy \\
& -((1-s)tt^\tau + I_{N \times N}) \frac{1}{(2\pi)^{N/2}} \int_{\mathbb{R}^N} \exp\left(-\frac{1}{2}\alpha_{y,t,s}^\tau \alpha_{y,t,s}\right) dy
\end{aligned}$$

which, by Cauchy's Integral Theorem,

$$\begin{aligned}
& = \exp\left(-i\sqrt{s}\langle t, x \rangle - \frac{1}{2}(1-s)|t|^2\right) \left(\frac{1}{(2\pi)^{N/2}} \int_{\mathbb{R}^N} (yy^\tau \exp\left(-\frac{1}{2}|y|^2\right)) dy\right. \\
& \quad - i\sqrt{1-s} \frac{1}{(2\pi)^{N/2}} \int_{\mathbb{R}^N} ty^\tau \exp\left(-\frac{1}{2}|y|^2\right) dy \\
& \quad - i\sqrt{1-s} \frac{1}{(2\pi)^{N/2}} \int_{\mathbb{R}^N} yt^\tau \exp\left(-\frac{1}{2}|y|^2\right) dy \\
& \quad \left. - ((1-s)tt^\tau + I_{N \times N}) \frac{1}{(2\pi)^{N/2}} \int_{\mathbb{R}^N} \exp\left(-\frac{1}{2}|y|^2\right) dy\right) \\
& = \exp\left(-i\sqrt{s}\langle t, x \rangle - \frac{1}{2}(1-s)|t|^2\right) \\
& \quad (\text{cov}(\Xi) - i\sqrt{1-s}t\mathbb{E}[\Xi]^\tau - i\sqrt{1-s}\mathbb{E}[\Xi]t^\tau - (1-s)tt^\tau - I_{N \times N})
\end{aligned}$$

and (30) follows. \square

Proposition 3.5.

$$\text{Hess}f_{e_t}(x) = -\frac{1}{2}tt^\tau \int_0^1 \exp\left(-i\sqrt{s}\langle t, x \rangle - \frac{1}{2}(1-s)|t|^2\right) ds. \quad (31)$$

In particular,

$$\begin{aligned}
& D_{\text{Hess}f_{e_t}}(x, y) \\
& = -\frac{1}{2}tt^\tau \int_0^1 e_{\sqrt{st}}(y) \left[e_{\sqrt{st}}(x-y) - 1\right] \exp\left(-\frac{1}{2}(1-s)|t|^2\right) ds. \quad (32)
\end{aligned}$$

Proof. Combining (27) and (30) gives (31). Also, (32) follows immediately from (31). \square

Proposition 3.1 and Proposition 3.5 lead to the following result, which contains an explicit formula for the quantity $\mathbb{E}[e_t(\Xi) - e_t(\Sigma_n)]$ without any reference to the Stein equation.

Proposition 3.6.

$$\begin{aligned}
& \mathbb{E}[e_t(\Xi) - e_t(\Sigma_n)] \quad (33) \\
& = -\frac{1}{2} \int_0^1 \left(\int_0^1 \sum_{k=1}^n \mathbb{E} \left[e_{\sqrt{st}} \left(\sum_{j \neq k} \Xi_{n,j} \right) \left[e_{\sqrt{st}}(r\Xi_{n,k}) - 1 \right] |\langle \Xi_{n,k}, t \rangle|^2 \right] dr \right)
\end{aligned}$$

$$- \sum_{k=1}^n \mathbb{E} \left[e^{\sqrt{st}} \left(\sum_{j \neq k} \Xi_{n,j}^0 \right) \left[e^{\sqrt{st}} (\Xi_{n,k}^0) - 1 \right] |\langle \Xi_{n,k}, t \rangle|^2 \right] e^{-\frac{1}{2}(1-s)|t|^2} ds.$$

Proof. Applying (22) gives

$$\begin{aligned} & \mathbb{E} [e_t(\Xi) - e_t(\Sigma_n)] \\ &= \sum_{k=1}^n \int_0^1 \mathbb{E} \left[\left\langle \Xi_{n,k}, D_{\text{Hess} f_{e_t}} \left(\sum_{j \neq k} \Xi_{n,j} + r \Xi_{n,k}, \sum_{j \neq k} \Xi_{n,j} \right) \Xi_{n,k} \right\rangle \right] dr \\ & \quad - \sum_{k=1}^n \mathbb{E} \left[\left\langle \Xi_{n,k}, D_{\text{Hess} f_{e_t}} \left(\sum_{k=1}^n \Xi_{n,k}^0, \sum_{j \neq k} \Xi_{n,j}^0 \right) \Xi_{n,k} \right\rangle \right] \end{aligned}$$

which, using (32) and the elementary equality $\langle x, tt^\top x \rangle = |\langle x, t \rangle|^2$,

$$\begin{aligned} &= \frac{1}{2} \int_0^1 \left(\int_0^1 \sum_{k=1}^n \mathbb{E} \left[e^{\sqrt{st}} \left(\sum_{j \neq k} \Xi_{n,j} \right) \left[e^{\sqrt{st}} (r \Xi_{n,k}) - 1 \right] |\langle \Xi_{n,k}, t \rangle|^2 \right] dr \right. \\ & \quad \left. - \sum_{k=1}^n \mathbb{E} \left[e^{\sqrt{st}} \left(\sum_{j \neq k} \Xi_{n,j}^0 \right) \left[e^{\sqrt{st}} (\Xi_{n,k}^0) - 1 \right] |\langle \Xi_{n,k}, t \rangle|^2 \right] \right) e^{-\frac{1}{2}(1-s)|t|^2} ds \end{aligned}$$

proving the desired formula. \square

Proposition 3.6 is crucial for the proof of Theorem 2.5. We need one more lemma.

Lemma 3.7. *Fix $t \in \mathbb{R}^N$, $r, s \in [0, 1]$ and $\epsilon > 0$. Then*

$$\sum_{k=1}^n \mathbb{E} \left[\left| e^{\sqrt{st}} (r \Xi_{n,k}) - 1 \right| |\Xi_{n,k}|^2 \right] \leq \epsilon N + 2 \sum_{k=1}^n \mathbb{E} \left[|\Xi_{n,k}|^2; |\langle \Xi_{n,k}, t \rangle| > \epsilon \right] \quad (34)$$

and

$$\sum_{k=1}^n \mathbb{E} \left[\left| e^{\sqrt{st}} (\Xi_{n,k}^0) - 1 \right| |\Xi_{n,k}|^2 \right] \leq \epsilon N + 2 \sum_{k=1}^n \mathbb{E} \left[|\Xi_{n,k}|^2; |\langle \Xi_{n,k}^0, t \rangle| > \epsilon \right]. \quad (35)$$

Proof. The calculation

$$\begin{aligned} & \sum_{k=1}^n \mathbb{E} \left[\left| e^{\sqrt{st}} (r \Xi_{n,k}) - 1 \right| |\Xi_{n,k}|^2 \right] \\ &= \sum_{k=1}^n \mathbb{E} \left[\left| \exp(-i\sqrt{sr} \langle t, \Xi_{n,k} \rangle) - 1 \right| |\Xi_{n,k}|^2; |\langle \Xi_{n,k}, t \rangle| \leq \epsilon \right] \\ & \quad + \sum_{k=1}^n \mathbb{E} \left[\left| \exp(-i\sqrt{sr} \langle t, \Xi_{n,k} \rangle) - 1 \right| |\Xi_{n,k}|^2; |\langle \Xi_{n,k}, t \rangle| > \epsilon \right] \\ &\leq \epsilon \sum_{k=1}^n \mathbb{E} \left[|\Xi_{n,k}|^2 \right] + 2 \sum_{k=1}^n \mathbb{E} \left[|\Xi_{n,k}|^2; |\langle \Xi_{n,k}, t \rangle| > \epsilon \right] \end{aligned}$$

(Lemma 2.2)

$$= \epsilon N + 2 \sum_{k=1}^n \mathbb{E} \left[|\Xi_{n,k}|^2; |\langle \Xi_{n,k}, t \rangle| > \epsilon \right]$$

proves (34). The proof of (35) is similar. \square

Proof of Theorem 2.5. Applying (33) gives

$$\begin{aligned} & |\phi_{\Xi}(t) - \phi_{\Sigma_n}(t)| \\ &= |\mathbb{E}[e_t(\Xi) - e_t(\Sigma_n)]| \\ &= \left| \frac{1}{2} \int_0^1 \left(\int_0^1 \sum_{k=1}^n \mathbb{E} \left[e_{\sqrt{st}} \left(\sum_{j \neq k} \Xi_{n,j} \right) \left[e_{\sqrt{st}}(r\Xi_{n,k}) - 1 \right] |\langle \Xi_{n,k}, t \rangle|^2 \right] dr \right. \right. \\ &\quad \left. \left. - \sum_{k=1}^n \mathbb{E} \left[e_{\sqrt{st}} \left(\sum_{j \neq k} \Xi_{n,j}^0 \right) \left[e_{\sqrt{st}}(\Xi_{n,k}^0) - 1 \right] |\langle \Xi_{n,k}, t \rangle|^2 \right] \right) e^{-\frac{1}{2}(1-s)|t|^2} ds \right| \end{aligned}$$

and this is, by the Cauchy-Schwarz Inequality,

$$\begin{aligned} &\leq \frac{1}{2} |t|^2 \int_0^1 \left(\int_0^1 \sum_{k=1}^n \mathbb{E} \left[\left| e_{\sqrt{st}}(r\Xi_{n,k}) - 1 \right| |\Xi_{n,k}|^2 \right] dr \right. \\ &\quad \left. + \sum_{k=1}^n \mathbb{E} \left[\left| e_{\sqrt{st}}(\Xi_{n,k}^0) - 1 \right| |\Xi_{n,k}|^2 \right] \right) e^{-\frac{1}{2}(1-s)|t|^2} ds \end{aligned}$$

which, by (34) and (35), for any $\epsilon > 0$,

$$\begin{aligned} &\leq 2\epsilon N + 2 \left(\sum_{k=1}^n \mathbb{E} \left[|\Xi_{n,k}|^2; |\langle \Xi_{n,k}, t \rangle| > \epsilon \right] \right. \\ &\quad \left. + \sum_{k=1}^n \mathbb{E} \left[|\Xi_{n,k}|^2; |\langle \Xi_{n,k}^0, t \rangle| > \epsilon \right] \right) \frac{1}{2} |t|^2 \int_0^1 e^{-\frac{1}{2}(1-s)|t|^2} ds \\ &\quad \text{(perform the change of variables } u = \frac{1}{2}(1-s)|t|^2) \\ &= 2\epsilon N + 2 \left(\sum_{k=1}^n \mathbb{E} \left[|\Xi_{n,k}|^2; |\langle \Xi_{n,k}, t \rangle| > \epsilon \right] \right. \\ &\quad \left. + \sum_{k=1}^n \mathbb{E} \left[|\Xi_{n,k}|^2; |\langle \Xi_{n,k}^0, t \rangle| > \epsilon \right] \right) \left(1 - \exp \left(-\frac{1}{2} |t|^2 \right) \right). \end{aligned}$$

Since

$$\begin{aligned} &\sup_{\epsilon > 0} \sup_{t \in \mathbb{R}^N} \limsup_{n \rightarrow \infty} \sum_{k=1}^n \mathbb{E} \left[|\Xi_{n,k}|^2; |\langle \mathbf{H}_{n,k}, t \rangle| > \epsilon \right] \\ &= \sup_{t \in \mathbb{R}^N} \limsup_{n \rightarrow \infty} \sum_{k=1}^n \mathbb{E} \left[|\Xi_{n,k}|^2; |\langle \mathbf{H}_{n,k}, t \rangle| > 1 \right] \end{aligned}$$

$$= L(\{\Xi_{n,k}\}, \{\mathbf{H}_{n,k}\}),$$

the previous calculation establishes (17), finishing the proof of Theorem 2.5. \square

REFERENCES

- [AMS82] Amir D., Mach J. and Saatkamp K. *Existence of Chebyshev centers, best n -nets and best compact approximants* Trans. Amer. Math. Soc. 271 (1982) 513–520
- [B90] Barbour, A. D. *Stein’s method for diffusion approximations*. Probab. Theory Related Fields 84 (1990), no. 3, 297–322.
- [BH84] Barbour, A. D.; Hall, P. *Stein’s method and the Berry-Esseen theorem*. Austral. J. Statist. 26 (1984), no. 1, 8–15.
- [B03] Bentkus, V. *On the dependence of the Berry-Esseen bound on dimension*. J. Statist. Plann. Inference 113 (2003), no. 2, 385–402.
- [BLV11] Berckmoes, B.; Lowen, R.; Van Casteren, J. *Approach theory meets probability theory*. Topology Appl. 158 (2011), no. 7, 836–852.
- [BLV11’] Berckmoes, B.; Lowen, R.; Van Casteren, J. *Distances on probability measures and random variables*. J. Math. Anal. Appl. 374 (2011) 412–428.
- [BLV13] Berckmoes, B.; Lowen, R.; Van Casteren J. *An isometric study of the Lindeberg-Feller CLT via Stein’s method* J. Math. Anal. Appl. 405 (2013) 484–498
- [BH10] Bhattacharya, R.N.; Holmes, S. *An exposition of Götze’s Estimation of the Rate of Convergence in the Multivariate Central Limit Theorem*. Technical Report, Stanford University, 2010.
- [CM08] Chatterjee, S.; Meckes, E. *Multivariate normal approximation using exchangeable pairs*. ALEA Lat. Am. J. Probab. Math. Stat. 4 (2008), 257–283.
- [CF] Chen L.H.Y.; Fang X. *Multivariate Normal Approximation by Stein’s Method: The Concentration Inequality Approach* arXiv:1111.4073.
- [CS01] Chen, L. H. Y.; Shao, Q.-M. *A non-uniform Berry-Esseen bound via Stein’s method*. Probab. Theory Related Fields 120 (2001), no. 2, 236–254.
- [E72] Edelman M. *The construction of an asymptotic center with a fixed-point property* Bull. Am. Math. Soc. 78 (1972) 206–208
- [F68] Feller, W. *On the Berry-Esseen theorem*. Z. Wahrscheinlichkeitstheorie und Verw. Gebiete 10 (1968) 261–268.
- [F71] Feller, W. *An introduction to probability theory and its applications. Vol. II*. Second edition John Wiley & Sons, Inc., New York-London-Sydney 1971.
- [GR96] Goldstein, L.; Rinott, Y. *Multivariate normal approximations by Stein’s method and size bias couplings*. J. Appl. Probab. 33 (1996), no. 1, 1–17.
- [G91] Götze, F. *On the rate of convergence in the multivariate CLT*. Ann. Probab. 19 (1991), no. 2, 724–739.
- [GJT02] Goudon, T.; Junca, S.; Toscani, G. *Fourier-based distances and Berry-Esseen like inequalities for smooth densities*. Monatsh. Math. 135 (2002), no. 2, 115–136.
- [L75] Loh, W.Y. *On the normal approximation for sums of mixing random variables*. Master Thesis, Department of Mathematics, University of Singapore (1975).
- [L97] Lowen, R. *Approach spaces. The missing link in the topology-uniformity-metric triad*. Oxford Mathematical Monographs. Oxford Science Publications. The Clarendon Press, Oxford University Press, New York, 1997.
- [L15] Lowen, R. *Index Analysis: Approach Theory at Work* Springer Monographs in Mathematics, Springer Verlag London, 2015.
- [M09] Meckes, E. *On Stein’s method for multivariate normal approximation*. High dimensional probability V: the Luminy volume, 153–178, Inst. Math. Stat. Collect., 5, Inst. Math. Statist., Beachwood, OH, 2009.
- [NPR10] Nourdin, I.; Peccati, G.; Réveillac, A. *Multivariate normal approximation using Stein’s method and Malliavin calculus*. Ann. Inst. Henri Poincaré Probab. Stat. 46 (2010), no. 1, 45–58.

- [RR09] Reinert, G.; Röllin, A. *Multivariate normal approximation with Stein's method of exchangeable pairs under a general linearity condition*. Ann. Probab. 37 (2009), no. 6, 2150–2173.
- [RR96] Rinott, Y.; Rotar, V. *A multivariate CLT for local dependence with $n^{-1/2} \log n$ rate and applications to multivariate graph related statistics*. J. Multivariate Anal. 56 (1996), no. 2, 333–350.
- [Sh10] Shevtsova, I. G. *Refinement of estimates for the rate of convergence in Lyapunov's theorem*. (Russian) Dokl. Akad. Nauk 435 (2010), no. 1, 26–28; translation in Dokl. Math. 82 (2010), no. 3, 862–864.
- [S11] Stroock, D. W. *Probability theory. An analytic view*. Second edition. Cambridge University Press, Cambridge, 2011.