The Annals of Probability
1981, Vol. 9, No. 4, 611–623

THE MULTIDIMENSIONAL CENTRAL LIMIT THEOREM FOR ARRAYS NORMED BY AFFINE TRANSFORMATIONS

BY MARJORIE G. HAHN[1] AND MICHAEL J. KLASS[2]

Tufts University and University of Calfornia, Berkeley

Let $X_{n1}, \cdots, X_{nk_n}$ be independent random vectors in $\mathbb{R}^d$. Necessary and sufficient conditions are found for the existence of linear operators A_n on $\mathbb{R}^d$ such that $\mathscr{L}(A_n(\sum_{j=1}^{k_n} X_{nj})) \to N(\vec{0}, I)$, where I is the $d \times d$ identity covariance matrix. These results extend the authors' previous work on sums of i.i.d. random vectors. The proof of the main theorem is constructive, yielding explicit centering vectors and norming linear operators.

0. Introduction. For each $n \geq 1$, let $X_{n1}, \cdots, X_{nk_n}$ be independent d-dimensional random vectors with partial sums $S_n = \sum_{j=1}^{k_n} X_{nj}$. This paper considers how and when S_n can be normalized so as to converge weakly to a d-variate normal limit. Normalizing S_n by constants or even componentwise will often result in either a subprobability limit distribution or else one which is not full, i.e., degenerate in that it concentrates on a proper subspace of $\mathbb{R}^d$. Hence the random vectors S_n which can be successfully normed in this manner to converge to a full d-variate normal limit is rather restrictive. To enlarge the class of normalizable distributions we norm by affine transformations.

We present necessary and sufficient conditions (NASC) for the existence of linear transformations T_n and vectors v_n such that $(T_n(S_n - v_n)) \to N(\vec{0}, I)$ (any limit $N(\vec{0}, \Sigma)$ may be obtained upon replacing T_n by $\Sigma^{1/2}T_n$, where $\Sigma^{1/2}$ denotes the positive definite symmetric square root of Σ). This result generalizes Hahn and Klass (1980a), where NASC were obtained in the i.i.d. mean zero case.

To achieve affine norming, a uniform symmetrized version of the 1-dimensional limit condition is required (see Remark 2 below for a discussion of the symmetrization). The uniformity is essential even for i.i.d. random vectors (see Example 4 of Hahn and Klass (1980)). Presumably, the uniformity constraint was unnoticed earlier because it is superfluous for prenormalized arrays.

Classical NASC when $d = 1$ can be found in, e.g., Gnedenko and Kolmogorov ((1968), page 121). Varadhan obtains an extension of these conditions to Hilbert space for u.a.n. random elements when $T_n = I$ (see, e.g., Parthasarathy (1967) page 200). Several authors have considered the case in which $E\|X_{nj}\|^2 < \infty$ for all n and j and where the T_n are derived from covariance operators. In Euclidian space this amounts to determining when it is possible to norm by $T_n = (\mathrm{Cov}(S_n))^{-1/2}$ and $v_n = ES_n$. Bhattacharya and Rao (1976), Corollary 18.2 page 183, give a multidimensional analogue of the Lindeberg-Feller conditions. A Hilbert space extension is provided by Kandelaki and Sazonov (1964).

However, even in the finite variance case on the real line expectations and standard deviations need not yield appropriate centering and norming constants. This occurs when ES_n does not adequately reflect the center of the S_n-distribution and $\sigma(S_n)$ does not reflect the magnitude of its range. An approach to handling this difficulty in $\mathbb{R}$ appears in Gnedenko and Kolmogorov ((1968) page 121). The problems are slightly compounded in higher dimensional spaces. Theorem 1 (below) provides a solution to these problems in Euclidian space.

Received May 4, 1979; revised May, 1980.
[1] Supported in part by NSF Grant MCS-78-02417.
[2] Supported in part by NSF Grant MCS-75-10376-A01
AMS 1979 *subject classification.* 60F05.
Key words and phrases. Central limit theorem, triangular array, multivariate normal, operator normalization, translated trimming.

To motivate our solution to the centering and norming problem, think of $\mu_n = \mathscr{L}(S_n)$ as a basis-free probability measure on $\mathbb{R}^d$. Corresponding to each coordinate system in $\mathbb{R}^d$ there is a random vector V_n with $\mathscr{L}(V_n) = \mu_n$. To produce a canonical random vector, at stage n we identify a preferred basis which is, in fact, an orthonormal basis (ONB). The canonical random vector representing μ_n is therefore obtained by application of a unitary operator to S_n. With this new random vector it is sufficient to use componentwise centering and norming. (Details for our 1-dimensional centering and norming appear in Section 1.) This is tantamount to applying a single affine transformation A_n of rather special type to the original random vector S_n. An earlier paper (Hahn and Klass (1980)) implicitly utilizes this norming concept to treat the i.i.d. mean zero case, where no centering difficulties arise. Related results on operator norming in Hilbert space appear in Hahn (1978).

In the i.i.d. mean zero case, norming by the sequence of constants $\{\sqrt{n}\}$ will suffice whenever $E\|X\|^2 < \infty$. Applied statisticians may therefore regard operator norming as little more than idle curiosity. In the present context, however, the finite second moment case is the general case (whence use of affine transformations is often essential). To see this, simply note that appropriate truncations will produce sums of independent bounded random vectors having the same affine normed limit laws as the original variates. (e.g., for triangular arrays, replace X_{nj} by $X'_{nj} \equiv X_{nj}I_{(\|X_{nj}\| \leq b_{nj})}$, where $\sum_{j=1}^{k_n} P(\|X_{nj}\| > b_{nj}) \to 0$. For sums $S_n = X_1 + \cdots + X_n$ with the property that $(S_n - v_n, \theta)$ does not converge weakly for any unit vector θ and centering sequence v_n, one may replace X_j by $X'_j = X_jI_{(\|X_j\| \leq b_j)}$ where $\sum_{j=1}^{n} P(\|X_j\| > b_j) < \infty$.

The following simple example illustrates the need for linear operators in the nonidentically distributed case. Operator norming is necessitated by the fact that there exist two separate directions whose norming constants have different growth rates.

EXAMPLE 1. Let $X_1, Y_1, X_2, Y_2, \cdots$ be independent mean zero random variables with finite second moments. Let $s_n^2(X) = \sum_{k=1}^n EX_k^2$, $s_n^2(Y) = \sum_{k=1}^n EY_k^2$. Assume $\lim_{n\to\infty} s_n(X)/s_n(Y) = 0$ and standard deviations give appropriate norming constants for the nth partial sums of the X_i's and Y_i's. Finally let $S_n = \sum_{k=1}^n (U_k, V_k)$ where $(U_k, V_k) = (X_k \cos\phi + Y_k \sin\phi, -X_k \sin\phi + Y_k \cos\phi)$ for a fixed $\phi \neq 0 \pmod{\pi/2}$. In order to obtain weak convergence to a nondegenerate 2-dimensional limit it is clear that S_n must be normalized componentwise along the preferred ONB $\{(\cos\tilde{\phi}, -\sin\tilde{\phi}), (-\sin\tilde{\phi}, -\cos\tilde{\phi})\}$ where $\tilde{\phi} \equiv \phi \pmod{\pi/2}$.

An easy example of this phenomena is obtained by letting $\mathscr{L}(X_k) = N(0, 1)$ and $\mathscr{L}(Y_k) = N(0, k)$.

More elaborate examples can be constructed in which the preferred ONB varies with n, (see Example 2 of Hahn and Klass 1980)).

As a notational convention, throughout the paper for any random vector X, $\bar{X}$ will denote an independent copy of X which is also independent of all other vectors in sight. Also let $X^s = X - \bar{X}$.

THEOREM 1. *Let $X_{n1}, \cdots, X_{nk_n}$ be a triangular array of rowwise independent d-dimensional random vectors. For any $\theta \in \mathbb{R}^d$ and $n \geq 1$, define*

$$a_n(\theta) = \inf\{a > 0: 2 \geq \textstyle\sum_{j=1}^{k_n} E((\langle X_{nj}^s, \theta\rangle/a)^2 \wedge 1)\}. \tag{1}$$

Let $S_n = \sum_{j=1}^{k_n} X_{nj}$. Then there exist linear transformations T_n and d-dimensional vectors v_n uch that

$$\mathscr{L}(T_n(S_n - v_n)) \to N(\vec{0}, I) \tag{2}$$

and

(3) *for every* $\epsilon > 0$,

$$\lim_{n\to\infty} \max_{1\leq j\leq k_n} \sup_{\|\theta\|=1} P(|\langle T_nX_{nj}^s, \theta\rangle| > \epsilon) = 0$$

iff for every $\epsilon > 0$,

(4) $$\lim_{n\to\infty} \sup_{\|\theta\|=1} \sum_{j=1}^{k_n} P(|\langle X_{nj}^s, \theta\rangle| > \epsilon a_n(\theta)) = 0$$

and there exists an n_0 *such that for all* $n \geq n_0$

(5) $$\inf_\theta \{a_n(\theta): \|\theta\| = 1\} > 0.$$

Moreover, whenever (5) *holds, the* $a_n(\theta)$ *are continuous for all* $n \geq n_0$. *Consequently, by compactness of the unit sphere there exists an orthonormal basis* $\{\theta_{nj}, j = 1, \cdots, d\}$ *such that*

$$a_n(\theta_{n1}) = \inf\{a_n(\theta): \|\theta\| = 1\}$$
$$a_n(\theta_{nj}) = \inf\{a_n(\theta): \|\theta\| = 1, \theta \perp \operatorname{span}\{\theta_{nk}, 1 \leq k \leq j-1\}\} \qquad \text{for } 1 < j \leq d.$$

The linear operators T_n *may be taken to be of the form*

(6) $$T_n x = \sum_{j=1}^d (\langle x, \theta_{nj}\rangle / a_n(\theta_{nj})) e_j.$$

Finally, the centering constants v_n *are determined by setting*

(7) $$\langle v_n, \theta_{nj}\rangle = \sum_{i=1}^{k_n} E\langle X_{ni}, \theta_{nj}\rangle\, A''_{nij}$$

where $A_{nij} = [\operatorname{med}(\langle X_{ni}, \theta_{nj}\rangle) - a_n(\theta_{nj}), \operatorname{med}(\langle X_{ni}, \theta_{nj}\rangle) + a_n(\theta_{nj})]$ *and* $Y''_A = YI_{(Y\notin A)} +$ *(midpoint of* $A)I_{(Y\notin A)}$.

REMARK 1. Notice that $a_n(\theta)$ is the largest real number satisfying the implicit relation

$$2a_n^2(\theta) = \sum_{j=1}^{k_n} E(\langle X_{nj}^s, \theta\rangle^2 \wedge a_n^2(\theta)).$$

Moreover, $a_n(\theta) > 0$ iff $\sum_{j=1}^{k_n} P(|\langle X_{nj}^s, \theta\rangle| > 0) > 2$.

REMARK 2. We give a word of caution to those who wish to restate the theorem in terms of unsymmetrized random vectors. Specialize to the 1-dimensional case. The purpose of symmetrization is to produce continuous easily computable norming constants which play the role of standard deviations. The following example illustrates that failure to symmetrize or recenter the variables can easily result in quantities totally inappropriate for norming. For $j = 1, 2, \cdots$ let $X_{2j} = Y_{2j} + b_j$ and $X_{2j-1} = Y_{2j-1} - b_j$. Then let a_n and c_n be the largest reals satisfying $a_n^2 = \sum_{j=1}^n E(X_j^2 \wedge a_n^2)$ and $c_n^2 = \sum_{j=1}^n E(Y_j^2 \wedge c_n^2)$. Norm $S_n = \sum_{j=1}^n X_j$ by a_n and $T_n = \sum_{j=1}^n Y_j$ by c_n. Suppose the Y_j's are i.i.d. and the b_j's grow so rapidly that $c_{2n}/a_{2n} \to 0$. Thus a_{2n} and c_{2n} are not asymptotic. However, since S_{2n} and T_{2n} are equal, both distributions require asymptotically equivalent norming sequences to obtain identical nondegenerate limits.

REMARK 3. The linear operators T_n are constructed in such a way that the marginals of $T_n(S_n - v_n)$ converge weakly at a uniform rate to those of the standard d-variate normal. This means, in particular, that the marginals must be asymptotically independent along the preferred orthonormal basis directions $\{\theta_{nj}, j = 1, \cdots, d\}$. This is, in fact, the main step of the proof.

REMARK 4. The norming constants $a_n(\theta)$ are essentially determined by $a_n(\theta_{n1}), \cdots, a_n(\theta_{nd})$. To see this, let

$$b_n^2(\theta) = \sum_{j=1}^d a_n^2(\theta_{nj})\, \langle \theta_{nj}, \theta\rangle^2.$$

Since $\theta_{n1}, \cdots, \theta_{nd}$ form an ONB, $\mathscr{L}(\langle S_n^s, \theta\rangle / b_n(\theta)) =$

$$\mathscr{L}(\textstyle\sum_{j=1}^d (\langle S_n^s, \theta_{nj}\rangle / b_n(\theta_{nj}))(b_n(\theta_{nj})\, \langle \theta_{nj}, \theta\rangle / b_n(\theta))) \to N(0, 2)$$

because

$$\mathscr{L}((\langle S_n^s, \theta_{n1}\rangle/b_n(\theta_{n1}), \cdots, \langle S_n^s, \theta_{nd}\rangle/b_n(\theta_{nd}))) \to N(\vec{0}, 2I)$$

and

$$(b_n(\theta_{n1})\langle\theta_{n1}, \theta\rangle/b_n(\theta), \cdots, b_n(\theta_{nd})\langle\theta_{nd}, \theta\rangle/b_n(\theta))$$

is a unit vector. Also $\mathscr{L}(\langle S_n^s, \theta\rangle/a_n(\theta)) \to N(0, 2)$. So, by the convergence-of-types theorem, $\lim_{n\to\infty} a_n(\theta)/b_n(\theta) = 1$. Since this argument works equally well for any sequence of θ's, the convergence is, in fact, uniform in θ.

REMARK 5. Theorem 1 suggests one method of constructing suitable affine norming transformations. It is possible to norm slightly more classically. To do so, at stage n simply t-trim (see Section 1 for an explanation of this procedure) and center the individual variables along the preferred ONB and then norm by the positive definite symmetric square root of the inverse covariance operator of the sum of the new variates. If one wishes to norm based on inverse covariance operators of modified variates, the truncation procedure adopted is crucial. For sums of precentered random vectors, t-trimming amounts to discarding mass from the individual summands (at stage n) lying outside a d-dimensional rectangle whose sides are parallel to the preferred ONB. An example to appear in Hahn and Klass (1981) shows that truncation outside spheres (instead of rectangles) can produce operators whose inverse square root is inappropriate for norming.

Since reference will often be made to the normal convergence criterion (NCC) for 1-dimensional (prenormalized) triangular arrays, we state it here for easy reference.

NCC. If W_{nj} are independent summands for $j = 1, \cdots, k_n$, then for every $\epsilon > 0$,

$$\mathscr{L}(\textstyle\sum_{j=1}^{k_n} W_{nj}) \to N(0, 1)$$

and

$$\lim_{n\to\infty} \max_{1\le j\le k_n} P(|W_{nj}| > \epsilon) = 0 \quad \text{(u.a.n. condition)}$$

iff for every $\epsilon > 0$ and a $\tau > 0$

$$\text{(i) } \lim_{n\to\infty} \textstyle\sum_{j=1}^{k_n} P(|W_{nj}| > \epsilon) = 0$$

$$\text{(ii) } \lim_{n\to\infty} \textstyle\sum_{j=1}^{k_n} EW_{nj}I_{(|W_{nj}|\le\tau)} = 0$$

$$\text{(iii) } \lim_{n\to\infty} \textstyle\sum_{j=1}^{k_n} \mathrm{Var}(W_{nj}I_{(|W_{nj}|\le\tau)}) = 1.$$

1. Centering and norming constants for arrays of independent random variables. The convergence-of-types theorem says that if there exists constants b_n and a_n and a random variable Z such that

$$\mathscr{L}\left(\frac{S_n - b_n}{a_n}\right) \to \mathscr{L}(Z)$$

then the same result holds when b_n and a_n are replaced by b_n' and a_n' if and only if

$$\lim_{n\to\infty} a_n/a'_n = 1 \qquad \text{and} \qquad \lim_{n\to\infty} (b_n - b_n')/a_n = 0.$$

Thus, once one sequence of norming and centering constants is known, the behavior of all others is determined. We now provide a general approach to norming and centering.

To construct an appropriate norming sequence $\{a_n\}$, note that a_n must be reflective of the "range" of the distribution of S_n. With this interpretation in mind, a_n must also be reflective of the "range" of S_n^s. Consideration of the symmetrized variables S_n^s avoids all centering difficulties. Indeed, Theorem 2 verifies that a canonical norming sequence $\{a_n\}$ may be constructed from sums of the symmetrized variables which comprise S_n^s.

The centering problem is more delicate. Ideally, we would like to center at expectations. To do so we need to find $\tilde{S}_n$ such that

$$(S_n - \tilde{S}_n)/a_n \to_{\Pr} 0 \quad \text{and} \quad \mathscr{L}((\tilde{S}_n - E\tilde{S}_n)/a_n) \to \mathscr{L}(Z).$$

$\tilde{S}_n$ is constructed by transforming the individual summands in S_n into bounded random variables via a method called *t-trimming*. *T*-trimming involves truncating a given random variable Y on a (bounded) set A and placing the excess mass at one or several points. For example, t-trimming Y on A and placing the excess mass associated with A^c at b yields

$$Y_{A,b} = YI_{(Y\in A)} + bI_{(Y\in A^c)}.$$

From the point of view of normal convergence, no more generality is achieved by utilizing several points rather than one. Throughout we will assume A is a bounded interval and

$$Y'_A = Y_{A,0} \quad \text{and} \quad Y''_A = Y_{A,\text{midpoint of } A}.$$

Y'_A is the usual truncated variable while Y''_A is the modification of Y we will need. Subscripts will be deleted whenever A is clear from the context.

Weak convergence of the partial sums of a triangular array $\{X_{nj}, j = 1, \cdots, k_n\}$ of random variables remains invariant under t-trimming if the sets A_{nj} satisfy

$$\lim_{n\to\infty} \textstyle\sum_{j=1}^{k_n} P(X_{nj} \notin A_{nj}) = 0. \tag{8}$$

The natural choice for A_{nj} might be an interval $[a_{nj} - l_n, a_{nj} + l_n]$ of prescribed length $2l_n$, where a_{nj} satisfies

$$P(X_{nj} \in [a_{nj} - l_n, a_{nj} + l_n]) = \sup_{y\in R} P(X_{nj} \in [y - l_n, y + l_n]).$$

Of course, l_n must be chosen to satisfy (8). Since such an a_{nj} is not easily identified, we notice that (8) requires the eventual coverage of every fixed quantile strictly between 0 and 1 of each X_{nj}, simultaneously for all j. Consequently, for convenience, we may let a_{nj} = med X_{nj}, the median of X_{nj} having smallest absolute value.

Since the purpose of t-trimming random variables is to produce variates for which expectations (and also standard deviations) yield appropriate centering (and norming) constants, the excess trimmed mass must be judiciously located and cannot be blindly placed at zero, as that might make expectations and standard deviations unreflective of the true "center" and "range" of the distribution of S_n. Theorem 2 shows that the excess mass may be placed at the median of X_{nj}, the midpoint of the interval used for t-trimming.

In the following example we have central limiting behavior. However, regardless of which intervals are used for t-trimming, if the resultant, mass is placed at zero (yielding standard truncation), the expectations and standard deviations are always of the wrong order of magnitude.

EXAMPLE 2. The idea is to construct $X_1, X_2, \cdots$ such that there exist constants a_n and b_n for which $\mathscr{L}((S_n - b_n)/a_n) \to N(0, 1)$. However, for any bounded intervals I_{nj} of common length l_n with $\sum_{j=1}^n P(X_j \notin I_{nj}) \to 0$, the terms $\sum_{j=1}^n EX'_{j,I_{nj}}$ are the wrong order for centering (we achieve this only for n even) and the terms $(\sum_{j=1}^n \operatorname{Var}(X'_{j,I_{nj}}))^{1/2}$ are the wrong order for norming.

In order to construct X_j, let p_j and c_j be constants such that

(i) $p_1 \geq p_2 \geq \cdots \geq 0$, $\sum_{j=1}^\infty p_j < \infty$;

(ii) $|c_j| > 1$ for $j = 1, 2, \cdots$ and $c_{2j-1} = -c_{2j} > 0$;

(iii) $\sum_{j=1}^{2n-1} c_{2j-1}(p_{2j-1} - p_{2j})/\sqrt{n} \sim c_{2n-1}(p_{2n-1} - p_{2n})/\sqrt{n}$;

(iv) $\sum_{j=1}^{2n-1} c_j^2 p_j \sim c_{2n-1}^2 p_{2n-1}$;

(v) $c_{2n-1}(p_{2n-1} - p_{2n})/\sqrt{n} \to \infty$ as $n \to \infty$.

A specific example is $p_j = (j + 1)^{-2}$, $c_{2j-1} = (2j)!$.

Let $Y_j, j = 1, 2, \cdots$ be independent with $P(Y_j = 1) = P(Y_j = -1) = 1/2$. Define $\lambda_j, j = 1, 2, \cdots$ to be independent of each other and of the Y_j with $p_j = P(\lambda_j = 0) = 1 - P(\lambda_j = 1)$. Now we take

$$X_j = \lambda_j(Y_j + c_j) = \begin{cases} 0 & \text{with probability } p_j \\ c_j + 1 & \text{with probability } (1 - p_j)/2. \\ c_j - 1 & \text{with probability } (1 - p_j)/2 \end{cases}$$

Since $\sum_{j=1}^{\infty} P(X_j \neq Y_j + c_j) = \sum_{j=1}^{\infty} P(\lambda_j = 0) = \sum_{j=1}^{\infty} p_j < \infty$, the partial sums $T_n = \sum_{j=1}^{n} (Y_j + c_j)$ and $S_n = \sum_{j=1}^{n} X_j$ are equivalent sequences. Moreover, $\mathscr{L}((T_n - c_n I_{(n \text{ odd})})/\sqrt{n}) \to N(0, 1)$ which therefore implies $\mathscr{L}((S_n - c_n I_{(n \text{ odd})})/\sqrt{n}) \to N(0, 1)$. Thus, by the convergence of types theorem we know the order of magnitude of the centering and norming constants for the sequence S_n.

Condition (i) implies $P(X_j = c_j + 1) = P(X_j = c_j - 1) = (1 - p_j)/2 \to \frac{1}{2}$ as $j \to \infty$. Consequently, if I_{nj} are intervals of common length l_n for which $\sum_{j=1}^{n} P(X_j \notin I_{nj}) \to 0$, for n and j large we must have $c_j + 1$ and $c_j - 1$ in I_{nj}. This, in particular, means that $X_j = X_j I_{(X_j \in I_{nj})}$ for n and j large. Thus, the example will be complete if we can show that ES_n is not an appropriate centering sequence and $(\text{Var}(S_n))^{1/2}$ is not an appropriate norming sequence.

Now

$$\begin{aligned} ES_{2n}/\sqrt{2n} &= \textstyle\sum_{j=1}^{2n} (1 - p_j)c_j/\sqrt{2n} \\ &= -\textstyle\sum_{j=1}^{n} c_{2j-1}(p_{2j-1} - p_j)/\sqrt{2n} \qquad \text{by (ii)}, \\ &\sim -c_{2n-1}(p_{2n-1} - p_{2n})/\sqrt{2n} \qquad \text{by (iii)}, \\ &\to -\infty \quad \text{as} \quad n \to \infty \quad \text{rather than 0, by (v)}. \end{aligned}$$

Thus, ES_{2n} is not an appropriate $2n$th centering constant.

Also,

$$\begin{aligned} \textstyle\sum_{j=1}^{2n-1} \text{Var}(X_j)/(2n - 1) &= \textstyle\sum_{j=1}^{2n-1} (1 - p_j)(1 + p_j c_j^2)/(2n - 1) \\ &\sim + \textstyle\sum_{j=1}^{2n-1} p_j c_j^2/(2n - 1) \\ &\sim 1 + c_{2n-1}^2 p_{2n-1}/(2n - 1) \qquad \text{by (iv)}, \\ &\to \infty \quad \text{as} \quad n \to \infty \quad \text{by (v) and (i)}. \end{aligned}$$

This phenomenon persists for n even. So $(\sum_{j=1}^{n} \text{Var}(X_j))^{1/2}$ is not an appropriate nth norming constant.

Even if we did not require $\sum_{j=1}^{n} P(X_j \notin I_{nj}) \to 0$, the above method of centering and norming by placing the excess mass at zero would still fail.

For notational convenience, S_n'' will denote $\sum_{j=1}^{k_n} X_{nj}''$ where the set on which each X_{nj} is t-trimmed may vary for each X_{nj} but is either clear from the context or will be explicitly specified.

THEOREM 2. *Let $X_{n1}, \cdots, X_{nk_n}$ be independent random variables with $S_n = \sum_{j=1}^{k_n} X_{nj}$. As usual, let a_n be the largest real number satisfying the implicit relation*

$$2a_n^2 = \textstyle\sum_{j=1}^{k_n} E((X_{nj}^s)^2 \wedge a_n^2).$$

There exist constants b_n and c_n such that

$$\mathscr{L}((S_n - b_n)/c_n) \to N(0, 1) \quad \textit{and} \quad \{X_{nj}^s/c_n\} \textit{ is u.a.n.} \tag{9}$$

iff *there exists $N > 0$ such that*

$$a_n > 0 \textit{ for all } n \geq N, \tag{10}$$

and

$$\mathscr{L}((S_n - ES_n'')/a_n) \to N(0, 1) \quad \textit{and} \quad \{X_{nj}^s/a_n\} \textit{ is u.a.n.} \tag{11}$$

where

$$S_n'' \equiv \textstyle\sum_{j=1}^{k_n} X_{nj,A_{nj}}'' \textit{ with } A_{nj} = [\text{med}(X_{nj}) - a_n, \text{med}(X_{nj}) + a_n].$$

PROOF. Sufficiency is obvious. For necessity, first assume $\text{med}(X_{nj}) = 0$. Now (9) implies

(12) $$\mathscr{L}(S_n^s \sqrt{2}c_n) \to N(0, 1).$$

Consequently, by the NCC, for every $\epsilon > 0$

(13) $$\lim_{n\to\infty} \sum_{j=1}^{k_n} P(| X_{nj}^s | > \epsilon c_n) = 0$$

which, since $\text{med}(X_{nj}) = 0$, immediately implies

(14) $$\lim_{n\to\infty} \sum_{j=1}^{k_n} P(| X_{nj} | > \epsilon c_n) = 0 \quad \text{for every} \quad \epsilon > 0.$$

The NCC also implies that for any $\epsilon > 0$,

$$\lim_{n\to\infty} \sum_{j=1}^{k_n} E \mid X_{nj}^s \mid {}^2 I_{(|X_{nj}| \le \epsilon c_n)}/2c_n^2 = 1$$

which together with (13) yields

(15) $$\lim_{n\to\infty} \sum_{j=1}^{k_n} E(| X_{nj}^s |^2 \wedge (\epsilon c_n)^2)/2c_n^2 = 1.$$

Varying ϵ is a neighborhood of 1 and applying a simple monotonicity argument to (15) imply both the existence of a_n and the fact that

(16) $$\lim_{n\to\infty} a_n/c_n = 1.$$

Hence, by (12), (14), and (16)

(17) $$\mathscr{L}((\sum_{j=1}^{k_n} X_{nj}'' - \sum_{j=1}^{k_n} \bar{X}_{nj}'')/\sqrt{2}\, a_n) \to N(0, 1)$$

where the variables are t-trimmed on $A_{nj} = [-a_n, a_n]$. Thus, in order to conclude (11), it suffices to show that

$$\mathscr{L}((S_n'' - ES_n'')/a_n) \to N(0, 1).$$

This will follow from verification of the conditions in the NCC.

$$\sum_{j=1}^{k_n} P(| X_{nj}'' - EX_{nj}'' | > 4\epsilon a_n)$$
$$\le \sum_{j=1}^{k_n} P(| X_{nj}'' | > 2\epsilon a_n) + \sum_{j=1}^{k_n} P(| EX_{nj}'' | > 2\epsilon a_n).$$

The first sum converges to 0 by (14) and (16). The second sum converges to 0 because for all n large

$$\begin{aligned} | EX_{nj}''| &= | EX_{nj}''(I_{(|X_{nj}| \le \epsilon a_n)} + I_{(|X_{nj}| > \epsilon a_n)}) | \\ &\le \epsilon a_n + a_n P(| X_{nj} | > \epsilon a_n) \\ &\le 2\epsilon a_n. \end{aligned}$$

Since $| X_{nj}''/a_n | \le 1$ and X_{nj}'' is independent of $\bar{X}_{nj}''$, (17) implies that

$$\lim_{n\to\infty} \sum_{j=1}^{k_n} \text{Var}(X_{nj}''/a_n) = 1.$$

Finally, since $E((S_n'' - ES_n'')/a_n) = 0$, the conditions of the NCC are satisfied, completing the proof for median zero random variables.

The general case can, of course, be reduced to the median zero case by letting $U_{nj} = X_{nj} - \text{med}(X_{nj})$ and $\bar{U}_{nj} = \bar{X}_{nj} - \text{med}(\bar{X}_{nj})$. If $S_n = \sum_{j=1}^{k_n} X_{nj}$ then this reduction yields centering constants b_n of the form

$$b_n = \sum_{j=1}^{k_n} b_{nj} = \sum_{j=1}^{k_n} \text{med}(X_{nj}) + \sum_{j=1}^{k_n} EU_{nj,[-a_n,a_n]}'' = \sum_{j=1}^{k_n} EX_{nj,A_{nj}}''$$

where $A_{nj} = [\text{med}(X_{nj}) - a_n, \text{med}(X_{nj}) + a_n]$.

The above technique amounts to centering at medians and then recentering the resultant by subtracting off truncated expectations, ultimately adding back all the subtracted quantities. The idea was exploited in this form in Klass (1980). It motivated the centering notion presented and refined here. However, an essentially equivalent centering method in fact already appears in Gnedenko and Kolmogorov ((1968), Theorem 2, page 121).

Theorem 2 together with the convergence of types theorem indicates that the excess mass may be placed at zero only when

$$\lim_{n\to\infty} \sum_{j=1}^{k_n} \text{med}(X_{nj}) P(| X_{nj} - \text{med}(X_{nj})| > a_n)/a_n = 0.$$

Furthermore, since

$$\lim_{n\to\infty} \sum_{j=1}^{k_n} P(| X_{nj} - \text{med}(X_{nj})| \geq a_n) = 0,$$

no more generality results from placing the excess mass at the two endpoints of the intervals A_{nj}.

2. Proof of Theorem 1. The basic preliminary to our proof of sufficiency is the following real variables lemma.

LEMMA. *Let A_n, B_n, C_n be real numbers with $A_n, B_n > 0$. Then*

$$\lim_{n\to\infty} \inf_\theta (A_n^2 \cos^2\theta + B_n^2 \sin^2\theta + C_n \sin 2\theta)/A_n^2 = 1 \tag{18}$$

iff

$$\liminf_{n\to\infty} B_n/A_n \geq 1 \tag{19}$$

and

$$\lim_{n\to\infty} C_n/A_n B_n = 0. \tag{20}$$

PROOF. *Necessity.* Assume (18). Let

$$\begin{aligned} F(n, \theta) &= A_n^{-2}(A_n^2 \cos^2\theta + B_n^2 \sin^2\theta + C_n \sin 2\theta) - 1 \\ &= (B_n^2 A_n^{-2} - 1)\sin^2\theta + C_n A_n^{-2} \sin 2\theta. \end{aligned}$$

Since,

$$\begin{aligned} 0 = \lim_{n\to\infty} \inf_\theta F(n, \theta) &\leq \liminf_{n\to\infty} F(n, \pi/2) \\ &= \liminf_{n\to\infty} (B_n^2 A_n^{-2} - 1), \end{aligned}$$

(19) holds.

Next suppose (20) is false. Then there exists $0 < \epsilon < 1$ and $n_1 < n_2 < \cdots$ such that $| C_{n_k}|/A_{n_k} B_{n_k} > 2\epsilon$. We may assume $C_{n_k} < 0$. Let $Q = \{k \geq 1: B_{n_k}/A_{n_k} \geq 1 + \epsilon\}$. If Q^c contains infinitely many positive integers

$$\begin{aligned} 0 \leq \liminf_{\substack{k\to\infty \\ k\in Q^c}} \inf_\theta F(n_k, \theta) &\leq \liminf_{\substack{k\to\infty \\ k\in Q^c}} F(n_k, \pi/4) \\ &\leq .5((1 + \epsilon)^2 - 1) - 2\epsilon < 0, \end{aligned}$$

a contradiction. Hence every sufficiently large integer belongs to Q. By redefining $\{n_k\}$ if necessary, we may suppose every $k \geq 1$ belongs to Q.

Thus there exists $0 < \theta_{n_k} \leq \pi/4$ such that

$$\tan \theta_{n_k} = \epsilon/(B_{n_k} A_{n_k}^{-1} - 1).$$

Since $\cos^2\theta_{n_k}$ is bounded away from zero,

$$\begin{aligned} 0 &\leq \liminf_{k\to\infty} F(n_k, \theta_{n_k}) \cos^{-2}\theta_{n_k} \\ &= \liminf_{k\to\infty} (2F(n_k, \theta_{n_k})/\sin 2\theta_{n_k}) \tan \theta_{n_k} \\ &= \liminf_{k\to\infty} (\epsilon(B_{n_k} A_{n_k}^{-1} + 1) + 2C_{n_k} A_{n_k}^{-2}) \tan \theta_{n_k} \\ &\leq \liminf_{k\to\infty} (\epsilon(B_{n_k} A_{n_k}^{-1} + 1) - 2\epsilon B_{n_k} A_{n_k}^{-1}) \tan \theta_{n_k} \\ &= -\epsilon^2 < 0, \qquad \text{a contradiction.} \end{aligned}$$

Hence (20) holds.

Sufficiency. Assume (19) and (20) hold. $F(n, 0) \equiv 0$ so if (18) is false, there exist $0 < \epsilon < 1$, $1 < n_1 < n_2 < \cdots$, and θ_{n_k} such that $F(n_k, \theta_{n_k}) < -2\epsilon$. In view of (19), $| C_{n_k} | A_{n_k}^{-2} > \epsilon$ for all k large. Applying (20) we see that $\lim_{k\to\infty} B_{n_k}/A_{n_k} = \infty$. Hence for all k large,

$$-2\epsilon > F(n_k, \theta_{n_k}) > (.5B_{n_k}^2 A_{n_k}^{-2} | \sin \theta_{n_k} | - 2 | C_{n_k} | A_{n_k}^{-2})| \sin \theta_{n_k} |.$$

Clearly, $| \sin \theta_{n_k} | \le 4 | C_{n_k} | B_{n_k}^{-2}$. Therefore,

$$-2\epsilon > -2 | C_{n_k} | A_{n_k}^{-2} 4 | C_{n_k} | B_{n_k}^{-2} = -8(C_{n_k}/A_{n_k} B_{n_k})^2,$$

which tends to 0 by (20), yielding a contradiction which proves (18).

We are now ready to prove Theorem 1.

Proof of Sufficiency in Theorem 1. For notational convenience we restrict to $X_{nj} = X_j$ and $k_n = n$. The following proof extends to arrays by merely changing the subscripts.

Suppose (4) and (5) hold and that $n \ge n_0$. Now (5) insures that $a_n(\theta) > 0$ for all $n \ge n_0$. $a_n(\theta)$ also satisfies the implicit relation in Remark 1. So the continuity of $\theta \to a_n(\theta)$ can be seen as follows: For fixed n, $M_n(y, \theta) \equiv \sum_{j=1}^n E(\langle X_j^s, \theta\rangle^2 y^{-2} \wedge 1)$ is jointly continuous in θ and y and strictly monotone in y for

$$y \geqq \sup_{\|\theta\|=1} \sup\{z \geqq 0: \textstyle\sum_{j=1}^n P(0 < | \langle X_j, \theta\rangle | \le z) = 0\}.$$

Consequently, for fixed θ^* and each $2\epsilon > 0$ sufficiently small, continuity and strict monotonicity in y implies the existence of $y_1(\epsilon) < a_n(\theta^*) < y_2(\epsilon)$ such that for $i = 1, 2$,

$$M_n(y_i(\epsilon), \theta^*) = 2(1 - (-1)^i \epsilon)$$

and

$$\lim_{\epsilon\downarrow 0} y_i(\epsilon) = a_n(\theta^*).$$

Continuity in θ of $M_n(y, \theta)$, implies the existence of $\delta_\epsilon > 0$ such that whenever $| \theta - \theta^* | < \delta_\epsilon$ we have $| M_n(y_i(\epsilon), \theta) - M_n(y_i(\epsilon), \theta^*)| < \epsilon$ for $i = 1, 2$. Hence, if $y \le y_1(\epsilon)$ and $| \theta - \theta^* | < \delta_\epsilon$, $M_n(y, \theta) \ge M_n(y_1(\epsilon), \theta) \ge M_n(y_1(\epsilon), \theta^*) - \epsilon = 2 + \epsilon$. Also if $y \ge y_2(\epsilon)$, $M_n(y, \theta) \le M_n(y_2(\epsilon), \theta) \le M_n(y_2(\epsilon), \theta^*) + \epsilon = 2 - \epsilon$. Since $M_n(a_n(\theta), \theta) \equiv 2$, it follows that $\theta \to \theta^*$ implies $a_n(\theta) \to a_n(\theta^*)$. As a result, the $a_n(\theta_{nj}), j = 1, \cdots, d$ and hence the T_n and d_n are well-defined.

We prove sufficiency using the particular T_n and d_n defined in (6) and (7). Conditions (4) and (5) easily imply (3). Indeed, for every $\epsilon > 0$,

$$\begin{aligned}
&\lim_{n\to\infty} \max_{1\le j\le n} \sup_{\|\theta\|=1} P(| \langle T_n X_j^s, \theta\rangle | > \epsilon) \\
&\quad = \lim_{n\to\infty} \max_{1\le j\le n} \sup_{\|\theta\|=1} P(| \textstyle\sum_{i=1}^d (\langle X_j^s, \theta_{ni}\rangle \langle e_i, \theta\rangle / a_n(\theta_{ni})) | > \epsilon) \\
&\quad \le \lim_{n\to\infty} \max_{1\le j\le n} \textstyle\sum_{i=1}^d P(| \langle X_j^s, \theta_{ni}\rangle | > \epsilon a_n(\theta_{ni})/\sqrt{d}) \qquad \text{by Cauchy-Schwarz} \\
&\quad \le \lim_{n\to\infty} \sup_{\|\theta\|=1} d \textstyle\sum_{j=1}^n P(| \langle X_j^s, \theta\rangle | > \epsilon a_n(\theta)/\sqrt{d}) \\
&\quad = 0 \text{ by (4).}
\end{aligned}$$

The proof that (4) and (5) imply (2) is more complicated. Our first step involves t-trimming and centering to reduce to uniformly bounded, mean zero components. By (4) there exist $\epsilon_n \to 0$ such that

$$(4') \qquad \lim_{n\to\infty} \sup_{\|\theta\|=1} \textstyle\sum_{j=1}^n P(| \langle X_j^s, \theta\rangle | > \epsilon_n a_n(\theta)) = 0.$$

Let $B_{nj\theta} = [\text{med}(\langle X_j, \theta\rangle) - \epsilon_n a_n(\theta), \text{med}(\langle X_j, \theta\rangle) + \epsilon_n a_n(\theta)]$ and let $A_{nj\theta}$ be defined as $B_{nj\theta}$ with ϵ_n replaced by 1. Define $Y_{nj\theta} = \langle X_j, \theta\rangle''_{B_{nj\theta}} - E\langle X_j, \theta\rangle''_{B_{nj\theta}}$.

Step 1. (2) is implied by (4), (5) and

(21) $$\mathscr{L}(\sum_{i=1}^{d} \sum_{j=1}^{n} (Y_{nj\theta_n}/a_n(\theta_n))e_i) \to N(\vec{0}, I).$$

Proof of Step 1. Let $\tilde{Y}_{nj\theta} = \langle X_j, \theta\rangle - E\langle X_j, \theta\rangle''_{B_{nj\theta}}$. Observe that (4′) implies

(22) $$\lim_{n\to\infty} P(\sum_{i=1}^{d} \sum_{j=1}^{n} (Y_{nj\theta_{ni}}/a_n(\theta_{ni}))e_i \neq \sum_{i=1}^{d} \sum_{j=1}^{n} (\tilde{Y}_{nj\theta_{ni}}/a_n(\theta_{ni}))e_i) = 0.$$

Consequently, (21) together with (22) imply

$$\mathscr{L}(\sum_{i=1}^{d} \sum_{j=1}^{n} (\tilde{Y}_{nj\theta_{ni}}/a_n(\theta_{ni}))e_i) \to N(\vec{0}, I).$$

Finally, applying Theorem 2 componentwise yields

$$\mathscr{L}(\sum_{i=1}^{d} \sum_{j=1}^{n} ((\langle X_j, \theta_{ni}\rangle - E\langle X_j, \theta_{ni}\rangle''_{A_{nj\theta_{ni}}})/a_n(\theta_{ni}))e_i) \to N(\vec{0}, I)$$

which is (2) with T_n and v_n as given in (6) and (7).

Since using symmetrized variates doubles the correlation we claim it suffices to prove, for all $1 \leqq i < k \leqq d$,

(23) $$\lim_{n\to\infty} \sum_{j=1}^{n} E(Y^s_{nj\theta_{ni}} Y^s_{nj\theta_{nk}})/a_n(\theta_{ni})a_n(\theta_{nk}) = 0.$$

Step 2. Assume (4), (5) and (23). Then condition (21) holds.

Proof of Step 2. Fix θ with $\|\theta\| = 1$ and let

$$\begin{aligned} Z_{n\theta} &= \langle \sum_{i=1}^{d} \sum_{j=1}^{n} (Y_{nj\theta_{ni}}/a_n(\theta_{ni}))e_i, \theta\rangle \\ &= \sum_{j=1}^{n} (\sum_{i=1}^{d} \langle\theta, e_i\rangle Y_{nj\theta_{ni}}/a_n(\theta_{ni})) \equiv \sum_{j=1}^{n} W_{nj\theta}. \end{aligned}$$

The Cramér-Wold device implies that verification of (21) is equivalent to proving that $\mathscr{L}(Z_{n\theta}) \to N(0, 1)$ for all θ with $\|\theta\| = 1$. Notice that $Z_{n\theta}$ is the sum of n independent mean zero random variables $W_{nj\theta}$, $1 \le j \le n$, for which

$$|W_{nj\theta}| \le 2\epsilon_n \sum_{j=1}^{d} |\langle\theta, e_j\rangle| \le 2\epsilon_n\sqrt{d}.$$

Hence, $\mathscr{L}(Z_{n\theta}) \to N(0, 1)$ provided $\lim_{n\to\infty} EZ^2_{n\theta} = 1$.

By independence,

(24) $$\begin{aligned} EZ^2_{n\theta} &= \sum_{j=1}^{n} EW^2_{nj\theta} = (1/2) \sum_{j=1}^{n} E(W^s_{nj\theta})^2 \\ &= (1/2) \sum_{i=1}^{d} (\langle\theta, e_i\rangle/a_n(\theta_{ni}))^2 \sum_{j=1}^{n} E(Y^s_{nj\theta_{ni}})^2 \\ &\quad + \sum_{1\le i<k\le d}(\langle\theta, e_i\rangle\langle\theta, e_k\rangle/a_n(\theta_{ni})a_n(\theta_{nk})) \sum_{j=1}^{n} EY^s_{nj\theta_{ni}} Y^s_{nj\theta_{nk}}. \end{aligned}$$

Notice that (4′), (5) and the NCC imply that for any unit vectors θ_n

$$\mathscr{L}(\sum_{j=1}^{n} \langle X^s_j, \theta_n\rangle/a_n(\theta_n)) \to N(0, 2).$$

Utilizing (4′) once again and the fact that $\text{med}(\langle X_j, \theta\rangle - \text{med}(\langle X_j, \theta\rangle)) = 0$ we see that $\langle X^s_j, \theta\rangle$ can be replaced by $Y^s_{nj\theta}$ because

$$\le 2 \lim_{n\to\infty} \sup_{\|\theta\|=1} P(\sum_{j=1}^{n} \langle X^s_j, \theta\rangle \neq \sum_{j=1}^{n} Y^s_{nj\theta})$$

(25) $$\le 2 \lim_{n\to\infty} \sup_{\|\theta\|=1} \sum_{j=1}^{n} P(|\langle X_j, \theta\rangle - \text{med}(\langle X_j, \theta\rangle)| > \epsilon_n a_n(\theta)) = 0.$$

Consequently, the first group of terms in (24) converges to 1 by the NCC. Hence, it suffices to verify (23).

Our next step is a preliminary to verifying (23).

Step 3. Let $V_{nj\theta} = \sum_{i=1}^{d} \langle\theta, \theta_{ni}\rangle Y^s_{nj\theta_{ni}}$. Then

(26) $$\liminf_{n\to\infty}\inf_{\|\theta\|=1} \sum_{j=1}^{n} E(V_{nj\theta}/a_n(\theta))^2 \geq 2$$

with equality achieved by $\theta \in \{\theta_{n1}, \cdots, \theta_{nd}\}$

Proof of Step 3. Notice that $Y^s_{nj\theta} = \langle X^s_j, \theta\rangle$ on

$$\Omega_{nj\theta} = \{ | \langle X_j, \theta\rangle - \text{med}(\langle X_j, \theta\rangle) | \leq \epsilon_n a_n(\theta), | \langle \bar{X}_j, \theta\rangle - \text{med}(\langle \bar{X}_j, \theta\rangle) | \leq \epsilon_n a_n(\theta)\}.$$

Since $\{\theta_{ni}, i = 1, \cdots, d\}$ form an orthonormal basis,

$$\langle X^s_j, \theta\rangle = \sum_{i=1}^{d} \langle X^s_j, \theta_{ni}\rangle\langle\theta_{ni}, \theta\rangle$$
$$= V_{nj\theta} \text{ on } \cap_{i=1}^{d} \Omega_{nj\theta_{ni}}.$$

Let $U_{nj\theta} = \langle X^s_j, \theta\rangle$. Lower bounding,

$$\sum_{j=1}^{n} EV^2_{nj\theta}/a^2_n(\theta) \geq \sum_{j=1}^{n} E((U^2_{nj\theta} \wedge a^2_n(\theta))/a^2_n(\theta))I_{(U_{nj\theta}=V_{nj\theta})}$$
$$= 2 - \sum_{j=1}^{n} E((U^2_{nj\theta} \wedge a^2_n(\theta))/a^2_n(\theta))I_{(U_{nj\theta}\neq V_{nj\theta})}$$
$$\geq 2 - \sum_{j=1}^{n} P(U_{nj\theta} \neq V_{nj\theta})$$
$$\geq 2 - \sum_{j=1}^{n} \sum_{i=1}^{d} P(\Omega^c_{nj\theta_{ni}}) \to 2 \quad \text{as } n \to \infty \quad \text{by (25)}.$$

To complete the proof notice that

$$\lim_{n\to\infty} \sum_{j=1}^{n} EV^2_{nj\theta_{ni}}/a^2_n(\theta_{ni}) = 2.$$

Step 4. (4) and (5) imply (23).

Proof of Step 4. The idea is to use Step 3 in order to obtain the validity of (18) in the Lemma with C_n/A_nB_n denoting the sum in (23).

Restricting θ to be a unit vector in the spanof θ_{n_i} and θ_{nk} for some fixed $1 \leq i < k \leq d$,

$$\liminf_{n\to\infty}\inf_\theta \frac{\sum_{r=i,k} \sum_{j=1}^{n} E(Y^s_{nj\theta_{nr}})^2\langle\theta, \theta_{nr}\rangle^2 + 2\sum_{j=1}^{n} EY^s_{nj\theta_{ni}}Y^s_{nj\theta_{nk}}\langle\theta, \theta_{ni}\rangle\langle\theta, \theta_{nk}\rangle}{\sum_{j=1}^{n} E(Y^s_{nj\theta_{ni}})^2}$$
$$= \liminf_{n\to\infty}\inf_\theta \sum_{j=1}^{n} EV^2_{nj\theta}/2a^2_n(\theta_{ni})$$
$$= \liminf_{n\to\infty}\inf_\theta \sum_{j=1}^{n} E(V^2_{nj\theta}/2a^2_n(\theta))(a^2_n(\theta)/a^2_n(\theta_{ni}))$$
$$= 1 \quad \text{since} \quad \geq 1 \text{ by step 3 and } \quad = 1 \text{ for } \theta = \theta_{ni}.$$

Invoking the Lemma with

$$A^2_n = \sum_{j=1}^{n} E(Y^s_{nj\theta_{ni}})^2 \sim 2a^2_n(\theta_{ni})$$
$$B^2_n = \sum_{j=1}^{n} E(Y^s_{nj\theta_{nk}})^2 \sim 2a^2_n(\theta_{nk})$$
$$C_n = \sum_{j=1}^{n} EY^s_{nj\theta_{ni}}Y^s_{nj\theta_{nk}}$$

completes the verification of (23).

Thus, $(T_n(S_n - v_n)) \to N(\vec{0}, I)$.

Since the proof of necessity is fairly analogous to the proof in the i.i.d. case which appears in Hahn and Klass (1978), we briefly sketch the proof indicating the modifications.

PROOF OF NECESSITY IN THEOREM 1. Assuming (2), $\mathscr{L}(T_nS^s_n) \to N(\vec{0}, 2I)$. The norming linear operators T_n can be written in the form $T_ne_i = \sum_{j=1}^{d} c_{ijn}e_j$ where $c_{ijn} = \langle T_ne_i, e_j\rangle$ for the standard orthonormal basis $\{e_i\}$. Each T_n is nonsingular for n sufficiently large and so has an associated set of unit vectors

$$\theta_{jn} = \sum_{i=1}^{d} Q_{jn}c_{ijn}e_i, \qquad j = 1, \cdots, d,$$

where $Q_{jn} = (\sum_{i=1}^{d} c_{ijn}^2)^{-1/2}$. Expressing S_n^s and T_n in terms of these unit vectors we obtain

$$\mathscr{L}(\sum_{k=1}^{d} (\langle S_n^s, \theta_{kn}\rangle/Q_{kn})e_k = \mathscr{L}(\sum_{k=1}^{d}\sum_{i=1}^{d} \langle S_n^s, e_i\rangle c_{ikn}e_k)$$

$$= \mathscr{L}(\sum_{k=1}^{d} \langle T_nS_n^s, e_k\rangle e_k) = \mathscr{L}(T_nS_n^s) \to N(\vec{0}, 2I).$$

Thus, utilizing a Lemma of Rao and the Lévy Continuity Theorem exactly as in Hahn and Klass (1978), Section 2, for any unit vectors ψ_n

$$\text{(27)} \qquad \mathscr{L}(\sum_{k=1}^{d} (\langle S_n^s, \theta_{kn}\rangle\langle e_k, \psi_n\rangle/Q_{kn})) \to N(0, 2).$$

Therefore (for all n sufficiently large) $\theta_{1n}, \cdots, \theta_{dn}$ are linearly independent. Consequently, for each unit vector φ, there exist unique constants $\beta_{in}(\varphi)$ such that

$$\varphi = \sum_{i=1}^{d} (\beta_{in}(\varphi)/Q_{in})\theta_{in}.$$

Let $b_n^2(\varphi) = \sum_{i=1}^{d} \beta_{in}^2(\varphi)$. For any sequence of unit vectors $\{\varphi_n\}$, since $\sum_{k=1}^{d} (\beta_{kn}(\phi_n)/b_n(\phi_n))e_k$ is a unit vector, (27) implies that

$$\mathscr{L}(\langle S_n^s, \varphi_n\rangle/b_n(\varphi_n)) = \mathscr{L}\left(\sum_{k=1}^{d} \frac{\beta_{kn}(\varphi_n)}{b_n(\varphi_n)} \frac{\langle S_n^s, \theta_{kn}\rangle}{Q_{kn}}\right) \to N(0, 2).$$

We now appeal to the following uniform version of the NCC for symmetrized random variables.

PROPOSITION. *For every $n \geq 1$, $\alpha \in J$ let $X_{n1\alpha}, \cdots, X_{nj_n\alpha}$ be independent random variables. Then for every sequence $\{\alpha_n\} \subset J$ and for every $\epsilon > 0$*

$$\mathscr{L}(\sum_{k=1}^{j_n} X_{nk\alpha_n}^s) \to N(0, 2)$$

and

$$\lim_{n\to\infty}\max_{1\le k\le n}\sup_{\alpha\in J}P(|X_{nk\alpha}^s| > \epsilon) = 0$$

iff *for every $\epsilon > 0$ and a $\tau > 0$*

$$\text{(i) } \lim_{n\to\infty}\sup_{\alpha\in J} \sum_{k=1}^{j_n} P(|X_{nj\alpha}^s| > \epsilon) = 0$$

$$\text{(ii) } \lim_{n\to\infty}\sup_{\alpha\in J} |\sum_{k=1}^{j_n} E|X_{nj\alpha}^s|^2 I_{(|X_{nj\alpha}^s|\le\tau)} - 2| = 0.$$

Letting $X_{nk\varphi}^s = \langle X_k^s, \varphi\rangle/b_n(\varphi)$, (i) of the Proposition yields (4) with $b_n(\varphi)$ in place of $a_n(\varphi)$ once the uniform u.a.n. condition is verified for $X_{nk\varphi}^s$. However, by (27),

$$\lim_{n\to\infty}\max_{1\le j\le n}\sup_{\|\varphi\|=1}P(|\sum_{k=1}^{d} (\langle X_j^s, \theta_{kn}\rangle/Q_{kn})\langle e_k, \varphi\rangle| > \epsilon) = 0$$

which immediately implies

$$\text{(28)} \qquad \lim_{n\to\infty}\max_{1\le j\le n}\sup_k P(|\langle X_j^s, \theta_{kn}\rangle/Q_{kn}| > \epsilon) = 0.$$

Combining (28) with

$$P(|\langle X_j^s, \varphi\rangle/b_n(\varphi)| > \epsilon) = P(|\sum_{k=1}^{d} \frac{\beta_{kn}(\varphi)}{b_n(\varphi)} \langle X_j^s, \theta_{kn}\rangle/Q_{kn}| > \epsilon)$$

$$\le \sum_{k=1}^{d} P(|\langle X_j^s, \theta_{kn}\rangle/Q_{kn}| > \epsilon/\sqrt{d}) \qquad \text{by Cauchy-Schwarz}$$

$$= d\,\sup_k P(|\langle X_j^s, \theta_{kn}\rangle/Q_{kn}| > \epsilon/\sqrt{d})$$

which tends to zero by (3) and thereby yields the uniform u.a.n. condition. Thus, (4) holds with $b_n(\varphi)$ in place of $a_n(\varphi)$. Finally by (i) and (ii), there exist $\epsilon_n \downarrow 0$ such that for all φ

$$\sum_{k=1}^{n} E((\langle X_k^s, \varphi\rangle/\tilde{b}_n(\varphi))^2 \wedge 1) \text{ is } \begin{cases} < 2 & \text{if } \tilde{b}_n(\varphi) > (1+\epsilon_n)b_n(\varphi) \\ > 2 & \text{if } \tilde{b}_n(\varphi) < (1-\epsilon_n)b_n(\varphi) \end{cases}.$$

This implies $\lim_{n\to\infty}\sup_{\|\varphi\|=1} | b_n(\varphi)/a_n(\varphi) - 1 | = 0$. Hence (4) is necessary.

To complete the proof notice that (5) is a consequence of the fact that there exists n_0 and $\gamma > 0$ such that for all $n \geq n_0$,

$$\inf_{\|\varphi\|=1} a_n(\varphi) \geq \inf_{\|\varphi\|=1} \gamma b_n(\varphi) \geq \inf_{\|\varphi\|=1} \max_{1\leq i\leq d} \beta_{in}^2(\varphi) \geq 1/d \min_{1\leq i\leq d} Q_{in} > 0. \qquad \square$$

Acknowledgment. The authors are grateful to the referee for bringing several references to their attention.

REFERENCES

BHATTACHARYA, R. N. and RAO, R. R. (1976). *Normal Approximations and Asymptotic Expansions.* Wiley, New York.

GNEDENKO, B. V. and KOLMOGOROV, A. N. (1968). *Limit Distributions For Sums of Independent Random Variables.* Addison-Wesley, Reading, Massachusetts.

HAHN, M. G. (1978). The generalized domain of attraction of a Gaussian law on Hilbert space. Proc. Second Int. Conf. Prob. in Banach Spaces. *Lecture Notes in Math.* **709** 125–144.

HAHN, M. G. and KLASS, M. J. (1980). Matrix normalization of sums of random vectors in the domain of attraction of the multivariate normal. *Ann. Probability* **8** 262–280.

HAHN, M. G. and KLASS, M. J. (1981). A survey of generalized domains of attraction and operator norming methods. Proc. Third Int. Conf. Prob. in Banach Spaces. *Lecture Notes in Math.* To appear.

KANDELAKI, N. P. and SAZONOV, V. V. (1964). On a central limit theorem for random elements with values in Hilbert space. *Theory of Probability Appl.* **9** 38–46.

KLASS, M. J. (1981). A method of approximating expectations of functions of sums of independent random variables. *Ann. Probability* **9** 413–428.

PARTHASARATHY, K. R. (1967). *Probability Measures on Metric Spaces.* Academic.

DEPARTMENT OF MATHEMATICS
TUFTS UNIVERSITY
MEDFORD, MASSACHUSETTS 02155

DEPARTMENT OF STATISTICS
UNIVERSITY OF CALIFORNIA, BERKELEY
BERKELEY, CALIFORNIA 94720