Elect. Comm. in Probab. **5** (2002) 63–76

ELECTRONIC COMMUNICATIONS in PROBABILITY

A LAW OF THE ITERATED LOGARITHM FOR THE SAMPLE COVARIANCE MATRIX

STEVEN J. SEPANSKI
Department of Mathematical Sciences, Saginaw Valley State University, University Center, MI 48710
Email: sepanski@svsu.edu

Submitted 5 August 2002 *Final version accepted* 24 January 2003

AMS 2000 subject classifications: Primary 60F15; secondary 60F05.
Law of the iterated logarithm, central limit theorem, generalized domain of attraction, extreme values, operator normalization, self normalization, t statistic.

Abstract

For a sequence of independent identically distributed Euclidean ran vectors, we prove a law of the iterated logarithm for the sample covariance matrix when $o(\log\log n)$ terms are omitted. The result is proved under the hypotheses that the random vectors belong to the generalized domain of attraction of the multivariate Gaussian law. As an application, we obtain a bounded law of the iterated logarithm for the multivariate t-statistic.

Introduction:

Let $X, X_1, X_2, \cdots$ be independent identically distributed (iid) random column vectors taking values in Euclidean space, $\mathbb{R}^d$. We use the standard inner product $\langle\cdot,\cdot\rangle$ and induced norm. Assume that the law of X is full. That is, the law of X is not supported on any $d-1$ dimensional hyperplane of $\mathbb{R}^d$. We say that the law of X, or merely X, is in the generalized domain of attraction of the multivariate Gaussian law (GDOAG) if there exist operators T_n and vectors b_n such that $T_nS_n - b_n \Rightarrow N(0,I)$. Here, $S_n = \sum_{i=1}^n X_i$, I is the identity on $\mathbb{R}^d$, $N(0,I)$ is the standard Gaussian distribution on $\mathbb{R}^d$ and $\Rightarrow$ denotes weak convergence. Since the standard Gaussian law is an affine transformation of any other Gaussian law, there is no loss of generality in assuming the limit is standard. Also, if the law of X is in the GDOAG then $E\|X\| < \infty$ and so, changing T_n accordingly, one can take $b_n = nEX$. See [11], Proposition 8.1.6. Therefore, there is no loss of generality in assuming that $EX = 0$. This will be assumed throughout the article. If the law of X is in the GDOAG we then have that there exist nonsingular operators, T_n, such that

$$T_nS_n \Rightarrow N(0,I). \tag{1}$$

In order to simplify the presentation we extend the definition of the normalizing operators to noninteger values. Define $T_x = T_{[x]}$, where $[\cdot]$ denotes the greatest integer function.

Meerschaert [10] uses regular variation in $\mathbb{R}^d$ to characterize GDOAs. The characterization using regular variation will be central to our approach. The main result we will use from Meerschaert [10] is that the norming operators in (1) can be chosen to satisfy

$$\lim_{x\to\infty} \|T_{bx}T_x^{-1} - b^{-1/2}I\| = 0 \tag{2}$$

if the law of X belongs to the GDOAG. Moreover, the convergence is uniform over compact subsets of $(0,\infty)$ Meerschaert [10], Theorem 3.1.

In this paper, we investigate the law of the iterated logarithm behavior of the sample covariance matrix $C_n = \sum_{i=1}^n X_i X_i^T$. It is shown that if X is in the GDOAG, there exist T_n satisfying (1) and a set $B_n \subset \{1, 2, \cdots, n\}$ with cardinality $r_n = o(\log\log n)$, such that

$$\frac{T_{n/\log\log n}(C_n - R_n)^{1/2}}{\sqrt{\log\log n}} \to I \quad a.s.$$

Here $(\cdot)^{1/2}$ is the symmetric square root and $R_n = \sum_{i\in B_n} X_i X_i^T$.

Moreover, from this one may obtain a self normalized bounded LIL. In particular, we will show that if X is in the GDOAG, then

$$\limsup \left\| \frac{C_n^{-1/2} S_n}{\sqrt{2\log\log n}} \right\| = 1 \quad a.s.$$

These results are multivariate analogues of some univariate results obtained by Griffin and Kuelbs [7] and Gine and Mason [6].

Results:

For $t > 0$, let $Lt = \max(1, \log_e t)$ and let $L_2 t = LLt$. Also, let $\alpha(t) = t/L_2 t$. For $x \in \mathbb{R}^d$ and $A \subset \mathbb{R}^d$, define the distance from x to A to be $d(x, A) = \inf_{y\in A} \|x - y\|$. For a sequence of vectors, $\{x_n\}$, denote the cluster set of $\{x_n\}$ by $C(\{x_n\}) = \{y : \liminf \|x_n - y\| = 0\}$. Write $x_n \twoheadrightarrow A$ if $d(x_n, A) \to 0$ and $C(\{x_n\}) = A$. Define $t_n = \frac{n}{L_2 n}$. Let $\bar{B} = \{x \in \mathbb{R}^d : \|x\| \le 1\}$.

We first describe what we mean by an extreme value. Assuming that X is in the GDOAG, with $EX = 0$, let T_n be as in (1) and (2). For each n, rearrange $X_1, \cdots, X_n$ as $X_n(1), \cdots, X_n(n)$ so that $\|T_{t_n} X_n(1)\| \ge \|T_{t_n} X_n(2)\| \ge \cdots \ge \|T_{t_n} X_n(n)\|$. Ties may be broken in any arbitrary manner. Next, define $S_n^{(r)} = S_n - \sum_{i=1}^r X_n(i)$ and $S_n^{(0)} = S_n$. Finally, for any sequence $r_n > 0$ define $R_n = \sum_{i=1}^{r_n} X_n(i) X_n(i)^T$. If $r_n = 0$, define $R_n = 0$.

Following are the statements of the main results contained in this article. Proofs and technical lemmas are contained in the subsequent section.

Theorem 1: Let X be in the GDOAG with $EX = 0$. There exist operators T_n, satisfying (1) and (2), and scalars $r_n = o(L_2 n)$ such that

$$\frac{T_{t_n}(C_n - R_n)T_{t_n}^*}{L_2 n} \to I \quad a.s.$$

If we can pass the operator square root through the limit in Theorem 1, we will obtain the following corollary. To do so requires the operators T_n to be positive and symmetric. The proof of Theorem 1 relies heavily on property (2). It is not clear whether one can obtain operators using the techniques of Meerschaert to yield operators that satisfy (2) and are positive and symmetric. On the other hand, the operators of Hahn and Klass, [8], are positive and symmetric, but may not satisfy (2). Therefore, we do not include this condition in the statement of the corollary. Billingsley's multivariate convergence of types theorem allows us to switch back and forth between the two approaches to normalizing operators, utilizing whichever property we need, regular variation or symmetry.

Corollary 2: Let X be in the GDOAG with $EX = 0$. There exist positive symmetric operators A_n satisfying (1) and scalars $r_n = o(L_2 n)$ such that

$$\frac{A_{t_n}(C_n - R_n)^{1/2}}{\sqrt{L_2 n}} \to I \quad a.s.$$

From Sepanski [15], wherein was proved a trimmed LIL with the operators T_n, and Corollary 2 above, we will obtain the following self normalized LIL.

Corollary 3: Let X be in the GDOAG with $EX = 0$. There exist $r_n = o(L_2 n)$ such that

$$\frac{(C_n - R_n)^{-1/2} S_n^{(r_n)}}{\sqrt{2L_2 n}} \twoheadrightarrow \bar{B} \quad a.s.$$

Theorem 4: If X is in the GDOAG with $EX = 0$, then

$$\limsup_n \left\| \frac{C_n^{-1/2} S_n}{\sqrt{2 \log \log n}} \right\| = 1 \ a.s.$$

Theorem 4 stands in comparison to the real valued Theorem 1 of Griffin and Kuelbs [7]. There they show that if X is in the Domain of Attraction of a Gaussian Law in $\mathbb{R}$, then

$$\frac{\sum_{i=1}^n (X_i - EX_i)}{\sqrt{2L_2 n \sum_{i=1}^n X_i^2}} \twoheadrightarrow [-1, 1] \quad a.s.$$

To see why results such as Theorem 1 and its corollaries may be true, consider first the case where $E\|X\|^2 < \infty$. In this case, the covariance matrix, $C = E(XX^T)$ exists. Moreover, by the classical case of the Central Limit Theorem, one may take $T_n = (nC)^{-1/2}$ in (1) and (2). In which case,

$$\frac{T_{t_n} C_n^{1/2}}{\sqrt{L_2 n}} = \frac{C^{-1/2} C_n^{1/2}}{\sqrt{n}} \to I \quad a.s.$$

by the Strong Law of Large Numbers. Hence, Corollary 2 holds with $r_n \equiv 0$, and therefore $R_n \equiv 0$. From this, and the classical LIL, one may infer the conclusions of the other three results.

The main impetus of this article is to extend the results to the case where X is in the GDOAG but $E\|X\|^2 = \infty$. Here, accounting for the influence of extreme terms is paramount. Although there may be extreme terms, they are few in number. In fact, they are on the order of $o(L_2 n)$. Also, note that Theorem 4 contains no trimming. Heuristically, this is due to the fact that when any extreme reappear in the sum, they are essentially cancelled out when they reappear in the *inverse* of the sample covariance matrix.

Proofs: We begin by constructing the sequence $\{r_n\}$. This construction mirrors that of Kuelbs and Ledoux [9]. Define $\Gamma(t) = \sup_{s \geq t} sP(\|T_s X\| \geq 1)$. By Gaussian convergence criteria (see, for example, Araujo and Gine, [1], Theorem 5.9), $\Gamma(t) \downarrow 0$, as $t \uparrow \infty$. We let

$$\xi_n = \left(L\Big(\frac{1}{\Gamma(\sqrt{n})}\Big)\right)^{-1/2}.$$

Gaussian convergence guarantees that $\xi_n \downarrow 0$. By basic calculus, $\xi_n L\Big(\frac{\xi_n}{\Gamma(\delta t_n)}\Big) \to \infty, \forall \delta > 0$.

Replacing ξ_n by anything larger does not alter the second of these, so we may assume, without loss of generality that $\xi_n \downarrow 0,\ \xi_n L\Big(\frac{\xi_n}{\Gamma(\delta t_n)}\Big) \to \infty,\ \forall \delta > 0$, and also that $r_n = [\xi_n L_2 n] \uparrow \infty$. Having defined the sequence r_n, we now proceed to the proof of Theorem 1.

Proof of Theorem 1: Fix $a > 1,\ \rho > 0$. Let $n_k = [a^k]$, and $I_k = (n_k, n_{k+1}]$. For $1 \leq j \leq n_{k+1}$, define the random matrices,

$$u_j = u_j(k,a,\rho) = X_j X_j^T I[\|T_{t_{n_k}} X_j\| \leq \rho]$$
$$w_j = w_j(k,a,\rho) = X_j X_j^T I[\|T_{t_{n_k}} X_j\| > \rho]$$

Next, for $n \in I_k$ define the random truncated covariance matrices

$$U_n = U_n(a,\rho) = \sum_{j=1}^{n} (u_j - Eu_j)$$
$$W_n = W_n(a,\rho) = \sum_{j=1}^{n} w_j$$

We analyze each of the sequences according to the following Lemmas. Note that $C_n - R_n = U_n + W_n - R_n + nEXX^T I[\|T_{t_{n_k}} X\| \leq \rho]$.

Lemma 5: Assume that X has mean zero and is in the GDOAG. For all $\rho > 0,\ a > 1, v > 0$, and for all unit vectors $\theta,\ \phi$,

$$\limsup_k \max_{n \in I_k} \left| \frac{\theta^T T_{t_{n_k}} U_n T^*_{t_{n_k}} \phi}{L_2 n} \right| \leq \frac{\rho v e^v}{2} a + \frac{\rho}{v} \quad a.s.$$

Lemma 6: Assume that X has mean zero and is in the GDOAG. For all $\rho > 0$, and $a > 1$,

$$\limsup_k \max_{n \in I_k} \left\| \frac{T_{t_{n_k}} (W_n - R_n) T^*_{t_{n_k}}}{L_2 n} \right\| = 0 \quad a.s.$$

Lemma 7: Assume that X has mean zero and is in the GDOAG. For all $\rho > 0$, and $a > 1$,

$$\limsup_k \max_{n \in I_k} \left\| t_n T_{t_{n_k}} E(XX^T I[\|T_{t_{n_k}} X\| \leq \rho]) T^*_{t_{n_k}} - I \right\| \leq a - 1$$

Assuming each lemma to be true for the time being and deferring the proofs of them until later, we continue with the proof of Theorem 1. Combining Lemmas 5-7 via the triangle inequality yields, for every pair of unit vectors ϕ, θ, and for every $v > 0,\ \rho > 0,\ a > 1$,

$$\limsup_k \max_{n \in I_k} \left| \frac{\theta^T T_{t_{n_k}} (C_n - R_n) T^*_{t_{n_k}} \phi}{L_2 n} - \langle \theta, \phi \rangle \right| \leq \frac{\rho v e^v}{2} a + \frac{\rho}{v} + a - 1 \quad a.s.$$

Note that the left side here no longer depends on ρ and v. First, we let $v = \sqrt{\rho}$, and then let $\rho \downarrow 0$. Then we have that for every pair of unit vectors ϕ, θ, and for every $a > 1$,

$$\limsup_k \max_{n \in I_k} \left| \frac{\theta^T T_{t_{n_k}} (C_n - R_n) T^*_{t_{n_k}} \phi}{L_2 n} - \langle \theta, \phi \rangle \right| \le a - 1 \quad a.s.$$

Intersecting over θ and ϕ in a countable dense subset of the unit sphere yields for every $a > 1$,

$$\limsup_k \max_{n \in I_k} \left\| \frac{T_{t_{n_k}} (C_n - R_n) T^*_{t_{n_k}}}{L_2 n} - I \right\| \le a - 1 \quad a.s.$$

Finally, we need to eliminate the dependence on $a > 1$ by eliminating the subsequence n_k from the left side. This essentially follows from (2).

First, observe that since t_n is eventually increasing we have that for $n \in I_k$, $t_{n_k} \le t_n \le t_{n_{k+1}}$. Therefore, dividing through by t_{n_k}, we have that $1 \le \frac{t_n}{t_{n_k}} \le \frac{t_{n_{k+1}}}{t_{n_k}}$. The latter converges to a and therefore is eventually bounded by $2a$. Hence, for large k, $\{\frac{t_n}{t_{n_k}}\}_{n \in I_k} \subset [1, 2a]$. Applying (2) over this compact set yields $\max_{n \in I_k} \|T_{t_n} T^{-1}_{t_{n_k}} - (\frac{t_n}{t_{n_k}})^{-1/2} I\| \le \sup_{b \in [1,2a]} \|T_{bt_{n_k}} T^{-1}_{t_{n_k}} - b^{-1/2} I\| \to 0$. From this we obtain

$$\limsup_k \max_{n \in I_k} \|T_{t_n} T^{-1}_{t_{n_k}}\| \le \limsup_k \max_{n \in I_k} \|T_{t_n} T^{-1}_{t_{n_k}} - (\frac{t_n}{t_{n_k}})^{-1/2} I\| + \limsup_k \max_{n \in I_k} (\frac{t_n}{t_{n_k}})^{-1/2} \le 1$$

Also by (2), and the fact that $\|T\| = \|T^*\|$, we obtain uniformly over $n \in I_k$

$$\begin{aligned} o(1) &= \left(T_{t_n} T^{-1}_{t_{n_k}} - (\frac{t_n}{t_{n_k}})^{-1/2} I\right) \left(T^{*}_{t_{n_k}}{}^{-1} T^*_{t_n} - (\frac{t_n}{t_{n_k}})^{-1/2} I\right) \\ &= \left(T_{t_n} T^{-1}_{t_{n_k}} T^{*}_{t_{n_k}}{}^{-1} T^*_{t_n} - \frac{t_{n_k}}{t_n} I\right) + \sqrt{\frac{t_{n_k}}{t_n}} \left(T_{t_n} T^{-1}_{t_{n_k}} - (\frac{t_n}{t_{n_k}})^{-1/2} I + T^{*}_{t_{n_k}}{}^{-1} T^*_{t_n} - (\frac{t_n}{t_{n_k}})^{-1/2} I\right) \\ &= \left(T_{t_n} T^{-1}_{t_{n_k}} T^{*}_{t_{n_k}}{}^{-1} T^*_{t_n} - \frac{t_{n_k}}{t_n} I\right) + O(1) o(1) \end{aligned}$$

Finally, we note that $1 \le \frac{t_n}{t_{n_k}} \le \frac{t_{n_{k+1}}}{t_{n_k}}$ implies $0 \le \max_{n \in I_k} (1 - \frac{t_{n_k}}{t_n}) \le 1 - \frac{t_{n_k}}{t_{n_{k+1}}} \to 1 - \frac{1}{a}$. Putting the three facts outlined above together yields,

$$\begin{aligned} &\limsup_k \max_{n \in I_k} \left\| \frac{T_{t_n} (C_n - R_n) T^*_{t_n}}{L_2 n} - I \right\| \\ &\le \limsup_k \max_{n \in I_k} \|T_{t_n} T^{-1}_{t_{n_k}}\|^2 \ \limsup_k \max_{n \in I_k} \left\| \frac{T_{t_{n_k}} (C_n - R_n) T^*_{t_{n_k}}}{L_2 n} - I \right\| \\ &\quad + \limsup_k \max_{n \in I_k} \left\| T_{t_n} T^{-1}_{t_{n_k}} T^{*}_{t_{n_k}}{}^{-1} T^*_{t_n} - \frac{t_{n_k}}{t_n} I \right\| + \limsup_k \max_{n \in I_k} \left\| (1 - \frac{t_{n_k}}{t_n}) I \right\| \\ &\le (1)(a-1) + 0 + \left(1 - \frac{1}{a}\right) \quad a.s. \end{aligned}$$

To complete the proof of Theorem 1 (subject to Lemmas 5-7) let $a \downarrow 1$ through a countable set.

Proof of Lemma 5: The proof uses classical blocking arguments and the Borel-Cantelli Lemma along the subsequence n_k. A maximal inequality (Ottaviani's) allows us to pass to the full sequence.

Fix unit vectors θ, ϕ. Let $\epsilon > 0$, $\xi > 0$ be given. Apply Ottaviani's inequality (see, for example, Dudley, [3], p.251) to obtain

$$P\Big(\max_{n\in I_k}\frac{\Big|\theta^T T_{t_{n_k}} U_n T^*_{t_{n_k}}\phi\Big|}{L_2 n} \ge (1+\xi)\epsilon\Big) \le P\Big(\max_{n\in I_k}\Big|\theta^T T_{t_{n_k}} U_n T^*_{t_{n_k}}\phi\Big| \ge (1+\xi)\epsilon L_2 n_k\Big) \\ \le 2P\Big(\Big|\theta^T T_{t_{n_k}} U_{n_{k+1}} T^*_{t_{n_k}}\phi\Big| \ge \epsilon L_2 n_k\Big) \qquad (3)$$

as long as $\max_{n\in I_k} P(|\sum_{j=n+1}^{n_{k+1}} \theta^T T_{t_{n_k}}(u_j - Eu_j)T^*_{t_{n_k}}\phi| \ge \epsilon L_2 n_k) \le \frac{1}{2}$. To verify that this is the case, at least for sufficiently large k, we apply Chebychev's inequality.

$$P\Big(\Big|\sum_{j=n+1}^{n_{k+1}}\theta^T T_{t_{n_k}}(u_j - Eu_j)T^*_{t_{n_k}}\phi\Big| \ge \epsilon L_2 n_k\Big)$$

$$\le \frac{1}{\epsilon^2 L_2 n_k^2} Var\Big(\sum_{j=n+1}^{n_{k+1}}\theta^T T_{t_{n_k}}(u_j - Eu_j)T^*_{t_{n_k}}\phi\Big)$$

$$\le \frac{\rho^2 n_{k+1}}{\epsilon^2 L_2 n_k^2} E\Big\|T_{t_{n_k}}X\Big\|^2 I[\|T_{t_{n_k}}X\| \le \rho]$$

$$= \frac{\rho^2}{\epsilon^2 L_2 n_k} \cdot \frac{n_{k+1}}{n_k} \cdot t_{n_k} E\Big\|T_{t_{n_k}}X\Big\|^2 I[\|T_{t_{n_k}}X\| \le \rho].$$

This converges to zero, as k goes to infinity, due to the fact that the first factor does, while the second factor converges to a and the third factor is bounded by the truncated variance condition of Gaussian convergence.

We will show that the last series in (3) is finitely summable over k. In order to do so, we require another Lemma. The lemma is based on an inequality of Pruitt [12]. For iid random vectors, Y_j and unit vectors θ, ϕ, define $Z_j(\theta,\phi) = \langle Y_j, \theta\rangle\langle Y_j, \phi\rangle I[\|Y_j\| \le \rho]$. Also, write $K(\theta,\phi,\rho) = \rho^{-2}EZ^2(\theta,\phi)$, and $M_n(\theta,\phi) = \sum_{j=1}^n Z_j(\theta,\phi)$.

Lemma 8: Let $Y_1, \cdots, Y_n$ be iid random vectors in Euclidean space, $\mathbb{R}^d$. For any $\rho > 0, s > 0, v > 0$, and for any unit vectors θ, ϕ

$$P\Big(\Big|M_n(\theta,\phi) - EM_n(\theta,\phi)\Big| > \frac{v}{2}e^{v\rho}n\rho K(\theta,\phi,\rho) + \frac{s\rho}{v}\Big) \le 2e^{-s}$$

Proof: By Taylor's expansion for the exponential function, if $x \in (-\rho^2, \rho^2)$ then $e^{ux} \le 1 + ux + \frac{u^2x^2}{2}e^{u\rho^2}$. Also, $1 + t \le e^t, \forall t$. Dropping subscripts, applying these inequalities to the random variable $Z(\theta,\phi)$, (note that $|Z(\theta,\phi)| \le \rho^2$), taking expectations, and using Jensen's inequality, we have that

$$Ee^{uZ(\theta,\phi)} \le 1 + uEZ(\theta,\phi) + \frac{u^2\rho^2e^{u\rho^2}}{2}K(\theta,\phi,\rho)$$

$$\le \exp\Big(uEZ(\theta,\phi) + \frac{u^2\rho^2e^{u\rho^2}}{2}K(\theta,\phi,\rho)\Big).$$

Below, we apply Markov's inequality in the second step and the previous inequality in the last step. We have the following.

$$P(M_n(\theta,\phi) > t)P(e^{uM_n(\theta,\phi)} > e^{ut})$$

$$\le e^{-ut}Ee^{uM_n(\theta,\phi)}$$

$$= e^{-ut}\Big(Ee^{uZ(\theta,\phi)}\Big)^n$$

$$\le \exp\Big(nuEZ(\theta,\phi) + n\frac{u^2\rho^2e^{u\rho^2}}{2}K(\theta,\phi,\rho) - ut\Big).$$

Take $t = nEZ(\theta,\phi) + n\frac{u\rho^2 e^{u\rho^2}}{2}K(\theta,\phi,\rho) + \frac{s}{u}$, $v = u\rho$. Observe that $nEZ(\theta,\phi) = EM_n(\theta,\phi)$, and then apply the same argument to $-M_n(\theta,\phi)$ to complete the proof of Lemma 8.

Continuing on with the proof of Lemma 5, we will apply Lemma 8 with $Y_j = T_{t_{n_k}}X_j$. We begin by analyzing

$$\begin{aligned} t_{n_k}K(\theta,\phi,\rho) &= t_{n_k}\rho^{-2}E\langle T_{t_{n_k}}X,\theta\rangle^2\langle T_{t_{n_k}}X,\phi\rangle^2 I[\|T_{t_{n_k}}X\| \le \rho] \\ &\le t_{n_k}E\langle T_{t_{n_k}}X,\theta\rangle^2 I[\|T_{t_{n_k}}X\| \le \rho] \to 1, \end{aligned}$$

by the basic convergence criteria for a *standard* Gaussian limit (see [11], Cor. 3.3.12). Using the above, and the fact that $n_{k+1}/n_k \to a$, we have that for all $\xi > 0$, and for sufficiently large k,

$$\frac{v\rho e^{v\rho}n_{k+1}K(\theta,\phi,\rho)}{2L_2 n_k} + \frac{(1+\xi)\rho}{v} \le \frac{v\rho e^{v\rho}}{2}a(1+\xi) + \frac{(1+\xi)\rho}{v}$$

Let the right hand side of this inequality be ϵ. Continuing the string of inequalities started in (3), by containment we then have that, for sufficiently large k,

$$\begin{aligned} &P\Big(\Big|\theta^T T_{t_{n_k}}U_{n_{k+1}}T^*_{t_{n_k}}\phi\Big| \ge \epsilon L_2 n_k\Big) \\ &= P\Big(\Big|M_{n_{k+1}}(\theta,\phi) - EM_{n_{k+1}}(\theta,\phi)\Big| \ge \epsilon L_2 n_k\Big) \\ &\le P\Big(\Big|M_{n_{k+1}}(\theta,\phi) - EM_{n_{k+1}}(\theta,\phi)\Big| \ge \frac{1}{2}n_{k+1}\rho v e^{v\rho}K(\theta,\phi,\rho) + \frac{L_2 n_k(1+\xi)\rho}{v}\Big) \\ &\le 2e^{-(1+\xi)L_2 n_k} \end{aligned}$$

This last series is finitely summable. Therefore, recalling the definition of ϵ, the above display, (3), and the Borel-Cantelli Lemma yield that for all $\xi > 0,\ \rho > 0,\ a > 1, v > 0$, and for all unit vectors $\theta,\ \phi$,

$$\limsup_k \max_{n\in I_k}\left|\frac{\theta^T T_{t_{n_k}}U_n T^*_{t_{n_k}}\phi}{L_2 n}\right| \le \epsilon(1+\xi) = \frac{\rho v e^{v\rho}}{2}a(1+\xi)^2 + \frac{(1+\xi)^2\rho}{v} \quad a.s.$$

Now let $\xi \downarrow 0$. Note that the left side does not depend on ξ. This completes the proof of Lemma 5.

Proof of Lemma 6: Recall that $X_n(1),\cdots,X_n(n)$ are an ordering of $X_1,\cdots,X_n$ under the operator T_{t_n}. Because in Lemma 6 we are normalizing by the operator $T_{t_{n_k}}$, we must also consider an ordering of $X_1,\cdots,X_n$ under the operator $T_{t_{n_k}}$. It is possible that the two orderings are different. However, due to (2), for $n \in I_k$, they are not significantly different. To this end, we let $X'_n(1),\cdots,X'_n(n)$ denote an ordering of $X_1,\cdots,X_n$ under $T_{t_{n_k}}$ such that $\|T_{t_{n_k}}X'_n(1)\| \ge \|T_{t_{n_k}}X'_n(2)\| \ge \cdots \|T_{t_{n_k}}X'_n(n)\|$. Recalling the definition of w_j,

$$\begin{aligned} \Big\|T_{t_{n_k}}(W_n - R_n)T^*_{t_{n_k}}\Big\| &= \Big\|T_{t_{n_k}}\Big(\sum_{j=1}^n w_j - \sum_{j=1}^{r_n}X_n(j)X_n(j)^T\Big)T^*_{t_{n_k}}\Big\| \\ &\le \Big\|\sum_{j=1}^n T_{t_{n_k}}X_jX_j^T T^*_{t_{n_k}}I[\|T_{t_{n_k}}X_j\| > \rho] - \sum_{j=1}^{r_n}T_{t_{n_k}}X'_n(j)X'_n(j)^T T^*_{t_{n_k}}\Big\| \\ &+ \Big\|\sum_{j=1}^{r_n}T_{t_{n_k}}X'_n(j)X'_n(j)^T T^*_{t_{n_k}} - \sum_{j=1}^{r_n}T_{t_{n_k}}X_n(j)X_n(j)^T T^*_{t_{n_k}}\Big\| \end{aligned} \tag{4}$$

We now try to bound each of the two terms in (4). For $n \in I_k$, let $p_n = \sum_{i=1}^n I[\|T_{t_{n_k}} X_i\| > \rho]$, and let $q_n = \sum_{i=1}^n I[\|T_{t_n} X_i\| > \rho a^{-1}]$. Note that p_n and q_n depend on k and are random. By (2), observe that

$$\max_{n\in I_k} \left\| T_{t_{n_k}} T_{t_n}^{-1} \right\| \le \max_{n\in I_k} \left\| T_{t_{n_k}} T_{t_n}^{-1} - \left(\frac{t_{n_k}}{t_n}\right)^{-1/2} I \right\| + \max_{n\in I_k} \left(\frac{t_{n_k}}{t_n}\right)^{-1/2} \to \sqrt{a}.$$

Therefore, for sufficiently large k, it can be made less than a, (since $a > \sqrt{a}$,). We assume that this holds for the rest of the proof. Therefore, $p_n \le q_n$, for k sufficiently large. Assume, for the moment, that $\max_{n\in I_k} q_n \le r_{n_k}$. Note that since $r_n \uparrow$, this also implies that $p_n \le q_n \le r_n$.

Observe that if $X_i \notin \{X_n'(1), \cdots, X_n'(p_n)\}$ then $\|T_{t_{n_k}} X_i\| \le \rho$. Otherwise we would have $\|T_{t_{n_k}} \; X_n'(1)\| \ge \|T_{t_{n_k}} \; X_n'(2)\| \ge \cdots \ge \|T_{t_{n_k}} \; X_n'(p_n)\| \ge \|T_{t_{n_k}} \; X_i\| > \rho$. This is contrary to the definition of p_n. Moreover, $\|T_{t_{n_k}} \; X_n'(i)\| > \rho$ for $i = 1, \cdots, p_n$. Similarly, if $X_i \notin \{X_n(1), \cdots, X_n(q_n)\}$ then $\|T_{t_n} X_i\| \le \rho/a$, and $\|T_{t_n} X_n(i)\| > \rho/a$ for $i = 1, \cdots, q_n$.

Consider the first term in (4). If at most r_n of $X_1, \cdots, X_{n_{k+1}}$ satisfy $\|T_{t_{n_k}} X_i\| > \rho$, then since $n \le n_{k+1}$, at most r_n of $X_1, \cdots, X_n$ satisfy $\|T_{t_{n_k}} X_i\| > \rho$. Such summands will then appear in both $\sum_{j=1}^n T_{t_{n_k}} X_j X_j T_{t_{n_k}}^* I[\|T_{t_{n_k}}\| > \rho]$ and $\sum_{j=1}^{r_n} T_{t_{n_k}} X_n'(j) X_n'(j)^T T_{t_{n_k}}^*$. They therefore cancel out in subtraction. If there are any other such summands, they appear only in the latter, not the former, because they will be zeroed out by the indicator. By the preceeding paragraph, such summands satisfy $\|T_{t_{n_k}} X_i X_i^T T_{t_{n_k}}^*\| = \|T_{t_{n_k}} X_i\| \;^2 \le \rho^2$. Furthermore, there are at most $r_n - p_n \le r_n$ such summands. Therefore,

$$\left\| \sum_{j=1}^n T_{t_{n_k}} X_j X_j^T T_{t_{n_k}}^* I[\|T_{t_{n_k}}\| > \rho] - \sum_{j=1}^{r_n} T_{t_{n_k}} X_n'(j) X_n'(j)^T T_{t_{n_k}}^* \right\| \le r_n \rho^2. \tag{5}$$

To handle the second term in (4) we have to consider the possibility that $X_1, \cdots, X_n$ may be ordered differently under T_{t_n} and $T_{t_{n_k}}$. However, as noted, by (2), the orderings are not significantly different.

By the definition of p_n and q_n, we have that for n sufficiently large, $\{X_n'(1), \cdots, X_n'(p_n)\} \subset \{X_n(1), \cdots, X_n(q_n)\}$. If not, $X_n'(i) \notin \{X_n(1), \cdots, X_n(q_n)\}$ for some $i = 1, \cdots, p_n$. In which case, as observed above, $\|T_{t_n} \; X_n'(i)\| \le \rho/a$. So, $\|T_{t_{n_k}} \; X_n'(i)\| \le \rho$. On the other hand, since $i \le p_n$, as observed above, $\|T_{t_{n_k}} \; X_n'(i)\| > \rho$. A contradiction.

Because of this containment, there is cancellation in the second term in (4). Indeed,

$$\begin{aligned}
&\sum_{j=1}^{r_n} T_{t_{n_k}} X_n(j) X_n(j)^T T_{t_{n_k}}^* - \sum_{j=1}^{r_n} T_{t_{n_k}} X_n'(j) X_n'(j)^T T_{t_{n_k}}^* \\
&\quad = \sum_1 T_{t_{n_k}} X_j X_j^T T_{t_{n_k}}^* - \sum_{j=p_n+1}^{r_n} T_{t_{n_k}} \; X_n'(j) X_n'(j)^T T_{t_{n_k}}^* + \sum_{j=q_n+1}^{r_n} T_{t_{n_k}} \; X_n(j) X_n(j)^T T_{t_{n_k}}^*.
\end{aligned}$$

Here, $\sum_1$ extends over all j such that $X_j \subset \{X_n(1), \cdots, X_n(q_n)\} \setminus \{X_n'(1), \cdots, X_n'(p_n)\}$. However, for any such X_j, $\|T_{t_{n_k}} X_j\| \le \rho a^{-1} < \rho$. Furthermore, the number of such possible j is $q_n - p_n \le r_n$. In the third summation, each summand is bounded because for $j > q_n$, $\|T_{t_{n_k}} \; X_n(j)\| \le \rho$. The number of such j is $r_n - q_n \le r_n$. For the middle term, if $j > p_n$, then $\|T_{t_{n_k}} \; X_n'(j)\| \le \rho$. The number of such j is $r_n - p_n \le r_n$. Therefore,

$$\left\| \sum_{j=1}^{r_n} T_{t_{n_k}} X_n(j) X_n(j)^T T_{t_{n_k}}^* - \sum_{j=1}^{r_n} T_{t_{n_k}} X_n'(j) X_n'(j)^T T_{t_{n_k}}^* \right\| \le 3 r_n \rho^2 \tag{6}$$

Summarizing, by combining (4-6) , and since $r_n \le r_{n_{k+1}}$, for $n \in I_k$ we have that

$$[\max_{n\in I_k} q_n \le r_{n_k}] \subset \Big[\max_{n\in I_k} \Big\| T_{t_{n_k}}(W_n - R_n)T^*_{t_{n_k}} \Big\| \le 4r_{n_{k+1}}\rho^2\Big]$$

Now, let $\epsilon > 0$. Since $r_n = o(L_2 n)$, and $n_{k+1}/n_k \to a$ we have that for sufficiently large k, $\frac{4r_{n_{k+1}}\rho^2}{L_2 n_k} < \epsilon$. We then obtain

$$\begin{aligned} P\Big(\max_{n\in I_k} \Big\| \frac{T_{t_{n_k}}(W_n - R_n)T^*_{t_{n_k}}}{L_2 n} \Big\| > \epsilon\Big) &\le P\Big(\max_{n\in I_k} \Big\| T_{t_{n_k}}(W_n - R_n)T^*_{t_{n_k}} \Big\| > \epsilon L_2 n_k\Big) \\ &\le P\Big(\max_{n\in I_k} \Big\| T_{t_{n_k}}(\sum_{j=1}^{n} w_j - \sum_{j=1}^{r_n} X_n(j)) \Big\| > 4r_{n_{k+1}}\rho^2\Big) \\ &\le P\Big(\max_{n\in I_k} q_n \ge r_{n_{k+1}} + 1\Big) \qquad (7) \end{aligned}$$

Our goal is now to show that the quantity in (7) is finitely summable. By (2), observe that

$$\max_{n\in I_k} \Big\| T_{t_n} T^{-1}_{t_{n_k}} \Big\| \le \max_{n\in I_k} \Big\| T_{t_n} T^{-1}_{t_{n_k}} - \Big(\frac{t_n}{t_{n_k}}\Big)^{-1/2} I \Big\| + \max_{n\in I_k} \Big(\frac{t_n}{t_{n_k}}\Big)^{-1/2} \to 1.$$

Therefore, for sufficiently large k, it can be made less than a, (since $a > 1$). Then,

$$\max_{n\in I_k} q_n = \max_{n\in I_k} \sum_{j=1}^{n} I\big[\|T_{t_n} X_j\| > \rho/a\big] \le \max_{n\in I_k} \sum_{j=1}^{n} I\big[\|T_{t_{n_k}} X_j\| > \frac{\rho}{a^2}\big] \le \sum_{j=1}^{n_{k+1}} I\big[\|T_{t_{n_k}} X_j\| > \frac{\rho}{a^2}\big].$$

Therefore, $\big[\max_{n\in I_k} q_n \ge r_{n_{k+1}} + 1\big] \subset \big[\sum_{j=1}^{n_{k+1}} I[\|T_{t_{n_k}} X_j\| > \frac{\rho}{a^2}] \ge r_{n_{k+1}} + 1\big]$ The proof is now reduced to showing that

$$\sum_k P\Big(\sum_{j=1}^{n_{k+1}} I[\|T_{t_{n_k}} X_j\| > \frac{\rho}{a^2}] \ge r_{n_{k+1}} + 1\Big) < \infty. \qquad (8)$$

The probability here is Binomial, so we need a bound on probabilities, $P(B \ge \alpha + 1)$, where B has a binomial distribution with parameters $n = n_{k+1}$, and $p = P(\|T_{t_{n_k}} X\| > \frac{\rho}{a^2})$. Take $\alpha = r_{n_{k+1}}$. The bound we use is from Feller, [4], p. 173, equation (10.9),

$$\begin{aligned} P(B \ge \alpha + 1) &= \frac{n!}{\alpha!(n-1-\alpha)!} \int_0^p t^\alpha (1-t)^{n-\alpha-1} dt \\ &\le \frac{n!}{\alpha!(n-1-\alpha)!} \frac{p^{\alpha+1}}{\alpha+1} \end{aligned}$$

Note that $\frac{n}{n-\alpha-1} \to 1$, as $k \to \infty$, since $r_n = o(L_2 n)$. Therefore, its square root is bounded by two eventually. Also, note that as $k \to \infty$, so do n, α, and $n - \alpha - 1$. Since each of the three factorials is going to infinity with k, we may apply Stirling's Formula to each. Hence for sufficiently large k,

$$\frac{n!}{\alpha!(n-1-\alpha)!} \le \frac{2}{e\sqrt{\pi}} \Big(\frac{n}{n-\alpha-1}\Big)^n n^{\alpha+1} \alpha^{-(\alpha+1/2)} \le \frac{2}{e\sqrt{\pi}} \exp\Big(\frac{n(\alpha+1)}{n-\alpha-1}\Big) n^{\alpha+1} \alpha^{-(\alpha+1/2)}$$

The last inequality here follows from the fact that $(1+\frac{x}{n})^n \le e^x \quad \forall x, n$.

Substituting this into the Binomial bound from Feller yields

$$
\begin{aligned}
P(B \ge \alpha+1) &\le \frac{2}{e\sqrt{\pi}} \exp\Big(\frac{n(\alpha+1)}{n-\alpha-1}\Big) n^{\alpha+1} \alpha^{-(\alpha+1/2)} \frac{p^{\alpha+1}}{\alpha+1} \\
&= \frac{2}{e\sqrt{\pi}} \frac{\sqrt{\alpha}}{\alpha+1} \Big(\alpha^{-1} np \exp(\frac{n}{n-\alpha-1})\Big)^{\alpha+1} \\
&\le \frac{2}{e\sqrt{\pi}} \Big(\alpha^{-1} np \exp(\frac{n}{n-\alpha-1})\Big)^{\alpha+1}
\end{aligned}
$$

The last inequality follows from the fact that $\frac{\sqrt{\alpha}}{\alpha+1} \to 0$, and so is eventually bounded above by one. Next, let $\delta = \frac{\rho^2}{a^6}$. By regular variation of T_t, (2), we have that $\|T_{t_{n_k}} T^{-1}_{\delta t_{n_k}}\| \to \sqrt{\delta}$, and so, since $a > 1$, $\|T_{t_{n_k}} X\| \le \|T_{t_{n_k}} T^{-1}_{\delta t_{n_k}}\| \, \|T_{\delta t_{n_k}} X\| \le a\sqrt{\delta}\|T_{\delta t_{n_k}} X\|$, for large k. So, $[\|T_{t_{n_k}} X\| > \frac{\rho}{a^2}] \subset [\|T_{\delta t_{n_k}} X\| > 1]$, by definition of δ. Recalling the construction of r_n and the definition of $\Gamma(s)$ at the outset of this section, we've shown that $p = P(\|T_{t_{n_k}} X\| > \frac{\rho}{a^2}) \le P(\|T_{\delta t_{n_k}} X\| > 1) \le \frac{1}{\delta t_{n_k}} \Gamma(\delta t_{n_k})$.

Next, fix $M > 0$, to be specified later. Since $\xi_n L(\frac{\xi_n}{\Gamma(\delta t_n)}) \to \infty$, we have that for large enough k, $\xi_{n_k} L(\frac{\xi_{n_k}}{\Gamma(\delta t_{n_k})}) \ge M$. A little algebra yields, $\Gamma(\delta t_{n_k}) \le \xi_{n_k} e^{-M/\xi_{n_k}}$.

Organizing the preceding paragraphs shows that the series in (8) is, for some constant $C > 0$ which may increase from one line to the next,

$$
\begin{aligned}
&\le C \sum_k \Big(r^{-1}_{n_{k+1}} n_{k+1} \; \frac{1}{\delta t_{n_k}} \Gamma(\delta t_{n_k}) \exp(\frac{n_{k+1}}{n_{k+1} - r_{n_{k+1}} - 1}) \Big)^{r_{n_{k+1}}+1} \\
&\le C \sum_k \Big(r^{-1}_{n_{k+1}} n_{k+1} \; \frac{1}{\delta t_{n_k}} \xi_{n_k} e^{-M/\xi_{n_k}} \exp(\frac{n_{k+1}}{n_{k+1} - r_{n_{k+1}} - 1}) \Big)^{r_{n_{k+1}}+1} \\
&= C \sum_k \Big(\frac{1}{\delta} \; \frac{n_{k+1}}{n_k} \; \frac{L_2 n_k \xi_{n_k}}{r_{n_{k+1}}} \; e^{-M/\xi_{n_k}} \exp(\frac{n_{k+1}}{n_{k+1} - r_{n_{k+1}} - 1}) \Big)^{r_{n_{k+1}}+1} \\
&\le C \sum_k \Big(\frac{a^2 e^a}{\delta} \; e^{-M/\xi_{n_k}} \Big)^{r_{n_{k+1}}+1} \qquad (9)
\end{aligned}
$$

To see (9) we bound the tail of the series as follows. $n_{k+1}/n_k \to a$, so eventually it is bounded by $a^{3/2}$. Also, since $\xi_{n_k} L_2 n_k \uparrow$, $\frac{\xi_{n_k} L_2 n_k}{r_{n_{k+1}}} \le \frac{\xi_{n_{k+1}} L_2 n_{k+1}}{[\xi_{n_{k+1}} L_2 n_{k+1}]} \to 1$, so eventually it is bounded by $a^{1/2}$. Finally, $\frac{n_{k+1}}{n_{k+1} - r_{n_k} - 1} \to 1$, so eventually it is bounded by a. Exponentiating, $\exp(\frac{n_{k+1}}{n_{k+1} - r_{n_k} - 1}) \le e^a$.

Next, given $a > 1$, and $\delta > 0$, there exists $x_0(a, \delta)$, such that for $x \ge x_0$, $e^{x/2} \ge \frac{a^2}{\delta} e^a$. Since $\xi_{n_k} \downarrow 0$, $\frac{M}{\xi_{n_k}} \ge x_0$ for large k. Hence, $\frac{a^2 e^a}{\delta} e^{-M/\xi_{n_k}} \le e^{-M/2\xi_{n_k}}$. Also, $r_{n_{k+1}} + 1 = [\xi_{n_{k+1}} L_2 n_{k+1}] + 1 \ge \xi_{n_{k+1}} L_2 n_{k+1} \ge \xi_{n_k} L_2 n_k$. So, $\frac{-M(r_{n_{k+1}}+1)}{2\xi_{n_k}} \le \frac{-M L_2 n_k}{2}$.

Utilizing the previous paragraph in (9), for some constants $C > 0$, which may increase from one inequality to the next,

$$
\sum_k \Big(\frac{a^2 e^a}{\delta} \; e^{-M/\xi_{n_k}} \Big)^{r_{n_k}+1} \le C \sum_k \exp\Big(\frac{-M(r_{n_k}+1)}{2\xi_{n_k}} \Big) \le C \sum_k \exp\Big(\frac{-M L_2 n_k}{2} \Big).
$$

This converges as long as $M > 2$. This shows the series in (7) converges and since $\epsilon > 0$ was arbitrary, this completes the proof of Lemma 6.
Proof of Lemma 7: First, we would like to replace the t_n with t_{n_k}. This can be done via the triangle inequality, and $t_n \uparrow$. Also, note that the quantity in the expectation is nothing but $u = u(k, a, \rho)$.

$$\begin{aligned} \left\| t_n T_{t_{n_k}} E u T^*_{t_{n_k}} - I \right\| &\le \left| \frac{t_n}{t_{n_k}} \right| \left\| t_{n_k} T_{t_{n_k}} E u T^*_{t_{n_k}} - I \right\| + \left| \frac{t_n}{t_{n_k}} - 1 \right| \\ &\le \frac{t_{n_{k+1}}}{t_{n_k}} \left\| t_{n_k} T_{t_{n_k}} E u T^*_{t_{n_k}} - I \right\| + \left| \frac{t_{n_{k+1}}}{t_{n_k}} - 1 \right| \\ &\to a \cdot 0 + a - 1 \end{aligned}$$

as $k \to \infty$, as we will show presently. The other two being obvious, we must show the second factor of the first term goes to zero. However, this follows from "Gaussian" convergence. Indeed, Sepanski [16] shows that under mean 0 and GDOAG, $T_n C_n T^*_n \to I$ in probability. Considering this as a triangular array $\{T_n X_i X_i^T T^*_n\}$ of random elements in the space of operators from $\mathbb{R}^d$ to $\mathbb{R}^d$, we have $T_n C_n T^*_n$ is shift convergent to the zero operator, when centered by I. On the other hand, (see Araujo and Gine, Theorem 5.9) if the triangular array is shift convergent it can be centered by the truncated means. That is $T_n C_n T^*_n - nET_n XX^T T^*_n I[\|T_n XX^T T^*_n\| \le \delta]$ is convergent to the zero operator for all $\delta > 0$. Replace δ with ρ^2 and apply convergence of types to yield $nET_n XX^T T^*_n I[\|T_n XX^T T^*_n\| \le \rho^2] - I \to 0$. Taking square roots inside the indicator, $(\|T_n XX^T T^*_n\| = \|T_n X\|^2)$ and applying this along the subsequence t_{n_k} completes the proof of Lemma 7.

Proof of Corollary 2: Let T_{t_n} be as in Theorem 1. Let A_n be the normalizing sequence constructed by Hahn and Klass. In particular, the key property we will need is that A_n satisfy (1) and are positive and symmetric. Since both A_n and T_n satisfy (1), we may apply Billingsley's multivariate Convergence of Types Theorem [2] to conclude that there exist orthogonal transformations P_n, and a sequence of linear operators $B_n \to I$ such that $T_n = B_n P_n A_n$. From this it follows that Theorem 1 holds for the sequence A_n, except that perhaps A_n may not satisfy (2). However, the almost sure limit does hold as claimed in the corollary. Indeed, since $A_{t_n} T^{-1}_{t_n} = P^*_{t_n} B^{-1}_{t_n}$, we have that

$$\begin{aligned} \left\| \frac{A_{t_n}(C_n - R_n)A_{t_n}}{L_2 n} - I \right\| &= \left\| A_{t_n} T^{-1}_{t_n} \Big[\frac{T_{t_n}(C_n - R_n)T^*_{t_n}}{L_2 n} \Big] T^{*-1}_{t_n} A_{t_n} - I \right\| \\ &= \left\| P^*_{t_n} B^{-1}_{t_n} \Big[\frac{T_{t_n}(C_n - R_n)T^*_{t_n}}{L_2 n} \Big] B^{*-1}_{t_n} P_{t_n} - I \right\| \\ &= \left\| B^{-1}_{t_n} \Big[\frac{T_{t_n}(C_n - R_n)T^*_{t_n}}{L_2 n} \Big] B^{*-1}_{t_n} - I \right\| \to 0 \quad a.s., \end{aligned}$$

since the outside factors converge to the identity and the inside factor converges to the identity almost surely by Theorem 1. Since A_{t_n} are positive and symmetric and since $C_n - R_n$ are positive and symmetric as well, $\frac{A_{t_n}(C_n - R_n)^{1/2}}{\sqrt{L_2 n}} \to I \quad a.s.$

Proof of Corollary 3: In Sepanski [15] it was shown that $\frac{T_{t_n} S_n^{(r_n)}}{\sqrt{2} L_2 n} \twoheadrightarrow \bar{B} \quad a.s.$, for operators satisfying (1) and (2). However, this will also hold for any other sequence of operators satisfying (1) by convergence of types and the fact that $\bar{B}$ is invariant under orthogonal transformations.

Hence, it holds with T_{t_n} replaced by the Hahn and Klass sequence of operators A_{t_n}. Now combine with Corollary (2) to obtain the result.

Proof of Theorem 4: The point here is to eliminate the trimming from Corollary 3, both from the normalizing operator, and from the partial sum. This can be done since the two types of trimming, or lack thereof, basically cancel each other out. First, we prove a simple result relating the norms of the trimmed operator to the untrimmed operator.

Lemma 9: Let $v_1, \cdots, v_n$ be vectors in $\mathbb{R}^d$. Suppose $S \subset \{1, \cdots, n\}$. Define $A = \sum_{i=1}^n v_i v_i^T$, and $B = \sum_{i \in S} v_i v_i^T$. Assume that A is invertible, then $\|A^{-1/2} B^{1/2}\| \leq 1$.
Proof: First, clearly A and B are nonnegative and symmetric, therefore each has a nonnegative and symmetric square root. We denote the unit sphere by S^{d-1}.

$$\begin{aligned}
\left\| A^{-1/2} B^{1/2} \right\|^2 &= \left\| (A^{-1/2} B^{1/2})(A^{-1/2} B^{1/2})^* \right\| \\
&= \left\| A^{-1/2} B A^{-1/2} \right\| \\
&= \sup_{\theta \in S^{d-1}} \left\langle A^{-1/2} B A^{-1/2} \theta, \theta \right\rangle \\
&= \sup_{\theta \in S^{d-1}} \left\langle B A^{-1/2} \theta, A^{-1/2} \theta \right\rangle \\
&= \sup_{\theta \in S^{d-1}} \left\langle B A^{-1/2} \Big(\frac{A^{1/2}\theta}{\|A^{1/2}\theta\|} \Big), A^{-1/2} \Big(\frac{A^{1/2}\theta}{\|A^{1/2}\theta\|} \Big) \right\rangle \\
&= \sup_{\theta \in S^{d-1}} \frac{\langle B\theta, \theta \rangle}{\|A^{1/2}\theta\|^2} \\
&= \sup_{\theta \in S^{d-1}} \frac{\sum_{i \in S} \langle v_i, \theta \rangle^2}{\sum_{i=1}^n \langle v_i, \theta \rangle^2} \leq 1.
\end{aligned}$$

From Lemma 9 we conclude that $\|C_n^{-1/2}(C_n - R_n)^{1/2}\| \leq 1$. Combining this with Corollary 3, we obtain

$$\frac{C_n^{-1/2} S_n^{(r_n)}}{\sqrt{2L_2 n}} \twoheadrightarrow C \subset \bar{B} \quad a.s. \tag{10}$$

Of course, C is nonrandom due to the zero-one law.

Next, we show that, up to a set of measure zero, $\frac{C_n^{-1/2} S_n^{(r_n)}}{\sqrt{2L_2 n}}$ and $\frac{C_n^{-1/2} S_n}{\sqrt{2L_2 n}}$ have the same cluster sets. We achieve this by showing the difference goes to zero with probability one.

Using the notation of Lemma 9, but denoting $\mathbf{b} = \sum_{i \in S} v_i$, and applying a similar argument, we see that

$$\begin{aligned}
\|A^{-1/2}\mathbf{b}\| &= \sup_{\theta \in S^{d-1}} |\langle A^{-1/2}\mathbf{b}, \theta \rangle| \\
&= \sup_{\theta \in S^{d-1}} \frac{|\langle \mathbf{b}, \theta \rangle|}{\|A^{1/2}\theta\|} \\
&= \sup_{\theta \in S^{d-1}} \frac{|\sum_{i \in S} \langle v_i, \theta \rangle|}{\Big(\sum_{i=1}^n \langle v_i, \theta \rangle^2 \Big)^{1/2}} \\
&\leq \sup_{\theta \in S^{d-1}} \frac{\sum_{i \in S} |\langle v_i, \theta \rangle|}{\Big(\sum_{i=1}^n \langle v_i, \theta \rangle^2 \Big)^{1/2}}
\end{aligned}$$

$$\leq \sup_{\theta \in S^{d-1}} \frac{\sqrt{\#S}\Big(\sum_{i\in S}\langle v_i,\theta\rangle^2\Big)^{1/2}}{\Big(\sum_{i=1}^n \langle v_i,\theta\rangle^2\Big)^{1/2}}$$
$$\leq \sqrt{\#S}$$

The second to last inequality follows from Cauchy-Schwarz. Here $\#S$ denotes the cardinality of S. Applying the above to $A = C_n$, and $\mathbf{b} = \sum_{i=1}^{r_n} X_n(i)$, yields $\|C_n^{-1/2}(S_n - S_n^{(r_n)})\| \leq \sqrt{r_n}$. From this we conclude that

$$\Big\|\frac{C_n^{-1/2}(S_n - S_n^{(r_n)})}{\sqrt{2L_2 n}}\Big\| \leq \sqrt{\frac{r_n}{2L_2 n}} \to 0. \tag{11}$$

(10) and (11) combine to yield

$$\frac{C_n^{-1/2} S_n}{\sqrt{2L_2 n}} \twoheadrightarrow C \subset \bar{B} \quad a.s. \tag{12}$$

This implies the upper bound in Theorem 4. To show the lower bound in Theorem 4, we appeal to Theorem 1 of Griffin and Kuelbs, [7]. First, observe that by Cauchy-Schwarz,

$$|\langle S_n,\theta\rangle| = |\langle C_n^{1/2} C_n^{-1/2} S_n,\theta\rangle| = |\langle C_n^{-1/2} S_n, C_n^{1/2}\theta\rangle| \leq \|C_n^{-1/2} S_n\|\,\|C_n^{1/2}\theta\| = \|C_n^{-1/2} S_n\|\Big(\sum_{i=1}^n \langle X_i,\theta\rangle^2\Big)^{1/2}$$

Dividing through this inequality by $\sqrt{2L_2 n}\Big(\sum_{i=1}^n \langle X_i,\theta\rangle^2\Big)^{1/2}$ yields for any unit vector θ,

$$1 \leq \limsup \frac{|\langle S_n,\theta\rangle|}{\sqrt{2L_2 n}\Big(\sum_{i=1}^n \langle X_i,\theta\rangle^2\Big)^{1/2}} \leq \limsup \frac{\|C_n^{-1/2} S_n\|}{\sqrt{2L_2 n}} \leq 1 \quad a.s.$$

The first inequality holds by Griffin and Kuelbs, [7], Theorem 1, which is applicable due to the fact that if X is in the GDOA of the multivariate Gaussian law then, for any θ, $\langle X,\theta\rangle$ is in the DOA of the univariate Gaussian law. The last inequality holds by (12).

REFERENCES

[1] Araujo, and Gine, E. (1980) *The Central Limit Theorem for Real and Banach Valued Random Variables.* John Wiley and Sons. New York.

[2] Billingsley, P. (1966) Convergence of types in k-space. *Z. Wahrsh. Verw. Gebiete.* **5** 175-179

[3] Dudley, R.M. (1989) *Real Analysis and Probability.* Chapman and Hall. New York.

[4] Feller, W.(1968) *An Introduction to Probability Theory and its Applications, Volume 1. Third Edition.* John Wiley and Sons. New York.

[5] Feller, W.(1968) An extension of the law of the iterated logarithm to variables without variance.*J. Math. Mechan.* **18**, 343-355.

[6] Gine, E. and Mason, D.M. (1998) On the LIL for self-normalized sums of iid random variables. *J. Theor. Probab.* **v. 11 no. 2**, 351-370.

[7] Griffin, P.S. and Kuelbs, J.D. (1989) Self normalized laws of the iterated logarithm. *Ann. Probab.* **17**, 1571-1601.

[8] 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. Probab.* **8**, 262-280.

[9] Kuelbs, J. and Ledoux M. (1987) Extreme values and the law of the iterated logarithm. *Prob. Theory and Rel. Fields.* **74**, 319-340.

[10] Meerschaert, M. (1994) Norming operators for generalized domains of attraction. *J. Theoret. Probab.* **7**, 793-798.

[11] Meerschaert, M.M. and Scheffler, H. P.(2001) *Limit Theorems for Sums of Independent Random Vectors.* Wiley. New York.

[12] Pruitt, W.E. (1981) General one sided laws of the iterated logarithm. *Ann. Probab.* **9**, 1-48.

[13] Sepanski, S. (1999) A compact Law of the Iterated Logarithm for random vectors in the generalized domain of attraction of the multivariate gaussian law. *J. Theoret. Probab.* **12**, 757-778

[14] Sepanski, S. (2001) Extreme Values and the multivariate compact Law of the Iterated Logarithm. *J. Theoret. Probab.* To appear.

[15] Sepanski, S. (2001) A law of the iterated logarithm for multivariate trimmed sums. Preprint.

[16] Sepanski, S. (1994) Asymptotics for multivariate t statistic and Hotelling's T^2 statistic under infinite second moments via bootstrapping. *J. Multivariate Analysis.* **v 49** 41-54